Introduction
to
MATRIX
ALGEBRA©

KAW

Introduction
to
MATRIX
ALGEBRA

Autar K. Kaw

University of South Florida

Autar K. Kaw
Professor & Jerome Krivanek Distinguished Teacher
Mechanical Engineering Department
University of South Florida, ENB 118
4202 E. Fowler Avenue
Tampa, FL 33620-5350.
Office: (813) 974-5626
Fax: (813) 974-3539
E-mail: kaw@eng.usf.edu
URL: http://www.eng.usf.edu/~kaw

Table of Contents

Chapter 1
Introduction

After reading this chapter, you should be able to

- *Know what a matrix is*
- *Identify special types of matrices*
- *When two matrices are equal*

What is a matrix?

Matrices are everywhere. If you have used a spreadsheet such as Excel or Lotus or written a table, you have used a matrix. Matrices make presentation of numbers clearer and make calculations easier to program. Look at the matrix below about the sale of tires in a Blowoutr'us store – given by quarter and make of tires.

$$\begin{array}{l} \\ \textbf{Tirestone} \\ \textbf{Michigan} \\ \textbf{Copper} \end{array} \begin{array}{c} \begin{array}{cccc} \textbf{Quarter 1} & \textbf{Quarter 2} & \textbf{Quarter 3} & \textbf{Quarter 4} \end{array} \\ \begin{bmatrix} 25 & 20 & 3 & 2 \\ 5 & 10 & 15 & 25 \\ 6 & 16 & 7 & 27 \end{bmatrix} \end{array}$$

If one wants to know how many Copper tires were sold in Quarter 4, we go along the row 'Copper' and column 'Quarter 4' and find that it is 27.

So what is a matrix?

A matrix is a rectangular array of elements. The elements can be symbolic expressions or numbers. Matrix [A] is denoted by

$$[A] = \begin{bmatrix} a_{11} & a_{12} & \ldots\ldots. & a_{1n} \\ a_{21} & a_{22} & \ldots\ldots. & a_{2n} \\ \vdots & & & \vdots \\ a_{m1} & a_{m2} & \ldots\ldots. & a_{mn} \end{bmatrix}$$

Row i of [A] has n elements and is $\begin{bmatrix} a_{i1} & a_{i2} \ldots . a_{in} \end{bmatrix}$ and

Column j of [A] has m elements and is $\begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix}$

Each matrix has rows and columns and this defines the size of the matrix. If a matrix [A] has m rows and n columns, the **size of the matrix** is denoted by m x n. The matrix [A] may also be denoted by $[A]_{mxn}$ to show that $[A]$ is a matrix with m rows and n columns.

Each entry in the matrix is called the **entry** or **element** of the matrix and is denoted by a_{ij} where i is the row number and j is the column number of the element.

The matrix for the tire sales example could be denoted by the matrix [A] as

$$[\mathrm{A}] = \begin{bmatrix} 25 & 20 & 3 & 2 \\ 5 & 10 & 15 & 25 \\ 6 & 16 & 7 & 27 \end{bmatrix}$$

There are 3 rows and 4 columns, so the size of the matrix is 3 x 4. In the above $[A]$ matrix, $\mathrm{a}_{34} = 27$.

What are the special types of matrices?

Vector: A vector is a matrix that has only one row or one column. There are two types of vectors – row vectors and column vectors.

Row vector: If a matrix has one row, it is called a row vector

$$[\mathrm{B}] = [\mathrm{b}_1 \ \mathrm{b}_2 \ldots\ldots \mathrm{b}_\mathrm{m}]$$

and ‘m’ is the dimension of the row vector.

Example

Give an example of a row vector.

Solution

[B] = [25   20   3   2   0] is an example of a row vector of dimension 5.

Column vector: If a matrix has one column, it is called a column vector

$$[C] = \begin{bmatrix} c_1 \\ \vdots \\ \vdots \\ c_n \end{bmatrix}$$

and n is the dimension of the vector.

Example

Give an example of a column vector.

Solution

$[C] = \begin{bmatrix} 25 \\ 5 \\ 6 \end{bmatrix}$ is an example of a column vector

of dimension 3.

Submatrix: If some row(s) or/and column(s) of a matrix [A] are deleted, the remaining matrix is called a submatrix of [A].

Example

Find some of the submatrices of the matrix

$$[A] = \begin{bmatrix} 4 & 6 & 2 \\ 3 & -1 & 2 \end{bmatrix}$$

Solution

$\begin{bmatrix} 4 & 6 & 2 \\ 3 & -1 & 2 \end{bmatrix}, \begin{bmatrix} 4 & 6 \\ 3 & -1 \end{bmatrix}, \begin{bmatrix} 4 & 6 & 2 \end{bmatrix}, [4], \begin{bmatrix} 2 \\ 2 \end{bmatrix}$ are all submatrices of [A]. Can you find other submatrices of [A]?

Square matrix: If the number of rows (m) of a matrix is equal to the number of columns (n) of the matrix, (m = n), it is called a square matrix. The entries a_{11}, a_{22}, . . . a_{nn} are called the **diagonal elements** of a square matrix. Sometimes the diagonal of the matrix is also called the **principal** or **main** of the matrix.

Example

Give an example of a square matrix.

Solution

$$[A] = \begin{bmatrix} 25 & 20 & 3 \\ 5 & 10 & 15 \\ 6 & 15 & 7 \end{bmatrix}$$

is a square matrix as it has same number of rows and columns, that is, three.

The diagonal elements of [A] are $a_{11} = 25$, $a_{22} = 10$, $a_{33} = 7$.

Upper triangular matrix: A mxn matrix for which $a_{ij} = 0$, i>j is called an upper triangular matrix. That is, all the elements below the diagonal entries are zero.

Example

Give an example of an upper triangular matrix.

Solution

$$[A] = \begin{bmatrix} 10 & -7 & 0 \\ 0 & -0.001 & 6 \\ 0 & 0 & 15005 \end{bmatrix}$$

is an upper triangular matrix.

Lower triangular matrix: A mxn matrix for which $a_{ij} = 0, j > i$ is called a lower triangular matrix. That is, all the elements above the diagonal entries are zero.

Example

Give an example of a lower triangular matrix.

Solution

$$[A] = \begin{bmatrix} 1 & 0 & 0 \\ 0.3 & 1 & 0 \\ 0.6 & 2.5 & 1 \end{bmatrix}$$

is a lower triangular matrix.

Diagonal matrix: A square matrix with all non-diagonal elements equal to zero is called a diagonal matrix, that is, only the diagonal entries of the square matrix can be non-zero, ($a_{ij} = 0, i \neq j$).

Example

Give examples of a diagonal matrix.

Solution

$$[A]=\begin{bmatrix}3 & 0 & 0\\ 0 & 2.1 & 0\\ 0 & 0 & 5\end{bmatrix}$$

is a diagonal matrix.

Any or all the diagonal entries of a diagonal matrix can be zero.

For example

$$[A]=\begin{bmatrix}3 & 0 & 0\\ 0 & 2.1 & 0\\ 0 & 0 & 0\end{bmatrix}$$

is also a diagonal matrix.

Identity matrix: A diagonal matrix with all diagonal elements equal to one is called an identity matrix, ($a_{ij} = 0$, $i \neq j$; and $a_{ii} = 1$ for all i).

.

Example

Give an example of an identity matrix.

Solution

$$[A]=\begin{bmatrix}1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{bmatrix}$$

is an identity matrix.

Zero matrix: A matrix whose all entries are zero is called a zero matrix, (a_{ij} = 0 for all i and j).

Example

Give examples of a zero matrix.

Solution

$$[A]=\begin{bmatrix}0 & 0 & 0\\0 & 0 & 0\\0 & 0 & 0\end{bmatrix}$$

$$[B]=\begin{bmatrix}0 & 0 & 0\\0 & 0 & 0\end{bmatrix}$$

$$[C]=\begin{bmatrix}0 & 0 & 0 & 0\\0 & 0 & 0 & 0\\0 & 0 & 0 & 0\end{bmatrix}$$

$$[D]=\begin{bmatrix}0 & 0 & 0\end{bmatrix}$$

are all examples of a zero matrix.

Tridiagonal matrices: A tridiagonal matrix is a square matrix in which all elements not on the major diagonal, the diagonal above the major diagonal and the diagonal below the major diagonal are zero.

Example

Give an example of a tridiagonal matrix.

Solution

$$[A]=\begin{bmatrix} 2 & 4 & 0 & 0 \\ 2 & 3 & 9 & 0 \\ 0 & 0 & 5 & 2 \\ 0 & 0 & 3 & 6 \end{bmatrix}$$

is a tridiagonal matrix.

Do non-square matrices have diagonal entries?

Yes, for a mxn matrix [A], the diagonal entries are $a_{11}, a_{22} ..., a_{k-1,k-1}, a_{kk}$ where k=min {m,n}.

Example

What are the diagonal entries of

$$[A]=\begin{bmatrix} 3.2 & 5 \\ 6 & 7 \\ 2.9 & 3.2 \\ 5.6 & 7.8 \end{bmatrix}$$

Solution

The diagonal elements of [A] are $a_{11} = 3.2$ and $a_{22} = 7$.

Diagonally Dominant Matrix: A nxn square matrix [A] is a diagonally dominant matrix if

$$|a_{ii}| \geq \sum_{\substack{j=1 \\ i \neq j}} | a_{ij} | \text{ for all i =1, 2, …, n}$$

and $|a_{ii}| > \sum_{\substack{j=1 \\ i \neq j}}^{n} | a_{ij} |$ for at least one i, that is, for each row, the absolute value of the diagonal element is greater than or equal to the sum of the absolute values of the rest of the elements of that row, and that the inequality is strictly greater than for at least one row. Diagonally dominant matrices are important in ensuring convergence in iterative schemes of solving simultaneous linear equations.

Example

Give examples of diagonally dominant matrices and not diagonally dominant matrices.

Solution

$$[A] = \begin{bmatrix} 15 & 6 & 7 \\ 2 & -4 & -2 \\ 3 & 2 & 6 \end{bmatrix}$$

is a diagonally dominant matrix

as

$$|a_{11}| = |15| = 15 \geq |a_{12}| + |a_{13}| = |6| + |7| = 13$$

$$|a_{22}| = |-4| = 4 \geq |a_{21}| + |a_{23}| = |2| + |2| = 4$$

$$|a_{33}| = |6| = 6 \geq |a_{31}| + |a_{32}| = |3| + |2| = 5$$

and for at least one row, that is Rows 1 and 3 in this case, the inequality is a strictly greater than inequality.

$$[A] = \begin{bmatrix} -15 & 6 & 9 \\ 2 & -4 & 2 \\ 3 & -2 & 5.001 \end{bmatrix}$$

is a diagonally dominant matrix

as

$$|a_{11}| = |-15| = 15 \geq |a_{12}| + |a_{13}| = |6| + |9| = 15$$

$$|a_{22}| = |-4| = 4 \geq |a_{21}| + |a_{23}| = |2| + |2| = 4$$

$$|a_{33}| = |5.001| = 5.001 \geq |a_{31}| + |a_{32}| = |3| + |-2| = 5$$

the inequalities are satisfied for all rows and it is satisfied strictly greater than for at least one row (in this case it is Row 3)

$$[A] = \begin{bmatrix} 25 & 5 & 1 \\ 64 & 8 & 1 \\ 144 & 12 & 1 \end{bmatrix}$$

is not diagonally dominant as

$$|a_{22}| = |8| = 8 \leq |a_{21}| + |a_{23}| = |64| + |1| = 65$$

When are two matrices considered to be equal?

Two matrices [A] and [B] are equal if the size of [A] and [B] is the same (number of rows and columns are same for [A] and [B]) and $a_{ij} = b_{ij}$ for all i and j.

Example

What would make

$[A] = \begin{bmatrix} 2 & 3 \\ 6 & 7 \end{bmatrix}$ to be equal to

$[B] = \begin{bmatrix} b_{11} & 3 \\ 6 & b_{22} \end{bmatrix}$,

Solution

The two matrices $[A]$ and $[B]$ would be equal if

$b_{11} = 2$, $b_{22} = 7$.

Key Terms

Matrix Vector Sub-matrix Square matrix

Upper triangular matrix Lower triangular matrix Diagonal matrix

Identity matrix Zero matrix Tridiagonal matrix

Diagonally dominant matrix Equal matrices.

Homework Assignment

1. Write an example of a row vector.
2. Write an example of a column vector.
3. Write an example of a square matrix.
4. Write an example of a diagonal matrix.
5. Write an example of a tridiagonal matrix.
6. Write an example of a identity matrix.
7. Write an example of a upper triangular matrix.
8. Write an example of a lower triangular matrix.
9. Are these matrices strictly diagonally dominant?

a) $[A]=\begin{bmatrix} 15 & 6 & 7 \\ 2 & -4 & 2 \\ 3 & 2 & 6 \end{bmatrix}$

b) $[A]=\begin{bmatrix} 5 & 6 & 7 \\ 2 & -4 & 2 \\ 3 & 2 & -5 \end{bmatrix}$

c) $[A]=\begin{bmatrix} 5 & 3 & 2 \\ 6 & -8 & 2 \\ 7 & -5 & 12 \end{bmatrix}$

Answer: a) Yes b) No c)No

10. Give an example of a diagonally dominant matrix with no zero elements.

11. If $[\mathrm{A}]=\begin{bmatrix} 4 & -1 \\ 0 & 2 \end{bmatrix}$

What are b_{11} and b_{12} in

$[\mathrm{B}]=\begin{bmatrix} \mathrm{b}_{11} & b_{12} \\ 0 & 4 \end{bmatrix}$

if [B] = 2[A].

Answers: 8, -2

12.Find all the submatrices of

$[A]=\begin{bmatrix} 10 & -7 & 0 \\ 0 & -0.001 & 6 \end{bmatrix}$

Chapter 2
Vectors

After reading this chapter, you should be able to

- *Know what a vector is*
- *How to add and subtract vectors*
- *How to find linear combination of vectors and their relationship to a set of equations*
- *Know what it means to have linearly independent set of vectors*
- *How to find the rank of a set of vectors*

What is a vector?

A vector is a collection of numbers in a definite order. If it is a collection of ‘n’ numbers, it is called a n-dimensional vector. So the vector $\vec{A}$ given by

$$\vec{A} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$$

is a n-dimensional column vector with n components, a_1, a_2 . . . , a_n. The above is a column vector. A row vector [B] is of the form

$\vec{B}$ = [b_1, b_2,. . . , b_n]

where $\vec{B}$ is a n-dimensional row vector with n components b_1, b_2, . . ., b_n.

Example

Give an example of a 3-dimensional column vector.

Solution

Assume a point in space is given by its (x,y,z) coordinates. Then if the value of x=3, y=2, z=5, the column vector corresponding to the location of the points is $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \\ 5 \end{bmatrix}$.

When are two vectors equal?

Two vectors $\vec{A}$ and $\vec{B}$ are equal if they are of the same dimension and if their corresponding components are equal.

Given

$$\vec{A} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$$

and

$$\vec{B} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

then $\vec{A} = \vec{B}$ if $a_i = b_i$, i = 1, 2, ..., n.

Example

What are the values of the unknown components in $\vec{B}$ if

$$\vec{A} = \begin{bmatrix} 2 \\ 3 \\ 4 \\ 1 \end{bmatrix} \text{ and } \vec{B} = \begin{bmatrix} b_1 \\ 3 \\ 4 \\ b_4 \end{bmatrix}$$

and $\vec{A} = \vec{B}$.

Solution

$$b_1 = 2, b_4 = 1.$$

How do you add two vectors?

Two vectors can be added only if they are of the same dimension and the addition is given by

$$[A] + [B] = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

$$= \begin{bmatrix} a_1 + b_1 \\ a_2 + b_2 \\ \vdots \\ a_n + b_n \end{bmatrix}$$

Example

Add the two vectors

$$\vec{A} = \begin{bmatrix} 2 \\ 3 \\ 4 \\ 1 \end{bmatrix} \text{ and } \vec{B} = \begin{bmatrix} 5 \\ -2 \\ 3 \\ 7 \end{bmatrix}$$

Solution

$$\vec{A}+\vec{B}=\begin{bmatrix}2\\3\\4\\1\end{bmatrix}+\begin{bmatrix}5\\-2\\3\\7\end{bmatrix}$$

$$=\begin{bmatrix}2+5\\3-2\\4+3\\1+7\end{bmatrix}$$

$$=\begin{bmatrix}7\\1\\7\\8\end{bmatrix}$$

Example

A store sells three brands of tires, Tirestone, Michigan and Cooper. In quarter 1, the sales are given by the vector

$$\vec{A}_1=\begin{bmatrix}25\\25\\6\end{bmatrix}$$

where the rows represent the three brands of tires sold – Tirestone, Michigan and Cooper. In quarter 2, the sales are given by

$$\vec{A}_2=\begin{bmatrix}20\\10\\6\end{bmatrix}$$

What is the total sale of each brand of tire in the first half year?

Solution

The total sales would be given by

$$\vec{C} = \vec{A}_1 + \vec{A}_2$$

$$= \begin{bmatrix} 25 \\ 5 \\ 6 \end{bmatrix} + \begin{bmatrix} 20 \\ 10 \\ 6 \end{bmatrix}$$

$$= \begin{bmatrix} 25+20 \\ 5+10 \\ 6+6 \end{bmatrix}$$

$$= \begin{bmatrix} 45 \\ 15 \\ 12 \end{bmatrix}$$

So number of Tirestone tires sold is 25, Michigan is 15 and Cooper is 12 in the first half year.

What is a null vector?

A null vector is where all the components are zero.

Example

Give an example of a null vector or zero vector is

Solution

The vector $\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$ is an example of a zero or null vector.

What is a unit vector?

A unit vector $\vec{U}$ is defined if

$$\vec{U} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}$$

where $\sqrt{u_1^2 + u_2^2 + u_3^2 + \ldots + u_n^2} = 1$

Example

Give examples of 3-dimensional unit column vectors.

Solution

Examples include

$$\begin{bmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \text{etc.}$$

How do you multiply a vector by a scalar?

If k is a scalar and $\vec{A}$ is a n-dimensional vector, then

$$k\vec{A} = k\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} ka_1 \\ ka_2 \\ \vdots \\ ka_n \end{bmatrix}$$

Example

What is $2\vec{A}$ if

$$\vec{A} = \begin{bmatrix} 25 \\ 20 \\ 5 \end{bmatrix}$$

Solution

$$2\vec{A} = 2\begin{bmatrix} 25 \\ 20 \\ 5 \end{bmatrix}$$

$$= \begin{bmatrix} (2)(25) \\ (2)(20) \\ (2)(5) \end{bmatrix}$$

$$= \begin{bmatrix} 50 \\ 40 \\ 10 \end{bmatrix}$$

Example

A store sells three brands of tires, Tirestone, Michigan and Cooper. In quarter 1, the sales are given by the vector

$$\vec{A} = \begin{bmatrix} 25 \\ 25 \\ 6 \end{bmatrix}$$

If the goal is to increase the sales of all tires by at least 25% in the next quarter, how many of each brand should be the goal of the store?

Solution

Since the goal is to increase the sales by 25%, one would multiply the $\vec{A}$ vector by 1.25,

$$\vec{B} = 1.25\begin{bmatrix} 25 \\ 25 \\ 6 \end{bmatrix}$$

$$= \begin{bmatrix} 31.25 \\ 31.25 \\ 7.5 \end{bmatrix}$$

Since the number of tires is an integer we can say that the goal of sales would be

$$\vec{B} = \begin{bmatrix} 32 \\ 32 \\ 8 \end{bmatrix}$$

What do you mean by a linear combination of vectors?

Given $\vec{A}_1$, $\vec{A}_2$, ………. , $\vec{A}_m$ as m vectors of same dimension n, then if k_1, k_2, ……. , k_m are scalars, then

$k_1 \vec{A}_1 + k_2 \vec{A}_2 + \ldots \ldots\ldots\ldots\ldots + k_m \vec{A}_m$

is a linear combination of the m vectors.

Example

Find the linear combinations

a) [A] – [B], and

b) [A] + [B] – 3[C], where

$$\vec{A} = \begin{bmatrix} 2 \\ 3 \\ 6 \end{bmatrix}, \vec{B} = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}, \vec{C} = \begin{bmatrix} 10 \\ 1 \\ 2 \end{bmatrix}$$

Solution

a) $\vec{A} - \vec{B} = \begin{bmatrix} 2 \\ 3 \\ 6 \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$

$$= \begin{bmatrix} 2-1 \\ 3-1 \\ 6-2 \end{bmatrix}$$

$$= \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix}$$

b) $\vec{A} + \vec{B} - 3\vec{C} = \begin{bmatrix} 2 \\ 3 \\ 6 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} - 3\begin{bmatrix} 10 \\ 1 \\ 2 \end{bmatrix}$

$$= \begin{bmatrix} 2+1-30 \\ 3+1-3 \\ 6+2-6 \end{bmatrix}$$

$$= \begin{bmatrix} -27 \\ 1 \\ 2 \end{bmatrix}$$

What do you mean by vectors being linearly independent?

A set of vectors $\vec{A}_1, \vec{A}_2, \ldots, \vec{A}_m$ are considered to be linearly independent if

$k_1 \vec{A}_1 + k_2 \vec{A}_2 + \ldots\ldots\ldots\ldots + k_m \vec{A}_m = \vec{0}$

has only one solution of $k_1 = k_2 = \ldots = k_m = 0$.

Example

Are the three vectors

$$\vec{A}_1 = \begin{bmatrix} 25 \\ 64 \\ 144 \end{bmatrix}, \vec{A}_2 = \begin{bmatrix} 5 \\ 8 \\ 12 \end{bmatrix}, \vec{A}_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

linearly independent.

Solution

Writing the linear combination of the three vectors

$$k_1\begin{bmatrix}25\\64\\144\end{bmatrix}+k_2\begin{bmatrix}5\\8\\12\end{bmatrix}+k_3\begin{bmatrix}1\\1\\1\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix}$$

gives

$$\begin{bmatrix}25k_1+5k_2+k_3\\64k_1+8k_2+k_3\\144k_1+12k_2+k_3\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix}$$

The only solution is $k_1 = k_2 = k_3 = 0$.

Example

Are the three vectors

$$\vec{A}_1=\begin{bmatrix}1\\2\\5\end{bmatrix},\vec{A}_2=\begin{bmatrix}2\\5\\7\end{bmatrix},A_3=\begin{bmatrix}6\\14\\24\end{bmatrix}$$

linearly independent?

Solution

By inspection,

$$\vec{A}_3=2\vec{A}_1+2\vec{A}_2$$

or

$$-2\vec{A}_1-2\vec{A}_2+\vec{A}_3=0$$

So the linear combination

$$k_1\vec{A}_1+k_2\vec{A}_2+k_3\vec{A}_3=\vec{0}$$

has a non-zero solution

$$k_1=-2,k_2=-2,k_3=1$$

Hence the set of vectors is linearly dependent.

Example

Are the three vectors

$$\vec{A}_1 = \begin{bmatrix} 25 \\ 64 \\ 89 \end{bmatrix}, \vec{A}_2 = \begin{bmatrix} 5 \\ 8 \\ 13 \end{bmatrix}, \vec{A}_3 = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$$

linearly independent.

Solution

Writing the linear combination of the three vectors

$$k_1 \begin{bmatrix} 25 \\ 64 \\ 89 \end{bmatrix} + k_2 \begin{bmatrix} 5 \\ 8 \\ 13 \end{bmatrix} + k_3 \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

gives

$$\begin{bmatrix} 25k_1 + 5k_2 + k_3 \\ 64k_1 + 8k_2 + k_3 \\ 89k_1 + 13k_2 + 2k_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

In addition to $k_1 = k_2 = k_3 = 0$, one can find other solutions for which k_1, k_2, k_3 are not equal to zero. For example $k_1 = 1$, $k_2 = -13$, $k_3 = 40$ is also a solution. This implies

$$1\begin{bmatrix} 25 \\ 64 \\ 89 \end{bmatrix} - 13 \begin{bmatrix} 5 \\ 8 \\ 13 \end{bmatrix} + 40 \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

So the linear combination that gives us a zero vector consists of non-zero constants.

Hence $\vec{A}_1, \vec{A}_2, \vec{A}_3$ are linearly dependent.

What do you mean by the rank of a set of vectors?

From a set of n-dimensional vectors, the maximum number of linearly independent vectors in the set is called the rank of the set of vectors. Note that the rank of the vectors can never be greater than its dimension.

Example

What is the rank of

$$\vec{A}_1 = \begin{bmatrix} 25 \\ 64 \\ 144 \end{bmatrix}, \vec{A}_2 = \begin{bmatrix} 5 \\ 8 \\ 12 \end{bmatrix}, \vec{A}_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

Solution:

Since we found in a previous example that $\vec{A}_1, \vec{A}_2$, and $\vec{A}_3$ are linearly independent, the rank of the set of vectors $\vec{A}_1, \vec{A}_2, \vec{A}_3$ is 3.

Example

What is the rank of

$$\vec{A} = \begin{bmatrix} 25 \\ 64 \\ 89 \end{bmatrix}, \vec{A}_2 = \begin{bmatrix} 5 \\ 8 \\ 13 \end{bmatrix}, \vec{A}_3 = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$$

Solution

Since we found that $\vec{A}_1, \vec{A}_2$ and $\vec{A}_3$ are not linearly independent, the rank of $\vec{A}_1, \vec{A}_2, \vec{A}_3$ is not 3, and hence is less than 3. Is it 2? Let us choose

$$\vec{A}_1 = \begin{bmatrix} 25 \\ 64 \\ 89 \end{bmatrix}, \vec{A}_2 = \begin{bmatrix} 5 \\ 8 \\ 13 \end{bmatrix}$$

Linear combination of $\vec{A}_1$ and $\vec{A}_2$ equal to zero has only one solution. So the rank is 2.

Prove that if a set of vectors contains the null vector, the set of vectors is linearly dependent.

Let $\vec{A}_1, \vec{A}_2, \ldots, \vec{A}_m$ be a set of n-dimensional vectors, then

$$k_1\vec{A}_1 + k_2\vec{A}_2 + \ldots + k_m\vec{A}_m = 0$$

is a linear combination of the 'm' vectors. Then assuming if $\vec{A}_1$ is the zero or null vector, any value of k_1 coupled with $k_2 = k_3 = \ldots = k_m = 0$ will satisfy the above equation.

Hence the set of vectors is linearly dependent as more than one solution exists.

Prove that if a set of vectors are linearly independent, then a subset of the m vectors also has to be linearly independent.

Let this subset be

$$\vec{A}_{a1}, \vec{A}_{a2}, \ldots, \vec{A}_{ap}$$

where p < m.

