5 2025-09-05 | Week 03 | Lecture 05

This lecture is based on section 1.2 in the textbook.

5.1 Matrices and matrix notation

> ***The nexus question of this lecture:*** *What is a matrix, and what are the fundamental algebraic operations we can do with it?*

A ***matrix*** is a rectangular array of objects, usually numbers, which are called ***entries***. If the number of rows and the number of columns are equal, the matrix is said to be a ***square matrix***.

For example,

$$\begin{bmatrix} 1 & 0 & 3 \\ 2 & 5 & -3 \end{bmatrix} \quad \text{or} \quad \underbrace{\begin{bmatrix} 1 & 0 & -7 & 5 \end{bmatrix}}_{\text{a row vector}} \quad \text{or} \quad \underbrace{\begin{bmatrix} 1 \\ 4 \\ 0 \end{bmatrix}}_{\text{a column vector}} \quad \text{or} \quad \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

A matrix with m rows and n columns takes the form

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

Such a matrix is said to be an $m \times n$ ***matrix***. The pair (m, n) is called the ***dimensions*** of the matrix (i.e., the number of rows and number of columns). If $m = n$, then the matrix is said to be a ***square matrix***.

The set of all $m \times n$ matrices with real entries is denoted

$$M_{m\times n}(\mathbb{R})$$

For example, in set notation

$$M_{2\times 3}(\mathbb{R}) = \left\{ \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} : a_{11}, a_{12}, a_{13}, a_{21}, a_{22}, a_{23} \in \mathbb{R} \right\}$$

Notation 14. It is common to denote a matrix compactly using the notation

$$A = [a_{ij}] \quad \text{or} \quad A = (a_{ij})$$

To denote the entry at row i, column j, we write either

$$\text{ent}_{ij}(A) \quad \text{or more simply,} \quad a_{ij}$$

For example, if

$$A = (a_{ij}) = \begin{bmatrix} -1 & 2 & 1 \\ 5 & 4 & -9 \\ 3 & -4 & 7 \end{bmatrix}$$

then $a_{23} = \text{ent}_{23}(A) = -9$, and $a_{21} = \text{ent}_{21} = 5$.

Vectors are a special case of matrices. An n-dimensonal vector is an $n \times 1$ matrices (a column matrix, typically).

5.2 Matrix operations: scaling, addition and multiplication

We can ***multiply matrices by a scalar***, in the obvious way: $10 \times \begin{bmatrix} 3 & 4 \\ -2 & 0 \end{bmatrix} = \begin{bmatrix} 30 & 40 \\ -20 & 0 \end{bmatrix}$.

We can ***add matrices***, also in the obvious way:

$$\begin{bmatrix} 3 & 4 \\ -2 & 0 \end{bmatrix} + \begin{bmatrix} 1 & 5 \\ 2 & 10 \end{bmatrix} = \begin{bmatrix} 4 & 9 \\ 0 & 10 \end{bmatrix}.$$

But note that we can only add two matrices if they have the same dimensons:

$$\begin{bmatrix} 5 & 2 \\ -2 & -1 \end{bmatrix} + \begin{bmatrix} -1 & 10 \\ 4 & 7 \\ 2 & 5 \end{bmatrix} = \text{undefined}.$$

We can also **multiply** matrices, but before defining matrix multiplication, it will be helpful to recall the notion of dot product. Suppose we have two vectors of the same length:

$$X = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix} \quad \text{and} \quad Y = \begin{bmatrix} y_1 & y_2 & \cdots & y_n \end{bmatrix},$$

then the ***dot product*** of X and Y, denoted $X \cdot Y$, is

$$X \cdot Y = x_1 y_1 + x_2 y_2 + \ldots + x_n y_n = \sum_{i=1}^{n} x_i y_i.$$

Matrix multiplication is defined in the following way:

