Matematika 1 **Izročki: Teden 1**

Ponovitev: matrike, sled, rang, lastne vrednosti,...

Polona Oblak

1. Sled

Sled matrike $A = [a_{i,j}] \in \mathbb{R}^{n\times n}$ (oznaka: $\mathrm{tr}(A)$) je vsota vseh njenih diagonalnih elementov

$$\mathrm{tr}(A) = \sum_{i=1}^{n} a_{i,i}.$$

Lastnosti sledi. Za matrike $A, B, P \in \mathbb{R}^{n\times n}$ velja

(1) $\mathrm{tr}(\alpha A) = \alpha\, \mathrm{tr}(A)$,
(2) $\mathrm{tr}(A + B) = \mathrm{tr}(A) + \mathrm{tr}(B)$,
(3) $\mathrm{tr}(A^{\mathsf{T}}) = \mathrm{tr}(A)$,
(4) $\mathrm{tr}(AB) = \mathrm{tr}(BA)$,
(5) $\mathrm{tr}(PAP^{-1}) = \mathrm{tr}(A)$ za vsako obrnljivo matriko P.

2. Rang

Rang matrike $A \in \mathbb{R}^{m\times n}$ (oznaka: $\mathrm{rank}(A) = \mathrm{rk}(A)$) je

- število pivotov, ki jih dobimo v (reducirani) vrstično stopničasti obliki matrike A po Gaussovi eliminaciji
- število linearno neodvisnih vrstic matrike A,
- dimenzija linearne ogrinjače vrstic matrike A,
- število linearno neodvisnih stolpcev matrike A,
- dimenzija linearne ogrinjače stolpcev matrike A,
- $\dim C(A)$,
- $n - \dim N(A)$,
- velikost največje obrnljive (kvadratne) podmatrike matrike A.

3. Podobnost

Matriki $A \in \mathbb{R}^{n\times n}$ in $B \in \mathbb{R}^{n\times n}$ sta *podobni*, če obstaja takšna obrnljiva matrika $P \in \mathbb{R}^{n\times n}$, da je

$$A = PBP^{-1}.$$

Podobne matrike imajo isto:

(1) sled,
(2) determinanto,

(3) karakteristični polinom,
(4) lastne vrednosti,
(5) rang.

4. Do naslednjega tedna

(1) Pokažite naslednje enakosti in neenakosti, ki veljajo za rang matrik:
   (a) $\operatorname{rank}(A^{\mathsf{T}}) = \operatorname{rank}(A)$,
   (b) $\operatorname{rank}(AB) \leq \operatorname{rank}(A)$ in $\operatorname{rank}(AB) \leq \operatorname{rank}(B)$,
   (c) $\operatorname{rank}(A + B) \leq \operatorname{rank}([A|B])$,
   (d) $\operatorname{rank}(A + B) \leq \operatorname{rank}(A) + \operatorname{rank}(B)$,
   (e) $\operatorname{rank}([A|B]) \leq \operatorname{rank}(A) + \operatorname{rank}(B)$, kjer je za $A \in \mathbb{R}^{m\times n}$ in $A \in \mathbb{R}^{m\times p}$ matrika $[A|B] \in \mathbb{R}^{m\times(n+p)}$ razširjena matrika matrik A in B,
   (f) $\operatorname{rank}(A \oplus B) = \operatorname{rank}(A) + \operatorname{rank}(B)$, kjer je za $A \in \mathbb{R}^{m\times n}$ in $A \in \mathbb{R}^{p\times q}$ matrika $A \oplus B = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \in \mathbb{R}^{(m+p)\times(n+q)}$ direktna vsota matrik A in B.

   (Pri tem naj bosta matriki A in B v posameznih alinejah primernih velikosti za množenje, seštevanje,...)
(2) Rešite kviz in preverite vaše znanje osnovnih pojmov linearne algebre.

5. Nadaljnje branje

(1) David A. Harville: Matrix Algebra From a Statistian's Perspective, Springer, 1997, razdelek 5.
(2) 3Blue1Brown, Essence of linear algebra

6. (Priporočljiva) domača naloga

(1) David A. Harville: Matrix Algebra From a Statistian's Perspective, Springer, 1997, stran 53 (Exercise 1, 2, 3).
(2) Dokažite, da je podobnost matrik tranzitivna lastnost.
(3) Dokažite, da imajo podobne matrike isti karakteristični polinom.

Matematika 1	**Izročki, 2.teden**

Schurov izrek, Frobeniusova norma matrike, izrek Eckarta in Younga

Polona Oblak

1. Schurov izrek in njegove posledice

Izrek 1 (Schur). *Naj ima matrika $A \in \mathbb{R}^{n \times n}$ lastne vrednosti $\lambda_1, \lambda_2, \ldots, \lambda_n$. Tedaj obstaja takšna ortogonalna matrika $Q \in \mathbb{R}^{n \times n}$, da je*

$$Q^\mathsf{T} A Q$$

zgornje trikotna $n \times n$ matrika z diagonalnimi elementi enakimi $\lambda_1, \lambda_2, \ldots, \lambda_n$.

Opazimo:

- V izreku 1 lahko zgornjo trikotnost nadomestimo s spodnjo trikotnostjo.
- Pri tem niti matrika Q niti $Q^\mathsf{T} A Q$ nista enolično določeni z matriko A.

Posledica 1. *Vsaka matrika $A \in \mathbb{R}^{n \times n}$ je podobna zgornje trikotni matriki.*

Posledica 2. *Vsaka simetrična matrika $A \in \mathbb{R}^{n \times n}$ je ortogonalno podobna diagonalni matriki.*

Posledica 3. *Če ima matrika $A \in \mathbb{R}^{n \times n}$ lastne vrednosti enake $\lambda_1, \lambda_2, \ldots, \lambda_n$, potem je*

$$\mathrm{tr}(A) = \lambda_1 + \lambda_2 + \ldots + \lambda_n$$

in

$$\det(A) = \lambda_1 \lambda_2 \ldots \lambda_n.$$

Posledica 4 (Cayley-Hamilton). *Če je $\Delta_A(x) = \det(A - xI_n)$ karakteristični polinom matrike A, potem velja $\Delta_A(A) = 0$.*

Izrek 2 (SVD - Razcep singularnih vrednosti). *Za vsako matriko $A \in \mathbb{R}^{m \times n}$, $m \geq n$, obstajajo takšna matrika*

$$\Sigma = \begin{bmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_n \\ & & & \\ & & & \end{bmatrix} \in \mathbb{R}^{m \times n}$$

ter ortogonalni matriki $U \in \mathbb{R}^{m \times m}$ in $V \in \mathbb{R}^{n \times n}$, da velja

$$A = U \Sigma V^\mathsf{T}.$$

2. Frobeniusova norma matrike

Za matriki $A \in \mathbb{R}^{m\times n}$ in $B \in \mathbb{R}^{m\times n}$ definiramo

$$\langle A, B\rangle = \mathrm{tr}(A^T B).$$

Izrek 3. *Za produkt* $\langle A, B\rangle\colon \mathbb{R}^{m\times n} \times \mathbb{R}^{m\times n} \to \mathbb{R}$ *velja za vse matrike* $A, B, C \in \mathbb{R}^{m\times n}$

(1) $\langle A, B\rangle = \langle B, A\rangle$,
(2) $\langle \alpha A + \beta B, C\rangle = \alpha\langle A, C\rangle + \beta\langle B, C\rangle$, *za vse* $\alpha, \beta \in \mathbb{R}$,
(3) $\langle A, A\rangle \geq 0$,
(4) $\langle A, A\rangle = 0$ *natanko tedaj, ko je* $A = 0$.

Zato $\langle A, B\rangle$ *imenujemo* skalarni produkt *matrik* A *in* B.

Izrek 4. *Za matrike* $A \in \mathbb{R}^{m\times n}$, $B \in \mathbb{R}^{m\times k}$ *in* $C \in \mathbb{R}^{k\times n}$ *velja*

$$\langle A, BC\rangle = \left\langle B^{\mathsf{T}} A, C\right\rangle = \left\langle AC^{\mathsf{T}}, B\right\rangle.$$

Frobeniusova norma matrike $A = [a_{i,}] \in \mathbb{R}^{m\times n}$ je definirana kot

$$||A||_F = ||A|| = \sqrt{\langle A, A\rangle} = \sqrt{\mathrm{tr}(A^T A)}.$$

Velja:

$$||A||_F^2 = \sum_{i=1}^{n}\sum_{j=1}^{m} a_{ij}^2 = \sum_{i=1}^{\min\{m,n\}} \sigma_i^2.$$

Izrek 5 (Eckart, Young)**.** *Naj bo* $A = U\Sigma V^{\mathsf{T}}$ *razcep singularnih vrednosti matrike* $A \in \mathbb{R}^{m\times n}$, $m \geq n$, *kjer* $U = [U^{(1)}\, U^{(2)} \ldots U^{(m)}] \in \mathbb{R}^{m\times m}$ *in* $V = [V^{(1)}\, V^{(2)} \ldots V^{(n)}] \in \mathbb{R}^{n\times n}$. *Potem je matrika* $A_k \in \mathbb{R}^{m\times n}$ *ranga* k, $k \leq n$, *ki je med vsemi matrikami ranga* k *v Frobeniusovi normi najbližje matriki* A, *enaka*

$$A_k = \sigma_1 U^{(1)}(V^{(1)})^{\mathsf{T}} + \sigma_2 U^{(2)}(V^{(2)})^{\mathsf{T}} + \ldots + \sigma_k U^{(k)}(V^{(k)})^{\mathsf{T}}$$

in velja $||A - A_k||_F = \sqrt{\sigma_{k+1}^2 + \ldots + \sigma_n^2}$. *(Velja torej* $||A - A_k||_F \leq ||A - X||_F$ *za vse matrike* $X \in \mathbb{R}^{m\times n}$, *za katere velja* $\mathrm{rank}(X) = k$.*)*

3. Nadaljnje branje

(1) Roger A. Horn and Charles R. Johnson, *Matrix Analysis*, Cambridge, 2006, razdelka 2.3 and 2.4.
(2) več matričnih norm: Roger A. Horn and Charles R. Johnson, *Matrix Analysis*, Cambridge, 2006, razdelek 5.6.
(3) Gilbert Strang, *Linear algebra and Learning from Data*, Wellesley-Cambridge Press, Wellesley, 2019, razdelek I.9.

4. Domača naloga

ϟ (1) Rešite kviz na spletni Učilnici.

⋆ (2) Uporabite Schurov izrek, da dokažete, da je rang kvadratne matrike A enak velikosti največje obrnljive (kvadratne) podmatrike matrike A.

⋆ (3) Dokažite Cayley-Hamiltonov izrek, posledica 4.

ϟ (4) Pokažite, da za simetrično matriko $A \in \mathbb{R}^{n\times n}$ velja

$$||A||_F = \sqrt{\lambda_1^2 + \ldots + \lambda_n^2},$$

kjer so $\lambda_1, \ldots, \lambda_n$ lastne vrednosti matrike A. Pokažite na primeru, da enakost ne velja za nesimetrične matrike.

ϟ (5) Naj bo $A = PDP^{-1}$ simetrična matrika. Pokažite, da je najboljša aproksimacija ranga ena matrike A v Frobeniusovi normi enaka $\lambda_1 v_1 v_1^{\mathsf{T}}$, kjer je λ_1 po absolutni vrednosti največja lastna vrednost matrike A, v_1 pa normirani lastni vektor, ki pripada λ_1.

⋆ (6) Roger A. Horn and Charles R. Johnson, *Matrix Analysis*, Cambridge, 2006, nalogi 2.3.P6 (str. 107) in 2.4.P2 (str. 124).

(Naloge, označene s ϟ preverjajo razumevanje osnovnih pojmov in so primeri nalog s teoretičnih izpitov. Naloge, označene s ⋆ so težje, širijo vaše znanje in dopolnjujejo odpredavano snov.)

5. Rešitve nalog

(2) Naj ima matrika $A \in \mathbb{R}^{n\times n}$ lastne vrednosti $\lambda_1, \dots, \lambda_n$. Po Schurovem izreku obstajata takšna ortogonalna matrika $Q \in \mathbb{R}^{n\times n}$ in takšna zgornje trikotna matrika $Z \in \mathbb{R}^{n\times n}$ z diagonalnimi elementi $\lambda_1, \dots, \lambda_n$, da je $Q^\mathsf{T} A Q = Z$.

Matrika Z je zgornje trikotna, zato lahko s permutacijo vrstic (množenje Z s permutacijsko matriko $P \in \mathbb{R}^{n\times n}$ z leve) in s permutacijo stolpcev (množenje Z s permutacijsko matriko $R \in \mathbb{R}^{n\times n}$ z desne) dosežemo, da bo matrika PZR oblike

$$PZR = \begin{bmatrix} Z_1 & X \\ 0 & Z_2 \end{bmatrix},$$

kjer je $Z_1 \in \mathbb{R}^{r\times r}$ obrnljiva zgornje trikotna matrika, $Z_2 \in \mathbb{R}^{(n-r)\times(n-r)}$ pa zgornje trikotna matrika z ničlami na diagonali. To pomeni, da ima matrika PZR obrnljivo $r\times r$ podmatriko Z_1 in da velja $\operatorname{rank}(Z) \geq \operatorname{rank} Z_1 = r$. Še več, ker ima Z_2 na diagonali ničelne vrednosti, ima vsaka $(r+1)\times(r+1)$ podmatrika matrike Z ničelno determinanto.