Then if this subset is linearly dependent, the linear combination

$$k_1\vec{A}_{a1} + k_2\vec{A}_{a2} + \ldots + k_p\vec{A}_{ap} = 0$$

has a non-trivial solution.

So

$$k_1\vec{A}_{a1} + k_2\vec{A}_{a2} + \ldots + k_p\vec{A} + 0\vec{A}_{a((p+1)} \ldots 0\vec{A}_{am} = 0_{ap}$$

has a non-trivial solution too, where $A_{a(p+1)}, \ldots, A_{am}$ are the rest of the (m-p) vectors. But this is a contradiction. So a subset of linearly independent vectors cannot be linearly dependent.

Prove that if a set of vectors is linearly dependent, then at least one vector can be written as a linear combination of others.

Let $\vec{A}_1, \vec{A}_2, \ldots, \vec{A}_m$ be linearly dependent, then there exists a set of numbers

$k_1, \ldots, k_m$ not all of which are zero for the linear combination

$$k_1\vec{A}_1 + k_2\vec{A}_2 + \ldots + k_m\vec{A}_m = 0$$

that one of non-zero values of $k_i, i = 1, \ldots, m$, is for let i = p, then

$$A_p = -\frac{k_2}{k_{p1}}\vec{A}_2 - \ldots - \frac{k_{p-1}}{k_1}A_{p-1} - \frac{k_{p+1}}{k_p}A_p - \ldots - \frac{k_m}{k_p}A_m.$$

and that proves the theorem.

Prove that if the dimension of a set of vectors is less than the number of vectors in the set, then the set of vectors is linearly dependent.

Can you prove it??

How can vectors be used to write simultaneous linear equations?

A set of m linear equations with n unknowns is written as

$$a_{11}x_1 + \ldots + a_{1n}x_n = c_1$$

$$a_{21}x_1 + \ldots + a_{2n}x_n = c_2$$

$$\vdots \qquad \vdots$$

$$\vdots \qquad \vdots$$

$$a_{m1}x_1 + \ldots + a_{mn}x_n = c_n$$

where

$x_1, x_2, \ldots, x_n$ are the unknowns, then in the vector notation they can be written as

$$x_1\vec{A}_1 + x_2\vec{A}_2 + \ldots + x_n\vec{A}_n = \vec{C}$$

where

$$\vec{A}_1 = \begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \end{bmatrix}$$

$$\vec{A}_2 = \begin{bmatrix} a_{12} \\ \vdots \\ a_{m2} \end{bmatrix}$$

$$\vec{A}_n = \begin{bmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{bmatrix}$$

The problem now becomes whether you can find the scalars x_1, ………, x_n such that the linear combination

$x_1\vec{A}_1 + ...+ x_n\vec{A}_n = \vec{C}$.

Example

Write

$$25x_1 + 5x_2 + x_3 = 106.8$$

$$64x_1 + 8x_2 + x_3 = 177.2$$

$$144x_1 + 12x_2 + x_3 = 279.2$$

as a linear combination of vectors.

Solution

$$\begin{bmatrix} 25x_1 & +5x_2 & +x_3 \\ 64x_1 & +8x_2 & +x_3 \\ 144x_1 & +12x_2 & +x_3 \end{bmatrix} = \begin{bmatrix} 106.8 \\ 177.2 \\ 279.2 \end{bmatrix}$$

$$x_1\begin{bmatrix} 25 \\ 64 \\ 144 \end{bmatrix} + x_2\begin{bmatrix} 5 \\ 8 \\ 12 \end{bmatrix} + x_3\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 106.8 \\ 177.2 \\ 279.2 \end{bmatrix}$$

What is the definition of the dot product of two vectors?

Let $\vec{A} = [a_1, a_2, \ldots, a_n]$ and $\vec{B} = [b_1, b_2, \ldots, b_n]$ be two n-dimensional vectors. Then the dot product of the two vectors $\vec{A}$ and $\vec{B}$ is defined as

$$\vec{A}\cdot\vec{B} = a_1b_1 + a_2b_2 + \ldots + a_nb_n = \sum_{i=1}^{n} a_ib_i$$

A dot product is also called an inner product or scalar.

Example

Find the dot product of the two vectors $\vec{A}$ = (4, 1, 2, 3) and $\vec{B}$ = (3, 1, 7, 2).

Solution

$\vec{A}\cdot\vec{B} = (4,1,2,3)\cdot(3,1,7,2)$

$= (4)(3)+(1)(1)+(2)(7)+(3)(2)$

$= 33$

Example

A product line needs three types of rubber as given in the table below.

Rubber Type	**Weight**	**Cost**
	lbs	$
A	200	20.23
B	250	30.56
C	310	29.12

How much is the total price of the rubber needed?

Solution

The weight vector is given by

$$\vec{W} = (200, 250, 310)$$

and the cost vector is given by

$$\vec{C} = (20.23, 30.56, 29.12).$$

The total cost of the rubber would be the dot product of $\vec{W}$ and $\vec{C}$.

$$\vec{W} \cdot \vec{C} = (200, 250, 310) \cdot (20.23, 30.56, 29.12)$$

$= (200)(20.23) + (250)(30.56) + (310)(29.12)$

$= 4046 + 7640 + 9027.2$

$= \$20713.2$

Key Terms

Vector Addition of vectors Subtraction of vectors

Unit vectors Null vector Scalar multiplication of vectors

Linear combination of vectors Linearly independent vectors

Rank Dot products.

Homework

1. For

$$\vec{A} = \begin{bmatrix} 2 \\ 9 \\ -7 \end{bmatrix}, \vec{B} = \begin{bmatrix} 3 \\ 2 \\ 5 \end{bmatrix}, \vec{C} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

find

a) $\vec{A} + \vec{B}$

b) $2\vec{A} - 3\vec{B} + \vec{C}$

2. Are

$$\vec{A} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \vec{B} = \begin{bmatrix} 1 \\ 2 \\ 5 \end{bmatrix}, \vec{C} = \begin{bmatrix} 1 \\ 4 \\ 25 \end{bmatrix}$$

linearly independent. What is the rank of the above set of vectors?

3. Are

$$\vec{A} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \vec{B} = \begin{bmatrix} 1 \\ 2 \\ 5 \end{bmatrix}, \vec{C} = \begin{bmatrix} 3 \\ 5 \\ 7 \end{bmatrix}$$

linearly independent. What is the rank of the above set of vectors?

4. If a set of vectors contains the null vector, the set of vectors is

A. linearly independent.

B. linearly dependent.

5. If a set of vectors is linearly independent, a subset of the vectors is

   A. linearly independent.

   B. linearly dependent.

6. If a set of vectors is linearly dependent, then

   A. at least one vector can be written as a linear combination of others.

   B. At least one vector is a null vector.

7. If the dimension of a set of vectors is less than the number of vectors in the set, then the set of vectors is

   A. linearly dependent.

   B. linearly independent.

8. Find the dot product of

   $\vec{A} = (2,1,2.5,3)$ and $\vec{B} = (-3,2,1,2.5)$

9. If $\vec{u}, \vec{v}, \vec{w}$ are three nonzero vector of 2-dimensions, then

   A. $\vec{u}, \vec{v}, \vec{w}$ are linearly independent

   B. $\vec{u}, \vec{v}, \vec{w}$ are linearly dependent

   C. $\vec{u}, \vec{v}, \vec{w}$ are unit vectors

   D. $k_1\vec{u} + k_2\vec{v} + k_3\vec{v} = \vec{0}$ has a unique solution.

Chapter 3
Binary Matrix Operations

After reading this chapter, you will be able to

- *Add, subtract and multiply matrices*
- *Learn rules of binary operations on matrices*

How do you add two matrices?

Two matrices [A] and [B] can be added only if they are the same size, then the addition is shown as

[C] = [A] + [B]

where

$c_{ij} = a_{ij} + b_{ij}$

Example

Add two matrices

$$[A] = \begin{bmatrix} 5 & 2 & 3 \\ 1 & 2 & 7 \end{bmatrix}$$

$$[B] = \begin{bmatrix} 6 & 7 & -2 \\ 3 & 5 & 19 \end{bmatrix}$$

Solution

$$[C] = [A] + [B]$$

$$= \begin{bmatrix} 5 & 2 & 3 \\ 1 & 2 & 7 \end{bmatrix} + \begin{bmatrix} 6 & 7 & -2 \\ 3 & 5 & 19 \end{bmatrix}$$

$$= \begin{bmatrix} 5+6 & 2+7 & 3-2 \\ 1+3 & 2+5 & 7+19 \end{bmatrix}$$

$$= \begin{bmatrix} 11 & 9 & 1 \\ 4 & 7 & 26 \end{bmatrix}$$

Example

Blowout r'us store has two locations 'A' and 'B', and their sales of tires are given by make (in rows) and quarters (in columns) as shown below.

$$[\mathrm{A}] = \begin{bmatrix} 25 & 20 & 3 & 2 \\ 5 & 10 & 15 & 25 \\ 6 & 16 & 7 & 27 \end{bmatrix}$$

$$[\mathrm{B}] = \begin{bmatrix} 20 & 5 & 4 & 0 \\ 3 & 6 & 15 & 21 \\ 4 & 1 & 7 & 20 \end{bmatrix}$$

where the rows represent sale of Tirestone, Michigan and Copper tires and the columns represent the quarter number - 1, 2, 3, 4. What are the total sales of the two locations by make and quarter?

Solution

$$[C] = [A] + [B]$$

$$= \begin{bmatrix} 25 & 20 & 3 & 2 \\ 5 & 10 & 15 & 25 \\ 6 & 16 & 7 & 27 \end{bmatrix} + \begin{bmatrix} 20 & 5 & 4 & 0 \\ 3 & 6 & 15 & 21 \\ 4 & 1 & 7 & 20 \end{bmatrix}$$

$$= \begin{bmatrix} (25+20) & (20+5) & (3+4) & (2+0) \\ (5+3) & (10+6) & (15+15) & (25+21) \\ (6+4) & (16+1) & (7+7) & (27+20) \end{bmatrix}$$

$$= \begin{bmatrix} 45 & 25 & 7 & 2 \\ 8 & 16 & 30 & 46 \\ 10 & 17 & 14 & 47 \end{bmatrix}$$

So if one wants to know the total number of Copper tires sold in quarter 4 in the two locations, we would look at Row 3 – Column 4 to give

$c_{34} = 47$.

How do you subtract two matrices?

Two matrices [A] and [B] can be subtracted only if they are the same size and the subtraction is given by

[D] = [A] – [B]

where

$d_{ij} = a_{ij} - b_{ij}$

Example

Subtract matrix [B] from matrix [A].

$$[A] = \begin{bmatrix} 5 & 2 & 3 \\ 1 & 2 & 7 \end{bmatrix}$$

$$[B] = \begin{bmatrix} 6 & 7 & -2 \\ 3 & 5 & 19 \end{bmatrix}$$

Solution

$$[C] = [A] - [B]$$

$$= \begin{bmatrix} 5 & 2 & 3 \\ 1 & 2 & 7 \end{bmatrix} - \begin{bmatrix} 6 & 7 & -2 \\ 3 & 5 & 19 \end{bmatrix}$$

$$= \begin{bmatrix} 5-6 & 2-7 & 3-(-2) \\ 1-3 & 2-5 & 7-19 \end{bmatrix}$$

$$= \begin{bmatrix} -1 & -5 & 5 \\ -2 & -3 & -12 \end{bmatrix}$$

Example

Blowout r'us store has two locations A and B and their sales of tires are given by make (in rows) and quarters (in columns) as shown below.

$$[\mathrm{A}] = \begin{bmatrix} 25 & 20 & 3 & 2 \\ 5 & 10 & 15 & 25 \\ 6 & 16 & 7 & 27 \end{bmatrix}$$

$$[\mathrm{B}] = \begin{bmatrix} 20 & 5 & 4 & 0 \\ 3 & 6 & 15 & 21 \\ 4 & 1 & 7 & 20 \end{bmatrix}$$

where the rows represent sale of Tirestone, Michigan and Copper tires and the columns represent the quarter number- 1, 2, 3, 4. How many more tires did store A sell than store B of each brand in each quarter?

Solution

[D]

$= [A] - [B]$

$$= \begin{bmatrix} 25 & 20 & 3 & 2 \\ 5 & 10 & 15 & 25 \\ 6 & 16 & 7 & 27 \end{bmatrix} - \begin{bmatrix} 20 & 5 & 4 & 0 \\ 3 & 6 & 15 & 21 \\ 4 & 1 & 7 & 20 \end{bmatrix}$$

$$= \begin{bmatrix} 25-20 & 20-5 & 3-4 & 2-0 \\ 5-3 & 10-6 & 15-15 & 25-21 \\ 6-4 & 16-1 & 7-7 & 27-20 \end{bmatrix}$$

$$= \begin{bmatrix} 5 & 15 & -1 & 2 \\ 2 & 4 & 0 & 4 \\ 2 & 15 & 0 & 7 \end{bmatrix}$$

So if you want to know how many more Copper Tires were sold in quarter 4 in Store A than Store B, $d_{34} = 7$. Note that $d_{13} = -1$ implying that store A sold 1 less Michigan tire than Store B in quarter 3.

How do I multiply two matrices?

Two matrices [A] and [B] can be multiplied only if the number of columns of [A] is equal to the number of rows of [B] to give

$$[C]_{mxn} = [A]_{mxp}[B]_{pxn}$$

If [A] is a mxp matrix and [B] is a pxn matrix, the resulting matrix [C] is a mxn matrix.

So how does one calculate the elements of [C] matrix?

$$\begin{aligned} c_{ij} &= \sum_{k=1}^{p} a_{ik} b_{kj} \\ &= a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots\ldots + a_{ip}b_{pj} \end{aligned}$$

for each $i = 1,\ 2, \ldots\ldots, m, and\ j = 1,\ 2, \ldots\ldots, n$.

To put it in simpler terms, the i^{th} row and j^{th} column of the [C] matrix in [C] = [A][B] is calculated by multiplying the i^{th} row of [A] by the j^{th} column of [B], that is,

$$\begin{aligned} c_{ij} &= \begin{bmatrix} a_{i1} & a_{i2} & \ldots\ldots a_{ip} \end{bmatrix} \begin{bmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ \vdots \\ b_{pj} \end{bmatrix} \\ &= \mathrm{a_{i1}\ b_{1j} + a_{i2}\ b_{2j} + \ldots\ldots + a_{ip}\ b_{pj}}. \\ &= \sum_{k=1}^{p} a_{ik} b_{kj} \end{aligned}$$

Example

Given

$$[A]=\begin{bmatrix}5 & 2 & 3\\1 & 2 & 7\end{bmatrix}$$

$$[B]=\begin{bmatrix}3 & -2\\5 & -8\\9 & -10\end{bmatrix}$$

find

$$[C]=[A][B]$$

Solution

c_{12} can be found by multiplying the first row of [A] by the second column of [B],

$$c_{12}=\begin{bmatrix}5 & 2 & 3\end{bmatrix}\begin{bmatrix}-2\\-8\\-10\end{bmatrix}$$

= (5)(-2) + (2)(-8) + (3)(-10)

= -56

Similarly, one can find the other elements of [C] to give

$$[C]=\begin{bmatrix}52 & -56\\76 & -88\end{bmatrix}$$

Example

Blowout r’us store location A and the sales of tires are given by make (in rows) and quarters (in columns) as shown below

$$[\mathrm{A}]=\begin{bmatrix}25 & 20 & 3 & 2\\5 & 10 & 15 & 25\\6 & 16 & 7 & 27\end{bmatrix}$$

where the rows represent sale of Tirestone, Michigan and Copper tires and the columns represent the quarter number - 1, 2, 3, 4. Find the per quarter sales of store A if following are the prices of each tire.

Tirestone = $33.25

Michigan = $40.19

Copper = $25.03

Solution

The answer is given by multiplying the price matrix by the quantity sales of store A. The price matrix is $[33.25 \quad 40.19 \quad 25.03]$, then the per quarter sales of store A would be given by

$$[C] = [33.25 \quad 40.19 \quad 25.03]\begin{bmatrix} 25 & 20 & 3 & 2 \\ 5 & 10 & 15 & 25 \\ 6 & 16 & 7 & 27 \end{bmatrix}$$

$$c_{ij} = \sum_{k=1}^{3} a_{ik} b_{kj}$$

$$c_{11} = \sum_{k=1}^{3} a_{1k} b_{k1}$$

$$= a_{11}b_{11} + a_{12}b_{21} + a_{13}b_{31}$$

$$= (33.25)(25) + (40.19)(5) + (25.03)(6)$$

= $1182.38

Similarly

$$c_{12} = \$1467.38,$$
$$c_{13} = \$877.81,$$
$$c_{14} = \$1747.06.$$

So each quarter sales of store A in dollars are given by the four columns of the row vector

$$[C] = [1182.38 \quad 1467.38 \quad 877.81 \quad 1747.06]$$

Remember since we are multiplying a 1 x 3 matrix by a 3 x 4 matrix, the resulting matrix is a 1 x 4 matrix.

What is a scalar product of a constant and a matrix?

If $[A]$ is a n x n matrix and k is a real number, then the scalar product of k and [A] is another matrix [B], where $b_{ij} = k\,a_{ij}$.

Example

Let $[A] = \begin{bmatrix} 2.1 & 3 & 2 \\ 5 & 1 & 6 \end{bmatrix}$. Find 2 [A]

Solution

$$[A] = \begin{bmatrix} 2.1 & 3 & 2 \\ 5 & 1 & 6 \end{bmatrix}$$

Then

$$2[A]$$

$$=2\begin{bmatrix} 2.1 & 3 & 2 \\ 5 & 1 & 6 \end{bmatrix}$$

$$=\begin{bmatrix} (2)(2.1) & (2)(3) & (2)(2) \\ (2)(5) & (2)(1) & (2)(6) \end{bmatrix}$$

$$=\begin{bmatrix} 4.2 & 6 & 4 \\ 10 & 2 & 12 \end{bmatrix}$$

What is a linear combination of matrices?

If $[A_1]$, $[A_2]$, ……, $[A_p]$ are matrices of the same size and k_1, k_2, ……., k_p are scalars, then

$k_1[A_1]+k_2[A_2]+........+k_p[A_p]$

is called a linear combination of $[A_1],[A_2]\cdots\cdots,[A_p]$.

Example

If

$$[A_1]=\begin{bmatrix}5 & 6 & 2\\ 3 & 2 & 1\end{bmatrix},[A_2]=\begin{bmatrix}2.1 & 3 & 2\\ 5 & 1 & 6\end{bmatrix},[A_3]=\begin{bmatrix}0 & 2.2 & 2\\ 3 & 3.5 & 6\end{bmatrix}$$

then find

$[A_1]+2[A_2]-0.5[A_3]$

Solution

$$=\begin{bmatrix}5 & 6 & 2\\ 3 & 2 & 1\end{bmatrix}+2\begin{bmatrix}2.1 & 3 & 2\\ 5 & 1 & 6\end{bmatrix}-0.5\begin{bmatrix}0 & 2.2 & 2\\ 3 & 3.5 & 6\end{bmatrix}$$

$$=\begin{bmatrix}5 & 6 & 2\\ 3 & 2 & 1\end{bmatrix}+\begin{bmatrix}4.2 & 6 & 4\\ 10 & 2 & 12\end{bmatrix}-\begin{bmatrix}0 & 1.1 & 1\\ 1.5 & 1.75 & 3\end{bmatrix}$$

$$=\begin{bmatrix}9.2 & 10.9 & 5\\ 11.5 & 2.25 & 10\end{bmatrix}$$

What are some of the rules of binary matrix operations?

Commutative law of addition

If [A] and [B] are mxn matrices, then

$[A]+[B]=[B]+[A]$

Associate law of addition

If [A], [B] and [C] all are mxn matrices, then

$[A]+([B]+[C])=([A]+[B])+[C]$

Associate law of multiplication

If [A], [B] and [C] are mxn, nxp and pxr size matrices, respectively, then

$$[A]([B][C]) = ([A][B])[C]$$

and the resulting matrix size on both sides is mxr.

Distributive law:

If [A] and [B] are mxn size matrices, and [C] and [D] are nxp size matrices

$$[A]([C]+[D]) = [A][C]+[A][D]$$

$$([A]+[B])[C] = [A][C]+[B][C]$$

and the resulting matrix size on both sides is mxp.

Example

Illustrate the associative law of multiplication of matrices using

$$[A] = \begin{bmatrix} 1 & 2 \\ 3 & 5 \\ 0 & 2 \end{bmatrix}, \quad [B] = \begin{bmatrix} 2 & 5 \\ 9 & 6 \end{bmatrix}, \quad [C] = \begin{bmatrix} 2 & 1 \\ 3 & 5 \end{bmatrix}$$

Solution

$$[B][C] = \begin{bmatrix} 2 & 5 \\ 9 & 6 \end{bmatrix}\begin{bmatrix} 2 & 1 \\ 3 & 5 \end{bmatrix} = \begin{bmatrix} 19 & 27 \\ 36 & 39 \end{bmatrix}$$

$$[A][B][C] = \begin{bmatrix} 1 & 2 \\ 3 & 5 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} 19 & 27 \\ 36 & 39 \end{bmatrix} = \begin{bmatrix} 91 & 105 \\ 237 & 276 \\ 72 & 78 \end{bmatrix}$$

$$[A][B]=\begin{bmatrix}1 & 2\\ 3 & 5\\ 0 & 2\end{bmatrix}\begin{bmatrix}2 & 5\\ 9 & 6\end{bmatrix}=\begin{bmatrix}20 & 17\\ 51 & 45\\ 18 & 12\end{bmatrix}$$

$$[A][B][C]=\begin{bmatrix}20 & 17\\ 51 & 45\\ 18 & 12\end{bmatrix}\begin{bmatrix}2 & 1\\ 3 & 5\end{bmatrix}=\begin{bmatrix}91 & 105\\ 237 & 276\\ 72 & 78\end{bmatrix}$$

The above illustrates the associate law of multiplication of matrices.

Is [A][B]=[B][A]?

First both operations $[A][B]$ and $[B][A]$ are only possible if $[A]$ and $[B]$ are square matrices of same size. Why? If $[A][B]$ exists, number of columns of $[A]$ has to be same as the number of rows of $[B]$ and if $[B][A]$ exists, number of columns of $[B]$ has to be same as the number of rows of $[A]$.

Even then in general $[A][B]\neq[B][A]$.

Example

Illustrate if $[A][B]=[B][A]$ for the following matrices

$$[A]=\begin{bmatrix}6 & 3\\ 2 & 5\end{bmatrix},\quad [B]=\begin{bmatrix}-3 & 2\\ 1 & 5\end{bmatrix}$$

Solution

$[A][B]$

$$=\begin{bmatrix}6 & 3\\ 2 & 5\end{bmatrix}\begin{bmatrix}-3 & 2\\ 1 & 5\end{bmatrix}$$

$$\begin{bmatrix}-15 & 27\\ -1 & 29\end{bmatrix}$$

$[B][A] =$

$$\begin{bmatrix} -3 & 2 \\ 1 & 5 \end{bmatrix}\begin{bmatrix} 6 & 3 \\ 2 & 5 \end{bmatrix}$$

$$= \begin{bmatrix} -14 & 1 \\ 16 & 28 \end{bmatrix}$$

$[A][B] \neq [B][A]$

Key Terms

Addition of matrices Subtraction of matrices Multiplication of matrices

Scalar product of matrices Linear combination of matrices

Rules of binary matrix operation.

Homework

1. For the following matrices

$$[A] = \begin{bmatrix} 3 & 0 \\ -1 & 2 \\ 1 & 1 \end{bmatrix}, \quad [B] = \begin{bmatrix} 4 & -1 \\ 0 & 2 \end{bmatrix}, \quad [C] = \begin{bmatrix} 5 & 2 \\ 3 & 5 \\ 6 & 7 \end{bmatrix}$$

find where possible

a) 4[A] + 5[C]

b) [A][B]

c) [A] – 2[C]

Answers a) $= \begin{bmatrix} 37 & 10 \\ 11 & 33 \\ 34 & 39 \end{bmatrix}$ **b)** $\begin{bmatrix} 12 & -3 \\ -4 & 5 \\ 4 & 1 \end{bmatrix}$ **c)** $\begin{bmatrix} -7 & -4 \\ -7 & -8 \\ -11 & -13 \end{bmatrix}$

2. Food orders are taken from two engineering departments for a takeout. The order is tabulated below.

Food order:
$$\begin{array}{c} \\ \text{Mechanical} \\ \text{Civil} \end{array} \begin{array}{c} \begin{array}{ccc} \scriptstyle\text{Chicken Sandwich} & \scriptstyle\text{Fries} & \scriptstyle\text{Drink} \end{array} \\ \begin{bmatrix} 25 & 35 & 25 \\ 21 & 20 & 21 \end{bmatrix} \end{array}$$

However they have a choice of buying this food from three different restaurants. Their prices for the three food items are tabulated below

Price Matrix:
$$\begin{array}{c} \\ \textit{Chicken Sandwich} \\ \textit{Fries} \\ \textit{Drink} \end{array} \begin{array}{c} \begin{array}{ccc} \scriptstyle\textit{McFat} & \scriptstyle\textit{Burcholestrol} & \scriptstyle\textit{Kentucky Sodium} \end{array} \\ \begin{bmatrix} 2.42 & 2.38 & 2.46 \\ 0.93 & 0.90 & 0.89 \\ 0.95 & 1.03 & 1.13 \end{bmatrix} \end{array}$$

Show how much each department will pay for their order at each restaurant. Which restaurant would be more economical to order from for each department?

Answer: The cost in dollars is 116.80, 116.75, 120.90 for the Mechanical Department at three fast food joints. So BurCholestrol is the cheapest for the Mechanical Department.

The cost in dollars is 89.37, 89.61, 93.19 for the Civil Department at three fast food joints. McFat is the cheapest for the Civil Department.

3. Given

$$[A] = \begin{bmatrix} 2 & 3 & 5 \\ 6 & 7 & 9 \\ 2 & 1 & 3 \end{bmatrix}, [B] = \begin{bmatrix} 3 & 5 \\ 2 & 9 \\ 1 & 6 \end{bmatrix}. [C] = \begin{bmatrix} 5 & 2 \\ 3 & 9 \\ 7 & 6 \end{bmatrix}$$

Illustrate the distributive law of binary matrix operations

$$[A]([B]+[C]) = [A][B]+[A][C].$$

4. Let [I] be a n x n identity matrix. Show that [A][I] = [I][A] = [A] for every n x n matrix [A].

Chapter 4
Unary Matrix Operations

After reading this chapter, you should be able to

- *Know what unary operations mean*
- *Find the transpose of a square matrix and it relationship to symmetric matrices*
- *How to find the trace of a matrix*
- *How to find the determinant of a matrix by the cofactor method*

Transpose of a matrix: Let [A] be a m x n matrix. Then [B] is the transpose of the [A] if $b_{ji} = a_{ij}$ for all i and j. That is, the i^{th} row and the j^{th} column element of [A] is the j^{th} row and i^{th} column element of [B]. Note, [B] would be a n x m matrix. The transpose of [A] is denoted by $[A]^T$.

