558 LECTURE NOTES IN ECONOMICS AND MATHEMATICAL SYSTEMS

Mario Faliva
Maria Grazia Zoia

Topics in Dynamic Model Analysis

Advanced Matrix Methods and Unit-Root Econometrics Representation Theorems

Springer

Lecture Notes in Economics and Mathematical Systems 558

Mario Faliva
Maria Grazia Zoia

Topics in Dynamic Model Analysis

Advanced Matrix Methods and Unit-Root Econometrics Representation Theorems

Springer

Authors

Prof. Mario Faliva
Full Professor of Econometrics and
Head of the Department of Econometrics
and Applied Mathematics
Faculty of Economics
Catholic University of Milan
Largo Gemelli, 1
I-20123 Milano, Italy
mario.faliva@unicatt.it

Prof. Maria Grazia Zoia
Associate Professor of Econometrics
Faculty of Economics
Catholic University of Milan
Largo Gemelli, 1
I-20123 Milano, Italy
maria.zoia@unicatt.it

Library of Congress Control Number: 2005931329

ISSN 0075-8442
ISBN-10 3-540-26196-6 Springer Berlin Heidelberg New York
ISBN-13 978-3-540-26196-4 Springer Berlin Heidelberg New York

Springer is a part of Springer Science+Business Media

springeronline.com

Typesetting: Camera ready by author
Cover design: *Erich Kirchner*, Heidelberg

Printed on acid-free paper 42/3130Jö 5 4 3 2 1 0

To Massimiliano

To Giulia and Sofia

Preface

Classical econometrics – which plunges its roots in economic theory with simultaneous equations models (SEM) as offshoots – and time series econometrics – which stems from economic data with vector autoregressive (VAR) models as offsprings – scour, like the Janus's facing heads, the flowing of economic variables so as to bring to the fore their autonomous and non-autonomous dynamics. It is up to the so-called final form of a dynamic SEM, on the one hand, and to the so-called representation theorems of (unit-root) VAR models, on the other, to provide informative closed form expressions for the trajectories, or time paths, of the economic variables of interest.

Should we look at the issues just put forward from a mathematical standpoint, the emblematic models of both classical and time series econometrics would turn out to be difference equation systems with *ad hoc* characteristics, whose solutions are attained via a final form or a representation theorem approach. The final form solution – algebraic technicalities apart – arises in the wake of classical difference equation theory, displaying besides a transitory autonomous component, an exogenous one along with a stochastic nuisance term. This follows from a properly defined matrix function inversion admitting a Taylor expansion in the lag operator because of the assumptions regarding the roots of a determinant equation peculiar to SEM specifications.

Such was the state of the art when, after Granger's seminal work, time series econometrics came into the limelight and (co)integration burst onto the stage. While opening up new horizons to the modelling of economic dynamics, this nevertheless demanded a somewhat sophisticated analytical apparatus to bridge the unit-root gap between SEM and VAR models.

Over the past two decades econometric literature has by and large given preferential treatment to the rôle and content of time series econometrics as such and as compared with classical econometrics. Meanwhile, a fascinating – although at time cumbersome – algebraic toolkit has taken shape in a sort of osmotic relationship with (co)integration theory advancements.

The picture just outlined, where lights and shadows – although not explicitly mentioned – still share out the scene, spurs us on to seek a deeper insight into several facets of dynamic model analysis, whence the idea of this monograph devoted to representation theorems and their analytical foundations.

The book is organised as follows.

Chapter 1 is designed to provide the reader with a self-contained treatment of matrix theory aimed at paving the way to a rigorous derivation of representation theorems later on. It brings together several results on generalized inverses, orthogonal complements, partitioned inversion rules (some of them new) and investigates the issue of matrix polynomial inversion about a pole (in its relationships with difference equation theory) via Laurent expansions in matrix form, with the notion of Schur complement and a newly found partitioned inversion formula playing a crucial rôle in the determination of coefficients.

Chapter 2 deals with statistical setting problems tailored to the special needs of this monograph . In particular, it covers the basic concepts on stochastic processes – both stationary and integrated – with a glimpse at cointegration in view of a deeper insight to be provided in the next chapter.

Chapter 3, after outlining a common frame of reference for classical and time series econometrics bridging the unit-root gap between structural and vector autoregressive models, tackles the issue of VAR specification and resulting processes, with the integration orders of the latters drawn from the rank characteristics of the formers. Having outlined the general setting, the central topic of representation theorems is dealt with, in the wake of time series econometrics tradition named after Granger and Johansen (to quote only the forerunner and the leading figure *par excellence*), and further developed along innovating directions thanks to the effective analytical toolkit set forth in Chapter 1.

The book is obviously not free from external influences and acknowledgement must be given to the authors, quoted in the reference list, whose works have inspired and stimulated the writing of this book.

We should like to express our gratitude to Siegfried Schaible for his encouragement about the publication of this monograph.

Our greatest debt is to Giorgio Pederzoli, who read the whole manuscript and made detailed comments and insightful suggestions.

We are also indebted to Wendy Farrar for her peerless checking of the text.

Finally we would like to thank Daniele Clarizia for his painstaking typing of the manuscript.

Milan, March 2005

Mario Faliva and Maria Grazia Zoia
Istituto di Econometria e Matematica
Università Cattolica, Milano

Contents

1 The Algebraic Framework of Unit-Root Econometrics

Time series econometrics is centred around the representation theorems from which one can evict the integration and cointegration characteristics of the solutions for the vector autoregressive (VAR) models.

Such theorems, along the path established by Engle and Granger and by Johansen and his school, have promoted the parallel development of an "ad hoc" analytical implement – although not always fully settled.

The present chapter, by reworking and expanding some recent contributions due to Faliva and Zoia, provides in an organic fashion an algebraic setting based upon several interesting results on inversion by parts and on Laurent series expansion for the reciprocal of a matrix polynomial in a deleted neighbourhood of a unitary root. Rigorous and efficient, such a technique allows for a quick and new reformulation of the representation theorems as it will become clear in Chapter 3.

1.1 Generalized Inverses and Orthogonal Complements

We begin by giving some definitions and theorems on generalized inverses. For these and related results see Rao and Mitra (1971), Pringle and Rayner (1971), S.R. Searle (1982).

Definition 1

A generalized inverse of a matrix $\boldsymbol{A}$ of order $m \times n$ is a matrix $\boldsymbol{A}^-$ of order $n \times m$ such that

$$\boldsymbol{A}\,\boldsymbol{A}^-\boldsymbol{A} = \boldsymbol{A} \tag{1}$$

The matrix $\boldsymbol{A}^-$ is not unique unless $\boldsymbol{A}$ is a square non-singular matrix.

We will adopt the following conventions

$$\boldsymbol{B} = \boldsymbol{A}^- \tag{2}$$

to indicate that $\boldsymbol{B}$ is a generalized inverse of $\boldsymbol{A}$;

$$A^{-} = B \tag{3}$$

to indicate that one possible choice for the generalized inverse of A is given by the matrix B.

Definiton 2

The Moore-Penrose generalized inverse of a matrix A of order $m \times n$ is a matrix A^g of order $n \times m$ such that

$$A\,A^g A = A \tag{4}$$

$$A^g A\,A^g = A^g \tag{5}$$

$$(A\,A^g)' = A\,A^g \tag{6}$$

$$(A^g A)' = A^g A \tag{7}$$

where A' stands for the transpose of A. The matrix A^g is unique.

Definition 3

A right inverse of a matrix A of order $m \times n$ and full row-rank is a matrix A_r^- of order $n \times m$ such that

$$A\ A_r^- = I \tag{8}$$

Theorem 1

The general expression of A_r^- is

$$A_r^- = H'\,(AH')^{-1} \tag{9}$$

where H is an arbitrary matrix of order $m \times n$ such that

$$det\,(AH') \neq 0 \tag{10}$$

Proof

For a proof see Rao and Mitra (1971, Theorem 2.1.1).

□

Remark

By taking $\boldsymbol{H} = \boldsymbol{A}$, we obtain

$$\boldsymbol{A}_r^- = \boldsymbol{A}'(\boldsymbol{A}\boldsymbol{A}')^{-1} \equiv \boldsymbol{A}^g, \tag{11}$$

a particularly useful form of right inverse.

Definition 4

A left inverse of a matrix $\boldsymbol{A}$ of order $m \times n$ and full column-rank is a matrix $\boldsymbol{A}_l^-$ of order $n \times m$ such that

$$\boldsymbol{A}_l^- \boldsymbol{A} = \boldsymbol{I} \tag{12}$$

Theorem 2

The general expression of $\boldsymbol{A}_l^-$ is

$$\boldsymbol{A}_l^- = (\boldsymbol{K}'\boldsymbol{A})^{-1}\boldsymbol{K}' \tag{13}$$

where $\boldsymbol{K}$ is an arbitrary matrix of order $m \times n$ such that

$$det\,(\boldsymbol{K}'\boldsymbol{A}) \neq 0 \tag{14}$$

Proof

For a proof see Rao and Mitra (1971, Theorem 2.1.1).

□

Remark

By letting $\boldsymbol{K} = \boldsymbol{A}$, we obtain

$$\boldsymbol{A}_l^- = (\boldsymbol{A}'\boldsymbol{A})^{-1}\boldsymbol{A}' \equiv \boldsymbol{A}^g \tag{15}$$

a particularly useful form of left inverse.

We will now introduce the notion of rank factorization.

Theorem 3

Any matrix $\boldsymbol{A}$ of order $m \times n$ and rank r may be factored as follows

$$\boldsymbol{A} = \boldsymbol{B}\boldsymbol{C}' \tag{16}$$

where $\boldsymbol{B}$ is of order $m \times r$, $\boldsymbol{C}$ is of order $n \times r$, and both $\boldsymbol{B}$ and $\boldsymbol{C}$ have rank equal to r.

Such a representation is known as a rank factorization of $\boldsymbol{A}$.

Proof

For a proof see Searle (1982, p. 194).

□

Theorem 4

Let a matrix $\boldsymbol{A}$ of order $m \times n$ and rank r be factored as in (16). Then $\boldsymbol{A}^-$ can be factored as

$$\boldsymbol{A}^- = (\boldsymbol{C}')_r^- \, \boldsymbol{B}_l^- \tag{17}$$

with the noteworthy relationship

$$\boldsymbol{C}'\boldsymbol{A}^-\boldsymbol{B} = \boldsymbol{I} \tag{18}$$

as a by-product.

In particular, the Moore-Penrose inverse $\boldsymbol{A}^g$ can be factored as

$$\boldsymbol{A}^g = (\boldsymbol{C}')^g\boldsymbol{B}^g = \boldsymbol{C}\,(\boldsymbol{C}'\boldsymbol{C})^{-1}\,(\boldsymbol{B}'\boldsymbol{B})^{-1}\,\boldsymbol{B}' \tag{19}$$

Proof

The proofs of both (17) and (18) are simple and are omitted. For a proof of (19) see Greville (1960).

□

We shall now introduce some further definitions and establish several results on orthogonal complements. For these and related results see Thrall and Tornheim (1957), Lancaster and Tismenetsky (1985), Lütkepohl (1996) and the already quoted references.

Definition 5

The row kernel, or null row space, of a matrix $\boldsymbol{A}$ of order $m \times n$ and rank r is the space of dimension $(m - r)$ of all solutions $\boldsymbol{x}$ of $\boldsymbol{x}'\,\boldsymbol{A}' = \boldsymbol{0}'$.

Definition 6

An orthogonal complement of a matrix $\boldsymbol{A}$ of order $m \times n$ and full column-rank is a matrix $\boldsymbol{A}_{\perp}$ of order $m \times (m-n)$ and full column-rank such that

$$\boldsymbol{A}'_{\perp}\boldsymbol{A} = \boldsymbol{0} \tag{20}$$

Remark

The matrix $\boldsymbol{A}_{\perp}$ is not unique. Indeed the columns of $\boldsymbol{A}_{\perp}$ form not only a spanning set, but even a basis for the row kernel of $\boldsymbol{A}$ and the other way around. In light of the foregoing, a general representation of the orthogonal complement of a matrix $\boldsymbol{A}$ is given by

$$\boldsymbol{A}_{\perp} = \boldsymbol{\Lambda V} \tag{21}$$

where $\boldsymbol{\Lambda}$ is a particular orthogonal complement of $\boldsymbol{A}$ and $\boldsymbol{V}$ is an arbitrary square non-singular matrix connecting the reference basis (namely, the $m-n$ columns of $\boldsymbol{\Lambda}$) to an another (namely, the $m-n$ columns of $\boldsymbol{\Lambda V}$) . The matrix $\boldsymbol{V}$ is usually referred to as a *transition matrix* between bases (cf. Lancaster and Tismenetsky, 1985, p. 98).

We shall adopt the following conventions:

$$\boldsymbol{\Lambda} = \boldsymbol{A}_{\perp} \tag{22}$$

to indicate that $\boldsymbol{\Lambda}$ is an orthogonal complement of $\boldsymbol{A}$;

$$\boldsymbol{A}_{\perp} = \boldsymbol{\Lambda} \tag{23}$$

to indicate that one possible choice for the orthogonal complement of $\boldsymbol{A}$ is given by the matrix $\boldsymbol{\Lambda}$.

The equality

$$(\boldsymbol{A}_{\perp})_{\perp} = \boldsymbol{A} \tag{24}$$

reads accordingly.

We now prove the following invariance theorem

Theorem 5

The expressions

$$\boldsymbol{A}_{\perp}(\boldsymbol{H}'\boldsymbol{A}_{\perp})^{-1} \tag{25}$$

$$C_\perp (B'_\perp K C_\perp)^{-1} B'_\perp \tag{26}$$

and the rank of the partitioned matrix

$$\begin{bmatrix} J & B_\perp \\ C'_\perp & 0 \end{bmatrix} \tag{27}$$

are invariant for any choice of $A_\perp$, $B_\perp$ and $C_\perp$, where A, B and C are full column-rank matrices of order $m \times n$, H is an arbitrary full column-rank matrix of order $m \times (m-n)$ such that

$$det\,(H'A_\perp) \neq 0 \tag{28}$$

and both J and K, of order m, are arbitrary matrices, except that

$$det\,(B'_\perp K C_\perp) \neq 0 \tag{29}$$

Proof

To prove the invariance of the matrix (25) we check that

$$A_{\perp 1}\,(H'A_{\perp 1})^{-1} - A_{\perp 2}\,(H'A_{\perp 2})^{-1} = 0 \tag{30}$$

where $A_{\perp 1}$ and $A_{\perp 2}$ are two choices of the orthogonal complement of A. After the arguments put forward to arrive at (21), the matrices $A_{\perp 1}$ and $A_{\perp 2}$ are linked by the relation

$$A_{\perp 2} = A_{\perp 1} V \tag{31}$$

for a suitable choice of the transition matrix V.

Substituting $A_{\perp 1} V$ for $A_{\perp 2}$ in the left-hand side of (30) yields

$$\begin{aligned} A_{\perp 1}\,(H'A_{\perp 1})^{-1} - A_{\perp 1} V (H'A_{\perp 1} V)^{-1} &= A_{\perp 1}\,(H'A_{\perp 1})^{-1} \\ &- A_{\perp 1} V V^{-1} (H'A_{\perp 1})^{-1} = 0 \end{aligned} \tag{32}$$

which proves the asserted invariance.

The proof of the invariance of the matrix (26) follows along the same lines as above by repeating for $B_\perp$ and $C_\perp$ the reasoning used for $A_\perp$.

The proof of the invariance of the rank of the matrix (27) follows upon noticing that

$$
\begin{aligned}
& r\left(\begin{bmatrix} J & B_{\perp 2} \\ C'_{\perp 2} & 0 \end{bmatrix}\right) = r\left(\begin{bmatrix} J & B_{\perp 1}V_1 \\ V'_2 C'_{\perp 1} & 0 \end{bmatrix}\right) \\
& = r\left(\begin{bmatrix} I & 0 \\ 0 & V'_2 \end{bmatrix}\begin{bmatrix} J & B_{\perp 1} \\ C'_{\perp 1} & 0 \end{bmatrix}\begin{bmatrix} I & 0 \\ 0 & V_1 \end{bmatrix}\right) = r\left(\begin{bmatrix} J & B_{\perp 1} \\ C'_{\perp 1} & 0 \end{bmatrix}\right)
\end{aligned}
\tag{33}
$$

where V_1 and V_2 are suitable choices of transition matrices.

□

The following theorem provides explicit expressions for orthogonal complements of matrix products, which find considerable use in the text.

Theorem 6

Let A and B be full column-rank matrices of order $l \times m$ and $m \times n$ respectively. Then the orthogonal complement of the matrix product AB can be expressed as

$$(AB)_{\perp} = [(A')^{g} B_{\perp}, A_{\perp}] \tag{34}$$

In particular if $l = m$, then the following holds

$$(AB)_{\perp} = (A')^{-1} B_{\perp} \tag{35}$$

Moreover, if C is any non-singular matrix of order m, then we can write

$$(BC)_{\perp} = B_{\perp} \tag{36}$$

Proof

Observe that

$$(AB)' [(A')^{g} B_{\perp}, A_{\perp}] = 0 \tag{37}$$

and that the block matrix

$$[(A')^{g} B_{\perp}, A_{\perp}, AB] \tag{38}$$

is square and of full rank. Hence the matrix $[(A')^{g} B_{\perp}, A_{\perp}]$ provides an explicit expression for the orthogonal complement of AB, according to Definition 6 (see also Faliva and Zoia, 2003).

The result (35) is established by straightforward computation.

The result (36) is easily proved and rests on the arguments underlying the representation (21) of orthogonal complements.

□

The next three theorems provide expressions for generalized and regular inverses of block matrices and related results of major interest for our analysis.

Theorem 7

Suppose that $\boldsymbol{A}$ and $\boldsymbol{B}$ are as in Theorem 6. Then

$$[(\boldsymbol{A}')^g \boldsymbol{B}_\perp, \boldsymbol{A}_\perp]^- = \begin{bmatrix} \boldsymbol{B}_\perp^g \boldsymbol{A}' \\ \boldsymbol{A}_\perp^g \end{bmatrix} \tag{39}$$

$$[(\boldsymbol{A}')^g \boldsymbol{B}_\perp, \boldsymbol{A}_\perp]^g = \begin{bmatrix} ((\boldsymbol{A}')^g \boldsymbol{B}_\perp)^g \\ \boldsymbol{A}_\perp^g \end{bmatrix} \tag{40}$$

Proof

The results follow from Definitions 1 and 2 by applying Theorems 3.1 and 3.4, Corollary 4, in Pringle and Rayner (1971) p. 38.

□

Theorem 8

The inverse of the composite matrix $[\boldsymbol{A}, \boldsymbol{A}_\perp]$ can be written as follows

$$[\boldsymbol{A}, \boldsymbol{A}_\perp]^{-1} = \begin{bmatrix} \boldsymbol{A}^g \\ \boldsymbol{A}_\perp^g \end{bmatrix} \tag{41}$$

which, in turns, leads to the noteworthy identity

$$\boldsymbol{A}\boldsymbol{A}^g + \boldsymbol{A}_\perp \boldsymbol{A}_\perp^g = \boldsymbol{I} \tag{42}$$

Proof

The proof is a by-product of Theorem 3.4, Corollary 4, in Pringle and Rayner (1971), and the identity (42) ensues from the commutative property of the inverse.

□

The following theorem provides a useful generalization of the identity (42).

Theorem 9

Let $\boldsymbol{A}$ and $\boldsymbol{B}$ be full column-rank matrices of order $m \times n$ and $m \times (m-n)$ respectively, such that the composite matrix $[\boldsymbol{A}, \boldsymbol{B}]$ is non singular. Then, the following identity

$$\boldsymbol{A}\,(\boldsymbol{B}_{\perp}'\boldsymbol{A})^{-1}\boldsymbol{B}_{\perp}' + \boldsymbol{B}\,(\boldsymbol{A}_{\perp}'\boldsymbol{B})^{-1}\boldsymbol{A}_{\perp}' = \boldsymbol{I} \tag{43}$$

holds true.

Proof

Observe that insofar as the square matrix $[\boldsymbol{A}, \boldsymbol{B}]$ is non-singular, both $\boldsymbol{B}_{\perp}'\boldsymbol{A}$ and $\boldsymbol{A}_{\perp}'\boldsymbol{B}$ are non-singular matrices also.

Furthermore, verify that

$$\begin{bmatrix} (\boldsymbol{B}_{\perp}'\boldsymbol{A})^{-1}\boldsymbol{B}_{\perp}' \\ (\boldsymbol{A}_{\perp}'\boldsymbol{B})^{-1}\boldsymbol{A}_{\perp}' \end{bmatrix} [\boldsymbol{A}, \quad \boldsymbol{B}] = \begin{bmatrix} \boldsymbol{I}_n & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I}_{m-n} \end{bmatrix} \tag{44}$$

This shows that $\begin{bmatrix} (\boldsymbol{B}_{\perp}'\boldsymbol{A})^{-1}\boldsymbol{B}_{\perp}' \\ (\boldsymbol{A}_{\perp}'\boldsymbol{B})^{-1}\boldsymbol{A}_{\perp}' \end{bmatrix}$ is the inverse of $[\boldsymbol{A}, \boldsymbol{B}]$. Hence the identity (43) follows from the commutative property of the inverse.

□

Let us now quote a few identities which can easily be proved because of Theorems 4 and 8

$$\boldsymbol{A}\boldsymbol{A}^{g} = \boldsymbol{B}\boldsymbol{B}^{g} \tag{45}$$

$$\boldsymbol{A}^{g}\boldsymbol{A} = (\boldsymbol{C}')^{g}\,\boldsymbol{C}' \tag{46}$$

$$\boldsymbol{I}_m - \boldsymbol{A}\boldsymbol{A}^{g} = \boldsymbol{I}_m - \boldsymbol{B}\boldsymbol{B}^{g} = \boldsymbol{B}_{\perp}\,(\boldsymbol{B}_{\perp})^{g} = (\boldsymbol{B}_{\perp}')^{g}\,\boldsymbol{B}_{\perp}' \tag{47}$$

$$\boldsymbol{I}_n - \boldsymbol{A}^{g}\boldsymbol{A} = \boldsymbol{I}_n - (\boldsymbol{C}')^{g}\,\boldsymbol{C}' = (\boldsymbol{C}_{\perp}')^{g}\boldsymbol{C}_{\perp}' = \boldsymbol{C}_{\perp}(\boldsymbol{C}_{\perp})^{g} \tag{48}$$

where $\boldsymbol{A}$, $\boldsymbol{B}$ and $\boldsymbol{C}$ are as in Theorem 3.

To conclude this section, let us observe that an alternative definition of orthogonal complement – which differs slightly from that of Definition 6 – may be more conveniently adopted for square singular matrices as indicated in the next definition.

Definition 7

Let A be a square matrix of order n and rank $r < n$. A left-orthogonal complement of A is a square matrix of order n and rank $n - r$, denoted by $A_l^{\perp}$, such that

$$A_l^{\perp} A = 0 \tag{49}$$

$$r([A_l^{\perp}, A]) = n \tag{50}$$

Analogously, a right-orthogonal complement of A is a square matrix of order n and rank $n - r$, denoted by $A_r^{\perp}$, such that

$$A\ A_r^{\perp} = 0 \tag{51}$$

$$r([A,\ A_r^{\perp}]) = n \tag{52}$$

Suitable choices for the matrices $A_l^{\perp}$ and $A_r^{\perp}$ turn out to be the idempotent matrices (see, e.g., Rao, 1973)

$$A_l^{\perp} = I - AA^g \tag{53}$$

$$A_r^{\perp} = I - A^g A \tag{54}$$

which will henceforth simply be denoted by $A_l^{\perp}$ and $A_r^{\perp}$, respectively, unless otherwise stated.

1.2 Partitioned Inversion: Classical and Newly Found Results

This section, after recalling classic results on partitioned inversion, presents newly found (see, in this regard, Faliva and Zoia, 2002a) inversion formulas which, like Pandora's box, provide the keys to an elegant and rigorous approach to unit-root econometrics main theorems, as shown in Chapter 3.

To begin with we recall the following classical result:

Theorem 1

Let A and D be square matrices of order m and n, respectively, and let B and C be full column-rank matrices of order $m \times n$.
Consider the partitioned matrix

$$P = \begin{bmatrix} A & B \\ C' & D \end{bmatrix} \tag{1}$$

Then anyone of the following sets of conditions is sufficient for the existence of P^{-1}

a) Both A and its Schur complement $E = D - C'A^{-1}B$ are non-singular matrices.

b) Both D and its Schur complement $F = A - BD^{-1}C'$ are non-singular matrices.

Moreover the results listed below hold true:

i) Under a), the partitioned inverse of P can be written as

$$P^{-1} = \begin{bmatrix} A^{-1} + A^{-1}BE^{-1}C'A^{-1} & -A^{-1}BE^{-1} \\ -E^{-1}C'A^{-1} & E^{-1} \end{bmatrix} \tag{2}$$

ii) Under b), the partitioned inverse of P can be written as

$$P^{-1} = \begin{bmatrix} F^{-1} & -F^{-1}BD^{-1} \\ -D^{-1}C'F^{-1} & D^{-1} + D^{-1}C'F^{-1}BD^{-1} \end{bmatrix} \tag{3}$$

Proof

The matrix P^{-1} exists insofar as (see Rao, 1973, p. 32)

$$det(P) = \begin{cases} det(A)\,det(E) \neq 0, \; under\ a) \\ det(D)\,det(F) \neq 0, \; under\ b) \end{cases} \tag{4}$$

The partitioned inversion formulas (2) and (3), under the assumptions a) and b), respectively, are standard results of the algebraic tool-kit of econometricians (see, e.g., Goldberger, 1964; Theil, 1971; Faliva, 1987).

□

We shall now establish the main result (see also Faliva and Zoia, 2002 a).

Theorem 2

Consider the block matrix

$$P = \begin{bmatrix} A & B \\ C' & 0 \end{bmatrix} \tag{5}$$

where A, B and C are as in Theorem 1.

The condition

$$det\,(B'_\perp A C_\perp) \neq 0 \tag{6}$$

is necessary and sufficient for the existence of P^{-1}.

Further, the following representations of P^{-1} hold

$$P^{-1} = \begin{bmatrix} H & (I - HA)(C')^g \\ B^g(I - AH) & B^g(AHA - A)(C')^g \end{bmatrix} \tag{7}$$

where

$$H = C_\perp (B'_\perp A C_\perp)^{-1} B'_\perp \tag{8}$$

and

$$P^{-1} = \begin{bmatrix} H & K(C'K)^{-1} \\ (\tilde{K}'B)^{-1}\tilde{K}' & -(\tilde{K}'B)^{-1}\tilde{K}'AK(C'K)^{-1} \end{bmatrix} \tag{9}$$

where

$$K = (A'\,B_\perp)_\perp \tag{10}$$

$$\tilde{K} = (AC_\perp)_\perp \tag{11}$$

Proof

Condition (6) follows from the rank identity (see Marsaglia and Styan, 1974, Theorem 19)

$$\begin{aligned} r\,(P) &= r\,(B) + r\,(C) + r\,[(I - BB^g)\,A\,(I - (C')^g\,C')] \\ &= n + n + r\,[(B'_\perp)^g B'_\perp A\; C_\perp (C_\perp)^g] = 2n + r\,(B'_\perp A C_\perp) \end{aligned} \tag{12}$$

where use has been made of the identities (47) and (48) of Section 1.1.

To prove (7), let the inverse of P be

$$P^{-1} = \begin{bmatrix} P_1 & P_2 \\ P_3 & P_4 \end{bmatrix} \tag{13}$$

where the blocks in $\boldsymbol{P}^{-1}$ are of the same order as the corresponding blocks in $\boldsymbol{P}$. Then, in order to express the blocks of the former in terms of the blocks of the latter, write $\boldsymbol{P}^{-1}\boldsymbol{P} = \boldsymbol{I}$ and $\boldsymbol{P}\boldsymbol{P}^{-1} = \boldsymbol{I}$ in partitioned form

$$\begin{bmatrix} \boldsymbol{P}_1 & \boldsymbol{P}_2 \\ \boldsymbol{P}_3 & \boldsymbol{P}_4 \end{bmatrix} \begin{bmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C}' & \boldsymbol{0} \end{bmatrix} = \begin{bmatrix} \boldsymbol{I}_m & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I}_n \end{bmatrix} \tag{14}$$

$$\begin{bmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C}' & \boldsymbol{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{P}_1 & \boldsymbol{P}_2 \\ \boldsymbol{P}_3 & \boldsymbol{P}_4 \end{bmatrix} = \begin{bmatrix} \boldsymbol{I}_m & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I}_n \end{bmatrix} \tag{15}$$

and equate block to block as follows

$$\begin{cases} \boldsymbol{P}_1\boldsymbol{A} + \boldsymbol{P}_2\boldsymbol{C}' = \boldsymbol{I}_m & (16) \\ \boldsymbol{P}_1\boldsymbol{B} = \boldsymbol{0} & (17) \\ \boldsymbol{P}_3\boldsymbol{A} + \boldsymbol{P}_4\boldsymbol{C}' = \boldsymbol{0} & (18) \\ \boldsymbol{P}_3\boldsymbol{B} = \boldsymbol{I}_n & (19) \end{cases}$$

$$\begin{cases} \boldsymbol{A}\boldsymbol{P}_1 + \boldsymbol{B}\boldsymbol{P}_3 = \boldsymbol{I}_m & (16') \\ \boldsymbol{A}\boldsymbol{P}_2 + \boldsymbol{B}\boldsymbol{P}_4 = \boldsymbol{0} & (17') \\ \boldsymbol{C}'\boldsymbol{P}_1 = \boldsymbol{0} & (18') \\ \boldsymbol{C}'\boldsymbol{P}_2 = \boldsymbol{I}_n & (19') \end{cases}$$

From (16) and (16′) we get

$$\boldsymbol{P}_2 = (\boldsymbol{C}')^g - \boldsymbol{P}_1\boldsymbol{A}\,(\boldsymbol{C}')^g = (\boldsymbol{I} - \boldsymbol{P}_1\boldsymbol{A})\,(\boldsymbol{C}')^g \tag{20}$$

$$\boldsymbol{P}_3 = \boldsymbol{B}^g - \boldsymbol{B}^g\boldsymbol{A}\boldsymbol{P}_1 = \boldsymbol{B}^g\,(\boldsymbol{I} - \boldsymbol{A}\boldsymbol{P}_1) \tag{21}$$

respectively.

From (17′), in light of (20) we can write

$$\boldsymbol{P}_4 = -\boldsymbol{B}^g\boldsymbol{A}\boldsymbol{P}_2 = -\boldsymbol{B}^g\boldsymbol{A}\,[(\boldsymbol{C}')^g - \boldsymbol{P}_1\boldsymbol{A}\,(\boldsymbol{C}')^g] = \boldsymbol{B}^g\,[\boldsymbol{A}\boldsymbol{P}_1\boldsymbol{A} - \boldsymbol{A}](\boldsymbol{C}')^g \tag{22}$$

Consider now the equation (17). Solving for $\boldsymbol{P}_1$ gives

$$\boldsymbol{P}_1 = \boldsymbol{V}\boldsymbol{B}'_\perp \tag{23}$$

for some $\boldsymbol{V}$. Substituting the right-hand side of (23) for $\boldsymbol{P}_1$ in (16) and postmultiplying both sides by $\boldsymbol{C}_\perp$ we get

$$V B'_{\perp} A C_{\perp} = C_{\perp} \tag{24}$$

which solved for V yields

$$V = C_{\perp} (B'_{\perp} A C_{\perp})^{-1} \tag{25}$$

in view of (6).

Substituting the right-hand side of (25) for V in (23) we obtain

$$P_1 = C_{\perp} (B'_{\perp} A C_{\perp})^{-1} B'_{\perp} \tag{26}$$

Hence, substituting the right-hand side of (26) for P_1 in (20), (21) and (22) the expressions of the other blocks are easily found.

The proof of (9) follows as a by-product of (7), in light of identity (43) of Section 1.1, upon noticing that, on the one hand

$$\begin{aligned} I - AH = I - (AC_{\perp}) (B'_{\perp} (AC_{\perp}))^{-1} B'_{\perp} = B ((AC_{\perp})'_{\perp}) B^{-1} (AC_{\perp})'_{\perp} \\ = B (\tilde{K}' B)^{-1} \tilde{K}' \end{aligned} \tag{27}$$

whereas, on the other hand,

$$I - HA = K (C'K)^{-1} C' \tag{28}$$

□

The following corollaries provide further results whose usefulness will soon become apparent.

Corollary 2.1

Should both assumptions a) and b) of Theorem 1 hold, then the following equality

$$(A - BD^{-1}C')^{-1} = A^{-1} + A^{-1}B (D - C'A^{-1}B)^{-1} C'A^{-1} \tag{29}$$

ensues.

Proof

Result (29) arises from equating the upper diagonal blocks of the right-hand sides of (2) and (3).

□

Corollary 2.2

Should both assumption a) of Theorem 1 with $D = 0$, and assumption (6) of Theorem 2 hold, then the equality

$$C_{\perp}\,(B_{\perp}'AC_{\perp})^{-1}B_{\perp}' = A^{-1} - A^{-1}B\,(C'A^{-1}B)^{-1}C'A^{-1} \tag{30}$$

ensues.

Proof

Result (30) arises from equating the upper diagonal blocks of the right-hand sides of (2) and (7) for $\boldsymbol{D} = \boldsymbol{0}$.

□

Corollary 2.3

By taking $\boldsymbol{D} = -\lambda\boldsymbol{I}$, let both assumption b) of Theorem 1 in a deleted neighbourhood of $\lambda = 0$, and assumption (6) of Theorem 2 hold. Then the following equality

$$C_{\perp}\,(B_{\perp}'AC_{\perp})^{-1}B_{\perp}' = \lim_{\lambda\to 0}\left\{\lambda\,(\lambda A+BC')^{-1}\right\} \tag{31}$$

ensues as $\lambda \to 0$.

Proof

To prove (31) observe that $\lambda^{-1}\,(\lambda\boldsymbol{A}+\boldsymbol{BC}')$ plays the rôle of Schur complement of $\boldsymbol{D} = -\lambda\boldsymbol{I}$ in the partitioned matrix

$$\begin{bmatrix} A & B \\ C' & -\lambda I \end{bmatrix} \tag{32}$$

whence

$$\left\{\lambda\,(\lambda A+BC')^{-1}\right\} = [I,\ 0]\begin{bmatrix} A & B \\ C' & -\lambda I \end{bmatrix}^{-1}\begin{bmatrix} I \\ 0 \end{bmatrix} \tag{33}$$

Taking the limit as $\lambda \to 0$ of both sides of (33) yields

$$\lim_{\lambda\to 0}\left\{\lambda\,(\lambda A+BC')^{-1}\right\} = [I,\ 0]\begin{bmatrix} A & B \\ C' & 0 \end{bmatrix}^{-1}\begin{bmatrix} I \\ 0 \end{bmatrix} \tag{34}$$

which eventually leads to (31) in view of (7).

□

1.3 Matrix Polynomials: Preliminaries

We start by introducing the following definitions

Definition 1

A matrix polynomial of degree K in the scalar argument z is an expression of the form

$$A(z) = \sum_{k=0}^{K} A_k z^k, \quad A_K \neq 0 \tag{1}$$

In the following we assume, unless otherwise stated, that $A_0, A_1, \ldots, A_K$ are square matrices of order n.

When $K = 1$ the matrix polynomial is said to be linear.

Definition 2

The scalar polynomial

$$\pi(z) = det\, A(z) \tag{2}$$

is referred to as the characteristic polynomial of $A(z)$.

Definition 3

The algebraic equation

$$\pi(z) = 0 \tag{3}$$

is referred to as the characteristic equation of the matrix polynomial $A(z)$.

Expanding the matrix polynomial $A(z)$ about $z = 1$ yields

$$A(z) = A(1) + \sum_{k=1}^{K} (1-z)^k (-1)^k \frac{1}{k!} A^{(k)}(1) \tag{4}$$

where

$$A^{(k)}(1) = \left. \frac{d^k A(z)}{d z^k} \right]_{z=1} = k! \sum_{j=k}^{K} \binom{j}{k} A_j \tag{5}$$

The dot matrix notation $\dot{A}(z)$, $\ddot{A}(z)$, $\dddot{A}(z)$ will be adopted for $k = 1, 2, 3$. For simplicity of notation, A, $\dot{A}$, $\ddot{A}$, $\dddot{A}$ will henceforth be written instead of $A(1)$, $\dot{A}(1)$, $\ddot{A}(1)$, $\dddot{A}(1)$.

The following truncated expansions of (4)

$$\boldsymbol{A}(z) = \boldsymbol{A} + (1-z)\,\boldsymbol{Q}(z) \tag{6}$$

$$\boldsymbol{A}(z) = \boldsymbol{A} - (1-z)\,\dot{\boldsymbol{A}} + (1-z)^2\,\boldsymbol{\Psi}(z) \tag{7}$$

where

$$\boldsymbol{Q}(z) = \sum_{k=1}^{K}(1-z)^{k-1}(-1)^k\,\frac{1}{k!}\boldsymbol{A}^{(k)}(1), \quad \boldsymbol{Q}(1) = -\dot{\boldsymbol{A}} \tag{8}$$

$$\boldsymbol{\Psi}(z) = \sum_{k=2}^{K}(1-z)^{k-2}\,(-1)^k\,\frac{1}{k!}\boldsymbol{A}^{(k)}(1), \quad \boldsymbol{\Psi}(1) = -\,\dot{\boldsymbol{Q}}(1) = \frac{1}{2}\,\ddot{\boldsymbol{A}} \tag{9}$$

are of special interest for the subsequent analysis.

We prove the following classical result.

Theorem 1

The characteristic polynomial $det\,\boldsymbol{A}(z)$ has a possibly multiple unit-root $z = 1$ if and only if

$$det\,\boldsymbol{A} = 0 \tag{10}$$

Proof

According to (6) the characteristic equation (3) can be exhibited in the form

$$det\,[(1-z)\,\boldsymbol{Q}(z) + \boldsymbol{A}] = 0 \tag{11}$$

which for $z = 1$ entails

$$det\,(\boldsymbol{A}) = 0 \Rightarrow r\,(\boldsymbol{A}) < n \tag{12}$$

and the other way around.

□

The next result sheds more light on the characteristic polynomial roots.

Theorem 2

We distinguish two possibilities

i) $z = 1$ is a simple root of the characteristic polynomial $det\,\boldsymbol{A}(z)$ if and only if

$$det\,\boldsymbol{A} = 0 \tag{13}$$

$$tr\,(A^{+}(1)\dot{A}\,) \neq 0 \tag{14}$$

where $A^{+}(1)$ denotes the adjoint matrix $A^{+}(z)$ of $A\ (z)$ evaluated at $z = 1$;

ii) $z = 1$ is a root of multiplicity two of the characteristic polynomial $det\,A\ (z)$ if and only if

$$det\,A = 0 \tag{15}$$

$$tr\,(A^{+}(1)\dot{A}\,) = 0 \tag{16}$$

$$tr\,(\dot{A}^{+}(1)\dot{A} \ + A^{+}(1)\ddot{A}\,) \neq 0 \tag{17}$$

where $\dot{A}^{+}(1)$ denotes the derivative of $\dot{A}^{+}(z)$ with respect to z evaluated at $z = 1$.

