Matematika 2 za fizike

Janez Mrčun

zapiski predavanj na FMF UL
2015-2024

Zadnja sprememba: 30. marec 2024

Kazalo

Poglavje 1

Matrike in determinanta

1.1. Matrike

Definicija 1.1. Naj bosta m in n naravni števili. *Matrika* kompleksnih števil velikosti $m \times n$ je urejena mn-terica A kompleksnih števil, katere komponente $A_{(i,j)} \in \mathbb{C}$ indeksiramo s pari $(i,j) \in \{1,2,\ldots,m\} \times \{1,2,\ldots,n\}$, zapišemo pa jo v obliki urejene pravokotne tabele

$$A = \begin{bmatrix} A_{(1,1)} & A_{(1,2)} & \cdots & A_{(1,n)} \\ A_{(2,1)} & A_{(2,2)} & \cdots & A_{(2,n)} \\ \vdots & \vdots & \ddots & \vdots \\ A_{(m,1)} & A_{(m,2)} & \cdots & A_{(m,n)} \end{bmatrix}$$

z m vrsticami in n stolpci.

Komentar 1.2. (1) Komponento $A_{(i,j)}$ takšne matrike A imenujemo (i,j)*-ta komponenta* matrike A in jo navadno krajše označimo z

$$A_{(i,j)} = A_{i,j} = A_{ij}.$$

Matrika A ima torej svojo (i,j)-to komponento A_{ij} zapisano v i-ti vrstici in j-tem stolpcu oziroma na *poziciji* (i,j). Matriko A na kratko zapišemo tudi kot

$$A = [A_{ij}]_{\substack{i=1,\ldots,m \\ j=1,\ldots,n}} = [A_{ij}].$$

Množico vseh matrik velikosti $m \times n$ kompleksnih števil označimo z

$$\mathrm{Mat}(m \times n, \mathbb{C}).$$

Matriki kompleksnih števil A na kratko rečemo tudi *kompleksna matrika*. Če so vse komponente matrike A realna (racionalna) števila, potem pravimo, da je A matrika realnih (racionalnih) števil oziroma da je A *realna (racionalna) matrika*. Množico vseh realnih matrik velikosti $m \times n$ označimo z $\mathrm{Mat}(m \times n, \mathbb{R})$, množico vseh racionalnih matrik velikosti $m \times n$ pa z $\mathrm{Mat}(m \times n, \mathbb{Q})$. Seveda velja

$$\mathrm{Mat}(m \times n, \mathbb{Q}) \subset \mathrm{Mat}(m \times n, \mathbb{R}) \subset \mathrm{Mat}(m \times n, \mathbb{C}).$$

Dostikrat bomo govorili o lastnostih, ki veljajo tako za kompleksne kot tudi za realne in racionalne matrike. V ta namen bomo z $\mathbb{F}$ označili obseg kompleksnih števil, obseg realnih števil ali obseg racionalnih števil. Števila iz izbranega obsega $\mathbb{F}$ imenujemo tudi *skalarji*. Krajše zapišemo tudi

$$\mathrm{Mat}(m \times n, \mathbb{F}) = \mathbb{F}^{m \times n}.$$

(2) Matrike velikosti $1 \times n$ imenujemo *vrstice*, matrike velikosti $m \times 1$ pa imenujemo *stolpci*. Za poljubno matriko $A \in \mathrm{Mat}(m \times n, \mathbb{F})$ označimo njeno i-to vrstico

$$A_{i\bullet} = \begin{bmatrix} A_{i1} & A_{i2} & \cdots & A_{in} \end{bmatrix} \in \mathrm{Mat}(1 \times n, \mathbb{F})$$

in njen j-ti stolpec

$$A_{\bullet j} = \begin{bmatrix} A_{1j} \\ A_{2j} \\ \vdots \\ A_{mj} \end{bmatrix} \in \mathrm{Mat}(m \times 1, \mathbb{F}),$$

za vse $i = 1, 2, \ldots, m$ in $j = 1, 2, \ldots, n$. Matrika A je *kvadratna*, če je $n = m$. Kvadratna matrika ima torej prav toliko vrstic kot stolpcev.

(3) Spomnimo se, da so elementi množice $\mathbb{F}^m$ urejene m-terice števil iz $\mathbb{F}$,

$$\mathbb{F}^m = \{(\alpha_1, \alpha_2, \ldots, \alpha_m) \; ; \; \alpha_1, \alpha_2, \ldots, \alpha_m \in \mathbb{F}\}.$$

Takšno urejeno m-terico $(\alpha_1, \alpha_2, \ldots, \alpha_m) \in \mathbb{F}^m$ standardno identificiramo s stolpcem, ki ima komponente $\alpha_1, \alpha_2, \ldots, \alpha_m$:

$$(\alpha_1, \alpha_2, \ldots, \alpha_m) = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_m \end{bmatrix}$$

V tem smislu torej velja $\mathbb{F}^m = \mathrm{Mat}(m \times 1, \mathbb{F})$. Posebej torej matrike velikosti 1×1 identificiramo s skalarji, $\mathbb{F} = \mathbb{F}^1 = \mathrm{Mat}(1 \times 1, \mathbb{F})$.

(4) Matrika A je *diagonalna*, če je kvadratna in je $A_{ij} = 0$ za vse i in j, za katere je $i \neq j$. Takšno diagonalno matriko A velikosti $n \times n$ lahko zapišemo na naslednji način:

$$A = \mathrm{diag}(A_{11}, A_{22}, \ldots, A_{nn}) = \begin{bmatrix} A_{11} & 0 & \cdots & 0 \\ 0 & A_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_{nn} \end{bmatrix}$$

(5) Matrika A je *simetrična*, če je kvadratna in je $A_{ij} = A_{ji}$ za vse i in j. Vsaka diagonalna matrika je simetrična. Matrika B je *antisimetrična*, če je kvadratna in je $B_{ij} = -B_{ji}$ za vse i in j. Za antisimetrično matriko B posebej torej velja $B_{ii} = 0$ za vse i.

(6) Matrika A je *zgornje-trikotna*, če je kvadratna in je $A_{ij} = 0$ za vse i in j, za katere je $i > j$. Zgornje-trikotna matrika A velikosti $n \times n$ je torej naslednje oblike:

$$A = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ 0 & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_{nn} \end{bmatrix}$$

Matrika B je *spodnje-trikotna*, če je kvadratna in je $B_{ij} = 0$ za vse i in j, za katere je $i < j$.

Zgled 1.3. (1) *Ničelna matrika* velikosti $m \times n$ je matrika

$$0 = 0_{m \times n} = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} \in \mathrm{Mat}(m \times n, \mathbb{F}),$$

ki ima vse komponente enake številu 0, torej $0_{ij} = 0$ za vse i in j.

(2) *Identična matrika* velikosti $n \times n$ je kvadratna matrika

$$\mathrm{I} = \mathrm{I}_{n\times n} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \in \mathrm{Mat}(n \times n, \mathbb{F}),$$

ki ima diagonalne komponente enake številu 1, vse ostale pa številu 0, torej

$$\mathrm{I}_{ij} = \begin{cases} 1 & ; \quad i = j \\ 0 & ; \quad i \neq j \end{cases}$$

za vse i in j. Identična matrika je torej diagonalna.

1.1.1. Računanje z matrikami. Za poljubni matriki iste velikosti $A, A' \in \mathrm{Mat}(m \times n, \mathbb{F})$ definiramo njuno *vsoto*, ki je matrika

$$A + A' \in \mathrm{Mat}(m \times n, \mathbb{F})$$

dana s predpisom $(A+A')_{ij} = A_{ij} + A'_{ij}$ za vse i in j. Podobno definiramo produkt poljubne matrike $A \in \mathrm{Mat}(m \times n, \mathbb{F})$ s poljubnim skalarjem $\alpha \in \mathbb{F}$, ki je matrika

$$\alpha A \in \mathrm{Mat}(m \times n, \mathbb{F})$$

dana s predpisom $(\alpha A)_{ij} = \alpha A_{ij}$ za vse i in j. Drugače povedano, vsota matrik in produkt matrike s skalarjem se izračunata z vsoto oziroma s produktom *po komponentah*.

Za poljubni matriki $A \in \mathrm{Mat}(m \times n, \mathbb{F})$ in $B \in \mathrm{Mat}(n \times p, \mathbb{F})$ definiramo njun *produkt*, ki je matrika

$$AB \in \mathrm{Mat}(m \times p, \mathbb{F})$$

dana s predpisom

$$(AB)_{ik} = \sum_{j=1}^{n} A_{ij} B_{jk}$$

za vse $i = 1, 2, \ldots, m$ in $k = 1, 2, \ldots, p$.

Zgled 1.4. Izračunamo lahko na primer:

$$\begin{bmatrix} 2 & 3 & 1 \\ 0 & 2 & 4 \end{bmatrix} + \begin{bmatrix} 0 & -1 & 1 \\ 1 & 1 & -1 \end{bmatrix} = \begin{bmatrix} 2 & 2 & 2 \\ 1 & 3 & 3 \end{bmatrix}$$

$$(1+\mathrm{i}) \begin{bmatrix} 2 & \mathrm{i} \\ -1+\mathrm{i} & -3\mathrm{i} \end{bmatrix} = \begin{bmatrix} 2+2\mathrm{i} & -1+\mathrm{i} \\ -2 & 3-3\mathrm{i} \end{bmatrix}$$

$$\begin{bmatrix} 1 & 3 & -1 \\ 7 & -2 & 5 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & -1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & -2 \\ 2 & 21 \end{bmatrix}$$

Definirali smo torej tri operacije na matrikah: *seštevanje matrik* (ki je dano z vsoto matrik), *množenje matrik s skalarji* (ki je dano s produktom matrike s skalarjem) in *množenje matrik* (ki je dano s produktom matrik). Ogledali si bomo osnovne lastnosti teh operacij. Iz definicije direktno sledi:

Trditev 1.5. *Za vse matrike $A, A', A'' \in \mathrm{Mat}(m \times n, \mathbb{F})$ in vse $\alpha, \alpha' \in \mathbb{F}$ velja*

(i) $(A + A') + A'' = A + (A' + A'')$,
(ii) $A + A' = A' + A$,
(iii) $A + 0 = A$,

(iv) $A + (-1)A = 0$,
(v) $\alpha(A + A') = \alpha A + \alpha A'$,
(vi) $(\alpha + \alpha')A = \alpha A + \alpha' A$,
(vii) $(\alpha\alpha')A = \alpha(\alpha' A)$ *in*
(viii) $1A = A$.

KOMENTAR 1.6. Lastnosti (i) pravimo *asociativnost* seštevanja, lastnosti (ii) pa *komutativnost* seštevanja. Točka (iii) pove, da je ničelna matrika *nevtralni element* za seštevanje. Po točki (iv) je matrika $-A = (-1)A$ *nasprotna* matrika matrike A. Označimo tudi operacijo *odštevanja* matrik, ki matrikama $A, A' \in \mathrm{Mat}(m \times n, \mathbb{F})$ priredi njuno *razliko* $A - A' = A + (-1)A'$. Ker za seštevanje veljajo lastnosti (i-iv), pravimo, da je množica $\mathrm{Mat}(m \times n, \mathbb{F})$ skupaj z operacijo seštevanja *Abelova grupa*.

Ker imamo na Abelovi grupi $\mathrm{Mat}(m \times n, \mathbb{F})$ definirano tudi operacijo množenja s skalarji, za katero veljajo lastnosti (v-viii), Abelovi grupi $\mathrm{Mat}(m \times n, \mathbb{F})$ pravimo *vektorski prostor* nad $\mathbb{F}$. Posebej je $\mathbb{F}^m = \mathrm{Mat}(m \times 1, \mathbb{F})$ vektorski prostor nad $\mathbb{F}$.

TRDITEV 1.7. *Za vse matrike* $A \in \mathrm{Mat}(m \times n, \mathbb{F})$, $B, B' \in \mathrm{Mat}(n \times p, \mathbb{F})$ *in* $C \in \mathrm{Mat}(p \times q, \mathbb{F})$ *ter za vsak* $\alpha \in \mathbb{F}$ *velja*
(i) $(AB)C = A(BC)$,
(ii) $\mathrm{I}A = A\,\mathrm{I} = A$,
(iii) $A(B + B') = AB + AB'$,
(iv) $(B + B')C = BC + B'C$ *in*
(v) $\alpha(AB) = (\alpha A)B = A(\alpha B)$.

DOKAZ. (i) Za vse i in l velja:

$$\begin{aligned}((AB)C)_{il} &= \sum_{k=1}^{p} (AB)_{ik} C_{kl} \\ &= \sum_{k=1}^{p} (\sum_{j=1}^{n} A_{ij} B_{jk}) C_{kl} \\ &= \sum_{j=1}^{n} A_{ij} (\sum_{k=1}^{p} B_{jk} C_{kl}) \\ &= \sum_{j=1}^{n} A_{ij} (BC)_{jl} = (A(BC))_{il}\end{aligned}$$

(ii) Za vse i in j velja

$$(\mathrm{I}A)_{ij} = \sum_{r=1}^{m} \mathrm{I}_{ir} A_{rj} = A_{ij},$$

saj je po definiciji identične matrike $\mathrm{I}_{ii} = 1$ in $\mathrm{I}_{ir} = 0$ za $r \neq i$. Prav tako vidimo, da je

$$(A\,\mathrm{I})_{ij} = \sum_{s=1}^{n} A_{is} \mathrm{I}_{sj} = A_{ij}.$$

Točke (iii-v) dokažemo podobno. □

KOMENTAR 1.8. (1) Lastnost (i) iz trditve pravi, da je množenje matrik asociativno, medtem ko lastnost (ii) pove, da je identična matrika nevtralni element za množenje matrik.

Operacija množenje matrik v splošnem ni komutativna. Za dani matriki $A \in \mathrm{Mat}(m \times n, \mathbb{F})$ in $B \in \mathrm{Mat}(n \times p, \mathbb{F})$ je produkt AB matrika velikosti $m \times p$, po drugi strani pa je produkt BA definiran le v primeru, da je $p = m$, in tedaj je BA matrika velikosti $n \times n$. A tudi v primeru, da je $m = n = p > 1$, lahko matriki A in B izberemo tako, da sta matriki AB in BA različni, na primer:

$$\begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}$$

(2) Produkt kvadratnih matrik velikosti $n \times n$ je spet kvadratna matrika velikosti $n \times n$. Množenje matrik je torej operacija na množici kvadratnih matrik $\mathrm{Mat}(n \times n, \mathbb{F})$. Ker je $\mathrm{Mat}(n \times n, \mathbb{F})$ vektorski prostor nad $\mathbb{F}$ in za množenje matrik veljajo lastnosti iz trditve 1.7, pravimo, da je $\mathrm{Mat}(n \times n, \mathbb{F})$ *algebra* nad $\mathbb{F}$.

Zgled 1.9. Produkt diagonalnih matrik velikosti $n \times n$ je prav tako diagonalna matrika velikosti $n \times n$ in velja

$$\mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)\,\mathrm{diag}(\mu_1, \mu_2, \ldots, \mu_n) = \mathrm{diag}(\lambda_1\mu_1, \lambda_2\mu_2, \ldots, \lambda_n\mu_n).$$

Odtod tudi vidimo, da diagonalne matrike enakih si velikosti med seboj komutirajo: če sta D in D' diagonalni matriki iste velikosti, potem je $DD' = D'D$.

1.1.2. Obrnljive matrike. Matrika $B \in \mathrm{Mat}(n \times n, \mathbb{F})$ je *inverz* matrike $A \in \mathrm{Mat}(n \times n, \mathbb{F})$, če velja $AB = BA = \mathrm{I}$. Matrika A ima lahko največ en inverz: če sta namreč B in B' dva inverza matrike A, potem velja

$$B = \mathrm{I}B = (B'A)B = B'(AB) = B'\mathrm{I} = B'.$$

Če ima kvadratna matrika A inverz, potem pravimo, da je *obrnljiva*, inverz matrike A pa označimo z

$$A^{-1}.$$

Matrika A^{-1} je torej kvadratna matrika enake velikosti kot matrika A in zanjo velja $AA^{-1} = A^{-1}A = \mathrm{I}$.

Zgled 1.10. (1) Najenostavnejši primer kvadratne matrike, ki ni obrnljiva, je kvadratna ničelna matrika. Tudi matrika

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

ni obrnljiva. Za poljubno matriko B velikosti 2×2 namreč lahko izračunamo, da je $(AB)_{11} = (AB)_{21}$, to pa že pomeni, da $AB \neq \mathrm{I}$. Matrika A torej ni obrnljiva.

(2) Identična matrika je obrnljiva in je sama sebi inverz, $\mathrm{I}^{-1} = \mathrm{I}$.

Trditev 1.11. *Diagonalna matrika*

$$A = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n) \in \mathrm{Mat}(n \times n, \mathbb{F})$$

je obrnljiva, če, in samo če, velja $\lambda_1\lambda_2\cdots\lambda_n \neq 0$. *V tem primeru je*

$$A^{-1} = \mathrm{diag}(\lambda_1^{-1}, \lambda_2^{-1}, \ldots, \lambda_n^{-1}).$$

Dokaz. Če je $\lambda_1\lambda_2\cdots\lambda_n \neq 0$, je matrika $\operatorname{diag}(\lambda_1^{-1}, \lambda_2^{-1}, \ldots, \lambda_n^{-1})$ res inverz matrike A, saj je

$$\begin{aligned}\operatorname{diag}(\lambda_1^{-1}, \lambda_2^{-1}, \ldots, \lambda_n^{-1})A &= A\operatorname{diag}(\lambda_1^{-1}, \lambda_2^{-1}, \ldots, \lambda_n^{-1})\\ &= \operatorname{diag}(\lambda_1\lambda_1^{-1}, \lambda_2\lambda_2^{-1}, \ldots, \lambda_n\lambda_n^{-1})\\ &= \operatorname{diag}(1, 1, \ldots, 1) = \mathrm{I}.\end{aligned}$$

Obratno, pa predpostavimo, da je $\lambda_i = 0$ za nek i. Za poljubno matriko B velikosti $n \times n$ je

$$(AB)_{ii} = (\operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)B)_{ii} = \lambda_i B_{ii} = 0,$$

zato matrika B ni inverz matrike A. Inverz matrike A torej ne obstaja. □

Trditev 1.12. *Na bosta matriki $A, B \in \operatorname{Mat}(n \times n, \mathbb{F})$ obrnljivi. Tedaj sta tudi matriki AB in A^{-1} obrnljivi in velja*

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

ter

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

Dokaz. Z uporabo asociativnosti množenja lahko preverimo, da velja

$$(B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}A)B = B^{-1}\mathrm{I}B = \mathrm{I}$$

in

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = A\mathrm{I}A^{-1} = \mathrm{I}.$$

Preostali del trditve sledi iz enakosti $A^{-1}A = AA^{-1} = \mathrm{I}$. □

Komentar 1.13. Množico obrnljivih matrik velikosti $n \times n$ s komponentami v $\mathbb{F}$ označimo z

$$\mathrm{GL}(n, \mathbb{F}).$$

Iz trditve sledi, da je množenje matrik operacija na množici $\mathrm{GL}(n, \mathbb{F})$. Vemo že, da je operacija množenja asociativna in da je identična matrika $\mathrm{I} \in \mathrm{GL}(n, \mathbb{F})$ nevtralni element za množenje. Iz trditve tudi sledi, da ima vsak element množice $\mathrm{GL}(n, \mathbb{F})$ svoj inverz v $\mathrm{GL}(n, \mathbb{F})$. Zaradi teh lastnosti pravimo, da je $\mathrm{GL}(n, \mathbb{F})$ z operacijo množenja matrik *grupa*. Grupo $\mathrm{GL}(n, \mathbb{F})$ imenujemo tudi *splošna linearna grupa* stopnje n. Za $n > 1$ grupa $\mathrm{GL}(n, \mathbb{F})$ ni Abelova.

1.1.3. Transponiranje, konjugiranje in hermitiranje matrik. Naj bo $A \in \operatorname{Mat}(m \times n, \mathbb{F})$. *Transponirana matrika* matrike A je matrika

$$A^{\mathsf{t}} \in \operatorname{Mat}(n \times m, \mathbb{F}),$$

dana s predpisom $(A^{\mathsf{t}})_{ji} = A_{ij}$ za vse i in j. *Konjugirana matrika* matrike A je matrika

$$\bar{A} \in \operatorname{Mat}(m \times n, \mathbb{F}),$$

ki jo dobimo tako, da konjugiramo vse komponente matrike A, torej $(\bar{A})_{ij} = \overline{A_{ij}}$ za vse i in j. *Hermitirana matrika* matrike A je matrika

$$A^{\mathsf{h}} = \bar{A}^{\mathsf{t}} \in \operatorname{Mat}(n \times m, \mathbb{F}).$$

Zgled 1.14. (1) Opazimo lahko, da je matrika A realna, če, in samo če, velja $\bar{A} = A$ oziroma, če, in samo če, je $A^{\mathsf{h}} = A^{\mathsf{t}}$. Transponiranje stolpce preslika v vrstice in obratno.

$$\begin{bmatrix} 2 & 3 & 5 \\ 0 & 1 & 7 \end{bmatrix}^{\mathsf{t}} = \begin{bmatrix} 2 & 0 \\ 3 & 1 \\ 5 & 7 \end{bmatrix}$$

$$\begin{bmatrix} 1+\mathrm{i} \\ -2 \end{bmatrix}^{\mathsf{h}} = \begin{bmatrix} 1-\mathrm{i} & -2 \end{bmatrix}$$

Če transponiramo, konjugiramo ali hermitiramo ničelno matriko, je rezultat spet ničelna matrika. Podobno velja za identično matriko,

$$\mathrm{I}^{\mathsf{t}} = \bar{\mathrm{I}} = \mathrm{I}^{\mathsf{h}} = \mathrm{I}.$$

(2) Iz definicije transponiranja tudi sledi, da je matrika A simetrična, če, in samo če, velja enakost $A^{\mathsf{t}} = A$. Še več, matrika B je antisimetrična, če, in samo če, velja enakost $B^{\mathsf{t}} = -B$.

(3) Matriko A, za katero velja $A^{\mathsf{h}} = A$, imenujemo *hermitska* matrika. Če za matriko B velja enakost $B^{\mathsf{h}} = -B$, potem pravimo, da je matrika B *antihermitska*.

Direktno iz definicije sledi, da za transponiranje, konjugiranje in hermitiranje matrik veljajo naslednje lastnosti:

Trditev 1.15. *Za vse matrike* $A, A' \in \mathrm{Mat}(m \times n, \mathbb{F})$ *in* $B \in \mathrm{Mat}(n \times p, \mathbb{F})$ *ter za vsak* $\alpha \in \mathbb{F}$ *velja*

(i) $(A^{\mathsf{t}})^{\mathsf{t}} = A$,
(ii) $(A + A')^{\mathsf{t}} = A^{\mathsf{t}} + A'^{\mathsf{t}}$,
(iii) $(\alpha A)^{\mathsf{t}} = \alpha A^{\mathsf{t}}$,
(iv) $(AB)^{\mathsf{t}} = B^{\mathsf{t}} A^{\mathsf{t}}$,
(v) $\bar{\bar{A}} = A$,
(vi) $\overline{A + A'} = \bar{A} + \overline{A'}$,
(vii) $\overline{\alpha A} = \bar{\alpha}\bar{A}$,
(viii) $\overline{AB} = \bar{A}\bar{B}$,
(ix) $(A^{\mathsf{h}})^{\mathsf{h}} = A$,
(x) $(A + A')^{\mathsf{h}} = A^{\mathsf{h}} + A'^{\mathsf{h}}$,
(xi) $(\alpha A)^{\mathsf{h}} = \bar{\alpha} A^{\mathsf{h}}$,
(xii) $(AB)^{\mathsf{h}} = B^{\mathsf{h}} A^{\mathsf{h}}$,
(xiii) $\overline{A^{\mathsf{t}}} = \bar{A}^{\mathsf{t}}$,
(xiv) $(A^{\mathsf{t}})^{\mathsf{h}} = (A^{\mathsf{h}})^{\mathsf{t}} = \bar{A}$ *in*
(xv) $(\bar{A})^{\mathsf{h}} = \overline{A^{\mathsf{h}}} = A^{\mathsf{t}}$.

Dokaz. (iv) Za vse i in k velja:

$$\begin{aligned} ((AB)^{\mathsf{t}})_{ki} &= (AB)_{ik} = \sum_{j=1}^{n} A_{ij} B_{jk} \\ &= \sum_{j=1}^{n} (A^{\mathsf{t}})_{ji} (B^{\mathsf{t}})_{kj} = \sum_{j=1}^{n} (B^{\mathsf{t}})_{kj} (A^{\mathsf{t}})_{ji} \\ &= (B^{\mathsf{t}} A^{\mathsf{t}})_{ki} \end{aligned}$$

Ostale točke pokažemo na podoben način. □

TRDITEV 1.16. *Za vsako obrnljivo matriko* $A \in \mathrm{GL}(n, \mathbb{F})$ *so tudi matrike* A^{t}, $\bar{A}$ *in* A^{h} *obrnljive in velja:*

$$(A^{\mathsf{t}})^{-1} = (A^{-1})^{\mathsf{t}}$$
$$\bar{A}^{-1} = \overline{A^{-1}}$$
$$(A^{\mathsf{h}})^{-1} = (A^{-1})^{\mathsf{h}}$$

DOKAZ. Trditev je posledica lastnosti iz trditve 1.15. Na primer, za hermitirano matriko velja

$$(A^{-1})^{\mathsf{h}} A^{\mathsf{h}} = (AA^{-1})^{\mathsf{h}} = \mathrm{I}^{\mathsf{h}} = \mathrm{I},$$
$$A^{\mathsf{h}} (A^{-1})^{\mathsf{h}} = (A^{-1}A)^{\mathsf{h}} = \mathrm{I}^{\mathsf{h}} = \mathrm{I}.$$

Na podoben način dokažemo trditev tudi za transponirano in za konjugirano matriko. □

1.1.4. Podmatrike in bločni zapis matrike. Naj bo $A \in \mathrm{Mat}(m \times n, \mathbb{F})$. *Podmatrika* matrike A je matrika, ki jo dobimo iz matrike A tako, da iz nje izbrišemo nekaj vrstic in nekaj stolpcev. Pri tem ne smemo izbrisati vseh vrstic ali vseh stolpcev, saj v tem primeru za rezultat ne bi dobili matrike, dopustno pa je, da ne izbrišemo nobene vrstice oziroma nobenega stolpca.

Bolj natančno, izberimo poljubni neprazni podmnožici $\mathcal{I} = \{i_1, i_2, \ldots, i_{m'}\} \subset \{1, 2, \ldots, m\}$, $i_1 < i_2 < \cdots < i_{m'}$, in $\mathcal{J} = \{j_1, j_2, \ldots, j_{n'}\} \subset \{1, 2, \ldots, n\}$, $j_1 < j_2 < \cdots < j_{n'}$. Podmatrika matrike A, sestavljena iz vrstic z indeksi iz $\mathcal{I}$ in iz stolpcev z indeksi iz $\mathcal{J}$, je matrika

$$\mathrm{sub}_{\mathcal{I}\mathcal{J}}(A) \in \mathrm{Mat}(m' \times n', \mathbb{F})$$

dana s predpisom

$$(\mathrm{sub}_{\mathcal{I}\mathcal{J}}(A))_{kl} = A_{i_k j_l}$$

za vse $k = 1, 2, \ldots, m'$ in $l = 1, 2, \ldots, n'$.

Še posebej pomemben primer je naslednji: Predpostavimo, da velja $m, n \geq 2$, in izberimo poljubni naravni števili $i \leq m$ ter $j \leq n$. Podmatriko matrike A, ki jo dobimo tako, da iz matrike A izbrišemo i-to vrstico in j-ti stolpec, označimo z

$$\mathrm{sub}^{ij}(A) \in \mathrm{Mat}((m-1) \times (n-1), \mathbb{F}).$$

Velja torej

$$\mathrm{sub}^{ij}(A) = \mathrm{sub}_{\{1,\ldots,i-1,i+1,\ldots,m\}\{1,\ldots,j-1,j+1,\ldots,n\}}(A).$$

ZGLED 1.17. (1) Naj bo

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

Primeri podmatrik matrike A so tedaj

$$\mathrm{sub}_{\{1,2\}\{2,4\}}(A) = \begin{bmatrix} 2 & 4 \\ 7 & 5 \end{bmatrix},$$

$$\mathrm{sub}^{12}(A) = \mathrm{sub}_{\{2\}\{1,3,4\}}(A) = \begin{bmatrix} 8 & 6 & 5 \end{bmatrix}$$

in

$$\mathrm{sub}_{\{1\}\{2\}}(A) = \begin{bmatrix} 2 \end{bmatrix}.$$

Naj bo $A \in \mathrm{Mat}(m \times n, \mathbb{F})$. Izberimo si naravna števila $m_1, m_2, \dots, m_{\mathrm{M}}$, za katera velja

$$m = m_1 + m_2 + \cdots + m_{\mathrm{M}}.$$

Pri tem je M naravno število, $1 \leq \mathrm{M} \leq m$. Število m smo s tem razdelili na M delov. Izberimo še naravna števila $n_1, n_2, \dots, n_{\mathrm{N}}$, ki nam število n razdelijo na N delov, torej

$$n = n_1 + n_2 + \cdots + n_{\mathrm{N}}.$$

Za poenostavitev zapisa označimo še $\mu_r = m_1 + \cdots + m_{r-1}$ za vse $r = 1, \dots, \mathrm{M}$ in $\nu_s = n_1 + \cdots + n_{s-1}$ za vse $s = 1, \dots, \mathrm{N}$. Posebej je $\mu_1 = \nu_1 = 0$.

Z izbiro razdelitev $m = m_1 + m_2 + \cdots + m_{\mathrm{M}}$ in $n = n_1 + n_2 + \cdots + n_{\mathrm{N}}$ razdelimo matriko A na NM *blokov*

$$A = \left[\begin{array}{ccc|c|ccc}
A_{11} & \cdots & A_{1n_1} & \cdots & A_{1(\nu_{\mathrm{N}}+1)} & \cdots & A_{1n} \\
\vdots & & \vdots & & \vdots & & \vdots \\
A_{m_1 1} & \cdots & A_{m_1 n_1} & \cdots & A_{m_1(\nu_{\mathrm{N}}+1)} & \cdots & A_{m_1 n} \\
\hline
\vdots & & \vdots & \ddots & \vdots & & \vdots \\
\hline
A_{(\mu_{\mathrm{M}}+1)1} & \cdots & A_{(\mu_{\mathrm{M}}+1)n_1} & \cdots & A_{(\mu_{\mathrm{M}}+1)(\nu_{\mathrm{N}}+1)} & \cdots & A_{(\mu_{\mathrm{M}}+1)n} \\
\vdots & & \vdots & & \vdots & & \vdots \\
A_{m1} & \cdots & A_{mn_1} & \cdots & A_{m(\nu_{\mathrm{N}}+1)} & \cdots & A_{mn}
\end{array}\right]$$

$$= \begin{bmatrix}
F(1,1) & F(1,2) & \cdots & F(1,\mathrm{N}) \\
F(2,1) & F(2,2) & \cdots & F(2,\mathrm{N}) \\
\vdots & \vdots & \ddots & \vdots \\
F(\mathrm{M},1) & F(\mathrm{M},2) & \cdots & F(\mathrm{M},\mathrm{N})
\end{bmatrix}$$

kjer je

$$F(r,s) = \begin{bmatrix}
A_{(\mu_r+1)(\nu_s+1)} & A_{(\mu_r+1)(\nu_s+2)} & \cdots & A_{(\mu_r+1)(\nu_s+n_s)} \\
A_{(\mu_r+2)(\nu_s+1)} & A_{(\mu_r+2)(\nu_s+2)} & \cdots & A_{(\mu_r+2)(\nu_s+n_s)} \\
\vdots & \vdots & \ddots & \vdots \\
A_{(\mu_r+m_r)(\nu_s+1)} & A_{(\mu_r+m_r)(\nu_s+2)} & \cdots & A_{(\mu_r+m_r)(\nu_s+n_s)}
\end{bmatrix}$$

za vse $r = 1, 2, \dots, \mathrm{M}$ in $s = 1, 2, \dots, \mathrm{N}$. Blok $F(r,s)$ je podmatrika matrike A velikosti $m_r \times n_s$, dana s predpisom

$$F(r,s)_{ij} = A_{(\mu_r+i)(\nu_s+j)}$$

za vse i in j.

Dobljeni zapis

$$A = \begin{bmatrix}
F(1,1) & F(1,2) & \cdots & F(1,\mathrm{N}) \\
F(2,1) & F(2,2) & \cdots & F(2,\mathrm{N}) \\
\vdots & \vdots & \ddots & \vdots \\
F(\mathrm{M},1) & F(\mathrm{M},2) & \cdots & F(\mathrm{M},\mathrm{N})
\end{bmatrix}$$

imenujemo *bločni zapis* matrike A glede na razdelitev vrstic $m = m_1 + m_2 + \cdots + m_{\mathrm{M}}$ in razdelitev stolpcev $n = n_1 + n_2 + \cdots + n_{\mathrm{N}}$.

Bločni zapis matrik nam dostikrat poenostavi zapis matrik in računanje z njimi. Z matrikami v bločnem zapisu namreč lahko računamo bločno. Najprej lahko opazimo, da je za poljuben skalar $\alpha \in \mathbb{F}$

$$\alpha A = \begin{bmatrix} \alpha F(1,1) & \alpha F(1,2) & \cdots & \alpha F(1,\textsc{n}) \\ \alpha F(2,1) & \alpha F(2,2) & \cdots & \alpha F(2,\textsc{n}) \\ \vdots & \vdots & \ddots & \vdots \\ \alpha F(\textsc{m},1) & \alpha F(\textsc{m},2) & \cdots & \alpha F(\textsc{m},\textsc{n}) \end{bmatrix}$$

bločni zapis matrike αA glede na razdelitvi $m = m_1 + m_2 + \cdots + m_{\textsc{m}}$ in $n = n_1 + n_2 + \cdots + n_{\textsc{n}}$. Naj bo $A' \in \mathrm{Mat}(m \times n, \mathbb{F})$ še ena matrika in naj bo

$$A' = \begin{bmatrix} F'(1,1) & F'(1,2) & \cdots & F'(1,\textsc{n}) \\ F'(2,1) & F'(2,2) & \cdots & F'(2,\textsc{n}) \\ \vdots & \vdots & \ddots & \vdots \\ F'(\textsc{m},1) & F'(\textsc{m},2) & \cdots & F'(\textsc{m},\textsc{n}) \end{bmatrix}$$

njen bločni zapis glede na razdelitvi $m = m_1 + m_2 + \cdots + m_{\textsc{m}}$ in $n = n_1 + n_2 + \cdots + n_{\textsc{n}}$. Tedaj je očitno

$$A + A' = \begin{bmatrix} F(1,1) + F'(1,1) & F(1,2) + F'(1,2) & \cdots & F(1,\textsc{n}) + F'(1,\textsc{n}) \\ F(2,1) + F'(2,1) & F(2,2) + F'(2,2) & \cdots & F(2,\textsc{n}) + F'(2,\textsc{n}) \\ \vdots & \vdots & \ddots & \vdots \\ F(\textsc{m},1) + F'(\textsc{m},1) & F(\textsc{m},2) + F'(\textsc{m},2) & \cdots & F(\textsc{m},\textsc{n}) + F'(\textsc{m},\textsc{n}) \end{bmatrix}$$

bločni zapis matrike $A + A'$ glede na razdelitvi $m = m_1 + m_2 + \cdots + m_{\textsc{m}}$ in $n = n_1 + n_2 + \cdots + n_{\textsc{n}}$.

Podobno pa velja tudi za množenje. Naj bo $B \in \mathrm{Mat}(n \times p, \mathbb{F})$, izberimo razdelitev $p = p_1 + p_2 + \cdots + p_{\textsc{p}}$ in naj bo

$$B = \begin{bmatrix} G(1,1) & G(1,2) & \cdots & G(1,\textsc{p}) \\ G(2,1) & G(2,2) & \cdots & G(2,\textsc{p}) \\ \vdots & \vdots & \ddots & \vdots \\ G(\textsc{n},1) & G(\textsc{n},2) & \cdots & G(\textsc{n},\textsc{p}) \end{bmatrix}$$

bločni zapis matrike B glede na razdelitvi $n = n_1 + n_2 + \cdots + n_{\textsc{n}}$ in $p = p_1 + p_2 + \cdots + p_{\textsc{p}}$. Produkt AB lahko tedaj zapišemo bločno kot

$$AB = \begin{bmatrix} H(1,1) & H(1,2) & \cdots & H(1,\textsc{p}) \\ H(2,1) & H(2,2) & \cdots & H(2,\textsc{p}) \\ \vdots & \vdots & \ddots & \vdots \\ H(\textsc{m},1) & H(\textsc{m},2) & \cdots & H(\textsc{m},\textsc{p}) \end{bmatrix}$$

glede na razdelitvi $m = m_1 + n_2 + \cdots + m_{\textsc{m}}$ in $p = p_1 + p_2 + \cdots + p_{\textsc{p}}$.

Ob teh predpostavkah in oznakah lahko izračunamo, da je

$$H(r,q) = \sum_{s=1}^{\textsc{n}} F(r,s)G(s,q). \tag{1.1}$$

za vse $r = 1, 2, \ldots, \textsc{m}$ in $q = 1, 2, \ldots, \textsc{p}$. Ta enakost nam pove, da lahko bločne matrike tudi bločno množimo na enak način kot množimo matrike.

Pokažimo, da velja enakost (1.1). Označimo $\psi_q = p_1 + p_2 + \cdots + p_{q-1}$, za vse $q = 1, 2, \ldots, \mathrm{P}$, in $\psi_1 = 0$. Za poljubne i in k izračunamo

$$\begin{aligned} H(r,q)_{ik} &= (AB)_{(\mu_r+i)(\psi_q+k)} = \sum_{j=1}^{n} A_{(\mu_r+i)j} B_{j(\psi_q+k)} \\ &= \sum_{s=1}^{\mathrm{N}} \sum_{l=1}^{n_s} A_{(\mu_r+i)(\nu_s+l)} B_{(\nu_s+l)(\psi_q+k)} \\ &= \sum_{s=1}^{\mathrm{N}} \sum_{l=1}^{n_s} F(r,s)_{il} G(s,q)_{lk} \\ &= \sum_{s=1}^{\mathrm{N}} (F(r,s)G(s,q))_{ik}. \end{aligned}$$

Zgled 1.18. Z vertikalnimi in horizontalnimi črtami bločno razdelimo matrike v enačbi:

$$\left[\begin{array}{rr|rr} 0 & 2 & 1 & 3 \\ 1 & -1 & 0 & 1 \\ \hline 2 & 1 & 0 & 1 \end{array}\right] \left[\begin{array}{rr} 0 & 5 \\ 2 & 2 \\ \hline 1 & -1 \\ 0 & 1 \end{array}\right] = \left[\begin{array}{rr} 5 & 6 \\ -2 & 4 \\ \hline 2 & 13 \end{array}\right]$$

To pomeni, da smo za vrstice prve matrike uporabili razdelitev $3 = 2 + 1$, za stolpce prve matrike in za vrstice druge matrike smo uporabili razdelitev $4 = 2 + 2$, za stolpce druge matrike pa imamo trivialno razdelitev $2 = 2$. Enakost (1.1) pove, da veljata tudi enakosti

$$\begin{bmatrix} 0 & 2 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 0 & 5 \\ 2 & 2 \end{bmatrix} + \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 5 & 6 \\ -2 & 4 \end{bmatrix}$$

in

$$\begin{bmatrix} 2 & 1 \end{bmatrix} \begin{bmatrix} 0 & 5 \\ 2 & 2 \end{bmatrix} + \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 13 \end{bmatrix},$$

ki jih seveda lahko tudi direktno preverimo.

1.1.5. Elementarne vrstične operacije in elementarne matrike. Dano matriko lahko preoblikujemo tako, da naredimo eno od naslednjih operacij na njenih vrsticah:

(1) Pomnožimo i-to vrstico matrike z neničelnim skalarjem $\alpha \in \mathbb{F} \setminus \{0\}$. To operacijo zapišemo s simbolom $(\mathscr{V}_i \leftarrow \alpha \mathscr{V}_i)$.
(2) Prištejemo β-kratnik j-te vrstice k i-ti vrstici, kjer je $\beta \in \mathbb{F}$ in $i \neq j$. Za to operacijo uporabljamo simbol $(\mathscr{V}_i \leftarrow \mathscr{V}_i + \beta \mathscr{V}_j)$.
(3) Zamenjamo i-to in j-to vrstico, kjer je $i \neq j$. Simbol za to operacijo je $(\mathscr{V}_i \leftrightarrow \mathscr{V}_j)$.

Naštete operacije imenujemo *elementarne vrstične operacije*.

Vsaki elementarni vrstični operaciji pridružimo pripadajočo kvadratno matriko poljubne izbrane velikosti $m \times m$ na naslednji način:

(i) Za vsak $i \leq m$ in za vsak $\alpha \in \mathbb{F} \setminus \{0\}$ je

$$[\mathscr{V}_i \leftarrow \alpha \mathscr{V}_i] = \begin{bmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & \alpha & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{bmatrix}$$

diagonalna matrika, katere (i,i)-ta komponenta je enaka skalarju α, ostale komponente pa so enake kot pri identični matriki (vse neoznačene komponente izven diagonale so enake številu 0).

Opazimo lahko, da je za poljubno matriko $A \in \mathrm{Mat}(m \times n, \mathbb{F})$ produkt matrik $[\mathscr{V}_i \leftarrow \alpha \mathscr{V}_i]A$ enak matriki, ki jo dobimo, če na matriki A uporabimo elementarno vrstično operacijo $(\mathscr{V}_i \leftarrow \alpha \mathscr{V}_i)$.

(ii) Za vse $i \leq m$ in $j \leq m$, $i \neq j$, in za vsak $\beta \in \mathbb{F}$ je

$$[\mathscr{V}_i \leftarrow \mathscr{V}_i + \beta \mathscr{V}_j] = \begin{bmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & \beta & & \\ & & & \ddots & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{bmatrix}$$

kvadratna matrika, ki ima (i,j)-to komponento enako skalarju β, vse ostale komponente pa so enake kot pri identični matriki (vse neoznačene komponente izven diagonale so enake številu 0).

Če na matriki $A \in \mathrm{Mat}(m \times n, \mathbb{F})$ uporabimo elementarno vrstično operacijo $(\mathscr{V}_i \leftarrow \mathscr{V}_i + \beta \mathscr{V}_j)$, je dobljena matrika enaka matričnemu produktu $[\mathscr{V}_i \leftarrow \mathscr{V}_i + \beta \mathscr{V}_j]A$.

(iii) Za vse $i \leq m$ in $j \leq m$, $i \neq j$, je

$$[\mathscr{V}_i \leftrightarrow \mathscr{V}_j] = \begin{bmatrix} 1 & & & & & & & & & & \\ & \ddots & & & & & & & & & \\ & & 1 & & & & & & & & \\ & & & 0 & & & & 1 & & & \\ & & & & 1 & & & & & & \\ & & & & & \ddots & & & & & \\ & & & & & & 1 & & & & \\ & & & 1 & & & & 0 & & & \\ & & & & & & & & 1 & & \\ & & & & & & & & & \ddots & \\ & & & & & & & & & & 1 \end{bmatrix}$$

kvadratna matrika, v kateri sta (i,i)-ta komponenta in (j,j)-ta komponenta enaki številu 0, (i,j)-ta komponenta in (j,i)-ta komponenta sta enaki številu 1, vse ostale komponente pa so enake kot pri identični matriki (neoznačene komponente izven diagonale so enake številu 0).

Če na matriki $A \in \mathrm{Mat}(m \times n, \mathbb{F})$ uporabimo elementarno vrstično operacijo $(\mathscr{V}_i \leftrightarrow \mathscr{V}_j)$, je dobljena matrika enaka matričnemu produktu $[\mathscr{V}_i \leftrightarrow \mathscr{V}_j]A$.

Matrike iz točk (i-iii) imenujemo *elementarne matrike*. Videli smo torej, da elementarne vrstične operacije na matriki A ustrezajo množenju matrike A z leve z ustreznimi elementarnimi matrikami. Z enostavnim izračunom lahko dokažemo naslednjo trditev:

TRDITEV 1.19. *Vse elementarne matrike so obrnljive. Inverz elementarne matrike je elementarna matrika, in sicer:*

$$[\mathscr{V}_i \leftarrow \alpha\mathscr{V}_i]^{-1} = [\mathscr{V}_i \leftarrow \tfrac{1}{\alpha}\mathscr{V}_i]$$

$$[\mathscr{V}_i \leftarrow \mathscr{V}_i + \beta\mathscr{V}_j]^{-1} = [\mathscr{V}_i \leftarrow \mathscr{V}_i - \beta\mathscr{V}_j]$$

$$[\mathscr{V}_i \leftrightarrow \mathscr{V}_j]^{-1} = [\mathscr{V}_i \leftrightarrow \mathscr{V}_j]$$

DEFINICIJA 1.20. Naj bosta $A, B \in \mathrm{Mat}(m \times n, \mathbb{F})$ poljubni matriki. Matrika B je *vrstično ekvivalentna* matriki A, če obstajajo takšne elementarne matrike $E^{(1)}, E^{(2)}, \dots, E^{(p)}$ velikosti $m \times m$, da je $B = E^{(p)} \cdots E^{(2)}E^{(1)}A$. Drugače povedano, matrika B je vrstično ekvivalentna matriki A tedaj, ko lahko matriko A s končnim zaporedjem elementarnih vrstičnih operacij preoblikujemo v matriko B.

KOMENTAR 1.21. (1) Posebej je vsaka matrika A vrstično ekvivalentna sama sebi, zato pravimo, da je vrstična ekvivalenca *refleksivna* relacija.

Vrstična ekvivalenca je tudi *simetrična* relacija, kar pomeni: če je matrika B vrstično ekvivalentna matriki A, potem je tudi matrika A vrstično ekvivalentna matriki B. Res, če je namreč $B = E^{(p)} \cdots E^{(2)}E^{(1)}A$ za neke elementarne matrike $E^{(1)}, E^{(2)}, \dots, E^{(p)}$, potem je $A = (E^{(1)})^{-1}(E^{(2)})^{-1} \cdots (E^{(p)})^{-1}B$, inverzi elementarnih matrik pa so spet elementarne matrike.

Vrstična ekvivalenca pa je tudi *tranzitivna* relacija, kar pomeni: če je matrika B vrstično ekvivalentna matriki A in matrika C vrstično ekvivalentna matriki B, potem je matrika C vrstično ekvivalentna matriki A. Res, če namreč velja $B = E^{(p)} \cdots E^{(2)}E^{(1)}A$ za neke elementarne matrike $E^{(1)}, E^{(2)}, \dots, E^{(p)}$, in če je $C = F^{(q)} \cdots F^{(2)}F^{(1)}B$ za neke elementarne matrike $F^{(1)}, F^{(2)}, \dots, F^{(q)}$, potem velja tudi $C = F^{(q)} \cdots F^{(2)}F^{(1)}E^{(p)} \cdots E^{(2)}E^{(1)}A$.

Ker je vrstična ekvivalenca torej refleksivna, simetrična in tranzitivna relacija, pravimo, da je *ekvivalenčna* relacija. Če sta si matriki A in B vrstično ekvivalentni, to običajno označimo z $A \sim B$.

(2) Ničelna matrika je vrstično ekvivalentna le sama sebi.

(3) Naj bosta A in B matriki velikosti $m \times n$, ki sta si vrstično ekvivalentni. Obstajajo torej elementarne matrike $E^{(1)}, E^{(2)}, \dots, E^{(p)}$, za katere je $B = E^{(p)} \cdots E^{(2)}E^{(1)}A$. Označimo $E = E^{(p)} \cdots E^{(2)}E^{(1)}$. Matrika E je torej produkt elementarnih matrik in je zato obrnljiva matrika, zanjo pa velja $B = EA$.

Iz zadnje enačbe sledi, da za vsak i velja

$$B_{i\bullet} = \sum_{j=1}^{m} E_{ij}A_{j\bullet}$$

Desna stran te enakosti je vsota s skalarji pomnoženih vrstic matrike A, takemu izrazu pa pravimo tudi *linearna kombinacija* vrstic matrike A. Če sta si torej matriki A in B vrstično ekvivalentni, potem je vsaka vrstica matrike B linearna kombinacija vrstic matrike A.

Komentar 1.22. Na podoben način lahko definiramo elementarne stolpčne operacije:

(1) Elementarna stolpčna operacija $(\mathscr{S}_i \leftarrow \alpha\mathscr{S}_i)$ pomnoži i-ti stolpec z neničelnim skalarjem $\alpha \in \mathbb{F}\setminus\{0\}$.
(2) Elementarna stolpčna operacija $(\mathscr{S}_i \leftarrow \mathscr{S}_i + \beta\mathscr{S}_j)$ prišteje β-kratnik j-tega stolpca k i-temu stolpcu, kjer je $\beta \in \mathbb{F}$ in $i \neq j$.
(3) Elementarna stolpčna operacija $(\mathscr{S}_i \leftrightarrow \mathscr{S}_j)$ zamenja i-ti in j-ti stolpec, kjer je $i \neq j$.

Stolpčne operacije na matriki lahko izvedemo tako, da matriko transponiramo, nato izvedemo ustrezno vrstično operacijo, rezultat pa nato še enkrat transponiramo. Rezultat operacije $(\mathscr{S}_i \leftarrow \alpha\mathscr{S}_i)$ na matriki A je torej produkt matrik

$$([\mathscr{V}_i \leftarrow \alpha\mathscr{V}_i]A^{\mathsf{t}})^{\mathsf{t}} = A[\mathscr{V}_i \leftarrow \alpha\mathscr{V}_i],$$

rezultat operacije $(\mathscr{S}_i \leftarrow \mathscr{S}_i + \beta\mathscr{S}_j)$ na matriki A je produkt matrik

$$([\mathscr{V}_i \leftarrow \mathscr{V}_i + \beta\mathscr{V}_j]A^{\mathsf{t}})^{\mathsf{t}} = A[\mathscr{V}_i \leftarrow \mathscr{V}_i + \beta\mathscr{V}_j]^{\mathsf{t}} = A[\mathscr{V}_j \leftarrow \mathscr{V}_j + \beta\mathscr{V}_i],$$

rezultat operacije $(\mathscr{S}_i \leftrightarrow \mathscr{S}_j)$ na matriki A pa je produkt matrik

$$([\mathscr{V}_i \leftrightarrow \mathscr{V}_j]A^{\mathsf{t}})^{\mathsf{t}} = A[\mathscr{V}_i \leftrightarrow \mathscr{V}_j].$$

Pri tem smo upoštevali, da velja $[\mathscr{V}_i \leftarrow \mathscr{V}_i + \beta\mathscr{V}_j]^{\mathsf{t}} = [\mathscr{V}_j \leftarrow \mathscr{V}_j + \beta\mathscr{V}_i]$ in da sta matriki $[\mathscr{V}_i \leftarrow \alpha\mathscr{V}_i]$ ter $[\mathscr{V}_i \leftrightarrow \mathscr{V}_j]$ simetrični.

Naj bosta $A, B \in \mathrm{Mat}(m \times n, \mathbb{F})$ poljubni matriki. Matrika B je *ekvivalentna* matriki A, če obstajajo takšne elementarne matrike $E^{(1)}, E^{(2)}, \ldots, E^{(p)}$ velikosti $m \times m$ in takšne elementarne matrike $F^{(1)}, F^{(2)}, \ldots, F^{(q)}$ velikosti $n \times n$, da je

$$B = E^{(p)} \cdots E^{(2)} E^{(1)} A F^{(1)} F^{(2)} \cdots F^{(q)}.$$

Drugače povedano, matrika B je ekvivalentna matriki A tedaj, ko lahko matriko A s končnim zaporedjem elementarnih vrstičnih operacij in končnim zaporedjem elementarnih stolpčnih operacij preoblikujemo v matriko B.

Podobno kot vrstična ekvivalenca je tudi ekvivalenca matrik ekvivalenčna relacija. To lahko brez težav preverimo na podoben način, kot smo to storili v komentarju 1.21(1) za vrstično ekvivalenco.

1.1.6. Vrstična kanonična forma matrike. Pokazali bomo, da z uporabo elementarnih vrstičnih operacij lahko matriko preoblikujemo v enostavnejšo obliko, ki je definirana na naslednji način:

Definicija 1.23. Matrika $A \in \mathrm{Mat}(m \times n, \mathbb{F})$ je v *vrstični kanonični formi*, če obstajajo celo število r, $0 \leq r \leq m$, in naravna števila $k_1, k_2, \ldots, k_r$, $k_1 < k_2 < \cdots < k_r \leq n$, tako da velja

(i) $A_{1k_1} = A_{2k_2} = \cdots = A_{rk_r} = 1$,
(ii) $A_{ij} = 0$ za vsa naravna števila i in j, za katera je $r < i \leq m$ in $j \leq n$, ter
(iii) $A_{ij} = 0$ za vsa naravna števila i in j, za katera je $i \leq r$ in $j < k_i$.

Če poleg tega tudi velja, da je

(iv) $A_{ik_l} = 0$ za vsa naravna števila i in l, za katera je $i < l \leq r$,

potem je A v *reducirani vrstični kanonični formi*.

Komentar 1.24. Če je matrika A v vrstični kanonični formi kot v definiciji, potem poziciji (l, k_l) v matriki A pravimo *pivot* matrike A, za vsa naravna števila $l \leq r$. Matrika A ima torej r pivotov, po eden od njih je v vsaki od prve do r-te vrstice. V prvih r vrsticah so komponente levo od pivotov enake številu 0, ostale vrstice pa so v celoti ničelne. Komponente matrike na pivotih so enake številu 1. Pivot, ki je v vrstici nad vrstico nekega drugega pivota, je v stolpcu, ki je levo od stolpca tega drugega pivota. Če je A v reducirani vrstični kanonični formi, potem so v stolpcih s pivoti tudi komponente nad pivoti enake številu 0.

Zgled 1.25. (1) Matrika

$$A = \begin{bmatrix} 0 & 1 & * & * & * & * & * & * & * \\ 0 & 0 & 0 & 1 & * & * & * & * & * \\ 0 & 0 & 0 & 0 & 1 & * & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

je v vrstični kanonični formi. Ima 5 pivotov, ki so v stolpcih z indeksi $2, 4, 5, 7, 9$. Tu smo s simbolom $*$ označili komponente, ki imajo lahko poljubne vrednosti.

(2) Matrike

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & 0 & 5 \\ 0 & 0 & 1 & 2 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & 2 & 5 \\ 0 & 1 & 3 & 6 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

so v reducirani vrstični kanonični formi.

(3) Matrike

$$\begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 5 \\ 0 & 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 7 & 3 & 4 \\ 0 & 1 & 2 & 5 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & 2 & 5 \\ 0 & 0 & 1 & 2 \end{bmatrix}$$

so v vrstični kanonični formi, a ne v reducirani vrstični kanonični formi.

(4) Matrike

$$\begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & 0 & 4 \\ 0 & 0 & 1 & 5 \\ 0 & 1 & 0 & 0 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & 2 & 3 \\ 1 & 0 & 1 & 0 \end{bmatrix}$$

niso v vrstični kanonični formi.

Trditev 1.26. *Naj bosta A in B matriki, ki sta si vrstično ekvivalentni.*

(i) Če sta matriki A in B obe v vrstični kanonični formi, potem je število pivotov matrike A enako številu pivotov matrike B. Še več, pivoti matrike A so v stolpcih z istimi indeksi kot pivoti matrike B.

(ii) Če sta matriki A in B obe v reducirani vrstični kanonični formi, potem sta si enaki.

Dokaz. Ker sta si matriki A in B vrstično ekvivalentni, sta obe iste velikosti $m \times n$. Če je ena od matrik A in B ničelna, je ničelna tudi druga in tedaj ni kaj dokazovati. Privzemimo torej, da sta matriki A in B obe neničelni.

Naj bosta matriki A in B obe v vrstični kanonični formi, naj bodo $k_1 < k_2 < \cdots < k_r$ indeksi stolpcev, v katerih so pivoti matrike A, in naj bodo $l_1 < l_2 <$

$\cdots < l_s$ indeksi stolpcev, v katerih so pivoti matrike B. Označimo še $k_0 = 0$ in $k_{r+1} = n + 1$.

(i) Najprej bomo dokazali, da je $\{l_1, \ldots, l_s\} \subset \{k_1 \ldots, k_r\}$. Vzemimo torej poljuben $p \in \{1, 2, \ldots, s\}$. Ker je $1 \leq l_p \leq n$, obstaja natanko en tak $q \in \{1, 2, \ldots, r+1\}$, da velja

$$0 = k_0 < k_1 < \cdots < k_{q-1} < l_p \leq k_q < \cdots < k_r < k_{r+1} = n + 1.$$

Ker sta si matriki A in B vrstično ekvivalentni, je p-ta vrstica matrike B linearna kombinacija vrstic matrike A (glej komentar 1.21(3)). Vendar pa so v matriki A vse vrstice pod r-to vrstico enake vrstici 0, zato mora biti p-ta vrstica matrike B linearna kombinacija prvih r vrstic matrike A. Drugače povedano, obstajajo takšni skalarji $\alpha_1, \alpha_2, \ldots, \alpha_r$, da je

$$B_{p\bullet} = \alpha_1 A_{1\bullet} + \alpha_2 A_{2\bullet} + \cdots + \alpha_r A_{r\bullet}. \tag{1.2}$$

Pokazali bomo, da je $\alpha_i = 0$ za vse $i < q$.

Oglejmo si enakost (1.2) v stolpcih z indeksi $k_1, k_2, \ldots, k_{q-1}$. Ker je pozicija (p, l_p) pivot v matriki B in ker je $k_1 < k_2 < \cdots < k_{q-1} < l_p$, velja $B_{pk_1} = B_{pk_2} = \cdots = B_{pk_{q-1}} = 0$. Ker pa je pozicija (i, k_i) pivot v matriki A za vse $i = 1, 2, \ldots, q-1$, je $A_{ik_i} = 1$ in $A_{(i+1)k_i} = \cdots = A_{rk_i} = 0$. Iz enakosti (1.2) zato sledi:

$$\begin{aligned}
0 &= \alpha_1 \\
0 &= \alpha_1 A_{1k_2} + \alpha_2 \\
&\;\;\vdots \\
0 &= \alpha_1 A_{1k_{q-1}} + \cdots + \alpha_{q-2} A_{(q-2)k_{q-1}} + \alpha_{q-1}
\end{aligned}$$

Odtod izračunamo, da je $\alpha_i = 0$ za vse $i < q$.

Iz enakosti (1.2) tako sledi

$$B_{p\bullet} = \alpha_q A_{q\bullet} + \alpha_{q+1} A_{(q+1)\bullet} + \cdots + \alpha_r A_{r\bullet} \tag{1.3}$$

Hkrati opazimo, da je $q \leq r$. V nasprotnem primeru bi namreč veljalo $B_{p\bullet} = 0$, kar pa ni res, saj je $B_{pl_p} = 1$. Enakost (1.3) si zdaj ogledamo v l_p-tem stolpcu,

$$B_{pl_p} = \alpha_q A_{ql_p} + \alpha_{q+1} A_{(q+1)l_p} + \cdots + \alpha_r A_{rl_p}.$$

Leva stran te enačbe je enaka številu 1, poleg tega pa je $A_{(q+1)l_p} = \cdots = A_{rl_p} = 0$, saj je pozicija (q, k_q) pivot matrike A in $l_p \leq k_q$. Odtod torej sledi

$$1 = \alpha_q A_{ql_p}.$$

Posebej to pomeni, da je $A_{ql_p} \neq 0$, a ker je $l_p \leq k_q$ in je pozicija (q, k_q) pivot matrike A, mora veljati $l_p = k_q$ in $\alpha_q = 1$.

S tem smo dokazali, da je $\{l_1, \ldots, l_s\} \subset \{k_1, \ldots, k_r\}$. Če v dokazu zamenjamo vlogi matrik A in B, na enak način pokažemo $\{k_1, \ldots, k_r\} \subset \{l_1, \ldots, l_s\}$. Sledi torej $\{l_1, \ldots, l_s\} = \{k_1, \ldots, k_r\}$ in posebej $s = r$.

(ii) Predpostavimo, da sta matriki A in B v reducirani vrstični kanonični formi. Po točki (i) že vemo, da je $s = r$ in $l_i = k_i$ za vse $i = 1, 2, \ldots, r$. Dokazati moramo, da je $B_{p\bullet} = A_{p\bullet}$ za vse $p \in \{1, 2, \ldots, m\}$. Če je $p > r$, je seveda $B_{p\bullet} = A_{p\bullet} = 0$. Predpostavimo torej, da je $p \leq r$.

Ob enakih oznakah kot v točki (i) v enakosti (1.3) zdaj velja $q = p$ in $\alpha_q = 1$, torej

$$B_{p\bullet} = A_{p\bullet} + \alpha_{p+1} A_{(p+1)\bullet} + \cdots + \alpha_r A_{r\bullet}. \tag{1.4}$$

Dokazati moramo, da je $\alpha_i = 0$ za vse $i > p$.

Oglejmo si enakost (1.4) v stolpcih z indeksi $k_{p+1}, k_{p+2}, \ldots, k_r$. Ker je matrika B v reducirani vrstični kanonični formi, je $B_{pk_{p+1}} = B_{pk_{p+2}} = \cdots = B_{pk_r} = 0$. Za vsak $i = p+1, p+2, \ldots, r$ je $A_{ik_i} = 1$, vse ostale komponente matrike A v k_i-tem stolpcu pa so enake številu 0, saj je tudi A v reducirani vrstični kanonični formi. Iz enakosti (1.4) v stolpcih z indeksi $k_{p+1}, k_{p+2}, \ldots, k_r$ tako dobimo:

$$\begin{aligned} 0 &= \alpha_{p+1} \\ 0 &= \alpha_{p+2} \\ &\vdots \\ 0 &= \alpha_r \end{aligned}$$

S tem je dokaz končan. □

1.1.7. Gaussova in Gauss-Jordanova eliminacija.

TRDITEV 1.27. *Vsaka matrika je vrstično ekvivalentna natanko eni matriki v reducirani vrstični kanonični formi.*

DOKAZ. Poljubno matriko $A \in \mathrm{Mat}(m \times n, \mathbb{F})$ z elementarnimi vrstičnimi operacijami lahko preoblikujemo v reducirano vrstično kanonično formo s spodaj opisanim algoritmom, ki se imenuje *Gauss-Jordanova eliminacija*.

Z algoritmom rekurzivno izračunamo matrike $A^{(0)}, A^{(1)}, \ldots, A^{(n)}$ in cela števila $r_0 \le r_1 \le \cdots \le r_n$ tako, da velja:

(i) $A^{(0)} = A$ in $r_0 = 0$,
(ii) matrika $A^{(j)}$ je dobljena iz matrike $A^{(j-1)}$ z elementarnimi vrstičnimi operacijami, za vse $j = 1, 2, \ldots, n$, in
(iii) matrika

$$\begin{bmatrix} A^{(j)}_{\bullet 1} & A^{(j)}_{\bullet 2} & \cdots & A^{(j)}_{\bullet j} \end{bmatrix},$$

sestavljena iz prvih j stolpcev matrike $A^{(j)}$, je v reducirani vrstični kanonični formi z r_j pivoti, za vse $j = 1, 2, \ldots, n$.

Kot začetni korak imamo $A^{(0)} = A$ in $r_0 = 0$. Predpostavimo, da je j naravno število, $j \le n$, in da smo v prvih $j-1$ korakih algoritma že izračunali matrike $A^{(0)}, A^{(1)}, \ldots, A^{(j-1)}$ in števila $r_0 \le r_1 \le \cdots \le r_{j-1}$ z želenimi lastnostmi. Matriko $A^{(j)}$ in število r_j tedaj izračunamo na naslednji način:

Če je $A^{(j-1)}_{ij} = 0$ za vsa naravna števila i, za katera je $r_{j-1} < i \le m$, potem vzamemo $A^{(j)} = A^{(j-1)}$ in $r_j = r_{j-1}$. V nasprotnem primeru pa lahko izberemo tako naravno število p_j, da je $r_{j-1} < p_j \le m$ in $A^{(j-1)}_{p_j j} \neq 0$. Definiramo $r_j = r_{j-1} + 1$ in označimo

$$\begin{aligned} P^{(j)} &= [\mathscr{V}_{r_j} \leftrightarrow \mathscr{V}_{p_j}] A^{(j-1)}, \\ Q^{(j)} &= [\mathscr{V}_{r_j} \leftarrow \tfrac{1}{P^{(j)}_{r_j j}} \mathscr{V}_{r_j}] P^{(j)}. \end{aligned}$$

S tem smo dosegli, da je (r_j, j)-ta komponenta matrike $Q^{(j)}$ enaka številu 1. Pozicija (r_j, j) je novi pivot.

Zdaj moramo še izničiti vse komponente nad in pod novim pivotom. To storimo tako, da matriko $Q^{(j)}$ z leve pomnožimo z elementarnimi matrikami

$$
\begin{gathered}
[\mathscr{V}_1 \leftarrow \mathscr{V}_1 - Q^{(j)}_{1j}\mathscr{V}_{r_j}],\\
\vdots\\
[\mathscr{V}_{r_j-1} \leftarrow \mathscr{V}_{r_j-1} - Q^{(j)}_{(r_j-1)j}\mathscr{V}_{r_j}],\\
[\mathscr{V}_{r_j+1} \leftarrow \mathscr{V}_{r_j+1} - Q^{(j)}_{(r_j+1)j}\mathscr{V}_{r_j}],\\
\vdots\\
[\mathscr{V}_m \leftarrow \mathscr{V}_m - Q^{(j)}_{mj}\mathscr{V}_{r_j}].
\end{gathered}
$$

Dobljeni rezultat je matrika $A^{(j)}$. Opazimo lahko, da z uporabljenimi vrstičnimi operacijami ne spremenimo stolpcev levo od j-tega stolpca.

Po n korakih dobimo matriko $A^{(n)}$, ki je v reducirani vrstični kanonični formi po točki (iii).

It trditve 1.26 sledi, da je vsaka matrika vrstično ekvivalentna največ eni matriki v vrstični kanonični formi. □

Komentar 1.28. Če v Gauss-Jordanovi eliminaciji izpustimo izničevanje komponent nad pivoti, za rezultat dobimo vrstično kanonično formo, ki ni nujno reducirana. Tako poenostavljenemu algoritmu pravimo *Gaussova eliminacija*.

Do (reducirane) vrstične forme bi lahko prišli tudi s kakšnim drugačnim zaporedjem elementarnih vrstičnih operacij. Če je matrika A' v (reducirani) vrstični kanonični formi in je vrstično ekvivalentna matriki A, potem pravimo, da je A' *(reducirana) vrstična kanonična forma matrike A*. Zadnja trditev torej pove, da ima vsaka matrika natanko eno reducirano vrstično kanonično formo.

Zgled 1.29. Uporabimo Gaussovo eliminacijo na konkretnem primeru matrike:

$$
\begin{array}{rl}
 & \begin{bmatrix} 1 & 2 & 5 & 0 & 1\\ 1 & 2 & 5 & 1 & 0\\ -3 & -6 & -16 & -3 & -2\\ -2 & -4 & -12 & -4 & -2 \end{bmatrix}\\
\begin{matrix}(\mathscr{V}_2 \leftarrow \mathscr{V}_2 - \mathscr{V}_1)\\ (\mathscr{V}_3 \leftarrow \mathscr{V}_3 + 3\mathscr{V}_1)\\ (\mathscr{V}_4 \leftarrow \mathscr{V}_4 + 2\mathscr{V}_1)\end{matrix} & \sim \begin{bmatrix} 1 & 2 & 5 & 0 & 1\\ 0 & 0 & 0 & 1 & -1\\ 0 & 0 & -1 & -3 & 1\\ 0 & 0 & -2 & -4 & 0 \end{bmatrix}\\
(\mathscr{V}_2 \leftrightarrow \mathscr{V}_3) & \sim \begin{bmatrix} 1 & 2 & 5 & 0 & 1\\ 0 & 0 & -1 & -3 & 1\\ 0 & 0 & 0 & 1 & -1\\ 0 & 0 & -2 & -4 & 0 \end{bmatrix}\\
(\mathscr{V}_2 \leftarrow -\mathscr{V}_2) & \sim \begin{bmatrix} 1 & 2 & 5 & 0 & 1\\ 0 & 0 & 1 & 3 & -1\\ 0 & 0 & 0 & 1 & -1\\ 0 & 0 & -2 & -4 & 0 \end{bmatrix}\\
(\mathscr{V}_4 \leftarrow \mathscr{V}_4 + 2\mathscr{V}_2) & \sim \begin{bmatrix} 1 & 2 & 5 & 0 & 1\\ 0 & 0 & 1 & 3 & -1\\ 0 & 0 & 0 & 1 & -1\\ 0 & 0 & 0 & 2 & -2 \end{bmatrix}
\end{array}
$$

$$(\mathscr{V}_4 \leftarrow \mathscr{V}_4 - 2\mathscr{V}_3) \qquad \sim \begin{bmatrix} 1 & 2 & 5 & 0 & 1 \\ 0 & 0 & 1 & 3 & -1 \\ 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Rezultat je v vrstični kanonični formi.

Definicija 1.30. Naj bo $A \in \mathrm{Mat}(m \times n, \mathbb{F})$ poljubna matrika. *Rang* matrike A je število neničelnih vrstic v reducirani vrstični kanonični formi matrike A. Rang matrike A označimo z

$$\mathrm{rank}(A).$$

Komentar 1.31. Rang matrike A je torej enak številu pivotov v (reducirani) vrstični kanonični formi matrike A. Ker je reducirana vrstična forma matrike A enolično določena, je rang matrike A dobro definiran. Iz definicije sledi, da imajo med seboj vrstično ekvivalentne matrike isti rang.

Rang matrike A je enak številu 0, če, in samo če, je matrika A ničelna. Rang neničelne matrike je naravno število. Če ima matrika A velikost $m \times n$, potem velja

$$0 \leq \mathrm{rank}(A) \leq \min\{n, m\}.$$

Matrika v vrstični kanonični formi ima namreč v vsakem stolpcu največ en pivot, zato rang matrike ne more biti večji od števila njenih stolpcev.

Zgled 1.32. (1) Z enim korakom Gaussove eliminacije dobimo:

$$A = \begin{bmatrix} 2 & 3 & 1 \\ 4 & 6 & 2 \end{bmatrix} \sim \begin{bmatrix} 2 & 3 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

Odtod vidimo, da je $\mathrm{rank}(A) = 1$. Postopka Gaussove eliminacije sicer še nismo v celoti končali, vendar tudi brez deljenja prve vrstice z 2 vidimo, da ima vrstična kanonična forma matrike en sam pivot, ki nastopa v prvem stolpcu.

(2) Z elementarnimi vrstičnimi operacijami izračunamo:

$$A = \begin{bmatrix} 1 & 1 & 0 & 2 & 3 \\ 1 & 1 & 1 & 0 & 2 \\ 0 & 2 & 1 & 2 & 1 \\ 1 & -1 & -1 & 0 & 2 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 0 & 2 & 3 \\ 0 & 0 & 1 & -2 & -1 \\ 0 & 2 & 1 & 2 & 1 \\ 0 & -2 & -1 & -2 & -1 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & 1 & 0 & 2 & 3 \\ 0 & 2 & 1 & 2 & 1 \\ 0 & 0 & 1 & -2 & -1 \\ 0 & -2 & -1 & -2 & -1 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 0 & 2 & 3 \\ 0 & 2 & 1 & 2 & 1 \\ 0 & 0 & 1 & -2 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Končna matrika sicer ni v vrstični kanonični formi, a je iz nje jasno razvidno, da ima vrstična kanonična forma matrike A tri pivote in sicer v stolpcih z indeksi $1, 2, 3$. Vidimo torej, da je $\mathrm{rank}(A) = 3$.

Vaja 1.33. Naj bo $A \in \mathrm{Mat}(m \times n, \mathbb{F})$ in $r = \mathrm{rank}(A)$. Z uporabo primernih stolpčnih operacij na reducirani vrstični kanonični formi matrike A pokaži, da je matrika A ekvivalentna matriki oblike:

$$\begin{bmatrix} \mathrm{I}_{r \times r} & 0_{r \times (n-r)} \\ 0_{(m-r) \times r} & 0_{(m-r) \times (n-r)} \end{bmatrix}$$

Pri tem zapisu privzamemo, da posameznega bloka v matriki ni, če je kateri od njegovih indeksov enak številu 0.

1.1.8. Računanje inverza z Gauss-Jordanovo eliminacijo.

TRDITEV 1.34. *Za poljubno kvadratno matriko* $A \in \mathrm{Mat}(n \times n, \mathbb{F})$ *so si ekvivalentne naslednje trditve:*

(i) *Matrika* A *obrnljiva.*
(ii) *Rang matrike* A *je enak številu* n.
(iii) *Matrika* A *se da zapisati kot produkt elementarnih matrik.*
(iv) *Reducirana vrstična kanonična forma matrike* A *je identična matrika* I.

DOKAZ. Naj bo A' reducirana vrstična kanonična forma matrike A. Posebej to pomeni, da je matrika A' vrstično ekvivalentna matriki A. Obstaja torej obrnljiva matrika E velikosti $n \times n$, ki je produkt elementarnih matrik in za katero velja $A' = EA$ (glej komentar 1.21(3)).

(iii⇒i) Ker so vse elementarne matrike obrnjive, je tudi vsak produkt elementarnih matrik obrnljiva matrika.

(i⇒iv) Če je matrika A obrnljiva, je obrnljiva tudi matrika A', saj je enaka produktu matrike A z obrnljivo matriko E. Posebej to pomeni, da je

$$A'_{n\bullet}((A')^{-1})_{\bullet n} = \mathrm{I}_{nn} = 1,$$

zato $A'_{n\bullet} \neq 0$. Matrika A' je torej kvadratna matrika v reducirani vrstični kanonični formi, ki ima zadnjo vrstico neničelno. Edina takšna matrika je identična matrika.

(iv⇒ii) Očitno je rang identične matrike velikosti $n \times n$ enak številu n.

(ii⇒iii) Ker je $\mathrm{rank}(A) = n$, je A' v reducirani vrstični kanonični formi z n pivoti, torej je $A' = \mathrm{I}$. Odtod dobimo $\mathrm{I} = EA$ in zato $A = E^{-1}$. Ker je matrika E produkt elementarnih matrik, je tudi njen inverz produkt elementarnih matrik. □

KOMENTAR 1.35. Dokazali smo torej, da je vsaka obrnljiva matrika produkt elementarnih matrik. Na podlagi tega lahko relaciji vrstične ekvivalence in ekvivalence matrik (glej komentarja 1.21 in 1.22) karakteriziramo na naslednji način:

(i) Matriki $A, B \in \mathrm{Mat}(m \times n, \mathbb{F})$ sta si vrstično ekvivalentni, če, in samo če, obstaja takšna matrika $E \in \mathrm{GL}(m, \mathbb{F})$, da je $B = EA$.

(ii) Matriki $A, B \in \mathrm{Mat}(m \times n, \mathbb{F})$ sta si ekvivalentni, če, in samo če, obstajata takšni matriki $E \in \mathrm{GL}(m, \mathbb{F})$ in $F \in \mathrm{GL}(n, \mathbb{F})$, da je $B = EAF$.

Trditev 1.34 posebej pove, da lahko z Gauss-Jordanovo eliminacijo ugotovimo, ali je dana kvadratna matrika obrnljiva ali ne. Pri tem postopku pa hkrati lahko izračunamo tudi inverz matrike, če ta obstaja.

Naj bo A poljubna kvadratna matrika velikosti $n \times n$ in naj bo A' reducirana vrstična kanonična forma matrike A. Vzemimo bločno zapisano matriko

$$\begin{bmatrix} A & \mathrm{I} \end{bmatrix}$$

velikosti $n \times 2n$ in na njej naredimo Gauss-Jordanovo eliminacijo. Produkt elementarnih matrik, ki smo jih uporabili pri tem postopku, je obrnljiva matrika E, matrika $E \begin{bmatrix} A & \mathrm{I} \end{bmatrix} = \begin{bmatrix} EA & E\mathrm{I} \end{bmatrix}$ pa je v reducirani vrstični kanonični formi. Iz definicije reducirane vrstične kanonične forme sledi, da je tudi matrika EA v reducirani vrstični kanonični formi, zato je $A' = EA$ oziroma

$$E \begin{bmatrix} A & \mathrm{I} \end{bmatrix} = \begin{bmatrix} EA & E\mathrm{I} \end{bmatrix} = \begin{bmatrix} A' & E \end{bmatrix}$$

Odtod vidimo:

(i) Če je $A' = \mathrm{I}$, je A obrnljiva in $A^{-1} = E$. Po Gauss-Jordanovi eliminaciji na matriki $\begin{bmatrix} A & \mathrm{I} \end{bmatrix}$ smo torej v zadnjih n stolpcih dobljene reducirane vrstične forme našli matriko A^{-1}.

(ii) Če $A' \neq \mathrm{I}$, potem matrika A ni obrnljiva.

ZGLED 1.36. Ali je matrika

$$A = \begin{bmatrix} 0 & 1 & 0 \\ -1 & -3 & 0 \\ 2 & 0 & 1 \end{bmatrix}$$

obrnljiva, in če je, kaj je njen inverz? Naredimo Gauss-Jordanovo eliminacijo na matriki A, razširjeni z identično matriko:

$$\left[\begin{array}{rrr|rrr} 0 & 1 & 0 & 1 & 0 & 0 \\ -1 & -3 & 0 & 0 & 1 & 0 \\ 2 & 0 & 1 & 0 & 0 & 1 \end{array}\right] \sim \left[\begin{array}{rrr|rrr} 1 & 3 & 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 2 & 0 & 1 & 0 & 0 & 1 \end{array}\right]$$
$$\sim \left[\begin{array}{rrr|rrr} 1 & 3 & 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & -6 & 1 & 0 & 2 & 1 \end{array}\right] \sim \left[\begin{array}{rrr|rrr} 1 & 3 & 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 6 & 2 & 1 \end{array}\right]$$
$$\sim \left[\begin{array}{rrr|rrr} 1 & 0 & 0 & -3 & -1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 6 & 2 & 1 \end{array}\right]$$

Matrika A je torej obrnljiva in velja:

$$A^{-1} = \begin{bmatrix} -3 & -1 & 0 \\ 1 & 0 & 0 \\ 6 & 2 & 1 \end{bmatrix}$$

VAJA 1.37. *Determinanta* matrike $A \in \mathrm{Mat}(2 \times 2)$ je število

$$\det(A) = A_{11}A_{22} - A_{12}A_{21} \in \mathbb{F}.$$

Pokaži, da je matrika A obrnljiva, če, in samo če, je $\det(A) \neq 0$, in da je tedaj

$$A^{-1} = \frac{1}{\det(A)} \begin{bmatrix} A_{22} & -A_{12} \\ -A_{21} & A_{11} \end{bmatrix}$$

inverz matrike A.

1.1.9. Reševanje sistemov linearnih enačb z Gaussovo eliminacijo. Matrike lahko med drugim uporabimo pri reševanju sistemov linearnih enačb. Sistem m linearnih enačb za n neznank $x_1, \ldots, x_n$ je sistem enačb:

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

Pri tem so $a_{11}, \ldots, a_{mn}, b_1, \ldots, b_m \in \mathbb{F}$ vnaprej dana števila. Rešitve takšnega sistema so urejene n-terice $(x_1, x_2, \ldots, x_n) \in \mathbb{F}^n$, ki zadoščajo vsem m enačbam hkrati.

KOMENTAR 1.38. (1) Dva sistema linearnih enačb za n neznank sta si *ekvivalentna*, če imata isto množico rešitev. Na primer, sistem dveh enačb za dve neznanki

$$\begin{aligned} 2x - 4y &= -2 \\ -x + 2y &= 1 \end{aligned}$$

je ekvivalenten sistemu dveh enačb za dve neznanki

$$\begin{aligned} -x + 2y &= 1 \\ 4x - 8y &= -4 \end{aligned}$$

in tudi sistemu ene enačbe za dve neznanki (bolje rečeno, eni sami enačbi za dve neznanki)

$$x - 2y = -1.$$

(2) Če za obseg skalarjev $\mathbb{F}$ izberemo kompleksna (ali realna, ali racionalna) števila, potem imenujemo sistem linearnih enačb (1.5) tudi sistem *kompleksnih* (ali *realnih*, ali *racionalnih*) linearnih enačb. Od izbora obsega skalarjev $\mathbb{F}$ je lahko odvisna množica rešitev sistema. Sistem realnih linearnih enačb je tudi sistem kompleksnih linearnih enačb in ima lahko kakšno kompleksno rešitev, ki ni realna.

