MA 110
Linear Algebra and Differential Equations
Lecture 01

Prof. Sudhir R. Ghorpade
Department of Mathematics
IIT Bombay
http://www.math.iitb.ac.in/∼srg/

Spring 2025

Basic Information (for the first half of this course)

Aim of the course

The first half of this course aims at an introduction to the fundamental concepts of linear algebra, including systems of linear equations, matrices, linear transformations, vector spaces, eigenvalues, and eigenvectors.

Course contents and References

See the booklet uploaded on the Moodle page of the course.

Lectures and Tutorials

Prior to the mid-sem, there will be seven weeks of instruction with 3 hours of lectures and 1 hour of tutorial every week. The slides of the lectures will be uploaded in the moodle page of the course.

Acknowledgement

I shall make use of the lecture notes prepared by Prof. B. V, Limaye of a similar course in the past.

Basic Information (contd..)

Grading Policy

There will be two common quizzes (scheduled on 22 January and 12 February 2025) and one mid-semester exam. Evaluation scheme for the second half of the course will be announced later. The grade will be based on the performance in both the first half and the second half of the course.)

The distribution of marks for the first half will be as follows:

Common Quiz 1	:	10 marks
Common Quiz 2	:	10 marks
Mid-sem exam	:	30 marks

Attendance Policy

Attendance in lectures and tutorials is COMPULSORY. Students who do not meet 80% attendance will be awarded the DX grade.

Basic Questions and Answers

Q. 1 What is Linear Algebra?

Ans. You will find out in this course!

Q. 2 Why should I study it?

Ans. Because it is beautiful and useful! In greater detail,

- Linear Algebra is one of the basic areas in Mathematics, having at least as great an impact as Calculus.
- Provides a vital arena where the interaction of Mathematics and machine computation is seen. Many of the problems studied in Linear Algebra are amenable to systematic and even algorithmic solutions, and this makes them implementable on computers.
- Many geometric topics are studied using Linear Algebra.
- Numerous Applications within and outside Mathematics. For example, Google page rank algorithm is based on notions and results from Linear Algebra.

Notation

- $\mathbb{N} := \{1, 2, 3, \ldots\}$
- $\mathbb{Z} := \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}$
- $\mathbb{R} :=$ the set of all real numbers

For $n \in \mathbb{N}$, let us consider the **Euclidean space**

$$\mathbb{R}^n := \{(x_1, \ldots, x_n) : x_j \in \mathbb{R} \text{ for } j = 1, \ldots, n\}.$$

We let $\mathbf{0} := (0, \ldots, 0)$. Also, for $\boldsymbol{x} := (x_1, \ldots, x_n)$ and $\boldsymbol{y} := (y_1, \ldots, y_n)$ in $\mathbb{R}^n$, and for $\alpha \in \mathbb{R}$, we define

$$\begin{array}{rrcl} \text{(sum)} & \boldsymbol{x} + \boldsymbol{y} & := & (x_1 + y_1, \ldots, x_n + y_n) \in \mathbb{R}^n, \\ \text{(scalar multiple)} & \alpha\boldsymbol{x} & := & (\alpha x_1, \ldots, \alpha x_n) \in \mathbb{R}^n, \\ \text{(scalar product)} & \boldsymbol{x} \cdot \boldsymbol{y} & := & x_1 y_1 + \cdots + x_n y_n \in \mathbb{R}. \end{array}$$

Matrices

Let $m, n \in \mathbb{N}$. An $m \times n$ **matrix A** with real entries is a rectangular array of real numbers arranged in m rows and n columns, written as follows:

$$\mathbf{A} := \begin{bmatrix} a_{11} & \cdots & a_{1k} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{j1} & \cdots & \boxed{a_{jk}} & \cdots & a_{jn} \\ \vdots & & \vdots & & \vdots \\ a_{m1} & \cdots & a_{mk} & \cdots & a_{mn} \end{bmatrix} = [a_{jk}],$$

where $a_{jk} \in \mathbb{R}$ is called the (j, k)th **entry** of **A** for $j = 1, \ldots, m$ and $k = 1, \ldots, n$.

