Matrices

Authored by Afsah Buraaq

June 11, 2025

Contents

Introduction

Before diving into Artificial Intelligence, we must first learn the language it speaks; and at the heart of that language lies something deceptively simple yet deeply powerful: **matrices**.

In the world around us, information is rarely scattered; it is structured. Whether it's the schedule of your classes, the pixels in an image, or the scores in a tournament, data often sits neatly in rows and columns. Matrices give us a way to capture this order. A matrix is nothing but a structured grid; a table of numbers; that lets us organize, compute, and communicate with machines efficiently.

But matrices are more than just tidy boxes of numbers. They are tools for transformation. In physics, they rotate and reflect objects. In economics, they help solve networks of equations. In Artificial Intelligence, they form the foundation for tasks like image recognition, natural language processing, and neural networks.

This chapter begins with the basics: what a matrix is, how to read and write one, and what kinds of matrices exist. From there, we'll explore how to perform operations like addition, scaling, and multiplication. We'll also discuss how matrices are used in real-world contexts, particularly in the systems that power modern AI.

By the end of this chapter, you will not only understand matrices; you will think in them.

Chapter 1

Understanding the Structure and Dimensions of a Matrix

In this subchapter, we will cover the following ideas:

- What exactly is a matrix and how it looks
- The vocabulary we use when talking about matrices; like elements, rows, and columns
- How we describe the size of a matrix; its dimensions

These are the essential first steps toward thinking in matrices. Once you know how to speak this language, you can begin to use it in meaningful ways.

What is a Matrix?

A matrix is a rectangular arrangement of objects—most often numbers—placed in rows and columns. Think of it like a neatly organized table where each box holds a value. These values are called the **elements** of the matrix.

We write a matrix by placing its elements inside large square brackets. The order of the numbers tells us how they are arranged: first by row; then by column.

Raw Idea

Imagine your school lunch menu for the week, where days go down the side (rows) and meals go across the top (columns). Monday's lunch is in the first row; first column. Wednesday's dessert might be in the third row; third column. This whole table? That's a matrix.

Formal Definition

A matrix is a rectangular array of values arranged in horizontal rows and vertical columns. Each value is called an element. A matrix is usually enclosed in brackets and may contain numbers, symbols, or expressions.

Matrix Vocabulary: Elements, Rows, Columns

To work with matrices, we use some standard words:

- **Element**: A single value in the matrix
- **Row**: A horizontal line of elements
- **Column**: A vertical line of elements

Each element in a matrix is identified by its position: the row it is in and the column it is in. For example, the element in the 2nd row and 3rd column is denoted as $a_{2,3}$.

Let's see an example matrix:

$$A = \begin{bmatrix} 1 & 3 & 5 \\ 7 & 9 & 11 \end{bmatrix}$$

Here:

- Matrix A has 2 rows and 3 columns
- The element $a_{1,2} = 3$; this means the value in row 1, column 2 is 3
- The element $a_{2,3} = 11$

Dimensions of a Matrix

The dimension of a matrix tells us how big it is: how many rows and how many columns it has.

We write the dimension of a matrix as:

$$\text{Number of rows} \times \text{Number of columns}$$

So, the matrix A above is a 2×3 matrix (2 rows; 3 columns).

Examples

$$B = \begin{bmatrix} 4 & 5 \\ 6 & 7 \end{bmatrix}$$

Matrix B is a 2×2 matrix.

$$C = \begin{bmatrix} 2 & 3 & 4 & 5 \end{bmatrix}$$

Matrix C is a 1×4 matrix (1 row; 4 columns). This is called a **row vector**.

$$D = \begin{bmatrix} 8 \\ 9 \\ 10 \end{bmatrix}$$

Matrix D is a 3×1 matrix (3 rows; 1 column). This is called a **column vector**.

Checkpoint: Quick Questions

Q1. What is the dimension of the matrix below?

$$E = \begin{bmatrix} 2 & 4 \\ 6 & 8 \\ 10 & 12 \end{bmatrix}$$

A) 3×2

B) 2×3

C) 2×2

D) 3×3

Answer: A)

Q2. What is the element $a_{2,1}$ in matrix A where

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

Answer: 4

Chapter 2

Types of Matrices

In this subchapter, we will explore different types of matrices and understand how each one is unique in its structure and use. These include:

- Row Matrix
- Column Matrix
- Square Matrix
- Zero Matrix
- Diagonal Matrix
- Identity Matrix

Each of these special kinds of matrices plays an important role in different areas of mathematics, computing, and Artificial Intelligence.

