Linear algebra

István Gaál and László Kozma

University of Debrecen
Institute of Mathematics

2005

Tartalomjegyzék

1 Determinants

1.1 Defining the determinants

DEFINITION. *Let* $\alpha_{ij}\quad 1 \le i \le m, 1 \le j \le n$ *be real numbers. Then*

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

is a matrix of type $m \times n$. *The diagonal of the matrix is* $a_{11}, a_{22}, a_{33}, \dots$. *The transposed of the matrix* A *is obtained by reflecting the elements to the diagonal:*

$$A^t = \begin{pmatrix} \alpha_{11} & \alpha_{21} & \dots & \alpha_{m1} \\ \alpha_{12} & \alpha_{22} & \dots & \alpha_{m2} \\ \vdots & \vdots & & \vdots \\ \alpha_{1n} & \alpha_{2n} & \dots & \alpha_{mn} \end{pmatrix}$$

NOTATION.

1. The matrix A with entries α_{ij} of type $m \times n$ is also denoted by $A = (\alpha_{ij})_{m\times n}$. The element in the i-th row and j-th column of the matrix A is also denoted by $(A)_{ij}$.

2. The i-th **row** $(1 \le i \le m)$ of the matrix A is

$$A_i = (\alpha_{i1}, \alpha_{i2}, \dots, \alpha_{in})$$

the j-th **column** $(1 \le j \le n)$

$$A^{(j)} = \begin{pmatrix} \alpha_{1j} \\ \alpha_{2j} \\ \vdots \\ \alpha_{mj} \end{pmatrix}.$$

Therefore

$$A = \begin{pmatrix} A_1 \\ A_2 \\ \vdots \\ A_m \end{pmatrix}$$

and

$$A = (A^{(1)}, A^{(2)}, \dots, A^{(n)}).$$

Definition. *An ordering $(i_1, i_2, \ldots, i_n)$ of $(1, 2, \ldots, n)$ is called a* **permutation** *of $(1, 2, \ldots, n)$.*

In the permutation $(i_1, i_2, \ldots, i_k, \ldots, i_\ell, \ldots, i_n)$ the numbers i_k and i_ℓ are in **inversion**, *if $k < \ell$ but $i_k > i_\ell$.*

$I(i_1, i_2, \ldots, i_n)$ denotes the total number of inversions in the permutation $(i_1, i_2, \ldots, i_n)$.

The permutation $(i_1, i_2, \ldots, i_n)$ is **even**, *if $I(i_1, i_2, \ldots, i_n)$ is even, otherwise it is an* **odd permutation**.

The determinant is a number corresponding to a matrix of type $n \times n$.

Definition. *Let $\alpha_{ij}(1 \le i, j \le n)$ be real numbers. The* **determinant** *of the square matrix*

$$A = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \ldots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \ldots & \alpha_{2n} \\ \vdots & \vdots & & \vdots \\ \alpha_{n1} & \alpha_{n2} & \ldots & \alpha_{nn} \end{pmatrix}$$

of type $n \times n$ is

$$D = |A| = \begin{vmatrix} \alpha_{11} & \alpha_{12} & \ldots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \ldots & \alpha_{2n} \\ \vdots & \vdots & & \vdots \\ \alpha_{n1} & \alpha_{n2} & \ldots & \alpha_{nn} \end{vmatrix} =$$

$$= \sum_{(i_1, i_2, \ldots, i_n) \in P_n} (-1)^{I(i_1, i_2, \ldots, i_n)} \alpha_{1i_1} \alpha_{2i_2} \ldots \alpha_{ni_n}$$

where P_n denotes the set of all permutations of $(1, 2, \ldots, n)$. If the matrix A is of type $n \times n$ then D is a **determinant of order** *n. .*

Lemma.**Sarrus rule**

$$\begin{vmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{vmatrix} = \alpha_{11}\alpha_{22} - \alpha_{21}\alpha_{12}$$

and

$$\begin{vmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{vmatrix} =$$

$$= \alpha_{11}\alpha_{22}\alpha_{33} + \alpha_{12}\alpha_{23}\alpha_{31} + \alpha_{13}\alpha_{21}\alpha_{32} - \alpha_{31}\alpha_{22}\alpha_{13} - \alpha_{32}\alpha_{23}\alpha_{11} - \alpha_{33}\alpha_{21}\alpha_{12}$$

Proof. Obvious by the definition of the determinant. □

Remark. Not valid for determinants of higher order.

1.2 Elementary properties of determinants

0. If all elements in a row of the matrix A are zero, then the determinant is zero.

1. The determinant of the transposed matrix is equal to the determinant of the original matrix.

Proof.

$$D = |A| = \begin{vmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ \vdots & \vdots & & \vdots \\ \alpha_{n1} & \alpha_{n2} & \dots & \alpha_{nn} \end{vmatrix} =$$

$$= \sum_{(i_1,i_2,\dots,i_n)\in P_n} (-1)^{I(i_1,i_2,\dots,i_n)} \alpha_{1i_1}\alpha_{2i_2}\dots\alpha_{ni_n}$$

and

$$D' = |A^t| = \begin{vmatrix} \alpha_{11} & \alpha_{21} & \dots & \alpha_{n1} \\ \alpha_{12} & \alpha_{22} & \dots & \alpha_{n2} \\ \vdots & \vdots & & \vdots \\ \alpha_{1n} & \alpha_{2n} & \dots & \alpha_{nn} \end{vmatrix} =$$

$$= \sum_{(j_1,j_2,\dots,j_n)\in P_n} (-1)^{I(j_1,j_2,\dots,j_n)} \alpha_{j_1 1}\alpha_{j_2 2}\dots\alpha_{j_n n}$$

□

2. If all elements in a row are sums of two numbers, then the determinant is the sum of two determinants:

$$\begin{vmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ \vdots & \vdots & & \vdots \\ \beta_{i1}+\gamma_{i1} & \beta_{i2}+\gamma_{i2} & \dots & \beta_{in}+\gamma_{in} \\ \vdots & \vdots & & \vdots \\ \alpha_{n1} & \alpha_{n2} & \dots & \alpha_{nn} \end{vmatrix} =$$

$$= \begin{vmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ \vdots & \vdots & & \vdots \\ \beta_{i1} & \beta_{i2} & \dots & \beta_{in} \\ \vdots & \vdots & & \vdots \\ \alpha_{n1} & \alpha_{n2} & \dots & \alpha_{nn} \end{vmatrix} + \begin{vmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ \vdots & \vdots & & \vdots \\ \gamma_{i1} & \gamma_{i2} & \dots & \gamma_{in} \\ \vdots & \vdots & & \vdots \\ \alpha_{n1} & \alpha_{n2} & \dots & \alpha_{nn} \end{vmatrix}$$

3. If we interchange two rows, then the sign of the determinant changes.

Proof.

$$|A| = \begin{vmatrix} A_1 \\ \vdots \\ A_k \\ \vdots \\ A_l \\ \vdots \\ A_n \end{vmatrix} =$$

$$= \sum_{(i_1,\ldots,i_k,\ldots,i_l\ldots,i_n)\in P_n} (-1)^{I(i_1,\ldots,i_k,\ldots,i_l\ldots,i_n)} \alpha_{1i_1}\ldots\alpha_{ki_k}\ldots\alpha_{li_l}\ldots\alpha_{ni_n},$$

on the other hand

$$|B| = \begin{vmatrix} A_1 \\ \vdots \\ A_l \\ \vdots \\ A_k \\ \vdots \\ A_n \end{vmatrix} =$$

$$= \sum_{(i_1,\ldots,i_l,\ldots,i_k\ldots,i_n)\in P_n} (-1)^{I(i_1,\ldots,i_l,\ldots,i_k\ldots,i_n)} \alpha_{1i_1}\ldots\alpha_{li_l}\ldots\alpha_{ki_k}\ldots\alpha_{ni_n}.$$

The permutations $(i_1,\ldots,i_l,\ldots,i_k,\ldots,i_n)$ and $(i_1,\ldots,i_k,\ldots,i_l,\ldots,i_n)$ are of opposite parity. □

4. If two rows are equal, then the determinant is zero.

5. If a constant can be extracted from all elements of a row, then this constant can be extracted from the determinant:

$$\begin{vmatrix} \alpha_{11} & \alpha_{12} & \ldots & \alpha_{1n} \\ \vdots & \vdots & & \vdots \\ \gamma\cdot\alpha_{i1} & \gamma\cdot\alpha_{i2} & \ldots & \gamma\cdot\alpha_{in} \\ \vdots & \vdots & & \vdots \\ \alpha_{n1} & \alpha_{n2} & \ldots & \alpha_{nn} \end{vmatrix} = \gamma\cdot \begin{vmatrix} \alpha_{11} & \alpha_{12} & \ldots & \alpha_{1n} \\ \vdots & \vdots & & \vdots \\ \alpha_{i1} & \alpha_{i2} & \ldots & \alpha_{in} \\ \vdots & \vdots & & \vdots \\ \alpha_{n1} & \alpha_{n2} & \ldots & \alpha_{nn} \end{vmatrix}$$

Proof.

$$\sum_{(j_1,\ldots,j_i,\ldots,j_n)\in P_n} (-1)^{I(j_1,\ldots,j_i,\ldots,j_n)} \alpha_{1j_1}\ldots(\gamma\cdot\alpha_{ji_j})\ldots,\alpha_{nj_n}$$

$$= \gamma \cdot \sum_{(j_1,\dots,j_i,\dots,j_n)\in P_n} (-1)^{I(j_1,\dots,j_i,\dots,j_n)} \alpha_{1j_1} \dots \alpha_{ji_j} \dots \alpha_{nj_n}$$

□

6. If the elements of a row are equal to the constant multiple of the elements of another row, then the determinant is zero.

$$\begin{vmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ \vdots & \vdots & & \vdots \\ \gamma\alpha_{j1} & \gamma\alpha_{j2} & \dots & \gamma\alpha_{jn} \\ \vdots & \vdots & & \vdots \\ \alpha_{j1} & \alpha_{j2} & \dots & \alpha_{jn} \\ \vdots & \vdots & & \vdots \\ \alpha_{n1} & \alpha_{n2} & \dots & \alpha_{nn} \end{vmatrix} = 0$$

7. The value of the determinant remains the same if we add a constant multiple of the elements of a row to the elements of another row.

$$\begin{vmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ \vdots & \vdots & & \vdots \\ \alpha_{i1}+\gamma\alpha_{j1} & \alpha_{i2}+\gamma\alpha_{j2} & \dots & \alpha_{in}+\gamma\alpha_{jn} \\ \vdots & \vdots & & \vdots \\ \alpha_{j1} & \alpha_{j2} & \dots & \alpha_{jn} \\ \vdots & \vdots & & \vdots \\ \alpha_{n1} & \alpha_{n2} & \dots & \alpha_{nn} \end{vmatrix} = \begin{vmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ \vdots & \vdots & & \vdots \\ \alpha_{i1} & \alpha_{i2} & \dots & \alpha_{in} \\ \vdots & \vdots & & \vdots \\ \alpha_{j1} & \alpha_{j2} & \dots & \alpha_{jn} \\ \vdots & \vdots & & \vdots \\ \alpha_{n1} & \alpha_{n2} & \dots & \alpha_{nn} \end{vmatrix}$$

Proof. By properties 2.,6. the left hand side determinant is

$$\begin{vmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ \vdots & \vdots & & \vdots \\ \alpha_{i1} & \alpha_{i2} & \dots & \alpha_{in} \\ \vdots & \vdots & & \vdots \\ \alpha_{j1} & \alpha_{j2} & \dots & \alpha_{jn} \\ \vdots & \vdots & & \vdots \\ \alpha_{n1} & \alpha_{n2} & \dots & \alpha_{nn} \end{vmatrix} + \gamma \cdot \begin{vmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ \vdots & \vdots & & \vdots \\ \alpha_{j1} & \alpha_{j2} & \dots & \alpha_{jn} \\ \vdots & \vdots & & \vdots \\ \alpha_{j1} & \alpha_{j2} & \dots & \alpha_{jn} \\ \vdots & \vdots & & \vdots \\ \alpha_{n1} & \alpha_{n2} & \dots & \alpha_{nn} \end{vmatrix}$$

and the second determinant is zero. □

8. The properties 0,2–7 are valid for columns, as well.

1.3 Expansion of the determinant

Definition. *Let* $A = (\alpha_{ij})$ *be a of type* $n \times n$ *with real elements* (α_{ij})*. The* **subdeterminant** D_{ij} *belonging to the element* α_{ij} *of* $D = |A|$ *is the determinant of order* $n-1$ *that we obtain from* D *by deleting the* i*-th row and* j*-th column from* D*.*

$$D_{ij} = \begin{vmatrix} \alpha_{11} & \dots & \alpha_{1,j-1} & \alpha_{1,j+1} & \dots & \alpha_{1n} \\ \vdots & & \vdots & \vdots & & \vdots \\ \alpha_{i-1,1} & \dots & \alpha_{i-1,j-1} & \alpha_{i-1,j+1} & \dots & \alpha_{i-1,n} \\ \alpha_{i+1,1} & \dots & \alpha_{i+1,j-1} & \alpha_{i+1,j+1} & \dots & \alpha_{i+1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ \alpha_{n1} & \dots & \alpha_{n,j-1} & \alpha_{n,j+1} & \dots & \alpha_{n,n} \end{vmatrix}$$

The **algebraic subdeterminant** *belonging to the element* α_{ij} *is*

$$A_{ij} = (-1)^{i+j} D_{ij}.$$

Lemma.

Proof.
I.

$$\begin{vmatrix} \alpha_{11} & 0 & \dots & 0 \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2n} \\ \vdots & \vdots & & \vdots \\ \alpha_{n1} & \alpha_{n2} & \dots & \alpha_{nn} \end{vmatrix} = \sum_{(i_1,i_2,\dots,i_n)\in P_n} (-1)^{I(i_1,i_2,\dots,i_n)} \alpha_{1,i_1}\alpha_{2i_2}\dots\alpha_{ni_n}$$

$$= \sum_{(1,i_2,\dots,i_n)\in P_n} (-1)^{I(1,i_2,\dots,i_n)} \alpha_{11}\alpha_{2i_2}\dots\alpha_{ni_n}$$

$$\alpha_{11} \sum_{(i_2,\dots,i_n)\in P_{n-1}} (-1)^{I(i_2,\dots,i_n)} \alpha_{2i_2}\dots\alpha_{ni_n} = \alpha_{11} \begin{vmatrix} \alpha_{22} & \dots & \alpha_{2n} \\ \vdots & & \vdots \\ \alpha_{n2} & \dots & \alpha_{nn} \end{vmatrix}$$

II.

$$\begin{vmatrix} \alpha_{11} & \dots & \alpha_{1,j-1} & \alpha_{1j} & \alpha_{1,j+1} & \dots & \alpha_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \alpha_{i-1,1} & \dots & \alpha_{i-1,j-1} & \alpha_{i-1,j} & \alpha_{i-1,j+1} & \dots & \alpha_{i-1,n} \\ 0 & \dots & 0 & \alpha_{ij} & 0 & \dots & 0 \\ \alpha_{i+1,1} & \dots & \alpha_{i+1,j-1} & \alpha_{i+1,j} & \alpha_{i+1,j+1} & \dots & \alpha_{i+1,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \alpha_{n1} & \dots & \alpha_{n,j-1} & \alpha_{nj} & \alpha_{n,j+1} & \dots & \alpha_{nn} \end{vmatrix}$$

$$(-1)^{j-1+i-1}\begin{vmatrix} \alpha_{ij} & 0 & \dots & 0 & 0 & \dots & 0 \\ \alpha_{1j} & \alpha_{11} & \dots & \alpha_{1,j-1} & \alpha_{1,j+1} & \dots & \alpha_{1n} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ \alpha_{i-1,j} & \alpha_{i-1,1} & \dots & \alpha_{i-1,j-1} & \alpha_{i-1,j+1} & \dots & \alpha_{i-1,n} \\ \alpha_{i+1,j} & \alpha_{i+1,1} & \dots & \alpha_{i+1,j-1} & \alpha_{i+1,j+1} & \dots & \alpha_{i+1,n} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ \alpha_{nj} & \alpha_{n1} & \dots & \alpha_{n,j-1} & \alpha_{n,j+1} & \dots & \alpha_{nn} \end{vmatrix}.$$

□

Theorem. **Expansion theorem**

$$|A| = \sum_{k=1}^{n} \alpha_{ik} A_{ik} = \sum_{k=1}^{n} \alpha_{ki} A_{ki}.$$

Proof.

$$\begin{vmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ \vdots & \vdots & & \vdots \\ \alpha_{i1} & \alpha_{i2} & \dots & \alpha_{in} \\ \vdots & \vdots & & \vdots \\ \alpha_{n1} & \alpha_{n2} & \dots & \alpha_{nn} \end{vmatrix} =$$

$$= \begin{vmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ \vdots & \vdots & & \vdots \\ \alpha_{i1} & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ \alpha_{n1} & \alpha_{n2} & \dots & \alpha_{nn} \end{vmatrix} + \begin{vmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ \vdots & \vdots & & \vdots \\ 0 & \alpha_{i2} & \dots & 0 \\ \vdots & \vdots & & \vdots \\ \alpha_{n1} & \alpha_{n2} & \dots & \alpha_{nn} \end{vmatrix} +$$

$$+ \dots + \begin{vmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & \alpha_{in} \\ \vdots & \vdots & & \vdots \\ \alpha_{n1} & \alpha_{n2} & \dots & \alpha_{nn} \end{vmatrix} = \alpha_{i1}A_{i1} + \alpha_{i2}A_{i2} + \dots + \alpha_{in}A_{in}.$$

□

Theorem. **Skew expansion theorem** *If $i \neq j$, then*

$$\sum_{k=1}^{n} \alpha_{ik} A_{jk} = \sum_{k=1}^{n} \alpha_{ki} A_{kj} = 0.$$

Proof.

$$\begin{vmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ \vdots & \vdots & & \vdots \\ \alpha_{i1} & \alpha_{i2} & \dots & \alpha_{in} \\ \vdots & \vdots & & \vdots \\ \alpha_{i1} & \alpha_{i2} & \dots & \alpha_{in} \\ \vdots & \vdots & & \vdots \\ \alpha_{n1} & \alpha_{n2} & \dots & \alpha_{nn} \end{vmatrix}$$

□

DEFINITION. **Upper triangular matrix**

$$\begin{pmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ 0 & \alpha_{22} & \dots & \alpha_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & \alpha_{nn} \end{pmatrix}$$

Lower triangular matrix

$$\begin{pmatrix} \alpha_{11} & 0 & \dots & 0 \\ \alpha_{21} & \alpha_{22} & \dots & 0 \\ \vdots & \vdots & & \vdots \\ \alpha_{n1} & \alpha_{n2} & \dots & \alpha_{nn} \end{pmatrix}$$

THEOREM.

$$\begin{vmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} & \dots & \alpha_{1n} \\ 0 & \alpha_{22} & \alpha_{23} & \dots & \alpha_{2n} \\ 0 & 0 & \alpha_{33} & \dots & \alpha_{3n} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & \alpha_{nn} \end{vmatrix} = \alpha_{11} \dots \alpha_{nn}$$

Calculating the determinant with elimination process: using the elementary properties of the calculation of the determinants we bring the matrix to an upper triangular form.

