Summary | Matrices

Introduction

Revise Matrices unit from G.C.E. (A/L) Combined Mathematics and G.C.E. (O/L) Mathematics.

Types of matrices

Square matrix

Number of columns equal to number of rows.

> ⓘ **Main diagonals of a square matrix**
>
> Formed by elements having equal subscripts.

Diagonal matrix

A square matrix whose only non-zero elements are main-diagonal elements. Denoted by D. Subset of triangular matrices.

Identity matrix or Unit matrix

A diagonal matrix whose diagonal elements are all equal to 1. Denoted by I. Subset of diagonal matrices.

Zero matrix / Null matrix

All elements are 0.

Column matrix (column vector)

Only 1 column.

Row matrix (row vector)

Only 1 row.

Triangular matrix

Upper triangular matrix or lower triangular matrix.

Upper triangular matrix

All elements below the main diagonal are 0. Subset of square matrices.

Lower triangular matrix

All elements above the main diagonal are 0. Subset of square matrices.

Matrix operations

Addition and subtraction

Order of the 2 matrices must be same. Matrix obtained by adding or subtracting corresponding elements.

Scalar multiplication

Matrix obtained by multiplying all elements by the scalar.

Matrix multiplication

Explained on a separate page: [Matrix multiplication]

> ⓘ **Note**
>
> Other operations are also defined in separate pages.

Matrix Multiplication

Defined only when the number of columns of the first matrix is equal to the number of rows of the second matrix.

Suppose $A = (a_{ij})_{m \times p}$ and $B = (b_{ij})_{p \times n}$.

$$A \times B = C = (c_{ij})_{m \times n} \ \text{ where } \ c_{ij} = \sum_{k=1}^{p} a_{ik} \times b_{kj}$$

> ⓘ **Note**
>
> - Generally $A \times B \neq B \times A$
> - $A \times B = 0 \not\Rightarrow A = 0 \vee B = 0$
> - $A \neq 0 \wedge B \neq 0 \not\Rightarrow A \times B \neq 0$

Properties of matrix multiplication

A, B, C, I (Identity) matrices must be chosen so that below-mentioned product matrices are defined.

1. Associative: $A(BC) = (AB)C$
2. Right distributive over addition: $(A + B)C = AC + BC$
3. Left distributive over addition: $C(A + B) = CA + CB$
4. $AI = IA = A$

Transpose

Matrix obtained from a given matrix by interchanging its rows and columns. Denoted by a superscript T, like A^T.

Properties

1. $(A^T)^T = A$
2. Distributive over addition: $(A + B)^T = A^T + B^T$
3. $(kA)^T = kA^T$
4. $(A \times B)^T = B^T \times A^T$

More Types of Matrices

Symmetric matrix

If $A = A^T$. Subset of square matrices.

Skew Symmetric matrix

If $A = -A^T$. Subset of square matrices. All elements in main diagonal are 0.

> ⓘ **Note**
>
> Any square matrix can be expressed as a sum of a symmetric matrix and a skew-symmetric matrix.

Complex conjucate of a matrix

Suppose $A = (\, a_{\text{ij}} \,)_{n \times n}$. Complex conjucate matrix of A is:

$$A^* = \overline{A^T} = (\, \overline{a_{\text{ji}}} \,)_{n \times n}$$

Hermitian matrix

A square matrix A is said to be a Hermitian matrix **iff** $A = \overline{A^T}$.

Skew Hermitian matrix

A square matrix A is said to be a Hermitian matrix **iff** $A = -\overline{A^T}$.

Determinant

Defined only for square matrices. Denoted by $|A|$.

For 2x2

$$|A| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

For higher order

Minor of an element

Suppose $A = (a_{ij})$.

Minor of an element a_{ij}, is the matrix obtained by deleting i^{th} row and j^{th} column of A. Denoted by M_{ij}.

Co-factor of an element

Suppose $A = (a_{ij})$.

Co-factor of an element a_{ij}, is defined as (commonly denoted as A_{ij}):

$$A_{ij} = (-1)^{i+j} |M_{ij}|$$

Definition

If $A = (a_{ij})_{n \times n}$ then the **determinant** of A is defined by:

$$|A| = \sum_{j=1}^{n} a_{ij} A_{ij}$$

where $1 \leq j \leq n$.

Properties of determinants

- $|A^T| = |A|$
- Every element of a row or column of a matrix is 0 then the value of its determinant is 0.
- If 2 columns or 2 rows of a matrix are identical then its determinant is 0.
- If A and B are two square matrices then $|AB| = |A||B|$.
- The value of the determinant of a matrix remains unchanged if a scalar multiple of a row or column is added to any other row or column.
- If a matrix B is obtained from a square matrix A by an interchange of two columns or rows: $|B| = -|A|$.
- If every entry in any row or column is multiplied by k, then the whole determinant is multiplied by k.

