Elementary Matrix Theory

Definition

A matrix A of order (or size) $m \times n$ is a rectangular array of mn elements arranged in m rows and n columns. The element in the i^{th} row and j^{th} column is called the (i, j)-entry of the matrix and is denoted by a_{ij}.

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

An $n \times n$ matrix is called a square matrix of order n. A matrix of order $m \times 1$ is called a column vector, and one of order $1 \times n$ is called a row vector.

Here, consider matrices whose elements are, in general, complex numbers. Two matrices A and B are equal when their corresponding elements are equal i.e., $a_{ij} = b_{ij}$ for all i, j (which implies that A and B are of the same order).

Matrix operations:

Addition: Matrices are added (or subtracted) element-wise: $[a_{ij}] + [b_{ij}] = [a_{ij} + b_{ij}]$, which is of same order.

A matrix $A = [a_{ij}]$ can also be multiplied by a scalar (complex number) k as $kA = k[a_{ij}] = [ka_{ij}]$.

Matrix Multiplication

An $m \times n$ matrix A can be multiplied with an $p \times q$ matrix B if and only if $n = p$. Their product is defined as $AB = \left[\sum_{k=1}^{n} a_{ik} b_{kj}\right]$.

Example

$$\begin{bmatrix} 2 & 3 \\ 5 & 7 \\ 11 & 13 \end{bmatrix} \times \begin{bmatrix} 1 & -1 \\ 4 & 6 \end{bmatrix} = \begin{bmatrix} 2\times1+3\times4 & 2\times(-1)+3\times6 \\ 5\times1+7\times4 & 5\times(-1)+7\times6 \\ 11\times1+13\times4 & 11\times(-1)+13\times6 \end{bmatrix} = \begin{bmatrix} 14 & 16 \\ 33 & 37 \\ 63 & 67 \end{bmatrix}$$

Matrix multiplication is associative, $A(BC)=(AB)C$; and distributive, $A(B+C)=AB+AC$, but not commutative (in general), $AB \neq BA$. Indeed, the product BA may not even be defined, as in the case of the matrices in the above example.

The Identity Matrix

The identity matrix I of order n is defined as the square matrix of order n with (i, j) -entry 1 if $i=j$ and 0 if $i \neq j$ for all $i, j=1,2\ldots n$.

$$I=\begin{bmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \ldots & 1 \end{bmatrix}$$

Multiplying any matrix by an identity matrix of appropriate order results in the same matrix, i.e., if A is any $m \times n$ matrix, then $I_m A = AI_n = A$. It is easy to show that the identity matrix I the only matrix with this property. For, suppose K is another such matrix, then $IK=K$, because of the property of I , and $IK=I$, because of the property of K , and thus, $I=K$.

Transpose and Trace

The transpose of an $m \times n$ matrix $A=[a_{ij}]_{m\times n}$ is defined as $A^T=[a_{ji}]_{n\times m}$.

Example

If $A=\begin{bmatrix} 1 & 5 & -1 & 4 \\ 0 & -3 & 1 & 7 \\ 9 & -2 & 5 & -1 \end{bmatrix}$, then $A^T=\begin{bmatrix} 1 & 0 & 9 \\ 5 & -3 & -2 \\ -1 & 1 & 5 \\ 4 & 7 & -1 \end{bmatrix}$.

Hermitian transpose

The conjugate transpose of A (also called the Hermitian transpose or transjugate) is matrix A^* obtained by replacing every element of A^T by its complex conjugate. Thus, if $A=[a_{ij}]_{m\times n}$, then $A^*=\overline{A^T}=\left[\overline{a_{ji}}\right]_{n\times m}$.

Example

If $A=\begin{bmatrix} 2+i & 4 & -3+2i \\ 1-i & 1+3i & -5 \\ 6i & 4-i & 1-2i \end{bmatrix}$, then $A^{*}=\begin{bmatrix} 2-i & 1+i & -6i \\ 4 & 1-3i & 4+i \\ -3+2i & -5 & 1+2i \end{bmatrix}$.

Symmetric, Skew-symmetric, Hermitian matrices

A square matrix A is said to be

(i) Symmetric if $A=A^{T}$

(ii) Skew-symmetric if $A=-A^{T}$

(iii) Hermitian of $A=A^{*}$

Note: The principal or main diagonal of a square matrix $A=[a_{ij}]$ of order n consists of the entries a_{ii} for every $i=1,2\ldots n$.

Trace of a matrix

The trace of a square matrix is the sum of all the elements on the main diagonal, and is denoted by $\operatorname{trace}(A)$ or $\operatorname{tr}(A)=\sum_{i=1}^{n} a_{ii}$.

Example

If $A=\begin{bmatrix} 9 & -1 & 1 \\ 3 & -2 & 1 \\ -4 & 7 & 5 \end{bmatrix}$, then $\operatorname{tr}(A)=a_{11}+a_{22}+a_{33}=9+(-2)+5=12$.

Properties

1. $(A+B)^{T}=A^{T}+B^{T}$ and $(A+B)^{*}=A^{*}+B^{*}$ for all $m\times n$ matrices A and B.
2. $(cA)^{T}=cA^{T}$ and $(cA)^{*}=c^{*}A^{*}$ for all matrices A and scalars c.
3. $\left(A^{T}\right)^{T}=(A^{*})^{*}=A$ for all matrices A.
4. $(AB)^{T}=B^{T}A^{T}$ and $(AB)^{*}=B^{*}A^{*}$ (provided the product AB is defined).
5. $\operatorname{tr}(A+B)=\operatorname{tr}(A)+\operatorname{tr}(B)$ for all $n\times n$ matrices A and B.
6. $\operatorname{tr}(cA)=c\operatorname{tr}(A)$ for all square matrices A and scalars c.
7. $\operatorname{tr}\left(A^{T}\right)=\operatorname{tr}(A)$ and $\operatorname{tr}(A^{*})=\overline{\operatorname{tr}(A)}$ for all square matrices A.
8. $\operatorname{tr}(AB)=\operatorname{tr}(BA)$ (provided the product AB is defined and is a square matrix).

Example

Let $A = \begin{bmatrix} 3 & -1 \\ 2 & 5 \end{bmatrix}$ and $B = \begin{bmatrix} -2 & 7 \\ 1 & 4 \end{bmatrix}$.

Then $AB = \begin{bmatrix} -7 & 17 \\ 1 & 34 \end{bmatrix}$ and $BA = \begin{bmatrix} 8 & 37 \\ 11 & 19 \end{bmatrix}$.

We observe that $\text{tr}(AB) = -7 + 34 = 27$ and $\text{tr}(BA) = 8 + 19 = 27$. Thus, although the products AB and BA, and indeed, even the main diagonal entries of the products are different, $\text{tr}(AB) = \text{tr}(BA)$.

Determinants

For every square matrix A, we can associate a unique real (or complex) number, called the determinant of A, and denoted by $|A|$ or $\det(A)$, having certain important properties.

Definition

1. For a 1×1 matrix $A = [a_{11}]$, the determinant is defined to be $|A| = a_{11}$.
2. For an $n \times n$ matrix $A = [a_{ij}]$ (where $n > 1$), the determinant is defined as

   $$|A| = \sum_{j=1}^{n} (-1)^{j+1} a_{1j} \left|A(1\,|\,j)\right|$$

   where $A(i\,|\,j)$ denotes the $(n-1)\times(n-1)$ matrix obtained by deleting the i^{th} row and j^{th} column of A (the determinant $\left|A(i\,|\,j)\right|$ is itself called the minor and is denoted by M_{ij}.

Example

1. Let $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$.

   Then $|A| = \begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix}$

   $= 1 \times 4 - 2 \times 3$

   $= -2$.

2. Let $A = \begin{bmatrix} 1 & 3 & 1 \\ -1 & 1 & 2 \\ 2 & 1 & -2 \end{bmatrix}$.

Then $|A| = \begin{vmatrix} 1 & 3 & 1 \\ -1 & 1 & 2 \\ 2 & 1 & -2 \end{vmatrix}$

$$= 1 \times \begin{vmatrix} 1 & 2 \\ 1 & -2 \end{vmatrix} - 3 \times \begin{vmatrix} -1 & 2 \\ 2 & -2 \end{vmatrix} + 1 \times \begin{vmatrix} -1 & 1 \\ 2 & 1 \end{vmatrix}$$

$$= 1\big(1 \times (-2) - 2 \times 1\big) - 3\big((-1) \times (-2) - 2 \times 2\big) + 1\big(-1 \times 1 - 1 \times 2\big)$$

$$= -4 + 6 - 3$$

$$= -1.$$

Inverse of a Matrix

For a square matrix A of order n, another matrix B of the same order is said to be an inverse of A if $AB = BA = I$, the identity matrix of order n.

A square matrix that has no inverse is said to be singular, and one that has an inverse is said to be non-singular or invertible.

Theorem

Every invertible matrix has a unique inverse.

Proof

Let A be an invertible matrix, and suppose B_1 and B_2 are inverses of A.

Then $B_2 = B_2 I$ (where I is the identity matrix of appropriate order)

$= B_2(AB_1)$ (as B_1 is an inverse of A)

$= (B_2 A)B_1$ (as matrix multiplication is associative)

$= IB_1$ (as B_2 is an inverse of A)

$= B_1$.

$\therefore B_2 = B_1$, i.e., the inverse of A is unique.

The unique inverse of A is denoted by A^{-1}.

Example

If $A=\begin{bmatrix}1 & 3 & 1\\ -1 & 1 & 2\\ 2 & 1 & -2\end{bmatrix}$, then $A^{-1}=\begin{bmatrix}4 & -7 & -5\\ -2 & 4 & 3\\ 3 & -5 & -4\end{bmatrix}$.

For, $AA^{-1}=\begin{bmatrix}1 & 3 & 1\\ -1 & 1 & 2\\ 2 & 1 & -2\end{bmatrix}\begin{bmatrix}4 & -7 & -5\\ -2 & 4 & 3\\ 3 & -5 & -4\end{bmatrix}=\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}$.

Properties

1. $\left(A^{-1}\right)^{-1}=A$ for all invertible matrices A.
2. $\left(AB\right)^{-1}=B^{-1}A^{-1}$ for all invertible matrices A and B of the same order.
3. $\left(A^{T}\right)^{-1}=\left(A^{-1}\right)^{T}$ for all invertible matrices A.

A square matrix is non-singular if and only if its determinant is non-zero. The inverse of a non-singular matrix A can be calculated in terms of determinants of A and its submatrices.

Definition

The cofactor C_{ij} of A corresponding to the (i,j)-entry is defined as $C_{ij}=(-1)^{i+j}M_{ij}$, where M_{ij} is the minor.

The cofactor matrix C of A is the matrix with (i,j)-entry as the cofactor C_{ij}, for every i, j.

The adjoint (or adjugate) of A, denoted by $\operatorname{adj} A$ is the transpose of the cofactor matrix C.

The inverse of a non-singular matrix A is given by the formula $A^{-1}=\dfrac{1}{|A|}\left(\operatorname{adj} A\right)$

Example:

Let $A=\begin{bmatrix}1 & 0 & 2\\ 2 & 3 & 1\\ 1 & 2 & 1\end{bmatrix}$.

The cofactor matrix is $C = \begin{bmatrix} 1 & -1 & 1 \\ 4 & -1 & -2 \\ -6 & 3 & 3 \end{bmatrix}$ and the determinant is $|A| = 3$.

Thus, $\operatorname{adj} A = C^T = \begin{bmatrix} 1 & 4 & -6 \\ -1 & -1 & 3 \\ 1 & -2 & 3 \end{bmatrix}$.

$$\therefore A^{-1} = \frac{\operatorname{adj} A}{|A|} = \frac{1}{3}\begin{bmatrix} 1 & 4 & -6 \\ -1 & -1 & 3 \\ 1 & -2 & 3 \end{bmatrix}.$$

Special Matrices

A matrix $A = [a_{ij}]$ be an $n \times n$ (real or complex) matrix is said to be

1. **diagonal** if $a_{ij} = 0$ for all $i \neq j$.
2. **tridiagonal** if $a_{ij} = 0$ for all i, j with $|i - j| > 1$, i.e., all the entries of A are zero except (possibly) for those on the main diagonal ($a_{11}, a_{22} \ldots a_{nn}$), the superdiagonal ($a_{12}, a_{23} \ldots a_{n-1,n}$), and the subdiagonal ($a_{21}, a_{32} \ldots a_{n,n-1}$).
3. **upper-triangular** if $a_{ij} = 0$ for $i > j$, i.e., all entries below the main diagonal are zero.
4. **lower-triangular** if $a_{ij} = 0$ for $i < j$, i.e., all entries above the main diagonal are zero.
5. **orthogonal** if all its entries are real and $A^T = A^{-1}$, i.e., $AA^T = A^T A = I$.
6. **unitary** if it has complex entries and $A^* = A^{-1}$, i.e., $AA^* = A^* A = I$.

