Chapter 1

Elements of Linear Algebra

1.1. Matrices

– Matrices. Basic definitions
– Types of Matrices
– Basic Matrix Operations (addition, scalar multiplication, matrix multiplication)
– Transposition of Matrices
– Inverse Matrix (definition)
– Application of Matrices

Matrices are essential in the study of linear algebra. These mathematical objects have wide application in engineering, linear programming, physics, economics, statistics, computer graphics. When one needs to store the data for multi-dimensional phenomena, matrix helps. Their advantage is better tractability. This is really a very important question which every engineer must know.

1.1.1. Basic Definitions

In 1850 the English mathematician *John Joseph Sylvester* introduced into mathematics the word “matrix” to describe arrays. Sylvester was known for giving fantastic names to mathematical objects. Matrix comes from the Latin word for womb, it is formed from *mater* meaning mother. [10] Sylvester used this name because he realized that a matrix could be the place, where something develops or arises. We'll understand a matrix in the sense of ordering.

John Joseph Sylvester

A matrix is a mathematical construction, which is a meaningful ordering of other mathematical constructions.

Definition 1.1.

A Matrix A *of size* $m \times n$ (read as "m *by* n") *is a rectangular array of* mn *real numbers* a_{ij} *with* m *rows and* n *columns*:

$$A = (a_{ij}) = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}.$$

In other words, a matrix *is a two dimensional array of scalars* with rows and columns.

Numbers a_{ij} are called *the elements* (*the entries*) of a matrix A. An element a_{ij} has two subscripts i and j, they indicate the «address» of this element: the first one designates the row of a matrix (i varies from 1 to m), and the second subscript j (j varies from 1 to n) designates the column of a matrix in which this element a_{ij} appears. This notation is essential to distinguish the elements of a matrix.

For example, the matrix

$$\begin{array}{c} \text{5 columns} \\ \downarrow\ \downarrow\ \downarrow\ \downarrow\ \downarrow \end{array}$$

$$A = \begin{pmatrix} 1 & -2 & 1 & 0 & 5 \\ 9 & -8 & 7 & 6 & 4 \end{pmatrix} \begin{array}{l} \leftarrow \\ \leftarrow \end{array} \text{2 rows} \tag{1.1}$$

has size 2×5; its element a_{23} (the element in row 2 and in column 3, we read as "*the element ' a ' two three* ") is 7 ($a_{23} = 7$), and the element a_{14} in the first row and in the fourth column is 0 ($a_{14} = 0$).

The row of a matrix A

$$(a_{i1} \quad a_{i2} \quad \dots \quad a_{in}) = \overleftarrow{a}_i$$

is called *the* i*-th row* of a matrix A.

The column

$$\begin{pmatrix} a_{1j} \\ a_{2j} \\ \cdots \\ a_{mj} \end{pmatrix} = \vec{a}_j$$

is called *the* j*-th column* of a matrix A.

For example, the first row of the matrix A, shown above in (1.1), is

$$\overleftarrow{a}_1 = \; 1 \; -2 \; 1 \; 0 \; 5$$

and the third column is

$$\vec{a}_3 = \begin{pmatrix} 1 \\ 7 \end{pmatrix}.$$

A matrix $\mathrm{A}_{m\times n}$ *can be considered as a row of* n *columns*:

$$\mathrm{A}_{m\times n} = \left(\begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix} \cdots \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix} \right) = \; \vec{a}_1 \;\; \vec{a}_2 \; \ldots \; \vec{a}_n \;\; ,$$

or as a column of m *rows*:

$$\mathrm{A}_{m\times n} = \begin{pmatrix} (a_{11} \;\; a_{12} \;\; \cdots \;\; a_{1n}) \\ (a_{21} \;\; a_{22} \;\; \cdots \;\; a_{2n}) \\ \vdots \\ (a_{m1} \;\; a_{m2} \;\; \cdots \;\; a_{mn}) \end{pmatrix} = \begin{pmatrix} \overleftarrow{a}_1 \\ \overleftarrow{a}_2 \\ \vdots \\ \overleftarrow{a}_m \end{pmatrix}.$$

So, we have

$$\mathrm{A}_{m\times n} = (a_{ij}) = (\vec{a}_j) = (\overleftarrow{a}_i).$$

Remark

A matrix of size 1×1 *that contains only one element* (a_{11}) *is identified with this element; thus, the matrix* (5) *can be thought of as* 5.

1.1.2. Types of Matrices

And now let us discuss some types of matrices.