Proof

Expanding $det\,A\ (z)$ about $z = 1$ yields

$$det\,A\ (z) = det\,A - (1-z)\left[\frac{\mathrm{d}\,det\,A(z)}{\mathrm{d}z}\right]_{z=1} + (1-z)^2\left[\frac{\mathrm{d}^2\,det\,A(z)}{\mathrm{d}z^2}\right]_{z=1}$$

$$+\ terms\ of\ higher\ powers\ of\ (1-z) \tag{18}$$

$$= det\,A - (1-z)\,tr\,(A^{+}(1)\dot{A}\,) + (1-z)^2\,tr\,(\dot{A}^{+}(1)\dot{A} + (A^{+}(1)\ddot{A}\,)$$

$$+\ terms\ of\ higher\ powers\ of\ (1-z)$$

by virtue of

$$\left[\frac{\mathrm{d}\,det\,A(z)}{\mathrm{d}z}\right] = \left[\frac{\mathrm{d}\,det\,A(z)}{\mathrm{d}\,vec\,A(z)}\right]' \frac{\mathrm{d}\,vec\,A(z)}{\mathrm{d}z} \tag{19}$$

$$= \left\{vec(A^{+}(z))'\right\}' vec\,\dot{A}(z) \ = tr\ \{A^{+}(z)\ \dot{A}\ (z)\}$$

$$\left[\frac{\mathrm{d}^2\,det\,A(z)}{\mathrm{d}z^2}\right] = \frac{\mathrm{d}\,tr\left\{A^{+}(z)\dot{A}(z)\right\}}{\mathrm{d}z} = tr\,(\,\dot{A}^{+}(z)\,\dot{A}^{+}(z) + A^{+}(z)\,\ddot{A}\ (z)) \tag{20}$$

where use has been made of matrix differentiation rules and *vec* vs. *trace* relationships (see, e.g., Faliva, 1975, 1987; Magnus and Neudecker, 1999).

In view of (18) both statements i) and ii) clearly hold true.

□

1.4 Matrix Polynomial Inversion by Laurent Expansion

In this section the reader will find the essentials of matrix polynomial inversion about a pole, a topic whose technicalities will extend over the forthcoming sections, to duly cover analytical demands of dynamic model econometrics in Chapter 3.

Theorem 1

Let the roots of the characteristic polynomial

$$\pi\,(z) = det\, \boldsymbol{A}(z) \tag{1}$$

lie either outside or on the unit circle and, in the latter case, be equal to one. Then the inverse of the matrix polynomial $\boldsymbol{A}(z)$ admits the Laurent expansion

$$\boldsymbol{A}^{-1}(z) = \underbrace{\sum_{j=1}^{H \le K} \frac{1}{(1-z)^j} \boldsymbol{N}_j}_{\text{principal part}} + \underbrace{\sum_{i=0}^{\infty} z^i\, \boldsymbol{M}_i}_{\text{regular part}} \tag{2}$$

in a deleted neighbourhood of $z = 1$, where the coefficient matrices $\boldsymbol{M}_i$ of the regular part consist of exponentially decreasing entries, and the coefficient matrices $\boldsymbol{N}_j$ of the principal part vanish if $\boldsymbol{A}$ is of full rank.

Proof

The statement of the theorem can be read as a matrix extension of classical results of Laurent series theory (see, e.g., Jeffrey, 1992; Markuscevich, 1965). A deeper insight into the subject will be gained through Theorem 4 at the end of this section.

□

For further analysis we will need the following

Definition 1

An isolated point z_0 of a (matrix) function $\boldsymbol{A}^{-1}(z)$ such that the Euclidian norm $\left\| \boldsymbol{A}^{-1}(z) \right\| \to \infty$ as $z = z_0$ is called a pole of $\boldsymbol{A}^{-1}(z)$.

If $z = z_0$ is not a pole of $\boldsymbol{A}^{-1}(z)$, the function $\boldsymbol{A}^{-1}(z)$ is olomorphic (analytical) in a neighbourhood of the point z_0.

Definition 2

The point z_0 is a pole of order H of the (matrix) function $A^{-1}(z)$ if and only if the principal part of the Laurent expansion of $A^{-1}(z)$ about z_0 contains a finite number of terms forming a polynomial of degree H in $(z_0 - z)^{-1}$, i.e. if and only if $A^{-1}(z)$ admits the Laurent expansion

$$A^{-1}(z) = \sum_{j=1}^{H} \frac{1}{(z_0 - z)^j} N_j + \sum_{i=0}^{\infty} z^i M_i, \quad N_H \neq 0 \tag{3}$$

in a deleted neighbourhood of z_0.

When $H = 1$ the pole located at z_0 is referred to as a simple pole.

Observe that, if (3) holds true, then both the matrix function $(z_0 - z)^H A^{-1}(z)$ and its derivatives have finite limits as z tends to z_0, the former N_H being a non null matrix.

Definition 3

The point z_0 is a zero of order H of the matrix polynomial $A(z)$ if and only if z_0 is a pole of order H of the meromorphic matrix function $A^{-1}(z)$ (see also Theorem 2 of Section 1.3).

The simplest form of the Laurent expansion (2) is

$$A^{-1}(z) = \frac{1}{(1-z)} N_1 + M(z) \tag{4}$$

which corresponds to the case of a simple pole at $z = 1$ where

$$N_1 = \lim_{z \to 1} [(1-z) A^{-1}(z)] \tag{5}$$

$$M(1) = -\lim_{z \to 1} \frac{\mathrm{d}[(1-z) A^{-1}(z)]}{\mathrm{d}z} \tag{6}$$

and $M(z)$ stands for $\sum_{i=0}^{\infty} z^i M_i$.

Theorem 2

The matrix N_1 is singular.

Proof

Since the equalities

$$A(z)\,A^{-1}(z) = I \Leftrightarrow [(1-z)\,Q\,(z) + A][\frac{1}{(1-z)}N_1 + M(z)] = I \tag{7}$$

$$A^{-1}(z)\,A(z) = I \Leftrightarrow [\frac{1}{(1-z)}N_1 + M(z)]\,[(1-z)\,Q\,(z) + A] = I \tag{8}$$

hold true in a deleted neighbourhood of $z = 1$, the term containing the negative power of $(1-z)$ in the left-hand sides of (7) and (8) must vanish. This occurs as long as N_1 satisfies the twin conditions

$$AN_1 = 0 \tag{9}$$

$$N_1A = 0 \tag{10}$$

which, in turn, entails the singularity of N_1 (we rule out the case of a null A).

□

Another case of the Laurent expansion (2) which turns out to be of prominent interest is

$$A^{-1}(z) = \sum_{j=1}^{2}\frac{1}{(1-z)^j}N_j + M\,(z) \tag{11}$$

which corresponds to a second order pole at $z = 1$, where

$$N_2 = \lim_{z\to 1}[(1-z)^2 A^{-1}(z)] \tag{12}$$

$$N_1 = -\lim_{z\to 1}\frac{\mathrm{d}[(1-z)^2\,A^{-1}(z)]}{\mathrm{d}z} \tag{13}$$

$$M(1) = \frac{1}{2}\lim_{z\to 1}\frac{\mathrm{d}^2[(1-z)^2\,A^{-1}(z)]}{\mathrm{d}z^2} \tag{14}$$

In this connection we have the following

Theorem 3

The matrix N_2 is singular

Proof

Since the equalities

$$A(z)\,A^{-1}(z) = I \Leftrightarrow$$
$$\Leftrightarrow [(1-z)^2\Psi(z) - (1-z)\,\dot{A} + A][\frac{1}{(1-z)^2}N_2 + \frac{1}{(1-z)}N_1 + M(z)] = I \tag{15}$$

$$A^{-1}(z)\,A(z) = I \Leftrightarrow$$
$$\Leftrightarrow [\frac{1}{(1-z)^2}N_2 + \frac{1}{(1-z)}N_1 + M(z)]\,[(1-z)\Psi(z) - (1-z)\,\dot{A} + A] = I \tag{16}$$

hold true in a deleted neighbourhood of $z = 1$, the terms containing the negative powers of $(1 - z)$ in the left-hand sides of (15) and (16) must vanish. This occurs provided N_2 and N_1 satisfy the following set of conditions

$$AN_2 = 0 \tag{17}$$

$$N_2 A = 0 \tag{18}$$

$$\dot{A}\,N_2 = AN_1 \tag{19}$$

$$N_2\,\dot{A} = N_1 A \tag{20}$$

Equalities (17) and (18), in turn, entail the singularity of N_2.

□

Finally, the next result leads to a deeper insight as far as the algebraic premises of expansion (2) are concerned.

Theorem 4

Under the assumptions of Theorem 1 about the roots of the characteristic polynomial $det\,A(z)$, in a deleted neighbourhood of $z = 1$ the matrix function $A^{-1}(z)$ admits the expansion

$$A^{-1}(z) = \sum_{j=1}^{H} \frac{1}{(1-z)^j} N_j + M\,(z) \tag{21}$$

where H is a non negative integer and

$$M\,(z) = \sum_{i=0}^{\infty} z^i\, M_i \tag{22}$$

with the entries of the coefficient matrices M_i decreasing exponentially.

Proof

First of all observe that, on the one hand, the factorization

$$det\, A(z) = k\,(1-z)^{a}\,\Pi_j(1-\frac{z}{z_j}) \tag{23}$$

holds for $det\, A(z)$, where α is a non negative integer, the $z_j' s$ denote the roots lying outside the unit circle ($|z_j| > 1$) and k is a suitably chosen scalar. On the other hand, the partial fraction expansion

$$\{det\, A\,(z)\}^{-1} = \sum_{i=1}^{a} \lambda_i \frac{1}{(1-z)^i} + \sum_j \mu_j \frac{1}{1-\dfrac{z}{z_j}} \tag{24}$$

holds for the reciprocal of $det\, A(z)$ accordingly, where the $\lambda_i' s$ and the $\mu_j' s$ are properly chosen coefficients, under the assumption that the roots $z_j' s$ are real and simple for algebraic convenience. Should some roots be complex and/or repeated, the expansion still holds with the addition of rational terms whose numerators are linear in z whereas the denominators are higher order polynomials in z (see, e.g. Jeffrey, 1992, p. 382). This, apart from algebraic burdening, does not ultimately affect the conclusions drawn in the theorem.

Insofar as $|z_j| > 1$, a power expansion of the form

$$\left(1-\frac{z}{z_j}\right)^{-1} = \sum_{k=0}^{\infty} (z_j)^{-k} z^k \tag{25}$$

holds for $|z| \le 1$.

This together with (24) lead to the conclusion that $\{det\, A(z)\}^{-1}$ can be written in the form

$$\begin{aligned}\{det\, A(z)\}^{-1} &= \sum_{i=1}^{a} \lambda_i \frac{1}{(1-z)^i} + \sum_j \mu_j \sum_{k=0}^{\infty} (z_j)^{-k} z^k \\ &= \sum_{i=1}^{a} \lambda_i \frac{1}{(1-z)^i} + \sum_{k=0}^{\infty} \eta_k z^k\end{aligned} \tag{26}$$

where the $\eta_k' s$ are exponentially decreasing weights depending on the $\mu_j' s$ and the $z_j' s$.

Now, provided $A^{-1}(z)$ exists in a deleted neighbourhood of $z = 1$, it can be expressed in the form

$$A^{-1}(z) = \{det\, A(z)\}^{-1} A^{+}(z) \tag{27}$$

where the adjoint matrix $A^{+}(z)$ can be expanded about $z = 1$ yielding

$$A^{+}(z) = A^{+}(1) - \dot{A}^{+}(1)(1 - z) + \textit{terms of higher powers of}\,(1 - z) \tag{28}$$

Substituting the right-hand sides of (26) and (28) for $\{det\, A(z)\}^{-1}$ and $A^{+}(z)$ respectively into (27), we can eventually express $A^{-1}(z)$ in the form (21), where the exponentially decay property of the regular part matrices M_i is a by-product of the aforesaid property of the coefficients η_k' s.

□

1.5 Matrix Polynomials and Difference Equation Systems

Insofar as the algebra of polynomial functions of the complex variable z and the algebra of polynomial functions of the lag operator L are isomorphic (see, e.g., Dhrymes, 1971, p. 23), the arguments developed in the previous sections provide an analytical tool-kit paving the way to find elegant closed-form solutions to finite difference equation systems which are of prominent interest in econometrics.

Indeed, a non homogeneous linear system of difference equations with constant coefficients can be conveniently written in operator form as follows

$$A\,(L)\,y_t = g_t \tag{1}$$

where g_t is a given real valued function commonly called *forcing* function in mathematical physics (see, e.g., Vladimirov, 1984, p. 38), L is the lag operator defined by the relations

$$Ly_t = y_{t-1}, \quad L^0 y_t = y_t, \quad L^K y_t = y_{t-K} \tag{2}$$

with K denoting an arbitrary integer and $A(L)$ is a matrix polynomial in the argument L, defined as

$$A(z) = \sum_{k=0}^{K} A_k\, L^k \tag{3}$$

where $A_0, A_1, \ldots, A_K$ are matrices of constant coefficients.

By replacing g_t by $\boldsymbol{0}$ we obtain the homogeneous equation corresponding to (1), otherwise known as reduced equation.

Any solution of the nonhomogeneous equation (1) will be referred to as a particular solution, whereas the general solution of the reduced equation will be referred to as the complementary solution. The latter turns out to depend on the roots z_j of the characteristic equation

$$det\, A(z) = 0 \tag{4}$$

via the solutions $\boldsymbol{h}_j$ of the generalized eigenvector problem

$$A(z_j)\, \boldsymbol{h}_j = \boldsymbol{0} \tag{5}$$

Before further investigating the issue of how to handle equation (1) some special purpose analytical tooling is needed.

As pointed out in Section 1.4, the following Laurent expansions hold for the meromorphic matrix function $A^{-1}(z)$ in a deleted neighbourhood of $z = 1$

$$A^{-1}(z) = \frac{1}{(1-z)} N_1 + M(z) \tag{6}$$

$$A^{-1}(z) = \frac{1}{(1-z)^2} N_2 + \frac{1}{(1-z)} N_1 + M(z) \tag{7}$$

under the case of a simple pole and a second order pole, located at $z = 1$, respectively.

Thanks to the said isomorphism, by replacing 1 by the identity operator I and z by the lag operator L, we obtain the counterparts of the expansions (6) and (7) in operator form, namely

$$A^{-1}(L) = \frac{1}{(I-L)} N_1 + M\,(L) \tag{8}$$

$$A^{-1}(L) = \frac{1}{(I-L)^2} N_2 + \frac{1}{(I-L)} N_1 + M(L) \tag{9}$$

Let us now introduce a few operators related to L which play a crucial rôle in the study of the difference equations we are primarily interested in. For these and related results see Elaydi (1996) and Mickens (1990).

Definition 1 – Backward difference operator

The backward difference operator, denoted by ∇, is defined by the relation

$$\nabla = I - L \tag{10}$$

Higher order operators ∇^K are defined as follows:

$$\nabla^K = (I - L)^K, \; K = 2, 3 \ldots \tag{11}$$

whereas $\nabla^0 = I$.

Definition 2 – Antidifference or indefinite sum operator

The antidifference operator, denoted by ∇^{-1} – otherwise known as indefinite sum operator and written as Σ – is defined as the operator such that the identity

$$(I - L)\, \nabla^{-1}\boldsymbol{x}_t = \boldsymbol{x}_t \tag{12}$$

holds true for arbitrary $\boldsymbol{x}_t$, which is tantamount to saying that ∇^{-1} acts as a right inverse of $I - L$.

Higher order operators, ∇^{-K}, are defined accordingly, i.e.

$$(I - L)^K\, \nabla^{-K} = I \tag{13}$$

In light of the identities (12) and (13), insofar as a K-order difference operator annihilates a $(K - 1)$-degree polynomial, the following hold

$$\nabla^{-1}\boldsymbol{0} = \boldsymbol{c} \tag{14}$$

$$\nabla^{-2}\boldsymbol{0} = \boldsymbol{c}t + \boldsymbol{d} \tag{15}$$

where $\boldsymbol{c}$ and $\boldsymbol{d}$ are arbitrary.

We now state without proof the well-known result of

Theorem 1

The general solution of the nonhomogeneous equation (1) consists of the sum of any particular solution of the given equation and of the complementary solution.

Because of the foregoing arguments, we are able to establish the following elegant results.

Theorem 2

A particular solution of the nonhomogeneous equation (1) can be expressed in operator form as

$$\boldsymbol{y}_t = \boldsymbol{A}^{-1}(L)\, \boldsymbol{g}_t \tag{16}$$

In particular, the following hold true

i)
$$y_t = N_1 \nabla^{-1} g_t + M(L) g_t = N_1 \sum_{\tau \le t} g_\tau + M(L) g_t \tag{17}$$

if $z = 1$ is a simple pole of $A^{-1}(z)$;

ii)
$$\begin{aligned} y_t &= N_2 \nabla^{-2} g_t + N_1 \nabla^{-1} g_t + M(L) g_t \\ &= N_2 \sum_{\vartheta \le t} \sum_{\tau \le \vartheta} g_\tau + N_1 \sum_{\tau \le t} g_\tau + M(L) g_t \\ &= N_2 \sum_{\tau \le t} (t+1-\tau) g_\tau + N_1 \sum_{\tau \le t} g_\tau + M(L) g_t \end{aligned} \tag{18}$$

if $z = 1$ is a second order pole of $A^{-1}(z)$.

Proof

Clearly, the right-hand side of (16) is a solution provided $A^{-1}(L)$ is a meaningful operator. Indeed, this is the case for $A^{-1}(L)$ as defined in (8) and in (9) for a simple and a second order pole at $z = 1$, respectively.

To prove the second part of the theorem observe first that in view of Definitions 1 and 2, the following operator identities hold

$$\frac{1}{1-L} = \nabla^{-1} = \sum \rightarrow \frac{1}{1-L} x_t = \sum_{\tau \le t} x_\tau \tag{19}$$

$$\frac{1}{(1-L)^2} = \nabla^{-2} = \sum \sum \rightarrow \frac{1}{(1-L)^2} x_t = \sum_{\vartheta \le t} \sum_{\tau \le \vartheta} x_\tau \tag{20}$$

where x_t is arbitrary. Further, simple sum-calculus rules show that

$$\sum_{\vartheta \le t} \sum_{\tau \le \vartheta} x_\tau = \sum_{\tau \le t} (t+1-\tau) x_\tau = \sum_{\tau \ge 0} (\tau+1) x_{t-\tau} \tag{21}$$

Thus, in view of expansions (8) and (9) and the foregoing identities, statements i) and ii) are easily established.

□

Theorem 3

The solution of the reduced equation

$$A(L) z_t = \mathbf{0} \tag{22}$$

corresponding to the unit root $z = 1$ can be written in operator form as

$$z_t = \boldsymbol{A}^{-1}(L)\,\boldsymbol{0} \tag{23}$$

where the operator $\boldsymbol{A}^{-1}(L)$ is defined as in (8) or in (9), depending upon the order (first vs. second, respectively) of the pole of $\boldsymbol{A}^{-1}(z)$ at $z = 1$.

Finally, the following closed-form expressions of the solution hold

$$z_t = \boldsymbol{N}_1\boldsymbol{c} \tag{24}$$

$$z_t = \boldsymbol{N}_2\boldsymbol{c}t + \boldsymbol{N}_2\boldsymbol{d} + \boldsymbol{N}_1\boldsymbol{c} \tag{25}$$

for a first and a second order pole respectively, with $\boldsymbol{c}$ and $\boldsymbol{d}$ arbitrary vectors.

Proof

The proof follows from arguments similar to those of Theorem 2 by making use of results (14) and (15) above.

□

Theorem 4

The solution of the reduced equation

$$\boldsymbol{A}(L)\,z_t = \boldsymbol{0} \tag{26}$$

corresponding to unit-roots is a polynomial of the same degree as the order, reduced by one, of the pole of $\boldsymbol{A}^{-1}(z)$ at $z = 1$.

Proof

Should $z = 1$ be either a simple or a second order pole of $\boldsymbol{A}^{-1}(z)$, then Theorem 3 trivially applies. The proof for a higher order pole follows along the same lines.

□

Two additional results, whose rôle will become apparent later on, are worth stating.

Theorem 5

Let $z = 1$ be a simple pole of $\boldsymbol{A}^{-1}(z)$ and

$$\boldsymbol{N}_1 = \boldsymbol{F}\boldsymbol{G}' \tag{27}$$

a rank factorization of $\boldsymbol{N}_1$. Then the following hold

$$\boldsymbol{F}_\perp' \boldsymbol{y}_t = \boldsymbol{F}_\perp' \boldsymbol{M}(L)\,\boldsymbol{g}_t \tag{28}$$

$$F'_{\perp} z_t = 0 \tag{29}$$

where y_t and z_t are as determined by (17) and (24) of Theorems 2 and 3, respectively.

Proof

The proof is simple and is omitted.

□

Theorem 6

Let $z = 1$ be a second order pole of $A^{-1}(z)$ and

$$N_2 = HK' \tag{30}$$

a rank factorization of N_2. Then the following hold

$$H'_{\perp} y_t = H'_{\perp} N_1 \sum_{\tau \le t} g_\tau + H'_{\perp} M(L)\, g_t \tag{31}$$

$$H'_{\perp} z_t = H'_{\perp} N_1 c \tag{32}$$

where y_t and z_t are as determined by (18) and (25) of Theorem 2 and 3, respectively.

Besides, should $[N_2, N_1]$ not be of full row-rank and

$$[N_2, N_1] = JL' \tag{33}$$

represent a rank factorization of the same, then the following hold

$$J'_{\perp} y_t = J'_{\perp} M(L)\, g_t \tag{34}$$

$$J'_{\perp} z_t = 0 \tag{35}$$

where y_t and z_t are as above.

Proof

The proof is simple and is omitted.

□

1.6 Matrix Coefficient Rank Properties vs. Pole Order in Matrix Polynomial Inversion

This section will be devoted to presenting several relationships between rank characteristics of the matrices in the Taylor expansion of a matrix polynomial, $A(z)$, about $z = 1$, and the order of the poles inherent in the Laurent expansion of its inverse, $A^{-1}(z)$, in a deleted neighbourhood of $z = 1$.

Basically, references will be made to Sections 1.3 and 1.4 for notational purposes as well as for relevant expansions.

Theorem 1

The inverse $A^{-1}(z)$ of the matrix polynomial $A(z)$ is an analytical (matrix) function about $z = 1$ if and only if

$$det\, A \neq 0 \tag{1}$$

Under (1), the point $z = 1$ is neither a pole of $A^{-1}(z)$ nor a zero of $A(z)$.

Proof

The theorem mirrors the concluding remark of the statement in Theorem 1 of Section 1.4. See also Theorem 1 of Section 1.3.

□

Theorem 2

The inverse, $A^{-1}(z)$, of the matrix polynomial $A(z)$ has a simple pole at $z = 1$ provided the following conditions are satisfied

i) $$det\, A = 0, \quad A \neq \mathbf{0} \tag{2}$$

ii) $$det \begin{bmatrix} -\dot{A} & B \\ C' & \mathbf{0} \end{bmatrix} \neq 0 \tag{3}$$

where B and C are full column-rank matrices obtained by rank factorization of A, i.e.

$$A = BC', \quad r(A) = r(B) = r(C) \tag{4}$$

Proof

From (6) of Section 1.3 and (4) above, it follows that

$$\frac{1}{(1-z)}A(z) = \frac{1}{(1-z)}[(1-z)\,Q\,(z) + BC'] \tag{5}$$

where $Q\,(z)$ is as defined in (8) of Section 1.3.

We notice now that the right-hand side of (5) corresponds to the Schur complement of the lower diagonal block, $(z-1)\,I$, in the partitioned matrix

$$P\,(z) = \begin{bmatrix} Q(z) & B \\ C' & (z-1)I \end{bmatrix} \tag{6}$$

Hence, by (3) of Theorem 1 of Section 1.2, the following holds

$$(1-z)\,A^{-1}(z) = [I \quad 0]\begin{bmatrix} Q(z) & B \\ C' & (z-1)I \end{bmatrix}^{-1}\begin{bmatrix} I \\ 0 \end{bmatrix} \tag{7}$$

provided $det\,P\,(z) \neq 0$.

By virtue of condition ii), by taking the limit of both sides of (7) as z tends to 1, the outcome would be

$$\lim_{z\to 1}[(1-z)\,A^{-1}(z)] = [I \;\; 0]\begin{bmatrix} Q(1) & B \\ C' & 0 \end{bmatrix}^{-1}\begin{bmatrix} I \\ 0 \end{bmatrix} = [I \;\; 0]\begin{bmatrix} -\dot{A} & B \\ C' & 0 \end{bmatrix}^{-1}\begin{bmatrix} I \\ 0 \end{bmatrix} \tag{8}$$

which, in view of Definition (2), together with (5), of Section 1.4 leads to conclude that $z = 1$ is a simple pole of $A^{-1}(z)$.

□

Corollary 2.1

The following statements are equivalent to condition ii)

a) $$det\,(B'_{\perp}\,\dot{A}\,C_{\perp}) \neq 0 \tag{9}$$

b) $$r\left(\begin{bmatrix} \dot{A} & A \\ A & 0 \end{bmatrix}\right) = n + r\,(A) \tag{10}$$

Proof

Equivalence of ii) and a) follows from Theorem 2 of Section 1.2. Equivalence of ii) and b) is easily proved upon noticing that

$$det \begin{bmatrix} -\dot{A} & B \\ C' & 0 \end{bmatrix} \neq 0 \Leftrightarrow r\left(\begin{bmatrix} -\dot{A} & B \\ C' & 0 \end{bmatrix}\right) = n + r\,(C) = n + r\,(A) \tag{11}$$

Indeed the following hold

$$\begin{aligned} r\left(\begin{bmatrix} -\dot{A} & B \\ C' & 0 \end{bmatrix}\right) &= r\left(\begin{bmatrix} -I & 0 \\ 0 & B \end{bmatrix}\begin{bmatrix} -\dot{A} & B \\ C' & 0 \end{bmatrix}\begin{bmatrix} I & 0 \\ 0 & -C' \end{bmatrix}\right) = r\left(\begin{bmatrix} \dot{A} & A \\ A & 0 \end{bmatrix}\right) \\ &= r\,(A) + r\,(A) + r((I - AA^{g})\,\dot{A}\,(I - A^{g}A)) = r\,(A) + r\,(B) \\ &\quad + r\,(B_{\perp}'\,\dot{A}\,C_{\perp}) = r\,(A) + r\,(B) + r(B_{\perp}) = r\,(A) + n \end{aligned} \tag{12}$$

in light of Theorem 19 in Marsaglia and Styan, 1974, and identities (47) and (48) of Section 1.1.

□

Theorem 3

The inverse $A^{-1}(z)$ of the matrix polynomial $A\,(z)$ has a second order pole at $z = 1$ provided the following conditions are satisfied

i) $$det\,A = 0, \quad A \neq 0 \tag{13}$$

ii) $$det\,(B_{\perp}'\,\dot{A}\,C_{\perp}) = 0, \quad B_{\perp}'\,\dot{A}\,C_{\perp} \neq 0 \tag{14}$$

where B and C are full column-rank matrices obtained by rank factorization (4) of A.

iii) $$det \begin{bmatrix} \tilde{A} & (B_{\perp}R_{\perp})_{\perp} \\ (C_{\perp}S_{\perp})_{\perp}' & 0 \end{bmatrix} \neq 0 \tag{15}$$

where R and S are full column-rank matrices obtained by rank factorization of $B_{\perp}'\,\dot{A}\,C_{\perp}$, i.e.

$$B_{\perp}'\,\dot{A}\,C_{\perp} = RS', \quad r\,(B_{\perp}'\,\dot{A}\,C_{\perp}) = r\,(S) = r\,(R) \tag{16}$$

and $\tilde{A}$ is the matrix

$$\tilde{A} = \frac{1}{2}\ddot{A} - \dot{A}A^{g}\dot{A} \tag{17}$$

Proof

From (7) of Section 1.3 and (4) above, it follows that

$$\frac{1}{(1-z)^2}A(z) = \frac{1}{(1-z)^2}[(1-z)^2\Psi(z) - (1-z)\dot{A} + BC'] \tag{18}$$

where $\Psi(z)$ is as defined in (9) of Section 1.3.

Pre and postmultiplying $\dot{A}$ by $(B')^g B' + (B'_\perp)^g B'_\perp = I$ and by $C(C)^g + C_\perp(C_\perp)^g = I$ (cf. identity (42) of Section 1.1) yields

$$\begin{aligned}\dot{A} &= (B'_\perp)^g B'_\perp \dot{A} C_\perp (C_\perp)^g + (B'_\perp)^g B'_\perp \dot{A}(C')^g C' + BB^g\dot{A} \\ &= (B'_\perp)^g RS'(C_\perp)^g + (I - BB^g)\dot{A}(C')^g C' + BB^g\dot{A}\end{aligned} \tag{19}$$

where use has been made of (16) above.

Substituting the right-hand side of (19) for $\dot{A}$ into (18) and putting

$$F = [(B'_\perp)^g R, \;\; B, \;\; \dot{A}(C')^g], \;\; G' = \begin{bmatrix} S'(C_\perp)^g \\ C' \\ B^g\dot{A} \end{bmatrix} \tag{20}$$

$$V(z) = \begin{bmatrix} (1-z)I & 0 & 0 \\ 0 & -I-(1-z)B^g\dot{A}(C')^g & (1-z)I \\ 0 & (1-z)I & 0 \end{bmatrix} \tag{21}$$

$$\Lambda(z) = (1-z)^2V^{-1}(z) = \begin{bmatrix} (1-z)I & 0 & 0 \\ 0 & 0 & (1-z)I \\ 0 & (1-z)I & (1-z)B^g\dot{A}(C')^g + I \end{bmatrix} \tag{22}$$

we can rewrite (18) in the form

$$\frac{1}{(1-z)^2}A(z) = \frac{1}{(1-z)^2}[(1-z)^2\Psi(z) - FV(z)G'] \tag{23}$$

We notice that the right-hand side of (23) corresponds to the Schur complement of the lower diagonal block, $\Lambda(z)$, of the partitioned matrix

$$\boldsymbol{P}(z) = \begin{bmatrix} \Psi(z) & \boldsymbol{F} \\ \boldsymbol{G}' & \Lambda(z) \end{bmatrix} \tag{24}$$

Hence, by (3) of Theorem 1 in Section 1.2, the following holds

$$(1-z)^2\boldsymbol{A}^{-1}(z) = [\boldsymbol{I} \quad \boldsymbol{0}] \begin{bmatrix} \Psi(z) & \boldsymbol{F} \\ \boldsymbol{G}' & \Lambda(z) \end{bmatrix}^{-1} \begin{bmatrix} \boldsymbol{I} \\ \boldsymbol{0} \end{bmatrix} \tag{25}$$

provided $det\, \boldsymbol{P}(z) \neq 0$.

Further, observe that

$$\boldsymbol{P}(1) = \begin{bmatrix} \Psi(1) & \boldsymbol{F} \\ \boldsymbol{G}' & \Lambda(1) \end{bmatrix} = \begin{bmatrix} \Psi(1) & (\boldsymbol{B}_\perp')^g\boldsymbol{R} & \boldsymbol{B} & \vdots & \dot{\boldsymbol{A}}(\boldsymbol{C}')^g \\ \boldsymbol{S}'\boldsymbol{C}_\perp^g & \boldsymbol{0} & \boldsymbol{0} & \vdots & \boldsymbol{0} \\ \boldsymbol{C}' & \boldsymbol{0} & \boldsymbol{0} & \vdots & \boldsymbol{0} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \boldsymbol{B}^g\dot{\boldsymbol{A}} & \boldsymbol{0} & \boldsymbol{0} & \vdots & \boldsymbol{I} \end{bmatrix} \tag{26}$$

$$= \begin{bmatrix} \frac{1}{2}\ddot{\boldsymbol{A}} & (\boldsymbol{B}_\perp\boldsymbol{R}_\perp)_\perp & \vdots & \dot{\boldsymbol{A}}(\boldsymbol{C}')^g \\ (\boldsymbol{C}_\perp\boldsymbol{S}_\perp)_\perp' & \boldsymbol{0} & \vdots & \boldsymbol{0} \\ \cdots & \cdots & \vdots & \cdots \\ \boldsymbol{B}^g\dot{\boldsymbol{A}} & \boldsymbol{0} & \vdots & \boldsymbol{I} \end{bmatrix}$$

in light of (9) of Section 1.3 and equalities

$$[(\boldsymbol{B}_\perp')^g\boldsymbol{R}, \quad \boldsymbol{B}] = (\boldsymbol{B}_\perp\boldsymbol{R}_\perp)_\perp \tag{27}$$

$$[(\boldsymbol{C}_\perp')^g\boldsymbol{S}, \quad \boldsymbol{C}] = (\boldsymbol{C}_\perp\boldsymbol{S}_\perp)_\perp \tag{28}$$

as per Theorem 6, recalling Theorem 5, of Section 1.1.

Now, since the matrix

$$\boldsymbol{J} = \begin{bmatrix} \frac{1}{2}\ddot{\boldsymbol{A}} & (\boldsymbol{B}_\perp\boldsymbol{R}_\perp)_\perp \\ (\boldsymbol{C}_\perp\boldsymbol{S}_\perp)_\perp' & \boldsymbol{0} \end{bmatrix} - \begin{bmatrix} \dot{\boldsymbol{A}}(\boldsymbol{C}')^g \\ \boldsymbol{0} \end{bmatrix} [\boldsymbol{B}^g\dot{\boldsymbol{A}} \quad \boldsymbol{0}] \tag{29}$$

$$= \begin{bmatrix} \frac{1}{2}\ddot{A} - \dot{A}A^g\dot{A} & (B_\perp R_\perp)_\perp \\ (C_\perp S_\perp)'_\perp & 0 \end{bmatrix}$$

corresponds to the Schur complement of the lower diagonal block I of the partitioned matrix on the right-hand side of (26), it follows that

$$det\,(P\,(1)) = det\,(J)\;det\,(I) = det\,(J) \tag{30}$$

which, in turn, entails that

$$det\,(P\,(1)) \neq 0 \tag{31}$$

by virtue of assumption iii).

In view of the foregoing, should we take the limit of both sides of (25) as z tends to 1, we would obtain

$$\begin{aligned} \lim_{z\to 1}[(1-z)^2 A^{-1}(z)] &= [I \quad 0] \begin{bmatrix} \frac{1}{2}\ddot{A} & F \\ G' & \Lambda(1) \end{bmatrix}^{-1} \begin{bmatrix} I \\ 0 \end{bmatrix} \\ &= [I \quad 0 \;\vdots\; 0] \begin{bmatrix} \frac{1}{2}\ddot{A} & (B_\perp R_\perp)_\perp & \vdots & \dot{A}(C')^g \\ (C_\perp S_\perp)'_\perp & 0 & \vdots & 0 \\ \cdots & \cdots & \vdots & \cdots \\ (B)^g\dot{A} & 0 & \vdots & I \end{bmatrix}^{-1} \begin{bmatrix} I \\ 0 \\ \cdots \\ 0 \end{bmatrix} \end{aligned} \tag{32}$$

whence, in view of (3) of Theorem1 in Section 1.2, we get

$$\begin{aligned} \lim_{z\to 1}[(1-z)^2 A^{-1}(z)] &= [I \quad 0]\, J^{-1} \begin{bmatrix} I \\ 0 \end{bmatrix} \\ &= [I \quad 0] \begin{bmatrix} \frac{1}{2}\ddot{A} - \dot{A}A^g\dot{A} & (B_\perp R_\perp)_\perp \\ (C_\perp S_\perp)'_\perp & 0 \end{bmatrix}^{-1} \begin{bmatrix} I \\ 0 \end{bmatrix} \end{aligned} \tag{33}$$

This, in view of Definition 2 together with (12) of Section 1.4, leads to conclude that $z = 1$ is a second order pole of $A^{-1}(z)$.

□

Corollary 3.1

The following statements are equivalent to condition iii)

a) $$det\,(R'_{\perp}\,B'_{\perp}\,\tilde{A}\,C_{\perp}S_{\perp}) \neq 0 \tag{34}$$

b) $$r\left(\begin{bmatrix} \tilde{A} & A_l^{\perp}\dot{A}A_r^{\perp}+A \\ A_l^{\perp}\dot{A}A_r^{\perp}+A & \boldsymbol{0} \end{bmatrix}\right) = n + r\,(A) + r\,(A_l^{\perp}\,\dot{A}\,A_r^{\perp}) \tag{35}$$

where $A_l^{\perp}$ and $A_r^{\perp}$ are as defined in (53) and (54) of Section 1.1.

Proof

Equivalence of iii) and a) follows from Theorem 2 of Section 1.2 given that Theorem 5 of Section 1.1 applies, bearing in mind (24) of Section 1.1.

Equivalence of iii) and b) is easily proved following an argument similar to that of Corollary 2.1, by observing that

$$\begin{aligned} det\left(\begin{bmatrix} \tilde{A} & (B_{\perp}R_{\perp})_{\perp} \\ (C_{\perp}S_{\perp})'_{\perp} & \boldsymbol{0} \end{bmatrix}\right) \neq 0 \Leftrightarrow r\left(\begin{bmatrix} \tilde{A} & (B_{\perp}R_{\perp})_{\perp} \\ (C_{\perp}S_{\perp})'_{\perp} & \boldsymbol{0} \end{bmatrix}\right) \\ = n + r\,(C_{\perp}S_{\perp})_{\perp} \end{aligned} \tag{36}$$

Indeed, the following hold:

$$\begin{aligned} & r\left(\begin{bmatrix} \tilde{A} & (B_{\perp}R_{\perp})_{\perp} \\ (C_{\perp}S_{\perp})'_{\perp} & \boldsymbol{0} \end{bmatrix}\right) = r\,(C_{\perp}S_{\perp})_{\perp} + r\,(B_{\perp}R_{\perp})_{\perp} \\ & + r\,\{[I-(B_{\perp}R_{\perp})_{\perp}\,(B_{\perp}R_{\perp})_{\perp}^{-}]\,\tilde{A}\,[I-((C_{\perp}S_{\perp})'_{\perp})^{-}(C_{\perp}S_{\perp})'_{\perp}]\} \\ & = r\,(C_{\perp}S_{\perp})_{\perp} + r\,(B_{\perp}R_{\perp})_{\perp} + r\,(R'_{\perp}\,B'_{\perp}\,\tilde{A}\,C_{\perp}S_{\perp}) \\ & = r\,(C_{\perp}S_{\perp})_{\perp} + r\,(B_{\perp}R_{\perp})_{\perp} + r\,(B_{\perp}R_{\perp}) = r\,(C_{\perp}S_{\perp})_{\perp} + n \end{aligned} \tag{37}$$

in light of Theorem 19 in Marsaglia and Styan, 1974, identities (47) and (48) of Section 1.1;

$$\begin{aligned} r\,(C_{\perp}S_{\perp})_{\perp} &= r\,((B_{\perp}R_{\perp})_{\perp}(C_{\perp}S_{\perp})'_{\perp}) = r\,((B'_{\perp})^{g}\,RS'\,C_{\perp}^{g} + A) \\ &= r\,(A_l^{\perp}\,\dot{A}\,A_r^{\perp}+A) = r\,(A_l^{\perp}\,\dot{A}\,A_r^{\perp}) + r\,(A) \end{aligned} \tag{38}$$

in light of (16), (27) and (28) above, (47), (48), (53) and (54) of Section 1.1 together with Theorem 14 in Marsaglia and Styan, 1974, and

$$
\begin{aligned}
& r\left(\begin{bmatrix} \tilde{A} & (B_\perp R_\perp)_\perp \\ (C_\perp S_\perp)'_\perp & 0 \end{bmatrix}\right) \\
&= r\left\{\begin{bmatrix} I & 0 \\ 0 & (B_\perp R_\perp)_\perp \end{bmatrix}\begin{bmatrix} \tilde{A} & (B_\perp R_\perp)_\perp \\ (C_\perp S_\perp)'_\perp & 0 \end{bmatrix}\begin{bmatrix} I & 0 \\ 0 & (C_\perp S_\perp)'_\perp \end{bmatrix}\right\} \\
&= r\left(\begin{bmatrix} \tilde{A} & (B_\perp R_\perp)_\perp (C_\perp S_\perp)'_\perp \\ (B_\perp R_\perp)_\perp (C_\perp S_\perp)'_\perp & 0 \end{bmatrix}\right) \\
&= r\left(\begin{bmatrix} \tilde{A} & (B'_\perp)^g RS'C_\perp^g + A \\ (B'_\perp)^g RS'C_\perp^g + A & 0 \end{bmatrix}\right) \\
&= r\left(\begin{bmatrix} \tilde{A} & A_l^\perp \dot{A} A_r^\perp + A \\ A_l^\perp \dot{A} A_r^\perp + A & 0 \end{bmatrix}\right)
\end{aligned} \tag{39}
$$

whence (35), because of (37) and (38).

