
STATE ESTIMATION IN
ELECTRIC POWER SYSTEMS
A Generalized Approach

THE KLUWER INTERNATIONAL SERIES IN ENGINEERING AND COMPUTER SCIENCE

Power Electronics and Power Systems

Series Editor

M. A. Pai

Other books in the series:

COMPUTATIONAL AUCTION MECHANISMS FOR RESTRUCTURED POWER
INDUSTRY OPERATIONS

Gerald B. Sheblé, ISBN: 0-7923-8475-X

ANALYSIS OF SUBSYNCHRONOUS RESONANCE IN POWER SYSTEMS

K.R. Padiyar, ISBN: 0-7923-8319-2

POWER SYSTEMS RESTRUCTURING: *Engineering and Economics*

Marija Ilic, Francisco Galiana, and Lester Fink, ISBN: 0-7923-8163-7

CRYOGENIC OPERATION OF SILICON POWER DEVICES

Ranbir Singh and B. Jayant Baliga, ISBN: 0-7923-8157-2

VOLTAGE STABILITY OF ELECTRIC POWER SYSTEMS, Thierry

Van Cutsem and Costas Vournas, ISBN: 0-7923-8139-4

AUTOMATIC LEARNING TECHNIQUES IN POWER SYSTEMS, Louis A.

Wehenkel, ISBN: 0-7923-8068-1

ENERGY FUNCTION ANALYSIS FOR POWER SYSTEM STABILITY,

M. A. Pai, ISBN: 0-7923-9035-0

ELECTROMAGNETIC MODELLING OF POWER ELECTRONIC

CONVERTERS, J. A. Ferreira, ISBN: 0-7923-9034-2

MODERN POWER SYSTEMS CONTROL AND OPERATION, A. S. Debs,

ISBN: 0-89838-265-3

RELIABILITY ASSESSMENT OF LARGE ELECTRIC POWER SYSTEMS,

R. Billington, R. N. Allan, ISBN: 0-89838-266-1

SPOT PRICING OF ELECTRICITY, F. C. Schweppe, M. C. Caramanis, R. D.

Tabors, R. E. Bohn, ISBN: 0-89838-260-2

INDUSTRIAL ENERGY MANAGEMENT: *Principles and Applications*,

Giovanni Petrecca, ISBN: 0-7923-9305-8

THE FIELD ORIENTATION PRINCIPLE IN CONTROL OF INDUCTION

MOTORS, Andrzej M. Trzynadlowski, ISBN: 0-7923-9420-8

FINITE ELEMENT ANALYSIS OF ELECTRICAL MACHINES, S. J. Salon,

ISBN: 0-7923-9594-8

STATE ESTIMATION IN ELECTRIC POWER SYSTEMS A Generalized Approach

A. MONTICELLI

University of Campinas, Unicamp



Kluwer Academic Publishers
Boston/Dordrecht/London

0-7923-2254-4

Distributors for North, Central and South America:

Kluwer Academic Publishers
101 Philip Drive
Assinippi Park
Norwell, Massachusetts 02061 USA
Tel: 781-871-6600
Fax: 781-871-6528
E-mail: kluwer@wkap.com

Distributors for all other countries:

Kluwer Academic Publishers Group
Distribution Centre
Post Office Box 322
3300 AH Dordrecht, THE NETHERLANDS
Tel: 31 78 6392 392



Fax: 31 78 6546 474

E-mail: orderdept@wkap.nl

Electronic Services: <http://www.wkap.nl>

Library of Congress Cataloging-in-Publication Data

A C.I.P. Catalogue record for this book is available
from the Library of Congress.

Copyright © 1999 by Kluwer Academic Publishers.

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, mechanical, photocopying, recording, or otherwise, without the prior written permission of the publisher, Kluwer Academic Publishers, 101 Philip Drive, Assinippi Park, Norwell, Massachusetts 02061

Printed on acid-free paper.

To Isadora

Contents

Foreword	xv
Acknowledgments	xvii
1. REAL-TIME MODELING OF POWER NETWORKS	1
1.1 Security Concepts	1
1.2 Network Model Builder	3
1.3 Conventional State Estimation	7
1.4 Economy-Security Control	8
1.5 Generalized State Estimation	10
1.6 Historical Notes and References	12
Chapter References	13
2. LEAST-SQUARES AND MINIMUM NORM PROBLEMS	15
2.1 Introduction	15
2.2 Linear Least-Squares Problem	16
2.2.1 Problem Formulation	16
2.2.2 Normal Equation	16
2.3 Linear Minimum-Norm Problem	19
2.3.1 Problem Formulation	19
2.3.2 Lagrangian Function	19
2.4 Observability and Controllability	22
2.4.1 Least-Squares Problem	22
2.4.2 Minimum-Norm Problem	23
2.5 Geometric Interpretation	24
2.5.1 Range and Null Spaces	24
2.5.2 Least-Squares Problem	26
2.5.3 Minimum-Norm Problem	26
2.6 Overdetermined Nonlinear Models	28
2.6.1 Gauss Newton Method	28
2.6.2 Newton Raphson Method	29
2.7 Historical Notes and References	33
2.8 Problems	33

Chapter References	36
3. DC STATE ESTIMATOR	39
3.1 Overview of the DC State Estimator	39
3.2 State Variables	41
3.3 Measurement Model	44
3.4 Solving the Normal Equation	49
3.5 Phase-Shift Estimation	52
3.6 Parameter Estimation	53
3.7 Physical Level Modeling	55
3.8 Historical Notes and References	60
3.9 Problems	60
Chapter References	61
4. POWER FLOW EQUATIONS	63
4.1 Network Branch Model	63
4.1.1 Transmission Line	63
4.1.2 Transformer	66
4.1.3 In-phase Transformer	66
4.1.4 Phase-Shifting Transformer	69
4.1.5 Unified Branch Model	70
4.2 Active and Reactive Power Flows	72
4.2.1 Transmission Line	72
4.2.2 In-phase Transformer	73
4.2.3 Phase-Shifting Transformer with $a_{km} = 1$	74
4.2.4 Unified Power Flow Equations	76
4.3 Nodal Formulation of the Network Equations	76
4.4 Basic Power Flow Problem	78
4.4.1 Problem Variables	78
4.4.2 Basic Bus Types	78
4.4.3 Problem Solvability	79
4.4.4 Equality and Inequality Constraints	79
4.5 Newton Raphson Method	81
4.5.1 Unidimensional Case	82
4.5.2 Quadratic Convergence	85
4.5.3 Multidimensional Case	85
4.6 $P\theta$ -QV Decoupling	89
4.7 Linearization	91
4.7.1 Transmission Line	91
4.7.2 Series Capacitor	93
4.7.3 In-Phase Transformer	94
4.7.4 Phase-Shifter	95
4.8 Matrix Formulation	97
4.9 DC Power Flow Model	99
4.10 Historical Notes and References	100
4.11 Problems	100

Chapter References	101
5. NETWORK REDUCTION AND GAUSS ELIMINATION	103
5.1 Bus Admittance Matrix	103
5.2 Network Reduction and Expansion	105
5.2.1 Network Reduction	105
5.2.2 Network Expansion	109
5.2.3 Paths	115
5.3 LDU Decomposition	116
5.3.1 Gauss Elimination	116
5.3.2 Elementary Matrices	117
5.3.3 Triangular Factorization	119
5.4 Using LDU Factors to Solve Linear Systems	122
5.5 Path Finding	123
5.6 Pivot Ordering to Preserve Sparsity	124
5.7 MDML and MLMD Ordering Schemes	126
5.8 Blocked Formulation of Newton Power Flow	129
5.9 Gain Matrix	131
5.10 Factorization of Rectangular Matrices	135
5.10.1 Partial and Complete Pivoting	135
5.10.2 Factorization	136
5.11 Matrix Inversion Lemma	138
5.12 Historical Notes and References	140
5.13 Problems	141
Chapter References	141
6. NETWORK TOPOLOGY PROCESSING	143
6.1 Conventional Topology Processing	143
6.1.1 Pre-Processing of Raw Data	144
6.1.2 Bus Section Processing	145
6.1.3 Network Connectivity Analysis	147
6.1.4 Topology Processor in Tracking Mode	147
6.1.5 Critical Remarks	148
6.2 Generalized Topology Processing	148
6.2.1 Extended Islands	149
6.2.2 Local State Estimation: Analysis of Good Data and Bad Data	151
6.3 Network Reduction	152
6.3.1 Meter Arrangements	153
6.3.2 NTP and Jacobian Reduction	154
6.4 Historical Notes and References	156
6.5 Problems	157
Chapter References	157
7. OBSERVABILITY ANALYSIS	161
7.1 Bus/Branch Network Model	161
7.1.1 Physical Islands and Observable Islands	162
7.1.2 Basic Topological Algorithm	164

7.1.3	Reference Angles and Triangular Factorization	165
7.1.4	Unobservable Networks	168
7.1.5	Basic Numerical Algorithm	171
7.1.6	Redundant Measurements	172
7.1.7	Transformable Injection Measurements	174
7.1.8	Observable Islands	175
7.1.9	Irrelevant Injection Measurements	176
7.1.10	Hybrid Algorithm	179
7.1.11	Solvability and Observability	181
7.2	Bus-Section/Switching-Device Network Model	183
7.2.1	Power Flows as State Variables	184
7.2.2	Extended Observable Islands	187
7.2.3	Extended Numerical Observability Algorithm	189
7.3	Measurement Addition to Improve Observability	191
7.3.1	Additional Measurements	191
7.3.2	Avoiding Contamination	192
7.3.3	Measurement Placement Algorithm for the Bus/Branch Model	192
7.3.4	Measurement Placement Algorithm for the Physical Level Model	195
7.4	Historical Notes and References	197
7.5	Problems	197
	Chapter References	198
8.	BASIC TECHNIQUES FOR BAD DATA PROCESSING	201
8.1	Review of the DC State Estimator	201
8.2	Covariance Matrices	202
8.2.1	Sensitivity Analysis	202
8.2.2	Covariances of Estimated Vectors	204
8.3	Normalized Residuals	208
8.3.1	Covariance Matrix of the Normalized Residuals	208
8.3.2	Correlation Coefficients	209
8.3.3	The Largest Normalized Residual Test	210
8.3.4	Dormant Measurements and Error Estimates	213
8.3.5	Updating the $J(\hat{\mathbf{x}})$ Index	215
8.4	Hypotheses Testing	218
8.4.1	The $J(\hat{\mathbf{x}})$ Detection Test	218
8.4.2	The \mathbf{r}^n Detection and Identification Test	222
8.5	Historical Notes and References	223
8.6	Problems	224
	Chapter References	225
9.	MULTIPLE BAD DATA PROCESSING TECHNIQUES	227
9.1	Estimation Residuals	227
9.1.1	Partition of Measurement Set	228
9.1.2	Distribution Function	228
9.1.3	Dormant Measurements and Error Estimates	231
9.2	Multiple Normalized Residuals	232
9.2.1	Distribution Function	232
9.2.2	Updating the $J(\hat{\mathbf{x}})$ Index	235

9.3	Hypotheses Testing	236
9.3.1	Detection Test	236
9.3.2	Identification Test	239
9.4	Testing Equality Constraint Hypotheses	243
9.5	Strategies for Processing Interacting Bad Data	247
9.5.1	Successive Largest Normalized Residual	248
9.5.2	Dormant Measurements and Perfect Measurements	249
9.5.3	Binary Tree Search	252
9.5.4	Tabu Search	256
9.5.5	Localized Solutions, Pocketing and Zooming	258
9.6	Robust Estimators	260
9.6.1	Non-quadratic Estimators	260
9.6.2	Least Absolute Value Estimator	261
9.6.3	Least Median of Squares Estimator	262
9.7	Historical Notes and References	263
9.8	Problems	264
	Chapter References	265
10.	AC STATE ESTIMATOR	267
10.1	Review of the Problem Formulation	267
10.2	Flow Measurements	268
10.2.1	Power Flow Measurements	268
10.2.2	Current Magnitude Measurements	273
10.3	Bus Injection Measurement	275
10.3.1	Power Injection Measurement	275
10.3.2	Current Magnitude Measurement	280
10.4	Historical Notes and References	281
10.5	Problems	282
	Chapter References	282
11.	ESTIMATION BASED ON MULTIPLE SCANS OF MEASUREMENTS	283
11.1	State Estimation	283
11.1.1	Tracking State Estimator	283
11.1.2	Linear Dynamic State Estimator	284
11.1.3	Non-Linear Dynamic State Estimator	287
11.2	Parameter Estimation	289
11.2.1	Parameter Estimation by State Augmentation	289
11.2.2	Alternate State-Parameter Estimation	295
11.2.3	Dynamic State-Parameter Estimation	300
11.2.4	Newton Raphson Method – Second Order Derivatives	301
11.2.5	Dormant Parameter Technique	304
11.3	Historical Notes and References	309
11.4	Problems	310
	Chapter References	311
12.	FAST DECOUPLED STATE ESTIMATOR	313
12.1	Decoupled Solution of Linear System of Equations	313

12.1.1	System of Linear Equations	314
12.1.2	3-Step Algorithm	314
12.2	Fast Decoupled Power Flow	316
12.2.1	Newton Raphson Iteration	316
12.2.2	Mid-Iteration Updating	317
12.2.3	Ignoring Branch Series Resistances	317
12.2.4	Extended BX Method	319
12.2.5	Extended XB Method	322
12.2.6	BX Method	323
12.2.7	XB Method	324
12.3	Decoupled Solution of Overdetermined Systems	325
12.3.1	Overdetermined System of Linear Equations	325
12.3.2	3-Step Algorithm	325
12.4	Fast Decoupled State Estimator	326
12.4.1	Gauss Newton Iteration	326
12.4.2	Mid-Iteration Updating	327
12.4.3	Ignoring Branch Series Resistances	328
12.4.4	Extended BX Method	331
12.4.5	Extended XB Method	333
12.4.6	BX Method	334
12.4.7	XB Method	339
12.5	Historical Notes and References	340
12.6	Problems	341
	Chapter References	342
13.	NUMERICALLY ROBUST STATE ESTIMATORS	343
13.1	Normal Equation	343
13.1.1	Basic formulation	344
13.1.2	Equality Constraints	344
13.2	Sparse Tableau Formulation	347
13.2.1	Basic formulation	347
13.2.2	Equality Constraints	350
13.3	Peters Wilkinson Method	351
13.4	Blocked Sparse Tableau	353
13.5	Mixed Pivoting for Indefinite Matrices	356
13.5.1	Basic Algorithm	356
13.5.2	Sparsity Considerations	357
13.6	Orthogonal Transformation Approach	357
13.6.1	Orthogonal Decomposition of Normal Equation	358
13.6.2	Semi-normal Equation – A Hybrid Approach	361
13.6.3	Sparsity Considerations	362
13.6.4	Observability Analysis	363
13.7	Historical Notes and References	365
13.8	Problems	366
	Chapter References	367
	Appendices	369

A–Statistical Properties of Estimated Quantities	369
A.1 Distribution of State Estimate	369
A.2 Rank of Weighted Sensitivity Matrix	370
A.2.1 Eigenvalues and Eigenvectors	370
A.2.2 Weighted Sensitivity Matrix	370
A.2.3 Trace of Covariance Matrix of Measurement Estimates	370
A.2.4 Diagonalization of the Weighted Sensitivity Matrix	371
A.3 The χ^2 Property of the Least Squares Index	372
A.3.1 The $J(\hat{\mathbf{x}})$ Performance Index	372
A.3.2 Orthogonal Transformation	372
A.3.3 Covariance Matrix of Transformed Residuals	372
A.3.4 The χ_{m-n}^2 Distribution	373
A.4 Testing Equality Constraint Hypotheses	373
A.5 Historical Notes and References	374
Chapter References	374
B–Givens Rotation	375
B.1 Orthogonal Matrices	375
B.1.1 2x2 Orthogonal Matrices	375
B.1.2 Rotations and Reflections	376
B.2 Givens Rotations	378
B.2.1 Standard Givens Rotations	378
B.2.2 Fast Givens Rotations	380
B.2.3 Triangular Factorization	382
B.2.4 The 2-multiplication Algorithm	383
B.3 Historical Notes and References	387
Chapter References	388
Index	389

Foreword

Since it was championed three decades ago by Schweppe and others, state estimation for the real-time modeling of the electric power transmission network has remained an extremely active and contentious area. At the last count, there have been around a thousand research and development publications on new and improved methods. More recently, these have promoted dynamic, distributed and non-WLS (Weighted Least Squares) approaches.

Regardless, as I write this, centralized single-snapshot WLS estimation of the transmission network is universally established in the electric power industry. It is installed, or is scheduled to be installed, in virtually all Energy Management Systems (EMS) in the world.

The present book could not be more timely. It uniquely provides, at advanced text book level, a comprehensive anatomy of modern electric power system WLS state estimation theory and practice. It deals with everything of significance in this field, from network and power flow basics, to observability theory, to matrix solution methods, to bad data analysis, and as far as emerging technologies such as combinatorial bad data analysis and multiple-snapshot estimation. The reader can obtain an excellent, realistic grasp of the area, and even, with meticulous attention to detail, can approach specialist level.

In the early classical days of power system state estimation, the problem was formulated simply as the statistical averaging of telemetered measurements, in a network of accurately known impedances and topology. The network states (bus voltage magnitudes and angles) were declared observable or not, based on a relatively simple analysis of telemetered flow, injection and voltage measurements. A power flow model of the network was thus constructed, with any defective measurements detected, identified, and corrected or eliminated. Even this idealized formulation was (and is) not trivially solvable, particularly in the presence of mixed measurements, multiple bad measurements and poor mathematical conditioning of the problem. And unfortunately, the state estimation problem is not this simple.

Over the years, the focus of power system state estimation has shifted towards a more holistic approach. This term, borrowed from the medical field, implies utilizing all the available data. The present book's author, Professor Alcir Monticelli, has been a prime mover in this trend. He and others have insisted that, in order to construct the best power flow model, all data sources must be regarded as subject to statistical error, including device statuses and impedance values. This is the underlying message of the book's title: Generalized State Estimation. A generalized approach takes into account that the network impedances and topology are not accurately known. And many

measurements, particularly in the external network, are just crude predictions. Likewise, all other potentially imprecise data sources, including transformer taps, phase shifter angles, voltage regulation set points, interchange schedules, equipment limits, and plausibility criteria (e.g. non-negative loads) have to be factored into the estimation process.

Electric power industry deregulation has transformed state estimation from an important application into a critical one. Most individual utilities transmission networks are being amalgamated into larger regional not for profit network infrastructures. Power transfers now take place over electrical distances and in directions for which the network was never designed. Competitive, monetary factors are the sole drivers of the process, within electrical operating limits. But the transmission network has finite capacity. Therefore the system operator has to make equitable, security-related, congestion management decisions to curtail or deny power transfer rights. This is a critical commercial issue. It has to be founded and justified on a precise model of the power system, derived from the state estimation process. Moreover, accurate state estimation is the foundation of Locational Marginal Pricing methodologies for transmission congestion management costing.

These regional networks can be huge (a recent such EMS specification calls for the supervision of 100,000 buses). Computationally efficient state estimation solutions are therefore essential. This book deals with all the related issues - sparse matrix solution technology, including ordering and network reduction, and the robust orthogonal transformation approaches. In principle, state estimation in electric power systems is ubiquitous. It is not limited to transmission level modeling. Modern Distribution Management Systems and, in some cases, more vertical network management, need state estimation to reach further into the lower voltage network. Even off-line studies, requiring the assemblage of a multi-area model from disparate data sources, can benefit from the data integrating and reconciling capabilities of state estimation.

The scope of the present book is close to monumental. It crystalizes 30 years of electric power system WLS state estimation theory and practice into the modern essentials, presented at advanced text book level. Alcir Monticelli's multiple, focused R&D contributions to state estimation are cumulatively the most important in the area. These techniques have been adopted by state estimation developers worldwide. Moreover, Monticelli has the credentials of being a designer and developer of industrial-grade state estimation software that is used in the USA, South America and elsewhere. He therefore brings important insights to the book and, perhaps as profoundly, omits irrelevancies and side issues. My colleagues and I have had the privilege of working with him for 22 years.

Brian Stott
PCA Corporation
Mesa, Arizona, USA

Acknowledgments

I am indebted to Professors Ruben Romero and Linda Gentry El-Dash, who read the manuscript and provided many corrections and useful comments. Thanks are also due to Eduardo Asada for his useful suggestions, and to all those outstanding individuals with whom I have worked in the past two decades, who have helped me understand state estimation and its applications better, including Ariovaldo Garcia, Felix F. Wu, Brian Stott, Ongun Alsaç and Narsi Vempati, as well as to Professor M. A. Pai, who is the series editor for the Kluwer series on power systems for his valuable assistance. Last but not least, I would like to thank my wife Stella for her encouragement and love. Without her help, the book could never have been completed.

1 REAL-TIME MODELING OF POWER NETWORKS

A real-time model is a quasi-static computer-based mathematical representation of the current conditions in a power network. This model is extracted at intervals from “snapshots” of real-time measurements (analog measurements and the status of switching devices) as well as from static network data (basic configuration and parameters). State estimation is the key function for obtaining such a real-time network model. A more complete understanding of state estimation is becoming more important than in the past due both to new modeling needs and to changes induced by deregulation. This chapter reviews the main concepts involved in state estimation and economy-security functions carried out by the Energy Management System (EMS).

1.1 SECURITY CONCEPTS

The concept of network security is associated with the probability of the maintenance of an adequate supply. The higher the security level, the lower the probability of loss of load. Security oriented control actions are thus aimed at avoiding blackouts and equipment damage.

A contingency is a loss of transmission equipment and/or generation units. Contingency analysis involves the analysis of possible contingencies that are likely to occur to detect the potentially harmful ones, i.e., contingencies that, if they occurred, would lead the system into a state of emergency. Both single and

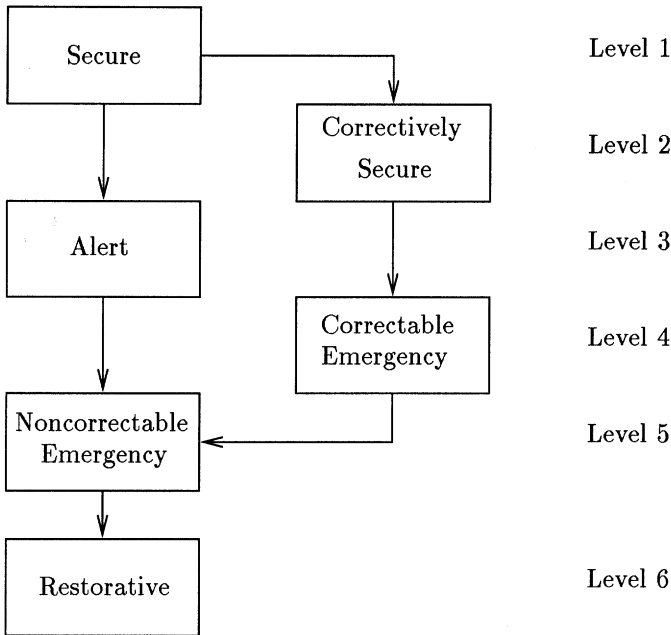


Figure 1.1. Power network static security levels. Involuntary transitions caused by contingencies (or accidental actions) are indicated by arrows.

multiple equipment outages are considered. There are two basic responses to an insecure contingency case (alert state): (a) modification of the pre-contingency state to eliminate the potential emergency in case the contingency really occurs, and (b) definition of a control strategy to manage the emergency should it occur.

Figure 1.1 illustrates a classification of the power network security levels which are needed to define such control actions. These security levels, along with the relevant control actions, are summarized below. The determination of the current security level of a system and the appropriate control action to be taken at each level are carried out by the various EMS economy-security functions. All these functions have in common the need to know the current system state provided by the state estimator.

- Level 1 (Secure). All load is supplied without operating limit violations. In the event of a contingency, no violations will occur. In this level, the network survives any of the contingencies postulated without the need to rely on any post-contingency action.
- Level 2 (Correctively Secure). All load is supplied without operating limit violations, as in Level 1, and contingency violations cause no loss of load as long as appropriate control action is taken to correct them (primarily for

active-power control). This level is more economical than Level 1, but relies on post-contingency actions being performed by the EMS. Such actions can be determined in advance by running an optimal power flow with security constraints and including pos-contingency rescheduling.

- Level 3 (Alert). All load is supplied without operating limit violations, as in Levels 1 and 2, although some violations caused by a contingency cannot be corrected without loss of load. The system can be brought back to Levels 1 or 2 by preventive rescheduling using an optimal power flow with contingency constraints.
- Level 4 (Correctable Emergency). All load is supplied, but operating limits are violated. These can be corrected without loss of load. The system can be brought back to Level 3 by corrective actions. At this level the long- or medium term limits used in Levels 1 through 3 can be violated, but not the short-term limits.
- Level 5 (Noncorrectable Emergency). All load is supplied but operating limits are violated and the situation cannot be corrected without loss of load. The amount and location of load loss can be optimized by an optimal power flow program.
- Level 6 (Restorative). No operating limits are violated, but loss of load has occurred. Restorative control aims to return the system to Levels 1 or 2.

The EMS runs real-time processes designed to maintain a designated security level at minimum operational cost. The maintenance of a designated security level is designed to avoid emergency conditions, and, if an emergency does occur, to guarantee the recovery of the network from the emergency conditions. These economy-security functions involves the integrated use of network monitoring, contingency analysis and optimal power flow, as illustrated in Fig. 1.2.

1.2 NETWORK MODEL BUILDER

Quasi-static network models are usually extracted from a single scan of measurements (a “snapshot” of system measurements) with new scans carried out every one or two seconds. Thus, the first step in real-time modeling is the establishment of a dynamic data base. This is performed by the data acquisition function, which processes both analog and status data. Analog measurements include power flows, power injections, voltage magnitudes, phase angles, and current magnitude measurements, whereas status data consists of data concerning network configuration. The static data base contains network basic configuration data along with parameter data (e.g. line impedances). If a state variable can be calculated from the available data (both dynamic and static), it is considered to be observable.

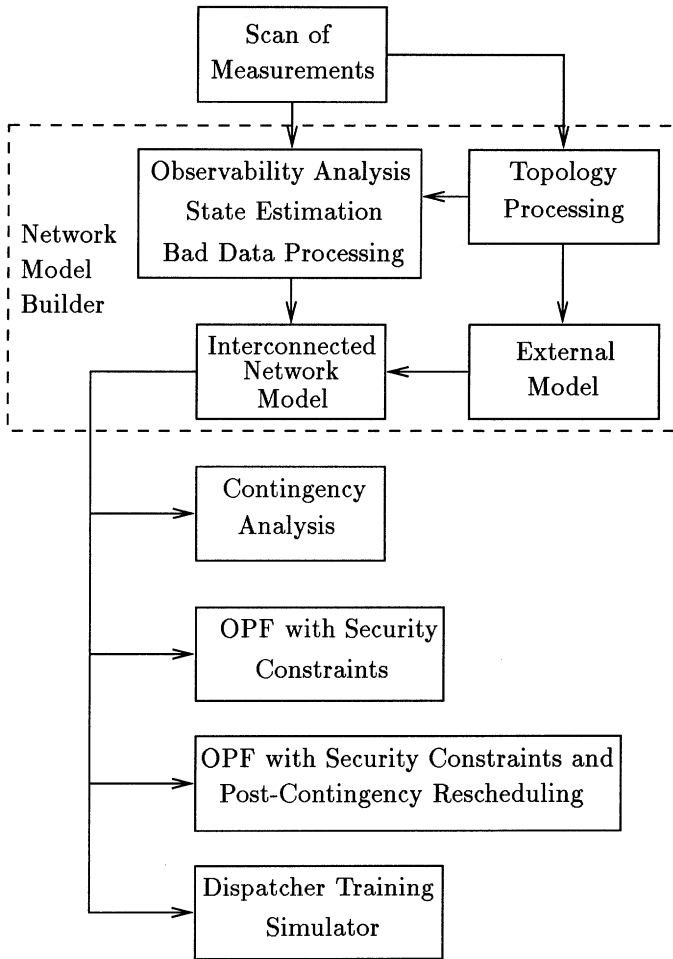


Figure 1.2. Real-time network analysis functions.

Conventional state estimation is performed on a bus/branch model of the same type that is used in power flow calculations. Network connectivity, however, is normally described in terms of bus-sections and switching-devices (physical level representation). The network topology processor is used to transform the bus-section/switching-device model into a bus/branch model.

Conventional topology processing is performed prior to state estimation and other related functions such as observability analysis and bad data processing; once the network topology is known, state estimation assumes that this topology is correct, and proceeds to estimate state and identify analog bad data whenever possible. In the conventional approach, the determination of

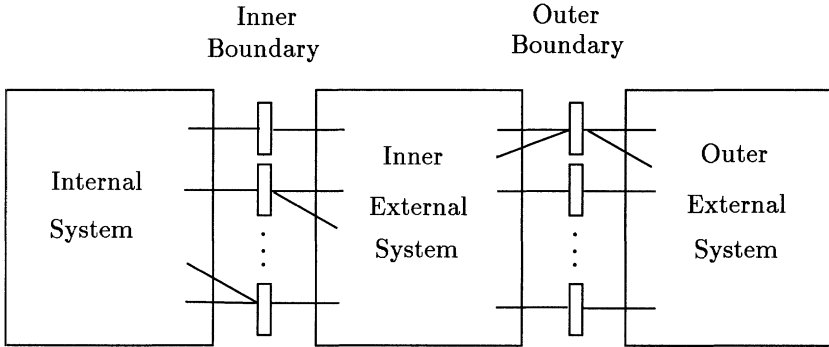


Figure 1.3. Network model.

the real time model from the available data set is decomposed into two main subproblems: the processing of logical data and the processing of analog data. The former consists of topology processing and the latter involves observability analysis and state estimation in which the network topology and parameters are considered as given, and analog data is processed using a bus/branch network model.

Figure 1.3 shows an interconnected system as viewed from a control center. The definition of internal system varies from case to case. Roughly speaking, it is the control area for which a specific control center is responsible. Ideally, this control area is observable, but this is not always the case. Parts of the control area can be permanently or temporarily unobservable, whereas parts of the network outside the control area, which are normally unobservable, can under certain circumstances be made observable by direct metering or data exchange.

Observable islands may be created due to the lack of appropriate telemetry, and these are handled with full state estimation including bad data analysis. State estimation can be extended to the rest of the internal system through the addition of pseudo-measurements based on load prediction and generation scheduling. In running state estimation for this augmented system, however, care must be taken to avoid corrupting the states estimated from telemetry data.

The external system can be replaced by an equivalent network or it can be maintained in the unreduced form. As suggested in Fig. 1.3, however, the best approach may be a compromise between the unreduced model (used for the inner external network) and the reduced equivalent (used for the outer external network). In this case, state estimation can be extended to the inner external system.

There are three principal approaches for building the model for the interconnected system:

- (a) the power flow based method,
- (b) the one-pass state estimation method, and
- (c) the two-pass state estimation method.

In the power flow based method, the external model (unreduced/reduced) is attached to the internal model by means of boundary matching injections. These injections are calculated by solving a power flow for the external network while treating the inner boundary buses as swing buses (with the values of the voltage magnitudes, V , and voltage angles, θ , being obtained by state estimation and used as target, or specified, values). The model thus obtained for the interconnected network correctly reproduces the conditions of the internal system, since possible external errors are absorbed by the boundary matching injections. These errors, however, may affect both contingency analysis and optimal power flow studies, because, although the base-case model is correct, the reactions of the external model to internal changes, such as a contingency, may be incorrect. One common criticism of this method relates the possible accumulation of external system modeling errors in boundary injections.

In the one-pass method, a single state estimation run is performed for the network as a whole. Power flow variables (e.g., target values for voltages and flows) and limits (e.g., MVAR limits) are treated as pseudo-measurements or as equality/inequality constraints. If a numerically robust estimator is used, these constraints can be enforced by assigning low weights to the corresponding pseudo-measurements. Alternatively, the set of pseudo-measurements can be made non-redundant, in which case even the presence of errors in the data will not corrupt the states estimated from telemetry, but special care must be taken with low impedance lines, since if neither the power flow in the low impedance branch nor the injections at its terminal nodes are “metered”, these variables may become numerically indeterminate, leading to abnormally large variances. One possible solution is to assign a pseudo-measurement to each low impedance branch; normally, this can be done without affecting overall measurement redundancy.

In the two-pass method, state estimation produces an initial estimate of the internal system states. The external model is then attached in two passes: (a) initial application of the power flow method, as above, to calculate branch power flows in the unobservable network; and (b) later state estimation using the estimated states as pseudo-measurements, for the internal system, with the power flows for the external system, run to match the two parts of the network model. The scheduled injections at the inner boundary nodes are used as target values. Zero injections for both the internal and external network are also treated as pseudo-measurements. Whenever available, external telemetry can also be included in the model. Depending on the relative weights used for external pseudo-measurements, the modeling errors will be dispersed throughout the external system, rather than accumulating at the inner boundary nodes.

1.3 CONVENTIONAL STATE ESTIMATION

As discussed above, the state estimator is used to build the model for the observable part of the network, and optionally to attach the external network model to it. With a proper redundancy level, state estimation can eliminate the effect of bad data and allow the temporary loss of measurement without significantly affecting the quality of the estimated values. State estimation is mainly used to filter redundant data, to eliminate incorrect measurements and to produce reliable state estimates, although, to a certain extent, it allows the determination of the power flows in parts of the network that are not directly metered. Both contingency analysis, optimal power flow, and dispatcher training simulator rely on the quality of the real-time network model obtained via state estimation, and even more so the new functions needed by the emerging energy markets.

To perform the tasks described above, a conventional state estimator normally involves the following set of state variables:

- Nodal voltage
 1. Voltage magnitude V_k at bus k
 2. Voltage angle θ_k at bus k
- Transformer turns ratio
 1. Magnitude of turns ratio t_{km} in transformer km
 2. Angle of turns ratio φ_{km} in transformer km

The state estimator provides estimates of the state variables based on a combination of measurements and pseudo-measurements of the following types:

- Measurements
 1. Voltage magnitude V_k at bus k
 2. Voltage angle θ_k at bus k
 3. Active power
 - (a) branch flow P_{km} in branch km
 - (b) branch-group flow $\sum P_{km}$ in a designated group of branches
 - (c) bus injection P_k into bus k
 4. Reactive power
 - (a) branch flow Q_{km} in branch km
 - (b) branch-group flow $\sum Q_{km}$ in a designated group of branches
 - (c) bus injection Q_k into bus k
 5. Current magnitude flow $|I_{km}|$ in branch km , and injection $|I_k|$ at bus k
 6. Magnitude of turns ratio t_{km} in transformer km

7. Angle of turns ratio φ_{km} in transformer km

■ Pseudo-measurements

1. Target voltage magnitude V_k^{sp} at bus k
2. Target voltage angle θ_k^{sp} at bus k
3. Active power
 - (a) target flow P_{km}^{sp} in branch km
 - (b) VAr limit Q_k^{lim} at bus k
4. Target reactive power flow Q_{km}^{sp} in branch km
5. Target current magnitude flow $|I_{km}^{sp}|$ in branch km , and injection $|I_k|$ at bus k
6. Tap limit t_{km}^{lim} in transformer km
7. Phase-shift limit φ_{km}^{lim} in phase-shifter km

State estimation can be formulated mathematically using the following constrained optimization problem:

$$\begin{aligned} & \text{Minimize} && f(\mathbf{z} - \mathbf{h}(\mathbf{x})) \\ & \text{subject to} && \mathbf{g}(\mathbf{x}) = \mathbf{0} \\ & && \mathbf{c}(\mathbf{x}) \leq \mathbf{0} \end{aligned}$$

where \mathbf{z} is a vector of measurements, $\mathbf{h}(\cdot)$ is a vector function relating measurements to state variables, $f(\cdot)$ is an objective function, and $\mathbf{g}(\cdot)$ and $\mathbf{c}(\cdot)$ are vector functions representing power flow quantities (Vectors and matrices will be denoted by boldface throughout.) In the weighted least-squares method the following quadratic objective function is used:

$$f(\mathbf{z} - \mathbf{h}(\mathbf{x})) = (\mathbf{z} - \mathbf{h}(\mathbf{x}))' \mathbf{W}(\mathbf{z} - \mathbf{h}(\mathbf{x}))$$

where the apostrophe denotes vector/matrix transposition, and \mathbf{W} is a diagonal matrix of weights (For telemetry, the inverses of measurement variances are normally used as weights). Equality and inequality constraints are used to represent target values and limits in the unobservable parts of the network. Non-quadratic objective functions such as the absolute value of the residuals can also be used, as well as a combination of quadratic/non-quadratic objective functions.

1.4 ECONOMY-SECURITY CONTROL

Economy-security control aims at maintaining a designated security level at minimum operational cost. Cost minimization involves not only a company's own generation and transmission, but also its transactions with other companies

connected to the power network. These economy-oriented actions normally conflict with network security and an integrated treatment of these two aspects of the problem is essential.

As discussed above, the notion of security is closely related to contingency analysis. The first step in contingency analysis is the establishment of a base-case power flow model of the interconnected network. This task is performed by the network model builder. Unlike state estimation, which can be run for the monitored part of a network only, even if external system information is not available, a contingency study affects and is affected by the external network. The network model should thus be extended to represent the relevant part of the external network.

The second step is the definition of a list of credible contingencies. Usually the vast majority of contingency cases will lead to no violations, so the list is shortened by contingency screening, which normally uses an approximate base-case power flow model to rank the contingencies by order of severity according to a performance (or severity) index. The final step in contingency analysis is contingency evaluation, which is performed by simulating each postulated contingency, in decreasing order of severity, using the base-case ac power flow model produced by the network model builder.

Contingency analysis can be conducted as a stand-alone function or as part of security constrained optimal power flow. Security constrained optimal power flow with post-contingency rescheduling can be formulated mathematically as follows:

$$\begin{array}{ll}
 \text{Minimize} & f(\mathbf{u}^0, \mathbf{x}^0) \\
 \text{subject to} & \mathbf{g}^i(\mathbf{u}^i, \mathbf{x}^i) = \mathbf{0} \\
 & \mathbf{h}^i(\mathbf{u}^i, \mathbf{x}^i) \leq \mathbf{0} \\
 & \mathbf{b}^{min} \leq \delta \mathbf{u}^i \leq \mathbf{b}^{max} \\
 & \mathbf{u}^i = \mathbf{u}^0 + \Delta \mathbf{u}^i + \delta \mathbf{u}^i \\
 \text{for} & i = 0, 1, 2, \dots, nc
 \end{array}$$

where the superscript i denotes a contingency case out of a list of nc contingencies, \mathbf{u} is a control vector, \mathbf{x} is the state vector, $\mathbf{g}(\cdot)$ and $\mathbf{h}(\cdot)$ are vector functions representing power flow quantities, $\delta \mathbf{u}^i$ are the EMS corrective control actions, \mathbf{b}^{min} and \mathbf{b}^{max} are the limits on these control actions, $\Delta \mathbf{u}^i$ is the response of the spontaneous automatic controls of the power system (local controls and controls that do not depend on EMS). If $\mathbf{b}^{min} = \mathbf{b}^{max} = \mathbf{0}$, no post-contingency rescheduling (corrective control) is taken into consideration, and the above problem is reduced to a regular optimal power flow problem with security constraints. If it is further assumed that $nc = 0$, the result is simple optimal power flow without security constraints.

Remarks:

- The time allowed for corrective actions (post-contingency rescheduling) and their rate of change dictate the limits \mathbf{b}^{min} and \mathbf{b}^{max} used in the formulation of the OPF with security constraints and corrective action.
- Level 1 security is achieved with $\mathbf{b}^{min} = \mathbf{b}^{max} = \mathbf{0}$.
- Level 2 security is obtained with the general formulation above, in which corrective actions are allowed, i.e., $\mathbf{b}^{min} \neq \mathbf{b}^{max}$.

1.5 GENERALIZED STATE ESTIMATION

Generalized state estimation is performed on a model in which parts of the network can be represented at the physical level, i.e., modeled in terms of bus-sections and switching devices. This allows the inclusion of measurements made on zero impedance branches and switching devices. The conventional states are augmented with new state variables. Observability analysis is extended to voltages at bus-sections and flows in switching devices, and if their values can be computed from the available measurements they are considered to be observable. The main advantage of the explicit modeling of switches, however, is facilitation of bad data analysis when topology errors are involved (incorrect status). Similarly, parameter estimation is performed by the addition of new states and pseudo-measurements.

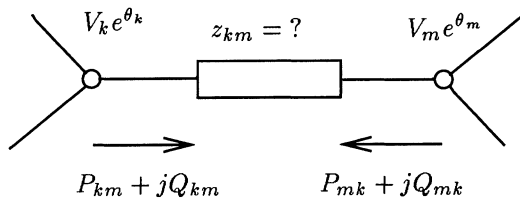


Figure 1.4. Branch with unknown impedance.

The idea behind the generalized state estimator can be understood with the help of the diagram shown in Fig. 1.4, where the branch impedance z_{km} is unknown and all branches incident to k and m are assumed to be known. In this case, Ohm's law cannot be used to relate the state variables $V_k e^{j\theta_k}$ and $V_m e^{j\theta_m}$, associated with the terminal nodes k and m , with the branch complex power flows $P_{km} + jQ_{km}$ and $P_{mk} + jQ_{mk}$. These power flows can be used as additional states, although they are not independent, since they are linked by the constraint

$$I_{km} + I_{mk} = 0 \quad \text{or} \quad \frac{P_{km} + jQ_{km}}{V_k e^{j\theta_k}} + \frac{P_{mk} + jQ_{mk}}{V_m e^{j\theta_m}} = 0$$

which can be expressed by the two following pseudo-measurements in a generalized estimator:

$$P_{km} V_m + \left(P_{mk} \cos(\theta_{km}) - Q_{mk} \sin(\theta_{km}) \right) V_k = 0 \quad (1.1)$$

$$Q_{km} V_m + \left(P_{mk} \sin(\theta_{km}) + Q_{mk} \cos(\theta_{km}) \right) V_k = 0 \quad (1.2)$$

A power injection measurement at node k can be expressed as the summation of the flow state variables $P_{km} + jQ_{km}$ and the flows in all other branches incident to k . Since only the flows in regular branches are functions of the nodal state variables, the unknown impedance will not form part of the measurement model. A similar analysis holds for power injection measurement at node m and power flow measurements made in the unknown impedance branch. Once the network state is estimated, the value of the unknown parameter can be computed from the estimates.

For a zero impedance branch, or a closed switch, the following is true:

$$V_k - V_m = 0 \quad (1.3)$$

$$\theta_k - \theta_m = 0 \quad (1.4)$$

In this case, the pseudo-measurements in Eqs. (1.1) and (1.2) above are reduced to the following:

$$P_{km} + P_{mk} = 0$$

$$Q_{km} + Q_{mk} = 0$$

Hence, rather than using P_{km} , Q_{km} , P_{mk} and Q_{mk} as additional states, only P_{km} and Q_{km} are used, and the pseudo-measurements $P_{km} + P_{mk} = 0$ and $Q_{km} + Q_{mk} = 0$ are replaced by Eqs. (1.3) and (1.4)

For open switches and switches with unknown status, the additional state variables are P_{km} and Q_{km} , as with closed switches. In the case of closed switches the pseudo-measurements are as follows:

$$P_{km} = 0$$

$$Q_{km} = 0$$

In the case of switches with unknown status, no pseudo-measurements are added, since there is no additional information. The application of the same principles to nodal shunt branches and shunt branches of a π equivalent model are discussed in Chap. 11. This approach allows modeling in the same central

state estimation algorithm the three main types of data, i.e., states, topology and parameters.

In the generalized state estimator, the set of state variables listed above is thus augmented to include the following states:

- Complex power flow
 1. Active power flow P_{km} and P_{mk}
 2. Reactive power flow Q_{km} and Q_{mk}

The set of measurements and pseudo-measurements is also augmented by adding the following:

- Measurements
 1. Active power flow P_{km}
 - (a) in switches
 - (b) in zero impedance branches
 - (c) in branches of unknown impedance
 2. Reactive power flow Q_{km}
 - (a) in switches
 - (b) in zero impedance branches
 - (c) in branches of unknown impedance
- Pseudo-measurements
 1. Voltage magnitude difference $V_k - V_m$ in closed switches
 2. Voltage angle difference $\theta_k - \theta_m$ in closed switches
 3. Active power flow P_{km} in an open switch
 4. Reactive power flow Q_{km} in an open switch
 5. Current magnitude difference $I_{km} + I_{mk}$
 6. Admittance difference ΔY_{km}^{sh} in π equivalent models

1.6 HISTORICAL NOTES AND REFERENCES

The weighted least-squares approach to problems of static state estimation in power networks based on a bus/branch model was introduced by Schweppe [1969-74]. This approach assumes that a bus-section/switching-device model has been processed by a network topology processor and that an error-free bus/branch model has been produced. Reviews of the state of the art in state estimation algorithms based on this modeling approach were published by Bose and Clements [1987] and Wu [1990]. A review of external system modeling was presented by Wu and Monticelli [1983]. More recently, the state of the art on

this subject was reviewed by the IEEE Task Force on External Network Modeling, chaired by Kato [1997]. The basic methodology for security control was created by Dy Liacco [1978]. Economy-security control functions and optimal power flows including contingency constraints and preventive control actions were discussed by Stott, Alsac, and Monticelli [1987]. A generalized state estimator with integrated state, status, and parameter estimation capabilities has recently been proposed by Alsac, Vempati, Stott, and Monticelli [1998]. The new role of state estimation and other advanced analytical functions in competitive energy markets was discussed by Shirmohammadi and coauthors [1998]. A comprehensive bibliography on state estimation from 1968-1989 was prepared by Coutto, Silva, and Falcão [1990].

References

- Alsac, O., Vempati, N., Stott, B., and Monticelli, A., "Generalized state estimation", IEEE Trans. on Power Systems, Vol. 13, No. 3, pp. 1069-1075, Aug. 1998.
- Bose, A. and Clements, K.A., "Real-time modeling of power networks", IEEE Proc., Special Issue on Computers in Power System Operations, Vol. 75, No. 12, pp. 1607-1622, Dec. 1987.
- Coutto, M.B., Silva, A.M.L., and Falcão, D.M., "Bibliography on power system state estimation (1968-1989)", IEEE Trans. Power Syst., Vol. 7, No. 3, pp. 950-961, Aug. 1990.
- Dy Liacco, "System security: The computer role", IEEE Spectrum, Vol. 16, No. 6, pp. 48-53, Jun. 1978.
- Kato, K., Chairman, IEEE Task Force, "External network modeling – recent practical experience", IEEE Trans. Power Syst., Vol. 9, No. 1, pp. 216-225, Nov. 1997
- Schweppe, F.C., Wildes, J., and Rom, D. , "Power system static state estimation: Parts I, II, and III", Power Industry Computer Conference, PICA, Denver, Colorado, June 1969.
- Schweppe, F.C. and Handschin, E., "Static state estimation in electric power systems", IEEE Proc., Vol. 62, pp. 972-983, July 1974.
- Shirmohammadi, D. et alli., "Transmission dispatch and congestion management in the emerging energy market structures", IEEE Trans. Power Syst., Vol. 13, No. 4, pp. 1466-1474, Nov. 1998
- Stott, B., Alsac, O., and Monticelli, A. , "Security analysis and optimization", IEEE Proc., Vol. 75, No. 12, pp. 1623-1644, Dec. 1987.
- Wu, F.F. and Monticelli, A., "A critical review on external network modeling for on-line security analysis", Int J Elec. Power and Energy Syst., Vol. 5, pp. 222-235, Oct. 1983.
- Wu, F.F. , "Power system state estimation: A survey", Int J Elec. Power and Energy Syst., Vol. 12, pp. 80-87, Jan. 1990.

2 LEAST-SQUARES AND MINIMUM NORM PROBLEMS

In this chapter, the solution and solvability conditions of over- and underdetermined systems of linear equations ($\mathbf{Ax} = \mathbf{b}$) and their application to circuit analysis problems are discussed. The related concepts of observability and controllability are reviewed. Nonlinear overdetermined system are also discussed.

2.1 INTRODUCTION

Most state estimation programs in practical use are formulated as overdetermined systems of non-linear equations and solved as weighted least-squares (WLS) problems. State estimation solvability conditions are determined by observability analysis: i.e., given a network and a set of analog measurements, observability analysis determines whether the state estimation problem is solvable or not. If such a problem is not solvable, this procedure identifies which states can be estimated. In this chapter the concept of observability in connection with weighted least squares estimators is thus introduced.

Controllability analysis, which is closely related to observability analysis, determines the solvability conditions of optimal power flows; these problems are formulated as optimization problems. Given a network and specified values for a set of controlled variables, controllability analysis determines whether it is possible to find a set of values for the control variables such that the desired values of the controlled variables are reached. It is thus necessary to understand

controllability in order to solve minimum norm problems, and the basic concept is developed here.

2.2 LINEAR LEAST-SQUARES PROBLEM

The least-squares solution provides one way of dealing with overdetermined systems of linear equations (systems with more equations than unknowns) of the type:

$$\mathbf{Ax} = \mathbf{b}, \quad (2.1)$$

where \mathbf{x} and \mathbf{b} are n and m vectors, respectively, with $n < m$. \mathbf{A} is an $m \times n$ matrix.

2.2.1 Problem Formulation

The least-squares problem for the overdetermined system in Eq. (2.1) is to find the n -vector \mathbf{x} for which the index $J(\mathbf{x})$, defined by

$$J(\mathbf{x}) = \frac{1}{2} (\mathbf{b} - \mathbf{Ax})' (\mathbf{b} - \mathbf{Ax}),$$

is minimized. The residual vector is defined as:

$$\mathbf{r} = \mathbf{b} - \mathbf{Ax}$$

The least-squares solution $\hat{\mathbf{x}}$ is such that

$$J(\hat{\mathbf{x}}) = \min_{\mathbf{x}} \mathbf{r}' \mathbf{r}$$

2.2.2 Normal Equation

The performance index $J(\mathbf{x})$ can be differentiated to obtain the first-order optimal conditions

$$\left. \frac{\partial J(\mathbf{x})}{\partial \mathbf{x}} \right|_{\mathbf{x}=\hat{\mathbf{x}}} = \mathbf{A}' \mathbf{A} \hat{\mathbf{x}} - \mathbf{A}' \mathbf{b} = \mathbf{0}$$

This yields the normal equation

$$\mathbf{A}' \mathbf{A} \hat{\mathbf{x}} = \mathbf{A}' \mathbf{b}$$

Solving this equation for $\hat{\mathbf{x}}$ thus gives:

$$\hat{\mathbf{x}} = (\mathbf{A}' \mathbf{A})^{-1} \mathbf{A}' \mathbf{b} \quad (2.2)$$

where $\hat{\mathbf{x}}$ is the state estimate.

The gain matrix \mathbf{G} and the pseudo-inverse of \mathbf{A} , are defined as

$$\mathbf{G} = \mathbf{A}' \mathbf{A} \quad (2.3)$$

$$\mathbf{A}^I = (\mathbf{A}' \mathbf{A})^{-1} \mathbf{A}'$$

It can then be verified that the pseudo-inverse \mathbf{A}^I satisfies the following conditions:

$$\mathbf{A} \mathbf{A}^I \mathbf{A} = \mathbf{A}$$

$$\mathbf{A}^I \mathbf{A} \mathbf{A}^I = \mathbf{A}^I$$

$$(\mathbf{A} \mathbf{A}^I)' = \mathbf{A} \mathbf{A}^I$$

$$(\mathbf{A}^I \mathbf{A})' = \mathbf{A}^I \mathbf{A}$$

which are known as the Moore-Penrose conditions. Finally, note that, if $m = n = \text{rank}(\mathbf{A})$, then $\mathbf{A}^I = \mathbf{A}^{-1}$.

Example 2.1:

Consider the circuit given in Fig. 2.1 where current injections I_1 and I_2 , and voltage E are unknown. Let $R_1=R_2=R_3=1.0 \Omega$. Meters A_1, A_2, A_3 , and V read 1.0 A, -3.2 A, 0.8 A, and 1.1 V, respectively. The problem is to determine the state of the circuit: nodal voltages v_1 and v_2 , and the voltage e across the voltage source. Positive sign conventions: positive currents flow from the node adjacent to the meter to the opposite node; positive voltage readings are indicated beside the voltage meter.

Kirchhoff's current law, KCL, allows the expression of the measured variables as functions of the state variables:

$$v_1 - v_2 - e = i_{1,2}^{meas} = 1.0$$

$$-v_1 = i_{3,1}^{meas} = -3.2$$

$$v_2 = i_{2,3}^{meas} = 0.8$$

$$e = e^{meas} = 1.1$$

In matrix form this gives:

$$\begin{pmatrix} 1 & -1 & -1 \\ -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ e \end{pmatrix} = \begin{pmatrix} 1.0 \\ -3.2 \\ 0.8 \\ 1.1 \end{pmatrix}$$

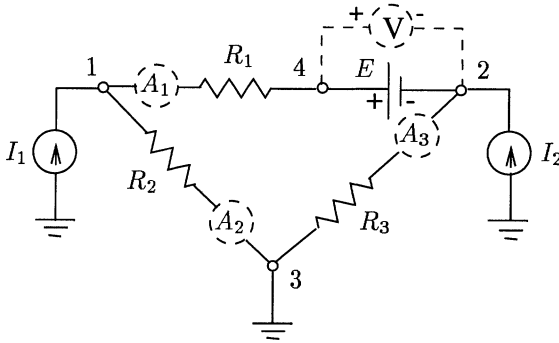


Figure 2.1. Circuit used to illustrate the formulation of the least-squares problem and observability analysis.

which is an overdetermined system (three unknowns and four equations) solvable as a least-squares problem. Since the right hand side consists of measured values, it may be affected by errors; hence, the four equations does not necessarily form a consistent set.

The gain matrix \mathbf{G} and its inverse for this problem are, respectively:

$$\mathbf{G} = \mathbf{A}'\mathbf{A} = \begin{pmatrix} 1 & -1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 & -1 \\ -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & 1 \\ -1 & 1 & 2 \end{pmatrix}$$

$$\mathbf{G}^{-1} = \frac{1}{4} \begin{pmatrix} 3 & 1 & 1 \\ 1 & 3 & -1 \\ 1 & -1 & 3 \end{pmatrix}$$

The pseudo-inverse of \mathbf{A} is given by:

$$\mathbf{A}^I = \mathbf{G}^{-1}\mathbf{A}' = \frac{1}{4} \begin{pmatrix} 3 & 1 & 1 \\ 1 & 3 & -1 \\ 1 & -1 & 3 \end{pmatrix} \begin{pmatrix} 1 & -1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{pmatrix}$$

$$\mathbf{A}^I = \frac{1}{4} \begin{pmatrix} 1 & -3 & 1 & 1 \\ -1 & -1 & 3 & -1 \\ -1 & -1 & -1 & 3 \end{pmatrix}$$

The least-squares estimate $\hat{\mathbf{x}}$ is then:

$$\begin{pmatrix} \hat{v}_1 \\ \hat{v}_2 \\ \hat{e} \end{pmatrix} = \mathbf{A}^I \begin{pmatrix} i_{12}^{meas} \\ i_{31}^{meas} \\ i_{23}^{meas} \\ e^{meas} \end{pmatrix} = \frac{1}{4} \begin{pmatrix} 1 & -3 & 1 & 1 \\ -1 & -1 & 3 & -1 \\ -1 & -1 & -1 & 3 \end{pmatrix} \begin{pmatrix} 1.0 \\ -3.2 \\ 0.8 \\ 1.1 \end{pmatrix} = \begin{pmatrix} 3.125 \\ 0.875 \\ 1.175 \end{pmatrix}$$

From these estimates, the estimated values of the measured variables follows:

$$\begin{pmatrix} \hat{i}_{12} \\ \hat{i}_{31} \\ \hat{i}_{23} \\ \hat{e} \end{pmatrix} = \mathbf{A}\hat{\mathbf{x}} = \begin{pmatrix} 1 & -1 & -1 \\ -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 3.125 \\ 0.875 \\ 1.175 \end{pmatrix} = \begin{pmatrix} 1.075 \\ -3.125 \\ 0.875 \\ 1.175 \end{pmatrix}$$

The measurement residuals, which give an idea of the quality of the measurements used in the estimation process, are as follows:

$$\mathbf{r} = \mathbf{b} - \mathbf{A}\hat{\mathbf{x}} = \begin{pmatrix} 1.0 \\ -3.2 \\ 0.8 \\ 1.1 \end{pmatrix} - \begin{pmatrix} 1.075 \\ -3.125 \\ 0.875 \\ 1.175 \end{pmatrix} = \begin{pmatrix} -0.075 \\ -0.075 \\ -0.075 \\ -0.075 \end{pmatrix}$$

2.3 LINEAR MINIMUM-NORM PROBLEM

Minimum-norm solutions are used to solve underdetermined systems of equations (systems with more unknowns than equations) of the type:

$$\mathbf{Ax} = \mathbf{b}, \quad (2.4)$$

where \mathbf{x} and \mathbf{b} are n and m vectors, respectively, with $n > m$. \mathbf{A} is an $m \times n$ matrix.

2.3.1 Problem Formulation

The minimum-norm problem for the underdetermined system in Eq. (2.4) is to find the n -vector \mathbf{x} such that $\mathbf{Ax} = \mathbf{b}$ and for which the 2-norm of \mathbf{x} , defined by

$$\|\mathbf{x}\|_2 = (\mathbf{x}'\mathbf{x})^{\frac{1}{2}},$$

is minimized. The minimum-norm solution is a vector $\hat{\mathbf{x}}$ that solves the following optimization problem:

$$\begin{array}{ll} \text{Minimize} & \|\mathbf{x}\|_2 \\ \text{subject to} & \mathbf{Ax} = \mathbf{b} \end{array} \quad (2.5)$$

2.3.2 Lagrangian Function

This optimization problem in Eq. (2.5) can be expressed by the following Lagrangian function:

$$\mathcal{L}(\mathbf{x}, \Lambda) = \frac{1}{2}\mathbf{x}'\mathbf{x} - \Lambda'(\mathbf{Ax} - \mathbf{b})$$

This function can be differentiated to obtain the following first-order optimal conditions

$$\left. \frac{\partial \mathcal{L}(\mathbf{x}, \Lambda)}{\partial \mathbf{x}} \right|_{(\hat{\mathbf{x}}; \hat{\Lambda})} = \hat{\mathbf{x}} - \mathbf{A}' \hat{\Lambda} = \mathbf{0} \rightarrow \hat{\mathbf{x}} = \mathbf{A}' \hat{\Lambda} \quad (2.6)$$

$$\left. \frac{\partial \mathcal{L}(\mathbf{x}, \Lambda)}{\partial \Lambda} \right|_{(\hat{\mathbf{x}}; \hat{\Lambda})} = \mathbf{A} \hat{\mathbf{x}} - \mathbf{b} = \mathbf{0} \rightarrow \mathbf{A} \hat{\mathbf{x}} = \mathbf{b} \quad (2.7)$$

Premultiplying Eq. (2.6) by \mathbf{A} and introducing Eq. (2.7) yields

$$\mathbf{A} \mathbf{A}' \hat{\Lambda} = \mathbf{b}$$

This yields the optimal Lagrange multipliers and the minimum-norm solution

$$\begin{aligned} \hat{\Lambda} &= (\mathbf{A} \mathbf{A}')^{-1} \mathbf{b} \\ \hat{\mathbf{x}} &= \mathbf{A}' (\mathbf{A} \mathbf{A}')^{-1} \mathbf{b} \end{aligned} \quad (2.8)$$

As with the least-squares problem, a pseudo-inverse matrix can also be defined:

$$\mathbf{A}^I = \mathbf{A}' (\mathbf{A} \mathbf{A}')^{-1}$$

Example 2.2:

Consider the circuit given in Fig. 2.2(a) where $R_1=R_2=R_3=1.0 \Omega$. I_1 , I_2 , and E are control variables. The solution sought are the values for the control variables, such that the constraints $i_{12}^{sp} = 1.0 \text{ A}$ and $i_{23}^{sp} = 1.0 \text{ A}$ are satisfied. Since there is more than one combination of the control variables that can satisfy the constraints on i_{12} and i_{23} , an attempt is made to find the minimum-norm solution.

The Norton equivalent circuit of Fig. 2.2(b) gives the following nodal equations:

$$\begin{pmatrix} R_1^{-1} + R_2^{-1} & -R_1^{-1} \\ -R_1^{-1} & R_1^{-1} + R_3^{-1} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} I_1 + E/R_1 \\ I_2 - E/R_1 \end{pmatrix}$$

where voltages v_1 and v_2 are state variables,

Introducing numerical data and inverting the coefficient matrix, yields:

$$\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} I_1 + E \\ I_2 - E \end{pmatrix}$$

The controlled variables can be expressed in terms of the state variables as follows:

$$\begin{pmatrix} i_{12} \\ i_{23} \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} + \begin{pmatrix} -E \\ 0 \end{pmatrix}$$

The final problem model is obtained by expressing the controlled variables as functions of the control variables:

$$\begin{pmatrix} i_{12} \\ i_{23} \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 1 & -1 & -1 \\ 1 & 2 & -1 \end{pmatrix} \begin{pmatrix} I_1 \\ I_2 \\ E \end{pmatrix}$$

This system is underdetermined (two equations and three unknowns) and can be solved as a minimum-norm problem. Matrix \mathbf{AA}' and its inverse are, respectively:

$$\mathbf{AA}' = \frac{1}{9} \begin{pmatrix} 1 & -1 & -1 \\ 1 & 2 & -1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -1 & 2 \\ -1 & -1 \end{pmatrix} = \frac{1}{9} \begin{pmatrix} 3 & 0 \\ 0 & 6 \end{pmatrix}$$

$$(\mathbf{AA}')^{-1} = \begin{pmatrix} 3 & 0 \\ 0 & 3/2 \end{pmatrix}$$

The pseudo-inverse of \mathbf{A} is given by:

$$\mathbf{A}^I = \mathbf{A}'(\mathbf{AA}')^{-1} = \frac{1}{3} \begin{pmatrix} 1 & 1 \\ -1 & 2 \\ -1 & -1 \end{pmatrix} \begin{pmatrix} 3 & 0 \\ 0 & 3/2 \end{pmatrix} = \begin{pmatrix} 1 & 1/2 \\ -1 & 1 \\ -1 & -1/2 \end{pmatrix}$$

The minimum-norm solution is then given by:

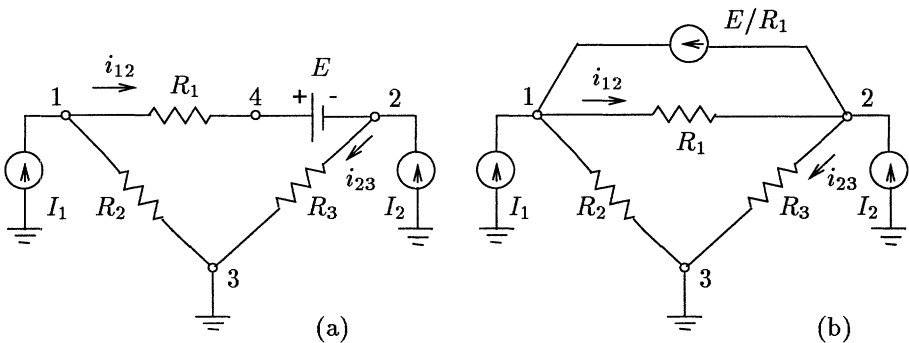


Figure 2.2. Circuit used to illustrate the formulation of a minimum-norm problem and controllability analysis. Control variables: I_1 , I_2 , and E . Controlled variables: i_{12} and i_{23} . (a) Original circuit. (b) Norton equivalent.

$$\begin{pmatrix} I_1 \\ I_2 \\ E \end{pmatrix} = \mathbf{A}^T \begin{pmatrix} i_{12}^{sp} \\ i_{23}^{sp} \end{pmatrix} = \begin{pmatrix} 1 & 1/2 \\ -1 & 1 \\ -1 & -1/2 \end{pmatrix} \begin{pmatrix} 1.0 \\ 1.0 \end{pmatrix} = \begin{pmatrix} 3/2 \\ 0 \\ -3/2 \end{pmatrix}$$

2.4 OBSERVABILITY AND CONTROLLABILITY

In both examples above, the problems are solvable (an observable case in Example 2.1, and a controllable case in Example 2.2). The present section discusses cases in which solvability is critical.

2.4.1 Least-Squares Problem

In the least-squares problem examined above, it was assumed that the $m \times n$ coefficient matrix \mathbf{A} , ($m > n$) has a full rank, $\text{rank}(\mathbf{A})=n$. If so, the gain matrix $\mathbf{A}'\mathbf{A}$ is nonsingular and the least-squares estimate can be obtained from Eq. (2.2). When this happens the circuit is considered to be observable, i.e., it is possible to estimate the circuit state from the available measurements. In the case studied in Example 2.1 (Fig. 2.1), even after the removal of one of the four meters (either A_1 , A_2 , A_3 , or V) the circuit remains observable. The simultaneous removal of two or more meters, however, would make the circuit unobservable, or only partially observable. If, for example, meters A_2 and V are removed, only states v_1 and v_2 remain observable; state e would become unobservable (and as a consequence it is impossible to calculate current i_{12}).

Figure 2.3 shows a circuit in which the resistances are known, but the voltages across the voltage sources are unknown. In case (a) there are two voltmeters that read the voltages across the two resistances. In case (b) there is only one voltmeter across resistance R_1 . In the former the circuit is observable because all the relevant quantities can be calculated from the available metered values. For example: the current through R_1 and E_2 can be calculated from meter V_1 ; and the current through R_2 from meter V_2 ; from these two currents it is possible to determine the current across voltage source E_1 . In the second case, however, only the current through R_1 and E_2 can be calculated; the current distribution between R_2 and E_1 cannot be determined, and so the circuit is only partially observable.

Example 2.3:

Consider the circuit shown in Fig. 2.3(a). Let the state be the voltages at Nodes 1 and 2. The voltage at Node 3 is considered to be a reference value, i.e., $v_3 = 0$. The model of the circuit expresses the measured quantities V_1 and V_2 in terms of the state variables v_1 and v_2 :

$$\begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} V_1^{meas} \\ V_2^{meas} \end{pmatrix}$$

Matrix \mathbf{A} has a full rank, so $\mathbf{A}'\mathbf{A}$ is non singular; thus the circuit is observable.

Now consider the circuit shown in Fig. 2.3(b) with the same set of state variables. In this case the model is given by:

$$(1 \quad -1) \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = (V_1^{meas})$$

In this case $\text{rank}(\mathbf{A}) = 1$ and so $\mathbf{A}'\mathbf{A}$ is singular; thus the circuit is unobservable.

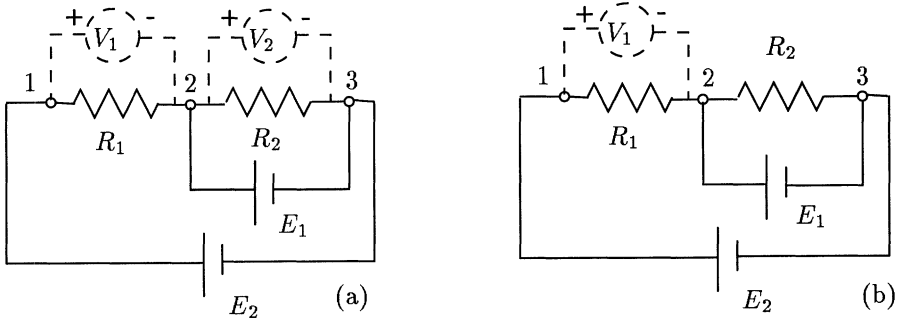


Figure 2.3. Circuit observability. (a) Observable case. (b) Unobservable case. V_1 and V_2 are metered values. R_1 and R_2 are known. E_1 and E_2 are unknown.

2.4.2 Minimum-Norm Problem

In the minimum-norm problem formulated in Sec. 2.3 it was assumed that the $m \times n$ coefficient matrix \mathbf{A} ($m < n$) has a full rank, $\text{rank}(\mathbf{A})=m$. As a consequence, matrix $\mathbf{A}\mathbf{A}'$ is nonsingular, and the minimum-norm solution can be obtained from Eq. (2.8). When this happens, the case is considered to be controllable, i.e., it is possible to find values for the control variables such that the specified constraints (set of target values) are satisfied. In Example 2.2 (Fig. 2.2), even after the removal of controller I_1 , the two specified variables remain controllable; this would also be true if only controller E was removed. The removal of controller I_2 , however, would make the problem uncontrollable.

Fig. 2.4 presents a circuit in which the resistances are known. Currents i_1 and i_2 are the controlled variables (with target values). In case (a) E_1 and E_2 are the control variables. In case (b) E is the control variable. In the first case the specified currents are both controllable, since values for both E_1 and E_2 can be found such that $i_1 = i_1^{sp}$ and $i_2 = i_2^{sp}$. In the second case there is a single control variable, and only one of the two constraints can be satisfied (either $i_1 = i_1^{sp}$ or $i_2 = i_2^{sp}$).

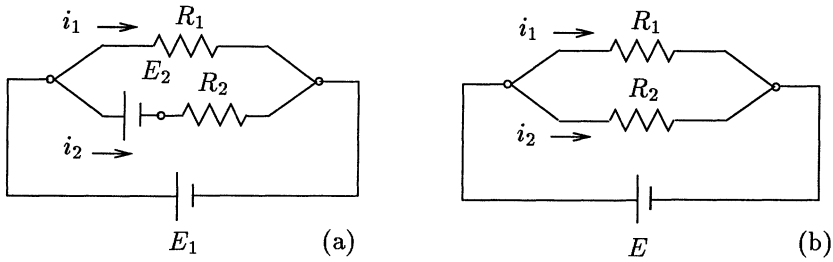


Figure 2.4. Circuit controllability. (a) Controllable case. (b) Uncontrollable case. E_1 and E_2 are control variables. i_1 and i_2 are controlled variables. R_1 and R_2 are known.

Example 2.4:

Assume that in the circuit shown in Fig. 2.4(a) $R_1 = R_2 = 1\Omega$. The model of the circuit expresses the controlled quantities i_1 and i_2 in terms of the control variables E_1 and E_2 :

$$\begin{pmatrix} 1 & 0 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} E_1 \\ E_2 \end{pmatrix} = \begin{pmatrix} i_1^{sp} \\ i_2^{sp} \end{pmatrix}$$

Matrix \mathbf{A} has a full rank so $\mathbf{A}\mathbf{A}'$ is non singular; thus the circuit is controllable.

Now consider the circuit shown in Fig. 2.4(b). for this example the model is given by:

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix} (E) = \begin{pmatrix} i_1^{sp} \\ i_2^{sp} \end{pmatrix}$$

In this case $\text{rank}(\mathbf{A}) = 1$, so $\mathbf{A}\mathbf{A}'$ is singular; thus the circuit is not controllable.

2.5 GEOMETRIC INTERPRETATION

2.5.1 Range and Null Spaces

Let \mathbf{A} be an $m \times n$ matrix. Define the range space of \mathbf{A} as

$$R(\mathbf{A}) = \{ \mathbf{y} \in \mathbb{R}^m \mid \mathbf{y} = \mathbf{A}\mathbf{x} \text{ for some } \mathbf{x} \in \mathbb{R}^n \}$$

The rank of matrix \mathbf{A} is defined as

$$\text{rank}(\mathbf{A}) = \dim[R(\mathbf{A})]$$

The rank of \mathbf{A} is equal to the number of independent rows (or columns); i.e., $\text{rank}(\mathbf{A}) = \text{rank}(\mathbf{A}')$. The null space of \mathbf{A} is defined as

$$N(\mathbf{A}) = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{A}\mathbf{x} = \mathbf{0}\}$$

The orthogonal complement of a subspace $S \subset \mathbb{R}^m$ is defined as

$$S^\perp = \{\mathbf{y} \in \mathbb{R}^m \mid \mathbf{y}'\mathbf{x} = 0 \text{ for all } \mathbf{x} \in S\}$$

It can now be shown that:

$$R(\mathbf{A})^\perp = N(\mathbf{A}')$$

Example 2.5:

Consider the 2×3 matrix

$$\mathbf{A} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{pmatrix}$$

The rank of \mathbf{A} is $n = 2$, and the range space is the following:

$$R(\mathbf{A}) = \{\mathbf{y} \in \mathbb{R}^3 \mid \mathbf{y} = \mathbf{A}\mathbf{x} \text{ for some } \mathbf{x} \in \mathbb{R}^2\}$$

$R(\mathbf{A})$ is two dimensional and hence a plane. This is illustrated in Fig. 2.5 (in this case $\text{rank}(\mathbf{A}) = \dim[R(\mathbf{A})] = 2$). While only two of the three rows are linearly independent, both columns of \mathbf{A} are linearly independent. Then the null space of \mathbf{A}' is

$$N(\mathbf{A}') = \{\mathbf{x} \in \mathbb{R}^3 \mid \mathbf{A}'\mathbf{x} = \mathbf{0}\}$$

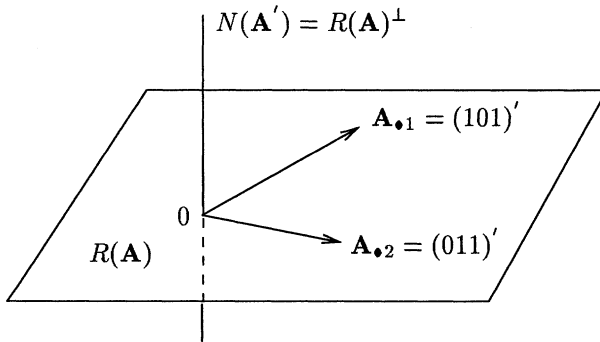


Figure 2.5. Orthogonal subspaces $R(\mathbf{A})$ and $N(\mathbf{A}')$ for matrix \mathbf{A} of Example 2.5.

If $\mathbf{x} \in N(\mathbf{A}')$, then \mathbf{x} is orthogonal to the vectors represented by the columns of \mathbf{A} (or the rows of \mathbf{A}') as illustrated in Fig. 2.5, that is, $\mathbf{x} \perp \mathbf{A}_{\bullet 1}$ and $\mathbf{x} \perp \mathbf{A}_{\bullet 2}$, where $\mathbf{A}_{\bullet 1}$ and $\mathbf{A}_{\bullet 2}$ are the columns of \mathbf{A} .

2.5.2 Least-Squares Problem

Consider the least-squares problem

$$\text{Min}_{\mathbf{x}} \|\mathbf{r}\|_2$$

where $\mathbf{r} = \mathbf{b} - \mathbf{A}\mathbf{x}$ is the residual vector. Figure 2.6(a) shows vector \mathbf{b} , the range space of \mathbf{A} , and the null space of \mathbf{A}' (orthogonal to $R(\mathbf{A})$). Since vector $\mathbf{b} - \mathbf{A}\hat{\mathbf{x}}$ has minimum-norm, the solution of the least-squares problem can be seen as the projection of \mathbf{b} onto $R(\mathbf{A})$:

$$\mathbf{A}\hat{\mathbf{x}} = \mathbf{A}\mathbf{A}'\mathbf{b} = P_A\mathbf{b}$$

2.5.3 Minimum-Norm Problem

Consider the minimum-norm problem $\text{min}_{\mathbf{x}} \|\mathbf{x}\|_2$ such that $\mathbf{A}\mathbf{x} = \mathbf{b}$. Figure 2.6(b) shows vector \mathbf{b} , the range space of \mathbf{A}' , and the null space of \mathbf{A} (orthogonal to $R(\mathbf{A}')$). The solution of the minimum-norm problem can be seen as the projection of \mathbf{b} onto the range space of \mathbf{A}' , that is:

$$\hat{\mathbf{x}} = \mathbf{A}'\mathbf{b} = P_{A'}\mathbf{b}$$

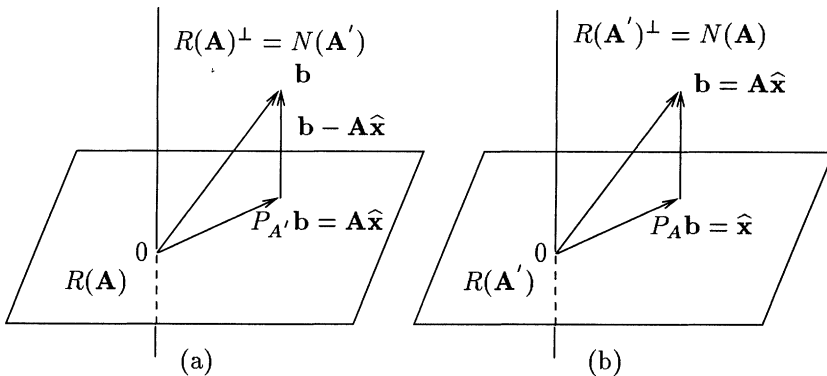


Figure 2.6. (a) Least-squares solution: orthogonal projection of \mathbf{b} onto $R(\mathbf{A})$. (b) Minimum-norm vector: orthogonal projection of \mathbf{b} onto $R(\mathbf{A}')$.

Example 2.6:

Consider the overdetermined system of linear equations,

$$\begin{pmatrix} 1 \\ 2 \end{pmatrix} (x_1) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

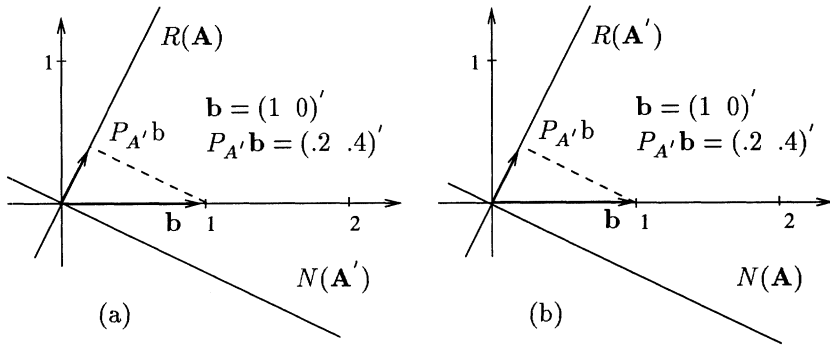


Figure 2.7. Geometric interpretation of Example 2.6. (a) Least-squares solution. (b) Minimum-norm solution.

and the underdetermined system:

$$(1 \ 2) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = (1)$$

The least-squares solution of the overdetermined system is given by:

$$\hat{\mathbf{x}} = (\hat{x}_1) = \left((1 \ 2) \begin{pmatrix} 1 \\ 2 \end{pmatrix} \right)^{-1} (1 \ 2) \begin{pmatrix} 1 \\ 0 \end{pmatrix} = (.2)$$

Then the projection of \mathbf{b} onto $R(\mathbf{A})$ is given by:

$$P_A \mathbf{b} = \mathbf{A} \hat{\mathbf{x}} = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \hat{x} = \begin{pmatrix} .2 \\ .4 \end{pmatrix}$$

as illustrated in Fig. 2.7(a).

The minimum-norm solution of the underdetermined system is then written as follows:

$$\hat{\mathbf{x}} = \begin{pmatrix} \hat{x}_1 \\ \hat{x}_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \left((1 \ 2) \begin{pmatrix} 1 \\ 2 \end{pmatrix} \right)^{-1} (1) = \begin{pmatrix} .2 \\ .4 \end{pmatrix}$$

The projection of \mathbf{b} onto $R(\mathbf{A}')$ is given by:

$$P_{A'} \mathbf{b} = \hat{\mathbf{x}} = \begin{pmatrix} .2 \\ .4 \end{pmatrix}$$

as illustrated in Fig. 2.7(b).

2.6 OVERDETERMINED NONLINEAR MODELS

Consider the nonlinear measurement model

$$\mathbf{z} = \mathbf{h}(\mathbf{x}) + \mathbf{e}$$

where \mathbf{z} is the measurement vector (m -vector), \mathbf{x} is the true state vector (n -vector, $n < m$), $\mathbf{h}(\cdot)$ is a nonlinear vector function relating measurements to states (m -vector), and \mathbf{e} is the measurement error vector (m -vector). The elements of \mathbf{e} are assumed to have zero mean; the corresponding variance matrix is denoted by \mathbf{R}_z .

The state estimation problem can be formulated as a minimization of

$$J(\mathbf{x}) = \frac{1}{2}(\mathbf{z} - \mathbf{h}(\mathbf{x}))' \mathbf{R}_z^{-1}(\mathbf{z} - \mathbf{h}(\mathbf{x})) \quad (2.9)$$

This weighted least-squares problem, WLS, extends the formulation presented in Subsec. 2.2.1 since now the inverse of the covariance matrix \mathbf{R}_z is used as a weighting matrix. The state estimate $\hat{\mathbf{x}}$ is obtained by the following iterative procedure:

$$\begin{aligned} \mathbf{G}(\mathbf{x}^\nu) \Delta \mathbf{x}^\nu &= -\mathbf{g}(\mathbf{x}^\nu) \\ \mathbf{x}^{\nu+1} &= \mathbf{x}^\nu + \Delta \mathbf{x}^\nu \end{aligned}$$

where $\mathbf{g}(\mathbf{x})$ is the gradient of $J(\mathbf{x})$, and $\mathbf{G}(\mathbf{x})$ is a gain matrix that depends on the method used to solve minimization problem (Gauss Newton or Newton Raphson).

2.6.1 Gauss Newton Method

Taylor expansion provides an approximation of the nonlinear vector function $\mathbf{h}(\mathbf{x})$:

$$\mathbf{h}(\mathbf{x} + \Delta \mathbf{x}) \simeq \mathbf{h}(\mathbf{x}) + \mathbf{H}(\mathbf{x}) \Delta \mathbf{x}$$

The minimization problem in Eq. (2.9) can then be rewritten as follows:

$$J(\Delta \mathbf{x}) = (\Delta \mathbf{z} - \mathbf{H}(\mathbf{x}) \Delta \mathbf{x})' \mathbf{R}_z^{-1} (\Delta \mathbf{z} - \mathbf{H}(\mathbf{x}) \Delta \mathbf{x})$$

where $\Delta \mathbf{z} = \mathbf{z} - \mathbf{h}(\mathbf{x})$ and $\mathbf{H}(\mathbf{x}) = \partial \mathbf{h} / \partial \mathbf{x}$ (the Jacobian matrix). The first-order optimal condition is

$$\frac{\partial J(\Delta \mathbf{x})}{\partial \Delta \mathbf{x}} = -\mathbf{H}'(\mathbf{x}) \mathbf{R}_z^{-1} (\Delta \mathbf{z} - \mathbf{H}(\mathbf{x}) \Delta \mathbf{x}) = 0$$

Hence, the linear least-squares solution can be expressed as follows:

$$\Delta \mathbf{x} = (\mathbf{H}'(\mathbf{x}) \mathbf{R}_z^{-1} \mathbf{H}(\mathbf{x}))^{-1} \mathbf{H}'(\mathbf{x}) \mathbf{R}_z^{-1} \Delta \mathbf{z}(\mathbf{x}) \quad (2.10)$$

2.6.2 Newton Raphson Method

The derivation of the the Gauss Newton method involved the transformation of the original minimization problem into a linear least-squares problem by linearizing the vector function $\mathbf{h}(\mathbf{x})$. For the Newton Raphson method, however, the optimality conditions are applied directly to the performance index $J(\mathbf{x})$ as expressed in Eq. (2.9).

First, $J(\mathbf{x})$ is rewritten as follows:

$$J(\mathbf{x}) = \frac{1}{2} \sum_{j=1}^m \left(\frac{z_j - h_j(\mathbf{x})}{\sigma_j} \right)^2$$

where σ_j is the (j, j) th element of the measurement error covariance matrix, \mathbf{R}_z . The first-order optimal condition for this model is

$$\mathbf{g}(\mathbf{x}) = \frac{\partial J(\mathbf{x})}{\partial \mathbf{x}} = - \sum_{j=1}^m \left(\frac{z_j - h_j(\mathbf{x})}{\sigma_j} \right) \frac{\partial h_j(\mathbf{x})}{\partial \mathbf{x}} = \mathbf{0}$$

where $\mathbf{g}(\mathbf{x})$ denotes the gradient of $J(\mathbf{x})$. The root of the nonlinear equation $\mathbf{g}(\mathbf{x}) = \mathbf{0}$ can be found using the Newton Raphson method. Taylor expansion approximates the gradient function:

$$bf\mathbf{g}(\mathbf{x} + \Delta\mathbf{x}) \simeq \mathbf{g}(\mathbf{x}) + \mathbf{G}(\mathbf{x}) \Delta\mathbf{x}$$

where $\mathbf{G}(\mathbf{x})$ is the Jacobian matrix of $\mathbf{g}(\mathbf{x})$ (or the Hessian matrix of $J(\mathbf{x})$):

$$\mathbf{G}(\mathbf{x}) = \frac{\partial \mathbf{g}(\mathbf{x})}{\partial \mathbf{x}} = \frac{\partial^2 J(\mathbf{x})}{\partial \mathbf{x}^2} = \sum_{j=1}^m \sigma_j^{-1} \frac{\partial h_j(\mathbf{x})}{\partial \mathbf{x}} \left(\frac{\partial h_j(\mathbf{x})}{\partial \mathbf{x}} \right)' - \sum_{j=1}^m \sigma_j^{-1} \Delta z_j \frac{\partial^2 h_j(\mathbf{x})}{\partial \mathbf{x}^2}$$

The least-squares solution can be expressed as follows:

$$\Delta\mathbf{x} = \mathbf{G}^{-1}(\mathbf{x}) \mathbf{H}'(\mathbf{x}) \mathbf{R}_z^{-1} \Delta\mathbf{z}(\mathbf{x})$$

Thus, since

$$\sum_{j=1}^m \sigma_j^{-1} \frac{\partial h_j(\mathbf{x})}{\partial \mathbf{x}} \left(\frac{\partial h_j(\mathbf{x})}{\partial \mathbf{x}} \right)' = \mathbf{H}'(\mathbf{x}) \mathbf{R}_z^{-1} \mathbf{H}(\mathbf{x})$$

the state vector correction can be written as follows:

$$\Delta\mathbf{x} = \left(\mathbf{H}'(\mathbf{x}) \mathbf{R}_z^{-1} \mathbf{H}(\mathbf{x}) - \sum_{j=1}^m \sigma_j^{-1} \Delta z_j \frac{\partial^2 h_j(\mathbf{x})}{\partial \mathbf{x}^2} \right)^{-1} \mathbf{H}'(\mathbf{x}) \mathbf{R}_z^{-1} \Delta\mathbf{z} \quad (2.11)$$

Note that if the term that depends on the second derivatives is ignored, Eq. (2.11) reduces to Eq. (2.10). Of course, second derivatives appear only in non-linear models. Even in these cases their effect on state estimation convergence will depend on how well the proposed model fits the data, that is, it depends on how close $\Delta \mathbf{z}$ gets to zero. In most situations, the impact of second derivatives on the convergence of static power system state estimation is negligible.

Example 2.7:

Three measurements were made in the dc network shown in Fig. 2.8: the current entering Node 1 ($I_1^{meas} = 1.19$ A); the power dissipated by the shunt resistance connected to Node 2 ($P_2^{meas} = 0.1$ W); and the voltage at Node 1 ($V_1^{meas} = 1.00$ V). The state of the system is defined as the nodal voltages V_1 and V_2 . For simplicity, a unit measurement error variance of $\sigma = 1$ is considered; hence the covariance matrix \mathbf{R}_z is a unit matrix.

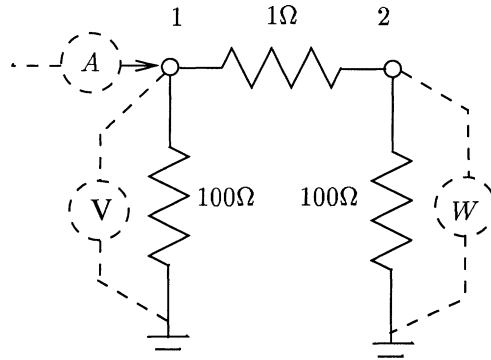


Figure 2.8. Network used in Example 2.7.

In this case the measurement model can be written as follows

$$\begin{aligned} I_1^{meas} &= 1.01V_1 - V_2 + e_1 \\ V_1^{meas} &= V_1 + e_2 \\ P_2^{meas} &= V_1V_2 - V_2^2 + e_3 \end{aligned}$$

where e_i are the measurement errors. Hence, the vector function $\mathbf{h}(V_1, V_2)$ and the corresponding Jacobian matrix, $\mathbf{H}(V_1, V_2)$, are as follows

$$\mathbf{h}(V_1, V_2) = \begin{pmatrix} h_1(V_1, V_2) \\ h_2(V_1, V_2) \\ h_3(V_1, V_2) \end{pmatrix} = \begin{pmatrix} 1.01V_1 - V_2 \\ V_1 \\ V_1V_2 - 1.01V_2^2 \end{pmatrix}$$

$$\mathbf{H}(V_1, V_2) = \begin{pmatrix} \partial h_1/\partial V_1 & \partial h_1/\partial V_2 \\ \partial h_2/\partial V_1 & \partial h_2/\partial V_2 \\ \partial h_3/\partial V_1 & \partial h_3/\partial V_2 \end{pmatrix} = \begin{pmatrix} 1.01 & -1.00 \\ 1.00 & 0 \\ V_2 & V_1 - 2.02V_2 \end{pmatrix}$$

The Hessian matrix obtained from the Newton Raphson method is as follows:

$$\begin{aligned} \mathbf{G}(V_1, V_2) &= \mathbf{H}'\mathbf{H} - (P_2^{meas} - h_3(V_1, V_2)) \begin{pmatrix} \partial h_3^2/\partial V_1^2 & \partial h_3^2/\partial V_1 V_2 \\ \partial h_3^2/\partial V_2 V_1 & \partial h_3^2/\partial V_2^2 \end{pmatrix} \\ &= \begin{pmatrix} 2.02 + V_2^2 & -1 + V_1 V_2 - 2.02V_2^2 \\ -1 + V_1 V_2 - 2.02V_2^2 & 1 + V_1^2 - 4.04V_1 V_2 + 4.08V_2^2 \end{pmatrix} + \\ &\quad -(P_2^{meas} - V_1 V_2 + 1.01V_2^2) \begin{pmatrix} 0 & 1 \\ 1 & -2.02 \end{pmatrix} \end{aligned}$$

If the initial state estimate is $V_1 = 1.0$ and $V_2 = 1.0$, the first iteration for both the Gauss Newton and the Newton Raphson methods is as follows:

Gauss Newton Iteration:

In this case, the state estimate update, (Eq. (2.10)), is

$$\begin{pmatrix} \Delta V_1 \\ \Delta V_2 \end{pmatrix} = \begin{pmatrix} 3.02 & -2.02 \\ -2.02 & 2.00 \end{pmatrix}^{-1} \begin{pmatrix} 1.01 & 1.00 & 1.00 \\ -1.00 & 0 & -1.02 \end{pmatrix} \begin{pmatrix} 1.19 - 1.01 \\ 1.00 - 1.00 \\ 0.10 - 0.01 \end{pmatrix}$$

$$\begin{pmatrix} \Delta V_1 \\ \Delta V_2 \end{pmatrix} = \begin{pmatrix} 0.0000 \\ -0.1084 \end{pmatrix}$$

Newton Raphson Iteration:

In this case, the state estimate update, (Eq. (2.11)), is

$$\begin{pmatrix} \Delta V_1 \\ \Delta V_2 \end{pmatrix} = \left[\begin{pmatrix} 3.02 & -2.02 \\ -2.02 & 2.00 \end{pmatrix} - 0.09 \begin{pmatrix} 0 & 1.00 \\ 1.00 & -2.02 \end{pmatrix} \right]^{-1}$$

$$\times \begin{pmatrix} 1.01 & 1.00 & 1.00 \\ -1.00 & 0 & -1.02 \end{pmatrix} \begin{pmatrix} 1.19 - 1.01 \\ 1.00 - 1.00 \\ 0.10 - 0.01 \end{pmatrix}$$

$$\begin{pmatrix} \Delta V_1 \\ \Delta V_2 \end{pmatrix} = \begin{pmatrix} 0.0109 \\ -0.0874 \end{pmatrix}$$

The iterative process is summarized in the following table. Notice that in this example the set of measurements is consistent (no bad data). Also, all

the parameters are correct. As could be expected, the Newton Raphson method presents quadratic convergence. Although the Gauss Newton method furnishes a convergence which is nearly as good as that obtained by the Newton Raphson method, it is not quadratic.

ν	<i>Gauss Newton</i>		<i>Newton Raphson</i>	
	ΔV_1	ΔV_2	ΔV_1	ΔV_2
1	0.0000175	-0.1083930	0.0109189	-0.0874364
2	-0.0005244	-0.0053171	-0.0101129	-0.0255036
3	0.0001086	0.0000651	-0.0001897	-0.0006930
4	0.0009777	0.0000003	-0.0000002	-0.0000007
5	0.0000068	0.0000111	0.0000000	0.0000000
6	0.0000000	-0.0000001	0.0000000	0.0000000

The following table shows the results obtained for this sample system, but with voltage measurement $V_1^{meas} = 0.6$ V, i.e., a measurement set with inconsistent (bad data). In this case the converged solution will present relatively large values for the elements of Δz . The Newton Raphson method maintains its quadratic performance. The convergence with the Gauss Newton method is affected negatively, but with a moderate tolerance, the method still performs remarkably well.

ν	<i>Gauss Newton</i>		<i>Newton Raphson</i>	
	ΔV_1	ΔV_2	ΔV_1	ΔV_2
1	-0.3998040	-0.5061760	-0.3906650	-0.4672600
2	0.0035555	-0.0191454	-0.0037684	-0.0522969
3	0.0010484	0.0024795	-0.0008267	-0.0034745
4	-0.0000677	-0.0002171	-0.0000013	-0.0000085
5	0.0000070	0.0000209	0.0000000	0.0000000
6	-0.0000006	-0.0000020	0.0000000	0.0000000

The final table shows a case where a parameter is wrong. The shunt resistance connected at Node 1 has been considered to be 1Ω (one hundred times smaller than the correct value). This shows that the effect of a wrong parameter on convergence is similar to the effect caused by analog errors, as presented above.

ν	<i>Gauss Newton</i>		<i>Newton Raphson</i>	
	ΔV_1	ΔV_2	ΔV_1	ΔV_2
1	-0.4013470	-0.2172450	-0.4122400	-0.2304920
2	0.0139478	0.0056331	-0.0061932	0.0127182
3	-0.0045380	-0.0096975	-0.0002772	-0.0006445
4	0.0016842	0.0043020	-0.0000002	-0.0000005
5	-0.0008293	-0.0020834	0.0000000	0.0000000
6	0.0003957	0.0009948	0.0000000	0.0000000

Remarks: The quadratic convergence of the Newton Raphson method applied to the power flow problem is discussed in Subsec. 4.5.2 in Chap. 4.

2.7 HISTORICAL NOTES AND REFERENCES

Power network static state estimation was originally formulated as a weighted least-squares problem by Schweppe, Wildes, and Rom [1969], and the concept of topological observability in power network state estimation was introduced as a partial requirement for state estimation solvability by Clements and Wollenberg [1975] and further developed by Krumpholz, Clements, and Davis [1980]. Practical solvability also depends on numerical aspects of state estimation, and a numerical approach to observability analysis which could take both topological and numerical aspects into consideration was suggested by Monticelli and Wu [1985]. A comprehensive treatment of numerical methods used to solve least-squares and minimum-norm problem is presented by Golub and Van Loan [1983]. The use of sparse matrix techniques to solve large linear systems of equations was addressed by Duff and Erisman [1992], while both the Gauss and the Givens methods are described by Watkins [1991]. Controllability analysis in optimal power flows has not been studied as thoroughly as observability analysis in state estimation. Some of the deficiencies of the treatment of optimal power flows pointed out by Tinney, Bright, Demaree, and Hugues [1988] and by Monticelli and Liu [1992] are in fact issues related to controllability. (The formulation used in Example 2.2 is similar to the one proposed by Stott and Robson [1978] for LP based optimal power flows.) The effect of the second order derivatives of the performance index $J(\cdot)$ on state estimation convergence rates was discussed by Van Amerongen [1995]. A detailed discussion about the quadratic convergence property can be found in Chap. 6 of the book by Dahlquist and Björck [1974]. The sparse tableau formulation discussed in Problems 7 and 8 was introduced by Hachtel [1976].

2.8 PROBLEMS

- 1. Figures 2.9(a) and (b) describe a situation with an unknown current I and resistance $R_1 = 1.0 \Omega$ and $R_2 = 2.0 \Omega$. Meter A reads 4.7 A and meter

V reads 5.2 V. Choose state variables for both cases and perform state estimation using the least-squares approach. (The positive conventions are the same ones used in Example 2.1: positive currents flow from the node adjacent to the meter to the opposite node; the positive sign convention for voltages is indicated beside the voltage meter.)

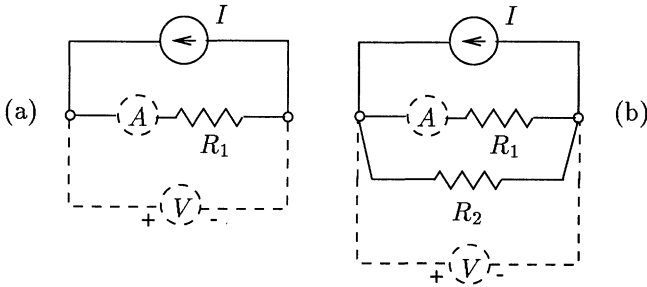


Figure 2.9. Networks used in Prob. 1.

- 2. Figures 2.10(a) and (b) present circuits in which $R_1=1.0 \Omega$ and $R_2=2.0 \Omega$. Let I_1 and I_2 be the control variables. The values of the control variables are to be determined such that: (i) in case (a) the current in R_1 is equal to 2.0 A. (ii) in case (b) the current in R_1 is equal to 2.0 A and that in R_2 to 3.0 A.

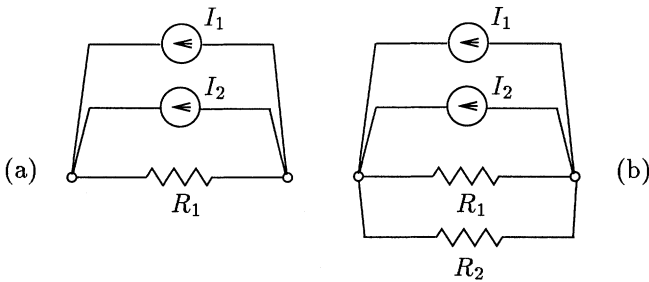


Figure 2.10. Networks used in Prob. 2.

- 3. Figure 2.11 presents a network with resistances of $R_1 = R_2 = R_3 = R_4 = R_5 = 1.0 \Omega$ and $R_6 = 2.0 \Omega$. E and I are unknown. The metered values are $V = 2.9 \text{ V}$, $I_1 = 3.2 \text{ A}$, $I_2 = -3.0 \text{ A}$, $I_3 = -3.7 \text{ A}$, $I_4 = -4.4 \text{ A}$, $I_5 = 4.1 \text{ A}$, $I_6 = -6.9 \text{ A}$. Define and estimate the state of the circuit. Compute the measurement residuals.

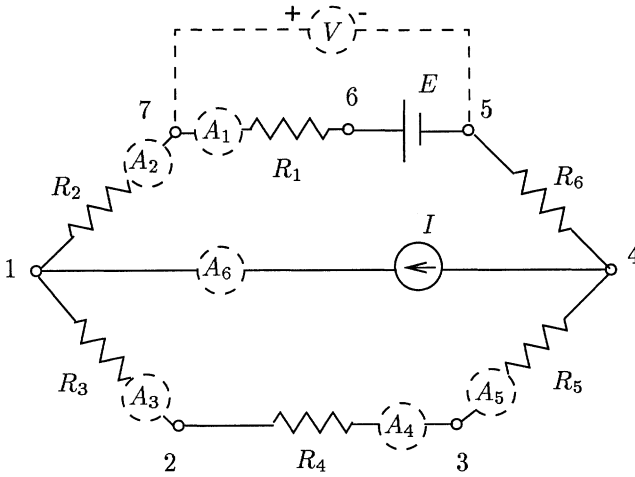


Figure 2.11. Network used in Probs. 3 and 4.

- 4. For the network in Fig. 2.11, assume that E and I are control variables. Determine whether current flow in the resistances R_5 and R_6 can be controlled simultaneously (use arbitrarily specified values for both currents).
- 5. Consider the network given in Fig. 2.2. Repeat Example 2.2, for the two situations given here: (i) E is now constant and equal to zero (only I_1 and I_2 remain as control variables); (ii) I_2 is constant and equal to zero (only I_1 and E remain as control variables).
- 6. For an overdetermined system $\mathbf{Ax} = \mathbf{b}$ with full rank, show that the vector of residuals, $\mathbf{r} = \mathbf{Ax} - \mathbf{b}$, is orthogonal to the range space of \mathbf{A} (as indicated in Fig. 2.6.), i.e., $\mathbf{A}'\mathbf{r} = \mathbf{0}$, and $\mathbf{r}'\mathbf{b} = 0$ if and only if $\mathbf{r}' = \mathbf{0}$, i.e., the vector of measurements is in the range space of \mathbf{A} .
- 7. Consider an overdetermined system $\mathbf{Ax} = \mathbf{b}$ with full rank. Show that the least-squares solution to this system is given by the following tableau equation (Hachtel method):

$$\begin{pmatrix} \mathbf{I} & \mathbf{A} \\ \mathbf{A}' & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{r} \\ \mathbf{x} \end{pmatrix} = \begin{pmatrix} \mathbf{b} \\ \mathbf{0} \end{pmatrix}$$

- 8. Show that the first order optimality conditions for the minimization problem

$$\text{Minimize} \quad \frac{1}{2} \mathbf{r}' \mathbf{r}$$

$$\begin{aligned} \text{subject to } \quad \mathbf{r} &= \mathbf{b} - \mathbf{A} \mathbf{x} \\ \mathbf{C} \mathbf{x} &= \mathbf{0} \end{aligned}$$

can be written as follows (Hachtel tableau):

$$\begin{pmatrix} \mathbf{I} & \mathbf{0} & \mathbf{A} \\ \mathbf{0} & \mathbf{0} & \mathbf{C} \\ \mathbf{A}' & \mathbf{C}' & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{r} \\ \Lambda \\ \mathbf{x} \end{pmatrix} = \begin{pmatrix} \mathbf{b} \\ \mathbf{0} \\ \mathbf{0} \end{pmatrix}$$

where Λ is the Lagrange multiplier associated with the constraint $\mathbf{C} \mathbf{x} = \mathbf{0}$.

- 9. Extend the geometric interpretation in Fig. 2.6(a) to the constrained least-squares problem below:

$$\begin{aligned} \text{Minimize} \quad & \frac{1}{2} \mathbf{r}' \mathbf{r} \\ \text{subject to} \quad & \mathbf{C} \mathbf{x} = \mathbf{0} \end{aligned}$$

References

- Clements, K.A. and Wollenberg, B.F., "An algorithm for observability determination in power system state estimation", paper A75 447-3, IEEE/PES Summer Meeting, San Francisco, CA, July 1975.
- Dahlquist, G. and Björck, Å., *Numerical Methods*, Prentice Hall Series in Automatic Computation, 1974.
- Krumpholz, G.R., Clements, K.A., and Davis, P.W., "Power system observability: A practical algorithm using network topology", IEEE Trans. Power App. Syst., Vol. 99, pp. 1534-1542, July/Aug. 1980.
- Duff, I.F. and Erisman, A.M., *Direct Methods for Sparse Matrices*, Oxford Science Publications, Oxford, 1992.
- Golub, G.H., Van Loan, C., *Matrix Computations*, 2nd Edition, John Hopkins University Press, 1989.
- Hachtel, G.D., "The sparse tableau approach to finite element assembly", *Sparse Matrix Computations*, pp. 349-363, 1976.
- Monticelli, A., Wu, F.F., "Network observability: Identification of observable islands and measurement placement", IEEE Transactions Power App. Syst., Vol. 104, No. 5, pp. 1035-1041, May 1985.
- Monticelli, A. and Wu, F.F., "Network observability: Theory", IEEE Trans. Power App. Systems, Vol. 104, No. 5, pp. 1042-1048, May 1985.
- Monticelli, A. and Liu, W.E., "Adaptive movement penalty method for the Newton optimal power flow", IEEE Trans. Power Syst., Vol. 7, No. 1, pp. 334-343, Feb. 1992.
- Schweppe, F.C., Wildes, J., and Rom, D., "Power system static state estimation: Parts I, II, and III", Power Industry Computer Conference, PICA, Denver, Colorado, June 1969.

- Stott, B. and Robson, E., "Power system security control calculations using linear programming, Parts I and II", IEEE Trans. Power App. Syst., Vol. 3, pp. 676-683, May 1988.
- Tinney, W.F., Bright, J.M., Demaree, K.D., and Hugues, B.A., "Some deficiencies in optimal power flow", IEEE Trans. Power Syst., Vol. 3, pp. 676-683, May 1988.
- Van Amerongen, R.A.M., "On convergence analysis and convergence enhancement of power-system least-squares state estimators", IEEE Trans. Power Syst., Vol. 10, No. 4, pp. 2038-2044, Nov. 1995.
- Watkins, D.S., *Fundamentals of Matrix Computations*, John Wiley and Sons, New York, 1991.

3 DC STATE ESTIMATOR

This chapter discusses the main characteristics of generalized state estimation problems using a dc power flow model. First the concept of state variable is introduced. Then measurement models are discussed. Finally state, status and parameter estimation are introduced.

3.1 OVERVIEW OF THE DC STATE ESTIMATOR

The dc state estimation problem is commonly formulated as an overdetermined system of linear equations (the measurement model), and solved as weighted least-squares (WLS) problem. The state estimation measurement model relates measurements to state variables.

$$\mathbf{z} = \mathbf{H}\mathbf{x} + \mathbf{e} \quad (3.1)$$

where

- \mathbf{x} is the n vector of the true states (unknown)
- \mathbf{z} is the m vector of measurements (known)
- \mathbf{H} is the $m \times n$ Jacobian matrix
- $\mathbf{H}\mathbf{x}$ is the m vector of linear functions linking measurements to states)

- \mathbf{e} is the m vector of random errors
- m is the number of measurements
- n is the number of state variables

The measurement residual vector is defined as

$$\mathbf{r} = \mathbf{z} - \mathbf{H}\mathbf{x} \quad (3.2)$$

An estimate of \mathbf{r} is the differences between the measured values, \mathbf{z} , and the corresponding estimated values $\hat{\mathbf{z}} = \mathbf{H}\hat{\mathbf{x}}$.

The weighted least-squares problem for the overdetermined system in Eq. (3.1) involves finding the n -vector \mathbf{x} that minimizes the index $J(\mathbf{x})$, defined as follows

$$J(\mathbf{x}) = (\mathbf{z} - \mathbf{H}\mathbf{x})' \mathbf{W} (\mathbf{z} - \mathbf{H}\mathbf{x}) \quad (3.3)$$

Matrix \mathbf{W} is a diagonal matrix whose elements are the measurement weights. Depending on the application, these weights may represent such entities as meter accuracy, reliability, or an engineering judgement expressing the relative importance one wishes to allocate to each individual measurement. Most commonly, \mathbf{W} is based on the reciprocals of the variance of measurement error

$$\mathbf{W} = \mathbf{R}_z^{-1} = \begin{pmatrix} \sigma_1^{-2} & & & \\ & \sigma_2^{-2} & & \\ & & \ddots & \\ & & & \sigma_m^{-2} \end{pmatrix} \quad (3.4)$$

where \mathbf{R}_z is the measurement covariance matrix.

The performance index $J(\mathbf{x})$ can be differentiated to obtain the first-order optimal conditions

$$\mathbf{G}\hat{\mathbf{x}} = \mathbf{H}' \mathbf{W} \mathbf{z} \quad (3.5)$$

where $\hat{\mathbf{x}}$ is the state estimate and $\mathbf{G} = \mathbf{H}' \mathbf{W} \mathbf{H}$ is the state estimation gain matrix.

Remarks: Solving Eq. (3.5) yields the state estimate $\hat{\mathbf{x}}$. The formula that furnishes the state estimate, $\hat{\mathbf{x}} = \mathbf{G}^{-1} \mathbf{H}' \mathbf{W} \mathbf{z}$, is the state estimator. The process of obtaining the estimate $\hat{\mathbf{x}}$ and related quantities is called state estimation.

3.2 STATE VARIABLES

Consider a network described by n_v problem variables. When the dc power flow model is used, these variables are bus voltage angles, branch active power flows, and bus active power loads and generations. These variables are connected by a set of n_e network equations (network model) that express the power balance at network buses and the relationship between nodal voltage angles and branch active power flows. Since the dc power flow is linear, the network model is a linear set of independent equations. Normally it is an undetermined set of equations: the difference $n_v - n_e$, between the number of problem variables and the order of the model (number of independent equations), is the number of degrees of freedom of this network model. (This is not to be confused with the number of degrees of freedom of the measurement model discussed in the next section.)

The set of problem variables can be divided into two subsets: a set of $(n_v - n_e)$ state variables and a set of dependent variables. The set of state variables has two principal characteristics: (a) it describes the system completely, in the sense that, if the states are known, all the remaining variables (the dependent set) can be determined using the network model equations; (b) the set of states is minimum, in the sense that if any of the state variables is removed from the set, property (a) does not hold.

These properties are illustrated by the following three examples.

Example 3.1:

Figure 3.1 shows a network with three buses and two short circuit branches (zero impedance branches). The choice of a set of state variables for this system is discussed below.

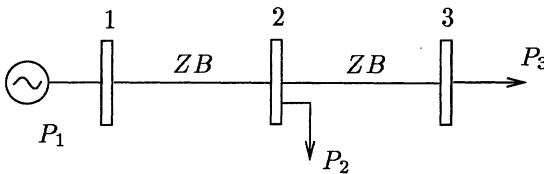


Figure 3.1. Three-bus system with two zero-impedance branches (ZB's).

- *Problem variables:* This dc power flow model has the following problem variables: $P_1, P_2, P_3, P_{12}, P_{23}, \theta_1, \theta_2$, and θ_3 . An angle reference is required, for example, $\theta_1 = 0^\circ$. Since the three buses are connected by zero-impedance branches, $\theta_2 = \theta_3 = 0^\circ$ is also true. Thus the bus voltage angles can be ignored. There are five variables remaining: P_1, P_2, P_3, P_{12} , and P_{23} .

- *Network model:* The network model establishes the active power balance at Buses 1, 2 and 3

$$\begin{aligned}P_1 - P_{12} &= 0 \\P_2 + P_{12} - P_{23} &= 0 \\P_3 + P_{23} &= 0\end{aligned}$$

Thus there is a set of three independent equations for the five problem variables.

- *State variables:* Power flows P_{12} and P_{23} can be selected as state variables since: (a) once their values are known, the other problem variables can be determined from the network model; and (b) they form a minimum set. The pair P_1 and P_{23} , as well as various other pairs, can also be considered state variables. Not all pairs of problem variables, however, can be selected as state variables: for example, P_{12} and P_1 are not states since, even when both values are known, the other variables will remain undetermined (because they cannot be determined from the network model equations).

Example 3.2:

Consider the two-bus system shown in Fig. 3.2.

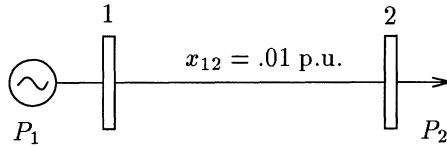


Figure 3.2. Two-bus system.

- *Problem variables:* P_1 , P_2 , P_{12} , and θ_2 ($\theta_1 = 0^\circ$ is the angle reference).
- *Network model:* The network model furnishes the relationship between bus voltage angles and the active power flow in the line, and establishes the active power balance at Buses 1 and 2:

$$\begin{aligned}-100 \theta_2 - P_{12} &= 0 \\P_1 - P_{12} &= 0 \\P_2 + P_{12} &= 0\end{aligned}$$

This gives a set of three independent equations relating the four problem variables.

- *State variables:* In this example any of the variables can be selected as a state variable, but in practice it is common to take θ_2 as state variable. Of course, all the remaining variables can then be determined in terms of θ_2 using the set network model equations.

Example 3.3:

Consider the three-bus system shown in Fig. 3.3.

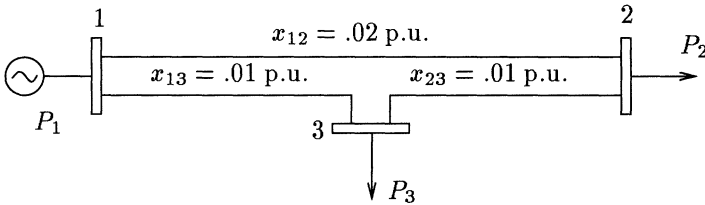


Figure 3.3. Three-bus system.

- *Problem variables:* $P_1, P_2, P_3, P_{12}, P_{13}, P_{23}, \theta_2$ and θ_3 ($\theta_1 = 0^\circ$).
- *Network model:*

$$\begin{aligned} P_1 - P_{12} - P_{13} &= 0 \\ P_2 + P_{12} - P_{23} &= 0 \\ P_3 + P_{13} + P_{23} &= 0 \\ P_{12} + 50 \theta_2 &= 0 \\ P_{13} + 100 \theta_3 &= 0 \\ P_{23} - 100 (\theta_2 - \theta_3) &= 0 \end{aligned}$$

- *State variables:* Since there are eight problem variables and six independent network equations, two variables can be selected as state variables – for example, θ_2 and θ_3 is a common choice.

3.3 MEASUREMENT MODEL

The network model discussed above relates the state variables to the dependent variables by means of the set of network equations. State estimation, however, is based on a set of measurements (both telemetered data and pseudo-measurements). The measurement model describes the measurement system by establishing the relationships between the measured variables and the state variables.

Measurements are divided into two categories: measurements of state variables and the measurements of dependent variables. Both types of measurements are represented in the measurement model: the former is trivial; whereas for the latter network model equations are used to write the measured variables in terms of the state variables. The number of degrees of freedom of the measurement model is the difference between the number of states and the rank (number of independent rows/columns) of the model.

The following examples illustrate how measurement models are built from the set of measurements using network model equations.

Example 3.4:

Consider the one-bus system shown in Fig. 3.4 (a) with the metering systems shown in parts (b), (c), and (d) of the figure.

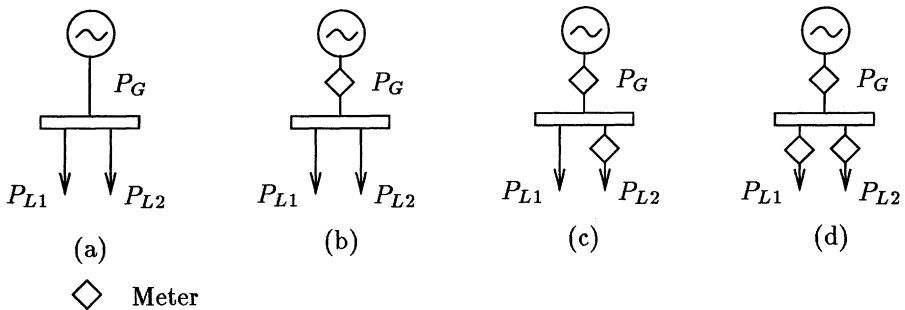


Figure 3.4. One-bus system with three different measurement sets.

- *Problem variables:* P_G , P_{L1} , and P_{L2}
- *Network model:* In this case the network model simply establishes the active power balance at the bus, i.e. $P_G + P_{L1} + P_{L2} = 0$.
- *State variables:* any two variables, P_G and P_{L1} , can be selected as state variables; once these states are known, the dependent variable P_{L2} can be computed from the network model equation.

- Measurement model:** The equations corresponding to the measurement model furnishes the metered variables in terms of the state variables. When the metered variable is a state variable, as for example P_G^{meas} , the result is $P_G = P_G^{meas}$. When the metered value is a dependent variable, it is expressed in terms of the state variables using the network model equations, as for example in the case of measurements P_{L2}^{meas} which yields $-P_{L1} - P_G = P_{L2}^{meas}$. The resulting measurement model for the example described in Fig. 3.4 (d) is

$$\begin{aligned} P_G &= P_G^{meas} \\ P_{L1} &= P_{L1}^{meas} \\ -P_{L1} - P_G &= P_{L2}^{meas} \end{aligned}$$

These equations can be rewritten in matrix form as follows

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ -1 & -1 \end{pmatrix} \begin{pmatrix} P_G \\ P_{L1} \end{pmatrix} = \begin{pmatrix} P_G^{meas} \\ P_{L1}^{meas} \\ P_{L2}^{meas} \end{pmatrix} \quad \text{Example (d)}$$

where the Jacobian matrix, \mathbf{H} , the vector of state variables, \mathbf{x} , and the measurement vector, \mathbf{z} , are, respectively,

$$\mathbf{H} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ -1 & -1 \end{pmatrix} \quad \mathbf{x} = \begin{pmatrix} P_G \\ P_{L1} \end{pmatrix} \quad \mathbf{z} = \begin{pmatrix} P_G^{meas} \\ P_{L1}^{meas} \\ P_{L2}^{meas} \end{pmatrix}$$

Note that in this case (Fig. 3.4 (d)) the measurement model is an overdetermined system of linear equations that can be solved as a least-squares problem. For cases (b) and (c) of Fig. 3.4 we have the following models, respectively,

$$\begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} P_G \\ P_{L1} \end{pmatrix} = \begin{pmatrix} P_G^{meas} \end{pmatrix} \quad \text{Example (b)}$$

$$\begin{pmatrix} 1 & 0 \\ -1 & -1 \end{pmatrix} \begin{pmatrix} P_G \\ P_{L1} \end{pmatrix} = \begin{pmatrix} P_G^{meas} \\ P_{L2}^{meas} \end{pmatrix} \quad \text{Example (c)}$$

- Remarks:** In the examples 3.4 (c) and 3.4 (d) both states P_G and P_{L1} are observable, i.e., they can be determined from the equations of the measurement model. In the example (b), however, only state P_G is observable; state P_{L1} is unobservable and the dependent variable P_{L2} is indeterminate.

Example 3.5:

Consider the two-bus system shown in Fig. 3.5 (a) with the three alternative metering systems illustrated in (b), (c), and (d) of the same figure.

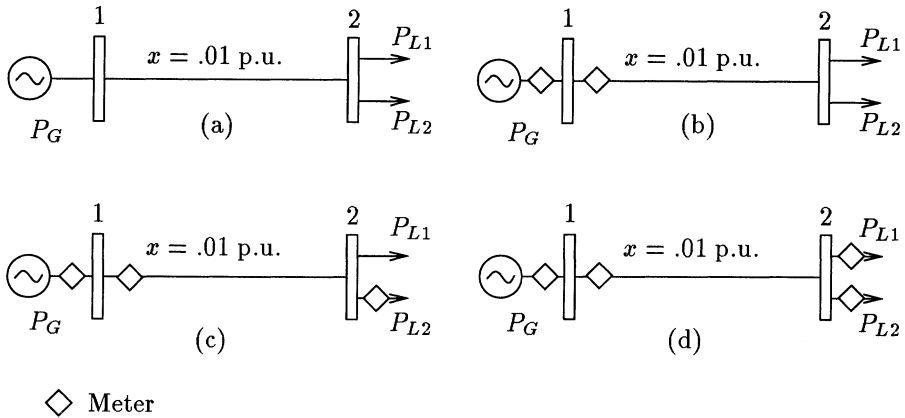


Figure 3.5. Two-bus system with three different metering systems.

- *Problem variables:* P_G , P_{L1} , P_{L2} , P_{12} , and θ_2 ($\theta_1 = 0^\circ$).
- *Network model:* In this case there are three network equations. Two equations determine the active power balance at the two buses; and the other equation relates the active power flow in the line to the bus voltage angles:

$$\begin{aligned} P_G - P_{12} &= 0 \\ P_{L1} + P_{L2} + P_{12} &= 0 \\ P_{12} + 100 \theta_2 &= 0 \end{aligned}$$

- *State variables:* Two of the problem variables can be chosen as state variables. If, for example, θ_2 and P_{L1} are selected as state variables and calculated, the dependent variables P_{12} , P_G , and P_{L2} can be computed from network model equations. Of course there are other possible sets of states, such as, P_{12} and P_{L1} , or P_{L1} and P_{L2} .
- *Measurement model:* As in the previous example, the equations of the measurement model express metered variables in terms of the state variables. For example, considering θ_2 and P_{L1} to be state variables, results in the following equations for the measurements of Fig. 3.4 (d):

$$\begin{aligned} P_{L1} &= P_{L1}^{meas} \\ 100 \theta_2 - P_{L1} &= P_{L2}^{meas} \\ -100 \theta_2 &= P_G^{meas} \\ -100 \theta_2 &= P_{12}^{meas} \end{aligned}$$

These equations can be rewritten in matrix form as follows

$$\begin{pmatrix} 0 & 1 \\ 100 & -1 \\ -100 & 0 \\ -100 & 0 \end{pmatrix} \begin{pmatrix} \theta_2 \\ P_{L1} \end{pmatrix} = \begin{pmatrix} P_{L1}^{meas} \\ P_{L2}^{meas} \\ P_G^{meas} \\ P_{12}^{meas} \end{pmatrix} \quad \text{Example (d)}$$

where the Jacobian matrix, \mathbf{H} , the state variable vector, \mathbf{x} , and the measurement vector \mathbf{z} are, respectively,

$$\mathbf{H} = \begin{pmatrix} 0 & 1 \\ 100 & -1 \\ -100 & 0 \\ -100 & 0 \end{pmatrix} \quad \mathbf{x} = \begin{pmatrix} \theta_2 \\ P_{L1} \end{pmatrix} \quad \mathbf{z} = \begin{pmatrix} P_{L1}^{meas} \\ P_{L2}^{meas} \\ P_G^{meas} \\ P_{12}^{meas} \end{pmatrix}$$

Thus the measurement model for this example (Fig. 3.5 (d)) is an overdetermined system of linear equations that can be solved as a least-squares problem. For the examples (b) and (c) of Fig. 3.5 we have the following models, respectively,

$$\begin{pmatrix} -100 & 0 \\ -100 & 0 \end{pmatrix} \begin{pmatrix} \theta_2 \\ P_{L1} \end{pmatrix} = \begin{pmatrix} P_G^{meas} \\ P_{12}^{meas} \end{pmatrix} \quad \text{Example (b)}$$

$$\begin{pmatrix} 100 & -1 \\ -100 & 0 \\ -100 & 0 \end{pmatrix} \begin{pmatrix} \theta_2 \\ P_{L1} \end{pmatrix} = \begin{pmatrix} P_{L2}^{meas} \\ P_G^{meas} \\ P_{12}^{meas} \end{pmatrix} \quad \text{Example (c)}$$

- *Remarks:* In the examples (c) and (d) both states θ_2 and P_{L2} are observable, i.e., they can be determined from the equations of the measurement model. In the example (b) the Jacobian matrix is singular (i.e. $\text{rank}(\mathbf{H})=1$), which means that states θ_2 and P_{L1} cannot be determined from the measurement model.
- *Alternative approach:* In the example (b) above an arbitrary value can be used for the state variable P_{L1} and θ_2 can then be estimated from the measurement model. In this case the measurement model can then be rewritten as

$$\begin{pmatrix} -100 \\ -100 \end{pmatrix} (\theta_2) = \begin{pmatrix} P_G^{meas} \\ P_{12}^{meas} \end{pmatrix} \quad \text{case (b)}$$

Thus, the problem is not entirely solvable, since we cannot determine P_{L1} and P_{L2} , although it is still possible to estimate the state θ_2 .

Example 3.6:

Consider the three-bus system shown in Fig. 3.6. The problem variables, the network model equations, and the state variables are the same as in Example 3.3.

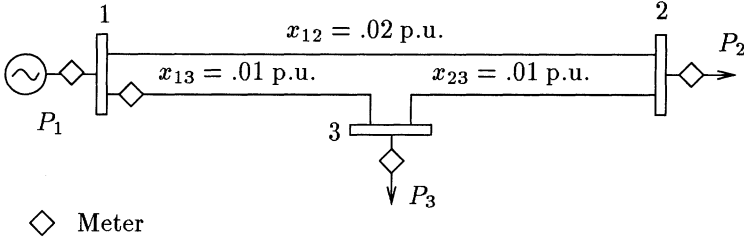


Figure 3.6. Three-bus system.

- *Measurement model:* In this case there are four measurements (P_1^{meas} , P_2^{meas} , P_3^{meas} , and P_{13}^{meas}) which are expressed in terms of the state variables (θ_2 and θ_3):

$$\begin{aligned} -50 \theta_2 - 100 \theta_3 &= P_1^{meas} \\ 150 \theta_2 - 100 \theta_3 &= P_2^{meas} \\ -100 \theta_2 + 200 \theta_3 &= P_3^{meas} \\ -100 \theta_3 &= P_{13}^{meas} \end{aligned}$$

These equations can be rewritten in matrix form as follows

$$\begin{pmatrix} -50 & -100 \\ 150 & -100 \\ -100 & 200 \\ 0 & -100 \end{pmatrix} \begin{pmatrix} \theta_2 \\ \theta_3 \end{pmatrix} = \begin{pmatrix} P_1^{meas} \\ P_2^{meas} \\ P_3^{meas} \\ P_{13}^{meas} \end{pmatrix}$$

where the Jacobian matrix, \mathbf{H} , the state variable vector, \mathbf{x} , and the measurement vector \mathbf{z} are as follows

$$\mathbf{H} = \begin{pmatrix} -50 & -100 \\ 150 & -100 \\ -100 & 200 \\ 0 & -100 \end{pmatrix} \quad \mathbf{x} = \begin{pmatrix} \theta_2 \\ \theta_3 \end{pmatrix} \quad \mathbf{z} = \begin{pmatrix} P_1^{meas} \\ P_2^{meas} \\ P_3^{meas} \\ P_{13}^{meas} \end{pmatrix}$$

3.4 SOLVING THE NORMAL EQUATION

The solution of normal equation $\mathbf{G} \hat{\mathbf{x}} = \mathbf{H}' \mathbf{W} \mathbf{z}$, where the gain matrix \mathbf{G} is $\mathbf{G} = \mathbf{H}' \mathbf{W} \mathbf{H}$, yields the state estimate $\hat{\mathbf{x}}$. In this section some of the examples discussed above will be used to illustrate how $\hat{\mathbf{x}}$ is obtained by solving the normal equation. Once $\hat{\mathbf{x}}$ is known, the estimated values of the measured quantities, $\hat{\mathbf{z}}$, and the estimation residuals, $\hat{\mathbf{r}}$, can be calculated, respectively, from,

$$\hat{\mathbf{z}} = \mathbf{H} \hat{\mathbf{x}}$$

$$\hat{\mathbf{r}} = \mathbf{z} - \mathbf{H} \hat{\mathbf{x}}$$

Example 3.7:

Consider the one-bus example of Fig. 3.4 (d) with the following measured values: $P_G^{meas} = 1.05$ p.u., $P_{L_1}^{meas} = -0.72$ p.u., and $P_{L_2}^{meas} = -0.29$ p.u., as well as variances $\sigma_G^2 = .004$, $\sigma_{L_1}^2 = .001$, $\sigma_{L_2}^2 = .001$.