Let $\boxed{\mathbb{R}^{m \times n}}$ denote the set of all $m \times n$ matrices with real entries. If $\mathbf{A} := [a_{jk}]$ and $\mathbf{B} := [b_{jk}]$ are in $\mathbb{R}^{m \times n}$, then we say $\mathbf{A} = \mathbf{B} \iff a_{jk} = b_{jk}$ for all $j = 1, \ldots, m$ and $k = 1, \ldots, n$.

Let $0 \leq r < m$, $0 \leq s < n$. By deleting r rows and s columns from $\mathbf{A}$, we obtain an $(m-r)\times(n-s)$ **submatrix** of $\mathbf{A}$.

An $n \times n$ matrix, that is, an element of $\mathbb{R}^{n\times n}$, is called a **square matrix** of size n.

- A square matrix $\mathbf{A} = [a_{jk}]$ is called **symmetric** if $a_{jk} = a_{kj}$ for all j, k.
- A square matrix $\mathbf{A} = [a_{jk}]$ is called **skew-symmetric** if $a_{jk} = -a_{kj}$ for all j, k.
- A square matrix $\mathbf{A} = [a_{jk}]$ is called a **diagonal matrix** if $a_{jk} = 0$ for all $j \neq k$.
- A diagonal matrix $\mathbf{A} = [a_{jk}]$ is called a **scalar matrix** if all diagonal entries of $\mathbf{A}$ are equal.

Two important scalar matrices are the **identity matrix $\mathbf{I}$** in which all diagonal elements are equal to 1, and the **zero matrix $\mathbf{O}$** in which all diagonal elements are equal to 0.

A square matrix $\mathbf{A} = [a_{jk}]$ is called **upper triangular** if $a_{jk} = 0$ for all $j > k$, and **lower triangular** if $a_{jk} = 0$ for all $j < k$.

Note: A matrix $\mathbf{A}$ is upper triangular as well as lower triangular if and only if $\mathbf{A}$ is a diagonal matrix.

Examples

The matrix $\begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{bmatrix}$ is symmetric, while the matrix

$\begin{bmatrix} 0 & 2 & 3 \\ -2 & 0 & 5 \\ -3 & -5 & 0 \end{bmatrix}$ is skew-symmetric.

Note: Every diagonal entry of a skew-symmetric matrix is 0 since $a_{jj} = -a_{jj} \implies a_{jj} = 0$ for $j = 1, \ldots, n$.

Examples

The matrix $\begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 4 \end{bmatrix}$ is diagonal, while $\begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$ is a scalar matrix.

The matrix $\begin{bmatrix} 2 & 1 & -1 \\ 0 & 3 & 1 \\ 0 & 0 & 4 \end{bmatrix}$ is upper triangular,

while the matrix $\begin{bmatrix} 2 & 0 & 0 \\ 1 & 3 & 0 \\ -1 & 1 & 4 \end{bmatrix}$ is lower triangular.

A **row vector** $\mathbf{a}$ of length n is a matrix with only one row consisting of n real numbers; it is written as follows:

$$\mathbf{a} = \begin{bmatrix} a_1 & \cdots & a_k & \cdots & a_n \end{bmatrix},$$

where $a_k \in \mathbb{R}$ for $k = 1, \ldots, n$. Here $\mathbf{a} \in \mathbb{R}^{1\times n}$.

A **column vector** $\mathbf{b}$ of length n is a matrix with only one column consisting of n real numbers; it is written as follows:

$$\mathbf{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_k \\ \vdots \\ b_n \end{bmatrix}, \text{ where } b_k \in \mathbb{R} \text{ for } k = 1, \ldots, n. \text{ Here } \mathbf{b} \in \mathbb{R}^{n\times 1}.$$

When $n = 1$, we may identify $[\alpha] \in \mathbb{R}^{1\times 1}$ with $\alpha \in \mathbb{R}$.