Row Matrix

Raw Idea

A row matrix is like a single shelf with boxes arranged left to right. All the data is in a single horizontal line.

Formal Definition

A row matrix is a matrix that has only one row and one or more columns. It is of order $1 \times n$.

Example:

$$R = \begin{bmatrix} 3 & 5 & 7 & 9 \end{bmatrix}$$

Matrix R is a row matrix of size 1×4.

Column Matrix

Raw Idea

Think of a column matrix as a vertical stack—like books on a tower—one on top of the other.

Formal Definition

A column matrix is a matrix that has only one column and one or more rows. It is of order $m \times 1$.

Example:

$$C = \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix}$$

Matrix C is a column matrix of size 3×1.

Square Matrix

Raw Idea

A square matrix is like a perfect chessboard—same number of rows and columns.

Formal Definition

A square matrix is a matrix that has the same number of rows and columns. That is, it is of order $n \times n$.

Example:

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

Matrix S is a square matrix of size 3×3.

Zero Matrix

Raw Idea

A zero matrix is like an empty container—every place is filled with zero.

Formal Definition

A zero matrix (or null matrix) is a matrix in which all elements are zero.

Example:

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

Matrix Z is a zero matrix of size 2×2.

Diagonal Matrix

Raw Idea

A diagonal matrix has numbers only across the diagonal—like a streak of light running from top-left to bottom-right.

Formal Definition

A diagonal matrix is a square matrix in which all elements except those on the main diagonal are zero.
The main diagonal runs from the top-left corner to the bottom-right corner of the matrix.

Example:

$$D = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & 9 \end{bmatrix}$$

Only the diagonal elements (4, 7, 9) are non-zero.

Identity Matrix

Raw Idea

An identity matrix is like the number 1 for multiplication—when you multiply with it, the other matrix stays the same.

Formal Definition

An identity matrix is a square matrix in which all the elements on the main diagonal are 1 and all other elements are 0. It is denoted by I.

Example:

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

Multiplying any matrix A with an identity matrix I of the same size gives back matrix A:

$$AI = IA = A$$

Checkpoint: Quick Questions

Q1. Which of the following is a diagonal matrix?

A) $\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$

B) $\begin{bmatrix} 3 & 0 \\ 0 & 7 \end{bmatrix}$

C) $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$

D) $\begin{bmatrix} 1 & 0 & 0 \end{bmatrix}$

Answer: B

Q2. What is the order of the identity matrix below?

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

A) 2×2

B) 1×3

C) 3×3

D) 4×4

Answer: C

Chapter 3

Matrix Notation and Indexing

In this subchapter, we will learn how to refer to specific elements in a matrix using notation and indexing. Understanding how to read and write the position of an element is essential for using matrices in code, in math problems, and in AI applications.

Here's what we'll cover:

- How to write elements as $A[i][j]$
- How to understand the position of an element using (row i, column j)
- How indexing can differ in maths and in programming; that is, 1-based vs 0-based indexing

Writing Elements as `A[i][j]`

Raw Idea

A matrix is like a city with rows as streets and columns as buildings. To find one house (element), you need both the street number and the building number. That's what $A[i][j]$ means—"go to row i, then go to column j."

Formal Definition

In a matrix A, the notation $A[i][j]$ or a_{ij} refers to the element located in the i-th row and j-th column.

Example:

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

- $A[1][2] = 7$
- $A[3][3] = 6$
- $A[2][1] = 9$

In mathematics, we write a_{ij} to mean the element at row i, column j.

Understanding Position: (Row i, Column j)

Every element in a matrix has a position, which we describe using a pair of numbers:

- The row number tells us how far down the matrix to go
- The column number tells us how far across

So, a_{31} means "go to the third row, first column."
Example:

$$B = \begin{bmatrix} 10 & 20 \\ 30 & 40 \\ 50 & 60 \end{bmatrix}$$

- $B[1][2] = 20$
- $B[2][1] = 30$
- $B[3][2] = 60$

The general rule is: First row, then column.

Common Confusion: 1-Based vs 0-Based Indexing

Raw Idea

Mathematicians and computer programmers are like two people starting a race from different starting lines. One begins counting from 1; the other from 0. This can lead to confusion!

Formal Explanation

- In mathematics, indexing starts from 1. So the top-left element is $a_{1,1}$.
- In programming languages like Python, indexing starts from 0. So the top-left element is `A[0][0]`.