Exercises

1. Classify the following matrices as symmetric, skew-symmetric, Hermitian, or none of these:

   a. $\begin{bmatrix} 1 & -3 \\ -3 & 4 \end{bmatrix}$

b. $\begin{bmatrix} e^x & e^{-ix} \\ e^{ix} & e^{-x} \end{bmatrix}$

c. $\begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$

d. $\begin{bmatrix} 2 & 1+i & 3 \\ 1-i & 0 & i \\ 3 & -i & -1 \end{bmatrix}$

e. $\begin{bmatrix} 0 & 1 & 1 \\ -1 & 0 & 1 \\ -1 & -1 & 0 \end{bmatrix}$

f. $\begin{bmatrix} 1 & 1 & 1 \\ -1 & 1 & 1 \\ -1 & -1 & 1 \end{bmatrix}$

g. $\begin{bmatrix} 2 & 1+i & 3 \\ 1-i & i & i \\ 3 & -i & -1 \end{bmatrix}$

2. Find the value of x such that $A = \begin{bmatrix} x^2 - 9 & x^2 - 2x - 1 \\ 4 - 2x & 0 \end{bmatrix}$ is

   a. symmetric
   b. skew-symmetric

3. Show that every skew-symmetric matrix $A = [a_{ij}]$ of order n has all main diagonal entries zero, i.e., $a_{11} = a_{22} = \ldots = a_{nn} = 0$.

4. Show that every Hermitian matrix has only real numbers along the main diagonal.

5. Find the determinant of the following:

   a. $\begin{bmatrix} 1 & -1 \\ 3 & 3 \end{bmatrix}$

   b. $\begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$

c. $\begin{bmatrix} e^x & e^{-ix} \\ e^{ix} & e^{-x} \end{bmatrix}$

d. $\begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$

e. $\begin{bmatrix} e^x & e^{2x} \\ e^x & 2e^{2x} \end{bmatrix}$

f. $\begin{bmatrix} 1 & 2 & -2 \\ -1 & 3 & 0 \\ 0 & -2 & 1 \end{bmatrix}$

6. Determine which of the following are invertible and find the inverse:

a. $\begin{bmatrix} 1 & -1 \\ -3 & 3 \end{bmatrix}$

b. $\begin{bmatrix} 1 & -1 \\ 3 & 3 \end{bmatrix}$

c. $\begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$

d. $\begin{bmatrix} 0 & 1 & 1 \\ -1 & 0 & 1 \\ -1 & -1 & 0 \end{bmatrix}$

e. $\begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$

f. $\begin{bmatrix} 1 & x & x^2 \\ 0 & 1 & x \\ 0 & 0 & 1 \end{bmatrix}$

g. $\begin{bmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{bmatrix}$, where a, b, c are any three complex numbers.

h. $\begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix}$

7. Show that the inverse of every 3×3 upper-triangular matrices with non-zero diagonal entries is also upper triangular.
   Hint: If the i^{th} row of A is multiplied by a non-zero constant p, then the i^{th} column of A^{-1} gets divided by p. Use this property to reduce A to a matrix of the same form as that in Exercise 6g.
8. Find the inverse of a general $n\times n$ diagonal matrix with non-zero diagonal entries.
9. Classify the following matrices as orthogonal, unitary, or neither:
   a. $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$
   b. $\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$
   c. $\begin{bmatrix} 1 & -1 \\ -1 & 0 \end{bmatrix}$
   d. $\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$
   e. $\begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}$
   f. $\begin{bmatrix} 1/\sqrt{2} & i/\sqrt{2} \\ 1/\sqrt{2} & -i/\sqrt{2} \end{bmatrix}$
   g. $\begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$

Elementary Row and Column Operations

A matrix is said to be in row-reduced echelon form (or reduced row echelon form) if it satisfies the following conditions:

1. In each row, the first non-zero entry (called the leading entry) should be1.
2. The leading entry of every row is to the left of the leading entry, if any, of all rows below it.

Note that zero rows (i.e., rows with all entries 0) are allowed, but the second condition implies that all zero rows must occur below all non-zero rows (i.e., rows with at least one non-zero entry).

<u>Example</u>:
Among the matrices given below, A and C are in row-reduced echelon form, while B and D are not (why?).

$$A = \begin{bmatrix} 1 & 0 & 2 & -1 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad B = \begin{bmatrix} 1 & 0 & 2 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 4 \end{bmatrix}$$

$$C = \begin{bmatrix} 1 & 2 & 0 & 0 & -1 & -5 \\ 0 & 1 & 0 & 4 & 3 & -3 \\ 0 & 0 & 0 & 1 & 1 & 2 \end{bmatrix} \quad D = \begin{bmatrix} 1 & 2 & 0 & 0 & -1 & -5 \\ 0 & 1 & 0 & 4 & 3 & -3 \\ 0 & 1 & 0 & 5 & 4 & -1 \end{bmatrix}.$$

The following three operations performed on a matrix are called elementary row operations:

1. **Interchange**: Interchange any two rows. $R_i \leftrightarrow R_j$ denotes interchanging of the i^{th} and j^{th} rows
2. **Scaling**: Multiply every entry of a row by the same non-zero scalar. $R_i \to kR_i$ denotes multiplying the i^{th} row by the scalar k.
3. **Row Addition**: Add a multiple of one row of the matrix to another row. $R_i \to R_i + kR_j$ denotes adding k times the j^{th} row, to the i^{th} row.

Note: These definitions are motivated by analogous operations that can be performed on a given system of linear equations to obtain equivalent systems of linear equations.

We can use elementary row operations to transform a matrix into its row-reduced echelon form.

<u>Example:</u>

$$A = \begin{bmatrix} 4 & 8 & 0 & 0 & -4 & -20 \\ -3 & -5 & 0 & 4 & 6 & 12 \\ 2 & 6 & 0 & 9 & 5 & -14 \end{bmatrix}$$

$$\sim\begin{bmatrix} 1 & 2 & 0 & 0 & -1 & -5 \\ -3 & -5 & 0 & 4 & 6 & 12 \\ 2 & 6 & 0 & 9 & 5 & -14 \end{bmatrix} R_1 \to R_1/4$$

$$\sim\begin{bmatrix} 1 & 2 & 0 & 0 & -1 & -5 \\ 0 & 1 & 0 & 4 & 3 & -3 \\ 0 & 2 & 0 & 9 & 7 & -4 \end{bmatrix} \begin{matrix} R_2 \to R_2 + 3R_1 \\ R_3 \to R_3 - 2R_1 \\ \end{matrix}$$

$$\sim\begin{bmatrix} 1 & 2 & 0 & 0 & -1 & -5 \\ 0 & 1 & 0 & 4 & 3 & -3 \\ 0 & 0 & 0 & 1 & 1 & 2 \end{bmatrix} R_3 \to R_3 - 2R_2$$

Elementary column operations can also be defined, similar to elementary row operations. Applying these operations, we can find the analogous column-reduced echelon form of a matrix.

Example

$$A = \begin{bmatrix} 4 & 8 & 0 & 0 & -4 & -20 \\ -3 & -5 & 0 & 4 & 6 & 12 \\ 2 & 6 & 0 & 9 & 5 & -14 \end{bmatrix}$$

$$\sim\begin{bmatrix} 4 & 0 & 0 & 0 & 0 & 0 \\ -3 & 1 & 0 & 4 & 3 & -3 \\ 2 & 2 & 0 & 9 & 7 & -4 \end{bmatrix} \begin{matrix} C_2 \to C_2 - 2C_1 \\ C_5 \to C_5 + C_1 \\ C_6 \to C_6 + 5C_1 \end{matrix}$$

$$\sim\begin{bmatrix} 4 & 0 & 0 & 0 & 0 & 0 \\ -3 & 1 & 0 & 0 & 0 & 0 \\ 2 & 2 & 0 & 1 & 1 & 2 \end{bmatrix} \begin{matrix} C_4 \to C_4 - 4C_2 \\ C_5 \to C_5 - 3C_2 \\ C_6 \to C_6 + 3C_2 \end{matrix}$$

$$\sim\begin{bmatrix} 4 & 0 & 0 & 0 & 0 & 0 \\ -3 & 1 & 0 & 0 & 0 & 0 \\ 2 & 2 & 0 & 1 & 0 & 0 \end{bmatrix} \begin{matrix} C_5 \to C_5 - C_4 \\ C_6 \to C_6 - 2C_4 \\ \end{matrix}$$

$$\sim\begin{bmatrix} 4 & 0 & 0 & 0 & 0 & 0 \\ -3 & 1 & 0 & 0 & 0 & 0 \\ 2 & 2 & 1 & 0 & 0 & 0 \end{bmatrix} C_3 \leftrightarrow C_4$$

Elementary Matrices

All the elementary row and column operations can be defined in terms of multiplication by special matrices called elementary matrices.

Definition

An elementary matrix E is a square matrix that generates an elementary row operation on a matrix A under the multiplication EA.

Example:

Let $E = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, $\quad A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$.

$$EA = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} = \begin{bmatrix} 4 & 5 & 6 \\ 1 & 2 & 3 \\ 7 & 8 & 9 \end{bmatrix}.$$

Thus, left-multiplication by E has the effect of interchanging the first and second rows of A.

Now,
$$AE = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 3 \\ 5 & 4 & 6 \\ 8 & 7 & 9 \end{bmatrix}.$$

Right-multiplication by E, therefore, interchanging the first and second columns of A.

Construction of Elementary Matrices

An elementary operation E of order $n \times n$ that performs a specified elementary row operation is obtained by performing the same operation on the identity matrix of equal order.

Example: Let $A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{bmatrix}$. To construct an elementary matrix, say E_1, that interchanges the first and third rows of A, consider $I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$. By interchanging the first and third rows of I, we obtain $E_1 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$.

Then, $E_1 A = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{bmatrix} = \begin{bmatrix} 9 & 10 & 11 & 12 \\ 5 & 6 & 7 & 8 \\ 1 & 2 & 3 & 4 \end{bmatrix}$.

To construct an elementary matrix E_2 that subtracts 5 times the first row of A from the second row of A, perform the same operation on I:

$$E_2 = \begin{bmatrix} 1 & 0 & 0 \\ -5 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Then $E_2 A = \begin{bmatrix} 1 & 0 & 0 \\ -5 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & -4 & -8 & -12 \\ 9 & 10 & 11 & 12 \end{bmatrix}$.

However, E_1 cannot be used for interchanging the first and third columns of A, as it is not possible to right-multiply A having order 3×4 by E_1 having order 3×3. Therefore, we need to obtain another elementary matrix, say E_3, from the identity matrix of order 4 by interchanging the first and third columns (or equivalently, rows).

Thus, $E_3 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$.

Then, $AE_3 = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 2 & 1 & 4 \\ 7 & 6 & 5 & 8 \\ 11 & 10 & 9 & 12 \end{bmatrix}$

Rank and Inverse Using Row Operations

The rank of a matrix is the order of its largest non-singular square submatrix. It is also possible to define and compute the rank in a different manner.

Definition

The row rank of a matrix is the number of non-zero rows in the echelon form of that matrix.

Example: The rank of $\begin{bmatrix} 1 & 2 & -4 & 2 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ is 3, since this is in the echelon form of the matrix and it contains three rows are non-zero.

The column rank of a matrix is the number of non-zero columns in the echelon form of that matrix.

