5. MATRIX CALCULUS

5.1. **Basic operations with matrices.**

Definition. The scheme

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix},$$

where $a_{ij} \in \mathbf{R}$, $i = 1, \dots, m$, $j = 1, \dots, n$, is called a *matrix of type* $m \times n$ (shortly, an m*-by-*n *matrix*). We write $(a_{ij})_{\substack{i=1..m \\ j=1..n}}$. An n-by-n matrix is called *square matrix of order* n. The set of all m-by-n matrices is denoted $M(m \times n)$.

Definition. Let

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

The n-tuple $(a_{i1}, a_{i2}, \dots, a_{in})$, where $i \in \{1, 2, \dots, m\}$, is called i*-th row* of the matrix $\mathbb{A}$.
The m-tuple $\begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix}$, where $j \in \{1, 2, \dots, n\}$, is called j*-th column* of the matrix $\mathbb{A}$.

Definition. We say that two matrices are equal, if they are of the same type and the corresponding elements are equal, i.e., if $\mathbb{A} = (a_{ij})_{\substack{i=1..m \\ j=1..n}}$ and $\mathbb{B} = (b_{uv})_{\substack{u=1..r \\ v=1..s}}$, then $\mathbb{A} = \mathbb{B}$ if and only if $m = r$, $n = s$ and $a_{ij} = b_{ij}$ for every $i \in \{1, \dots, m\}$, $j \in \{1, \dots, n\}$.

Definition. Let $\mathbb{A}, \mathbb{B} \in M(m \times n)$, $\mathbb{A} = (a_{ij})_{\substack{i=1..m \\ j=1..n}}$, $\mathbb{B} = (b_{ij})_{\substack{i=1..m \\ j=1..n}}$, $\lambda \in \mathbf{R}$. The *sum of* $\mathbb{A}$ *and* $\mathbb{B}$ is defined by

$$\mathbb{A} + \mathbb{B} = \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \dots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + b_{22} & \dots & a_{2n} + b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} + b_{m1} & a_{m2} + b_{m1} & \dots & a_{mn} + b_{mn} \end{pmatrix}.$$

Product of a real number λ *and the matrix* $\mathbb{A}$ is defined by

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

Proposition 5.1 (basic properties)**.**

- $\forall \mathbb{A}, \mathbb{B}, \mathbb{C} \in M(m \times n)\colon \mathbb{A} + (\mathbb{B} + \mathbb{C}) = (\mathbb{A} + \mathbb{B}) + \mathbb{C}$, *(associativity)*
- $\forall \mathbb{A}, \mathbb{B} \in M(m \times n)\colon \mathbb{A} + \mathbb{B} = \mathbb{B} + \mathbb{A}$, *(commutativity)*
- $\exists! \mathbb{O} \in M(m \times n)\ \forall \mathbb{A} \in M(m \times n)\colon \mathbb{A} + \mathbb{O} = \mathbb{A}$ *($\mathbb{O}$ is the zero matrix, all it entries are zero),*
- $\forall \mathbb{A} \in M(m \times n)\ \exists \mathbb{C}_{\mathbb{A}} \in M(m \times n)\colon \mathbb{A} + \mathbb{C}_{\mathbb{A}} = \mathbb{O}$ *(the matrix $C_{\mathbb{A}}$ is usually denoted $-\mathbb{A}$ and equals $(-1) \cdot \mathbb{A}$),*
- $\forall \mathbb{A} \in M(m \times n)\ \forall \lambda, \mu \in \mathbf{R}\colon (\lambda\mu)\mathbb{A} = \lambda(\mu\mathbb{A})$,
- $\forall \mathbb{A} \in M(m \times n)\colon 1 \cdot \mathbb{A} = \mathbb{A}$,
- $\forall \mathbb{A} \in M(m \times n)\ \forall \lambda, \mu \in \mathbf{R}\colon (\lambda + \mu)\mathbb{A} = \lambda\mathbb{A} + \mu\mathbb{A}$,
- $\forall \mathbb{A}, \mathbb{B} \in M(m \times n)\ \forall \lambda \in \mathbf{R}\colon \lambda(\mathbb{A} + \mathbb{B}) = \lambda\mathbb{A} + \lambda\mathbb{B}$.

Definition. Let $\mathbb{A} \in M(m \times n)$, $\mathbb{A} = (a_{is})_{\substack{i=1..m \\ s=1..n}}$, $\mathbb{B} \in M(n \times k)$, $\mathbb{B} = (b_{sj})_{\substack{s=1..n \\ j=1..k}}$. Then *the product of matrices* $\mathbb{A}$ and $\mathbb{B}$ is defined as $\mathbb{A}\mathbb{B} \in M(m \times k)$, $\mathbb{A}\mathbb{B} = (c_{ij})_{\substack{i=1..m \\ j=1..k}}$, where

$$c_{ij} = \sum_{s=1}^{n} a_{is} b_{sj}.$$

Remark.