Example

Find the transpose of

$$[A]=\begin{bmatrix} 25 & 20 & 3 & 2 \\ 5 & 10 & 15 & 25 \\ 6 & 16 & 7 & 27 \end{bmatrix}$$

Solution

The transpose of [A] is

$$[A]^T = \begin{bmatrix} 25 & 5 & 6 \\ 20 & 10 & 16 \\ 3 & 15 & 7 \\ 2 & 25 & 27 \end{bmatrix}$$

Note, the transpose of a row vector is a column vector and the transpose of a column vector is a row vector.

Also, note that the transpose of a transpose of a matrix is the matrix itself, that is, $\left([A]^T\right)^T = [A]$. Also, $(A+B)^T = A^T + B^T; (cA)^T = cA^T$.

Symmetric matrix: A square matrix [A] with real elements where $a_{ij} = a_{ji}$ for i=1,….n and j=1,…,n is called a symmetric matrix. This is same as, if $[A] = [A]^T$, then [A] is a symmetric matrix.

Example

Give an example of a symmetric matrix.

Solution

$$[\mathrm{A}] = \begin{bmatrix} 21.2 & 3.2 & 6 \\ 3.2 & 21.5 & 8 \\ 6 & 8 & 9.3 \end{bmatrix}$$

is a symmetric matrix as $a_{12} = a_{21} = 3.2;\ \mathrm{a}_{13} = a_{31} = 6$ and $a_{23} = a_{32} = 8$.

What is a skew-symmetric matrix?

A nxn matrix is skew symmetric for which $a_{ij} = -a_{ji}$ for all i and j. This is same as $[A] = -[A]^T$.

Example

Give an example of a skew-symmetric matrix.

Solution

$$\begin{bmatrix} 0 & 1 & 2 \\ -1 & 0 & -5 \\ -2 & 5 & 0 \end{bmatrix}$$

is skew-symmetric as

$a_{12} = -a_{21} = 1$; $a_{13} = -a_{31} = 2$; $a_{23} = -a_{32} = -5$. Since $a_{ii} = -a_{ii}$ only if $a_{ii} = 0$, all the diagonal elements of a skew symmetric matrix have to be zero.

Trace of a matrix: The trace of a nxn matrix [A] is the sum of the diagonal entries of [A], that is,

$$tr[A] = \sum_{i=1}^{n} a_{ii}$$

Example

Find the trace of

$$[A] = \begin{bmatrix} 15 & 6 & 7 \\ 2 & -4 & 2 \\ 3 & 2 & 6 \end{bmatrix}$$

Solution

$$tr[A] = \sum_{i=1}^{3} a_{ii}$$

$$= (15) + (-4) + (6)$$

$$= 17$$

Example

The sale of tires are given by make (rows) and quarters (columns) for Blowout r'us store location A.

$$[A] = \begin{bmatrix} 25 & 20 & 3 & 2 \\ 5 & 10 & 15 & 25 \\ 6 & 16 & 7 & 27 \end{bmatrix}$$

where the rows represent sale of Tirestone, Michigan and Cooper tires, and the columns represent the quarter number 1,2,3,4.

Find the total yearly revenue of Store A if the prices of tires vary by quarters as follows.

$$[B] = \begin{bmatrix} 33.25 & 30.01 & 35.02 & 30.05 \\ 40.19 & 38.02 & 41.03 & 38.23 \\ 25.03 & 22.02 & 27.03 & 22.95 \end{bmatrix}$$

where the rows represent the cost of each tire made by Tirestone, Michigan and Cooper, the columns represent the quarter numbers.

Solution

To find the total sales of store A for the whole year, we need to find the sales of each brand of tire for the whole year and then add the total sales. To do so, we need to rewrite the price matrix so that the quarters are in rows and the brand names are in the columns, that is find transpose of $[B]$.

$$[C] = [B]^T = \begin{bmatrix} 33.25 & 30.01 & 35.02 & 30.05 \\ 40.19 & 38.02 & 41.03 & 38.23 \\ 25.03 & 22.02 & 27.03 & 22.95 \end{bmatrix}^T$$

$$[C] = \begin{bmatrix} 33.25 & 40.19 & 25.03 \\ 30.01 & 38.02 & 22.02 \\ 35.02 & 41.03 & 27.03 \\ 30.05 & 38.23 & 22.95 \end{bmatrix}$$

Recognize now that if we find [A] [C], we get

$$[D] = [A][C] = \begin{bmatrix} 25 & 20 & 3 & 2 \\ 5 & 10 & 15 & 25 \\ 6 & 16 & 7 & 27 \end{bmatrix} \begin{bmatrix} 33.25 & 40.19 & 25.03 \\ 30.01 & 38.02 & 22.02 \\ 35.02 & 41.03 & 27.03 \\ 30.05 & 38.23 & 22.95 \end{bmatrix}$$

$$= \begin{bmatrix} 1597 & 1965 & 1193 \\ 1743 & 2152 & 1325 \\ 1736 & 2169 & 1311 \end{bmatrix}$$

The diagonal elements give the sales of each brand of tire for the whole year, that is

d_{11} = \$1597 (Tirestone sales)

d_{22} = \$2152 (Michigan sales)

d_{33} = \$1311 (Cooper sales)

The total yearly sales of all three brands of tires are $= \sum_{i=1}^{3} d_{ii}$

= 1597 + 2152 + 1311

= \$5060

and this is the trace of the matrix [D].

Define the determinant of a matrix.

A determinant of a square matrix is a single unique real number corresponding to a matrix. For a matrix [A], determinant is denoted by |A| or det(A). So **do not** use [A] and |A| interchangeably.

For a 2 x 2 matrix,

$$[A] = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

det (A) = $a_{11}\, a_{22} - a_{12}\, a_{21}$

How does one calculate the determinant of any square matrix?

Let [A] be n x n matrix. The minor of entry a_{ij} is denoted by M_{ij} and is defined as the determinant of the (n - 1) x (n – 1) submatrix of [A], where the submatrix is obtained by deleting the i^{th} row and j^{th} column of the matrix [A]. The determinant is then given by

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} M_{ij} \quad \textit{for} \text{ any } i = 1, 2, \cdots, n$$

or

$$\det(\mathrm{A}) = \sum_{\mathrm{i}=1}^{\mathrm{n}} (-1)^{\mathrm{i}+\mathrm{j}} \mathrm{a}_{\mathrm{ij}} \mathrm{M}_{\mathrm{ij}} \text{ for any } \mathrm{j} = 1, 2, \cdots, \mathrm{n}$$

coupled with that

$\det(A) = a_{11} \;\; \textit{for a } 1 \times 1 \textit{ matrix } [A]$, as we can always reduce the determinant of a matrix to determinants of 1 x 1 matrices. The number $(-1)^{i+j} M_{ij}$ is called the **cofactor** of a_{ij} and is denoted by C_{ij}. The above equation for the determinant can then be written as

$$\det(A) = \sum_{j=1}^{n} a_{ij} C_{ij} \quad \textit{for any } i = 1, 2, \cdots, n$$

or

$$\det(A) = \sum_{i=1}^{n} a_{ij} C_{ij} \quad \textit{for any } j = 1, 2, \cdots, n$$

The only reason why determinants are not generally calculated using this method is that it becomes computationally intensive. For a n x n matrix, it requires arithmetic operations proportional to n!.

Example

Find the determinant of

$$[A] = \begin{bmatrix} 25 & 5 & 1 \\ 64 & 8 & 1 \\ 144 & 12 & 1 \end{bmatrix}$$

Solution

Method 1

$$\det(A) = \sum_{j=1}^{3} (-1)^{i+j} a_{ij} M_{ij} \quad \textit{for any } i = 1, 2, 3$$

Let i = 1 in the formula

$$\det(A) = \sum_{j=1}^{3} (-1)^{1+j} a_{1j} M_{1j}$$

$$= (-1)^{1+1} a_{11} M_{11} + (-1)^{1+2} a_{12} M_{12} + (-1)^{1+3} a_{13} M_{13}$$

$$= a_{11} M_{11} - a_{12} M_{12} + a_{13} M_{13}$$

$$M_{11} = \begin{vmatrix} 25 & 5 & 1 \\ 64 & 8 & 1 \\ 144 & 12 & 1 \end{vmatrix}$$

$$= \begin{vmatrix} 8 & 1 \\ 12 & 1 \end{vmatrix}$$

$$= -4$$

$$M_{12} = \begin{vmatrix} 25 & 5 & 1 \\ 64 & 8 & 1 \\ 144 & 12 & 1 \end{vmatrix}$$

$$= \begin{vmatrix} 64 & 1 \\ 144 & 1 \end{vmatrix}$$

$$= -80$$

$$M_{13} = \begin{vmatrix} 25 & 5 & 1 \\ 64 & 8 & 1 \\ 144 & 12 & 1 \end{vmatrix}$$

$$= \begin{vmatrix} 64 & 8 \\ 144 & 12 \end{vmatrix}$$

$$= -384$$

det(A) $= a_{11} M_{11} - a_{12} M_{12} + a_{13} M_{13} = 25(-4) - 5(-80) + 1(-384)$

$$= -100 + 400 - 384$$

$= -84$

Also for i=1,

$$\det(A) = \sum_{j=1}^{3} a_{1j} C_{1j}$$

$$C_{11} = (-1)^{1+1} M_{11}$$

$$= M_{11}$$

$$= -4$$

$$C_{12} = (-1)^{1+2} M_{12}$$

$$= -M_{21}$$

$$= 80$$

$$C_{13} = (-1)^{1+3} M_{13}$$

$$= M_{13}$$

$$= -384$$

$$\det(A) = a_{11}C_{11} + a_{21}C_{21} + a_{31}C_{31}$$

$$= (25)(-4) + (5)(80) + (1)(-384)$$

$$= -100 + 400 - 384$$

$$= -84$$

Method 2

$\det(A) = \sum_{i=1}^{3} (-1)^{i+j} a_{ij} M_{ij}$ for any j = 1,2,3.

Let j=2 in the formula

$$\det(A) = \sum_{i=1}^{3} (-1)^{i+2} a_{i2} M_{i2}$$

$$= (-1)^{1+2} a_{12} M_{12} + (-1)^{2+2} a_{22} M_{22} + (-1)^{3+2} a_{32} M_{32}$$

$$= -a_{12} M_{12} + a_{22} M_{22} - a_{32} M_{32}$$

$$M_{12} = \begin{vmatrix} 25 & 5 & 1 \\ 64 & 8 & 1 \\ 144 & 12 & 1 \end{vmatrix}$$

$$= \begin{vmatrix} 64 & 1 \\ 144 & 1 \end{vmatrix}$$

$$= -80$$

$$M_{22} = \begin{vmatrix} 25 & 5 & 1 \\ 64 & 8 & 1 \\ 144 & 12 & 1 \end{vmatrix}$$

$$= \begin{vmatrix} 25 & 1 \\ 144 & 1 \end{vmatrix}$$

$$= -119$$

$$M_{32} = \begin{vmatrix} 25 & 5 & 1 \\ 64 & 8 & 1 \\ 144 & 12 & 1 \end{vmatrix}$$

$$= \begin{vmatrix} 25 & 1 \\ 64 & 1 \end{vmatrix}$$

$$= -39$$

det(A) $= -a_{12}M_{12} + a_{22}M_{22} - a_{32}M_{32}$ = -5(-80) + 8(-119) – 12(-39)

= 400 – 952 + 468

= -84.

In terms of cofactors for j=1,

$$\det(A) = \sum_{j=1}^{3} a_{i2}C_{i2}$$

$$C_{12} = (-1)^{1+2} M_{12}$$

$$= -M_{12}$$

$$= 80$$

$$C_{22} = (-1)^{2+2} M_{22}$$

$$= M_{22}$$

$$= -119$$

$$C_{32} = (-1)^{3+2} M_{32}$$

$$= -M_{32}$$

$$= 39$$

$$\det(A) = a_{12}C_{12} + a_{22}C_{22} + a_{32}C_{32}$$

$$= (5)(80) + (8)(-119) + (12)(39)$$

$$= 400 - 952 + 468 = -84$$

Is there a relationship between det (AB), and det (A) and det (B)?

Yes, if [A] and [B] are square matrices of same size, then

det (AB) = det (A) det (B).

Are there some other theorems that are important in finding the determinant?

<u>Theorem 1</u>: If a row or a column in a nxn matrix [A] is zero, then det (A) =0

<u>Theorem 2</u>: Let [A] be a nxn matrix. If a row is proportional to another row, then det(A) = 0.

<u>Theorem 3</u>: Let [A] be a nxn matrix. If a column is proportional to another column, then det (A) = 0

<u>Theorem 4</u>: Let [A] be a nxn matrix. If a column or row is multiplied by k to result in matrix [B]. Then det(B)=k det(A).

Example

What is the determinant of

$$[A]=\begin{bmatrix}0 & 2 & 6 & 3\\0 & 3 & 7 & 4\\0 & 4 & 9 & 5\\0 & 5 & 2 & 1\end{bmatrix}$$

Solution

Since one of the columns (first column in the above example) of [A] is a zero, det(A) = 0.

Example

What is the determinant of

$$[A]=\begin{bmatrix}2 & 1 & 6 & 4\\3 & 2 & 7 & 6\\5 & 4 & 2 & 10\\9 & 5 & 3 & 18\end{bmatrix}$$

Solution

det(A) is zero because the fourth column

$$\begin{bmatrix}4\\6\\10\\18\end{bmatrix}$$

is 2 times the first column

$$\begin{bmatrix}2\\3\\5\\9\end{bmatrix}$$

Example

If the determinant of

$$[A] = \begin{bmatrix} 25 & 5 & 1 \\ 64 & 8 & 1 \\ 144 & 12 & 1 \end{bmatrix}$$

is –84, then what is the determinant of

$$[B] = \begin{bmatrix} 25 & 10.5 & 1 \\ 64 & 16.8 & 1 \\ 144 & 24.2 & 1 \end{bmatrix}$$

Solution

Since the second column of [B] is 2.1 times the second column of [A],

$$\begin{aligned} \det(B) &= 2.1 \det(A) \\ &= (2.1)(-84) \\ &= -176.4 \end{aligned}$$

Example

Given the determinant of

$$[A] = \begin{bmatrix} 25 & 5 & 1 \\ 64 & 8 & 1 \\ 144 & 12 & 1 \end{bmatrix}$$

is -84, what is the determinant of

$$[B] = \begin{bmatrix} 25 & 5 & 1 \\ 0 & -4.8 & -1.56 \\ 144 & 12 & 1 \end{bmatrix}$$

Solution

Since [B] is simply obtained by subtracting the second row of [A] by 2.56 times the first row of [A], det(B) = det(A) = -84.

Example

What is the determinant of

$$[A]=\begin{bmatrix}25 & 5 & 1\\ 0 & -4.8 & -1.56\\ 0 & 0 & 0.7\end{bmatrix}$$

Solution

Since [A] is an upper triangular matrix

$$\det(A)=\prod_{i=1}^{3}a_{ii}$$

$$=(a_{11})(a_{22})(a_{33})$$

$$=(25)(-4.8)(0.7)$$

$$= -84.$$

Key Terms

Transpose Symmetric Skew symmetric Trace Determinant

Homework

1. Let [A] = $\begin{bmatrix}25 & 3 & 6\\ 7 & 9 & 2\end{bmatrix}$. Find $[A]^T$

Answer: $\begin{bmatrix}25 & 7\\ 3 & 9\\ 6 & 2\end{bmatrix}$

2. If [A] and [B] are two nxn symmetric matrices, show that [A]+[B] is also symmetric.

Hint: Let $[C]=[A]+[B]$

$c_{ij} = a_{ij} + b_{ij}$ **for all i, j.**

and $c_{ji} = a_{ji} + b_{ji}$ **for all i, j.**

$c_{ji} = a_{ij} + b_{ij}$ **as** $[A]$ **and** $[B]$ **are symmetric**

Hence

$c_{ji} = c_{ij}$.

3. Give an example of a 4x4 symmetric matrix.

4. What is the trace of

$$[A]=\begin{bmatrix} 7 & 2 & 3 & 4 \\ -5 & -5 & -5 & -5 \\ 6 & 6 & 7 & 9 \\ -5 & 2 & 3 & 10 \end{bmatrix}$$

Answer: 19

5. For

$$[\mathrm{A}]=\begin{bmatrix} 10 & -7 & 0 \\ -3 & 2.099 & 6 \\ 5 & -1 & 5 \end{bmatrix}$$

Find the determinant of [A] using the cofactor method.

Answer: -150.05

6. det (3[A]) of a nxn matrix is

   a. 3 det (A)

   b. det (A)

c. 3^n det (A)

d. 9 det (A)

7. For a 5x5 matrix [A], the first row is interchanged with the fifth row, the determinant of the resulting matrix [B] is

A. –det (A)

B. det (A)

C. 5 det (A)

D. 2 det (A)

8. $\det\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}$ is

A. 0

B. 1

C. –1

D. ∞

9. If [A] is a nxn matrix and is invertible, then $\det(A^T)$ is equal to

A. $\det(A^{-1})$

B. det(A)

C. –det(A)

D. $-\det(A^{-1})$

10. Without using the cofactor method of finding determinants, find the determinant of

$$[A]=\begin{bmatrix} 0 & 0 & 0 \\ 2 & 3 & 5 \\ 6 & 9 & 2 \end{bmatrix}$$

11. Without using the cofactor method of finding determinants, find the determinant of

$$[A] = \begin{bmatrix} 0 & 0 & 2 & 3 \\ 0 & 2 & 3 & 5 \\ 6 & 7 & 2 & 3 \\ 6.6 & 7.7 & 2.2 & 3.3 \end{bmatrix}$$

12. Without using the cofactor method of finding determinants, find the determinant of

$$[A] = \begin{bmatrix} 5 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 2 & 5 & 6 & 0 \\ 1 & 2 & 3 & 9 \end{bmatrix}$$

13. Given the matrix

$$[A] = \begin{bmatrix} 125 & 25 & 5 & 1 \\ 512 & 64 & 8 & 1 \\ 1157 & 89 & 13 & 1 \\ 8 & 4 & 2 & 1 \end{bmatrix}$$

and

det(A) =-32400

find the determinant of

a) $[A] = \begin{bmatrix} 125 & 25 & 5 & 1 \\ 512 & 64 & 8 & 1 \\ 1141 & 81 & 9 & -1 \\ 8 & 4 & 2 & 1 \end{bmatrix}$

b) $[B] = \begin{bmatrix} 125 & 25 & 5 & 1 \\ 1157 & 89 & 13 & 1 \\ 512 & 64 & 8 & 1 \\ 8 & 4 & 2 & 1 \end{bmatrix}$

c) $[C] = \begin{bmatrix} 125 & 25 & 5 & 1 \\ 1157 & 89 & 13 & 1 \\ 8 & 4 & 2 & 1 \\ 512 & 64 & 8 & 1 \end{bmatrix}$

d) $[D]=\begin{bmatrix} 125 & 25 & 5 & 1 \\ 512 & 64 & 8 & 1 \\ 1157 & 89 & 13 & 1 \\ 16 & 8 & 4 & 2 \end{bmatrix}$

Answer:a) –32400 b) 32400 c) –32400 d) –64800

Chapter 5
System of Equations

After reading this chapter, you will be able to

- *Setup simultaneous linear equations in matrix form and vice-versa*
- *Understand the concept of inverse of a matrix*
- *Know the difference between consistent and inconsistent system of linear equations*
- *Learn that system of linear equations can have a unique solution, no solution or infinite solutions*

Matrix algebra is used for solving system of equations. Can you illustrate this concept?

Matrix algebra is used to solve a system of simultaneous linear equations. In fact, for many mathematical procedures such as solution of set of nonlinear equations, interpolation, integration, and differential equations, the solutions reduce to a set of simultaneous linear equations. Let us illustrate with an example.

Example

The upward velocity of a rocket is given at three different times on the following table

Time, t	**Velocity, v**
s	m/s
5	106.8

8	177.2
12	279.2

The velocity data is approximated by a polynomial as

$v(t) = at^2 + bt + c$, $5 \le \mathrm{t} \le 12.$

Set up the equations in matrix form to find the coefficients a, b, c of the velocity profile.

Solution

The polynomial is going through three data points $(t_1, v_1), (t_2, v_2),$ and (t_3, v_3) where from the above table

$t_1 = 5, v_1 = 106.8$

$t_2 = 8, v_2 = 177.2$

$t_3 = 12, v_3 = 279.2$

Requiring that $v(t) = at^2 + bt + c$ passes through the three data points gives

$v(t_1) = v_1 = at_1^2 + bt_1 + c$

$v(t_2) = v_2 = at_2^2 + bt_2 + c$

$v(t_3) = v_3 = at_3^2 + bt_3 + c$

Substituting the data $(t_1, v_1), (t_2, v_2), (t_3, v_3)$ gives

$a(5^2) + b(5) + c = 106.8$

$a(8^2) + b(8) + c = 177.2$

$a(12^2) + b(12) + c = 279.2$

or

$25a + 5b + c = 106.8$

$64a + 8b + c = 177.2$

$144a + 12b + c = 279.2$

This set of equations can be rewritten in the matrix form as

$$\begin{bmatrix} 25a + & 5b + & c \\ 64a + & 8b + & c \\ 144a + & 12b + & c \end{bmatrix} = \begin{bmatrix} 106.8 \\ 177.2 \\ 279.2 \end{bmatrix}$$

The above equation can be written as a linear combination as follows

$$a\begin{bmatrix} 25 \\ 64 \\ 144 \end{bmatrix} + b\begin{bmatrix} 5 \\ 8 \\ 12 \end{bmatrix} + c\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 106.8 \\ 177.2 \\ 279.2 \end{bmatrix}$$

and further using matrix multiplications gives

$$\begin{bmatrix} 25 & 5 & 1 \\ 64 & 8 & 1 \\ 144 & 12 & 1 \end{bmatrix}\begin{bmatrix} \mathrm{a} \\ \mathrm{b} \\ \mathrm{c} \end{bmatrix} = \begin{bmatrix} 106.8 \\ 177.2 \\ 279.2 \end{bmatrix}$$

The above is an illustration of why matrix algebra is needed. The complete solution to the set of equations is given later in this chapter.

For a general set of "m" linear equations and "n" unknowns,

$$a_{11}x_1 + a_{22}x_2 + \cdots\cdots + a_{1n}x_n = c_1$$

$$a_{21}x_1 + a_{22}x_2 + \cdots\cdots + a_{2n}x_n = c_2$$

……………………………………

…………………………………….

$$a_{m1}x_1 + a_{m2}x_2 + \dots\dots\dots + a_{mn}x_n = c_m$$

can be rewritten in the matrix form as

$$\begin{bmatrix} a_{11} & a_{12} & . & . & a_{1n} \\ a_{21} & a_{22} & . & . & a_{2n} \\ \vdots & & & & \vdots \\ \vdots & & & & \vdots \\ a_{m1} & a_{m2} & . & . & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ . \\ . \\ x_n \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \\ . \\ . \\ c_m \end{bmatrix}$$

Denoting the matrices by $[A]$, $[X]$, and $[C]$, the system of equation is

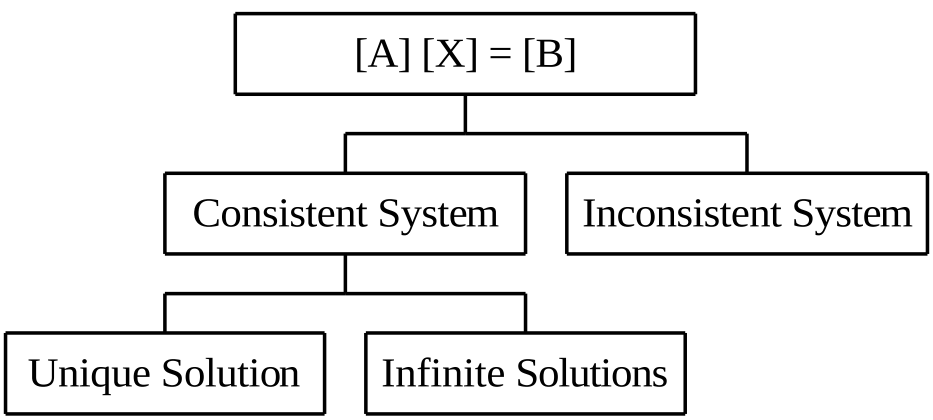
$[A]\ [X]=[C]$, where $[A]$ is called the coefficient matrix, $[C]$ is called the right hand side vector and $[X]$ is called the solution vector.

Sometimes $[A]\ [X]=[C]$ systems of equations is written in the augmented form. That is

$$[\mathrm{A} \vdots \mathrm{C}] = \left[\begin{array}{cccc:c} \mathrm{a}_{11} & \mathrm{a}_{12} & \ldots\ldots & \mathrm{a}_{1\mathrm{n}} & c_1 \\ \mathrm{a}_{21} & \mathrm{a}_{22} & \ldots\ldots & \mathrm{a}_{2\mathrm{n}} & c_2 \\ \vdots & & & \vdots & \\ \vdots & & & \vdots & \\ \mathrm{a}_{\mathrm{m}1} & \mathrm{a}_{\mathrm{m}2} & \ldots\ldots & a_{mn} & c_n \end{array}\right]$$

A system of equations can be consistent or inconsistent. What does that mean?

A system of equations [A] [X]=[C] is consistent if there is a solution, and it is inconsistent if there is no solution. However, consistent system of equations does not mean a unique solution, that is, a consistent system of equation may have a unique solution or infinite solutions.



Example

Give examples of consistent and inconsistent system of equations.

Solution

a) The system of equations

$$\begin{bmatrix} 2 & 4 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 6 \\ 4 \end{bmatrix}$$

is a consistent system of equations as it has a unique solution, that is,

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

b) The system of equations

$$\begin{bmatrix} 2 & 4 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 6 \\ 3 \end{bmatrix}$$

is also a consistent system of equations but it has infinite solutions as given as follows.

Expanding the above set of equations,

$$2x + 4y = 6$$
$$x + 2y = 3$$

you can see that they are the same equation. Hence any combination of (x, y) that satisfies

$$2x + 4y = 6$$

is a solution. For example $(x, y) = (1,1)$ is a solution and other solutions include $(x, y) = (0.5,1.25)$, $(x, y) = (0,\ 1.5)$ and so on.

c) The system of equations

$$\begin{bmatrix} 2 & 4 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 6 \\ 4 \end{bmatrix}$$

is inconsistent as no solution exists.