Example $\quad A = \begin{bmatrix} 2 & -1 & 3 \\ 1 & 0 & 1 \\ 0 & 2 & -1 \\ 1 & 1 & 4 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 1 \\ 2 & -1 & 3 \\ 0 & 2 & -1 \\ 1 & 1 & 4 \end{bmatrix} R_1 \leftrightarrow R_2$

$$\sim \begin{bmatrix} 1 & 0 & 1 \\ 0 & -1 & 1 \\ 0 & 2 & -1 \\ 0 & 1 & 3 \end{bmatrix} \begin{matrix} R_2 \to R_2 + 2R_1 \\ R_4 \to R_4 - R_1 \end{matrix}$$

$$\sim \begin{bmatrix} 1 & 0 & 1 \\ 0 & -1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 4 \end{bmatrix} \begin{matrix} R_3 \to R_3 + 2R_2 \\ R_4 \to R_4 + R_2 \\ \\ \\ \end{matrix}$$

$$\sim \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{matrix} R_2 \to -R_2 \\ R_4 \to R_4 - 4R_3 \end{matrix}$$

As there are three non-zero rows in the row-reduced echelon form of A, the row rank of A is 3. Now, we find the column rank.

$$A = \begin{bmatrix} 2 & -1 & 3 \\ 1 & 0 & 1 \\ 0 & 2 & -1 \\ 1 & 1 & 4 \end{bmatrix}$$

$$\sim \begin{bmatrix} -1 & 2 & 3 \\ 0 & 1 & 1 \\ 2 & 0 & -1 \\ 1 & 1 & 4 \end{bmatrix} \begin{matrix} C_1 \leftrightarrow C_2 \\ \\ \\ \\ \end{matrix}$$

$$\sim \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 1 \\ 2 & 4 & 5 \\ 1 & 3 & 7 \end{bmatrix} \begin{matrix} C_2 \to C_2 + 2C_1 \\ C_3 \to C_3 + 3C_1 \\ \\ \\ \end{matrix}$$

$$\sim \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 4 & 1 \\ 1 & 3 & 4 \end{bmatrix} \begin{matrix} C_3 \to C_3 - C_2 \\ \\ \\ \\ \end{matrix}$$

As there are three non-zero columns in the column-reduced echelon form of A, the column rank of A is 3. We see that the row rank and column rank of A are equal. Indeed, this is true in general, as stated in the following theorem.

Theorem

The row rank of a matrix is the same as the column rank, and the common value is the rank of the matrix.

Inverse using elementary operations

It is also possible to find the inverse of a non-singular matrix using elementary row (or column) operations. Consider a square matrix A of order n, and the identity matrix I of the same order. We define an **augmented matrix** $[A|I]$ by appending the columns of I to the columns of A, on the right. Then we reduce A in $[A|I]$ to the identity matrix of order n by performing a series of elementary row operations on the augmented matrix. When A has been reduced to the identity matrix, I would have been reduced to A^{-1}. If it is not possible to reduce A to the identity matrix, then its row reduced form contains at least one zero row, which means that A is singular and hence has no inverse.

Example: Find the inverse of the matrix A $= \begin{bmatrix} 1 & 0 & 2 \\ 2 & -1 & 3 \\ 4 & 1 & 8 \end{bmatrix}$

Solution: $[A \mid I] = \begin{array}{c} \left[\begin{array}{ccc:ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 2 & -1 & 3 & 0 & 1 & 0 \\ 4 & 1 & 8 & 0 & 0 & 1 \end{array}\right] \end{array} \begin{array}{l} R_1 \\ R_2 \\ R_3 \end{array}$

$$[A \mid I] = \left[\begin{array}{ccc:ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & -1 & -1 & -2 & 1 & 0 \\ 0 & 1 & 0 & -4 & 0 & 1 \end{array}\right]$$

$$\sim \left[\begin{array}{ccc:ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & -1 & -1 & -2 & 1 & 0 \\ 0 & 0 & -1 & -6 & 1 & 1 \end{array}\right]$$

$$\sim \left[\begin{array}{ccc:ccc} 1 & 0 & 0 & -11 & 2 & 2 \\ 0 & -1 & 0 & 4 & 0 & -1 \\ 0 & 0 & 1 & 6 & -1 & -1 \end{array}\right]$$

$$\sim \left[\begin{array}{ccc:ccc} 1 & 0 & 0 & -11 & 2 & 2 \\ 0 & 1 & 0 & -4 & 0 & 1 \\ 0 & 0 & 1 & 6 & -1 & -1 \end{array}\right].$$

Therefore $A^{-1} = \begin{bmatrix} -11 & 2 & 2 \\ -4 & 0 & 1 \\ 6 & -1 & -1 \end{bmatrix}$.

Example: Find the inverse of the matrix $A = \begin{bmatrix} 4 & 1 & 2 \\ 0 & 1 & 0 \\ 8 & 4 & 5 \end{bmatrix}$

Solution:

Writing the given matrix side by side with the unit matrix of order 3, we have

$$[A \mid I] = \left[\begin{array}{ccc:ccc} 4 & 1 & 2 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 8 & 4 & 5 & 0 & 0 & 1 \end{array}\right] \begin{array}{l} R_1 \\ R_2 \\ R_3 \end{array}$$

Perform the row transformations: $R_1 \to \frac{R_1}{4}$ and $R_3 \to R_3 - 2R_1$

$$[A \mid I] = \left[\begin{array}{ccc:ccc} 1 & \frac{1}{4} & \frac{1}{2} & \frac{1}{4} & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 2 & 1 & -2 & 0 & 1 \end{array}\right]$$

Perform the row transformations: $R_3 \to R_3 - 2R_2$

$$[A \mid I] = \left[\begin{array}{ccc:ccc} 1 & \frac{1}{4} & \frac{1}{2} & \frac{1}{4} & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & -2 & -2 & 1 \end{array}\right]$$

Perform the row transformations: $R_1 \to R_1 - \frac{1}{4}R_2$

$$[A \mid I] = \left[\begin{array}{ccc:ccc} 1 & 0 & \frac{1}{2} & \frac{1}{4} & -\frac{1}{4} & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & -2 & -2 & 1 \end{array}\right]$$

Perform the row transformations: $R_1 \to R_1 - \frac{1}{2} R_3$

$$[A \mid I] = \left[\begin{array}{ccc:ccc} 1 & 0 & 0 & \frac{5}{4} & \frac{3}{4} & -\frac{1}{2} \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & -2 & -2 & 1 \end{array}\right] = [\, I \mid B]$$

Hence the inverse of the given matrix is B = A^{-1} = $\begin{bmatrix} \frac{5}{4} & \frac{3}{4} & -\frac{1}{2} \\ 0 & 1 & 0 \\ -2 & -2 & 1 \end{bmatrix}$.

Determine the inverse of the following matrices by using elementary row operations:

1. $A = \begin{bmatrix} 1 & 1 & 0 \\ 2 & -2 & 1 \\ 1 & -1 & 0 \end{bmatrix}$ Ans. $A^{-1} = \begin{bmatrix} 1/2 & 0 & 1/2 \\ 1/2 & 0 & -1/2 \\ 0 & 1 & -2 \end{bmatrix}$

2. $A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & -1 & 1 \\ 1 & -1 & 0 \end{bmatrix}$ Ans. $A^{-1} = \begin{bmatrix} 1/2 & 0 & 1/2 \\ 1/2 & 0 & -1/2 \\ 0 & 1 & -1 \end{bmatrix}$

3. $A=\begin{bmatrix} 1 & 3 & 1 \\ -1 & 1 & 2 \\ 2 & 1 & -2 \end{bmatrix}$ Ans. $A^{-1}=\begin{bmatrix} 4 & -7 & -5 \\ -2 & 4 & 3 \\ 3 & -5 & -4 \end{bmatrix}$

4. $A=\begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$ Ans. $A^{-1}=\begin{bmatrix} -1 & 1 & 1 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{bmatrix}$

5. $A=\begin{bmatrix} 5 & 8 & 1 \\ 0 & 2 & 1 \\ 4 & 3 & -1 \end{bmatrix}$ $Ans\ \ A^{-1}=\begin{bmatrix} 5 & -11 & -6 \\ -4 & 9 & 5 \\ 8 & -17 & -10 \end{bmatrix}$

Rank of a matrix and solving system of equations

The rank of a matrix A is the maximum number of linearly independent row vectors (or column vectors) in A. A simple application of rank comes from the fact that a linear system of equations AX=B has a solution if and only if the rank of the coefficient matrix A is the same as the rank of the matrix [A|B] where [A|B] denotes the coefficient matrix augmented by column vector B.

The rank of the matrix $\begin{bmatrix} 1 & 2 & 3 & 2 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ is 3, since all the three rows are non-zero.

The rank of the matrix $\begin{bmatrix} 1 & 4 & 2 & 9 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ is 2, since two rows are non-zero.

Problem: Find the rank of the following matrix using elementary row transformations.

A = $\begin{pmatrix} 2 & 1 & 3 & 1 \\ 4 & 0 & 2 & 6 \\ 0 & 1 & 2 & 2 \end{pmatrix}$.

Solution: Clearly, rank (A) $\leq$ min $\{3, 4\} = 3$.

Apply row transformations,

Perform, $R_2 \to R_2 - 2R_1$,

$$A \sim \begin{pmatrix} 2 & 1 & 3 & 1 \\ 0 & -2 & -4 & 4 \\ 0 & 1 & 2 & 2 \end{pmatrix}.$$

$R_3 \to R_3 + (1/2)R_3$,

$$\Rightarrow A \sim \begin{pmatrix} 2 & 1 & 3 & 1 \\ 0 & -2 & -4 & 4 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

$\Rightarrow$ Rank(A) = 2

Example: Find the rank of $\begin{bmatrix} 1 & 2 & 3 & 2 \\ 2 & 3 & 5 & 1 \\ 1 & 3 & 4 & 5 \end{bmatrix}_{3\times 4}$ using elementary row transformation.

The rank of A $\leq$ min {3, 4} = 3.

$$A = \begin{array}{c|c} \begin{bmatrix} 1 & 2 & 3 & 2 \\ 2 & 3 & 5 & 1 \\ 1 & 3 & 4 & 5 \end{bmatrix} & \begin{matrix} R_1 \\ R_2 \\ R_3 \end{matrix} \end{array}$$

[Firstly we use the leading entry in the first row 1 to make the leading entries in second and third rows to zero].

Perform, $\begin{matrix} R_2 \to R_2 - 2R_1, \\ R_3 \to R_3 - R_1 \end{matrix}$ then $A \approx \begin{bmatrix} 1 & 2 & 3 & 2 \\ 0 & -1 & -1 & -3 \\ 0 & 1 & 1 & 3 \end{bmatrix}$

Perform, $R_2 \to (-1)R_2$ then $A \approx \begin{bmatrix} 1 & 2 & 3 & 2 \\ 0 & 1 & 1 & 3 \\ 0 & 1 & 1 & 3 \end{bmatrix}$

Perform, $R_3 \to R_3 - R_2$ then $A \approx \begin{bmatrix} 1 & 2 & 3 & 2 \\ 0 & 1 & 1 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

The above matrix is in the echelon form having two non-zero rows. Hence the rank of A is 2.

Example: Determine the rank of the matrix $\begin{pmatrix} 1 & 2 & 3 \\ 1 & 4 & 2 \\ 2 & 6 & 5 \end{pmatrix}$.

Solution: Apply $R_2 \to R_2 - R_1$, $R_3 \to R_3 - 2R_1$

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & 2 & -1 \\ 0 & 2 & -1 \end{pmatrix}$$

Apply $R_3 \to R_3 - R_1 \sim \begin{pmatrix} 1 & 2 & 3 \\ 0 & 2 & -1 \\ 0 & 0 & 0 \end{pmatrix}$.

Therefore the rank of the matrix is 2.

Exercises

Determine the rank of the following matrices by using elementary row operations:

1. $\begin{bmatrix} 1 & -2 & 0 \\ 1 & 1 & 1 \\ 2 & -1 & 1 \end{bmatrix}$ Ans. Rank = 2

2. $\begin{bmatrix} 1 & 1 & 0 \\ 2 & -2 & 1 \\ 1 & -1 & 0 \end{bmatrix}$ Ans. Rank = 3

3. $\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & 2 \\ 3 & 1 & 3 & 4 \end{bmatrix}$ Ans. Rank = 2

4. $\begin{bmatrix} 1 & 3 & 1 & 1 \\ -1 & 1 & 2 & 2 \\ 2 & 1 & -2 & 1 \\ 1 & 2 & -2 & 2 \end{bmatrix}$ Ans. Rank = 4

5. $\begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$ Ans. Rank = 3

Linear Equations-Consistency:

Definitions: A system of *m* linear equations in *n* unknowns $x_1, x_2, \ldots, x_n$ is a set of equations of the form

$$a_{11} x_1 + a_{12} x_2 + \ldots + a_{1n} x_n = b_1$$

$$a_{21} x_1 + a_{22} x_2 + \ldots + a_{2n} x_n = b_2$$

$$\ldots \qquad \ldots \qquad \ldots \qquad \text{\_\_\_\_\_\_\_\_} \quad (*) \text{ say.}$$

$$\ldots \qquad \ldots \qquad \ldots$$

$$a_{m1} x_1 + a_{m2} x_2 + \ldots + a_{mn} x_n = b_m.$$

The a_{ij} are given numbers, which are called the *coefficients* of the system. The b_i are also given numbers.