Operations on Matrices

Let $m, n \in \mathbb{N}$, and let $\mathbf{A} := [a_{jk}]$ and $\mathbf{B} := [b_{jk}]$ be $m \times n$ matrices. Then the $m \times n$ matrix $\mathbf{A} + \mathbf{B} := [a_{jk} + b_{jk}]$ is called the **sum** of $\mathbf{A}$ and $\mathbf{B}$. Also, if $\alpha \in \mathbb{R}$, then the $m \times n$ matrix $\alpha\mathbf{A} := [\alpha a_{jk}]$ is called the **scalar multiple** of $\mathbf{A}$ by α.

These operations follow the usual rules: $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$, $(\mathbf{A} + \mathbf{B}) + \mathbf{C} = \mathbf{A} + (\mathbf{B} + \mathbf{C})$, which we write as $\mathbf{A} + \mathbf{B} + \mathbf{C}$, $\alpha(\mathbf{A} + \mathbf{B}) = \alpha\mathbf{A} + \alpha\mathbf{B}$, $(\alpha + \beta)\mathbf{A} = \alpha\mathbf{A} + \beta\mathbf{A}$ and $\alpha(\beta\mathbf{A}) = (\alpha\beta)\mathbf{A}$, which we write as $\alpha\beta\mathbf{A}$. Also, we write $(-1)\mathbf{A}$ as $-\mathbf{A}$, and $\mathbf{A} + (-\mathbf{B})$ as $\mathbf{A} - \mathbf{B}$.

The **transpose** of an $m \times n$ matrix $\mathbf{A} := [a_{jk}]$ is the $n \times m$ matrix $\mathbf{A}^\mathsf{T} := [a_{kj}]$ (in which the rows and the columns of $\mathbf{A}$ are interchanged).

Clearly, $(\mathbf{A}^\mathsf{T})^\mathsf{T} = \mathbf{A}$, $(\mathbf{A} + \mathbf{B})^\mathsf{T} = \mathbf{A}^\mathsf{T} + \mathbf{B}^\mathsf{T}$ and $(\alpha\mathbf{A})^\mathsf{T} = \alpha\mathbf{A}^\mathsf{T}$.

Note: A square matrix $\mathbf{A}$ is symmetric $\iff \mathbf{A}^\mathsf{T} = \mathbf{A}$.

In particular, the preceding operations can be performed on row vectors, and also on column vectors since they are particular types of matrices.

Notice that the sum $\mathbf{a}_1 + \mathbf{a}_2$ of two row vectors $\mathbf{a_1}$ and $\mathbf{a_2}$ of the same length follows the parallelogram law, and so does the sum $\mathbf{b}_1 + \mathbf{b}_2$ of two column vectors $\mathbf{b}_1$ and $\mathbf{b}_2$ of the same length. (Note: All vectors 'originate' from the zero vector.)

Also, note that the transpose of a row vector is a column vector, and vice versa.

We shall often write a column vector $\mathbf{b} := \begin{bmatrix} b_1 \\ \vdots \\ b_k \\ \vdots \\ b_n \end{bmatrix} \in \mathbb{R}^{n\times 1}$, as $\begin{bmatrix} b_1 & \cdots & b_k & \cdots & b_n \end{bmatrix}^{\mathsf{T}}$ in order to save space.

Let $m, n \in \mathbb{N}$. Let $\alpha_1, \ldots, \alpha_m \in \mathbb{R}$.

If $\mathbf{a}_1, \ldots, \mathbf{a}_m \in \mathbb{R}^{1\times n}$, then

$$\alpha_1\mathbf{a}_1 + \cdots + \alpha_m\mathbf{a}_m \in \mathbb{R}^{1\times n}$$

is called a **(finite) linear combination** of $\mathbf{a}_1, \ldots, \mathbf{a}_m$.

Similarly, if $\mathbf{b}_1, \ldots, \mathbf{b}_m \in \mathbb{R}^{n\times 1}$, then

$$\alpha_1\mathbf{b}_1 + \cdots + \alpha_m\mathbf{b}_m \in \mathbb{R}^{n\times 1}$$

is a **(finite) linear combination** of $\mathbf{b}_1, \ldots, \mathbf{b}_m$.

In particular, for $k = 1, \ldots, n$, consider the column vector $\mathbf{e}_k := \begin{bmatrix} 0 & \cdots & 1 & \cdots & 0 \end{bmatrix}^{\mathsf{T}} \in \mathbb{R}^{n \times 1}$, where the kth entry is 1 and all other entries are 0.