2 Matrices

2.1 Operations with matrices

Definition. *The matirices $A = (\alpha_{ij}) \in \mathcal{M}_{m\times n}$ and $B = (\beta_{ij}) \in \mathcal{M}_{m\times n}$ are* **equal**, *if they are of the same type and $\alpha_{ij} = \beta_{ij}, 1 \le i \le m, 1 \le j \le n$.*

Definition. **The sum** *of the matrices $A = (\alpha_{ij}) \in \mathcal{M}_{m\times n}$ and $B = (\beta_{ij}) \in \mathcal{M}_{m\times n}$ of the same type is $C = (\gamma_{ij}) \in \mathcal{M}_{m\times n}$, where*

$$\gamma_{ij} = \alpha_{ij} + \beta_{ij} \quad (1 \le i \le m, 1 \le j \le n).$$

Theorem. *$(\mathcal{M}_{m\times n}, +)$ is a commutative group.*

Definition. *If $A = (\alpha_{ij}) \in \mathcal{M}_{m\times n}$ $\lambda \in \mathbb{R}$ then λA is a matrix of type $m \times n$ such that*

$$(\lambda A)_{ij} = \lambda \alpha_{ij} \quad (1 \le i \le m, 1 \le j \le n).$$

Theorem. *$\mathcal{M}_{m\times n}$ is a vector space over $\mathbb{R}$.*

Definition. *The* **product** *of $A = (\alpha_{ij}) \in \mathcal{M}_{m\times n}$ and $B = (\beta_{ij}) \in \mathcal{M}_{n\times k}$ is $C = (\gamma_{ij}) \in \mathcal{M}_{m\times k}$ with*

$$\gamma_{ij} = \sum_{k=1}^{n} \alpha_{ik}\beta_{kj}.$$

Theorem. *If $A \in \mathcal{M}_{m\times n}, B \in \mathcal{M}_{n\times k}, C \in \mathcal{M}_{k\times l}$, then*

$$A(BC) = (AB)C$$

Proof. Let $A = (\alpha_{ij}), B = (\beta_{ij}), C = (\gamma_{ij})$, then

$$(A(BC))_{ij} = \sum_{g=1}^{n} \alpha_{ig}(BC)_{gj} = \sum_{g=1}^{n} \alpha_{ig} \left(\sum_{h=1}^{k} \beta_{gh}\gamma_{hj} \right)$$

$$= \sum_{g=1}^{n}\sum_{h=1}^{k} \alpha_{ig}\beta_{gh}\gamma_{hj} = \sum_{h=1}^{k}\sum_{g=1}^{n} \alpha_{ig}\beta_{gh}\gamma_{hj}$$

$$= \sum_{h=1}^{k} \left(\sum_{g=1}^{n} \alpha_{ig}\beta_{gh} \right) \gamma_{hj} = \sum_{h=1}^{k} (AB)_{ih}\gamma_{hj} = ((AB)C)_{ij}$$

□

THEOREM. $(\mathcal{M}_{n\times n}, \cdot)$ *is a semigroup.*

THEOREM. $(\mathcal{M}_{n\times n}, +, \cdot)$ *is a non-commutative ring with unit element.*

Proof.

$$(A(B+C))_{ij} = \sum_{k=1}^{n} \alpha_{ik}(\beta_{kj}+\gamma_{kj}) = \sum_{k=1}^{n} \alpha_{ik}\beta_{kj} + \sum_{k=1}^{n} \alpha_{ik}\gamma_{kj} =$$

$$= (AB)_{ij} + (AC)_{ij} = (AB+AC)_{ij}.$$

The **unit matrix** is

$$E = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}$$

□

THEOREM. $\mathcal{M}_{n\times n}$ *is an algebra over* $\mathbb{R}$.

Proof.

$$\lambda(AB) = (\lambda A)B = A(\lambda B),$$

□

THEOREM. **Multiplication theorem of determinants**
If $A, B \in \mathcal{M}_{n\times n}$ *then*

$$|AB| = |A|\ |B|.$$

Proof.

$$|C| = \begin{vmatrix} \alpha_{11} & \dots & \alpha_{1n} & 0 & \dots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ \alpha_{n1} & \dots & \alpha_{nn} & 0 & \dots & 0 \\ -1 & \dots & 0 & \beta_{11} & \dots & \beta_{1n} \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \dots & -1 & \beta_{n1} & \dots & \beta_{nn} \end{vmatrix}.$$

$|C| = |A|\ |B|$.

$$|C_1| = \begin{vmatrix} 0 & \dots & 0 & (AB)_{11} & \dots & (AB)_{1n} \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & (AB)_{n1} & \dots & (AB)_{nn} \\ -1 & \dots & 0 & \beta_{11} & \dots & \beta_{1n} \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \dots & -1 & \beta_{n1} & \dots & \beta_{nn} \end{vmatrix}$$

$$|C_1| = |AB|(-1)^n(-1)^{(n+1)+\ldots+(2n)+1+\ldots+n}$$

□

2.2 The inverse of matrices

DEFINITION. *The matrix $A \in \mathcal{M}_{n\times n}$ is* **invertible** *, if there exist $B \in \mathcal{M}_{n\times n}$ such that*

$$AB = BA = E$$

The matrix B is the **inverse** *of A.*

NOTATION. $B = A^{-1}$.

THEOREM. *The matrix $A \in \mathcal{M}_{n\times n}$ is invertible if and only if it is regular, that is $|A| \neq 0$*

Proof. Assume $AB = E$. Then

$$|A||B| = |AB| = |E| = 1$$

hence $|A| \neq 0$.

Assume $A = (\alpha_{ij})$ is regular. Let $B = (\beta_{ij})$:

$$\beta_{ij} = \frac{A_{ji}}{|A|} \quad (1 \leq i, j \leq n),$$

$$(AB)_{ij} = \sum_{k=1}^{n} \alpha_{ik}\beta_{kj} = \sum_{k=1}^{n} \alpha_{ik}\frac{A_{jk}}{|A|} = \frac{1}{|A|}\sum_{k=1}^{n} \alpha_{ik}A_{jk} = \frac{1}{|A|}(\delta_{ij}|A|),$$

$$(BA)_{ij} = \sum_{k=1}^{n} \beta_{ik}\alpha_{kj} = \sum_{k=1}^{n} \frac{A_{ki}}{|A|}\alpha_{kj} = \frac{1}{|A|}\sum_{k=1}^{n} A_{ki}\alpha_{kj} = \frac{1}{|A|}(\delta_{ij}|A|).$$

□

2.3 Calculation of the inverse matrix with elimination

DEFINITION. **Elementary row operations on a matrix***:*
1. *Multiplication of a row by $\lambda \neq 0$.*
2. *Addition of the λ-multiple of a row to another row.*
3. *Interchanging rows*

THEOREM. *Let $A \in \mathcal{M}_{n\times n}$ and let $E \in \mathcal{M}_{n\times n}$ be the unit matrix. If the matrix $(A|E)$ of type $n \times (2n)$ can be transformed by elementary row operations to the form $(E|B)$, then A is invertible its inverse is B.*

2.4 Further properties of matrix operations

THEOREM. *The regular matrices of type $n \times n$ form a non-commutative group for multiplication*

THEOREM. *If $A, B \in \mathcal{M}_{m\times n}, \lambda \in \mathbb{R}$, then*

$$(A+B)^t = A^t + B^t$$

$$(\lambda A)^t = \lambda A^t.$$

Ha $A \in \mathcal{M}_{m\times n}, B \in \mathcal{M}_{n\times k}$ akkor

$$(AB)^t = B^t A^t.$$

Proof.

$$((AB)^t)_{ij} = (AB)_{ji} = \sum_{l=1}^{n} \alpha_{jl}\beta_{li} = \sum_{l=1}^{n} (B^t)_{il}(A^t)_{lj} = (B^t A^t)_{ij}$$

□

THEOREM. *If $A, B \in \mathcal{M}_{n\times n}$ are regular, then A^t and AB are also regular and*

$$(A^t)^{-1} = (A^{-1})^t,$$

$$(AB)^{-1} = B^{-1}A^{-1}.$$

Proof.

$$(A^{-1})^t A^t = (AA^{-1})^t = E^t = E.$$

$$(B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}A)B = B^{-1}B = E.$$

□

3 Vector spaces

3.1 The concept of a vector space

DEFINITION. *A non empty set V is a* **vector space** *over $\mathbb{R}$ if addition is defined and $(V,+)$ is a commutative group, further for any $\lambda \in \mathbb{R}$ and $a \in V$ there exist a $\lambda a \in V$, such that*

$$\begin{aligned} \lambda(a+b) &= \lambda a + \lambda b \\ (\lambda+\mu)a &= \lambda a + \mu a \\ (\lambda\mu)a &= \lambda(\mu a) = \mu(\lambda a) \\ 1a &= a \end{aligned}$$

for any $\lambda, \mu \in \mathbb{R}$ and any $a, b \in V$.

REMARK.

1.
$$\begin{aligned} \lambda \cdot \underline{0} &= \underline{0} \\ 0 \cdot a &= \underline{0} \\ \lambda \cdot a &= \underline{0} \Longleftrightarrow \lambda = 0 \quad \text{vagy} \quad a = \underline{0} \\ (-a) &= (-1) \cdot a \end{aligned}$$

EXAMPLES.

1. Vectorspace of directed segments in the plain or in the space.
2. n-tuples of real numbers: $\mathbb{R}^n$

$$\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} + \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} = \begin{pmatrix} \alpha_1 + \beta_1 \\ \vdots \\ \alpha_n + \beta_n \end{pmatrix},$$

$$\lambda \cdot \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} = \begin{pmatrix} \lambda\alpha_1 \\ \vdots \\ \lambda\alpha_n \end{pmatrix}.$$

3. Polynomials of degree at most n with real coefficients: $\mathbb{R}_n[x]$.
4. All polynomials with real coefficients: $\mathbb{R}[x]$
5. Matrices of type $m \times n$

3.2 Subspaces

Definition. *A non-empty subset L of the vector space V is a* **subspace** *if it is itself a vectorspace with the operations in V.*

Theorem. **Criterion for subspaces**
Let L be a non-empty subset of the vectorspace V. L is a subspace in V if and only if

$$\forall a, b \in L : \quad a - b \in L \tag{1}$$
$$\forall \lambda \in \mathbb{T}, \forall a \in L : \quad \lambda a \in L \tag{2}$$

Examples.

1. In the vectorspace of directed segments all vectors, the representatives of which, starting from a fixed point O and showing to any point of a straight line through O from a subspace.
2. In $\mathbb{R}^4$ the following sets are subspaces:
$$\{(x_1, 0, x_2, 0) \mid x_1, x_2 \in \mathbb{R}\}$$
$$\{(x_1, x_2, x_3, x_4) \mid x_1, x_2, x_3, x_4 \in R, x_1 - x_2 + 3x_3 = 0\}.$$
3. $\mathbb{R}_3[x]$ is a subspace in $\mathbb{R}_5[x]$ and in $\mathbb{R}[x]$,
4. In $\mathcal{M}_{3\times 2}$ all matrices $A = (a_{ij})$ with $a_{11} = a_{31} = 0$ form a subspace.
5. In $\mathcal{M}_{n\times n}$ the symmetric matrices form a subspace: $(A^t = A)$.

Theorem. *Let V be a vector space, Γ an arbitrary index set. $L_\gamma, (\gamma \in \Gamma)$ subspaces in V. Then*
$$\bigcap_{\gamma \in \Gamma} L_\gamma$$
is a subspace in V.

We use the Criterion for subspaces.

Definition. *Let V be a vector space over $\mathbb{R}$, $a_1, \ldots, a_n \in V$ and $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$. Then $\lambda_1 a_1 + \ldots + \lambda_n a_n \in V$ is the* **linear combination** *of $a_1, \ldots, a_n$ with coefficients $\lambda_1, \ldots, \lambda_n$.*

Definition. *Let V be a vector space, H a non-empty subset in V. The* **subspace generated (spanned) by** *H is the smallest subspace of V containing H.*

Remark.

1. $\mathcal{L}(H)$ is also called the linear closure of H.
2. $\mathcal{L}(H)$ is the intersection of all subspaces containing H.

THEOREM. *$\mathcal{L}(H)$ is the set of all linear combinations composed of the vectors of H:*

$$\mathcal{L}(H) = \{\lambda_1 a_1 + \ldots + \lambda_n a_n \mid \lambda_i \in \mathbb{T}, a_i \in H, 1 \le i \le n, n = 1, 2, \ldots\}.$$

DEFINITION. *The subset H of the vectorspace V is a* **generating system** *of V if*

$$\mathcal{L}(H) = V.$$

DEFINITION. *A vector space is finitely generated if it admits a finite generating set.*

EXAMPLES.

1. In $\mathbb{T}^3$ $\{(1,1,0)^t, (1-2,1)^t, (1,3,-1)^t, (3,1,7)^t\}$ is a generating set.
2. In $\mathbb{T}_2[x]$ $\{x^2+x, x^2-2x+1, x^2+3x-1, 3x^2+x+7\}$ is a generating set.
3. $\mathbb{T}[x]$ is not finitely generated.

3.3 Linear dependence, independence, basis, dimension

DEFINITION. *The vectors $a_1, \ldots, a_n$ are* **linearly independent** *if a linear combination*

$$\lambda_1 a_1 + \ldots + \lambda_n a_n = 0$$

can only be zero if

$$\lambda_1 = \ldots = \lambda_n = 0$$

In the opposite case $a_1, \ldots, a_n$ are **linearly dependent**.

An infinite set of vectors is linearly independent if any finite subset of it is linearly independent.

REMARK. A vector set containing the zero vector is always linearly dependent.

DEFINITION. *A linearly independent generating set is called a* **basis**.

REMARK. Any vectorspace admits a basis.

DEFINITION. *A set of vectors H is* **maximal linearly independent**, *if*
1. H is linearly independent
2. for any $a \in V$ the set $\{a\} \cup H$ is linearly dependent.

Theorem. *The vectorset $(a_1, \dots, a_n)$ is a basis if and only if it is maximal linearly independent.*

Theorem. *In a finitely generated vector space all bases are of the same cardinality (consist of the same number of vectors.)*

Proof. Indirect proof:
$(a_1, \dots, a_m)$ are $(b_1, \dots, b_n)$ bases with $m < n$.

$$b_1 = \lambda_1 a_1 + \ldots + \lambda_m a_m. \tag{3}$$

If $\lambda_m \neq 0$ then

$$a_m = \frac{1}{\lambda_m} b_1 - \frac{\lambda_1}{\lambda_m} a_1 \ldots - \frac{\lambda_{m-1}}{\lambda_m} a_{m-1}. \tag{4}$$

We show that $(b_1, a_1, \dots, a_{m-1})$ is a basis.

$$\nu b_1 + \nu_1 a_1 + \ldots + \nu_{m-1} a_{m-1} = 0,$$

$$(\nu\lambda_1 + \nu_1)a_1 + \ldots + (\nu\lambda_{m-1} + \nu_{m-1})a_{m-1} + \nu\lambda_m a_m = 0.$$

then by $\lambda_m \neq 0$ we have $\nu = 0$, therefore $\nu_1 = \ldots = \nu_{m-1} = 0$

$(b_1, a_1, \dots, a_{m-1})$ is obviously a generating system.

Similarly, $(b_1, b_2, a_1, \dots, a_{m-2})$, etc, $(b_1, \dots, b_{m-1}, a_1)$, $(b_1, \dots, b_m)$ are bases. But then $(b_1, \dots, b_m, \dots, b_n)$ is linearly dependent which is a contradiction. □

Definition. *The common cardinality of bases of a vectors space is the* **dimension** *of the vector space.*

Notation. $\dim V$

Remark.

1. If V has dimension n, then any set of n linearly independent vectors is a basis.
2. If V has dimension n, then any linearly independent system of less than n vectors can be extended to a basis.
3. If L is a subspace in V, then $0 \le \dim L \le \dim V$.

Examples.

1. In the plain any two nonlinear verctors form a basis.
2. In $\mathbb{R}^n$ $((1,0,\dots,0)^t, (0,1,\dots,0)^t, (0,0,\dots,1)^t)$ is a basis. We call it the **natural basis**.
3. In $\mathbb{R}^3$ $((1,2,1)^t, (3,-1,2)^t, (1,0,1)^t)$ is also a basis.
4. In $\mathbb{R}_n[x]$ $(1, x, x^2, \dots, x^n)$ is a basis.

Theorem. *If $(a_1, \ldots, a_n)$ is a basis in V, then for any $a \in V$ there uniquely exist $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ with*

$$a = \lambda_1 a_1 + \ldots + \lambda_n a_n.$$

Definition. *These $\lambda_1, \ldots, \lambda_n$ are called the coordinates of $a \in V$ with respect to the basis $(a_1, \ldots, a_n)$.*

Remark. We write the coordinates as column vectors. If a has coordinates $(\lambda_1, \ldots, \lambda_n)^t$, and b has coordinates $(\mu_1, \ldots, \mu_n)^t$, then by

$$a + b = (\lambda_1 + \mu_1)a_1 + \ldots + (\lambda_n + \mu_n)a_n$$

$$\nu a = (\nu\lambda_1)a_1 + \ldots + (\nu\lambda_n)a_n$$

$a+b$ has coordinates $(\lambda_1+\mu_1, \ldots, \lambda_n+\mu_n)^t$, and νa has coordinates $(\nu\lambda_1, \ldots, \nu\lambda_n)^t$.

3.4 Linear mappings of vector spaces

Definition. *Let V_1, V_2 be vector spaces over $\mathbb{R}$. $\varphi : V_1 \to V_2$ is* **linear**, *if it is additive and homogeneous:*

$$\begin{aligned} \forall a, b \in V_1 : \quad \varphi(a+b) = &\quad \varphi(a) + \varphi(b) \\ \forall a \in V_1, \forall \lambda \in \mathbb{T} : \quad \varphi(\lambda a) = &\quad \lambda\varphi(a) \end{aligned}$$

Notation. The set of linear mappings $\varphi : V_1 \to V_2$ will be denoted by $L(V_1, V_2)$.

Theorem. **Basic theorems on linear mappings. 1** *Let $(e_1, \ldots, e_n)$ be a basis in V_1. If $\varphi, \psi : V_1 \to V_2$ are linear mappings and $\varphi(e_i) = \psi(e_i), 1 \leq i \leq n$, then for any $a \in V$ we have $\varphi(a) = \psi(a)$.*

Proof. Let $a = \lambda_1 e_1 + \ldots + \lambda_n e_n$. Then

$$\varphi(a) = \lambda_1\varphi(e_1) + \ldots + \lambda_n\varphi(e_n) = \lambda_1\psi(e_1) + \ldots + \lambda_n\psi(e_n) = \psi(a).$$

□

Theorem. **Basic theorems on linear mappings. 2** *Let $(e_1, \ldots, e_n)$ be a basis in V_1 and let $a_1, \ldots, a_n \in V_2$ be arbitrary vectors in V_2. There uniquely exist a linear mapping $\varphi : V_1 \to V_2$ with $\varphi(e_i) = a_i, 1 \leq i \leq n$.*

Proof. For any $a = \lambda_1 e_1 + \ldots + \lambda_n e_n$ let

$$\varphi(a) = \lambda_1 a_1 + \ldots + \lambda_n a_n.$$

This is a lineaar mapping. Let $a = \sum_{i=1}^{n} \lambda_i e_i, b = \sum_{i=1}^{n} \mu_i e_i, \nu \in \mathbb{T}$. Then

$$\varphi(a+b) = \varphi\left(\sum_{i=1}^{n}(\lambda_i+\mu_i)e_i\right) = \sum_{i=1}^{n}(\lambda_i+\mu_i)a_i = \sum_{i=1}^{n}\lambda_i a_i + \sum_{i=1}^{n}\mu_i a_i = \varphi(a)+\varphi(b),$$

on the other hand

$$\varphi(\nu a) = \varphi\left(\sum_{i=1}^{n}(\nu\lambda_i)e_i\right) = \sum_{i=1}^{n}(\nu\lambda_i)a_i = \nu\sum_{i=1}^{n}\lambda_i a_i = \nu\varphi(a).$$

The uniqueness follows from the preceeding theorem. □

Definition. *Let V_1, V_2 be vector spaces over $\mathbb{R}$. The mapping $\varphi : V_1 \to V_2$ is* **isomorphic** *if it is bijective (that is injective and surjective). The vector spaces V_1, V_2 are called* **isomorphic** *if there exist an isomorphic mapping $\varphi : V_1 \to V_2$*

Notation. $V_1 \cong V_2$

Theorem. *The vector spaces V_1, V_2 are isomorphic if and only if $\dim V_1 = \dim V_2$.*

Proof.

I. Assume V_1 and V_2 are isomorphic, that is there exist an isomorphic mapping $\varphi : V_1 \to V_2$. If $(e_1, \ldots, e_n)$ is a basis in V_1, then $(\varphi(e_1), \ldots, \varphi(e_n))$ is a basis in V_2.

$$\lambda_1\varphi(e_1) + \ldots + \lambda_n\varphi(e_n) = 0$$

$$\varphi(\lambda_1 e_1 + \ldots + \lambda_n e_n) = 0$$

$\lambda_1 e_1 + \ldots + \lambda_n e_n = 0$, whence $\lambda_1 = \ldots = \lambda_n = 0$.

If $b \in V_2$ then there exist $a \in V_1$, with $\varphi(a) = b$. If $a = \lambda_1 e_1 + \ldots + \lambda_n e_n$ then az

$$b = \varphi(a) = \lambda_1\varphi(e_1) + \ldots + \lambda_n\varphi(e_n),$$

hence $(\varphi(e_1), \ldots, \varphi(e_n))$ is a generating system in V_2.