If the b_i 's are all zero, then (*) is called a *homogeneous system*. If at least one b_i is not zero, then (*) is called a *non-homogeneous system*.

A *solution* of (*) is a set of numbers $x_1, x_2, .., x_n$ which satisfy all the m equations. If the system (*) is homogeneous, it has at least one trivial solution $x_1 = 0, x_2 = 0 \ldots x_n = 0$.

From the definition of matrix multiplication, we can write the above system (*) as

$$AX = B$$

where $A = \begin{bmatrix} a_{11} & a_{12} & \ldots. & a_{1n} \\ a_{21} & a_{22} & \ldots. & a_{2n} \\ \ldots. & \ldots. & \ldots. & \ldots. \\ \ldots. & \ldots. & \ldots. & \ldots. \\ a_{m1} & a_{m2} & \ldots. & a_{mn} \end{bmatrix}_{m \times n}$ is the coefficient matrix

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}_{n \times 1} \quad \text{and} \quad B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}_{m \times 1}$$

Note that X has n elements where as B has m elements.

The matrix $\left[A \mid B\right]$ or [A:B] = $\left[\begin{array}{cccc:c} a_{11} & a_{12} & ..... & a_{1n} & b_1 \\ a_{21} & a_{22} & ..... & a_{2n} & b_2 \\ .... & .... & .... & .... & .... \\ .... & .... & .... & .... & .... \\ a_{m1} & a_{m2} & .... & a_{mn} & b_m \end{array}\right]$

is called the *augmented* matrix. The matrix $\left[A \mid B\right]$ determines the system (*) completely, since it contains all the given numbers appearing in (*).

Note: The matrix equation AX = B need not always have a solution. It may have *no solution* or a *unique solution* or an *infinite* number of solutions.

Definition: A system of equations having no solution is called an *inconsistent system*. A system of equations having one or more solution is called a *consistent system*.

Observations: Consider a system of non-homogeneous linear equations AX = B.

i) if rank A $\neq$ rank $\left[A \mid B\right]$, then system is inconsistent.

ii) if rank A = rank $\left[A \mid B\right]$ = number of unknowns, then system has a unique solution.

iii) if rank A = rank $\left[A \mid B\right]$ < number of unknowns, then system has an infinite number of solutions.

Gauss Elimination Method

This is the elementary elimination method and it reduces the system of equations to an equivalent upper triangular system which can be solved by back substitution. The method is quite general and is well-adapted for computer operations.

Here we shall explain it by considering a system of three equations for sake of simplicity.

Consider the equations

$$a_1 x + b_1 y + c_1 z = d_1$$

$$a_2 x + b_2 y + c_2 z = d_2 \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(i)}$$

$$a_3x + b_3y + c_3z = d_3$$

Step I: To eliminate x from second and third equations.

We eliminate x from the second equation by subtracting $\left(\frac{a_2}{a_1}\right)$ times the first equation from the second equation. Similarly we eliminate x from the third equation by eliminating $\left(\frac{a_3}{a_1}\right)$ times the first equation from the third equation. We thus, get the new system

$$a_1 x + b_1 y + c_1 z = d_1$$

$$b'_2 y + c'_2 z = d'_2 \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(ii)}$$

$$b'_3 y + c'_3 z = d'_3$$

Here the first equation is called the pivotal equation and a_1 is called the first pivot element.

Step II: To eliminate y from third equation in (ii).

We eliminate y from the third equation of (ii) by subtracting $\left(\frac{b'_3}{b'_2}\right)$ times the second equation from the third equation. We thus, get the new system

$$a_1 x + b_1 y + c_1 z = d_1$$

$$b'_2 y + c'_2 z = d'_2 \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots.. \text{(iii)}$$

$$c''_3 z = d''_3$$

Here the second equation is the pivotal equation and b'_2 is the new pivot element.

Step III: To evaluate the unknowns

The values of x, y, z are found from the reduced system (iii) by back substitution.

This also can be derived from the augmented matrix of the system (i),

$$[A|B] = \left[\begin{array}{ccc:c} a_1 & b_1 & c_1 & d_1 \\ a_2 & b_2 & c_2 & d_2 \\ a_3 & b_3 & c_3 & d_1 \end{array}\right] \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(iv)}$$

Using elementary row transformation, reduce the co-efficient matrix A to an upper triangular matrix.

$$\text{That is., } [A|B] \simeq \left[\begin{array}{ccc:c} a_1 & b_1 & c_1 & d_1 \\ 0 & b'_2 & c'_2 & d'_2 \\ 0 & 0 & c''_3 & d''_3 \end{array}\right] \qquad \ldots\ldots\ldots\ldots\ldots\ldots.. \text{(v)}$$

From (v) the system of linear equations is equivalent to

$$a_1 x + b_1 y + c_1 z = d_1$$

$$b'_2 y + c'_2 z = d'_2$$

$$c_3'' z = d_3''$$

By back substitution we get z, y and x constituting the exact solution of the system (i).

Note: The method will fail if one of the pivot elements a_1, b'_2 or c_3'' vanishes. In such cases the method can be modified by rearranging the rows so that the pivots are non-zero. If this is impossible, then the matrix is singular and the equations have no solution.

Example: Test for consistency and solve:

$5x + 3y + 7z = 4$, $3x + 26y + 2z = 9$, $7x + 2y + 10z = 5$.

Solution: We have $\left[\begin{array}{ccc:c} 5 & 3 & 7 & 4 \\ 3 & 26 & 2 & 9 \\ 7 & 2 & 10 & 5 \end{array}\right]$

Apply $R_1 \to 3R_1$, $R_2 \to 5R_2$

$$\left[\begin{array}{ccc|c} 15 & 9 & 21 & 12 \\ 15 & 130 & 10 & 45 \\ 7 & 2 & 10 & 5 \end{array}\right]$$

Apply $R_2 \to R_2 - R_1$

$$\left[\begin{array}{ccc|c} 15 & 9 & 21 & 12 \\ 0 & 121 & -11 & 33 \\ 7 & 2 & 10 & 5 \end{array}\right]$$

Apply $R_1 \to (7/8)R_1$, $R_3 \to 5R_3$, $R_2 \to (1/11)R_1$

$$\left[\begin{array}{ccc|c} 35 & 21 & 49 & 28 \\ 0 & 11 & -1 & 3 \\ 35 & 10 & 50 & 25 \end{array}\right]$$

Apply $R_3 \to R_3 - R_1 + R_2$, $R_1 \to (1/7)R_1$

$$\left[\begin{array}{ccc|c} 5 & 3 & 7 & 4 \\ 0 & 11 & -1 & 3 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Therefore the rank of the coefficient matrix and the augmented matrix both are 2. Hence the equations are consistent. The given system is equivalent to

$5x + 3y + 7z = 4$, $11y - z = 3$. Therefore $x = 7/11$, $y = 3/11$, and $z = 0$.

Example: Solve by Gauss elimination method $5x - y + z = 10$, $2x + 4y = 12$, $x + y + 5z = -1$.

Solution: The augmented matrix is $\begin{bmatrix} 1 & 1 & 5 & -1 \\ 2 & 4 & 0 & 12 \\ 5 & -1 & 1 & 10 \end{bmatrix}$

Apply $R_2 \to R_2 - 2R_1$, $R_3 \to R_3 - 5R_1$

$$\sim \begin{bmatrix} 1 & 1 & 5 & -1 \\ 0 & 2 & -10 & 14 \\ 0 & -6 & -24 & 15 \end{bmatrix}$$

Apply $R_2 \rightarrow (1/2)R_2$

$$\sim \begin{bmatrix} 1 & 1 & 5 & -1 \\ 0 & 1 & -5 & 7 \\ 0 & -6 & -24 & 15 \end{bmatrix}$$

Apply $R_3 \rightarrow R_3 + 6R_2$

$$\sim \begin{bmatrix} 1 & 1 & 5 & -1 \\ 0 & 1 & -5 & 7 \\ 0 & 0 & -54 & 57 \end{bmatrix}$$

Apply $R_3 \rightarrow (-1/54)R_3$

$$\begin{bmatrix} 1 & 1 & 5 & -1 \\ 0 & 1 & -5 & 7 \\ 0 & 0 & 1 & -1.0556 \end{bmatrix}$$

Therefore x = 2.556, y = 1.722, z = -1.056

Exercises and Answers

1. Solve the system

$$x - z - 2 = 0$$

$$y - z + 2 = 0$$

$$x + y - 2z = 0$$

Answer: Rank (A) = Rank (A:B) = 2 < the number of unknowns = 3. Therefore the system is consistent, and have infinite number of solutions, which involve n – r = 3 – 2 = 1 constant (s).

Gauss Jordon Method:

Example: Balance the following chemical equation with the help of Gauss Jordan Method

$C_2H_6 + O_2 \rightarrow CO_2 + H_2O$

Solution: We have to find positive integers x, y, z, w such that

(x) C_2H_6 +(y) $O_2 \rightarrow$ (z) CO_2 + (w) H_2O.

Because the number of atoms of each element must be same on each side of the above equation, we get

Carbon (C): 2x = z,

Hydrogen (H): 6x = 2w,

Oxygen (O): 2y = 2z + w.

This gives us the system

$$\begin{aligned} 2x + 0y - z + 0w &= 0 \\ 6x + 0y + 0z - 2w &= 0 \\ 0x + 2y - 2z - w &= 0 \end{aligned}$$

Gauss Jordan elimination gives

$$\left[\begin{array}{cccc:c} 1 & 0 & 0 & -\frac{1}{3} & 0 \\ 0 & 1 & 0 & -\frac{7}{6} & 0 \\ 0 & 0 & 1 & -\frac{2}{3} & 0 \end{array}\right]$$

We get x = (1/3)k; y = (7/6)k; z = (2/3)k; w = k. Take k = 6. Thus we get the balanced chemical equation as follows:

$2C_2H_6 + 7O_2 \rightarrow 4CO_2 + 6H_2O$.

Example: Using Gauss–Jordan method, find the inverse of the matrix

$$A = \begin{bmatrix} 4 & 1 & 2 \\ 0 & 1 & 0 \\ 8 & 4 & 5 \end{bmatrix}$$

Solution:

Writing the given matrix side by side with the unit matrix of order 3, we have

$$[A \mid I\,] = \left[\begin{array}{ccc:ccc} 4 & 1 & 2 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 8 & 4 & 5 & 0 & 0 & 1 \end{array}\right] \begin{array}{l} R_1 \\ R_2 \\ R_3 \end{array}$$

Perform the row transformations

$$\simeq \left[\begin{array}{ccc:ccc} 1 & \frac{1}{4} & \frac{1}{2} & \frac{1}{4} & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 2 & 1 & -2 & 0 & 1 \end{array}\right] \begin{array}{l} R_1 \to \frac{R_1}{4} \\ R_2 \\ R_3 \to R_3 - 2\,R_1 \end{array}$$

$$\simeq \left[\begin{array}{ccc:ccc} 1 & \frac{1}{4} & \frac{1}{2} & \frac{1}{4} & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & -2 & -2 & 1 \end{array}\right] \begin{array}{l} R_1 \\ R_2 \\ R_3 \to R_3 - 2\,R_2 \end{array}$$

$$\simeq \left[\begin{array}{ccc:ccc} 1 & 0 & \frac{1}{2} & \frac{1}{4} & -\frac{1}{4} & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & -2 & -2 & 1 \end{array}\right] \begin{array}{l} R_1 \to R_1 - \frac{1}{4}\,R_2 \\ R_2 \\ R_3 \end{array}$$

$$\simeq \left[\begin{array}{ccc:ccc} 1 & 0 & 0 & \frac{5}{4} & \frac{3}{4} & -\frac{1}{2} \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & -2 & -2 & 1 \end{array}\right] \begin{array}{l} R_1 \to R_1 - \frac{1}{2}\,R_3 \\ R_2 \\ R_3 \end{array}$$

$$= [\,I \mid B]$$

Hence the inverse of the given matrix is

$$B = A^{-1} = \begin{bmatrix} \frac{5}{4} & \frac{3}{4} & -\frac{1}{2} \\ 0 & 1 & 0 \\ -2 & -2 & 1 \end{bmatrix}$$

Example: Solve the following system of equations by Gauss Jordon method.

$$\begin{aligned} 3x + y + 2z &= 3 \\ 2x - 3y - z &= -3 \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(i)} \\ x + 2y + z &= 4 \end{aligned}$$