□

1.7 Closed-Forms of Laurent Expansion Coefficient Matrices

In this section closed-form expressions for the matrices of Laurent expansions of matrix polynomial inverse about a simple and a second order pole, are derived.

We also present a collection of useful properties and worthwhile relationships, as by-products of the main results, which pave the way to obtaining special expansions with either truncated or annihilated principal parts via pole order-reduction or removal.

Notation and matrix qualifications of the previous section apply unless otherwise stated.

The simple pole case is dealt with in the following theorem

Theorem 1

Let the inverse, $A^{-1}(z)$, of the matrix polynomial

$$A(z) = (1 - z)\,Q(z) + A \tag{1}$$

have a simple pole at the point $z = 1$, as per Theorem 2 of Section 1.6, and suppose the Laurent expansion

$$A^{-1}(z) = \frac{1}{(1-z)} N_1 + M(z) \tag{2}$$

holds accordingly, in a deleted neighbourhood of $z = 1$.

Then the following closed-form representations hold for N_1 and $M(1)$ respectively:

$$N_1 = P_1 = -C_\perp(B'_\perp \dot{A} C_\perp)^{-1} B'_\perp \tag{3}$$

$$M(1) = -\frac{1}{2} P_1 \ddot{A} P_1 + P_2 P_3 = -\frac{1}{2} N_1 \ddot{A} N_1 + (I + N_1 \dot{A}) A^g (I + \dot{A} N_1) \tag{4}$$

where

$$\begin{bmatrix} P_1 & P_2 \\ P_3 & P_4 \end{bmatrix} = \begin{bmatrix} -\dot{A} & B \\ C' & 0 \end{bmatrix}^{-1} \tag{5}$$

Proof

In view of (5) of Section 1.4, the matrix N_1 is given by

$$N_1 = \lim_{z \to 1} [(1-z) A^{-1}(z)] \tag{6}$$

which, by virtue of (8) of Section 1.6, can be written as

$$N_1 = [I \quad 0] \begin{bmatrix} -\dot{A} & B \\ C' & 0 \end{bmatrix}^{-1} \begin{bmatrix} I \\ 0 \end{bmatrix} = P_1 \tag{7}$$

whence, according to the partitioned inversion (7) of Theorem 2, Section 1.2, the elegant closed-form solution

$$N_1 = -C_\perp(B'_\perp \dot{A} C_\perp)^{-1} B'_\perp \tag{8}$$

ensues.

In view of (6) of Section 1.4 the matrix $M(1)$ is given by

$$M(1) = -\lim_{z \to 1} \frac{d[(1-z)A^{-1}(z)]}{dz} \tag{9}$$

which, in light of (7) of Section 1.6 can be rewritten as

$$\boldsymbol{M}(1) = -\lim_{z\to 1}\frac{\mathrm{d}}{\mathrm{d}z}\left([\boldsymbol{I} \quad \boldsymbol{0}]\begin{bmatrix}\boldsymbol{Q}(z) & \boldsymbol{B}\\ \boldsymbol{C}' & (z-1)\boldsymbol{I}\end{bmatrix}^{-1}\begin{bmatrix}\boldsymbol{I}\\ \boldsymbol{0}\end{bmatrix}\right) \tag{10}$$

Differentiating, taking the limit as z tends to 1 and making use of the aforementioned partitioned inversion formula, simple computations lead to

$$\boldsymbol{M}(1) =$$

$$= -\lim_{z\to 1}\left([-\boldsymbol{I} \quad \boldsymbol{0}]\begin{bmatrix}\boldsymbol{Q}(z) & \boldsymbol{B}\\ \boldsymbol{C}' & (z-1)\boldsymbol{I}\end{bmatrix}^{-1}\begin{bmatrix}\dot{\boldsymbol{Q}}(z) & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{I}\end{bmatrix}\begin{bmatrix}\boldsymbol{Q}(z) & \boldsymbol{B}\\ \boldsymbol{C}' & (z-1)\boldsymbol{I}\end{bmatrix}^{-1}\begin{bmatrix}\boldsymbol{I}\\ \boldsymbol{0}\end{bmatrix}\right)$$

$$= [\boldsymbol{I} \quad \boldsymbol{0}]\begin{bmatrix}-\dot{\boldsymbol{A}} & \boldsymbol{B}\\ \boldsymbol{C}' & \boldsymbol{0}\end{bmatrix}^{-1}\begin{bmatrix}-\frac{1}{2}\ddot{\boldsymbol{A}} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{I}\end{bmatrix}\begin{bmatrix}-\dot{\boldsymbol{A}} & \boldsymbol{B}\\ \boldsymbol{C}' & \boldsymbol{0}\end{bmatrix}^{-1}\begin{bmatrix}\boldsymbol{I}\\ \boldsymbol{0}\end{bmatrix}$$

$$= [\boldsymbol{P}_1, \quad \boldsymbol{P}_2]\begin{bmatrix}-\frac{1}{2}\ddot{\boldsymbol{A}} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{I}\end{bmatrix}\begin{bmatrix}\boldsymbol{P}_1\\ \boldsymbol{P}_3\end{bmatrix} \tag{11}$$

$$= [\boldsymbol{N}_1, \quad (\boldsymbol{I} + \boldsymbol{N}_1\dot{\boldsymbol{A}})(\boldsymbol{C}')^g]\begin{bmatrix}-\frac{1}{2}\ddot{\boldsymbol{A}} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{I}\end{bmatrix}\begin{bmatrix}\boldsymbol{N}_1\\ \boldsymbol{B}^g(\boldsymbol{I} + \dot{\boldsymbol{A}}\boldsymbol{N}_1)\end{bmatrix}$$

$$= -\frac{1}{2}\boldsymbol{N}_1\ddot{\boldsymbol{A}}\boldsymbol{N}_1 + (\boldsymbol{I} + \boldsymbol{N}_1\dot{\boldsymbol{A}})\boldsymbol{A}^g(\boldsymbol{I} + \dot{\boldsymbol{A}}\boldsymbol{N}_1)$$

which completes the proof.

□

Corollary 1.1

The following results are true for the $n \times n$ matrix $\boldsymbol{N}_1$

i) $$r(\boldsymbol{N}_1) = n - r \tag{12}$$

where r is written instead of $r(\boldsymbol{A})$ for notational convenience.

ii) The null row-space of $\boldsymbol{N}_1$ is spanned by the r linearly independent columns of the matrix $\boldsymbol{C}$ of the rank factorization $\boldsymbol{A} = \boldsymbol{B}\boldsymbol{C}'$.

As a by-product of ii) we have the following pole free expansion

$$\boldsymbol{C}'\boldsymbol{A}^{-1}(z) = \boldsymbol{C}'\boldsymbol{M}(z) \tag{13}$$

Proof

The proof of i) rests on the rank equalities

$$r(N_1) = r(C_\perp) = n - r(C) = n - r(A) = n - r \tag{14}$$

which ensue from rank factorization and orthogonal complement rank properties.

To prove ii) observe that, in view of (8), the null row-spaces of N_1 and $C_\perp$ are the same. Hence, insofar as the r columns of C form a basis for the null row-space of $C_\perp$, they span the null row-space of N_1 as claimed.

Finally, expansion (13) follows by premultiplying the right-hand side of (2) by C', in light of (8). Given that C' is orthogonal to N_1 the term in $(1-z)^{-1}$ vanishes, thus removing the pole once located at $z = 1$, and the matrix function $C'A^{-1}(z)$ is analytical about $z = 1$.

□

Corollary 1.2

The following statements hold

i)
$$tr(N_1\dot{A}) = r - n \tag{15}$$

ii)
$$N_1\dot{A}N_1 = -N_1 \Rightarrow \dot{A} = -N_1^- \tag{16}$$

iii)
$$\begin{cases} M(1)B = (I + N_1\dot{A})(C')^g = P_2 \\ C'M(1) = B^g(I + \dot{A}N_1) = P_3 \end{cases} \tag{17}$$

iv)
$$AM(1)A = A \Rightarrow M(1) = A^- \tag{18}$$

Proof

Proof of i) follows by checking that

$$\begin{aligned} tr(N_1\dot{A}) &= -tr(C_\perp(B_\perp'\dot{A}C_\perp)^{-1}B_\perp'\dot{A}) \\ &= -tr((B_\perp'\dot{A}C_\perp)^{-1}(B_\perp'\dot{A}C_\perp)) = -tr(I_{n-r}) = r - n \end{aligned} \tag{19}$$

Proof of ii) follows from (3) by simple computation and from Definition 1 of Section 1.1.

Proof of iii) follows from (4) by simple computation in view of (19) of Section 1.1.

Proof of iv) follows from (4), in view of (3), by straightforward computations.

□

The next theorem deals with the case of a second order pole.

Theorem 2

Let the inverse, $A^{-1}(z)$, of the matrix polynomial

$$A\,(z) = (1-z)^2\,\Psi(z) - (1-z)\,\dot{A} + A \tag{20}$$

have a second order pole at the point $z = 1$, as per Theorem 3 of Section 1.6, and assume that the Laurent expansion

$$A^{-1}(z) = \frac{1}{(1-z)^2}N_2 + \frac{1}{(1-z)}N_1 + M\,(z) \tag{21}$$

holds accordingly, in a deleted neighbourhood of $z = 1$.

Then, the following closed-form representations hold for N_2 and N_1

$$N_2 = P_1 = C_\perp S_\perp (R'_\perp\, B'_\perp\, \tilde{A}\, C_\perp S_\perp)^{-1} R'_\perp\, B'_\perp \tag{22}$$

$$\begin{aligned} N_1 &= P_1\,\tilde{\tilde{A}}\,P_1 + P_2 U_2 B^g \dot{A}\, P_1 + P_1 \dot{A}\,(C')^g U'_2 P_3 - P_2 U_1 U'_1 P_3 \\ &= [N_2,\;\; I - N_2\tilde{A}\,] \begin{bmatrix} \tilde{\tilde{A}} & \dot{A}A^g \\ A^g\dot{A} & -C_\perp(B'_\perp\dot{A}C_\perp)^g B'_\perp \end{bmatrix} \begin{bmatrix} N_2 \\ I - \tilde{A}N_2 \end{bmatrix} \end{aligned} \tag{23}$$

where

$$\tilde{A} = \frac{1}{2}\ddot{A} - \dot{A}\,A^g\,\dot{A} \tag{24}$$

$$\tilde{\tilde{A}} = \frac{1}{6}\dddot{A} - \dot{A}\,A^g\,\dot{A}\,A^g\,\dot{A} \tag{25}$$

$$\begin{bmatrix} P_1 & P_2 \\ P_3 & P_4 \end{bmatrix} = \begin{bmatrix} \tilde{A} & (B_\perp R_\perp)_\perp \\ (C_\perp S_\perp)'_\perp & 0 \end{bmatrix}^{-1} \tag{26}$$

Here $(B_\perp R_\perp)_\perp$ and $(C_\perp S_\perp)_\perp$ stand for $[(B'_\perp)^g R,\;\; B]$ and $[(C'_\perp)^g S,\;\; C]$, rspectively, and

$$\boldsymbol{U}_1 = \begin{bmatrix} \boldsymbol{I} \\ \boldsymbol{0} \end{bmatrix}, \quad \boldsymbol{U}_2 = \begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{I} \end{bmatrix} = \boldsymbol{U}_{1\perp} \tag{27}$$

are selection matrices such that

$$(\boldsymbol{B}_\perp \boldsymbol{R}_\perp)_\perp \boldsymbol{U}_1 = (\boldsymbol{B}'_\perp)^g \boldsymbol{R}, \quad (\boldsymbol{B}_\perp \boldsymbol{R}_\perp)_\perp \boldsymbol{U}_2 = \boldsymbol{B} \tag{28}$$

Proof

In view of (12) of Section 1.4, the matrix $\boldsymbol{N}_2$ is given by

$$\boldsymbol{N}_2 = \lim_{z \to 1} [(1-z)^2 \boldsymbol{A}^{-1}(z)] \tag{29}$$

which , in light of (33) of Section 1.6, can be expressed as

$$\boldsymbol{N}_2 = [\boldsymbol{I} \quad \boldsymbol{0}] \begin{bmatrix} \frac{1}{2}\ddot{\boldsymbol{A}} - \dot{\boldsymbol{A}}\boldsymbol{A}^g\dot{\boldsymbol{A}} & (\boldsymbol{B}_\perp \boldsymbol{R}_\perp)_\perp \\ (\boldsymbol{C}_\perp \boldsymbol{S}_\perp)'_\perp & \boldsymbol{0} \end{bmatrix}^{-1} \begin{bmatrix} \boldsymbol{I} \\ \boldsymbol{0} \end{bmatrix} \tag{30}$$

whence, according to the partitioned inversion formula (7), Theorem 2, Section 1.2, the elegant closed form

$$\boldsymbol{N}_2 = \boldsymbol{C}_\perp \boldsymbol{S}_\perp (\boldsymbol{R}'_\perp \boldsymbol{B}'_\perp \tilde{\boldsymbol{A}} \boldsymbol{C}_\perp \boldsymbol{S}_\perp)^{-1} \boldsymbol{R}'_\perp \boldsymbol{B}'_\perp \tag{31}$$

ensues.

In view of (13) of Section 1.4, the matrix $\boldsymbol{N}_1$ is given by:

$$\boldsymbol{N}_1 = -\lim_{z \to 1} \frac{\mathrm{d}[(1-z)^2 \boldsymbol{A}^{-1}(z)]}{\mathrm{d}z} \tag{32}$$

which, in light of (23) and (25) of Section 1.6, can be written as

$$\begin{aligned} \boldsymbol{N}_1 &= -\lim_{z \to 1} \frac{\mathrm{d}[(1-z)^2 \boldsymbol{\Psi}(z) - \boldsymbol{F}\boldsymbol{V}(z)\boldsymbol{G}']^{-1}}{\mathrm{d}z} \\ &= -\lim_{z \to 1} \frac{\mathrm{d}}{\mathrm{d}z} \left\{ [\boldsymbol{I} \quad \boldsymbol{0}] \begin{bmatrix} \boldsymbol{\Psi}(z) & \boldsymbol{F} \\ \boldsymbol{G}' & \boldsymbol{\Lambda}(z) \end{bmatrix}^{-1} \begin{bmatrix} \boldsymbol{I} \\ \boldsymbol{0} \end{bmatrix} \right\} \end{aligned} \tag{33}$$

Before differentiating and taking the limit, some notational short cuts are introduced in order to simplify the computations. In particular, by denoting $[(B'_{\perp})^g R,\ B]$ and $[(C'_{\perp})^g S,\ C]$ by $(B_{\perp}R_{\perp})_{\perp}$ and $(C_{\perp}S_{\perp})_{\perp}$ respectively, with U_1 and U_2 as defined in (27), the matrix in the right-hand side of (33) can be written as follows

$$[I \quad 0]\begin{bmatrix} \Psi(z) & F \\ G' & \Lambda(z) \end{bmatrix}^{-1}\begin{bmatrix} I \\ 0 \end{bmatrix}$$

$$= [I \quad 0 \ \vdots \ 0]\left[\begin{array}{cc:c} \Psi(z) & (B_{\perp}R_{\perp})_{\perp} & \dot{A}(C')^g \\ (C_{\perp}S_{\perp})'_{\perp} & (1-z)U_1U'_1 & (1-z)U_2 \\ \hdashline B^g\dot{A} & (1-z)U'_2 & \Theta(z) \end{array}\right]^{-1}\begin{bmatrix} I \\ 0 \\ \hdashline 0 \end{bmatrix} \tag{34}$$

where

$$\Theta(z) = (1-z)\, B^g \dot{A}\, (C')^g + I \tag{35}$$

and the dotted lines indicate convenient partitionings of the matrices above.

Since $\Theta(z)$ is non-singular in a neighbourhood of $z = 1$, partitioned inversion formula (3) of Theorem 1, Section 1.2, applies and, referring to (25) of Section 1.6, it follows that

$$\begin{aligned}(1-z)^2 A^{-1}(z) &= [I \quad 0]\left\{\begin{bmatrix} \Psi(z) & (B_{\perp}R_{\perp})_{\perp} \\ (C_{\perp}S_{\perp})'_{\perp} & (1-z)\,U_1U'_1 \end{bmatrix} - \begin{bmatrix} \dot{A}(C')^g \\ (1-z)\,U_2 \end{bmatrix}\Theta^{-1}(z)\left[B^g\dot{A} \quad (1-z)U'_2\right]\right\}^{-1}\begin{bmatrix} I \\ 0 \end{bmatrix} \\ &= [I \quad 0]\ \{\Omega_0(z) + (1-z)\,\Omega_1(z) + (1-z)^2\,\Omega_2(z)\}^{-1}\begin{bmatrix} I \\ 0 \end{bmatrix} \\ &= [I \quad 0]\,\Omega^{-1}(z)\begin{bmatrix} I \\ 0 \end{bmatrix}\end{aligned} \tag{36}$$

where

$$\boldsymbol{\Omega}(z) = \boldsymbol{\Omega}_0(z) + (1-z)\,\boldsymbol{\Omega}_1(z) + (1-z)^2\,\boldsymbol{\Omega}_2(z) \tag{37}$$

$$\boldsymbol{\Omega}_0(z) = \begin{bmatrix} \boldsymbol{\Psi}(z) - \dot{\boldsymbol{A}}(\boldsymbol{C}')^g \boldsymbol{\Theta}^{-1}(z)\boldsymbol{B}^g\dot{\boldsymbol{A}} & (\boldsymbol{B}_\perp \boldsymbol{R}_\perp)_\perp \\ (\boldsymbol{C}_\perp \boldsymbol{S}_\perp)'_\perp & \boldsymbol{0} \end{bmatrix} \tag{38}$$

$$\boldsymbol{\Omega}_1(z) = \begin{bmatrix} \boldsymbol{0} & -\dot{\boldsymbol{A}}(\boldsymbol{C}')^g \boldsymbol{\Theta}^{-1}(z)\,\boldsymbol{U}'_2 \\ -\boldsymbol{U}_2\boldsymbol{\Theta}^{-1}(z)\boldsymbol{B}^g\dot{\boldsymbol{A}} & \boldsymbol{U}_1\boldsymbol{U}'_1 \end{bmatrix} \tag{39}$$

$$\boldsymbol{\Omega}_2(z) = \begin{bmatrix} \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & -\boldsymbol{U}_2\boldsymbol{\Theta}^{-1}(z)\,\boldsymbol{U}'_2 \end{bmatrix} \tag{40}$$

In particular, we have

$$\boldsymbol{\Omega}(1) = \boldsymbol{\Omega}_0(1) = \begin{bmatrix} \tilde{\boldsymbol{A}} & (\boldsymbol{B}_\perp \boldsymbol{R}_\perp)_\perp \\ (\boldsymbol{C}_\perp \boldsymbol{S}_\perp)'_\perp & \boldsymbol{0} \end{bmatrix} \tag{41}$$

Differentiating both sides of (36) yields

$$\frac{\mathrm{d}[(1-z)^2 \boldsymbol{A}^{-1}(z)]}{\mathrm{d}z} = -\,[\boldsymbol{I} \quad \boldsymbol{0}]\,\boldsymbol{\Omega}^{-1}(z)\,\dot{\boldsymbol{\Omega}}(z)\,\boldsymbol{\Omega}^{-1}(z)\begin{bmatrix} \boldsymbol{I} \\ \boldsymbol{0} \end{bmatrix} \tag{42}$$

Now, observe that

$$\dot{\boldsymbol{\Omega}}(z) = \dot{\boldsymbol{\Omega}}_0(z) - \boldsymbol{\Omega}_1(z) + \textit{terms in } (1-z) \text{ e } (1-z)^2 \tag{43}$$

and therefore by simple computations it follows that

$$\dot{\boldsymbol{\Omega}}(1) = \dot{\boldsymbol{\Omega}}_0(1) - \boldsymbol{\Omega}_1(1) = \begin{bmatrix} \tilde{\tilde{\boldsymbol{A}}} & \dot{\boldsymbol{A}}(\boldsymbol{C}')^g \boldsymbol{U}'_2 \\ \boldsymbol{U}_2\boldsymbol{B}^g\dot{\boldsymbol{A}} & -\boldsymbol{U}_1\boldsymbol{U}'_1 \end{bmatrix} \tag{44}$$

because of

$$\dot{\boldsymbol{\Psi}}(1) = \frac{1}{6}\dddot{\boldsymbol{A}} \tag{45}$$

$$\dot{\boldsymbol{\Theta}}^{-1}(1) = -\,\boldsymbol{\Theta}^{-1}(1)\,\dot{\boldsymbol{\Theta}}(1)\,\boldsymbol{\Theta}^{-1}(1) = -\,\dot{\boldsymbol{\Theta}}(1) = \boldsymbol{B}^g\,\dot{\boldsymbol{A}}\,(\boldsymbol{C}')^g \tag{46}$$

where the symbols $\dot{\boldsymbol{\Omega}}_0(1)$, $\dot{\boldsymbol{\Psi}}(1)$, $\dot{\boldsymbol{\Theta}}(1)$ indicate the derivative of the matrices $\boldsymbol{\Omega}_0(z)$, $\boldsymbol{\Psi}(z)$ and $\boldsymbol{\Theta}(z)$ evaluated at $z = 1$.

Combining (32) with (42) and making use of (41) and (44), the matrix N_1 turns out to be expressible in the following way

$$N_1 = [I \quad 0] \begin{bmatrix} \tilde{A} & (B_\perp R_\perp)_\perp \\ (C_\perp S_\perp)'_\perp & 0 \end{bmatrix}^{-1} \begin{bmatrix} \tilde{\tilde{A}} & \dot{A}(C')^g U'_2 \\ U_2 B^g \dot{A} & -U_1 U'_1 \end{bmatrix} \cdot \begin{bmatrix} \tilde{A} & (B_\perp R_\perp)_\perp \\ (C_\perp S_\perp)'_\perp & 0 \end{bmatrix}^{-1} \begin{bmatrix} I \\ 0 \end{bmatrix} \tag{47}$$

Now, one can verify that

$$\begin{aligned} [I \quad 0] \begin{bmatrix} \tilde{A} & (B_\perp R_\perp)_\perp \\ (C_\perp S_\perp)'_\perp & 0 \end{bmatrix}^{-1} &= [N_2 \quad (I - N_2\tilde{A})(C_\perp S_\perp)'^g_\perp] \\ &= [N_2 \quad I - N_2\tilde{A}] \begin{bmatrix} I & 0 \\ 0 & (C_\perp S_\perp)'^g_\perp \end{bmatrix} \end{aligned} \tag{48}$$

$$\begin{aligned} \begin{bmatrix} \tilde{A} & (B_\perp R_\perp)_\perp \\ (C_\perp S_\perp)'_\perp & 0 \end{bmatrix}^{-1} \begin{bmatrix} I \\ 0 \end{bmatrix} &= \begin{bmatrix} N_2 \\ (B_\perp R_\perp)^g_\perp (I - \tilde{A}N_2) \end{bmatrix} \\ &= \begin{bmatrix} I & 0 \\ 0 & (B_\perp R_\perp)^g_\perp \end{bmatrix} \begin{bmatrix} N_2 \\ I - \tilde{A}N_2 \end{bmatrix} \end{aligned} \tag{49}$$

in view of the partitioned inversion formula (7) of Theorem 2, Section 1.2. Besides, by virtue of (40) of Section 1.1, the following results are easy to verify

$$((C_\perp S_\perp)'_\perp)^g U_2 = (C')^g, \quad U'_2\,(B_\perp R_\perp)^g_\perp = B^g \tag{50}$$

$$((C_\perp S_\perp)'_\perp)^g U_1 U'_1\,(B_\perp R_\perp)^g_\perp = (S' C^g_\perp)^g\,((B'_\perp)^g R)^g \tag{51}$$

Indeed, the matrix $(S' C^g_\perp)^g$, bearing in mind (42) of Section 1.1, can be expressed as

$$\begin{aligned} (S' C^g_\perp)^g &= (C'_\perp)^g S\,[S' C^g_\perp (C'_\perp)^g S]^{-1} \\ &= C_\perp (C'_\perp C_\perp)^{-1} S\,[S' (C'_\perp C_\perp)^{-1} S]^{-1} \\ &= C_\perp [(S')^g S' + S_\perp S^g_\perp](C'_\perp C_\perp)^{-1} S\,[S'(C'_\perp C_\perp)^{-1} S]^{-1} \end{aligned} \tag{52}$$

$$= C_\perp(S')^g + C_\perp S_\perp S_\perp^g (C_\perp' C_\perp)^{-1} S\,[S'(C_\perp' C_\perp)^{-1} S]^{-1}$$

and $((B_\perp')^g R)^g$ can be likewise expressed as

$$((B_\perp')^g R)^g = R^g B_\perp' + [R'(B_\perp' B_\perp)^{-1} R]^{-1} R'(B_\perp' B_\perp)^{-1} R_\perp R_\perp^g B_\perp' \tag{52'}$$

This together with

$$(I - N_2 \tilde{A})\, C_\perp S_\perp = 0 \tag{53}$$

yield the equality

$$(I - N_2 \tilde{A})\,(S' C_\perp^g)^g = (I - N_2 \tilde{A})\, C_\perp (S')^g \tag{54}$$

as well as the equality

$$((B_\perp')^g R)^g (I - \tilde{A} N_2) = R^g B_\perp' (I - \tilde{A} N_2) \tag{55}$$

In view of the foregoing, after proper substitutions and some computations, the following closed-form expression for N_1 is obtained

$$N_1 = [N_2, \;\; I - N_2 \tilde{A}\,] \begin{bmatrix} \tilde{\tilde{A}} & \dot{A} A^g \\ A^g \dot{A} & -C_\perp (S')^g R^g B_\perp' \end{bmatrix} \begin{bmatrix} N_2 \\ I - \tilde{A} N_2 \end{bmatrix} \tag{56}$$

which eventually leads to the right-hand side of (23), upon noting that

$$C_\perp(S')^g R^g B_\perp' = C_\perp (RS')^g B_\perp' = C_\perp (B_\perp' \dot{A} C_\perp)^g B_\perp' \tag{57}$$

by (19) of Section 1.1 and (16) of Section 1.6. The preceding closed-form of N_1 can be drawn from (47) in light of (26).

□

Corollary 2.1

The following results hold for the n matrix N_2

i)
$$r(N_2) = n - r_1 - r_2 \tag{58}$$

where r_1 and r_2 are written instead of $r(A)$ and $r(B_\perp' \dot{A} C_\perp)$ respectively, for notational convenience.

ii) The row-kernel of N_2 is spanned by the $r_1 + r_2$ linearly independent columns of an arbitrary orthogonal complement of $C_\perp S_\perp$, say $(C_\perp S_\perp)_\perp = [(C_\perp')^g S, \;\; C]$.

As a by-product of ii) the following expansion about the simple pole $z = 1$ can be obtained

$$(C_\perp S_\perp)'_\perp A^{-1}(z) = \frac{1}{(1-z)}(C_\perp S_\perp)'_\perp N_1 + (C_\perp S_\perp)'_\perp M(z) \tag{59}$$

Proof

The proof of i) rests on the rank equalities

$$r(N_2) = r(C_\perp S_\perp) = r(S_\perp) = r(C_\perp) - r(B'_\perp \dot{A} C_\perp) = n - r_1 - r_2 \tag{60}$$

after a reasoning similar to that in the proof of Corollary 1.1.

To prove ii) observe that, in view of (31), the $r_1 + r_2$ columns of $(C_\perp S_\perp)_\perp$ form a basis for the null row-space of N_2, and as long as $[(C'_\perp)^g S,\ C]$ is a possible choice of $(C_\perp S_\perp)_\perp$, its columns span the row-kernel of N_2 as claimed.

Finally, expansion (59) follows by premultiplying the right-hand side of (21) by $(C_\perp S_\perp)'_\perp$ in light of (31). Given that $(C_\perp S_\perp)'_\perp$ is orthogonal to N_2, the term in $(1-z)^{-2}$ of $(C_\perp S_\perp)'_\perp A^{-1}(z)$ vanishes, a pole order-reduction occurs and the matrix function $(C_\perp S_\perp)'_\perp A^{-1}(z)$ exhibits a simple pole at $z = 1$.

□

Corollary 2.2

The following results hold

i) $$r[N_2, N_1] = n - r_1 + r_3 \tag{61}$$

where r_3 stands for $r(B' \dot{A} C_\perp S_\perp)$ to simplify the notation, and r_1 is as in Corollary 2.1.

ii) Should

$$r_3 < r_1 \tag{62}$$

hold true and

$$VW' = B^g \dot{A} C_\perp S_\perp,\ r(V) = r(W) = r_3 \tag{63}$$

be a rank factorization of $B^g \dot{A} C_\perp S_\perp$, then the row-kernel of $[N_2, N_1]$ is spanned by the $r_1 - r_3$ columns of the matrix $CV_\perp$.

Under ii) we have the following pole free expansion

$$V'_\perp C' A^{-1}(z) = V'_\perp C' M(z) \tag{64}$$

Proof

The proof of i) rests on several rank relationships given here below, which can be established by applying onward and backward, in turn, Theorem 19 of Marsaglia and Styan, 1974.

First, observe that, by onward application of the said theorem, we can write

$$r\,[\boldsymbol{N}_2, \boldsymbol{N}_1] = r\,(\boldsymbol{N}_2) + r\,((\boldsymbol{I} - \boldsymbol{N}_2\tilde{\boldsymbol{A}})\,\boldsymbol{N}_1) \tag{65}$$

upon noting that $\tilde{\boldsymbol{A}}$ is a generalized inverse of $\boldsymbol{N}_2$ as a straightforward computation shows.

Next, refer to (56) and check that

$$\begin{aligned}(\boldsymbol{I} - \boldsymbol{N}_2\tilde{\boldsymbol{A}})\,\boldsymbol{N}_1 &= [\boldsymbol{0},\ \boldsymbol{I} - \boldsymbol{N}_2\tilde{\boldsymbol{A}}]\begin{bmatrix}\tilde{\tilde{\boldsymbol{A}}} & \dot{\boldsymbol{A}}\boldsymbol{A}^g \\ \boldsymbol{A}^g\dot{\boldsymbol{A}} & \boldsymbol{C}_\perp(\boldsymbol{S}')^g\boldsymbol{R}^g\boldsymbol{B}'_\perp\end{bmatrix}\begin{bmatrix}\boldsymbol{N}_2 \\ \boldsymbol{I} - \tilde{\boldsymbol{A}}\boldsymbol{N}_2\end{bmatrix} \\ &= (\boldsymbol{I} - \boldsymbol{N}_2\tilde{\boldsymbol{A}})\,\boldsymbol{K}\end{aligned} \tag{66}$$

where

$$\boldsymbol{K} = \boldsymbol{A}^g\dot{\boldsymbol{A}}\,\boldsymbol{N}_2 + \boldsymbol{C}_\perp(\boldsymbol{S}')^g\,\boldsymbol{R}^g\,\boldsymbol{B}'_\perp(\boldsymbol{I} - \tilde{\boldsymbol{A}}\,\boldsymbol{N}_2) \tag{67}$$

This and (65), by backward application of the said theorem, yield

$$r\,[\boldsymbol{N}_2, \boldsymbol{N}_1] = r\,(\boldsymbol{N}_2) + r\,((\boldsymbol{I} - \boldsymbol{N}_2\tilde{\boldsymbol{A}})\boldsymbol{N}_1) = r\,([\boldsymbol{N}_2, \boldsymbol{K}]) \tag{68}$$

Now, refer to (31) and observe that

$$\begin{aligned}\boldsymbol{I} - \tilde{\boldsymbol{A}}\,\boldsymbol{N}_2 &= \boldsymbol{I} - \tilde{\boldsymbol{A}}\,\boldsymbol{C}_\perp\boldsymbol{S}_\perp(\boldsymbol{R}'_\perp\,\boldsymbol{B}'_\perp\,\tilde{\boldsymbol{A}}\,\boldsymbol{C}_\perp\boldsymbol{S}_\perp)^{-1}\,\boldsymbol{R}'_\perp\,\boldsymbol{B}'_\perp \\ &= (\boldsymbol{B}_\perp\boldsymbol{R}_\perp)_\perp\{[\tilde{\boldsymbol{A}}\boldsymbol{C}_\perp\boldsymbol{S}_\perp]'_\perp(\boldsymbol{B}_\perp\boldsymbol{R}_\perp)_\perp\}^{-1}[\tilde{\boldsymbol{A}}\boldsymbol{C}_\perp\boldsymbol{S}_\perp]'_\perp\end{aligned} \tag{69}$$

in view of (43) of Section 1.1.

When Theorem 6 of Section 1.1 is applied to the right-hand side of (69), we can choose

$$(\boldsymbol{B}_\perp\boldsymbol{R}_\perp)_\perp = [(\boldsymbol{B}'_\perp)^g\boldsymbol{R},\ \boldsymbol{B}] \tag{70}$$

which, after proper substitutions in (67) and some simple computations, yields

$$\boldsymbol{K} = \boldsymbol{A}^g\dot{\boldsymbol{A}}\,\boldsymbol{C}_\perp\boldsymbol{S}_\perp\boldsymbol{T} + [\boldsymbol{C}_\perp(\boldsymbol{S}')^g,\ \boldsymbol{0}]\,\boldsymbol{\vartheta} \tag{71}$$

where

$$\boldsymbol{T} = (\boldsymbol{R}'_\perp\,\boldsymbol{B}'_\perp\,\tilde{\boldsymbol{A}}\,\boldsymbol{C}_\perp\boldsymbol{S}_\perp)^{-1}\,\boldsymbol{R}'_\perp\,\boldsymbol{B}'_\perp \tag{72}$$

$$\vartheta = \{ [\tilde{A} C_{\perp} S_{\perp}]'_{\perp} (B_{\perp} R_{\perp})_{\perp} \}^{-1} [\tilde{A} C_{\perp} S_{\perp}]'_{\perp} \tag{73}$$

are full row-rank matrices.

In turn, this together with (68) lead to the relations below

$$\begin{aligned} r\,[N_2, N_1] &= r\,([C_{\perp} S_{\perp} T, A^g \dot{A}\, C_{\perp} S_{\perp} T + [C_{\perp} (S')^g, 0]\, \vartheta]) \\ &= r\,([C_{\perp} S_{\perp}, A^g \dot{A}\, C_{\perp} S_{\perp}, [C_{\perp} (S')^g, 0]]\, H) \\ &= r\,([C_{\perp} S_{\perp}, A^g \dot{A}\, C_{\perp} S_{\perp}, C_{\perp} (S')^g]) \\ &= r\,([C_{\perp} [S_{\perp}, S\,(S'S)^{-1}], A^g \dot{A}\, C_{\perp} S_{\perp}]) \\ &= r\,([C_{\perp}, A^g \dot{A}\, C_{\perp} S_{\perp}]\, \Xi) = r\,([C_{\perp}, A^g \dot{A}\, C_{\perp} S_{\perp}]) \end{aligned} \tag{74}$$

where

$$H = \begin{bmatrix} T & 0 \\ 0 & T \\ 0 & \vartheta \end{bmatrix} \tag{75}$$

$$\Xi = \begin{bmatrix} [S_{\perp}, & S(S'S)^{-1}] & 0 \\ & 0 & I \end{bmatrix} \tag{76}$$

are full row-rank matrices.

In light of what we have seen so far, by onward application of the above mentioned theorem once more, we eventually arrive at

$$\begin{aligned} r\,(N_2, N_1) &= r\,([C_{\perp}, A^g \dot{A}\, C_{\perp} S_{\perp}]) \\ &= r\,(C_{\perp}) + r\,([I - (C_{\perp} C_{\perp}^g)]\, A^g \dot{A}\, C_{\perp} S_{\perp}) \\ &= r\,(C_{\perp}) + r\,(A^g \dot{A}\, C_{\perp} S_{\perp}) = n - r\,(A) + r\,(A^g \dot{A}\, C_{\perp} S_{\perp}) \\ &= n - r\,(A) + r\,(B' \dot{A}\, C_{\perp} S_{\perp}) \end{aligned} \tag{77}$$

by taking into account that

$$(I - C_{\perp} C_{\perp}^g)\, A^g = A^g \tag{78}$$

in view of (48) of Section 1.1.

To prove ii) observe that, by (17) and (19) of Section 1.4, the following hold

$$C'N_2 = 0 \tag{79}$$

$$C'N_1 = B^g B C' N_1 = B^g A N_1 = B^g \dot{A}\, N_2 \tag{80}$$

whence

$$r(C'N_1) = r(B'\dot{A}N_2) \tag{81}$$

together with the rank factorization (63) as a by-product, under the rank condition (62).

The foregoing, in turn, entails that

$$V_\perp' C'[N_2, N_1] = [0, 0] \tag{82}$$

which is tantamount to saying that the $r_1 - r_3$ columns of $CV_\perp$ form a spanning set, and more precisely a basis, for the row kernel of the block matrix $[N_2, N_1]$.

Because of (82) the terms in $(1-z)^{-1}$ and in $(1-z)^{-2}$ of $V_\perp' C'A^{-1}(z)$ vanish, thus removing the pole located at $z = 1$ and making the matrix function $V_\perp' C'A^{-1}(z)$ analytical about $z = 1$.

□

Corollary 2.3

The following statements hold true

i)
$$tr(N_2\dot{A}) = 0 \tag{83}$$

ii)
$$N_2\tilde{A}N_2 = N_2 \;\Rightarrow\; \tilde{A} = N_2^- \tag{84}$$

iii)
$$(C_\perp S_\perp)_\perp' N_1 (B_\perp R_\perp)_\perp = -U_1 U_1' \tag{85}$$

where U_1 is as defined in (27) and $(C_\perp S_\perp)_\perp$ and $(B_\perp R_\perp)_\perp$ denote $[(C_\perp')^g S, \; C]$ and $[(B_\perp')^g R, \; B]$ respectively.

iv)
$$AN_1A = 0 \tag{86}$$

v)
$$AM(1)A = A + \dot{A}N_2\dot{A} \tag{87}$$

vi)
$$A(M(1) - N_1N_2^gN_1)A = A \Rightarrow M(1) - N_1N_2^gN_1 = A^- \tag{88}$$

Proof

Point i): the proof follows by checking that

$$\begin{aligned} tr(N_2\dot{A}) &= tr(C_\perp S_\perp (R_\perp' B_\perp' \tilde{A} C_\perp S_\perp)^{-1} R_\perp' B_\perp' \dot{A} \\ &= tr((R_\perp' B_\perp' \tilde{A} C_\perp S_\perp)^{-1} R_\perp' B_\perp' \dot{A} C_\perp S_\perp) \end{aligned} \tag{89}$$

$$= tr\,((\boldsymbol{R}'_{\perp}\,\boldsymbol{B}'_{\perp}\,\tilde{\boldsymbol{A}}\,\boldsymbol{C}_{\perp}\boldsymbol{S}_{\perp})^{-1}\,\boldsymbol{R}'_{\perp}\boldsymbol{R}\boldsymbol{S}'\boldsymbol{S}_{\perp}) = tr\,(\boldsymbol{0}) = 0$$

in light of (16) of Section 1.6 (cf. also Theorem 2 of Section 1.3).

Point ii): the proof ensues from (31) after elementary computations. The implication concerning the generalized inverse is trivial.