ZGLED 1.39. Enačba ravnine v $\mathbb{R}^3$ je primer realne linearne enačbe za tri neznanke. Rešitve sistema dveh takšnih enačb so točke v preseku ustreznih ravnin. Ta presek je lahko prazen (če sta si ravnini vzporedni in se ne sekata), lahko je premica (če si ravnini nista vzporedni), lahko pa je tudi cela ravnina (če sta si ravnini enaki).

Sistem linearnih enačb lahko zapišemo kot enačbo med matrikami. Sistemu (1.5) pridružimo naslednje matrike:

(i) *matrika sistema* (1.5) je matrika

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \in \mathrm{Mat}(m \times n, \mathbb{F}),$$

(ii) *stolpec desnih strani sistema* (1.5) je stolpec

$$B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} \in \mathrm{Mat}(m \times 1, \mathbb{F}),$$

(iii) *stolpec neznank sistema* (1.5) pa je stolpec

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \in \mathrm{Mat}(n \times 1, \mathbb{F}).$$

(iv) *Razširjena matrika sistema* (1.5) je matrika

$$C = \begin{bmatrix} A & B \end{bmatrix} \in \mathrm{Mat}(m \times (n+1), \mathbb{F}).$$

Sistem linearnih enačb (1.5) s temi matrikami zapišemo kot eno samo matrično enačbo:

$$AX = B$$

Stolpec neznank $X \in \mathrm{Mat}(n \times 1, \mathbb{F}) = \mathbb{F}^n$ reši to matrično enačbo, če, in samo če, n-terica njegovih komponent $(x_1, x_2, \ldots, x_n) \in \mathbb{F}^n$ reši sistem linearnih enačb (1.5).

TRDITEV 1.40. *Naj bo $AX = B$ tak sistem n linearnih enačb za n neznank, da je matrika A obrnljiva. Tedaj ima sistem $AX = B$ eno samo rešitev in sicer $X = A^{-1}B$.*

DOKAZ. Ker je $A(A^{-1}B) = (AA^{-1})B = B$, je $A^{-1}B$ res rešitev sistema. Če je X poljubna rešitev sistema, sledi $X = (A^{-1}A)X = A^{-1}(AX) = A^{-1}B$. □

ZGLED 1.41. Dan je sistem n linearnih enačb za n neznank naslednje oblike:

$$\begin{aligned}
x_1 + a_{12}x_2 + \cdots + a_{1(n-1)}x_{n-1} + \quad a_{1n}x_n &= b_1 \\
x_2 + \cdots + a_{2(n-1)}x_{n-1} + \quad a_{2n}x_n &= b_2 \\
&\vdots \\
x_{n-1} + a_{(n-1)n}x_n &= b_{n-1} \\
x_n &= b_n
\end{aligned}$$

Drugače povedano, matrika sistema je obrnljiva in v vrstični kanonični formi. Po prejšni trditvi vemo, da ima sistem natanko eno rešitev. Ker pa je matrika sistema v vrstični kanonični formi, lahko rešitev zelo enostavno izračunamo, in sicer tako, da vrednosti neznank izračunamo po vrsti od zadnje proti prvi, in sicer iz enačb od zadnje proti prvi:

$$\begin{aligned}
x_n &= b_n \\
x_{n-1} &= b_{n-1} - a_{(n-1)n}x_n \\
&\vdots \\
x_2 &= b_2 - a_{23}x_3 - \cdots - a_{2n}x_n \\
x_1 &= b_1 - a_{12}x_2 - \cdots - a_{1(n-1)}x_{n-1} - a_{1n}x_n
\end{aligned}$$

Homogen sistem linearnih enačb je sistem linearnih enačb oblike $AX = 0$, torej tak, v katerem je stolpec desnih strani ničelen. Za poljuben sistem linearnih enačb $AX = B$ imamo *pridruženi homogen sistem* linearnih enačb $AX = 0$ z isto matriko sistema. Rešitve sistema linearnih enačb in pridruženega homogenega sistema so v tesni zvezi.

TRDITEV 1.42. *Naj bo X^{part} rešitev sistema linearnih enačb $AX = B$. Tedaj so rešitve sistema $AX = B$ natanko vsi stolpci oblike*

$$X^{\mathrm{part}} + X^{\mathrm{hom}},$$

kjer je X^{hom} poljubna rešitev pridruženega homogenega sistema.

DOKAZ. Ker je

$$A(X^{\mathrm{part}} + X^{\mathrm{hom}}) = AX^{\mathrm{part}} + AX^{\mathrm{hom}} = B + 0 = B,$$

je vsota $X^{\mathrm{part}} + X^{\mathrm{hom}}$ res rešitev sistema $AX = B$. Obratno, naj bo X poljubna rešitev sistema $AX = B$. Tedaj je

$$A(X - X^{\mathrm{part}}) = AX - AX^{\mathrm{part}} = B - B = 0,$$

torej je $X^{\text{hom}} = X - X^{\text{part}}$ rešitev pridruženega homogenega sistema. □

Komentar 1.43. Izbrani rešitvi X^{part} sistema $AX = B$ običajno pravimo *partikularna rešitev* sistema. Vsota $X^{\text{part}} + X^{\text{hom}}$, v kateri X^{hom} označuje poljubno rešitev pridruženega homogenega sistema, je *splošna* rešitev sistema $AX = B$, saj s takšnimi vsotami dobimo vse rešitve sistema.

Trditev 1.44. *Naj bo $AX = 0$ homogen sistem linearnih enačb.*

(i) Sistem $AX = 0$ ima vsaj eno rešitev, namreč trivialno rešitev $X = 0$.

(ii) Če sta X' in X'' dve rešitvi sistema $AX = 0$, potem je tudi $\alpha' X' + \alpha'' X''$ rešitev sistema $AX = 0$, za vse $\alpha', \alpha'' \in \mathbb{F}$.

(iii) Če ima homogen sistem kakšno neničelno rešitev, potem ima neskončno rešitev.

Dokaz. Velja $A(\alpha' X' + \alpha'' X'') = \alpha' AX' + \alpha'' AX'' = 0$. S tem smo dokazali točko (ii). Če ima homogen sistem kakšno neničelno rešitev X', potem je po točki (ii) tudi $\alpha' X'$ rešitev tega sistema, za vsak α', takšnih rešitev pa je neskončno. □

Kot posledico zadnjih dveh trditev lahko ugotovimo, da ima v splošnem sistem linearnih enačb lahko nič, eno ali neskončno rešitev. V nadaljevanju bomo opisali, kako z metodo Gaussove eliminacije ugotovimo, za katero od teh treh možnosti gre, in tudi izračunamo vse rešitve danega sistema linearnih enačb.

Naslednja trditev nam pove, kako lahko preoblikujemo sistem linearnih enačb brez da bi spremenili njegovo množico rešitev.

Trditev 1.45. *Naj bo $AX = B$ sistem m linearnih enačb za n neznank in naj bo E obrnljiva matrika velikosti $m \times m$. Tedaj sta si sistema linearnih enačb $AX = B$ in $(EA)X = EB$ ekvivalentna.*

Dokaz. Če je $AX = B$, potem velja $(EA)X = E(AX) = EB$. Obratno, če je $(EA)X = EB$, potem $AX = (E^{-1}E)(AX) = E^{-1}(EA)X = E^{-1}(EB) = B$. □

Naj bo zdaj $AX = B$ poljuben sistem m linearnih enačb za n neznank. Naj bo

$$C = \begin{bmatrix} A & B \end{bmatrix}$$

razširjena matrika tega sistema. Z Gaussovo (oziroma Gauss-Jordanovo) eliminacijo lahko matriko C preoblikujemo v njeno (reducirano) vrstično kanonično formo. Pri tem smo torej matriko C z leve pomnožili s takšno obrnljivo matriko E, da je matrika

$$EC = E \begin{bmatrix} A & B \end{bmatrix} = \begin{bmatrix} EA & EB \end{bmatrix}$$

v (reducirani) vrstični kanonični formi. Odtod sledi, da je tudi matrika EA v (reducirani) vrstični kanonični formi, matrika EC pa je razširjena matrika sistema linearnih enačb $(EA)X = EB$, ki ima iste rešitve kot prvotni sistem $AX = B$. Ker je matrika EA v (reducirani) vrstični kanonični formi, lahko te rešitve iz sistema $(EA)X = EB$ enostavno preberemo, kot bomo videli v dokazu naslednjega izreka.

Opazimo lahko, da elementarne vrstične operacije, katerih produkt je matrika E in ki jih naredimo na matriki C, ustrezajo operacijam na linearnih enačbah našega sistema: celotno enačbo lahko pomnožimo z neničelnim skalarjem, večkratnik ene enačbe lahko prištejemo k drugi enačbi, in lahko zamenjamo dve enačbi med seboj, pa se množica rešitev sistema ne bo spremenila.

Naslednji izrek nam pove, da je število rešitev sistema $AX = B$ določeno z rangom matrike A, rangom matrike C in številom neznank n. Spomnimo se, da velja $0 \leq \operatorname{rank}(A) \leq \min\{m, n\}$ in $0 \leq \operatorname{rank}(C) \leq \min\{m, n+1\}$. Ker je reducirana vrstična kanonična forma matrike A sestavljena iz prvih n stolpcev reducirane vrstične kanonične forme matrike C, je bodisi $\operatorname{rank}(C) = \operatorname{rank}(A)$ ali pa $\operatorname{rank}(C) = \operatorname{rank}(A) + 1$.

Izrek 1.46. *Naj bo $AX = B$ sistem m linearnih enačb za n neznank in naj bo $C = \begin{bmatrix} A & B \end{bmatrix}$ razširjena matrika tega sistema. Tedaj je bodisi* $\operatorname{rank}(C) = \operatorname{rank}(A)+1$ *ali pa* $\operatorname{rank}(C) = \operatorname{rank}(A) \leq n$.

(i) *Če je* $\operatorname{rank}(C) = \operatorname{rank}(A) + 1$, *potem sistem $AX = B$ nima rešitev.*
(ii) *Če je* $\operatorname{rank}(C) = \operatorname{rank}(A) = n$, *potem ima sistem $AX = B$ natanko eno rešitev.*
(iii) *Če je* $\operatorname{rank}(C) = \operatorname{rank}(A) < n$, *potem ima sistem neskončno rešitev, ki jih lahko parametriziramo z* $(n - \operatorname{rank}(A))$ *neodvisnih parametrov.*

Dokaz. Ker se z uporabo elementarnih vrstičnih operacij na matriki C množica rešitev sistema ne spremeni, lahko brez izgube splošnosti predpostavimo, da je matrika C v reducirani vrstični kanonični formi. Naj bo $r = \operatorname{rank}(C)$ in naj ima C pivote v stolpcih z indeksi $k_1, k_2, \ldots, k_r$,

$$1 \leq k_1 < k_2 < \cdots < k_r \leq n + 1.$$

Posebej je tudi matrika A v reducirani vrstični kanonični formi. Odtod direktno razberemo: če je $k_r = n + 1$, potem velja $r = \operatorname{rank}(C) = \operatorname{rank}(A) + 1$, sicer pa je $k_r \leq n$ in $\operatorname{rank}(A) = r = \operatorname{rank}(C)$.

(i) Če je $\operatorname{rank}(C) = \operatorname{rank}(A) + 1$, je $k_r = n + 1$. Tedaj je r-ta vrstica matrike C enaka vrstici $\begin{bmatrix} 0 & \cdots & 0 & 1 \end{bmatrix}$, kar pomeni, da se r-ta enačba v sistemu glasi

$$0X = 1.$$

Sistem torej nima rešitev.

(ii) Če je $r = \operatorname{rank}(C) = \operatorname{rank}(A) = n$, je matrika A v reducirani vrstični kanonični formi s pivoti v vseh svojih stolpcih. To pomeni, da je matrika C naslednje oblike:

$$C = \begin{bmatrix} \mathrm{I}_{n\times n} & B' \\ 0_{(m-n)\times n} & 0_{(m-n)\times 1} \end{bmatrix}$$

Zadnjih $m - n$ enačb v sistemu je torej odveč, saj so trivialne, torej enake enačbi $0X = 0$. Preostale enačbe sestavljajo sistem n enačb za n neznank,

$$\mathrm{I}X = B',$$

iz katerega direktno preberemo edino rešitev $X = B'$.

(iii) Predpostavimo, da je $r = \operatorname{rank}(C) = \operatorname{rank}(A) < n$. Kot običajno označimo $A = [a_{ij}]$, $X = (x_1, x_2, \ldots, x_n)$ in $B = (b_1, b_2, \ldots, b_m)$. Izberimo indekse $1 \leq p_1 < p_2 < \cdots < p_{n-r} \leq n$ tako, da je $\{p_1, p_2, \ldots, p_{n-r}\} \cup \{k_1, k_2, \ldots, k_r\} = \{1, 2, \ldots, n\}$. To pomeni, da so $p_1, p_2, \ldots, p_{n-r}$ indeksi natanko tistih stolpcev v matriki A, v katerih A nima pivotov.

Za $(n - r)$ prostih parametrov vzamemo neznanke

$$x_{p_1}, x_{p_2}, \ldots, x_{p_{n-r}}.$$

Tem neznankam lahko torej določimo poljubne vrednosti. Zdaj prepišemo sistem linearnih enačb v novo, ekvivalentno obliko tako, da izpustimo zadnjih $m-r$ enačb, ki so trivialne, izbrane proste parametre pa prenesemo na desno stran. Tako dobimo naslednji sistem enačb:

$$\begin{aligned} x_{k_1} &= b_1 - a_{1p_1}x_{p_1} - a_{1p_2}x_{p_2} - \cdots - a_{1p_{n-r}}x_{p_{n-r}} \\ x_{k_2} &= b_2 - a_{2p_1}x_{p_1} - a_{2p_2}x_{p_2} - \cdots - a_{2p_{n-r}}x_{p_{n-r}} \\ &\vdots \\ x_{k_r} &= b_r - a_{rp_1}x_{p_1} - a_{rp_2}x_{p_2} - \cdots - a_{rp_{n-r}}x_{p_{n-r}} \end{aligned}$$

Vrednosti neznank oziroma parametrov $x_{p_1}, x_{p_2}, \ldots, x_{p_{n-r}}$ torej izberemo poljubno, vrednosti preostalih neznanke $x_{k_1}, x_{k_2}, \ldots, x_{k_r}$ pa so s tem enolično določene in jih direktno preberemo iz dobljenega sistema linearnih enačb. □

Komentar 1.47. Podobno kot pri računanju ranga lahko tudi pri reševanju sistemov linearnih enačb uporabimo Gaussovo eliminacijo, s katero matriko C z elementarnimi vrstičnimi operacijami preoblikujemo v vrstično kanonično formo, ki ni nujno reducirana. V tem primeru moramo na koncu postopka neznanke izračunati postopoma od zadnje proti prvi, iz dobljenih preoblikovanih enačb od zadnje proti prvi, podobno kot smo to naredili v zgledu 1.41.

Zgled 1.48. (1) Rešimo sistem linearnih enačb:

$$\begin{aligned} x + y + z + t &= 2 \\ 2x + y + z - t &= 0 \\ y + z + 3t &= 0 \end{aligned}$$

Razširjeno matriko sistema preoblikujemo v vrstično kanonično formo z Gaussovo eliminacijo:

$$\left[\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 2 \\ 2 & 1 & 1 & -1 & 0 \\ 0 & 1 & 1 & 3 & 0 \end{array}\right] \sim \left[\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 2 \\ 0 & -1 & -1 & -3 & -4 \\ 0 & 1 & 1 & 3 & 0 \end{array}\right] \sim \left[\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 2 \\ 0 & 1 & 1 & 3 & 4 \\ 0 & 0 & 0 & 0 & -4 \end{array}\right]$$

Odtod vidimo, da je rang matrike sistema enak številu 2, rang razširjene matrike sistema pa je enak številu 3. Sistem torej nima rešitev.

(2) Za sistem

$$\begin{aligned} x + y - z - t &= 0 \\ x + y + z - t &= 2 \\ x - 2y + z + t &= 1 \\ x + y + z - 3t &= 0 \end{aligned}$$

izračunamo

$$\left[\begin{array}{cccc|c} 1 & 1 & -1 & -1 & 0 \\ 1 & 1 & 1 & -1 & 2 \\ 1 & -2 & 1 & 1 & 1 \\ 1 & 1 & 1 & -3 & 0 \end{array}\right] \sim \left[\begin{array}{cccc|c} 1 & 1 & -1 & -1 & 0 \\ 0 & 0 & 2 & 0 & 2 \\ 0 & -3 & 2 & 2 & 1 \\ 0 & 0 & 2 & -2 & 0 \end{array}\right]$$

$$\sim \left[\begin{array}{cccc|c} 1 & 1 & -1 & -1 & 0 \\ 0 & -3 & 2 & 2 & 1 \\ 0 & 0 & 2 & 0 & 2 \\ 0 & 0 & 0 & -2 & -2 \end{array}\right] \sim \left[\begin{array}{cccc|c} 1 & 1 & -1 & -1 & 0 \\ 0 & -3 & 2 & 2 & 1 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{array}\right]$$

Rang matrike sistema je enak številu 4, rang razširjene matrike sistema je prav tako enak številu 4, pa tudi neznanke so štiri. Sistem ima torej natanko eno rešitev.

Preoblikovan sistem je torej naslednji:

$$\begin{aligned} x + y - z - t &= 0 \\ -3y + 2z + 2t &= 1 \\ z &= 1 \\ t &= 1 \end{aligned}$$

Ta sistem je ekvivalenten prvotnemu sistemu. Zdaj izračunamo neznanke po vrsti od zadnje proti prvi, in sicer postopoma iz enačb od zadnje proti prvi:

$$\begin{aligned} t &= 1 \\ z &= 1 \\ y &= -\tfrac{1}{3}(1 - 2z - 2t) = 1 \\ x &= -y + z + t = 1 \end{aligned}$$

(3) Oglejmo si sistem realnih linearnih enačb:

$$\begin{aligned} x - y + z - t &= 1 \\ x - y + 2z + 3t &= 0 \\ x - y - z - 9t &= 3 \end{aligned}$$

Uporabimo Gaussovo eliminacijo na razširjeni matriki sistema:

$$\left[\begin{array}{rrrr|r} 1 & -1 & 1 & -1 & 1 \\ 1 & -1 & 2 & 3 & 0 \\ 1 & -1 & -1 & -9 & 3 \end{array}\right] \sim \left[\begin{array}{rrrr|r} 1 & -1 & 1 & -1 & 1 \\ 0 & 0 & 1 & 4 & -1 \\ 0 & 0 & -2 & -8 & 2 \end{array}\right]$$
$$\sim \left[\begin{array}{rrrr|r} 1 & -1 & 1 & -1 & 1 \\ 0 & 0 & 1 & 4 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

Rang matrike sistema je enak rangu razširjene matrike sistema, ki je 2. Ker imamo štiri neznanke, ima sistem neskončno rešitev, ki jih lahko parametriziramo z dvema neodvisnima parametroma. Ker ima vrstična kanonična forma razširjene matrike sistema pivote v prvem in tretjem stolpcu, ta dva pa pripadata neznankama x in z, preostali neznanki y in t izberemo za neodvisna parametra.

Prvotni sistem enačb je torej ekvivalenten sistemu dveh enačb:

$$\begin{aligned} x - y + z - t &= 1 \\ z + 4t &= -1 \end{aligned}$$

Vrednosti neznank y in t sta lahko poljubni, po izbiri njunih vrednosti pa lahko enolično izračunamo vrednosti neznank x in z, in sicer postopoma iz enačb od zadnje proti prvi:

$$\begin{aligned} z &= -1 - 4t \\ x &= 1 + y - z + t = 1 + y - (-1 - 4t) + t = 2 + y + 5t \end{aligned}$$

Množica rešitev sistema je torej

$$\{(2 + y + 5t, y, -1 - 4t, t)\ ;\ y, t \in \mathbb{R}\}.$$

1.2. Determinanta matrike

Spoznali smo že determinanto kvadratne matrike velikosti 2×2 in videli, da nam ta koristi pri karakterizaciji obrnljivost matrike ter pri računanju inverza takšne matrike. Ta rezultat bomo razširili na kvadratne matrike poljubne velikosti. V ta namen bomo najprej raziskali nekaj osnovnih lastnosti permutacij.

1.2.1. Simetrična grupa permutacij. Permutacija prvih n naravnih števil je bijekcija $\sigma : \{1, 2, \ldots, n\} \to \{1, 2, \ldots, n\}$. Množico vseh takšnih permutacij označimo z

$$\mathrm{Sym}(n).$$

Ta množica ima $n!$ elementov. Kompozicija dveh permutacij prvih n naravnih števil je spet permutacija prvih n naravnih števil, zato je kompozicija operacija na množici $\mathrm{Sym}(n)$. Spomnimo se, da je kompozicija asociativna operacija. *Identična permutacija* prvih n naravnih števil, ki je dana z identiteto $\mathrm{id} : \{1, 2, \ldots, n\} \to \{1, 2, \ldots, n\}$, je nevtralni element za kompozicijo. Ker je poljubna permutacij $\sigma \in \mathrm{Sym}(n)$ bijekcija, ima inverz, ki je spet permutacija $\sigma^{-1} \in \mathrm{Sym}(n)$. Iz naštetih lastnosti sledi, da je množica $\mathrm{Sym}(n)$ grupa za kompozicijo. Grupi $\mathrm{Sym}(n)$ pravimo *simetrična grupa* stopnje n. V literaturi se za grupo $\mathrm{Sym}(n)$ pogosto uporabljata tudi oznaki S_n in Σ_n.

Poljubno permutacijo $\sigma \in \mathrm{Sym}(n)$ lahko zapišemo v tako imenovani dvovrstični obliki:

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(n) \end{pmatrix}$$

V drugi vrstici nastopajo vsa naravna števila med 1 in n, seveda vsako natanko enkrat. Ta zapis nam ponudi drugačen, zelo nazoren pogled na permutacijo: permutacijo σ gledamo lahko kot razvrstitev naravnih števil med 1 in n v vrstni red $\sigma(1), \sigma(2), \ldots, \sigma(n)$.

Zgled 1.49. (1) Izračunamo lahko

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

in

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

Komentar 1.50. (1) Za poljubno permutacijo $\sigma \in \mathrm{Sym}(n)$ imamo preslikavi

$$\mathrm{L}_\sigma : \mathrm{Sym}(n) \to \mathrm{Sym}(n)$$

in

$$\mathrm{R}_\sigma : \mathrm{Sym}(n) \to \mathrm{Sym}(n),$$

ki se imenujeta *leva* oziroma *desna translacija* in sta definirani s predpisoma $\mathrm{L}_\sigma(\delta) = \sigma \circ \delta$ in $\mathrm{R}_\sigma(\delta) = \delta \circ \sigma$. Posebej je $\mathrm{L}_{\mathrm{id}} = \mathrm{R}_{\mathrm{id}} = \mathrm{id}$. Direkten izračun pokaže, da za poljubni permutaciji $\sigma, \sigma' \in \mathrm{Sym}(n)$ velja $\mathrm{L}_\sigma \circ \mathrm{L}_{\sigma'} = \mathrm{L}_{\sigma\circ\sigma'}$ in $\mathrm{R}_\sigma \circ \mathrm{R}_{\sigma'} = \mathrm{R}_{\sigma'\circ\sigma}$. Odtod sledi, da sta preslikavi L_σ in R_σ bijekciji z inverzoma $\mathrm{L}_\sigma^{-1} = \mathrm{L}_{\sigma^{-1}}$ in $\mathrm{R}_\sigma^{-1} = \mathrm{R}_{\sigma^{-1}}$. Tudi preslikava $\mathrm{Sym}(n) \to \mathrm{Sym}(n)$, $\delta \mapsto \delta^{-1}$, je bijekcija, saj je sama sebi inverz.

(2) Naj bo $\sigma \in \mathrm{Sym}(n)$. Definiramo $\sigma^0 = \mathrm{id}$, $\sigma^1 = \sigma$ in induktivno $\sigma^{m+1} = \sigma \circ \sigma^m$ za vsak $m \in \mathbb{N}$. Za vsako negativno celo število p definiramo $\sigma^p = (\sigma^{-1})^{|p|}$. Na ta način smo definirali

$$\sigma^k \in \mathrm{Sym}(n)$$

za vsak $k \in \mathbb{Z}$. Za vse $k, l \in \mathbb{Z}$ velja $\sigma^k \circ \sigma^l = \sigma^{k+l}$ in $(\sigma^k)^l = \sigma^{kl}$.

Izberimo k različnih naravnih števil $i_1, i_2, \ldots, i_k \in \{1, 2, \ldots, n\}$. Tem številom priredimo permutacijo

$$\begin{pmatrix} i_1 & i_2 & \cdots & i_k \end{pmatrix} \in \mathrm{Sym}(n),$$

ki i_k preslika v i_1, ki i_l preslika v i_{l+1}, za vse $l = 1, 2, \ldots, k-1$, in ki vsa ostala števila preslika sama vase. Permutaciji $\begin{pmatrix} i_1 & i_2 & \cdots & i_k \end{pmatrix}$ pravimo *cikel* dolžine k. Zapis cikla ni enoličen, saj velja $\begin{pmatrix} i_1 & i_2 & \cdots & i_k \end{pmatrix} = \begin{pmatrix} i_2 & \cdots & i_k & i_1 \end{pmatrix} = \cdots = \begin{pmatrix} i_k & i_1 & \cdots & i_{k-1} \end{pmatrix}$.

Cikel dolžine 1 je kar identična permutacija. Ciklom dolžine 2 pravimo tudi *transpozicije*. Transpozicije oblike $\begin{pmatrix} i & i+1 \end{pmatrix}$ imenujemo tudi *sosedske transpozicije*. Dva cikla $\begin{pmatrix} i_1 & i_2 & \cdots & i_k \end{pmatrix}$ in $\begin{pmatrix} j_1 & j_2 & \cdots & j_s \end{pmatrix}$ sta *disjunktna*, če je

$$\{i_1, i_2, \ldots, i_k\} \cap \{j_1, j_2, \ldots, j_s\} = \emptyset.$$

Z malo razmisleka lahko ugotovimo:

TRDITEV 1.51. *Vsako permutacijo lahko zapišemo kot kompozicijo paroma disjunktnih ciklov in tudi kot kompozicijo transpozicij. Vsako transpozicijo lahko zapišemo kot kompozicijo lihega števila sosedskih transpozicij.*

DOKAZ. Naj bo $\sigma \in \mathrm{Sym}(n)$ poljubna permutacija. Za poljubno naravno število $s \leq n$ si oglejmo zaporedje naravnih števil

$$(s, \sigma(s), \sigma^2(s), \sigma^3(s), \ldots).$$

Ker v tem zaporedju nastopajo le naravna števila med 1 in n, se morajo nekatera v njem ponavljati. Naj bo k prvo naravno število, za katero je $\sigma^k(s) = \sigma^l(s)$ za neko nenegativno celo število $l < k$. Drugače povedano, števila $s, \sigma(s), \ldots, \sigma^{k-1}(s)$ so med seboj različna, eno izmed njih pa je enako številu $\sigma^k(s)$. Tedaj je $l = 0$. Res, če bi veljalo $l > 0$, potem bi imeli $\sigma^{k-l}(s) = \sigma^0(s) = s$, kar pa bi bilo protislovje, saj so števila $s, \sigma(s), \ldots, \sigma^{k-1}(s)$ med seboj različna. Velja torej, da je $\sigma^k(s) = s$. S tem smo dobili cikel

$$\zeta_s = \begin{pmatrix} s & \sigma(s) & \sigma^2(s) & \cdots & \sigma^{k-1}(s) \end{pmatrix} \in \mathrm{Sym}(n),$$

ki se ujema s permutacijo σ na množici $C_s = \{s, \sigma(s), \sigma^2(s), \ldots, \sigma^{k-1}(s)\}$. Za vsak $s' \in C_s$ je $\zeta_s = \zeta_{s'}$, za poljuben $s'' \in \{1, 2, \ldots, n\} \setminus C_s$ pa velja $C_s \cap C_{s''} = \emptyset$. Odtod sledi, da lahko izberemo takšna števila $s_1, s_1, \ldots, s_p \in \{1, 2, \ldots, n\}$, da so množice $C_{s_1}, C_{s_2}, \ldots, C_{s_p}$ paroma disjunktne in da je njihova unija enaka množici $\{1, 2, \ldots, n\}$. Tedaj je $\sigma = \zeta_{s_1} \circ \zeta_{s_2} \circ \cdots \circ \zeta_{s_p}$.

Permutacijo σ lahko zapišemo kot kompozicijo transpozicij, ker lahko vsak cikel zapišemo kot kompozicijo transpozicij,

$$\begin{pmatrix} i_1 & i_2 & \cdots & i_k \end{pmatrix} = \begin{pmatrix} i_1 & i_2 \end{pmatrix} \circ \begin{pmatrix} i_2 & i_3 \end{pmatrix} \circ \cdots \circ \begin{pmatrix} i_{k-2} & i_{k-1} \end{pmatrix} \circ \begin{pmatrix} i_{k-1} & i_k \end{pmatrix}.$$

Poljubna transpozicija $\begin{pmatrix} i & j \end{pmatrix}$, $i < j$, se da zapisati kot kompozicija lihega števila sosedskih transpozicij: na primer, če vzamemo

$$\delta = \begin{pmatrix} i & i+1 \end{pmatrix} \circ \begin{pmatrix} i+1 & i+2 \end{pmatrix} \circ \cdots \circ \begin{pmatrix} j-2 & j-1 \end{pmatrix} \circ \begin{pmatrix} j-1 & j \end{pmatrix}$$

in

$$\delta' = \begin{pmatrix} j-1 & j-2 \end{pmatrix} \circ \begin{pmatrix} j-2 & j-3 \end{pmatrix} \circ \cdots \circ \begin{pmatrix} i+2 & i+1 \end{pmatrix} \circ \begin{pmatrix} i+1 & i \end{pmatrix},$$

potem velja $\begin{pmatrix} i & j \end{pmatrix} = \delta \circ \delta'$. □

KOMENTAR 1.52. Če sta cikla $\begin{pmatrix} i_1 & i_2 & \cdots & i_k \end{pmatrix}$ in $\begin{pmatrix} j_1 & j_2 & \cdots & j_s \end{pmatrix}$ disjunktna, potem komutirata, torej

$$\begin{pmatrix} i_1 & i_2 & \cdots & i_k \end{pmatrix} \circ \begin{pmatrix} j_1 & j_2 & \cdots & j_s \end{pmatrix} = \begin{pmatrix} j_1 & j_2 & \cdots & j_s \end{pmatrix} \circ \begin{pmatrix} i_1 & i_2 & \cdots & i_k \end{pmatrix}.$$

Po prejšnji trditvi lahko vsako permutacijo $\sigma \in \mathrm{Sym}(n)$ zapišemo kot kompozicijo paroma disjunktnih ciklov $\sigma = \zeta_{s_1} \circ \zeta_{s_2} \circ \cdots \circ \zeta_{s_p}$. Cikle $\zeta_{s_1}, \zeta_{s_2}, \ldots, \zeta_{s_p}$ lahko izberemo tako, da je vsota njihovih dolžin enaka številu n. Drugače povedano, vsako naravno število $i \leq n$ nastopi v zapisu natanko enega od ciklov $\zeta_{s_1}, \zeta_{s_2}, \ldots, \zeta_{s_p}$. V tem primeru krajše zapišemo

$$\sigma = \zeta_{s_1} \zeta_{s_2} \cdots \zeta_{s_p},$$

ta zapis pa je enoličen do vrstnega reda ciklov natančno.

ZGLED 1.53. Permutacijo zapišimo kot kompozicijo ciklov in kompozicijo transpozicij:

$$\begin{aligned} \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 4 & 6 & 5 & 3 & 1 & 2 & 7 \end{pmatrix} &= \begin{pmatrix} 1 & 4 & 3 & 5 \end{pmatrix} \begin{pmatrix} 2 & 6 \end{pmatrix} \begin{pmatrix} 7 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 4 \end{pmatrix} \circ \begin{pmatrix} 4 & 3 \end{pmatrix} \circ \begin{pmatrix} 3 & 5 \end{pmatrix} \circ \begin{pmatrix} 2 & 6 \end{pmatrix} \end{aligned}$$

Vsaka od dobljenih transpozicij je nadalje tudi kompozicija lihega števila sosedskih transpozicij:

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

$$\begin{pmatrix} 4 & 3 \end{pmatrix} = \begin{pmatrix} 4 & 3 \end{pmatrix}$$

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

$$\begin{pmatrix} 2 & 6 \end{pmatrix} = \begin{pmatrix} 2 & 3 \end{pmatrix} \circ \begin{pmatrix} 3 & 4 \end{pmatrix} \circ \begin{pmatrix} 4 & 5 \end{pmatrix} \circ \begin{pmatrix} 5 & 6 \end{pmatrix} \circ \begin{pmatrix} 5 & 4 \end{pmatrix} \circ \begin{pmatrix} 4 & 3 \end{pmatrix} \circ \begin{pmatrix} 3 & 2 \end{pmatrix}$$

Za poljubno permutacijo $\sigma \in \mathrm{Sym}(n)$ definirajmo

$$\mathrm{sgn}(\sigma) = \frac{1}{(n-1)!(n-2)!\cdots 2!1!} \prod_{1 \leq i < j \leq n} (\sigma(j) - \sigma(i)).$$

Število $\mathrm{sgn}(\sigma)$ je torej s konstanto $(n-1)!(n-2)!\cdots 2!1!$ deljen produkt razlik $(\sigma(j) - \sigma(i))$, za vse pare naravnih števil (i, j), za katere je $1 \leq i < j \leq n$. Očitno velja $\mathrm{sgn}(\sigma) \neq 0$. Definicija števila $\mathrm{sgn}(\sigma)$ na prvi pogled izgleda nekoliko zapletena, a hitro bomo videli, da je rezultat lahko le bodisi 1 ali pa -1, seveda v odvisnosti od tega, koliko faktorjev $(\sigma(j) - \sigma(i))$ v zgornjem produktu ima negativen predznak.

TRDITEV 1.54. *Naj bosta $\sigma, \sigma' \in \mathrm{Sym}(n)$ permutaciji in naj bo $\tau \in \mathrm{Sym}(n)$ transpozicija. Tedaj je*

(i) $\mathrm{sgn}(\sigma) \in \{1, -1\}$,
(ii) $\mathrm{sgn}(\mathrm{id}) = 1$,
(iii) $\mathrm{sgn}(\sigma \circ \sigma') = \mathrm{sgn}(\sigma)\,\mathrm{sgn}(\sigma')$ *in*
(iv) $\mathrm{sgn}(\tau) = -1$.

Dokaz. Direktno lahko izračunamo, da je $\operatorname{sgn}(\mathrm{id}) = 1$, torej velja točka (ii). Da bi dokazali še ostale točke se najprej prepričajmo, da velja

$$\operatorname{sgn}\big(\sigma \circ \big(r \quad r+1\big)\big) = -\operatorname{sgn}(\sigma),$$

za vsak $r = 1, 2, \ldots, n-1$. Res, produkta, s katerima sta definirani števili $\operatorname{sgn}(\sigma)$ in $\operatorname{sgn}\big(\sigma \circ \big(r \quad r+1\big)\big)$, imata večino faktorjev med seboj enakih, razlika je le v faktorjih pri $i = r, r+1$ in $j = r, r+1$. V definiciji števila $\operatorname{sgn}(\sigma)$ tako nastopajo faktorji

$$\prod_{i=1}^{r-1}(\sigma(r) - \sigma(i)) \prod_{i=1}^{r-1}(\sigma(r+1) - \sigma(i))$$

in

$$(\sigma(r+1) - \sigma(r)) \prod_{j=r+2}^{n}(\sigma(j) - \sigma(r)) \prod_{j=r+2}^{n}(\sigma(j) - \sigma(r+1)),$$

v definiciji števila $\operatorname{sgn}\big(\sigma \circ \big(r \quad r+1\big)\big)$ pa se zamenjata vlogi r in $r+1$, zato tam nastopajo faktorji

$$\prod_{i=1}^{r-1}(\sigma(r+1) - \sigma(i)) \prod_{i=1}^{r-1}(\sigma(r) - \sigma(i)),$$

in

$$(\sigma(r) - \sigma(r+1)) \prod_{j=r+2}^{n}(\sigma(j) - \sigma(r+1)) \prod_{j=r+2}^{n}(\sigma(j) - \sigma(r)).$$

Odtod vidimo, da je kvocient $\operatorname{sgn}\big(\sigma \circ \big(r \quad r+1\big)\big) / \operatorname{sgn}(\sigma)$ enak številu -1.

Zapišimo zdaj σ' kot kompozicijo sosedskih transpozicij, $\sigma' = \tau_1 \circ \tau_2 \circ \cdots \circ \tau_k$. Po pravkar dokazani formuli vidimo, da je

$$\begin{aligned}
\operatorname{sgn}(\sigma \circ \sigma') &= \operatorname{sgn}(\sigma \circ \tau_1 \circ \tau_2 \circ \cdots \circ \tau_k) \\
&= (-1)\operatorname{sgn}(\sigma \circ \tau_1 \circ \tau_2 \circ \cdots \circ \tau_{k-1}) \\
&= (-1)^2 \operatorname{sgn}(\sigma \circ \tau_1 \circ \tau_2 \circ \cdots \circ \tau_{k-2}) \\
&= \cdots \\
&= (-1)^{k-1} \operatorname{sgn}(\sigma \circ \tau_1) \\
&= (-1)^k \operatorname{sgn}(\sigma).
\end{aligned}$$

Če v to enakost vstavimo $\sigma = \mathrm{id}$, dobimo, da je

$$\operatorname{sgn}(\sigma') = (-1)^k.$$

Ker to velja za poljubno permutacijo σ', je s tem dokazana točka (i). Iz zgornjih dveh enakosti zdaj sledi

$$\operatorname{sgn}(\sigma \circ \sigma') = (-1)^k \operatorname{sgn}(\sigma) = \operatorname{sgn}(\sigma') \operatorname{sgn}(\sigma) = \operatorname{sgn}(\sigma) \operatorname{sgn}(\sigma'),$$

torej velja tudi točka (iii). Za dokaz točke (iv) se moramo le spomniti, da lahko vsako transpozicijo zapišemo kot kompozicijo lihega števila sosedskih transpozicij. □

Komentar 1.55. Številu $\operatorname{sgn}(\sigma)$ pravimo tudi *predznak* ali *signatura* permutacije σ. S tem je definirana preslikava

$$\operatorname{sgn} : \operatorname{Sym}(n) \to \{1, -1\}.$$

Množica $\{1, -1\}$ je (Abelova) grupa za množenje, in ker za preslikavo sgn veljata lastnosti (ii) in (iii) iz trditve 1.54, pravimo, da je sgn *homomorfizem grup*.

Predznak permutacije σ je torej lahko izračunati: preštejemo, za koliko parov (i, j), kjer je $1 \leq i < j \leq n$, je $\sigma(i) > \sigma(j)$. Takim parom pravimo *inverzije* permutacije σ, številu inverzij permutacije σ pa pravimo tudi *indeks* permutacije σ in ga označimo z $\operatorname{ind}(\sigma)$. Indeks permutacije je nenegativno celo število. Edina permutacija, ki ima indeks 0, je identiteta. Zveza med indeksom in predznakom permutacije je torej

$$\operatorname{sgn}(\sigma) = (-1)^{\operatorname{ind}(\sigma)}.$$

Če je $\operatorname{sgn}(\sigma) = 1$ oziroma če je $\operatorname{ind}(\sigma)$ sodo število, potem pravimo, da je permutacija σ *soda*. Če je $\operatorname{sgn}(\sigma) = -1$ oziroma če je $\operatorname{ind}(\sigma)$ liho število, potem pravimo, da je permutacija σ *liha*.

Iz prejšnje trditve direktno sledi:

POSLEDICA 1.56. *Permutacija $\sigma \in \operatorname{Sym}(n)$ je soda, če, in samo če, jo lahko zapišemo kot kompozicijo sodega števila transpozicij. Permutacija $\sigma \in \operatorname{Sym}(n)$ je liha, če, in samo če, jo lahko zapišemo kot kompozicijo lihega števila transpozicij.*

ZGLED 1.57. (1) Preštejemo inverzije v permutaciji

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 5 & 1 & 6 & 3 & 4 \end{pmatrix}$$

in ugotovimo, da je $\operatorname{ind}(\sigma) = 6$. Permutacija σ je torej soda.

(2) Če permutacijo $\sigma \in \operatorname{Sym}(n)$ zapišemo kot kompozicijo transpozicij

$$\sigma = \tau_1 \circ \tau_2 \circ \cdots \circ \tau_k,$$

potem je

$$\sigma^{-1} = \tau_k \circ \cdots \circ \tau_2 \circ \tau_1.$$

Posebej to pomeni, da je $\operatorname{sgn}(\sigma^{-1}) = \operatorname{sgn}(\sigma)$.

Za poljubno transpozicijo $\tau \in \operatorname{Sym}(n)$ je

$$\operatorname{sgn}(\mathrm{L}_\tau(\sigma)) = \operatorname{sgn}(\tau \circ \sigma) = \operatorname{sgn}(\tau) \operatorname{sgn}(\sigma) = -\operatorname{sgn}(\sigma)$$

in

$$\operatorname{sgn}(\mathrm{R}_\tau(\sigma)) = \operatorname{sgn}(\sigma \circ \tau) = \operatorname{sgn}(\sigma) \operatorname{sgn}(\tau) = -\operatorname{sgn}(\sigma).$$

To pomeni, da translaciji L_τ in R_τ obe bijektivno preslikata množico sodih permutacij v množico lihih permutacij in množico lihih permutacij v množico sodih permutacij. Posebej odtod sledi, da je število lihih permutacij iz $\operatorname{Sym}(n)$ enako številu sodih permutacij v $\operatorname{Sym}(n)$, če je $n \geq 2$. Grupa $\operatorname{Sym}(1)$ ima le en element, to je identiteta, in ta permutacija je soda.

1.2.2. Definicija determinante. Spoznali smo že determinanto matrike velikosti 2×2 in videli, da je determinanta takšne matrike različna od 0, če, in samo če, je matrika obrnljiva. To definicijo zdaj razširimo na poljubne kvadratne matrike. *Determinanta* kvadratne matrike $A \in \operatorname{Mat}(n \times n, \mathbb{F})$ je število

$$\det(A) = \sum_{\sigma \in \operatorname{Sym}(n)} \operatorname{sgn}(\sigma) A_{1\sigma(1)} A_{2\sigma(2)} \cdots A_{n\sigma(n)} \in \mathbb{F}.$$

V definiciji je torej $n!$ sumandov, vsak od sumandov $\operatorname{sgn}(\sigma) A_{1\sigma(1)} A_{2\sigma(2)} \cdots A_{n\sigma(n)}$ pa je s predznakom permutacije pomnožen produkt n komponent matrike, ki so izbrane tako, da je v vsaki vrstici in tudi v vsakem stolpcu natanko ena od teh komponent iz produkta. S tem smo za vsako naravno število n definirali preslikavo

$$\det : \mathrm{Mat}(n \times n, \mathbb{F}) \to \mathbb{F}.$$

V nadaljevanju bomo spoznali nekatere osnovne lastnosti determinante, ki nam bodo potrdile, da je zgoraj podana definicija determinante upravičena. Med drugim bomo videli, da so obrnljive natanko vse poljubno velike kvadratne matrike z neničelno determinanto. Videli bomo, da lahko determinanto izračunamo tudi z Gaussovo eliminacijo in da na ta način determinanto matriko večje velikosti izračunamo hitreje in enostavneje.

Komentar 1.58. Determinanto matrike $A \in \mathrm{Mat}(n \times n, \mathbb{F})$, kjer je $n \geq 2$, včasih zapišemo na naslednji način:

$$\det(A) = \det \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{bmatrix} = \begin{vmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{vmatrix}$$

Pri tem pa moramo paziti, da tega zapisa z navpičnimi črtami ne pomešamo z zapisom za absolutno vrednost števila, ki z determinanto nima nič skupnega. To je tudi razlog, da ta zapis za determinanto uporabljamo izključno v primeru $n \geq 2$.

Zgled 1.59. Oglejmo si determinanto matrik majhnih velikosti. Za $n = 1$ imamo

$$\det[A_{11}] = A_{11},$$

za $n = 2$ velja

$$\det \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \begin{vmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{vmatrix} = A_{11}A_{22} - A_{12}A_{21},$$

za $n = 3$ pa imamo

$$\begin{aligned} \det \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix} &= \begin{vmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{vmatrix} \\ &= A_{11}A_{22}A_{33} + A_{12}A_{23}A_{31} + A_{13}A_{21}A_{32} \\ &\quad -A_{11}A_{23}A_{32} - A_{13}A_{22}A_{31} - A_{12}A_{21}A_{33}. \end{aligned}$$

1.2.3. Osnovne lastnosti determinante.

Trditev 1.60. *Za vsak matriko $A \in \mathrm{Mat}(n \times n, \mathbb{F})$ velja*

$$\det(A) = \det(A^{\mathsf{t}}) = \sum_{\sigma \in \mathrm{Sym}(n)} \mathrm{sgn}(\sigma) A_{\sigma(1)1} A_{\sigma(2)2} \cdots A_{\sigma(n)n}.$$

Dokaz. Determinanto transponirane matrike izračunamo po definiciji in ob tem upoštevamo, da je preslikava $\mathrm{Sym}(n) \to \mathrm{Sym}(n)$, $\sigma \mapsto \sigma^{-1}$, bijekcija, ki ohranja predznak permutacij. S tem dobimo

$$\begin{aligned}
\det(A^{\mathsf{t}}) &= \sum_{\sigma\in \mathrm{Sym}(n)} \mathrm{sgn}(\sigma) A^{\mathsf{t}}_{1\sigma(1)} A^{\mathsf{t}}_{2\sigma(2)} \cdots A^{\mathsf{t}}_{n\sigma(n)} \\
&= \sum_{\sigma\in \mathrm{Sym}(n)} \mathrm{sgn}(\sigma) A_{\sigma(1)1} A_{\sigma(2)2} \cdots A_{\sigma(n)n} \\
&= \sum_{\sigma\in \mathrm{Sym}(n)} \mathrm{sgn}(\sigma^{-1}) A_{1\sigma^{-1}(1)} A_{2\sigma^{-1}(2)} \cdots A_{n\sigma^{-1}(n)} \\
&= \sum_{\delta\in \mathrm{Sym}(n)} \mathrm{sgn}(\delta) A_{1\delta(1)} A_{2\delta(2)} \cdots A_{n\delta(n)} \\
&= \det(A),
\end{aligned}$$

pri čemer smo pisali $\delta = \sigma^{-1}$. □

TRDITEV 1.61. *Če je matrika $A \in \mathrm{Mat}(n \times n, \mathbb{F})$ zgornje-trikotna ali spodnje-trikotna, potem je*

$$\det(A) = A_{11} A_{22} \cdots A_{nn}.$$

Posebej za diagonalne matrike velja

$$\det(\mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)) = \lambda_1 \lambda_2 \cdots \lambda_n$$

in za identično matriko

$$\det(\mathrm{I}) = 1.$$

DOKAZ. Predpostavimo najprej, da je A spodnje-trikotna. Naj bo $\sigma \in \mathrm{Sym}(n)$ poljubna permutacija. Če σ ni identiteta, lahko najdemo naravno število $p \leq n$, tako da je $\sigma(i) = i$ za vse $i < p$ in $\sigma(p) \neq p$. Tedaj je $\sigma(p) > p$, kajti σ je bijekcija. Ker je A spodnje-trikotna, odtod sledi $A_{p\sigma(p)} = 0$. Ugotovili smo torej, da je

$$A_{1\sigma(1)} A_{2\sigma(2)} \cdots A_{n\sigma(n)} = 0$$

za vsako permutacijo $\sigma \neq \mathrm{id}$, s tem pa smo trditev dokazali za spodnje-trikotne matrike.

Če je matrika A zgornje-trikotna, potem je matrika A^{t} spodnje-trikotna, in po že dokazanem dobimo

$$\det(A) = \det(A^{\mathsf{t}}) = A^{\mathsf{t}}_{11} A^{\mathsf{t}}_{22} \cdots A^{\mathsf{t}}_{nn} = A_{11} A_{22} \cdots A_{nn}.$$

Vsaka diagonalna matrika je zgornje-trikotna. □

V nadaljevanju si bomo ogledali, kako se determinanta matrike spremeni, če na vrsticah ali stolpcih matrike naredimo določene operacije. Prva lastnost sledi enostavno iz distributivnosti računanja s skalarji:

TRDITEV 1.62 (Multilinearnost determinante). *Naj bo $A \in \mathrm{Mat}(n \times n, \mathbb{F})$ poljubna kvadratna matrika.*

(i) Za vsako naravno število $i \leq n$, za vsako vrstico $B \in \mathrm{Mat}(1 \times n, \mathbb{F})$ in za vse skalarje $\alpha, \beta \in \mathbb{F}$ velja enakost:

$$\det \begin{bmatrix} A_{1\bullet} \\ \vdots \\ A_{(i-1)\bullet} \\ \alpha A_{i\bullet} + \beta B \\ A_{(i+1)\bullet} \\ \vdots \\ A_{n\bullet} \end{bmatrix} = \alpha \det \begin{bmatrix} A_{1\bullet} \\ \vdots \\ A_{(i-1)\bullet} \\ A_{i\bullet} \\ A_{(i+1)\bullet} \\ \vdots \\ A_{n\bullet} \end{bmatrix} + \beta \det \begin{bmatrix} A_{1\bullet} \\ \vdots \\ A_{(i-1)\bullet} \\ B \\ A_{(i+1)\bullet} \\ \vdots \\ A_{n\bullet} \end{bmatrix}$$

(ii) Za vsako naravno število $j \leq n$, za vsak stolpec $C \in \mathrm{Mat}(n \times 1, \mathbb{F})$ in za vse skalarje $\alpha, \beta \in \mathbb{F}$ velja

$$\begin{aligned} &\det \begin{bmatrix} A_{\bullet 1} & \cdots & A_{\bullet (j-1)} & \alpha A_{\bullet j} + \beta C & A_{\bullet (j+1)} & \cdots & A_{\bullet n} \end{bmatrix} \\ &= \alpha \det \begin{bmatrix} A_{\bullet 1} & \cdots & A_{\bullet (j-1)} & A_{\bullet j} & A_{\bullet (j+1)} & \cdots & A_{\bullet n} \end{bmatrix} \\ &\quad + \beta \det \begin{bmatrix} A_{\bullet 1} & \cdots & A_{\bullet (j-1)} & C & A_{\bullet (j+1)} & \cdots & A_{\bullet n} \end{bmatrix}. \end{aligned}$$

DOKAZ. Enakost (i) velja, ker je

$$\begin{aligned} &\sum_{\sigma \in \mathrm{Sym}(n)} \mathrm{sgn}(\sigma) A_{1\sigma(1)} \cdots A_{(i-1)\sigma(i-1)} (\alpha A_{i\sigma(i)} + \beta B_{1\sigma(i)}) A_{(i+1)\sigma(i+1)} \cdots A_{n\sigma(n)} \\ &= \alpha \sum_{\sigma \in \mathrm{Sym}(n)} \mathrm{sgn}(\sigma) A_{1\sigma(1)} \cdots A_{(i-1)\sigma(i-1)} A_{i\sigma(i)} A_{(i+1)\sigma(i+1)} \cdots A_{n\sigma(n)} \\ &\quad + \beta \sum_{\sigma \in \mathrm{Sym}(n)} \mathrm{sgn}(\sigma) A_{1\sigma(1)} \cdots A_{(i-1)\sigma(i-1)} B_{1\sigma(i)} A_{(i+1)\sigma(i+1)} \cdots A_{n\sigma(n)}. \qquad \square \end{aligned}$$

Dokaz točke (ii) je podoben. Na drug način lahko točko (ii) dokažemo tudi kot direktno posledico točke (ii) in trditve 1.60.

KOMENTAR 1.63. Zaradi lastnosti iz 1.62(i) pravimo, da je determinanta *multilinearna* funkcija vrstic matrike. Lastnosti determinante, ki se nanašajo na vrstice, pa veljajo tudi za stolpce: determinanta se namreč pri transponiranju ne spremeni, transponiranje pa vrstice spremeni v stolpce in obratno. Determinanta je zato tudi multilinearna funkcija stolpcev matrike, kar je vsebina trditve 1.62(ii.)

Trditev 1.62(i) v posebnem primeru $\beta = 0$ nam pove: Če eno od vrstico matrike pomnožimo s poljubnim skalarjem, se pri tem determinanta matrike pomnoži z istim skalarjem.

Podobno trditev 1.62(ii) v posebnem primeru $\beta = 0$ pravi: Če enega od stolpcev matrike pomnožimo s poljubnim skalarjem, se pri tem determinanta matrike pomnoži z istim skalarjem.

POSLEDICA 1.64. *Naj bo $A \in \mathrm{Mat}(n \times n, \mathbb{F})$ poljubna kvadratna matrika in naj bo $\alpha \in \mathbb{F}$ poljuben skalar. Tedaj je*

$$\det(\alpha A) = \alpha^n \det(A).$$

DOKAZ. Uporabimo trditev 1.62(i) pri $\beta = 0$ in sicer n-krat, za vsako vrstico posebej. $\square$

POSLEDICA 1.65. *Naj bo* $A \in \mathrm{Mat}(n \times n, \mathbb{F})$ *poljubna kvadratna matrika. Če je ena od vrstic matrike* A *ničelna, potem je* $\det(A) = 0$. *Če je eden od stolpcev matrike* A *ničelen, potem je* $\det(A) = 0$.

DOKAZ. Prvi del trditve je posledica trditve 1.62(i), kjer za i vzamemo indeks ničelne vrstice in $\alpha = \beta = 0$. Podobno vidimo, da je drugi del trditve posledica trditve 1.62(ii). □

TRDITEV 1.66. *Naj bo* $A \in \mathrm{Mat}(n \times n, \mathbb{F})$ *poljubna kvadratna matrika.*

(i) Za vsako naravno število $i \leq n$ *in vsak skalar* $\alpha \in \mathbb{F} \setminus \{0\}$ *velja*

$$\det([\mathscr{V}_i \leftarrow \alpha \mathscr{V}_i]A) = \det(A[\mathscr{V}_i \leftarrow \alpha \mathscr{V}_i]) = \alpha \det(A).$$

Posebej je $\det([\mathscr{V}_i \leftarrow \alpha \mathscr{V}_i]) = \alpha$.

(ii) Za poljubni različni naravni števili $i \leq n$ *in* $j \leq n$ *ter za vsak* $\beta \in \mathbb{F}$ *velja*

$$\det([\mathscr{V}_i \leftarrow \mathscr{V}_i + \beta \mathscr{V}_j]A) = \det(A[\mathscr{V}_j \leftarrow \mathscr{V}_j + \beta \mathscr{V}_i]) = \det(A).$$

Posebej je $\det([\mathscr{V}_i \leftarrow \mathscr{V}_i + \beta \mathscr{V}_j]) = 1$.

(iii) Za poljubni različni naravni števili $i \leq n$ *in* $j \leq n$ *velja*

$$\det([\mathscr{V}_i \leftrightarrow \mathscr{V}_j]A) = \det(A[\mathscr{V}_i \leftrightarrow \mathscr{V}_j]) = -\det(A).$$

Posebej je $\det([\mathscr{V}_i \leftrightarrow \mathscr{V}_j]) = -1$.

(iv) Če je ena od vrstic matrike A *večkratnik druge vrstice matrike* A*, potem je* $\det(A) = 0$. *Posebej, če ima matrika* A *dve enaki si vrstici, potem je* $\det(A) = 0$.

(v) Če je eden od stolpcev matrike A *večkratnik drugega stolpca matrike* A*, potem je* $\det(A) = 0$. *Posebej, če ima matrika* A *dva enaka si stolpca, potem je* $\det(A) = 0$.

DOKAZ. Točka (i) je direktna posledica trditve 1.62, v kateri vzamemo $\beta = 0$. Dokažimo zdaj točko (iii). Predpostavimo lahko, da je $i < j$. Označimo matriko $C = [\mathscr{V}_i \leftrightarrow \mathscr{V}_j]A$. Po definiciji determinante velja

$$\begin{aligned}
\det(C) &= \sum_{\sigma \in \mathrm{Sym}(n)} \mathrm{sgn}(\sigma) C_{1\sigma(1)} \cdots C_{i\sigma(i)} \cdots C_{j\sigma(j)} \cdots C_{n\sigma(n)} \\
&= \sum_{\sigma \in \mathrm{Sym}(n)} \mathrm{sgn}(\sigma) A_{1\sigma(1)} \cdots A_{j\sigma(i)} \cdots A_{i\sigma(j)} \cdots A_{n\sigma(n)} \\
&= \sum_{\sigma \in \mathrm{Sym}(n)} \mathrm{sgn}(\sigma) A_{1\sigma(1)} \cdots A_{i\sigma(j)} \cdots A_{j\sigma(i)} \cdots A_{n\sigma(n)}.
\end{aligned}$$

Naj bo $\tau = \begin{pmatrix} i & j \end{pmatrix}$ in označimo $\delta = \sigma \circ \tau$. Spomnimo se, da je $\mathrm{sgn}(\delta) = -\mathrm{sgn}(\sigma)$ in da je $\mathrm{R}_\tau : \mathrm{Sym}(n) \to \mathrm{Sym}(n)$ bijekcija. Iz zgornje enačbe sledi

$$\det(C) = \sum_{\delta \in \mathrm{Sym}(n)} (-\mathrm{sgn}(\delta)) A_{1\delta(1)} \cdots A_{i\delta(i)} \cdots A_{j\delta(j)} \cdots A_{n\delta(n)} = -\det(A).$$

Na koncu še izračunamo

$$\begin{aligned}
\det(A[\mathscr{V}_i \leftrightarrow \mathscr{V}_j]) &= \det((A[\mathscr{V}_i \leftrightarrow \mathscr{V}_j])^{\mathsf{t}}) \\
&= \det([\mathscr{V}_i \leftrightarrow \mathscr{V}_j]A^{\mathsf{t}}) \\
&= -\det(A^{\mathsf{t}}) = -\det(A).
\end{aligned}$$

Dokažimo zdaj točko (iv). Predpostavimo, da je i-ta vrstica matrike A γ-kratnik njene j-te vrstice, $i \neq j$. Če je $\gamma = 0$, potem je i-ta vrstica ničelna in je

$\det(A) = 0$ po trditvi 1.65. Naj bo torej $\gamma \neq 0$. Tedaj sta si i-ta in j-ta vrstica matrike $[\mathscr{V}_j \leftarrow \gamma \mathscr{V}_j]A$ enaki in zato velja

$$[\mathscr{V}_i \leftrightarrow \mathscr{V}_j][\mathscr{V}_j \leftarrow \gamma \mathscr{V}_j]A = [\mathscr{V}_j \leftarrow \gamma \mathscr{V}_j]A.$$

Odtod sledi

$$\begin{aligned} \gamma \det(A) &= \det([\mathscr{V}_j \leftarrow \gamma \mathscr{V}_j]A) \\ &= \det([\mathscr{V}_i \leftrightarrow \mathscr{V}_j][\mathscr{V}_j \leftarrow \gamma \mathscr{V}_j]A) \\ &= -\det([\mathscr{V}_j \leftarrow \gamma \mathscr{V}_j]A) \\ &= -\gamma \det(A). \end{aligned}$$

Sledi torej $\det(A) = 0$. Točko (v) dokažemo podobno.

(ii) Po trditvi 1.62(i) velja

$$\det([\mathscr{V}_i \leftarrow \mathscr{V}_i + \beta \mathscr{V}_j]A) = \det(A) + \beta \det(A'),$$

kjer je A' matrika, za katero je $A'_{i\bullet} = \beta A_{j\bullet}$ in $A'_{k\bullet} = A_{k\bullet}$ za vse $k \neq i$. Posebej to pomeni, da je i-ta vrstica matrike A' večkratnik njene j-te vrstice, zato po točki (iv) sledi, da je $\det(A') = 0$. Preostanek trditve iz točke (ii) dobimo s transponiranjem:

$$\begin{aligned} \det(A[\mathscr{V}_j \leftarrow \mathscr{V}_j + \beta \mathscr{V}_i]) &= \det((A[\mathscr{V}_j \leftarrow \mathscr{V}_j + \beta \mathscr{V}_i])^{\mathsf{t}}) \\ &= \det([\mathscr{V}_i \leftarrow \mathscr{V}_i + \beta \mathscr{V}_j]A^{\mathsf{t}}) \\ &= \det(A^{\mathsf{t}}) = \det(A). \end{aligned} \qquad \square$$

Točke (i-iii) trditve 1.66 natančno povedo, kako se determinanta spremeni, če matriko A spremenimo s kakšno od elementarnih vrstičnih ali elementarnih stolpčnih operacij. Velja torej:

(i) Če eno vrstico matrike ali en stolpec matrike pomnožimo s poljubnim skalarjem, se pri tem determinanta matrike pomnoži z istim skalarjem.

(ii) Če eni vrstici matrike prištejemo večkratnik neke druge vrstice te matrike, ali če enemu stolpcu matrike prištejemo večkratnik nekega drugega stolpca te matrike, se pri tem determinanta matrike ne spremeni.

(iii) Če v matriki zamenjamo med seboj dve vrstici ali če zamenjamo med seboj dva stolpca, se pri tem determinanta matrike pomnoži z -1.

Na kratko bi lahko zapisali, da velja $\det(EA) = \det(AE) = \det(E)\det(A)$ za vsako elementarno matriko E ustrezne velikosti.

Na podlagi tega vidimo, da determinanto lahko izračunamo z Gaussovo eliminacijo. Res, naj bo $A \in \mathrm{Mat}(n \times n, \mathbb{F})$ poljubna kvadratna matrika. Z metodo Gaussove eliminacije lahko najdemo takšne elementarne matrike $E^{(1)}, E^{(2)}, \ldots, E^{(p)}$ velikosti $n \times n$, da je matrika

$$A' = E^{(p)} \cdots E^{(2)} E^{(1)} A$$

v vrstični kanonični formi. Odtod sledi

$$\det(A') = \det(E^{(p)} \cdots E^{(2)} E^{(1)} A) = \det(E^{(p)}) \cdots \det(E^{(2)}) \det(E^{(1)}) \det(A).$$

Ker je kvadratna matrika A' v vrstični kanonični formi, je zgornje-trikotna, njena determinanta pa je bodisi 0 ali 1. Determinante elementarnih matrik so neničelne in jih poznamo. S tem smo torej izračunali

$$\det(A) = \frac{\det(A')}{\det(E^{(p)}) \cdots \det(E^{(2)}) \det(E^{(1)})}.$$

Poleg tega vidimo, da je $\det(A) \neq 0$, če, in samo če, je $\det(A') \neq 0$. Zadnje velja natanko tedaj, ko je $\operatorname{rank}(A) = n$, torej tedaj, ko je matrika A obrnljiva. S tem smo dokazali izrek:

IZREK 1.67. *Kvadratna matrika je obrnljiva, če, in samo če, je njena determinanta neničelna.*

Na podoben način dokažemo naslednjo pomembno lastnost determinante:

TRDITEV 1.68. *Za poljubni kvadratni matriki $A, B \in \operatorname{Mat}(n \times n, \mathbb{F})$ velja*

$$\det(AB) = \det(A)\det(B).$$

DOKAZ. Z metodo Gaussove eliminacije poiščemo takšne elementarne matrike $E^{(1)}, E^{(2)}, \ldots, E^{(p)}$ velikosti $n \times n$, da je matrika $A' = E^{(p)} \cdots E^{(2)} E^{(1)} A$ v reducirani vrstični kanonični formi. Odtod sledi $A = (E^{(1)})^{-1}(E^{(2)})^{-1} \cdots (E^{(p)})^{-1} A'$, zato

$$\begin{aligned}\det(A) &= \det((E^{(1)})^{-1}(E^{(2)})^{-1} \cdots (E^{(p)})^{-1} A') \\ &= \det((E^{(1)})^{-1}) \det((E^{(2)})^{-1}) \cdots \det((E^{(p)})^{-1}) \det(A')\end{aligned}$$

in splošneje

$$\begin{aligned}\det(AB) &= \det((E^{(1)})^{-1}(E^{(2)})^{-1} \cdots (E^{(p)})^{-1} A'B) \\ &= \det((E^{(1)})^{-1}) \det((E^{(2)})^{-1}) \cdots \det((E^{(p)})^{-1}) \det(A'B).\end{aligned}$$

Zdaj imamo dve možnosti:

(i) Če je A obrnljiva, je $A' = \mathrm{I}$ in $\det(A') = 1$, zato velja

$$\begin{aligned}\det(AB) &= \det((E^{(1)})^{-1}) \det((E^{(2)})^{-1}) \cdots \det((E^{(p)})^{-1}) \det(\mathrm{I}B) \\ &= \det(A)\det(B).\end{aligned}$$

(ii) Če A ni obrnljiva, je $\operatorname{rank}(A) < n$ in torej $A'_{n\bullet} = 0$. Odtod sledi, da je $(A'B)_{n\bullet} = 0$, zato velja $\det(A'B) = 0$ in posledično $\det(AB) = 0$. Ker je hkrati $\det(A) = 0$, tudi v tem primeru velja $\det(AB) = \det(A)\det(B)$. □

POSLEDICA 1.69. *Naj bosta $A, B \in \operatorname{Mat}(n \times n, \mathbb{F})$ kvadratni matriki.*
(i) Če je $AB = \mathrm{I}$, potem sta matriki A in B obrnljivi in velja $B = A^{-1}$.
(ii) Če je matrika A obrnljiva, potem velja

$$\det(A^{-1}) = \frac{1}{\det(A)}.$$

DOKAZ. Ker je $\det(A)\det(B) = \det(AB) = \det(\mathrm{I}) = 1$, imata matriki A in B neničelni determinanti in sta zato obrnljivi. Poleg tega vidimo, da je

$$B = A^{-1}AB = A^{-1}\mathrm{I} = A^{-1}.$$

S tem smo dokazali točko (i). Točka (ii) sledi iz enakosti

$$\det(A)\det(A^{-1}) = \det(AA^{-1}) = \det(\mathrm{I}) = 1.$$ □

ZGLED 1.70. (1) Z Gaussovo eliminacijo izračunamo:

$$\begin{vmatrix} 1 & 2 & 3 \\ -1 & 1 & 0 \\ 2 & 4 & 2 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 3 \\ 0 & 3 & 3 \\ 0 & 0 & -4 \end{vmatrix} = -12$$

$$\begin{vmatrix} 1 & 1 & 1 & 0 \\ 2 & 2 & 1 & 1 \\ 1 & 2 & 2 & 1 \\ 0 & 1 & 2 & 1 \end{vmatrix} = \begin{vmatrix} 1 & 1 & 1 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 1 & 2 & 1 \end{vmatrix} = - \begin{vmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 1 & 0 \end{vmatrix}$$
$$= - \begin{vmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 1 \end{vmatrix} = 1$$

(2) Če je matrika $A \in \mathrm{Mat}(n \times n, \mathbb{F})$ antisimetrična in če je število n liho, potem velja
$$\det(A) = \det(-A^{\mathsf{t}}) = (-1)^n \det(A) = -\det(A),$$
in zato $\det(A) = 0$.

(3) Za poljubno matriko $A \in \mathrm{Mat}(n \times n, \mathbb{F})$ direktno iz definicije sledi, da velja $\det(\bar{A}) = \overline{\det(A)}$. Če je matrika A hermitska, potem velja
$$\det(A) = \det(A^{\mathsf{h}}) = \det(\bar{A}^{\mathsf{t}}) = \overline{\det(A)},$$
in zato je $\det(A)$ realno število tudi v primeru, ko je $\mathbb{F} = \mathbb{C}$.

(4) Naj bo $A \in \mathrm{Mat}(n \times n, \mathbb{F})$ antihermitska matrika. Tedaj je
$$\det(A) = \det(-A^{\mathsf{h}}) = (-1)^n \det(\bar{A}^{\mathsf{t}}) = (-1)^n \overline{\det(A)}.$$
Če je število n sodo, potem je $\det(A)$ realno število, če pa je n liho število, iz enakosti sledi, da je realni del števila $\det(A)$ enak številu 0.

(5) Realna matrika Q je *ortogonalna*, če je kvadratna in zanjo velja $Q^{\mathsf{t}}Q = \mathrm{I}$. Vsaka ortogonalna matrika Q je obrnljiva in velja $Q^{-1} = Q^{\mathsf{t}}$, torej tudi $QQ^{\mathsf{t}} = \mathrm{I}$. Za poljubno ortogonalno matriko Q je
$$\det(Q)^2 = \det(Q^{\mathsf{t}}) \det(Q) = \det(Q^{\mathsf{t}}Q) = \det(\mathrm{I}) = 1,$$
torej $\det(Q) \in \{1, -1\}$. Množica vseh ortogonalnih matrik velikosti $n \times n$ je grupa za množenje, ki jo označimo z
$$\mathrm{O}(n)$$
in imenujemo *ortogonalna grupa* stopnje n.

(6) Kompleksna matrika U je *unitarna*, če je kvadratna in zanjo velja $U^{\mathsf{h}}U = \mathrm{I}$. Vsaka unitarna matrika U je obrnljiva in velja $U^{-1} = U^{\mathsf{h}}$, torej tudi $UU^{\mathsf{h}} = \mathrm{I}$. Za poljubno unitarno matriko U je
$$\overline{\det(U)} \det(U) = \det(U^{\mathsf{h}}) \det(U) = \det(U^{\mathsf{h}}U) = \det(\mathrm{I}) = 1,$$
torej $|\det(U)| = 1$. Množica vseh unitarnih matrik velikosti $n \times n$ je grupa za množenje, ki jo označimo z
$$\mathrm{U}(n)$$
in imenujemo *unitarna grupa* stopnje n. Opazimo lahko, da so ortogonalne matrike ravno realne unitarne matrike.

(7) Matrika $A \in \mathrm{Mat}(n \times n, \mathbb{F})$ je *nilpotentna*, če obstaja takšno naravno število k, da je $A^k = 0$. Tu smo z A^k označili produkt k kopij matrike A. Iz enakosti $A^k = 0$ sledi
$$\det(A)^k = \det(A^k) = \det(0) = 0,$$
torej je determinanta vsake nilpotentne matrike enaka številu 0.

VAJA 1.71. (1) Naj bo $A \in \mathrm{Mat}(n \times n, \mathbb{F})$ kvadratna matrika, ki je bločno oblike:

$$A = \begin{bmatrix} F(1,1) & F(1,2) & \cdots & F(1,\textsc{n}) \\ 0 & F(2,2) & \cdots & F(2,\textsc{n}) \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & F(\textsc{n},\textsc{n}) \end{bmatrix}$$

Pri tem naj bo $F(r,r) \in \mathrm{Mat}(n_r \times n_r, \mathbb{F})$ kvadratna matrika, za vse $r = 1, 2, \ldots, \textsc{n}$, in seveda $n = n_1 + n_2 + \cdots + n_{\textsc{n}}$. V tem primeru pravimo, da je matrika A *bločno zgornje-trikotna*. Pokaži, da velja

$$\det(A) = \det(F(1,1)) \det(F(2,2)) \cdots \det(F(\textsc{n},\textsc{n})).$$

Matrika A je torej obrnljiva, če, in samo če, so matrike $F(1,1), F(2,2), \ldots, F(\textsc{n},\textsc{n})$ vse obrnljive.

Če v takšni matriki A velja, da je tudi $F(r,s) = 0$ za $r < s$, potem je A *bločno diagonalna*. Pokaži, da za obrnljivo bločno diagonalno matriko A takšne oblike velja:

$$A^{-1} = \begin{bmatrix} F(1,1)^{-1} & 0 & \cdots & 0 \\ 0 & F(2,2)^{-1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & F(\textsc{n},\textsc{n})^{-1} \end{bmatrix}$$

(2) Pokaži, da je za poljubne skalarje $\alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbb{F}$ *Vandermondeova determinanta*

$$\begin{vmatrix} 1 & \alpha_1 & \alpha_1^2 & \cdots & \alpha_1^{n-1} \\ 1 & \alpha_2 & \alpha_2^2 & \cdots & \alpha_2^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \alpha_n & \alpha_n^2 & \cdots & \alpha_n^{n-1} \end{vmatrix}$$

enaka produktu $\prod_{1 \le i < j \le n} (\alpha_j - \alpha_i)$.

1.2.4. Razvoj determinante po vrstici ali stolpcu. Naj bo A kvadratna matrika velikosti $n \times n$ s komponentami v $\mathbb{F}$. V primeru, da je $n \geq 2$, smo za poljubni naravni števili $i \leq n$ in $j \leq n$ z

$$\mathrm{sub}^{ij}(A) \in \mathrm{Mat}((n-1) \times (n-1), \mathbb{F})$$

označili podmatriko matrike A, ki jo dobimo tako, da iz matrike A izbrišemo i-to vrstico in j-ti stolpec. Determinanto te podmatrike

$$\det(\mathrm{sub}^{ij}(A)) \in \mathbb{F}$$

imenujemo tudi (i,j)-ti *minor* matrike A. Če tega pomnožimo z $(-1)^{i+j}$, dobimo (i,j)-ti *kofaktor* matrike A, ki ga označimo z

$$\mathrm{co}_{ij}(A) = (-1)^{i+j} \det(\mathrm{sub}^{ij}(A)).$$

Na ta način smo definirali novo matriko

$$\mathrm{co}(A) \in \mathrm{Mat}(n \times n, \mathbb{F}),$$

matriko kofaktorjev matrike A, ki je podana z $(\mathrm{co}(A))_{ij} = \mathrm{co}_{ij}(A)$ za vse i in j. Če je $n = 1$ in je torej velikost matrike A enaka 1×1, posebej definiramo $\mathrm{co}_{11}(A) = 1$ in torej $\mathrm{co}(A) = \begin{bmatrix} 1 \end{bmatrix} \in \mathrm{Mat}(1 \times 1, \mathbb{F})$.

V nadaljevanju bomo spoznali, da lahko determinanto matrike na enostaven način izračunamo, če poznamo kofaktorje matrike A. Na ta način dobimo rekurzivno formulo za izračun determinante, s katero determinanto matrike velikosti $n \times n$ izračunamo s pomočjo kofaktorjev, ki so determinante matrik nižje velikosti. Ta formula je posebno uporabna, če ima matrika v neki vrstici ali stolpcu veliko ničel.

TRDITEV 1.72. *Naj bo $A \in \mathrm{Mat}(n \times n, \mathbb{F})$ kvadratna matrika.*
(i) Za vsako naravno število $i \leq n$ velja

$$\det(A) = \sum_{k=1}^{n} A_{ik} \operatorname{co}_{ik}(A).$$

(ii) Za vsako naravno število $j \leq n$ velja

$$\det(A) = \sum_{k=1}^{n} A_{kj} \operatorname{co}_{kj}(A).$$

KOMENTAR 1.73. Formuli iz točke (i) pravimo *razvoj determinante po i-ti vrstici*, formuli iz točke (ii) pa *razvoj determinante po j-tem stolpcu.*

DOKAZ. Najprej so oglejmo dva posebna primera:
(a) Predpostavimo najprej, da je n-ta vrstica matrike oblike

$$A_{n\bullet} = \begin{bmatrix} 0 & \cdots & 0 & A_{nn} \end{bmatrix}.$$

V definiciji determinante bodo tako neničelni le tisti sumandi, ki ustrezajo permutaciji $\sigma \in \mathrm{Sym}(n)$, za katero je $\sigma(n) = n$. Takšna permutacija je torej dana s permutacijo δ prvih $(n-1)$ naravnih števil, ki ima isti indeks in torej tudi isti predznak kot permutacija σ. Odtod sledi

$$\begin{aligned}
\det(A) &= \sum_{\sigma \in \mathrm{Sym}(n)} \operatorname{sgn}(\sigma) A_{1\sigma(1)} \cdots A_{(n-1)\sigma(n-1)} A_{n\sigma(n)} \\
&= \sum_{\substack{\sigma \in \mathrm{Sym}(n) \\ \sigma(n)=n}} \operatorname{sgn}(\sigma) A_{1\sigma(1)} \cdots A_{(n-1)\sigma(n-1)} A_{nn} \\
&= A_{nn} \sum_{\delta \in \mathrm{Sym}(n-1)} \operatorname{sgn}(\delta) A_{1\delta(1)} \cdots A_{(n-1)\delta(n-1)} \\
&= A_{nn} \det(\operatorname{sub}^{nn}(A)).
\end{aligned}$$

(b) Predpostavimo zdaj, da je i-ta vrstica matrike A oblike

$$A_{i\bullet} = \begin{bmatrix} 0 & \cdots & 0 & A_{ik} & 0 & \cdots & 0 \end{bmatrix}.$$

Drugače povedano, matrika A ima v i-ti vrstici vse komponente enake številu 0 razen morda tisto v k-tem stolpcu. Matriko A lahko z zamenjavami vrstic in stolpcev spremenimo v matriko B, ki je oblike iz točke (a). Res, vzemimo

$$B = [\mathscr{V}_n \leftrightarrow \mathscr{V}_{n-1}] \cdots [\mathscr{V}_{i+1} \leftrightarrow \mathscr{V}_i] A [\mathscr{V}_k \leftrightarrow \mathscr{V}_{k+1}] \cdots [\mathscr{V}_{n-1} \leftrightarrow \mathscr{V}_n].$$

Odtod dobimo $\det(A) = (-1)^{n-i}(-1)^{n-k} \det(B) = (-1)^{i+k} \det(B)$. Poleg tega vidimo, da je $B_{nn} = A_{ik}$ in da je $\operatorname{sub}^{nn}(B) = \operatorname{sub}^{ik}(A)$. Če torej točko (a) uporabimo na matriki B, dobimo

$$\begin{aligned}\det(A) &= (-1)^{i+k}\det(B) = (-1)^{i+k}B_{nn}\det(\mathrm{sub}^{nn}(B)) \\ &= (-1)^{i+k}A_{ik}\det(\mathrm{sub}^{ik}(A)) \\ &= A_{ik}\,\mathrm{co}_{ik}(A).\end{aligned}$$

(i) Naj bo zdaj A poljubna kvadratna matrika. Za vsak $k = 1, 2, \ldots, n$ naj bo

$$Y^{(k)} = \begin{bmatrix} 0 & \cdots & 0 & A_{ik} & 0 & \cdots & 0 \end{bmatrix} \in \mathrm{Mat}(1 \times n, \mathbb{F})$$

vrstica, ki ima vse komponente enake številu 0 razen morda tiste v k-tem stolpcu, ki je enaka številu A_{ik}. Označimo še z $A^{(k)} \in \mathrm{Mat}(n \times n, \mathbb{F})$ matriko, za katero je $A^{(k)}_{i\bullet} = Y^{(k)}$ in $A^{(k)}_{l\bullet} = A_{l\bullet}$ za vsak $l \neq i$. Vsaka od matrik $A^{(k)}$ je oblike iz točke (b). Ker je

$$A_{i\bullet} = \sum_{k=1}^{n} Y^{(k)},$$

z uporabo multilinearnosti determinante in enakosti iz točke (b) izračunamo

$$\det(A) = \sum_{k=1}^{n}\det(A^{(k)}) = \sum_{k=1}^{n} A_{ik}\,\mathrm{co}_{ik}(A).$$

(ii) S transponiranjem formulo iz točke (ii) prevedemo na formulo iz točke (i), ki smo jo že dokazali:

$$\begin{aligned}\det(A) = \det(A^{\mathsf{t}}) &= \sum_{k=1}^{n} A^{\mathsf{t}}_{jk}\,\mathrm{co}_{jk}(A^{\mathsf{t}}) \\ &= \sum_{k=1}^{n} A_{kj}\,\mathrm{co}_{kj}(A)\end{aligned}$$

Pri tem smo uporabili, da je

$$\begin{aligned}\mathrm{co}_{jk}(A^{\mathsf{t}}) &= (-1)^{j+k}\det(\mathrm{sub}^{jk}(A^{\mathsf{t}})) = (-1)^{j+k}\det((\mathrm{sub}^{kj}(A))^{\mathsf{t}}) \\ &= (-1)^{k+j}\det(\mathrm{sub}^{kj}(A)) = \mathrm{co}_{kj}(A).\end{aligned}$$ □

Zgled 1.74. Determinanto matrike velikosti 4×4 najprej razvijemo po tretji vrstici, dobljeni determinanti matrik velikosti 3×3 pa nato razvijemo prvo po prvi vrstici in drugo po prvem stolpcu:

$$\begin{aligned}\begin{vmatrix} 1 & 0 & -1 & 0 \\ 0 & 5 & 1 & 2 \\ 2 & 0 & 0 & 3 \\ 3 & 2 & 1 & 4 \end{vmatrix} &= 2\begin{vmatrix} 0 & -1 & 0 \\ 5 & 1 & 2 \\ 2 & 1 & 4 \end{vmatrix} - 3\begin{vmatrix} 1 & 0 & -1 \\ 0 & 5 & 1 \\ 3 & 2 & 1 \end{vmatrix} \\ &= 2\begin{vmatrix} 5 & 2 \\ 2 & 4 \end{vmatrix} - 3\left(\begin{vmatrix} 5 & 1 \\ 2 & 1 \end{vmatrix} + 3\begin{vmatrix} 0 & -1 \\ 5 & 1 \end{vmatrix}\right) \\ &= 2(20-4) - 3((5-2) + 3(0+5)) = -22\end{aligned}$$

1.2.5. Cramerjevo pravilo. Kot posledico formule za razvoj determinante po vrstici oziroma stolpcu dobimo:

TRDITEV 1.75. *Za poljubno obrnljivo matriko* $A \in \mathrm{GL}(n, \mathbb{F})$ *velja*

$$A^{-1} = \frac{1}{\det(A)} (\mathrm{co}(A))^{\mathrm{t}}.$$

DOKAZ. Za poljubna i in j je

$$(A(\mathrm{co}(A))^{\mathrm{t}})_{ij} = \sum_{k=1}^{n} A_{ik} (\mathrm{co}(A)^{\mathrm{t}})_{kj} = \sum_{k=1}^{n} A_{ik} \,\mathrm{co}_{jk}(A)$$

Če je $i = j$, po formuli za razvoj determinante matrike A po i-ti vrstici velja

$$(A(\mathrm{co}(A))^{\mathrm{t}})_{ii} = \sum_{k=1}^{n} A_{ik} \,\mathrm{co}_{ik}(A) = \det(A).$$

Predpostavimo zdaj, da je $i \neq j$. Naj bo B matrika velikosti $n \times n$, dana s predpisom $B_{j\bullet} = A_{i\bullet}$ in $B_{l\bullet} = A_{l\bullet}$ za vsak $l \neq j$. Matrika B ima j-to vrstico enako i-ti, zato je $\det(B) = 0$. Poleg tega velja $\mathrm{co}_{jk}(B) = \mathrm{co}_{jk}(A)$ za vsak k. Po formuli za razvoj determinante matrike B po j-ti vrstici zdaj dobimo

$$(A(\mathrm{co}(A))^{\mathrm{t}})_{ij} = \sum_{k=1}^{n} A_{ik} \,\mathrm{co}_{jk}(A) = \sum_{k=1}^{n} B_{jk} \,\mathrm{co}_{jk}(B) = \det(B) = 0.$$

S tem smo dokazali, da je $A(\mathrm{co}(A))^{\mathrm{t}} = \det(A)\mathrm{I}$. □

ZGLED 1.76. Naj bo

$$A = \begin{bmatrix} 1 & 0 & 2 \\ 1 & -1 & 1 \\ 0 & 2 & 3 \end{bmatrix}$$

dana matrika. Če izračunamo kofaktorje matrike A, dobimo:

$$\mathrm{co}(A) = \begin{bmatrix} -5 & -3 & 2 \\ 4 & 3 & -2 \\ 2 & 1 & -1 \end{bmatrix}$$

Iz kofaktorjev lahko izračunamo tudi determinanto matrike A, recimo z razvojem po prvi vrstici:

$$\det(A) = \mathrm{co}_{11}(A) + 2\,\mathrm{co}_{13}(A) = -5 + 4 = -1.$$

Matrika A je torej obrnljiva, njen inverz pa je matrika:

$$A^{-1} = -(\mathrm{co}(A))^{\mathrm{t}} = \begin{bmatrix} 5 & -4 & -2 \\ 3 & -3 & -1 \\ -2 & 2 & 1 \end{bmatrix}$$

Iz prejšnje trditve dobimo *Cramerjevo pravilo* za reševanje sistemov linearnih enačb:

TRDITEV 1.77 (Cramerjevo pravilo). *Naj bo* $AX = B$ *sistem* n *linearnih enačb za* n *neznank* $x_1, x_2, \ldots, x_n$*, za katerega velja* $\det(A) \neq 0$*. Tedaj ima sistem linearnih enačb* $AX = B$ *eno samo rešitev in sicer*

$$x_j = \frac{\det \begin{bmatrix} A_{\bullet 1} & \cdots & A_{\bullet (j-1)} & B & A_{\bullet (j+1)} & \cdots & A_{\bullet n} \end{bmatrix}}{\det(A)}$$

za vse $j = 1, 2, \ldots, n$.