1. A matrix $O_{m\times n}$, which *all elements* are equal to *zero* , is called *a Zero* (*a Null*) *matrix*, for instance,

$$O_{2\times 1} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad O_{2\times 2} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad O_{3\times 5} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

The Zero matrix O acts the same as does the number zero when doing arithmetic with real numbers.

2. A matrix A is said to be *a Square matrix* when it has *the same number of rows and columns* $(m = n)$; about a $n \times n$ matrix we speak as of *a square matrix of order* n (we denote as A_n):

$$A_n = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n-1,1} & a_{n-1,2} & a_{n-1,3} & \cdots & a_{n-1,n} \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} \end{pmatrix}.$$

The diagonal elements $(i = j)$ from the top left corner to the bottom right corner

$$a_{11}, a_{22}, \ldots, a_{nn}$$

of a square matrix A (A_n) form *the Main* (*Principal*) *Diagonal* of this matrix, and the elements

$$a_{n1}, a_{n-1,2}, \ldots, a_{1n}$$

form *the Off Diagonal* of a matrix A.

The Trace of a square matrix A is the sum of diagonal elements:

$$\mathrm{tr}A = a_{11} + a_{22} + \ldots + a_{nn}.$$

For example, the matrix [6]

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

is a square matrix of the third order (A_3); the elements $1, 5, 9$ form the main diagonal, and the elements $7, 5, 3$ form the off diagonal; $\mathrm{tr}A = 1 + 5 + 9 = 15$.

3. A square matrix, which all elements below (resp. above) the main diagonal are all zero, is called *an Upper* (resp. *a Lower*) *Triangular matrix*:

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix}, \qquad \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}.$$

4. A square matrix, that has only nonzero elements on *the main diagonal* and zeros wherever, is known as *a Diagonal matrix*:

$$\mathrm{D}(a_{11}, a_{22}, \ldots, a_{nn}) = \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix}.$$

When the elements a_{ii}, $i = 1, \ldots, n$, of a diagonal matrix D are all equal, it is called a *scalar matrix*.

For example,

$$\begin{pmatrix} 7 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & 7 \end{pmatrix}$$

is a scalar matrix.

5. A special case of a Diagonal matrix is *the Identity matrix* (it is usually represented by the letter I).

All diagonal elements of the Identity matrix are equal to one:

$$\mathrm{I}_n = \mathrm{D}(1,1,\dots,1) = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}$$

(the subscript n may be left out).

For example,

$$\mathrm{I}_1 = 1 \ , \ \mathrm{I}_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \ \mathrm{I}_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Elements of the Identity matrix can be represented by *the Kronecker Delta symbol*:

$$\mathrm{I} = (\delta_{ij}) = \begin{cases} 1, \text{if } i = j, \\ 0, \text{ otherwise.} \end{cases}$$

6. A matrix of size $1 \times n$ (has only one row)

$$a_{11} \quad a_{12} \quad \dots \quad a_{1n} \quad = \overleftarrow{a}$$

is called *a row matrix*.

An alternative and frequently used name for a row matrix is *a row vector*.

7. A matrix that has only one column (a matrix of size $m \times 1$)

$$\begin{pmatrix} a_{11} \\ a_{21} \\ \cdots \\ a_{m1} \end{pmatrix} = \vec{a}$$

is called *a column matrix* (*a column vector*).

1.1.3. Basic Matrix Operations

We will now introduce the basic arithmetic operations with matrices.

Equality, Addition, Scalar Multiplication of Matrices

Let us look at such operations with matrices.

Consider two matrices:

$$\mathrm{A}_{m\times n} = (a_{ij}),$$

$$\mathrm{B}_{m\times n} = (b_{ij}).$$

Definition 1.2.

Matrices A *and* B *are said to be equal if they posseses the same number of rows and the same number of columns and if all pairs of corresponding elements are equal:*

$$\mathrm{A} = \mathrm{B} \Leftrightarrow a_{ij} = b_{ij}$$

for all possible $i = \overline{1, m}$ *and* $j = \overline{1, n}$.

For example,

$$\begin{pmatrix} 2 & 0 \\ 3 & 5 \end{pmatrix} = \begin{pmatrix} 0+2 & 3-3 \\ 3\cdot 1 & 15:3 \end{pmatrix}.$$

Matrix equality has the following properties:

1) *each matrix is equal to itself*: $\mathrm{A} = \mathrm{A}$ (*reflexivity*);

2) *if* $\mathrm{A} = \mathrm{B}$, *then* $\mathrm{B} = \mathrm{A}$ (*symmetry*);

3) *for matrices of the same size we have*:

if $\mathrm{A} = \mathrm{B}$ *and* $\mathrm{B} = \mathrm{C}$, *then* $\mathrm{A} = \mathrm{C}$ (*transitivity*).