If $\mathbf{b} = \begin{bmatrix} b_1 & \cdots & b_k & \cdots & b_n \end{bmatrix}^{\mathsf{T}}$ is any column vector of length n, then it follows that $\mathbf{b} = b_1\mathbf{e}_1 + \cdots + b_k\mathbf{e}_k + \cdots + b_n\mathbf{e}_n$, which is a linear combination of $\mathbf{e}_1, \ldots, \mathbf{e}_n$. The vectors $\mathbf{e}_1, \ldots, \mathbf{e}_n$ are known as the **basic column vectors** in $\mathbb{R}^{n \times 1}$.

Let $\mathbf{A} := [a_{jk}] \in \mathbb{R}^{m \times n}$.

Then $\mathbf{a}_j := \begin{bmatrix} a_{j1} & \cdots & a_{jn} \end{bmatrix} \in \mathbb{R}^{1 \times n}$ is called the jth **row vector of A** for $j = 1, \ldots, m$, and we write $\mathbf{A} = \begin{bmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_m \end{bmatrix}$.

Also, $\mathbf{c}_k := \begin{bmatrix} a_{1k} & \cdots & a_{mk} \end{bmatrix}^{\mathsf{T}}$ is called the kth **column vector of A** for $k = 1, \ldots, n$, and we write $\mathbf{A} = \begin{bmatrix} \mathbf{c}_1 & \cdots & \mathbf{c}_n \end{bmatrix}$.

Examples

Let $\mathbf{A} := \begin{bmatrix} 2 & 1 & -1 \\ 0 & 3 & 1 \end{bmatrix}$ and $\mathbf{B} := \begin{bmatrix} 1 & 0 & 2 \\ -1 & 4 & 1 \end{bmatrix}$.

Then $\mathbf{A} + \mathbf{B} := \begin{bmatrix} 3 & 1 & 1 \\ -1 & 7 & 2 \end{bmatrix}$ and $5\mathbf{A} = \begin{bmatrix} 10 & 5 & -5 \\ 0 & 15 & 5 \end{bmatrix}$.

The row vectors of $\mathbf{A}$ are $\begin{bmatrix} 2 & 1 & -1 \end{bmatrix}$ and $\begin{bmatrix} 0 & 3 & 1 \end{bmatrix}$.

The column vectors of $\mathbf{A}$ are $\begin{bmatrix} 2 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ 3 \end{bmatrix}$ and $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$.

Also, $\mathbf{A}^{\mathsf{T}} = \begin{bmatrix} 2 & 0 \\ 1 & 3 \\ -1 & 1 \end{bmatrix}$.

Matrix Multiplication

Before we discuss when and how a product **AB** of **A** by **B** may be defined, we define the product of a row vector **a** by a column vector **b** of the same length. Since every matrix can be written in terms of row vectors as well as in terms of column vectors, we shall then consider the product **AB**.

Let $n \in \mathbb{N}$, $\mathbf{a} := \begin{bmatrix} a_1 & \cdots & a_n \end{bmatrix} \in \mathbb{R}^{1\times n}$ and $\mathbf{b} := \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} \in \mathbb{R}^{n\times 1}$. Define the **product** of a row vector **a** with a column vector **b** as follows:

$$\mathbf{ab} = \begin{bmatrix} a_1 & \cdots & a_n \end{bmatrix} \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} := a_1 b_1 + \cdots + a_n b_n \in \mathbb{R}.$$

(This is like the scalar product of two vectors in $\mathbb{R}^n$.)