Let's look at an example matrix and its indices in both styles:
Matrix:

$$C = \begin{bmatrix} 11 & 12 & 13 \\ 21 & 22 & 23 \\ 31 & 32 & 33 \end{bmatrix}$$

Position	Math Indexing (1-based)	Python Indexing (0-based)
Top-left	$a_{1,1}$	`C[0][0]`
Center	$a_{2,2}$	`C[1][1]`
Bottom-right	$a_{3,3}$	`C[2][2]`

Always be careful:

- Use a_{ij} in math class
- Use `A[i][j]` with i, j starting from 0 in code

Checkpoint: Quick Questions

Q1. In matrix

$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$, what is the value of $A[2][1]$?]

A) 2

B) 4

C) 5

D) 6

Answer: B

Q2. In Python

, if `M = [[9, 8], [7, 6]]`, what is `M[1][0]`?]

A) 9

B) 7

C) 8

D) 6

Answer: B

Chapter 4

Matrix Addition and Subtraction

In this subchapter, we will learn how to add and subtract matrices. These operations follow specific rules and are different from how we usually add or subtract numbers.

Here's what we will cover:

- When can we add or subtract matrices?
- How do we add elements one by one?
- How does subtraction work similarly?

Once we learn this, we'll be able to handle matrix equations, model problems in AI, and even describe networks using matrix sums.

Conditions for Matrix Addition or Subtraction

Raw Idea

You can't add apples and oranges—and you can't add or subtract matrices unless they're the same shape. Think of two grids that must "fit" over each other to be combined.

Formal Rule

Two matrices can only be added or subtracted if they have the same dimensions—that is, the same number of rows and columns.
Let A and B both be matrices of order $m \times n$. Then:

- $A + B$ and $A - B$ are defined
- Each element is added or subtracted position by position

Adding Corresponding Elements

Raw Idea

You go box by box and add the numbers that are in the same place—like adding the top-left of one with the top-left of the other, and so on.

Formal Definition

Let:

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \quad B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$$

Then the sum $A + B$ is:

$$A + B = \begin{bmatrix} a_{11} + b_{11} & a_{12} + b_{12} \\ a_{21} + b_{21} & a_{22} + b_{22} \end{bmatrix}$$

Example:

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

$$A + B = \begin{bmatrix} 2+1 & 4+3 \\ 6+5 & 8+7 \end{bmatrix} = \begin{bmatrix} 3 & 7 \\ 11 & 15 \end{bmatrix}$$

Subtraction: Same Method, Different Sign

Raw Idea

Subtraction is just like addition—but instead of combining, you find the difference at each position.

Formal Definition

If matrices A and B have the same order $m \times n$, then:

$$A - B = \begin{bmatrix} a_{11} - b_{11} & a_{12} - b_{12} \\ \vdots & \vdots \\ a_{mn} - b_{mn} & \end{bmatrix}$$

Example:

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

$$A - B = \begin{bmatrix} 9-3 & 7-2 \\ 4-1 & 6-5 \end{bmatrix} = \begin{bmatrix} 6 & 5 \\ 3 & 1 \end{bmatrix}$$

Checkpoint: Quick Questions

Q1. Matrix Addition

Given:

$$X = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \quad Y = \begin{bmatrix} 4 & 3 \\ 2 & 1 \end{bmatrix}$$

Then what is $X + Y$?

A) $\begin{bmatrix} 5 & 5 \\ 5 & 5 \end{bmatrix}$

B) $\begin{bmatrix} 3 & 5 \\ 1 & 3 \end{bmatrix}$

C) $\begin{bmatrix} -3 & -1 \\ 1 & 3 \end{bmatrix}$

D) Not defined

Answer: A

Q2. Matrix subtraction is defined only when:

A) Both matrices are square

B) Matrices are of the same size

C) One is a row matrix

D) Matrices have even elements

Answer: B

Chapter 5

Scalar Multiplication of a Matrix

In this subchapter, we'll understand how to multiply a matrix by a single number—called a scalar. This operation is like stretching or shrinking a matrix while keeping its shape the same.

We'll explore:

- What scalar multiplication means
- What happens when each element is scaled
- Where this concept appears in real-world AI problems

Multiplying Each Element by a Number

Raw Idea

Imagine turning up the brightness on a photo—every pixel becomes more intense, but the shape of the photo doesn't change. That's what multiplying a matrix by a scalar does. Every element gets brighter, duller, or flipped, depending on the number.