Definition 1.3.

The Sum of two matrices A and B of **! N.B.** *the same size* $(m \times n)$ *is called a* **!** $m \times n$ *matrix* $\mathrm{A}+\mathrm{B}$, *each element of it is equal to the sum of the elements in the corresponding positions of the matrices* A *and* B:

$$\mathrm{A}+\mathrm{B} = (a_{ij} + b_{ij}).$$

For example,

$$\begin{pmatrix} -1 & 3 & 0 \\ 0 & 1 & -2 \end{pmatrix} + \begin{pmatrix} 1 & 0 & 3 \\ -5 & 2 & 4 \end{pmatrix} = \begin{pmatrix} (-1)+1 & 3+0 & 0+3 \\ 0+(-5) & 1+2 & (-2)+4 \end{pmatrix} = \begin{pmatrix} 0 & 3 & 3 \\ -5 & 3 & 2 \end{pmatrix}.$$

Definition 1.4.

If A *is a* $m \times n$ *matrix and* $\alpha \in \mathbb{R}$ *is a scalar, then a Scalar Multiple of* A *by* α *is a* $m \times n$ *matrix* $\alpha\mathrm{A}$ *given by*

$$\alpha\mathrm{A} = (\alpha \cdot a_{ij}).$$

N.B. *Every element of a given matrix* A *is multiplied by a scalar* α.

For example,

$$3 \cdot \begin{pmatrix} 1 & 3 \\ -5 & 0 \\ -1 & 4 \end{pmatrix} = \begin{pmatrix} 3 \cdot 1 & 3 \cdot 3 \\ 3 \cdot (-5) & 3 \cdot 0 \\ 3 \cdot (-1) & 3 \cdot 4 \end{pmatrix} = \begin{pmatrix} 3 & 9 \\ -15 & 0 \\ -3 & 12 \end{pmatrix}.$$

Remark

The Subtraction $A - B$ *of two matrices* A *and* B *of the same size is defined as the sum*

$$A - B = A + (-1) \cdot B.$$

Addition and multiplication by a scalar are called *the linear operations*.

Addition of matrices and scalar multiplication of a matrix by a scalar both have the following properties.

For matrices $A_{m\times n}, B_{m\times n}, C_{m\times n}$ *and scalars* $\alpha, \beta \in \mathbb{R}$ *we have*:

1) $A + B = B + A$;

2) $A + (B + C) = (A + B) + C$;

3) $A + O_{m\times n} = A$

   ($O_{m\times n}$ is *a Zero matrix*);

4) $A + (-A) = O_{m\times n}$

   (($-A = (-1) \cdot A$ is the *Additive Inverse of a matrix* A);

5) $1 \cdot A = A$;

6) $(\alpha + \beta) \cdot A = \alpha \cdot A + \beta \cdot A$;

7) $\alpha \cdot (A + B) = \alpha \cdot A + \alpha \cdot B$;

8) $\alpha \cdot (\beta \cdot A) = (\alpha\beta) \cdot A$.

Matrix Multiplication

Matrix multiplication is a common binary operation we come across in engineering and mathematics. We see it a lot in machine learning algorithms.

Scalar multiplication, as we know, is quite simple. Multiplying two matrices together is a little trickier. As with addition, there are restrictions on what kinds of matrices can be multiplied together.

At first let us discuss the product of a row by a column.

The product of a row

$$\overleftarrow{a} = (a_1, a_2, \ldots, a_n)$$

by *a column*

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{pmatrix}$$

(*row "length" equals column "height"*)

is **N.B.** *a scalar* (i.e., as we know, a 1×1 matrix), that is *the sum of the element-by-element products of a row by a column* (*dot product*):

$$(a_1, a_2, \ldots, a_n) \cdot \begin{pmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{pmatrix} = (a_1 x_1 + a_2 x_2 + \ldots + a_n x_n),$$

$$\overleftarrow{a} \cdot \vec{x} = a_1 x_1 + a_2 x_2 + \ldots + a_n x_n = \sum_{k=1}^{n} a_k x_k$$

(in this case the brackets are omitted).

For example,

$$-1 \;\; 2 \;\; 5 \;\; \cdot \begin{pmatrix} 2 \\ 0 \\ 7 \end{pmatrix} = (-1) \cdot 2 + 2 \cdot 0 + 5 \cdot 7 = 33.$$

Let's discuss now the product of a matrix by a column.

At first consider the example

(*"the length" of each row of given matrix must equal to "the height" of the column*).