Let $m \in \mathbb{N}$ and $\mathbf{A} \in \mathbb{R}^{m\times n}$. Then $\mathbf{A} = \begin{bmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_m \end{bmatrix}$, where $\mathbf{a}_1, \dots, \mathbf{a}_m \in \mathbb{R}^{1\times n}$. Recalling that $\mathbf{b} \in \mathbb{R}^{n\times 1}$, we define

$$\mathbf{Ab} = \begin{bmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_m \end{bmatrix} \mathbf{b} := \begin{bmatrix} \mathbf{a}_1\mathbf{b} \\ \vdots \\ \mathbf{a}_m\mathbf{b} \end{bmatrix} \in \mathbb{R}^{m\times 1}.$$

Finally, let $p \in \mathbb{N}$ and $\mathbf{B} \in \mathbb{R}^{n\times p}$. Then $\mathbf{B} = [\mathbf{b}_1 \cdots \mathbf{b}_p]$, where $\mathbf{b}_1, \dots, \mathbf{b}_p \in \mathbb{R}^{n\times 1}$. Noting that $\mathbf{Ab}_1, \dots, \mathbf{Ab}_p$ belong to $\mathbb{R}^{m\times 1}$, we define

$$\mathbf{AB} = \mathbf{A}[\mathbf{b}_1 \cdots \mathbf{b}_p] := [\mathbf{Ab}_1 \cdots \mathbf{Ab}_p] \in \mathbb{R}^{m\times p}.$$

Thus

$$\mathbf{AB} = \begin{bmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_m \end{bmatrix} [\mathbf{b}_1 \cdots \mathbf{b}_p] = \begin{bmatrix} \mathbf{a}_1\mathbf{b}_1 & \cdots & \mathbf{a}_1\mathbf{b}_p \\ \vdots & & \vdots \\ \mathbf{a}_m\mathbf{b}_1 & \cdots & \mathbf{a}_m\mathbf{b}_p \end{bmatrix} \in \mathbb{R}^{m\times p}.$$

So if $m, n, p \in \mathbb{N}$, $\mathbf{A} := [a_{jk}] \in \mathbb{R}^{m\times n}$ and $\mathbf{B} := [b_{jk}] \in \mathbb{R}^{n\times p}$, then $\mathbf{AB} \in \mathbb{R}^{m\times p}$, and for $j = 1, \ldots, m$; $k = 1 \ldots, p$,

$$\mathbf{AB} = [c_{jk}], \quad \text{where } c_{jk} := \mathbf{a}_j\mathbf{b}_k = \sum_{\ell=1}^{n} a_{j\ell}b_{\ell k}.$$

Note that the (j, k)th entry of $\mathbf{AB}$ is a product of the jth row vector of $\mathbf{A}$ with the kth column vector of $\mathbf{B}$ as shown below:

$$\begin{bmatrix} & & & & \\ a_{j1} & \cdots & a_{j\ell} & \cdots & a_{jn} \\ & & & & \end{bmatrix} \begin{bmatrix} b_{1k} \\ \vdots \\ b_{\ell k} \\ \vdots \\ b_{nk} \end{bmatrix}$$

Clearly, the product $\mathbf{AB}$ is defined only when the number of columns of $\mathbf{A}$ is equal to the number of rows of $\mathbf{B}$.

Note: $\mathbf{AI} = \mathbf{A}, \mathbf{IA} = \mathbf{A}, \mathbf{AO} = \mathbf{O}$ and $\mathbf{OA} = \mathbf{O}$ whenever these products are defined.

Examples

(i) Let $\mathbf{A} := \begin{bmatrix} 2 & 1 & -1 \\ 0 & 3 & 1 \end{bmatrix}_{2\times 3}$ and $\mathbf{B} := \begin{bmatrix} 1 & 6 & 0 & 2 \\ 2 & -1 & 1 & -2 \\ 2 & 0 & -1 & 1 \end{bmatrix}_{3\times 4}$.

Then $\mathbf{AB} = \begin{bmatrix} 2 & 11 & 2 & 1 \\ 8 & -3 & 2 & -5 \end{bmatrix}_{2\times 4}$.

(ii) Both products **AB** and **BA** are defined $\iff$ the number of columns of **A** is equal to the number rows of **B** and the number columns of **B** is equal to the number of rows of **A**, that is, when $\mathbf{A} \in \mathbb{R}^{m\times n}$ and $\mathbf{B} \in \mathbb{R}^{n\times m}$.

In general, $\mathbf{AB} \neq \mathbf{BA}$. For example, if $\mathbf{A} := \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ and

$\mathbf{B} := \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, then $\mathbf{AB} := \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, while $\mathbf{BA} := \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$.