The corresponding measurement model, obtained in Example 3.4, is

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ -1 & -1 \end{pmatrix} \begin{pmatrix} P_G \\ P_{L_1} \end{pmatrix} = \begin{pmatrix} P_G^{meas} \\ P_{L_1}^{meas} \\ P_{L_2}^{meas} \end{pmatrix} = \begin{pmatrix} 1.05 \\ -0.72 \\ -0.29 \end{pmatrix}$$

The gain matrix ($\mathbf{G} = \mathbf{H}' \mathbf{W} \mathbf{H}$) is

$$\mathbf{G} = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} 250 & 0 & 0 \\ 0 & 1000 & 0 \\ 0 & 0 & 1000 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ -1 & -1 \end{pmatrix} = \begin{pmatrix} 1250 & 1000 \\ 1000 & 2000 \end{pmatrix}$$

Thus the state estimate ($\hat{\mathbf{x}} = \mathbf{G}^{-1} \mathbf{H}' \mathbf{W} \mathbf{z}$) is

$$\hat{\mathbf{x}} = 10^{-4} \begin{pmatrix} 13.3 & -6.67 \\ -6.67 & 8.33 \end{pmatrix} \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} 250 & 0 & 0 \\ 0 & 1000 & 0 \\ 0 & 0 & 1000 \end{pmatrix} \begin{pmatrix} 1.05 \\ -0.72 \\ -0.29 \end{pmatrix}$$

$$\hat{\mathbf{x}} = \begin{pmatrix} \hat{P}_G \\ \hat{P}_{L_1} \end{pmatrix} = \begin{pmatrix} 1.0233 \\ -0.7267 \end{pmatrix}$$

The estimated values of the measured quantities ($\hat{\mathbf{z}} = \mathbf{H} \hat{\mathbf{x}}$) and the estimation residuals ($\hat{\mathbf{r}} = \mathbf{z} - \mathbf{H} \hat{\mathbf{x}}$) are as follows

$$\hat{\mathbf{z}} = \begin{pmatrix} \hat{P}_G \\ \hat{P}_{L_1} \\ \hat{P}_{L_2} \end{pmatrix} = \begin{pmatrix} 1.0233 \\ -0.7267 \\ -0.2967 \end{pmatrix}$$

$$\hat{\mathbf{r}} = \begin{pmatrix} P_G^{meas} - \hat{P}_G \\ P_{L_1}^{meas} - \hat{P}_{L_1} \\ P_{L_2}^{meas} - \hat{P}_{L_2} \end{pmatrix} = \begin{pmatrix} 0.0267 \\ 0.0067 \\ 0.0067 \end{pmatrix}$$

Example 3.8:

Consider the three-bus example of Fig. 3.6 with the following measured values: $P_1^{meas} = 3.90$ p.u., $P_2^{meas} = -4.07$ p.u., $P_3^{meas} = -.04$ p.u., and $P_{13}^{meas} = 2.04$ p.u., as well as variances $\sigma_1^2 = .004$ p.u., $\sigma_2^2 = .004$ p.u., $\sigma_3^2 = .001$, and $\sigma_{13}^2 = .002$.

The corresponding measurement model, obtained in Example 3.6, is rewritten here

$$\begin{pmatrix} -50 & -100 \\ 150 & -100 \\ -100 & 200 \\ 0 & -100 \end{pmatrix} \begin{pmatrix} \theta_2 \\ \theta_3 \end{pmatrix} = \begin{pmatrix} P_1^{meas} \\ P_2^{meas} \\ P_3^{meas} \\ P_{13}^{meas} \end{pmatrix} = \begin{pmatrix} 3.90 \\ -4.07 \\ -.04 \\ 2.04 \end{pmatrix}$$

The gain matrix ($\mathbf{G} = \mathbf{H}' \mathbf{W} \mathbf{H}$) is

$$\mathbf{G} = \begin{pmatrix} -50 & 150 & -100 & 0 \\ -100 & -100 & 200 & -100 \end{pmatrix} \begin{pmatrix} 250 & 0 & 0 & 0 \\ 0 & 250 & 0 & 0 \\ 0 & 0 & 1000 & 0 \\ 0 & 0 & 0 & 500 \end{pmatrix} \begin{pmatrix} -50 & -100 \\ 150 & -100 \\ -100 & 200 \\ 0 & -100 \end{pmatrix}$$

$$\mathbf{G} = \begin{pmatrix} 16250000 & -22500000 \\ -22500000 & 50000000 \end{pmatrix}$$

Thus the state estimate, $\hat{\mathbf{x}}$, is as follows

$$\hat{\mathbf{x}} = \mathbf{G}^{-1} \mathbf{H}' \mathbf{W} \mathbf{z} = \begin{pmatrix} \hat{\theta}_2 \\ \hat{\theta}_3 \end{pmatrix} = \begin{pmatrix} -.0400 \\ -0.0201 \end{pmatrix}$$

The estimated values of the measured quantities ($\hat{\mathbf{z}} = \mathbf{H} \hat{\mathbf{x}}$) and the estimation residuals ($\hat{\mathbf{r}} = \mathbf{z} - \mathbf{H} \hat{\mathbf{x}}$) are as follows

$$\hat{\mathbf{z}} = \begin{pmatrix} \hat{P}_1 \\ \hat{P}_2 \\ \hat{P}_3 \\ \hat{P}_{13} \end{pmatrix} = \begin{pmatrix} 4.01 \\ -3.99 \\ -0.023 \\ 2.01 \end{pmatrix}$$

$$\hat{\mathbf{r}} = \begin{pmatrix} P_1^{meas} - \hat{P}_1 \\ P_2^{meas} - \hat{P}_2 \\ P_3^{meas} - \hat{P}_3 \\ P_{13}^{meas} - \hat{P}_{13} \end{pmatrix} = \begin{pmatrix} -0.111 \\ -0.082 \\ -0.017 \\ 0.029 \end{pmatrix}$$

Example 3.9:

Consider the three-bus example of Fig. 3.7 with the following measured values: $P_1^{meas} = 5.10$ p.u., $P_2^{meas} = -4.87$ p.u., $P_3^{meas} = -0.05$ p.u., and $P_{13}^{meas} = 2.03$ p.u., as well as variances $\sigma_1^2 = .004$ p.u., $\sigma_2^2 = .004$ p.u., $\sigma_3^2 = .001$, and $\sigma_{13}^2 = .002$. Consider also a phase-shift of $\varphi = -0.02$ radians.

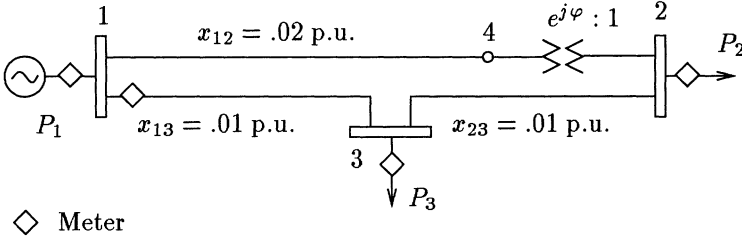


Figure 3.7. Four-bus system with ideal phase-shifter.

In this case the measurement model is

$$\begin{aligned} 50 (\theta_1 - (\theta_2 + \varphi)) + 100 (\theta_1 - \theta_3) &= P_1^{meas} \\ 50 ((\theta_2 + \varphi) - \theta_1) + 100 (\theta_2 - \theta_3) &= P_2^{meas} \\ 100 (\theta_3 - \theta_1) + 100 (\theta_3 - \theta_2) &= P_3^{meas} \\ 100 (\theta_1 - \theta_3) &= P_{13}^{meas} \end{aligned}$$

Considering that $\theta_1 = 0^\circ$ and $\varphi = -0.02$ radians, the measurement model can be rewritten as,

$$\begin{pmatrix} -50 & -100 \\ 150 & -100 \\ -100 & 200 \\ 0 & -100 \end{pmatrix} \begin{pmatrix} \theta_2 \\ \theta_3 \end{pmatrix} = \begin{pmatrix} P_1^{meas} - 1.0 \\ P_2^{meas} + 1.0 \\ P_3^{meas} \\ P_{13}^{meas} \end{pmatrix} = \begin{pmatrix} 4.10 \\ -3.87 \\ -0.05 \\ 2.03 \end{pmatrix}$$

The state estimate, $\hat{\mathbf{x}}$, is then

$$\hat{\mathbf{x}} = \mathbf{G}^{-1} \mathbf{H}' \mathbf{W} \mathbf{z} = \begin{pmatrix} \hat{\theta}_2 \\ \hat{\theta}_3 \end{pmatrix} = \begin{pmatrix} -0.0399 \\ -0.0203 \end{pmatrix}$$

The estimated values of the measured quantities ($\hat{\mathbf{z}} = \mathbf{H} \hat{\mathbf{x}}$) are as follows

$$\hat{\mathbf{z}} = \begin{pmatrix} \hat{P}_1 \\ \hat{P}_2 \\ \hat{P}_3 \\ \hat{P}_{13} \end{pmatrix} = \begin{pmatrix} 5.02 \\ -4.95 \\ -0.070 \\ 2.03 \end{pmatrix}$$

3.5 PHASE-SHIFT ESTIMATION

The previous example considered a phase-shifter whose phase-shift is known. There are situations, however, in which the phase-shift varies according to the operating conditions of the network and it may be necessary to obtain a reliable estimate of the current value. There are several different approaches to this problem. The most commonly adopted practice consists of including the phase-shift as an additional state variable in the state estimation problem. It should be noted that when no direct measurement of the phase-shift is available, it may still be possible to estimate its current value from other measurements. And in those cases in which the direct measurement of the phase-shift is possible, the state estimator will yield a more accurate estimate. Besides, assuming there is measurement redundancy, state estimation is useful for cross-checking the measured shift and the other measurements that form part of the measurement model.

Example 3.10:

Consider the case in which the phase shift angle φ of the network shown in Fig. 3.7 is unknown. The formulation of the measurement model follows the same procedure as in Example 3.9, except that φ appears now as an extra state, whereas above it was treated as a given.

In Example 3.9 the measurement model was originally written as

$$\begin{aligned} 50 (\theta_1 - (\theta_2 + \varphi)) + 100 (\theta_1 - \theta_3) &= P_1^{meas} \\ 50 ((\theta_2 + \varphi) - \theta_1) + 100 (\theta_2 - \theta_3) &= P_2^{meas} \\ 100 (\theta_3 - \theta_1) + 100 (\theta_3 - \theta_2) &= P_3^{meas} \\ 100 (\theta_1 - \theta_3) &= P_{13}^{meas} \end{aligned}$$

Now, considering φ as a state, this model can be rewritten as

$$\begin{pmatrix} -50 & -100 & -50 \\ 150 & -100 & 50 \\ -100 & 200 & 0 \\ 0 & -100 & 0 \end{pmatrix} \begin{pmatrix} \theta_2 \\ \theta_3 \\ \varphi \end{pmatrix} = \begin{pmatrix} P_1^{meas} \\ P_2^{meas} \\ P_3^{meas} \\ P_{13}^{meas} \end{pmatrix} = \begin{pmatrix} 5.10 \\ -4.87 \\ -0.05 \\ 2.03 \end{pmatrix}$$

The state estimate, $\hat{\mathbf{x}}$, and the estimated values of the measured quantities ($\hat{\mathbf{z}} = \mathbf{H} \hat{\mathbf{x}}$) are as follows

$$\begin{aligned} \hat{\mathbf{x}} &= \mathbf{G}^{-1} \mathbf{H}' \mathbf{W} \mathbf{z} = \begin{pmatrix} \hat{\theta}_2 \\ \hat{\theta}_3 \\ \hat{\varphi} \end{pmatrix} = \begin{pmatrix} -0.0399 \\ -0.0203 \\ -0.0199 \end{pmatrix} \\ \hat{\mathbf{z}} &= \begin{pmatrix} \hat{P}_1 \\ \hat{P}_2 \\ \hat{P}_3 \\ \hat{P}_{13} \end{pmatrix} = \begin{pmatrix} 5.02 \\ -4.95 \\ -0.070 \\ 2.03 \end{pmatrix} \end{aligned}$$

- *Remarks:* Note that the actual value for φ used in Example 3.9 was $-.020$. In the present case, however, it was considered to be an unknown parameter which was included in the state estimation problem as an additional state variable. Although φ is not metered directly it was possible to estimate its value. If a direct measure of φ were available it could be included in the measurement model as follows

$$\begin{pmatrix} -50 & -100 & -50 \\ 150 & -100 & 50 \\ -100 & 200 & 0 \\ 0 & -100 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \theta_2 \\ \theta_3 \\ \varphi \end{pmatrix} = \begin{pmatrix} P_1^{meas} \\ P_2^{meas} \\ P_3^{meas} \\ P_{13}^{meas} \\ \varphi^{meas} \end{pmatrix}$$

3.6 PARAMETER ESTIMATION

The procedure adopted to estimate phase shift angles discussed in the previous section can be extended to the estimation of other types of parameters, such as branch series reactances. This, however, makes the measurement model into a nonlinear model. For example, consider the active power flow measurement P_{13} in the system of Fig. 3.6. The contribution of this particular measurement to the measurement model is the equation $P_{12} + x_{13}^{-1} \theta_3 = 0$; if both θ_3 and x_{13} are state variables, the model becomes nonlinear. Not only would this increase the computational complexity, but might lead to singularities (or severe ill-conditioning) during state estimation.

A simpler alternative is thus presented which eliminates the unknown (or suspect) parameter to be estimated from the measurement model. By doing this, even those examples in which the initial parameter value is far from the correct value will not cause convergence or ill-conditioning problems in relation to the state estimator (see Chap. 11).

Example 3.11:

Consider the network of Fig. 3.8 in which Branch 1 – 3 has an unknown reactance. The measurement model formulation follows the same procedure as in Example 3.6, except that now parameter x_{13} cannot appear in the model since it is unknown.

- *Problem variables:* $P_1, P_2, P_3, P_{12}, P_{13}, P_{23}, \theta_2$ and θ_3 ($\theta_1 = 0^\circ$).
- *Network model:*

$$P_1 - P_{12} - P_{13} = 0$$

$$P_2 + P_{12} - P_{23} = 0$$

$$P_3 + P_{13} + P_{23} = 0$$

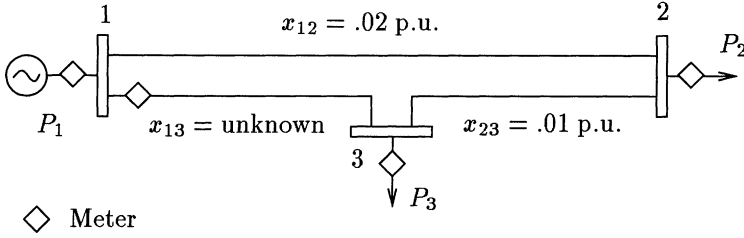


Figure 3.8. Three-bus system with an unknown reactance for one branch.

$$\begin{aligned} P_{12} - x_{12}^{-1} (-\theta_2) &= 0 \\ P_{23} - x_{23}^{-1} (\theta_2 - \theta_3) &= 0 \end{aligned}$$

Note that equation $P_{13} - x_{13}^{-1} (-\theta_3) = 0$ which was include in the measurement model derived in Example 3.6, does not form part of the current model. As a result there is one extra degree of freedom in the model.

- *State variables:* Since there are eight problem variables and five independent network equations, it is possible to select three variables as state variables. Thus, θ_2 , θ_3 , and P_{13} could form a set of possible state variables (This is a minimum set, and completely determines all the remaining problem variables.)
- *Measurement model:* In this example four measurements (P_1^{meas} , P_2^{meas} , P_3^{meas} , and P_{13}^{meas}) are expressed in terms of the state variables (θ_2 , θ_3 , and P_{13}) according the following model:

$$\begin{pmatrix} -x_{12}^{-1} & 0 & 1 \\ x_{12}^{-1} + x_{23}^{-1} & -x_{23}^{-1} & 0 \\ -x_{23}^{-1} & x_{23}^{-1} & -1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \theta_2 \\ \theta_3 \\ P_{13} \end{pmatrix} = \begin{pmatrix} P_1^{meas} \\ P_2^{meas} \\ P_3^{meas} \\ P_{13}^{meas} \end{pmatrix}$$

Example 3.12:

Consider again the three-bus example in Fig. 3.8 with the same set of measurements used in Example 3.8, i.e., $P_1^{meas} = 3.90$ p.u., $P_2^{meas} = -4.07$ p.u., $P_3^{meas} = -.04$ p.u., and $P_{13}^{meas} = 2.04$ p.u., as well as variances $\sigma_1^2 = .004$ p.u., $\sigma_2^2 = .004$ p.u., $\sigma_3^2 = .001$, and $\sigma_{13}^2 = .002$.

The corresponding measurement model can be written as follows

$$\begin{pmatrix} -50 & 0 & 1 \\ 150 & -100 & 0 \\ -100 & 100 & -1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \theta_2 \\ \theta_3 \\ P_{13} \end{pmatrix} = \begin{pmatrix} 3.90 \\ -4.07 \\ -0.04 \\ 2.04 \end{pmatrix}$$

The state estimate, $\hat{\mathbf{x}}$, is given by

$$\hat{\mathbf{x}} = \mathbf{G}^{-1} \mathbf{H}' \mathbf{W} \mathbf{z} = \begin{pmatrix} \hat{\theta}_2 \\ \hat{\theta}_3 \\ \hat{P}_{13} \end{pmatrix} = \begin{pmatrix} -0.0391 \\ -0.0188 \\ 2.04 \end{pmatrix}$$

The estimated values of the measured quantities ($\hat{\mathbf{z}} = \mathbf{H} \hat{\mathbf{x}}$) are

$$\hat{\mathbf{z}} = \begin{pmatrix} \hat{P}_1 \\ \hat{P}_2 \\ \hat{P}_3 \\ \hat{P}_{13} \end{pmatrix} = \begin{pmatrix} 3.99 \\ -3.98 \\ -0.017 \\ 2.04 \end{pmatrix}$$

A first estimate of the reactance x_{13} is obtained as follows

$$\hat{x}_{13} = (\hat{\theta}_1 - \hat{\theta}_3) / \hat{P}_{13} = (0 + 0.0188) / 2.04 = 0.00922 \text{ p.u.}$$

Note that the actual value used in Example 3.8 is $x_{13} = 0.01$ p.u. Of course this estimate can be improved by (a) increasing the measurement redundancy; and (b) performing recursive state estimation runs under different operating conditions considering multiple scans of measurements (see Chap. 11).

3.7 PHYSICAL LEVEL MODELING

Most of the literature on static state estimation deals with network models of the type used in off-line power flow studies: the so called bus/branch model or network planning model. For many years this was also true of most of the practical implementations of state estimation. Only recently has the industry recognized the need for representation at the physical level, that is, the representation of a network as is – at the bus-section/switching-device level.

Conventional estimation first processes topological data in order to convert network data from the physical level to the power flow model level. To do this, it is necessary to assume that the data processed by the network topology processor is correct, although this is not always the case as a number of situations have shown. Thus, at least for parts of the network, it may be appropriate to use a physical level model. In this section the extension of the dc state estimation equations to represent bus-sections and switching devices explicitly is discussed.

In Example 3.1 the modeling of zero impedance branches (ZB's) by the introduction of the power flow in the ZB as an additional state variable and by the imposition of zero voltage drop across the ZB as a pseudo-measurement, was shown. This is necessary since when a zero impedance branch is involved the current does not follow Ohm's law and thus cannot be expressed in terms of the voltages (states) of the terminal buses. The same thing happens with closed breakers which are in fact zero impedance branches; hence, in the generalized state estimator, they will be modeled by a pseudo-measurement stating that

the voltage difference across the closed breaker is equal to zero, and by an additional state (the active power flow through the breaker).

An open breaker can be modeled the same way, with the additional state variable being the same as for the closed breaker; and the pseudo-measurement now representing the fact that the current is zero when the breaker is open. This concept can also be extended to a breaker with an unknown status, in which case the extra power flow state variable is simply added to the state estimation problem; no pseudo-measurement is included (neither $P_{km} = 0$ nor $\theta_{km} = 0$) since the status of the breaker is unknown.

Example 3.13:

Consider the system of Fig. 3.9 with the following telemetered data: $P_1^{meas} = 1.58$ p.u. and $P_2^{meas} = -1.49$ p.u. The corresponding variances are $\sigma_1^2 = .001$ and $\sigma_2^2 = .001$. Breaker 1 – 3 is explicitly modeled by imposing the pseudo-measurements $\theta_{13} = 0$ and $P_3 = 0$, with pseudo-measurement variances $\sigma_{\theta_{13}}^2 = \sigma_{P_3}^2 = .00001$ arbitrarily assigned.

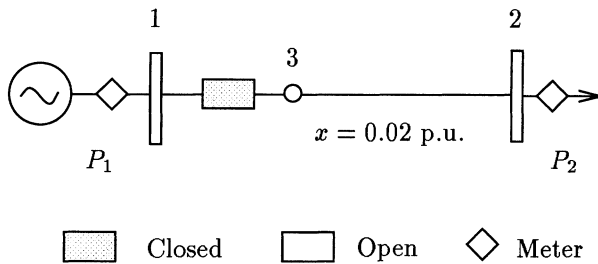


Figure 3.9. Physical-level network representation.

Thus the weighting matrix is

$$\mathbf{W} = \begin{pmatrix} 100000 & 0 & 0 & 0 \\ 0 & 100000 & 0 & 0 \\ 0 & 0 & 1000 & 0 \\ 0 & 0 & 0 & 1000 \end{pmatrix}$$

The measurement model assuming a reference angle $\theta_1 = 0^\circ$ is then

$$\begin{pmatrix} 0 & -1 & 0 \\ -50 & 50 & -1 \\ 0 & 0 & 1 \\ 50 & -50 & 0 \end{pmatrix} \begin{pmatrix} \theta_2 \\ \theta_3 \\ P_{13} \end{pmatrix} = \begin{pmatrix} \theta_{13}^{pseudo} \\ P_3^{pseudo} \\ P_1^{meas} \\ P_2^{meas} \end{pmatrix} = \begin{pmatrix} 0.0 \\ 0.0 \\ 1.58 \\ -1.49 \end{pmatrix}$$

The state estimate, $\hat{\mathbf{x}}$, is then

$$\hat{\mathbf{x}} = \mathbf{G}^{-1} \mathbf{H}' \mathbf{W} \mathbf{z} = \begin{pmatrix} \hat{\theta}_2 \\ \hat{\theta}_3 \\ \hat{P}_{13} \end{pmatrix} = \begin{pmatrix} -0.0307 \\ 0.0 \\ 1.54 \end{pmatrix}$$

Note that the estimated value for the power flow in the breaker is $P_{13} = 1.54$ p.u., which is consistent with the hypothesis that Breaker 1 – 3 is actually closed.

The estimated values of the measured quantities, $\hat{\mathbf{z}}$, and the estimated residuals, $\hat{\mathbf{r}}$, are as follows

$$\hat{\mathbf{z}} = \begin{pmatrix} \hat{\theta}_{13} \\ \hat{P}_3 \\ \hat{P}_1 \\ \hat{P}_2 \end{pmatrix} = \begin{pmatrix} 0.0 \\ -0.0005 \\ 1.54 \\ -1.53 \end{pmatrix}$$

$$\hat{\mathbf{r}} = \begin{pmatrix} \theta_{13}^{pseudo} - \hat{\theta}_{13} \\ P_3^{pseudo} - \hat{P}_3 \\ P_1^{meas} - \hat{P}_1 \\ P_2^{meas} - \hat{P}_2 \end{pmatrix} = \begin{pmatrix} 0.000 \\ 0.0004 \\ 0.045 \\ 0.045 \end{pmatrix}$$

Note that the residual associated with pseudo-measurement $\theta_{13}^{pseudo} = 0$ is zero since this measurement is critical (i.e. it is not a redundant measurement). The residual associated with the other pseudo-measurement, $P_3 = 0$, is smaller than the residuals associated with the telemetered data, since its weighting factor is ten times greater than those of the telemetered data.

Example 3.14:

For the same situation in the previous example, consider now that the status of Breaker 1 – 3 is unknown (see Fig. 3.10). All the other data remain the same as above.

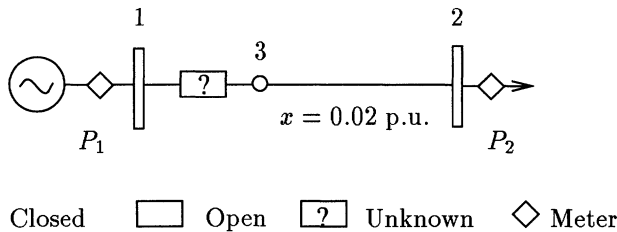


Figure 3.10. Physical-level network representation: breaker with an unknown status

The pseudo-measurement $\theta_{13}^{pseudo} = 0$ representing the closed breaker (as in the previous example) now has to be dropped from the model. Since the power flow P_{13} will be retained as a state variable, some indirect information about

the breaker status will still be available from the estimated value for this state variable.

Another difference concerns angle references, since there are now two separate networks, the first formed by Node 1 and the second by Nodes 2 and 3. Variable θ_2 will also be dropped from the model since it serves as the reference angle for the island formed by Nodes 2 and 3. The θ_1 has already been dropped, since it serves as the reference angle for the island formed by Node 1.

The resulting measurement model is then given by

$$\begin{pmatrix} 50 & -1 \\ 0 & 1 \\ -50 & 0 \end{pmatrix} \begin{pmatrix} \theta_3 \\ P_{13} \end{pmatrix} = \begin{pmatrix} P_3^{pseudo} \\ P_1^{meas} \\ P_2^{meas} \end{pmatrix} = \begin{pmatrix} 0.0 \\ 1.58 \\ -1.49 \end{pmatrix}$$

The state estimate, $\hat{\mathbf{x}}$, is

$$\hat{\mathbf{x}} = \mathbf{G}^{-1} \mathbf{H}' \mathbf{W} \mathbf{z} = \begin{pmatrix} \hat{\theta}_3 \\ \hat{P}_{13} \end{pmatrix} = \begin{pmatrix} 0.0307 \\ 1.54 \end{pmatrix}$$

Note that the estimated value for the power flow in the breaker is $P_{13} = 1.54$ p.u., which suggests that the breaker is most probably closed.

The estimated values of the measured quantities, $\hat{\mathbf{z}}$, are

$$\hat{\mathbf{z}} = \begin{pmatrix} \hat{P}_3 \\ \hat{P}_1 \\ \hat{P}_2 \end{pmatrix} = \begin{pmatrix} -0.0004 \\ 1.54 \\ -1.53 \end{pmatrix}$$

The estimated residuals, $\hat{\mathbf{r}}$, are

$$\hat{\mathbf{r}} = \begin{pmatrix} P_3^{pseudo} - \hat{P}_3 \\ P_1^{meas} - \hat{P}_1 \\ P_2^{meas} - \hat{P}_2 \end{pmatrix} = \begin{pmatrix} 0.0004 \\ 0.045 \\ 0.045 \end{pmatrix}$$

Example 3.15:

Figure 3.11 presents a somewhat more complex situation. The measured values are $P_3^{meas} = 1.55$ p.u., $P_4^{meas} = -0.62$ p.u., and $P_{65}^{meas} = -0.88$ p.u., and the corresponding variances are $\sigma_3^2 = .001$, $\sigma_4^2 = .001$ and $\sigma_{P_{65}}^2 = .001$. A pseudo-measurement variance $\sigma_{P_3^{pseudo}}^2 = .00001$ is also assigned.

In this case there are three separate islands, so three angular references are needed: θ_1 for the island formed by Node 1, θ_2 for the island formed by Nodes 2, 3, and 4, and θ_5 for the island formed by Nodes 5 and 6 (The automatic assignment of reference angles during observability analysis is discussed in Chap. 7.) Dropping the states θ_1 , θ_2 , and θ_5 (reference angles) yields the following measurement model

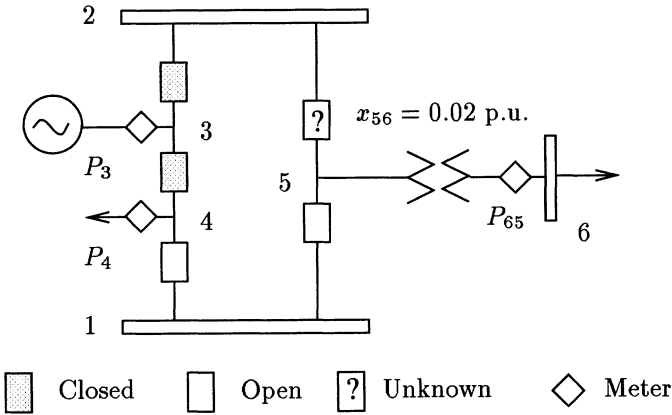


Figure 3.11. Physical-level network representation.

$$\begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & -50 & 0 & -1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & 1 \\ 0 & 0 & 0 & -1 & 0 & 0 & 0 & -1 \\ 0 & 0 & 50 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \theta_3 \\ \theta_4 \\ \theta_6 \\ P_{14} \\ P_{15} \\ P_{23} \\ P_{25} \\ P_{34} \end{pmatrix} = \begin{pmatrix} P_{14}^{pseudo} = 0 \\ P_{15}^{pseudo} = 0 \\ \theta_{23}^{pseudo} = 0 \\ \theta_{34}^{pseudo} = 0 \\ P_1^{pseudo} = 0 \\ P_2^{pseudo} = 0 \\ P_5^{pseudo} = 0 \\ P_3^{meas} = 1.55 \\ P_4^{meas} = -0.62 \\ P_{65}^{meas} = -0.88 \end{pmatrix}$$

The state estimate, $\hat{\mathbf{x}}$, is thus

$$\hat{\mathbf{x}} = \begin{pmatrix} \theta_3 \\ \theta_4 \\ \theta_6 \\ P_{14} \\ P_{15} \\ P_{23} \\ P_{25} \\ P_{34} \end{pmatrix} = \begin{pmatrix} 0.0 \\ 0.0 \\ -0.0179 \\ -0.00005 \\ -0.00005 \\ -0.897 \\ 0.897 \\ 0.637 \end{pmatrix}$$

Note that the flows in the breakers assumed to be closed are $P_{23} = -0.897$ p.u. and $P_{34} = 0.637$ p.u., while the flows in the open breakers are $P_{14} = -0.00005$ p.u. and $P_{15} = -0.00005$ p.u.; moreover, the breaker with a status considered to be unknown has flow $P_{25} = 0.897$ p.u. This suggests that the actual status of Breaker 2 – 5 is probably closed (See Chap. 9 for a more comprehensive analysis of equality constraint hypotheses testing.)

The estimated values of the measured quantities, $\hat{\mathbf{z}}$, are then

$$\hat{\mathbf{z}} = \begin{pmatrix} \hat{P}_{14} \\ \hat{P}_{15} \\ \hat{\theta}_{23} \\ \hat{\theta}_{34} \\ \hat{P}_1 \\ \hat{P}_2 \\ \hat{P}_5 \\ \hat{P}_3 \\ \hat{P}_4 \\ \hat{P}_{65} \end{pmatrix} = \begin{pmatrix} -0.00005 \\ -0.00005 \\ 0.0 \\ 0.0 \\ -0.00011 \\ -0.00016 \\ -0.00016 \\ 1.53 \\ -0.637 \\ -0.897 \end{pmatrix}$$

These results too are consistent with the hypothesis that Breaker 2 – 5 is actually closed, since there is a continuous power trajectory between Node 3 (or 4) and Node 6.

3.8 HISTORICAL NOTES AND REFERENCES

A power network static state estimation problem based on a power flow model (bus/branch model) was first formulated by Schweppe [1969], and the weighted least squares approach was proposed for solving it. This solution implicitly assumes that the physical level model, based on a bus-section/switching-device representation, has been successfully processed by the network topology processor and that the resulting model is exact. A review of the state of the art in state estimation algorithms based on the modeling approach of Schweppe (the bus/branch approach) was published by Bose and Clements [1987]. A radically different path was suggested by Irving and Sterling [1982, 1983]: (a) modeling at the bus-section/switching-device level; and (b) an LP based solution approach. The modeling of zero impedance branches for conventional WLS state estimators was proposed by Monticelli and Garcia [1991] and a WLS state estimator for networks modeled at the bus-section/switching-device level (physical level model) was introduced by Monticelli [1993]. A generalized state estimator with integrated state, status, and parameter estimation capability has recently been proposed by Alsaç, Vempati, Stott, and Monticelli [1998].

3.9 PROBLEMS

- 1. For the network in Fig. 3.12, define a set of state variables and determine the measurement model.
- 2. For the network in Fig. 3.6, establish a measurement model considering: (a) P_{12} and P_{13} as state variables; (b) P_1 and P_3 as state variables. Compare the results with those obtained for Example 3.6. Discuss the pros and cons of each of the three choices of state variables.

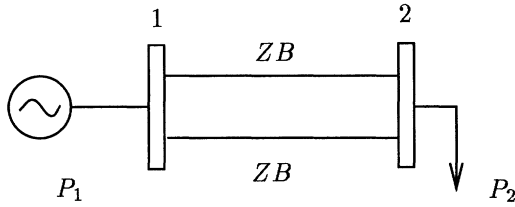


Figure 3.12. Two-bus system with two parallel zero-impedance branches (ZB's).

- 3. Consider the network in Fig. 3.5 (a), but with two meters added to measure loads P_{L1} and P_{L2} . Establish the measurement model for the resulting system considering both P_{L1} and P_{L2} as state variables. Repeat the problem merging the two measurements into a single injection measurement and considering the voltage angle θ_2 as a state variable. Compare and discuss the results.
- 4. For the network shown in Fig. 3.1, shown (a) that P_2 and P_3 form a possible pair of state variables; and (b) that P_3 and P_{23} cannot be a pair of state variables.
- 5. Consider the phase-shift estimation in Example 3.10 again, but rather than including φ as an additional state variable, remove the branch where the shifter is located from the network model (i.e. remove the shifter along with the reactance x_{12}). In this new situation Branch 1–2 will be modeled by a new state P_{12} . (a) Formulate the new measurement model. (b) Solve the state estimation problem using the same set of measurements as in Example 3.10. (c) Obtain an estimate of φ from the results in (b).
- 6. Analyze the impact of the arbitrary values chosen for the reference angles (θ_1 , θ_2 , and θ_5) on the estimates obtained in Example 3.15.
- 7. Analyze the impact of the weighting factors assigned to the pseudo-measurements in Example 3.15. Compute the states and the pseudo-measurement estimates considering the following: (a) $\sigma_{pseudo}^2 = .0001$; (b) $\sigma_{pseudo}^2 = .000001$. Compare the pseudo-measurement estimates with the corresponding standard deviation for (a) and (b).

References

- Alsaç, O., Vempati, N., Stott, B., and Monticelli, A., "Generalized state estimation", IEEE Trans. on Power Systems, Vol. 13, No. 3, pp. 1069-1075, August 1998.

- Bose, A. and Clements, K.A., "Real-time modeling of power networks", IEEE Proc., Special Issue on Computers in Power System Operations, Vol. 75, No. 12, pp 1607-1622, Dec. 1987.
- Irving, M.R., and Sterling, M.J.H., "Substation data validation", IEE Proc., Vol. 129, pt C, No. 3, pp. 119-122, May 1982.
- Irving, M.R., and Sterling, M.J.H., discussion of the paper: Bonanomi, P. and Gramberg, G., "Power system data validation and state calculation by network search techniques", IEEE Trans. PAS, Vol. 102, pp 238-249, January 1983.
- Monticelli, A. and Garcia, A., "Modeling zero-impedance branches in power-system state estimation". IEEE Transactions on Power Systems, Vol. 6, No. 4, pp. 1561-1570, Nov. 1991.
- Monticelli, A., "Modeling circuit breakers in weighted least squares state estimation", IEEE Transactions on Power Systems, Vol. 8, No. 3, pp 1143-1149, Aug. 1993.
- Schweppe, F.C., Wildes, J., and Rom, D. , "Power system static state estimation: Parts I, II, and III", Power Industry Computer Conference, PICA, Denver, Colorado, June 1969.

4 POWER FLOW EQUATIONS

This chapter reviews the power flow equations used in both power flow calculations and state estimation. The derivation of models of the main power network components is presented. Solvability conditions (“observability/controllability”) for the power flow problem are also discussed.

4.1 NETWORK BRANCH MODEL

4.1.1 Transmission Line

The equivalent π model of a transmission line presented in Fig. 4.1 is defined by three complex parameters: series impedance z_{km} ; shunt admittances y_{km}^{sh} and y_{mk}^{sh} . Branch series impedance and the corresponding admittance are as follows:

$$z_{km} = r_{km} + jx_{km}$$

$$y_{km} = z_{km}^{-1} = g_{km} + jb_{km}$$

Where the series conductance g_{km} and the series susceptance b_{km} are as follows:

$$g_{km} = \frac{r_{km}}{r_{km}^2 + x_{km}^2}$$

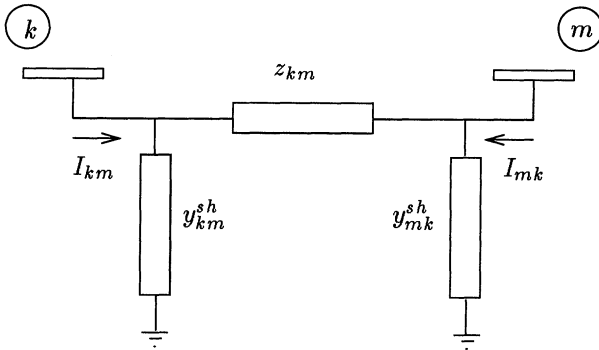


Figure 4.1. Equivalent π model of a transmission line.

$$b_{km} = -\frac{x_{km}}{r_{km}^2 + x_{km}^2}$$

In general, shunt admittance is expressed as

$$y_{km}^{sh} = g_{km}^{sh} + jb_{km}^{sh}$$

For actual transmission line sections r_{km} and x_{km} have positive values. This means that g_{km} is positive and b_{km} is negative (inductive susceptance); the shunt susceptance b_{km}^{sh} and the shunt conductance g_{km}^{sh} are both positive in real line sections. Note, however, that in certain types of reduced models (reduced equivalent networks), these parameters may assume values with signs that are different from those observed in actual transmission lines (e.g., in Ward-type equivalent networks where negative series reactances are observed).

The complex currents I_{km} and I_{mk} (Fig. 4.1) can be expressed as functions of the complex voltage at the branch terminal buses k and m :

$$I_{km} = y_{km}(E_k - E_m) + y_{km}^{sh} E_k \quad (4.1)$$

$$I_{mk} = y_{mk}(E_m - E_k) + y_{mk}^{sh} E_m \quad (4.2)$$

where the complex voltages are:

$$E_k = V_k e^{j\theta_k}$$

$$E_m = V_m e^{j\theta_m}$$

Example 4.1:

The series impedance of a 138 kV transmission line is

$$z = r + jx = 0.0062 + j0.0360 \text{ p.u.}$$

The total shunt susceptance (double the susceptance that appears in the equivalent π model) is

$$b^{sh} = 0.0104 \text{ p.u.}$$

and the shunt resistance is ignored. Series conductance and series susceptances are

$$g = \frac{r}{r^2 + x^2} = \frac{0.0062}{0.0062^2 + 0.0360^2} = 4.64 \text{ p.u.}$$

$$b = \frac{-x}{r^2 + x^2} = \frac{-0.0360}{0.0062^2 + 0.0360^2} = -27.0 \text{ p.u.}$$

The x/r ratio and the b/b^{sh} ratio are as follows:

$$\frac{x}{r} = \frac{0.0360}{0.0062} = 5.8$$

$$\frac{b}{b^{sh}} = \frac{-27.0}{0.0104} = -2596$$

Example 4.2:

The series impedance and the total shunt susceptance of a 750 kV transmission line are

$$z = 0.00072 + j0.0175 \text{ p.u.}$$

$$b^{sh} = 8.77 \text{ p.u.}$$

As in the previous example, this value is the double of the susceptance that appears in the π -model.

Series conductance and susceptance are

$$g = 2.35 \text{ p.u.}$$

$$b = -57.0 \text{ p.u.}$$

And the x/r ratio and the b/b^{sh} ratio are as follows:

$$\frac{x}{r} = 24.3$$

$$\frac{b}{b^{sh}} = -6.5$$

Remarks: Note that the 750 kV line has a much higher x/r ratio than the 138 kV line and, at the same time, a much smaller (in magnitude) b/b^{sh} ratio. Higher x/r ratios in general mean better decoupling between active and reactive parts of the power flow problem, while smaller $|b/b^{sh}|$ may indicate the need for some sort of compensation, either of the series, the shunt, or both.

4.1.2 Transformer

Figure 4.2 shows an equivalent model of a transformer formed by an ideal transformer on the primary side with turns ratio t_{km} and a series impedance z_{km} which represents resistive losses and the leakage reactance. Network data are usually formatted as in (b), and although both representations are equivalent, (a) leads to simpler power flow expressions and will then be used in the following; the conversion of data to this format is trivial since $t_{km} = 1/\bar{t}_{km}$.

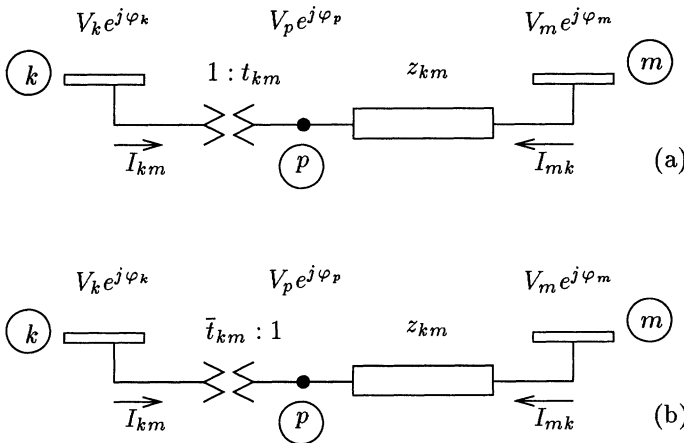


Figure 4.2. Transformer model with complex ratio $t_{km} = a_{km}e^{j\varphi_{km}}$ ($\bar{t}_{km} = a_{km}^{-1}e^{-j\varphi_{km}}$)

4.1.3 In-phase Transformer

Figure 4.3 shows an in-phase transformer model indicating the voltage at the intermediate node p . In this model the ideal voltage magnitude ratio is

$$\frac{V_p}{V_k} = a_{km}$$

Since $\theta_k = \theta_p$, this is also the ratio between the complex voltages at nodes k and p ,

$$\frac{E_p}{E_k} = \frac{V_p e^{j\theta_p}}{V_k e^{j\theta_k}} = a_{km} \tag{4.3}$$

There are no power losses (neither active nor reactive) in the ideal transformer (the $k - p$ part of the transformer model). This yields

$$E_k I_{km}^* + E_p I_{mk}^* = 0 \tag{4.4}$$

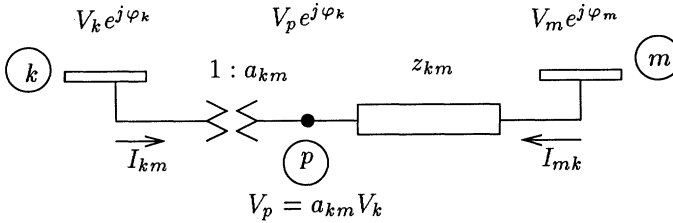


Figure 4.3. In-phase transformer model.

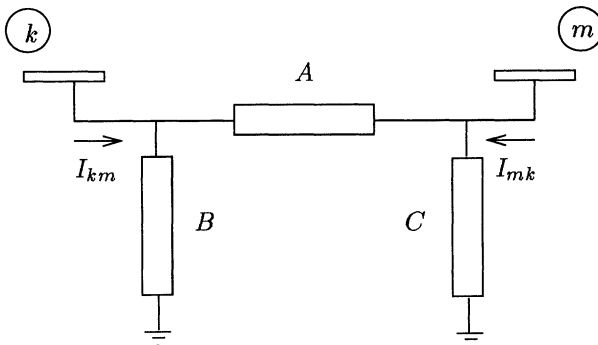


Figure 4.4. Equivalent π model for in-phase transformer.

Then applying Eqs. (4.3) and (4.4):

$$\frac{I_{km}}{I_{mk}} = -\frac{|I_{km}|}{|I_{mk}|} = -a_{km},$$

i.e., the complex currents I_{km} and I_{mk} are out of phase by 180° .

Figure 4.4 represents the equivalent π model for the in-phase transformer in Fig. 4.3. Parameters A , B , and C of this model can be obtained by identifying the coefficients of the expressions for the complex currents I_{km} and I_{mk} associated with the models of Figs. 4.3 and 4.4. Figure 4.3 gives

$$I_{km} = -a_{km}y_{km}(E_m - E_p) = (a_{km}^2y_{km})E_k + (-a_{km}y_{km})E_m \quad (4.5)$$

$$I_{mk} = y_{km}(E_m - E_p) = (-a_{km}y_{km})E_k + (y_{km})E_m \quad (4.6)$$

And Fig. 4.4 provides the following:

$$I_{km} = (A + B)E_k + (-A)E_m \quad (4.7)$$

$$I_{mk} = (-A)E_k + (A + C)E_m \quad (4.8)$$

Identifying the coefficients of E_k and E_m from expressions (4.5)-(4.6) and (4.7)-(4.8) yields

$$A = a_{km}y_{km}$$

$$B = a_{km}(a_{km} - 1)y_{km}$$

$$C = (1 - a_{km})y_{km}$$

Example 4.3:

A 138/69 kV in-phase transformer with a series resistance of zero, a 0.23 p.u. series reactance, and a p.u. turns-ratio of 1 : 1.030 (from the model in Fig. 4.3) will give the following equivalent π model parameters

$$A = a_{km}y_{km} = 1.030 (0.230)^{-1} = -j4.48 \text{ p.u.}$$

$$B = a_{km}(a_{km} - 1)y_{km} = 1.030 (1.030 - 1.000) (0.230)^{-1} = -j0.13 \text{ p.u.}$$

$$C = (1 - a_{km})y_{km} = (1.000 - 1.030) (0.230)^{-1} = j0.13 \text{ p.u.}$$

Hence, since A , B and C denote admittances, A and B are inductive, and C is capacitive.

Example 4.4:

A 500/750 kV in-phase transformer with series resistance of zero, a 0.00623 p.u. series reactance, and p.u. turns ratio of 1 : 0.950 (from the model in Fig. 4.3) will give the following parameters for the equivalent π model:

$$A = -j152.5 \text{ p.u.}$$

$$B = j7.62 \text{ p.u.}$$

$$C = -j8.03 \text{ p.u.}$$

i.e., parameter B is capacitive and parameter C is inductive (762 MVar and -803 MVar, respectively, assuming nominal voltage magnitudes and a 100 MVA base).

4.1.4 Phase-Shifting Transformer

Phase-shifting transformers, such as the one represented in Fig. 4.5, are used to control active power flows; the control variable is the phase angle and the controlled variable can be, among other possibilities, the active power flow in the branch where the shifter is placed.

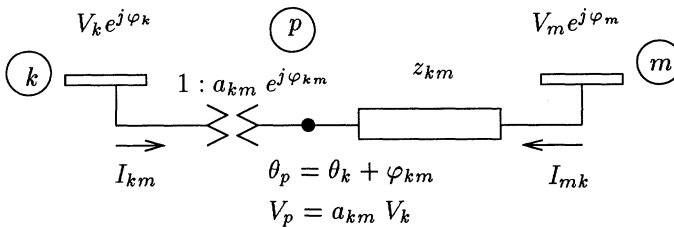


Figure 4.5. Phase-shifting transformer with $t = a e^{j\varphi}$.

A phase-shifting transformer affects both the phases and magnitudes of complex voltages E_k and E_p , without changing their ratio, i.e.,

$$\frac{E_p}{E_k} = t_{km} = a_{km} e^{j\varphi_{km}} \tag{4.9}$$

Thus, $\theta_p = \theta_k + \varphi_{km}$ and $V_p = a_{km} V_k$, and this yields, using Eq. (4.4) and Eq. (4.9),

$$\frac{I_{km}}{I_{mk}} = -t_{km}^* = -a_{km} e^{-j\varphi_{km}}$$

As with in-phase transformers, the complex currents I_{km} and I_{mk} can be expressed in terms of complex voltages at the shifter terminal nodes:

$$I_{km} = -t_{km}^* y_{km} (E_m - E_p) = (y_{km}) E_k + (-t_{km}^* y_{km}) E_m$$

$$I_{mk} = y_{km} (E_m - E_p) = (-t_{km} y_{km}) E_k + (y_{km}) E_m$$

There is no way to determine parameters A , B , and C of the equivalent π model from these equations, since the coefficient $-t_{km}^* y_{km}$ of E_m in the equation of I_{km} , differs from the coefficient $-t_{km} y_{km}$ of E_k in the equation of I_{mk} , as long as there is a nonzero phase shift.

Example 4.5:

A 230/138 kV transformer (Fig. 4.5) has a series resistance of zero, a 0.0127 p.u. series reactance, and a complex turns ratio of $1 : 1.007 e^{j30^\circ}$ (Y - Δ connection).

This transformer can be seen to consist of a series connection of two transformers: an ideal in-phase transformer with transformation ratio of $1 : 1.007$ (constant voltage phase) and a phase-shifting transformers with a complex turns ratio of $1 : e^{j30^\circ}$ (constant voltage magnitude) and series reactance of 0.0127 p.u.

4.1.5 Unified Branch Model

The expressions for the complex currents I_{km} and I_{mk} for both transformers and shifters, derived above, depend on the side the tap is located; i.e., they are not symmetrical. It is possible, however, to develop unified complex current expressions which can be used for lines, transformers, and shifters, regardless of the side on which the tap is located (or even for cases in which there are taps on both sides of the model). Consider initially the model in Fig. 4.6 in which shunt elements have been temporarily ignored and $t_{km} = a_{km} e^{j\varphi_{km}}$ and $t_{mk} = a_{mk} e^{j\varphi_{mk}}$. In this case

$$I_{km} = t_{km}^* I_{pq} = t_{km}^* (E_p - E_q) y_{km} = t_{km}^* (t_{km} E_k - t_{mk} E_m) y_{km}$$

$$I_{mk} = t_{mk}^* I_{qp} = t_{mk}^* (E_q - E_p) y_{km} = t_{mk}^* (t_{mk} E_m - t_{km} E_k) y_{km}$$

This yields the following symmetrical expressions:

$$I_{km} = (a_{km}^2 E_k - t_{km}^* t_{mk} E_m) y_{km}$$

$$I_{mk} = (a_{mk}^2 E_m - t_{mk}^* t_{km} E_k) y_{km}$$

(These expressions are symmetrical in the sense that if k and m are interchanged, as in the expression for I_{km} , the result is the expression for I_{mk} , and vice-versa.)

Figure 4.7 shows the unified branch model. All the devices studied above can be derived from this general model by establishing the appropriate definitions of the parameters that appear in the unified model. Thus, for instance, if $t_{km} = t_{mk} = 1.0$ is assumed, the result is a equivalent π model of a transmission line; or, if the shunt elements are ignored, and $t_{km} = 1.0$ and $t_{mk} = a_{mk} e^{-j\varphi_{mk}}$ is assumed, then the result is a phase-shifting transformer with tap located on the bus- m side. The general expressions for I_{km} and I_{mk} can be obtained from the model in Fig. 4.7:

$$I_{km} = (a_{km}^2 E_k - t_{km}^* t_{mk} E_m) y_{km} + y_{km}^{sh} a_{km}^2 E_k$$

$$I_{mk} = (a_{mk}^2 E_m - t_{mk}^* t_{km} E_k) y_{km} + y_{mk}^{sh} a_{mk}^2 E_m$$

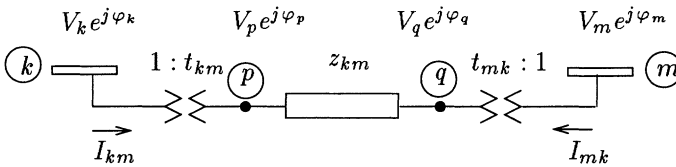


Figure 4.6. Transformer symmetrical model.

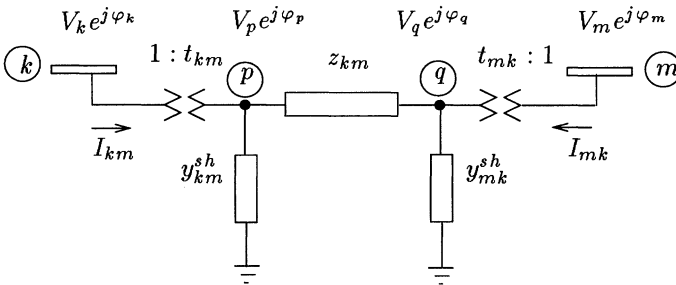


Figure 4.7. Unified branch model (π -model).

4.2 ACTIVE AND REACTIVE POWER FLOWS

In this section the expressions for the active and reactive power flows (P_{km} and Q_{km}) in transmission lines, transformers, phase-shifting transformers, and unified branch models are derived.

4.2.1 Transmission Line

Consider the complex current I_{km} in a transmission line

$$I_{km} = y_{km}(E_k - E_m) + jb_{km}^{sh} E_k;$$

The conjugate of the complex power flow ($S_{km}^* = P_{km} - jQ_{km}$) is

$$S_{km}^* = E_k^* I_{km} = y_{km} V_k e^{-j\theta_k} (V_k e^{j\theta_k} - V_m e^{j\theta_m}) + jb_{km}^{sh} V_k^2 \quad (4.10)$$

The expressions for P_{km} and Q_{km} can be determined by identifying the corresponding coefficients of the real and the imaginary parts of (4.10). This yields

$$P_{km} = V_k^2 g_{km} - V_k V_m g_{km} \cos\theta_{km} - V_k V_m b_{km} \sin\theta_{km} \quad (4.11)$$

$$Q_{km} = -V_k^2 (b_{km} + b_{km}^{sh}) + V_k V_m b_{km} \cos\theta_{km} - V_k V_m g_{km} \sin\theta_{km} \quad (4.12)$$

Active and reactive power flowing in opposite directions, P_{mk} and Q_{mk} , can be obtained in the same way, resulting in:

$$P_{mk} = V_m^2 g_{km} - V_k V_m g_{km} \cos\theta_{km} + V_k V_m b_{km} \sin\theta_{km}$$

$$Q_{mk} = -V_m^2 (b_{km} + b_{km}^{sh}) + V_k V_m b_{km} \cos\theta_{km} + V_k V_m g_{km} \sin\theta_{km}$$

The active and reactive power losses in the line are given by:

$$P_{km} + P_{mk} = g_{km}(V_k^2 + V_m^2 - 2V_k V_m \cos\theta_{km}) = g_{km}|E_k - E_m|^2$$

$$\begin{aligned} Q_{km} + Q_{mk} &= -b_{km}^{sh}(V_k^2 + V_m^2) - b_{km}(V_k^2 + V_m^2 - 2V_k V_m \cos\theta_{km}) \\ &= -b_{km}^{sh}(V_k^2 + V_m^2) - b_{km}|E_k - E_m|^2 \end{aligned}$$

Note that $|E_k - E_m|$ represents the magnitude of the voltage drop across the line, $g_{km}|E_k - E_m|^2$ represents the active power losses, $-b_{km}|E_k - E_m|^2$ represents reactive power losses; and $-b_{km}^{sh}(V_k^2 + V_m^2)$ represents the reactive power

generated by the shunt elements of the equivalent π model (assuming actual transmission line sections, i.e., with $b_{km} < 0$ and $b_{km}^{sh} > 0$).

Example 4.6:

A 750 kV transmission line section has a series impedance of $0.00072 + j0.0175$ p.u., a total shunt admittance of 8.775 p.u, a voltage magnitude at the terminal buses of 0.984 p.u. and 0.962 p.u., and a voltage angle spread of 22° .

The active and reactive power flows in the line are obtained by applying Eqs. (4.11)-(4.12), where $V_k = 0.984$ p.u., $V_m = 0.962$ p.u., and $\theta_{km} = 22^\circ$. The series impedance and admittances are as follows:

$$z_{km} = 0.00072 + j0.0175 \text{ p.u.}$$

$$y_{km} = g_{km} + jb_{km} = z_{km}^{-1} = 2.347 - j57.05 \text{ p.u.}$$

The π model shunt admittances (100 MVA base) are:

$$b_{km}^{sh} = 8.775/2 = 4.387 \text{ p.u.}$$

The active and reactive power flows can then be expressed as

$$P_{km} = 0.984^2 \cdot 2.347 - 0.984 \cdot 0.962 \cdot 2.347 \cos 22^\circ + 0.984 \cdot 0.962 \cdot 57.05 \sin 22^\circ$$

$$Q_{km} = 0.984^2 \cdot 52.66 - 0.984 \cdot 0.962 \cdot 57.05 \cos 22^\circ - 0.984 \cdot 0.962 \cdot 2.347 \sin 22^\circ$$

This yields the following results:

$$P_{km} = 2044 \text{ MW} \quad ; \quad Q_{km} = 8.4 \text{ MVAr}$$

4.2.2 In-phase Transformer

The complex current I_{km} in an in-phase transformer is expressed as in Eq. (4.5)

$$I_{km} = a_{km} y_{km} (a_{km} E_k - E_m)$$

The conjugate complex power flow ($S_{km}^* = P_{km} - jQ_{km}$) is given by

$$S_{km}^* = E_k^* I_{km} = y_{km} a_{km} V_k e^{-j\theta_k} (a_{km} V_k e^{j\theta_k} - V_m e^{j\theta_m}) \quad (4.13)$$

Separating the real and imaginary parts of this latter expression yields the active and reactive power flow equations:

$$P_{km} = (a_{km} V_k)^2 g_{km} - a_{km} V_k V_m g_{km} \cos \theta_{km} - a_{km} V_k V_m b_{km} \sin \theta_{km} \quad (4.14)$$

$$Q_{km} = -(a_{km} V_k)^2 b_{km} + a_{km} V_k V_m b_{km} \cos \theta_{km} - a_{km} V_k V_m g_{km} \sin \theta_{km} \quad (4.15)$$

These same expressions can be obtained by comparing Eqs. (4.13) and (4.10); in Eq. (4.13) the term $jb_{km}^{sh} V_k^2$ is not present, and V_k is replaced by $a_{km} V_k$. Hence, the expressions for the active and reactive power flows on in-phase transformers are the same expressions derived for a transmission line, except for two modifications: ignore b_{km}^{sh} , and replace V_k with $a_{km} V_k$.

Example 4.7:

A 500/750 kV transformer with a tap ratio of 1.050 : 1.0 on the 500 kV side (see Fig. 4.2 (b)), has negligible series resistance and a leakage reactance of 0.00623 p.u., terminal voltage magnitudes of 1.023 p.u. e 0.968 p.u., and an angle spread of 5.3° .

The active and reactive power flows in the transformer are given by Eqs. (4.14)-(4.15), where $V_k = 1.023$ p.u., $V_m = 0.968$ p.u., $\theta_{km} = 5.3^\circ$, and $a_{km} = 1.0/1.050 = 0.9524$. The series reactance and susceptance are as follows:

$$x_{km} = 0.00623 \text{ p.u.}$$

$$b_{km} = x_{km}^{-1} = -160.51 \text{ p.u.}$$

The active and reactive power flows can then be expressed as

$$P_{km} = 1.023 \ 0.9524 \ 0.968 \ 160.51 \ \sin 5.3^\circ$$

$$Q_{km} = (0.9524 \ 1.023)^2 \ 160.51 - (0.9524 \ 1.023) \ 0.968 \ 160.51 \ \cos 5.3^\circ$$

This yields

$$P_{km} = 1398 \text{ MW}$$

$$Q_{km} = 163 \text{ MVAR}$$

4.2.3 Phase-Shifting Transformer with $a_{km} = 1$

The complex current I_{km} in a phase-shifting transformer with $a_{km} = 1$ is as follows (see Fig. 4.5):

$$I_{km} = y_{km}(E_k - e^{-j\varphi_{km}} E_m) = y_{km} e^{-j\varphi_{km}} (E_k e^{j\varphi_{km}} - E_m) \quad (4.16)$$

The conjugate complex power flow ($S_{km}^* = P_{km} - jQ_{km}$) is thus

$$S_{km}^* = E_k^* I_{km} = y_{km} V_k e^{-j(\theta_k + \varphi_{km})} (V_k e^{j(\theta_k + \varphi_{km})} - V_m e^{j\theta_m}) \quad (4.17)$$

Separating the real and imaginary parts of this expression, yields the active and reactive power flow equations, P_{km} and Q_{km} , respectively:

$$P_{km} = V_k^2 g_{km} - V_k V_m g_{km} \cos(\theta_{km} + \varphi_{km}) - V_k V_m b_{km} \sin(\theta_{km} + \varphi_{km}) \quad (4.18)$$

$$Q_{km} = -V_k^2 b_{km} + V_k V_m b_{km} \cos(\theta_{km} + \varphi_{km}) - V_k V_m g_{km} \sin(\theta_{km} + \varphi_{km}) \quad (4.19)$$

As with transmission lines, these expressions could have been obtained by inspection by comparing Eqs. (4.10) and (4.17): in Eq. (4.17), the term $jb_{km}^s V_k^2$ is not present, and θ_k is replaced by $\theta_k + \varphi_{km}$. Hence, the expressions for the active and reactive power flows in phase-shifting transformers are the same expressions derived for the transmission line, albeit with two modifications: ignore b_{km}^s and replace θ_{km} with $\theta_{km} + \varphi_{km}$.

Example 4.8:

A $\Delta - Y$, 230/138 kV transformer presents a 30.0° phase angle (not compensated by other transformers in the network). Series resistance is neglected and series reactance is 0.0997 p.u. Terminal voltage magnitudes are 0.882 p.u. and 0.989 p.u., and the total angle spread is -16.6° .

The active and reactive power flows in this phase-shifting transformer are obtained from Eqs. (4.18)-(4.19), where $V_k = 0.882$ p.u., $V_m = 0.989$ p.u., $\theta_{km} = -16.6^\circ$, and $\varphi_{km} = 30.0^\circ$. The series reactance and series susceptance are thus

$$x_{km} = 0.0997 \text{ p.u.}$$

$$b_{km} = x_{km}^{-1} = -10.03 \text{ p.u.}$$

The active and reactive power flows can then be expressed as:

$$P_{km} = 0.882 \cdot 0.989 \cdot 10.03 \sin(-16.6^\circ + 30.0^\circ)$$

$$Q_{km} = 0.882^2 \cdot 10.03 - 0.882 \cdot 0.989 \cdot 10.03 \cos(-16.6^\circ + 30.0^\circ)$$

This yields

$$P_{km} = 203 \text{ MW}$$

$$Q_{km} = -70.8 \text{ MVA}_r$$

4.2.4 Unified Power Flow Equations

The expressions for active and reactive power flows on transmission lines, in-phase transformers, and phase-shifting transformers can be expressed in the following unified forms:

$$\begin{aligned}
 P_{km} = & (a_{km} V_k)^2 g_{km} \\
 & - (a_{km} V_k)(a_{mk} V_m) g_{km} \cos(\theta_{km} + \varphi_{km} - \varphi_{mk}) \\
 & - (a_{km} V_k)(a_{mk} V_m) b_{km} \sin(\theta_{km} + \varphi_{km} - \varphi_{mk}) \quad (4.20)
 \end{aligned}$$

$$\begin{aligned}
 Q_{km} = & - (a_{km} V_k)^2 (b_{km} + b_{km}^{sh}) \\
 & + (a_{km} V_k)(a_{mk} V_m) b_{km} \cos(\theta_{km} + \varphi_{km} - \varphi_{mk}) \\
 & - (a_{km} V_k)(a_{mk} V_m) g_{km} \sin(\theta_{km} + \varphi_{km} - \varphi_{mk}) \quad (4.21)
 \end{aligned}$$

Where, for transmission lines like the one represented in Fig. 4.1, $a_{km} = a_{mk} = 1$ and $\varphi_{km} = \varphi_{mk} = 0$; for in-phase transformers such as the one represented in Fig. 4.3, $y_{km}^{sh} = y_{mk}^{sh} = 0$, $a_{mk} = 1$, and $\varphi_{km} = \varphi_{mk} = 0$; and for a phase-shifting transformers such as the one in Fig. 4.5, $y_{km}^{sh} = y_{mk}^{sh} = 0$, $a_{mk} = 1$, and $\varphi_{mk} = 0$.

4.3 NODAL FORMULATION OF THE NETWORK EQUATIONS

The net complex current injection at a network bus (see Fig. 4.8) is related to the current flows in the branches incident to the bus. Applying Kirchoff's Current Law (KCL) yields

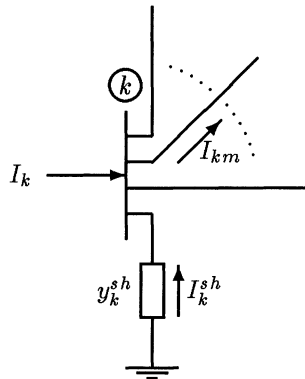


Figure 4.8. Generic bus with positive sign conventions for currents and power flows.

$$I_k + I_k^{sh} = \sum_{m \in \Omega_k} I_{km}; \quad \text{for } k = 1, \dots, N \quad (4.22)$$

where k is a generic node, m is a node adjacent to k , Ω_k is the set of nodes adjacent to k , and N is the number of nodes in the network.

The complex current I_{km} in the unified branch model (Fig. 4.7) is

$$I_{km} = (a_{km}^2 E_k - t_{km}^* t_{mk} E_m) y_{km} + y_{km}^{sh} a_{km}^2 E_k, \quad (4.23)$$

where $t_{km} = a_{km} e^{j\varphi_{km}}$ and $t_{mk} = a_{mk} e^{j\varphi_{mk}}$.

Equations (4.22) and (4.23) yield

$$I_k = \left(y_k^{sh} + \sum_{m \in \Omega_k} a_{km}^2 (y_{km}^{sh} + y_{km}) \right) E_k - \sum_{m \in \Omega_k} (t_{km}^* t_{mk} y_{km}) E_m,$$

for $k = 1, \dots, N$. This expression can be rewritten in matrix form as

$$\mathbf{I} = \mathbf{Y}\mathbf{E}$$

where

- \mathbf{I} is the injection vector with elements I_k ($k = 1, N$)
- \mathbf{E} is the nodal voltage vector with elements $E_k = V_k e^{j\theta_k}$
- $\mathbf{Y} = \mathbf{G} + j\mathbf{B}$ is the nodal admittance matrix, with the following elements:

$$Y_{km} = -t_{km}^* t_{mk} y_{km}$$

$$Y_{kk} = y_k^{sh} + \sum_{m \in \Omega_k} a_{km}^2 (y_{km}^{sh} + y_{km})$$

For large, practical networks, this matrix is usually very sparse. The degree of sparsity (percentage of zero elements) normally increases with the dimensions of the network: e.g., a network with 1000 buses and 1500 branches, typically presents a degree of sparsity greater than 99 % (i.e., less than 1 % of the matrix elements have non-zero values).

The k th component of \mathbf{I} , I_k , is given by the following:

$$I_k = Y_{kk} E_k + \sum_{m \in \Omega_k} Y_{km} E_m = \sum_{m \in K} Y_{km} E_m, \quad (4.24)$$

where K is the set of buses adjacent to Bus k , including Bus k , and Ω_k is the set of buses adjacent to Bus k , excluding Bus k .

Now, considering that $Y_{km} = G_{km} + jB_{km}$ and $E_m = V_m e^{j\theta_m}$, Eq. (4.24) can be rewritten as

$$I_k = \sum_{m \in K} (G_{km} + jB_{km})(V_m e^{j\theta_m}) \quad (4.25)$$

The conjugate complex power injection at Bus k is

$$S_k^* = P_k - jQ_k = E_k^* I_k \quad (4.26)$$

Applying Eqs. (4.25) and (4.26) and considering that $E_k^* = V_k e^{-j\theta_k}$, this gives the following:

$$S_k^* = V_k e^{-j\theta_k} \sum_{m \in K} (G_{km} + jB_{km})(V_m e^{j\theta_m}) \quad (4.27)$$

The expressions for active and reactive power injections are obtained by identifying the real and imaginary parts of Eq.(4.27), yielding

$$P_k = V_k \sum_{m \in K} V_m (G_{km} \cos\theta_{km} + B_{km} \sin\theta_{km}) \quad (4.28)$$

$$Q_k = V_k \sum_{m \in K} V_m (G_{km} \sin\theta_{km} - B_{km} \cos\theta_{km}) \quad (4.29)$$

4.4 BASIC POWER FLOW PROBLEM

4.4.1 Problem Variables

The power flow problem can be formulated as a set of non-linear algebraic equality/inequality constraints. These constraints represent both Kirchhoff's laws and network operation limits. In the basic formulation of the power flow problem, four variables are associated to each Bus k (network node):

- V_k - voltage magnitude (Bus k)
- θ_k - voltage angle
- P_k - net active power (algebraic sum of generation and load)
- Q_k - net reactive power

4.4.2 Basic Bus Types

Depending on which of the above four variables are known (given) and which ones are unknown (to be calculated), three basic types of buses can be defined:

- PQ bus: P_k and Q_k are specified, and V_k and θ_k are calculated;
- PV bus: P_k and V_k are specified, and Q_k and θ_k are calculated;

- $V\Theta$ (reference) bus: V_k and θ_k are specified, and P_k and Q_k are calculated.

PQ and PV buses are normally used to represent load and generation buses in power flow calculations. Synchronous condensers are also treated as PV buses. The $V\Theta$ bus has a double function in the basic formulation of the power flow problem: (a) it serves as the voltage angle reference; and (b) since active power losses are unknown in advance, the active power generation of the $V\Theta$ bus is used to balance generation, load and losses. Other possible bus types are: P, V, and PQV. The use of multiple $V\Theta$ buses may also be required for certain applications. In more general cases, the values are not limited to the specified set of bus variables (P, Q, V , and θ), and branch related variables can also be specified.

4.4.3 Problem Solvability

The major problem in the definition of bus type (bus classification) is to guarantee that the resulting set of power flow equations contains the same number of equations and unknowns, as are normally necessary for solvability, although not always sufficient. In terms of state estimation terminology, this means a minimally observable network in which there are as many unknowns as “measurements” (the “measurements”, of course, being the target values indicated by the corresponding bus type). Moreover, the “measurements” are evenly spread throughout the network. If correctly followed during data preparation, this procedure avoids the need for a detailed solvability analysis before data is fed into the power flow program. This is not true for state estimation, however, since data has to be processed “as is”; even if data is missing in parts of the network (unobservable parts) and redundant in others. Hence, there is no way to balance equations and unknowns as is done in power flow calculations.

Example 4.9:

Figure 4.9 shows a 5-bus network with four transmission lines and two transformers. Bus 1 and Bus 3 are PV buses, for which both the active power injection, P , and the voltage magnitude V are given. Bus 5 is also a generation bus, but here it has been chosen as the reference bus: the voltage angle, θ , and the voltage magnitude, V , are specified accordingly at this bus; the active power injection at this bus will be calculated in such a way as to balance generation, load and losses. Bus 2 is a transition bus in which both P and Q are equal to zero (a PQ bus). Bus 4 is a typical load bus to which is also connected a shunt susceptance: since this shunt is modeled as part of the network, the bus is also classified as a PQ bus.

4.4.4 Equality and Inequality Constraints

Equations (4.28) and (4.29) can be rewritten as follows:

$$P_k = \sum_{m \in \Omega_k} P_{km}(V_k, V_m, \theta_k, \theta_m)$$

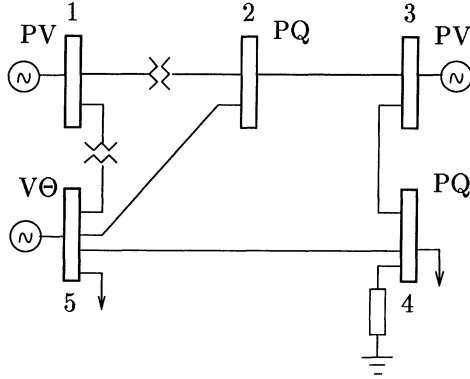


Figure 4.9. 5-bus system.

$$Q_k + Q_k^{sh}(V_k) = \sum_{m \in \Omega_k} Q_{km}(V_k, V_m, \theta_k, \theta_m)$$

where

- $k = 1, \dots, NB$ (NB is the number of buses in the network)
- Ω_k : set of buses adjacent to Bus k
- V_k, V_m : voltage magnitudes at terminal buses of Branch $k - m$
- θ_k, θ_m : voltage angles at terminal buses of Branch $k - m$
- P_{km} : active power flow from Bus k to Bus m
- Q_{km} : reactive power flow from Bus k to Bus m
- Q_k^{sh} : component of reactive power injection due to the shunt element at Bus k ($Q_k^{sh} = b_k^{sh} V_k^2$, where b_k^{sh} is shunt susceptance).

A set of inequality constraints imposes operating limits on variables such as the reactive power injections at PV buses and voltage magnitudes at PQ buses:

$$\begin{aligned} V_k^{min} &\leq V_k \leq V_k^{max} \\ Q_k^{min} &\leq Q_k \leq Q_k^{max} \end{aligned}$$

When a limit is satisfied, nothing happens, but if a limit is violated, its status is changed and it is enforced as an equality constraint at the limiting value. This normally requires a change in bus type: if, for example, a Q limit of a PV bus is violated, the bus is transformed into a PQ bus (Q is specified and V becomes a problem unknown). A similar procedure is adopted for backing-off,

whenever appropriate. What is crucial is that bus type changes must not affect solvability (“observability”). Various other types of limits are also considered in practical implementations, including branch current flows, branch power flows, active power generation levels, transformer taps, phase shifter angles, and area interchanges (Note that the strategies normally used for coordinating bus type changes implicitly account for the “controllability” analysis in the power flow problems.)

Example 4.10:

Consider the 5-bus network in Fig. 4.9, with the following sign conventions: power flows leaving a bus and power injections entering a bus are positive. The application of KCL to the network buses yields

$$\begin{aligned}
 P_1 &= P_{1,2} + P_{1,5} \\
 Q_1 &= Q_{1,2} + Q_{1,5} \\
 P_2 &= P_{2,1} + P_{2,3} + P_{2,5} \\
 Q_2 &= Q_{2,1} + Q_{2,3} + Q_{2,5} \\
 P_3 &= P_{3,2} + P_{3,4} \\
 Q_3 &= Q_{3,2} + Q_{3,4} \\
 P_4 &= P_{4,3} + P_{4,5} \\
 Q_4 + Q_4^{sh} &= Q_{4,3} + Q_{4,5} ; \\
 P_5 &= P_{5,1} + P_{5,3} + P_{5,4} \\
 Q_5 &= Q_{5,1} + Q_{5,3} + Q_{5,4}
 \end{aligned}$$

These equations can be expressed in terms of the network state variables (bus voltage magnitudes and angles). A power flow program solves this system of equations for the state variables subject to a number of inequality constraints such as:

$$\begin{aligned}
 Q_1^{min} &\leq Q_1 \leq Q_1^{max} \\
 Q_3^{min} &\leq Q_3 \leq Q_3^{max} \\
 Q_5^{min} &\leq Q_5 \leq Q_5^{max} \\
 V_2^{min} &\leq V_2 \leq V_2^{max} \\
 V_4^{min} &\leq V_4 \leq V_4^{max} \\
 P_{1,2}^{min} &\leq P_{1,2} \leq P_{1,2}^{max} \\
 P_{3,4}^{min} &\leq P_{3,4} \leq P_{3,4}^{max}
 \end{aligned}$$

4.5 NEWTON RAPHSON METHOD

A system of nonlinear algebraic equations can be written as

$$\mathbf{f}(\mathbf{x}) = \mathbf{0} \tag{4.30}$$

where \mathbf{x} is an n vector of unknowns and \mathbf{f} is an n vector function of \mathbf{x} . Given an appropriate starting value, \mathbf{x}^0 , the Newton Raphson method solves this vector equation by generating the following sequence:

$$\begin{aligned} \mathbf{J}(\mathbf{x}^\nu)\Delta\mathbf{x}^\nu &= -\mathbf{f}(\mathbf{x}^\nu) \\ \mathbf{x}^{\nu+1} &= \mathbf{x}^\nu + \Delta\mathbf{x}^\nu \end{aligned}$$

where $\mathbf{J}(\mathbf{x}^\nu) = \partial\mathbf{f}(\mathbf{x})/\partial\mathbf{x}$ is the Jacobian matrix.

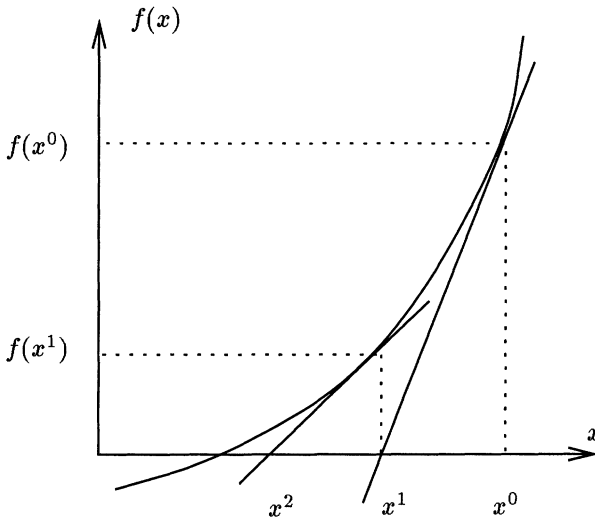


Figure 4.10. (a) Newton Raphson method.

4.5.1 Unidimensional Case

When the system is unidimensional Eq. (4.30) becomes:

$$f(x) = 0 \quad (4.31)$$

where x is the unknown and $f(x)$ is a scalar function of x . Fig. 4.10 illustrates a simple case in which there is a single solution to Eq. (4.31). Under these circumstances, the following algorithm can be used to find the solution to Eq. 4.31:

- i) Set $\nu = 0$ and choose an appropriate starting value x^0 ;
- ii) Compute $f(x^\nu)$;

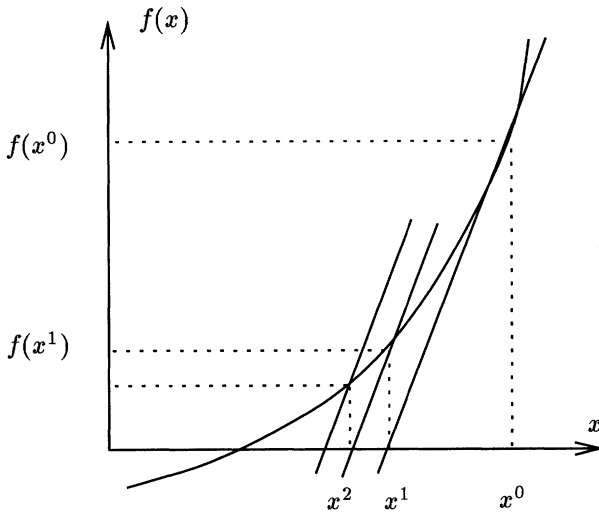


Figure 4.11. (a) Dishonest Newton Raphson method.

- iii) Compare $f(x^\nu)$ with specified tolerance ε ;
 if $|f(x^\nu)| \leq \varepsilon$, then $x = x^\nu$ is the solution to Eq. 4.31;
 Otherwise, if $|f(x^\nu)| > \varepsilon$, go to the next step;
- iv) Linearize $f(x)$ at the current solution point $[x^\nu; f(x^\nu)]$, as shown in Fig. 4.10. That is, $f(x^\nu + \Delta x^\nu) \cong f(x^\nu) + f'(x^\nu)\Delta x^\nu$, where $f'(x^\nu)$ is calculated at x^ν ;
- v) Solve $f(x^\nu) + f'(x^\nu)\Delta x^\nu = 0$ for Δx^ν , and update the solution estimate, $x^{\nu+1} = x^\nu + \Delta x^\nu$, where $\Delta x^\nu = -f(x^\nu)/f'(x^\nu)$;
- vi) Update iteration counter $\nu + 1 \rightarrow \nu$ and go to Step (ii).

The Dishonest Newton Raphson method is illustrated in Fig. 4.11. In this case, at Step (iv) of the algorithm, a constant derivative is assigned and $f'(x^\nu) = f'(x^0)$. Although the number of iterations required for convergence usually increases, it is not necessary to recalculate the derivatives for each iteration. When only limited accuracy is required, the overall performance of the dishonest version may be better than that of the full Newton Raphson method.

Example 4.11:

A 750 kV transmission line section has a series impedance of $z_{12}^{se} = 0.00072 + j0.0175$ p.u. The total line shunt admittance is $2b_{12}^{sh} = 8.775$ p.u. The voltage magnitude at the terminal buses is $V_1 = 0.984$ p.u. and $V_2 = 0.962$ p.u. The

active power generation at Bus 1 is $P_1 = 2044$ MW. The angle spread across the line must be calculated. Assume that $\epsilon = 0.005$.

The active power flow in the line is given by Eq. 4.11:

$$P_{12} = V_1^2 g_{12} - V_1 V_2 g_{12} \cos \theta_{12} - V_1 V_2 b_{12} \sin \theta_{12}$$

where $V_1 = 0.984$ p.u., $V_2 = 0.962$ p.u. The π model shunt admittances (100 MVA base) consists of:

$$b_{12}^{sh} = 8.775/2 = 4.387 \text{ p.u.}$$

The series susceptance and series conductance are as follows:

$$y_{12} = g_{12} + jb_{12} = z_{12}^{-1} = 2.347 - j57.05 \text{ p.u.}$$

Considering $\theta_2 = 0^\circ$ (Bus 2 is the reference bus), the active power injection at Bus 1 can be written as

$$P_1 = 20.44 = 2.273 - 2.221 \cos \theta_1 + 54.00 \sin \theta_1$$

In this case, Eq. (4.31) assumes the form of

$$f(\theta_1) = 18.17 + 2.221 \cos \theta_1 - 54.00 \sin \theta_1 = 0$$

Although in this case θ_1 could be determined analytically, for the sake of illustration it will be solved using the Newton Raphson algorithm described above, i.e.

$$\begin{aligned} f'(\theta_1^v) \Delta \theta_1^v &= -f(\theta_1^v) \\ \theta_1^{v+1} &= \theta_1^v + \Delta \theta_1^v \end{aligned}$$

where, $f'(\theta_1) = -2.222 \sin \theta_1 - 54.00 \cos \theta_1$.

■ First iteration: Let $\theta_1^0 = 0$, then:

$$f(\theta_1^0) = 20.39 \quad ; \quad f'(\theta_1^0) = -54.00 \quad ;$$

$$\Delta \theta_1^0 = -\frac{f(\theta_1^0)}{f'(\theta_1^0)} = -\frac{20.39}{-54.00} = 0.378 \quad ; \quad \theta_1^1 = \theta_1^0 + \Delta \theta_1^0 = 0.378$$

■ Second iteration:

$$f(\theta_1^1) = 0.325 \quad ; \quad f'(\theta_1^1) = -51.02 \quad ;$$

$$\Delta \theta_1^1 = -\frac{f(\theta_1^1)}{f'(\theta_1^1)} = -\frac{0.325}{-51.02} = 0.006 \quad ; \quad \theta_1^2 = \theta_1^1 + \Delta \theta_1^1 = 0.384$$

■ Third iteration:

$$f(\theta_1^2) = 0.0004 < \epsilon$$

The angle spread across the line is then $\theta_{12} = 22.0^\circ$.

4.5.2 Quadratic Convergence

Close to the solution point, x^* , the Newton Raphson method normally presents a property called quadratic convergence. This can be proved for the unidimensional case discussed above if it is assumed that x^* is a simple (not a multiple) root and that its first and the second derivatives are continuous.

Hence, $f'(x^*) = 0$, and for any x in a certain neighborhood of x^* , $f'(x) \neq 0$. If ϵ_ν denotes the error at the ν -th iteration, i.e.,

$$\epsilon_\nu = x^* - x^\nu$$

the Taylor series expansion about x^ν yields:

$$\begin{aligned} f(x^*) &= f(x^\nu + \epsilon_\nu) \\ &= f(x^\nu) + f'(x^\nu)\epsilon_\nu + 1/2 f''(\bar{x})\epsilon_\nu^2 \\ &= 0 \end{aligned}$$

where $\bar{x} \in [x^\nu, x^*]$. Dividing by $f'(x^\nu)$, this expression can be rewritten as follows:

$$\frac{f(x^\nu)}{f'(x^\nu)} + \epsilon_\nu + 1/2 \frac{f''(\bar{x})}{f'(x^\nu)} \epsilon_\nu^2 = 0$$

Since,

$$\frac{f(x^\nu)}{f'(x^\nu)} + \epsilon_\nu = \frac{f(x^\nu)}{f'(x^\nu)} + x^* - x^\nu = x^* - x^{\nu+1} = \epsilon_{\nu+1}$$

the following relationship between ϵ_ν and $\epsilon_{\nu+1}$ results:

$$\frac{\epsilon_{\nu+1}}{\epsilon_\nu^2} = -\frac{1}{2} \frac{f''(\bar{x})}{f'(x^\nu)}$$

In the vicinity of the root, i.e., as $x^\nu \rightarrow x^*$, $\bar{x} \rightarrow x^*$. Thus it results:

$$\frac{|\epsilon_{\nu+1}|}{\epsilon_\nu^2} = \frac{1}{2} \frac{f''(x^*)}{f'(x^*)}$$

A practical illustration of quadratic convergence is given in Example 2.7 in Chap. 2.

4.5.3 Multidimensional Case

Consider now the n -dimensional case

$$\mathbf{f}(\mathbf{x}) = \mathbf{0}$$

where

$$\mathbf{f}(\mathbf{x}) = \begin{pmatrix} f_1(\mathbf{x}) \\ f_2(\mathbf{x}) \\ \vdots \\ f_n(\mathbf{x}) \end{pmatrix} \quad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

The Newton Raphson method applied to solve Eq. (4.30) follows basically the same steps as those applied to the unidimensional case above, except that in Step (iv), Jacobian matrix $\mathbf{J}(\mathbf{x}^\nu)$ is used, and the linearization of $\mathbf{f}(\mathbf{x})$ at \mathbf{x}^ν is given by the Taylor expansion:

$$\mathbf{f}(\mathbf{x}^\nu + \Delta\mathbf{x}^\nu) \cong \mathbf{f}(\mathbf{x}^\nu) + \mathbf{J}(\mathbf{x}^\nu)\Delta\mathbf{x}^\nu$$

where the Jacobian matrix has the general form:

$$\mathbf{J} = \frac{\partial \mathbf{f}}{\partial \mathbf{x}} = \begin{array}{|c|c|c|c|} \hline \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \dots & \frac{\partial f_1}{\partial x_n} \\ \hline \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \dots & \frac{\partial f_2}{\partial x_n} \\ \hline \dots & \dots & \dots & \dots \\ \hline \frac{\partial f_n}{\partial x_1} & \frac{\partial f_n}{\partial x_2} & \dots & \frac{\partial f_n}{\partial x_n} \\ \hline \end{array}$$

The correction vector $\Delta\mathbf{x}$ is the solution to

$$\mathbf{f}(\mathbf{x}^\nu) + \mathbf{J}(\mathbf{x}^\nu)\Delta\mathbf{x}^\nu = 0$$

Note that this is the linearized version of the original problem $\mathbf{f}(\mathbf{x}^\nu + \Delta\mathbf{x}^\nu) = 0$.

The Newton Raphson algorithm for the n-dimensional case is thus as follows:

- i) Set $\nu = 0$ and choose an appropriate starting value \mathbf{x}^0 ;
- ii) Compute $\mathbf{f}(\mathbf{x}^\nu)$.
- iii) Test convergence: if $|f_i(\mathbf{x}^\nu)| \leq \epsilon$, for $i = 1, n$, \mathbf{x}^ν is the solution; Otherwise, go to (iv).
- iv) Compute the Jacobian matrix $\mathbf{J}(\mathbf{x}^\nu)$.
- v) Update the solution:

$$\mathbf{x}^{\nu+1} = \mathbf{x}^\nu + \Delta\mathbf{x}^\nu \tag{4.24}$$

$$\Delta\mathbf{x}^\nu = -\mathbf{J}^{-1}(\mathbf{x}^\nu)\mathbf{f}(\mathbf{x}^\nu)$$

vi) Set $v + 1 \rightarrow v$ and go to (ii).

Example 4.12:

Consider the case where the series impedance and total shunt susceptance of a 750 kV transmission line, z_{12} and b_{12}^{sh} , are as follows:

$$z_{12} = 0.00072 + j0.0175 \text{ p.u.}$$

$$b_{12}^{sh} = 4.387 \text{ p.u.}$$

The voltage magnitude at Bus 2 is $V_2 = 0.962$ p.u. The active and reactive power generation at Bus 1 are $P_1 = 2044$ MW and $Q_1 = 8.4$ MVar, respectively. The voltage magnitude at Bus 1 and the angle spread across the line are to be calculated. Assume that $\epsilon = 0.005$ p.u..

The active and reactive power flows in the line are given by Eqs. (4.11)-(4.12):

$$P_{12} = V_1^2 g_{12} - V_1 V_2 g_{12} \cos \theta_{12} - V_1 V_2 b_{12} \sin \theta_{12}$$

$$Q_{12} = -V_1^2 (b_{12} + b_{12}^{sh}) + V_1 V_2 b_{12} \cos \theta_{12} - V_1 V_2 g_{12} \sin \theta_{12}$$

where $g_{12} = 2.347$ p.u., $b_{12} = -57.05$ p.u., $b_{12}^{sh} = 4.387$ p.u., $V_2 = 0.962$ p.u., $P_1 = 20.44$ p.u., and $Q_1 = 0.084$ p.u.

Considering $\theta_2 = 0^\circ$ (Bus 2 is the reference bus), the active and the reactive power injections at Bus 1 can be written as:

$$P_1 = 2.347 V_1^2 - 2.258 V_1 \cos \theta_1 + 54.88 V_1 \sin \theta_1$$

$$Q_1 = 52.66 V_1^2 - 54.88 V_1 \cos \theta_1 - 2.258 V_1 \sin \theta_1$$

Equation (4.30) then assumes the following form:

$$f_1(V_1, \theta_1) = 20.44 - 2.347 V_1^2 + 2.258 V_1 \cos \theta_1 - 54.88 V_1 \sin \theta_1 = 0$$

$$f_2(V_1, \theta_1) = 0.084 - 52.66 V_1^2 + 54.88 V_1 \cos \theta_1 + 2.258 V_1 \sin \theta_1 = 0$$

The Newton Raphson iteration (Eq. (4.24) for the 2-dimension case is

$$\begin{pmatrix} \partial f_1 / \partial \theta_1 |^\nu & \partial f_1 / \partial V_1 |^\nu \\ \partial f_2 / \partial \theta_1 |^\nu & \partial f_2 / \partial V_1 |^\nu \end{pmatrix} \cdot \begin{pmatrix} \Delta \theta_1^\nu \\ \Delta V_1^\nu \end{pmatrix} = - \begin{pmatrix} f_1(V_1^\nu, \theta_1^\nu) \\ f_2(V_1^\nu, \theta_1^\nu) \end{pmatrix}$$

$$\begin{pmatrix} \theta_1^{\nu+1} \\ V_1^{\nu+1} \end{pmatrix} = \begin{pmatrix} \theta_1^\nu \\ V_1^\nu \end{pmatrix} + \begin{pmatrix} \Delta \theta_1^\nu \\ \Delta V_1^\nu \end{pmatrix}$$

where,

$$\begin{aligned}
\partial f_1 / \partial \theta_1 &= -2.258 V_1 \sin \theta_1 - 54.88 V_1 \cos \theta_1 \\
\partial f_1 / \partial V_1 &= -4.694 V_1 + 2.258 \cos \theta_1 - 54.88 \sin \theta_1 \\
\partial f_2 / \partial \theta_1 &= -54.88 V_1 \sin \theta_1 + 2.258 V_1 \cos \theta_1 \\
\partial f_2 / \partial V_1 &= -105.32 V_1 + 54.88 \cos \theta_1 + 2.258 \sin \theta_1
\end{aligned}$$

- First iteration: Let $\theta_1^0 = 0.$ and $V_1^0 = 0.$,

$$\begin{pmatrix} -54.88 & -2.436 \\ 2.258 & -50.44 \end{pmatrix} \cdot \begin{pmatrix} \Delta \theta_1^0 \\ \Delta V_1^0 \end{pmatrix} = - \begin{pmatrix} 20.35 \\ 2.304 \end{pmatrix}$$

$$\begin{pmatrix} \theta_1^1 \\ V_1^1 \end{pmatrix} = \begin{pmatrix} 0.0 \\ 1.0 \end{pmatrix} + \begin{pmatrix} 0.3681 \\ 0.0622 \end{pmatrix} = \begin{pmatrix} 0.3681 \\ 1.0622 \end{pmatrix}$$

- Second iteration:

$$\begin{pmatrix} -55.25 & -22.63 \\ -18.74 & -59.85 \end{pmatrix} \cdot \begin{pmatrix} \Delta \theta_1^1 \\ \Delta V_1^1 \end{pmatrix} = - \begin{pmatrix} -0.9441 \\ -4.076 \end{pmatrix}$$

$$\begin{pmatrix} \theta_1^2 \\ V_1^2 \end{pmatrix} = \begin{pmatrix} 0.3681 \\ 1.0622 \end{pmatrix} + \begin{pmatrix} 0.0124 \\ -0.0720 \end{pmatrix} = \begin{pmatrix} 0.3805 \\ 0.9902 \end{pmatrix}$$

- Third iteration:

$$\begin{pmatrix} -51.29 & -22.93 \\ -18.10 & -52.49 \end{pmatrix} \cdot \begin{pmatrix} \Delta \theta_1^2 \\ \Delta V_1^2 \end{pmatrix} = - \begin{pmatrix} 0.0357 \\ -0.2608 \end{pmatrix}$$

$$\begin{pmatrix} \theta_1^3 \\ V_1^3 \end{pmatrix} = \begin{pmatrix} 0.3805 \\ 0.9902 \end{pmatrix} + \begin{pmatrix} 0.00345 \\ -0.00616 \end{pmatrix} = \begin{pmatrix} 0.3839 \\ 0.9840 \end{pmatrix}$$

- Fourth iteration:

$$f_1(V_1^3, \theta_1^3) = 0.0011 < \epsilon$$

$$f_2(V_1^3, \theta_1^3) = -0.0020 < \epsilon$$

The voltage at the sending bus is $V_1 = 0.9840$ p.u. and the angle spread across the line is $\theta_{12} = 22.0^\circ$.

4.6 P-θ-QV DECOUPLING

The ac power flow problem above involves four variables associated with each network bus:

- V_k is the voltage magnitude (Bus k)
- θ_k is the voltage angle
- P_k is the net active power (generation - load)
- Q_k is the net reactive power

For transmission systems, a strong coupling is normally observed between P and θ as well as between Q and V . This property is used here to derive a linear approximation called dc load flow (or dc power flow). This linear model relates the active power P to the bus voltage angle θ .

Let us consider a π model of a transmission line where the series resistance and the shunt admittance are both zero. In this case, the active and reactive power flows are given by the following simplified expressions (Eqs. (4.11)-(4.12)):

$$P_{km} = V_k V_m x_{km}^{-1} \sin \theta_{km}$$

$$Q_{km} = -V_k V_m x_{km}^{-1} \cos \theta_{km} + V_k^2 x_{km}^{-1}$$

where x_{km} is the series reactance of the line.

The sensitivities between power flows P_{km} and Q_{km} and the state variables V e θ are given by

$$\frac{\partial P_{km}}{\partial \theta_k} = V_k V_m x_{km}^{-1} \cos \theta_{km} \quad \frac{\partial P_{km}}{\partial V_k} = V_m x_{km}^{-1} \sin \theta_{km}$$

$$\frac{\partial Q_{km}}{\partial \theta_k} = V_k V_m x_{km}^{-1} \sin \theta_{km} \quad \frac{\partial Q_{km}}{\partial V_k} = 2V_k x_{km}^{-1} - V_m x_{km}^{-1} \cos \theta_{km}$$

When $\theta_{km} = 0^\circ$, perfect decoupling conditions are observed, i.e.

$$\frac{\partial P_{km}}{\partial \theta_k} = V_k V_m x_{km}^{-1} \quad \frac{\partial P_{km}}{\partial V_k} = 0$$

$$\frac{\partial Q_{km}}{\partial \theta_k} = 0 \quad \frac{\partial Q_{km}}{\partial V_k} = 2V_k x_{km}^{-1} - V_m x_{km}^{-1}$$

As illustrated in Fig. 4.12, in the usual range of operations (relatively small voltage angles), a strong coupling between real power and voltage angle as well as between reactive power and voltage magnitude results; while a much

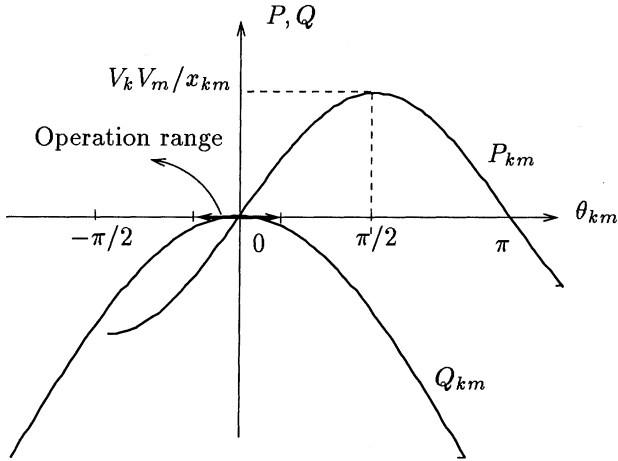


Figure 4.12. $P - \delta$ and $Q - \delta$ curves for a line with a series resistance and a shunt admittance of zero and considering terminal voltages $V_k = V_m = 1.0$ p.u..

weaker coupling between reactive power and voltage angle, and between voltage magnitude and real power also exists. Notice, however, that for larger angles this is no longer true. In the neighborhood of $\theta_{km} = 90^\circ$, there is strong coupling between P and V as well as between Q and θ .

Example 4.13:

A 750 kV transmission line has 0.0175 p.u. series reactance (the series resistance and the shunt admittance are ignored in this example); terminal bus voltage magnitudes of 0.984 p.u. and 0.962 p.u., and an angle spread of 10° are also given. The four sensitivities derived above assume the following values in this case:

$$\begin{aligned} \frac{\partial P_{km}}{\partial \theta_k} &= \frac{V_k V_m \cos \theta_{km}}{x_{km}} = \frac{0.984 \cdot 0.962 \cdot \cos 10^\circ}{0.0175} = 54.1 \\ \frac{\partial P_{km}}{\partial V_k} &= \frac{V_m \sin \theta_{km}}{x_{km}} = \frac{0.962 \cdot \sin 10^\circ}{0.0175} = 9.5 \\ \frac{\partial Q_{km}}{\partial \theta_k} &= \frac{V_k V_m \sin \theta_{km}}{x_{km}} = \frac{0.984 \cdot 0.962 \cdot \sin 10^\circ}{0.0175} = 9.4 \\ \frac{\partial Q_{km}}{\partial V_k} &= \frac{2V_k - V_m \cos \theta_{km}}{x_{km}} = \frac{2 \cdot 0.984 - 0.962 \cdot \cos 10^\circ}{0.0175} = 58.3 \end{aligned}$$

4.7 LINEARIZATION

In this section, the dc power flow equations are derived.

4.7.1 Transmission Line

Consider again expressions for the active power flows (P_{km} and P_{mk}) in a transmission line:

$$P_{km} = V_k^2 g_{km} - V_k V_m g_{km} \cos \theta_{km} - V_k V_m b_{km} \sin \theta_{km} \quad (4.25)$$

$$P_{mk} = V_m^2 g_{km} - V_k V_m g_{km} \cos \theta_{km} + V_k V_m b_{km} \sin \theta_{km} \quad (4.26)$$

These equations can be used to ascertain real power losses in a transmission line

$$P_{km} + P_{mk} = g_{km}(V_k^2 + V_m^2 - 2V_k V_m \cos \theta_{km}) \quad (4.27)$$

If the terms corresponding to the active power losses are ignored in Eqs. 4.25 and 4.26, the result is:

$$P_{km} = -P_{mk} = -V_k V_m b_{km} \sin \theta_{km} \quad (4.28)$$

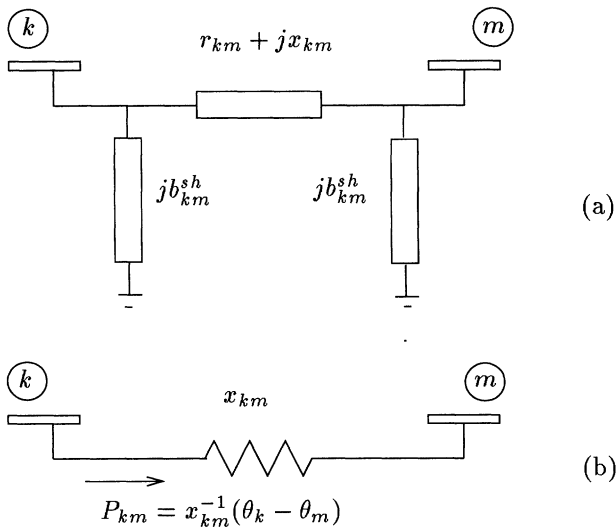


Figure 4.13. Transmission line. (a) Equivalent π model. (b) dc power flow model.

The following additional approximations are often valid

$$V_k \cong V_m \cong 1 \text{ p.u.}$$

$$\sin\theta_{km} \cong \theta_{km}$$

$$b_{km} \cong -\frac{1}{x_{km}}$$

Using these approximations to simplify the expression for the active power flow P_{km} yields

$$P_{km} = x_{km}^{-1} \theta_{km} = \frac{\theta_k - \theta_m}{x_{km}}$$

This equation is analogous to Ohm's Law applied to a resistor carrying a dc current: P_{km} is the dc current; θ_k and θ_m are the dc voltages at the resistor terminals; and x_{km} is the resistance.

Example 4.14:

A 750 kV transmission line has a $0.00072 + j0.0175$ p.u. series impedance. The measured terminal bus voltage magnitudes are 0.984 p.u. and 0.962 p.u., and the angle spread is 22° . The active power flow is given by the following nonlinear expression (Eq. 4.11):

$$P_{km} = V_k^2 g_{km} - V_k V_m g_{km} \cos\theta_{km} - V_k V_m b_{km} \sin\theta_{km}$$

where, $V_k = 0.984$ p.u., $V_m = 0.962$ p.u., $\theta_{km} = 22^\circ$. The series admittance is

$$y_{km} = g_{km} + jb_{km} = z_{km}^{-1} = 2.347 - j57.05 \text{ p.u.}$$

Inserting the numerical values yields

$$P_{km} = 0.984^2 \cdot 2.347 - 0.984 \cdot 0.962 \cdot 2.347 \cos 22^\circ +$$

$$+ 0.984 \cdot 0.962 \cdot 57.05 \sin 22^\circ$$

$$P_{km} = 20.44 \text{ p.u.} = 2044 \text{ MW}$$

In contrast, the linearized power flow equation gives the following:

$$P_{km} = x_{km}^{-1} \theta_{km} = \frac{0.384}{0.0175} = 21.94 \text{ p.u.} = 2194 \text{ MW}$$

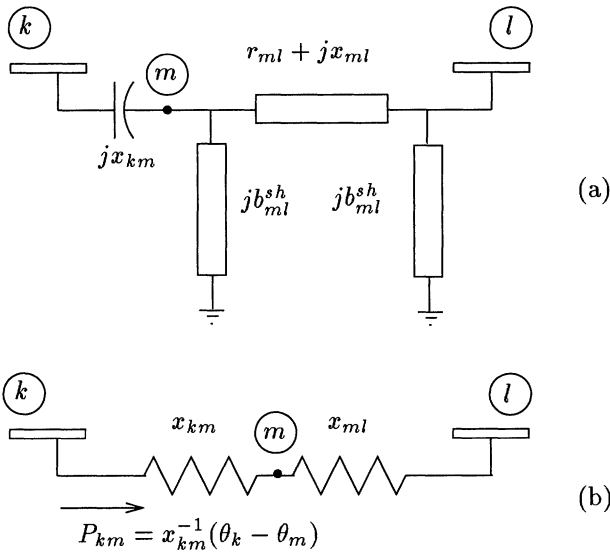


Figure 4.14. Transmission line with series compensation. (a) Series capacitor and a equivalent π model. (b) dc power flow model for the compensated line.

4.7.2 Series Capacitor

For a given voltage angle spread, the activer power flow in a transmission line decreases with the line reactance (and series reactance normally increases with line length). Series compensation aims at reducing the effective electric length of the line: a series capacitor is connected in series with the line. If, for example, a 40% compensation corresponds to a capacitor with a susceptance of 40% of the original line susceptance, the resulting susceptance of the compensated line becomes 60% of the original value.

Example 4.15:

A 50 % series compensation is introduced in a 750 kV transmission line with a series impedance of $0.00072 + j0.0175$ p.u. The measured bus voltage magnitudes at the capacitor terminals are 1.001 p.u. and 0.984 p.u., and the angle spread is -11° . The active power flow in the series capacitor is given by the non-linear expression:

$$P_{km} = -V_k V_m b_{km} \sin \theta_{km}$$

where, $V_k = 1.001$ p.u., $V_m = 0.984$ p.u., and $\theta_{km} = 11^\circ$. The capacitor series reactance and susceptance are the following:

$$x_{km} = -0.50 \ 0.0175 \text{ p.u.} = -0.00875 \text{ p.u.}$$

$$b_{km} = -x_{km}^{-1} = -(-0.00875)^{-1} = 114.3 \text{ p.u.}$$

The active power flow is given by the following equation:

$$P_{km} = -1.001 \ 0.962 \ 114.3 \ \sin(-11^\circ) = 21.00 \text{ p.u.} = 2100 \text{ MW}$$

In contrast, the dc power flow model results in the following:

$$P_{km} = x_{km}^{-1} \theta_{km} = \frac{-0.19}{-0.00875} = 21.94 \text{ p.u.} = 2194 \text{ MW}$$

4.7.3 In-Phase Transformer

Let us consider again expressions for active power flows (P_{km} and P_{mk}) in an in-phase transformer (Eq. 4.14):

$$P_{km} = a_{km}^2 V_k^2 g_{km} - a_{km} V_k V_m g_{km} \cos \theta_{km} - a_{km} V_k V_m b_{km} \sin \theta_{km}$$

Neglecting the terms associated with losses and introducing the same approximations used for transmission lines yields

$$P_{km} = a_{km} x_{km}^{-1} \theta_{km},$$

where further approximating $a_{km} \cong 1$, yields the same expression derived for transmission lines, i.e.:

$$P_{km} = x_{km}^{-1} \theta_{km}$$

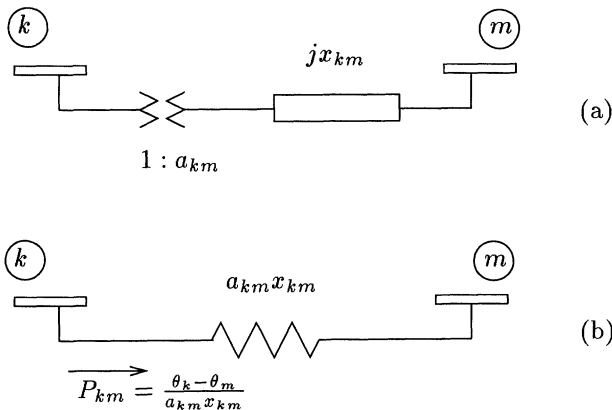


Figure 4.15. In-phase transformers: (a) transformer model comprising ideal transformer and series reactance; (b) dc power flow model.

Example 4.16:

A 500/750 kV transformer with a tap ratio of 1.050 : 1.0 on the low voltage side, has a leakage reactance of 0.00623 p.u., and series resistance and shunt admittance of zero, terminal voltages of 1.050 p.u. and 0.968 p.u., and an angle spread of 5.3°.

The active power flow is given by the non-linear expression:

$$P_{km} = a_{km}^2 V_k^2 g_{km} - a_{km} V_k V_m g_{km} \cos \theta_{km} - a_{km} V_k V_m b_{km} \sin \theta_{km}$$

where, $V_k = 1.050$ p.u., $V_m = 0.986$ p.u., $\theta_{km} = 5.3^\circ$, and $a_{km} = 1.0/1.050 = 0.9524$. The series reactance and susceptance are the following:

$$x_{km} = 0.00623 \text{ p.u.}$$

$$b_{km} = -x_{km}^{-1} = -(0.00623)^{-1} = -160.51 \text{ p.u.}$$

Hence,

$$\begin{aligned} P_{km} &= 0.9524 \cdot 1.050 \cdot 0.968 \cdot 160.51 \cdot \sin 5.3^\circ \\ &= 14.34 \text{ p.u.} = 1435 \text{ MW} \end{aligned}$$

In contrast, the dc power flow model, gives

$$P_{km} = a_{km} x_{km}^{-1} \theta_{km} = \frac{0.952 \cdot 0.0925}{0.00623} = 14.10 \text{ p.u.} = 1410 \text{ MW}$$

4.7.4 Phase-Shifter

Let us consider again the expression for the active power flow P_{km} in a phase-shifting transformers of the type represented in Fig. 4.5 with $a_{km} = 1$ (Eq. 4.18):

$$P_{km} = V_k^2 g_{km} - V_k V_m g_{km} \cos(\theta_{km} + \varphi_{km}) - V_k V_m b_{km} \sin(\theta_{km} + \varphi_{km})$$

As with transmission lines and in-phase transformers, if the terms associated with active power losses are ignored and $V_k = V_m = 1$ p.u. and $b_{km} = -x_{km}^{-1}$, the result is:

$$P_{km} = x_{km}^{-1} \sin(\theta_{km} + \varphi_{km})$$

If $(\theta_{km} + \varphi_{km}) \ll \pi/2$ linear approximation can be used:

$$P_{km} = x_{km}^{-1} (\theta_{km} + \varphi_{km}) \quad (4.29)$$

Note that P_{km} has two components, the first depending on the terminal bus voltage angles ($x_{km}^{-1}\theta_{km}$) and other depending only on the phase-shifting transformers angle ($x_{km}^{-1}\varphi_{km}$). If φ_{km} is considered to be a constant, Eq. 4.29 can be represented by the linearized model shown in Fig. 4.16, where the constant part of the active power flow ($x_{km}^{-1}\varphi_{km}$) appears as an extra load on the terminal Bus k and an extra generation on the terminal Bus m (or vice-versa, when $\varphi_{km} < 0$).

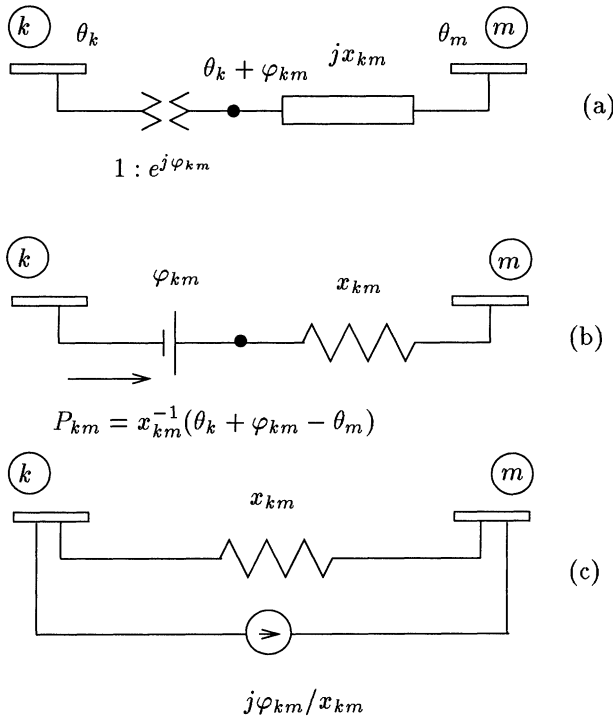


Figure 4.16. Phase-shifting transformer with $a_{km} = 1$. (a) Phase-shifting transformer model. (b) Thévenin dc power flow model. (c) Norton dc power flow model.

Example 4.17:

A 230/138 kV transformer introduces a 30° phase rotation ($\Delta - Y$ connection) and works as a phase-shifting transformers. The series resistance and the shunt admittance are ignored. The series reactance is 0.0997 p.u., and the terminal bus voltage magnitudes are equal to 0.882 p.u. and 0.989 p.u. The total angle spread is -16.6° . The active power flow in the shifter is

$$P_{km} = -V_k V_m b_{km} \sin(\theta_{km} + \varphi_{km})$$

where, $V_k = 0.882$ p.u., $V_m = 0.989$ p.u., $\theta_{km} = -16.6^\circ$ and $\varphi_{km} = 30^\circ$. The series susceptance is

$$b_{km} = -x_{km}^{-1} = (-0.0997)^{-1} = -10.03 \text{ p.u.}$$

Hence,

$$P_{km} = 0.882 \cdot 0.989 \cdot 10.03 \cdot \sin(-16.6^\circ + 30.0^\circ)$$

$$P_{km} = 2.03 \text{ p.u.} = 203 \text{ MW}$$

In contrast, the dc power flow model results in the following:

$$P_{km} = x_{km}^{-1}(\theta_{km} + \varphi_{km}) = \frac{0.23}{0.0997} = 2.35 \text{ p.u.} = 235 \text{ MW}$$

4.8 MATRIX FORMULATION

In this section, the dc model developed above is expressed in the form $\mathbf{I} = \mathbf{Y}\mathbf{E}$. According to the dc model, the active power flow in a branch is given by

$$P_{km} = x_{km}^{-1} \theta_{km}$$

where x_{km} is the series reactance of the branch (parallel equivalent of all the circuits existing in the branch).

The active power injection at Bus k is given by

$$P_k = \sum_{m \in \Omega_k} x_{km}^{-1} \theta_{km} = \left(\sum_{m \in \Omega_k} x_{km}^{-1} \right) \theta_k + \sum_{m \in \Omega_k} (-x_{km}^{-1} \theta_m)$$

for $k = 1, N$, where N is the number of buses in the network. This can be put into the matrix form as follows:

$$\mathbf{P} = \mathbf{B}' \boldsymbol{\Theta} \quad (4.30)$$

where

- $\boldsymbol{\Theta}$ is the vector of the voltage angles θ_k
- \mathbf{P} is the vector of the net active power injections P_k
- \mathbf{B}' is the nodal admittance matrix with the following elements:

$$B'_{km} = -x_{km}^{-1}$$

$$B'_{kk} = \sum_{m \in \Omega_k} x_{km}^{-1}$$

Matrix \mathbf{B}' in Eq. 4.30 is singular, i.e., with a determinant equal to zero (throughout the apostropher indicates transposition, except for the \mathbf{B}' matrix and the related matrix used in the fast decoupled power flow, where it simply reads prime). Since power losses have been ignored, the sum of the components of \mathbf{P} is equal to zero. This means that the rows of \mathbf{B}' are linearly dependent. One of the equations in the system is then removed, and the bus associated with that row is chosen as the angle reference bus ($\theta_k = 0^\circ$).

In forming matrix \mathbf{B}' , in-phase transformers and phase-shifting transformers are treated like transmission lines. The phase-shifting transformers also contribute to the construction of the independent vector \mathbf{P} with the Norton equivalent injections shown in Fig. 4.16 (c).

Example 4.18:

Consider the network given in Fig. 4.17 in which the reference angle is $\theta_1 = 0^\circ$.

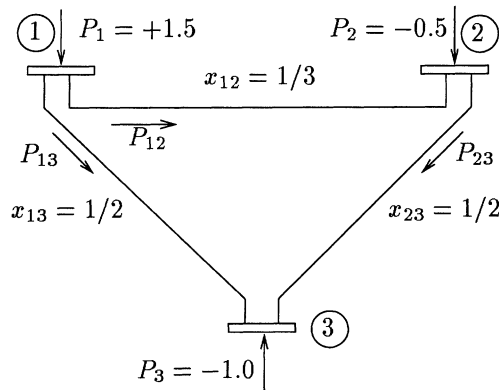


Figure 4.17. 3-bus network. (Active power in p.u.; Branch reactance in 10^{-2} p.u.)

In this case, the matrices \mathbf{B}' and $(\mathbf{B}')^{-1}$ are

$$\mathbf{B}' = \begin{matrix} & \begin{matrix} 2 & 3 \end{matrix} \\ \begin{matrix} 2 \\ 3 \end{matrix} & \begin{pmatrix} 500 & -200 \\ -200 & 400 \end{pmatrix} \end{matrix}$$

$$(\mathbf{B}')^{-1} = 10^{-2} \begin{pmatrix} 1/4 & 1/8 \\ 1/8 & 5/16 \end{pmatrix}$$

The nodal voltage angles (in radians) are

$$\Theta = (\mathbf{B}')^{-1} \mathbf{P} = 10^{-2} \begin{pmatrix} 1/4 & 1/8 \\ 1/8 & 5/16 \end{pmatrix} \cdot \begin{pmatrix} -0.5 \\ -1.0 \end{pmatrix}$$

$$\Theta = \begin{pmatrix} \theta_2 \\ \theta_3 \end{pmatrix} = \begin{pmatrix} -0.00250 \\ -0.00375 \end{pmatrix}$$

The power flows in the transmission lines are

$$P_{12} = x_{12}^{-1} \theta_{12} = 300 \cdot 0.0025 = 0.75 \text{ p.u.}$$

$$P_{13} = x_{13}^{-1} \theta_{13} = 200 \cdot 0.00375 = 0.75 \text{ p.u.}$$

$$P_{23} = x_{23}^{-1} \theta_{23} = 200 \cdot 0.00125 = 0.25 \text{ p.u.}$$

4.9 DC POWER FLOW MODEL

The linearized model $\mathbf{P} = \mathbf{B}' \Theta$ can be interpreted as the model for a network of resistors fed by dc current sources where \mathbf{P} is the vector of nodal current injections, Θ is the vector of nodal dc voltages, and \mathbf{B}' is the nodal conductance matrix, as illustrated in Fig. 4.18.

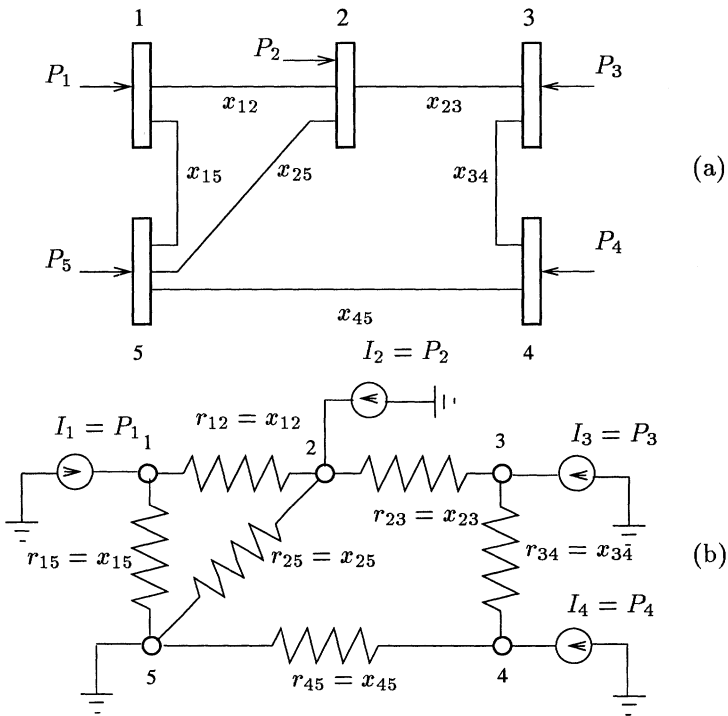


Figure 4.18. 6-bus network. (a) ac power network. (b) dc power flow model.

Example 4.19:

Consider the 5-bus network in Fig. 4.18. Taking Bus 5 as the reference bus, the linearized model $\mathbf{P} = \mathbf{B}'\Theta$ is as follows:

$$\begin{pmatrix} P_1 \\ P_2 \\ P_3 \\ P_4 \end{pmatrix} = \begin{pmatrix} x_{12}^{-1} + x_{15}^{-1} & -x_{12}^{-1} & 0 & 0 \\ -x_{12}^{-1} & x_{12}^{-1} + x_{23}^{-1} + x_{25}^{-1} & -x_{2,3}^{-1} & 0 \\ 0 & -x_{23}^{-1} & x_{23}^{-1} + x_{34}^{-1} & -x_{34}^{-1} \\ 0 & 0 & -x_{34}^{-1} & x_{34}^{-1} + x_{45}^{-1} \end{pmatrix} \cdot \begin{pmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \\ \theta_4 \end{pmatrix}$$

As with nodal admittance matrices in general, the main diagonal element (k, k) is given by the sum of all susceptances incident to node k , with a negative sign, and the off-diagonal element (k, m) is given by the susceptance of Branch $k - m$.

4.10 HISTORICAL NOTES AND REFERENCES

DC network analyzers were analog computers built out of resistor and batteries (dc network) and used as approximate models for ac power networks ($P\theta$ model). AC network analyzers, however, were more accurate, since they could represent both active and reactive power behavior. Early digital computer power flow programs were designed to mimic the old network analyzer procedures (e.g., the Gauss-Seidel method is analogous to the procedures used to tune a network analyzer by adjusting iteratively the power balance of one node at a time). Van Ess and Griffin [1961] proposed the Newton Raphson method to solve the power flow problem and Tinney and Hart [1967] later introduced sparse matrix techniques as a means of implementing the Newton Raphson method for large networks. Early developments of digital power flow calculations were reviewed by Tinney and Powell [1971] and Stott [1974], and the latter used dc power flow to initiate the Newton Raphson method [1971]. The $P\theta$ half-iteration of the fast decoupled power flow is (Stott and Alsac [1974]) is closely related to the dc power flow. The book by Grainger and Stevenson [1994] describes the basic power network models. A more detailed discussion about the Newton Raphson method and its quadratic convergence property can be found in Chap. 6 of the book by Dahlquist and Björck [1974].

4.11 PROBLEMS

- 1. Consider the three bus network in Fig. 4.19. Line $k - l$ has a series impedance of $z_{kl} = 0.00072 + j0.0175$ p.u. and a total shunt admittance of $2b_{kl}^{sh} = 877.5$ MVar (100 MVA base). Transformer $m - k$ has a negligible series resistance, a leakage reactance $x_{mk} = 0.00623$ p.u., and a transform tap $a_{mk} = 0.960$. The shunt susceptance of Bus k is $b_k^{sh} = -300$ MVar (it is a shunt reactor). The voltage magnitude at the three buses is known to be: $V_l = 1.020$ p.u., $V_k = 0.970$ p.u., $V_m = 1.000$ p.u. Voltage angles are

$\theta_l = 0^\circ$, $\theta_k = 16.0^\circ$, and $\theta_m = 20.0^\circ$. Compute the active and reactive power injections in buses k , l , and m .

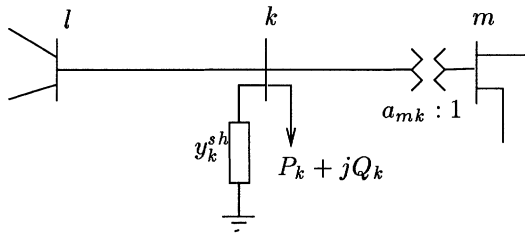


Figure 4.19. Network used in Probs. 1 and 6.

- 2. Form the blocked structure of the Jacobian matrix for the network shown in Fig. 4.9 (see Sec.5.8 in Chap. 5).
- 3. Verify the quadratic convergence for the systems in Examples 4.11 and 4.12 (Obtain a converged solution considering a tolerance of 10^{-8} p.u.)
- 4. A 500/750 kV transformer with a tap ratio of 1.050:1.000 in the 500 kV side (see Fig. 4.3), has negligible series resistance and a leakage reactance of 0.00623 p.u. Assuming the complex power injection on the 750 kV side is $9.50 + j2.40$ p.u., and the voltage magnitude on the 500 kV side is known to be 1.000 p.u., determine the voltage magnitude on the 750 kV side and the angle spread across the transformer. Use the Newton Raphson method and consider a convergence tolerance of 10^{-4} p.u. for the active and reactive power mismatches.
- 5. Repeat Prob. 4 using the fast decoupled power flows, versions XB and BX, as described in Chap. 12. Observe how the quadratic convergence is affected.
- 6. Repeat Prob. 1 considering now that the complex power injections at buses k and m are known (i.e., they are considered to be PQ buses with injection values given by the results obtained in Prob. 1) and that the voltage angle and voltage magnitude at the bus l is known to be $\theta_l = 0^\circ$ and $V_l = 1.020$ (slack bus). Use (a) the Newton Raphson method, and (b) the dishonest Newton Raphson method with a constant Jacobian matrix computed at flat-start ($V = 1.0$ p.u. and $\theta = 0^\circ$). Consider a convergence tolerance of 10^{-3} p.u. for the active and reactive power mismatches.

References

Dahlquist, G. and Björck, Å, *Numerical Methods*, Prentice Hall Series in Automatic Computation, 1974.

- Grainger, J. and Stevenson W., *Power System Analysis*, McGraw-Hill, 1994.
- Stott, B., "Effective starting process for Newton-Raphson load flows", Proc. of IEE, Vol. 118, pp. 983-987, Aug. 1971.
- Stott, B., "Review of load flow calculations methods", Proc. of IEEE 62, pp. 916-929, 1974.
- Stott, B. and Alsac, O., "Fast Decoupled Load Flow", IEEE Trans. Power App. Syst., Vol. 93, pp. 859-869, 1974.
- Tinney, W. F. and Hart, C. E., "Power flow solution by Newton's method", IEEE Trans. Power App. Syst., Vol. 86, pp. 1449-1456, 1967.
- Tinney, W. F. and Powell, W. L., "Notes on Newton-Raphson method for solution of AC power flow problem", IEEE Short Course, Power Systems Planning, 1971.
- Van Ess, J. E. and Griffin, J. H., "Elimination methods for load flow studies", AIEE Transactions, Vol. 80, pp. 299-304, 1961.

5 NETWORK REDUCTION AND GAUSS ELIMINATION

In this chapter, sparse vector and sparse matrix techniques are discussed. The principal properties on which sparse techniques are based are first presented using network concepts such as network reduction and network expansion. Gauss elimination is then introduced so these ideas can be used for solving large sparse network problems.

5.1 BUS ADMITTANCE MATRIX

If \mathbf{I} is the vector of nodal current injections, \mathbf{E} the vector of nodal voltages, and \mathbf{Y} the nodal admittance matrix, then

$$\mathbf{I} = \mathbf{Y}\mathbf{E}$$

The complex current injection at bus k is given by:

$$I_k = Y_{kk}E_k + \sum_{m \in \Omega_k} Y_{km}E_m$$

where Ω_k is the set of all buses adjacent to bus k . This means that the k th row of matrix \mathbf{Y} has non-zero elements (k, m) for $m = k$ and for $m \in \Omega_k$. Consider, for example, a 1000-bus network with a generic bus k connected to three adjacent buses. To this bus corresponds a row in the \mathbf{Y} matrix with only four non-zero elements; the remaining 996 elements will be zero. Hence

most of the elements of the nodal admittance matrix of a large network are zero. The number of non zero elements increases only close to linearly, whereas the number of zeros approximately increases with the square of the number of network nodes (assuming that, for large networks, the average number of connections per bus does not vary with the size of the network).

The degree of sparsity of a matrix is defined as the percentage of non-zero elements in the matrix. For the nodal admittance matrix of a network with n buses and m branches, the degree of sparsity is given by:

$$S = \frac{n^2 - (n + 2m)}{n^2} \cdot 100 \%$$

The larger the size of the system, the closer S moves to 100. Thus, for example: for $n = 100$ and $m = 150$, $S = 96.0\%$; for $n = 1000$ and $m = 1500$, $S = 99.6\%$; and for $n = 10000$ and $m = 15000$, $S = 99.96\%$.

Example 5.1:

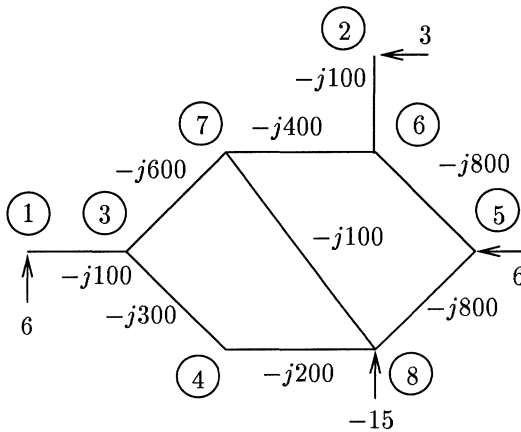


Figure 5.1. 8-bus network in Example 5.1 (Admittances are in p.u.)