Dokaz. Za vsako naravno število $j \leq n$ označimo

$$A^{(j)} = \begin{bmatrix} A_{\bullet 1} & \cdots & A_{\bullet (j-1)} & B & A_{\bullet (j+1)} & \cdots & A_{\bullet n} \end{bmatrix}.$$

Opazimo lahko, da velja $\mathrm{co}_{ij}(A) = \mathrm{co}_{ij}(A^{(j)})$ za vsak i. Ker je determinanta matrike sistema A neničelna, je matrika A obrnljiva in sistem ima eno samo rešitev $X = A^{-1}B$. Iz trditve 1.75 lahko za vsak j eksplicitno izračunamo

$$\begin{aligned}
\det(A)\, x_j &= \det(A)(A^{-1}B)_{j1} = ((\mathrm{co}(A))^{\mathrm{t}} B)_{j1} \\
&= \sum_{i=1}^{n} ((\mathrm{co}(A))^{\mathrm{t}})_{ji} B_{i1} = \sum_{i=1}^{n} \mathrm{co}_{ij}(A) B_{i1} \\
&= \sum_{i=1}^{n} A_{ij}^{(j)}\, \mathrm{co}_{ij}(A^{(j)}) \\
&= \det(A^{(j)}),
\end{aligned}$$

kjer smo v zadnjem koraku uporabili formulo za razvoj determinante matrike $A^{(j)}$ po j-tem stolpcu. □

Zgled 1.78. Oglejmo si sistem treh linearnih enačb za tri neznanke:

$$\begin{aligned}
x + y + z &= 1 \\
x - y - 2z &= 2 \\
2x - y + z &= 0
\end{aligned}$$

Determinanto matrike sistema

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & -2 \\ 2 & -1 & 1 \end{bmatrix}$$

izračunamo na primer z razvojem po prvi vrstici,

$$\det(A) = \begin{vmatrix} -1 & -2 \\ -1 & 1 \end{vmatrix} - \begin{vmatrix} 1 & -2 \\ 2 & 1 \end{vmatrix} + \begin{vmatrix} 1 & -1 \\ 2 & -1 \end{vmatrix} = -3 - 5 + 1 = -7.$$

Determinanta je neničelna, zato je matrika A obrnljiva. Sistem linearnih enačb ima torej eno samo rešitev, ki jo izračunamo po Cramerjevem pravilu:

$$x = -\frac{1}{7} \begin{vmatrix} 1 & 1 & 1 \\ 2 & -1 & -2 \\ 0 & -1 & 1 \end{vmatrix} = -\frac{1}{7} \left(\begin{vmatrix} -1 & -2 \\ -1 & 1 \end{vmatrix} - 2 \begin{vmatrix} 1 & 1 \\ -1 & 1 \end{vmatrix} \right) = 1$$

$$y = -\frac{1}{7} \begin{vmatrix} 1 & 1 & 1 \\ 1 & 2 & -2 \\ 2 & 0 & 1 \end{vmatrix} = -\frac{1}{7} \left(2 \begin{vmatrix} 1 & 1 \\ 2 & -2 \end{vmatrix} + \begin{vmatrix} 1 & 1 \\ 1 & 2 \end{vmatrix} \right) = 1$$

$$z = -\frac{1}{7} \begin{vmatrix} 1 & 1 & 1 \\ 1 & -1 & 2 \\ 2 & -1 & 0 \end{vmatrix} = -\frac{1}{7} \left(2 \begin{vmatrix} 1 & 1 \\ -1 & 2 \end{vmatrix} + \begin{vmatrix} 1 & 1 \\ 1 & 2 \end{vmatrix} \right) = -1$$

Pri tem smo prvo determinanto razvili po prvem stolpcu, drugo in tretjo pa po tretji vrstici.

1.2.6. Poddeterminante matrike. Naj bo $A \in \mathrm{Mat}(m \times n, \mathbb{F})$ poljubna matrika. Matrika A torej ni nujno kvadratna, in če ni kvadratna, nima determinante. Vendar pa ima matrika A prav gotovo podmatrike, ki so kvadratne, in te imajo svoje determinante. Determinantam kvadratnih podmatrik matrike A pravimo *poddeterminante* matrike A.

Bolj natančno, za vsako naravno število k so $(k \times k)$-*poddeterminante* matrike A determinante kvadratnih podmatrik matrike A velikosti $k \times k$. Matrika A ima torej lahko več različnih $(k \times k)$-poddeterminant, za vsak $k = 1, 2, \ldots, \min\{m, n\}$, seveda pa nima $(p \times p)$-poddeterminant za $p > \min\{m, n\}$.

Zgled 1.79. Kvadratne podmatrike matrike

$$A = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 3 & 1 \end{bmatrix}$$

imajo lahko velikost bodisi 1×1 ali pa 2×2. Matrika A ima štiri različne podmatrike velikosti 1×1:

$$\begin{bmatrix} 1 \end{bmatrix} \quad \begin{bmatrix} 2 \end{bmatrix} \quad \begin{bmatrix} 0 \end{bmatrix} \quad \begin{bmatrix} 3 \end{bmatrix}$$

Podmatrike matrike A velikosti 2×2 so tri, in sicer:

$$\begin{bmatrix} 2 & 0 \\ 3 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}$$

Matrika A ima zato (1×1)-poddeterminante $0, 1, 2, 3$ in (2×2)-poddeterminante $1, 2, 3$.

Če poznamo poddeterminante matrike, lahko iz njih razberemo rang matrike:

Trditev 1.80. *Naj bo $A \in \mathrm{Mat}(m \times n, \mathbb{F})$ poljubna neničelna matrika in naj bo r njen rang. Tedaj velja:*

(i) Obstaja $(r \times r)$-poddeterminanta matrike A, ki je različna od 0.

(ii) Za vsako naravno število k, ki je večje od r, so vse $(k \times k)$-poddeterminante matrike A enake številu 0.

Dokaz. Najprej lahko opazimo, da trditvi (i) in (ii) veljata v primeru, da je A v vrstični kanonični formi. Tedaj je namreč

$$\det(\mathrm{sub}_{\{1,\ldots,r\}\{k_1,\ldots,k_r\}}(A)) = 1,$$

kjer so $k_1, \ldots, k_r$ indeksi stolpcev s pivoti v A, za $k > r$ pa ima vsaka podmatrika velikosti $k \times k$ ničelno zadnjo vrstico in zato ničelno determinanto.

Če je A poljubna matrika, jo lahko z elementarnimi vrstičnimi operacijami preoblikujemo v vrstično kanonično formo. Dovolj je torej dokazati, da za vsako naravno število k velja: če ima matrika A vse $(k \times k)$-poddeterminante enake številu 0, potem ima tudi matrika EA vse $(k \times k)$-poddeterminante enake številu 0, za vsako elementarno matriko E.

Predpostavimo torej, da so vse $(k \times k)$-poddeterminante matrike A enake številu 0. Glede na tip elementarne matrike E imamo zdaj tri možnosti:

(1) Če je $E = [\mathscr{V}_i \leftarrow \alpha \mathscr{V}_i]$ za nek indeks $i \leq m$ in nek skalar $\alpha \in \mathbb{F} \setminus \{0\}$, potem je vsaka $(k \times k)$-poddeterminanta matrike EA enaka bodisi neki $(k \times k)$-poddeterminanti matrike A ali pa neki z α pomnoženi $(k \times k)$-poddeterminanti matrike A. Vse $(k \times k)$-poddeterminante matrike EA so torej enake številu 0.

(2) Če je $E = [\mathscr{V}_i \leftrightarrow \mathscr{V}_j]$ za neka indeksa $i \leq m$ in $j \leq m$, $i \neq j$, potem je vsaka $(k \times k)$-poddeterminanta matrike EA do predznaka natančno enaka neki

$(k \times k)$-poddeterminanti matrike A. Vse $(k \times k)$-poddeterminante matrike EA so zato enake številu 0.

(3) Predpostavimo, da je $E = [\mathscr{V}_i \leftarrow \mathscr{V}_i + \beta \mathscr{V}_j]$ za neka indeksa $i \leq m$ in $j \leq m$, $i \neq j$, in nek skalar $\beta \in \mathbb{F}$. Vzemimo poljubno podmatriko matrike A velikosti $k \times k$, torej podmatriko oblike

$$\mathrm{sub}_{\mathcal{I}\mathcal{J}}(A)$$

za neki podmnožici $\mathcal{I} \subset \{1, 2, \ldots, m\}$ in $\mathcal{J} \subset \{1, 2, \ldots, n\}$, ki imata obe po k elementov. Spet imamo tri možnosti:

(a) Če $i \notin \mathcal{I}$, potem je $\mathrm{sub}_{\mathcal{I}\mathcal{J}}(EA) = \mathrm{sub}_{\mathcal{I}\mathcal{J}}(A)$ in zato

$$\det(\mathrm{sub}_{\mathcal{I}\mathcal{J}}(EA)) = \det(\mathrm{sub}_{\mathcal{I}\mathcal{J}}(A)) = 0.$$

(b) Če velja $i, j \in \mathcal{I}$, potem matriko $\mathrm{sub}_{\mathcal{I}\mathcal{J}}(EA)$ dobimo iz matrike $\mathrm{sub}_{\mathcal{I}\mathcal{J}}(A)$ tako, da eni od vrstic matrike $\mathrm{sub}_{\mathcal{I}\mathcal{J}}(A)$ prištejemo β-kratnik neke druge vrstice matrike $\mathrm{sub}_{\mathcal{I}\mathcal{J}}(A)$. Odtod sledi

$$\det(\mathrm{sub}_{\mathcal{I}\mathcal{J}}(EA)) = \det(\mathrm{sub}_{\mathcal{I}\mathcal{J}}(A)) = 0.$$

(c) V primeru, da je $i \in \mathcal{I}$ in $j \notin \mathcal{I}$, iz multilinearnosti determinante sledi

$$\det(\mathrm{sub}_{\mathcal{I}\mathcal{J}}(EA)) = \det(\mathrm{sub}_{\mathcal{I}\mathcal{J}}(A)) + \beta \det(\mathrm{sub}_{\mathcal{I}\mathcal{J}}([\mathscr{V}_i \leftrightarrow \mathscr{V}_j]A)) = 0,$$

saj velja $\det(\mathrm{sub}_{\mathcal{I}\mathcal{J}}([\mathscr{V}_i \leftrightarrow \mathscr{V}_j]A)) = 0$ po točki (2). □

Posledica 1.81. *Za vsako matriko* $A \in \mathrm{Mat}(m \times n, \mathbb{F})$ *je* $\mathrm{rank}(A) = \mathrm{rank}(A^{\mathsf{t}})$.

Dokaz. Pri transponiranju matrike se tudi njene podmatrike transponirajo, zato ima matrika A^{t} enake poddeterminante kot matrika A. Trditev torej sledi iz trditve 1.80. □

Zgled 1.82. Matrika

$$A = \begin{bmatrix} 1 & 0 & 2 \\ 2 & 1 & 4 \\ 3 & 2 & 6 \end{bmatrix}$$

ima ničelno determinanto, saj je njen tretji stolpec dvakratnik njenega prvega stolpca. Ob tem ima matrika A neničelno (2×2)-poddeterminanto

$$\det(\mathrm{sub}^{33}(A)) = 1.$$

Odtod lahko zaključimo, da je $\mathrm{rank}(A) = 2$.

Vaja 1.83. (1) Naj bo p naravno število in naj bo $A \in \mathrm{Mat}(m \times n, \mathbb{F})$ takšna matrika, da so vse $(p \times p)$- poddeterminante matrike A enake številu 0. Pokaži, da so tedaj tudi vse $(k \times k)$-poddeterminante matrike A enake številu 0, za vsako naravno število $k > p$.

(2) Pokaži, da za vsako matriko $\mathrm{Mat}(m \times n, \mathbb{C})$ velja $\mathrm{rank}(A) = \mathrm{rank}(A^{\mathsf{h}})$.

Poglavje 2

Končno-dimenzionalni vektorski prostori

V prvem poglavju smo spoznali osnovne lastnosti matrik. Med drugim smo povedali, da zaradi lastnosti seštevanja matrik in množenja matrik s skalarji množici matrik velikosti $m \times n$ pravimo vektorski prostor. V tem razdelku bomo spoznali še nekatere druge primere vektorskih prostorov, nekatere splošne lastnosti vektorskih prostorov in linearne preslikave med vektorskimi prostori, s katerimi vektorske prostore lahko med seboj primerjamo.

Tudi v tem poglavju bomo z $\mathbb{F}$ označili izbran obseg, ki je bodisi obseg racionalnih števil $\mathbb{Q}$, obseg realnih števil $\mathbb{R}$ ali obseg kompleksnih števil $\mathbb{C}$. Elementom tega komutativnega obsega bomo rekli skalarji. Večina rezultatov, ki jih bomo spoznali, velja sicer tudi v primeru, da je $\mathbb{F}$ kakšen drug komutativen obseg karakteristike 0, vendar pa se s takimi obsegi v okviru naših predavanj ne bomo ukvarjali.

2.1. Vektorski prostori in linearne preslikave

Definicija 2.1. *Vektorski prostor* nad $\mathbb{F}$ je množica V, opremljena z operacijo *seštevanja*, ki je funkcija

$$\mathsf{V} \times \mathsf{V} \to \mathsf{V}, \qquad (v, w) \mapsto v + w,$$

in z operacijo *množenja s skalarji*, ki je funkcija

$$\mathbb{F} \times \mathsf{V} \to \mathsf{V}, \qquad (\alpha, v) \mapsto \alpha v,$$

tako da velja:

(i) $u + (v + w) = (u + v) + w$ za vse $u, v, w \in \mathsf{V}$,
(ii) $v + w = w + v$ za vse $v, w \in \mathsf{V}$,
(iii) obstaja tak element $0 \in \mathsf{V}$, da je $v + 0 = v$ za vsak $v \in \mathsf{V}$,
(iv) za vsak $v \in \mathsf{V}$ obstaja tak element $-v \in \mathsf{V}$, da je $v + (-v) = 0$,
(v) $\alpha(v + w) = \alpha v + \alpha w$ za vse $v, w \in \mathsf{V}$ in vsak $\alpha \in \mathbb{F}$,
(vi) $(\alpha + \beta)v = \alpha v + \beta v$ za vsak $v \in \mathsf{V}$ in vse $\alpha, \beta \in \mathbb{F}$,
(vii) $(\alpha\beta)v = \alpha(\beta v)$ za vsak $v \in \mathsf{V}$ vse $\alpha, \beta \in \mathbb{F}$, in
(viii) $1v = v$ za vsak $v \in \mathsf{V}$.

Komentar 2.2. Elementom vektorskega prostora običajno pravimo *vektorji*. Operacija seštevanja poljubnima dvema vektorjema v in w pridruži njuno *vsoto* $v + w$, ki je spet vektor. Operacija množenja s skalarji poljubnemu skalarju α in poljubnemu vektorju v priredi vektor αv, ki ga imenujemo *produkt* vektorja *v s skalarjem α*. Točke (i-iv) povedo, da je vektorski prostor V za seštevanje Abelova grupa. Kot v vsaki Abelovi grupi je nevtralni element za seštevanje 0 iz točke (iii), ki mu pravimo tudi *vektor nič*, en sam. Prav tako je za vsak vektor $v \in \mathsf{V}$ tudi njemu *nasprotni vektor* $-v$ iz točke (iv) en sam. Vektorje lahko tudi *odštevamo*, in sicer za poljubna vektorja v in w izračunamo njuno *razliko* po predpisu $v - w = v + (-w)$.

Vektorski prostor nad $\mathbb{C}$ (oziroma $\mathbb{R}$ ali $\mathbb{Q}$) imenujemo tudi *kompleksen* (oziroma *realen* ali *racionalen*) vektorski prostor.

ZGLED 2.3. (1) Komutativen obseg $\mathbb{F}$ je hkrati tudi vektorski prostor nad $\mathbb{F}$, pri čemer sta seštevanje in množenje s skalarji dana s seštevanjem in množenjem v $\mathbb{F}$. Množica $\{0\}$ z enim samim elementom $0 \in \mathbb{F}$ je prav tako vektorski prostor nad $\mathbb{F}$, ki ga bomo označili tudi z $\mathbb{F}^0 = \{0\}$.

(2) V prvem poglavju smo videli, da je množica matrik $\mathrm{Mat}(m \times n, \mathbb{F})$ vektorski prostor nad $\mathbb{F}$. Posebej je $\mathbb{F}^n = \mathrm{Mat}(n \times 1)$ vektorski prostor nad $\mathbb{F}$. Tako je $\mathbb{R}^n$ realen vektorski prostor. Najbolje poznamo seveda realen vektorski prostor $\mathbb{R}^3$.

(3) Naj bodo $\mathsf{V}_1, \mathsf{V}_2, \ldots, \mathsf{V}_k$ vektorski prostori nad $\mathbb{F}$. Kartezični produkt

$$\mathsf{V}_1 \times \mathsf{V}_2 \times \cdots \times \mathsf{V}_k$$

je množica, katere elementi so vse urejene k-terice $(v_1, v_2, \ldots, v_k)$, kjer je $v_i \in \mathsf{V}_i$ za vsak $i = 1, 2, \ldots, k$. Na tej množici definiramo operaciji seštevanja in množenja s skalarji *po komponentah*, torej

$$(v_1, v_2, \ldots, v_k) + (v_1', v_2', \ldots, v_k') = (v_1 + v_1', v_2 + v_2', \ldots, v_k + v_k')$$

in

$$\alpha(v_1, v_2, \ldots, v_k) = (\alpha v_1, \alpha v_2, \ldots, \alpha v_k)$$

za vse urejene k-terice $(v_1, v_2, \ldots, v_k), (v_1', v_2', \ldots, v_k') \in \mathsf{V}_1 \times \mathsf{V}_2 \times \cdots \times \mathsf{V}_k$ in za vse skalarje $\alpha \in \mathbb{F}$. Lahko je preveriti, da je s tema dvema operacijama

$$\mathsf{V}_1 \times \mathsf{V}_2 \times \cdots \times \mathsf{V}_k$$

vektorski prostor nad $\mathbb{F}$, ki mu pravimo *produkt vektorskih prostorov* $\mathsf{V}_1, \mathsf{V}_2, \ldots, \mathsf{V}_k$. Primer takšnega produkta je vektorski prostor $\mathbb{F}^n$, ki je enak produktu n kopij vektorskega prostora $\mathbb{F}$,

$$\mathbb{F}^n = \mathbb{F} \times \mathbb{F} \times \cdots \times \mathbb{F}.$$

(4) Naj bo V poljuben vektorski prostor nad $\mathbb{F}$, naj bo Ω poljubna množica in naj bo V^Ω množica vseh funkcij iz Ω v V. Na množici V^Ω imamo operaciji seštevanja in množenja s skalarji definirani *po točkah*: za poljubni funkciji $f, g \in \mathsf{V}^\Omega$ in poljuben skalar $\alpha \in \mathbb{F}$ sta $f + g$ in αf funkciji iz Ω v V, dani s predpisoma

$$(f + g)(\omega) = f(\omega) + g(\omega)$$

in

$$(\alpha f)(\omega) = \alpha f(\omega)$$

za vsak $\omega \in \Omega$. Direktno po definiciji lahko preverimo, da je glede na ti dve operaciji množica V^Ω vektorski prostor nad $\mathbb{F}$. Ničelni element v vektorskem prostoru V^Ω je *ničelna preslikava*, ki je konstantna preslikava z vrednostjo 0.

TRDITEV 2.4. *Naj bo V vektorski prostor nad $\mathbb{F}$. Tedaj za vsak vektor $v \in \mathsf{V}$ in za vsak skalar $\alpha \in \mathbb{F}$ velja:*

(i) $0v = 0$,
(ii) $(-1)v = -v$,
(iii) *produkt vektorja 0 s skalarjem α je enak vektorju 0, in*
(iv) *če je $\alpha v = 0$ in $\alpha \neq 0$, potem je $v = 0$.*

Dokaz. (i) Z uporabo definicije vektorskega prostora izračunamo

$$0v = 0v + 1v - 1v = (0+1)v - 1v = 1v - 1v = 0.$$

(ii) Iz točke (i) dobimo

$$(-1)v + v = (-1)v + 1v = (-1+1)v = 0v = 0,$$

in ker je nasprotni vektor vektorja v en sam, mora biti $-v = (-1)v$.

(iii) Produkt vektorja 0 s skalarjem α dobimo iz enakosti

$$\alpha 0 = \alpha 0 + \alpha 0 - \alpha 0 = \alpha(0+0) - \alpha 0 = \alpha 0 - \alpha 0 = 0.$$

(iv) Če je $\alpha v = 0$ in $\alpha \neq 0$, potem iz točke (iii) sledi

$$v = 1v = (\alpha^{-1}\alpha)v = \alpha^{-1}(\alpha v) = \alpha^{-1}0 = 0. \qquad \square$$

Definicija 2.5. Naj bo V vektorski prostor nad $\mathbb{F}$. *Vektorski podprostor* prostora V je neprazna podmnožica U prostora V, za katero velja:

(i) za poljubna vektorja $u, u' \in \mathsf{U}$ je $u + u' \in \mathsf{U}$, in
(ii) za vsak vektor $u \in \mathsf{U}$ in za vsak skalar $\alpha \in \mathbb{F}$ je $\alpha u \in \mathsf{U}$.

Komentar 2.6. Če za podmnožico $\mathsf{U} \subset \mathsf{V}$ velja lastnost (i), potem pravimo, da je podmnožica U *zaprta za seštevanje*. Če velja točka (ii), potem pravimo, da je podmnožica U *zaprta za množenje s skalarji*.

Vsak vektorski podprostor U vektorskega prostora V je posebej tudi vektorski prostor za operaciji, ki jih dobimo z zožitvijo operacij vektorskega prostora V na vektorski podprostor U. Natančneje, seštevanje $\mathsf{V} \times \mathsf{V} \to \mathsf{V}$ zožimo na množico $\mathsf{U} \times \mathsf{U}$, po točki (i) pa ta zožitev slika v množico U in nam da operacijo seštevanja na množici U. Podobno množenje s skalarji $\mathbb{F} \times \mathsf{V} \to \mathsf{V}$ zožimo na množico $\mathbb{F} \times \mathsf{U}$, po točki (ii) pa tudi ta zožitev slika v množico U in nam da operacijo množenja s skalarji na množici U. Ob tem lahko hitro preverimo, da tako definirani operaciji na U zadoščata vsem pogojem iz definicije vektorskega prostora. Na primer, ker množica U ni prazna, obstaja nek vektor $u \in \mathsf{U}$, po točki (ii) iz definicije vektorskega podprostora pa je tedaj tudi vektor $0 = 0u$ v možici U. Podobno po točki (ii) iz definicije vektorskega podprostora sledi, da je tudi nasprotni vektor $-u = (-1)u$ poljubnega vektorja $u \in \mathsf{U}$ spet v množici U.

Zgled 2.7. (1) Vsak vektorski prostor V ima trivialni vektorski podprostor $\{0\}$, pa tudi cel V je vektorski podprostor vektorskega prostora V.

(2) Premica v $\mathbb{R}^3$ je vektorski podprostor vektorskega prostora $\mathbb{R}^3$, če, in samo če, gre skozi izhodišče. Podobno je ravnina v $\mathbb{R}^3$ vektorski podprostor vektorskega prostora $\mathbb{R}^3$, če, in samo če, gre skozi izhodišče.

Splošneje, po trditvi 1.44 je množica rešitev homogenega sistema linearnih enačb za n neznank iz $\mathbb{F}$ vektorski podprostor vektorskega prostora $\mathbb{F}^n$. Velja pa tudi obratno: če sistem linearnih enačb za n neznank ni homogen, potem vektor 0 ni rešitev tega sistema in zato množica rešitev tega sistema ni vektorski podprostor vektorskega prostora $\mathbb{F}^n$.

(3) Za poljubno podmnožico $\mathcal{D} \subset \mathbb{R}^n$ je množica $C(\mathcal{D}, \mathbb{R}^m)$ vseh zveznih vektorskih funkcij $\mathcal{D} \to \mathbb{R}^m$ za operacije po točkah vektorski podprostor realnega vektorskega prostora $(\mathbb{R}^m)^{\mathcal{D}}$.

(4) Za poljubno odprto podmnožico $\mathcal{D} \subset \mathbb{R}^n$ in za vsak $k \in \mathbb{N}$ je množica $C^k(\mathcal{D}, \mathbb{R}^m)$ vseh k-krat zvezno odvedljivih vektorskih funkcij $\mathcal{D} \to \mathbb{R}^m$ za operacije po točkah vektorski podprostor realnega vektorskega prostora $C(\mathcal{D}, \mathbb{R}^m)$.

Podobno je tudi množica $C^\infty(\mathcal{D}, \mathbb{R}^m)$ vseh gladkih vektorskih funkcij $\mathcal{D} \to \mathbb{R}^m$ vektorski podprostor realnega vektorskega prostora $C(\mathcal{D}, \mathbb{R}^m)$.

(5) Množica $\mathbb{F}[t]$ vseh polinomov v spremenljivki t in s koeficienti v $\mathbb{F}$ je podmnožica vektorskega prostora $\mathbb{F}^\mathbb{F}$, saj polinome s koeficienti v $\mathbb{F}$ lahko gledamo kot funkcije $\mathbb{F} \to \mathbb{F}$. Ker je vsota dveh polinomov spet polinom in ker je produkt polinoma s skalarjem prav tako spet polinom, je $\mathbb{F}[t]$ za operacije po točkah vektorski prostor nad $\mathbb{F}$.

Za poljubno nenegativno celo število s označimo z

$$\mathbb{F}_s[t]$$

množico vseh polinomov v spremenljivki t in s koeficienti v $\mathbb{F}$, katerih stopnja je največ s. Opazimo lahko, da je tudi ta množica zaprta za seštevanje in množenje s skalarji, zato je $\mathbb{F}_s[t]$ vektorski podprostor vektorskega prostora $\mathbb{F}[t]$.

2.1.1. Linearne kombinacije. Naj bodo $v_1, v_2, \dots, v_k$ vektorji iz vektorskega prostora V nad $\mathbb{F}$. Izraz oblike

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k,$$

kjer so $\alpha_1, \alpha_2, \dots, \alpha_k \in \mathbb{F}$ poljubni skalarji, imenujemo *linearna kombinacija* vektorjev $v_1, v_2, \dots, v_k$. Rezultat oziroma vrednost takšne linearne kombinacije je seveda vektor iz vektorskega prostora V. Vektor $v \in \mathsf{V}$ se da zapisati kot linearna kombinacija (ali, krajše rečeno, vektor v je linearna kombinacija) vektorjev $v_1, v_2, \dots, v_k$, če obstajajo takšni skalarji $\alpha_1, \alpha_2, \dots, \alpha_k \in \mathbb{F}$, da je $v = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k \in \mathsf{V}$. Množico vrednosti vseh linearnih kombinacij vektorjev $v_1, v_2, \dots, v_k$ imenujemo tudi *linearna ogrinjača* vektorjev $v_1, v_2, \dots, v_k$ v vektorskem prostoru V in jo označimo

$$\begin{aligned}&\mathrm{Span}\{v_1, v_2, \dots, v_k\} = \mathrm{Span}_\mathsf{V}\{v_1, v_2, \dots, v_k\}\\ &= \{\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k \in \mathsf{V}\ ;\ \alpha_1, \alpha_2, \dots, \alpha_k \in \mathbb{F}\} \subset \mathsf{V}.\end{aligned}$$

Definicijo linearne ogrinjače lahko nekoliko razširimo. Za poljubno neprazno podmnožico $\mathcal{O} \subset \mathsf{V}$, ki je lahko tudi neskončna, naj bo

$$\begin{aligned}&\mathrm{Span}(\mathcal{O}) = \mathrm{Span}_\mathsf{V}(\mathcal{O})\\ &= \{\alpha_1 u_1 + \cdots + \alpha_p u_p \in \mathsf{V}\ ;\ p \in \mathbb{N},\ u_1, \dots, u_p \in \mathcal{O},\ \alpha_1, \dots, \alpha_p \in \mathbb{F}\} \subset \mathsf{V}\end{aligned}$$

množica vseh tistih vektorjev iz V, ki jih na kakšen način lahko zapišemo kot linearno kombinacijo vektorjev iz množice $\mathcal{O}$. Množico $\mathrm{Span}_\mathsf{V}(\mathcal{O})$ imenujemo *linearna ogrinjača* podmnožice $\mathcal{O}$ v vektorskem prostoru V. Seveda je linearna ogrinjača vektorjev $v_1, v_2, \dots, v_k$ enaka linearni ogrinjači podmnožice $\{v_1, v_2, \dots, v_k\}$. Posebej se dogovorimo, da je $\mathrm{Span}(\emptyset) = \mathrm{Span}_\mathsf{V}(\emptyset) = \{0\}$.

TRDITEV 2.8. *Naj bo $\mathcal{O}$ poljubna podmnožica vektorskega prostora V nad $\mathbb{F}$. Linearna ogrinjača množice $\mathcal{O}$ je najmanjši vektorski podprostor vektorskega prostora V, ki ima $\mathcal{O}$ za svojo podmnožico. Podmnožica $\mathcal{O}$ je vektorski podprostor vektorskega prostora V, če, in samo če, velja $\mathrm{Span}(\mathcal{O}) = \mathcal{O}$.*

DOKAZ. Direkten izračun pokaže, da se da vsota dveh linearnih kombinacij vektorjev iz $\mathcal{O}$ zapisati kot linearna kombinacija vektorjev iz $\mathcal{O}$ in da se s skalarjem pomnožena linearna kombinacija vektorjev iz $\mathcal{O}$ prav tako da zapisati kot linearna kombinacija vektorjev iz $\mathcal{O}$. Tako smo se prepričali, da je $\mathrm{Span}(\mathcal{O})$ vektorski podprostor vektorskega prostora V. Očitno velja $\mathcal{O} \subset \mathrm{Span}(\mathcal{O})$, saj vsak vektor $u \in \mathcal{O}$ lahko zapišemo kot linearno kombinacijo $u = 1u$. Preostali del trditve sledi iz dejstva, da vsak vektorski podprostor prav gotovo vsebuje rezultate vseh linearnih kombinacij svojih vektorjev. □

Vaja 2.9. Naj bosta $\mathcal{O}$ in $\mathcal{O}'$ takšni podmnožici vektorskega prostora V nad $\mathbb{F}$, da je $\mathcal{O} \subset \mathcal{O}'$. Pokaži, da tedaj velja:

(i) $\mathrm{Span}(\mathcal{O}) \subset \mathrm{Span}(\mathcal{O}')$,
(ii) $\mathrm{Span}(\mathrm{Span}(\mathcal{O})) = \mathrm{Span}(\mathcal{O})$, in
(iii) če je $\mathrm{Span}(\mathcal{O}) = \mathsf{V}$, potem je tudi $\mathrm{Span}(\mathcal{O}') = \mathsf{V}$.

Definicija 2.10. Naj bo V vektorski prostor nad $\mathbb{F}$. Vektorji $v_1, v_2, \ldots, v_k \in \mathsf{V}$ *generirajo* (ali *razpenjajo*) vektorski prostor V, če je $\mathrm{Span}\{v_1, v_2, \ldots, v_k\} = \mathsf{V}$. Podmnožica $\mathcal{O} \subset \mathsf{V}$ *generira* (ali *razpenja*) vektorski prostor V, če je $\mathrm{Span}(\mathcal{O}) = \mathsf{V}$.

Z drugimi besedami, vektorji $v_1, v_2, \ldots, v_k \in \mathsf{V}$ generirajo vektorski prostor V, če, in samo če, lahko vsak vektor iz V zapišemo kot linearno kombinacijo vektorjev $v_1, v_2, \ldots, v_k$. Pravimo tudi, da tedaj vektorji $v_1, v_2, \ldots, v_k$ sestavljajo *ogrodje* vektorskega prostora V.

Trditev 2.11. *Vektorji $v_1, v_2, \ldots, v_k \in \mathbb{F}^n$ generirajo vektorski prostor $\mathbb{F}^n$, če, in samo če, velja*

$$\mathrm{rank}\begin{bmatrix} v_1 & v_2 & \cdots & v_k \end{bmatrix} = n.$$

V tem primeru torej velja $k \geq n$.

Dokaz. Spomnimo se, da so vektorji v $\mathbb{F}^n$ urejene n-terice oziroma stolpci. Stolpci $v_1, v_2, \ldots, v_k$ nam skupaj sestavljajo matriko

$$A = \begin{bmatrix} v_1 & v_2 & \cdots & v_k \end{bmatrix} \in \mathrm{Mat}(n \times k, \mathbb{F}).$$

Za poljuben stolpec $X \in \mathrm{Mat}(k \times 1, \mathbb{F})$, $X = [x_j]$, tedaj velja

$$AX = x_1 v_1 + x_2 v_2 + \cdots + x_k v_k.$$

Zapisati nek vektor $v \in \mathbb{F}^n$ kot linearno kombinacijo vektorjev $v_1, v_2, \ldots, v_k$ pomeni torej nič drugega kot poiskati rešitev sistema linearnih enačb $AX = v$.

Če je rang matrike A enak številu n, potem je tudi rang razširjene matrike sistema $AX = v$ enak številu n, saj ima ta matrika n vrstic. Po izreku 1.46 ima tedaj sistem rešitev za vsak $v \in \mathbb{F}^n$, zato vektorji $v_1, v_2, \ldots, v_k \in \mathbb{F}^n$ generirajo vektorski prostor $\mathbb{F}^n$.

Predpostavimo zdaj, da število $\mathrm{rank}(A)$ ni enako številu n. To pomeni, da je $\mathrm{rank}(A) < n$, saj je število vrstic matrike A enako številu n. Izberimo takšno obrnljivo matriko $E \in \mathrm{GL}(n, \mathbb{F})$, da je matrika EA v reducirani vrstični kanonični formi. Ker je $\mathrm{rank}(A) < n$, je zadnja vrstica matrike EA ničelna. Za vektor $\mathrm{e}_n = (0, 0, \ldots, 0, 1) \in \mathbb{F}^n$ torej sistem $(EA)X = \mathrm{e}_n$ nima rešitev, odtod pa sledi, da tudi sistem $AX = E^{-1}\mathrm{e}_n$ nima rešitev. V tem primeru vektorji $v_1, v_2, \ldots, v_k \in \mathbb{F}^n$ torej ne generirajo vektorskega prostora $\mathbb{F}^n$. □

Zgled 2.12. (1) Cel vektorski prostor V seveda generira vektorski prostor V nad $\mathbb{F}$. Prazna množica generira trivialni vektorski podprostor $\{0\} \subset \mathsf{V}$. Za poljuben vektor $v \in \mathsf{V}$ v vektorskem prostoru V nad $\mathbb{F}$ je

$$\mathrm{Span}\{v\} = \{\alpha v \; ; \; \alpha \in \mathbb{F}\},$$

zato pogosto krajše označimo $\mathrm{Span}\{v\} = \mathbb{F}v$. Posebej za $v = 0$ velja $\mathrm{Span}\{0\} = \{0\}$.

(2) Naj bodo $u = (1, 0)$, $v = (2, 1)$ in $w = (3, 1)$ vektorji v $\mathbb{R}^2$. Matrika

$$A = \begin{bmatrix} u & v & w \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 1 \end{bmatrix}$$

ima rang enak številu 2, saj je $\det(\mathrm{sub}_{\{1,2\}\{1,2\}}(A)) = 1$. Odtod sledi, da vektorji u, v, w generirajo vektorski prostor $\mathbb{R}^2$.

2.1.2. Linearna neodvisnost in baze vektorskega prostora. V vektorskem prostoru V nad $\mathbb{F}$ so vektorji $v_1, v_2, \ldots, v_k \in \mathsf{V}$ med seboj *linearno odvisni*, če obstajajo takšni skalarji $\alpha_1, \alpha_2, \ldots, \alpha_k \in \mathbb{F}$, da je vsaj eden od teh skalarjev neničelen in da velja

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k = 0.$$

V vektorskem prostoru V nad $\mathbb{F}$ so vektorji $v_1, v_2, \ldots, v_k$ med seboj *linearno neodvisni*, če niso med seboj linearno odvisni.

V vektorskem prostoru V nad $\mathbb{F}$ je podmnožica $\mathcal{O} \subset \mathsf{V}$ *linearno odvisna*, če obstajajo takšno naravno število p, takšni paroma različni vektorji $u_1, u_2, \ldots, u_p \in \mathcal{O}$ in takšni skalarji $\alpha_1, \alpha_2, \ldots, \alpha_p \in \mathbb{F}$, da je vsaj eden od teh skalarjev neničelen in da velja

$$\alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_p u_p = 0.$$

V vektorskem prostoru V nad $\mathbb{F}$ je podmnožica $\mathcal{O} \subset \mathsf{V}$ *linearno neodvisna*, če ni linearno odvisna.

Komentar 2.13. (1) Če so torej vektorji $v_1, v_2, \ldots, v_k \in \mathsf{V}$ med seboj linearno neodvisni in če so $\alpha_1, \alpha_2, \ldots, \alpha_k \in \mathbb{F}$ takšni skalarji, da velja

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k = 0,$$

potem odtod sledi, da je $\alpha_1 = \alpha_2 = \cdots = \alpha_k = 0$. Velja pa tudi obratno: če ima enačba

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k = 0$$

za neznane skalarje $\alpha_1, \alpha_2, \ldots, \alpha_k \in \mathbb{F}$ le trivialno rešitev, torej $\alpha_1 = \alpha_2 = \cdots = \alpha_k = 0$, potem so vektorji $v_1, v_2, \ldots, v_k \in \mathsf{V}$ med seboj linearno neodvisni.

(2) Linearni kombinaciji

$$0v_1 + 0v_2 + \cdots + 0v_k$$

pravimo tudi *trivialna* linearna kombinacija vektorjev $v_1, v_2, \ldots, v_k \in \mathsf{V}$. Rezultat trivialne linearne kombinacije je seveda vektor 0. Linearna kombinacija

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k$$

je *netrivialna*, če ni trivialna, ali drugače povedano, če je vsaj eden od skalarjev $\alpha_1, \alpha_2, \ldots, \alpha_k$ različen od nič.

Lahko bi torej rekli, da so vektorji $v_1, v_2, \ldots, v_k \in \mathsf{V}$ med seboj linearno neodvisni, če, in samo če, je rezultat vsake netrivialne linearne kombinacije teh vektorjev neničelen.

(3) Paroma različni vektorji $v_1, v_2, \ldots, v_k \in \mathsf{V}$ so med seboj linearno neodvisni, če, in samo če, je množica $\{v_1, v_2, \ldots, v_k\}$ linearno neodvisna. Če sta si med vektorji $v_1, v_2, \ldots, v_k$ vsaj dva enaka, potem so vektorji $v_1, v_2, \ldots, v_k$ med seboj linearno odvisni. Res, če velja $v_i = v_j$ za neka indeksa i in j, $1 \leq i < j \leq k$, potem lahko vektor 0 zapišemo kot netrivialno linearno kombinacijo

$$0v_1 + \cdots + 0v_{i-1} + 1v_i + 0v_{i+1} + \cdots + 0v_{j-1} + (-1)v_j + 0v_{j+1} + \cdots + 0v_k.$$

Če je vsaj eden izmed vektorjev $v_1, v_2, \ldots, v_k$ enak vektorju 0, potem so vektorji $v_1, v_2, \ldots, v_k$ med seboj linearno odvisni. Če namreč velja $v_i = 0$ nek indeks i, $1 \leq i \leq k$, potem lahko vektor 0 zapišemo kot netrivialno linearno kombinacijo

$$0v_1 + \cdots + 0v_{i-1} + 1v_i + 0v_{i+1} + \cdots + 0v_k.$$

(4) Prazna množica je po definiciji linearno neodvisna. Množica z enim samim vektorjem $v_1 \in \mathsf{V}$ je linearno odvisna, če, in samo če, je $v_1 = 0$.

(5) Če je $\mathcal{O}$ linearno neodvisna podmnožica vektorskega prostora V nad $\mathbb{F}$, potem je v vektorskem prostoru V tudi vsaka podmnožica množice $\mathcal{O}$ linearno neodvisna.

Trditev 2.14. *V vektorskem prostoru V nad $\mathbb{F}$ so vektorji $v_1, v_2, \ldots, v_k \in \mathsf{V}$ med seboj linearno odvisni, če, in samo če, lahko enega od njih zapišemo kot linearno kombinacijo ostalih.*

Dokaz. Naj bodo vektorji $v_1, v_2, \ldots, v_k$ med seboj linearno odvisni. Obstajajo torej skalarji $\alpha_1, \alpha_2, \ldots, \alpha_k \in \mathbb{F}$, tako da je

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k = 0$$

in da je vsaj eden od teh skalarjev, recimo α_i, različen od 0. Tedaj lahko vektor v_i zapišemo kot linearno kombinacijo

$$v_i = \Big(-\frac{\alpha_1}{\alpha_i}\Big)v_1 + \cdots + \Big(-\frac{\alpha_{i-1}}{\alpha_i}\Big)v_{i-1} + \Big(-\frac{\alpha_{i+1}}{\alpha_i}\Big)v_{i+1} + \cdots + \Big(-\frac{\alpha_k}{\alpha_i}\Big)v_k.$$

Obratno, predpostavimo zdaj, da je eden od vektorjev $v_1, v_2, \ldots, v_k$, recimo v_i, vrednost neke linearne kombinacije ostalih. Velja torej

$$v_i = \beta_1 v_1 + \cdots + \beta_{i-1} v_{i-1} + \beta_{i+1} v_{i+1} + \cdots + \beta_k v_k$$

za neke skalarje $\beta_1, \ldots, \beta_{i-1}, \beta_{i+1}, \ldots, \beta_k \in \mathbb{F}$. Tedaj imamo netrivialno linearno kombinacijo

$$\beta_1 v_1 + \cdots + \beta_{i-1} v_{i-1} + (-1)v_i + \beta_{i+1} v_{i+1} + \cdots + \beta_k v_k,$$

katere vrednost je enaka vektorju 0. □

Trditev 2.15. *V vektorskem prostoru V nad $\mathbb{F}$ so vektorji $v_1, v_2, \ldots, v_k \in \mathsf{V}$ med seboj linearno neodvisni, če, in samo če, je vsak vektor $v \in \mathsf{V}$ rezultat največ ene linearne kombinacije vektorjev $v_1, v_2, \ldots, v_k$.*

Dokaz. Če je vsak vektor $v \in \mathsf{V}$ rezultat največ ene linearne kombinacije vektorjev $v_1, v_2, \ldots, v_k$, je torej vektor $v = 0$ rezultat le trivialne linearne kombinacije vektorjev $v_1, v_2, \ldots, v_k$. To pomeni, da so vektorji $v_1, v_2, \ldots, v_k$ med seboj linearno neodvisni.

Predpostavimo zdaj, da so vektorji $v_1, v_2, \ldots, v_k$ med seboj linearno neodvisni in da je

$$v = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k = \beta_1 v_1 + \beta_2 v_2 + \cdots + \beta_k v_k$$

za neke skalarje $\alpha_1, \beta_1, \alpha_2, \beta_2, \ldots, \alpha_k, \beta_k \in \mathbb{F}$. Odtod lahko izračunamo, da velja

$$(\alpha_1 - \beta_1)v_1 + (\alpha_2 - \beta_2)v_2 + \cdots + (\alpha_k - \beta_k)v_k = 0.$$

Zaradi medsebojne linearne neodvisnosti vektorjev $v_1, v_2, \ldots, v_k$ mora biti ta linearna kombinacija trivialna, torej $\alpha_i = \beta_i$ za vse $i = 1, 2, \ldots, k$. □

TRDITEV 2.16. *V vektorskem prostoru $\mathbb{F}^n$ so vektorji $v_1, v_2, \dots, v_k \in \mathbb{F}^n$ med seboj linearno neodvisni, če, in samo če, velja*

$$\operatorname{rank} \begin{bmatrix} v_1 & v_2 & \cdots & v_k \end{bmatrix} = k.$$

V tem primeru torej velja $k \leq n$.

DOKAZ. Vektorji oziroma stolpci $v_1, v_2, \dots, v_k \in \mathbb{F}^n = \mathrm{Mat}(n \times 1, \mathbb{F})$ skupaj sestavljajo matriko

$$A = \begin{bmatrix} v_1 & v_2 & \cdots & v_k \end{bmatrix} \in \mathrm{Mat}(n \times k, \mathbb{F}).$$

Za poljuben stolpec $X \in \mathrm{Mat}(k \times 1, \mathbb{F})$, $X = [x_j]$, velja

$$AX = x_1 v_1 + x_2 v_2 + \cdots + x_k v_k.$$

Zapisati ničelni vektor kot linearno kombinacijo vektorjev $v_1, v_2, \dots, v_k$ pomeni torej nič drugega kot poiskati rešitev homogenega sistema linearnih enačb $AX = 0$.

Če je rang matrike A enak številu k, ima po izreku 1.46 sistem $AX = 0$ natanko eno rešitev, torej $X = 0$. To pomeni, da so vektorji $v_1, v_2, \dots, v_k$ med seboj linearno neodvisni.

Predpostavimo, da sta števili $\operatorname{rank}(A)$ in k različni. Iz izreka 1.46 sledi, da ima sistem $AX = 0$ tedaj neskončno rešitev, in posebej, da ima tudi netrivialne rešitve. Vektorji $v_1, v_2, \dots, v_k$ so torej v tem primeru med seboj linearno odvisni. □

ZGLED 2.17. (1) Vektorja v, w v $\mathbb{R}^3$ sta med seboj linearno odvisna, če, in samo če, je eden od njiju večkratnik drugega. To je res natanko tedaj, ko sta si vektorja v, w vzporedna oziroma kolinearna, ali ekvivalentno, ko je vektorski produkt $v \times w$ enak vektorju 0.

(2) Vektorji u, v, w v $\mathbb{R}^3$ so med seboj linearno odvisni, če, in samo če, ima matrika $\begin{bmatrix} u & v & w \end{bmatrix}$ ničelno determinanto, to pa je res natanko tedaj, ko je mešani produkt $(u \times v) \cdot w$ enak številu 0 oziroma ko so vektorji u, v, w koplanarni.

(3) Naj bodo $u = (1, 1, 2)$, $v = (1, 0, 1)$ in $w = (0, 1, 1)$ vektorji v $\mathbb{R}^3$. Ker je determinanta matrike

$$A = \begin{bmatrix} u & v & w \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 2 & 1 & 1 \end{bmatrix}$$

enaka številu 0, je $\operatorname{rank}(A) < 3$ in so zato vektorji u, v, w med seboj linearno odvisni. Po drugi strani pa je rang matrike

$$B = \begin{bmatrix} u & v \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 2 & 1 \end{bmatrix}$$

enak številu 2, saj je $\det(\mathrm{sub}_{\{1,2\}\{1,2\}}(B)) = \mathrm{co}_{33}(A) = -1$. Vektorja v in w sta torej med seboj linearno neodvisna. Podobno vidimo, da sta vektorja u, w med seboj linearno neodvisna, saj je $\mathrm{co}_{32}(A) = -1$. Prav tako sta vektorja v, w med seboj linearno neodvisna, saj velja $\mathrm{co}_{31}(A) = 1$.

(4) V realnem vektorskem prostoru $C^\infty(\mathbb{R})$ vseh realnih gladkih funkcij ene realne spremenljivke $\mathbb{R} \to \mathbb{R}$ imamo za vsak $n \in \mathbb{N}$ funkcijo f_n dano s prepisom $f_n(x) = \mathrm{e}^{nx}$, za vsak $x \in \mathbb{R}$. Tedaj je množica funkcij

$$\{f_n \; ; \; n \in \mathbb{N}\} \subset C^\infty(\mathbb{R})$$

linearno neodvisna. Res, pa recimo, da je k naravno število in da so $\alpha_1, \alpha_2, \ldots, \alpha_k$ takšne realne konstante, da velja

$$\alpha_1 f_1 + \alpha_2 f_2 + \cdots + \alpha_k f_k = 0.$$

To pomeni, da je

$$\alpha_1 \mathrm{e}^x + \alpha_2 \mathrm{e}^{2x} + \cdots + \alpha_k \mathrm{e}^{kx} = 0$$

za vsak $x \in \mathbb{R}$. Če to enačbo p-krat odvajamo in v rezultat vstavimo $x = 0$, dobimo enakost

$$\alpha_1 + 2^p \alpha_2 + \cdots + k^p \alpha_k = 0$$

za vsak $p = 0, 1, 2, \ldots$ Matrika A velikosti $k \times k$, dana s predpisom $A_{ij} = j^{(i-1)}$, je obrnljiva. Z indukcijo na k namreč lahko dokažemo, da je njena determinanta neničelna (gre za poseben primer Vandermondeove determinante). Odtod sledi $\alpha_1 = \alpha_2 = \cdots = \alpha_k = 0$.

Trditev 2.18. *Naj bo* V *vektorski prostor nad* $\mathbb{F}$. *Če je* $\mathcal{O} \subset \mathsf{V}$ *linearno neodvisna podmnožica vektorskega prostora* V *in če je* v *vektor iz množice* $\mathsf{V} \setminus \mathrm{Span}(\mathcal{O})$, *potem je tudi* $\mathcal{O} \cup \{v\}$ *linearno neodvisna podmnožica vektorskega prostora* V.

Dokaz. Naj bodo $u_1, u_2, \ldots, u_p \in \mathcal{O}$ paroma različni vektorji iz množice $\mathcal{O}$ in naj bodo $\alpha_1, \alpha_2, \ldots, \alpha_p, \beta \in \mathbb{F}$ takšni skalarji, da je

$$\alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_p u_p + \beta v = 0.$$

Če bi bil skalar β različen od 0, bi odtod sledilo, da je vektor v rezultat linearne kombinacije vektorjev $u_1, u_2, \ldots, u_p$,

$$v = \Big(-\frac{\alpha_1}{\beta}\Big) u_1 + \Big(-\frac{\alpha_2}{\beta}\Big) u_2 + \cdots + \Big(-\frac{\alpha_p}{\beta}\Big) u_p,$$

to pa je v protislovju s predpostavko, da vektor v ni v linearni ogrinjači $\mathrm{Span}(\mathcal{O})$. Velja torej $\beta = 0$ in torej

$$\alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_p u_p = 0.$$

Ker je množica $\mathcal{O}$ linearno neodvisna, odtod sledi $\alpha_1 = \alpha_2 = \cdots = \alpha_p = 0$. $\square$

Definicija 2.19. Naj bo V vektorski prostor nad $\mathbb{F}$. Vektorji $v_1, v_2, \ldots, v_n \in \mathsf{V}$ *sestavljajo bazo* vektorskega prostora V, če generirajo vektorski prostor V in so v vektorskem prostoru V med seboj linearno neodvisni.

Komentar 2.20. Če vektorji $v_1, v_2, \ldots, v_n \in \mathsf{V}$ sestavljajo bazo vektorskega prostora V, potem urejeno n-terico vektorjev $v_1, v_2, \ldots, v_n$ imenujemo *baza* vektorskega prostora V in jo zapišemo

$$\mathscr{B} = [v_1, v_2, \ldots, v_n].$$

Takšni bazi $\mathscr{B} = [v_1, v_2, \ldots, v_n]$ včasih bolj natančno pravimo tudi *urejena končna baza.* Baze namreč lahko podamo tudi kot množice in ne kot urejene n-terice. Obstaja tudi pojem neskončne baze vektorskega prostora, vendar se z njim v okviru teh predavanj ne bomo ukvarjali. V okviru teh predavanj bodo vse naše baze končne in urejene, zato tega ne bomo vsakič posebej poudarjali in bomo končnim urejenim bazam na kratko rekli *baze* vektorskih prostorov.

Posledica 2.21. *Vektorji* $v_1, v_2, \ldots, v_n \in \mathsf{V}$ *sestavljajo bazo vektorskega prostora* V *nad* $\mathbb{F}$, *če, in samo če, lahko vsak vektor* $v \in \mathsf{V}$ *na en sam način zapišemo kot linearno kombinacijo vektorjev* $v_1, v_2, \ldots, v_n \in \mathsf{V}$.

DOKAZ. Direktna posledica trditve 2.15 in definicije 2.10. □

KOMENTAR 2.22. Če je $\mathscr{B} = [v_1, v_2, \ldots, v_n]$ baza vektorskega prostora V in če nek vektor $v \in \mathsf{V}$ zapišemo kot linearno kombinacijo teh baznih vektorjev,

$$v = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n,$$

potem tej linearni kombinaciji pravimo *razvoj* vektorja v *po bazi* $\mathscr{B}$. Razvoj po bazi torej vedno obstaja in je enoličen, kar pomeni, da so za vsak vektor v skalarji $\alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbb{F}$ enolično določeni. Ti skalarji, ki jih imenujemo tudi *koordinate* vektorja v v bazi $\mathscr{B}$, sestavljajo vektor

$$[v]_{\mathscr{B}} = (\alpha_1, \alpha_2, \ldots, \alpha_n) \in \mathbb{F}^n,$$

ki ga imenujemo *koordinatni vektor* vektorja v v bazi $\mathscr{B}$.

TRDITEV 2.23. *V vektorskem prostoru $\mathbb{F}^n$ vektorji $v_1, v_2, \ldots, v_k \in \mathbb{F}^n$ sestavljajo bazo, če, in samo če, velja*

$$\operatorname{rank}\begin{bmatrix} v_1 & v_2 & \cdots & v_k \end{bmatrix} = k = n.$$

DOKAZ. Direktna posledica trditve 2.16 in trditve 2.11. □

KOMENTAR 2.24. Trditev 2.23 posebej torej pove, da je vsaka baza vektorskega prostora $\mathbb{F}^n$ sestavljena iz n vektorjev, ki so stolpci neke obrnljive matrike velikosti $n \times n$. S predpisom $[v_1, v_2, \ldots, v_n] \mapsto \begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix}$ je podana bijekcija med množico vseh baz vektorskega prostora $\mathbb{F}^n$ in splošno linearno grupo $\mathrm{GL}(n, \mathbb{F})$. Po definiciji koordinatni vektor $[v]_{\mathscr{B}}$ danega vektorja $v \in \mathbb{F}^n$ v bazi $\mathscr{B} = [v_1, v_2, \ldots, v_n]$ reši sistem n linearnih enačb za n neznank

$$\begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix} [v]_{\mathscr{B}} = v.$$

Če želimo torej vektor v razviti po bazi $\mathscr{B}$ in s tem izračunati koordinatni vektor $[v]_{\mathscr{B}}$, moramo le rešiti ta sistem linearnih enačb.

ZGLED 2.25. (1) Naj bodo $u = (1, 3, 0)$, $v = (0, 2, 1)$ in $w = (1, 0, 2)$ vektorji v $\mathbb{R}^3$. Determinanta matrike

$$\begin{bmatrix} u & v & w \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 \\ 3 & 2 & 0 \\ 0 & 1 & 2 \end{bmatrix}$$

je enaka številu 7, zato vektorji u, v, w sestavljajo bazo vektorskega prostora $\mathbb{R}^3$.

Izračunajmo razvoj vektorja $(2, 5, 3) \in \mathbb{R}^3$ po bazi $\mathscr{B} = [u, v, w]$. Rešiti moramo torej sistem linearnih enačb

$$\begin{bmatrix} u & v & w \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} = \begin{bmatrix} 2 \\ 5 \\ 3 \end{bmatrix}$$

za neznane skalarje α, β, γ, in ko to storimo (na primer s Cramerjevim pravilom), dobimo $\alpha = \beta = \gamma = 1$. S tem smo dobili razvoj po bazi

$$(2, 5, 3) = 1u + 1v + 1w$$

oziroma koordinatni vektor $[(2, 5, 3)]_{\mathscr{B}} = (1, 1, 1)$.

(2) V vektorskem prostoru $\mathbb{F}^n$ imamo vektorje:

$$\begin{aligned} \mathrm{e}_1 &= (1, 0, \ldots, 0) \\ \mathrm{e}_2 &= (0, 1, \ldots, 0) \\ &\vdots \\ \mathrm{e}_n &= (0, 0, \ldots, 1) \end{aligned}$$

Pripadajoča matrika $\begin{bmatrix} \mathrm{e}_1 & \mathrm{e}_2 & \cdots & \mathrm{e}_n \end{bmatrix}$ je identična matrika, torej obrnljiva. Odtod sledi, da vektorji $\mathrm{e}_1, \mathrm{e}_2, \ldots, \mathrm{e}_n$ sestavljajo bazo vektorskega prostora $\mathbb{F}^n$, ki jo imenujemo *standardna baza* vektorskega prostora $\mathbb{F}^n$ in označimo

$$\mathscr{E} = \mathscr{E}_n = [\mathrm{e}_1, \mathrm{e}_2, \ldots, \mathrm{e}_n].$$

Vektorjem $\mathrm{e}_1, \mathrm{e}_2, \ldots, \mathrm{e}_n$ pravimo *standardni bazni vektorji* vektorskega prostora $\mathbb{F}^n$.

Vsak vektor $(\alpha_1, \alpha_2, \ldots, \alpha_n) \in \mathbb{F}^n$ lahko torej zapišemo na en sam način kot linearno kombinacijo standardnih baznih vektorjev, in sicer

$$(\alpha_1, \alpha_2, \ldots, \alpha_n) = \alpha_1 \mathrm{e}_1 + \alpha_2 \mathrm{e}_2 + \cdots + \alpha_n \mathrm{e}_n.$$

(3) Vsak polinom $f \in \mathbb{F}[t]$ lahko na en sam način zapišemo v obliki

$$f(t) = \alpha_0 1 + \alpha_1 t + \alpha_2 t^2 + \cdots + \alpha_s t^s,$$

kjer so $\alpha_0, \alpha_1, \alpha_2, \ldots, \alpha_s \in \mathbb{F}$. Odtod sledi, da je podmnožica polinomov

$$\{t^k \ ; \ k \in \mathbb{N} \cup \{0\}\}$$

v vektorskem prostoru $\mathbb{F}[t]$ linearno neodvisna in da polinomi

$$1, t, t^2, \ldots, t^s$$

sestavljajo bazo vektorskega prostora $\mathbb{F}_s[t]$. Bazo $[1, t, t^2, \ldots, t^s]$ imenujemo *standardna baza* vektorskega prostora $\mathbb{F}_s[t]$ polinomov stopnje največ s.

TRDITEV 2.26. *Naj podmnožica $\mathcal{O} \subset \mathsf{V}$ generira vektorski prostor V nad $\mathbb{F}$ in naj bodo $v_1, v_2, \ldots, v_k \in \mathsf{V}$ med seboj linearno neodvisni vektorji v vektorskem prostoru V. Če so za vsak vektor $u \in \mathcal{O} \setminus \{v_1, v_2, \ldots, v_k\}$ vektorji $v_1, v_2, \ldots, v_k, u$ med seboj linearno odvisni v vektorskem prostoru V, potem vektorji $v_1, v_2, \ldots, v_k$ sestavljajo bazo vektorskega prostora V.*

DOKAZ. Po trditvi 2.18 sledi, da je $\mathcal{O} \subset \mathrm{Span}\{v_1, v_2, \ldots, v_k\}$, odtod pa

$$\mathsf{V} = \mathrm{Span}(\mathcal{O}) \subset \mathrm{Span}(\mathrm{Span}\{v_1, v_2, \ldots, v_k\}) = \mathrm{Span}\{v_1, v_2, \ldots, v_k\}. \qquad \square$$

KOMENTAR 2.27. Vektorski prostor lahko ima kakšno končno (urejeno) bazo ali pa ne. Če vektorski prostor $\mathsf{V} \neq \{0\}$ nima končne baze, potem po prejšni trditvi, v kateri vzamemo $\mathcal{O} = \mathsf{V}$, sledi, da ima vektorski prostor V števno neskončno linearno neodvisno podmnožico.

Naslednja trditev nam pove, da vsako končno neprazno podmnožico vektorskega prostora lahko *reduciramo do baze* linearne ogrinjače te podmnožice:

TRDITEV 2.28. *Naj bo V vektorski prostor nad $\mathbb{F}$. Če vektorji $v_1, v_2, \ldots, v_s \in \mathsf{V}$ generirajo vektorski prostor V, potem lahko izberemo takšna naravna števila $1 \leq i_1 < i_2 < \cdots < i_n \leq s$, da vektorji*

$$v_{i_1}, v_{i_2}, \ldots, v_{i_n}$$

sestavljajo bazo vektorskega prostora V.

DOKAZ. Množica $\{v_1, v_2, \ldots, v_s\}$ ima končno mnogo linearno neodvisnih podmnožic in vsaka od teh ima končno elementov. Izberemo torej lahko takšno linearno neodvisno podmnožico

$$\mathcal{O} \subset \{v_1, v_2, \ldots, v_s\},$$

da je vsaka podmnožica množice $\{v_1, v_2, \ldots, v_s\}$, ki ima več elementov kot množica $\mathcal{O}$, linearno odvisna. Po trditvi 2.26 odtod sledi, da je $\mathrm{Span}(\mathcal{O}) = \mathsf{V}$. Naravna števila $1 \le i_1 < i_2 < \cdots < i_n \le k$ izberemo tako, da so vektorji $v_{i_1}, v_{i_2}, \ldots, v_{i_n}$ paroma različni in da je $\mathcal{O} = \{v_{i_1}, v_{i_2}, \ldots, v_{i_n}\}$. □

ZGLED 2.29. Naj bodo $u = (1, 2)$, $v = (1, 1)$ in $w = (3, 3)$ vektorji v $\mathbb{R}^2$. Ker je rang matrike

$$A = \begin{bmatrix} 1 & 1 & 3 \\ 2 & 1 & 3 \end{bmatrix}$$

enak številu 2, vektorji u, v, w generirajo vektorski prostor $\mathbb{R}^2$. Ker je

$$\det(\mathrm{sub}_{\{1,2\}\{1,2\}}(A)) \neq 0,$$

vektorja u, v sestavljata bazo vektorskega prostora $\mathbb{R}^2$. Podoben izračun nam pokaže, da vektorja u, w sestavljata bazo, da pa vektorja v, w ne sestavljata baze vektorskega prostora $\mathbb{R}^2$.

VAJA 2.30. Naj bo n naravno število in naj bodo $v_1, v_2, \ldots, v_{n-1}$ vektorji v $\mathbb{F}^n$. Označimo

$$A = \begin{bmatrix} v_1 & v_2 & \cdots & v_{n-1} & 0 \end{bmatrix} \in \mathrm{Mat}(n \times n, \mathbb{F})$$

in

$$w = \left(\overline{\mathrm{co}_{1n}(A)}, \overline{\mathrm{co}_{2n}(A)}, \ldots, \overline{\mathrm{co}_{nn}(A)}\right) \in \mathbb{F}^n.$$

Pokaži:

(i) Če so vektorji $v_1, v_2, \ldots, v_{n-1}$ linearno neodvisni, potem je $w \neq 0$.
(ii) Če je $w \neq 0$, potem vektorji $v_1, v_2, \ldots, v_{n-1}, w$ sestavljajo bazo vektorskega prostora $\mathbb{F}^n$.
(iii) Za vsak $j = 1, 2, \ldots, n-1$ velja $w^{\mathsf{h}} v_j = 0$.

2.1.3. Linearne preslikave. Strukturi vektorskega prostora, ki je dana s seštevanjem in množenjem s skalarji, včasih pravimo tudi *linearna struktura*. Vektorske prostore lahko med seboj primerjamo s preslikavami, ki ohranjajo njihovo linearno strukturo. Takšnim preslikavam pravimo *linearne preslikave*.

DEFINICIJA 2.31. Naj bosta V in W vektorska prostora nad $\mathbb{F}$. Preslikava $T : \mathsf{V} \to \mathsf{W}$ je *linearna*, če velja

(i) $T(v + v') = T(v) + T(v')$ in
(ii) $T(\alpha v) = \alpha T(v)$

za vse vektorje $v, v' \in \mathsf{V}$ in vsak $\alpha \in \mathbb{F}$. Množico vseh linearnih preslikav iz V v W označimo z

$$\mathrm{Lin}(\mathsf{V}, \mathsf{W}) = \mathrm{Lin}_{\mathbb{F}}(\mathsf{V}, \mathsf{W}).$$

KOMENTAR 2.32. Linearnim preslikavam včasih pravimo tudi *linearne transformacije*. Lastnosti (i) iz definicije pravimo *aditivnost*, lastnosti (ii) pa *homogenost* preslikave T. Za linearno preslikavo $T \in \mathrm{Lin}(\mathsf{V}, \mathsf{W})$ in za poljuben $v \in \mathsf{V}$ pogosto pišemo krajše $Tv = T(v)$. Iz definicije direktno sledi, da je $T(0) = 0$ in da za poljubne vektorje $v_1, v_2, \ldots, v_k \in \mathsf{V}$ in za vse skalarje $\alpha_1, \alpha_2, \ldots, \alpha_k \in \mathbb{F}$ velja

$$T(\alpha_1 v_1 + \alpha_2 v_2 + \cdots \alpha_k v_k) = \alpha_1 T v_1 + \alpha_2 T v_2 + \cdots \alpha_k T v_k.$$

Linearna preslikava T torej preslika linearno kombinacijo vektorjev $v_1, v_2, \ldots, v_k$ v linearno kombinacijo vektorjev $Tv_1, Tv_2, \ldots, Tv_k$.

ZGLED 2.33. (1) Za vsak vektorski prostor V nad $\mathbb{F}$ je identiteta $\mathrm{id} = \mathrm{id}_{\mathsf{V}} : \mathsf{V} \to \mathsf{V}$ linearna preslikava. Za poljubna vektorska prostora V in W nad $\mathbb{F}$ je ničelna preslikava $0 = 0_{\mathsf{W},\mathsf{V}} : \mathsf{V} \to \mathsf{W}$, $v \mapsto 0$, linearna preslikava.

(2) Naj bo dan vektor $u \in \mathbb{R}^n$ s komponentami $u = (\beta_1, \beta_2, \ldots, \beta_n)$. Definirajmo preslikavo $T : \mathbb{R}^n \to \mathbb{R}$ s skalarnim produktom

$$T(v) = u \cdot v = \alpha_1\beta_1 + \alpha_2\beta_2 + \cdots + \alpha_n\beta_n,$$

za vsak $v = (\alpha_1, \alpha_2, \ldots, \alpha_n) \in \mathbb{R}^n$. Iz osnovnih lastnosti skalarnega produkta direktno sledi, da je T linearna preslikava.

(3) Naj bo dan polinom $g \in \mathbb{F}_p[t]$ stopnje največ p. Preslikava $T : \mathbb{F}_s[t] \to \mathbb{F}_{s+p}[t]$, dana s produktom polinomov

$$T(f) = fg$$

za vsak $f \in \mathbb{F}_s[t]$, je linearna.

(4) Naj bo $\mathcal{D}$ odprta podmnožica realnih števil in k naravno število. Preslikava $\frac{d}{dx} : C^k(\mathcal{D}) \to C^{k-1}(\mathcal{D})$ je dana z odvodom

$$\frac{d}{dx}(f) = \frac{df}{dx},$$

za vsako funkcijo $f \in C^k(\mathcal{D})$. Iz osnovnih lastnosti odvajanja sledi, da je preslikava $\frac{d}{dx}$ linearna.

(5) Definirajmo preslikavo

$$\mathrm{tr} : \mathrm{Mat}(n \times n, \mathbb{F}) \to \mathbb{F}$$

s predpisom

$$\mathrm{tr}(A) = A_{11} + A_{22} + \cdots + A_{nn},$$

za vsako matriko $A \in \mathrm{Mat}(n \times n, \mathbb{F})$. Ni se težko prepričati, da je ta preslikava linearna. Vrednost $\mathrm{tr}(A)$ imenujemo *sled* matrike A, in včasih tudi označimo $\mathrm{sled}(A) = \mathrm{tr}(A)$.

Za poljubni matriki $B \in \mathrm{Mat}(m \times n, \mathbb{F})$ in $C \in \mathrm{Mat}(n \times m, \mathbb{F})$ velja

$$\mathrm{tr}(BC) = \mathrm{tr}(CB).$$

Res, velja namreč

$$\mathrm{tr}(BC) = \sum_{i=1}^{m} (BC)_{ii} = \sum_{i=1}^{m} \sum_{j=1}^{n} B_{ij}C_{ji} = \sum_{j=1}^{n} \sum_{i=1}^{m} C_{ji}B_{ij} = \mathrm{tr}(CB).$$

(6) Za poljubno matriko $A \in \mathrm{Mat}(m \times n, \mathbb{F})$ in za poljuben vektor oziroma stolpec $v \in \mathbb{F}^n = \mathrm{Mat}(n \times 1, \mathbb{F})$ je matrični produkt Av spet stolpec oziroma vektor iz $\mathrm{Mat}(m \times 1, \mathbb{F}) = \mathbb{F}^m$. Matrika A nam da torej preslikavo, ki poljubnemu vektorju $v \in \mathbb{F}^n$ priredi matrični produkt $Av \in \mathbb{F}^m$. Iz osnovnih lastnosti množenja matrik sledi, da je tako definirana preslikava $\mathbb{F}^n \to \mathbb{F}^m$ linearna. Opazimo lahko, da za vsak standardni bazni vektor $\mathrm{e}_j \in \mathbb{F}^n$, $j = 1, 2, \ldots, n$, velja $A\mathrm{e}_j = A_{\bullet j}$.

(7) Naj bo $\mathrm{rot}_\varphi : \mathbb{R}^2 \to \mathbb{R}^2$ preslikava, dana z rotacijo v ravnini za kot φ v pozitivni smeri okoli izhodišča. Če poljuben vektor $v = (x, y) \in \mathbb{R}^2$ zapišemo v

polarnem zapisu, dobimo $v = (r\cos\vartheta, r\sin\vartheta)$ in

$$\begin{aligned}\mathrm{rot}_\varphi(v) &= (r\cos(\vartheta+\varphi), r\sin(\vartheta+\varphi))\\ &= r(\cos\vartheta\cos\varphi - \sin\vartheta\sin\varphi, \cos\vartheta\sin\varphi + \sin\vartheta\cos\varphi).\end{aligned}$$

Če to zapišemo v matrični obliki, dobimo

$$\begin{aligned}\mathrm{rot}_\varphi(v) &= r\begin{bmatrix}\cos\vartheta\cos\varphi - \sin\vartheta\sin\varphi\\ \cos\vartheta\sin\varphi + \sin\vartheta\cos\varphi\end{bmatrix}\\ &= \begin{bmatrix}\cos\varphi & -\sin\varphi\\ \sin\varphi & \cos\varphi\end{bmatrix}\begin{bmatrix}r\cos\vartheta\\ r\sin\vartheta\end{bmatrix}\\ &= \begin{bmatrix}\cos\varphi & -\sin\varphi\\ \sin\varphi & \cos\varphi\end{bmatrix}v.\end{aligned}$$

Rotacija rot_φ je torej dana z matriko

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

in je zato linearna po točki (6). Opazimo lahko, da je R_φ ortogonalna matrika z determinanto 1.

(8) Preslikava $\mathrm{Mat}(m\times n,\mathbb{F}) \to \mathrm{Mat}(n\times m,\mathbb{F})$, ki je podana s transponiranjem $A \mapsto A^{\mathsf{t}}$, je linearna.

(9) Naj bo $\mathsf{V}_1\times\mathsf{V}_2\times\cdots\times\mathsf{V}_k$ produkt vektorskih prostorov $\mathsf{V}_1,\mathsf{V}_2,\ldots,\mathsf{V}_k$ nad $\mathbb{F}$. Za vsak $i = 1,2,\ldots,k$ imamo preslikavo $\mathrm{pr}_i : \mathsf{V}_1\times\mathsf{V}_2\times\cdots\times\mathsf{V}_k \to \mathsf{V}_i$, dano s predpisom

$$\mathrm{pr}_i(v_1, v_2, \ldots, v_k) = v_i$$

za vse $(v_1, v_2, \ldots, v_k) \in \mathsf{V}_1\times\mathsf{V}_2\times\cdots\times\mathsf{V}_k$. Ta preslikava je linearna in jo imenujemo *projekcija* na i-to komponento. Imamo tudi preslikavo $\mathrm{vl}_i : \mathsf{V}_i \to \mathsf{V}_1\times\mathsf{V}_2\times\cdots\times\mathsf{V}_k$, ki poljuben vektor $v\in\mathsf{V}_i$ preslika v k-terico

$$\mathrm{vl}_i(v) = (0,\ldots,0,v,0,\ldots,0),$$

katere i-ta komponenta je enaka vektorju v, ostale komponente pa so enake vektorju 0. Tudi ta preslikava, ki jo imenujemo *vložitev* na i-to komponento, je linearna.

Naj bodo dane linearne preslikave $T_i : V_i \to \mathsf{W}_i$ med vektorskimi prostori, za vse $i = 1,2,\ldots,k$. Definirajmo novo preslikavo

$$T_1\times T_2\times\cdots\times T_k : \mathsf{V}_1\times\mathsf{V}_2\times\cdots\times\mathsf{V}_k \to \mathsf{W}_1\times\mathsf{W}_2\times\cdots\times\mathsf{W}_k$$

s predpisom

$$(T_1\times T_2\times\cdots\times T_k)(v_1, v_2, \ldots, v_k) = (T_1v_1, T_2v_2, \ldots, T_kv_k).$$

Tako definirana preslikava $T_1\times T_2\times\cdots\times T_k$ je spet linearna in jo imenujemo *produkt* preslikav $T_1, T_2, \ldots, T_k$.

(10) Preslikava $\mathbb{R}\to\mathbb{R}$ je linearna preslikava vektorskih prostorov nad $\mathbb{R}$, če, in samo če, je dana s predpisom $x\mapsto ax$ za neko konstanto $a\in\mathbb{R}$. Realen polinom stopnje 1 (ki mu običajno pravimo tudi linearna realna funkcija) je torej linearna preslikava vektorskih prostorov nad $\mathbb{R}$ le v primeru, ko je njegova vrednost v točki 0 enaka številu 0.

TRDITEV 2.34. *Naj bo $[v_1, v_2, \ldots, v_n]$ baza vektorskega prostora V nad $\mathbb{F}$. Za poljuben vektorski prostor W nad $\mathbb{F}$ in za poljubne izbrane vektorje $w_1, w_2, \ldots, w_n \in \mathsf{W}$ obstaja natanko ena linearna preslikava $T : \mathsf{V}\to\mathsf{W}$, za katero je $Tv_i = w_i$ za vsak $i = 1,2,\ldots,n$.*

Dokaz. Vzemimo poljuben vektor $v \in \mathsf{V}$ in ga razvijmo po bazi $[v_1, v_2, \ldots, v_n]$,

$$v = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n.$$

Zdaj definiramo

$$T(v) = \alpha_1 w_1 + \alpha_2 w_2 + \cdots + \alpha_n w_n \in \mathsf{W}.$$

S tem smo dobili preslikavo $T : \mathsf{V} \to \mathsf{W}$. Lahko je preveriti, da je preslikava T linearna in da velja $Tv_i = w_i$ za vsak $i = 1, 2, \ldots, n$.

Če je $S \in \mathrm{Lin}(\mathsf{V}, \mathsf{W})$ poljubna linearna preslikava, za katero je $Sv_i = w_i$, potem mora za vsak vektor $v = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n \in \mathsf{V}$ veljati

$$\begin{aligned} Sv &= S(\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n) = \alpha_1 Sv_1 + \alpha_2 Sv_2 + \cdots + \alpha_n Sv_n \\ &= \alpha_1 w_1 + \alpha_2 w_2 + \cdots + \alpha_n w_n = Tv. \end{aligned}$$ □

Naj bosta V in W vektorska prostora nad $\mathbb{F}$. Spomnimo se, da je množica W^{V} vseh funkcij iz $\mathsf{V} \to \mathsf{W}$ vektorski prostor nad $\mathbb{F}$, in sicer za operaciji po točkah (zgled 2.3(4)). Izkaže se, da je množica linearnih preslikav $\mathrm{Lin}(\mathsf{V}, \mathsf{W}) \subset \mathsf{W}^{\mathsf{V}}$ zaprta za seštevanje in za množenje s skalarji, kar pomeni, da je množica $\mathrm{Lin}(\mathsf{V}, \mathsf{W})$ vektorski podprostor vektorskega prostora W^{V}.

Trditev 2.35. *Naj bodo* U, V, W *in* Z *vektorski prostori nad* $\mathbb{F}$.

