Matrices

What are Matrices?

A **matrix** is a rectangular array of numbers, symbols, or expressions arranged in **rows and columns**. It is used in various mathematical and computational applications, including statistics, machine learning, and linear algebra.

Notation & Structure

A matrix is typically denoted by an **uppercase letter**, such as **A**, and its elements are represented using subscripts.

For example, a $m \times n$ **matrix** (where m is the number of rows and n is the number of columns) is written as:

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mn} \end{bmatrix}$$

where a_{ij} represents the element in the **i th row** and **j th column**.

Types of Matrices

1. **Row Matrix** – A matrix with only one row.

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

2. **Column Matrix** – A matrix with only one column.

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

3. **Square Matrix** – A matrix with an equal number of rows and columns (m = n).

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

4. **Diagonal Matrix** – A square matrix where all non-diagonal elements are zero.

$$A = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 8 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

5. **Identity Matrix (II)** – A square matrix with **1s** on the diagonal and **0s** elsewhere.

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

6. **Zero Matrix** – A matrix where all elements are **0**.

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

7. **Upper Triangular Matrix (U) -** A **square matrix** is called **upper triangular** if **all the elements below the main diagonal are zero**.

$$U = \begin{bmatrix} 2 & 3 & 1 \\ 0 & 5 & 4 \\ 0 & 0 & 6 \end{bmatrix}$$

8. **Lower Triangular Matrix (L) -** A **square matrix** is called **lower triangular** if **all the elements above the main diagonal are zero**.

$$L = \begin{bmatrix} 7 & 0 & 0 \\ 2 & 5 & 0 \\ 3 & 8 & 1 \end{bmatrix}$$

Basic Matrix Operations

1. **Addition & Subtraction**

   Matrices of the same dimensions can be added or subtracted element-wise.

2. **Multiplication**

   - **Scalar Multiplication**: Multiply each element by a scalar.
   - **Matrix Multiplication**: Multiply rows of the first matrix by columns of the second matrix (only possible if the number of columns in the first matrix equals the number of rows in the second).

3. **Transpose (A^T)**

   Flips the matrix over its diagonal, switching rows and columns.

4. **Determinant & Inverse**

   - **Determinant**: A scalar value that provides information about a matrix's properties.
   - **Inverse (A^{-1})**: Exists only for **square, non-singular matrices** ($\det(A) \neq 0$).

Applications of Matrices

- **Linear Algebra** – Solving systems of equations.
- **Machine Learning & AI** – Representing datasets and transformations.
- **Computer Graphics** – Image processing and transformations.
- **Probability & Statistics** – Covariance matrices, Markov chains.
- **Physics & Engineering** – Quantum mechanics, electrical circuits.

Gaussian Elimination

Gaussian Elimination is a method for solving **systems of linear equations** by transforming the system's augmented matrix into an **upper triangular form** (row echelon form) using **elementary row operations**. Once in this form, the system can be solved easily using **back-substitution**.

Steps of Gaussian Elimination

For a system of linear equations:

Ax = b

where A is a coefficient matrix, x is the unknowns vector, and b is the constants vector.

We transform the augmented matrix [A | b] into an **upper triangular form** through the following steps:

Step 1: Convert to Augmented Matrix

Convert the system into an augmented matrix form:

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} & |b_1 \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} & |b_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & |\vdots \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mn} & |b_m \end{bmatrix}$$

Step 2: Forward Elimination

- Make the first column below the leading coefficient (**pivot**) zero by subtracting multiples of the first row from the lower rows.
- Move to the second pivot and repeat until we get an upper triangular matrix.

Step 3: Back-Substitution

- Solve for the last variable first.
- Substitute it into the second-last equation, and so on, until all unknowns are found.

What is Row Echelon Form (REF)?

Row Echelon Form is a type of matrix form used in **Gaussian elimination** to solve systems of linear equations. It simplifies the system step-by-step until it's easy to solve using **back-substitution**.

A matrix is in Row Echelon Form (REF) if:

1. All **nonzero rows** are above any rows of **all zeros**.
2. The **leading entry** (first non-zero number from the left) in each nonzero row is **to the right** of the leading entry in the row above it.
3. All entries **below** a leading entry (pivot) are **zeros**.

Example

This matrix is in **row echelon form**:

$$\begin{bmatrix} 1 & 2 & -1 & 0 \\ 0 & 3 & 2 & 1 \\ 0 & 0 & 5 & -2 \end{bmatrix}$$

Not in REF Example

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

Why?

Why?

- Each leading number moves to the right as we go down rows
- Everything **below** each leading number is **zero**
- All zero rows (if any) would go to the bottom

This is **not** in row echelon form because the first row has a **zero** as the first element, while the second row starts with a **leading 1** — violating the "pivot moves to the right" rule.

What is Reduced Row Echelon Form (RREF)?

Reduced Row Echelon Form (also called **row canonical form**) is a specific form of a matrix where solving systems of linear equations becomes very straightforward — like reading the answers directly off the matrix.