Figure 5.1 shows an 8-bus, 9-branch network. The corresponding admittance matrix ($\mathbf{Y} = \mathbf{G} + j\mathbf{B}$) is

$$\mathbf{Y} = \underset{j}{\overset{i}{\begin{matrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ \begin{pmatrix} -100 & & 100 & & & & & \\ & -100 & & & & 100 & & \\ 100 & & -1000 & 300 & & & 600 & \\ & & 300 & -500 & & & & 200 \\ & & & & -1600 & 800 & & 800 \\ & 100 & & & 800 & -1300 & 400 & \\ & & 600 & & & 400 & -1100 & 100 \\ & & & 200 & 800 & & 100 & -1100 \end{pmatrix} & \end{matrix}}}$$

5.2 NETWORK REDUCTION AND EXPANSION

This section shows how the basic concepts needed to derive the main sparse matrix and sparse vector techniques can be understood from the study of network reduction (or equivalencing) and the reverse operation, called network expansion.

5.2.1 Network Reduction

Consider the situation represented in Fig. 5.2 (a) where k is a generic node and m is a node adjacent to k , i.e. $m \in \Omega_k$; Node k is to be eliminated. Figure 5.2 shows the original network and the reduced network. If it is assumed that the network is modeled by $\mathbf{I} = \mathbf{Y}\mathbf{E}$ it follows

$$\begin{aligned}
 Y_{11}E_1 + \dots + Y_{1k}E_k + \dots + Y_{1n}E_n &= I_1 \\
 \dots & \\
 Y_{m1}E_1 + \dots + Y_{mk}E_k + \dots + Y_{mn}E_n &= I_m \quad (5.1) \\
 \dots & \\
 Y_{k1}E_1 + \dots + Y_{kk}E_k + \dots + Y_{kn}E_n &= I_k \\
 \dots & \\
 Y_{n1}E_1 + \dots + Y_{nk}E_k + \dots + Y_{nn}E_n &= I_n
 \end{aligned}$$

In terms of the $\mathbf{I} = \mathbf{Y}\mathbf{E}$ model, the elimination of Node k proceeds as follows. First, E_k is obtained from the k th equation, resulting in the following

$$E_k = Y_{kk}^{-1}(I_k - \sum_{m \in \Omega_k} Y_{km}E_m) \quad (5.2)$$

Second, E_k is eliminated from the remaining equations (equations m , with $m \in \Omega_k$). Finally, deleting the k th equation, results:

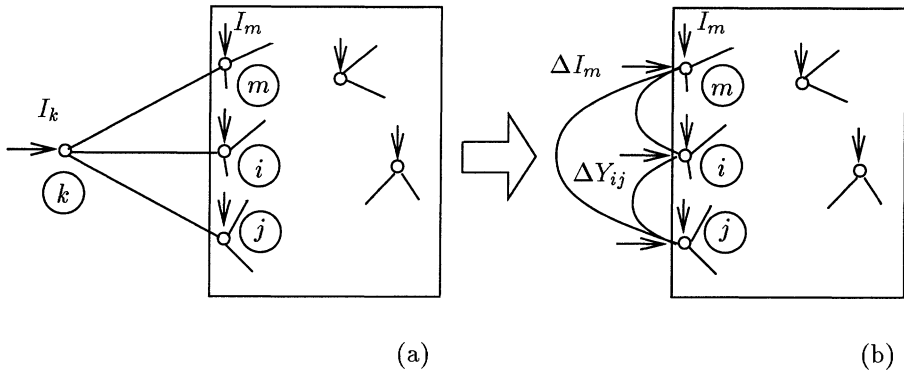


Figure 5.2. Network reduction by elimination of Node k . ΔY_{ij} , $i, j \in \Omega_k$, is the change in the admittance of Branch $i - j$ due to elimination of Node k . ΔI_m is the change in current injection at Node $m \in \Omega_k$, due to the distribution of current I_k on Node m (current divider).

$$\begin{aligned}
 Y_{11}^r E_1 + \dots + 0 + \dots + Y_{1n}^r E_n &= I_1^r \\
 \dots & \\
 Y_{m1}^r E_1 + \dots + 0 + \dots + Y_{mn}^r E_n &= I_m^r \quad (5.3) \\
 \dots & \\
 Y_{n1}^r E_1 + \dots + 0 + \dots + Y_{nn}^r E_n &= I_n^r
 \end{aligned}$$

Note that not all the coefficients Y_{ij}^r in Eq. (5.3) are different from the respective coefficients in Eq. (5.1). Actually, only the coefficients for which $i, j \in \Omega_k$ are modified according to Eq. (5.4). This is also true for the equivalent current injections I_m^r : the changes are limited to nodes for which $m \in \Omega_k$ (Eq. (5.5)).

In the reduced network the equivalent branch admittances connect the nodes originally adjacent to Node k :

$$Y_{ij}^r = Y_{ij} - \frac{Y_{ik} Y_{kj}}{Y_{kk}} \quad (5.4)$$

and the equivalent nodal current injections are as follows:

$$I_m^r = I_m - \frac{Y_{mk}}{Y_{kk}} I_k \quad (5.5)$$

Remarks: Note that the elimination of a node from a network corresponds to the application of Gauss elimination to the column of \mathbf{Y} associated with the

eliminated node (the pivot column). For symmetric matrices, regarding the calculation effort, it makes no difference performing elimination by rows or by columns. The elimination by columns, however, is more closely related to the elimination of a node and its incident branches from the network (network reduction). As for asymmetric matrices, or incident-symmetric matrices, it is computationally more efficient to perform elimination by rows.

Example 5.2:

For the network in Fig. 5.1, consider first the simultaneous elimination of Nodes 1 and 3. The elimination of Node 1 creates no new admittances since the neighborhood of Node 1 contains a single Node (Node 3); current injection I_1 is thus transferred to Node 3 according to Eq. (5.5):

$$I_3^r = I_3 - \frac{Y_{31}}{Y_{11}} I_1 = 0 - \frac{(j100.)}{(-j100.)} 6. = 6. \text{ p.u.}$$

The elimination of Node 3, however, creates a new admittance between Nodes 4 and 7 according Eq. (5.4):

$$Y_{47}^r = Y_{47} - \frac{Y_{43}Y_{37}}{Y_{33}} = j0. - \frac{(j300.)(j600.)}{(-j900.)} = j200. \text{ p.u.}$$

(Notice that the updated value $Y_{33} = -j900. \text{ p.u.}$ has been used.) The elimination of Node 3 makes current I_3 to be distributed between nodes 4 and 7, according to Eq. (5.5):

$$I_4^r = I_4 - \frac{Y_{43}}{Y_{33}} I_3 = 0 - \frac{(j300.)}{(-j900.)} 6. = 2. \text{ p.u.}$$

$$I_7^r = I_7 - \frac{Y_{73}}{Y_{33}} I_3 = 0 - \frac{(j600.)}{(-j900.)} 6. = 4. \text{ p.u.}$$

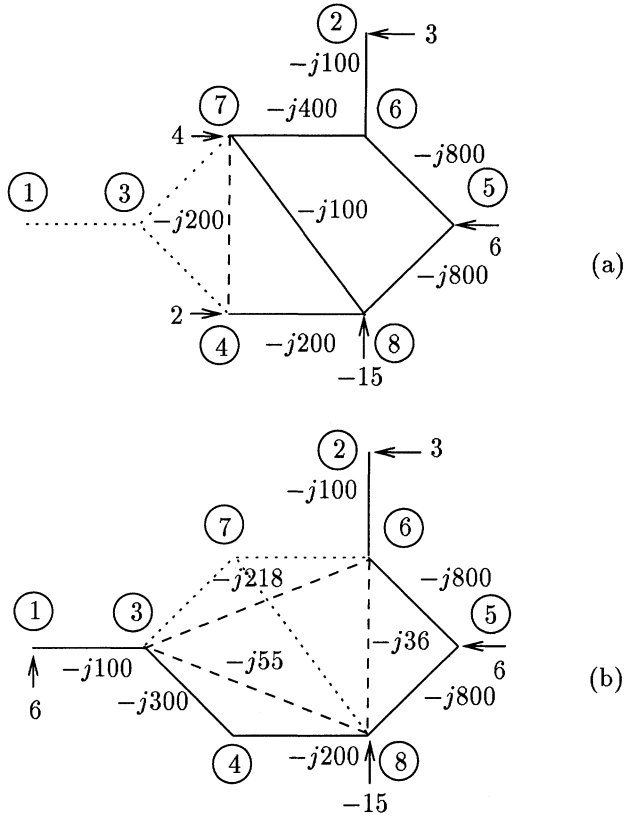
(This has the effect of a the current divider.) The resulting reduced network is given in Fig. 5.3 (a).

Now consider the elimination of Node 7 from the original network (5.1). This is a case of standard $Y-\Delta$ transformation. The resulting network is shown in Fig. 5.3 (b). The transformed admittances (Eq. (5.4)) and the transformed current injections (Eq. (5.5)) are as follows:

$$Y_{36}^r = Y_{36} - \frac{Y_{37}Y_{76}}{Y_{77}} = j0. - \frac{(j600.)(j400.)}{(-j1100.)} = j218. \text{ p.u.}$$

$$Y_{38}^r = Y_{38} - \frac{Y_{37}Y_{78}}{Y_{77}} = j0. - \frac{(j600.)(j100.)}{(-j1100.)} = j55. \text{ p.u.}$$

$$Y_{68}^r = Y_{68} - \frac{Y_{67}Y_{78}}{Y_{77}} = j0. - \frac{(j400.)(j100.)}{(-j1100.)} = j36. \text{ p.u.}$$



Key: ——— Branch of the original network
 Branch eliminated by the reduction.
 - - - - - Branch created by the reduction.

Figure 5.3. 8-bus network in Example 5.1. (a) Network after elimination of nodes 1 and 2. (b) Network after elimination of Node 7.

$$I_3^r = I_3 - \frac{Y_{37}}{Y_{77}} I_7 = 0. - \frac{(j600.)}{(-j1100.)} 0. = 0. \text{ p.u.}$$

$$I_6^r = I_6 - \frac{Y_{67}}{Y_{77}} I_7 = 0. - \frac{(j400.)}{(-j1100.)} 0. = 0. \text{ p.u.}$$

$$I_8^r = I_8 - \frac{Y_{87}}{Y_{77}} I_7 = -15. - \frac{(j100.)}{(-j1100.)} 0. = -15. \text{ p.u.}$$

5.2.2 Network Expansion

Consider a situation in which the determination of voltage E_k of the unreduced network (Fig. 5.2(a)) is difficult, but the network can be reduced. This reduced network can then be solved for nodal voltages E_m , $m \in \Omega_k$, and Eq. (5.2) used to compute the desired voltage E_k . The first operation, i.e., determining the reduced network, is called reduction; the second operation, obtaining the voltage of the original network from the data of the reduced network, is called expansion.

But, what if the reduced network is still difficult to solve? The same strategy can be applied recursively to the reduced network until it does become simple enough, if necessary to a single node (the reference node). This simplified problem can then be solved and its results used to solve progressively less reduced networks as one proceeds back up the chain.

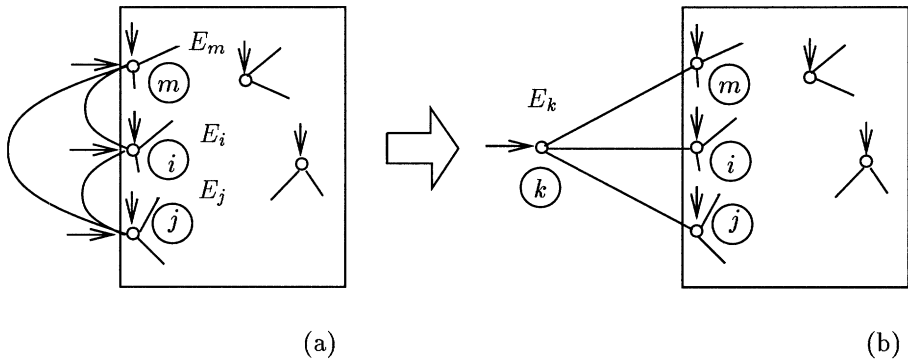
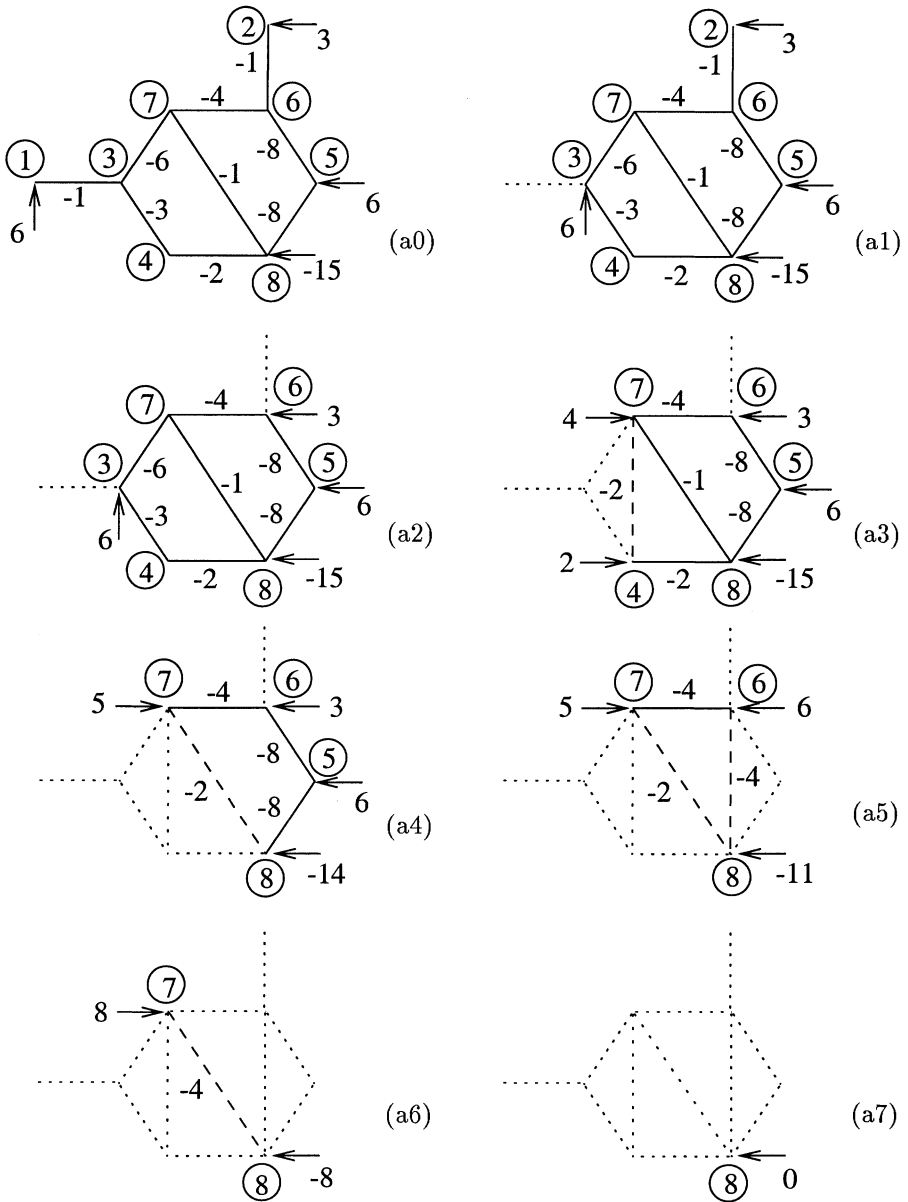


Figure 5.4. Network expansion. Recovery of voltage at Node k . E_k is calculated as a function of the voltages E_m , $m \in \Omega_k$.

Example 5.3:

Consider again the 8-bus network in Example 5.1 (Fig. 5.1). Figure 5.5 shows the reduction process, step-by-step. (The corresponding expansion of this same (reduced) network is shown in Fig. 5.7.) Reduction is carried out according to node numbering. The equivalent admittances and current injections at each step of the reduction process are presented in Fig. 5.6.

Consider, for example, the reduction from (a5) to (a6) in Fig. 5.5. From the data in Fig. 5.6 we have: $Y_{76}^5 = j400$. p.u., $Y_{68}^5 = j400$. p.u., $Y_{78}^5 = j200$. p.u., $Y_{66}^5 = -(Y_{67}^5 + Y_{68}^5) = -j800$. p.u., $I_6^5 = 6$. p.u., $I_7^5 = 5$. p.u., and $I_8^5 = -11$. p.u. The elimination of Node 6 modifies the admittance of Branch 7 – 8, according to Eq. (5.4), and the value of the current injection I_7 , given by Eq. (5.5), i.e.:



Key: ——— Branch of the original network
 Branch eliminated by the reduction.
 - - - - - Branch created by the reduction.

Figure 5.5. Network reduction. During reduction nodal equivalent injections and branch equivalent admittances are computed. Currents in p.u. (Susceptances in 10^2 p.u.)

		Branch susceptance											Current injection								
		13	26	34	37	47	48	56	58	67	68	78	1	2	3	4	5	6	7	8	
Reduction step	0	-1	-1	-3	-6	0	-2	-8	-8	-4	0	-1	6	3	0	0	6	0	0	-15	
	1		-1	-3	-6	0	-2	-8	-8	-4	0	-1		3	6	0	6	0	0	-15	
	2			-3	-6	0	-2	-8	-8	-4	0	-1			6	0	6	3	0	-15	
	3					-2	-2	-8	-8	-4	0	-1				2	6	3	4	-15	
	4							-8	-8	-4	0	-2					6	3	5	-14	
	5									-4	-4	-2						6	5	-11	
	6												-4							8	-8
	7																				0

Figure 5.6. Evolution of branch susceptance (in 10^2 p.u.) and node current injection (in p.u.) throughout reduction (a0 through a7 in Fig 5.5).

$$Y_{78}^6 = Y_{78}^5 - \frac{Y_{76}^5 Y_{68}^5}{Y_{66}^5} = (j200.) - \frac{(j400.)(j400.)}{(-j800.)} = j400. \text{ p.u.}$$

$$I_7^6 = I_7^5 - \frac{Y_{67}^5}{Y_{66}^5} I_6^5 = 8. \text{ p.u.}$$

Similarly, $I_8^6 = -8. \text{ p.u.}$ The updated values of Y_{78} , I_7 , and I_8 are entered in the 6th row of Fig. 5.6.

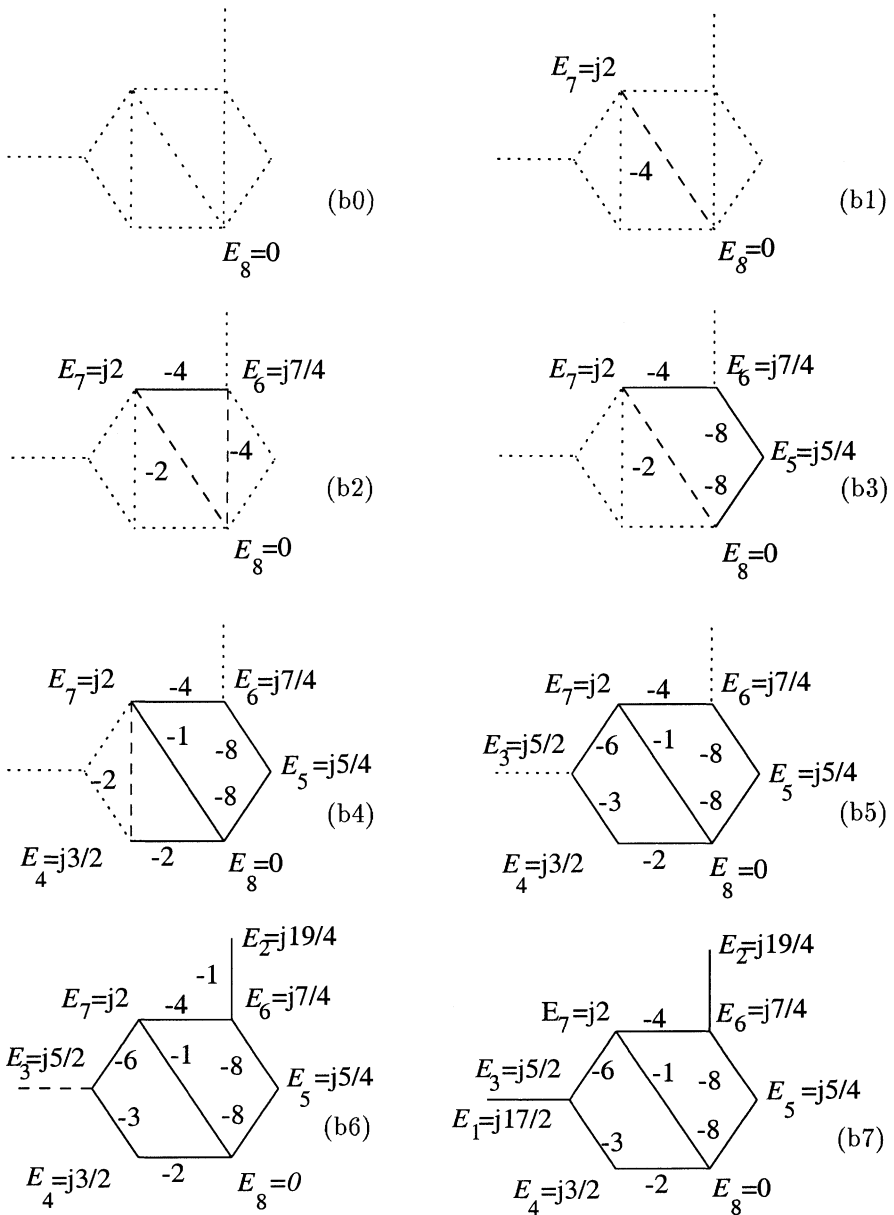
The results of the expansion phase are summarized in Fig 5.8. In this figure, the first row corresponds to a trivial network formed by a single node, i.e., the reference node with voltage $E_8 = 0$. The second row is obtained by applying Eq. (5.2) to the network (b1) of Fig. 5.7. The reduction phase (see Fig. 5.6) has given: $I_7^6 = 8. \text{ p.u.}$, $Y_{78}^6 = j400 \text{ p.u.}$, $Y_{77}^6 = -Y_{78}^6 = -j400 \text{ p.u.}$ Then, applying Eq. (5.2) gives the voltage E_7 of the expanded network:

$$E_7 = \frac{(I_7^6 - Y_{78}^6 E_8)}{Y_{77}^6} = j0.02 \text{ p.u.}$$

All the nodal voltages are obtained the same way (i.e. by applying Eq. (5.2)) as one proceed backwards from Node 8 to Node 1 (see Fig. 5.7).

Example 5.4:

Consider Example 5.3 from which the solution is already known. The change produced in nodal voltage E_6 when nodal current injection I_4 is changed can



Key: ——— Branch of the original network
 Branch eliminated by the reduction.
 - - - - - Branch created by the reduction.

Figure 5.7. Network expansion. During expansion, voltages of previously eliminated nodes are computed in term of the voltages of the reduced network. (Voltages in 10^{-2} p.u.; susceptances in 10^2 p.u.)

Nodal voltage

		1	2	3	4	5	6	7	8
Expansion step	0	-	-	-	-	-	-	-	0
	1	-	-	-	-	-	-	j2	0
	2	-	-	-	-	-	j7/4	j2	0
	3	-	-	-	-	j5/4	j7/4	j2	0
	4	-	-	-	j3/2	j5/4	j7/4	j2	0
	5	-	-	j5/2	j3/2	j5/4	j7/4	j2	0
	6	-	j19/4	j5/2	j3/2	j5/4	j7/4	j2	0
	7	j17/2	j19/4	j5/2	j3/2	j5/4	j7/4	j2	0

Figure 5.8. Evolution of nodal voltages (in 10^{-2} p.u.) during expansion process.

also be determined. Figure 5.9 presents the relevant parts of Figs. 5.5 and 5.7 needed to compute ΔE_6 in terms of ΔI_4 .

Since a linear model is being used, all the current injections are considered equal to zero, except ΔI_4 and ΔI_8 which will undergo change (Node 8 is the reference/slack node.) Consider initially the network in Fig. 5.9 (a3). Figure 5.5 has given: $Y_{47}^3 = j200.$, $Y_{48}^3 = j200.$, $Y_{44}^3 = -(Y_{47}^3 + Y_{48}^3) = -j400.$ and $Y_{78}^3 = j200.$ Now, from Eq. (5.5), the division of current ΔI_4 between Nodes 7 and 8 when Node 4 is eliminated is computed (see Fig. 5.9 (a3)):

$$\Delta I_7 = 0. - \frac{Y_{74}^3}{Y_{44}^3} \Delta I_4 = 0. + \frac{j200.}{j400.} 2.0 = 1.0 \text{ p.u.}$$

This value of ΔI_7 will remain the same when Nodes 5 and 6 are eliminated, since these nodes have zero incremental injections (See the network in Fig. 5.9 (a6)). The effect of this ΔI_7 on ΔE_7 can be determined by consideration of the network of 5.9 (b1). The voltage ΔE_7 is given by Eq. (5.2)

$$\Delta E_7 = \frac{\Delta I_7 - Y_{78}^6 E_8}{Y_{77}^6} = \frac{1.0 - (j400.)0.}{-j400.} = j0.0025 \text{ p.u.}$$

where $Y_{78}^6 = j400.$ and $Y_{77}^6 = -Y_{78}^6 = -j400.$ (These values are given by the network in Fig. 5.9 (b1)). Finally, consideration of the network in Fig. 5.9 (b2) leads to the determination of voltage ΔE_6 . Figure 5.5 has given: $Y_{67}^5 = j400.$, $Y_{68}^5 = j400.$, $Y_{66}^5 = -(Y_{67}^5 + Y_{68}^5) = -j800.$, and $\Delta I_6 = 0.$ Voltage ΔE_6 is then obtained using Eq. (5.2):

$$\Delta E_6 = \frac{\Delta I_6 - (Y_{67}^5 E_7 + Y_{66}^5 E_7)}{Y_{66}^5} = \frac{0. - (j400. j0.0025)}{-j800.} = j0.00125 \text{ p.u.}$$

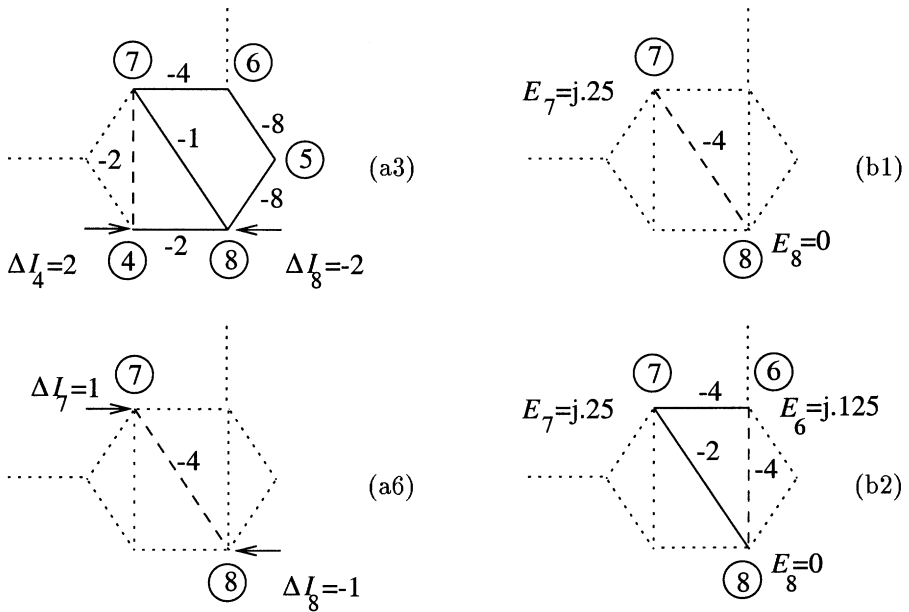


Figure 5.9. Partial networks used to compute the effect of ΔI_4 on ΔE_6 . (Susceptances in 10^2 p.u.; voltages in 10^{-2} p.u.)

		Current injection							
		1	2	3	4	5	6	7	8
Reduction step	0	-	-	-	-	-	-	-	-
	1	-	-	-	-	-	-	-	-
	2	-	-	-	-	-	-	-	-
	3	-	-	-	2	-	-	-	-2
	4	-	-	-	-	-	-	-	-
	5	-	-	-	-	-	-	-	-
	6	-	-	-	-	-	-	1	-1
	7	-	-	-	-	-	-	-	0

		Nodal voltage							
		1	2	3	4	5	6	7	8
expansion step	0	-	-	-	-	-	-	-	0
	1	-	-	-	-	-	-	$j/4$	0
	2	-	-	-	-	-	$j/8$	$j/4$	0
	3	-	-	-	-	-	-	-	-
	4	-	-	-	-	-	-	-	-
	5	-	-	-	-	-	-	-	-
	6	-	-	-	-	-	-	-	-
	7	-	-	-	-	-	-	-	-

Figure 5.10. Changes in current injection and on nodal voltage produced by $\Delta I_4 = 2$ p.u. (Currents in p.u.; voltages in 10^{-2} p.u.) All the remaining data are the same as in Fig 5.3.

The results are summarized in Fig. 5.10. Part (a) of the figure shows how the current increment ΔI_4 is distributed to other network nodes, with only Nodes 7 and 8 being affected in this example. Part (b) of the figure shows the steps needed to compute ΔE_6 . In this specific example, visualization of the relevant parts of the network reduction/expansion process used to perform the desired calculations was easy, but the concept of paths is generally necessary and is described below.

5.2.3 Paths

The previous example illustrates the important fact that, once a initial solution is known, solution updating when nodal injection is changed, normally requires the consideration of only parts of the original network.

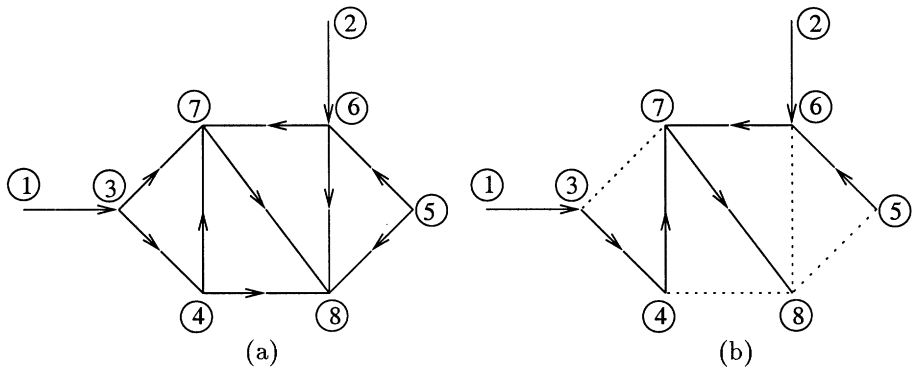


Figure 5.11. (a) Nodal injections distribution during the network reduction. (b) Paths connecting each eliminated node to the downstream nodes that are affected by the current injection of an eliminated node.

For example, in the case of the 8-bus network analyzed above, the affected parts of the network are shown in Figure 5.11 (a). This figure shows how nodal current injection propagates throughout the network as reduction proceeds. For example, the injection I_1 is initially transferred to Node 3; it is then distributed between Nodes 4 and 7 by a current divider, and so on.

When a generic Node k is eliminated, its current is distributed among adjacent nodes, i.e., among the nodes $m \in \Omega_k$. These nodes will eventually be eliminated. To keep track of the impact of the current I_k , however, it is only necessary to know which node of this neighborhood was eliminated first. Figure 5.11 (b) shows the paths associated with this example; each segment of these paths point to the first node eliminate in a neighborhood. These paths form a spanning tree rooted at the last node to be eliminated (the reference node in this case), with a single path connecting the root to any other node

of the network. Such paths are the basis for developing efficient sparse vector techniques.

Consider, for example, a change ΔI_4 , as in Example 5.4 (a). Path $4 \rightarrow 7 \rightarrow 8$, in Fig. 5.11 (b), shows the nodes affected by ΔI_4 when reduction is carried out; Nodes 1, 2, 3, 5, and 6 are not affected. On the other hand, path $6 \rightarrow 7 \rightarrow 8$, shows that the voltages ΔE_8 and ΔE_7 are needed to compute ΔE_6 ; solution involves determination of ΔE_7 from $\Delta E_8 = 0$, and then computation of ΔE_6 from ΔE_7 and ΔE_8 , as in Example 5.4.

5.3 LDU DECOMPOSITION

The properties discussed above are formalized below in terms of the LDU factorization of the nodal admittance matrix. Network reduction and expansion are then associated with forward and back substitutions performed by the Gauss elimination method.

5.3.1 Gauss Elimination

Consider a system $\mathbf{Ax} = \mathbf{b}$ of n linear equations with n unknowns. Assume that \mathbf{A} is nonsingular. The system is solved by applying a sequence of Gauss eliminations that transform $\mathbf{Ax} = \mathbf{b}$ into a system $\mathbf{Ux} = \mathbf{y}$. The systems $\mathbf{Ax} = \mathbf{b}$ and $\mathbf{Ux} = \mathbf{y}$ are equivalent since they have the same solution \mathbf{x} . The appropriate sequence of Gauss eliminations successively yields

$$\begin{pmatrix} A_{11} & A_{12} & A_{13} & \dots & A_{1n} \\ A_{21} & A_{22} & A_{23} & \dots & A_{2n} \\ A_{31} & A_{32} & A_{33} & \dots & A_{3n} \\ \dots & \dots & \dots & \dots & \dots \\ A_{n1} & A_{n2} & A_{n3} & \dots & A_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \cdot \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ \cdot \\ b_n \end{pmatrix} \quad (5.6)$$

$$\begin{pmatrix} A'_{11} & A'_{12} & A'_{13} & \dots & A'_{1n} \\ 0 & A_{22} & A_{23} & \dots & A_{2n} \\ 0 & 0 & A_{33} & \dots & A_{3n} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & A'_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \cdot \\ x_n \end{pmatrix} = \begin{pmatrix} b'_1 \\ b'_2 \\ b'_3 \\ \cdot \\ b'_n \end{pmatrix} \quad (5.7)$$

$$\begin{pmatrix} 1 & U_{12} & U_{13} & \dots & U_{1n} \\ 0 & 1 & U_{23} & \dots & U_{2n} \\ 0 & 0 & 1 & \dots & U_{3n} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \cdot \\ x_n \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ \cdot \\ y_n \end{pmatrix} \quad (5.8)$$

The transformed system of Eq. (5.8) can then be solved for \mathbf{x} by backward substitution.

Example 5.5:

Consider the 8-bus network in Fig. 5.1. Figure 5.12 gives the column-wise sequence of Gauss eliminations used to triangularize the corresponding bus susceptance matrix (all the values are given in 10^2 p.u.). Note the one-to-one correspondence between the steps used to triangularize the coefficient matrix and the network reduction steps as shown in Fig. 5.5.

5.3.2 Elementary Matrices

The transformation of a system $\mathbf{Ax} = \mathbf{b}$ into an equivalent system $\mathbf{Ux} = \mathbf{y}$ by Gauss elimination requires a series of elementary operations. These operations can be seen as premultiplications of the coefficient matrix \mathbf{A} , or the augmented matrix $[\mathbf{A}|\mathbf{b}]$, by elementary matrices of the following types

$$\mathbf{T}_{ij} = \begin{array}{c} \begin{array}{|c|} \hline j \\ \hline \end{array} \begin{array}{|c|} \hline 1 \\ \hline \end{array} \begin{array}{|c|} \hline \vdots \\ \hline \end{array} \begin{array}{|c|} \hline c_{ij} \\ \hline \end{array} \begin{array}{|c|} \hline \vdots \\ \hline \end{array} \begin{array}{|c|} \hline 1 \\ \hline \end{array} \\ \hline \end{array} ; \quad \mathbf{T}_{ii} = \begin{array}{c} \begin{array}{|c|} \hline i \\ \hline \end{array} \begin{array}{|c|} \hline 1 \\ \hline \end{array} \begin{array}{|c|} \hline \vdots \\ \hline \end{array} \begin{array}{|c|} \hline c_{ii} \\ \hline \end{array} \begin{array}{|c|} \hline \vdots \\ \hline \end{array} \begin{array}{|c|} \hline 1 \\ \hline \end{array} \\ \hline \end{array} \quad (5.9)$$

If the (i, j) th element of the transformed matrix \mathbf{A} is to become zero at some intermediate stage of the elimination process, \mathbf{A} can be premultiplied by the elementary matrix \mathbf{T}_{ij} with $c_{ij} = -A_{ij}/A_{jj}$, where A_{ij} and A_{jj} are the current values of elements (i, j) and (j, j) , respectively. When \mathbf{T}_{ij} , $i < j$, is applied to the set of linear equations given in Eqs. (5.6)-(5.8), it has the effect of replacing the j th row with a new row given by the linear combination of rows i and j : row i is multiplied by c_{ij} and row j is multiplied by one. Note that these are the same elementary operations needed to transform a unit matrix into the elementary matrix \mathbf{T}_{ij} .

Similarly, to transform diagonal element A_{ii} into a unit, matrix \mathbf{A} is premultiplied by the elementary matrix \mathbf{T}_{ii} , with $c_{ii} = 1/A_{ii}$.

The transformation of Eq. (5.6) into Eq. (5.8) by Gauss elimination can be described by the following sequence of elementary transformations:

$$\mathbf{T}^q(\mathbf{T}^{q-1}(\dots(\mathbf{T}^3(\mathbf{T}^2(\mathbf{T}^1\mathbf{A})))) = \mathbf{I}_n \quad (5.10)$$

$$\mathbf{T}^q(\mathbf{T}^{q-1}(\dots(\mathbf{T}^3(\mathbf{T}^2(\mathbf{T}^1\mathbf{b})))) = \mathbf{x} \quad (5.11)$$

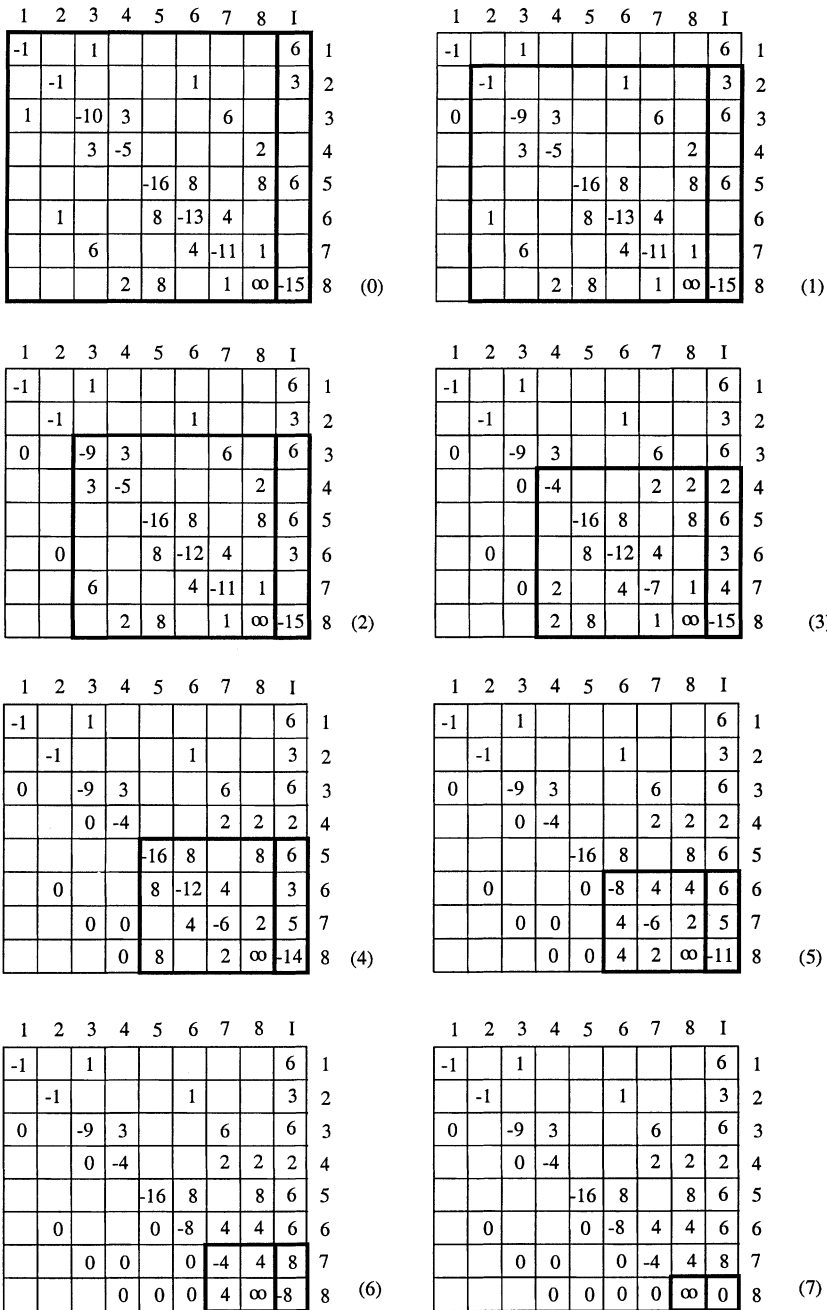


Figure 5.12. (a) 8-bus network in Fig. 5.1. Reduction applied to the susceptance matrix augmented with the current vector ($[A|b]$).

Note that the same operations are applied to the coefficient matrix **A** and to the independent vector **b**.

Example 5.6:

Consider the 8-bus network in Example 5.5. Figure 5.13 gives three examples of elementary matrices used to transform the susceptance matrix of the network given in Fig. 5.1 into an upper triangular matrix (The triangularization process is summarized in Fig. 5.12.)

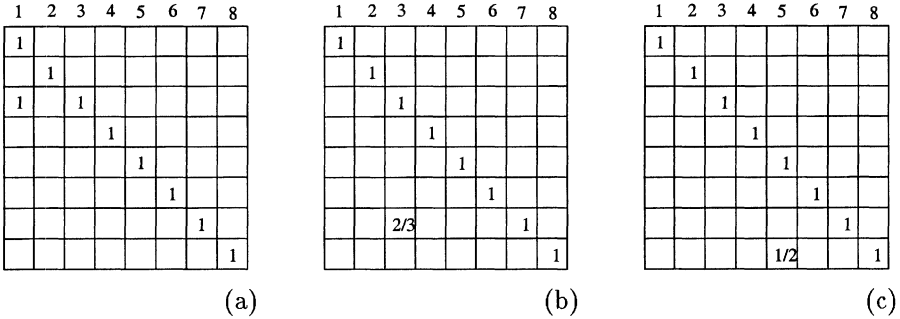


Figure 5.13. Examples of elementary matrices used in Fig. 5.12. (a) Matrix used to zeroize element (3,1). (b) Matrix used to zeroize element (7,3). (c) Matrix used to zeroize element (8,5).

5.3.3 Triangular Factorization

Equations (5.10) and (5.11) can be used to write the inverse matrix **A**⁻¹ as

$$\mathbf{A}^{-1} = \mathbf{T}^q \mathbf{T}^{q-1} \dots \mathbf{T}^3 \mathbf{T}^2 \mathbf{T}^1 \tag{5.12}$$

Now, calling $[\mathbf{T}_k]^{-1} = \mathbf{T}_I^k$, makes it possible to express the matrix **A** as

$$\mathbf{A} = \mathbf{T}_I^1 \mathbf{T}_I^2 \mathbf{T}_I^3 \dots \mathbf{T}_I^{q-1} \mathbf{T}_I^q \tag{5.13}$$

where the inverses of the elementary matrices **T**_{ij} and **T**_{ii} are given by

$$\mathbf{T}_{ij}^{-1} = \begin{array}{|c|} \hline \begin{array}{c} j \\ 1 \\ 1 \\ \vdots \\ -c_{ij} \\ \vdots \\ 1 \end{array} \\ \hline \end{array} ; \mathbf{T}_{ii}^{-1} = \begin{array}{|c|} \hline \begin{array}{c} i \\ 1 \\ 1 \\ \vdots \\ c_{ii}^{-1} \\ \vdots \\ 1 \end{array} \\ \hline \end{array} \quad (5.14)$$

Thus, inverting the elementary matrix \mathbf{T}_{ij} amounts to changing the sign of the (i, j) th element of \mathbf{T}_{ij} (i.e., changing from c_{ij} to $-c_{ij}$). The inverse of the elementary matrix \mathbf{T}_{ii} is obtained by inverting the diagonal element c_{ii} .

If the elementary operations in Eqs. (5.10) and (5.11) are performed, first, to zeroize the elements of the lower (left) triangle, second, to normalize the elements of the main diagonal, and third, to zeroize the elements of the upper (right) triangle, then Eq. (5.13) can be rewritten as:

$$\mathbf{A} = [\mathbf{L}^1 \mathbf{L}^2 \dots \mathbf{L}^{n_1}] [\mathbf{D}^1 \mathbf{D}^2 \dots \mathbf{D}^n] [\mathbf{U}^1 \mathbf{U}^2 \dots \mathbf{U}^{n_u}] \quad (5.15)$$

where \mathbf{L}^i , \mathbf{D}^i , and \mathbf{U}^i are applied to the lower triangle, the main diagonal, and the upper triangle, respectively.

\mathbf{L}_{km} and \mathbf{L}_{ij} can be considered to be two elementary matrices corresponding to operations performed on the lower triangle; moreover the element (k, m) is zeroized before the element (i, j) , i.e., the elementary matrix \mathbf{L}_{km} is used before \mathbf{L}_{ij} is. In this case

$$\mathbf{L}_{km} \mathbf{L}_{ij} = \mathbf{L}_{km} + \mathbf{L}_{ij} - \mathbf{I}_n = \begin{array}{|c|} \hline \begin{array}{c} m \quad j \\ 1 \\ 1 \\ \vdots \\ -c_{km} \\ \vdots \\ -c_{ij} \\ \vdots \\ 1 \end{array} \\ \hline \end{array} \quad (5.16)$$

This means that the product $\mathbf{L}_{km} \mathbf{L}_{ij}$ is the result of the superposition of the elementary matrices \mathbf{L}_{km} and \mathbf{L}_{ij} . This property is valid for both the column-wise elimination and the row-wise elimination. Note, however, that it does not

necessarily hold if \mathbf{L}_{km} and \mathbf{L}_{ij} are multiplied in the reverse order. This simple property has a fundamental impact on sparse matrix techniques, as will be seen below.

This property is extended to all the elementary \mathbf{L} matrices as follows

$$\mathbf{L}^1 \mathbf{L}^2 \dots \mathbf{L}^{n_l} = \mathbf{I}_n + \sum_{p=1}^{n_l} (\mathbf{L}^p - \mathbf{I}_n) = \mathbf{L} \tag{5.17}$$

The product of all the \mathbf{L}^p matrices (the elementary matrices used to zeroize the lower triangle), is also a lower triangle matrix obtained by the superposition of the elementary matrices \mathbf{L}^p . Similar results are obtained for the elementary matrices \mathbf{U}^p :

$$\mathbf{U}^1 \mathbf{U}^2 \dots \mathbf{U}^{n_u} = \mathbf{I}_n + \sum_{p=1}^{n_u} (\mathbf{U}^p - \mathbf{I}_n) = \mathbf{U} \tag{5.18}$$

where \mathbf{U} is an upper triangular matrix. The elementary matrices \mathbf{D}^p can be written as follows

$$\mathbf{D}^1 \mathbf{D}^2 \dots \mathbf{D}^n = \mathbf{D} \tag{5.19}$$

Note that the matrices \mathbf{L} and \mathbf{U} partially retain the sparsity of the original coefficient matrix \mathbf{A} . Non-zero elements occur both in positions in which \mathbf{A} has non-zero elements and in positions previously filled with zeros (the so called fill-in elements). In terms of network reduction, fill-ins correspond to branches that are created during the reduction process. For example, Branch 4 – 7 created by the elimination of Node 3 corresponds to the fill-in elements (4, 7) and (7, 4) shown in Fig. 5.16 (b).

The sparsity of \mathbf{L} and \mathbf{U} is based on the superposition property illustrated by Eq. (5.16). However, there is no guarantee that the inverses \mathbf{L}^{-1} and \mathbf{U}^{-1} will retain sparsity; they will generally be less sparse than than \mathbf{L} and \mathbf{U} , and may even be full, because the superposition property does not generally holds for \mathbf{L}^{-1} and \mathbf{U}^{-1} , since $\mathbf{L}^{-1} = \mathbf{L}_I^{n_l} \mathbf{L}_I^{n_l-1} \dots \mathbf{L}_I^1$ and $\mathbf{U}^{-1} = \mathbf{U}_I^{n_u} \mathbf{U}_I^{n_u-1} \dots \mathbf{U}_I^1$, i.e., the product of elementary matrices appear in the reverse order.

Finally, introducing Eqs. (5.17), (5.18) and (5.19) into (5.15), yields the factorized form of \mathbf{A} :

$$\mathbf{A} = \mathbf{L} \mathbf{D} \mathbf{U}$$

If \mathbf{A} is symmetric, then $\mathbf{L} = \mathbf{U}^t$. In this case all the numerical operations used to obtain \mathbf{L} already contain the data necessary to obtain both \mathbf{D} and \mathbf{U} .

Example 5.7:

Consider the 8-bus network in Examples 5.1 through 5.6. Figure 5.14 gives the factors \mathbf{L} , \mathbf{D} , and \mathbf{U} of the network susceptance matrix (susceptances in 10^2 p.u.).

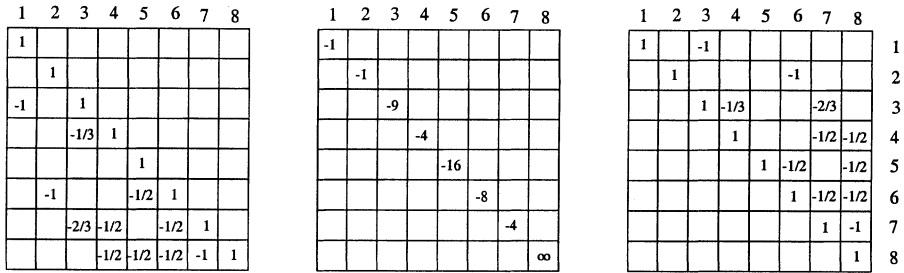


Figure 5.14. L, D, and U factors of the susceptance matrix B (see Fig. 5.1).

For example, the element (4, 3) of the matrix L corresponds to the elementary matrix used to zeroize the element (4, 3) of the coefficient matrix. This elementary operation, along with elementary operation L_{7,4}, corresponds to the elimination of Node 4 of the 8-bus network.

5.4 USING LDU FACTORS TO SOLVE LINEAR SYSTEMS

Consider that the coefficient matrix A of a linear system of equations Ax = b can be decomposed as A = LDU. Then Ax = b can be written successively as

$$\begin{aligned}
 \mathbf{Ax} &= \mathbf{LDUx} = \mathbf{b} \\
 \mathbf{L}^{-1}\mathbf{Ax} &= \mathbf{DUx} = \mathbf{L}^{-1}\mathbf{b} = \mathbf{b}' \\
 \mathbf{D}^{-1}\mathbf{L}^{-1}\mathbf{Ax} &= \mathbf{Ux} = \mathbf{D}^{-1}\mathbf{L}^{-1}\mathbf{b} = \mathbf{b}'' \\
 \mathbf{U}^{-1}\mathbf{D}^{-1}\mathbf{L}^{-1}\mathbf{Ax} &= \mathbf{x} = \mathbf{U}^{-1}\mathbf{D}^{-1}\mathbf{L}^{-1}\mathbf{b} = \mathbf{b}''' \quad (5.20)
 \end{aligned}$$

Then, Eq. (5.20) gives

$$\mathbf{x} = \mathbf{U}^{-1}(\mathbf{D}^{-1}(\mathbf{L}^{-1}\mathbf{b}))$$

which is equivalent to solving the following three problems in sequence

$$\begin{aligned}
 \mathbf{b}' &= \mathbf{L}^{-1}\mathbf{b} \\
 \mathbf{b}'' &= \mathbf{D}^{-1}\mathbf{b}' \\
 \mathbf{x} &= \mathbf{b}''' = \mathbf{U}^{-1}\mathbf{b}''
 \end{aligned}$$

As mentioned above, the matrices L⁻¹ and U⁻¹ are generally less sparse than matrices L and U. Thus, L⁻¹ and U⁻¹ are not computed explicitly. Instead, they are normally retained in the product form:

$$\begin{aligned} \mathbf{b}' &= (\mathbf{L}_I^{n_l} (\mathbf{L}_I^{n_l-1} \dots (\mathbf{L}_I^2 (\mathbf{L}_I^1 \mathbf{b}))) \dots) \\ \mathbf{b}'' &= (\mathbf{D}_I^n (\mathbf{D}_I^{n-1} \dots (\mathbf{D}_I^2 (\mathbf{D}_I^1 \mathbf{b}')) \dots)) \\ \mathbf{x} &= (\mathbf{U}_I^{n_u} (\mathbf{U}_I^{n_u-1} \dots (\mathbf{U}_I^2 (\mathbf{U}_I^1 \mathbf{b}'')) \dots)) \end{aligned}$$

where $\mathbf{L}_I^p = (\mathbf{L}^p)^{-1}$, $\mathbf{D}_I^p = (\mathbf{D}^p)^{-1}$ and $\mathbf{U}_I^p = (\mathbf{U}^p)^{-1}$ are given by Eq. (5.14), and $\mathbf{U}_I^p = (\mathbf{U}^p)^{-1}$ is given by an expression similar to the one used for \mathbf{L}^p . This means that all the information needed to transform the independent vector \mathbf{b} into the solution vector \mathbf{x} , is contained in factor matrices \mathbf{L} , \mathbf{D} and \mathbf{U} .

The elementary operations used for solving $\mathbf{Ax} = \mathbf{b}$ are summarized below:

- Matrix \mathbf{L} (Computation of \mathbf{b}'):

Add to the current content of the i th position of \mathbf{b} , the product of the current content of the j th position by $-l_{ij}$, i.e., $(b_i \leftarrow b_i - b_j l_{ij})$. The result is denoted by \mathbf{b}' .

- Matrix \mathbf{D} (Computation of \mathbf{b}''):

Divide the current contents of each position of vector \mathbf{b}' by d_{ii} . The result is denoted by \mathbf{b}'' , its components represented by $b_i'' \leftarrow b_i' / d_{ii}$.

- Matrix \mathbf{U} (Computation of \mathbf{b}''')

Add to the current content of the i th position of \mathbf{b}'' , the product of the current content of the j th position by $-u_{ij}$, i.e., $(b_i'' \leftarrow b_i'' - b_j'' u_{ij})$. The solution vector is \mathbf{x} .

Example 5.8:

Figure 5.15 gives all the elementary operations used for solving the system $\mathbf{YE} = \mathbf{I}$, corresponding to the 8-bus network in Example 5.5, using the \mathbf{LDU} factors.

5.5 PATH FINDING

The discussion of the role of paths in terms of network reduction presented in the first part of this chapter shows how injections of eliminated nodes affect nodes located downstream in the elimination process. The concept of paths is now presented in terms of sparsity pattern of the triangular factors of the nodal admittance matrix.

When the pivot (k, k) is processed, elements (i, j) of the coefficient matrix, with $i, j \in \Omega_k$, are affected; this is also true for the elements $i \in \Omega_k$ of the independent vector. Eventually the pivots (i, i) , with $i \in \Omega_k$, will be processed as well. Thus it is sufficient to know a single node of Ω_k (the pivot processed

		I										E																						
		0	1	2	3	4	5	6	7	8	9	10	11	1	2	3	4	5	6	7	8	9	10	11										
1		6	6	6	6	6	6	6	6	6	6	6	6	-6	-6	-6	-6	-6	-6	-6	-6	-6	-6	-6	-17/2									
2		3	3	3	3	3	3	3	3	3	3	3	3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-19/4	-19/4								
3		0	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6								
4		0	0	0	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2								
5		6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6								
6		0	0	3	3	3	3	3	3	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6								
7		0	0	0	0	4	5	5	5	5	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8								
8		-15	-15	-15	-15	-15	-14	-14	-11	-11	-8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0								
		3,1	6,2	4,3	7,3	7,4	8,4	6,5	8,5	7,6	8,6	8,7	1	2	3	4	5	6	7	8	7,8	6,8	6,7	5,8	5,6	4,8	4,7	3,7	3,4	2,6	1,3			
		1	1	1/3	2/3	1/2	1/2	1/2	1/2	1/2	1/2	1	-1	-1	-1/9	-1/4	1/16	-1/8	-1/4	∞	1	1/2	1/2	1/2	1/2	1/2	1/2	1/2	2/3	1/3	1	1		
		L_{ij}											D_i											U_{ji}										

Figure 5.15. (a) 8-bus network. Solution of $\mathbf{YE} = \mathbf{I}$ using \mathbf{LDU} factors.

first) to keep track of the impact of the (k, k) th pivoting on the remaining pivots.

Example 5.9:

Figure 5.16 shows (a) the 8-node network used in the previous examples, including the branches created during the network reduction process (Branches 4 – 7 and 6 – 8); (b) the sparsity structure of matrix $\mathbf{A} = \mathbf{LDU}$, including the fill-in elements (the elements (4, 7), (7, 4), (6, 8), and (8, 6)); (c) paths; (d) determination of paths from the sparsity pattern of the \mathbf{L} matrix.

Example 5.10:

Consider that the solution vector \mathbf{E} obtained in Examples 5.5 and 5.9 is known. The change produced in nodal voltage E_6 when the nodal current injection at Node 4 is changed by $\Delta I_4 = 2$, can be determined considering the paths given in Fig. 5.16. The calculations are summarized in Fig. 5.17.

5.6 PIVOT ORDERING TO PRESERVE SPARSITY

The sparsity of the triangular factors \mathbf{L} and \mathbf{U} of the nodal admittance matrix \mathbf{Y} , and of other matrices with a similar structure, depends on the order in which node elimination is performed. This section describes ordering schemes used to preserve sparsity. When a node k with a neighborhood $m \in \Omega_k$ is eliminated, the admittances of Branches $i - j$, with $i, j \in \Omega_k$, are affected; for those branches with zero admittances, a new admittance is introduced (fill-in element). Thus, the chance that a new branch is created by a pivot operation depends basically of the number of nodes in the neighborhood of the eliminated

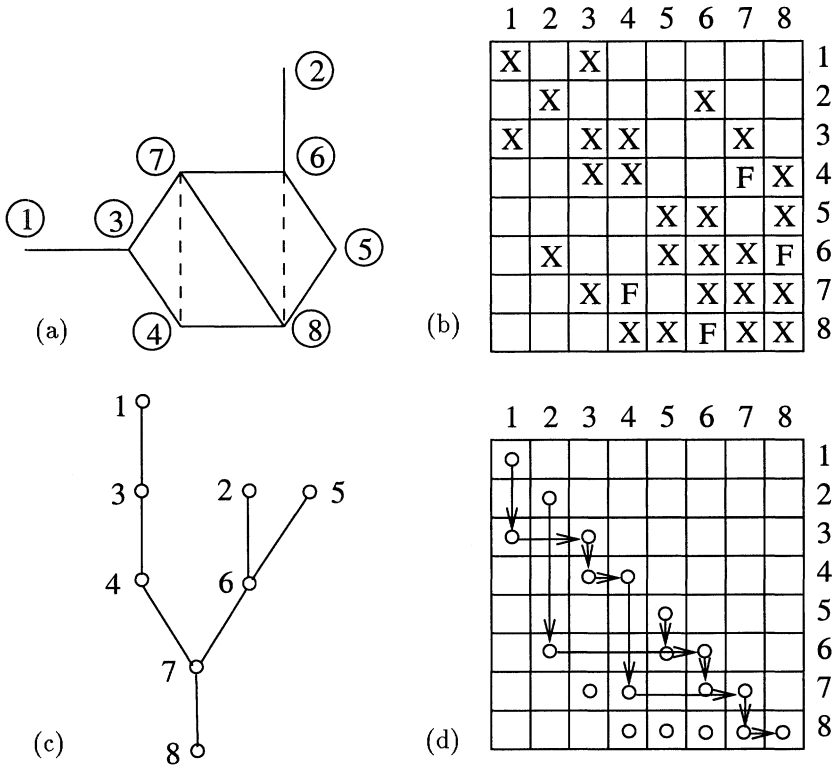


Figure 5.16. (a) 8-bus network. (b) Y matrix sparsity pattern including fill-ins. (c) Paths. (d) Paths indicated on L matrix structure.

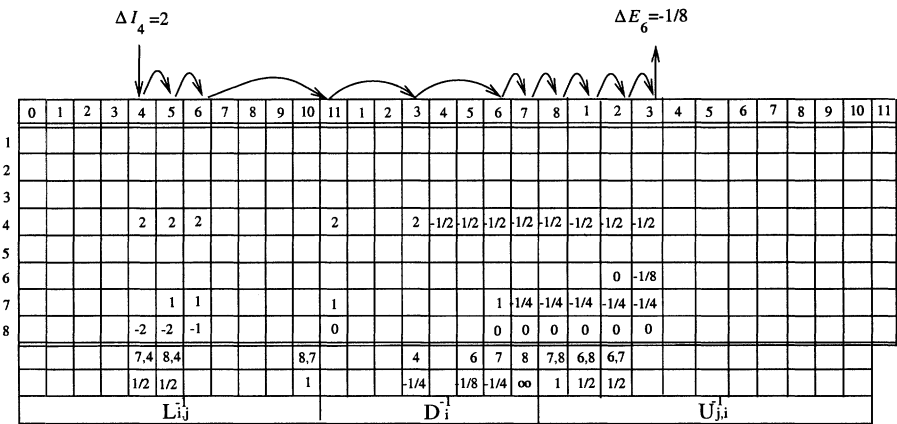


Figure 5.17. (a) Example 5.10. Calculation of the correction ΔE_6 due to change ΔI_4 using LDU factors (according to the paths $4 \rightarrow 7 \rightarrow 8$ and $6 \rightarrow 7 \rightarrow 8$).

node and the strategies normally used to preserve sparsity are related to node degree (or valence), i.e., the number of adjacent branches connected to a node. The three classic ordering schemes (Tinney's) are summarized in the following:

- Scheme-1

Node degrees are determined for the unreduced (original) network, or matrix. Nodes with lower degrees are eliminated first. Tie breaking is arbitrary. This scheme is simple though not very effective. It can be used as an initial order to be improved by other schemes.

- Scheme-2

Node degrees are determined for the current reduce network, or matrix. Node with minimum degree is selected for elimination. Tie breaking is arbitrary. This scheme has been widely used in practice.

- Scheme-3

The next node to be eliminated is the one that will produce the minimum number of fill-ins in the reduced network, or matrix. Tie breaking is arbitrary. This scheme does not necessarily yields better results than Scheme-2.

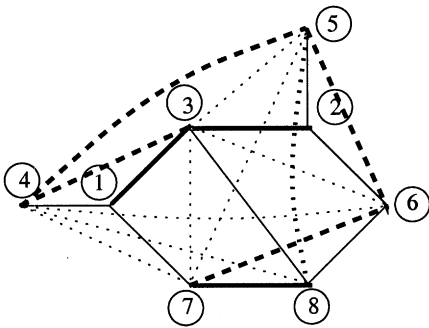
Example 5.11:

The effect of node ordering on the 8-bus network and on the corresponding \mathbf{Y} -matrix is illustrated in Figs. 5.18 through 5.20. The elements already present in the original network are represented in the structure of the \mathbf{Y} -matrix by X's; the new elements (fill-in elements) are indicated by F's. The original branches of the network are represented by solid lines and the fill-in elements are represented by dotted lines. The branches that form part of paths are represented by darker lines.

Fig. 5.18 shows a case of arbitrary pivot ordering: the effect on sparsity is dramatic. There are 24 fill-ins (12 added branches). In this case the longest path has six branches; The average path length is $27/8$. Figure 5.19 shows the ordering yielded by Scheme-1; there are six fill-ins (three new branches); in this case the longest path has only three branches; The average path length is $16/8$. Finally, Fig. 5.20 shows the result obtained with Scheme-2; there are only four fill-ins (two branches). The longest path, however, has four branches; The average path length is $18/8$.

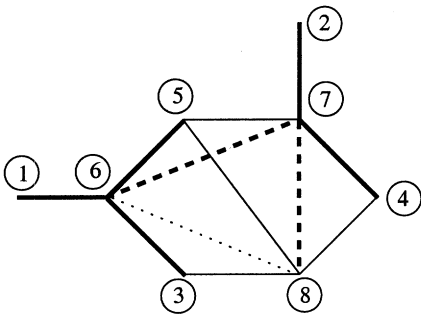
5.7 MDML AND MLMD ORDERING SCHEMES

The examples discussed above (Figs. 5.19 and 5.20) show that the minimum number of fill-ins do not necessarily leads to minimum path lengths. For repeated applications of the sparse triangular factors with sparse vectors, as in contingency analysis and bad data processing, however, it may be desirable to



	1	2	3	4	5	6	7	8	
1	X		X	X			X		1
2		X	X		X	X			2
3	X	X	X	F	F	F	F	X	3
4	X		F	X	F	F	F	F	4
5		X	F	F	X	F	F	F	5
6		X	F	F	F	X	F	X	6
7	X		F	F	F	F	X	X	7
8			X	F	F	X	X	X	8

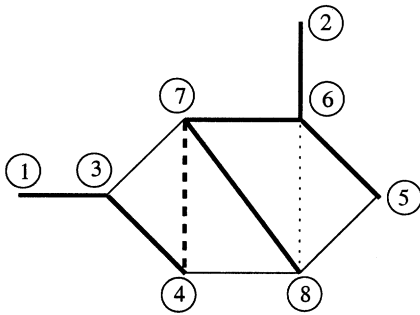
Figure 5.18. Effect of ordering on the sparsity and on path lengths: Arbitrary ordering.



	1	2	3	4	5	6	7	8	
1	X					X			1
2		X					X		2
3			X			X		X	3
4				X			X	X	4
5					X	X	X	X	5
6	X		X		X	X	F	F	6
7		X		X	X	F	X	F	7
8			X	X	X	F	F	X	8

Figure 5.19. The effect of ordering on the sparsity and on path lengths: Ordering according to Scheme-1.

reduce average path lengths. The MDML (Minimum Degree Minimum Level) scheme can be used to this end. This scheme simply introduces a tie breaking rule to Scheme-2; whenever two or more candidate nodes have the same degree, the node to be eliminated first is the one with minimum current path length. Thus, in determining the best node ordering one has to keep track of the of the current path lengths of all the nodes still to be processed. This algorithm is both effective and easy to implement from an existing Scheme-2 code. The



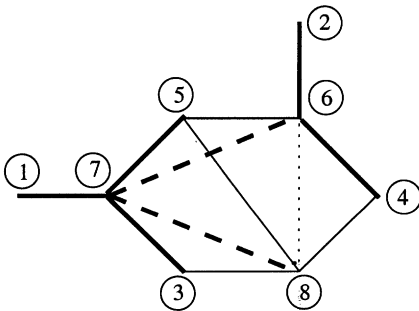
	1	2	3	4	5	6	7	8	
1	X		X						1
2		X				X			2
3	X		X	X			X		3
4			X	X			F	X	4
5					X	X		X	5
6		X			X	X	X	F	6
7			X	F		X	X	X	7
8				X	X	F	X	X	8

Figure 5.20. The effect of ordering on the sparsity and on path lengths: Ordering according to Scheme-2.

MLMD (Minimum Level Minimum Degree) scheme uses path lengths as the main criterion and consider node degrees for tie breaking.

Example 5.12:

Figure 5.21 shows the result obtained with Scheme-MLMD; there are six fill-ins (three branches). The longest path has three branches; The average path length is 15/8.



	1	2	3	4	5	6	7	8	
1	X						X		1
2		X				X			2
3			X				X	X	3
4				X		X		X	4
5					X	X	X	X	5
6		X		X	X	X	F	F	6
7	X		X		X	F	X	F	7
8			X	X	X	F	F	X	8

Figure 5.21. Effect of ordering on the sparsity and on path lengths: Ordering according to Scheme MLMD.

Remarks: For this network, no significant differences in sparsity patterns can be observed, and the examples above are thus considered only as illustrations.

More convincing results can be obtained with larger, more complex systems (see references).

5.8 BLOCKED FORMULATION OF NEWTON POWER FLOW

The Jacobian matrix equation for the Newton Raphson power flow method can be written as follows:

$$\begin{pmatrix} \mathbf{H} & \mathbf{N} \\ \mathbf{M} & \mathbf{L} \end{pmatrix} \begin{pmatrix} \Delta\Theta \\ \Delta\mathbf{V} \end{pmatrix} = \begin{pmatrix} \Delta\mathbf{P} \\ \Delta\mathbf{Q} \end{pmatrix}$$

where

$$\begin{aligned} H_{km} &= \partial P_k / \partial \theta_m & N_{km} &= \partial P_k / \partial V_m \\ M_{km} &= \partial Q_k / \partial \theta_m & L_{km} &= \partial Q_k / \partial V_m \end{aligned}$$

and the elements of $\Delta\mathbf{P}$ are defined for each PV and each PQ bus, the elements of $\Delta\mathbf{Q}$ are defined for each PQ bus, the elements of $\Delta\Theta$ are calculated for each PV and each PQ bus, and the elements of $\Delta\mathbf{V}$ are computed for each PQ bus. Thus the Jacobian matrix equations is an $n \times n$ system where $n = n_{PV} + 2n_{PQ}$, with n_{PV} being the number of PV buses and n_{PQ} the number of PQ buses.

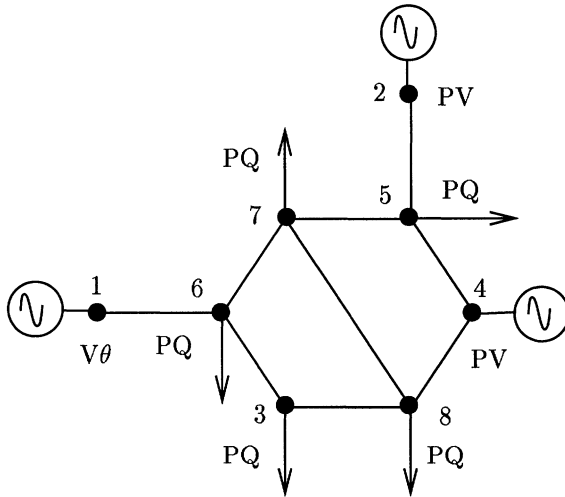


Figure 5.22. 8-bus network.

The Jacobian matrix equation for the network in Fig. 5.22 is given below. For each PV-bus k ($k = 2$ and $k = 4$), ΔQ_k is not defined, and the corresponding Jacobian element L_{kk} is set at ∞ . For the slack bus (bus-1) neither

ΔP_1 nor ΔQ_1 are defined and the elements H_{11} and L_{11} are also set at ∞ . The addition of such large numbers has the effect of eliminating the unwanted equations, but without affecting the structure of the Jacobian matrix.

The sparsity pattern of the Jacobian matrix is the same as that of the \mathbf{Y} -bus matrix since there is a one-to-one correspondence between the complex elements of the \mathbf{Y} -matrix, $\mathbf{Y}_{km} = \mathbf{G}_{km} + j\mathbf{B}_{km}$, and the 2×2 blocks

$$\begin{pmatrix} \partial P_k / \partial \theta_m & \partial P_k / \partial V_m \\ \partial Q_k / \partial \theta_m & \partial Q_k / \partial V_m \end{pmatrix}$$

\mathbf{LDL}' factorization is used to solve the Jacobian matrix equation as is done with the \mathbf{Y} -matrix, except that for the former all operations are performed using block-arithmetic. This includes the processing of 1×1 , 1×2 , 2×1 , and 2×2 blocks, as illustrated by the Jacobian matrix for the system in Fig. 5.22.

The Jacobian matrix is perhaps the simplest example of a matrix that has the same sparsity pattern as the \mathbf{Y} -matrix. Other examples, involving larger blocks, are found in optimal power flow methods and three-phase power flows.

	1	2	3	4	5	6	7	8		
1	∞N $M \infty$					$H N$ $M L$			$\Delta \theta_1$ ΔV_1	ΔP_1 ΔQ_1
2		$H N$ $M \infty$			$H N$ $M L$				$\Delta \theta_2$ ΔV_2	ΔP_2 ΔQ_2
3			$H N$ $M L$			$H N$ $M L$		$H N$ $M L$	$\Delta \theta_3$ ΔV_3	ΔP_3 ΔQ_3
4				$H N$ $M \infty$	$H N$ $M L$			$H N$ $M L$	$\Delta \theta_4$ ΔV_4	ΔP_4 ΔQ_4
5		$H N$ $M L$		$H N$ $M L$	$H N$ $M L$		$H N$ $M L$		$\Delta \theta_5$ ΔV_5	ΔP_5 ΔQ_5
6	$H N$ $M L$		$H N$ $M L$			$H N$ $M L$	$H N$ $M L$		$\Delta \theta_6$ ΔV_6	ΔP_6 ΔQ_6
7					$H N$ $M L$	$H N$ $M L$	$H N$ $M L$	$H N$ $M L$	$\Delta \theta_7$ ΔV_7	ΔP_7 ΔQ_7
8			$H N$ $M L$	$H N$ $M L$			$H N$ $M L$	$H N$ $M L$	$\Delta \theta_8$ ΔV_8	ΔP_8 ΔQ_8

An alternative for adding large numbers to the unwanted pivots is to remove the corresponding rows and columns from the system. In this case the Jacobian matrix equation is written as above.

	2	3	4	5	6	7	8		
2	<i>H</i>			<i>H N</i>				$\Delta\theta_2$	ΔP_2
3		<i>H N</i>			<i>H N</i>		<i>H N</i>	$\Delta\theta_3$	ΔP_3
		<i>M L</i>			<i>M L</i>		<i>M L</i>	ΔV_3	ΔQ_3
4			<i>H</i>	<i>H N</i>			<i>H N</i>	$\Delta\theta_4$	ΔP_4
5	<i>H</i>		<i>H</i>	<i>H N</i>		<i>H N</i>		$\Delta\theta_5$	ΔP_5
	<i>M</i>		<i>M</i>	<i>M L</i>		<i>M L</i>		ΔV_5	ΔQ_5
6		<i>H N</i>			<i>H N</i>	<i>H N</i>		$\Delta\theta_6$	ΔP_6
		<i>M L</i>			<i>M L</i>	<i>M L</i>		ΔV_6	ΔQ_6
7				<i>H N</i>	<i>H N</i>	<i>H N</i>	<i>H N</i>	$\Delta\theta_7$	ΔP_7
				<i>M L</i>	<i>M L</i>	<i>M L</i>	<i>M L</i>	ΔV_7	ΔQ_7
8		<i>H N</i>	<i>H</i>			<i>H N</i>	<i>H N</i>	$\Delta\theta_8$	ΔP_8
		<i>M L</i>	<i>M</i>			<i>M L</i>	<i>M L</i>	ΔV_8	ΔQ_8

5.9 GAIN MATRIX

The gain matrix for Gauss Newton solution to normal equations is

$$\mathbf{G} = \mathbf{H}'\mathbf{W}\mathbf{H}$$

where \mathbf{H} is the measurement Jacobian matrix and \mathbf{W} is the weighting diagonal matrix. This gain matrix can be rewritten as follows:

$$\mathbf{G} = \sum_{i=1}^m w_i \mathbf{h}_i' \mathbf{h}_i$$

where \mathbf{h}_i is the i th row of the Jacobian matrix \mathbf{H} and w_i is the (i, i) th element of the weighting matrix \mathbf{W} . The expression above suggests that to form the gain matrix \mathbf{G} , or its structure, measurements can be processed one at a time, in any order. Only the column order matters, i.e., the order of the elements in a row.

For example, considering a dc model for the network in Fig. 5.23, the power flow measurement P_{37} gives the following contribution to the gain matrix:

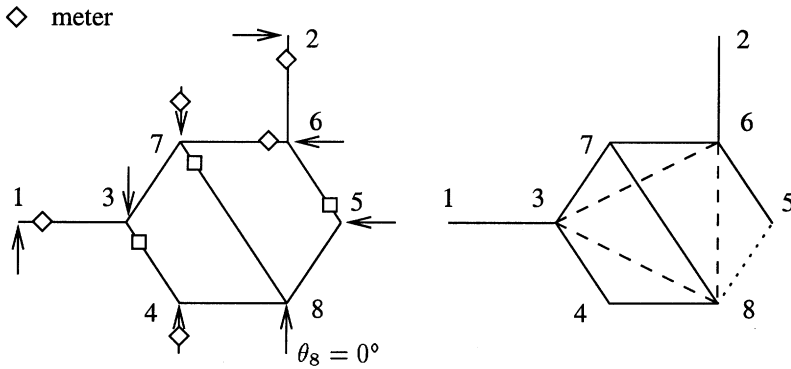


Figure 5.23. 8-bus network in Fig. 5.1 with both power flow and power injection measurements.

$$\Delta \mathbf{G}(P_{37}) = \begin{matrix} & \begin{matrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \end{matrix} & \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \times & 0 & 0 & 0 & \times & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \times & 0 & 0 & 0 & \times & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \end{matrix}$$

And the contribution of injection measurement P_4 is as follows:

$$\Delta \mathbf{G}(P_4) = \begin{matrix} & \begin{matrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \end{matrix} & \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \times & \times & 0 & 0 & 0 & \times \\ 0 & 0 & \times & \times & 0 & 0 & 0 & \times \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \times & \times & 0 & 0 & 0 & \times \end{pmatrix} \end{matrix}$$

The branch 3 – 8 (represented by a dashed line in Fig. 5.23) results from the squaring of injection measurement P_4 . In the gain matrix, this corresponds to the second order neighbors (3, 8) and (8, 3).

For a system with only power flow measurements, the gain matrix of the WLS state estimator presents the same sparsity pattern as the corresponding \mathbf{Y} matrix (the \mathbf{Y} matrix of a network formed by the branches with the given power flow measurements). The presence of bus injection measurements, however,

negatively affects sparsity: an injection measurement contributes to the gain matrix structure with elements corresponding to the branches connecting the node at which the measurement is taken and its neighborhood in all possible ways. Hence, for a system with injection measurements at all nodes, the gain has the same structure (and sparsity) as the square of the corresponding matrix \mathbf{Y} .

For the system illustrated in Fig. 5.23, for example, injection measurement at Node 2 creates the fill-in elements (3, 6), (3, 7) and (6, 7) in the gain matrix (Note that the element (6, 7) is also created by the flow measurement 6 – 7). In the same way, injection measurement at Node 4 is responsible for the fill-in elements (3, 4), (3, 8) and (4, 8) in the gain matrix (as above, the element (3, 4) is also created by the corresponding flow measurement).

Note that although the gain matrix is generally less sparse than the \mathbf{Y} matrix (and may in certain circumstances even have the same sparsity pattern as \mathbf{Y}^2), it is still very sparse for large networks, and justifies the use of sparse techniques. One reason for the good sparsity of the gain matrix triangular factors is that, at least partially, the elements created by injection measurements coincide with fill-ins that would be created anyway by the Gauss elimination process (network reduction).

Example 5.13:

Figure 5.23 (a) shows the network in Fig. 5.1 with six power flow meters (placed in Branches 1 – 3, 2 – 6, 3 – 4, 5 – 6, 7 – 8, and 6 – 7) and two power injection meters (located at Nodes 4 and 7). Part (b) of the same figure shows the graph representing the structure of the gain matrix: the solid lines represent branches of the original network which are adjacent to at least one measurement, whereas the dashed lines denote branches that are added due to the injection measurements

As mentioned above, the neighborhood of a node with an injection measurement is connected in all the possible ways. The dotted lines correspond to irrelevant branches, that is, branches of the power network with no adjacent measurements (They do not form part of the state estimation measurement model.)

The corresponding gain matrix of dc power flow model is as follows (for simplicity, the weighting factors were set at one):

$$\mathbf{G} = 10^4 \begin{pmatrix}
 1 & & & & & & & \\
 & 1 & & & & & & \\
 & & 1 & & & & & \\
 & -1 & & 55 & -24 & & & \\
 & & & -24 & 34 & & & \\
 & & & & & 64 & -64 & \\
 & & -1 & 24 & & -64 & 97 & -60 & 4 \\
 & & & -66 & & & -60 & 137 & -11 \\
 & & & 12 & -10 & & 4 & -11 & \infty
 \end{pmatrix}$$

Example 5.14:

Figure 5.24 (a) shows the 8-bus network in the previous example, with the same set of measurements. The structure of the corresponding gain matrix \mathbf{G} is given in part (b) of the figure. Part (c) shows the paths, and part (d) illustrates the connection between paths and the lower-triangle sparse factor of the gain matrix. Note that due to the existence of power injection measurements the gain matrix is less sparse than the corresponding bus admittance matrix, as observed above. There are only two fill-ins, however.

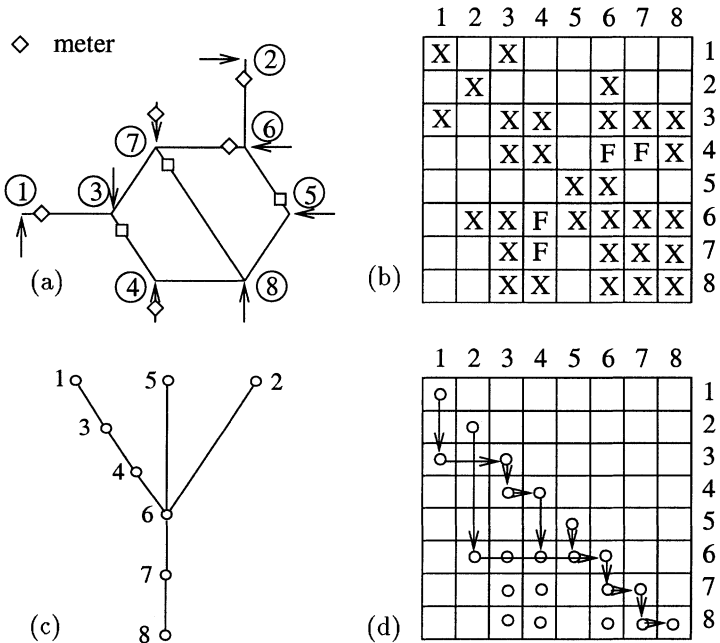


Figure 5.24. (a) 8-bus sample network. (b) Structure of the gain matrix \mathbf{G} . (c) Paths. (d) Paths shown on the structure of the \mathbf{L} matrix ($\mathbf{G} = \mathbf{L}\mathbf{L}'$).

5.10 FACTORIZATION OF RECTANGULAR MATRICES

In this section **LDU** factorization is extended to rectangular matrices. This type of factorization is also used in Chap. 7 to determine a basic set of measurements and in Chap. 13 in connection with the Peters Wilkinson method (In both cases the Jacobian matrix is factorized as $\mathbf{H} = \mathbf{LU}$.)

5.10.1 Partial and Complete Pivoting

Pivoting is normally used for both enhancing numerical stability and preserving sparsity. In previous sections, symmetrical pivoting was used to preserve sparsity of the gain matrix in the method involving normal equations (see the Tinney schemes). The Hachtel and Peter Wilkinson methods, however, may require more flexible pivoting strategies, as discussed below (see also Chap. 13).

Figure 5.25 shows three types of interchanges that are normally used for pivoting. The interchange of rows i and k of a matrix \mathbf{A} , indicated in Fig. 5.25(a), can be obtained by premultiplying this matrix by a permutation matrix \mathbf{P}_r , i.e.,

$$\mathbf{P}_r \mathbf{A} = \widehat{\mathbf{A}}$$

where $\widehat{\mathbf{A}}$ is the transformed matrix and \mathbf{P}_r is obtained from a unit matrix by performing the same row interchange that is to be made in the matrix \mathbf{A} . Similarly, to interchange columns i and m of \mathbf{A} , a permutation matrix \mathbf{P}_c is used, i.e., $\mathbf{A} \mathbf{P}_c = \widehat{\mathbf{A}}$. Complete pivoting is then obtained as follows:

$$\mathbf{P}_r \mathbf{A} \mathbf{P}_c = \widehat{\mathbf{A}}$$

When complete pivoting is required to improve numerical stability, at each stage of the elimination process the unprocessed element with the largest ab-

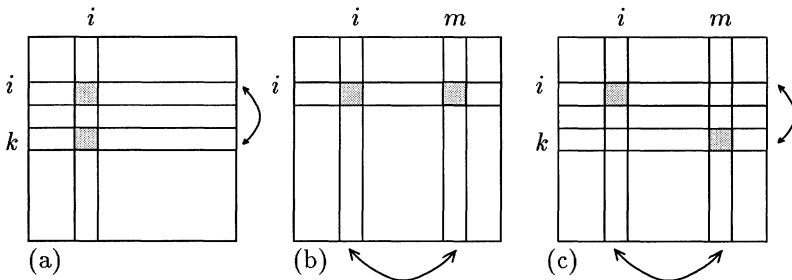


Figure 5.25. (a) Partial pivoting by row interchange. (b) Partial pivoting by column interchange. (c) Complete pivoting.

solute value is chosen as the pivot. Both row and column interchanges are normally necessary, and the permutation matrices \mathbf{P}_r and \mathbf{P}_c are obtained as the product of a sequence of elementary permutation matrices.

5.10.2 Factorization

Consider an overdetermined system of linear equations $\mathbf{A}\mathbf{x} = \mathbf{b}$ where the $m \times n$ matrix \mathbf{A} has full rank. The interchanged coefficient matrix $\hat{\mathbf{A}}$ can be transformed into a unit upper trapezoidal $m \times n$ matrix \mathbf{L} by the application of elementary operations as in Eq. (5.10), i.e.:

$$\hat{\mathbf{A}} [\mathbf{U}_1^1 \mathbf{U}_1^2 \dots \mathbf{U}_1^{n_u}] [\mathbf{D}_1^1 \mathbf{D}_1^2 \dots \mathbf{D}_1^n] = \mathbf{L} \tag{5.21}$$

where the elementary matrices \mathbf{U}_1^i are used to zeroize the upper triangle of the matrix $\hat{\mathbf{A}}$, whereas the elementary matrices \mathbf{D}_1^i are used to normalize the diagonal elements. Hence, the matrix $\hat{\mathbf{A}}$ can be written as follows:

$$\hat{\mathbf{A}} = \mathbf{L} [\mathbf{D}^n \dots \mathbf{D}^2 \mathbf{D}^1] [\mathbf{U}^{n_u} \dots \mathbf{U}^2 \mathbf{U}^1]$$

or

$$\hat{\mathbf{A}} = \mathbf{L}\mathbf{D}\mathbf{U} \tag{5.22}$$

where \mathbf{U} is an upper triangular $n \times n$ matrix. Factorization is illustrated in Fig. 5.26.

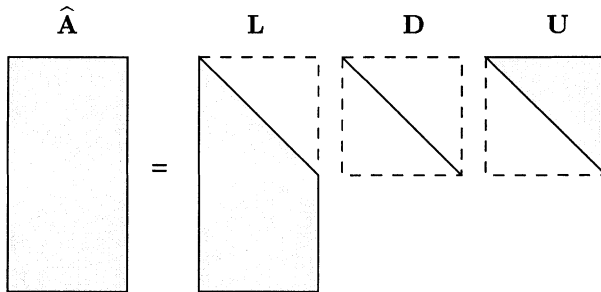


Figure 5.26. Factorization of a rectangular matrix. ($\hat{\mathbf{A}}$ is a permutation of an $m \times n$ matrix \mathbf{A} , \mathbf{L} is a unit lower trapezoidal $m \times n$ matrix, \mathbf{D} is a diagonal $n \times n$ matrix, and \mathbf{U} is an upper triangular $n \times n$ matrix.)

Remarks:

- To obtain the trapezoidal matrix \mathbf{L} both column and row permutations are normally necessary as indicated above. If, however, off-diagonal pivots are

allowed, a factorization $\mathbf{A} = \widehat{\mathbf{L}}\widehat{\mathbf{D}}\widehat{\mathbf{U}}$ can be obtained where $\widehat{\mathbf{L}}$ is a row permutation of a unit trapezoidal $m \times n$ matrix and $\widehat{\mathbf{U}}$ is a column permutation of an upper triangular $n \times n$ matrix. If the corresponding permutations are performed in $\widehat{\mathbf{L}}$ and $\widehat{\mathbf{U}}$, they will be transformed into the matrices \mathbf{L} and \mathbf{U} that appear in the decomposition in Eq. (5.22).

- To preserve the elements that have already been zeroized, elementary operations \mathbf{U}_i^j in Eq. (5.21) can be performed either row-wise from left to right or column-wise from top to bottom.
- To preserve sparsity interchanges are usually necessary. Pivot selection depends on row/column counts, i.e., the number of nonzero elements in a row/column. In the classical Markowitz scheme the pivot selected is the nonzero element which yields the lowest product of row and column counts. Another scheme is the Duff criterion for selecting the pivotal column as the one with the minimum current count, with the nonzero pivotal element within it having the lowest number of nonzeros in its row.

Example 5.15:

Consider the system in Fig. 5.23 with power flow measurement variances of 10^{-2} , power injection variances of 10^{-8} and angle-reference pseudo-measurement variance of 10^{16} . For simplicity, all branch reactances are considered to be 1 p.u. The weighted Jacobian matrix for this example can be written as:

$$\mathbf{H} = \begin{matrix} P_{13}^{(1)} \\ P_{34}^{(4)} \\ P_{78}^{(7)} \\ P_{67}^{(6)} \\ P_{56}^{(3)} \\ P_{26}^{(2)} \\ P_4^{(5)} \\ P_7 \\ \theta_8^{(8)} \end{matrix} \begin{pmatrix} \theta_1^{(1)} & \theta_2^{(2)} & \theta_3^{(5)} & \theta_4^{(4)} & \theta_5^{(3)} & \theta_6^{(6)} & \theta_7^{(7)} & \theta_8^{(8)} \\ 10 & & -10 & & & & & \\ & & 10 & -10 & & & & \\ & & & & & & 10 & -10 \\ & & & & & 10 & -10 & \\ & & & & 10 & -10 & & \\ & 10 & & & & -10 & & \\ & & -100 & 200 & & & & -100 \\ & & -100 & & & -100 & 300 & -100 \\ & & & & & & & 10^8 \end{pmatrix}$$

The pivoting order is indicated by the superscripts: for example, the first pivot is $(P_{13}^1; \theta_1^1)$, the second pivot is $(P_{26}^2; \theta_2^2)$, and so on. This matrix can be factorized as $\mathbf{H} = \widehat{\mathbf{L}}\widehat{\mathbf{U}}$, where

$$\hat{\mathbf{L}} = \begin{matrix} P_{13} \\ P_{34} \\ P_{78} \\ P_{67} \\ P_{56} \\ P_{26} \\ P_4 \\ P_7 \\ \theta_8 \end{matrix} \begin{pmatrix} \theta_1 & \theta_2 & \theta_3 & \theta_4 & \theta_5 & \theta_6 & \theta_7 & \theta_8 \\ 1 & & & & & & & \\ & & & 1 & & & & \\ & & & & & & 1 & \\ & & & & & 1 & & \\ & & & & 1 & & & \\ & 1 & & & & & & \\ & & 1 & 20 & & & & \\ & & -1 & & -10 & 20 & & \\ & & & & & & & 1 \end{pmatrix}$$

$$\hat{\mathbf{U}} = \begin{matrix} \theta_1 \\ \theta_2 \\ \theta_5 \\ \theta_4 \\ \theta_3 \\ \theta_6 \\ \theta_7 \\ \theta_8 \end{matrix} \begin{pmatrix} \theta_1 & \theta_2 & \theta_5 & \theta_4 & \theta_3 & \theta_6 & \theta_7 & \theta_8 \\ 10 & & -10 & & & & & \\ & 10 & & & & & -10 & \\ & & 100 & & & & & -1 \\ & & -10 & 10 & & & & \\ & & & & 10 & -10 & & \\ & & & & & 10 & -10 & \\ & & & & & & 10 & -10 \\ & & & & & & & 10^8 \end{pmatrix}$$

Using the pivot order indicated in matrix \mathbf{H} above, the matrix $\hat{\mathbf{L}}$ can be rewritten in the following trapezoidal form:

$$\mathbf{L} = \begin{matrix} P_{13}^{(1)} \\ P_{26}^{(2)} \\ P_{56}^{(3)} \\ P_{34}^{(4)} \\ P_4^{(5)} \\ P_{67}^{(6)} \\ P_{78}^{(7)} \\ \theta_8^{(8)} \\ P_7 \end{matrix} \begin{pmatrix} \theta_1^{(1)} & \theta_2^{(2)} & \theta_5^{(3)} & \theta_4^{(4)} & \theta_3^{(5)} & \theta_6^{(6)} & \theta_7^{(7)} & \theta_8^{(8)} \\ 1 & & & & & & & \\ & 1 & & & & & & \\ & & 1 & & & & & \\ & & & 1 & & & & \\ & & & 20 & 1 & & & \\ & & & & & 1 & & \\ & & & & & & 1 & \\ & & & & & & & 1 \\ & & & & -1 & -10 & 20 & \end{pmatrix}$$

5.11 MATRIX INVERSION LEMMA

Consider an $(m + n) \times (m + n)$ matrix partitioned as follows:

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} \tag{5.23}$$

where \mathbf{A} is an $n \times n$ matrix and \mathbf{D} is an $m \times m$ matrix.

If \mathbf{A} has an inverse, the following block elementary transformations can be performed (as in Eq. (5.10)):

$$\begin{pmatrix} \mathbf{I}_n & \mathbf{0} \\ -\mathbf{CA}^{-1} & \mathbf{I}_m \end{pmatrix} \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{0} & \mathbf{D} - \mathbf{CA}^{-1}\mathbf{B} \end{pmatrix}$$

$$\begin{pmatrix} \mathbf{A}^{-1} & \mathbf{0} \\ \mathbf{0} & (\mathbf{D} - \mathbf{CA}^{-1}\mathbf{B})^{-1} \end{pmatrix} \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{0} & \mathbf{D} - \mathbf{CA}^{-1}\mathbf{B} \end{pmatrix} = \begin{pmatrix} \mathbf{I}_n & \mathbf{A}^{-1}\mathbf{B} \\ \mathbf{0} & \mathbf{I}_m \end{pmatrix}$$

$$\begin{pmatrix} \mathbf{I}_n & -\mathbf{A}^{-1}\mathbf{B} \\ \mathbf{0} & \mathbf{I}_m \end{pmatrix} \begin{pmatrix} \mathbf{I}_n & \mathbf{A}^{-1}\mathbf{B} \\ \mathbf{0} & \mathbf{I}_m \end{pmatrix} = \begin{pmatrix} \mathbf{I}_n & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_m \end{pmatrix}$$

Hence, as in Eq. (5.12), the corresponding partitioned inverse matrix can be written as follows:

$$\begin{aligned} \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix}^{-1} &= \begin{pmatrix} \mathbf{I}_n & -\mathbf{A}^{-1}\mathbf{B} \\ \mathbf{0} & \mathbf{I}_m \end{pmatrix} \begin{pmatrix} \mathbf{A}^{-1} & \mathbf{0} \\ \mathbf{0} & (\mathbf{D} - \mathbf{CA}^{-1}\mathbf{B})^{-1} \end{pmatrix} \begin{pmatrix} \mathbf{I}_n & \mathbf{0} \\ -\mathbf{CA}^{-1} & \mathbf{I}_m \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{A}^{-1} + \mathbf{A}^{-1}\mathbf{B}(\mathbf{D} - \mathbf{CA}^{-1}\mathbf{B})^{-1}\mathbf{CA}^{-1} & -\mathbf{A}^{-1}\mathbf{B}(\mathbf{D} - \mathbf{CA}^{-1}\mathbf{B})^{-1} \\ -(\mathbf{D} - \mathbf{CA}^{-1}\mathbf{B})^{-1}\mathbf{CA}^{-1} & (\mathbf{D} - \mathbf{CA}^{-1}\mathbf{B})^{-1} \end{pmatrix} \end{aligned} \tag{5.24}$$

Alternatively, elementary operations can be applied to zeroize first the upper triangular block and then the lower triangular block, resulting in the following factorization:

$$\begin{pmatrix} \mathbf{I}_n & -\mathbf{BD}^{-1} \\ \mathbf{0} & \mathbf{I}_m \end{pmatrix} \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} = \begin{pmatrix} \mathbf{A} - \mathbf{BD}^{-1}\mathbf{C} & \mathbf{0} \\ \mathbf{C} & \mathbf{D} \end{pmatrix}$$

$$\begin{pmatrix} (\mathbf{A} - \mathbf{BD}^{-1}\mathbf{C})^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{D}^{-1} \end{pmatrix} \begin{pmatrix} \mathbf{A} - \mathbf{BD}^{-1}\mathbf{C} & \mathbf{0} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} = \begin{pmatrix} \mathbf{I}_n & \mathbf{0} \\ \mathbf{D}^{-1}\mathbf{C} & \mathbf{I}_m \end{pmatrix}$$

$$\begin{pmatrix} \mathbf{I}_n & \mathbf{0} \\ -\mathbf{D}^{-1}\mathbf{C} & \mathbf{I}_m \end{pmatrix} \begin{pmatrix} \mathbf{I}_n & \mathbf{0} \\ \mathbf{D}^{-1}\mathbf{C} & \mathbf{I}_m \end{pmatrix} = \begin{pmatrix} \mathbf{I}_n & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_m \end{pmatrix}$$

In this case, the partitioned inverse matrix can be expressed as follows:

$$\begin{aligned} \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix}^{-1} &= \begin{pmatrix} \mathbf{I}_n & \mathbf{0} \\ -\mathbf{D}^{-1}\mathbf{C} & \mathbf{I}_m \end{pmatrix} \begin{pmatrix} (\mathbf{A} - \mathbf{BD}^{-1}\mathbf{C})^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{D}^{-1} \end{pmatrix} \begin{pmatrix} \mathbf{I}_n & -\mathbf{BD}^{-1} \\ \mathbf{0} & \mathbf{I}_m \end{pmatrix} \\ &= \begin{pmatrix} (\mathbf{A} - \mathbf{BD}^{-1}\mathbf{C})^{-1} & -(\mathbf{A} - \mathbf{BD}^{-1}\mathbf{C})^{-1}\mathbf{BD}^{-1} \\ -\mathbf{D}^{-1}\mathbf{C}(\mathbf{A} - \mathbf{BD}^{-1}\mathbf{C})^{-1} & \mathbf{D}^{-1} + \mathbf{D}^{-1}\mathbf{C}(\mathbf{A} - \mathbf{BD}^{-1}\mathbf{C})^{-1}\mathbf{BD}^{-1} \end{pmatrix} \end{aligned} \tag{5.25}$$

Identifying the corresponding submatrices in Eqs.(5.24)-(5.25) results in the following:

$$(\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C})^{-1} = \mathbf{A}^{-1} + \mathbf{A}^{-1}\mathbf{B}(\mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B})^{-1}\mathbf{C}\mathbf{A}^{-1} \quad (5.26)$$

$$\mathbf{A}^{-1}\mathbf{B}(\mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B})^{-1} = (\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C})^{-1}\mathbf{B}\mathbf{D}^{-1}$$

$$(\mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B})^{-1}\mathbf{C}\mathbf{A}^{-1} = \mathbf{D}^{-1}\mathbf{C}(\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C})^{-1}$$

$$(\mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B})^{-1} = \mathbf{D}^{-1} + \mathbf{D}^{-1}\mathbf{C}(\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C})^{-1}\mathbf{B}\mathbf{D}^{-1}$$

The result expressed in Eq. (5.26) is known as the matrix inversion lemma and has been used in various power network algorithms such as the compensation method applied to contingency analysis and bad data processing (in state estimation). It is also important in deriving the equations of the Kalman filter and related parameter estimation algorithms (see Chap. 11).

5.12 HISTORICAL NOTES AND REFERENCES

The classic Ward equivalent method is an extension of the Norton equivalent method adapted to power flow calculations (Ward [1949]). The relationship between network reduction and Gauss elimination applied to sparse network matrices was clarified by Tinney, Powell and Peterson [1973]. Various reviews of equivalents have also been written, including two papers by Deckmann, Pizzolante, Monticelli, Stott and Alsaç [1980] and a paper by Wu and Monticelli [1983]. The concept of paths for exploring vector sparsity was introduced by Tinney, Brandwajn, and Chan [1985], and an adaptive approach using such techniques in power flow equivalencing was proposed by Tinney and Bright [1987]. More recently, a critical analysis of practical experience with external power flow equivalent models was presented Kato [1994].

Sparse Gauss elimination applied to \mathbf{Y} bus type network matrices was introduced in the classical paper by Tinney and Hart [1967]. An efficient method for performing solution updating that can be used in problems such as contingency analysis and bad data processing, which is a generalization of the compensation approach introduced by Tinney [1972], was then proposed by Alsaç, Stott, and Tinney [1983]. The sparse inverse matrix method was used to compute the variance of estimated quantities in WLS estimators by Broussolle [1978]. The application of sparse vector techniques (paths) to localized solution update was proposed by Bacher and Tinney [1989]. A new way of obtaining a sparse-matrix inverse (sparse- \mathbf{Z}) using paths that can be used to compute normalized residuals, was then devised by Enns, Tinney, and Alvarado [1990]. A general approach to partial refactorization was introduced by Zhang and Tinney [1995].

Numerical stability and sparsity preservation were discussed and tested for normal equations, orthogonal methods, and those of Hachtel and Peters Wilkinson by Duff and Reid [1976]. Two general books are quite helpful in the understanding of matrix factorization methods. The book by Watkins [1991] gives a general introduction to methods commonly used to solve systems of linear equations, and the one by Duff and Erisman [1992] deals with sparse matrix techniques in a variety of applications.

5.13 PROBLEMS

- 1. Consider the network in Fig. 5.1. Reduce it by eliminating Nodes 2 and 6. Solve the reduced network for the nodal voltages, and compare the results with those shown in the 7th row of Fig. 5.8.
- 2. Consider that the nodal voltages of the network presented in Fig. 5.1 are known (see Example 5.3). Update the solution when Branch 2-6 is removed from the network: (a) by inspection; (b) using reduction and expansion as in Example 5.3.
- 3. Consider the following elimination order for the network in Fig. 5.1: 4, 1, 2, 3, 5, 6, 7, 8. Draw the graph showing the new nodal injection distributions and the path graph (see Fig. 5.11).
- 4. For the same elimination order in Prob. 3, update the nodal voltages when the current injection in Node 2 changes by $\Delta I_2 = -4$.
- 5. Use the paths obtained in Prob. 3 to determine the minimum number of operations needed to update the nodal voltages when: (a) current injection at Node 1 is changed; (b) Branch 5-6 is removed from the network.
- 6. In Example 5.15, (a) write the matrix \mathbf{U} using the pivoting order used in factorization and (b) determine the permutation matrices \mathbf{P}_r and \mathbf{P}_c that transform the matrices $\hat{\mathbf{L}}$ and $\hat{\mathbf{U}}$ into the matrices \mathbf{L} and \mathbf{U} .
- 7. For the network shown in Fig. 5.24, Example 5.14, form the blocked gain matrix of the Gauss Newton state estimator (see Sec. 5.8).

References

- Alsaç O., Stott, B., and Tinney, W. F., "Sparsity-oriented compensation methods for modified network solutions", IEEE Trans. Power Syst., Vol. 102, pp. 1050-1060, May 1983.
- Bacher, R. and Tinney, W.F., "Faster local-power flow solutions - the zero mismatch approach", IEEE Trans. Power Syst., Vol. 4, pp. 1345-1354, Nov. 1989.
- Betancourt, R., "An efficient heuristic ordering algorithm for partial matrix refactorization", IEEE Trans. Power Syst., Vol. 3, No. 3, pp. 1181-1187, Nov. 1989.

- Broussolle, F., "State estimation in power systems: Detecting bad data through the sparse inverse matrix method", *IEEE Trans. Power App. and Syst.*, Vol. 97, No. 3, pp. 678-682, May/June 1978.
- Deckmann, S., Pizzolante, A., Monticelli, A., Stott, O., and Alsaç, O., "Studies on power system load flow equivalencing", *IEEE Trans. Power App. Syst.*, Vol. 99, pp. 2301-2310, 1980.
- Deckmann, S., Pizzolante, A., Monticelli, A., Stott, O., and Alsaç, O., "Numerical Testing on power system load flow equivalencing", *IEEE Trans. Power App. Syst.*, Vol. 99, pp. 2292-2300, 1980.
- Duff, I.F. and Reid, J.K., "A comparison of some methods for the solution of sparse overdetermined systems of linear equations", *J. Inst. Maths Applies*, Vol. 17, pp. 267-280, 1976.
- Duff, I.F. and Erisman, A.M., *Direct Methods for Sparse Matrices*, Oxford Science Publications, Oxford, 1992.
- Enns, M.K., Tinney, W.F., and Alvarado, F.L., "Sparse-matrix inverse factors", *IEEE Trans. Power Syst.*, Vol. 5, pp. 466-473, May 1990.
- Kato K., Chairman, External Network Modeling Task Force, "External network modeling -Recent practical experience", *IEEE Trans. Power Syst.*, Vol. 9, pp. 216-228, 1994.
- Tinney, W. F. and Hart, C. E., "Power flow solution by Newton's method", *IEEE Trans. Power App. Syst.*, Vol. 86, pp. 1449-1456, 1967.
- Tinney, W. F., "Compensation methods with ordered triangular factorization", *IEEE Trans. Power Syst.*, Vol. 91, pp. 123-127, Jan/Feb 1972.
- Tinney, W. F., Powell, W.L., and Peterson, N.M., "Sparsity oriented network reduction", *Proc. of PICA Conf.*, pp. 385-390, Minneapolis, 1973.
- Tinney, W. F., Brandwajn, V., and Chan, S.M., "Sparse Vector Methods", *IEEE Trans. Power App. Syst.*, Vol. 104, pp. 295-301, 1985.
- Tinney, W. F. and Bright, J.M., "Adaptive reductions for power flow equivalents", *IEEE Trans. Power Syst.*, Vol. 2, pp. 351-360, 1987.
- Zhang, Y. and Tinney, W.F., "Partial refactorization with unrestricted topology changes", *IEEE Trans. Power Syst.*, Vol. 10, pp. 1361-1368, Aug. 1995.
- Watkins, D.S., *Fundamentals of Matrix Computations*, John Wiley and Sons, New York, 1991.
- Ward, J.B., "Equivalent circuits for power flow studies", *AIEE Trans. Power App. Syst.*, Vol. 68, pp. 373-382, 1949.
- Wu, F.F. and Monticelli, A., "Critical review of external network modeling for online security analysis", *Int. J. Elect. Power and Energy Syst.*, Vol. 5, No. 4, pp. 222-235, 1983.

6 NETWORK TOPOLOGY PROCESSING

This chapter describes the basic features of network topology processors (NTP), both conventional and generalized. A number of EMS functions use network connectivity information produced by NTPs, although here the main interest is in state estimation. What makes the state estimation application different from other network applications is that not only the network connectivity, but also the location of metering devices in the identified power network, must be considered.

6.1 CONVENTIONAL TOPOLOGY PROCESSING

The network topology processor determines both the connectivity of the electrical network and the location of metering devices in the identified power network. The data-base is assumed to contain a complete description of the network model and the location of metering devices in terms of bus-sections and switching-devices (breakers and switches). The status of switching devices can either be telemetered or entered manually by system operators. Conventional topology processing is performed before state estimation and other related functions such as observability analysis and bad data processing; once the network topology is known, state estimation assumes that this topology is correct and proceeds to estimate the state and identify analog bad data whenever possible (i.e., when meter redundancy allows this).

In the conventional approach, the state estimator uses a bus/branch model of the same type used in power flow calculations. However, actual data gathered from the field is not in a bus-branch format, since network connectivity is normally described in terms of bus-sections and switching-devices (physical level representation). One of the functions of the network topology processor is to transform the bus-section/switching-device model into the bus/branch model required by the state estimator and for other network analysis functions. A second function performed by NTP is the assignment of metering devices to the components of the bus/branch network model identified.

In more conventional implementations, the real time modeling of a power network usually follows a 6-step procedure involving:

1. Data gathering;
2. Network topology processing;
3. Observability analysis;
4. State estimation (central algorithm);
5. Processing of bad data (state re-estimation);
6. Real time power network modeling (power flow model);

Step 1 assumes a bus-section/switching-device model. Steps 2 and 3 assume that switching device status is correct. Step 4 additionally assumes that the parameters are correct. Step 5 processes bad data assuming that they are caused by analog measurements. In this phase, topological errors usually appear in the form of analog bad data. The sequential removal of these bits of data from the neighborhood of the topological error may eventually reduce local redundancy to zero. When this happens, the remaining set of analog data becomes consistent. But the topology error remains. A similar analysis can be conducted for parameter errors.

Thus, in the conventional approach, the more general problem of extracting a real time model from the available data set is decomposed into two main subproblems: processing of logical data (the status of switching devices), and processing of analog data (e.g., power flow, power injection and voltage magnitude measurements). During topology processing, the status of breaker/switch is processed using a bus-section/switching-device network model of the type illustrated in Fig. 6.1. During observability analysis and state estimation, the network topology and parameters are considered as given, and analog data is processed using the bus/branch network model (Fig. 6.2). Data consistency checking is performed, in the conventional approach, with the logical data checked by the topology processor and the analog data checked by the state estimator.

6.1.1 Pre-Processing of Raw Data

The first task of the network topology processor is to convert analog raw measurements to the appropriate units. Then, certain simple tests and checks are

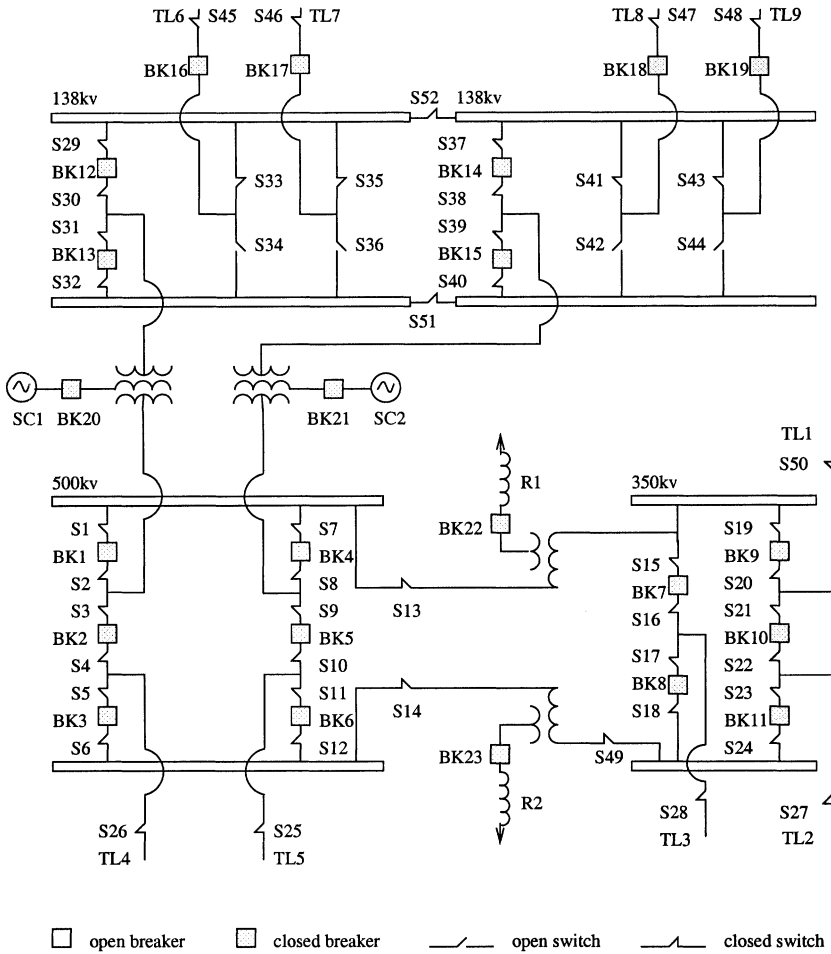


Figure 6.1. Bus-section/switching-device model of a substation with three voltage levels. (BK represents circuit breakers; S represents switches; TL represents transmission lines; SC represents synchronous condensers; and R represents shunt reactors.)

performed: verification of operating limits, verification of limits on the rate of change of operating variables, testing of non-zero flows in open switching devices, and testing of non-zero voltage differences across closed switching devices.

6.1.2 Bus Section Processing

In this phase, the connectivity of a bus section group is determined. Consider the substation shown in Fig. 6.1. If, for example, all the 500 kV switching

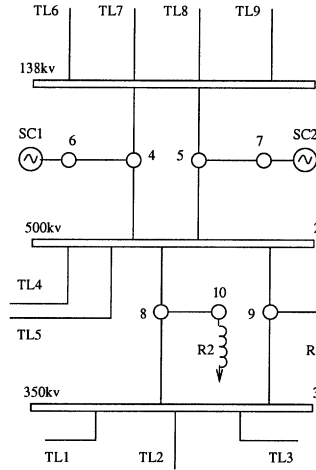


Figure 6.2. Bus/branch model of the substation presented in Fig. 5.1.

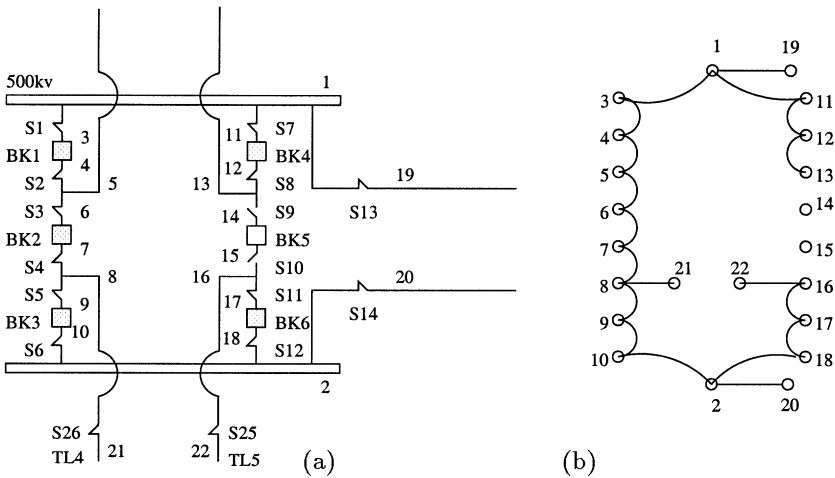


Figure 6.3. (a) 500 kV bus section group of the substation illustrated in Fig. 5.1, with Switches S9 and S10, and Breaker BK5 open. (b) Graph with nodes representing bus sections and branches representing closed switching devices.

devices (breakers and switches) are closed, then the bus sections connected by these switching devices will merge into a single bus; the same would happen for devices at the 350 kV and 138 kV voltage levels. Such a system of bus sections that forms a bus when all switching devices are closed is denoted as a bus

section group. Bus section groups are determined at start-up and the relevant data structures are retained in the data base for later use by the tracking mode of the NTP.

Switching devices are not the only aspect which must be considered however. Substations may have branch devices (such as transmission lines, transformers, phase shifters, and series devices), shunt devices (such as capacitors, reactors, synchronous condensers, static VAR compensators, loads and generators), and metering devices (such as power and current flow meters, power and current injection meters, and voltage magnitude meters) as well. The connection of these devices in a bus/branch model requires the determination of the network buses.

Bus section processing consists of merging bus sections of a bus section group into one or more network buses (buses of the bus/branch network model). For example, Fig. 6.3 (a) shows the 500 kV bus section group of the substation presented in Fig. 6.1; part (b) of the figure shows the connectivity graph associated with the bus-section group. A simple tree search through the graph reveals the network buses in the bus-section group.

Once such network buses are formed, data structures (pointers and links) are built to associate them with branch and shunt devices.

6.1.3 Network Connectivity Analysis

Connectivity analysis identifies energized, de-energized and grounded electrical islands. The NTP creates two data structures: a bus-to-branch data structure and a network-to-device data structure. Consider, for example, the substation illustrated in Fig. 6.1: for the switch and breaker status indicated in the figure, the substation can be reduced to the bus/branch network model given in Fig. 6.2.

Figure 6.8 provides a few examples of possible power meter arrangements. In processing network topology, part of the original model is reduced, and the meters originally associated with devices that are eliminated from the model have to be re-assigned (whenever possible) to the newly created network components (branches and buses). This aspect of network topology processing is critical for the correct execution of state estimation and related functions.

6.1.4 Topology Processor in Tracking Mode

In the tracking mode, the network topology processor updates the parts of the power network affected by switching device status changes. Only bus section groups where changes have occurred are processed and the associated data structures are updated accordingly. Switching device status changes can modify both the way bus sections are grouped into network buses and the association of branch and shunt devices with network buses. The location of metering devices may be affected as well. Thus, the new data structures relating network buses to various devices (branch, shunt and metering) are compared with the corresponding structures saved from the previous run. Simple plausibility

checks are then performed for changed bus section groups using KCL and KVL (Kirchhoff's laws) .

There are cases in which changes in a bus section group also affect network connectivity (e.g., cases when a bus splits or a branch device switches from one bus to other). In these cases, the data structures describing network connectivity and network islands are updated by NTP in tracking mode.

The benefits of the tracking mode are not limited to the topology processor itself: in the case of minor changes, or when no changes occur, matrix structures (including the optimal pivoting order) used in other applications can be reused up to a certain point, although eventually the cumulative effect of changes will require a complete matrix refactorization.

6.1.5 *Critical Remarks*

If topological errors, whether due to telemetry or operator entry errors, pass undetected through configuration analysis, one of the basic assumptions of state estimation will no longer be valid, and the state estimator will yield incorrect results. Processing topology errors using the conventional bus/branch network model is not always effective. This is especially true during the implementation phase of the state estimator and its related functions, when the number and frequency of errors can be relatively high. A more robust state estimation approach has been developed, the generalized state estimation, to cope with these critical cases. In this approach, an integrated estimation of state, status and parameters is made. The generalized topology processor described in the next section is a critical part of this generalized estimator.

6.2 GENERALIZED TOPOLOGY PROCESSING

In addition to identifying energized, de-energized and grounded electrical island, as a conventional NTP does, a generalized topology processor identifies extended islands in which switching devices that appear in the data base as being open can be explicitly represented. Unknown and suspect status can also be represented explicitly in the model. For example, the status of switching devices of a bus section group in which one or more changes have occurred since last time, can be considered as suspect. Such explicit representation facilitates bad data processing, since possible status error will appear in state estimation as such, and thus will not be disguised as errors in the analog measurements.

The generalized NTP is also able to run a local weighted least squares state estimator at the bus-section/switching-device level for areas containing suspected status or analog data, such as a bus section group or a substation. Depending on the results of the state estimation, bad analog/status data can be removed or, if redundancy does not permits a safe decision, the suspect area is kept in the bus-section/switching-device model level for further evaluation during state estimation of the entire network.

6.2.1 Extended Islands

The concept of extended islands is crucial to generalized state estimation. In certain cases, an incorrect status leads a conventional NTP to identify two islands where there is in fact only a single network. It can also lead to the identification of a de-energized island where the island is actually part of a larger, energized network. To cope with these situations the concept of extended island is introduced here (see also Chap. 7).

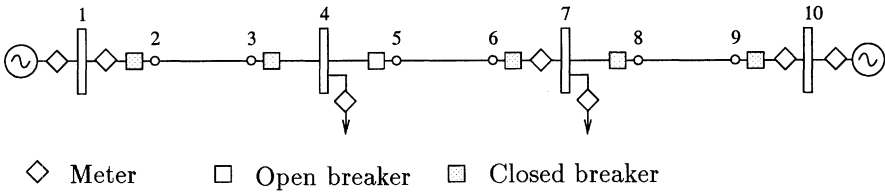


Figure 6.4. The status of Breaker 4 – 5 is incorrect (it is closed not open). Conventional NTP identifies two separate islands.

Figure 6.4 illustrates the case in which the telemetered status of Breaker 4 – 5 is incorrect: it reads open, whereas in the field, the breaker is actually closed. A conventional NTP would identify two separate islands as indicated in the figure. Assuming that the actual flow in the line 5 – 6 is not zero, and that all the analog data are correct, bad analog data will be detected in both islands, which will lead to a completely incorrect network model. If, however, the open breaker 4 – 5 is retained in an extended model, then this incorrect status can be detected and identified by the generalized state estimator applied to the extended island; the correct model will then be identified without the misclassification of good analog data as being bad.

Example 6.1

Consider the system represented in Fig. 6.5 (This system has been studied in Example 3.15, Chap. 3). The measured values are as follows: $P_G^{meas} = 1.55$ p.u., $P_L^{meas} = -0.62$ p.u., and $P_{65}^{meas} = -0.88$ p.u. The corresponding variances are $\sigma_{P_G}^2 = .001$, $\sigma_{P_L}^2 = .001$ and $\sigma_{P_{65}}^2 = .001$, with the pseudo-measurement variances $\sigma_{pseudo}^2 = .00001$ arbitrarily assigned.

A conventional NTP would identify three separate islands in this case: an island formed by Node 1, a second island formed by Nodes 2, 3, and 4, and a third island formed by Nodes 5 and 6. A generalized dc power state estimation model, however, can solve the entire system as a single island, and, at the end, decide which is the correct configuration as shown in the following.

In this case the measurement model is given by (see Example 3.15):

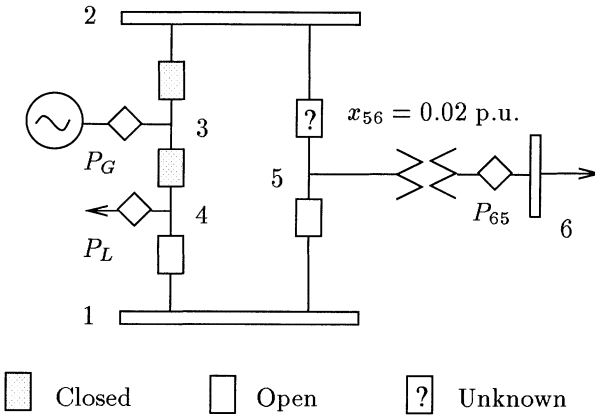


Figure 6.5. Bus-section/switching-device model with a breaker with an unknown status (Example 6.1).

$$\begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & -50 & 0 & -1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & 1 \\ 0 & 0 & 0 & -1 & 0 & 0 & 0 & -1 \\ 0 & 0 & 50 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \theta_3 \\ \theta_4 \\ \theta_6 \\ P_{14} \\ P_{15} \\ P_{23} \\ P_{25} \\ P_{34} \end{pmatrix} = \begin{pmatrix} P_{14}^{pseudo} = 0.0 \\ P_{15}^{pseudo} = 0.0 \\ \theta_{23}^{pseudo} = 0.0 \\ \theta_{34}^{pseudo} = 0.0 \\ P_1^{pseudo} = 0.0 \\ P_2^{pseudo} = 0.0 \\ P_5^{pseudo} = 0.0 \\ P_G^{meas} = 1.55 \\ P_L^{meas} = -0.62 \\ P_{65}^{meas} = -0.88 \end{pmatrix}$$

The state estimate and the estimated values of the measured quantities as follows

$$\hat{\mathbf{x}} = \begin{pmatrix} \theta_3 \\ \theta_4 \\ \theta_6 \\ P_{14} \\ P_{15} \\ P_{23} \\ P_{25} \\ P_{34} \end{pmatrix} = \begin{pmatrix} 0.0 \\ 0.0 \\ -0.0179 \\ -0.00005 \\ -0.00005 \\ -0.897 \\ 0.897 \\ 0.637 \end{pmatrix} ; \quad \hat{\mathbf{z}} = \begin{pmatrix} \hat{P}_{14} \\ \hat{P}_{15} \\ \hat{\theta}_{23} \\ \hat{\theta}_{34} \\ \hat{P}_1 \\ \hat{P}_2 \\ \hat{P}_5 \\ \hat{P}_G \\ \hat{P}_L \\ \hat{P}_{65} \end{pmatrix} = \begin{pmatrix} -0.00005 \\ -0.00005 \\ 0.0 \\ 0.0 \\ -0.00011 \\ -0.00016 \\ -0.00016 \\ 1.533 \\ -0.637 \\ -0.897 \end{pmatrix}$$

The standard deviation of the pseudo measurements $P_{14}^{pseudo} = 0$ and $P_{15}^{pseudo} = 0$ (the active power flows in the open breakers) was assumed to be $\sigma_{pseudo} =$

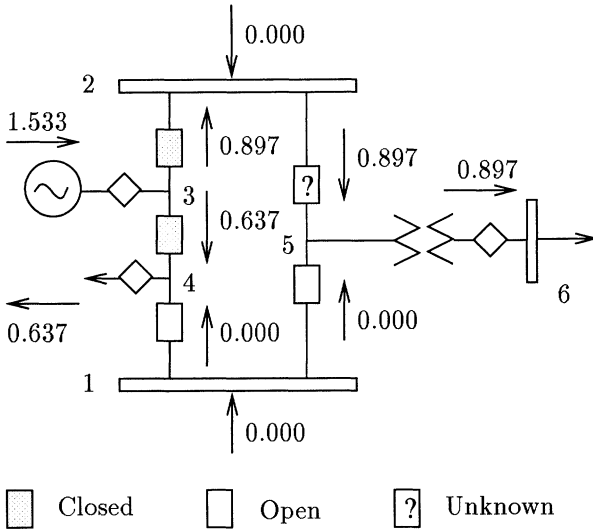


Figure 6.6. Active power flow distribution for the network shown in Fig. 6.5 (Example 6.1).

.00316 p.u. In the measurement model, we have assumed that these power flows are equal to zero. Dividing flows P_{14} and P_{15} by the standard deviation σ_{pseudo} yields

$$\tilde{P}_{14} = 0.0161 ; \quad \tilde{P}_{15} = 0.0161$$

This suggests that the zero flow hypothesis (open breaker) is probably correct, since the estimated flows are less than two percent of the assigned standard deviation value. On the other hand, the active power flow in the breaker 2 – 5 is $P_{25} = 0.897$ p.u., thus indicating that the breaker is actually open (See Chap. 9 for a more formal discussion about hypothesis testing.)

Figure 6.6 shows the estimated active power flows in the network. This results have been approximated considering a 10^{-3} p.u. tolerance. Figure 6.7 presents the corresponding bus/branch model.

6.2.2 Local State Estimation: Analysis of Good Data and Bad Data

In areas where status changes have occurred since the previous time, or where status information is doubtful for any other reason, local state estimation using the bus-section/switching-device model is performed. This area can be a bay, a substation or a bus section group, depending on the implementation and the execution mode of the NTP. Local state estimation includes observability analysis and bad data processing. Since the network of interest is modeled at the bus-section/switching-device level, the state estimator is able to process

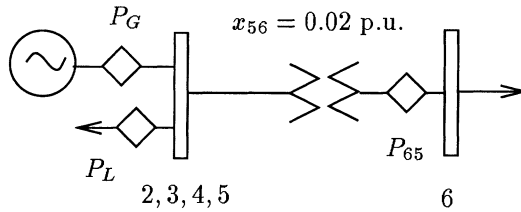


Figure 6.7. Bus/branch model for the network shown in Fig. 6.6.

both bad analog and bad status data, at least as long as local redundancy permits.

Local state estimation can also be viewed as a systematic approach to the performance of plausibility tests. Local state estimation implicitly verifies whether KCL and KVL are satisfied. This includes both network flow and voltage consistency checks. Moreover, other more sophisticated checks are also performed, such as the automatic checking of the algebraic sum of all the currents entering a given area (say, a bus section group or an entire substation). Even when redundancy is insufficient for the identification of bad data, if any is present, and this redundant data is consistent (i.e., there are no significant residuals), something can be said about the good data. This idea has been used to assign grades to measurements according to the results of the local state estimator. Of course, the quality of the processed data cannot be guaranteed, even in cases where they are entirely consistent, it is possible for them to be consistently incorrect. Despite this drawback, the results of the analysis of good-bad data can be useful for directing the actual bad data processing for the network as a whole (see combinatorial search procedures in Chap. 9).