(i) Za poljubni linearni preslikavi $T, T' \in \mathrm{Lin}(\mathsf{V}, \mathsf{W})$ *in za poljuben skalar* $\beta \in \mathbb{F}$ *je vsota* $T + T' : \mathsf{V} \to \mathsf{W}$, *dana s predpisom*

$$(T + T')(v) = Tv + T'v$$

za vsak $v \in \mathsf{V}$, *linearna preslikava, poleg tega pa je tudi produkt s skalarjem* $\beta T : \mathsf{V} \to \mathsf{W}$, *dan s predpisom*

$$(\beta T)(v) = \beta(Tv)$$

za vsak $v \in \mathsf{V}$, *linearna preslikava. Za tako definirani operaciji je množica* $\mathrm{Lin}(\mathsf{V}, \mathsf{W})$ *vektorski prostor nad* $\mathbb{F}$.

(ii) Če sta $T \in \mathrm{Lin}(\mathsf{V}, \mathsf{W})$ *in* $S \in \mathrm{Lin}(\mathsf{W}, \mathsf{Z})$ *linearni preslikavi, potem je tudi njuna kompozicija*

$$S \circ T : \mathsf{V} \to \mathsf{Z}$$

linearna preslikava. Za kompozicijo linearnih preslikav velja

$$\begin{aligned} &S \circ (T + T') = S \circ T + S \circ T', \\ &(T + T') \circ R = T \circ R + T' \circ R \text{ in} \\ &\beta(S \circ T) = (\beta S) \circ T = S \circ (\beta T) \end{aligned}$$

za vse $R \in \mathrm{Lin}(\mathsf{U}, \mathsf{V})$, $T, T' \in \mathrm{Lin}(\mathsf{V}, \mathsf{W})$ *in* $S \in \mathrm{Lin}(\mathsf{W}, \mathsf{Z})$ *ter za vsak* $\beta \in \mathbb{F}$.

(iii) Če je $T \in \mathrm{Lin}(\mathsf{V}, \mathsf{W})$ *bijektivna linearna preslikava, potem je tudi njen inverz* $T^{-1} : \mathsf{W} \to \mathsf{V}$ *linearna preslikava.*

Dokaz. (i) Za poljubne vektorje $v, v' \in \mathsf{V}$ in za vsak skalar $\alpha \in \mathbb{F}$ velja

$$\begin{aligned} (T + T')(v + v') &= T(v + v') + T'(v + v') = Tv + Tv' + T'v + T'v' \\ &= (T + T')(v) + (T + T')(v'), \\ (T + T')(\alpha v) &= T(\alpha v) + T'(\alpha v) = \alpha(Tv) + \alpha(T'v) = \alpha(T + T')(v), \end{aligned}$$

poleg tega pa tudi

$$\begin{aligned}(\beta T)(v+v') &= \beta(T(v+v')) = \beta(Tv+Tv') \\ &= \beta(Tv)+\beta(Tv') = (\beta T)(v)+(\beta T)(v'), \\ (\beta T)(\alpha v) &= \beta(T(\alpha v)) = \beta\alpha(Tv) = \alpha(\beta T)(v).\end{aligned}$$

(ii) Za poljubne vektorje $v, v' \in \mathsf{V}$ in za vsak skalar $\alpha \in \mathbb{F}$ velja

$$\begin{aligned}(S\circ T)(v+v') &= S(T(v+v')) = S(Tv+Tv') \\ &= S(Tv)+S(Tv') = (S\circ T)(v)+(S\circ T)(v'), \\ (S\circ T)(\alpha v) &= S(T(\alpha v)) = S(\alpha(Tv)) = \alpha S(Tv) = \alpha(S\circ T)(v).\end{aligned}$$

(iii) Za poljubne vektorje $w, w' \in \mathsf{W}$ in za vsak skalar $\alpha \in \mathbb{F}$ velja

$$\begin{aligned}T^{-1}(w+w') &= T^{-1}\big(T(T^{-1}(w))+T(T^{-1}(w'))\big) \\ &= T^{-1}\big(T(T^{-1}(w)+T^{-1}(w'))\big) = T^{-1}(w)+T^{-1}(w'), \\ T^{-1}(\alpha w) &= T^{-1}\big(\alpha T(T^{-1}(w))\big) = T^{-1}\big(T(\alpha T^{-1}(w))\big) = \alpha T^{-1}(w).\end{aligned}$$

□

Komentar 2.36. (1) Kompozicijo $S\circ T$ linearnih preslikav običajno na kratko označimo z ST.

(2) Linearna preslikava $T \in \mathrm{Lin}(\mathsf{V},\mathsf{W})$ je *monomorfizem* vektorskih prostorov, če je injektivna. Linearna preslikava T je *epimorfizem* vektorskih prostorov, če je surjektivna. Linearna preslikava T je *izomorfizem* vektorskih prostorov, če je bijekcija.

Vektorska prostora V in W nad $\mathbb{F}$ sta *izomorfna*, če obstaja kakšen izomorfizem vektorskih prostorov $\mathsf{V}\to\mathsf{W}$. Ni težko preveriti, da je izomorfnost ekvivalenčna relacija med vektorskimi prostori. Tako refleksivnost sledi iz dejstva, da je identična preslikava linearni izomorfizem, simetričnost sledi iz trditve 2.35(iii), tranzitivnost pa iz trditve 2.35(ii). Če sta vektorska prostora V in W izomorfna, potem to zapišemo s simbolom $\mathsf{V}\cong\mathsf{W}$.

(3) *Endomorfizmi* vektorskega prostora V so linearne preslikave $\mathsf{V}\to\mathsf{V}$. Množico takšnih endomorfizmov označimo

$$\mathrm{End}(\mathsf{V}) = \mathrm{End}_{\mathbb{F}}(\mathsf{V}) = \mathrm{Lin}(\mathsf{V},\mathsf{V}).$$

Endomorfizmom vektorskega prostora V pravimo včasih tudi *linearni operatorji* na vektorskem prostoru V. V vektorskem prostoru $\mathrm{End}(\mathsf{V})$ imamo operacijo kompozicije, ki je seveda asociativna in za katero je identiteta $\mathrm{id} = \mathrm{id}_{\mathsf{V}} \in \mathrm{End}(\mathsf{V})$ nevtralni element. Ker poleg tega za kompozicijo veljajo še lastnosti it trditve 2.35(ii), je $\mathrm{End}(\mathsf{V})$ algebra nad $\mathbb{F}$.

Avtomorfizmi vektorskega prostora V so endomorfizni $\mathsf{V}\to\mathsf{V}$, ki so tudi izomorfizmi. Množico vseh avtomorfizmov vektorskega prostora V označimo

$$\mathrm{Aut}(\mathsf{V}) = \mathrm{Aut}_{\mathbb{F}}(\mathsf{V}) = \mathrm{GL}(\mathsf{V}).$$

Množica $\mathrm{Aut}(\mathsf{V})$ je grupa za kompozicijo, ki jo imenujemo *grupa avtomorfizmov* vektorskega prostora V.

Trditev 2.37. *Naj bo $T \in \mathrm{Lin}(\mathsf{V},\mathsf{W})$ linearna preslikava med vektorskima prostoroma nad $\mathbb{F}$.*

(i) Za vsak vektorski podprostor U vektorskega prostora V je njegova slika

$$T(\mathsf{U}) = \{Tv\ ;\ v\in\mathsf{U}\}$$

vektorski podprostor vektorskega prostora W.

(ii) Za vsak vektorski podprostor Z *vektorskega prostora* W *je njegova praslika*

$$T^{-1}(\mathsf{Z}) = \{v \in \mathsf{V}\ ;\ Tv \in \mathsf{Z}\}$$

vektorski podprostor vektorskega prostora V.

DOKAZ. Za poljubna vektorja Tv in Tv' v sliki $T(\mathsf{U})$ je njuna vsota $Tv + Tv' = T(v + v')$ tudi v sliki $T(\mathsf{U})$, zaradi aditivnosti preslikave T. Ostalo dokažemo podobno. □

DEFINICIJA 2.38. Naj bo $T \in \mathrm{Lin}(\mathsf{V}, \mathsf{W})$ linearna preslikava med vektorskima prostoroma nad $\mathbb{F}$. *Jedro* linearne preslikave T je vektorski podprostor

$$\ker(T) = T^{-1}(\{0\}) \subset \mathsf{V}$$

vektorskega prostora V. *Slika* linearne preslikave T je vektorski podprostor

$$\mathrm{im}(T) = T(V) \subset \mathsf{W}$$

vektorskega prostora W.

TRDITEV 2.39. *Linearna preslikava* $T \in \mathrm{Lin}(\mathsf{V}, \mathsf{W})$ *med vektorskima prostoroma nad* $\mathbb{F}$ *je monomorfizem, če, in samo če, je* $\ker(T) = \{0\}$.

DOKAZ. Če je T monomorfizem, je injektivna in zato je $0 \in \mathsf{V}$ edini vektor, ki se preslika v $0 \in \mathsf{W}$. Velja torej $\ker(T) = \{0\}$. Obratno, če je $\ker(T) = \{0\}$ in če za vektorja $v, v' \in \mathsf{V}$ velja $Tv = Tv'$, potem velja $T(v - v') = Tv - Tv' = 0$. To pomeni, da je $v - v' \in \ker(T)$, torej $v - v' = 0$. □

VAJA 2.40. Naj bo $T : \mathsf{V} \to \mathsf{W}$ linearna preslikava med vektorskima prostoroma nad $\mathbb{F}$ in naj bo $w \in \mathsf{W}$ dan vektor. Pokaži, da za vsak vektor $v \in T^{-1}(\{w\})$ velja $T^{-1}(\{w\}) = v + \ker(T) = \{v + u\ ;\ u \in \ker(T)\}$.

TRDITEV 2.41. *Naj bo* $T \in \mathrm{Lin}(\mathsf{V}, \mathsf{W})$ *linearna preslikava med vektorskima prostoroma nad* $\mathbb{F}$.

(i) Če vektorji $v_1, v_2, \ldots, v_s \in \mathsf{V}$ *generirajo vektorski prostor* V*, potem vektorji* $Tv_1, Tv_2, \ldots, Tv_s$ *generirajo vektorski prostor* $\mathrm{im}(T)$.

(ii) Če je T *monomorfizem in če so vektorji* $v_1, v_2, \ldots, v_k \in \mathsf{V}$ *med seboj linearno neodvisni, potem so vektorji* $Tv_1, Tv_2, \ldots, Tv_k \in \mathsf{W}$ *med seboj linearno neodvisni.*

(iii) Če je T *izomorfizem in če vektorji* $v_1, v_2, \ldots, v_n \in \mathsf{V}$ *sestavljajo bazo vektorskega prostora* V*, potem vektorji* $Tv_1, Tv_2, \ldots, Tv_n \in \mathsf{W}$ *sestavljajo bazo vektorskega prostora* W.

(iv) Če obstaja takšna baza $[v_1, v_2, \ldots, v_n]$ *vektorskega prostora* V*, da vektorji* $Tv_1, Tv_2, \ldots, Tv_n \in \mathsf{W}$ *sestavljajo bazo vektorskega prostora* W*, potem je* T *izomorfizem.*

DOKAZ. (i) Poljuben vektor iz slike $\mathrm{im}(T)$ je oblike Tv za nek vektor $v \in \mathsf{V}$. Ker vektorji $v_1, v_2, \ldots, v_s \in \mathsf{V}$ generirajo vektorski prostor V, lahko zapišemo

$$v = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_s v_s$$

za neke skalarje $\alpha_1, \alpha_2, \ldots, \alpha_s \in \mathbb{F}$, odtod pa sledi

$$Tv = \alpha_1 Tv_1 + \alpha_2 Tv_2 + \cdots + \alpha_s Tv_s.$$

(ii) Naj bodo $\alpha_1, \alpha_2, \ldots, \alpha_k \in \mathbb{F}$ takšni skalarji, da je
$$\alpha_1 Tv_1 + \alpha_2 Tv_2 + \cdots + \alpha_k Tv_k = 0.$$
Velja torej
$$T(\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k) = 0,$$
a ker je jedro preslikave T trivialno, odtod dobimo
$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k = 0.$$
Vektorji $v_1, v_2, \ldots, v_k$ so med seboj linearno neodvisni, zato sledi $\alpha_1 = \alpha_2 = \cdots = \alpha_k = 0$.

(iii) Ta del trditve sledi direktno iz točk (i) in (ii).

(iv) Iz točke (i) sledi, da je T epimorfizem. Dokazati moramo še, da je T monomorfizem. Vzemimo poljuben vektor $v \in \ker(T)$ in ga zapišimo kot linearno kombinacijo
$$v = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n,$$
za neke skalarje $\alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbb{F}$. Odtod dobimo
$$\alpha_1 Tv_1 + \alpha_2 Tv_2 + \cdots + \alpha_n Tv_n = Tv = 0,$$
ker pa so vektorji $Tv_1, Tv_2, \ldots, Tv_n \in \mathsf{W}$ med seboj linearno neodvisni, sledi $\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$. Velja torej tudi $v = 0$ in zato je $\ker(T) = \{0\}$. □

Komentar 2.42. Naj bo $T \in \mathrm{Lin}(\mathsf{V}, \mathsf{W})$ linearna preslikava med vektorskima prostoroma in naj bo $\mathscr{B} = [v_1, v_2, \ldots, v_n]$ baza vektorskega prostora V. Če vektorji $Tv_1, Tv_2, \ldots, Tv_n$ sestavljajo bazo vektorskega prostora W, potem pravimo, da linearna preslikava T *preslika bazo* $\mathscr{B}$ *v bazo* $T(\mathscr{B}) = [Tv_1, Tv_2, \ldots, Tv_n]$. Po točkah (iii) in (iv) iz zadnje trditve je to res, če, in samo če, je T izomorfizem.

Zgled 2.43. Naj bo $\mathscr{B} = [v_1, v_2, \ldots, v_n]$ baza vektorskega prostora V nad $\mathbb{F}$. Preslikava $[\,\bullet\,]_{\mathscr{B}} : \mathsf{V} \to \mathbb{F}^n$, ki vsakemu vektorju $v \in \mathsf{V}$ priredi njegov koordinatni vektor $[v]_{\mathscr{B}}$ v bazi $\mathscr{B}$, je linearna in enolično določena s pogojem $[v_i]_{\mathscr{B}} = \mathrm{e}_i$ za vsak $i = 1, 2, \ldots, n$. Ta preslikava torej preslika bazo $\mathscr{B}$ v standardno bazo $\mathscr{E}$ vektorskega prostora $\mathbb{F}^n$, zato je po trditvi 2.41(iv) izomorfizem.

V zgledu 2.33(6) smo videli, da nam poljubna matrika $A \in \mathrm{Mat}(m \times n, \mathbb{F})$ definira linearno preslikavo $\mathbb{F}^n \to \mathbb{F}^m$, ki vektorju $v \in \mathbb{F}^n = \mathrm{Mat}(n \times 1, \mathbb{F})$ priredi matrični produkt $Av \in \mathbb{F}^m = \mathrm{Mat}(m \times 1, \mathbb{F})$. Označimo to linearno preslikavo z
$$\ell(A) : \mathbb{F}^n \to \mathbb{F}^m, \qquad \ell(A)(v) = Av.$$
Izkaže se, da je tako definirana preslikava
$$\ell : \mathrm{Mat}(m \times n, \mathbb{F}) \to \mathrm{Lin}(\mathbb{F}^n, \mathbb{F}^m), \qquad A \mapsto \ell(A),$$
izomorfizem vektorskih prostorov in da produkt matrik preko tega izomorfizma ustreza kompoziciji pripadajočih linearnih preslikav:

Trditev 2.44. *Za vsa naravna števila m, n in p velja:*

(i) preslikava $\ell : \mathrm{Mat}(m \times n, \mathbb{F}) \to \mathrm{Lin}(\mathbb{F}^n, \mathbb{F}^m)$ je izomorfizem vektorskih prostorov nad $\mathbb{F}$,

(ii) za poljubni matriki $A \in \mathrm{Mat}(m \times n, \mathbb{F})$ in $B \in \mathrm{Mat}(n \times p, \mathbb{F})$ je
$$\ell(AB) = \ell(A) \circ \ell(B),$$
(iii) $\ell(\mathrm{I}_{n \times n}) = \mathrm{id}_{\mathbb{F}^n}$, ter

(iv) matrika $C \in \mathrm{Mat}(n \times n, \mathbb{F})$ je obrnljiva, če, in samo če, je preslikava $\ell(C) \in \mathrm{Lin}(\mathbb{F}^n, \mathbb{F}^n)$ izomorfizem, in v tem primeru je

$$\ell(C^{-1}) = (\ell(C))^{-1}.$$

DOKAZ. (i) Iz osnovnih lastnosti množenja matrik direktno sledi, da je preslikava ℓ linearna. Poljubni linearni preslikavi $T : \mathbb{F}^n \to \mathbb{F}^m$ lahko priredimo matriko $\begin{bmatrix} Te_1 & Te_2 & \cdots & Te_n \end{bmatrix} \in \mathrm{Mat}(m \times n, \mathbb{F})$. Lahko je preveriti, da smo s tem definirali inverz preslikave ℓ, kar posebej pomeni, da je ℓ bijekcija.

(ii) Za poljuben vektor $v \in \mathbb{F}^n$ velja

$$\ell(AB)(v) = (AB)v = A(Bv) = \ell(A)(Bv) = \ell(A)(\ell(B)(v)) = (\ell(A) \circ \ell(B))(v).$$

(iii) Za poljuben vektor $v \in \mathbb{F}^n$ je

$$\ell(\mathrm{I}_{n\times n})(v) = \mathrm{I}_{n\times n} v = v = \mathrm{id}_{\mathbb{F}^n}(v).$$

(iv) Če je matrika C obrnljiva, potem velja

$$\mathrm{id}_{\mathbb{F}^n} = \ell(\mathrm{I}_{n\times n}) = \ell(CC^{-1}) = \ell(C) \circ \ell(C^{-1}),$$

prav tako pa tudi

$$\mathrm{id}_{\mathbb{F}^n} = \ell(\mathrm{I}_{n\times n}) = \ell(C^{-1}C) = \ell(C^{-1}) \circ \ell(C),$$

kar pomeni, da je $\ell(C)$ bijekcija z inverzom $\ell(C^{-1})$.

Obratno, pa predpostavimo, da je $\ell(C)$ izomorfizem. Tedaj je inverz $\ell(C)^{-1} : \mathbb{F}^n \to \mathbb{F}^n$ linearna preslikava, ki je po točki (i) dana z natančno eno matriko $D \in \mathrm{Mat}(n \times n, \mathbb{F})$, torej $\ell(D) = \ell(C)^{-1}$. Ker velja

$$\ell(CD) = \ell(C) \circ \ell(D) = \ell(C) \circ \ell(C)^{-1} = \mathrm{id}_{\mathbb{F}^n} = \ell(\mathrm{I}_{n\times n})$$

in ker je preslikava ℓ izomorfizem, odtod sledi $CD = \mathrm{I}_{n\times n}$, zato je matrika C obrnljiva in $C^{-1} = D$. $\square$

Glede na izomorfizem ℓ je torej vsaka linearna preslikava $T : \mathbb{F}^n \to \mathbb{F}^m$ dana z matriko $\begin{bmatrix} Te_1 & Te_2 & \cdots & Te_n \end{bmatrix} \in \mathrm{Mat}(m \times n, \mathbb{F})$, kvadratne matrike velikosti $n \times n$ ustrezajo endomorfizmom vektorskega prostora $\mathbb{F}^n$, po točki (iv) iz trditve pa velja tudi

$$\ell(\mathrm{GL}(n, \mathbb{F})) = \mathrm{Aut}(\mathbb{F}^n) = \mathrm{GL}(\mathbb{F}^n).$$

2.1.4. Dimenzija vektorskega prostora.

DEFINICIJA 2.45. Vektorski prostor V nad $\mathbb{F}$ je *končno-dimenzionalen*, če ima kakšno končno urejeno bazo ali pa če ima le en element. Vektorski prostor je *neskončno-dimenzionalen*, če ni končno-dimenzionalen.

ZGLED 2.46. Vektorski prostor $\mathbb{F}^n$ je končno-dimenzionalen, saj ima standardno bazo $\mathscr{E} = [e_1, e_2, \ldots, e_n]$, ki je seveda končna. Če je vektorski prostor V nad $\mathbb{F}$ izomorfen vektorskemu prostoru $\mathbb{F}^n$, potem je prav tako končno-dimenzionalen: v tem primeru lahko izberemo izomorfizem $T : \mathbb{F}^n \to \mathsf{V}$, po trditvi 2.41(iii) pa vektorji $Te_1, Te_2, \ldots, Te_n$ sestavljajo končno urejeno bazo vektorskega prostora V.

Vektorski prostor je končno-dimenzionalen, če, in samo če, je generiran s kakšno svojo končno podmnožico, saj po trditvi 2.28 vsako končno neprazno podmnožico, ki generira vektorski prostor, lahko reduciramo do baze tega vektorskega prostora.

Trditev 2.47. *Naj bo* V *končno-dimenzionalen vektorski prostor nad* $\mathbb{F}$*. Tedaj obstaja natanko eno nenegativno celo število* n*, da je vektorski prostor* V *izomorfen vektorskemu prostoru* $\mathbb{F}^n$ *nad* $\mathbb{F}$*. Vsaka baza vektorskega prostora* V *je sestavljena iz* n *vektorjev.*

Dokaz. Če ima V le en sam element, potem je $n = 0$ in imamo izomorfizem med V in $\mathbb{F}^n = \{0\}$. Predpostavimo zdaj, da ima V končno urejeno bazo $\mathscr{B} = [v_1, v_2, \ldots, v_n]$, kjer je $n \geq 1$. Tedaj je preslikava $[\,\bullet\,]_{\mathscr{B}} : \mathsf{V} \to \mathbb{F}^n$ izomorfizem, po zgledu 2.43. Preostali del trditve je posledica trditve 2.23. Res, za poljuben izomorfizem $T : \mathsf{V} \to \mathbb{F}^m$ bi vektorji $Tv_1, Tv_2, \ldots, Tv_n$ sestavljali bazo vektorskega prostora $\mathbb{F}^m$, to pa je po trditvi 2.23 mogoče le v primeru, da je $m = n$. $\square$

Komentar 2.48. (1) Če ima vektorski prostor V bazo, ki je sestavljena iz n vektorjev, potem številu n pravimo *dimenzija* vektorskega prostora V. Po zadnji trditvi je dimenzija končno-dimenzionalnega vektorskega prostora enolično določena, saj so vse njegove baze sestavljene iz enakega števila vektorjev. Posebej definiramo, da je dimenzija vektorskega prostora z enim samim elementom enaka številu 0. Dimenzijo končno-dimenzionalnega vektorskega prostora V nad $\mathbb{F}$ označimo z

$$\dim \mathsf{V} = \dim_{\mathbb{F}} \mathsf{V} \in \mathbb{N} \cup \{0\}.$$

Seveda velja $\dim_{\mathbb{F}}(\mathbb{F}^n) = n$. Po zadnji trditvi je vsak končno-dimenzionalen vektorski prostor V nad $\mathbb{F}$ izomorfen vektorskemu prostoru $\mathbb{F}^{\dim \mathsf{V}}$,

$$\mathsf{V} \cong \mathbb{F}^{\dim \mathsf{V}}.$$

Ta izomorfizem je odvisen od izbire baze vektorskega prostora V.

Posebej torej vidimo, da sta dva končno-dimenzionalna vektorska prostora nad $\mathbb{F}$ izomorfna, če, in samo če, imata isto dimenzijo.

(2) Z izbiro baze $\mathscr{B}$ vektorskega prostora V nad $\mathbb{F}$ dobimo izomorfizem $[\,\bullet\,]_{\mathscr{B}} : \mathsf{V} \to \mathbb{F}^n$, preko katerega računanje v vektorji v V prevedemo na računanje s koordinatami teh vektorjev v bazi $\mathscr{B}$. Vektorski prostori pa pogosto nimajo vnaprej izbrane baze, pa tudi če jo imajo, je morda za računanje primernejša kakšna druga baza. Smiselno je torej obvladati računanje v poljubni bazi in poznati zvezo med rezultati izračuna v dveh različnih bazah.

(3) Poljuben izomorfizem $T : \mathbb{F}^n \to \mathsf{V}$ vektorskih prostorov nad $\mathbb{F}$ preslika standardno bazo $\mathscr{E} = [\mathrm{e}_1, \mathrm{e}_2, \ldots, \mathrm{e}_n]$ vektorskega prostora $\mathbb{F}^n$ v bazo

$$\mathscr{B}_T = [T\mathrm{e}_1, T\mathrm{e}_2, \ldots, T\mathrm{e}_n]$$

vektorskega prostora V. Pri tem je torej $T^{-1} = [\,\bullet\,]_{\mathscr{B}_T}$, zato je res tudi obratno, torej da je izomorfizem T določen z izbiro baze $\mathscr{B}_T$. Vidimo torej, da smo s tem dobili naravno bijektivno preslikavo med izomorfizmi vektorskih prostorov $\mathbb{F}^n \to \mathsf{V}$ in bazami vektorskega prostora V:

$$\begin{aligned} \{\text{izomorfizmi } \mathbb{F}^n \to \mathsf{V}\} &\leftrightarrow \{\text{baze vektorskega prostora } \mathsf{V}\} \\ T &\mapsto [T\mathrm{e}_1, T\mathrm{e}_2, \ldots, T\mathrm{e}_n] \end{aligned}$$

V posebnem primeru, ko je $\mathsf{V} = \mathbb{F}^n$, imamo torej bijektivno preslikavo:

$$\begin{aligned} \mathrm{GL}(n, \mathbb{F}) &\leftrightarrow \{\text{baze vektorskega prostora } \mathbb{F}^n\} \\ P &\mapsto [P\mathrm{e}_1, P\mathrm{e}_2, \ldots, P\mathrm{e}_n] \end{aligned}$$

Pri tem lahko opazimo, da so vektorji $P\mathrm{e}_1, P\mathrm{e}_2, \ldots, P\mathrm{e}_n$ ravno stolpci v matriki P, torej $P = \begin{bmatrix} P\mathrm{e}_1 & P\mathrm{e}_2 & \cdots & P\mathrm{e}_n \end{bmatrix}$.

Za poljubno bazo $\mathscr{B} = [v_1, v_2, \ldots, v_n]$ vektorskega prostora V pripadajoči izomorfizem $([\,\bullet\,]_{\mathscr{B}})^{-1} : \mathbb{F}^n \to \mathsf{V}$ preslika standardno bazo $\mathscr{E}$ v bazo $\mathscr{B}$. Včasih označimo enostavno $([\,\bullet\,]_{\mathscr{B}})^{-1} = [v_1, v_2, \ldots, v_n]$. Velja torej $[v_1, v_2, \ldots, v_n]([v]_{\mathscr{B}}) = v$ za vsak $v \in \mathsf{V}$, in posebej $[v_1, v_2, \ldots, v_n](\mathrm{e}_j) = v_j$ za vse $j = 1, 2, \ldots, n$.

(4) Dimenzija vektorskega prostora je odvisna od izbranih skalarjev. Na primer, ker je vsako realno število tudi kompleksno število, je vsak kompleksen vektorski prostor V hkrati tudi realen vektorski prostor, saj množenje s skalarji v V lahko zožimo na realne skalarje. Pri tem pa velja

$$\dim_{\mathbb{R}} \mathsf{V} = 2 \dim_{\mathbb{C}} \mathsf{V}.$$

Če namreč vektorji $v_1, v_2, \ldots, v_n$ sestavljajo bazo kompleksnega vektorskega prostora V, potem vektorji $v_1, \mathrm{i}v_1, v_2, \mathrm{i}v_2, \ldots, v_n, \mathrm{i}v_n$ sestavljajo bazo realnega vektorskega prostora V. Drugačen primer je realen vektorski prostor $\mathbb{R}$, ki ima dimenzijo nad realnimi skalarji enako številu 1, je pa hkrati tudi neskončno-dimenzionalen vektorski prostor nad $\mathbb{Q}$.

ZGLED 2.49. (1) Vektorski prostor $\mathbb{F}_s[t]$ ima bazo, sestavljeno iz polinomov $1, t, t^2, \ldots, t^s$, zato je $\dim \mathbb{F}_s[t] = s + 1$.

(2) Standardna bazna matrika je matrika, ki ima eno od komponent enako številu 1, vse ostale njene komponente pa so enake številu 0. Takih standardnih baznih matrik velikosti $m \times n$ je mn in sestavljajo bazo vektorskega prostora $\mathrm{Mat}(m \times n, \mathbb{F})$. Vektorski prostor $\mathrm{Mat}(m \times n, \mathbb{F})$ je torej končno-dimenzionalen in $\dim \mathrm{Mat}(m \times n, \mathbb{F}) = mn$.

TRDITEV 2.50. *Naj bo V končno-dimenzionalen vektorski prostor nad $\mathbb{F}$ dimenzije n.*

(i) Če vektorji $v_1, v_2, \ldots, v_s \in \mathsf{V}$ generirajo vektorski prostor V, potem je $s \geq n$.

(ii) Če vektorji $v_1, v_2, \ldots, v_n \in \mathsf{V}$ generirajo vektorski prostor V, potem tudi sestavljajo bazo vektorskega prostora V.

(iii) Če so vektorji $v_1, v_2, \ldots, v_k \in \mathsf{V}$ v vektorskem prostoru V med seboj linearno neodvisni, potem je $k \leq n$.

(iv) Če so vektorji $v_1, v_2, \ldots, v_n \in \mathsf{V}$ v vektorskem prostoru V med seboj linearno neodvisni, potem sestavljajo bazo vektorskega prostora V.

DOKAZ. Ker je vektorski prostor V izomorfen vektorskemu prostoru $\mathbb{F}^n$, je trditev dovolj dokazati za primer $\mathsf{V} = \mathbb{F}^n$. V tem primeru pa trditev sledi direktno iz trditev 2.11,2.16 in 2.23. □

POSLEDICA 2.51. *Naj bo V končno-dimenzionalen vektorski prostor nad $\mathbb{F}$ in naj bo U njegov vektorski podprostor. Tedaj velja:*

(i) *vektorski prostor U je končno-dimenzionalen,*
(ii) $\dim \mathsf{U} \leq \dim \mathsf{V}$*, in*
(iii) *če je* $\dim \mathsf{U} = \dim \mathsf{V}$*, potem je* $\mathsf{U} = \mathsf{V}$.

DOKAZ. Če so vektorji $u_1, u_2, \ldots, u_k \in \mathsf{U}$ med seboj linearno neodvisni v vektorskem prostoru U, potem so med seboj linearno neodvisni tudi v vektorskem prostoru V. Odtod sledita točki (i) in (ii). Točka (iii) sledi iz trditve 2.50(iv). □

KOMENTAR 2.52. (1) Naj bo V končno-dimenzionalen vektorski prostor nad $\mathbb{F}$ dimenzije n. Za vsak vektor $v_0 \in \mathsf{V}$ in za vsak vektorski podprostor U dimenzije k vektorskega prostora V označimo

$$v_0 + \mathsf{U} = \{v_0 + u \;;\; u \in \mathsf{U}\} \subset \mathsf{V}.$$

Podmnožicam vektorskega prostora V takšne oblike pravimo *afini podprostori* dimenzije k vektorskega prostora V nad $\mathbb{F}$. Afin podprostor je vektorski podprostor, če, in samo če, vsebuje vektor 0.

Za poljubno linearno preslikavo $T \in \mathrm{Lin}(\mathsf{V}, \mathsf{W})$ med končno-dimenzionalnima vektorskima prostoroma nad $\mathbb{F}$ in za vsak vektor $w \in \mathsf{W}$ je podmnožica $T^{-1}(\{w\})$ afin podprostor vektorskega prostora V. Za poljuben $v \in T^{-1}(\{w\})$ je namreč $T^{-1}(\{w\}) = v + \ker(T)$ (vaja 2.40). Posebej je torej množica rešitev sistema linearnih enačb za n neznank iz $\mathbb{F}$ afin podprostor vektorskega prostora $\mathbb{F}^n$.

(2) Afini podprostori dimenzije 1 vektorskega prostora $\mathbb{R}^3$ so natanko vse premice v $\mathbb{R}^3$. Vektorski podprostori dimenzije 1 vektorskega prostora $\mathbb{R}^3$ so natanko vse premice v $\mathbb{R}^3$, ki gredo skozi izhodišče.

Afini podprostori dimenzije 2 vektorskega prostora $\mathbb{R}^3$ so natanko vse ravnine v $\mathbb{R}^3$. Vektorski podprostori dimenzije 2 vektorskega prostora $\mathbb{R}^3$ so natanko vse ravnine v $\mathbb{R}^3$, ki gredo skozi izhodišče.

(3) Naj bo V realen končno-dimenzionalen vektorski prostor nad $\mathbb{F}$ dimenzije n. Afinim podprostorom dimenzije 1 vektorskega prostora V pravimo tudi *premice* v vektorskem prostoru V, afinim podprostorom dimenzije 2 vektorskega prostora V pa pravimo tudi *ravnine* v vektorskem prostoru V. Afinim podprostorom dimenzije $n-1$ vektorskega prostora V pravimo tudi *hiperravnine* v vektorskem prostoru V.

Ena od posledic trditve 2.50 je tudi ta, da lahko v končno-dimenzionalnem vektorskem prostoru medsebojno linearno neodvisne vektorje vedno *dopolnimo do baze*:

Posledica 2.53. *Naj bo V končno-dimenzionalen vektorski prostor nad $\mathbb{F}$ dimenzije $n \geq 1$ in naj podmnožica $\mathcal{O} \subset \mathsf{V}$ generira vektorski prostor V. Če so vektorji $v_1, v_2, \ldots, v_k \in \mathsf{V}$ v vektorskem prostoru V med seboj linearno neodvisni, potem obstajajo takšni vektorji $v_{k+1}, v_{k+2}, \ldots, v_n \in \mathcal{O}$, da vektorji*

$$v_1, v_2, \ldots, v_k, v_{k+1}, v_{k+2}, \ldots, v_n$$

sestavljajo bazo vektorskega prostora V.

Dokaz. Vektorje $v_{k+1}, v_{k+2}, \ldots, v_n \in \mathcal{O}$ izberemo rekurzivno in sicer tako, da za vsak $j = 1, 2, \ldots, n-k$ izberemo poljuben vektor

$$v_{k+j} \in \mathcal{O} \setminus \mathrm{Span}\{v_1, \ldots, v_k, v_{k+1}, \ldots, v_{k+j-1}\}.$$

To lahko storimo, saj je množica $\mathcal{O} \setminus \mathrm{Span}\{v_1, \ldots, v_k, v_{k+1}, \ldots, v_{k+j-1}\}$ neprazna: če bi namreč veljalo $\mathcal{O} \subset \mathrm{Span}\{v_1, \ldots, v_k, v_{k+1}, \ldots, v_{k+j-1}\}$, bi odtod sledilo

$$\begin{aligned}\mathsf{V} = \mathrm{Span}(\mathcal{O}) &\subset \mathrm{Span}(\mathrm{Span}\{v_1, \ldots, v_k, v_{k+1}, \ldots, v_{k+j-1}\}) \\ &= \mathrm{Span}\{v_1, \ldots, v_k, v_{k+1}, \ldots, v_{k+j-1}\},\end{aligned}$$

kar pa ni mogoče zaradi trditve 2.50(i). Po trditvi 2.18 so tako konstruirani vektorji $v_1, \ldots, v_k, v_{k+1}, \ldots, v_n$ med seboj linearno neodvisni, iz trditve 2.50(iv) pa sledi, da sestavljajo bazo vektorskega prostora V. $\square$

Zgled 2.54. (1) Naj bodo vektorji $v_1, v_2, \ldots, v_k \in \mathbb{F}^n$ med seboj linearno neodvisni. Želimo jih dopolniti do baze prostora $\mathbb{F}^n$. Matrika

$$A = \begin{bmatrix} v_1 & v_2 & \cdots & v_k \end{bmatrix}$$

ima rang enak številu k, zato ima neko neničelno $(k\times k)$-poddeterminanto. Recimo, da so $i_1, i_2, \ldots, i_k$ takšni paroma različni indeksi vrstic, da je

$$\det(\mathrm{sub}_{\{i_1,i_2,\ldots,i_k\}\{1,2,\ldots,k\}}(A)) \neq 0.$$

Zdaj naj bodo $p_1, p_2, \ldots, p_{n-k}$ tista naravna števila, za katera je $\{i_1, i_2, \ldots, i_k\} \cup \{p_1, p_2, \ldots, p_{n-k}\} = \{1, 2, \ldots, n\}$. Ni se težko prepričati, da je tedaj determinanta matrike

$$\begin{bmatrix} v_1 & v_2 & \cdots & v_k & \mathrm{e}_{p_1} & \mathrm{e}_{p_2} & \cdots & \mathrm{e}_{p_{n-k}} \end{bmatrix}$$

neničelna, zato vektorji

$$v_1, v_2, \ldots, v_k, \mathrm{e}_{p_1}, \mathrm{e}_{p_2}, \ldots, \mathrm{e}_{p_{n-k}}$$

sestavljajo bazo vektorskega prostora $\mathbb{F}^n$.

(2) Dana sta vektorja $u = (1, 2, 0)$ in $v = (1, 2, 1)$ v $\mathbb{R}^3$. Za matriko

$$A = \begin{bmatrix} u & v \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 2 & 2 \\ 0 & 1 \end{bmatrix}$$

je $\det(\mathrm{sub}_{\{1,3\}\{1,2\}}(A)) = 1$, zato sta vektorja u in v med seboj linearno neodvisna. Vzemimo $w = \mathrm{e}_2 = (0, 1, 0)$. Po formuli za razvoj determinante po tretjem stolpcu sledi, da je $\det \begin{bmatrix} u & v & w \end{bmatrix} \neq 0$, zato vektorji u, v, w sestavljajo bazo vektorskega prostora $\mathbb{R}^3$.

Posledica 2.55. *Naj bo $T \in \mathrm{Lin}(\mathsf{V}, \mathsf{W})$ linearna preslikava med vektorskima prostoroma nad $\mathbb{F}$, pri čemer je vektorski prostor V končno-dimenzionalen. Naj bo U vektorski podprostor vektorskega prostora V in naj bo Z vektorski podprostor vektorskega prostora W. Tedaj velja:*

(i) *vektorski prostor* $\mathrm{im}(T)$ *je končno-dimenzionalen,*
(ii) $\dim T(\mathsf{U}) \leq \dim \mathsf{U}$,
(iii) *če je T monomorfizem, potem je* $\dim T(\mathsf{U}) = \dim \mathsf{U}$*,*
(iv) *če je T epimorfizem, potem je* $\dim T^{-1}(\mathsf{Z}) \geq \dim \mathsf{Z}$*, in*
(v) *če je T izomorfizem, potem je* $\dim T^{-1}(\mathsf{Z}) = \dim \mathsf{Z}$*.*

Dokaz. Trditev je direktna posledica trditve 2.41 in trditve 2.50. □

Vaja 2.56. Naj bodo $\mathsf{V}_1, \mathsf{V}_2, \ldots, \mathsf{V}_k$ končno-dimenzionalni vektorski prostori nad $\mathbb{F}$. Pokaži, da je tedaj tudi vektorski prostor $\mathsf{V}_1 \times \mathsf{V}_2 \times \cdots \times \mathsf{V}_k$ končno-dimenzionalen in da velja

$$\dim(\mathsf{V}_1 \times \mathsf{V}_2 \times \cdots \times \mathsf{V}_k) = \dim \mathsf{V}_1 + \dim \mathsf{V}_2 + \cdots + \dim \mathsf{V}_k.$$

2.1.5. Presek, vsota in direktna vsota podprostorov. Naj bosta U in W vektorska podprostora vektorskega prostora V nad $\mathbb{F}$. Tedaj je *presek podprostorov*

$$\mathsf{U} \cap \mathsf{W}$$

vektorski podprostor vektorskih prostorov V, U in W. Množica

$$\mathsf{U} + \mathsf{W} = \{u + w \ ; \ u \in \mathsf{U}, \ w \in \mathsf{W}\}$$

je vektorski podprostor prostora V, ki ga imenujemo *vsota podprostorov* U in V.

Definicijo preseka in vsote vektorskih podprostorov lahko posplošimo: Naj bo $(\mathsf{V}_i)_{i\in\mathcal{I}}$ družina vektorskih podprostorov vektorskega prostora V, indeksirana z indeksno množico $\mathcal{I}$. Tedaj je *presek podprostorov*

$$\bigcap_{i\in\mathcal{I}} \mathsf{V}_i$$

spet vektorski podprostor vektorskega prostora V. Množica

$$\sum_{i\in\mathcal{I}} \mathsf{V}_i = \{v_1 + v_2 + \cdots + v_p \ ; \ p \in \mathbb{N}, \ v_1, v_2, \ldots, v_p \in \bigcup_{i\in\mathcal{I}} \mathsf{V}_i\}$$

je vektorski podprostor prostora V, ki ga imenujemo *vsota podprostorov* $(\mathsf{V}_i)_{i\in\mathcal{I}}$. Vsota $\sum_{i\in\mathcal{I}} \mathsf{V}_i$ je najmanjši vektorski podprostor vektorskega prostora V, ki ima vse vektorske podprostore V_i, $i \in \mathcal{I}$, za svoje podmnožice, torej

$$\sum_{i\in\mathcal{I}} \mathsf{V}_i = \mathrm{Span}(\bigcup_{i\in\mathcal{I}} \mathsf{V}_i).$$

Če je $\mathcal{I} = \{1, 2, \ldots, k\}$, potem zapišemo tudi

$$\bigcap_{i\in\mathcal{I}} \mathsf{V}_i = \bigcap_{i=1}^{k} \mathsf{V}_i = \mathsf{V}_1 \cap \mathsf{V}_2 \cap \cdots \cap \mathsf{V}_k$$

in

$$\sum_{i\in\mathcal{I}} \mathsf{V}_i = \sum_{i=1}^{k} \mathsf{V}_i = \mathsf{V}_1 + \mathsf{V}_2 + \cdots + \mathsf{V}_k.$$

Vaja 2.57. Pokaži, da za poljubne vektorje $v_1, v_2, \ldots, v_k \in \mathsf{V}$ iz vektorskega prostora V velja $\mathrm{Span}\{v_1, v_2, \ldots, v_k\} = \mathbb{F}v_1 + \mathbb{F}v_2 + \cdots + \mathbb{F}v_k$, kjer smo označili $\mathbb{F}v_i = \mathrm{Span}\{v_i\}$ za vsak i.

Trditev 2.58. *Za poljubna vektorska podprostora* U *in* W *končno-dimenzionalnega vektorskega prostora* V *nad* $\mathbb{F}$ *velja*

$$\dim \mathsf{U} + \dim \mathsf{W} = \dim(\mathsf{U} + \mathsf{W}) + \dim(\mathsf{U} \cap \mathsf{W}).$$

Dokaz. Izberimo bazo $[v_1, v_2, \ldots, v_p]$ vektorskega podprostora $\mathsf{U} \cap \mathsf{W}$, ki jo nato lahko z vektorji $u_1, u_2, \ldots, u_k \in \mathsf{U}$ dopolnimo do baze vektorskega prostora U, z vektorji $w_1, w_2, \ldots, w_s \in \mathsf{W}$ pa do baze vektorskega prostora W. Vektorji

$$v_1, v_2, \ldots, v_p, u_1, u_2, \ldots, u_k, w_1, w_2, \ldots, w_s$$

zdaj sestavljajo bazo vektorskega prostora $\mathsf{U}+\mathsf{W}$. Res, očitno ti vektorji generirajo vektorski prostor $\mathsf{U} + \mathsf{W}$. Dokazati moramo še, da so tudi med seboj linearno neodvisni. Naj bodo $\alpha_1, \ldots, \alpha_p, \beta_1, \ldots, \beta_k, \gamma_1, \ldots, \gamma_s \in \mathbb{F}$ takšni skalarji, da je

$$\alpha_1 v_1 + \cdots + \alpha_p v_p + \beta_1 u_1 + \cdots + \beta_k u_k + \gamma_1 w_1 + \cdots + \gamma_s w_s = 0.$$

Označimo še $v = \alpha_1 v_1 + \cdots + \alpha_p v_p$, $u = \beta_1 u_1 + \cdots + \beta_k u_k$ in $w = \gamma_1 w_1 + \cdots + \gamma_s w_s$. Tako velja $v + u + w = 0$, in ker je $v + u \in \mathsf{U}$, odtod sledi $w \in \mathsf{U}$. Ker je hkrati $w \in \mathsf{W}$, sledi torej $w \in \mathsf{U} \cap \mathsf{W}$. Na simetričen način pokažemo, da je tudi $u \in \mathsf{U} \cap \mathsf{W}$. Velja torej $v, u, w \in \mathsf{U} \cap \mathsf{W}$.

Vektor u je linearna kombinacija vektorjev $u_1, u_2, \ldots, u_k$, a ker leži v preseku $\mathsf{U} \cap \mathsf{W}$, ga lahko zapišemo tudi kot linearno kombinacijo vektorjev $v_1, v_2, \ldots, v_p$. Ti dve linearni kombinaciji si morata biti enaki, saj vektorji $v_1, v_2, \ldots, v_p, u_1, u_2, \ldots, u_k$ sestavljajo bazo vektorskega prostora U, to pa je mogoče le, če sta obe ti dve linearni kombinaciji trivialni, zato je $\beta_1 = \cdots = \beta_k = 0$ in $u = 0$. Na simetričen način spet lahko vidimo, da velja $\gamma_1 = \cdots = \gamma_s = 0$ in $w = 0$. Odtod sledi tudi $v = 0$, in ker so vektorji $v_1, v_2, \ldots, v_p$ med seboj linearno neodvisni, velja tudi $\alpha_1 = \cdots = \alpha_p = 0$.

Za konec dokaza le preberemo dimenzije,

$$\begin{aligned}\dim \mathsf{U} + \dim \mathsf{W} &= (p+k) + (p+s) = (p+k+s) + p \\ &= \dim(\mathsf{U} + \mathsf{W}) + \dim(\mathsf{U} \cap \mathsf{W}).\end{aligned}$$ □

Naj bosta U in W vektorska podprostora vektorskega prostora V nad $\mathbb{F}$. Če velja $\mathsf{U} + \mathsf{W} = \mathsf{V}$ in hkrati $\mathsf{U} \cap \mathsf{W} = \{0\}$, potem pravimo, da je vektorski prostor V *direktna vsota* svojih vektorskih podprostorov U in W, in to s simbolom zapišemo

$$\mathsf{V} = \mathsf{U} \oplus \mathsf{W}.$$

Če to velja, potem tudi pravimo, da sta si vektorska podprostora U in W *komplementarna* v vektorskem prostoru V. Lahko bi tudi rekli, da je tedaj v vektorskem prostoru V vektorski podprostor U *komplementaren* vektorskemu podprostoru W, ali ekvivalentno, da je vektorski podprostor W komplementaren vektorskemu podprostoru U.

To definicijo lahko nekoliko posplošimo. Naj bodo $\mathsf{V}_1, \mathsf{V}_2, \ldots, \mathsf{V}_k$ vektorski podprostori vektorskega prostora V. Če velja $\mathsf{V}_1 + \mathsf{V}_2 + \cdots + \mathsf{V}_k = \mathsf{V}$ in hkrati $\mathsf{V}_i \cap (\mathsf{V}_1 + \cdots + \mathsf{V}_{i-1} + \mathsf{V}_{i+1} + \cdots + \mathsf{V}_k) = \{0\}$ za vsak $i = 1, 2, \ldots, k$, potem pravimo, da je vektorski prostor V *direktna vsota* svojih vektorskih podprostorov $\mathsf{V}_1, \mathsf{V}_2, \ldots, \mathsf{V}_k$ in to s simbolom zapišemo

$$\mathsf{V} = \mathsf{V}_1 \oplus \mathsf{V}_2 \oplus \cdots \oplus \mathsf{V}_k.$$

Če je $k = 2$, se ta definicija seveda ujema z definicijo direktne vsote dveh podprostorov. V posebnem primeru, ko je $k = 1$, se dogovorimo, da je vsak vektorski prostor V tudi direktna vsota enega samega svojega vektorskega podprostora, namreč podprostora V.

TRDITEV 2.59. *Vektorski prostor V nad $\mathbb{F}$ je direktna vsota svojih vektorskih podprostorov $\mathsf{V}_1, \mathsf{V}_2, \ldots, \mathsf{V}_k$, če, in samo če, za vsak vektor $v \in \mathsf{V}$ obstaja natanko ena urejena k-terica $(v_1, v_2, \ldots, v_k) \in \mathsf{V}_1 \times \mathsf{V}_2 \times \cdots \times \mathsf{V}_k$, za katero je*

$$v = v_1 + v_2 + \cdots + v_k.$$

KOMENTAR 2.60. S predpisom $(v_1, v_2, \ldots, v_k) \mapsto v_1 + v_2 + \cdots + v_k$ je definirana linearna preslikava vektorskih prostorov $\mathsf{V}_1 \times \mathsf{V}_2 \times \cdots \times \mathsf{V}_k \to \mathsf{V}$. Trditev nam pove, da je ta preslikava izomorfizem, če, in samo če, velja $\mathsf{V} = \mathsf{V}_1 \oplus \mathsf{V}_2 \oplus \cdots \oplus \mathsf{V}_k$.

DOKAZ. Naj bo $T : \mathsf{V}_1 \times \mathsf{V}_2 \times \cdots \times \mathsf{V}_k \to \mathsf{V}$ linearna preslikava, dana s predpisom $T(v_1, v_2, \ldots, v_k) = v_1 + v_2 + \cdots + v_k$.

Predpostavimo najprej, da je vektorski prostor V direktna vsota vektorskih podprostorov $\mathsf{V}_1, \mathsf{V}_2, \ldots, \mathsf{V}_k$. Ker je tedaj $\mathsf{V} = \sum_{i=1}^k \mathsf{V}_i$, je preslikava T surjektivna. Če je $(v_1, v_2, \ldots, v_k) \in \ker(T)$, velja $v_1 + v_2 + \cdots + v_k = 0$. Za vsak i je tedaj $v_i = -v_1 - \cdots - v_{i-1} - v_{i+1} - \cdots - v_k$, odtod pa sledi

$$v_i \in \mathsf{V}_i \cap (\mathsf{V}_1 + \cdots + \mathsf{V}_{i-1} + \mathsf{V}_{i+1} + \cdots + \mathsf{V}_k) = \{0\}$$

in torej $v_i = 0$. S tem smo dokazali, da je jedro linearne preslikave T trivialno, zato je T tudi injektivna.

Obratno, zdaj predpostavimo, da je preslikava T izomorfizem. Ker je preslikava T surjektivna, sledi $\mathsf{V} = \sum_{i=1}^k \mathsf{V}_i$. Naj bo

$$v \in \mathsf{V}_i \cap (\mathsf{V}_1 + \cdots + \mathsf{V}_{i-1} + \mathsf{V}_{i+1} + \cdots + \mathsf{V}_k).$$

Vektor v torej lahko zapišemo kot vsoto $v = u_1 + \cdots + u_{i-1} + u_{i+1} + \cdots + u_k$, kjer je $u_j \in \mathsf{V}_j$ za vsak $j \neq i$. Posebej to pomeni, da je

$$T(u_1, \ldots, u_{i-1}, -v, u_{i+1}, \ldots, u_k) = 0,$$

iz injektivnosti preslikave T pa sledi $v = 0$. Dokazali smo torej, da za vsak i velja

$$\mathsf{V}_i \cap (\mathsf{V}_1 + \cdots + \mathsf{V}_{i-1} + \mathsf{V}_{i+1} + \cdots + \mathsf{V}_k) = \{0\}. \qquad \square$$

Vaja 2.61. (1) Pokaži, da neničelni vektorji $v_1, v_2, \ldots, v_n \in \mathsf{V}$ sestavljajo bazo vektorskega prostora V, če, in samo če, je vektorski prostor V direktna vsota svojih vektorskih podprostorov $\mathbb{F}v_1, \mathbb{F}v_2, \ldots, \mathbb{F}v_n$.

(2) Naj bo končno-dimenzionalni vektorski prostor V direktna vsota svojih vektorskih podprostorov $\mathsf{V}_1, \mathsf{V}_2, \ldots, \mathsf{V}_k$. Tedaj velja

$$\dim \mathsf{V} = \dim \mathsf{V}_1 + \dim \mathsf{V}_2 + \cdots + \dim \mathsf{V}_k.$$

Če je $[v_1^{(i)}, v_2^{(i)}, \ldots, v_{n_i}^{(i)}]$ baza vektorskega prostora V_i, za vsak $i = 1, 2, \ldots, k$, potem je

$$[v_1^{(1)}, \ldots, v_{n_1}^{(1)}, v_1^{(2)}, \ldots, v_{n_2}^{(2)}, \ldots, v_1^{(k)}, \ldots, v_{n_k}^{(k)}]$$

baza vektorskega prostora V.

(3) Naj bosta U in W vektorska podprostora vektorskega prostora V. Prepričaj se, da je $\mathsf{V} = \mathsf{U} \oplus \mathsf{W}$, če, in samo če, je $\mathsf{V} = \mathsf{W} \oplus \mathsf{U}$.

(4) Naj bodo $\mathsf{V}_1, \mathsf{V}_2, \mathsf{V}_3$ vektorski podprostori vektorskega prostora V. Dokaži, da je $\mathsf{V} = \mathsf{V}_1 \oplus \mathsf{V}_2 \oplus \mathsf{V}_3$, če, in samo če, velja $\mathsf{V}_1 + \mathsf{V}_2 = \mathsf{V}_1 \oplus \mathsf{V}_2$ in $\mathsf{V} = (\mathsf{V}_1 \oplus \mathsf{V}_2) \oplus \mathsf{V}_3$.

Trditev 2.62. *Naj bo V končno-dimenzionalen vektorski prostor nad $\mathbb{F}$ in U njegov vektorski podprostor. Tedaj obstaja vektorski podprostor W vektorskega prostora V, ki je komplementaren vektorskemu podprostoru U v vektorskem prostoru V, torej $\mathsf{V} = \mathsf{U} \oplus \mathsf{W}$.*

Dokaz. Izberimo bazo $[u_1, u_2, \ldots, u_p]$ vektorskega prostora U, jo dopolnimo do baze vektorskega prostora V z vektorji $w_1, w_2, \ldots, w_{n-p} \in \mathsf{V}$, in vzemimo $\mathsf{W} = \mathrm{Span}\{w_1, w_2, \ldots, w_{n-p}\}$. $\square$

Zgled 2.63. Naj do Σ ravnina v $\mathbb{R}^3$, ki gre izhodišče. Komplementarni vektorski podprostori vektorskega prostora Σ so premice v $\mathbb{R}^3$, ki gredo skozi izhodišče in ki niso vzporedne ravnini Σ.

Trditev 2.64. *Naj bosta V in W vektorska prostora nad $\mathbb{F}$. Predpostavimo, da je vektorski prostor V direktna vsota svojih vektorskih podprostorov $\mathsf{V}_1, \mathsf{V}_2, \ldots, \mathsf{V}_k$, torej $\mathsf{V} = \mathsf{V}_1 \oplus \mathsf{V}_2 \oplus \cdots \oplus \mathsf{V}_k$. Za poljubne linearne preslikave $T_i : \mathsf{V}_i \to \mathsf{W}$, $i = 1, 2, \ldots, k$, tedaj obstaja natanko ena takšna linearna preslikava $T : \mathsf{V} \to \mathsf{W}$, da je $T|_{\mathsf{V}_i} = T_i$ za vsak i. Za poljuben vektor $v \in \mathsf{V}$ je*

$$Tv = T_1 v_1 + T_2 v_2 + \cdots + T_k v_k,$$

kjer je $(v_1, v_2, \ldots, v_k) \in \mathsf{V}_1 \times \mathsf{V}_2 \times \cdots \times \mathsf{V}_k$ tista enolično določena k-terica vektorjev, za katero velja $v = v_1 + v_2 + \cdots + v_k$.

Dokaz. Ni težko preveriti, da je s predpisom iz trditve definirana linearna preslikava z želenimi lastnostmi. Po drugi strani pa je predpis iz trditve posledica linearnosti preslikave T in zahteve, da velja $T|_{\mathsf{V}_i} = T_i$ za vsak i, zato je linearna preslikava, ki zadošča tej zahtevi, ena sama. $\square$

ZGLED 2.65. Naj bosta V in W vektorska prostora nad $\mathbb{F}$. Predpostavimo, da je vektorski prostor V direktna vsota svojih vektorskih podprostorov $\mathsf{V}_1, \mathsf{V}_2, \ldots, \mathsf{V}_k$, torej $\mathsf{V} = \mathsf{V}_1 \oplus \mathsf{V}_2 \oplus \cdots \oplus \mathsf{V}_k$, in da je tudi vektorski prostor W direktna vsota svojih vektorskih podprostorov $\mathsf{W}_1, \mathsf{W}_2, \ldots, \mathsf{W}_k$, torej $\mathsf{W} = \mathsf{W}_1 \oplus \mathsf{W}_2 \oplus \cdots \oplus \mathsf{W}_k$.

Naj bodo $T_i \in \mathrm{Lin}(\mathsf{V}_i, \mathsf{W}_i)$ linearne preslikave, za vse $i = 1, 2, \ldots, k$. Po prejšnji trditvi obstaja natanko ena linearna preslikava $T : \mathsf{V} \to \mathsf{W}$, za katero velja $T|_{\mathsf{V}_i} = T_i : \mathsf{V}_i \to \mathsf{W}_i \subset \mathsf{W}$. To linearno preslikavo označimo tudi

$$T_1 \oplus T_2 \oplus \cdots \oplus T_k : \mathsf{V}_1 \oplus \mathsf{V}_2 \oplus \cdots \oplus \mathsf{V}_k \to \mathsf{W}_1 \oplus \mathsf{W}_2 \oplus \cdots \oplus \mathsf{W}_k$$

in jo imenujemo *direktna vsota* linearnih preslikav $T_1, T_2, \ldots, T_k$. Za poljuben vektor $v \in \mathsf{V}$ velja torej

$$(T_1 \oplus T_2 \oplus \cdots \oplus T_k)v = T_1 v_1 + T_2 v_2 + \cdots + T_k v_k,$$

kjer je $(v_1, v_2, \ldots, v_k) \in \mathsf{V}_1 \times \mathsf{V}_2 \times \cdots \times \mathsf{V}_k$ tista k-terica vektorjev, za katero velja $v = v_1 + v_2 + \cdots + v_k$.

Če je $S \in \mathrm{Lin}(\mathsf{V}, \mathsf{W})$ poljubna linearna preslikava, za katero velja $S(V_i) \subset \mathsf{W}_i$ za vsak i, potem velja

$$S = S_1 \oplus S_2 \oplus \cdots \oplus S_k,$$

kjer je linearna preslikava $S_i \in \mathrm{Lin}(\mathsf{V}_i, \mathsf{W}_i)$ dana z zožitvijo $S|_{\mathsf{V}_i} : \mathsf{V}_i \to \mathsf{W}_i \subset \mathsf{W}$, za vsak i.

2.1.6. Rang in defekt linearne preslikave. V prvem poglavju smo med drugim spoznali pojem ranga matrike in se naučili, kako se rang matrike izračuna. Naslednja trditev nam da geometrijsko interpretacijo ranga matrike:

TRDITEV 2.66. *Za vsako matriko* $A \in \mathrm{Mat}(m \times n, \mathbb{F})$ *je*

$$\mathrm{rank}(A) = \dim(\mathrm{im}(\ell(A))).$$

DOKAZ. Ker standardni bazni vektorji $\mathrm{e}_1, \mathrm{e}_2, \ldots, \mathrm{e}_n$ sestavljajo bazo vektorskega prostora $\mathbb{F}^n$, trditev 2.41(i) pove, da slike $\ell(A)\mathrm{e}_1, \ell(A)\mathrm{e}_2, \ldots, \ell(A)\mathrm{e}_n$ generirajo vektorski prostor $\mathrm{im}(\ell(A))$. Spomnimo se, da je $\ell(A)\mathrm{e}_j = A\mathrm{e}_j = A_{\bullet j}$ za vsak $j = 1, 2, \ldots, n$.

Po trditvi 1.80 ima matrika A vsaj eno neničelno $(r \times r)$-poddeterminanto, $r = \mathrm{rank}(A)$. Obstajajo torej naravna števila $1 \le i_1 < i_2 < \cdots < i_r \le m$ in $1 \le j_1 < j_2 < \cdots < j_r \le n$, da velja

$$\det(\mathrm{sub}_{\{i_1, i_2, \ldots, i_r\}\{j_1, j_2, \ldots, j_r\}}(A)) \ne 0.$$

Po trditvi 2.16 sledi, da so stolpci $A_{\bullet j_1}, A_{\bullet j_2}, \ldots, A_{\bullet j_r}$ med seboj linearno neodvisni. Ker je vsaka $((r+1) \times (r+1))$-poddeterminanta matrike A enaka številu 0, spet po trditvi 2.16 sledi, da je poljubnih $r + 1$ stolpcev matrike A med seboj linearno odvisnih. Po trditvi 2.26 odtod sledi, da stolpci $A_{\bullet j_1}, A_{\bullet j_2}, \ldots, A_{\bullet j_r}$ sestavljajo bazo vektorskega prostora $\mathrm{im}(\ell(A))$. $\square$

Zadnja trditev nam je motivacija za naslednjo definicijo:

DEFINICIJA 2.67. Naj bo $T \in \mathrm{Lin}(\mathsf{V}, \mathsf{W})$ linearna preslikava med vektorskima prostoroma nad $\mathbb{F}$, pri čemer je vektorski prostor V končno-dimenzionalen. *Rang* linearne preslikave T je število

$$\mathrm{rank}(T) = \dim(\mathrm{im}(T)),$$

defekt linearne preslikave T pa je število

$$\mathrm{null}(T) = \dim(\ker(T)).$$

Med rangom in defektom linearne preslikave je tesna zveza:

TRDITEV 2.68. *Naj bo $T \in \mathrm{Lin}(\mathsf{V}, \mathsf{W})$ linearna preslikava med vektorskima prostoroma nad $\mathbb{F}$, pri čemer je vektorski prostor V končno-dimenzionalen. Naj bo U komplementaren vektorski podprostor vektorskega podprostora $\ker(T)$ v vektorskem prostoru V. Tedaj je zožitev*

$$T|_{\mathsf{U}} : \mathsf{U} \to \mathrm{im}(T)$$

izomorfizem vektorskih prostorov in velja

$$\dim \mathsf{V} = \mathrm{rank}(T) + \mathrm{null}(T).$$

DOKAZ. Velja torej $\mathsf{V} = \mathsf{U} \oplus \ker(T)$. To posebej pomeni, da je $\mathsf{U} \cap \ker(T) = \{0\}$, zato je $\ker(T|_{\mathsf{U}}) = \{0\}$. Poljuben vektor iz slike $\mathrm{im}(T)$ je oblike Tv za nek vektor $v \in \mathsf{V}$, ki ga na en sam način lahko zapišemo kot vsoto

$$v = u + v',$$

kjer je $u \in \mathsf{U}$ in $v' \in \ker(T)$. Odtod sledi $Tv = Tu + Tv' = Tu$, zato je $Tv \in \mathrm{im}(T|_{\mathsf{U}})$. S tem smo dokazali, da je $\mathrm{im}(T|_{\mathsf{U}}) = \mathrm{im}(T)$. □

KOMENTAR 2.69. Defekt poljubne matrike $A \in \mathrm{Mat}(m \times n, \mathbb{F})$ definiramo kot

$$\mathrm{null}(A) = n - \mathrm{rank}(A) = \dim(\ker(\ell(A))) = \mathrm{null}(\ell(A)).$$

ZGLED 2.70. (1) Naj bo $AX = 0$ homogen sistem m linearnih enačb za n neznank. Množica rešitev tega sistema je vektorski podprostor $\ker(\ell(A))$ dimenzije $n - \mathrm{rank}(A)$. Če je $\mathrm{rank}(A) = n$, ima sistem eno samo rešitev $X = 0$. Predpostavimo zdaj, da je $\mathrm{rank}(A) < n$. Kot običajno označimo $X = (x_1, x_2, \ldots, x_n)$, in naj bodo $p_1, p_2, \ldots, p_{n-r}$ indeksi natanko tistih stolpcev v matriki A, v katerih A nima pivotov, $r = \mathrm{rank}(A)$. Kot vemo iz izreka 1.46, lahko neznanke

$$x_{p_1}, x_{p_2}, \ldots, x_{p_{n-r}}$$

vzamemo za proste parametre, vrednosti ostalih neznank pa so z izbiro vrednosti teh prostih parametrov enolično določene. Za vsak $j = 1, 2, \ldots, n - r$ imamo torej natanko eno rešitev $X^{(j)}$ sistema $AX = 0$, za katero je

$$\mathrm{sub}_{\{p_1, p_2, \ldots, p_{n-r}\}\{1\}}(X^{(j)}) = \mathrm{e}_j.$$

Vektorji $X^{(1)}, X^{(2)}, \ldots, X^{(n-r)}$ sestavljajo bazo vektorskega prostora rešitev sistema $AX = 0$.

(2) V vektorskem prostoru $\mathbb{R}_2[t]$ realnih polinomov stopnje največ 2 imamo standardno bazo $[1, t, t^2]$ in njej pridruženi izomorfizem vektorskih prostorov $\Psi : \mathbb{R}^3 \to \mathbb{R}_2[t]$, dan s predpisom $\Psi \mathrm{e}_1 = 1$, $\Psi \mathrm{e}_2 = t$ in $\Psi \mathrm{e}_3 = t^2$, oziroma $\Psi(a, b, c) = a + bt + ct^2$.

Naj bo linearna preslikava $T : \mathbb{R}_2[t] \to \mathbb{R}$ dana s predpisom

$$T(f) = f(0) - \frac{df}{dt}(1).$$

Realen polinom $f(t) = a + bt + ct^2 = \Psi(a, b, c)$ je v jedru preslikave T, če, in samo če, velja

$$f(0) - \frac{df}{dt}(1) = a - b - 2c = 0.$$

Dobili smo torej eno homogeno linearno enačbo za tri realne neznanke a, b, c. Če izberemo vrednosti neznank b in c poljubno, je s tem vrednost neznanke a enolično določena. Po zgledu (1) vektorja

$$(1, 1, 0),\ (2, 0, 1)$$

sestavljata bazo vektorskega prostora U vseh rešitev homogene linearne enačbe $a-b-2c = 0$ v vektorskem prostoru $\mathbb{R}^3$. Ker je Ψ izomorfizem, ker je $\Psi(\mathsf{U}) = \ker(T)$ in ker velja $\Psi(1, 1, 0) = 1 + t$ ter $\Psi(2, 0, 1) = 2 + t^2$, odtod sledi, da je

$$[1 + t, 2 + t^2]$$

baza vektorskega prostora $\ker(T)$.

Trditev 2.71. *Naj bo $T \in \mathrm{Lin}(\mathsf{V}, \mathsf{W})$ linearna preslikava med vektorskima prostoroma nad $\mathbb{F}$, pri čemer je vektorski prostor V končno-dimenzionalen, in naj bosta $\Psi \in \mathrm{Lin}(\mathsf{U}, \mathsf{V})$ ter $\Theta \in \mathrm{Lin}(\mathsf{W}, \mathsf{Z})$ izomorfizma vektorskih prostorov. Tedaj velja*

(i) $\mathrm{im}(\Theta T \Psi) = \Theta(\mathrm{im}(T))$,
(ii) $\ker(\Theta T \Psi) = \Psi^{-1}(\ker(T))$,
(iii) $\mathrm{rank}(\Theta T \Psi) = \mathrm{rank}(T)$ *in*
(iv) $\mathrm{null}(\Theta T \Psi) = \mathrm{null}(T)$.

Dokaz. (i) Če je $z \in \mathrm{im}(\Theta T \Psi)$, je $z = \Theta T \Psi u$ za nek vektor $u \in \mathsf{U}$. Odtod sledi, da je $T\Psi u \in \mathrm{im}(T)$ in zato je $z \in \Theta(\mathrm{im}(T))$. Obratno, vsak vektor $z' \in \Theta(\mathrm{im}(T))$ je oblike $z' = \Theta T v$ za nek $v \in \mathsf{V}$, odtod pa sledi, da je $z' \in \Theta T \Psi(\Psi^{-1} v) \in \mathrm{im}(\Theta T \Psi)$.

(ii) Če je $u \in \ker(\Theta T \Psi)$, sledi $T(\Psi u) = \Theta^{-1}(\Theta T \Psi u) = \Theta^{-1} 0 = 0$, zato je res $\Psi u \in \ker(T)$. Obratno, če je $u' \in \Psi^{-1}(\ker(T))$, to pomeni, da je $T(\Psi u') = 0$ in zato $\Theta T \Psi u' = 0$.

Točki (iii) in (iv) sta posledici točk (i) in (ii) ter posledice 2.55. $\square$

Tudi relacijo ekvivalence matrik lahko na naraven način posplošimo na linearne preslikave. Naj bosta $T \in \mathrm{Lin}(\mathsf{V}, \mathsf{W})$ in $S \in \mathrm{Lin}(\mathsf{U}, \mathsf{Z})$ linearni preslikavi med vektorskima prostoroma nad $\mathbb{F}$. Linearna preslikava S je ekvivalentna linearni preslikavi T, če obstajata takšna izomorfizma vektorskih prostorov $\Psi \in \mathrm{Lin}(\mathsf{U}, \mathsf{V})$ in $\Theta \in \mathrm{Lin}(\mathsf{W}, \mathsf{Z})$, da je

$$S = \Theta T \Psi.$$

Ni težko preveriti, da je tako definirana ekvivalenca linearnih preslikav ekvivalenčna relacija.

Trditev 2.72. *Naj bosta $T \in \mathrm{Lin}(\mathsf{V}, \mathsf{W})$ in $S \in \mathrm{Lin}(\mathsf{U}, \mathsf{Z})$ linearni preslikavi med končno-dimenzionalnima vektorskima prostoroma nad $\mathbb{F}$. Linearni preslikavi T in S sta si ekvivalentni, če, in samo če, velja $\dim \mathsf{U} = \dim \mathsf{V}$, $\dim \mathsf{Z} = \dim \mathsf{W}$ in $\mathrm{rank}(T) = \mathrm{rank}(S)$.*

Dokaz. Po trditvi 2.71(iii) imata med seboj ekvivalentni linearni preslikavi isti rang. Dokažimo sedaj še obratno. Predpostavimo, da je $r = \mathrm{rank}(S) = \mathrm{rank}(T)$. Označimo $n = \dim \mathsf{V} = \dim \mathsf{U}$, $m = \dim \mathsf{W} = \dim \mathsf{Z}$ in $d = \mathrm{null}(T)$. Po trditvi 2.68 je tedaj tudi $d = n - r = \mathrm{null}(S)$.

Izberimo bazo $[u_1, u_2, \ldots, u_d]$ jedra $\ker(S)$ in jo dopolnimo do baze vektorskega prostora U z vektorji $u_{d+1}, u_{d+1}, \ldots, u_n \in \mathsf{U}$. Izberimo še bazo $[v_1, v_2, \ldots, v_d]$ jedra $\ker(T)$, in to bazo dopolnimo do baze vektorskega prostora V, denimo z vektorji

$v_{d+1}, v_{d+2}, \ldots, v_n \in \mathsf{V}$. Zdaj definiramo izomorfizem $\Psi \in \mathrm{Lin}(\mathsf{U}, \mathsf{V})$ na bazi s predpisom

$$\Psi u_j = v_j$$

za vse $j = 1, 2, \ldots, n$. Ker smo podali predpis za linearno preslikavo $\Psi : \mathsf{U} \to \mathsf{V}$ na vektorjih iz baze vektorskega prostora V, je linearna preslikava Ψ s tem enolično določena (trdtev 2.34). Ker linearna preslikava Ψ preslika bazo v bazo, je izomorfizem (trditev 2.41(iv)).

Označimo $w_i = Tv_{d+i} \in \mathsf{W}$ in $z_i = Su_{d+i} \in \mathsf{Z}$, za vsa naravna števila $i \leq r = n - d$. Vektorski podprostor $\mathsf{V}_1 = \mathrm{Span}\{v_{d+1}, v_{d+2}, \ldots, v_n\}$ je komplementaren jedru $\ker(T)$ v vektorskem prostoru V, zato je po trditvi 2.68 zožitev $T|_{\mathsf{V}_1} : \mathsf{V}_1 \to \mathrm{im}(T)$ izomorfizem. Odtod sledi, da je $[w_1, w_2, \ldots, w_r]$ baza slike $\mathrm{im}(T)$. To bazo dopolnimo do baze vektorskega prostora W z vektorji $w_{r+1}, w_{r+2}, \ldots, w_m \in \mathsf{W}$. Podobno vidimo, da je $[z_1, z_2, \ldots, z_r]$ baza slike $\mathrm{im}(S)$. To bazo dopolnimo do baze vektorskega prostora Z, recimo z vektorji $z_{r+1}, z_{r+2}, \ldots, z_m \in \mathsf{Z}$. Zdaj definiramo izomorfizem $\Theta \in \mathrm{Lin}(\mathsf{W}, \mathsf{Z})$ na bazi s predpisom

$$\Theta w_i = z_i$$

za vse $i = 1, 2, \ldots, m$. Kot prej linearna preslikava Ψ je tudi linearna preslikava Θ enolično določena z vrednostmi na bazi in je izomorfizem, ker preslika bazo v bazo.

Zdaj je vsak $j = 1, 2, \ldots, d$

$$\Theta T \Psi u_j = \Theta T v_j = \Theta 0 = 0 = S u_j,$$

za $j = d + 1, d + 2, \ldots, n$ pa velja

$$\Theta T \Psi u_j = \Theta T v_j = \Theta w_{j-d} = z_{j-d} = S u_j.$$

Linearni preslikavi $\Theta T \Psi$ in S se torej ujemata na vektorjih iz baze $[u_1, u_2, \ldots, u_n]$ vektorskega prostora U, zato sta si enaki (trditev 2.34). $\square$

2.1.7. Koordinatna matrika linearne preslikave. Naj bo $T \in \mathrm{Lin}(\mathsf{V}, \mathsf{W})$ linearna preslikava med končno-dimenzionalnima vektorskima prostoroma nad $\mathbb{F}$, in naj bo $n = \dim \mathsf{V} \geq 1$ ter $m = \dim \mathsf{W} \geq 1$. Izberimo bazo $\mathscr{B} = [v_1, v_2, \ldots, v_n]$ vektorskega prostora V in bazo $\mathscr{B}' = [w_1, w_2, \ldots, w_m]$ vektorskega prostora W.

Spomnimo se, da nam baza $\mathscr{B}$ določa izomorfizem vektorskih prostorov

$$[\,\bullet\,]_{\mathscr{B}} : \mathsf{V} \to \mathbb{F}^n,$$

ki je enolično določen na bazi s predpisom $[v_j]_{\mathscr{B}} = \mathrm{e}_j$ za vse $j = 1, 2, \ldots, n$. Za poljuben $v \in \mathsf{V}$ je $[v]_{\mathscr{B}}$ koordinatni vektor vektorja v v bazi $\mathscr{B}$, torej stolpec, katerega komponente dobimo z razvojem po bazi

$$v = \big([v]_{\mathscr{B}}\big)_{11} v_1 + \big([v]_{\mathscr{B}}\big)_{21} v_2 + \cdots \big([v]_{\mathscr{B}}\big)_{n1} v_n.$$

Zdaj bomo pokazali, da na podoben način lahko linearni preslikavi T priredimo *koordinatno matriko* v bazah $\mathscr{B}$ in $\mathscr{B}'$.

Za poljuben $j = 1, 2, \ldots, n$ razvijmo vektor Tv_j po bazi $\mathscr{B}'$. Imamo torej

$$Tv_j = a_{1j} w_1 + a_{2j} w_2 + \cdots + a_{mj} w_m$$

za natanko določene skalarje $a_{1j}, a_{2j}, \ldots, a_{mj} \in \mathbb{F}$. *Koordinatna matrika* linearne preslikave T v bazah $\mathscr{B}$ in $\mathscr{B}'$ je matrika

$$[T]^{\mathscr{B}}_{\mathscr{B}'} \in \mathrm{Mat}(m \times n, \mathbb{F})$$

s komponentami

$$\big([T]^{\mathscr{B}}_{\mathscr{B}'}\big)_{ij} = a_{ij}$$

za vse $i = 1, 2, \ldots, m$ in vse $j = 1, 2, \ldots, n$. Stolpci koordinatne matrike $[T]^{\mathscr{B}}_{\mathscr{B}'}$ so torej koordinatni vektorji vektorjev Tv_j v bazi $\mathscr{B}'$,

$$[T]^{\mathscr{B}}_{\mathscr{B}'} = \big[[Tv_1]_{\mathscr{B}'} \quad [Tv_2]_{\mathscr{B}'} \quad \cdots \quad [Tv_n]_{\mathscr{B}'} \big] .$$

TRDITEV 2.73. *Naj bodo* V*,* W *in* Z *končno-dimenzionalni vektorski prostori nad* $\mathbb{F}$ *pozitivnih dimenzij, naj bo* $\mathscr{B}$ *baza vektorskega prostora* V*, naj bo* $\mathscr{B}'$ *baza vektorskega prostora* W *in naj bo* $\mathscr{B}''$ *baza vektorskega prostora* Z*. Za poljubni linearni preslikavi* $T \in \operatorname{Lin}(\mathsf{V}, \mathsf{W})$ *in* $S \in \operatorname{Lin}(\mathsf{W}, \mathsf{Z})$ *velja:*

$$[ST]^{\mathscr{B}}_{\mathscr{B}''} = [S]^{\mathscr{B}'}_{\mathscr{B}''} [T]^{\mathscr{B}}_{\mathscr{B}'}$$

Poleg tega za vsak vektor $v \in \mathsf{V}$ *velja:*

$$[Tv]_{\mathscr{B}'} = [T]^{\mathscr{B}}_{\mathscr{B}'} [v]_{\mathscr{B}}$$

Koordinatna matrika $[\mathrm{id}_{\mathsf{V}}]^{\mathscr{B}}_{\mathscr{B}}$ *identične preslikave* id_{V} *je identična matrika.*

DOKAZ. Označimo $\mathscr{B} = [v_1, v_2, \ldots, v_n]$ in $\mathscr{B}' = [w_1, w_2, \ldots, w_m]$. Komponente koordinatnega vektorja $[Tv]_{\mathscr{B}'}$ vektorja Tv po bazi $\mathscr{B}'$ dobimo z razvojem po bazi

$$Tv = \sum_{i=1}^{m} \big([Tv]_{\mathscr{B}'}\big)_{i1} w_i,$$

po drugi strani pa lahko izračunamo

$$\begin{aligned} Tv &= T\Big(\sum_{j=1}^{n} \big([v]_{\mathscr{B}}\big)_{j1} v_j \Big) = \sum_{j=1}^{n} \big([v]_{\mathscr{B}}\big)_{j1} T v_j \\ &= \sum_{j=1}^{n} \big([v]_{\mathscr{B}}\big)_{j1} \sum_{i=1}^{m} \big([T]^{\mathscr{B}}_{\mathscr{B}'}\big)_{ij} w_i \\ &= \sum_{i=1}^{m} \Big(\sum_{j=1}^{n} \big([T]^{\mathscr{B}}_{\mathscr{B}'}\big)_{ij} \big([v]_{\mathscr{B}}\big)_{j1} \Big) w_i \\ &= \sum_{i=1}^{m} \big([T]^{\mathscr{B}}_{\mathscr{B}'} [v]_{\mathscr{B}}\big)_{i1} w_i. \end{aligned}$$

Odtod sledi, da mora veljati $\big([Tv]_{\mathscr{B}'}\big)_{i1} = \big([T]^{\mathscr{B}}_{\mathscr{B}'}[v]_{\mathscr{B}}\big)_{i1}$ za vsak i. S tem smo dokazali drugo enakost iz trditve.