It's a stricter version of the **row echelon form** (REF) used in **Gauss-Jordan Elimination**.

A matrix is in RREF if it satisfies all of the following:

1. **All zero rows** (if any) are **at the bottom**.
2. The **leading entry** (first non-zero number from the left) in every non-zero row is **1**.
3. The **leading 1** in each row is **to the right** of the leading 1 in the row above it.
4. Each **leading 1** is the **only non-zero entry** in its column.

Example

This matrix is in **Reduced Row Echelon Form**:

$$\begin{bmatrix} 1 & 0 & 0 & 5 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 3 \end{bmatrix}$$

Interpretation:

- It represents the system: $x=5$, $y=-2$, $z=3$

What is Matrix Rank?

The **rank** of a matrix is the **maximum number of linearly independent rows or columns** in the matrix. It tells you the **"dimension" of the information** in the matrix — basically, how many rows or columns are *not redundant*.

More formally:

For an m × n matrix:

- **Rank** = the number of **leading 1s** (pivots) in its **Row Echelon Form** (REF) or **Reduced Row Echelon Form** (RREF).
- It also equals the number of **linearly independent rows** or **columns**.

What is a Matrix Transpose?

The **transpose** of a matrix is a new matrix that is formed by **flipping the rows and columns** of the original matrix.

If you have a matrix A, its **transpose** is written as A^T .

Key Properties:

1. **Double transpose gives the original matrix:**
   $(A^T)^T = A$
2. **Transpose of a sum:**
   $(A + B)^T = A^T + B^T$
3. **Transpose of a product:**
   $(AB)^T = B^T A^T$
   (Note: The order reverses!)
4. If A is a **symmetric matrix**, then:
   $A = A^T$

Symmetric Matrix

A matrix A is **symmetric** if:

$A^T = A$

In other words, the matrix is **equal to its transpose**.

This means that the entries are **mirrored across the main diagonal**.

Example:

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{bmatrix}$$

Check:

- $A_{12} = A_{21}$
- $A_{13} = A_{31}$
- $A_{23} = A_{32}$

✔ So this is a **symmetric matrix**.

Skew-Symmetric (Antisymmetric) Matrix

A matrix A is **skew-symmetric** if:

$A^T = -A$

This means:

- All **diagonal elements must be 0**, because $a_{ii} = -a_{ii} \Rightarrow a_{ii} = 0$
- The entries **above the diagonal are negatives** of those below.

Example:

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

Check:

- $A_{12} = -A_{21}$
- $A_{13} = -A_{31}$
- Diagonal: all 0

✔ This is a **skew-symmetric matrix**.

What is the Determinant?

The **determinant** is a special number that you can calculate **only for square matrices** (same number of rows and columns). It tells you important things about the matrix, like:

- Whether it's **invertible** (only if determinant ≠ 0)
- The **area/volume** scaling factor of a transformation
- Whether rows/columns are **linearly independent**

We write it as:

$$\det(A) \quad \text{or} \quad |A|$$

1. Determinant of a 2×2 Matrix

For:

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

$$\det(A) = ad - bc$$

2. Determinant of a 3×3 Matrix

For:

$$A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$$

Use the formula:

$$\det(A) = a(ei - fh) - b(di - fg) + c(dh - eg)$$

This is called **cofactor expansion** along the first row.

3. For Bigger Matrices (4×4, etc.)

You use **cofactor expansion** recursively, or use **row operations** (like in Gauss elimination) to reduce it to upper triangular form — then multiply the diagonal elements.

What Is a Cofactor Matrix?

The **cofactor matrix** of a square matrix comprises the **cofactors** of each element of the original matrix.

Each **cofactor** is a signed minor:

$$C_{ij} = (-1)^{i+j} \cdot M_{ij}$$

Where:

- M_{ij} is the **minor** of the element a_{ij} (i.e., the determinant of the smaller matrix you get by deleting the i-th row and j-th column).
- $(-1)^{i+j}$ is the sign factor that alternates across the matrix.

Example with a 3×3 Matrix

Let's say:

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

Find the minors (determinants of 2×2 matrices)

Let's find C_{11} (top-left element's cofactor):

- Remove 1st row and 1st column →

  Submatrix:

  $$\begin{bmatrix} 4 & 5 \\ 0 & 6 \end{bmatrix} \Rightarrow \det = (4)(6) - (5)(0) = 24$$

- Apply sign: $C_{11} = (+1) \cdot 24 = 24$

Do this for **each element**, applying the sign pattern:

$$\begin{bmatrix} + & - & + \\ - & + & - \\ + & - & + \end{bmatrix}$$

After computing all, you get the full **cofactor matrix**.

What is the Adjoint (Adjugate) of a Matrix?

The **adjoint** (also called **adjugate**) of a square matrix is the **transpose of its cofactor matrix**.

$$\mathrm{adj}(A) = (\text{Cofactor matrix of } A)^T$$

So it's made up of the **cofactors** of each element, but then **rows and columns are swapped** (transposed).