II. Conversely, assume that $\dim V_1 = \dim V_2$. Let $(e_1, \ldots, e_n)$ be a basis in V_1–ben, and $(f_1, \ldots, f_n)$ be a basis in V_2. Then the mapping $\varphi : V_1 \to V_2$ with $\varphi(e_i) = f_i \ \ (i = 1, \ldots, n)$ is isomorphic. □

Theorem. *Let V be a vector space of dimension n with a basis $(e_1, \ldots, e_n)$. The mapping assigning to a vector $a \in V$ the coordinate column vector of a with respect to this basis, is an isomorphic mapping of V onto $\mathbb{R}^n$.*

Consequence. *Any vector space of dimension n is isomorphic to $\mathbb{R}^n$.*

3.5 Transformation of basis and coordinates

Definition. *Let $(e) = (e_1, \ldots, e_n)$ and $(f) = (f_1, \ldots, f_n)$ be bases in V. The* **transition matrix** *$(e) \to (f)$ is a matrix $S = (\alpha_{ij}) \in \mathcal{M}_{n\times n}$ having the coordinates of f_j with respect to the basis (e) in its j-th column $(1 \leq j \leq n)$, that is*

$$f_j = \sum_{k=1}^{n} \alpha_{kj} e_k$$

Notation. $(e) \xrightarrow{S} (f)$

Theorem. *Let $(e) = (e_1, \ldots, e_n)$ and $(f) = (f_1, \ldots, f_n)$ be bases in V. Denote by S the transition matrix $(e) \to (f)$ and by T the transition matrix $(f) \to (e)$. Then S, T are regular and $T = S^{-1}$.*

Proof. $S = (\alpha_{ij}), T = (\beta_{ij})$. For any j $(1 \leq j \leq n)$

$$f_j = \sum_{k=1}^{n} \alpha_{kj} e_k = \sum_{k=1}^{n} \alpha_{kj} \left(\sum_{l=1}^{n} \beta_{lk} f_l \right) = \sum_{k=1}^{n} \sum_{l=1}^{n} \alpha_{kj} \beta_{lk} f_l$$

$$= \sum_{l=1}^{n} \left(\sum_{k=1}^{n} \beta_{lk} \alpha_{kj} \right) f_l == \sum_{l=1}^{n} (TS)_{lj} f_l$$

whence $(TS)_{lj} = \delta_{lj} (1 \leq l, j \leq n)$, $TS = E$. □

Theorem. *Let $(e) = (e_1, \ldots, e_n)$ and $(f) = (f_1, \ldots, f_n)$ be bases in V, let S be the transition matrix $(e) \to (f)$. Let $a = x_1 e_1 + \ldots + x_n e_n = y_1 f_1 + \ldots + y_n f_n$,*

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

Then

$$Y = S^{-1} X.$$

Proof.

$$\sum_{k=1}^{n} x_k e_k = a = \sum_{j=1}^{n} y_j f_j = \sum_{j=1}^{n} y_j \left(\sum_{k=1}^{n} \alpha_{kj} e_k \right) =$$

$$= \sum_{j=1}^{n} \sum_{k=1}^{n} (\alpha_{kj} y_j e_k) = \sum_{k=1}^{n} \left(\sum_{j=1}^{n} \alpha_{kj} y_j \right) e_k$$

whnece

$$\sum_{j=1}^{n} \alpha_{kj} y_j = x_k \quad (1 \leq k \leq n)$$

that is $X = SY$. □

3.6 Rank of a vector system, rank of a matrix

DEFINITION. *The rank of the system of vectors $a_1, \dots, a_n \in V$ is defined as. $\dim \mathcal{L}(a_1, \dots, a_n)$.*

NOTATION. $\rho(a_1, \dots, a_n)$

THEOREM. *The rank of a system of vectors does not change if we*
1. multiply a vector by $\lambda \neq 0$
2. add the λ-multiple of a vector to another vector
3. delete vectors that are linearly depending on the remaining vectors.
4. interchange the order of vectors.

DEFINITION. *The* **rank of a matrix** *is defined as the rank of its system of row vectors.*

NOTATION. $\rho(A)$

DEFINITION. *A* **non-vanishing minor** *is a subdeterminant of non-zero value.*
A **non-vanishing minor of maximal order** *is a non-vanishing minor such that there is no non-vanishing minors of larger order.*

THEOREM. *Let D be a non-vanishing minor of maximal order of the matrix A. Then the rows of A appearing in D form a maximal linearly independent system in the system of row vectors of A.*

Proof. Assume that the maximal non-vanishing minor D is in the upper left corner of the matrix. Denote by r the rank of D.

We show that
I. $A_1, \dots, A_r$ are linearly independent in $\mathbb{T}^n$
II. $A_s = \lambda_1 A_1 + \dots + \lambda_r A_r$, for every $r+1 \leq s \leq m$, with some scalars $\lambda_1, \dots, \lambda_r$ depending on s.

I. If $A_1, \dots, A_r$ were linearly dependent in $\mathbb{T}^n$–ben, then a similar linear connection would hold also for their first r components. This would imply that the rows of D are linearly dependent, which would contradict to $D \neq 0$.

II. Assume $A = (\alpha_{ij})$ and let $r+1 \leq s \leq m$. The equation $A_s = \lambda_1 A_1 + \dots + \lambda_r A_r$ yields for components:

$$\lambda_1 \alpha_{1k} + \dots + \lambda_r \alpha_{rk} = \alpha_{sk} \quad (1 \leq k \leq n). \tag{5}$$

where the scalars $\lambda_1, \dots, \lambda_n$ are the same for all k. We show that there exist such scalars. For this purpose consider following the determinant of order $r+1$:

$$D(s,k) = \begin{vmatrix} \alpha_{11} & \dots & \alpha_{1r} & \alpha_{1k} \\ \vdots & & \vdots & \vdots \\ \alpha_{r1} & \dots & \alpha_{rr} & \alpha_{rk} \\ \alpha_{s1} & \dots & \alpha_{sr} & \alpha_{sk} \end{vmatrix}.$$

This determinant contains the entries of D in its upper left corner. If $1 \leq k \leq r$ then $D(s,k)$ has two equal rows. If $r+1 \leq k \leq n$ then $D(s,k)$ is a monor of A of order $(r+1)$. In both cases we have $D(s,k) = 0$. Expand $D(s,k)$ using its last column:

$$0 = \alpha_{1k}\mu_1 + \ldots + \alpha_{rk}\mu_k + \alpha_{sk}D$$

where the coefficients do not depend on k. This way we obtain the constants of (5), taking $\lambda_i = -\mu_i/D \ \ (1 \leq i \leq r)$. □

Theorem. *The rank of a matrix is equal to the common order of its maximal non-vanishing minors.*

Consequence. *The rank of a matrix is equal to the rank of it transposed.*

Consequence. *The rank of a matrix is equal to the rank of its system of column vectors.*

3.7 Calculation of the rank of the matrix by elimination method

see tutorial

3.8 Sum and direct sum of subspaces

Definition. *The sum of the subspaces L_1, L_2 of V is*

$$L_1 + L_2 = \{l_1 + l_2 \mid l_1 \in L_1, l_2 \in L_2\}$$

Theorem. *$L_1 + L_2$ is a subspace of V.*

Proof. Let $l_1 + l_2, l_1' + l_2' (l_1, l_1' \in L_1; l_2, l_2' \in L_2)$ be arbitrary vectors in $L_1 + L_2$, let $\lambda \in \mathbb{T}$. Then

$$(l_1 + l_2) - (l_1' + l_2') = (l_1 - l_1') + (l_2 - l_2') \in L_1 + L_2,$$

$$\lambda(l_1 + l_2) = (\lambda l_1) + (\lambda l_2) \in L_1 + L_2,$$

□

Theorem. *If L_1, L_2 are subspaces in V, then $L_1 + L_2$ is just the subspace spanned by $L_1 \cup L_2$.*

Proof. Obviously $L_1 \subseteq L_1 + L_2$ and $L_2 \subseteq L_1 + L_2$, hence $L_1 \cup L_2 \subseteq L_1 + L_2$. Further we know that $L_1 + L_2$ is a subspace, therefore it also contains the subspace generated by $L_1 \cup L_2$.

On the other hand, for any $l_1 \in L_1, l_2 \in L_2$ the vector $l_1 + l_2$ is contained in the subspace generated by $L_1 \cup L_2$. This implies that also $L_1 + L_2$ is contained in the subspace generated by $L_1 \cup L_2$. □

REMARK. If $L_1, \dots, L_k$ are subspaces in V, then their sum is the set of sums $l_1 + \dots + l_k$, where $l_i \in L_i, 1 \le i \le k$. $L_1 + \dots + L_k$ is also a subspace.

DEFINITION. *The sum $L_1 + L_2$ of the subspaces L_1, L_2 is a* **direct sum** *if $L_1 \cap L_2 = \{0\}$.*

NOTATION. $L_1 \oplus L_2$

REMARK. If $L_1, \dots, L_k$ are subspaces in V, then their sum is a direct sum, if for any $1 \le i \le k$ we have

$$L_i \bigcap \left(\sum_{\substack{j=1 \\ j \neq i}}^{k} L_j \right) = \{0\}$$

THEOREM. *Let L_1, L_2 be subspaces in V. Then the following properties are equivalent:*
1. $L_1 + L_2$ is a direct sum.
2. Any vector of $L_1 + L_2$ can be uniquely represented in the form $a = l_1 + l_2$ with $l_1 \in L_1, l_2 \in L_2$.
3. $dim\,(L_1 + L_2) = dim\,L_1 + dim\,L_2$.

Proof. We show that 1→2, 2→3, 3→1.

1→2. If the vector $a \in L_1 + L_2$ had two different representation $a = l_1 + l_2 = l_1' + l_2'$ $(l_1, l_1' \in L_1; l_2, l_2' \in L_2)$, then the non-zero vector $l_1 - l_1' = l_2' - l_2$ would lay in $L_1 \cap L_2$. But $L_1 \cap L_2 = \{0\}$, therefore $l_1 = l_1', l_2 = l_2'$, hence the representation is unique.

2→3. Let $(e_1, \dots, e_k)$ be a basis in L_1, $(f_1, \dots, f_l)$ a basis in L_2. We show that $(e_1, \dots, e_k, f_1, \dots, f_l)$ is a basis in $L_1 + L_2$, that is 3 holds. Since any vector of $L_1 + L_2$ is of the form $l_1 + l_2$ $(l_1 \in L_1; l_2 \in L_2)$ hence $(e_1, \dots, e_k, f_1, \dots, f_l)$ is obviously a generating system in $L_1 + L_2$. If these vectors were linearly dependent the we had

$$\lambda_1 e_1 + \dots + \lambda_k e_k + \mu_1 f_1 + \dots + \mu_l f_l = 0.$$

This can only hold if non of $l_1 = \lambda_1 e_1 + \dots + \lambda_k e_k \in L_1$ and $l_2 = \mu_1 f_1 + \dots + \mu_l f_l \in L_2$ are non-zero. (Otherwise both were zero and all λ_i, μ_i were also zero.) For these vectore we have $l_1 + l_2 = 0$. Now if any vector a has the representation $a = l_1' + l_2' \in L_1 + L_2$ $(l_1' \in L_1; l_2' \in L_2)$ then $a = (l_1 + l_1') + (l_2 + l_2')$ would be a different representation, contradicting 2.

3→1. Assume that $L_1 \cap L_2$ is fo dimension $k \ge 0$.We show that 3 implies $k = 0$, that is 1 holds. Let $(e_1, \dots, e_k)$ be a basis of $L_1 \cap L_2$. As an extension of this basis we obtain a basis $(e_1, \dots, e_k, f_1, \dots, f_l)$ of L_1 $(l \ge 0)$ and a basis $(e_1, \dots, e_k, g_1, \dots, g_m)$ of L_2 $(m \ge 0)$. We have $\dim L_1 = k + l, \dim L_2 = k + m$.

Obviously $(e_1, \dots, e_k, f_1, \dots, f_l, g_1, \dots, g_m)$ is a generating system of $L_1 + L_2$ therefore $\dim(L_1 + L_2) \leq k + l + m$. By 3 we have for the dimensions

$$2k + l + m = \dim L_1 + \dim L_2 = \dim(L_1 + L_2) \leq k + l + m$$

which can only be satisfied with $k = 0$. □

3.9 Quotient space of a vector space

Definition. *Let V be a vector space, L a subspace in V, $a \in V$. The* **coset** *of a is*

$$a + L = \{a + l \mid l \in L\}$$

The vector a is a representing element of the coset.

Theorem. *Let $a, b \in V$ the cosets $a+L$ and $b+L$ are equal if and only if $a-b \in L$.*

Proof. If $a - b \in L$, then there is an $l \in L$, such that $a - b = l$. If $x \in a + L$ then $x = a + l_1$ $(l_1 \in L)$. Then $x = a + l_1 = b + l + l_1 \in b + L$. On the other hand, assume that $y \in b + L$. Then $y = b + l_2$ $(l_2 \in L)$, $y = b + l_2 = a - l + l_2 \in a + L$. Therefore $a + L = b + L$.

If $a + L = b + L$, then there are $l_1, l_2 \in L$ with $a + l_1 = b + l_2$, which implies that $a - b = l_2 - l_1 \in L$. □

Remark. Equivalence relation: $a \equiv b$ if $a - b \in L$

Theorem. *The cosets of L form a partition of V compatible with the operations.*

Proof.
$a \in a + L$
If $x \in a + L$, $x \in b + L$ then $a + l_1 = x = b + l_2$, hence $a - b = l_2 - l_1 \in L$ which implies $a + L = b + L$
If $a + l_1 \in a + L, b + l_2 \in b + L$ then
$(a + l_1) + (b + l_2) = (a + b) + (l_1 + l_2) \in (a + b) + L$
$\lambda(a + l_1) = \lambda a + \lambda l_1 \in (\lambda a) + L$ □

Theorem. *The cosets $\{a + L | a \in V\}$ form a vectors space with the following operations:*

$$\begin{aligned} (a + L) + (b + L) &= (a + b) + L \\ \lambda(a + L) &= (\lambda a) + L \end{aligned}$$

where $a, b \in V, \lambda \in \mathbb{T}$ are arbitrary.

Remark.

1. The result of the operations is independent from the representing elements of the cosets.

Proof. $0 + L$, $(-a) + L$.

$$(\lambda + \mu)(a + L) = ((\lambda + \mu)a) + L = (\lambda a + \mu a) + L =$$

$$= (\lambda a + L) + (\mu a + L) = \lambda(a + L) + \mu(a + L).$$

□

DEFINITION. *The vectors space of the cosets $\{a + L | a \in V\}$ is called the quotient space of V with respect to L.*

NOTATION. V/L

THEOREM. *The codimension of L in V is*

$$dim\ (V/L) = dim V - dim L.$$

Proof. Let $(e_1, \ldots, e_k)$ be a basis in L. Extend it to a basis $(e_1, \ldots, e_k, e_{k+1}, \ldots, e_n)$ of V. We show that $(e_{k+1} + L, \ldots, e_n + L)$ is a basis in V/L.

If a linear combination of these cosets were zero, then we would have

$$0 + L = \lambda_{k+1}(e_{k+1} + L) + \ldots + \lambda_n(e_n + L) = (\lambda_{k+1}e_{k+1} + \ldots + \lambda_n e_n) + L$$

that is $\lambda_{k+1}e_{k+1} + \ldots + \lambda_n e_n \in L$. Then there would be constants $\lambda_1, \ldots, \lambda_k$ with

$$\lambda_{k+1}e_{k+1} + \ldots + \lambda_n e_n = \lambda_1 e_1 + \ldots + \lambda_k e_k.$$

Using the linear independence of $e_1, \ldots, e_k, e_{k+1}, \ldots, e_n$ this implies that all coefficients are zero, that is $\lambda_{k+1} = \ldots = \lambda_n = 0$, whence $\{e_{k+1} + L, \ldots, e_n + L\}$ are linearly independent in V/L.

Let $a = \lambda_1 e_1 + \ldots + \lambda_k e_k + \lambda_{k+1}e_{k+1} + \ldots + \lambda_n e_n$ be an arbitrary vector in V. Then

$$a + L = (\lambda_{k+1}e_{k+1} + \ldots + \lambda_n e_n) + L = \lambda_{k+1}(e_{k+1} + L) + \ldots + \lambda_n(e_n + L)$$

using $a - (\lambda_{k+1}e_{k+1} + \ldots + \lambda_n e_n) = \lambda_1 e_1 + \ldots + \lambda_k e_k \in L$. Therefore $(e_{k+1} + L, \ldots, e_n + L)$ is a generating system in V/L. □

4 Systems of linear equations

4.1 General properties

Definition. *Let $A = (\alpha_{ij}) \in \mathcal{M}_{m\times n}$ be a matrix, $b = (\beta_1, \ldots, \beta_m)^t \in \mathbb{R}^m$ a column vector. The system of equations*

$$\begin{aligned} \alpha_{11}x_1 + \ldots + \alpha_{1n}x_n &= \beta_1 \\ &\vdots \\ \alpha_{m1}x_1 + \ldots + \alpha_{mn}x_n &= \beta_m \end{aligned}$$

is called a **system of linear equations**. *The matrix A is the* **coefficient matrix** *the vector b is the* **constant vector**

$$(A|b) = \begin{pmatrix} \alpha_{11} & \ldots & \alpha_{1n} & \beta_1 \\ \vdots & & \vdots & \vdots \\ \alpha_{m1} & \ldots & \alpha_{mn} & \beta_m \end{pmatrix}$$

is the **augmented matrix**. *The vectors*

$$a_i = \begin{pmatrix} \alpha_{1i} \\ \vdots \\ \alpha_{mi} \end{pmatrix}$$

are the **columns** *of the coefficient matrix ($1 \le i \le n$). $X = (x_1, \ldots, x_n)^t$ is the* **vector of unknowns**.

$(\gamma_1, \ldots, \gamma_n)^t \in \mathbb{R}^n$ is a particular solution if substituting $x_1 = \gamma_1, \ldots, x_n = \gamma_n$ we have equations in the system.

The system of linear equations is **consistent** *if it has particular solutions,* **inconsistent** *otherwise.*

A consistent system of linear equations is said to be **uniquely determined** *if there is only one solution, otherwise it is called* **indetermined**.

The **general solution** *of the system of linear equations is the set of all particular solutions.*

Definition. *The* **matrix form** *of the system is*

$$AX = b,$$

the **vector form** *is*

$$x_1a_1 + \ldots + x_na_n = b$$

Theorem. **Rank criterion 1, Kronecker–Capelli theorem** *The system of linear equations is consistent if and only if*

$$\rho(A) = \rho(A|b).$$

REMARK.
$$\rho(a_1, \ldots, a_n) = \rho(a_1, \ldots, a_n, b).$$

THEOREM. **Rank criterion 2** *A consistent system of linear equations is uniquely determined if and only if*
$$\rho(A) = n.$$

DEFINITION. *The system of linear equations is* **homogeneous**, *if* $\beta_1 = \ldots = \beta_m = 0$, **inhomogeneous**, *otherwise.*

THEOREM. *The set of all particular solutions of a homogeneous system of linear equations is a subspace in* $\mathbb{R}^n$ *of dimension* $n - \rho(A)$.

REMARK.

1. This is called **solution space**

Proof. Assume $\rho(A) = r$ and $(a_1, \ldots, a_r)$ is a basis in $\mathcal{L}(a_1, \ldots, a_n)$.

For any $\lambda_{r+1}, \ldots, \lambda_n$ there uniquely exist $\lambda_1, \ldots, \lambda_r$ such that $(\lambda_1, \ldots, \lambda_r, \lambda_{r+1}, \ldots, \lambda_n)^t$, is a solution. $\lambda_1 a_1 + \ldots + \lambda_r a_r + \lambda_{r+1} a_{r+1} + \ldots + \lambda_n a_n = 0$.

Assume
$$\begin{aligned}
f_{r+1} &= (\alpha_{r+1,1}, \ldots, \alpha_{r+1,r}, 1, 0, 0, \ldots, 0)^t \\
f_{r+2} &= (\alpha_{r+2,1}, \ldots, \alpha_{r+2,r}, 0, 1, 0, \ldots, 0)^t \\
f_{r+3} &= (\alpha_{r+3,1}, \ldots, \alpha_{r+3,r}, 0, 0, 1, \ldots, 0)^t \\
&\ldots \\
f_n &= (\alpha_{n1}, \ldots, \alpha_{nr}, 0, 0, 0, \ldots, 1)^t
\end{aligned}$$

are solutions. These vectors form a basis in the solution space. □

THEOREM. *The set of particluar solutions of a consistent inhomogeneous system of linear equations* $Ax = b$ *is a coset of the form*
$$c + H = \{c + h \mid h \in H\}$$
where H *is the solution space of* $Ax = 0$ *and* c *is any particular solution of* $Ax = b$.