Matrika $Z = P^\mathsf{T} A R^\mathsf{T}$ ima tako tudi obrnljivo $r \times r$ podmatriko (vendar morda ne več v zgornjem levem vogalu) in posledično jo ima tudi matrika $A = QZQ^\mathsf{T}$ (saj je Q obrnljiva). Ker sta matriki A in Z podobni, je torej $\operatorname{rank}(A) = \operatorname{rank}(Z) \geq r$. Zatorej je $\operatorname{rank}(A)$ enak velikosti največje obrnljive podmatrike matrike A (=največji možni r).

Matematika 1	**Izročki: 3. teden**

Kroneckerjev produkt

Polona Oblak

1. Kroneckerjev (tenzorski) produkt

Kroneckerjev produkt (tudi *tenzorski produkt*) matrik $A = [a_{ij}] \in \mathbb{R}^{m\times n}$ in $B \in \mathbb{R}^{p\times q}$ je $mp \times nq$ matrika

$$A \otimes B = \begin{bmatrix} a_{11}B & a_{12}B & \dots & a_{1n}B \\ a_{21}B & a_{22}B & \dots & a_{2n}B \\ \vdots & \vdots & & \vdots \\ a_{m1}B & a_{m2}B & \dots & a_{mn}B \end{bmatrix} \in \mathbb{R}^{mp\times nq}.$$

Če so matrike A, B, C in D primernih velikosti, potem veljajo naslednje enakosti.

(1) $0 \otimes A = A \otimes 0 = 0$
(2) $\alpha \otimes A = A \otimes \alpha = \alpha A$ za vsak $\alpha \in \mathbb{R}$
(3) $(\alpha A) \otimes B = A \otimes (\alpha B) = \alpha(A \otimes B)$
(4) $(A + B) \otimes C = A \otimes C + B \otimes C$ in $A \otimes (B + C) = A \otimes B + A \otimes C$
(5) $(A \otimes B)^{\mathsf{T}} = A^{\mathsf{T}} \otimes B^{\mathsf{T}}$
(6) $(A \otimes B) \otimes C = A \otimes (B \otimes C)$.
(7) $(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$.
(8) $(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$
(9) $\operatorname{tr}(A \otimes B) = \operatorname{tr}(A)\operatorname{tr}(B)$
(10) $\operatorname{rang}(A \otimes B) = \operatorname{rang}(A)\operatorname{rang}(B)$
(11) Če ima matrika $A \in \mathbb{R}^{n\times n}$ lastne vrednosti $\lambda_1, \dots, \lambda_m$ in ima matrika B lastne vrednosti $\mu_1, \dots, \mu_n$, potem je množica lastnih vrednosti matrike $A \otimes B$ enaka
$$\{\lambda_i\mu_j;\ \lambda_i \text{ lastna vrednost } A, \mu_j \text{ lastna vrednost } B\}.$$
(12) Če $A \in \mathbb{R}^{n\times n}$ in $B \in \mathbb{R}^{m\times m}$, potem je $\det(A \otimes B) = (\det A)^m(\det B)^n$.

2. Preslikava vec

Za matriko $A \in \mathbb{R}^{m\times n}$ označimo *vektorizacijo matrike A* kot

$$\operatorname{vec}(A) = \begin{bmatrix} A^{(1)} \\ A^{(2)} \\ \vdots \\ A^{(n)} \end{bmatrix} \in \mathbb{R}^{mn}.$$

vec je preslikava iz $\mathbb{R}^{m\times n}$ v $\mathbb{R}^{mn}$.

Izrek 1. *Za matrike* $A \in \mathbb{R}^{m\times n}$, $B \in \mathbb{R}^{n\times p}$ *in* $C \in \mathbb{R}^{p\times r}$ *velja*

$$\operatorname{vec}(ABC) = (C^{\mathsf{T}} \otimes A) \operatorname{vec}(B).$$

3. Nadaljnje branje

(1) David A. Harville: Matrix Algebra From a Statistian's Perspective, Springer, 1997, razdelki 16.1., 16.2. in 16.3.
(2) Roger A. Horn and Charles R. Johnson, *Topics in Matrix Analysis*, Cambridge, 2006, razdelek 4.2.
(3) Charles F. Van Loan:The ubiquitous Kronecker product, Journal of Computational and Applied Mathematics 123 (2000) 85-100.

4. Domača naloga

(1) Rešite kviz na spletni Učilnici.
(2) David A. Harville: Matrix Algebra From a Statistian's Perspective, Springer, 1997, stran 368 (Exercises 1, 2, 3, 4, 14, 15, 16).

Matematika 1	**Izročki: 4. teden**

Pozitivno semidefinitne matrike. Razcep Choleskega.

Polona Oblak

1. Pozitivno semidefinitne matrike

Spomnimo se, da ima simetrična matrika $A \in \mathbb{R}^{n\times n}$ vse lastne vrednosti realne.

Simetrični matriki $A \in \mathbb{R}^{n\times n}$ pravimo

(1) *pozitivno semidefinitna*, če je $x^{\mathsf{T}}Ax \geq 0$ za vse $x \in \mathbb{R}^n$.
(2) *pozitivno definitna*, če je $x^{\mathsf{T}}Ax > 0$ za vse neničelne $x \in \mathbb{R}^n$.
(3) *negativno semidefinitna*, če je $x^{\mathsf{T}}Ax \leq 0$ za vse $x \in \mathbb{R}^n$.
(4) *negativno definitna*, če je $x^{\mathsf{T}}Ax < 0$ za vse neničelne $x \in \mathbb{R}^n$.
(5) *nedefinitna*, če je $x^{\mathsf{T}}Ax > 0$ za nekatere $x \in \mathbb{R}^n$ in $y^{\mathsf{T}}Ay < 0$ za nekatere $y \in \mathbb{R}^n$.

Izrek 1. *Simetrična matrika $A \in \mathbb{R}^{n\times n}$ je*

(1) *pozitivno semidefinitna natanko tedaj, ko ima vse lastne vrednosti nenegativne,*
(2) *pozitivno definitna natanko tedaj, ko ima vse lastne vrednosti pozitivne,*
(3) *negativno semidefinitna natanko tedaj, ko ima vse lastne vrednosti nepozitivne,*
(4) *negativno definitna natanko tedaj, ko ima vse lastne vrednosti negativne,*
(5) *nedefinitna natanko tedaj, ko ima tako pozitivne kot negativne lastne vrednosti.*

Izrek 2. (1) *Simetrična matrika $A \in \mathbb{R}^{n\times n}$, $\mathrm{rank}(A) = r$, je pozitivno semidefinitna natanko tedaj, ko obstaja takšna matrika $B \in \mathbb{R}^{n\times r}$, $\mathrm{rank}(B) = r$, da je $A = BB^{\mathsf{T}}$.*
(2) *Simetrična matrika $A \in \mathbb{R}^{n\times n}$ je pozitivno definitna natanko tedaj, ko obstaja takšna obrnljiva matrika $B \in \mathbb{R}^{n\times n}$, da je $A = BB^{\mathsf{T}}$.*

Izrek 3 (Sylvester)**.** *Simetrična matrika A je pozitivno definitna natanko tedaj, ko so determinante vseh vodilnih glavnih podmatrik matrike A pozitivne.*

Simetrična matrika A je negativno definitna natanko tedaj, ko je determinanta vsake $k \times k$ vodilne glavne podmatrike A pozitivna, če je k sodo število, ter negativna, če je k liho število.

Izrek 4 (Razcep Choleskega)**.** *Obrnljiva matrika $A \in \mathbb{R}^{n\times n}$ ima razcep Choleskega*

$$A = LL^{\mathsf{T}},$$

kjer je $L \in \mathbb{R}^{n\times n}$ spodnje trikotna matrika, natanko tedaj, ko je A simetrična in pozitivno definitna.

2. Nadaljnje branje

ϟ (1) David A. Harville: Matrix Algebra From a Statistian's Perspective, Springer, 1997, Poglavje 14.

ϟ (2) Gilbert Strang, *Linear algebra and Learning from Data*, Wellesley-Cambridge Press, Wellesley, 2019, razdelek I.7.

⋆ (3) Roger A. Horn and Charles R. Johnson, *Matrix Analysis*, Cambridge, 2006, Podpoglavja 7.0, 7.1, 7.2.

⋆ (4) Giorgio Giorgi: Various Proofs of the Sylvester Criterion for Quadratic Forms, Journal of Mathematics Research; Vol 9, No 6, 2017.

3. Domača naloga

ϟ (1) Rešite kviz na spletni Učilnici.

ϟ (2) Poiščite linearno spremembo spremenljivk $u = f(x, y, z)$, $v = g(x, y, z)$ and $w = h(x, y, z)$, v kateri ima

$$Q(x, y, z) = 3x^2 + 4xy - 4xz - 2yz$$

diagonalno obliko.

ϟ (3) Če sta matriki $A \in \mathbb{R}^{n\times n}$ in $B \in \mathbb{R}^{m\times m}$ simetrični pozitivno semidefinitni matriki, pokažite, da je tudi $A \otimes B$ pozitivno semidefinitna matrika.

ϟ (4) Če je simetrična matrika A pozitivno semidefinitna, pokažite, da je tudi A^k pozitivno semidefinitna za vsak $k \in \mathbb{N}$.

ϟ (5) Naj bo $A = B^\mathsf{T} B$ takšna pozitivno semidefinitna matrika velikosti $n \times n$, da za nek vektor $x \in \mathbb{R}^n$ velja $x^\mathsf{T} A x = 0$. Pokažite, da je $Bx = 0$ in nato sklepajte, da velja tudi $Ax = 0$.

ϟ (6) Pokažite, da je bločna matrika $\begin{bmatrix} A & B \\ B^\mathsf{T} & C \end{bmatrix}$ (simetrična) pozitivno semidefinitna natanko tedaj, ko je matrika $\begin{bmatrix} A & -B \\ -B^\mathsf{T} & C \end{bmatrix}$ pozitivno semidefinitna. Pri tem je $A \in \mathbb{R}^{n\times n}, B \in \mathbb{R}^{n\times m}$ ter $C \in \mathbb{R}^{m\times m}$.

ϟ (7) Naj bo $A \in \mathbb{R}^{n\times n}$ simetrična pozitivno semidefinitna matrika z lastnimi vrednostmi $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$.
   (a) Dokažite, da velja $\vec{x}^\mathsf{T} A \vec{x} \leq \lambda_1 ||\vec{x}||^2$ za vsak $\vec{x} \in \mathbb{R}^n$.
   (b) Dokažite, da velja $\max_{\vec{x}\in\mathbb{R}^n} \frac{\vec{x}^\mathsf{T} A \vec{x}}{\vec{x}^\mathsf{T} \vec{x}} = \lambda_1$.
   (c) Dokažite, da je tudi matrika $A - \lambda_1 I_n$ pozitivno semidefinitna.

(8) Naj bosta $A, B \in \mathbb{R}^{n\times n}$ simetrični pozitivno semidefinitni matriki.

ϟ (a) Pokažite, da obstaja natanko pozitivno semidefinitna matrika K, da velja $K^2 = A$.

⋆ (b) Pokažite, da obstaja natanko ena pozitivno semidefinitna matrika K, da velja $K^2 = A$. Pravimo tudi, da je $K = \sqrt{A}$ koren matrike A.

ϟ (c) Pokažite, da je $\langle A, B\rangle = ||\sqrt{A}\sqrt{B}||_F^2$.

⋆ (9) Preberite in razumite vsaj enega od dokazov Sylvestrovega izreka.

(Naloge, označene s ϟ preverjajo razumevanje osnovnih pojmov in so primeri nalog s teoretičnih izpitov. Naloge, označene s ⋆ so težje, širijo vaše znanje in dopolnjujejo odpredavano snov.)

Matematika 1	Izročki: 5. teden

Vektorski prostor in podprostor. Baza.

Polona Oblak

1. Vektorski prostor

Realni vektorski prostor V je množica *vektorjev* $v \in V$, za katere imamo definirani dve notranji operaciji

- seštevanje vektorjev ($u, v \in V \Longrightarrow u+v \in V$) in
- množenje vektorjev z realnimi števili ($v \in V$, $\alpha \in \mathbb{R} \Longrightarrow \alpha v = \alpha \cdot v \in V$),

z lastnostmi

(VP1) $u+v = v+u$ in $(u+v)+w = u+(v+w)$,
(VP2) obstaja *ničelni vektor* $\mathbf{0}$ in velja $v+\mathbf{0} = \mathbf{0}+v = v$,
(VP3) za vsak $v \in V$ obstaja *nasprotni vektor* $-v$, za katerega velja $v+(-v) = (-v)+v = \mathbf{0}$,
(VP4) $1 \cdot v = v$ za vsak $v \in V$,
(VP5) $(\alpha\beta) \cdot v = \alpha \cdot (\beta \cdot v)$,
(VP6) $(\alpha+\beta) \cdot v = \alpha \cdot v + \beta \cdot v$,
(VP7) $\alpha \cdot (u+v) = \alpha \cdot u + \alpha \cdot v$,

za poljubne $u, v, w \in V$ in $\alpha, \beta \in \mathbb{R}$.

Pri tem bomo, kot običajno v matematiki, simbol $\cdot$ pri produktu vektorja s številom večinoma izpuščali. Torej bomo pisali tudi $\alpha v = \alpha \cdot v$.