To check the Inverse of a Matrix via Adjoint

$$A^{-1} = \frac{1}{|A|} \cdot \mathrm{adj}(A)$$

Here:

- |A| = **Determinant of A**
- $\mathrm{adj}(A)$ = **Adjoint** of A

So when does A^{-1} exist?

For A^{-1} to exist:

$$|A| \neq 0$$

If the **determinant is zero**, then:

- Division by zero occurs → ❌ **Inverse does not exist.**
- The matrix is called **singular**.

So, how do you use this?

To check if A^{-1} exists:

1. **Compute the determinant** |A|
2. ✅ If $|A| \neq 0$, inverse exists.
3. ❌ If $|A| = 0$, inverse does **not** exist.

Cramer's Rule — Solving a System of Linear Equations using Determinants

Cramer's Rule is a **formula-based method** to solve a system of **linear equations** (when the number of equations equals the number of unknowns), using **determinants** of matrices.

Suppose you have a system like this:

$A \cdot \vec{x} = \vec{b}$

Where:

- A is an n×n **coefficient matrix**
- $\vec{x}$ is the column vector of **unknowns**
- $\vec{b}$ is the **constant** column vector

Cramer's Rule says:

If $|A| \neq 0$, then the solution $\vec{x} = [x_1, x_2, \ldots, x_n]^T$ is given by:

$x_i = \frac{|A_i|}{|A|}, \quad \text{for } i = 1, 2, \ldots, n$

Where:

- |A| is the **determinant** of the original matrix.
- A_i is the matrix obtained by **replacing the i-th column of A** with the vector $\vec{b}$.

What is Matrix Multiplication?

Matrix multiplication (also called the **dot product** of two matrices) involves **combining two matrices** to produce a third matrix — but **not** by multiplying element by element (that's called the Hadamard product). Instead, it's done by combining **rows and columns** using **dot products**.

Conditions to Multiply Matrices

To multiply matrix A and matrix B:

- The **number of columns of A** must **equal** the **number of rows of B**.

If:

- A is of size m×n
- B is of size n×p

Then the result C = AB will be of size:

- m×p

How It Works (Step-by-Step)

Let's say:

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \quad B = \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix}$$

To compute AB, take **each row of A** and **dot it with each column of B**:

$$AB = \begin{bmatrix} (1 \times 5 + 2 \times 7) & (1 \times 6 + 2 \times 8) \\ (3 \times 5 + 4 \times 7) & (3 \times 6 + 4 \times 8) \end{bmatrix} = \begin{bmatrix} 19 & 22 \\ 43 & 50 \end{bmatrix}$$

Example with Sizes:

Let A be a **2×3** matrix, and B be a **3×2** matrix:

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}, \quad B = \begin{bmatrix} 7 & 8 \\ 9 & 10 \\ 11 & 12 \end{bmatrix}$$

Then AB will be a **2×2** matrix:

$$AB = \begin{bmatrix} (1 \times 7 + 2 \times 9 + 3 \times 11) & (1 \times 8 + 2 \times 10 + 3 \times 12) \\ (4 \times 7 + 5 \times 9 + 6 \times 11) & (4 \times 8 + 5 \times 10 + 6 \times 12) \end{bmatrix} = \begin{bmatrix} 58 & 64 \\ 139 & 154 \end{bmatrix}$$

Important Notes

- **Matrix multiplication is not commutative**:

  $AB \neq BA \text{ (in general)}$

- But it **is associative**:

  (AB)C = A(BC)

- And **distributive** over addition:

  A(B + C) = AB + AC

Real-life Uses:

- **AI & Machine Learning**: Activations, weights, transformations
- **Computer Graphics**: Rotating, scaling, translating objects
- **Statistics**: Linear regression, covariance matrices

- **Quantum mechanics**, physics, and more!

What is the Inverse of a Matrix?

The **inverse** of a square matrix A is another matrix, denoted A^{-1}, such that:

$$A \cdot A^{-1} = A^{-1} \cdot A = I$$

Where I is the **identity matrix** (like 1 for regular numbers — it doesn't change anything when multiplied).

✅ Only Square Matrices Can Have Inverses:

A matrix must be:

- **Square** (same number of rows and columns), and
- **Non-singular** (i.e., its **determinant ≠ 0**)

If a matrix doesn't meet these conditions, **it has no inverse**.

What is Gauss-Jordan Elimination?

It's a method of using **row operations** to transform a matrix into a **Reduced Row Echelon Form (RREF)**.

For solving:

- AX = B, we use: [A ∣ B]⟶[I ∣ X]

For inverse:

- Start with [A ∣ I]
- Apply row operations until:
  $[A|I] \longrightarrow [I|A^{-1}]$

Allowed Row Operations

You can:

1. **Swap** two rows
2. **Multiply** a row by a nonzero constant

3. **Add/Subtract** a multiple of one row to another

These operations don't change the **solution** to the system or the matrix's identity.

The goal of Gauss-Jordan:

Transform matrix into:

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

Which is the **identity matrix** (for square matrices).