THEOREM.**Cramer's rule** *Assume that in* $Ax = b$ *we have*
1. a square coefficient matrix, $m = n$,
2. $|A| \neq 0$.

Then the system of equations is consistent, uniquely determined, and its solution is

$$x_k = \frac{\Delta_k}{|A|} \quad (k = 1, \ldots, n)$$

where Δ_k is the determinant of the matrix obtained from A by replacing its k-th column with b:

$$\Delta_k = \begin{vmatrix} \alpha_{11} & \ldots & \alpha_{1,k-1} & \beta_1 & \alpha_{1,k+1} & \ldots & \alpha_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \alpha_{n1} & \ldots & \alpha_{n,k-1} & \beta_n & \alpha_{n,k+1} & \ldots & \alpha_{nn} \end{vmatrix}.$$

Proof. $Ac = b$, $c = A^{-1}b$.

$$\gamma_k = (A^{-1}b)_k = \sum_{i=1}^{n} (A^{-1})_{ki}\beta_i = \sum_{i=1}^{n} \frac{A_{ik}}{|A|}\beta_i = \frac{1}{|A|}\sum_{i=1}^{n} A_{ik}\beta_i = \frac{\Delta_k}{|A|}$$

□

4.2 Gaussian elimination

Equivalent operations (the set of solutions remains the same) on a system of linear equations:
1. Multiply an equation by $\lambda \neq 0$
2. Add the λ-multiple of an equation to another equation
3. Omit an equation which is the linear combination of the remaining equations.
4. Interchange equations
5. Interchange unknowns together with their coefficients.

Definition. *A matrix $A = (\alpha_{ij}) \in \mathcal{M}_{k\times n}$ is of* **trapezoid** *form if*
$\alpha_{ii} \neq 0$ if $k \leq n$ for all i, and
$\alpha_{ij} = 0$ for all $i > j$.

Example.

$$\begin{pmatrix} 12 & 31 & -1 & 4 & 7 \\ 0 & 23 & 4 & -5 & 0 \\ 0 & 0 & 27 & 2 & 2 \end{pmatrix} \quad \begin{pmatrix} 78 & 17 & 2 \\ 0 & 34 & 21 \\ 0 & 0 & 67 \end{pmatrix}$$

Theorem. *If $Ax = b$ is consistent that it can be transformed to trapesoid form by equivalent operations:*

$$\begin{aligned} \alpha_{11}x_1 + \alpha_{12}x_2 + \ldots + \alpha_{1k}x_k + \alpha_{1,k+1}x_{k+1} + \ldots + \alpha_{1n}x_n &= \beta_1 \\ \alpha_{22}x_2 + \ldots + \alpha_{2k}x_k + \alpha_{2,k+1}x_{k+1} + \ldots + \alpha_{2n}x_n &= \beta_2 \\ \ldots & \\ \alpha_{kk}x_k + \alpha_{k,k+1}x_{k+1} + \ldots + \alpha_{kn}x_n &= \beta_k \end{aligned}$$

where $k \leq n, \alpha_{ii} \neq 0, 1 \leq i \leq k$.

For any values $x_{k+1} = \gamma_{k+1}, \ldots, x_n = \gamma_n$ *of the* **free variables** $x_{k+1}, \ldots, x_n$ *there exist uniquely* $\gamma_1, \ldots, \gamma_k$ *such that*
$(\gamma_1, \ldots, \gamma_k, \gamma_{k+1}, \ldots, \gamma_n)^t$ *is a solution. Further all solutions are obtained this way.*

5 Linear mappings and transformations

5.1 Linear mappings on vectors spaces

Definition. *Let V_1, V_2 be vector spaces over $\mathbb{R}$ and let $\varphi : V_1 \to V_2$ be a linear mapping. The* **image** *of φ is*

$$\varphi(V_1) = \{\varphi(a) \mid a \in V_1\}$$

and the **kernel** *of φ is*

$$\mathrm{Ker}\varphi = \{a \in V_1 \mid \varphi(a) = 0\}.$$

Remark. $\varphi(V_1) \subseteq V_2$ and $\mathrm{Ker}\varphi \subseteq V_1$.

Theorem. *The image is a subspace in V_2 and the kernel is a subspace in V_1.*

Theorem. **Theorem on homomorphisms**
Let V_1, V_2 be vector spaces over $\mathbb{R}$ and let $\varphi : V_1 \to V_2$ be a linear mapping. Then

$$V_1/\mathrm{Ker}\varphi \cong \varphi(V_1).$$

Proof. The natural homomorphism:

$$F(a + \mathrm{Ker}\varphi) = \varphi(a)$$

is an isomorphism $V_1/\mathrm{Ker}\varphi \to \varphi(V_1)$.

$$\begin{aligned} F((a + \mathrm{Ker}\varphi) + (b + \mathrm{Ker}\varphi)) &= F((a + b) + \mathrm{Ker}\varphi) = \varphi(a + b) = \varphi(a) + \varphi(b) \\ &= F(a + \mathrm{Ker}\varphi) + F(b + \mathrm{Ker}\varphi), \\ F(\lambda(a + \mathrm{Ker}\varphi)) &= F(\lambda a + \mathrm{Ker}\varphi) = \varphi(\lambda a) = \lambda\varphi(a) = \lambda F(a + \mathrm{Ker}\varphi). \end{aligned}$$

If $F(a + \mathrm{Ker}\varphi) = F(b + \mathrm{Ker}\varphi)$, then $\varphi(a) = \varphi(b)$, $\varphi(a - b) = 0$, hence $a - b \in \mathrm{Ker}\varphi$, but then $a + \mathrm{Ker}\varphi = b + \mathrm{Ker}\varphi$.

Any element of $\varphi(V_1)$ is of type $\varphi(a)$ with $(a \in V_1)$, and $F(a + \mathrm{Ker}\varphi) = \varphi(a)$. □

Remark. φ is surjective if and only if $\varphi(V_1) = V_2$.

Theorem. *$\varphi : V_1 \to V_2$ is injective if and only if* $\mathrm{Ker}\varphi = \{0\}$.

Notation. The rank of φ is

$$\rho(\varphi) = \dim \varphi(V_1).$$

Consequence. *If $\varphi : V_1 \to V_2$ is linear, then*

$$dim\, V_1 = dim\, \mathrm{Ker}\varphi + dim\, \varphi(V_1). \tag{6}$$

5.2 Linear transformations

DEFINITION. *Let V be a vector space, $\varphi : V \to V$ a linear mapping. Then φ is a* **linear transformation** *(or linear operator) on V.*

NOTATION. The set of all linear transformations on V is denoted by τ_V.

EXAMPLES.

1. The identical mapping, the zero mapping and $\varphi(a) = \lambda a$ are linear transformations.
2. The orthogonal projection on a plane or line is a linear transformation.
3. For any matrix A of type $n \times n$ the mapping $\varphi(x) = Ax$ is a linear transformation on $\mathbb{R}^n$.

THEOREM. *A linear transformation $\varphi \in \tau_V$ is injective if and only if it is surjective.*

Proof. φ is injective if and only if $\mathrm{Ker}\varphi = \{0\}$. On the other hand

$$\dim V = \dim \mathrm{Ker}\varphi + \dim \varphi(V)$$

therefore the kernel is of dimension 0 if and only if the image is of dimension $\dim V$, that is $\varphi(V) = V$. □

DEFINITION. *Let $(e) = (e_1, \dots, e_n)$ be a basis of V and let $\varphi \in \tau_V$. The* **matrix of the linear transformation** *φ in the basis (e) is a matrix $A = (\alpha_{ij}) \in \mathcal{M}_{n\times n}$ which contains the coordinates of $\varphi(e_i)$ in the basis (e) in its i-th column ($1 \le i \le n$), that is*

$$\varphi(e_i) = \sum_{j=1}^{n} \alpha_{ji} e_j.$$

THEOREM. *Let V be a vector space with basis $(e) = (e_1, \dots, e_n)$ and let $\varphi \in \tau_V$, with matrix A in the basis (e). Let $a = x_1e_1 + \ldots + x_ne_n \in V$, $\varphi(a) = y_1e_1 + \ldots + y_ne_n$, and*

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

Then

$$Y = AX.$$

Proof.

$$\sum_{j=1}^{n} y_j e_j = \varphi(a) = \varphi\left(\sum_{i=1}^{n} x_i e_i\right) = \sum_{i=1}^{n} x_i \varphi(e_i) = \sum_{i=1}^{n}\left(x_i \sum_{j=1}^{n} \alpha_{ji} e_j\right)$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_{ji} x_i e_j = \sum_{j=1}^{n}\left(\sum_{i=1}^{n} \alpha_{ji} x_i\right) e_j = \sum_{j=1}^{n} (AX)_j e_j.$$

□

THEOREM. *Let V be a vector space and let $(e) = (e_1, \ldots, e_n)$ and $(f) = (f_1, \ldots, f_n)$ be bases in V. Denote by S the matirx of the basis transformation $(e) \to (f)$. Denote by A the matrix of $\varphi \in \tau_V$ in the basis (e) and by B its matrix in the basis (f). Then*

$$B = S^{-1} A S.$$

Proof. Let $A = (\alpha_{ij}), B = (\beta_{ij}), S = (\gamma_{ij})$. For any j $(1 \leq j \leq n)$ we have

$$\varphi(f_j) = \varphi\left(\sum_{i=1}^{n} \gamma_{ij} e_i\right) = \sum_{i=1}^{n} \gamma_{ij} \varphi(e_i) = \sum_{i=1}^{n} \gamma_{ij}\left(\sum_{k=1}^{n} \alpha_{ki} e_k\right)$$

$$= \sum_{i=1}^{n}\sum_{k=1}^{n} \gamma_{ij} \alpha_{ki} e_k = \sum_{k=1}^{n}\left(\sum_{i=1}^{n} \alpha_{ki} \gamma_{ij}\right) e_k = \sum_{k=1}^{n} (AS)_{kj} e_k$$

On the other hand

$$\varphi(f_j) = \sum_{i=1}^{n} \beta_{ij} f_i = \sum_{i=1}^{n} \beta_{ij}\left(\sum_{k=1}^{n} \gamma_{ki} e_k\right) = \sum_{i=1}^{n}\sum_{k=1}^{n} \beta_{ij} \gamma_{ki} e_k$$

$$= \sum_{k=1}^{n}\left(\sum_{i=1}^{n} \gamma_{ki} \beta_{ij}\right) e_k = \sum_{k=1}^{n} (SB)_{kj} e_k.$$

Comparing these results we get $AS = SB$, whence $B = S^{-1}AS$. □

LEMMA. *Let V_1, V_2 be vector spaces, $\varphi : V_1 \to V_2$ an isomorphic mapping. Then for any $\{a_1, \ldots, a_k\} \subseteq V_1$ we have*

$$\rho(a_1, \ldots, a_k) = \rho(\varphi(a_1), \ldots, \varphi(a_k)).$$

THEOREM. *Let V be a vector space, $(e) = (e_1, \ldots, e_n)$ a basis in V, let $\varphi \in \tau_V$, and denote by A the matrix of φ in the basis (e). Then*

$$\rho(\varphi) = \rho(A).$$

REMARK. $\rho(\varphi)$ is independent of the basis.

DEFINITION. *Let V be a vector space, $\varphi, \psi \in \tau_V, \lambda \in \mathbb{T}$. For any $a \in V$ let*

$$\begin{aligned}(\varphi+\psi)(a) &= \varphi(a)+\psi(a)\\ (\varphi\circ\psi)(a) &= \varphi(\psi(a))\\ (\lambda\varphi)(a) &= \lambda\varphi(a).\end{aligned}$$

THEOREM. *$\varphi+\psi$, $\varphi\circ\psi$, $\lambda\varphi$ are also linear transformations.*

THEOREM. *τ_V is an algebra over $\mathbb{T}$.*

Proof.

$$(\varphi\circ(\psi\circ\rho))(a) = \varphi((\psi\circ\rho)(a)) = \varphi(\psi(\rho(a))) = (\varphi\circ\psi)(\rho(a)) = ((\varphi\circ\psi)\circ\rho)(a)$$

□

THEOREM. *Let V be a vector space with basis $(e) = (e_1, \ldots, e_n)$, let $\varphi \in \tau_V$. Let A and B be the matrices of φ and ψ in the basis (e), respectively. Then the matrices of $\varphi+\psi$, $\varphi\circ\psi$ and $\lambda\varphi$ in the basis (e) are $A+B, AB, \lambda A$, respectively.*

Proof. Let $A = (\alpha_{ij}), B = (\beta_{ij})$.

$$(\varphi+\psi)(e_i) = \varphi(e_i)+\psi(e_i) = \sum_{j=1}^n \alpha_{ji}e_j + \sum_{j=1}^n \beta_{ji}e_j = \sum_{j=1}^n (\alpha_{ji}+\beta_{ji})e_j =$$

$$= \sum_{j=1}^n (A+B)_{ji}e_j,$$

$$(\varphi\circ\psi)(e_i) = \varphi(\psi(e_i)) = \varphi\left(\sum_{j=1}^n \beta_{ji}e_j\right) = \sum_{j=1}^n \beta_{ji}\varphi(e_j) = \sum_{j=1}^n \beta_{ji}\left(\sum_{k=1}^n \alpha_{kj}e_k\right)$$

$$= \sum_{j=1}^n\sum_{k=1}^n (\beta_{ji}\alpha_{kj}e_k) = \sum_{k=1}^n \left(\sum_{j=1}^n \alpha_{kj}\beta_{ji}\right) e_k = \sum_{k=1}^n (AB)_{ki}e_k,$$

$$(\lambda\varphi)(e_i) = \lambda(\varphi(e_i)) = \lambda\sum_{j=1}^k \alpha_{ji}e_j = \sum_{j=1}^k \lambda\alpha_{ji}e_j = \sum_{j=1}^k (\lambda A)_{ji}e_j,$$

□

DEFINITION. *Let A_1, A_2 be algebras over the field $\mathbb{T}$. A_1 is isomorphic to A_2, if there is a mapping $F : A_1 \to A_2$ which is bijective, linear, moreover*

$$\forall a, b \in A_1 : F(ab) = F(a)F(b).$$

THEOREM. *Let V be a vector space over $\mathbb{T}$ with basis $(e) = (e_1, \ldots, e_n)$. Denote by $F : \tau_V \to \mathcal{M}_{n\times n}$ the mapping for which for any $\varphi \in \tau_V$ $F(\varphi)$ is the matrix of of φ in the basis (e). Then F is an isomorphic mapping of τ_V onto $\mathcal{M}_{n\times n}$.*

CONSEQUENCE.

$$\dim \tau_V = n^2.$$

5.3 Similar matrices

DEFINITION. *The matrices $A, B \in \mathcal{M}_{n\times n}$ are* **similar**, *if there is a regular matrix $S \in \mathcal{M}_{n\times n}$, such that*

$$B = S^{-1}AS.$$

THEOREM. *Similarity of matrices is an equivalence relation.*

Proof.
$A = E^{-1}AE$,
If $B = S^{-1}AS$, then $A = (S^{-1})^{-1}B(S^{-1})$
If $B = S^{-1}AS$ and $C = T^{-1}BT$, then $C = (ST)^{-1}A(ST)$. □

THEOREM. *Similar matrices have the same rank and determinant.*

REMARK. Similar matrices have also an equal characteristic polynomial.

Proof.

$$\rho(B) = \rho(S^{-1}AS) = \rho(S(S^{-1}AS)) = \rho(AS) = \rho((AS)^t) = \rho(S^tA^t)$$

$$= \rho((S^t)^{-1}(S^tA^t)) = \rho(A^t) = \rho(A).$$

$$|B| = |S^{-1}AS| = |S^{-1}|\;|A|\;|S| = |S|^{-1}\;|A|\;|S| = |A|.$$

□

5.4 Automorphisms

DEFINITION. *If V is a vector space and $\varphi \in \tau_V$ is bijective then it is an automorphism.*

THEOREM. *Let $\varphi \in \tau_V$.*
1. φ is an automorphism if and only if it has a regular matrix in any basis.
2. If φ is automorphic, then φ^{-1} is also automorphic and if φ has matrix A in the basis $(e) = (e_1, \dots, e_n)$, then φ^{-1} has matrix A^{-1} in this basis.

THEOREM. *The following statements are pairwise equivalent.*
1. φ is automorphic
2. φ is injective
3. $\mathrm{Ker}(\varphi) = \{0\}$
4. φ is surjectiv
5. $\varphi(V) = V$
6. $\rho(\varphi) = n$
7. φ has a regular matrix in any basis.

5.5 Invariant subspaces of linear transformations

DEFINITION. *The subspace L is an invariant subspace of $\varphi \in \tau_V$ if*

$$\forall a \in L : \varphi(a) \in L.$$

REMARK. $V, \{0\}, \varphi(V), \mathrm{Ker}(\varphi)$ are always invariant.

THEOREM. *Any subspace of V is invariant subspace of $\varphi \in \tau_V$ if and only if there is a $\lambda \in T$ such that*

$$\forall a \in V : \varphi(a) = \lambda a.$$

DEFINITION. *Let L be a subspace in the vector space V, let $\varphi \in \tau_V$. Assume that L is an invariant subspace of φ. The restriction of φ on L is a linear transformation $\varphi/L : L \to L$ such that*

$$\forall a \in L : \quad (\varphi/L)(a) = \varphi(a).$$

THEOREM. *Let L, M be invariant subspaces of φ with $L \oplus M = V$. Let $(e_1, \dots, e_k)$ be a basis in L, denote by A the matrix of φ/L is this basis. Let $(f_1, \dots, f_l)$ be a* *basis in M, denote by B the matrix of φ/M in this basis. Then the matrix of φ in the basis $(e_1, \dots, e_k, f_1, \dots, f_l)$ of V is*

$$\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$$

6 Spectral theory of linear transformations

6.1 Eigenvalue, eigenvector

DEFINITION. *Let V be a vector space over $\mathbb{T}$ and let $\varphi \in \tau_V$. If*

$$\varphi(a) = \lambda a$$

for some vector $0 \neq a \in V$ and scalar $\lambda \in T$, then λ is an eigenvalue of φ and a is an eigenvector of φ.

THEOREM. *If λ is an eigenvalue of φ, then then vectors a with $\varphi(a) = \lambda a$ form a subspace in V.*

DEFINITION. *If λ is an eigenvalue of φ, then the subspace*

$$L_\lambda = \{a \mid \varphi(a) = \lambda a\}$$

is called the eigenspace corresponding to λ.

REMARK.

1. L_λ contains the eigenvectors corresponding to λ and the zero vector.
2. Let A be the matrix of φ in a basis of V. Denote by $X_0 \in \mathbb{T}^n$ the coordinates of $a \in V$ in this basis. $\varphi(a) = \lambda a$ is equivalent to $AX_0 = \lambda X_0$, that is $(A - \lambda E)X_0 = 0$. Therefore the coordinate n-tuples of the eigenvectors are the solutions of the system of homogeneous linear equations

   $$(A - \lambda E)X = 0$$

3. The eigenspace is an invariant subspace of φ.

THEOREM. *Let V be a vector space, $\varphi \in \tau_V$. The eigenvectors corresponding to pairwise distinct eigenvalues of φ form a linearly independent vector system.*

Proof. $\lambda_1, \ldots, \lambda_k$, $a_1, \ldots, a_k$
proof by induction:

$$\mu_1 a_1 + \ldots + \mu_{k-1} a_{k-1} + \mu_k a_k = 0 \tag{7}$$

$$\lambda_1 \mu_1 a_1 + \ldots + \lambda_{k-1} \mu_{k-1} a_{k-1} + \lambda_k \mu_k a_k = 0,$$

$$\lambda_k \mu_1 a_1 + \ldots + \lambda_k \mu_{k-1} a_{k-1} + \lambda_k \mu_k a_k = 0.$$

$$(\lambda_1 - \lambda_k)\mu_1 a_1 + \ldots + (\lambda_{k-1} - \lambda_k)\mu_{k-1} a_{k-1} = 0.$$

By induction we obtain $\mu_1 = \ldots = \mu_{k-1} = 0$, hence $\mu_k a_k = 0$. $\square$

THEOREM. *The sum of eigenspaces corresponding to distinct eigenvalues of $\varphi \in \tau_V$ is a direct sum.*

6.2 Characteristic polynomial

DEFINITION. *The characteristic polynomial of a matrix $A \in \mathcal{M}_{n\times n}$ with elements in $\mathbb{T}$ is*

$$f(x) = |A - xE|$$

where $E \in \mathcal{M}_{n\times n}$ is the unit matrix.