Point iii): the result follows from (47) upon noting that

$$(\boldsymbol{C}_{\perp}\boldsymbol{S}_{\perp})'_{\perp}\,[\boldsymbol{I} \quad \boldsymbol{0}] = [\boldsymbol{0} \quad \boldsymbol{I}]\begin{bmatrix} \tilde{\boldsymbol{A}} & (\boldsymbol{B}_{\perp}\boldsymbol{R}_{\perp})_{\perp} \\ (\boldsymbol{C}_{\perp}\boldsymbol{S}_{\perp})'_{\perp} & \boldsymbol{0} \end{bmatrix} \tag{90}$$

$$\begin{bmatrix} \boldsymbol{I} \\ \boldsymbol{0} \end{bmatrix}(\boldsymbol{B}_{\perp}\boldsymbol{R}_{\perp})_{\perp} = \begin{bmatrix} \tilde{\boldsymbol{A}} & (\boldsymbol{B}_{\perp}\boldsymbol{R}_{\perp})_{\perp} \\ (\boldsymbol{C}_{\perp}\boldsymbol{S}_{\perp})'_{\perp} & \boldsymbol{0} \end{bmatrix}\begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{I} \end{bmatrix} \tag{91}$$

Point iv): the result is obvious in view of (18) and (19) of Section 1.4.

To prove the result v), first observe that the following identity

$$\boldsymbol{A}\,(z)\,\boldsymbol{A}^{-1}(z)\,\boldsymbol{A}\,(z) = \boldsymbol{A}\,(z) \tag{92}$$

is trivially satisfied in a deleted neighbourhood of $z = 1$.

Then, substituting the right-hand side of (20) and (21) for $\boldsymbol{A}\,(z)$ and $\boldsymbol{A}^{-1}(z)$ in (92), making use of (17), (18) , (19) and (20) of Section 1.4 together with (86) above , after simple computations we obtain

$$\begin{aligned} -\dot{\boldsymbol{A}}\,\boldsymbol{N}_2\,\dot{\boldsymbol{A}} \;+ \boldsymbol{A}\boldsymbol{M}\,(z)\,\boldsymbol{A} + \textit{terms in positive powers of}\,(1-z) \\ = \boldsymbol{A} - (1-z)\;\dot{\boldsymbol{A}} \;+ (1-z)^2\,\boldsymbol{\Psi}(z) \end{aligned} \tag{93}$$

Expanding $\boldsymbol{M}\,(z)$ about $z = 1$, that is to say

$$\boldsymbol{M}\,(z) = \boldsymbol{M}\,(1) + \textit{terms in positive powers of}\,(1-z) \tag{94}$$

and substituting the right-hand side of (94) for $\boldsymbol{M}\,(z)$ in (93), collecting like powers and equating term by term, we at the end get for the constant terms the equality

$$-\dot{\boldsymbol{A}}\,\boldsymbol{N}_2\,\dot{\boldsymbol{A}} \;+ \boldsymbol{A}\boldsymbol{M}\,(1)\,\boldsymbol{A} = \boldsymbol{A} \tag{95}$$

Point vi): Because of (19) and (20) of Section 1.4, identity (87) can be restated as follows

$$\boldsymbol{A}\boldsymbol{M}\,(1)\,\boldsymbol{A} = \boldsymbol{A} + \;\dot{\boldsymbol{A}}\,\boldsymbol{N}_2\boldsymbol{N}_2^{g}\boldsymbol{N}_2\dot{\boldsymbol{A}} \;= \boldsymbol{A} + \boldsymbol{A}\boldsymbol{N}_1\,\boldsymbol{N}_2^{g}\boldsymbol{N}_1\boldsymbol{A} \tag{96}$$

which, rearranging terms, eventually leads to (88).

□

2 The Statistical Setting

This chapter introduces the basic notions regarding the multivariate stochastic processes. In particular, the reader will find the definitions of stationarity and of integration which are of special interest for the subsequent developments. The second part deals with principle stationary processes. The third section shows the way to integrated processes and takes a glance at cointegration. The last section deals with integrated and cointegrated processes and related topics of major interest. An appendix on the rôle of cointegration completes this chapter.

2.1 Stochastic Processes: Preliminaries

The notion of stochastic process is a dynamic extension of the notion of random variable. Broadly speaking a random process is a process running along in time and controlled by probabilistic laws. It can be properly defined as a family, an ordered sequence, of random variables y_t, where the order is given by the (discrete) time variable t.

As a mirror image of the foregoing reading key, we can look at a stochastic process as a complex of like mechanisms, whose outcomes – to be identified with the notion of time series – exhibit distinguishing features and discrepancies which can be explained on a statistical basis.

By a multivariate stochastic process we mean a random vector, say

$$\underset{(1,\, n)}{\boldsymbol{y}_t'} = [y_{t1}, y_{t2}, \ldots, y_{tn}] \tag{1}$$

whose elements are scalar random processes.

In order to properly specify a stochastic process, the distribution functions of its elements, pairs of elements, ..., k-ples of elements, for any k, should be given and satisfy the so-called symmetry and compatibility conditions (see, e.g., Yaglom, 1962).

In practise, a short cut simplification is usually adopted and reference is made to the lower-order moments of the process, basically the mean and autocovariance functions that we are going to introduce.

Denoting by E the averaging operator, otherwise known as expectation operator, the (unconditional) mean vector of the process is defined as

$$E(y_t) \tag{2}$$

while the autocovariance matrices are defined as

$$E\{(y_t - E(y_t))\,(y_\tau - E(y_\tau))'\} \tag{3}$$

It is evident that formula (3) describes a family of functions when the pair of indices t and τ varies.

Restricting the attention to the principal moments, namely the mean vector and the autocovariance matrices, paves the way to the various notions of stationarity which enjoy prominent interest in econometrics.

In this connection, let us give the following definitions

Definition 1

A stochastic process is called *stationary* insofar as – at least to some extent – it exhibits characteristics of permanence and satisfies statistical properties which are not affected by a shift in the time origin, which in turn grants some sort of temporal homogeneity (see, e.g., Blanc-Lapierre and Fortet, 1953; Papoulis, 1965).

The notion of stationary can actually assumes a plurality of facets: the ones reported below are of particular interest for the subsequent analysis.

Definition 2

A process y_t is said to be *stationary in mean* if

$$E(y_t) = \mu \tag{4}$$

where μ is a time-invariant vector.

Remark

If a process y_t is stationary in mean, the difference process ∇y_t is itself a stationary process, whose mean is a null vector and vice versa.

Definition 3

A process y_t is said to be *covariance stationary* if (3) depends only on the temporal lag $\tau - t$ of the argument processes.

Definition 4

A process $\boldsymbol{y}_t$ is said to be *stationary in the wide sense,* or *weakly stationary*, when both stationary in mean and in covariance.

For a covariance stationary n-dimensional process the matrix

$$\boldsymbol{\Gamma}(h) = E\{(\boldsymbol{y}_t - \boldsymbol{\mu})(\boldsymbol{y}_{t+h} - \boldsymbol{\mu})'\} \tag{5}$$

represents the *autocovariance matrix* of order h. It easy to see that for real processes the following holds

$$\boldsymbol{\Gamma}(-h) = \boldsymbol{\Gamma}'(h) \tag{6}$$

The *autocorrelation matrix* $\boldsymbol{P}(h)$ *of order* h is the matrix defined as follows

$$\boldsymbol{P}(h) = \boldsymbol{D}^{-1}\,\boldsymbol{\Gamma}(h)\,\boldsymbol{D}^{-1} \tag{7}$$

where $\boldsymbol{D}$ is the diagonal matrix

$$\boldsymbol{D} = \begin{bmatrix} \sqrt{\gamma_{11}(0)} & 0 & 0 & 0 \\ 0 & \sqrt{\gamma_{22}(0)} & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \sqrt{\gamma_{nn}(0)} \end{bmatrix} \tag{8}$$

whose diagonal entries are the standard error of the elements of the vector $\boldsymbol{y}_t$.

The foregoing covers what really matters about stationarity for our purposes. Moving to non stationary processes, we are mainly interested in the class of so-called integrated processes, which we are going to define.

Definition 5

An *integrated process* of order d – written as $I(d)$ – where d is a positive integer, is a process ζ_t such that it must be differenced d times in order to recover stationarity.

As a by-product of the operator identity

$$\nabla^0 = \boldsymbol{I} \tag{9}$$

a process $I(0)$ is trivially stationary.

2.2 Principal Multivariate Stationary Processes

This section displays the outline of principle stochastic processes and derives the closed-forms of their first and second moments.

We give the following definitions

Definition 1

A *white noise* of dimension n, written as $WN_{(n)}$, is a process ε_t with

$$E(\varepsilon_t) = \boldsymbol{0} \tag{1}$$

$$E(\varepsilon_t\, \varepsilon'_\tau) = \delta_{\tau - t}\, \boldsymbol{\Sigma} \tag{2}$$

where $\boldsymbol{\Sigma}$ denotes a positive definite time-invariant dispersion matrix, and δ_v is the (discrete) unitary function, that is to say

$$\begin{cases} \delta_v = 1 & if \;\; v = 0 \\ \delta_v = 0 & otherwise \end{cases} \tag{3}$$

The autocovariance matrices of the process turn out to be given by

$$\Gamma_\varepsilon(h) = \delta_h \boldsymbol{\Sigma} \tag{4}$$

with the corollary that the following noteworthy relation holds for the autocovariance matrix of composite vectors (cf. Faliva and Zoia, 1999, p. 23)

$$E = \left\{ \begin{bmatrix} \varepsilon_t \\ \varepsilon_{t-1} \\ \vdots \\ \varepsilon_{t-q} \end{bmatrix} [\varepsilon'_{t+h}, \varepsilon'_{t-1+h}, \ldots, \varepsilon'_{t-q+h}] \right\} = \boldsymbol{D}_h \otimes \boldsymbol{\Sigma} \tag{5}$$

where $\boldsymbol{D}_h$ is a matrix given by

$$\boldsymbol{D}_h = \begin{cases} \boldsymbol{I}_{q+1} & if \;\; h = 0 \\ \boldsymbol{J}^h & if \;\; 1 \le h \le q \\ (\boldsymbol{J}')^{|h|} & if \;\; -q \le h \le -1 \\ \boldsymbol{0}_{q+1} & if \;\; |h| > \mathrm{q} \end{cases} \tag{6}$$

Here $\boldsymbol{J}$ denotes the first unitary super diagonal matrix (of order $q + 1$), defined as

$$\underset{(q+1,\,q+1)}{\boldsymbol{J}} = [j_{nm}], \quad \text{with } j_{nm} = \begin{cases} 1 & \text{if } m = n+1 \\ 0 & \text{if } m \neq n+1 \end{cases} \tag{7}$$

while $\boldsymbol{J}^h$ and $(\boldsymbol{J}')^h$, stand for, respectively, the unitary super and sub diagonal matrices of order $h = 1, 2, \ldots$.

Definition 2

A *vector moving-average process* of order q, denoted by VMA (q), is a multivariate process specified as follows

$$\underset{(n,1)}{\boldsymbol{y}_t} = \boldsymbol{\mu} + \sum_{j=0}^{q} \boldsymbol{M}_j\, \boldsymbol{\varepsilon}_{t-j}, \quad \boldsymbol{\varepsilon}_t \sim WN_{(n)} \tag{8}$$

where $\boldsymbol{\mu}$ and $\boldsymbol{M}_0, \boldsymbol{M}_1, \ldots, \boldsymbol{M}_q$ are, respectively, a constant vector and constant matrices.

In operator form this process can be expressed as

$$\boldsymbol{y}_t = \boldsymbol{\mu} + \boldsymbol{M}(L)\, \boldsymbol{\varepsilon}_t, \quad \boldsymbol{M}(L) = \sum_{j=0}^{q} \boldsymbol{M}_j L^j \tag{9}$$

where L is the lag operator.

A VMA(q) process is weakly stationary, as the following formulas show

$$E(\boldsymbol{y}_t) = \boldsymbol{\mu} \tag{10}$$

$$\boldsymbol{\Gamma}(h) = \begin{cases} \sum\limits_{j=0}^{q} \boldsymbol{M}_j \boldsymbol{\Sigma} \boldsymbol{M}'_j & \text{if} \quad h = 0 \\ \sum\limits_{j=0}^{q-h} \boldsymbol{M}_j \boldsymbol{\Sigma} \boldsymbol{M}'_{j+h} & \text{if} \quad 1 \le h \le q \\ \sum\limits_{j=0}^{q-|h|} \boldsymbol{M}_{j+|h|} \boldsymbol{\Sigma} \boldsymbol{M}'_j & \text{if} \quad -q \le h \le -1 \\ \boldsymbol{0} & \text{if} \quad |h| > q \end{cases} \tag{11}$$

The proof of (10) is straightforward in view of the properties of the expectation operator and of (1) above.

The proof of (11) can easily be obtained upon noting that

$$y_t = \mu + [M_0, M_1, ..., M_q] \begin{bmatrix} \varepsilon_t \\ \varepsilon_{t-1} \\ \vdots \\ \varepsilon_{t-q} \end{bmatrix} \tag{12}$$

which in view of (5) and (6) leads to

$$\begin{aligned} \Gamma(h) &= E\left\{[M_{0,} M_{1,} ..., M_q] \begin{bmatrix} \varepsilon_t \\ \varepsilon_{t-1} \\ \vdots \\ \varepsilon_{t-q} \end{bmatrix} [\varepsilon'_{t+h}, \varepsilon'_{t-1+h}, ..., \varepsilon'_{t-q+h}] \begin{bmatrix} M'_0 \\ M'_1 \\ \vdots \\ M'_q \end{bmatrix}\right\} \\ &= [M_0, M_1, ..., M_q]\, E\left\{\begin{bmatrix} \varepsilon_t \\ \varepsilon_{t-1} \\ \vdots \\ \varepsilon_{t-q} \end{bmatrix} [\varepsilon'_{t+h}, \varepsilon'_{t-1+h}, ..., \varepsilon'_{t-q+h}]\right\} \begin{bmatrix} M'_0 \\ M'_1 \\ \vdots \\ M'_q \end{bmatrix} \\ &= [M_0, M_1, ..., M_q]\, (D_h \otimes \Sigma) \begin{bmatrix} M'_0 \\ M'_1 \\ \vdots \\ M'_q \end{bmatrix} \end{aligned} \tag{13}$$

whence (11).

It is also worth mentioning the staked version of the autocovariance matrix of order zero, namely

$$vec\, \Gamma(0) = \sum_{j=0}^{q} (M_j \otimes M_j)\ vec\, \Sigma \tag{14}$$

The first and second differences of a white noise process happen to play some rôle in time series econometrics and as such are worth mentioning here.

Actually, such processes can be viewed as special cases of VMA processes, and enjoy the weak stationarity property accordingly, as the following definitions show.

Definition 3

Let the process $\boldsymbol{y}_t$ be specified as a VMA(1) as follows

$$\boldsymbol{y}_t = \boldsymbol{M}\boldsymbol{\varepsilon}_t - \boldsymbol{M}\boldsymbol{\varepsilon}_{t-1} \tag{15}$$

or, equivalently, as a first difference of a $WN_{(n)}$ process, namely

$$\boldsymbol{y}_t = \boldsymbol{M}\nabla\boldsymbol{\varepsilon}_t \tag{16}$$

The following hold for the first and second moments of $\boldsymbol{y}_t$

$$E(\boldsymbol{y}_t) = \boldsymbol{0} \tag{17}$$

$$\Gamma(h) = \begin{cases} 2\boldsymbol{M}\boldsymbol{\Sigma}\,\boldsymbol{M}' & if\ h = 0 \\ -\boldsymbol{M}\boldsymbol{\Sigma}\,\boldsymbol{M}' & if\ |h| = 1 \\ \boldsymbol{0} & otherwise \end{cases} \tag{18}$$

as a by-product of (10) and (11) above.

Such a process can be referred to as an $I(-1)$ process upon the operator identity

$$\nabla = \nabla^{-(-1)} \tag{19}$$

Definition 4

Let the process $\boldsymbol{y}_t$ be specified as a VMA(2) as follows

$$\boldsymbol{y}_t = \boldsymbol{M}\boldsymbol{\varepsilon}_t - 2\boldsymbol{M}\boldsymbol{\varepsilon}_{t-1} + \boldsymbol{M}\boldsymbol{\varepsilon}_{t-2} \tag{20}$$

or, equivalently, as a second difference of a $WN_{(n)}$ process, namely

$$\boldsymbol{y}_t = \boldsymbol{M}\nabla^2\boldsymbol{\varepsilon}_t \tag{21}$$

The following hold for the first and second moments of $\boldsymbol{y}_t$

$$E(\boldsymbol{y}_t) = \boldsymbol{0} \tag{22}$$

$$\Gamma(h) = \begin{cases} 6M\Sigma M' & \text{if } h = 0 \\ -4M\Sigma M' & \text{if } |h| = 1 \\ M\Sigma M' & \text{if } |h| = 2 \\ \mathbf{0} & \text{otherwise} \end{cases} \tag{23}$$

again as a by-product of (10) and (11) above.

Such a process can be read as an $I(-2)$ process upon the operator identity

$$\nabla^2 = \nabla^{-(-2)} \tag{24}$$

Remark

Should q tends to ∞, the VMA(q) process as defined by (8) is referred to as an *infinite causal* – i.e. unidirectional from the present backward to the past – *moving average,* (10) and (14) are still meaningful expressions, and stationarity is maintained accordingly provided both $\lim_{q\to\infty}\sum_{i=0}^{q} M_i$ and $\lim_{q\to\infty}\sum_{i=0}^{q} M_i \otimes M_i$ exist as matrices with finite entries.

Definition 5

A *vector autoregressive process* of order p, written as VAR (p), is a multivariate process y_t specified as follows

$$\underset{(n,1)}{y_t} = \eta + \sum_{j=1}^{p} A_j y_{t-j} + \varepsilon_t, \quad \varepsilon_t \sim WN_{(n)} \tag{25}$$

where η and $A_1, A_2, \ldots, A_p$, are a constant vector and constant matrices, respectively.

Such a process can be rewritten in operator form as follows

$$A(L)\, y_t = \eta + \varepsilon_t, \quad A(L) = I_n - \sum_{j=1}^{p} A_j L^j \tag{26}$$

and it turns out to be stationary provided all roots of the characteristic equation

$$det\, A(z) = 0 \tag{27}$$

lie outside the unit circle (see, e.g., Lütkepohl, 1991). In this circumstance, the polynomial matrix $A^{-1}(z)$ is an analytical (matrix) function about $z = 1$ according to Theorem 1 of Section 1.6, and the process admits a causal VMA (∞) representation, namely

$$y_t = \omega + \sum_{\tau=0}^{\infty} C_\tau \; \varepsilon_{t-\tau} \tag{28}$$

where the matrices C_τ are polynomials in the matrices A_j and the vector ω depends on both the vector η and the matrices C_τ. Indeed the following hold

$$A^{-1}(L) = C\,(L) = \sum_{\tau=0}^{\infty} C_\tau L^\tau \tag{29}$$

$$\omega = A^{-1}(L)\ \eta = (\sum_{\tau=0}^{\infty} C_\tau\ )\,\eta \tag{30}$$

and the expressions of the matrices C_τ can be obtained, by virtue of the isomorphism between polynomials in the lag operator L and in a complex variable z, from the identity

$$\begin{aligned} I &= (C_0 + C_1 z + C_2 z^2 + \ldots)\,(I - A_1 z + \ldots - A_p z^p) \\ &= C_0 + (C_1 - C_0 A_1)\, z + (C_2 - C_1 A_1 - C_0 A_2)\, z^2 \ldots, + \end{aligned} \tag{31}$$

which implies the relashionship

$$\begin{cases} I = C_0 \\ 0 = C_1 - C_0 A_1 \\ 0 = C_2 - C_1 A_1 - C_0 A_2 \\ \ldots \end{cases} \tag{32}$$

The following recursive equations ensue as a by-product

$$C_\tau = \sum_{j=1}^{\tau} C_{\tau-j} A_j \tag{33}$$

The case $p = 1$, which we are going to examine in some details, is of special interest not so much in itself but because of the isomorphic relationship between polynomial matrices and companion matrices (see, e.g., Banjeree et al., 1993; Lancaster and Tismenesky, 1985) which paves the way to bringing a VAR model of arbitrary order back to an equivalent first order VAR model, after a proper reparametrization.

With this premise, consider a first order VAR model specified as follows

$$y_t = \eta + Ay_{t-1} + \varepsilon_t \quad \varepsilon_t \sim \mathrm{WN}_{(n)} \tag{34}$$

where A stands for A_1.

The stationarity condition in this case entails that the matrix A is stable, i.e. all its eigenvalues lie inside the unit circle.

The noteworthy expansion (see, e.g., Faliva, 1987, p. 77)

$$(I - A)^{-1} = I + \sum_{h=1}^{\infty} A^h \tag{35}$$

holds accordingly, and the related expansions

$$(I - Az)^{-1} = I + \sum_{h=1}^{\infty} A^h z^h \Leftrightarrow (I - AL)^{-1} = I + \sum_{h=1}^{\infty} A^h L^h \tag{36}$$

$$[I_{n^2} - A \otimes A]^{-1} = I_{n^2} + \sum_{h=1}^{\infty} A^h \otimes A^h \tag{37}$$

where $|z| \le 1$, ensue as by-products.

By virtue of (36) the VMA (∞) representation of the process (34) takes the form

$$y_t = \omega + \varepsilon_t + \sum_{\tau=1}^{\infty} A^{\tau} \varepsilon_{t-\tau} \tag{38}$$

where

$$\omega = (I - A)^{-1} \eta \tag{39}$$

and the principle moments of the process may be derived accordingly. For what concerns the mean vector, taking expectations of both sides of (38) yields

$$E(y_t) = \omega \tag{40}$$

As far as the autocovariances are concerned, observe first that the following remarkable staked form of the autocovariance of order zero

$$vec\,\Gamma(0) = (I_{n^2} - A \otimes A)^{-1}\, vec\,\Sigma \tag{41}$$

holds true because of (37) as a special case of (14) once M_0 is replaced by I, and M_j is replaced by A^j respectively, and we let q tend to ∞.

Bearing in mind (11) and letting q tend to ∞, simple computations lead to find the following expressions for the higher order autocovariance matrices

$$\Gamma(h) = \Gamma(0)(A')^h \quad for\ h > 0 \tag{42}$$

$$\Gamma(h) = A^{|h|}\, \Gamma(0) \quad for\ h < 0 \tag{43}$$

whence the recursive formulas

$$\Gamma(h) = \Gamma(h-1)\, A' \qquad for\ h > 0 \tag{44}$$

$$\Gamma(h) = A\Gamma'(h-1) \qquad for\ h < 0 \tag{45}$$

follow as a by-product.

The extensions of the conclusions just drawn to higher order VAR processes, rest on the aforementioned companion-form analogue.

The stationary condition on the roots of the characteristic polynomial quoted for a VAR model has a mirror image in the so-called invertibility condition of a VMA model. In this connection we give the following definition

Definition 6

A VMA process is *invertible* if all roots of the characteristic equation

$$det\, M(z) = 0 \tag{46}$$

lie outside the unit circle. In this case the matrix $M^{-1}(z)$ is an analytical matrix function about $z = 1$ by Theorem 1 of Section 1.6, and therefore the process admits a (unique) representation as a function of its past, in the form of a VAR model.

Emblematic examples of non invertible VMA processes are given in Definitions 3 and 4 above.

One should be aware of the fact that it is immaterial to draw a distinction between invertible and non invertible processes for what concerns stationarity.

The property of invertibility is clearly related to the possibility of making predictions since it allows the process y_t to be specified as a convergent function of past random variables.

Should a VMA process be invertible according to Definition 6 above, the following VMA vs. VAR representation holds

$$\boldsymbol{y}_t = \boldsymbol{\mu} + \sum_{j=0}^{q} \boldsymbol{M}_j\, \boldsymbol{\varepsilon}_{t-j} \Rightarrow \boldsymbol{G}(L)\, \boldsymbol{y}_t = \boldsymbol{\nu} + \boldsymbol{\varepsilon}_t \tag{47}$$

where

$$\boldsymbol{\nu} = \boldsymbol{M}^{-1}(L)\, \boldsymbol{\mu} \tag{48}$$

$$\boldsymbol{G}(L) = \sum_{\tau=0}^{\infty} \boldsymbol{G}_\tau L^\tau = \boldsymbol{M}^{-1}(L) \tag{49}$$

The matrices $\boldsymbol{G}_\tau$ may be obtained through the recursive equations

$$\boldsymbol{G}_\tau = \boldsymbol{M}_\tau - \sum_{j=1}^{\tau-1} \boldsymbol{G}_{\tau-j}\, \boldsymbol{M}_j, \quad \boldsymbol{G}_0 = \boldsymbol{M}_0 = \boldsymbol{I} \tag{50}$$

which are the mirror image of the recursive equations (33) and can be obtained in a similar manner.

Taking $q = 1$ in formula (8) yields a VMA (1) model specified as

$$\underset{(n,\,1)}{\boldsymbol{y}_t} = \boldsymbol{\mu} + \boldsymbol{M}\boldsymbol{\varepsilon}_{t-1} + \boldsymbol{\varepsilon}_t, \quad \boldsymbol{\varepsilon}_t \sim WN_{(n)} \tag{51}$$

where $\boldsymbol{M}$ stands for $\boldsymbol{M}_1$.

The following hold for the first and second moments in light of (10) and (11)

$$E(\boldsymbol{y}_t) = \boldsymbol{\mu} \tag{52}$$

$$\boldsymbol{\Gamma}(h) = \begin{cases} \boldsymbol{\Sigma} + \boldsymbol{M}\boldsymbol{\Sigma}\boldsymbol{M}' & \text{if } h = 0 \\ \boldsymbol{\Sigma}\boldsymbol{M}' & \text{if } h = 1 \\ \boldsymbol{M}\boldsymbol{\Sigma} & \text{if } h = -1 \\ \boldsymbol{0} & \text{if } h > 1 \end{cases} \tag{53}$$

The invertibilty condition in this case entails that the matrix $\boldsymbol{M}$ is stable, that is to say all its eigenvalues lie inside the unit circle.

The following noteworthy expansions

$$(\boldsymbol{I} + \boldsymbol{M})^{-1} = \boldsymbol{I} + \sum_{\tau=1}^{\infty} (-1)^{\tau} \boldsymbol{M}^{\tau} \tag{54}$$

$$(\boldsymbol{I} + \boldsymbol{M}z)^{-1} = \boldsymbol{I} + \sum_{\tau=1}^{\infty} (-1)^{\tau} \boldsymbol{M}^{\tau} z^{\tau} \Leftrightarrow (\boldsymbol{I} + \boldsymbol{M}L)^{-1} = \boldsymbol{I} + \sum_{\tau=1}^{\infty} (-1)^{\tau} \boldsymbol{M}^{\tau} L^{\tau} \tag{55}$$

where $|z| \le 1$, hold for the same arguments as (35) and (36) above.

As a consequence of (55), the VAR representation of the process (51) takes the form

$$\boldsymbol{y}_t + \sum_{\tau=1}^{\infty} (-1)^{\tau} \boldsymbol{M}^{\tau} \boldsymbol{y}_{t-\tau} = \boldsymbol{\nu} + \boldsymbol{\varepsilon}_t \tag{56}$$

where

$$\boldsymbol{\nu} = (\boldsymbol{I} + \boldsymbol{M})^{-1} \boldsymbol{\mu} \tag{57}$$

Let us now introduce VARMA models which engender processes combining the characteristics of both VMA and VAR specifications.

Definition 7

A *vector autoregressive moving-average process* of orders p and q (where p is the order of the autoregressive component and q is the order of the moving-average component) – written as VARMA(p, q) – is a multivariate process $\boldsymbol{y}_t$ specified as follows

$$\underset{(n,1)}{\boldsymbol{y}_t} = \boldsymbol{\eta} + \sum_{j=1}^{p} \boldsymbol{A}_j \boldsymbol{y}_{t-j} + \sum_{j=0}^{q} \boldsymbol{M}_j \boldsymbol{\varepsilon}_{t-j}, \quad \boldsymbol{\varepsilon}_t \sim WN_{(n)} \tag{58}$$

where $\boldsymbol{\eta}$, $\boldsymbol{A}_j$ and $\boldsymbol{M}_j$ are a constant vector and constant matrices, respectively.

In operator form the process can be written as follows

$$\boldsymbol{A}(L)\, \boldsymbol{y}_t = \boldsymbol{\eta} + \boldsymbol{M}(L)\, \boldsymbol{\varepsilon}_t, \; \boldsymbol{A}(L) = \boldsymbol{I}_n - \sum_{j=1}^{p} \boldsymbol{A}_j L^j, \; \boldsymbol{M}(L) = \sum_{j=0}^{q} \boldsymbol{M}_j L^j \tag{59}$$

The process is stationary if all roots of the characteristic equation of its autoregressive part, i.e.

$$\det \boldsymbol{A}(z) = 0 \tag{60}$$

lie outside the unit circle. When this is the case, the matrix $A^{-1}(z)$ is an analytical function in a neighbourhood of $z = 1$ by Theorem 1 in Section 1.6 and the process admits a causal VMA (∞) representation, namely

$$y_t = \omega + \sum_{\tau=0}^{\infty} C_\tau \varepsilon_{t-\tau} \tag{61}$$

where the matrices C_τ are polynomials in the matrices A_j and M_j while the vector ω depends on both the vector η and the matrices A_j. Indeed, the following hold

$$\omega = A^{-1}(L)\,\eta \tag{62}$$

$$C(L) = \sum_{\tau=0}^{\infty} C_\tau L^\tau = A^{-1}(L)\,M(L) \tag{63}$$

which, in turn, leads to the recursive formulas

$$C_\tau = M_\tau + \sum_{j=1}^{\tau} A_j C_{\tau-j}\,,\; C_0 = M_0 = I \tag{64}$$

As far as the invertibility property is concerned, reference must be made to the VMA component of the process. The process is invertible if all roots of the characteristic equation

$$det\, M(z) = 0 \tag{65}$$

lie outside the unit circle. Then again the matrix $M^{-1}(L)$ is an analytical function in a neighbourhood of $z = 1$ by Theorem 1 in Section 1.6, and the VARMA process admits a VAR (∞) representation such as

$$G(L)\, y_t = \nu + \varepsilon_t \tag{66}$$

where

$$\nu = M^{-1}(L)\,\eta \tag{67}$$

$$G(L) = \sum_{\tau=0}^{\infty} G_\tau L^\tau = M^{-1}(L)\,A(L) \tag{68}$$

and the matrices G_τ may be computed through the recursive equations

$$G_\tau = M_\tau + A_\tau - \sum_{j=1}^{\tau-1} M_{\tau-j}\, G_j,\;\; G_0 = M_0 = I \tag{69}$$

Taking $p = q = 1$ in formula (58) yields a VARMA (1,1) specified in this way

$$y_t = \eta + Ay_{t-1} + \varepsilon_t + M\varepsilon_{t-1}, \quad A \neq -M \tag{70}$$

where A and M stand for A_1 and M_1 respectively, and the parameter requirement $A \neq -M$ is introduced in order to rule out the degenerate case of a first order dynamic model collapsing into that of order zero.

In this case the stationary condition is equivalent to assuming that the matrix A is stable whereas the invertibilty condition requires the stability of matrix M.

Under stationarity, the following holds

$$y_t = (I - A)^{-1}\eta + (I + \sum_{\tau=1}^{\infty} A^{\tau} L^{\tau})(I + ML)\varepsilon_t \tag{71}$$

which tallies with the VMA (∞) representation (61) once we put

$$\omega = (I + \sum_{\tau=1}^{\infty} A^{\tau})\eta \tag{72}$$

$$C_\tau = \begin{cases} I & if \quad \tau = 0 \\ A + M & if \quad \tau = 1 \\ A^{\tau-1}(A + M) & if \quad \tau > 1 \end{cases} \tag{73}$$

Under invertibility, the following holds

$$(I + ML)^{-1}(y_t - Ay_{t-1}) = (I + M)^{-1}\eta + \varepsilon_t \tag{74}$$

which tallies with the VAR(∞) representation (66) once we put

$$v = \left(I + \sum_{\tau=1}^{\infty} (-1)^{\tau} M^{\tau} \right)\eta \tag{75}$$

$$G_\tau = \begin{cases} I & if \quad \tau = 0 \\ -M - A & if \quad \tau = 1 \\ -(-1)^{\tau-1}M^{\tau-1}(M + A) & if \quad \tau > 1 \end{cases} \tag{76}$$

In order to derive the autocovariance matrices of a general n-dimensional VARMA (p, q) one may transform the model in a $n\,(p+q)$-dimensional VAR (1) by virtue of the already mentioned companion form analogue.

So far we have considered only VAR and VARMA models, whose characteristic polynomial roots lie outside the unit circle.

Nevertheless, the case of a possibly repeated unit-root is worth considering also. As a matter of fact, this proves to stand as a gateway bridging the gap between stationarity and integrated processes as the next section will clarify.

2.3 The Source of Integration and the Seeds of Cointegration

In this section we set out two theorems which bring to the fore the link between the unit-roots of a VAR model and the integration order of the engendered process and disclose the two-faced nature of the model solution with cointegration finally appearing on stage.

Theorem 1

The order of integration of the process $\boldsymbol{y}_t$ generated by a VAR model

$$\boldsymbol{A}(L)\,\boldsymbol{y}_t = \boldsymbol{\eta} + \boldsymbol{\varepsilon}_t \tag{1}$$

whose characteristic polynomial has a possibly repeated unit-root, is the same as the degree of the principal part, i.e. the order of the pole, in the Laurent expansion for $\boldsymbol{A}^{-1}(z)$ in a deleted neighbourhood of $z = 1$.

Proof

A particular solution of the operational equation (1) is given by

$$\boldsymbol{y}_t = \boldsymbol{A}^{-1}(L)\,(\boldsymbol{\varepsilon}_t + \boldsymbol{\eta}) \tag{2}$$

By virtue of the isomorphism existing between the polynomials in the lag operator L and in a complex variable z (see, e.g., Dhrymes, 1971, p. 23), the following holds

$$\boldsymbol{A}^{-1}(z) \Leftrightarrow \boldsymbol{A}^{-1}(L) \tag{3}$$

and the paired expansions

$$\sum_{j=1}^{K} \frac{1}{(1-z)^j} \boldsymbol{N}_j + \sum_{i=0}^{\infty} z^i \boldsymbol{M}_i \Leftrightarrow \sum_{j=1}^{K} \frac{1}{(I-L)^j} \boldsymbol{N}_j + \sum_{i=0}^{\infty} L^i \boldsymbol{M}_i \tag{4}$$

where K stands for the order of the pole of $A^{-1}(z)$ at $z = 1$, are also true.

Because of (4) and by making use of sum-calculus identities such as

$$(I - L)^{-j} = \nabla^{-j} \quad j = 0, 1, 2, \ldots \tag{5}$$

where

$$\nabla^{-1} = \sum_{\tau \le t}, \ \nabla^{-2} = \sum_{\vartheta \le t} \ \sum_{\tau \le \vartheta} \tag{6}$$

the right-hand side of (2) can be given the informative expression

$$\begin{aligned} A^{-1}(L)\,(\varepsilon_t + \eta) &= (N_1 \nabla^{-1} + N_2 \nabla^{-2} + \ldots + N_K \nabla^{-K} + \sum_{j=0}^{\infty} M_j \ L^j (\varepsilon_t + \eta) \\ &= N_1 \sum_{\tau \le t} \varepsilon_\tau + N_2 \sum_{\vartheta \le t} \ \sum_{\tau \le \vartheta} \varepsilon_\tau + \ldots + \sum_{j=0}^{\infty} M_j \ \varepsilon_{t-j} + N_1 \sum_{\tau \le t} \eta \\ &\quad + N_2 \sum_{\tau \le t} (t + 1 - \tau)\,\eta + \ldots + \sum_{j=0}^{\infty} M_j \ \eta \end{aligned} \tag{7}$$

By inspection of (7) the conclusion is easily drawn that the process engendered by the VAR model (1) is composed – stationary components apart – of integrated processes of progressive order.

Hence, the overall effect is that the solution $\boldsymbol{y}_t$ turns out to be an integrated process itself, whose order is the same as the order of the pole of $A^{-1}(z)$, that is to say

$$\boldsymbol{y}_t \sim I\,(K) \tag{8}$$

□

Theorem 2

Let $z = 1$ be a possibly repeated root of the characteristic polynomial $det\,\boldsymbol{A}\,(z)$ of the VAR model

$$\boldsymbol{A}\,(L)\,\boldsymbol{y}_t = \eta + \varepsilon_t \tag{9}$$

and its solution $\boldsymbol{y}_t$ be, correspondingly, an integrated process, say $\boldsymbol{y}_t \sim I\,(d)$ for some $d > 0$.

Furthermore, let

$$\boldsymbol{A} = \boldsymbol{B}\boldsymbol{C}' \tag{10}$$

be a rank factorization of the singular matrix $\boldsymbol{A}\,(1) = \boldsymbol{A}$.

Then the following decomposition holds

$$\boldsymbol{y}_t = \underbrace{(\boldsymbol{C}'_\perp)^g\, \boldsymbol{C}'_\perp \boldsymbol{y}_t}_{\text{maintained integrated component}} + \underbrace{(\boldsymbol{C}')^g\, \boldsymbol{C}'\boldsymbol{y}_t}_{\text{degenerate integrated component}} \tag{11}$$

where the maintained and degenerate components enjoy the integration properties

$$(\boldsymbol{C}'_\perp)^g\, \boldsymbol{C}'_\perp \boldsymbol{y}_t \sim I\,(d) \tag{12}$$

$$(\boldsymbol{C}')^g\, \boldsymbol{C}'\boldsymbol{y}_t \sim I\,(\delta), \quad \delta \le d-1 \tag{13}$$

respectively.

The notion of cointegration fits with the process $\boldsymbol{y}_t$ accordingly.

Proof

In light of (6) of Section 1.3 and of isomorphism between polynomials in a complex variable z and in the lag operator L, the VAR model (9) can be rewritten in the more convenient form

$$\boldsymbol{Q}\,(L)\,\nabla \boldsymbol{y}_t + \boldsymbol{B}\boldsymbol{C}'\boldsymbol{y}_t = \boldsymbol{\eta} + \boldsymbol{\varepsilon}_t \tag{14}$$

where $\boldsymbol{Q}\,(z)$ is as defined by (8) of Section 1.3, and $\boldsymbol{B}$ and $\boldsymbol{C}$ are defined in (10).

Upon noting that

$$\boldsymbol{y}_t \sim I\,(d) \Rightarrow \nabla \boldsymbol{y}_t \sim I\,(d-1) \Rightarrow \boldsymbol{Q}\,(L)\,\nabla \boldsymbol{y}_t \sim I\,(\delta),\, \delta \le d-1 \tag{15}$$

the conclusion

$$\boldsymbol{C}'\boldsymbol{y}_t \sim I\,(\delta) \Leftrightarrow (\boldsymbol{C}')^g\, \boldsymbol{C}'\boldsymbol{y}_t \sim I\,(\delta) \tag{16}$$

is easily drawn, given that

$$\boldsymbol{B}\boldsymbol{C}'\boldsymbol{y}_t = -\,\boldsymbol{Q}\,(L)\nabla \boldsymbol{y}_t + \boldsymbol{\eta} + \boldsymbol{\varepsilon}_t \Leftrightarrow \boldsymbol{C}'\boldsymbol{y}_t = -\,\boldsymbol{B}^g\,\boldsymbol{Q}\,(L)\nabla \boldsymbol{y}_t + \boldsymbol{B}^g\boldsymbol{\eta} + \boldsymbol{B}^g\boldsymbol{\varepsilon}_t \tag{17}$$

by (14) and the integration order of $-\,\boldsymbol{B}^g\,\boldsymbol{Q}\,(L)\nabla \boldsymbol{y}_t + \boldsymbol{B}^g\boldsymbol{\eta} + \boldsymbol{B}^g\boldsymbol{\varepsilon}_t$ is at most that of $\boldsymbol{Q}\,(L)\nabla \boldsymbol{y}_t$, namely $\delta \le d-1$.