6.3 NETWORK REDUCTION

This section provides an analyses of the relationship between network topology processing and network reduction. As indicated above, the NTP performs the following three principal tasks: (1) it eliminates selected switching devices from the model (although it is possible for part of the network to be retained in the bus-section/switching-device level); (2) it merges bus sections of selected bus section groups into network buses; and (3) it assigns metering devices to the components of the reduced model. These tasks can be interpreted by the reduction of the Jacobian matrix of the WLS state estimator, as Gauss elimination transforms the Jacobian matrix representing the bus-section/switching-device model into one representing the bus/branch model.

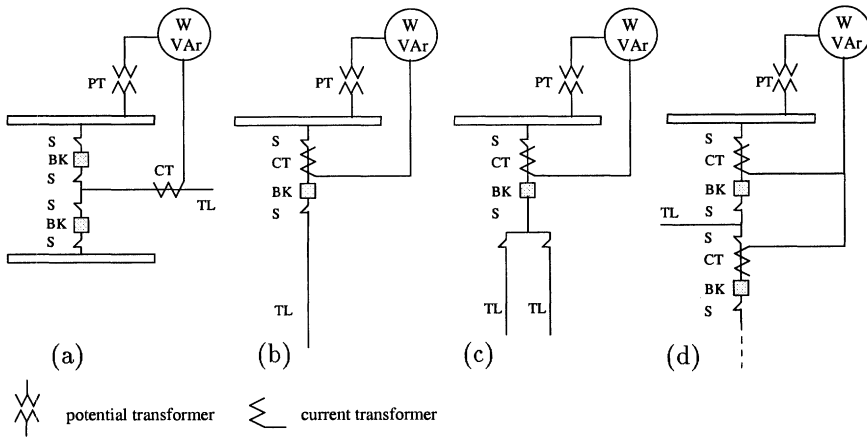


Figure 6.8. Examples of MW-MVAr meter arrangements.

6.3.1 Meter Arrangements

Although in conventional state estimators, meters are normally associated with buses and branches, their location in practice is given in terms of bus sections and switching devices, as illustrated in Fig. 6.8. If the NTP is used to reduce the network in Fig. 6.8 (a), all the bus sections will be merged into a single bus, with the flow meter naturally being assigned to the transmission line TL. The situation in Fig. 6.8 (b) requires a minor modification, in which the NTP will transfer the meter (located in a zero impedance branch connecting the breaker to the bus) to the adjacent transmission line TL. In the situation in Fig. 6.8 (c), if the NTP eliminates the switching devices (one breaker and three switches), then the power flow meter will be assigned as reading the sum of the flows in the two transmission lines (i.e., it becomes a group flow measurement in the bus-branch model). Figure 6.8 (d), represents a different situation, with the sum of the flows in the two breakers showing the power flow in a single line (i.e., a group flow measurement at the physical level) such that the reduction of the network to a bus/branch model by the NTP will lead to the change of the meter to indicate the power flow in the transmission line TL.

Figure 6.9 summarizes a few practical rules that can be used to reduce a network model from the bus-section/switching-device level to the bus/branch level. The next section will show how these topological rules can automatically be taken into account by applying Gauss elimination to the Jacobian matrix of a WLS state estimator (in fact, only the structure of the Jacobian matrix is needed, although the the numerical Jacobian matrix is used for clarity).

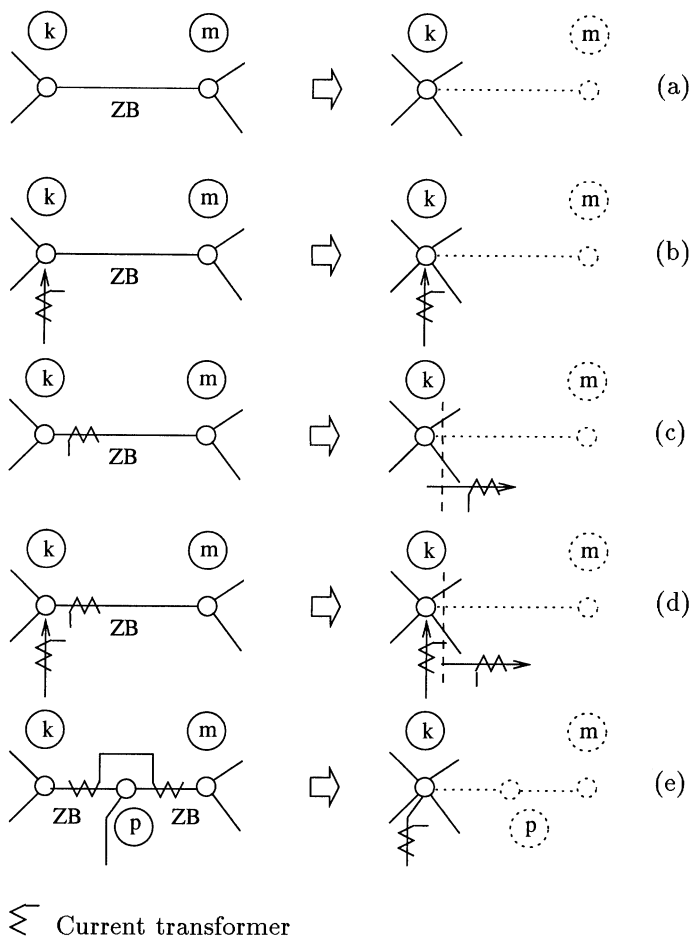


Figure 6.9. Effect of network topological reduction on measurement assignment. Dotted lines represent the eliminated part of the network. ZB stands for a zero impedance branch (e.g., a closed breaker).

6.3.2 NTP and Jacobian Reduction

Consider the elimination of a ZB (zero impedance branch) from a network as shown in Fig. 6.9 (a). For simplicity, consider also that a Jacobian matrix of a dc state estimation model is used. Two state variables and two pseudo-measurements are needed to model the closed breaker $k - m$ (denoted by ZB) in the original model. The two states are the voltage angle at the eliminated node (θ_m) and a power flow representing the state of the closed breaker (P_{km}). The two pseudo measurements consist of a zero injection measurement ($P_m = 0$)

and an angle difference measurement ($\theta_{km} = 0$). Eliminating the breaker from the model, removes the two states and the two pseudo measurements from the measurement dc model. This operation is represented by the Gauss elimination process as shown in Fig. 6.10.

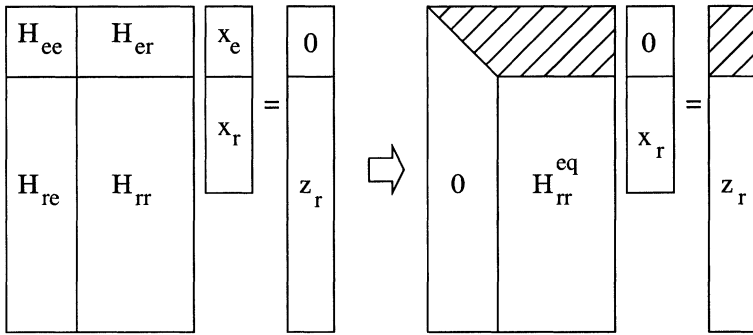


Figure 6.10. Elimination of a closed breaker (a ZB) from a bus-section/switching-device model by applying Gauss elimination to the Jacobian matrix (state variables $\mathbf{x}_e^T = [\theta_k; P_{km}]$ and pseudo-measurements $\mathbf{z}_e^T = [P_k; \theta_{km}] = [0; 0]$).

Example 6.2

Consider the network shown in Fig. 6.11 (a). For the sake of simplicity assume that the reactance of line 2-3 is equal to 1 p.u., the Jacobian matrix of the dc state estimator is given by:

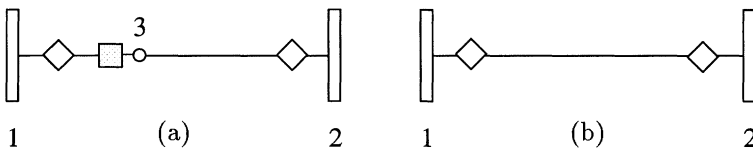


Figure 6.11. (a) network with a explicitly modeled breaker, with flow meter 1-3 located on a zero impedance branch connecting the breaker to Bus 1. (b) Reduced network model obtained by the elimination of the breaker from the model.

$$\begin{pmatrix} -1 & 0 & 1 & 0 \\ 1 & -1 & 0 & -1 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} \theta_3 \\ P_{13} \\ \theta_1 \\ \theta_2 \end{pmatrix} = \begin{pmatrix} \theta_{13}^{pseudo} = 0 \\ P_3^{pseudo} = 0 \\ P_{13}^{meas} \\ P_{23}^{meas} \\ \theta_1^{ref} = 0 \end{pmatrix}$$

where $\theta_1, \theta_2, \theta_3$ and P_{13} are state variables, $\theta_{13} = 0$ and $P_3 = 0$ are pseudo measurements related to the explicit modeling of breaker 1-3, P_{13}^{meas} and P_{23}^{meas} are telemetered values, and $\theta_1 = 0$ is the reference angle. Taking element (1, 1) as the pivot, and zeroizing the elements in the first column by Gauss elimination, gives:

$$\begin{pmatrix} -1 & 0 & 1 & 0 \\ 0 & -1 & 1 & -1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} \theta_3 \\ P_{13} \\ \theta_1 \\ \theta_2 \end{pmatrix} = \begin{pmatrix} \theta_{13}^{pseudo} = 0 \\ P_3^{pseudo} = 0 \\ P_{13}^{meas} \\ P_{21}^{eq} \\ \theta_1^{ref} = 0 \end{pmatrix}$$

Now, taking element (2, 2) as the second pivot, and zeroizing the elements in the second column below the pivot element, gives:

$$\begin{pmatrix} -1 & 0 & 1 & 0 \\ 0 & -1 & 1 & -1 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} \theta_3 \\ P_{13} \\ \theta_1 \\ \theta_2 \end{pmatrix} = \begin{pmatrix} \theta_{13}^{pseudo} = 0 \\ P_3^{pseudo} = 0 \\ P_{12}^{eq} \\ P_{21}^{eq} \\ \theta_1^{ref} = 0 \end{pmatrix}$$

The final reduced model corresponding to the network in Fig. 6.11 (b) is

$$\begin{pmatrix} 1 & -1 \\ -1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix} = \begin{pmatrix} P_{12}^{eq} \\ P_{21}^{eq} \\ \theta_1^{ref} \end{pmatrix}$$

Remarks: This is an exact model provided that the assumption made about the status of the eliminated breaker is correct. Notice that the independent vector representing the measurements of the reduced model (\mathbf{z}_r) is not affected by the reduction process since the pseudo-measurements eliminated are equal to zero.

6.4 HISTORICAL NOTES AND REFERENCES

The need for a network topology processor in real-time network modeling was first recognized by Sasson, Ehrmann, Lynch, and van Slyck [1973] and Dy Liacco, Ramarao, and Weiner [1973]. A matrix formulation for the determination of network topology was later suggested by Goderya, Metwally, and Mansour [1980], and a depth-first algorithm proposed for determining network connectivity (Bertran and Corbella [1982]), with further developments presented by Gahagan, Hunt, and Bose [1986]. A LP based data validation approach based on a bus-section/switching-device substation model was proposed by Irving and Sterling [1982, 1983]. The concept tracking topology processor was introduced by Prais and Bose [1988] and further developed by Yehsakul and Dabbaghchi [1995], who introduced the concept of network connectivity local

update. A comprehensive set of rules for consistency checking and plausibility testing was introduced by Bonanomi and Gramberg [1983] and Singh and Glavitsch [1991]. The use of conventional WLS state estimators in modeling short circuit branches (Monticelli and Garcia [1991]), and such an estimator using networks modeled at the bus-section/switching-device level (Monticelli [1993]) have also been proposed, as well as a generalized state estimator with a network topology processor capable of performing both local state estimation and the pre-processing of bad analog and bad status data (Alsaç, Vempati, Stott, and Monticelli [1998]). A review of the state of the art in real-time modeling and a discussion of network topology processing can be found in Bose and Clements [1987].

6.5 PROBLEMS

- 1. Consider the substation represented in Fig. 6.1. Update the bus/branch model in Fig. 6.2 when switches S51 and S52 are open.
- 2. Consider the network shown in Fig. 6.12: (i) form the Jacobian matrix for the bus-section/switching-device model given in Part (a) of the figure; (ii) reduce the Jacobian matrix by Gauss elimination to obtain the models shown in Parts (b) and (c) of the figure. Use a dc state estimation model, and, for simplicity, consider that the line reactances and the measurement variances are equal to 1 p.u.

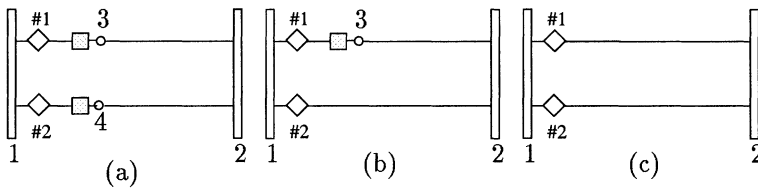


Figure 6.12. (a) Network with two explicitly modeled breakers and two flow meters located on the zero impedance branches connecting the breakers to Bus 1. Part (b) shows the reduced network model obtained by the elimination of Breaker 1 – 4. Part (c) shows the final bus/branch model obtained by the elimination of both breakers.

- 3. Repeat Prob. 2 for the network represented in Fig. 6.13. What should be done to retain the flow meter 2-4 in the reduced model?

References

Alsaç, O., Vempati, N., Stott, B., and Monticelli, A., “Generalized state estimation”, *IEEE Trans. on Power Systems*, Vol. 13, No. 3, pp. 1069-1075, Aug. 1998.

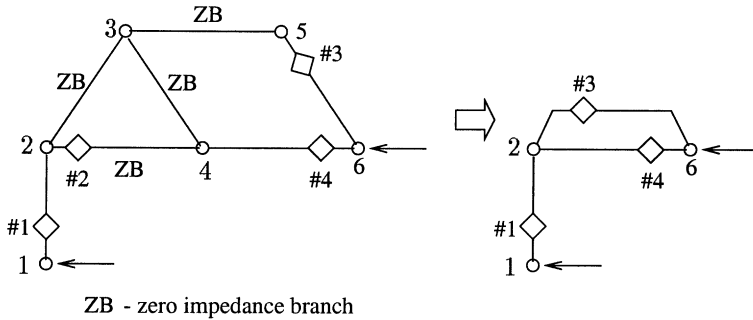


Figure 6.13. Network models for Prob. 3. (a) Unreduced model. (b) Reduced model after elimination of node 3, 4, and 5.

- Bertran, M. and Corbella, X, "On the validation and analysis of a new method for power network connectivity determination", *IEEE Trans. Power App. and Syst.*, Vol. 101, pp. 316-324, Feb. 1982.
- Bonanomi, P. and Gramberg, G., "Power system data validation and state calculation by network search techniques", *IEEE Trans. on Power App. and Syst.*, Vol. 102, pp 238-249, Jan. 1983.
- Bose, A. and Clements, K.A., "Real-time modeling of power networks", *IEEE Proc.*, Special Issue on Computers in Power System Operations, Vol. 75, No. 12, pp 1607-1622, Dec. 1987.
- Dy Liacco, T.E., Ramarao, K.A., and Weiner, A.W., "Network status analysis for real-time systems", in 8th PICA Conf. Proc., pp 356-362, June 1973.
- Gahagan, M.J., Hunt, J.R., and Bose, A., "Power system model development for the control center security analysis functions", *IEEE Trans. Power Systems*, Vol. 1, pp 308-313, Aug. 1986.
- Goderya, F., Metwally, A.A., and Mansour, O., "Fast detection and identification of islands in power networks", *IEEE Trans. on Power App. and Syst.*, Vol. 99, No. 1, pp. 217-221, Jan./Feb. 1980.
- Irving, M.R., and Sterling, M.J.H., "Substation data validation", *IEE Proc.*, Vol. 129, pt C, No. 3, pp. 119-122, May 1982.
- Irving, M.R., and Sterling, M.J.H., discussion of the paper by Bonanomi and Gramberg [1982].
- Monticelli, A. and Garcia, A., "Modeling zero impedance branches in power system state estimation", *IEEE Trans. on Power Systems*, Vol. 6, No. 4, pp. 1061-1069, July 1993.
- Monticelli, A., "Modeling circuit breakers in weighted least squares state estimation", *IEEE Trans. on Power Systems*, Vol. 8, No. 3, pp 1143-1149, Aug. 1993.

- Prais, M. and Bose, A., "A topology processor that tracks network modifications over time", IEEE Trans. on Power System, Vol. 3, No. 3, pp 992-998, Aug. 1988.
- Sasson, A.M., Ehrmann, S.T. Lynch, P., and van Slyck, L.S. "Automatic power system network topology determination", IEEE Trans. Power App. Syst., Vol. 92, pp 610-618, March/April 1973.
- Singh, N. and Glavitsch H., "Detection and Identification of Topological Errors in on-line power system analysis", IEEE Trans. on Power Syst., Vol. 6, pp. 324-331, Feb. 1991.
- Yehsakul, P.D. and Dabbaghchi, I., "A topology-based algorithm for tracking network connectivity", IEEE Trans. on Power Systems, Vol. 10, No. 1, pp. 339-364, Feb. 1995.

7 OBSERVABILITY ANALYSIS

The ability to perform state estimation depends on whether sufficient measurements are well distributed throughout the system. If enough measurements are available to make state estimation possible, the network is considered to be observable. If a network is not observable, it is still useful to know which portion has a state which can be estimated (observable islands). In this chapter, three observability analysis methodologies for state estimation are discussed: topological, numerical, and hybrid, and the concept of extended observable islands is presented. Network modeling at the physical level, which is an essential characteristic of generalized state estimation, is also discussed. In this type of model, short-circuit branches and open/closed breakers and switches can be explicitly represented.

7.1 BUS/BRANCH NETWORK MODEL

Bus/branch network models are those models most commonly used in power flow studies. In state estimation, the use of this type of model assumes that the network topology processor has reduced the physical level model to a model formed of buses and branches.

In more conventional implementations, observability analysis is performed prior to state estimation to determine whether enough real-time measurements are available to make state estimation possible and, if not, which part(s) of

the network contain(s) states which can still be estimated, as well as where pseudo-measurements (based on data prediction) can be added to improve observability.

Usually systems are designed to be observable under most operating conditions, although temporary unobservability may still occur due to unanticipated network topology changes or failures in the telecommunication system.

7.1.1 Physical Islands and Observable Islands

When the bus/branch model of a network is adopted, the following definitions apply:

- *Definition 1:* A physical island is a connected part of a network, with branches representing transmission lines, transformers and series capacitors.
- *Definition 2:* An observable island is an island for which all branch flows can be calculated from the available measurements, independent of the values adopted for angular reference.

A network is considered to be observable if all flow in the network can be observed by some sort of indication in the set of measurements. In other words, whenever there is a nonzero flow in the network, at least one of the measurements should read nonzero. This is equivalent to saying that a network is observable if, whenever all measurements are equal to zero implies that all flows are zero. When a network is not observable, it is possible to have all measurements equal to zero, although some nonzero flow is still present in the network (due to the arbitrary values assigned to the reference angles which are added to make the state estimation problem solvable). In such a case, those branches having nonzero flows are considered to be unobservable.

In a linearized dc measurement model $\mathbf{z} = \mathbf{H}\Theta + \mathbf{e}$ for a network modeled at the bus-branch level, the above definition can be formalized as follows: A network is said to be observable if for all Θ such that $\mathbf{H}\Theta = \mathbf{0}$, the power flows in the network are zero. Any state for which $\mathbf{H}\Theta = \mathbf{0}$ and for which there are nonzero power flows in the network is an unobservable state, and a branch carrying nonzero power flow is considered to be an unobservable branch.

Example 7.1:

Figure 7.1(a) shows a system formed by a single physical island, with part (b) showing a system formed by two physical islands.

Example 7.2:

Figure 7.2(a) shows an observable island. Using a dc model, and choosing Bus-1 as the angular reference bus makes it possible to compute the states (voltage angles) of all the network nodes from the measurements available (assuming that all branch reactances are known and have non-zero and finite values).

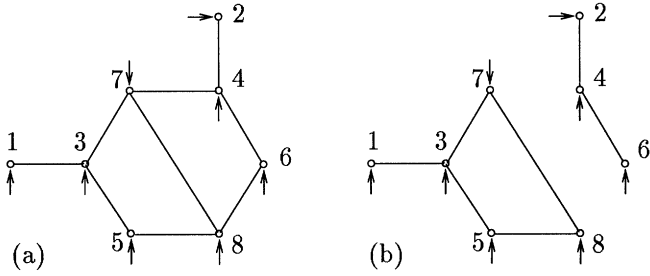


Figure 7.1. Eight bus system. (a) System composed of a single physical island. (b) System formed by two physical islands.

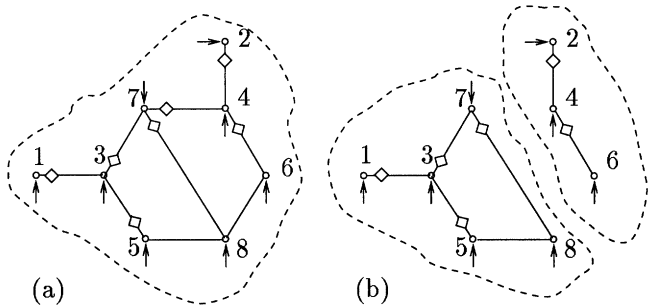


Figure 7.2. Eight bus system. (a) System comprising a single observable island. (b) System comprising two observable islands.

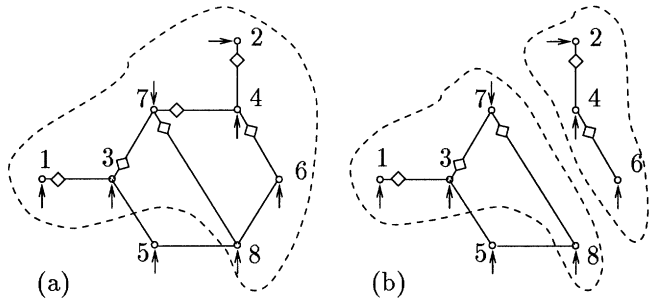


Figure 7.3. Eight bus system with observable branches encircled by dashed lines. (a) Part of the physical island is an observable island. (b) Two physical islands: one partially observable island (on the left); and the other observable (on the right).

Figure 7.2(b) shows two physical islands that are also observable, while Figure 7.3 (a) shows a physical island that contains one observable island (encircled by the dashed line), one unobservable node (Node 5), and two unobservable

branches (3 – 5 and 5 – 8). Figure 7.3(b) depicts a system with a partially observable island on the left, and an observable one, on the right.

7.1.2 Basic Topological Algorithm

A basic topological algorithm will now be presented which can be used to build observable islands in networks modeled at the bus/branch level, with only power flow measurements being performed in these networks.

Algorithm:

1. Initialize nb islands, each comprising a single node, where nb is the number of buses in the network.
2. Select a branch with a flow meter and, if the two terminal nodes of the branch belong to different islands, merge the islands.
3. Stop when all flow measurements have been processed.
4. Declare as observable all the power flows in branches belonging to the same island, and as unobservable the flows in branches connecting different islands.
5. Declare as observable all the power injections at nodes whose incident branch flows are observable; and as unobservable the power injections at boundary nodes, i.e. nodes at which at least one of the incident branch flows is unobservable.

Example 7.3:

Consider the system represented in Fig. 7.3(b). In this case the algorithm yields the following results:

1. Each node is initialized as an island.
2. Select Branch 1 – 3 and merge Nodes 1 and 3 in a single island. Select Branch 2 – 4 and merge Nodes 2 and 4. Select Branch 3 – 7 and form Island {1, 3, 7}. Select Branch 4 – 6 and form Island {2, 4, 6}. Select Branch 7 – 8 and form Island {1, 3, 7, 8}.
3. Now, all the flow measurements have been processed. Yet Node 5 remains as a single bus island.
4. Branches {1 – 3, 3 – 7, 7 – 8}, Branches {2 – 4, 4 – 6} are observable, while Branches 3 – 5 and 5 – 8 are unobservable, since they connect different islands.
5. Injections {1, 7} and {2, 4, 6} are observable, while those at Nodes 3, 5, and 8 are unobservable, since they are boundary nodes.

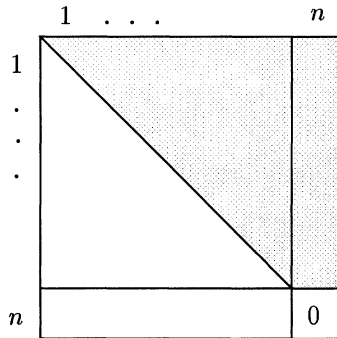
7.1.3 Reference Angles and Triangular Factorization

For an entirely observable network, the triangular factorization of the gain matrix \mathbf{G} leads to a \mathbf{U} -factor ($\mathbf{G} = \mathbf{U}'\mathbf{U}$) in which the final pivot is zero (i.e., the gain matrix is singular). This situation requires the pseudo-measurement of a reference angle; this can be added to any bus in the observable network. Since the reference angle is assigned an arbitrary value, the estimated node angles will vary accordingly. The estimated power flows, however, will be uniquely determined by the measurements. If the reference angle is added beforehand, no zero pivot will occur during triangular factorization. In dealing with partially observable systems, however, it is more convenient to delay the assignment of reference angles.

Theorem:

Assume a dc measurement model $\mathbf{z} = \mathbf{H}\Theta + \mathbf{e}$, where \mathbf{H} is an $m \times n$ matrix, without voltage angle measurements (no reference angles). The following statements are thus equivalent:

- (i) The network is observable.
- (ii) If $\bar{\mathbf{H}}$ is obtained from \mathbf{H} by deleting any column, then $\bar{\mathbf{H}}$ has full rank.
- (iii) Triangular factorization reduces the gain matrix $\mathbf{G} = \mathbf{H}'\mathbf{H}$ into the following form



where the shaded area corresponds to possibly nonzero elements.

Proof:

- (i) \iff (ii). The branch power flows $P_{km} = x_{km}^{-1}(\theta_k - \theta_m)$ are zero if and only if all nodal voltage angles are equal to the reference angle, i.e., $\theta_k = \alpha$, for $k = 1, 2, \dots, n$, where α is a real number. Hence, statement (i) is equivalent to statement (a): $\mathbf{H}\Theta = \mathbf{0}$ iff $\Theta = \alpha\mathbf{1}$, where $\mathbf{1} = (1, 1, \dots, 1)'$. Now it is possible to show that (a) \iff (ii):

1. (a) \implies (ii). Let $\bar{\mathbf{H}}$ be obtained from \mathbf{H} by deleting the k th column \mathbf{h} . Supposing $\bar{\mathbf{H}}\bar{\boldsymbol{\Theta}} = \mathbf{0}$, if $\boldsymbol{\Theta} = (\bar{\boldsymbol{\Theta}}', 0)'$, then $\bar{\boldsymbol{\Theta}} = \mathbf{0}$, hence $\bar{\mathbf{H}}$ has full rank.
 2. (ii) \implies (a). Since the column sum of \mathbf{H} is always zero, $\bar{\mathbf{H}}\mathbf{1} = -\mathbf{h}$. Thus, from the normal equations, $(\bar{\mathbf{H}}'\bar{\mathbf{H}})^{-1}\bar{\mathbf{H}}'\mathbf{h} = -\mathbf{1}$. Now if $\mathbf{H}\boldsymbol{\Theta} = \mathbf{0}$ or $\bar{\mathbf{H}}\bar{\boldsymbol{\Theta}} + \mathbf{h}\theta_k = \mathbf{0}$, then $\bar{\boldsymbol{\Theta}} = -(\bar{\mathbf{H}}'\bar{\mathbf{H}})^{-1}\bar{\mathbf{H}}'\mathbf{h}\theta_k = \mathbf{1}\theta_k$.
- (ii) \iff (iii). For the partition $\mathbf{H} = (\bar{\mathbf{H}}, \mathbf{h})$, the gain matrix \mathbf{G} can be written as follows:

$$\mathbf{G} = \mathbf{H}'\mathbf{H} = \begin{pmatrix} \bar{\mathbf{H}}'\bar{\mathbf{H}} & \bar{\mathbf{H}}'\mathbf{h} \\ \mathbf{h}'\bar{\mathbf{H}} & \mathbf{h}'\mathbf{h} \end{pmatrix}$$

The matrix $\bar{\mathbf{H}}'\bar{\mathbf{H}}$ is nonsingular if and only if triangular factorization reduces it to a triangular matrix. Applying block Gauss elimination to the partitioned gain matrix above transforms the (n, n) th element $\mathbf{h}'\mathbf{h}$ into $\mathbf{h}'\mathbf{h} - \mathbf{h}'\bar{\mathbf{H}}(\bar{\mathbf{H}}'\bar{\mathbf{H}})^{-1}\bar{\mathbf{H}}'\mathbf{h} = \mathbf{h}'\mathbf{h} + \mathbf{h}'\bar{\mathbf{H}}\mathbf{1} = \mathbf{h}'\mathbf{h} - \mathbf{h}'\mathbf{h} = 0$

Remarks:

- The matrix $\bar{\mathbf{H}}$ has full rank if and only if the gain matrix $\bar{\mathbf{H}}'\bar{\mathbf{H}}$ is nonsingular. This is the condition required if the state estimator is to have a unique solution. Thus this theorem implies that a network is observable if and only if the state estimation problem can be solved with a unique solution.
- When voltage angle measurements are given, statement (ii) will be replaced by: \mathbf{H} has full rank.
- Note that, provided node ordering is maintained (the columns of \mathbf{H} have the same order as the columns of \mathbf{G}), the \mathbf{U} -factor that results from the orthogonal factorization of the Jacobian matrix, $\mathbf{H} = \mathbf{Q}\mathbf{U}$, is the same one obtained by the factorization of the gain matrix, $\mathbf{H}'\mathbf{H} = \mathbf{U}'\mathbf{U}$; hence the previous properties can be easily extended to orthogonal methods.

Example 7.4:

Consider the system presented in Fig. 7.2(a); a linear power flow model (dc model) is assumed, with all branch reactances arbitrarily set at one. Measurement weighting factors are also set at one. As indicated below, the triangular factorization of the gain matrix yields a \mathbf{U} -factor for which the final pivot is zero. The gain matrix \mathbf{G} and the corresponding triangular factor \mathbf{U} are as follows ($\mathbf{G} = \mathbf{U}'\mathbf{U}$):

$$\mathbf{H} = \begin{matrix} P_{13} \\ P_{24} \\ P_{37} \\ P_{46} \\ P_{53} \\ P_{74} \\ P_{78} \end{matrix} \begin{pmatrix} \theta_1 & \theta_2 & \theta_6 & \theta_5 & \theta_8 & \theta_4 & \theta_3 & \theta_7 \\ 1 & & & & & & -1 & \\ & 1 & & & & -1 & & \\ & & -1 & & & 1 & 1 & -1 \\ & & & 1 & & & -1 & \\ & & & & -1 & & & 1 \\ & & & & & -1 & & 1 \end{pmatrix}$$

$$\mathbf{G} = \begin{matrix} \theta_1 \\ \theta_2 \\ \theta_6 \\ \theta_5 \\ \theta_8 \\ \theta_4 \\ \theta_3 \\ \theta_7 \end{matrix} \begin{pmatrix} \theta_1 & \theta_2 & \theta_6 & \theta_5 & \theta_8 & \theta_4 & \theta_3 & \theta_7 \\ 1 & & & & & & -1 & \\ & 1 & & & & -1 & & \\ & & 1 & & & -1 & & \\ & & & 1 & & & -1 & \\ & & & & 1 & & & -1 \\ & -1 & -1 & & & 3 & & -1 \\ -1 & & & -1 & & & 3 & -1 \\ & & & & -1 & -1 & -1 & 3 \end{pmatrix}$$

$$\mathbf{U} = \begin{matrix} \theta_1 \\ \theta_2 \\ \theta_6 \\ \theta_5 \\ \theta_8 \\ \theta_4 \\ \theta_3 \\ \theta_7 \end{matrix} \begin{pmatrix} \theta_1 & \theta_2 & \theta_6 & \theta_5 & \theta_8 & \theta_4 & \theta_3 & \theta_7 \\ 1 & & & & & & -1 & \\ & 1 & & & & -1 & & \\ & & 1 & & & -1 & & \\ & & & 1 & & & -1 & \\ & & & & 1 & & & -1 \\ & & & & & 1 & & -1 \\ & & & & & & 1 & -1 \\ & & & & & & & 0 \end{pmatrix}$$

Example 7.5:

Consider the system presented in Fig. 7.2(b) (dc model), with all the branch reactances and weighting factors set at one. In this case the triangular factor \mathbf{U} has two zero pivots, which will require two reference angles, one per island. The gain matrix \mathbf{G} and the corresponding triangular factor \mathbf{U} are as follows ($\mathbf{G} = \mathbf{U}'\mathbf{U}$):

$$\mathbf{G} = \begin{matrix} \theta_1 \\ \theta_2 \\ \theta_6 \\ \theta_5 \\ \theta_8 \\ \theta_4 \\ \theta_3 \\ \theta_7 \end{matrix} \begin{pmatrix} \theta_1 & \theta_2 & \theta_6 & \theta_5 & \theta_8 & \theta_4 & \theta_3 & \theta_7 \\ 1 & & & & & & -1 & \\ & 1 & & & & -1 & & \\ & & 1 & & & -1 & & \\ & & & 1 & & & -1 & \\ & & & & 1 & & & -1 \\ & -1 & -1 & & & 2 & & \\ -1 & & & -1 & & & 3 & -1 \\ & & & & -1 & & -1 & 2 \end{pmatrix}$$

$$\mathbf{U} = \begin{matrix} & \theta_1 & \theta_2 & \theta_6 & \theta_5 & \theta_8 & \theta_4 & \theta_3 & \theta_7 \\ \theta_1 & \left(\begin{array}{cccccccc} 1 & & & & & & & -1 & \\ & 1 & & & & & -1 & & \\ & & 1 & & & & -1 & & \\ \theta_5 & & & 1 & & & -1 & & \\ \theta_8 & & & & 1 & & & & -1 \\ \theta_4 & & & & & 1 & & & \\ \theta_3 & & & & & & 0 & & \\ \theta_7 & & & & & & & 1 & -1 \\ & & & & & & & & 0 \end{array} \right) \end{matrix}$$

7.1.4 Unobservable Networks

When a given network is not observable, triangular factorization with complete pivoting reduces $\mathbf{H}'\mathbf{H}$ to the following form:

$$\begin{array}{|c|c|} \hline \text{shaded triangle} & \\ \hline \mathbf{0} & \\ \hline \end{array} = \begin{array}{|c|} \hline \Theta_a \\ \hline \Theta_b \\ \hline \end{array} = \begin{array}{|c|} \hline \mathbf{0} \\ \hline \end{array} \tag{7.1}$$

For any arbitrary Θ_b , for example $\Theta_b = (0, 1, 2, \dots)'$, solving the upper half of this equation yields Θ_a ; then (Θ_a, Θ_b) is an unobservable state.

An alternative way of obtaining the same unobservable state (Θ_a, Θ_b) is by (i) replacing the diagonal elements of the lower right matrix with 1's, (ii) replacing the corresponding right-hand side of Eq. (7.1) with arbitrary values, for example, $\Theta_b = (0, 1, 2, \dots)'$, and (iii) solving the resulting equation.

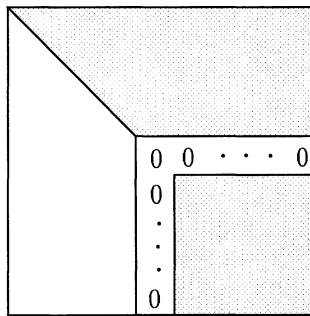
$$\begin{array}{|c|c|} \hline \text{shaded triangle} & \\ \hline \begin{matrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{matrix} & \\ \hline \end{array} = \begin{array}{|c|} \hline \Theta_a \\ \hline \Theta_b \\ \hline \end{array} = \begin{array}{|c|} \hline \mathbf{0} \\ \hline 0 \\ \hline 1 \\ \hline \vdots \\ \hline \vdots \\ \hline \end{array} \tag{7.2}$$

Note that Eq. (7.2) is identical to a normal equation with the pseudo-measurement of the voltage angles at buses corresponding to Θ_b present and all other measurements set at zero.

Triangular factorization with complete pivoting involves permutations of rows and columns of the gain matrix $\mathbf{H}'\mathbf{H}$ and the corresponding reordering of the vector Θ . Since the solution of a large system by triangular factorization involves the ordering of the matrix based on sparsity considerations, it is preferable not to have two different orders. The theorem below shows that the reordering for complete pivoting indicated in Eqs. (7.1) and (7.2) is not really necessary to obtain an unobservable state.

Theorem:

In the triangular factorization of the gain matrix $\mathbf{G} = \mathbf{H}'\mathbf{H}$, if a zero pivot is encountered, then the remaining row and column are all zeros, i.e., \mathbf{G} is reduced to the form:

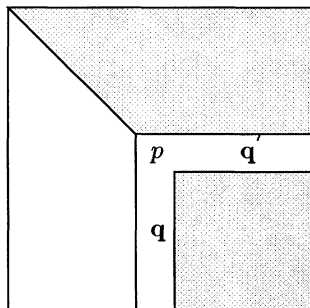


Proof:

Consider the following partition of the Jacobian matrix: $\mathbf{H} = (\mathbf{H}_1\mathbf{h}_2\mathbf{H}_3)$. The gain matrix can be written as follows:

$$\mathbf{G} = \mathbf{H}'\mathbf{H} = \begin{pmatrix} \mathbf{H}'_1\mathbf{H}_1 & \mathbf{H}'_1\mathbf{h}_2 & \mathbf{H}'_1\mathbf{H}_3 \\ \mathbf{h}'_2\mathbf{H}_1 & \mathbf{h}'_2\mathbf{h}_2 & \mathbf{h}'_2\mathbf{H}_3 \\ \mathbf{H}'_3\mathbf{H}_1 & \mathbf{H}'_3\mathbf{h}_2 & \mathbf{H}'_3\mathbf{H}_3 \end{pmatrix}$$

The triangular factorization reduces \mathbf{G} to



where

$$\begin{aligned} p &= \mathbf{h}'_2 \mathbf{h}_2 - \mathbf{h}'_2 \mathbf{H}_1 (\mathbf{H}'_1 \mathbf{H}_1)^{-1} \mathbf{H}'_1 \mathbf{h}_2 \\ \mathbf{q}' &= \mathbf{h}'_2 \mathbf{H}_3 - \mathbf{h}'_2 \mathbf{H}_1 (\mathbf{H}'_1 \mathbf{H}_1)^{-1} \mathbf{H}'_1 \mathbf{H}_3 \end{aligned}$$

In this situation the gain submatrix $\mathbf{H}'_1 \mathbf{H}_1$ is nonsingular, and a zero pivot $p = 0$ was encountered, i.e.,

$$0 = (\mathbf{h}'_2 - \mathbf{h}'_2 \mathbf{H}_1 (\mathbf{H}'_1 \mathbf{H}_1)^{-1} \mathbf{H}'_1) \mathbf{h}_2 = \mathbf{r}'_2 \mathbf{h}_2$$

where \mathbf{r}_2 is the “residual” vector $\mathbf{r}_2 = \mathbf{h}_2 - \mathbf{H}\Psi_2$ and Ψ_2 is a solution to the normal equation $(\mathbf{H}'_1 \mathbf{H}_1)^{-1} \mathbf{H}'_1 \Psi_2 = \mathbf{H}'_1 \mathbf{h}_2$. Since \mathbf{r}_2 is orthogonal to the “measurement” vector \mathbf{h}_2 (i.e., $\mathbf{r}'_2 \mathbf{h}_2 = 0$), the vector of residuals \mathbf{r}_2 is a null vector (see both Fig 2.6(a) and Prob. 6 in Chap. 2).

Remarks: The question of whether the arbitrary values assigned for the reference-angle pseudo-measurements would affect state estimation results may arise. The answer is no, since the residuals corresponding to the actual measurements are not affected by the values of the reference-angles. This can be shown by considering all the actual measurements as zero, assigning arbitrary values the pseudo-measurements, and computing the resulting residuals by solving the normal equation (see Prob. 7 at the end of the chapter).

Example 7.6:

Consider the system presented in Fig. 7.3(b) for a dc model with all branch reactances and weighting factors set at one. The triangular factor \mathbf{U} here has three zero pivots, which requires three reference angles (one for the observable island $\{1, 3, 7, 8\}$, one for Island $\{2, 4, 6\}$, and the third for the single node island $\{5\}$). Since the reference angles are assigned arbitrary values, the flows in Branches 3 – 5 and 5 – 8 are undetermined, and hence are unobservable. The gain matrix \mathbf{G} and the triangular factor \mathbf{U} are as follows ($\mathbf{G} = \mathbf{U}'\mathbf{U}$):

$$\mathbf{G} = \begin{matrix} & \theta_1 & \theta_2 & \theta_6 & \theta_5 & \theta_8 & \theta_4 & \theta_3 & \theta_7 \\ \theta_1 & \left(\begin{array}{cccccccc} 1 & & & & & & & -1 & \\ & 1 & & & & & -1 & & \\ & & 1 & & & & -1 & & \\ & & & 0 & & & & & \\ & & & & 1 & & & & -1 \\ \theta_4 & & -1 & -1 & & 2 & & & \\ \theta_3 & -1 & & & & & 2 & -1 & \\ \theta_7 & & & & -1 & & -1 & 2 & \end{array} \right) \end{matrix}$$

$$\mathbf{U} = \begin{matrix} & \theta_1 & \theta_2 & \theta_6 & \theta_5 & \theta_8 & \theta_4 & \theta_3 & \theta_7 \\ \theta_1 & \left(\begin{array}{cccccccc} 1 & & & & & & & -1 & \\ & 1 & & & & & -1 & & \\ & & 1 & & & & -1 & & \\ \theta_5 & & & 0 & & & & & \\ \theta_8 & & & & 1 & & & & -1 \\ \theta_4 & & & & & 0 & & & \\ \theta_3 & & & & & & & 1 & -1 \\ \theta_7 & & & & & & & & 0 \end{array} \right) \end{matrix}$$

7.1.5 Basic Numerical Algorithm

The basic topological algorithm discussed above corresponds to a numerical algorithm which will be presented here. As with the topological method, this algorithm allows the determination of observable islands in networks modeled at the bus/branch level in which only power flow measurements are made.

Algorithm:

1. Form gain matrix \mathbf{G}_θ .
2. Perform triangular factorization of \mathbf{G}_θ , introducing θ pseudo-measurements whenever a zero pivot is encountered. If only one zero pivot occurs, stop.
3. Solve the dc state estimator equation for θ , considering all measured values equal to zero except for the θ pseudo-measurements, which are assigned the arbitrary values $\theta_0, \theta_1, \theta_2$, and so on.
4. Evaluate branch flows and remove all branches with non-zero flows.
5. Form islands by grouping the nodes connected by branches with zero flows.

Remarks:

- In Step 1 of the algorithm, the matrix \mathbf{G} can be built using either the original network parameters (line reactances) and measurement weightings, or using reactances and weightings arbitrarily set to random values in the vicinity of unity (for example, in the range from 0.9 to 1.1). In the first case the algorithm will reproduce the conditions that will prevail during state estimation solution; in the second case the algorithm will mimic the topological algorithm. The random values are introduced to avoid numerical coincidences which can affect the result of the analysis (This will be discussed later in this chapter.) When the original parameter and weighting values are used, the numerical algorithm is said to be in the numerical mode ; when the parameter and weighting values are arbitrarily set, it is said that the numerical algorithm is in the topological mode .

- In Step 3 of the algorithm, arbitrary numbers are assigned as angular references for observable islands. Since these values are arbitrary, they will generally produce non-zero flows through the branches connecting any two observable islands, although it may happen that an unobservable branch will show a zero or negligible flow, for a given choice of angles. Such a problem can be avoided by the adoption of any of the three strategies: (a) using random numbers rather than integers (as in the case of parameters and weights discussed above, these random numbers can be generated in the vicinity of the integer sequence given in Step 3 of the algorithm); (b) performing various generations of random reference angles until the number of non-zero flows stabilizes before proceeding to Step 4; (c) cycling through the algorithm to make sure no unobservable flow survives. In most cases, the strategy (a) will be sufficient; the other two strategies will serve as backups.
- This numerical algorithm can easily be extended for use in orthogonal state estimators. As noted above, the \mathbf{U} -factor yielded by orthogonal factorization is the same as that obtained by Gauss elimination (provided the column order of matrix \mathbf{H} is the same as the pivot order used in the \mathbf{G} matrix).
- The power flows used in Step 4 do not have to be computed explicitly, as long as the voltage angles at the terminal nodes of each branch are checked.

Example 7.7:

Consider the system represented in Fig. 7.3(b). In this case the numerical algorithm yields the following:

1. The gain matrix \mathbf{G}_θ is the same as in Example 7.6.
2. The triangular factor \mathbf{U} is the same as in Example 7.6. Three voltage angle measurements are assigned during factorization: θ_4^{ref} , θ_5^{ref} , and θ_7^{ref} ,
3. The solution of the dc state estimator considers all measured values to be zero, except for the arbitrarily assigned θ pseudo-measurements.
4. Branches 3 – 5 and 5 – 8 will present non-zero power flows (unobservable branches). All the other (observable) branches will have zero flows, regardless of the values of the reference angles .
5. Two observable islands are formed: $\{1, 3, 7, 8\}$ and $\{2, 4, 6\}$.

7.1.6 Redundant Measurements

The placement of an additional power flow (or power injection) measurement to a branch (or to a node) with all the adjacent branches belonging to the same observable island, does not expand this observable island. Such measurements

are redundant and have no impact on network observability, although they are crucial for the detection and the identification of bad data.

Example 7.8:

Consider the addition of measurements P_{73}^{meas} and P_4^{meas} to the system formed by two physical islands represented in Fig. 7.3(b). Both measurements are redundant and thus have no impact on the assignment of reference angles during the triangular factorization of the gain matrix \mathbf{G} . The observable islands are given in Fig. 7.4 and the corresponding \mathbf{U} -factor is ($\mathbf{G} = \mathbf{U}'\mathbf{U}$) as follows:

$$\mathbf{G} = \begin{matrix} & \theta_1 & \theta_2 & \theta_6 & \theta_5 & \theta_8 & \theta_4 & \theta_3 & \theta_7 \\ \theta_1 & \left(\begin{array}{cccccccc} 1 & & & & & & & -1 & \\ & 2 & 1 & & & & -3 & & \\ & & 1 & 2 & & & -3 & & \\ \theta_5 & & & & 0 & & & & \\ \theta_8 & & & & & 1 & & & -1 \\ \theta_4 & & -3 & -3 & & & 6 & & \\ \theta_3 & -1 & & & & & & 3 & -2 \\ \theta_7 & & & & & -1 & & -2 & 3 \end{array} \right) \end{matrix}$$

$$\mathbf{U} = \begin{matrix} & \theta_1 & \theta_2 & \theta_6 & \theta_5 & \theta_8 & \theta_4 & \theta_3 & \theta_7 \\ \theta_1 & \left(\begin{array}{cccccccc} 1.00 & & & & & & & & -1.00 \\ & 1.41 & 0.71 & & & & -2.12 & & \\ & & 1.22 & & & & -1.22 & & \\ \theta_5 & & & 0.00 & & & & & \\ \theta_8 & & & & 1.00 & & & & -1.00 \\ \theta_4 & & & & & 1.00 & & & \\ \theta_3 & & & & & & 0.00 & & 1.41 & -1.41 \\ \theta_7 & & & & & & & & & 0.00 \end{array} \right) \end{matrix}$$

Note that the number of zero pivots remains the same.

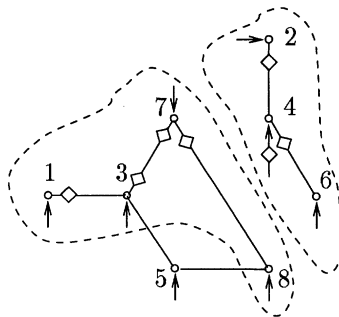


Figure 7.4. Network in Fig. 7.3(b) with additional measurements (P_{73}^{meas} and P_4^{meas})

7.1.7 Transformable Injection Measurements

If all except one of the branches incident to a node with an injection measurement are observable, the unobservable branch is made observable by the injection measurement. There are cases, however, in which the application of the transformation principle is not obvious. This normally happens when more than one unobservable branch is connected to a node with injection measurement. Such a situation may lead to the need to undertake a search procedure (using tentative measurement/branch assignments and backtracking) in order to find an assignment that yields a spanning tree. If the system is actually observable, this will eventually happen, but if it is unobservable, numerous trials will have to be made before reaching the conclusion that no such tree will be formed and that the system is actually unobservable. The number of attempts may have to be very large, however, as this procedure involves a combinatorial problem (A clever implementation of the topological method would try to minimize the number of such trials in such way that most attempts would be made only implicitly.) The combinatorial nature of the assignment problem is further illustrated in the examples that follow.

Example 7.9:

Consider the situation given in Fig. 7.5(a). First the flow measurements are assigned to the respective branches. Next the injection measurements 3, 4, and 7 are assigned, as indicated. This yields the spanning tree 1 – 3, 3 – 5, 3 – 7, 7 – 4, 7 – 8, 4 – 2, 4 – 6, i.e., the system is observable. Note that in this case there is only one possible measurement/branch assignment. In the situation in Fig. 7.5(b), however, there are various possibilities. Even though the system is observable, it is impossible to determine this when the measurements are assigned as in the figure, since two observable islands are formed (see Prob. 1 at the end of the chapter).

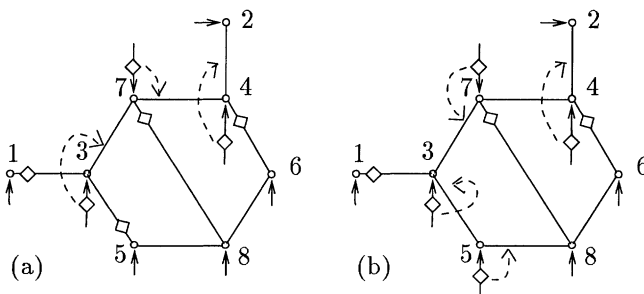


Figure 7.5. Observable eight bus network. (a) Injection measurement assignment yielding a spanning tree. (b) Tentative injection measurement assignment not yielding a spanning tree, even though the system is observable.

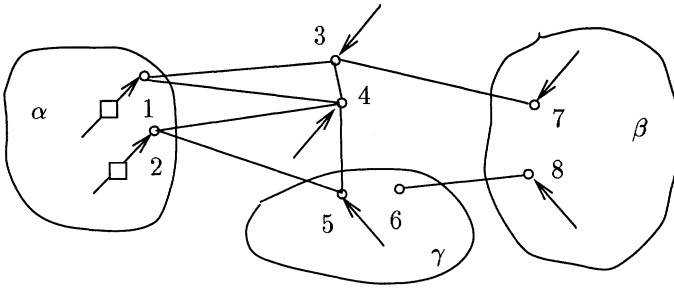


Figure 7.6. Three observable islands connected by an unobservable network.

7.1.8 Observable Islands

Consider an unobservable state Θ for the example given in Fig. 7.6 obtained from the solution of $\mathbf{G}\Theta = \mathbf{0}$ by using angle pseudo-measurements as suggested in Eq. (7.2). The state Θ can be arranged so that the components having identical values are grouped together i.e., $\Theta = (\Theta_\alpha, \Theta_\beta, \Theta_\gamma)'$ where $\Theta_\alpha = (\theta_\alpha, \theta_\alpha, \dots)'$, etc. The subnetwork α consists of the nodes in Θ_α together with the branches connecting them. Similarly for subnetworks β and γ . Line flow measurements in the same subnetwork, and injection measurements for which all the nodes connected to the injection node belong to the same subnetwork, are also grouped together. Note that it is impossible to have line flow measurements for lines connecting different subnetworks, for if this were possible the θ values for these two components would have to be the same. Therefore the remaining measurements are those injections for which the nodes connected to them do not belong to the same subnetwork (Injections 1 and 2 in Fig. 7.6). According to this grouping the matrix \mathbf{H} becomes:

$$\begin{bmatrix}
 \mathbf{H}_\alpha & & & \\
 & \mathbf{H}_\beta & & \\
 & & \mathbf{H}_\gamma & \\
 \mathbf{H}_{\alpha\beta\gamma} & & &
 \end{bmatrix}
 \begin{bmatrix}
 \Theta_\alpha \\
 \Theta_\beta \\
 \Theta_\gamma
 \end{bmatrix}
 =
 \begin{bmatrix}
 \mathbf{0}
 \end{bmatrix}
 \tag{7.3}$$

The rows of \mathbf{H}_α , \mathbf{H}_β , and \mathbf{H}_γ correspond to the subnetworks α , β , and γ in Fig. 7.6, respectively, and the rows of $\mathbf{H}_{\alpha\beta\gamma}$ corresponds to the unobservable network (branches that are not encircled in Fig. 7.6).

Although the network in Fig. 7.6 is unobservable, the subnetworks α , β , and γ are observable. For subnetwork α , for example, $\Theta_\alpha = (\theta_\alpha, \theta_\alpha, \dots)'$ is a solution to $\mathbf{H}_\alpha \Theta_\alpha = \mathbf{0}$ and the corresponding flows are all zero. The same is true for the subnetworks β and γ .

7.1.9 Irrelevant Injection Measurements

At times the tentative assignment of injection measurements to branches does not lead to a spanning tree, even after all possible variations of assignment have been taken into consideration. In such cases the network is considered unobservable, but it is still useful to identify these observable islands for which state can be estimated, as discussed above. If no further measurements are added to the network, injection measurements tentatively assigned, albeit unsuccessfully, are deemed irrelevant. They still may be temporary, since due the addition of more measurements, such as from load/generation prediction, these previously irrelevant measurements may again become relevant, making the entire network observable. This is the case of the injection measurements 1 and 2 in Fig. 7.6, which correspond to the Jacobian submatrix $\mathbf{H}_{\alpha\beta\gamma}$ in Eq. (7.3).

Such irrelevant (discardable) measurements can be handled by both observability algorithms discussed above (topological and numerical). The modified version of the numerical algorithm can be expanded as follows:

Algorithm:

1. Initialization
 - (a) Initialize the measurement set of interest as consisting of all available measurements.
 - (b) Initialize the power network of interest as consisting of all branches incident to at least one measurement.
2. Form gain matrix \mathbf{G}_θ and perform triangular factorization $\mathbf{G}_\theta = \mathbf{U}'\mathbf{U}$.
3. Introduce θ pseudo-measurements whenever a zero pivot is encountered. If only one zero pivot occurs, stop. The system is then entirely observable.
4. Solve the dc state estimator equation for θ by assuming that all the measured values are equal to zero, except for the θ pseudo-measurements, which have been assigned arbitrary values $\theta_0, \theta_1, \theta_2$, and so on.
5. Update system.

- (a) Remove from the power network of interest all branches with non-zero flows.
 - (b) Update the measurement set of interest by removing power injection measurements from buses adjacent to at least one of the branches removed in the previous step.
 - (c) If modifications have been made, update the triangular factor \mathbf{U} and proceed to Step 3.
6. Form islands with nodes connected by branches with zero flows.

Example 7.10:

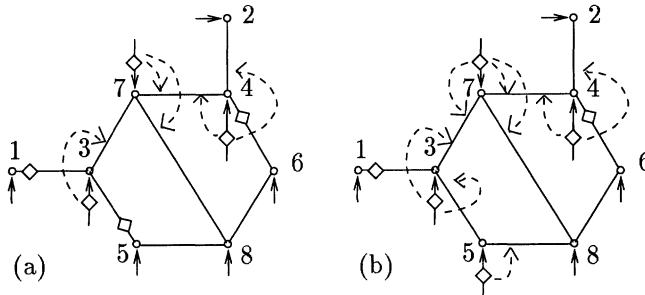


Figure 7.7. Observable islands and possible injection measurement assignments. (a) Injection measurements at Nodes 4 and 7 are irrelevant. (b) Injection measurements at Nodes 3, 4 and 7 are irrelevant.

Consider now the two situations presented in Fig. 7.7, which result from the removal of the measurement 7 – 8 from the networks in Fig. 7.5. In both, the system becomes only partially observable.

In the first (a) the injection measurement at Node 3 is assigned to Branch 3 – 7 and the injection measurement at Node 4 is assigned to Branch 2 – 4. Two different options are now available for assigning measurement 7, but neither of them yields a spanning tree due to the formation of two observable island ($\{1, 3\}$ and $\{4, 6\}$). Under these circumstances, the injection measurements at Nodes 4 and 7 cannot be used to expand the observable islands any further, and they have become irrelevant.

Applying the numerical algorithm to situation (a) gives the following:

1. Initialization

(a) Measurement set of interest: $\{P_{13}, P_{46}, P_{53}, P_3, P_4, P_7\}$.

(b) Network of interest: $\{1-3, 2-4, 3-5, 3-7, 4-6, 4-7, 5-8, 6-8, 7-8\}$.

2. The gain matrix \mathbf{G}_θ and the triangular factor \mathbf{U} are as follows:

$$\mathbf{G} = \begin{matrix} & \theta_1 & \theta_2 & \theta_6 & \theta_5 & \theta_8 & \theta_4 & \theta_3 & \theta_7 \\ \theta_1 & \left(\begin{array}{ccccccc} 2 & & & 1 & & & -4 & 1 \\ & 1 & 1 & & & -3 & & 1 \\ & & 1 & 2 & & -4 & & 1 \\ 1 & & & & 2 & & -4 & 1 \\ & & & & & 1 & 1 & 1 & -3 \\ & & -3 & -4 & & 1 & 11 & 1 & -6 \\ -4 & & & & -4 & 1 & 1 & 12 & -6 \\ 1 & 1 & 1 & 1 & -3 & -6 & -6 & -6 & 11 \end{array} \right) \end{matrix}$$

$$\mathbf{U} = \begin{matrix} & \theta_1 & \theta_2 & \theta_6 & \theta_5 & \theta_8 & \theta_4 & \theta_3 & \theta_7 \\ \theta_1 & \left(\begin{array}{ccccccc} 1.41 & & & 0.71 & & & -2.83 & 0.71 \\ & 1.00 & 1.00 & & & -3.00 & & 1.00 \\ & & & 1.00 & & -1.00 & & \\ \theta_5 & & & & 1.22 & & -1.63 & 0.41 \\ \theta_8 & & & & & 1.00 & 1.00 & 1.00 & -3.00 \\ \theta_4 & & & & & & 0.00 & & \\ \theta_3 & & & & & & & 0.58 & -0.58 \\ \theta_7 & & & & & & & & 0.00 \end{array} \right) \end{matrix}$$

3. Two zero pivots are found: Nodes 4 and 7. Hence θ_4 and θ_7 are assigned arbitrary values, in this case, for simplicity, $\theta_4 = 0.$ and $\theta_7 = 1.0,$ and the corresponding pivots of the \mathbf{U} -factor are set at one.
4. The solution of the dc state estimator equation for $\theta,$ considering all the measured values equal to zero, except θ_4 and $\theta_7,$ yields: $\theta_1 = 1.0, \theta_2 = -1.0, \theta_3 = 1.0, \theta_4 = 0., \theta_5 = 1.0, \theta_6 = 0., \theta_7 = 1.0$ and $\theta_8 = 2.0$
5. Update system.
 - (a) Branches 2 – 4, 4 – 7, 5 – 8, 6 – 8, and 7 – 8, are removed from the network of interest. The updated network is {1 – 3, 3 – 5, 3 – 7, 4 – 6}.
 - (b) Injection measurements adjacent to the removed branches are removed. The updated measurement set is $\{P_{13}, P_{46}, P_{53}, P_3\}.$
 - (c) The reduced system is presented in Fig. 7.8 and the updated \mathbf{U} -factor is

$$\mathbf{U} = \begin{matrix} & \theta_1 & \theta_2 & \theta_6 & \theta_5 & \theta_8 & \theta_4 & \theta_3 & \theta_7 \\ \theta_1 & 1.41 & & & 0.71 & & & -2.83 & 0.71 \\ \theta_2 & & 0.00 & & & & & & \\ \theta_6 & & & 1.00 & & & -1.00 & & \\ \theta_5 & & & & 1.22 & & & -1.63 & 0.41 \\ \theta_8 & & & & & 0.00 & & & \\ \theta_4 & & & & & & 0.00 & & \\ \theta_3 & & & & & & & 0.58 & -0.58 \\ \theta_7 & & & & & & & & 0.00 \end{matrix}$$

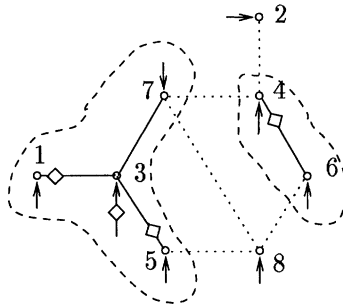


Figure 7.8. Observable islands of the network in Fig. 7-7(a).

3. Four zero pivots appear in the \mathbf{U} -factor: Nodes 2, 8, 4 and 7. Hence θ_2 , θ_8 , θ_4 , and θ_7 are assigned arbitrary values: for example, $\theta_2 = 0.$, $\theta_8 = 1.$, $\theta_4 = 2.$, and $\theta_7 = 3.$ The corresponding pivots of the \mathbf{U} -factor are set at one.
4. The state estimate is as follows: $\theta_1 = 3.0$, $\theta_2 = 0.0$, $\theta_3 = 3.0$, $\theta_4 = 2.0$, $\theta_5 = 3.0$, $\theta_6 = 2.0$, $\theta_7 = 3.0$, and $\theta_8 = 1.0$.
5. All the remaining branches reveals zero power flow readings. No additional branches are removed.
6. Two islands and two isolated nodes are identified, as shown in Fig. 7.8.

Case (b) requires some more work since various options must be considered by the topological algorithm (see Prob. 1 at the end of the chapter). Two islands are eventually formed ($\{1, 3\}$ and $\{4, 6\}$), and all injection measurements are reduced to the status of being irrelevant.

7.1.10 Hybrid Algorithm

The basic topological algorithm with injection measurement transformation is quick and easy when processing (a) flow measurements and (b) injection measurements for which all except one of the incident branches are observable. For

those injection measurements for which branch assignment is not straightforward, the numerical algorithm based on triangular factorization can be used. The resulting hybrid algorithm exploits the best of both approaches: a basic topological algorithm with simple injection conversion to obtain one or more islands which are as large as possible, and a numerical algorithm for application to the reduced system. Only the boundary nodes of the islands obtained via the topological algorithm are retained for the numerical analysis. A tree of angle difference pseudo-measurements are then associated with these nodes in order to take into account the effect of the measurements that have already been processed by the topological algorithm. Such measurements, in addition to the unprocessed injection measurements (measurements for which branch assignment was not straightforward), are then treated with the numerical algorithm.

Example 7.11:

The system shown in Fig. 7.9 will now be used to illustrate this hybrid approach. Processing the power flow measurements (P_{13} , P_{24} , P_{46} , and P_{53}) and the transformable injection measurement (P_3), leaves two islands: $\{1, 3, 5, 7\}$ and $\{2, 4, 6\}$. The reduced system is shown in Fig. 7.9(b). Two voltage angle difference pseudo-measurements are added to the reduced system to represent the measurements which have already been processed. This reduced system can be seen to be observable since the corresponding triangular factor U has a single zero pivot, and this is assigned a pseudo-measurement for the reference angle.

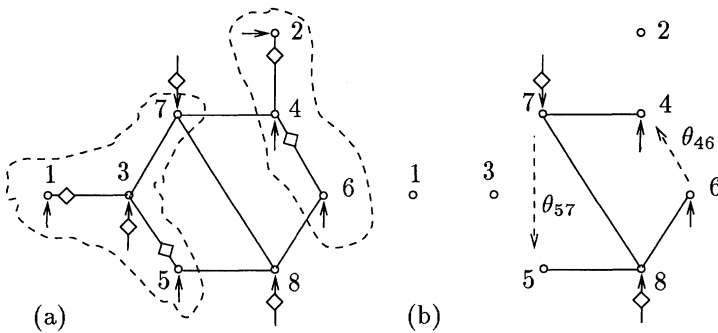


Figure 7.9. (a) Observable 8-node system. (b) Reduced system obtained after reduction of flow and transformable injection measurement.

7.1.11 Solvability and Observability

Although the existence of a spanning tree is a necessary condition for state estimation solvability, it is not a sufficient condition. This section presents some such situations in which topological observability does not guarantee state estimation solvability.

In the previous discussions, a known relationship between the voltage angle difference and the active power flow in each branch of the network was assumed (In the dc power flow model this is expressed as $P_{km} = \theta_{km}/x_{km}$.) This assumption breaks down in cases in which: (a) the branch reactance, x_{km} , is unknown; (b) the branch represents a short-circuit in which case $x_{km} = 0$ (e.g. a closed switching device); and (c) the branch represents an open circuit (e.g. an open switching device). In these cases, the resulting state estimation problem may be unsolvable, even when the network is topologically observable.

Consider, for example, the situation represented in Fig. 7.10(a): the system is both topologically and numerically observable. Compare this to the modified situation in Fig. 7.10(b), where Branch 2–3 is a short circuit. The power flow in this branch cannot be determined from the voltages at the terminal nodes. Hence, both the power flow and the terminal node injections are indeterminate, i.e., P_{23} , P_2 , and P_3 are unobservable, and the system is as a whole unobservable.

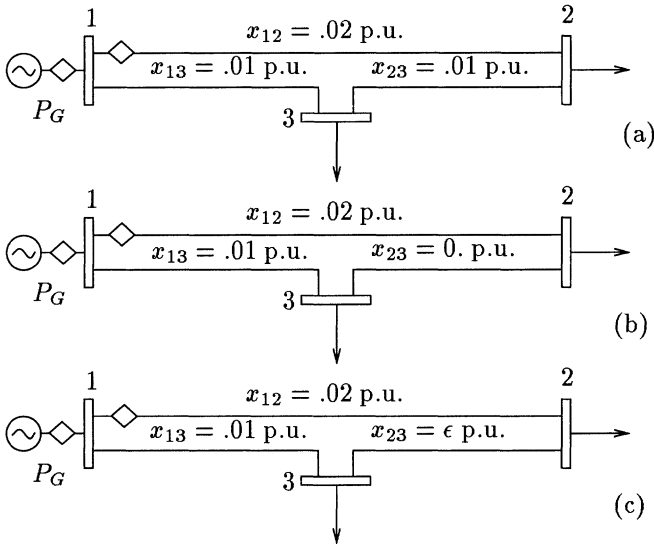


Figure 7.10. Three-bus system.

Since in practice state estimation problems are solved by computers with finite precision, small branch impedances may become an issue: when the angle voltage difference across a small impedance becomes smaller than a certain crit-

ical value, the corresponding power flow estimate becomes uncertain. If neither of the terminal injections are metered, there will be more than one combination of injection and power flow estimates which will satisfy the measurement model. This situation also leads to numerical unobservability.

Consider, for example, the system represented in Fig. 7.10(c) in which Branch 2 – 3 has a small reactance ϵ p.u. Assuming that the branch reactances are $x_{12} = 0.02$ p.u., $x_{13} = 0.01$ p.u., and $x_{23} = \epsilon$ p.u. and that the variance of the measurements are $\sigma_{P_1}^2 = \sigma_{P_{12}}^2 = 0.001$ p.u. and $\sigma_{\theta_1}^2 = 0.$, the measurement model ($\mathbf{z} = \mathbf{H}\mathbf{x} + \mathbf{e}$) can be written as follows

$$\begin{pmatrix} \theta_1^{ref} \\ P_{12}^{meas} \\ P_1^{meas} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 50 & -50 & 0 \\ 150 & -50 & -100 \end{pmatrix} \begin{pmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \end{pmatrix} + \begin{pmatrix} e_1 \\ e_2 \\ e_3 \end{pmatrix}$$

The variance of the power flow estimate P_{23} is given by

$$Var\{P_{23}\} = \mathbf{H}_{23} \mathbf{G}^{-1} \mathbf{H}'_{23} = 1.0 \times 10^{-6} \times \epsilon^{-2}$$

where

$$\mathbf{G} = 10^6 \begin{pmatrix} \infty & -10 & -15 \\ -10 & 5 & 5 \\ -15.0 & 5 & 10 \end{pmatrix} \quad \mathbf{H}'_{23} = \begin{pmatrix} 0 \\ \epsilon^{-1} \\ -\epsilon^{-1} \end{pmatrix}$$

i.e., the variance may assume an arbitrarily large value as reactance ϵ tends to work as a short-circuit. This means that although the network is formally observable, the power flow in Branch 2–3 cannot be determined with acceptable precision; this is also true for the estimates of the power injections at Nodes 2 and 3.

Other cases involve numerical coincidences. Consider, for example, the dc model for the system in Fig. 7.11. In this case, there are four state variables ($\theta_k, k = 1, 2, 3, 4$) and four measurements ($\theta_1^{ref} = 0, P_{12}^{meas}, P_1^{meas}$, and P_2^{meas}). The corresponding measurement model is

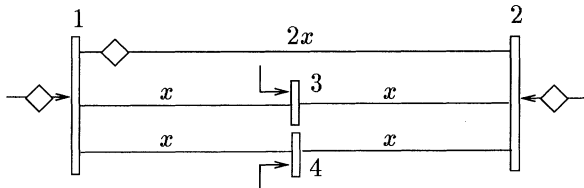


Figure 7.11. Case of numerical coincidence making state estimation unsolvable although the system is topologically observable.

$$\begin{pmatrix} \theta_1^{ref} \\ P_{12}^{meas} \\ P_1^{meas} \\ P_2^{meas} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ x^{-1} & -x^{-1} & 0 & 0 \\ 2.5x^{-1} & -0.5x^{-1} & -x^{-1} & -x^{-1} \\ -0.5x^{-1} & 2.5x^{-1} & -x^{-1} & -x^{-1} \end{pmatrix} \begin{pmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \\ \theta_4 \end{pmatrix} + \begin{pmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \end{pmatrix}$$

The rank of the Jacobian matrix is 3, since Columns 3 and 4 are linearly dependent. Hence, the gain matrix is singular, and the system is unsolvable (although it is observable in the topological sense).

Example 7.12:

Figure 7.12 illustrates a practical situation in which numerical coincidence drastically affects the results of observability analysis. The figure shows an arrangement of two identical three-winding transformers with two measurements (#1 and #2) and two zero injection pseudo-measurements (#3 and #4). The system is observable from the topological point of view; however, state estimation is unsolvable: two zero pivots occur during triangular factorization of the gain matrix (or of the Jacobian matrix if orthogonal transformations are used). Measurements #3 and #4 are thus considered irrelevant, and the branches represented by dashed lines in Fig. 7.12(b) are unobservable.

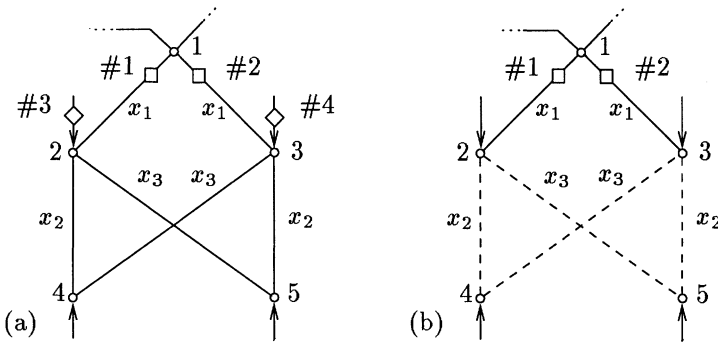


Figure 7.12. Three-winding transformer arrangement. (a) Topologically observable and numerically unsolvable system. (b) Reduced network resulting from numerical observability analysis.

7.2 BUS-SECTION/SWITCHING-DEVICE NETWORK MODEL

Thus far in this chapter, networks modeled at the bus/branch level (power flow model) have been considered. The main characteristic of this type of model is that the network branches follow Ohm's Law; when the dc state estimation

model is used, this means that the angle spread and the power flow in Branch $k - m$ are related by the equation $\theta_{km} = x_{km}^{-1} P_{km}$, where x_{km} is the branch reactance. This is not the case, however, when parts of the network are modeled at the physical level, in which case devices for which Ohm's law does not apply, are represented explicitly.

The impact of modeling parts of a system at the physical level on network observability and in state estimation solvability is discussed next. Figure Fig. 7.13 shows an example of such a model: in this situation, bus sections and switching devices are explicitly represented in the state estimation measurement model $\mathbf{z} = \mathbf{H}\mathbf{x} + \mathbf{e}$.

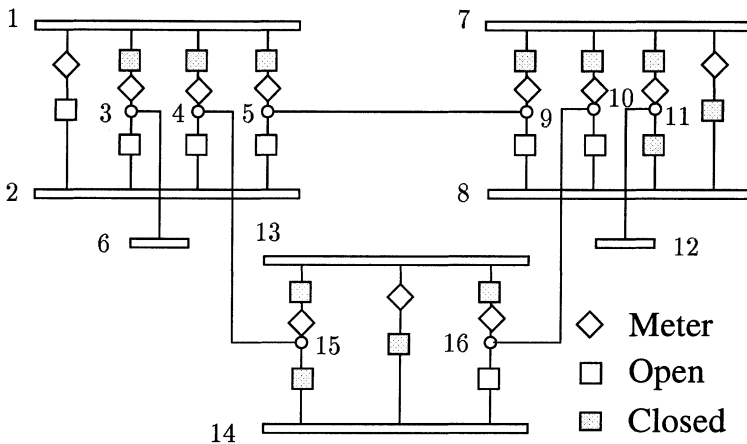


Figure 7.13. Network modeled at the physical level.

7.2.1 Power Flows as State Variables

There are cases, such as in the system illustrated in Fig. 7.10(b), in which, even when the states (voltage angles) at the terminal nodes of a short circuit branch are known, it is not possible to calculate the power flow in the branch. The same is true still for closed switching devices, switching devices with unknown status, and branches with unknown impedances. Open switching devices are also a problem, since even knowing that there is no power flow across the device, leaves the corresponding voltage angle spread indeterminate.

This type of situation has led to a generalization of the concept of state variable. The power flow through these devices is considered as a state variable, and additional information is added, either in terms of pseudo-measurements or equality constraints. For a dc model, zero impedance branches and closed switching devices, such pseudo-measurement (or equality constraint) is given by $\theta_{km} = 0$, while for open switching devices, $P_{km} = 0$. Switching devices with

unknown status and branches with unknown impedance introduce no extra pseudo-measurements into the measurement model.

Power flow measurement P_{km}^{meas} is directly expressed in terms of the corresponding state variables, P_{km} , not involving angle state variables (θ), which guarantees that the series impedance of the device will not appear in the model:

$$P_{km}^{meas} = P_{km} + e_{P_{km}}$$

where P_{km} represents the true value, and the variable $e_{P_{km}}$ denotes the measurement error.

The power injection measurements at terminal Nodes k and m are obtained by the summation of the power flows in the incident branches, as in the usual state estimation formulation; the only difference is that the power flow across the branch $k - m$ is directly expressed in terms of the state variable P_{km} , while the flows across the normal branches (lines and transformers) are given as usual in terms of the state variables θ :

$$P_k^{meas} = \sum_{j \neq m} P_{kj}(\theta_{kj}) + P_{km} + e_{P_k}$$

$$P_m^{meas} = \sum_{j \neq k} P_{mj}(\theta_{mj}) + P_{mk} + e_{P_m}$$

where P_{km} , P_{mk} , θ_{kj} and θ_{mj} denote the true values, and the variables e_{P_k} and e_{P_m} represent the measurement errors.

The Jacobian matrix of the dc state estimator must be modified accordingly. Considering the inclusion of a single flow state on Branch $k - m$, and considering that this flow and that the power injections at both terminal nodes, P_k and P_m , are metered, the augmented Jacobian matrix can be written as follows:

$$\mathbf{H}_{P\theta} = \begin{bmatrix} \overbrace{\theta_k \quad \theta_m}^{\text{angle states}} & \overbrace{P_{km}}^{\text{flow state}} \\ \mathbf{H}_{P\theta} & 0 \\ \mathbf{h}_{P_k} & 1 \\ \mathbf{h}_{P_m} & -1 \\ 0 & 1 \end{bmatrix} \begin{matrix} P_k \\ P_m \\ P_{km} \end{matrix}$$

Since power flow P_{km} is independent of the angle state variables, the additional column is independent of the other columns of the Jacobian matrix. Note that

this would not hold if Ohm's law was applicable to the branch with a flow considered to be a state variable.

Example 7.13:

Consider the three-bus system represented in Fig. 7.10(b). The network is only partially observable for the given set of measurements (P_1 and P_{12}). In order to make the system entirely observable it is necessary to measure at least one of the following: P_2 , P_3 , or P_{23} ; or a combination of them. For example, adding measurement P_3 to the system, results in the following Jacobian matrix and corresponding \mathbf{U} -factor. Since the \mathbf{H} matrix is already in the triangular form, no factorization is necessary.

$$\mathbf{H} = \begin{matrix} & P_{23} & \theta_3 & \theta_2 & \theta_1 \\ \begin{matrix} P_3 \\ P_1 \\ P_{12} \\ \theta_1 \end{matrix} & \begin{pmatrix} -1 & 1 & -1 & \\ & -1 & -1 & 2 \\ & & -1 & 1 \\ & & & 1 \end{pmatrix} \end{matrix} \quad \mathbf{U} = \begin{matrix} & P_{23} & \theta_3 & \theta_2 & \theta_1 \\ \begin{matrix} P_{23} \\ \theta_3 \\ \theta_2 \\ \theta_1 \end{matrix} & \begin{pmatrix} 1 & -1 & 1 & \\ & 1 & 1 & -2 \\ & & 1 & -1 \\ & & & -1 \end{pmatrix} \end{matrix}$$

Note that for simplicity, except for the reactance of Branch 2 – 3 (a zero impedance branch), weighting factors and branch reactances are set at one. Hence, the system is observable and state estimation can be solved.

For the sake of illustration, now consider the elimination of measurement P_3^{meas} and pseudo-measurement θ_1 from the measurement set. In this case, the Jacobian matrix and the corresponding \mathbf{U} -factor become

$$\mathbf{H} = \begin{matrix} & P_{23} & \theta_3 & \theta_2 & \theta_1 \\ \begin{matrix} P_1 \\ P_{12} \end{matrix} & \begin{pmatrix} & -1 & -1 & 2 \\ & & -1 & 1 \end{pmatrix} \end{matrix} \quad \mathbf{U} = \begin{matrix} & P_{23} & \theta_3 & \theta_2 & \theta_1 \\ \begin{matrix} P_{23} \\ \theta_3 \\ \theta_2 \\ \theta_1 \end{matrix} & \begin{pmatrix} 0 & & & \\ & 1 & 1 & -2 \\ & & 1 & -1 \\ & & & 0 \end{pmatrix} \end{matrix}$$

which suggests the need for at least two extra measurements in order to avoid the two zero pivots that appear in the triangular factor.

Example 7.14:

Figure 7.14 shows a ring-bus in which the injection at Node 4 is zero, the injections at Nodes 1 and 3 are unmetered, and the injection at Node 2 is metered. Breakers 1 – 2 and 2 – 3 are closed and Breakers 3 – 4 and 4 – 1 are open. Power flow through Breaker 2 – 1 is metered. The Jacobian matrix \mathbf{H} and the \mathbf{U} -factor are as follows

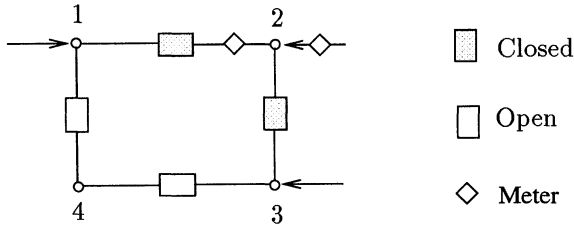


Figure 7.14. Observable ring bus.

$$\mathbf{H} = \begin{matrix} & \theta_1 & \theta_2 & \theta_3 & \theta_4 & P_{23} & P_{12} & P_{34} & P_{41} \\ \theta_{12} & \left(\begin{array}{cccccccc} 1 & -1 & & & & & & & \\ & 1 & -1 & & & & & & \\ & & & & & 1 & -1 & & \\ P_2 & & & & & & -1 & & \\ P_{21} & & & & & & & & \\ P_{34} & & & & & & & 1 & \\ P_4 & & & & & & & -1 & 1 \\ P_{41} & & & & & & & & 1 \end{array} \right) \end{matrix}$$

$$\mathbf{U} = \begin{matrix} & \theta_1 & \theta_2 & \theta_3 & \theta_4 & P_{23} & P_{12} & P_{34} & P_{41} \\ \theta_1 & \left(\begin{array}{cccccccc} 1 & -1 & & & & & & & \\ \theta_2 & & 1 & -1 & & & & & \\ \theta_3 & & & 0 & & & & & \\ \theta_4 & & & & 0 & & & & \\ P_{23} & & & & & 1 & -1 & & \\ P_{12} & & & & & & 1 & & \\ P_{34} & & & & & & & 1 & -1 \\ P_{41} & & & & & & & & 1 \end{array} \right) \end{matrix}$$

In this case the \mathbf{U} -factor has two zero pivots: at the positions (3, 3) and (4, 4), indicating that two reference angles are needed: one at Node 3 and another at Node 4. If these are provided, the system becomes entirely observable: all node angles and breaker flows can be calculated from the set of measured values available. The entire system can then be treated as a single island, including Node 4 and open Breakers 3 – 4 and 4 – 1. The advantage of doing this is the facilitation of the processing of status errors, which may involve breakers with status incorrectly considered as being open.

7.2.2 Extended Observable Islands

As discussed above, the inclusion of breakers, switches, zero impedance branches, and branches with unknown impedances in the generalized state estimation

model, motivates the following extensions of the concepts of islands and observable islands (see also Definitions 1 and 2 at the beginning of the chapter):

- *Definition 3:* An island is a contiguous part of a network with bus sections as nodes and lines, transformers, open switches, closed switches, and switches with unknown status as branches.
- *Definition 4:* An observable island is an island for which all branch flows can be calculated from the available measurements independent of the values adopted for reference pseudo-measurements.

Example 7.15:

Considering a conventional network topology processor and in view of Definition 2 (presented at the beginning of the chapter), the network of Fig. 7.15(a) is unobservable.

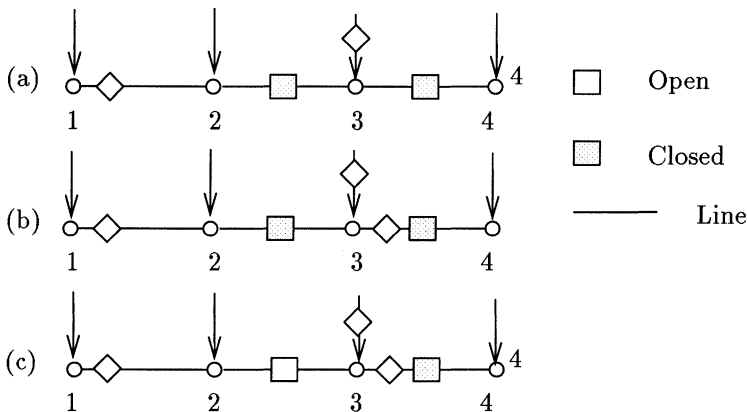


Figure 7.15. (a) Unobservable island. (b) Observable islands. (c) Extended observable island.

According to Definition 4, the system in Fig. 7.15(a) is unobservable. If a reference angle is provided at Node 1, and knowing the value P_{12}^{meas} , since the angular differences across closed breakers (short-circuits) are zero, all the nodal voltage angles can be estimated. The flow through the closed breakers is indeterminate, however, since any value can be assigned to power injections at Nodes 2 and 4, as long as $P_2 + P_4 = 0$; hence, the power flows P_{23} and P_{34} are indeterminate. Even though the nodal state variables are known the state of the system (i.e., all the relevant information about the system) is not completely known. Observability analysis can treat the power flows 2-3 and 3-4 as unobservable flows, and drop the corresponding states from the state estimation.

If an additional flow measurement is added, as indicated in Fig. 7.15(b), the flows through Branches 2 – 3 and 3 – 4, and the injections at Nodes 3 and 4 would become susceptible to estimation, and all the relevant data about the network would be known.

The situation depicted in Fig. 7.15(c) shows an observable network. Notice that the flow through the open switch is known (zero); two angular references, at, for example, Buses 1 and 4 are required, but the arbitrary values adopted for θ_1 and θ_4 do not affect the power flows in the network. Of course, this situation could be considered as two separate physical islands, but, as mentioned above, it may be useful to maintain the system as a single island in order to be able to process bad status data, since, open breaker status may at times be considered suspect data; i.e., the correct status could be closed rather than open.

7.2.3 Extended Numerical Observability Algorithm

A modified version of the numerical observability algorithm designed to handle networks that are totally or partially represented at the physical level will now be discussed. The basic modifications involves the addition of new state variables (power flow state variables) and new pseudo-measurements (to express information about the status of switches and zero impedance branches). As with the algorithm discussed above for the bus/branch model, this modified algorithm is also based on the presence of zero pivots that may occur during triangular factorization of the gain matrix (or the orthogonal factorization of the Jacobian matrix). The difference is that the gain matrix (or the Jacobian matrix) in the present situation will include additional information (pseudo-measurements) as well as new states (the flow states), and since power flows are also state variables, zero pivots may occur in connection with these variables as well. When this happens, state variables corresponding to these zero pivots are added as pseudo-measurements with arbitrary values, just as in the bus/branch model. Irrelevant measurements are again identified, i.e., injection measurements with incident branches with estimated flows being a function of the arbitrary values assigned to the pseudo-measurements added (non zero flows), are considered irrelevant.

Algorithm:

1. Initialization
 - (a) Initialize the measurement set of interest as consisting of all available measurements and pseudo-measurements representing switches and short circuits.
 - (b) Initialize the power network of interest as consisting of all branches incident to at least one measurement or pseudo-measurement.
2. Form gain matrix \mathbf{G} and perform triangular factorization $\mathbf{G} = \mathbf{U}'\mathbf{U}$.

3. Introduce angle/flow pseudo-measurements whenever a zero pivot is encountered.
4. Solve the dc state estimator equation for the angle/flow state variables considering all the measured values equal to zero, except for the added angle/flow pseudo-measurements that are assigned arbitrary values.
5. System update.
 - (a) Remove from the power network of interest all the branches with non-zero flows.
 - (b) Update the measurement set of interest by removing power injection measurements adjacent to the removed branches along with the corresponding pseudo-measurements.
 - (c) If modifications have been made, update the triangular factor \mathbf{U} and go to 3.
6. Form islands with nodes connected by branches with zero flows.

Example 7.16:

This algorithm will now be applied to the system in Fig. 7.15(a).

1. Initialization
 - (a) Initial measurement set of interest: $\{P_{12}, P_3, \theta_{23}, \theta_{34}\}$.
 - (b) Initial power network of interest: $\{1-2, 2-3, 3-4\}$.
2. The \mathbf{H} matrix and the corresponding triangular factor \mathbf{U} are as follows

$$\mathbf{H} = \begin{array}{c} \begin{array}{cccccc} & \theta_1 & \theta_4 & P_{23} & P_{34} & \theta_2 & \theta_3 \\ P_{12} & \left(\begin{array}{cccccc} 1 & & & & -1 & \\ & -1 & & & & 1 \\ & & -1 & 1 & & \\ \theta_{34} & & & & & 1 & \\ P_3 & & & & & & 1 & -1 \\ \theta_{23} & & & & & & & -1 \end{array} \right) \end{array} \end{array}$$

$$\mathbf{U} = \begin{array}{c} \begin{array}{cccccc} & \theta_1 & \theta_4 & P_{23} & P_{34} & \theta_2 & \theta_3 \\ \theta_1 & \left(\begin{array}{cccccc} 1 & & & & -1 & \\ \theta_4 & & 1 & & & -1 \\ P_{23} & & & 1 & -1 & \\ P_{34} & & & & 0 & \\ \theta_2 & & & & & 1 & -1 \\ \theta_3 & & & & & & 0 \end{array} \right) \end{array} \end{array}$$

3. A flow pseudo-measurement is added for P_{34} and an angle pseudo-measurement is added for θ_3 .

4. The solution of the dc state estimator, considering all the measurements and pseudo-measurements of the set of interest reading zero, yields arbitrary (non-zero) flows in Branches 2 – 3 and 3 – 4, and zero flow in Branch 1 – 2..
5. Update system.
 - (a) Branches 2 – 3 and 3 – 4 are removed from the network of interest.
 - (b) Injection measurement P_3 and the pseudo-measurements θ_{23} and θ_{34} are removed from the measurement set of interest.
 - (c) The updated \mathbf{U} -factor is

$$\mathbf{U} = \begin{matrix} & \theta_1 & \theta_2 \\ \theta_1 & \left(\begin{array}{cc} 1 & -1 \\ \theta_2 & 0 \end{array} \right) \end{matrix}$$

3. A single reference angle at Bus 2 is required for this reduced system.
4. $\hat{P}_{12} = 0$ regardless of the value assigned to the reference angle θ_2 .
5. System update is not required.
6. Branch 1 – 2 forms an observable island.

7.3 MEASUREMENT ADDITION TO IMPROVE OBSERVABILITY

Meter placement algorithms can be used both for the design of a metering system (planning phase) and for the addition of pseudo-measurements based on load and generation prediction as part of a real-time network model builder. Simpler meter placement algorithms aim at making the network observable, although other objectives such as bad data detectability and identifiability may also be considered. In this section only the addition of pseudo-measurements in a real-time environment is discussed.

The addition of pseudo-measurements will be considered with the system shown in Fig. 7.6. This system consists of three observable islands and eight unobservable branches. The injection measurements 1 and 2 were temporarily considered irrelevant since they were of no help in further expanding the three observable islands. The measurement placement procedure adds a set of pseudo-measurements to make the whole system observable, and at this stage, the injection measurements 1 and 2 can be reintroduced into the measurement set, along with the pseudo-measurements required for observability.

7.3.1 Additional Measurements

To make a network such as that illustrated in Fig. 7.6 observable, it is sufficient to add injection measurements at all boundary nodes, as well as at the isolated

nodes of the unobservable part of the network (Nodes 3 through 8 in the example case, since Nodes 1 and 2 already have injection measurements). A Bus Scheduler function can provide the predicted data for both the unobservable part of the network and the boundary buses wherever measured values are not available (measurements that previously have been considered irrelevant).

7.3.2 Avoiding Contamination

There are two basic ways of protecting previously-determined observable states (observable islands shown in Fig. 7.6) from being corrupted by the predicted data introduced to make the whole network observable: (a) making sure the set of measurements added is a non-redundant set (with zero residuals in state estimation); or (b) using very high weightings for the observable parts (see Sec. 1.2 in Chap 1). In the first case, the observability algorithm could be used to guide the addition of pseudo-measurements (predicted data). In the second case, one must assume the robustness of the central state estimation algorithm for dealing with high weightings and possible interferences with small branch impedances.

7.3.3 Measurement Placement Algorithm for the Bus/Branch Model

This algorithm selects an additional set of pseudo-measurements to make the network barely observable (observable without redundancy). Thus these measurements will not contaminate the state estimation of the observable islands, even with normal weightings for the additional measurements. Measurement placement is performed sequentially by adding one pseudo-measurement at a time. Candidate nodes for additional pseudo-measurement of injections to make the whole network observable are those nodes with at least one unobservable branch. The additional pseudo-measurements will eventually make unobservable branches observable, thus coalescing observable islands (and isolated unobservable nodes) into larger islands. This process is carried out by updating the observable islands after each iteration (meter placement) until the whole network becomes one observable island. Thus the algorithm for meter placement uses the same basic steps as the algorithm for observability analysis. The computational requirements of the algorithm depend on the method of updating triangular factors of the gain matrix after new measurements are added. Two options are available: sparse factor updating (normal equation) or Givens rotations (orthogonal factorization).

Algorithm:

1. Form the gain matrix \mathbf{G} and perform triangular factorization $\mathbf{G} = \mathbf{U}'\mathbf{U}$.
2. Introduce angle pseudo-measurements whenever a zero pivot is encountered. If only one zero pivot occurs, stop, and the system is observable.

3. Solve the dc state estimator equation for the angle state variables, considering all measured values to be zero, except the added angle pseudo-measurements that are assigned arbitrary values.
4. Place meter.
 - (a) Determine the set of nodes that have no injection measurements, but with incident branches having at least one arbitrary (non zero) flow. If no candidate nodes are found, stop. Else.
 - (b) Introduce an injection pseudo-measurement at one of the candidate nodes and update the \mathbf{U} -factor.
 - (c) Solve the dc state estimator equation as in Step 3. Compute the residuals for all the angle pseudo-measurements. Drop one of the angle pseudo-measurements (with non zero residual) that has become redundant through the addition of an injection pseudo-measurement and update the \mathbf{U} -factor.
5. Return to Step 3.

Remarks:

- When the part of the network made barely observable by the above algorithm contains relatively short lines in which neither power flow nor terminal node injections are metered, convergence problems can occur. The algorithm can then be modified by forcing the addition of injection pseudo-measurements at these nodes.
- The above algorithm can also be extended to allow placement of flow measurements (in addition to injection measurements). Power flow measurements can be added to unobservable branches in Step 3.a. The placement of flow measurements on short lines normally helps avoid convergence problems.
- As discussed earlier in the chapter, making the system observable does not guarantee that state estimation will be successful. The variances of the power flow estimates can be used as an indication of the need for additional measurement placement (redundant measurements) or as a guide for measurement replacement (in which case, the system would remain barely observable). If redundant measurements are added, they should be given reduced weights in order to avoid contamination of telemetered data.
- Steps 4.b and 4.c can be merged into a single step. Power injections at candidate nodes are necessarily unobservable. Hence the corresponding power injection estimate will be a function of at least one of the angle reference pseudo-measurements. Then the injection measurement and a related angle pseudo-measurement are swapped and the \mathbf{U} -factor is updated accordingly in a single operation.

Example 7.17:

This algorithm will now be applied to the system in Fig. 7.7 (a).

1. The gain matrix G_θ and the triangular factor U are the same as in Example 7.10. The triangular factor is

$$U = \begin{matrix} & \theta_1 & \theta_2 & \theta_6 & \theta_5 & \theta_8 & \theta_4 & \theta_3 & \theta_7 \\ \theta_1 & \left(\begin{array}{cccccccc} 1.41 & & & & & & & & \\ & 1.00 & 1.00 & & & & & & \\ & & 1.00 & & & & & & \\ & & & 1.22 & & & & & \\ & & & & 1.00 & 1.00 & 1.00 & & \\ & & & & & 0.00 & & & \\ & & & & & & & 0.58 & -0.58 \\ & & & & & & & & 0.00 \end{array} \right) \end{matrix}$$

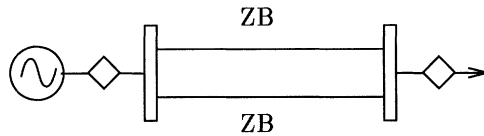
2. Two zero pivots are found: Nodes 4 and 7. Hence, θ_4 and θ_7 are assigned arbitrary values: here, $\theta_4 = 0.$ and $\theta_7 = 1.0.$ The corresponding pivots of the U -factor are set at one.
3. The solution of the dc state estimator equation for θ , considering all the measured values, except θ_4 and θ_7 which assume arbitrary values, equal to zero, yields: $\theta_1 = 1., \theta_2 = -1., \theta_3 = 1., \theta_4 = 0., \theta_5 = 1., \theta_6 = 0., \theta_7 = 1.,$ and $\theta_8 = 2.$
4. Meter placement.
 - (a) The set of nodes that have no injection measurements and with incident branches having at least one arbitrary (non zero) flow is identified: $\{2, 5, 6, 8\}.$
 - (b) Introduce an injection pseudo-measurement at Node 2. The updated U -factor is

$$U = \begin{matrix} & \theta_1 & \theta_2 & \theta_6 & \theta_5 & \theta_8 & \theta_4 & \theta_3 & \theta_7 \\ \theta_1 & \left(\begin{array}{cccccccc} 1.41 & & & & & & & & \\ & 1.41 & 0.71 & & & & & & \\ & & 1.22 & & & & & & \\ & & & 1.22 & & & & & \\ & & & & 1.00 & 1.00 & 1.00 & & \\ & & & & & 0.58 & & & \\ & & & & & & & 0.58 & -0.58 \\ & & & & & & & & 0.00 \end{array} \right) \end{matrix}$$

- (c) The new solution of the dc state estimator indicates that both angle pseudo-measurements have become redundant, since they have non-zero residuals, and pseudo-measurement θ_4 is dropped.
5. Return to Step 3.
 3. The new solution of the dc state estimator indicates that all branch flows are zero, and the system has become observable.

7.3.4 Measurement Placement Algorithm for the Physical Level Model

When a network is modeled at the physical level, the addition of injection measurements may not be enough to make the network observable (as, for example, the system shown in Fig. 7.16 in which there are two zero impedance branches connected in parallel; see also Prob. 6 at the end of the chapter). The following algorithm is an extension of the numerical algorithm above, and was designed to cope with systems modeled at the physical level.



ZB – Zero impedance branch

Figure 7.16. Unobservable system requiring at least one flow measurement to become observable.

Algorithm:

1. Form gain matrix \mathbf{G} and perform triangular factorization $\mathbf{G} = \mathbf{U}'\mathbf{U}$.
2. Introduce angle/flow pseudo-measurements whenever a zero pivot is encountered. If only one zero pivot occurs, stop; the system is observable.
3. Solve the dc state estimator equation for the angle/flow state variables considering all measured values to be equal to zero, except for the added angle/flow pseudo-measurements that are assigned arbitrary values.
4. Place meter.
 - (a) Determine the set of unobservable power flow states and the set of unobservable injections (unmetered injections with at least one incident

unobservable power flow). If there are neither unobservable injections nor unobservable flow states, stop. Else.

(b) Select Meter

- i. If the set of unobservable injections is not empty,
 - A. Introduce a injection pseudo-measurement at one of the nodes with unobservable injections and update the \mathbf{U} -factor.
 - B. Solve the dc state estimator equation as in Step 3. Compute the residuals for all the angle pseudo-measurements. Drop one of the angle pseudo-measurements that has become redundant (that has a non zero residual) by the addition of the injection pseudo-measurement and update the \mathbf{U} -factor.
- ii. Otherwise, if the set of unobservable flow states is not empty,
 - A. Introduce a flow pseudo-measurement in one of the branches modeled with flow state variables and update the \mathbf{U} -factor.

5. Return to Step 3.

Example 7.18:

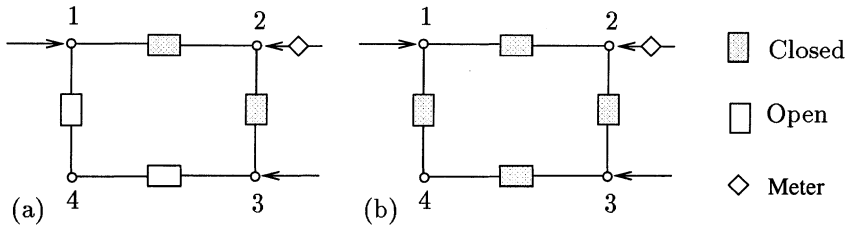


Figure 7.17. Unobservable ring buses.

The two examples given in Fig. 7.17 will now be considered. In case (a), Breakers 1 – 2 and 2 – 3 are closed and Breakers 1 – 4 and 3 – 4 are open. Injections at Nodes 1 and 3 are unmetered. Injection at Node 2 is metered and injection at Node 4 is zero (pseudo-measurement). The system is unobservable since power flows in the closed breakers can be assigned arbitrary values, as long as $P_2 = P_{21} + P_{23}$, and still be consistent with the given measurements (injection at Node 2 and pseudo-measurement at Node 4). The injections at Nodes 1 and 3 are found to be unobservable and become candidates for meter placement: one possibility for the procedure would be the placement of an injection measurement at Node 1 using the algorithm above. The flow in the closed breakers and the injection at Node 3 then become observable.

In case (b), all the breakers are closed. Injections at Nodes 1 and 3 are unmetered. The injection at Node 2 is metered and at Node 4 is zero (pseudo-measurement), as in the previous case. The system is unobservable since power

flows in the closed breakers can be assigned arbitrary values, as long as $P_2 = P_{21} + P_{23}$ and $0 = P_{41} + P_{43}$, and still be consistent with the given measurements (injection at Node 2 and pseudo-measurement at Node 4). The injections at Nodes 1 and 3 are declared unobservable and become candidates for meter placement. As in the previous case, one possibility is the placement of an injection measurement at Node 1 again using the above algorithm. In this case the flow in the closed breakers remain unobservable, but the injection at Node 3 becomes observable. Since all injections are observable, the algorithm selects one of the closed breaker flows as a new pseudo-measurements, and the system finally becomes observable (The quality of the estimated flows in the remaining closed breakers that form part of the ring will depend on the quality of this pseudo-measurement.)

7.4 HISTORICAL NOTES AND REFERENCES

The topological observability approach was introduced as a partial requirement for state estimation solvability by Clements and Wollenberg [1975], and further developed by Krumpholz, Clements, and Davis [1980] and Clements, Krumpholz, and Davis [1981,82,83]. An observability analysis method based on matroid intersections was developed by Quintana, Simões-Costa, and Mandel [1982]. A critical evaluation of the topological approach was presented by Allemong, Irisarri, and Sasson [1980], and a numerical approach proposed to observability analysis which could take into account both topological and numerical aspects (Monticelli and Wu [1985]). This numerical approach was extended to the orthogonal estimators (Monticelli and Wu [1986]), to the normal equation method with equality constraint (Wu, Liu, and Lun [1988]) and to the blocked sparse matrix formulation (Nucera, Brandwajn and Gilles [1993]). A hybrid (topological/numerical) approach was proposed by Contaxis and Korres [1988]. A review of the principal observability analysis and meter placement algorithms was presented by Clements [1990]. The impact of modeling short circuit branches in conventional WLS state estimators and in observability analysis was discussed by Monticelli and Garcia [1992] and a WLS state estimator for networks modeled at the physical level was introduced by Monticelli [1993]. The role of irrelevant measurements (discardable measurements) in observability analysis, both numerical and topological, was discussed by Monticelli, Garcia, and Slutsker [1992]. A comprehensive discussion about rank deficient matrices can be found in Steward [1973].

7.5 PROBLEMS

- 1. Consider the system in Fig. 7.5(b). Find an injection measurement assignment that yields a spanning tree (a connected tree incorporating all network nodes).

- 2. Consider the system in Fig. 7.5(b) again. Now find all the possible injection measurement assignments and determine the largest observable islands.
- 3. Consider the situation represented in Fig. 7.7(b). Determine the observable islands and the irrelevant injection measurements using the numerical observability algorithm given in Subsec. 7.1.9. Repeat the procedure using different sets of values for the reference voltage angles. Compare and discuss the results.
- 4. Extend the numerical observability algorithm in Subsec. 7.1.9 to include the features discussed in the remarks of Subsec. 7.1.5. Apply the resulting algorithms to the situations given in Fig. 7.7.
- 5. Study the observability of the system in Fig. 7.16. Introduce pseudo-measurements to make the system observable with zero redundancy. Solve the state estimation problem considering the injection measurements to be equal to zero or assigning an arbitrary value for the added pseudo-measurement. Determine which parts of the network are affected by the pseudo-measurement added.
- 6. Study the observability of the systems given in Fig. 7.15. Whenever necessary, use the meter placement algorithm described in Subsec. 7.3.4 to make the system observable.
- 7. Consider the state estimation model $\mathbf{z} = \mathbf{H}\Theta + \mathbf{e}$. Suppose that l pseudo-measurements are added to make the network solvable, and all other measurements are equal to zero. Show that the residuals $\mathbf{r} = \mathbf{0}$. Hint: see the theorem in Subsec. 7.1.4 in this chapter and Prob. 6 in Chap 2.
- 8. Write the normal equation corresponding to Eq. 7.2 indicating the relevant parts of matrices \mathbf{H} and \mathbf{W} , and of vectors \mathbf{x} and \mathbf{z} .

References

- Allemon, J.J., Irisarri, G.D. and Sasson, A.M., "An examination of solvability for state estimation algorithms", Paper A80-008-3, IEEE PES Winter Meeting, New York, NY, 1980.
- Clements, K.A. and Wollenberg, B.F., "An algorithm for observability determination in power system state estimation", paper A75 447-3, IEEE/PES Summer Meeting, San Francisco, CA, July 1975.
- Clements, K.A., Krumpholz, G.R. and Davis, P.W., "Power system state estimation residual analysis: An algorithm using network topology", IEEE Trans. Power Apparatus and Systems, Vol. 100, pp. 1779-1787, April, 1981.
- Clements, K.A., Krumpholz, G.R. and Davis, P.W., "Power system state estimation with measurement deficiency: An observability/measurement placement algorithm", IEEE Trans. Power Apparatus and Systems, Vol. 101, No. 9, pp. 3044-3052, Sept. 1982.

- Clements, K.A., Krumpholz, G.R. and . Davis, P.W., "Power system state estimation with measurement deficiency: An algorithm that determines the maximal observable network", IEEE Trans. Power Apparatus and Systems, Vol. 102, No. 7, pp. 2012-2020, Aug. 1983.
- Clements, K.A., "Observability methods and optimal meter placement", Int. J. Elec. Power, Vol. 12, No. 2, pp. 89-93, April 1990.
- Contaxis, G.C. and Korres, G.N., "A reduced model for power system observability analysis and restoration", IEEE Trans. in Power Systems, Vol. 3, No. 4, pp. 1411-1417, Nov. 1988.
- Krumpholz, G. R., Clements, K.A. and Davis, P.W., "Power system observability analysis: A practical algorithm using network topology", IEEE Trans. Power Apparatus and Systems, Vol. 99, pp.1534-1542, July/Aug., 1980.
- Monticelli, A. and Wu, F.F., "A method that combines internal state estimation and external network modeling", IEEE Trans. PAS, Vol. 104, No. 1, Jan. 1983.
- Monticelli, A. and Wu, F.F., "Network observability: Theory", IEEE Trans. PAS, Vol. 104, No. 5, pp. 1035-1041, May 1985.
- Monticelli, A. and Wu, F.F., "Network observability: Identification of observable islands and measurement placement", IEEE Trans. on Power Apparatus and System, Vol. 104, No. 5, pp. 1042-1048, May 1985.
- Monticelli, A. and Wu, F.F., "Observability analysis for orthogonal transformation based state estimation", IEEE Trans. on Power Systems, Vol. 1, No. 1, pp. 201-206, Feb. 1986.
- Monticelli, A. , "The impact of modeling short circuit branches in state estimation", Paper 92 WM 186-7-on Power Systems, 1992 Winter Meeting, New York, Jan. 1992.
- Monticelli, A., Garcia, A. and Slutsker, I., "Handling Discardable Measurements in Power System State Estimation", IEEE Trans. on Power Systems, Vol. 7, pp. 1341-1349, Aug. 1992.
- Monticelli, A., "Modeling Circuit Breakers in Weighted Least Squares State Estimation", IEEE Trans. on Power Systems, Vol. 8, No. 3, pp. 1143-1149, Aug. 1993.
- Nucera, R.R., Brandwajn, V., and Gilles, M.L., "Observability analysis and bad data - Analysis using augmented blocked matrices", IEEE Trans. Power Syst., Vol. 8, No. 2, pp. 426-433, May 1993.
- Quintana, V.H., Simões-Costa, A, and Mandel, A., "Power system topological observability using a direct graph-theoretic approach", IEEE Trans. on Power Systems, Vol. 101, No. 3, pp. 617-626, 1982.
- Steward, G.W., *Introduction to Matrix Computations*, Academic Press, 1973.
- Wu, F.F., Liu, E.H.E, and Lun, S.M. "Observability analysis and bad data-processing for state estimation with equality constraints", IEEE Trans. Power Syst., Vol. 3, No. 2, pp. 541-578, May 1988.

8 BASIC TECHNIQUES FOR BAD DATA PROCESSING

When a state estimation model fails to yield estimates with a degree of accuracy compatible with the standard deviations of the quantities estimated one must conclude either that the measured quantities contain spurious data or that the model is unfit to explain the measured quantities, or that both are true. This chapter deals with cases in which the estimation model is assumed to be fixed and correct and that the set of measured quantities contains bad data (measurements that are grossly in error). The basic techniques normally used for detecting and identifying bad data are discussed, including the $J(\mathbf{x})$ -test, the \mathbf{r}^n -test, and error estimation techniques; these techniques are basic in that they can be successfully used for processing both single and multiple noninteracting bad data. Techniques for treating interacting bad data and status/parameter errors are discussed in Chap. 9.

8.1 REVIEW OF THE DC STATE ESTIMATOR

The dc state estimator model discussed in Chap. 3 relates the measurement vector, \mathbf{z} , to the state vector, \mathbf{x} , i.e.,

$$\mathbf{z} = \mathbf{H} \mathbf{x} + \mathbf{e} \quad (8.1)$$

where \mathbf{x} is the n vector of the true states (unknown), \mathbf{z} is the m vector of measurement (known), \mathbf{H} is the $m \times n$ Jacobian matrix, \mathbf{e} is the m vector of random errors (in the absence of gross errors, elements e_i are assumed to be normal and independent, each with a mean of zero and a variance of σ_i^2), m is the number of measurements, and n is the number of state variables.

The estimate $\hat{\mathbf{x}}$ of the unknown state vector \mathbf{x} is obtained by minimizing the weighted least-squares index, defined as

$$J(\mathbf{x}) = (\mathbf{z} - \mathbf{H}\mathbf{x})' \mathbf{W} (\mathbf{z} - \mathbf{H}\mathbf{x}) \quad (8.2)$$

where the weighting matrix \mathbf{W} is usually the inverse of the covariance matrix of the measurements, \mathbf{R}_z .

The condition of optimality is that the gradient of $J(\mathbf{x})$ vanishes at the optimal solution $\hat{\mathbf{x}}$, i.e.,

$$\mathbf{G} \hat{\mathbf{x}} - \mathbf{H}' \mathbf{W} \mathbf{z} = \mathbf{0}$$

where $\mathbf{G} = \mathbf{H}' \mathbf{W} \mathbf{H}$. This yields

$$\hat{\mathbf{x}} = \mathbf{G}^{-1} \mathbf{H}' \mathbf{W} \mathbf{z} \quad (8.3)$$

An estimate $\hat{\mathbf{z}}$ of the measurement vector \mathbf{z} is given by

$$\hat{\mathbf{z}} = \mathbf{H} \hat{\mathbf{x}} \quad (8.4)$$

The vector of residuals is defined as $\mathbf{r} = \mathbf{z} - \mathbf{H}\mathbf{x}$; an estimate of \mathbf{r} is given by

$$\hat{\mathbf{r}} = \mathbf{z} - \mathbf{H} \hat{\mathbf{x}} \quad (8.5)$$

8.2 COVARIANCE MATRICES

This section summarizes the derivation of the covariance matrices used in the analysis of bad data. These covariance matrices are obtained from the sensitivity matrices $\partial\hat{\mathbf{x}}/\partial\mathbf{z}$, $\partial\hat{\mathbf{z}}/\partial\mathbf{z}$ and $\partial\hat{\mathbf{r}}/\partial\mathbf{z}$.

8.2.1 Sensitivity Analysis

Assume that the state estimate $\hat{\mathbf{x}}$ has already been calculated from Eq. (8.3). Sensitivity analysis is used to determine the impact $\Delta\hat{\mathbf{x}}$ caused on the state estimate by an arbitrary perturbation $\Delta\mathbf{z}$ introduced in the measurement vector. In this case the sensitivity matrix $\partial\hat{\mathbf{x}}/\partial\mathbf{z}$ is given by

$$\partial\hat{\mathbf{x}}/\partial\mathbf{z} = \mathbf{G}^{-1} \mathbf{H}' \mathbf{W} \quad (8.6)$$

since Eq. (8.3) is linear with respect to both $\hat{\mathbf{x}}$ and \mathbf{z} , having the following incremental form

$$\Delta \hat{\mathbf{x}} = \mathbf{G}^{-1} \mathbf{H}' \mathbf{W} \Delta \mathbf{z} \quad (8.7)$$

Equation (8.4) gives

$$\Delta \hat{\mathbf{z}} = \mathbf{H} \Delta \hat{\mathbf{x}} = \mathbf{H} \mathbf{G}^{-1} \mathbf{H}' \mathbf{W} \Delta \mathbf{z} \quad (8.8)$$

which yields the sensitivity matrix $\partial \hat{\mathbf{z}} / \partial \mathbf{z}$

$$\partial \hat{\mathbf{z}} / \partial \mathbf{z} = \mathbf{H} \mathbf{G}^{-1} \mathbf{H}' \mathbf{W} = \mathbf{H} \partial \hat{\mathbf{x}} / \partial \mathbf{z} \quad (8.9)$$

Now, writing Eq. (8.5) in the incremental form, and using Eq. (8.6), yields

$$\Delta \hat{\mathbf{r}} = \Delta \mathbf{z} - \mathbf{H} \Delta \hat{\mathbf{x}} = (\mathbf{I} - \mathbf{H} \mathbf{G}^{-1} \mathbf{H}' \mathbf{W}) \Delta \mathbf{z} \quad (8.10)$$

where \mathbf{I} is an $m \times m$ identity matrix. This implies that the sensitivity matrix $\partial \hat{\mathbf{r}} / \partial \mathbf{z}$ is as follows

$$\mathbf{S} = \partial \hat{\mathbf{r}} / \partial \mathbf{z} = \mathbf{I} - \mathbf{H} \mathbf{G}^{-1} \mathbf{H}' \mathbf{W} = \mathbf{I} - \partial \hat{\mathbf{z}} / \partial \mathbf{z} \quad (8.11)$$

Remarks:

- For the linear estimation model in Eq. (8.1), Eqs. (8.10)-(8.11) yield the following residual vector estimate

$$\hat{\mathbf{r}} = \mathbf{S} \mathbf{z} = \mathbf{S} (\mathbf{H} \mathbf{x} + \mathbf{e}) \quad (8.12)$$

where \mathbf{x} is the true state vector (unknown) and \mathbf{e} is the vector of measurement errors. Since

$$\mathbf{S} \mathbf{H} \mathbf{x} = (\mathbf{I} - \mathbf{H} \mathbf{G}^{-1} \mathbf{H}' \mathbf{W}) \mathbf{H} \mathbf{x} = \mathbf{0}$$

then

$$\hat{\mathbf{r}} = \mathbf{S} \mathbf{e} \quad (8.13)$$

This equation can now be rewritten in the following weighted form:

$$\mathbf{r}^w = \mathbf{S}^w \mathbf{e}^w \quad (8.14)$$

where $\mathbf{r}^w = \mathbf{R}_z^{-1/2} \hat{\mathbf{r}}$ is the weighted residual vector, $\mathbf{S}^w = \mathbf{R}_z^{-1/2} \mathbf{S} \mathbf{R}_z^{1/2}$ is the weighted sensitivity matrix, $\mathbf{e}^w = \mathbf{R}_z^{-1/2} \mathbf{e}$ is the weighted vector of measurement errors, and $\mathbf{R}_z = \mathbf{W}^{-1}$ is the measurement covariance matrix (a diagonal matrix).

- Note that, since

$$(\mathbf{H} \mathbf{G}^{-1} \mathbf{H}' \mathbf{W}) \hat{\mathbf{r}} = \mathbf{0}$$

(see Prob. 1 at the end of the chapter) it follows that

$$\mathbf{S} \hat{\mathbf{r}} = (\mathbf{I} - \mathbf{H} \mathbf{G}^{-1} \mathbf{H}' \mathbf{W}) \hat{\mathbf{r}} = \hat{\mathbf{r}} \quad (8.15)$$

As above, this expression can be rewritten in weighted form, i.e.,

$$\mathbf{r}^w = \mathbf{S}^w \mathbf{r}^w \quad (8.16)$$

- If \mathbf{u}_j is a vector of zeros, except that the j th element is equal to one, then $\mathbf{S}_{\bullet j}^w = \mathbf{S}^w \mathbf{u}_j$. Equation (8.15) thus yields $\mathbf{S}_{\bullet j}^w = \mathbf{S}^w \mathbf{S}_{\bullet j}^w$, with the j th element of $\mathbf{S}_{\bullet j}^w$ being

$$\mathbf{S}_{jj}^w = (\mathbf{S}_{\bullet j}^w)' \mathbf{S}_{\bullet j}^w \quad (8.17)$$

8.2.2 Covariances of Estimated Vectors

If \mathbf{y} is an m_y vector of random variables with an expected value of $E(\mathbf{y}) = \bar{\mathbf{y}}$, \mathbf{v} is an m_v vector of random variables with an expected value $E(\mathbf{v}) = \bar{\mathbf{v}}$, and assuming these vectors are related by a known $m_y \times m_v$ transformation matrix \mathbf{A} :

$$\mathbf{y} = \mathbf{A} \mathbf{v} \quad (8.18)$$

Then the covariance matrix \mathbf{R}_y of the transformed vector \mathbf{y} is

$$\mathbf{R}_y = E\{(\mathbf{y} - \bar{\mathbf{y}})(\mathbf{y} - \bar{\mathbf{y}})'\} = E\{\mathbf{A}(\mathbf{v} - \bar{\mathbf{v}})(\mathbf{v} - \bar{\mathbf{v}})' \mathbf{A}'\} = \mathbf{A} \mathbf{R}_v \mathbf{A}' \quad (8.19)$$

where $\mathbf{R}_v = E\{(\mathbf{v} - \bar{\mathbf{v}})(\mathbf{v} - \bar{\mathbf{v}})'\}$ is the covariance matrix of vector \mathbf{v} .

In view of Eq. (8.7), Eqs. (8.18)-(8.19) can be used to obtain the covariance matrix of the state estimate:

$$\mathbf{R}_{\hat{\mathbf{x}}} = (\mathbf{G}^{-1} \mathbf{H}' \mathbf{W}) \mathbf{R}_z (\mathbf{G}^{-1} \mathbf{H}' \mathbf{W})' \quad (8.20)$$

which, considering $\mathbf{W} = \mathbf{R}_z^{-1}$, yields

$$\mathbf{R}_{\hat{\mathbf{x}}} = \mathbf{G}^{-1} \quad (8.21)$$

The covariance matrix of the estimate of the measurement vector, $\hat{\mathbf{z}}$, is given by

$$\mathbf{R}_{\hat{\mathbf{z}}} = \mathbf{H} \mathbf{R}_{\hat{\mathbf{x}}} \mathbf{H}' = \mathbf{H} \mathbf{G}^{-1} \mathbf{H}' \quad (8.22)$$

In the same way, the covariance matrix of the estimate of the residual vector, $\hat{\mathbf{r}}$, is obtained:

$$\mathbf{R}_{\hat{\mathbf{r}}} = (\mathbf{I} - \mathbf{H} \mathbf{G}^{-1} \mathbf{H}' \mathbf{W}) \mathbf{R}_z (\mathbf{I} - \mathbf{H} \mathbf{G}^{-1} \mathbf{H}' \mathbf{W})' \quad (8.23)$$

This yields

$$\mathbf{R}_{\hat{\mathbf{r}}} = \mathbf{R}_z - \mathbf{H} \mathbf{G}^{-1} \mathbf{H}' = (\mathbf{I} - \mathbf{H} \mathbf{G}^{-1} \mathbf{H}' \mathbf{R}_z^{-1}) \mathbf{R}_z = \mathbf{S} \mathbf{R}_z \quad (8.24)$$

The covariance matrix of the residuals results from the difference between the covariance matrix of the measurements and that of the measurement estimates, i.e.

$$\mathbf{R}_{\hat{\mathbf{r}}} = \mathbf{R}_z - \mathbf{R}_{\hat{\mathbf{z}}} \quad (8.25)$$

This property is important in bad data detection and identification. For any critical (non redundant) measurement, the variance of the measurement estimate will be the same as the variance of the measurement itself; hence, the variance of the corresponding residual will be equal to zero (Note also that in this case state estimation yields a measurement estimate that will be equal to the measured value, and thus the corresponding estimation residual will be equal to zero.)

Example 8.1:

Consider the three-bus example in Fig. 8.1 with measurements of $P_1^{meas} = 3.90$ p.u., $P_2^{meas} = -4.07$ p.u., $P_3^{meas} = -0.04$ p.u., and $P_{13}^{meas} = 2.04$ p.u., and

variances of $\sigma_{P_1}^2 = .004$ p.u., $\sigma_{P_2}^2 = .004$ p.u., $\sigma_{P_3}^2 = .001$ p.u., and $\sigma_{P_{13}}^2 = .002$ p.u.

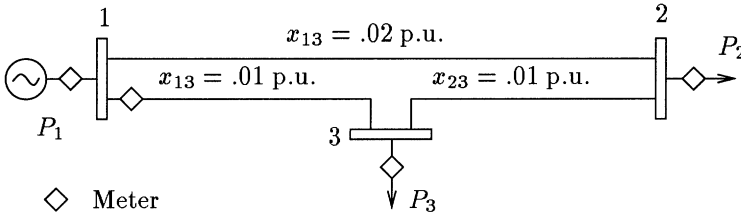


Figure 8.1. Three-bus system.

If θ_2 and θ_3 are considered to be state variables, the Jacobian matrix is as follows (see Example 3.6, Chap. 3):

$$\mathbf{H} = \begin{matrix} & \theta_2 & \theta_3 \\ \begin{matrix} P_1 \\ P_2 \\ P_3 \\ P_{13} \end{matrix} & \begin{pmatrix} -50 & -100 \\ 150 & -100 \\ -100 & 200 \\ 0 & -100 \end{pmatrix} \end{matrix}$$

The covariance matrix of the measurement vector is

$$\mathbf{R}_z = \begin{pmatrix} \sigma_{P_1}^2 & 0 & 0 & 0 \\ 0 & \sigma_{P_2}^2 & 0 & 0 \\ 0 & 0 & \sigma_{P_3}^2 & 0 \\ 0 & 0 & 0 & \sigma_{P_{13}}^2 \end{pmatrix} = 10^{-3} \begin{pmatrix} 4 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}$$

and that of the state estimate is

$$\mathbf{R}_{\hat{\mathbf{x}}} = \mathbf{G}^{-1} = (\mathbf{H}' \mathbf{R}_z^{-1} \mathbf{H})^{-1} = 10^{-7} \begin{pmatrix} 1.633 & 0.735 \\ 0.735 & 0.531 \end{pmatrix}$$

The covariance matrix of the measurement estimate is

$$\mathbf{R}_{\hat{\mathbf{z}}} = \mathbf{H} \mathbf{R}_{\hat{\mathbf{x}}} \mathbf{H}' = 10^{-3} \begin{pmatrix} 1.67 & -1.43 & -0.24 & 0.90 \\ -1.43 & 2.00 & -0.57 & -0.57 \\ -0.24 & -0.57 & 0.81 & -0.33 \\ 0.90 & -0.57 & -0.33 & 0.53 \end{pmatrix}$$

and that of the residuals is

$$\mathbf{R}_{\hat{\mathbf{r}}} = \mathbf{R}_z - \mathbf{R}_{\hat{\mathbf{z}}} = 10^{-3} \begin{pmatrix} 2.33 & 1.43 & 0.24 & -0.90 \\ 1.43 & 2.00 & 0.57 & 0.57 \\ 0.24 & 0.57 & 0.18 & 0.33 \\ -0.90 & 0.57 & 0.33 & 1.47 \end{pmatrix}$$

Remarks: Note that in this case the number of residual degrees of freedom is $\nu = m - n = 2$; thus, the rank of the covariance matrix $\mathbf{R}_{\hat{\mathbf{r}}}$ is also $\nu = 2$. This means that arbitrary values can be assigned to two of the residuals, whereas the other two will be determined by Eqs. (8.10) and (8.11), i.e.,

$$\hat{\mathbf{r}} = \left(\mathbf{I} - \mathbf{H} \mathbf{G}^{-1} \mathbf{H}' \mathbf{W} \right) \mathbf{z}$$

Example 8.2:

Consider again the three-bus system discussed above now with the following set of measurements: $P_2^{meas} = -4.07$ p.u., $P_{13}^{meas} = 2.04$ p.u. and $P_{31}^{meas} = -1.90$ p.u. The corresponding variances being: $\sigma_{P_2}^2 = .004$ p.u., $\sigma_{P_{13}}^2 = .002$ p.u., and $\sigma_{P_{31}}^2 = .002$ p.u.

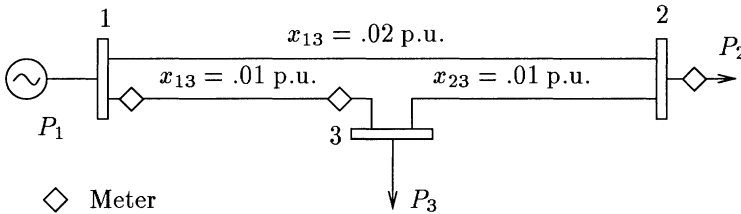


Figure 8.2. Modified three-bus system.

In this case, the Jacobian matrix is

$$\mathbf{H} = \begin{matrix} & \theta_2 & \theta_3 \\ \begin{matrix} P_2 \\ P_{13} \\ P_{31} \end{matrix} & \begin{pmatrix} 150 & -100 \\ 0 & -100 \\ 0 & 100 \end{pmatrix} \end{matrix}$$

The covariance matrix of the measurement vector is

$$\mathbf{R}_{\mathbf{z}} = \begin{pmatrix} \sigma_{P_2}^2 & 0 & 0 \\ 0 & \sigma_{P_{13}}^2 & 0 \\ 0 & 0 & \sigma_{P_{31}}^2 \end{pmatrix} = 10^{-3} \begin{pmatrix} 4 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

and that of the state estimate is

$$\mathbf{R}_{\hat{\mathbf{x}}} = \mathbf{G}^{-1} = (\mathbf{H}' \mathbf{R}_{\mathbf{z}}^{-1} \mathbf{H})^{-1} = 10^{-7} \begin{pmatrix} 2.222 & 0.667 \\ 0.667 & 1.000 \end{pmatrix}$$

The covariance matrix of the measurement estimate is

$$\mathbf{R}_{\hat{\mathbf{z}}} = \mathbf{H} \mathbf{R}_{\hat{\mathbf{x}}} \mathbf{H}' = 10^{-3} \begin{matrix} & P_2 & P_{13} & P_{31} \\ P_2 & \begin{pmatrix} 4 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{pmatrix} \\ P_{13} & & & \\ P_{31} & & & \end{matrix}$$

and that of the residuals is

$$\mathbf{R}_{\hat{\mathbf{r}}} = \mathbf{R}_{\mathbf{z}} - \mathbf{R}_{\hat{\mathbf{z}}} = 10^{-3} \begin{matrix} & P_2 & P_{13} & P_{31} \\ P_2 & \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix} \\ P_{13} & & & \\ P_{31} & & & \end{matrix}$$

Remarks: In this case, measurement P_2 is critical so that the variance of the corresponding estimate, \hat{P}_2 , is the same as that of the original measurement, P_2 , which makes the variance of the associated residual equal to zero. As noted above, for critical measurements in general, the estimated value is always equal to the measured value, which makes the corresponding residual equal to zero.

8.3 NORMALIZED RESIDUALS

This section shows how under certain conditions such as those found for both single or multiple noninteracting bad data, grossly erroneous measurements can be identified using the largest normalized residual test (the largest \mathbf{r}^n -test) .