We are taking the matrix (4×3!)

$$\begin{pmatrix} -1 & 2 & 5 \\ 1 & -3 & 7 \\ 0 & 4 & 3 \\ 9 & 0 & 6 \end{pmatrix}$$

and the column (!3×1)

$$\begin{pmatrix} 2 \\ 0 \\ 7 \end{pmatrix}.$$

N.B. *Each row of the matrix we multiply by the given column*.

As a result we obtain the column (4×1):

$$\begin{pmatrix} -1 & 2 & 5 \\ 1 & -3 & 7 \\ 0 & 4 & 3 \\ 9 & 0 & 6 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ 0 \\ 7 \end{pmatrix} = \begin{pmatrix} (-1)\cdot 2 + 2\cdot 0 + 5\cdot 7 \\ 1\cdot 2 + (-3)\cdot 0 + 7\cdot 7 \\ 0\cdot 2 + 4\cdot 0 + 3\cdot 7 \\ 9\cdot 2 + 0\cdot 0 + 6\cdot 7 \end{pmatrix} = \begin{pmatrix} 33 \\ 51 \\ 21 \\ 60 \end{pmatrix}. \tag{1.2}$$

In general we have:

$$\mathrm{A}_{m\times n} \cdot \vec{x}_{n\times 1} = \vec{b}_{m\times 1},$$

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}, \tag{1.3}$$

$$b_i = \overleftarrow{a}_i \cdot \vec{x} = a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n = \sum_{k=1}^{n} a_{ik}x_k, \; i = \overline{1,m}.$$

We may write the equality (1.3) in the form

$$\begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} \tag{1.4}$$

(*row picture*),
or in the form

$$x_1 \cdot \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} + x_2 \cdot \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix} + \cdots + x_n \cdot \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} \tag{1.5}$$

(*column picture*).

Thus, the product of the matrix

$$\begin{pmatrix} -1 & 2 & 5 \\ 1 & -3 & 7 \\ 0 & 4 & 3 \\ 9 & 0 & 6 \end{pmatrix}$$

by the column

$$\begin{pmatrix} 2 \\ 0 \\ 7 \end{pmatrix}$$

we may find as

$$\begin{pmatrix} -1 & 2 & 5 \\ 1 & -3 & 7 \\ 0 & 4 & 3 \\ 9 & 0 & 6 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ 0 \\ 7 \end{pmatrix} = 2 \cdot \begin{pmatrix} -1 \\ 1 \\ 0 \\ 9 \end{pmatrix} + 0 \cdot \begin{pmatrix} 2 \\ -3 \\ 4 \\ 0 \end{pmatrix} + 7 \cdot \begin{pmatrix} 5 \\ 7 \\ 3 \\ 6 \end{pmatrix} = \begin{pmatrix} 33 \\ 51 \\ 21 \\ 60 \end{pmatrix}.$$

The left part in (1.5) (x_1 *times the first column* $\vec{a}_1$ *of the matrix* A *plus* x_2 *times the second column* $\vec{a}_2$ *of* A *plus, etc., plus* x_n *times the* n*- th column* $\vec{a}_n$ *of* A) is called *a linear combination of columns* $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n$ *with the coefficients* $x_1, x_2, \ldots, x_n$. And we read the equality (1.5) as: *the column*

$$\vec{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

is the linear combination of the columns

$$\vec{a}_1 = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix}, \vec{a}_2 = \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix}, \ldots, \vec{a}_n = \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix}.$$

And now we are discussing the product of two matrices.

N.B. *For the product* $A \cdot B$ *of two matrices* A *and* B *to be defined, the number of columns of the matrix* A *must be equal to the number of rows of the matrix* B ($\Leftrightarrow$ *the column dimension of the matrix* A *equals the row dimension of the matrix* B). *Such matrices* A *and* B *are called conformable for multiplication.*

That is, for matrices

$$A_{2\times 3} = \begin{pmatrix} 2 & -1 & 5 \\ 0 & 3 & -7 \end{pmatrix} \quad \text{and} \quad B_{2\times 5} = \begin{pmatrix} 1 & 4 & 7 & 10 & 13 \\ 2 & -4 & 6 & -8 & -1 \end{pmatrix}$$

their product cannot be defined $((2\times 3)\cdot(2\times 5)$ since the inside numbers are different).

We may visualize matrix multiplication in a XY-grid for validating the feasibility of multiplication [9].

We arrange both matrices in two of the XY-grid quadrants (first matrix $A_{2\times3}$ (yellow color) in the corner of the third quadrant Q3 and the second matrix $B_{2\times5}$ (green color) in the corner of the first quadrant Q1) (Fig. 1.1).

Let us imagine rays of light originating from each of the four edges of these two matrices. These light rays overlap in two sections: in quadrant Q2 and in quadrant Q4.