Insofar as a drop of integration order occurs when moving from the parent process $\boldsymbol{y}_t$ to its component $(\boldsymbol{C}')^g\, \boldsymbol{C}'\boldsymbol{y}_t$, the latter is a degenerate process with respect to the former.

The analysis of the degenerate component $(\boldsymbol{C}')^g\, \boldsymbol{C}'\boldsymbol{y}_t$ being accomplished, let us examine the complementary component $(\boldsymbol{C}'_\perp)^g\, \boldsymbol{C}'_\perp \boldsymbol{y}_t$.

To this end, observe that by virtue of (42) of Section 1.1, the following identity

$$I = (C'_{\perp})^{g}\, C'_{\perp} + (C')^{g}\, C' \tag{18}$$

holds true and, in turn, leads us to split y_t into two components, as shown in (11).

Since the following integration properties

$$y_t \sim I\,(d) \tag{19}$$

$$(C')^{g}\, C' y_t \sim I\,(\delta) \tag{20}$$

hold in light of the foregoing, the conclusion that the component $(C'_{\perp})^{g}\, C'_{\perp} y_t$ maintains the integration order inherent in the parent process y_t, that is to say

$$(C'_{\perp})^{g}\, C'_{\perp} y_t \sim I\,(d) \tag{21}$$

is eventually drawn.

Finally, in light of (20) and (21), with (19) as a benchmark, the seeds of the concept of cointegration – whose notion and rôle will come to the fore in the next section and in Chapter 3 – are sown.

□

2.4 A Glance at Integrated and Cointegrated Processes

We will introduce the basic notions concerning both integrated and cointegrated processes along with some related results.

Definition 1

A n-dimensional *random-walk* is a multivariate $I\,(1)$ process ξ_t defined after the property

$$\underset{(n,\,1)}{\nabla \xi_t} = \varepsilon_t \quad \varepsilon_t \sim WN_{(n)} \tag{1}$$

The following representations

$$\xi_t = \sum_{\tau \le t} \varepsilon_\tau \tag{2}$$

$$= \xi_0 + \sum_{\tau = 1}^{t} \varepsilon_\tau \tag{2'}$$

hold accordingly, where ξ_0 stands for an initial condition vector, independent from ε_t, $t > 0$, and assumed to have zero mean and finite second moments (see, e.g., Hatanaka, 1996).

The process, while stationary in mean, namely

$$E(\xi_t) = 0 \tag{3}$$

is not covariance stationary, because of

$$E(\xi_t \xi_t') = E(\xi_0 \xi_0') + \Gamma_\varepsilon(0)\, t \tag{4}$$

as a simple computation shows.

Definition 2

A *random walk with drift* is a multivariate $I(1)$ process ξ_t defined as follows

$$\nabla \xi_t = \mu + \varepsilon_t \quad \varepsilon_t \sim WN_{(n)} \tag{5}$$

where μ is a drift vector.

The following representations

$$\xi_t = k\mu + \mu t + \sum_{\tau \le t} \varepsilon_\tau \tag{6}$$

$$= \xi_0 + \mu t + \sum_{\tau=1}^{t} \varepsilon_\tau \tag{6'}$$

hold true, where k and ξ_0 are a scalar and a random vector, respectively, depending on the initial condition and independent from ε_t, $t > 0$. Moreover, ξ_0 is assumed to have finite first and second moments.

The process is neither stationary in mean nor covariance stationary, as simple computations show. In fact

$$E(\xi_t) = E(\xi_0) + \mu t \tag{7}$$

$$V(\xi_t) = V(\xi_0) + \Gamma_\varepsilon(0)\, t \tag{8}$$

where V stands for covariance matrix.

The notion of random walk can be generalized to cover processes whose k-order difference, $k > 1$, leads back to a white noise process.

In this connection, we give the following definition (see also Hansen and Johansen, 1998, p. 110).

Definition 3

By a *cumulated random walk* we mean a multivariate $I(2)$ process defined after the property

$$\nabla^2 \xi_t = \varepsilon_t \quad \varepsilon_t \sim WN_{(n)} \tag{9}$$

The following representations

$$\xi_t = \sum_{\vartheta \le t} \sum_{\tau \le \vartheta} \varepsilon_\tau \tag{10}$$

$$= \sum_{\tau \le t} (t + 1 - \tau)\, \varepsilon_\tau \tag{10'}$$

$$= \sum_{\tau \le 0} (\tau + 1)\, \varepsilon_{t-\tau} \tag{10''}$$

hold true, and the analysis of the process can be carried out along the same line as in Definition 1.

Cumulated random walks with drift can be likewise defined along the lines traced in Definition 2.

Inasmuch as an analogue signal vs. noise (in system theory) and trend vs. disturbances (in time series analysis) is established and noise as well as disturbances stand for non systematic nuisance components, the term signal or trend fits in with any component which exhibits either a regular time path or evolving stochastic swings. Whence the notions of deterministic and stochastic trends which follow.

Definition 4

The term *deterministic trend* will be henceforth used to indicate polynomial functions in the time variable, namely

$$f_t = at + bt^2 + \ldots + dt^r \tag{11}$$

where r is a positive integer and $a, b \ldots, d$ denote parameters.

Linear and quadratic deterministic trends turn out to be of major interest for time series econometrics owing to their connection with random walks with drifts.

Definition 5

By a *stochastic trend* we mean a vector φ_t defined as follows

$$\varphi_t = \sum_{\tau=1}^{t} \varepsilon_\tau , \quad \varepsilon_t \sim WN_{(n)} \tag{12}$$

Upon noting that

$$\nabla \varphi_t = \varepsilon_t \tag{13}$$

the notion of stochastic trend turns out to mirror that of random walk.

Remark

If reference is made to a cumulated random walk, as specified by (9), we can analogously define a second order stochastic trend in this manner

$$\varphi_t = \sum_{\vartheta=1}^{t} \sum_{\tau=1}^{\vartheta} \varepsilon_\tau , \quad \varepsilon_t \sim WN_{(n)} \tag{14}$$

Should a drift enter the underlying random-walk specification, a trend mixing stochastic and deterministic features would occur.

The foregoing offers a first glance at integrated processes and related topics.

A deeper insight into the subject matter, resting on VAR models with unit-roots, will be gained in next chapter, especially Sections 3.4 and 3.5 which are devoted to the so-called representation theorems.

When dealing with several integrated processes, the question may be raised as to whether it would be possible to recover stationarity – besides trivially differencing the said processes – by some sort of a clearing-house-like mechanism, able to lead non stationarities to balance each others out, at least to some extent.

This idea is at the root of cointegration theory which looks for those linear forms of stochastic processes with preassigned integration orders which turn out to be more stationary – possibly, stationary *tout court* – than the original ones.

Here below we will give a few basic notions about cointegration, postponing a closer scrutiny of this fascinating topic to Chapter 3.

Definition 6

The components of a multivariate integrated process y_t form a *cointegrated system* of order (d, b) – with d and b non negative integer numbers such that $d \geq b$ – and we write

$$y_t \sim CI\,(d, b) \tag{15}$$

if the following conditions are fulfilled

i) the n scalar random processes which represent the elements of the vector $\boldsymbol{y}_t$ are integrated of order d, which is tantamount to saying that

$$\underset{(n,1)}{\boldsymbol{y}_t} \sim I(d) \tag{16}$$

ii) there exist one or more (linearly independent) vectors $\boldsymbol{\alpha}$ neither null nor proportional to an elementary vector, such that the linear form

$$\underset{(1,1)}{x_t} = \boldsymbol{\alpha}'\boldsymbol{y}_t \tag{17}$$

is integrated of order $d - b$, i.e.

$$x_t \sim I(d-b) \tag{18}$$

The vectors $\boldsymbol{\alpha}$ are called *cointegration vectors*. The number of cointegration vectors, which are linearly independent, identifies the so-called *cointegration rank* for the process $\boldsymbol{y}_t$.

The basic idea of cointegration is that of describing the stable relations of the economy through linear relations which are more stationary than the variables under consideration.

Observe, in particular, that the class of $CI(1, 1)$ processes is that of $I(1)$ processes which by cointegration give rise to stationary processes.

Definition (6) can be extended to the case of a possibly different order of integration for the components of the vector $\boldsymbol{y}_t$ (see, e.g., Charenza and Deadman, 1992).

In practice, conditions i) and ii) can be reformulated in this way

i) the variables $y_{t1}, y_{t2}, \ldots, y_{tn}$, which represent the elements of the vector $\boldsymbol{y}_t$, are integrated of (possibly) different orders d_h $(h = 1, 2, ..., K)$, with $d_1 > d_2, \ldots, > d_K \geq b$, and these orders are, at least, equal pairwise. By defining the integration order of a vector as the highest integration order of its components, we will simply write

$$\boldsymbol{y}_t \sim I(d_1) \tag{19}$$

ii) For every subset of (two or more) elements of the vector $\boldsymbol{y}_t$, integrated of the same order, there exists al least one cointegration vector by which we obtain – through a linear combination of the previous ones – a variable that is integrated of an order corresponding to that of another subset of (two or more) elements of $\boldsymbol{y}_t$.

As a result there will exist one or more linearly independent vectors $\boldsymbol{\alpha}$ (encompassing the weights of the said linear combinations), neither null nor proportional to an elementary vector, such that the linear form

$$\underset{(1,1)}{x_t} = \alpha' y_t \tag{20}$$

is integrated of order $d_1 - b$, i.e.

$$x_t \sim I(d_1 - b) \tag{21}$$

Let us finally introduce the notion of polynomial cointegration (see, e.g., Johansen, 1995).

Definition 7

The components of a multivariate stochastic process y_t integrated of order $d \geq 2$ form a *polynomially cointegrated system* of order (d, b), where b is a non negative integer satisfying the condition $b \leq d$, and we write

$$y_t \sim PCI(d, b) \tag{22}$$

if there exist vectors α and β_k $(1 \leq k \leq d - b + 1)$ – at least one of them, besides α, neither null nor proportional to an elementary vector – such that the linear form in levels and differences

$$\underset{(1,1)}{z_t} = \alpha' y_t + \sum_{k=1}^{d-b+1} \beta_k' \nabla^k y_t \tag{23}$$

is an integrated process of order $d - b$, i.e.

$$z_t \sim I(d - b) \tag{24}$$

Observe, in particular, that the class of $PCI(2, 2)$ processes is that of $I(2)$ processes which by polynomial cointegration give rise to stationary processes.

Cointegration is actually a cornerstone of time series eeconometrics as the next chapter will show. A quick glance at the rôle of cointegration, in connection with the notion of stochastic trends, will be cast in the appendix of this chapter.

Appendix. Integrated Processes, Stochastic Trends and Rôle of Cointegration

Let

$$\underset{(2,1)}{\xi_t} = \begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix} \sim I(1) \tag{A1}$$

be a bivariate process integrated of order 1, specified as follows

$$\xi_t = A\vartheta_t + \eta_t \tag{A2}$$

where ϑ_t is a vector of stochastic trends

$$\underset{(2,1)}{\vartheta_t} = \sum_{\tau=1}^{t} \varepsilon_\tau , \quad \varepsilon_t \sim WN_{(2)} \tag{A3}$$

and η_t is a bivariate process which is covariance stationary with a null mean, which is tantamount to saying that

$$\underset{(2,1)}{\eta_t} \sim I(0) \tag{A4}$$

Let us suppose that the matrix $A = [a_{ij}]$ is such that $a_{ij} \neq 0$ for $i, j = 1, 2$ and let us assume this matrix to be singular, i.e.

$$r(A) = 1 \tag{A5}$$

Then it follows that

i) the matrix has a null eigenvalue associated with a (left) eigenvector $\boldsymbol{p}'$ such that

$$\boldsymbol{p}'A = \boldsymbol{0}' \tag{A6}$$

ii) the matrix can be factored into two non–null vectors, in terms of the representation

$$A = \underset{(2,1)}{\boldsymbol{b}} \underset{(1,2)}{\boldsymbol{c}'} , \boldsymbol{b}'\boldsymbol{b} \neq 0, \boldsymbol{c}'\boldsymbol{c} \neq 0 \tag{A7}$$

Now, according to (A7), formula (A2) can be rewritten as

$$\xi_t = \boldsymbol{b}\boldsymbol{c}'\vartheta_t + \eta_t \tag{A8}$$

where

$$c'\vartheta_t \sim I(1) \tag{A9}$$

Then, by premultiplying both sides of (A8) by p' we get

$$p'\xi_t = p'bc'\vartheta_t + p'\eta_t = p'\eta_t \sim I(0) \tag{A10}$$

since from formulas (A6) and (A7) it follows that

$$p'b = 0 \Rightarrow p'bc'\vartheta_t = 0 \tag{A11}$$

Finally, by virtue of (A1) and (A10) the conclusion that

$$\xi_t \sim CI(1, 1) \tag{A12}$$

is easily drawn.

Considering the above results we realize that

i) the process ξ_t is integrated of first order owing to the presence of a stochastic trend via the process $c'\vartheta_t$, which plays the rôle of a common trend (cf. Stock and Watson, 1988) and turns out to influence both the components of ξ_t through the (non-null) elements of b;

ii) the vector p (left eigenvector associated with the null eigenvalue of the matrix A) is a cointegration vector for ξ_t since $p'\xi_t$ is stationary;

iii) the cointegrability of ξ_t relies crucially on the annihilation of (common) trends, according to (A11) above.

The very meaning of cointegration is thus that of making immaterial or at least weakening the rôle of the non stationary components.

3 Econometric Dynamic Models: from Classical Econometrics to Time Series Econometrics

3.1 Macroeconometric Structural Models Versus VAR Models

According to the so-called time series econometrics, the typical assumption of classical econometrics about the determinant rôle played by economic theory in model specification is refuted. Therefore, the core of econometric modelling rests crucially on VAR specifications with the addition of integration and cointegration analysis to overcome the problem of non stationary variables and detect possibly stable economic relationships from available data.

This implies that the conceptual frame based upon the interaction among economic theory, mathematics and empirical evidence – provided with the pertinent statistical reading key – which characterizes classical econometrics, leads to a mirror reinterpretation within the time series econometrics. The implication is essentially an overturning between the *specific rôle* of empirical evidence and the *guide rôle* of economic theory.

Thus, whereas the empirical evidence plays a complementary rôle in comparison with economic theory within classical econometrics – about which a *iuris tantum* presumption of *a priori* reliability does indeed exist, although not explicitly expressed – in time series econometrics the perspective is in a certain way overturned. Here are the data – that is the empirical evidence – to outline the frame of reference, while economic theory intervenes with an ancillary rôle to check *a posteriori* the coherence of the results obtained through statistical methods, according to principles accepted by economic theory.

In light of these brief considerations, it is possible to single out the common aspects as well as the distinctive features of the above mentioned approaches to econometric modelling. One could then understand the *modus operandi* of econometric research within both perspectives, when the common denominators are provided by economic theory and by empirical evidence, even though with different hierarchical rôles.

Given these preliminaries, the reader will find in this section a comparison – restricted to the essentials features – between the dynamic specification in the context of VAR modelling and in that of structural econometrics, with a characterization which rests on the assumption that the roots of the characteristic polynomial associated with the model lie inside or on the unitary circle, and with the reduced form as a unifying frame of reference.

Whereas, on the one hand, the reference to the reduced form (as an element of connection) could lead to a restrictive reading key of VAR models, in subordinate terms with respect to structural models, this meaning is no longer applicable when the comparison is made about the nature and the rôle of the roots of the relative characteristic polynomial.

As a starting point it may be convenient to consider the following general *primary form* for the model

$$\underset{(n,1)}{\boldsymbol{y}_t} = \Gamma \boldsymbol{y}_t + \Gamma^*(L)\,\boldsymbol{y}_t + A^*(L)\,\boldsymbol{x}_t + \boldsymbol{\varepsilon}_t \tag{1}$$

where

$$\Gamma^*(L) = \sum_{k=1}^{K} \Gamma_k L^k \tag{2}$$

and

$$A^*(L) = \sum_{r=0}^{R} A_r L^r \tag{3}$$

The notation reflects the one currently used in econometric literature (see, e.g., Faliva, 1987). Here the vectors $\boldsymbol{y}$ and $\boldsymbol{x}$ denote the endogenous and the exogenous variables respectively, $\boldsymbol{\varepsilon}$ represents a white noise vector of disturbances, whereas Γ, Γ_s and A_r stand for matrices of parameters.

Next, we consider first the point of view of classical econometrics and then that of time series econometrics.

The distinctive features of *structural models* are

i) $$\Gamma * \boldsymbol{I} = \boldsymbol{0}, \tag{4}$$

where the symbol $*$ denotes the Hadamard product for matrices (see, e.g., Faliva, 1987, p.86).

ii) Γ, Γ_k $(k = 1, 2, ..., K)$ and A_r $(r = 0, 1, ..., R)$ are sparse matrices, specified according to the economic theory, *ex ante* with respect to model estimation and validation.

While formula (1) expresses the so-called *structural form* of the model, which arises from the transposition of the economic theory into a model, the *secondary (reduced) form* of the model is given by

$$\boldsymbol{y}_t = \boldsymbol{P}(L)\,\boldsymbol{y}_t + \Pi(L)\,\boldsymbol{x}_t + \boldsymbol{\mu}_t \tag{5}$$

where

$$\boldsymbol{\mu}_t = (\boldsymbol{I} - \Gamma)^{-1}\boldsymbol{\varepsilon}_t \tag{6}$$

$$\boldsymbol{P}(L) = \sum_{k=1}^{K} \boldsymbol{P}_k L^k \tag{7}$$

$$\Pi(L) = \sum_{r=0}^{R} \Pi_r L^r \tag{8}$$

with

$$\boldsymbol{P}_k = (\boldsymbol{I}_L - \Gamma)^{-1}\,\Gamma_k,\; k = 1, 2, \ldots, K \tag{9}$$

$$\Pi_r = (\boldsymbol{I}_L - \Gamma)^{-1}\boldsymbol{A}_r,\; r = 0, 1, \ldots, R \tag{10}$$

In a more compact form model (5) may be written as follows

$$\boldsymbol{A}(L)\,\boldsymbol{y}_t = \Pi(L)\,\boldsymbol{x}_t + \boldsymbol{\mu}_t \tag{11}$$

with $\boldsymbol{A}(L)$ defined as

$$\boldsymbol{A}(L) = \boldsymbol{I} - \boldsymbol{P}(L) \tag{12}$$

The spectrum of the characteristic polynomial

$$|\boldsymbol{A}(z)| = det\,[\boldsymbol{I} - \boldsymbol{P}(z)] \tag{13}$$

plays a crucial rôle in the analysis. As a matter of fact, the assumption that all its roots lie outside the unitary circle is indeed a main feature of structural models.

Starting from the reduced form in formula (11), it is possible to obtain, through suitable computations (cf. Faliva, 1987, p. 167), the so–called *final form* of the model, namely

$$\boldsymbol{y}_t = \boldsymbol{H}\boldsymbol{\lambda}^{(t)} + [\boldsymbol{A}(L)]^{-1}\,\Pi(L)\,\boldsymbol{x}_t + [\boldsymbol{A}(L)]^{-1}\,\boldsymbol{\mu}_t \tag{14}$$

Here $\boldsymbol{\lambda}^{(t)}$ denotes the vector

$$\underset{(nK,1)}{\lambda^{(t)}} = \begin{bmatrix} \lambda_1^t \\ \lambda_2^t \\ \vdots \\ \lambda_{nK}^t \end{bmatrix} \tag{15}$$

whose elements are the t-th powers of the solutions $\lambda_1, \lambda_2, \ldots, \lambda_{nK}$ (which are all assumed to be distinct, in order to simplify the formulas) of the equation

$$det\,(\lambda^K \boldsymbol{I}_n - \sum_{k=1}^{K} \lambda^{K-k} \boldsymbol{P}_k) = 0 \tag{16}$$

and $\boldsymbol{H}$ is a matrix whose columns $\boldsymbol{h}_i$ are the non trivial solutions of the homogeneous systems

$$(\lambda_i^K \boldsymbol{I} - \lambda_i^{K-1} \boldsymbol{P}_1 - \ldots - \lambda_i \boldsymbol{P}_{K-1} - \boldsymbol{P}_K)\, \boldsymbol{h}_i = \boldsymbol{0}_n \quad i = 1, 2, \ldots, nK \tag{17}$$

The term $\boldsymbol{H}\lambda^{(t)}$, in the right side of (14), reflects the dynamics of the endogenous variables of inside origin (so–called autonomous component), not due to exogenous or casual factors, which corresponds to the general solution of the homogeneous equation

$$A\,(L)\, \boldsymbol{y}_t = 0 \tag{18}$$

The last two terms in the second member of (14) represent a particular solution of the non-homogeneous equation (11).

The term

$$[A\,(L)]^{-1}\, \Pi\,(L)\, \boldsymbol{x}_t = \sum_{\tau=0}^{\infty} \boldsymbol{K}_\tau \boldsymbol{x}_{t-\tau} \tag{19}$$

reflects the deterministic dynamics, due to exogenous factors (so-called exogenous component), while the term

$$[A\,(L)]^{-1}\, \mu_t = \sum_{\tau=0}^{\infty} C_\tau \mu_{t-\tau} \tag{20}$$

reflects the dynamics induced by casual factors (so-called stochastic component), which assumes the form of a causal moving average *VMA* (∞) of a multivariate white noise, namely a stationary process.

The complete reading key of (14) is illustrated in the following scheme

$$\boldsymbol{y}_t = \underbrace{\underbrace{\underbrace{\boldsymbol{H}\lambda^{(t)}}_{\text{autonomous component}}}_{\text{complementary solution}} + \underbrace{\underbrace{[\boldsymbol{I}-\boldsymbol{P}(L)]^{-1}\,\Pi(L)\,\boldsymbol{x}_t}_{\text{exogeneous component}} + \underbrace{[\boldsymbol{I}-\boldsymbol{P}(L)]^{-1}\mu_t}_{\text{VMA component (stationary process)}}}_{\text{particular solution}}}_{\text{general solution of the non-homogeneous equation}} \tag{21}$$

It is worth mentioning that, in this context, the autonomous component assumes a transitory character which is uninfluential in the long run. This is because the scalars $\lambda_1, \lambda_2, \ldots, \lambda_{nK}$ are the reciprocals of the roots of the characteristic polynomial (13) and as such lie inside the unit circle. As a result, the component $\boldsymbol{H}\lambda^{(t)}$ may be neglected when t is quite high, leading to a more concise representation

$$\boldsymbol{y}_t = \underbrace{[\boldsymbol{I}-\boldsymbol{P}(L)]^{-1}\,\Pi(L)\,\boldsymbol{x}_t}_{\text{exogeneous component}} + \underbrace{[\boldsymbol{I}-\boldsymbol{P}(L)]^{-1}\mu_t}_{\text{VMA component (stationary process)}} \tag{22}$$

What we have seen so far are the salient points in the analysis of linear dynamic models from the viewpoint of classical econometrics.

As far as time series econometrics is concerned, the distinctive features of the *VAR model* are

i) $$\Gamma = \boldsymbol{0} \Rightarrow \Gamma_k = \boldsymbol{P}_k, \quad k = 1, 2, \ldots, K \tag{23}$$

ii) Γ_k ($k = 1, 2, \ldots, K$), are full matrices, in absence of an economic informative theory, *ex ante* with respect to model estimation and validation.

iii) $$\boldsymbol{A}^*(L) = \boldsymbol{0} \Rightarrow \boldsymbol{A}_r = \Pi_r = \boldsymbol{0}, \quad r = 0, 1, \ldots, R \tag{24}$$

As long as the distinction between endogenous and exogenous variables is no longer drawn, all relevant variables – *de facto* – turn out to be treated as endogenous.

Here the *primary and secondary forms* are the same: the model in fact is automatically specified in reduced form, in light of (23) and of (24), i.e.

$$\underset{(n,\,1)}{\boldsymbol{y}_t} = \boldsymbol{P}(L)\,\boldsymbol{y}_t + \varepsilon_t \tag{25}$$

where

$$\boldsymbol{P}(L) = \Gamma^*(L) \tag{26}$$

with the further qualification

$$\Pi(L) = \boldsymbol{A}^*(L) = \boldsymbol{0} \tag{27}$$

The specification of the VAR model, according to (25) and in view of (7), assumes the form

$$\boldsymbol{y}_t = \sum_{k=1}^{K} \boldsymbol{P}_k \boldsymbol{y}_{t-k} + \boldsymbol{\varepsilon}_t \tag{28}$$

According to (12) the following representation holds

$$\boldsymbol{A}(L)\boldsymbol{y}_t = \boldsymbol{\varepsilon}_t \tag{29}$$

The solution of the operational equation (29) – which is the counterpart to the notion of final form of classical econometrics – is the object of the so-called *representation theorems*, and can be given a form such as

$$\boldsymbol{y}_t = \boldsymbol{H}\boldsymbol{\lambda}^{(t)} + \boldsymbol{k}_0 + \boldsymbol{k}_1 t + \boldsymbol{N}_2 \sum_{\tau \le t}(t+1-\tau)\,\boldsymbol{\varepsilon}_\tau + \boldsymbol{N}_1 \sum_{\tau \le t} \boldsymbol{\varepsilon}_\tau + \sum_{i=0}^{\infty} \boldsymbol{M}_i \boldsymbol{\varepsilon}_{t-i} \tag{30}$$

whose rationale will become clear from the subsequent Sections 3.4 and 3.5, which are concerned with specifications of prominent interest for econometrics involving processes integrated up to the second order.

In formula (30) the term $\boldsymbol{H}\boldsymbol{\lambda}^{(t)}$, analogously to what was pointed out for (14), represents the autonomous component which is of transitory character and corresponds to the solution of the homogeneous equation

$$\boldsymbol{A}(L)\boldsymbol{y}_t = \boldsymbol{0} \tag{31}$$

inherent to the roots of the characteristic polynomial which lie outside the unitary circle. Conversely, the term $\boldsymbol{k}_0 + \boldsymbol{k}_1 t$ represents the autonomous component (so-called deterministic trend) which has permanent character insofar as it is inherent to unitary roots.

The other terms represent a particular solution of the non- homogeneous equation (29). Specifically, the term $\sum_{i=0}^{\infty} \boldsymbol{M}_i \boldsymbol{\varepsilon}_{t-i}$ is a causal moving-average process – whose analogy with (20) is evident –associated with the regular part of Laurent expansion of $\boldsymbol{A}^{-1}(z)$ in a deleted neighbourhood of $z = 1$.

The term

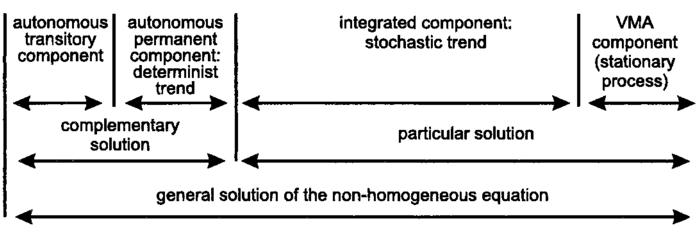
$$N_2 \sum_{\tau \le t} (t+1-\tau)\,\varepsilon_\tau + N_1 \sum_{\tau \le t} \varepsilon_\tau \tag{32}$$

on the other hand, reflects the (random) dynamics, i.e. the stochastic trend or integrated component associated with the principal part of the Laurent expansion of $A^{-1}(z)$ in a deleted neighbourhood of $z = 1$, where $z = 1$ is meant to be a second order pole of $A^{-1}(z)$.

As it will become evident in Sections 3.4 and 3.5, the cointegration relations of the model turn out to be associated with the left eigenvectors corresponding to the null eigenvalues of N_2 and N_1.

The complete reading key of (30) is illustrated in the following scheme

$$y_t = H\lambda^{(t)} + k_0 + k_1 t + N_1 \sum_{\tau \le t} \varepsilon_\tau + N_2 \sum_{\tau \le t} (t+1-\tau)\,\varepsilon_\tau + \sum_{i=0}^{\infty} M_i \varepsilon_{t-i}$$



(33)

3.2 Basic VAR Specifications and Engendered Processes

The validity of VAR specifications to actually grasp the dynamics of economic variables rests on *ad hoc* rank qualifications of the parameter matrices in the reference model.

Before going into the matter in due depth and eventually tackling the major issues of representation theorems content and meaning, let us get an insight into the general setting of unit-root econometrics.

Definition 1

A vector autoregressive (VAR) model

$$\underset{(n,n)}{A(L)}\ \underset{(n,1)}{y_t} = \underset{(n,1)}{\varepsilon_t} + \underset{(n,1)}{\eta}, \quad \varepsilon_t \sim WN_{(n)} \tag{1}$$

where η is a vector of constants (*drift* vector) and

$$A\,(L) = \sum_{k=0}^{K} A_k\, L^k, \quad A_0 = I, A_K \neq 0 \tag{2}$$

is a matrix polynomial whose characteristic polynomial

$$\pi\,(z) = det\, A\,(z) \tag{3}$$

can be factored as

$$\pi\,(z) = (1 - z)^{\alpha}\, \tilde{\pi}\,(z) \tag{4}$$

where $\alpha \geq 0$ is a non negative integer and $\tilde{\pi}\,(z)$ has all roots outside the unit circle, will be referred to as a *basic* VAR model of order K and dimension n.

Definition 2

VAR models can also be specified in terms of both levels and differences by resorting to representations such as

$$Q\,(L)\,\nabla y_t + A y_t = \varepsilon_t + \eta \tag{5}$$

$$\Psi\,(L)\,\nabla^2 y_t - \dot{A}\,\nabla y_t + A y_t = \varepsilon_t + \eta \tag{6}$$

where the symbols have the same meaning as in (8) and (9) of Section 1.3. Such representations are referred to as *error-correction models* (ECM).

The following propositions summarize the fundamental features of VAR-based econometric modelling. Here the proofs, when not derived as by-products of the results presented in Chapters 1 and 2, are justified by material to be found in later sections.

Proposition 1

A basic VAR model, as per Definition 1, engenders a stationary process, i.e.

$$y_t \sim I\,(0) \tag{7}$$

whenever

$$r\,(A\,) = n \tag{8}$$

or, otherwise stated, whenever

$$\alpha = 0 \tag{9}$$

Proof

The proposition ensues from Theorem 1 of Section 1.6 together with Theorem 1 of Section 2.3, after the arguments developed therein.

Thereafter the matrix $\boldsymbol{A}$, even if singular, will always be assumed to be non null.

□

Proposition 2

A basic VAR model, as per Definition 1, engenders an integrated process, i.e.

$$y_t \sim I(d) \tag{10}$$

where d is a positive integer, whenever

$$r(\boldsymbol{A}) < n \tag{11}$$

or, otherwise stated, whenever

$$\alpha > 0 \tag{12}$$

Proof

The proof follows the same line as the proof of Proposition 1.

□

Proposition 3

A basic VAR model, as per Definition 1, engenders a first order integrated process, i.e.

$$y_t \sim I(1) \tag{13}$$

if

$$det(\boldsymbol{A}) = 0 \tag{14}$$

$$det(\boldsymbol{B}_\perp' \dot{\boldsymbol{A}} \boldsymbol{C}_\perp) \neq 0 \tag{15}$$

where $\boldsymbol{B}_\perp$ and $\boldsymbol{C}_\perp$ denote the orthogonal complements of the matrices $\boldsymbol{B}$ and $\boldsymbol{C}$ of the rank factorization

$$\boldsymbol{A} = \boldsymbol{B}\boldsymbol{C}' \tag{16}$$

Proof

The statement ensues jointly from Theorem 2, in view of Corollary 2.1, of Section 1.6 as well as Theorem 1 of Section 2.3, after the arguments developed therein.

□

Proposition 4

Under the assumptions of Proposition 3, the twin processes $\boldsymbol{C}'_{\perp}\boldsymbol{y}_t$ and $\boldsymbol{C}'\boldsymbol{y}_t$ are integrated of first order and stationary respectively, i.e.

$$\boldsymbol{C}'_{\perp}\boldsymbol{y}_t \sim I(1) \tag{17}$$

$$\boldsymbol{C}'\boldsymbol{y}_t \sim I(0) \tag{18}$$

which, in turn, entails the cointegration property

$$\boldsymbol{y}_t \sim CI(1, 1) \tag{19}$$

to hold true for the process $\boldsymbol{y}_t$.

Proof

The proposition mirrors the twin statements (12) and (13) of Theorem 2 in Section 2.3 once we take $d = 1$ and $\delta = d - 1 = 0$. Indeed, writing the VAR model in the ECM form (5), making use of (16) and rearranging term, we get

$$\boldsymbol{B}\boldsymbol{C}'\boldsymbol{y}_t = \boldsymbol{\varepsilon}_t + \boldsymbol{\eta} - \boldsymbol{Q}(L)\nabla\boldsymbol{y}_t \tag{20}$$

Insofar as $\boldsymbol{y}_t \sim I(1)$, the following holds true

$$\boldsymbol{Q}(L)\nabla\boldsymbol{y}_t \sim I(0) \tag{21}$$

which, in turn, entails

$$\boldsymbol{B}\boldsymbol{C}'\boldsymbol{y}_t \sim I(0) \Rightarrow \boldsymbol{C}'\boldsymbol{y}_t \sim I(0) \tag{22}$$

in view of (20).

For what concerns (17), the result is an outcome of the representation theorem for $I(1)$ processes to which Section 3.4 will thereinafter be devoted. After (17) and (18), the conclusion about the cointegrated nature of $\boldsymbol{y}_t$, as per (19), is trivially drawn.

□

Proposition 5

A basic VAR model, as per Definition 1, engenders a second order integrated process, i.e.

$$y_t \sim I(2) \tag{23}$$

if

$$det(A) = 0 \tag{24}$$

$$det(B'_\perp \dot{A} C_\perp) = 0, (B'_\perp \dot{A} C_\perp) \neq \mathbf{0} \tag{25}$$

$$det(R'_\perp B'_\perp \tilde{A} C_\perp S_\perp) \neq 0 \tag{26}$$

where the matrices $B_\perp$ and $C_\perp$ have the same meaning as in Proposition 3, the matrices $R_\perp$ and $S_\perp$ denote the orthogonal complements of R and S in the rank factorization

$$B'_\perp \dot{A} C_\perp = RS' \tag{27}$$

and the matrix $\tilde{A}$ is given by

$$\tilde{A} = \frac{1}{2}\ddot{A} - \dot{A} A^g \dot{A} \tag{28}$$

Proof

The proposition ensues jointly from Theorem 3, in view of Corollary 3.1, of Section 1.6 and Theorem 1 of Section 2.3, in accordance with the arguments developed therein.

□

Proposition 6

Under the assumptions of Proposition 5, the twin processes $S'_\perp C'_\perp y_t$ and $(C_\perp S_\perp)'_\perp y_t$ are integrated of the second and first order, respectively, i.e.

$$S'_\perp C'_\perp y_t \sim I(2) \tag{29}$$

$$(C_\perp S_\perp)'_\perp y_t \sim I(1) \tag{30}$$

which in turn entails the cointegration property

$$y_t \sim CI\,(2, 1) \tag{31}$$

to hold true for the process y_t.

The stronger cointegration property

$$y_t \sim CI\,(2, 0) \tag{32}$$

may also holds true, under the circumstances of Corollary 2.2 of Section 1.7.

Proof

An analogy – although partial – between this proposition and Theorem 2 in Section 2.3 can be drawn bearing in mind the representation (28), Section 1.6, of the orthogonal complement of the product $C_{\perp}S_{\perp}$.

Actually, results (29) and (30) are spin-offs of the representation theorem for I (2) processes to which Section 3.5 will thereinafter be devoted.

After (29) and (30), the conclusion about the cointegrated nature of y_t, as per (31), is trivially drawn. The proof of the second part of the theorem rests on the Corollary 2.2 of Section 1.7 as well as the arguments of the subsequent Section 3.5.

□

3.3 A Sequential Rank Criterion for the Integration Order of a VAR Solution

The following theorem provides a chain rule for the integration order of a process generated by a VAR model on the basis of the rank characteristics of its matrix coefficients.

Theorem 1

Consider a basic VAR model, as per Definition 1 of Section 3.2

$$A\,(L)\,y_t = \varepsilon_t \quad \varepsilon_t \sim WN_{(n)} \tag{1}$$

where the symbols have the same meaning as in the said section, the matrices Γ, $\tilde{\Gamma}$ and Λ are defined as follows

$$\Gamma = [\,A_l^{\perp}\,\dot{A}\,A_r^{\perp} \quad A\,], \quad \tilde{\Gamma} = \begin{bmatrix} A_l^{\perp}\dot{A}A_r^{\perp} \\ A \end{bmatrix} \tag{2}$$

$$\Lambda = (I - \Gamma\Gamma^{g})\,\tilde{A}\,(I - \tilde{\Gamma}^{g}\,\tilde{\Gamma}) \tag{3}$$

and $A_l^{\perp}$ and $A_r^{\perp}$ are as in Definition 7 of Section 1.1.

The following results hold true

i) if

$$r(A) = n \tag{4}$$

then

$$y_t \sim I(0) \tag{5}$$

whereas if

$$r(A) < n \tag{6}$$

then

$$y_t \sim I(d), \quad d > 0 \tag{7}$$

ii) Under rank condition (6), if

$$r([\dot{A}, A]) = n \tag{8}$$

then

$$y_t \sim I(d), \quad d \geq 1 \tag{9}$$

whereas if

$$r([\dot{A}, A]) < n \tag{10}$$

then

$$y_t \sim I(d), \quad d \geq 2 \tag{11}$$

iii)Under rank condition (8), if

$$r(\Gamma) = n \tag{12}$$

then

$$y_t \sim I(1) \tag{13}$$

whereas if

$$r(\Gamma) < n \tag{14}$$

then

$$y_t \sim I(d), \quad d > 1 \tag{15}$$

iv)Under rank condition (14), if

$$r([\tilde{A}, \Gamma]) = n \tag{16}$$

then

$$y_t \sim I(d), \quad d \geq 2 \tag{17}$$

whereas if

$$r([\tilde{A}, \Gamma]) < n \tag{18}$$

then

$$y_t \sim I(d), \quad d \geq 3 \tag{19}$$

v) Under rank condition (16) if

$$r([\Lambda, \quad \Gamma]) = n \tag{20}$$

then

$$y_t \sim I(2) \tag{21}$$

whereas if

$$r([\Lambda, \quad \Gamma]) < n \tag{22}$$

then

$$y_t \sim I(d), \quad d > 2 \tag{23}$$

Proof

To prove i) refer back to Theorem 1 of Section 1.6 and Theorem 1 of Section 2.3 (compare also with Propositions 1 and 2 of the foregoing section).

To prove point ii) refer back to Theorem 2 and 3, as well as to Corollary 2.1 of Section 1.6, and also to Theorem 1 of Section 2.3 (compare also with Propositions 3 and 5 of the foregoing section). Then, observe that

$$r\left(\begin{bmatrix} -\dot{A} & B \\ C' & 0 \end{bmatrix}\right) = r\left[\begin{bmatrix} -I & 0 \\ 0 & B \end{bmatrix}\begin{bmatrix} -\dot{A} & B \\ C' & 0 \end{bmatrix}\begin{bmatrix} I & 0 \\ 0 & -C' \end{bmatrix}\right] = r\left(\begin{bmatrix} \dot{A} & A \\ A & 0 \end{bmatrix}\right) \tag{24}$$

whence

$$r\left(\begin{bmatrix} -\dot{A} & B \\ C' & 0 \end{bmatrix}\right) = n + r(A) \Leftrightarrow r\left(\begin{bmatrix} \dot{A} & A \\ A & 0 \end{bmatrix}\right) = n + r(A) \Rightarrow r([\dot{A}, A]) = n \tag{25}$$

but not necessarily the other way around.