Izrek 1. *Naj bo V vektorski prostor. Potem velja*

(1) *V vsebuje ničelni vektor $\mathbf{0}$,*
(2) *v vsakem vektorskem prostoru V je ničelni vektor $\mathbf{0}$ en sam,*
(3) *$\alpha \cdot \mathbf{0} = \mathbf{0}$ za vsak $\alpha \in \mathbb{R}$,*
(4) *$0 \cdot v = \mathbf{0}$ za vsak $v \in V$.*

Za vektorje $v_1, v_2, \ldots, v_n \in V$ in skalarje $\alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbb{R}$ imenujemo vektor

$$\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n$$

linearna kombinacija vektorjev $v_1, v_2, \ldots, v_n$.

Denimo, ničelni vektor $\mathbf{0} = 0 \cdot v_1 + 0 \cdot v_2 + \ldots + 0 \cdot v_n$ je linearna kombinacija poljubnih vektorjev $v_1, v_2, \ldots, v_n \in V$. Linearno kombinacijo z izključno ničelnimi koeficienti imenujemo *trivialna* linearna kombinacija.

2. VEKTORSKI PODPROSTOR

Če je podmnožica U vektorskega prostora V

(VPP1) zaprta za seštevanje ($u, v \in U \Longrightarrow u + v \in U$) in

(VPP2) zaprta za množenje vektorjev z realnimi števili ($v \in U$, $\alpha \in \mathbb{R} \Longrightarrow \alpha v \in U$),

potem jo imenujemo *vektorski podprostor* prostora V.

Izrek 2. *Podmnožica U vektorskega prostora V je vektorski podprostor natanko tedaj, ko je poljubna linearna kombinacija $\alpha u + \beta v$ vektorjev $u, v \in U$ tudi vsebovana v U.*

Vsak vektorski podprostor po (VPP2) vsebuje tudi vektor $0 \cdot v = \mathbf{0}$. Zatorej podmnožica vektorskega prostora, ki ne vsebuje ničelnega vektorja, ne more biti vektorski podprostor.

Ker lastnosti (VP1)-(VP7) veljajo za poljubne elemente vektorskega prostora V, veljajo tudi za vse elemente vektorskega podprostora U v V. Poleg tega je vektorski podprostor po definiciji zaprt za seštevanje in množenje s števili. Zatorej je vsak vektorski podprostor hkrati tudi vektorski prostor.

Linearna ogrinjača $\mathcal{L}\{v_1, v_2, \ldots, v_n\}$ vektorjev $v_1, v_2, \ldots, v_n$ je množica vseh linearnih kombinacij vektorjev $v_1, v_2, \ldots, v_n$.

Ker je linearna kombinacija linearnih kombinacij vektorjev $v_1, v_2, \ldots, v_n \in V$ zopet linearna kombinacija vektorjev $v_1, v_2, \ldots, v_n$, je po Izreku 2 linearna ogrinjača $\mathcal{L}\{v_1, v_2, \ldots, v_n\}$ linearni podprostor v V. Pravimo, da vektorji $v_1, v_2, \ldots, v_n$ *napenjajo* prostor $\mathcal{L}\{v_1, v_2, \ldots, v_n\}$.

Ne le, da je linearna ogrinjača vektorski prostor. Velja celo več.

Izrek 3. *Linearna ogrinjača vektorjev $v_1, v_2, \ldots, v_n$ vektorskega prostora V je najmanjši vektorski podprostor v V, ki vsebuje vektorje $v_1, v_2, \ldots, v_n$.*

3. BAZA VEKTORSKEGA PROSTORA

Vektorji $v_1, v_2, \ldots, v_n \in V$ so *linearno odvisni*, ko obstaja vektor v_k, ki je linearna kombinacija ostalih $v_1, v_2 \ldots, v_{k-1}, v_{k+1}, \ldots, v_n$:

$$v_k = \alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_{k-1} v_{k-1} + \alpha_{k+1} v_{k+1} + \ldots + \alpha_n v_n,$$

kjer $\alpha_i \in \mathbb{R}$.

Vektorji $v_1, v_2, \ldots, v_n \in V$ so *linearno neodvisni*, če niso linearno odvisni. Ekvivalentno, $v_1, v_2, \ldots, v_n \in V$ so linearno neodvisni, če je njihova trivialna linearna kombinacija edina njihova linearna kombinacija, ki je enaka ničelnemu vektorju $\mathbf{0}$. Z drugimi besedami, $v_1, v_2, \ldots, v_n \in V$ so linearno neodvisni, če iz

$$\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n = \mathbf{0}$$

sledi

$$\alpha_1 = \alpha_2 = \ldots = \alpha_n = 0.$$

Če množica vektorjev vsebuje ničelni vektor $\mathbf{0}$, potem lahko ničelni vektor izrazimo kot trivialno linearno kombinacijo ostalih vektorjev. Zato je vsaka množica, ki vsebuje ničelni vektor, linearno odvisna.

Naj vektorji $u_1, u_2, \ldots, u_m$ napenjajo vektorski prostor $V = \mathcal{L}\{u_1, u_2, \ldots, u_m\}$. Če so $u_1, u_2, \ldots, u_m$ linearno odvisni, potem obstaja podmnožica $\{v_1, v_2, \ldots, v_n\} \subseteq \{u_1, u_2, \ldots, u_m\}$, ki prav tako napenja prostor V. Najmanjšo takšno podmnožico bomo imenovali *baza* vektorskega prostora. Dobimo jo tako, da iz množice $\{u_1, u_2, \ldots, u_m\}$ izberemo čim manj vektorjev (torej bodo linearno neodvisni), a bodo še vedno napenjali prostor V.

Množica vektorjev $\{v_1, v_2, \ldots, v_n\}$ je *baza* vektorskega prostora V, če

(B1) so $v_1, v_2, \ldots, v_n$ linearno neodvisni in
(B2) $v_1, v_2, \ldots, v_n$ napenjajo prostor V.

Izrek 4. *Vsak vektorski prostor ima nešteto baz. Vse baze vektorskega prostora imajo enako število vektorjev.*

Dimenzija prostora V je enaka moči (poljubne) baze prostora V. Označimo jo z $\dim V$.

Izrek 5. *Za vsako bazo vektorskega prostora V je zapis poljubnega vektorja $v \in V$ kot linearna kombinacija baznih vektorjev vedno enoličen.*

4. Nadaljnje branje

- (1) David Poole, Linear algebra, a modern introduction, Thomson, 2006, Podpoglavji 6.1 in 6.2.
- (2) Tomaž Košir, Linearna algebra, poglavje VI: Vektorski prostori.

5. Domača naloga

- (1) Rešite kviz na Učilnici.
- (2) Drži ali ne drži?
  - (a) Množica vseh zgornje trikotnih $n \times n$ matrik vektorski podprostor v $\mathbb{R}^{n\times n}$.
  - (b) Množica vseh 3×3 matrik z vsemi diagonalnimi elementi enakimi 0 je vektorski podprostor v $\mathbb{R}^{3\times 3}$.
  - (c) Množica vseh 4×4 matrik, ki imajo vse vrstice enake, je vektorski podprostor v $\mathbb{R}^{4\times 4}$.
  - (d) Ravnina v $\mathbb{R}^3$, podana z enačbo $x + 2y + 3z = 4$, je vektorski podprostor v $\mathbb{R}^3$.
  - (e) Če so $x, y, z \in \mathbb{R}^3$ linearno odvisno vektorji, potem je linearna ogrinjača $\mathcal{L}\{x, y, z\}$ ravnina v $\mathbb{R}^3$ skozi koordinatno izhodišče.

(f) Če je $\{v_1, v_2, \ldots, v_7\}$ ortogonalna množica v vektorskem prostoru V dimenzije 7 in v_i neničelni vektorji, potem je $v_1, v_2, \ldots, v_7\}$ baza prostora V.
(g) Vsaka linearno neodvisna množica vektorjev v vektorskem prostoru dimenzije 9 vsebuje vsaj 9 elementov.
(h) Vsaka baza prostora $\mathbb{R}^{2\times 4}$ ima največ 4 elemente.
(i) Če je U linearna ogrinjača vektorjev $v_1, v_2, \ldots, v_k$, potem vektorji $v_1, v_2, \ldots, v_k$ tvorijo bazo prostora U.

ϟ (3) Katere od naslednjih množic realnih $n \times n$ matrik so vektorski podprostori v $\mathbb{R}^{n\times n}$? Za vsak podprostor določite tudi bazo.
(a) Matrike, ki imajo prvo vrstico ničelno.
(b) Matrike, ki imajo vsoto elementov v vsaki vrstici enako 1.
(c) Vse matrike C, za katere velja $C^2 = I$.
(d) Vse matrike D, ki so rešitve sistema $D\vec{x} = 0$. Bazo določite le v posebnem primeru, ko je $n = 2$ in $\vec{x} = [1\ 2]^{\mathsf{T}}$.
(e) Vse matrike, katerih elementi so nenegativna realna števila.
(f) Vse matrike F, za katere velja $F = F^T$.
(g) Vse matrike G, za katere velja $G = -G^T$.
(h) Vse matrike H, za katere velja $\operatorname{rank} H = n$.
(i) Vse matrike, katerih vse vrstice so med seboj enake.
(j) Vse matrike X, katerih produkt z vnaprej dano matriko J je enak ničelni matriki. Bazo določite le v posebnem primeru, ko je $n = 2$ in $\vec{J} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$.

ϟ (4) Naj bo V vektorski prostor ter $U, W \subseteq V$ vektorska prostora v V. Pokažite, da je tudi $U \cap W$ vektorski prodprostor v V.

ϟϟ (5) Pokažite Izrek 1.

ϟ (6) Naj $\mathcal{C}[0, 2\pi]$ označuje vektorski prostor vseh zveznih funkcij na intervalu $[0, 1]$. Pokažite, da so vektorji

$$f(x) = 1,\ g(x) = \cos(2x),\ h(x) = \cos^2 x,$$

linearno odvisni v $\mathcal{C}[0, 1]$.

ϟϟ (7) Naj bo $\mathcal{V}$ množica vseh simetričnih matrik oblike

$$\begin{bmatrix} A & S \\ -S & A \end{bmatrix},$$

kjer je $A \in \mathbb{R}^{n\times n}$ simetrična matrika ($A^{\mathsf{T}} = A$), $S \in \mathbb{R}^{n\times n}$ pa poševno simetrična ($S^{\mathsf{T}} = -S$).
(a) Pokažite so, da je $\mathcal{V}$ vektorski prostor.
(b) Poiščite bazo prostora $\mathcal{V}$ in določite $\dim \mathcal{V}$.
(c) Dokažite, da je karakteristični polinom vsake matrike v $\mathcal{V}$ kvadrat.

ϟ (8) David Poole, Linear algebra, a modern introduction, Thomson, 2006, stran 446 (Exercises 24-50), strani 460-463 (Exercises 1-58).

(Naloge, označene s $\lightning$ preverjajo razumevanje osnovnih pojmov in so primeri nalog s teoretičnih izpitov. Naloga, označena s $\lightning\lightning$ je nekolika težja.)

Matematika 1 **Izročki: 6. teden**

Linearne preslikave

Polona Oblak

1. Definicija

Naj bosta V in U vektorska prostora. Preslikava $\tau\colon V \to U$ je *linearna preslikava*, če velja

(LP1) $\tau(v+u) = \tau(v) + \tau(u)$ za vsaka $v, u \in V$ in
(LP2) $\tau(\alpha v) = \alpha\tau(v)$ za vsak $v \in V$ in vsak $\alpha \in \mathbb{R}$.

Izrek 1. *Preslikava $\tau\colon V \to U$ je linearna natanko tedaj, ko velja*

$$\tau(\alpha v + \beta u) = \alpha\tau(v) + \beta\tau(u) \tag{1}$$

za vse $v, u \in V$ ter vse $\alpha, \beta \in \mathbb{R}$.

Trditev 1. *Za poljubno linearno preslikavo $\tau\colon V \to U$ velja $\tau(\mathbf{0}_V) = \mathbf{0}_U$.*

2. Operacije z linearnimi preslikavami

Naj bodo $\tau, \psi\colon V \to U$ ter $\theta\colon U \to W$ linearne preslikave in naj bo $\gamma \in \mathbb{R}$.

(1) *Vsota $\tau + \psi\colon V \to U$* je preslikava, definirana s predpisom
$$(\tau + \psi)(v) = \tau(v) + \psi(v).$$
(2) *Produkt s skalarjem $\gamma\tau\colon V \to U$* je preslikava, definirana s predpisom
$$(\gamma\tau)(v) = \gamma\,\tau(v).$$
(3) *Kompozitum $\theta \circ \tau$* je preslikava $\theta \circ \tau\colon V \to W$ definirana s predpisom
$$(\theta \circ \tau)(v) = \theta(\tau(v)).$$

Izrek 2. *Vsota, produkt s skalarjem in kompozitum linearnih preslikav so linearne preslikave.*

Posledica 1. *Množica vseh linearnih preslikav iz vektorskega prostora V v vektorski prostor U je vektorski prostor.*

3. Matrike linearne preslikave

Naj bosta V in U vektorska prostora dimenzij m in n ter $\tau\colon V \to U$ linearna preslikava. Izberimo bazi $\mathcal{B} = \{b_1, b_2, \ldots, b_m\}$ prostora V in $\mathcal{C} = \{c_1, c_2, \ldots, c_n\}$ prostora U.