Prva enakost iz trditve zdaj sledi s trikratno uporabo pravkar dokazanega, saj za vsak $v \in \mathsf{V}$ velja

$$[S]^{\mathscr{B}'}_{\mathscr{B}''} [T]^{\mathscr{B}}_{\mathscr{B}'} [v]_{\mathscr{B}} = [S]^{\mathscr{B}'}_{\mathscr{B}''} [Tv]_{\mathscr{B}'} = [STv]_{\mathscr{B}''} = [ST]^{\mathscr{B}}_{\mathscr{B}''} [v]_{\mathscr{B}}.$$

Ker je preslikava $[\,\bullet\,]_{\mathscr{B}} : \mathsf{V} \to \mathbb{F}^n$ izomorfizem, je vsak vektor iz $\mathbb{F}^n$ oblike $[v]_{\mathscr{B}}$ za natanko določen vektor $v \in \mathsf{V}$. Zadnja enakost torej pomeni, da sta si matriki $[S]^{\mathscr{B}'}_{\mathscr{B}''}[T]^{\mathscr{B}}_{\mathscr{B}'}$ in $[ST]^{\mathscr{B}}_{\mathscr{B}''}$ enaki kot linearni preslikavi, kar pa pomeni, da sta si enaki tudi kot matriki. Enakost $[\mathrm{id}_{\mathsf{V}}]^{\mathscr{B}}_{\mathscr{B}} = \mathrm{I}_{n\times n}$ sledi direktno iz definicije. $\square$

POSLEDICA 2.74. *Naj bosta* V *in* W *končno-dimenzionalna vektorska prostora nad* $\mathbb{F}$*, naj bo* $n = \dim \mathsf{V} \geq 1$ *ter* $m = \dim \mathsf{W} \geq 1$*, naj bo* $\mathscr{B}$ *baza vektorskega prostora* V *in naj bo* $\mathscr{B}'$ *baza vektorskega prostora* W*. Preslikava*

$$[\,\bullet\,]^{\mathscr{B}}_{\mathscr{B}'} : \operatorname{Lin}(\mathsf{V}, \mathsf{W}) \to \operatorname{Mat}(m \times n, \mathbb{F}),$$

ki poljubni linearni preslikavi $T : \mathsf{V} \to \mathsf{W}$ *priredi njeno koordinatno matriko* $[T]^{\mathscr{B}}_{\mathscr{B}'}$ *v bazah* $\mathscr{B}$ *in* $\mathscr{B}'$*, je izomorfizem vektorskih prostorov nad* $\mathbb{F}$*.*

DOKAZ. Direktno po definiciji lahko preverimo, da je preslikava $[\,\bullet\,]^{\mathscr{B}}_{\mathscr{B}'}$ linearna. Iz trditve 2.73 sledi, da za vsako linearno preslikavo $T \in \mathrm{Lin}(\mathsf{V},\mathsf{W})$ velja

$$\ell\big([T]^{\mathscr{B}}_{\mathscr{B}'}\big) \circ [\,\bullet\,]_{\mathscr{B}} = [\,\bullet\,]_{\mathscr{B}'} \circ T$$

oziroma da komutira diagram linearnih preslikav:

$$\begin{tikzcd} \mathsf{V} \arrow[r, "T"] \arrow[d, "{[\,\bullet\,]_{\mathscr{B}}}"'] & \mathsf{W} \arrow[d, "{[\,\bullet\,]_{\mathscr{B}'}}"] \\ \mathbb{F}^n \arrow[r, "{\ell([T]^{\mathscr{B}}_{\mathscr{B}'})}"'] & \mathbb{F}^m \end{tikzcd}$$

Ker sta vertikalni preslikavi v tem diagramu izomorfizma, velja torej

$$\ell\big([T]^{\mathscr{B}}_{\mathscr{B}'}\big) = [\,\bullet\,]_{\mathscr{B}'} \circ T \circ [\,\bullet\,]^{-1}_{\mathscr{B}}.$$

Definiramo pa lahko tudi preslikavo $\mathrm{Mat}(m \times n, \mathbb{F}) \to \mathrm{Lin}(\mathsf{V},\mathsf{W})$ s predpisom

$$A \mapsto [\,\bullet\,]^{-1}_{\mathscr{B}'} \circ \ell(A) \circ [\,\bullet\,]_{\mathscr{B}}.$$

Direktno lahko preverimo, da je s tem podan inverz preslikave $[\,\bullet\,]^{\mathscr{B}}_{\mathscr{B}'}$, kar posebej pomeni, da je preslikava $[\,\bullet\,]^{\mathscr{B}}_{\mathscr{B}'}$ bijekcija. □

POSLEDICA 2.75. *Za poljubna končno-dimenzionalna vektorska prostora* V *in* W *nad* $\mathbb{F}$ *je vektorski prostor* $\mathrm{Lin}_{\mathbb{F}}(\mathsf{V},\mathsf{W})$ *končno-dimenzionalen in velja*

$$\dim_{\mathbb{F}} \mathrm{Lin}_{\mathbb{F}}(\mathsf{V},\mathsf{W}) = (\dim_{\mathbb{F}} \mathsf{V})(\dim_{\mathbb{F}} \mathsf{W}).$$

DOKAZ. Če je $\mathsf{V} = \{0\}$ ali $\mathsf{W} = \{0\}$, je očitno tudi $\mathrm{Lin}(\mathsf{V},\mathsf{W}) = \{0\}$, sicer pa trditev sledi iz posledice 2.74 in dejstva, da je $\dim \mathrm{Mat}(m \times n, \mathbb{F}) = mn$. □

POSLEDICA 2.76. *Naj bosta* V *in* W *končno-dimenzionalna vektorska prostora nad* $\mathbb{F}$ *pozitivnih dimenzij, naj bo* $\mathscr{B}$ *baza vektorskega prostora* V *in naj bo* $\mathscr{B}'$ *baza vektorskega prostora* W. *Tedaj za vsako linearno preslikavo* $T \in \mathrm{Lin}(\mathsf{V},\mathsf{W})$ *velja:*

(i) $\mathrm{im}(\ell([T]^{\mathscr{B}}_{\mathscr{B}'})) = [\mathrm{im}(T)]_{\mathscr{B}'}$
(ii) $\ker(\ell([T]^{\mathscr{B}}_{\mathscr{B}'})) = [\ker(T)]_{\mathscr{B}}$
(iii) $\mathrm{rank}[T]^{\mathscr{B}}_{\mathscr{B}'} = \mathrm{rank}(T)$
(iv) $\mathrm{null}[T]^{\mathscr{B}}_{\mathscr{B}'} = \mathrm{null}(T)$
(v) T *je monomorfizem, če, in samo če, je* $\mathrm{rank}[T]^{\mathscr{B}}_{\mathscr{B}'} = \dim(\mathsf{V})$,
(vi) T *je epimorfizem, če, in samo če, je* $\mathrm{rank}[T]^{\mathscr{B}}_{\mathscr{B}'} = \dim(\mathsf{W})$,
(vii) T *je izomorfizem, če, in samo če, je matrika* $[T]^{\mathscr{B}}_{\mathscr{B}'}$ *obrnljiva, in v tem primeru velja:*

$$\big([T]^{\mathscr{B}}_{\mathscr{B}'}\big)^{-1} = [T^{-1}]^{\mathscr{B}'}_{\mathscr{B}}$$

KOMENTAR 2.77. Oznaka $[\ker(T)]_{\mathscr{B}}$ v točki (ii) iz trditve pomeni sliko jedra $\ker(T)$ z izomorfizmom $[\,\bullet\,]_{\mathscr{B}}$, torej $[\ker(T)]_{\mathscr{B}} = [\,\bullet\,]_{\mathscr{B}}(\ker(T))$. V točki (i) imamo podobno $[\mathrm{im}(T)]_{\mathscr{B}'} = [\,\bullet\,]_{\mathscr{B}'}(\mathrm{im}(T))$.

DOKAZ. Po trditvi 2.73 velja enakost

$$\ell([T]^{\mathscr{B}}_{\mathscr{B}'}) = [\,\bullet\,]_{\mathscr{B}'} \circ T \circ [\,\bullet\,]^{-1}_{\mathscr{B}}.$$

Točke (i-iv) so zato direktne posledice trditve 2.71, v katerih vzamemo $\Theta = [\,\bullet\,]_{\mathscr{B}'}$ in $\Psi = [\,\bullet\,]^{-1}_{\mathscr{B}}$. Točki (v) in (vi) sledita iz točk (iii) in (iv). Ekvivalenca iz točke (vii) je posledica točk (v) in (vi). Za dokaz zadnje enakosti iz točke (vii) izračunamo

$$[T^{-1}]^{\mathscr{B}'}_{\mathscr{B}}[T]^{\mathscr{B}}_{\mathscr{B}'} = [T^{-1}T]^{\mathscr{B}}_{\mathscr{B}} = [\mathrm{id}_{\mathsf{V}}]^{\mathscr{B}}_{\mathscr{B}} = \mathrm{I}.$$ □

Zgled 2.78. (1) Poljuben vektor $v = (\alpha_1, \alpha_2, \ldots, \alpha_n) \in \mathbb{F}^n$ lahko zapišemo v obliki

$$v = \alpha_1 \mathrm{e}_1 + \alpha_2 \mathrm{e}_2 + \cdots + \alpha_n \mathrm{e}_n$$

in s tem smo dobili razvoj vektorja v po standardni bazi $\mathscr{E} = \mathscr{E}_n = [\mathrm{e}_1, \mathrm{e}_2, \ldots, \mathrm{e}_n]$ vektorskega prostora $\mathbb{F}^n$. To pomeni, da je za vsak vektor $v \in \mathbb{F}^n$

$$[v]_{\mathscr{E}} = v.$$

Odtod sledi, da tudi za vsako matriko $A \in \mathrm{Mat}(m \times n, \mathbb{F})$ velja

$$[\ell(A)]_{\mathscr{E}_m}^{\mathscr{E}_n} = \begin{bmatrix} [A\mathrm{e}_1]_{\mathscr{E}_m} & [A\mathrm{e}_2]_{\mathscr{E}_m} & \cdots & [A\mathrm{e}_n]_{\mathscr{E}_m} \end{bmatrix} = \begin{bmatrix} A\mathrm{e}_1 & A\mathrm{e}_2 & \cdots & A\mathrm{e}_n \end{bmatrix} = A,$$

zato je $[\,\bullet\,]_{\mathscr{E}_m}^{\mathscr{E}_n}$ inverz izomorfizma $\ell : \mathrm{Mat}(m \times n, \mathbb{F}) \to \mathrm{Lin}(\mathbb{F}^n, \mathbb{F}^m)$.

Ker je izomorfizem $[\,\bullet\,]_{\mathscr{E}_m}^{\mathscr{E}_n}$ določen s standardnima bazama, preko tega izomorfizma oziroma preko njegovega inverza ℓ pogosto kar identificiramo vektorska prostora $\mathrm{Lin}(\mathbb{F}^n, \mathbb{F}^m)$ in $\mathrm{Mat}(m \times n, \mathbb{F})$, preslikavo $\ell(A)$ pa označimo enostavno z A. Sliki oziroma jedru linearne preslikave $\ell(A)$ zato pravimo enostavno kar slika oziroma jedro matrike A in označimo $\ker(A) = \ker(\ell(A))$ oziroma $\mathrm{im}(A) = \mathrm{im}(\ell(A))$.

(2) Naj bosta $\mathscr{B}$ in $\mathscr{B}'$ dve bazi vektorskega prostora V nad $\mathbb{F}$. Iz trditve 2.73 sledi, da velja

$$[\mathrm{id}_{\mathsf{V}}]_{\mathscr{B}'}^{\mathscr{B}} [v]_{\mathscr{B}} = [v]_{\mathscr{B}'}$$

za vsak vektor $v \in \mathsf{V}$. Z drugimi besedami, preslikava $\ell\big([\mathrm{id}_{\mathsf{V}}]_{\mathscr{B}'}^{\mathscr{B}}\big)$ preslika koordinatni vektor vektorja v v bazi $\mathscr{B}$ v koordinatni vektor vektorja v v bazi $\mathscr{B}'$. Zaradi te lastnosti matriko

$$[\mathrm{id}_{\mathsf{V}}]_{\mathscr{B}'}^{\mathscr{B}}$$

imenujemo *prehodna matrika* z baze $\mathscr{B}$ na bazo $\mathscr{B}'$. Iz trditve 2.73 tudi sledi, da je ta matrika obrnljiva in da je njen inverz prehodna matrika z baze $\mathscr{B}'$ na bazo $\mathscr{B}$:

$$\big([\mathrm{id}_{\mathsf{V}}]_{\mathscr{B}'}^{\mathscr{B}}\big)^{-1} = [\mathrm{id}_{\mathsf{V}}^{-1}]_{\mathscr{B}}^{\mathscr{B}'} = [\mathrm{id}_{\mathsf{V}}]_{\mathscr{B}}^{\mathscr{B}'}$$

Prav tako iz trditve 2.73 sledi, da za poljuben endomorfizem $T \in \mathrm{End}(\mathsf{V})$ velja

$$[T]_{\mathscr{B}'}^{\mathscr{B}'} = [\mathrm{id}_{\mathsf{V}} \circ T \circ \mathrm{id}_{\mathsf{V}}]_{\mathscr{B}'}^{\mathscr{B}'} = [\mathrm{id}_{\mathsf{V}}]_{\mathscr{B}'}^{\mathscr{B}} [T]_{\mathscr{B}}^{\mathscr{B}} [\mathrm{id}_{\mathsf{V}}]_{\mathscr{B}}^{\mathscr{B}'} = [\mathrm{id}_{\mathsf{V}}]_{\mathscr{B}'}^{\mathscr{B}} [T]_{\mathscr{B}}^{\mathscr{B}} \big([\mathrm{id}_{\mathsf{V}}]_{\mathscr{B}'}^{\mathscr{B}}\big)^{-1}.$$

Z enakostjo izomorfizmov $\ell\big([\mathrm{id}_{\mathsf{V}}]_{\mathscr{B}'}^{\mathscr{B}}\big)[\,\bullet\,]_{\mathscr{B}} = [\,\bullet\,]_{\mathscr{B}'}$ je prehodna matrika $[\mathrm{id}_{\mathsf{V}}]_{\mathscr{B}'}^{\mathscr{B}}$ enolično določena. Ekvivalentno lahko to enakost zapišemo v obliki

$$([\,\bullet\,]_{\mathscr{B}})^{-1} \ell\big([\mathrm{id}_{\mathsf{V}}]_{\mathscr{B}}^{\mathscr{B}'}\big) = ([\,\bullet\,]_{\mathscr{B}'})^{-1}.$$

(3) Naj bo $\mathscr{B} = [v_1, v_2, \ldots, v_n]$ baza vektorskega prostora $\mathbb{F}^n$. Če izračunamo prehodno matriko z baze $\mathscr{B}$ na standardno bazo $\mathscr{E}$ vektorskega prostora $\mathbb{F}^n$, dobimo

$$[\mathrm{id}]_{\mathscr{E}}^{\mathscr{B}} = \begin{bmatrix} [v_1]_{\mathscr{E}} & [v_2]_{\mathscr{E}} & \cdots & [v_n]_{\mathscr{E}} \end{bmatrix} = \begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix}.$$

Prehodna matrika $P = [\mathrm{id}]_{\mathscr{E}}^{\mathscr{B}} \in \mathrm{GL}(n, \mathbb{F})$ je torej matrika, katere stolpci so vektorji iz baze $\mathscr{B} = [v_1, v_2, \ldots, v_n]$, torej matrika, za katero je $P\mathrm{e}_i = v_i$ za vsak i. Drugače povedano, baza $\mathscr{B}$ je baza, pridružena obrnljivi matriki P, kot v komentarju 2.48(3).

(4) Na vektorskem prostoru realnih polinomov stopnje največ 3 imamo endomorfizem $\frac{d}{dt} : \mathbb{R}_3[t] \to \mathbb{R}_3[t]$, dan z odvajanjem, torej

$$\frac{d}{dt}(f) = \frac{df}{dt}$$

za vsak $f \in \mathbb{R}_3[t]$. Na vektorskem prostoru $\mathbb{R}_3[t]$ imamo tudi standardno bazo

$$\mathscr{B} = [1, t, t^2, t^3].$$

Ker velja $\frac{d}{dt}(1) = 0$ in $\frac{d}{dt}(t^k) = kt^{k-1}$ za vsak $k = 1, 2, 3$, odtod sledi:

$$\left[\tfrac{d}{dt}\right]_{\mathscr{B}}^{\mathscr{B}} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

(5) Naj bo $\Sigma \subset \mathbb{R}^3$ ravnina, dana z enačbo

$$x + 2y + z = 0,$$

in naj bo $T : \mathbb{R}^3 \to \mathbb{R}^3$ pravokotna projekcija na ravnino Σ, torej preslikava, ki poljubnemu vektorju $v \in \mathbb{R}^3$ priredi tisti vektor iz ravnine Σ, ki je najmanj oddaljen od vektorja v. Drugače povedano, vektor Tv je tisti vektor iz ravnine Σ, za katerega je vektor $v - Tv$ pravokoten na ravnino Σ.

Najprej zapišimo eksplicitno formulo, po kateri lahko izračunamo vektor Tv. Naj bo v_1 normala ravnine Σ, torej $v_1 = (1, 2, 1)$. Pravokotna projekcija vektorja v na normalo v_1 je dana s predpisom

$$\mathrm{pr}_{v_1}^{\perp}(v) = \left(\frac{v \cdot v_1}{v_1 \cdot v_1}\right) v_1 = \tfrac{1}{6}(v \cdot v_1)v_1,$$

zato je vektor Tv dan s predpisom

$$Tv = v - \mathrm{pr}_{v_1}^{\perp}(v) = v - \tfrac{1}{6}\big(v \cdot (1, 2, 1)\big)(1, 2, 1).$$

Iz te formule in iz osnovnih lastnosti skalarnega produkta je jasno, da je preslikava $T : \mathbb{R}^3 \to \mathbb{R}^3$ linearna.

Kot vsaka linearna preslikava $\mathbb{R}^3 \to \mathbb{R}^3$ je tudi preslikava T dana z matriko $[T]_{\mathscr{E}}^{\mathscr{E}}$ glede na standardno bazo $\mathscr{E}$ vektorskega prostora $\mathbb{R}^3$. Da bi to matriko izračunali, moramo izračunati vektorje $T\mathrm{e}_1$, $T\mathrm{e}_2$ in $T\mathrm{e}_3$:

$$\begin{aligned} T\mathrm{e}_1 &= (1, 0, 0,) - \tfrac{1}{6}\big((1, 0, 0,) \cdot (1, 2, 1)\big)(1, 2, 1) = \tfrac{1}{6}(5, -2, -1) \\ T\mathrm{e}_2 &= (0, 1, 0,) - \tfrac{1}{6}\big((0, 1, 0,) \cdot (1, 2, 1)\big)(1, 2, 1) = \tfrac{1}{6}(-2, 2, -2) \\ T\mathrm{e}_3 &= (0, 0, 1,) - \tfrac{1}{6}\big((0, 0, 1,) \cdot (1, 2, 1)\big)(1, 2, 1) = \tfrac{1}{6}(-1, -2, 5) \end{aligned}$$

Odtod dobimo:

$$[T]_{\mathscr{E}}^{\mathscr{E}} = \frac{1}{6}\begin{bmatrix} 5 & -2 & -1 \\ -2 & 2 & -2 \\ -1 & -2 & 5 \end{bmatrix}$$

Matrika $[T]_{\mathscr{E}}^{\mathscr{E}}$ je sicer simetrična, a sam pogled na njene komponente nam ne pove prav dosti o preslikavi, ki jo predstavlja. Razlog za to je ta, da tudi standardna baza nima kakšne jasne geometrijske zveze s preslikavo T. Poskusimo torej izbrati drugačno bazo, ki je bolj naravna glede na geometrijski pomen preslikave T. Naraven kandidat za bazni vektor je seveda normala v_1 ravnine Σ. Vektor $v_1 = (1, 2, 1)$ lahko dopolnimo do baze vektorskega prostora $\mathbb{R}^3$ z vektorji, ki ležijo v ravnini in so torej pravokotni na normalo, na primer $v_2 = (1, 0, -1)$ in $v_3 = (-1, 1, -1)$. Ker je determinanta matrike

$$P = \begin{bmatrix} v_1 & v_2 & v_3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & -1 \\ 2 & 0 & 1 \\ 1 & -1 & -1 \end{bmatrix}$$

enaka številu 6, vektorji v_1, v_2, v_3 res sestavljajo bazo

$$\mathscr{B} = [v_1, v_2, v_3]$$

vektorskega prostora $\mathbb{R}^3$. Matrika P je prehodna matrika iz baze $\mathscr{B}$ na standardno bazo $\mathscr{E}$, torej $P = [\mathrm{id}]_{\mathscr{E}}^{\mathscr{B}}$. Prehodna matrika iz standardne baze $\mathscr{E}$ na bazo $\mathscr{B}$ je inverz matrike P. Če ta inverz izračunamo (na primer po trditvi 1.75), dobimo:

$$[\mathrm{id}]_{\mathscr{B}}^{\mathscr{E}} = P^{-1} = \frac{1}{6}\begin{bmatrix} 1 & 2 & 1 \\ 3 & 0 & -3 \\ -2 & 2 & -2 \end{bmatrix}$$

Zdaj lahko izračunamo tudi koordinatno matriko $[T]_{\mathscr{B}}^{\mathscr{B}}$ po formuli

$$[T]_{\mathscr{B}}^{\mathscr{B}} = [\mathrm{id}]_{\mathscr{B}}^{\mathscr{E}}[T]_{\mathscr{E}}^{\mathscr{E}}[\mathrm{id}]_{\mathscr{E}}^{\mathscr{B}} = P^{-1}[T]_{\mathscr{E}}^{\mathscr{E}}P,$$

še lažje pa jo je izračunati direktno po definiciji: Ker je T pravokotna projekcija na ravnino, preslikava T preslika normalo ravnine v 0, vektorje v ravnini pa same vase, zato je $Tv_1 = 0$, $Tv_2 = v_2$ in $Tv_3 = v_3$ in odtod:

$$[T]_{\mathscr{B}}^{\mathscr{B}} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

2.2. Endomorfizmi vektorskih prostorov

V tem razdelku se bomo osredotočili na študij strukture endomorfizmov vektorskih prostorov. Skozi cel razdelek naj bo obseg skalarjev $\mathbb{F}$ enak $\mathbb{R}$ ali $\mathbb{C}$, in naj bo V poljuben končno-dimenzionalen vektorski prostor nad $\mathbb{F}$ dimenzije $n \geq 1$. Spomnimo se, da so endomorfizmi vektorskega prostora V linearne preslikave $\mathsf{V} \to \mathsf{V}$, torej $\mathrm{End}(\mathsf{V}) = \mathrm{Lin}(\mathsf{V}, \mathsf{V})$. Koordinatne matrike endomorfizmov so kvadratne in za vsako kvadratno matriko $A \in \mathrm{Mat}(n \times n, \mathbb{F})$ je $\ell(A) \in \mathrm{End}(\mathbb{F}^n)$.

2.2.1. Podobnost, determinanta in karakteristični polinom. Koordinatne matrike istega endomorfizma v različnih bazah so med seboj v zvezi, ki smo jo spoznali v zgledu 2.78(2). To nam da motivacija za naslednjo definicijo:

Definicija 2.79. Naj bosta $A, B \in \mathrm{Mat}(n \times n, \mathbb{F})$ kvadratni matriki. Matrika B je *podobna* matriki A, če obstaja takšna obrnljiva matrika $P \in \mathrm{GL}(n, \mathbb{F})$, da velja $B = P^{-1}AP$.

Komentar 2.80. Lahko je preveriti, da je podobnost matrik ekvivalenčna relacija. Za dokaz refleksivnosti namreč vzamemo $P = \mathrm{I}$, za dokaz simetričnosti upoštevamo, da je inverz obrnljive matrike spet obrnljiva matrika, medtem ko tranzitivnost sledi iz dejstva, da je kompozicija obrnljivih matrik spet obrnljiva matrika. Če sta si matriki podobni, sta si očitno tudi ekvivalentni.

Trditev 2.81. *Naj bosta* $A, B \in \mathrm{Mat}(n \times n, \mathbb{F})$ *kvadratni matriki.*

(i) Matrika B je podobna matriki A, če, in samo če, obstaja takšna baza $\mathscr{B}$ vektorskega prostora $\mathbb{F}^n$, da je $[\ell(A)]_{\mathscr{B}}^{\mathscr{B}} = B$.

(ii) Če sta si kvadratni matriki A in B podobni, potem velja $\det(A) = \det(B)$ *in* $\mathrm{tr}(A) = \mathrm{tr}(B)$.

Dokaz. (i) Če je matrika A podobna matriki B, potem obstaja takšna obrnljiva matrika $P \in \mathrm{GL}(n, \mathbb{F})$, da je $B = P^{-1}AP$. Stolpci matrike P sestavljajo bazo $\mathscr{B}$ vektorskega prostora $\mathbb{F}^n$ in velja $P = [\mathrm{id}]_{\mathscr{E}}^{\mathscr{B}}$, kot v zgledu 2.78(3). Pri tem je

$$[\ell(A)]_{\mathscr{B}}^{\mathscr{B}} = [\mathrm{id}]_{\mathscr{B}}^{\mathscr{E}}[\ell(A)]_{\mathscr{E}}^{\mathscr{E}}[\mathrm{id}]_{\mathscr{E}}^{\mathscr{B}} = P^{-1}AP = B.$$

Obratno, predpostavimo zdaj, da je $[\ell(A)]_{\mathscr{B}}^{\mathscr{B}} = B$ za neko bazo $\mathscr{B}$ vektorskega prostora $\mathbb{F}^n$. Tedaj velja

$$B = [\ell(A)]_{\mathscr{B}}^{\mathscr{B}} = [\mathrm{id}]_{\mathscr{B}}^{\mathscr{E}}[\ell(A)]_{\mathscr{E}}^{\mathscr{E}}[\mathrm{id}]_{\mathscr{E}}^{\mathscr{B}} = [\mathrm{id}]_{\mathscr{B}}^{\mathscr{E}} A [\mathrm{id}]_{\mathscr{E}}^{\mathscr{B}},$$

in če vzamemo $P = [\mathrm{id}]_{\mathscr{E}}^{\mathscr{B}}$, imamo torej $B = P^{-1}AP$.

(ii) Če je $B = P^{-1}AP$ za neko obrnljivo matriko $P \in \mathrm{GL}(n, \mathbb{F})$, potem sledi

$$\begin{aligned}\det(B) &= \det(P^{-1}AP) = \det(P^{-1})\det(A)\det(P)\\ &= (\det(P))^{-1}\det(A)\det(P) = \det(A)\end{aligned}$$

in

$$\mathrm{tr}(B) = \mathrm{tr}(P^{-1}AP) = \mathrm{tr}(APP^{-1}) = \mathrm{tr}(A). \qquad \square$$

V splošnem koordinatno matriko linearne preslikave izračunamo glede na dve bazi, bazo domene in bazo kodomene linearne preslikave. Ker sta si domena in kodomena endomorfizma $T \in \mathrm{End}(\mathsf{V}) = \mathrm{Lin}(\mathsf{V}, \mathsf{V})$ enaki, lahko tudi bazi, v katerih izračunamo koordinatno matriko, izberemo enaki, in to največkrat tudi naredimo. Če je $\mathscr{B}$ baza vektorskega prostora V, potem kvadratni koordinatni matriki $[T]_{\mathscr{B}}^{\mathscr{B}}$ pravimo krajše koordinatna matrika endomorfizma T v bazi $\mathscr{B}$, in včasih tudi označimo krajše $[T]_{\mathscr{B}}^{\mathscr{B}} = [T]_{\mathscr{B}}$. Iz zgleda 2.78(2) dobimo:

Trditev 2.82. *Naj bosta $\mathscr{B}$ in $\mathscr{B}'$ dve bazi vektorskega prostora V nad $\mathbb{F}$ in naj bo $T \in \mathrm{End}(\mathsf{V})$ endomorfizem. Tedaj sta si matriki $[T]_{\mathscr{B}}^{\mathscr{B}}$ in $[T]_{\mathscr{B}'}^{\mathscr{B}'}$ podobni.*

Trditev 2.83. *Naj bo $T \in \mathrm{End}(\mathsf{V})$ endomorfizem in naj bo $\Theta \in \mathrm{Lin}(\mathsf{V}, \mathsf{W})$ poljuben izomorfizem vektorskih prostorov nad $\mathbb{F}$. Naj bo $\mathscr{B}$ baza vektorskega prostora V in naj bo $\mathscr{B}'$ baza vektorskega prostora W. Tedaj sta si matriki $[T]_{\mathscr{B}}^{\mathscr{B}}$ in $[\Theta T \Theta^{-1}]_{\mathscr{B}'}^{\mathscr{B}'}$ podobni.*

Dokaz. Matrika $P = [\Theta]_{\mathscr{B}'}^{\mathscr{B}}$ je obrnljiva z inverzom $P^{-1} = [\Theta^{-1}]_{\mathscr{B}}^{\mathscr{B}'}$, ob tem pa velja

$$P^{-1}[\Theta T \Theta^{-1}]_{\mathscr{B}'}^{\mathscr{B}'}P = P^{-1}[\Theta]_{\mathscr{B}'}^{\mathscr{B}}[T]_{\mathscr{B}}^{\mathscr{B}}[\Theta^{-1}]_{\mathscr{B}}^{\mathscr{B}'}P = [T]_{\mathscr{B}}^{\mathscr{B}}. \qquad \square$$

Definicija 2.84. Naj bo $T \in \mathrm{End}(\mathsf{V})$ endomorfizem vektorskega prostora V. *Determinanta* endomorfizma T je determinanta njegove koordinatne matrike v poljubni bazi $\mathscr{B}$ vektorskega prostora V. Determinanto endomorfizma T označimo

$$\det(T) = \det[T]_{\mathscr{B}}^{\mathscr{B}} \in \mathbb{F}.$$

Sled endomorfizma T je sled njegove koordinatne matrike v poljubni bazi $\mathscr{B}$ vektorskega prostora V. Sled endomorfizma T označimo

$$\mathrm{tr}(T) = \mathrm{tr}[T]_{\mathscr{B}}^{\mathscr{B}} \in \mathbb{F}.$$

Komentar 2.85. (1) Po trditvi 2.81 in po trditvi 2.82 je definicija smiselna, saj sta skalarja $\det[T]_{\mathscr{B}}^{\mathscr{B}}$ in $\mathrm{tr}[T]_{\mathscr{B}}^{\mathscr{B}}$ neodvisna od izbire baze $\mathscr{B}$.

(2) Za poljubno kvadratno matriko $A \in \mathrm{Mat}(n \times n, \mathbb{F})$ seveda velja $\det(\ell(A)) = \det(A)$ in $\mathrm{tr}(\ell(A)) = \mathrm{tr}(A)$, saj je $A = [\ell(A)]_{\mathscr{E}}^{\mathscr{E}}$.

(3) Za poljubna endomorfizma $T, S \in \mathrm{End}(\mathsf{V})$ in za vsak skalar $\alpha \in \mathbb{F}$ velja

$$\det(ST) = \det(S)\det(T)$$

in

$$\det(\alpha T) = \alpha^n \det(T),$$

kjer je $n = \dim \mathsf{V}$. Res, za poljubno bazo $\mathscr{B}$ vektorskega prostora V je

$$\begin{aligned}\det(ST) &= \det[ST]^{\mathscr{B}}_{\mathscr{B}} = \det\left([S]^{\mathscr{B}}_{\mathscr{B}}[T]^{\mathscr{B}}_{\mathscr{B}}\right)\\ &= \det[S]^{\mathscr{B}}_{\mathscr{B}}\det[T]^{\mathscr{B}}_{\mathscr{B}}\\ &= \det(S)\det(T)\end{aligned}$$

in

$$\det(\alpha T) = \det[\alpha T]^{\mathscr{B}}_{\mathscr{B}} = \det(\alpha[T]^{\mathscr{B}}_{\mathscr{B}}) = \alpha^n \det[T]^{\mathscr{B}}_{\mathscr{B}} = \alpha^n \det(T).$$

Na podoben način lahko pokažemo, da je $\operatorname{tr} : \operatorname{End}(\mathsf{V}) \to \mathbb{F}$ linearna preslikava.

POSLEDICA 2.86. *Za vsak endomorfizem $T \in \operatorname{End}(\mathsf{V})$ so si ekvivalentne naslednje trditve:*

(i) *T je avtomorfizem,*
(ii) $\operatorname{rank}(T) = \dim \mathsf{V}$*,*
(iii) $\operatorname{null}(T) = 0$*,*
(iv) $\det(T) \neq 0$*.*

DOKAZ. Po posledici 2.76(vii) je endomorfizem T izomorfizem, če, in samo če, je pripadajoča koordinatna matrika v poljubni bazi obrnljiva, oziroma, če, in samo če, je determinanta koordinatne matrike neničelna. Preostale ekvivalence iz trditve sledijo iz enakosti $\operatorname{rank}(T) + \operatorname{null}(T) = \dim \mathsf{V}$ (trditev 2.68). □

DEFINICIJA 2.87. Naj bo $T \in \operatorname{End}(\mathsf{V})$ endomorfizem vektorskega prostora V. *Karakteristični polinom* endomorfizma T je polinom p_T s koeficienti iz $\mathbb{F}$, dan s predpisom

$$\mathrm{p}_T(t) = \det(T - t\,\mathrm{id}_{\mathsf{V}})$$

za vsak $t \in \mathbb{F}$.

KOMENTAR 2.88. (1) Karakteristični polinom p_A poljubne kvadratne matrike $A \in \operatorname{Mat}(n \times n, \mathbb{F})$ definiramo kot karakteristični polinom endomorfizma $\ell(A)$:

$$\begin{aligned}\mathrm{p}_A(t) &= \mathrm{p}_{\ell(A)}(t) = \det(\ell(A) - t\,\mathrm{id}_{\mathsf{V}}) = \det(\ell(A - t\mathrm{I}))\\ &= \det(A - t\mathrm{I}) = \begin{vmatrix} (A_{11} - t) & A_{12} & \cdots & A_{1n} \\ A_{21} & (A_{22} - t) & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & (A_{nn} - t) \end{vmatrix}\end{aligned}$$

Odtod lahko po definiciji determinante matrike razberemo, da je $\mathrm{p}_A(t)$ v resnici polinom s koeficienti v $\mathbb{F}$ stopnje natanko n. Čelni koeficient tega polinoma pri potenci t^n je enak številu $(-1)^n$, koeficient pri potenci t^{n-1} je $(-1)^{n-1}\operatorname{tr}(A)$, vrednost tega polinoma v točki $t = 0$ pa je $\det(A)$.

Za poljuben endomorfizem $T \in \operatorname{End}(\mathsf{V})$ determinanto endomorfizma $T - t\,\mathrm{id}_{\mathsf{V}}$ izračunamo v poljubni bazi $\mathscr{B}$ vektorskega prostora V, zato enako velja tudi za karakteristični polinom, torej

$$\mathrm{p}_T(t) = \det[T - t\,\mathrm{id}_{\mathsf{V}}]^{\mathscr{B}}_{\mathscr{B}} = \det([T]^{\mathscr{B}}_{\mathscr{B}} - t[\mathrm{id}_{\mathsf{V}}]^{\mathscr{B}}_{\mathscr{B}}) = \det([T]^{\mathscr{B}}_{\mathscr{B}} - t\mathrm{I}) = \mathrm{p}_{[T]^{\mathscr{B}}_{\mathscr{B}}}(t).$$

Odtod sledi, da je p_T polinom stopnje $n = \dim \mathsf{V}$ s koeficienti v $\mathbb{F}$ in da je oblike

$$\mathrm{p}_T(t) = (-t)^n + \operatorname{tr}(T)(-t)^{n-1} + \cdots + \det(T).$$

(2) Naj bosta $A, B \in \mathrm{Mat}(n \times n, \mathbb{F})$ kvadratni matriki. Če je matrika B podobna matriki A, potem velja $\mathrm{p}_B = \mathrm{p}_A$. Res, v tem primeru obstaja takšna obrnljiva matrika $P \in \mathrm{GL}(n, \mathbb{F})$, da je $B = P^{-1}AP$, zato je

$$B - t\mathrm{I} = P^{-1}AP - t\mathrm{I} = P^{-1}AP - tP^{-1}\mathrm{I}P = P^{-1}(A - t\mathrm{I})P,$$

kar pomeni da je tedaj matrika $B - t\mathrm{I}$ podobna matriki $A - t\mathrm{I}$. Podobne si matrike pa imajo isto determinanto, zato sledi $\mathrm{p}_B(t) = \det(B - t\mathrm{I}) = \det(A - t\mathrm{I}) = p_A(t)$.

(3) Spomnimo se nekaj osnovnih lastnosti polinomov. Naj bo f poljuben kompleksen polinom stopnje $s \geq 1$ v spremenljivki t in naj bo λ kompleksno število. Polinom f lahko brez ostanka delimo s polinomom $(t - \lambda)$, če, in samo če, je število λ ničla polinoma f. Najdemo lahko takšno nenegativno celo število $k \leq s$, da lahko polinom f brez ostanka delimo s polinomom $(t-\lambda)^k$, ne pa tudi s polinomom $(t - \lambda)^{k+1}$. Drugače povedano, polinom f lahko zapišemo kot produkt

$$f(t) = (t - \lambda)^k g(t),$$

kjer je g tak polinom stopnje $s - k$, da je $g(\lambda) \neq 0$. Število k imenujemo *kratnost* števila λ kot ničle polinoma f, in ga bomo označili $\mathrm{kr}_f(\lambda)$. Če je $\mathrm{kr}_f(\lambda) = 0$, potem λ ni ničla polinoma f. Če velja $\mathrm{kr}_f(\lambda) = 1$, potem je λ *enostavna ničla*, če pa je $\mathrm{kr}_f(\lambda) > 1$, potem je λ *večkratna ničla* polinoma f.

Vsak kompleksen polinom stopnje $s \geq 1$ ima največ s kompleksnih ničel, po osnovnem izreku algebre pa ima vsaj eno kompleksno ničlo. Če je kompleksen polinom f oblike

$$f(t) = a_s t^s + a_{s-1}t^{s-1} + \cdots + a_1 t + a_0$$

za neka kompleksna števila $a_0, a_1, \ldots, a_s \in \mathbb{C}$, kjer je $a_s \neq 0$, potem ga do vrstnega reda faktorjev natančno lahko na en sam način zapišemo v razcepljeni obliki kot produkt linearnih členov

$$f(t) = a_s(t - \lambda_1)^{k_1}(t - \lambda_2)^{k_2} \cdots (t - \lambda_m)^{k_m},$$

kjer so $\lambda_1, \lambda_2, \ldots, \lambda_m$ vse paroma različne kompleksne ničle polinoma f in kjer je $k_i = \mathrm{kr}_f(\lambda_i)$ za vsak i. Posebej je torej $\sum_{\lambda \in \mathbb{C}} \mathrm{kr}_f(\lambda) = s$.

Poljuben realen polinom lahko gledamo tudi kot kompleksen polinom. Če je f realen polinom, potem ima lahko ničlo λ, ki ni realna, vendar takšna nastopa v paru s konjugirano ničlo $\bar{\lambda}$ tako, da je $\mathrm{kr}_f(\lambda) = \mathrm{kr}_f(\bar{\lambda})$. Produkt linearnih faktorjev $(t - \lambda)(t - \bar{\lambda})$ je v tem primeru nerazcepen realen kvadratni polinom z negativno diskriminanto. Realen polinom f zato lahko zapišemo v obliki

$$f(t) = a_s(t - \eta_1)^{k_1} \cdots (t - \eta_p)^{k_p}(t^2 + b_1 t + c_1)^{r_1} \cdots (t^2 + b_q t + c_q)^{r_q},$$

kjer so $\eta_1, \ldots, \eta_p$ vse paroma različne realne ničle polinoma f, kjer so kvadratni faktorji $(t^2 + b_1 t + c_1), \ldots, (t^2 + b_q t + c_q)$ paroma različni realni polinomi z negativno diskriminanto, kjer je $k_i = \mathrm{kr}_f(\eta_i)$ za vsak i in kjer so $r_1, \ldots, r_q$ naravna števila. Realen polinom f je razcepen na realne linearne faktorje, ko so vse njegove kompleksne ničle realne, torej tedaj, ko je $q = 0$.

2.2.2. Lastne vrednosti in lastni vektorji. Naj bo $T \in \mathrm{End}(\mathsf{V})$ endomorfizem. Vektorski podprostor $\mathsf{U} \subset \mathsf{V}$ je *T-invarianten*, če velja $T(\mathsf{U}) \subset \mathsf{U}$. Natančno poznavanje T-invariantnih vektorskih podprostorov nam praviloma pomeni dobro geometrijsko razumevanje endomorfizma T. Še posebej zanimivi in enostavni so seveda enodimenzionalni T-invariantni vektorski podprostori. Če je v poljuben neničelen vektor iz takšnega enodimenzionalnega T-invariantnega vektorskega podprostora U, potem mora biti slika Tv spet element vektorskega podprostora U in zato skalarni večkratnih vektorja v. To nam da motivacijo za naslednjo definicijo:

DEFINICIJA 2.89. Naj bo $T \in \mathrm{End}(\mathsf{V})$ endomorfizem. Skalar $\lambda \in \mathbb{F}$ je *lastna vrednost* endomorfizma T, če obstaja takšen neničelen vektor $v \in \mathsf{V}$, da je

$$Tv = \lambda v.$$

TRDITEV 2.90. *Za vsak endomorfizem* $T \in \mathrm{End}(\mathsf{V})$ *in vsak skalar* $\lambda \in \mathbb{F}$ *so si naslednje trditve ekvivalentne:*

(i) λ *je lastna vrednost endomorfizma* T*,*
(ii) $\ker(T - \lambda\,\mathrm{id}) \neq \{0\}$,
(iii) $(T - \lambda\,\mathrm{id})$ *ni avtomorfizem,*
(iv) $\mathrm{p}_T(\lambda) = 0$.

DOKAZ. Vektor $v \in \mathsf{V}$ reši enačbo $Tv = \lambda v$, če, in samo če, je $(T - \lambda\,\mathrm{id})v = 0$, oziroma, če, in samo če, je $v \in \ker(T - \lambda\,\mathrm{id})$. Ob tem je $\ker(T - \lambda\,\mathrm{id}) = \{0\}$, če, in samo če, je $(T-\lambda\,\mathrm{id})$ avtomorfizem, oziroma, če, in samo če, je $\det(T-\lambda\,\mathrm{id}) \neq 0$. □

KOMENTAR 2.91. Ker ima karakteristični polinom stopnjo $n = \dim \mathsf{V}$ in zato največ n ničel, ima endomorfizem $T \in \mathrm{End}(\mathsf{V})$ največ n lastnih vrednosti. Kratnost poljubnega skalarja $\lambda \in \mathbb{F}$ kot ničle karakterističnega polinoma endomorfizma T imenujemo *algebraična kratnost* števila λ za endomorfizem T, in označimo

$$\mathrm{akr}_T(\lambda) = \mathrm{kr}_{\mathrm{p}_T}(\lambda).$$

Skalar λ je torej lastna vrednost endomorfizma T, če, in samo če, je $\mathrm{akr}_T(\lambda) > 0$.

Dimenzijo vektorskega prostora $\ker(T - \lambda\,\mathrm{id})$ imenujemo *geometrična kratnost* skalarja λ za endomorfizem T, in označimo

$$\mathrm{gkr}_T(\lambda) = \dim(\ker(T - \lambda\,\mathrm{id})).$$

Ponovno velja, da je skalar λ lastna vrednost endomorfizma T, če, in samo če, je $\mathrm{gkr}_T(\lambda) > 0$. Za vsak skalar $\lambda \in \mathbb{F}$ krajše zapišemo

$$\mathrm{E}_T(\lambda) = \ker(T - \lambda\,\mathrm{id}).$$

Vektorski prostor $\mathrm{E}_T(\lambda)$ je T-invarianten in zožitev $T|_{\mathrm{E}_T(\lambda)} : \mathrm{E}_T(\lambda) \to \mathrm{E}_T(\lambda)$ je enaka endomorfizmu $\lambda\,\mathrm{id}_{\mathrm{E}_T(\lambda)}$. Če skalar λ ni lastna vrednost endomorfizma T, potem je $\mathrm{E}_T(\lambda) = \{0\}$. Če je λ lastna vrednost endomorfizma T, potem vektorski prostor $\mathrm{E}_T(\lambda)$ imenujemo *lastni podprostor* endomorfizma T za lastno vrednost λ, neničelne vektorje iz vektorskega prostora $\mathrm{E}_T(\lambda)$ pa imenujemo *lastni vektorji* endomorfizma T za lastno vrednost λ. Lastni vektor endomorfizma T za lastno prednost λ je torej vsak neničelni vektor $v \in \mathsf{V}$, za katerega velja

$$Tv = \lambda v,$$

takšnih vektorjev pa je za vsako lastno vrednost neskončno.

Množico vseh lastnih vrednosti endomorfizma T označimo z $\sigma(T)$ in jo imenujemo *spekter* endomorfizma T. Velja torej $\sigma(T) = \{\lambda \in \mathbb{F}\,;\, \ker(T - \lambda\,\mathrm{id}_{\mathsf{V}}) \neq \{0\}\}$.

Lastne vektorje, lastne vrednosti z njihovimi algebraičnimi in geometričnimi kratnostmi, lastne podprostore ter spekter poljubne matrike $A \in \mathrm{Mat}(n \times n, \mathbb{F})$

definiramo kot lastne vektorje, lastne vrednosti z njihovimi algebraičnimi in geometričnimi kratnostmi, lastne podprostore oziroma spekter pripadajočega endomorfizma $\ell(A)$, in označimo

$$\begin{aligned} \mathrm{akr}_A(\lambda) &= \mathrm{akr}_{\ell(A)}(\lambda), & \mathrm{gkr}_A(\lambda) &= \mathrm{gkr}_{\ell(A)}(\lambda), \\ \mathrm{E}_A(\lambda) &= \mathrm{E}_{\ell(A)}(\lambda), & \sigma(A) &= \sigma(\ell(A)). \end{aligned}$$

TRDITEV 2.92. *Naj bodo $\lambda_1, \lambda_2, \ldots, \lambda_k$ paroma različne lastne vrednosti endomorfizma $T \in \mathrm{End}(\mathsf{V})$ in naj bodo $v_1, v_2, \ldots, v_k \in \mathsf{V}$ takšni neničelni vektorji, da je v_i lastni vektor za lastno vrednost λ_i za vsak $i = 1, 2, \ldots, k$. Tedaj so vektorji $v_1, v_2, \ldots, v_k$ med seboj linearno neodvisni.*

DOKAZ. Trditev bomo dokazali z indukcijo na k. Za $k = 1$ trditev sledi iz dejstva, da je vektor v_1 neničelen. Predpostavimo zdaj, da je $k \geq 2$ in da so vektorji $v_1, \ldots, v_{k-1}$ med seboj linearno neodvisni. Naj bodo $\alpha_1, \ldots, \alpha_{k-1}, \alpha_k \in \mathbb{F}$ skalarji, za katere je

$$\alpha_1 v_1 + \cdots + \alpha_{k-1} v_{k-1} + \alpha_k v_k = 0.$$

Odtod sledi

$$\alpha_1 \lambda_k v_1 + \cdots + \alpha_{k-1} \lambda_k v_{k-1} + \alpha_k \lambda_k v_k = 0,$$

poleg tega pa tudi

$$\begin{aligned} \alpha_1 \lambda_1 v_1 + \cdots + \alpha_{k-1} \lambda_{k-1} v_{k-1} + \alpha_k \lambda_k v_k &= \alpha_1 T v_1 + \cdots + \alpha_{k-1} T v_{k-1} + \alpha_k T v_k \\ &= T(\alpha_1 v_1 + \cdots + \alpha_{k-1} v_{k-1} + \alpha_k v_k) \\ &= T(0) = 0. \end{aligned}$$

Če zadnji dve enakosti odštejemo, dobimo

$$\alpha_1(\lambda_k - \lambda_1) v_1 + \cdots + \alpha_{k-1}(\lambda_k - \lambda_{k-1}) v_{k-1} = 0.$$

Po indukcijski predpostavki so vektorji $v_1, \ldots, v_{k-1}$ med seboj linearno neodvisni, zato odtod sledi

$$\alpha_i(\lambda_k - \lambda_i) = 0$$

za vsak $i = 1, \ldots, k-1$. Ker so lastne vrednosti $\lambda_1, \lambda_2, \ldots, \lambda_k$ paroma različne, velja torej $\alpha_i = 0$ za vsak $i = 1, \ldots, k-1$. Ker je tedaj $\alpha_k v_k = 0$ in ker je $v_k \neq 0$, je tudi $\alpha_k = 0$. □

POSLEDICA 2.93. *Naj bodo $\lambda_1, \lambda_2, \ldots, \lambda_m \in \mathbb{F}$ vse paroma različne lastne vrednosti endomorfizma $T \in \mathrm{End}(\mathsf{V})$. Tedaj je*

$$\sum_{\lambda \in \mathbb{F}} \mathrm{E}_T(\lambda) = \sum_{\lambda \in \sigma(T)} \mathrm{E}_T(\lambda) = \mathrm{E}_T(\lambda_1) \oplus \mathrm{E}_T(\lambda_2) \oplus \cdots \oplus \mathrm{E}_T(\lambda_m).$$

DOKAZ. Ker je $\mathrm{E}_T(\lambda) = \{0\}$ za $\lambda \notin \sigma(T) = \{\lambda_1, \lambda_2, \ldots, \lambda_m\}$, je vektorski podprostor $\sum_{\lambda \in \mathbb{F}} \mathrm{E}_T(\lambda) \subset \mathsf{V}$ enak vsoti $\mathrm{E}_T(\lambda_1) + \mathrm{E}_T(\lambda_2) + \cdots + \mathrm{E}_T(\lambda_m)$. Iz trditve 2.92 neposredno sledi, da je ta vsota direktna. □

TRDITEV 2.94. *Za vsak endomorfizem $T \in \mathrm{End}(\mathsf{V})$ in za vsak izomorfizem vektorskih prostorov $\Theta \in \mathrm{Lin}(\mathsf{V}, \mathsf{W})$ velja $\det(\Theta T \Theta^{-1}) = \det(T)$, $\mathrm{tr}(\Theta T \Theta^{-1}) = \mathrm{tr}(T)$, $\mathrm{p}_{\Theta T \Theta^{-1}} = \mathrm{p}_T$ in $\sigma(\Theta T \Theta^{-1}) = \sigma(T)$, velja pa tudi $\mathrm{akr}_{\Theta T \Theta^{-1}}(\lambda) = \mathrm{akr}_T(\lambda)$, $\mathrm{gkr}_{\Theta T \Theta^{-1}}(\lambda) = \mathrm{gkr}_T(\lambda)$ in $\mathrm{E}_{\Theta T \Theta^{-1}}(\lambda) = \Theta(\mathrm{E}_T(\lambda))$ za vsak $\lambda \in \mathbb{F}$.*

Dokaz. Prvi dve enakosti sledita direktno iz trditve 2.83 in trditve 2.81(ii). Ostalo sledi sledi iz enakosti

$$\begin{aligned}\mathrm{p}_{\Theta T\Theta^{-1}}(t) &= \det(\Theta T\Theta^{-1} - t\,\mathrm{id}_{\mathsf{W}}) = \det(\Theta(T - t\,\mathrm{id}_{\mathsf{V}})\Theta^{-1}) \\ &= \det(T - t\,\mathrm{id}_{\mathsf{V}}) = \mathrm{p}_T(t),\end{aligned}$$

$$\begin{aligned}\mathrm{E}_{\Theta T\Theta^{-1}}(\lambda) &= \ker(\Theta T\Theta^{-1} - t\,\mathrm{id}_{\mathsf{W}}) = \ker(\Theta(T - t\,\mathrm{id}_{\mathsf{V}})\Theta^{-1}) \\ &= \Theta(\ker(T - t\,\mathrm{id}_{\mathsf{V}})) = \Theta(\mathrm{E}_T(\lambda)).\end{aligned}$$

□

Zgled 2.95. (1) Če je $A \in \mathrm{Mat}(n \times n, \mathbb{F})$ zgornje-trikotna ali spodnje-trikotna matrika, potem velja

$$\mathrm{p}_A(t) = (A_{11} - t)(A_{22} - t)\cdots(A_{nn} - t),$$

zato so $A_{11}, A_{22}, \ldots, A_{nn}$ natanko vse lastne vrednosti matrike A in velja

$$\mathrm{akr}_A(\lambda) = \#\{i \in \{1, 2, \ldots, n\}\ ;\ A_{ii} = \lambda\}$$

za vsak $\lambda \in \mathbb{F}$.

(2) Naj bo $A \in \mathrm{Mat}(n \times n, \mathbb{F})$ diagonalna matrika,

$$A = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n).$$

Karakteristični polinom matrike A je enak polinomu

$$\mathrm{p}_A(t) = (\lambda_1 - t)(\lambda_2 - t)\cdots(\lambda_n - t)$$

in $\lambda_1, \lambda_2, \ldots, \lambda_n$ so natanko vse lastne vrednosti matrike A. Ker velja $A\mathrm{e}_i = \lambda_i \mathrm{e}_i$, je standardni bazni vektor e_i lastni vektor pri lastni vrednosti λ_i, za vsak i. Za poljuben $\lambda \in \mathbb{F}$ si ogledamo rešitve homogenega sistema $(A - \lambda \mathrm{I})v = 0$ in vidimo, da je

$$\mathrm{E}_A(\lambda) = \mathrm{Span}(\{\mathrm{e}_i\ ;\ i \in \{1, 2, \ldots, n\},\ \lambda_i = \lambda\}),$$

$$\mathrm{gkr}_A(\lambda) = \mathrm{akr}_A(\lambda) = \#\{i \in \{1, 2, \ldots, n\}\ ;\ \lambda_i = \lambda\}.$$

Trditev 2.96. *Naj bo $T \in \mathrm{End}(\mathsf{V})$ endomorfizem in naj bo $\mathscr{B} = [v_1, v_2, \ldots, v_n]$ baza vektorskega prostora V. Koordinatna matrika $A = [T]_{\mathscr{B}}^{\mathscr{B}}$ endomorfizma T v bazi $\mathscr{B}$ je diagonalna, če, in samo če, so vektorji $v_1, v_2, \ldots, v_n$ lastni vektorji endomorfizma T. V tem primeru so komponente $A_{11}, A_{22}, \ldots, A_{nn}$ natanko vse lastne vrednosti endomorfizma T, pri čemer velja*

$$\mathrm{p}_T(t) = (A_{11} - t)(A_{22} - t)\cdots(A_{nn} - t),$$

$$\mathrm{E}_T(\lambda) = \mathrm{Span}(\{v_i\ ;\ i \in \{1, 2, \ldots, n\},\ A_{ii} = \lambda\})$$

in

$$\mathrm{gkr}_T(\lambda) = \mathrm{akr}_T(\lambda) = \#\{i \in \{1, 2, \ldots, n\}\ ;\ A_{ii} = \lambda\}$$

za vsak skalar $\lambda \in \mathbb{F}$.

Dokaz. Trditev je direktna posledica zgleda 2.95 in trditve 2.94, v kateri vzamemo $\Theta = [\,\bullet\,]_{\mathscr{B}} : \mathsf{V} \to \mathbb{F}^n$. □

2.2.3. Diagonalizabilnost.

Definicija 2.97. Endomorfizem $T \in \mathrm{End}(\mathsf{V})$ je *diagonalizabilen*, če obstaja takšna baza $\mathscr{B}$ vektorskega prostora V, da je matrika $[T]^{\mathscr{B}}_{\mathscr{B}}$ diagonalna.

Komentar 2.98. Če poiščemo takšno bazo $\mathscr{B}$ vektorskega prostora V, da je koordinatna matrika $[T]^{\mathscr{B}}_{\mathscr{B}}$ diagonalna, in to koordinatno matriko tudi izračunamo, potem pravimo, da smo endomorfizem T *diagonalizirali*. Za diagonalizacijo endomorfizma T je po trditvi 2.96 torej potrebno najti bazo vektorskega prostora V, ki je sestavljena iz lastnih vektorjev endomorfizma T.

Matrika $A \in \mathrm{Mat}(n \times n, \mathbb{F})$ je diagonalizabilna tedaj, ko je endomorfizem $\ell(A)$ diagonalizabilen. Po trditvi 2.81(i) je kvadratna matrika A diagonalizabilna, če, in samo če, je podobna kakšni diagonalni matriki, torej, če, in samo če, obstaja takšna obrnljiva matrika $P \in \mathrm{GL}(n, \mathbb{F})$, da je matrika $P^{-1}AP$ diagonalna. V tem primeru so stolpci matrike P lastni vektorji matrike A. Stolpci matrike P torej sestavljajo takšno bazo $\mathscr{B}$ vektorskega prostora $\mathbb{F}^n$, da je $D = [\ell(A)]^{\mathscr{B}}_{\mathscr{B}} = P^{-1}AP$ diagonalna matrika, pri čemer je $P = [\mathrm{id}]^{\mathscr{B}}_{\mathscr{E}}$ in velja

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

Trditev 2.99. *Za poljuben endomorfizem $T \in \mathrm{End}(\mathsf{V})$ so si naslednje trditve ekvivalentne:*

(i) *endomorfizem T je diagonalizabilen,*
(ii) *obstaja baza vektorskega prostora V, sestavljena iz lastnih vektorjev endomorfizma T,*
(iii) $\sum_{\lambda \in \mathbb{F}} \mathrm{E}_T(\lambda) = \mathsf{V}$,
(iv) $\sum_{\lambda \in \mathbb{F}} \mathrm{gkr}_T(\lambda) = \dim_{\mathbb{F}} \mathsf{V}$.

Dokaz. Ekvivalenca med točkama (i) in (ii) sledi iz trditve 2.96, ostalo pa sledi iz posledice 2.93. □

Komentar 2.100. Če je endomorfizem $T \in \mathrm{End}(\mathsf{V})$ diagonalizabilen, potem je torej vektorski prostor V direktna vsota lastnih podprostorov

$$\mathsf{V} = \mathrm{E}_T(\lambda_1) \oplus \mathrm{E}_T(\lambda_2) \oplus \cdots \oplus \mathrm{E}_T(\lambda_m),$$

kjer so $\lambda_1, \lambda_2, \ldots, \lambda_m \in \mathbb{F}$ vse paroma različne lastne vrednosti endomorfizma T (glej posledico 2.93). Vsak lastni podprostor $\mathrm{E}_T(\lambda_i)$ je T-invarianten, zožitev $T|_{\mathrm{E}_T(\lambda_i)} : \mathrm{E}_T(\lambda_i) \to \mathrm{E}_T(\lambda_i)$ pa je enaka endomorfizmu $\lambda_i \, \mathrm{id}_{\mathrm{E}_T(\lambda_i)}$. Odtod sledi, da je endomorfizem T direktna vsota

$$T = \lambda_1 \, \mathrm{id}_{\mathrm{E}_T(\lambda_1)} \oplus \lambda_2 \, \mathrm{id}_{\mathrm{E}_T(\lambda_2)} \oplus \cdots \oplus \lambda_m \, \mathrm{id}_{\mathrm{E}_T(\lambda_m)}.$$

Ni pa se težko prepričati, da velja tudi obratno: Če za endomorfizem $T \in \mathrm{End}(\mathsf{V})$ obstajajo takšni netrivialni T-invariantni vektorski podprostori $\mathsf{V}_1, \mathsf{V}_2, \ldots, \mathsf{V}_k$ vektorskega prostora V, da je

$$\mathsf{V} = \mathsf{V}_1 \oplus \mathsf{V}_2 \oplus \cdots \oplus \mathsf{V}_k$$

in da je

$$T = \eta_1 \, \mathrm{id}_{\mathsf{V}_1} \oplus \eta_2 \, \mathrm{id}_{\mathsf{V}_2} \oplus \cdots \oplus \eta_k \, \mathrm{id}_{\mathsf{V}_k}$$

za neke skalarje $\eta_1, \eta_2, \ldots, \eta_k$, potem je endomorfizem T diagonalizabilen in velja $\sigma(T) = \{\eta_1, \eta_2, \ldots, \eta_k\}$.

IZREK 2.101. *Naj bo $T \in \mathrm{End}(\mathsf{V})$ poljuben endomorfizem končno-dimenzionalenega vektorskega prostora V nad $\mathbb{F}$ pozitivne dimenzije.*

(i) Za vsak $\lambda \in \mathbb{F}$ velja $\mathrm{gkr}_T(\lambda) \leq \mathrm{akr}_T(\lambda)$.

(ii) Če je $\mathbb{F} = \mathbb{C}$, potem je endomorfizem T diagonalizabilen, če, in samo če, velja $\mathrm{akr}_T(\lambda) = \mathrm{gkr}_T(\lambda)$ za vsak $\lambda \in \mathbb{C}$.

(iii) Če je $\mathbb{F} = \mathbb{R}$, potem je endomorfizem T diagonalizabilen, če, in samo če, je karakteristični polinom p_T razcepen na realne linearne faktorje in velja $\mathrm{akr}_T(\lambda) = \mathrm{gkr}_T(\lambda)$ za vsak $\lambda \in \mathbb{R}$.

DOKAZ. (i) Naj bodo $\lambda_1, \lambda_2, \ldots, \lambda_m \in \mathbb{F}$ vse paroma različne lastne vrednosti endomorfizma $T \in \mathrm{End}(\mathsf{V})$. Če je $\sum_{\lambda\in\mathbb{F}} \mathrm{E}_T(\lambda) = \mathsf{V}$, potem je endomorfizem T diagonalizabilen in točka (i) sledi direktno iz trditve 2.96, sicer pa izberimo komplementaren podprostor U vektorskega podprostora $\sum_{\lambda\in\mathbb{F}} \mathrm{E}_T(\lambda)$ v vektorskem prostoru V. Po posledici 2.93 tedaj velja

$$\mathsf{V} = \mathrm{E}_T(\lambda_1) \oplus \mathrm{E}_T(\lambda_2) \oplus \cdots \oplus \mathrm{E}_T(\lambda_m) \oplus \mathsf{U}.$$

Označimo $n_i = \mathrm{gkr}_T(\lambda_i)$ in izberimo bazo $[v_1^{(i)}, v_2^{(i)}, \ldots, v_{n_i}^{(i)}]$ vektorskega prostora $\mathrm{E}_T(\lambda_i)$, za $i = 1, 2, \ldots, m$. Izberimo še neko bazo $[u_1, u_2, \ldots, u_q]$ vektorskega prostora U. S tem smo dobili bazo

$$\mathscr{B} = [v_1^{(1)}, \ldots, v_{n_1}^{(1)}, v_1^{(2)}, \ldots, v_{n_2}^{(2)}, \ldots, v_1^{(m)}, \ldots, v_{n_m}^{(m)}, u_1, \ldots, u_q]$$

vektorskega prostora V. Koordinatna matrika endomorfizma T v tej bazi je tedaj naslednje bločne oblike:

$$[T]_{\mathscr{B}}^{\mathscr{B}} = \begin{bmatrix} \lambda_1 \mathrm{I}_{n_1\times n_1} & 0 & \cdots & 0 & * \\ 0 & \lambda_2 \mathrm{I}_{n_2\times n_2} & \cdots & 0 & * \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda_m \mathrm{I}_{n_m\times n_m} & * \\ 0 & 0 & \cdots & 0 & * \end{bmatrix}$$

Iz te oblike matrike je jasno, da je

$$\mathrm{p}_T(t) = (\lambda_1 - t)^{n_1}(\lambda_2 - t)^{n_2} \cdots (\lambda_m - t)^{n_m} g(t)$$

za nek polinom g, zato je $\mathrm{gkr}_T(\lambda_i) = n_i \leq \mathrm{akr}_T(\lambda_i)$ za vsak i.

Točki (ii) in (iii) sta posledici točke (i) in trditve 2.99(iv). □

Direktna posledica izreka je:

POSLEDICA 2.102. *Vsak endomorfizem $T \in \mathrm{End}(\mathsf{V})$ končno-dimenzionalnega vektorskega prostora V nad $\mathbb{F}$ dimenzije $n = \dim_{\mathbb{F}} \mathsf{V}$, ki ima n različnih lastnih vrednosti v $\mathbb{F}$, je diagonalizabilen.*

ZGLED 2.103. (1) Karakteristični polinom realne matrike

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

je enak polinomu

$$\mathrm{p}_A(t) = \begin{vmatrix} 2-t & 1 \\ 1 & 2-t \end{vmatrix} = (2-t)^2 - 1 = t^2 - 4t + 3 = (t-1)(t-3)$$

in ima torej dve ničli, ki sta obe realni. Matrika A ima torej dve lastni vrednosti, $\lambda_1 = 1$ in $\lambda_2 = 3$, in obe imata algebraično kratnost enako številu 1. Ker geometrična kratnost ne more biti večja od algebraične, odtod sledi $\mathrm{gkr}_A(1) = \mathrm{akr}_A(1) = 1$

in $\mathrm{gkr}_A(3) = \mathrm{akr}_A(3) = 1$. Matrika A je torej diagonalizabilna. Zdaj matriko A diagonalizirajmo, torej poiščimo takšno bazo $\mathscr{B}$ vektorskega prostora $\mathbb{R}^2$, da je matrika $[\ell(A)]_{\mathscr{B}}^{\mathscr{B}}$ diagonalna. Kot vemo, je potrebno takšno bazo sestaviti iz lastnih vektorjev matrike A.

Pri lastni vrednosti $\lambda_1 = 1$ je lastni podprostor $\mathrm{E}_A(1)$ sestavljen iz rešitev homogenega sistema enačb:

$$\begin{aligned} x + y &= 0 \\ x + y &= 0 \end{aligned}$$

Odtod izračunamo, da je $\mathrm{E}_A(1) = \{(-y, y)\ ;\ y \in \mathbb{R}\}$. V enodimenzionalnem vektorskem prostoru $\mathrm{E}_A(1)$ izberemo bazo oziroma bazni vektor, na primer $v_1 = (-1, 1)$. Vektor v_1 je torej lastni vektor pri lastni vrednosti $\lambda_1 = 1$, vsi lastni vektorji pri tej lastni vrednosti pa so oblike αv_1 za poljubno neničelno realno število α.

Pri lastni vrednosti $\lambda_1 = 3$ je lastni podprostor $\mathrm{E}_A(3)$ sestavljen iz rešitev homogenega sistema enačb:

$$\begin{aligned} -x + y &= 0 \\ x - y &= 0 \end{aligned}$$

Odtod izračunamo, da je $\mathrm{E}_A(1) = \{(y, y)\ ;\ y \in \mathbb{R}\}$. V enodimenzionalnem vektorskem prostoru $\mathrm{E}_A(1)$ izberemo bazo oziroma bazni vektor, na primer $v_2 = (1, 1)$. Vektor v_2 je torej lastni vektor pri lastni vrednosti $\lambda_2 = 3$, vsi lastni vektorji pri tej lastni vrednosti pa so oblike αv_2 za poljubno neničelno realno število α.

Na ta način smo dobili bazo

$$\mathscr{B} = [v_1, v_2] = [(-1, 1), (1, 1)]$$

vektorskega prostora $\mathbb{R}^2$, ki je sestavljena iz lastnih vektorjev matrike A. Prehodna matrika z baze $\mathscr{B}$ na standardno bazo $\mathscr{E}$ je obrnljiva matrika:

$$P = [\mathrm{id}]_{\mathscr{E}}^{\mathscr{B}} = \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix}$$

Če izračunamo inverz matrike P, dobimo:

$$P^{-1} = [\mathrm{id}]_{\mathscr{B}}^{\mathscr{E}} = \frac{1}{2} \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix}$$

Matrika

$$D = [\ell(A)]_{\mathscr{B}}^{\mathscr{B}} = P^{-1}AP = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}$$

je diagonalna, na diagonali ima lastne vrednosti matrike A, in velja $A = PDP^{-1}$.

(2) Karakteristični polinom matrike

$$A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

je enak polinomu

$$\mathrm{p}_A(t) = \begin{vmatrix} -t & -1 \\ 1 & -t \end{vmatrix} = t^2 + 1 = (t - \mathrm{i})(t + \mathrm{i})$$

in ima torej dve ničli, ki pa nista realni. Matrika A kot realna matrika nima (realnih) lastnih vrednosti in ni diagonalizabilna nad realnimi skalarji. Vendar pa je vsaka realna matrika poseben primer kompleksne matrike. Matrika A je torej tudi kompleksna matrika in ima kot taka dve lastni vrednosti, $\lambda_1 = \mathrm{i}$ in $\lambda_2 = -\mathrm{i}$, in obe imata algebraično kratnost enako številu 1. Odtod sledi, da sta tudi geometrični kratnosti teh dveh lastnih vrednosti enaki številu 1 in da je matrika A diagonalizabilna nad kompleksnimi skalarji. Drugače povedano, z matriko A je v standardni bazi podan endomorfizem vektorskega prostora $\mathbb{R}^2$, ki ni diagonalizabilen, a z isto matriko je v standardni bazi podan tudi endomorfizem vektorskega prostora $\mathbb{C}^2$, ki pa je diagonalizabilen.

Za lastno vrednosti $\lambda_1 = \mathrm{i}$ lahko izberemo na primer lastni vektor $v_1 = (1, -\mathrm{i})$, ki sestavlja bazo kompleksnega vektorskega podprostora $\mathrm{E}_A(\mathrm{i}) \subset \mathbb{C}^2$. Podobno si za lastno vrednosti $\lambda_2 = -\mathrm{i}$ lahko izberemo lastni vektor $v_2 = (\mathrm{i}, -1)$, ki sestavlja bazo kompleksnega vektorskega podprostora $\mathrm{E}_A(-\mathrm{i}) \subset \mathbb{C}^2$. Vektorja v_1, v_2 sestavljata bazo $\mathscr{B}$ kompleksnega vektorskega prostora $\mathbb{C}^2$, matrika

$$P = \begin{bmatrix} 1 & \mathrm{i} \\ -\mathrm{i} & -1 \end{bmatrix}$$

pa je prehodna matrika z baze $\mathscr{B}$ na standardno bazo $\mathscr{E}$. Matrika

$$D = [\ell(A)]_{\mathscr{B}}^{\mathscr{B}} = \begin{bmatrix} \mathrm{i} & 0 \\ 0 & -\mathrm{i} \end{bmatrix}$$

je diagonalna in $A = PDP^{-1}$.

(3) Karakteristični polinom matrike

$$A = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}$$

je enak polinomu

$$\mathrm{p}_A(t) = \begin{vmatrix} 2-t & 1 \\ 0 & 2-t \end{vmatrix} = (2-t)^2$$

in ima torej eno samo ničlo $\lambda_1 = 2$, ki je realna in ima kratnost 2. Število 2 je torej edina lastna vrednosti matrike A in je algebraične kratnosti 2. Lastni podprostor $\mathrm{E}_A(2)$ je sestavljen iz rešitev sistema linearnih enačb

$$\begin{aligned} 0x + 1y &= 0 \\ 0x + 0y &= 0 \end{aligned}$$

in je zato enodimenzionalen, $\mathrm{E}_A(2) = \{(x, 0)\ ;\ x \in \mathbb{F}\}$. To velja tako v primeru $\mathbb{F} = \mathbb{R}$, ko gledamo matriko A kot realno matriko oziroma kot matriko endomorfizma realnega vektorskega prostora $\mathbb{R}^2$, kot tudi v primeru, ko gledamo matriko A kot kompleksno matriko oziroma kot matriko endomorfizma kompleksnega vektorskega prostora $\mathbb{C}^2$. V obeh primerih je

$$\mathrm{gkr}_A(2) = 1 < 2 = \mathrm{akr}_A(2),$$

zato matrika A ni diagonalizabilna niti nad skalarji $\mathbb{R}$ niti nad skalarji $\mathbb{C}$.

(4) V zgledu 2.78(5) smo si ogledali endomorfizem $T : \mathbb{R}^3 \to \mathbb{R}^3$, ki je podan kot pravokotna projekcija na ravnino $\Sigma \subset \mathbb{R}^3$. Pri tem ravnina Σ vsebuje izhodišče in je torej dvodimenzionalni vektorski podprostor prostora $\mathbb{R}^3$. Izbrali smo bazo $\mathscr{B} = [v_1, v_2, v_3]$ vektorskega prostora $\mathbb{R}^3$ tako, da je v_1 normala ravnine Σ, vektorja v_2 in v_3 pa sta pravokotna na to normalo in zato ležita v ravnini Σ. Izračunali smo, da je koordinatna matrika $[T]_{\mathscr{B}}^{\mathscr{B}}$ diagonalna, in sicer

$$[T]_{\mathscr{B}}^{\mathscr{B}} = \mathrm{diag}(0, 1, 1).$$

V zgledu 2.78(5) smo torej že diagonalizirali endomorfizem T. Baza $\mathscr{B} = [v_1, v_2, v_3]$ je sestavljena iz lastnih vektorjev endomorfizma T, iz matrike $[T]_{\mathscr{B}}^{\mathscr{B}}$ pa lahko vidimo, da je

$$\mathrm{p}_T(t) = -t(1-t)^2,$$

da sta $\lambda_1 = 0$ in $\lambda_2 = 1$ edini lastni vrednosti endomorfizma T in da velja $\mathrm{akr}_T(0) = \mathrm{gkr}_T(0) = 1$ ter $\mathrm{akr}_T(1) = \mathrm{gkr}_T(1) = 2$.

(5) Endomorfizem $T \in \mathrm{End}(\mathsf{V})$ je *nilpotenten*, če obstaja takšno naravno število k, da je $T^k = 0$. Tu smo s T^k označili kompozicijo k kopij endomorfizma T. S tem smo posplošili pojem nilpotentne matrike, ki smo ga spoznali v prvem poglavju. Za nilpotenten endomorfizem T lahko izberemo najmanjše takšno naravno število k, da je $T^k = 0$, in to število k imenujemo *red* nilpotenta T.

Naj bo $T \in \mathrm{End}(\mathsf{V})$ poljuben nilpotent reda k. Najprej lahko opazimo, da je $\mathrm{null}(T) > 0$, saj bi bil sicer T avtomorfizem in bi bil zato tudi T^k avtomorfizem, kar pa ni res. Odtod sledi, da je vektorski prostor $\mathrm{E}_T(0) = \ker(T)$ netrivialen, zato je število 0 lastna vrednosti endomorfizma T.

Koordinatna matrika $[T]_{\mathscr{B}}^{\mathscr{B}}$ v poljubni bazi $\mathscr{B}$ vektorskega prostora V je prav tako nilpotentna reda k, saj je $[T^p]_{\mathscr{B}}^{\mathscr{B}} = ([T]_{\mathscr{B}}^{\mathscr{B}})^p$ za vsako naravno število p. Velja torej $([T]_{\mathscr{B}}^{\mathscr{B}})^k = 0$. Matriko $[T]_{\mathscr{B}}^{\mathscr{B}}$ lahko gledamo kot kompleksno matriko, ne glede na to, ali je $\mathbb{F} = \mathbb{R}$ ali $\mathbb{F} = \mathbb{C}$. Če za nek neničelen vektor $u \in \mathbb{C}^n$ in neko število $\lambda \in \mathbb{C}$ velja $[T]_{\mathscr{B}}^{\mathscr{B}} u = \lambda u$, odtod sledi

$$0 = \left([T]_{\mathscr{B}}^{\mathscr{B}}\right)^k u = \lambda^k u$$

in zato $\lambda = 0$. To pomeni, da je 0 edina lastna vrednost kompleksne matrike $[T]_{\mathscr{B}}^{\mathscr{B}}$, zato je 0 tudi edina lastna vrednost endomorfizma T, $\mathrm{akr}_T(0) = \dim(V) = n$ in

$$\mathrm{p}_T(t) = (-t)^n.$$

Če je $k = 1$, je nilpotent T enak nič in je seveda diagonalizabilen. Če je $k \geq 2$, je T neničeln in $\mathrm{gkr}_T(0) < n = \mathrm{akr}_T(0)$, zato nilpotent T tedaj ni diagonalizabilen.

(6) Endomorfizem $T \in \mathrm{End}(\mathsf{V})$ je *projektor*, če velja $T^2 = T$. Projektorje na naraven način konstruiramo iz parov komplementarnih podprostorov vektorskega prostora V. Če sta namreč U in W komplementarna vektorska podprostora vektorskega prostora V, potem po trditvi 2.64 obstaja natanko en tak endomorfizem $\mathrm{pr}_{\mathsf{U}}^{\mathsf{W}} \in \mathrm{End}(\mathsf{V})$, da je $\mathrm{pr}_{\mathsf{U}}^{\mathsf{W}}|_{\mathsf{U}} = \mathrm{id}$ in $\mathrm{pr}_{\mathsf{U}}^{\mathsf{W}}|_{\mathsf{W}} = 0$. Tedaj je $(\mathrm{pr}_{\mathsf{U}}^{\mathsf{W}})^2 = \mathrm{pr}_{\mathsf{U}}^{\mathsf{W}}$ in $\mathrm{pr}_{\mathsf{U}}^{\mathsf{W}}$ je torej projektor. Dobljeni projektor $\mathrm{pr}_{\mathsf{U}}^{\mathsf{W}}$ imenujemo *projektor na vektorski podprostor* U *vzdolž vektorskega podprostora* W.

Za poljuben projektor $T \in \mathrm{End}(\mathsf{V})$ velja

$$\mathsf{V} = \mathrm{im}(T) \oplus \ker(T).$$

Res, vsak vektor $v \in \mathsf{V}$ lahko zapišemo kot vsoto $v = Tv + (v - Tv)$, kjer je $Tv \in \mathrm{im}(T)$ in $(v - Tv) \in \ker(T)$, saj je $T(v - Tv) = Tv - T^2 v = Tv - Tv = 0$. Poleg tega je tudi $\mathrm{im}(T) \cap \ker(T) = \{0\}$, saj je poljuben vektor $w \in \mathrm{im}(T) \cap \ker(T)$ oblike $w = Tv$ za nek vektor $v \in \mathsf{V}$ in velja $0 = Tw = T^2 v = Tv = w$, torej je $w = 0$. Za endomorfizem T očitno velja $T|_{\mathrm{im}(T)} = \mathrm{id}$ in $T|_{\ker(T)} = 0$. Poljuben projektor $T \in \mathrm{End}(\mathsf{V})$ je torej projektor na vektorski podprostor $\mathrm{im}(T)$ vzdolž vektorskega podprostora $\ker(T)$, torej

$$T = \mathrm{pr}_{\mathrm{im}(T)}^{\ker(T)}.$$

Označimo $r = \mathrm{rank}(T)$ in $d = \mathrm{null}(T)$. Ker je $T|_{\mathrm{im}(T)} = \mathrm{id}$ in $T|_{\ker(T)} = 0$, je $\mathrm{im}(T) \subset \mathrm{E}_T(1)$ in $\ker(T) \subset \mathrm{E}_T(0)$. Posebej je torej $r \leq \mathrm{gkr}_T(1)$ in $d \leq \mathrm{gkr}_T(0)$.

A ker je hkrati $r + d = n$ in $\mathrm{gkr}_T(1) + \mathrm{gkr}_T(0) \leq n$, mora veljati $r = \mathrm{gkr}_T(1)$ in $d = \mathrm{gkr}_T(0)$, oziroma $\mathrm{im}(T) = \mathrm{E}_T(1)$ in $\ker(T) = \mathrm{E}_T(0)$. To posebej pomeni, da je endomorfizem T diagonalizabilen in da je

$$\mathrm{p}_T(t) = (-t)^d (1-t)^r.$$

V posebnem primeru, ko je $r = n$, je $T = \mathrm{id}$ in tedaj je 1 edina lastna vrednost endomorfizma T. V primeru, ko je $r = 0$, je $T = 0$ in 0 je edina lastna vrednost endomorfizma T. Opazimo lahko, da sta koordinatni matriki identitete in ničelnega endomorfizma diagonalni v vsaki bazi vektorskega prostora V.

Če je $0 < r < n$, potem ima T natanko dve lastni vrednosti, in sicer 1 in 0. Če v tem primeru izberemo bazo $[u_1, u_2, \ldots, u_r]$ vektorskega podprostora $\mathrm{im}(T)$ in bazo $[w_1, w_2, \ldots, w_d]$ vektorskega podprostora $\ker(T)$, potem je

$$\mathscr{B} = [u_1, \ldots, u_r, w_1, \ldots, w_d]$$

baza vektorskega prostora V, sestavljena iz lastnih vektorjev endomorfizma T. Koordinatna matrika endomorfizma T v bazi $\mathscr{B}$ je diagonalna,

$$[T]_{\mathscr{B}}^{\mathscr{B}} = \mathrm{diag}(1, \ldots, 1, 0, \ldots, 0),$$

pri tem pa na diagonali te matrike število 1 nastopa r-krat, število 0 pa d-krat.

Vaja 2.104. (1) Naj bo $T \in \mathrm{End}(\mathsf{V})$ diagonalizabilen endomorfizem vektorskega prostora V in naj bodo $\lambda_1, \lambda_2, \ldots, \lambda_m \in \mathbb{F}$ natanko vse paroma različne lastne vrednosti endomorfizma T. S pomočjo Vandermondeove determinante pokaži, da za vsak T-invarianten vektorski podprostor U vektorskega prostora V velja

$$\mathsf{U} = (\mathsf{U} \cap \mathrm{E}_T(\lambda_1)) \oplus (\mathsf{U} \cap \mathrm{E}_T(\lambda_2)) \oplus \cdots \oplus (\mathsf{U} \cap \mathrm{E}_T(\lambda_m)).$$

Če je vektorski prostor V direktna vsota svojih T-invariantnih vektorskih podprostorov $\mathsf{U}_1, \mathsf{U}_2, \ldots, \mathsf{U}_k$, potem je vektorski prostor V tudi direktna vsota presekov $\mathsf{U}_i \cap \mathrm{E}_T(\lambda_j)$, $i = 1, \ldots, k$, $j = 1, \ldots, m$.