Hence, d is possibly equal to one, as per (9), under (8), whereas this is no longer possible, as per (11), under (10).

The proof of iii) rests on Theorems 14 and 19 in Marsaglia and Styan (1974), Definition 7 of Section 1.1, yielding the rank equalities

$$\begin{aligned} r\left(\begin{bmatrix} -\dot{A} & B \\ C' & 0 \end{bmatrix}\right) &= r\left(\begin{bmatrix} \dot{A} & A \\ A & 0 \end{bmatrix}\right) \\ &= r(A) + r(A) + r\left((I - AA^g)\,\dot{A}\,(I - A^gA)\right) \\ &= r(A) + r(A) + r\left(A_l^{\perp}\dot{A}A_r^{\perp}\right) = r(A) + r(\Gamma) \end{aligned} \tag{26}$$

Hence, after (12) the following holds

$$r\left(\begin{bmatrix} -\dot{A} & B \\ C' & 0 \end{bmatrix}\right) = n + r(A) \tag{27}$$

and (13) follows accordingly, in view of the theorems quoted in proving ii), whereas the circumstance (15) occurs under (14).

To prove iv) refer, on the one hand, back to Theorem 3 – along with its corollary – of Section 1.6 and Theorem 1 of Section 2.3 (compare also with Proposition 5 of the foregoing section) and, on the other hand, to the proof of ii), by replacing

$$\begin{bmatrix} -I & 0 \\ 0 & B \end{bmatrix}, \begin{bmatrix} -\dot{A} & B \\ C' & 0 \end{bmatrix}, \begin{bmatrix} I & 0 \\ 0 & -C' \end{bmatrix} \tag{28}$$

with

$$\begin{bmatrix} I & 0 & 0 \\ 0 & (B_{\perp}')^g R & 0 \\ 0 & 0 & B \end{bmatrix}, \begin{bmatrix} \tilde{A} & (B_{\perp}')^g R & B \\ S'(C_{\perp})^g & 0 & 0 \\ C' & 0 & 0 \end{bmatrix}, \begin{bmatrix} I & 0 & 0 \\ 0 & S'(C_{\perp})^g & 0 \\ 0 & 0 & C' \end{bmatrix} \tag{29}$$

respectively, after the equalities

$$(B_{\perp}R_{\perp})_{\perp} = [(B_{\perp}')^g R, B] \tag{30}$$

$$(C_{\perp}S_{\perp})_{\perp} = [(C_{\perp}')^g S, C] \tag{31}$$

$$(B_{\perp}')^g R S'(C_{\perp})^g = A_l^{\perp}\dot{A}A_r^{\perp} \tag{32}$$

because of (27) and (28) of Section 1.6 and the pairs (47)-(53) and (48)-(54) of Section 1.1.

Next observe that

$$
\begin{aligned}
& r\left(\begin{bmatrix} \tilde{A} & (B_\perp R_\perp)_\perp \\ (C_\perp S_\perp)'_\perp & 0 \end{bmatrix}\right) \\
&= r\left(\begin{bmatrix} I & 0 & 0 \\ 0 & (B'_\perp)^g R & 0 \\ 0 & 0 & B \end{bmatrix}\begin{bmatrix} \tilde{A} & (B'_\perp)^g R & B \\ S'(C_\perp)^g & 0 & 0 \\ C' & 0 & 0 \end{bmatrix}\begin{bmatrix} I & 0 & 0 \\ 0 & S'(C_\perp)^g & 0 \\ 0 & 0 & C' \end{bmatrix}\right) \\
&= r\begin{bmatrix} \tilde{A} & \Gamma \\ \tilde{\Gamma} & 0 \end{bmatrix}
\end{aligned}
\tag{33}
$$

whence

$$
\begin{aligned}
& r\left(\begin{bmatrix} \tilde{A} & (B_\perp R_\perp)_\perp \\ (C_\perp S_\perp)'_\perp & 0 \end{bmatrix}\right) = n + r\,((C_\perp S_\perp)_\perp) \Leftrightarrow \\
& \Leftrightarrow r\begin{bmatrix} \tilde{A} & \Gamma \\ \tilde{\Gamma} & 0 \end{bmatrix} = n + r\,((C_\perp S_\perp)_\perp) \Rightarrow r\,([\tilde{A} \quad \Gamma]) = n
\end{aligned}
\tag{34}
$$

but not necessarily the other way around, given that

$$
r\,(\Gamma) = r\,(\tilde{\Gamma}\,) = r\,(C_\perp S_\perp)_\perp \tag{35}
$$

Hence d is possibly equal to two, as per (17), under (16) whereas this is no longer possible, as per (19) under (18).

The proof of point v) proceeds along the same lines as the proof of point iii) by replacing $A_r^\perp$, $\dot{A}$, $A_l^\perp$, A respectively, with $(I - \tilde{\Gamma}^g\,\tilde{\Gamma}\,)$, $\tilde{A}$, $(I - \Gamma\Gamma^g)$, Γ.

In this connection, observe that

$$
\begin{aligned}
& r\left(\begin{bmatrix} \tilde{A} & (B_\perp R_\perp)_\perp \\ (C_\perp S_\perp)'_\perp & 0 \end{bmatrix}\right) = r\left(\begin{bmatrix} \tilde{A} & \Gamma \\ \tilde{\Gamma} & 0 \end{bmatrix}\right) \\
&= r\,(\Gamma) + r\,(\Gamma) + r\,((I - \Gamma\Gamma^g)\,\tilde{A}\,(I - \tilde{\Gamma}^g\,\tilde{\Gamma}\,)) \\
&= r\,(\Gamma) + r\,([\Lambda \quad \Gamma])
\end{aligned}
\tag{36}
$$

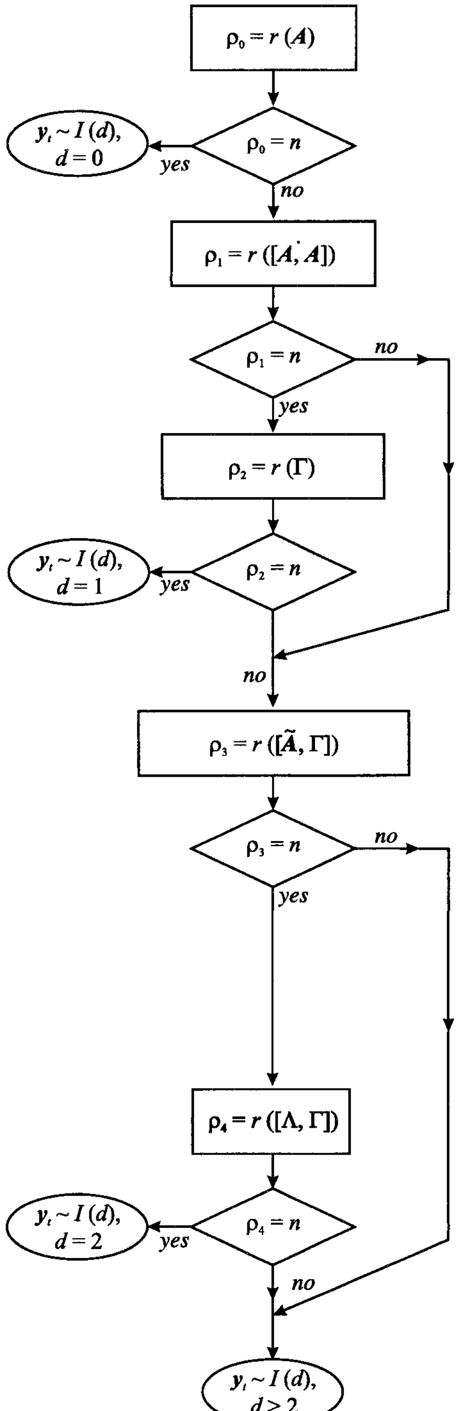
Hence, after (20) the following holds

$$r\left(\begin{bmatrix} \tilde{A} & (B_{\perp}R_{\perp})_{\perp} \\ (C_{\perp}S_{\perp})'_{\perp} & 0 \end{bmatrix}\right) = n + r\,(C_{\perp}S_{\perp})_{\perp} \tag{37}$$

and (21) ensues accordingly, in view of the theorems quoted in proving iv), whereas the circumstances (23) occurs under (22).

□

The sequential procedure for integration order identification inherent in the theorem can be given an enlightening visualization via the decision chart of the next page.


$\rho_0 = r(A)$
$\rho_0 = n$
yes
$y_t \sim I(d)$, $d = 0$
no
$\rho_1 = r([\dot{A}, A])$
$\rho_1 = n$
no
yes
$\rho_2 = r(\Gamma)$
$\rho_2 = n$
yes
$y_t \sim I(d)$, $d = 1$
no
$\rho_3 = r([\tilde{A}, \Gamma])$
$\rho_3 = n$
no
yes
$\rho_4 = r([\Lambda, \Gamma])$
$\rho_4 = n$
yes
$y_t \sim I(d)$, $d = 2$
no
$y_t \sim I(d)$, $d > 2$

3.4 Representation Theorems for Processes *I* (1)

Representation theorems – whose rôle in unit-root econometrics mirrors that of the final form for the dynamic models of structural econometrics – are concerned with the closed form solutions of VAR models in presence of unit-roots, with the inherent reading keys in terms of integrated vs. stationary components of the solutions and cointegration effects as possible offsprings. Such theorems, after Granger's seminal work and the major contributions due to the school named after Johansen, stand as a milestone of the so-called time series econometrics.

Even if the way is, by and large, paved, the underlying analytical setting still presents some subtle facets, which have actually hindered, in some respects, a fully satisfactory treatment of the whole matter.

The remaining of this chapter will be expressly devoted to representation theorems with the aim of shedding proper light on the subject after the arguments developed so far. The clarity of the statements and the fluent structure of the proofs are indebted to the innovative as well as rigorous algebraic apparatus drawn up in the first chapter. Hence, an elegant reappraisal of classical results is combined with original contributions, widening and enriching both the content and the significance of the theorems presented.

This section concentrates on $I\,(1)$ processes, while the next covers $I\,(2)$ processes.

Theorem 1

Consider a VAR model specified as follows

$$\underset{(n,\,n)}{A\,(L)}\boldsymbol{y}_t = \boldsymbol{\varepsilon}_t + \boldsymbol{\eta}, \quad \boldsymbol{\varepsilon}_t \sim WN_{(n)} \tag{1}$$

where $\boldsymbol{\eta}$ is a drift vector and

$$A\,(L) = \sum_{j=0}^{p} A_j L^j, \quad A_0 = I, A_p \neq \boldsymbol{0} \tag{2}$$

is a matrix polynomial whose characteristic polynomial $det A(z)$ is assumed to have a possibly repeated unit-root with all other roots lying outside the unit circle.

Let

$$det \begin{bmatrix} -\dot{A} & B \\ C' & 0 \end{bmatrix} \neq 0 \tag{3}$$

where B and C are defined as per the rank factorization

$$A = BC' \tag{4}$$

of the singular matrix $A\,(1) = A \neq 0$.

Moreover define

$$\begin{bmatrix} P_1 & P_2 \\ P_3 & P_4 \end{bmatrix} = \begin{bmatrix} -\dot{A} & B \\ C' & 0 \end{bmatrix}^{-1} \tag{5}$$

Then, the following representation holds for the process engendered by the model (1) above

$$y_t = k_0 + k_1 t + N_1 \sum_{\tau \leq t} \varepsilon_\tau + M\,(L)\,\varepsilon_t \tag{6}$$

where

$$N_1 = P_1 = -\,C_\perp (B'_\perp\, \dot{A}\, C_\perp)^{-1} B'_\perp \tag{7}$$

$$M\,(L) = \sum_{i=0}^{\infty} M_i\, L^i, \quad M\,(1) = \sum_{i=0}^{\infty} M_i \tag{8}$$

$$\begin{aligned} M\,(1) = -\frac{1}{2} P_1 \ddot{A}\, P_1 + P_2 P_3 = -\frac{1}{2} N_1 \ddot{A}\, N_1 \\ + (I + N_1 \dot{A}\,)\, A^g (I + \dot{A}\, N_1) \end{aligned} \tag{9}$$

$$k_0 = N_1 \nu + M\,(1)\,\eta = C_\perp \tilde{\nu} + M\,(1)\,\eta \tag{10}$$

$$k_1 = N_1 \eta \tag{11}$$

M_0, M_1, …, are coefficient matrices whose entries decrease at an exponential rate, and both ν and $\tilde{\nu}$ denote arbitrary vectors.

Solution (6) represents an integrated process which is inherently cointegrated. Indeed, the following results hold true

i) $$y_t \sim I\,(1) \Rightarrow \nabla y_t \sim I\,(0) \tag{12}$$

ii)
$$C' y_t \sim I(0) \Rightarrow y_t \sim CI(1, 1) \tag{13}$$

Proof

The relationship (1) is nothing but a linear difference equation whose general solution can be represented as (see Theorem 1, Section 1.5)

$$y_t = \left\{ \begin{matrix} \textit{complementary} \\ \textit{solution} \end{matrix} \right\} + \left\{ \begin{matrix} \textit{particular solution of the} \\ \textit{non-homogeneous equation} \end{matrix} \right\} \tag{14}$$

As far as the complementary solution, i.e. the (general) solution of the reduced equation

$$\boldsymbol{A}(L)\, \boldsymbol{y}_t = \boldsymbol{0} \tag{15}$$

is concerned, we have to distinguish between a permanent component associated with a (possibly repeated) unit-root and a transitory component associated with the other roots of the characteristic polynomial

$$|\boldsymbol{A}(z)| = \det \boldsymbol{A}(z) \tag{16}$$

By referring back to Theorem 3 of Section 1.5, the permanent component can be expressed as follows

$$\boldsymbol{\zeta} = \boldsymbol{N}_1 \boldsymbol{\nu} \tag{17}$$

or, in view of (7), as

$$\boldsymbol{\zeta} = \boldsymbol{C}_\perp \tilde{\boldsymbol{\nu}} \tag{18}$$

by taking

$$\tilde{\boldsymbol{\nu}} = -(\boldsymbol{B}'_\perp \dot{\boldsymbol{A}} \boldsymbol{C}_\perp)^{-1} \boldsymbol{B}'_\perp \boldsymbol{\nu} \tag{19}$$

The transitory component of the complementary solution can be expressed as follows

$$\tilde{\boldsymbol{\zeta}}_t = \boldsymbol{H} \boldsymbol{\lambda}^{(t)} \tag{20}$$

where the symbols have the same meaning as in (14)-(17) of Section 3.1 and reference should be made to all roots of the characteristic equation

$$\det \boldsymbol{A}(z) = 0 \tag{21}$$

except for the unit-roots.

Given that the elements of $\boldsymbol{\lambda}^{(t)}$ decrease at an exponential rate, the contribution of component $\tilde{\boldsymbol{\zeta}}_t$ turns out to be ultimately immaterial, and as such it is ignored by the closed-form solution of equation (1) in the right-hand side of (6).

As far as the search for a particular solution for the non-homogeneous equation (1) in the statement of the theorem is concerned, we can refer back to the proof of Theorem 1 of Section 2.3 and write

$$y_t = A^{-1}(L)(\eta + \varepsilon_t) \tag{22}$$

accordingly.

Since, by virtue of hypothesis (3), $A^{-1}(z)$ has a simple pole in $z = 1$ (see Theorem 2 of Section 1.6), in light of (4) of Section 1.4 and of the isomorphism between matrix polynomials in a complex variable z and in the lag operator L, the following Laurent expansion holds

$$A^{-1}(L) = \frac{1}{(I-L)} N_1 + M(L) \tag{23}$$

which implies

$$y_t = \frac{1}{(I-L)} N_1 (\eta + \varepsilon_t) + M(L)(\eta + \varepsilon_t) \tag{24}$$

with N_1 given by (3) in Theorem 1 of Section 1.7.

Because of the formal relationships (see (5) and (6) of Section 2.3)

$$\frac{1}{(I-L)} = \nabla^{-1} = \sum_{\tau \le t} \tag{25}$$

we derive from (24) the elegant closed-form solution

$$y_t = M(1)\,\eta + N_1\eta t + N_1 \sum_{\tau \le t} \varepsilon_\tau + \sum_{j=0}^{\infty} M_j\, \varepsilon_{t-j} \tag{26}$$

where $M(1)$ is given by (4) in Theorem 1 of Section 1.7.

Combining the particular solution, in the right-hand side of (26) of the non-homogeneous equation (1) with the permanent component in the right-hand sides of either (17) or (18), of the complementary solution, we eventually get for the process y_t the representation

$$y_t = \zeta + M(1)\,\eta + N_1\eta t + N_1 \sum_{\tau \le t} \varepsilon_\tau + \sum_{j=0}^{\infty} M_j\, \varepsilon_{t-j} \tag{27}$$

which tallies with (6), by virtue of (10) and (11).

With respect to results *i*) and *ii*), their proofs rest on the following considerations.

Result *i)* – By inspection of (6) we deduce that y_t is the resultant of a drift component k_0, of a deterministic linear trend component $k_1 t$, of a first order stochastic trend component $N_1 \sum_{\tau \le t} \varepsilon_\tau$, and of a VMA ($\infty$) component in the white noise argument ε_t. Therefore, the solution y_t displays the connotation of a first order integrated process, and consequently ∇y_t qualifies as a stationary process.

Result *ii)* – It ensues from (6), in view of (7), through premultiplication of both sides by C'. Because of the orthogonality of C' with N_1 and in view of (11), the terms of $C'y_t$ involving both deterministic and stochastic trends, namely $C'k_1 t$ and $C'N_1 \sum_{\tau \le t} \varepsilon_\tau$, disappear. The non stationary terms being annihilated, the resulting process $C'y_t$ turns out to be stationary.

As long as $y_t \sim I(1)$ and also $C'y_t \sim I(0)$, the solution process y_t turns out to be cointegrated and we can write $y_t \sim CI(1, 1)$, accordingly.

□

The following corollaries highlight some interesting results about the stationary processes ∇y_t and $C'y_t$ associated with the integrated process y_t, i.e. the solution of the VAR model (1). To pave the way to deriving the intended results, we will first define the partitioned matrices

$$\tilde{P}(z) = \begin{bmatrix} P_1(z) & P_2(z) \\ P_3(z) & P_4(z) \end{bmatrix} = \begin{bmatrix} Q(z) & B \\ C' & 0 \end{bmatrix}^{-1} \tag{28}$$

$$\Pi(z) = \begin{bmatrix} \Pi_1(z) & \Pi_2(z) \\ \Pi_3(z) & \Pi_4(z) \end{bmatrix} = \begin{bmatrix} Q(z) & B \\ C' & -(1-z)I \end{bmatrix}^{-1} \tag{29}$$

Both (28) and (29) are meaningful expressions and the matrix functions $\tilde{P}(z)$ and $\Pi(z)$ are matrix polynomials themselves thanks to the Cayley-Hamilton theorem (see, e.g., Rao and Mitra, 1971) provided the inverses in the right-hand sides of (28) and (29) exist, which actually occurs in a neighbourhood of $z = 1$ under the assumptions of Theorem 1 above.

Now, let us prove the following

Lemma

The following relation holds among the blocks of the matrices $\Pi(z)$ and $\tilde{P}(z)$

$$\begin{bmatrix} \Pi_1(z) & \Pi_2(z) \\ \Pi_3(z) & \Pi_4(z) \end{bmatrix} = \begin{bmatrix} P_1(z)+(1-z)P_2(z)[I-(1-z)P_4(z)]^{-1}P_3(z) & P_2(z)[I-(1-z)P_4(z)]^{-1} \\ [I-(1-z)P_4(z)]^{-1}P_3(z) & [I-(1-z)P_4(z)]^{-1}P_4(z) \end{bmatrix} \tag{30}$$

In particular, we have

$$\Pi(1) = \tilde{P}(1) \tag{31}$$

Proof

Upon noting that

$$\begin{bmatrix} Q(z) & B \\ C' & -(1-z)I \end{bmatrix} = \begin{bmatrix} Q(z) & B \\ C' & 0 \end{bmatrix} - \begin{bmatrix} 0 \\ I \end{bmatrix} (1-z)\, I\, [0, \;\; I] \tag{32}$$

straightforward application of result (29) of Section 1.2 implies that

$$\Pi(z) = \tilde{P}(z) + \tilde{P}(z)\begin{bmatrix} 0 \\ I \end{bmatrix} \left\{ \frac{1}{1-z} I - [0, \;\; I]\tilde{P}(z)\begin{bmatrix} 0 \\ I \end{bmatrix} \right\}^{-1} [0, \;\; I]\; \tilde{P}(z) \tag{33}$$

which is tantamount to saying that

$$\begin{bmatrix} \Pi_1(z) & \Pi_2(z) \\ \Pi_3(z) & \Pi_4(z) \end{bmatrix} = \begin{bmatrix} P_1(z) & P_2(z) \\ P_3(z) & P_4(z) \end{bmatrix} + (1-z)\begin{bmatrix} P_2(z) \\ P_4(z) \end{bmatrix} \{I-(1-z)P_4(z)\}^{-1}[P_3(z), \;\; P_4(z)] \tag{34}$$

Now, a simple computation shows that

$$\Pi_3(z) = P_3(z) + (1-z)\, P_4(z)\, [I-(1-z)\, \mathbf{P}_4(z)]^{-1}P_3(z) \tag{35}$$

$$= P_3(z) + \{I - [I-(1-z)\, P_4(z)]\}[I-(1-z)\, P_4(z)]^{-1}P_3(z)$$

$$= \boldsymbol{P}_3(z) + [\boldsymbol{I} - (1-z)\,\boldsymbol{P}_4(z)]^{-1}\boldsymbol{P}_3(z) - \boldsymbol{P}_3(z) = [\boldsymbol{I} - (1-z)\,\boldsymbol{P}_4(z)]^{-1}\boldsymbol{P}_3(z)$$

and, in a similar fashion, that

$$\Pi_2(z) = \boldsymbol{P}_2(z)\,[\boldsymbol{I} - (1-z)\,\boldsymbol{P}_4(z)]^{-1} \tag{36}$$

$$\Pi_4(z) = [\boldsymbol{I} - (1-z)\,\boldsymbol{P}_4(z)]^{-1}\,\boldsymbol{P}_4(z) = \boldsymbol{P}_4(z)\,[\boldsymbol{I} - (1-z)\,\boldsymbol{P}_4(z)]^{-1} \tag{37}$$

The expression for the leading diagonal block

$$\Pi_1(z) = \boldsymbol{P}_1(z) + (1-z)\,\boldsymbol{P}_2(z)\,[\boldsymbol{I} - (1-z)\,\boldsymbol{P}_4(z)]^{-1}\boldsymbol{P}_3(z) \tag{38}$$

is easily obtained.

Then, in light of the foregoing, equality (30) is verified.

Proof of (31) is trivial.

□

We now present the aforementioned corollaries.

Corollary 1.1

Alternative VMA representations of the stationary process $\nabla \boldsymbol{y}_t$ are

$$\nabla \boldsymbol{y}_t = \Pi_1(L)\,(\eta + \varepsilon_t) = N_1\eta + \Pi_1(L)\,\varepsilon_t \tag{39}$$

$$= \delta + \Xi\,(L)\,\varepsilon_t \tag{39'}$$

where $\Pi_1(L)$ is obtained from the leading diagonal block of (29) by replacing z with L, while δ and $\Xi\,(L)$ are given by

$$\delta = N_1\eta \tag{40}$$

$$\Xi\,(L) = \sum_{j=0}^{\infty} \Xi_j L^j = \boldsymbol{M}\,(L)\,\nabla + N_1 \tag{41}$$

The operator relationship

$$\boldsymbol{M}\,(L)\nabla = \Pi_1(L) - N_1 \tag{42}$$

holds accordingly.

Furthermore the following statements are true

i) the matrix polynomial $\Xi\,(z)$ has a simple zero at $z = 1$;

ii) $\nabla \boldsymbol{y}_t$ is a non invertible VMA process;

iii) $$E(\nabla \boldsymbol{y}_t) = N_1\eta \tag{43}$$

Proof

From (7) of Section 1.6, because of the isomorphism between matrix polynomials in a complex variables z and in the lag operator L, this noteworthy relationship holds true

$$A^{-1}(L)\nabla = [\boldsymbol{I},\ \boldsymbol{0}]\begin{bmatrix} \boldsymbol{Q}(L) & \boldsymbol{B} \\ \boldsymbol{C}' & -\nabla \boldsymbol{I} \end{bmatrix}^{-1}\begin{bmatrix} \boldsymbol{I} \\ \boldsymbol{0} \end{bmatrix} \tag{44}$$

whence, in view of (22) and by virtue of (29), the following VMA representation of $\nabla \boldsymbol{y}_t$

$$\begin{aligned}\nabla \boldsymbol{y}_t = [\boldsymbol{I},\ \boldsymbol{0}]\begin{bmatrix} \boldsymbol{Q}(L) & \boldsymbol{B} \\ \boldsymbol{C}' & -\nabla \boldsymbol{I} \end{bmatrix}^{-1}\begin{bmatrix} \boldsymbol{I} \\ \boldsymbol{0} \end{bmatrix}(\eta + \varepsilon_t) &= \Pi_1(L)\,(\eta + \varepsilon_t) \\ &= \boldsymbol{N}_1\eta + \Pi_1(L)\,\varepsilon_t\end{aligned} \tag{45}$$

is obtained in a straightforward manner, upon noting that

$$\Pi_1(1) = \boldsymbol{N}_1 \tag{46}$$

according to (31).

The VMA representation (39′) follows from (6) by elementary computations.

The equality (42) is shown to be true by comparing the right-hand sides of (39) and (39’) in light of (41).

For what concerns results i)-iii), their proofs rest on the following considerations:

Result i) – The matrix polynomial

$$\Xi\,(z) = (1 - z)\,\boldsymbol{M}\,(z) + \boldsymbol{N}_1 \tag{47}$$

has a simple zero at $z = 1$, according to Definition 3 of Section 1.4 and by virtue of Theorem 2, along with Corollary 2.1, of Section 1.6. Indeed the following hold

$$\Xi\,(1) = \boldsymbol{N}_1 \Rightarrow det\,\Xi\,(1) = 0 \tag{48}$$

$$\dot{\Xi}\,(1) = -\boldsymbol{M}\,(1) \tag{49}$$

$$det\,(\boldsymbol{D}_\perp'\,\dot{\Xi}\,(1)\,\boldsymbol{E}_\perp) \neq 0 \Leftrightarrow det\begin{bmatrix} -\dot{\Xi}(1) & \boldsymbol{D} \\ \boldsymbol{E}' & \boldsymbol{0} \end{bmatrix} \neq 0 \tag{50}$$

recalling (18) of Section 1.7, and taking

$$D_{\perp} = -C, \quad E_{\perp} = B \tag{51}$$

where D and E are defined as per a rank factorization of N_1, namely

$$N_1 = DE' \tag{52}$$

$$D = -C_{\perp}(B'_{\perp}\dot{A}C_{\perp})^{-1}, \quad E = B_{\perp} \tag{53}$$

Result ii) – The conclusion is easily drawn by inspection of (39′) because of i) above.

Result iii) – The proof is straightforward by taking the expected value of both sides of (39′) in light of (40).

□

Corollary 1.2

Alternative VMA and VARMA representations, respectively, of the stationary process

$$\gamma_t = C'y_t \tag{54}$$

are

a)
$$\gamma_t = \Pi_3(L)(\eta + \varepsilon_t) = P_3(1)\eta + \Pi_3(L)\varepsilon_t \tag{55}$$

$$= C'M(1)\eta + C'M(L)\varepsilon_t \tag{55′}$$

b)
$$(-P_4(L)\nabla + I)\gamma_t = P_3(1)\eta + P_3(L)\varepsilon_t \tag{56}$$

where, $\Pi_3(L)$, $P_3(L)$ and $P_4(L)$ stand for the homologous blocks of (28) and (29) with z replaced by L.

The following relationships

$$P_3(1) = C'M(1) \tag{57}$$

$$P_3(L) = C'M(L) \tag{58}$$

hold as by-products.

Furthermore, one can establish the equality

$$E\{\gamma_t\} = C'M(1)\eta = B^g(I + \dot{A}N_1)\eta \tag{59}$$

Proof

Proof of a) – Because of

$$\boldsymbol{C}'\,[\boldsymbol{I},\ \ \boldsymbol{0}] = [\boldsymbol{0},\ \ \boldsymbol{I}]\begin{bmatrix} \boldsymbol{Q}(L) & \boldsymbol{B} \\ \boldsymbol{C}' & -\nabla \boldsymbol{I} \end{bmatrix} + \nabla \boldsymbol{I}\,[\boldsymbol{0},\ \ \boldsymbol{I}] \tag{60}$$

the following conclusion

$$\begin{aligned} \boldsymbol{C}'\boldsymbol{A}^{-1}(L)\,\nabla &= \boldsymbol{C}'\,[\boldsymbol{I},\ \ \boldsymbol{0}]\begin{bmatrix} \boldsymbol{Q}(L) & \boldsymbol{B} \\ \boldsymbol{C}' & -\nabla \boldsymbol{I} \end{bmatrix}^{-1}\begin{bmatrix} \boldsymbol{I} \\ \boldsymbol{0} \end{bmatrix} = \\ = \nabla \boldsymbol{I}\,[\boldsymbol{0},\ \ \boldsymbol{I}]&\begin{bmatrix} \boldsymbol{Q}(L) & \boldsymbol{B} \\ \boldsymbol{C}' & -\nabla \boldsymbol{I} \end{bmatrix}^{-1}\begin{bmatrix} \boldsymbol{I} \\ \boldsymbol{0} \end{bmatrix} = \Pi_3(L)\,\nabla \Rightarrow \boldsymbol{C}'\boldsymbol{A}^{-1}(L) = \Pi_3(L) \end{aligned} \tag{61}$$

is easily drawn in light of (44) and (29). Hence, the VMA representation (55) follows, also bearing in mind (31).

The VMA representation (55’) follows from (6) by elementary computations in view of the orthogonality of $\boldsymbol{C}'$ with $\boldsymbol{N}_1$.

Proof of b) – In view of (35), by replacing z with L and $(1 - z)$ with ∇ as usual, the VMA representation (54) can be rewritten as

$$\gamma_t = \{\boldsymbol{I} - \boldsymbol{P}_4(L)\nabla\}^{-1}\boldsymbol{P}_3(L)\,(\eta + \varepsilon_t) \tag{62}$$

Then, premultiplying both sides of (62) by the operator $\boldsymbol{I} - \boldsymbol{P}_4(L)\,\nabla$ the desired representation (56) is easily established.

Eventually, equalities (57) and (58) are proven to be true by comparing the right-hand sides of (55) and (55′).

Finally, as far as (59) is concerned, the proof is straightforward by taking the expected value of both sides of (55′) and by making use of (17) of Section1.7.

□

What is claimed in Theorem 1 and in its corollaries, both reflect and extend the content of the basic representation theorem of time series econometrics.

This theorem can likewise be given a dual version (see, e.g., Banjeree et al. 1993; Johansen, 1995), which originates from a VMA model for the difference process $\nabla \boldsymbol{y}_t$, which in turn underlies a VAR model for the parent process $\boldsymbol{y}_t$ whose integration and cointegration properties can be eventually gathered.

Theorem 2

Consider two stochastic processes ξ_t and y_t, being the former defined as the finite difference of the latter, i.e.

$$\underset{(n,1)}{\xi_t} = \nabla y_t \tag{63}$$

Let ξ_t be stationary and assume a VMA(∞) representation such as

$$\xi_t = \Xi\,(L)\,(\eta + \varepsilon_t) \tag{64}$$

whose parent matrix polynomial $\Xi\,(z) = \Xi_0 + \sum_{i=1}^{\infty} \Xi_i\, z^i$ in the complex argument z is characterized by a first order zero at $z = 1$ and by coefficient matrices Ξ_i with exponentially decreasing entries.

Then the companion process y_t admits a VAR generating model, namely

$$A\,(L)\,y_t = \eta + \varepsilon_t \tag{65}$$

whose parent matrix polynomial $A\,(z)$ in the complex argument z has a first order zero at $z = 1$ and whose characteristic polynomial $det\,A(z)$ has, besides a (possibly multiple) unit-root, all other roots lying outside the unit circle.

The engendered process y_t enjoys the integration and cointegration properties

$$\left.\begin{array}{r} y_t \sim I\,(1) \\ C'y_t \sim I\,(0) \end{array}\right\} \Rightarrow y_t \sim CI\,(1,\,1) \tag{66}$$

where C' is defined as per a rank factorization of $A = A\,(1)$, such as

$$A = BC' \tag{67}$$

Proof

In view of (6) and (8) of Section 1.3 and of the isomorphism between polynomials in a complex variable z and in the lag operator L, we obtain the paired expansions

$$\Xi\,(1) = \tilde{\Phi}\,(z)\,(1 - z) + \Xi\,(1) \Leftrightarrow \Xi\,(L) = \tilde{\Phi}\,(L)\nabla + \Xi\,(1) \tag{68}$$

where

$$\tilde{\Phi}(z) = \sum_{k \geq 1} (-1)^k (1-z)^{k-1} \frac{1}{k!} \Xi^{(k)}(1) \Leftrightarrow \tilde{\Phi}(L) = \sum_{k \geq 1} (-1)^k \frac{1}{k!} \Xi^{(k)}(1) \nabla^{k-1} \tag{69}$$

$$\tilde{\Phi}(1) = -\dot{\Xi}(1) \tag{70}$$

Then, in light of Definition 3 of Section 1.4 and by virtue of Theorem 2 and Corollary 2.1 of Section 1.6, the following hold true

$$det\, \Xi(1) = 0 \tag{71}$$

$$det \left(\begin{bmatrix} -\dot{\Xi}(1) & \tilde{\boldsymbol{B}} \\ \tilde{\boldsymbol{C}}' & \boldsymbol{0} \end{bmatrix} \right) \neq 0 \Leftrightarrow det\, (\tilde{\boldsymbol{B}}'_{\perp} \dot{\Xi}(1) \tilde{\boldsymbol{C}}_{\perp}) \neq 0 \tag{72}$$

where $\tilde{\boldsymbol{B}}$ and $\tilde{\boldsymbol{C}}$ are defined by a rank factorization of $\Xi(1)$, namely

$$\Xi(1) = \tilde{\boldsymbol{B}} \tilde{\boldsymbol{C}}' \tag{73}$$

Given this premise, expansion (4) of Section 1.4 holds for $\Xi^{-1}(z)$ in a deleted neighbourhood of the simple pole $z = 1$ and we can accordingly write

$$\Xi^{-1}(z) = \frac{1}{(1-z)} \tilde{\boldsymbol{N}}_1 + \tilde{\boldsymbol{M}}(z) \Leftrightarrow \Xi^{-1}(L) = \tilde{\boldsymbol{N}}_1 \nabla^{-1} + \tilde{\boldsymbol{M}}(L) \tag{74}$$

where

$$\tilde{\boldsymbol{N}}_1 = -\tilde{\boldsymbol{C}}_{\perp} (\tilde{\boldsymbol{B}}'_{\perp} \dot{\Xi}(1) \tilde{\boldsymbol{C}}_{\perp})^{-1} \tilde{\boldsymbol{B}}'_{\perp} \tag{75}$$

$$\tilde{\boldsymbol{M}}(1) = -\frac{1}{2} \tilde{\boldsymbol{N}}_1 \ddot{\Xi}(1) \tilde{\boldsymbol{N}}_1 + (\boldsymbol{I} + \tilde{\boldsymbol{N}}_1 \dot{\Xi}(1)) \Xi^{g}(1) (\boldsymbol{I} + \dot{\Xi}(1) \tilde{\boldsymbol{N}}_1) \tag{76}$$

in view of Theorem 1 of Section 1.7.

Applying the operator $\Xi^{-1}(L)$ to both sides of (64) yields

$$\Xi^{-1}(L)\, \xi_t = \eta + \varepsilon_t \Rightarrow \boldsymbol{A}(L)\, \boldsymbol{y}_t = \eta + \varepsilon_t \tag{77}$$

namely the VAR representation (65), whose parent matrix polynomial

$$\boldsymbol{A}(z) = \Xi^{-1}(z)(1-z) = (1-z)\tilde{\boldsymbol{M}}(z) + \tilde{\boldsymbol{N}}_1 \tag{78}$$

turns out to have a first order zero at $z = 1$.

This is because of

i) $$\boldsymbol{A} = \tilde{\boldsymbol{N}}_1 \Rightarrow det\, \boldsymbol{A} = 0 \tag{79}$$

whence the rank factorization

$$A = BC' \tag{80}$$

where B and C can conveniently be chosen as

$$B = -\tilde{C}_{\perp}(\tilde{B}'_{\perp}\Xi(1)\tilde{C}_{\perp})^{-1}, \quad C = \tilde{B}_{\perp} \tag{81}$$

ii) The matrices $B_{\perp}$ and $C_{\perp}$, in light of (81), can therefore be chosen as

$$B_{\perp} = \tilde{C}, \quad C_{\perp} = \tilde{B} \tag{82}$$

iii)
$$\dot{A} = -\tilde{M}(1) \tag{83}$$

iv)
$$det(B'_{\perp}\dot{A}C_{\perp}) \neq 0 \Leftrightarrow det\left(\begin{bmatrix} -\dot{A} & B \\ C' & 0 \end{bmatrix}\right) \neq 0 \tag{84}$$

upon noting that, according to (82) and (83) above and (18) of Section 1.7, the following identity holds

$$B'_{\perp}\dot{A}C_{\perp} = -\tilde{C}'\tilde{M}(1)\tilde{B} = -I \tag{85}$$

The proof of what has been claimed about the roots of the characteristic polynomial $det\, A(z)$ rests on the arguments of Theorem 4 of Section 1.4.

Finally, insofar as $A(z)$ has a first order zero at $z = 1$ in light of (79) and (84) above, $A^{-1}(z)$ has a first order pole at $z = 1$.

Hence, by Theorem 1 of Section 2.3 the VAR model (65) engenders an integrated process y_t of the first order, i.e. $y_t \sim I(1)$.

Indeed, according to (78), the following Laurent expansions holds

$$A(L)^{-1} = \frac{1}{(I-L)}N_1 + M(L) = \Xi(L)\nabla^{-1} = \frac{1}{(I-L)}\Xi(1) + \tilde{\Phi}(L) \tag{86}$$

Then, in light of (65), (73) and (86) the following holds true

$$\begin{aligned} y_t &= A(L)^{-1}(\eta + \varepsilon_t) = [\Xi(1)\nabla^{-1} + \tilde{\Phi}(L)](\eta + \varepsilon_t) \\ &= \sum_{i=0}^{\infty}\tilde{\Phi}_i\,\varepsilon_{t-i} + \tilde{B}\tilde{C}'\sum_{\tau \le t}\varepsilon_{\tau} + \tilde{B}\tilde{C}'\eta t + k \end{aligned} \tag{87}$$

where k is a drift vector.

Moreover, because of representation (87), the matrix $\tilde{B}_{\perp} = C$ plays the rôle of matrix of the cointegration vectors since

$$\tilde{B}'_{\perp} y_t = C' y_t \sim I(0) \tag{88}$$

Finally, the cointegration property, $y_t \sim CI(1, 1)$, holds as a by-product. □

3.5 Representation Theorems for Processes *I* (2)

Following the same lines of reasoning as in the previous section, we will provide a neat and rigorous formulation of the representation theorem for $I(2)$ processes, to be followed – as corollaries – by some noteworthy related results.

To conclude we will show how to derive the dual form of this theorem.