Look at Q2 now. While examining the overlapped region in Q2 we can determine if these two matrices can be multiplied.

If the overlapped region in quadrant Q2 *is a square then given matrices can be multiplied.*

In our case we have the 2×3 rectangle (pink color, see Fig. 1.1) in quadrant Q2. It means that we cannot multiply given matrices $A_{2\times3}$ and $B_{2\times5}$.

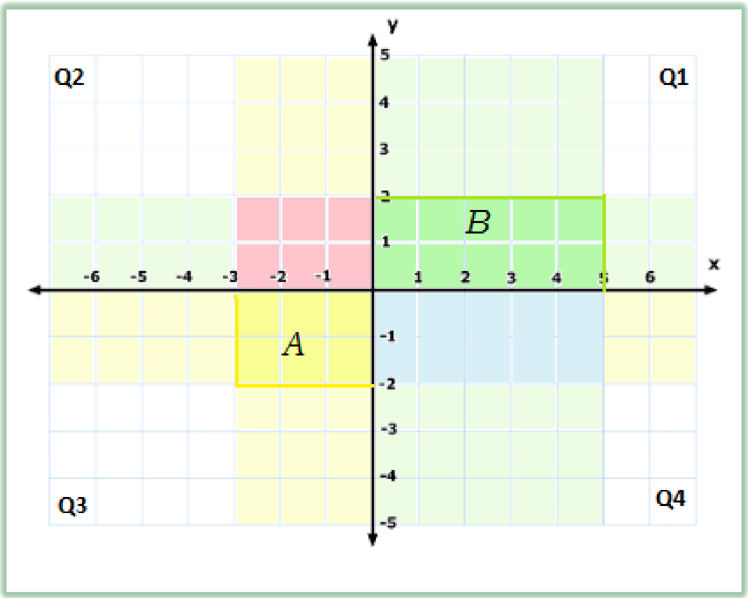

Fig. 1.1

And for matrices $A_{2\times3}$ (yellow color, in quadrant Q3) and $B_{3\times3}$ (green color, in the corner of quadrant Q1) (Fig. 1.2) we obtain a pink 3×3 square in quadrant Q2 (see Fig. 1.2). Thus these matrices are conformable for multiplication.

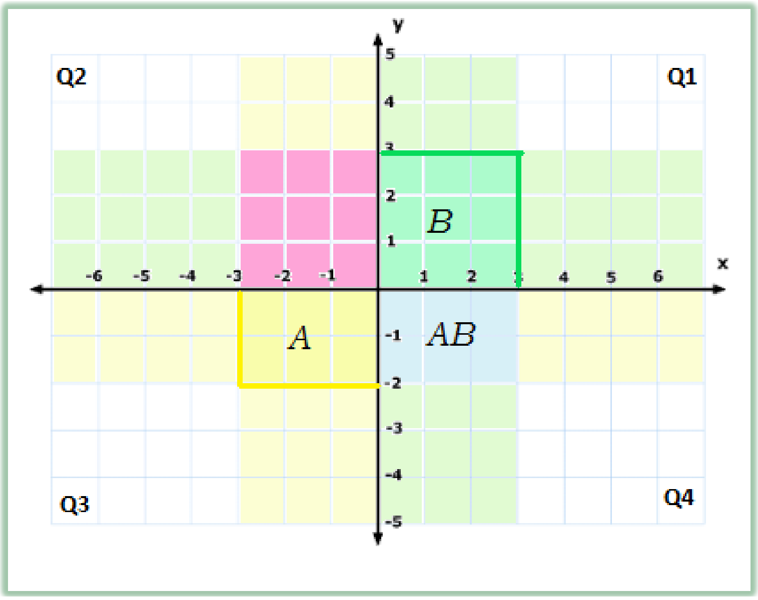

Fig. 1.2

And then we are determining the shape of the product of these matrices. Let's look at quadrant Q4. We see a blue overlapped region. The shape of the overlapped region determines the size (2×3) of the product of the matrices $A_{2\times3}$ and $B_{3\times3}$. Thus, the matrix-product $(AB)_{2\times3}$ has two rows and three columns.

N.B. Rule: *to multiply a* $(m\times k)$ *matrix by a* $(k\times n)$ *matrix, the* (k -s) *must be the same, and the result is a* $(m\times n)$ *matrix*:

$$(m\times k)\cdot(k\times n)\mapsto(m\times n).$$

Definition 1.5.