The system (i) can be written as AX = B

where $A = \begin{bmatrix} 3 & 1 & 2 \\ 2 & -3 & -1 \\ 1 & 2 & 1 \end{bmatrix}$, $X = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ and $B = \begin{bmatrix} 3 \\ -3 \\ 4 \end{bmatrix}$

Now $AX = B \Rightarrow X = A^{-1} B$ $\ldots\ldots\ldots\ldots\ldots\ldots$ (ii)

Writing the given matrix side by side with the unit matrix of order 3, we have

$$[A \mid I] = \begin{array}{c} \left[\begin{array}{ccc:ccc} 3 & 1 & 2 & 1 & 0 & 0 \\ 2 & -3 & -1 & 0 & 1 & 0 \\ 1 & 2 & 1 & 0 & 0 & 1 \end{array}\right] \end{array} \begin{array}{l} R_1 \\ R_2 \\ R_3 \end{array}$$

Continue row transformations (please refer the earlier lesson for inverse computation), we obtain

$$A^{-1} = \begin{bmatrix} -\frac{1}{8} & \frac{3}{8} & \frac{5}{8} \\ -\frac{3}{8} & \frac{1}{8} & \frac{7}{8} \\ \frac{7}{8} & -\frac{5}{8} & -\frac{11}{8} \end{bmatrix}$$

Equation (ii) becomes

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -\frac{1}{8} & \frac{3}{8} & \frac{5}{8} \\ -\frac{3}{8} & \frac{1}{8} & \frac{7}{8} \\ \frac{7}{8} & -\frac{5}{8} & -\frac{11}{8} \end{bmatrix} \begin{bmatrix} 3 \\ -3 \\ 4 \end{bmatrix}$$

$$= \begin{bmatrix} \frac{-3}{8} - \frac{9}{8} + \frac{20}{8} \\ \frac{-9}{8} - \frac{3}{8} + \frac{28}{8} \\ \frac{21}{8} + \frac{15}{8} - \frac{44}{8} \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$$

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix}_{3\times 1} = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}_{3\times 1}$$

This implies that x = 1, y = 2, z = –1.

ITERATIVE METHODS:

We shall now describe the iterative or indirect methods, which start from an approximation to the true solution and if convergent, derive a sequence of closer approximations – the cycle of computations being repeated till the required accuracy is obtained. This means that in a direct method the amount of computation is fixed, while in an iterative method the amount of computation depends on the accuracy required. In the case of matrices with a large number of zero elements, it will be advantageous to use iterative methods which preserve these elements.

Jacobi's Method:

Let the system be given by

$$\left.\begin{array}{l} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \ldots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \ldots + a_{2n}x_n = b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + \ldots + a_{3n}x_n = b_1 \\ \qquad\qquad\qquad\qquad . \\ \qquad\qquad\qquad\qquad . \\ a_{n1}x_2 + a_{n2}x_2 + a_{n3}x_3 + \ldots + a_{nn}x_n = b_n \end{array}\right\}$$

in which the diagonal elements a_{ii} do not vanish. If this is not the case, then the equations should be rearranged so that this condition is satisfied. Now, we rewrite the system above as

$$x_1 = \frac{b_1}{a_{11}} - \frac{a_{12}}{a_{11}}x_2 - \frac{a_{13}}{a_{11}}x_3 - \ldots - \frac{a_{1n}}{a_{11}}x_n$$

$$x_2 = \frac{b_2}{a_{22}} - \frac{a_{21}}{a_{22}}x_1 - \frac{a_{23}}{a_{22}}x_2 - \ldots - \frac{a_{2n}}{a_{22}}x_n$$

.

.

$$x_n = \frac{b_n}{a_{nn}} - \frac{a_{n1}}{a_{nn}}x_1 - \frac{a_{n2}}{a_{nn}}x_2 - \ldots - \frac{a_{n,n-1}}{a_{nn}}x_{n-1}$$

Let $x_1^{(0)}, x_2^{(0)}, \ldots\ldots, x_n^{(0)}$ be the initial approximation to the solution.

If $x_1^{(k)}, x_2^{(k)}, \ldots\ldots, x_n^{(k)}$ denotes the approximation to the solution at k^{th} iteration, then the solution at $(k+1)^{th}$ iteration is given by

$$x_1^{(k+1)} = \frac{b_1}{a_{11}} - \frac{a_{12}}{a_{11}}x_2^{(k)} - \frac{a_{13}}{a_{11}}x_3^{(k)} - \ldots - \frac{a_{1n}}{a_{11}}x_n^{(k)}$$

$$x_2^{(k+1)} = \frac{b_2}{a_{22}} - \frac{a_{21}}{a_{22}}x_1^{(k)} - \frac{a_{23}}{a_{22}}x_2^{(k)} - \ldots - \frac{a_{2n}}{a_{22}}x_n^{(k)}$$

.

.

$$x_n^{(k+1)} = \frac{b_n}{a_{nn}} - \frac{a_{n1}}{a_{nn}}x_1^{(k)} - \frac{a_{n2}}{a_{nn}}x_2^{(k)} - \ldots - \frac{a_{n,n-1}}{a_{nn}}x_{n-1}^{(k)}$$

This method is due to Jacobi and is called the method of simultaneous displacements.

Gauss Seidel Method:

In Jacobi's method, to obtain the values at the $(k+1)^{th}$ iteration, only k^{th} iteration values are used. A simple modification to the above method sometimes yields faster convergence and is described below. Here, along with the k^{th} iteration values, all available $(k + 1)^{th}$ iteration values are used. It is clear, therefore, that this method uses an improved component as soon as it is available and it is called the method of successive displacements, or the Gauss – Seidel method.

The Jacobi and Gauss – Seidel methods converge, for any choice of the first approximation $x_j^{(0)}$ (j= 1, 2, …, n), if every equation of the system satisfies the condition that sum of the absolute values of the coefficients a_{ij}/a_{ii} is almost equal to, or in at least one equation less than unity, i.e. provided that

$$\sum_{j=1,\, j\neq 1}^{n} \left|a_{ij} / a_{ii}\right| \leq 1, \ (i = 1, 2, \ldots, n)$$

where the "<" sign should be valid in the case of 'at least' one equation. A matrix A satisfying the above condition is known as a diagonally dominant matrix. It can be shown that the Gauss – Seidel method converges twice as fast as the Jacobi method. The working of the methods is illustrated in the following examples.

Example: We consider the equations.

$$3x + 20y - z = -18$$

$$2x - 3y + 20z = 25$$

$$20x + y - 2z = 17$$

After partial pivoting we can rewrite the system as follows:

$$\begin{aligned} 20x + y - 2z &= 17 \\ 3x + 20y - z &= -18 \qquad\qquad -(1) \\ 2x - 3y + 20z &= 25 \end{aligned}$$

Solve for x, y, z respectively

$$x = \frac{1}{20}[17 - y + 2z]$$

$$y = \frac{1}{20}[-18 - 3x + z] \qquad \text{— (2)}$$

$$z = \frac{1}{20}[25 - 2x + 3y]$$

we start from an initial approximation

$x = x^{(0)} = 0, \quad y = y^{(0)} = 0, \quad z = z^{(0)} = 0.$

Substituting these on the right hand sides of the equations (2), we get

$$x = x^{(1)} = \frac{17}{20} = 0.85$$

$$y = y^{(1)} = \frac{-18}{20} = -0.9$$

$$z = z^{(1)} = \frac{25}{20} = 1.25,$$

where $(x^{(1)}, y^{(1)}, z^{(1)}) = (0.85, -0.9, 1.25)$ are the first approximated values.

Putting these values on the right hand sides of the equations (2) we obtain

$$x = x^{(2)} = \frac{1}{20}[17 - (-0.9) + 2 \times 1.25] = 1.0200$$

$$y = y^{(2)} = \frac{1}{20}[-18 - 3 \times 0.85 + 1.25] = -0.965$$

$$z = z^{(2)} = \frac{1}{20}[25 - 2 \times 0.85 + 3(-0.9)] = 1.1515$$

The second approximated values of x, y, z are

$(x^{(2)}, y^{(2)}, z^{(2)}) = (1.0200, -0.965, 1.515).$

Substituting these values on the right hand sides of the equations (2), we have

$$x = x^{(3)} = 1.0134$$

$$y = y^{(3)} = -0.9954$$

$$z = z^{(3)} = 1.0032$$

Continuing this process, we get

The values in the 5th and 6th iterations being practically the same, we can stop the iterations. Hence the solution is

Iteration number (n)	Variables		
	x	y	z
3	1.0134	-0.9954	1.0032
4	1.0009	-1.0018	0.9993
5	1.000	-1.0002	0.9996
6	1.0000	-1.0000	1.0000

Therefore $x = 1, y = -1, z = 1.$

Example: Find the solution to the following system of equations

$$83x + 11y - 4z = 95$$

$$7x + 52y + 13z = 104$$

$$3x + 8y + 29z = 71$$

using Jacobi's iterative method for the first five iterations.

Solution: The given system is diagonally dominant.

Rewrite the given system as

$$\left.\begin{aligned} x &= \frac{95}{83} - \frac{11}{83}y + \frac{4}{83}z \\ y &= \frac{104}{52} - \frac{7}{52}x - \frac{13}{52}z \\ z &= \frac{71}{29} - \frac{3}{29}x - \frac{8}{29}y \end{aligned}\right\rangle \qquad \text{..........(1)}$$

Take the initial approximation vector as $x = x^{(0)} = 0, \; y = y^{(0)} = 0, \; z = z^{(0)} = 0.$

Then the first approximation is

$$\begin{pmatrix} x^{(1)} \\ y^{(1)} \\ z^{(1)} \end{pmatrix} = \begin{pmatrix} 1.1446 \\ 2.0000 \\ 2.4483 \end{pmatrix} \qquad \text{.........(2)}$$

Now using equation (1), the second approximation is computed from the equations as

$$\left.\begin{aligned} x^{(2)} &= 1.1446 - 0.1325y^{(1)} + 0.0482z^{(1)} \\ y^{(2)} &= 2.0 - 0.1346x^{(1)} + 0.25z^{(1)} \\ z^{(2)} &= 2.4483 - 0.1035x^{(1)} - 0.2759z^{(1)} \end{aligned}\right\rangle \qquad \text{...............(3)}$$

Substituting equation (2) into equation (3), we get the second approximation as

$$\begin{pmatrix} x^{(2)} \\ y^{(2)} \\ z^{(2)} \end{pmatrix} = \begin{pmatrix} 0.9976 \\ 1.2339 \\ 1.7424 \end{pmatrix} \qquad \text{.........(4)}$$

Proceed in a similar way, we get the third, fourth and fifth approximations to the required solution and they are tabulated as below.

Iteration number (n)	Variables		
	x	y	z
1	1.1446	2.0000	2.4483
2	0.9976	1.2339	1.7424
3	1.0651	1.4301	2.0046
4	1.0517	1.3555	1.9435
5	1.0587	1.3726	1.9665

Example: Apply Gauss – Seidal iteration method to solve the equations

$$3x_1 + 20x_2 - x_3 = -18$$

$$2x_1 - 3x_2 + 20x_3 = 25$$

$$20x_1 + x_2 - 2x_3 = 17$$

We write the equations in the form (after partial pivoting),

$$20x_1 + x_2 - 2x_3 = 17$$

$3x_1 + 20x_2 - x_3 = -18$

$20x_1 - x_2 + 2x_3 = 17$

and $x_1 = \frac{1}{20}[17 - x_2 + 2x_3]$ – (i)

$x_2 = \frac{1}{20}[-18 - 3x_1 + x_3]$ – (ii)

$x_3 = \frac{1}{20}[25 - 2x_1 + 3x_2]$ – (iii)

Initial approximation: $x_1 = x_1^{(0)} = 0$, $x_2 = x_2^{(0)} = 0$, $x_3 = x_3^{(0)} = 0$.

First iteration:

Taking $x_2 = x_2^{(0)} = 0$, $x_3 = x_3^{(0)} = 0$ in (i), we get $x_1^{(1)} = 0.85$

Taking $x_1 = x_1^{(1)} = 0.85$, $x_3 = x_3^{(0)} = 0$ in (ii), we have

$$x_2^{(1)} = \frac{1}{20}[-18 - 3 \times 0.85 + 0] = -1.0275$$

Taking $x_1 = x_1^{(1)} = 0.85$, $x_2 = x_2^{(1)} = -1.0275$ in (iii), we obtain

$$x_3^{(1)} = \frac{1}{20}[25 - 2 \times 0.85 + 3 \times (-1.0275)] = 1.0109$$

Therefore $x_1^{(1)} = 0.85$, $x_2^{(1)} = -1.0275$, $x_3^{(1)} = 1.0109$.