Denimo, da poznamo slike $\tau(b_1), \tau(b_2), \ldots, \tau(b_m)$ vektorjev baze $\mathcal{B}$ preko preslikave τ. Izberimo poljubni vektor $v \in V$ in ga zapišimo kot linearno kombinacijo

$$v = \beta_1 b_1 + \beta_2 b_2 + \ldots + \beta_m b_m$$

baznih vektorjev množice $\mathcal{B}$. Ker je τ linearna preslikava, lahko nato sliko $\tau(v)$ vektorja izračunamo kot

$$\tau(v) = \tau\left(\beta_1 b_1 + \beta_2 b_2 + \ldots + \beta_m b_m\right) = \beta_1 \tau(b_1) + \beta_2 \tau(b_2) + \ldots + \beta_m \tau(b_m).$$

Tako smo ugotovili, da lahko s pomočjo slik baznih vektorjev izračunamo sliko poljubnega vektorja $v \in V$.

Sedaj pa razvijmo slike vektorjev baze $\mathcal{B}$, ki se nahajajo v vektorskem prostoru U, po bazi $\mathcal{C}$. Torej, za $j = 1, \ldots, m$, vsako sliko $\tau(b_j)$ zapišimo kot linearno kombinacijo vektorjev množice $\mathcal{C} = \{c_1, c_2, \ldots, c_n\}$:

$$\tau(b_j) = \alpha_{1j} c_1 + \alpha_{2j} c_2 + \ldots + \alpha_{nj} c_n.$$

S tem smo dobili $m \cdot n$ enolično določnih koeficientov α_{ij}, $i = 1, 2, \ldots, n$, $j = 1, \ldots, m$. Naj bo $A_{\tau,\mathcal{B},\mathcal{C}} = [\alpha_{ij}]$ matrika reda $n \times m$ sestavljena iz koeficientov v razvoju slik baznih vektorjev $\mathcal{B}$ po bazi $\mathcal{C}$.

Tako definirano matriko

$$A_{\tau,\mathcal{B},\mathcal{C}} = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \ldots & \alpha_{1m} \\ \alpha_{21} & \alpha_{22} & \ldots & \alpha_{2m} \\ \vdots & \vdots & & \vdots \\ \alpha_{n1} & \alpha_{n2} & \ldots & \alpha_{nm} \end{bmatrix}$$

imenujemo *matrika linearne preslikave τ iz baze $\mathcal{B}$ v bazo $\mathcal{C}$*.

V njej j-ti stolpec predstavlja koeficiente, s katerimi se slika j-tega vektorja baze $\mathcal{B}$ izraža kot linearna kombinacija vektorjev baze $\mathcal{C}$.

Matrika $A_{\tau,\mathcal{B},\mathcal{C}}$ nam torej preslika koeficiente v razvoju vektorja po bazi $\mathcal{B}$ v koeficiente, če sliko vektorja razvijemo po bazi $\mathcal{C}$. Povedano formalno, če je $v = \sum_{j=1}^{m} \beta_j b_j$ in $\tau(v) = \sum_{i=1}^{n} \gamma_i c_i$, potem lahko koeficiente γ_j dobimo tudi z matričnim množenjem

$$\begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \\ \vdots \\ \\ \gamma_n \end{bmatrix} = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \ldots & \alpha_{1m} \\ \alpha_{21} & \alpha_{22} & \ldots & \alpha_{2m} \\ \\ \vdots & \vdots & & \vdots \\ \\ \alpha_{n1} & \alpha_{n2} & \ldots & \alpha_{nm} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_m \end{bmatrix}. \tag{2}$$

Posebej opozorimo, da je matrika $A_{\tau,\mathcal{B},\mathcal{C}}$ odvisna od izbire baz $\mathcal{B}$ in $\mathcal{C}$ vektorskih prostorov V in U. Če bi si izbrali drugačni bazi (ali vsaj eno od njiju), potem bi isti preslikavi priredili drugo matriko (ki ustreza koeficientom v razvoju po drugih bazah).

Dogovorimo se, da če bomo iskali matriko preslikave $\tau\colon V \to V$ iz baze $\mathcal{B}$ v bazo $\mathcal{B}$, bomo pri tem opustili enega od indeksov, t.j. $A_{\tau,\mathcal{B}} = A_{\tau,\mathcal{B},\mathcal{B}}$.

Izrek 3. *Naj bodo $\tau, \psi\colon V \to U$ ter $\theta\colon U \to W$ linearne preslikave in naj bo $\alpha \in \mathbb{R}$.*

(1) *Matrika, ki ustreza vsoti preslikav $\tau + \psi$, je enaka vsoti matrik posameznih preslikav.*

$$A_{\tau+\psi,\mathcal{B},\mathcal{C}} = A_{\tau,\mathcal{B},\mathcal{C}} + A_{\psi,\mathcal{B},\mathcal{C}}$$

(2) *Matrika, ki ustreza produktu s skalarjem $\alpha\tau$, je enaka večkratniku matrike preslikave.*

$$A_{\alpha\tau,\mathcal{B},\mathcal{C}} = \alpha A_{\tau,\mathcal{B},\mathcal{C}}$$

(3) *Matrika, ki ustreza kompozitumu preslikav, je enaka produktu matrik posameznih preslikav.*

$$A_{\theta\circ\tau,\mathcal{B},\mathcal{D}} = A_{\psi,\mathcal{C},\mathcal{D}} \cdot A_{\theta,\mathcal{B},\mathcal{C}}$$

(4) *Matrika, ki ustreza inverzu obrnljive preslikave, je enaka inverzu matrike te preslikave. Torej, če je ψ obrnljiva preslikava, je obrnljiva tudi matrika $A_{\psi,\mathcal{B},\mathcal{C}}$. Velja*

$$A_{\psi^{-1},\mathcal{C},\mathcal{B}} = \left(A_{\psi,\mathcal{B},\mathcal{C}}\right)^{-1}.$$

Denimo, da poznamo matriko $A = A_{\tau,\mathcal{B},\mathcal{C}}$ linearne preslikave $\tau\colon V \to U$ iz baze $\mathcal{B}$ v $\mathcal{C}$. Radi bi zapisali matriko $A' = A_{\tau,\mathcal{B}',\mathcal{C}'}$ iste linearne preslikave τ, a iz neke (morda) druge baze $\mathcal{B}'$ v (morda) drugo bazo $\mathcal{C}'$.

Če si ogledamo spodnji diagram matrik, ki utrezajo preslikavam, potem poznamo "zeleno" matriko $A = A_{\tau,\mathcal{B},\mathcal{C}}$, želimo pa izračunati "rdečo" matriko $A' = A_{\tau,\mathcal{B}',\mathcal{C}'}$ linearne preslikave τ.

$$\begin{tikzcd}
(V, \mathcal{B}) \arrow[r, "A"] \arrow[d, shift left, "P"] & (U, \mathcal{C}) \arrow[d, "Q"'] \\
(V, \mathcal{B}') \arrow[r, "A' = QAP^{-1}"'] \arrow[u, shift left, "P^{-1}"] & (U, \mathcal{C}')
\end{tikzcd}$$

Pri tem matrika $P = \mathcal{I}_{V,\mathcal{B},\mathcal{B}'}$ ustreza identični preslikavi prostora V vase iz baze $\mathcal{B}$ v bazo $\mathcal{B}'$. Takšno matriko dobimo s koeficienti pri razvoju vektorjev baze $\mathcal{B}$ po vektorjih baze $\mathcal{B}'$. Podobno je $Q = \mathcal{I}_{U,\mathcal{C},\mathcal{C}'}$ ustreza identični preslikavi prostora U vase iz baze $\mathcal{C}$ v bazo $\mathcal{C}'$ in je sestavljena iz koeficientov pri razvoju vektorjev baze $\mathcal{C}$ po vektorjih baze $\mathcal{C}'$. Če se v diagramu sprehodimo po puščicah, vidimo, da bomo morali matriko A' sestaviti kot kompozitum preslikav. Zatorej je A' enak produktu pripadajočih matrik, t.j.

$$A' = QAP^{-1}.$$

4. Lastne vrednosti linearne preslikave

Neničelnemu vektorju $v \in V$ pravimo *lastni vektor* linearne preslikave $\tau\colon V \to V$, če velja

$$\tau(v) = \lambda v.$$

Številu λ pravimo *lastna vrednost* linearne preslikave τ.

Izrek 4. *Vsaka lastna vrednost linearne preslikave τ je tudi lastna vrednost poljubne matrike A_τ, ki pripada preslikavi τ. Vse matrike, ki pripadajo dani linearni preslikavi τ imajo enake lastne vrednosti.*

Pravimo, da je linearno preslikavo $\tau\colon V \to V$ mogoče *diagonalizirati*, če obstaja baza, v kateri pripada preslikavi τ diagonalna matrika.

Izrek 5. *Linearno preslikavo $\tau\colon V \to V$ je mogoče diagonalizirati natanko tedaj, ko obstaja baza prostora V sestavljena iz lastnih vektorjev preslikave τ.*

5. Jedro in slika linearne preslikave

Naj bo $\tau\colon V \to U$ linearna preslikava vektorskega prostora V v vektorski prostor U.

Definicija 1. Jedro *linearne preslikave τ je množica $\ker(\tau)$ vseh vektorjev $v \in V$, za katere velja*

$$\tau(v) = 0.$$

Definicija 2. Slika *linearne preslikave je množica $\mathrm{im}(\tau) = \{\tau(v)\colon v \in V\} \subseteq U$.*

Izrek 6. *Jedro $\ker\tau$ linearne preslikave $\tau\colon V \to U$ je vektorski podprostor v V, slika $\mathrm{im}\,\tau$ pa vektorski podprostor v U.*

Izrek 7. *Naj bo $\tau\colon V \to U$ linearna preslikava iz vektorskega prostora V v vektorski prostor U.*

(1) *τ je injektivna natanko tedaj, ko je $\ker\tau = \{0\}$.*
(2) *τ je surjektivna natanko tedaj, ko je $\mathrm{im}\,\tau = U$.*

Naj bo A_τ matrika, ki pripada linearni preslikavi $\tau\colon V \to U$ iz baze $\mathcal{B}$ v bazo $\mathcal{C}$. Po definiciji je jedro linearne preslikave τ množica vseh vektorjev v, za katere velja $\tau(v) = 0$. Če razvijemo vektor v po bazi $\mathcal{B}$, so koeficienti v razvoju določeni s stolpcem $x = [\beta_1, \beta_2, \ldots, \beta_m]^T$. Spomnimo se enakosti (2). Ker ima slika $\tau(v) = 0$ v razvoju po bazi $\mathcal{C}$ vse koeficiente enake 0, velja $A_\tau x = 0$. (Z drugimi besedami, vektorji koeficientov v razvoju vektorjev iz $\ker\tau$ po bazi $\mathcal{B}$ ustrezajo natanko ničelnemu prostoru matrike A_τ.)

Podobno si pomagamo z enakostjo (2), da bi določili $\operatorname{im}\tau$. Po definiciji je $\operatorname{im}\tau$ množica vseh slik linearne preslikave, torej množica vseh linearnih kombinacij slik vektorjev iz baze $\mathcal{B}$. Z drugimi besedami,

$$\begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \\ \vdots \\ \\ \gamma_n \end{bmatrix} = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \ldots & \alpha_{1m} \\ \alpha_{21} & \alpha_{22} & \ldots & \alpha_{2m} \\ \\ \vdots & \vdots & & \vdots \\ \\ \alpha_{n1} & \alpha_{n2} & \ldots & \alpha_{nm} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_m \end{bmatrix} = A^{(1)}\beta_1 + A^{(2)}\beta_2 + \ldots + A^{(m)}\beta_m,$$

kjer z A^j označimo j-ti stolpec matrike $A_\tau = [\alpha_{ij}]$. Torej so vektorji koeficientov $[\gamma_1, \gamma_2, \ldots, \gamma_n]^T$ natanko vse linearne kombinacije stolpcev matrike A_τ. Sledi, da linearna ogrinjača stolpcev natanko določa koeficiente vektorjev v $\operatorname{im}\tau$ pri razvoju po bazi $\mathcal{C}$. (Z drugimi besedami, koeficienti pri razvoju vektorjev iz $\operatorname{im}\tau$ po bazi $\mathcal{C}$ ustrezajo stolpčnemu prostoru matrike A_τ.)

Izrek 8. *Naj bo $\tau\colon V \to U$ linearna preslikava in naj bo $A = A_{\tau,\mathcal{B},\mathcal{C}}$ matrika, ki pripada preslikavi τ. Potem je*

(1) $\dim(\operatorname{im}(\tau)) = \operatorname{rank}(A)$,
(2) $\dim(\ker(\tau)) + \dim(\operatorname{im}(\tau)) = \dim(V)$.