Composition

$$\begin{vmatrix} a & b & c_1 + c_2 \\ d & e & f_1 + f_2 \\ g & h & i_1 + i_2 \end{vmatrix} = \begin{vmatrix} a & b & c_1 \\ d & e & f_1 \\ g & h & i_1 \end{vmatrix} + \begin{vmatrix} a & b & c_2 \\ d & e & f_2 \\ g & h & i_2 \end{vmatrix}$$

In relation with eigenvalues

For a $n \times n$ matrix A with n number of eigenvalues:

$$|A| = \prod_{i=1}^{n} \lambda_i$$

Adjoint

Suppose $A = (a_{ij})_{n \times n}$.

$$\mathrm{adj} A = (A_{ij})^T_{n \times n}$$

Where A_{ij} is the co-factor of a_{ij}.

Properties

Suppose A is a $n \times n$ matrix.

- $\mathrm{adj}(I) = I$
- $\mathrm{adj}(cA) = c^{n-1} \mathrm{adj}(A)$
- $\mathrm{adj}(A^T) = (\mathrm{adj}(A))^T$
- $\mathrm{adj}(A)\,A = A\,\mathrm{adj}(A) = |A| I$

Note

For a 2×2 matrix, $\mathrm{adj}(\mathrm{adj}(A)) = A$.

Inverse

Suppose A and B are square matrices of the same order. If $AB = BA = I$ then B is called the inverse of A and is denoted by A^{-1}.

$$A^{-1} = \frac{\text{adj } A}{|A|}$$

Singular or Non-singular

A square matrix is singular **iff** $|A| = 0$. Otherwise its non-singular or invertible.

Properties of Inverse

- $(AB)^{-1} = B^{-1}A^{-1}$
- $(A^T)^{-1} = (A^{-1})^T$
- $A\,(\text{adj}A) = (\text{adj}A)\,A = |A|I$

Elementary Transformations

- Interchange of any columns or rows
- Addition of multiple of any row or column to any other row or column
- Multiplication of each element of a column or a row by a non-zero constant

When a matrix B is obtained by applying elementary transformations to a matrix A, then A **is equivalent to** B. Denoted by $A \approx B$.

Theorem

The elementary row operations that reduce a given matrix A to the identity matrix, also transform the identity matrix to the inverse of A.

Augmented Matrix

Two matrices are written as a single matrix with a vertical line in-between. Denoted by $(A|B)$.
Example:

$$\left[\begin{array}{cc|c} 1 & 2 & 3 \\ 4 & 5 & 6 \end{array}\right]$$

Inverse using elementary row transformations

Let A be a square matrix with order $n \times n$.

- Start with $(A_{n\times n}|I_n)$
- Repeatedly add **row** transformations (not column) to both of the matrices until the LHS becomes an identity matrix.
  - Convert all elements outside the main diagonal to 0.
  - Convert elements on the main diagonal to 1 by multiplying by a constant.
- When LHS is an identity matrix, RHS is A^{-1}.

> ⚠ **TODO**
>
> What about singular matrices?

Echelon Form

A matrix is in row echelon form (or just "row echelon" form) **iff**:

- All rows having only zero entries are at the bottom.
- For all row that does not contain entirely zeros, the first non-zero entry is 1.
- For 2 successive non-zero rows, the leading 1 in the higher row is further left than the leading 1 in the lower row.

The process of reducing the augmented matrix to row Echelon form is known as **Gaussian elimination**.

Column echelon form

A matrix A is in column echelon form if A^{T} is in row echelon form.

System of Linear Equations

Any system of linear equations can be represented in matrix notation as shown below.

- $a_{11}x + a_{12}y + a_{13}z = b_1$
- $a_{21}x + a_{22}y + a_{23}z = b_2$
- $a_{31}x + a_{32}y + a_{33}z = b_3$

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} \implies AX = B$$

2 types based on B:

- $= 0$: Homogeneous system
- $\neq 0$: Non-homogeneous system

Consistent

When the system of equations has at least 1 solution. Otherwise inconsistent.

Rank

Number of non-zero rows of row echelon form of a matrix A. Denoted by $\text{Rank } A$.

> **Note**
>
> $\text{Rank } A \leq \text{Rank } (A|B)$ is always true.

Relation with non-homogenous system of solutions

Consider the system: $A_{n\times n}X_{n\times 1} = Bn \times 1$.