(2) Če za endomorfizma $S, T \in \mathrm{End}(\mathsf{V})$ velja $ST = TS$, potem pravimo, da endomorfizma S in T *komutirata*. Pokaži, da je v tem primeru vektorski podprostor $\mathrm{E}_S(\lambda)$ T-invarianten, za vsak skalar $\lambda \in \mathbb{F}$.

(3) Naj bosta S in T diagonalizabilna endomorfizma vektorskega prostora V. Pokaži, da endomorfizma S in T komutirata, če, in samo če, obstaja takšna baza $\mathscr{B}$ vektorskega prostora V, da sta matriki $[S]_{\mathscr{B}}^{\mathscr{B}}$ in $[T]_{\mathscr{B}}^{\mathscr{B}}$ obe diagonalni.

(4) Diagonaliziraj matriko:

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 \\ 2 & -2 & 3 \end{bmatrix}$$

2.2.4. Schurov izrek in Cayley-Hamiltonov izrek.

Izrek 2.105 (Schurov izrek). *Naj bo $T \in \mathrm{End}(\mathsf{V})$ endomorfizem končno-dimenzionalenega kompleksnega vektorskega prostora V pozitivne dimenzije. Tedaj obstaja takšna baza $\mathscr{B}$ vektorskega prostora V, da je koordinatna matrika $[T]_{\mathscr{B}}^{\mathscr{B}}$ zgornje-trikotna.*

Dokaz. Izrek bomo dokazali z indukcijo na dimenzijo $n = \dim \mathsf{V}$ vektorskega prostora V. Če je $n = 1$, ni kaj dokazovati, saj je vsaka matrika velikosti 1×1 diagonalna in tudi zgornje-trikotna.

Predpostavimo zdaj, da je $n > 1$ in da izrek velja za vse endomorfizme kompleksnih vektorskih prostorov dimenzije $n-1$. Karakteristični polinom endomorfizma T ima vsaj eno kompleksno ničlo λ_1, ki je torej lastna vrednost endomorfizma T. Za to lastno vrednost obstaja vsaj en lastni vektor $v_1 \in \mathsf{V}$. Izberimo poljuben vektorski podprostor U vektorskega prostora V, ki je komplementaren vektorskemu podprostoru $\mathbb{C}v_1$. To pomeni, da je

$$\mathsf{V} = \mathbb{C}v_1 \oplus \mathsf{U}$$

in da je $\dim \mathsf{U} = n-1$. Naj bo $R : \mathsf{V} \to \mathsf{U}$ tista linearna preslikava, za katero je $R|_{\mathbb{C}v_1} = 0$ in $R|_{\mathsf{U}} = \mathrm{id}$. Definirajmo endomorfizem $S \in \mathrm{End}(\mathsf{U})$ s predpisom

$$S = R \circ T|_U : \mathsf{U} \to \mathsf{U}.$$

Po indukcijski predpostavki lahko najdemo takšno bazo $\mathscr{B}' = [v_2, v_3, \ldots, v_n]$ vektorskega prostora U, da je koordinatna matrika $B = [S]_{\mathscr{B}'}^{\mathscr{B}'}$ zgornje-trikotna. Na ta način smo dobili bazo $\mathscr{B} = [v_1, v_2, \ldots, v_n]$ vektorskega prostora V. Pokazati moramo še, da je koordinatna matrika $A = [T]_{\mathscr{B}}^{\mathscr{B}}$ res zgornje-trikotna.

Za poljuben $j = 2, 3, \ldots, n$ velja

$$Tv_j = A_{1j}v_1 + A_{2j}v_2 + \cdots + A_{nj}v_n,$$

zato je

$$\begin{aligned} Sv_j &= (R \circ T|_U)(v_j) = R(Tv_j) \\ &= R(A_{1j}v_1 + A_{2j}v_2 + \cdots + A_{nj}v_n) \\ &= A_{1j}Rv_1 + A_{2j}Rv_2 + \cdots + A_{nj}Rv_n \\ &= A_{2j}v_2 + \cdots + A_{nj}v_n. \end{aligned}$$

Ker po drugi strani iz definicije matrike B sledi

$$Sv_j = B_{1(j-1)}v_2 + \cdots + B_{(n-1)(j-1)}v_n,$$

smo s tem dokazali, da je $A_{ij} = B_{(i-1)(j-1)}$ za vse $i, j = 2, \ldots, n$. Poleg tega velja $[Tv_1]_{\mathscr{B}} = \lambda_1 \mathrm{e}_1$, saj je $Tv_1 = \lambda_1 v_1$. Matrika A je torej bločno oblike

$$A = \begin{bmatrix} \lambda_1 & C \\ 0 & B \end{bmatrix}$$

za neko kompleksno matriko C velikosti $1 \times (n-1)$. Ker je B zgornje-trikotna, je takšna tudi matrika A. $\square$

Direktna posledica Schurovega izreka in zgleda 2.95(1) je:

POSLEDICA 2.106. *Naj bo $T \in \mathrm{End}(\mathsf{V})$ endomorfizem končno-dimenzionalenega kompleksnega vektorskega prostora V dimenzije $n \geq 1$ in naj bodo $\lambda_1, \lambda_2, \ldots, \lambda_m$ vse paroma različne lastne vrednosti endomorfizma T. Tedaj je*

$$\det(T) = \lambda_1^{\mathrm{akr}_T(\lambda_1)} \lambda_2^{\mathrm{akr}_T(\lambda_2)} \cdots \lambda_m^{\mathrm{akr}_T(\lambda_m)}$$

in

$$\mathrm{tr}(T) = \mathrm{akr}_T(\lambda_1)\lambda_1 + \mathrm{akr}_T(\lambda_2)\lambda_2 + \cdots + \mathrm{akr}_T(\lambda_m)\lambda_m.$$

Naj bo $T \in \mathrm{End}(\mathsf{V})$ endomorfizem. Za poljubno naravno število k označimo z $T^k = TT \cdots T$ kompozicijo k kopij endomorfizma T. Posebej definiramo tudi $T^0 = \mathrm{id}_{\mathsf{V}}$. Za poljuben polinom $f \in \mathbb{F}[t]$ oblike $f(t) = a_0 + a_1 t + a_2 t^2 + \cdots + a_s t^s$ lahko definiramo endomorfizem

$$f(T) = a_0 \,\mathrm{id}_{\mathsf{V}} + a_1 T + a_2 T^2 + \cdots + a_s T^s \in \mathrm{End}(\mathsf{V}),$$

ki mu pravimo *vrednost polinoma f v endomorfizmu T*. S tem smo definirali preslikavo

$$\mathrm{ev}_T : \mathbb{F}[t] \to \mathrm{End}(\mathsf{V}), \qquad T \mapsto f(T).$$

Ni težko preveriti, da je ta preslikava linearna, da velja $\mathrm{ev}_T(1) = \mathrm{id}_{\mathsf{V}}$ in da je

$$\mathrm{ev}_T(fg) = \mathrm{ev}_T(f)\,\mathrm{ev}_T(g)$$

za poljubna polinoma $f, g \in \mathbb{F}[t]$. Vrednost produkta polinomov v endomorfizmu T je torej enaka kompoziciji vrednosti teh dveh polinomov v endomorfizmu T.

Podobno lahko vrednost polinoma f definiramo tudi v poljubni kvadratni matriki $A \in \mathrm{Mat}(n \times n, \mathbb{F})$, kot matriko

$$f(A) = a_0\mathrm{I} + a_1 A + a_2 A^2 + \cdots + a_s A^s \in \mathrm{Mat}(n \times n, \mathbb{F}).$$

Opazimo, da za poljubno matriko $P \in \mathrm{GL}(n, \mathbb{F})$ velja

$$f(PAP^{-1}) = Pf(A)P^{-1}$$

in da za poljubno bazo $\mathscr{B}$ vektorskega prostora V velja

$$[f(T)]^{\mathscr{B}}_{\mathscr{B}} = f\big([T]^{\mathscr{B}}_{\mathscr{B}}\big).$$

Izrek 2.107 (Cayley-Hamiltonov izrek). *Za vsak endomorfizem T končnodimenzionalnega vektorskega prostora V pozitivne dimenzije velja*

$$\mathrm{p}_T(T) = 0.$$

Posebej za vsako kvadratno matriko A velja $\mathrm{p}_A(A) = 0$.

Dokaz. Naj bo $n = \dim \mathsf{V}$, naj bo $\mathscr{B}$ poljubna baza vektorskega prostora V in naj bo $A = [T]^{\mathscr{B}}_{\mathscr{B}}$. Ker je $[\mathrm{p}_T(T)]^{\mathscr{B}}_{\mathscr{B}} = \mathrm{p}_T([T]^{\mathscr{B}}_{\mathscr{B}}) = \mathrm{p}_A(A)$, je dovolj dokazati, da je $\mathrm{p}_A(A) = 0$.

Po Schurovem izreku obstajata takšna matrika $P \in \mathrm{GL}(n, \mathbb{C})$ in takšna zgornje trikotna matrika matrika $B \in \mathrm{Mat}(n \times n, \mathbb{C})$, da je $A = PBP^{-1}$. Ker je $\mathrm{p}_A(A) = \mathrm{p}_B(PBP^{-1}) = P\,\mathrm{p}_B(B)P^{-1}$, je dovolj dokazati, da je $\mathrm{p}_B(B) = 0$.

Ker je matrika B zgornje trikotna, za vsak $j = 1, 2, \ldots, n$ velja

$$B\mathrm{e}_j = B_{1j}\mathrm{e}_1 + B_{2j}\mathrm{e}_2 + \cdots + B_{(j-1)j}\mathrm{e}_{j-1} + B_{jj}\mathrm{e}_j \tag{2.1}$$

oziroma

$$(B - B_{jj}\mathrm{I})\mathrm{e}_j = B_{1j}\mathrm{e}_1 + B_{2j}\mathrm{e}_2 + \cdots + B_{(j-1)j}\mathrm{e}_{j-1}. \tag{2.2}$$

Označimo $\mathrm{U}_j = \mathrm{Span}_{\mathbb{C}^n}\{\mathrm{e}_1, \mathrm{e}_2, \ldots, \mathrm{e}_j\}$ in posebej $\mathsf{U}_0 = \{0\}$, torej da imamo

$$\{0\} = \mathsf{U}_0 \subset \mathsf{U}_1 \subset \cdots \subset \mathsf{U}_{n-1} \subset \mathsf{U}_n = \mathbb{C}^n.$$

Iz enakosti (2.1) sledi

$$\{Bu\ ;\ u \in \mathsf{U}_j\} \subset \mathsf{U}_j \tag{2.3}$$

podobno pa zaradi enakosti (2.1) in (2.2) velja

$$\{(B - B_{jj}\mathrm{I})u\ ;\ u \in \mathsf{U}_j\} \subset \mathsf{U}_{j-1}, \tag{2.4}$$

za vsak $j = 1, 2, \ldots, n$.

Ker je matrika B zgornje-trikotna, je $\mathrm{p}_B(t) = (B_{11} - t)(B_{22} - t)\cdots(B_{nn} - t)$, zato imamo

$$\mathrm{p}_B(B) = (-1)^n(B - B_{11}\mathrm{I})(B - B_{22}\mathrm{I})\cdots(B - B_{nn}\mathrm{I}).$$

Z indukcijo bomo dokazali, da velja

$$(B - B_{11}\mathrm{I})(B - B_{22}\mathrm{I})\cdots(B - B_{jj}\mathrm{I})u = 0 \tag{2.5}$$

za vsak $u \in \mathsf{U}_j$ in za vsak $j = 1, 2, \ldots, n$. Če je $j = 1$, je enakost (2.5) direktna posledica enakosti (2.4). Predpostavimo zdaj, da je $1 < j \leq n$ in da je

$$(B - B_{11}\mathrm{I}) \cdots (B - B_{(j-1)(j-1)}\mathrm{I})w = 0$$

za vse $w \in \mathsf{U}_{j-1}$. Iz enakosti (2.4) za vsak $u \in \mathsf{U}_j$ sledi, da vektor $v = (B - B_{jj}\mathrm{I})u$ leži v podprostoru U_{j-1}, po indukcijski hipotezi pa je zato

$$(B - B_{11}\mathrm{I}) \cdots (B - B_{jj}\mathrm{I})u = (B - B_{11}\mathrm{I}) \cdots (B - B_{(j-1)(j-1)}\mathrm{I})v = 0.$$

Enakost (2.5) velja torej za vse $j = 1, 2, \ldots, n$. Posebej nam pri $j = n$ ta enakost pove, da je $\mathrm{p}_B(B) = 0$. $\square$

ZGLED 2.108. Naj bo T poljuben endomorfizem končno-dimenzionalnega kompleksnega vektorskega prostora V dimenzije $n = \dim \mathsf{V}$, in naj bo

$$\mathrm{p}_T(t) = (-t)^n + \mathrm{tr}(T)(-t)^{n-1} + c_{n-2}(-t)^{n-2} + \cdots + c_1(-t)^1 + \det(T)$$

karakteristični polinom endomorfizma T. Cayley-Hamiltonov izrek nam pove, da je $\mathrm{p}_T(T) = 0$ oziroma da velja

$$(-T)^n + \mathrm{tr}(T)(-T)^{n-1} + c_{n-2}(-T)^{n-2} + \cdots + c_1(-T)^1 + \det(T)\,\mathrm{id}_{\mathsf{V}} = 0.$$

Odtod sledi

$$\begin{aligned}\det(T)\,\mathrm{id}_{\mathsf{V}} &= -(-T)^n - \mathrm{tr}(T)(-T)^{n-1} - c_{n-2}(-T)^{n-2} - \cdots - c_1(-T)^1\\ &= T\big((-T)^{n-1} + \mathrm{tr}(T)(-T)^{n-2} + c_{n-2}(-T)^{n-3} + \cdots + c_1T^0\big).\end{aligned}$$

Če je T avtomorfizem, smo odtod dobili novo formulo za izračun inverza, namreč

$$T^{-1} = \frac{1}{\det(T)}\Big((-T)^{n-1} + \mathrm{tr}(T)(-T)^{n-2} + c_{n-2}(-T)^{n-3} + \cdots + c_1T^0\Big).$$

Vzemimo na primer kvadratno matriko

$$A = \begin{bmatrix} 1 & 1 & 0\\ 0 & 1 & 1\\ 0 & 0 & 1\end{bmatrix}$$

s karakterističnim polinomom

$$\mathrm{p}_A(t) = (1-t)^3 = -t^3 + 3t^2 - 3t + 1.$$

Matrika A ima determinanto 1 in je obrnljiva. Izračunajmo še potence matrike A:

$$A^2 = \begin{bmatrix} 1 & 2 & 1\\ 0 & 1 & 2\\ 0 & 0 & 1\end{bmatrix} \qquad A^3 = \begin{bmatrix} 1 & 3 & 3\\ 0 & 1 & 3\\ 0 & 0 & 1\end{bmatrix}$$

Zdaj lahko z direktnim izračunom preverimo, da je res

$$\mathrm{p}_A(A) = -A^3 + 3A^2 - 3A + \mathrm{I} = 0,$$

po zgoraj izpeljani formuli pa lahko izračunamo tudi inverz matrike A:

$$\begin{aligned}A^{-1} = A^2 - 3A + 3\mathrm{I} &= \begin{bmatrix} 1 & 2 & 1\\ 0 & 1 & 2\\ 0 & 0 & 1\end{bmatrix} - 3\begin{bmatrix} 1 & 1 & 0\\ 0 & 1 & 1\\ 0 & 0 & 1\end{bmatrix} + 3\begin{bmatrix} 1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}\\ &= \begin{bmatrix} 1 & -1 & 1\\ 0 & 1 & -1\\ 0 & 0 & 1\end{bmatrix}\end{aligned}$$

2.2.5. Funkcije diagonalizabilnih endomorfizmov. Za poljuben polinom $f \in \mathbb{F}[t]$ in za vsako diagonalno matriko $D = \operatorname{diag}(\eta_1, \eta_2, \ldots, \eta_n) \in \operatorname{Mat}(n \times n, \mathbb{F})$ je $D^p = \operatorname{diag}(\eta_1^p, \eta_2^p, \ldots, \eta_n^p)$, za vsak $p \in \mathbb{N} \cup \{0\}$, in zato tudi

$$f(D) = \operatorname{diag}(f(\eta_1), f(\eta_2), \ldots, f(\eta_n)).$$

Če je $A \in \operatorname{Mat}(n \times n, \mathbb{F})$ matrika, ki je podobna matriki D, potem velja $A = PDP^{-1}$ za neko obrnljivo matriko $P \in \operatorname{GL}(n, \mathbb{F})$. V tem primeru dobimo

$$f(A) = Pf(D)P^{-1} = P \operatorname{diag}(f(\eta_1), f(\eta_2), \ldots, f(\eta_n))P^{-1}.$$

Spoznali smo torej, da je vrednost $f(A)$ enostavno izračunati, če znamo diagonalizirati matriko A. Podobno lahko naredimo za vsak diagonalizabilen endomorfizem $T \in \operatorname{End}(\mathsf{V})$. Za tak endomorfizem izberimo bazo $\mathscr{B}$, da je matrika $[T]_{\mathscr{B}}^{\mathscr{B}}$ diagonalna, recimo $[T]_{\mathscr{B}}^{\mathscr{B}} = \operatorname{diag}(\eta_1, \eta_2, \ldots, \eta_n)$. Tedaj je

$$[f(T)]_{\mathscr{B}}^{\mathscr{B}} = f\big([(T)]_{\mathscr{B}}^{\mathscr{B}}\big) = \operatorname{diag}(f(\eta_1), f(\eta_2), \ldots, f(\eta_n))$$

oziroma

$$f(T) = [\,\bullet\,]_{\mathscr{B}}^{-1} \circ \ell\big(\operatorname{diag}(f(\eta_1), f(\eta_2), \ldots, f(\eta_n))\big) \circ [\,\bullet\,]_{\mathscr{B}}.$$

Posebej velja $\sigma(f(T)) = f(\sigma(T))$.

Pravkar povedano lahko nekoliko posplošimo. Recimo, da je $T \in \operatorname{End}(\mathsf{V})$ poljuben endomorfizem in da so $\mathsf{V}_1, \mathsf{V}_2, \ldots, \mathsf{V}_k$ takšni T-invariantni vektorski podprostori vektorskega prostora V, da je

$$\mathsf{V} = \mathsf{V}_1 \oplus \mathsf{V}_2 \oplus \cdots \oplus \mathsf{V}_k.$$

Tedaj je

$$T = T_1 \oplus T_2 \oplus \cdots \oplus T_k,$$

kjer je T_i zožitev $T|_{\mathsf{V}_i} : \mathsf{V}_i \to \mathsf{V}_i$ za vsak $i = 1, 2, \ldots, k$. Ker je tedaj

$$T^p = T_1^p \oplus T_2^p \oplus \cdots \oplus T_k^p$$

za vsak p, velja tudi

$$f(T) = f(T_1) \oplus f(T_2) \oplus \cdots \oplus f(T_k).$$

Predpostavimo zdaj, da je endomorfizem T diagonalizabilen. Tedaj je $\mathsf{V} = \mathrm{E}_T(\lambda_1) \oplus \mathrm{E}_T(\lambda_2) \oplus \cdots \oplus \mathrm{E}_T(\lambda_m)$, kjer so $\lambda_1, \lambda_2, \ldots, \lambda_m \in \mathbb{F}$ natanko vse paroma različne lastne vrednosti endomorfizma T. Ker je tedaj

$$T = \lambda_1 \operatorname{id}_{\mathrm{E}_T(\lambda_1)} \oplus \lambda_2 \operatorname{id}_{\mathrm{E}_T(\lambda_2)} \oplus \cdots \oplus \lambda_m \operatorname{id}_{\mathrm{E}_T(\lambda_m)}$$

in ker je očitno $f(\lambda_i \operatorname{id}_{\mathrm{E}_T(\lambda_i)}) = f(\lambda_i) \operatorname{id}_{\mathrm{E}_T(\lambda_i)}$, je torej

$$f(T) = f(\lambda_1) \operatorname{id}_{\mathrm{E}_T(\lambda_1)} \oplus f(\lambda_2) \operatorname{id}_{\mathrm{E}_T(\lambda_2)} \oplus \cdots \oplus f(\lambda_m) \operatorname{id}_{\mathrm{E}_T(\lambda_m)}.$$

Opazimo lahko, da je zadnja formula smiselna pravzaprav za poljubno funkcijo $f : \sigma(T) \to \mathbb{F}$. To nam je motivacija za naslednjo definicijo:

Definicija 2.109. Naj bo $T \in \operatorname{End}(\mathsf{V})$ poljuben diagonalizabilen endomorfizem in naj bodo $\lambda_1, \lambda_2, \ldots, \lambda_m \in \mathbb{F}$ natanko vse paroma različne lastne vrednosti endomorfizma T. Za poljubno funkcijo $f : \sigma(T) \to \mathbb{F}$ definiramo *vrednost* $f(T) \in \operatorname{End}(\mathsf{V})$ funkcije f v endomorfizmu T s predpisom

$$f(T) = f(\lambda_1) \operatorname{id}_{\mathrm{E}_T(\lambda_1)} \oplus f(\lambda_2) \operatorname{id}_{\mathrm{E}_T(\lambda_2)} \oplus \cdots \oplus f(\lambda_m) \operatorname{id}_{\mathrm{E}_T(\lambda_m)}.$$

KOMENTAR 2.110. (1) Ker je direktna vsota komutativna, je definicija vrednosti $f(T)$ neodvisna od izbranega vrstnega reda lastnih vrednosti $\lambda_1, \lambda_2, \dots, \lambda_m$.

(2) Če je f polinom, je vrednost $f(T)$ po zadnji definiciji enaka vednosti polinoma f v endomorfizmu T, ki smo jo spoznali v prejšnjem razdelku.

(3) Za poljubno diagonalizabilno matriko $A \in \mathrm{Mat}(n \times n, \mathbb{F})$ in za vsako funkcijo $f : \sigma(A) \to \mathbb{F}$ podobno definiramo vrednost funkcije f v matriki A kot matriko

$$f(A) = Pf(D)P^{-1} = P\,\mathrm{diag}(f(\eta_1), f(\eta_2), \dots, f(\eta_n))P^{-1},$$

kjer je $P \in \mathrm{GL}(n, \mathbb{F})$ takšna obrnljiva matrika in $D = \mathrm{diag}(\eta_1, \eta_2, \dots, \eta_n)$ takšna diagonalna matrika, da velja $D = P^{-1}AP$. Opazimo lahko, da je $f(A)$ matrika endomorfizma $f(\ell(A))$ v standardni bazi, torej $f(\ell(A)) = \ell(f(A))$, zato je definicija neodvisna od izbire prehodne matrike P.

(4) Za poljubno funkcijo $f : \sigma(T) \to \mathbb{F}$ vrednost $f(T)$ izračunamo na naslednji način: Izberemo bazo $\mathscr{B}$, sestavljeno iz lastnih vektorjev endomorfizma T. Matrika $D = [T]_{\mathscr{B}}^{\mathscr{B}}$ je torej diagonalna, $D = \mathrm{diag}(\eta_1, \eta_2, \dots, \eta_n)$, in $T = [\,\bullet\,]_{\mathscr{B}}^{-1} \circ \ell(D) \circ [\,\bullet\,]_{\mathscr{B}}$. Tedaj je

$$f(T) = [\,\bullet\,]_{\mathscr{B}}^{-1} \circ f(\ell(D)) \circ [\,\bullet\,]_{\mathscr{B}} = [\,\bullet\,]_{\mathscr{B}}^{-1} \circ \ell\big(\,\mathrm{diag}(f(\eta_1), f(\eta_2), \dots, f(\eta_n))\big) \circ [\,\bullet\,]_{\mathscr{B}}.$$

Matrika

$$[f(T)]_{\mathscr{B}}^{\mathscr{B}} = \mathrm{diag}(f(\eta_1), f(\eta_2), \dots, f(\eta_n))$$

je torej diagonalna, odtod pa sledi, da je tudi endomorfizem $f(T)$ diagonalizabilen.

(5) Direktno iz definicije sledi, da za poljubni funkciji $f, g : \sigma(T) \to \mathbb{F}$, za vsako funkcijo $h : \sigma(f(T)) \to \mathbb{F}$, za vsak skalar $\alpha \in \mathbb{F}$ in za vsako bazo $\mathscr{B}$ vektorskega prostora V velja:

(i) $\sigma(f(T)) = f(\sigma(T))$,
(ii) $(f + g)(T) = f(T) + g(T)$,
(iii) $(\alpha f)(T) = \alpha f(T)$,
(iv) $(fg)(T) = f(T)g(T) = g(T)f(T)$,
(v) $(h \circ f)(T) = h(f(T))$,
(vi) $\mathrm{id}_{\mathbb{F}}(T) = T$,
(vii) $1(T) = \mathrm{id}_{\mathsf{V}}$ in
(viii) $[f(T)]_{\mathscr{B}}^{\mathscr{B}} = f\big([T]_{\mathscr{B}}^{\mathscr{B}}\big)$.

V točki (vii) smo z 1 označili konstantno funkcijo $\mathbb{F} \to \mathbb{F}$ z vrednostjo $1 \in \mathbb{F}$.

ZGLED 2.111. Karakteristični polinom realne matrike

$$A = \begin{bmatrix} 2 & 0 \\ -2 & 1 \end{bmatrix}$$

je enak polinomu $\mathrm{p}_A(t) = (2 - t)(1 - t)$, zato ima matrika A dve različni lastni vrednosti in je diagonalizabilna. Lastna vektorja $v_1 = (1, -2)$ in $v_2 = (0, 1)$ matrike A sestavljata bazo $\mathscr{B}$ vektorskega prostora $\mathbb{R}^2$. Zapišemo lahko prehodno matriko P iz baze $\mathscr{B}$ na standardno bazo in njej inverzno matriko:

$$P = \begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix} \qquad P^{-1} = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}$$

Diagonalna matrika

$$D = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$$

je podobna matriki A in zanjo velja $A = PDP^{-1}$. Zdaj lahko izračunamo:

$$A^{42} = (PDP^{-1})^{42} = PD^{42}P^{-1} = P \begin{bmatrix} 2^{42} & 0 \\ 0 & 1 \end{bmatrix} P^{-1} = \begin{bmatrix} 2^{42} & 0 \\ 2(1-2^{42}) & 1 \end{bmatrix}$$

$$A^{1/2} = (PDP^{-1})^{1/2} = PD^{1/2}P^{-1} = P \begin{bmatrix} \sqrt{2} & 0 \\ 0 & 1 \end{bmatrix} P^{-1} = \begin{bmatrix} \sqrt{2} & 0 \\ 2(1-\sqrt{2}) & 1 \end{bmatrix}$$

$$(A^{1/2})^2 = \begin{bmatrix} \sqrt{2} & 0 \\ 2(1-\sqrt{2}) & 1 \end{bmatrix} \begin{bmatrix} \sqrt{2} & 0 \\ 2(1-\sqrt{2}) & 1 \end{bmatrix} = A$$

$$\mathrm{e}^A = \mathrm{e}^{PDP^{-1}} = P\mathrm{e}^D P^{-1} = P \begin{bmatrix} \mathrm{e}^2 & 0 \\ 0 & \mathrm{e} \end{bmatrix} P^{-1} = \begin{bmatrix} \mathrm{e}^2 & 0 \\ 2\mathrm{e}(1-\mathrm{e}) & \mathrm{e} \end{bmatrix}$$

Komentar 2.112. V zadnjem zgledu smo s pomočjo diagonalizacije matrike izračunali vrednost e^A eksponentne funkcije z osnovo e v matriki A. Matriko e^A pa lahko splošneje definiramo za poljubno kvadratno matriko $A \in \mathrm{Mat}(n \times n, \mathbb{C})$ kot vsoto vrste matrik

$$\mathrm{e}^A = \sum_{k=0}^{\infty} \frac{1}{k!} A^k.$$

Izkaže se, da ta vrsta po komponentah konvergira in da je njena vsota e^A obrnljiva matrika. Če je matrika A diagonalizabilna, je vsota te vrste enaka vrednosti eksponentne funkcije z osnovo e v matriki A, izračunani na prej opisan način s pomočjo diagonalizacije. Preslikava

$$\exp : \mathrm{Mat}(n \times n, \mathbb{C}) \to \mathrm{GL}(n, \mathbb{C}), \qquad A \mapsto \exp(A) = \mathrm{e}^A,$$

se imenuje *eksponentna preslikava* splošne linearne grupe $\mathrm{GL}(n, \mathbb{C})$. Če matriki $A, B \in \mathrm{Mat}(n \times n, \mathbb{C})$ med seboj komutirata (torej, če je $AB = BA$), potem velja

$$\mathrm{e}^A \mathrm{e}^B = \mathrm{e}^{A+B}.$$

POGLAVJE 3

Vektorski prostori s skalarnim produktom

V tem poglavju bomo spoznali osnovne lastnosti vektorskih prostorov s skalarnim produktom in linearnih preslikav med takšnimi prostori. Skalarni produkt na poljubnem vektorskem prostoru je operacija, ki ima podobne lastnosti kot običajen standardni skalarni produkt na vektorskem prostoru $\mathbb{R}^3$. Takšnih skalarnih produktov na poljubnem netrivialnem vektorskem prostoru V je neskončno, izbira enega od njih pa nam omogoča, da lahko izračunamo na primer dolžino vektorja ali pa razdaljo med dvema vektorjema *glede na* ta izbrani skalarni produkt. Tudi v tem poglavju bomo z $\mathbb{F}$ označili izbran obseg, ki je bodisi obseg realnih števil $\mathbb{R}$ ali pa obseg kompleksnih števil $\mathbb{C}$.

3.1. Skalarni produkt

DEFINICIJA 3.1. Naj bo V vektorski prostor nad $\mathbb{F}$. Preslikavi

$$\mathsf{V} \times \mathsf{V} \to \mathbb{F}, \qquad (v, w) \mapsto \langle v, w \rangle,$$

pravimo *skalarni produkt* na vektorskem prostoru V, če za vse vektorje $u, v, w \in \mathsf{V}$ in za vsak skalar $\alpha \in \mathbb{F}$ velja

(i) $\langle v, w \rangle = \overline{\langle w, v \rangle}$,
(ii) $\langle u + v, w \rangle = \langle u, w \rangle + \langle v, w \rangle$,
(iii) $\langle \alpha v, w \rangle = \alpha \langle v, w \rangle$,
(iv) $\langle v, v \rangle \geq 0$ in
(v) če je $\langle v, v \rangle = 0$, potem je $v = 0$.

KOMENTAR 3.2. (1) V točki (i) iz definicije nastopa konjugirana vrednosti skalarnega produkta $\langle v, w \rangle$, zaradi te lastnosti pa pravimo, da je skalarni produkt *konjugirano simetričen.* Iz točke (i) sledi, da je skalarni produkt $\langle v, v \rangle$ realno število. Če je $\mathbb{F} = \mathbb{R}$, potem točka (i) pove, da je $\langle v, w \rangle = \langle w, v \rangle$, zato je skalarni produkt na realnem vektorskem prostoru *simetričen* oziroma komutativen.

Zaradi lastnosti (ii) in (iii) iz definicije pravimo, da je skalarni produkt *linearen v prvem faktorju.* Zaradi lastnosti (iv) in (v) je skalarni produkt *pozitivno definiten.* Iz točk (i-iii) sledi, da velja tudi

(a) $\langle v, u + w \rangle = \langle v, u \rangle + \langle v, w \rangle$ in
(b) $\langle v, \beta w \rangle = \bar{\beta} \langle v, w \rangle$

za vse vektorje $u, v, w \in \mathsf{V}$ in za vsak skalar $\beta \in \mathbb{F}$. Zaradi teh dveh lastnosti pravimo, da je skalarni produkt *konjugirano linearen v drugem faktorju.* Iz linearnosti v prvem faktorju in iz konjugirane linearnosti v drugem faktorju sledi, da je $\langle 0, w \rangle = \langle v, 0 \rangle = 0$ za vse vektorje $v, w \in \mathsf{V}$.

Vektorski prostor nad $\mathbb{F}$ *s skalarnim produktom* je vektorski prostor V nad $\mathbb{F}$, opremljen z nekim izbranim skalarnim produktom $\mathsf{V} \times \mathsf{V} \to \mathbb{F}$, $(v, w) \mapsto \langle v, w \rangle$.

Naj bo V vektorski prostor s skalarnim produktom. Za vsak vektor $v \in \mathsf{V}$ število

$$\|v\| = \langle v, v\rangle^{1/2} \geq 0$$

imenujemo *norma* (ali *dolžina*) vektorja v. Iz naštetih lastnosti skalarnega produkta očitno sledi, da je $\|v\| \geq 0$, da je ničelni vektor edini vektor z normo 0 in da velja $\|\alpha v\| = |\alpha|\|v\|$, za vsak vektor $v \in \mathsf{V}$ in za vsak skalar $\alpha \in \mathbb{F}$. Preslikavi $\mathsf{V} \to \mathbb{R}$, $v \mapsto \|v\|$, pravimo *norma inducirana s skalarnim produktom* na V. *Razdalja* med vektorjema $v, w \in \mathsf{V}$ je norma njune razlike $d(v, w) = \|v - w\|$.

(2) Naj bo V vektorski prostor nad $\mathbb{F}$ s skalarnim produktom. Po analogiji z vektorskim prostorom $\mathbb{R}^3$ definiramo, da je vektor $v \in \mathsf{V}$ *pravokoten* (ali *ortogonalen*) na vektor $w \in \mathsf{V}$, če je $\langle v, w\rangle = 0$. V tem primeru označimo $v \perp w$. Relacija pravokotnosti je simetrična, saj iz enakosti $\langle v, w\rangle = 0$ sledi tudi $\langle w, v\rangle = \overline{\langle v, w\rangle} = 0$. Ničelni vektor $0 \in \mathsf{V}$ je pravokoten na vse vektorje iz vektorskega prostora V. Zaradi pozitivne definitnosti skalarnega produkta je ničelni vektor edini vektor iz vektorskega prostora V, ki je pravokoten sam nase.

Naj bosta $\mathcal{O}$ in $\mathcal{O}'$ dve podmnožici vektorskega prostora V. Če je vsak vektor iz množice $\mathcal{O}$ pravokoten na vsak vektor iz množice $\mathcal{O}'$, potem pravimo, da sta množici *med seboj pravokotni*, in označimo $\mathcal{O} \perp \mathcal{O}'$.

Pravokotna projekcija vektorja $v \in \mathsf{V}$ na neničelni vektor $w \in \mathsf{V}$ je vektor

$$\mathrm{pr}_w^{\perp}(v) = \frac{\langle v, w\rangle}{\langle w, w\rangle} w.$$

Pravokotna projekcija $\mathrm{pr}_w^{\perp}(v)$ je natanko tisti skalarni večkratnik vektorja w, za katerega je vektor $v - \mathrm{pr}_w^{\perp}(v)$ pravokoten na vektor w, kar nam pokaže izračun

$$\begin{aligned}\langle v - \mathrm{pr}_w^{\perp}(v), w\rangle &= \left\langle v - \frac{\langle v, w\rangle}{\langle w, w\rangle} w, w \right\rangle = \langle v, w\rangle + \left\langle -\frac{\langle v, w\rangle}{\langle w, w\rangle} w, w \right\rangle \\ &= \langle v, w\rangle - \frac{\langle v, w\rangle}{\langle w, w\rangle}\langle w, w\rangle = \langle v, w\rangle - \langle v, w\rangle = 0.\end{aligned}$$

(3) Naj bo U vektorski podprostor vektorskega prostora V s skalarnim produktom. Zožitev skalarnega produkta $\mathsf{V} \times \mathsf{V} \to \mathbb{F}$ na podmnožico $\mathsf{U} \times \mathsf{U} \subset \mathsf{V} \times \mathsf{V}$ je skalarni produkt na vektorskem prostoru U, zato je glede na to zožitev tudi U vektorski prostor s skalarnim produktom.

VAJA 3.3. Naj bo $\mathcal{O}$ podmnožica vektorskega prostora V s skalarnim produktom in naj bo vektor $v \in \mathsf{V}$ pravokoten na vsak vektor iz podmnožice $\mathcal{O}$. Pokaži, da je tedaj vektor v pravokoten na vsak vektor iz vektorskega podprostora $\mathrm{Span}(\mathcal{O})$.

ZGLED 3.4. (1) Na vektorskem prostoru $\mathbb{R}^n$ imamo *standardni skalarni produkt*, ki je za poljubna vektorja $v = (\alpha_1, \alpha_2, \ldots, \alpha_n) \in \mathbb{R}^n$ in $w = (\beta_1, \beta_2, \ldots, \beta_n) \in \mathbb{R}^n$ dan s predpisom

$$\langle v, w\rangle = v \cdot w = \alpha_1\beta_1 + \alpha_2\beta_2 + \cdots + \alpha_n\beta_n = v^{\mathsf{t}} w = w^{\mathsf{t}} v.$$

Norma, inducirana s tem skalarnim produktom, je standardna Evklidska norma

$$\|v\| = |v| = \sqrt{\alpha_1^2 + \alpha_2^2 + \cdots + \alpha_n^2}.$$

Praviloma bomo vedno privzeli, da je vektorski prostor $\mathbb{R}^n$ opremljen s standardnim skalarnim produktom, razen če ne bomo posebej poudarili nasprotno.

(2) Tudi na vektorskem prostoru $\mathbb{C}^n$ imamo *standardni skalarni produkt*, ki je za poljubna vektorja $v = (\alpha_1, \alpha_2, \ldots, \alpha_n) \in \mathbb{C}^n$ in $w = (\beta_1, \beta_2, \ldots, \beta_n) \in \mathbb{C}^n$ podan s predpisom

$$\langle v, w \rangle = v \cdot w = \alpha_1 \bar{\beta}_1 + \alpha_2 \bar{\beta}_2 + \cdots + \alpha_n \bar{\beta}_n = v^{\mathsf{t}} \bar{w} = w^{\mathsf{h}} v.$$

Norma vektorja v glede na ta skalarni produkt je enaka številu

$$\|v\| = |v| = \sqrt{|\alpha_1|^2 + |\alpha_2|^2 + \cdots + |\alpha_n|^2}.$$

Praviloma bomo vedno privzeli, da je vektorski prostor $\mathbb{C}^n$ opremljen s standardnim skalarnim produktom, razen če ne bomo posebej poudarili nasprotno.

(3) Naj bo $[\![a, b]\!] \subset \mathbb{R}$ zaprt interval od a do b, $a < b$, in naj bo $\rho : [\![a, b]\!] \to \mathbb{R}$ povsod pozitivna zvezna realna funkcija. Za poljubni realni zvezni funkciji $f, g : [\![a, b]\!] \to \mathbb{R}$ naj bo

$$\langle f, g \rangle = \int_a^b f(t) g(t) \rho(t)\, dt.$$

Ni težko preveriti, da smo s tem definirali skalarni produkt na realnem vektorskem prostoru $C([\![a, b]\!])$ vseh realnih zveznih funkcij na intervalu $[\![a, b]\!]$. Glede na ta skalarni produkt je norma funkcije f dana s predpisom:

$$\|f\| = \left(\int_a^b f(t)^2 \rho(t)\, dt \right)^{1/2}$$

Na enak način lahko definiramo skalarni produkt na vektorskem prostoru $\mathbb{R}[t]$ vseh realnih polinomov, saj je vsak realen polinom v spremenljivki t tudi realna zvezna funkcija. Pozitivna definitnost tako definiranega skalarnega produkta v tem primeru sledi iz dejstva, da ima vsak neničelen polinom le končno število ničel. Dobljeni skalarni produkt lahko zožimo tudi na vektorski podprostor $\mathbb{R}_s[t] \subset \mathbb{R}[t]$ vseh realnih polinomov stopnje največ s, za vsako nenegativno celo število s.

(4) Naj bo $[\![a, b]\!] \subset \mathbb{R}$ zaprt interval od a do b, $a < b$, in naj bo $C([\![a, b]\!], \mathbb{C})$ kompleksen vektorski prostor vseh kompleksnih zveznih funkcij na intervalu $[\![a, b]\!]$.

Določeni integral kompleksne zvezne funkcije $f \in C([\![a, b]\!], \mathbb{C})$ definiramo kot vsoto integralov

$$\int_a^b f(t)\, dt = \int_a^b \Re(f)(t)\, dt + \mathrm{i} \int_a^b \Im(f)(t)\, dt,$$

kjer sta $\Re(f)$ in $\Im(f)$ realni oziroma imaginarni del funkcije f, torej realni zvezni funkciji na intervalu $[\![a, b]\!]$, za kateri velja $f = \Re(f) + \mathrm{i}\Im(f)$. Tako definirana preslikava $C([\![a, b]\!], \mathbb{C}) \to \mathbb{C}$, $f \mapsto \int_a^b f(t)\, dt$, je linearna preslikava kompleksnih vektorskih prostorov.

Naj bo $\rho : [\![a, b]\!] \to \mathbb{R}$ povsod pozitivna zvezna realna funkcija. Za poljubni funkciji $f, g \in C([\![a, b]\!], \mathbb{C})$ naj bo

$$\langle f, g \rangle = \int_a^b f(t) \overline{g(t)} \rho(t)\, dt.$$

S tem predpisom smo definirali skalarni produkt na kompleksnem vektorskem prostoru $C([\![a, b]\!], \mathbb{C})$, glede na katerega je norma funkcije f dana s predpisom:

$$\|f\| = \left(\int_a^b |f(t)|^2 \rho(t)\, dt \right)^{1/2}$$

Na enak način lahko definiramo skalarni produkt na vektorskem prostoru $\mathbb{C}[t]$ vseh kompleksnih polinomov v spremenljivki t. Dobljeni skalarni produkt lahko zožimo tudi na vektorski podprostor $\mathbb{C}_s[t] \subset \mathbb{C}[t]$ vseh kompleksnih polinomov stopnje največ s, za vsako nenegativno celo število s.

3.1.1. Osnovne lastnosti skalarnega produkta.

TRDITEV 3.5. *Naj bo* V *vektorski prostor nad* $\mathbb{F}$ *s skalarnim produktom. Če za vektorja* $v, w \in \mathsf{V}$ *velja, da je* $\langle u, v\rangle = \langle u, w\rangle$ *za vsak vektor* $u \in \mathsf{V}$*, potem je* $v = w$.

DOKAZ. Za vsak vektor $u \in \mathsf{V}$ velja $\langle u, v\rangle = \langle u, w\rangle$ in zato $\langle u, v - w\rangle = 0$. Posebej lahko vzamemo $u = v - w$ in dobimo $\langle v - w, v - w\rangle = 0$, zaradi pozitivne definitnosti skalarnega produkta pa odtod sledi $v - w = 0$. $\square$

TRDITEV 3.6 (Pitagorov izrek). *Naj bo* V *vektorski prostor nad* $\mathbb{F}$ *s skalarnim produktom. Če sta vektorja* $v, w \in \mathsf{V}$ *med seboj pravokotna, potem je* $\|v + w\|^2 = \|v\|^2 + \|w\|^2$.

DOKAZ. Ker sta vektorja v in w med seboj pravokotna, je $\langle v, w\rangle = \langle w, v\rangle = 0$ in zato

$$\begin{aligned}\|v + w\|^2 &= \langle v + w, v + w\rangle \\ &= \langle v, v\rangle + \langle v, w\rangle + \langle w, v\rangle + \langle w, w\rangle \\ &= \langle v, v\rangle + \langle w, w\rangle \\ &= \|v\|^2 + \|w\|^2.\end{aligned}$$

$\square$

TRDITEV 3.7. *Naj bo* V *vektorski prostor nad* $\mathbb{F}$ *s skalarnim produktom. Za poljubna vektorja* $v, w \in \mathsf{V}$ *velja*

(i) $\|v + w\|^2 + \|v - w\|^2 = 2\|v\|^2 + 2\|w\|^2$, *(paralelogramsko pravilo)*
(ii) $\big|\,\|v\| - \|w\|\,\big| \leq \|v + w\| \leq \|v\| + \|w\|$, *(trikotniška neenakost)*
(iii) $|\langle v, w\rangle| \leq \|v\|\|w\|$. *(Cauchy-Schwarzova neenakost)*

Enakost $|\langle v, w\rangle| = \|v\|\|w\|$ *velja, če, in samo če, sta vektorja* $v, w \in \mathsf{V}$ *med seboj linearno odvisna.*

DOKAZ. Točko (i) dobimo z izračunom

$$\begin{aligned}\|v + w\|^2 + \|v - w\|^2 &= \langle v + w, v + w\rangle + \langle v - w, v - w\rangle \\ &= \langle v, v\rangle + \langle v, w\rangle + \langle w, v\rangle + \langle w, w\rangle + \langle v, v\rangle - \langle v, w\rangle - \langle w, v\rangle + \langle w, w\rangle \\ &= 2\langle v, v\rangle + 2\langle w, w\rangle = 2\|v\|^2 + 2\|w\|^2.\end{aligned}$$

Dokažimo zdaj točko (iii). Naj bosta $v, w \in \mathsf{V}$ poljubna vektorja. Če je $w = 0$, potem neenakost (iii) trivialno velja. Predpostavimo torej lahko, da je vektor w neničelen. Označimo $u = v - \mathrm{pr}_w^{\perp}(v)$. Ker je vektor u pravokoten na vektor $\mathrm{pr}_w^{\perp}(v)$, po Pitagorovem izreku velja

$$\begin{aligned}\|v\|^2 = \|u + \mathrm{pr}_w^{\perp}(v)\|^2 = \|u\|^2 + \|\mathrm{pr}_w^{\perp}(v)\|^2 &= \|u\|^2 + \left\|\frac{\langle v, w\rangle}{\langle w, w\rangle} w\right\|^2 \\ = \|u\|^2 + \left|\frac{\langle v, w\rangle}{\langle w, w\rangle}\right|^2 \|w\|^2 &= \|u\|^2 + \frac{|\langle v, w\rangle|^2}{\|w\|^2}\end{aligned}$$

in zato

$$|\langle v, w\rangle|^2 = \|v\|^2\|w\|^2 - \|u\|^2\|w\|^2 \leq \|v\|^2\|w\|^2,$$

odtod pa sledi neenakost iz točke (iii). Hkrati odtod vidimo, da velja enakost $|\langle v, w\rangle| = \|v\|\|w\|$, če, in samo če, je $u = 0$, kar je res tedaj, ko je vektor v skalarni večkratnik vektorja w.

Iz pravkar dokazane točke (iii) sledi $|\Re(\langle v, w\rangle)| \leq |\langle v, w\rangle| \leq \|v\|\|w\|$, kjer smo z $\Re(\langle v, w\rangle)$ označili realni del števila $\langle v, w\rangle$. Odtod, in iz enakosti

$$\begin{aligned}\|v+w\|^2 &= \langle v+w, v+w\rangle = \langle v, v\rangle + \langle v, w\rangle + \langle w, v\rangle + \langle w, w\rangle \\ &= \langle v, v\rangle + \langle v, w\rangle + \overline{\langle v, w\rangle} + \langle w, w\rangle = \|v\|^2 + 2\Re(\langle v, w\rangle) + \|w\|^2,\end{aligned}$$

zdaj dobimo neenakosti:

$$\begin{aligned}\|v+w\|^2 &\leq \|v\|^2 + 2\|v\|\|w\| + \|w\|^2 = (\|v\| + \|w\|)^2 \\ \|v+w\|^2 &\geq \|v\|^2 - 2\|v\|\|w\| + \|w\|^2 = (\|v\| - \|w\|)^2\end{aligned}$$

Tako smo dokazali tudi točko (ii). □

Z enostavnim izračunom, v katerem uporabimo le osnovne lastnosti iz definicije skalarnega produkta, izpeljemo tudi naslednjo trditev. Rezultat te trditve imenujemo tudi *polarizacijska enačba.* Ta enačba nam pove, da je skalarni produkt enolično določen s sebi pridruženo normo. Pri uporabi te trditve moramo biti nekoliko pozorni, saj ima polarizacijska enačba za skalarni produkt na kompleksnem vektorskem prostoru nekoliko drugačno obliko kot za skalarni produkt na realnem vektorskem prostoru.

TRDITEV 3.8 (Polarizacijska enačba). *Naj bo V vektorski prostor nad $\mathbb{F}$ s skalarnim produktom.*

(i) Če je $\mathbb{F} = \mathbb{R}$, potem za poljubna vektorja $v, w \in \mathsf{V}$ velja

$$\langle v, w\rangle = \tfrac{1}{4}\left(\|v+w\|^2 - \|v-w\|^2\right).$$

(ii) Če je $\mathbb{F} = \mathbb{C}$, potem za poljubna vektorja $v, w \in \mathsf{V}$ velja

$$\langle v, w\rangle = \tfrac{1}{4}\left(\|v+w\|^2 - \|v-w\|^2 + \mathrm{i}\|v+\mathrm{i}w\|^2 - \mathrm{i}\|v-\mathrm{i}w\|^2\right).$$

3.1.2. Ortonormirane baze. Naj bo V vektorski prostor nad $\mathbb{F}$ s skalarnim produktom. Vektor $u \in \mathsf{V}$ je *normiran*, če je $\|u\| = 1$. Poljuben neničelen vektor $v \in \mathsf{V}$ lahko *normiramo* tako, da ga pomnožimo s skalarjem $1/\|v\|$. Tako dobljen vektor

$$u = \frac{1}{\|v\|}v$$

je normiran, poleg tega pa zanj velja $\mathbb{F}u = \mathbb{F}v$.

Vektorji $v_1, v_2, \ldots, v_k \in \mathsf{V}$ sestavljajo *ortonormiran sistem* vektorskega prostora V, če velja

$$\langle v_i, v_j\rangle = \begin{cases} 1 & ; \quad i = j \\ 0 & ; \quad i \neq j \end{cases}$$

za vse $i, j = 1, 2, \ldots, k$. To je res torej natanko tedaj, ko so vektorji $v_1, v_2, \ldots, v_k$ vsi normirani in med seboj paroma pravokotni.

Če vektorji $v_1, v_2, \ldots, v_n \in \mathsf{V}$ sestavljajo ortonormiran sistem in hkrati sestavljajo bazo vektorskega prostora V, potem pravimo, da sestavljajo *kompleten ortonormiran sistem* oziroma *ortonormirano bazo* vektorskega prostora V. V tem primeru je baza $\mathscr{B} = [v_1, v_2, \ldots, v_n]$ vektorskega prostora V *ortonormirana.*

ZGLED 3.9. Glede na standardni skalarni produkt je standardna baza $\mathscr{E} = [\mathrm{e}_1, \mathrm{e}_2, \ldots, \mathrm{e}_n]$ vektorskega prostora $\mathbb{F}^n$ je ortonormirana.

TRDITEV 3.10. *Naj bo* V *končno-dimenzionalen vektorski prostor nad* $\mathbb{F}$ *pozitivne dimenzije in naj bo* $\mathscr{B}$ *baza vektorskega prostora* V*. Tedaj obstaja natanko en skalarni produkt na vektorskem prostoru* V*, glede na katerega je baza* $\mathscr{B}$ *ortonormirana.*

DOKAZ. Naj bo $\mathscr{B} = [v_1, v_2, \dots, v_n]$. Če je baza $\mathscr{B}$ ortonormirana glede na nek skalarni produkt $\mathsf{V} \times \mathsf{V} \to \mathbb{F}$, $(v, w) \mapsto \langle v, w \rangle$, potem za ta skalarni produkt velja

$$\begin{aligned}\langle v, w\rangle &= \left\langle \sum_{i=1}^{n} \big([v]_{\mathscr{B}}\big)_{i1} v_i, \sum_{j=1}^{n} \big([w]_{\mathscr{B}}\big)_{j1} v_j \right\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n} \big([v]_{\mathscr{B}}\big)_{i1} \overline{\big([w]_{\mathscr{B}}\big)_{j1}} \langle v_i, v_j\rangle \\ &= \sum_{i=1}^{n} \big([v]_{\mathscr{B}}\big)_{i1} \overline{\big([w]_{\mathscr{B}}\big)_{i1}}.\end{aligned}$$

S tem je dokazana enoličnost skalarnega produkta, glede na katerega je baza $\mathscr{B}$ ortonormirana. Po drugi strani pa lahko zadnjo enakost uporabimo za definicijo preslikave $\mathsf{V} \times \mathsf{V} \to \mathbb{F}$, za katero direktno preverimo, da je skalarni produkt na vektorskem prostoru V z želenimi lastnostmi. $\square$

TRDITEV 3.11. *Naj bo* V *vektorski prostor nad* $\mathbb{F}$ *s skalarnim produktom. Če vektorji* $v_1, v_2, \dots, v_k \in \mathsf{V}$ *sestavljajo ortonormiran sistem, potem so med seboj linearno neodvisni in za vsak vektor* $v \in \operatorname{Span}\{v_1, v_2, \dots, v_k\}$ *velja*

$$v = \sum_{i=1}^{k} \langle v, v_i\rangle v_i$$

in

$$\|v\|^2 = \sum_{i=1}^{k} |\langle v, v_i\rangle|^2.$$

DOKAZ. Naj bodo $\alpha_1, \alpha_2, \dots, \alpha_k \in \mathbb{F}$ takšni skalarji, da je $v = \sum_{i=1}^{k} \alpha_i v_i$. Ker vektorji $v_1, v_2, \dots, v_k$ sestavljajo ortonormiran sistem, je

$$\begin{aligned}\langle v, v_i\rangle &= \langle \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k, v_i\rangle \\ &= \alpha_1\langle v_1, v_i\rangle + \alpha_2\langle v_2, v_i\rangle + \cdots + \alpha_k\langle v_k, v_i\rangle = \alpha_i\end{aligned}$$

za vsak $i = 1, 2, \dots, k$. Če je $v = 0$, potem sledi $\alpha_1 = \alpha_2 = \cdots = \alpha_k = 0$. Poleg tega velja $\|v\|^2 = \big\langle \sum_{i=1}^{k} \langle v, v_i\rangle v_i, \sum_{j=1}^{k}\langle v, v_j\rangle v_j \big\rangle = \sum_{i=1}^{k} |\langle v, v_i\rangle|^2$. $\square$

Iz zadnje trditve in iz trditve 2.50(iv) sledi:

POSLEDICA 3.12. *Naj bo* V *končno-dimenzionalen vektorski prostor nad* $\mathbb{F}$ *s skalarnim produktom dimenzije* $n \geq 1$*. Če vektorji* $v_1, v_2, \dots, v_n \in \mathsf{V}$ *sestavljajo ortonormiran sistem, potem sestavljajo ortonormirano bazo vektorskega prostora* V *in za vsak vektor* $v \in \mathsf{V}$ *velja* $v = \sum_{i=1}^{n} \langle v, v_i\rangle v_i$ *in* $\|v\|^2 = \sum_{i=1}^{n} |\langle v, v_i\rangle|^2$.

KOMENTAR 3.13. Če je torej $\mathscr{B} = [v_1, v_2, \dots, v_n]$ ortonormirana baza vektorskega prostora V s skalarnim produktom, potem za vsak vektor $v \in \mathsf{V}$ velja $([v]_{\mathscr{B}})_{i1} = \langle v, v_i\rangle$, za vse $i = 1, 2, \dots, n$.

VAJA 3.14. (1) Naj bo V vektorski prostor nad $\mathbb{F}$ s skalarnim produktom in naj bodo $v_1, v_2, \dots, v_k \in \mathsf{V}$ neničelni vektorji, ki so med seboj paroma pravokotni.

Pokaži, da so tedaj vektorji $v_1, v_2, \dots, v_k$ med seboj linearno neodvisni in da za vsak vektor $v \in \mathrm{Span}\{v_1, v_2, \dots, v_k\}$ velja

$$v = \sum_{i=1}^{k} \frac{\langle v, v_i\rangle}{\langle v_i, v_i\rangle} v_i.$$

(2) Naj bo $\mathscr{B} = [v_1, v_2, \dots, v_n]$ baza vektorskega prostora V nad $\mathbb{F}$ in naj bo $\mathscr{B}' = [w_1, w_2, \dots, w_m]$ ortonormirana baza vektorskega prostora W nad $\mathbb{F}$ s skalarnim produktom. Pokaži, da za vsako linearno preslikavo $T : \mathsf{V} \to \mathsf{W}$ velja $\left([T]_{\mathscr{B}'}^{\mathscr{B}}\right)_{ij} = \langle Tv_j, w_i\rangle$, za vse $i = 1, 2, \dots, m$ in $j = 1, 2, \dots, n$.

3.1.3. Gram-Schmidtova ortonormalizacija.

TRDITEV 3.15. *Naj bo* V *vektorski prostor nad* $\mathbb{F}$ *s skalarnim produktom in naj vektorji* $u_1, u_2, \dots, u_r \in \mathsf{V}$ *sestavljajo ortonormiran sistem. Za vsak vektor* $v \in \mathsf{V} \setminus \mathrm{Span}\{u_1, u_2, \dots, u_r\}$ *je vektor*

$$v - \sum_{i=1}^{r} \langle v, u_i\rangle u_i$$

element množice $\mathsf{V} \setminus \mathrm{Span}\{u_1, u_2, \dots, u_r\}$ *in pravokoten na vse vektorje iz vektorskega podprostora* $\mathrm{Span}\{u_1, u_2, \dots, u_r\}$.

DOKAZ. Označimo $u = \sum_{i=1}^{r} \langle v, u_i\rangle u_i$. Ker je $u \in \mathrm{Span}\{u_1, u_2, \dots, u_r\}$ in ker v ni element vektorskega podprostora $\mathrm{Span}\{u_1, u_2, \dots, u_r\}$, odtod sledi $v - u \in \mathsf{V} \setminus \mathrm{Span}\{u_1, u_2, \dots, u_r\}$. Za vsak $j = 1, 2, \dots, r$ velja

$$\langle v - u, u_j\rangle = \langle v, u_j\rangle - \sum_{i=1}^{r} \langle v, u_i\rangle\langle u_i, u_j\rangle = \langle v, u_j\rangle - \langle v, u_j\rangle = 0.$$

Vektor $v - u$ je torej pravokoten na vse vektorje $u_1, u_2, \dots, u_r$, zato je pravokoten na vse vektorje iz linearne ogrinjače $\mathrm{Span}\{u_1, u_2, \dots, u_r\}$. □

IZREK 3.16 (Gram-Schmidtova ortonormalizacija). *Naj bo* V *vektorski prostor nad* $\mathbb{F}$ *s skalarnim produktom. Za poljubne dane vektorje* $v_1, v_2, \dots, v_k \in \mathsf{V}$*, ki so med seboj linearno neodvisni, lahko rekurzivno izračunamo vektorje* $u_1, u_2, \dots, u_k \in \mathsf{V}$ *s predpisom*

$$u_1 = \frac{1}{\|v_1\|} v_1,$$
$$u_j = \frac{1}{\|v_j - \sum_{i=1}^{j-1} \langle v_j, u_i\rangle u_i\|} \left(v_j - \sum_{i=1}^{j-1} \langle v_j, u_i\rangle u_i \right)$$

za $j = 2, 3, \dots, k$*. Dobljeni vektorji* $u_1, u_2, \dots, u_k$ *sestavljajo ortonormiran sistem, ob tem pa velja*

$$\mathrm{Span}\{u_1, u_2, \dots, u_p\} = \mathrm{Span}\{v_1, v_2, \dots, v_p\}$$

za vsak $p = 1, 2, \dots, k$*.*

DOKAZ. Izrek bomo dokazali z indukcijo na število k. Za $k = 1$ izrek očitno velja. Naj bo torej $k \geq 2$ in predpostavimo, da smo že izračunali vektorje $u_1, u_2, \dots, u_{k-1}$ po rekurzivni definiciji iz izreka, da ti vektorji sestavljajo ortonormiran sistem in da je

$$\mathrm{Span}\{u_1, u_2, \dots, u_p\} = \mathrm{Span}\{v_1, v_2, \dots, v_p\}$$

za vsak $p = 1, 2, \ldots, k-1$. Ker so vektorji $v_1, v_2, \ldots, v_k$ med seboj linearno neodvisni, velja $v_k \in \mathsf{V} \setminus \mathrm{Span}\{v_1, v_2, \ldots, v_{k-1}\} = \mathsf{V} \setminus \mathrm{Span}\{u_1, u_2, \ldots, u_{k-1}\}$. Po trditvi 3.15 sledi, da je vektor

$$w = v_k - \sum_{i=1}^{k-1} \langle v_k, u_i \rangle u_i$$

element množice $\mathsf{V} \setminus \mathrm{Span}\{u_1, u_2, \ldots, u_{k-1}\}$ in pravokoten na vse vektorje iz vektorskega podprostora $\mathrm{Span}\{u_1, u_2, \ldots, u_{k-1}\}$. Posebej to pomeni, da je vektor w neničelen, zato je deljenje z njegovo normo v rekurzivni formuli iz trditve dobro definirano. Ker je vektor w pravokoten na vse vektorje iz vektorskega podprostora $\mathrm{Span}\{u_1, u_2, \ldots, u_{k-1}\}$ in ker je

$$u_k = \frac{1}{\|w\|} w,$$

vektorji $u_1, u_2, \ldots, u_{k-1}, u_k$ sestavljajo ortonormiran sistem. Ker sta vektorska podprostora $\mathrm{Span}\{u_1, u_2, \ldots, u_{k-1}, u_k\}$ in $\mathrm{Span}\{v_1, v_2, \ldots, v_{k-1}, v_k\}$ oba dimenzije k in ker iz definicije sledi $\mathrm{Span}\{u_1, u_2, \ldots, u_{k-1}, u_k\} \subset \mathrm{Span}\{v_1, v_2, \ldots, v_{k-1}, v_k\}$, mora veljati tudi $\mathrm{Span}\{u_1, u_2, \ldots, u_{k-1}, u_k\} = \mathrm{Span}\{v_1, v_2, \ldots, v_{k-1}, v_k\}$. □

Komentar 3.17. V izreku opisana Gram-Schmidtova ortonormalizacija je torej rekurzivni algoritem, ki poljubnim med seboj linearno neodvisnim vektorjem $v_1, v_2, \ldots, v_k \in \mathsf{V}$ priredi vektorje $u_1, u_2, \ldots, u_k$, ki sestavljajo ortonormiran sistem in za katere je

$$\mathrm{Span}\{u_1, u_2, \ldots, u_k\} = \mathrm{Span}\{v_1, v_2, \ldots, v_k\}.$$

Če pri tem vektorji $v_1, v_2, \ldots, v_p$ že sestavljajo ortonormiran sistem, za nek $p \in \{1, 2, \ldots, k\}$, potem iz rekurzivne formule sledi, da velja $u_j = v_j$ za vsak $j = 1, 2, \ldots, p$.

Posledica 3.18. *Vsak končno-dimenzionalen vektorski prostor s skalarnim produktom dimenzije $n \geq 1$ ima vsaj eno ortonormirano bazo.*

Dokaz. Izberemo poljubne vektorje $v_1, v_2, \ldots, v_n \in \mathsf{V}$, ki sestavljajo bazo vektorskega prostora V, in na teh vektorjih uporabimo algoritem Gram-Schmidtove ortonormalizacije. Dobljeni vektorji $u_1, u_2, \ldots, u_n \in \mathsf{V}$ sestavljajo ortonormirano bazo vektorskega prostora V. □

Iz izreka med drugim tudi sledi, da vektorje, ki tvorijo ortonormiran sistem v končno-dimenzionalnem vektorskem prostoru s skalarnim produktom, lahko vedno dopolnimo do ortonormirane baze tega vektorskega prostora.

Posledica 3.19. *Naj bo V končno-dimenzionalen vektorski prostor s skalarnim produktom dimenzije $n \geq 1$ in naj vektorji $v_1, v_2, \ldots, v_k \in \mathsf{V}$ sestavljajo ortonormiran sistem. Tedaj obstajajo takšni vektorji $v_{k+1}, v_{k+2}, \ldots, v_n \in \mathsf{V}$, da je $[v_1, \ldots, v_k, v_{k+1}, \ldots, v_n]$ ortonormirana baza vektorskega prostora V.*

Dokaz. Ker vektorji $v_1, v_2, \ldots, v_k$ sestavljajo ortonormiran sistem, so linearno neodvisni, zato jih lahko dopolnimo do baze vektorskega prostora V, recimo z vektorji $w_1, w_2, \ldots, w_{n-k} \in \mathsf{V}$. Na vektorjih $v_1, \ldots, v_k, w_1, \ldots, w_{n-k}$ uporabimo Gram-Schmidtovo ortonormalizacijo in dobimo vektorje $u_1, \ldots, u_k, u_{k+1}, \ldots, u_n$, ki sestavljajo orotnormirano bazo vektorskega prostora V. Ker vektorji $v_1, v_2, \ldots, v_k$ sestavljajo ortonormiran sistem, velja $u_i = v_i$ za vsak $i = 1, 2, \ldots, k$. Za indekse $i = k+1, k+2, \ldots, n$ definiramo $v_i = u_i$. □

3.1.4. Ortogonalni komplement. Naj bo V vektorski prostor nad $\mathbb{F}$ s skalarnim produktom. *Ortogonalni komplement* $\mathcal{O}^\perp$ poljubne podmnožice $\mathcal{O} \subset \mathsf{V}$ v vektorskem prostoru V je množica vseh vektorjev iz vektorskega prostora V, ki so pravokotni na vse vektorje iz množice $\mathcal{O}$, torej

$$\mathcal{O}^\perp = \{v \in \mathsf{V}\,;\ \langle v, u\rangle = 0 \text{ za vsak } u \in \mathcal{O}\} \subset \mathsf{V}.$$

Ker je skalarni produkt linearen v prvem faktorju, je ortogonalni komplement $\mathcal{O}^\perp$ podmnožice $\mathcal{O} \subset \mathsf{V}$ vektorski podprostor vektorskega prostora V, očitno pa velja tudi $\mathcal{O} \subset (\mathcal{O}^\perp)^\perp$ in $\mathcal{O}^\perp = \mathrm{Span}(\mathcal{O})^\perp$.

Vaja 3.20. Pokaži, da v poljubnem vektorskem prostoru V s skalarnim produktom velja $\{0\}^\perp = \mathsf{V}$ in $\mathsf{V}^\perp = \{0\}$.

Trditev 3.21. *Za poljuben vektorski podprostor U končno-dimenzionalnega vektorskega prostora V s skalarnim produktom velja $\mathsf{V} = \mathsf{U} \oplus \mathsf{U}^\perp$ in $\mathsf{U} = (\mathsf{U}^\perp)^\perp$.*

Dokaz. Če je u vektor iz preseka $\mathsf{U} \cap \mathsf{U}^\perp$, je pravokoten sam nase, zaradi pozitivne definitnosti skalarnega produkta pa odtod sledi, da je $u = 0$. Dokazali smo torej, da je $\mathsf{U} \cap \mathsf{U}^\perp = \{0\}$.

Izberimo poljubno ortonormirano bazo $[u_1, u_2, \ldots, u_p]$ vektorskega prostora U in jo dopolnimo do ortonormirane baze vektorskega prostora V, recimo z vektorji $w_1, w_2, \ldots, w_s \in \mathsf{V}$. Za vsak $j = 1, 2, \ldots, s$ je vektor w_j pravokoten na vse vektorje $u_1, u_2, \ldots, u_p$, zato je pravokoten tudi na vse vektorje iz vektorskega prostora U in torej velja $w_j \in \mathsf{U}^\perp$. Poljuben vektor $v \in \mathsf{V}$ lahko zapišemo kot linearno kombinacijo

$$v = \alpha_1 u_1 + \cdots + \alpha_p u_p + \beta_1 w_1 + \cdots + \beta_s w_s$$

za neke skalarje $\alpha_1, \ldots, \alpha_p, \beta_1, \ldots, \beta_s \in \mathbb{F}$, pri tem pa je $\alpha_1 u_1 + \cdots + \alpha_p u_p \in \mathsf{U}$ in $\beta_1 w_1 + \cdots + \beta_s w_s \in \mathsf{U}^\perp$. S tem smo torej dokazali, da velja tudi $\mathsf{U} + \mathsf{U}^\perp = \mathsf{V}$.

Ker torej velja $\mathsf{V} = \mathsf{U} \oplus \mathsf{U}^\perp$, je posebej $\dim \mathsf{U}^\perp = \dim \mathsf{V} - \dim \mathsf{U}$. Odtod sledi

$$\dim(\mathsf{U}^\perp)^\perp = \dim \mathsf{V} - (\dim \mathsf{V} - \dim \mathsf{U}) = \dim \mathsf{U},$$

ker pa je hkrati $\mathsf{U} \subset (\mathsf{U}^\perp)^\perp$, mora veljati $\mathsf{U} = (\mathsf{U}^\perp)^\perp$. □

Naj bo U vektorski podprostor končno-dimenzionalnega vektorskega prostora V s skalarnim produktom. Projektor na vektorski podprostor U vzdolž ortogonalnega komplementa $\mathsf{U}^\perp$ označimo z

$$\mathrm{pr}^\perp_{\mathsf{U}} : \mathsf{V} \to \mathsf{V}$$

in ga imenujemo *pravokotna projekcija na vektorski podprostor* U vektorskega prostora V. Pravokotna projekcija $\mathrm{pr}^\perp_{\mathsf{U}}$ je torej tisti endomorfizem vektorskega prostora V, za katerega velja $\mathrm{pr}^\perp_{\mathsf{U}}|_{\mathsf{U}} = \mathrm{id}$ in $\mathrm{pr}^\perp_{\mathsf{U}}|_{\mathsf{U}^\perp} = 0$. Ker je $\mathrm{pr}^\perp_{\mathsf{U}}$ projektor, je $(\mathrm{pr}^\perp_{\mathsf{U}})^2 = \mathrm{pr}^\perp_{\mathsf{U}}$.

Naj bo $[u_1, u_2, \ldots, u_r]$ poljubna ortonormirana baza vektorskega prostora U. Poljuben vektor $v \in \mathsf{V}$ lahko zapišemo kot vsoto

$$v = \sum_{i=1}^r \langle v, u_i\rangle u_i + \Big(v - \sum_{i=1}^r \langle v, u_i\rangle u_i\Big).$$

Pri tem je očitno $\sum_{i=1}^r \langle v, u_i\rangle u_i \in \mathsf{U}$, medtem ko je $v - \sum_{i=1}^r \langle v, u_i\rangle u_i \in \mathsf{U}^\perp$ po trditvah 3.11 in 3.15. Odtod sledi, da je

$$\mathrm{pr}^\perp_{\mathsf{U}}(v) = \sum_{i=1}^r \langle v, u_i\rangle u_i$$

za vsak vektor $v \in \mathsf{V}$. Posebej lahko iz te formule razberemo, da za poljuben neničelen vektor $w \in \mathsf{V}$ velja $\mathrm{pr}_w^\perp = \mathrm{pr}_{\mathbb{F}w}^\perp$.

Vektor $\mathrm{pr}_{\mathsf{U}}^\perp(v) \in \mathsf{U}$ je izmed vseh vektorjev iz vektorskega podprostora U najbližji vektorju v. Res, ker je $v - \mathrm{pr}_{\mathsf{U}}^\perp(v) \in \mathsf{U}^\perp$ in ker za poljuben vektor $u \in \mathsf{U}$ velja $\mathrm{pr}_{\mathsf{U}}^\perp(v) - u \in \mathsf{U}$, po Pitagorovem izreku dobimo

$$\|v-u\|^2 = \|v - \mathrm{pr}_{\mathsf{U}}^\perp(v) + \mathrm{pr}_{\mathsf{U}}^\perp(v) - u\|^2 = \|v - \mathrm{pr}_{\mathsf{U}}^\perp(v)\|^2 + \|\,\mathrm{pr}_{\mathsf{U}}^\perp(v) - u\|^2.$$

Odtod sledi, da je vrednost $\|v-u\|^2$ minimalna natanko tedaj, ko je $u = \mathrm{pr}_{\mathsf{U}}^\perp(v)$. Število

$$d(v, \mathsf{U}) = \|v - \mathrm{pr}_{\mathsf{U}}^\perp(v)\|$$

imenujemo *razdalja med vektorjem v in vektorskim podprostorom* U.

3.1.5. Rieszov reprezentacijski izrek. Naj bo V vektorski prostor nad $\mathbb{F}$. Linearnim preslikavam $\mathsf{V} \to \mathbb{F}$ pravimo tudi *linearni funkcionali* na vektorskem prostoru V. Vektorski prostor vseh linearnih funkcionalov na vektorskem prostoru V običajno krajše označimo z

$$\mathsf{V}^\vee = \mathrm{Lin}_{\mathbb{F}}(\mathsf{V}, \mathbb{F}) = \mathrm{Lin}(\mathsf{V}, \mathbb{F})$$

in ga imenujemo tudi *dualni vektorski prostor* oziroma *dual* vektorskega prostora V. V literaturi se za dualni vektorski prostor $\mathsf{V}^\vee$ pogosto uporabljata tudi oznaki V' in V^*. Če je vektorski prostor V končno-dimenzionalen, potem je po posledici 2.75 tudi dual $\mathsf{V}^\vee$ končno-dimenzionalen in velja $\dim \mathsf{V}^\vee = \dim \mathsf{V}$.

Komentar 3.22. (1) Naj bo V končno-dimenzionalen vektorski prostor nad $\mathbb{F}$. Ker imata vektorska prostora V in $\mathsf{V}^\vee$ isto dimenzijo, sta izomorfna, a izomorfizem med njima je v splošnem odvisen od izbire baze.

Naj bo $\mathscr{B} = [v_1, v_2, \ldots, v_n]$ baza vektorskega prostora V. Za vsak $j = 1, 2, \ldots, n$ naj bo $\phi_j : \mathsf{V} \to \mathbb{F}$ tisti linearen funkcional, za katerega velja:

$$\phi_j(v_i) = \begin{cases} 1 & ; \quad i = j \\ 0 & ; \quad i \neq j \end{cases}$$

Po trditvi 2.34 je s tem linearen funkcional ϕ_j enolično določen. Za vsak vektor $v \in \mathsf{V}$ je

$$\phi_j(v) = \phi_j\Big(\sum_{i=1}^n \big([v]_{\mathscr{B}}\big)_{i1} v_i\Big) = \sum_{i=1}^n \big([v]_{\mathscr{B}}\big)_{i1} \phi_j(v_i) = \big([v]_{\mathscr{B}}\big)_{j1}.$$

Za poljuben linearen funkcional $\phi \in \mathsf{V}^\vee$ in za vsak vektor $v \in \mathsf{V}$ velja

$$\begin{aligned} \phi(v) = \phi\Big(\sum_{i=1}^n \big([v]_{\mathscr{B}}\big)_{i1} v_i\Big) &= \sum_{i=1}^n \big([v]_{\mathscr{B}}\big)_{i1} \phi(v_i) \\ &= \sum_{i=1}^n \phi_i(v)\phi(v_i) = \Big(\sum_{i=1}^n \phi(v_i)\phi_i\Big)(v), \end{aligned}$$

kar pomeni, da je linearen funkcional ϕ vrednost linearne kombinacije linearnih funkcionalov $\phi_1, \phi_2, \ldots, \phi_n$, torej

$$\phi = \sum_{i=1}^n \phi(v_i)\phi_i.$$

S tem smo pokazali, da linearni funkcionali $\phi_1, \phi_2, \ldots, \phi_n$ generirajo vektorski prostor $\mathsf{V}^\vee$. Ker je $\dim \mathsf{V}^\vee = n$, so linearni funkcionali $\phi_1, \phi_2, \ldots, \phi_n$ med seboj tudi linearno neodvisni. To lahko vidimo tudi direktno: če so namreč $\alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbb{F}$ takšni skalarji, da je $\sum_{j=1}^n \alpha_j \phi_j = 0$, potem za vsak $i = 1, 2, \ldots, n$ velja

$$0 = \Big(\sum_{j=1}^n \alpha_j \phi_j\Big)(v_i) = \sum_{j=1}^n \alpha_j \phi_j(v_i) = \alpha_i.$$

Linearni funkcionali $\phi_1, \phi_2, \ldots, \phi_n$ torej sestavljajo bazo dualnega vektorskega prostora $\mathsf{V}^\vee$, ki jo označimo z $\mathscr{B}^\vee$ in imenujemo *dualna baza baze* $\mathscr{B}$.

Z izbiro baze $\mathscr{B} = [v_1, v_2, \ldots, v_n]$ vektorskega prostora V je torej določena njena dualna baza $\mathscr{B}^\vee = [\phi_1, \phi_2, \ldots, \phi_n]$, s tem pa je določen tudi izomorfizem vektorskih prostorov $\mathsf{V} \to \mathsf{V}^\vee$ kot tisti izomorfizem, ki preslika bazo $\mathscr{B}$ v dualno bazo $\mathscr{B}^\vee$.

(2) Naj bo V končno-dimenzionalen vektorski prostor nad $\mathbb{F}$. Po točki (1) sta vektorska prostora V in $\mathsf{V}^\vee$ izomorfna, prav tako pa sta izomorfna vektorska prostora $\mathsf{V}^\vee$ in $(\mathsf{V}^\vee)^\vee$. Vektorski prostor V je torej tudi izomorfen svojemu *drugemu dualu*

$$\mathsf{V}^{\vee\vee} = (\mathsf{V}^\vee)^\vee = \mathrm{Lin}(\mathrm{Lin}(\mathsf{V}, \mathbb{F}), \mathbb{F}).$$

Vendar pa velja še nekoliko več: med vektorskim prostorom V in njegovim drugim dualom $\mathsf{V}^{\vee\vee}$ obstaja naravni izomorfizem $\delta = \delta_{\mathsf{V}} : \mathsf{V} \to \mathsf{V}^{\vee\vee}$, za katerega lahko zapišemo predpis brez da bi za to morali izbrati bazo vektorskega prostora V. Za poljuben vektor $v \in \mathsf{V}$ namreč lahko definiramo

$$\delta(v)(\phi) = \phi(v),$$

za vsak $\phi \in \mathsf{V}^\vee$. Ni težko preveriti, da je s tem res dobro definirana linearna preslikava $\delta \in \mathrm{Lin}(\mathsf{V}, \mathsf{V}^{\vee\vee})$. Preslikava δ je izomorfizem, saj preslika poljubno bazo $\mathscr{B}$ vektorskega prostora V v bazo $\mathscr{B}^{\vee\vee} = (\mathscr{B}^\vee)^\vee$ drugega duala $\mathsf{V}^{\vee\vee}$.

(3) Za poljubno linearno preslikavo $T \in \mathrm{Lin}(\mathsf{V}, \mathsf{W})$ med vektorskimi prostori imamo njej *dualno linearno preslikavo*

$$T^\vee : \mathsf{W}^\vee \to \mathsf{V}^\vee, \qquad \psi \mapsto \psi T.$$

Tako definirana preslikava $\mathrm{Lin}(\mathsf{V}, \mathsf{W}) \to \mathrm{Lin}(\mathsf{W}^\vee, \mathsf{V}^\vee)$, $T \mapsto T^\vee$, je linearna ter velja $\mathrm{id}_{\mathsf{V}}^\vee = \mathrm{id}_{\mathsf{V}^\vee}$ in $(ST)^\vee = T^\vee S^\vee$, kjer je $S \in \mathrm{Lin}(\mathsf{W}, \mathsf{Z})$ še ena poljubna linearna preslikava med vektorskimi prostori.

Zgled 3.23. (1) Naj bo V vektorski prostor nad $\mathbb{F}$ s skalarnim produktom. Za poljuben vektor $v \in \mathsf{V}$ je preslikava $\langle\,\bullet\,, v\rangle : \mathsf{V} \to \mathbb{F}$, $u \mapsto \langle u, v\rangle$, linearen funkcional na vektorskem prostoru V, torej $\langle\,\bullet\,, v\rangle \in \mathsf{V}^\vee$.

(2) Naj bosta V in W poljubna vektorska prostora nad $\mathbb{F}$. Preslikava $T : \mathsf{V} \to \mathsf{W}$ je *konjugirano linearna*, če za poljubna vektorja $v, v' \in \mathsf{V}$ in za vsak skalar $\alpha \in \mathbb{F}$ velja

$$T(v + v') = T(v) + T(v')$$

in

$$T(\alpha v) = \bar{\alpha} T(v).$$

Če je $\mathbb{F} = \mathbb{R}$, potem je preslikava konjugirano linearna, če, in samo če, je linearna. Če je $\mathbb{F} = \mathbb{C}$, potem v splošnem konjugirano linearne preslikave niso linearne nad kompleksnimi skalarji. Ker pa velja $\mathbb{R} \subset \mathbb{C}$, je vsak kompleksen vektorski prostor posebej tudi realen vektorski prostor, vsaka konjugirano linearna preslikava med vektorskimi prostori pa je linearna preslikava nad realnimi skalarji, torej linearna preslikava realnih vektorskih prostorov.