Formal Definition

Let A be a matrix and k a scalar (a constant number). The scalar multiplication of A by k is a new matrix kA obtained by multiplying each element of A by k:

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \quad kA = \begin{bmatrix} k \cdot a_{11} & k \cdot a_{12} \\ k \cdot a_{21} & k \cdot a_{22} \end{bmatrix}$$

Concept of Scaling

Multiplying by:

- A positive number greater than 1 $\quad \Rightarrow$ Stretches the matrix
- A number between 0 and 1 $\quad \Rightarrow$ Shrinks the matrix
- A negative number $\quad \Rightarrow$ Flips signs (like a mirror), and may stretch or shrink

Example:

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

$$4A = \begin{bmatrix} 4 \cdot 2 & 4 \cdot (-1) \\ 4 \cdot 0 & 4 \cdot 3 \end{bmatrix} = \begin{bmatrix} 8 & -4 \\ 0 & 12 \end{bmatrix}$$

Checkpoint: Quick Questions

Q1. Scalar Multiplication

Given:

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \quad k = 0$$

What is kA?

A) Same as A

B) A zero matrix

C) Matrix with negative signs

D) Not defined

Answer: B

Q2. What does scalar multiplication change in a matrix?

A) The size

B) The number of rows

C) The individual values only

D) The shape

Answer: C

Chapter 6

Transpose and Multiplication of Matrices

6.1 Transpose of a Matrix

This section introduces the idea of flipping a matrix over its diagonal—an operation called the **transpose**.

You will learn:

- What transpose means and how it's written
- How to swap rows with columns
- Where transpose is useful in matrix algebra and AI

Raw Idea

Imagine a mirror placed along the diagonal of a matrix. The elements reflect across it—what was a row becomes a column, and what was a column becomes a row.

Formal Definition

If A is an $m \times n$ matrix, then the transpose of A, written as A^T, is the $n \times m$ matrix obtained by switching the rows with columns.

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

Switching Rows with Columns

- Row 1 of A becomes Column 1 of A^T
- Row 2 of A becomes Column 2 of A^T
- …and so on

Another Example:

$$A = \begin{bmatrix} 7 & 8 & 9 \end{bmatrix} \quad \Rightarrow \quad A^T = \begin{bmatrix} 7 \\ 8 \\ 9 \end{bmatrix}$$

6.2 Matrix Multiplication

In this section, we uncover the most important matrix operation of all: **matrix multiplication**. Unlike scalar multiplication, this is not about changing values—but about combining two matrices to create something new.

We will learn:

- When matrix multiplication is possible
- How to multiply using dot products
- Why the order matters (non-commutative)
- How AI and real-world problems use this operation

When is Matrix Multiplication Possible?

Raw Idea

You can only multiply two matrices when the *inner numbers match*—just like puzzle pieces clicking together.

Formal Definition

If A is of size $m \times n$ and B is of size $n \times p$, then AB is defined and has size $m \times p$.
The number of **columns in the first matrix** must equal the number of **rows in the second**.

Dot-Product of Rows and Columns

Raw Idea

To get one number in the new matrix, take a row from the first matrix and a column from the second. Multiply their matching parts and add them up.
It's like comparing two lists item-by-item and summing the result.

Formal Definition

Given:

$$A = \begin{bmatrix} a_{11} & a_{12} \end{bmatrix}, \quad B = \begin{bmatrix} b_{11} \\ b_{21} \end{bmatrix}$$

Then:

$$AB = \begin{bmatrix} a_{11} \cdot b_{11} + a_{12} \cdot b_{21} \end{bmatrix}$$

Example

Let:

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

Then:

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

Is Matrix Multiplication Commutative?

No. $\boxed{AB \neq BA}$ in general. The order of multiplication matters.

Real-World Example: Data Transformation

Let's say you have a data matrix where rows are students and columns are subjects. You can multiply it by a matrix that transforms or combines subjects—say, to get an overall score or rating.

Matrix multiplication makes it possible to:

- Apply filters in image processing
- Combine features in data
- Perform weighted sums in neural networks

Checkpoint: Quick Questions

Q1. When is matrix multiplication possible?

A) Same number of rows

B) Same number of columns

C) Columns of first = rows of second

D) Rows of first = rows of second

Answer: C

Chapter 7

Applications and Deeper Matrix Ideas

7.1 Applications of Matrices in Real Life and AI

Why are we learning matrices at all? Because they show up everywhere—from photos on your screen to the neurons in a robot's brain.