Second iteration:

$$x_1^{(2)} = \frac{1}{20}[17 - 3(-1.0275) + 2 \times 1.0109] = 1.0025$$

$$x_2^{(2)} = \frac{1}{20}[-18 - 3 \times (-1.0025) + 1.0109] = -0.9998$$

$$x_3^{(2)} = \frac{1}{20}[25 - 2 \times 1.0025 + 3 \times (-0.9998)] = 0.9998$$

Similarly in third iteration, we get $x_1^{(3)} = 1.0000$, $x_2^{(3)} = -1.0000$, $x_3^{(3)} = 1.0000$

The values in the 2nd and 3rd iterations being practically the same, we can stop the iterations. Hence the solution is $x_1 = 1$, $x_2 = -1$, $x_3 = 1$.

Example: Find the solution of the following system of equations

$$x - \frac{1}{4}y - \frac{1}{4}z = \frac{1}{2}$$
$$-\frac{1}{4}x + y - \frac{1}{4}w = \frac{1}{2}$$
$$-\frac{1}{4}x + z - \frac{1}{4}w = \frac{1}{2}$$
$$-\frac{1}{4}y - \frac{1}{4}z + w = \frac{1}{2}$$

using Gauss-Seidel method and perform the first five iterations.

Solution: The given system of equations can be rewritten as

$$\left.\begin{aligned} x &= 0.5 + 0.25y + 0.25z \\ y &= 0.5 + 0.25x + 0.25w \\ z &= 0.25 + 0.25x + 0.25w \\ w &= 0.25 + 0.25y + 0.25z \end{aligned}\right\} \quad \ldots\ldots \quad (1)$$

Taking the initial approximation as y = z = 0 on the right side of the equation (1) we get $x^{(1)} = 0.5$.

Taking z = 0 and w = 0 and the current value of x, we get

$y^{(1)} = 0.5 + (0.25)(0.5) + 0 = 0.625$ from the second equation of (1).

Now taking w = 0 and the current value of x, we get

$z^{(3)} = 0.25 + (0.25)(0.5) + 0 = 0.375$ from the third equation of (1).

Lastly, using the current values of y and z, the fourth equation of (1) gives

$w^{(1)} = 0.25 + (0.25)(0.625) + (0.25)(0.375) = 0.5$.

The Gauss-Seidal iterations for the given set of equations can be written as

$$\left.\begin{aligned} x^{(r+1)} &= 0.5 + 0.25y^{(r)} + 0.25z^{(r)} \\ y^{(r+1)} &= 0.5 + 0.25x^{(r+1)} + 0.25w^{(r)} \\ z^{(r+1)} &= 0.25 + 0.25x^{(r+1)} + 0.25w^{(r)} \\ w^{(r+1)} &= 0.25 + 0.25y^{(r+1)} + 0.25z^{(r+1)} \end{aligned}\right\} \quad \ldots\ldots \quad (1)$$

Now, by Gauss-Seidel procedure, the second and the subsequent approximations can be obtained and the sequence of the first five approximations is tabulated below.

Iteration number (n)	Variables			
	x	y	z	w
1	0.5	0.625	0.375	0.5
2	0.75	0.8125	0.5625	0.59375
3	0.84375	0.85938	0.60938	0.61719
4	0.86719	0.87110	0.62110	0.62305
5	0.87305	0.87402	0.62402	0.62451

Eigenvalues and Eigenvectors

Introduction

A problem which arises frequently in applications of linear algebra and matrix theory is that of finding the values of scalar parameter λ for which there exists vectors $x \neq 0$ satisfying

$$Ax = \lambda x \qquad (1)$$

where A is any square matrix of order n. Such a problem is called as an Eigenvalue problem or Characteristic value problem. If $x \neq 0$ satisfies (1) for a given λ, then A operating on x yields a scalar multiple of x.

Clearly, $x = 0$ is one solution of (1) for any λ. But we are looking for vectors $x \neq 0$ which satisfy (1).

$\therefore Ax = \lambda x \Leftrightarrow (A - \lambda I)x = 0$

Where I is the n^{th} order identity matrix. The determinant of this matrix $(A - \lambda I)$ equated to zero,

i.e,
$$|A - \lambda I| = \begin{vmatrix} a_{11} - \lambda & a_{12} & .... & a_{1n} \\ a_{21} & a_{22} - \lambda & .... & a_{2n} \\ .... & .... & .... & .... \\ a_{n1} & a_{n2} & .... & a_{nn} - \lambda \end{vmatrix} = 0 \qquad (2)$$

is called as the characteristic equation of the matrix A. On expanding the determinant, the characteristic equation takes the form

$$(-1)^n \lambda^n + k_1\lambda^{n-1} + k_2\lambda^{n-2} + ... + k_n = 0 \qquad (3)$$

Where the k_i's are expressible in terms of the elements a_{ij} of the matrix A.

The roots of the characteristic equation of A are called as the Eigenvalues or latent roots or the characteristic roots of the matrix A.

The vectors $x \neq 0$ which satisfy (1) are called the Eigenvectors or latent vectors of A. The determinant $|A - \lambda I|$, which is a polynomial in λ, is called as the characteristic polynomial of A.

The matrix equation (2) represents n homogenous linear equations

$$\begin{aligned}
&(a_{11}-\lambda)x_1 + a_{12}x_2 + ... + a_{1n}x_n = 0 \\
&a_{21}x_1 + (a_{22}-\lambda)x_2 + ... + a_{2n}x_n = 0 \\
&.... \qquad .... \qquad .... \qquad .... \\
&a_{n1}x_1 + a_{n2}x_2 + ... + (a_{nn}-\lambda)x_n = 0
\end{aligned} \qquad (4)$$

which will have a non-trivial solution if and only if the coefficient matrix is singular, i.e, if $|A - \lambda I| = 0$.

The characteristic equation of A has n roots (need not be distinct) and corresponding to each root, the equation (4) will have a non-zero solution, which is the eigenvector or latent vector.

Example: Find the eigenvalues and eigenvectors of A = $\begin{bmatrix} 5 & 4 \\ 1 & 2 \end{bmatrix}$.

Solution: The characteristic equation is $|A - \lambda I| = 0$

i.e, $\begin{vmatrix} 5-\lambda & 4 \\ 1 & 2-\lambda \end{vmatrix} = 0$ or $\lambda^2 - 7\lambda + 6 = 0$

or $(\lambda - 6)(\lambda - 1) = 0$

$\therefore \lambda = 6,1$

If $x = [x_1 \;\; x_2]$ is an eigenvector corresponding to the eigenvalue λ, then

$$(A - \lambda I)x = \begin{bmatrix} 5-\lambda & 4 \\ 1 & 2-\lambda \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0.$$

Corresponding to $\lambda = 6$, we have $\begin{bmatrix} -1 & 4 \\ 1 & -4 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0$ which gives only one dependent equation $-x_1 + 4x_2 = 0$.

$\therefore \dfrac{x_1}{4} = \dfrac{x_2}{1}$ giving the eigenvector $(4,1)$.

Corresponding to $\lambda = 1$, we have $\begin{bmatrix} 4 & 4 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0$ which gives only one dependent equation $x_1 + x_2 = 0$.

$\therefore \dfrac{x_1}{1} = \dfrac{x_2}{-1}$ giving the eigenvector $(1,-1)$.

Example 3: Find all eigen values and the corresponding eigen vectors of the matrix

$$A = \begin{bmatrix} 8 & -6 & 2 \\ -6 & 7 & -4 \\ 2 & -4 & 3 \end{bmatrix}$$

The characteristic equation of A is $|A - \lambda I| = 0$ where $\lambda I = \begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{bmatrix}$.

That is, $\begin{vmatrix} 8-\lambda & -6 & 2 \\ -6 & 7-\lambda & -4 \\ 2 & -4 & 3-\lambda \end{vmatrix} = 0.$

Expanding we get

$$(8 - \lambda)\,[(7-\lambda)(3-\lambda) - 16] + 6\,[-6(3-\lambda) + 8] + 2\,[24 - 2(7-\lambda)] = 0$$

$$\Rightarrow \; -\lambda^3 + 18\lambda^2 - 45\lambda = 0$$

That is, $\lambda^3 - 18\lambda^2 + 45\lambda = 0$, on simplification.

$$\Rightarrow \lambda(\lambda - 3)(\lambda - 15) = 0$$

Therefore $\lambda = 0, 3, 15$ are the eigen values of A.

If x, y, z be the components of an eigen vector corresponding to the eigen value λ, then we have

$$(A - \lambda I)\, X = \begin{bmatrix} 8-\lambda & -6 & 2 \\ -6 & 7-\lambda & -4 \\ 2 & -4 & 3-\lambda \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

That is,

$$\begin{aligned} (8-\lambda)x - 6y + 2z &= 0 \\ -6x + (7-\lambda)y - 4z &= 0 \\ 2x - 4y + (3-\lambda)z &= 0 \end{aligned}$$

<u>Case I:</u> Let $\lambda = 0$ we have

$$8x - 6y + 2z = 0 \qquad \text{(i)}$$

$$-6x + 7y - 4z = 0 \qquad \text{(ii)}$$

$$2x - 4y + 3z = 0 \qquad \text{(iii)}$$

Applying the rule of cross multiplication for (i) and (ii)

$$\frac{x}{\begin{vmatrix} -6 & 2 \\ 7 & -4 \end{vmatrix}} = \frac{-y}{\begin{vmatrix} 8 & 2 \\ -6 & -4 \end{vmatrix}} = \frac{z}{\begin{vmatrix} 8 & -6 \\ -6 & 7 \end{vmatrix}}$$

$$\frac{x}{24-14} = \frac{-y}{-32+12} = \frac{z}{56-36}$$

$$\frac{x}{10} = \frac{-y}{-20} = \frac{z}{20} \quad \text{or} \quad \frac{x}{1} = \frac{y}{2} = \frac{z}{2}$$

Therefore (x, y, z) are proportional to (1, 2, 2) and we can write $x = 1k$, $y = 2k$, $z = 2k$ where $k\ (\neq 0)$ is an arbitrary constant. However it is enough to keep the values of (x, y, z) in the simplest form $x = 1$, $y = 2$, $z = 2$ (putting $k = 1$). These values satisfy all the equations simultaneously.

Thus the eigen vector X_1 corresponding to the eigen value $\lambda = 0$ is $\begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}$.

<u>Case II:</u> Let $\lambda = 3$ and the corresponding equations are

$$5x - 6y + 2z = 0 \qquad -(iv)$$

$$-6x + 4y - 4z = 0 \qquad -(v)$$

$$2x - 4y - 1z = 0 \qquad -(vi)$$

From (iv) and (v) we have

$$\frac{x}{\begin{vmatrix} -6 & 2 \\ 4 & -4 \end{vmatrix}} = \frac{-y}{\begin{vmatrix} 5 & 2 \\ -6 & -4 \end{vmatrix}} = \frac{z}{\begin{vmatrix} 5 & -6 \\ -6 & 4 \end{vmatrix}}$$

That is, $\frac{x}{16} = \frac{-y}{-8} = \frac{z}{-16}$

That is, $\frac{x}{2} = \frac{-y}{1} = \frac{z}{-2}$

Thus $X_2 = \begin{pmatrix} 2 \\ 1 \\ -2 \end{pmatrix}$ is the eigen vector corresponding to $\lambda = 3$.

<u>Case III:</u> Let $\lambda = 15$ and the associated equations are

$$-7x - 6y + 2z = 0 \qquad (vii)$$

$$-6x - 8y - 4z = 0 \qquad (viii)$$

$$2x - 4y - 12z = 0 \qquad (ix)$$

From (vii) and (viii) we have

$$\frac{x}{40} = \frac{-y}{40} = \frac{z}{20}$$

That is, $\frac{x}{2} = \frac{y}{-2} = \frac{z}{1}$

Thus $X_3 = \begin{pmatrix} 2 \\ -2 \\ 1 \end{pmatrix}$ is the eigen vector corresponding to $\lambda = 15$.

Therefore $\lambda = 0, 3, 15$ are the eigen values of A and $\begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}$, $\begin{pmatrix} 2 \\ 1 \\ -2 \end{pmatrix}$ and $\begin{pmatrix} 2 \\ -2 \\ 1 \end{pmatrix}$ are the corresponding eigen vectors.