Naj bo V vektorski prostor nad $\mathbb{F}$ s skalarnim produktom. Za poljuben vektor $u \in \mathsf{V}$ je tedaj preslikava $\langle u, \bullet\,\rangle : \mathsf{V} \to \mathbb{F}$, $v \mapsto \langle u, v\rangle$, konjugirano linearna.

IZREK 3.24 (Rieszov reprezentacijski izrek). *Naj bo* V *končno-dimenzionalen vektorski prostor nad* $\mathbb{F}$ *s skalarnim produktom. Preslikava* $\mathsf{V} \to \mathsf{V}^\vee$*,* $v \mapsto \langle\,\bullet\,, v\rangle$*, je konjugirano linearna bijekcija. Za poljuben linearen funkcional* $\phi \in \mathsf{V}^\vee$ *torej obstaja natanko en tak vektor* $v \in \mathsf{V}$*, da je* $\phi(u) = \langle u, v\rangle$ *za vsak vektor* $u \in \mathsf{V}$*.*

DOKAZ. Iz osnovnih lastnosti skalarnega produkta direktno sledi, da je preslikava $T : \mathsf{V} \to \mathsf{V}^\vee$, $T(v) = \langle\,\bullet\,, v\rangle$, konjugirano linearna. Če je $T(v) = 0$ za nek vektor $v \in \mathsf{V}$, je $\langle u, v\rangle = 0$ za vsak vektor $u \in \mathsf{V}$ in posebej $\langle v, v\rangle = 0$, odtod pa sledi $v = 0$. Velja torej $T^{-1}(\{0\}) = \{0\}$.

Preslikava T je linearna nad realnimi skalarji in ima kot takšna trivialno jedro. Ker imata vektorska prostora V in $\mathsf{V}^\vee$ isto dimenzijo kot vektorska prostora nad $\mathbb{F}$, imata isto dimenzijo tudi kot realna vektorska prostora. Odtod sledi, da je preslikava T izomorfizem realnih vektorskih prostorov in torej bijekcija. □

KOMENTAR 3.25. Če je ϕ linearen funkcional na vektorskem prostoru V s skalarnim produktom in če je $v \in \mathsf{V}$ tak vektor, da velja $\phi(u) = \langle u, v\rangle$ za vsak vektor $u \in \mathsf{V}$, potem pravimo, da vektor v *reprezentira funkcional* ϕ. Rieszov izrek torej pravi, da poljuben linearen funkcional na končno-dimenzionalnem vektorskem prostoru V s skalarnim produktom lahko reprezentiramo z natanko enim vektorjem iz vektorskega prostora V.

ZGLED 3.26. (1) Ker je $(\mathbb{F}^n)^\vee = \mathrm{Lin}(\mathbb{F}^n, \mathbb{F})$, so linearni funkcionali na vektorskem prostoru $\mathbb{F}^n$ glede na standardno bazo dani z vrsticami velikosti $1 \times n$. Naj bo $\Lambda \in \mathrm{Mat}(1 \times n, \mathbb{F})$ takšna vrstica. Tedaj je $\Lambda^{\mathsf{h}} \in \mathrm{Mat}(n \times 1, \mathbb{F}) = \mathbb{F}^n$ stolpec oziroma vektor v $\mathbb{F}^n$, standardni skalarni produkt tega vektorja s poljubnim vektorjem $u \in \mathbb{F}^n = \mathrm{Mat}(n \times 1, \mathbb{F})$ pa je

$$\langle u, \Lambda^{\mathsf{h}}\rangle = (\Lambda^{\mathsf{h}})^{\mathsf{h}} u = \Lambda u = \ell(\Lambda)(u).$$

To pomeni, da je linearen funkcional $\ell(\Lambda)$, ki ustreza vrstici Λ, glede na standardni skalarni produkt po Rieszovem izreku reprezentiran z vektorjem Λ^{h}.

(2) Naj bo $\phi \in \mathsf{V}^\vee$ linearen funkcional na končno-dimenzionalnem vektorskem prostoru V nad $\mathbb{F}$ s skalarnim produktom. Naj bo $\mathscr{B} = [v_1, v_2, \ldots, v_n]$ ortonormirana baza vektorskega prostora V. Za poljuben vektor $u \in \mathsf{V}$ je

$$\begin{aligned} \phi(u) &= \phi\Big(\sum_{j=1}^n \langle u, v_j\rangle v_j\Big) = \sum_{j=1}^n \langle u, v_j\rangle \phi(v_j) \\ &= \Big\langle \sum_{j=1}^n \langle u, v_j\rangle v_j, \sum_{i=1}^n \overline{\phi(v_i)} v_i \Big\rangle = \Big\langle u, \sum_{i=1}^n \overline{\phi(v_i)} v_i \Big\rangle. \end{aligned}$$

Funkcional ϕ je torej reprezentiran z vektorjem $\sum_{i=1}^n \overline{\phi(v_i)} v_i \in \mathsf{V}$.

(3) Na vektorskem prostoru $\mathbb{R}_2[t]$ realnih polinomov stopnje največ dva izberimo skalarni produkt, dan s predpisom

$$\langle f, g\rangle = \int_{-1}^1 f(t) g(t)\, dt$$

za vse polinome $f, g \in \mathbb{R}_2[t]$.

Poiščimo kakšno ortonormirano bazo vektorskega prostora $\mathbb{R}_2[t]$ z izbranim skalarnim produktom. Na polinomih $f_1(t) = 1$, $f_2(t) = t$, $f_3(t) = t^2$, ki sestavljajo bazo vektorskega prostora $\mathbb{R}_2[t]$, uporabimo Gram-Schmidtovo ortonormalizacijo:

Najprej izračunamo $\|f_1\|^2 = \int_{-1}^{1} dt = 2$, zato je

$$h_1(t) = \frac{1}{\|f_1\|} f_1 = \tfrac{1}{\sqrt{2}}$$

prvi polinom, ki ga dobimo po Gram-Schmidtovem algoritmu.

V drugem koraku izračunamo $\langle f_2, h_1 \rangle = \int_{-1}^{1} t \tfrac{1}{\sqrt{2}} \, dt = 0$. Odtod sledi, da velja $f_2 - \langle f_2, h_1 \rangle h_1 = f_2$ in $\|f_2 - \langle f_2, h_1 \rangle h_1\|^2 = \|f_2\|^2 = \int_{-1}^{1} t^2 \, dt = \tfrac{2}{3}$, zato je

$$h_2(t) = \frac{1}{\|f_2 - \langle f_2, h_1 \rangle h_1\|} \big(f_2 - \langle f_2, h_1 \rangle h_1\big)(t) = \tfrac{\sqrt{3}}{\sqrt{2}} t$$

drugi polinom, ki ga dobimo po Gram-Schmidtovem algoritmu.

V tretjem koraku najprej izračunamo $\langle f_3, h_1 \rangle = \int_{-1}^{1} t^2 \tfrac{1}{\sqrt{2}} \, dt = \tfrac{\sqrt{2}}{3}$ in $\langle f_3, h_2 \rangle = \int_{-1}^{1} t^2 \tfrac{\sqrt{3}}{\sqrt{2}} t \, dt = 0$. Odtod sledi, da je $\big(f_3 - \langle f_3, h_1 \rangle h_1 - \langle f_3, h_2 \rangle h_2\big)(t) = t^2 - \tfrac{1}{3}$ in $\|f_3 - \langle f_3, h_1 \rangle h_1 - \langle f_3, h_2 \rangle h_2\|^2 = \int_{-1}^{1} (t^2 - \tfrac{1}{3})^2 \, dt = \tfrac{8}{45}$. Tretji polinom, ki ga dobimo z Gram-Schmidtovo ortonormalizacijo, je torej

$$\begin{aligned} h_3(t) &= \frac{1}{\|f_3 - \langle f_3, h_1 \rangle h_1 - \langle f_3, h_2 \rangle h_2\|} \big(f_3 - \langle f_3, h_1 \rangle h_1 - \langle f_3, h_2 \rangle h_2\big)(t) \\ &= \tfrac{\sqrt{45}}{\sqrt{8}} (t^2 - \tfrac{1}{3}) = \tfrac{\sqrt{5}}{2\sqrt{2}} (3t^2 - 1). \end{aligned}$$

Dobili smo ortonormirano bazo

$$[h_1, h_2, h_3] = \Big[\tfrac{1}{\sqrt{2}}, \tfrac{\sqrt{3}}{\sqrt{2}} t, \tfrac{\sqrt{5}}{2\sqrt{2}} (3t^2 - 1) \Big]$$

vektorskega prostora $\mathbb{R}_2[t]$ glede na izbrani skalarni produkt.

Vektorski prostor $\mathsf{U} = \mathbb{R}_1[t]$ je vektorski podprostor vektorskega prostora $\mathbb{R}_2[t]$, pri čemer velja

$$\mathsf{U} = \mathrm{Span}\{1, t\} = \mathrm{Span}\{h_1, h_2\}.$$

Polinoma h_1, h_2 sestavljata ortonormirano bazo vektorskega prostora U. Izračunajmo pravokotno projekcijo polinoma $g \in \mathbb{R}_2[t]$, $g(t) = t^2 + t$, na vektorski podprostor U. Ker velja $\langle g, h_1 \rangle = \int_{-1}^{1} (t^2 + t) \tfrac{1}{\sqrt{2}} \, dt = \tfrac{\sqrt{2}}{3}$ in $\langle g, h_2 \rangle = \int_{-1}^{1} (t^2 + t) \tfrac{\sqrt{3}}{\sqrt{2}} t \, dt = \tfrac{\sqrt{2}}{\sqrt{3}}$, odtod dobimo

$$\mathrm{pr}_{\mathsf{U}}^{\perp}(g)(t) = \big(\langle g, h_1 \rangle h_1 + \langle g, h_2 \rangle h_2\big)(t) = \tfrac{\sqrt{2}}{3} \tfrac{1}{\sqrt{2}} + \tfrac{\sqrt{2}}{\sqrt{3}} \tfrac{\sqrt{3}}{\sqrt{2}} t = t + \tfrac{1}{3}.$$

Tako velja $\big(g - \mathrm{pr}_{\mathsf{U}}^{\perp}(g)\big)(t) = t^2 - \tfrac{1}{3}$ in

$$\|g - \mathrm{pr}_{\mathsf{U}}^{\perp}(g)\|^2 = \int_{-1}^{1} (t^2 - \tfrac{1}{3})^2 \, dt = \tfrac{8}{45},$$

zato je razdalja med polinomom g in vektorskim podprostorom U enaka številu

$$d(g, \mathsf{U}) = \tfrac{2\sqrt{2}}{3\sqrt{5}}.$$

Na vektorskem prostoru $\mathbb{R}_2[t]$ imamo linearen funkcional $\phi : \mathbb{R}_2[t] \to \mathbb{R}$, dan s predpisom

$$\phi(f) = \frac{df}{dt}(0).$$

Izračunamo lahko $\phi(h_1) = 0$, $\phi(h_2) = \tfrac{\sqrt{3}}{\sqrt{2}}$ in $\phi(h_3) = 0$. Odtod sledi, da je linearen funkcional ϕ reprezentiran s polinomom

$$\big(\phi(h_1) h_1 + \phi(h_2) h_2 + \phi(h_3) h_3\big)(t) = \tfrac{3}{2} t.$$

To pomeni, da za vsak polinom $f \in \mathbb{R}_2[t]$ velja

$$\frac{df}{dt}(0) = \tfrac{3}{2}\int_{-1}^{1} tf(t)\,dt.$$

3.2. Preslikave med vektorskimi prostori s skalarnim produktom

3.2.1. Adjungirana preslikava. Naj bosta V in W končno-dimenzionalna vektorska prostora nad $\mathbb{F}$ s skalarnim produktom in naj bo $T \in \mathrm{Lin}(\mathsf{V},\mathsf{W})$ linearna preslikava. Za poljuben vektor $w \in \mathsf{W}$ je preslikava $\langle T\bullet, w\rangle : \mathsf{V} \to \mathbb{F}$, $v \mapsto \langle Tv, w\rangle$, linearen funkcional na vektorskem prostoru V. Po Rieszovem izreku lahko ta funkcional reprezentiramo z natanko enim vektorjem iz vektorskega prostora V, ki ga bomo označili z $T^*w \in \mathsf{V}$. Drugače povedano, vektor $T^*w \in \mathsf{V}$ je tisti enolično določen vektor iz vektorskega prostora V, za katerega je

$$\langle Tv, w\rangle = \langle v, T^*w\rangle$$

za vsak vektor $v \in \mathsf{V}$. S tem smo definirali preslikavo

$$T^* : \mathsf{W} \to \mathsf{V}, \qquad w \mapsto T^*w,$$

ki jo imenujemo *adjungirana preslikava* linearne preslikave T.

TRDITEV 3.27. *Naj bosta V in W končno-dimenzionalna vektorska prostora nad $\mathbb{F}$ s skalarnim produktom. Za vsako linearno preslikavo $T \in \mathrm{Lin}(\mathsf{V},\mathsf{W})$ je njej adjungirana preslikava $T^* : \mathsf{W} \to \mathsf{V}$ linearna.*

DOKAZ. Naj bosta $w, w' \in \mathsf{W}$ poljubna vektorja in $\alpha \in \mathbb{F}$ poljuben skalar. Za vsak vektor $v \in \mathsf{V}$ velja

$$\begin{aligned}\langle v, T^*(w + w')\rangle &= \langle Tv, w + w'\rangle = \langle Tv, w\rangle + \langle Tv, w'\rangle \\ &= \langle v, T^*w\rangle + \langle v, T^*w'\rangle = \langle v, T^*w + T^*w'\rangle\end{aligned}$$

in

$$\langle v, T^*(\alpha w)\rangle = \langle Tv, \alpha w\rangle = \bar{\alpha}\langle Tv, w\rangle = \bar{\alpha}\langle v, T^*w\rangle = \langle v, \alpha T^*w\rangle,$$

odtod pa po trditvi 3.5 sledi $T^*(w + w') = T^*w + T^*w'$ in $T^*(\alpha w) = \alpha T^*w$. □

Adjungirana preslikava linearne preslikave $T : \mathsf{V} \to \mathsf{W}$ je torej linearna preslikava $T^* : \mathsf{W} \to \mathsf{V}$, ki je enolično določena s pogojem, da velja

$$\langle Tv, w\rangle = \langle v, T^*w\rangle$$

za vse vektorje $v \in \mathsf{V}$ in za vse vektorje $w \in \mathsf{W}$. Opazimo lahko, da v zadnji enakosti na desni nastopa skalarni produkt na vektorskem prostoru V, na levi pa skalarni produkt na vektorskem prostoru W.

TRDITEV 3.28. *Naj bodo V, W in Z končno-dimenzionalni vektorski prostori nad $\mathbb{F}$ s skalarnim produktom. Za poljubne linearne preslikave $T, T' \in \mathrm{Lin}(\mathsf{V},\mathsf{W})$ in $S \in \mathrm{Lin}(\mathsf{W},\mathsf{Z})$ ter za vsak skalar $\alpha \in \mathbb{F}$ velja:*

(i) $(T + T')^* = T^* + T'^*$,
(ii) $(\alpha T)^* = \bar{\alpha}T^*$,
(iii) $(T^*)^* = T$,
(iv) $(ST)^* = T^*S^*$,
(v) $\mathrm{id}^* = \mathrm{id}$,
(vi) $\ker(T^*) = (\mathrm{im}(T))^\perp$ *in*
(vii) $\mathrm{im}(T^*) = (\ker(T))^\perp$.

KOMENTAR 3.29. Preslikavi $\mathrm{Lin}(\mathsf{V},\mathsf{W}) \to \mathrm{Lin}(\mathsf{W},\mathsf{V})$, $T \mapsto T^*$, pravimo tudi *adjungiranje*. Iz točk (i-iii) sledi, da je adjungiranje $\mathrm{Lin}(\mathsf{V},\mathsf{W}) \to \mathrm{Lin}(\mathsf{W},\mathsf{V})$ konjugirano linearna bijekcija, katere inverz je adjungiranje $\mathrm{Lin}(\mathsf{W},\mathsf{V}) \to \mathrm{Lin}(\mathsf{V},\mathsf{W})$.

DOKAZ. Za poljubna vektorja $v \in \mathsf{V}$ in $w \in \mathsf{W}$ velja

$$\begin{aligned}\langle v, (T+T')^*w\rangle &= \langle (T+T')v, w\rangle = \langle Tv + T'v, w\rangle \\ &= \langle Tv, w\rangle + \langle T'v, w\rangle = \langle v, T^*w\rangle + \langle v, T'^*w\rangle \\ &= \langle v, T^*w + T'^*w\rangle = \langle v, (T^* + T'^*)w\rangle\end{aligned}$$

in

$$\begin{aligned}\langle v, (\alpha T)^*w\rangle &= \langle (\alpha T)v, w\rangle = \alpha\langle Tv, w\rangle \\ &= \alpha\langle v, T^*w\rangle = \langle v, (\bar{\alpha}T^*)w\rangle,\end{aligned}$$

po trditvi 3.5 pa sledi $(T+T')^*w = (T^* + T'^*)w$ in $(\alpha T)^*w = (\bar{\alpha}T^*)w$. S tem smo dokazali točki (i) in (ii).

Ker za vse vektorje $v, v' \in \mathsf{V}$ in $w \in \mathsf{W}$ velja

$$\langle w, (T^*)^*v\rangle = \langle T^*w, v\rangle = \overline{\langle v, T^*w\rangle} = \overline{\langle Tv, w\rangle} = \langle w, Tv\rangle$$

in

$$\langle v, \mathrm{id}^*(v')\rangle = \langle \mathrm{id}(v), v'\rangle = \langle v, v'\rangle = \langle v, \mathrm{id}(v')\rangle,$$

po trditvi 3.5 sledita tudi točki (iii) in (v). Podobno točka (iv) sledi iz enakosti

$$\begin{aligned}\langle v, (ST)^*z\rangle &= \langle (ST)v, z\rangle = \langle S(Tv), z\rangle = \langle Tv, S^*z\rangle \\ &= \langle v, T^*(S^*z)\rangle = \langle v, (T^*S^*)z\rangle\end{aligned}$$

ki velja za vsak vektor $v \in \mathsf{V}$ in vsak vektor $z \in \mathsf{Z}$.

(vi) Vzemimo poljuben vektor $w \in \mathsf{W}$. Če je $w \in \ker(T^*)$, potem je za poljuben vektor $v \in \mathsf{V}$

$$\langle Tv, w\rangle = \langle v, T^*w\rangle = 0,$$

zato je vektor w pravokoten na vse vektorje iz slike $\mathrm{im}(T)$. S tem smo dokazali, da velja $\ker(T^*) \subset (\mathrm{im}(T))^\perp$. Obratno, če je $w \in (\mathrm{im}(T))^\perp$, velja

$$\langle v, T^*w\rangle = \langle Tv, w\rangle = 0$$

za vsak vektor $v \in \mathsf{V}$. Vektor T^*w je torej v tem primeru pravokoten na vse vektorje iz vektorskega prostora V, zato mora biti enak vektorju 0. S tem smo dokazali, da je tudi $\ker(T^*) \supset (\mathrm{im}(T))^\perp$.

Točka (vii) sledi iz pravkar dokazanih točk (vi) in (ii), saj je

$$\mathrm{im}(T^*) = ((\mathrm{im}(T^*))^\perp)^\perp = (\ker((T^*)^*))^\perp = (\ker(T))^\perp. \qquad \square$$

TRDITEV 3.30. *Naj bosta V in W vektorska prostora nad $\mathbb{F}$ s skalarnim produktom, oba končno-dimenzionalna s pozitivno dimenzijo, naj bo $\mathscr{B}$ ortonormirana baza vektorskega prostora V in naj bo $\mathscr{B}'$ ortonormirana baza vektorskega prostora W. Za poljubno linearno preslikavo $T \in \mathrm{Lin}(\mathsf{V},\mathsf{W})$ velja*

$$[T^*]^{\mathscr{B}'}_{\mathscr{B}} = \left([T]^{\mathscr{B}}_{\mathscr{B}'}\right)^{\mathsf{h}}.$$

Posebej, če je $\mathbb{F} = \mathbb{R}$, potem velja $[T^*]^{\mathscr{B}'}_{\mathscr{B}} = ([T]^{\mathscr{B}}_{\mathscr{B}'})^{\mathsf{t}}$.

DOKAZ. Naj bo $\mathscr{B} = [v_1, v_2, \dots, v_n]$ in $\mathscr{B}' = [w_1, w_2, \dots, w_m]$. Označimo $A = [T]^{\mathscr{B}}_{\mathscr{B}'}$ in $B = [T^*]^{\mathscr{B}'}_{\mathscr{B}}$. Za vse $i = 1, 2, \dots, m$ in $j = 1, 2, \dots, n$ je tedaj

$$\begin{aligned} A_{ij} &= \langle A_{1j} w_1 + A_{2j} w_2 + \cdots + A_{mj} w_m, w_i \rangle = \langle T v_j, w_i \rangle = \langle v_j, T^* w_i \rangle \\ &= \langle v_j, B_{1i} v_1 + B_{2i} v_2 + \cdots + B_{ni} v_n \rangle = \overline{B_{ji}}. \end{aligned}$$ □

ZGLED 3.31. (1) Za endomorfizem $\ell(A)$ vektorskega prostora $\mathbb{F}^n$ s standardnim skalarnim produktom, ki je dan s kvadratno matriko A, velja $\ell(A)^* = \ell(A^{\mathsf{h}})$.

(2) Naj bo V vektorski prostor nad $\mathbb{F}$. Vsakemu vektorju $v \in \mathsf{V}$ lahko priredimo linearno preslikavo $\bullet v : \mathbb{F} \to \mathsf{V}$, $\alpha \mapsto \alpha v$. Obratno, vsaka preslikava $T \in \mathrm{Lin}(\mathbb{F}, \mathsf{V})$ je oblike $T = \bullet(T(1))$. Odtod sledi, da je linearna preslikava

$$\mathsf{V} \to \mathrm{Lin}(\mathbb{F}, \mathsf{V}), \qquad v \mapsto \bullet v,$$

izomorfizem vektorskih prostorov. Preko tega izomorfizma identificiramo vektorska prostora V in $\mathrm{Lin}(\mathbb{F}, \mathsf{V})$, zato bi lahko linearno preslikavo $\bullet v$ označili enostavno z v.

Predpostavimo zdaj, da je vektorski prostor V končno-dimenzionalen in da je opremljen s skalarnim produktom. Na vektorskem prostoru $\mathbb{F}$ imamo standardni skalarni produkt. Za vsak vektor $v \in \mathsf{V}$ je linearni preslikavi $\bullet v : \mathbb{F} \to \mathsf{V}$ pridružena adjungirana preslikava $(\bullet v)^* : \mathsf{V} \to \mathbb{F}$, ki je linearen funkcional na vektorskem prostoru V. Za vsak vektor $u \in \mathsf{V}$ in za vsak skalar $\alpha \in \mathbb{F}$ velja

$$\langle u, (\bullet v)(\alpha) \rangle = \langle u, \alpha v \rangle = \langle u, v \rangle \bar{\alpha} = \langle \langle u, v \rangle, \alpha \rangle = \langle \langle \bullet, v \rangle (u), \alpha \rangle,$$

zato je $\langle \bullet, v \rangle^* = \bullet v$ in $(\bullet v)^* = \langle \bullet, v \rangle$. Posebej je $(\bullet \alpha)^* = \langle \bullet, \alpha \rangle = \bullet \bar{\alpha} \in \mathrm{Lin}(\mathbb{F}, \mathbb{F})$.

Poljubnemu linearnemu funkcionalu $\phi : \mathsf{V} \to \mathbb{F}$ na vektorskem prostoru V je pridružena adjungirana preslikava $\phi^* : \mathbb{F} \to \mathsf{V}$. Ker za vsak $u \in \mathsf{V}$ velja

$$\phi(u) = \langle \phi(u), 1 \rangle = \langle u, \phi^*(1) \rangle,$$

je linearen funkcional ϕ po Rieszovem izreku reprezentiran z vektorjem $\phi^*(1)$.

KOMENTAR 3.32. Pojem adjungirane preslikave je zelo pomemben v kvantni mehaniki, kjer je sicer običajna nekoliko drugačna, tako imenovana Diracova *bra-ket* notacija. Za poljubno linearno preslikavo $T \in \mathrm{Lin}(\mathsf{V}, \mathsf{W})$ med končno-dimenzionalnima vektorskima prostoroma s skalarnim produktom in poljubne vektorje $u, v \in \mathsf{V}$ označimo naslednje linearne preslikave:

$$\begin{aligned} |v\rangle &= \bullet v \\ \langle u| &= \langle \bullet, u \rangle \\ \langle u|v\rangle &= \langle u| \circ |v\rangle = \bullet \langle v, u \rangle \\ T^\dagger &= T^* \end{aligned}$$

S kompozicijo takšnih linearnih preslikav, za $v \in \mathsf{V}$ in $w \in \mathsf{W}$, dobimo na primer:

$$\begin{aligned} T|v\rangle &= T \circ |v\rangle = \bullet(Tv) \\ \langle w|T &= \langle w| \circ T = \langle T\bullet, w \rangle \\ \langle w|T|v\rangle &= \langle w| \circ T \circ |v\rangle = \bullet \langle Tv, w \rangle \end{aligned}$$

Ob tem pa velja:

$$\begin{aligned} |v\rangle^\dagger &= \langle v| \\ \langle w|^\dagger &= |w\rangle \\ \langle w|T|v\rangle^\dagger &= \langle v|T^\dagger|w\rangle \end{aligned}$$

Naj bo $f \in \mathbb{F}[t]$ poljuben polinom s koeficienti iz $\mathbb{F}$ v spremenljivki t,

$$f(t) = a_s t^s + a_{s-1} t^{s-1} + \cdots + a_1 t + a_0.$$

Konjugiran polinom polinoma f je polinom $\bar{f} \in \mathbb{F}[t]$, dan s predpisom

$$\bar{f}(t) = \overline{a_s} t^s + \overline{a_{s-1}} t^{s-1} + \cdots + \overline{a_1} t + \overline{a_0}.$$

Za poljubna polinoma $f, g \in \mathbb{F}[t]$ in vsak skalar $\alpha \in \mathbb{F}$ očitno velja $\overline{f+g} = \bar{f} + \bar{g}$, $\overline{\alpha f} = \bar{\alpha}\bar{f}$ in $\overline{fg} = \bar{f}\bar{g}$. V primeru $\mathbb{F} = \mathbb{R}$ je seveda $\bar{f} = f$.

Za poljuben skalar $\lambda \in \mathbb{F}$ je

$$\bar{f}(\lambda) = \overline{f(\bar{\lambda})}.$$

Če velja $f(t) = (t-\lambda)^k h(t)$ za neko naravno število k in nek polinom $h \in \mathbb{F}[t]$, potem je tudi $\bar{f}(t) = (t - \bar{\lambda})^k \bar{h}(t)$. Odtod sledi, da je $\mathrm{kr}_{\bar{f}}(\lambda) = \mathrm{kr}_f(\bar{\lambda})$.

Posledica 3.33. *Naj bo* V *končno-dimenzionalen vektorski prostor nad* $\mathbb{F}$ *s skalarnim produktom in naj bo* $n = \dim \mathsf{V} \geq 1$. *Naj bo* $T \in \mathrm{End}(\mathsf{V})$ *endomorfizem vektorskega prostora* V. *Tedaj je* $\mathrm{p}_{T^*} = \overline{\mathrm{p}_T}$, $\det(T^*) = \overline{\det(T)}$ *in* $\mathrm{tr}(T^*) = \overline{\mathrm{tr}(T)}$, *velja pa tudi*

$$\sigma(T^*) = \overline{\sigma(T)} = \{\bar{\lambda}\ ;\ \lambda \in \sigma(T)\}.$$

Poleg tega za vsak skalar $\lambda \in \mathbb{F}$ *velja* $\mathrm{akr}_{T^*}(\lambda) = \mathrm{akr}_T(\bar{\lambda})$, $\mathrm{gkr}_{T^*}(\lambda) = \mathrm{gkr}_T(\bar{\lambda})$ *in*

$$\mathrm{E}_{T^*}(\lambda) = \big(\mathrm{im}(T - \bar{\lambda}\,\mathrm{id})\big)^{\perp}.$$

Dokaz. Izberimo poljubno ortonormirano bazo $\mathscr{B}$ vektorskega prostora V in označimo z $A = [T]^{\mathscr{B}}_{\mathscr{B}}$ koordinatno matriko endomorfizma T. Po trditvi 3.30 velja $[T^*]^{\mathscr{B}}_{\mathscr{B}} = A^{\mathsf{h}}$ in za vsak $t \in \mathbb{F}$ lahko izračunamo

$$\begin{aligned}
\mathrm{p}_{T^*}(t) &= \det(T^* - t\,\mathrm{id}) = \det(A^{\mathsf{h}} - t\mathrm{I}) = \det((A - \bar{t}\mathrm{I})^{\mathsf{h}}) \\
&= \overline{\det(A - \bar{t}\mathrm{I})} = \overline{\mathrm{p}_T(\bar{t}\,)} = \overline{\mathrm{p}_T}(t).
\end{aligned}$$

Odtod sledi

$$\mathrm{akr}_{T^*}(\lambda) = \mathrm{kr}_{\mathrm{p}_{T^*}}(\lambda) = \mathrm{kr}_{\overline{\mathrm{p}_T}}(\lambda) = \mathrm{kr}_{\mathrm{p}_T}(\bar{\lambda}) = \mathrm{akr}_T(\bar{\lambda}),$$

zato je tudi $\sigma(T^*) = \overline{\sigma(T)}$.

Iz trditve 3.28(vi) dobimo

$$\mathrm{E}_{T^*}(\lambda) = \ker(T^* - \lambda\,\mathrm{id}) = \ker((T - \bar{\lambda}\,\mathrm{id})^*) = \big(\,\mathrm{im}(T - \bar{\lambda}\,\mathrm{id})\big)^{\perp}$$

in odtod

$$\begin{aligned}
\mathrm{gkr}_{T^*}(\lambda) &= \dim(\mathrm{E}_{T^*}(\lambda)) = \dim\big(\,\mathrm{im}(T - \bar{\lambda}\,\mathrm{id})\big)^{\perp} \\
&= n - \dim(\mathrm{im}(T - \bar{\lambda}\,\mathrm{id})) = \dim(\ker(T - \bar{\lambda}\,\mathrm{id})) = \mathrm{gkr}_T(\bar{\lambda}).
\end{aligned}$$ □

Trditev 3.34. *Naj bo* V *končno-dimenzionalen vektorski prostor nad* $\mathbb{F}$ *s skalarnim produktom in naj bo* $T \in \mathrm{End}(\mathsf{V})$ *endomorfizem vektorskega prostora* V. *Ortogonalni komplement* $\mathsf{U}^{\perp}$ *poljubnega* T*-invariantnega vektorskega podprostora* U *vektorskega prostora* V *je* T^**-invarianten.*

Dokaz. Vzemimo poljuben vektor $v \in \mathsf{U}^{\perp}$. Za vsak vektor $u \in \mathsf{U}$ je $Tu \in \mathsf{U}$ in zato

$$\langle u, T^* v\rangle = \langle Tu, v\rangle = 0.$$

To pomeni, da velja $T^* v \in \mathsf{U}^{\perp}$. □

3.2.2. Linearne izometrije. Naj bosta V in W vektorska prostora nad $\mathbb{F}$ s skalarnim produktom. Linearna preslikava $T \in \mathrm{Lin}(\mathsf{V}, \mathsf{W})$ je *linearna izometrija*, če velja

$$\|Tv - Tv'\| = \|v - v'\|$$

za vse vektorje $v, v' \in \mathsf{V}$. Linearna izometrija torej po definiciji ohranja razdalje med vektorji. Če je preslikava T izomorfizem in linearna izometrija, potem pravimo, da je *izometrični izomorfizem*. Najpreprostejši primer izometričnega izomorfizma je identična preslikava $\mathrm{id} : \mathsf{V} \to \mathsf{V}$.

Vaja 3.35. (1) Pokaži, da je kompozicija linearnih izometrij spet linearna izometrija in da je inverz izometričnega izomorfizma spet izometrični izomorfizem.

(2) Naj bosta V in W vektorska prostora nad $\mathbb{F}$ s skalarnim produktom in naj bo $T \in \mathrm{Lin}(\mathsf{V}, \mathsf{W})$ linearna izometrija. Pokaži, da je tedaj T monomorfizem. Če sta ob tem vektorska prostora končno-dimenzionalna in velja $\dim \mathsf{V} = \dim \mathsf{W}$, potem je T izomorfizem.

Naslednja trditev pove, da linearna preslikava ohranja razdalje oziroma normo, če, in samo če, ohranja skalarni produkt:

Trditev 3.36. *Naj bosta V in W vektorska prostora nad $\mathbb{F}$ s skalarnim produktom. Za poljubno linearno preslikavo $T \in \mathrm{Lin}(\mathsf{V}, \mathsf{W})$ so si ekvivalentne naslednje trditve:*

(i) *preslikava T je linearna izometrija,*
(ii) *$\|Tv\| = \|v\|$ za vsak vektor $v \in \mathsf{V}$,*
(iii) *$\langle Tv, Tv' \rangle = \langle v, v' \rangle$ za vse vektorje $v, v' \in \mathsf{V}$.*

Dokaz. Očitno je, da iz (iii) sledi (i) in da iz (i) sledi (ii). Točka (iii) sledi iz točke (ii) po polarizacijski enačbi. □

Trditev 3.37. *Naj bosta V in W končno-dimenzionalna vektorska prostora nad $\mathbb{F}$ s skalarnim produktom in naj bo $T \in \mathrm{Lin}(\mathsf{V}, \mathsf{W})$ linearna preslikava.*

(i) Če je T izometrični izomorfizem in če je $[v_1, v_2, \ldots, v_n]$ ortonormirana baza vektorskega prostora V, potem je $[Tv_1, Tv_2, \ldots, Tv_n]$ ortonormirana baza vektorskega prostora W.

(ii) Če obstaja takšna ortonormirana baza $[v_1, v_2, \ldots, v_n]$ vektorskega prostora V, da je $[Tv_1, Tv_2, \ldots, Tv_n]$ ortonormirana baza vektorskega prostora W, potem je T izometrični izomorfizem.

Dokaz. Točka (i) je direktna posledica trditve 3.36. Dokažimo še točko (ii). Označimo $\mathscr{B} = [v_1, v_2, \ldots, v_n]$. Za poljubna vektorja $v, v' \in \mathsf{V}$ velja

$$\begin{aligned}
\langle Tv, Tv' \rangle &= \left\langle T\Big(\sum_{i=1}^n \langle v, v_i \rangle v_i\Big), T\Big(\sum_{j=1}^n \langle v', v_j \rangle v_j\Big) \right\rangle \\
&= \left\langle \sum_{i=1}^n \langle v, v_i \rangle Tv_i, \sum_{j=1}^n \langle v', v_j \rangle Tv_j \right\rangle \\
&= \sum_{i=1}^n \sum_{j=1}^n \langle v, v_i \rangle \overline{\langle v', v_j \rangle} \langle Tv_i, Tv_j \rangle = \sum_{i=1}^n \langle v, v_i \rangle \overline{\langle v', v_i \rangle} \\
&= \sum_{i=1}^n \sum_{j=1}^n \langle v, v_i \rangle \overline{\langle v', v_j \rangle} \langle v_i, v_j \rangle
\end{aligned}$$

$$= \left\langle \sum_{i=1}^{n} \langle v, v_i \rangle v_i, \sum_{j=1}^{n} \langle v', v_j \rangle v_j \right\rangle = \langle v, v' \rangle. \qquad \square$$

TRDITEV 3.38. *Naj bosta* V *in* W *končno-dimenzionalna vektorska prostora nad* $\mathbb{F}$ *s skalarnim produktom in naj bo* $T \in \mathrm{Lin}(\mathsf{V}, \mathsf{W})$ *linearna preslikava.*

(i) Preslikava T *je linearna izometrija, če, in samo če, velja* $T^*T = \mathrm{id}$.

(ii) Preslikava T *je je izometrični izomorfizem, če, in samo če, je izomorfizem in velja* $T^{-1} = T^*$.

DOKAZ. Iz trditve 3.36 sledi, da je preslikava T linearna izometrija, če, in samo če, velja $\langle v, v' \rangle = \langle Tv, Tv' \rangle = \langle v, T^*Tv' \rangle$ za vse vektorje $v, v' \in \mathsf{V}$. Po trditvi 3.5 je to res natanko tedaj, ko je $T^*T = \mathrm{id}$. $\square$

Spomnimo se, da je matrika U unitarna, če je kvadratna in zanjo velja $U^{\mathsf{h}}U = \mathrm{I}$. Vsaka unitarna matrika U je obrnljiva in velja $U^{-1} = U^{\mathsf{h}}$. Realnim unitarnim matrikam pravimo ortogonalne matrike. Vsaka ortogonalna matrika Q je obrnljiva in velja $Q^{-1} = Q^{\mathsf{t}}$.

TRDITEV 3.39. *Naj bosta* V *in* W *končno-dimenzionalna vektorska prostora nad* $\mathbb{F}$ *s skalarnim produktom in naj bo* $T \in \mathrm{Lin}(\mathsf{V}, \mathsf{W})$ *linearna preslikava.*

(i) Če je preslikava T *izometrični izomorfizem, če je* $\mathscr{B}$ *ortonormirana baza vektorskega prostora* V *in če je* $\mathscr{B}'$ *ortonormirana baza vektorskega prostora* W*, potem je matrika* $[T]^{\mathscr{B}}_{\mathscr{B}'}$ *unitarna.*

(ii) Če obstajata takšna ortonormirana baza $\mathscr{B}$ *vektorskega prostora* V *in takšna ortonormirana baza* $\mathscr{B}'$ *vektorskega prostora* W*, da je matrika* $[T]^{\mathscr{B}}_{\mathscr{B}'}$ *unitarna, potem je* T *izometrični izomorfizem.*

DOKAZ. (i) Ker je T izomorfizem, je $\dim \mathsf{V} = \dim \mathsf{W}$, zato je matrika $[T]^{\mathscr{B}}_{\mathscr{B}'}$ kvadratna. Ker je T linearna izometrija, velja $T^*T = \mathrm{id}$, odtod pa sledi

$$([T]^{\mathscr{B}}_{\mathscr{B}'})^{\mathsf{h}}[T]^{\mathscr{B}}_{\mathscr{B}'} = [T^*]^{\mathscr{B}'}_{\mathscr{B}}[T]^{\mathscr{B}}_{\mathscr{B}'} = [T^*T]^{\mathscr{B}}_{\mathscr{B}} = [\mathrm{id}]^{\mathscr{B}}_{\mathscr{B}} = \mathrm{I}.$$

(ii) Ker je matrika $[T]^{\mathscr{B}}_{\mathscr{B}'}$ unitarna, je posebej kvadratna, zato velja $\dim \mathsf{V} = \dim \mathsf{W}$. Ker velja

$$[T^*T]^{\mathscr{B}}_{\mathscr{B}} = [T^*]^{\mathscr{B}'}_{\mathscr{B}}[T]^{\mathscr{B}}_{\mathscr{B}'} = ([T]^{\mathscr{B}}_{\mathscr{B}'})^{\mathsf{h}}[T]^{\mathscr{B}}_{\mathscr{B}'} = \mathrm{I},$$

sledi $T^*T = \mathrm{id}$. $\square$

KOMENTAR 3.40. (1) Izometričnim izomorfizmom med realnimi vektorskimi prostori s skalarnim produktom pravimo tudi *ortogonalne* preslikave. Množica $\mathsf{O}(\mathsf{V}) \subset \mathrm{End}(\mathsf{V})$ vseh ortogonalnih endomorfizmov realnega vektorskega prostora V je grupa za kompozicijo.

Izometričnim izomorfizmom med kompleksnimi vektorskimi prostori s skalarnim produktom pravimo tudi *unitarne* preslikave. Množica $\mathrm{U}(\mathsf{V}) \subset \mathrm{End}(\mathsf{V})$ vseh unitarnih endomorfizmov kompleksnega vektorskega prostora V je grupa za kompozicijo.

(2) Po trditvi 3.39 je matrika $U \in \mathrm{Mat}(n \times n, \mathbb{C})$ unitarna, če, in samo če, je unitarna linearna preslikava $\ell(U) : \mathbb{C}^n \to \mathbb{C}^n$ glede na standardni skalarni produkt. To je res natanko tedaj, ko stolpci matrike U sestavljajo ortonormirano bazo vektorskega prostora $\mathbb{C}^n$ s standardnim skalarnim produktom. Grupo $\mathrm{U}(\mathbb{C}^n)$ tako lahko glede na standardno bazo identificiramo z unitarno grupo $\mathrm{U}(n)$ stopnje n.

Podobno je realna matrika $Q \in \mathrm{Mat}(n \times n, \mathbb{R})$ ortogonalna, če, in samo če, je ortogonalna linearna preslikava $\ell(Q) : \mathbb{R}^n \to \mathbb{R}^n$ glede na standardni skalarni produkt. To velja natanko tedaj, ko stolpci matrike Q sestavljajo ortonormirano bazo vektorskega prostora $\mathbb{R}^n$ s standardnim skalarnim produktom. Grupo $\mathrm{O}(\mathbb{R}^n)$ tako lahko glede na standardno bazo identificiramo z ortogonalno grupo $\mathrm{O}(n)$ stopnje n.

TRDITEV 3.41. *Naj bo* V *končno-dimenzionalen vektorski prostor pozitivne dimenzije nad* $\mathbb{F}$ *s skalarnim produktom in naj bo* $T \in \mathrm{End}(\mathsf{V})$ *linearna izometrija. Tedaj je preslikava* T *unitarna oziroma ortogonalna,* $|\det(T)| = 1$ *in* $|\lambda| = 1$ *za vsak* $\lambda \in \sigma(T)$.

DOKAZ. Za $\lambda \in \sigma(T)$ obstaja neničelni vektor $v \in \mathsf{V}$, za katerega je $Tv = \lambda v$. Ker je T linearna izometrija, velja

$$\|v\| = \|Tv\| = \|\lambda v\| = |\lambda|\|v\|,$$

odtod pa sledi $|\lambda| = 1$. Preostali del trditve trditve sledi iz enakosti

$$|\det(T)|^2 = \overline{\det(T)}\det(T) = \det(T^*)\det(T) = \det(T^*T) = \det(\mathrm{id}) = 1. \qquad \square$$

Naj bo V končno-dimenzionalen vektorski prostor pozitivne dimenzije nad $\mathbb{F}$ s skalarnim produktom. Endomorfizem $T \in \mathrm{End}(\mathsf{V})$ je *unitarno diagonalizabilen* (v primeru $\mathbb{F} = \mathbb{C}$) oziroma *ortogonalno diagonalizabilen* (v primeru $\mathbb{F} = \mathbb{R}$), če obstaja takšna ortonormirana baza $\mathscr{B}$ vektorskega prostora V, da je matrika $[T]_{\mathscr{B}}^{\mathscr{B}}$ diagonalna. Drugače povedano, endomorfizem T je unitarno (oziroma ortogonalno) diagonalizabilen natanko tedaj, ko obstaja ortonormirana baza vektorskega prostora V, ki je sestavljena iz lastnih vektorjev endomorfizma T.

KOMENTAR 3.42. (1) Matrika $B \in \mathrm{Mat}(n \times n, \mathbb{C})$ je *unitarno podobna* matriki $A \in \mathrm{Mat}(n \times n, \mathbb{C})$, če obstaja takšna unitarna matrika $U \in \mathrm{U}(n)$, da je $B = U^{\mathsf{h}}AU$. Ni težko preveriti, da je tako definirana unitarna podobnost ekvivalenčna relacija.

Matrika $A \in \mathrm{Mat}(n \times n, \mathbb{C})$ je unitarno diagonalizabilna, kadar je pripadajoči endomorfizem $\ell(A)$ vektorskega prostora $\mathbb{C}^n$ s standardnim skalarnim produktom unitarno diagonalizabilen, kar pa je res, če, in samo če, je matrika A unitarno podobna diagonalni kompleksni matriki. V tem primeru torej obstaja takšna unitarna matrika $U \in \mathrm{U}(n)$, da je matrika $U^{\mathsf{h}}AU$ diagonalna. Stolpci matrike U so tedaj lastni vektorji matrike A in sestavljajo ortonormirano bazo vektorskega prostora $\mathbb{C}^n$ glede na standardni skalarni produkt.

(2) Realna matrika $B \in \mathrm{Mat}(n \times n, \mathbb{R})$ je *ortogonalno podobna* realni matriki $A \in \mathrm{Mat}(n \times n, \mathbb{R})$, če obstaja takšna ortogonalna matrika $Q \in \mathrm{O}(n)$, da je $B = Q^{\mathsf{t}}AQ$. Ni težko preveriti, da je tudi ortogonalna podobnost ekvivalenčna relacija.

Matrika $A \in \mathrm{Mat}(n \times n, \mathbb{R})$ je ortogonalno diagonalizabilna, kadar je pripadajoči endomorfizem $\ell(A)$ vektorskega prostora $\mathbb{R}^n$ s standardnim skalarnim produktom ortogonalno diagonalizabilen, kar pa je res, če, in samo če, je matrika A ortogonalno podobna diagonalni realni matriki. V tem primeru torej obstaja takšna ortogonalna matrika $Q \in \mathrm{O}(n)$, da je matrika $Q^{\mathsf{t}}AQ$ diagonalna. Stolpci matrike Q so tedaj lastni vektorji matrike A in sestavljajo ortonormirano bazo vektorskega prostora $\mathbb{R}^n$ glede na standardni skalarni produkt.

VAJA 3.43. Naj bo $T \in \mathrm{End}(\mathsf{V})$ unitarni oziroma ortogonalni endomorfizem končno-dimenzionalnega vektorskega prostora V nad $\mathbb{F}$ s skalarnim produktom. Pokaži, da je ortogonalni komplement $\mathsf{U}^{\perp}$ poljubnega T-invariantnega vektorskega podprostora $\mathsf{U} \subset \mathsf{V}$ prav tako T-invarianten.

ZGLED 3.44. (1) Naj bo Q poljubna ortogonalna matrika velikosti 2×2. Determinanta matrike Q je bodisi 1 ali -1. Ker je prvi stolpec matrike Q normiran vektor, je

$$Q_{\bullet 1} = (\cos\varphi, \sin\varphi)$$

za nek kot $\varphi \in [\![0, 2\pi)\!)$. Ker je drugi stolpec matrike Q normiran in pravokoten na prvi stolpec, velja

$$Q_{\bullet 2} = \pm(-\sin\varphi, \cos\varphi).$$

Če je $\det(Q) = 1$, je torej

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

matrika rotacije za kot φ v ravnini. Če je $\det(Q) = -1$, potem je

$$Q = \begin{bmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi & -\cos\varphi \end{bmatrix} = R_\varphi \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

matrika kompozicije zrcaljenja preko abscisne osi in rotacije za kot φ.

(2) Naj bo U vektorski podprostor končno-dimenzionalnega vektorskega prostora V nad $\mathbb{F}$ s skalarnim produktom. Direktno lahko izračunamo, da je endomorfizem

$$\mathrm{ref}_U^\perp = 2\,\mathrm{pr}_U^\perp - \mathrm{id}_{\mathsf{V}} \in \mathrm{End}(V)$$

izometrija, da velja $(\mathrm{ref}_U^\perp)^2 = \mathrm{id}_{\mathsf{V}}$, $\sigma(\mathrm{ref}_U^\perp) \subset \{1, -1\}$ in

$$\mathrm{E}_{\mathrm{ref}_U^\perp}(1) = \mathsf{U}\,, \qquad \mathrm{E}_{\mathrm{ref}_U^\perp}(-1) = \mathsf{U}^\perp.$$

Izometričnemu izomorfizmu $\mathrm{ref}_U^\perp$ pravimo *(pravokotno) zrcaljenje* čez vektorski podprostor U.

3.2.3. Normalni endomorfizmi. Naj bo V končno-dimenzionalen vektorski prostor nad $\mathbb{F}$ s skalarnim produktom. Endomorfizem $T \in \mathrm{End}(\mathsf{V})$ je *normalen*, če zanj velja $T^*T = TT^*$.

Matrika $A \in \mathrm{Mat}(n \times n, \mathbb{F})$ je normalna, kadar velja $A^{\mathsf{h}}A = AA^{\mathsf{h}}$, kar pa je res, če, in samo če, je pripadajoči endomorfizem $\ell(A)$ vektorskega prostora $\mathbb{F}^n$ s standardnim skalarnim produktom normalen.

ZGLED 3.45. Vsaka diagonalna matrika je normalna. Vsaka unitarna oziroma ortogonalna matrika je normalna. Vsak unitaren oziroma ortogonalen endomorfizem končno-dimenzionalnega vektorskega prostora s skalarnim produktom je normalen.

Iz trditve 3.30 direktno sledi:

TRDITEV 3.46. *Naj bo $T \in \mathrm{End}(\mathsf{V})$ endomorfizem končno-dimenzionalnega vektorskega prostora nad $\mathbb{F}$ s skalarnim produktom.*

(i) Če je endomorfizem T normalen, potem je za vsako ortonormirano bazo $\mathscr{B}$ vektorskega prostora V matrika $[T]_{\mathscr{B}}^{\mathscr{B}}$ normalna.

(ii) Če je $\mathscr{B}$ takšna ortonormirana baza vektorskega prostora V, da je matrika $[T]_{\mathscr{B}}^{\mathscr{B}}$ normalna, potem je endomorfizem T normalen.

TRDITEV 3.47. *Naj bo V končno-dimenzionalen vektorski prostor nad $\mathbb{F}$ s skalarnim produktom in naj bo $T \in \mathrm{End}(\mathsf{V})$ normalen endomorfizem vektorskega prostora V. Ortogonalni komplement $\mathrm{E}_T(\lambda)^\perp$ je T-invarianten vektorski podprostor, za vsak skalar $\lambda \in \mathbb{F}$. Za poljubna različna skalarja $\lambda, \mu \in \mathbb{F}$ velja $\mathrm{E}_T(\mu) \subset \mathrm{E}_T(\lambda)^\perp$.*

Dokaz. Najprej opazimo, da velja

$$\mathrm{E}_T(\lambda)^{\perp} = (\ker(T - \lambda\,\mathrm{id}))^{\perp} = \mathrm{im}((T - \lambda\,\mathrm{id})^*) = \mathrm{im}(T^* - \bar{\lambda}\,\mathrm{id}).$$

Poljuben vektor $v \in \mathrm{E}_T(\lambda)^{\perp}$ je torej oblike $v = (T^* - \bar{\lambda}\,\mathrm{id})u$ za nek vektor $u \in \mathsf{V}$. Odtod sledi

$$Tv = T(T^* - \bar{\lambda}\,\mathrm{id})u = (TT^* - \bar{\lambda}T)u = (T^*T - \bar{\lambda}T)u = (T^* - \bar{\lambda}\,\mathrm{id})Tu,$$

zato je $Tv \in \mathrm{im}(T^* - \bar{\lambda}\,\mathrm{id}) = \mathrm{E}_T(\lambda)^{\perp}$.

Poljuben vektor $w \in \mathrm{E}_T(\mu)$ lahko zapišemo kot vsoto $w = w' + w''$, kjer je $w' \in \mathrm{E}_T(\lambda)$ in $w'' \in \mathrm{E}_T(\lambda)^{\perp}$. Zdaj velja $\mu w' + \mu w'' = \mu w = Tw = Tw' + Tw'' = \lambda w' + Tw''$, odtod pa sledi

$$(\lambda - \mu)w' = \mu w'' - Tw'' \in \mathrm{E}_T(\lambda) \cap \mathrm{E}_T(\lambda)^{\perp} = \{0\},$$

zato velja $w' = 0$ in torej $w = w'' \in \mathrm{E}_T(\lambda)^{\perp}$. □

Izrek 3.48. *Naj bo* V *končno-dimenzionalen kompleksen vektorski prostor pozitivne dimenzije s skalarnim produktom. Vsak normalen endomorfizem vektorskega prostora* V *je unitarno diagonalizabilen.*

Dokaz. Trditev bomo dokazali z indukcijo na dimenzijo $n = \dim \mathsf{V}$. V primeru $n = 1$ ni kaj dokazovati. Naj bo torej $n \geq 2$ in predpostavimo, da izrek velja za vse vektorske prostore dimenzije manjše od n.

Ker je V kompleksen vektorski prostor, ima endomorfizem T vsaj eno lastno vrednost $\lambda \in \mathbb{C}$. Lastni podprostor $\mathrm{E}_T(\lambda)$ je T-invarianten, po trditvi 3.47 pa je ortogonalni komplement $\mathrm{E}_T(\lambda)^{\perp}$ prav tako T-invarianten. Po trditvi 3.34 sta tedaj vektorska podprostora $\mathrm{E}_T(\lambda)$ in $\mathrm{E}_T(\lambda)^{\perp}$ tudi T^*-invariantna. Direktno iz definicije adjungirane preslikave vidimo, da za zožitev $T|_{\mathrm{E}_T(\lambda)^{\perp}} : \mathrm{E}_T(\lambda)^{\perp} \to \mathrm{E}_T(\lambda)^{\perp}$ velja $(T|_{\mathrm{E}_T(\lambda)^{\perp}})^* = T^*|_{\mathrm{E}_T(\lambda)^{\perp}}$, odtod pa sledi, da je

$$T|_{\mathrm{E}_T(\lambda)^{\perp}} \in \mathrm{End}(\mathrm{E}_T(\lambda)^{\perp})$$

normalen endomorfizem vektorskega prostora $\mathrm{E}_T(\lambda)^{\perp}$. Ker je $\dim \mathrm{E}_T(\lambda)^{\perp} < n$, po indukcijski hipotezi obstajajo ortonormirana baza

$$[u_1, u_2, \ldots, u_q]$$

vektorskega prostora $\mathrm{E}_T(\lambda)^{\perp}$, ki je sestavljena iz lastnih vektorjev endomorfizma $T|_{\mathrm{E}_T(\lambda)^{\perp}}$. Posebej so vektorji $u_1, u_2, \ldots, u_q$ tudi lastni vektorji endomorfizma T. Izberimo še vektorje

$$v_1, v_2, \ldots, v_p \in \mathrm{E}_T(\lambda),$$

ki sestavljajo ortonormirano bazo vektorskega prostora $\mathrm{E}_T(\lambda)$. Tedaj je

$$[v_1, v_2, \ldots, v_p, u_1, u_2, \ldots, u_q]$$

ortonormirana baza vektorskega prostora V, sestavljena iz lastnih vektorjev endomorfizma T. □

Komentar 3.49. Za vsak normalen endomorfizem $T \in \mathrm{End}(\mathsf{V})$ kompleksnega vektorskega prostora V s skalarnim produktom je torej vektorski prostor V direktna vsota lastnih podprostorov,

$$\mathsf{V} = \mathrm{E}_T(\lambda_1) \oplus \mathrm{E}_T(\lambda_2) \oplus \cdots \oplus \mathrm{E}_T(\lambda_m),$$

kjer so $\lambda_1, \lambda_2, \ldots, \lambda_m \in \mathbb{C}$ vse paroma različne lastne vrednosti endomorfizma T. Po trditvi 3.47 je $\mathrm{E}_T(\lambda_i) \subset \mathrm{E}_T(\lambda_j)^\perp$ za poljubna različna indeksa i in j. Drugače povedano, lastni podprostori $\mathrm{E}_T(\lambda_1), \mathrm{E}_T(\lambda_2), \ldots, \mathrm{E}_T(\lambda_m)$ so med seboj paroma pravokotni. Odtod tudi sledi, da za vsak $i = 1, 2, \ldots, m$ velja

$$\mathrm{E}_T(\lambda_i)^\perp = \mathrm{E}_T(\lambda_1) \oplus \cdots \oplus \mathrm{E}_T(\lambda_{i-1}) \oplus \mathrm{E}_T(\lambda_{i+1}) \oplus \cdots \oplus \mathrm{E}_T(\lambda_m).$$

Zaradi tega pravimo, da je direktna vsota $\mathsf{V} = \mathrm{E}_T(\lambda_1) \oplus \mathrm{E}_T(\lambda_2) \oplus \cdots \oplus \mathrm{E}_T(\lambda_m)$ *ortogonalna*.

TRDITEV 3.50. *(i) Vsaka realna matrika, ki je ortogonalno podobna normalni realni matriki, je normalna.*

(ii) Vsaka kompleksna matrika, ki je unitarno podobna normalni kompleksni matriki, je normalna.

(iii) Kvadratna kompleksna matrika je normalna, če, in samo če, je unitarno podobna diagonalni matriki.

DOKAZ. Naj bo matrika $A \in \mathrm{Mat}(n \times n, \mathbb{C})$ normalna in naj bo matrika $B \in \mathrm{Mat}(n \times n, \mathbb{C})$ unitarno podobna matriki A. Tedaj obstaja takšna unitarna matrika $U \in \mathrm{U}(n)$, da je $B = U^\mathsf{h} A U$. Ker je matrika A normalna, velja $A^\mathsf{h} A = A A^\mathsf{h}$, odtod pa sledi

$$\begin{aligned} B^\mathsf{h} B &= (U^\mathsf{h} A U)^\mathsf{h} (U^\mathsf{h} A U) = U^\mathsf{h} A^\mathsf{h} U U^\mathsf{h} A U = U^\mathsf{h} A^\mathsf{h} A U \\ &= U^\mathsf{h} A A^\mathsf{h} U = U^\mathsf{h} A U U^\mathsf{h} A^\mathsf{h} U = (U^\mathsf{h} A U)(U^\mathsf{h} A U)^\mathsf{h} = B B^\mathsf{h}. \end{aligned}$$

Matrika B je torej normalna. S tem smo dokazali točko (ii). Točko (i) dokažemo podobno. Točka (iii) sledi direktno iz točke (ii) in iz izreka 3.48. □

VAJA 3.51. Naj bosta V in W končno-dimenzionalna vektorska prostora nad $\mathbb{F}$ s skalarnim produktom in naj bo $\Theta \in \mathrm{Lin}(\mathsf{V}, \mathsf{W})$ izometrični izomorfizem. Pokaži, da je za vsak normalen endomorfizem $T \in \mathrm{End}(\mathsf{V})$ tudi endomorfizem $\Theta T \Theta^* \in \mathrm{End}(\mathsf{W})$ normalen.

TRDITEV 3.52. *Naj bo* V *končno-dimenzionalen vektorski prostor nad* $\mathbb{F}$ *pozitivne dimenzije. Za poljuben diagonalizabilen endomorfizem* $T \in \mathrm{End}(\mathsf{V})$ *obstaja skalarni produkt na vektorskem prostoru* V*, glede na katerega je* T *normalen endomorfizem.*

DOKAZ. Naj bo $\mathscr{B}$ baza vektorskega prostora V, sestavljena iz lastnih vektorjev endomorfizma T. Obstaja natanko en skalarni produkt na vektorskem prostoru V, glede na katerega je baza $\mathscr{B}$ ortonormirana. Tedaj je matrika $[T]_{\mathscr{B}}^{\mathscr{B}}$ diagonalna in torej normalna, odtod pa sledi, da je endomorfizem T normalen glede na izbrani skalarni produkt. □

ZGLED 3.53. (1) Naj bo $Q \in \mathrm{O}(n)$ ortogonalna matrika. Realno matriko Q lahko gledamo tudi kot kompleksno unitarno matriko. Vse kompleksne lastne vrednosti matrike Q imajo absolutno vrednost 1. Za poljubno kompleksno lastno vrednost λ matrike Q je tudi njena konjugirana vrednost lastna vrednost matrike Q. Še več, če je $w \in \mathbb{C}^n$ lastni vektor matrike Q za lastno vrednost λ, potem velja $Qw = \lambda w$ in zato tudi $Q\bar{w} = \bar{\lambda}\bar{w}$, torej je konjugiran vektor $\bar{w} \in \mathbb{C}^n$ lastni vektor matrike Q za lastno vrednost $\bar{\lambda}$. Če vektorje $w_1, w_2, \ldots, w_k \in \mathbb{C}^n$ izberemo tako, da ti sestavljajo ortonormirano bazo kompleksnega lastnega podprostora matrike Q pri lastni vrednosti λ, potem vektorji $\bar{w}_1, \bar{w}_2, \ldots, \bar{w}_k \in \mathbb{C}^n$ sestavljajo ortonormirano bazo kompleksnega lastnega podprostora matrike Q pri lastni vrednosti $\bar{\lambda}$.

Ker je matrika Q ortogonalna, je tudi normalna, zato je po izreku 3.48 unitarno diagonalizabilna kot kompleksna matrika. Po zgoraj povedanem lahko izberemo ortonormirano bazo $\mathscr{B}$ vektorskega prostora $\mathbb{C}^n$, ki je oblike

$$\mathscr{B} = [v_1^+, \ldots, v_p^+, w_1, \bar{w}_1, \ldots, w_s, \bar{w}_s, v_1^-, \ldots, v_q^-]$$

in ob tem velja:

(i) vektorji $v_1^+, \ldots, v_p^+ \in \mathbb{R}^n$ sestavljajo bazo realnega lastnega podprostora matrike Q pri lastni vrednosti 1,
(ii) vektorji $v_1^-, \ldots, v_q^- \in \mathbb{R}^n$ sestavljajo bazo realnega lastnega podprostora matrike Q pri lastni vrednosti -1, in
(iii) za vsak indeks $j = 1, 2, \ldots, s$ je vektor w_j lastni vektor za neko kompleksno lastno vrednost λ_j matrike Q z negativnim imaginarnim delom.

Iz točke (iii) sledi, da je za vsak indeks $j = 1, 2, \ldots, s$ vektor $\bar{w}_j$ lastni vektor za kompleksno lastno vrednost $\bar{\lambda}_j$ matrike Q. Pri tem seveda velja $p, s, q \in \{0, 1, \ldots, n\}$ in $p + 2s + q = n$. Če je katero od števil p, s, q enako številu 0, to pomeni, da ustrezni vektorji v bazi $\mathscr{B}$ ne nastopajo. Koordinatna matrika endomorfizma $\ell(Q)$ v ortonormirani bazi $\mathscr{B}$ je diagonalna kompleksna matrika oblike

$$[\mathrm{id}]^{\mathscr{E}}_{\mathscr{B}}\, Q\, [\mathrm{id}]^{\mathscr{B}}_{\mathscr{E}} = [\ell(Q)]^{\mathscr{B}}_{\mathscr{B}} = \mathrm{diag}(1, \ldots, 1, \lambda_1, \bar{\lambda}_1, \ldots, \lambda_s, \bar{\lambda}_s, -1, \ldots, -1).$$

Za vsak $j = 1, 2, \ldots, s$ naj bo $\varphi_j \in (\!| 0, \pi |\!)$ tisti kot, za katerega je $\lambda_j = \mathrm{e}^{-\mathrm{i}\varphi_j}$, in naj bosta

$$u_j = \tfrac{1}{\sqrt{2}}(w_j + \bar{w}_j), \qquad u_j' = \tfrac{1}{\mathrm{i}\sqrt{2}}(w_j - \bar{w}_j)$$

normirani realni in normirani imaginarni del vektorja w_j. Posebej to pomeni, da sta $u_j, u_j' \in \mathbb{R}^n$ vektorja z realnimi komponentami in da velja

$$w_j = \tfrac{1}{\sqrt{2}}(u_j + \mathrm{i}u_j'), \qquad \bar{w}_j = \tfrac{1}{\sqrt{2}}(u_j - \mathrm{i}u_j').$$

Ker vektorja $w_j, \bar{w}_j$ sestavljata ortonormiran sistem, tudi vektorja u_j, u_j' sestavljata ortonormiran sistem, kar lahko enostavno preverimo. Poleg tega velja

$$\mathrm{Span}_{\mathbb{C}}\{u_j, u_j'\} = \mathrm{Span}_{\mathbb{C}}\{w_j, \bar{w}_j\}.$$

Tako smo dobili novo ortonormirano bazo

$$\mathscr{B}' = [v_1^+, \ldots, v_p^+, u_1, u_1', \ldots, u_s, u_s', v_1^-, \ldots, v_q^-]$$

vektorskega prostora $\mathbb{C}^n$, ki pa je sestavljena iz vektorjev z realnimi komponentami. Baza $\mathscr{B}'$ je zato tudi ortonormirana baza realnega vektorskega prostora $\mathbb{R}^n$.

Izračunajmo koordinatno matriko endomorfizma $\ell(Q)$ v ortonormirani bazi $\mathscr{B}'$. Za vsak indeks $j = 1, 2, \ldots, n$ velja

$$\begin{aligned}
Qu_j &= \tfrac{1}{\sqrt{2}}(Qw_j + Q\bar{w}_j) = \tfrac{1}{\sqrt{2}}(\lambda_j w_j + \bar{\lambda}_j \bar{w}_j)\\
&= \tfrac{1}{\sqrt{2}}\big(\lambda_j \tfrac{1}{\sqrt{2}}(u_j + \mathrm{i}u_j') + \bar{\lambda}_j \tfrac{1}{\sqrt{2}}(u_j - \mathrm{i}u_j')\big)\\
&= \frac{1}{2}\big((\lambda_j + \bar{\lambda}_j)u_j + \mathrm{i}(\lambda_j - \bar{\lambda}_j)u_j'\big)\\
&= \Re(\lambda_j)u_j - \Im(\lambda_j)u_j' = \cos(\varphi_j)u_j + \sin(\varphi_j)u_j'.
\end{aligned}$$

Podobno dobimo

$$\begin{aligned}
Qu_j' &= \tfrac{1}{\mathrm{i}\sqrt{2}}(Qw_j - Q\bar{w}_j) = \tfrac{1}{\mathrm{i}\sqrt{2}}(\lambda_j w_j - \bar{\lambda}_j \bar{w}_j)\\
&= \tfrac{1}{\mathrm{i}\sqrt{2}}\big(\lambda_j \tfrac{1}{\sqrt{2}}(u_j + \mathrm{i}u_j') - \bar{\lambda}_j \tfrac{1}{\sqrt{2}}(u_j - \mathrm{i}u_j')\big)
\end{aligned}$$

$$\begin{aligned}
&= \frac{1}{2\mathrm{i}}\big((\lambda_j - \bar{\lambda}_j)u_j + \mathrm{i}(\lambda_j + \bar{\lambda}_j)u_j'\big) \\
&= \Im(\lambda_j)u_j + \Re(\lambda_j)u_j' = -\sin(\varphi_j)u_j + \cos(\varphi_j)u_j'.
\end{aligned}$$

Odtod sledi, da je koordinatna matrika $[\ell(Q)]_{\mathscr{B}'}^{\mathscr{B}'}$ bločno diagonalna oblike

$$[\mathrm{id}]_{\mathscr{B}'}^{\mathscr{E}}\, Q\, [\mathrm{id}]_{\mathscr{E}}^{\mathscr{B}'} = [\ell(Q)]_{\mathscr{B}'}^{\mathscr{B}'} = \begin{bmatrix} \mathrm{I}_{p\times p} & & & & \\ & R_{\varphi_1} & & & \\ & & \ddots & & \\ & & & R_{\varphi_s} & \\ & & & & -\mathrm{I}_{q\times q} \end{bmatrix}$$

kjer so $R_{\varphi_1}, \dots, R_{\varphi_s}$ rotacijske matrike:

$$R_{\varphi_j} = \begin{bmatrix} \cos\varphi_j & -\sin\varphi_j \\ \sin\varphi_j & \cos\varphi_j \end{bmatrix}$$

Pokazali smo torej, da je vsaka ortogonalna matrika Q ortogonalno podobna bločno diagonalni realni matriki takšne oblike. Če je katero od števil p, s, q enako številu 0, potem v zgornjem bločnem zapisu ustrezen blok ne nastopa.

Ob tem lahko še opazimo, da je $\mathrm{I}_{2\times 2} = R_0$ rotacijska matrika za kot 0 in da je $-\mathrm{I}_{2\times 2} = R_\pi$ rotacijska matrika za kot π. Matriko $[\ell(Q)]_{\mathscr{B}'}^{\mathscr{B}'}$ lahko torej zapišemo tudi v obliki

$$\begin{bmatrix} \mathrm{I}_{p'\times p'} & & & & \\ & R_{\vartheta_1} & & & \\ & & \ddots & & \\ & & & R_{\vartheta_r} & \\ & & & & -\mathrm{I}_{q'\times q'} \end{bmatrix}$$

kjer sta p' in q' števili iz možice $\{1, 0\}$, kjer je $p' + 2r + q' = n$ in kjer je

$$(\vartheta_1, \dots, \vartheta_r) = (0, \dots, 0, \varphi_1, \dots, \varphi_s, \pi, \dots, \pi).$$

Ob tem velja $\det(Q) = (-1)^q = (-1)^{q'}$.

Če je $\det(Q) = 1$, potem torej ortogonalna matrika Q predstavlja kombinacijo rotacij v med seboj pravokotnih ravninah. Če je $\det(Q) = -1$, potem ortogonalna matrika Q predstavlja kombinacijo rotacij v med seboj pravokotnih ravninah in zrcaljena čez hiperravnino. Če je n liho število in je $\det(Q) = 1$, potem je $p' = 1$ in $q' = 0$. Če je n liho število in je $\det(Q) = -1$, potem je $p' = 0$ in $q' = 1$.

(2) Naj bo $Q \in \mathrm{O}(3)$ ortogonalna matrika z determinanto 1. Po točki (1) je število 1 lastna vrednost matrike Q in obstaja takšna ortonormirana baza $\mathscr{B}' = [v_1, v_2, v_3]$ vektorskega prostora $\mathbb{R}^3$, da je

$$[\ell(Q)]_{\mathscr{B}'}^{\mathscr{B}'} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\varphi & -\sin\varphi \\ 0 & \sin\varphi & \cos\varphi \end{bmatrix}$$

za nek kot $\varphi \in [\![\, 0, \pi \,]\!]$. Drugače povedano, matrika Q predstavlja rotacijo v vektorskem prostoru $\mathbb{R}^3$ za kot φ okoli osi $\mathbb{R}v_1$.

Vaja 3.54. Naj bo $u \in \mathbb{R}^3$ normiran vektor in naj bo $T : \mathbb{R}^3 \to \mathbb{R}^3$ linearna preslikava, dana s predpisom

$$Tv = (u \cdot v)u + u \times v.$$

Pokaži, da je T rotacija okoli osi $\mathbb{R}u$ za kot $\pi/2$.

3.2.4. Sebi-adjungirani endomorfizmi. Naj bo V končno-dimenzionalen vektorski prostor nad $\mathbb{F}$ s skalarnim produktom. Endomorfizem $T \in \mathrm{End}(\mathsf{V})$ je *sebi-adjungiran*, če zanj velja $T^* = T$.

Če je endomorfizem $T \in \mathrm{End}(\mathsf{V})$ sebi-adjungiran, potem je tudi normalen, saj je tedaj $T^*T = TT = TT^*$.

Zgled 3.55. (1) Endomorfizem $\ell(A)$ vektorskega prostora $\mathbb{F}^n$ s standardnim skalarnim produktom, dan s kvadratno matriko A, je sebi-adjungiran, če, in samo če, je matrika A hermitska. Za realno kvadratno matriko A je to res natanko tedaj, ko je matrika A simetrična.

(2) Naj bo V končno-dimenzionalen vektorski prostor nad $\mathbb{F}$ s skalarnim produktom in naj bo $T \in \mathrm{End}(\mathsf{V})$ projektor. Opazimo lahko, da je tedaj tudi adjungiran endomorfizem $T^* \in \mathrm{End}(\mathsf{V})$ projektor, saj velja $(T^*)^2 = (T^2)^* = T^*$. Ob tem je

$$\ker(T^*) = (\mathrm{im}(T))^\perp, \qquad \mathrm{im}(T^*) = (\ker(T))^\perp.$$

Projektor T je zato sebi-adjungiran, če, in samo če, velja $\ker(T) = (\mathrm{im}(T))^\perp$.

Iz trditve 3.30 direktno sledi:

Trditev 3.56. *Naj bo T endomorfizem končno-dimenzionalnega vektorskega prostora nad $\mathbb{F}$ s skalarnim produktom.*

(i) Če je endomorfizem T sebi-adjungiran, potem je za vsako ortonormirano bazo $\mathscr{B}$ vektorskega prostora V matrika $[T]^{\mathscr{B}}_{\mathscr{B}}$ hermitska.

(ii) Če je $\mathscr{B}$ takšna ortonormirana baza vektorskega prostora V, da je matrika $[T]^{\mathscr{B}}_{\mathscr{B}}$ hermitska, potem je endomorfizem T sebi-adjungiran.

Trditev 3.57. *Naj bosta T in S sebi-adjungirana endomorfizma končno-dimenzionalnega vektorskega prostora V nad $\mathbb{F}$ s skalarnim produktom in naj bo $\alpha \in \mathbb{R}$ poljubno realno število. Tedaj sta tudi endomorfzma $S+T$ in αT sebi-adjungirana. Sebi-adjungirani endomorfizmi vektorskega prostora V sestavljajo realen vektorski prostor.*

Dokaz. Velja $(S+T)^* = S^* + T^* = S + T$ in $(\alpha T)^* = \bar{\alpha} T^* = \alpha T$. □

Vaja 3.58. Pokaži, da velja:

(i) Vsaka realna matrika, ki je ortogonalno podobna simetrični realni matriki, je simetrična.

(ii) Vsaka kompleksna matrika, ki je unitarno podobna hermitski kompleksni matriki, je hermitska.

(iii) Naj bosta V in W končno-dimenzionalna vektorska prostora nad $\mathbb{F}$ s skalarnim produktom in naj bo $\Theta \in \mathrm{Lin}(\mathsf{V}, \mathsf{W})$ izometrični izomorfizem. Pokaži, da je za vsak sebi-adjungiran endomorfizem $T \in \mathrm{End}(\mathsf{V})$ tudi endomorfizem $\Theta T \Theta^* \in \mathrm{End}(\mathsf{W})$ sebi-adjungiran.

Trditev 3.59. *Za vsak sebi-adjungiran endomorfizem T končno-dimenzionalnega vektorskega prostora V nad $\mathbb{F}$ pozitivne dimenzije s skalarnim produktom velja:*

(i) *za vsak vektor $v \in \mathsf{V}$ je skalarni produkt $\langle Tv, v\rangle$ realno število,*
(ii) *vse lastne vrednosti endomorfizma T so realne,*
(iii) *determinanta in sled endomorfizma T sta realni števili, in*
(iv) *karakteristični polinom p_T endomorfzma T je realen in razcepen na realne linearne faktorje.*

DOKAZ. (i) Velja $\langle Tv, v\rangle = \langle v, T^*v\rangle = \langle v, Tv\rangle = \overline{\langle Tv, v\rangle}$.

(ii) Za poljubno lastno vrednosti $\lambda \in \sigma(T)$ obstaja neničelen vektor $v \in \mathrm{E}_T(\lambda)$, in ker velja

$$\lambda\langle v, v\rangle = \langle \lambda v, v\rangle = \langle Tv, v\rangle = \langle v, T^*v\rangle = \langle v, Tv\rangle = \langle v, \lambda v\rangle = \bar{\lambda}\langle v, v\rangle,$$

sledi $\lambda = \bar{\lambda}$.

(iv) Izberimo ortonormirano bazo $\mathscr{B}$ vektorskega prostora V. Karakteristični polinom p_T je enak karakterističnemu polinomu p_A matrike $A = [T]_{\mathscr{B}}^{\mathscr{B}}$. Matrika A je hermitska in ima zato po točki (ii) same realne lastne vrednosti. Odtod sledi, da ima karakteristični polinom p_A same realne ničle in je zato realen ter razcepen na realne linearne faktorje.

Točka (iii) je direktna posledica točke (iv). □

Ker je vsak sebi-adjungiran endomorfizem normalen, iz trditve 3.59(ii) in izreka 3.48 direktno sledi:

POSLEDICA 3.60. *Naj bo* V *končno-dimenzionalen kompleksen vektorski prostor pozitivne dimenzije s skalarnim produktom. Za poljuben sebi-adjungiran endomorfizem* $T \in \mathrm{End}(\mathsf{V})$ *obstaja takšna ortonormirana baza* $\mathscr{B}$ *vektorskega prostora* V*, da je matrika* $[T]_{\mathscr{B}}^{\mathscr{B}}$ *realna diagonalna matrika.*

IZREK 3.61. *Naj bo* V *končno-dimenzionalen realen vektorski prostor pozitivne dimenzije s skalarnim produktom. Vsak sebi-adjungiran endomorfizem* $T \in \mathrm{End}(\mathsf{V})$ *je ortogonalno diagonalizabilen.*

DOKAZ. Dokaz je podoben dokazu izreka 3.48 in ga dokažemo z indukcijo na dimenzijo $n = \dim \mathsf{V}$. V primeru $n = 1$ ni kaj dokazovati. Naj bo torej $n \geq 2$ in predpostavimo, da izrek velja za vsak vektorski prostor dimenzije manjše od n.

Po trditvi 3.59(ii) ima endomorfizem T vsaj eno realno lastno vrednost $\lambda \in \mathbb{R}$. Ker je endomorfizem T sebi-adjungiran, je tudi normalen. Lastni podprostor $\mathrm{E}_T(\lambda)$ je T-invarianten, po trditvi 3.47 pa je ortogonalni komplement $\mathrm{E}_T(\lambda)^\perp$ prav tako T-invarianten. Iz definicije adjungirane preslikave vidimo, da za zožitev $T|_{\mathrm{E}_T(\lambda)^\perp} : \mathrm{E}_T(\lambda)^\perp \to \mathrm{E}_T(\lambda)^\perp$ velja $(T|_{\mathrm{E}_T(\lambda)^\perp})^* = T^*|_{\mathrm{E}_T(\lambda)^\perp} = T|_{\mathrm{E}_T(\lambda)^\perp}$, odtod pa sledi, da je

$$T|_{\mathrm{E}_T(\lambda)^\perp} \in \mathrm{End}(\mathrm{E}_T(\lambda)^\perp)$$

sebi-adjungiran endomorfizem vektorskega prostora $\mathrm{E}_T(\lambda)^\perp$. Ker je $\dim \mathrm{E}_T(\lambda)^\perp < n$, po indukcijski hipotezi obstaja ortonormirana baza $[u_1, u_2, \ldots, u_q]$ vektorskega prostora $\mathrm{E}_T(\lambda)^\perp$, sestavljena iz lastnih vektorjev endomorfizma $T|_{\mathrm{E}_T(\lambda)^\perp}$. Vektorji $u_1, u_2, \ldots, u_q$ so tudi lastni vektorji endomorfizma T. Izberemo še vektorje $v_1, v_2, \ldots, v_p \in \mathrm{E}_T(\lambda)$, ki sestavljajo ortonormirano bazo vektorskega prostora $\mathrm{E}_T(\lambda)$, in dobimo ortonormirano bazo

$$[v_1, v_2, \ldots, v_p, u_1, u_2, \ldots, u_q]$$

vektorskega prostora V, ki je sestavljena iz lastnih vektorjev endomorfizma T. □

Iz posledice 3.60 in izreka 3.61 direktno sledi:

POSLEDICA 3.62. *(i) Vsaka realna simetrična matrika je ortogonalno podobna realni diagonalni matriki.*

(ii) Vsaka kompleksna hermitska matrika je unitarno podobna realni diagonalni matriki.

3.2.5. Pozitivno definitni endomorfizmi. Naj bo V končno-dimenzionalen vektorski prostor nad $\mathbb{F}$ s skalarnim produktom. Spomnimo se, da za vsak sebi-adjungiran endomorfizem $T \in \operatorname{End}(\mathsf{V})$ in za vsak vektor $v \in \mathsf{V}$ velja $\langle Tv, v\rangle \in \mathbb{R}$. Endomorfizem $T \in \operatorname{End}(\mathsf{V})$ je *pozitivno definiten*, če je sebi-adjungiran in velja

$$\langle Tv, v\rangle > 0$$

za vsak neničelen vektor $v \in \mathsf{V}$. V tem primeru zapišemo $T > 0$.

Matrika $A \in \operatorname{Mat}(n\times n, \mathbb{F})$ je pozitivno definitna, kadar je pripadajoči endomorfizem $\ell(A)$ vektorskega prostora $\mathbb{F}^n$ s standardnim skalarnim produktom pozitivno definiten, kar pa je res, če, in samo če, je matrika A hermitska in velja $v^{\mathsf{h}}Av > 0$ za vsak neničelen vektor $v \in \mathbb{F}^n$. V tem primeru zapišemo $A > 0$.

Trditev 3.63. *(i) Vsaka realna matrika, ki je ortogonalno podobna pozitivno definitni realni matriki, je pozitivno definitna.*

(ii) Vsaka kompleksna matrika, ki je unitarno podobna pozitivno definitni kompleksni matriki, je pozitivno definitna.