We'll explore:

- Image processing
- Data tables in machine learning
- Neural networks
- Information storage

Matrices in Image Processing

Raw Idea

A black-and-white image is just a matrix of numbers, where each number is a shade of gray. Color images? A matrix for each color—red, green, blue.

Real-World Use: We apply filters, detect edges, sharpen or blur—all through matrix multiplication and convolution.

Matrices in Data Tables (Features × Samples)

Imagine a table where:

- Each row is a person
- Each column is a detail: height, age, score, etc.

That's a **feature matrix**. It forms the input of many AI systems.

Neural Networks Use Matrices

Every layer of a neural network is just a series of matrix multiplications:

$$\text{Inputs} \times \text{Weights} = \text{Output}$$

Multiply, activate, pass on. Behind that "AI magic" is matrix multiplication, transposes, and more.

Storing Info in Computers

In memory and programming, we store matrices as 2D arrays:

```
data = [[1, 2], [3, 4]]
```

This is how spreadsheets, pixel data, and even game maps are represented!

Checkpoint

Q1. What is the role of matrices in neural networks?

A) Storing images

B) Multiplying inputs and weights

C) Representing code

D) None of the above

Answer: B

7.2 Determinants and Inverse

In this section, we'll explore the **determinant**, a single number that tells us something deep about a square matrix—its invertibility, its scale, even whether it's safe to divide by it.

We'll also cover the **inverse of a matrix**, a key tool in solving equations.

Determinant of a Matrix

Raw Idea

Think of the determinant as the DNA of a matrix. It tells you if the matrix behaves nicely—or if it's dangerous (non-invertible).

Formal Definition

If:

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

Then:

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

A matrix is invertible (non-singular) if and only if $\det(A) \neq 0$.

Inverse of a Matrix

Raw Idea

An inverse matrix is like the ”undo” button. If matrix A does something to data, then A^{-1} undoes it.

Formal Definition

For a 2×2 matrix A,

$$A^{-1} = \frac{1}{\det(A)} \cdot \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

Only exists if $\det(A) \neq 0$.

Checkpoint

Q1. If $\det(A) = 0$, what does that mean?

A) A is invertible

B) A is not invertible

C) A is square

D) A is zero

Answer: B

7.3 Eigenvalues and Eigenvectors

This is the heart of advanced matrix theory—and the foundation of many AI algorithms.

We will learn:

- What eigenvectors and eigenvalues mean

- Their role in understanding transformations
- Why AI loves them

What are Eigenvectors?

Raw Idea

An eigenvector is a direction that a matrix transformation doesn't change. It might stretch or shrink it, but the direction remains the same.

Formal Definition

If A is a square matrix, then:

$$A\vec{v} = \lambda\vec{v}$$

Where:

- $\vec{v}$ is the eigenvector
- λ is the eigenvalue

Example (Conceptual)

Imagine spinning and stretching a piece of rubber. Most directions get all jumbled—but some directions just get longer. Those special directions are eigenvectors.

In AI

Eigenvectors are used in:

- PCA (Principal Component Analysis) for dimensionality reduction
- Understanding patterns in data
- Creating stable systems in differential equations

Checkpoint

Q1. In $A\vec{v} = \lambda\vec{v}$, what does λ represent?

A) A matrix

B) A rotation

C) A scalar (eigenvalue)

D) A determinant

Answer: C

7.4 Special Matrix Properties

Now we look at matrices that are "special"—those with symmetry, structure, or geometric meaning.

We will learn:

- Symmetric matrices
- Orthogonal matrices
- Identity matrix revisited

Symmetric Matrices

Raw Idea

A symmetric matrix is like a mirror—whatever's on the top-left is the same as the bottom-right.

Formal Definition

A matrix A is symmetric if:

$$A = A^T$$

Orthogonal Matrices

Raw Idea

Orthogonal matrices are like perfect gymnasts—they rotate or flip things without stretching them.

Formal Definition

A matrix Q is orthogonal if:

$$Q^T Q = I$$

This means its transpose is also its inverse.

Why These Matter in AI

- Symmetric matrices often come from distance measures and graph structures.
- Orthogonal matrices help in optimization and numerical stability.

Checkpoint

Q1. Which of the following is true for orthogonal matrices?

A) $Q^TQ = Q$

B) $Q^TQ = I$

C) $Q + Q^T = I$

D) $Q^2 = 0$

Answer: B