Theorem 1

Consider a VAR model specified as follows

$$\underset{(n,n)}{A(L)} y_t = \varepsilon_t + \eta, \quad \varepsilon_t \sim WN_{(n)} \tag{1}$$

where η is a drift vector and

$$A(L) = \sum_{j=0}^{p} A_j L^j, \quad A_0 = I, A_p \neq 0 \tag{2}$$

is a matrix polynomial whose characteristic polynomial $\det A(z)$ is assumed to have a multiple unit-root with all other roots lying outside the unit circle.

Let

a)
$$\det \begin{bmatrix} -\dot{A} & B \\ C' & 0 \end{bmatrix} = 0 \tag{3}$$

where B and C are defined as per the rank factorization

$$A = BC' \tag{4}$$

of the singular matrix $A(1) = A \neq 0$;

b)
$$\det \begin{bmatrix} \tilde{A} & (B_{\perp} R_{\perp})_{\perp} \\ (C_{\perp} S_{\perp})_{\perp} & 0 \end{bmatrix} \neq 0 \tag{5}$$

where

$$\tilde{A} = \frac{1}{2}\ddot{A} - \dot{A}A^{g}\dot{A} \tag{6}$$

and R and S are defined as per the rank factorization

$$\tilde{B}'_{\perp}\,\dot{A}\,C_{\perp} = RS' \tag{7}$$

of the singular matrix $\tilde{B}'_{\perp}\,\dot{A}\,C_{\perp} \neq 0$ with $(B_{\perp}R_{\perp})_{\perp}$ and $(C_{\perp}S_{\perp})_{\perp}$ standing for $[(\tilde{B}'_{\perp})^{g}R,\ \ B]$ and $[(C'_{\perp})^{g}S,\ \ C]$, respectively.

Further, define

$$\begin{bmatrix} J_1 & J_2 \\ J_3 & J_4 \end{bmatrix} = \begin{bmatrix} \tilde{A} & (B_{\perp}R_{\perp})_{\perp} \\ (C_{\perp}S_{\perp})'_{\perp} & 0 \end{bmatrix}^{-1} \tag{8}$$

$$\begin{bmatrix} P_1 & P_2 \\ P_3 & P_4 \end{bmatrix} = \begin{bmatrix} \frac{1}{2}\ddot{A} & \vdots & (B_{\perp}R_{\perp})_{\perp} & & \dot{A}(C')^{g} \\ \cdots & & \cdots & & \cdots \\ (C_{\perp}S_{\perp})'_{\perp} & \vdots & 0 & \vdots & 0 \\ B^{g}\dot{A} & \vdots & 0 & \vdots & I \end{bmatrix}^{-1} \tag{9}$$

Then, the following representation holds for the process engendered by the model (1) above

$$y_t = k_0 + k_1 t + k_2 t^2 + N_1 \sum_{\tau \le t} \varepsilon_\tau + N_2 \sum_{\tau \le t} (t+1-\tau)\varepsilon_\tau + \sum_{i=0}^{\infty} M_i\, \varepsilon_{t-i} \tag{10}$$

where

$$N_2 = P_1 \tag{11}$$

$$= J_1 = C_{\perp}S_{\perp}(R'_{\perp}\,B'_{\perp}\,\tilde{A}\,C_{\perp}S_{\perp})^{-1}R'_{\perp}\,B'_{\perp} \tag{11'}$$

$$N_1 = \frac{1}{6}P_1\dddot{A}P_1 + P_2\overline{\Lambda}P_3 \tag{12}$$

$$= [N_2,\ I - N_2\tilde{A}\,] \begin{bmatrix} \tilde{\tilde{A}} & \dot{A}A^g \\ A^g\dot{A} & -C_\perp(B_\perp'\dot{A}C_\perp)^g B_\perp' \end{bmatrix} \begin{bmatrix} N_2 \\ I - \tilde{A}N_2 \end{bmatrix} \tag{12'}$$

where

$$\overline{\Lambda} = \Lambda(1) - \Lambda(0) \tag{13}$$

$$\tilde{\tilde{A}} = \frac{1}{6}\dddot{A} - \dot{A}A^g\dot{A}A^g\dot{A} \tag{14}$$

$$M(L) = \sum_{i=0}^{\infty} M_i L^i,\quad M(1) = \sum_{i=0}^{\infty} M_i \tag{15}$$

$$k_0 = N_1 v + N_2 w + M(1)\eta \tag{16}$$

$$k_1 = N_1\eta + N_2(v + \frac{1}{2}\eta) \tag{17}$$

$$k_2 = \frac{1}{2}N_2\eta \tag{18}$$

$\Lambda(z)$ is as defined in (22) of Section 1.6, $M_0, M_1, \ldots,$ are coefficient matrices whose entries decrease at an exponential rate, v and w are arbitrary vectors.

The solution (10) represents an integrated process which is inherently cointegrated. Indeed, the following results hold true

i) $$y_t \sim I(2) \Rightarrow \nabla^2 y_t \sim I(0) \tag{19}$$

ii) $$(C_\perp S_\perp)_\perp' y_t \sim I(1) \Rightarrow y_t \sim CI(2, 1) \tag{20}$$

iii) $$B^g \dot{A}(C_\perp')^g C_\perp' \nabla y_t - C' y_t \sim I(0) \Rightarrow y_t \sim PCI(2, 2) \tag{21}$$

iv) $$V_\perp' C' y_t \sim I(0) \Rightarrow y_t \sim CI(2, 2) \tag{22}$$

provided that

$$r(B'\dot{A}\,C_\perp S_\perp) < r(B) \tag{23}$$

with the rank factorization

$$B' \dot{A}\, C_{\perp} S_{\perp} = VW' \tag{24}$$

as a by-product.

Proof

The structure of the proof mirrors that of the representation theorem stated in the previous section.

Once again relationship (1) reads as a linear difference equation, whose general solution can be partitioned as (see Theorem 1 of Section 1.5).

$$y_t = \begin{Bmatrix} \textit{complementary} \\ \textit{solution} \end{Bmatrix} + \begin{Bmatrix} \textit{particular solution of the} \\ \textit{non-homogeneous equation} \end{Bmatrix} \tag{25}$$

As usual, for what concerns the complementary solution, i.e. the (general) solution of the reduced equation

$$A\,(L)\, y_t = 0 \tag{26}$$

we have to distinguish between a permanent component associated with the unit-roots and a transitory component associated with the roots of the characteristic polynomial

$$\det A\,(z) = 0 \tag{27}$$

lying outside the unit circle.

According to Theorem 3 of Section 1.5, the permanent component turns out to take the form

$$\xi_t = N_1 v + N_2 w + N_2 v t \tag{28}$$

where v and w are arbitrary vectors, whereas the transitory component takes the same form as in formula (20) of Section 3.4 and is likewise ignored by the closed-form solution of equation (1) in the right-hand side of (10).

When searching for a particular solution for the non-homogeneous equation (1) in the statement of the theorem, by resorting to Theorem 1 of Section 2.3 we get

$$y_t = A^{-1}(L)\,(\eta + \varepsilon_t) \tag{29}$$

Since, by virtue of the hypotheses (3) and (5), $A^{-1}(z)$ has a second order pole in $z = 1$ (see Theorem 3, together with Corollary 3.1, of Section 1.6), in light of (3) and (11) of Section 1.4 and of the isomorphism between matrix polynomials in a complex variable z and in the lag operator L, the following Laurent expansion holds

$$A^{-1}(L) = \frac{1}{(I-L)^2} N_2 + \frac{1}{(I-L)} N_1 + M(L) \tag{30}$$

which implies the solution

$$y_t = \frac{1}{(I-L)^2} N_2 (\eta + \varepsilon_t) + \frac{1}{(I-L)} N_1 (\eta + \varepsilon_t) + M(L)(\eta + \varepsilon_t) \tag{31}$$

For what concerns the expressions of the coefficient matrices N_2 and N_1, as given by (11)-(11′) and (12)-(12′) respectively, their rationale rests on the following arguments.

Equality (11) follows from (32) of Section 1.6 in light of (29) of Section 1.7 because of (9) above, whereas equality (11') mirrors (22) in Theorem 2 of Section 1.7 with J_1 written for P_1.

To prove equality (12) reference must be made, on the one hand, to (33) of Section 1.7 and, on the other, to result (49) of the lemma hereinafter established, bearing in mind (9) above as well as (42), (43) and (47) below.

Conversely, equality (12′) tallies with (23) in Theorem 2 of Section 1.7.

By virtue of sum-calculus identities (see (5) and (6) of Section 2.3)

$$\frac{1}{(I-L)} = \nabla^{-1} = \sum_{\tau \le t}, \quad \frac{1}{(I-L)^2} = \nabla^{-2} = \sum_{\vartheta \le t}\sum_{\tau \le \vartheta} = \sum_{\tau \le t}(t+1-\tau) \tag{32}$$

we get from (30) the elegant closed-form solution

$$\begin{aligned} y_t = {} & M(1)\eta + (N_1 + \frac{1}{2}N_2)\eta t + \frac{1}{2}N_2\eta t^2 + N_1 \sum_{\tau \le t} \varepsilon_\tau \\ & + N_2 \sum_{\tau \le t}(t+1-\tau)\,\varepsilon_\tau + \sum_{i=0}^{\infty} M_i\, \varepsilon_{t-i} \end{aligned} \tag{33}$$

Combining the particular solution in the right-hand side of (33) of the non-homogeneous equation (1), with the permanent component in the right-hand side of (28) of the complementary solution, we eventually get for the process y_t the representation

$$\begin{aligned} y_t = {} & \xi_t + M(1)\,\eta + (N_1 + \frac{1}{2}N_2)\,\eta t + \frac{1}{2}N_2\eta t^2 + N_1 \sum_{\tau \le t} \varepsilon_\tau \\ & + N_2 \sum_{\tau \le t}(t+1-\tau)\,\varepsilon_\tau + \sum_{i=0}^{\infty} M_i\, \varepsilon_{t-i} \end{aligned} \tag{34}$$

which tallies with (10), in view of (16), (17) and (18).

As far as results *i)*-*iv)* are concerned, their proofs rest on the following considerations.

Result *i)* – By inspection of (10) we deduce that $\boldsymbol{y}_t$ is the resultant of a drift component $\boldsymbol{k}_0$, of deterministic linear and quadratic trend components, $\boldsymbol{k}_1 t$ and $\boldsymbol{k}_2 t^2$ respectively, of first and second order stochastic trend components, $N_1 \sum_{\tau \le t} \varepsilon_\tau$ and $N_2 \sum_{\tau \le t} (t+1-\tau)\, \varepsilon_\tau$ respectively, and of a VMA(∞) component in the white noise argument ε_t. The overall effect is that of a second order integrated process $\boldsymbol{y}_t$ whence $\nabla^2 \boldsymbol{y}_t$ qualifies as a stationary process as a by-product.

Result *ii)* – It ensues from (10), in view of (11′), through premultiplication on both sides by $(C_\perp S_\perp)'_\perp$. Because of the orthogonality of $(C_\perp S_\perp)'_\perp$ with N_2 and in view of (18), the terms of $(C_\perp S_\perp)'_\perp \boldsymbol{y}_t$ involving quadratic deterministic as well as second order stochastic trends, namely $\boldsymbol{k}_2 t^2$ and $N_2 \sum_{\tau \le t} (t+1-\tau)\, \varepsilon_\tau$, cancel out. The higher order non stationary components being annihilated, the integration order of the resulting process lessens, that is $(C_\perp S_\perp)'_\perp \boldsymbol{y}_t \sim I\,(1)$.

Since $\boldsymbol{y}_t \sim I\,(2)$ and $(C_\perp S_\perp)'_\perp \boldsymbol{y}_t \sim I\,(1)$, the cointegration property $\boldsymbol{y}_t \sim CI\,(2, 1)$ holds true.

Result *iii)* – To prove (21) observe that

a) on the one hand, premultiplying both sides of (10) by C' yields

$$C'\boldsymbol{y}_t = C'\boldsymbol{k}_0 + C'N_1 \eta t + C'N_1 \nabla^{-1} \varepsilon_t + C'M\,(L)\, \varepsilon_t \tag{35}$$

because of the orthogonality of C' with N_2 and in view of (17) and (18).

b) On the other hand, differencing both sides of (10) gives

$$\nabla \boldsymbol{y}_t = \boldsymbol{k}_1 - \boldsymbol{k}_2 + N_2 \eta t + N_1 \varepsilon_t + N_2 \nabla^{-1} \varepsilon_t + M\,(L)\, \nabla \varepsilon_t \tag{36}$$

By inspection of the right-hand sides of (35) and (36) it is apparent that both $C'\boldsymbol{y}_t$ and $\nabla \boldsymbol{y}_t$ are $I\,(1)$ processes, since both expressions contains first order stochastic as well as deterministic trends. This suggests the possibility of arriving at a stationary process by a linear combination of the said processes whose non stationary component balance out.

Now, let D be a matrix of the same dimension as C' and consider the linear form

$$C'\boldsymbol{y}_t + D\nabla \boldsymbol{y}_t = (C'N_1 + DN_2)\, \eta t + (C'N_1 + DN_2)\, \nabla^{-1} \varepsilon_t \\ + \textit{drifts and stationary components} \tag{37}$$

The deterministic and stochastic trends in the right side of (37) vanish provided the following equality holds

$$\boldsymbol{C}'\boldsymbol{N}_1 + \boldsymbol{D}\boldsymbol{N}_2 = \boldsymbol{0} \tag{38}$$

which occurs, in light of (80) of Section 1.7, if

$$\boldsymbol{D} = -\boldsymbol{B}^g \dot{\boldsymbol{A}} \tag{39}$$

or likewise if

$$\boldsymbol{D} = -\boldsymbol{B}^g \dot{\boldsymbol{A}}\,(\boldsymbol{C}_\perp')^g \boldsymbol{C}_\perp' \tag{40}$$

thanks to the identity

$$(\boldsymbol{C}_\perp')^g \boldsymbol{C}_\perp' \boldsymbol{N}_2 = \boldsymbol{N}_2 \tag{41}$$

Hence both $\boldsymbol{C}'\boldsymbol{y} - \boldsymbol{B}^g \dot{\boldsymbol{A}}\,\nabla \boldsymbol{y}_t$ and $\boldsymbol{C}'\boldsymbol{y}_t - \boldsymbol{B}^g \dot{\boldsymbol{A}}\,(\boldsymbol{C}_\perp')^g \boldsymbol{C}_\perp' \nabla \boldsymbol{y}_t$ are free of non stationary components, *iii*) holds accordingly and the parent process $\boldsymbol{y}_t$, solution of the VAR model (1), turns out to be polynomially cointegrated.

Result *iv*) – This can be deduced from (10) through a premultiplication of both sides by $\boldsymbol{V}_\perp' \boldsymbol{C}'$, under the rank condition (23), after the arguments put forward in Corollary 2.2 of Section 1.7.

Insofar as $\boldsymbol{V}_\perp' \boldsymbol{C}'$ turns out to be orthogonal with both $\boldsymbol{N}_2$ and $\boldsymbol{N}_1$ as per (82) of Section 1.7, the terms of $\boldsymbol{V}_\perp' \boldsymbol{C}'\boldsymbol{y}_t$ involving both deterministic and stochastic trends of second as well as first order, namely $\boldsymbol{k}_1 t$, $\boldsymbol{k}_2 t^2$, $\boldsymbol{N}_2 \sum_{\tau \le t} (t+1-\tau)\,\boldsymbol{\varepsilon}_\tau$ and $\boldsymbol{N}_1 \sum_{\tau \le t} \boldsymbol{\varepsilon}_\tau$ disappear. The non stationary terms being annihilated, the resulting process $\boldsymbol{V}_\perp' \boldsymbol{C}'\boldsymbol{y}_t$ turns out to be stationary.

Thus, on the one hand, we have $\boldsymbol{y}_t \sim I\,(2)$, on the other, under the rank condition (23) and the rank factorization (24), we get $\boldsymbol{V}_\perp' \boldsymbol{C}'\boldsymbol{y}_t \sim I\,(0)$. Hence the sharper cointegration property $\boldsymbol{y}_t \sim CI\,(2, 2)$ holds true.

□

The corollaries which follow provide an insight into the stationary processes obtained either by differencing or cointegrating the solution of the VAR model (1) of Theorem 1.

Let us first introduce the partitioned matrices

$$\widetilde{\boldsymbol{P}}(z) = \begin{bmatrix} \boldsymbol{P}_1(z) & \boldsymbol{P}_2(z) \\ \boldsymbol{P}_3(z) & \boldsymbol{P}_4(z) \end{bmatrix} = \begin{bmatrix} \boldsymbol{\Psi}(z) & \boldsymbol{F} \\ \boldsymbol{G}' & \boldsymbol{\Lambda}(1) \end{bmatrix}^{-1} \tag{42}$$

$$\Pi(z) = \begin{bmatrix} \Pi_1(z) & \Pi_2(z) \\ \Pi_3(z) & \Pi_4(z) \end{bmatrix} = \begin{bmatrix} \Psi(z) & F \\ G' & \Lambda(z) \end{bmatrix}^{-1} \tag{43}$$

where $\Psi(z)$, F, G and $\Lambda(z)$ are as defined by (9) of Section 1.3 and by (20) and (22) of Section 1.6.

Both (42) and (43) are meaningful expressions and the matrix functions $\tilde{P}(z)$ and $\Pi(z)$ are matrix polynomials themselves thanks to the Cayley-Hamilton theorem (see, e.g., Rao and Mitra, 1971) provided the inverses in the right-hand sides of (42) and (43) exist, which actually occurs in a neighbourhood of $z = 1$ on the basis of the arguments of Theorem 1 above.

Further, matrix $\Lambda(z)$ can be expressed as follows

$$\Lambda(z) = \Lambda(1) - (1 - z)\overline{\Lambda} \tag{44}$$

where $\overline{\Lambda}$ is as defined by (13), and enjoys the properties

$$det\, \Lambda(0) \neq 0 \tag{45}$$

$$det\, (\overline{\Lambda}) \neq 0 \tag{46}$$

as simple computations show.

The following lemma proves useful in paving the way to the representation theorem we are interested in.

Lemma

The following relation holds among the blocks of the matrices $\Pi(z)$ and $\tilde{P}(z)$

$$\begin{aligned} &\begin{bmatrix} \Pi_1(z) & \Pi_2(z) \\ \Pi_3(z) & \Pi_4(z) \end{bmatrix} \\ &= \begin{bmatrix} P_1(z) + \lambda P_2(z)\overline{\Lambda}[I - \lambda P_4(z)\overline{\Lambda}]^{-1}P_3(z) & P_2(z)[I - \lambda\overline{\Lambda}P_4(z)]^{-1} \\ [I - \lambda P_4(z)\overline{\Lambda}]^{-1}P_3(z) & P_4(z)[I - \lambda\overline{\Lambda}P_4(z)]^{-1} \end{bmatrix} \end{aligned} \tag{47}$$

where λ is written for $1 - z$.

In particular we have

$$\Pi(1) = \tilde{P}(1) \tag{48}$$

$$\dot{\Pi}_1(1) = \frac{1}{6} P_1(1)\ddot{A} P_1(1) - P_2(1)\overline{\Lambda} P_3(1) \tag{49}$$

where $\dot{\Pi}_1(1)$ denotes the derivative of $\Pi_1(z)$, evaluated at $z = 1$

Proof

The proof of (47) follows along the same lines as the proof of (30) of Section 3.4, upon noting that

$$\begin{bmatrix} \Psi(z) & F \\ G' & \Lambda(z) \end{bmatrix} = \begin{bmatrix} \Psi(z) & F \\ G' & \Lambda(1) \end{bmatrix} - \begin{bmatrix} 0 \\ I \end{bmatrix} (1-z)\overline{\Lambda}\,[0 \quad I] \tag{50}$$

In view of (46) and of the non singularity of $\tilde{P}(z)$ in a neighbourhood of $z = 1$, inversion formula (29) of Section 1.2 applies and we can therefore write

$$\Pi(z) = \tilde{P}(z) + \tilde{P}(z)\begin{bmatrix} 0 \\ I \end{bmatrix} \left\{ \frac{1}{(1-z)}\overline{\Lambda}^{-1} - [0 \quad I]\tilde{P}(z)\begin{bmatrix} 0 \\ I \end{bmatrix} \right\}^{-1} [0 \quad I]\tilde{P}(z) \tag{51}$$

whence in particular, by simple computation, we get

$$\Pi_1(z) = P_1(z) + (1-z)\, P_2(z)\overline{\Lambda}\,[I - (1-z)\, P_4(z)\overline{\Lambda}\,]^{-1} P_3(z) \tag{52}$$

$$\begin{aligned} \Pi_3(z) &= P_3(z) + (1-z)\, P_4(z)\overline{\Lambda}\,[I - (1-z)\, P_4(z)\overline{\Lambda}\,]^{-1} P_3(z) \\ &= P_3(z) + \{I - [I - (1-z)\, P_4(z)\overline{\Lambda}\,]\}[I - (1-z)\, P_4(z)\overline{\Lambda}\,]^{-1} P_3(z) \\ &= P_3(z) + [I - (1-z)\, P_4(z)\overline{\Lambda}\,]^{-1} P_3(z) - P_3(z) \\ &= [I - (1-z)\, P_4(z)\overline{\Lambda}\,]^{-1} P_3(z) \end{aligned} \tag{53}$$

In a similar fashion, we get the expressions for the remaining blocks, i.e.

$$\Pi_2(z) = P_2(z)\, [I - (1-z)\, \overline{\Lambda} P_4(z)]^{-1} \tag{54}$$

$$\Pi_4(z) = [I - (1-z)\, P_4(z)\overline{\Lambda}\,]^{-1} P_4(z) = P_4(z)\, [I - (1-z)\, \overline{\Lambda} P_4(z)]^{-1} \tag{55}$$

The proof of (48) is straightforward in light of (47).

For what concerns (49), taking the derivative with respect to z of both sides of (52) yields

$$\dot{\Pi}_1(z) = \dot{P}_1(z) - P_2(z)\overline{\Lambda}\,[I - (1-z)\, P_4(z)\overline{\Lambda}\,]^{-1} P_3(z) \tag{56}$$

$$+ (1 - z) \frac{d\{P_2(z)\overline{\Lambda}[I - (1 - z)P_4(z)\overline{\Lambda}]^{-1}P_3(z)\}}{dz}$$

which eventually leads to (49) by replacing z with 1, given that

$$\dot{P}_1(z) = [I \quad 0] \frac{d\begin{bmatrix} \Psi(z) & F \\ G' & \Lambda(1) \end{bmatrix}^{-1}}{dz} \begin{bmatrix} I \\ 0 \end{bmatrix}$$

$$= - [I \quad 0] \begin{bmatrix} P_1(z) & P_2(z) \\ P_3(z) & P_4(z) \end{bmatrix} \frac{d\begin{bmatrix} \Psi(z) & F \\ G' & \Lambda(1) \end{bmatrix}}{dz} \begin{bmatrix} P_1(z) & P_2(z) \\ P_3(z) & P_4(z) \end{bmatrix} \begin{bmatrix} I \\ 0 \end{bmatrix} \tag{57}$$

$$= - [I \quad 0] \begin{bmatrix} P_1(z) & P_2(z) \\ P_3(z) & P_4(z) \end{bmatrix} \begin{bmatrix} \dot{\Psi}(z) & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} P_1(z) & P_2(z) \\ P_3(z) & P_4(z) \end{bmatrix} \begin{bmatrix} I \\ 0 \end{bmatrix}$$

$$= - P_1(z)\dot{\Psi}(z)P_1(z)$$

and that

$$\dot{\Psi}(1) = \frac{1}{6}\dddot{A} \tag{58}$$

according to (45) of Section 1.7.

□

We will now establish the aforementioned corollaries.

Corollary 1.1

Alternative VMA representations of the stationary process $\nabla^2 y_t$ are

$$\nabla^2 y_t = \Pi_1(L)(\eta + \varepsilon_t) = N_2\eta + \Pi_1(L)\,\varepsilon_t \tag{59}$$

$$= \delta + \Xi(L)\,\varepsilon_t \tag{59'}$$

where $\Pi_1(L)$ is obtained from the leading diagonal block of (43) by replacing z with L, while δ and $\Xi(L)$ are given by

$$\delta = N_2\eta \tag{60}$$

$$\Xi(L) = \sum_{j=0}^{\infty} \Xi_j L^j = M(L)\nabla^2 + N_1\nabla + N_2 \tag{61}$$

Moreover, the operator relationship

$$M(L)\,\nabla^2 = \Pi_1(L) - N_1\nabla - N_2 \tag{62}$$

holds.

Furthermore, the following statements are true

i) the matrix polynomial $\Xi\,(L)$ has a second order zero at $z = 1$;

ii) $\nabla^2 y_t$ is a non invertible VMA process;

iii) $E\,(\nabla^2 y_t) = N_2\eta$

Proof

The proof is similar to that of Corollary 1.1 of Section 3.4.

Thanks to (25) of Section 1.6 and by virtue of the isomorphism between polynomials in a complex variable z and in the lag operator L, the following equality proves to be true

$$A^{-1}(L)\nabla^2 = [I \quad 0]\begin{bmatrix} \Psi(L) & F \\ G' & \Lambda(L) \end{bmatrix}^{-1}\begin{bmatrix} I \\ 0 \end{bmatrix} \tag{63}$$

This, in view of (29) and by virtue of (43), leads to the intended representation for $\nabla^2 y_t$, namely

$$\begin{aligned}\nabla^2 y_t = [I \quad 0]\begin{bmatrix} \Psi(L) & F \\ G' & \Lambda(L) \end{bmatrix}^{-1}\begin{bmatrix} I \\ 0 \end{bmatrix}(\eta + \varepsilon_t) &= \Pi_1(L)\,(\eta + \varepsilon_t) \\ &= N_2\eta + \Pi_1(L)\,\varepsilon_t \end{aligned} \tag{64}$$

upon noting that

$$\Pi_1\,(1) = N_2 \tag{65}$$

in light of (29) of Section 1.7.

The VMA representation (59′), as well as (60) and (61), follows from (10) by elementary computations.

The equality (62) proves true by comparing the right-hand sides of (59) and (59′) in view of (61).

As far as statements i)-iii), are concerned, their proofs rest on the arguments developed here below.

Result i) – The matrix polynomial

$$\Xi\,(z) = (1 - z)^2 M(z) + (1 - z)\,N_1 + N_2 \tag{66}$$

has a second order zero at $z = 1$, according to Definition 3 of Section 1.4 and by virtue of Theorem 3, along with Corollary 3.1, of Section 1.6. Indeed the following hold

$$\Xi(1) = \boldsymbol{N}_2 \Rightarrow det\, \Xi(1) = 0 \tag{67}$$

$$\dot{\Xi}(1) = -\boldsymbol{N}_1, \quad \ddot{\Xi}(1) = \frac{1}{2}\boldsymbol{M}_1 \tag{68}$$

$$det\,(\boldsymbol{D}_\perp' \,\dot{\Xi}(1)\, \boldsymbol{E}_\perp) = 0 \Leftrightarrow det \begin{bmatrix} -\dot{\Xi}(1) & \boldsymbol{D} \\ \boldsymbol{E}' & \boldsymbol{0} \end{bmatrix} = 0 \tag{69}$$

$$det\,(\boldsymbol{U}_\perp' \,\boldsymbol{D}_\perp' \,\tilde{\Xi}(1)\, \boldsymbol{E}_\perp \boldsymbol{U}_\perp) \neq 0 \tag{70}$$

where

a) the matrices $\boldsymbol{D}$ and $\boldsymbol{E}$ are defined as per a rank factorization of $\Xi(1)$, namely

$$\Xi(1) = \boldsymbol{N}_2 = \boldsymbol{D}\boldsymbol{E}', \quad \boldsymbol{D} = \boldsymbol{C}_\perp \boldsymbol{S}_\perp (\boldsymbol{R}_\perp' \,\boldsymbol{B}_\perp' \,\tilde{\boldsymbol{A}}\, \boldsymbol{C}_\perp \boldsymbol{S}_\perp)^{-1}, \quad \boldsymbol{E} = \boldsymbol{B}_\perp \boldsymbol{R}_\perp \tag{71}$$

and thus the matrices $\boldsymbol{D}_\perp$ and $\boldsymbol{E}_\perp$ can be chosen as

$$\boldsymbol{D}_\perp = (\boldsymbol{C}_\perp \boldsymbol{S}_\perp)_\perp, \quad \boldsymbol{E}_\perp\, (\boldsymbol{B}_\perp \boldsymbol{R}_\perp)_\perp \tag{72}$$

b) The matrix $\boldsymbol{U}$ is defined as per a rank factorization of $\boldsymbol{D}_\perp' \,\dot{\Xi}(1)\, \boldsymbol{E}_\perp$, namely

$$\boldsymbol{D}_\perp' \,\dot{\Xi}(1)\, \boldsymbol{E}_\perp = \boldsymbol{U}\boldsymbol{U}' \tag{73}$$

In light of (85) of Corollary 2.3, Section 1.7, the matrix $\boldsymbol{U}$ coincides with the selection matrix $\boldsymbol{U}_1$ of (27) of the same section, whence the following holds

$$\boldsymbol{D}_\perp \boldsymbol{U}_\perp = \boldsymbol{C}, \quad \boldsymbol{E}_\perp \boldsymbol{U}_\perp = \boldsymbol{B} \tag{74}$$

by virtue of (28) of Section 1.7.

c) The matrix $\tilde{\Xi}$ is defined as follows

$$\tilde{\Xi} = \frac{1}{2}\ddot{\Xi}(1) - \dot{\Xi}(1)\, \Xi^g(1)\, \dot{\Xi}(1) = \boldsymbol{M}(1) - \boldsymbol{N}_1 \boldsymbol{N}_2^g \boldsymbol{N}_1 \tag{75}$$

and consequently

$$\boldsymbol{D}_\perp' \,\boldsymbol{U}_\perp' \,\tilde{\Xi}(1)\, \boldsymbol{E}_\perp \boldsymbol{U}_\perp = \boldsymbol{C}'(\boldsymbol{M}(1) - \boldsymbol{N}_1 \boldsymbol{N}_2^g \boldsymbol{N}_1)\, \boldsymbol{B} = \boldsymbol{I} \tag{76}$$

by virtue of (88) of Section 1.7.

Result ii) – The assertion simply follows from representation (59′), taking into account i) above.

Result iii) – The proof is straightforward, taking the expected value of both sides of (59′) in light of (60).

□

Corollary 1.2

VMA and VARMA representations, respectively, of the stationary process

$$\gamma_t = \begin{bmatrix} S'C_\perp^g \nabla \\ B^g \dot{A}(C'_\perp)^g C'_\perp \nabla - C' \\ C'\nabla \end{bmatrix} y_t \tag{77}$$

are given by

a) $$\gamma_t = \Pi_3(L)\,(\eta + \varepsilon_t) = P_3(1)\,\eta + \Pi_3(L)\,\varepsilon_t \tag{78}$$

b) $$(-P_4(L)\,\overline{\Lambda}\,\nabla + I)\,\gamma_t = P_3\,(1)\,\eta + P_3(L)\,\varepsilon_t \tag{79}$$

where, $\Pi_3(L)$, $P_3(L)$ and $P_4(L)$ stand for the homologous blocks of (42) and (43) with z replaced by L.

Further, we have

$$E\,\{\gamma_t\} = P_3(1)\,\eta \tag{80}$$

Proof

The proof is analogous to the proof of Corollary 1.2 of the previous section.

To begin with, observe how the claimed stationarity of the process γ_t as defined by (77) is nothing but a by-product of what has been pointed out in Theorem 1 about the integration and cointegration properties of the process y_t (see statements *i)* and *ii)* with $(C_\perp S_\perp)_\perp$ replaced by $[(C'_\perp)^g S,\ \ C]$).

Proof of a) – Since

$$G'\,[I\ \ 0] = [0\ \ I] \begin{bmatrix} \Psi(z) & F \\ G' & \Lambda(z) \end{bmatrix} - \Lambda\,(z)\,[0\ \ I] \tag{81}$$

the following result is easily established – thanks to the isomorphism between matrix polynomials in z and in L – in view of (63) and (43) above and by virtue of (22) of Section 1.6

$$\begin{aligned} G'A^{-1}(L)\nabla^2 &= G'\,[I \quad 0]\begin{bmatrix} \Psi(L) & F \\ G' & \Lambda(L) \end{bmatrix}^{-1}\begin{bmatrix} I \\ 0 \end{bmatrix} \\ &= -\Lambda(L)\,[0 \quad I]\begin{bmatrix} \Psi(L) & F \\ G' & \Lambda(L) \end{bmatrix}^{-1}\begin{bmatrix} I \\ 0 \end{bmatrix} \\ &= -\Lambda(L)\,\Pi_3(L) \Rightarrow V(L)\,G'A^{-1}(L) = \Pi_3(L) \end{aligned} \tag{82}$$

Hence, the VMA representation (78) is easily established by recalling (48) and observing that

$$\gamma_t = V(L)\,G'y_t \tag{83}$$

Proof of (b) – In light of (53), by replacing z with L and $(1-z)$ with ∇ as usual, the VMA representation (78) can be rewritten as

$$\gamma_t = \{I - P_4(L)\overline{\Lambda}\,\nabla\}^{-1}P_3(L)\,(\eta + \varepsilon_t) \tag{84}$$

whence, premultiplying both sides by the operator $I - P_4(L)\overline{\Lambda}\,\nabla$, the intended representation (79) is easily established.

The proof of (80) is trivial because of (78).

□

Also for the representation theorem established in this section there exists a dual version. It originates from a specification of a stationary second difference process $\nabla^2 y_t$ through a VMA model characterized so as to be the mirror image of a VAR model whose solution is the parent process y_t which enjoys particular integration and cointegration properties.

Theorem 2

Consider two stochastic processes ξ_t and y_t, the former being defined as the second difference of the latter, i.e.

$$\underset{(n,1)}{\xi_t} = \nabla^2 y_t \tag{85}$$

Let ξ_t be stationary and assume a VMA(∞) representation such as

$$\xi_t = \Xi(L)\,(\eta + \varepsilon_t) \tag{86}$$

whose parent matrix polynomial $\Xi(z) = \Xi_0 + \sum_{i=1}^{\infty} \Xi_i z^i$ in the complex argument z is characterized by a second order zero at $z = 1$ and by coefficient matrices Ξ_i with exponentially decreasing entries.

Then the companion process $\boldsymbol{y}_t$ admits a VAR generating model, namely

$$\boldsymbol{A}(L)\boldsymbol{y}_t = \boldsymbol{\eta} + \boldsymbol{\varepsilon}_t \tag{87}$$

whose parent matrix polynomial $\boldsymbol{A}(z)$ in the complex argument z has a second order zero at $z = 1$ and whose characteristic polynomial $det\,\boldsymbol{A}(z)$ has, besides a repeated unit-root, all other roots lying outside the unit circle.

The engendered process $\boldsymbol{y}_t$ enjoys the following major integration and cointegration properties

$$\left.\begin{array}{r} \boldsymbol{y}_t \sim I(2) \\ (\boldsymbol{C}_\perp \boldsymbol{S}_\perp)'_\perp \boldsymbol{y}_t \sim I(1) \end{array}\right\} \Rightarrow \boldsymbol{y}_t \sim CI(2, 1) \tag{88}$$

where $\boldsymbol{C}$ and $\boldsymbol{S}$ are defined through rank factorizations of $\boldsymbol{A}(1) = \boldsymbol{A}$ and $\boldsymbol{B}'_\perp \dot{\boldsymbol{A}} \boldsymbol{C}_\perp$ respectively, namely

$$\boldsymbol{A} = \boldsymbol{B}\boldsymbol{C}' \tag{89}$$

$$\boldsymbol{B}'_\perp \dot{\boldsymbol{A}} \boldsymbol{C}_\perp = \boldsymbol{R}\boldsymbol{S}' \tag{90}$$

Proof

In light of (7) and (9) of Section 1.3 together with the isomorphism between polynomials in a complex variable z and in the lag operator L, we obtain the paired expansions

$$\Xi(z) = \tilde{\Psi}(z)(1-z)^2 - \dot{\Xi}(1)(1-z) + \Xi(1) \Leftrightarrow \Xi(L) = \tilde{\Psi}(L)\nabla^2 - \dot{\Xi}(1)\nabla + \Xi(1) \tag{91}$$

where

$$\begin{aligned} \tilde{\Psi}(z) &= \sum_{k\geq 2}(-1)^k (1-z)^{k-2} \frac{1}{k!}\Xi^{(k)}(1) \Leftrightarrow \\ \Leftrightarrow \tilde{\Psi}(L) &= \sum_{k\geq 2}(-1)^k \frac{1}{k!}\Xi^{(k)}(1)\,\nabla^{k-2} \end{aligned} \tag{92}$$

with

$$\tilde{\Psi}(1) = \frac{1}{2}\ddot{\Xi}(1) \tag{93}$$

Then, in view of Definition 3 of Section 1.4 and by virtue of Theorem 3 and Corollary 3.1 of Section 1.6, the following hold true

$$det\, \Xi(1) = 0 \tag{94}$$

$$det\left(\begin{bmatrix} -\dot{\Xi}(1) & \tilde{\boldsymbol{B}} \\ \tilde{\boldsymbol{C}}' & \boldsymbol{0} \end{bmatrix}\right) = 0 \Leftrightarrow det\,(\tilde{\boldsymbol{B}}'_{\perp}\,\dot{\Xi}(1)\tilde{\boldsymbol{C}}_{\perp}) = 0 \tag{95}$$

$$det\left(\begin{bmatrix} \tilde{\Xi} & (\tilde{\boldsymbol{B}}_{\perp}\tilde{\boldsymbol{R}}_{\perp})_{\perp} \\ (\tilde{\boldsymbol{C}}_{\perp}\tilde{\boldsymbol{S}}_{\perp})'_{\perp} & \boldsymbol{0} \end{bmatrix}\right) \neq 0 \Leftrightarrow det\,(\tilde{\boldsymbol{R}}'_{\perp}\,\tilde{\boldsymbol{B}}'_{\perp}\,\tilde{\Xi}\,\tilde{\boldsymbol{C}}_{\perp}\,\tilde{\boldsymbol{S}}_{\perp}) \neq 0 \tag{96}$$

where $\tilde{\boldsymbol{B}}$ and $\tilde{\boldsymbol{C}}$, on the one hand, and $\tilde{\boldsymbol{R}}$ and $\tilde{\boldsymbol{S}}$ on the other, are defined through rank factorizations of $\Xi(1)$ and $\tilde{\boldsymbol{B}}'_{\perp}\,\dot{\Xi}(1)\tilde{\boldsymbol{C}}_{\perp}$, respectively, that is to say

$$\Xi(1) = \tilde{\boldsymbol{B}}\,\tilde{\boldsymbol{C}}' \tag{97}$$

$$\tilde{\boldsymbol{B}}'_{\perp}\,\dot{\Xi}(1)\tilde{\boldsymbol{C}}_{\perp} = \tilde{\boldsymbol{R}}\,\tilde{\boldsymbol{S}}' \tag{98}$$

and where

$$\tilde{\Xi} = \frac{1}{2}\ddot{\Xi}(1) - \dot{\Xi}(1)\,\Xi^{g}(1)\dot{\Xi}(1) \tag{99}$$

Given this premise, expansion (11) of Section 1.4 holds for $\Xi^{-1}(z)$ in a deleted neighbourhood of the second order pole $z = 1$ and hence we can write

$$\begin{aligned} \Xi^{-1}(z) &= \frac{1}{(1-z)^2}\tilde{\boldsymbol{N}}_2 + \frac{1}{(1-z)}\tilde{\boldsymbol{N}}_1 + \tilde{\boldsymbol{M}}(z) \Leftrightarrow \\ &\Leftrightarrow \Xi^{-1}(L) = \tilde{\boldsymbol{N}}_2\,\nabla^{-2} + \tilde{\boldsymbol{N}}_1\,\nabla^{-1} + \tilde{\boldsymbol{M}}(L) \end{aligned} \tag{100}$$

where

$$\tilde{\boldsymbol{N}}_2 = \tilde{\boldsymbol{C}}_{\perp}\,\tilde{\boldsymbol{S}}_{\perp}\,\{\tilde{\boldsymbol{R}}'_{\perp}\,\tilde{\boldsymbol{B}}'_{\perp}\,\tilde{\Xi}\,\tilde{\boldsymbol{C}}_{\perp}\,\tilde{\boldsymbol{S}}_{\perp}\}^{-1}\tilde{\boldsymbol{R}}'_{\perp}\,\tilde{\boldsymbol{B}}'_{\perp} \tag{101}$$

$$\tilde{\boldsymbol{N}}_1 = [\,\tilde{\boldsymbol{N}}_2 \quad \boldsymbol{I} - \tilde{\boldsymbol{N}}_2\,\tilde{\Xi}\,]\,\cdot \tag{102}$$

$$\cdot \begin{bmatrix} \tilde{\tilde{\Xi}} & \dot{\Xi}(1)\,\Xi^g(1) \\ \Xi^g(1)\,\dot{\Xi}(1) & -\tilde{C}_\perp(\tilde{B}'_\perp\dot{\Xi}(1)\,\tilde{C}_\perp)^g\,\tilde{B}'_\perp \end{bmatrix} \begin{bmatrix} \tilde{N}_2 \\ I - \tilde{\Xi}\tilde{N}_2 \end{bmatrix}$$

$$\tilde{\tilde{\Xi}} = \frac{1}{6}\,\dddot{\Xi}(1) - [\dot{\Xi}(1)\,\Xi^g(1)]^2\,\dot{\Xi}(1) \tag{103}$$

in view of Theorem 2 of Section 1.7.