The product of two conformable for multiplication matrices $\mathrm{A}_{m\times k}$ *and* $\mathrm{B}_{k\times n}$ *is the third matrix* $\mathrm{C}_{m\times n}$:

$$\mathrm{A}_{m\times k}\cdot \mathrm{B}_{k\times n}=\mathrm{C}_{m\times n}=(c_{ij}),$$

each element c_{ij}, $i=\overline{1,m}$, $j=\overline{1,n}$, *is obtained by multiplying component–wise the elements of the* i*-th row of the matrix* A *by the elements of the* j*-th column of the matrix* B (*a row – column multiplication*) (Fig. 1.3):

$$c_{ij}=a_{i1}b_{1j}+a_{i2}b_{2j}+\ldots+a_{ik}b_{kj}=\sum_{s=1}^{k}a_{is}b_{sj}.$$

$$\overleftarrow{a}_i\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1k}\\ \vdots & \vdots & \vdots & \vdots\\ a_{i1} & a_{i2} & \cdots & a_{ik}\\ \vdots & \vdots & \ddots & \vdots\\ a_{m1} & a_{m2} & \cdots & a_{mk}\end{pmatrix}\cdot\overset{\vec{b}_j}{\begin{pmatrix} b_{11} & \cdots & b_{1j} & \cdots & b_{1n}\\ b_{21} & \cdots & b_{2j} & \cdots & b_{2n}\\ \vdots & \ddots & \vdots & \ddots & \vdots\\ b_{k1} & \cdots & b_{kj} & \cdots & b_{kn}\end{pmatrix}}=$$

$$=\overset{j\text{th col.}}{\begin{pmatrix} c_{11} & \cdots & c_{1j} & \cdots & c_{1n}\\ \vdots & \ddots & \vdots & \ddots & \vdots\\ c_{i1} & \cdots & c_{ij} & \cdots & c_{in}\\ \vdots & \ddots & \vdots & \ddots & \vdots\\ c_{m1} & \cdots & c_{mj} & \cdots & c_{mn}\end{pmatrix}}i\text{th row}$$

Fig. 1.3

For example,

$$A_{2\times 3}\cdot B_{3\times 3}=\begin{pmatrix}-1 & 3 & 0\\ 0 & 1 & 7\end{pmatrix}\cdot\begin{pmatrix}1 & 0 & 3\\ -5 & 2 & 0\\ -1 & 4 & -3\end{pmatrix}=$$

$$=\begin{pmatrix}(-1)\cdot 1+3\cdot(-5)+0\cdot(-1) & (-1)\cdot 0+3\cdot 2+0\cdot 4 & (-1)\cdot 3+3\cdot 0+0\cdot(-3)\\ 0\cdot 1+1\cdot(-5)+7\cdot(-1) & 0\cdot 0+1\cdot 2+7\cdot 4 & 0\cdot 3+1\cdot 0+7\cdot(-3)\end{pmatrix}=$$

$$=\begin{pmatrix}-16 & 6 & -3\\ -12 & 30 & -21\end{pmatrix}=C_{2\times 3}$$

(see Fig. 1.4).

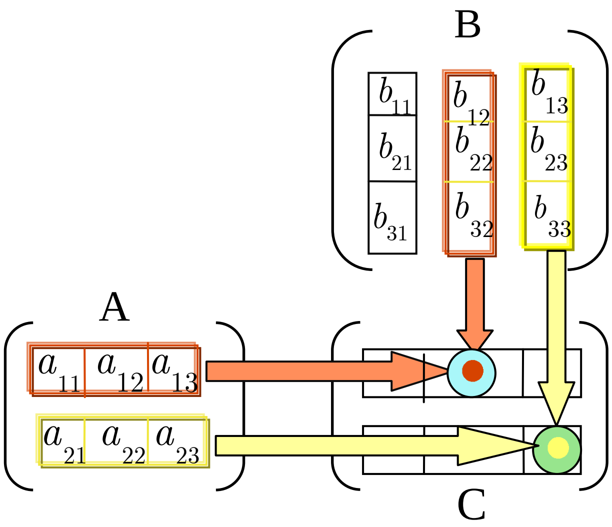

Fig.1.4

Multiplying any matrix A *by a square matrix on either side results in a matrix of the same size as* A, *provided that the sizes of the matrices are such that the multiplication is allowed.*

If a square matrix is the Identity matrix I, *then the result is the original matrix* A:

$$\mathrm{A}_{2\times 3}\cdot \mathrm{I}_3=\begin{pmatrix}-1 & 3 & 0\\ 0 & 1 & 7\end{pmatrix}\cdot\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}=\begin{pmatrix}-1 & 3 & 0\\ 0 & 1 & 7\end{pmatrix}=\mathrm{A},$$

$$\mathrm{I}_2\cdot \mathrm{A}_{2\times 3}=\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}\cdot\begin{pmatrix}-1 & 3 & 0\\ 0 & 1 & 7\end{pmatrix}=\begin{pmatrix}-1 & 3 & 0\\ 0 & 1 & 7\end{pmatrix}=\mathrm{A}.$$

So the Identity matrix for matrices is what the number one is for real numbers.