Example: Find the eigen values and eigen vectors of the matrix A = $\begin{bmatrix} 3 & 1 & -1 \\ 2 & 2 & -1 \\ 2 & 2 & 0 \end{bmatrix}$

Solution: The characteristic equation is $|A - \lambda I| = 0$

$$\Rightarrow \left\| \begin{bmatrix} 3 & 1 & -1 \\ 2 & 2 & -1 \\ 2 & 2 & 0 \end{bmatrix} - \begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{bmatrix} \right\| = 0 = 0$$

$$\Rightarrow \begin{vmatrix} \lambda - 3 & -1 & 1 \\ -2 & \lambda - 2 & 1 \\ -2 & -2 & \lambda \end{vmatrix} = 0$$

$\Rightarrow \lambda^3 - 5\lambda^2 + 8\lambda - 4 = 0.$

$\Rightarrow \lambda = 1, 2, 2$

Thus the characteristic values are 1 and 2.

Part (ii): Suppose the characteristic vector corresponding to the characteristic value 1 is α. Then $(A - I)\alpha = 0$

$$\Rightarrow \left\{ \begin{bmatrix} 3 & 1 & -1 \\ 2 & 2 & -1 \\ 2 & 2 & 0 \end{bmatrix} - \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right\} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} = 0$$

$$\Rightarrow \begin{bmatrix} 2 & 1 & -1 \\ 2 & 1 & -1 \\ 2 & 2 & -1 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

The set of equations is given below

$2\alpha_1 + \alpha_2 - \alpha_3 = 0$ ……………..(i)

$2\alpha_1 + \alpha_2 - \alpha_3 = 0$ ……………...(ii)

$2\alpha_1 + 2\alpha_2 - \alpha_3 = 0$ ……………..(iii)

Use (i) and (iii).

$2\alpha_1 + 2\alpha_2 - \alpha_3 = 0$ ………….(iii)

$2\alpha_1 + \alpha_2 - \alpha_3 = 0$ …………..(i)

$\alpha_2 = 0$

Thus $\alpha_2 = 0$. Substituting this in (i), we get $2\alpha_1 + \alpha_2 - \alpha_3 = 0$ ………..(i))

$2\alpha_1 - \alpha_3 = 0$

$$\Rightarrow \frac{\alpha_1}{1} = \frac{\alpha_3}{2}.$$

If $\alpha_1 = 1$, then we get

$\alpha_1 = 1, \alpha_2 = 0, \alpha_3 = 2.$

$\alpha = (\alpha_1, \alpha_2, \alpha_3) = (1, 0, 2)$

$\alpha = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}$ is the characteristic vector corresponding to the characteristic value1.

Part (iii): Suppose a characteristic vector corresponding to the characteristic value 2 is α. Then $(A - 2I)\alpha = 0$

$$\Rightarrow \left\{\begin{bmatrix} 3 & 1 & -1 \\ 2 & 2 & -1 \\ 2 & 2 & 0 \end{bmatrix} - \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}\right\}\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} = 0$$

$$\Rightarrow \begin{bmatrix} 1 & 1 & -1 \\ 2 & 0 & -1 \\ 2 & 2 & -2 \end{bmatrix}\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

The related system of equations is

$\alpha_1 + \alpha_2 - \alpha_3 = 0$ …………..…(i)

$2\alpha_1 - \alpha_3 = 0$ ……………….(ii)

$\Rightarrow \alpha_3 = 2\alpha_1$

$2\alpha_1 + 2\alpha_2 - 2\alpha_3 = 0$ ………(iii)

(or $\alpha_1 + \alpha_2 - \alpha_3 = 0$) $\Rightarrow \alpha_3 = \alpha_1 + \alpha_2$

Now we use (i) and (ii).

Now $2\alpha_1 = \alpha_1 + \alpha_2 \Rightarrow \alpha_1 = \alpha_2$.

So $\alpha_1 = \alpha_2$ and $\alpha_3 = 2\alpha_1$

If $\alpha_1 = 1$, then $\alpha_2 = 1$ and $\alpha_3 = 2$.

Thus $\alpha = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$ is a characteristic vector corresponding to the characteristic value 2.

Example: Find all the eigen roots and the eigen vector corresponding to the least eigen root of the matrix A = $\begin{bmatrix} 6 & -2 & 2 \\ -2 & 3 & -1 \\ 2 & -1 & 3 \end{bmatrix}$.

Solution: The characteristic equation of A is $|A - \lambda I| = 0$.

This implies $\begin{vmatrix} 6-\lambda & -2 & 2 \\ -2 & 3-\lambda & -1 \\ 2 & -1 & 3-\lambda \end{vmatrix} = 0.$

Expanding we obtain, λ^3-12 λ^2 + 36 λ -32 = 0. …………………(1)

By inspection, λ = 2 is a root. Then (λ - 2) is a factor of the (1).

Therefore dividing equation (1) with (λ - 2). We get that λ^2 -10 λ + 16 = 0.

This means (λ -2)(λ - 8) = 0. Therefore the eigen values are λ = 2, 2, 8.

We find the eigen vector corresponding the eigen value λ = 2.

When λ = 2, the set of equations are

4x -2y + 2z = 0

-2x + y - z = 0

2x- y + z = 0

The above set of equations represents a single independent equation, 2x - y + z = 0 and hence we can choose two variables z = k_1 and y = k_2.

Therefore X_1 = $\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \frac{k_2 - k_1}{2} \\ k_2 \\ k_1 \end{pmatrix}$ is the eigen vector corresponding to λ = 2 where

k_1, k_2 are not simultaneously equal to zero.

Problems

Find the eigenvalues and eigenvectors of the matrices:

a) $\begin{bmatrix} 2 & 1 & -1 \\ 1 & 1 & -2 \\ -1 & -2 & 1 \end{bmatrix}$ b) $\begin{bmatrix} 2 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 2 \end{bmatrix}$

Properties of Eigenvalues

1. Any square matrix A and its transpose A' have the same eigenvalues.
2. The eigenvalues of a triangular matrix are just the diagonal elements of the matrix.
3. The eigenvalues of an idempotent matrix are either zero or unity.
4. The sum of the eigenvalues of a matrix is the sum of the elements of the principal diagonal.
5. The product of the eigenvalues of a matrix A is equal to its determinant.
6. If λ is an eigenvalue of a matrix A, then $\frac{1}{\lambda}$ is the eigenvalue of A^{-1}.
7. If λ is an eigenvalue of an orthogonal matrix, then $\frac{1}{\lambda}$ is also its eigenvalue.
8. If $\lambda_1, \lambda_2, ...\lambda_n$ are the eigenvalues of a matrix A, then A^m has the eigenvalues $\lambda_1^m, \lambda_2^m, ...\lambda_n^m$ where m is a positive integer.

Cayley-Hamilton Theorem

Every square matrix satisfies its own characteristic equation.

i.e if the characteristic equation for the n^{th} order square matrix A is

$|A - \lambda I| = (-1)^n \lambda^n + k_1 \lambda^{n-1} + k_2 \lambda^{n-2} + ... + k_n = 0$, then

$$(-1)^n A^n + k_1 A^{n-1} + k_2 A^{n-2} + ... + k_n I = 0 \qquad \text{......................} \ (5)$$

Multiplying (5) by A^{-1}, we get $(-1)^n A^{n-1} + k_1 A^{n-2} + ... + k_{n-1} I + k_n A^{-1} = 0$.

$$\therefore A^{-1} = -\frac{1}{k_n}\left[(-1)^n A^{n-1} + k_1 A^{n-2} + ... + k_{n-1} I\right].$$

This result gives the inverse of A in terms of $n-1$ powers of A and is considered as a practical method for the computation of the inverse of the large matrices.

From the above example, since $\lambda^3 - 18\lambda^2 + 45\lambda = 0$ is the characteristic equation of the matrix $A = \begin{bmatrix} 8 & -6 & 2 \\ -6 & 7 & -4 \\ 2 & -4 & 3 \end{bmatrix}$ we have by Cayley-Hamilton Theorem, $A^3 - 18A^2 + 45A = O$ where O is the null matrix of order 3.

Inverse of a square matrix: (using Cayley – Hamilton theorem):

If A is a square matrix of order 3 (for convenience) then its characteristic equation $|A - \lambda I| = 0$ can be put in the form $\lambda^3 + k_1\lambda^2 + k_2\lambda + k_3 = 0$.

Then by Cayley – Hamilton theorem, we have

$$A^3 + k_1A^2 + k_2A + k_3I = O$$

Post multiplying by A^{-1}, we have

$$A^2 + k_1A + k_2I + k_3A^{-1} = O, \text{ since } A^{-1}A = I, IA = A$$

Therefore $A^{-1} = -\frac{1}{k_3}[A^2 + k_1A + k_2I]$.

Example: Apply Cayley-Hamilton theorem compute the inverse of the matrix $A = \begin{pmatrix} 0 & 1 & 2 \\ 1 & 2 & 3 \\ 3 & 1 & 1 \end{pmatrix}$ and verify the answer.

Solution: The characteristic equation of A is $|A - \lambda I| = 0$.

That is, $\begin{vmatrix} 0-\lambda & 1 & 2 \\ 1 & 2-\lambda & 3 \\ 3 & 1 & 1-\lambda \end{vmatrix} = 0$ ……………………..(1)

On expanding, we get $\lambda^3 - 3\lambda^2 - 8\lambda + 2 = 0$.

Applying Cayley-Hamilton theorem, $A^3 - 3A^2 - 8A + 2I = O$.

Post multiplying with A^{-1} we have,

$A^2 - 3A - 8I + 2A^{-1} = O$. This implies $A^{-1} = \frac{-1}{2}(A^2 - 3A - 8I)$ ………(2)

Now $A^2 = A \cdot A = \begin{pmatrix} 0 & 1 & 2 \\ 1 & 2 & 3 \\ 3 & 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 & 2 \\ 1 & 2 & 3 \\ 3 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 7 & 4 & 5 \\ 11 & 8 & 11 \\ 4 & 6 & 10 \end{pmatrix}$

Thus (ii), becomes

$$\frac{-1}{2}\left\{\begin{bmatrix} 7 & 4 & 5 \\ 11 & 8 & 11 \\ 4 & 6 & 10 \end{bmatrix} - \begin{bmatrix} 0 & 3 & 6 \\ 3 & 6 & 9 \\ 9 & 3 & 3 \end{bmatrix} - \begin{bmatrix} 8 & 0 & 0 \\ 0 & 8 & 0 \\ 0 & 0 & 8 \end{bmatrix}\right\}.$$

Therefore

$$A^{-1} = \frac{-1}{2}\begin{bmatrix} -1 & 1 & -1 \\ 8 & -6 & 2 \\ -5 & 3 & -1 \end{bmatrix}.$$

It can be verified that $AA^{-1} = I$.

Example: Verify Cayley – Hamilton theorem and compute the inverse of the matrix

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 3 & 2 & 3 \\ 1 & 1 & 2 \end{bmatrix}$$

The characteristic equation of A is $|A - \lambda I| = 0$. Therefore

$$|A - \lambda I| = \begin{vmatrix} 1-\lambda & 1 & 1 \\ 3 & 2-\lambda & 3 \\ 1 & 1 & 2-\lambda \end{vmatrix} = 0$$

Therefore $\lambda^3 - 5\lambda^2 + \lambda + 1 = 0$, is the characteristic equation of A.

We have to show that $A^3 - 5A^2 + A + I = O$ (replacing λ by A).

$$A^2 = A.\ A = \begin{bmatrix} 1 & 1 & 1 \\ 3 & 2 & 3 \\ 1 & 1 & 2 \end{bmatrix}\begin{bmatrix} 1 & 1 & 1 \\ 3 & 2 & 3 \\ 1 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 5 & 4 & 6 \\ 12 & 10 & 15 \\ 6 & 5 & 8 \end{bmatrix}$$

$$A^3 = A.A^2 = \begin{bmatrix} 1 & 1 & 1 \\ 3 & 2 & 3 \\ 1 & 1 & 2 \end{bmatrix} \begin{bmatrix} 5 & 4 & 6 \\ 12 & 10 & 15 \\ 6 & 5 & 8 \end{bmatrix} = \begin{bmatrix} 23 & 19 & 29 \\ 57 & 47 & 72 \\ 29 & 24 & 37 \end{bmatrix}$$

L.H.S. of $A^3 - 5A^2 + A + I$ becomes,

$$\begin{bmatrix} 23 & 19 & 27 \\ 57 & 47 & 72 \\ 29 & 24 & 37 \end{bmatrix} - 5 \begin{bmatrix} 5 & 4 & 6 \\ 12 & 10 & 15 \\ 6 & 5 & 8 \end{bmatrix} + \begin{bmatrix} 1 & 1 & 1 \\ 3 & 2 & 3 \\ 1 & 1 & 2 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 23-25+1+1 & 19-20+1+0 & 29-30+1+0 \\ 57-60+3+0 & 47-50+2+1 & 72-75+3+0 \\ 29-30+1+0 & 24-25+1+0 & 37-40+2+1 \end{bmatrix}$$

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

Therefore $A^3 - 5A^2 + A + I = 0$

Hence Cayley – Hamilton theorem is verified.