- If $\mathbb{A}$ is a 1-by-n matrix and $\mathbb{B}$ is an n-by-1 matrix, then $\mathbb{A}\mathbb{B}$ is a 1-by-1 matrix. Such a matrix is usually viewed as a number.
- Let $\mathbb{A} \in M(m \times n)$, $\mathbb{B} \in M(n \times k)$, $i \in \{1, \ldots, m\}$ and $j \in \{1, \ldots, k\}$. Then:
  - The entry of $\mathbb{A}\mathbb{B}$ with coordinates ij is equal to the product of the ith row of $\mathbb{A}$ and the jth column of $\mathbb{B}$.
  - The ith row of $\mathbb{A}\mathbb{B}$ is the product of ith row of $\mathbb{A}$ and $\mathbb{B}$.
  - The jth column of $\mathbb{A}\mathbb{B}$ is the product of $\mathbb{A}$ and the jth column of $\mathbb{B}$.

Theorem 5.2 (properties of matrix multiplication)**.** *Let $m, n, k, l \in \mathbf{N}$. Then we have:*

(i) $\forall \mathbb{A} \in M(m \times n)\ \forall \mathbb{B} \in M(n \times k)\ \forall \mathbb{C} \in M(k \times l)\colon \mathbb{A}(\mathbb{B}\mathbb{C}) = (\mathbb{A}\mathbb{B})\mathbb{C}$, *(associativity)*

(ii) $\forall \mathbb{A} \in M(m \times n)\ \forall \mathbb{B}, \mathbb{C} \in M(n \times k)\colon \mathbb{A}(\mathbb{B} + \mathbb{C}) = \mathbb{A}\mathbb{B} + \mathbb{A}\mathbb{C}$, *(left distributivity)*

(iii) $\forall \mathbb{A}, \mathbb{B} \in M(m \times n)\ \forall \mathbb{C} \in M(n \times k)\colon (\mathbb{A} + \mathbb{B})\mathbb{C} = \mathbb{A}\mathbb{C} + \mathbb{B}\mathbb{C}$, *(right distributivity)*

(iv) $\exists! \mathbb{I} \in M(n \times n)\ \forall \mathbb{A} \in M(n \times n)\colon \mathbb{I}\mathbb{A} = \mathbb{A}\mathbb{I} = \mathbb{A}$. *(identity matrix $\mathbb{I}$)*

Remark. Warning! Matrix multiplication is not commutative. Firstly, it may happen that $\mathbb{A}\mathbb{B}$ is defined and $\mathbb{B}\mathbb{A}$ is not defined. Secondly, it may happen that both products $\mathbb{A}\mathbb{B}$ and $\mathbb{B}\mathbb{A}$ do exist but they are of different types. And, finally, even if $\mathbb{A}$ and $\mathbb{B}$ are square matrices of the same order, $\mathbb{A}\mathbb{B}$ may differ from $\mathbb{B}\mathbb{A}$.

Remark. The identity matrix $\mathbb{I} \in M(n \times n)$ has the following entries: The entries on the main diagonal are equal to 1, all the other entries are 0.

Moreover, if $\mathbb{B} \in M(m \times n)$, then $\mathbb{B}\mathbb{I} = \mathbb{B}$; if $\mathbb{C} \in M(n \times k)$, then $\mathbb{I}\mathbb{C} = \mathbb{C}$.

Definition. *Transpose* of a matrix

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

is a matrix defined by

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

i.e., if $\mathbb{A} = (a_{ij})_{\substack{i=1..m \\ j=1..n}}$, then $\mathbb{A}^T = (b_{uv})_{\substack{u=1..n \\ v=1..m}}$, where $b_{uv} = a_{vu}$ for each $u \in \{1, \dots, n\}$, $v \in \{1, 2, \dots, m\}$.

Theorem 5.3 (transpose of a matrix – properties)**.** *We have*

(i) $\forall \mathbb{A} \in M(m \times n)\colon \left(\mathbb{A}^T\right)^T = \mathbb{A}$*,*

(ii) $\forall \mathbb{A}, \mathbb{B} \in M(m \times n)\colon (\mathbb{A} + \mathbb{B})^T = \mathbb{A}^T + \mathbb{B}^T$*,*

(iii) $\forall \mathbb{A} \in M(m \times n)\ \forall \mathbb{B} \in M(n \times k)\colon (\mathbb{A}\mathbb{B})^T = \mathbb{B}^T \mathbb{A}^T$*.*