8.3.1 Covariance Matrix of the Normalized Residuals

For a noncritical (redundant) measurement z_i , the normalized residual r_i^n can be defined as the ratio of the residual estimate, $\hat{r}_i = z_i - \hat{z}_i$, to the standard deviation of the residual, ρ_{ii} . Thus, the vector of normalized residuals is given by

$$\mathbf{r}^n = (\mathbf{diag}(\mathbf{R}_{\hat{\mathbf{r}}}))^{-1/2} \hat{\mathbf{r}} \quad (8.26)$$

where $\mathbf{diag}(\mathbf{R}_{\hat{\mathbf{r}}})$ is the matrix formed by the diagonal elements of the covariance matrix $\mathbf{R}_{\hat{\mathbf{r}}}$. The covariance matrix of the normalized residuals can then be written as follows

$$\mathbf{R}_{\mathbf{r}^n} = (\mathbf{diag}(\mathbf{R}_{\hat{\mathbf{r}}}))^{-1/2} \mathbf{R}_{\hat{\mathbf{r}}} (\mathbf{diag}(\mathbf{R}_{\hat{\mathbf{r}}}))^{-1/2} \quad (8.27)$$

or

$$\mathbf{R}_{\mathbf{r}^n} = \begin{pmatrix} 1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & 1 & \cdot & \frac{\rho_{ij}^2}{\rho_{ii}\rho_{jj}} & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \frac{\rho_{ij}^2}{\rho_{ii}\rho_{jj}} & \cdot & 1 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 1 \end{pmatrix}$$

where ρ_{ij}^2 is the (i, j) th element of $\mathbf{R}_{\hat{\mathbf{r}}}$.

8.3.2 Correlation Coefficients

If \bar{r}_i and \bar{r}_j are the expected values of residuals r_i and r_j , respectively, the corresponding variances are given by

$$E\{(r_i - \bar{r}_i)^2\} = \rho_{ii}^2$$

$$E\{(r_j - \bar{r}_j)^2\} = \rho_{jj}^2$$

Since for any scalar a , the expected value

$$E\{(a(r_i - \bar{r}_i) - (r_j - \bar{r}_j))^2\}$$

is nonnegative, it follows that

$$(\rho_{ii}^2) a^2 - (2 \rho_{ij}^2) a + \rho_{jj}^2 \geq 0$$

where the discriminant is nonpositive (i.e., there are no real roots):

$$(2 \rho_{ij}^2)^2 - 4 \rho_{ii}^2 \rho_{jj}^2 \leq 0$$

This yields

$$\frac{|\rho_{ij}^2|}{\rho_{ii} \rho_{jj}} \leq 1$$

where ρ_{ii}^2 and ρ_{jj}^2 are the variances of residual r_i and r_j , respectively.

The coefficient of correlation between the residual estimates \hat{r}_i and \hat{r}_j is defined as

$$\gamma_{ij} = \frac{\rho_{ij}^2}{\rho_{ii} \rho_{jj}}$$

and these coefficients γ_{ij} , $i, j = 1, \dots, m$, are used for writing the covariance matrix of the normalized residuals, i.e.

$$\mathbf{R}_{r^n} = \begin{pmatrix} 1 & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ . & . & 1 & . & \gamma_{ij} & . & . \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ . & . & \gamma_{ij} & . & 1 & . & . \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & 1 \end{pmatrix}$$

Since $\gamma_{ij}^2 \leq 1$, this is a diagonal-dominant matrix.

Example 8.3:

Consider the system discussed in Example 8.1. The covariance matrix of the normalized residuals here is given by

$$\begin{aligned} \mathbf{R}_{r^n} &= (\text{diag}(\mathbf{R}_{\hat{r}}))^{-1/2} \mathbf{R}_{\hat{r}} (\text{diag}(\mathbf{R}_{\hat{r}}))^{-1/2} \\ &= \begin{pmatrix} 1.00 & 0.66 & 0.37 & -0.49 \\ 0.66 & 1.00 & 0.94 & 0.33 \\ 0.37 & 0.94 & 1.00 & 0.63 \\ -0.49 & 0.33 & 0.63 & 1.00 \end{pmatrix} \end{aligned}$$

All four measurements here are redundant and the variances of the measurement estimates are smaller than those of the corresponding actual measurements, i.e., the variances of the residual are positive (see Eq. (8.25)).

For the system in Example 8.2, there is one nonredundant measurement P_2 , although the other two measurements P_{13} and P_{31} are redundant. The variance of the residual of P_2 is zero, although that of the two latter (P_{13} and P_{31}) are positive. In this case, the normalization of the covariance matrix of the residuals, expressed by Eq. (8.26), is limited to the redundant measurements:

$$\mathbf{R}_{r^n} = \begin{matrix} & P_2 & P_{13} & P_{31} \\ \begin{matrix} P_2 \\ P_{13} \\ P_{31} \end{matrix} & \begin{pmatrix} * & * & * \\ * & 1 & 1 \\ * & 1 & 1 \end{pmatrix} \end{matrix}$$

where the elements of the row/column associated with the nonredundant measurement are indicated by *.

8.3.3 The Largest Normalized Residual Test

Now consider a different situation in which all the measurements are perfect except measurement z_i which consists of bad data (a grossly erroneous measurement). If $z_i^{meas} = z_i^{true} + b_i\sigma_i$, where b_i is the magnitude of the bad data measured in standard deviations σ_i , represents this situation, the measurement model can be written as

$$\mathbf{z} = \mathbf{H} \mathbf{x}^{true} + b_i \sigma_i \mathbf{u}_i \tag{8.28}$$

where \mathbf{u}_i is a null vector, except for the i th element, which is a unit. In this case the vector of estimation residuals is given by

$$\hat{\mathbf{r}} = \mathbf{S} \mathbf{z} = (\mathbf{I} - \mathbf{H} \mathbf{G}^{-1} \mathbf{H}' \mathbf{W}) \mathbf{z} = \mathbf{R}_{\hat{\mathbf{r}}} \mathbf{W} \mathbf{z}$$

Since the other measurements (unperturbed measurements) are assumed to be perfect, this yields

$$\hat{\mathbf{r}} = \mathbf{R}_{\hat{\mathbf{r}}} \mathbf{W} (b_i \sigma_i \mathbf{u}_i) = b_i \sigma_i^{-1} \mathbf{R}_{\hat{\mathbf{r}}} \mathbf{u}_i$$

where $\mathbf{R}_{\hat{\mathbf{r}}} \mathbf{u}_i$ is the i th column in the covariance matrix of the residuals, $\mathbf{R}_{\hat{\mathbf{r}}}$.

The vector of normalized residuals is given by

$$\mathbf{r}^n = (\text{diag}(\mathbf{R}_{\hat{\mathbf{r}}}))^{-1/2} \hat{\mathbf{r}} = b_i \sigma_i^{-1} (\text{diag}(\mathbf{R}_{\hat{\mathbf{r}}}))^{-1/2} \mathbf{R}_{\hat{\mathbf{r}}} \mathbf{u}_i$$

which can be rewritten as

$$\mathbf{r}^n = b_i \sigma_i^{-1} \begin{pmatrix} \rho_{11}^2 \rho_{11}^{-1} \\ \cdot \\ \cdot \\ \rho_{ii} \\ \cdot \\ \cdot \\ \rho_{jj}^2 \rho_{jj}^{-1} \\ \cdot \\ \cdot \\ \rho_{mi}^2 \rho_{mm}^{-1} \end{pmatrix} \tag{8.29}$$

The ratio of the magnitude of the normalized residual r_j^n , $j \neq i$, to the magnitude of the normalized residual associated with the measurement affected by bad data, r_i^n , is given by

$$\frac{|r_j^n|}{|r_i^n|} = \frac{|\rho_{ji}^2|}{\rho_{jj} \rho_{ii}} = |\gamma_{ji}| \leq 1$$

This finally yields

$$|r_j^n| \geq |r_i^n| \quad \text{for } j = 1, \dots, m$$

This means that the wrong measurement is responsible for the largest normalized residual. Although other measurements with the same residual magnitude may exist, none will have a residual larger than that of the i th measurement (the measurement assumed to be bad data).

Example 8.4:

Consider the three-bus example in Fig. 8.1 with perfect measurements, $P_1^{meas} = 4.00$ p.u., $P_2^{meas} = -4.00$ p.u., $P_3^{meas} = 0.00$ p.u., and $P_{13}^{meas} = 2.00$

p.u., and variances $\sigma_{P_1}^2 = .004$ p.u., $\sigma_{P_2}^2 = .004$ p.u., $\sigma_{P_3}^2 = .001$ p.u., and $\sigma_{P_{13}}^2 = .002$ p.u.

The measurement vector estimate is given by

$$\hat{\mathbf{z}} = \mathbf{H} \hat{\mathbf{x}} = \mathbf{H} \mathbf{G}^{-1} \mathbf{H}' \mathbf{W} \mathbf{z} = \mathbf{R}_{\hat{\mathbf{z}}} \mathbf{W} \mathbf{z}$$

Using the covariance matrix of the measurement estimates obtained in Example 8.1, yields

$$\hat{\mathbf{z}} = 10^{-3} \begin{pmatrix} 1.67 & -1.43 & -0.24 & 0.90 \\ -1.43 & 2.00 & -0.57 & -0.57 \\ -0.24 & -0.57 & 0.81 & -0.33 \\ 0.90 & -0.57 & -0.33 & 0.53 \end{pmatrix} \begin{pmatrix} 250 & 0 & 0 & 0 \\ 0 & 250 & 0 & 0 \\ 0 & 0 & 1000 & 0 \\ 0 & 0 & 0 & 500 \end{pmatrix} \begin{pmatrix} 4.00 \\ -4.00 \\ 0.00 \\ 2.00 \end{pmatrix}$$

Then the measurement and residual vector estimates are, respectively,

$$\hat{\mathbf{z}} = \begin{pmatrix} 4.00 \\ -4.00 \\ 0.00 \\ 2.00 \end{pmatrix} \quad \text{and} \quad \hat{\mathbf{r}} = \mathbf{z} - \hat{\mathbf{z}} = \begin{pmatrix} 0.00 \\ 0.00 \\ 0.00 \\ 0.00 \end{pmatrix}$$

If P_1 is a bad data of 5.00 p.u. instead of the correct value of 4.00 p.u., the estimation residuals, the weighted residuals and the normalized residuals will be, respectively,

$$\hat{\mathbf{r}} = \begin{pmatrix} 0.58 \\ 0.36 \\ 0.06 \\ -0.22 \end{pmatrix}, \quad \mathbf{r}^w = \begin{pmatrix} 9.20 \\ 5.65 \\ 1.94 \\ -5.02 \end{pmatrix} \quad \text{and} \quad \mathbf{r}^n = \begin{pmatrix} 12.06 \\ 7.99 \\ 4.52 \\ -5.86 \end{pmatrix}$$

Note that the measurement affected, P_1 , is represented not only by the largest residual, but also by the largest weighted residual, and the largest normalized residual. This is not always the case however. If the bad data had been for measurement P_3 , with a reading of 1.00 p.u. rather than the correct value of 0.00 p.u., the following residuals would result:

$$\hat{\mathbf{r}} = \begin{pmatrix} 0.24 \\ 0.57 \\ 0.18 \\ 0.32 \end{pmatrix}, \quad \mathbf{r}^w = \begin{pmatrix} 3.87 \\ 9.04 \\ 5.81 \\ 7.30 \end{pmatrix} \quad \text{and} \quad \mathbf{r}^n = \begin{pmatrix} 5.08 \\ 12.78 \\ 13.55 \\ 8.52 \end{pmatrix}$$

In this case, neither the residual itself nor the weighted residual would indicate the bad data (P_3), although the normalized residual would still suggest its existence, as is to be expected from the property discussed above.

Example 8.5:

Consider now the three-bus example in Fig. 8.2 with the following measurements (affected by normal, random errors): $P_2^{meas} = -4.07$ p.u., $P_{13}^{meas} = 2.04$

p.u. and $P_{31}^{meas} = -1.90$ p.u. The corresponding variances are $\sigma_{P_2}^2 = .004$ p.u., $\sigma_{P_{13}}^2 = .002$ p.u., and $\sigma_{P_{31}}^2 = .002$ p.u.

The measurement vector estimate is given by

$$\hat{\mathbf{z}} = \mathbf{R}_{\hat{\mathbf{z}}} \mathbf{W} \mathbf{z} = 10^{-3} \begin{pmatrix} 4 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{pmatrix} \begin{pmatrix} 250 & 0 & 0 \\ 0 & 500 & 0 \\ 0 & 0 & 500 \end{pmatrix} \begin{pmatrix} -4.07 \\ 2.04 \\ -1.90 \end{pmatrix} = \begin{pmatrix} -4.07 \\ 1.97 \\ -1.97 \end{pmatrix}$$

(The covariance matrix $\mathbf{R}_{\hat{\mathbf{z}}}$ was obtained in Example 8.2)

The estimation residuals, the weighted residuals and the normalized residuals are as follows

$$\hat{\mathbf{r}} = \begin{pmatrix} 0.00 \\ 0.07 \\ 0.07 \end{pmatrix}, \quad \mathbf{r}^w = \begin{pmatrix} 0.00 \\ 1.57 \\ 1.57 \end{pmatrix} \quad \text{and} \quad \mathbf{r}^n = \begin{pmatrix} * \\ 2.21 \\ 2.21 \end{pmatrix}$$

Remarks: Note that the residual of measurement P_2 is zero, and, since it is a nonredundant measurement, the variance of the corresponding residual is also equal to zero. Thus a normalized residual is not defined for this measurement. Measurements P_{13} and P_{31} have the same residual magnitude for this meter configuration, even if one of them (P_{13} or P_{31}) consists of bad data, both residuals will be the same. In cases such as this, the available redundancy is said to allow for bad data detection although its identification is not possible.

8.3.4 Dormant Measurements and Error Estimates

Consider the linear state estimator presented in Eq. (8.1). The sensitivity relation (8.12) can be written as

$$\hat{\mathbf{r}} = \mathbf{S} \mathbf{z}$$

If z_j is a suspect measurement, the j th equation of $\hat{\mathbf{r}} = \mathbf{S} \mathbf{z}$ is as follows

$$\hat{r}_j = \mathbf{S}_{j\bullet} \mathbf{z} \quad \text{for} \quad j = 1, \dots, m$$

If a correction c_j is added to the suspect measured value, the result is a new measurement vector:

$$\mathbf{z}^{new} = \mathbf{z} + c_j \mathbf{u}_j$$

where, except for the j th element, which is equal to one, \mathbf{u}_j is a null vector. The new value of the j th residual is

$$\hat{r}_j^{new} = \mathbf{S}_{j\bullet} (\mathbf{z} + c_j \mathbf{u}_j)$$

An attempt must now be made to determine the value of the correction c_j which will make the suspect measurement dormant, i.e., eliminate its effect on

the new state estimate. To do this, \hat{r}_j^{new} is assumed to be zero in the above equation, yielding

$$0 = \mathbf{S}_{j\bullet} (\mathbf{z} + c_j \mathbf{u}_j) = \hat{r}_j + c_j S_{jj}$$

where S_{jj} is the (j, j) th element of \mathbf{S} . Assuming that $S_{jj} \neq 0$, which is always true for redundant measurements, this finally yields

$$c_j = -\frac{\hat{r}_j}{S_{jj}} \quad (8.30)$$

This correction can now be used as an error estimate, i.e.,

$$\hat{e}_j = -c_j = \hat{b}_j \sigma_j = \frac{\hat{r}_j}{S_{jj}} = \frac{\sigma_j^2 \hat{r}_j}{\rho_{jj}^2} \quad (8.31)$$

$$\hat{b}_j = \frac{\hat{r}_j}{\sigma_j S_{jj}} = \frac{\sigma_j \hat{r}_j}{\rho_{jj}^2} \quad (8.32)$$

where, in view of Eq. (8.24), $\mathbf{R}_{\hat{\mathbf{r}}} = \mathbf{S} \mathbf{R}_{\mathbf{z}}$ and $\rho_{jj}^2 = S_{jj} \sigma_j^2$.

Remarks:

- Consider the following least-squares estimator

$$\begin{aligned} J(e_j) &= (\hat{\mathbf{r}} - \mathbf{S}_{\bullet j} e_j)' \mathbf{R}_{\mathbf{z}}^{-1} (\hat{\mathbf{r}} - \mathbf{S}_{\bullet j} e_j) \\ &= (\hat{\mathbf{r}}^w - \mathbf{S}_{\bullet j}^w e_j^w)' (\hat{\mathbf{r}}^w - \mathbf{S}_{\bullet j}^w e_j^w) \end{aligned} \quad (8.33)$$

where $\hat{\mathbf{r}}$ is the residual vector, $\mathbf{S}_{\bullet j}$ is the j th column of \mathbf{S} , and e_j is the gross error estimate best explaining the residual vector obtained. Optimality conditions are the same as in Eq. 8.3, i.e.

$$\hat{e}_j^w = ((\mathbf{S}_{\bullet j}^w)' \mathbf{S}_{\bullet j}^w)^{-1} (\mathbf{S}_{\bullet j}^w)' \hat{\mathbf{r}}$$

In view of Eq. (8.15), $r_j^w = (\mathbf{S}_{\bullet j}^w)' \hat{\mathbf{r}}$, and in view of Eq. 8.17, $S_{jj}^w = (\mathbf{S}_{\bullet j}^w)' \mathbf{S}_{\bullet j}^w$. The above equation thus yields the following error estimate:

$$\hat{e}_j^w = \frac{r_j^w}{S_{jj}^w} \quad (8.34)$$

- These results can be interpreted in statistical terms using the measurement model

$$\mathbf{z} = \mathbf{H} \mathbf{x}^{true} + \mathbf{e} + b_j \sigma_j \mathbf{u}_j$$

where each element e_j of the error vector \mathbf{e} follows a distribution $N(0, \sigma_i^2)$. In this case, Eq. (8.32) becomes $\hat{b}_j = \frac{\sigma_j}{\rho_{jj}^2} E\{r_j\}$, where \hat{b}_j is the total bias estimate and $E\{r_j\}$ is the expected value of the residual. Since it is assumed that $E\{e_j\} = 0$, then $b_j = 0$ implies that $E\{r_j\} = 0$. In practice, only a single evaluation of the residual \hat{r}_j is available and this estimate can be used as an approximation for the expected value of the residual in the presence of a biased measurement, $E\{r_j\}$. This turns out to be equivalent to estimating the total error affecting measurement z_j rather than estimating the bias $b_j \sigma_j$.

- Although Eqs. (8.31)-(8.32) involve only the diagonal elements of matrix \mathbf{S} , the error estimate \hat{e}_j is exact for cases where bad data consists of single or noninteracting multiple bad data, i.e., if state estimation is run with the corrected measurement z_j^{new} , the corresponding residual will be zero ($\hat{r}_j^{new} = 0$), as is to be expected from the above derivations.

8.3.5 Updating the $J(\hat{\mathbf{x}})$ Index

Consider that the correction in Eq. (8.30) has actually been made. Now calculate the impact of this correction on the performance index. The index corresponding to the original state estimate is

$$J(\hat{\mathbf{x}}) = (\mathbf{r}^w)' \mathbf{r}^w$$

where \mathbf{r}^w is the vector of weighted residuals. The new value for the index is as follows

$$J(\hat{\mathbf{x}}^{new}) = (\mathbf{r}^w + \Delta \mathbf{r}^w)' (\mathbf{r}^w + \Delta \mathbf{r}^w)$$

In view of Eq. (8.34), the correction c_j^w is given by:

$$c_j^w = -\frac{r_j^w}{S_{jj}^w}$$

This yields the following change in the vector of residuals:

$$\Delta \mathbf{r}^w = \mathbf{S}_{\bullet j}^w c_j^w = -\mathbf{S}_{\bullet j}^w \frac{r_j^w}{S_{jj}^w}$$

Hence,

$$J(\hat{\mathbf{x}}^{new}) = (\mathbf{r}^w)' \mathbf{r}^w - \frac{2r_j^w}{S_{jj}^w} (\mathbf{r}^w)' \mathbf{S}_{\bullet j}^w + \frac{(r_j^w)^2}{(S_{jj}^w)^2} (\mathbf{S}_{\bullet j}^w)' \mathbf{S}_{\bullet j}^w$$

Since

$$(\mathbf{r}^w)' \mathbf{S}_{\bullet j}^w = r_j^w$$

and

$$(\mathbf{S}_{\bullet j}^w)' \mathbf{S}_{\bullet j}^w = S_{jj}^w$$

this yields,

$$J(\hat{\mathbf{x}}^{new}) = J(\hat{\mathbf{x}}) - \frac{(r_j^w)^2}{S_{jj}^w}$$

Finally, considering that

$$S_{jj}^w = S_{jj} = \rho_{jj}^2 \sigma_j^{-2}$$

and

$$r_j^n = \sigma_j^{-1} \hat{r}_j$$

the following is obtained:

$$J(\hat{\mathbf{x}}^{new}) = J(\hat{\mathbf{x}}) - (r_j^n)^2$$

Remarks: For all cases of single or non-interacting multiple bad data, the removal of the data presenting the largest normalized residual will eliminate the effect of the suspect data and provide the maximum reduction in the performance index. Hence, in these cases, detecting and identifying bad data by the largest normalized residual criterion turns out to be the same as selecting the measurement whose removal would produce the largest reduction in the performance index.

Example 8.6:

Consider the three-bus example in Fig. 8.1 with the following set of measurements: $P_1^{meas} = 3.90$ p.u., $P_2^{meas} = -4.07$ p.u., $P_3^{meas} = -1.04$ p.u., and $P_{13}^{meas} = 2.04$ p.u., and variances $\sigma_{P_1}^2 = .004$ p.u., $\sigma_{P_2}^2 = .004$ p.u., $\sigma_{P_3}^2 = .001$ p.u., and $\sigma_{P_{13}}^2 = .002$ p.u.

The value of the objective function at the optimal estimate is

$$J(\hat{\mathbf{x}}) = 223.0$$

and the estimated residuals are

$$\hat{\mathbf{r}} = \mathbf{S}\mathbf{z} = \begin{pmatrix} 0.5816 & 0.3571 & 0.2449 & -0.4490 \\ 0.3571 & 0.5000 & 0.5714 & 0.2857 \\ 0.0612 & 0.1428 & 0.1837 & 0.1632 \\ -0.2245 & 0.1428 & 0.3265 & 0.7347 \end{pmatrix} \begin{pmatrix} 3.90 \\ -4.07 \\ -1.04 \\ 2.04 \end{pmatrix} = \begin{pmatrix} -0.356 \\ -0.654 \\ -0.201 \\ -0.298 \end{pmatrix}$$

The normalized residuals are

$$\mathbf{r}^n = (\text{diag}(\mathbf{R}_{\hat{\mathbf{r}}}))^{-1/2} \hat{\mathbf{r}} = (-7.377 \quad -14.614 \quad -14.802 \quad -7.767)'$$

Now, considering that $\sigma_3 = \sigma_{P_3} = 0.0316$ p.u., $\rho_{33}^2 = \rho_{P_3, P_3}^2 = 0.184$ p.u. (see Example 8.1), and that $\hat{r}_3 = \hat{r}_{P_3} = -0.201$ p.u., Eqs. (8.31)-(8.32) yields

$$\hat{e}_3 = \frac{\hat{r}_3}{S_{33}} = \frac{-0.201}{0.184} = -1.09 \text{ p.u.}$$

$$\hat{b}_3 = \frac{\hat{e}_3}{\sigma_3} = \frac{-1.09}{0.0316} = -34.5$$

Remarks:

- The true value of measurement P_3 is unknown. The value affected by random, normal error is -0.04 p.u., but when both random and gross errors are involved, it is -1.04 : the error estimate is -1.09 . Thus, the measurement estimate, assuming that this is actually involves bad data, and that all the remaining data are good data, is given by

$$\tilde{P}_3 = P_3^{meas} - \hat{e}_3 = -1.04 - (-1.09) = 0.05 \text{ p.u.}$$

- Performing state estimation with corrected measurements, yields

$$J(\hat{\mathbf{x}}^{new}) = J(\hat{\mathbf{x}}) - (r_3^n)^2 = 3.9$$

$$\mathbf{r}^{n,new} = (-1.831 \quad -0.658 \quad -0.000 \quad 1.536)'$$

- Running state estimation for the original set of measurements has yielded the following residuals

$$\mathbf{r}^n = (-7.38, ; -14.61, ; -14.80, ; -7.78)'$$

Since P_3^{meas} presents the largest normalized residual it was considered to be bad data. The residual of P_2^{meas} is very close to the largest normalized residual, however. What if it had been classified as bad data instead? In this case, the error estimate would be $\hat{b}_2 = -20.6$ ($\hat{e}_2 = -1.30$ p.u.). Performing state estimation with the corrected measurement, $\tilde{P}_2 = P_2^{meas} - \hat{e}_2$, would yield:

$$J(\hat{\mathbf{x}}^{new}) = J(\hat{\mathbf{x}}) - (r_2^n)^2 = 9.4$$

$$\mathbf{r}^{n,new} = (2.302 \quad 0.000 \quad -1.024 \quad -2.896)'$$

This shows that the choice of P_3 as bad data leads to a better solution, both in terms of overall fitness, index $J(\hat{\mathbf{x}})$, and maximum normalized residual in the remaining set of measurements.

8.4 HYPOTHESES TESTING

In state estimation, the two main problems are to obtain estimates $\hat{\mathbf{x}}$ and $\hat{\mathbf{z}}$ and to test hypotheses concerning them, i.e., to check whether these estimates are compatible in accuracy to their standard deviations. This section discusses two fundamental hypothesis testing approaches: the $J(\hat{\mathbf{x}})$ -test and the \mathbf{r}^n -test. The former is based on the χ^2 distribution, and the latter on the standard normal distribution.

8.4.1 The $J(\hat{\mathbf{x}})$ Detection Test

If x_i , $i = 1, \dots, m$, are random independent variables which follow a normal distribution with mean zero and unit variance (distribution $N(0, 1)$), the χ_m^2 distribution, i.e., the chi-square distribution with m degrees of freedom, is the distribution of the random variable y defined as $y = \sum_{i=1}^m x_i^2$. It can be shown that the χ_m^2 distribution has mean m and variance $2m$. When random variables x_i are constrained by n independent equations, then y follows a χ_{m-n}^2 distribution with the number of degrees of freedom reduced by the number of constraints.

In view of the Central Limit Theorem, the sum of a large number of random variables that follow any distribution with bounded variance, approximates a normal distribution. As the number of degrees of freedom, ν , increases, the χ_ν^2 distribution also tends to a normal distribution. Figure 8.3 illustrates the χ_ν^2 probability density function for three different cases: $\nu = 2$, $\nu = 4$, and $\nu = 8$.

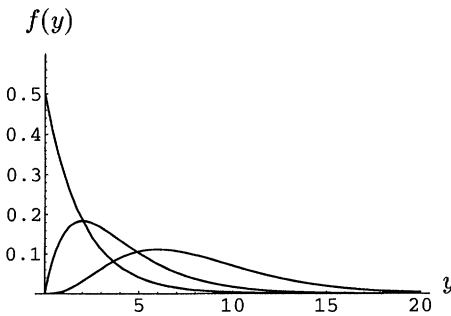


Figure 8.3. Probability density function for the χ_ν^2 distribution, chi-square distribution with ν degrees of freedom. (From left to right: $\nu = 2$, $\nu = 4$, and $\nu = 8$.)

Assuming that the measurement errors e_i , $i = 1, \dots, m$, are normal and independent, each having mean zero and variance σ_i^2 , $N(0, \sigma_i^2)$, then the performance index $J(\hat{\mathbf{x}})$, is

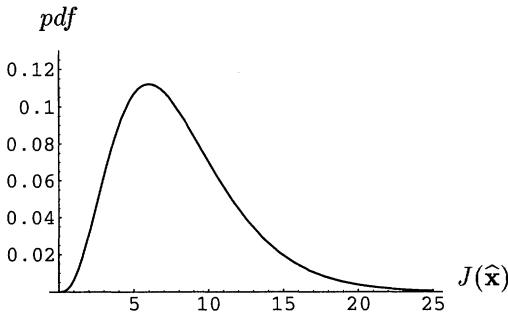


Figure 8.4. Probability density function for the χ^2_8 distribution (chi-square distribution with 8 degrees of freedom): $\mu = 8$ (mean); $\sigma = 4$ (standard deviation); $C = \chi^2_{8,0.95} = 15.5$ (95 % quantile, i.e. $\text{Prob}(\chi^2 > 15.5) = 0.05$).

$$J(\hat{\mathbf{x}}) = \sum_{i=1}^m \left(\frac{z_i^{meas} - \hat{z}_i}{\sigma_i} \right)^2$$

and follows a χ^2_{m-n} distribution, i.e., a chi-square distribution with $m - n$ degrees of freedom, with m being the number of measurements and n the number of state variables. Under these conditions, the expected value of $J(\hat{\mathbf{x}})$ and its variance are as follows: $E\{J(\hat{\mathbf{x}})\} = m - n$ and $E\{[J(\hat{\mathbf{x}}) - (m - n)]^2\} = 2(m - n)$. Note that for an observable system with $m = n$, the state estimate $\hat{\mathbf{x}}$ fits the measurement model perfectly, i.e., all the residuals will be zero and $J(\hat{\mathbf{x}}) = 0$.

A state estimation run yields only one observation $J(\hat{\mathbf{x}})$ of the random variable $J(\mathbf{x})$. Based on this observation, the decision as to whether or not it actually belongs to the hypothesized χ^2 -distribution must be made. The hypotheses made regarding the behavior of the random variables e_i will also be tested – indirectly – for normal distribution, $N(0, \sigma_i^2)$, or the effect of gross errors or biases.

The first step is the consideration of a null hypothesis H_0 and an alternative hypothesis H_1 : $H_0, E\{J(\hat{\mathbf{x}})\} = m - n$; and $H_1, E\{J(\hat{\mathbf{x}})\} > m - n$. The alternative hypothesis suggests the way to perform the test, i.e.,

- If $J(\hat{\mathbf{x}}) > C$ then reject the hypothesis H_0 .
- If $J(\hat{\mathbf{x}}) \leq C$ then accept the hypothesis H_0 .

where C is a constant to be determined. If α is the significance level of the test, then constant C is given by:

$$C = \chi_{m-n, 1-\alpha}^2$$

This means that if H_0 is true, the probability of $J(\hat{\mathbf{x}}) > C$ is α (or $\alpha \times 100\%$), i.e.,

$$\int_0^C f(t) dt = 1 - \alpha \quad \text{with} \quad f(t) = \frac{t^{\frac{\nu}{2}-1} e^{-\frac{t}{2}}}{2^{\frac{\nu}{2}} \Gamma(\frac{\nu}{2})}$$

where $f(t)$ is the probability density function of the χ_ν^2 distribution, with $\nu = m - n$ degrees of freedom, and Γ is the Gamma function. (Figure 8.4 gives the basic data concerning a chi-square distribution with $\nu = 8$ degrees of freedom.)

It should be noted that the significance level α is also equivalent to the probability of a false alarm, or the probability of the occurrence of a Type 1 error:

- Type 1 error : reject H_0 when it is actually true
- Type 2 error : accept H_0 when it is actually false

Example 8.7:

Consider the three-bus example in Fig. 8.1 in which $m = 4$, $n = 2$ and $\nu = m - n = 2$. The corresponding probability density function, pdf , is presented in Fig. 8.5: for a 2.5 % probability of false alarm ($\alpha = 0.025$), the distribution χ_2^2 results in $C = 7.4$. The measurement variances are $\sigma_{P_1}^2 = .004$ p.u., $\sigma_{P_2}^2 = .004$ p.u., $\sigma_{P_3}^2 = .001$ p.u., and $\sigma_{P_{13}}^2 = .002$ p.u. Four different sets of measurements (in p.u.) are given here:

- Case 1: $P_1^{meas} = 4.08$, $P_2^{meas} = -3.96$, $P_3^{meas} = -0.02$, and $P_{13}^{meas} = 2.02$.

In this case $J(\hat{\mathbf{x}}) = 1.1 < 7.4$, and the null hypothesis is accepted (i.e. no bad data are detected). Note that $J(\hat{\mathbf{x}})$ is even smaller than the expected value $m - n = 2$. The normalized residual vector is

$$\mathbf{r}^n = (0.97 \quad 0.96 \quad 0.75 \quad -0.11)'$$

If the measurement with the largest normalized residual were to be considered suspect, the corresponding error estimate would be (Eqs. (8.31)-(8.32))

$$\hat{b}_1 = \frac{\sigma_1}{\rho_{11}} r_1^n = 1.28 \quad \text{and} \quad \hat{e}_1 = \hat{b}_1 \sigma_1 = 0.08 \text{ p.u.}$$

- Case 2: $P_1^{meas} = 4.18$, $P_2^{meas} = -3.96$, $P_3^{meas} = -0.02$, and $P_{13}^{meas} = 2.02$.

In this case $J(\hat{\mathbf{x}}) = 4.9 < 7.4$, and as above, the null hypothesis is accepted. The normalized residual vector is

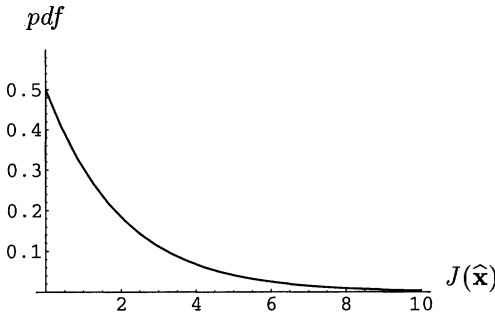


Figure 8.5. Probability density function for the χ^2_2 distribution: $\mu = 2$ (mean); $\sigma = 2$ (standard deviation); $\chi^2_{2,0.975} = 7.4$ (i.e. $\text{Prob}(\chi^2 > 7.4) = 0.025$).

$$\mathbf{r}^n = (2.18, 1.76, 1.20, -0.69)'$$

If, for example, measurement P_1 were held to be spurious, the corresponding error estimate would be

$$\hat{b}_1 = 2.86 \quad \text{and} \quad \hat{e}_1 = 0.18 \text{ p.u.}$$

- Case 3: $P_1^{meas} = 4.28, P_2^{meas} = -3.96, P_3^{meas} = -0.02,$ and $P_{13}^{meas} = 2.02;$
 In this case $J(\hat{\mathbf{x}}) = 11.6 > 7.4$. The null hypothesis is rejected because bad data are detected. Note that for $\nu = 2, \text{Prob}(J(\hat{\mathbf{x}}) > 11.6) = 0.0030$ or 0.30 %. The normalized residual vector is

$$\mathbf{r}^n = (3.38, 2.56, 1.66, -1.28)'$$

and the error estimate corresponding to the measurement with the largest residual is

$$\hat{b}_1 = 4.44 \quad \text{and} \quad \hat{e}_1 = 0.28 \text{ p.u.}$$

- Case 4: $P_1^{meas} = 4.38, P_2^{meas} = -3.96, P_3^{meas} = -0.02,$ and $P_{13}^{meas} = 2.02.$
 In this case $J(\hat{\mathbf{x}}) = 21.2 > 7.4$, which also implies the presence of bad data. In this case $\text{Prob}(J(\hat{\mathbf{x}}) > 21.2) = 0.000025$ or 0.0025 %. The normalized residual vector is

$$\mathbf{r}^n = (4.59, 3.35, 2.11, -1.86)'$$

and the error estimate corresponding to the measurement with the largest residual is

$$\hat{b}_1 = 6.02 \quad \text{and} \quad \hat{e}_1 = 0.38 \text{ p.u.}$$

8.4.2 The \mathbf{r}^n Detection and Identification Test

As above, the components of the vector of measurement errors, e_i , $i = 1, \dots, m$, are assumed to be normal and independent, each having mean zero and variance σ_i^2 . Since the corresponding residual estimates \hat{r}_i are given by linear combinations of the measurement errors, i.e., $\hat{r}_i = \mathbf{S}_{i\bullet} \mathbf{e}$, they are also normally distributed with mean zero and variance ρ_{ii}^2 , where $\rho_{ii}^2 = \mathbf{R}_{\hat{\mathbf{r}}}(i, i)$ (the diagonal elements of the covariance matrix of the residuals). Hence, the marginal distribution of each normalized residuals $r_i^n = \rho_{ii}^{-1} \hat{r}_i$, is normal with mean zero and unit variance, i.e., $E\{r_i^n\} = 0$ and $E\{(r_i^n)^2\} = 1$.

A state estimation run yields only one observation \hat{r}_i of the random variable r_i . Based on this observation a decision must be made as to whether this observation actually belongs to the hypothesized normal distribution or is affected by gross errors.

Two hypotheses are considered: $H_0: E\{r_i^n\} = 0$ and $H_1: E\{|r_i^n|\} > 0$. H_0 is rejected if $|r_i^n| > C$, where C is a constant to be determined. The significance level of the test is given by α and corresponds to the probability of false alarm. Hence,

$$C = K_{\frac{\alpha}{2}}$$

This means that α is the probability that $|r_i^n| > C$, i.e.,

$$\int_{-C}^C f(t) dt = 1 - \alpha \quad \text{with} \quad f(t) = \frac{e^{-\frac{t^2}{2}}}{\sqrt{2\pi}}$$

where $f(t)$ is the probability density function of the standard normal distribution with mean zero and unit variance. Hence C is a function of α . In order to use the tables for the cumulative standard normal distribution function it is more convenient to write the above equation as,

$$\int_{-\infty}^C f(t) dt = 1 - \frac{\alpha}{2}$$

For example, for $\alpha = 0.025$ (2.5 %), $C = K_{0.0125} = 2.24$, and for $\alpha = 0.01$ (1.0 %), $C = K_{0.005} = 2.57$.

Example 8.8:

Consider again the three-bus example in Fig. 8.1. For a 2.5 % probability of false alarm ($\alpha = 0.025$), the distribution $N(0, 1)$ gives $C = 2.24$. The measurement variances are: $\sigma_{P_1}^2 = .004$ p.u., $\sigma_{P_2}^2 = .004$ p.u., $\sigma_{P_3}^2 = .001$ p.u.,

and $\sigma_{P_{13}}^2 = .002$ p.u. Four different sets of measurements (in p.u.) are analyzed below:

- Case 1: $P_1^{meas} = 4.08$, $P_2^{meas} = -3.96$, $P_3^{meas} = -0.02$, and $P_{13}^{meas} = 2.02$.
In this case,

$$\mathbf{r}^n = (0.97, 0.96, 0.75, -0.11)'$$

The largest normalized residual is $|r_1| = 0.97 < 2.24$. Hence the null hypothesis is accepted, as no bad data have been detected.

- Case 2: $P_1^{meas} = 4.18$, $P_2^{meas} = -3.96$, $P_3^{meas} = -0.02$, and $P_{13}^{meas} = 2.02$.
In this case,

$$\mathbf{r}^n = (2.18, 1.76, 1.20, -0.69)'$$

The largest normalized residual is $|r_1| = 2.18 < 2.24$. Hence the null hypothesis is again accepted, as no bad data were detected.

- Case 3: $P_1^{meas} = 4.28$, $P_2^{meas} = -3.96$, $P_3^{meas} = -0.02$, and $P_{13}^{meas} = 2.02$;
In this case,

$$\mathbf{r}^n = (3.38, 2.56, 1.66, -1.28)'$$

The largest normalized residual is $|r_1| = 3.38 > 2.24$. Hence, the null hypothesis is rejected, and measurement P_1^{meas} is considered to be bad data.

- Case 4: $P_1^{meas} = 4.38$, $P_2^{meas} = -3.96$, $P_3^{meas} = -0.02$, and $P_{13}^{meas} = 2.02$.
In this case,

$$\mathbf{r}^n = (4.59, 3.35, 2.11, -1.86)'$$

The largest normalized residual is $|r_1| = 4.59 > 2.24$. Hence, the null hypothesis is again rejected, and measurement P_1^{meas} is considered to be bad data.

8.5 HISTORICAL NOTES AND REFERENCES

The estimation of measurement errors based on the residual vector and on elements of sensitivity matrix \mathbf{S} , and the correction of measured values for reestimation purposes was introduced by Aboytes and Cory [1975], and the approach further studied by Duran [1977] and Monticelli, Garcia and Abreu [1979] for the sequential processing of bad data ordered by the largest normalized residual criterion. Error estimates were used in conjunction with normalized residuals

for the detection/identification of bad data (Monticelli and Garcia [1983]). The interpretation of the residual estimates as the result of a state estimation based on parts of the sensitivity relation $\mathbf{S} \hat{\mathbf{r}} = \mathbf{e}$ was introduced by Xiang, Wang, and Yu [1981, 82] and applied to an hypotheses testing procedure designed to detect and identify interacting multiple bad data (Mili, Van Cutsem, and Ribbens-Pavella [1984]). The largest normalized residual method was extended to the normal equations approach with equality constraints by Wu, Liu and Lun [1988] and to the blocked sparse matrix approach by Nucera, Brandwajn and Gilles [1993].

8.6 PROBLEMS

- 1. Consider a linear least-squares state estimation with residuals given by

$$\hat{\mathbf{r}} = \mathbf{S} \mathbf{z} = (\mathbf{I} - \mathbf{H} \mathbf{G}^{-1} \mathbf{H}' \mathbf{W}) \mathbf{z}$$

Show that

$$(\mathbf{H} \mathbf{G}^{-1} \mathbf{H}' \mathbf{W}) \hat{\mathbf{r}} = \mathbf{0}$$

- 2. Show that sensitivity matrix in Eq. 8.11,

$$\mathbf{S} = \partial \hat{\mathbf{r}} / \partial \mathbf{z} = \mathbf{I} - \mathbf{H} \mathbf{G}^{-1} \mathbf{H}' \mathbf{W}$$

is an idempotent matrix, i.e.,

$$\mathbf{S}^2 = \mathbf{S}$$

- 3. Verify that the two alternative formulae for the performance index $J(e_j)$ in Eq. (8.33) are equivalent.
- 4. Verify that for an observable system with zero redundancy the residual sensitivity matrix \mathbf{S} is a null matrix.
- 5. Consider a dc power flow model for the system of Fig. 8.6 in which all the reactances are 0.01 p.u. and all the power flow measurements read 1.0 p.u., except for one of the measurements between Buses 1 and 2 which reads $(1.0 + \epsilon)$ p.u., where ϵ is a variable error. All the meters have error variances of 0.03 p.u. (a) Determine the bad data detection threshold for the $J(\hat{\mathbf{x}})$ considering a probability of false alarm of 5.0 %. (b) determine the magnitude of the maximum error ϵ that will pass undetected through the $J(\hat{\mathbf{x}})$ -test defined above. (c) Consider the fact that this error does not propagate through the whole system and define a new detection threshold for the subsystem formed only by the flow measurements placed between Buses 1 and 2. (d) Discuss how this type of behavior may affect the performance of the $J(\hat{\mathbf{x}})$ -test for large networks.

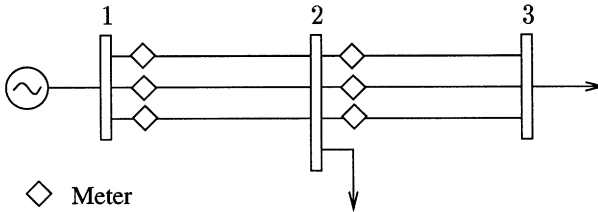


Figure 8.6. Three-bus system used to illustrate error propagation.

References

- Aboytes, F. and Cory, B.J., "Identification of measurement and configuration errors in static estimation", Proc. 9th Power Industry Computer Application Conference, New Orleans, June 1975, pp. 298-302.
- Bose, A. and Clements, K.A., "Real-time modeling of power networks", IEEE Proc., Special Issue on Computers in Power System Operations, Vol. 75, No. 12, pp. 1607-1622, Dec. 1987.
- Duran, H., "Surrogate measurements make faster state estimation optimal and general", Paper A 77 598-6, IEEE PES Summer Meeting, Mexico City, 1977.
- Garcia, A., Monticelli, A., and Abreu, P., "Fast decoupled state estimation and bad data processing", IEEE Trans. Power App. and Syst., Vol. 98, No. 5, pp. 1645-1652, Sept./Oct. 1979.
- Mili, L. Van Cutsem, Th., and Ribbens-Pavella, M., "Hypothesis testing identification: A new method for bad data analysis in power system state estimation", IEEE Trans. Power App. and Syst., Vol. 103, No. 11, pp. 3239-3252, Nov. 1984.
- Monticelli, A. and Garcia, A., "Reliable bad data processing for real time state estimation", IEEE Trans. Power App. and Syst., Vol. 102, No. 5, pp. 1126-1139, May 1983.
- Nucera, R.R., Brandwajn, V., and Gilles, M.L., "Observability analysis and bad data - Analysis using augmented blocked matrices", IEEE Trans. Power Syst., Vol. 8, No. 2, pp. 426-433, May 1993.
- Xiang Nian-De, Wang Shi-Ying, and Yu Er-Keng, "Estimation and identification of multiple bad data in power system state estimation", Proc. 7th Power System Computation Conference, PSCC, Lausanne, July 1981.
- Xiang Nian-De, Wang Shi-Ying, and Yu Er-Keng, "A new approach for detection and identification of multiple bad data in power system state estimation", IEEE Trans. Power App. and Syst., Vol. 101, No. 2, pp. 454-462, Feb. 1982.
- Wu, F.F., Liu, E.H.E, and Lun, S.M. "Observability analysis and bad data-processing for state estimation with equality constraints", IEEE Trans. Power Syst., Vol. 3, No. 2, pp. 541-578, May 1988.

9 MULTIPLE BAD DATA PROCESSING TECHNIQUES

This chapter discusses techniques for processing interacting multiple bad data. This discussion includes the generalization of the concepts presented in the previous chapter in connection with single and multiple noninteracting bad data. Parameter and status errors as well as their interaction with analog bad data are also discussed.

9.1 ESTIMATION RESIDUALS

Consider the linear state estimation measurement model

$$\mathbf{z} = \mathbf{H} \mathbf{x} + \mathbf{e}$$

Assuming that the measurement errors e_i , $i = 1, \dots, m$, are normal, independent with zero mean and variance σ_i^2 , the vector of estimation residuals

$$\hat{\mathbf{r}} = \mathbf{S} \mathbf{e}$$

follows a ν -variate normal distribution; $\nu = m - n$ is the number of degrees of freedom, where m is the number of measurements and n is the number of state variables.

9.1.1 Partition of Measurement Set

Let \mathbf{r}^o denote a $o \times 1$ vector of residuals, with $o \geq n$, corresponding to a selected set of measurements that assures network observability (If, for example, $o = n$, the system is minimally observable, i.e., there are no redundant measurements.) And $\hat{\mathbf{r}}_k$ denotes a $k \times 1$ vector of residuals, with $k \leq m - n$, corresponding to a set of redundant measurements. Given this partition, the sensitivity relation $\hat{\mathbf{r}} = \mathbf{S}\mathbf{e}$ can be written as

$$\begin{pmatrix} \hat{\mathbf{r}}_k \\ \hat{\mathbf{r}}_o \end{pmatrix} = \begin{pmatrix} \mathbf{S}_{k\bullet} \\ \mathbf{S}_{o\bullet} \end{pmatrix} (\mathbf{e})$$

where submatrices $\mathbf{S}_{k\bullet}$ and $\mathbf{S}_{o\bullet}$ are formed by the rows of the sensitivity matrix corresponding to measurements k and o , respectively.

Since the sensitivity matrix \mathbf{S} is idempotent, i.e., $\mathbf{S} = \mathbf{S}^2$, then

$$\begin{pmatrix} \mathbf{S}_{kk} & \mathbf{S}_{ko} \\ \mathbf{S}_{ok} & \mathbf{S}_{oo} \end{pmatrix} = \begin{pmatrix} \mathbf{S}_{k\bullet} \\ \mathbf{S}_{o\bullet} \end{pmatrix} \begin{pmatrix} \mathbf{S}_{\bullet k} & \mathbf{S}_{\bullet o} \end{pmatrix}$$

where $\mathbf{S}_{\bullet k}$ and $\mathbf{S}_{\bullet o}$ are formed by the columns of the matrix \mathbf{S} , corresponding to measurements k and o , respectively.

Hence, $\mathbf{S}_{kk} = \mathbf{S}_{k\bullet}\mathbf{S}_{\bullet k}$. A similar relation holds for the weighted sensitivity matrix $\mathbf{S}^w = \mathbf{R}_z^{-1/2} \mathbf{S} \mathbf{R}_z^{1/2}$, i.e., $\mathbf{S}_{kk}^w = \mathbf{S}_{k\bullet}^w \mathbf{S}_{\bullet k}^w$ ($\mathbf{R}_z = \mathbf{W}^{-1}$ is the measurement covariance matrix).

9.1.2 Distribution Function

The vector of multiple residuals $\hat{\mathbf{r}}_k$ follows a k -variate normal distribution with the following probability density function,

$$f(\hat{\mathbf{r}}_k) = (2\pi)^{-\frac{k}{2}} |\text{Var}\{\hat{\mathbf{r}}_k\}|^{-\frac{1}{2}} e^{-\frac{1}{2}[\hat{\mathbf{r}}_k]' (\text{Var}\{\hat{\mathbf{r}}_k\})^{-1} \hat{\mathbf{r}}_k}$$

where $|\text{Var}\{\hat{\mathbf{r}}_k\}|$ is the determinant of the covariance matrix of $\hat{\mathbf{r}}_k$, and $\text{Var}\{\hat{\mathbf{r}}_k\}$ is given by,

$$\begin{aligned} \text{Var}\{\hat{\mathbf{r}}_k\} &= E\{\hat{\mathbf{r}}_k (\hat{\mathbf{r}}_k)'\} = \mathbf{S}_{k\bullet} E\{\mathbf{e} \mathbf{e}'\} \mathbf{S}_{\bullet k}' = \mathbf{S}_{k\bullet} \mathbf{R}_z \mathbf{S}_{\bullet k}' \\ &= \mathbf{R}_{z,k} \mathbf{S}_{k\bullet}^w (\mathbf{S}_{k\bullet}^w)'\mathbf{R}_{z,k}^{1/2} = \mathbf{R}_{z,k}^{1/2} \mathbf{S}_{kk}^w \mathbf{R}_{z,k}^{1/2} = \mathbf{S}_{kk} \mathbf{R}_{z,k} \\ &= \mathbf{R}_{kk} \end{aligned}$$

where \mathbf{R}_{kk} is the submatrix of the covariance matrix of residuals corresponding to the k redundant measurements, i.e.,

$$\mathbf{R}_{\hat{\mathbf{r}}} = \begin{pmatrix} \mathbf{R}_{kk} & \mathbf{R}_{ko} \\ \mathbf{R}_{ok} & \mathbf{R}_{oo} \end{pmatrix}$$

and $\mathbf{R}_{z,k}$ is the diagonal matrix of the variances of the k measurements.

Example 9.1:

Consider the three-bus example in Fig. 9.2 with variances of $\sigma_{P_1}^2 = .004$ p.u., $\sigma_{P_2}^2 = .004$ p.u., $\sigma_{P_3}^2 = .001$ p.u., $\sigma_{P_{12}}^2 = .002$ p.u., and $\sigma_{P_{13}}^2 = .002$ p.u.

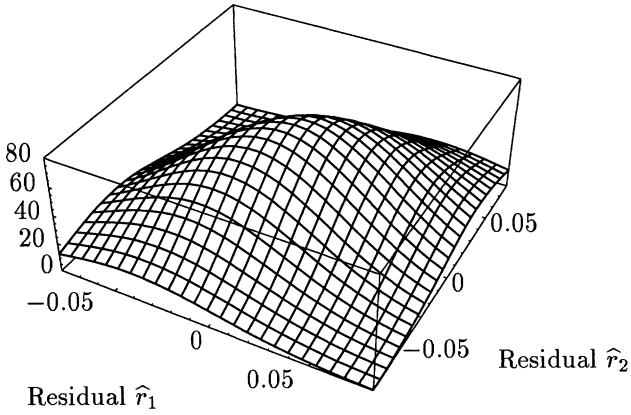


Figure 9.1. Joint marginal distribution function for the vector of multiple residuals $\hat{\mathbf{r}}_k$ in Example 9.1.

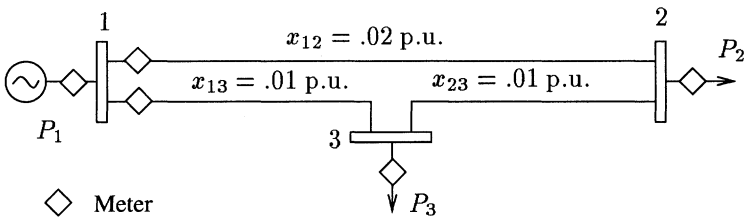


Figure 9.2. Three-bus system.

The Jacobian matrix for θ_2 and θ_3 as the state variables, and the covariance matrix of the measurement vector are, as follows:

$$\mathbf{H} = \begin{matrix} P_1 \\ P_2 \\ P_3 \\ P_{12} \\ P_{13} \end{matrix} \begin{pmatrix} \theta_2 & \theta_3 \\ -50 & -100 \\ 150 & -100 \\ -100 & 200 \\ -50 & 0 \\ 0 & -100 \end{pmatrix}$$

$$\mathbf{R}_z = 10^{-3} \begin{pmatrix} 4 & 0 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{pmatrix}$$

The covariance matrix of the residuals

$$\mathbf{R}_{\hat{\mathbf{r}}} = \mathbf{R}_z - \mathbf{H} (\mathbf{H}' \mathbf{R}_z^{-1} \mathbf{H})^{-1} \mathbf{H}'$$

in this example is as follows:

$$\mathbf{R}_{\hat{\mathbf{r}}} = 10^{-3} \begin{pmatrix} 2.576 & 1.152 & 0.271 & -0.644 & -0.779 \\ 1.152 & 2.305 & 0.542 & 0.711 & 0.440 \\ 0.271 & 0.542 & 0.186 & -0.067 & 0.338 \\ -0.644 & 0.711 & -0.067 & 1.661 & -0.305 \\ -0.779 & 0.440 & 0.338 & -0.305 & 1.525 \end{pmatrix}$$

Consider, for example, the following partition of the measurement set:

$$\text{Set}\{o\} = \{P_3, P_{12}, P_{13}\}$$

$$\text{Set}\{k\} = \{P_1, P_2\}$$

In this case matrix \mathbf{R}_{kk} and its inverse are, as follows:

$$\mathbf{R}_{kk} = 10^{-3} \begin{pmatrix} 2.576 & 1.152 \\ 1.152 & 2.305 \end{pmatrix} \quad \mathbf{R}_{kk}^{-1} = \begin{pmatrix} 500 & -250 \\ -250 & 558.8 \end{pmatrix}$$

and the probability density function, illustrated in Fig. 9.1, is given by

$$f(\hat{r}_1, \hat{r}_2) = \frac{465.7}{2\pi} e^{-\frac{1}{2}[500 \hat{r}_1^2 + 558.8 \hat{r}_2^2 - 500 \hat{r}_1 \hat{r}_2]}$$

where \hat{r}_1 and \hat{r}_2 are the elements of the multiple residual vector $\mathbf{r}_k = (\hat{r}_1; \hat{r}_2)'$.

Remarks: In this case the eigenvalues of matrix \mathbf{R}_{kk}^{-1} are

$$\lambda_1 = 781.1$$

$$\lambda_2 = 277.7$$

and the corresponding eigenvectors are

$$\mathbf{v}_1 = (-0.6645, 0.7472)'$$

$$\mathbf{v}_2 = (0.7472, 0.6645)'$$

The estimated residuals tend to align along eigenvector \mathbf{v}_2 , as shown in Fig. 9.1; eigenvector \mathbf{v}_1 provides the direction that passes through the flat regions in the front and in the back of the figure where the probability of finding residuals is relatively lower.

9.1.3 Dormant Measurements and Error Estimates

Consider now that all the measurements, except those in the k -set, are perfect, i.e., $\mathbf{e} = (\mathbf{e}_k \ \mathbf{e}_o)'$, with $\mathbf{e}_o = \mathbf{0}$. In this case the sensitivity relation $\hat{\mathbf{r}} = \mathbf{S}\mathbf{z}$ can be written as

$$\begin{pmatrix} \hat{\mathbf{r}}_k \\ \hat{\mathbf{r}}_o \end{pmatrix} = \begin{pmatrix} \mathbf{S}_{kk} & \mathbf{S}_{ko} \\ \mathbf{S}_{ok} & \mathbf{S}_{oo} \end{pmatrix} \begin{pmatrix} \mathbf{e}_k \\ \mathbf{0} \end{pmatrix}$$

As with the situations involving single and multiple non-interacting bad data, discussed in the previous chapter, an error estimate is given by

$$\hat{\mathbf{e}}_k = \mathbf{S}_{kk}^{-1} \hat{\mathbf{r}}_k \tag{9.1}$$

This error estimate can be used to correct the measurements in the k -set, yielding

$$\begin{pmatrix} \mathbf{S}_{kk} & \mathbf{S}_{ko} \\ \mathbf{S}_{ok} & \mathbf{S}_{oo} \end{pmatrix} \begin{pmatrix} \mathbf{e}_k - \mathbf{S}_{kk}^{-1} \hat{\mathbf{r}}_k \\ \mathbf{0} \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \end{pmatrix}$$

The correction in Eq. (9.1) zeroizes the estimation residuals corresponding to the k measurements. And since all the other measurements are assumed to be perfect, the residuals of the measurements of the o -set become zero as well. In general, if the o -set measurements are affected by errors, then only the residuals corresponding to the k -set measurements become zero after measurement correction.

Example 9.2:

Consider again the system presented in Example 9.1 (Fig. 9.2), with measurements $P_1^{meas} = 3.33$ p.u., $P_2^{meas} = -4.60$ p.u., $P_3^{meas} = -0.02$ p.u., $P_{12}^{meas} = 2.02$ p.u. and $P_{13}^{meas} = 1.96$ p.u., and variances $\sigma_{P_1}^2 = .004$ p.u., $\sigma_{P_2}^2 = .004$ p.u., $\sigma_{P_3}^2 = .001$ p.u., $\sigma_{P_{12}}^2 = .002$ p.u., and $\sigma_{P_{13}}^2 = .002$ p.u. Consider also the following partition of the measurement set :

$$\text{Set}\{o\} = \{P_3, P_{12}, P_{13}\}$$

$$\text{Set}\{k\} = \{P_1, P_2\}$$

In this case, the sensitivity matrix

$$\mathbf{S} = \mathbf{I}_5 - \mathbf{H} (\mathbf{H}' \mathbf{R}_z^{-1} \mathbf{H})^{-1} \mathbf{H}' \mathbf{R}_z^{-1}$$

is

$$\mathbf{S} = \begin{pmatrix} 0.644 & 0.288 & 0.271 & -0.322 & -0.389 \\ 0.288 & 0.576 & 0.542 & 0.355 & 0.220 \\ 0.067 & 0.135 & 0.186 & -0.033 & 0.169 \\ -0.161 & 0.177 & -0.067 & 0.830 & -0.152 \\ -0.194 & 0.110 & 0.338 & -0.152 & 0.762 \end{pmatrix}$$

The vector of estimation residuals $\hat{\mathbf{r}} = \mathbf{S} \mathbf{z}$ and the sensitivity submatrix \mathbf{S}_{kk} , are as follows:

$$\hat{\mathbf{r}} = \begin{pmatrix} -0.600 \\ -0.551 \\ -0.137 \\ 0.025 \\ 0.024 \end{pmatrix}$$

$$\mathbf{S}_{kk} = \begin{pmatrix} 0.644 & 0.288 \\ 0.288 & 0.576 \end{pmatrix}$$

The error estimate is determined in Eq. (9.1),

$$\hat{\mathbf{e}}_k = \mathbf{S}_{kk}^{-1} \hat{\mathbf{r}}_k = \begin{pmatrix} 2.000 & -1.000 \\ -1.000 & 2.235 \end{pmatrix} \begin{pmatrix} -0.600 \\ -0.551 \end{pmatrix} = \begin{pmatrix} -0.650 \\ -0.631 \end{pmatrix}$$

and the corrected measurements are as follows:

$$P_1^{new} = P_1^{meas} - \hat{e}_1 = 3.33 - (-0.65) = 3.98 \text{ p.u.}$$

$$P_2^{new} = P_2^{meas} - \hat{e}_2 = -4.60 - (-0.63) = -3.97 \text{ p.u.}$$

9.2 MULTIPLE NORMALIZED RESIDUALS

In this section the largest normalized residual criterion, developed for situations involving single and multiple non-interacting bad data, is extended to interacting multiple bad data.

9.2.1 Distribution Function

Again, consider a partition (o, k) of the measurement set, with $o + k = m$, $o \geq n$, such that the system is observable with the measurements of the o -set. Since the inverse of the covariance matrix \mathbf{R}_{kk}^{-1} is real and symmetric, it can be factorized as follows (similarity transformation):

$$\mathbf{R}_{kk}^{-1} = \mathbf{A}' \text{diag}(\lambda) \mathbf{A} \quad (9.2)$$

where \mathbf{A} is an orthogonal $k \times k$ matrix formed by the eigenvectors of \mathbf{R}_{kk}^{-1} and $\mathbf{diag}(\lambda)$ is a diagonal matrix with the corresponding eigenvalues.

The vector of multiple normalized residuals is

$$\mathbf{r}_k^n = \mathbf{diag}(\lambda^{1/2}) \mathbf{A} \hat{\mathbf{r}}_k \tag{9.3}$$

and its variance matrix is

$$\begin{aligned} \text{Var}\{\mathbf{r}_k^n\} &= E\{\mathbf{r}_k^n (\mathbf{r}_k^n)'\} = E\{\mathbf{diag}(\lambda^{1/2}) \mathbf{A} \hat{\mathbf{r}}_k \hat{\mathbf{r}}_k' \mathbf{A}' \mathbf{diag}(\lambda^{1/2})\} \\ &= \mathbf{diag}(\lambda^{1/2}) \mathbf{A} E\{\hat{\mathbf{r}}_k \hat{\mathbf{r}}_k'\} \mathbf{A}' \mathbf{diag}(\lambda^{1/2}) \\ &= \mathbf{diag}(\lambda^{1/2}) \mathbf{A} \mathbf{R}_{kk} \mathbf{A}' \mathbf{diag}(\lambda^{1/2}) \\ &= \mathbf{diag}(\lambda^{1/2}) \mathbf{A} (\mathbf{A}' \mathbf{diag}(\lambda^{-1}) \mathbf{A}) \mathbf{A}' \mathbf{diag}(\lambda^{1/2}) \\ &= \mathbf{I}_k \end{aligned}$$

(Note that $\mathbf{A}^{-1} = \mathbf{A}'$; see Appendix A) Hence, the elements of \mathbf{r}_k^n are normal, with zero mean and unit variance, and they are uncorrelated. The corresponding probability density function is

$$f(\mathbf{r}_k^n) = (2\pi)^{-\frac{k}{2}} e^{-\frac{1}{2}[(\mathbf{r}_k^n)'\mathbf{r}_k^n]}$$

Remarks: The decomposition presented in Eq. (9.2) is not unique, as matrix \mathbf{R}_{kk}^{-1} can be decomposed as $\mathbf{R}_{kk}^{-1} = \mathbf{U}'\mathbf{U}$, where \mathbf{U} is an upper triangular matrix, in which case the multiple normalized residual is defined as $\mathbf{r}_k^n = \mathbf{U}\hat{\mathbf{r}}_k$. Note, however, that the variance and the Euclidean norm of the multiple normalized residuals remains the same. Hence, the detection of bad data (discussed below) is based on the Euclidean norm rather than the elements of the multiple normalized residual vector taken individually.

Example 9.3:

Returning to Example 9.1 (Fig. 9.2) with measurements $P_1^{meas} = 3.97$ p.u., $P_2^{meas} = -4.05$ p.u., $P_3^{meas} = -0.05$ p.u., $P_{12}^{meas} = 2.04$ p.u. and $P_{13}^{meas} = 1.98$ p.u., and variances $\sigma_{P_1}^2 = .004$ p.u., $\sigma_{P_2}^2 = .004$ p.u., $\sigma_{P_3}^2 = .001$ p.u., $\sigma_{P_{12}}^2 = .002$ p.u., and $\sigma_{P_{13}}^2 = .002$ p.u. Consider also the following partition of the measurement set: Set $\{o\} = \{P_3, P_{12}, P_{13}\}$ and Set $\{k\} = \{P_1, P_2\}$.

The vector of estimation residuals is

$$\hat{\mathbf{r}} = \begin{pmatrix} -0.0523 \\ -0.0547 \\ -0.0228 \\ 0.0355 \\ 0.0379 \end{pmatrix}$$

The inverse of the covariance matrix \mathbf{R}_{kk} , is

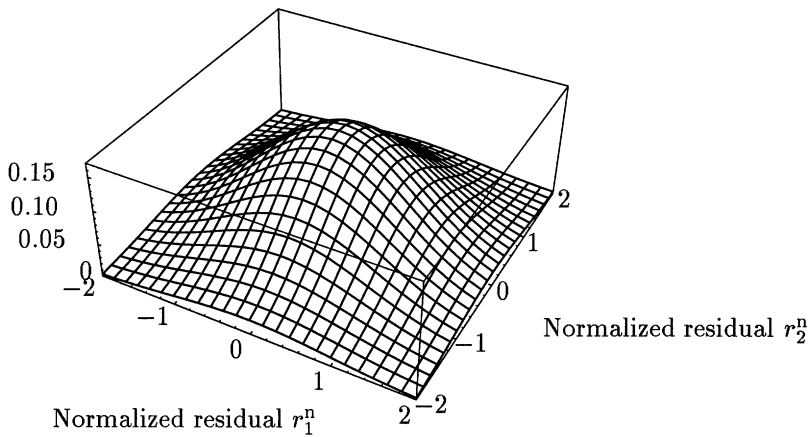


Figure 9.3. Joint marginal distribution function of the vector of normalized residuals $\mathbf{r}^n = (r_1^n \ r_2^n)'$ (2-variate standard normal distribution with zero mean and unit covariance matrix).

$$\mathbf{R}_{kk}^{-1} = 10^3 \begin{pmatrix} 2.576 & 1.152 \\ 1.152 & 2.305 \end{pmatrix}^{-1} = \begin{pmatrix} 500 & -250 \\ -250 & 558.8 \end{pmatrix}$$

The decomposition of matrix \mathbf{R}_{kk}^{-1} (Eq. (9.2)) yields

$$\mathbf{R}_{kk}^{-1} = \begin{pmatrix} -0.6645 & 0.7472 \\ 0.7472 & 0.6645 \end{pmatrix} \begin{pmatrix} 781.1 & 0 \\ 0 & 277.6 \end{pmatrix} \begin{pmatrix} -0.6645 & 0.7472 \\ 0.7472 & 0.6645 \end{pmatrix}$$

Hence, the vector of multiple normalized residuals (Eq. (9.3)) is as follows:

$$\begin{aligned} \mathbf{r}_k^n &= \begin{pmatrix} r_1^n \\ r_2^n \end{pmatrix} = \mathbf{diag}(\lambda^{1/2}) \mathbf{A} \begin{pmatrix} \hat{r}_1 \\ \hat{r}_2 \end{pmatrix} \\ &= \begin{pmatrix} 27.94 & 0 \\ 0 & 16.66 \end{pmatrix} \begin{pmatrix} -0.6645 & 0.7472 \\ 0.7472 & 0.6645 \end{pmatrix} \begin{pmatrix} -0.0523 \\ -0.0547 \end{pmatrix} \end{aligned}$$

$$= \begin{pmatrix} -0.170 \\ -1.258 \end{pmatrix}$$

The corresponding probability density function, illustrated in Fig. 9.3, is given by

$$f(r_1^n, r_2^n) = \frac{1}{2\pi} e^{-\frac{1}{2}[(r_1^n)^2 + (r_2^n)^2]}$$

For $\mathbf{r}_k^n = (-0.170; -1.258)'$, the value is $f(r_1^n, r_2^n) = 0.071$. Considering the inconsistent set (a set with bad data) of measurements $P_1^{meas} = 3.87$ p.u., $P_2^{meas} = -4.15$ p.u., $P_3^{meas} = -0.05$ p.u., $P_{12}^{meas} = 2.04$ p.u. and $P_{13}^{meas} = 1.98$ p.u., gives $\mathbf{r}_k^n = (-0.244; -3.376)'$ and $f(r_2^n, r_2^n) = 0.0005$.

9.2.2 Updating the $J(\hat{\mathbf{x}})$ Index

Consider the correction of errors in Eq. (9.1). The impact of this correction on the performance index $J(\hat{\mathbf{x}})$ is given by

$$\begin{aligned} J(\hat{\mathbf{x}} + \Delta\hat{\mathbf{x}}) &= (\mathbf{r}^w + \Delta\mathbf{r}^w)'(\mathbf{r}^w + \Delta\mathbf{r}^w) \\ &= (\mathbf{r}^w)' \mathbf{r}^w + 2(\mathbf{r}^w)' \Delta\mathbf{r}^w + (\Delta\mathbf{r}^w)' \Delta\mathbf{r}^w \end{aligned}$$

Given Eq. (9.1), the correction vector that zeroizes the vector of multiple residuals, \mathbf{r}_k , is given by

$$\mathbf{c}_k = -(\mathbf{S}_{kk}^w)^{-1} \mathbf{r}_k^w$$

And the impact of this correction on the vector of weighted estimation residuals is represented by

$$\Delta\mathbf{r}^w = -\mathbf{S}_{\bullet k}^w (\mathbf{S}_{kk}^w)^{-1} \mathbf{r}_k^w$$

Hence, the corrected performance index can be successively written as follows:

$$\begin{aligned} J(\hat{\mathbf{x}} + \Delta\hat{\mathbf{x}}) &= (\mathbf{r}^w)' \mathbf{r}^w - 2(\mathbf{r}^w)' \mathbf{S}_{\bullet k}^w \mathbf{e}_k^w + (\mathbf{e}_k^w)' \mathbf{S}_{kk}^w \mathbf{e}_k^w \\ &= (\mathbf{r}^w)' \mathbf{r}^w - 2(\mathbf{r}^w)' \mathbf{S}_{\bullet k}^w (\mathbf{S}_{kk}^w)^{-1} \mathbf{r}_k^w + (\mathbf{r}_k^w)' (\mathbf{S}_{kk}^w)^{-1} \mathbf{r}_k^w \\ &= (\mathbf{r}^w)' \mathbf{r}^w - 2(\mathbf{r}_k^w)' (\mathbf{S}_{kk}^w)^{-1} \mathbf{r}_k^w + (\mathbf{r}_k^w)' (\mathbf{S}_{kk}^w)^{-1} \mathbf{r}_k^w \\ &= (\mathbf{r}^w)' \mathbf{r}^w - (\mathbf{r}_k^w)' (\mathbf{S}_{kk}^w)^{-1} \mathbf{r}_k^w = J(\hat{\mathbf{x}}) - \hat{\mathbf{r}}_k' \mathbf{R}_{kk}^{-1} \hat{\mathbf{r}}_k \\ &= J(\hat{\mathbf{x}}) - \|\mathbf{r}_k\|^2 \end{aligned}$$

Example 9.4:

Consider the 3-bus system in Fig. 9.2 with the same measurement set and the same partition as in Example 9.3. In this case the vector of estimation residuals $\hat{\mathbf{r}}$ and the performance index $J(\hat{\mathbf{x}})$ are, as follows:

$$\hat{\mathbf{r}} = \mathbf{S} \mathbf{z} = \begin{pmatrix} -0.0523 \\ -0.0547 \\ -0.0228 \\ 0.0355 \\ 0.0379 \end{pmatrix}$$

$$J(\hat{\mathbf{x}}) = \hat{\mathbf{r}}' \mathbf{R}_z^{-1} \hat{\mathbf{r}} = 3.31$$

The inverse of covariance matrix \mathbf{R}_{kk} is

$$\mathbf{R}_{kk}^{-1} = \begin{pmatrix} 500 & -250 \\ -250 & 558.8 \end{pmatrix}$$

Hence,

$$\| \mathbf{r}_k \|^2 = \hat{\mathbf{r}}_k' \mathbf{R}_{kk}^{-1} \hat{\mathbf{r}}_k = 1.61$$

Finally, making measurements P_1 and P_2 dormant, yields the following updated index

$$J(\hat{\mathbf{x}} + \Delta\hat{\mathbf{x}}) = J(\hat{\mathbf{x}}) - \| \mathbf{r}_k \|^2 = 3.31 - 1.61 = 1.70$$

9.3 HYPOTHESES TESTING

Figure 9.3 suggests that the Euclidean norm of the vector of the multiple normalized residuals can be used to test hypotheses about the state estimation model: assuming a correct model and normal, independent measurement errors, the probability of observing a given \mathbf{r}_k^n depends only on its magnitude (the distance from the origin in terms in Fig. 9.3). This idea can be used to develop detection and identification tests for cases of interacting multiple bad data.

9.3.1 Detection Test

Since the elements of the vector of multiple normalized residuals, \mathbf{r}_k^n , are normal and independent, with zero mean and unit variance, the square of its Euclidean norm,

$$\| \mathbf{r}_k^n \|^2 = (\mathbf{r}_k^n)' \mathbf{r}_k^n = (\hat{\mathbf{r}}_k)' \mathbf{R}_{kk}^{-1} \hat{\mathbf{r}}_k,$$

follows a χ_k^2 distribution, i.e., a chi-square distribution with k degrees of freedom.

A state estimation run yields an observation $\| \mathbf{r}_k^n \|^2$. Based on this value a decision must be made as to whether or not it actually belongs to the hypothesized χ^2 -distribution. To make this decision, the null hypothesis H_0 ($E\{\| \mathbf{r}_k^n \|^2\} = k$) and the alternative hypothesis H_1 ($E\{\| \mathbf{r}_k^n \|^2\} > k$)

must be considered. It is the alternative hypothesis which suggests a means of performing the test:

- If $\| \mathbf{r}_k^n \|^2 > C$, then reject Hypothesis H_0 .
- If $\| \mathbf{r}_k^n \|^2 \leq C$, then accept Hypothesis H_0 .

where C is determined in terms of the significance level α (or the probability of false alarm). Since $\| \mathbf{r}_k^n \|^2$ follows a χ_k^2 -distribution, C can be expressed as follows:

$$C = \chi_{k,1-\alpha}^2$$

All the measurements, except those in the k -set, are assumed to be perfect, i.e., $\mathbf{e} = (\mathbf{e}_k \ \mathbf{e}_o)'$, $\mathbf{e}_o = \mathbf{0}$. Hence,

$$\begin{aligned} \| \mathbf{r}_k^n \|^2 &= (\hat{\mathbf{r}}_k)' \mathbf{R}_{kk}^{-1} \hat{\mathbf{r}}_k = (\hat{\mathbf{r}}_k)' \mathbf{R}_{z,k}^{-1} \mathbf{S}_{kk}^{-1} \hat{\mathbf{r}}_k = (\hat{\mathbf{e}}_k)' \mathbf{S}'_{kk} \mathbf{R}_{z,k}^{-1} \hat{\mathbf{e}}_k \\ &= (\hat{\mathbf{e}}_k^w)' \mathbf{S}_{kk}^w \hat{\mathbf{e}}_k^w = (\hat{\mathbf{e}}^w)' \mathbf{S}^w \hat{\mathbf{e}}^w = (\hat{\mathbf{e}}^w)' \mathbf{S}^w \mathbf{S}^w \hat{\mathbf{e}}^w = (\hat{\mathbf{r}}^w)' \hat{\mathbf{r}}^w \\ &= \hat{\mathbf{r}}' \mathbf{R}_z^{-1} \hat{\mathbf{r}} \\ &= J(\hat{\mathbf{x}}) \end{aligned}$$

In this case, the square of the Euclidean norm of the vector of multiple normalized residuals corresponding to the measurements affected by errors, $\| \mathbf{r}_k^n \|^2$, is equal to the performance index $J(\hat{\mathbf{x}})$ computed for the entire network.

If the k redundant measurements are corrected using the error estimates given in Eq. (9.1), or if they are removed from the measurement set, the new value of the $J(\hat{\mathbf{x}})$ index will be zero. Hence, the removal of the measurements in error causes the maximum reduction in $J(\hat{\mathbf{x}})$. Since the updated index is $J(\hat{\mathbf{x}}) - \| \mathbf{r}_k^n \|^2 = 0$, the Euclidean norm of the multiple normalized residual corresponding to the bad data is maximum. As mentioned in the previous chapter in connection with single and multiple non-interacting bad data, however, although no other multiple normalized residual can have an Euclidean norm larger than that corresponding to the bad data, other sets of suspect data may present multiple normalized residuals with the same, or nearly the same, Euclidean norm. When this happens there is more than one plausible explanation for the inconsistencies in the data being dealt with.

Example 9.5:

Consider the 3-bus system in Fig. 9.2 with the same measurement set given in Example 9.2. In this case, the vector of estimation residuals $\hat{\mathbf{r}}$ and the performance index $J(\hat{\mathbf{x}})$ are as follows:

$$\hat{\mathbf{r}} = \mathbf{S} \mathbf{z} = \begin{pmatrix} -0.600 \\ -0.551 \\ -0.137 \\ 0.025 \\ 0.024 \end{pmatrix}$$

$$J(\hat{\mathbf{x}}) = \hat{\mathbf{r}}' \mathbf{R}_z^{-1} \hat{\mathbf{r}} = 185.8$$

Suspect data are now identified. Assume, for example, the measurements P_1 and P_2 are suspect. The following partition of the measurement set is thus made:

$$\begin{aligned} \text{Set}\{o\} &= \{P_3, P_{12}, P_{13}\} \\ \text{Set}\{k\} &= \{P_1, P_2\} \end{aligned}$$

The vector of multiple residuals, $\hat{\mathbf{r}}_k$, and the inverse of the corresponding covariance matrix, \mathbf{R}_{kk} , are as follows:

$$\begin{aligned} \hat{\mathbf{r}}_k &= \begin{pmatrix} -0.600 \\ -0.551 \end{pmatrix} \\ \mathbf{R}_{kk}^{-1} &= \begin{pmatrix} 500 & -250 \\ -250 & 558.8 \end{pmatrix} \end{aligned}$$

Finally, the Euclidean norm of the multiple normalized residual, $\mathbf{r}_k^n = (r_1^n \ r_2^n)$, is computed:

$$\|\mathbf{r}_k\|^2 = \hat{\mathbf{r}}_k' \mathbf{R}_{kk}^{-1} \hat{\mathbf{r}}_k = 184.7$$

In the absence of bad data, $\|\mathbf{r}_k\|^2$ follows a χ_k^2 distribution ($k=2$). Considering a probability of false alarm of $\alpha = 0.05$, the detection threshold is $C = \chi_{2,0.95}^2 = 7.4$. Since $\|\mathbf{r}_k\|^2 = 184.7 > 7.4$, the H_0 hypothesis is rejected and bad data is assumed to have been detected.

Since P_1 and P_2 represents bad data, they are made dormant, yielding:

$$J(\hat{\mathbf{x}} + \Delta\hat{\mathbf{x}}) = J(\hat{\mathbf{x}}) - \|\mathbf{r}_k\|^2 = 185.8 - 184.7 = 1.1$$

After the elimination of the suspect measurements, or the elimination of their effect, a system with three measurements and one degree of freedom remains. For a probability of false alarm of $\alpha = 0.05$, the detection threshold of the $J(\hat{\mathbf{x}})$ -test is $\chi_{1,0.95}^2 = 3.8$. Since $J(\hat{\mathbf{x}} + \Delta\hat{\mathbf{x}}) = 1.1 < 3.8$, flagging P_1 and P_2 as bad data would be a plausible explanation for the inconsistency observed in the data, although it is not necessarily the only one.

Remarks:

- If the suspect set is considered to be $\{P_2, P_3\}$, then

$$\|\mathbf{r}_k\|^2 = \hat{\mathbf{r}}_k' \mathbf{R}_{kk}^{-1} \hat{\mathbf{r}}_k = 133.0$$

Since $133.0 > \chi_{2,0.95}^2 = 7.4$, the null hypothesis H_0 is rejected as above, although the corrected index is

$$J(\hat{\mathbf{x}} + \Delta\hat{\mathbf{x}}) = J(\hat{\mathbf{x}}) - \|\mathbf{r}_k\|^2 = 185.8 - 133.0 = 52.8 > 3.8$$

which shows that flagging these two measurements (P_2 and P_3) as suspect data does not provide a plausible explanation for the inconsistency observed in the data.

- Various other sets of suspect measurements, such as $\{P_1\}$ or $\{P_4, P_5\}$, also fail to provide plausible explanations for errors observed.
- The suspect set $\{P_2, P_3, P_4\}$ does provide a plausible explanation, however, since it yields

$$J(\hat{\mathbf{x}} + \Delta\hat{\mathbf{x}}) < \chi_{1,0.95}^2$$

This explanation, however, involves the measurement P_4 , which does not contribute to the further reduction of the performance index in comparison with the reduction yielded by the suspect set $\{P_2, P_3\}$.

9.3.2 Identification Test

The identification of interacting multiple bad data problem is treated here as a combinatorial optimization problem, using an approach which borrows the framework of the Decision Theory. This optimization considers not only the accuracy of the measurements, but also their reliability. As discussed above, there are cases in which various plausible explanations for an inconsistency observed in data are possible. In these cases, it is important to distinguish between the reliability of a measurement and its precision. Reliability refers to the probability of failure of a meter (in fact, of the whole data acquisition system associated with a given meter). Precision refers to the magnitude of error when that meter is working. For example, a meter designed for great accuracy can still be quite unreliable (due to frequent break downs). The weighted least squares estimator incorporates the precision of measurements by means of the weighting factors (the \mathbf{W} matrix); measurement reliability, however, is not explicitly considered in the WLS formulation.

The decision as to whether the i th measurement is bad is denoted by one of the following:

$$\begin{aligned} d_i = 0, & \text{ if the } i\text{-th measurement is a bad data} \\ d_i = 1, & \text{ if the } i\text{-th measurement is good} \end{aligned}$$

Each possible combination of good and bad measurements is associated with a decision vector

$$\mathbf{d} = (d_1, d_2, d_3, \dots, d_m)$$

with $d_i = 0$ or 1. Hence, for an observable system with m measurements there are 2^m possible decision vectors.

Each decision vector corresponds to a set of suspect data (measurements with $d_i = 0$). Any given decision vector is a plausible explanation for the observed data inconsistencies if, after the removal of the suspect data or of their effect on state estimation, no further bad data are detected. If, for example, the $J(\hat{\mathbf{x}})$ -test is used, this means that $J(\hat{\mathbf{x}} + \Delta\hat{\mathbf{x}}) < C$, where $\hat{\mathbf{x}} + \Delta\hat{\mathbf{x}}$ is the new state estimate and C is the appropriate detection threshold. Hence, in any practical problem, a number of plausible explanations are likely to exist. The next step is to find the most plausible explanation, or the set of most plausible explanations, according to a pre-specified criterion.

Meters can be up or down. Let p_i be the probability that meter i is up and $q_i = 1 - p_i$ the probability that it is down. Let U be the set of all meters (and the corresponding part of the data gathering system) that is up at any given time t and D the set that is down. The probability associated with this particular decision vector \mathbf{d} is

$$Prob(\mathbf{d}) = \prod_{i \in U} p_i \prod_{j \in D} q_j \quad (9.4)$$

According to the maximum likelihood criterion, the optimal decision \mathbf{d} maximizes $Prob(\mathbf{d})$, or equivalently maximizes $\log(Prob(\mathbf{d}))$.

It is now assumed that (a) p_i is close to one, and (b) the number of elements in U is usually much larger than that in D . The following is then true:

$$\log(Prob(\mathbf{d})) = \sum_{i \in U} \log(p_i) + \sum_{j \in D} \log(q_j) \simeq \sum_{j \in D} \log(q_j) \quad (9.5)$$

If all meters are assumed to have the same reliability, i.e., q_j is constant for all j , a solution that maximizes $Prob(\mathbf{d})$ also maximizes the objective function

$$F(\mathbf{d}) = \sum_{i=1}^m (1 - d_i) \quad (9.6)$$

Hence, given the above assumptions, the optimal solution (or solutions if there is more than one) is the one that leads to the identification of the minimum number of bad measurements.

The problem of the identification of interacting multiple bad data can be formulated as follows. For any given decision vector \mathbf{d} , $N(\mathbf{d})$ denotes the network and corresponding meter configuration, $J(\hat{\mathbf{x}}(\mathbf{d}))$ is the performance index, and $C(\mathbf{d})$ is the detection threshold. The following combinatorial problem can thus be written:

$$\begin{aligned} & \text{Minimize} && F(\mathbf{d}) \\ & \text{subject to} && N(\mathbf{d}) \text{ is observable} \\ & && J(\hat{\mathbf{x}}(\mathbf{d})) < C(\mathbf{d}) \end{aligned} \quad (9.7)$$

Example 9.6:

Consider the two-bus system in Fig. 9.4, with measurements $P_{12}^1 = 1.02$ p.u., $P_{12}^2 = 1.04$ p.u., $P_{21}^3 = -0.97$ p.u., and $P_{21}^4 = -1.01$ p.u., and variances $\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \sigma_4^2 = .001$ p.u.

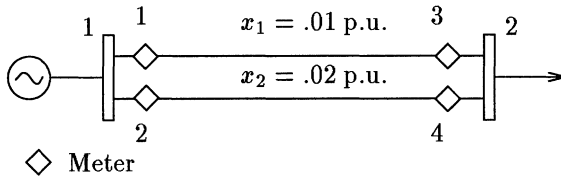


Figure 9.4. Two-bus system.

In this case, the vector of estimation residuals $\hat{\mathbf{r}}$, the vector of normalized residuals \mathbf{r}^n , and the performance index $J(\hat{\mathbf{x}})$ are as follows:

$$\hat{\mathbf{r}} = \mathbf{S} \mathbf{z} = \begin{pmatrix} -0.186 \\ 0.437 \\ 0.236 \\ -0.407 \end{pmatrix}$$

$$\mathbf{r}^n = [\mathbf{diag}(\mathbf{R}_{\hat{\mathbf{r}}})]^{-1/2} \hat{\mathbf{r}} = \begin{pmatrix} -7.59 \\ 14.56 \\ 9.63 \\ -13.56 \end{pmatrix}$$

$$J(\hat{\mathbf{x}}) = \hat{\mathbf{r}}' \mathbf{R}_{\hat{\mathbf{z}}}^{-1} \hat{\mathbf{r}} = 446.9$$

Applying the largest normalized residual criterion for the successive elimination of suspect measurements leads to the removal of the measurements 2 and 4. The resulting measurement set guarantees observability and presents no inconsistencies, and thus this is a plausible explanation for the observed errors.

The elimination of the measurements 2 and 4 also implies the removal of the 0.02 p.u. reactance from the measurement model. Hence, another plausible explanation would be that this reactance is actually wrong: if its value is corrected to 0.01 p.u., no inconsistencies would remain and the model would fit the available data.

Yet another plausible explanation would be to consider measurements 1 and 3 as bad data. A fourth possibility would be to consider the 0.01 p.u. reactance as bad data.

The removal of any subset of three measurements is also a plausible explanation, although this is not an optimal one. The removal of all four measure-

ments, is of course unfeasible, since the system would then become unobservable.

Remarks: Note that, although the weightings are the same, the residuals and the normalized residuals of the flow measurements located in the 0.02 p.u. line are smaller than the corresponding residuals of the 0.01 p.u. line. This would not happen if the two lines had the same reactance, however; in the weighted least squares estimator, power flow measurements in short lines (electrically) are implicitly more heavily weighted than measurements in long ones.

Example 9.7:

Consider the three-bus system in Fig. 9.5, with measurements $P_1^{meas} = 4.06$ p.u., $P_2^{meas} = -5.00$ p.u., $P_3^{meas} = -0.02$ p.u., $P_{12}^{meas} = 2.02$ p.u., $P_{21}^{meas} = -2.50$ p.u., and $P_{13}^{meas} = 1.96$ p.u., and variances $\sigma_{P_1}^2 = .004$ p.u., $\sigma_{P_2}^2 = .004$ p.u., $\sigma_{P_3}^2 = .001$ p.u., $\sigma_{P_{12}}^2 = .002$ p.u., $\sigma_{P_{21}}^2 = .002$ p.u., and $\sigma_{P_{13}}^2 = .002$ p.u. All meters are assumed to have equal reliability. The state variables are θ_1 and θ_2 as above.

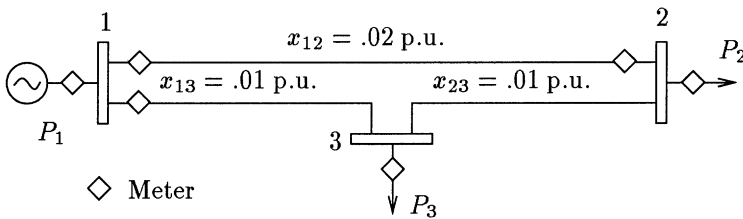


Figure 9.5. Three-bus system.

In this case, the vector of estimation residuals $\hat{\mathbf{r}}$, the vector of normalized residuals \mathbf{r}^n , and the performance index $J(\hat{\mathbf{x}})$ are as follows:

$$\hat{\mathbf{r}} = \mathbf{S} \mathbf{z} = \begin{pmatrix} -0.332 \\ -0.475 \\ -0.151 \\ -0.209 \\ -0.270 \\ -0.203 \end{pmatrix}$$

$$\mathbf{r}^n = [\mathbf{diag}(\mathbf{R}_{\hat{\mathbf{r}}})]^{-1/2} \hat{\mathbf{r}} = \begin{pmatrix} -6.34 \\ -9.46 \\ -11.06 \\ -5.06 \\ -6.54 \\ -5.14 \end{pmatrix}$$

$$J(\hat{\mathbf{x}}) = \hat{\mathbf{r}}' \mathbf{R}_z^{-1} \hat{\mathbf{r}} = 186.4$$

There is only one optimal solution to the optimization problem in Eq. (9.7): the removal of measurements P_2^{meas} and P_{21}^{meas} . The corresponding multiple normalized residual vector is

$$\mathbf{r}_k^n = \begin{pmatrix} r_{P_2}^n \\ r_{P_{21}}^n \end{pmatrix} = \begin{pmatrix} -12.43 \\ 5.45 \end{pmatrix}$$

which has a Euclidean norm of

$$\|\mathbf{r}_k\|^2 = \hat{\mathbf{r}}_k' \mathbf{R}_{kk}^{-1} \hat{\mathbf{r}}_k = 184.5$$

Hence, the corrected performance index is

$$J(\hat{\mathbf{x}} + \Delta\hat{\mathbf{x}}) = J(\hat{\mathbf{x}}) - \|\mathbf{r}_k\|^2 = 186.4 - 184.5 = 1.9$$

An alternative plausible explanation (although with higher cost) consists of flagging measurements P_1^{meas} , P_{12}^{meas} and P_{13}^{meas} as bad data. In this case, the multiple normalized residual vector is

$$\mathbf{r}_k^n = \begin{pmatrix} r_{P_1}^n \\ r_{P_{12}}^n \\ r_{P_{13}}^n \end{pmatrix} = \begin{pmatrix} 13.27 \\ 10.97 \\ -3.05 \end{pmatrix}$$

which has a Euclidean norm of

$$\|\mathbf{r}_k\|^2 = \hat{\mathbf{r}}_k' \mathbf{R}_{kk}^{-1} \hat{\mathbf{r}}_k = 186.4$$

Hence, the corrected performance index is

$$J(\hat{\mathbf{x}} + \Delta\hat{\mathbf{x}}) = J(\hat{\mathbf{x}}) - \|\mathbf{r}_k\|^2 = 186.4 - 186.4 = 0.0$$

9.4 TESTING EQUALITY CONSTRAINT HYPOTHESES

Equality constraints in state estimation can be represented as either hard or soft. The former consists of equality constraints explicitly added to the minimization problem of state estimation; the latter consists of equality constraints represented by means of pseudo measurements, accompanied by appropriate weightings (the higher the weighting, the better the representation). When nonlinear models are used, both methods can suffer from convergence problems when the constraint represents false information such as reporting a breaker status as open when it is actually closed in the field, or using a bus as a transition bus (zero injection) when it is actually a load bus. This section discusses an alternative approach to testing hypotheses about equality constraints in which no equality constraint is added to the state estimation problem, neither as an additional constraint nor as a pseudo-measurement.

Take the null hypothesis

$$\mathbf{Ax} = \mathbf{a}$$

where \mathbf{x} is an n -vector of state variables, \mathbf{a} is a constant p -vector, and \mathbf{A} is a constant $p \times n$ matrix. When testing this hypothesis, the expected value of $\mathbf{A}\hat{\mathbf{x}}$ and the corresponding covariance matrix are as follows:

$$E\{\mathbf{A}\hat{\mathbf{x}}\} = \mathbf{a}$$

$$\text{Cov}\{\mathbf{A}\hat{\mathbf{x}}\} = E\{(\mathbf{A}\hat{\mathbf{x}} - \mathbf{a})(\mathbf{A}\hat{\mathbf{x}} - \mathbf{a})'\} = \mathbf{A}(\mathbf{H}'\mathbf{W}\mathbf{H})^{-1}\mathbf{A}'$$

where $\mathbf{W} = \mathbf{R}_z^{-1}$. Since $\hat{\mathbf{x}}$ is n -variate normal, if $\text{Rank}(\mathbf{A}) = p \leq n$, then $\mathbf{A}\hat{\mathbf{x}} - \mathbf{a}$ is p -variate normal with covariance given by the above expression, and the performance index is expressed as

$$J(\mathbf{A}\hat{\mathbf{x}} - \mathbf{a}) = (\mathbf{A}\hat{\mathbf{x}} - \mathbf{a})' (\mathbf{A}(\mathbf{H}'\mathbf{W}\mathbf{H})^{-1}\mathbf{A}')^{-1} (\mathbf{A}\hat{\mathbf{x}} - \mathbf{a})$$

This follows a χ_p^2 distribution, i.e., chi-square distribution with p degrees of freedom. The statistic defined by $J(\mathbf{A}\hat{\mathbf{x}} - \mathbf{a})$ can then be used to test the null hypothesis $\mathbf{Ax} = \mathbf{a}$.

Example 9.8:

Consider the network in Fig. 0.1. The branch reactances are 0.01 p.u. Assume for simplicity a linear dc power flow model is used. The status of the breaker 1 – 3 is unknown, thus neither $P_{13} = 0$ nor $\theta_{13} = 0$ are added as constraints (or pseudo-measurements). The Jacobian matrix is as follows:

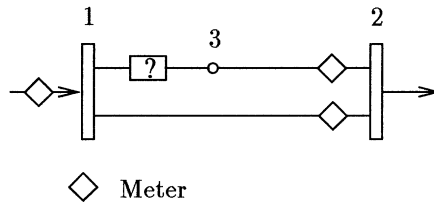


Figure 0.1. Hypothesis testing of the status Breaker 1 – 3.

$$\mathbf{H} = \begin{matrix} & \theta_2 & \theta_3 & P_{13} \\ \begin{matrix} P_1 \\ P_{23} \\ P_{21} \\ P_3 \end{matrix} & \begin{pmatrix} -100. & 0. & 1. \\ 100. & -100. & 0. \\ 100. & 0. & 0. \\ -100. & 100. & -1. \end{pmatrix} \end{matrix}$$

The weighting matrix and the measurement vector are assumed to be as follows:

$$\mathbf{W} = \begin{pmatrix} 700 & 0 & 0 & 0 \\ 0 & 700 & 0 & 0 \\ 0 & 0 & 700 & 0 \\ 0 & 0 & 0 & 10000 \end{pmatrix}$$

$$\mathbf{z} = \begin{pmatrix} 2.06 \\ -0.96 \\ -0.98 \\ 0.00 \end{pmatrix}$$

The state estimate $\hat{\mathbf{x}}$ is given by,

$$\hat{\mathbf{x}} = (\mathbf{H}' \mathbf{W} \mathbf{H})^{-1} \mathbf{H}' \mathbf{W} \mathbf{z} = \begin{pmatrix} -0.0102 \\ -0.0002 \\ 1.0018 \end{pmatrix}$$

The coefficient matrices \mathbf{A} corresponding to the equality constraints $P_{13} = 0$ and $\theta_{13} = 0$ are given below:

$$\mathbf{A}_{\text{open}} = P_{13} \begin{pmatrix} \theta_2 & \theta_3 & P_{13} \\ 0 & 0 & 1 \end{pmatrix}$$

$$\mathbf{A}_{\text{closed}} = \theta_{13} \begin{pmatrix} \theta_2 & \theta_3 & P_{13} \\ 0 & -1 & 0 \end{pmatrix}$$

The detection threshold for $\alpha = 0.025$ is $\chi_{1,0.025}^2 = 5.0$ (chi-square distribution with one degree of freedom). When the breaker is considered open, the null hypotheses is rejected, since

$$J(\mathbf{A}_{\text{open}} \hat{\mathbf{x}}) = 1007.9 > 5.0$$

but when the breaker is considered closed, the null hypothesis is accepted since

$$J(\mathbf{A}_{\text{closed}} \hat{\mathbf{x}}) = 0.14 \leq 5.0$$

Example 9.9:

An interesting situation in which neither of the two hypothesis (neither open nor closed) will be accepted, i.e.,

$$J(\mathbf{A} \hat{\mathbf{x}}) > \chi_{1,0.025}^2$$

in both tests, is considered in the following. This would be the case if the following values were obtained for \mathbf{z} :

$$\mathbf{z} = \begin{pmatrix} 2.06 \\ -0.80 \\ -0.98 \\ 0.00 \end{pmatrix}$$

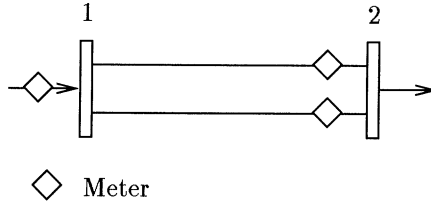


Figure 9.7. Reduced model obtained when the closed breaker hypothesis is accepted.

In this case, The results would be

$$J(\mathbf{A}_{\text{closed}}\hat{\mathbf{x}}) = 11.3 > 5.0$$

and

$$J(\mathbf{A}_{\text{open}}\hat{\mathbf{x}}) = 809.1 > 5.0$$

and both hypotheses would be rejected. The closed breaker hypotheses could still be accepted, since

$$J(\mathbf{A}_{\text{closed}}\hat{\mathbf{x}}) \ll J(\mathbf{A}_{\text{open}}\hat{\mathbf{x}}),$$

a decision which would produce the reduced model given in Fig. 9.7. Performing state estimation for this system with the given set of measurements would yield

$$J(\hat{\mathbf{x}}) = 29.6$$

$$\mathbf{r}^n = \begin{pmatrix} 4.28 \\ 5.31 \\ 0.10 \end{pmatrix}$$

which clearly indicates the presence of bad data. Hence, the proposed set of measurements (i.e., $\mathbf{z} = (2.06 \ -0.80 \ -0.98 \ 0.00)'$) should not be used for hypotheses testing concerning the status of Breaker 1 – 3, and this is what the double negative results of the equality constraint hypothesis testing are saying. Thus, before testing the hypothesis about the breaker status, analog bad data has to be corrected.

Example 9.10:

In the example above, the status of Breaker 1 – 3 is considered unknown. In terms of the measurement model, this means excluding the pseudo-measurements corresponding to the breaker status ($P_{13} = 0$ when the breaker is open and $\theta_{13} = 0$ when it is closed). There is, however, one more pseudo-measurement that forms part of the breaker model: the bus injection pseudo-measurement $P_3 = 0$. The hypotheses testing for the previous example will thus

be repeated incorporating this pseudo-measurement. The two following cases will then be considered: open breaker, represented by $P_{13} = 0$ and $P_3 = 0$; and closed breaker, represented by $\theta_{13} = 0$ and $P_3 = 0$. The equality constraints are then dropped from the model and used after estimation for hypotheses testing. With these modifications, the measurement model, the weighting matrix, and the measurement vector become the following:

$$\mathbf{H} = \begin{matrix} & \theta_2 & \theta_3 & P_{13} \\ \begin{matrix} P_1 \\ P_{23} \\ P_{21} \end{matrix} & \begin{pmatrix} -100. & 0. & 1. \\ 100. & -100. & 0. \\ 100. & 0. & 0. \end{pmatrix} \end{matrix}$$

$$\mathbf{W} = \begin{pmatrix} 700 & 0 & 0 \\ 0 & 700 & 0 \\ 0 & 0 & 700 \end{pmatrix}$$

$$\mathbf{z} = \begin{pmatrix} 2.06 \\ -0.96 \\ -0.98 \end{pmatrix}$$

Note that in this case all measurements are critical. The state estimate $\hat{\mathbf{x}}$ is given by

$$\hat{\mathbf{x}} = (\mathbf{H}'\mathbf{W}\mathbf{H})^{-1}\mathbf{H}'\mathbf{W}\mathbf{z} = \begin{pmatrix} -0.0098 \\ -0.0002 \\ 1.0800 \end{pmatrix}$$

The coefficient matrices \mathbf{A} corresponding to the open and closed breaker cases are

$$\mathbf{A}_{\text{open}} = \begin{matrix} & \theta_2 & \theta_3 & P_{13} \\ \begin{matrix} P_{13} \\ P_3 \end{matrix} & \begin{pmatrix} 0 & 0 & 1 \\ -100 & 100 & -1 \end{pmatrix} \end{matrix}$$

$$\mathbf{A}_{\text{closed}} = \begin{matrix} & \theta_2 & \theta_3 & P_{13} \\ \begin{matrix} \theta_{13} \\ P_3 \end{matrix} & \begin{pmatrix} 0 & -1 & 0 \\ -100 & 100 & -1 \end{pmatrix} \end{matrix}$$

The detection threshold for $\alpha = 0.025$ is $\chi_{2,0.025}^2 = 7.4$ (chi-square distribution with two degrees of freedom). The null hypotheses with the breaker open is rejected since $J(\mathbf{A}_{\text{open}}\hat{\mathbf{x}}) = 1053 > 7.4$, whereas if the breaker is considered closed, the null hypotheses is accepted since $J(\mathbf{A}_{\text{closed}}\hat{\mathbf{x}}) = 3.5 \leq 7.4$.

9.5 STRATEGIES FOR PROCESSING INTERACTING BAD DATA

The identification of interacting multiple bad data involves a difficult combinatorial problem (see Eq. (9.7) above). The situation is further complicated by the presence of parameter and topology errors. Although simple heuristic

methods are available that show good performance in most cases, there is no general method for the treatment of the identification problem in its more general form (interacting multiple analog, status, and parameter errors). In this section, the successive largest normalized residual method will be discussed, and a combinatorial search method based on a binary tree presented. Then a brief presentation of the main characteristics of the Tabu Search approach is made, followed by a discussion of local solutions and concepts of pocketing and zooming.

9.5.1 Successive Largest Normalized Residual

This method consists of the repeated application of the largest normalized residual method to cases with multiple bad data. When multiple bad data are noninteracting, the largest normalized residual still corresponds to a bad measurement, and even cases of interacting bad data can be solved by this method, as long as the interacting bad data are not conforming. The most difficult cases are those in which the errors in the interacting bad data are in close agreement. When this happens, a correct measurement may present the largest normalized residual whereas a measurement that is actually wrong may yield a small normalized residual or no residual at all. The following example provides some insights into cases of interacting bad data that are both interacting and conforming.

Example 9.11:

Consider the three-bus system in Fig. 9.5 discussed in Example 9.7, in which $P_1^{meas} = 4.06$ p.u., $P_2^{meas} = -5.00$ p.u., $P_3^{meas} = -0.02$ p.u., $P_{12}^{meas} = 2.02$ p.u., $P_{21}^{meas} = -2.50$ p.u., and $P_{13}^{meas} = 1.96$ p.u.

In this case the vector of normalized residuals \mathbf{r}^n is

$$\mathbf{r}^n = \begin{pmatrix} -6.34 \\ -9.46 \\ -11.06 \\ -5.06 \\ -6.54 \\ -5.14 \end{pmatrix}$$

As noted in Example 9.7, in this instance there is a single optimal solution to the optimization problem formulated in Eq. (9.7): the removal of measurements P_2^{meas} and P_{21}^{meas} . The corresponding error estimates, obtained from Eq. (9.1), are -0.99 and -0.59 , respectively. These measurements are both interacting and conforming: the largest normalized residual corresponds to measurement P_3^{meas} , which is actually good data, although the repeated removal of the measurement with the largest normalized residual would eventually identify measurements P_2^{meas} and P_{21}^{meas} as bad data (i.e., these two, besides P_3^{meas}).

Another situation to consider would be the measurement set: $P_1^{meas} = 4.06$ p.u., $P_2^{meas} = -5.00$ p.u., $P_3^{meas} = -0.02$ p.u., $P_{12}^{meas} = 2.02$ p.u., $P_{21}^{meas} = -1.50$ p.u., and $P_{13}^{meas} = 1.96$ p.u., in which the gross error in measurement P_{21}^{meas} has the same magnitude as above but the reverse sign. In this case, the error is $+0.50$ p.u., whereas it was -0.50 p.u. previously. The vector of normalized residuals is

$$\mathbf{r}^n = \begin{pmatrix} -1.09 \\ -15.52 \\ -8.95 \\ -1.55 \\ -14.13 \\ -1.84 \end{pmatrix}$$

The largest normalized residual actually corresponds to bad data (measurement P_{21}^{meas}), and the repeated removal of the measurements with the largest normalized residuals would lead to the correct solution. Although the bad data are interacting, as above, they are no longer conforming, and the repeated largest normalized residual method is sufficient; measurements P_2^{meas} and P_{21}^{meas} are identified as being bad.

9.5.2 Dormant Measurements and Perfect Measurements

Consider again the sensitivity relation $\hat{\mathbf{r}} = \mathbf{S}\mathbf{z}$ with the partition discussed in Sec. 9.1:

$$\begin{pmatrix} \hat{\mathbf{r}}_k \\ \hat{\mathbf{r}}_o \end{pmatrix} = \begin{pmatrix} \mathbf{S}_{kk} & \mathbf{S}_{ko} \\ \mathbf{S}_{ok} & \mathbf{S}_{oo} \end{pmatrix} \begin{pmatrix} \mathbf{e}_k \\ \mathbf{e}_o \end{pmatrix}$$

where k denotes a set of redundant measurements and o a set of measurements that guarantees network observability. It is known that when the measurement vector is corrected by the error estimate given in Eq. (9.1), the result is

$$\mathbf{z}^{new} = \mathbf{z} - \mathbf{S}_{kk}^{-1} \hat{\mathbf{r}}_k,$$

The redundant measurements k become dormant, and will have no effect on the state estimation results. This is equivalent to zeroizing the weights of these measurements in the weighted least squares state estimator. The residuals corresponding to the corrected measurement are also zero.

Now, consider what happens when the k measurements are perfect, with weights equivalent to ∞ . This can be simulated rewriting the sensitivity relation above as follows:

$$\begin{pmatrix} \Delta \mathbf{z}_k \\ \hat{\mathbf{r}}_o^{new} \end{pmatrix} = \begin{pmatrix} \mathbf{S}_{kk} & \mathbf{S}_{ko} \\ \mathbf{S}_{ok} & \mathbf{S}_{oo} \end{pmatrix} \begin{pmatrix} \mathbf{e}_k + \Delta \mathbf{z}_k \\ \mathbf{e}_o \end{pmatrix}$$

where $\Delta \mathbf{z}_k$ represents fictitious corrections imposed on the measurements of the k -set such that the corresponding residual becomes $\hat{\mathbf{r}}_k = \Delta \mathbf{z}_k$, i.e., the compo-

ment of the residual $\hat{\mathbf{r}}_k$ due to the errors \mathbf{e}_k and \mathbf{e}_o , is zeroized. Considering that

$$\hat{\mathbf{r}}_k = \mathbf{S}_{kk}\mathbf{e}_k + \mathbf{S}_{ko}\mathbf{e}_o$$

the first row of this equation yields

$$\Delta z_k = \hat{\mathbf{r}}_k + \mathbf{S}_{kk}\Delta z_k$$

Hence, introducing corrections into the measurement vector yields

$$\Delta \mathbf{z}_k = (\mathbf{I}_k - \mathbf{S}_{kk})^{-1}\hat{\mathbf{r}}_k \quad (9.8)$$

where \mathbf{I}_k is the unit matrix of order k , is equivalent to modeling the k measurements as equality constraints (or as pseudo-measurements with ∞ weightings in a WLS estimator).

Example 9.12:

Consider the two-bus system in Fig. 9.4 with measurements $P_{12}^1 = 1.02$ p.u., $P_{12}^2 = 1.04$ p.u., $P_{21}^3 = -0.97$ p.u., and $P_{21}^4 = -1.01$ p.u., and variances $\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \sigma_4^2 = .001$ p.u.

The measurement vector and the corresponding estimate are as follows:

$$\mathbf{z} = \begin{pmatrix} 1.02 \\ 1.04 \\ -0.97 \\ -1.01 \end{pmatrix}$$

$$\hat{\mathbf{z}} = \begin{pmatrix} 1.21 \\ 0.60 \\ -1.21 \\ -0.60 \end{pmatrix}$$

The vector of residuals is

$$\hat{\mathbf{r}} = \begin{pmatrix} -0.19 \\ 0.44 \\ 0.24 \\ -0.41 \end{pmatrix}$$

and the vector of normalized residuals is

$$\mathbf{r}^n = \begin{pmatrix} -7.6 \\ 14.6 \\ 9.6 \\ -13.6 \end{pmatrix}$$

The removal of the measurement with the largest normalized residual (Measurement 2) can be simulated by making this measurement dormant. The modified measurement (Eq. (9.1)) is,

$$P_{12}^{2,new} = P_{12}^{2,meas} - S_{22}^{-1} \hat{r}_2 = 1.04 - 0.90^{-1} \times 0.44 = 0.55$$

This new measurement vector and the corresponding estimate are as follows: respectively,

$$\mathbf{z} = \begin{pmatrix} 1.02 \\ 0.55 \\ -0.97 \\ -1.01 \end{pmatrix}$$

$$\hat{\mathbf{z}} = \begin{pmatrix} 1.11 \\ 0.55 \\ -1.11 \\ -0.55 \end{pmatrix}$$

The new vector of residuals is

$$\hat{\mathbf{r}} = \begin{pmatrix} -0.09 \\ 0.00 \\ 0.14 \\ -0.46 \end{pmatrix}$$

and the new vector of normalized residuals is

$$\mathbf{r}^n = \begin{pmatrix} -3.63 \\ 0.00 \\ 5.67 \\ -15.2 \end{pmatrix}$$

The largest normalized residual resulting corresponds to the measurement 4.

Rather than removing the measurement 2, however, the effect of transforming it into a perfect observation (i.e., an equality constraint) can be simulated. The modified measurement (Eq. (9.8)) will be as follows:

$$P_{12}^{2,new} = P_{12}^{2,meas} + (1 - S_{22})^{-1} \hat{r}_2 = 1.04 + (1 - 0.90)^{-1} \times 0.44 = 5.41$$

The modified measurement vector and the corresponding estimate are as follows:

$$\mathbf{z} = \begin{pmatrix} 1.02 \\ 5.41 \\ -0.97 \\ -1.01 \end{pmatrix}$$

$$\hat{\mathbf{z}} = \begin{pmatrix} 2.08 \\ 1.04 \\ -2.08 \\ -1.04 \end{pmatrix}$$

Note that in this case the estimate corresponding to Measurement 2 is the original measured value, i.e., $P_{12}^{2,meas} = 1.04$ (now considered to be a perfect measurement) and also that the new residuals are computed considering this value:

$$\hat{\mathbf{r}} = \begin{pmatrix} -1.06 \\ 0.00 \\ 1.11 \\ 0.03 \end{pmatrix}$$

$$\mathbf{r}^n = \begin{pmatrix} -43.3 \\ 0.00 \\ 45.3 \\ 1.00 \end{pmatrix}$$

The largest normalized residual now corresponds to Measurement 3, and the repeated application of the largest normalized residual criterion leads to the removal of Measurements 3 and 1.

Remarks: The techniques discussed above (dormant/perfect measurements) are very helpful in the practical implementation of combinatorial methods used to identify multiple bad data that are both interacting and conforming. A suspect measurement, or a set of suspect measurements, can be made either dormant or perfect, and this will affect the identification of the remaining bad data, yielding alternative plausible explanations for inconsistencies found in the measurement set under consideration.

9.5.3 Binary Tree Search

Figure 9.8 shows a binary tree that can be used systematically to generate decision vectors. As noted above, there are a total of 2^m possible decision vectors in a system with m measurements (and pseudo-measurements). At each node of the tree, a decision as to whether a particular measurement is good or bad must be made. At level j , j measurements have been classified as being either good or bad. At level m , all measurement have been studied (2^m decision vectors). Some of these decision vectors are feasible, i.e., they satisfy the constraints of the optimization problem in Eq. (9.7), so that after the removal of the bad data in \mathbf{d} , the resulting network will be observable and no other bad data will be detected.

There are several ways of solving the optimization problem in Eq. (9.7). When there is a single optimal solution, the determination of only that solution may be desirable. But when multiple optimal solutions are possible, the determination of all of them may be preferred. Alternatively, a set of all the solutions within a certain distance from the optimal ones may be of interest, such as when the optimal solution involves the removal of two measurements, and all the plausible solutions involving the removal of up to three measurements may be of interest. Of course, the greater the number of plausible solutions sought, the bigger is the computational effort necessary. All of these alternative

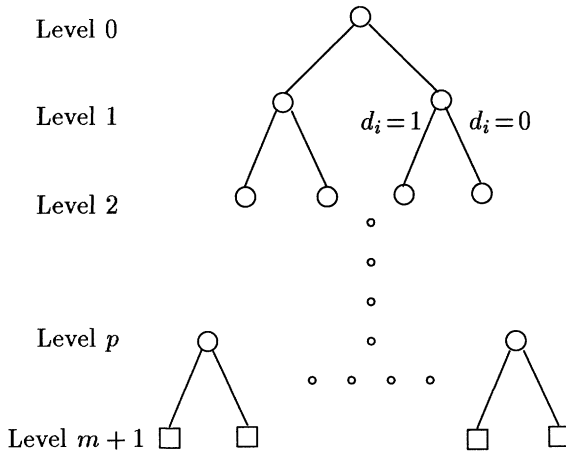


Figure 9.8. Generic binary decision tree used to identify multiple bad data in a system with m measurements.

formulations can be determined by the branch-and-bound algorithm presented below. The underlying support of this method consists of a series of strategies determining when to grow (branching) and when to stop growing the tree (bounding).

The detection test (e.g., the r^n test) is conducted at each decision node using the assumption that all undeclared measurements are good. If bad data are detected, the next measurement to be declared is the one which has the largest normalized residual among the undeclared measurements. Two new decision nodes are created: the b -successor and the g -successor. The b -successor corresponds to the decision to consider the measurement under consideration to be bad, and the g -successor corresponds to considering the measurement to be good data.

The state estimation problem involving the decision node currently under consideration after all measurements identified as bad have been removed, is known as the current candidate problem. The set of all minimum cost feasible solutions obtained so far at any given stage of the search process is an incumbent solution set. These decision nodes (or vectors) which may lead to solutions that are as good as or better than the incumbent solutions are included on the list of open nodes; those nodes which do not satisfy these conditions are put on the list of closed nodes (the corresponding parts of the tree are then pruned). Whenever a new incumbent solution is found, the list of open nodes is consulted and the nodes leading to a solution with greater costs than those involved for the new incumbent are deleted (pruned) from the list.

In most cases, bad data have large normalized residuals, even when these are not the largest. The inclusion of g -successors widens the search space. Choosing to move along the g -successor essentially suggests the rejection of the idea

that the largest normalized residual refers to a bad data, and a request for more information about the remaining measurements is made (a strategy which is also helpful in handling bad data which is conforming). Obtaining more information requires additional computation. Therefore the branch-and-bound algorithm maintains record of how many times more information is sought; in the algorithm, nr denotes the number of times the g -successor can appear in decision vectors. This constant can be specified beforehand (as in the algorithm below) or can be determined on an ad-hoc basis considering, for example, the number of bad data flagged by the repeat largest normalized residual method. Note that the branch-and-bound method supersedes the repeat largest normalized residual method: if $nr = 0$, all the g -successors are ignored, and the branch-and-bound method will behave just like the repeat largest normalized residual method. On the other hand, setting $nr = \infty$, results in a comprehensive search.

Branch and Bound Algorithm:

1. Initialization:
 - (a) Define a detection test (e.g., the $J(\hat{\mathbf{x}})$ -test or the \mathbf{r}^n -test).
 - (b) Put the zero cost decision vector $\mathbf{d} = (0, 0, \dots, 0)'$ on the list of open nodes (all measurements are considered (declared) good).
 - (c) Initialize the cost of the incumbent solution as $m - n$ (the number of degrees of freedom).
 - (d) Specify nr (the penalty for requesting more information).
2. Visit the list of open nodes:
 - (a) If the list is empty, terminate the search: the set of incumbent solutions maintains the optimal solutions (plausible explanations of minimum cost).
 - (b) Otherwise, choose from the list, the decision vector with minimum cost and form the corresponding current candidate problem.
3. Conduct the detection test for the current candidate problem.
 - (a) If no bad data are detected, then:
 - i. If the cost of the current candidate problem is less than that of the incumbent:
 - A. Make the current candidate problem the new incumbent.
 - B. Prune the tree by removing from the list of open nodes all decision nodes leading to solutions whose costs are greater than that of the new incumbent and return to 2.
 - ii. If the cost of the current candidate problem is equal to the cost of the incumbent, then add the current candidate problem to the set of incumbents and return to 2.

- (b) Otherwise, if bad data are detected, then:
- i. If the cost of the current candidate problem is less than the cost of the incumbent(s), then
 - A. The b -successor becomes the new current candidate problem.
 - B. If the number of g -successors in the current decision vector is less than nr , then put the g -successor on the list of open nodes.
 - C. Return to 3.
 - ii. Otherwise, put the current candidate problem on the list of closed nodes and return to 3.

A number of improvements can be introduced into the above algorithm. For example, when a measurement appears as a suspect for bad data for first time in the search (for example, when it has the largest normalized residual) it is possible to determine whether it actually consists of interacting bad data. To do this, the neighborhood of the suspect measurement is determined (using either numerical sensitivities or a topological method based on adjacency). The removal of the suspect measurement is simulated by making it dormant. The residuals of the other measurements in the neighborhood are then reestimated. If no bad data are detected, the suspect measurement is assumed to be noninteracting, and considered to be bad data. (See also Probs. 4 and 4 at the end of the chapter.)

Example 9.13:

Consider the three-bus system in Fig. 9.5 (Example 9.7) in which $P_1^{meas} = 4.06$ p.u., $P_2^{meas} = -5.00$ p.u., $P_3^{meas} = -0.02$ p.u., $P_{12}^{meas} = 2.02$ p.u., $P_{21}^{meas} = -2.50$ p.u., and $P_{13}^{meas} = 1.96$ p.u. Assume that all the measurements are equally reliable, and select the r^n -test with a detection threshold of 3.0. The penalty for requesting more information should be set at $nr = 2$.

Figure 9.9 summarizes the search performed by the branch-and-bound algorithm. Initially the algorithm finds the plausible explanation that consists of the removal of measurements P_3 , P_{12} , and P_{21} (Measurements 3, 4, and 5, respectively). This solution corresponds to Node 6 in the figure. At this stage, Nodes 3, 5, and 7 are on the list of open nodes. Next, Node 3 is branched, eventually leading a the plausible explanation consisting of declaring the measurements P_2 and P_{21} (Measurements 2 and 5) as bad data. This solution, corresponding to Node 12 in the figure, becomes the new incumbent. As a consequence, Node 7 is pruned, as it cannot lead to solutions that are as good as or better than the incumbent solution. Next, Node 9 is branched: the g -successor is ignored since the limit on nr will have been violated. All the nodes remaining to be explored lead to closed nodes. Hence, the optimal solution consists of the removal of two bad data: Measurements P_2 and P_{21} .

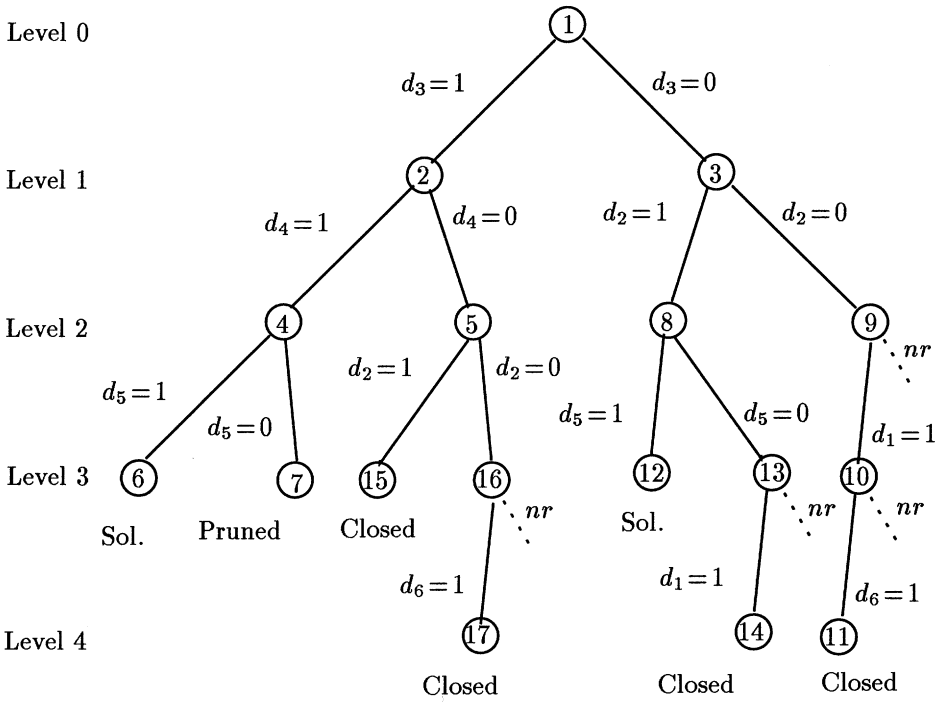


Figure 9.9. Binary decision tree from the application of the branch-and-bound algorithm to the situation in Example 9.13.

9.5.4 Tabu Search

The most typical feature of Tabu Search, TS, algorithms is the direct exclusion of search alternatives temporarily classed as forbidden or tabu. As a consequence, the use of memory becomes an essential part of the method, i.e., records of the alternatives visited must be maintained. A decision vector, containing decisions on analogs, statuses and parameters, can be obtained by an heuristic method (for example, the successive application of the largest normalized residual criterion or that of largest weighted residual) and some or all of the suspect measurements of the decision vector are then classified as good or bad for a certain period of time (tabu tenure).

The basic search mechanism proceeds as follows. A neighborhood structure is defined for each decision vector; the algorithm then moves to the best decision vector in this neighborhood, i.e., it switches to the vector presenting the greatest likelihood (minimum cost). Normally only the most attractive neighbors are evaluated as otherwise the problem could become too large. Other mechanisms of tabu search are intensification and diversification. These mechanisms allow for more comprehensive explorations of attractive regions which may lead to

a local optimal decision vector and movement to previously unvisited regions (which is important to avoid being trapped by local minima), respectively.

Decision vectors found during the search are stored, and their neighborhoods are then explored more thoroughly, and the local optimal decision vector closest to these vectors is found during the intensification phase. Intensification can also be performed based on "building blocks" present in elite decision vectors which have been stored in memory. A mechanism of path relinking can be used to produce new attractive configurations utilizing such building blocks; in fact path-relinking is a means of performing both intensification and diversification, depending on the amplitude of the changes introduced by the mechanism in the current decision vector.

Diversification involves a temporary change in the rules used for finding new decision vectors, and decisions already made about bad data (analogs, breakers and parameters) can be reversed: such variables then become tabu-active, i.e., they cannot be considered to be bad again for a certain period of time. This usually changes the decision vector in a way that forces the algorithm to visit unexplored regions.

Tabu search, more than simulated annealing and genetic algorithms, is especially suited for incorporating strategies from non-optimal identification algorithms. As has been mentioned above, one of the main difficulties found in using TS regards the need for reducing the size of the neighborhood of a decision vector without sacrificing quality, since TS tries to find the best decision in the neighborhood before making a move during the intensification phase. In dealing with practical problems, dramatic reductions in neighborhood sizes are normally required, and this must be done adaptively. Approximate algorithms can be extremely useful at this stage. For instance, when the current decision vector is infeasible, i.e., it passes neither the observability test nor the detection test, approximate methods based on weighted or normalized residuals (single residuals) can provide lists of suspect data.

In the tabu search approach, memory structures are used to direct the search. Four different dimensions are considered in memory structures: recency, frequency, quality, and influence. Recency-based memory is one of the most important in tabu search: it is a type of short term memory which keeps a record of decision vector attributes that have been changed during the most recent moves made by the algorithm. The information contained in this memory allows the labeling of specific attributes as tabu-active for solutions recently visited; this feature avoids unnecessarily revisiting. The frequency dimension is linked to long term memory mechanisms maintaining a record of the frequency with which certain specific decisions appear during the search process. This information is used later on in diversifying the search by changing the selection rules in a way that configurations containing still unused features can be visited by the algorithm. The quality dimension refers to the costs associated with different decision vectors searched by the algorithm. Certain groups of individual decisions can greatly improve a candidate decision vector when performed together, and these building blocks can be used as part of the

path relinking strategy, allowing the creation of new decision vectors from high quality building blocks (located in elite decision vectors stored in long term memory). Finally, the influence dimension accounts for the impact of an individual decision on a detection test, such as the impact on the performance index $J(\hat{\mathbf{x}})$.

Example 9.14:

Consider the two-bus system in Fig. 9.4 discussed in Example 9.6. By the repeated application of the largest weighted residual criterion, Measurements 2 and 4 are declared bad data. One of the measurements will now be reclassified as good data and a tabu created. The repeated application of the largest weighted residual criterion to the original set of data modified only by the tabu now reveals Measurements 1 and 3 to be bad data. Hence, there are two optimal plausible explanations. Consideration of all analog measurements as good data (tabus), will lead to the investigation of possible errors in line parameter or in breaker status. An error in either one of the two line reactances would also lead to a plausible explanation, whereas errors in breaker status would yield less probable plausible solutions because such a solution would also require the consideration of least two analog measurements as bad.

9.5.5 Localized Solutions, Pocketing and Zooming

Normally gross errors propagate poorly in state estimation. Therefore localization methods become attractive as the size of the network increases. Two different ways of exploring localization are discussed here: pocketing and zooming. Pockets are relatively small parts of the network affected by bad data. Zoomed-in areas are parts of the network modeled at the physical level: these can be pre-established selections or ad-hoc decisions based on the results of state estimation.

Bad data analysis is normally performed in observable islands with a certain degree of measurement redundancy. (These areas are called regions of interest.) The repeated application of a bad data identification method may be used to determine the parts of the region of interest affected by bad data: once suspect measurements are identified, their neighborhood is located using the topological algorithm below.

This algorithm uses the Jacobian matrix structure to form pockets: a measurement corresponds to a row entry in the Jacobian matrix and columns of the nonzero elements in a row define the states that are adjacent to the measurement. A state corresponds to a column entry in the Jacobian matrix, and the rows of the nonzero elements in the column define the states that are adjacent to the measurement. Pockets are formed around a predefined set of suspect measurements through the alternate use of these two properties.

Measurement/State Adjacency Algorithm

1. Initialize:
 - (a) Initialize the set of suspect measurements using an approximate method (e.g., repeat largest weighted residual).
 - (b) Initialize the set of adjacent states as an empty set.
 - (c) Define $nmax$ (the desired expansion level).
 - (d) Make $nexp = 0$ (the number of current expansions).
2. Perform state expansion:
 - (a) Using the Jacobian matrix structure, add to the set of adjacent states, the state variables adjacent to the current suspect measurement set.
 - (b) If $nexp = nmax$ go to 4.
3. Perform measurement expansion:
 - (a) Update the expansion counter: $nexp \leftarrow nexp + 1$.
 - (b) Using the Jacobian matrix structure, expand the current suspect measurement set by adding those measurements that are functions of the state variables of the set of adjacent states.
 - (c) Go to 2.
4. Define pockets based on the sets of measurement and adjacent states.

This algorithm automatically merges overlapping pockets. A bad data pocket identified by the algorithm refers to a specific portion of the region of interest that is electrically local to the suspect data. Note that the pocket is normally, but not always, comprised of contiguous devices.

