Week 2

We use two kinds of vectors written either as vector-columns, denoted by lower case letters in bold font (as \( \mathbf{x} \)), or as vector-rows, denoted by lower case letters with arrows above it (as \( \vec{x} \)). The norm, or length, or magnitude, of a vector \( \mathbf{x} \) is a positive number

\[
\|\mathbf{x}\| = (\mathbf{x}, \mathbf{x})^{1/2} = \left( \sum_{k=1}^{n} x_k \bar{x}_k \right)^{1/2} = \left( \sum_{k=1}^{n} |x_k|^2 \right)^{1/2}.
\]

Two vectors \( \mathbf{x} \) and \( \mathbf{y} \) are said to be orthogonal if

\[
\mathbf{x} \cdot \mathbf{y} = 0.
\]

A matrix is a rectangular array of objects or entries, written in rows and columns. The term “matrix” was first mentioned in the mathematical literature in a 1850 paper by James Joseph Sylvester (1814 – 1897). The rectangular array usually enclosed either in square brackets:

\[
\mathbf{A} = [a_{ij}] = \begin{bmatrix}
  a_{11} & a_{12} & \cdots & a_{1n} \\
  a_{21} & a_{22} & \cdots & a_{2n} \\
  \vdots & \vdots & \cdots & \vdots \\
  a_{m1} & a_{m2} & \cdots & a_{mn}
\end{bmatrix}
\]

or in parenthesis

\[
\mathbf{A} = (a_{ij}) = \begin{pmatrix}
  a_{11} & a_{12} & \cdots & a_{1n} \\
  a_{21} & a_{22} & \cdots & a_{2n} \\
  \vdots & \vdots & \cdots & \vdots \\
  a_{m1} & a_{m2} & \cdots & a_{mn}
\end{pmatrix},
\]

which consists of \( m \) rows and \( n \) columns of \( mn \) objects (or entries) chosen from a given set. In this case we speak about \( m \times n - matrix \) (pronounced “m by n”). In the symbol \( a_{ij} \), representing a typical entry (or element), the first subscript \( (i) \) denotes the row and the second subscript \( (j) \) denotes the column occupied by the entry.

The number of rows and the number of columns are called its dimensions of the given matrix. When two matrices have the same dimensions, we say that they are of the same shape or of the same order or of the same size. Unless otherwise indicated, matrices will be denoted by capital letters in bold font.

Two matrices \( \mathbf{A} = [a_{ij}] \) and \( \mathbf{B} = [b_{ij}] \) are equal, written \( \mathbf{A} = \mathbf{B} \), if and only if they have the same dimensions and the corresponding entries are equal. Thus, an equality between two \( m \times n \) matrices \( \mathbf{A} \) and \( \mathbf{B} \) entails equalities between \( mn \) pairs of elements: \( a_{11} = b_{11}, \ a_{12} = b_{12}, \) and so on.
Two matrices of the same size may be added or subtracted. If \( A = [a_{ij}] \) and \( B = [b_{ij}] \) are both \( m \times n \) matrices, then their sum, written \( A + B \), or difference, written \( A - B \), is obtained by adding or subtracting the corresponding entries.

A matrix that is obtained from a \( m \times n \) matrix \( A = [a_{ij}] \) by interchanging rows and columns is called the transpose of \( A \) and usually denoted by \( A^T \) or \( A^t \) or even \( A' \). Thus \( A^T = [a_{ji}] \). A matrix is called symmetric if \( A^T = A \); that is, \( a_{ij} = a_{ji} \).

The complex conjugate of the matrix \( A = [a_{ij}] \), denoted by \( \overline{A} \), is the matrix obtained from \( A \) by replacing each element \( a_{ij} = \alpha + i\beta \) by its conjugate \( \overline{a_{ij}} = \alpha - i\beta \). The adjoint of the \( m \times n \) matrix \( A \) is the transpose of its conjugate matrix and is denoted by \( A^H \), that is, \( A^H = \overline{A}^T \). A matrix is called self-adjoint or Hermitian if \( A^H = A \); that is, \( \overline{a_{ij}} = a_{ji} \).

Note that if \( \lambda \) is a real-valued matrix, then its adjoint is just its transposition because the complex conjugate operation does not change real entries. If \( A \) and \( B \) are two \( n \times n \) matrices and \( \lambda \) is a (complex) scalar, then the following properties hold:

- \( (\overline{A}) = A \), \( (A^T)^T = A \), \( (A^H)^H = A \).
- \( (A + B) = \overline{A} + \overline{B} \), \( (A + B)^T = A^T + B^T \), \( (A + B)^H = A^H + B^H \).
- \( (\lambda A) = \lambda \overline{A} \), \( (\lambda A)^T = \lambda A^T \), \( (\lambda A)^H = \overline{\lambda} A^H \).