- $|A| \neq 0 \iff \text{Rank } A = \text{Rank } (A|B) = n \iff \text{unique solution exists}$
- $|A| = 0 \implies \text{no solution} \vee \text{infinitely many solutions}$
- $\text{Rank } A < \text{Rank } (A|B) \implies \text{no solutions}$
- $\text{Rank } A = \text{Rank } (A|B) < n \implies \text{infinitely many solutions}$

Solutions of Homogenous Systems

Consider the system:

$$A_{m\times n}X_{n\times 1} = O_{m\times 1}$$

Any homogenous system is consistent, because $X = O$ is always a solution.

- $\text{Rank } A = \text{Rank } (A|B) = n \iff \text{unique solution exists}$
- $\text{Rank } A = \text{Rank } (A|B) < n \implies \text{infinitely many solutions}$

Solution of Non-homogenous Systems

Method 1: Direct approach

Used when coefficient matrix A is invertible. It means the system has a unique set of solutions.

$$AX = B \implies X = A^{-1}B$$

Method 2: Cramer's Rule

Let $AX = B$, where A is the coefficient matrix and $X = (x_i)_{n\times 1}$.

$$x_i = \frac{|A_i|}{|A|}$$

Where A_i is the matrix obtained by replacing ith column in matrix A by B.

Method 3: Reducing to Echelon Form

Start with $(A|B)$. Convert the LHS to echelon form using elementary row transformations. The solution can be found now. If a contradiction is encountered while solving the equation, that means the system has no solutions.

Eigenvalues & Eigenvectors

Definitions

Characteristic Polynomial

Let A be a $n \times n$ matrix.

$$p(\lambda) = |A - \lambda I|$$

Eigenvalues

Roots of the equation $p(\lambda) = 0$ are the eigenvalues of A.

> **Note**
>
> - Determinant of a matrix can be written in terms of all of its eigenvalues.
> - If λ is an eigenvalue of A, then λ^2 is an eigenvalue of A^2

Eigenvectors

The column vectors satisfying the equation $(A - \lambda_i I)X_i$.

Normalized eigenvectors

An eigenvector with the magnitude (norm) of 1. Normalizing factor k of any eigenvector is:

$$\frac{1}{k} = \sqrt{\sum_{i=1}^{n} X_i^2}$$

Norm

Norm of a column or row matrix $W_{n\times n}$ is denoted by $||W||$ and defined as:

$$||W|| = \sqrt{\sum_{i=1}^{n} w_i^2}$$

Algebraic Multiplicity

If the characteristic polynomial consists of a factor of the form $(\lambda - \lambda_i)^r$ and $(\lambda - \lambda_i)^{r+1}$ is not a factor of the characteristic polynomial then r is the algebraic multiplicity of the eigenvalue λ.

Spectrum

Set of all eigenvalues.

Spectral Radius

$$R = \max\left\{|\lambda_i| \; where \; \lambda_i \in \text{Spectrum}\right\}$$

Linear Independence of Eigenvectors

Suppose $X_1, X_2, X_3, \ldots, X_n$ is a set of eigenvectors. $k_1, k_2, k_3, \ldots, k_n$ is a set of scalars.

All those eigenvectors are independent **iff**:

$$k_1X_1 + k_2X_2 + k_3X_3 + \cdots + k_nX_n = 0 \implies k_1 = k_2 = k_3 = \cdots = k_n = 0$$

For special matrices

Real symmetric matrix

Suppose A is a symmetric matrix with all real entries. Then:

- The eigenvalues of A are all real: $\forall\lambda \in S_A, (\lambda_i \in \mathbb{R})$
- The eigenvectors of A (corresponding to distinct values of λ) are mutually orthogonal

Upper triangular matrix

The eigenvalues are the diagonal entries.

Orthogonal

Consider 2 column matrices v_1 and v_2:

$$v_1 = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \land \; v_2 = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$

Product

The product of v_1 and v_2 is defined as:

$$v_1 \cdot v_2 = \sum_{k=1}^{n} a_k b_k = v_2 \cdot v_1 = v_1^T v_2$$

Orthogonal vectors

v_1 and v_2 are orthogonal **iff** $v_1 \cdot v_2 = 0$.

For a set of n column vectors, they are orthogonal **iff** they are pairwise orthogonal. That is:

$$\forall i, j \in \{1, \ldots, n\} \land i \neq j, (v_i \cdot v_j = 0)$$

> ⓘ **Note**
>
> v_1, v_2 are orthogonal $\implies$ v_1, v_2 are linearly independent.
>
> Converse is **not** true.

Orthogonal matrix

For a square matrix A with real entries, it is orthogonal **iff** $A^{-1} = A^T$.

A matrix is orthogonal **iff** sum of the squared elements of any row or column is 1.