Combinatorial search algorithms are generally time consuming, although execution times vary widely depending on the complexity of the case under consideration; this may be critical in a real-time environment. This is especially true when modeling at the physical level, due to the relatively large number of bus sections and switches. The correct identification of interacting multiple (and conforming) bad data constitutes a difficult combinatorial problem. As discussed above, some sort of search will normally be necessary to determine the most plausible explanation for the inconsistencies presented by a given set of measurements. Whatever search procedure is adopted, it will be limited to the pocket. A common cause of multiple bad data that are both interacting and conforming is the presence of topological or parameter errors, with the former being best analyzed by modeling the network at the physical level.

9.6 ROBUST ESTIMATORS

Robust estimators are designed to reduce the influence of bad data on state estimation results. In the previous sections, methods for detecting, identifying and removing bad data have been discussed. These methods are based on the weighted least squares estimator. WLS estimation gives relatively high weight to measurements with large residuals, positive or negative. Hence, the methods for processing bad data are normally based on the measurement residuals (weighted and/or normalized residuals). One key feature of the robust estimators is to reduce the weight given to bad data. In this section, methods that in principle have the inherent property of eliminating bad data, will be discussed (the so called robust estimators). These estimators are not to be confused with the numerically robust estimators discussed in Chap. 13.

The solution to the WLS problem is sensitive to bad data (outliers), since the larger the residual the bigger its influence on the objective function (quadratic objective function; see Fig. 9.10). In this sense, the WLS estimator is not robust. This can be illustrated by the unidimensional situation, in which case the WLS estimator yields the mean value of the given measurements (observations). Assuming a finite number of measurements, and that all measurements except one are correct, if the bad data tends to ∞ , so will the mean value, which is the state estimate in this case. In an ideally robust estimator this would not happen, however, since the state estimate would not be sensitive to changes occurring in a single measurement.

9.6.1 Non-quadratic Estimators

The unconstrained state estimation problem can be formulated mathematically as the following minimization problem:

$$\underset{\hat{\mathbf{x}}}{\text{Minimize}} \quad \rho(\mathbf{r}^w)$$

where $\rho(\cdot)$ is a scalar objective function, and

$$\mathbf{r}^w = \mathbf{R}_z^{-1/2}(\mathbf{z} - \mathbf{h}(\mathbf{x}))$$

is the vector of weighted residuals, \mathbf{R}_z being the measurement covariance matrix (a diagonal matrix) and $\mathbf{h}(\cdot)$ a vector function that relates measurements to state variables. In the WLS method discussed above the following quadratic objective function is adopted (Fig. 9.10(a)):

$$\rho(\mathbf{r}^w) = (\mathbf{r}^w)' \mathbf{r}^w = \sum_{i=1}^m \sigma_i^{-1} (z_i - h_i(\mathbf{x}))^2$$

To reduce the influence of large residuals, ρ can be defined as a quadratic function only for small weighted residuals and as a function with constant or decreasing derivative for large weighted residuals (see Fig. 9.10(b)-(c)). Since the measurements with large weighted residuals are not known a priori, the

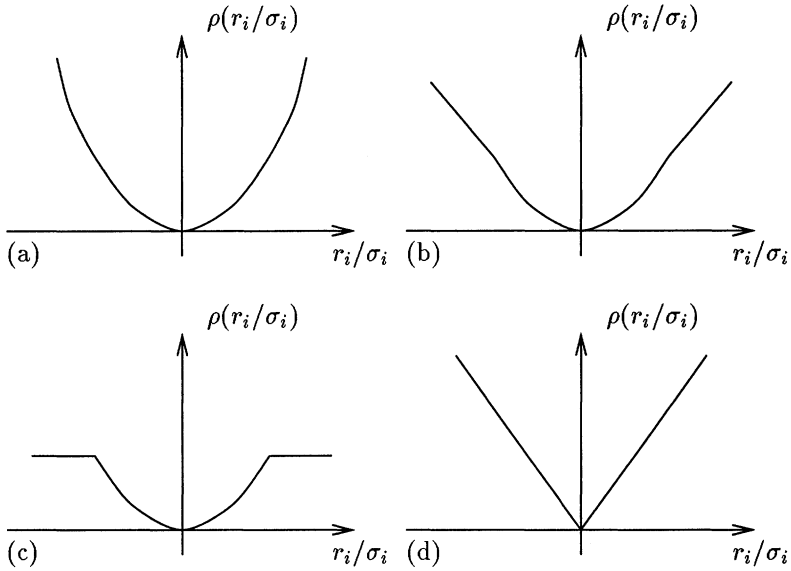


Figure 9.10. Objective functions. (a) Quadratic. (b) Quadratic-linear. (c) Quadratic-constant. (d) Absolute-value.

process is normally initialized with a regular WLS estimator, and the measurements with large residuals have their weights gradually reduced, according to a non-quadratic objective function. The problem then becomes how to select the set of suspect measurements. If this modification on weights is based purely on the magnitude of weighted residuals, good data may be rejected whereas bad data can be classified as good data. One alternative is to use the normalized residuals, which will work for cases of single and multiple non-conforming bad data, as discussed above. Whatever the choice, however, the combinatorial nature of the bad data problem will always be present in cases of conforming bad data, since in these cases a bad data can initially appear as having small or even zero residual.

9.6.2 Least Absolute Value Estimator

The least absolute value estimator, LAV, (or weighted least absolute value, WLAV) can be seen as a non-quadratic estimator in which the range of the quadratic objective function is reduced to zero and a linear objective is used instead, as illustrated in Fig. 9.10(d). This objective function can be written as follows:

$$\rho(\mathbf{r}^w) = \sum_{i=1}^m |r_i^w| = \sum_{i=1}^m \sigma_i^{-1} |z_i - h_i(\mathbf{x})|$$

The solution to this minimization problem is given by n basic measurements which fits perfectly the state estimate (these measurements have zero residuals). Only the remaining $m - n$ measurements, the non-basic ones, can present non-zero residuals. In the unidimensional example discussed above, the least absolute value estimator yields a state estimate equal to one of the correct values, regardless of the magnitude of the bad data. For large and complex systems, however, things in general are more complicated, due both to the computational effort and to the existence of leverage points which makes unwanted bad data to be selected as good data.

Leverage points are measurements that have significant influence on state estimation, regardless of the relatively low weights they have in a least absolute value objective function. These measurements normally occur in connection with low impedance branches and nodal injection measurements, i.e., (a) flows in low impedance branches, (b) injections at nodes adjacent to a low impedance branch, and (c) injections into nodes with a large number of incident branches. In practical terms, this means that the designed automatic bad data rejection capability of the LAV estimator does not work in all cases. Although various methods to identify leverage points and to reduce their effect on state estimation have been suggested in the literature, this still is an active research area. Alternatively, modeling low impedance branches as zero impedance branches with the corresponding through-flows considered as additional state variables, as in the generalized state estimation approach, may help to minimize the occurrence of leverage points.

9.6.3 Least Median of Squares Estimator

The least median of squares estimator, LMS, represents a radical change in comparison with all the previous objective functions, since in this case the objective is not based on the sum of the squared residuals or of their absolute values but on a single residual, i.e., the median squared residual. Mathematically this estimator can be formulated as follows:

$$\text{Minimize}_{\hat{\mathbf{x}}} \quad \text{median}_i \quad (r_i/\sigma_i)^2$$

i.e., for all alternative state estimates $\hat{\mathbf{x}}$, it is sought the one presenting the minimum median squared residual. Since nonzero residuals occur only when there are redundant measurements, i.e., $m > n$, the objective function above is used only for $n = 1$ or $n = 2$, whereas for multidimensional states ($n > 2$) the median is given by the following expression:

$$md = \frac{m}{2} + \frac{n+1}{2}$$

Hence, the problem can be rewritten as follows:

$$\text{Minimize}_{\hat{\mathbf{x}}} \quad (r_{md}/\sigma_{md})^2$$

As originally proposed in the literature, samples with n measurements each, for which the network is minimally observable, are drawn, and for each one of these the state is calculated and the weighted residuals of the remaining $m - n$ measurements are computed. The optimal solution is given by the sample that minimizes the above objective function. Clearly this is a combinatorial search problem of the same type discussed above, and thus combinatorial techniques such as the branch-and-bound and tabu search, along with the appropriate heuristic methods to guide the search, can be used to make the search for the optimal solution (or solutions) more systematical (see Prob. 8 at the end of the chapter).

9.7 HISTORICAL NOTES AND REFERENCES

The importance of the residual sensitivity matrix in the processing of bad data was first recognized by Handschin, Schweppe, Kohlas, and Fiechter [1975], and a heuristic approach for processing multiple analog, structure and parameter errors was proposed. Ma Zhi-quiang [1981] extended the use of the residual sensitivity matrix for the detection and identification of multiple bad data. A new interpretation of the residual estimates which are seen as the result of a state estimation based on parts of the sensitivity relation $\mathbf{S} \hat{\mathbf{r}} = \mathbf{e}$ was introduced by Xiang, Wang, and Yu [1981, 82, 84] (In this approach the detection test relies upon a $J(\hat{\mathbf{x}})$ defined for a reduced set of suspect measurements.) The updating of the $J(\hat{\mathbf{x}})$ index was used in the hypothesis testing approach proposed by Mili, Van Cutsem, and Ribbens-Pavella [1984]. The use of bad data pockets for processing interacting analog and status errors was proposed by Alsaç, Vempati, Stott, and Monticelli [1998]. The combinatorial optimization approach to bad data processing was proposed by Monticelli, Wu and Yen [1986]. A review of bad data detection and identification methods was presented by Koglin, Neisius, Beißler and Schmitt [1990]. A comprehensive discussion about Tabu Search can be found in the book by Glover and Laguna [1997].

Robust estimators were introduced in statistical analysis by Huber [1964] and first applied to power system state estimation by Merrill and Schweppe [1971]. The least absolute value estimator was introduced by Irving, Owen and Sterling [1978], and its bad data rejection capability was studied by Kotiuga and Vidyasagar [1982] and Falcão and Assis [1988]. The least absolute value estimator was extended to the identification of network topology (status of switching devices) by Abur, Kim, and Celik [1995] and Singh and Alvarado [1995], and a method for dealing with leverage measurements was proposed (Abur, Magnano and Alvarado [1997]). The least median of squares estimator was proposed by Mili, Phaniraj, and Rousseeuw [1991]. A critical overview of the actual bad data rejection capabilities of a variety of robust estimators is given in the Chapter 1 of the book by Rousseeuw and Leroy [1987], where the least trimmed squares estimator, LTS, is suggested as a means to overcome the

slow convergence rate of the least median of squares approach for regression analysis.

9.8 PROBLEMS

- 1. Consider the three bus system in Fig. 9.2 with the same set of data used in Examples 9.1 and 9.2. Formulate and solve a state estimation problem for testing the hypothesis $P_3 = 0$ (equality constraint).
- 2. Consider the three bus system in Fig. 9.5 with the same set of data used in Example 9.7. (a) Consider measurements P_2 and P_{21} as suspect data and rerun state estimation considering their weighting factors divided by 100. (b) Compare with the results obtained by making the two measurements dormant using Eq. 9.1 to correct the measured values. (c) Simulate the effect of considering measurements P_2 and P_{21} as perfect data by solving the WLS state estimation with their weighting factors multiplied by 100. (d) Compare with the results obtained for the same situation using Eq. (9.8) to simulate the effect of imposing these two measurements as equality constraints.
- 3. Consider the system given in Fig. 9.5 with the same measurements as in Example 9.7. Remove all suspect measurements, one at a time, by the repeated application of the largest normalized residual criterion. Consider a probability of false alarm of 1.0 %. Compare with the results given in Example 9.7.
- 4. The branch-and-bound algorithm discussed in Sec. 9.5.3 finds the set of optimal solutions considering that all the measurements have the same reliability (objective function provided by Eq. (9.6)). Modify the algorithm to handle the objective function in Eq. (9.4).
- 5. The branch-and-bound algorithm discussed in Sec. 9.5.3 finds the set of optimal solutions when all measurements have the same reliability (objective function of Eq. (9.5)). Modify the algorithm to allow the determination of all feasible solutions with costs equal to the optimal cost minus one.
- 6. Extend Eqs. (9.1) and (9.8) such that in the set k both dormant and perfect measurements can be included simultaneously (i.e., having both measurements that are simulated as being dormant and measurements that are simulated as being perfect).
- 7. Adapt the branch-and-bound algorithm discussed in Sec. 9.5.3 to use the least median square residual as objective function.
- 8. Consider again the system given in Fig. 9.5 with the same measurements as in Example 9.7. Find (a) the weighted least absolute value and (b) least median of squares solution to this problem.

References

- Abur, A., Kim, H., and Celik, M. K., "Identifying the unknown circuit-breaker statuses in power network", *IEEE Trans. on Power Syst.*, Vol. 10, No. 4, pp. 2029-2037, Nov. 1995.
- Abur, A., Magnano, F. H., and Alvarado, F.L. "Elimination of leverage measurements via matrix stretching", *Int J Elec. Power and Energy Syst.*, Vol. 19, pp. 557-562, Nov. 1997.
- Alsaç, O., Vempati, N., Stott, B., and Monticelli, A., "Generalized state estimation", *IEEE Trans. on Power Syst.*, Vol. 13, No. 3, pp. 1069-1075, Aug. 1998.
- Singh, H. and Alvarado F. L., "Network topology determination using least absolute value state estimation", *IEEE Trans. on Power Syst.*, Vol. 10, No. 3, pp. 1159-1165, Aug 1995.
- Falcão, D. and Assis, S. M., "Linear-programming state estimation - error analysis and gross error identification", *IEEE Trans. on Power Syst.*, Vol. 3, No. 3, pp. 809-815, Aug. 1988.
- Glover, F. and Laguna, M., *Tabu Search*, Kluwer Academic Publishers, 1997.
- Handschin, E., Schweppe, F. C., Kohlas, J., and Fiechter, A., "Bad data analysis for power systems state estimation", *IEEE Trans. Power App. and Syst.*, Vol. 94, No. 2, pp. 329-337, March/April 1975.
- Huber, P. J., "Robust estimation of a location parameter", *Ann. Math. Statist.*, Vol. 35, pp. 73-101, 1964,
- Irving, M. R., Owen, R. C., and Sterling, M., "Power system state estimation using linear programming", *Proc. IEE*, Vol. 125, pp. 978-885, Sept. 1978.
- Koglin, H-J, Neisius, Th., Beißler, G., and Schmitt, K. D., "Bad data detection and identification", *Int. J. Elec. Power*, Vol. 12, No. 2, pp. 94-103, April 1990.
- Kotiuga, W. W. and Vidyasagar, M., "Bad data rejection properties of weighted least absolute value techniques applied to static state estimation", *IEEE Trans. Power App. and Syst.*, Vol. 101, No. 4, pp. 844-853, April 1982.
- Ma Zhi-qiang, "Bad data reestimation-identification using residual sensitivity matrix", *Proc. 7th Power System Computation Conf., PSCC*, pp. 1056-1050, Lausanne, July 1981.
- Merril, H. M. and Schweppe, F. C., "Bad data suppression in power system state estimation", *IEEE Trans. Power App. and Syst.*, Vol. 90, No. 6, pp. 2718-2725, Nov./Dec. 1971.
- Mili, L., Van Cutsem, Th., and Ribbens-Pavella, M., "Hypothesis testing identification: A new method for bad data analysis in power system state estimation", *IEEE Trans. Power App. and Syst.*, Vol. 103, No. 11, pp. 3239-3252, Nov. 1984.
- Mili, L., Phaniraj, V., and Rousseeuw, P. J., "Least median of squares estimation in power-systems", *IEEE Trans. Power App. and Syst.*, Vol. 6, No. 2, pp. 511-523, May 1991.

- Monticelli, A. , Wu, F .F. , and Yen, M., "Multiple bad data identification for state estimation by combinatorial optimization", IEEE Trans Power Distr., Vol. 1, No. 3, pp. 361-369, July 1986.
- Rousseeuw, P. J. and Leroy, A. M., *Robust Regression and Outlier Detection*, John Wiley & Sons, 1987.
- Scheppe, F. C., Wildes, J., and Rom, D., "Power system static estate estimation: Parts I, II, and III", Power Industry Computer Conference, PICA, Denver, Colorado, June 1969.
- Xiang Nian-De, Wang Shi-Ying, and Yu Er-Keng, "Estimation and identification of multiple bad data in power system state estimation", Proc. 7th Power System Computation Conference, PSCC, Lausanne, July 1981.
- Xiang Nian-De, Wang Shi-Ying, and Yu Er-Keng, "A new approach for detection and identification of multiple bad data in power system state estimation", IEEE Trans. Power App. and Syst., Vol. 101, No.. 2, pp. 454-462, Feb. 1982.
- Xiang Nian-De, Wang Shi-Ying, and Yu Er-Keng, "An application of estimation-identification approach of multiple bad data in power system state estimation", IEEE Trans. Power App. and Syst., Vol. 103, No. 2, pp. 225-233, Feb. 1984.

10 AC STATE ESTIMATOR

This chapter reviews the principal characteristics of the Jacobian matrix associated with the measurement model equations of a WLS state estimator based on an ac (nonlinear) network model. (The dc (linear) model was discussed in Chap. 3.)

10.1 REVIEW OF THE PROBLEM FORMULATION

The ac estimator is based on the nonlinear measurement model

$$\mathbf{z} = \mathbf{h}(\mathbf{x}) + \mathbf{e} \quad (10.1)$$

where \mathbf{z} is the measurement vector (m -vector), \mathbf{x} is the true state vector (n -vector, $n < m$), $\mathbf{h}(\cdot)$ is a nonlinear vector function relating measurements to states (m -vector), and \mathbf{e} is the measurement error vector (m -vector). The elements of \mathbf{e} are assumed to have zero mean; the corresponding variance matrix is denoted by \mathbf{R}_z .

The optimality conditions are applied to the performance index $J(\mathbf{x})$ which is expressed as follows:

$$J(\mathbf{x}) = \frac{1}{2} \sum_{j=1}^m \left(\frac{z_j - h_j(\mathbf{x})}{\sigma_j} \right)^2$$

where σ_j is the (j, j) th element of the measurement error covariance matrix, \mathbf{R}_z . The first-order optimal condition for this model can be written as:

$$\mathbf{g}(\mathbf{x}) = \frac{\partial J(\mathbf{x})}{\partial \mathbf{x}} = - \sum_{j=1}^m \left(\frac{z_j - h_j(\mathbf{x})}{\sigma_j} \right) \frac{\partial h_j(\mathbf{x})}{\partial \mathbf{x}} = \mathbf{0} \quad (10.2)$$

where $\mathbf{g}(\mathbf{x})$ denotes the gradient of $J(\mathbf{x})$. The root of the nonlinear equation $\mathbf{g}(\mathbf{x}) = \mathbf{0}$ can be found using the Newton Raphson method. By Taylor expansion the gradient function can be approximated by

$$\mathbf{g}(\mathbf{x} + \Delta \mathbf{x}) \simeq \mathbf{g}(\mathbf{x}) + \mathbf{G}(\mathbf{x}) \Delta \mathbf{x} \quad (10.3)$$

where $\mathbf{G}(\mathbf{x})$ is the Hessian matrix of $J(\mathbf{x})$:

$$\mathbf{G}(\mathbf{x}) = \frac{\partial^2 J(\mathbf{x})}{\partial \mathbf{x}^2} = \mathbf{H}'(\mathbf{x}) \mathbf{R}_z^{-1} \mathbf{H}(\mathbf{x}) - \sum_{j=1}^m \Delta z \frac{\partial^2 h_j(\mathbf{x})}{\partial \mathbf{x}^2}$$

In the Gauss Newton method the term that depends on the second derivatives is ignored and the gain matrix becomes (see Sec. 2.6 in Chap. 2):

$$\mathbf{G}(\mathbf{x}) = \mathbf{H}'(\mathbf{x}) \mathbf{R}_z^{-1} \mathbf{H}(\mathbf{x})$$

The state estimate $\hat{\mathbf{x}}$ is then obtained by the following iterative procedure:

$$\begin{aligned} \left(\mathbf{H}'(\mathbf{x}^\nu) \mathbf{R}_z^{-1} \mathbf{H}(\mathbf{x}^\nu) \right) \Delta \hat{\mathbf{x}}^\nu &= \mathbf{H}'(\mathbf{x}^\nu) \mathbf{R}_z^{-1} \Delta z(\mathbf{x}^\nu) \\ \hat{\mathbf{x}}^{\nu+1} &= \hat{\mathbf{x}}^\nu + \Delta \hat{\mathbf{x}}^\nu \end{aligned}$$

where $\mathbf{H}(\mathbf{x})$ is the Jacobian matrix of $\mathbf{g}(\mathbf{x})$.

Remark: For a state estimation problem solved by the Gauss Newton method see Example 12.6 in Chap. 12.

10.2 FLOW MEASUREMENTS

In this section, the contribution of flow measurement to the measurement Jacobian matrix is discussed and a unified branch model for analyzing them is considered. Three types of flow measurements are analyzed: active power flow measurements (P_{km}^{meas}), reactive power flow measurements (Q_{km}^{meas}), and current magnitude measurements ($|I_{km}|^{meas}$).

10.2.1 Power Flow Measurements

The expressions for active and reactive power flows in the unified branch model in Fig. 10.1 are as follows:

$$P_{km} = a_{km}^2 V_k^2 g_{km} - a_{km} V_k a_{mk} V_m g_{km} \cos(\theta_{km} + \varphi_{km} - \varphi_{mk}) +$$

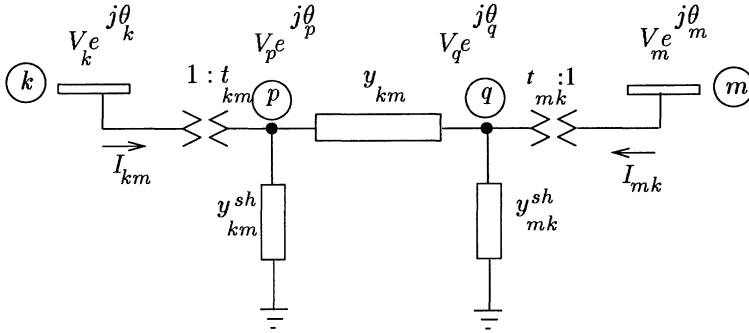


Figure 10.1. Unified branch model ($t = ae^{j\varphi}$).

$$-a_{km}V_k a_{mk}V_m b_{km} \sin(\theta_{km} + \varphi_{km} - \varphi_{mk})$$

$$Q_{km} = -a_{km}^2 V_k^2 (b_{km} + b_{km}^{sh}) + a_{km}V_k a_{mk}V_m b_{km} \cos(\theta_{km} + \varphi_{km} - \varphi_{mk}) + \\ -a_{km}V_k a_{mk}V_m g_{km} \sin(\theta_{km} + \varphi_{km} - \varphi_{mk})$$

where for transmission lines, $a_{km} = a_{mk} = 1$ and $\varphi_{km} = \varphi_{mk} = 0$; for in-phase transformers with tap on the k -bus side, $y_{km}^{sh} = y_{mk}^{sh} = 0$, $\varphi_{km} = \varphi_{mk} = 0$, and $a_{mk} = 1$; and for pure phase-shifters with regulators on the k -bus side, $y_{km}^{sh} = y_{mk}^{sh} = 0$, $a_{km} = a_{mk} = 1$, and $\varphi_{mk} = 0$.

Now, considering P_{km} and Q_{km} as metered, and V_k , V_m , θ_k , and θ_m as state variables, the corresponding Jacobian matrix elements would be the following:

$$\frac{\partial P_{km}}{\partial \theta_k} = a_{km}V_k a_{mk}V_m g_{km} \sin(\theta_{km} + \varphi_{km} - \varphi_{mk}) + \\ -a_{km}V_k a_{mk}V_m b_{km} \cos(\theta_{km} + \varphi_{km} - \varphi_{mk})$$

$$\frac{\partial P_{km}}{\partial \theta_m} = -a_{km}V_k a_{mk}V_m g_{km} \sin(\theta_{km} + \varphi_{km} - \varphi_{mk}) + \\ a_{km}V_k a_{mk}V_m b_{km} \cos(\theta_{km} + \varphi_{km} - \varphi_{mk})$$

$$\frac{\partial P_{km}}{\partial V_k} = 2 a_{km}^2 V_k g_{km} - a_{km} a_{mk} V_m g_{km} \cos(\theta_{km} + \varphi_{km} - \varphi_{mk}) + \\ -a_{km} a_{mk} V_m b_{km} \sin(\theta_{km} + \varphi_{km} - \varphi_{mk})$$

$$\frac{\partial P_{km}}{\partial V_m} = -a_{km} a_{mk} V_k g_{km} \cos(\theta_{km} + \varphi_{km} - \varphi_{mk}) + \\ -a_{km} a_{mk} V_k b_{km} \sin(\theta_{km} + \varphi_{km} - \varphi_{mk})$$

$$\frac{\partial Q_{km}}{\partial \theta_k} = -a_{km}V_k a_{mk}V_m b_{km} \sin(\theta_{km} + \varphi_{km} - \varphi_{mk}) + \\ -a_{km}V_k a_{mk}V_m g_{km} \cos(\theta_{km} + \varphi_{km} - \varphi_{mk})$$

$$\frac{\partial Q_{km}}{\partial \theta_m} = a_{km}V_k a_{mk}V_m b_{km} \sin(\theta_{km} + \varphi_{km} - \varphi_{mk}) +$$

$$\begin{aligned}
& + a_{km} V_k a_{mk} V_m g_{km} \cos(\theta_{km} + \varphi_{km} - \varphi_{mk}) \\
\frac{\partial Q_{km}}{\partial V_k} &= -2 a_{km}^2 V_k (b_{km} + b_{km}^{sh}) + a_{km} a_{mk} V_m b_{km} \cos(\theta_{km} + \varphi_{km} - \varphi_{mk}) + \\
& - a_{km} a_{mk} V_m g_{km} \sin(\theta_{km} + \varphi_{km} - \varphi_{mk}) \\
\frac{\partial Q_{km}}{\partial V_m} &= a_{km} a_{mk} V_k b_{km} \cos(\theta_{km} + \varphi_{km} - \varphi_{mk}) + \\
& - a_{km} a_{mk} V_k g_{km} \sin(\theta_{km} + \varphi_{km} - \varphi_{mk})
\end{aligned}$$

There are situations in which transformer tap magnitudes and phase shifts must be estimated. In these cases, one possible approach consists of including tap magnitudes and phase shifts as additional state variables. For example, if a_{km} and φ_{km} of the unified branch model represented in Fig. 10.1 are included as extra states (whereas a_{mk} and φ_{mk} are assumed to have constant, known values), two new columns have to be added to the Jacobian matrix; in this case the contributions corresponding to power flow measurements P_{km} and Q_{km} are as follows:

$$\begin{aligned}
\frac{\partial P_{km}}{\partial a_{km}} &= 2 a_{km} V_k^2 - a_{mk} V_k V_m g_{km} \cos(\theta_{km} + \varphi_{km} - \varphi_{mk}) + \\
& - a_{mk} V_k V_m b_{km} \sin(\theta_{km} + \varphi_{km} - \varphi_{mk}) \\
\frac{\partial Q_{km}}{\partial a_{km}} &= -2 a_{km} V_k^2 (b_{km} + b_{km}^{sh}) + a_{mk} V_k V_m b_{km} \cos(\theta_{km} + \varphi_{km} - \varphi_{mk}) + \\
& - a_{mk} V_k V_m g_{km} \sin(\theta_{km} + \varphi_{km} - \varphi_{mk}) \\
\frac{\partial P_{km}}{\partial \varphi_{km}} &= a_{km} a_{mk} V_k V_m g_{km} \sin(\theta_{km} + \varphi_{km} - \varphi_{mk}) + \\
& - a_{km} a_{mk} V_k V_m b_{km} \cos(\theta_{km} + \varphi_{km} - \varphi_{mk}) \\
\frac{\partial Q_{km}}{\partial \varphi_{km}} &= -a_{km} a_{mk} V_k V_m b_{km} \sin(\theta_{km} + \varphi_{km} - \varphi_{mk}) + \\
& - a_{km} a_{mk} V_k V_m g_{km} \cos(\theta_{km} + \varphi_{km} - \varphi_{mk})
\end{aligned}$$

Example 10.1

A 750 kV transmission line section has a series impedance of $0.00072 + j0.0175$ p.u., a total shunt admittance of 877.5 MVar, a voltage magnitude at the terminal buses of 0.984 p.u. and 0.962 p.u., and voltage angle spread of 22° (100 MVA base).

For a transmission line $a_{km} = a_{mk} = 1$ and $\varphi_{km} = \varphi_{mk} = 0$; hence, the contributions of the power flow measurements (P_{km} and Q_{km}) to the Jacobian matrix (see Fig. 10.2) are as follows:

$$\begin{aligned}
\frac{\partial P_{km}}{\partial \theta_k} &= V_k V_m g_{km} \sin \theta_{km} - V_k V_m b_{km} \cos \theta_{km} \\
\frac{\partial P_{km}}{\partial \theta_m} &= -V_k V_m g_{km} \sin \theta_{km} + V_k V_m b_{km} \cos \theta_{km}
\end{aligned}$$

$$\begin{aligned} \frac{\partial P_{km}}{\partial V_k} &= 2 V_k g_{km} - V_m g_{km} \cos \theta_{km} - V_m b_{km} \sin \theta_{km} \\ \frac{\partial P_{km}}{\partial V_m} &= -V_k g_{km} \cos \theta_{km} - V_k b_{km} \sin \theta_{km} \\ \frac{\partial Q_{km}}{\partial \theta_k} &= -V_k V_m b_{km} \sin \theta_{km} - V_k V_m g_{km} \cos \theta_{km} \\ \frac{\partial Q_{km}}{\partial \theta_m} &= V_k V_m b_{km} \sin \theta_{km} + V_k V_m g_{km} \cos \theta_{km} \\ \frac{\partial Q_{km}}{\partial V_k} &= -2 V_k (b_{km} + b_{km}^{sh}) + V_m b_{km} \cos \theta_{km} - V_m g_{km} \sin \theta_{km} \\ \frac{\partial Q_{km}}{\partial V_m} &= V_k b_{km} \cos \theta_{km} - V_k g_{km} \sin \theta_{km} \end{aligned}$$

where $V_k = 0.984$ p.u., $V_m = 0.962$ p.u., $\theta_{km} = 22^\circ$, $g_{km} = 2.347$ p.u., $b_{km} = -57.05$ p.u., and $b_{km}^{sh} = 4.387$ p.u.

θ_k	θ_m	V_k	V_m	
\vdots	\vdots	\vdots	\vdots	
$\dots \frac{\partial P_k}{\partial \theta_k} \dots$	$\frac{\partial P_k}{\partial \theta_m} \dots$	$\dots \frac{\partial P_k}{\partial V_k} \dots$	$\dots \frac{\partial P_k}{\partial V_m} \dots$	P_{km}^{meas}
\vdots	\vdots	\vdots	\vdots	
$\dots \frac{\partial Q_k}{\partial \theta_k} \dots$	$\frac{\partial Q_k}{\partial \theta_m} \dots$	$\dots \frac{\partial Q_k}{\partial V_k} \dots$	$\dots \frac{\partial Q_k}{\partial V_m} \dots$	Q_{km}^{meas}
\vdots	\vdots	\vdots	\vdots	

Figure 10.2. Transmission line (Contribution of power flow measurement to Jacobian matrix).

Now, introducing the numerical values, yields:

$$\begin{aligned} \frac{\partial P_{km}}{\partial \theta_k} &= 50.9 & \frac{\partial P_{km}}{\partial \theta_m} &= -50.9 & \frac{\partial P_{km}}{\partial V_k} &= 23.1 & \frac{\partial P_{km}}{\partial V_m} &= 18.9 \\ \frac{\partial Q_{km}}{\partial \theta_k} &= 18.2 & \frac{\partial Q_{km}}{\partial \theta_m} &= -18.2 & \frac{\partial Q_{km}}{\partial V_k} &= 51.9 & \frac{\partial Q_{km}}{\partial V_m} &= -52.9 \end{aligned}$$

For the sake of comparison, at flat-start these derivatives are as follows:

$$\begin{aligned} \frac{\partial P_{km}}{\partial \theta_k} &= 57.0 & \frac{\partial P_{km}}{\partial \theta_m} &= -57.0 & \frac{\partial P_{km}}{\partial V_k} &= 2.35 & \frac{\partial P_{km}}{\partial V_m} &= -2.35 \\ \frac{\partial Q_{km}}{\partial \theta_k} &= -2.35 & \frac{\partial Q_{km}}{\partial \theta_m} &= 2.35 & \frac{\partial Q_{km}}{\partial V_k} &= 48.3 & \frac{\partial Q_{km}}{\partial V_m} &= -57.0 \end{aligned}$$

This shows that the coupling elements vary significantly with the angle spread, even for lines with very high x/r ratios, as illustrated here.

Example 10.2

A 500/750 kV transformer with tap ratio of 1.050 : 1.0 referred to the low voltage side, has a negligible series resistance and a leakage reactance of 0.00623 p.u., terminal voltage magnitudes of 1.023 p.u. and 0.968 p.u., and angle spread of 5.3°.

In this case $a_{km} = 1./1.050$, $a_{mk} = 1.0$, and $\varphi_{km} = \varphi_{mk} = 0$. The contributions of the power flow measurements (P_{km} and Q_{km}) to the Jacobian matrix (see Fig. 10.3) would be:

$$\begin{aligned} \frac{\partial P_{km}}{\partial \theta_k} &= a_{km} V_k V_m g_{km} \sin \theta_{km} - a_{km} V_k V_m b_{km} \cos \theta_{km} \\ \frac{\partial P_{km}}{\partial \theta_m} &= -a_{km} V_k V_m g_{km} \sin \theta_{km} + a_{km} V_k V_m b_{km} \cos \theta_{km} \\ \frac{\partial P_{km}}{\partial V_k} &= 2 a_{km}^2 V_k g_{km} - a_{km} V_m g_{km} \cos \theta_{km} - a_{km} V_m b_{km} \sin \theta_{km} \\ \frac{\partial P_{km}}{\partial V_m} &= -a_{km} V_k g_{km} \cos \theta_{km} - a_{km} V_k b_{km} \sin \theta_{km} \\ \frac{\partial P_{km}}{\partial a_{km}} &= 2 a_{km} V_k^2 - V_k V_m g_{km} \cos \theta_{km} - V_k V_m b_{km} \sin \theta_{km} \\ \frac{\partial Q_{km}}{\partial \theta_k} &= -a_{km} V_k V_m b_{km} \sin \theta_{km} - a_{km} V_k V_m g_{km} \cos \theta_{km} \\ \frac{\partial Q_{km}}{\partial \theta_m} &= a_{km} V_k V_m b_{km} \sin \theta_{km} + a_{km} V_k V_m g_{km} \cos \theta_{km} \\ \frac{\partial Q_{km}}{\partial V_k} &= -2 a_{km}^2 V_k (b_{km} + b_{km}^{sh}) + \\ &\quad + a_{km} V_m b_{km} \cos \theta_{km} - a_{km} V_m g_{km} \sin \theta_{km} \\ \frac{\partial Q_{km}}{\partial V_m} &= a_{km} V_k b_{km} \cos \theta_{km} - a_{km} V_k g_{km} \sin \theta_{km} \\ \frac{\partial Q_{km}}{\partial a_{km}} &= -2 a_{km} V_k^2 (b_{km} + b_{km}^{sh}) + \\ &\quad + V_k V_m b_{km} \cos \theta_{km} - V_k V_m g_{km} \sin \theta_{km} \end{aligned}$$

where $V_k = 1.023$ p.u., $V_m = 0.968$ p.u., $\theta_{km} = 5.3^\circ$, $g_{km} = 0.0$ p.u., and $b_{km} = -160.5$ p.u.

The following contributions to the Jacobian matrix would result:

$$\begin{aligned} \frac{\partial P_{km}}{\partial \theta_k} &= 151. & \frac{\partial P_{km}}{\partial \theta_m} &= -151. \\ \frac{\partial P_{km}}{\partial V_k} &= 13.7 & \frac{\partial P_{km}}{\partial V_m} &= 14.4 & \frac{\partial P_{km}}{\partial a_{km}} &= 16.7 \end{aligned}$$

θ_k	θ_m	V_k	V_m	a_{km}	
\vdots	\vdots	\vdots	\vdots	\vdots	
$\dots \frac{\partial P_k}{\partial \theta_k} \dots$	$\frac{\partial P_k}{\partial \theta_m} \dots$	$\frac{\partial P_k}{\partial V_k} \dots$	$\frac{\partial P_k}{\partial V_m} \dots$	$\frac{\partial P_k}{\partial a_{km}}$	P_{km}^{meas}
\vdots	\vdots	\vdots	\vdots	\vdots	
$\dots \frac{\partial Q_k}{\partial \theta_k} \dots$	$\frac{\partial Q_k}{\partial \theta_m} \dots$	$\frac{\partial Q_k}{\partial V_k} \dots$	$\frac{\partial Q_k}{\partial V_m} \dots$	$\frac{\partial Q_k}{\partial a_{km}}$	Q_{km}^{meas}
\vdots	\vdots	\vdots	\vdots	\vdots	
$\dots 0 \dots$	$0 \dots$	$0 \dots$	$0 \dots$	$1 \dots$	a_{km}^{meas}
\vdots	\vdots	\vdots	\vdots	\vdots	

Figure 10.3. In-phase transformer (Contributions of power flow and tap magnitude measurements to Jacobian matrix).

$$\frac{\partial Q_{km}}{\partial \theta_k} = 14.0 \quad \frac{\partial Q_{km}}{\partial \theta_m} = -14.0$$

$$\frac{\partial Q_{km}}{\partial V_k} = 151. \quad \frac{\partial Q_{km}}{\partial V_m} = -156. \quad \frac{\partial Q_{km}}{\partial a_{km}} = -162.$$

10.2.2 Current Magnitude Measurements

Consider again the unified branch model in Fig. 10.1. The current magnitude $|I_{km}|$ can be expressed as:

$$|I_{km}| = \frac{(P_{km}^2 + Q_{km}^2)^{1/2}}{V_k}$$

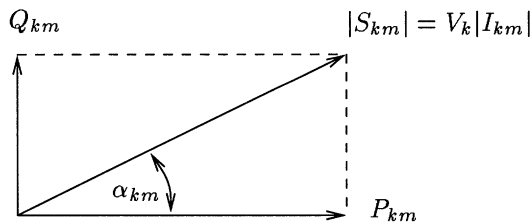


Figure 10.4. Current magnitude measurements do not carry information on power factor angle α_{km} .

θ_k	θ_m	V_k	V_m	
\vdots	\vdots	\vdots	\vdots	
$\dots \frac{\partial I_{km} }{\partial \theta_k} \dots \frac{\partial I_{km} }{\partial \theta_m} \dots \frac{\partial I_{km} }{\partial V_k} \dots \frac{\partial I_{km} }{\partial V_m} \dots$				$ I_{km} ^{meas}$
\vdots	\vdots	\vdots	\vdots	

Figure 10.5. Transmission line (Contributions of current magnitude measurement to Jacobian matrix).

The corresponding Jacobian matrix elements are:

$$\begin{aligned} \frac{\partial |I_{km}|}{\partial \theta_k} &= \frac{P_{km} \frac{\partial P_{km}}{\partial \theta_k} + Q_{km} \frac{\partial Q_{km}}{\partial \theta_k}}{V_k |I_{km}|} = \cos \alpha_{km} \frac{\partial P_{km}}{\partial \theta_k} + \sin \alpha_{km} \frac{\partial Q_{km}}{\partial \theta_k} \\ \frac{\partial |I_{km}|}{\partial \theta_m} &= \frac{P_{km} \frac{\partial P_{km}}{\partial \theta_m} + Q_{km} \frac{\partial Q_{km}}{\partial \theta_m}}{V_k |I_{km}|} = \cos \alpha_{km} \frac{\partial P_{km}}{\partial \theta_m} + \sin \alpha_{km} \frac{\partial Q_{km}}{\partial \theta_m} \\ \frac{\partial |I_{km}|}{\partial V_k} &= \frac{P_{km} \frac{\partial P_{km}}{\partial V_k} + Q_{km} \frac{\partial Q_{km}}{\partial V_k}}{V_k |I_{km}|} = \cos \alpha_{km} \frac{\partial P_{km}}{\partial V_k} + \sin \alpha_{km} \frac{\partial Q_{km}}{\partial V_k} \\ \frac{\partial |I_{km}|}{\partial V_m} &= \frac{P_{km} \frac{\partial P_{km}}{\partial V_m} + Q_{km} \frac{\partial Q_{km}}{\partial V_m}}{V_k |I_{km}|} = \cos \alpha_{km} \frac{\partial P_{km}}{\partial V_m} + \sin \alpha_{km} \frac{\partial Q_{km}}{\partial V_m} \end{aligned}$$

where $\frac{\partial P_{km}}{\partial \theta_k}$, $\frac{\partial P_{km}}{\partial \theta_m}$, $\frac{\partial P_{km}}{\partial V_k}$, and $\frac{\partial P_{km}}{\partial V_m}$ are the same as for the power flow measurement contributions given above.

Under certain circumstances tap magnitudes and/or phase shifts have to be estimated. If, for example, φ_{km} and a_{km} of the unified model in Fig. 10.1 are included as state variables, the following extra terms are added to the Jacobian matrix:

$$\begin{aligned} \frac{\partial |I_{km}|}{\partial \varphi_{km}} &= \frac{P_{km} \frac{\partial P_{km}}{\partial \varphi_{km}} + Q_{km} \frac{\partial Q_{km}}{\partial \varphi_{km}}}{V_k |I_{km}|} = \cos \alpha_{km} \frac{\partial P_{km}}{\partial \varphi_{km}} + \sin \alpha_{km} \frac{\partial Q_{km}}{\partial \varphi_{km}} \\ \frac{\partial |I_{km}|}{\partial a_{km}} &= \frac{P_{km} \frac{\partial P_{km}}{\partial a_{km}} + Q_{km} \frac{\partial Q_{km}}{\partial a_{km}}}{V_k |I_{km}|} = \cos \alpha_{km} \frac{\partial P_{km}}{\partial a_{km}} + \sin \alpha_{km} \frac{\partial Q_{km}}{\partial a_{km}} \end{aligned}$$

Example 10.3

Consider again the 750 kV transmission line section in Example 10.1, for which $V_k = 0.984$ p.u., $V_m = 0.962$ p.u., $\theta_{km} = 22^\circ$, $g_{km} = 2.347$ p.u., $b_{km} = -57.05$ p.u., and $b_{km}^{sh} = 4.387$ p.u. In terms of the unified branch model in Fig. 10.1, $a_{km} = a_{mk} = 1.0$ and $\varphi_{km} = \varphi_{mk} = 0.0$.

The active and reactive power flows in this line section under the given conditions are 2044 MW and 8.4 MVar, respectively (see Example 4.6 in Chap. 4). Thus, the power factor angle α is close to zero. In this case, a good initial estimate of the current magnitude measurement contributions to the Jacobian matrix would be as follows:

$$\begin{aligned} \frac{\partial |I_{km}|}{\partial \theta_k} &\simeq \frac{\partial P_{km}}{\partial \theta_k} = 50.9 & \frac{\partial |I_{km}|}{\partial \theta_m} &\simeq \frac{\partial P_{km}}{\partial \theta_m} = -50.9 \\ \frac{\partial |I_{km}|}{\partial V_k} &\simeq \frac{\partial P_{km}}{\partial V_k} = 23.1 & \frac{\partial |I_{km}|}{\partial V_m} &\simeq \frac{\partial P_{km}}{\partial V_m} = 18.9 \end{aligned}$$

This is not always the case, however. For light load conditions, the reactive power flow may become relatively important and so the power factor angle may be closer to 90° . In these cases, the initialization of the power factor angle is crucial, and the possibility of multiple solutions may be an issue. When current measurements are not critical (i.e., are redundant), one possible strategy consists of performing one or two iterations of the state estimator in order to obtain an initial estimate of the power flows on the line (or the power factor angles) before activating current measurements. On the other hand, when current magnitude measurements are critical, safeguards must be provided to avoid convergence to undesired solutions (e.g., information from previous state estimation runs and/or forecast data may be used to do this).

10.3 BUS INJECTION MEASUREMENT*10.3.1 Power Injection Measurement*

The active and reactive power injections at a generic Bus k are given by:

$$\begin{aligned} P_k &= V_k \sum_{m \in K} V_m (G_{km} \cos \theta_{km} + B_{km} \sin \theta_{km}) \\ Q_k &= V_k \sum_{m \in K} V_m (G_{km} \sin \theta_{km} - B_{km} \cos \theta_{km}) \end{aligned}$$

where G_{km} and B_{km} are the elements of the real and imaginary parts of the nodal admittance matrix ($\mathbf{Y} = \mathbf{G} + j\mathbf{B}$):

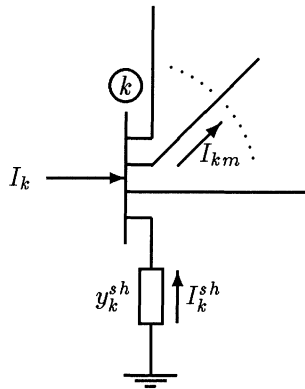


Figure 10.6. Network bus with positive sign conventions for currents and power flows.

$$\begin{aligned}
 Y_{km} &= G_{km} + jB_{km} = -t_{km}^* t_{mk} y_{km} \\
 Y_{kk} &= G_{kk} + jB_{kk} = y_k^{sh} + \sum_{m \in \Omega_k} a_{km}^2 (y_{km}^{sh} + y_{km})
 \end{aligned}$$

where $t = ae^{j\varphi}$.

Example 10.4

Consider the three bus network in Fig. 10.7. Line $k-l$ has a series impedance of $z_{kl} = 0.00072 + j0.0175$ p.u. and a total shunt admittance of $2b_{kl}^{sh} = 877.5$ MVar (100 MVA base). Transformer $m-k$ has a negligible series resistance, a leakage reactance $x_{mk} = 0.00623$ p.u., and a transform tap $a_{mk} = 0.960$ (the tap 1 : a_{mk} is placed on the m side of the transformer). The shunt susceptance of Bus k is $b_k^{sh} = -300$ MVar (it is a shunt reactor). The voltage magnitude at the three buses is known to be: $V_l = 1.020$ p.u., $V_k = 0.970$ p.u., $V_m = 1.000$ p.u. Voltage angles are $\theta_l = 0^\circ$, $\theta_k = 16.0^\circ$, and $\theta_m = 20.0^\circ$.

The parameters for the transmission line $k-l$ model are: $g_{kl} = 2.347$ p.u., $b_{kl} = -57.05$ p.u., and $b_{kl}^{sh} = 4.387$ p.u., and for the transformer $m-k$ they are: $g_{mk} = 0.0$ p.u., $b_{mk} = -160.5$ p.u., and $a_{mk} = 0.960$. For the shunt reactor $b_k^{sh} = -3.0$ p.u.

Matrices \mathbf{G} and \mathbf{B} are as follows:

$$\mathbf{G} = \begin{pmatrix} G_{ll} & G_{lk} & G_{lm} \\ G_{kl} & G_{kk} & G_{km} \\ G_{ml} & G_{mk} & G_{mm} \end{pmatrix}$$

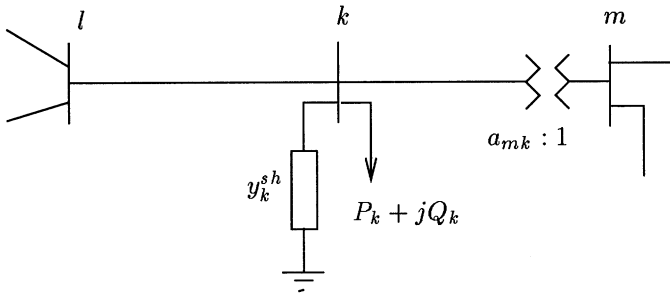


Figure 10.7. Network for Example 10.4. Active and reactive power injection at Bus k are measured (P_k^{meas} and Q_k^{meas} , respectively).

$$\begin{aligned}
 &= \begin{pmatrix} g_{kl} & -g_{kl} & 0 \\ -g_{kl} & g_{kl} & 0 \\ 0 & 0 & 0 \end{pmatrix} \\
 &= \begin{pmatrix} 2.347 & -2.347 & 0 \\ -2.347 & 2.347 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\
 \mathbf{B} &= \begin{pmatrix} B_{ll} & B_{lk} & B_{lm} \\ B_{kl} & B_{kk} & B_{km} \\ B_{ml} & B_{mk} & B_{mm} \end{pmatrix} \\
 &= \begin{pmatrix} b_{kl} & -b_{kl} & 0 \\ -b_{kl} & b_{kl}^{sh} + b_k^{sh} + b_{kl} + b_{mk} & -a_{mk} b_{mk} \\ 0 & -a_{mk} b_{mk} & a_{mk}^2 b_{mk} \end{pmatrix} \\
 &= \begin{pmatrix} -57.0 & 57.0 & 0 \\ 57.0 & -216.2 & 154.1 \\ 0 & 154.1 & -147.9 \end{pmatrix}
 \end{aligned}$$

The expressions for the active and reactive power injections at Bus k of the network given in Fig. 10.7 are

$$P_k = V_k^2 G_{kk} + V_k V_m (G_{km} \cos \theta_{km} + B_{km} \sin \theta_{km}) +$$

θ_l	θ_k	θ_m	V_l	V_k	V_m	
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	
$\dots \frac{\partial P_k}{\partial \theta_l} \dots \frac{\partial P_k}{\partial \theta_k} \dots \frac{\partial P_k}{\partial \theta_m} \dots \frac{\partial P_k}{\partial V_l} \dots \frac{\partial P_k}{\partial V_k} \dots \frac{\partial P_k}{\partial V_m} \dots$						P_k^{meas}
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	
$\dots \frac{\partial Q_k}{\partial \theta_l} \dots \frac{\partial Q_k}{\partial \theta_k} \dots \frac{\partial Q_k}{\partial \theta_m} \dots \frac{\partial Q_k}{\partial V_l} \dots \frac{\partial Q_k}{\partial V_k} \dots \frac{\partial Q_k}{\partial V_m} \dots$						Q_k^{meas}
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	

Figure 10.8. Contribution of active and reactive power injection measurements to Jacobian matrix.

$$\begin{aligned}
 &+V_k V_l (G_{kl} \cos \theta_{kl} + B_{kl} \sin \theta_{kl}) \\
 Q_k = &-V_k^2 B_{kk} + V_k V_m (G_{km} \sin \theta_{km} - B_{km} \cos \theta_{km}) + \\
 &+V_k V_l (G_{kl} \sin \theta_{kl} - B_{kl} \cos \theta_{kl})
 \end{aligned}$$

Power injection measurements P_k^{meas} and Q_k^{meas} are adjacent to state variables $V_l, V_k, V_m, \theta_l, \theta_k,$ and θ_m as shown in Fig. 10.8. Hence, each of these two measurements contributes with six elements to the Jacobian matrix. The contributions related to measurement P_k^{meas} are the following:

$$\begin{aligned}
 \frac{\partial P_k}{\partial \theta_k} &= V_k V_m (-G_{km} \sin \theta_{km} + B_{km} \cos \theta_{km}) + \\
 &+V_k V_l (-G_{kl} \sin \theta_{kl} + B_{kl} \cos \theta_{kl}) \\
 \frac{\partial P_k}{\partial \theta_l} &= V_k V_l (G_{kl} \sin \theta_{kl} - B_{kl} \cos \theta_{kl}) \\
 \frac{\partial P_k}{\partial \theta_m} &= V_k V_m (G_{km} \sin \theta_{km} - B_{km} \cos \theta_{km}) \\
 \frac{\partial P_k}{\partial V_k} &= 2V_k G_{kk} + V_m (G_{km} \cos \theta_{km} + B_{km} \sin \theta_{km}) + \\
 &+V_l (G_{kl} \cos \theta_{kl} + B_{kl} \sin \theta_{kl}) \\
 \frac{\partial P_k}{\partial V_l} &= V_k (G_{kl} \cos \theta_{kl} + B_{kl} \sin \theta_{kl}) \\
 \frac{\partial P_k}{\partial V_m} &= V_k (G_{km} \cos \theta_{km} + B_{km} \sin \theta_{km})
 \end{aligned}$$

The contributions related to measurement Q_k^{meas} are

$$\begin{aligned}\frac{\partial Q_k}{\partial \theta_k} &= V_k V_m (G_{km} \cos \theta_{km} + B_{km} \sin \theta_{km}) + \\ &\quad + V_k V_l (G_{kl} \cos \theta_{kl} + B_{kl} \sin \theta_{kl}) \\ \frac{\partial Q_k}{\partial \theta_l} &= V_k V_l (-G_{kl} \cos \theta_{kl} - B_{kl} \sin \theta_{kl}) \\ \frac{\partial Q_k}{\partial \theta_m} &= V_k V_m (-G_{km} \cos \theta_{km} - B_{km} \sin \theta_{km}) \\ \frac{\partial Q_k}{\partial V_k} &= -2V_k B_{kk} + V_m (G_{km} \sin \theta_{km} - B_{km} \cos \theta_{km}) + \\ &\quad + V_l (G_{kl} \sin \theta_{kl} - B_{kl} \cos \theta_{kl}) \\ \frac{\partial Q_k}{\partial V_l} &= V_k (G_{kl} \sin \theta_{kl} - B_{kl} \cos \theta_{kl}) \\ \frac{\partial Q_k}{\partial V_m} &= V_k (G_{km} \sin \theta_{km} - B_{km} \cos \theta_{km})\end{aligned}$$

Now, considering that $V_l = 1.020$ p.u., $V_k = 0.970$ p.u., $V_m = 1.000$ p.u., $\theta_{km} = -4.0^\circ$, $\theta_{kl} = 16.0^\circ$, $G_{kk} = 2.347$ p.u., $G_{km} = 0$ p.u., $G_{kl} = -2.347$ p.u., $B_{kk} = -216.2$ p.u., $B_{km} = 154.1$ p.u., and $B_{kl} = 57.0$ p.u., the results are as follows:

$$\begin{aligned}\frac{\partial P_k}{\partial \theta_k} &= 204.0 & \frac{\partial P_k}{\partial \theta_l} &= -54.9 & \frac{\partial P_k}{\partial \theta_m} &= -149.1 \\ \frac{\partial P_k}{\partial V_k} &= 7.54 & \frac{\partial P_k}{\partial V_l} &= 13.1 & \frac{\partial P_k}{\partial V_m} &= -10.4\end{aligned}$$

and,

$$\begin{aligned}\frac{\partial Q_k}{\partial \theta_k} &= 2.90 & \frac{\partial Q_k}{\partial \theta_l} &= -13.3 & \frac{\partial Q_k}{\partial \theta_m} &= 10.4 \\ \frac{\partial Q_k}{\partial V_k} &= 249. & \frac{\partial Q_k}{\partial V_l} &= -53.8 & \frac{\partial Q_k}{\partial V_m} &= -149.1\end{aligned}$$

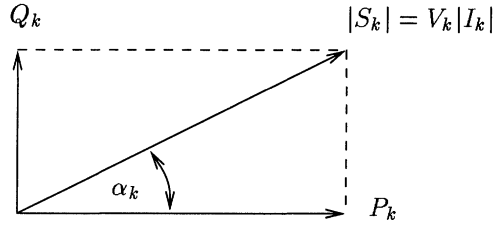


Figure 10.9. Bus current magnitude measurements do not carry information on power factor angle α_k .

10.3.2 Current Magnitude Measurement

Consider the generic bus model in Fig. 10.6. The magnitude of the current injection $|I_k|$ is given by:

$$|I_k| = \frac{(P_k^2 + Q_k^2)^{1/2}}{V_k} \tag{10.4}$$

The corresponding Jacobian matrix elements are

$$\begin{aligned} \frac{\partial |I_k|}{\partial \theta_k} &= \frac{P_{km} \frac{\partial P_{km}}{\partial \theta_k} + Q_{km} \frac{\partial Q_{km}}{\partial \theta_k}}{V_k |I_k|} = \cos \alpha_{km} \frac{\partial P_{km}}{\partial \theta_k} + \sin \alpha_{km} \frac{\partial Q_{km}}{\partial \theta_k} \\ \frac{\partial |I_k|}{\partial \theta_m} &= \frac{P_{km} \frac{\partial P_{km}}{\partial \theta_m} + Q_{km} \frac{\partial Q_{km}}{\partial \theta_m}}{V_k |I_k|} = \cos \alpha_{km} \frac{\partial P_{km}}{\partial \theta_m} + \sin \alpha_{km} \frac{\partial Q_{km}}{\partial \theta_m} \\ \frac{\partial |I_k|}{\partial V_k} &= \frac{P_{km} \frac{\partial P_{km}}{\partial V_k} + Q_{km} \frac{\partial Q_{km}}{\partial V_k}}{V_k |I_k|} = \cos \alpha_{km} \frac{\partial P_{km}}{\partial V_k} + \sin \alpha_{km} \frac{\partial Q_{km}}{\partial V_k} \\ \frac{\partial |I_k|}{\partial V_m} &= \frac{P_{km} \frac{\partial P_{km}}{\partial V_m} + Q_{km} \frac{\partial Q_{km}}{\partial V_m}}{V_k |I_k|} = \cos \alpha_{km} \frac{\partial P_{km}}{\partial V_m} + \sin \alpha_{km} \frac{\partial Q_{km}}{\partial V_m} \end{aligned}$$

where $\frac{\partial P_{km}}{\partial \theta_k}$, $\frac{\partial P_{km}}{\partial \theta_m}$, $\frac{\partial P_{km}}{\partial V_k}$, and $\frac{\partial P_{km}}{\partial V_m}$ are the same as for the power flow measurement contributions given above.

The expressions above can be easily extended to cases in which tap magnitudes (a_{km}) and phase shifts (φ_{km}) of branches incident to Bus k are used as additional state variables (in the same way it was done for branch current magnitude measurements).

Example 10.5

Consider again the 3-bus network shown in Fig. 10.7. In order to compute the contributions to the Jacobian matrix it is necessary to have either an estimate of the bus current magnitude ($|I_k|$) or an estimate of the power factor

θ_l	θ_k	θ_m	V_l	V_k	V_m	
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	
$\dots \frac{\partial I_k }{\partial \theta_l} \dots \frac{\partial I_k }{\partial \theta_k} \dots \frac{\partial I_k }{\partial \theta_m} \dots \frac{\partial I_k }{\partial V_l} \dots \frac{\partial I_k }{\partial V_k} \dots \frac{\partial I_k }{\partial V_m} \dots$						$ I_k ^{meas}$
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	

Figure 10.10. Contributions of bus current magnitude measurement to Jacobian matrix.

angle (α_k). For the sake of illustration, the latter will be assigned the value of zero:

$$\begin{aligned} \frac{\partial |I_k|}{\partial \theta_k} &\simeq \frac{\partial P_k}{\partial \theta_k} = 204.0 & \frac{\partial |I_k|}{\partial \theta_l} &\simeq \frac{\partial P_k}{\partial \theta_l} = -54.9 & \frac{\partial |I_k|}{\partial \theta_m} &\simeq \frac{\partial P_k}{\partial \theta_m} = -149.1 \\ \frac{\partial |I_k|}{\partial V_k} &\simeq \frac{\partial P_k}{\partial V_k} = 7.54 & \frac{\partial |I_k|}{\partial V_l} &\simeq \frac{\partial P_k}{\partial V_l} = 13.1 & \frac{\partial |I_k|}{\partial V_m} &\simeq \frac{\partial P_k}{\partial V_m} = -10.4 \end{aligned}$$

As with branch current magnitude measurements, the reactive power component may be relatively large. When this is the case, the initialization of power factor angles is crucial since multiple solutions may occur. When current measurements are not critical measurements, information from other sources can be used to estimate either the power factor angles or the current magnitudes (depending on the expression used to evaluate the contribution to the Jacobian matrix). The use of critical bus current magnitude measurements requires certain additional information from previous state estimation runs, and/or forecast data, and there is always a risk of producing an incorrect solution.

10.4 HISTORICAL NOTES AND REFERENCES

The basic properties of the Jacobian matrix were reviewed by Schweppe and Handschin [1974], and the impact of current magnitude measurements on $P - \theta$ observability analysis was discussed by Abur and Exposito [1995]. These authors also investigated the role of current magnitude measurement on the existence of multiple solutions (Abur and Exposito [1997]). A general treatment of the measurement Jacobian matrix considering a wide variety of types of measurement can also be found in Alsac, Vempati, Stott and Monticelli [1998]. The books by Wood and Wollenberg [1995] and Grainger and Stevenson [1994]

dedicate a chapter to weighted least squares state estimation with examples showing how the Jacobian matrices are established.

10.5 PROBLEMS

- 1. Consider again the 750 kV in Example 10.1. The series impedance is $0.00072 + j0.0175$ p.u.; the total shunt admittance is 877.5 MVar. Assume now that the voltage magnitudes at the terminal buses are 0.991 p.u. and 0.966 p.u.; and the voltage angle spread is 19° (100 MVA base). (a) Determine the active and reactive power flows and the magnitude of the current flow in the line. Considering that both power flows and the current magnitude are metered correctly (no errors), (c) determine the contributions of P_{km}^{meas} , Q_{km}^{meas} , and $|I_{km}|^{meas}$ to measurement Jacobian matrix. (d) Compare these results with those obtained in Example 10.3, where the power factor angle was close to zero.
- 2. Consider again the three bus network in Example 10.4 (Fig. 10.7). Line $k-l$ has a series impedance of $z_{kl} = 0.00072 + j0.0175$ p.u. and a total shunt admittance of $2b_{kl}^{sh} = 877.5$ MVar (100 MVA base). The transformer $m-k$ presents negligible series resistance, a leakage reactance $x_{mk} = 0.00623$ p.u., and a transform tap $a_{mk} = 0.960$. The shunt susceptance of Bus k is $b_k^{sh} = -300$ MVar (it represents a shunt reactor). The voltage magnitude at the three buses is known to be: $V_l = 1.020$ p.u., $V_k = 0.970$ p.u., and $V_m = 1.000$ p.u. The voltage angles are $\theta_l = 0^\circ$, $\theta_k = 16.0^\circ$, and $\theta_m = 20.0^\circ$. (a) Determine the power injections P_k and Q_k at Node k . (b) Determine the current magnitude $|I_k|$. (c) Determine the contribution to Jacobian matrix of measurement $|I_k|^{meas}$ (Consider that P_k and Q_k are known). (d) Compare these results with those obtained in Example 10.5, where the power factor angle was considered to be zero.

References

- Abur, A. and Exposito, A.G., "Algorithm for determining phase-angle observability in the presence of line-current-magnitude measurements", IEE Proc. Trans. Dist. Pt. D, Vol. 142, pp. 453-458, Sept. 1995.
- Abur, A. and Exposito, A.G., "Detecting multiple solutions in state estimation in the presence of current magnitude measurements", IEEE Trans. Power Syst., Vol. 12, pp. 370-375, Feb. 1997.
- Alsaç, O., Vempati, N., Stott, B., and Monticelli, A., "Generalized state estimation", IEEE Trans. on Power Systems, Vol. 13, No. 3, pp. 1069-1075, Aug. 1998.
- Scheppe, F.C. and Handschin, E.J., "Static state estimation in electric power systems", Proc. IEEE, vol. 62, pp 972-983, July 1974.
- Grainger, J. and Stevenson W., *Power System Analysis*, McGraw-Hill, 1994.
- Wood, A.J. and Wollenberg, B.F., *Power system Operation and Control*, 2nd Ed., John Wiley & Sons, 1995.

11 ESTIMATION BASED ON MULTIPLE SCANS OF MEASUREMENTS

Most of the discussion of the preceding chapters dealt with the time invariant case in which state estimates are extracted from a single scan of measurements. This chapter discusses recursive estimation methods based on a time sequence of “snapshots” of system measurements. Both tracking and dynamic methods are analyzed, and applications to parameter estimation are presented. Parameter estimation, along with topology estimation, is a crucial part of generalized state estimation. In this chapter, the technique used to model switches with unknown status (topology estimation), is extended to deal with branches with unknown parameters (parameter estimation).

11.1 STATE ESTIMATION

11.1.1 Tracking State Estimator

Each scan of measurements is described by

$$\mathbf{z}(t_i) = \mathbf{h}(\mathbf{x}(t_i)) + \mathbf{e}(t_i)$$

for $i = 0, 1, 2, \dots$, where $\mathbf{h}(\cdot)$ is a nonlinear vector function, $\mathbf{z}(t_i)$ is the measurement vector corresponding to time t_i , $\mathbf{x}(t_i)$ is the state vector at t_i , and $\mathbf{e}(t_i)$ is a zero mean, orthogonal (uncorrelated in time) vector with variance given by

$$E\{\mathbf{e}(t_i)\mathbf{e}'(t_j)\} = \begin{cases} \mathbf{R}_z(t_i), & t_i = t_j \\ \mathbf{0}, & t_i \neq t_j \end{cases} \quad (11.1)$$

If the scan interval, $t_i - t_{i-1}$, is assumed to be small, the correction between consecutive estimates, $\mathbf{x}(t_i) - \mathbf{x}(t_{i-1})$, will be small enough that the state estimate update can be obtained in a single iteration, i.e.,

$$\mathbf{G}(t_i) (\mathbf{x}(t_i) - \mathbf{x}(t_{i-1})) = \mathbf{H}(\mathbf{x}(t_{i-1})) \mathbf{R}_z^{-1} (\mathbf{z}(t_i) - \mathbf{h}(\mathbf{x}(t_{i-1})))$$

This gives the normal equation solution for the scan of measurements $\mathbf{z}(t_i)$ using the estimate obtained from the previous scan, $\hat{\mathbf{x}}(t_{i-1})$, as the initial solution and considering convergence is obtained in a single iteration. Of course this process can be extended to allow more than one iteration per scan, either using a fixed number of iterations or until convergence is attained (for example, until the magnitudes of the state corrections become smaller than a given tolerance level).

Note that the tracking state estimator does not make assumptions about the time behavior of the state vector $\mathbf{x}(t_i)$ overtime. Hence, it is not necessary to define a transition matrix relating the states at consecutive scan times. This is not the case with dynamic estimators, which are based on mathematical models for the behavior of the state vector overtime.

11.1.2 Linear Dynamic State Estimator

In this section, the measurement model is assumed to be written in the following linear form:

$$\mathbf{z}(t_i) = \mathbf{H}(t_i) \mathbf{x}(t_i) + \mathbf{e}(t_i) \quad (11.2)$$

Normally, it is not possible to define a good dynamic model for the quasi-static conditions for which state estimation is performed in an energy management system. A simplified dynamic model commonly used in practice is provided in the following difference equation:

$$\mathbf{x}(t_i) = \mathbf{x}(t_{i-1}) + \mathbf{v}(t_i) \quad (11.3)$$

where $\mathbf{v}(t_i)$ has a mean of zero and a variance of

$$E\{\mathbf{v}(t_i)\mathbf{v}'(t_j)\} = \begin{cases} \mathbf{R}_v(t_i), & t_i = t_j \\ \mathbf{0}, & t_i \neq t_j \end{cases} \quad (11.4)$$

i.e., $\mathbf{v}(t_i)$, $i = 0, 1, 2, \dots$ is a discrete time process with orthogonal, zero mean random variables (This type of process is also called a wide sense white process.)

The dynamic estimator extracts state estimate updates from two types of "measurements": measurements obtained at scan t_i , as modeled in Eqs. (11.2)-

(11.1) and the pseudo-measurements given by the previous state estimate modeled in Eqs. (11.3)-(11.4). Hence, an augmented measurement vector $\tilde{\mathbf{z}}(t_i)$ can be defined as follows:

$$\tilde{\mathbf{z}}(t_i) = \begin{pmatrix} \mathbf{z}(t_i) \\ \bar{\mathbf{x}}(t_i) \end{pmatrix} = \begin{pmatrix} \mathbf{z}(t_i) \\ \hat{\mathbf{x}}(t_{i-1}) \end{pmatrix} \tag{11.5}$$

Note that in Eq. (11.5) $\bar{\mathbf{x}}(t_i)$ denotes the predicted state at t_i according to the simplified dynamic model (11.3).

Finally the augmented measurement model can be written as a static measurement model

$$\tilde{\mathbf{z}}(t_i) = \tilde{\mathbf{H}}(t_i) \mathbf{x}(t_i) + \tilde{\mathbf{e}}(t_i)$$

where the augmented error vector is defined as

$$\tilde{\mathbf{e}}(t_i) = \begin{pmatrix} \mathbf{e}(t_i) \\ \boldsymbol{\epsilon}(t_i) \end{pmatrix} = \begin{pmatrix} \mathbf{e}(t_i) \\ \mathbf{x}(t_i) - \bar{\mathbf{x}}(t_i) \end{pmatrix}$$

The prediction error $\boldsymbol{\epsilon}(t_i) = \mathbf{x}(t_i) - \bar{\mathbf{x}}(t_i)$ has the following covariance matrix:

$$\mathbf{R}_{\bar{\mathbf{x}}}(t_i) = \mathbf{R}_{\hat{\mathbf{x}}}(t_{i-1}) + \mathbf{R}_{\mathbf{v}}(t_i)$$

Hence, the covariance matrix of the augmented error vector is

$$\mathbf{R}_{\tilde{\mathbf{z}}}(t_i) = E\{\tilde{\mathbf{e}}(t_i)\tilde{\mathbf{e}}'(t_i)\} = \begin{pmatrix} \mathbf{R}_{\mathbf{z}}(t_i) & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_{\bar{\mathbf{x}}}(t_i) \end{pmatrix}$$

The corresponding augmented Jacobian matrix is

$$\tilde{\mathbf{H}}(t_i) = \begin{pmatrix} \mathbf{H}(t_i) \\ \mathbf{I} \end{pmatrix} \tag{11.6}$$

Now the augmented performance index is defined as follows:

$$J(\mathbf{x}(t_i)) = \frac{1}{2} \left(\mathbf{z}(t_i) - \mathbf{H}(t_i) \mathbf{x}(t_i) \right)' \mathbf{R}_{\mathbf{z}}^{-1} \left(\mathbf{z}(t_i) - \mathbf{H}(t_i) \mathbf{x}(t_i) \right) + \frac{1}{2} \left(\bar{\mathbf{x}}(t_i) - \mathbf{x}(t_i) \right)' \mathbf{R}_{\bar{\mathbf{x}}}^{-1} \left(\bar{\mathbf{x}}(t_i) - \mathbf{x}(t_i) \right)$$

and the weighted least square solution is

$$\hat{\mathbf{x}}(t_i) = \left(\tilde{\mathbf{H}}'(t_i) \mathbf{R}_{\tilde{\mathbf{z}}}^{-1}(t_i) \tilde{\mathbf{H}}(t_i) \right)^{-1} \tilde{\mathbf{H}}'(t_i) \mathbf{R}_{\tilde{\mathbf{z}}}^{-1}(t_i) \tilde{\mathbf{z}}(t_i)$$

In view of Eqs. (11.5)-(11.6), this equation can be successively rewritten as

$$\begin{aligned} \hat{\mathbf{x}}(t_i) &= \left(\begin{pmatrix} \mathbf{H}'(t_i) & \mathbf{I} \end{pmatrix} \begin{pmatrix} \mathbf{R}_z^{-1}(t_i) & 0 \\ 0 & \mathbf{R}_{\bar{\mathbf{x}}}^{-1}(t_i) \end{pmatrix} \begin{pmatrix} \mathbf{H}(t_i) \\ \mathbf{I} \end{pmatrix} \right)^{-1} \times \\ &\quad \begin{pmatrix} \mathbf{H}'(t_i) & \mathbf{I} \end{pmatrix} \begin{pmatrix} \mathbf{R}_z^{-1}(t_i) & 0 \\ 0 & \mathbf{R}_{\bar{\mathbf{x}}}^{-1}(t_i) \end{pmatrix} \begin{pmatrix} \mathbf{z}(t_i) \\ \hat{\mathbf{x}}(t_{i-1}) \end{pmatrix} \\ \hat{\mathbf{x}}(t_i) &= \left(\mathbf{R}_{\bar{\mathbf{x}}}^{-1}(t_i) + \mathbf{H}'(t_i) \mathbf{R}_z^{-1}(t_i) \mathbf{H}(t_i) \right)^{-1} \times \\ &\quad \left(\mathbf{H}'(t_i) \mathbf{R}_z^{-1}(t_i) \mathbf{z}(t_i) + \mathbf{R}_{\bar{\mathbf{x}}}^{-1}(t_i) \hat{\mathbf{x}}(t_{i-1}) \right) \end{aligned} \quad (11.7)$$

Now, in view of the matrix inversion lemma (see Sec. 5.11, Chap. 5), the following is true:

$$\left(\mathbf{R}_{\bar{\mathbf{x}}}^{-1} + \mathbf{H}' \mathbf{R}_z^{-1} \mathbf{H} \right)^{-1} = \mathbf{R}_{\bar{\mathbf{x}}} - \mathbf{R}_{\bar{\mathbf{x}}} \mathbf{H}' \left(\mathbf{R}_z + \mathbf{H} \mathbf{R}_{\bar{\mathbf{x}}} \mathbf{H}' \right)^{-1} \mathbf{H} \mathbf{R}_{\bar{\mathbf{x}}}$$

Note that, for simplicity, in the expression above and in the following expressions it is assumed that all matrices are evaluated at t_i . Introducing this expression into Eq. (11.7) yields the following:

$$\begin{aligned} \hat{\mathbf{x}}(t_i) &= \hat{\mathbf{x}}(t_{i-1}) \\ &\quad + \left(\mathbf{R}_{\bar{\mathbf{x}}} \mathbf{H}' \mathbf{R}_z^{-1} - \mathbf{R}_{\bar{\mathbf{x}}} \mathbf{H}' \left(\mathbf{R}_z + \mathbf{H} \mathbf{R}_{\bar{\mathbf{x}}} \mathbf{H}' \right)^{-1} \mathbf{H} \mathbf{R}_{\bar{\mathbf{x}}} \mathbf{H}' \mathbf{R}_z^{-1} \right) \mathbf{z}(t_i) \\ &\quad - \mathbf{R}_{\bar{\mathbf{x}}} \mathbf{H}' \left(\mathbf{R}_z + \mathbf{H} \mathbf{R}_{\bar{\mathbf{x}}} \mathbf{H}' \right)^{-1} \mathbf{H} \hat{\mathbf{x}}(t_{i-1}) \end{aligned}$$

The coefficient of $\mathbf{z}(t_i)$ in the expression above can be rewritten as follows:

$$\begin{aligned} &\mathbf{R}_{\bar{\mathbf{x}}} \mathbf{H}' \mathbf{R}_z^{-1} - \mathbf{R}_{\bar{\mathbf{x}}} \mathbf{H}' \left(\mathbf{R}_z + \mathbf{H} \mathbf{R}_{\bar{\mathbf{x}}} \mathbf{H}' \right)^{-1} \mathbf{H} \mathbf{R}_{\bar{\mathbf{x}}} \mathbf{H}' \mathbf{R}_z^{-1} \\ &= \mathbf{R}_{\bar{\mathbf{x}}} \mathbf{H}' \left(\mathbf{I} - \left(\mathbf{R}_z + \mathbf{H} \mathbf{R}_{\bar{\mathbf{x}}} \mathbf{H}' \right)^{-1} \mathbf{H} \mathbf{R}_{\bar{\mathbf{x}}} \mathbf{H}' \right) \mathbf{R}_z^{-1} \\ &= \mathbf{R}_{\bar{\mathbf{x}}} \mathbf{H}' \left(\mathbf{R}_z + \mathbf{H} \mathbf{R}_{\bar{\mathbf{x}}} \mathbf{H}' \right)^{-1} \left(\left(\mathbf{R}_z + \mathbf{H} \mathbf{R}_{\bar{\mathbf{x}}} \mathbf{H}' \right) - \mathbf{H} \mathbf{R}_{\bar{\mathbf{x}}} \mathbf{H}' \right) \mathbf{R}_z^{-1} \\ &= \mathbf{R}_{\bar{\mathbf{x}}} \mathbf{H}' \left(\mathbf{R}_z + \mathbf{H} \mathbf{R}_{\bar{\mathbf{x}}} \mathbf{H}' \right)^{-1} \end{aligned}$$

The expression for the new state estimate can thus be expressed as follows:

$$\hat{\mathbf{x}}(t_i) = \hat{\mathbf{x}}(t_{i-1}) + \mathbf{K}(t_i) \left(\mathbf{z}(t_i) - \mathbf{H}(t_i) \hat{\mathbf{x}}(t_{i-1}) \right)$$

where $\mathbf{K}(t_i)$ is the pseudo-inverse matrix, expressed as

$$\mathbf{K}(t_i) = \mathbf{R}_{\bar{\mathbf{x}}}(t_i) \mathbf{H}'(t_i) \left(\mathbf{R}_z(t_i) + \mathbf{H}(t_i) \mathbf{R}_{\bar{\mathbf{x}}}(t_i) \mathbf{H}'(t_i) \right)^{-1}$$

The covariance matrix of the updated state estimate is the inverse of the gain matrix for the augmented system, $\tilde{\mathbf{G}}(t_i) = \tilde{\mathbf{H}}(t_i) \mathbf{R}_z^{-1}(t_i) \tilde{\mathbf{H}}'(t_i)$:

$$\mathbf{R}_{\hat{\mathbf{x}}}(t_i) = \tilde{\mathbf{G}}^{-1}(t_i) = \left(\mathbf{R}_{\bar{\mathbf{x}}}^{-1}(t_i) + \mathbf{H}'(t_i) \mathbf{R}_{\mathbf{z}}^{-1}(t_i) \mathbf{H}(t_i) \right)^{-1}$$

Using the matrix inversion lemma as above, this yields:

$$\mathbf{R}_{\hat{\mathbf{x}}}(t_i) = \left(\mathbf{I} - \mathbf{K}(t_i) \mathbf{H}(t_i) \right) \mathbf{R}_{\bar{\mathbf{x}}}(t_i)$$

Kalman Filter

1. Initialize the state vector $\hat{\mathbf{x}}(t_0)$ and the corresponding covariance matrix $\mathbf{R}_{\hat{\mathbf{x}}}(t_0)$, and set the scan time $t_i = t_1$.
2. Compute the covariance matrix of the predicted state for the t_i th scan of measurements: $\mathbf{R}_{\bar{\mathbf{x}}}(t_i) = \mathbf{R}_{\hat{\mathbf{x}}}(t_{i-1}) + \mathbf{R}_{\mathbf{v}}(t_i)$.
3. Update the pseudo-inverse matrix $\mathbf{K}(t_i)$,

$$\mathbf{K}(t_i) = \mathbf{R}_{\bar{\mathbf{x}}}(t_i) \mathbf{H}'(t_i) \left(\mathbf{R}_{\mathbf{z}}(t_i) + \mathbf{H}(t_i) \mathbf{R}_{\bar{\mathbf{x}}}(t_i) \mathbf{H}'(t_i) \right)^{-1}$$

4. Update the covariance matrix of the state estimate, $\mathbf{R}_{\hat{\mathbf{x}}}(t_i)$,

$$\mathbf{R}_{\hat{\mathbf{x}}}(t_i) = \left(\mathbf{I} - \mathbf{K}(t_i) \mathbf{H}(t_i) \right) \mathbf{R}_{\bar{\mathbf{x}}}(t_i)$$

5. Update the state estimate $\hat{\mathbf{x}}(t_i)$,

$$\hat{\mathbf{x}}(t_i) = \hat{\mathbf{x}}(t_{i-1}) + \mathbf{K}(t_i) \left(\mathbf{z}(t_i) - \mathbf{H}(t_i) \hat{\mathbf{x}}(t_{i-1}) \right)$$

6. Advance the scan time, $t_i \leftarrow t_i + 1$, and return to Step 2.

11.1.3 Non-Linear Dynamic State Estimator

The algorithm above can be extended to the nonlinear model

$$\begin{aligned} \mathbf{z}(t_i) &= \mathbf{h}(\mathbf{x}(t_i)) + \mathbf{e}(t_i) \\ \mathbf{x}(t_i) &= \mathbf{x}(t_{i-1}) + \mathbf{v}(t_i) \end{aligned}$$

where $\mathbf{h}(\cdot)$ is a nonlinear vector function of the state vector $\mathbf{x}(t_i)$. This problem is solved by successive linearizations, to each linear problem being applied the algorithm above. The resulting filter is as follows:

Kalman Bucy Filter

1. Initialization

- (a) Define the number of iterations per snapshot of measurements ν^{max} .
- (b) Initialize the state vector $\hat{\mathbf{x}}^{(\nu)}(t_0)$ and the corresponding covariance matrix $\mathbf{R}_{\mathbf{x}}^{(\nu)}(t_0)$.
- (c) Set the scan time $t_i = t_1$.

2. Start processing the t_i th scan of measurements

- (a) Initialize the state estimate corresponding to scan t_i as $\hat{\mathbf{x}}^{(0)}(t_i) = \hat{\mathbf{x}}^{(\nu)}(t_{i-1})$.
- (b) Initialize the covariance matrix of the state estimate corresponding to scan t_i as the covariance of the predicted state error: $\mathbf{R}_{\mathbf{x}}^{(0)}(t_i) = \mathbf{R}_{\mathbf{x}}^{(\nu)}(t_{i-1}) + \mathbf{R}_{\mathbf{v}}(t_i)$.
- (c) Set the iteration count $\nu = 1$.

3. Perform the ν th iteration at scan t_i ,

- (a) Update pseudo-inverse matrix $\mathbf{K}^{(\nu)}(t_i)$:

$$\mathbf{K}^{(\nu)}(t_i) = \mathbf{R}_{\mathbf{x}}^{(\nu-1)}(t_i) \mathbf{H}'_{(\nu)}(t_i) \left(\mathbf{R}_{\mathbf{z}}(t_i) + \mathbf{H}_{(\nu)}(t_i) \mathbf{R}_{\mathbf{x}}^{(\nu-1)}(t_i) \mathbf{H}'_{(\nu)}(t_i) \right)^{-1}$$

- (b) Update the covariance matrix of the state estimate, $\hat{\mathbf{R}}_{\mathbf{x}}^{(\nu)}(t_i)$:

$$\mathbf{R}_{\mathbf{x}}^{(\nu)}(t_i) = \left(\mathbf{I} - \mathbf{K}^{(\nu)}(t_i) \mathbf{H}_{(\nu)}(t_i) \right) \mathbf{R}_{\mathbf{x}}^{(\nu-1)}(t_i)$$

- (c) Update the state estimate $\hat{\mathbf{x}}^{(\nu)}(t_i)$:

$$\hat{\mathbf{x}}^{(\nu)}(t_i) = \hat{\mathbf{x}}^{(\nu-1)}(t_i) + \mathbf{K}^{(\nu)}(t_i) \left(\mathbf{z}(t_i) - \mathbf{H}_{(\nu)}(t_i) \hat{\mathbf{x}}^{(\nu-1)}(t_i) \right)$$

- 4. If $\nu < \nu^{max}$, update $\nu \leftarrow \nu + 1$ and return to Step 3; otherwise advance time scan, $t_i \leftarrow t_i + 1$, and return to Step 2.

11.2 PARAMETER ESTIMATION

This section discusses three methods for network parameter estimation. The first approach is an approximate version of the Kalman Bucy filter and is based on the augmented state vector that is formed by the usual state variables and the parameters to be estimated. The second method is based on the decomposition of the state/parameter estimation problem into two sub-problems: state estimation and parameter estimation. The third method is a dynamic state/parameter estimator which is also based on the Kalman Bucy filter and which yields estimates of an augmented state vector. In all the three methods, it is assumed that a set of N observations is available (N scan of measurements). It is also assumed that an initial parameter vector with $\mathbf{p}^{(0)}$ is given, along with its respective covariance matrix \mathbf{R}_p . Although most of the material presented in this section is directly applicable to both on-line and off-line parameter estimation, the practical experience with on-line parameter estimation is very limited and the on-line update of network parameters must be considered with caution.

11.2.1 Parameter Estimation by State Augmentation

Each scan of measurements is modeled as follows:

$$\mathbf{z}(t_i) = \mathbf{h}(\mathbf{x}(t_i), \mathbf{p}) + \mathbf{e}(t_i) \quad i = 0, 1, 2, \dots, N$$

where $\mathbf{h}(\cdot)$ is a nonlinear vector function, $\mathbf{z}(t_i)$ is the measurement vector corresponding to time t_i , $\mathbf{x}(t_i)$ is the state vector at t_i , and $\mathbf{e}(t_i)$ is a vector with mean of zero and variance \mathbf{R}_z , and \mathbf{p} is an unknown constant parameter vector. Both the initial estimate of the parameter vector, $\hat{\mathbf{p}}^0$, and its a priori covariance matrix, \mathbf{R}_p , are known.

An augmented performance index is defined as follows:

$$J(\mathbf{x}(t_i), \mathbf{p}) = \frac{1}{2} \left(\mathbf{z}(t_i) - \mathbf{h}(\mathbf{x}(t_i), \mathbf{p}) \right)' \mathbf{R}_z^{-1}(t_i) \left(\mathbf{z}(t_i) - \mathbf{h}(\mathbf{x}(t_i), \mathbf{p}) \right) + \frac{1}{2} \left(\bar{\mathbf{p}}(t_i) - \mathbf{p} \right)' \mathbf{R}_p^{-1}(t_i) \left(\bar{\mathbf{p}}(t_i) - \mathbf{p} \right)$$

where the augmented state vector and the augmented measurement vector are as follows:

$$\tilde{\mathbf{x}}(t_i) = \begin{pmatrix} \mathbf{x}(t_i) \\ \mathbf{p} \end{pmatrix}$$

$$\tilde{\mathbf{z}}(t_i) = \begin{pmatrix} \mathbf{z}(t_i) \\ \bar{\mathbf{p}}(t_i) \end{pmatrix}$$

The corresponding Jacobian matrix can be written is

$$\tilde{\mathbf{H}}(t_i) = \begin{pmatrix} \mathbf{H}_x(t_i) & \mathbf{H}_p(t_i) \\ \mathbf{0} & \mathbf{I} \end{pmatrix}$$

whereas the augmented gain matrix is

$$\tilde{\mathbf{G}}(t_i) = \begin{pmatrix} \mathbf{H}'_{\mathbf{x}}(t_i) \mathbf{R}_{\mathbf{z}}^{-1}(t_i) \mathbf{H}_{\mathbf{x}}(t_i) & \mathbf{H}'_{\mathbf{x}}(t_i) \mathbf{R}_{\mathbf{z}}^{-1}(t_i) \mathbf{H}_{\mathbf{p}}(t_i) \\ \mathbf{H}'_{\mathbf{p}}(t_i) \mathbf{R}_{\mathbf{z}}^{-1}(t_i) \mathbf{H}_{\mathbf{x}}(t_i) & \mathbf{H}'_{\mathbf{p}}(t_i) \mathbf{R}_{\mathbf{z}}^{-1}(t_i) \mathbf{H}_{\mathbf{p}}(t_i) + \mathbf{R}_{\mathbf{p}}^{-1}(t_i) \end{pmatrix}$$

In this case, the normal equation

$$\tilde{\mathbf{G}}(t_i) \Delta \tilde{\mathbf{x}}(t_i) = \tilde{\mathbf{H}}'(t_i) \mathbf{R}_{\mathbf{z}}^{-1}(t_i) \Delta \tilde{\mathbf{z}}(t_i)$$

can then be written as follows:

$$\begin{aligned} & \begin{pmatrix} \mathbf{H}'_{\mathbf{x}}(t_i) \mathbf{R}_{\mathbf{z}}^{-1}(t_i) \mathbf{H}_{\mathbf{x}}(t_i) & \mathbf{H}'_{\mathbf{x}}(t_i) \mathbf{R}_{\mathbf{z}}^{-1}(t_i) \mathbf{H}_{\mathbf{p}}(t_i) \\ \mathbf{H}'_{\mathbf{p}}(t_i) \mathbf{R}_{\mathbf{z}}^{-1}(t_i) \mathbf{H}_{\mathbf{x}}(t_i) & \mathbf{H}'_{\mathbf{p}}(t_i) \mathbf{R}_{\mathbf{z}}^{-1}(t_i) \mathbf{H}_{\mathbf{p}}(t_i) + \mathbf{R}_{\mathbf{p}}^{-1}(t_i) \end{pmatrix} \begin{pmatrix} \Delta \mathbf{x}(t_i) \\ \Delta \mathbf{p}(t_i) \end{pmatrix} \\ & = \begin{pmatrix} \mathbf{H}'_{\mathbf{x}}(t_i) & \mathbf{0} \\ \mathbf{H}'_{\mathbf{p}}(t_i) & \mathbf{I} \end{pmatrix} \begin{pmatrix} \mathbf{R}_{\mathbf{z}}^{-1}(t_i) & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_{\mathbf{p}}^{-1}(t_i) \end{pmatrix} \begin{pmatrix} \Delta \mathbf{z}(t_i) \\ \hat{\mathbf{p}}(t_{i-1}) - \mathbf{p}(t_i) \end{pmatrix} \end{aligned} \quad (11.8)$$

The covariance matrix of the augmented state estimate error, $(\mathbf{x}(t_i) - \hat{\mathbf{x}}(t_i), \mathbf{p}(t_i) - \hat{\mathbf{p}}(t_i))'$, can be decomposed as follows

$$\mathbf{R}_{\tilde{\mathbf{x}}}(t_i) = \begin{pmatrix} \mathbf{R}_{\hat{\mathbf{x}}}(t_i) & \mathbf{R}_{\hat{\mathbf{x}}, \hat{\mathbf{p}}}(t_i) \\ \mathbf{R}_{\hat{\mathbf{p}}, \hat{\mathbf{x}}}(t_i) & \mathbf{R}_{\hat{\mathbf{p}}}(t_i) \end{pmatrix}$$

In view of Eq. (11.8), this can be rewritten as

$$\mathbf{R}_{\tilde{\mathbf{x}}}(t_i) = \begin{pmatrix} \mathbf{H}'_{\mathbf{x}}(t_i) \mathbf{R}_{\mathbf{z}}^{-1}(t_i) \mathbf{H}_{\mathbf{x}}(t_i) & \mathbf{H}'_{\mathbf{x}}(t_i) \mathbf{R}_{\mathbf{z}}^{-1}(t_i) \mathbf{H}_{\mathbf{p}}(t_i) \\ \mathbf{H}'_{\mathbf{p}}(t_i) \mathbf{R}_{\mathbf{z}}^{-1}(t_i) \mathbf{H}_{\mathbf{x}}(t_i) & \mathbf{H}'_{\mathbf{p}}(t_i) \mathbf{R}_{\mathbf{z}}^{-1}(t_i) \mathbf{H}_{\mathbf{p}}(t_i) + \mathbf{R}_{\mathbf{p}}^{-1}(t_i) \end{pmatrix}^{-1}$$

Inverting this blocked matrix (Eq. (5.35), Sec. 5.11, Chap. 5) yields the covariance matrix associated with the errors of the parameter estimate vector:

$$\begin{aligned} \mathbf{R}_{\hat{\mathbf{p}}}(t_i) &= \left(\mathbf{H}'_{\mathbf{p}}(t_i) \mathbf{R}_{\mathbf{z}}^{-1}(t_i) \mathbf{H}_{\mathbf{p}}(t_i) + \mathbf{R}_{\mathbf{p}}^{-1}(t_i) \right. \\ & \left. - \mathbf{H}'_{\mathbf{p}}(t_i) \mathbf{R}_{\mathbf{z}}^{-1}(t_i) \mathbf{H}_{\mathbf{x}}(t_i) (\mathbf{H}'_{\mathbf{x}}(t_i) \mathbf{R}_{\mathbf{z}}^{-1}(t_i) \mathbf{H}_{\mathbf{x}}(t_i))^{-1} \mathbf{H}'_{\mathbf{x}}(t_i) \mathbf{R}_{\mathbf{z}}^{-1}(t_i) \mathbf{H}_{\mathbf{p}}(t_i) \right)^{-1} \end{aligned}$$

Hence:

$$\begin{aligned} \mathbf{R}_{\hat{\mathbf{p}}}^{-1}(t_i) &= \mathbf{H}'_{\mathbf{p}}(t_i) \mathbf{R}_{\mathbf{z}}^{-1}(t_i) \mathbf{H}_{\mathbf{p}}(t_i) + \mathbf{R}_{\mathbf{p}}^{-1}(t_i) \\ & - \mathbf{H}'_{\mathbf{p}}(t_i) \mathbf{R}_{\mathbf{z}}^{-1}(t_i) \mathbf{H}_{\mathbf{x}}(t_i) (\mathbf{H}'_{\mathbf{x}}(t_i) \mathbf{R}_{\mathbf{z}}^{-1}(t_i) \mathbf{H}_{\mathbf{x}}(t_i))^{-1} \mathbf{H}'_{\mathbf{x}}(t_i) \mathbf{R}_{\mathbf{z}}^{-1}(t_i) \mathbf{H}_{\mathbf{p}}(t_i) \end{aligned}$$

Finally, note that the covariance matrix of the parameter estimate to be used for the following scan of measurements is

$$\mathbf{R}_{\mathbf{p}}(t_{i+1}) = \mathbf{R}_{\hat{\mathbf{p}}}(t_i)$$

Remarks: For systems with high measurement redundancies, the following approximation can be made:

$$\mathbf{R}_p^{-1}(t_{i+1}) \simeq \mathbf{R}_p^{-1}(t_i) + \mathbf{H}_p'(t_i)\mathbf{R}_z^{-1}(t_i)\mathbf{H}_p(t_i)$$

This approximation is partly justified by the fact that when the measurement redundancy tends to ∞ , the covariance matrix of the measurement estimate errors, $\mathbf{R}_z^{-1} = \mathbf{H}_x [\mathbf{H}_x' \mathbf{R}_z^{-1} \mathbf{H}_x]^{-1} \mathbf{H}_x'$, tends to a null matrix.

Although the \mathbf{R}_p matrix can always be initialized as a diagonal matrix, the updated matrix given by the formulae above (both the exact and approximate expressions) will turn the covariance matrix into a full matrix for the next iteration. The following approximation can be used in order to keep the covariance matrix of the parameter error estimates as a diagonal matrix throughout the estimation process:

$$\mathbf{R}_p^{-1}(t_{i+1}) \simeq \mathbf{R}_p^{-1}(t_i) + \text{diag}(\mathbf{H}_p'(t_i)\mathbf{R}_z^{-1}(t_i)\mathbf{H}_p(t_i)) + \mathbf{D} \quad (11.9)$$

where matrix \mathbf{D} is a constant matrix used to guarantee that the covariance matrix remains positive definite.

Example 11.1:

Consider the one-bus system given in Fig. 11.1. The capacitor has a known fixed susceptance $b = 1.0$ p.u. The reactive part of the load is modeled by

$$Q_L = d + fV^2$$

with $d = 0.95$ p.u. and $f = 0.50$ p.u. (see Prob. 3 at the end of the chapter). The objective is to correct (estimate) parameters d and f of the reactive load model. Three measurements are performed: the bus voltage magnitude V ; the reactive load Q_L ; and the capacitor reactive power Q_C . Three measurements scans are considered: t_1 , t_2 , and t_3 . The measured values and the corresponding estimates are summarized in the following table:

Scan time	Measurement (p.u.)			Estimate (p.u.)		
	Q_L	Q_C	V	V	d	f
t_1	1.90	0.905	0.950	0.9506	1.2226	0.7465
t_2	2.06	1.008	1.005	1.0071	1.2470	0.7767
t_3	2.08	1.092	1.045	1.0445	1.2438	0.7732

The state of the system is represented by the voltage magnitude V (for simplicity the active power part of the problem is ignored). The state vector is augmented by the addition of parameters d and f from the load model. The measurement model can then be written as follows:

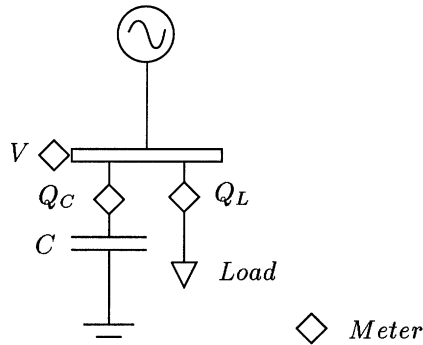


Figure 11.1. One-bus system used in Example 11.1

$$\begin{aligned}
 Q_L^{meas} &= d + f V^2 + e_1 \\
 Q_C^{meas} &= b V^2 + e_2 \\
 V_2^{meas} &= V + e_3
 \end{aligned}$$

Two pseudo-measurements relating parameters d and f to their previous estimates are also included in the model. Hence, the vector function $\mathbf{h}(V, d, f)$ and the corresponding Jacobian matrix, $\mathbf{H}(V, d, f)$, are as follows:

$$\mathbf{h}(V, d, f) = \begin{pmatrix} h_1(V, d, f) \\ h_2(V, d, f) \\ h_3(V, d, f) \\ h_4(V, d, f) \\ h_5(V, d, f) \end{pmatrix} = \begin{pmatrix} d + f V^2 \\ b V^2 \\ V \\ d \\ f \end{pmatrix}$$

$$\mathbf{H}(V, d, f) = \tilde{\mathbf{H}}(t_i) = \begin{matrix} & \mathbf{x} & \mathbf{p} \\ \mathbf{z} & \begin{pmatrix} \mathbf{H}_x(t_i) & \mathbf{H}_p(t_i) \\ \mathbf{0} & \mathbf{I} \end{pmatrix} \end{matrix}$$

$$= \begin{matrix} & V & d & f \\ \begin{matrix} Q_L \\ Q_C \\ V \\ d \\ f \end{matrix} & \begin{pmatrix} \partial h_1 / \partial V & \partial h_1 / \partial d & \partial h_1 / \partial f \\ \partial h_2 / \partial V & \partial h_2 / \partial d & \partial h_2 / \partial f \\ \partial h_3 / \partial V & \partial h_3 / \partial d & \partial h_3 / \partial f \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \end{matrix}$$

$$\begin{matrix} & V & d & f \\ Q_L & 2fV & 1 & V^2 \\ Q_C & 2bV & 0 & 0 \\ = V & 1 & 0 & 0 \\ d & 0 & 1 & 0 \\ f & 0 & 0 & 1 \end{matrix}$$

where, in terms of the partition of variables used above, $\tilde{\mathbf{x}}$ is the augmented state given by (V, d, f) , with \mathbf{x} being the state V and \mathbf{p} the parameters d and f .

The covariance matrix of measurement errors (assumed to be a constant matrix) and the initial covariance matrix of the parameter estimate errors are as follows:

$$\mathbf{R}_z = \begin{pmatrix} 10^{-2} & 0 & 0 \\ 0 & 10^{-2} & 0 \\ 0 & 0 & 10^{-3} \end{pmatrix} \quad \mathbf{R}_p(t_0) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Scan t_1 :

The initial augmented state is $V^{(0)} = 1.0$, $d^{(0)} = 0.95$, and $f^{(0)} = 0.50$ and the parameter error covariance matrix used for this scan is $\mathbf{R}_p(t_1) = \mathbf{R}_p(t_0)$. The correction in the state estimate is given by the solution to Eq. (11.8), which in this case can be written as

$$\begin{pmatrix} \Delta V \\ \Delta d \\ \Delta f \end{pmatrix}^{(1)} = \begin{pmatrix} 1500 & 100 & 100 \\ 100 & 101 & 100 \\ 100 & 100 & 101 \end{pmatrix}^{-1} \begin{pmatrix} 1.0 & 2.0 & 1.0 & 0 & 0 \\ 1.0 & 0 & 0 & 1.0 & 0 \\ 1.0 & 0 & 0 & 0 & 1.0 \end{pmatrix} \\
 \times \begin{pmatrix} 100 & 0 & 0 & 0 & 0 \\ 0 & 100 & 0 & 0 & 0 \\ 0 & 0 & 1000 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0.450 \\ -0.095 \\ -0.050 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} -0.0491 \\ 0.2483 \\ 0.2483 \end{pmatrix}$$

$$\begin{pmatrix} V \\ d \\ f \end{pmatrix}^{(1)} = \begin{pmatrix} V \\ d \\ f \end{pmatrix}^{(0)} + \begin{pmatrix} \Delta V \\ \Delta d \\ \Delta f \end{pmatrix}^{(1)} = \begin{pmatrix} 1.0000 \\ 0.9500 \\ 0.5000 \end{pmatrix} + \begin{pmatrix} -0.0491 \\ 0.2483 \\ 0.2483 \end{pmatrix} = \begin{pmatrix} 0.9509 \\ 1.1983 \\ 0.7483 \end{pmatrix}$$

Similarly, the second iteration yields

$$\begin{pmatrix} V \\ d \\ f \end{pmatrix}^{(2)} = \begin{pmatrix} V \\ d \\ f \end{pmatrix}^{(1)} + \begin{pmatrix} \Delta V \\ \Delta d \\ \Delta f \end{pmatrix}^{(2)} = \begin{pmatrix} 0.9509 \\ 1.1983 \\ 0.7483 \end{pmatrix} + \begin{pmatrix} -0.0003 \\ 0.0243 \\ -0.0018 \end{pmatrix} = \begin{pmatrix} 0.9506 \\ 1.2226 \\ 0.7465 \end{pmatrix}$$

All the state corrections of the third iteration are smaller than the 10^{-2} p.u. tolerance. Hence, the next scan of measurements is undertaken.

Scan t_2 :

The updated covariance matrix of the parameter errors is given by Eq. (11.9)

$$\mathbf{R}_p^{-1}(t_2) = \mathbf{R}_p^{-1}(t_1) + \text{diag}\left(\mathbf{H}_p'(t_1) \mathbf{R}_z^{-1} \mathbf{H}_p(t_1)\right) + \mathbf{D}$$

where

$$\mathbf{R}_p^{-1}(t_1) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \mathbf{H}_p(t_1) = \begin{pmatrix} 1 & 0.90 \\ 0 & 0 \end{pmatrix}$$

Assuming $\mathbf{D} = \mathbf{0}$, this yields

$$\mathbf{R}_p^{-1}(t_1) = \begin{pmatrix} 101 & 0 \\ 0 & 83 \end{pmatrix}$$

After two iterations the new estimates are

$$\begin{pmatrix} V \\ d \\ f \end{pmatrix}^{(2)} = \begin{pmatrix} 1.0071 \\ 1.2470 \\ 0.7767 \end{pmatrix}$$

Scan t_3 :

The updated covariance matrix of the parameter errors is

$$\mathbf{R}_p^{-1}(t_3) = \begin{pmatrix} 201 & 0 \\ 0 & 186 \end{pmatrix}$$

After two iterations the new estimates are

$$\begin{pmatrix} V \\ d \\ f \end{pmatrix}^{(2)} = \begin{pmatrix} 1.0445 \\ 1.2438 \\ 0.7732 \end{pmatrix}$$

Example 11.2:

A linear version of the above algorithm in which a single iteration per scan of measurements is performed, will now be presented. In this case, the normal equation, Eq. (11.8), can be rewritten as

$$\begin{pmatrix} \mathbf{H}_x'(t_i) \mathbf{R}_z^{-1}(t_i) \mathbf{H}_x(t_i) & \mathbf{H}_x'(t_i) \mathbf{R}_z^{-1}(t_i) \mathbf{H}_p(t_i) \\ \mathbf{H}_p'(t_i) \mathbf{R}_z^{-1}(t_i) \mathbf{H}_x(t_i) & \mathbf{H}_p'(t_i) \mathbf{R}_z^{-1}(t_i) \mathbf{H}_p(t_i) + \mathbf{R}_p^{-1}(t_i) \end{pmatrix} \begin{pmatrix} \Delta \mathbf{x}(t_i) \\ \Delta \mathbf{p}(t_i) \end{pmatrix} \\ = \begin{pmatrix} \mathbf{H}_x'(t_i) \\ \mathbf{H}_p'(t_i) \end{pmatrix} \mathbf{R}_z^{-1}(t_i) \Delta \mathbf{z}(t_i)$$

Consider the one-bus system given in Fig. 11.1 with the same set of data as in Example 11.1. The measured values and the corresponding estimates obtained with the augmented state parameter estimator, used in the tracking mode, are summarized in the following table:

Scan time	Measurement (p.u.)			Estimate (p.u.)		
	Q_L	Q_C	V	V	d	f
t_1	1.90	0.905	0.950	0.9509	1.1983	0.7483
t_2	2.06	1.008	1.005	1.0087	1.2325	0.7860
t_3	2.08	1.092	1.045	1.0448	1.2305	0.7834

11.2.2 Alternate State-Parameter Estimation

For this alternate procedure, as with the state augmentation method discussed above, each scan of measurements is modeled as

$$\mathbf{z}(t_i) = \mathbf{h}(\mathbf{x}(t_i), \mathbf{p}) + \mathbf{e}(t_i) \quad i = 0, 1, 2, \dots, N$$

where $\mathbf{z}(t_i)$ is the measurement vector corresponding to time t_i , $\mathbf{x}(t_i)$ is the state vector at t_i , and $\mathbf{e}(t_i)$ is a zero mean vector with variance \mathbf{R}_z , and \mathbf{p} is an unknown constant parameter vector. Both the initial estimate of the parameter vector, $\hat{\mathbf{p}}^0$, and its a-priori covariance matrix, \mathbf{R}_p , are known.

We define the performance index by

$$J(\mathbf{x}(t_i), \mathbf{p}) = \frac{1}{2} \left(\mathbf{z}(t_i) - \mathbf{h}(\mathbf{x}(t_i), \mathbf{p}) \right)' \mathbf{R}_z^{-1} \left(\mathbf{z}(t_i) - \mathbf{h}(\mathbf{x}(t_i), \mathbf{p}) \right)$$

for each scan, considering a fixed value for \mathbf{p} . In this case the normal equation is the same as for a regular non-linear state estimator, i.e.,

$$\mathbf{H}'_{\mathbf{x}}(t_i) \mathbf{R}_z^{-1} \mathbf{H}_{\mathbf{x}}(t_i) \Delta \mathbf{x}(t_i) = \mathbf{H}'_{\mathbf{x}}(t_i) \mathbf{R}_z^{-1} \Delta \mathbf{z}(t_i)$$

This equation is used iteratively to minimize $J(\mathbf{x}(t_i), \mathbf{p})$ for each scan. Then a new parameter estimate is obtained by minimizing the following:

$$J(\mathbf{p}) = \sum_{i=1}^N J(\hat{\mathbf{x}}(t_i), \mathbf{p})$$

The optimal state $\hat{\mathbf{p}}$ is then used to compute a new set of state estimates $\hat{\mathbf{x}}(t_i), i = 1, 2, \dots, N$. The minimization of $J(\mathbf{p})$ can be performed using the Kalman Bucy filter considering now \mathbf{p} as a state vector. The corresponding normal equation which is solved at each iteration for the scan of measurements t_i is

$$\begin{aligned} & \left(\mathbf{H}'_{\mathbf{p}}(t_i) \mathbf{R}_z^{-1}(t_i) \mathbf{H}_{\mathbf{p}}(t_i) + \mathbf{R}_p^{-1}(t_i) \right) \Delta \mathbf{p}(t_i) \\ & = \mathbf{H}'_{\mathbf{p}}(t_i) \mathbf{R}_z^{-1}(t_i) \Delta \mathbf{z}(t_i) + \mathbf{R}_p^{-1}(t_i) \left(\mathbf{p}(t_i) - \hat{\mathbf{p}}(t_{i-1}) \right) \end{aligned} \quad (11.10)$$

where the updating formula for the covariance matrix of the error in parameter estimates is the same as in the previous algorithm (the augmented state approach), i.e.,

$$\mathbf{R}_p^{-1}(t_{i+1}) = \mathbf{H}_p'(t_i)\mathbf{R}_z^{-1}(t_i)\mathbf{H}_p(t_i) + \mathbf{R}_p^{-1}(t_i)$$

The same diagonal matrix approximations presented above can also be used for this alternate procedure.

Example 11.3:

Consider the one-bus system presented in Fig. 11.1 with the same set of data as in Example 11.1. Three scan measurements and the respective estimates obtained with the alternate method are summarized in the following table:

Scan time	Measurement (p.u.)			Estimate (p.u.)		
	Q_L	Q_C	V	V	d	f
t_1	1.90	0.905	0.950	0.9512	1.2221	0.7462
t_2	2.06	1.008	1.005	1.0080	1.2465	0.7764
t_3	2.08	1.092	1.045	1.0439	1.2436	0.7729

The calculations involved to obtain this table will now be explained. In the alternate method, the problem is decomposed into two estimation subproblems: state estimation and parameter estimation. In the state estimation subproblem, the state of the system is represented by the voltage magnitude V ; whereas in the parameter estimation subproblem the states are the load parameters d and f . For each scan of measurements, the two subproblems are alternately solved until convergence is attained (Tolerances of 10^{-3} p.u. have been used for all estimated quantities.)

The state estimator measurement model can be written as follows:

$$\begin{aligned} Q_L^{meas} &= d + f V^2 + e_1 \\ Q_C^{meas} &= b V^2 + e_2 \\ V_2^{meas} &= V + e_3 \end{aligned}$$

Hence, the vector function $\mathbf{h}(V)$ and the corresponding Jacobian matrix, $\mathbf{H}(V)$, are the following:

$$\mathbf{h}(V) = \begin{pmatrix} h_1(V) \\ h_2(V) \\ h_3(V) \end{pmatrix} = \begin{pmatrix} d + f V^2 \\ b V^2 \\ V \end{pmatrix}$$

$$\mathbf{H}(V) = \begin{matrix} & V & & V \\ \begin{matrix} Q_L \\ Q_C \\ V \end{matrix} & \begin{pmatrix} \partial h_1 / \partial V \\ \partial h_2 / \partial V \\ \partial h_3 / \partial V \end{pmatrix} & = & \begin{matrix} Q_L \\ Q_C \\ V \end{matrix} \begin{pmatrix} 2 f V \\ 2 b V \\ 1 \end{pmatrix} \end{matrix}$$

The only measurement that depends on parameters d and f is Q_L^{meas} .

$$Q_L^{meas} = d + f V^2 + e_1$$

Two pseudo-measurements relating parameters d and f to their previous estimates are also added to the measurement model of the parameter estimator. Hence, the vector function $\mathbf{h}(d, f)$ and the corresponding Jacobian matrix, $\mathbf{H}(d, f)$, are written as bellow:

$$\mathbf{h}(d, f) = \begin{pmatrix} h_1(d, f) \\ h_4(d, f) \\ h_5(d, f) \end{pmatrix} = \begin{pmatrix} d + f V^2 \\ d \\ f \end{pmatrix}$$

$$\mathbf{H}(d, f) = \begin{matrix} & d & f \\ \begin{matrix} Q_L \\ d \\ f \end{matrix} & \begin{pmatrix} \partial h_1 / \partial d & \partial h_1 / \partial f \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \end{matrix} = \begin{matrix} & d & f \\ \begin{matrix} Q_L \\ d \\ f \end{matrix} & \begin{pmatrix} 1 & V^2 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \end{matrix}$$

The covariance matrices are the same as in the previous example, i.e.,

$$\mathbf{R}_z = \begin{pmatrix} 10^{-2} & 0 & 0 \\ 0 & 10^{-2} & 0 \\ 0 & 0 & 10^{-3} \end{pmatrix} \quad \mathbf{R}_p(t_0) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Note that in the above formulation the Jacobian and the weighting matrix corresponding to the parameter estimation subproblem can be written as follows:

$$\mathbf{H} = \begin{matrix} & \mathbf{p} \\ \begin{matrix} Q_L \\ \mathbf{p} \end{matrix} & \begin{pmatrix} \mathbf{H}_p \\ \mathbf{I}_2 \end{pmatrix} \end{matrix}, \quad \mathbf{W} = \begin{matrix} & Q_L & \mathbf{p} \\ \begin{matrix} Q_L \\ \mathbf{p} \end{matrix} & \begin{pmatrix} w_1 & 0 \\ 0 & \mathbf{R}_p^{-1} \end{pmatrix} \end{matrix}$$

and the corresponding gain matrix can then be written as in Eq. 11.10, i.e.,

$$\mathbf{G} = \mathbf{H}'_p w_1 \mathbf{H}_p + \mathbf{R}_p^{-1}$$

Here, only the weighting w_1 appears (it replaces the inverse of the measurement covariance matrix, \mathbf{R}_z^{-1}), since, as noted above, the only telemetry measurement that depends on the parameters d and f is Q_L^{meas} .

Scan t_1 :

The initial state is $V^{(0)} = 1.0$ and the parameter error covariance matrix for this scan is $\mathbf{R}_p(t_1) = \mathbf{R}_p(t_0)$. The state estimate update correction is given by the solution to Eq. (11.10), which can be written as follows:

$$\Delta V^{(1)} = (1500)^{-1} \begin{pmatrix} 1.0 & 2.0 & 1.0 \end{pmatrix} \begin{pmatrix} 100 & 0 & 0 \\ 0 & 100 & 0 \\ 0 & 0 & 1000 \end{pmatrix} \begin{pmatrix} 0.450 \\ -0.095 \\ -0.050 \end{pmatrix} = -0.0160$$

$$V^{(1)} = V^{(0)} + \Delta V^{(1)} = 1.0000 - 0.0160 = 0.9840$$

The initial parameters are $d^{(0)} = 0.95$ and $f^{(0)} = 0.50$. The corrections of the parameter estimates are furnished by the solution to Eq. (11.10) that can be written as

$$\begin{aligned} \begin{pmatrix} \Delta d \\ \Delta f \end{pmatrix}^{(1)} &= \begin{pmatrix} 101.0 & 96.83 \\ 96.83 & 94.75 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 1 & 0 \\ 0.9683 & 0 & 1 \end{pmatrix} \\ &\times \begin{pmatrix} 100 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0.4659 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0.2392 \\ 0.2316 \end{pmatrix} \end{aligned}$$

$$\begin{pmatrix} d \\ f \end{pmatrix}^{(1)} = \begin{pmatrix} d \\ f \end{pmatrix}^{(0)} + \begin{pmatrix} \Delta d \\ \Delta f \end{pmatrix}^{(1)} = \begin{pmatrix} 0.9500 \\ 0.5000 \end{pmatrix} + \begin{pmatrix} 0.2392 \\ 0.2316 \end{pmatrix} = \begin{pmatrix} 1.1892 \\ 0.7316 \end{pmatrix}$$

Similarly, the second iteration of the state estimation subproblem yields

$$\begin{aligned} \Delta V^{(2)} &= (1594.6)^{-1} \begin{pmatrix} 1.4398 & 1.968 & 1.0 \end{pmatrix} \\ &\times \begin{pmatrix} 100 & 0 & 0 \\ 0 & 100 & 0 \\ 0 & 0 & 1000 \end{pmatrix} \begin{pmatrix} 0.02393 \\ -0.06326 \\ -0.034 \end{pmatrix} = -0.02891 \end{aligned}$$

$$V^{(2)} = V^{(1)} + \Delta V^{(2)} = 0.9840 - 0.0289 = 0.9551$$

The second iteration of the parameter subproblem is as follows:

$$\begin{aligned} \begin{pmatrix} \Delta d \\ \Delta f \end{pmatrix}^{(2)} &= \begin{pmatrix} 101.0 & 91.22 \\ 96.75 & 84.21 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 1 & 0 \\ 0.9122 & 0 & 1 \end{pmatrix} \\ &\times \begin{pmatrix} 100 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0.04341 \\ -0.2392 \\ -0.2316 \end{pmatrix} = \begin{pmatrix} 0.02891 \\ 0.01296 \end{pmatrix} \end{aligned}$$

$$\begin{pmatrix} d \\ f \end{pmatrix}^{(2)} = \begin{pmatrix} d \\ f \end{pmatrix}^{(1)} + \begin{pmatrix} \Delta d \\ \Delta f \end{pmatrix}^{(2)} = \begin{pmatrix} 1.1892 \\ 0.7316 \end{pmatrix} + \begin{pmatrix} 0.02891 \\ 0.01296 \end{pmatrix} = \begin{pmatrix} 1.2181 \\ 0.7446 \end{pmatrix}$$

A converged solution is attained after the third iteration of both subproblems:

$$V^{(3)} = V^{(2)} + \Delta V^{(3)} = 0.9551 - 0.0039 = 0.9512$$

$$\begin{pmatrix} d \\ f \end{pmatrix}^{(3)} = \begin{pmatrix} d \\ f \end{pmatrix}^{(2)} + \begin{pmatrix} \Delta d \\ \Delta f \end{pmatrix}^{(3)} = \begin{pmatrix} 1.2181 \\ 0.7446 \end{pmatrix} + \begin{pmatrix} 0.00399 \\ 0.00163 \end{pmatrix} = \begin{pmatrix} 1.2221 \\ 0.7462 \end{pmatrix}$$

Scan t_2 :

After two iterations, the state estimate is

$$V^{(2)} = 1.0080$$

The updated covariance matrix of the parameter errors is given by Eq. (11.9):

$$\mathbf{R}_p^{-1}(t_2) = \begin{pmatrix} 101 & 0 \\ 0 & 82.87 \end{pmatrix}$$

After one iteration, the new parameter estimates are

$$\begin{pmatrix} d \\ f \end{pmatrix}^{(1)} = \begin{pmatrix} 1.2465 \\ 0.7764 \end{pmatrix}$$

Scan t_3 :

After one iteration, the state estimate is

$$V^{(1)} = 1.0439$$

The updated covariance matrix of the parameter errors is given by Eq. (11.9):

$$\mathbf{R}_p^{-1}(t_3) = \begin{pmatrix} 201 & 0 \\ 0 & 186.1 \end{pmatrix}$$

After one iteration, the new parameter estimates are

$$\begin{pmatrix} d \\ f \end{pmatrix}^{(1)} = \begin{pmatrix} 1.2436 \\ 0.7729 \end{pmatrix}$$

Remarks: In this example, for each scan of measurements, the state/parameter subproblems were solve alternately, one iteration of each subproblem at a time. It is also possible to solve each subproblem until convergence is attained, and then switch to the other subproblem, and so on, until both subproblems converge. The same is true regarding the dormant technique (see Example 11.5).

11.2.3 *Dynamic State-Parameter Estimation*

The augmented performance index will now be defined:

$$\begin{aligned}
 J(\mathbf{x}(t_i), \mathbf{p}(t_i)) &= \frac{1}{2} \left(\mathbf{z}(t_i) - \mathbf{h}(\mathbf{x}(t_i), \mathbf{p}(t_i)) \right)' \mathbf{R}_z^{-1} \left(\mathbf{z}(t_i) - \mathbf{h}(\mathbf{x}(t_i), \mathbf{p}(t_i)) \right) \\
 &+ \frac{1}{2} \left(\bar{\mathbf{x}}(t_i) - \mathbf{x}(t_i) \right)' \bar{\mathbf{R}}_x^{-1} \left(\bar{\mathbf{x}}(t_i) - \mathbf{x}(t_i) \right) \\
 &+ \frac{1}{2} \left(\bar{\mathbf{p}}(t_i) - \mathbf{p}(t_i) \right)' \bar{\mathbf{R}}_p^{-1} \left(\bar{\mathbf{p}}(t_i) - \mathbf{p}(t_i) \right)
 \end{aligned}$$

The inclusion of the predicted values $\bar{\mathbf{x}}$ as additional pseudo-measurements helps in filtering bad analog data that may arise during parameter estimation; it also improves observability conditions in situations where the meter configuration may change during the process (e.g., due to temporary unavailability of certain measurements). This is especially important in real-time parameter estimation.

The definitions of both augmented state and measurement vectors are similar to those used for the method of parameter estimation by state augmentation described in Subsec. 11.2.1, the only difference being that here the predicted state $\bar{\mathbf{x}}$ is also used as a pseudo-measurement along with the current parameter estimate $\bar{\mathbf{p}}$. The normal equation for the augmented performance index can then be written as follows:

$$\begin{aligned}
 &\begin{pmatrix} \mathbf{H}'_x(t_i)\mathbf{R}_z^{-1}(t_i)\mathbf{H}_x(t_i) + \mathbf{R}_x^{-1}(t_i) & \mathbf{H}'_x(t_i)\mathbf{R}_z^{-1}(t_i)\mathbf{H}_p(t_i) \\ \mathbf{H}'_p(t_i)\mathbf{R}_z^{-1}(t_i)\mathbf{H}_x(t_i) & \mathbf{H}'_p(t_i)\mathbf{R}_z^{-1}(t_i)\mathbf{H}_p(t_i) + \mathbf{R}_p^{-1}(t_i) \end{pmatrix} \begin{pmatrix} \Delta\mathbf{x}(t_i) \\ \Delta\mathbf{p}(t_i) \end{pmatrix} \\
 &= \begin{pmatrix} \mathbf{H}'_x(t_i) & \mathbf{0} \\ \mathbf{H}'_p(t_i) & \mathbf{I} \end{pmatrix} \begin{pmatrix} \mathbf{R}_z^{-1}(t_i) & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_p^{-1}(t_i) \end{pmatrix} \begin{pmatrix} \Delta\mathbf{z}(t_i) \\ \mathbf{p}(t_i) - \hat{\mathbf{p}}(t_{i-1}) \end{pmatrix}
 \end{aligned}$$

The updating of the covariance matrices of the state/parameter estimate errors is a crucial aspect of combined state/parameter estimation. For most practical situations the rate of change of network parameters (if any) is much slower than that of the state variables, and in these cases, during a sequence of scans, the covariance matrix of the state estimate error can be maintained constant, while updating only the covariance matrix of the parameter estimate error (it is assumed that the parameter estimates become more accurate as calculation proceeds). There are certain types of parameters, however, such as certain variable transformer taps, which may present a behavior as volatile as that of the regular state variables, and these parameters are then treated as normal states.

The covariance matrix of the augmented state estimate error, $(\mathbf{x}(t_i) - \hat{\mathbf{x}}(t_i), \mathbf{p}(t_i) - \hat{\mathbf{p}}(t_i))'$, can be decomposed as follows:

$$\hat{\mathbf{R}}(t_i) = \begin{pmatrix} \mathbf{R}_{\hat{\mathbf{x}}}(t_i) & \mathbf{R}_{\hat{\mathbf{x}},\hat{\mathbf{p}}}(t_i) \\ \mathbf{R}_{\hat{\mathbf{p}},\hat{\mathbf{x}}}(t_i) & \mathbf{R}_{\hat{\mathbf{p}}}(t_i) \end{pmatrix}$$

In view of Eq. (11.8), this can be rewritten as

$$\widehat{\mathbf{R}}(t_i) = \begin{pmatrix} \mathbf{H}'_{\mathbf{x}}(t_i)\mathbf{R}_z^{-1}(t_i)\mathbf{H}_{\mathbf{x}}(t_i) + \mathbf{R}_p^{-1}(t_i) & \mathbf{H}'_{\mathbf{x}}(t_i)\mathbf{R}_z^{-1}(t_i)\mathbf{H}_{\mathbf{p}}(t_i) \\ \mathbf{H}'_{\mathbf{p}}(t_i)\mathbf{R}_z^{-1}(t_i)\mathbf{H}_{\mathbf{x}}(t_i) & \mathbf{H}'_{\mathbf{p}}(t_i)\mathbf{R}_z^{-1}(t_i)\mathbf{H}_{\mathbf{p}}(t_i) + \mathbf{R}_p^{-1}(t_i) \end{pmatrix}^{-1}$$

Inverting this blocked matrix yields the covariance matrix associated with the vector of parameter estimate errors as follows:

$$\begin{aligned} \mathbf{R}_p^{-1}(t_{i+1}) &= \mathbf{H}'_{\mathbf{p}}(t_i)\mathbf{R}_z^{-1}(t_i)\mathbf{H}_{\mathbf{p}}(t_i) + \mathbf{R}_p^{-1}(t_i) + \\ &- \mathbf{H}'_{\mathbf{p}}(t_i)\mathbf{R}_z^{-1}(t_i)\mathbf{H}_{\mathbf{x}}(t_i) \left(\mathbf{H}'_{\mathbf{x}}(t_i)\mathbf{R}_z^{-1}(t_i)\mathbf{H}_{\mathbf{x}}(t_i) + \mathbf{R}_p^{-1}(t_i) \right)^{-1} \mathbf{H}'_{\mathbf{x}}(t_i)\mathbf{R}_z^{-1}(t_i)\mathbf{H}_{\mathbf{p}}(t_i) \end{aligned}$$

In actual practice this expression has been used for small subnetworks without the diagonal approximation (although this further approximation is necessary for larger networks).

The gain matrix can then be written as

$$\widehat{\mathbf{R}}(t_i) = \begin{pmatrix} \mathbf{G}_{\mathbf{xx}} & \mathbf{G}_{\mathbf{px}} \\ \mathbf{G}_{\mathbf{xp}} & \mathbf{G}_{\mathbf{pp}} \end{pmatrix}$$

where, given Eq. (11.8),

$$\begin{aligned} \mathbf{G}_{\mathbf{xx}} &= \mathbf{H}'_{\mathbf{x}}(t_i)\mathbf{R}_z^{-1}(t_i)\mathbf{H}_{\mathbf{x}}(t_i) + \mathbf{R}_p^{-1}(t_i) \\ \mathbf{G}_{\mathbf{xp}} &= \mathbf{H}'_{\mathbf{x}}(t_i)\mathbf{R}_z^{-1}(t_i)\mathbf{H}_{\mathbf{p}}(t_i) \\ \mathbf{G}_{\mathbf{px}} &= \mathbf{H}'_{\mathbf{p}}(t_i)\mathbf{R}_z^{-1}(t_i)\mathbf{H}_{\mathbf{x}}(t_i) \\ \mathbf{G}_{\mathbf{pp}} &= \mathbf{H}'_{\mathbf{p}}(t_i)\mathbf{R}_z^{-1}(t_i)\mathbf{H}_{\mathbf{p}}(t_i) + \mathbf{R}_p^{-1}(t_i) \end{aligned} \quad (11.11)$$

11.2.4 Newton Raphson Method – Second Order Derivatives

In Chap. 2 we have shown that the Newton-Raphson method takes into account the second order derivatives that appear in the gain matrix, which can be written as

$$\mathbf{G} = \mathbf{H}' \mathbf{W} \mathbf{H} - \sum_{j=1}^m \sigma_j^{-1} \Delta z_j \frac{\partial^2 h_j(\mathbf{x})}{\partial \mathbf{x}^2} \quad (11.12)$$

Hence, the gain submatrices above (Eq. (11.11)), can be rewritten as follows:

$$\begin{aligned} \mathbf{G}_{\mathbf{xx}} &= \mathbf{H}'_{\mathbf{x}}(t_i)\mathbf{R}_z^{-1}(t_i)\mathbf{H}_{\mathbf{x}}(t_i) + \mathbf{R}_p^{-1}(t_i) - \sum_{j=1}^m \sigma_j^{-1} \Delta z_j \frac{\partial^2 h_j(\mathbf{x})}{\partial \mathbf{x}^2} \\ \mathbf{G}_{\mathbf{xp}} &= \mathbf{H}'_{\mathbf{x}}(t_i)\mathbf{R}_z^{-1}(t_i)\mathbf{H}_{\mathbf{p}}(t_i) - \sum_{j=1}^m \sigma_j^{-1} \Delta z_j \frac{\partial^2 h_j(\mathbf{x})}{\partial \mathbf{x} \partial \mathbf{p}} \end{aligned}$$

$$\mathbf{G}_{\mathbf{p}\mathbf{x}} = \mathbf{H}'_{\mathbf{p}}(t_i)\mathbf{R}_{\mathbf{z}}^{-1}(t_i)\mathbf{H}_{\mathbf{x}}(t_i) - \sum_{j=1}^m \sigma_j^{-1} \Delta z_j \frac{\partial^2 h_j(\mathbf{x})}{\partial \mathbf{p} \partial \mathbf{x}}$$

$$\mathbf{G}_{\mathbf{p}\mathbf{p}} = \mathbf{H}'_{\mathbf{p}}(t_i)\mathbf{R}_{\mathbf{z}}^{-1}(t_i)\mathbf{H}_{\mathbf{p}}(t_i) + \mathbf{R}_{\mathbf{p}}^{-1}(t_i) - \sum_{j=1}^m \sigma_j^{-1} \Delta z_j \frac{\partial^2 h_j(\mathbf{x})}{\partial \mathbf{p}^2}$$

Second order terms can be used in connection with any one of the three parameter estimation methods discussed in this chapter and these will provide the same potential benefits observed with a conventional state estimator (see Sec. 2.6, Chap. 2). More specifically, these second order derivative have been implemented in practice as part of a localized dynamic estimator of the type discussed in Subsec. 11.2.3 (See Historical Notes and References at the end of the chapter).