Note that $\mathbf{BA} = \mathbf{O}$, while $\mathbf{AB} = \mathbf{B} \neq \mathbf{O}$. Since $\mathbf{A} \neq \mathbf{I}$,

we see that the so-called cancellation law does not hold.

Important Remark

Let $\mathbf{A} := [a_{jk}] \in \mathbb{R}^{m \times n}$, and let $\mathbf{e}_1, \ldots, \mathbf{e}_n$ be the basic column vectors in $\mathbb{R}^{n \times 1}$. Then for $k = 1, \ldots, n$,

$$\mathbf{A}\,\mathbf{e}_k = \begin{bmatrix} a_{1k} \\ \vdots \\ a_{jk} \\ \vdots \\ a_{mk} \end{bmatrix}, \text{ which is the } k\text{th column of } \mathbf{A}.$$

This follows from our definition of matrix multiplication.

It follows that if $\mathbf{B} \in \mathbb{R}^{m \times n}$, then

$\boxed{\mathbf{A} = \mathbf{B} \text{ if and only if } \mathbf{A}\mathbf{e}_k = \mathbf{B}\mathbf{e}_k \text{ for each } k = 1, \ldots, n.}$

In particular, $\mathbf{A} = \mathbf{I} \iff \mathbf{A}\mathbf{e_k} = \mathbf{e_k}$ for each $k = 1, \ldots, n$.

Let $\mathbf{A}$ be an $m \times n$ matrix and $\mathbf{B}$ be an $n \times p$ matrix. We have defined the matrix product $\mathbf{AB}$ in terms of the m row vectors of $\mathbf{A}$ and the p column vectors of $\mathbf{B}$. Now let us write $\mathbf{A}$ in terms of its n column vectors $\mathbf{c}_1, \ldots, \mathbf{c}_n$, and we write $\mathbf{B}$ in terms of its n row vectors $\mathbf{d}_1, \ldots, \mathbf{d}_n$. Then

$$\mathbf{A} = \begin{bmatrix} \mathbf{c}_1 & \cdots & \mathbf{c}_n \end{bmatrix} \text{ and } \mathbf{B} = \begin{bmatrix} \mathbf{d}_1 \\ \vdots \\ \mathbf{d}_n \end{bmatrix}, \text{ where } \mathbf{c}_\ell = \begin{bmatrix} a_{1\ell} \\ \vdots \\ a_{j\ell} \\ \vdots \\ a_{m\ell} \end{bmatrix},$$

$$\mathbf{d}_\ell = \begin{bmatrix} b_{\ell 1} & \cdots & b_{\ell k} & \cdots & b_{\ell p} \end{bmatrix} \text{ for } \ell = 1, \ldots, n.$$

Hence

$$\sum_{\ell=1}^{n} \mathbf{c}_\ell \mathbf{d}_\ell = \sum_{\ell=1}^{n} \left[a_{j\ell} b_{\ell k} \right] = \left[\sum_{\ell=1}^{n} a_{j\ell} b_{\ell k} \right] = \mathbf{AB}.$$

Thus

$$\mathbf{AB} = \begin{bmatrix} \mathbf{c}_1 & \cdots & \mathbf{c}_n \end{bmatrix} \begin{bmatrix} \mathbf{d}_1 \\ \vdots \\ \mathbf{d}_n \end{bmatrix} = \mathbf{c}_1\mathbf{d}_1 + \cdots + \mathbf{c}_n\mathbf{d}_n \in \mathbb{R}^{m\times p}.$$

In particular, if $p = 1$, that is, if $\mathbf{B} := \begin{bmatrix} d_1 \\ \vdots \\ d_n \end{bmatrix} \in \mathbb{R}^{n\times 1}$ is a column vector, then

$$\mathbf{AB} = \begin{bmatrix} \mathbf{c}_1 & \cdots & \mathbf{c}_n \end{bmatrix} \begin{bmatrix} d_1 \\ \vdots \\ d_n \end{bmatrix} = d_1\mathbf{c}_1 + \cdots + d_n\mathbf{c}_n \in \mathbb{R}^{m\times 1}.$$

Properties of Matrix Multiplication

Consider matrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ and $\alpha \in \mathbb{R}$. Then it is easy to see that $(\mathbf{A}+\mathbf{B})\mathbf{C} = \mathbf{AC}+\mathbf{BC}$, $\mathbf{C}(\mathbf{A}+\mathbf{B}) = \mathbf{CA}+\mathbf{CB}$ and $(\alpha\mathbf{A})\mathbf{B} = \alpha\mathbf{AB} = \mathbf{A}(\alpha\mathbf{B})$, if sums & products are well-defined.