Applying the operator $\Xi^{-1}(L)$ to both sides of (86) yields

$$\Xi^{-1}(L)\,\xi_t = \eta + \varepsilon_t \Rightarrow A\,(L)\,y_t = \eta + \varepsilon_t \tag{104}$$

that is the VAR representation (87), whose parent matrix polynomial

$$A\,(z) = \Xi^{-1}(z)\,(1-z)^2 = (1-z)^2\,\tilde{M}\,(z) + (1-z)\,\tilde{N}_1 + \tilde{N}_2 \tag{105}$$

is characterized by a second-order zero at $z = 1$.

Indeed, the following results hold true:

i)
$$A = \tilde{N}_2 \Rightarrow det\,A = 0 \tag{106}$$

whence the rank factorization (89), with B and C which can be conveniently chosen as

$$B = \tilde{C}_\perp\,\tilde{S}_\perp\,\{\,\tilde{R}'_\perp\,\tilde{B}'_\perp\,\tilde{\Xi}\,\tilde{C}_\perp\,\tilde{S}_\perp\,\}^{-1},\quad C = \tilde{B}_\perp\,\tilde{R}_\perp \tag{107}$$

ii) The matrices $B_\perp$ and $C_\perp$ can be accordingly chosen as

$$B_\perp = (\tilde{C}_\perp\,\tilde{S}_\perp)_\perp = [(\tilde{C}'_\perp)^g\tilde{S}\,,\;\;\tilde{C}\,],\;\; C_\perp = (\tilde{B}_\perp\,\tilde{R}_\perp)_\perp = [(\tilde{B}'_\perp)^g\tilde{R}\,,\;\;\tilde{B}\,] \tag{108}$$

iii)
$$\dot{A} = -\,\tilde{N}_1 \tag{109}$$

whence, in light of (108) and (109) and by virtue of (85) of Section 1.7, the following proves to be true

$$B'_\perp\,\dot{A}\,C_\perp = (\tilde{C}_\perp\tilde{S}_\perp)'_\perp\;\tilde{N}_1\,(\tilde{B}_\perp\,\tilde{R}_\perp)_\perp = U_1 U'_1 \tag{110}$$

where U_1 is as defined in (27) of Section 1.7.

iv) In light of (110) the matrix $B'_\perp\,\dot{A}\,C_\perp$ is singular, whence the rank factorization (90) occurs with $R = S = U_1$ and we can choose $R_\perp = S_\perp = U_2$ accordingly, where U_2 is as defined in (27) of Section 1.7.

v)
$$\ddot{A} = 2\,\tilde{M}\,(1) \tag{111}$$

$$\tilde{A} = \frac{1}{2}\ddot{A} - \dot{A}A^g\dot{A} = \tilde{M}(1) - \tilde{N}_1\tilde{N}_2^g\tilde{N}_1 \tag{112}$$

because of (105), (106) and (109).

vi) The matrix $R'_\perp B'_\perp \tilde{A} C_\perp S_\perp$ is non-singular. In fact the following holds

$$\begin{aligned} & R'_\perp B'_\perp \tilde{A} C_\perp S_\perp \\ &= U'_2(\tilde{C}_\perp \tilde{S}_\perp)_\perp(\tilde{M}(1) - \tilde{N}_1\tilde{N}_2^g\tilde{N}_1)(\tilde{B}_\perp \tilde{R}_\perp)_\perp U_2 \\ &= \tilde{C}'(\tilde{M}(1) - \tilde{N}_1\tilde{N}_2^g\tilde{N}_1)\tilde{B} = I \end{aligned} \tag{113}$$

in view of (28) and (88) of Section 1.7.

The proof of what has been claimed about the roots of the characteristic polynomial $det\, A(z)$ relies upon the arguments of Theorem 4 of Section 1.4.

Now, note that, insofar as $A(z)$ has a second order zero at $z = 1$, $A^{-1}(z)$ has a second order pole at $z = 1$. According to (105) the following Laurent expansions holds

$$\begin{aligned} A^{-1}(L) &= \sum_{j=1}^{2}\frac{1}{(I-L)^j}N_j + M(L) = \Xi(L)\nabla^{-2} \\ &= \tilde{\Psi}(L) - \dot{\Xi}(1)\nabla^{-1} + \Xi(1)\nabla^{-2} \end{aligned} \tag{114}$$

Then, from (87) and (114) it follows that

$$\begin{aligned} y_t &= A(L)^{-1}(\varepsilon_t + \eta) = (\tilde{\Psi}(L) - \dot{\Xi}(1)\nabla^{-1} + \Xi(1)\nabla^{-2})(\varepsilon_t + \eta) \\ &= \tilde{k}_0 + \tilde{k}_1 t + \tilde{B}\tilde{C}'\tilde{k}_2 t^2 - \dot{\Xi}(1)\sum_{\tau \le t}\varepsilon_\tau \\ &\quad + \tilde{B}\tilde{C}'\sum_{\tau \le t}(t+1-\tau)\varepsilon_\tau + \sum_{i=0}^{\infty}\tilde{\Psi}_i\varepsilon_{t-i} \end{aligned} \tag{115}$$

where $\tilde{k}_0$, $\tilde{k}_1$ and $\tilde{k}_2$ are vectors of constants.

Looking at (115) it is easy to realize that premultiplication by $\tilde{B}'_\perp$ leads to the elimination of both quadratic deterministic trends and cumulated random walks, which is tantamount to saying that

$$\tilde{B}'_\perp y_t \sim I(1) \Rightarrow y_t \sim CI(2, 1) \tag{116}$$

with $\tilde{B}_\perp = (C_\perp U_2)_\perp = (C_\perp S_\perp)_\perp$ playing the rôle of matrix of the cointegrating vectors.

□

3.6 A Unified Representation Theorem

By following the path already established in the foregoing sections we proceed now to establish a unified representation theorem which will shed extra light on VAR modelling offsprings, by bridging the virtual gap between representation theorems tailored on $I\,(1)$ and $I\,(2)$ processes, respectively.

Theorem

Let

$$\underset{(n,\,n)}{A(L)}\boldsymbol{y}_t = \boldsymbol{\eta} + \boldsymbol{\varepsilon}_t \tag{1}$$

be a VAR model whose characteristic polynomial has all roots lying outside the unit circle, except possibly for one or more unit-roots, and whose parametric specification is based upon the following rank factorizations and properties

$$\boldsymbol{A} = \boldsymbol{B}\boldsymbol{C}', \boldsymbol{A} \neq \boldsymbol{0} \tag{2}$$

$$\boldsymbol{B}_\perp' \,\dot{\boldsymbol{A}}\, \boldsymbol{C}_\perp = \boldsymbol{R}\boldsymbol{S}' \tag{3}$$

$$\boldsymbol{B}^g \dot{\boldsymbol{A}}\, \boldsymbol{C}_\perp \boldsymbol{S}_\perp = \boldsymbol{V}\boldsymbol{W}' \tag{4}$$

$$r\,\{\, \boldsymbol{R}_\perp' \,\boldsymbol{B}_\perp' (\frac{1}{2}\,\ddot{\boldsymbol{A}} - \,\dot{\boldsymbol{A}}\boldsymbol{A}^g \dot{\boldsymbol{A}}\,)\, \boldsymbol{C}_\perp \boldsymbol{S}_\perp\} = n - r\,(\boldsymbol{A}) - r\,(\boldsymbol{B}_\perp' \,\dot{\boldsymbol{A}}\, \boldsymbol{C}_\perp) \tag{5}$$

Then the solution $\boldsymbol{y}_t$ of model (1) enjoys the integration property

$$\boldsymbol{y}_t \sim I\,(d), \quad 0 \leq d \leq 2 \tag{6}$$

where

a) $$d = 0 \quad \textit{if det}\, \boldsymbol{A} \neq 0 \tag{7}$$

b) $$d = 1 \quad \textit{if det}\, \boldsymbol{A} = 0 \quad \textit{but det}\,(\boldsymbol{B}_\perp' \,\dot{\boldsymbol{A}}\, \boldsymbol{C}_\perp) \neq 0 \tag{8}$$

c) $$d = 2 \quad \textit{if det}\,(\boldsymbol{B}_\perp' \,\dot{\boldsymbol{A}}\, \boldsymbol{C}_\perp) = 0 \tag{9}$$

The process $\boldsymbol{y}_t$ enjoys the cointegration properties

i) $$\boldsymbol{C}'\boldsymbol{y}_t \sim I\,(d-1) \textit{ under } (\text{b}) \textit{ and } (\text{c}) \tag{10}$$

ii) $$S' C_{\perp}^{g} y_t \sim I(1) \; under \; (c) \tag{11}$$

iii) $$(B^{g} \dot{A} \; (C_{\perp}')^{g} C_{\perp}' \nabla - C') y_t \sim I(0) \; under \; (c) \tag{12}$$

iv) $$V_{\perp}' C' y_t \sim I(0) \; under \; (c) \tag{13}$$

Finally, y_t engendered by the VAR model (1) has the representation

$$y_t = k_0 + k_1 t + k_2 t^2 + N_1 \sum_{\tau \le t} \varepsilon_\tau + N_2 \sum_{\tau \le t} (t+1-\tau)\varepsilon_\tau + \sum_{i=0}^{\infty} M_i \, \varepsilon_{t-i} \tag{14}$$

with

$$N_2 = \begin{cases} C_{\perp} S_{\perp} (R_{\perp}' B_{\perp}' \tilde{A} C_{\perp} S_{\perp})^{-1} R_{\perp}' B_{\perp}' & under \, (c) \quad (15) \\ 0 & otherwise \quad (15') \end{cases}$$

$$N_1 = \begin{cases} -C_{\perp} (B_{\perp}' \dot{A} C_{\perp})^{g} B_{\perp}' + K(N_2) & under \, (c) \quad (16) \\ -C_{\perp} (B_{\perp}' \dot{A} C_{\perp})^{-1} B_{\perp}' & under \, (b) \quad (16') \\ 0 & under \, (a) \quad (16'') \end{cases}$$

and

$$\begin{aligned} K(N_2) = {} & N_2 \tilde{\tilde{A}} N_2 + (I - N_2 \tilde{A}) A^{g} \dot{A} N_2 + N_2 \dot{A} A^{g} (I - \tilde{A} N_2) \\ & + N_2 \tilde{A} C_{\perp} (B_{\perp}' \dot{A} C_{\perp})^{g} B_{\perp}' + C_{\perp} (B_{\perp}' \dot{A} C_{\perp})^{g} B_{\perp}' \tilde{A} N_2 \\ & - N_2 \tilde{A} C_{\perp} (B_{\perp}' \dot{A} C_{\perp})^{-1} B_{\perp}' \tilde{A} N_2 \end{aligned} \tag{17}$$

with $\tilde{A}$ and $\tilde{\tilde{A}}$ as defined in (24) and (25) of Section 1.7.

Proof

The proof of a) rests on Theorem 1 of Section 1.6, from the algebraic standpoint, on a by-product of Theorem 1 of Section 2.3 and eventually on Proposition 1 of Section 3.2, from the econometric standpoint.

The proof of b) rests on Theorem 2, along with the Corollary 2.1, of Section 1.6, from the algebraic standpoint, and on Theorem 1 of Section 2.3 as well as Proposition 3 of Section 3.2, from the econometric standpoint.

The proof of c), bearing the rank hypothesis (5) in mind, rests on Theorem 3, along with Corollary 3.1, of Section 1.6, from the algebraic standpoint, and on Theorem 1 of Section 2.3 and eventually Proposition 5 of Section 3.2, from the econometric standpoint.

As far as the proofs of statements i)-iv) are concerned, we will proceed in this way.

To prove (10), first refer back, to one hand, to Result ii) in Corollary 1.1 and to Result ii) in Corollary 2.1 of Section 1.7 for the algebraic rationale; and then refer to Propositions 4 and 6 of Section 3.2 for the econometric reading key we are primarily interested in.

To prove (11), reference must be made again to Result ii) in Corollary 2.1 of Section 1.7, on the one hand, and to Proposition 6 of Section 3.2 on the other.

Statements (12) and (13) coincide with statements (21) and (22) in Theorem 1 of Section 3.5 and are proved accordingly.

Moving to representation (14) its proof hinges on Theorems 2 and 3 of Section 1.5 and follows the lines traced in establishing (10) in Theorem 1 of the previous section.

For what concerns the two-fold form of N_2, notice how the closed-form (15) mirrors (22) of Section 1.7 as well as (11′) of Section 3.5 and its validity relies upon the reasoning put forward therein under the condition (5). Here N_2 becomes a null matrix, according to (15′), under the circumstances b) and a) insofar as either the pair $\boldsymbol{R}_\perp, \boldsymbol{S}_\perp$ or both the pairs $\boldsymbol{B}_\perp, \boldsymbol{C}_\perp$ and $\boldsymbol{R}_\perp, \boldsymbol{S}_\perp$ respectively, turn out to be made of empty matrices (see the Appendix at the end of the chapter in this connection). References to Theorems 1 and 2 of Section 1.6 as well as to their econometric mirror images represented by Propositions 1 and 3 of Section 3.2, can likewise be appropriate in this connection.

Finally, for the three-fold form of N_1 observe that

- The closed-form (16) is nothing but an algebraic rearrangement of (23) in Section 1.7, as well as of (12′) in Section 3.5, and its validity rests on the arguments advanced therein under the condition (5).
- The closed-form (16′) mirrors (3) in Section 1.7, as well as (7) in Section 3.4, its validity being supported by the arguments developed therein under the assumption (3). Indeed, the expression (16′) reads as a special form of (16) occurring insofar as the pair $\boldsymbol{R}_\perp, \boldsymbol{S}_\perp$ is made up of empty matrices, N_2 becomes a null matrix accordingly, and $\boldsymbol{B}'_\perp \dot{\boldsymbol{A}} \boldsymbol{C}_\perp$ is nonsingular, whence the identity

$$(\boldsymbol{B}'_\perp \dot{\boldsymbol{A}} \boldsymbol{C}_\perp)^g = (\boldsymbol{B}'_\perp \dot{\boldsymbol{A}} \boldsymbol{C}_\perp)^{-1} \tag{18}$$

- The matrix N_1 degenerates to a null matrix under the circumstance a) insofar as both the pairs $\boldsymbol{R}_\perp, \boldsymbol{S}_\perp$ and $\boldsymbol{B}_\perp, \boldsymbol{C}_\perp$ turn out to be made up of empty matrices.

References to Theorems 1 and 2 of Section 1.6, as well to their econometric counterparts given by Propositions 1 and 3 of Section 3.2, are likewise appropriate in this connection.

□

Appendix. Empty Matrices

By an empty matrix we mean a matrix whose number of rows or of columns is equal to zero (see, e.g., Chipman and Rao, 1964).

Some formal rules proves useful when operating with empty matrices.

Let $\boldsymbol{B}$ and $\boldsymbol{C}$ be empty matrices of order $n \times 0$ and $0 \times p$, respectively, and $\boldsymbol{A}$ be a matrix of order $m \times n$.

We then assume the following formal rules of calculus

i) $$\boldsymbol{AB} = \underset{(m,\,0)}{\boldsymbol{D}}\text{, namely an empty matrix} \tag{A1}$$

ii) $$\boldsymbol{BC} = \underset{(n,\,p)}{\boldsymbol{0}}\text{, namely a null matrix} \tag{A2}$$

iii) $$\boldsymbol{B'B} = (\boldsymbol{B'B})^g = \underset{(0,\,0)}{\boldsymbol{0}}\text{, namely the empty null matrix} \tag{A3}$$

iv) $$\boldsymbol{B}^g = \boldsymbol{B'} \tag{A4}$$

v) $$r(\boldsymbol{B}) = 0 \tag{A5}$$

The notion of empty matrix paves the way to some noteworthy extensions of the algebra of orthogonal complement.

Let $\boldsymbol{B}$ be an empty matrix of order $n \times 0$ and $\boldsymbol{A}$ be a non-singular matrix of order n. Then, we will agree upon the following formal relationships

a) $$\boldsymbol{B}_{\perp} = \boldsymbol{A}\text{, namely an arbitrary non-singular matrix} \tag{A6}$$

b) $$\boldsymbol{A}_{\perp} = \boldsymbol{B}\text{, namely an empty matrix.} \tag{A7}$$

References

Baltagi BH (ed) (2003) A companion to theoretical econometrics. Blackwell, Malden, Ma

Banerjee A, Dolado J, Galbraith JW, Hendry DF (1993) Co-integration, error-correction and the econometric analysis of non stationary data. Oxford University Press, New York

Barnett S (1990) Matrices: methods and applications. Clarendon Press, Oxford

Bellman R (1970) Introduction to matrix analysis. 2nd edn, McGraw Hill, New York

Blanc-Lapierre A, Fortet R (1953) Théorie des fonctions aléatoires. Masson, Paris

Charenza WW, Deadman DF (1992) New directions in econometric practise. Edgar Publ Lim, Aldershot, UK

Chipman JS, Rao MM (1964) The treatment of linear restrictions in regression analysis. Econometrica 32: 198-209

Cogan EJ, Norman RZ (1963) Handbook of calculus, difference and differential equations. 2nd edn, Prentice-Hall, Englewood Cliffs, NJ

Cramer H, Leadbetter MR (1967) Stationary and related stochastic processes. Wiley, New York

Davidson R, MacKinnon JG (1993) Estimation and inference in econometrics. Oxford University Press, New York

Dhrymes PJ (1971) Distributed lags: problems of estimation and formulation. Holden Day, San Francisco

Elaydi SN (1996) An introduction to difference equations. Springer, New York

Engle RF, Granger CWJ (1987) Co-integration and error correction: representation, estimation and testing. Econometrica 55: 251-276

Engle RF, Granger CWJ (eds) (1991) Long-run economic relationships: readings in cointegration. Oxford University Press, New York

Engsted T, Johansen S (1999) Granger's representation theorem and multicointegration. In: Engle RF, White H (eds) Cointegration, causality and forecasting. Oxford University Press, New York, pp 200-211

Faliva M (1975) Alcune regole di derivazione in termini matriciali. Statistica 35: 329-340

Faliva M (1987) Econometria: principi e metodi. UTET, Torino

Faliva M, Zoia MG (1999) Econometria delle serie storiche. Giappichelli, Torino

Faliva M, Zoia MG (2002 a) On a partitioned inversion formula having useful applications in econometrics. Econometric Theory 18: 525-530

Faliva M, Zoia MG (2002 b) Matrix polynomials and their inversion: the algebraic framework of unit-root econometrics representation theorems. Statistica 62: 187-202

Faliva M, Zoia MG (2003) A new proof of the representation theorem for I(2) processes. J of Interdiscipl Math 6: 331-347

Goldberger AS (1964) Econometric theory. Wiley, New York

Greville TNE (1960) Some applications of the pseudoinverse of a matrix. Siam Rev 2: 15-22

Haldrup N, Salmon M (1998) Representations of I(2) cointegrated systems using the Smith-McMillan form. J Econometrics 84: 303-325

Hamilton JD (1994) Time series analysis. Princeton University Press, Princeton, NJ

Hansen PR, Johansen S (1998) Workbook on cointegration. Oxford University Press, New York

Hargreaves CP (ed) (1994) Non stationary time series analysis and cointegration. Oxford University Press, New York

Harvey AC (1993) Time series models. 2nd edn, Harvester Wheatsheaf, New York

Hatanaka M (1996) Time-series-based econometrics. Oxford University Press, New York

Jeffrey A (1992) Complex analysis and applications. CRC Press, London

Johansen S (1991) Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models. Econometrica 59: 1551-80

Johansen S (1992) A representation of vector autoregressive processes integrated of order 2. Econometric Theory 8: 188-202

Johansen S (1995) Likelihood-based inference in cointegrated vector autoregressive models. Oxford University Press, New York

Klein LR (1983) Lectures in econometrics. North-Holland, Amsterdam

La Cour L (1998) A parametric characterization of integrated vector autoregressive (VAR) processes. Econometric Theory 14: 187-199

Lancaster P, Tismenetsky M (1985) The theory of matrices. 2nd edn, Academic Press, New York

Lütkepohl H (1991) Introduction to multiple time series analysis. Springer, Berlin

Lütkepohl H (1996) Handbook of matrices. Wiley, New York

Lütkepohl H, Krätzig M (2004) Applied time series econometrics. Cambridge University Press, New York

Magnus JR, Neudecker H (1999) Matrix differential calculus with applications to statistics and econometrics. Wiley, New York

Markuscevich AI (1965) Theory of a complex variable. Prentice Hall, London

Marsaglia G, Styan GPH (1974) Equalities and inequalities for ranks of matrices. Linear and Multilinear Algebra 2: 269-292

Mickens RE (1990) Difference equations. Theory and applications. 2nd edn, Van Nostrand Reinhold, New York

Miller KS (1987) Some eclectic matrix theory. Krieger Publ Co, Malabar, Florida

Papoulis A (1965) Probability, random variables and stochastic processes. McGraw Hill, New York

Pringle RM, Rayner AA (1971) Generalized inverse matrices with applications to statistics. Griffin, London

Rao CR (1973) Linear statistical inference and its applications. 2nd edn, Wiley, New York

Rao CR, Mitra SK (1971) Generalized inverse of matrices and its applications. Wiley, New York

Searle SR (1982) Matrix algebra useful for statistics. Wiley, New York

Stock JH, Watson MW (1988) Testing for common trends. J Amer Stat Ass 83: 1097-1107

Theil H (1971) Principles of econometrics. North Holland, Amsterdam

Thrall RM, Tornheim L (1957) Vector spaces and matrices. Wiley, New York

Vladimirov VS (1984) Equations of mathematical physics. MIR Publ, Moscow

Yaglom AM (1962) Stationary random functions. Prentice Hall, Englewood Cliffs, NJ

Zoia MG (2003) Rank-based diagnostics for the integration order of a VAR solution. Statistica & Applicazioni 1: 56-67

Notational Conventions, Symbols and Acronyms

The following notational conventions will be used throughout the text:

- Bold lower case letters (Roman or Greek) indicate vectors.
- Bold upper case letters (Roman or Greek) indicate matrices.
- Both notations $[\boldsymbol{A} \quad \boldsymbol{B}]$ and $[\boldsymbol{A} \ , \boldsymbol{B}]$ will be used, depending on convenience, for column-wise partitioned matrices
- Both notations $\begin{bmatrix} \boldsymbol{A} & \boldsymbol{B} \\ & \\ \boldsymbol{C} & \boldsymbol{D} \end{bmatrix}$ and $\begin{bmatrix} \boldsymbol{A} & \vdots & \boldsymbol{B} \\ \cdots & & \cdots \\ \boldsymbol{C} & \vdots & \boldsymbol{D} \end{bmatrix}$ will be used, depending on convenience, for block matrices

Symbols and Acronyms

	Meaning	*Section*
$\boldsymbol{A}^-$	generalized inverse of $\boldsymbol{A}$	1.1 (Definition 1)
$\boldsymbol{A}'$	transpose of $\boldsymbol{A}$	1.1
$\boldsymbol{A}^g$	Moore-Penrose inverse of $\boldsymbol{A}$	1.1 (Definition 2)
$\boldsymbol{A}_r^-$	right inverse of $\boldsymbol{A}$	1.1 (Definition 3)
$\boldsymbol{A}_l^-$	left inverse of $\boldsymbol{A}$	1.1 (Definition 4)
$det\,(\boldsymbol{A})$	determinant of $\boldsymbol{A}$	1.1
$r\,(\boldsymbol{A})$	rank of $\boldsymbol{A}$	1.1
$\boldsymbol{A}_\perp$	orthogonal complement of $\boldsymbol{A}$	1.1 (Definition 6)
$\boldsymbol{A}_l^\perp$	left orthogonal complement of $\boldsymbol{A}$	1.1 (Definition 7)
$\boldsymbol{A}_r^\perp$	right orthogonal complement of $\boldsymbol{A}$	1.1 (ibid.)
$\boldsymbol{A}\,(z)$	matrix polynomial in the scalar argument z	1.3 (Definition 1)
$\dot{\boldsymbol{A}}(z), \ddot{\boldsymbol{A}}(z), \dddot{\boldsymbol{A}}(z)$	dot notation for derivatives	1.3
$\boldsymbol{A}, \dot{\boldsymbol{A}}, \ddot{\boldsymbol{A}}, \dddot{\boldsymbol{A}}$	short notation for $\boldsymbol{A}(1), \dot{\boldsymbol{A}}(1), \ddot{\boldsymbol{A}}(1), \dddot{\boldsymbol{A}}(1)$	1.3
$tr\,\boldsymbol{A}$	trace of $\boldsymbol{A}$	1.3
$\boldsymbol{A}^+$	adjoint of $\boldsymbol{A}$	1.3
$vec\,\boldsymbol{A}$	staked form of $\boldsymbol{A}$	1.3
L	lag operator	1.5

∇	backward difference operator	1.5 (Definition 1)
∇^{-1}	antidifference operator	1.5 (Definition 2)
Σ	indefinite sum operator	1.5 (ibid.)
E	expectation operator	2.1
$\Gamma(h)$	autocovariance matrix of order h	2.1
$I(d)$	integrated process of order d (d positive integer)	2.1 (Definition 5)
$I(0)$	stationary process	2.1
$WN_{(n)}$	n-dimensional white noise	2.2 (Definition 1)
$\varepsilon_t \sim WN$	ε_t is a white noise	
δ_v	discrete unitary function	2.2
$\otimes$	Kronecker matrix product	2.2
VMA (q)	vector moving average process of order q	2.2 (Definition 2)
VAR (p)	vector autoregressive process of order p	2.2 (Definition 5)
VARMA (p, q)	vector autoregressive moving average process of order (p, q)	2.2 (Definition 7)
$CI(d, b)$	cointegrated system of order (d, b)	2.4 (Definition 6)
$PCI(d, b)$	polynomially cointegrated system of order (d, b)	2.4 (Definition 7)

List of Definitions

Section 1.1

Section 1.3

Section 1.4

Section 1.5

Section 2.1

Section 2.2

Section 2.4

Section 3.2

List of Theorems, Corollaries and Propositions

Section 1.1

Section 1.2

Section 1.3

Section 1.4

Section 1.5

Section 1.6

Section 1.7

Section 2.3

Section 3.2

Section 3.3

Section 3.4

Section 3.5

Section 3.6

Lecture Notes in Economics and Mathematical Systems

For information about Vols. 1–470
please contact your bookseller or Springer-Verlag

Vol. 471: N. H. M. Wilson (Ed.), Computer-Aided Transit Scheduling. XI, 444 pages. 1999.

Vol. 472: J.-R. Tyran, Money Illusion and Strategic Complementarity as Causes of Monetary Non-Neutrality. X, 228 pages. 1999.

Vol. 473: S. Helber, Performance Analysis of Flow Lines with Non-Linear Flow of Material. IX, 280 pages. 1999.

Vol. 474: U. Schwalbe, The Core of Economies with Asymmetric Information. IX, 141 pages. 1999.

Vol. 475: L. Kaas, Dynamic Macroeconomics with Imperfect Competition. XI, 155 pages. 1999.

Vol. 476: R. Demel, Fiscal Policy, Public Debt and the Term Structure of Interest Rates. X, 279 pages. 1999.

Vol. 477: M. Théra, R. Tichatschke (Eds.), Ill-posed Variational Problems and Regularization Techniques. VIII, 274 pages. 1999.

Vol. 478: S. Hartmann, Project Scheduling under Limited Resources. XII, 221 pages. 1999.

Vol. 479: L. v. Thadden, Money, Inflation, and Capital Formation. IX, 192 pages. 1999.

Vol. 480: M. Grazia Speranza, P. Stähly (Eds.), New Trends in Distribution Logistics. X, 336 pages. 1999.

Vol. 481: V. H. Nguyen, J. J. Strodiot, P. Tossings (Eds.). Optimation. IX, 498 pages. 2000.

Vol. 482: W. B. Zhang, A Theory of International Trade. XI, 192 pages. 2000.

Vol. 483: M. Königstein, Equity, Efficiency and Evolutionary Stability in Bargaining Games with Joint Production. XII, 197 pages. 2000.

Vol. 484: D. D. Gatti, M. Gallegati, A. Kirman, Interaction and Market Structure. VI, 298 pages. 2000.

Vol. 485: A. Garnaev, Search Games and Other Applications of Game Theory. VIII, 145 pages. 2000.

Vol. 486: M. Neugart, Nonlinear Labor Market Dynamics. X, 175 pages. 2000.

Vol. 487: Y. Y. Haimes, R. E. Steuer (Eds.), Research and Practice in Multiple Criteria Decision Making. XVII, 553 pages. 2000.

Vol. 488: B. Schmolck, Ommitted Variable Tests and Dynamic Specification. X, 144 pages. 2000.

Vol. 489: T. Steger, Transitional Dynamics and Economic Growth in Developing Countries. VIII, 151 pages. 2000.

Vol. 490: S. Minner, Strategic Safety Stocks in Supply Chains. XI, 214 pages. 2000.

Vol. 491: M. Ehrgott, Multicriteria Optimization. VIII, 242 pages. 2000.

Vol. 492: T. Phan Huy, Constraint Propagation in Flexible Manufacturing. IX, 258 pages. 2000.

Vol. 493: J. Zhu, Modular Pricing of Options. X, 170 pages. 2000.

Vol. 494: D. Franzen, Design of Master Agreements for OTC Derivatives. VIII, 175 pages. 2001.

Vol. 495: I. Konnov, Combined Relaxation Methods for Variational Inequalities. XI, 181 pages. 2001.

Vol. 496: P. Weiß, Unemployment in Open Economies. XII, 226 pages. 2001.

Vol. 497: J. Inkmann, Conditional Moment Estimation of Nonlinear Equation Systems. VIII, 214 pages. 2001.

Vol. 498: M. Reutter, A Macroeconomic Model of West German Unemployment. X, 125 pages. 2001.

Vol. 499: A. Casajus, Focal Points in Framed Games. XI, 131 pages. 2001.

Vol. 500: F. Nardini, Technical Progress and Economic Growth. XVII, 191 pages. 2001.

Vol. 501: M. Fleischmann, Quantitative Models for Reverse Logistics. XI, 181 pages. 2001.

Vol. 502: N. Hadjisavvas, J. E. Martínez-Legaz, J.-P. Penot (Eds.), Generalized Convexity and Generalized Monotonicity. IX, 410 pages. 2001.

Vol. 503: A. Kirman, J.-B. Zimmermann (Eds.), Economics with Heterogenous Interacting Agents. VII, 343 pages. 2001.

Vol. 504: P.-Y. Moix (Ed.),The Measurement of Market Risk. XI, 272 pages. 2001.

Vol. 505: S. Voß, J. R. Daduna (Eds.), Computer-Aided Scheduling of Public Transport. XI, 466 pages. 2001.

Vol. 506: B. P. Kellerhals, Financial Pricing Models in Con-tinuous Time and Kalman Filtering. XIV, 247 pages. 2001.

Vol. 507: M. Koksalan, S. Zionts, Multiple Criteria Decision Making in the New Millenium. XII, 481 pages. 2001.

Vol. 508: K. Neumann, C. Schwindt, J. Zimmermann, Project Scheduling with Time Windows and Scarce Resources. XI, 335 pages. 2002.

Vol. 509: D. Hornung, Investment, R&D, and Long-Run Growth. XVI, 194 pages. 2002.

Vol. 510: A. S. Tangian, Constructing and Applying Objective Functions. XII, 582 pages. 2002.

Vol. 511: M. Külpmann, Stock Market Overreaction and Fundamental Valuation. IX, 198 pages. 2002.

Vol. 512: W.-B. Zhang, An Economic Theory of Cities.XI, 220 pages. 2002.

Vol. 513: K. Marti, Stochastic Optimization Techniques. VIII, 364 pages. 2002.

Vol. 514: S. Wang, Y. Xia, Portfolio and Asset Pricing. XII, 200 pages. 2002.

Vol. 515: G. Heisig, Planning Stability in Material Requirements Planning System. XII, 264 pages. 2002.

Vol. 516: B. Schmid, Pricing Credit Linked Financial Instruments. X, 246 pages. 2002.

Vol. 517: H. I. Meinhardt, Cooperative Decision Making in Common Pool Situations. VIII, 205 pages. 2002.

Vol. 518: S. Napel, Bilateral Bargaining. VIII, 188 pages. 2002.

Vol. 519: A. Klose, G. Speranza, L. N. Van Wassenhove (Eds.), Quantitative Approaches to Distribution Logistics and Supply Chain Management. XIII, 421 pages. 2002.

Vol. 520: B. Glaser, Efficiency versus Sustainability in Dynamic Decision Making. IX, 252 pages. 2002.

Vol. 521: R. Cowan, N. Jonard (Eds.), Heterogenous Agents, Interactions and Economic Performance. XIV, 339 pages. 2003.

Vol. 522: C. Neff, Corporate Finance, Innovation, and Strategic Competition. IX, 218 pages. 2003.

Vol. 523: W.-B. Zhang, A Theory of Interregional Dynamics. XI, 231 pages. 2003.

Vol. 524: M. Frölich, Programme Evaluation and Treatment Choise. VIII, 191 pages. 2003.

Vol. 525: S. Spinler, Capacity Reservation for Capital-Intensive Technologies. XVI, 139 pages. 2003.

Vol. 526: C. F. Daganzo, A Theory of Supply Chains. VIII, 123 pages. 2003.

Vol. 527: C. E. Metz, Information Dissemination in Currency Crises. XI, 231 pages. 2003.

Vol. 528: R. Stolletz, Performance Analysis and Optimization of Inbound Call Centers. X, 219 pages. 2003.

Vol. 529: W. Krabs, S. W. Pickl, Analysis, Controllability and Optimization of Time-Discrete Systems and Dynamical Games. XII, 187 pages. 2003.

Vol. 530: R. Wapler, Unemployment, Market Structure and Growth. XXVII, 207 pages. 2003.

Vol. 531: M. Gallegati, A. Kirman, M. Marsili (Eds.), The Complex Dynamics of Economic Interaction. XV, 402 pages, 2004.

Vol. 532: K. Marti, Y. Ermoliev, G. Pflug (Eds.), Dynamic Stochastic Optimization. VIII, 336 pages. 2004.

Vol. 533: G. Dudek, Collaborative Planning in Supply Chains. X, 234 pages. 2004.

Vol. 534: M. Runkel, Environmental and Resource Policy for Consumer Durables. X, 197 pages. 2004.

Vol. 535: X. Gandibleux, M. Sevaux, K. Sörensen, V.T'kindt (Eds.), Metaheuristics for Multiobjective Optimisation. IX, 249 pages. 2004.

Vol. 536: R. Brüggemann, Model Reduction Methods for Vector Autoregressive Processes. X, 218 pages. 2004.

Vol. 537: A. Esser, Pricing in (In)Complete Markets. XI, 122 pages, 2004.

Vol. 538: S. Kokot, The Econometrics of Sequential Trade Models. XI, 193 pages. 2004.

Vol. 539: N. Hautsch, Modelling Irregularly Spaced Financial Data. XII, 291 pages. 2004.

Vol. 540: H. Kraft, Optimal Portfolios with Stochastic Interest Rates and Defaultable Assets. X, 173 pages. 2004.

Vol. 541: G.-y. Chen, X. Huang, X. Yang, Vector Optimization. X, 306 pages. 2005.

Vol. 542: J. Lingens, Union Wage Bargaining and Economic Growth. XIII, 199 pages. 2004.

Vol. 543: C. Benkert, Default Risk in Bond and Credit Derivatives Markets. IX, 135 pages. 2004.

Vol. 544: B. Fleischmann, A. Klose, Distribution Logistics. X, 284 pages. 2004.

Vol. 545: R. Hafner, Stochastic Implied Volatility. XI, 229 pages. 2004.

Vol. 546: D. Quadt, Lot-Sizing and Scheduling for Flexible Flow Lines. XVIII, 227 pages. 2004.

Vol. 547: M. Wildi, Signal Extraction. XI, 279 pages. 2005.

Vol. 548: D. Kuhn, Generalized Bounds for Convex Multistage Stochastic Programs. XI, 190 pages. 2005.

Vol. 549: G. N. Krieg, Kanban-Controlled Manufacturing Systems. IX, 236 pages. 2005.

Vol. 550: T. Lux, S. Reitz, E. Samanidou, Nonlinear Dynamics and Heterogeneous Interacting Agents. XIII, 327 pages. 2005.

Vol. 551: J. Leskow, M. Puchet Anyul, L. F. Punzo, New Tools of Economic Dynamics. XIX, 392 pages. 2005.

Vol. 552: C. Suerie, Time Continuity in Discrete Time Models. XVIII, 229 pages. 2005.

Vol. 553: B. Mönch, Strategic Trading in Illiquid Markets. XIII, 116 pages. 2005.

Vol. 554: R. Foellmi, Consumption Structure and Macroeconomics. IX, 152 pages. 2005.

Vol. 555: J. Wenzelburger, Learning in Economic Systems with Expectations Feedback (planned) 2005.

Vol. 556: R. Branzei, D. Dimitrov, S. Tijs, Models in Cooperative Game Theory. VIII, 135 pages. 2005.

Vol. 557: S. Barbaro, Equity and Efficiency Considerations of Public Higer Education. XII, 128 pages. 2005.

Vol. 558: M. Faliva, M. G. Zoia, Topics in Dynamic Model Analysis. X, 144 pages. 2005.

Vol. 559: M. Schulmerich, Real Options Valuation. XVI, 357 pages. 2005.

Vol. 560: A. von Schemde, Index and Stability in Bimatrix Games. X, 151 pages. 2005.

Vol. 561: H. Bobzin, Principles of Network Economics. XX, 390 pages. 2006. (planned)

Vol. 562: T. Langenberg, Standardization and Expectations. IX, 122 pages. 2006. (planned)

Vol. 563: A. Seeger, Recent Advances in Optimization. XV, 455 pages. 2006. (planned)

Vol. 564: P. Mathieu, B. Beaufils, O. Brandouy, (Eds.) Artificial Economics. XIII, 237 pages. 2005.

Vol. 565: W. Lemke, Term Structure Modeling and Estimation in a State Space Framework. IX, 233 pages. 2006. (planned)

Vol. 566: M. Genser, A Structural Framework for the Pricing of Corporate Securities. XIX, 185 pages. 2006.