Remark

N.B. *In arithmetic we are used to*

$$a\cdot b=b\cdot a$$

(*the commutative law of multiplication*). *But this is not generally true for matrices. When we change the order of multiplication* $\mathrm{A}\cdot\mathrm{B}$ *of two matrices, the answer is* (*usually*) *different;*

the product $\mathrm{B}\cdot\mathrm{A}$ *may not even be conformable for multiplication, even though* $\mathrm{A}\cdot\mathrm{B}$ *is perfectly legal.*

Let us see how changing the order affects multiplication:

$$\mathrm{A}\cdot\mathrm{B}=\begin{pmatrix}1 & 3\\ 3 & 9\end{pmatrix}\cdot\begin{pmatrix}3 & -3\\ -1 & 1\end{pmatrix}=\begin{pmatrix}0 & 0\\ 0 & 0\end{pmatrix}$$

(*we have two* ! *non-zero matrices, and the product of these matrices is zero matrix*, for numbers we always have: $a\cdot b=0 \Leftrightarrow a=0$ or $b=0$);

$$\mathrm{B}\cdot\mathrm{A}=\begin{pmatrix}3 & -3\\ -1 & 1\end{pmatrix}\cdot\begin{pmatrix}1 & 3\\ 3 & 9\end{pmatrix}=\begin{pmatrix}-6 & -18\\ 2 & 6\end{pmatrix},$$

the answers are different, so

$$\mathrm{A}\cdot\mathrm{B}\neq\mathrm{B}\cdot\mathrm{A}.$$

While matrix multiplication does not commute, the trace of a product of matrices does not depend on the order of multiplication:

$$\mathrm{tr}(\mathrm{A} \cdot \mathrm{B}) = 0 + 0 = 0 = (-6) + 6 = \mathrm{tr}(\mathrm{B} \cdot \mathrm{A}).$$

Multiplication of matrices (*only for conformable matrices* A, B, C) *has the following properties*:

1) $\mathrm{A} \cdot (\mathrm{B} \cdot \mathrm{C}) = (\mathrm{A} \cdot \mathrm{B}) \cdot \mathrm{C}$;

2) $\mathrm{A} \cdot (\mathrm{B} + \mathrm{C}) = \mathrm{A} \cdot \mathrm{B} + \mathrm{A} \cdot \mathrm{C}$,
   $(\mathrm{A} + \mathrm{B}) \cdot \mathrm{C} = \mathrm{A} \cdot \mathrm{C} + \mathrm{B} \cdot \mathrm{C}$;

3) $\alpha(\mathrm{A} \cdot \mathrm{B}) = (\alpha\mathrm{A}) \cdot \mathrm{B} = \mathrm{A} \cdot (\alpha\mathrm{B})$, $\alpha \in \mathbb{R}$;

4) $\mathrm{A}_{m \times n} \cdot \mathrm{I}_n = \mathrm{I}_m \cdot \mathrm{A}_{m \times n} = \mathrm{A}$;

5) $\mathrm{A}_{m \times k} \cdot \mathrm{O}_{k \times n} = \mathrm{O}_{m \times n}$, $\mathrm{O}_{s \times m} \cdot \mathrm{A}_{m \times n} = \mathrm{O}_{s \times n}$.

The powers n of a **N.B.** *necessarily square matrix* A are defined as

$$\mathrm{A}^n = \underbrace{\mathrm{A} \cdot \mathrm{A} \cdot \ldots \cdot \mathrm{A}}_{n}.$$

It follows that *matrix polynomials* are also valid. As a result, any polynomial equation can be evaluated on a matrix.

If

$$f(x) = a_n x^n + \ldots + a_2 x^2 + a_1 x + a_0,$$

$$\mathrm{A} = \mathrm{A}_n,$$

then

$$f(\mathrm{A}) = a_n \mathrm{A}^n + \ldots + a_2 \mathrm{A}^2 + a_1 \mathrm{A} + a_0 \mathrm{I}$$

N.B. ($\mathrm{I} = \mathrm{I}_n -$ Identity matrix, $\mathrm{A}_n^0 = \mathrm{I}_n$, $a_0 \to a_0 \cdot \mathrm{I}$).