The square \( n \times n \) matrix

\[
\begin{bmatrix}
1 & 0 & 0 & \cdots & 0 \\
0 & 1 & 0 & \cdots & 0 \\
\vdots & \vdots & \ddots & \cdots & \vdots \\
0 & 0 & 0 & \cdots & 1
\end{bmatrix}
\]

is denoted by the symbol \( I \) (or \( I_n \) when there is a need to emphasize the dimensions of the matrix) and called the identity matrix (or unit matrix). The matrix with all entries being zero is denoted by \( 0 \) and called the zero matrix.

The product \( AB \) of two matrices is defined whenever the number of columns of the first matrix \( A \) is the same as the number of rows of the second matrix \( B \). If \( A \) is an \( m \times n \) matrix and \( B \) is an \( n \times r \) matrix, then the product \( C = AB \) is an \( m \times r \) matrix whose element \( c_{ij} \) in the \( i \)th row and \( j \)th column is defined as the inner or scalar product of the \( i \)th row of \( A \) and \( j \)th column of \( B \). Namely,

\[
c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj} = \sum_{k=1}^{n} a_{ik}b_{kj}.
\]

It is easy to show that matrix multiplication satisfies the associative law \((AB)C = A(BC)\) and the distributive law \(A(B + C) = AB + AC\), but, generally speaking, it might not satisfy the commutative law, that is, it may happen that \( AB \neq BA \).

Some properties of matrix multiplication differ from the corresponding properties of numbers and we emphasize some of them.

- The multiplication of matrices may not commute even for square matrices. Moreover the product \( AB \) of two matrices \( A \) and \( B \) may exists but their inverse product \( BA \) may not. In general,

\[
AB \neq BA.
\]
• $AB = 0$ does not generally imply $A = 0$ or $B = 0$ or $BA = 0$.

• $AB = AC$ does not generally imply $B = C$.

The identity matrix, $I$, possesses the property that $IA = AI = A$ for any square matrix $A$. However, there may exist a matrix $B \neq I$ such that $BA = AB = A$ for a particular matrix $A$.

The following properties of matrix multiplication hold:

\[
(AB) = A \cdot B, \quad (AB)^T = B^T A^T, \quad (AB)^H = B^H A^H.
\]

A square $n \times n$ matrix $A$ is called an unitary (or isometric) matrix if

\[
||Ax|| = ||x|| \quad \text{for all $n$-vectors $x$}.
\]

A square $n \times n$ matrix $A$ is called normal if $AA^H = A^HA$.

Self-adjoint (Hermitian) matrices and unitary matrices are normal. The trace of an $n \times n$ matrix $A = [a_{ij}]$, denoted by $\text{tr} (A)$ or $\text{tr} A$, is the sum of its diagonal elements, that is, $\text{tr} (A) = a_{11} + a_{22} + \cdots + a_{nn}$.

Whenever $\alpha$ is a number (may be complex or real) and $A$ and $B$ are square matrices of the same dimensions, the following identities hold:

• $\text{tr} (A + B) = \text{tr} (A) + \text{tr} (B)$;

• $\text{tr} (\alpha A) = \alpha \text{tr} (A)$;

• $\text{tr} (AB) = \text{tr} (BA)$.

A matrix $A(t)$ is said to be continuous on an interval $|\alpha, \beta|$ if each element of $A$ is a continuous function on the given interval. With matrix functions we can operate in similar way as with functions. For example, we define the integral of a matrix function as

\[
\int_{\alpha}^{\beta} A(t) \; dt = \left[ \int_{\alpha}^{\beta} a_{ij}(t) \; dt \right].
\]

The derivative of $A(t)$ is defined as

\[
\frac{dA}{dt} = \left[ \frac{da_{ij}(t)}{dt} \right].
\]

Many properties and formulas from Calculus extend to matrix functions; in particular,

\[
\frac{d}{dt} (AB) = A \frac{dB}{dt} + \frac{dA}{dt} B,
\]

\[
\frac{d}{dt} (A + B) = \frac{dA}{dt} + \frac{dB}{dt}, \quad \int (A(t) + B(t)) \; dt = \int A(t) \; dt + \int B(t) \; dt,
\]

\[
\frac{d}{dt} (C A(t)) = C \frac{dA}{dt}, \quad \int (C A(t)) \; dt = C \int A(t) \; dt,
\]

where $C$ is a constant matrix.