Properties

- $\det A = \pm 1$
- A is invertible, non-singular
- $A^{-1} = A^T$
- It is diagonalizable over $\mathbb{C}$ (may not be, over $\mathbb{R}$)
- $\text{rank } A = \text{order } A$
- Product of 2 orthogonal matrices of the same order is also an orthogonal matrix
- The columns or rows of an orthogonal matrix form an orthogonal set of vectors

Orthonormal

For a set of n column vectors, they are orthonormal **iff**:

- They are pairwise orthogonal **AND**
- For all n column vectors their norm is 1 $\forall i \in \{1, \ldots, n\}, ||v_i|| = 1$

Trace

Suppose $A = (a_{ij})_{n \times n}$ is an square matrix. Trace of A is the sum of the diagonal entries.

$$\text{trace}(A) = \text{Tr}(A) = \sum_{i=1}^{n} a_{ii}$$

Trace is also equal to the sum of eigenvalues.

$$\text{trace}(A) = \sum \lambda_i \text{ where } \lambda_i \in \text{spectrum of } A$$

Diagonalization

Similar matrices

2 square matrices A and B of the same order, are similar **iff** there exists an invertible matrix P such that:

$$B = P^{-1}AP$$

Similarity of 2 matrices is commutative.

Similar matrices have the set of eigenvalues.

> ⓘ **Note**
>
> If A and B are similar, then A^2 and B^2 are similar.

Definition

A matrix A is **diagonalizable** if it is similar to a diagonal matrix.

$$\exists\, D, P \ \text{ s.t. } D = P^{-1}AP$$

Here:

- D is a diagonal matrix
- P is an invertible matrix

Steps

- Find eigenvalues of $A_{n\times n}: \lambda_1, \lambda_2, \ldots, \lambda_n$
- Find corresponding eigenvectors: $X_1, X_2, \ldots, X_n$
- Construct P by joining the eigenvectors as columns

$$P = (X_1 X_2 \dots X_n)_{n\times n} \quad \wedge \quad D = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$$

Note

Order of those eigenvectors is **not** a problem. Here the matrix P differs based on the order, and hence is not unique.

Real symmetric matrix

Suppose $A_{n\times n}$ is a **real symmetric matrix** . If it has **distinct** eigenvalues then it has n **mutually orthogonal linearly-independent** eigenvectors.

Hence the diagonalizing matrix P (formed by using the normalized eigenvectors) is an **orthogonal matrix**.

Uses

Finding integer powers

Suppose $A_{n\times n}$ is diagonalizable, and $k \in \mathbb{R}$.

$$A = P^{-1}DP \implies A^k = P^{-1}D^kP$$

Cayley-Hamilton Theorem

If $p(\lambda)$ is the characteristic polynomial of the matrix $A_{n\times n}$, then $p(A) = O$

Uses

- Easily compute the inverse of a matrix
- Easily express higher powers of a matrix in terms of its lower powers

Matrix Norms

Let $A_{n \times n}$. A norm of A is denoted by $||A||$.

Definitions

Suppose $A = (a_{ij})_{m \times n}$ for all the definitions below.

1-norm

Maximum of the absolute column sums.

$$\|A\|_1 = \max \left\{ \sum_{i=1}^{m} |a_{ij}| \;;\; j \in [1, n] \right\}$$

2-norm

Square root of the sum of all elements squared. Aka. Euclidean norm, or Frobenius norm. Defined for non-square matrices as well.

$$\left(\|A\|_2 \right)^2 = \left(\|A\|_E \right)^2 = \sum_{i=1}^{m} \sum_{j=1}^{n} (a_{ij})^2$$

Infinity norm

Maximum of the absolute row sums.

$$\|A\|_\infty = \max \left\{ \sum_{j=1}^{n} |a_{ij}| \;;\; i \in [1, m] \right\}$$

> ⓘ **Note**
>
> For any matrix $X \in \mathbb{R}^n$:
>
> $$\|X\|_\infty \leq \|X\|_2 \leq \|X\|_1$$

Vector norm

Norm defined for column vectors.

Induced norm

Aka. operator norm, subordinate norm.

Suppose $A = (a_{ij})_{m\times n}$. The induced norm is defined for A with respect to a given norm, $\|\|$.

$$\|A\|_{\text{ind}} = \max_{\|X\|=1} \|AX\|$$

Properties of Norms

Works for all types of norms.

Suppose A, B are $m \times n$ ordered.

1. $\|A\| \geq 0$
2. $\|A\| = 0 \iff A = 0$
3. $\|kA\| = |k| \times \|A\|$
4. $\|A + B\| \leq \|A\| + \|B\|$ (triangle inequality)
5. $\|AB\| \leq \|A\| \times \|B\|$

Unit Ball

A unit ball in $\mathbb{R}^n$ with respect to a norm $\|\|$.

$$\{X \mid X \in \mathbb{R}^n,\ \|X\| \leq 1\}$$

Unit disc

When $n = 2$, unit ball is also called the unit disc.