Posledica 2. *Naj bo $\tau\colon V \to U$ linearna preslikava,* $\dim V = \dim U = n$ *in naj bo A neka matrika, ki pripada τ. Naslednje trditve so ekvivalentne:*

(1) *τ je bijektivna.*
(2) *τ je injektivna.*
(3) *τ je surjektivna.*
(4) *A je obrnljiva.*
(5) $\ker\tau = \{\mathbf{0}\}$.
(6) $N(A) = \{\mathbf{0}\}$.
(7) $\operatorname{im}\tau = U$.
(8) $C(A) = \mathbb{R}^n$.
(9) *Rang matrike A je n.*
(10) *Vrstice matrike A so linearno neodvisne.*
(11) *Vrstice matrike A razpenjajo $\mathbb{R}^n$.*
(12) *Vrstice matrike A tvorijo bazo $\mathbb{R}^n$.*
(13) *Stolpci matrike A so linearno neodvisni.*
(14) *Stolpci matrike A razpenjajo $\mathbb{R}^n$.*
(15) *Stolpci matrike A tvorijo bazo $\mathbb{R}^n$.*

(16) $\det A \neq 0$.
(17) *Homogeni sistem enačb $Ax = 0$ ima le trivialno rešitev.*
(18) *Sistem enačb $Ax = b$ ima rešitev za vsak $b \in \mathbb{R}^n$.*

6. Nadaljnje branje

(1) David Poole, Linear algebra, a modern introduction, Thomson, 2006, Razdelki 6.4.-6.7.
(2) Tomaž Košir, Linearna algebra, poglavje VII: Linearne preslikave.

7. Domača naloga

ϟ (1) Rešite kviz na Učilnici.
ϟ (2) Dokažite izreka 2 and 7 ter posledico 2.
ϟ (3) Naj bosta $u, v \in V$ linearno neodvisna lastna vektorja linearne preslikave $\tau\colon V \to V$. Če je $u + v$ tudi lastni vektor za τ, potem pokažite, da u in v pripadata isti lastni vrednosti.
ϟ (4) Dokažite, da nobena linearna preslikava $\tau : \mathbb{R}^{2\times 3} \to \mathbb{R}_4[x]$ ni injektivna.
ϟ (5) Definirajmo preslikavo $\tau : \mathbb{R}^{3\times 3} \to \mathbb{R}$ s predpisom $\tau(A) = \operatorname{tr}(A)$.
 (a.) Pokažite, da je τ linearna.
 (b.) Zapišite matriko, ki pripada τ v standardnih bazah $\mathbb{R}^{3\times 3}$ in $\mathbb{R}$.
 (c.) Določite $\ker \tau$ in $\operatorname{im} \tau$.
ϟ (6) Definirajmo preslikavo $\tau : \mathbb{R}_1[x] \to \mathbb{R}_2[x]$ s predpisom $\tau(p(x)) = x^2 p'(x)$.
 (a) Pokažite, da je τ linearna.
 (b) Zapišite matriko, ki ustreza preslikavi τ v standardnih bazah $\mathbb{R}_1[x]$ in $\mathbb{R}_2[x]$.
 (c) Določite $\ker \tau$ in $\operatorname{im} \tau$.
ϟ (7) Aleksandra Franc, Rešene naloge iz linearne algebre, naloge poglavja 5.
ϟ (8) David Poole, Linear algebra, a modern introduction, Thomson, 2006,
 (a) strani 484-485, naloge 1-36,
 (b) strani 499-500, naloge 1-20,
 (c) strani 516-518, naloge 1-44,
 (d) strani 536-537, review questions 1-20.

(Naloge, označene s ϟ preverjajo razumevanje osnovnih pojmov in so primeri nalog s teoretičnih izpitov.)

Matematika 1	**Izročki: 7. teden**

Izometrije

Polona Oblak

1. (Nelinearne) izometrije

Izometrija je (*morda nelinearna*) preslikava $\mathcal{A}\colon \mathbb{R}^n \to \mathbb{R}^n$, za katero velja

$$||x-y|| = ||\mathcal{A}(x) - \mathcal{A}(y)||,$$

za vse $x, y \in \mathbb{R}^n$. To pomeni, da izometrije ohranjajo razdalje.

Izrek 1. *Za linearno preslikavo $\tau\colon \mathbb{R}^n \to \mathbb{R}^n$ so naslednje trditve ekvivalentne:*

(1) *τ je izometrija.*
(2) *$||x|| = ||\tau(x)||$ za vse $x \in \mathbb{R}^n$.*
(3) *$x^{\mathsf{T}} y = \tau(x)^{\mathsf{T}} \tau(y)$ za vse $x, y \in \mathbb{R}^n$.*
(4) *Matrika, ki pripada τ, je ortogonalna.*

Linearno izometrijo imenujemo ortogonalna *preslikava.*

Izrek 2. *Izometrijo $\mathcal{A} : \mathbb{R}^n \to \mathbb{R}^n$ lahko enolično zapišemo kot*

$$\mathcal{A}v = Qv + a,$$

kjer je $Q \in \mathbb{R}^{n\times n}$ ortogonalna matrika in $a \in \mathbb{R}^n$.

2. Ortogonalne preslikave = linearne izometrije

Izrek 3. *Vsaka ortogonalna preslikava $\tau\colon \mathbb{R}^2 \to \mathbb{R}^2$ je ena od naslednjih:*

(1) *Zrcaljenje čez premico $y = kx$, kjer $k = \tan\frac{\varphi}{2}$ in*

$$Z_\varphi = \begin{pmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi & -\cos\varphi \end{pmatrix}$$

je matrika, ki pripada τ v standardni bazi $\mathbb{R}^2$.

Lastni vrednosti zrcaljenja sta 1 in -1, pripadajoča lastna vektorja sta v smeri premice, čez katero zrcalimo (pri lastni vrednosti 1), ter v smeri vektorja, pravokotnega na premico (pri lastni vrednosti -1).

(2) *Rotacija za kot φ okoli koordinatnega izhodišča v pozitivni smeri in*

$$R_\varphi = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}$$

je matrika, ki pripada τ v standardni bazi $\mathbb{R}^2$.

Lastni vrednosti rotacije sta $\cos\varphi + i\sin\varphi$ *in* $\cos\varphi - i\sin\varphi$*, pripadajoča lastna vektorja pa imata kompleksne vrednosti (razen v primeru, ko je* $\varphi = k\pi$ *za* $k \in \mathbb{Z}$*).*

Izrek 4. *Vsaka ortogonalna preslikava* $\tau\colon \mathbb{R}^3 \to \mathbb{R}^3$ *je ena od naslednjih:*

(1) *Zrcaljenje čez ravnino* Σ *skozi koordinatno izhodišče* $(0, 0, 0)$*. V tem primeru ima* τ *dvojno lastno vrednost 1 in enojno lastno vrednost* -1*. Pri tem je ravnina zrcaljenja* $\Sigma = \mathcal{L}\{a, b\}$ *napeta na lastna vektorja* a *in* b *pri lastni vrednosti* 1*. Matrika, ki pripada* τ *v bazi* $\{a, b, n\}$*, je enaka*
$$A_\tau = \left[\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \hline 0 & 0 & -1 \end{array}\right],$$
kjer je n *normala na ravnino* Σ*.*

(2) *Rotacija okoli premice* p *skozi koordinatno izhodišče* $(0, 0, 0)$ *za kot* φ*. Pri tem ima* τ *lastne vrednosti* 1*,* $\cos\varphi + i\sin\varphi$ *in* $\cos\varphi - i\sin\varphi$ *in je os rotacije* p *napeta na lastni vektor* a *pri lastni vrednosti 1. Matrika, ki pripada* τ*, je tako v bazi* $\{a, b, c\}$ *enaka*
$$A_\tau = \left[\begin{array}{c|cc} 1 & 0 & 0 \\ \hline 0 & \cos\varphi & -\sin\varphi \\ 0 & \sin\varphi & \cos\varphi \end{array}\right],$$
kjer sta vektorja b *in* c *pravokotna na* a*.*

(3) *Zrcalni zasuk, kjer je os rotacije* p *pravokotna na ravnino zrcaljenja* Σ*. Pri tem ima* τ *lastne vrednosti* -1*,* $\cos\varphi + i\sin\varphi$ *in* $\cos\varphi - i\sin\varphi$ *in je os rotacije* $p = \mathcal{L}\{n\}$ *napeta na lastni vektor* n *pri lastni vrednosti* -1*, kjer je* n *normala na* Σ*. Tako je matrika, ki pripada* τ*, v bazi* $\{n, a, b\}$ *enaka*
$$A_\tau = \left[\begin{array}{c|cc} -1 & 0 & 0 \\ \hline 0 & \cos\varphi & -\sin\varphi \\ 0 & \sin\varphi & \cos\varphi \end{array}\right],$$
kjer je $\Sigma = \mathcal{L}\{a, b\}$*.*

3. Nadaljnje branje

ϟ (1) Joseph B. Kadane, Principles of Uncertainty, strani 198-201.

⋆ (2) Tomaž Košir, Linearna algebra, Razdelek 5.

⋆ (3) (še bolj geometrijski pogled) Walter Meyer, Geometry and Its Applications, 2006, Poglavje 4.

4. Domača naloga

ϟ (1) Rešite kviz na Učilnici.

ϟ (2) Zapišite primer izometrije $\mathcal{A} : \mathbb{R}^3 \to \mathbb{R}^3$, ki ni identiteta, vendar pa je $\mathcal{A}^3$ identična preslikava.

ϟϟ (3) Dokažite, da je vsaka rotacija v $\mathbb{R}^2$ kompozitum dveh zrcaljenj.

⋆ (4) Dokažite, da je vsaka rotacija v $\mathbb{R}^n$ kompozitum dveh zrcaljenj.

$\lightning$ (5) David Poole, Linear algebra, a modern introduction, Thomson, 2006, strani 374-375, exercises 29-32, 34.

(Naloge, označene s $\lightning$ preverjajo razumevanje osnovnih pojmov in so primeri nalog s teoretičnih izpitov, naloga označena s $\lightning\lightning$ pa je nekoliko težja. Naloga, označena s ⋆ je težja, širi vaše znanje in dopolnjuje odpredavano snov.)

Matematika 1 **Izročki: 9. teden**

Funkcije in vektorske funkcije več spremenljivk, ponovitev

Polona Oblak

1. Funkcije več spremenljivk

1.1. Definicija. *Funkcija več spremenljivk*

$$\begin{aligned} f\colon \quad & \mathcal{D}_f \subseteq \mathbb{R}^n && \to \mathbb{R}, \\ \mathbf{x} = (x_1, x_2, \ldots, x_n) & && \mapsto f(x_1, x_2, \ldots, x_n) \end{aligned}$$

je funkcija, ki predpiše realno vrednost $f(\mathbf{x}) = f(x_1, x_2, \ldots, x_n) \in \mathbb{R}$ vsaki točki $\mathbf{x} = (x_1, x_2, \ldots, x_n) \in \mathcal{D}_f \subseteq \mathbb{R}^n$. Množici $\mathcal{D}_f$ pravimo *definicijsko območje* funkcije f.

V primeru, ko je $n = 2$, je graf funkcije $f = f(x, y)\colon \mathcal{D}_f \subseteq \mathbb{R}^2 \to \mathbb{R}$ ploskev

$$\Gamma_f = \{(x, y, f(x, y))\colon (x, y) \in \mathcal{D}_f\}$$

v $\mathbb{R}^3$. *Nivojska krivulja* (ali *nivojnica*) funkcije $f = f(x, y)$ je množica vseh točk $(x, y) \in \mathcal{D}_f$, za katere velja $f(x, y) = c$ za dano realno realno število $c \in \mathbb{R}$. Tako vsaka točka $(x, y) \in \mathcal{D}_f$ leži na natanko eni nivojski krivulji in zato se definicijsko območje $\mathcal{D}_f$ razsloji na nivojske krivulje.

1.2. Odvodi. *Parcialni odvod* funkcije $f\colon \mathbb{R}^n \to \mathbb{R}$ v točki $\mathbf{a} = (a_1, a_2, \ldots, a_n)$ *po spremenljivki x_i* definiramo kot

$$f_{x_i}(\mathbf{a}) = \frac{\partial f}{\partial x_i}(\mathbf{a}) = \lim_{h \to 0} \frac{f(a_1, \ldots, a_{i-1}, a_i + h, a_{i+1}, \ldots, a_n) - f(\mathbf{a})}{h}.$$

Tako nam torej parcialni odvod funkcije f po x_i v točki $\mathbf{a} = (a_1, a_2, \ldots, a_n)$ pove relativno spremembo funkcijske vrednosti pri zelo majhni spremembi spremenljivke x_i, kjer so ostale spremenljivke fiksne.

Vektor

$$(\operatorname{grad} f)(\mathbf{a}) = (f_{x_1}(\mathbf{a}), f_{x_2}(\mathbf{a}), \ldots, f_{x_n}(\mathbf{a}))$$

imenujemo *gradient* funkcije f v točki $\mathbf{a}$.

Smerni odvod funkcije $f\colon \mathbb{R}^n \to \mathbb{R}$ v točki $\mathbf{a} = (a_1, a_2, \ldots, a_n)$ *v smeri vektorja $\vec{e}$* je enak

$$f_{\vec{e}}(\mathbf{a}) = (\operatorname{grad} f)(\mathbf{a}) \cdot \frac{\vec{e}}{||\vec{e}||} = \sum_{i=1}^{n} f_{x_i}(\mathbf{a}) \frac{e_i}{||\vec{e}||}.$$

Za funkcijo $f\colon \mathbb{R}^n \to \mathbb{R}$ velja:

(1) Gradient funkcije f v točki $\mathbf{a}$ kaže v smeri najhitrjšega naraščanja funkcije f v točki $\mathbf{a}$.
(2) V primeru $n = 2$ je gradient funkcije $f = f(x, y)$ v točki $\mathbf{a}$ pravokoten na nivojsko krivuljo v tej točki.