REMARK. If $A = (\alpha_{ij})$, then the determinant

$$f(x) = |A - xE| = \begin{vmatrix} \alpha_{11} - x & \alpha_{12} & \dots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} - x & \dots & \alpha_{2n} \\ \vdots & \vdots & & \vdots \\ \alpha_{n1} & \alpha_{n2} & \dots & \alpha_{nn} - x \end{vmatrix}$$

is a polynomial of x of degree n.

THEOREM. **Cayley–Hamilton theorem** *Every square matrix is the root of its characteristic polynomial.*

REMARK.

$$f(x) = \alpha_n x^n + \alpha_{n-1} x^{n-1} + \ldots + \alpha_1 x + \alpha_0$$
$$f(A) = \alpha_n A^n + \alpha_{n-1} A^{n-1} + \ldots + \alpha_1 A + \alpha_0 E = 0$$

THEOREM. *The caharcteristic polynomials of similar matrices are equal.*

Proof. $B = S^{-1}AS$. Akkor

$$|B - xE| = |S^{-1}AS - xE| = |S^{-1}AS - xS^{-1}ES| = |S^{-1}(A - xE)S|$$
$$= |S^{-1}||A - xE||S| = |A - xE|.$$

□

DEFINITION. *Let $\varphi \in \tau_V$ and denote by A the matrix of φ in the basis $(e_1, \ldots, e_n)$ of V. The characteristic polynomial of φ is*

$$f(x) = |A - xE|.$$

REMARK. The characteristic polynomial is independent from the basis.

DEFINITION. *The characteristic roots of $\varphi \in \tau_V$ are the roots in T of its characteristic polynomial.*

REMARK. If $\mathbb{T} = \mathbb{C}$, then all roots of the characteristic polynomial are in $\mathbb{C}$. If $\mathbb{T} = \mathbb{R}$ then in general not all roots are in $\mathbb{R}$.

THEOREM. *$\lambda \in \mathbb{T}$ is an eigenvalue of $\varphi \in \tau_V$ if and only if it is a characterictic root of φ.*

Proof. $\lambda \in \mathbb{T}$ is eigenvalue if and only if there is a non-zero vector $a \in V$ with

$$\varphi(a) = \lambda a.$$

$$(A - \lambda E)X = 0$$

□

DEFINITION. *Let $\lambda \in \mathbb{T}$ be an eigenvalue of φ.*
The algebraic multplicity of λ yields the multiplicity if λ as a root of the characteristic polynomial: $\mathrm{mult}\lambda$.
The geometric multiplicity of λ is the dimension of its eigenspace. $\dim L_\lambda$.

THEOREM. *If λ is an eigenvalue of $\varphi \in \tau_V$, then*

$$\dim L_\lambda \leq \mathrm{mult}\lambda.$$

Proof. Let $(e_1, \dots, e_k)$ be a basis of L_λ. Extend this basis to be a basis of V: $(e_1, \dots, e_k, e_{k+1}, \dots, e_n)$.

$$\begin{pmatrix} \lambda & \dots & 0 & \alpha_{1,k+1} & \dots & \alpha_{1n} \\ \vdots & & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & \lambda & \alpha_{k,k+1} & \dots & \alpha_{kn} \\ 0 & \dots & 0 & \alpha_{k+1,k+1} & \dots & \alpha_{k+1,n} \\ \vdots & & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & 0 & \alpha_{n,k+1} & \dots & \alpha_{nn} \end{pmatrix}.$$

The characteristic polynomial of the matrix is divisible by $(\lambda - x)^k$. Therefore $\mathrm{mult}\lambda \geq k = \dim L_\lambda$. □

6.3 The spectrum of linear transformations

DEFINITION. *The spectrum of $\varphi \in \tau_V$ is the set of its eigenvalues, each taken with multiplicity according to its algebraic multiplicity. The spectrum of φ is complete if it consits of $\dim V = n$ elements.*

NOTATION. $\mathrm{Sp}\varphi$

REMARK. If $\mathbb{T} = \mathbb{C}$ then $\mathrm{Sp}\varphi$ is always complete, if $\mathbb{T} = \mathbb{R}$ then not always.

THEOREM. *There exists a basis of V consisting of the eigenvectors if $\varphi \in \tau_V$ if and only if the following two conditions are satisfied:*

1. $\mathrm{Sp}\varphi$ *is complete*
2. For each eigenvalue λ *of* φ *we have* $\mathrm{mult}\lambda = \dim L_\lambda$.

THEOREM. *Any linear transformation in a vector space over* $\mathbb{R}$ *or* $\mathbb{C}$ *admits an invariant subspace of dimension at most two.*

Proof.
$$|A - (\alpha + i\beta)E| = 0,$$
$$(A - (\alpha + i\beta)E)X = 0$$
$$0 = (A - (\alpha + i\beta))(X_0 + iY_0) = (AX_0 - \alpha X_0 + \beta Y_0) + i(AY_0 - \alpha Y_0 - \beta X_0).$$
$$AX_0 = \alpha X_0 - \beta Y_0, \quad AY_0 = \alpha Y_0 + \beta X_0$$
$$\varphi(a) = \alpha a - \beta b, \quad \varphi(b) = \alpha b + \beta a$$

□

6.4 Nilpotent operators

Definition. *Let V be a vector space, $\varphi \in \tau_V$. φ is* **nilpotent** *if there is a positive integer r such that $\varphi^r = \mathcal{O}$, the zero operator. The smalles such r is called the* **index of nilpotency** *of φ.*

Theorem. *If $\varphi \in \tau_V$ is a nilpotent operator of index r, and for $e \in V$ we have $\varphi^{r-1}(e) \neq 0$, then*
I. $(e, \varphi(e), \ldots, \varphi^{r-1}(e))$ are linearly independent
II. $M = \mathcal{L}(e, \varphi(e), \ldots, \varphi^{r-1}(e))$ is a φ–invariant subspace and there is a φ–invariant subspace N with $M \oplus N = V$.

Theorem. *If $\varphi \in \tau_V$ is nilpotent of index r, then there are positive integers $r_1, \ldots, r_k$ and vectors $e_1, \ldots, e_k$ such that*
1. $r_1 \geq \ldots \geq r_k$
2. $(e_1, \varphi(e_1), \ldots, \varphi^{r_1-1}(e_1), \ldots, e_k, \varphi(e_k), \ldots, \varphi^{r_k-1}(e_k))$ is a basis in V
3. $\varphi^{r_1}(e_1) = \ldots = \varphi^{r_k}(e_k) = 0$.

Remark. φ/M_j has the following matrix in the basis $(e_j, \varphi(e_j), \ldots, \varphi^{r_j-1}(e_j))$ of the subspace M_j:

$$\begin{pmatrix} 0 & 0 & 0 & \ldots & 0 & 0 \\ 1 & 0 & 0 & \ldots & 0 & 0 \\ 0 & 1 & 0 & \ldots & 0 & 0 \\ 0 & 0 & 1 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & 1 & 0 \end{pmatrix}$$

According to the theorem there is a basis of V in which the matrix of φ consists of such blocks along the main diagonal.

6.5 Jordan normal form

In this section we assume that V is a vector space over the complex numbers, or the spectrum of φ is complete if V is over the real numbers.

Theorem. *For any $\varphi \in \tau_V$ there are invariant subspaces R, N such that φ/N is nilpotent, φ/R is automorph and $V = R \oplus N$.*

Theorem. *Let $\lambda_1, \ldots, \lambda_p$ be the distinct eigen values of $\varphi \in \tau_V$ with (algebraic) multiplicities $m_1, \ldots, m_p$, respectively. Then there are subspaces $M_1, \ldots, M_p$ in V, such that*
1. $V = M_1 \oplus \ldots \oplus M_p$
2. $\dim M_j = m_j (1 \leq j \leq p)$
3. M_j is a φ–invariant subspace of V $(1 \leq j \leq p)$
4. $\varphi - \lambda_j \varepsilon$ is nilpotent on M_j–n $(1 \leq j \leq p)$, where ε is the identical transformation.

Remark. φ_j/M_j is nilpotent, therefore there is a basis of M_j in which the matrix of φ_j/M_j consits of blocks of type

$$\begin{pmatrix} 0 & 0 & 0 & \dots & 0 & 0 \\ 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & 0 \end{pmatrix}$$

Since $\varphi = \varphi_j + \lambda_j \varepsilon$, hence in the same basis the matrix of φ/M_j consists of blocks of type

$$\begin{pmatrix} \lambda_j & 0 & 0 & \dots & 0 & 0 \\ 1 & \lambda_j & 0 & \dots & 0 & 0 \\ 0 & 1 & \lambda_j & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & \lambda_j \end{pmatrix}$$

These are called **blocks of Jordan type**. Constract such a basis in all subspaces M_j. The union of these bases is a basis in V. In this basis the matrix of φ consists of Jordan type blocks along the main diagonal. Such a matrix is said to have a **Jordan normal form**.

7 Linear, bilinear and quardatic forms

7.1 Linear forms

Definition. *The linear mapping $\ell: V \to \mathbb{T}$ is called* **linear form**.

7.2 Bilinear forms

Definition. *The mapping $L: V \times V \to \mathbb{T}$ is a* **bilinear form** *if it is linear in both variables, that is*

$$L(\lambda_1 x_1 + \lambda_2 x_2, y) = \lambda_1 L(x_1, y) + \lambda_2 L(x_2, y)$$

$$L(x, \lambda_1 y_1 + \lambda_2 y_2) = \lambda_1 L(x, y_1) + \lambda_2 L(x, y_2)$$

for any $x, x_1, x_2, y, y_1, y_2 \in V$ and $\lambda_1, \lambda_2 \in \mathbb{T}$

If $(b_1, \ldots, b_n)$ is a basis in V-ben, then the matrix of the bilinear form L in this basis is the matrix of type $n \times n$ consisting of the elements $\alpha_{ij} = L(b_i, b_j)$. If the coordinate columns of the vectors $x, y \in V$ in the basis $(b_1, \ldots, b_n)$ are $X, Y \in \mathbb{T}^n$, respectively, then we have

$$L(x, y) = X^t AY,$$

because

$$L(x, y) = L(\sum_{i=1}^{n} x_i b_i, \sum_{j=1}^{n} y_j b_j) = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i y_j L(b_i, b_j) =$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} x_i y_j \alpha_{ij} = X^t BY.$$

Remark. The matrix of a bilinear form is unique in a given basis.

Statement. *Let $(b_1, \ldots, b_n)$ and $(f_1, \ldots, f_n)$ be bases in V, assume that the transition matrix of $(b) \to (f)$ is S. If the bilinear form L has matrix A and B in the above bases, respectively, then*

$$B = S^t AS$$

Proof. Let $S = (s_{ij}), A = (\alpha_{ij}), B = (\beta_{ij})$. then

$$f_j = \sum_{k=1}^{n} s_{kj} b_k.$$

We have

$$\beta_{ij} = L(f_i, f_j) = L(\sum_{k=1}^{n} s_{ki} b_k, \sum_{l=1}^{n} s_{li} b_l) =$$

$$= \sum_{k=1}^{n}\sum_{l=1}^{n} s_{ki}s_{li}L(b_k, b_l) = \sum_{k=1}^{n}\sum_{l=1}^{n} s_{ki}\alpha_{kl}s_{li}.$$

□

As a consequence, the rank of the matrix of a bilinear form is the same in any basis. This is called the **rank of the bilinear form**.

DEFINITION. *A bilinear form $L: V \times V \to \mathbb{T}$ is* **symmetric** *if $L(x, y) = L(y, x)$ for any $x, y \in V$.*

The bilinear form L is symmetric if and only if its matrix in any basis is symmetric.

DEFINITION. *If $L: V \times V \to \mathbb{T}$ is a symmetric bilinear form, then $Q(x) = L(x, x)$, $Q: V \to \mathbb{T}$ is a* **quadratic form**.

The quadratic form uniquely determines the bilinear form, since for any $x, y \in V$-re

$$\begin{aligned} Q(x+y) = L(x+y, x+y) &= L(x,x) + 2L(x,y) + L(y,y) \\ &= Q(x) + 2L(x,y) + Q(y) \end{aligned}$$

therefore

$$L(x, y) = \frac{1}{2}\left(Q(x+y) - Q(x) - Q(y)\right)$$

.

STATEMENT. *If $(b_1, \ldots, b_n)$ is a basis in V, then any quadratic form can be represented as $Q(x) = \sum_{i,j=1}^{n} \alpha_{ij}x_ix_j$ with a symmetric matrix $A = (\alpha_{ij})$, where $(x_1, \ldots, x_n)$ denotes the coordinate column of the vector x in the basis $(b_1, \ldots, b_n)$. Conversely, for any symmetric matrix A, the above formula defines a quadratic form.*

7.3 Canonical form

A **canonical basis** of a bilinear form L is a basis in which the bilinear form admits the **canonical form**

$$L(x, y) = \lambda_1 x_1 y_1 + \ldots + \lambda_n x_n y_n$$

.

A **canonical basis** of a quadratic form L is a basis in which the quadratic form admits the **canonical form**

$$Q(x_1, \ldots, x_n) = \lambda_1 x_1^2 + \ldots + \lambda_n x_n^2$$

.

THEOREM. **Lagrange's theorem.** *Any symmetric bilinear form admits a canonical basis.*

As a consequence any quadratic form admits a canonical basis.

THEOREM.**Sylvester theorem, inertia law** *In any canonical form*

$$L(x,y) = \lambda_1 x_1 y_1 + \ldots + \lambda_n x_n y_n$$

of a bilinear form the positive, negative and zero coefficients among $\lambda_1, \ldots, \lambda_n$ are independent of the canonical basis.

Practical method to transform a quadratic foem into a canonical form

Lagrange method

Let $Q(x) = Q(x_1, \ldots, x_n) = \sum_{i,j=1}^{n} \alpha_{ij} x_i x_j$ be a quadratic form.

A) If $\alpha_{ii} = 0$ for all $i = 1, \ldots, n$, but e.g. $\alpha_{12} \neq 0$. Then perform the coordinate transformation

$$\begin{aligned}
x_1' &= \frac{1}{2}(x_1 + x_2) \\
x_2' &= \frac{1}{2}(x_1 - x_2) \\
x_i' &= x_i, \qquad \text{ha} \qquad i = 3, \ldots, n
\end{aligned}$$

B) If e.g. $\alpha_{11} \neq 0$, then let

$$\begin{aligned}
Q(x_1, \ldots, x_n) &= \alpha_{11} \left(x_1^2 + 2 \sum_{i=2}^{n} \frac{\alpha_{i1}}{\alpha_{11}} x_1 x_i \right) + \sum_{i,j=2}^{n} \alpha_{ij} x_i x_j \\
&= \alpha_{11} \left(x_1 + \sum_{i=2}^{n} \frac{\alpha_{i1}}{\alpha_{11}} x_i \right)^2 - \alpha_{11} \sum_{i=2}^{n} \frac{\alpha_{i1}^2}{\alpha_{11}^2} x_i^2 \\
&\quad -2 \sum_{i,j=2}^{n} \frac{\alpha_{i1}\alpha_{j1}}{\alpha_{11}} x_i x_j + \sum_{i,j=2}^{n} \alpha_{ij} x_i x_j
\end{aligned}$$

Introducing the transformation

$$\begin{aligned}
x_1' &= x_1 + \sum_{i=2}^{n} \frac{\alpha_{i1}}{\alpha_{11}} x_i \\
x_i' &= x_i \qquad \text{ha} \qquad i = 2, \ldots, n
\end{aligned}$$

we obtain

$$Q(x_1, \ldots, x_n) = \alpha_{11} x_1'^2 + \tilde{Q}(x_2', \ldots, x_n').$$

THEOREM. **Jacobi theorem**.
Let $L(x, y) = \sum_{i,j,=1}^n \alpha_{ij} x_i y_j$ *be a symmetric bilinear form. If the minors* $\Delta_i = \begin{vmatrix} \alpha_{11} & \ldots & \alpha_{i1} \\ \vdots & & \vdots \\ \alpha_{i1} & \ldots & \alpha_{ii} \end{vmatrix}$ *are all non-zero, then there exists a basis, in which* L *has the canonical form*

$$L(x, y) = \frac{\Delta_0}{\Delta_1} x_1 y_1 + \ldots + \frac{\Delta_{n-1}}{\Delta_n} x_n y_n,$$

where $(x_1, \ldots, x_n)$ *and* $(y_1, \ldots, y_n)$ *are the coordinate columns of the vectors* x *and* y *corresponding to the basis, and* $\Delta_0 = 1$.

CONSEQUENCE. **Jacobi theorem on quadratic forms**.
Let $Q(x_1, \ldots, x_n) = \sum\limits_{i,j,=1}^n \alpha_{ij} x_i x_j$ *be a quadratic form. If the minors* $\Delta_i = \begin{vmatrix} \alpha_{11} & \ldots & \alpha_{i1} \\ \vdots & & \vdots \\ \alpha_{i1} & \ldots & \alpha_{ii} \end{vmatrix}$ *are all non-zero, then there exists a basis in which* Q *has the canonical form*

$$Q(y_1, \ldots, y_n) = \frac{\Delta_0}{\Delta_1} y_1^2 + \ldots + \frac{\Delta_{n-1}}{\Delta_n} y_n^2,$$

where $(y_1, \ldots, y_n)$ *denotes the coordinate column of the vector* y *in this basis and* $\Delta_0 = 1$.

DEFINITION. *The quadratic form* Q *is* **positive definite**, *if* $Q(x) > 0$ *for all* $x \neq 0$.

A quadratic form

$$Q(x_1, \ldots, x_n) = \lambda_1 x_1^2 + \ldots + \lambda_n x_n^2$$

is obviously positive definite if and only if $\lambda_1 > 0, \ldots, \lambda_n > 0$.

The following theorem gives a criterion for this property without using canonical forms.

THEOREM. *A quadratic form is positive definite if and only if all minors* Δ_i $(i = 1, \ldots, n)$ *of its matrix are positive.*

8 Inner product spaces

8.1 The concept of an inner product space

DEFINITION. *An* **inner product** *(or* **scalar product***) on a real vector space is a symmetric bilinear form L such that the quadratic form $Q(x) = L(x, x)$ is positive definite.*
Notation: $(x, y) = L(x, y)$ *An* **inner product space** *(or* **Euclidean space***) is a real vector space equipped with an inner product. The* **norm of a vector** *is*

$$||x|| = \sqrt{(x, x)}$$

STATEMENT. **Cauchy-Bunyakovszkij-Schwarz inequality**
For any vectors x, y of an Euclidean space E we have

$$|(x, y)| \leq ||x|| \cdot ||y||.$$

Proof. For any $\lambda \in \mathbb{R}$ we have

$$(x + \lambda y, x + \lambda y) \geq 0$$

that is

$$(x, x) + 2\lambda(x, y) + \lambda^2(y, y) \geq 0$$

The discriminant of the second degree polynomial is $D \leq 0$:

$$4(x, y)^2 - (x, x)(y, y) \leq 0.$$

□

DEFINITION. *For any non-zero vectors x, y, the* **angle** *of the vectors is the α with*

$$\cos\alpha = \frac{(x, y)}{||x|| \cdot ||y||}.$$

STATEMENT. **Minkowski inequality** *For any $x, y \in E$ we have*

$$||x + y|| \leq ||x|| + ||y||.$$

Proof.

$$\begin{aligned} ||x+y||^2 = (x+y, x+y) = (x, x) + 2(x, y) + (y, y) &\leq ||x||^2 2|(x, y)| + ||y||^2 \\ &\leq ||x||^2 + 2||x|| \cdot ||y|| + ||y||^2 = (||x|| + ||y||)^2 \end{aligned}$$

□

REMARK.

1. $d(x, y) = ||x - y||$ is a metric on E.

8.2 Orthogonality

Definition. *The vectors x, y in the Euclidean space E are orthogonal if $(x, y) = 0$. Notation: $x \perp y$. The vector $x \in E$ is* **normed** *(unit vector) if $||x|| = 1$. A system $e_1, e_2, \ldots, e_k$ of vectors in E is an* **orthonormal vectors system**, *if it consists of pairwise orthogonal unit vectors.*

The orthonormal property can be expressed as

$$(e_i, e_j) = \delta_{ij} \quad (1 \leq i, j \leq n).$$

Statement. *An orthonormed vector system not containing the null vector is linearly independent.*

Statement. **Gram–Schmidt orthogonalization algorithm**
For any basis $(b_1, b_2, \ldots, b_n) \in E$ there exists an orthonormed basis $(e_1, e_2, \ldots, e_n)$ such that $\mathcal{L}(e_1, \ldots, e_k) = \mathcal{L}(b_1, \ldots, b_k)$ is satisfied for $k = 1, 2, \ldots, n$. The vectors $e_1, e_2, \ldots, e_n$ are unique up to sign.