To find the inverse of A

We have $A^3 - 5A^2 + A + I = O$. Post multiplying by A^{-1} we have

$$A^2 - 5A + I + A^{-1} = O$$

Therefore $A^{-1} = -A^2 + 5A - I$

$$\Rightarrow A^{-1} = -\begin{bmatrix} 5 & 4 & 6 \\ 12 & 10 & 15 \\ 6 & 5 & 8 \end{bmatrix} + 5 \begin{bmatrix} 1 & 1 & 1 \\ 3 & 2 & 3 \\ 1 & 1 & 2 \end{bmatrix} - \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \Rightarrow A^{-1} = \begin{bmatrix} -1 & 1 & -1 \\ 3 & -1 & 0 \\ -1 & 0 & 1 \end{bmatrix}$$

Example: Using Cayley-Hamilton theorem, find A^8, if $A = \begin{bmatrix} 1 & 2 \\ 2 & -1 \end{bmatrix}$.

Solution: The characteristic equation of A is given by $\begin{vmatrix} 1-\lambda & 2 \\ 2 & -1-\lambda \end{vmatrix} = 0$ or $\lambda^2 - 5 = 0$.

$$A^8=(A^2)^4=(A^2-5I+5I)^4$$
$$=(A^2-5I)^4+4(A^2-5I)^3 5I+6(A^2-5I)^2(5I)^2+4(A^2-5I)(5I)^3+(5I)^4$$

By Cayley-Hamilton theorem, $A^2-5I=0$.

$\therefore A^8=(5I)^4=625I$.

Example: If $A=\begin{bmatrix} 2 & 2 & 1 \\ 1 & 3 & 1 \\ 1 & 2 & 2 \end{bmatrix}$, then find the eigenvalues of A^{-1} and hence, find its determinant.

Solution: The eigenvalues of A are given by the roots of its characteristic equation, i.e $|A-\lambda I|=\lambda^3-7\lambda^2+11\lambda-5=0$.

$\therefore \lambda=5,1,1$.

If λ is an eigenvalue of A, then $\frac{1}{\lambda}$ is an eigenvalue of A^{-1}.

Therefore, the eigenvalues of A^{-1} are $\frac{1}{5},1,1$.

$\therefore |A^{-1}|=\frac{1}{5}$.

<u>Problems:</u>

1. Verify Cayley-Hamilton theorem for the following matrices and find its inverse: a) $\begin{bmatrix} 2 & -1 & 1 \\ -1 & 2 & -1 \\ 1 & -1 & 2 \end{bmatrix}$ b) $\begin{bmatrix} 7 & 2 & -2 \\ -6 & -1 & 2 \\ 6 & 2 & -1 \end{bmatrix}$ c) $\begin{bmatrix} 3 & 2 & 4 \\ 4 & 3 & 2 \\ 2 & 4 & 3 \end{bmatrix}$

2. If $A=\begin{bmatrix} 2 & -1 & 2 \\ -1 & 2 & -1 \\ 1 & -1 & 2 \end{bmatrix}$, find A^4.

3. Find the eigenvalues of A^2-2A+I where $A=\begin{bmatrix} 2 & 3 & 4 \\ 0 & 4 & 2 \\ 0 & 0 & 3 \end{bmatrix}$.

Iterative method to obtain largest eigenvalue of a matrix

(Rayleigh Power method)

Consider the matrix equation $Ax = \lambda x$ which gives the following homogenous system of equations

$$\begin{aligned}
&(a_{11}-\lambda)x_1 + a_{12}x_2 + \ldots.. + a_{1n}x_n = 0 \\
&a_{21}x_1 + (a_{22}-\lambda)x_2 + \ldots.. + a_{2n}x_n = 0 \\
&\ldots. \quad \ldots. \quad \ldots. \quad \ldots. \\
&a_{n1}x_1 + a_{n2}x_2 + \ldots.. + (a_{nn}-\lambda)x_n = 0
\end{aligned}$$

In many applications, it is required to compute numerically the largest eigenvalue and the corresponding eigenvector.

We start with a column vector x which is as near the solution as possible and evaluate Ax, which is written as $\lambda^{(1)}x^{(1)}$ after normalisation, that is, we obtain $\lambda^{(1)}$ from Ax which is largest in magnitude. Hence, the resulting vector $x^{(1)}$ will have largest component unity. This gives the first approximation $\lambda^{(1)}$ to the largest eigenvalue and $x^{(1)}$ to the eigenvector. Similarly, we evaluate $Ax^{(1)} = \lambda^{(2)}x^{(2)}$ which gives the second approximation. We repeat this process till $x^{(r)} - x^{(r-1)}$ becomes negligible. Then, $\lambda^{(r)}$ will be the largest eigenvalue of A and $x^{(r)}$, the corresponding eigenvector.

This iterative method to find the dominant eigenvalue of a matrix is known as Rayleigh's power method.

For the initial approximation, we can take the vector $x = [1 \quad 0 \quad 0]'$, if it is not explicitly provided.

Note: To find the smallest eigenvalue of A, apply the Rayleigh's power method to A^{-1} and obtain its largest eigenvalue. Since $\dfrac{1}{\lambda}$ is an eigenvalue of A^{-1} if λ is an eigenvalue of A, the resultant largest eigenvalue of A^{-1} is the smallest eigenvalue of A.

Example: Determine the largest eigen value and the corresponding eigen vector of the matrix A = $\begin{bmatrix} 2 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 2 \end{bmatrix}$ by power method taking the initial vector as [1, 1, 1]^t.

Solution: Given $X^{(0)} = (1, 1, 1)^t = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ is the initial eigen vector

$$AX^{(0)} = \begin{bmatrix} 2 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} = 2 \begin{bmatrix} 1 \\ 0 \\ 0.5 \end{bmatrix} = \lambda^{(1)} X^{(1)}$$

(Since 2 is the largest value in $\begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}$)

$$AX^{(1)} = \begin{bmatrix} 2 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0.5 \end{bmatrix} = \begin{bmatrix} 2.5 \\ 0 \\ 2 \end{bmatrix} = 2 \cdot 5 \begin{bmatrix} 1 \\ 0 \\ 0.8 \end{bmatrix} = \lambda^{(2)} X^{(2)}$$

$$AX^{(2)} = \begin{bmatrix} 2 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0.8 \end{bmatrix} = \begin{bmatrix} 2.8 \\ 0 \\ 2.6 \end{bmatrix} = 2 \cdot 8 \begin{bmatrix} 1 \\ 0 \\ 0.93 \end{bmatrix} = \lambda^{(3)} X^{(3)}$$

$$AX^{(3)} = \begin{bmatrix} 2 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0.93 \end{bmatrix} = \begin{bmatrix} 2.93 \\ 0 \\ 2.86 \end{bmatrix} = 2 \cdot 93 \begin{bmatrix} 1 \\ 0 \\ 0.98 \end{bmatrix} = \lambda^{(4)} X^{(4)}$$

$$AX^{(4)} = \begin{bmatrix} 2 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0.98 \end{bmatrix} = \begin{bmatrix} 2.98 \\ 0 \\ 2.96 \end{bmatrix} = 2 \cdot 98 \begin{bmatrix} 1 \\ 0 \\ 0.99 \end{bmatrix} = \lambda^{(5)} X^{(5)}$$

$$AX^{(5)} = \begin{bmatrix} 2 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0.99 \end{bmatrix} = \begin{bmatrix} 2.99 \\ 0 \\ 2.98 \end{bmatrix} = 2.99 \begin{bmatrix} 1 \\ 0 \\ 0.997 \end{bmatrix} = \lambda^{(6)} X^{(6)}$$

$$AX^{(6)} = \begin{bmatrix} 2 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0.997 \end{bmatrix} = \begin{bmatrix} 2.997 \\ 0 \\ 2.994 \end{bmatrix} = 2.997 \begin{bmatrix} 1 \\ 0 \\ 0.999 \end{bmatrix} = \lambda^{(7)} X^{(7)}.$$

Therefore we can conclude that the largest eigen value is approximately 3 and the corresponding eigen vector is $(1, 0, 1)^t$.

Example: Determine the largest eigen value and the corresponding eigen vector of the matrix A $= \begin{bmatrix} 6 & 2 & 2 \\ -2 & 3 & -1 \\ 2 & -1 & 3 \end{bmatrix}$ by power method taking the initial vector as $[1, 1, 1]^t$.

Solution: Given $X^{(0)} = (1, 1, 1)^t = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ is the initial eigen vector

$$A\,X^{(0)} = \begin{bmatrix} 6 & 2 & 2 \\ -2 & 3 & -1 \\ 2 & -1 & 3 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 6 \\ 0 \\ 4 \end{bmatrix} = 6\begin{bmatrix} 1 \\ 0 \\ 0.67 \end{bmatrix} = \lambda^{(1)} X^{(1)}$$

(Since 6 is the largest value in $\begin{pmatrix} 6 \\ 0 \\ 4 \end{pmatrix}$)

$$AX^{(1)} = \begin{bmatrix} 6 & 2 & 2 \\ -2 & 3 & -1 \\ 2 & -1 & 3 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \\ 0.67 \end{bmatrix} = \begin{bmatrix} 7.34 \\ -2.67 \\ 4.01 \end{bmatrix} = 7.34\begin{bmatrix} 1.0 \\ -0.36 \\ 0.55 \end{bmatrix} = \lambda^{(2)} X^{(2)}$$

$$AX^{(2)} = \begin{bmatrix} 6 & 2 & 2 \\ -2 & 3 & -1 \\ 2 & -1 & 3 \end{bmatrix}\begin{bmatrix} 1 \\ -0.36 \\ 0.55 \end{bmatrix} = \begin{bmatrix} 7.82 \\ -3.63 \\ 4.01 \end{bmatrix} = 7.82\begin{bmatrix} 1.0 \\ -0.46 \\ 0.51 \end{bmatrix} = \lambda^{(3)} X^{(3)}$$

$$AX^{(3)} = \begin{bmatrix} 6 & 2 & 2 \\ -2 & 3 & -1 \\ 2 & -1 & 3 \end{bmatrix}\begin{bmatrix} 1 \\ -0.45 \\ 0.51 \end{bmatrix} = \begin{bmatrix} 7.94 \\ -3.89 \\ 4.01 \end{bmatrix} = 7.94\begin{bmatrix} 1.0 \\ -0.69 \\ 0.51 \end{bmatrix} = \lambda^{(4)} X^{(4)}$$

$$AX^{(4)} = \begin{bmatrix} 6 & 2 & 2 \\ -2 & 3 & -1 \\ 2 & -1 & 3 \end{bmatrix}\begin{bmatrix} 1 \\ -0.49 \\ 0.51 \end{bmatrix} = \begin{bmatrix} 7.98 \\ -3.97 \\ 3.99 \end{bmatrix} = 7.98\begin{bmatrix} 1 \\ -0.5 \\ 0.5 \end{bmatrix} = \lambda^{(5)} X^{(5)}$$

$$AX^{(5)} = \begin{bmatrix} 6 & 2 & 2 \\ -2 & 3 & -1 \\ 2 & -1 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ -0.5 \\ 0.5 \end{bmatrix} = \begin{bmatrix} 8 \\ -4 \\ 4 \end{bmatrix} = 8 \begin{bmatrix} 1 \\ -0.5 \\ 0.5 \end{bmatrix} = \lambda^{(6)} X^{(6)}.$$

Therefore the largest eigen value is $\lambda = \lambda^{(6)} = 8$ and the corresponding eigen vector $X = X^{(6)} = (1, -0.5, 0.5)^t$.

<u>Problems:</u>

1. Find the largest eigenvalue and the corresponding eigenvector of the matrix $\begin{bmatrix} 1 & -3 & 2 \\ 4 & 4 & -1 \\ 6 & 3 & 5 \end{bmatrix}$ using the initial approximation $\begin{bmatrix} 1 & 0 & 0 \end{bmatrix}'$. Carry out 4 iterations.

2. Find the dominant eigenvalue and the corresponding eigenvector of the matrix $\begin{bmatrix} 6 & -2 & 2 \\ -2 & 3 & -1 \\ 2 & -1 & 3 \end{bmatrix}$. Take an appropriate initial approximation and carry out 4 iterations.