Example 11.4:

Consider the one-bus system presented in Fig. 11.2. For a given scan the measured values are: $V^{meas} = 0.950$ p.u. and $Q_C^{meas} = 0.920$ p.u. The corresponding variances are $\sigma_V^2 = 10^{-3}$ p.u. and $\sigma_{Q_C}^2 = 10^{-2}$ p.u. The initial estimate of the capacitor susceptance is $b = 1.5$ p.u., and the corresponding variance is $\sigma_b^2 = 1.0$ p.u. The problem is to estimate b . The predicted state is $\bar{V} = 0.948$ p.u. with variance $\sigma_{\bar{V}}^2 = 10^{-2}$ p.u.

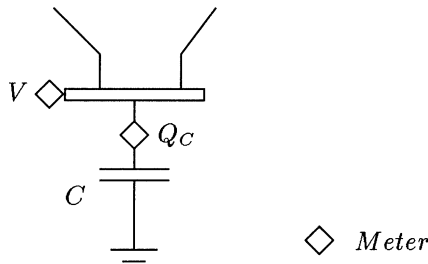


Figure 11.2. One-bus system used in Example 11.4

The measurement model can be written as

$$Q_C^{meas} = b V^2 + e_2$$

$$V^{meas} = V + e_3$$

In addition, the following pseudo-measurements are considered

$$\bar{V} = V + e_3$$

$$\bar{b} = b + e_4$$

Hence, the vector function $\mathbf{h}(V, b)$ and the corresponding Jacobian matrix, $\mathbf{H}(V, b)$, are as follows:

$$\mathbf{h}(V, b) = \begin{pmatrix} h_1(V, b) \\ h_2(V, b) \\ h_3(V, b) \\ h_4(V, b) \end{pmatrix} = \begin{pmatrix} b V^2 \\ V \\ V \\ b \end{pmatrix}$$

$$\mathbf{H}(V, b) = \begin{matrix} & V & b \\ Q_C & \begin{pmatrix} \partial h_1/\partial V & \partial h_1/\partial b \\ \partial h_2/\partial V & \partial h_2/\partial b \\ \partial h_3/\partial V & \partial h_3/\partial b \\ \partial h_4/\partial V & \partial h_4/\partial b \end{pmatrix} & \begin{matrix} V & b \\ V & V^2 \\ V & 0 \\ b & 0 \\ b & 1 \end{matrix} \end{matrix} = \begin{matrix} & V & b \\ Q_C & \begin{pmatrix} 2 b V & V^2 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \end{matrix}$$

The covariance matrices are the same as for the previous example, i.e.,

$$\mathbf{R}_z = \begin{pmatrix} 10^{-2} & 0 & 0 \\ 0 & 10^{-3} & 0 \\ 0 & 0 & 10^{-2} \end{pmatrix} \quad \mathbf{R}_p = (1)$$

And the gain matrix, including the second order derivatives, is

$$\mathbf{G} = \begin{pmatrix} 400 b^2 V^2 + 1100 & 200 b V^3 \\ 200 b V^3 & 100 V^4 + 1 \end{pmatrix} - \sigma_{Q_C}^{-1} (Q_C^{meas} - b V^2) \begin{pmatrix} 2 b & 2 V \\ 2 V & 0 \end{pmatrix}$$

If the initial state is $V^{(0)} = 0.950$, the correction of the state/parameter estimates is given by

$$\begin{pmatrix} \Delta V \\ \Delta b \end{pmatrix}^{(1)} = \begin{pmatrix} 1925.3 & 265.45 \\ 265.45 & 82.451 \end{pmatrix}^{-1} \begin{pmatrix} 2.85 & 1.0 & 1.0 & 0.0 \\ 0.9025 & 0.0 & 0.0 & 1.0 \end{pmatrix} \\ \times \begin{pmatrix} 100 & 0 & 0 & 0 \\ 0 & 1000 & 0 & 0 \\ 0 & 0 & 100 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -0.43375 \\ 0.0 \\ -0.002 \\ 0.0 \end{pmatrix} = \begin{pmatrix} 0.002067 \\ -0.4814 \end{pmatrix}$$

$$V^{(1)} = V^{(0)} + \Delta V^{(1)} = 0.9500 + 0.0021 = 0.9521$$

$$b^{(1)} = b^{(0)} + \Delta b^{(1)} = 1.5000 - 0.4814 = 1.0186$$

Similarly, the second iteration yields

$$V^{(2)} = V^{(1)} + \Delta V^{(2)} = 0.9521 - 0.0032 = 0.9489$$

$$b^{(2)} = b^{(1)} + \Delta b^{(2)} = 1.0186 + 0.0089 = 1.0275$$

The convergence process is summarized in the following table where the results of the Gauss Newton method (ignoring second order derivatives) are

also presented. Since no extreme inconsistencies are found in the data set, both methods show comparable behavior (quadratic convergence).

ν	<i>Gauss-Newton</i>		<i>Newton-Raphson</i>	
	ΔV	Δb	ΔV	Δb
1	-0.0015311	-0.4700040	0.0020675	-0.4814370
2	0.0004166	-0.0023792	-0.0031691	0.0089391
3	0.0000025	-0.0000086	-0.0000103	0.0001057
4	0.0000000	0.0000000	0.0000000	0.0000003

11.2.5 Dormant Parameter Technique

In some cases, parameter errors are such that both the state estimation convergence and the solution can be negatively affected, regardless of the method used for parameter estimation. The dormant parameter technique makes it possible to remove these parameters from the state estimation measurement model. In this case, the parameter estimation is performed separately as it is in the alternate state/parameter estimation method described above.

Consider, for example, the one-bus system in Fig. 11.2. The estimates obtained for this system, if convergence is attained, depend on the weighting factors and on the level of inconsistency in the data set. This can be easily seen if the weights given to the voltage magnitude measurements in the previous example are changed or the error of the initial parameter estimate are modified. Hence, it might be preferable to make the suspect parameter dormant and obtain a more accurate estimate before including the parameter in the estimation model for further refinement, although this might not be necessary.

In the dc power flow model, using the branch power flow as an additional state variable is sufficient to make a branch reactance dormant. When a shunt element such as the one studied in the previous example is involved, consideration of the reactive power flow in the element as a state variable is sufficient. More complex network elements such as a transmission line π equivalent model requires the consideration of additional constraints (pseudo-measurements) in addition to the inclusion of flow state variables.

Consider, for example, the equivalent π model in Fig. 11.3, where, the series branch impedance is to be estimated. In this case, power flows P'_{km} , Q'_{km} , P'_{mk} , and Q'_{mk} are considered to be additional states. The terminal power flows P_{km} , Q_{km} , P_{mk} , and Q_{mk} are then expressed in terms of the new state variables rather than as a function of the terminal bus voltages; as a consequence, the series branch impedance will not appear in the measurement model, and this can be written as:

$$P_{km} = P_{kk} + P'_{km} \quad \text{and} \quad Q_{km} = Q_{kk} + Q'_{km}$$

$$P_{mk} = P_{mm} + P'_{mk} \quad \text{and} \quad Q_{mk} = Q_{mm} + Q'_{mk}$$

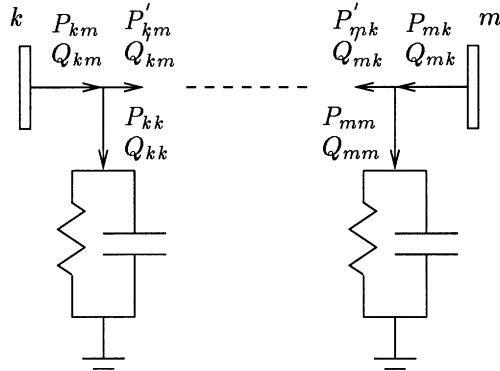


Figure 11.3. Dormant parameter technique applied to the estimation of the series impedance of a π equivalent model.

Notice that in the above equations, the power flows P_{kk} , Q_{kk} , P_{mm} , and Q_{mm} are written in terms of the shunt parameters, as usual. The bus injection measurements at Buses k and m are expressed in terms of the terminal power flows as described above.

These added states are not entirely independent, so it is necessary to include their relationships in the model. In order to do this, the following constraint relating the complex currents in the series branch must be considered:

$$I'_{km} + I'_{mk} = 0$$

Expressing the currents in terms of the corresponding active and reactive power flows, gives the following:

$$\frac{P'_{km} + jQ'_{km}}{V_k e^{j\theta_k}} + \frac{P'_{mk} + jQ'_{mk}}{V_m e^{j\theta_m}} = 0$$

which yields the two following constraints (pseudo-measurements):

$$P'_{km} V_m + \left(P'_{mk} \cos(\theta_{km}) - Q'_{mk} \sin(\theta_{km}) \right) V_k = 0$$

$$Q'_{km} V_m + \left(P'_{mk} \sin(\theta_{km}) + Q'_{mk} \cos(\theta_{km}) \right) V_k = 0$$

Now consider the situation represented in Fig. 11.4, which shows a π equivalent model in which the shunt elements are made dormant. The state variables and measurement model are the same as those in the previous problem, i.e.,

$$P_{km} = P_{kk} + P'_{km} \quad \text{and} \quad Q_{km} = Q_{kk} + Q'_{km}$$

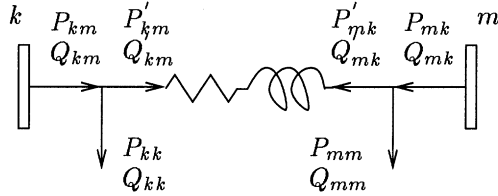


Figure 11.4. Dormant parameter technique applied to the estimation of the shunt admittance of a balanced π equivalent model.

$$P_{mk} = P_{mm} + P'_{mk} \quad \text{and} \quad Q_{mk} = Q_{mm} + Q'_{mk}$$

The constraints linking the state variables are different, however. In this balanced π model, what must be considered for estimating the shunt elements is

$$y_{kk}^{sh} = y_{mm}^{sh}$$

Expressing the shunt admittances in terms of the corresponding active and reactive power flows, gives

$$\frac{P_{kk} + jQ_{kk}}{V_k^2} = \frac{P_{mm} + jQ_{mm}}{V_m^2}$$

and this yields the two following constraints (pseudo-measurements):

$$P_{kk} V_m^2 - P_{mm} V_k^2 = 0$$

$$Q_{kk} V_m^2 - Q_{mm} V_k^2 = 0$$

Example 11.5:

Consider the one-bus system given in Fig. 11.1 with the same set of data as in Example 11.1. Three scan measurements and the respective estimates obtained using the dormant parameter technique are summarized in the following table.

Scan time	Measurement (p.u.)			Estimate (p.u.)			
	Q_L	Q_C	V	V	Q_L	d	f
t_1	1.90	0.905	0.950	0.9507	1.9000	1.2230	0.7465
t_2	2.06	1.008	1.005	1.0047	2.0600	1.2486	0.7782
t_3	2.08	1.092	1.045	1.0447	2.0800	1.2443	0.7735

The calculations performed to obtain this table will now be presented. Just as was done for the alternate method, the problem is decomposed into two estimation problems (state and parameter estimation). The difference now is that in the state estimation subproblem the state has two state variables: V and Q_L (since Q_L is used as an additional state, the load parameters d and f will not form part of the state estimation model). In the parameter estimation subproblem, the states are the load parameters d and f .

The state estimator measurement model can be written as

$$\begin{aligned} Q_L^{meas} &= Q_L + e_1 \\ Q_C^{meas} &= b V^2 + e_2 \\ V^{meas} &= V + e_3 \end{aligned}$$

Hence, the vector function $\mathbf{h}(V, Q_L)$ and the corresponding Jacobian matrix, $\mathbf{H}(V, Q_L)$, are as follows:

$$\mathbf{h}(V, Q_L) = \begin{pmatrix} h_1(V, Q_L) \\ h_2(V, Q_L) \\ h_3(V, Q_L) \end{pmatrix} = \begin{pmatrix} Q_L \\ b V^2 \\ V \end{pmatrix}$$

$$\mathbf{H}(V, Q_L) = \begin{matrix} & V & Q_L & & V & Q_L \\ \begin{matrix} Q_L \\ Q_C \\ V \end{matrix} & \begin{pmatrix} \partial h_1/\partial V & \partial h_1/\partial Q_L \\ \partial h_2/\partial V & \partial h_2/\partial Q_L \\ \partial h_3/\partial V & \partial h_3/\partial Q_L \end{pmatrix} & = & \begin{matrix} Q_L \\ Q_C \\ V \end{matrix} & \begin{pmatrix} 0 & 1 \\ 2 b V & 0 \\ 1 & 0 \end{pmatrix} \end{matrix}$$

Now the corresponding parameter estimation subproblem can be formulated. The state estimator described above yields \hat{Q}_L and \hat{V} . Given the load model, the following is true:

$$\hat{Q}_L = d + f \hat{V}^2$$

Two pseudo-measurements relating parameters d and f to their previous estimates are also added to the measurement model of the parameter estimator. Hence, the vector function $\mathbf{h}(d, f)$ and the corresponding Jacobian matrix, $\mathbf{H}(d, f)$, are written as

$$\mathbf{h}(d, f) = \begin{pmatrix} h_1(d, f) \\ h_4(d, f) \\ h_5(d, f) \end{pmatrix} = \begin{pmatrix} d + f \hat{V}^2 \\ d \\ f \end{pmatrix}$$

$$\mathbf{H}(d, f) = \begin{matrix} & d & f & & d & f \\ \begin{matrix} Q_L \\ d \\ f \end{matrix} & \begin{pmatrix} \partial h_1/\partial d & \partial h_1/\partial f \\ 1 & 0 \\ 0 & 1 \end{pmatrix} & = & \begin{matrix} Q_L \\ d \\ f \end{matrix} & \begin{pmatrix} 1 & \hat{V}^2 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \end{matrix}$$

The covariance matrices are the same as for the previous example, i.e.,

$$\mathbf{R}_z = \begin{pmatrix} 10^{-2} & 0 & 0 \\ 0 & 10^{-2} & 0 \\ 0 & 0 & 10^{-3} \end{pmatrix} \quad \mathbf{R}_p(t_0) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Scan t_1 :

The initial state is $V^{(0)} = 1.0$ and $Q_L^{(0)} = 0.0$. The parameter error covariance matrix for this scan is $\mathbf{R}_p(t_1) = \mathbf{R}_p(t_0)$. The correction of the state estimate is given by

$$\begin{aligned} \begin{pmatrix} \Delta Q_L \\ \Delta V \end{pmatrix}^{(1)} &= \begin{pmatrix} 1400 & 0 \\ 0 & 100 \end{pmatrix}^{-1} \begin{pmatrix} 0.0 & 2.0 & 1.0 \\ 1 & 0 & 0 \end{pmatrix} \\ &\times \begin{pmatrix} 100 & 0 & 0 \\ 0 & 100 & 0 \\ 0 & 0 & 1000 \end{pmatrix} \begin{pmatrix} 1.90 \\ -0.095 \\ -0.050 \end{pmatrix} = \begin{pmatrix} -0.0493 \\ 1.90 \end{pmatrix} \end{aligned}$$

$$V^{(1)} = V^{(0)} + \Delta V^{(1)} = 1.0000 - 0.0493 = 0.9507$$

$$Q_L^{(1)} = Q_L^{(0)} + \Delta Q_L^{(1)} = 0.0000 + 1.9000 = 1.9000$$

Similarly, the second iteration yields

$$V^{(2)} = V^{(1)} + \Delta V^{(2)} = 0.9507 - 0.0004 = 0.9503$$

$$Q_L^{(2)} = Q_L^{(1)} + \Delta Q_L^{(2)} = 1.9000 + 0.0000 = 1.9000$$

The initial parameter values are $d^{(0)} = 0.95$ and $f^{(0)} = 0.50$. The corrections of the parameter estimates are as follows:

$$\begin{aligned} \begin{pmatrix} \Delta d \\ \Delta f \end{pmatrix}^{(1)} &= \begin{pmatrix} 101.0 & 90.32 \\ 90.32 & 82.57 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 1 & 0 \\ 0.9032 & 0 & 1 \end{pmatrix} \\ &\times \begin{pmatrix} 100 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0.4984 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0.2730 \\ 0.2465 \end{pmatrix} \end{aligned}$$

$$\begin{pmatrix} d \\ f \end{pmatrix}^{(1)} = \begin{pmatrix} d \\ f \end{pmatrix}^{(0)} + \begin{pmatrix} \Delta d \\ \Delta f \end{pmatrix}^{(1)} = \begin{pmatrix} 0.9500 \\ 0.5000 \end{pmatrix} + \begin{pmatrix} 0.2730 \\ 0.2465 \end{pmatrix} = \begin{pmatrix} 1.2230 \\ 0.7465 \end{pmatrix}$$

Scan t_2 :

After one iteration, the state estimate is

$$V^{(1)} = 1.00471 \quad Q_L = 2.06$$

The updated covariance matrix of the parameter errors is obtained from Eq. (11.9):

$$\mathbf{R}_p^{-1}(t_2) = \begin{pmatrix} 101 & 0 \\ 0 & 83 \end{pmatrix}$$

After two iterations, the new parameters estimates are

$$\begin{pmatrix} d \\ f \end{pmatrix}^{(2)} = \begin{pmatrix} 1.2486 \\ 0.7782 \end{pmatrix}$$

Scan t_3 :

After one iteration, the state estimate is

$$V^{(1)} = 1.0447$$

The updated covariance matrix of the parameter errors is obtained from Eq. (11.9):

$$\mathbf{R}_p^{-1}(t_3) = \begin{pmatrix} 201 & 0 \\ 0 & 184 \end{pmatrix}$$

After one iteration, the new parameters estimates are

$$\begin{pmatrix} d \\ f \end{pmatrix}^{(1)} = \begin{pmatrix} 1.2443 \\ 0.7735 \end{pmatrix}$$

Remarks: This approach is very similar to the alternate state/parameter estimation method (see Remarks at the end of Example 11.3). The only difference is that the parameter being estimated is made dormant during the state estimation phase, at least until an acceptable estimate is obtained. The main benefit is that the suspect parameter will not affect convergence (see Prob. 1 at the end of the chapter).

11.3 HISTORICAL NOTES AND REFERENCES

The literature available on the topics discussed in this chapter is extensive, although the same cannot be said about successful practical implementations. The tracking state estimator presented in Sec. 11.1 was first suggested in Masiello and Schweppe [1971], and Schweppe [1973] also discussed state augmentation in Chap. 14 of his book. The basic ideas of tracking and dynamic state estimators are discussed in Handschin [1972]. The use of dynamic state estimation to facilitate bad data identification was described by Silva, Coutto and Cantera [1987]. A review of the main developments in dynamic state estimation and hierarchical state estimation was presented by Rousseaux, Van

Cutsem and Dy Liacco [1990]. The parameter estimation approach based on state augmentation was developed by Debs [1974], who with Larson [1970] proposed the simplified dynamic model in Eq. (11.3). A dynamic estimator with both state and parameter prediction and including second order derivatives was suggested by Slutsker, Mokhtaru, and Clements [1998], and the development of the topic in Subsec. 11.2.3 draws on this account. The importance of second order derivatives in the presence of erroneous data was emphasized by Van Amerongen [1995]. All the derivations presented in this chapter were made by extending the normal equation from a single scan to multiple scans. A rigorous theoretical background for discrete time filtering can be found in Bucy [1994]. Although for all the simple examples presented in this chapter explicit inverses of the corresponding gain matrices have been used, this is not appropriate in practical applications, where the direct factorization of the Jacobian matrix by Givens rotations is recommended.

11.4 PROBLEMS

- 1. Solve the problem presented in Example 11.1 considering the exact error covariance matrix update formula

$$\mathbf{R}_p^{-1}(t_{i+1}) = \mathbf{H}_p'(t_i)\mathbf{R}_z^{-1}(t_i)\mathbf{H}_p(t_i) + \mathbf{R}_p^{-1}(t_i) - \mathbf{H}_p'(t_i)\mathbf{R}_z^{-1}(t_i)\mathbf{H}_x(t_i)\left(\mathbf{H}_x'(t_i)\mathbf{R}_z^{-1}(t_i)\mathbf{H}_x(t_i)\right)^{-1}\mathbf{H}_x'(t_i)\mathbf{R}_z^{-1}(t_i)\mathbf{H}_p(t_i)$$

(Note that Example 11.1 was developed using a diagonal (definite positive) approximation for the parameter error covariance matrix $\mathbf{R}_p^{-1}(t_{i+1})$.)

- 2. Study the effect of the weighting factors associated with the voltage magnitude measurement in Example 11.1.
- 3. Extend the state/parameter estimation formulation of Example 11.1 to a more general situation in which the load is modeled as

$$P = a + bV + cV^2$$

$$Q = d + eV + fV^2$$

Determine the conditions required for the identifiability of the load model parameters (a , b , c , d , e , and f).

- 4. Three measurements are made in the dc network shown in Fig. 11.5: current entering Node 1 ($I_1^{meas} = 0.119$ A); the power dissipated by the shunt resistance connected to Node 2 ($P_2^{meas} = 0.1$ W). The state of the system is defined as the nodal voltages V_1 and V_2 . For simplicity the measurement error variance is $\sigma = 1$. Estimate the series resistance R_{12} using the dormant

parameter technique described in Subsec. 11.2.5, for which no initial estimate is known. (Hint: (a) Consider the power flows in series Branch 1 – 2 as state variables; and (b) find a constraint (pseudo-measurement) relating these two state variables).

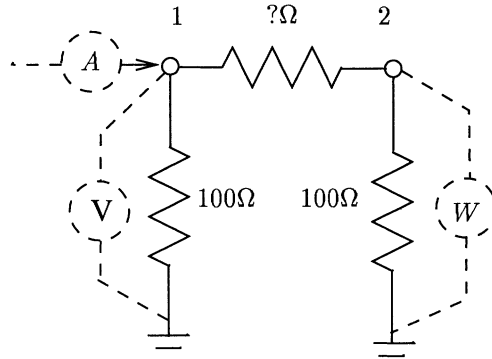


Figure 11.5. Network used for Prob. 4 (the series resistance R_{12} is unknown.)

References

- Alsaç, O., Vempati, N., Stott, B., and Monticelli, A., "Generalized state estimation", *IEEE Trans. on Power Systems*, Vol. 13, No. 3, pp. 1069-1075, Aug. 1998.
- Bucy, R.S., *Lectures on Discrete Time Filtering*, Springer-Verlag, 1994.
- Clements, K.A. and Pinglee, R.J., "Treatment of parameter uncertainty in power system state estimation", Paper C-74-311-7, *IEEE PES Summer Power Meeting*, Anaheim, California, July, 1974.
- Debs, A.S., "Estimation of steady-state power system model parameters", *IEEE Trans. on Power App. and Systems*, Vol. 93, pp. 1260-1268, Sept/Oct. 1974.
- Debs, A.S. and Larson, R.E., "A dynamic estimator for tracking the state of a power system", *IEEE Trans. on Power App. and Systems*, Vol. 89, No. 5, pp. 1670-1678, Sept./Oct. 1970.
- Handschin, E., "Real time data processing using state estimation in electric power systems", in *Real Time Control of Electric Power Systems*, Handschin, E., Editor, pp. 29-57, Elsevier Publishing Company, 1972.
- Masiello, R.D. and Schweppe, F.C., "A tracking static state estimator", *IEEE Trans. on Power App. and Systems*, Vol. 90, pp. 1025-1033, Mar./Apr. 1971.
- Rousseaux, P., Van Cutsem, Th., and Dy Liacco, T. E., "Whither dynamic state estimation", *Int. J. Elec. Power*, Vol. 12, No. 2, pp. 105-116, April 1990.

- Slutsker, I.W., Mokhtaru, S., and Clements, K.A., "Real time recursive parameter estimation in energy management systems", IEEE Trans. on Power Systems, Vol. 11, No. 3, pp. 1393-1399, aug. 1998.
- Schweppe, F.C., *Uncertain Dynamic Systems*, Prentice Hall, 1973.
- Schweppe, F.C., Discussion of Paper: Debs, A.S, "Estimation of steady-state power system model parameters", IEEE Trans. on Power App. and Systems, Vol. 93, pp. 1260-1268, Sept/Oct. 1974.
- Schweppe, F.C. and Handschin, E.J., "Static state estimation in electric power systems", Proc. IEEE, Vol. 62, No. 7, pp. 972-982, July 1974.
- Silva, A.M.L., Coutto, M.B., Cantera, J.M.C., "An efficient dynamic state estimation algorithm including bad data-processing", IEEE Trans. on Power Systems, Vol. 2. No. 4, pp. 1050-1058, Nov. 1987.
- Van Amerongen, R.A.M., "On convergence analysis and convergence enhancement of power-system least-squares state estimators", IEEE Trans. Power Syst., Vol. 10, No. 4, pp. 2038-2044, Nov. 1995.

12 FAST DECOUPLED STATE ESTIMATOR

This chapter discusses the relationships between fast decoupled state estimation methods with the full Gauss Newton method. The role of Jacobian coupling submatrices is analyzed, and the effects of the relevant approximations involved in the derivation of fast decoupled state estimation methods from full Gauss Newton equations are described. Before discussing fast decoupled state estimation, however, fast decoupled power flow methods are introduced and its relationships with the Newton Raphson method are analyzed. (The Gauss Newton method applied to state estimation is discussed in Chap. 2. The Newton Raphson method applied to the power flow problem is introduced in Chap.3, whereas its application to state estimation and parameter estimation are presented in Chaps. 4 and 11, respectively.)

12.1 DECOUPLED SOLUTION OF LINEAR SYSTEM OF EQUATIONS

This section presents the solution of a system of linear equations by partition. This solution approach can be seen as a kind of “decoupling” in which no approximations are made regarding the relative magnitudes of diagonal and off-diagonal submatrices that form part of the coefficient matrix of the linear system being solved.

12.1.1 System of Linear Equations

Consider the following system of linear equations:

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} \begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix} = \begin{pmatrix} \mathbf{a} \\ \mathbf{b} \end{pmatrix} \quad (12.1)$$

where \mathbf{x} and \mathbf{a} are n -vectors and \mathbf{y} and \mathbf{b} are m -vectors. Assuming \mathbf{A} has an inverse, the following transformation can be performed (block Gauss elimination):

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{0} & \mathbf{D} - \mathbf{CA}^{-1}\mathbf{B} \end{pmatrix} \begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix} = \begin{pmatrix} \mathbf{a} \\ \mathbf{b} - \mathbf{CA}^{-1}\mathbf{a} \end{pmatrix}$$

Now assuming that $\mathbf{D} - \mathbf{CA}^{-1}\mathbf{B}$ has an inverse, \mathbf{y} can be determined as follows:

$$\mathbf{y} = (\mathbf{D} - \mathbf{CA}^{-1}\mathbf{B})^{-1}(\mathbf{b} - \mathbf{CA}^{-1}\mathbf{a})$$

Finally, \mathbf{x} results from the first row in Eq. (12.1), i.e.,

$$\mathbf{x} = \mathbf{A}^{-1}(\mathbf{a} - \mathbf{B}\mathbf{y})$$

12.1.2 3-Step Algorithm

The solution procedure above suggests the following 3-step algorithm to obtain \mathbf{x} and \mathbf{y} :

- Step 1: Determine temporary value of \mathbf{x} :

$$\mathbf{x}^{tmp} = \mathbf{A}^{-1}\mathbf{a}$$

- Step 2: Calculate \mathbf{y} :

$$\mathbf{D}_{eq} = \mathbf{D} - \mathbf{CA}^{-1}\mathbf{B}$$

$$\mathbf{b}^{eq} = \mathbf{b} - \mathbf{CA}^{-1}\mathbf{a}$$

$$\mathbf{y} = \mathbf{D}_{eq}^{-1}\mathbf{b}^{eq}$$

- Step 3: Calculate \mathbf{x} :

$$\mathbf{x} = \mathbf{x}^{tmp} - \mathbf{A}^{-1}\mathbf{B}\mathbf{y}$$

Remarks: Rather than solving Eq. (12.1) for \mathbf{x} and \mathbf{y} directly, which would require the inversion (or the triangular factorization) of a coefficient matrix of order $m+n$, the 3-step algorithm above can be used; this requires the inversion (or the factorization) of two lower order matrices (an $n \times n$ matrix, \mathbf{A} , and an

$m \times m$ equivalent matrix, $\mathbf{D}_{eq} = \mathbf{D} - \mathbf{CA}^{-1}\mathbf{B}$). As will be seen, in the case of the fast decoupled power flow, both \mathbf{b}^{eq} and \mathbf{D}_{eq} have interesting interpretations: the mid-iteration updating (the block Gauss-Seidel scheme), and the use of the branch reactance to replace the branch susceptance, respectively.

Example 12.1:

Consider the 2-bus system illustrated in Fig 12.1. In this case, the Jacobian matrix computed for $V = 1$ p.u. and $\theta = 0^\circ$ becomes:

$$\begin{bmatrix} \Delta P_2 \\ \Delta Q_2 \end{bmatrix} = \begin{bmatrix} -b & g \\ -g & -b \end{bmatrix} \begin{bmatrix} \Delta \theta_2 \\ \Delta V_2 \end{bmatrix} \tag{12.2}$$

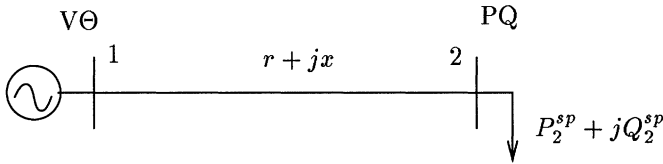


Figure 12.1. Two-bus system.

These equations can be written (BX version) as

$$\begin{bmatrix} \Delta P_2 \\ \Delta Q_2 - \frac{g}{b} \Delta P_2 \end{bmatrix} = \begin{bmatrix} -b & g \\ 0 & -b - \frac{g^2}{b} \end{bmatrix} \begin{bmatrix} \Delta \theta_2 \\ \Delta V_2 \end{bmatrix}$$

and, since

$$\frac{1}{x} = -b - \frac{g^2}{b}$$

and

$$\Delta Q_2(\theta_2 + \Delta \theta_2, V_2) \simeq \Delta Q_2(\theta_2, V_2) - \frac{g}{b} \Delta P_2$$

it results the following:

$$\begin{bmatrix} \Delta P_2 \\ \Delta Q_2 - \frac{g}{b} \Delta P_2 \end{bmatrix} = \begin{bmatrix} -b & g \\ 0 & \frac{1}{x} \end{bmatrix} \begin{bmatrix} \Delta \theta_2 \\ \Delta V_2 \end{bmatrix}$$

This then yields the following 3-step algorithm

- Step 1: Determine temporary value of $\Delta \theta_2$:

$$\Delta \theta_2^{tmp} = -b^{-1} \Delta P_2$$

- Step 2: Calculate ΔV_2 :

$$\Delta V_2 = x (\Delta Q_2 - \frac{g}{b} \Delta P_2) \simeq x \Delta Q_2 (\theta_2 + \Delta \theta_2^{tmp}, V)$$

- Step 3: Calculate correct value of $\Delta \theta_2$:

$$\Delta \theta_2 = \Delta \theta_2^{tmp} + \frac{g}{b} \Delta V_2$$

Note that in Step 2 there are two alternatives for performing mismatch computation; one using $x(\Delta Q_2 - \frac{g}{b} \Delta P_2)$, another using the most recent estimate of the angle θ_2 (mid-iteration angle update). Note also that the same operations performed to zeroize the coupling element $-g$ also transform element $-b$ into $1/x$ and alter the mismatch ΔQ_2 (from ΔQ_2 to $\Delta Q_2 - (g/b)\Delta P_2$). Thus, there is a connection between apparently disconnected factors such as mid-iteration angle updating, zeroizing coupling elements, and replacing $-b$ by $1/x$ in the reactive subproblem, as is normal in fast decoupled power flow methods.

Similar comments could be made regarding the following arrangement of the Newton Raphson iteration equation (XB version):

$$\begin{bmatrix} \Delta P_2 - \frac{g}{b} \Delta Q_2 \\ \Delta Q_2 \end{bmatrix} = \begin{bmatrix} \frac{1}{x} & 0 \\ -g & -b \end{bmatrix} \begin{bmatrix} \Delta \theta_2 \\ \Delta V_2 \end{bmatrix}$$

12.2 FAST DECOUPLED POWER FLOW

In this section the 3-step procedure discussed in the previous section is applied to the calculation of an iteration of the Newton Raphson method.

12.2.1 Newton Raphson Iteration

In the Newton Raphson method the following system of linear equations is solved for $\Delta \Theta$ and $\Delta \mathbf{V}$:

$$\begin{bmatrix} \Delta \mathbf{P} \\ \Delta \mathbf{Q} \end{bmatrix} = \begin{bmatrix} \mathbf{H} & \mathbf{N} \\ \mathbf{M} & \mathbf{L} \end{bmatrix} \begin{bmatrix} \Delta \Theta \\ \Delta \mathbf{V} \end{bmatrix} \quad (12.3)$$

Assuming that \mathbf{H} has an inverse, premultiplying the $\Delta \mathbf{P}$ equations by \mathbf{MH}^{-1} and adding the resulting equations to the $\Delta \mathbf{Q}$ equations, leads to the following system of equations:

$$\begin{bmatrix} \Delta \mathbf{P} \\ \Delta \mathbf{Q} - \mathbf{MH}^{-1} \Delta \mathbf{P} \end{bmatrix} = \begin{bmatrix} \mathbf{H} & \mathbf{N} \\ \mathbf{0} & \mathbf{L} - \mathbf{MH}^{-1} \mathbf{N} \end{bmatrix} \begin{bmatrix} \Delta \Theta \\ \Delta \mathbf{V} \end{bmatrix} \quad (12.4)$$

which is equivalent to the system in Eq.(12.3), since the solution vectors $\Delta \Theta$ and $\Delta \mathbf{V}$ are the same. In the same way, assuming that \mathbf{L} has an inverse, results in the following system:

$$\begin{bmatrix} \Delta \mathbf{P} - \mathbf{N}\mathbf{L}^{-1}\Delta \mathbf{Q} \\ \Delta \mathbf{Q} \end{bmatrix} = \begin{bmatrix} \mathbf{H} - \mathbf{N}\mathbf{L}^{-1}\mathbf{M} & \mathbf{0} \\ \mathbf{M} & \mathbf{L} \end{bmatrix} \begin{bmatrix} \Delta \Theta \\ \Delta \mathbf{V} \end{bmatrix}$$

which gives the same solution as was obtained from the systems in Eqs. (12.3) and (12.4).

As in the previous section, the equivalent matrices \mathbf{H}_{eq} and \mathbf{L}_{eq} are defined as follows:

$$\mathbf{H}_{eq} = \mathbf{H} - \mathbf{N}\mathbf{L}^{-1}\mathbf{M} \quad (12.5)$$

$$\mathbf{L}_{eq} = \mathbf{L} - \mathbf{M}\mathbf{H}^{-1}\mathbf{N} \quad (12.6)$$

12.2.2 Mid-Iteration Updating

Submatrices \mathbf{M} and \mathbf{N} of the Jacobian matrix are the same matrices that appear in the Taylor series expansion of the mismatch functions $\Delta \mathbf{P}(\cdot)$ and $\Delta \mathbf{Q}(\cdot)$, i.e.,

$$\Delta \mathbf{Q}(\mathbf{V}, \Theta + \mathbf{H}^{-1}\Delta \mathbf{P}) \simeq \Delta \mathbf{Q}(\mathbf{V}, \Theta) - \mathbf{M}\mathbf{H}^{-1}\Delta \mathbf{P}(\mathbf{V}, \Theta) \quad (12.7)$$

$$\Delta \mathbf{P}(\mathbf{V} + \mathbf{L}^{-1}\Delta \mathbf{Q}, \Theta) \simeq \Delta \mathbf{P}(\mathbf{V}, \Theta) - \mathbf{N}\mathbf{L}^{-1}\Delta \mathbf{Q}(\mathbf{V}, \Theta) \quad (12.8)$$

If, for example, in Eq. (12.4), $\Delta \Theta^{tmp} = \mathbf{H}^{-1}\Delta \mathbf{P}$ is calculated first, and the updated angle $\Theta + \Delta \Theta^{tmp}$ is used for computing the reactive mismatch $\Delta \mathbf{Q}$ as indicated in Eq. (12.7), then the effect of the coupling matrix \mathbf{M} is implicitly taken into account, as with a normal Newton Raphson iteration.

12.2.3 Ignoring Branch Series Resistances

A critical aspect in decoupled methods is how to calculate the equivalent matrices \mathbf{H}_{eq} and \mathbf{L}_{eq} efficiently. In the following derivations, the rows and columns corresponding to the $\mathbf{V}\Theta$ and $\mathbf{P}\mathbf{V}$ buses are not removed from the Jacobian matrix; instead, large numbers are added to the corresponding main diagonal elements. Thus, the Jacobian matrix in Eq. (12.3) will be a $2n \times 2n$ matrix comprising four $n \times n$ submatrices, where n is the number of buses.

Let Jacobian submatrices \mathbf{H} , \mathbf{N} , \mathbf{M} , and \mathbf{L} be evaluated at flat-start ($V = 1$ p.u. and $\theta = 0^\circ$), and \mathbf{C} be the branch-node incidence matrix (bus incidence matrix), $\mathbf{diag}(b)$ be the diagonal matrix of the branch susceptances (primitive susceptance matrix), and $\mathbf{diag}(g)$ be the diagonal matrix of branch conductances (primitive conductance matrix). Under these conditions, the equivalent submatrix \mathbf{L}_{eq} can be written as follows:

$$\begin{aligned} \mathbf{L}_{eq} &= \mathbf{L} - \mathbf{M}\mathbf{H}^{-1}\mathbf{N} \\ &= -\mathbf{C}'\mathbf{diag}(b)\mathbf{C} - \mathbf{C}'\mathbf{diag}(g)\mathbf{C}[\mathbf{C}'\mathbf{diag}(b)\mathbf{C}]^{-1}\mathbf{C}'\mathbf{diag}(g)\mathbf{C} \end{aligned} \quad (12.9)$$

There are two important specific cases for which Eq. (12.9) can be easily evaluated: radial networks and networks with uniform x/r ratios. These cases will be considered first.

Consider a network formed by branches with uniform x/r ratios. Assuming then that $\alpha = x/r$ results the following:

$$\mathbf{C}' \mathbf{diag}(b)\mathbf{C} = -\alpha \mathbf{C}' \mathbf{diag}(g)\mathbf{C}$$

and Eq. (12.9) can then be written as

$$\begin{aligned} \mathbf{L}_{eq} &= -\mathbf{C}' \mathbf{diag}(b)\mathbf{C} - \alpha^{-1} \mathbf{C}' \mathbf{diag}(g)\mathbf{C} \\ &= -\mathbf{C}' [\mathbf{diag}(b) + \mathbf{diag}(g^2/b)]\mathbf{C} \\ &= \mathbf{C}' \mathbf{diag}(1/x)\mathbf{C} \end{aligned}$$

where $\mathbf{diag}(g^2/b)$ and $\mathbf{diag}(1/x)$ are diagonal matrices with (i, i) th elements (the elements associated with Branch i) given by g_i^2/b_i and $1/x_i$, respectively. Thus, \mathbf{L}^{eq} has the same structure as matrix \mathbf{L} ; the only difference between \mathbf{L} and \mathbf{L}^{eq} is that at flat-start branch susceptance b are replaced by $1/x$. This is equivalent to saying that when forming \mathbf{L}^{eq} , the series resistances are ignored (i.e., the series impedance is considered to be $z = jx$).

For radial networks, even with nonuniform x/r ratios, the branch-node incidence matrix \mathbf{C} is nonsingular, which allows the simplification of Eq. (12.9) as follows:

$$\begin{aligned} \mathbf{L}_{eq} &= -\mathbf{C}' [\mathbf{diag}(b) + \mathbf{diag}(g^2/b)]\mathbf{C} \\ &= \mathbf{C}' \mathbf{diag}(1/x)\mathbf{C} \end{aligned}$$

This is the same result obtained for a network with uniform x/r ratios.

When the network is neither radial nor has uniform r/x ratio, although the matrix \mathbf{L}^{eq} will be generally full, the fill-in elements can be ignored in most practical applications. This means that

$$\mathbf{L}_{eq} = \mathbf{C}' \mathbf{diag}(1/x)\mathbf{C} \quad (12.10)$$

will still be a good approximation for \mathbf{L}_{eq} in these cases.

The derivations above can be extended to the matrix \mathbf{H}_{eq} . In this case, however, an extra complication has to be dealt with: as indicated in Eq. (12.5), the representation of shunts, tap ratios, and PV buses in matrix \mathbf{L} will affect the elements of \mathbf{H}_{eq} . Thus, in order to get an expression similar to the one in Eq. (12.10), the effect of those complicating factors (PV buses, shunts, and taps) will have to be ignored (this in addition to assuming that the system is radial or has constant r/x ratios). This seems to be the only relevant difference

between the standard fast decoupled load flow (XB method) and the general purpose version (BX method).

Example 12.2:

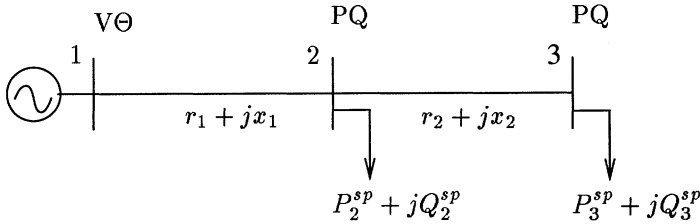


Figure 12.2. Three-bus system.

Consider the three bus system depicted in Fig. 12.2. At flat-start, the correction matrix $\mathbf{MH}^{-1}\mathbf{N}$ is given by the following:

$$\mathbf{MH}^{-1}\mathbf{N} = \begin{pmatrix} -g_1 - g_2 & g_2 \\ g_2 & -g_2 \end{pmatrix} \begin{pmatrix} b_1^{-1} & b_1^{-1} \\ b_1^{-1} & b_1^{-1} + b_2^{-1} \end{pmatrix} \begin{pmatrix} g_1 + g_2 & -g_2 \\ -g_2 & g_2 \end{pmatrix}$$

where $g = r/r^2 + x^2$ and $b = -x/r^2 + x^2$. The equivalent matrix \mathbf{L}_{eq} is thus given by:

$$\mathbf{L}_{eq} = \mathbf{L} - \mathbf{MH}^{-1}\mathbf{N} = \begin{pmatrix} \frac{1}{x_1} + \frac{1}{x_2} & -\frac{1}{x_2} \\ -\frac{1}{x_2} & \frac{1}{x_2} \end{pmatrix}$$

This means that \mathbf{L}_{eq} can be obtained directly from \mathbf{L} by replacing the branch series susceptances, b , with the inverses of the corresponding branch series reactances and changing the signs: i.e., with $-1/x$.

12.2.4 Extended BX Method

Consider the ν th iteration of the Newton Raphson method in which the following approximated Jacobian matrix is used (submatrices \mathbf{H} and \mathbf{L} are computed at flat-start and maintained constant):

$$\begin{pmatrix} \Delta\mathbf{P}^\nu \\ \Delta\mathbf{Q}^\nu \end{pmatrix} = \begin{pmatrix} \mathbf{H}^0 & \mathbf{N}^\nu \\ \mathbf{M}^\nu & \mathbf{L}^0 \end{pmatrix} \begin{pmatrix} \Delta\Theta^\nu \\ \Delta\mathbf{V}^\nu \end{pmatrix} \tag{12.11}$$

Using the procedure described above, gives the equivalent linear system:

$$\begin{pmatrix} \Delta\mathbf{P}^\nu \\ \Delta\mathbf{Q}^\nu - \mathbf{M}^\nu(\mathbf{H}^0)^{-1}\Delta\mathbf{P}^\nu \end{pmatrix} = \begin{pmatrix} \mathbf{H}^0 & \mathbf{N}^\nu \\ \mathbf{0} & \mathbf{L}^0 - \mathbf{M}^\nu(\mathbf{H}^0)^{-1}\mathbf{N}^\nu \end{pmatrix} \begin{pmatrix} \Delta\Theta^\nu \\ \Delta\mathbf{V}^\nu \end{pmatrix}$$

Consider now, the following definitions:

$$\begin{aligned}\mathbf{B}' &= \mathbf{H}^0 \\ \Delta \mathbf{Q}_{eq}^\nu &= \Delta \mathbf{Q}^\nu - \mathbf{M}^\nu (\mathbf{H}^0)^{-1} \Delta \mathbf{P}^\nu \\ \mathbf{L}_{eq}^\nu &= \mathbf{L}^0 - \mathbf{M}^\nu (\mathbf{H}^0)^{-1} \mathbf{N}^\nu\end{aligned}\quad (12.12)$$

(In this book the symbol $'$ is used to denote the transpose of a matrix. The only exception is the matrix \mathbf{B}' above, the B-prime matrix of the fast decoupled power flow, which is also used in the dc power flow in Chap. 4.)

Alternatively, the approximation given by Eq. (12.7) can be used to define $\Delta \mathbf{Q}_{eq}^\nu$, i.e.,

$$\Delta \mathbf{Q}_{eq}^\nu \simeq \Delta \mathbf{Q}(\Theta^\nu + \Delta \Theta^{tmp}, \mathbf{V}^\nu) \quad (12.13)$$

On the other hand, \mathbf{L}_{eq}^ν can be approximated as follows:

$$\mathbf{L}_{eq}^\nu = \mathbf{B}''$$

where \mathbf{B}'' is the \mathbf{L} matrix computed at flat-start and with branch susceptances b replaced by $-1/x$ as discussed above.

This then yields the following 3-step algorithm:

Algorithm for the ν th iteration

- Step 1: Determine temporary value of $\Delta \Theta^\nu$ by solving:

$$\Delta \mathbf{P}^\nu = \mathbf{B}' \Delta \Theta^{tmp}$$

- Step 2: Determine $\Delta \mathbf{V}^\nu$ by solving:

$$\Delta \mathbf{Q}_{eq}^\nu = \mathbf{B}'' \Delta \mathbf{V}^\nu$$

- Step 3: Determine the correct value of $\Delta \Theta^\nu$ by solving:

$$\Delta \mathbf{P}^\nu - \mathbf{N}^\nu \Delta \mathbf{V}^\nu = \mathbf{B}' \Delta \Theta^\nu$$

Example 12.3:

Consider the two systems given in Fig. 12.3. A summary of the computations involved in the first iteration of the Extended BX algorithm for cases (a) and (b) follows:

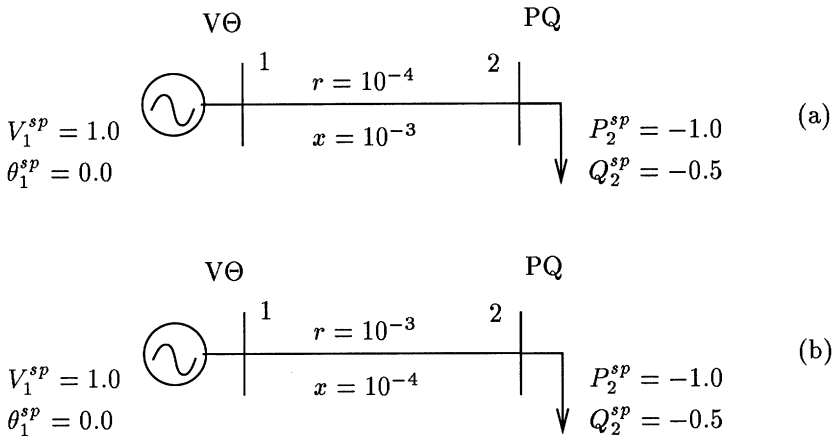


Figure 12.3. Examples used to illustrate the effect of the x/r ratio on the decoupled algorithms (all data in p.u.)

Case (a):

$$b = \frac{-10^{-3}}{10^{-8} + 10^{-6}} = -990 \quad \text{and} \quad g = \frac{-10^{-4}}{10^{-8} + 10^{-6}} = 99$$

Step1: $-1.0 = 990 \Delta\theta_2^{tmp} \Rightarrow \Delta\theta_2^{tmp} = -0.00101$

Step2: $-0.5 - 99 \cdot 1.01 \cdot 10^{-3} = 10^3 \Delta V_2 \Rightarrow \Delta V_2 = -0.00059$

Step3: $-1.0 + 99 \cdot 6 \cdot 10^{-4} = 990 \Delta\theta_2 \Rightarrow \Delta\theta_2 = -0.00095$

Case (b):

$$b = \frac{-10^{-4}}{10^{-6} + 10^{-8}} = -99 \quad \text{and} \quad g = \frac{-10^{-3}}{10^{-6} + 10^{-8}} = 990$$

Step1: $-1.0 = 99 \Delta\theta_2^{tmp} \Rightarrow \Delta\theta_2^{tmp} = -0.0101$

Step2: $-0.5 - 990 \cdot 1.01 \cdot 10^{-2} = 10^4 \Delta V_2 \Rightarrow \Delta V_2 = -0.00105$

Step3: $-1.0 + 990 \cdot 1.05 \cdot 10^{-3} = 99 \Delta\theta_2 \Rightarrow \Delta\theta_2 = -0.000399$

Remarks:

- In the first case (a) there is a relatively high x/r ratio ($x/r = 10.$). In this case, $\Delta\theta_2 \simeq \Delta\theta_2^{tmp}$, which means that the third step may be ignored (the remaining correction will be made in the following iteration).
- In the second case (b) there is a very low x/r ratio ($x/r = 0.1$). In this case $\Delta\theta_2 \ll \Delta\theta_2^{tmp}$, which means that the third step may be essential for the convergence of the algorithm. It is not difficult to imagine a case in which

$\Delta\theta_2^{tmp} > 90^\circ$; in this case, the approximation in Eq. (12.13) would not work, and Eq. (12.12) must be used instead.

12.2.5 Extended XB Method

Consider the ν th iteration of the Newton Raphson method using the same approximated Jacobian matrix as above. By proper transformations, Eq. (12.11) can be written as follows:

$$\begin{pmatrix} \Delta\mathbf{P}^\nu - \mathbf{N}^\nu(\mathbf{L}^0)^{-1}\Delta\mathbf{Q}^\nu \\ \Delta\mathbf{Q}^\nu \end{pmatrix} = \begin{pmatrix} \mathbf{H}^0 - \mathbf{N}^\nu(\mathbf{L}^0)^{-1}\mathbf{M}^\nu & \mathbf{0} \\ \mathbf{M}^\nu & \mathbf{L}^0 \end{pmatrix} \begin{pmatrix} \Delta\Theta^\nu \\ \Delta\mathbf{V}^\nu \end{pmatrix}$$

Consider now, the following definitions:

$$\begin{aligned} \mathbf{B}'' &= \mathbf{L}^0 \\ \Delta\mathbf{P}_{eq}^\nu &= \Delta\mathbf{P}^\nu - \mathbf{N}^\nu(\mathbf{L}^0)^{-1}\Delta\mathbf{Q}^\nu \\ \mathbf{H}_{eq}^\nu &= \mathbf{H}^0 - \mathbf{N}^\nu(\mathbf{L}^0)^{-1}\mathbf{M}^\nu \end{aligned} \quad (12.14)$$

As with the Extended BX method, an alternative would be to use the approximation given in Eq. (12.8) for $\Delta\mathbf{P}_{eq}^\nu$, i.e.,

$$\Delta\mathbf{P}_{eq}^\nu \simeq \Delta\mathbf{P}(\Theta^\nu, \mathbf{V}^\nu + \Delta\mathbf{V}^{tmp}) \quad (12.15)$$

On the other hand, \mathbf{H}_{eq}^ν can be approximated by

$$\mathbf{H}_{eq}^\nu = \mathbf{B}'$$

where \mathbf{B}' is the \mathbf{H} matrix computed at flat-start and branch susceptances b replaced by $-1/x$ as above.

This then yields the following 3-step algorithm

Algorithm for the ν th iteration

- Step 1: Determine temporary value of $\Delta\mathbf{V}^\nu$ by solving:

$$\Delta\mathbf{Q}^\nu = \mathbf{B}'' \Delta\mathbf{V}^{tmp}$$

- Step 2: Determine $\Delta\Theta^\nu$ by solving:

$$\Delta\mathbf{P}_{eq}^\nu = \mathbf{B}' \Delta\Theta^\nu$$

- Step 3: Determine correct value of $\Delta\mathbf{V}^\nu$ by solving:

$$\Delta\mathbf{Q}^\nu - \mathbf{M}^\nu \Delta\Theta^\nu = \mathbf{B}'' \Delta\mathbf{V}^\nu$$

Example 12.4:

Consider again the two systems given in Fig. 12.3. A summary of the computations involved in the first iteration of the Extended BX algorithm for Cases (a) and (b) follows:

Case (a):

$$\text{Step1: } -0.5 = 990 \Delta V_2^{tmp} \Rightarrow \Delta V_2^{tmp} = -0.000505$$

$$\text{Step2: } -1.0 + 99 \cdot 5.05 \cdot 10^{-4} = 1000 \Delta \theta_2 \Rightarrow \Delta \theta_2 = -0.00095$$

$$\text{Step3: } -0.5 - 99 \cdot 9.5 \cdot 10^{-4} = 990 \Delta V_2 \Rightarrow \Delta V_2 = -0.00060$$

Case (b):

$$\text{Step1: } -0.5 = 99 \Delta V_2^{tmp} \Rightarrow \Delta V_2^{tmp} = -0.00505$$

$$\text{Step2: } -1.0 + 990 \cdot 5.05 \cdot 10^{-3} = 10^4 \Delta \theta_2 \Rightarrow \Delta \theta_2 = 0.0004$$

$$\text{Step3: } -0.5 + 990 \cdot 4.0 \cdot 10^{-4} = 99 \Delta V_2 \Rightarrow \Delta V_2 = -0.00105$$

Remarks:

- In the first case (a) where $x/r = 10.$, $\Delta V_2 \simeq \Delta V_2^{tmp}$, which means that the third step may be ignored, as the remaining correction will be made in the following iteration.
- In the second case (b) where $x/r = 0.1$, $\Delta V_2 \ll \Delta V_2^{tmp}$, which means that the third step may be essential for the convergence of the algorithm. Note that when the approximation given by Eq. (12.15) is used, the effect of non-linearity may become important, and as with Extended BX algorithm above, Eq. (12.14) may provide a better solution.

12.2.6 BX Method

As illustrated in Example 12.3 (a), the additional voltage angle correction made in Step 3 of Extended BX algorithm is relatively small for networks with high x/r ratios. In these cases, it is advantageous, and safe, to leave the correction of Step 3 to be made together with the calculation of the estimate $\Delta \theta$ in Step 1 of the following iteration. In such cases the removal of Step 3 of Extended BX algorithm, and making the temporary estimate of Step 1 permanent leads to the following two-step algorithm, known as the BX version of fast decoupled power flow:

Algorithm for the ν th iteration

- Step 1: Determine $\Delta \Theta^\nu$ by solving:

$$\Delta \mathbf{P}^\nu = \mathbf{B}' \Delta \Theta^\nu$$

- Step 2: Determine $\Delta \mathbf{V}^\nu$ by solving:

$$\Delta \mathbf{Q}_{eq}^\nu = \mathbf{B}'' \Delta \mathbf{V}^\nu$$

where, \mathbf{B}' is the \mathbf{H} matrix computed at flat-start, \mathbf{B}'' is the \mathbf{L} matrix computed at flat-start with branch susceptances b replaced by $-1/x$, and $\Delta \mathbf{Q}_{eq}^\nu = \Delta \mathbf{Q}(\Theta^\nu + \Delta \Theta^\nu, \mathbf{V}^\nu)$.

12.2.7 XB Method

As illustrated in Example 12.4 (a), the additional voltage magnitude correction made in Step 3 of Extended XB algorithm is relatively small for networks with high x/r ratios. In such cases the correction made in Step 3 can be made together with the calculation of estimate $\Delta \mathbf{V}$ in Step 1 of the following iteration. The removal of Step 3 of the Extended XB algorithm, and making the temporary estimate in Step 1 permanent leads to the following two-step algorithm:

Algorithm for the ν th iteration

- Step 1: Determine $\Delta \mathbf{V}^\nu$ by solving:

$$\Delta \mathbf{Q}^\nu = \mathbf{B}'' \Delta \mathbf{V}^\nu$$

- Step 2: Determine $\Delta \Theta^\nu$ by solving:

$$\Delta \mathbf{P}_{eq}^\nu = \mathbf{B}' \Delta \Theta^\nu$$

where, \mathbf{B}' is the \mathbf{H} matrix computed at flat start with branch susceptances b replaced by $-1/x$, \mathbf{B}'' is the \mathbf{L} matrix computed at flat-start, and $\Delta \mathbf{P}_{eq}^\nu = \Delta \mathbf{P}(\Theta^\nu, \mathbf{V}^\nu + \Delta \mathbf{V}^\nu)$.

It is known that dc power flow usually provides a good starting point for dc power flows. Based on this, the XB version of the fast decoupled power flow commonly used in practice inverts the order in which correction $\Delta \mathbf{V}^\nu$ and $\Delta \Theta^\nu$ are performed, resulting in the following algorithm:

Inverted algorithm for the ν th iteration

- Step 1: Determine $\Delta \Theta^\nu$ by solving:

$$\Delta \mathbf{P}^\nu = \mathbf{B}' \Delta \Theta^\nu$$

- Step 2: Determine $\Delta \mathbf{V}^\nu$ by solving:

$$\Delta \mathbf{Q}_{eq}^\nu = \mathbf{B}'' \Delta \mathbf{V}^\nu$$

where, \mathbf{B}' and \mathbf{B}'' , are the same as defined above, and $\Delta \mathbf{Q}_{eq}^\nu = \Delta \mathbf{Q}(\Theta^\nu + \Delta \Theta^\nu, \mathbf{V}^\nu)$.

12.3 DECOUPLED SOLUTION OF OVERDETERMINED SYSTEMS

This section extends the decoupled solution approach to overdetermined systems of linear equations.

12.3.1 Overdetermined System of Linear Equations

Consider the following overdetermined system of linear equations:

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} \begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix} = \begin{pmatrix} \mathbf{a} \\ \mathbf{b} \end{pmatrix} \quad (12.16)$$

where \mathbf{x} is an n_x -vector, \mathbf{y} is an n_y -vector, \mathbf{a} is an m_a -vector, and \mathbf{b} is an m_b -vector, with $m > n$, where $n = n_x + n_y$ and $m = m_a + m_b$. It is assumed that the system has an exact solution, i.e., the estimation residuals are zero. Assuming further that \mathbf{A} has full rank, the following transformation can be performed (block Gauss elimination):

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{0} & \mathbf{D} - \mathbf{CA}^I \mathbf{B} \end{pmatrix} \begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix} = \begin{pmatrix} \mathbf{a} \\ \mathbf{b} - \mathbf{CA}^I \mathbf{a} \end{pmatrix}$$

where $\mathbf{A}^I = (\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'$ is the pseudo-inverse of \mathbf{A} .

Now assuming that $\mathbf{D} - \mathbf{CA}^I \mathbf{B}$ has full rank, the least-square estimate \mathbf{y} can be determined as follows:

$$\hat{\mathbf{y}} = (\mathbf{D} - \mathbf{CA}^I \mathbf{B})^I (\mathbf{b} - \mathbf{CA}^I \mathbf{a})$$

where $(\mathbf{D} - \mathbf{CA}^I \mathbf{B})^I$ is the pseudo-inverse of $(\mathbf{D} - \mathbf{CA}^I \mathbf{B})$.

The least-square estimate \mathbf{x} results from the first row in Eq. (12.16):

$$\hat{\mathbf{x}} = \mathbf{A}^I (\mathbf{a} - \mathbf{B}\hat{\mathbf{y}})$$

12.3.2 3-Step Algorithm

The solution procedure above suggest the following 3-step algorithm to obtain estimates $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$:

- Step 1: Determine temporary value of \mathbf{x} :

$$\hat{\mathbf{x}}^{tmp} = \mathbf{A}^I \mathbf{a} = (\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}' \mathbf{a}$$

- Step 2: Estimate $\hat{\mathbf{y}}$:

$$\mathbf{D}_{eq} = \mathbf{D} - \mathbf{CA}^I \mathbf{B}$$

$$\mathbf{b}^{eq} = \mathbf{b} - \mathbf{CA}^I \mathbf{a}$$

$$\hat{\mathbf{y}} = \mathbf{D}_{eq}^I \mathbf{b}^{eq} = (\mathbf{D}'_{eq}\mathbf{D}_{eq})^{-1}\mathbf{D}'_{eq}\mathbf{b}^{eq}$$

- Step 3: Estimate $\hat{\mathbf{x}}$:

$$\hat{\mathbf{x}} = \hat{\mathbf{x}}^{tmp} - \mathbf{A}^T \mathbf{B} \hat{\mathbf{y}}$$

Remarks:

- Hence, the estimates $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ are obtained by the above decoupled procedure by solving three reduced size estimators, one $m_a \times n_x$ estimator in Step 1, one $m_b \times n_y$ estimator in Step 2, and another $m_a \times n_x$ estimator in Step 3. The direct solution of the original system would require solving a $(m_a + m_b) \times (n_x + n_y)$ estimator. As with the fast decoupled power flow, both \mathbf{b}^{eq} and \mathbf{D}_{eq} have interesting interpretations: the mid-iteration updating (the block Gauss-Seidel scheme), and the use of the branch reactance to replace the branch susceptance, respectively.
- The derivations above have been made assuming that the estimation residuals are zero, i.e., it is assumed that the “measurements” vectors \mathbf{a} and \mathbf{b} are perfect. In state estimation, however, measurements are affected by errors. When this happens, the above procedure introduces approximations in the measurement model, and the results yielded by the 3-step procedure are approximations to the ones that would be obtained by solving the system as a whole. Hence, it is said that this is a case of model decoupling, unlike the fast decoupled power flow, in which case only the algorithm is decoupled but the model remains exact. All tests reported in literature, however, agree that these approximations do not affect the quality of the fast decoupled state estimation.

12.4 FAST DECOUPLED STATE ESTIMATOR

In this section the 3-step procedure discussed in the previous section is applied to the calculation of an iteration of the Newton Raphson method.

12.4.1 Gauss Newton Iteration

In the Gauss Newton method the following overdetermined system of linear equations is solved for $\Delta\Theta$ and $\Delta\mathbf{V}$:

$$\begin{bmatrix} \mathbf{H}_{P\Theta} & \mathbf{H}_{PV} \\ \mathbf{H}_{Q\Theta} & \mathbf{H}_{QV} \end{bmatrix} \begin{bmatrix} \Delta\Theta \\ \Delta\mathbf{V} \end{bmatrix} = \begin{bmatrix} \Delta\mathbf{z}_P \\ \Delta\mathbf{z}_Q \end{bmatrix} \quad (12.17)$$

The pseudo-inverse of the Jacobian submatrix $\mathbf{H}_{P\Theta}$ is given by the following:

$$\mathbf{H}_{P\Theta}^I = (\mathbf{H}_{P\Theta}' \mathbf{H}_{P\Theta})^{-1} \mathbf{H}_{P\Theta}'$$

Premultiplying $\Delta\mathbf{z}_P$ in Eq. (12.17) by $-\mathbf{H}_{Q\Theta} \mathbf{H}_{P\Theta}^I$, and adding the result to $\Delta\mathbf{z}_Q$, yields the following transformed system:

$$\begin{bmatrix} \mathbf{H}_{P\Theta} & \mathbf{H}_{PV} \\ \mathbf{0} & \tilde{\mathbf{H}}_{QV} \end{bmatrix} \begin{bmatrix} \Delta\Theta \\ \Delta\mathbf{V} \end{bmatrix} = \begin{bmatrix} \Delta\mathbf{z}_P \\ \Delta\tilde{\mathbf{z}}_Q \end{bmatrix} \quad (12.18)$$

where

$$\begin{aligned} \tilde{\mathbf{H}}_{QV} &= \mathbf{H}_{QV} - \mathbf{H}_{Q\Theta} \mathbf{H}_{P\Theta}^I \mathbf{H}_{PV} \\ \tilde{\mathbf{z}}_Q &= \mathbf{z}_Q - \mathbf{H}_{Q\Theta} \mathbf{H}_{P\Theta}^I \mathbf{z}_P \end{aligned} \quad (12.19)$$

Equations (12.17) and (12.18) will be equivalent provided that the measurements have no errors (perfect measurements). Equation (12.18) can be further transformed into the following decoupled form:

$$\begin{bmatrix} \mathbf{H}_{P\Theta} & \mathbf{0} \\ \mathbf{0} & \tilde{\mathbf{H}}_{QV} \end{bmatrix} \begin{bmatrix} \Delta\Theta \\ \Delta\mathbf{V} \end{bmatrix} = \begin{bmatrix} \Delta\tilde{\mathbf{z}}_P \\ \Delta\tilde{\mathbf{z}}_Q \end{bmatrix} \quad (12.20)$$

where

$$\tilde{\mathbf{z}}_P = \mathbf{z}_P - \mathbf{H}_{PV} \tilde{\mathbf{H}}_{QV}^I \tilde{\mathbf{z}}_Q$$

and $\tilde{\mathbf{H}}_{QV}^I$ is the pseudo-inverse of $\tilde{\mathbf{H}}_{QV}$ given by:

$$\tilde{\mathbf{H}}_{QV}^I = (\tilde{\mathbf{H}}_{QV}' \tilde{\mathbf{H}}_{QV})^{-1} \tilde{\mathbf{H}}_{QV}'$$

Similarly, the following dual decoupled version can be derived:

$$\begin{bmatrix} \tilde{\mathbf{H}}_{P\Theta} & \mathbf{0} \\ \mathbf{0} & \mathbf{H}_{QV} \end{bmatrix} \begin{bmatrix} \Delta\Theta \\ \Delta\mathbf{V} \end{bmatrix} = \begin{bmatrix} \Delta\tilde{\mathbf{z}}_P \\ \Delta\tilde{\mathbf{z}}_Q \end{bmatrix} \quad (12.21)$$

where

$$\begin{aligned} \tilde{\mathbf{H}}_{P\Theta} &= \mathbf{H}_{P\Theta} - \mathbf{H}_{PV} \mathbf{H}_{QV}^I \mathbf{H}_{Q\Theta} \\ \tilde{\mathbf{z}}_P &= \mathbf{z}_P - \mathbf{H}_{PV} \mathbf{H}_{QV}^I \mathbf{z}_Q \\ \tilde{\mathbf{z}}_Q &= \mathbf{z}_Q - \mathbf{H}_{Q\Theta} \tilde{\mathbf{H}}_{P\Theta}^I \tilde{\mathbf{z}}_P \end{aligned}$$

where \mathbf{H}_{QV}^I and $\tilde{\mathbf{H}}_{P\Theta}^I$ are the pseudo-inverses of \mathbf{H}_{QV} and $\tilde{\mathbf{H}}_{P\Theta}$ respectively.

12.4.2 Mid-Iteration Updating

The Jacobian submatrix $\mathbf{H}_{Q\Theta}$ above is the same matrix that appears in the Taylor series expansion of the function $\mathbf{h}_Q(\cdot)$, and by that the reactive mismatch $\tilde{\mathbf{z}}_Q$ can be rewritten as follows:

$$\begin{aligned} \Delta\tilde{\mathbf{z}}_Q &= \Delta\mathbf{z}_Q(\mathbf{V}, \Theta) - \mathbf{H}_{Q\Theta} \Delta\Theta^{tmp} \\ &= \Delta\mathbf{z}_Q(\mathbf{V}, \Theta + \Delta\Theta^{tmp}) \end{aligned}$$

This corresponds to the mid-iteration voltage angle updating normally used in both fast decoupled power flows and state estimators.

Similarly, for the dual algorithm, the following relation holds true:

$$\begin{aligned} \Delta \tilde{\mathbf{z}}_{\mathbf{P}} &= \Delta \mathbf{z}_{\mathbf{P}}(\mathbf{V}, \Theta) - \mathbf{H}_{\mathbf{P}\mathbf{V}} \Delta \mathbf{V}^{tmp} \\ &= \Delta \mathbf{z}_{\mathbf{P}}(\mathbf{V} + \Delta \mathbf{V}^{tmp}, \Theta) \end{aligned}$$

12.4.3 Ignoring Branch Series Resistances

The equivalent matrix $\tilde{\mathbf{H}}_{\mathbf{Q}\mathbf{V}}$ can be approximated by a matrix with the same structure as the original matrix $\mathbf{H}_{\mathbf{Q}\mathbf{V}}$ in which branch susceptances b_{km} are replaced by $-1/x_{km}$, similarly to the fast decoupled power flow discussed above. Assuming all measurements are performed in pairs, for both (a) networks with uniform r/x ratios and (b) minimally observable systems with branch flow measurements only, it can be shown that the replacement of b by $-1/x$ does not involve approximations. As for the other cases, test results show that the approximations are of the same order as those obtained for the fast decoupled power flow.

Equation (12.17) can be rewritten in the following augmented form:

$$\begin{bmatrix} \mathbf{H}_{\Theta\Theta} & \mathbf{0} \\ \mathbf{H}_{\mathbf{P}\Theta} & \mathbf{H}_{\mathbf{P}\mathbf{V}} \\ \mathbf{H}_{\mathbf{Q}\Theta} & \mathbf{H}_{\mathbf{Q}\mathbf{V}} \\ \mathbf{0} & \mathbf{H}_{\mathbf{V}\mathbf{V}} \end{bmatrix} \begin{bmatrix} \Delta \Theta \\ \Delta \mathbf{V} \end{bmatrix} = \begin{bmatrix} \Delta \mathbf{z}_{\Theta} \\ \Delta \mathbf{z}_{\mathbf{P}} \\ \Delta \mathbf{z}_{\mathbf{Q}} \\ \Delta \mathbf{z}_{\mathbf{V}} \end{bmatrix} \quad (12.22)$$

where the active measurement vector $\mathbf{z}_{\mathbf{P}}$ has been partitioned into as vector of power flow and power injection measurements and another vector of voltage angle measurements, and similarly for the reactive measurement vector $\mathbf{z}_{\mathbf{Q}}$.

Firstly, consider cases in which the network has an uniform r/x ratio, denoted by $\alpha = r/x$. At flat-start, Eq. (12.22) can then be written as follows:

$$\begin{bmatrix} \mathbf{H}_{\Theta\Theta}^0 & \mathbf{0} \\ \mathbf{H}_{\mathbf{P}\Theta}^0 & \alpha \mathbf{H}_{\mathbf{P}\Theta}^0 \\ -\alpha \mathbf{H}_{\mathbf{P}\Theta}^0 & \mathbf{H}_{\mathbf{Q}\mathbf{V}}^{sh} + \mathbf{H}_{\mathbf{P}\Theta}^0 \\ \mathbf{0} & \mathbf{H}_{\mathbf{V}\mathbf{V}}^0 \end{bmatrix} \begin{bmatrix} \Delta \Theta \\ \Delta \mathbf{V} \end{bmatrix} = \begin{bmatrix} \Delta \mathbf{z}_{\Theta} \\ \Delta \mathbf{z}_{\mathbf{P}} \\ \Delta \mathbf{z}_{\mathbf{Q}} \\ \Delta \mathbf{z}_{\mathbf{V}} \end{bmatrix}$$

where $\mathbf{H}_{\mathbf{Q}\mathbf{V}}^{sh}$ gives the contributions of the shunt elements, and $\mathbf{H}_{\mathbf{P}\Theta}^0$ depends only on branch susceptances. The transformed matrix $\tilde{\mathbf{H}}_{\mathbf{Q}\mathbf{V}}^{sh}$ can then be written as follows:

$$\begin{aligned} \tilde{\mathbf{H}}_{\mathbf{Q}\mathbf{V}} &= \mathbf{H}_{\mathbf{Q}\mathbf{V}}^{sh} + \mathbf{H}_{\mathbf{P}\Theta}^0 + \alpha \mathbf{H}_{\mathbf{P}\Theta}^0 \mathbf{H}_{\mathbf{P}\Theta}^I \alpha \mathbf{H}_{\mathbf{P}\Theta}^0 \\ &= \mathbf{H}_{\mathbf{Q}\mathbf{V}}^{sh} + \mathbf{H}_{\mathbf{P}\Theta}^0 (1 + \alpha^2) \end{aligned} \quad (12.23)$$

Since $x_{km}^{-1} = -b_{km}(1 + \alpha^2)$, this result implies that the transformed matrix $\tilde{\mathbf{H}}_{\mathbf{Q}\mathbf{V}}$ can be obtained by substituting $-1/x_{km}$ for b_{km} in the original matrix $\mathbf{H}_{\mathbf{Q}\mathbf{V}}$.

Consider now cases in which the network is minimally observable by a set of branch flow measurements (forming a spanning tree). Under this condition, at flat-start, Eq. (12.22) can then be written as follows:

$$\begin{bmatrix} \mathbf{H}_{\Theta\Theta}^0 & \mathbf{0} \\ \text{diag}(b) \mathbf{C} & -\text{diag}(g) \mathbf{C} \\ \text{diag}(g) \mathbf{C} & \mathbf{H}_{\mathbf{QV}}^{sh} + \text{diag}(b) \mathbf{C} \\ \mathbf{0} & \mathbf{H}_{\mathbf{VV}}^0 \end{bmatrix} \begin{bmatrix} \Delta\Theta \\ \Delta\mathbf{V} \end{bmatrix} = \begin{bmatrix} \Delta\mathbf{z}_\Theta \\ \Delta\mathbf{z}_P \\ \Delta\mathbf{z}_Q \\ \Delta\mathbf{z}_V \end{bmatrix}$$

where $\text{diag}(b)$ and $\text{diag}(g)$ are diagonal matrices with branch susceptances and conductances, respectively, and \mathbf{C} is a node-branch incidence matrix. The transformed matrix $\tilde{\mathbf{H}}_{\mathbf{QV}}$ can then be written as follows:

$$\begin{aligned} \tilde{\mathbf{H}}_{\mathbf{QV}} &= \mathbf{H}_{\mathbf{QV}}^{sh} + \text{diag}(b) \mathbf{C} \\ &\quad + \text{diag}(g) \mathbf{C} (\mathbf{C}' \text{diag}(b)^2 \mathbf{C})^{-1} \mathbf{C}' \text{diag}(b) \text{diag}(g) \mathbf{C} \\ &= \mathbf{H}_{\mathbf{QV}}^{sh} + \mathbf{C} (\text{diag}(b) + \text{diag}(b)^{-1} \text{diag}(g)^2) \\ &= \mathbf{H}_{\mathbf{QV}}^{sh} - \text{diag}(1/x) \mathbf{C} \end{aligned} \tag{12.24}$$

where $\text{diag}(x)$ is a diagonal matrix with branch reactances.

Equations (12.23) and (12.24) show that, at least for the cases of uniform r/x ratios and spanning trees of flow measurements, $-x_{km}$ should be substituted for b_{km} in forming matrix $\tilde{\mathbf{H}}_{\mathbf{QV}}$.

Remarks: Although similar derivations can be made for the dual method, there is an additional complication, since in this case shunt elements and voltage magnitude measurements will affect the elements of $\tilde{\mathbf{H}}_{\mathbf{P}\theta}$. Hence, in order to obtain results similar to those expressed in Eqs. (12.23) and (12.24), it is necessary to ignore the shunts and to assume that there is a single voltage magnitude measurement. Under these assumptions the two versions of the fast decoupled estimator become perfectly symmetrical.

Example 12.5:

Consider the 2-bus system depicted in Fig. 12.4 where the following measurements are performed.

$$\mathbf{z} = \begin{pmatrix} \theta_1^{ref} \\ P_{12}^{meas} \\ P_{21}^{meas} \\ Q_{12}^{meas} \\ Q_{21}^{meas} \\ V_1^{meas} \\ V_2^{meas} \end{pmatrix}$$

The weighting factors associated with the flow measurements are equal to one, and the relative weights of the voltage measurements are w_1 and w_2 . The linearized measurement model computed at flat-start is as follows:

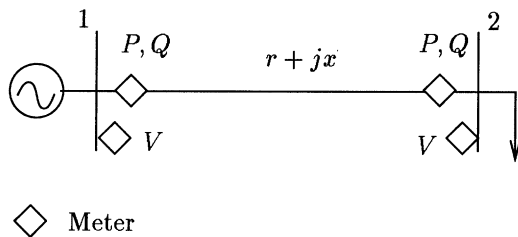


Figure 12.4. Two-bus system.

$$\begin{pmatrix} \infty & 0 & 0 & 0 \\ -b & b & g & -g \\ b & -b & -g & g \\ -b & b & g & -g \\ g & -g & b & -b \\ -g & g & -b & b \\ 0 & 0 & w_1^{1/2} & 0 \\ 0 & 0 & 0 & w_2^{1/2} \end{pmatrix} \begin{pmatrix} \Delta\theta_1 \\ \Delta\theta_2 \\ \Delta V_1 \\ \Delta V_2 \end{pmatrix} = \begin{pmatrix} \infty(\theta_1^{ref} - \theta_1) \\ P_{12}^{meas} - P_{12}(\mathbf{V}, \Theta) \\ P_{21}^{meas} - P_{21}(\mathbf{V}, \Theta) \\ Q_{12}^{meas} - Q_{12}(\mathbf{V}, \Theta) \\ Q_{21}^{meas} - Q_{21}(\mathbf{V}, \Theta) \\ w_1^{1/2}(V_1^{meas} - V_1) \\ w_2^{1/2}(V_2^{meas} - V_2) \end{pmatrix}$$

where

$$\begin{aligned} \mathbf{H}_{P\Theta}^0 &= \begin{pmatrix} b & -b \\ -b & b \end{pmatrix} & \mathbf{H}_{PV}^0 &= \begin{pmatrix} -g & g \\ g & -g \end{pmatrix} \\ \mathbf{H}_{Q\Theta}^0 &= \begin{pmatrix} g & -g \\ -g & g \end{pmatrix} & \mathbf{H}_{QV}^0 &= \begin{pmatrix} b & -b \\ -b & b \end{pmatrix} \end{aligned}$$

The transformed Jacobian submatrix $\tilde{\mathbf{H}}_{QV}$ is

$$\begin{aligned} \tilde{\mathbf{H}}_{QV} &= \mathbf{H}_{QV}^0 - \mathbf{H}_{Q\Theta}^0 \mathbf{H}_{P\Theta}^I \mathbf{H}_{PV}^0 \\ &= \begin{pmatrix} -1/x & 1/x \\ 1/x & -1/x \end{pmatrix} \end{aligned}$$

and the transformed reactive measurement mismatch is

$$\begin{aligned} \Delta\tilde{z}_Q &= \Delta z_Q(\mathbf{V}, \Theta) - \mathbf{H}_{Q\Theta} \mathbf{H}_{P\Theta}^I \Delta z_P \\ &\simeq \Delta z_Q(\mathbf{V}, \Theta + \mathbf{H}_{P\Theta}^I \Delta z_P) \end{aligned}$$

These equations illustrate that the matrix operations performed to zeroize the coupling submatrix $\mathbf{H}_{Q\Theta}$, (a) have transformed the submatrix \mathbf{H}_{QV} into a new submatrix $\tilde{\mathbf{H}}_{QV}$ in which the branch susceptances b have been replaced by $1/x$,

and (b) have yielded new reactive measurement mismatches that are computed at the updated angle $\Theta + \mathbf{H}_{\mathbf{P}\Theta}^I \Delta \mathbf{z}_{\mathbf{P}}$.

Remarks:

This example also serves to illustrate the additional approximations involved in the dual fast decoupled estimator (the XB version). Consider the following two cases: (a) $w_1 \rightarrow \infty$ and $w_2 = 0$; and (b) $w_1 \rightarrow \infty$ and $w_2 \rightarrow \infty$. In the first case, the matrix $\tilde{\mathbf{H}}_{\mathbf{P}\Theta}$ is

$$\tilde{\mathbf{H}}_{\mathbf{P}\Theta} = \begin{pmatrix} 1/x \\ -1/x \end{pmatrix}$$

and in the second case is

$$\tilde{\mathbf{H}}_{\mathbf{P}\Theta} = \begin{pmatrix} -b \\ b \end{pmatrix}$$

which shows that for the XB version the replacement of b by $-1/x$ is affected by the weights given to the voltage magnitude measurements.

12.4.4 Extended BX Method

Consider the ν th iteration of the Gauss Newton method in which the following approximate Jacobian matrix is used (submatrices $\mathbf{H}_{\mathbf{P}\Theta}^0$ and $\mathbf{H}_{\mathbf{Q}\mathbf{V}}^0$ are computed at flat-start):

$$\begin{bmatrix} \mathbf{H}_{\mathbf{P}\Theta}^0 & \mathbf{H}_{\mathbf{P}\mathbf{V}}^\nu \\ \mathbf{H}_{\mathbf{Q}\Theta}^\nu & \mathbf{H}_{\mathbf{Q}\mathbf{V}}^0 \end{bmatrix} \begin{bmatrix} \Delta \Theta^\nu \\ \Delta \mathbf{V}^\nu \end{bmatrix} = \begin{bmatrix} \Delta \mathbf{z}_{\mathbf{P}}^\nu \\ \Delta \mathbf{z}_{\mathbf{Q}}^\nu \end{bmatrix} \quad (12.25)$$

Premultiplying $\Delta \mathbf{z}_{\mathbf{P}}^\nu$ by $-\mathbf{H}_{\mathbf{Q}\Theta}^\nu \mathbf{H}_{\mathbf{P}\Theta}^I$ (where $\mathbf{H}_{\mathbf{P}\Theta}^I$ is the pseudo-inverse of $\mathbf{H}_{\mathbf{P}\Theta}^0$), and adding the result to $\Delta \mathbf{z}_{\mathbf{Q}}^\nu$, yields the following transformed system:

$$\begin{bmatrix} \mathbf{H}_{\mathbf{P}\Theta}^0 & \mathbf{H}_{\mathbf{P}\mathbf{V}}^\nu \\ \mathbf{0} & \tilde{\mathbf{H}}_{\mathbf{Q}\mathbf{V}}^\nu \end{bmatrix} \begin{bmatrix} \Delta \Theta^\nu \\ \Delta \mathbf{V}^\nu \end{bmatrix} = \begin{bmatrix} \Delta \mathbf{z}_{\mathbf{P}}^\nu \\ \Delta \tilde{\mathbf{z}}_{\mathbf{Q}}^\nu \end{bmatrix}$$

where

$$\begin{aligned} \tilde{\mathbf{H}}_{\mathbf{Q}\mathbf{V}}^\nu &= \mathbf{H}_{\mathbf{Q}\mathbf{V}}^0 - \mathbf{H}_{\mathbf{Q}\Theta}^\nu \mathbf{H}_{\mathbf{P}\Theta}^I \mathbf{H}_{\mathbf{P}\mathbf{V}}^\nu \\ \tilde{\mathbf{z}}_{\mathbf{Q}}^\nu &= \mathbf{z}_{\mathbf{Q}}^\nu - \mathbf{H}_{\mathbf{Q}\Theta}^\nu \mathbf{H}_{\mathbf{P}\Theta}^I \mathbf{z}_{\mathbf{P}}^\nu \end{aligned} \quad (12.26)$$

The pseudo-inverse $\mathbf{H}_{\mathbf{P}\Theta}^I = ((\mathbf{H}_{\mathbf{P}\Theta}^0)' \mathbf{H}_{\mathbf{P}\Theta}^0)^{-1} (\mathbf{H}_{\mathbf{P}\Theta}^0)'$, and the Jacobian submatrices $\mathbf{H}_{\mathbf{Q}\Theta}^\nu$ and $\mathbf{H}_{\mathbf{P}\mathbf{V}}^\nu$ that appear in the expression for $\tilde{\mathbf{H}}_{\mathbf{Q}\mathbf{V}}^\nu$ are now computed at flat-start, although in computing the transformed residual vector $\tilde{\mathbf{z}}_{\mathbf{Q}}^\nu$ it is used the Jacobian submatrix $\mathbf{H}_{\mathbf{Q}\Theta}^\nu$ calculated at ν . The transformed matrix $\mathbf{H}_{\mathbf{Q}\mathbf{V}}^0$ is further approximated by replacing the branch susceptances, b , by the inverses of the branch reactances, $-1/x$. This then yields the following 3-step algorithm:

Algorithm for the ν th iteration

- Step 1: Determine temporary value of $\Delta\Theta^\nu$ as follows:

$$\Delta\Theta_{tmp}^\nu = \mathbf{H}_{P\Theta}^I \Delta\mathbf{z}_P(\mathbf{V}^\nu, \Theta^\nu)$$

- Step 2: Determine $\Delta\mathbf{V}^\nu$ as follows:

$$\begin{aligned}\Delta\tilde{\mathbf{z}}_Q &= \Delta\mathbf{z}_Q(\mathbf{V}^\nu, \Theta^\nu) - \mathbf{H}_{Q\Theta}^\nu \Delta\Theta_{tmp}^\nu \\ \Delta\mathbf{V}^\nu &= \tilde{\mathbf{H}}_{QV}^I \Delta\tilde{\mathbf{z}}_Q\end{aligned}$$

- Step 3: Determine the complementary angle $\Delta\Theta_{comp}^\nu$ correction and the final correct value of $\Delta\Theta^\nu$ as follows:

$$\begin{aligned}\Delta\Theta_{comp}^\nu &= -\mathbf{H}_{P\Theta}^I \mathbf{H}_{PV} \Delta\mathbf{V}^\nu \\ \Delta\Theta^\nu &= \Delta\Theta_{tmp}^\nu + \Delta\Theta_{comp}^\nu\end{aligned}$$

Remarks:

- Note that the correction $\Delta\Theta_{tmp} = \mathbf{H}_{P\Theta}^I \Delta\mathbf{z}_P$ and $\Delta\Theta_{comp} = -\mathbf{H}_{P\Theta}^I \mathbf{H}_{PV} \Delta\mathbf{V}$ computed in steps (i) and (ii) correspond to the first blocked equation given by Eq. (12.18), i.e.,

$$\mathbf{H}_{P\Theta} \Delta\Theta + \mathbf{H}_{PV} \Delta\mathbf{V} = \Delta\mathbf{z}_P$$

- Step 1 performs a $P\Theta$ state estimation by solving the following overdetermined system

$$\mathbf{H}_{P\Theta} \Delta\Theta = \Delta\mathbf{z}_P$$

This step is suggested by Eq. (12.26), or (12.19), as a necessary intermediate step for calculating $\Delta\tilde{\mathbf{z}}_Q$.

- Step 2 follows from the reactive part of Eq. (12.18), i.e.,

$$\tilde{\mathbf{H}}_{QV} \Delta\mathbf{V} = \Delta\tilde{\mathbf{z}}_Q$$

- Step 3 follows from the active part of Eq. (12.18), i.e.,

$$\mathbf{H}_{P\Theta} \Delta\Theta = \Delta\tilde{\mathbf{z}}_P = \Delta\mathbf{z}_P - \mathbf{H}_{PV} \Delta\mathbf{V}$$

The angle corrections corresponding to $\Delta\mathbf{z}_P$ are computed in Step 1, and the complementary corrections that correspond to $-\mathbf{H}_{PV} \Delta\mathbf{V}$, are computed in Step 3.

12.4.5 Extended XB Method

Consider again the ν th iteration of the Gauss Newton method as in Eq. (12.25). Premultiplying $\Delta \mathbf{z}_Q^\nu$ by $-\mathbf{H}_{P\nu}^\nu \mathbf{H}_{Q\nu}^I$ (where $\mathbf{H}_{Q\nu}^I$ is the pseudo-inverse of $\mathbf{H}_{Q\nu}^0$), and adding the result to $\Delta \mathbf{z}_P^\nu$, yields the following transformed system:

$$\begin{bmatrix} \tilde{\mathbf{H}}_{P\Theta}^\nu & \mathbf{0} \\ \mathbf{H}_{Q\Theta}^\nu & \mathbf{H}_{Q\nu}^I \end{bmatrix} \begin{bmatrix} \Delta \Theta^\nu \\ \Delta \mathbf{V}^\nu \end{bmatrix} = \begin{bmatrix} \Delta \tilde{\mathbf{z}}_P^\nu \\ \Delta \mathbf{z}_Q^\nu \end{bmatrix}$$

where

$$\begin{aligned} \tilde{\mathbf{H}}_{P\Theta}^\nu &= \mathbf{H}_{P\Theta}^0 - \mathbf{H}_{P\nu}^\nu \mathbf{H}_{Q\nu}^I \mathbf{H}_{Q\Theta}^\nu \\ \tilde{\mathbf{z}}_P^\nu &= \mathbf{z}_P^\nu - \mathbf{H}_{P\nu}^\nu \mathbf{H}_{Q\nu}^I \mathbf{z}_Q^\nu \end{aligned} \quad (12.27)$$

The pseudo-inverse $\mathbf{H}_{Q\nu}^I = ((\mathbf{H}_{Q\nu}^0)' \mathbf{H}_{Q\nu}^0)^{-1} (\mathbf{H}_{Q\nu}^0)'$, and the Jacobian submatrices $\mathbf{H}_{P\nu}^\nu$ and $\mathbf{H}_{Q\Theta}^\nu$ that appear in the expression for $\tilde{\mathbf{H}}_{P\Theta}^\nu$ are now computed at flat-start, whereas the Jacobian submatrix $\mathbf{H}_{P\nu}^\nu$ that appears in the modified residual vector $\tilde{\mathbf{z}}_P^\nu$ is computed at ν . The transformed matrix $\mathbf{H}_{P\Theta}^I$ is further approximated by replacing the branch susceptances, b , by the inverses of the branch reactances, $-1/x$. This then yields the following 3-step algorithm:

Algorithm for the ν th iteration

- Step 1: Determine temporary value of $\Delta \mathbf{V}_{tmp}^\nu$ as follows:

$$\Delta \mathbf{V}_{tmp}^\nu = \mathbf{H}_{Q\nu}^I \Delta \mathbf{z}_Q(\mathbf{V}^\nu, \Theta^\nu)$$

- Step 2: Determine $\Delta \Theta^\nu$ as follows:

$$\begin{aligned} \Delta \tilde{\mathbf{z}}_P &= \Delta \mathbf{z}_P(\mathbf{V}^\nu, \Theta^\nu) - \mathbf{H}_{Q\Theta}^\nu \Delta \mathbf{V}_{tmp}^\nu \\ \Delta \Theta^\nu &= \tilde{\mathbf{H}}_{P\Theta}^I \Delta \tilde{\mathbf{z}}_P \end{aligned}$$

- Step 3: Determine the complementary angle $\Delta \mathbf{V}_{comp}^\nu$ correction and the final correct value of $\Delta \mathbf{V}^\nu$ as follows:

$$\begin{aligned} \Delta \mathbf{V}_{comp}^\nu &= -\mathbf{H}_{Q\nu}^I \mathbf{H}_{Q\Theta}^\nu \Delta \mathbf{V}^\nu \\ \Delta \mathbf{V}^\nu &= \Delta \mathbf{V}_{tmp}^\nu + \Delta \mathbf{V}_{comp}^\nu \end{aligned}$$

12.4.6 BX Method

For the 3-step primal decoupled procedure above, consider the computations performed in Step 3 of the ν th iteration, and in Step 1 of the $(\nu+1)$ th iteration, i.e.,

$$\begin{aligned}\Delta\Theta_{comp}^\nu &= -\mathbf{H}_{\mathbf{P}\Theta}^I \mathbf{H}_{\mathbf{P}\mathbf{V}} \Delta\mathbf{V}^\nu \\ \Theta^{\nu+1} &= \Theta^\nu + \Delta\Theta_{tmp}^\nu + \Delta\Theta_{comp}^\nu\end{aligned}$$

$$\begin{aligned}\Delta\Theta_{tmp}^{\nu+1} &= \mathbf{H}_{\mathbf{P}\Theta}^I \Delta\mathbf{z}_{\mathbf{P}}(\mathbf{V}^{\nu+1}, \Theta^{\nu+1}) \\ \Theta^{\nu+1} &= \Theta^{\nu+1} + \Delta\Theta_{tmp}^{\nu+1}\end{aligned}$$

By combining the angle calculation performed in Step 3 of the ν th iteration with the angle calculation of the $(\nu+1)$ th iteration into a single iteration (a single forward/backward substitution), yields:

$$\begin{aligned}\Delta\Theta_{comp}^\nu + \Delta\Theta_{tmp}^{\nu+1} &= \mathbf{H}_{\mathbf{P}\Theta}^I (\Delta\mathbf{z}_{\mathbf{P}}(\mathbf{V}^{\nu+1}, \Theta^{\nu+1}) - \mathbf{H}_{\mathbf{P}\mathbf{V}} \Delta\mathbf{V}^\nu) \\ &\simeq \mathbf{H}_{\mathbf{P}\Theta}^I (\Delta\mathbf{z}_{\mathbf{P}}(\mathbf{V}^{\nu+1}, \Theta^\nu + \Delta\Theta_{tmp}^\nu) - \mathbf{H}_{\mathbf{P}\Theta} \Delta\Theta_{comp}^\nu - \mathbf{H}_{\mathbf{P}\mathbf{V}} \Delta\mathbf{V}^\nu)\end{aligned}$$

Since (see Remarks below),

$$\mathbf{H}_{\mathbf{P}\Theta}^I (\mathbf{H}_{\mathbf{P}\Theta} \Delta\Theta_{comp}^\nu + \mathbf{H}_{\mathbf{P}\mathbf{V}} \Delta\mathbf{V}^\nu) = \mathbf{0}$$

it follows that

$$\Delta\Theta_{comp}^\nu + \Delta\Theta_{tmp}^{\nu+1} \simeq \mathbf{H}_{\mathbf{P}\Theta}^I (\Delta\mathbf{z}_{\mathbf{P}}(\mathbf{V}^{\nu+1}, \Theta^\nu + \Delta\Theta_{tmp}^\nu))$$

which shows that the combined angle correction can be computed in a single forward-backward solution using the active power mismatch $\Delta\mathbf{z}_{\mathbf{P}}$ computed at $(\mathbf{V}^{\nu+1}, \Theta^{\nu+1})$. This then yields the following 2-step algorithm:

Algorithm for the ν th iteration

- Step 1: Update angle estimate as follows:

$$\begin{aligned}\Delta\Theta^\nu &= \mathbf{H}_{\mathbf{P}\Theta}^I \Delta\mathbf{z}_{\mathbf{P}}(\mathbf{V}^\nu, \Theta^\nu) \\ \Theta^{\nu+1} &= \Theta^\nu + \Delta\Theta^\nu\end{aligned}$$

- Step 2: Update voltage magnitude estimate as follows:

$$\begin{aligned}\Delta\mathbf{V}^\nu &= \tilde{\mathbf{H}}_{\mathbf{Q}\mathbf{V}}^I (\Delta\mathbf{z}_{\mathbf{Q}}(\mathbf{V}^\nu, \Theta^\nu) - \mathbf{H}_{\mathbf{Q}\Theta}^\nu \Delta\Theta^\nu) \\ \mathbf{V}^{\nu+1} &= \mathbf{V}^\nu + \Delta\mathbf{V}^\nu\end{aligned}$$

Remarks:

- The pseudo-inverse matrices $\mathbf{H}_{\mathbf{P}\Theta}^I$ and $\tilde{\mathbf{H}}_{\mathbf{Q}\mathbf{V}}^I$ are computed at flat-start. The matrix $\tilde{\mathbf{H}}_{\mathbf{Q}\mathbf{V}}^I$ is obtained from $\tilde{\mathbf{H}}_{\mathbf{Q}\mathbf{V}}$ by substituting $-1/x$ for b , as discussed above.
- From Eq. (12.17) it follows:

$$\mathbf{H}_{\mathbf{P}\Theta}\Delta\Theta^\nu + \mathbf{H}_{\mathbf{P}\mathbf{V}}\Delta\mathbf{V}^\nu \simeq \Delta\mathbf{z}_{\mathbf{P}}(\mathbf{V}^\nu, \Theta^\nu)$$

Replacing $\Delta\Theta^\nu$ by $\Delta\Theta_{imp}^\nu + \Delta\Theta_{comp}^\nu$, it follows:

$$\mathbf{H}_{\mathbf{P}\Theta}(\Delta\Theta_{imp}^\nu + \Delta\Theta_{comp}^\nu) + \mathbf{H}_{\mathbf{P}\mathbf{V}}\Delta\mathbf{V}^\nu \simeq \Delta\mathbf{z}_{\mathbf{P}}(\mathbf{V}^\nu, \Theta^\nu)$$

In view that $\Delta\Theta_{imp}^\nu = \mathbf{H}_{\mathbf{P}\Theta}^I\Delta\mathbf{z}_{\mathbf{P}}(\mathbf{V}^\nu, \Theta^\nu)$ this expression can be rewritten as follows:

$$\mathbf{H}_{\mathbf{P}\Theta}\Delta\Theta_{comp}^\nu + \mathbf{H}_{\mathbf{P}\mathbf{V}}\Delta\mathbf{V}^\nu \simeq (\mathbf{I} - \mathbf{H}_{\mathbf{P}\Theta}\mathbf{H}_{\mathbf{P}\Theta}^I)\Delta\mathbf{z}_{\mathbf{P}}(\mathbf{V}^\nu, \Theta^\nu)$$

where \mathbf{I} is an identity matrix. Premultiplying the right-hand side of this expression by $\mathbf{H}_{\mathbf{P}\Theta}^I$ yields the following:

$$\begin{aligned} & \mathbf{H}_{\mathbf{P}\Theta}^I(\mathbf{I} - \mathbf{H}_{\mathbf{P}\Theta}\mathbf{H}_{\mathbf{P}\Theta}^I)\Delta\mathbf{z}_{\mathbf{P}}(\mathbf{V}^\nu, \Theta^\nu) \\ &= (\mathbf{H}_{\mathbf{P}\Theta}'\mathbf{H}_{\mathbf{P}\Theta})^{-1}\mathbf{H}_{\mathbf{P}\Theta}'(\mathbf{I} - \mathbf{H}_{\mathbf{P}\Theta}(\mathbf{H}_{\mathbf{P}\Theta}'\mathbf{H}_{\mathbf{P}\Theta})^{-1}\mathbf{H}_{\mathbf{P}\Theta}')\Delta\mathbf{z}_{\mathbf{P}}(\mathbf{V}^\nu, \Theta^\nu) \\ &= \mathbf{0} \end{aligned}$$

- In Step 2 the approximation

$$\Delta\mathbf{z}_{\mathbf{Q}}(\mathbf{V}^\nu, \Theta^\nu) - \mathbf{H}_{\mathbf{Q}\Theta}^\nu\Delta\Theta^\nu \simeq \Delta\mathbf{z}_{\mathbf{Q}}(\mathbf{V}^\nu, \Theta^\nu + \Delta\Theta^\nu)$$

can be used instead, although the linear approximation may be more effective for networks with large r/x ratios, as with the fast decoupled power flows.

Example 12.6:

A 138 kV transmission line section has a series impedance of $0.0062 + j0.036$ p.u., a total shunt admittance of 0.0104 p.u. The following measurements are available: $P_{12}^{meas} = 1.711$, $P_{21}^{meas} = -1.614$, $Q_{12}^{meas} = 0.2245$, $Q_{21}^{meas} =$

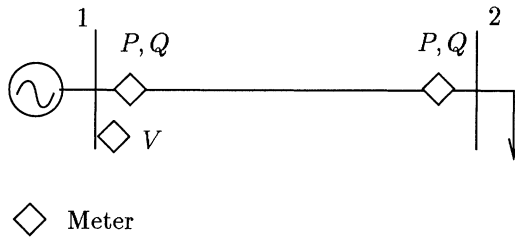


Figure 12.5. Two-bus system.

-0.1566 , $V_1^{meas} = 1.016$. This problem is now solved using the following algorithms: (a) the honest Gauss Newton method, (b) dishonest Gauss Newton method, and (c) the BX fast decoupled state estimator.

Honest Gauss Newton method:

- First Iteration: The Jacobian matrix and the vector of residuals (measurement mismatches) computed at flat-start are as follows:

$$\mathbf{H}^0 = \begin{matrix} & \theta_2 & V_1 & V_2 \\ \begin{matrix} P_{12} \\ P_{21} \\ Q_{12} \\ Q_{21} \\ V_1 \end{matrix} & \begin{pmatrix} -26.98 & 4.65 & -4.65 \\ 26.98 & -4.65 & 4.65 \\ 4.65 & 26.97 & -26.98 \\ -4.65 & -26.98 & 26.97 \\ 0 & 1 & 0 \end{pmatrix} & \Delta \mathbf{z}^0 = \begin{matrix} P_{12} \\ P_{21} \\ Q_{12} \\ Q_{21} \\ V_1 \end{matrix} & \begin{pmatrix} 1.711 \\ -1.614 \\ 0.2297 \\ -0.1514 \\ 0.016 \end{pmatrix} \end{matrix}$$

The state correction, given by the solution of the normal equations, and the updated state are as follows:

$$\Delta \mathbf{x}^0 = \begin{pmatrix} \Delta \theta_2 \\ \Delta V_1 \\ \Delta V_2 \end{pmatrix} = \begin{pmatrix} -0.0587 \\ 0.0158 \\ -0.00133 \end{pmatrix}$$

$$\mathbf{x}^1 = \mathbf{x}^0 + \Delta \mathbf{x}^0 = \begin{pmatrix} 0.000 \\ 1.000 \\ 1.000 \end{pmatrix} + \begin{pmatrix} -0.0587 \\ 0.0158 \\ -0.00133 \end{pmatrix} = \begin{matrix} \theta_2 \\ V_1 \\ V_2 \end{matrix} \begin{pmatrix} -0.0587 \\ 1.0158 \\ 0.9987 \end{pmatrix}$$

- Second Iteration:

$$\mathbf{H}^1 = \begin{matrix} & \theta_2 & V_1 & V_2 \\ P_{12} & \begin{pmatrix} -27.60 & 6.39 & -3.10 \end{pmatrix} \\ P_{21} & \begin{pmatrix} 27.04 & -6.21 & 2.96 \end{pmatrix} \\ Q_{12} & \begin{pmatrix} 3.10 & 27.63 & -27.63 \end{pmatrix} \\ Q_{21} & \begin{pmatrix} -6.31 & -26.62 & 26.79 \end{pmatrix} \\ V_1 & \begin{pmatrix} 0 & 1 & 0 \end{pmatrix} \end{matrix} \quad \Delta \mathbf{z}^1 = \begin{matrix} P_{12} \\ P_{21} \\ Q_{12} \\ Q_{21} \\ V_1 \end{matrix} \begin{pmatrix} 0.0171 \\ 0.0623 \\ -0.0114 \\ -0.0123 \\ 0.00016 \end{pmatrix}$$

The state correction and the updated state are as follows:

$$\Delta \mathbf{x}^0 = \begin{matrix} \Delta \theta_2 \\ \Delta V_1 \\ \Delta V_2 \end{matrix} \begin{pmatrix} 0.000861 \\ 0.000352 \\ 0.000486 \end{pmatrix}$$

$$\mathbf{x}^2 = \mathbf{x}^1 + \Delta \mathbf{x}^1 = \begin{pmatrix} -0.0587 \\ 1.0158 \\ 0.9987 \end{pmatrix} + \begin{pmatrix} 0.0008617 \\ 0.000352 \\ 0.000486 \end{pmatrix} = \begin{matrix} \theta_2 \\ V_1 \\ V_2 \end{matrix} \begin{pmatrix} -0.0578 \\ 1.0162 \\ 0.9992 \end{pmatrix}$$

Dishonest Gauss Newton method:

- First Iteration: The Jacobian matrix and the measurement mismatches computed at flat-start are the same as for the honest Gauss Newton method. The updated state is as follows:

$$\mathbf{x}^1 = \mathbf{x}^0 + \Delta \mathbf{x}^1 = \begin{matrix} \theta_2 \\ V_1 \\ V_2 \end{matrix} \begin{pmatrix} -0.0587 \\ 1.0158 \\ 0.9987 \end{pmatrix}$$

- Second Iteration: The Jacobian matrix is the same as in the previous iteration. The vector of residuals computed at \mathbf{x}^1 is

$$\Delta \mathbf{z}^1 = \begin{matrix} P_{12} \\ P_{21} \\ Q_{12} \\ Q_{21} \\ V_1 \end{matrix} \begin{pmatrix} 0.0171 \\ 0.0623 \\ -0.2878 \\ 0.2641 \\ 0.00016 \end{pmatrix}$$

The state correction and the updated state are as follows:

$$\Delta \mathbf{x}^0 = \begin{matrix} \Delta \theta_2 \\ \Delta V_1 \\ \Delta V_2 \end{matrix} \begin{pmatrix} -0.000898 \\ 0.000212 \\ 0.01029 \end{pmatrix}$$

$$\mathbf{x}^1 = \mathbf{x}^0 + \Delta \mathbf{x}^0 = \begin{pmatrix} -0.0587 \\ 1.0158 \\ 0.9987 \end{pmatrix} + \begin{pmatrix} -0.000898 \\ 0.000212 \\ 0.01029 \end{pmatrix} = \begin{matrix} \theta_2 \\ V_1 \\ V_2 \end{matrix} \begin{pmatrix} -0.0596 \\ 1.0160 \\ 1.0090 \end{pmatrix}$$

Fast Decoupled State Estimator (BX version):

- **First P Θ Iteration:** The Jacobian submatrix $\mathbf{H}_{P\Theta}$ and the vector of residuals $\Delta \mathbf{z}_P$ are as follows:

$$\mathbf{H}_{P\Theta}^0 = \begin{matrix} \theta_2 \\ P_{12} \\ P_{21} \end{matrix} \begin{pmatrix} -26.98 \\ 26.98 \end{pmatrix} \quad \Delta \mathbf{z}_P^0 = \begin{matrix} P_{12} \\ P_{21} \end{matrix} \begin{pmatrix} 1.711 \\ -1.614 \end{pmatrix}$$

The angle correction, given by the solution of the normal equations, and the updated angle are as follows:

$$\begin{aligned} \Delta \theta_2^0 &= -0.06163 \\ \theta_2^1 &= \theta_2^0 + \Delta \theta_2^0 = 0.000 - 0.06163 = -0.06163 \end{aligned}$$

- **First QV Iteration:** The Jacobian submatrix \mathbf{H}_{QV} and the vector of residuals $\Delta \mathbf{z}_Q$ are as follows:

$$\tilde{\mathbf{H}}_{QV}^0 = \begin{matrix} V_1 & V_2 \\ Q_{12} \\ Q_{21} \\ V_1 \end{matrix} \begin{pmatrix} 27.77 & -27.78 \\ -27.78 & 27.77 \\ 1 & 0 \end{pmatrix} \quad \Delta \tilde{\mathbf{z}}^0 = \begin{matrix} Q_{12} \\ Q_{21} \\ V_1 \end{matrix} \begin{pmatrix} 0.4646 \\ -0.4888 \\ 0.015 \end{pmatrix}$$

The voltage magnitude corrections and the updated magnitudes are as follows:

$$\Delta \mathbf{V}^0 = \begin{matrix} \Delta V_1 \\ \Delta V_2 \end{matrix} \begin{pmatrix} 0.0150 \\ -0.00211 \end{pmatrix}$$

$$\mathbf{V}^1 = \mathbf{V}^0 + \Delta \mathbf{V}^0 = \begin{pmatrix} 1.0000 \\ 1.0000 \end{pmatrix} + \begin{pmatrix} 0.0150 \\ -0.00211 \end{pmatrix} = \begin{matrix} V_1 \\ V_2 \end{matrix} \begin{pmatrix} 1.0150 \\ 0.9979 \end{pmatrix}$$

- Second P Θ Iteration: The Jacobian submatrix $\mathbf{H}_{P\Theta}$ is the same as above. The vector of residuals $\Delta \mathbf{z}_P$ is

$$\Delta \mathbf{z}_P^1 = \begin{pmatrix} P_{12} & -0.00619 \\ P_{21} & 0.1396 \end{pmatrix}$$

The angle correction and the updated angle are as follows:

$$\begin{aligned} \Delta \theta_2^1 &= 0.003734 \\ \theta_2^2 &= \theta_2^1 + \Delta \theta_2^1 = -0.06163 + 0.00373 = -0.05790 \end{aligned}$$

- Second QV Iteration: The Jacobian submatrix \mathbf{H}_{QV} is the same as above. The vector of residuals $\Delta \mathbf{z}_Q$ is

$$\Delta \tilde{\mathbf{z}}^1 = \begin{pmatrix} Q_{12} & -0.01368 \\ Q_{21} & -0.00741 \\ V_1 & 0.00095 \end{pmatrix}$$

The voltage magnitude corrections and the updated magnitudes are

$$\Delta \mathbf{V}^1 = \begin{pmatrix} \Delta V_1 & 0.000994 \\ \Delta V_2 & 0.001107 \end{pmatrix}$$

$$\mathbf{V}^2 = \mathbf{V}^1 + \Delta \mathbf{V}^1 = \begin{pmatrix} 1.0150 \\ 0.9979 \end{pmatrix} + \begin{pmatrix} 0.000994 \\ 0.001107 \end{pmatrix} = \begin{pmatrix} V_1 & 1.0160 \\ V_2 & 0.9990 \end{pmatrix}$$

Remarks: Hence, the quality of the results obtained with the fast decoupled state estimator, although not as good as those obtained with the honest Gauss Newton method, are better than those obtained with the dishonest Gauss Newton method. This is consistent with the derivations above.

12.4.7 XB Method

With the additional assumptions discussed earlier, the following XB version of the fast decoupled state estimator can be derived:

Algorithm for the ν th iteration

- Step 1: Update voltage magnitude estimate as follows:

$$\begin{aligned}\Delta \mathbf{V}^\nu &= \mathbf{H}_{\mathbf{QV}}^T \Delta \mathbf{z}_{\mathbf{Q}}(\mathbf{V}^\nu, \Theta^\nu) \\ \mathbf{V}^{\nu+1} &= \mathbf{V}^\nu + \Delta \mathbf{V}^\nu\end{aligned}$$

- Step 2: Update angle estimate as follows:

$$\begin{aligned}\Delta \Theta^\nu &= \tilde{\mathbf{H}}_{\mathbf{P}\Theta}^T (\Delta \mathbf{z}_{\mathbf{P}}(\mathbf{V}^\nu, \Theta^\nu) - \mathbf{H}_{\mathbf{P}\Theta}^\nu \Delta \Theta^\nu) \\ \Theta^{\nu+1} &= \Theta^\nu + \Delta \Theta^\nu\end{aligned}$$

Remarks: As observed above in connection with the fast decoupled power flow, version XB, it might be better to perform these two iterations in reverse order to benefit from the good initialization normally given by the dc estimator (iteration P Θ).

12.5 HISTORICAL NOTES AND REFERENCES

The Newton Raphson method for solving the power flow problem was first proposed by Van Ess and Griffin [1961], and sparse matrix techniques as a means of implementing Newton Raphson method for large networks was introduced by Tinney and Hart [1967]. Stott and Alsac [1974] proposed the fast decoupled power flow method (the so called XB version). Early developments in digital power flow calculations were reviewed by Tinney and Powell [1971] and Stott [1974]. An “exact” decoupling method (the CRIC method) was introduced by Carpentier [1986], and a hybrid version for the fast decoupled power flow method for systems where branches with critical r/x ratios are solved by the Newton Raphson method was suggested by Rajjicic and Bose [1988]. A new version was later presented for the fast decoupled power flow (BX version) (Van Amerongen [1989]), and a new decoupling theory capable of explaining the behavior of both XB and BX versions was proposed by Monticelli, Garcia and Saavedra [1990].

The extension of the fast decoupling principle to state estimation was first suggested by Horisberger, Richard, and Rossier [1976] who used a decoupled algorithm in which no approximations were introduced into the the right-hand side of the normal equation (algorithm-decoupling). The performance of the method was not comparable with that obtained with the fast decoupled power flow, however. Model decoupled estimators, which introduce the decoupling approximation into the Jacobian matrix (which affects both sides of the normal equations), rather than only into the gain matrix, were studied by a number of

authors, including Aschmoneit, Denzel, Graf, and Schellstede [1976], Garcia, Monticelli, and Abreu [1979], and Allemong, Radu, and Sasson [1982]. These papers have shown that the key point for the success of the fast decoupled estimators is the introduction of the same approximations in both sides of the coupled normal equation, and that the errors associated with such model decoupling are acceptable for state estimation purposes. A theory supporting fast decoupled state estimators was suggested (Monticelli and Garcia [1990]).

12.6 PROBLEMS

- 1. Using the Extended BX method (Subsec. 12.2.4), determine the converged solutions for the power flow problems (a) and (b) in Example 12.3. Consider convergence tolerances of $\epsilon = 0.001$ p.u. for the mismatches ΔP and ΔQ .
- 2. Repeat Prob. 1 using the BX method (Subsec. 12.2.6).
- 3. The network in Fig. 12.6 has two 138 kV lines with impedances $z_{12} = 0.0062 + j0.036$ p.u. and $z_{23} = 0.0031 + j0.018$ p.u., and shunt susceptances of $b_{12}^{sh} = 0.052$ p.u. and $b_{23}^{sh} = 0.026$ p.u. For the meter arrangement indicated in parts (a) and (b) of the figure, determine (i) the Jacobian submatrix $\mathbf{H}_{\mathbf{QV}}^0$ (at flat-start), (ii) the transformed matrix $\tilde{\mathbf{H}}_{\mathbf{QV}}^0 = \mathbf{H}_{\mathbf{QV}}^0 - \mathbf{H}_{\mathbf{Q}\Theta}^0 \mathbf{H}_{\mathbf{P}\Theta}^T \mathbf{H}_{\mathbf{P}\mathbf{V}}^0$, and (iii) the approximate transformed matrix $\tilde{\mathbf{H}}_{\mathbf{QV}}^0$ obtained from the matrix $\mathbf{H}_{\mathbf{QV}}^0$ by replacing b by $-1/x$.

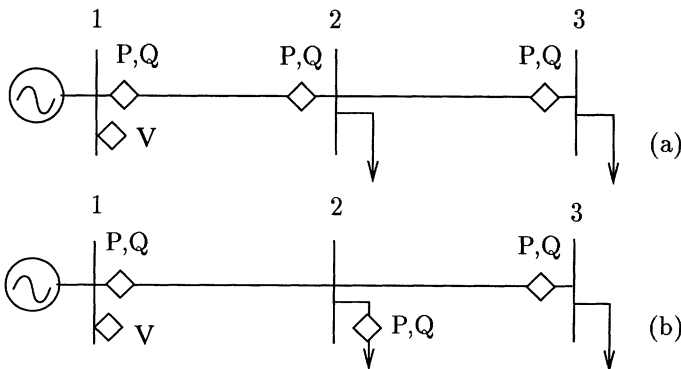


Figure 12.6. Networks used in Prob. 3.

- 4. Repeat Example 12.6 for the XB version of the fast decoupled state estimator.

References

- Allemong, J.J., Radu, L., and Sasson, A.M., "A fast and reliable state estimator algorithm for the AEP's new control center", *IEEE Trans. Power App. and Syst. Syst.*, Vol. 101, pp. 933-944, April 1982.
- Van Amerongen, R.A.M., "A general-purpose version of the fast decoupled load flow", *IEEE Trans. Power App. Syst.*, Vol. 4, pp. 760-770, May 1989.
- Aschmoneit, F., Denzel, D., Graf, R., and Schellstede, G., "Developments of an optimal state estimator and implementation in a real-time computer system", *CIGRE Meeting, Paris, 1976*.
- Carpentier, J.L., "CRIC: A new active reactive decoupling process in load flows, optimal power flows and system control", *IFAC Symposium on Power Systems and Power Plant Control, Beijing, 1986*.
- Horisberger, H.P., Richard, J.C., and Rossier, C., "A fast decoupled static state estimator for electric power systems", *IEEE Trans. Power App. Syst.*, Vol. 95, No. 1, pp. 208-215, Jan./Feb. 1976.
- Garcia, A., Monticelli, A., and Abreu P., "Fast decoupled state estimation and bad data processing", *IEEE Trans. Power App. Syst.*, Vol. 98, No.5, pp. 1645-1652, Sept./Oct. 1979.
- Monticelli, A. and Garcia, A., "Fast decoupled state estimators", *IEEE Trans. Power Syst.*, Vol. 5, No. 2, pp. 556-564, May 1990.
- Monticelli, A., Garcia, A., and Saavedra, O. R., "Fast decoupled load flow – Hypothesis, derivations and testing", *IEEE Trans. Power Syst.*, Vol. 5, No. 4, pp. 1425-1431, Nov. 1990.
- Rajicic, D. and Bose, A., "A Modification to the Fast Decoupled Power Flow for Network with High R/X Ratios", *IEEE Trans. Power App. Syst.*, Vol. 3, pp. 743-746, May 1988.
- Stott, B. and Alsac, O., "Fast Decoupled Load Flow", *IEEE Trans. Power App. Syst.*, Vol. 93, pp. 859-869, 1974.
- Stott, B., "Review of load flow calculations methods", *Proceedings of IEEE*, Vol. 62, pp. 916-929, 1974.
- Tinney, W. F. and Hart, C. E., "Power flow solution by Newton's method", *IEEE Trans. Power App. Syst.*, Vol. 86, pp. 1449-1456, 1967.
- Tinney, W. F. and Powell, W. L., "Notes on Newton-Raphson method for solution of AC power flow problem", *IEEE Short Course, Power Systems Planning, 1971*.
- Van Ess, J. E. and Griffin, J. H., "Elimination methods for load flow studies", *AIEE Transactions*, Vol. 80, pp. 299-304, 1961.

13

NUMERICALLY ROBUST STATE ESTIMATORS

This chapter discusses the numerical condition of the weighted least-squares state estimation problem and reviews the main techniques developed to enhance numerical robustness (Numerical robustness is not to be confused with robust estimators, or Huber estimators, discussed in Chap. 9.)

13.1 NORMAL EQUATION

Consider the state estimation measurement model

$$\mathbf{z} = \mathbf{h}(\mathbf{x}) + \mathbf{e}$$

where $\mathbf{h}(\cdot)$ is a nonlinear vector function (an m -vector), $\mathbf{z}(t_i)$ is a measurement (an m -vector), \mathbf{x} is the true state vector (an n -vector), \mathbf{e} is a zero mean error vector with variance \mathbf{R}_z (an $m \times m$ diagonal matrix), m is the number of measurements and n is the number of state variables. This model can be rewritten in the following weighted form:

$$\mathbf{z}^w = \mathbf{h}^w(\mathbf{x}) + \mathbf{e}^w$$

where $\mathbf{z}^w = \mathbf{R}_z^{-1/2} \mathbf{z}$ is the weighted measurement vector and \mathbf{e}^w is a zero mean vector with unit covariance matrix. In this chapter the weighted form of the measurement model will be used and, for simplicity, the superscript w will be dropped from all equations.

The problem is to determine the estimate $\hat{\mathbf{x}}$ that best fits the measurement model. The weighted least-squares solution for the problem can be obtained by minimizing the following index:

$$J(\mathbf{x}) = \frac{1}{2} \mathbf{r}' \mathbf{r}$$

where $\mathbf{r} = \mathbf{z} - \mathbf{h}(\mathbf{x})$ is the weighted residual vector. This performance index $J(\mathbf{x})$ can be differentiated to obtain the first-order optimal conditions:

$$\mathbf{H}'(\mathbf{x}) (\mathbf{z} - \mathbf{h}(\mathbf{x})) = \mathbf{0} \quad (13.1)$$

where $\mathbf{H} = \partial \mathbf{h}(\mathbf{x}) / \partial \mathbf{x}$ is the Jacobian matrix.

13.1.1 Basic formulation

A least-squares solution to Eq. (13.1) can be obtained using the Gauss Newton method, in which second order-derivatives are ignored, resulting the following iterative procedure:

$$\begin{aligned} \mathbf{H}'(\mathbf{x}^\nu) \mathbf{H}(\mathbf{x}^\nu) \Delta \mathbf{x}^\nu &= \mathbf{H}'(\mathbf{x}^\nu) \mathbf{r}(\mathbf{x}^\nu) \\ \mathbf{x}^{\nu+1} &= \mathbf{x}^\nu + \Delta \mathbf{x}^\nu \end{aligned} \quad (13.2)$$

for $k = 0, 1, 2, \dots$ until convergence is attained. If the system is observable, the gain matrix $\mathbf{G} = \mathbf{H}'\mathbf{H}$ which appears in the normal equation above is positive definite. Under these circumstances no changing of the pivot order is necessary during the triangular factorization of the gain matrix to improve numerical stability. Ordering is then performed only for the enhancement of sparsity. The numerical performance of the normal equation approach can be negatively affected by (a) the presence of injection measurements, (b) the use of large weighting factors (e.g., to enforce equality constraints), and (c) the presence of very low impedance branches. These problems are related to the fact that the gain matrix is obtained by squaring the Jacobian matrix (see Eq. (13.2)), which makes the condition number worse. Most of the methods developed to improve the numerical robustness of the weighted least-square estimators are thus designed to “unsquare” the normal equation.

13.1.2 Equality Constraints

Consider now the constrained problem:

$$\begin{aligned} \text{Minimize} \quad & J(\mathbf{x}) = \frac{1}{2} \mathbf{r}' \mathbf{r} \\ \text{subject to} \quad & \mathbf{c}(\mathbf{x}) = \mathbf{0} \end{aligned}$$

where $\mathbf{c}(\mathbf{x}) = \mathbf{0}$ represents a set of nonlinear constraints such as zero power injections in transition buses. This optimization problem can be expressed by the following Lagrangian function:

$$\mathcal{L}(\mathbf{x}, \Lambda) = \frac{1}{2} \mathbf{r}'(\mathbf{x}) \mathbf{r}(\mathbf{x}) - \Lambda' \mathbf{c}(\mathbf{x})$$

This function can be differentiated to obtain the Karush-Kuhn-Tucker (KKT) first-order necessary conditions for an optimal solution, yielding the following system of nonlinear equations:

$$\partial \mathcal{L} / \partial \mathbf{x} = -\mathbf{H}'(\mathbf{x}) \mathbf{r}(\mathbf{x}) - \mathbf{C}'(\mathbf{x}) \Lambda = \mathbf{0} \tag{13.3}$$

$$\partial \mathcal{L} / \partial \Lambda = -\mathbf{c}(\mathbf{x}) = \mathbf{0} \tag{13.4}$$

where

$$\mathbf{H}(\mathbf{x}) = \frac{\partial \mathbf{h}(\mathbf{x})}{\partial \mathbf{x}}; \quad \mathbf{C}(\mathbf{x}) = \frac{\partial \mathbf{c}(\mathbf{x})}{\partial \mathbf{x}}$$

If the Gauss Newton method is used to solve this system of nonlinear equations iteratively, the following Taylor series expansions are made at the ν th iteration:

$$\mathbf{r}(\mathbf{x}) \simeq \mathbf{r}^\nu - \mathbf{H}(\mathbf{x}^\nu) \Delta \mathbf{x}^\nu$$

$$\mathbf{c}(\mathbf{x}) = \mathbf{c}(\mathbf{x}^\nu) + \mathbf{C}(\mathbf{x}^\nu) \Delta \mathbf{x}^\nu$$

In view of these linear approximations, Eqs. (13.3) and (13.4) can be rewritten as follows:

$$\begin{pmatrix} \mathbf{H}'(\mathbf{x}^\nu) \mathbf{H}(\mathbf{x}^\nu) & -\mathbf{C}'(\mathbf{x}^\nu) \\ -\mathbf{C}(\mathbf{x}^\nu) & \mathbf{0} \end{pmatrix} \begin{pmatrix} \Delta \mathbf{x}^\nu \\ \Lambda^{\nu+1} \end{pmatrix} = \begin{pmatrix} \mathbf{H}'(\mathbf{x}^\nu) \mathbf{r}(\mathbf{x}^\nu) \\ \mathbf{c}(\mathbf{x}^\nu) \end{pmatrix} \tag{13.5}$$

Remarks:

- The advantages of modeling equality constraints as above rather than using pseudo-measurements with very high weights are the following: (a) in Eq. (13.5) the \mathbf{C} matrix is not squared, whereas in Eq. (13.2) these derivatives would be squared, and (b) in Eq. (13.5) no weights are assigned to the equality constraints, so they can be enforced without deteriorating the condition of the coefficient matrix.
- The equality constrained formulation presents two main disadvantages: (a) the coefficient matrix in Eq. (13.5) is indefinite, and (b) if the constraint being enforced involves a bad data, state estimation convergence and error analysis may be negatively affected. If pivot ordering is used only to improve the sparsity of an indefinite system, such as that in Eq. (13.5), zero (or very small) pivots may occur during factorization. Various strategies have thus been proposed in the literature to cope with indefiniteness: (a) delayed

pivoting to delay processing of rows/columns with small pivot elements, (b) blocked sparse matrices, and (c) mixed 1×1 and 2×2 pivoting. The validity of the latter can be mathematically proven (see Historical Notes and References at the end of the chapter); the other two strategies generally require very careful implementation and numerical checking to make sure the method is really applicable to the problem being solved.

Example 13.1:

The network in Fig. 13.1 is now used to illustrate how the numerical condition of the gain matrix is affected by the representation of zero injection (perfect information) in state estimation models. For simplicity, the variance of the flow measurement P_{21} is assumed to be 1 p.u. The zero injection $P_2 = 0$ is first treated as a pseudo-measurement (to which a variance of σ p.u. is arbitrarily assigned) and then as an equality constraint. A dc model with angle reference $\theta_1 = 0^\circ$ is used.

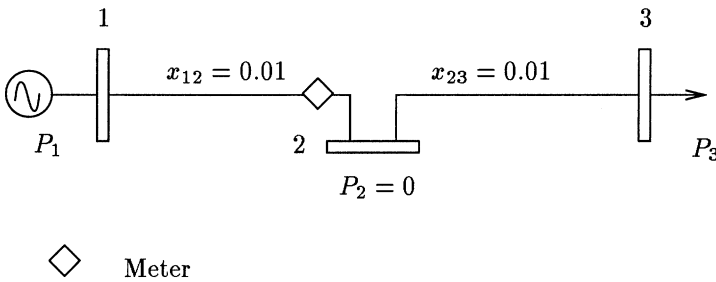


Figure 13.1. Three-bus system (Reactances in p.u.)

When $P_2 = 0$ is treated as a pseudo-measurement, the following Jacobian and gain matrices result; the weight assigned to P_{21} is one, and that assigned to the pseudo-measurement is σ^{-2} :

$$\mathbf{H} = 10^2 \begin{matrix} & \theta_2 & \theta_3 \\ P_{21} & \begin{pmatrix} 1 & 0 \\ 2\sigma^{-1} & -1\sigma^{-1} \end{pmatrix} \\ P_2 & \end{matrix}$$

$$\mathbf{G} = \mathbf{H}'\mathbf{H} = 10^4 \begin{matrix} & \theta_2 & \theta_3 \\ \theta_2 & \begin{pmatrix} 1 + 4\sigma^{-2} & -2\sigma^{-2} \\ -2\sigma^{-2} & \sigma^{-2} \end{pmatrix} \\ \theta_3 & \end{matrix}$$

The numerical condition of this gain matrix is degraded as the σ approaches zero. If, for example, seven-digit floating-point arithmetic is used, the gain matrix will be numerically singular for $\sigma = 10^{-4}$. This means that the problem becomes unsolvable.