(3) Smerni odvod $f_{\vec{e}}(\mathbf{a})$ je relativna sprememba funkcijske vrednosti $f(\mathbf{a})$ ob majhnem premiku iz točke $\mathbf{a}$ v smeri vektorja $\vec{e}$. Zato velja:
- Če je $f_{\vec{e}}(\mathbf{a}) > 0$, potem f ob majhnem pomiku iz točke $\mathbf{a}$ v smeri vektorja $\vec{e}$ narašča.
- Če je $f_{\vec{e}}(\mathbf{a}) < 0$, potem f ob majhnem pomiku iz točke $\mathbf{a}$ v smeri vektorja $\vec{e}$ pada.

1.3. **Linearna aproksimacija.** Vrednost funkcije $f\colon \mathbb{R}^n \to \mathbb{R}$ lahko v točki blizu $\mathbf{a}$ ocenimo z vrednostjo $f(\mathbf{a})$. Najprej si nazorno oglejmo, kako to naredimo za funkcijo dveh spremenljivk. Denimo, da želimo približno oceniti vrednost funkcije $f\colon \mathbb{R}^2 \to \mathbb{R}$ v točki blizu točke (a,b). Točke, ki so "blizu"točki (a,b) so oblike $(a+h, b+k)$, kjer sta h in k dovolj majhni realni števili. Tangentna ravnina $\Sigma_{(a,b)}$ na graf $z = f(x,y)$ v točki $(a,b,f(a,b))$ vsebuje vektorja $[1,0,f_x(a,b)]^T$ ter $[0,1,f_y(a,b)]^T$. Normala na ravnino $\Sigma_{(a,b)}$ je tako enaka

$$\begin{bmatrix} 1 \\ 0 \\ f_x(a,b) \end{bmatrix} \times \begin{bmatrix} 0 \\ 1 \\ f_y(a,b) \end{bmatrix} = \begin{bmatrix} -f_x(a,b) \\ -f_y(a,b) \\ 1 \end{bmatrix},$$

ker pa ravnina vsebuje tudi točko $(a,b,f(a,b))$, je enačba tangentne ravnine enaka

$$-f_x(a,b)x - f_y(a,b)y + z = -f_x(a,b)a - f_y(a,b)b + f(a,b).$$

Če izrazimo tretjo koordinato z točke (x,y,z) na tangentni ravnini $\Sigma_{(a,b)}$, dobimo enačbo tangentne ravnine

$$z = f(a,b) + (x-a)f_x(a,b) + (y-b)f_y(a,b).$$

Če torej želimo izračunati približno funkcijsko vrednost $f(a+h, b+k)$, jo lahko ocenimo z višino točke $(a+h, b+k, z)$ na tangentni ravnini $\Sigma_{(a,b)}$. Torej je

$$f(a+h, b+k) \cong f(a,b) + f_x(a,b)h + f_y(a,b)k.$$

Pri tem bi lahko zadnja sumanda zapisali tudi kot skalarni produkt gradienta funkcije f v točki (a,b) in vektorja pomika (h,k) iz točke (a,b). Tako lahko linearno aproksimacijo funkcije dveh spremenljivk posplošimo na funkcije več spremenljivk. Za dano funkcijo $f\colon \mathbb{R}^n \to \mathbb{R}$ lahko v točki $\mathbf{a}+\mathbf{h}$ blizu $\mathbf{a}$ njeno funkcijsko vrednost ocenimo s formulo

$$f(\mathbf{a}+\mathbf{h}) \cong f(\mathbf{a}) + (\operatorname{grad} f)(\mathbf{a}) \cdot \mathbf{h}.$$

1.4. **Višji odvodi.** Parcialne odvode drugega reda izračunamo s parcialnim odvajanjem parcialnih odvodov prvega reda. Definiramo jih kot

$$f_{x_i x_j}(\mathbf{a}x) = \frac{\partial^2 f}{\partial x_j \partial x_i}(\mathbf{x}) = \frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_i}(\mathbf{a}x)\right).$$

$n \times n$ matriko

$$H_f(\mathbf{x}) = \left[\frac{\partial^2 f}{\partial x_j \partial x_i}(\mathbf{x})\right]_{i,j,=1,\dots,n}$$

imenujemo *Hessejeva matrika* funkcije f v točki x. Če sta pri tem $f_{x_i x_j}$ in $f_{x_j x_i}$ zvezni funkciji, potem sta omenjena druga parcialna odvoda enaka. Zato je v primeru, ko so vsi parcialni odvodi $f_{x_i x_j}$ zvezni, Hessejeva matrika $H_f(x,y)$ matrika.

Če druge parcialne odvode še naprej odvajamo, dobimo parcialne odvode višjih redov. Če so zvezni, so neodvisni od vrstnega reda odvajanja. V tem primeru za funkcijo f dveh spremeljivk dobimo štiri različne parcialne odvode tretjega reda: $\frac{\partial^3 f}{\partial x^3}(a,b)$, $\frac{\partial^3 f}{\partial y \partial x^2}(a,b) = \frac{\partial^3 f}{\partial x \partial y \partial x}(a,b) = \frac{\partial^3 f}{\partial x^2 \partial y}(a,b)$, $\frac{\partial^3 f}{\partial y^2 \partial x}(a,b) = \frac{\partial^3 f}{\partial y \partial x \partial y}(a,b) = \frac{\partial^3 f}{\partial x \partial y^2}(a,b)$ in $\frac{\partial^3 f}{\partial y^3}(a,b)$.

2. Vektorske funkcije več spremenljivk

2.1. Definicija. *Vektorska funkcija*

$$\begin{aligned} F\colon \quad \mathcal{D}_F \subseteq \mathbb{R}^n &\to \mathbb{R}^m, \\ \mathbf{x} &\mapsto [f_1(\mathbf{x})\ f_2(\mathbf{x})\ \ldots\ f_m(\mathbf{x})]^\mathsf{T} \end{aligned}$$

je m-terica funkcij več spremenljivk.

2.2. Odvodi. *Jacobijeva matrika* vektorske funkcije $F\colon \mathcal{D}_F \subseteq \mathbb{R}^n \to \mathbb{R}^m$ je $m \times n$ matrika prvih odvodov funkcij $f_1, \ldots, f_m$:

$$J_F(\mathbf{x}) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(\mathbf{x}) & \frac{\partial f_1}{\partial x_2}(\mathbf{x}) & \ldots & \frac{\partial f_1}{\partial x_n}(\mathbf{x}) \\ \frac{\partial f_2}{\partial x_1}(\mathbf{x}) & \frac{\partial f_2}{\partial x_2}(\mathbf{x}) & \ldots & \frac{\partial f_2}{\partial x_n}(\mathbf{x}) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(\mathbf{x}) & \frac{\partial f_m}{\partial x_2}(\mathbf{x}) & \ldots & \frac{\partial f_m}{\partial x_n}(\mathbf{x}) \end{bmatrix}.$$

Absolutna vrednost determinante Jacobijeve matrike vektorske funkcije $F\colon \mathcal{D}_F \subseteq \mathbb{R}^n \to \mathbb{R}^n$ pove, za kakšen faktor funkcija lokalno raztegne prostornine.

3. Nadaljnje branje in gledanje

(1) James Stewart, Calculus, early transcendentals, 2016, Poglavja 14.1-14.6.
(2) Khan Academy: Unit: Derivatives of multivariable functions.

4. Domača naloge

(1) Rešite kviz na Učilnici.
(2) James Stewart, Calculus, early transcendentals, 2016, stran 945, exercise 5-6, 13-29, 35-48, stran 1020, exercise 1-6.

Matematika 1	**Izročki: 9. teden**

Večkratni integrali

Polona Oblak

1. Dvojni integralna pravokotniku

Formalno definiramo dvojni integral

$$\iint\limits_R f(x,y)\,dxdy,$$

funkcije $f\colon r \to \mathbb{R}$ na pravokotniku $R = [a,b]\times[c,d]$ na naslednji način. Najprej razdelimo pravokotnik R na mn manjših pravokotničkov R_{ij} s stranicama dolžin $\Delta x = \frac{b-a}{n}$ ter $\Delta y = \frac{d-c}{m}$. V vsakem pravokotničku R_{ij} izberemo poljubno točko (x_{ij}^*, y_{ij}^*). Prostornina telesa pod grafom $z = f(x,y)$ nad pravokotnikom R_{ij} je tako približno enaka

$$f(x_{ij}^*, y_{ij}^*)\Delta x\Delta y.$$

Dvojni integral funkcije f na pravokotniku $R = [a,b]\times[c,d]$ tako definiramo kot

$$\iint\limits_R f(x,y)\,dxdy = \lim_{m,n\to\infty} \sum_{j=1}^{m}\sum_{i=1}^{n} f(x_{ij}^*, y_{ij}^*)\Delta x\Delta y,$$

če ta limita obstaja.

Izkaže se, da je integral $\iint_R f(x,y)\,dxdy$ neodvisen od izbire točke $(x_{ij}^*, y_{ij}^*) \in R_{ij}$. Če je f nenegativna funkcija, je $\iint_R f(x,y)\,dxdy$ enak prostornini telesa, ki je omejen s pravokotnikom R ter grafom $z = f(x,y)$.

Izrek 1 (Fubini). *Če je $f\colon R \to \mathbb{R}$ zvezna funkcija na pravokotniku $R = [a,b]\times[c,d] \subseteq \mathbb{R}^2$, potem*

$$\iint\limits_R f(x,y)\,dxdy = \int_c^d \left(\int_a^b f(x,y)\,dx\right)dy = \int_a^b \left(\int_c^d f(x,y)\,dy\right)dx.$$

2. Dvojni integral

Če je $D \subseteq \mathbb{R}^2$ neko omejeno območje ter $f\colon D \to \mathbb{R}$ zvezna funkcija, izberimo tak pravokotnik R, da velja $D \subseteq R$. Sedaj definiramo *dvojni integral funkcije f na območju D* kot

$$\iint\limits_D f(x,y)\,dxdy = \iint\limits_R F(x,y)\,dxdy\,,$$

kjer

$$F(x,y)=\begin{cases} f(x,y), & (x,y)\in D\\ 0, & (x,y)\notin D. \end{cases}$$

Sedaj lahko zapišimo tudi Fubinijev izrek za nepravokotna območja.

Izrek 2 (Fubini). (1) *Če je* $D=\{(x,y);\ a\le x\le b$ *in* $\varphi_1(x)\le y\le \varphi_2(x)\}\subseteq \mathbb{R}^2$ *in* $f\colon D\to\mathbb{R}$ *zvezna funkcija, potem je*

$$\iint\limits_D f(x,y)\,dxdy=\int_a^b\left(\int_{\varphi_1(x)}^{\varphi_2(x)} f(x,y)\,dy\right)dx.$$

(2) *Če je* $D=\{(x,y);\ \vartheta_1(y)\le x\le \vartheta_2(y)$ *in* $c\le y\le d\}\subseteq\mathbb{R}^2$ *in* $f\colon D\to\mathbb{R}$ *zvezna funkcija, potem je*

$$\iint\limits_D f(x,y)\,dxdy=\int_c^d\left(\int_{\vartheta_1(y)}^{\vartheta_2(y)} f(x,y)\,dx\right)dy.$$

3. Trojni integral...

... je definiran na podoben način. Po Fubinijevem izreku lahko tudi trojne integrale na nekaterih območjih zapišemo kot trikratne integrale.

4. Menjava spremenljivk

Spomnimo se izreka prejšnjega tedna:

Izrek 3. *Naj bo* $f\colon D\to\mathbb{R}$ *zvezna funkcija na* $D\subseteq\mathbb{R}^2$. *Če je* $x=\varphi(u,v)$, $y=\vartheta(u,v)$, *takšna menjava spremenljivk, da je* $\det J_{\varphi,\vartheta}\neq 0$, *potem*

$$\iint\limits_D f(x,y)\,dxdy=\int_D f(\varphi(u,v),\vartheta(u,v))\left|\det J_{\varphi,\vartheta}\right|\,dudv.$$

Podobno, če je $f\colon D\to\mathbb{R}$ *zvezna funkcija na* $D\subseteq\mathbb{R}^3$ *ter* $x=\varphi(u,v,w)$, $y=\vartheta(u,v,w)$, $z=\psi(u,v,w)$ *takšna menjava spremenljivk, da je* $\det J_{\varphi,\vartheta,\psi}\neq 0$, *potem evlja*

$$\iiint\limits_D f(x,y,x)\,dxdydz=\int_D f(\varphi(u,v,w),\vartheta(u,v,w),\psi(u,v,w))\left|\det J_{\varphi,\vartheta,\psi}\right|\,dudvdw.$$

Nekaj primerov menjave spremenljivk:

(1) *Polarne koordinate* v $\mathbb{R}^2$ so podane z

$$x=r\cos\varphi,\ y=r\sin\varphi,$$

$r\ge 0$, $\varphi\in[0,2\pi]$, in velja $|\det J_{\text{polar}}|=r$.

(2) *Cilindrične koordinate* v $\mathbb{R}^3$ so podane z

$$x = r\cos\varphi,\ y = r\sin\varphi,\ z = z,$$

$r \geq 0$, $\varphi \in [0, 2\pi]$, $z \in \mathbb{R}$, in velja $|\det J_{\text{cylindric}}| = r$.

(3) *Sferične koordinate* v $\mathbb{R}^3$ so podane z

$$x = r\cos\varphi\cos\vartheta,\ y = r\sin\varphi\cos\vartheta,\ z = r\sin\vartheta.$$

$r \geq 0$, $\varphi \in [0, 2\pi]$, $\vartheta \in [-\frac{\pi}{2}, \frac{\pi}{2}]$ in velja $|\det J_{\text{spherical}}| = r^2\cos\vartheta$.