How can one distinguish between a consistent and inconsistent system of equations?

A system of equations [A] [X] = [C] is

- consistent if the rank of A is equal to the rank of the augmented matrix $[A \vdots C]$,
- inconsistent if the rank of A is less then the rank of the augmented matrix $[A \vdots C]$.

But, what do you mean by rank of a matrix?

The rank of a matrix is defined as the order of the largest square submatrix whose determinant is <u>not</u> zero.

Example

What is the rank of

$$[A]=\begin{bmatrix} 3 & 1 & 2 \\ 2 & 0 & 5 \\ 1 & 2 & 3 \end{bmatrix}$$

Solution

The largest square submatrix possible is of order 3 and that is [A] itself. Since det(A) = -25 $\neq$ 0, the rank of [A] = 3.

Example:

What is the rank of

$$[A]=\begin{bmatrix} 3 & 1 & 2 \\ 2 & 0 & 5 \\ 5 & 1 & 7 \end{bmatrix}$$

Solution:

The largest square submatrix of [A] is of order 3, and is [A] itself. Since det(A) = 0, the rank of [A] is less than 3. The next largest submatrix would be a 2x2 matrix. One of the submatrices of [A] is

$$[B]=\begin{bmatrix}3 & 1\\ 2 & 0\end{bmatrix}$$

and det [B] = -2 $\neq$ 0. Hence the rank of [A] is 2.

Example

How do I now use the concept of rank to find if

$$\begin{bmatrix}25 & 5 & 1\\ 64 & 8 & 1\\ 144 & 12 & 1\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix}=\begin{bmatrix}106.8\\ 177.2\\ 279.2\end{bmatrix}$$

is a consistent or inconsistent system of equations?

Solution

The coefficient matrix is

$$[A]=\begin{bmatrix}25 & 5 & 1\\ 64 & 8 & 1\\ 144 & 12 & 1\end{bmatrix}$$

and the right hand side vector

$$[C]=\begin{bmatrix}106.8\\ 177.2\\ 279.2\end{bmatrix}.$$

The augmented matrix is

$$[B]=\begin{bmatrix}25 & 5 & 1 & \vdots & 106.8\\ 64 & 8 & 1 & \vdots & 177.2\\ 144 & 12 & 1 & \vdots & 279.2\end{bmatrix}$$

Since there are no square submatrices of order 4 as [B] is a 4x3 matrix, the rank of [B] is at most 3. So let us look at the square submatrices of [B] of order 3 and if any of these square submatrices have determinant not equal to zero, then the rank is 3. For example, a submatrix of the augmented matrix [B] is

$$[D] = \begin{bmatrix} 25 & 5 & 1 \\ 64 & 8 & 1 \\ 144 & 12 & 1 \end{bmatrix}$$

has det (D) = -84 $\neq$ 0.

Hence the rank of the augmented matrix [B] is 3. Since [A]=[D] , the rank of [A] =3. Since the rank of augmented matrix [B] = rank of coefficient matrix [A], the system of equations is consistent.

Example

Use the concept of rank of matrix to find if

$$\begin{bmatrix} 25 & 5 & 1 \\ 64 & 8 & 1 \\ 89 & 13 & 2 \end{bmatrix} \begin{bmatrix} \mathrm{x}_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 106.8 \\ 177.2 \\ 284.0 \end{bmatrix}$$

is consistent or inconsistent?

Solution

The coefficient matrix is given by

$$[A] = \begin{bmatrix} 25 & 5 & 1 \\ 64 & 8 & 1 \\ 89 & 13 & 2 \end{bmatrix}$$

and the right hand side

$$[C] = \begin{bmatrix} 106.8 \\ 177.2 \\ 284.0 \end{bmatrix}$$

The augmented matrix is

$$[B]=\begin{bmatrix}25 & 5 & 1 & 106.8\\64 & 8 & 1 & 177.2\\89 & 13 & 2 & 284.0\end{bmatrix}$$

Since there are no square submatrices of order 4 as [B] is a 4x3 matrix, the rank of the augmented [B] is at most 3. So let us look at square submatrices of the augmented matrix [B] of order 3 and see if any of these submatrices have determinant not equal to zero, then the rank is 3. For example a submatrix of the augmented matrix [B] is

$$[D]=\begin{bmatrix}25 & 5 & 1\\64 & 8 & 1\\89 & 13 & 2\end{bmatrix}$$

has det(D) = 0. This means, we need to explore other square submatrices of order 3 of the augmented matrix [B].

That is,

$$[E]=\begin{bmatrix}5 & 1 & 106.8\\8 & 1 & 177.2\\13 & 2 & 284.0\end{bmatrix}$$

det(E) = 0,

$$[F]=\begin{bmatrix}25 & 5 & 106.8\\64 & 8 & 177.2\\89 & 13 & 284.0\end{bmatrix}$$

det(F) = 0,

and

$$[G]=\begin{bmatrix}25 & 1 & 106.8\\64 & 1 & 177.2\\89 & 2 & 284.0\end{bmatrix}$$

det(G) = 0.

All the square submatrices of order 3 of the augmented matrix [B] have a zero determinant. So the rank of the augmented matrix [B] is less than 3. Is the rank of [B] = 2? A 2x2 submatrix of the augmented matrix [B] is

$$[H]=\begin{bmatrix} 25 & 5 \\ 64 & 8 \end{bmatrix}$$

and

det(H) = -120?0

So the rank of the augmented matrix [B] is 2.

Now we need to find the rank of the coefficient matrix [A].

$$[A]=\begin{bmatrix} 25 & 5 & 1 \\ 64 & 8 & 1 \\ 89 & 13 & 2 \end{bmatrix}$$

and

det(A) = 0.

So the rank of the coefficient matrix [A] is less than 3. A submatrix of the coefficient matrix [A] is

$$[J]=\begin{bmatrix} 5 & 1 \\ 8 & 1 \end{bmatrix}$$

det (J) = -3 $\neq$ 0

So the rank of the coefficient matrix [A] is 2.

Hence, rank of the coefficient matrix [A] = rank of the augmented matrix [B]. So the system of equations [A] [X] = [B] is consistent.

Example

Use the concept of rank to find if

$$\begin{bmatrix} 25 & 5 & 1 \\ 64 & 8 & 1 \\ 89 & 13 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 106.8 \\ 177.2 \\ 280.0 \end{bmatrix}$$

is consistent or inconsistent.

Solution

The augmented matrix is

$$[B] = \begin{bmatrix} 25 & 5 & 1 & 106.8 \\ 64 & 8 & 4 & 177.2 \\ 89 & 13 & 2 & 280.0 \end{bmatrix}$$

Since there are no square submatrices of order 4 as the augmented matrix [B] is a 4x3 matrix, the rank of the augmented matrix [B] is at most 3. So let us look at square submatrices of the augmented matrix (B) of order 3 and see if any of the 3x3 submatrices have a determinant not equal to zero. For example a submatrix of order 3 of [B]

$$[D] = \begin{bmatrix} 25 & 5 & 1 \\ 64 & 8 & 1 \\ 89 & 13 & 2 \end{bmatrix}$$

det(D) = 0

So it means, we need to explore other square submatrices of the augmented matrix [B],

$$[E] = \begin{bmatrix} 5 & 1 & 106.8 \\ 8 & 1 & 177.2 \\ 13 & 2 & 280.0 \end{bmatrix}$$

det(E) $\neq$ 12.0 $\neq$ 0.

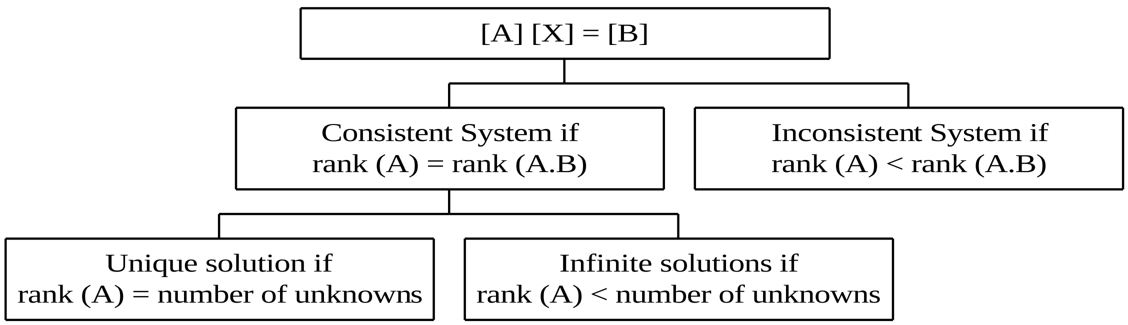
So rank of the augmented matrix [B] = 3.

The rank of the coefficient matrix [A] = 2 from the previous example.

Since rank of the coefficient matrix [A] < rank of the augmented matrix [B], the system of equations is inconsistent. Hence no solution exists for [A] [X] = [C].

If a solution exists, how do we know whether it is unique?

In a system of equations [A] [X] = [C] that is consistent, the rank of the coefficient matrix [A] is same as the augmented matrix [A|C]. If in addition, the rank of the coefficient matrix [A] is same as the number of unknowns, then the solution is unique; if the rank of the coefficient matrix [A] is less than the number of unknowns, then infinite solutions exist.



Example:

We found that the following system of equations

$$\begin{bmatrix} 25 & 5 & 1 \\ 64 & 8 & 1 \\ 144 & 12 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 106.8 \\ 177.2 \\ 279.2 \end{bmatrix}$$

is a consistent system of equations. Does the system of equations have a unique solution or does it have infinite solutions.

Solution

The coefficient matrix is

$$[A] = \begin{bmatrix} 25 & 5 & 1 \\ 64 & 8 & 1 \\ 144 & 12 & 1 \end{bmatrix}$$

and the right hand side

$$[C] = \begin{bmatrix} 106.8 \\ 177.2 \\ 279.2 \end{bmatrix}$$

While finding the whether the above equations were consistent in an earlier example, we found that

rank of the coefficient matrix (A) = rank of augmented matrix $[A \vdots C] = 3$

The solution is unique as the number of unknowns = 3 = rank of (A).

Example:

We found that the following system of equations

$$\begin{bmatrix} 25 & 5 & 1 \\ 64 & 8 & 1 \\ 89 & 13 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 106.8 \\ 177.2 \\ 284.0 \end{bmatrix}$$

is a consistent system of equations. Is the solution unique or does it have infinite solutions.

Solution

While finding the whether the above equations were consistent, we found that

rank of coefficient matrix [A] = rank of augmented matrix $(A \vdots C) = 2$

Since rank of [A] = 2 < number of unknowns =3, infinite solutions exist.

If we have more equations than unknowns in [A] [X] = [C], does it mean the system is inconsistent?

No, it depends on the rank of the augmented matrix $[A \vdots C]$ and the rank of [A].

a) For example

$$\begin{bmatrix} 25 & 5 & 1 \\ 64 & 8 & 1 \\ 144 & 12 & 1 \\ 89 & 13 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 106.8 \\ 177.2 \\ 279.2 \\ 284.0 \end{bmatrix}$$

is consistent, since

rank of augmented matrix = 3

rank of coefficient matrix = 3.

Now since rank of (A) = 3 = number of unknowns, the solution is not only consistent but also unique.

b) For example

$$\begin{bmatrix} 25 & 5 & 1 \\ 64 & 8 & 1 \\ 144 & 12 & 1 \\ 89 & 13 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 106.8 \\ 177.2 \\ 279.2 \\ 280.0 \end{bmatrix}$$

is inconsistent, since

rank of augmented matrix = 4

rank of coefficient matrix = 3.

c) For example

$$\begin{bmatrix} 25 & 5 & 1 \\ 64 & 8 & 1 \\ 50 & 10 & 2 \\ 89 & 13 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 106.8 \\ 177.2 \\ 213.6 \\ 280.0 \end{bmatrix}$$

is consistent, since

rank of augmented matrix = 2

rank of coefficient matrix = 2.

But since the rank of [A] = 2 < the number of unknowns = 3, infinite solutions exist.

Consistent system of equations can only have a unique solution or infinite solutions. Can a system of equations have a finite (more than one but not infinite) number of solutions?

No, you can only have a unique solution or infinite solutions. Let us suppose [A] [X]=[C] has two solutions [Y] and [Z] so that

[A] [Y]=[C]

[A] [Z]=[C]

If r is a constant, then from the two equations

$$r[A][\mathrm{Y}] = r[C]$$

$$(1-r)[A][\mathrm{Z}] = (1-r)[C]$$

Adding the above two equations gives

$$r[A][\mathrm{Y}] + (1-r)[A][\mathrm{Z}] = r[C] + (1-r)[C]$$

$$[A](r[Y] + (1-r)[Z]) = [C]$$

Hence

$$r[Y] + (1-r)[Z]$$

is a solution to

$$[A][\mathrm{X}] = [C].$$

Since r is any scalar, there are infinite solutions for [A] [X] =[C] of the form r[Y]+(1-r)[Z].

Can you divide two matrices?

If [A] [B]=[C] is defined, it might seem intuitive that [A]= $\frac{[C]}{[B]}$, but matrix division is not defined. However an inverse of a matrix can be defined for certain types of square matrices. The inverse of a square matrix [A], if existing, is denoted by $[A]^{-1}$ such that $[A][A]^{-1} = [I] = [A]^{-1}[A]$.

In other words, let [A] be a square matrix. If [B] is another square matrix of same size such that [B][A] = [I], then [B] is the inverse of [A]. [A] is then called to be **invertible** or **nonsingular**. If $[A]^{-1}$ does not exist, [A] is called to be **noninvertible** or **singular**.

If [A] and [B] are two nxn matrices such that [B] [A] = [I], then these statements are also true

a) [B] is the inverse of [A]

b) [A] is the inverse of [B]

c) [A] and [B] are both invertible

d) [A] [B]=[I].

e) [A] and [B] are both nonsingular

f) all columns of [A] or [B]are linearly independent

g) all rows of [A] or [B] are linearly independent.

Example

Show if

$$[B] = \begin{bmatrix} 3 & 2 \\ 5 & 3 \end{bmatrix} \text{ is the inverse of } [\mathrm{A}] = \begin{bmatrix} -3 & 2 \\ 5 & -3 \end{bmatrix}$$

Solution

$[B][A]$

$$= \begin{bmatrix} 3 & 2 \\ 5 & 3 \end{bmatrix} \begin{bmatrix} -3 & 2 \\ 5 & -3 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

= [I]

Since [B] [A] =[I], [B] is the inverse of [A] and [A] is the inverse of [B]. But we can also show that

$$[A][B]$$

$$= \begin{bmatrix} -3 & 2 \\ 5 & -3 \end{bmatrix} \begin{bmatrix} 3 & 2 \\ 5 & 3 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = [\mathrm{I}]$$

to show that [A] is the inverse of [B].

Can I use the concept of the inverse of a matrix to find the solution of a set of equations [A] [X] = [C]?

Yes, if the number of equations is same as the number of unknowns, the coefficient matrix [A] is a square matrix.

Given

[A][X] = [C]

Then, if $[A]^{-1}$ exists, multiplying both sides by $[A]^{-1}$.

$[A]^{-1}$ [A][X] = $[A]^{-1}$ [C]

[I][X] = $[A]^{-1}$[C]

[X] = $[A]^{-1}$ [C]

This implies that if we are able to find $[A]^{-1}$, the solution vector of [A][X] = [C] is simply a multiplication of $[A]^{-1}$ and the right hand side vector, [C].

How do I find the inverse of a matrix?

If [A] is a n x n matrix, then $[A]^{-1}$ is a n x n matrix and according to the definition of inverse of a matrix

$[A][A]^{-1}$ = [I].

Denoting

$$[A] = \begin{bmatrix} a_{11} & a_{12} & \cdot & \cdot & a_{1n} \\ a_{21} & a_{22} & \cdot & \cdot & a_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \cdot & \cdot & a_{nn} \end{bmatrix}$$

$$[A]^{-1} = \begin{bmatrix} a'_{11} & a'_{12} & \cdot & \cdot & a'_{1n} \\ a'_{21} & a'_{22} & \cdot & \cdot & a'_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a'_{n1} & a'_{n2} & \cdot & \cdot & a'_{nn} \end{bmatrix}$$

$$[I] = \begin{bmatrix} 1 & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & 1 & & & & 0 \\ 0 & & \cdot & & & \cdot \\ \cdot & & & 1 & & \cdot \\ \cdot & & & & \cdot & \cdot \\ 0 & \cdot & \cdot & \cdot & \cdot & 1 \end{bmatrix}$$

Using the definition of matrix multiplication, the first column of the $[A]^{-1}$ matrix can then be found by solving

$$\begin{bmatrix} a_{11} & a_{12} & \cdot & \cdot & a_{1n} \\ a_{21} & a_{22} & \cdot & \cdot & a_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \cdot & \cdot & a_{nn} \end{bmatrix} \begin{bmatrix} a'_{11} \\ a'_{21} \\ \cdot \\ \cdot \\ a'_{n1} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ \cdot \\ \cdot \\ 0 \end{bmatrix}$$

Similarly, one can find the other columns of the $[A]^{-1}$ matrix by changing the right hand side accordingly.

Example

The upward velocity of the rocket is given by

Time, t	Velocity
s	**m/s**
5	106.8
8	177.2
12	279.2

In an earlier example, we wanted to approximate the velocity profile by $v(t) = at^2 + bt + c, \quad 5 \le 8 \le 12$

We found that the coefficients a, b, c are given by

$$\begin{bmatrix} 25 & 5 & 1 \\ 64 & 8 & 1 \\ 144 & 12 & 1 \end{bmatrix} \begin{bmatrix} \mathrm{a} \\ \mathrm{b} \\ \mathrm{c} \end{bmatrix} = \begin{bmatrix} 106.8 \\ 177.2 \\ 279.2 \end{bmatrix}$$

First find the inverse of

$$[A] = \begin{bmatrix} 25 & 5 & 1 \\ 64 & 8 & 1 \\ 144 & 12 & 1 \end{bmatrix}$$

and then use the definition of inverse to find the coefficients $a, b, c.$

Solution

If $[A]^{-1} = \begin{bmatrix} a'_{11} & a'_{12} & a'_{13} \\ a'_{21} & a'_{22} & a'_{23} \\ a'_{31} & a'_{32} & a'_{33} \end{bmatrix}$

is the inverse of [A],

Then

$$\begin{bmatrix} 25 & 5 & 1 \\ 64 & 8 & 1 \\ 144 & 12 & 1 \end{bmatrix} \begin{bmatrix} a'_{11} & a'_{12} & a'_{13} \\ a'_{21} & a'_{22} & a'_{23} \\ a'_{31} & a'_{32} & a'_{33} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

gives three sets of equations

$$\begin{bmatrix} 25 & 5 & 1 \\ 64 & 8 & 1 \\ 144 & 12 & 1 \end{bmatrix} \begin{bmatrix} a'_{11} \\ a'_{21} \\ a'_{31} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} 25 & 5 & 1 \\ 64 & 8 & 1 \\ 144 & 12 & 1 \end{bmatrix} \begin{bmatrix} a'_{12} \\ a'_{22} \\ a'_{32} \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} 25 & 5 & 1 \\ 64 & 8 & 1 \\ 144 & 12 & 1 \end{bmatrix} \begin{bmatrix} a'_{13} \\ a'_{23} \\ a'_{33} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Solving the above three sets of equations separately gives

$$\begin{bmatrix} a'_{11} \\ a'_{21} \\ a'_{31} \end{bmatrix} = \begin{bmatrix} 0.04762 \\ -0.9524 \\ 4.571 \end{bmatrix}$$

$$\begin{bmatrix} a'_{12} \\ a'_{22} \\ a'_{32} \end{bmatrix} = \begin{bmatrix} -0.08333 \\ 1.417 \\ -5.000 \end{bmatrix}$$

$$\begin{bmatrix} a'_{13} \\ a'_{23} \\ a'_{33} \end{bmatrix} = \begin{bmatrix} 0.03571 \\ -0.4643 \\ 1.429 \end{bmatrix}$$

Hence

$$[A]^{-1} = \begin{bmatrix} 0.04762 & -0.08333 & 0.03571 \\ -0.9524 & 1.417 & -0.4643 \\ 4.571 & -5.000 & 1.429 \end{bmatrix}$$

Now

$$[A][\mathrm{X}]=[C]$$

where

$$[X]=\begin{bmatrix} a \\ b \\ c \end{bmatrix}$$

$$[C]=\begin{bmatrix} 106.8 \\ 177.2 \\ 279.2 \end{bmatrix}$$

Using the definition of $[A]^{-1}$,

$$[A]^{-1}[\mathrm{A}][\mathrm{X}]=[A]^{-1}[\mathrm{C}]$$

$$[\mathrm{X}]=[A]^{-1}[\mathrm{C}]$$

$$=\begin{bmatrix} -0.04762 & -0.08333 & 0.03571 \\ -0.9524 & 1.417 & -0.4643 \\ 4.571 & -5.000 & 1.429 \end{bmatrix}\begin{bmatrix} 106.8 \\ 177.2 \\ 279.2 \end{bmatrix}$$

$$\begin{bmatrix} \mathrm{a} \\ \mathrm{b} \\ \mathrm{c} \end{bmatrix}=\begin{bmatrix} 0.2900 \\ 19.70 \\ 1.050 \end{bmatrix}$$

So

$$v(t)=0.2900t^2+19.70t+1.050,\ 5\le \mathrm{t}\le 12$$

Is there another way to find the inverse of a matrix?

For finding inverse of small matrices, inverse of an invertible matrix can be found by

$$[A]^{-1}=\frac{1}{\det(A)}adj(A)$$

where

$$adj(A) = \begin{bmatrix} C_{11} & C_{12} & \cdots & C_{1n} \\ C_{21} & C_{22} & & C_{2n} \\ \vdots & & & \\ C_{n1} & C_{n2} & - & C_{nn} \end{bmatrix}^T$$

where C_{ij} are the cofactors of a_{ij}. The matrix $\begin{bmatrix} C_{11} & C_{12} & \cdots & C_{1n} \\ C_{22} & C_{22} & \cdots & C_{2n} \\ \vdots & & & \vdots \\ C_{n1} & \cdots & \cdots & C_{nn} \end{bmatrix}$ itself is called the matrix of cofactors from [A]. Cofactors are defined in Chapter 4.

Example

Find the inverse of

$$[A] = \begin{bmatrix} 25 & 5 & 1 \\ 64 & 8 & 1 \\ 144 & 12 & 1 \end{bmatrix}$$

Solution

From the example in Chapter 4, we found

$$\det(A) = -84.$$

Next we need to find the adjoint of [A]. The cofactors of A are found as follows.

The minor of entry a_{11} is

$$M_{11} = \begin{vmatrix} 25 & 5 & 1 \\ 64 & 8 & 1 \\ 144 & 12 & 1 \end{vmatrix}$$

$$= \begin{vmatrix} 8 & 1 \\ 12 & 1 \end{vmatrix}$$

$$= -4$$

The cofactors of entry a_{11} is

$$C_{11} = (-1)^{1+1} M_{11} = M_{11} = -4$$

The minor of entry a_{12} is

$$M_{12} = \begin{vmatrix} 25 & 5 & 1 \\ 64 & 8 & 1 \\ 144 & 12 & 1 \end{vmatrix}$$

$$= \begin{vmatrix} 64 & 1 \\ 144 & 1 \end{vmatrix}$$

$$= -80$$

The cofactor of entry a_{12} is

$$C_{12} = (-1)^{1+2} M_{12}$$

$$= -M_{12}$$

$$= 80$$

Similarly

$$C_{13} = 384$$

$$C_{21} = 7$$

$$C_{22} = -129$$

$$C_{23} = 420$$

$$C_{31} = -3$$

$$C_{32} = 39$$

$$C_{33} = -120$$

Hence the matrix of cofactors of [A] is

$$[C] = \begin{bmatrix} -4 & 80 & -384 \\ 7 & -129 & 420 \\ -3 & 39 & -120 \end{bmatrix}$$

The adjoint of matrix [A] is $[C]^T$,

$$adj(A) = [C]^T = \begin{bmatrix} -4 & 7 & -3 \\ 80 & -129 & 39 \\ -384 & 420 & -120 \end{bmatrix}$$

Hence

$$[A]^{-1} = \frac{1}{\det(A)} adj(A)$$

$$= \frac{1}{-84}\begin{bmatrix} -4 & 7 & -3 \\ 80 & -129 & 39 \\ -384 & 420 & -120 \end{bmatrix}$$

$$= \begin{bmatrix} 0.4762 & -0.08333 & 0.03571 \\ -0.9524 & 1.417 & -0.4643 \\ 4.571 & -5.000 & 1.429 \end{bmatrix}$$

If the inverse of a square matrix [A] exists, is it unique?

Yes, the inverse of a square matrix is unique, if it exists. The proof is as follows. Assume that the inverse of [A] is [B] and if this inverse is not unique, then let another inverse of [A] exist called [C].

[B] is inverse of [A], then

[B][A] = [I]

Multiply both sides by [C],

[B][A][C] = [I][C]

[B][A][C] = [C]

Since [C] is inverse of [A], [A][C] = [I]

[B][I] = [C]

[B] = [C]

This shows that [B] and [C] are the same. So inverse of [A] is unique.

Key Terms

Consistent system	Inconsistent system
Infinite solutions	Unique solution
Rank	Inverse

Homework

1. Express $\begin{bmatrix} 3 & -1 & 2 \\ 4 & 3 & 7 \\ 2 & -1 & 5 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \\ 4 \end{bmatrix}$

as a system of linear equations.

Answer:
$$\begin{aligned} 3x_1 - x_2 + 2x_3 &= 2 \\ 4x_1 + 3x_2 + 7x_3 &= -1 \\ 2x_1 - x_2 + 5x_3 &= 4. \end{aligned}$$

2. Express the system of linear equations

a) in matrix form [A][X] = [C] and

b) in the augmented form.