Proof. Set $e_1 = \frac{b_1}{||b_1||}$. If $e_1, \ldots, e_k$ are known then let

$$e'_{k+1} = b_{k+1} - (b_{k+1}, e_1)e_1 - \ldots - (b_{k+1}, e_k)e_k,$$

and

$$e_{k+1} = \frac{e'_{k+1}}{||e'_{k+1}||}.$$

□

Remark. If $(e_1, \ldots, e_n)$ is an orthonormed basis in E and $x = x_1e_1 + \ldots + x_ne_n$, then $x_i = (x, e_i)$ for $i = 1, \ldots, n$-re, since

$$(x, e_i) = (x_1e_1 + \ldots + x_ne_n, e_i) = x_1(e_1, e_i) + \ldots + x_n(e_n, e_i) = x_i(e_i, e_i) = x_i.$$

Further, if $x = x_1e_1 + \ldots + x_ne_n$ and $y = y_1e_1 + \ldots + y_ne_n$, then

$$(x, y) = (x, y_1e_1 + \ldots + y_ne_n) = y_1(x, e_1) + \ldots + y_n(x, e_n) = x_1y_1 + \ldots + x_ny_n,$$

and

$$||x||^2 = x_1^2 + \ldots + x_n^2.$$

Definition. *The Euclidean spaces E_1 és E_2 are called* **isomorphic** *if there exists a bijective linear mapping $\varphi: E_1 \to E_2$ such that*

$$(\varphi(x), \varphi(y)) = (x, y) \qquad (x, y \in E_1).$$

STATEMENT. *Two Euclidean spaces are isomorphic if and only if their dimensions are equal.*

Proof. Take orthonormed bases $(e) \in E_1$ and $(f) \in E_2$ is both Euclidean spaces. The mapping with $\varphi(e_i) = f_i$ is an isomorphic mapping of E_1 onto E_2. □

DEFINITION. *The* **orthogonal complement** *of a subspace L of an Euclidean space E is*

$$L^\perp = \{x \in E \mid x \perp y, \quad \forall y \in L\}.$$

REMARK. $L^\perp$ is a subspace of E since if $x_1, x_2 \in L^\perp, \lambda_1, \lambda_2 \in \mathbb{R}$, then $\lambda_1 x_1 + \lambda_2 x_2 \in L^\perp$ is also satisfied, since

$$(\lambda_1 x_1 + \lambda_2 x_2, y) = \lambda_1(x_1, y) + \lambda_2(x_2, y) = 0.$$

STATEMENT. *For any subspace L of an Euclidean space E we have*

$$L \oplus L^\perp = E,$$

and

$$(L^\perp)^\perp = L.$$

REMARK.

1. According to the Theorem for any $x \in E$ there uniquely exist $x' \in L$ and $x'' \in L^\perp$ with $x = x' + x''$. The vector x' is called the **orthogonal projection** of x on the subspace L. The mapping $p_L : E \to E, x \mapsto x'$ is linear with the idempotent property $p_L^2 = p_L$.

STATEMENT. **Bessel inequatily**
If $(e_1, \dots, e_k)$ is an orthonormed vector system in the Eucliedean space E, then for any $x \in E$ we have

$$(x, e_1)^2 + \ldots + (x, e_k)^2 \leq ||x||^2.$$

Here for any $x \in E$ equality is satisfied if and only if $(e_1, \dots, e_k)$ is a basis in E.

REMARK.

1. If $k = \dim E$ then we have equality for any $x \in E$. This is called **Parseval equation**.

Definition. *The* **distance** *of the vector x from the subspace L is*

$$d(x, L) = ||x - x'||$$

Remark.

$$d(x, L) = \min\{||x - y|| \ \mid y \in L\}$$

since

$$||x - y||^2 = ||(x - x') + (x' - y)||^2 = ||x - x'||^2 + ||x' - y||^2 \geq ||x - x'||^2.$$

8.3 Complex inner product spaces (unitary spaces)

Definition. *Let U be a complex vector space. The mapping $\ell\colon U \to \mathbb{C}$ is called* **conjugated linear form** *if for any $\lambda, \mu \in \mathbb{C}$ és $x, y \in U$ we have*

$$\ell(\lambda x + \mu y) = \bar{\lambda}\ell(x) + \bar{\mu}\ell(y)$$

Remark. If $(b_1, \ldots, b_n)$ is a basis in the vector space U, then there uniquely exist $\alpha_1, \ldots, \alpha_n \in \mathbb{C}$ such that for any $x = x_1 b_1 + \ldots x_n b_n \in U$ we have

$$\ell(x) = \alpha_1 \overline{x}_1 + \ldots + \alpha_n \overline{x}_n.$$

Definition. *Let U be a complex vector space. The mapping $L\colon U \times U \to \mathbb{C}$ is a* **Hermitian bilinear form**, *if it is linear in the first variable and conjugated linear in the second variable. A Hermitian bilinear form L is* **Hermitian-symmetric**, *if for any $x, y \in U$ we have*

$$L(y, x) = \overline{L(x, y)}.$$

Remark.

1. If $(b_1, \ldots, b_n)$ is a basis in U, $X = (x_1, \ldots, x_n)^t \in \mathbb{C}^n$ and $Y = (y_1, \ldots, y_n)^t \in \mathbb{C}^n$ are the coordinates of the vectors x and y in this basis, repsectively, then

$$L(x, y) = \sum_{j=1}^{n} \sum_{k=1}^{n} \alpha_{jk} x_j \overline{y}_k = XA\overline{Y}^t$$

where $A = (\alpha_{jk})$ is the coefficient matirx of L with $\alpha_{jk} = L(b_j, b_k)$.

2. We call the basis $(b_1, \ldots, b_n)$ canonical, if

$$L(x, y) = \sum_{j=1}^{n} \lambda_j x_j \overline{y}_j.$$

Definition. *If L is a Hermitian-symmetric, Hermitian-bilinear form, then $Q: U \to \mathbb{C}$ with*

$$Q(x) = L(x, x) \quad (x \in U)$$

is a **quadratic form** *on U.*

Remark.

- $Q(x)$ is real, since $L(x, x) = \overline{L(x, x)}$.
- If $(b_1, \ldots, b_n)$ is a basis in U, then

$$Q(x) = \sum_{j=1}^{n} \sum_{k=1}^{n} \alpha_{jk} x_j \overline{x}_k = XA\overline{X}^t.$$

  The basis is canonical, if

$$Q(x) = \sum_{j=1}^{n} \lambda_j |x_j|^2.$$

  Here $\lambda_j = L(b_j, b_j)$ is real and also for any $x \in U$ the value $Q(x)$ is real.

Definition. *A* **complex inner product** *is a Hermitian-symmetric Hermitian bilinear form $L(x, y)$ such that $Q(x) = L(x, x)$ is a positive definite quadratic form.*
Notation: $(x, y) = L(x, y)$
A **complex inner product space** *(or unitary space) is a complex vector spaces equipped with an inner product.*

Examples.

1. $\mathbb{C}^n$ is a complex inner product space with

$$(x, y) = x_1 \overline{y}_1 + \ldots + x_n \overline{y}_n$$

   for any $x = (x_1, \ldots, x_n)^t$, $y = (y_1, \ldots, y_n)^t \in \mathbb{C}^n$.

2. If f, g are complex valued continuos functions on the interval $[a, b]$, then

$$(f, g) = \int_a^b f(x)\overline{g(x)}\, dx$$

   is an inner product on the vector space of these functions.

Remark.

The majority of statements for Euclidean spaces remain valid also in a complex inner product space.

- In an orthonormed basis we have

$$\begin{aligned} (x, y) &= x_1\bar{y}_1 + \ldots + x_n\bar{y}_n. \\ \|x\| &= |x_1|^2 + \ldots + |x_n|^2. \end{aligned}$$

- In a complex inner product space we can not define the angle of the vectors, only the orthogonality by $(x, y) = 0$.
- The Gram-Schmidt orthogonalization process works the same way as in the real case.
- The Cauchy-Bunyakovszkij-Schwarz and the Minkowski inequalities remain valid.

9 Transformations in inner product spaces

9.1 Forms represented by inner products

STATEMENT. *If W is a real or complex inner product space, and ℓ is a linear form on W, then there uniquely exists $a \in W$, such that*

$$\ell(x) = (x, a) \quad (x \in W).$$

If ℓ is a conjugated linear form, then there uniquely exists $a \in W$, with

$$\ell(x) = (a, x) \quad (x \in W).$$

THEOREM. *If W is a real (or complex) inner product space, then for any bilinear (Hermitian bilinear) form there uniquely exist linear transformations $\varphi, \psi \in \tau_W$ such that for any $x, y \in W$ we have*

$$L(x, y) = (\varphi(x), y) = (x, \psi(y)).$$

On the other hand for any $\varphi, \psi \in \tau_W$

$$L(x, y) = (\varphi(x), y)$$

and

$$L(x, y) = (x, \psi(y))$$

are bilinear (Hermitian bilinear) forms on W.

9.2 Adjoints of transformations

DEFINITION. *Let W be a real (complex) inner product space, and let $\varphi, \psi \in \tau_W$. If*

$$(\varphi(x), y) = (x, \psi(y))$$

for any $x, y \in W$, then ψ is called the **adjoint** *of φ.*
Notation: φ^*.

Note that according to the above statement for any φ its adjoint uniquely exists. The adjoint is defined by

$$(\varphi(x), y) = (x, \varphi^*(y))$$

for any $x, y \in W$.

STATEMENT. *Let W be a real (complex) inner product space, $\varphi\psi \in \tau_W$, $\lambda \in \mathbb{R}$ ($\lambda \in \mathbb{C}$). Then*

$$(\varphi + \psi)^* = \varphi^* + \psi^*$$

$$(\lambda\varphi)^* = \overline{\lambda}\varphi^*$$
$$(\varphi \circ \psi)^* = \psi^* \circ \varphi^*$$
$$\varphi^{**} = \varphi.$$

Further, in any orthonormed basis the matrix of φ^* is the transposed conjugated of the matrix of φ.

Proof.

$$(x, (\varphi + \psi)^*(y)) = ((\varphi + \psi)(x), y) =$$
$$(\varphi(x) + \psi(x), y) = (\varphi(x), y) + (\psi(x), y) =$$
$$= (x, \varphi^*(y)) + (x, \psi^*(y)) = (x, (\varphi^* + \psi^*)(y))$$

$$(x, (\lambda\varphi)^*(x)) = ((\lambda\varphi)(x), y) = \lambda(\varphi(x), y) =$$
$$= \lambda(x, \varphi^*(y)) = (x, \overline{\lambda}\varphi^*(y)) = (x, (\overline{\lambda}\varphi^*)(x)),$$
$$(x, (\varphi \circ \psi)^*(y)) = ((\varphi \circ \psi)(x), y) = (\varphi(\psi(x)), y) =$$
$$= (\psi(x), \varphi^*(y)) = (x, \psi^*(\varphi^*(y))) = (x, (\psi^* \circ \varphi^*)(y)),$$
$$(x, \varphi^{**}(y)) = (\varphi^*(x), y) = \overline{(y, \varphi^*(x))} = \overline{(\varphi(y), x)} = (x, \varphi(y)).$$

$$(\varphi(e_k), e_l) = (\sum_{j=1}^{n} \alpha_{jk} e_j, e_l) = \sum_{j=1}^{n} \alpha_{jk}(e_j, e_l) = \sum_{j=1}^{n} \alpha_{jk}\delta_{jl} = \alpha_{lk},$$

$$(\varphi(e_k), e_l) = (e_k, \varphi^*(e_l)) = (e_k, \sum_{j=1}^{n} \beta_{jl} e_j) = \sum_{j=1}^{n} \overline{\beta}_{jl}(e_k, e_j) = \sum_{j=1}^{n} \overline{\beta}_{jl}\delta_{kj} = \overline{\beta}_{kl}$$

whence $B = \overline{A}^t$. □

DEFINITION. *Let W be a real (complex) inner product space, let $\varphi \in \tau_W$. We call φ*

- **symmetric (selfadjoint)**, *if $\varphi^* = \varphi$,*
- **orthogonal (unitary),** *if $\varphi^* = \varphi^{-1}$*
- **normal,** *if $\varphi \circ \varphi^* = \varphi^* \circ \varphi$*

We use similar names for matrices.

DEFINITION. *The* **adjoint** *of a matrix A of type $n \times n$ with real (complex) entries is*

$$A^* = \overline{A}^t.$$

DEFINITION. *The matrix A of type $n \times n$ with real (complex) entries is called*

- **symmetric (self adjoint)** *if $A^* = A$*
- **orthogonal (unitary),** *if $A^* = A^{-1}$*
- **normal,** *if $A \circ A^* = A^* \circ A$*

CONSEQUENCE. *The rank of φ is equal to the rank of φ^*
λ is an eigenvalues of φ if and only if $\overline{\lambda}$ is an eigenvalue of φ^*. For any transformation φ the ranks and defects of φ and its conjugate are equal. The eigenvalues of φ^* are the conjugates of the eigenvalues of φ.*

Proof.

$$\operatorname{rg}\varphi = \operatorname{rg} A = \operatorname{rg}\overline{A}^t = \operatorname{rg}\varphi^*.$$

$$p_\varphi(\lambda) = \det(A - \lambda E) = \det(A - \lambda E)^t =$$

$$\det(A^t - \lambda E) = \overline{\det(\overline{A} - \overline{\lambda} E)} = \overline{p_{\varphi^*}(\overline{\lambda})}.$$

□

STATEMENT.

$$Ker\,\varphi^* = (Im\,\varphi)^\perp$$

$$Ker\,\varphi = (Im\,\varphi^*)^\perp.$$

9.3 Selfadjoint transformations

φ is selfadjoint if and only if for any $x, y \in W$ we have

$$(\varphi(x), y) = (x, \varphi(y)).$$

In an orthonormed basis it yields $A = \overline{A}^t$ for the matrices of φ and φ^*.

EXAMPLES.

1. Let W be a real (complex) inner product space, let $\lambda \in \mathbb{R}$ ($\lambda \in \mathbb{C}$). $\varphi_\lambda \colon W \to W$, $x \mapsto \lambda x$ is a self adjoint transformation.
2. The orthogonal projection on a subspace of W is a self adjoint transformation. $W = L \oplus L^\perp$, $x = x' + x''$, $y = y' + y''$ ($x', y' \in L, x'', y'' \in L^\perp$),
$$(p_L(x), y) = (x', y' + y'') = (x', y') + (x', y'') =$$
$$= (x', y') + (x'', y') = (x' + x'', y') = (x, p_L(y))$$
since $(x', y'') = (x'', y') = 0$.
3. If the matrix of a linear transformation in an orthonormed basis is a diagonal matrix with real entries, then it is a self adjoint transformation.

STATEMENT. *The roots of the characteristic polynomial of a symmetric (self adjoint) transformation on a real (complex) inner product space are all real numbers.*

CONSEQUENCE. *The spectrum of a symmetric (self adjoint) transformation is complete.*
Proof.
We use the precess of complexification

$$\lambda(x, x) = (\varphi'(x), x) = (x, (\varphi')^*(x)) = (x, \lambda x) = \overline{\lambda}(x, x)$$

whence $\lambda = \overline{\lambda}$. □

STATEMENT. *The eigenvectors belonging to distinct eigenvalues of a symmetric (self adjoint) tranformation on a real (complex) inner product space are orthogonal.*

Proof.

$$\varphi(x) = \lambda x, \qquad \varphi(y) = \mu y, \qquad \lambda \neq \mu.$$
$$\lambda(x, y) = (\lambda x, y) = (\varphi(x), y) = (x, \varphi(y)) = (x, \mu y) = \overline{\mu}(x, y) = \mu(x, y)$$

□

LEMMA. *If $H \subset W$ is an invariant subspace of a symmetric (self adjoint) transformation on a real (complex) inner product space, then $H^\perp$ is also invariant.*

THEOREM. **Structure theorem.**
If φ is a symmetric (self adjoint) transformation on a real (complex) inner product space, then there exists an orthonormed basis consisting of eigenvectors of φ.

THEOREM. **Principal axis theorem**
Any quadratic form in a real (complex) inner product space admits an orthonormed canonical basis. The coefficients of the quadratic form in such a canonical form are just the eigenvalues of the matrix of the quadratic form belonging to the orhonormed basis.

STATEMENT. *Any real (complex) valued symmetric (self adjoint) matrix is similar to a diagonal matrix with real entries.*

THEOREM.**Spectral representation theorem**
Any symmetric (self adjoint) transformation φ on a real (complex) inner product space can be represented in the form

$$\varphi = \lambda_1\pi_1 + \ldots + \lambda_k\pi_k$$

with orthogonal projections $\pi_1, \ldots, \pi_k$ which are permutable with φ (that is $\varphi \circ \pi_i = \pi_i \circ \varphi$), with $\pi_i \circ \pi_j = 0$, if $i \neq j$, and $\pi_1 + \ldots + \pi_k = id.$, and with distinct real numbers $\lambda_1, \ldots, \lambda_k$. Conversly, any such transformation is symmetric (self adjoint).

9.4 Orthogonal and unitary transformations

STATEMENT. *If L is an invariant subspace of an orthogonal (unitary) transformation on a real (complex) inner product space. Then the orthogonal complement $L^\perp$ of L is also invariant.*

STATEMENT. *Let W be a real (complex) inner product space, and let $\varphi: W \to W$. The following statements are pairwise equivalent:*

1. *φ is orthogonal (unitary)*
2. *φ is inner product preserving, that is for any $x, y \in W$ we have $(\varphi(x), \varphi(y)) = (x, y)$.*
3. *φ is norm preserving, that is for any $x \in W$ we have $\|\varphi(x)\| = \|x\|$.*
4. *φ maps any orthonormed basis into an orthonormed basis.*

REMARK.

1. An orthogonal (unitary) transformation can only have eigenvalues of absolute value 1. If $\varphi(a) = \lambda a$, then by

$$\|a\| = \|\varphi(a)\| = \|\lambda a\| = |\lambda| \, \|a\|,$$

we have $|\lambda| = 1$.

2. Let $(e_1, \ldots, e_n)$ be an orthonormed basis, denote by $S = (s_{ij})$ the transition matrix $(e_1, \ldots, e_n) \to (f_1, \ldots, f_n)$.

$$(f_i, f_j) = (\sum_{k=1}^{n} s_{ki} e_k, \sum_{l=1}^{n} s_{lj} e_l) =$$

$$= \sum_{k=1}^{n} \sum_{l=1}^{n} s_{ki} \overline{s}_{lj} (e_k, e_l) = \sum_{k=1}^{n} \sum_{l=1}^{n} s_{ki} \overline{s}_{lj} \delta_{kl} = \sum_{k=1}^{n} s_{ki} \overline{s}_{kj}$$

$$= \sum_{k=1}^{n} (S^T)_{ik} \overline{s}_{kj} = (S^T \overline{S})_{ij} = \overline{(S^* S)_{ij}}$$

Therefore (f) is orthonormed if and only if S is orthogonal (unitary).

Statement. *A matrix of type $n \times n$ with real (complex) entries is orthogonal (unitary) if and only if its columns are pairwise orthogonal vectors of norm 1 in $\mathbb{R}^n$ (in $\mathbb{C}^n$). The same holds for the rows. An orthogonal (unitary) matrix has determinant of absolute value 1.*

Proof.

$$(A^{(i)}, A^{(j)}) = \sum_{k=1}^{n} a_{ki} \overline{a}_{kj}.$$

$$1 = \det(E) = \det(\overline{A}^t A) = \det(\overline{A}) \det(A) = \overline{\det(A)} \det(A) = |\det(A)|^2$$

□

9.5 Orthogonal transformations of Euclidean spaces

Statement. *In the two dimensional plane all orthogonal transformations are the identity, reflection onto a line, central reflection on the origin and rotation around the origin.*

Rotation:

$$M_\varphi = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}.$$

Statement. *If φ is an orthogonal transformation on the Euclidean space E, then*

$$E = L_1 \oplus \ldots \oplus L_k$$

where L_i are pairwise orthogonal invariant subspaces of φ of dimensions 1 or 2.