5. Nadaljnje branje

(1) James Stewart, Calculus, early transcendentals, 2016, Poglavje 15.

(2) Wolfram Demonstrations: Approximating a Double Integral with Cuboids.

6. Domača naloga

ϟ (1) James Stewart, Calculus, early transcendentals, 2016, strani 1022-23, Exercises 3-52.

ϟ (2) Naj bo (u, v) novi koordinatni sistem v $\mathbb{R}^2$, definiran z

$$x = e^u\cos v, y = e^u\sin v,$$

kjer $u \in \mathbb{R}$ and $v \in [0, 2\pi]$. Določite determinanto Jacobijeve matrike $J_{u,v} = \frac{\partial(x,y)}{\partial(u,v)}$.

ϟ (3) Naj bodo (u, v, w) nove koordinate v $\mathbb{R}^3$, definirane s predpisom

$$x = u\cos v, y = 2u\sin v, z = 3w.$$

kjer je $u \geq 0$, $v \in [0, 2\pi]$ in $w \in \mathbb{R}$. Izračunajte determinanto Jacobijeve matrike $J_{u,v,w} = \frac{\partial(x,y,z)}{\partial(u,v,w)}$.

ϟ (4) Naj bo (u, v, w) koordinatni sistem v $\mathbb{R}^3$ definiran z

$$x = 3ue^v, y = 2we^u, z = u.$$

Določite determinanto Jacobijeve matrike $J_{u,v,w} = \frac{\partial(x,y,z)}{\partial(u,v,w)}$.

ϟ (5) Naj bo $a \in \mathbb{R}$ neničelno število. Za pišite naslednji integral s spremembo v polarni koordinatni sistem

$$\int_0^a \left(\int_0^{\sqrt{a^2-y^2}} f(x,y)dx \right) dy.$$

ϟ (6) Zamenjajte vrstni red integracije:

$$\int_0^1 \left(\int_0^{\sqrt{y}} f(x,y)dx \right) dy = \int_?^? \left(\int_?^? f(x,y)dy \right) dx$$

ϟ (7) Zapišite integral

$$\int_{-1}^{1}\left(\int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}}\left(\int_{\sqrt{x^2+y^2}}^{1} f(x,y,z)dz\right)dx\right)dy$$

v cilindričnih koordinatah.

(Naloge, označene s ϟ preverjajo razumevanje osnovnih pojmov in so primeri nalog s teoretičnih izpitov.)

Matematika 1	**Izročki: 10. teden**

Optimizacija. 1. del.

Polona Oblak

1. Klasifikacija lokalnih ekstremov

Naj bo $f\colon \mathbb{R}^n \to \mathbb{R}$ ter a v definicijskem območju funkcije f.

Če za vse točke $\mathbf{x} \neq \mathbf{a}$, ki so "dovolj blizu"točke a (tj. $\|\mathbf{x}-\mathbf{a}\| < \varepsilon$ za nek dovolj majhen ε) velja $f(\mathbf{x}) < f(\mathbf{a})$, potem pravimo, da ima funkcija f v točki a *lokalni maksimum*.

Če za vse točke $\mathbf{x} \neq \mathbf{a}$, ki so "dovolj blizu"točke a (tj. $\|\mathbf{x}-\mathbf{a}\| < \varepsilon$ za nek dovolj majhen ε) velja $f(\mathbf{x}) > f(\mathbf{a})$, potem pravimo, da ima funkcija f v točki a *lokalni minimum*.

Če je funkcija f zvezno parcialno odvedljiva, potem je jasno, da ima lahko lokalne ekstreme le v stacionarnih točkah. Torej je potreben pogoj za lokalni ekstrem funkcije f v točki a:

$$(\operatorname{grad} f)(\mathbf{a}) = \mathbf{0},$$

kar pomeni, da moramo lokalne ekstreme iskati zgolj med stacionarnimi točkami.

Izrek 1. *Naj bo* a *stacionarna točka dvakrat parcialno zvezno odvedljive funkcije* $f\colon \mathbb{R}^n \to \mathbb{R}$*.*

(1) *Če so vse lastne vrednosti matrike* $H_f(\mathbf{a})$ *pozitivne, ima* f *v* a *lokalni minimum.*
(2) *Če so vse lastne vrednosti matrike* $H_f(\mathbf{a})$ *negativne, ima* f *v* a *lokalni maksimum.*
(3) *Če so vse lastne vrednosti matrike* $H_f(\mathbf{a})$ *neničelne, vendar različno predznačene, lokalnega ekstrema v* a *ni.*
(4) *Če je kakšna od lastnih vrednosti matrike* $H_f(\mathbf{a})$ *enaka* 0*, o lokalnih ekstremih funkcije* f *v točki* a *iz matrike* $H_f(\mathbf{a})$ *ne moremo sklepati.*

2. Lokalni ekstremi

Pogosto naletimo na problem iskanja ekstremalnih vrednosti funkcije $f\colon \mathbb{R}^n \to \mathbb{R}$ pri pogojih

$$g_1(\mathbf{x}) = g_2(\mathbf{x}) = \ldots = g_m(\mathbf{x}) = 0.$$

Izkaže se, da lahko lokalni ekstremi funkcije f pri pogojih $g_i(\mathbf{x}) = 0$, $i = 1, \dots, m$, nastopijo le v stacionarnih točkah funkcije

$$L = f - \lambda_1 g_1 - \cdots - \lambda_m g_m,$$

ki je funkcija $n+m$ spremenljivk x_1, x_2, ..., x_n, λ_1, λ_2, ..., λ_m.

Funkciji L pravimo *Lagrangeova funkcija*, novim spremenljivkam λ_1, λ_2, ..., λ_m pa *Lagrangevi multiplikatorji*.

Omenjeni pogoj ni zadosten. Nekatere kritične točke funkcije L so ekstremalne točke funkcije $f\colon \mathbb{R}^n \to \mathbb{R}$ pod pogoji $g_1(\mathbf{x}) = g_2(\mathbf{x}) = \ldots = g_m(\mathbf{x}) = 0$, ostale pa ne.

3. Nadaljnje branje

(1) James Stewart, Calculus, early transcendentals, 2016, Razdelek 14.8.
(2) Wolfram Demonstrations: The Geometry Of Lagrange Multipliers.

4. Domača naloga

ϟ (1) Rešite kviz na Spletni učilnici.
ϟ (2) James Stewart, Calculus, early transcendentals, 2016, strani 940-941, vse naloge.
ϟ (3) Naj bo $f\colon \mathbb{R}^3 \to \mathbb{R}^3$ dvakrat zvezno odvedljiva funkcija treh spremeljivk. Za vsako od točk $P, R \in \mathbb{R}^3$ določite in utemeljite, ali sta lokalni minimum, lokalni maksimum, ali nista ekstremalni točki ali pa iz danih podatkov ne moremo ugotoviti, katerega tipa sta točki.
  (a.) $f_x(P) = f_y(P) = f_z(P) = 0$, $f_{xx}(P) = 3, f_{yy}(P) = -1, f_{xy}(P) = 0$.
  (b.) $f_x(R) = f_y(R) = f_z(R) = 1$, $f_{xx}(R) = f_{yy}(R) = f_{zz}(R) = 1, f_{xy}(R) = f_{yz}(R) = f_{xz}(R) = 0$.
ϟ (4) Ali obstaja takšna dvakrat zvezno odvedljiva funkcija $f\colon \mathbb{R}^2 \to \mathbb{R}$, da je
$$\frac{\partial^2 f}{\partial x^2}(0,0) = 4, \frac{\partial^2 f}{\partial x \partial y}(0,0) = 1, \frac{\partial^2 f}{\partial y^2}(0,0) = 2,$$
in ima v točki $(x, y) = (0, 0)$
  (a) lokalni minumum? (Če da, zapišite primer. Če ne, razložite, zakaj ne.)
  (b) lokalni maksimum? (Če da, zapišite primer. Če ne, razložite, zakaj ne.)

(Naloge, označene s ϟ preverjajo razumevanje osnovnih pojmov in so primeri nalog s teoretičnih izpitov.)

Matematika 1	**Izročki: 11. teden**

Optimizacija. 2. del.
Odvodi vektorskih funkcij. Vezani ekstremi v vektorski obliki.

Polona Oblak

Naj bo $F\colon \mathbb{R}^n \to \mathbb{R}^m$, $F(\mathbf{x}) = \begin{bmatrix} f_1(\mathbf{x}) \\ f_2(\mathbf{x}) \\ \vdots \\ f_m(\mathbf{x}) \end{bmatrix}$ vektorska funkcija n spremenljivk $\mathbf{x} = (x_1, \ldots, x_n)$. Da bomo poudarili, da sta lahko n in m večja od 1, bomo označili vektor spremenljivk $\mathbf{x} = (x_1, \ldots, x_n)$ tudi kot $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ in vektorsko funkcijo kot $F = \vec{F}$.

Spomnimo se, da je *odvod vektorske funkcije $\vec{F}$ po vektorju spremenljivk $\vec{x}$* je definiran kot

$$\frac{\partial \vec{F}}{\partial \vec{x}} = J_{\vec{F}}(\mathbf{x}) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(\mathbf{x}) & \frac{\partial f_1}{\partial x_2}(\mathbf{x}) & \ldots & \frac{\partial f_1}{\partial x_n}(\mathbf{x}) \\ \frac{\partial f_2}{\partial x_1}(\mathbf{x}) & \frac{\partial f_2}{\partial x_2}(\mathbf{x}) & \ldots & \frac{\partial f_2}{\partial x_n}(\mathbf{x}) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(\mathbf{x}) & \frac{\partial f_m}{\partial x_2}(\mathbf{x}) & \ldots & \frac{\partial f_m}{\partial x_n}(\mathbf{x}). \end{bmatrix} \tag{1}$$

Drugi odvod funkcije $f\colon \mathbb{R}^n \to \mathbb{R}$ (tu $m = 1$) pa kot

$$\frac{\partial^2 f}{\partial \vec{x}^2} = \frac{\partial}{\partial \vec{x}} \left(\frac{\partial f}{\partial \vec{x}} \right)^{\mathsf{T}}.$$

Funkcija $f\colon \mathbb{R}^n \to \mathbb{R}$ je *konveksna* na $\mathcal{D}$, če velja

$$f(t\,\mathbf{x} + (1-t)\,\mathbf{y}) \le t\,f(\mathbf{x}) + (1-t)\,f(\mathbf{y})$$

za vse $\mathbf{x}, \mathbf{y} \in \mathcal{D}$ in za vse $t \in [0, 1]$. Funkcija f je *konkavna* na $\mathcal{D}$, če je funkcija $-f$ konveksna na $\mathcal{D}$.

Izrek 1. *Dvakrat zvezno odvedljiva funkcija $f\colon \mathcal{D} \subseteq \mathbb{R}^n \to \mathbb{R}$ je konveksna natanko tedaj, ko je $\frac{\partial^2 f}{\partial \vec{x}^2}$ pozitivno semidefinitna matrika na $\mathcal{D}$, in je konkavna natanko tedaj, ko je $\frac{\partial^2 f}{\partial \vec{x}^2}$ negativno semidefinitna na $\mathcal{D}$.*

Nekaj pravil:

(1) $\frac{\partial \vec{x}}{\partial \vec{x}} = I_n$

(2) Če je $A \in \mathbb{R}^{m \times n}$, potem $\frac{\partial A\vec{x}}{\partial \vec{x}} = A$.

(3) Če je $\vec{a} \in \mathbb{R}^n$, potem $\frac{\partial \vec{a}^\mathsf{T} \vec{x}}{\partial \vec{x}} = \vec{a}^\mathsf{T}$.
(4) Če je $A \in \mathbb{R}^{n\times n}$, potem $\frac{\partial(\vec{x}^\mathsf{T} A\vec{x})}{\partial \vec{x}} = \vec{x}^\mathsf{T}(A + A^\mathsf{T})$.
(5) Če je $A \in \mathbb{R}^{n\times n}$ simetrična matrika, potem velja $\frac{\partial(\vec{x}^\mathsf{T} A\vec{x})}{\partial \vec{x}} = 2\vec{x}^\mathsf{T} A$.
(6) $\frac{\partial ||\vec{x}||^2}{\partial \vec{x}} = 2\vec{x}^\mathsf{T}$.
(7) Če $\vec{z} = \vec{z}(\vec{x})$ in $\vec{y} = \vec{y}(\vec{x})$, potem $\frac{\partial(\vec{y}^\mathsf{T} \vec{z})}{\partial \vec{x}} = \vec{z}^\mathsf{T} \frac{\partial \vec{y}}{\partial \vec{x}} + \vec{y}^\mathsf{T} \frac{\partial \vec{z}}{\partial \vec{x}}$.
(8) Če $G\colon \mathcal{D}_G \subseteq \mathbb{R}^m \to \mathbb{R}^n$ in $F\colon \mathcal{D}_F \subseteq \mathbb{R}^n \to \mathbb{R}^p$ in $H = F \circ G$, potem $\frac{\partial \vec{H}}{\partial \vec{x}} = \frac{\partial(\vec{F}\circ\vec{G})}{\partial \vec{x}} = \frac{\partial \vec{F}}{\partial \vec{G}}(\vec{G}(\vec{x})) \cdot \frac{\partial \vec{G}}{\partial \vec{x}}$.