Matrix multiplication also satisfies the **associative law**:

Proposition

Let $m, n, p, q \in \mathbb{N}$. If $\mathbf{A} \in \mathbb{R}^{m\times n}$, $\mathbf{B} \in \mathbb{R}^{n\times p}$ and $\mathbf{C} \in \mathbb{R}^{p\times q}$, then $\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C}$ (which we shall write as $\mathbf{ABC}$).

Proof. Let $\mathbf{A} := [a_{jk}]$, $\mathbf{B} := [b_{jk}]$ and $\mathbf{C} := [c_{jk}]$. Also, let $(\mathbf{AB})\mathbf{C} := [\alpha_{jk}]$ and $\mathbf{A}(\mathbf{BC}) := [\beta_{jk}]$. Then

$$\alpha_{jk} = \sum_{i=1}^{p}\left(\sum_{\ell=1}^{n} a_{j\ell}b_{\ell i}\right)c_{ik} = \sum_{\ell=1}^{n} a_{j\ell}\left(\sum_{i=1}^{p} b_{\ell i}c_{ik}\right) = \beta_{jk}$$

for $j = 1, \ldots, m$ and $k = 1, \ldots, q$. Hence the result. $\square$

Also, the transpose of a product is the product of the transposes in the reverse order:

Proposition

Let $m, n, p \in \mathbb{N}$. If $\mathbf{A} \in \mathbb{R}^{m\times n}$ and $\mathbf{B} \in \mathbb{R}^{n\times p}$, then $(\mathbf{AB})^\mathsf{T} = \mathbf{B}^\mathsf{T}\mathbf{A}^\mathsf{T}$.

Proof. Let $\mathbf{A} := [a_{jk}]$, $\mathbf{B} := [b_{jk}]$ and $\mathbf{AB} := [c_{jk}]$.

Also, let $\mathbf{A}^\mathsf{T} := [a'_{jk}]$, $\mathbf{B}^\mathsf{T} := [b'_{jk}]$ and $(\mathbf{AB})^\mathsf{T} := [c'_{jk}]$. Then

$$c_{jk} = \sum_{\ell=1}^{n} a_{j\ell} b_{\ell k} \quad \text{and so} \quad c'_{jk} = c_{kj} = \sum_{\ell=1}^{n} a_{k\ell} b_{\ell j}$$

for $j = 1, \ldots, m$; $k = 1, \ldots, p$. Suppose $\mathbf{B}^\mathsf{T}\mathbf{A}^\mathsf{T} := [d_{jk}]$. Then

$$d_{jk} = \sum_{\ell=1}^{n} b'_{j\ell} a'_{\ell k} = \sum_{\ell=1}^{n} b_{\ell j} a_{k\ell} = c'_{jk}$$

for $j = 1, \ldots, m$; $k = 1, \ldots, p$. Hence the result. □

Matrix Multiplication Revisited

Let $m, n, p \in \mathbb{N}$, $\mathbf{A} \in \mathbb{R}^{m \times n}$ and $\mathbf{B} \in \mathbb{R}^{n \times p}$. For $j = 1, \ldots, m$, let $\mathbf{a}_j := \begin{bmatrix} a_{j1} & \cdots & a_{jn} \end{bmatrix}$ be the jth row of $\mathbf{A}$, and for $k = 1, \ldots, p$, let $\mathbf{b}_k := \begin{bmatrix} b_{1k} \\ \vdots \\ b_{nk} \end{bmatrix}$ be the kth column of $\mathbf{B}$. Then $\mathbf{A} = \begin{bmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_m \end{bmatrix}$, in terms of its rows. Also, $\mathbf{B} = \begin{bmatrix} \mathbf{b}_1 & \cdots & \mathbf{b}_p \end{bmatrix}$, in terms of its columns. Note: $\mathbf{AB}$ has m rows and p columns.