Consider now the representation of $P_2 = 200 \theta_2 - 100 \theta_3 = 0$. In this case, $\mathbf{C} = (200; -100)$ and the the vector Λ has a single element λ_{P_2} (the Lagrange multiplier associated with the equality constraint $P_2 = 0$). The augmented matrix in Eq. (13.5) can then be written as follows:

$$\begin{pmatrix} \mathbf{H}'\mathbf{H} & -\mathbf{C}' \\ -\mathbf{C} & 0 \end{pmatrix} = 10^2 \begin{matrix} & \theta_2 & \theta_3 & \lambda_{P_2} \\ \theta_2 & \begin{pmatrix} 1 & 0 & -2 \end{pmatrix} \\ \theta_3 & \begin{pmatrix} 0 & 0 & 1 \end{pmatrix} \\ \lambda_{P_2} & \begin{pmatrix} -2 & 1 & 0 \end{pmatrix} \end{matrix}$$

The augmented matrix above is a nonsingular matrix which can be rearranged and then triangularized as follows:

$$10^2 \begin{matrix} & \theta_2 & \lambda_{P_2} & \theta_3 \\ \theta_2 & \begin{pmatrix} 1 & -2 & 0 \\ -2 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \\ \lambda_{P_2} & & & \\ \theta_3 & & & \end{matrix} \rightarrow 10^2 \begin{matrix} & \theta_2 & \lambda_{P_2} & \theta_3 \\ \theta_2 & \begin{pmatrix} 1 & -2 & 0 \\ 0 & -4 & 1 \\ 0 & 0 & 1/4 \end{pmatrix} \\ \lambda_{P_2} & & & \\ \theta_3 & & & \end{matrix}$$

Hence, the representation of the zero injection $P_2 = 0$ as an equality constraint avoids the potential numerical problems observed with the normal equation approach, although the difficulties related to the occurrence of zero pivots will require further discussion since the pivot reordering as above is not always possible.

13.2 SPARSE TABLEAU FORMULATION

The unsquared representation of equality constraints was discussed above. In this section the unsquared approach is extended to regular measurements.

13.2.1 Basic formulation

The basic formulation deals with normal equation (without equality constraints) in an “unsquared” form. After the ν th iteration the updated vector of residuals can be written as

$$\mathbf{r}^{\nu+1} = \mathbf{z} - \mathbf{h}(\mathbf{x}^\nu)$$

This residual vector can be expanded using the linear terms of the Taylor series:

$$\mathbf{r}^{\nu+1} \simeq \Gamma^{\nu+1} = \mathbf{r}^\nu - \mathbf{H}(\mathbf{x}^\nu) \Delta \mathbf{x}^\nu \tag{13.6}$$

where $\Gamma^{\nu+1}$ is a linear approximation to $\mathbf{r}^{\nu+1}$ after the ν th iteration.

Since the columns of $\mathbf{H}(\mathbf{x}^\nu)$ are orthogonal to the current estimate of the residual vector $\mathbf{r}^{\nu+1}$ (see Fig. 2.6 (a), Chap. 2), the following holds true:

$$\mathbf{H}'(\mathbf{x}^\nu) \Gamma^{\nu+1} = \mathbf{0} \tag{13.7}$$

Equations (13.6)-(13.7) can then be written in the tableau form as follows:

$$\begin{pmatrix} \mathbf{I} & \mathbf{H}(\mathbf{x}^\nu) \\ \mathbf{H}'(\mathbf{x}^\nu) & \mathbf{0} \end{pmatrix} \begin{pmatrix} \Gamma^{\nu+1} \\ \Delta \mathbf{x}^\nu \end{pmatrix} = \begin{pmatrix} \mathbf{r}(\mathbf{x}^\nu) \\ \mathbf{0} \end{pmatrix} \quad (13.8)$$

Remarks:

- Note that if this system is factorized in the order indicated above, elimination using the identity matrix as a pivot leads to the following reduced system:

$$\begin{pmatrix} \mathbf{I} & \mathbf{H}(\mathbf{x}^\nu) \\ \mathbf{0} & -\mathbf{H}'(\mathbf{x}^\nu)\mathbf{H}(\mathbf{x}^\nu) \end{pmatrix} \begin{pmatrix} \Gamma^{\nu+1} \\ \Delta \mathbf{x}^\nu \end{pmatrix} = \begin{pmatrix} \mathbf{r}(\mathbf{x}^\nu) \\ -\mathbf{H}'(\mathbf{x}^\nu) \mathbf{r}(\mathbf{x}^\nu) \end{pmatrix} \quad (13.9)$$

Although Eqs. (13.8) and (13.9) are mathematically equivalent in the sense that if the computations were performed with infinite precision they would provide identical numerical solutions, when finite precision is involved, the results will depend on the pivot ordering, as illustrated in the example below.

- It can be verified that the tableau in Eq. (13.8) formulates the first order optimal conditions for the following minimization problem:

$$\begin{aligned} \text{Minimize} \quad & J(\mathbf{r}) = \frac{1}{2} \mathbf{r}' \mathbf{r} \\ \text{subject to} \quad & \mathbf{r} = \mathbf{z} - \mathbf{h}(\mathbf{x}) \end{aligned}$$

In this case, the Lagrangian function is

$$\mathcal{L}(\mathbf{r}, \Gamma) = \frac{1}{2} \mathbf{r}' \mathbf{r} - \Gamma' (\mathbf{r} - \mathbf{z} + \mathbf{h}(\mathbf{x}))$$

and thus

$$\begin{aligned} \partial \mathcal{L} / \partial \Gamma &= -\mathbf{r} + \mathbf{z} - \mathbf{h}(\mathbf{x}) = \mathbf{0} \\ \partial \mathcal{L} / \partial \mathbf{x} &= -\mathbf{H}'(\mathbf{x}) \Gamma = \mathbf{0} \\ \partial \mathcal{L} / \partial \mathbf{r} &= \mathbf{r} - \Gamma = \mathbf{0} \end{aligned} \quad (13.10)$$

The last equation shows that the multiplier Γ associated with the constraint $\mathbf{r} = \mathbf{z} + \mathbf{h}(\mathbf{x})$ is equal to the residual vector \mathbf{r} . The first equation can then be rewritten as follows:

$$-\Gamma + \mathbf{z} - \mathbf{h}(\mathbf{x}) = \mathbf{0} \quad (13.11)$$

If the updated values $\Gamma^{\nu+1}$ and $\mathbf{h}(\mathbf{x}^{\nu+1}) \simeq \mathbf{h}(\mathbf{x}^{\nu}) + \mathbf{H}(\mathbf{x}^{\nu}) \Delta \mathbf{x}^{\nu}$ are assumed to solve Eqs. (13.10) and (13.11), these equations can be rewritten as follows (Gauss Newton method):

$$\begin{aligned} \Gamma^{\nu+1} + \mathbf{H}(\mathbf{x}^{\nu}) \Delta \mathbf{x}^{\nu} &= \mathbf{r}(\mathbf{x}^{\nu}) \\ \mathbf{H}'(\mathbf{x}^{\nu}) \Gamma^{\nu+1} &= \mathbf{0} \end{aligned}$$

This yields Eq. (13.8).

Example 13.2:

The system in Fig. 13.2 is represented by a dc model with $\theta_1 = 0^\circ$. For simplicity, both measurement error variances are considered to be $\sigma = 1$ p.u.

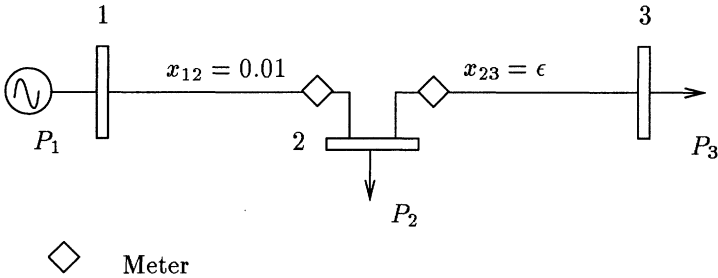


Figure 13.2. Three-bus system with a low impedance branch (Reactances in p.u.)

In this case, the Jacobian matrix and the gain matrix for the normal equation method are as follows:

$$\mathbf{H} = \begin{matrix} & \theta_2 & \theta_3 \\ \begin{matrix} P_{12} \\ P_{23} \end{matrix} & \begin{pmatrix} 10^2 & 0 \\ \epsilon^{-1} & -\epsilon^{-1} \end{pmatrix} \end{matrix}$$

$$\mathbf{G} = \begin{matrix} & \theta_2 & \theta_3 \\ \begin{matrix} \theta_2 \\ \theta_3 \end{matrix} & \begin{pmatrix} 10^4 + \epsilon^{-2} & -\epsilon^{-2} \\ -\epsilon^{-2} & \epsilon^{-2} \end{pmatrix} \end{matrix}$$

If, for example, computations are performed with 7-digit accuracy, the gain matrix above will be numerically singular for $\epsilon < 10^{-6}$, since, in this case, $10^4 + \epsilon^{-2} \simeq \epsilon^{-2}$.

The vector of Lagrange multipliers is

$$\Gamma = \begin{pmatrix} \gamma_{P_{21}} \\ \gamma_{P_{23}} \end{pmatrix}$$

where $\gamma_{P_{21}}$ corresponds to the residual $P_{21}^{meas} - \hat{P}_{21}$ and $\gamma_{P_{23}}$ to $P_{23}^{meas} - \hat{P}_{23}$. In this case the tableau formulation is as follows:

$$\begin{pmatrix} \mathbf{I} & \mathbf{H} \\ \mathbf{H}' & \mathbf{0} \end{pmatrix} = \begin{matrix} \gamma_{P_{21}} & \gamma_{P_{23}} & \theta_2 & \theta_3 \\ \gamma_{P_{21}} \begin{pmatrix} 1 & 0 & 10^2 & 0 \\ 0 & 1 & \epsilon^{-1} & -\epsilon^{-1} \\ 10^2 & \epsilon^{-1} & 0 & 0 \\ 0 & -\epsilon^{-1} & 0 & 0 \end{pmatrix} \end{matrix}$$

This matrix can be reordered and triangularized as follows:

$$\begin{pmatrix} 1 & 10^2 & 0 & 0 \\ 10^2 & 0 & \epsilon^{-1} & 0 \\ 0 & \epsilon^{-1} & 1 & -\epsilon^{-1} \\ 0 & 0 & -\epsilon^{-1} & 0 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 10^2 & 0 & 0 \\ 0 & -10^4 & \epsilon^{-1} & 0 \\ 0 & 0 & 1 + 10^{-4}\epsilon^{-1} & -\epsilon^{-1} \\ 0 & 0 & 0 & \frac{\epsilon^{-2}}{(1+10^{-4}\epsilon^{-1})} \end{pmatrix}$$

It can be easily verified that with the same 7-digit floating-point precision as above, for $\epsilon = 10^{-8}$, the numerical values for the above triangular matrix can be computed without difficulty, whereas under these circumstances, the gain matrix of the normal equation approach would become singular, as shown above.

13.2.2 Equality Constraints

Consider the following minimization problem:

$$\begin{aligned} \text{Minimize} \quad & J(\mathbf{r}) = \frac{1}{2} \mathbf{r}' \mathbf{r} \\ \text{subject to} \quad & \mathbf{r} = \mathbf{z} - \mathbf{h}(\mathbf{x}) \\ & \mathbf{c}(\mathbf{x}) = \mathbf{0} \end{aligned}$$

The corresponding Lagrangian function is

$$\mathcal{L}(\mathbf{x}, \mathbf{r}, \boldsymbol{\Lambda}) = \frac{1}{2} \mathbf{r}' \mathbf{r} - \boldsymbol{\Lambda}' \mathbf{c}(\mathbf{x}) - \boldsymbol{\Gamma}' (\mathbf{r} - \mathbf{z} + \mathbf{h}(\mathbf{x}))$$

The KKK first order necessary conditions for an optimal solution are expressed by the following augmented tableau (Hachtel tableau):

$$\begin{pmatrix} \mathbf{I} & \mathbf{0} & \mathbf{H}(\mathbf{x}^\nu) \\ \mathbf{0} & \mathbf{0} & \mathbf{C}(\mathbf{x}^\nu) \\ \mathbf{H}'(\mathbf{x}^\nu) & \mathbf{C}'(\mathbf{x}^\nu) & \mathbf{0} \end{pmatrix} \begin{pmatrix} \boldsymbol{\Gamma}^{\nu+1} \\ \boldsymbol{\Lambda}^{\nu+1} \\ \boldsymbol{\Delta} \mathbf{x}^\nu \end{pmatrix} = \begin{pmatrix} \mathbf{r}(\mathbf{x}^\nu) \\ -\mathbf{c}(\mathbf{x}^\nu) \\ \mathbf{0} \end{pmatrix} \tag{13.12}$$

13.3 PETERS WILKINSON METHOD

Consider an $m \times n$ Jacobian matrix $\mathbf{H}(\mathbf{x}^\nu)$ having rank n . The Peters Wilkinson method avoids the direct formation (squaring) of the gain matrix $\mathbf{H}'(\mathbf{x}^\nu)\mathbf{H}(\mathbf{x}^\nu)$ by using the following decomposing of the Jacobian matrix (see Sec. 5.10 in Chap. 5):

$$\mathbf{H} = \mathbf{L}\mathbf{U}$$

where \mathbf{L} is a unit lower trapezoidal $m \times n$ matrix and \mathbf{U} is an upper triangular $m \times n$ matrix (a nonsingular matrix). Since pivoting is usually necessary due to numerical and sparsity considerations, a permutation of \mathbf{H} is in fact factorized:

$$\mathbf{P}_r\mathbf{H}\mathbf{P}_c = \hat{\mathbf{H}} = \mathbf{L}\mathbf{U}$$

where \mathbf{P}_r performs row permutations on \mathbf{H} and \mathbf{P}_c performs column permutations. Peters Wilkinson factorization is shown in Fig. 13.3.

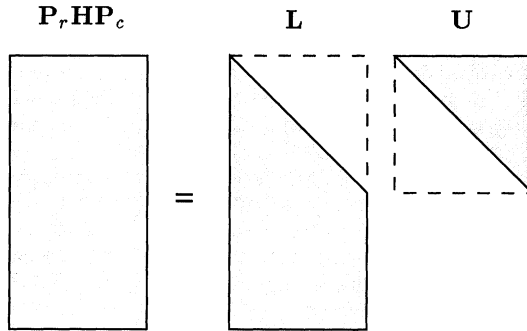


Figure 13.3. Peters Wilkinson factorization.

The normal equation

$$\hat{\mathbf{H}}(\mathbf{x}^\nu)' \hat{\mathbf{H}}(\mathbf{x}^\nu) \Delta \mathbf{x}^\nu = \hat{\mathbf{H}}'(\mathbf{x}^\nu) \mathbf{r}(\mathbf{x}^\nu)$$

can then be rewritten as follows:

$$\mathbf{U}' \mathbf{L}' \mathbf{L} \mathbf{U} \Delta \mathbf{x}^\nu = \mathbf{U}' \mathbf{L}' \mathbf{r}(\mathbf{x}^\nu)$$

Since \mathbf{U} is nonsingular, the result is

$$\mathbf{L}' \mathbf{L} \mathbf{U} \Delta \mathbf{x}^\nu = \mathbf{L}' \mathbf{r}(\mathbf{x}^\nu)$$

This equation can be solved in two stages, as follows:

$$\mathbf{L}' \mathbf{L} \mathbf{y} = \mathbf{L}' \mathbf{r}$$

$$U\Delta x = y$$

The first stage involves the solution of transformed normal equation, where the gain matrix $L'L$ is usually sparser than the original gain matrix (although not necessarily); the second stage involves a simple backward substitution using the triangular factor U .

Example 13.3:

Consider the system in Fig. 13.4 with power flow measurement variances of 10^{-2} , power injection variances of 10^{-4} and a variance of angle-reference pseudo-measurement of 10^{16} . For simplicity, all branch reactances are considered to be 1 p.u.

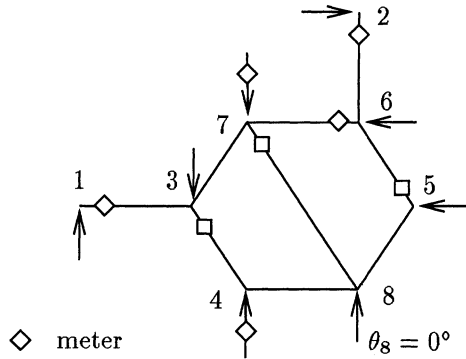


Figure 13.4. 8-bus system.

In this case, the weighted Jacobian matrix P_rHP_c can be written as:

$$\hat{H} = P_rHP_c = \begin{matrix} P_{13}^{(1)} \\ P_{26}^{(2)} \\ P_{56}^{(3)} \\ P_{34}^{(4)} \\ P_4^{(5)} \\ P_{67}^{(6)} \\ P_{78}^{(7)} \\ \theta_8^{(8)} \\ P_7 \end{matrix} \begin{pmatrix} \theta_1^{(1)} & \theta_2^{(2)} & \theta_5^{(3)} & \theta_4^{(4)} & \theta_3^{(5)} & \theta_6^{(6)} & \theta_7^{(7)} & \theta_8^{(8)} \\ 10 & & & & -10 & & & \\ & 10 & & & & & -10 & \\ & & 10 & & & & -10 & \\ & & & -10 & 10 & & & \\ & & & 200 & -100 & & & -100 \\ & & & & & 10 & -10 & \\ & & & & & & 10 & -10 \\ & & & & & & & 10^8 \\ & & & & -100 & -100 & 300 & -100 \end{pmatrix}$$

where the pivoting order is indicated by superscripts. The first pivot is $(P_{13}^1; \theta_1^1)$, the second is $(P_{26}^2; \theta_2^2)$, etc.

$$\mathbf{L} = \begin{matrix} P_{13}^{(1)} \\ P_{26}^{(2)} \\ P_{56}^{(3)} \\ P_{34}^{(4)} \\ P_4^{(5)} \\ P_{67}^{(6)} \\ P_{78}^{(7)} \\ \theta_8^{(8)} \\ P_7 \end{matrix} \begin{pmatrix} \theta_1^{(1)} & \theta_2^{(2)} & \theta_5^{(3)} & \theta_4^{(4)} & \theta_3^{(5)} & \theta_6^{(6)} & \theta_7^{(7)} & \theta_8^{(8)} \\ 1 & & & & & & & \\ & 1 & & & & & & \\ & & 1 & & & & & \\ & & & 1 & & & & \\ & & & & -20 & 1 & & \\ & & & & & & 1 & \\ & & & & & & & 1 \\ & & & & & & & & 1 \\ & & & & & -1 & -10 & 20 & \\ & & & & & & & & & 1 \end{pmatrix}$$

$$\mathbf{U} = \begin{matrix} \theta_1^{(1)} \\ \theta_2^{(2)} \\ \theta_5^{(3)} \\ \theta_4^{(4)} \\ \theta_3^{(5)} \\ \theta_6^{(6)} \\ \theta_7^{(7)} \\ \theta_8^{(8)} \end{matrix} \begin{pmatrix} \theta_1^{(1)} & \theta_2^{(2)} & \theta_5^{(3)} & \theta_4^{(4)} & \theta_3^{(5)} & \theta_6^{(6)} & \theta_7^{(7)} & \theta_8^{(8)} \\ 10 & & & & -10 & & & \\ & 10 & & & & & -10 & \\ & & 10 & & & & -10 & \\ & & & -10 & 10 & & & \\ & & & & 100 & & & -100 \\ & & & & & 10 & -10 & \\ & & & & & & 10 & -10 \\ & & & & & & & & 10^8 \end{pmatrix}$$

The gain matrix of the transformed problem, $\mathbf{L}'\mathbf{L}$, is

$$\mathbf{L}'\mathbf{L} = \begin{matrix} \theta_1^{(1)} \\ \theta_2^{(2)} \\ \theta_5^{(3)} \\ \theta_4^{(4)} \\ \theta_3^{(5)} \\ \theta_6^{(6)} \\ \theta_7^{(7)} \\ \theta_8^{(8)} \end{matrix} \begin{pmatrix} \theta_1^{(1)} & \theta_2^{(2)} & \theta_5^{(3)} & \theta_4^{(4)} & \theta_3^{(5)} & \theta_6^{(6)} & \theta_7^{(7)} & \theta_8^{(8)} \\ 1 & & & & & & & \\ & 1 & & & & & & \\ & & 1 & & & & & \\ & & & 401 & -20 & & & \\ & & & -20 & 2 & 10 & -20 & \\ & & & & 10 & 101 & -200 & \\ & & & & -20 & -200 & 401 & \\ & & & & & & & & 1 \end{pmatrix}$$

13.4 BLOCKED SPARSE TABLEAU

This method is a variation of the Hachtel method discussed above. The measurement vector is partitioned as follows:

$$\mathbf{z} = \begin{pmatrix} \mathbf{z}_b \\ \mathbf{z}_n \end{pmatrix}$$

where \mathbf{z}_b includes both branch flow measurements and nodal voltage measurements, and \mathbf{z}_n contains the nodal injection measurements. The Jacobian matrix is partitioned accordingly, i.e.:

$$\mathbf{H} = \begin{pmatrix} \mathbf{H}_b \\ \mathbf{H}_n \end{pmatrix}$$

The Hachtel tableau in Eq. (13.12) can then be rewritten as follows:

$$\begin{pmatrix} \mathbf{I}_b & \mathbf{0} & \mathbf{0} & \mathbf{H}_b(\mathbf{x}^\nu) \\ \mathbf{0} & \mathbf{I}_n & \mathbf{0} & \mathbf{H}_n(\mathbf{x}^\nu) \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{C}(\mathbf{x}^\nu) \\ \mathbf{H}'_b(\mathbf{x}^\nu) & \mathbf{H}'_n(\mathbf{x}^\nu) & \mathbf{C}'(\mathbf{x}^\nu) & \mathbf{0} \end{pmatrix} \begin{pmatrix} \Gamma_b^{\nu+1} \\ \Gamma_n^{\nu+1} \\ \Lambda^{\nu+1} \\ \Delta \mathbf{x}^\nu \end{pmatrix} = \begin{pmatrix} \mathbf{r}_b(\mathbf{x}^\nu) \\ \mathbf{r}_n(\mathbf{x}^\nu) \\ -\mathbf{c}(\mathbf{x}^\nu) \\ \mathbf{0} \end{pmatrix}$$

Applying Gauss elimination to zeroize the submatrix $\mathbf{H}'_b(\mathbf{x}^\nu)$ results in the following tableau:

$$\begin{pmatrix} \mathbf{I}_n & \mathbf{0} & \mathbf{H}_n(\mathbf{x}^\nu) \\ \mathbf{0} & \mathbf{0} & \mathbf{C}(\mathbf{x}^\nu) \\ \mathbf{H}'_n(\mathbf{x}^\nu) & \mathbf{C}'(\mathbf{x}^\nu) & -\mathbf{H}'_b(\mathbf{x}^\nu)\mathbf{H}_b(\mathbf{x}^\nu) \end{pmatrix} \begin{pmatrix} \Gamma_n^{\nu+1} \\ \Lambda^{\nu+1} \\ \Delta \mathbf{x}^\nu \end{pmatrix} = \begin{pmatrix} \mathbf{r}_n(\mathbf{x}^\nu) \\ -\mathbf{c}(\mathbf{x}^\nu) \\ -\mathbf{H}'_b(\mathbf{x}^\nu)\mathbf{r}_b(\mathbf{x}^\nu) \end{pmatrix}$$

If there is only one injection measurement or constraint per node, this tableau can be further arranged to have the same block structure of the \mathbf{Y} -matrix.

Remarks:

- The use of blocked matrices improves the speed of factorization significantly, since pivot ordering is based on the sparsity of the \mathbf{Y} -matrix, and efficient methods for dealing with this are well known (see Tinney schemes in Chap. 5).
- The selection of pivots based purely on sparsity, however, may lead to singular block pivots (a block with a zero determinant, which, of course, includes the case of a zero scalar pivot). Hence, both the identification of singularities and the modification of the blocking (or of the ordering) scheme is still necessary.
- Notice also that the results provided by numerical observability analysis can be drastically affected if singularity occurs for reasons other than unobservability (topological or numerical), as may be the case with the blocking approach. As a rule, reliable factorization methods are always necessary when indefinite matrices such as those in a Hachtel tableau are used.

Example 13.4:

Consider the system in Fig. 13.4. In the following, first the Hachtel tableau is written as above, and next in the blocked form.

	4	7	1	2	3	4	5	6	7	8		
4	1				H	H					γ_4	r_4
7		0			C			C	C	C	λ_7	$-P_7$
1			G		G						$\Delta\theta_1$	= $\mathbf{H}'_b \mathbf{r}_b$
2				G				G			$\Delta\theta_2$	
3	H	C			G	G					$\Delta\theta_3$	
4	H				G	G					$\Delta\theta_4$	
5							G	G			$\Delta\theta_5$	
6		C		G			G	G	G		$\Delta\theta_6$	
7		C						G	G	G	$\Delta\theta_7$	
8	H	C							G	G	$\Delta\theta_8$	

In this case, there is one injection measurement at node 4 and a single equality constraint corresponding to the junction-node 7 (zero injection). The multiplier corresponding to the residual r_4 is γ_4 , and the multiplier of the equality constraint is λ_7 .

The resulting blocked tableau is as follows:

	1	2	3	4	5	6	7	8		
1	G		G						$\Delta\theta_1$	$\mathbf{h}'_1 \mathbf{r}_b$
2		G				G			$\Delta\theta_2$	$\mathbf{h}'_2 \mathbf{r}_b$
3	G		G	G	H			C	$\Delta\theta_3$	$\mathbf{h}'_3 \mathbf{r}_b$
4			G	G	H				$\Delta\theta_4$	$\mathbf{h}'_4 \mathbf{r}_b$
			H	H	1			H	γ_4	r_4
5					G	G			$\Delta\theta_5$	$\mathbf{h}'_5 \mathbf{r}_b$
6		G			G	G	G	C	$\Delta\theta_6$	$\mathbf{h}'_6 \mathbf{r}_b$
7						G	G	C	$\Delta\theta_7$	$\mathbf{h}'_7 \mathbf{r}_b$
			C			C	C	0	λ_7	$-P_7$
8				H			G	C	$\Delta\theta_8$	$\mathbf{h}'_8 \mathbf{r}_b$

where \mathbf{h}_k is the k th column of the Jacobian matrix \mathbf{H}_b . Note that not all the elements corresponding to the sparsity pattern of the \mathbf{Y} -matrix are filled in. This can, however, happen if more injection measurements and/or equality constraints are added. When all injections are measured, for example, the structure of the blocked tableau is the same as that of the \mathbf{Y} -matrix, with all blocks being 2×2 .

13.5 MIXED PIVOTING FOR INDEFINITE MATRICES

The following algorithm overcomes the non-definiteness of the Hachtel tableau by allowing the use of 2×2 pivots in addition to regular 1×1 pivots.

13.5.1 Basic Algorithm

A 2×2 pivot is used whenever a zero, or very small, pivot is found in the course of factorization using the Tinney 2 scheme.

If the indefinite matrix is nonsingular, it can be proven that:

- If only symmetrical pivoting is used, 1×1 pivots with value zero may occur during factorization;
- Whenever such a zero pivot is found, it is always possible to find a nonsingular 2×2 pivot so that factorization can proceed. In this case, all operations are performed using block arithmetic (see Sec. 5.8, Chap 5).

The factorization algorithm using mixed 1×1 and 2×2 pivots can be expressed as follows:

Algorithm:

For $\nu=1,n$

1. Select the next pivotal row based in the Tinney 2 scheme (the k th row is selected and d_k is the pivot).
2. If $|d_k| \leq \delta$, then:
 - (a) then,
 - i. carry out 1×1 pivoting as usual
 - ii. $\nu \leftarrow \nu + 1$
 - (b) else,
 - i. select a companion row based on the extended Tinney 2 scheme (the i th row is selected and d_i is the corresponding diagonal element)
 - ii. perform 2×2 pivoting to zeroize the elements of the k th and i th columns
 - iii. $\nu \leftarrow \nu + 2$.

Remarks:

- The algorithm above supersedes the usual factorization based on the Tinney 2 scheme, and, as a rule, most of the pivots will be 1×1 .
- When a 2×2 pivot occurs, it usually has either $d_k = 0$ or both $d_k = 0$ and $d_i = 0$, i.e.,

$$\begin{pmatrix} 0 & x \\ x & d_i \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} 0 & x \\ x & 0 \end{pmatrix}$$

- A 2×2 pivot may be used for numerical reasons; if so, $d_k \neq 0$, i.e.,

$$\begin{pmatrix} d_k & x \\ x & d_i \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} d_k & x \\ x & 0 \end{pmatrix}$$

13.5.2 Sparsity Considerations

Consider that the rows k and i have been selected as pivotal, with $d_k = 0$ and $d_i \neq 0$. An upper-bound for the number of fill-ins is given by the following expression:

$$ub = (n_k + n_i - n_{ki} - 2)^2 - (n_i - n_{ki})^2$$

where n_k and n_i are the numbers of nonzero elements in the rows k and i , including the elements in the 2×2 pivot, and n_{ki} is the number of elements that appear in the same column in both rows (coincidences). This expression does not include the fill-ins that may occur in the pivotal rows. (Note that, in the case of a 1×1 pivot, this expression is reduced to $ub = n_k^2$, which is the same as in the Tinney 2 scheme.)

The direct utilization of this formula would require checking the coincidences between the elements in the pivotal row and the elements in all the remaining rows. In practice, a simplified version of the upper bound, which basically ignores the effect of column coincidences, is used. Hence, the companion row is selected based on the least $r_k + r_i$ criterion (extended Tinney 2 scheme).

Remarks: At least two reports of the use of the Hachtel method with mixed 1×1 and 2×2 pivots in conjunction with an extended Tinney 2 scheme can be found in the literature (see Historical Notes and References at the end of the chapter): the first using the Harwell subroutine MA27AD, and other a specialized code.

13.6 ORTHOGONAL TRANSFORMATION APPROACH

Consider an $m \times n$ Jacobian matrix $\mathbf{H}(\mathbf{x}^\nu)$ with rank n . The orthogonal transformation method avoids squaring the gain matrix $\mathbf{H}'(\mathbf{x}^\nu)\mathbf{H}(\mathbf{x}^\nu)$ by using the

following decomposing of the Jacobian matrix (see Subsec. B.2.3 in Appendix B):

$$\mathbf{H} = \mathbf{Q}'\mathbf{U}$$

where \mathbf{Q} is an orthogonal $m \times m$ matrix and \mathbf{U} is an upper trapezoidal $m \times n$ matrix. Since pivoting is usually necessary for sparsity reasons, a permutation of \mathbf{H} can be factorized, i.e.:

$$\mathbf{P}_r\mathbf{H}\mathbf{P}_c = \hat{\mathbf{H}} = \mathbf{Q}'\mathbf{U}$$

where \mathbf{P}_r performs row permutations on \mathbf{H} , and \mathbf{P}_c performs column permutations (see the Peters Wilkinson method above). Factorization is shown in Fig. 13.5.

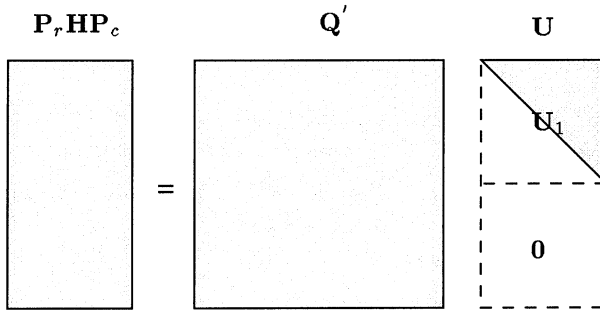


Figure 13.5. Orthogonal factorization of a permuted Jacobian matrix.

13.6.1 Orthogonal Decomposition of Normal Equation

The normal equation

$$\hat{\mathbf{H}}(\mathbf{x}^\nu)' \hat{\mathbf{H}}(\mathbf{x}^\nu) \Delta \mathbf{x}^\nu = \hat{\mathbf{H}}'(\mathbf{x}^\nu) \mathbf{r}(\mathbf{x}^\nu)$$

can be rewritten as follows:

$$\mathbf{U}'\mathbf{Q}\mathbf{Q}'\mathbf{U}\Delta\mathbf{x}^\nu = \mathbf{U}'\mathbf{Q}\mathbf{r}(\mathbf{x}^\nu)$$

Since \mathbf{Q} is orthogonal, i.e. $\mathbf{Q}'\mathbf{Q} = \mathbf{Q}\mathbf{Q}' = \mathbf{I}_m$, the result is

$$\mathbf{U}'\mathbf{U}\Delta\mathbf{x}^\nu = \mathbf{U}'\mathbf{Q}\mathbf{r}(\mathbf{x}^\nu)$$

In view of the partition in Fig. 13.6, this expression can be rewritten as follows:

$$\mathbf{U}'_1\mathbf{U}_1\Delta\mathbf{x}^\nu = \mathbf{U}'_1\mathbf{Q}_1\mathbf{r}(\mathbf{x}^\nu) \tag{13.13}$$

And since \mathbf{U}_1 is nonsingular, the result is the following:

$$\mathbf{U}_1 \Delta \mathbf{x}^\nu = \mathbf{Q}_1 \mathbf{r}(\mathbf{x}^\nu)$$

This equation can be solved in two stages, as follows:

$$\begin{aligned} \mathbf{y}_1 &= \mathbf{Q}_1 \mathbf{r}(\mathbf{x}^\nu) \\ \mathbf{U}_1 \Delta \mathbf{x}^\nu &= \mathbf{y}_1 \end{aligned}$$

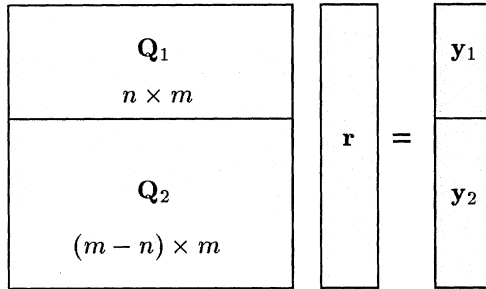


Figure 13.6. Transformation of the independent vector.

Remarks:

- The performance index $J(\Delta \mathbf{x})$ of the linearized least-squares problem can be written as follows:

$$J(\Delta \mathbf{x}) = (\mathbf{r} - \mathbf{H}\Delta \mathbf{x})' (\mathbf{r} - \mathbf{H}\Delta \mathbf{x})$$

Since \mathbf{Q} is orthogonal, this index can be rewritten as follows:

$$J(\Delta \mathbf{x}) = (\mathbf{r} - \mathbf{H}\Delta \mathbf{x})' \mathbf{Q}' \mathbf{Q} (\mathbf{r} - \mathbf{H}\Delta \mathbf{x})$$

Hence,

$$J(\Delta \mathbf{x}) = (\mathbf{Q}\mathbf{r} - \mathbf{Q}\mathbf{H}\Delta \mathbf{x})' (\mathbf{Q}\mathbf{r} - \mathbf{Q}\mathbf{H}\Delta \mathbf{x})$$

In view of the orthogonal factorization, $\mathbf{Q}\mathbf{H} = \mathbf{U}$, the result is as follows:

$$J(\Delta \mathbf{x}) = (\mathbf{Q}\mathbf{r} - \mathbf{U}\Delta \mathbf{x})' (\mathbf{Q}\mathbf{r} - \mathbf{U}\Delta \mathbf{x})$$

Now, considering the partition $\mathbf{y} = (\mathbf{y}_1, \mathbf{y}_2)'$, and the corresponding partitions of \mathbf{U} and \mathbf{Q} , the following is the result:

$$J(\Delta \mathbf{x}) = \left(\begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_1 \end{pmatrix} - \begin{pmatrix} \mathbf{U}_1 \\ \mathbf{0} \end{pmatrix} \Delta \mathbf{x} \right)' \left(\begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_1 \end{pmatrix} - \begin{pmatrix} \mathbf{U}_1 \\ \mathbf{0} \end{pmatrix} \Delta \mathbf{x} \right)$$

Or

$$J(\Delta \mathbf{x}) = (\mathbf{y}_1 - \mathbf{U}_1 \Delta \mathbf{x})' (\mathbf{y}_1 - \mathbf{U}_1 \Delta \mathbf{x}) + \mathbf{y}_2' \mathbf{y}_2$$

And, since $\mathbf{y}_1 = \mathbf{U}_1 \Delta \mathbf{x}$, the final result is

$$J(\Delta \mathbf{x}) = \mathbf{y}_2' \mathbf{y}_2$$

Hence, although \mathbf{y}_2 is not used in the computation of the state estimate correction $\Delta \mathbf{x}$, it is an important by-product of the process, especially since this relationship can be used for testing for bad data (the chi-square hypothesis test).

- The derivation above shows that the Euclidean norm of the estimation residual is invariant under an orthogonal transformation $\mathbf{QH} = \mathbf{U}$. (See Appendix B for an interpretation of orthogonal transformations as rotations or reflections, i.e., transformations that preserve the Euclidean norm.)

Basic Algorithm:

1. Determine optimal pivot order
 - (a) Form the structure of the Jacobian matrix \mathbf{H} and the gain matrix \mathbf{G} .
 - (b) Order the gain matrix \mathbf{G} using the Tinney 2 scheme (Equivalent to determining the permutation matrix \mathbf{P}_c .)
 - (c) Determine the upper triangular factor \mathbf{U} by performing symbolic factorization of the permuted gain matrix $\mathbf{P}_c \mathbf{G} \mathbf{P}_c$.
 - (d) Order the rows of the Jacobian matrix (Equivalent to determining the permutation matrix \mathbf{P}_r .)
2. Perform numerical factorization
 - (a) Form the permuted Jacobian matrix $\hat{\mathbf{H}}$
 - (b) Compute the gain matrix factor \mathbf{U}_1 by processing the rows of $\hat{\mathbf{H}}$ using orthogonal transformations and save the orthogonal matrix \mathbf{Q} in product form (see Appendix B).
3. Update state estimate
 - (a) Compute the measurement residual vector $\mathbf{r} = \mathbf{z} - \mathbf{h}(\mathbf{x})$

(b) Compute transformed residual subvectors \mathbf{y}_1 and \mathbf{y}_2 :

$$\begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{pmatrix} = \mathbf{Q}\mathbf{r}$$

(c) Solve $\mathbf{U}_1\Delta\mathbf{x} = \mathbf{y}_1$ for $\Delta\mathbf{x}$

(d) Update state estimate $\mathbf{x} \leftarrow \mathbf{x} + \Delta\mathbf{x}$

(e) If $|\Delta\mathbf{x}| \leq \epsilon$, terminate

(f) Update iteration count $\nu \leftarrow \nu + 1$. If $\nu > \nu_{max}$, terminate; otherwise, go to Step (2) for honest Newton method, else go to Step(3) (dishonest Newton method).

13.6.2 Semi-normal Equation – A Hybrid Approach

If the orthogonal decomposition $\widehat{\mathbf{H}} = \mathbf{Q}'\mathbf{U}$ is applied only to the left side of the normal equation above, the following results are obtained:

$$\mathbf{U}'\mathbf{Q}\mathbf{Q}'\mathbf{U}\Delta\mathbf{x}^\nu = \widehat{\mathbf{H}}'\mathbf{r}(\mathbf{x}^\nu)$$

Since \mathbf{Q} is orthogonal, and \mathbf{U} is upper trapezoidal, the result is

$$\mathbf{U}'_1\mathbf{U}_1\Delta\mathbf{x}^\nu = \widehat{\mathbf{H}}'\mathbf{r}(\mathbf{x}^\nu) \tag{13.14}$$

This equation can then be solved as in the normal equation approach, except that the triangular factor \mathbf{U}_1 is obtained via orthogonal transformations.

Remarks: Comparing the approximate semi-normal equation given in Eq. (13.14) with the full orthogonal method given in Eq. (13.13), the potential source for errors is in the right-hand side of Eq. (13.13). The corrected semi-normal equation method given in the following is aimed at correcting these errors:

$$\begin{aligned} \mathbf{U}'_1\mathbf{U}_1\Delta\mathbf{x}^\nu &= \widehat{\mathbf{H}}'\mathbf{r}(\mathbf{x}^\nu) \\ \mathbf{r}_c &= \mathbf{r}(\mathbf{x}^\nu) - \widehat{\mathbf{H}}\Delta\mathbf{x}^\nu \\ \mathbf{U}'_1\mathbf{U}_1\Delta\mathbf{x}_c &= \widehat{\mathbf{H}}'\mathbf{r}_c \\ \Delta\mathbf{x}^\nu &\leftarrow \Delta\mathbf{x}^\nu + \Delta\mathbf{x}_c \end{aligned}$$

For linear estimators, it is shown in the literature that this correction, under mild conditions, leads to an estimator as accurate as a full orthogonal method. For nonlinear models such as the one used in this section, the situation is even better, since the above correction can be made together with the next iteration, i.e., the new residuals will contain both the errors due to nonlinearities and those due to the approximations introduced in the right-hand side of Eq. (13.14), and hence the correction step is performed only implicitly in the nonlinear method.

13.6.3 Sparsity Considerations

Consider the factorization of \mathbf{H} with permutations, as above:

$$\mathbf{P}_r \mathbf{H} \mathbf{P}_c = \mathbf{Q}' \mathbf{U}$$

Since the permutation matrices are such that $\mathbf{P} = \mathbf{P}' = \mathbf{P}^{-1}$, the Jacobian matrix can be written as follows:

$$\mathbf{H} = \mathbf{P}_r \mathbf{Q}' \mathbf{U} \mathbf{P}_c$$

The gain matrix can thus be written as

$$\mathbf{H}' \mathbf{H} = \mathbf{P}_c' \mathbf{U}' \mathbf{Q} \mathbf{P}_r' \mathbf{P}_r \mathbf{Q}' \mathbf{U} \mathbf{P}_c$$

In view that $\mathbf{P}_r' \mathbf{P}_r = \mathbf{I}$ and $\mathbf{Q}' \mathbf{Q} = \mathbf{I}$, the result is

$$\mathbf{H}' \mathbf{H} = \mathbf{P}_c' \mathbf{U}' \mathbf{U} \mathbf{P}_c$$

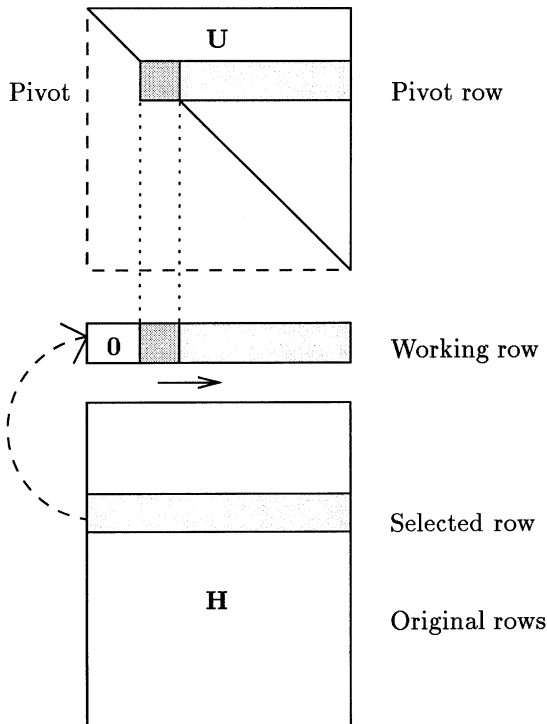


Figure 13.7. Orthogonal factorization of the Jacobian matrix (The elements of the working row are zeroed from left to right.).

Hence, the factor \mathbf{U} is unique, except for the signs of the rows, i.e., the same \mathbf{U} is obtained regardless of the row order.

The fact that only the column order of \mathbf{H} affects the structure of \mathbf{U} is used in the basic algorithm above, where columns and rows of \mathbf{H} are ordered independently. First, the structure of $\mathbf{H}'\mathbf{H}$ is analyzed and a column order is determined that approximately maximizes the sparsity of \mathbf{U} . This can be done forming the structure of $\mathbf{H}'\mathbf{H}$ and performing symbolic factorization, ignoring the numerical values. This requires the use of a dynamic scheme for keeping the current structure of the matrix introducing fill-ins as they appear. Once the near-optimal column order is known, the structure of the gain matrix can be formed using a static storage scheme and symbolic factorization is performed to determine the final structure of \mathbf{U} .

Although row order does not affect the structure of \mathbf{U} , for cases with a wide range of weights, it might be advisable to process the highly weighted rows first. Aside from that, since the gain matrix is positive definite (observability is assumed) and considering that the orthogonal factorization is stable, the determination of the best column/row order is made based only on sparsity considerations.

Also, row order may dramatically affect the number of intermediate fill-ins. Thus a near-optimal row order is sought for minimum factorization effort. Given a column order obtained via the Tinney 2 scheme, rows are processed according to the increasing column index of the right-most non-zero element in each row. This tends to yield a staircase structure. The number of nonzeros in each row is used as a tie-breaker.

13.6.4 Observability Analysis

Extending the results of Chap. 7 to the orthogonal method, it can be shown that when the network is observable, the matrix \mathbf{U} will have the form given in Fig. 13.8(a), i.e., a single reference angle is needed. On the other hand, when the network is not observable, after processing all available measurements

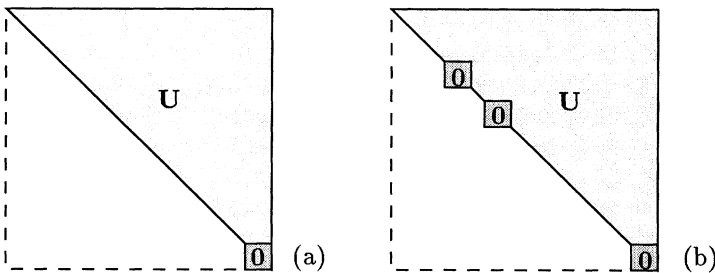


Figure 13.8. (a) Observable network (a single reference angle). (b) Unobservable network (multiple reference angles).

(rows, as indicated in Fig. 13.7), the matrix \mathbf{U} will have the form given in Fig. 13.8(b), where multiple zero pivots are shown. These additional zero pivots indicate the need for more than one θ pseudo-measurement; as a consequence, some of the network branches will read nonzero power flows, which means that they are unobservable (These flows are arbitrary since they are not entirely determined by the measured values.)

Example 13.5:

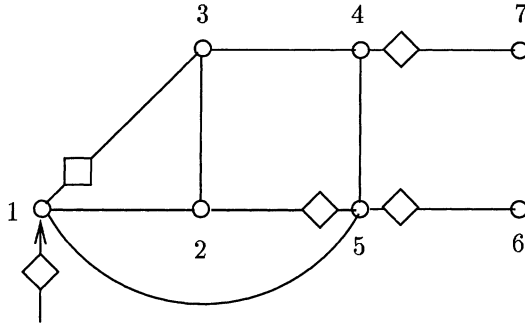


Figure 13.9. System of Example 13-5.

Consider a dc estimation model for the network in Fig. 13.9 with all branch reactances and all weights set at one for simplicity. The Jacobian matrix is as follows:

$$\mathbf{H} = \begin{matrix} P_{13} \\ P_{47} \\ P_{52} \\ P_{56} \\ P_1 \end{matrix} \begin{pmatrix} \theta_1 & \theta_2 & \theta_3 & \theta_4 & \theta_5 & \theta_6 & \theta_7 \\ 1 & & -1 & & & & \\ & & & 1 & & & -1 \\ & -1 & & & 1 & & \\ & & & & 1 & -1 & \\ 3 & -1 & -1 & & -1 & & \end{pmatrix}$$

The orthogonal factorization of \mathbf{H} yields the following permuted factor \mathbf{UP}_c :

$$\mathbf{UP}_c = \begin{matrix} \theta_6 \\ \theta_4 \\ \theta_7 \\ \theta_2 \\ \theta_5 \\ \theta_1 \\ \theta_3 \end{matrix} \begin{pmatrix} \theta_6 & \theta_4 & \theta_7 & \theta_2 & \theta_5 & \theta_1 & \theta_3 \\ 1 & & & & -1 & & \\ & 1 & -1 & & & & \\ & & 0 & & 1 & & \\ & & & \sqrt{2} & & -3/\sqrt{2} & 1/\sqrt{2} \\ & & & & \sqrt{2} & -3/\sqrt{2} & 1/\sqrt{2} \\ & & & & & 1 & -1 \\ & & & & & & 0 \end{pmatrix}$$

Two zero pivots occur at positions (θ_7, θ_7) and (θ_3, θ_3) .

13.7 HISTORICAL NOTES AND REFERENCES

The standard approach to the solution of weighted least-squares state estimation in power systems is the iterative normal equation method. Occasional ill-conditioning has been experienced with this method in connection with the use of widely different weighting factors, the presence of a large number of injection measurements, and the representation of low impedance branches which are incident to regular branches. All these problems are somehow related to the squared form of the gain matrix ($\mathbf{H}'\mathbf{H}$).

The first attempt to “unsquare” the gain matrix was presented in a paper by Aschmoneit, Peterson and Adrian [1977], which proposed the representation of zero injections as equality constraints rather than as pseudo-measurements with relatively high weights. To cope with the indefiniteness of the augmented gain matrix (zero pivots) these authors proposed a delayed pivoting scheme. Alternative approaches to reach a solution based on orthogonal transformation were proposed in two papers by Simões-Costa and Quintana [1982], the first one based on a column-wise Householder transformation and the other on row-wise Givens rotations. An early implementation of Givens rotations in a production grade state estimator was reported by ESCA [1984]. A hybrid approach based on semi-normal equation was then proposed by Monticelli and Wu [1985-86] (For a theoretical discussion about the semi-normal equation and the corrected semi-normal equation, see Björck [1987], where it is shown that under mild conditions the corrected semi-normal equation method can be as accurate as a full orthogonal method.) An efficient ordering scheme to preserve sparsity and minimize the number of intermediate fill-ins, along with a modified Givens rotations method (the 2-multiplication scheme), were developed by Vempati, Slutsker, and Tinney [1991-92]. The incorporation of equality constraints in orthogonal state estimators was described by Van Amerongen [1991]. The method of Peters Wilkinson was applied to power system state estimation by Gu, Clements, Krumpholtz, and Davis [1983].

A further step towards “unsquaring” the gain matrix was made by the estimators based on a Hachtel tableau (Hachtel [1976]), in which not only the equality constraints, but all measurements are included in an augmented gain matrix (Gjelsvik, Aam, and Holten [1985]). This method was further studied by Wu, Liu, and Lun [1988], who extended the normalized residuals approach to Hachtel state estimators. A specialized algorithm for mixed 1×1 and 2×2 pivoting was presented by Machado, Azevedo, and Monticelli [1991] (A formal discussion about the mixed pivot approach can be found in Bunch and Parlett [1971].) The use of blocked matrices was simultaneously suggested by Nucera and Gilles [1991] and Alvarado and Tinney [1991], and an extension of the blocked matrix approach to numerical observability analysis and bad data processing was introduced by Nucera and Gilles [1993]. An overview of numerically robust least-squares methods was presented by Duff and Reid [1976] and comparative studies of numerically robust estimators for power networks was published by Holten, Gjelsvik, Aam, Wu and Liu [1988].

13.8 PROBLEMS

- 1. Compute the condition number of the gain matrix in Examples 13-1 and 13.2.
- 2. Consider the two situations depicted in Fig. 13.10 in which, for simplicity, all reactance and all measurement variances are set at one. Compute the condition number of the gain matrix for the normal equation approach in both cases. Discuss the impact of presence of injection measurements on the numerical condition of the gain matrix. Use a dc model.

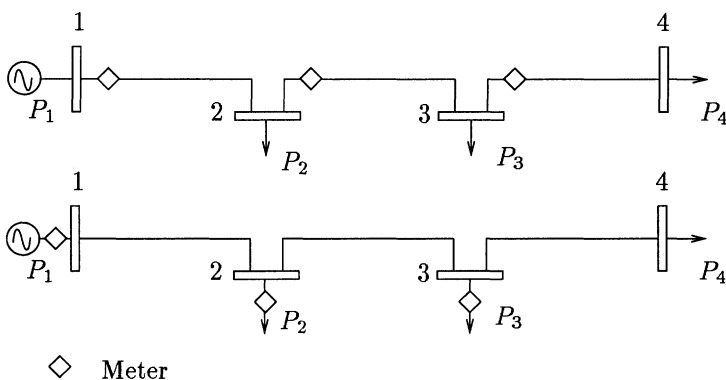


Figure 13.10. Four-bus systems used in Prob. 2.

- 3. Show how Eq. (13.5) can be obtained from Eq. (13.12).
- 4. Form the tableau in Eq. (13.12) for the system discussed in Example 13.1. Compute the condition number of this matrix and compare it with that of the gain matrix in Eq. 13.2 for the same situation.
- 5. Show that the tableau in Eq. (13.12) gives the KKK necessary conditions for the following minimization problem:

$$\begin{aligned}
 &\text{Minimize} && J(\mathbf{r}) = \frac{1}{2} \mathbf{r}' \mathbf{r} \\
 &\text{subject to} && \mathbf{r} = \mathbf{z} - \mathbf{h}(\mathbf{x}) \\
 &&& \mathbf{c}(\mathbf{x}) = \mathbf{0}
 \end{aligned}$$

- 6. Apply Peters Wilkinson decomposition to the problem in Example 13.2.
- 7. For the system in Example 13.3, form the gain matrices $\widehat{\mathbf{H}}' \widehat{\mathbf{H}}$ and $\mathbf{L}' \mathbf{L}$, where \mathbf{L} is a unit trapezoidal matrix, and obtain the respective \mathbf{LDU} factors. Compare the sparsity of the triangular factors.

- 8. Extend the blocked tableau in Example 13.4 to the ac model (see Sec. 5.8 in Chap 5). Consider that all measurements are performed in pairs, including the voltage measurement at Node 1.
- 9. Repeat Example 13.4, considering now that Nodes 4 and 5 represent junctions (zero injection).
- 10. Determine the column permutation matrix used in Example 13.5.
- 11. Determine the unobservable branches and the discardable measurements (if any) for the network in Fig. 13.9.
- 12. Calculate the factors \mathbf{Q}' and \mathbf{U} for the Jacobian matrix in Example 13.3 considering: (a) column order according to Tinney 2 scheme, (b) row order according to the increasing column index of the right-most non-zero element in each row, and (c) tie-breaker according to the number of nonzeros in each row.

References

- Aschmoneit, F.C., Peterson, N. M., and Adrian, E.C., "State estimation with equality constraints", PICA'77 Conf. Proc., pp. 427-430, May 1977.
- Alvarado, F.L. and Tinney, W.F., "State estimation augmented blocked matrices", IEEE Trans. Power Syst., Vol. 5, No. 3, pp. 911-921, Aug. 1990.
- Björck, Å, "Stability analysis of the method of seminormal equations for linear least squares problems", Linear Algebra and its Applications. Vol 88-89, pp. 31-48, 1987.
- Bunch, J.R. and Parlett, B.N., "Direct methods for solving symmetric indefinite systems of linear equations", SIAM J. Numer. Anal., Vol. 8, No. 4, pp. 639-655, Dec. 1971.
- Contribution to power system state estimation and transient stability analysis*, Prepared by ESCA Corporation for the U.S. Department of Energy, Contract DOE/ET/29362-1, Feb. 1984.
- Duff, I.F. and Reid, J.K., "A comparison of some methods for the solution of sparse overdetermined systems of linear equations", J. Inst. Maths Applics, Vol. 17, pp.267-280, 1976.
- Gjelsvik, A., Aam, S., and Holten, L., "Hachtel's augmented matrix method: A rapid method for improving numerical stability in power system state estimation", IEEE Trans. Power App. Syst., Vol. 104, pp. 2987-2993, Nov. 1985.
- Gu, J.W., Clements, K.A., Krumpholtz, G.R., and Davis, P.W., "The solution of ill-conditioned power system state estimation problems via the method of Peters and Wilkinson", IEEE Trans. Power App. Syst., Vol. 104, pp. 3473-3480, Oct. 1983.
- Hachtel, G.D., "The sparse tableau approach to finite element assembly", Sparse Matrix Computations, pp. 349-363, 1976.

- Holten, L., Gjelsvik, A., Aam, S., Wu, F.F., and Liu, W.H.E., "Comparison of different methods for state estimation", *IEEE Trans. Power Syst.*, Vol. 3, No. 4, pp. 1798-1806, Nov. 1988.
- Machado, P.A., Azevedo, G.P., and Monticelli, A., "A mixed pivoting approach to the factorization of indefinite matrices in power system state estimation", *IEEE Trans. Power Syst.*, Vol. 6, No. 2, pp. 676-682, May 1991.
- Monticelli, A., Murari, C.A.F., and Wu F.F., "A Hybrid state estimator: Solving normal equations by orthogonal transformations", *IEEE Trans. Power Syst.*, Vol. 105, pp. 3460-3468, Dec. 1985.
- Monticelli, A. and Wu, F.F., "Observability analysis for orthogonal transformation based state estimation", *IEEE Trans. Power Syst.*, Vol. 1, No. 1, pp. 201-206, Feb. 1986.
- Nucera, R.R. and Gilles, M.L., "A blocked sparse-matrix formulation for the solution of equality-constrained state estimation", *IEEE Trans. Power Syst.*, Vol. 6, No. 1, pp. 214-224, Feb. 1991.
- Nucera, R.R., Brandwajn, V., and Gilles, M.L., "Observability analysis and bad data - Analysis using augmented blocked matrices", *IEEE Trans. Power Syst.*, Vol. 8, No. 2, pp. 426-433, May 1993.
- Quintana, V.H., Simões-Costa, A., and Mier, M., "Bad data detection and identification techniques using estimation orthogonal methods", *IEEE Trans Power Appl. Syst.*, Vol. 101, No. 9, pp. 3356-3364, Sept., 1982.
- Simões-Costa, A. and Quintana, V.H., "A robust numerical technique for power systems state estimation", *IEEE Trans Power Appl. Syst.*, Vol. 100, No. 2, pp. 691-698, Feb., 1981.
- Simões-Costa, A. and Quintana, V.H., "An orthogonal row processing algorithm for power system sequential state estimation", *IEEE Trans Power Appl. Syst.*, Vol. 100, No. 8, pp. 3791-3800, Aug., 1981.
- Vempati, N., Slutsker, I.W. and Tinney, W.F., "Enhancements to givens rotations for power system state estimation", *IEEE Trans. on Power System*, Vol. 6, No. 2, pp 842-849, May, 1991.
- Vempati, N., Slutsker, I.W., and Tinney, W. F., "Orthogonal sparse vector methods", *IEEE Trans. Power Syst.*, Vol. 7, No. 2, pp. 926-932, May 1992.
- Van Amerongen, R.A.M., "On the exact incorporation of virtual measurements in orthogonal-transformation based state estimation procedures", *Int. J. Elec. Power*, Vol. 13, No. 3, pp. 167-174, Jun. 1991.
- Wu, F.F., Liu, E.H.E, and Lun, S.M., "Observability analysis and bad data-processing for state estimation with equality constraints", *IEEE Trans. Power Syst.*, Vol. 3, No. 2, pp. 541-578, May 1988.

Appendix A

Statistical Properties of Estimated Quantities

This appendix reviews the main statistics used for the processing of bad data.

A.1 DISTRIBUTION OF STATE ESTIMATE

Consider the linear state estimation model,

$$\mathbf{z} = \mathbf{H} \mathbf{x} + \mathbf{e}$$

where \mathbf{x} is the n vector of the true states, \mathbf{z} is the m vector of measurements, \mathbf{H} is the $m \times n$ Jacobian matrix, \mathbf{e} is the m vector of errors, m is number of measurements, and n is the number of state variables.

The estimate $\hat{\mathbf{x}}$ is given by

$$\hat{\mathbf{x}} = \mathbf{G}^{-1} \mathbf{H}' \mathbf{R}_z^{-1} \mathbf{z}$$

where \mathbf{R}_z^{-1} is the diagonal matrix of measurement error variances.

In the absence of gross errors, the elements of \mathbf{z} are normal and independent with variances of σ_i^2 . Thus, given the linear relation between $\hat{\mathbf{x}}$ and \mathbf{z} above, the state estimate $\hat{\mathbf{x}}$ follows an n -variate normal distribution, with the following probability density function:

$$f(\hat{\mathbf{x}}) = (2\pi)^{-\frac{n}{2}} |\mathbf{R}_{\hat{\mathbf{x}}}|^{-\frac{1}{2}} e^{-\frac{1}{2} [(\hat{\mathbf{x}}-\mathbf{x})' \mathbf{R}_{\hat{\mathbf{x}}}^{-1} (\hat{\mathbf{x}}-\mathbf{x})]}$$

where

$$\mathbf{R}_{\hat{\mathbf{x}}} = \mathbf{G}^{-1} = (\mathbf{H}' \mathbf{R}_z^{-1} \mathbf{H})^{-1}$$

and $|\mathbf{R}_{\hat{\mathbf{x}}}|$ denotes the determinant of the covariance matrix of the state estimate.

A.2 RANK OF WEIGHTED SENSITIVITY MATRIX

A.2.1 Eigenvalues and Eigenvectors

Let \mathbf{A} be an $m \times m$ matrix. A nonzero vector \mathbf{v}_i is an eigenvector of \mathbf{A} if $\mathbf{A}\mathbf{v}_i$ is a multiple of \mathbf{v}_i , that is

$$\mathbf{A}\mathbf{v}_i = \lambda_i \mathbf{v}_i$$

where the scalar λ_i is the eigenvalue of \mathbf{A} associated with \mathbf{v}_i . Let \mathbf{C} be a nonsingular matrix. The matrix \mathbf{B} , defined as

$$\mathbf{B} = \mathbf{C}^{-1} \mathbf{A} \mathbf{C}$$

has the same eigenvalues as the original matrix \mathbf{A} ; and this transformation is called similarity transformation: if $\mathbf{A}\mathbf{v}_i = \lambda_i \mathbf{v}_i$, then $\mathbf{C}^{-1} \mathbf{A} \mathbf{C} \mathbf{C}^{-1} \mathbf{v}_i = \lambda_i \mathbf{C}^{-1} \mathbf{v}_i$, i.e., λ_i is also an eigenvalue of \mathbf{B} ; the corresponding eigenvector is $\mathbf{C}^{-1} \mathbf{v}_i$.

Let $\mathbf{diag}(\lambda)$ denote a diagonal matrix whose elements are the eigenvalues λ_i , $i = 1, \dots, m$, and \mathbf{V} denote a matrix whose columns are the corresponding eigenvectors \mathbf{v}_i . If matrix \mathbf{A} is real and symmetric, all the eigenvalues are real and it is always possible to find a set of independent (orthogonal) eigenvectors, i.e., it is possible to form a matrix \mathbf{V} such that $\mathbf{V}^{-1} = \mathbf{V}'$. Under these conditions, the following is obtained:

$$\mathbf{A}\mathbf{V} = \mathbf{V} \mathbf{diag}(\lambda)$$

$$\mathbf{V}' \mathbf{A} \mathbf{V} = \mathbf{diag}(\lambda)$$

Matrix \mathbf{A} has thus been diagonalized by a similarity transformation in which the transformation matrix is formed by a set of orthogonal eigenvectors. This expression makes it possible to write the matrix \mathbf{A} in the following factorized form:

$$\mathbf{A} = \mathbf{V} \mathbf{diag}(\lambda) \mathbf{V}'$$

A.2.2 Weighted Sensitivity Matrix

Let $\mathbf{r}^w = \mathbf{R}_z^{-1/2} \hat{\mathbf{r}}$ denote the vector of weighted residuals with elements $r_i^w = \sigma_i^{-1} \hat{r}_i$, $i = 1, \dots, m$, where m is the number of measurements. In Chap. 8 it is shown that the vector of weighted residuals \mathbf{r}^w is a linear function of the weighted vector of measurement errors $\mathbf{e}^w = \mathbf{R}_z^{-1/2} \mathbf{e}$, i.e., $\mathbf{r}^w = \mathbf{S}^w \mathbf{e}^w$, where $\mathbf{S}^w = \mathbf{R}_z^{-1/2} \mathbf{S} \mathbf{R}_z^{1/2}$ is the weighted sensitivity matrix.

A.2.3 Trace of Covariance Matrix of Measurement Estimates

The covariance matrix of the weighted measurement estimates is given by

$$\mathbf{R}_{\hat{\mathbf{z}}^w} = \mathbf{H}^w ((\mathbf{H}^w)' \mathbf{H}^w)^{-1} (\mathbf{H}^w)'$$

Let \mathbf{A} and \mathbf{B} be two rectangular matrices with appropriate dimensions. Since $tr\{\mathbf{AB}\} = tr\{\mathbf{BA}\}$, the trace of covariance matrix $\mathbf{R}_{\hat{\mathbf{z}}^w}$ can be written as

$$\begin{aligned} tr\{\mathbf{R}_{\hat{\mathbf{z}}^w}\} &= tr\{((\mathbf{H}^w)' \mathbf{H}^w)^{-1} (\mathbf{H}^w)' \mathbf{H}^w\} \\ &= tr\{\mathbf{I}_n\} \\ &= n \end{aligned}$$

where n is the number of state variables.

A.2.4 Diagonalization of the Weighted Sensitivity Matrix

Matrix \mathbf{S}^w can be factorized as

$$\mathbf{S}^w = \mathbf{V} \mathbf{diag}(\lambda) \mathbf{V}'$$

Since \mathbf{S}^w is idempotent, this yields

$$\mathbf{V} \mathbf{diag}(\lambda) \mathbf{V}' = \mathbf{V} (\mathbf{diag}(\lambda))^2 \mathbf{V}'$$

This implies that $\mathbf{diag}(\lambda) = \mathbf{diag}(\lambda)^2$, that is, $\lambda_i = \lambda_i^2$, for $i = 1, \dots, m$. This is true only when the eigenvalues of \mathbf{S}^w are either 0 or 1. Hence, the trace of the sensitivity matrix \mathbf{S}^w is

$$\begin{aligned} tr\{\mathbf{S}^w\} &= tr\{\mathbf{V} \mathbf{diag}(\lambda) \mathbf{V}'\} \\ &= tr\{\mathbf{diag}(\lambda) \mathbf{V}' \mathbf{V}\} \\ &= tr\{\mathbf{diag}(\lambda)\} \end{aligned}$$

On the other hand, the trace of \mathbf{S}^w can also be expressed as

$$\begin{aligned} tr\{\mathbf{S}^w\} &= tr\{\mathbf{I}_m - \mathbf{R}_{\hat{\mathbf{z}}^w}\} \\ &= m - n \end{aligned}$$

Hence, $tr\{\mathbf{diag}(\lambda)\} = m - n$. Since the eigenvalues of \mathbf{S}^w are either 0 or 1, the matrix $\mathbf{diag}(\lambda)$ assumes the form

$$\mathbf{diag}(\lambda) = \begin{pmatrix} \mathbf{I}_{m-n} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}$$

where \mathbf{I}_{m-n} is an $(m-n) \times (m-n)$ unit matrix. Hence the rank of the weighted sensitivity matrix \mathbf{S}^w is $m - n$.

A.3 THE χ^2 PROPERTY OF THE LEAST SQUARES INDEX

In this section, the performance index

$$J(\hat{\mathbf{x}}) = (\mathbf{r}^w)' \mathbf{r}^w$$

is shown to follow a χ_{m-n}^2 distribution, i.e., a chi-square distribution with $m-n$ degrees of freedom.

A.3.1 The $J(\hat{\mathbf{x}})$ Performance Index

Since the sensitivity matrix \mathbf{S}^w is symmetric and idempotent, i.e., $\mathbf{S}^w = (\mathbf{S}^w)'$ and $\mathbf{S}^w \mathbf{S}^w = \mathbf{S}^w$, the performance index can be written as

$$\begin{aligned} J(\hat{\mathbf{x}}) &= (\mathbf{r}^w)' \mathbf{r}^w \\ &= (\mathbf{e}^w)' \mathbf{S}^w \mathbf{e}^w \end{aligned}$$

A.3.2 Orthogonal Transformation

The performance index $J(\hat{\mathbf{x}})$ can be rewritten in terms of the orthogonal transformation discussed above, i.e.,

$$\begin{aligned} J(\hat{\mathbf{x}}) &= (\mathbf{e}^w)' \mathbf{S}^w \mathbf{e}^w \\ &= (\mathbf{e}^w)' \mathbf{V} \mathbf{diag}(\lambda) \mathbf{V}' \mathbf{e}^w \end{aligned}$$

Let \mathbf{d} denote the transformed vector $\mathbf{V}' \mathbf{e}^w$. This vector can be factorized as,

$$\mathbf{d} = \begin{pmatrix} \mathbf{d}_r \\ \mathbf{d}_o \end{pmatrix}$$

where \mathbf{d}_r is a $m-n$ vector (corresponding to $m-n$ redundant measurements) and \mathbf{d}_o is an n vector (corresponding to a set of measurements that makes the system minimally observable).

Hence, the performance index can be rewritten as,

$$\begin{aligned} J(\hat{\mathbf{x}}) &= \mathbf{d}' \mathbf{diag}(\lambda) \mathbf{d} \\ &= (\mathbf{d}'_r \quad \mathbf{d}'_o) \begin{pmatrix} \mathbf{I}_{m-n} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{d}_r \\ \mathbf{d}_o \end{pmatrix} \\ &= \mathbf{d}'_r \mathbf{d}_r \end{aligned}$$

A.3.3 Covariance Matrix of Transformed Residuals

Now, the elements of \mathbf{d}_r can be shown to be uncorrelated, i.e., we show that $Cov\{d_i, d_j\} = 0$, for $i \neq j$.

$$\begin{aligned}
 \text{Cov}\{\mathbf{d}\} &= E\{\mathbf{d} \mathbf{d}'\} \\
 &= E\{\mathbf{V}' \mathbf{e}^w (\mathbf{e}^w)' \mathbf{V}\} \\
 &= \mathbf{V}' E\{\mathbf{e}^w (\mathbf{e}^w)'\} \mathbf{V} \\
 &= \mathbf{V}' \mathbf{I}_m \mathbf{V} \\
 &= \mathbf{I}_m
 \end{aligned}$$

A.3.4 The χ^2_{m-n} Distribution

Since each weighted residual follows an $N(0, 1)$ distribution, the transformed residual vector \mathbf{d} is multivariate normal. From the results above, it follows that the elements of \mathbf{d}_r are normal, with mean zero and unit variance, as well as uncorrelated. Hence, the performance index

$$J(\hat{\mathbf{x}}) = \mathbf{d}'_r \mathbf{d}_r = \sum_{i=1}^{m-n} d_i^2$$

follows a χ^2_{m-n} distribution (chi-square with $m - n$ degrees of freedom).

A.4 TESTING EQUALITY CONSTRAINT HYPOTHESES

Consider the null hypothesis $\mathbf{A}\hat{\mathbf{x}} = \mathbf{a}$, where $\hat{\mathbf{x}}$ is the n vector of state estimates, \mathbf{a} is a p vector and \mathbf{A} is a $p \times n$ matrix, as discussed in Chapter 9. The error vector $\mathbf{A}\hat{\mathbf{x}} - \mathbf{a}$ is p -variate normal with mean zero and covariance given by

$$\text{Cov}\{\mathbf{A}\hat{\mathbf{x}} - \mathbf{a}\} = \mathbf{A} (\mathbf{H}\mathbf{W}\mathbf{H}')^{-1} \mathbf{A}'$$

The performance index

$$J(\mathbf{A}\hat{\mathbf{x}} - \mathbf{a}) = (\mathbf{A}\hat{\mathbf{x}} - \mathbf{a})' (\text{Cov}\{\mathbf{A}\hat{\mathbf{x}} - \mathbf{a}\})^{-1} (\mathbf{A}\hat{\mathbf{x}} - \mathbf{a})$$

can be shown to follow a χ^2_p distribution. The inverse of the above covariance matrix can be factorized as

$$(\text{Cov}\{\mathbf{A}\hat{\mathbf{x}} - \mathbf{a}\})^{-1} = \mathbf{B}' \text{diag}(\lambda) \mathbf{B}$$

where \mathbf{B} is orthogonal (i.e., $\mathbf{B}\mathbf{B}' = \mathbf{I}_p$), and $\text{diag}(\lambda)$ is the diagonal matrix of the p eigenvalues of the factorized matrix. Let \mathbf{b} denote the transformed vector $\text{diag}(\lambda)^{1/2} \mathbf{B}(\mathbf{A}\hat{\mathbf{x}} - \mathbf{a})$, where $\text{diag}(\lambda)$ is an $p \times p$ diagonal matrix with the eigenvalues of matrix $(\text{Cov}\{\mathbf{A}\hat{\mathbf{x}} - \mathbf{a}\})^{-1}$. The performance index $J(\mathbf{A}\hat{\mathbf{x}} - \mathbf{a})$ can thus be written as

$$J(\mathbf{A}\hat{\mathbf{x}} - \mathbf{a}) = \mathbf{b}' \mathbf{b}$$

The variance of \mathbf{b} is given by

$$\begin{aligned}
\text{Var}\{\mathbf{b}\} &= E\{\mathbf{b}\mathbf{b}'\} \\
&= \text{diag}(\lambda)^{1/2} \mathbf{B} E\{(\mathbf{A}\mathbf{x} - \mathbf{a})(\mathbf{A}\mathbf{x} - \mathbf{a})'\} \mathbf{B}' \text{diag}(\lambda)^{1/2} \\
&= \text{diag}(\lambda)^{1/2} \mathbf{B} \text{Cov}\{\mathbf{A}\hat{\mathbf{x}} - \mathbf{a}\} \mathbf{B}' \text{diag}(\lambda)^{1/2} \\
&= \text{diag}(\lambda)^{1/2} \mathbf{B} (\mathbf{B}' \text{diag}(\lambda) \mathbf{B})^{-1} \mathbf{B}' \text{diag}(\lambda)^{1/2} \\
&= \text{diag}(\lambda)^{1/2} \mathbf{B} \mathbf{B}' \text{diag}(\lambda^{-1}) \mathbf{B} \mathbf{B}' \text{diag}(\lambda)^{1/2} \\
&= \mathbf{I}_p
\end{aligned}$$

Hence, the elements of \mathbf{b} are normal and independent, with mean zero and unit variance. This implies that the statistic $J(\mathbf{A}\hat{\mathbf{x}} - \mathbf{a})$ follows a χ_p^2 distribution.

A.5 HISTORICAL NOTES AND REFERENCES

The use of the residual sensitivity matrix for the detection and identification of bad data and a formula for performing correction on the index $J(\hat{\mathbf{x}})$ when a set of pre-specified suspect data is removed were proposed by Ma Zhi-quiang [1981]. The special properties of the residual sensitivity matrix (for example, $\text{rank}(\mathbf{S}) = m - n$) were investigated by Xiang, Wang, and Yu [1981,82,84] and a method for processing multiple bad data was developed. The $J(\hat{\mathbf{x}})$ index has been shown to become less sensitive to spurious measurements of constant magnitude as the system size increases Monticelli and Garcia [1983]. A review of bad data processing techniques, including the $J(\hat{\mathbf{x}})$ -test, was made by Bose and Clements [1987].

References

- Bose, A. and Clements, K.A., "Real-time modeling of power networks", IEEE Proc., Special Issue on Computers in Power System Operations, vol. 75, No. 12, pp 1607-1622, Dec. 1987.
- Ma Zhi-quiang, "Bad data reestimation-identification using residual sensitivity matrix", Proc. 7th Power System Computation Conf., PSCC, pp. 1056-1050, Lausanne, July 1981.
- Monticelli, A. and Garcia, A., "Reliable bad data processing for real time state estimation", IEEE Trans. Power App. and Syst., Vol. 102, No. 5, pp. 1126-1139, May 1983.
- Xiang Nian-De, Wang Shi-Ying, and Yu Er-Keng, "Estimation and identification of multiple bad data in power system state estimation", Proc. 7th Power System Computation Conf, PSCC, pp. 1061-1065, Lausanne, July 1981.
- Xiang Nian-De, Wang Shi-Ying, and Yu Er-Keng, "A new approach for detection and identification of multiple bad data in power system state estimation", IEEE Trans. Power App. and Syst., Vol. 101, No. 2, pp. 454-462, Feb. 1982.

Appendix B

Givens Rotation

This appendix discusses both standard and fast Givens rotations.

B.1 ORTHOGONAL MATRICES

A $n \times n$ matrix \mathbf{Q} is said to be orthogonal if:

$$\mathbf{Q}\mathbf{Q}' = \mathbf{I}_n$$

where \mathbf{I}_n is the identity matrix. This means that \mathbf{Q} has an inverse and

$$\mathbf{Q}^{-1} = \mathbf{Q}'$$

Since

$$\mathbf{Q}^{-1}\mathbf{Q} = \mathbf{Q}\mathbf{Q}^{-1} = \mathbf{I}_n$$

then

$$\mathbf{Q}\mathbf{Q}' = \mathbf{Q}'\mathbf{Q} = \mathbf{I}_n$$

B.1.1 2×2 Orthogonal Matrices

Let \mathbf{Q} be a 2×2 orthogonal matrix

$$\mathbf{Q} = \begin{pmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{pmatrix}$$

Since $\mathbf{Q}\mathbf{Q}' = \mathbf{Q}'\mathbf{Q} = \mathbf{I}_n$, then

$$q_{11}^2 + q_{12}^2 = q_{11}^2 + q_{21}^2 \rightarrow q_{12} = \pm q_{21}$$

$$q_{11}q_{12} + q_{21}q_{22}^2 \rightarrow q_{11} = \mp q_{22}$$

And since $q_{11}^2 + q_{22}^2 = 1$ it can be assumed that $q_{11} = \cos\theta$ for a given θ without any loss of generality.

It can thus be concluded that the orthogonal \mathbf{Q} matrix can assume either one of the following forms:

$$\mathbf{Q} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$$

or

$$\mathbf{Q} = \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}$$

B.1.2 Rotations and Reflections

Consider the following linear transformation

$$\begin{pmatrix} \hat{x}_1 \\ \hat{x}_2 \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \quad (\text{B.1})$$

In complex form, it can be expressed as the following:

$$\begin{aligned} \hat{x}_1 + j\hat{x}_2 &= (x_1 \cos\theta + x_2 \sin\theta) + j(-x_1 \sin\theta + x_2 \cos\theta) \\ \hat{x}_1 + j\hat{x}_2 &= (x_1 + jx_2)(\cos\theta - jsin\theta) \end{aligned}$$

And in polar form:

$$\hat{r}e^{j\hat{\phi}} = re^{j\phi} e^{-j\alpha}$$

where $\hat{r}e^{j\hat{\phi}} = \hat{x}_1 + j\hat{x}_2$, $re^{j\phi} = x_1 + jx_2$, and $e^{-j\alpha} = \cos\theta - jsin\theta$.

Figure B.1 shows that transformation (B.1) can be interpreted as a rotation. Hence, the orthogonal matrix

$$\mathbf{Q} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \quad (\text{B.2})$$

is called a rotation matrix, or a rotator.

Consider now the linear transformation

$$\begin{pmatrix} \hat{x}_1 \\ \hat{x}_2 \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \quad (\text{B.3})$$

In complex form we have

$$\hat{x}_1 + j\hat{x}_2 = (x_1 \cos\theta + x_2 \sin\theta) + j(x_1 \sin\theta - x_2 \cos\theta)$$

Taking the complex conjugate results

$$\hat{x}_1 - j\hat{x}_2 = (x_1 \cos\theta + x_2 \sin\theta) - j(x_1 \sin\theta - x_2 \cos\theta)$$

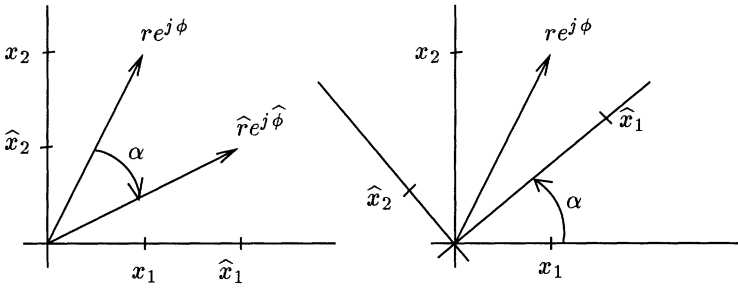


Figure B.1. Rotation.

$$\hat{x}_1 - j\hat{x}_2 = (x_1 + jx_2)(\cos\theta - jsin\theta)$$

In polar form this gives

$$\hat{r}e^{-j\hat{\phi}} = re^{j\phi} e^{-j\alpha}$$

where $\hat{r}e^{j\hat{\phi}} = \hat{x}_1 + j\hat{x}_2$, $re^{j\phi} = x_1 + jx_2$, and $e^{-j\alpha} = \cos\theta - jsin\theta$.

Figure B.2 shows that transformation (B.3) can be interpreted as a reflection. Hence the orthogonal matrix

$$\mathbf{Q} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$$

is called a reflection matrix, or a reflector.

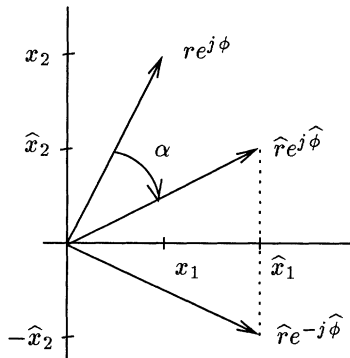


Figure B.2. Reflection.

B.2 GIVENS ROTATIONS

B.2.1 Standard Givens Rotations

If $c = \cos\theta$ and $s = \sin\theta$, the rotator defined in (B.2) can be applied to zeroize the second element of a vector $(a \ b)'$:

$$\begin{pmatrix} c & s \\ -s & c \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \hat{a} \\ 0 \end{pmatrix} \quad (\text{B.4})$$

where

$$\hat{a} = (a^2 + b^2)^{1/2} \quad \text{and} \quad -sa + cb = 0$$

Then

$$c = a/(a^2 + b^2)^{1/2} \quad \text{and} \quad s = b/(a^2 + b^2)^{1/2} \quad (\text{B.5})$$

If \mathbf{A} be a 2-row matrix

$$\mathbf{A} = \begin{pmatrix} a_1 & a_2 & a_3 & \dots & a_n \\ b_1 & b_2 & b_3 & \dots & b_n \end{pmatrix}$$

The problem is to zeroize the first element of the second row to give the following:

$$\hat{\mathbf{A}} = \begin{pmatrix} \hat{a}_1 & \hat{a}_2 & \hat{a}_3 & \dots & \hat{a}_n \\ 0 & \hat{b}_2 & \hat{b}_3 & \dots & \hat{b}_n \end{pmatrix}$$

This transformation can be performed using Eq. (B.4), as for the vector case discussed above. The difference is that now elements a_j and b_j , for $j = 2, 3, \dots, n$, must be updated by:

$$\hat{a}_j = ca_j + sb_j \quad \text{and} \quad \hat{b}_j = -sa_j + cb_j \quad (\text{B.6})$$

Hence, the transformed matrix $\hat{\mathbf{A}}$ is determined in two steps: first, c and s are evaluated from Eq. (B.5); second, elements \hat{a}_j and \hat{b}_j are calculated using Eq. (B.6). Four multiplications and two additions are required for each column j , for $j = 3, 4, \dots, n$ (4-multiplication scheme).

Example B.1:

Consider the matrix \mathbf{A} taken from Example 2.1, Chap. 2:

$$\begin{pmatrix} 1 & -1 & -1 \\ -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

This matrix can be put into upper triangular form by a sequence of three Givens rotations:

First rotation: zeroizes element b_{21}

$$\mathbf{Q}^{(1)}\mathbf{A}^{(0)} = \mathbf{A}^{(1)}$$

$$\begin{pmatrix} c & s & 0 & 0 \\ -s & c & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 & -1 \\ -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \sqrt{2} & -\sqrt{2}/2 & -\sqrt{2}/2 \\ 0 & -\sqrt{2}/2 & -\sqrt{2}/2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

where

$$c = \frac{a_{11}^{(0)}}{\sqrt{(a_{11}^{(0)})^2 + (a_{21}^{(0)})^2}} = \frac{\sqrt{2}}{2} \quad \text{and} \quad s = \frac{a_{21}^{(0)}}{\sqrt{(a_{11}^{(0)})^2 + (a_{21}^{(0)})^2}} = -\frac{\sqrt{2}}{2}$$

Second rotation: zeroizes element b_{32}

$$\mathbf{Q}^{(2)}\mathbf{A}^{(1)} = \mathbf{A}^{(2)}$$

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & c & s & 0 \\ 0 & -s & c & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \sqrt{2} & -\sqrt{2}/2 & -\sqrt{2}/2 \\ 0 & -\sqrt{2}/2 & -\sqrt{2}/2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \sqrt{2} & -\sqrt{2}/2 & -\sqrt{2}/2 \\ 0 & \sqrt{3}/\sqrt{2} & \sqrt{6}/6 \\ 0 & 0 & \sqrt{3}/3 \\ 0 & 0 & 1 \end{pmatrix}$$

where

$$c = \frac{a_{22}^{(1)}}{\sqrt{(a_{22}^{(1)})^2 + (a_{32}^{(1)})^2}} = -\frac{\sqrt{3}}{3} \quad \text{and} \quad s = \frac{a_{32}^{(1)}}{\sqrt{(a_{22}^{(1)})^2 + (a_{32}^{(1)})^2}} = \frac{\sqrt{2}}{\sqrt{3}}$$

Third rotation: zeroizes element b_{43}

$$\mathbf{Q}^{(3)}\mathbf{A}^{(2)} = \mathbf{A}^{(3)}$$

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & c & s \\ 0 & 0 & -s & c \end{pmatrix} \begin{pmatrix} \sqrt{2} & -\sqrt{2}/2 & -\sqrt{2}/2 \\ 0 & \sqrt{3}/\sqrt{2} & \sqrt{6}/6 \\ 0 & 0 & \sqrt{3}/3 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \sqrt{2} & -\sqrt{2}/2 & -\sqrt{2}/2 \\ 0 & \sqrt{3}/\sqrt{2} & \sqrt{6}/6 \\ 0 & 0 & 2\sqrt{3}/3 \\ 0 & 0 & 0 \end{pmatrix}$$

where

$$c = \frac{a_{33}^{(2)}}{\sqrt{(a_{33}^{(2)})^2 + (a_{43}^{(2)})^2}} = \frac{1}{2} \quad \text{and} \quad s = \frac{a_{43}^{(2)}}{\sqrt{(a_{33}^{(2)})^2 + (a_{43}^{(2)})^2}} = \frac{\sqrt{3}}{2}$$

This gives:

$$\mathbf{U} = \begin{pmatrix} \sqrt{2} & -\sqrt{2}/2 & -\sqrt{2}/2 \\ 0 & \sqrt{3}/\sqrt{2} & \sqrt{6}/6 \\ 0 & 0 & 2\sqrt{3}/3 \\ 0 & 0 & 0 \end{pmatrix}$$

Hence:

$$\mathbf{U}'\mathbf{U} = \mathbf{A}'\mathbf{A} = \begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & 1 \\ -1 & 1 & 2 \end{pmatrix}$$

B.2.2 Fast Givens Rotations

Let us assume for a moment that $c \neq 0$. Then the rotator \mathbf{Q} can be factorized as

$$\begin{pmatrix} c & s \\ -s & c \end{pmatrix} = \begin{pmatrix} c & 0 \\ 0 & c \end{pmatrix} \begin{pmatrix} 1 & t \\ -t & 1 \end{pmatrix} \quad (\text{B.7})$$

where $t = s/c = \sin\theta/\cos\theta = \tan\theta$ (the c -transformation).

If, however, $c = 0$ (or is very small) the s -transformation can be used instead:

$$\begin{pmatrix} c & s \\ -s & c \end{pmatrix} = \begin{pmatrix} s & 0 \\ 0 & s \end{pmatrix} \begin{pmatrix} 1 & k \\ -k & 1 \end{pmatrix}$$

where $k = c/s = \cos\theta/\sin\theta = \cotan\theta$ ($\cotan\theta$ is the cotangent of θ).

The factorized rotator can be applied to the matrix \mathbf{A} above. The factor involving c (or s) is maintained and that involving t (or k) is used to update the elements of \mathbf{A} , which will require only two multiplications and two additions for each updated column (2-multiplication scheme). Of course, multiplying the transformed matrix by the c -factor (or the s -factor) will give the same results as above. This is not, however, of interest, since it would mean a return to a 4-multiplication transformation. The actual benefits of the factorized form above can be illustrated by the following example.

Let \mathbf{A} be a 3-row matrix:

$$\mathbf{A} = \begin{pmatrix} a_1 & a_2 & a_3 & \dots & a_n \\ b_1 & b_2 & b_3 & \dots & b_n \\ e_1 & e_2 & e_3 & \dots & e_n \end{pmatrix}$$

The problem is to zeroize the first element of the second row, and the first two elements of the third row:

$$\hat{\mathbf{A}} = \begin{pmatrix} \hat{a}_1 & \hat{a}_2 & \hat{a}_3 & \dots & \hat{a}_n \\ 0 & \hat{b}_2 & \hat{b}_3 & \dots & \hat{b}_n \\ 0 & 0 & \hat{e}_3 & \dots & \hat{e}_n \end{pmatrix}$$

Assume that a c -rotator (B.7) is used to zeroize element b_1 :

$$\begin{pmatrix} c^{(1)} & 0 \\ 0 & c^{(1)} \end{pmatrix} \begin{pmatrix} 1 & t^{(1)} \\ -t^{(1)} & 1 \end{pmatrix} \begin{pmatrix} a_1^{(0)} & a_2^{(0)} & a_3^{(0)} & \dots & a_n^{(0)} \\ b_1^{(0)} & b_2^{(0)} & b_3^{(0)} & \dots & b_n^{(0)} \end{pmatrix} = \\ \begin{pmatrix} c^{(1)} & 0 \\ 0 & c^{(1)} \end{pmatrix} \begin{pmatrix} a_1^{(1)} & a_2^{(1)} & a_3^{(1)} & \dots & a_n^{(1)} \\ 0 & b_2^{(1)} & b_3^{(1)} & \dots & b_n^{(1)} \end{pmatrix}$$

where

$$-a_1^{(0)}t^{(1)} + b_1^{(0)} = 0 \Rightarrow t^{(1)} = \frac{b_1^{(0)}}{a_1^{(0)}} \Rightarrow (c^{(1)})^2 = \frac{1}{(1 + (t^{(1)})^2)}$$

Notice that if $a_1^{(0)} = 0$, the rotation is 90° rotation, i.e., the two rows being processed are interchanged (in this case, it is not necessary to compute $t^{(1)}$ and $c^{(1)}$). After the first rotation, the transformed matrix is maintained in factor form:

$$\begin{pmatrix} d_1^{(1)} & 0 & 0 \\ 0 & d_2^{(1)} & 0 \\ 0 & 0 & d_3^{(1)} \end{pmatrix} \begin{pmatrix} a_1^{(1)} & a_2^{(1)} & a_3^{(1)} & \dots & a_n^{(1)} \\ 0 & b_2^{(1)} & b_3^{(1)} & \dots & b_n^{(1)} \\ e_1^{(0)} & e_2^{(0)} & e_3^{(0)} & \dots & e_n^{(0)} \end{pmatrix} \quad (\text{B.8})$$

where $d_1^{(1)} = c_1^{(1)}$, $d_2^{(1)} = c_1^{(1)}$, and $d_3^{(1)} = 1$.

Next, the element e_1 must be zeroized. Assuming a c -transformation, the result is

$$\begin{pmatrix} c^{(2)} & 0 \\ 0 & c^{(2)} \end{pmatrix} \begin{pmatrix} 1 & t^{(2)} \\ -t^{(2)} & 1 \end{pmatrix} \begin{pmatrix} d_1^{(1)} & 0 \\ 0 & d_3^{(1)} \end{pmatrix} \begin{pmatrix} a_1^{(1)} & a_2^{(1)} & \dots & a_n^{(1)} \\ e_1^{(0)} & e_2^{(0)} & \dots & e_n^{(0)} \end{pmatrix} = \\ \begin{pmatrix} d_1^{(1)}c^{(2)} & 0 \\ 0 & d_3^{(1)}c^{(2)} \end{pmatrix} \begin{pmatrix} 1 & t^{(2)}\frac{d_3^{(1)}}{d_1^{(1)}} \\ -t^{(2)}\frac{d_1^{(1)}}{d_3^{(1)}} & 1 \end{pmatrix} \begin{pmatrix} a_1^{(1)} & a_2^{(1)} & \dots & a_n^{(1)} \\ e_1^{(0)} & e_2^{(0)} & \dots & e_n^{(0)} \end{pmatrix} = \\ \begin{pmatrix} a_1^{(2)} & a_2^{(2)} & \dots & a_n^{(2)} \\ 0 & e_2^{(1)} & \dots & e_n^{(1)} \end{pmatrix}$$

where

$$-a_1^{(1)}t^{(2)}\frac{d_1^{(1)}}{d_3^{(1)}} + e_1^{(0)} = 0 \Rightarrow t^{(2)} = \frac{e_1^{(0)}d_3^{(1)}}{a_1^{(1)}d_1^{(1)}} \Rightarrow (c^{(2)})^2 = \frac{1}{(1 + (t^{(2)})^2)}$$

After the second rotation, the result is

$$\begin{pmatrix} d_1^{(2)} & 0 & 0 \\ 0 & d_2^{(2)} & 0 \\ 0 & 0 & d_3^{(2)} \end{pmatrix} \begin{pmatrix} a_1^{(2)} & a_2^{(2)} & a_3^{(2)} & \dots & a_n^{(2)} \\ 0 & b_2^{(1)} & b_3^{(1)} & \dots & b_n^{(1)} \\ 0 & e_2^{(1)} & e_3^{(1)} & \dots & e_n^{(1)} \end{pmatrix}$$

where $d_1^{(2)} = d_1^{(1)}c^{(2)}$, $d_2^{(2)} = d_2^{(1)}$, and $d_3^{(2)} = d_3^{(1)}c^{(2)}$.

Finally the element e_2 is zeroized. Assuming a c -transformation, the result is

$$\begin{aligned} & \begin{pmatrix} c^{(3)} & 0 \\ 0 & c^{(3)} \end{pmatrix} \begin{pmatrix} 1 & t^{(3)} \\ -t^{(3)} & 1 \end{pmatrix} \begin{pmatrix} d_2^{(2)} & 0 \\ 0 & d_3^{(2)} \end{pmatrix} \begin{pmatrix} b_2^{(1)} & b_3^{(1)} & \dots & b_n^{(1)} \\ e_2^{(1)} & e_3^{(1)} & \dots & e_n^{(1)} \end{pmatrix} = \\ & \begin{pmatrix} d_2^{(2)}c^{(3)} & 0 \\ 0 & d_3^{(2)}c^{(3)} \end{pmatrix} \begin{pmatrix} 1 & t^{(3)}\frac{d_3^{(2)}}{d_2^{(2)}} \\ -t^{(3)}\frac{d_2^{(2)}}{d_3^{(2)}} & 1 \end{pmatrix} \begin{pmatrix} b_2^{(1)} & b_3^{(1)} & \dots & b_n^{(1)} \\ e_2^{(1)} & e_3^{(1)} & \dots & e_n^{(1)} \end{pmatrix} = \\ & \begin{pmatrix} b_2^{(2)} & b_3^{(2)} & \dots & b_n^{(2)} \\ 0 & e_3^{(2)} & \dots & e_n^{(2)} \end{pmatrix} \end{aligned}$$

where

$$-b_2^{(1)}t^{(3)}\frac{d_2^{(2)}}{d_3^{(2)}} + e_2^{(1)} = 0 \Rightarrow t^{(3)} = \frac{e_2^{(1)}d_3^{(2)}}{b_2^{(1)}d_2^{(2)}} \Rightarrow (c^{(3)})^2 = \frac{1}{(1 + (t^{(3)})^2)}$$

Then, after the third rotation, the final result is obtained:

$$\begin{pmatrix} d_1^{(3)} & 0 & 0 \\ 0 & d_2^{(3)} & 0 \\ 0 & 0 & d_3^{(3)} \end{pmatrix} \begin{pmatrix} a_1^{(2)} & a_2^{(2)} & a_3^{(2)} & \dots & a_n^{(2)} \\ 0 & b_2^{(1)} & b_3^{(1)} & \dots & b_n^{(1)} \\ 0 & 0 & e_3^{(2)} & \dots & e_n^{(2)} \end{pmatrix} \tag{B.9}$$

where $d_1^{(3)} = d_1^{(2)}$, $d_2^{(3)} = d_2^{(2)}c^{(3)}$, and $d_3^{(3)} = d_3^{(2)}c^{(3)}$. At this point, the multiplications by the diagonal factors $d_i^{(3)}$ can be made, which will require $3(n - 1)$ multiplications. By not performing the d -factor multiplications during the process $8n - 5$ multiplications were avoided. The fast rotation scheme thus requires $5n - 4$ fewer multiplications than does the original scheme .

B.2.3 Triangular Factorization

Consider a $m \times n$ matrix \mathbf{A} with full rank and $m > n$. A sequence of rotations can be used to triangularize \mathbf{A} as indicated below:

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ a_{31} & a_{32} & a_{23} & \dots & a_{3n} \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} \\ \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & a_{m2} & \dots & a_{mn} \end{pmatrix} \Rightarrow \begin{pmatrix} u_{11} & u_{12} & u_{13} & \dots & u_{1n} \\ 0 & u_{22} & u_{23} & \dots & u_{2n} \\ 0 & 0 & u_{23} & \dots & u_{3n} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & u_{nn} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix}$$

\mathbf{A} is thus premultiplied by a sequence of orthogonal matrices such that:

$$\mathbf{Q}^r \dots \mathbf{Q}^3 \mathbf{Q}^2 \mathbf{Q}^1 \mathbf{A} = \mathbf{QA} = \mathbf{U} \tag{B.10}$$

where \mathbf{U} is upper triangular, as indicated above, \mathbf{Q}^k are $m \times m$ plane rotators, and \mathbf{Q} is orthogonal (the product of orthogonal matrices). The \mathbf{Q}^k rotators differ from unity matrices only in positions (i, i) , (i, j) , (j, i) , and (j, j) , which assume values c , s , $-s$, and c , respectively. The rotations are performed in such a way that the zeros already obtained are not affected by the new rotations, either by processing the non zero elements by rows (up \Rightarrow down) or by columns (right \Rightarrow left).

Since \mathbf{Q} is orthogonal, the factor form of \mathbf{A} is given by

$$\mathbf{A} = \mathbf{Q}' \mathbf{U} \tag{B.11}$$

B.2.4 The 2-multiplication Algorithm

Now take a closer look at an individual rotation. Assume that in performing the rotations indicated in Eq. (B.10) the transformed matrix is maintained in factorized form as in Eqs. (B.8 - B.9), i.e.:

$$\mathbf{A} = \mathbf{D}\mathbf{B}$$

where \mathbf{D} is diagonal. Consider also that the use of a row-wise scheme (rotation by row). Starting with \mathbf{B} as a unit matrix, and processing a row at one time, as indicated in Fig. B.3:

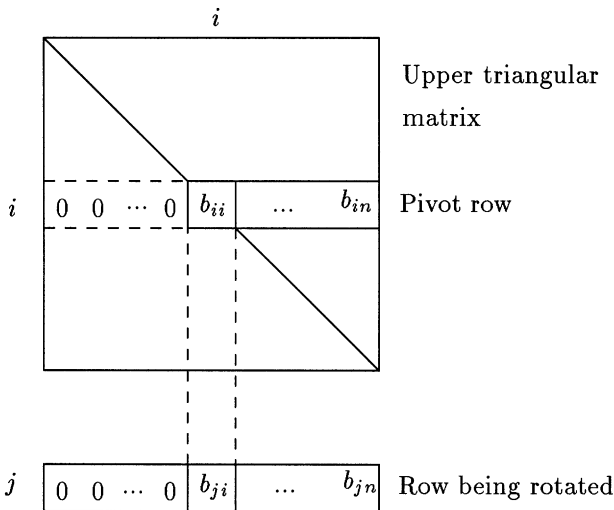


Figure B.3. Rotation of a row in upper triangular matrix \mathbf{B} , $\mathbf{A} = \mathbf{D}\mathbf{B}$.

To zeroize b_{ji} in Fig. B.3, the only two rows that are important are the pivot row and the row being processed. Thus c and t must be determined such that:

$$\begin{aligned} & \begin{pmatrix} c & 0 \\ 0 & c \end{pmatrix} \begin{pmatrix} 1 & t \\ -t & 1 \end{pmatrix} \begin{pmatrix} d_i & 0 \\ 0 & d_j \end{pmatrix} \begin{pmatrix} b_{ii} & b_{i,i+1} & b_{i,i+2} & \dots & b_{in} \\ b_{ji} & b_{j,i+1} & b_{j,i+2} & \dots & b_{jn} \end{pmatrix} = \\ & \begin{pmatrix} cd_i & 0 \\ 0 & cd_j \end{pmatrix} \begin{pmatrix} 1 & td_j/d_i \\ -td_i/d_j & 1 \end{pmatrix} \begin{pmatrix} b_{ii} & b_{i,i+1} & b_{i,i+2} & \dots & b_{in} \\ b_{ji} & b_{j,i+1} & b_{j,i+2} & \dots & b_{jn} \end{pmatrix} = \\ & \begin{pmatrix} \widehat{d}_i & 0 \\ 0 & \widehat{d}_j \end{pmatrix} \begin{pmatrix} \widehat{b}_{ii} & \widehat{b}_{i,i+1} & \widehat{b}_{i,i+2} & \dots & \widehat{b}_{in} \\ 0 & \widehat{b}_{j,i+1} & \widehat{b}_{j,i+2} & \dots & \widehat{b}_{jn} \end{pmatrix} \end{aligned}$$

where $\widehat{d}_i = cd_i$ and $\widehat{d}_j = cd_j$. Since

$$-b_{ii}t \frac{d_i}{d_j} + b_{ji} = 0$$

the required values of t and c are given by

$$t = \frac{b_{ji}d_j}{b_{ii}d_i} \quad \text{and} \quad c^2 = \frac{1}{1+t^2}$$

As mentioned above, if $b_{ii} = 0$ the two rows are simply interchanged (a 90° rotation). If $c = 0$, or $c \ll 1$, the s -transformation is used:

$$\begin{aligned} & \begin{pmatrix} s & 0 \\ 0 & s \end{pmatrix} \begin{pmatrix} k & 1 \\ -1 & k \end{pmatrix} \begin{pmatrix} d_i & 0 \\ 0 & d_j \end{pmatrix} \begin{pmatrix} b_{ii} & b_{i,i+1} & b_{i,i+2} & \dots & b_{in} \\ b_{ji} & b_{j,i+1} & b_{j,i+2} & \dots & b_{jn} \end{pmatrix} = \\ & \begin{pmatrix} sd_j & 0 \\ 0 & sd_i \end{pmatrix} \begin{pmatrix} kd_i/d_j & 1 \\ -1 & kd_j/d_i \end{pmatrix} \begin{pmatrix} b_{ii} & b_{i,i+1} & b_{i,i+2} & \dots & b_{in} \\ b_{ji} & b_{j,i+1} & b_{j,i+2} & \dots & b_{jn} \end{pmatrix} = \\ & \begin{pmatrix} \widehat{d}_i & 0 \\ 0 & \widehat{d}_j \end{pmatrix} \begin{pmatrix} \widehat{b}_{ii} & \widehat{b}_{i,i+1} & \widehat{b}_{i,i+2} & \dots & \widehat{b}_{in} \\ 0 & \widehat{b}_{j,i+1} & \widehat{b}_{j,i+2} & \dots & \widehat{b}_{jn} \end{pmatrix} \end{aligned}$$

where $\widehat{d}_i = sd_j$ and $\widehat{d}_j = sd_i$. Since

$$-b_{ii} + b_{ji}k \frac{d_j}{d_i} = 0$$

the required values of k and s are given by

$$k = \frac{b_{ii}d_i}{b_{ji}d_j} \quad \text{and} \quad s^2 = \frac{1}{1+k^2}$$

Summarizing, in the c -version of the fast Givens rotation, the following two transformations in factor matrices \mathbf{D} and \mathbf{B} (where $\mathbf{A} = \mathbf{DB}$) are performed:

$$\begin{pmatrix} c & 0 \\ 0 & c \end{pmatrix} \begin{pmatrix} d_i & 0 \\ 0 & d_j \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 1 & p \\ -q & 1 \end{pmatrix} \begin{pmatrix} b_{ii} & b_{i,i+1} & b_{i,i+2} & \dots & b_{in} \\ b_{ji} & b_{j,i+1} & b_{j,i+2} & \dots & b_{jn} \end{pmatrix}$$

where $q = td_i/d_j = b_{ji}/b_{ii}$ and $p = td_j/d_i = b_{ji}d_j^2/b_{ii}d_i^2$, whereas in the s -transformation, the results are

$$\begin{pmatrix} 0 & s \\ s & 0 \end{pmatrix} \begin{pmatrix} d_i & 0 \\ 0 & d_j \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} p & 1 \\ -1 & q \end{pmatrix} \begin{pmatrix} b_{ii} & b_{i,i+1} & b_{i,i+2} & \dots & b_{in} \\ b_{ji} & b_{j,i+1} & b_{j,i+2} & \dots & b_{jn} \end{pmatrix}$$

where $p = kd_i/d_j = b_{ii}d_i^2/b_{ji}d_j^2$ and $q = kd_j/d_i = b_{ii}/b_{ji}$.

Notice that since only the squares of d_i and d_j are needed for updating p , they are initialized in the squared form and updated as such (by multiplying either by $c^2 = 1/(1 + t^2)$ or $s^2 = 1/(1 + k^2)$); thus, t and k are obtained in squared form.

The main steps of the fast Givens method are summarized by the algorithm given below. It is assumed that the matrix being factorized has a full rank (The pivot processing order is such that $b_{ii} \neq 0$.)

- i) Compute $q = b_{ji}/b_{ii}$, $p = qd_i^2/d_j^2$, and $t^2 = pq$;
 - ii) If $t^2 \leq 1$, then
 - a) $c^2 \leftarrow 1/(1 + t^2)$;
 - b) Update the factor matrix **D**-squared:
 $d_i^2 \leftarrow c^2d_i^2$, and $d_j^2 \leftarrow c^2d_j^2$
 - c) Update the pivot row and the row being rotated (matrix **B**):
 $b_{il} \leftarrow b_{il} + pb_{jl}$ and $b_{jl} \leftarrow -qb_{il} + b_{jl}$, for $l = i, i + 1, \dots, n$
- Otherwise, if $t^2 > 1$, then
- a) $p \leftarrow 1/p$, $q \leftarrow 1/q$, $t^2 \leftarrow pq$,
 - b) $s^2 \leftarrow 1/(1 + k^2)$;
 - c) Update the factor matrix **D**-squared:
 $tmp \leftarrow d_i^2$, $d_i^2 \leftarrow s^2d_j^2$, and $d_j^2 \leftarrow s^2tmp$;
 - d) Update the pivot row and the row being rotated (matrix **B**):
 $b_{il} \leftarrow pb_{il} + b_{jl}$ and $b_{jl} \leftarrow -b_{il} + qb_{jl}$, for $l = i, i + 1, \dots, n$

Example B.2:

Reconsider the matrix **A** studied in Example B.1. In order to apply the fast Givens algorithm, **A** is initially factorized as follows:

$$\mathbf{A} = \mathbf{D}^{(0)}\mathbf{B}^{(0)} = \begin{pmatrix} 1 & -1 & -1 \\ -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$(\mathbf{D}^{(0)})^2 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad \mathbf{B}^{(0)} = \begin{pmatrix} 1 & -1 & -1 \\ -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

First rotation: zeroizes element b_{21}

Step i:

$$q = b_{21}/b_{11} = -1, \quad p = qd_2^2/d_1^2 = -1, \quad t^2 = pq = 1$$

Since $t^2 \leq 1 \Rightarrow c$ - transformation is used

Step ii:

$$c^2 \leftarrow 1/(1+t^2) = 1/2, \quad d_1^2 \leftarrow c^2 d_1^2 = 1/2, \quad d_2^2 \leftarrow c^2 d_2^2 = 1/2$$

$$b_{11} \leftarrow b_{11} + pb_{21} = 2, \quad b_{12} \leftarrow b_{12} + pb_{22} = -1, \quad b_{13} \leftarrow b_{13} + pb_{23} = -1$$

$$b_{21} \leftarrow 0, \quad b_{22} \leftarrow -qb_{22} + b_{12} = -1, \quad b_{23} \leftarrow -qb_{23} + b_{13} = -1$$

Then the updated factors become

$$(\mathbf{D}^{(1)})^2 = \begin{pmatrix} 1/2 & 0 & 0 & 0 \\ 0 & 1/2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad \mathbf{B}^{(1)} = \begin{pmatrix} 2 & -1 & -1 \\ 0 & -1 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Second rotation: zeroizes element b_{32}

Step i:

$$q = b_{32}/b_{22} = -1, \quad p = qd_3^2/d_2^2 = -2, \quad t^2 = pq = 2$$

Since $t^2 > 1 \Rightarrow s$ - transformation is used

Step ii:

$$p \leftarrow 1/p = -1/2 \quad q \leftarrow 1/q = -1 \quad k^2 \leftarrow pq = 1/2$$

$$s^2 \leftarrow 1/(1+k^2) = 2/3, \quad d_2^2 \leftarrow s^2 d_2^2 = 2/3, \quad d_3^2 \leftarrow s^2 d_3^2 = 1/3$$

$$b_{22} \leftarrow pb_{22} + b_{32} = 3/2, \quad b_{23} \leftarrow pb_{23} + b_{33} = 1/2$$

$$b_{32} \leftarrow 0, \quad b_{33} \leftarrow -b_{23} + qb_{33} = 1$$

Then the updated factors become:

$$(\mathbf{D}^{(2)})^2 = \begin{pmatrix} 1/2 & 0 & 0 & 0 \\ 0 & 2/3 & 0 & 0 \\ 0 & 0 & 1/3 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad \mathbf{B}^{(2)} = \begin{pmatrix} 2 & -1 & -1 \\ 0 & 3/2 & 1/2 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

Third rotation: zeroizes element b_{43}

Step i:

$$q = b_{43}/b_{33} = 1, \quad p = qd_4^2/d_3^2 = 3, \quad t^2 = pq = 3$$

Since $t^2 > 1 \Rightarrow s$ - transformation is used

Step ii:

$$p \leftarrow 1/p = 1/3 \quad q \leftarrow 1/q = 1 \quad k^2 \leftarrow pq = 1/3$$

$$s^2 \leftarrow 1/(1+k^2) = 3/4, \quad d_3^2 \leftarrow s^2 d_4^2 = 3/4, \quad d_4^2 \leftarrow s^2 d_3^2 = 1/4$$

$$b_{33} \leftarrow pb_{33} + b_{43} = 4/3$$

$$b_{43} \leftarrow 0$$

Then the updated factors become:

$$(\mathbf{D}^{(3)})^2 = \begin{pmatrix} 1/2 & 0 & 0 & 0 \\ 0 & 2/3 & 0 & 0 \\ 0 & 0 & 3/4 & 0 \\ 0 & 0 & 0 & 1/4 \end{pmatrix} \quad \text{and} \quad \mathbf{B}^{(3)} = \begin{pmatrix} 2 & -1 & -1 \\ 0 & 3/2 & 1/2 \\ 0 & 0 & 4/3 \\ 0 & 0 & 0 \end{pmatrix}$$

The final result is:

$$\mathbf{U} = \mathbf{A}^{(3)} = \mathbf{D}^{(3)}\mathbf{B}^{(3)} = \begin{pmatrix} \sqrt{2} & -\sqrt{2}/2 & -\sqrt{2}/2 \\ 0 & \sqrt{3}/\sqrt{2} & \sqrt{6}/6 \\ 0 & 0 & 2\sqrt{3}/3 \\ 0 & 0 & 0 \end{pmatrix}$$

which is the same result obtained in Example B.1.

B.3 HISTORICAL NOTES AND REFERENCES

A square-root free version for the Givens rotation was developed by Gentleman [1973] and improved by Hammarling [1974]. Both of these approaches to Givens rotation are reviewed by Golub and Van Loan [1989]. The use of orthogonal transformation methods in the power system static state estimation problem was introduced by Simões-Costa and Quintana [1981]. Various methods for the solution of systems of linear equations are presented by Watkins [1991]. Sparse matrix methods in general are discussed by Duff and Erisman [1992]. Efficient sparse vector/matrix techniques based on Givens rotation were developed by Vempati, Slutsker, and Tinney [1991-92].

References

- Duff, I.F. and Erisman, A.M., *Direct Methods for Sparse Matrices*, Oxford Science Publications, Oxford, 1992.
- Gentleman, W.M.; "Least squares computations by givens transformations without square roots", *J. Inst. Math. Applic.*, No. 12, pp. 329-336, 1973.
- Golub, G.H., Van Loan, C., *Matrix Computations*, John Hopkins University Press, 1989 (2nd Edition).
- Hammarling, S., "A note on modifications to the givens plane rotations", *J. Inst. Math. Applic.*, No. 13, pp. 215-218, 1974.
- Simões-Costa, A. and Quintana, V.H., "An orthogonal row processing algorithm for power system sequential state estimation", *IEEE Trans Power Appl. Syst.*, Vol. 100, No. 8, pp. 3791-3800, Aug., 1981.
- Vempati, N., Slutsker, I.W. and Tinney, W.F., "Enhancements to givens rotations for power system state estimation", *IEEE Trans. on Power System*, Vol. 6, No. 2, pp 842-849, May, 1991.
- Vempati, N., Slutsker, I.W., Tinney, W. F., "Orthogonal sparse vector methods", *IEEE Trans. Power Syst.*, Vol. 7, pp. 926-932, May 1992.
- Watkins, D.S., *Fundamentals of Matrix Computations*, John Wiley and Sons, New York, 1991.

Index

- ac state estimator, 267
- alert state, 3

- backward substitution, 109
- bad data
 - detection and identification, 201
 - basic methods, 201
 - branch and bound algorithm, 254
 - decision vector, 240
 - dormant measurement, 213, 249
 - error estimate, 213
 - hypothesis testing, 218, 236
 - interacting and conforming, 248
 - interacting and noninteracting, 227
 - largest normalized residual, 211
 - leverage points, 262
 - localized solutions, 258
 - perfect measurement, 249
 - pocketing, 258
 - robust estimator, 260
 - single, multiple, 227
 - suspect set, 240
 - tabu search, 256
 - the r^n test, 208, 222
 - the $J(\hat{\mathbf{x}})$ test, 218
 - zooming, 258
- binary tree, 248, 252
- branch and bound algorithm, 254, 263
- branch power flow equations
 - in-phase transformer, 73
 - phase-shifting transformers, 74
 - transmission line, 72
 - unified branch model, 76
- bus admittance matrix
 - expansion, 109
 - paths, 115
 - reduction, 105
 - sparsity considerations, 103
- bus section group, 145
- bus/branch model
 - meter placement, 192
 - network topology processing, 144
- BX method
 - fast decoupled power flow, 323
 - fast decoupled state estimator, 334

- Central Limit Theorem, 218
- chi-square distribution, 218, 372, 373
- combinatorial search
 - bad data, 240, 248, 259
 - least median of squares, 263
 - observability, 174
 - transformable injections, 174
- conforming bad data, 248
- contingency analysis, 1
- controllability
 - minimum-norm problem, 15, 23
 - power flow problem, 81
- correlation coefficients, 209
- covariance matrix
 - estimated vectors, 204
 - measurement estimate, 205
 - normalized residuals, 208
 - redundant measurements, 228
 - residuals, 205
 - state estimate, 205
 - trace, 370
- critical measurement, 205, 208
- current magnitude measurement, 275

- dc power flow model
 - in-phase transformer, 94
 - $P\theta$ -QV decoupling, 89
 - phase-shifting transformers, 95
 - series capacitor, 93
 - transmission line, 91
- dc state estimator
 - bad data analysis, 201
- decoupling
 - fast decoupled power flow, 316

- fast decoupled state estimator, 326
- overdetermined systems, 325
- system of linear equations, 314
- degrees of freedom
 - χ^2_v distribution, 218
 - measurement model, 44
 - network model, 41
- detection test
 - multiple errors, 236
 - the \mathbf{r}^n test, 253
 - the $J(\hat{\mathbf{x}})$ test, 218
- distribution function
 - joint marginal distribution, 229
 - state estimate, 369
 - vector of multiple residuals, 228
- diversification, 257
- dormant measurement, 213, 231
- dormant parameter, 53, 304
- dynamic state estimation
 - formulation, 284
 - Kalman Bucy filter, 288
 - Kalman filter, 287
- economy-security control, 8
- eigenvalues and eigenvectors, 230, 370
- elementary matrix, 117
- emergency state, 3
- Energy Management System, 1
- energy market, 7, 13
- equality constraints, 243
- equality-inequality constraints, 79
- error estimate, 213, 231
- Extended BX method
 - fast decoupled power flow, 319
 - fast decoupled state estimator, 331
- extended observable islands
 - network topology processing, 149
 - observability analysis, 187
- Extended XB method
 - fast decoupled power flow, 322
 - fast decoupled state estimator, 333
- external network model
 - one-pass approach, 6
 - power flow approach, 6
 - two-pass approach, 6
- false alarm
 - probability, 220, 237
- fast decoupled power flow
 - approximations, 317
 - BX method, 323
 - Extended BX method, 319
 - Extended XB method, 322
 - mid iteration updating, 317
 - Newton Raphson iteration, 316
 - XB method, 324
- fast decoupled state estimator
 - approximations, 328
 - BX method, 334
 - Extended BX method, 331
 - Extended XB method, 333
 - Gauss Newton iteration, 326
 - mid-iteration updating, 327
 - XB method, 339
- finite precision, 181
- forward substitution, 105
- gain matrix
 - factorization, 165
 - linear least-squares, 17
 - second order derivatives, 29, 268
 - singular, 165
 - sparse matrix, 131
- Gamma function, 220
- Gauss elimination
 - LDU factors, 122
 - pivot ordering, 124
 - sparse matrix, 116
- Gauss Newton method
 - dishonest, 337
 - equality constraints, 345
 - Extended BX method, 331
 - Extended XB method, 333
 - fast decoupled state estimator, 326
 - honest, 336
 - normal equation, 344
 - state estimation, 268
 - Taylor expansion, 28
- generalized state estimation, 10, 39, 55, 148, 149, 161, 188, 283
- generalized topology processing, 148
- genetic algorithms, 257
- geometric interpretation
 - least-squares problem, 26
 - minimum-norm problem, 26
- Givens rotation
 - 2-multiplication scheme, 380, 383
 - 4-multiplication scheme, 378
 - fast rotation, 382, 384
 - orthogonal matrices, 375
 - rotations and reflections, 376
 - the standard method, 378
 - triangular factorization, 382
- gross error, 222
- Hachtel tableau, 350, 354
- hybrid algorithm, 179
- hypothesis testing
 - alternative hypothesis, 219
 - equality constraints, 243, 373
 - interacting bad data, 236
 - multiple bad data, 236
 - noninteracting bad data, 218
 - null hypothesis, 219
 - significance level, 220

- single bad data, 218
- idempotent matrix
 - sensitivity matrix, 228
 - weighted sensitivity, 371
- identification of bad data
 - multiple interacting errors, 239
 - redundant measurements, 173
 - the r^n test, 222
- in-phase transformer
 - contribution to Jacobian matrix, 272
 - dc (linearized) model, 94
 - power flow equations, 73
 - tap estimation, 270
- injection measurement
 - irrelevant, 176
 - transformable, 174
 - transformations, 179
- irrelevant measurement, 176
- $J(\hat{\mathbf{x}})$ index
 - normal equation, 16, 344
 - update, 215, 235
- Jacobian matrix
 - current measurement, 275
 - flow state, 185
 - in-phase transformer flow, 272
 - injection measurement, 274, 278
 - network topology processing, 154
 - normal equation, 268
 - pockets, 258
 - power flow problem, 86
 - power flow equations, 82
 - rank deficiency, 183
 - transmission line flow, 270
- Kalman Bucy filter, 288
- Kalman filter, 287
- Karush-Kuhn-Tucker conditions
 - equality constraints, 345
 - Hachtel tableau, 350
- Kirchhoff's current law, 76
- Lagrange multipliers
 - equality constraints, 347
 - example, 349
 - minimum-norm problem, 20
- Lagrangian function
 - Hachtel tableau, 348
 - Hachtel tableau with equality constraints, 350
 - minimum-norm problem, 19
 - normal equation with equality constraints, 344
- largest normalized residual
 - criterion, 248
 - repeated application, 248
- single bad data, 210
 - when it fails, 248
- LDU factors
 - Gauss elimination, 122
- least absolute value estimator, 261
- least median of squares estimator, 262
- least-squares index, 202
- least-squares problem
 - formulation, 16
 - geometric interpretation, 26
 - Moore-Penrose conditions, 17
 - observability, 22
 - overdetermined systems, 15
 - pseudo-inverse matrix, 17
 - solvability, 15
- leverage points, 262
- local state estimation
 - network topology processing, 151
 - pocketing and zooming, 258
- matrix formulation
 - bus admittance matrix, 103
 - dc (linearized) model, 97
- matrix inversion lemma
 - derivation, 138
 - dynamic state estimation, 286
 - Kalman filter, 287
- MDML and MLMD schemes, 126
- measurement
 - branch assignment, 174
 - critical, 205, 228
 - current magnitude, 273, 275
 - dormant, 231
 - error estimate, 231
 - irrelevant, 176
 - model, 184
 - noncritical, 208
 - power flow, 268
 - power injection, 275
 - redundant, 173, 228, 237
- measurement model
 - concept, 44
 - degrees of freedom, 44
 - linear, 227
 - rank, 44
- measurement set partitioning, 228
- measurement/state adjacency algorithm, 259
- meter arrangement, 153
- meter placement
 - additional measurements, 191
 - avoiding contamination, 192
 - bus/branch model, 192
 - observability, 191
 - physical level model, 195
- minimum-norm problem
 - controllability, 15, 23
 - geometric interpretation, 26

- Lagrange multipliers, 20
- problem formulation, 19
- pseudo-inverse matrix, 20
- Moore-Penrose conditions, 17
- multiple normalized residuals
 - definition, 233
 - Euclidean norm, 237
 - probability density function, 233
- network expansion
 - backward substitution, 109
 - paths, 115
- network model
 - bus-section/switching device, 55
 - bus-section/switching-device, 183
 - bus/branch, 161
 - nodal formulation, 76
 - phase-shifting transformers, 69
 - physical level, 55, 148, 161
 - state variables, 41
 - transmission line, 63
 - unified branch model, 70, 268
- network model builder, 3
- network reduction
 - forward substitution, 105
 - network topology processing, 152
 - paths, 115
- network security, 1
- network topology processing
 - bus section processing, 145
 - bus/branch model, 144
 - connectivity analysis, 147
 - conventional approach, 143
 - extended observable islands, 149
 - generalized topology processing, 148
 - Jacobian reduction, 154
 - local state estimation, 151
 - meter arrangement, 153
 - network reduction, 152
 - pre-processing of raw data, 144
 - tracking, 147
- network topology processor, 161
- Newton Raphson method
 - fast decoupled power flow, 316
 - parameter estimation, 301
 - power flow problem, 81
 - state estimation, 29, 267
- nodal formulation, 76
- node ordering, 166
- noncritical measurement, 208
- noninteracting bad data, 201
- normal distribution, 218
- normal equation
 - "unsquaring", 344
 - dc state estimator, 49
 - least-squares problem, 16, 344
 - parameter estimation, 290
 - tracking state estimator, 283
- normalized residuals
 - definition, 208
 - largest, 210
- numerical coincidences, 171
- numerical observability
 - basic algorithm, 171
 - bus/branch, 171
 - extended algorithm, 189
 - irrelevant measurements, 189
 - numerical mode, 171
 - topological mode, 171
- observability
 - least-squares problem, 22
 - physical level, 184
 - power flow problem solvability, 79
- observable island
 - definition, 162, 188
 - extended, 187
- Ohm's law (dc power flow)
 - flow state, 186
- open access
 - see energy market, 13
- ordering to preserve sparsity
 - MDML and MLMD schemes, 126
 - Tinney scheme, 124
- orthogonal transformation
 - estimator, 357
 - factorization, 172
 - Givens rotation, 375
 - index $J(\hat{\mathbf{x}})$, 372
 - observability analysis, 166
 - rank deficiency, 183
- overdetermined systems
 - linear model, 16
 - nonlinear models, 28, 267
- $P\theta$ -QV decoupling
 - dc power flow model, 89
 - decoupled solution, 316
- parameter estimation
 - alternate approach, 295
 - dormant technique, 53, 304
 - dynamic estimation, 300
 - multiple scans, 289
 - state augmentation, 289
- paths
 - network reduction/expansion, 115
 - sparse matrix, 123
- performance index
 - normal equation, 16, 344
 - update, 215, 235
- Peters Wilkinson method, 351
- phase-shifting transformer
 - ac model, 69
 - dc (linearized) model, 95
 - phase shift estimation, 52, 270

- power flow equations, 74
- physical island
 - definition, 162, 188
- physical level model
 - least-squares estimator, 55
 - meter placement, 195
 - multiple bad data, 259
 - network model, 161
 - network topology processing, 148
- pivot ordering
 - Gauss elimination, 124
 - MDML and MLMD schemes, 126
 - Tinney scheme, 124
- pocketing, 258
- power flow problem
 - basic bus types, 78
 - basic formulation, 78
 - equality-inequality constraints, 79
 - Newton Raphson method, 81
 - power flow equations, 63
 - solvability, 79
- probability of false alarm, 220
- pseudo-inverse matrix
 - least-squares problem, 17
 - minimum-norm problem, 20
- pseudo-measurement
 - angle difference, 180
 - meter placement, 192
- pseudo-measurement
 - load prediction, 5
 - one-pass method, 6
 - state estimator, 7
 - two-pass method, 6
- quadratic convergence, 32, 85
- quasi-static network models, 3
- range and null spaces, 24
- raw data, 144
- redundant measurements, 172
- reference angle, 165
- restorative state, 3
- robust estimator
 - least absolute value, 261
 - least median of squares, 262
 - non-quadratic estimators, 260
 - objective functions, 261
 - the influence of outliers, 260
- rotations and reflections, 376
- secure state, 2
- security levels, 2
- sensitivity analysis, 202
- sensitivity matrix
 - $\partial \hat{\mathbf{r}} / \partial \mathbf{z}$, 203
 - $\partial \hat{\mathbf{x}} / \partial \mathbf{z}$, 202
 - $\partial \hat{\mathbf{z}} / \partial \mathbf{z}$, 203
- idempotent, 228
- sensitivity relation, 228
- series capacitor
 - dc (linearized) model, 93
- simulated annealing, 257
- small branch impedances, 181
- solvability
 - least-squares problem, 15
 - topological observability, 181
- solvability and observability, 15, 181
- spanning tree
 - measurement assignment, 174
 - solvability and observability, 181
- sparse matrix
 - bus admittance matrix, 103
 - elementary matrix, 117
 - gain matrix, 131
 - Gauss elimination, 116
 - paths, 123
 - triangular factorization, 119
- standard normal distribution, 222
- state variables
 - definition, 41
 - power flow, 184
- system of linear equations
 - decoupling, 314, 325
 - three-step algorithm, 314, 325
- tabu search, 256, 263
- Taylor expansion, 86, 268, 317, 327, 345, 347
- three-step algorithm
 - fast decoupled estimator, 332, 333
 - fast decoupled power flow, 320, 322
 - system of linear equations, 314, 325
- topological observability
 - basic algorithm, 164, 179
 - bus/branch, 164
 - solvability, 181
- topology processing (see network topology processing), 143
- tracking state estimator
 - network topology processing, 147
 - normal equation, 283
- transmission line
 - ac model, 63
 - contribution to Jacobian matrix, 270
 - dc (linearized) model, 91
 - power flow equations, 72
- triangular factorization
 - Gauss elimination, 119
 - Givens rotation, 382
 - observability analysis, 165
- two-step algorithm
 - fast decoupled estimator, 334, 340
 - fast decoupled power flow, 323, 324
- unified branch model

- ac model, 70
- contribution to Jacobian matrix, 268
- power flow equations, 76
- weighted least-squares problem, 28
- weighted residuals, 203, 235
- weighted sensitivity matrix
 - diagonalization, 371
 - multiple bad data, 228
- weighting matrix, 202
- XB method
 - fast decoupled power flow, 324
 - fast decoupled state estimator, 339
- zero pivot, 165, 171
- zooming, 258