Primer 1. *Naj bo $A \in \mathbb{R}^{m\times n}$, $\operatorname{rank} A = m$, $\vec{b} \in \mathbb{R}^m$. Poiščite tisto rešitev sistema $A\vec{x} = \vec{b}$, ki ima najmanjšo normo. (Rešitev)*

1. Nadaljnje branje

(1) David A. Harville: Matrix Algebra From a Statistician's Perspective, Springer, 1997, Poglavje 15.
(2) J.E. Gentle: Matrix Algebra, Theory, Computations, in Applications in Statistics, Springer, 2017, Poglavje 4.
(3) K.B. Petersen, M.S. Pedersen: The Matrix Cookbook.

(Pazite na to, da so v vseh treh zgoraj omenjenih referencah odvodi definirani kot transponiranke matrike (1).)

2. Domača naloga

ϟ (1) Denimo, da funkcija $f\colon \mathbb{R}^2 \to \mathbb{R}$ doseže maksimalno vrednost na krivulji $x^2+2y^2 = 3$ v točki $(1, 1)$. Ali sta $(\operatorname{grad} f)(1, 1)$ in $[1, 2]^T$ linearno odvisna?

ϟ (2) Naj bosta $\vec{a}, \vec{b} \in \mathbb{R}^n$ in funkciji $f, g\colon \mathbb{R}^n \to \mathbb{R}$ definirani kot $f(\vec{x}) = \vec{x}^\mathsf{T}\vec{a}$ in $g(\vec{x}) = \vec{x}^\mathsf{T}\vec{b}$. Izračunajte $\frac{\partial(f(\vec{x})g(\vec{x}))}{\partial \vec{x}}$.

ϟ (3) Naj bo $A \in \mathbb{R}^{n\times n}$ matrika z lastnostjo $A^\mathsf{T} = -A$. Izračunajte $\frac{\partial}{\partial \vec{x}}\left(\vec{x}^\mathsf{T} A\vec{x} + ||5\vec{x}||^2\right)$.

ϟ (4) Za matriko $A \in \mathbb{R}^{n\times n}$ izračunajte $\frac{\partial ||A\vec{x}||^2}{\partial \vec{x}}$.

ϟ (5) Naj bodo $B, C, D \in \mathbb{R}^{n\times n}$ in $\vec{b}, \vec{d} \in \mathbb{R}^n$. Za funkcijo

$$f(\vec{x}) = (B\vec{x} + \vec{b})^\mathsf{T} C(D\vec{x} + \vec{d})$$

izračunajte $\frac{\partial f}{\partial \vec{x}}$ in $\frac{\partial^2 f}{\partial \vec{x}^2}$.

ϟ (6) Naj bo $A \in \mathbb{R}^{n\times n}$ simetrična pozitivno definitna matrika ter $\vec{b} \in \mathbb{R}^n$. Poiščite in klasificirajte lokalne ekstreme funkcije $f\colon \mathbb{R}^n \to \mathbb{R}$, definirane s predpisom

$$f(\vec{x}) = \frac{1}{2}\vec{x}^\mathsf{T} A\vec{x} - \vec{x}^\mathsf{T}\vec{b}.$$

(7) Pokažite, da so naslednje funkcije konveksne:

ϟ (a) $f(x, y) = e^y - \log x$,

ϟ (b) $g(\vec{x}) = \frac{1}{2}\vec{x}^\mathsf{T} A\vec{x}$, kjer je A pozitivno semidefinitna matrika,

ϟ (c) $h(\vec{x}) = ||\vec{x}||^2$,

ϟ (d) $k(\vec{x}) = \vec{a}^\mathsf{T}\vec{x}$, kjer je $\vec{a} \in \mathbb{R}^n$.

ϟ (8) Za dane vektorje $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_k \in \mathbb{R}^n$ poiščite takšne $\vec{x} \in \mathbb{R}^n$, da bo povprečna kvadratna razdalja med vektorjem $\vec{x}$ in vektorji $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_k$ najmanjša možna.

(9) Naj bosta $\vec{a} \in \mathbb{R}^n$ ter $b \in \mathbb{R}$.

ϟ (a) Poiščite najmanjšo in največjo vrednost $\vec{a}^\mathsf{T}\vec{x}$ za vse vektorje $\vec{x} \in \mathbb{R}^n$ s predpisano dolžino $||\vec{x}|| = b$.

ϟ (b) Geometrijsko razložite svojo rešitev iz (9a) v primeru, ko je $n = 3$.

ϟϟ (10) Za pozitivno semidefinitno matriko $A \in \mathbb{R}^{n\times n}$, simetrično matriko $B \in \mathbb{R}^{n\times n}$ ter neničelni $b \in \mathbb{R}$ poiščite najmanjšo in največjo vrednost kvadratne forme $\vec{x}^\mathsf{T} A\vec{x}$ pri pogoju $\vec{x}^\mathsf{T} B\vec{x} = b$.

ϟ (11) Naj bodo dani vektorji in matriki

$$P = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}, \vec{q} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \vec{r} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \text{ in } \vec{b} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Poiščite najmanjšo vrednost funkcije $\vec{x}^\mathsf{T} P\vec{x} + \vec{q}^\mathsf{T}\vec{x} + \vec{r}$ pri pogoju, da $\vec{x}$ reši linearni sistem $A\vec{x} = \vec{b}$.

(Naloge, označene s ϟ preverjajo razumevanje osnovnih pojmov in so primeri nalog s teoretičnih izpitov. Naloga, označena s ϟϟ je nekolika težja.)

Matematika 1	**Izročki: 12. in 13. teden**

Optimizacija. 3. del.
Prirejena funkcija. Karush-Kuhn-Tuckerjevi pogoji.

Polona Oblak

Naj bodo $f, g_i, h_j \colon \mathbb{R}^n \to \mathbb{R}$ dane funkcije več spremenljivk. Radi bi našli rešitev naslednjega problema

$$(P\star) \qquad \begin{aligned} &\text{minimizirajmo}_{\vec{x}}\ f(\vec{x}) \\ &\text{pri pogojih } g_i(\vec{x}) \leq 0 \text{ za } i = 1, 2, \ldots, m \\ &\qquad\qquad h_j(\vec{x}) = 0 \text{ za } j = 1, 2, \ldots, r. \end{aligned}$$

Definirajmo še množice $\mathcal{D}_{g_i} = \{x \in \mathbb{R}^n;\ g_i(x) \leq 0\}$ za $i = 1, 2, \ldots, m$, $\mathcal{D}_{h_j} = \{x \in \mathbb{R}^n;\ h_j(x) = 0\}$ za $j = 1, 2, \ldots, r$, in

$$\mathcal{D} = \mathcal{D}_f \cap (\cap_{i=1}^m \mathcal{D}_{g_i}) \cap \left(\cap_{j=1}^r \mathcal{D}_{h_j}\right).$$

Sedaj lahko problem ($P\star$) zapišemo ekvivalentno kot

$$(P\star) \qquad \min_{\vec{x} \in \mathcal{D}} f(\vec{x}).$$

Definirajmo Lagrangevo funkcijo

$$\begin{aligned} L(\vec{x}, \vec{\lambda}, \vec{\mu}) &= f(\vec{x}) - \vec{\lambda}^{\mathsf{T}} \vec{G}(\vec{x}) - \vec{\mu}^{\mathsf{T}} \vec{H}(\vec{x}) = \\ &= f(\vec{x}) - \sum_{i=1}^{m} \lambda_i g_i(\vec{x}) - \sum_{j=1}^{r} \mu_j h_j(\vec{x}), \end{aligned}$$

kjer je

$$\vec{G}(\vec{x}) = \begin{bmatrix} g_1(\vec{x}) \\ \vdots \\ g_m(\vec{x}) \end{bmatrix}, \vec{H}(\vec{x}) = \begin{bmatrix} h_1(\vec{x}) \\ \vdots \\ h_r(\vec{x}) \end{bmatrix}, \vec{\lambda} = \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_m \end{bmatrix}, \text{ in } \vec{\mu} = \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_r \end{bmatrix}.$$

Funkcijo

$$K(\vec{\lambda}, \vec{\mu}) = \inf_{\vec{x} \in \mathcal{D}} L(\vec{x}, \vec{\lambda}, \vec{\mu}) = \inf_{\vec{x} \in \mathcal{D}} \left\{ f(\vec{x}) - \vec{\lambda}^{\mathsf{T}} \vec{G}(\vec{x}) - \vec{\mu}^{\mathsf{T}} \vec{H}(\vec{x}) \right\}$$

imenujemo *prirejena funkcija* problema ($P\star$). Pri tem spremenljivke $\vec{\lambda}$ in $\vec{\mu}$ imenujemo *prirejene spremenljivke*. Opazimo:

(1) $K(\vec{\lambda}, \vec{\mu})$ je konkavna funkcija (neodvisno od lastnosti funkcij f, g_i, h_j originalnega problema).
(2) Če je $\lambda_i \leq 0$ za $i = 1, 2, \ldots, m$, potem velja $K(\vec{\lambda}, \vec{\mu}) \leq \min_{\vec{x} \in \mathcal{D}} f(x)$ za vse $\vec{\lambda} \leq \vec{0}$ ter vse $\vec{\mu}$.

Problem

$$(D\star)\qquad \begin{aligned} &\text{maksimizirajmo}_{\vec{\lambda},\vec{\mu}}\ K(\vec{\lambda},\vec{\mu})\\ &\text{pri pogoju}\ \ \lambda_i \le 0 \text{ za } i=1,2,\ldots,m \end{aligned}$$

imenujemo *prirejeni problem* problema ($P\star$).

Označimo z $\vec{x}^*$ vektor iz $\mathcal{D}$, ki reši problem ($P\star$) in z $\vec{\lambda}^*$ in $\vec{\mu}^*$ prirejeni spremenljivki, ki rešita prirejeni problem ($D\star$). Naj bo torej $p^* = f(\vec{x}^*)$ rešitev problema ($P\star$) in $d^* = K(\vec{\lambda}^*, \vec{\mu}^*)$ rešitev problema ($D\star$). Potem iz (2) sledi

$$d^* \le p^*.$$

Če

- je ($P\star$) linearni program (t,.j. f je linearna in h_j so afine funkcije), ali če
- so f, g_i konveksne funkcije in h_j afine,

potem velja $d^* = p^*$.

V primeru, ko je $d^* = p^*$, sledi, da morajo $\vec{x}^*$, $\vec{\lambda}^*$ in $\vec{\mu}^*$ zadoščati *Karush–Kuhn–Tuckerjevim pogojem*:

$$(\text{KKT})\qquad \begin{aligned} \frac{\partial L(\vec{x},\vec{\lambda}^*,\vec{\mu}^*)}{\partial \vec{x}}(\vec{x}^*) &= 0,\\ g_i(\vec{x}^*) &\le 0 \ \text{ za } i=1,2,\ldots,m,\\ h_j(\vec{x}^*) &= 0 \ \text{ za } j=1,2,\ldots,r,\\ \lambda_i^* &\le 0 \ \text{ za } i=1,2,\ldots,m,\\ \lambda_i^* g_i(\vec{x}^*) &= 0 \ \text{ za } i=1,2,\ldots,m. \end{aligned}$$

1. Nadaljnje branje

(1) Wilhelm Forst, Dieter Hoffmann: Optimization - Theory in Practice, Springer, 2010.

2. Domača naloga

ϟ (1) Območje A v $\mathbb{R}^n$ je konveksno, če je za poljubni točki $x, y \in A$ tudi točka $tx + (1-t)y \in A$. Pokažite, da je presek konveksnih množic konveksna množica.

ϟ (2) Naj bodo $g_i \colon \mathcal{D} \subseteq \mathbb{R}^n \to \mathbb{R}$, $i = 1,2,\ldots,m$, konveksne funkcije na konveksni množici $\mathcal{D}$. Pokažite, da je tedaj območje

$$\mathcal{D}_0 = \{x \in \mathcal{D};\ g_i(x) \le 0 \text{ za } i = 1,2,\ldots,m\}$$

konveksna množica.

ϟ (3) Rešite naslednji problem:

$$\begin{aligned} \text{minimizirajte}_{x,y} \quad & (x-2)^2 + 2(y-1)^2 \\ \text{pri pogojih} \quad & x + 4y \leq 3, \\ & y - x \leq 0. \end{aligned}$$

ϟ (4) Zapišite Karush–Kuhn–Tuckerjeve pogoje naslednjih problemov:

(a)

$$\begin{aligned} \text{minimizirajte}_{x,y,z} \quad & x + y^2 + z^2 \\ \text{pri pogojih} \quad & x^2 + 2y^2 \leq 4 \\ & x + y + z = 1. \end{aligned}$$

(b)

$$\begin{aligned} \text{minimizirajte}_{x,y,z} \quad & xyz \\ \text{pri pogojih} \quad & x^2 + 2y^2 \leq 4 \\ & x^2 + y^2 + z^2 = 3 \end{aligned}$$

(c)

$$\begin{aligned} \text{minimizirajte}_{x,y,z} \quad & 2x + 2y^2 + xz \\ \text{pri pogojih} \quad & x^2 + y^2 = 1 \\ & x + z = 0 \\ & xy \geq 0 \end{aligned}$$

(Naloge, označene s ϟ preverjajo razumevanje osnovnih pojmov in so primeri nalog s teoretičnih izpitov.)