Rows of **AB**

Fix $j \in \{1, \ldots, m\}$, and consider the jth row of $\mathbf{AB}$, namely, $\mathbf{a}_j\mathbf{B} = [\mathbf{a}_j\mathbf{b}_1 \cdots \mathbf{a}_j\mathbf{b}_p]$. For $k = 1, \ldots, p$, the kth entry of the jth row of $\mathbf{AB}$ is $\mathbf{a}_j\mathbf{b}_k = a_{j1}b_{1k} + \cdots + a_{jn}b_{nk}$, where $b_{1k}, \ldots, b_{nk}$ are the kth entries of the n row vectors of $\mathbf{B}$.

Thus we see that for $j = 1, \ldots, m$, the jth row of **AB** is a linear combination of the n row vectors of **B** with coefficients $a_{j1}, \ldots, a_{jn}$ provided by the jth row of **A**.

Columns of **AB**

Fix $k \in \{1, \ldots, p\}$, and consider the kth column of **AB**, namely, $\mathbf{A}\mathbf{b}_k = \begin{bmatrix} \mathbf{a}_1\mathbf{b}_k \\ \vdots \\ \mathbf{a}_m\mathbf{b}_k \end{bmatrix}$. For $j = 1, \ldots, m$, the jth entry of the kth column of **AB** is $\mathbf{a}_j\mathbf{b}_k = a_{j1}b_{1k} + \cdots + a_{jn}b_{nk}$, that is, $b_{1k}a_{j1} + \cdots + b_{nk}a_{jn}$, where $a_{j1}, \ldots, a_{jn}$ are the jth entries of the n columns of **A**.

Thus we see that for $k = 1, \ldots, n$, the kth column of **AB** is a linear combination of the n column vectors of **A** with coefficients $b_{1k}, \ldots, b_{nk}$ provided by the kth column of **B**.

These descriptions of the rows and columns of **AB** are useful.

Example

As we have seen,

$$\begin{bmatrix} 2 & 1 & -1 \\ 0 & 3 & 1 \end{bmatrix} \begin{bmatrix} 1 & 6 & 0 & 2 \\ 2 & -1 & 1 & -2 \\ 2 & 0 & -1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 11 & 2 & 1 \\ 8 & -3 & 2 & -5 \end{bmatrix}, \quad \text{where}$$

$$\begin{aligned}
\begin{bmatrix} 2 & 11 & 2 & 1 \end{bmatrix} &= 2\begin{bmatrix} 1 & 6 & 0 & 2 \end{bmatrix} + 1\begin{bmatrix} 2 & -1 & 1 & -2 \end{bmatrix} \\
&\quad -1\begin{bmatrix} 2 & 0 & -1 & 1 \end{bmatrix}, \\
\begin{bmatrix} 8 & -3 & 2 & -5 \end{bmatrix} &= 0\begin{bmatrix} 1 & 6 & 0 & 2 \end{bmatrix} + 3\begin{bmatrix} 2 & -1 & 1 & -2 \end{bmatrix} \\
&\quad +1\begin{bmatrix} 2 & 0 & -1 & 1 \end{bmatrix}, \\
\begin{bmatrix} 2 \\ 8 \end{bmatrix} &= 1\begin{bmatrix} 2 \\ 0 \end{bmatrix} + 2\begin{bmatrix} 1 \\ 3 \end{bmatrix} + 2\begin{bmatrix} -1 \\ 1 \end{bmatrix}, \\
\begin{bmatrix} 11 \\ -3 \end{bmatrix} &= 6\begin{bmatrix} 2 \\ 0 \end{bmatrix} - 1\begin{bmatrix} 1 \\ 3 \end{bmatrix} + 0\begin{bmatrix} -1 \\ 1 \end{bmatrix}, \quad \text{etc.}
\end{aligned}$$

.