Remark

It is often convenient to partition a matrix into smaller matrices called blocks, like so:

$$\begin{pmatrix} 1 & 3 & 5 & 7 & 1 \\ 2 & 4 & 6 & 8 & 0 \\ 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 & 0 \end{pmatrix} = \left(\begin{array}{cccc|c} 1 & 3 & 5 & 7 & 1 \\ 2 & 4 & 6 & 8 & 0 \\ 1 & 1 & 1 & 1 & 1 \\ \hline 0 & 1 & 2 & 3 & 0 \end{array}\right) = \left(\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right).$$

Here

$$A_{11} = \begin{pmatrix} 1 & 3 & 5 & 7 \\ 2 & 4 & 6 & 8 \\ 1 & 1 & 1 & 1 \end{pmatrix}, \quad A_{12} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix},$$

$$A_{21} = \begin{pmatrix} 0 & 1 & 2 & 3 \end{pmatrix}, \quad A_{22} = (0).$$

The blocks of a block matrix must fit together to form a rectangle.

There are many ways to cut up a matrix into blocks. Often the elements of the matrix will suggest a useful way to divide the matrix into blocks. For example, if there are large blocks of zeros in a matrix, or blocks that look like an Identity matrix, it can be useful to partition the matrix accordingly.

Matrix operations on block matrices can be carried out by treating the blocks as matrix elements.

For example, let us find the product of the matrices

$$\mathrm{A} \cdot \mathrm{B} = \begin{pmatrix} 5 & 2 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 0 & 0 & 8 & 3 \\ 0 & 0 & 5 & 2 \end{pmatrix} \cdot \begin{pmatrix} 1 & -2 & -1 & 0 & 0 \\ -2 & 5 & 2 & 0 & 0 \\ 0 & 0 & 0 & 2 & -3 \\ 0 & 0 & 0 & -5 & 8 \end{pmatrix}.$$

We'll first divide given matrices into blocks:

$$\mathrm{A} = \begin{pmatrix} 5 & 2 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 0 & 0 & 8 & 3 \\ 0 & 0 & 5 & 2 \end{pmatrix} = \left(\begin{array}{cc|cc} 5 & 2 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ \hline 0 & 0 & 8 & 3 \\ 0 & 0 & 5 & 2 \end{array}\right) = \left(\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right),$$

$$\mathrm{B} = \begin{pmatrix} 1 & -2 & -1 & 0 & 0 \\ -2 & 5 & 2 & 0 & 0 \\ 0 & 0 & 0 & 2 & 3 \\ 0 & 0 & 0 & 5 & 8 \end{pmatrix} = \left(\begin{array}{ccc|cc} 1 & -2 & -1 & 0 & 0 \\ -2 & 5 & 2 & 0 & 0 \\ \hline 0 & 0 & 0 & 2 & -3 \\ 0 & 0 & 0 & -5 & 8 \end{array}\right) = \left(\begin{array}{c|c} B_{11} & B_{12} \\ \hline B_{21} & B_{22} \end{array}\right)$$

(we grouped the zero elements into blocks, all blocks are conformable for multiplication matrices).

Consider now blocks as elements of the matrices, we have

$$\mathrm{A} \cdot \mathrm{B} = \left(\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right) \cdot \left(\begin{array}{c|c} B_{11} & B_{12} \\ \hline B_{21} & B_{22} \end{array}\right) = \begin{pmatrix} A_{11} \cdot B_{11} + A_{12} \cdot B_{21} & A_{11} \cdot B_{12} + A_{12} \cdot B_{22} \\ A_{21} \cdot B_{11} + A_{22} \cdot B_{21} & A_{21} \cdot B_{12} + A_{22} \cdot B_{22} \end{pmatrix}.$$

Computing the individual blocks, we get

$$A_{11} \cdot B_{11} + A_{12} \cdot B_{21} = \begin{pmatrix} 5 & 2 \\ 2 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & -2 & -1 \\ -2 & 5 & 2 \end{pmatrix} + \mathrm{O}_{2\times 3} = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix},$$

$$A_{21} \cdot B_{11} + A_{22} \cdot B_{21} = \mathrm{O}_{2\times 3},$$

$$A_{21} \cdot B_{12} + A_{22} \cdot B_{22} = \mathrm{O}_{2\times 2} + \begin{pmatrix} 8 & 3 \\ 5 & 2 \end{pmatrix} \cdot \begin{pmatrix} 2 & -3 \\ -5 & 8 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

$$A_{12} \cdot B_{12} + A_{12} \cdot B_{22} = \mathrm{O}_{2\times 2}.$$

Assembling these pieces into a block matrix gives:

$$\mathrm{A} \cdot \mathrm{B} = \begin{pmatrix} 1 & 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

And this is the product of our matrices.