$4x_1 - 3x_3 + x_4 = 1$

$5x_1 + x_2 - 8x_4 = 3$

$2x_1 - 5x_2 + 9x_3 - x_4 = 0$

$3x_2 - x_3 + 7x_4 = 2$

Answer: a) $\begin{bmatrix} 4 & 0 & -3 & 1 \\ 5 & 1 & 0 & -8 \\ 2 & -5 & 9 & -1 \\ 0 & 3 & -1 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ 0 \\ 2 \end{bmatrix}$

b) $\begin{bmatrix} 4 & 0 & -3 & 1 & 1 \\ 5 & 1 & 0 & -8 & 3 \\ 2 & -5 & 9 & -1 & 0 \\ 0 & 3 & -1 & 7 & 2 \end{bmatrix}$

3. For a set of equations [A] [X] = [B], a unique solution exists if

A. rank (A) = rank $(A \vdots B)$

B. rank (A) = rank $(A \vdots B)$ and rank (A) = number of unknowns

C. rank (A) = rank $(A \vdots B)$ and rank (A) = number of rows of (A).

4. Rank of

$$A = \begin{bmatrix} 4 & 4 & 4 & 4 \\ 4 & 4 & 4 & 4 \\ 4 & 4 & 4 & 4 \\ 4 & 4 & 4 & 4 \end{bmatrix}$$

is

A. 1

B. 2

C. 3

D. 4

5. A 3x4 matrix can have a rank of at most

A. 3

B. 4

C. 12

D. 5

6. If [A][X] = [C] has a unique solution, where the order of [A] is 3x3, [X] is 3x1, then the rank of [A] is

A. 2

B. 3

C. 4

D. 5

7. Show if the following system of equations is consistent or inconsistent. If they are consistent, determine if the solution would be unique or infinite ones exist.

$$\begin{bmatrix} 1 & 2 & 5 \\ 7 & 3 & 9 \\ 8 & 5 & 14 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 8 \\ 19 \\ 27 \end{bmatrix}$$

Answer: consistent; infinite solutions

8. Show if the following system of equations is consistent or inconsistent. If they are consistent, determine if the solution would be unique or infinite ones exist.

$$\begin{bmatrix} 1 & 2 & 5 \\ 7 & 3 & 9 \\ 8 & 5 & 14 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 8 \\ 19 \\ 28 \end{bmatrix}$$

Answer: inconsistent

9. Show if the following system of equations is consistent or inconsistent. If they are consistent, determine if the solution would be unique or infinite ones exist.

$$\begin{bmatrix} 1 & 2 & 5 \\ 7 & 3 & 9 \\ 8 & 5 & 13 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 8 \\ 19 \\ 28 \end{bmatrix}$$

Answer: consistent; unique

10. For what (which) value(s) of 'a' does the following system have zero, one, infinitely many solutions.

$$x_1 + x_2 + x_3 = 4$$

$$x_3 = 2$$

$$\left(a^2 - 4\right)x_1 + x_3 = a - 2$$

Answer: if $a \neq +2 \; or -2,$ **then there will be a unique solution**

if $a = +2 \text{ or - } 2,$ **then there will be no solution.**

Possibility of infinite solutions does not exist.

11. Find if

$$[A] = \begin{bmatrix} 5 & -2.5 \\ -2 & 3 \end{bmatrix} \; and \; [B] = \begin{bmatrix} 0.3 & 0.25 \\ 0.2 & 0.5 \end{bmatrix}$$

are inverse of each other.

Answer: Yes

12. Find if

$$[A] = \begin{bmatrix} 5 & 2.5 \\ 2 & 3 \end{bmatrix} \; and \; [B] = \begin{bmatrix} 0.3 & -0.25 \\ 0.2 & 0.5 \end{bmatrix}$$

are inverse of each other.

Answer: No

13. Find the

a. cofactor matrix

b. adjoint matrix

of

$$[A] = \begin{bmatrix} 3 & 4 & 1 \\ 2 & -7 & -1 \\ 8 & 1 & 5 \end{bmatrix}$$

14. Find $[A]^{-1}$ using any method

$$[A]=\begin{bmatrix} 3 & 4 & 1 \\ 2 & -7 & -1 \\ 8 & 1 & 5 \end{bmatrix}$$

Answer: $$[A]^{-1}=\begin{bmatrix} 2.931x10^{-1} & 1.638x10^{-1} & -2.586x10^{-2} \\ 1.552x10^{-1} & -6.034x10^{-2} & -4.310x10^{-2} \\ -5.000x10^{-1} & -2.500x10^{-1} & 2.500x10^{-1} \end{bmatrix}$$

15. Prove that if [A] and [B] are both invertible and are square matrices of same order, then

$([A][B])^{-1}=[B]^{-1}[A]^{-1}$

Hint:

$$([A][B])^{-1}=[B]^{-1}[A]^{-1}$$

Let $[C]=[A][B]$

$$\begin{aligned} [C][B]^{-1} &= [A][B][B]^{-1} \\ &= [A][I] \\ &= [A] \end{aligned}$$

Again $[C]=[A][B]$

$$\begin{aligned} [A]^{-1}[C] &= [A]^{-1}[A][B] \\ &= [I][B] \\ &= [B] \end{aligned}$$

So

$$[C][B]^{-1}=[A]-----(1)$$
$$[A]^{-1}[C]=[B]-----(2)$$

From (1) and (2)

$$[C][B]^{-1}[A]^{-1}[C] = [A][B]$$

$$[A][B][B]^{-1}[A]^{-1}[A][B] = [A][B]$$

$$[A]^{-1}[A][B][B]^{-1}[A]^{-1}[A][B] = \left[A^{-1}\right][A][B]$$

$$[B][B]^{-1}\left[A^{-1}\right][A][B] = [B]$$

$$\left[B^{-1}\right][B][B]^{-1}\left[A^{-1}\right][A][B] = [B]^{-1}[B]$$

$$[B]^{-1}\left[A^{-1}\right][A][B] = [I].$$

16. What is the inverse of a square diagonal matrix? Does it always exist?

Hint: Inverse of a square nxn diagonal matrix $[A]$ is

$$[A]^{-1} = \begin{bmatrix} \frac{1}{a_{11}} & 0 & \cdots & 0 \\ 0 & \frac{1}{a_{22}} & \cdots & 0 \\ 0 & & & \vdots \\ \vdots & \cdots & \cdots & \frac{1}{a_{nn}} \end{bmatrix}$$

So inverse exists only if $a_{ii} \neq 0$ for all i.

17. [A] and [B] are square matrices. If [A][B] = [0] and [A] is invertible, show[B] = 0.

Hint: $[A][B] = [0]$

$$\left[A^{-1}\right][A][B] = [A]^{-1}[0]$$

18. If [A] [B] [C]=[I], where [A], [B] and [C] are of the same size, show that [B] is invertible.

19. Prove if [B] is invertible,

$[A][B]^{-1} = [B]^{-1}[A]$ if and only if $[A][B] = [B][A]$

Hint: Multiply by $[B]^{-1}$ on both sides,

$$[A][B][B]^{-1} = [B]^{-1}[A][B]^{-1}$$

20. For

$$[\mathrm{A}] = \begin{bmatrix} 10 & -7 & 0 \\ -3 & 2.099 & 6 \\ 5 & -1 & 5 \end{bmatrix}$$

$$[\mathrm{A}]^{-1} = \begin{bmatrix} -0.1099 & -0.2333 & 0.2799 \\ -0.2999 & -0.3332 & 0.3999 \\ 0.04995 & 0.1666 & 6.664\times10^{-5} \end{bmatrix}$$

Show $\det(A) = \dfrac{1}{\det(A^{-1})}$.

Chapter 6
Gaussian Elimination

After reading this chapter, you will be able to

- *Solve a set of simultaneous linear equations using Naïve Gauss Elimination*
- *Learn the pitfalls of Naïve Gauss Elimination method*
- *Understand the effect of round off error on a solving set of linear equation by Naïve Gauss Elimination Method*
- *Learn how to modify Naïve Gauss Elimination method to Gaussian Elimination with Partial Pivoting Method to avoid pitfalls of the former method.*
- *Find the determinant of a square matrix using Guassian Elimination*
- *Understand the relationship between determinant of coefficient matrix and the solution of simultaneous linear equations.*

How are a set of equations solved numerically?

One of the most popular techniques for solving simultaneous linear equations is the Gaussian elimination method. The approach is designed to solve a general set of n equations and n unknowns

$$a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + ... + a_{1n}x_n = b_1$$

$$a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + ... + a_{2n}x_n = b_2$$

$$\begin{array}{cc} . & . \\ . & . \\ . & . \end{array}$$

$$a_{n1}x_1 + a_{n2}x_2 + a_{n3}x_3 + ... + a_{nn}x_n = b_n$$

Gaussian elimination consists of two steps

1. <u>Forward Elimination of Unknowns:</u> In this step, the unknown is eliminated in each equation starting with the first equation. This way, the equations are "reduced" to one equation and one unknown in each equation.

2. <u>Back Substitution:</u> In this step, starting from the last equation, each of the unknowns is found.

Forward Elimination of Unknowns: In the first step of forward elimination, the first unknown, x_1 is eliminated from all rows below the first row. The first equation is selected as the pivot equation to eliminate x_1. So, to eliminate x_1 in the second equation, one divides the first equation by a_{11} (hence called the pivot element) and then multiply it by a_{21}. That is, same as multiplying the first equation by a_{21}/ a_{11} to give

$$a_{21}x_1 + \frac{a_{21}}{a_{11}}a_{12}x_2 + ... + \frac{a_{21}}{a_{11}}a_{1n}x_n = \frac{a_{21}}{a_{11}}b_1$$

Now, this equation can be subtracted from the second equation to give

$$\left(a_{22} - \frac{a_{21}}{a_{11}}a_{12}\right)x_2 + ... + \left(a_{2n} - \frac{a_{21}}{a_{11}}a_{1n}\right)x_n = b_2 - \frac{a_{21}}{a_{11}}b_1$$

or

$$a'_{22}x_2 + ... + a'_{2n}x_n = b'_2$$

where

$$a'_{22} = a_{22} - \frac{a_{21}}{a_{11}}a_{12}$$

$$\vdots$$

$$a'_{2n} = a_{2n} - \frac{a_{21}}{a_{11}}a_{1n}$$

This procedure of eliminating x_1, is now repeated for the third equation to the n^{th} equation to reduce the set of equations as

$$a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + ... + a_{1n}x_n = b_1$$

$$a'_{22}x_2 + a'_{23}x_3 + ... + a'_{2n}x_n = b'_2$$

$$a'_{32}x_2 + a'_{33}x_3 + ... + a'_{3n}x_n = b'_3$$

$$.\qquad .\qquad .$$

$$.\qquad .\qquad .$$

$$.\qquad .\qquad .$$

$$a'_{n2}x_2 + a'_{n3}x_3 + ... + a'_{nn}x_n = b'_n$$

This is the end of the first step of forward elimination. Now for the second step of forward elimination, we start with the second equation as the pivot equation and a'_{22} as the pivot element. So, to eliminate x_2 in the third equation, one divides the second equation by a'_{22} (the pivot element) and then multiply it by a'_{32}. That is, same as multiplying the second equation by a'_{32}/a'_{22} and subtracting from the third equation. This makes the coefficient of x_2 zero in the third equation. The same procedure is now repeated for the fourth equation till the nth equation to give

$$a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + ... + a_{1n}x_n = b_1$$

$$a'_{22}x_2 + a'_{23}x_3 + ... + a'_{2n}x_n = b'_2$$

$$a''_{33}x_3 + ... + a''_{3n}x_n = b''_3$$

$$.\qquad .$$

$$.\qquad .$$

$$.\qquad .$$

$$a''_{n3}x_3 + ... + a''_{nn}x_n = b''_n$$

The next steps of forward elimination are conducted by using the third equation as a pivot equation and so on. That is, there will be a total of (n-1) steps of forward elimination. At the end of (n-1) steps of forward elimination, we get a set of equations that look like

$$a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + ... + a_{1n}x_n = b_1$$

$$a'_{22}x_2 + a'_{23}x_3 + ... + a'_{2n}x_n = b'_2$$

$$a''_{33}x_3 + ... + a''_n x_n = b''_3$$

$$\begin{array}{cc} . & . \\ . & . \\ . & . \end{array}$$

$$a_{nn}^{(n-1)} x_n = b_n^{(n-1)}$$

<u>Back Substitution:</u> Now the equations are solved starting from the last equation as it has only one unknown.

$$x_n = \frac{b_n^{(n-1)}}{a_{nn}^{(n-1)}}$$

Then the second last equation, that is the $(n-1)^{th}$ equation, has two unknowns - x_n and x_{n-1}, but x_n is already known. This reduces the $(n-1)^{th}$ equation also to one unknown. Back substitution hence can be represented for all equations by the formula

$$x_i = \frac{b_i^{(i-1)} - \sum_{j=i+1}^{n} a_{ij}^{(i-1)} x_j}{a_{ii}^{(i-1)}} \quad \text{for } i = n-1, n-2, \ldots, 1$$

and

$$x_n = \frac{b_n^{(n-1)}}{a_{nn}^{(n-1)}}$$

Example

The upward velocity of a rocket is given at three different times in the following table

Time, t	**Velocity, v**
s	m/s
5	106.8
8	177.2

12	279.2

The velocity data is approximated by a polynomial as

$$v(t) = a_1 t^2 + a_2 t + a_3 , \qquad 5 \le t \le 12.$$

The coefficients a_1, a_2, a_3 for the above expression were found in Chapter 5 to be given by

$$\begin{bmatrix} 25 & 5 & 1 \\ 64 & 8 & 1 \\ 144 & 12 & 1 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} 106.8 \\ 177.2 \\ 279.2 \end{bmatrix}$$

Find the values of a_1, a_2, a_3 using Naïve Guass Elimination. Find the velocity at $t = 6, 7.5, 9, 11$ seconds.

Solution

Forward Elimination of Unknowns: Since there are three equations, there will be two steps of forward elimination of unknowns.

First step: Divide Row 1 by 25 and then multiply it by 64, and then subtract the result from Row 2 gives

$$\begin{bmatrix} 25 & 5 & 1 \\ 0 & -4.8 & -1.56 \\ 144 & 12 & 1 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} 106.81 \\ -96.21 \\ 279.2 \end{bmatrix}$$

Divide Row 1 by 25 and then multiply it by 144, and then subtract the result from Row 3 gives

$$\begin{bmatrix} 25 & 5 & 1 \\ 0 & -4.8 & -1.56 \\ 0 & -16.8 & -4.76 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} 106.8 \\ -96.21 \\ -336.0 \end{bmatrix}$$

Second step: We now divide Row 2 by –4.8 and then multiply by –16.8, and then subtract the result from Row 3, giving

$$\begin{bmatrix} 25 & 5 & 1 \\ 0 & -4.8 & -1.56 \\ 0 & 0 & 0.7 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} 106.8 \\ -96.21 \\ 0.735 \end{bmatrix}$$

<u>Back substitution</u>: From the third equation

$$0.7a_3 = 0.735$$

$$a_3 = \frac{0.735}{0.7}$$

$$= 1.050$$

Substituting the value of a_3 in the second equation,

$$-4.8a_2 - 1.56a_3 = -96.21$$

$$a_2 = \frac{-96.21 + 1.56a_3}{-4.8}$$

$$= \frac{-96.21 + 1.56(1.050)}{-4.8}$$

$$= 19.70$$

Substituting the value of a_2 and a_3 in the first equation,

$$25a_1 + 5a_2 + a_3 = 106.8$$

$$a_1 = \frac{106.8 - 5a_2 - a_3}{25}$$

$$= \frac{106.8 - 5(19.70) - 1.050}{25}$$

$$= 0.2900$$

Hence the solution vector is

$$\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} 0.2900 \\ 19.70 \\ 1.050 \end{bmatrix}$$

The polynomial that passes through the three data points is then

$$v(t) = a_1 t^2 + a_2 t + a_3$$

$$= 0.2900t^2 + 19.70t + 1.050,\ 5 \le t \le 12$$

Since we want to find the velocity at $t = 6, 7.5, 9 \text{ and } 11$ seconds, we could simply substitute each value of t in $v(t) = 0.2900t^2 + 19.70t + 1.050$ and find the corresponding velocity. For example, at $t = 6$

$$v(6) = 0.2900(6)^2 + 19.70(6) + 1.050$$
$$= 129.69\ m/s.$$

However we could also find all the needed values of velocity at t = 6, 7.5, 9 and 11 seconds using matrix multiplication.

$$v(t) = \begin{bmatrix} 0.29 & 19.7 & 1.05 \end{bmatrix} \begin{bmatrix} t^2 \\ t \\ 1 \end{bmatrix}$$

So if we want to find $v(6), v(7.5), v(9), v(11),$ it is given by

$$\begin{bmatrix} v(6) \, v(7.5) \, v(9) \, v(11) \end{bmatrix}$$

$$= \begin{bmatrix} 0.29 & 19.7 & 1.05 \end{bmatrix} \begin{bmatrix} 6^2 & 7.5^2 & 9^2 & 11^2 \\ 6 & 7.5 & 9 & 11 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 0.2900 & 19.70 & 1.050 \end{bmatrix} \begin{bmatrix} 36 & 56.25 & 81 & 121 \\ 6 & 7.5 & 9 & 11 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 129.7 & 165.1 & 201.8 & 252.8 \end{bmatrix}$$

$$v(6) = 129.7\ m/s$$

$$v(7.5) = 165.1\ m/s$$

$$v(9) = 201.8\ m/s$$

$$v(11) = 252.8\ m/s$$

Example

Use Naïve Gauss Elimination to solve

$10x_1 - 7x_2 = 7$

$-3x_1 + 2.099x_2 + 6x_3 = 3.901$

$5x_1 - x_2 + 5x_3 = 6$

Use <u>six</u> significant digits with chopping in your calculations.

Solution

Working in the matrix form

$$\begin{bmatrix} 10 & -7 & 0 \\ -3 & 2.099 & 6 \\ 5 & -1 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 7 \\ 3.901 \\ 6 \end{bmatrix}$$

<u>Forward Elimination of Unknowns</u>

Dividing Row 1 by 10 and multiplying by –3, that is, multiplying Row 1 by -0.3, and subtract it from Row 2 would eliminate a_{21},

$$\begin{bmatrix} 10 & -7 & 0 \\ 0 & -0.001 & 6 \\ 5 & -1 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 7 \\ 6.001 \\ 6 \end{bmatrix}$$

Again dividing Row 1 by 10 and multiplying by 5, that is, multiplying Row 1 by 0.5, and subtract it from Row 3 would eliminate a_{31},

$$\begin{bmatrix} 10 & -7 & 0 \\ 0 & -0.001 & 6 \\ 0 & 2.5 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 7 \\ 6.001 \\ 2.5 \end{bmatrix}$$

This is the end of the first step of forward elimination.

Now for the second step of forward elimination, we would use Row 2 as the pivot equation and eliminate Row 3 – Column 2. Dividing Row 2 by –0.001 and multiplying by 2.5, that is multiplying Row 2 by –2500, and subtracting from Row 3 gives

$$\begin{bmatrix} 10 & -7 & 0 \\ 0 & -0.001 & 6 \\ 0 & 0 & 15005 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 7 \\ 6.001 \\ 15005 \end{bmatrix}$$

This is the end of the forward elimination steps.

<u>Back substitution</u>

We can now solve the above equations by back substitution. From the third equation,

$$15005x_3 = 15005$$

$$x_3 = \frac{15005}{15005}$$

$$= 1.$$

Substituting the value of x_3 in the second equation

$$-0.001x_2 + 6x_3 = 6.001$$

$$x_2 = \frac{6.001 - 6x_3}{-0.001}$$

$$= \frac{6.001 - 6(1)}{-0.001}$$

$$= \frac{6.001 - 6}{-0.001}$$

$$= \frac{0.001}{-0.001}$$

$$= -1$$

Substituting the value of x_3 and x_2 in the first equation,

$$10x_1 - 7x_2 + 0x_3 = 7$$

$$x_1 = \frac{7 + 7\,x_2 - 0x_3}{10}$$

$$= \frac{7 + 7(-1) - 0(1)}{10}$$

$$= 0$$

Hence the solution is

$$[X] = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}$$

Are there any pitfalls of Naïve Gauss Elimination Method?

Yes, there are two pitfalls of Naïve Gauss Elimination method.

<u>Division by zero</u>: It is possible that division by zero may occur during forward elimination steps. For example for the set of equations

$$10x_2 - 7x_3 = 7$$

$$6x_1 + 2.099x_2 - 3x_3 = 3.901$$

$$5x_1 - x_2 + 5x_3 = 6$$

during the first forward elimination step, the coefficient of x_1 is zero and hence normalization would require division by zero.

<u>Round-off error</u>: Naïve Gauss Elimination Method is prone to round-off errors. This is true when there are large numbers of equations as errors propagate. Also, if there is subtraction of numbers from each other, it may create large errors. See the example below.

Example

Remember the previous example where we used Naïve Gauss Elimination to solve

$$10x_1 - 7x_2 = 7$$

$$-3x_1 + 2.099x_2 + 6x_3 = 3.901$$

$$5x_1 - x_2 + 5x_3 = 6$$

using six significant digits with chopping in your calculations. Repeat the problem, but now use five significant digits with chopping in your calculations.

Solution

Writing in the matrix form

$$\begin{bmatrix} 10 & -7 & 0 \\ -3 & 2.099 & 6 \\ 5 & -1 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 7 \\ 3.901 \\ 6 \end{bmatrix}$$

Forward Elimination of Unknowns

Dividing Row 1 by 10 and multiplying by –3, that is, multiplying Row 1 by -0.3, and subtract it from Row 2 would eliminate a_{21},

$$\begin{bmatrix} 10 & -7 & 0 \\ 0 & -0.001 & 6 \\ 5 & -1 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 7 \\ 6.001 \\ 6 \end{bmatrix}$$

Again dividing Row 1 by 10 and multiplying by 5, that is, multiplying the Row 1 by 0.5, and subtract it from Row 3 would eliminate a_{31},

$$\begin{bmatrix} 10 & -7 & 0 \\ 0 & -0.001 & 6 \\ 0 & 2.5 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 7 \\ 6.001 \\ 2.5 \end{bmatrix}$$

This is the end of the first step of forward elimination.

Now for the second step of forward elimination, we would use Row 2 as the pivoting equation and eliminate Row 3 – Column 2. Dividing Row 2 by –0.001 and multiplying by 2.5, that is, multiplying Row 2 by –2500, and subtract from Row 3 gives

$$\begin{bmatrix} 10 & -7 & 0 \\ 0 & -0.001 & 6 \\ 0 & 0 & 15005 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 7 \\ 6.001 \\ 15004 \end{bmatrix}$$

This is the end of the forward elimination steps.

Back substitution

We can now solve the above equations by back substitution. From the third equation,

$$15005x_3 = 15004$$

$$x_3 = \frac{15004}{15005}$$

$$= 0.99993$$

Substituting the value of x_3 in the second equation

$$-0.001x_2 + 6x_3 = 6.001$$

$$x_2 = \frac{6.001 - 6x_3}{-0.001}$$

$$= \frac{6.001 - 6(0.99993)}{-0.001}$$

$$= \frac{6.001 - 5.9995}{-0.001}$$

$$= \frac{0.0015}{-0.001}$$

$$= -1.5$$

Substituting the value of x_3 and x_2 in the first equation,

$$10x_1 - 7x_2 + 0x_3 = 7$$

$$x_1 = \frac{7 + 7x_2 - 0x_3}{10}$$

$$= \frac{7 + 7(-1.5) - 0(1)}{10}$$

$$= \frac{7 - 10.5 - 0}{10}$$

$$= \frac{-3.5}{10}$$

$$= -0.3500$$

Hence the solution is

$$[X] = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

$$= \begin{bmatrix} -0.35 \\ -1.5 \\ 0.99993 \end{bmatrix}$$

Compare this with the exact solution of

$$[X] = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

$$= \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}$$

What are the techniques for improving Naïve Gauss Elimination Method?

As seen in the example, round off errors were large when five significant digits were used as opposed to six significant digits. So one way of decreasing round off error would be to use more significant digits, that is, use double or quad precision. However, this would not avoid division by zero errors in Naïve Gauss Elimination. To avoid division by zero as well as reduce (not eliminate) round off error, Gaussian Elimination with partial pivoting is the method of choice.

How does Gaussian elimination with partial pivoting differ from Naïve Gauss elimination?

The two methods are the same, except in the beginning of each step of forward elimination, a row switching is done based on the following criterion. If there are n equations, then there are $(n-1)$ forward elimination steps. At the beginning of the k^{th} step of forward elimination, one finds the maximum of

$$|a_{kk}|, |a_{k+1,k}|, \ldots\ldots\ldots\ldots, |a_{nk}|$$

Then if the maximum of these values is $|a_{pk}|$ in the p^{th} row, k=p= n, then switch rows p and k.

The other steps of forward elimination are the same as Naïve Gauss elimination method. The back substitution steps stay exactly the same as Naïve Gauss Elimination method.

Example

In the previous two examples, we used Naïve Gauss Elimination to solve

$$10x_1 - 7x_2 = 7$$

$$-3x_1 + 2.099x_2 + 6x_3 = 3.901$$

$$5x_1 - x_2 + 5x_3 = 6$$

using five and six significant digits with chopping in the calculations. Using <u>five</u> significant digits with chopping, the solution found was

$$[X] = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

$$= \begin{bmatrix} -0.35 \\ -1.5 \\ 0.99993 \end{bmatrix}$$

This is different from the exact solution

$$[X] = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

$$= \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}$$

Find the solution using Gaussian elimination with partial pivoting using five significant digits with chopping in your calculations.

Solution

$$\begin{bmatrix} 10 & -7 & 0 \\ -3 & 2.099 & 6 \\ 5 & -1 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 7 \\ 2.901 \\ 6 \end{bmatrix}$$

Forward Elimination of Unknowns

Now for the first step of forward elimination, the absolute value of first column elements are

$|10|, |-3|, |5|$

or

10, 3, 5

So the largest absolute value is in the Row 1. So as per Gaussian Elimination with partial pivoting, the switch is between Row 1 and Row 1 to give

$$\begin{bmatrix} 10 & 7 & 0 \\ -3 & 2.099 & 6 \\ 5 & -1 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 7 \\ 3.901 \\ 6 \end{bmatrix}$$

Dividing Row 1 by 10 and multiplying by –3, that is, multiplying the Row 1 by -0.3, and subtract it from Row 2 would eliminate a_{21},

$$\begin{bmatrix} 10 & -7 & 0 \\ 0 & -0.001 & 6 \\ 5 & -1 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 7 \\ 6.001 \\ 6 \end{bmatrix}$$

Again dividing Row 1 by 10 and multiplying by 5, that is, multiplying the Row 1 by 0.5, and subtract it from Row 3 would eliminate a_{31},

$$\begin{bmatrix} 10 & -7 & 0 \\ 0 & -0.001 & 6 \\ 0 & 2.5 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 7 \\ 6.001 \\ 2.5 \end{bmatrix}$$

This is the end of the first step of forward elimination.