REMARK. In a suitable basis the matrix of an orthogonal transformation is a **quasi-diagonal** matrix

$$M_\varphi = \begin{pmatrix} \begin{matrix} \cos\alpha_1 & -\sin\alpha_1 \\ \sin\alpha_1 & \cos\alpha_1 \end{matrix} & 0 & \dots & & 0 \\ 0 & \dots & & & \\ \vdots & & \begin{matrix} \cos\alpha_s & -\sin\alpha_s \\ \sin\alpha_s & \cos\alpha_s \end{matrix} & & \vdots \\ & & & \lambda_1 & \\ \vdots & & & & \dots \quad 0 \\ 0 & \dots & & & \lambda_{n-2s} \end{pmatrix}.$$

STATEMENT. *If A is an orthogonal matrix then there is an orthogonal matrix S such that S^tAS is quasi-diagonal.*

9.6 Normal transformations of complex inner product spaces

STATEMENT. *In a complex inner product space any linear transformation can be uniquely represented in the from $\varphi = \varphi_1 + i\varphi_2$, where φ_1, φ_2 are self adjoint transformations. φ is normal if and only if $\varphi_1 \circ \varphi_2 = \varphi_2 \circ \varphi_2$.*

STATEMENT. *If φ_1 and φ_2 are symmetric (self adjoint) transformations of a real (complex) inner product space with $\varphi_1 \circ \varphi_2 = \varphi_2 \circ \varphi_2$, then there exists an orthonormed basis consisting of common eigenvectors of φ_1 and φ_2.*

THEOREM.**Structure theorem**
In a complex inner product space a linear transformation admits an orthonormed basis consisting of its eigenvectors if and only if it is normal.

STATEMENT. *Let φ be a normal transformation in a complex inner product space. φ is self-adjoint if and only if its eigenvalues are real. φ is unitary if and only if its eigenvalues are of absolute value 1.*

9.7 Polar representation theorem

DEFINITION. *The symmetric (self-adjoint) linear transformation φ of the euclidean (unitary) space W is* **positive definite** *if for any non-zero $x \in W$ we have $(\varphi(x), x) > 0)$.*

THEOREM. *A symmetric (self-adjoint) linear transformation φ of the euclidean (unitary) space W is positive definite if and only if all eigenvalues of φ are positive.*

THEOREM. *If φ is a positive definite linear transformation of the euclidean (unitary) space W, then there is a positive definite linear transformation ψ of W, such that*

$$\varphi = \psi^2.$$

THEOREM.**Polar representation theorem**
For any invertible linear transformation φ of the euclidean (unitary) space W there exist a positive definite linear transformation ψ and an orthogonal (unitary) transformation τ such that

$$\varphi = \psi \circ \tau.$$

10 Curves of second order

Definition. *Let $a_{ij}(1 \le i, j \le 3)$ be real numbers with $a_{ij} = a_{ji}(1 \le i, j \le 3)$ and $a_{11}^2 + a_{22}^2 + a_{33}^2 > 0$. The set of points (x, y) in the plain satisfying the equation*

$$a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2 + 2a_{13}x_1 + 2a_{23}x_2 + a_{33} = 0$$

is called **curve of second order**.

The symmetric matrix

$$A = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{pmatrix}$$

is the matrix of the curve.
The symmetric matrix

$$A_{33} = \begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{pmatrix}$$

is the kernel matrix of the curve. Let

$$X = \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}.$$

Then the equation of the curve can be writte as

$$X^t A X = 0$$

or

$$\sum_{i=1}^{2}\sum_{k=1}^{2} a_{ik}x_ix_k + \sum_{i=1}^{2} a_{i3}x_i + a_{33} = 0.$$

10.1 Curves of second order and lines

Let $x_i = d_i + tv_i (i = 1, 2)$ be a line with a non-zero vector (v_1, v_2). The intersections of the curve and the line are at the parameters t satisfying

$$t^2\left(\sum_{i=1}^{2}\sum_{k=1}^{2} a_{ik}v_iv_k\right) + 2t\left(\sum_{i=1}^{2}\sum_{k=1}^{2} a_{ik}d_iv_k + \sum_{i=1}^{2} a_{i3}v_i\right) + \left(\sum_{i=1}^{2}\sum_{k=1}^{2} a_{ik}d_id_k + \sum_{i=1}^{2} a_{i3}d_i + a_{33}\right) = 0 \tag{8}$$

Definition. (v_1, v_2) *is an* **asymptote direction** *if*

$$\sum_{i=1}^{2}\sum_{k=1}^{2} a_{ik}v_iv_k = 0.$$

$$a_{11}v_1^2 + 2a_{12}v_1v_2 + a_{22}v_2^2 = 0$$
$$a_{11}\left(\frac{v_1}{v_2}\right)^2 + 2a_{12}\left(\frac{v_1}{v_2}\right) + a_{22} = 0$$
$$D = 4a_{12} - 4a_{11}a_{22} = -4|A_{33}|$$

DEFINITION. *The curve is*
elliptic *if it admits no asymptote directions* $\iff |A_{33}| > 0$
hiperbolic *if it admits 2 asymptote directionss* $\iff |A_{33}| < 0$
parabolic *if it admits 1 asymptote directions* $\iff |A_{33}| = 0$

DEFINITION. *P is the* **center** *of the curve if for any point A of the curve there exists a point B of the curve such that P is the midpoint of the segment AB. The curve is* **central** *if it admits exactly one center.*

THEOREM. *The curve is central* $\iff |A_{33}| \neq 0$

10.2 Diameter of the curve

Assume the line $x_i = d_i + tv_i (i = 1, 2)$ has two common points P_1, P_2 with the curve, corresponding to the parameters t_1, t_2. The segment P_1P_2 is a **chord** of the curve. If $D(d_1, d_2)$ is the midpoint of the segment P_1P_2 then for the parameters t_1, t_2 we have $t_1 + t_2 = 0$ whence

$$\sum_{i=1}^{2}\sum_{k=1}^{2} a_{ik}d_iv_k + \sum_{i=1}^{2} a_{i3}v_i = 0$$

that is (d_1, d_2) satisfies

$$(a_{11}v_1 + a_{12}v_2)x_1 + (a_{21}v_1 + a_{22}v_2)x_1 + (a_{13}v_1 + a_{23}v_2) = 0$$

which is the equation of a line. This line is called the **diameter conjugated to the direction** (v_1, v_2). The above argument implies that the diameter is a line.

THEOREM. *If the curve is central, then its diameters pass through the center.*
If the curve is not central, the its diameters are parallel to each other.

10.3 The principal axis of the curve

DEFINITION. *If the direction is orthogonal to the diameter corresponding to it, then the diameter is called the* **principal axis** *of the curve.*

THEOREM. *The principal axes of the curve are exactly the diameters conjugated to the eigenvectors corresponding to the non-zero eigenvalues of the kernel matrix.*

THEOREM. *A curve of second order admits 1 or 2 principal axes.*

10.4 Classification of curves of second order

The curve is called **degenerate** if $|A| = 0$.

THEOREM. **Principal axix transformation**
There is an orthogonal transformation of the plane bringing the equation of the curve to the shape

$$\lambda_1 y_1^2 + \lambda_2 y_2^2 + 2c_{13}y_1 + 2c_{23}y_2 + c_{33} = 0$$

where λ_1, λ_2 are the eigenvalues of the kernel matrix.

Elliptic curves

circle	$x_1^2 + x_2^2 = a^2$
point circle	$x_1^2 + x_2^2 = 0$
imaginary circle	$x_1^2 + x_2^2 = -a^2$
ellipse	$\dfrac{x_1^2}{a^2} + \dfrac{x_2^2}{b^2} = 1$
point ellipse	$\dfrac{x_1^2}{a^2} + \dfrac{x_2^2}{b^2} = 0$
imaginary ellipse	$\dfrac{x_1^2}{a^2} + \dfrac{x_2^2}{b^2} = -1$

Hiperbolic curves

hyperbola	$\dfrac{x_1^2}{a^2} - \dfrac{x_2^2}{b^2} = 1$
intersecting lines	$\dfrac{x_1^2}{a^2} - \dfrac{x_2^2}{b^2} = 0$

Parabolic curves

parabola	$x_1^2 = 2px_2$
parallel lines	$x_1^2 = a^2$
identical lines	$x_1^2 = 0$
a pair of imaginary parallel lines	$x_1^2 = -a^2$

11 Függelék

11.1 Algebrai alapfogalmak

Ebben a fejezetben felsoroljuk azokat az algebrai alapfogalmakat, melyeket a jegyzetben felhasználunk. Minden esetben adunk példákat, egy közismert egyszerű példát és egy lineáris algebrai jellegűt.

DEFINITION. *A H nem üres halmazon $*$* **művelet***, ha bármely $a, b \in H$–hoz egyértelműen létezik egy $c \in H$, melyre*

$$c = a * b.$$

Ha a H halmazon a $*$ művelet értelmezve van, akkor $(H, *)$–ot algebrai struktúrának nevezzük.

REMARK. A művelet H–n tulajdonképpen egy $H \times H \to H$ leképezés.

DEFINITION. *Az $(F, *)$ algebrai struktúra* **félcsoport***, ha a művelet* **asszociatív***:*

$$\forall a, b, c \in F : \quad a * (b * c) = (a * b) * c.$$

EXAMPLES.

1. Félcsoportot alkot a természetes számok halmaza a szorzásra nézve.
2. Félcsoportot alkotnak a valós elemű 2×2 típusú mátrixok a szorzásra nézve.

DEFINITION. *A $(G, *)$ algebrai struktúra* **csoport***, ha*

- *$(G, *)$ félcsoport*
- *létezik* **neutrális elem***, azaz olyan $e \in G$, melyre*

$$\forall a \in G : \quad a * e = e * a = a$$

- *és minden elemnek van* **inverz***e, azaz*

$$\forall a \in G \;\; \exists a' \in G : \quad a * a' = a' * a = e.$$

EXAMPLES.

1. Csoportot alkotnak az egész számok az összeadásra nézve.
2. Csoportot alkotnak a 2×2 típusú valós elemű reguláris mátrixok a szorzásra nézve.

REMARK.

1. A $(G,*)$ csoport **kommutatív csoport, vagy Abel–féle csoport**, ha $(G,*)$ olyan csoport, melyen a művelet kommutatív:

$$\forall a,b \in G: \quad a*b = b*a$$

   Az előző példák közül az első kommutatív, a második nem kommutatív csoport.

2. Ha a művelet összeadás, akkor a neutrális elemet **nullelem**nek, ha a művelet szorzás, akkor a neutrális elemet **egységelem**nek nevezzük.

DEFINITION. *Az $(R,+,*)$ algebrai struktúra* **gyűrű**, *ha*

- $(R,+)$ *kommutatív csoport*
- $(R,*)$ *félcsoport*
- *érvényes a* **disztributivitás**, *azaz*

$$\forall a,b,c \in R: \quad \begin{aligned} a*(b+c) &= (a*b)+(a*c) \\ (a+b)*c &= (a*c)+(b*c). \end{aligned}$$

EXAMPLES.

1. Az egész számok halmaza gyűrű az összeadás és szorzás műveletére nézve.
2. A valós elemű 3×3 típusú mátrixok gyűrűt alkotnak a mátrixokon szokásos összeadás és szorzás műveletére nézve.

DEFINITION. *A $(T,+,*)$ algebrai struktúra* **test**, *ha*

- $(T,+)$ *kommutatív csoport*
- $(T\setminus\{0\},*)$ *kommutatív csoport, ahol 0 jelöli a neutrális elemet a + műveletre nézve*
- *teljesül a disztributivitás.*

EXAMPLE. A racionális számok testet alkotnak a szokásos összeadás és szorzás műveletére nézve.

DEFINITION. *A $(T,+,*)$ test* **karakterisztikája nulla**, *ha nem létezik olyan n természetes szám, hogy minden $t \in T$-re $n \cdot t = 0$, ahol $n \cdot t$ az n tagú $t + \ldots + t$ összeget jelöli.*

EXAMPLE. A racionális, valós, komplex számok teste nullkarakterisztikájú.

DEFINITION. *A V nem üres halmaz* **vektortér** *a T test felett, ha*

- *$(V, +)$ kommutatív csoport*
- *bármely $a, b \in V$ vektorok és $\lambda, \mu \in T$ skalárok esetén fennállnak az alábbi azonosságok:*

$$
\begin{aligned}
\lambda(a+b) &= \lambda a + \lambda b \\
(\lambda+\mu)a &= \lambda a + \mu a \\
(\lambda\mu)a &= \lambda(\mu a) = \mu(\lambda a) \\
1a &= a
\end{aligned}
$$

ahol 1 a T test multiplikatív egységeleme.

EXAMPLES.

1. A valós számok halmaza vektortér a racionális számok teste felett.
2. A valós elemű számhármasok vektorteret alkotnak a valós számok teste felett az alábbi műveletekre nézve:

$$
\begin{aligned}
(x_1, x_2, x_3) + (y_1, y_2, y_3) &= (x_1 + y_1, x_2 + y_2, x_3 + y_3) \\
\lambda(x_1, x_2, x_3) &= (\lambda x_1, \lambda x_2, \lambda x_3)
\end{aligned}
$$

DEFINITION. *Az A nem üres halmaz* **algebra** *a T test felett, ha*

- *A vektortér T felett*
- *$(A, +, *)$ gyűrű*
- *Bármely $a, b \in A$ és $\lambda \in T$ esetén fennáll*

$$\lambda(a * b) = (\lambda a) * b = a * (\lambda b).$$

REMARK. Természetesen a vektortéren és a gyűrűben értelmezett összeadás ugyanaz a művelet.

EXAMPLES.

1. A komplex számok halmaza algebra a racionális számok teste felett.
2. A valós elemű, 3×3 típusú mátrixok algebrát alkotnak a valós számok teste felett.

11.2 Alapvető tudnivalók permutációkról

Ebben a fejezetben összefoglaljuk a permutációk azon tulajdonságait, melyeket a determinánsokról szóló fejezetben felhasználunk.

DEFINITION. *Az $(1, 2, \ldots, n)$ számok egy sorrendjét ezen számok egy* **permutáció***jának nevezzük.*

NOTATION.

1. Az $(1, 2, \ldots, n)$ összes permutációinak halmazát P_n–nel jelöljük.
2. Ha $\pi = (i_1, i_2, \ldots, i_n)$ az $(1, 2, \ldots, n)$ egy permutációja, akkor azt a jelölést is szokás alkalmazni, hogy

$$\pi = \begin{pmatrix} 1 & 2 & \ldots & n \\ i_i & i_2 & \ldots & i_n \end{pmatrix}$$

   A felső sorban a számoknak nem feltétlenül nagyság szerint növekvő sorrendben kell következniük.

REMARK.

1. A permutáció tekinthető az $(1, 2, \ldots, n)$ halmaz önmagára történő kölcsönösen egyértelmű leképezésének is.
2. Az $(1, 2, \ldots, n)$ számok összes permutációinak száma $n!$.

DEFINITION. *Az $(1, 2, \ldots, n)$ számok $(i_1, i_2, \ldots, i_n)$ permutációjában i_k* **inverzió***ban áll i_l–lel, ha $k < l$ de $i_k > i_l$ $(1 \le k < l \le n)$.*

NOTATION. Az $(i_1, i_2, \ldots, i_n)$ összes inverzióinak a számát $I(i_1, i_2, \ldots, i_n)$–nel jelöljük.

EXAMPLE. Az $(5, 1, 4, 2, 3)$ permutáció inverzióinak száma 6.

DEFINITION. *Egy* **permutáció páros***, ha összes inverzióinak száma páros, egyébként* **páratlan**.

LEMMA. *Bármely $(i_1, i_2, \ldots, i_n)$ permutáció létrehozható az $(1, 2, \ldots, n)$–ből kiindulva, csak elempárok egymásutáni cseréjével.*

Proof. Az i_1–et az első számmal megcserélve elérhetjük, hogy az első helyre kerüljön. Hasonló módon hozzuk az i_2–t a második helyre, stb. □

LEMMA. *Két elem cseréjénel az inverziók számának paritása ellenkezőjére változik.*

Proof. Tekintsük az $(i_1, \ldots, i_k, \ldots, i_l, \ldots, i_n)$ permutációt és benne cseréljük fel az i_k és i_l számokat, igy az $(i_1, \ldots, i_l, \ldots, i_k, \ldots, i_n)$ permutációt kapjuk. A

cserekor i_k–nak ellenkezőjére változik az inverziója az i_k és i_l között lévő s db számmal. (ha nem voltak inverzióban, akkor inverzióban lesznek a csere után, és ha eddig inverzióban voltak, akkor a csere után nem lesznek). Ugyancsak, i_l–nek ellenkezőjére változik az inverziója az i_k és i_l között lévő s db számmal. Végül, i_k és i_l egymás közötti inverziója is megváltozik. Más elemek közötti inverzióban nem történik változás. Így összesen $2s + 1$ változás történik az inverziók számában.
□

DEFINITION. *A*

$$\pi = \begin{pmatrix} 1 & 2 & \dots & n \\ i_1 & i_2 & \dots & i_n \end{pmatrix}, \quad \rho = \begin{pmatrix} 1 & 2 & \dots & n \\ j_1 & j_2 & \dots & j_n \end{pmatrix}$$

permutációk szorzata

$$\begin{pmatrix} 1 & 2 & \dots & n \\ j_{i_1} & j_{i_2} & \dots & j_{i_n} \end{pmatrix}.$$

EXAMPLE.

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 4 & 1 & 5 & 3 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 3 & 5 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 2 & 4 & 1 & 5 \end{pmatrix}$$

REMARK.

1. A permutációk szorzása nem kommutatív.
2. A permutációk ezen műveletre nézve csoportot alkotnak, melynek egységeleme az $(1, 2, \dots, n)$ identikus permutáció.
3. Egy π permutáció **inverzén** azt a ρ permutációt értjük, melyre $\pi\rho$ az identikus permutáció.

LEMMA. *Azonos paritású permutációk szorzata páros, ellenkező paritású permutációk szorzata páratlan.*

Proof. A $\pi\rho$ szorzat elvégzésekor a π permutációban szereplő sorrendet rendezzük tovább a ρ permutációnak megfelelően. Ha a π permutációt az $(1, 2, \dots, n)$ permutációból k db elempárcserével lehet létrehozni, a ρ permutációt l db elempár cserével, akkor a $\pi\rho$ permutáció létrehozható $k+l$ db elempár cseréjével. Ha k, l paritása azonos, akkor $k + l$ páros, egyébként $k + l$ páratlan. □

REMARK. Könnyen ellenőrizhető, hogy a

$$\pi = \begin{pmatrix} 1 & 2 & \dots & n \\ i_1 & i_2 & \dots & i_n \end{pmatrix}$$

permutáció inverze

$$\pi^{-1} = \begin{pmatrix} i_1 & i_2 & \dots & i_n \\ 1 & 2 & \dots & n \end{pmatrix}$$

LEMMA. *A permutációnak a paritása megegyezik inverzének paritásával.*

Proof. A szorzatuk az egységpermutáció, ami páros, tehát az előző Lemma értelmében vagy mindkettő páros, vagy mindkettő páratlan. □

11.3 MAPLE: lineáris algebrai programcsomag

Ezen fejezet célja a MAPLE komputeralgebra programcsomag lineáris algebrával kapcsolatos eljárásainak áttekintése. Ezen eljárások igen hatékony eszközt adnak felhasználóik kezébe a lineáris algebrai feladatok numerikus megoldására.

11.3.1 A Maple általános használata

Először megismerkedünk a MAPLE programcsomag használatának néhány általános vonásával. Tudnivaló, hogy a MAPLE–t lehet interaktív módon használni és lehet programot is írni MAPLE utasítások felhasználásával. Mi itt csakis a MAPLE interakív használatának bemutatására szorítkozunk. Leírásunk korántsem teljes, csak ízelítőt ad a lehetséges felhasználási területekből.

A MAPLE programcsomagot konkrét installációjától függően legtöbbször a

```
maple
```

paranccsal lehet betölteni. Ezután a képernyőn megjelenik a MAPLE cég emblémája, és egy > prompt a sor elején, mely után lehet a parancsokat beírni. A parancs általában egy ; jellel záródik. A parancs beírása után az Enter billentyűt nyomjuk meg, melyre az adott parancs végrehajtódik, és az eredmény a képernyőn megjelenik. Amennyiben a parancsot a : jellel zárjuk, a parancs végrehajtódik, de az eredmény nem jelenik meg a képernyőn.

Kilépés:

```
> quit;
```

Remark.