Matrix Multiplication

If $A = (a_{ij})$ is a $p \times n$ matrix and $B = (b_{ij})$ is an $n \times q$ matrix, then we can can think of A and B as

$$A = \begin{bmatrix} \text{---}\, A_1 \,\text{---} \\ \text{---}\, A_2 \,\text{---} \\ \vdots \\ \text{---}\, A_p \,\text{---} \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} | & | & & | \\ B_1 & B_2 & \ldots & B_q \\ | & | & & | \end{bmatrix}$$

where $A_1, \ldots, A_p$ are the $1 \times n$ row vectors

$$A_i = \begin{bmatrix} a_{i1} & a_{i2} & \cdots & a_{in} \end{bmatrix} \quad (i = 1, \ldots, p)$$

and $B_1, \ldots, B_q$ are $n \times 1$ column vectors:

$$B_j = \begin{bmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{nj} \end{bmatrix} \quad (j = 1, 2, \ldots, q)$$

Then the product $P = AB$ is a $p \times q$ matrix $P = (p_{ij})$, whose entries are

$$\begin{aligned} p_{ij} &= A_i \cdot B_j \\ &= \sum_{k=1}^{n} a_{ik} b_{kj}. \end{aligned}$$

For example,

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

Dimensionality requirement for matrix multiplication: we can only multiply two matrices if their dimensions match in the right way. If A is a $p \times n$ matrix, and B is an $\tilde{n} \times q$ matrix, then the product AB is defined if and only if $n = \tilde{n}$. That is,

$$AB \text{ is } \begin{cases} \text{a } p \times q \text{ matrix} & \text{if} \quad n = \tilde{n} \\ \text{undefined} & \text{if} \quad n \neq \tilde{n} \end{cases}$$

Properties of matrix multiplication: Perhaps surprisingly, despite matrix multiplication's complicated definition, it nonetheless behaves sort of like regular multiplication in that the ***associative property*** and ***distributed property*** both hold. That is,

$$ABC = (AB)C = A(BC) \quad \text{(associative property)}$$

and

$$\begin{aligned} A(B+C) &= AB + BC \quad \text{(left distributive property)} \\ (B+C)A &= BA + CA \quad \text{(right distributive property)} \end{aligned}$$

It's not obvious why the properties always hold; they require proof. For now, we will take them as a given.

5.3 Special classes of matrices

5.3.1 Diagonal matrices

If $A = (a_{ij})$ is a square (i.e., $n \times n$) matrix, then the entries $a_{11}, a_{22}, a_{33}, \ldots, a_{nn}$ are called the ***diagonal entries***. A square matrix of the form

$$A = \begin{bmatrix} a_{11} & 0 & 0 & \ldots & 0 \\ 0 & a_{22} & 0 & \ldots & 0 \\ 0 & 0 & a_{33} & \ldots & 0 \\ \vdots & \vdots & & \ddots & \\ 0 & 0 & 0 & \ldots & a_{nn} \end{bmatrix}.$$

is called a ***diagonal*** matrix, and the notation used for such matrices is $A = \text{diag}(a_{11}, \ldots, a_{nn})$.

5.3.2 Identity matrix

An ***identity matrix*** is a diagonal matrix with 1's on its diagonal (and 0's everywhere else). For example,

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \text{et cetera}$$

Identity matrices are always square. The notation for an $n \times n$ identity matrix is I_n, or more simply I.

An identity matrix has the property that when you multiply it by another matrix, it doesn't change the other matrix. For example,

$$\begin{bmatrix} 5 & 6 \\ 2 & 7 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 5 & 6 \\ 2 & 7 \end{bmatrix}$$

It's like multiplying a number by 1.

5.3.3 Zero matrices

A matrix whose entries are all zeros is called a ***zero matrix***, like

$$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{et cetera}$$

The book uses the notation $\mathbf{O}_{m\times n}$ to denote an $m \times n$ zero matrix, or sometimes even just $\mathbf{O}$.