Now for the second step of forward elimination, the absolute value of the second column elements below the Row 2 is

$|-0.001|$, $|2.5|$

or

0.001, 2.5

So the largest absolute value is in Row 3. So the Row 2 is switched with the Row 3 to give

$$\begin{bmatrix} 10 & -7 & 0 \\ 0 & 2.5 & 5 \\ 0 & -0.001 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 7 \\ 2.5 \\ 6.001 \end{bmatrix}$$

Dividing row 2 by 2.5 and multiplying by –0.001, that is multiplying by 0.001/2.5=-0.0004, and then subtracting from Row 3 gives

$$\begin{bmatrix} 10 & -7 & 0 \\ 0 & 2.5 & 5 \\ 0 & 0 & 6.002 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 7 \\ 2.5 \\ 6.002 \end{bmatrix}$$

<u>Back substitution</u>

$$6.002x_3 = 6.002$$

$$x_3 = \frac{6.002}{6.002}$$

$$=1$$

Substituting the value of x_3 in Row 2

$$2.5x_2 + 5x_3 = 2.5$$

$$x_2 = \frac{2.5 - 5x_2}{2.5}$$

$$= \frac{2.5 - 5(1)}{2.5}$$

$$= \frac{2.5-5}{2.5}$$

$$= \frac{-2.5}{2.5}$$

$$= -1$$

Substituting the value of x_3 and x_2 in Row 1

$$10x_1 - 7x_2 + 0x_2 = 7$$

$$x_1 = \frac{7 + 7x_2 - 0x_3}{10}$$

$$= \frac{7 + 7(-1) - 0(1)}{10}$$

$$= \frac{7 - 7 - 0}{10}$$

$$= \frac{0}{10}$$

$$= 0$$

So the solution is

$$[X] = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}$$

This, in fact, is the exact solution. By coincidence only, in this case, the round off error is fully removed.

Can we use Naïve Gauss Elimination methods to find the determinant of a square matrix?

One of the more efficient ways to find the determinant of a square matrix is by taking advantage of the following two theorems on a determinant of matrices coupled with Naïve Gauss elimination.

Theorem 1: Let [A] be a nxn matrix. Then, if [B] is a matrix that results from adding or subtracting a multiple of one row to another row, then det (B) = det (A). (The same is true for column operations also).

Theorem 2: Let [A] be a nxn matrix that is upper triangular, lower triangular or diagonal, then det(A) = $a_{11} * a_{22} * \ldots\ldots * a_{nn} = \prod_{i=1}^{n} a_{ii}$

This implies that if we apply the forward elimination steps of Naive Gauss Elimination method, the determinant of the matrix stays the same according the Theorem 1. Then since at the end of the forward elimination steps, the resulting matrix is upper triangular, the determinant will be given by Theorem 2.

Example

Find the determinant of

$$[A]=\begin{bmatrix} 25 & 5 & 1 \\ 64 & 8 & 1 \\ 144 & 12 & 1 \end{bmatrix}$$

Solution

Remember earlier in this chapter, we conducted the steps of forward elimination of unknowns using Naïve Gauss Elimination method on [A] to give

$$[B]=\begin{bmatrix} 25 & 5 & 1 \\ 0 & -4.8 & -1.56 \\ 0 & 0 & 0.7 \end{bmatrix}$$

According to Theorem 2

det (A) = det (B)

$$= (25)(-4.8)(0.7)$$
$$= -84.00$$

What if I cannot find the determinant of the matrix using Naive Gauss Elimination method, for example, if I get division by zero problems during Naïve Gauss Elimination method?

Well, you can apply Gaussian Elimination with partial pivoting. However, the determinant of the resulting upper triangular matrix may differ by a sign. The following theorem applies in addition to the previous two to find determinant of a square matrix.

Theorem 3: Let [A] be a nxn matrix. Then, if [B] is a matrix that results from switching one row with another row, then det (B) = - det (A).

Example

Find the determinant of

$$[A]=\begin{bmatrix} 10 & -7 & 0 \\ -3 & 2.099 & 6 \\ 5 & -1 & 5 \end{bmatrix}$$

Remember from that at the end of the forward elimination steps of Gaussian elimination with partial pivoting, we obtained

$$[B]=\begin{bmatrix} 10 & -7 & 0 \\ 0 & 2.5 & 5 \\ 0 & 0 & 6.002 \end{bmatrix}$$

$$\det(B)=(10)(2.5)(6.002)=150.05$$

Since rows were switched once during the forward elimination steps of Gaussian elimination with partial pivoting,

$$\det(\mathrm{A})=-\det(\mathrm{B})$$

$$=-150.05.$$

Prove det (A) = $\dfrac{1}{\det\left(\mathrm{A}^{-1}\right)}$**.**

Proof

$[A][A]^{-1} = [I]$

$$\det\left(A\,A^{-1}\right) = \det(I)$$
$$\det(A)\det\left(A^{-1}\right) = 1$$
$$\det(A) = \frac{1}{\det\left(A^{-1}\right)}.$$

If [A] is a nxn matrix and det (A) ≠ 0, what other statements are equivalent to it?

1. [A] is invertible.

2. $[A]^{-1}$ exists.

3. [A] [X] = [C] has a unique solution.

4. [A] [X] = [0] solution is $[X] = \vec{0}$.

5. $[A]\,[A]^{-1} = [I] = [A]^{-1}\,[A]$.

Key Terms

Naïve Gauss Elimination   Partial Pivoting   Determinant

Homework

1. The goal of forward elimination steps in Naïve Gauss elimination method is to reduce the coefficient matrix

    A. to a diagonal matrix

    B. to an upper triangular matrix

    C. to a lower triangular matrix

    D. to an identity matrix

2. Using a computer with four significant digits with chopping, use Naïve Gauss elimination to solve

$$4x_1 + x_2 - x_3 = -2$$
$$5x_1 + x_2 + 2x_3 = 4$$
$$6x_1 + x_2 + x_3 = 6$$

Answer: (3, -13, 1)

3. Using a computer with four significant digits with chopping, use Gaussian Elimination with partial pivoting to solve

$$
\begin{aligned}
4x_1 + x_2 - x_3 &= -2 \\
5x_1 + x_2 + 2x_3 &= 4 \\
6x_1 + x_2 + x_3 &= 6
\end{aligned}
$$

Answer: (2.995, -12.98, 1.001)

4. For

$$
[\mathrm{A}] = \begin{bmatrix} 10 & -7 & 0 \\ -3 & 2.099 & 6 \\ 5 & -1 & 5 \end{bmatrix}
$$

Find the determinant of [A] using forward elimination step of Naïve Gauss Elimination method.

Answer: -150.05

Chapter 7
LU Decomposition

After reading this chapter, you will be able to

- *Learn when LU Decomposition is numerically more efficient than Gaussian Elimination*
- *Decompose a nonsingular matrix into LU*
- *Show how LU decomposition is used to find matrix inverse.*

I hear about LU Decomposition used as a method to solve a set of simultaneous linear equations? What is it and why do we need to learn different methods of solving a set of simultaneous linear equations?

We already studied two numerical methods of finding the solution to simultaneous linear equations – Naïve Gauss Elimination and Gaussian Elimination with Partial Pivoting. Then, why do we need to learn another method? To appreciate why LU Decomposition could be a better choice than the Gauss Elimination techniques in some cases, let us discuss first what LU Decomposition is about.

For any nonsingular matrix $[A]$ on which one can conduct Naïve Gauss Elimination forward elimination steps, one can always write it as

$$[A]=[L][U]$$

where

[L] = lower triangular matrix

[U] = upper triangular matrix

Then if one is solving a set of equations

[A] [X] = [C],

then

$[L][U][X]=[C]$ $\qquad\qquad$ $([A]=[L][U])$

Multiplying both side by $[L]^{-1}$,

$[L]^{-1}[L][U][X]=[L]^{-1}[C]$

$[I][U][X]=[L]^{-1}[C]$ $\qquad\qquad$ $([L]^{-1}[L]=[I])$

$[U][X]=[L]^{-1}[C]$ $\qquad\qquad$ $([I][U]=[U])$

Let

$[L]^{-1}[C]=[Z]$

then

$$[L][\mathrm{Z}]=[C] \qquad (1)$$

and

$$[U][\mathrm{X}]=[Z] \qquad (2)$$

So we can solve equation (1) first for $[Z]$ and then use equation (2) to calculate $[X]$.

This is all exciting but this looks more complicated than the Gaussian elimination techniques!! I know but I cannot tease you any longer. So here we go!

Without proof, the computational time required to decompose the $[A]$ matrix to [L] [U] form is proportional to $\frac{n^3}{3}$, where n is the number of equations (size of $[A]$ matrix). Then to solve the $[L][\mathrm{Z}]=[C]$, the computational time is proportional to $\frac{n^2}{2}$. Then to solve the $[U][\mathrm{X}]=[C]$, the computational time is proportional to $\frac{n^2}{2}$. So the total computational time to solve a set of equations by LU decomposition is proportional to $\frac{n^3}{3}+n^2$.

In comparison, Gaussian elimination is computationally more efficient. It takes a computational time proportional to $\frac{n^3}{3}+\frac{n^2}{2}$, where the computational time for forward elimination is proportional to $\frac{n^3}{3}$ and for the back substitution the time is proportional to $\frac{n^2}{2}$.

This has confused me further! Gaussian elimination takes less time than LU Decomposition method and you are trying to convince me then LU Decomposition has its place in solving linear equations! Yes, it does.

Remember in trying to find the inverse of the matrix $[A]$ in Chapter 5, the problem reduces to solving 'n' sets of equations with the 'n' columns of the identity matrix as the RHS vector. For calculations of each column of the inverse of the $[A]$ matrix, the coefficient matrix $[A]$ matrix in the set of equation $[A][\mathrm{X}]=[C]$ does not change. So if we use LU Decomposition method, the $[A]=[L][\mathrm{U}]$ decomposition needs to be done only once and the use of equations (1) and (2) still needs to be done 'n' times.

So the total computational time required to find the inverse of a matrix using LU decomposition is proportional to $\frac{n^3}{3}+n(n^2)=\frac{4n^3}{3}$.

In comparison, if Gaussian elimination method were applied to find the inverse of a matrix, the time would be proportional to

$$n\left(\frac{n^3}{3}+\frac{n^2}{2}\right)=\frac{n^4}{3}+\frac{n^3}{2}.$$

For large values of n

$$\frac{n^4}{3}+\frac{n^3}{2}\rangle\rangle\frac{4n^3}{3}$$

Are you now convinced now that LU decomposition has its place in solving systems of equations? We are now ready to answer other questions - how do I find LU matrices for a nonsingular matrix [A] and how do I solve equations (1) and (2).

How do I decompose a non-singular matrix [A], that is, how do I find $[A]=[L][\mathrm{U}]$?

If forward elimination steps of Naïve Gauss elimination methods can be applied on a nonsingular matrix, then $[A]$ can be decomposed into L U as

$$[A]=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

$$=\begin{bmatrix} 1 & 0 & \cdots & 0 \\ \ell_{21} & 1 & & \vdots \\ \ell_{31} & & \ddots & 0 \\ \vdots & & & 1 \\ \ell_{n1} & \cdots & \cdots & \end{bmatrix}\begin{bmatrix} \mathrm{u}_{11} & \mathrm{u}_{12} & \cdots & \mathrm{u}_{1\mathrm{n}} \\ 0 & \ddots & & \\ \vdots & & \ddots & 0 \\ & & \cdots & \mathrm{u}_{\mathrm{nn}} \\ 0 & 0 & & \end{bmatrix}$$

1. The elements of the $[U]$ matrix are exactly the same as the coefficient matrix one obtains at the end of the forward elimination steps in Naïve Gauss Elimination.
2. The lower triangular matrix $[L]$ has 1 in its diagonal entries. The non zero elements on the non-diagonal elements in $[L]$ are multipliers that made the corresponding entries zero in the upper triangular matrix $[U]$ during forward elimination.

Let us look at this using the same example as used in Naïve Gaussian elimination.

Example

Find the LU decomposition of the matrix

$$[A]=\begin{bmatrix} 25 & 5 & 1 \\ 64 & 8 & 1 \\ 144 & 12 & 1 \end{bmatrix}$$

Solution

$$[A]=[L][U]=\begin{bmatrix}1 & 0 & 0\\ \ell_{21} & 1 & 0\\ \ell_{31} & \ell_{32} & 1\end{bmatrix}\begin{bmatrix}u_{11} & u_{12} & u_{13}\\ 0 & u_{22} & u_{23}\\ 0 & 0 & u_{33}\end{bmatrix}$$

The $[U]$ matrix is the same as found at the end of the forward elimination of Naïve Gauss elimination method, that is

$$[U]=\begin{bmatrix}25 & 5 & 1\\ 0 & -4.8 & -1.56\\ 0 & 0 & 0.7\end{bmatrix}$$

To find ℓ_{21} and ℓ_{31}, what multiplier was used to make the a_{21} and a_{31} elements zero in the first step of forward elimination of Naïve Gauss Elimination Method It was

$$\ell_{21}=\frac{64}{25}$$

$$=2.56$$

$$\ell_{31}=\frac{144}{25}$$

$$=5.76$$

To find ℓ_{32}, what multiplier was used to make a_{32} element zero. Remember a_{32} element was made zero in the second step of forward elimination. The $[A]$ matrix at the beginning of the second step of forward elimination was

$$\begin{bmatrix}25 & 5 & 1\\ 0 & -4.8 & -1.56\\ 0 & -16.8 & -4.76\end{bmatrix}$$

So

$$\ell_{32}=\frac{-16.8}{-4.8}$$

$$=3.5$$

Hence

$$[L]=\begin{bmatrix}1 & 0 & 0\\ 2.56 & 1 & 0\\ 5.76 & 3.5 & 1\end{bmatrix}$$

Confirm

$$[L][U]=[A].$$

$$[L][U]=\begin{bmatrix}1 & 0 & 0\\ 2.56 & 1 & 0\\ 5.76 & 3.5 & 1\end{bmatrix}\begin{bmatrix}25 & 5 & 1\\ 0 & -4.8 & -1.56\\ 0 & 0 & 0.7\end{bmatrix}$$

$$=\begin{bmatrix}25 & 5 & 1\\ 64 & 8 & 1\\ 144 & 12 & 1\end{bmatrix}$$

Example

Use LU decomposition method to solve the following simultaneous linear equations.

$$\begin{bmatrix}25 & 5 & 1\\ 64 & 8 & 1\\ 144 & 12 & 1\end{bmatrix}\begin{bmatrix}a_1\\ a_2\\ a_3\end{bmatrix}=\begin{bmatrix}106.8\\ 177.2\\ 279.2\end{bmatrix}$$

Solution

Recall that

$$[A][X]=[C]$$

and if

$$[A]=[L][U]$$

then first solving

$$[L][Z]=[C]$$

and then

$[U][X]=[Z]$

gives the solution vector $[X]$.

Now in the previous example, we showed

$$[A]=[L][U]=\begin{bmatrix} 1 & 0 & 0 \\ 2.56 & 1 & 0 \\ 5.76 & 3.5 & 1 \end{bmatrix}\begin{bmatrix} 25 & 5 & 1 \\ 0 & -4.8 & -1.56 \\ 0 & 0 & 0.7 \end{bmatrix}$$

First solve

$[L][Z]=[C]$

$$\begin{bmatrix} 1 & 0 & 0 \\ 2.56 & 1 & 0 \\ 5.76 & 3.5 & 1 \end{bmatrix}\begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix}=\begin{bmatrix} 106.8 \\ 177.2 \\ 279.2 \end{bmatrix}$$

to give

$z_1 = 106.8$

$2.56z_1 + z_2 = 177.2$

$5.76 z_1 + 3.5 z_2 + z_3 = 279.2$

Forward substitution starting from the first equation gives

$z_1 = 106.8$

$z_2 = 177.2 - 2.56 z_1$

$\quad = 177.2 - 2.56 (106.8)$

$\quad = -96.21$

$z_3 = 279.2 - 5.76 z_1 - 3.5z_2$

$\quad = 279.2 - 5.76(106.8) - 3.5(-96.21)$

$\quad = 0.735$

Hence

$$[Z] = \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix}$$

$$= \begin{bmatrix} 106.8 \\ -96.21 \\ 0.735 \end{bmatrix}$$

This matrix is same as the right hand side obtained at the end of the forward elimination steps of Naïve Gauss elimination method. Is this a coincidence?

Now solve

$$[U][X] = [Z]$$

$$\begin{bmatrix} 25 & 5 & 1 \\ 0 & -4.8 & -1.56 \\ 0 & 0 & 0.7 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} 106.8 \\ -96.21 \\ 0.735 \end{bmatrix}$$

$$25a_1 + 5a_2 + a_3 = 106.8$$

$$-4.8a_2 - 1.56a_3 = -96.21$$

$$0.7a_3 = 0.735$$

From the third equation

$$0.7a_3 = 0.735$$

$$a_3 = \frac{0.735}{0.7}$$

$$= 1.050$$

Substituting the value of a_3 in the second equation,

$$-4.8a_2 - 1.56a_3 = -96.21$$

$$a_2 = \frac{-96.21 + 1.56a_3}{-4.8}$$

$$= \frac{-96.21 + 1.56(1.050)}{-4.8}$$

$= 19.70$

Substituting the value of a_2 and a_3 in the first equation,

$$25a_1 + 5a_2 + a_3 = 106.8$$

$$a_1 = \frac{106.8 - 5a_2 - a_3}{25}$$

$$= \frac{106.8 - 5(19.70) - 1.050}{25}$$

$$= 0.2900$$

The solution vector is

$$\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} 0.2900 \\ 19.70 \\ 1.050 \end{bmatrix}$$

How do I find the inverse of a square matrix using LU Decomposition?

A matrix $[B]$ is the inverse of $[A]$ if $[A][B] = [I] = [B][A]$. How can we use LU decomposition to find inverse of the matrix? Assume the first column of $[B]$ (the inverse of [A] is

$[b_{11}\, b_{12} \ldots\ldots\ldots\ldots b_{n1}]^T$

then from the above definition of inverse and definition of matrix multiplication.

$$[A]\begin{bmatrix} b_{11} \\ b_{21} \\ \vdots \\ b_{n1} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Similarly the second column of $[B]$ is given by

$$[A]\begin{bmatrix} b_{12} \\ b_{22} \\ \vdots \\ b_{n2} \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}$$

Similarly, all columns of $[B]$ can be found by solving n different sets of equations with the column of the right hand sides being the n columns of the identity matrix.

Example

Use LU decomposition to find the inverse of

$$[A] = \begin{bmatrix} 25 & 5 & 1 \\ 64 & 8 & 1 \\ 144 & 12 & 1 \end{bmatrix}$$

Solution

Knowing that

$$[A] = [L][U] = \begin{bmatrix} 1 & 0 & 0 \\ 2.56 & 1 & 0 \\ 5.76 & 3.5 & 1 \end{bmatrix}\begin{bmatrix} 25 & 5 & 1 \\ 0 & -4.8 & -1.56 \\ 0 & 0 & 0.7 \end{bmatrix}$$

We can solve for the first column of $[B] = [A]^{-1}$ by solving for

$$\begin{bmatrix} 25 & 5 & 1 \\ 64 & 8 & 1 \\ 144 & 12 & 1 \end{bmatrix}\begin{bmatrix} b_{11} \\ b_{21} \\ b_{31} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

First solve

$[L][Z] = [C]$, that is

$$\begin{bmatrix} 1 & 0 & 0 \\ 2.56 & 1 & 0 \\ 5.76 & 3.5 & 1 \end{bmatrix}\begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

to give

$z_1 = 1$

$2.56z_1 + z_2 = 0$

$5.76z_1 + 3.5z_2 + z_3 = 0$

Forward substitution starting from the first equation gives

$\mathrm{z}_1 = 1$

$$\begin{aligned}
\mathrm{z}_2 &= 0 - 2.56\mathrm{z}_1 \\
&= 0 - 2.56(1) \\
&= -2.56
\end{aligned}$$

$$\begin{aligned}
z_3 &= 0 - 5.76z_1 - 3.5z_2 \\
&= 0 - 5.76(1) - 3.5(-2.56) \\
&= 3.2
\end{aligned}$$

Hence

$$[Z] = \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix}$$

$$= \begin{bmatrix} 1 \\ -2.56 \\ 3.2 \end{bmatrix}$$

Now solve

$$[U][X] = [Z]$$

that is

$$\begin{bmatrix} 25 & 5 & 1 \\ 0 & -4.8 & -1.56 \\ 0 & 0 & 0.7 \end{bmatrix} \begin{bmatrix} b_{11} \\ b_{21} \\ b_{31} \end{bmatrix} = \begin{bmatrix} 1 \\ -2.56 \\ 3.2 \end{bmatrix}$$

$$25b_{11} + 5b_{21} + b_{31} = 1$$

$$-4.8b_{21} - 1.56b_{31} = -2.56$$

$$0.7b_{31} = 3.2$$

Backward substitution starting from the third equation gives

$$b_{31} = \frac{3.2}{0.7}$$

$$= 4.571$$

$$b_{21} = \frac{-2.56 + 1.560b_{31}}{-4.8}$$

$$= \frac{-2.56 + 1.560(4.571)}{-4.8}$$

$$= -0.9524$$

$$b_{11} = \frac{1 - 5b_{21} - b_{31}}{25}$$

$$= \frac{1 - 5(-0.9524) - 4.571}{25}$$

$$= 0.04762$$

Hence the first column of the inverse of [A] is

$$\begin{bmatrix} b_{11} \\ b_{21} \\ b_{31} \end{bmatrix} = \begin{bmatrix} 0.04762 \\ -0.9524 \\ 4.571 \end{bmatrix}$$

Similarly by solving

$$\begin{bmatrix} 25 & 5 & 1 \\ 64 & 8 & 1 \\ 144 & 12 & 1 \end{bmatrix} \begin{bmatrix} b_{12} \\ b_{22} \\ b_{32} \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} b_{12} \\ b_{22} \\ b_{32} \end{bmatrix} = \begin{bmatrix} -0.08333 \\ 1.417 \\ -5.000 \end{bmatrix}$$

and solving

$$\begin{bmatrix} 25 & 5 & 1 \\ 64 & 8 & 1 \\ 144 & 12 & 1 \end{bmatrix} \begin{bmatrix} \mathrm{b}_{13} \\ \mathrm{b}_{23} \\ \mathrm{b}_{33} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

gives

$$\begin{bmatrix} b_{13} \\ b_{23} \\ b_{33} \end{bmatrix} = \begin{bmatrix} 0.03571 \\ -0.4643 \\ 1.429 \end{bmatrix}$$

Hence

$$[A]^{-1} = \begin{bmatrix} 0.4762 & 0.08333 & 0.0357 \\ -0.9524 & 1.417 & -0.4643 \\ 4.571 & -5.050 & 1.429 \end{bmatrix}$$

Can you confirm the following for the above example? $[A][A]^{-1} = [I] = [A]^{-1}[A]$

Key Terms

LU decomposition Inverse

Homework

1. Show that LU decomposition is computationally more efficient way of finding the inverse of a square matrix than using Gaussian elimination.
2. LU decomposition method is computationally more efficient than Naïve Gauss elimination for
   A. Solving a single set of simultaneous linear equations
   B. Solving multiple sets of simultaneous linear equations with different coefficient matrices.

C. Solving multiple sets of simultaneous linear equations with same coefficient matrix but different right hand sides.

D. Solving less than ten simultaneous linear equations.

3. It one decomposes a symmetric matrix [A] to a LU form, then

A. $[L] = [U]^T$

B. $[U]^T = [L]^T$

A. [L] = [U]

D. [L] = [I]

4. Use LU decomposition to solve

$$
\begin{aligned}
4x_1 + x_2 - x_3 &= -2 \\
5x_1 + x_2 + 2x_3 &= 4 \\
6x_1 + x_2 + x_3 &= 6
\end{aligned}
$$

Answer: (3, -13, 1)

5. Find the inverse of

$$
[A] = \begin{bmatrix} 3 & 4 & 1 \\ 2 & -7 & -1 \\ 8 & 1 & 5 \end{bmatrix}
$$

using LU decomposition

Answer: $[\mathrm{A}]^{-1} = \begin{bmatrix} 2.931\mathrm{x}10^{-1} & 1.638\mathrm{x}10^{-1} & -2.586\mathrm{x}10^{-2} \\ 1.552\mathrm{x}10^{-1} & -6.034\mathrm{x}10^{-2} & -4.310\mathrm{x}10^{-2} \\ -5.000\mathrm{x}10^{-1} & -2.500\mathrm{x}10^{-1} & 2.500\mathrm{x}10^{-1} \end{bmatrix}$

6. Show that the nonsingular matrix $[A] = \begin{bmatrix} 0 & 2 \\ 2 & 0 \end{bmatrix}$ cannot be decomposed into LU form.

Hint: Try to find the unknowns in

$$
\begin{bmatrix} 1 & 0 \\ \ell_{21} & 1 \end{bmatrix} \begin{bmatrix} u_{11} & u_{12} \\ 0 & u_{22} \end{bmatrix} = \begin{bmatrix} 0 & 2 \\ 2 & 0 \end{bmatrix}
$$

Do you see any inconsistencies? Understand that $\begin{bmatrix} 0 & 2 \\ 2 & 0 \end{bmatrix}$ **is nonsingular.**

Chapter 8
Gauss-Siedel Method

After reading this chapter, you will be able to

- *Solve a set of equations using Gauss-Siedel method*
- *Learn the advantages and pitfalls of Gauss-Siedel method*
- *Understand under what conditions Gauss-Siedel method always converges*

Why do we need another method to solve a set of simultaneous linear equations?

In certain cases, such as when a system of equations is large, iterative methods of solving equations such as Gauss-Siedel method are more advantageous. Elimination methods such as Gaussan elimination, are prone to round off errors for a large set of equations. Iterative methods, such as Guass-Siedel method, allow the user the control of the round-off error. Also if the physics of the problem are well known for faster convergence, initial guesses needed in iterative methods can be made more judiciously.

You convinced me, so what is the algorithm for Gauss-Siedel method? Given a general set of *n* equations and *n* unknowns, we have

$$a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + ... + a_{1n}x_n = b_1$$

$$a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + ... + a_{2n}x_n = b_2$$

$$\begin{matrix} \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \end{matrix}$$

$$a_{n1}x_1 + a_{n2}x_2 + a_{n3}x_3 + ... + a_{nn}x_n = b_n$$

If the diagonal elements are non-zero, each equation is rewritten for the corresponding unknown, that is, the first equation is rewritten with x_1 on the left hand side, second equation is rewritten with x_2 on the left hand side and so on as follows

$$x_1 = \frac{c_1 - a_{12}x_2 - a_{13}x_3 \ldots\ldots - a_{1n}x_n}{a_{11}}$$

$$x_2 = \frac{c_2 - a_{21}x_1 - a_{23}x_3 \ldots\ldots - a_{2n}x_n}{a_{22}}$$

$$\vdots$$

$$\vdots$$

$$x_{n-1} = \frac{c_{n-1} - a_{n-1,1}x_1 - a_{n-1,2}x_2 \ldots\ldots - a_{n-1,n-2}x_{n-2} - a_{n-1,n}x_n}{a_{n-1,n-1}}$$

$$x_n = \frac{c_n - a_{n1}x_1 - a_{n2}x_2 - \ldots\ldots - a_{n,n-1}x_{n-1}}{a_{nn}}$$

These equations can be rewritten in the summation form as

$$x_1 = \frac{c_1 - \sum\limits_{\substack{j=1 \\ j\neq 1}}^{n} a_{1j}x_j}{a_{11}}$$

$$x_2 = \frac{c_2 - \sum\limits_{\substack{j=1 \\ j\neq 2}}^{n} a_{2j}x_j}{a_{22}}$$

.