1. *HELP* funkció aktiválása. A DOS-os verzióban F1-gyel a Windown-os verzióban a megfelelő ikonra való kattintással történik. A *help menüben* magától értetődő módon lehet keresni a kívánt információt. Célszerű mindenképpen kipróbálni. A *help* funkciót lehet konkrét eljárások nevével is hívni, ekkor a konkrét eljárás használatáról ad információt. Pl. a *help* hívása a *det* eljárásra vonatkozóan:

   ```
   > help(det);
   ```

2. *Visszalapozás a képernyőn.* A DOS-os verzióban F5-tel, a Windows-os verzióban a szokásos módon történik. Akkor használjuk, ha az eredmény túl hosszú, vagy korábbi eredményre vagyunk kiváncsiak.

3. *Parancsok megismétlése, módosítása* A DOS-os verzióban a fölfelé és lefelé mutató kurzor nyilak segítségével, a Windows-os verzióban az egérrel lehet a már beadott parancsok között keresni. Ha a parancs, melyet használni akarunk nem sokban tér el egy korábbitól, akkor azt előkeresve és módosítva gyorsabban célhoz érünk, mint újra a teljes parancsot begépelve.

4. A " jel a legutoljára kiszámított eredményre utal, "" az azelőttire, stb.

Szimbólikus változóknak betűvel kezdődő, betűkből, számokból és egyéb (nem műveleti) jelekből álló összefüggő jelsorozatokat használunk. A kis és nagy betűk között különbség tevődik. A már értéket kapott szimbólumok értékeit (ha azok nem vektorok vagy mátrixok) kiírathatjuk úgy, hogy a parancs sorba a szimbólum neve mögé ;–t teszünk. Mátrixok vagy vektorok értékeit a *print* utasítással írathatjuk ki. Az értékadó utasítást := –val jelöljük.

```
> a:=1/2+3/7+2/11;
                                     171
                                a := ---
                                     154

> 154*a;

                                   171
```

A MAPLE a számokat pontos közönséges tört alakban tárolja, amíg a tizedes tört alak kiszámítását az *evalf* paranccsal elő nem írjuk:

```
> evalf(a);
                              1.110389610
```

A tizedes törtek 10 jegy pontosságúak, ami azonban a *Digits* változó megfelelő beállításával változtatható:

```
> Digits:=40;
                              Digits := 40
> evalf(a);
                1.110389610389610389610389610389610389610
> Digits:=10;
                              Digits := 10
> h:=tan(3/7*Pi);
                            h := tan(3/7 Pi)
> evalf(h);
                              4.381286277
```

A *Pi,E,I* rendre a $\pi, e, \sqrt{-1}$ szimbólumokat jelölik.

Az alapműveletek jele a szokásos $+, *, /$, a hatványozás jele a "háztető ".

```
> p:=2^(2^6)+1;
                        p := 18446744073709551617
```

Egész számokat az *ifaktor*, egész együtthatós polinomokat a *faktor* eljárással bonthatunk szorzatra.

```
> ifactor(p);
                        (67280421310721) (274177)
```

```
> f:=x^4-1;
                                   4
                             f := x  - 1

> factor(f);
                                          2
                      (x - 1) (x + 1) (x  + 1)
```

A *normal* eljárás törteket hoz egyszerűsített alakra.

```
> g:=(x^2-y^2)/(x-y)^3;
                                   2    2
                                  x  - y
                            g := --------
                                        3
                                 (x - y)

> g:=normal(g);
                                   x + y
                            g := --------
                                        2
                                 (x - y)
```

Algebrai kifejezések egyszerűsítésére használhatjuk a *simplify* eljárást is, ennek előnye, hogy megadhatóak azon összefüggések, melyeket az egyszerűsítésnél fel lehet használni. Például, ha x, az x^2+x+1 polinom gyöke, akkor az x^5+x^3+1 polinom egyszerűsítése:

```
> simplify(x^5+x^3+1,{x^2+x+1=0});

                            1 - x
```

A szorzat alakok kiszámítása, a szorzás elvégzése az *expand* eljárással lehetséges.

```
> expand((x-y)*(x+y));

                             2    2
                            x  - y
```

A *subs* eljárás alkalmas algebrai kifejezések helyettesítési értékeinek kiszámítására.

```
> f:= x*y;
                            f:= x y
> f1:=subs(x=4,y=5,f);
                            f1:=20
> f2:=subs(x=3,y=7,f);
                            f2:=21
```

Komplex számok használata esetén az imaginárius egységet az I jelöli.

```
> a:=3+2*I;b:=1+5*I;
> c:=a/b;
> rc:=Re(c);
> ic:=Im(c);
```

Bonyolultabb komplex értékű kifejezéseket az *evalc* eljárással lehet kiértékeltetni.

Polinomok valós gyökeit az *fsolve* eljárás számítja ki, mely a *complex* paraméterrel kiegészítve a komplex gyököket is kiszámítja. Minden esetben az eljárás első paramétere a polinom, második pedig a változó szimbóluma. A gyököket eltárolhatjuk egy változóba is, ekkor ez a változó egy vektor lesz, melynek komponensei a megfelelő sorszámú gyököket tartalmazzák.

```
> f:=3*x^5-2*x^2+1;
> fsolve(f,x);
> g:=fsolve(f,x,complex);
> g[1];
```

A MAPLE–ben egyszerű módon lehet **függvényeket** megadni. Például azt a kétváltozós függvényt, melynek értéke az (i, j) helyen $(-1)^{i+j}$ így lehet definiálni:

```
> f:=(i,j)->(-1)^(i+j);
```

Ezek után a függvény hívása:

```
> f(2,3);
                                  - 1
```

11.3.2 Alapvető utasításelemek

Az *if* utasítás a szokásos kétféle alakban használatos (*else* ággal vagy anélkül). Lezárása *fi*–vel történik. Az *if, then, else, fi* szavak közötti esetleg több utasítás egy blokkot alkot.

```
> if a<12 then b:=3 fi;
> if a<12 then b:=3 else b:=4 fi;
```

A ciklikus utásítások legegyszerűbb formája a *for* ciklus, melyet *od* zár le. Ha nem adjuk meg a kezdőindexet (*from*) és a lépésközt (*by*), akkor a számlálás 1–től 1–esével történik. A *do* és *od* szavak közötti esetleg több utasítás alkotja a ciklus magját, ezek minden lépésben végrehajtódnak.

```
> for i from -1 by 2 to 14 do
  ...........
> od;
```

```
> k:=1;
> for j to 6 do k:=k*j od;
```

11.3.3 Lineáris algebra programcsomag

A MAPLE programcsomag tartalmaz számos olyan eljárást, melyek lineáris algebrai jellegű számítások elvégzését teszik lehetővé. Szemben az eddig tárgyalt eljárásokkal, ezek a MAPLE behívásakor nem kerülnek betöltésre. Használatuk csak akkor lehetséges, ha a

```
> with(linalg);
```

paranccsal külön betöltjük őket.

Adatstruktúrák

A lineáris algebrában leggyakrabban vektorokkal és mátrixokkal számolunk. Ezeket lehet definiálni vagy a méreteik megadásával és az elemek értékének megadásával, vagy közvetlenül az elemek megadásával. Mátrix elemeinek felsorolásakor az értékadás sorfolytonosan történik. A vektorok ill. mátrixok elemeire a megfelelő sor és oszlopindexek szögletes zárójelbeni megadásával hivatkozhatunk. Vektorokat vagy mátrixokat a *print* eljárás ír ki a képernyőre. Ennek megfelelően az alábbi vektor illetve mátrix definíciók egyenértékűek.

```
> v:=vector(3);v[1]:=-1;v[2]:=7;v[3]:=5;
> v:=vector([-1,7,5]);

> K:=matrix(2,3);
> K[1,1]:=1;K[1,2]:=-1;K[1,3]:=2;K[2,1]:=5;K[2,2]:=6;K[2,3]:=7;
> K:=matrix(2,3,[1,-1,2,5,6,7]);
> K:=matrix([[1,-1,2],[5,6,7]]);
```

Alapműveletek

Additív műveleteket egymással egyező típusú vektorokon vagy mátrixokon lehet végezni, és az eredmény is ugyanolyan típusú lesz. Mátrixok szorzásakor az ismert szabályok érvényesülnek a szorzandó és a szorzat méretére vonatkozóan.

Azonos típusú vektorok vagy mártixok összeadása:

```
> C:=add(A,B);
```

Azonos típusú A, B vektorok vagy mártixok lineáris kombinációja $c1, c2$ együtthatókkal:

```
> C:=add(A,B,c1,c2);
```

Az A vektor vagy mátrix elemeinek szorzása x–szel:

```
> scalarmul(A,x);
```

Az A mátrix i–edik sorának szorzása x–szel:

```
> mulrow(A,i,x);
```

Az A mátrix i–edik oszlopának szorzása x–szel:

```
> mulcol(A,i,x);
```

Egy A mátrix i–edik sorának x–szeresét hozzáadni a j–edik sorához:

```
> addrow(A,i,j,x);
```

Egy A mátrix i–edik oszlopának x–szeresét hozzáadni a j–edik oszlopához:

```
> addcol(A,i,j,x);
```

Az $n \times n$ típusú egységmátrix c-szeresének előállítása:

```
> band([c],n);
```

Olyan $n \times n$ típusú mátrix előállítása, melyben a főátló alatt $c1$, a főátlóban $c2$, a főátló fölött $c3$ áll:

```
> band([c1,c2,c3],n);
```

A B mátrixba az A mátrix bemásolása úgy, hogy az A bal felső sarka a B (m, n)–edik elemén lesz:

```
> copyinto(A,B,m,n);
```

Az A mátrixban az r,...,s sorok törlése:

```
> delrows(A,r..s);
```

Az A mátrixban az r,...,s oszlopok törlése:

```
> delcols(A,r..s);
```

Az A mátrix bővitése m sorral és n oszloppal, az új helyeket x–szel feltöltve:

```
> extend(A,m,n,x);
```

Az A mátrixban az r–edik sor és az s–edik oszlop elhagyása:

```
> minor(A,r,s);
```

Az A mátrix i–edik sora:

```
> row(A,i);
```

Az A mátrix i–edik oszlopa:

```
> col(A,i);
```

A B mátrix összeállítása a $v1, \ldots, vk$ vektorokból:

```
> B:=concat(v1,...,vk);
```

A v vektor dimenziója:

```
> vectdim(v);
```

Az A mátrix sorainak száma:

```
> rowdim(A);
```

Az A mátrix oszlopainak száma:

```
> coldim(A);
```

Az A mátrix transzponáltja:

```
> transpose(A);
```

Az A mátrix nyoma (azaz főátlóbeli elemeinek összege):

```
> trace(A);
```

Mátrixok vagy vektorok szorzása. A művelet csak megfelelő méretű mátrixok vagy vektorok esetén használható. Több összeszorzandó mátrix vagy vektor is megadható paraméterként.

```
> multiply(A,B);
> multiply(A,B,C);
```

Az A mátrix rangja:

```
> rank(A);
```

Az A négyzetes mátrix adjungáltja. Az adjungált mátrix (i,j)–edik eleme az a_{ji}–hez tartozó algebrai aldetermináns ($A * adj(A) = det(A) * E$).

```
> adj(A);
```

Az A négyzetes mátrix determinánsa:

```
> det(A);
```

Az A négyzetes mátrix inverze:

```
> inverse(A);
```

Az A vektor vagy mátrix normája. Második paraméterként megadható a norma típusa. Ez mátrixoknál $1, 2, frobenius, infinity$ lehet, vektoroknál pozitív egész szám, $frobenius, infinity$ lehet. Külön specializáció nélkül a végtelen norma kerül kiszámításra.

```
> norm(A);
> norm(A,infinity);
> norm(A,1);
> norm(A,2);
```

Az $i_1, \ldots, i_k$ számokhoz tartozó Vandermonde–féle mátrix:

```
> vandermonde([i1,...,ik]);
```

Mátrixokkal vagy vektorokkal végzett alapműveletek kiértékelése. Az *evalm* eljárás alkalmazásával az összeadás, skalárral való szorzás és szorzás műveleteinek kiszámítását nagyban egyszerűsíthetjük. Például, ha A és B azonos típusú négyzetes mátrixok, akkor $A^2 + A * B - 2 * A + 5 * B$ egy utasítással kiszámítható. Jegyezzük meg, hogy a szorzás műveletét ekkor &* jelöli:

```
> evalm(A^2+A&*B-2*A+5*B);
```

Vektorterek

A vektorterek és alterek bázisait halmazba lehet foglalni, és az eljárások paramétereként a bázisvektorok felsorolása helyett az őket tartalmazó halmazokat is meg lehet adni. Az eljárások egy részénél az eredmény is egy vektorhalmaz, melynek adott sorszámú elemére mint komponensére lehet hivatkozni:

```
a:=vector([1,4,2]);
b:=vector([-1,2,-7]);
alter:={a,b};
alter[1];
alter[2];
```

A $v_1, \ldots, v_k$ (azonos dimenziójú) vektorok által generált altér bázisa:

```
> basis({v1,...,vk});
```

vagy

```
> w:={v1,...,vk};
> t:=basis(w);
```

a bázis első eleme:

```
> t[1];
```

A $\{v_1, \ldots, v_k\}$ és a $\{w_1, \ldots, w_l\}$ bázissal rendelkező alterek összegének a bázisa:

```
> sumbasis({v1,...,vk},{w1,...,wl});
```

vagy

```
> s:={v1,...,vk};
> r:={w1,...,wl};
> t:=sumbasis(s,r);
```

a bázis első eleme:

```
> t[1];
```

A $\{v_1, \ldots, v_k\}$ és a $\{w_1, \ldots, w_l\}$ bázissal rendelkező alterek metszetének a bázisa:

```
> intbasis({v1,...,vk},{w1,...,wl});
```

vagy

```
> s:={v1,...,vk};
> r:={w1,...,wl};
> t:=intbasis(s,r);
```

a bázis első eleme:

```
> t[1];
```

Az A mátrix sorai által generált altér bázisa. Ha második paraméterként aposztrófok között megadunk még egy változót, akkor abban az altér dimenziója tárolódik:

```
> rowspace(A);
> rowspace(A,'dim');
```

a bázis tárolása, az elos" báziselem és a dimenzió kiíratása:

```
> s:=rowspace(A,'dim');
> s[1];
> dim;
```

Az A mátrix oszlopai által generált altér bázisa. Ha második paraméterként aposztrófok között megadunk még egy változót, akkor abban az altér dimenziója tárolódik:

```
> colspace(A);
> colspace(A,'dim');
> s:=colspace(A,'dim');
> s[1];
> dim;
```

Lineáris egyenletrendszerek

Az A négyzetes mátrix felső háromszög alakra hozása sorokkal végzett elemi átalakításokkal. Ha aposztrófok között második és harmadik paramétereket is megadunk, akkor az ott szereplő változókba a mátrix rangja illetve determinánsa tárolódik:

```
> gausselim(A,'rang','determinans');
> rang;
> determinans;
```

Az $Ax = b$ lineáris egyenletrendszer (paraméteres megoldása):

```
> linsolve(A,b);
```

Lineáris transzformációk

Az A négyzetes mátrix nulltere. A nulltér egy bázisa egy vektorhalmazban kerül tárolásra. Ha aposztrófok között még egy paramétert megadunk, abban a nulltér dimenziója tárolódik:

```
> kernel(A);
```

vagy

```
> s:=kernel(A,'dim');
> s[1];
> dim;
```

Az A (komplex) mátrix Jordan–féle normál alakja. Ha aposztrófok között egy második paramétet is megadunk, abban a bázistranszformáció P mátrixa kerül tárolásra, melyre fennáll $P^{-1}JP = A$.

```
> jordan(A);
> J:=jordan(A,'P');
```

$n \times n$ típusú x sajátértéket tartalmazó Jordan blokk előállítása, melyben a főátlóban x, felette 1, máshol 0 van:

```
> JordanBlock(x,n);
```

Remark. A mi tárgyalásunktól eltérően a MAPLE által használt Jordan–féle blokkok nem a főátló alatt, hanem fölötte tartalmaznak 1–eseket, és a Jordan–féle normálalak is ilyen típusú blokkokból épül fel. Belátható, hogy ez elvi eltérést nem jelent, a bázisvektorok megfelelő sorbarendezésével egyik alakból megkapható a másik.

Karakterisztikus polinom, sajátérték, sajátvektor

Az A négyzetes mátrix karakterisztikus mátrixa, x változóval, az $xE-A$ mátrix, ahol E az egységmátrix:

```
> charmat(A,x);
```

Remark. Mint látható, a MAPLE az általunk használt $A - xE$ mátrix helyett az $xE - A$ mátrixot használja. Ez azt eredményezi, az A mátrix karakterisztikus polinomja esetleg előjelben eltérhet az általunk használt polinomtól, de más (elvi) eltérés nincs.

Az A négyzetes mátrix karakterisztikus polinomja ($det(x * E - A)$). Ehhez természetesen eljuthatunk úgy is, hogy a karakterisztikus mátrixnak képezzük a determinánsát.

```
> charpoly(A,x);
```

vagy másképpen

```
> B:=charmat(A,x);
> d:=det(B);
```

vagy

```
> d:=det(charmat(A,x));
```

Az A négyzetes mátrix minimálpolinomja, az a legkisebb fokú polinom, melynek a mátrix gyöke (ez mindig a karakterisztikus polinom osztója):

```
> minpoly(A,x);
```

Az A négyzetes mátrix sajátértékeinek kiszámítása. A sajátértékek egy vektorban kerülnek tárolásra. Ehhez el lehet jutni a karakterisztikus polinom megoldása útján is:

```
> eigenvals(A);
> lambda:=eigenvals(A);
> lambda[1];
> evalc(lambda[1]);
> evalf(lambda[2]);
```

vagy másképpen

```
> f:=charpoly(A,x);
> fsolve(f,x);
> fsolve(f,x,complex);
```

Az A négyzetes mátrix sajátvektorainak kiszámítása. Az eredmény blokkok formájában jelenik meg, minden blokk tartalmazza a sajátértéket, annak multiplicitását, és a hozzátartozó sajátaltér egy bázisát. Ehhez a karakterisztikus mátrix nullterének kiszámítása útján is el lehet jutni.

```
> eigenvects(A);
```

vagy másképpen

```
> nullspace(charmat(A,ei));
```

ahol ei egy sajátértéke az A mátrixnak.

Euklideszi terek

Annak eldöntése, hogy az A mátrix ortogonális–e (eredmény logikai típusú, true vagy false):

```
> orthog(A);
```

3 dimenziós vektorok külső szorzatának kiszámítása:

```
> crossprod(a,b);
```

Azonos dimenziójú a, b vektorok belső (kompozíciós) szorzata:

```
> dotprod(a,b);
> innerprod(a,b);
```

Az x, y vektorok belső szorzata az A mátrixra vonatkozóan ($\sum_{i=1}^{n}\sum_{j=1}^{n} A_{ij}x_iy_j$):

```
> innerprod(x,A,y);
```

A $v_1, \ldots, v_k$ lineárisan független vektorrendszer ortogonalizálása, normálás nélkül a Gram–Schmidt–féle eljárással. A vektorok szeletei által generált alterekre szokásos feltételt a kiszámított vektorok általában más sorrendben elégítik ki.

```
> GramSchmidt({v1,...,vk});
```

vagy

```
> s:={v1,...,vk};
> u:=GramSchmidt(s);
> u[1];
```

Remark. Mint látható, a fenti eljárás csak ortogonalizálja a vektorokat, de nem normálja. Ortonormált bázis kiszámításához a vektorokat el kell osztani hosszukkal.

æ

Ajánlott irodalom

[1] *Bélteky Károly:* Analitikus geometria és lineáris algebra. Tankönyvkiadó, 1987.

[2] *D.K. Fagyejev – I.Sz. Szominszkij:* Felsőfokú algebrai feladatok. Műszaki Könyvkiadó, 1973.

[3] *Fried Ervin:* Klasszikus és lineáris algebra. Tankönyvkiadó, 1979.

[4] *I.M. Gelfand:* Előadások a lineáris algebrából. Akadémiai Kiadó, 1955.

[5] *Hajós György:* Bevezetés a geometriába. Tankönyvkiadó, 1972.

[6] *P.R. Halmos:* Véges dimenziós vektorterek. Műszaki Könyvkiadó, 1984.

[7] *Kovács Zoltán:* Feladatgyűjtemény lineáris algebra gyakorlatokhoz. Kossuth Egyetemi Kiadó, 1998.

[8] *A.G. Kuros:* Felsőbb algebra. Tankönyvkiadó, 1968.

[9] *Rózsa Pál:* Lineáris algebra és alkalmazásai. Műszaki Könyvkiadó, 1974.

Tárgymutató