.

.

$$x_{n-1} = \frac{c_{n-1} - \sum\limits_{\substack{j=1 \\ j\neq n-1}}^{n} a_{n-1,j}x_j}{a_{n-1,n-1}}$$

$$x_n = \frac{c_n - \sum_{\substack{j=1 \\ j \neq n}}^{n} a_{nj} x_j}{a_{nn}}$$

Hence for any row 'i',

$$x_i = \frac{c_i - \sum_{\substack{j=1 \\ j \neq i}}^{n} a_{ij} x_j}{a_{ii}}, i = 1,2,\ldots,n.$$

Now to find x_i's, one assumes an initial guess for the x_i's and then use the rewritten equations to calculate the new guesses. Remember, one always uses the most recent guesses to calculate x_i. At the end of each iteration, one calculates the absolute relative approximate error for each x_i as

$$|e_a|_i = \left| \frac{x_i^{new} - x_i^{old}}{x_i^{new}} \right| x100$$

where x_i^{new} is the recently obtained value of x_i, and x_i^{old} is the previous value of x_i.

When the absolute relative approximate error for each x_i is less than the prespecified tolerance, the iterations are stopped.

Example

The upward velocity of a rocket is given at three different times in the following table

Time, t	**Velocity, v**
S	m/s
5	106.8
8	177.2
12	279.2

The velocity data is approximated by a polynomial as

$$v(t) = a_1 t^2 + a_2 t + a_3 \,, \qquad 5 \le t \le 12.$$

The coefficients a_1, a_2, a_3 for the above expression were found in Chapter 5 to be given by

$$\begin{bmatrix} 25 & 5 & 1 \\ 64 & 8 & 1 \\ 144 & 12 & 1 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} 106.8 \\ 177.2 \\ 279.2 \end{bmatrix}$$

Find the values of a_1, a_2, a_3 using Guass-Siedal Method. Assume an initial guess of the solution as

$$\begin{bmatrix} a_1 & a_2 & a_3 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 5 \end{bmatrix}.$$

Solution

Rewriting the equations gives

$$a_1 = \frac{106.8 - 5a_2 - a_3}{25}$$

$$a_2 = \frac{177.2 - 64a_1 - a_3}{8}$$

$$a_3 = \frac{279.2 - 144a_1 - 12a_2}{1}$$

Iteration #1

Given the initial guess of the solution vector as

$$\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 5 \end{bmatrix}$$

we get

$$a_1 = \frac{106.8 - 5(2) - (5)}{25}$$

$$= 3.6720$$

$$a_2 = \frac{177.2 - 64(3.6720) - (5)}{8}$$

$= -7.8510$

$$a_3 = \frac{279.2 - 144(3.6720) - 12(-7.8510)}{1}$$

$= -155.36$

The absolute relative approximate error for each x_i then is

$$|e_a|_1 = \left|\frac{3.6720 - 1.0000}{3.6720}\right| x100$$

$=72.76\%$

$$|e_a|_2 = \left|\frac{-7.8510 - 2.0000}{-7.8510}\right| x100$$

$=125.47\%$

$$|e_a|_3 = \left|\frac{-155.36 - 5.0000}{-155.36}\right| x100$$

$= 103.22\%$

At the end of the first iteration, the guess of the solution vector is

$$\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} 3.6720 \\ -7.8510 \\ -155.36 \end{bmatrix}$$

and the maximum absolute relative approximate error is 125.47%.

Iteration #2

The estimate of the solution vector at the end of iteration #1 is

$$\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} 3.6720 \\ -7.8510 \\ -155.36 \end{bmatrix}$$

Now we get

$$a_1 = \frac{106.8 - 5(-7.8510) - 155.36}{25}$$

$= 12.056$

$$a_2 = \frac{177.2 - 64(12.056) - 155.36}{8}$$

$= -54.882$

$$a_3 = \frac{279.2 - 144(12.056) - 12(-54.882)}{1}$$

$= -798.34$

The absolute relative approximate error for each x_i then is

$$|\in_a|_1 = \left|\frac{12.056 - 3.6720}{12.056}\right| x 100$$

$= 69.542\%$

$$|\in_a|_2 = \left|\frac{-54.882 - (-7.8510)}{-54.882}\right| x 100$$

$= 85.695\%$

$$|\in_a|_3 = \left|\frac{-798.34 - (-155.36)}{-798.34}\right| x 100$$

$= 80.54\%$.

At the end of second iteration the estimate of the solution is

$$\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} 12.056 \\ -54.882 \\ -798.34 \end{bmatrix}$$

and the maximum absolute relative approximate error is 85.695%.

Conducting more iterations gives the following values for the solution vector and the corresponding absolute relative approximate errors.

Iteration	a_1	$\lvert\in_a\rvert_1$ %	a_2	$\lvert\in_a\rvert_2$ %	a_3	$\lvert\in_a\rvert_3$ %

1	3.672	72.767	-7.8510	125.47	-155.36	103.22
2	12.056	67.542	-54.882	85.695	-798.34	80.540
3	47.182	74.448	-255.51	78.521	-3448.9	76.852
4	193.33	75.595	-1093.4	76.632	-14440	76.116
5	800.53	75.850	-4577.2	76.112	-60072	75.962
6	3322.6	75.907	-19049	75.971	-249580	75.931

As seen in the above table, the solution is not converging to the true solution of

$a_1 = 0.29048$

$a_2 = 19.690$

$a_3 = 1.0858$

The above system of equations does not seem to converge? Why?

Well, a pitfall of most iterative methods is that they may or may not converge. However, certain class of systems of simultaneous equations do always converge to a solution using Gauss-Siedal method. This class of system of equations is where the coefficient matrix [A] in [A][X] = [C] is diagonally dominant, that is

$$|a_{ii}| \geq \sum_{\substack{j=1 \\ j\neq i}}^{n} |a_{ij}| \quad \text{for all 'i'}$$

and $|a_{ii}| \rangle \sum_{\substack{j=1 \\ j\neq i}}^{n} |a_{ij}|$ for at least one 'i'.

If a system of equations has a coefficient matrix that is not diagonally dominant, it may or may not converge. Fortunately, many physical systems that result in simultaneous linear equations have diagonally dominant coefficient matrices, which then assures convergence for iterative methods such as Gauss-Siedal method of solving simultaneous linear equations.

Example

Given the system of equations.

$$12x_1 + 3x_2 - 5x_3 = 1$$

$$x_1 + 5x_2 + 3x_3 = 28$$

$$3x_1 + 7x_2 + 13x_3 = 76$$

find the solution. Given

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

as the initial guess.

Solution

The coefficient matrix

$$[A] = \begin{bmatrix} 12 & 3 & -5 \\ 1 & 5 & 3 \\ 3 & 7 & 13 \end{bmatrix}$$

is diagonally dominant as

$$|a_{11}| = |12| = 12 \geq |a_{12}| + |a_{13}| = |3| + |-5| = 8$$

$$|a_{22}| = |5| = 5 \geq |a_{21}| + |a_{23}| = |1| + |3| = 4$$

$$|a_{33}| = |13| = 13 \geq |a_{31}| + |a_{32}| = |3| + |7| = 10$$

and the inequality is strictly greater than for at least one row. Hence the solution should converge using Gauss-Siedal method.

Rewriting the equations, we get

$$x_1 = \frac{1 - 3x_2 + 5x_3}{12}$$

$$x_2 = \frac{28 - x_1 - 3x_3}{5}$$

$$x_3 = \frac{76 - 3x_1 - 7x_2}{13}$$

Assuming an initial guess of

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

<u>Iteration 1:</u>

$$x_1 = \frac{1 - 3(0) + 5(1)}{12}$$

$$= \quad 0.50000$$

$$x_2 = \frac{28 - (0.5) - 3(1)}{5}$$

$$= \quad 4.9000$$

$$x_3 = \frac{76 - 3(0.50000) - 7(4.9000)}{13}$$

$$= \quad 3.0923$$

The absolute relative approximate error at the end of first iteration is

$$|\in_a|_1 = \left|\frac{0.50000 - 1.0000}{0.50000}\right| x100$$

$$= \; 67.662\%$$

$$|\in_a|_2 = \left|\frac{4.9000 - 0}{4.9000}\right| x100$$

$$= \; 100.000\%$$

$$|\in_a|_3 = \left|\frac{3.0923 - 1.0000}{3.0923}\right| x100$$

$$= 67.662\%$$

The maximum absolute relative approximate error is 100.000%

<u>Iteration 2:</u>

$$x_1 = \frac{1 - 3(4.9000) + 5(3.0923)}{12}$$

$$= \quad 0.14679$$

$$x_2 = \frac{28 - (0.14679) - 3(3.0923)}{5}$$

$$= \quad 3.7153$$

$$x_3 = \frac{76 - 3(0.14679) - 7(4.900)}{13}$$

$$= \quad 3.8118$$

At the end of second iteration, the absolute relative approximate error is

$$\left|\in_a\right|_1 = \left|\frac{0.14679 - 0.50000}{0.14679}\right| x100$$

$$= 240.62\%$$

$$\left|\in_a\right|_2 = \left|\frac{3.7153 - 4.9000}{3.7153}\right| x100$$

$$= 31.887\%$$

$$\left|\in_a\right|_3 = \left|\frac{3.8118 - 3.0923}{3.8118}\right| x100$$

$$= 18.876\%.$$

The maximum absolute relative approximate error is 240.62%. This is greater than the value of 67.612% we obtained in the first iteration. Is the solution diverging? No, as you conduct more iterations, the solution converges as follows.

Iteration	a_1	$\left|e_a\right|_1$	a_2	$\left|e_a\right|_2$	a_3	$\left|e_a\right|_3$
1	0.50000	67.662	4.900	100.00	3.0923	67.662

2	0.14679	240.62	3.7153	31.887	3.8118	18.876
3	0.74275	80.23	3.1644	17.409	3.9708	4.0042
4	0.94675	21.547	3.0281	4.5012	3.9971	0.65798
5	0.99177	4.5394	3.0034	0.82240	4.0001	0.07499
6	0.99919	0.74260	3.0001	0.11000	4.0001	0.00000

This is close to the exact solution vector of

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ 4 \end{bmatrix}$$

Example

Given the system of equations

$$3x_1 + 7x_2 + 13x_3 = 76$$

$$x_1 + 5x_2 + 3x_3 = 28$$

$$12x_1 + 3x_2 - 5x_3 = 1$$

find the solution using Gauss-Siedal method. Use $[x_1, x_2, x_3] = [1 \quad 0 \quad 1]$ as the initial guess.

Solution

Rewriting the equations, we get

$$x_1 = \frac{76 - 7x_2 - 13x_3}{3}$$

$$x_2 = \frac{28 - x_1 - 3x_3}{5}$$

$$x_3 = \frac{1 - 12x_1 - 3x_2}{-5}$$

Assuming an initial guess of

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

the next six iterative values are given in the table below

Iteration	a_1	$\left\vert \in_a \right\vert_1$	a_2	$\left\vert \in_a \right\vert_2$	a_3	$\left\vert \in_a \right\vert_3$
1	21.000	110.71	0.80000	100.00	5.0680	98.027
2	-196.15	109.83	14.421	94.453	-462.30	110.96
3	-1995.0	109.90	-116.02	112.43	4718.1	109.80
4	-20149	109.89	1204.6	109.63	-47636	109.90
5	$2.0364\text{x}10^5$	109.90	-12140	109.92	$4.8144\text{x}10^5$	109.89
6	$-2.0579\text{x}10^5$	1.0990	$1.2272\text{x}10^5$	109.89	$-4.8653\text{x}10^6$	109.89

You can see that this solution is not converging and the coefficient matrix is not diagonally dominant. The coefficient matrix

$$[A] = \begin{bmatrix} 3 & 7 & 13 \\ 1 & 5 & 3 \\ 12 & 3 & -5 \end{bmatrix}$$

is not diagonally dominant as

$$|a_{11}| = |3| = 3 \leq |a_{12}| + |a_{13}| = |7| + |13| = 20$$

Hence Gauss-Siedal method may or may not converge.

However, it is the same set of equations as the previous example and that converged. The only difference is that we exchanged first and the third equation with each other and that made the coefficient matrix not diagonally dominant.

So it is possible that a system of equations can be made diagonally dominant if one exchanges the equations with each other. But it is not possible for all cases. For example, the following set of equations.

$$x_1 + x_2 + x_3 = 3$$

$$2x_1 + 3x_2 + 4x_3 = 9$$

$$x_1 + 7x_2 + x_3 = 9$$

can not be rewritten to make the coefficient matrix diagonally dominant.

Homework

1. In a system of equation [A] [X] = [C], if [A] is diagonally dominant, then Gauss-Siedal method

   A. always converges

   B. may or may not converge

   C. always diverges

2. In a system of equations [A] [X] = [C], if [A] is not diagonally dominant, then Gauss-Siedal method

   A. Always converges

   B. May or may not converge

   C. Always diverges.

3. In a system of equations [A] [X] = [C], if [A] is not diagonally dominant, the system of equations can always be rewritten to make it diagonally dominant.

   A. True

   B. False

4. Solve the following system of equations using Gauss-Siedal method

   $$12x_1 + 7x_2 + 3x_3 = 2$$

   $$x_1 + 5x_2 + x_3 = -5$$

   $$2x_1 + 7x_2 - 11x_3 = 6$$

Conduct 3 iterations, calculate the maximum absolute relative approximate error at the end of each iteration and choose $\begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} = \begin{bmatrix} 1 & 3 & 5 \end{bmatrix}$ as your initial guess.

Answer: $\begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} = \begin{bmatrix} 0.90666 & -1.0115 & -1.0243 \end{bmatrix}$

$$\begin{bmatrix} |\in_a|_1 & |\in_a|_2 & |\in_a|_3 \end{bmatrix} = \begin{bmatrix} 65.001\% & 10.564\% & 17.099\% \end{bmatrix}$$

5. Solve the following system of equations using Gauss-Siedal Method

$$x_1 + 5x_2 + x_3 = 5$$

$$12x_1 + 7x_2 + 3x_3 = 2$$

$$2x_1 + 7x_2 - 11x_3 = 6$$

Conduct 3 iterations, calculate the maximum absolute relative approximate error at the end of each iteration and choose $\begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} = \begin{bmatrix} 1 & 3 & 5 \end{bmatrix}$ as your initial guess.

Answer: $\begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} = \begin{bmatrix} -1163.7 & 1947.6 & 1027.2 \end{bmatrix}$

$$\begin{bmatrix} |\in_a|_1 & |\in_a|_2 & |\in_a|_3 \end{bmatrix} = \begin{bmatrix} 89.156\% & 89.139\% & 89.183\% \end{bmatrix}$$

Chapter 9
Adequacy of Solutions

After reading this chapter, you will be able to

- *Know the difference between ill conditioned and well conditioned system of equations*
- *Define the norm of a matrix*
- *Relate the norm of the matrix and of its inverse to the ill or well conditioning of the matrix, that is, how much trust can you have in the solution of the matrix.*

What does it mean by ill conditioned and well-conditioned system of equations?

A system of equations is considered to be **well conditioned** if a small change in the coefficient matrix or a small change in the right hand side results in a small change in the solution vector.

A system of equations is considered to be **ill conditioned** if a small change in the coefficient matrix or a small change in the right hand side results in a large change in the solution vector.

Example

Is this system of equations well conditioned?

$$\begin{bmatrix} 1 & 2 \\ 2 & 3.999 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 4 \\ 7.999 \end{bmatrix}$$

Solution

The solution to the above set of equations is

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

Make a small change in the right hand side vector of the equations

$$\begin{bmatrix} 1 & 2 \\ 2 & 3.999 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 4.001 \\ 7.998 \end{bmatrix}$$

gives

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -3.999 \\ 4.000 \end{bmatrix}$$

Make a small change in the coefficient matrix of the equations

$$\begin{bmatrix} 1.001 & 2.001 \\ 2.001 & 3.998 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 4 \\ 7.999 \end{bmatrix}$$

gives

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 3.994 \\ 0.001388 \end{bmatrix}$$

This system of equation “looks” ill conditioned as a small change in the coefficient matrix or the right hand side resulted in a large change in the solution vector.

Example

Is this system of equations well conditioned?

$$\begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 4 \\ 7 \end{bmatrix}$$

Solution

The solution to the above equations is

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

Make a small change in the right hand side vector of the equations.

$$\begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 4.001 \\ 7.001 \end{bmatrix}$$

gives

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1.999 \\ 1.001 \end{bmatrix}$$

Make a small change in the coefficient matrix of the equations.

$$\begin{bmatrix} 1.001 & 2.001 \\ 2.001 & 3.001 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 4 \\ 7 \end{bmatrix}$$

gives

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2.003 \\ 0.997 \end{bmatrix}$$

This system of equation "looks" well conditioned as small changes in the coefficient matrix or the right hand side resulted in small changes in the solution vector.

So what if the system of equations is ill conditioning or well conditioning?

Well, if a system of equations is ill conditioned, we cannot trust the solution as much. Remember the velocity problem in Chapter 5. The values in the coefficient matrix are squares of time, etc. For example if instead of $a_{11} = 25$, you used $a_{11} = 24.99$, would you want it to make a huge difference in the solution vector. If it did, would you trust the solution?

Later we will see how much (quantifiable terms) we can trust the solution to a system of equations. Every invertible square matrix has a **condition number** and coupled with the **machine epsilon**, we can quantify how many significant digits one can trust in the solution.

To calculate condition number of an invertible square matrix, I need to know what norm of a matrix means. How is the norm of a matrix defined?

Just like the determinant, the norm of a matrix is a simple unique scalar number. However, norm is always positive and is defined for all matrices – square or rectangular; invertible or noninvertible square matrices.

One of the popular definitions of a norm is the row sum norm (also called uniform-matrix norm). For a mxn matrix [A], the row sum norm of [A] is defined as

$$\|A\|_\infty = \max_{1 \le i \le m} \sum_{j=1}^{n} |a_{ij}|$$

that is, find the sum of the absolute of the elements of each row of the matrix [A]. The maximum out of the 'm' such values is the row sum norm of the matrix [A].

Example

Find the row sum norm of the following matrix [A].

$$A = \begin{bmatrix} 10 & -7 & 0 \\ -3 & 2.099 & 6 \\ 5 & -1 & 5 \end{bmatrix}$$

Solution

$$\|A\|_\infty = \max_{1 \le i \le 3} \sum_{j=1}^{3} |a_{ij}|$$

$$\max\left[\left(|10| + |-7| + |0|\right), \left(|-3| + |2.099| + |6|\right), \left(|5| + |-1| + |5|\right)\right]$$

$$= \max\left[(10 + 7 + 0), (3 + 2.099 + 6), (5 + 1 + 5)\right]$$

$$= \max\left[17, 11.099, 11\right]$$

$$= 17.$$

How is norm related to the conditioning of the matrix?

Let us start answering this question using an example. Go back to the "ill conditioned" system of equations,

$$\begin{bmatrix} 1 & 2 \\ 2 & 3.999 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 4 \\ 7.999 \end{bmatrix}$$

that gives the solution as

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

Denoting the above set of equations as

$$[A][X] = [C]$$

$$\|X\|_\infty = 2$$

$$\|C\|_\infty = 7.999$$

Making a small change in the right hand side,

$$\begin{bmatrix} 1 & 2 \\ 2 & 3.999 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 4.001 \\ 7.998 \end{bmatrix}$$

gives

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -3.999 \\ 4.000 \end{bmatrix}$$

Denoting the above set of equations by

$$[A][X'] = [C']$$

and the change in right hand side vector

$$[\Delta C] = [C'] - [C]$$

and the change in the solution vector as

$$[\Delta X] = [X'] - [X]$$

then

$$[\Delta C] = \begin{bmatrix} 4.001 \\ 7.998 \end{bmatrix} - \begin{bmatrix} 4 \\ 7.999 \end{bmatrix}$$

$$= \begin{bmatrix} 0.001 \\ -0.001 \end{bmatrix}$$

and

$$[\Delta X] = \begin{bmatrix} -3.999 \\ 4.000 \end{bmatrix} - \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} -5.999 \\ 3.000 \end{bmatrix}$$

then

$$\|\Delta C\|_\infty = 0.001$$

$$\|\Delta X\|_\infty = 5.999$$

Relative change in the norm of the solution vector is

$$\frac{\|\Delta X\|_\infty}{\|X\|_\infty} = \frac{5.999}{2}$$

$$= 2.9995$$

Relative change in the norm of the right hand side vector is

$$\frac{\|\Delta C\|_\infty}{\|C\|_\infty} = \frac{0.001}{7.999}$$

$$= 1.250 \times 10^{-4}$$

See the small relative change in the right hand side vector of 1.250 x 10^{-4} results in a large relative change in the solution vector as 2.995.

In fact, the ratio between the relative change in the norm of the solution vector to the relative change in the norm of the right hand side vector is

$$\frac{\|\Delta X\|_\infty / \|X\|_\infty}{\|\Delta C\|_\infty / \|C\|_\infty}$$

$$= \frac{2.9995}{1.250 \times 10^{-4}}$$

$$= 23957$$

Let us now go back to the "well-conditioned" system of equations.

$$\begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 4 \\ 7 \end{bmatrix}$$

gives

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

Denoting the system of equations as

$$[A][X] = [C]$$

$$\|X\|_\infty = 2$$

$$\|C\|_\infty = 7$$

Making a small change in the right hand side vector

$$\begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 4.001 \\ 7.001 \end{bmatrix}$$

gives

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1.999 \\ 1.001 \end{bmatrix}$$

Denoting the above set of equations as

$$[A][X'] = [C']$$

and the change in the right hand side vector

$$[\Delta C] = [C'] - [C]$$

and the change in the solution vector as

$$[\Delta X] = [X'] - [X]$$

then

$$[\Delta C] = \begin{bmatrix} 4.001 \\ 7.001 \end{bmatrix} - \begin{bmatrix} 4 \\ 7 \end{bmatrix} = \begin{bmatrix} 0.001 \\ 0.001 \end{bmatrix}$$

and

$$[\Delta X] = \begin{bmatrix} 1.999 \\ 1.001 \end{bmatrix} - \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} -0.001 \\ 0.001 \end{bmatrix}$$

then

$\|\Delta C\|_\infty = 0.001$

$\|\Delta X\|_\infty = 0.001$

Relative change in the norm of solution vector is

$$\frac{\|\Delta X\|_\infty}{\|X\|_\infty}$$

$$= \frac{0.001}{2}$$

$$= 5 \times 10^{-4}$$

Relative change in the norm of the right hand side vector is

$$\frac{\|\Delta C\|_\infty}{\|C\|_\infty}$$

$$= \frac{0.001}{7}$$

$$=1.429 \times 10^{-4}$$

See a small relative change in the right hand side vector norm of 1.429 x 10^{-4} results in a small relative change in the solution vector norm of 5 x 10^{-4}.

In fact, the ratio between the relative change in the norm of the solution vector to the relative change in the norm of the right hand side vector is

$$\frac{\|\Delta X\|_\infty / \|X\|_\infty}{\|\Delta C\|_\infty / \|C\|_\infty}$$

$$= \frac{5 \times 10^{-4}}{1.429 \times 10^{-4}}$$

$$=3.5$$

What are some of the properties of norms?

1. For a matrix [A], $\|A\| \geq 0$
2. For a matrix [A] and a scalar k, $\|kA\| = |k|\|A\|$

3. For two matrices [A] and [B] of same order,

$$\|A+B\| \le \|A\| + \|B\|$$

4. For two matrices [A] and [B] that can be multiplied as [A] [B],

$$\|AB\| \le \|A\| \, \|B\|$$

Is there a general relationship that exists between $\|\Delta X\|/\|X\|$ and $\|\Delta C\|/\|C\|$ **or between** $\|\Delta X\|/\|X\|$ **and** $\|\Delta A\|/\|A\|$**? If so, it could help us identify well-conditioned and ill conditioned system of equations.**

If there is such a relationship, will it help us quantify the conditioning of the matrix, that is, tell us how many significant digits we could trust in the solution of a system of simultaneous linear equations?

There is a relationship that exists between

$\dfrac{\|\Delta X\|}{\|X\|}$ and $\dfrac{\|\Delta C\|}{\|C\|}$, and between

$\dfrac{\|\Delta X\|}{\|X\|}$ and $\dfrac{\|\Delta A\|}{\|A\|}$.

These relationships are

$$\frac{\|\Delta X\|}{\|X+\Delta X\|} \le \|A\| \, \|A^{-1}\| \frac{\|\Delta C\|}{\|C\|}.$$

$$\frac{\|\Delta X\|}{\|X\|} \le \|A\| \, \|A^{-1}\| \frac{\|\Delta A\|}{\|A\|}$$

Looking at the above two inequalities, it shows that the relative change in the norm of the right hand side vector or the coefficient matrix can be amplified by as much as $\|A\| \, \|A^{-1}\|$.

This number $\|A\| \, \|A^{-1}\|$ is called the **condition number** of the matrix and coupled with the machine epsilon, we can quantify the accuracy of the solution of [A][X] = [C].

Prove for $[A][X]=[C]$

$$\frac{\|\Delta X\|}{\|X+\Delta X\|} \leq \|A\|\,\|A^{-1}\|\frac{\|\Delta A\|}{\|A\|}.$$

Proof

Let

$$[A][X]=[C] \qquad (1)$$

Then if [A] is changed to $[A']$, then[X] will change to $[X']$, such

that

$$[A'][X']=[C] \qquad (2)$$

From equations (1) and (2),

$$[A][X]=[A'][X']$$

Denoting change in [A] and [X] matrices as

$$[\Delta A]=[A']-[A]$$

$$[\Delta X]=[X']-[X]$$

then

$$[A][X]=([A]+[\Delta A])([X]+[\Delta X])$$

Expanding the above expression

$$[A][X]=[A][X]+[A][\Delta X]+[\Delta A][X]+[\Delta A][\Delta X]$$

$$0=[A][\Delta X]+[\Delta A]([X]+[\Delta X])$$

$$-[A][\Delta X]=[\Delta A]([X]+[\Delta X])$$

$$[\Delta X]=-[A]^{-1}[\Delta A]([X]+[\Delta X])$$

Applying the theorem of norms that norm of multiplied matrices is less than the multiplication of the individual norms of the matrices,

$$\|\Delta X\| \leq \|A^{-1}\|\|\Delta A\|\|X+\Delta X\|$$

Multiplying both sides by $\|A\|$

$$\|A\|\|\Delta X\| \le \|A\|\|A^{-1}\|\|\Delta A\|\|X+\Delta X\|$$

$$\frac{\|\Delta X\|}{\|X+\Delta X\|} \le \|A\|\|A^{-1}\|\frac{\|\Delta A\|}{\|A\|}$$

How do I use the above theorems to find how many significant digits are correct in my solution vector?

Relative error in solution vector is <= Cond (A)* relative error in right hand side.

Possible relative error in the solution vector is <= Cond (A) * $\in_{mach}$

Hence Cond (A)*$\in_{mach}$ should give us the number of significant digits at least correct in our solution by comparing it with $0.5\text{x}10^{-m}$.

Example

How many significant digits can I trust in the solution of the following system of equations?

$$\begin{bmatrix} 1 & 2 \\ 2 & 3.999 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$$

Solution

$$[A] = \begin{bmatrix} 1 & 2 \\ 2 & 3.999 \end{bmatrix}$$

$$[A]^{-1} = \begin{bmatrix} -3999.31 & 2000.1 \\ 2000.1 & -1000.1 \end{bmatrix}$$

$$\|A\|_\infty = 5.999$$

$$\|A^{-1}\|_\infty = 5999.4$$

$$\text{Cond}(A) = \|A\|_\infty \|A^{-1}\|_\infty$$

$$= 5.999 \times 5999.4$$

$= 35990$

Assuming single precision with 24 bits used in the mantissa for real numbers, the machine epsilon

$\in_{mach} = 2^{1-24}$

$= 0.119209 \times 10^{-6}$

$Cond(A) * \in_{mach}$

$= 35990 \times 0.119209 \times 10^{-6}$

$= 0.4290 x 10^{-2}$

Comparing it with 0.5×10^{-m}

$0.5 \times 10^{-m} < 0.4290 \times 10^{-2}$

$m \leq 2$

So two significant digits are at least correct in the solution vector.

Example

How many significant digits can I trust in the solution of the following system of equations?

$$\begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 4 \\ 7 \end{bmatrix}$$

Solution

For

$$[A] = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}$$

it can be shown

$$[A]^{-1} = \begin{bmatrix} -3 & 2 \\ 2 & -1 \end{bmatrix}$$

Then

$\|A\|_\infty = 5$,

$\|A^{-1}\|_\infty = 5$.

Cond (A) = $\|A\|_\infty \|A^{-1}\|_\infty$

$=5\times 5$

$=25$

Assuming single precision with 24 bits of mantissa for real numbers, the machine epsilon

$\in_{\text{mach}} = 2^{1-24}$

$= 0.119209\times 10^{-6}$

$Cond(A)^* \in_{mach}$

$= 25\times 0.119209\times 10^{-6}$

$= 0.2980 x 10^{-5}$

Comparing it with 0.5 x 10^{-m}

$0.5\times 10^{-m} < 0.2980\times 10^{-5}$

$m \leq 5$

So five significant digits are at least correct in the solution vector.

Key Terms

Ill conditioned	Well conditioned
Norm	Condition number
Machine epsilon	Significant digits

Homework

1. The adequacy of the solution of simultaneous linear equations depends on

A. Condition number

B. Machine epsilon

C. Product of condition number and machine epsilon

D. Norm of the matrix.

2. If a system of equations [A] [X] = [C] is ill conditioned, then

A. det(A) = 0

B. Cond(A) = 1

C. Cond(A) is large.

D. $\|A\|$ is large.

3. If Cond(A) = 10^4 and $\in_{mach} = 0.119 \times 10^{-6}$, then in [A] [X] = [C], at least these many significant digits are correct in your solution,

A. 3

B. 2

C. 1

D. 0

4. Make a small change in the coefficient matrix to

$$\begin{bmatrix} 1 & 2 \\ 2 & 3.999 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 4 \\ 7.999 \end{bmatrix}$$

and find

$$\frac{\|\Delta X\|_\infty / \|X\|_\infty}{\|\Delta A\|_\infty / \|A\|_\infty}$$

Is it a large or small number? How is this number related to the condition number of the matrix?

5. Make a small change in the coefficient matrix to

$$\begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 4 \\ 7 \end{bmatrix}$$

and find

$$\frac{\left\|\Delta X\right\|_{\infty} / \left\|X\right\|_{\infty}}{\left\|\Delta A\right\|_{\infty} / \left\|A\right\|_{\infty}}$$

Is it a large or a small number? Compare your results with the previous problem. How is this number related to the condition number of the matrix?

6. Prove

$$\frac{\left\|\Delta X\right\|}{\left\|X\right\|} \le \left\|A\right\|\left\|A^{-1}\right\|\frac{\left\|\Delta C\right\|}{\left\|C\right\|}$$

7. For

$$[A] = \begin{bmatrix} 10 & -7 & 0 \\ -3 & 2.099 & 6 \\ 5 & -1 & 5 \end{bmatrix}$$

gives

$$[A]^{-1} = \begin{bmatrix} -0.1099 & -0.2333 & 0.2799 \\ -0.2999 & -0.3332 & 0.3999 \\ 0.04995 & 0.1666 & 6.664\times 10^{-5} \end{bmatrix}$$

a) What is the condition number of [A]?

b) How many significant digits can we at least trust in the solution of $[A][X] = [C]$ if

$\in_{mach} = 0.1192\times 10^{-6}$

c)Without calculating the inverse of the matrix [A], can you estimate the condition number of [A] using the theorem in Problem#6?

Answer: a) $\left\|A\right\| = 17$

$\left\|A^{-1}\right\| = 1.033$

Cond (A) = 17.56

b) 5

c) Try different values of right hand side of C= [±1   ±1   ±1]^T with signs chosen randomly. Then $||A^{-1}|| = ||X||$ obtained from solving equation set [A] [X] = [C] as $||C|| =1$.

8. Prove that the Cond(A) >=1.

Hint:

We know that

$||A\ B|| = ||A||\ ||B||$

then if $[B] = [A]^{-1}$,

$||\ A\ A^{-1}|| = ||A||\ ||A^{-1}||$

$||I|| = ||A||\ ||A^{-1}||$

$1 = ||A||\ ||A^{-1}||$

$||A||\ ||A^{-1}|| \geq 1$

$\text{Cond}\ (A) \geq 1.$

Chapter 10
Eigenvalues and Eigenvectors

After reading this chapter, you will be able to

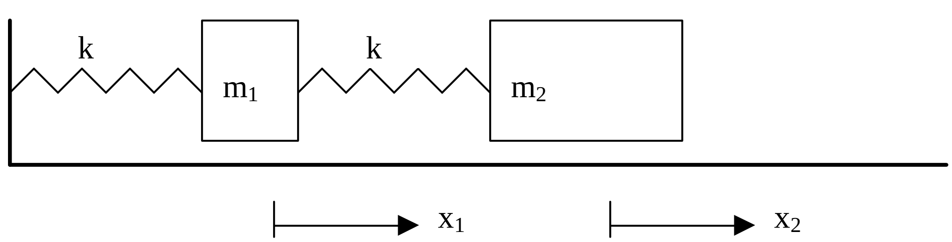
1. *Know the definition of eigenvalues and eigenvectors of a square matrix*
2. *Find eigenvalues and eigenvectors of a square matrix*
3. *Relate eigenvalues to the singularity of a square matrix*
4. *Use the power method to numerically find in magnitude the largest eigenvalue of a square matrix and the corresponding eigenvector.*

What does eigenvalue mean?

The word eigenvalue comes from the German word "Eigenwert" where Eigen means "characteristic" and Wert means "value". But what the word means is not on your mind! You want to know why do I need to learn about eigenvalues and eigenvectors. Once I give you an example of the application of eigenvalues and eigenvectors, you will want to know how to find these eigenvalues and eigenvectors. That is the motive of this chapter of the linear algebra primer.

Give me a physical example application of eigenvalues and eigenvectors?

Look at the spring-mass system as shown in the picture below.



Assume each of the two mass-displacements to be denoted by x_1 and x_2, and let us assume each spring has the same spring constant 'k'. Then by applying Newton's 2nd and 3rd law of motion to develop a force-balance for each mass we have

$$m_1 \frac{d^2 x_1}{dt^2} = -kx_1 + k(x_2 - x_1)$$

$$m_2 \frac{d^2 x_2}{dt^2} = -k(x_2 - x_1)$$

Rewriting the equations, we have

$$m_1 \frac{d^2 x_1}{dt^2} - k(-2x_1 + x_2) = 0$$

$$m_2 \frac{d^2 x_2}{dt^2} - k(x_1 - x_2) = 0$$

$$Let\ m_1 = 10, m_2 = 20, k = 15$$

$$10 \frac{d^2 x_1}{dt^2} - 15(-2x_1 + x_2) = 0$$

$$20 \frac{d^2 x_2}{dt^2} - 15(x_1 - x_2) = 0$$

From vibration theory, the solutions can be of the form

$$x_i = A_i Sin\ (wt - \emptyset)$$

where

A_i = amplitude of the vibration of mass 'i'

w = frequency of vibration

$\emptyset$ = phase shift

then

$$\frac{d^2 x_i}{dt^2} = -A_i w^2 Sin(wt - \emptyset)$$

Substituting x_i and $\frac{d^2 x_i}{dt^2}$ in equations,

$-10A_1 w^2 - 15(-2A_1 + A_2) = 0$

$-20A_2 w^2 - 15(A_1 - A_2) = 0$

gives

$(-10w^2 + 30)A_1 - 15A_2 = 0$

$-15A_1 + (-20w^2 + 15)A_2 = 0$

or

$(-w^2 + 3)A_1 - 1.5A_2 = 0$

$-0.75A_1 + (-w^2 + 0.75)A_2 = 0$

In matrix form, these equations can be rewritten as

$$\begin{bmatrix} -w^2+3 & -1.5 \\ -0.75 & -w^2+0.75 \end{bmatrix}\begin{bmatrix} A_1 \\ A_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} 3 & -1.5 \\ -0.75 & 0.75 \end{bmatrix}\begin{bmatrix} A_1 \\ A_2 \end{bmatrix} - w^2\begin{bmatrix} A_1 \\ A_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Let $w^2 = \lambda$

$$[A] = \begin{bmatrix} 3 & -1.5 \\ -0.75 & 0.75 \end{bmatrix}$$

$$[X] = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix}$$

$[A]\,[X] - \lambda\,[X] = 0$

$[A]\,[X] = \lambda\,[X]$

In the above equation, 'λ' is the eigenvalue and [X] is the eigenvector corresponding to λ. As you can see that if we know 'λ' for the above example, we can calculate the natural frequency of the vibration

$w = \sqrt{l}$

Why are they important? Because you do not want to have a forcing force on the spring-mass system close to this frequency as it would make the amplitude A_i very large and make the system unstable.

What is the general definition of eigenvalues and eigenvectors of a square matrix?

If [A] is a nxn matrix, then $[X] \neq \vec{0}$ is an eigenvector of [A] if

$[A][X] = \lambda[X]$

where λ is a scalar and [X]… 0. The scalar λ is called the eigenvalue of [A] and [X] is called the eigenvector corresponding to the eigenvalue l .

How do I find eigenvalues of a square matrix?

To find the eigenvalues of a nxn matrix [A], we have

$[A][X] = \lambda[X]$

$[A][X] - \lambda[X] = 0$

$[A][X] - \lambda[I][X] = 0$

$([A] - \lambda[I])[X] = 0$

Now for the above set of equations to have a nonzero solution,

$\det([A] - \lambda[I]) = 0$

This left hand side can be expanded to give a polynomial in λ solving the above equation would give us values of the eigenvalues. The above equation is called the characteristic equation of [A].

For a nxn [A] matrix, the characteristic polynomial of A is of degree n as follows

$\det([A] - \lambda[I]) = 0$

gives

$?^n + c_1 \lambda^{n-1} + - - + c_n = 0$

Hence this polynomial can have n roots.

Example

Find the eigenvalues of the physical problem discussed in the beginning of this chapter, that is, find the eigenvalues of the matrix

$$[A]=\begin{bmatrix} 3 & -1.5 \\ -0.75 & 0.75 \end{bmatrix}$$

Solution

$$[A]-l[I]=\begin{bmatrix} 3-l & -1.5 \\ -0.75 & 0.75-l \end{bmatrix}$$

$$\det([A]-l[I])=(3-l)(0.75-l)-(-0.75)(-1.5)=0$$

$$2.25-0.75l-3l+l^2-1.125=0$$

$$l^2-3.75l+1.125=0$$

$$l=\frac{-(-3.75)\pm\sqrt{(-3.75)^2-4(1)(1.125)}}{2(1)}$$

$$=\frac{3.75\pm 3.092}{2}$$

$$=3.421,0.3288$$

So the eigenvalues are 3.421 and 0.3288.

Example

Find the eigenvectors of

$$A=\begin{bmatrix} 3 & -1.5 \\ -0.75 & 0.75 \end{bmatrix}$$

Solution

The eigenvalues have already been found in the previous example

$l_1 = 3.421, l_2 = 0.3288$

Let

$$[X] = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

be the eigenvector corresponding to

$l_1 = 3.421$

Hence

$([A] - \lambda_1 [I])[X] = 0$

$$\left\{ \begin{bmatrix} 3 & -1.5 \\ -0.75 & 0.75 \end{bmatrix} - 3.421 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right\} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0$$

$$\begin{bmatrix} -0.421 & -1.5 \\ -0.75 & -2.671 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

If

$x_1 = s$

Then

$-0.421s - 1.5x_2 = 0$

$x_2 = -0.2807\ s$

The eigenvector corresponding to $l_1 = 3.421$ then is

$$[X] = \begin{bmatrix} s \\ -0.2807s \end{bmatrix} = s \begin{bmatrix} 1 \\ -0.2807 \end{bmatrix}.$$

The eigenvector corresponding to $l_1 = 3.421$ is $\begin{bmatrix} 1 \\ -0.2807 \end{bmatrix}$

Similarly, the eigenvector corresponding to $l_2 = 0.3288$ is

$$\begin{bmatrix} 1 \\ 1.781 \end{bmatrix}$$

Example

Find the eigenvalues and eigenvectors of

$$[A] = \begin{bmatrix} 1.5 & 0 & 1 \\ -0.5 & 0.5 & -0.5 \\ -0.5 & 0 & 0 \end{bmatrix}$$

Solution

The characteristic equation is given by

det ($[A]$ - $\lambda[I]$) = 0

$$\det \begin{bmatrix} 1.5-l & 0 & 1 \\ -0.5 & 0.5-l & -0.5 \\ -0.5 & 0 & -l \end{bmatrix} = 0$$

$$(1.5-l)[(0.5-l)(-l)-(-0.5)(0)]+(1)[(-0.5)(0)-(-0.5)(0.5-l)]=0$$

$$-l^3+2l^2-1.25l+0.25=0$$

The roots of the above equation are

$$l = 0.5, 0.5, 1.0$$

Note that there are eigenvalues that are repeated. Since there are only two distinct eigenvalues, there are only two eigenspaces. But corresponding to λ = 0.5 there should be two eigenvectors that form a basis for the eigenspace.

To find the eigenspaces, let

$$[X] = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

Given $[(A-lI)][X]=0$

$$\begin{bmatrix} 1.5 - l & 0 & 1 \\ -0.5 & 0.5 - l & -0.5 \\ -0.5 & 0 & -l \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

For $\lambda = 0.5$,

$$\begin{bmatrix} 1 & 0 & 1 \\ -0.5 & 0 & -0.5 \\ -0.5 & 0 & -0.5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Solving this system gives

$x_1 = a$, $x_2 = b$, $x_3 = -a$

So

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} a \\ b \\ -a \end{bmatrix} = \begin{bmatrix} a \\ 0 \\ -a \end{bmatrix} + \begin{bmatrix} 0 \\ b \\ 0 \end{bmatrix}$$

$$= a\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} + b\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

So the vectors $\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$ form a basis for the eigenspace for the eigenvalue $\lambda=0.5$.

For $\lambda = 1$,

$$\begin{bmatrix} 0.5 & 0 & 1 \\ -0.5 & -0.5 & -0.5 \\ -0.5 & 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Solving this system gives

$x_1 = a$, $x_2 = -0.5a$, $x_3 = -0.5a$

The eigenvector corresponding to $\lambda = 1$ is

$$\begin{bmatrix} a \\ -0.5a \\ -0.5a \end{bmatrix} = a\begin{bmatrix} 1 \\ -0.5 \\ -0.5 \end{bmatrix}$$

Hence the vector $\begin{bmatrix} 1 \\ -0.5 \\ -0.5 \end{bmatrix}$ is a basis for the eigenspace for the eigenvalue of $\lambda = 1$.

What are some of the theorems of eigenvalues and eigenvectors?

Theorem 1: If [A] is a nxn triangular matrix – upper triangular, lower triangular or diagonal, the eigenvalues of [A] are the diagonal entries of [A].

Theorem 2: l = 0 is an eigenvalue of [A] if [A] is a singular (noninvertible) matrix.

Theorem 3: [A] and $[A]^T$ have the same eigenvalues.

Theorem 4: Eigenvalues of a symmetric matrix are real.

Theorem 5: Eigenvectors of a symmetric matrix are orthogonal, but only for distinct eigenvalues.

Theorem 6: $|\det(A)|$ *is the product of the absolute values of the eigenvalues of [A].*

Example

What are the eigenvalues of

$$[A] = \begin{bmatrix} 6 & 0 & 0 & 0 \\ 7 & 3 & 0 & 0 \\ 9 & 5 & 7.5 & 0 \\ 2 & 6 & 0 & -7.2 \end{bmatrix}$$

Solution

Since the matrix [A] is a lower triangular matrix, the eigenvalues of [A] are the diagonal elements of [A]. The eigenvalues are

$$l_1 = 6, l_2 = 3, l_3 = 7.5, l_4 = -7.2$$

Example

One of the eigenvalues of

$$[A] = \begin{bmatrix} 5 & 6 & 2 \\ 3 & 5 & 9 \\ 2 & 1 & -7 \end{bmatrix}$$

is zero. Is [A] invertible?

Solution

$l = 0$ is an eigenvalue of [A], that implies [A] is singular and is not invertible.

Example

Given the eigenvalues of

$$[A] = \begin{bmatrix} 2 & -3.5 & 6 \\ 3.5 & 5 & 2 \\ 8 & 1 & 8.5 \end{bmatrix}$$

are

$$l_1 = -1.546, l_2 = 12.33, l_3 = 4.711$$

What are the eigenvalues if

$$[B] = \begin{bmatrix} 2 & 3.5 & 8 \\ -3.5 & 5 & 1 \\ 6 & 2 & 8.5 \end{bmatrix}$$

Solution

Since [B] = $[A]^T$, the eigenvalues of [A] and [B] are the same. Hence eigenvalues of [B] also are

$$-l_1 = 1.546,\ l_2 = 12.33,\ l_3 = 4.711$$

Example

Given the eigenvalues of

$$[A] = \begin{bmatrix} 2 & -3.5 & 6 \\ 3.5 & 5 & 2 \\ 8 & 1 & 8.0 \end{bmatrix}$$

are

$$l_1 = -1.546,\ l_2 = 12.33,\ l_3 = 4.711$$

Calculate the magnitude of the determinant of the matrix.

Solution

Since

$$\begin{aligned} |\det[A]| &= |l_1||l_2||l_3| \\ &= |-1.546||12.33||4.711| \\ &= 89.80 \end{aligned}$$

How does one find eigenvalues and eigenvectors numerically?

One of the most common methods used for finding eigenvalues and eigenvectors is the power method. It is used to find the largest eigenvalue in absolute sense. Note that if this largest eigenvalues is repeated, this method will not work. Also this eigenvalue needs to be distinct. The method is as follows:

1. Assume a guess $[X^{(0)}]$ for the eigenvector in

$[A][X] = \lambda[X]$

equation. One of the entries of $\left[X^{(0)}\right]$ needs to be unity.

2. Find

$[Y]^{(1)} = [A]\ [X^{(0)}]$

3. Scale $[Y^{(1)}]$ so that the chosen unity component remains unity.

$[Y^{(1)}] = \lambda^{(1)}\ [X^{(1)}]$

4. Repeat steps (2) and (3) with $[X] = [X^{(1)}]$ to get $\left[X^{(2)}\right]$.

5. Repeat the setups 2 and 3 until the value of the eigenvalue converges. If $\in_s$ is the pre-specified percentage relative error tolerance to which you would like the answer to converge to, keep iterating until

$$\left|\frac{l^{(i+1)} - l^{(i)}}{l^{(i+1)}}\right| \times 100 \leq e_s$$

where the left hand side of the above inequality is the definition of absolute percentage relative approximate error, denoted generally by $|e_a|$. A prespecified percentage relative tolerance of $0.5\times10^{2-m}$ implies at least 'm' significant digits are current in your answer. When the system converges, the value of λ is the largest (in absolute value) eigenvalue of [A].

Example

Using the power method, find the largest eigenvalue and the corresponding eigenvector of

$$[A] = \begin{bmatrix} 1.5 & 0 & 1 \\ -0.5 & 0.5 & -0.5 \\ -0.5 & 0 & 0 \end{bmatrix}$$

Solution

Assume

$$[X^{(0)}] = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

$$[A][X^{(0)}] = \begin{bmatrix} 1.5 & 0 & 1 \\ -0.5 & 0.5 & -0.5 \\ -0.5 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} 2.5 \\ -0.5 \\ -0.5 \end{bmatrix}$$

$$Y^{(1)} = 2.5 \begin{bmatrix} 1 \\ -0.2 \\ -0.2 \end{bmatrix}$$

$\lambda^{(1)} = 2.5$

We will choose the first element of $[X^{(0)}]$ to be unity.

$$[X^{(1)}] = \begin{bmatrix} 1 \\ -0.2 \\ -0.2 \end{bmatrix}$$

$$[A][X^{(1)}] = \begin{bmatrix} 1.5 & 0 & 1 \\ -0.5 & 0.5 & -0.5 \\ -0.5 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ -0.2 \\ -0.2 \end{bmatrix}$$

$$= \begin{bmatrix} 1.3 \\ -0.5 \\ -0.5 \end{bmatrix}$$

$$[Y^{(2)}] = 1.3 \begin{bmatrix} 1 \\ -0.3846 \\ -0.3846 \end{bmatrix}$$

$\lambda^{(2)} = 1.3$

$$[X^{(2)}] = \begin{bmatrix} 1 \\ -0.3846 \\ -0.3846 \end{bmatrix}$$

The absolute relative approximate error in the eigenvalues is

$$|\epsilon_a| = \left|\frac{l^{(2)} - l^{(1)}}{l^{(2)}}\right| x100$$

$$= \left|\frac{1.3-1.5}{1.3}\right| x100$$

=92.307%

Conducting further iterations, the values of $l^{(i)}$ and the corresponding eigenvectors is given in the table below

i	$l^{(i)}$	**[X$^{(i)}$]**	$\|\epsilon_a\|$ **(%)**
1	2.5	$\begin{bmatrix} 1 \\ -0.2 \\ -0.2 \end{bmatrix}$	———
2	1.3	$\begin{bmatrix} 1 \\ -0.3846 \\ -0.3846 \end{bmatrix}$	92.307
3	1.1154	$\begin{bmatrix} 1 \\ -0.44827 \\ -0.44827 \end{bmatrix}$	16.552
4	1.0517	$\begin{bmatrix} 1 \\ -0.47541 \\ -0.47541 \end{bmatrix}$	6.0529
5	1.02459	$\begin{bmatrix} 1 \\ -0.48800 \\ -0.48800 \end{bmatrix}$	1.2441

The exact value of the eigenvalue is

$$\lambda = 1$$

and the corresponding eigenvector is

$$[X] = \begin{bmatrix} 1 \\ -0.5 \\ -0.5 \end{bmatrix}$$

Key Terms

Eigenvalue   Eigenvectors   Power method

Homework

1. The eigenvalues ‘λ’ of matrix [A] are found by solving the equation(s)?
   a) $|A - lI| = 0$
   b) $[A][X] = [I]$
   c) $[A][X] - l[I] = \vec{0}$
   d) $|A| = 0$

2. Find the eigenvalues and eigenvectors of

$$[A] = \begin{bmatrix} 10 & 9 \\ 2 & 3 \end{bmatrix}$$

using the determinant method

Answer: (12,1), $\begin{bmatrix} 0.9762 \\ 0.2169 \end{bmatrix}$**,** $\begin{bmatrix} 0.8381 \\ 0.8381 \end{bmatrix}$

3. Find the eigenvalues and eigenvectors of

$$[A] = \begin{bmatrix} 4 & 0 & 1 \\ -2 & 0 & 1 \\ 2 & 0 & 1 \end{bmatrix}$$

using the determinant method

Answer: (0,4,5615,0.43845), $\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0.87193 \\ -0.27496 \\ 0.48963 \end{bmatrix}, \begin{bmatrix} -0.27816 \\ 3.5284 \\ 0.99068 \end{bmatrix}$

4. Find the eigenvalues of these matrices by inspection

$$\begin{bmatrix} 2 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 6 \end{bmatrix}$$

$$\begin{bmatrix} 3 & 5 & 7 \\ 0 & -2 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 2 & 0 & 0 \\ 3 & 5 & 0 \\ 2 & 1 & 6 \end{bmatrix}$$

Answer: a) 2,-3,6 b) 3,-2,0 c) 2,5,6

5. Prove if λ is an eigenvalue of [A], then $\frac{1}{l}$ is an eigenvalue of $[A]^{-1}$.

6. Prove that square matrices [A] and $[A]^T$ have the same eigenvalues.

7. Show that $|\det(A)|$ is the product of the absolute values of the eigenvalues of [A].

8. Find the largest eigenvalue in magnitude and its corresponding vector by using the power method

$$[A] = \begin{bmatrix} 4 & 0 & 1 \\ -2 & 0 & 1 \\ 2 & 0 & 1 \end{bmatrix}$$

Start with an initial guess of the eigenvector as $\begin{bmatrix} 1 \\ -0.5 \\ 0.5 \end{bmatrix}$

Answer:4.5615, $\begin{bmatrix} 1 \\ -0.31534 \\ 0.56154 \end{bmatrix}$ after 4 iterations