Dokaz. Naj bo matrika $A \in \operatorname{Mat}(n \times n, \mathbb{C})$ pozitivno definitna in naj bo matrika $B \in \operatorname{Mat}(n \times n, \mathbb{C})$ unitarno podobna matriki A. Tedaj obstaja takšna unitarna matrika $U \in \mathrm{U}(n)$, da je $B = U^{\mathsf{h}}AU$. Ker je matrika A pozitivno definitna, je hermitska, zato je tudi matrika B hermitska. Poleg tega za poljuben neničelen vektor $v \in \mathbb{C}^n$ velja

$$v^{\mathsf{h}}Bv = v^{\mathsf{h}}U^{\mathsf{h}}AUv = (Uv)^{\mathsf{h}}A(Uv) > 0,$$

saj je matrika A pozitivno definitna in $Uv \neq 0$. S tem smo dokazali točko (ii). Točko (i) dokažemo podobno. $\square$

Trditev 3.64. *Naj bo T endomorfizem končno-dimenzionalnega vektorskega prostora nad $\mathbb{F}$ s skalarnim produktom.*

(i) Če je endomorfizem T pozitivno definiten, potem je za vsako ortonormirano bazo $\mathscr{B}$ vektorskega prostora V matrika $[T]^{\mathscr{B}}_{\mathscr{B}}$ pozitivno definitna.

(ii) Če je $\mathscr{B}$ takšna ortonormirana baza vektorskega prostora V, da je matrika $[T]^{\mathscr{B}}_{\mathscr{B}}$ pozitivno definitna, potem je endomorfizem T pozitivno definiten.

Dokaz. Obe točki sta posledici enakosti

$$\begin{aligned}
\langle Tv, v\rangle &= \left\langle T\Big(\sum_{j=1}^{n}\langle v, v_j\rangle v_j\Big), \sum_{i=1}^{n}\langle v, v_i\rangle v_i\right\rangle = \left\langle \sum_{j=1}^{n}\langle v, v_j\rangle Tv_j, \sum_{i=1}^{n}\langle v, v_i\rangle v_i\right\rangle \\
&= \sum_{j=1}^{n}\sum_{i=1}^{n}\langle v, v_j\rangle\overline{\langle v, v_i\rangle}\langle Tv_j, v_i\rangle = \sum_{j=1}^{n}\sum_{i=1}^{n}\big([v]_{\mathscr{B}}\big)_{j1}\overline{\big([v]_{\mathscr{B}}\big)_{i1}}\big([T]^{\mathscr{B}}_{\mathscr{B}}\big)_{ij} \\
&= \sum_{i=1}^{n}\overline{\big([v]_{\mathscr{B}}\big)_{i1}}\sum_{j=1}^{n}\big([T]^{\mathscr{B}}_{\mathscr{B}}\big)_{ij}\big([v]_{\mathscr{B}}\big)_{j1} = \sum_{i=1}^{n}\overline{\big([v]_{\mathscr{B}}\big)_{i1}}\big([T]^{\mathscr{B}}_{\mathscr{B}}[v]_{\mathscr{B}}\big)_{i1} \\
&= \big([v]_{\mathscr{B}}\big)^{\mathsf{h}}[T]^{\mathscr{B}}_{\mathscr{B}}[v]_{\mathscr{B}}
\end{aligned}$$

in trditve 3.56. $\square$

Vaja 3.65. (1) Naj bosta V in W končno-dimenzionalna vektorska prostora nad $\mathbb{F}$ s skalarnim produktom in naj bo $\Theta \in \operatorname{Lin}(\mathsf{V}, \mathsf{W})$ izometrični izomorfizem. Pokaži, da je za vsak pozitivno definiten endomorfizem $T \in \operatorname{End}(\mathsf{V})$ tudi endomorfizem $\Theta T\Theta^* \in \operatorname{End}(\mathsf{W})$ pozitivno definiten.

(2) Naj bosta T in S pozitivno definitna endomorfizma končno-dimenzionalnega vektorskega prostora V nad $\mathbb{F}$ s skalarnim produktom in naj bo α poljubno pozitivno realno število. Pokaži, da sta tedaj tudi endomorfzma $S+T$ in αT pozitivno definitna.

(3) Naj bo T pozitivno definiten endomorfizem končno-dimenzionalnega vektorskega prostora V nad $\mathbb{F}$ s skalarnim produktom. Pokaži, da je tedaj T avtomorfizem.

TRDITEV 3.66. *Naj bo T endomorfizem končno-dimenzionalnega vektorskega prostora V nad $\mathbb{F}$ s skalarnim produktom. Endomorfizem T je pozitivno definiten, če, in samo če, je sebi-adjungiran in so vse njegove lastne vrednosti pozitivna realna števila. V tem primeru je determinanta endomorfizma T pozitivna.*

DOKAZ. Predpostavimo najprej, da je endomorfizem T pozitivno definiten. Naj bo λ poljubna lastna vrednosti endomorfizma T in naj bo $v \in \mathrm{E}_T(\lambda)$ neničelen vektor. Ker velja $\lambda\langle v,v\rangle = \langle \lambda v, v\rangle = \langle Tv, v\rangle > 0$, sledi $\lambda > 0$.

Dokažimo se obratni del trditve. Predpostavimo, da je endomorfizem T sebi-adjungiran in da so vse njegove lastne vrednosti pozitivne. Po posledici 3.60 obstaja ortonormirana baza $[v_1, v_2, \ldots, v_n]$ vektorskega prostora V, sestavljena iz lastnih vektorjev endomorfizma T. Za vsak $i = 1, 2, \ldots, n$ naj bo λ_i tista lastna vrednosti endomorfizma T, za katero je $Tv_i = \lambda_i v_i$. Po predpostavki je torej $\lambda_i > 0$. Za poljuben neničelen vektor $v \in \mathsf{V}$ tedaj velja

$$\begin{aligned}\langle Tv, v\rangle &= \left\langle T\Big(\sum_{j=1}^{n}\langle v, v_j\rangle v_j\Big), \sum_{i=1}^{n}\langle v, v_i\rangle v_i\right\rangle = \left\langle \sum_{j=1}^{n}\langle v, v_j\rangle Tv_j, \sum_{i=1}^{n}\langle v, v_i\rangle v_i\right\rangle \\ &= \left\langle \sum_{j=1}^{n}\langle v, v_j\rangle \lambda_j v_j, \sum_{i=1}^{n}\langle v, v_i\rangle v_i\right\rangle = \sum_{i=1}^{n}\lambda_i|\langle v, v_i\rangle|^2 > 0. \qquad \square\end{aligned}$$

TRDITEV 3.67. *Naj bo T sebi-adjungiran endomorfizem končno-dimenzionalnega vektorskega prostora V nad $\mathbb{F}$ s skalarnim produktom. Če je U tak vektorski podprostor vektorskega prostora V, da je $\langle Tu, u\rangle > 0$ za vsak neničelen vektor $u \in \mathsf{U}$, potem velja $\sum_{\lambda>0}\mathrm{akr}_T(\lambda) \geq \dim \mathsf{U}$.*

DOKAZ. Označimo $p = \sum_{\lambda>0}\mathrm{akr}_T(\lambda)$. Izberimo ortonormirano bazo $\mathscr{B} = [v_1, v_2, \ldots, v_n]$ vektorskega prostora V, sestavljeno iz lastnih vektorjev endomorfizma T. Za vsak $i = 1, 2, \ldots, n$ naj bo λ_i tista lastna vrednosti endomorfizma T, za katero je $Tv_i = \lambda_i v_i$. Vektorje v bazi $\mathscr{B}$ lahko uredimo tako, da je $\lambda_i > 0$ za vsak indeks i, za katerega je $i \leq p$, in da je $\lambda_i \leq 0$ za vsak indeks i, za katerega velja $i > p$. Dokazati moramo, da je $p \geq \dim \mathsf{U}$.

Pa predpostavimo obratno, da je torej $p < \dim \mathsf{U}$. V tem primeru bi veljalo

$$\dim \mathrm{Span}\{v_{p+1}, \ldots, v_n\} + \dim \mathsf{U} > (n-p) + p = n$$

in zato $\mathrm{Span}\{v_{p+1}, \ldots, v_n\} \cap \mathsf{U} \neq \{0\}$, to pa nas pripelje do protislovja, saj za poljuben neničelen vektor $u \in \mathrm{Span}\{v_{p+1}, \ldots, v_n\} \cap \mathsf{U}$ tedaj velja

$$\begin{aligned}\langle Tu, u\rangle &= \left\langle T\Big(\sum_{j=p+1}^{n}\langle u, v_j\rangle v_j\Big), \sum_{i=p+1}^{n}\langle u, v_i\rangle v_i\right\rangle = \left\langle \sum_{j=p+1}^{n}\langle u, v_j\rangle Tv_j, \sum_{i=p+1}^{n}\langle u, v_i\rangle v_i\right\rangle \\ &= \left\langle \sum_{j=p+1}^{n}\langle u, v_j\rangle \lambda_j v_j, \sum_{i=p+1}^{n}\langle u, v_i\rangle v_i\right\rangle = \sum_{i=p+1}^{n}\lambda_i|\langle u, v_i\rangle|^2 \leq 0. \qquad \square\end{aligned}$$

TRDITEV 3.68 (Sylvesterov kriterij). *Matrika* $A \in \mathrm{Mat}(n \times n, \mathbb{F})$ *je pozitivno definitna, če, in samo če, je hermitska in velja*

$$\det \mathrm{sub}_{\{1,\ldots,k\}\{1,\ldots,k\}}(A) > 0$$

za vsak $k = 1, 2, \ldots, n$.

DOKAZ. Če je matrika A hermitska, potem so očitno hermitske tudi matrike

$$A^{(k)} = \mathrm{sub}_{\{1,\ldots,k\}\{1,\ldots,k\}}(A),$$

za $k = 1, 2, \ldots, n$. Predpostavimo, da je matrika A pozitivno definitna. Za vsak neničelen vektor $v' \in \mathbb{F}^k$ je tudi vektor $v = (v', 0) \in \mathbb{F}^k \times \{0\}^{n-k} \subset \mathbb{F}^n$ neničelen in glede na standardni skalarni produkt velja

$$\langle A^{(k)} v', v' \rangle = \langle Av, v \rangle > 0.$$

To pomeni, da je tudi matrika $A^{(k)}$ pozitivno definitna, zato je njena determinanta pozitivna, za vsak $k = 1, 2, \ldots, n$.

Obratno implikacijo iz trditve bomo dokazali z indukcijo na število n. V primeru $n = 1$ je trditev očitna. Naj bo torej $n > 1$ in predpostavimo, da trditev velja za vse matrike velikosti $(n-1) \times (n-1)$, da je matrika A hermitska in da so poddeterminante $\det\left(A^{(1)}\right), \ldots, \det\left(A^{(n)}\right)$ matrike A vse pozitivne. Ker je

$$\det \mathrm{sub}_{\{1,\ldots,k\}\{1,\ldots,k\}}\left(A^{(n-1)}\right) = \det \mathrm{sub}_{\{1,\ldots,k\}\{1,\ldots,k\}}(A) = \det\left(A^{(k)}\right) > 0$$

za vsak $k = 1, 2, \ldots, n-1$, je po indukcijski predpostavki matrika $A^{(n-1)}$ pozitivno definitna. Odtod sledi, da za vsak neničelen vektor $u = (u', 0) \in \mathbb{F}^{(n-1)} \times \{0\} \subset \mathbb{F}^n$ glede na standardni skalarni produkt velja

$$\langle Au, u \rangle = \langle A^{(n-1)} u', u' \rangle > 0.$$

Trditev 3.67, v kateri vzamemo $\mathsf{U} = \mathbb{F}^{(n-1)} \times \{0\} \subset \mathbb{F}^n$, nam zdaj pove, da je

$$\sum_{\lambda > 0} \mathrm{akr}_A(\lambda) \geq n - 1.$$

Ker je ob tem $\det(A) = \det\left(A^{(n)}\right) > 0$, ima matrika A vse svoje lastne vrednosti pozitivne. Po trditvi 3.66 je matrika A pozitivno definitna. □

KOMENTAR 3.69. (1) Poddeterminante

$$\det \mathrm{sub}_{\{1,\ldots,k\}\{1,\ldots,k\}}(A)$$

matrike $A \in \mathrm{Mat}(n \times n, \mathbb{F})$, $k = 1, 2, \ldots, n$, ki nastopajo v Sylvesterovem kriteriju za pozitivno definitnost 3.68, včasih imenujemo tudi *glavne poddeterminante* ali *glavni minorji* matrike A.

(2) Naj bo $T \in \mathrm{End}(\mathsf{V})$ endomorfizem končno-dimenzionalnega vektorskega prostora nad $\mathbb{F}$ s skalarnim produktom.

Endomorfizem T je *negativno definiten*, če je sebi-adjungiran in za vsak neničelen vektor $v \in \mathsf{V}$ velja $\langle Tv, v \rangle < 0$. V tem primeru zapišemo $T < 0$.

Endomorfizem T je *pozitivno semi-definiten*, če je sebi-adjungiran in za vsak vektor $v \in \mathsf{V}$ velja $\langle Tv, v \rangle \geq 0$. V tem primeru zapišemo $T \geq 0$.

Endomorfizem T je *negativno semi-definiten*, če je sebi-adjungiran in za vsak vektor $v \in \mathsf{V}$ velja $\langle Tv, v \rangle \leq 0$. V tem primeru zapišemo $T \leq 0$.

Na podoben način, kot smo dokazali trditev 3.66, se lahko prepričamo, da je endomorfizen T negativno definiten (oziroma pozitivno semi-definiten, oziroma negativno semi-definiten), če, in samo če, je sebi-adjungiran in so vse njegove lastne vrednosti negativna (oziroma nenegativna, oziroma nepozitivna) realna števila.

Matrika $A \in \mathrm{Mat}(n \times n, \mathbb{F})$ je negativno definitna (oziroma pozitivno semi-definitna, oziroma negativno semi-definitna), kadar je tak pripadajoči endomorfizem $\ell(A)$ vektorskega prostora $\mathbb{F}^n$ s standardnim skalarnim produktom. V tem primeru zapišemo $A < 0$ (oziroma $A \geq 0$, oziroma $A \leq 0$).

Vaja 3.70. (1) Pokaži, da je matrika $A \in \mathrm{Mat}(n \times n, \mathbb{F})$ negativno definitna, če, in samo če, je hermitska in velja

$$(-1)^k \det \mathrm{sub}_{\{1,\ldots,k\}\{1,\ldots,k\}}(A) > 0$$

za vsak $k = 1, 2, \ldots, n$.

(2) Naj bo T pozitivno definiten endomorfizem končno-dimenzionalnega vektorskega prostora V nad $\mathbb{F}$ s skalarnim produktom. Pokaži, da obstaja pozitivno definiten endomorfizem $S \in \mathrm{End}(\mathsf{V})$, za katerega velja $S^2 = T$ in $ST = TS$.

3.3. Kvadratne forme

V tem razdelku bomo uporabili izrek o diagonalizaciji sebi-adjungiranih endomorfizmov pri študiju kvadratnih form. Privzeli bomo, da so naši skalarji realna števila, torej $\mathbb{F} = \mathbb{R}$.

3.3.1. Simetrične bilinearne forme in njihove kvadratne forme. Naj bo V končno-dimenzionalen realen vektorski prostor. *Bilinearna forma* na vektorskem prostoru V je preslikava $g : \mathsf{V} \times \mathsf{V} \to \mathbb{R}$, za katero velja

(i) $g(u + v, w) = g(u, w) + g(v, w)$,
(ii) $g(v, u + w) = g(v, u) + g(v, w)$ in
(iii) $g(\alpha v, w) = \alpha g(v, w) = g(v, \alpha w)$

za vse vektorje $u, v, w \in \mathsf{V}$ in za vse skalarje $\alpha \in \mathbb{F}$. Drugače povedano, bilinearna forma je linearna tako v prvi spremenljivki kot tudi v drugi spremenljivki. Bilinearna forma g na vektorskem prostoru V je *simetrična*, če velja še

(iv) $g(v, w) = g(w, v)$

za vse vektorje $v, w \in \mathsf{V}$.

Preslikava $f : \mathsf{V} \to \mathbb{R}$ je *kvadratna forma* na vektorskem prostoru V, če obstaja takšna simetrična bilinerana forma g na vektorskem prostoru V, da je

$$f(u) = g(u, u)$$

za vsak vektor $u \in \mathsf{V}$. V tem primeru pravimo, da je kvadratna forma f prirejena simetrični bilinearni formi g.

Zgled 3.71. (1) Vsak skalarni produkt na realnem vektorskem prostoru je simetrična bilinearna forma.

(2) Vsota dveh simetričnih bilinearnih form na vektorskem prostoru V je spet simetrična bilinearna forma, prav tako pa je produkt simetrične bilinearne forme s poljubnim realnim številom simetrična bilinearna forma. Množica simetričnih bilinearnih form na vektorskem prostoru V je torej realen vektorski prostor za operacije po točkah.

(3) Naj bo V končno-dimenzionalen realen vektorski prostor s skalarnim produktom. Poljubnemu sebi-adjungiranemu endomorfizmu $T \in \mathrm{End}(\mathsf{V})$ priredimo simetrično bilinearno formo $\Gamma_T : \mathsf{V} \times \mathsf{V} \to \mathbb{R}$, ki je dana s predpisom

$$\Gamma_T(v, w) = \langle Tv, w \rangle$$

za vse vektorje $v, w \in \mathsf{V}$. Kvadratno formo, prirejeno simetrični bilinearni formi Γ_T, označimo z $\Phi_T : \mathsf{V} \to \mathbb{R}$, torej

$$\Phi_T(u) = \Gamma_T(u,u) = \langle Tu, u\rangle$$

za vsak vektor $u \in \mathsf{V}$.

(4) Naj bo g poljubna simetrična bilinearna forma na vektorskem prostoru $\mathbb{R}^n$. Definirajmo realno matriko A velikosti $n \times n$ s predpisom $A_{ij} = g(\mathrm{e}_j, \mathrm{e}_i)$ za vse $i, j = 1, 2, \ldots, n$. Ker je bilinearna forma g simetrična, je matrika A simetrična, torej velja

$$A_{ij} = A_{ji} = g(\mathrm{e}_j, \mathrm{e}_i) = g(\mathrm{e}_i, \mathrm{e}_j).$$

Za vse vektorje $v = (\alpha_1, \alpha_2, \ldots, \alpha_n) \in \mathbb{R}^n$ in $w = (\beta_1, \beta_2, \ldots, \beta_n) \in \mathbb{R}^n$ velja

$$\begin{aligned} g(v,w) &= g\Big(\sum_{j=1}^n \alpha_j \mathrm{e}_j, \sum_{i=1}^n \beta_i \mathrm{e}_i\Big) = \sum_{j=1}^n \sum_{i=1}^n \alpha_j \beta_i g(\mathrm{e}_j, \mathrm{e}_i) \\ &= \sum_{i=1}^n \beta_i \sum_{j=1}^n A_{ij}\alpha_j = \sum_{i=1}^n \beta_i \big(Av\big)_{i1} = w^{\mathsf{t}} A v. \end{aligned}$$

Simetrično bilinearno formo g lahko torej zapišemo v obliki

$$g(v,w) = w^{\mathsf{t}} A v = \sum_{i=1}^n \sum_{j=1}^n \beta_i A_{ij} \alpha_j = \sum_{1 \le i,j \le n} A_{ij}\beta_i\alpha_j.$$

Ker je število $w^{\mathsf{t}}Av$ standardni skalarni produkt vektorjev Av in w, je $g = \Gamma_A$.

Naj bo $f : \mathbb{R}^n \to \mathbb{R}$ kvadratna forma, prirejena simetrični bilinearni formi g. Za vsak vektor $u = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ tedaj velja

$$f(u) = f(x_1, x_2, \ldots, x_n) = \sum_{1 \le i,j \le n} A_{ij} x_i x_j = \sum_{i=1}^n A_{ii} x_i^2 + 2 \sum_{1 \le i < j \le n} A_{ij} x_i x_j.$$

Funkciji takšne oblike pravimo realen *homogen polinom* stopnje 2 v spremenljivkah $x_1, x_2, \ldots, x_n$.

Po izreku 3.61 obstaja ortonormirana baza $\mathscr{B}$ vektorskega prostora $\mathbb{R}^n$, sestavljena iz lastnih vektorjev matrike A. Prehodna matrika $Q = [\mathrm{id}]^{\mathscr{B}}_{\mathscr{E}}$ iz baze $\mathscr{B}$ na standardno bazo je ortogonalna, matrika $D = Q^{\mathsf{t}}AQ$ pa je diagonalna,

$$D = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n).$$

Tedaj velja $Q^{\mathsf{t}}v = [\mathrm{id}]^{\mathscr{E}}_{\mathscr{B}}[v]_{\mathscr{E}} = [v]_{\mathscr{B}}$ in podobno $Q^{\mathsf{t}}w = [w]_{\mathscr{B}}$. Če označimo $\tilde{v} = Q^{\mathsf{t}}v = (\tilde{\alpha}_1, \tilde{\alpha}_2, \ldots, \tilde{\alpha}_n)$ in $\tilde{w} = Q^{\mathsf{t}}w = (\tilde{\beta}_1, \tilde{\beta}_2, \ldots, \tilde{\beta}_n)$, potem je

$$g(v,w) = g(Q\tilde{v}, Q\tilde{w}) = (Q\tilde{w})^{\mathsf{t}} A Q\tilde{v} = (Q\tilde{w})^{\mathsf{t}} Q D Q^{\mathsf{t}} Q \tilde{v} = \tilde{w}^{\mathsf{t}} D \tilde{v} = \sum_{i=1}^n \lambda_i \tilde{\beta}_i \tilde{\alpha}_i.$$

Če koordinatni vektor $[u]_{\mathscr{B}}$ vektorja $u = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ v bazi $\mathscr{B}$ označimo z $\tilde{u} = [u]_{\mathscr{B}} = Q^{\mathsf{t}}u = (\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_n)$, potem dobimo

$$\begin{aligned} f(u) = f(x_1, x_2, \ldots, x_n) &= \sum_{1 \le i,j \le n} A_{ij} x_i x_j \\ &= f(Q\tilde{u}) = f(Q(\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_n)) = \sum_{i=1}^n \lambda_i \tilde{x}_i^2. \end{aligned}$$

Bilinearno formo g in njeno kvadratno formo f smo tako zapisali v novih koordinatah, ki so dane z ortonormirano bazo $\mathscr{B}$. V teh novih koordinatah sta g in f v *diagonalni obliki*.

(5) Vsak realen homogen polinom f stopnje 2 v spremenljivkah $x_1, x_2, \ldots, x_n$ lahko zapišemo v obliki

$$f(x_1, x_2, \ldots, x_n) = \sum_{1 \le i \le j \le n} c_{ij} x_i x_j,$$

za neka realna števila c_{ij}. Tedaj lahko definiramo matriko A velikosti $n \times n$ s predpisom $A_{ii} = c_{ii}$ in $A_{ij} = A_{ji} = c_{ij}/2$ za vse indekse i in j, $i < j$. Matrika A je simetrična in f je kvadratna forma simetrične bilinearne forme Γ_A, torej $f = \Phi_A$.

(6) Funkcija $f : \mathbb{R}^2 \to \mathbb{R}$, dana s predpisom

$$f(x, y) = 3x^2 - 2xy + 3y^2,$$

je homogen polinom stopnje 2, ki mu pridružimo simetrično matriko:

$$A = \begin{bmatrix} 3 & -1 \\ -1 & 3 \end{bmatrix}$$

Za poljuben vektor $u = (x, y) \in \mathbb{R}^2$ tako velja:

$$f(u) = f(x, y) = u^{\mathrm{t}} A u = \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} 3 & -1 \\ -1 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

Izračunamo lahko, da je $\sigma(A) = \{2, 4\}$, $\mathrm{E}_A(2) = \mathbb{R}(1, 1)$ in $\mathrm{E}_A(4) = \mathbb{R}(-1, 1)$. Lastna vektorja $v_1 = \frac{1}{\sqrt{2}}(1, 1)$ in $v_2 = \frac{1}{\sqrt{2}}(-1, 1)$ matrike A sestavljata ortonormirano bazo $\mathscr{B}$ vektorskega prostora $\mathbb{R}^2$. Prehodna matrika

$$Q = [\mathrm{id}]^{\mathscr{B}}_{\mathscr{E}} = \tfrac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$$

je ortogonalna in predstavlja rotacijo v ravnini za kot $\pi/4$, ob tem pa je

$$D = Q^{\mathrm{t}} A Q = \mathrm{diag}(2, 4)$$

diagonalna matrika. Za vsak vektor $u = (x, y) \in \mathbb{R}^2$ označimo $\tilde{u} = (\tilde{x}, \tilde{y}) = Q^{\mathrm{t}} u$, torej

$$\tilde{x} = \frac{x + y}{\sqrt{2}}, \qquad \tilde{y} = \frac{-x + y}{\sqrt{2}},$$

oziroma

$$x = \frac{\tilde{x} - \tilde{y}}{\sqrt{2}}, \qquad y = \frac{\tilde{x} + \tilde{y}}{\sqrt{2}}.$$

V novih koordinatah velja

$$f(x, y) = 3x^2 - 2xy + 3y^2 = f\Big(\frac{\tilde{x} - \tilde{y}}{\sqrt{2}}, \frac{\tilde{x} + \tilde{y}}{\sqrt{2}}\Big) = f(Q(\tilde{x}, \tilde{y})) = 2\tilde{x}^2 + 4\tilde{y}^2.$$

Tako na primer lahko vidimo, da je množica rešitev enačbe $3x^2 - 2xy + 3y^2 = 8$ v ravnini $\mathbb{R}^2$ enaka množici rešitev enačbe $2\tilde{x}^2 + 4\tilde{y}^2 = 8$, torej elipsa, ki ima osi v smereh lastnih vektorjev v_1 in v_2.

Trditev 3.72. *Naj bo* V *končno-dimenzionalen realen vektorski prostor s skalarnim produktom. Preslikava*

$$\{\textit{sebi-adjungirani endomorfizmi } \mathsf{V} \to \mathsf{V}\} \to \{\textit{simetrične bilinearne forme na } \mathsf{V}\},$$

ki poljubnemu sebi-adjungiranemu endomorfizmu $T : \mathsf{V} \to \mathsf{V}$ *priredi simetrično bilinearno formo* $\Gamma_T : \mathsf{V} \times \mathsf{V} \to \mathbb{R}$, $\Gamma_T(v, w) = \langle Tv, w\rangle$, *je izomorfizem realnih vektorskih prostorov.*

DOKAZ. Preslikava $T \mapsto \Gamma_T$ je očitno linearna in injektivna. Naj bo g poljubna simetrična bilinearna forma na vektorskem prostoru V. Za poljuben vektor $v \in \mathsf{V}$ je $g(v, \bullet) : \mathsf{V} \to \mathbb{R}$, $w \mapsto g(v, w)$, linearen funkcional, zato po Rieszovem izreku obstaja natanko en vektor $T_g(v) \in \mathsf{V}$, tako da velja

$$g(v, w) = \langle w, T_g(v)\rangle$$

za vsak vektor $w \in \mathsf{V}$. Ni težko preveriti, da smo s tem definirali endomorfizem $T_g \in \mathrm{End}(\mathsf{V})$, ki je sebi adjungiran in za katerega velja $\Gamma_{T_g} = g$. $\square$

3.3.2. Diagonalna oblika kvadratne forme. Naj bo g poljubna simetrična bilinearna forma na končno-dimenzionalnem realnem vektorskem prostoru V dimenzije $n \geq 1$. Naj bo $\mathscr{B} = [v_1, v_2, \ldots, v_n]$ baza vektorskega prostora V. *Koordinatna matrika* oziroma *Sylvesterova matrika* simetrične bilinearne forme g v bazi $\mathscr{B}$ je simetrična realna matrika $[g]_{\mathscr{B}}$ velikosti $n \times n$, dana s predpisom

$$\big([g]_{\mathscr{B}}\big)_{ij} = g(v_j, v_i)$$

za vse indekse $i, j = 1, 2, \ldots, n$. Če za poljubna vektorja $v, w \in \mathsf{V}$ označimo $[v]_{\mathscr{B}} = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ in $[w]_{\mathscr{B}} = (\beta_1, \beta_2, \ldots, \beta_n)$, potem dobimo

$$\begin{aligned} g(v, w) &= g\Big(\sum_{j=1}^{n} \alpha_j v_j, \sum_{i=1}^{n} \beta_i v_i\Big) = \sum_{j=1}^{n}\sum_{i=1}^{n} \alpha_j\beta_i g(v_j, v_i) = \sum_{1\leq i,j\leq n} g(v_j, v_i)\beta_i\alpha_j \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} \beta_i\big([g]_{\mathscr{B}}\big)_{ij}\alpha_j = ([w]_{\mathscr{B}})^{\mathsf{t}}[g]_{\mathscr{B}}[v]_{\mathscr{B}}. \end{aligned}$$

Naj bo f kvadratna forma, prirejena simetrični bilinearni formi g. Če za poljuben vektor $u \in \mathsf{V}$ označimo $[u]_{\mathscr{B}} = (x_1, x_2, \ldots, x_n)$, potem velja

$$f(u) = \sum_{1\leq i,j\leq n} g(v_j, v_i)x_i x_j.$$

Simetrična bilinearna forma g in njena kvadratna forma f imata *v bazi* $\mathscr{B}$ *diagonalno obliko*, če velja $g(v_j, v_i) = 0$ za vse indekse $i, j \in \{1, 2, \ldots, n\}$, $i \neq j$, torej tedaj, ko je koordinatna marika $[g]_{\mathscr{B}}$ diagonalna. V tem primeru velja:

$$g(v, w) = \sum_{i=1}^{n} g(v_i, v_i)\beta_i\alpha_i$$

$$f(u) = \sum_{i=1}^{n} g(v_i, v_i)x_i^2$$

POSLEDICA 3.73. *Naj bo* g *simetrična bilinearna forma na končno-dimenzionalnem realnem vektorskem prostoru* V *dimenzije* $n \geq 1$ *s skalarnim produktom.*

(i) Obstaja ortonormirana baza vektorskega prostora V, *v kateri ima simetrična bilinearna forma* g *diagonalno obliko.*

(ii) Naj bo $\mathscr{B} = [v_1, v_2, \ldots, v_n]$ *poljubna ortonormirana baza vektorskega prostora* V, *v kateri ima simetrična bilinearna forma* g *diagonalno obliko. Če je* $T \in \mathrm{End}(\mathsf{V})$ *tisti sebi-adjungiran endomorfizem, za katerega velja* $\Gamma_T = g$, *potem je*

$$[T]_{\mathscr{B}}^{\mathscr{B}} = \mathrm{diag}\,\big(g(v_1, v_1), g(v_2, v_2), \ldots, g(v_n, v_n)\big).$$

DOKAZ. Po trditvi 3.72 obstaja sebi-adjungiran endomorfizem $T \in \mathrm{End}(\mathsf{V})$, za katerega je $\Gamma_T = g$.

(i) Po izreku 3.61 lahko izberemo ortonormirano bazo $\mathscr{B} = [v_1, v_2, \ldots, v_n]$ vektorskega prostora V, ki je sestavljena iz lastnih vektorjev endomorfizma T. Za vsak $j = 1, 2, \ldots, n$ je torej $Tv_j = \lambda_j v_j$ za ustrezno lastno vrednosti $\lambda_j \in \sigma(T)$. Ker velja

$$g(v_j, v_i) = \langle Tv_j, v_i \rangle = \langle \lambda_j v_j, v_i \rangle = \lambda_j \langle v_j, v_i \rangle = 0$$

za vse indekse $i, j \in \{1, 2, \ldots, n\}$, $i \neq j$, ima simetrična bilinearna forma g v ortonormirani bazi $\mathscr{B}$ diagonalno obliko.

(ii) Naj bo zdaj $\mathscr{B} = [v_1, v_2, \ldots, v_n]$ poljubna ortonormirana baza vektorskega prostora V, v kateri ima simetrična bilinearna forma g diagonalno obliko. Naj bo $j \in \{1, 2, \ldots, n\}$. Ker je

$$\langle Tv_j, v_i \rangle = g(v_j, v_i) = 0$$

za vsak indeks $i = 1, 2, \ldots, n$, $i \neq j$, je $Tv_j \in \{v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_n\}^\perp = \mathbb{R}v_j$. Odtod sledi, da je v_j lastni vektor endomorfizma T. Pokazali smo torej, da je baza $\mathscr{B}$ sestavljena iz lastnih vektorjev endomorfizma T, kar pomeni, da je matrika $[T]_{\mathscr{B}}^{\mathscr{B}}$ diagonalna. Ker je $Tv_j = \sum_{i=1}^n \big([T]_{\mathscr{B}}^{\mathscr{B}}\big)_{ij} v_i = \big([T]_{\mathscr{B}}^{\mathscr{B}}\big)_{jj} v_j$, sledi tudi

$$\big([T]_{\mathscr{B}}^{\mathscr{B}}\big)_{jj} = \langle Tv_j, v_j \rangle = g(v_j, v_j). \qquad \square$$

TRDITEV 3.74. *Naj bo g simetrična bilinearna forma na končno-dimenzionalnem realnem vektorskem prostoru V dimenzije $n \geq 1$. Obstaja takšna baza $\mathscr{B} = [v_1, v_2, \ldots, v_n]$ vektorskega prostora V, da za vse indekse $i \in \{1, 2, \ldots, n\}$ velja*

$$g(v_i, v_i) \in \{1, 0, -1\}$$

in da za vse indekse $i, j \in \{1, 2, \ldots, n\}$, $i \neq j$, velja

$$g(v_i, v_j) = 0.$$

Posebej to pomeni, da ima simetrična linearna forma g v bazi $\mathscr{B}$ diagonalno obliko.

DOKAZ. Izberimo poljuben skalarni produkt na vektorskem prostoru V. Po posledici 3.73 obstaja baza $\mathscr{B}' = [v_1', v_2', \ldots, v_n']$ vektorskega prostora V, v kateri ima simetrična bilinearna forma g diagonalno obliko. Zdaj za vsak $i = 1, 2, \ldots, n$ definirano:

$$v_i = \begin{cases} |g(v_i', v_i')|^{-\frac{1}{2}}\, v_i' & ; \quad g(v_i', v_i') \neq 0 \\ v_i' & ; \quad g(v_i', v_i') = 0 \end{cases}$$

Tedaj je $\mathscr{B} = [v_1, v_2, \ldots, v_n]$ baza vektorskega prostora V z želenimi lastnostmi. $\square$

TRDITEV 3.75 (Sylvesterov zakon inercije). *Naj bo g simetrična bilinearna forma na končno-dimenzionalnem realnem vektorskem prostoru V dimenzije $n \geq 1$. Če sta $\mathscr{B} = [v_1, v_2, \ldots, v_n]$ in $\mathscr{B}' = [v_1', v_2', \ldots, v_n']$ dve bazi vektorskega prostora V, v katerih ima simetrična bilinearna forma g diagonalno obliko, potem velja:*

$$\#\{i \in \mathbb{N}\,;\ i \leq n,\ g(v_i, v_i) > 0\} = \#\{i \in \mathbb{N}\,;\ i \leq n,\ g(v_i', v_i') > 0\}$$
$$\#\{i \in \mathbb{N}\,;\ i \leq n,\ g(v_i, v_i) < 0\} = \#\{i \in \mathbb{N}\,;\ i \leq n,\ g(v_i', v_i') < 0\}$$
$$\#\{i \in \mathbb{N}\,;\ i \leq n,\ g(v_i, v_i) = 0\} = \#\{i \in \mathbb{N}\,;\ i \leq n,\ g(v_i', v_i') = 0\}$$

DOKAZ. Brez izgube splošnosti lahko predpostavimo, da so vektorji v bazah $\mathscr{B}$ in $\mathscr{B}'$ urejeni tako, da velja $g(v_i, v_i) > 0$ za vse indekse $i \leq p$ ter $g(v_i, v_i) \leq 0$ za vse indekse $i > p$, in da podobno velja $g(v_i', v_i') > 0$ za vse indekse $i \leq q$ ter $g(v_i', v_i') \leq 0$ za vse indekse $i > q$.

Najprej bomo pokazali, da velja $p \leq q$. To je očitno res, če je bodisi $p = 0$ ali pa $q = n$. Predpostavimo torej, da je $p > 0$ in da je $q < n$. Dokazali bomo, da so tedaj vektorji $v_1, \ldots, v_p, v'_{q+1}, \ldots, v'_n$ med seboj linearno neodvisni. Naj bodo $\alpha_1, \ldots, \alpha_p, \beta_1, \ldots, \beta_{n-q} \in \mathbb{R}$ takšni skalarji, da je

$$\alpha_1 v_1 + \cdots + \alpha_p v_p + \beta_1 v'_{q+1} + \cdots + \beta_{n-q} v'_n = 0,$$

in označimo $u = \alpha_1 v_1 + \cdots + \alpha_p v_p = -(\beta_1 v'_{q+1} + \cdots + \beta_{n-q} v'_n)$. Zdaj velja

$$g(u, u) = g(\alpha_1 v_1 + \cdots + \alpha_p v_p, \alpha_1 v_1 + \cdots + \alpha_p v_p) = \sum_{i=1}^{p} \alpha_i^2 g(v_i, v_i) \geq 0$$

in hkrati

$$\begin{aligned} g(u, u) &= g\big(-(\beta_1 v'_{q+1} + \cdots + \beta_{n-q} v'_n), -(\beta_1 v'_{q+1} + \cdots + \beta_{n-q} v'_n)\big) \\ &= \sum_{i=1}^{n-q} \beta_i^2 g(v'_{q+i}, v'_{q+i}) \leq 0, \end{aligned}$$

zato je $g(u, u) = 0$. Velja torej $\sum_{i=1}^{p} \alpha_i^2 g(v_i, v_i) = 0$ in zato $\alpha_1 = \cdots = \alpha_p = 0$. Odtod sledi $u = 0$, torej $\beta_1 v'_{q+1} + \cdots + \beta_{n-q} v'_n = 0$, zato tudi $\beta_1 = \cdots = \beta_{n-q} = 0$. Vektorji $v_1, \ldots, v_p, v'_{q+1}, \ldots, v'_n$ so torej med seboj linearno neodvisni. Posebej to pomeni, da je $p + (n - q) \leq n$ oziroma $p \leq q$.

Če zamenjamo vlogi baz $\mathscr{B}$ in $\mathscr{B}'$, nam enak argument pove, da je tudi $q \leq p$. Velja torej enakost $p = q$. S tem smo dokazali, da velja prva od treh enakosti iz trditve. Drugo enakost dobimo iz prve tako, da jo uporabimo na simetrični bilinearni formi $-g$. Tretja enakost je direktna posledica prvih dveh. $\square$

Komentar 3.76. Za poljubno simetrično bilinearno formo g na končno-dimenzionalnem vektorskem prostoru V pozitivne dimenzije označimo

$$\begin{aligned} \operatorname{ind}^+(g) &= \#\{i \in \mathbb{N}\ ;\ i \leq n,\ g(v_i, v_i) > 0\}, \\ \operatorname{ind}^-(g) &= \#\{i \in \mathbb{N}\ ;\ i \leq n,\ g(v_i, v_i) < 0\}, \\ \operatorname{ind}^0(g) &= \#\{i \in \mathbb{N}\ ;\ i \leq n,\ g(v_i, v_i) = 0\}, \end{aligned}$$

kjer je $\mathscr{B} = [v_1, v_2, \ldots, v_n]$ poljubna baza vektorskega prostora V, v kateri ima simetrična bilinearna forma g diagonalno obliko. Po Sylvesterovem zakonu o inerciji so ta števila neodvisna od izbire takšne baze $\mathscr{B}$. Število $\operatorname{ind}^+(g)$ imenujemo *indeks pozitivnosti*, število $\operatorname{ind}^-(g)$ pa *indeks negativnosti* simetrične bilinearne forme g. Urejenemu paru $\big(\operatorname{ind}^+(g), \operatorname{ind}^-(g)\big)$ pravimo tudi *signatura*, vsoti $\operatorname{rank}(g) = \operatorname{ind}^+(g) + \operatorname{ind}^-(g)$ pa *rang* simetrične bilinearne forme g.

Če je vektorski prostor V opremljen s skalarnim produktom in je $T \in \operatorname{End}(\mathsf{V})$ tisti sebi-adjungiran endomorfizem, za katerega je $\Gamma_T = g$, potem po posledici 3.73(ii) sledi

$$\begin{aligned} \operatorname{ind}^+(g) &= \sum_{\lambda > 0} \operatorname{akr}_T(\lambda), \\ \operatorname{ind}^-(g) &= \sum_{\lambda < 0} \operatorname{akr}_T(\lambda), \\ \operatorname{ind}^0(g) &= \operatorname{akr}_T(0), \\ \operatorname{rank}(g) &= \operatorname{rank}(T). \end{aligned}$$

VAJA 3.77. (1) Matrika $B \in \mathrm{Mat}(n \times n, \mathbb{R})$ je po definiciji *kongruentna* matriki $A \in \mathrm{Mat}(n \times n, \mathbb{R})$, če obstaja takšna obrnljiva matrika $P \in \mathrm{GL}(n, \mathbb{R})$, da je $B = P^{\mathsf{t}} A P$. Pokaži, da je kongruenca matrik ekvivalenčna relacija.

(2) Naj bo g simetrična bilinearna forma na končno-dimenzionalnem realnem vektorskem prostoru V dimenzije $n \geq 1$. Naj bosta $\mathscr{B}$ in $\mathscr{B}'$ dve bazi vektorskega prostora V. Pokaži, da velja

$$[g]_{\mathscr{B}'} = \left([\mathrm{id}]_{\mathscr{B}}^{\mathscr{B}'}\right)^{\mathsf{t}} [g]_{\mathscr{B}} [\mathrm{id}]_{\mathscr{B}}^{\mathscr{B}'}$$

in da sta si torej matriki $[g]_{\mathscr{B}'}$ in $[g]_{\mathscr{B}}$ kongruentni.

(3) Pokaži, da je vsaka realna simetrična matrika kongruentna diagonalni matriki, katere komponente so števila iz množice $\{1, 0, -1\}$.

3.3.3. Krivulje, ploskve in hiperploskve drugega reda. Realen polinom stopnje 2 v n spremenljivkah $x_1, x_2, \ldots, x_n$ je funkcija $f : \mathbb{R}^n \to \mathbb{R}$ oblike

$$f(x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n} a_{ii} x_i^2 + 2 \sum_{1 \leq i < j \leq n} a_{ij} x_i x_j + \sum_{i=1}^{n} b_i x_i + c_0,$$

kjer so $a_{11}, a_{12}, \ldots, a_{nn}, b_1, \ldots, b_n, c_0$ realne konstante in kjer je vsaj ena od konstant $a_{11}, a_{12}, \ldots, a_{nn}$ neničelna. *Homogeni del* stopnje 2 polinoma f,

$$\sum_{i=1}^{n} a_{ii} x_i^2 + 2 \sum_{1 \leq i < j \leq n} a_{ij} x_i x_j,$$

je kvadratna forma na vektorskem prostoru $\mathbb{R}^n$. Množica ničel $f^{-1}(\{0\}) \subset \mathbb{R}^n$ polinoma f je množica rešitev $(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ enačbe

$$\sum_{i=1}^{n} a_{ii} x_i^2 + 2 \sum_{1 \leq i < j \leq n} a_{ij} x_i x_j + \sum_{i=1}^{n} b_i x_i + c_0 = 0. \tag{3.1}$$

Enačbi takšne oblike pravimo tudi *splošna enačba drugega reda* za realne neznanke $x_1, x_2, \ldots, x_n$, njeni množici rešitev $\mathcal{S} \subset \mathbb{R}^n$ pa *hiperploskev drugega reda* v $\mathbb{R}^n$.

Splošni enačbi drugega reda (3.1) priredimo simetrično realno matriko

$$A \in \mathrm{Mat}(n \times n, \mathbb{R}),$$

dano s predpisom

$$A_{ij} = A_{ji} = a_{ij}$$

za vse $i, j \in \{1, 2, \ldots, n\}$, $i \leq j$, ter vektor oziroma stolpec

$$B = (b_1, b_2, \ldots, b_n) = \begin{bmatrix} b_1 & b_2 & \cdots & b_n \end{bmatrix}^{\mathsf{t}} \in \mathrm{Mat}(n \times 1, \mathbb{R}).$$

Poleg tega označimo z X vektor oziroma stolpec neznank

$$X = (x_1, x_2, \ldots, x_n) = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix}^{\mathsf{t}} \in \mathrm{Mat}(n \times 1, \mathbb{R}).$$

Splošno enačbo drugega reda (3.1) lahko zdaj zapišemo v matrični obliki kot enačbo

$$X^{\mathsf{t}} A X + B^{\mathsf{t}} X + c_0 = 0 \tag{3.2}$$

za neznani stolpec $X \in \mathrm{Mat}(n \times 1, \mathbb{R})$.

Simetrična bilinearna forma Γ_A, ki pripada simetrični matriki A, ima svoj indeks pozitivnosti in svoj indeks negativnosti. Na kratko označimo $r^+ = \mathrm{ind}^+(\Gamma_A)$, $r^- = \mathrm{ind}^-(\Gamma_A)$ in $r = \mathrm{rank}(\Gamma_A) = r^+ + r^-$. Velja torej $0 \leq r^+ \leq n$, $0 \leq r^- \leq n$ in $1 \leq r \leq n$. Simetrična realna matrika A je ortogonalno diagonalizabilna, zato lahko izberemo ortonormirano bazo $\mathscr{B}$ vektorskega prostora $\mathbb{R}^n$, sestavljeno iz lastnih vektorjev matrike A. Pri tem lahko ortonormirano bazo $\mathscr{B} = [v_1, v_2, \ldots, v_n]$ izberemo tako, da za vsak $i = 1, 2, \ldots, n$ velja:

(i) če je $1 \leq i \leq r^+$, potem vektor v_i pripada pozitivni lastni vrednosti matrike A,
(ii) če je $r^+ < i \leq r$, potem vektor v_i pripada negativni lastni vrednosti matrike A,
(iii) če je $r < i \leq n$, potem vektor v_i pripada lastni vrednosti 0 matrike A,
(iv) če je $r + 2 \leq i \leq n$, potem je vektor v_i pravokoten na vektor B.

Res, z ustrezno izbiro vrstnega reda vektorjev v bazi $\mathscr{B}$ dosežemo, da veljajo točke (i-iii). Pogoj iz točke (iv) je avtomatično izpolnjen, če je $r+2 > n$. Če je $r+2 \leq n$, potem izberemo normiran lastni vektor $v_{r+1} \in \mathrm{E}_A(0)$ tako, da je

$$\mathrm{pr}^{\perp}_{\mathrm{E}_A(0)}(B) \in \mathbb{R}v_{r+1}.$$

Ob takšni izbiri je izpolnjena tudi točka (iv).

Naj bo $Q = [\mathrm{id}]^{\mathscr{B}}_{\mathscr{E}}$ prehodna matrika iz baze $\mathscr{B}$ na standardno bazo. Matrika Q je torej ortogonalna matrika s stolpci $v_1, v_2, \ldots, v_n$. Realna matrika $D = Q^{\mathsf{t}}AQ$ je diagonalna,

$$D = Q^{\mathsf{t}}AQ = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n).$$

Ker za bazo $\mathscr{B}$ veljajo točke (i-iii), je med lastnimi vrednostmi $\lambda_1, \lambda_2, \ldots, \lambda_n$ matrike A prvih r^+ pozitivnih, naslednjih r^- je negativnih, preostale pa so enake številu 0. Označimo

$$\tilde{X} = Q^{\mathsf{t}}X = (\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_n)$$

in

$$\tilde{B} = Q^{\mathsf{t}}B = (\tilde{b}_1, \tilde{b}_2, \ldots, \tilde{b}_n).$$

Če je $r + 2 \leq n$, potem po točki (iv) velja $\tilde{b}_{r+2} = \cdots = \tilde{b}_n = 0$. Ob teh oznakah je

$$\begin{aligned} X^{\mathsf{t}}AX + B^{\mathsf{t}}X + c_0 &= X^{\mathsf{t}}QDQ^{\mathsf{t}}X + B^{\mathsf{t}}QQ^{\mathsf{t}}X + c_0 \\ &= (Q^{\mathsf{t}}X)^{\mathsf{t}}D(Q^{\mathsf{t}}X) + (Q^{\mathsf{t}}B)^{\mathsf{t}}(Q^{\mathsf{t}}X) + c_0 \\ &= \tilde{X}^{\mathsf{t}}D\tilde{X} + \tilde{B}^{\mathsf{t}}\tilde{X} + c_0. \end{aligned}$$

V novih koordinatah, danih z bazo $\mathscr{B}$, oziroma za nove neznanke $\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_n$, ima tako splošna enačba drugega reda (3.1) obliko

$$\tilde{X}^{\mathsf{t}}D\tilde{X} + \tilde{B}^{\mathsf{t}}\tilde{X} + c_0 = 0$$

oziroma

$$\sum_{i=1}^{n} \lambda_i \tilde{x}_i^2 + \sum_{i=1}^{n} \tilde{b}_i \tilde{x}_i + c_0 = 0.$$

Definirajmo naravno število s na naslednji način: če je $r < n$ in je $\tilde{b}_{r+1} \neq 0$, potem naj bo $s = r + 1$, sicer pa naj bo $s = r$. Ker za bazo $\mathscr{B}$ veljata točki (iii) in (iv), zadnjo enačbo lahko zapišemo v obliki

$$\sum_{i=1}^{r} \lambda_i \tilde{x}_i^2 + \sum_{i=1}^{s} \tilde{b}_i \tilde{x}_i + c_0 = 0. \tag{3.3}$$

To enačbo lahko še nekoliko poenostavimo. Za vsak indeks $i = 1, 2, \ldots, r$ lahko zapišemo

$$\lambda_i \tilde{x}_i^2 + \tilde{b}_i \tilde{x}_i = \lambda_i \Big(\tilde{x}_i + \frac{\tilde{b}_i}{2\lambda_i}\Big)^2 - \frac{\tilde{b}_i^2}{4\lambda_i} = \lambda_i \Big(\tilde{x}_i - \xi_i\Big)^2 - \frac{\tilde{b}_i^2}{4\lambda_i},$$

kjer smo označili $\xi_i = -\tilde{b}_i/(2\lambda_i)$. Naj bo $\hat{c}_0 = c_0 - \sum_{i=1}^{r} \tilde{b}_i^2/(4\lambda_i)$. Za vsak indeks i, za katerega je $s+1 \leq i \leq n$, izberimo $\xi_i \in \mathbb{R}$ poljubno. Če je $s = r+1$, potem naj bo $\xi_{r+1} = -\hat{c}_0/\tilde{b}_{r+1}$. Tako smo določili vse komponente vektorja $\Xi = (\xi_1, \xi_2, \ldots, \xi_n)$. Če je $s = r$, potem enačbo (3.3) lahko z novimi oznakami zapišemo kot

$$\sum_{i=1}^{r} \lambda_i(\tilde{x}_i - \xi_i)^2 + \hat{c}_0 = 0,$$

medtem ko v primeru $s = r + 1$ enačba (3.3) dobi obliko

$$\sum_{i=1}^{r} \lambda_i(\tilde{x}_i - \xi_i)^2 + \tilde{b}_{r+1}(\tilde{x}_{r+1} - \xi_{r+1}) = 0.$$

Uvedemo lahko še nove neznanke $\hat{x}_i = \tilde{x}_i - \xi_i$ za vse $i = 1, 2, \ldots, n$ in dobimo enakost

$$\sum_{i=1}^{r} \lambda_i \hat{x}_i^2 + \hat{c}_0 = 0 \tag{3.4}$$

v primeru $s = r$ oziroma enakost

$$\sum_{i=1}^{r} \lambda_i \hat{x}_i^2 + \tilde{b}_{r+1}\hat{x}_{r+1} = 0 \tag{3.5}$$

v primeru $s = r + 1$.

(1) Predpostavimo, da velja $r = s$. Če je $\hat{c}_0 = 0$, potem označimo $\gamma = 1$, sicer pa naj bo $\gamma = |\hat{c}_0|^{-\frac{1}{2}}$. Enačbo (3.4) pomnožimo s pozitivnim številom γ^2 in dobimo

$$\sum_{1 \leq i \leq r^+} \big(\gamma \lambda_i^{\frac{1}{2}}\big)^2 \hat{x}_i^2 - \sum_{r^+ < i \leq r} \big(\gamma |\lambda_i|^{\frac{1}{2}}\big)^2 \hat{x}_i^2 + \gamma^2 \hat{c}_0 = 0.$$

Naj bo $\varrho_i = 1/(\gamma|\lambda_i|^{\frac{1}{2}})$ za vsak $i = 1, 2, \ldots, r$ in naj bo $\epsilon_0 = -\gamma^2 \hat{c}_0$. S temi oznakami enačbo drugega reda (3.4) zapišemo v *standardni obliki*

$$\sum_{1 \leq i \leq r^+} \hat{x}_i^2/\varrho_i^2 - \sum_{r^+ < i \leq r} \hat{x}_i^2/\varrho_i^2 = \epsilon_0, \tag{3.6}$$

v kateri so $\varrho_1, \varrho_2, \ldots, \varrho_r$ pozitivna števila in kjer je število ϵ_0 enako številu 1, številu -1 ali številu 0. Za neznanke $\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_n$ torej velja

$$\sum_{1 \leq i \leq r^+} (\tilde{x}_i - \xi_i)^2/\varrho_i^2 - \sum_{r^+ < i \leq r} (\tilde{x}_i - \xi_i)^2/\varrho_i^2 = \epsilon_0.$$

(2) Predpostavimo, da je $s = r + 1$. V tem primeru enačbo (3.5) delimo z absolutno vrednostjo $|\tilde{b}_{r+1}|$, torej dobimo

$$\sum_{1 \leq i \leq r^+} \big(\lambda_i^{\frac{1}{2}}/|\tilde{b}_{r+1}|^{\frac{1}{2}}\big)^2 \hat{x}_i^2 - \sum_{r^+ < i \leq r} \big(|\lambda_i|^{\frac{1}{2}}/|\tilde{b}_{r+1}|^{\frac{1}{2}}\big)^2 \hat{x}_i^2 + \big(\tilde{b}_{r+1}/|\tilde{b}_{r+1}|\big)\hat{x}_{r+1} = 0,$$

in definiramo $\varrho_i = |\tilde{b}_{r+1}/\lambda_i|^{\frac{1}{2}}$ za vsak $i = 1, 2, \ldots, r$. Naj bo $\epsilon = \tilde{b}_{r+1}/|\tilde{b}_{r+1}|$. S temi oznakami enačbo drugega reda (3.5) zapišemo v *standardni obliki*

$$\sum_{1 \leq i \leq r^+} \hat{x}_i^2/\varrho_i^2 - \sum_{r^+ < i \leq r} \hat{x}_i^2/\varrho_i^2 + \epsilon \hat{x}_{r+1} = 0, \tag{3.7}$$

v kateri so $\varrho_1, \varrho_2, \ldots, \varrho_r$ pozitivna števila in kjer je število ϵ enako številu 1 ali številu -1. Za neznanke $\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_n$ tako velja

$$\sum_{1 \le i \le r^+} (\tilde{x}_i - \xi_i)^2/\varrho_i^2 - \sum_{r^+ < i \le r} (\tilde{x}_i - \xi_i)^2/\varrho_i^2 + \epsilon(\tilde{x}_{r+1} - \xi_{r+1}) = 0.$$

Komentar 3.78. (1) Ker je Q ortogonalna matrika, je njena determinanta enaka številu 1 ali številu -1. V resnici lahko izberemo matriko Q tako, da je njena determinanta enaka številu 1. V nasprotnem primeru lahko namreč prvi stolpec ortogonalne matrike Q pomnožimo s številom -1 in s tem spremenimo predznak njene determinante, spremenjena matrika pa je še vedno ortogonalna in še vedno so njeni stolpci lastni vektorji matrike A. Ortogonalna matrika Q z determinanto 1 nam predstavlja rotacijo.

Koordinate, dane z bazo $\mathscr{B} = [Q\mathrm{e}_1, Q\mathrm{e}_2, \ldots, Q\mathrm{e}_n]$, so torej dobljene iz standardnih koordinat z rotacijo koordinatnega sistema z matriko Q, takšna pa je tudi zveza med vektorjem neznank X in vektorjem neznank $\tilde{X} = Q^{\mathsf{t}}X$. Ker je

$$\hat{X} = (\hat{x}_1, \hat{x}_2, \ldots, \hat{x}_n) = (\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_n) - (\xi_1, \xi_2, \ldots, \xi_n) = \tilde{X} - \Xi = Q^{\mathsf{t}}X - \Xi,$$

je zveza med vektorjem neznank X in vektorjem neznank $\hat{X}$ dana s kombinacijo rotacije koordinatnega sistema in translacije za vektor Ξ. Pripadajoča preslikava $F : \mathbb{R}^n \to \mathbb{R}^n$,

$$F(x_1, x_2, \ldots, x_n) = Q^{\mathsf{t}}(x_1, x_2, \ldots, x_n) - (\xi_1, \xi_2, \ldots, \xi_n) = (\hat{x}_1, \hat{x}_2, \ldots, \hat{x}_n),$$

sicer v splošnem ni linearna kot preslikava vektorskih prostorov, je pa kombinacija ortogonalne linearne preslikave ter translacije, zato ohranja razdalje med vektorji. Preslikava F je bijekcija, njen inverz pa je prav tako kombinacija translacije in rotacije, saj velja

$$(x_1, x_2, \ldots, x_n) = F^{-1}(\hat{x}_1, \hat{x}_2, \ldots, \hat{x}_n) = Q\big((\hat{x}_1, \hat{x}_2, \ldots, \hat{x}_n) + (\xi_1, \xi_2, \ldots, \xi_n)\big).$$

Tudi inverz F^{-1} seveda ohranja razdalje med vektorji.

Naj bo $\hat{\mathcal{S}} \subset \mathbb{R}^n$ hiperploskev drugega reda, ki jo dobimo kot množico vseh rešitev $\hat{X}$ enačbe (3.4) oziroma (3.5). Hiperploskev drugega reda $\mathcal{S} \subset \mathbb{R}^n$, ki je sestavljena iz rešitev X prvotne enačbe (3.1), dobimo z ustrezno rotacijo in translacijo hiperploskve $\hat{\mathcal{S}} \subset \mathbb{R}^n$, torej $\mathcal{S} = F^{-1}(\hat{\mathcal{S}})$. Ker preslikavi F in F^{-1} ohranjata razdalje, imata hiperploskvi $\mathcal{S}$ in $\hat{\mathcal{S}}$ enake geometrične lastnosti.

Točka $Q\Xi$ je *center* hiperploskve $\mathcal{S}$. Če je $s < n$, vektor Ξ ni enolično določen, zato ima v tem primeru hiperploskev $\mathcal{S}$ več centrov.

(2) Definirajmo nove neznanke $x_1', x_2', \ldots, x_n'$ s predpisom:

$$x_i' = \begin{cases} \hat{x}_i/\varrho_i & ; \quad 1 \le i \le r \\ \hat{x}_i & ; \quad r < i \le n \end{cases}$$

Nove neznanke x_i' smo torej dobili iz neznank $\hat{x}_i$ tako, da smo jih pomnožili s pozitivnimi konstantami, s tem pa smo povzročili razteg ali skrčitev koordinat v ustreznih smereh. Takšne deformacije sicer ne ohranjajo razdalj, pač pa le raztegnejo ali skrčijo množico rešitev v ustrezni smeri. Tip množice rešitev se pri tem ne spremeni. Z novimi oznakami lahko enačbo (3.6) oziroma (3.7) prepišemo v novo ekvivalentno, še enostavnejšo obliko

$$\sum_{1 \le i \le r^+} {x_i'}^2 - \sum_{r^+ < i \le r} {x_i'}^2 = \epsilon_0$$

v primeru $s = r$ oziroma

$$\sum_{1 \le i \le r^+} {x'_i}^2 - \sum_{r^+ < i \le r} {x'_i}^2 + \epsilon x'_{r+1} = 0$$

v primeru $s = r + 1$. Zamenjava predznaka števila $\epsilon \in \{1, -1\}$ le zrcali množico rešitev preko vektorskega podprostora $\{(x'_1, x'_2, \ldots, x'_n) \in \mathbb{R}^n \ ; \ x'_{r+1} = 0\}$, zato ta ne spremeni tipa množice rešitev. Tip množice rešitev je tako odvisen le od števil r^+, r in s ter od števila $\epsilon_0 \in \{1, -1, 0\}$.

(4) Množica rešitev $\mathcal{S}$ enačbe drugega reda (3.1) se ne spremeni, če to enačbo pomnožimo s poljubno neničelno konstantno. Pri tem se sicer spremeni enačbi prirejena matrika A in njene lastne vrednosti, vendar pa se razmerja med lastnimi vrednostmi ohranijo, prav tako pa se ohranijo lastni podprostori matrike A. Lastni podprostori matrike A imajo tesno zvezo s simetrijami množice $\mathcal{S}$. Enodimenzionalnim vektorskim podprostorom lastnih podprostorov matrike A pravimo tudi *glavne osi* hiperploskve drugega reda $\mathcal{S}$.

(5) Če je $s < n$, potem v enačbi (3.3) neznanke $\tilde{x}_{s+1}, \ldots, \tilde{x}_n$ sploh ne nastopajo. Hiperploskev drugega reda $\tilde{\mathcal{S}}$ v vektorskem prostoru $\mathbb{R}^n$, sestavljena iz rešitev $(\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_n)$ enačbe (3.3), je tedaj oblike

$$\tilde{\mathcal{S}} = \tilde{\mathcal{S}}^\flat \times \mathbb{R}^{n-s},$$

kjer je $\tilde{\mathcal{S}}^\flat$ hiperploskev drugega reda v vektorskem prostoru $\mathbb{R}^s$, sestavljena iz rešitev $(\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_s)$ enačbe (3.3). Podobno velja tudi za množico rešitev $\hat{\mathcal{S}}$ enačb (3.6) oziroma (3.7).

Zgled 3.79. (1) V vektorskem prostoru $\mathbb{R}$ je enačba drugega reda kar polinom druge stopnje v eni spremenljivki. Tipi hiperploskev drugega reda so v tem trivialnem primeru lahko:

(i) dve točki: $x^2/a^2 = 1$
(ii) prazna množica: $x^2/a^2 = -1$
(iii) ena točka: $x^2/a^2 = 0$

Pri tem je a pozitivna konstanta. Za hiperploskvi drugega reda iz točk (i) in (ii) pravimo, da sta *nedegenerirani*, hiperploskev drugega reda iz točke (iii) pa je *degenerirana*.

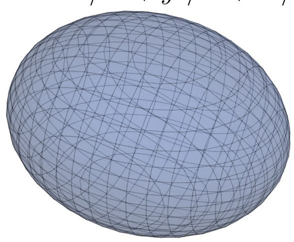
(2) Hiperploskvam drugega reda v $\mathbb{R}^2$ pravimo tudi *krivulje drugega reda*. Tipi krivulj drugega reda so naslednji:

(i) elipsa: $x^2/a^2 + y^2/b^2 = 1$
(ii) prazna množica *(imaginarna elipsa)*: $x^2/a^2 + y^2/b^2 = -1$
(iii) hiperbola: $x^2/a^2 - y^2/b^2 = 1$
(iv) parabola: $x^2/a^2 + \epsilon y = 0$
(v) točka *(par imaginarnih sekajočih se premic)*: $x^2/a^2 + y^2/b^2 = 0$
(vi) par sekajočih se premic: $x^2/a^2 - y^2/b^2 = 0$
(vii) par vzporednih si premic: $x^2/a^2 = 1$
(viii) prazna množica *(par imaginarnih vzporednih si premic)*: $x^2/a^2 = -1$
(ix) premica *(par enakih si premic)*: $x^2/a^2 = 0$

Pri tem sta a in b pozitivni konstanti in $\epsilon \in \{1, -1\}$. Krivulje drugega reda iz točk (i-iv) imenujemo *nedegenerirane*, preostale pa so *degenerirane*. V enačbah iz (vii-ix) neznanka y sploh ne nastopa, zato so to primeri produktov premice z ustrezno hiperploskvijo drugega reda v vektorskem prostoru $\mathbb{R}$.

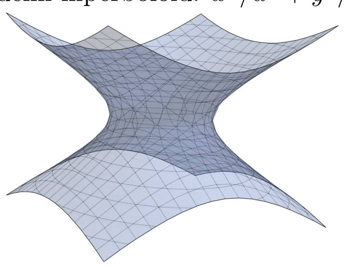
(3) Hiperploskvam drugega reda v $\mathbb{R}^3$ pravimo tudi *ploskve drugega reda.* Tipi ploskev drugega reda so naslednji:

(i) elipsoid: $x^2/a^2 + y^2/b^2 + z^2/c^2 = 1$

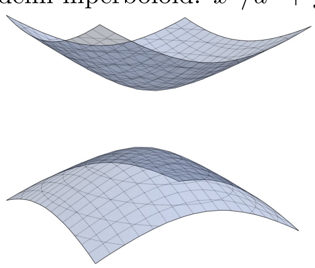


(ii) prazna množica *(imaginarni elipsoid)*: $x^2/a^2 + y^2/b^2 + z^2/c^2 = -1$

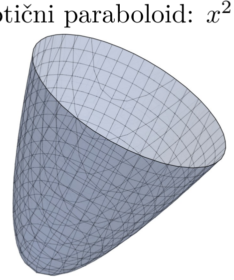
(iii) enodelni hiperboloid: $x^2/a^2 + y^2/b^2 - z^2/c^2 = 1$



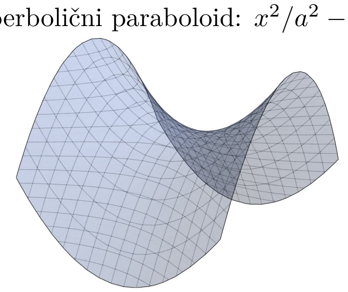
(iv) dvodelni hiperboloid: $x^2/a^2 + y^2/b^2 - z^2/c^2 = -1$



(v) eliptični paraboloid: $x^2/a^2 + y^2/b^2 + \epsilon z = 0$

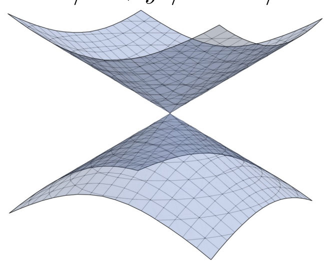


(vi) hiperbolični paraboloid: $x^2/a^2 - y^2/b^2 + \epsilon z = 0$

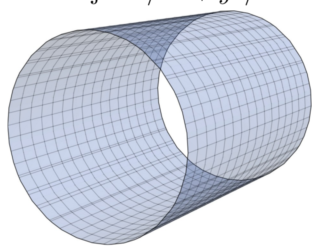


(vii) točka *(imaginarni stožec)*: $x^2/a^2 + y^2/b^2 + z^2/c^2 = 0$

(viii) stožec: $x^2/a^2 + y^2/b^2 - z^2/c^2 = 0$

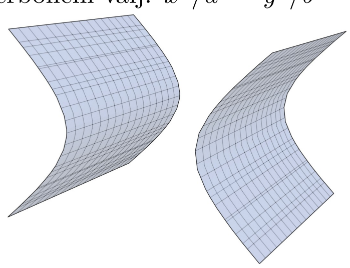


(ix) eliptični valj: $x^2/a^2 + y^2/b^2 = 1$

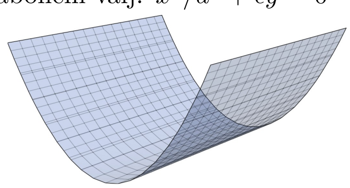


(x) prazna množica *(imaginarni eliptični valj)*: $x^2/a^2 + y^2/b^2 = -1$

(xi) hiperbolični valj: $x^2/a^2 - y^2/b^2 = 1$

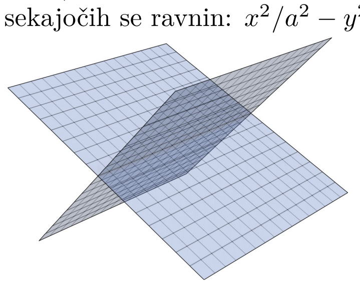


(xii) parabolični valj: $x^2/a^2 + \epsilon y = 0$

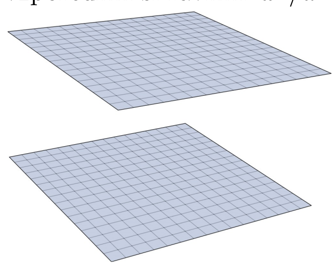


(xiii) premica *(par imaginarnih sekajočih se ravnin)*: $x^2/a^2 + y^2/b^2 = 0$

(xiv) par sekajočih se ravnin: $x^2/a^2 - y^2/b^2 = 0$

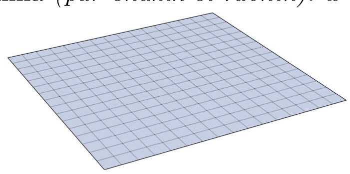


(xv) par vzporednih si ravnin: $x^2/a^2 = 1$



(xvi) prazna možica *(par imaginarnih vzporednih si ravnin)*: $x^2/a^2 = -1$

(xvii) ravnina *(par enakih si ravnin)*: $x^2/a^2 = 0$



Pri tem so a, b in c pozitivne konstante in $\epsilon \in \{1, -1\}$. Ploskve drugega reda iz točk (i-vi) imenujemo *nedegenerirane*, preostale pa so *degenerirane*. V enačbah iz (ix-xvii) neznanka z ne nastopa, zato so to primeri produktov premice z neko krivuljo drugega reda.

VAJA 3.80. Naj bo $\mathcal{S} \subset \mathbb{R}^3$ množica rešitev enačbe

$$6x^2 + 3y^2 - 12xy - 12yz + 4x + 2y - 4z = 0.$$

Pokaži, da je $\mathcal{S}$ enodelni hiperboloid z glavnimi osmi $\mathbb{R}(2, -2, 1)$, $\mathbb{R}(2, 1, -2)$ in $\mathbb{R}(1, 2, 2)$ ter s centrom v točki $\frac{1}{3}(-2, -1, 2)$.

Poglavje 4

Vektorske funkcije več spremenljivk

4.1. Metrični prostori

Pri predmetu Matematika 1 smo spoznali pojem zveznosti funkcije ene in več spremenljivk. Pri tem smo videli, da je za definicijo zveznosti ključen pojem razdalje med dvema točkama v vektorskem prostoru $\mathbb{R}$ oziroma $\mathbb{R}^n$. V prejšnjem poglavju smo definirali razdaljo med dvema vektorjema v poljubnem vektorskem prostoru s skalarnim produktom. V tem razdelku bomo spoznali, da lahko razdaljo med elementi smiselno definiramo tudi na poljubni množici.

Definicija 4.1. Naj bo $\mathcal{M}$ poljubna množica. Preslikavi $d : \mathcal{M} \times \mathcal{M} \to \mathbb{R}$ pravimo *metrika* na množici $\mathcal{M}$, če za vse elemente $x, y, z \in \mathcal{M}$ velja:

(i) $d(x, y) = d(y, x)$,
(ii) $d(x, z) \leq d(x, y) + d(y, z)$,
(iii) $d(x, y) \geq 0$,
(iv) $d(x, x) = 0$ in
(v) če $x \neq y$, potem je $d(x, y) > 0$.

Metrični prostor je množica $\mathcal{M}$, opremljena z izbrano metriko $d : \mathcal{M} \times \mathcal{M} \to \mathbb{R}$.

Elementom metričnega prostora običajno pravimo *točke*. Za nas najpomembnejši primeri metričnih prostorov bodo tisti, kjer je metrika podana z neko *normo* na vektorskem prostoru:

Definicija 4.2. Naj bo V vektorski prostor nad $\mathbb{F}$. Preslikavi $\| \bullet \| : \mathsf{V} \to \mathbb{R}$, $v \mapsto \|v\|$, pravimo *norma* na vektorskem prostoru V, če za vse vektorje $v, w \in \mathsf{V}$ in za vsak skalar $\alpha \in \mathbb{F}$ velja:

(i) $\|\alpha v\| = |\alpha| \|v\|$,
(ii) $\|v + w\| \leq \|v\| + \|w\|$,
(iii) $\|v\| \geq 0$,
(iv) $\|0\| = 0$ in
(v) če $v \neq 0$, potem je $\|v\| > 0$.

Normiran vektorski prostor je vektorski prostor V, opremljena z izbrano normo $\| \bullet \| : \mathsf{V} \to \mathbb{R}$.

Zgled 4.3. (1) Norma $\| \bullet \| : \mathsf{V} \to \mathbb{R}$ na vektorskem prostoru V nam inducira metriko $d : \mathsf{V} \times \mathsf{V} \to \mathbb{R}$ na V, ki je da na s predpisom

$$d(v, w) = \|v - w\|$$

za vsak $v, w \in \mathsf{V}$. Vsak normiran vektorski prostor je torej tudi metrični prostor glede na to inducirano metriko.

(2) Naj bo V vektorski prostor nad $\mathbb{F}$ s skalarnim produktom. Norma na V, inducirana s skalarnim produktom, je res norma s smislu zgornje definicije. Posebej

to pomeni, da je vsak vektorski prostor s skalarnim produktom tudi normiran vektorski prostor in torej tudi metrični prostor.

(3) Vektorski prostor $\mathbb{F}^n$ je opremljen s standardnim skalarnim produktom in s pridruženo standardno normo $|\bullet|$, ki jo včasih označimo tudi z $\|\bullet\|_2$. Ta standardna norma na vektorskem prostoru $\mathbb{F}^n$ inducira standardno metriko na $\mathbb{F}^n$, ki jo označimo z d_2. Vektorski prostor $\mathbb{F}^n$ je torej normiran vektorski prostor in tudi metrični prostor. Za poljubna vektorja $v = (\alpha_1, \alpha_2, \ldots, \alpha_n) \in \mathbb{F}^n$ in $w = (\beta_1, \beta_2, \ldots, \beta_n) \in \mathbb{F}^n$ je

$$\|v\|_2 = |v| = \sqrt{|\alpha_1|^2 + |\alpha_2|^2 + \cdots + |\alpha_n|^2}$$

in

$$d_2(v, w) = |v - w| = \sqrt{|\alpha_1 - \beta_1|^2 + |\alpha_2 - \beta_2|^2 + \cdots + |\alpha_n - \beta_n|^2}.$$

Na vektorskem prostoru $\mathbb{F}^n$ pa imamo tudi drugačne norme. Naj bo $p \in \mathbb{R}$ realno število, $p \geq 1$. Za vsak vektor $v = (\alpha_1, \alpha_2, \ldots, \alpha_n) \in \mathbb{F}^n$ definirajmo

$$\|v\|_p = \big(|\alpha_1|^p + |\alpha_2|^p + \cdots + |\alpha_n|^p\big)^{1/p}$$

in

$$\|v\|_\infty = \max\{|\alpha_1|, |\alpha_2|, \ldots, |\alpha_n|\}.$$

Ni težko preveriti, da so tako definirane preslikave $\|\bullet\|_\infty$ in $\|\bullet\|_p$, $p \geq 1$, res norme na vektorskem prostoru $\mathbb{F}^n$ in da za vsak vektor $v \in \mathbb{F}^n$ velja $\|v\|_\infty \leq \|v\|_p \leq n^{1/p}\|v\|_\infty$.

(4) Ker je vektorski prostor matrik $\mathrm{Mat}(m \times n, \mathbb{F})$ naravno izomorfen vektorskemu prostoru $\mathbb{F}^{nm}$, lahko po vzoru iz točke (3) tudi na njem definiramo norme $\|\bullet\|_\infty$ in $\|\bullet\|_p$, $p \geq 1$, s predpisom

$$\|A\|_p = \Big(\sum_{i=1}^{m}\sum_{j=1}^{n} |A_{ij}|^p\Big)^{1/p}$$

in

$$\|A\|_\infty = \max\{|A_{11}|, |A_{12}|, \ldots, |A_{mn}|\}$$

za vsako matriko $A \in \mathrm{Mat}(m \times n, \mathbb{F})$. Vendar pa imamo na vektorskem prostoru $\mathrm{Mat}(m \times n, \mathbb{F})$ še drugačno normo, ki je bistveno povezana z operacijo množenja na matrikah. Za vsako matriko $A \in \mathrm{Mat}(m \times n, \mathbb{F})$ definiramo

$$\|A\| = \sup\{|Au| \,;\, u \in \mathbb{F}^n,\ |u| = 1\}.$$

Ta definicija je smiselna, saj je množica $\{|Au| \,;\, u \in \mathbb{F}^n,\ |u| = 1\}$ omejena. Res, za vsak normiran vektor $u = (\gamma_1, \gamma_2, \ldots, \gamma_n) \in \mathbb{F}^n$ velja $|\gamma_j| \leq |u| = 1$ in

$$|Au| = |\sum_{j=1}^{n} \gamma_j A\mathrm{e}_j| \leq \sum_{j=1}^{n} |\gamma_j||A\mathrm{e}_j| \leq \sum_{j=1}^{n} |A\mathrm{e}_j|.$$

Poleg tega iz definicije sledi, da velja:

(i) preslikava $\|\bullet\|$ je norma na vektorskem prostoru $\mathrm{Mat}(m \times n, \mathbb{F})$,
(ii) $|Av| \leq \|A\||v|$,
(iii) $\|A\| = \inf\{\kappa \in \mathbb{R} \,;\, \kappa \geq 0,\ |Av| \leq \kappa|v| \text{ za vsak } v \in \mathbb{F}^n\}$,
(iv) $\|AB\| \leq \|A\|\|B\|$ in
(v) $\|A\|_\infty \leq \|A\| \leq \|A\|_2$.

za vse matrike $A \in \mathrm{Mat}(m \times n, \mathbb{F})$ in $B \in \mathrm{Mat}(n \times k, \mathbb{F})$ ter za vsak vektor $v \in \mathbb{F}^n$. Normo $\| \bullet \|$ imenujemo tudi *operatorska norma* na vektorskem prostoru matrik. Ker velja točka (iv), pravimo, da je operatorska norma *submultiplikativna*.

Prepričajmo se, da trditve (i-v) res držijo. Če je $v = 0$, trditev (ii) seveda velja, sicer pa vzemimo $u = v/|v|$ in ocenimo

$$|Av| = |Au||v| \leq \|A\||v|,$$

kar je neenakost iz točke (ii). Posebej to pomeni, da je število $\|A\|$ element množice

$$\mathcal{D} = \{\kappa \in \mathbb{R} \ ; \ \kappa \geq 0, \ |Av| \leq \kappa |v| \text{ za vsak } v \in \mathbb{F}^n\}.$$

Po drugi strani pa je vsako število $\kappa \in \mathcal{D}$ zgornja meja za množico

$$\{|Au| \ ; \ u \in \mathbb{F}^n, \ |u| = 1\},$$

zato je $\|A\| \leq \kappa$. S tem smo dokazali trditev iz točke (iii).

Dokažimo zdaj točko (i). Za vsak skalar $\alpha \in \mathbb{F}$ in za poljubno matriko $A \in \mathrm{Mat}(m \times n, \mathbb{F})$ velja:

$$\begin{aligned} \|\alpha A\| &= \sup\{|\alpha A u| \ ; \ u \in \mathbb{F}^n, \ |u| = 1\} = \sup\{|\alpha||Au| \ ; \ u \in \mathbb{F}^n, \ |u| = 1\} \\ &= |\alpha| \sup\{|Au| \ ; \ u \in \mathbb{F}^n, \ |u| = 1\} = |\alpha|\|A\|. \end{aligned}$$

Za poljubni matriki $A, A' \in \mathrm{Mat}(m \times n, \mathbb{F})$ in vsak vektor $v \in \mathbb{F}^n$ velja tudi

$$|(A + A')v| = |Av + A'v| \leq |Av| + |A'v| \leq \|A\||v| + \|A'\||v| = (\|A\| + \|A'\|)|v|,$$

zato je $\|A + A'\| \leq \|A\| + \|A'\|$ po točki (iii). Očitno je $\|A\| \geq 0$ in $\|0\| = 0$. Če je $\|A\| = 0$, velja $Av = 0$ za vsak vektor $v \in \mathbb{F}^n$, zato je tedaj $A = 0$. S tem je točka (i) dokazana.

Točka (iv) je posledica točke (iii) in neenakosti

$$|ABw| \leq \|A\||Bw| \leq \|A\|\|B\||w|,$$

ki velja za vsak vektor $w \in \mathbb{F}^k$. Prvi del točke (v) sledi iz ocene

$$|A_{ij}| \leq |A\mathrm{e}_j| \leq \|A\|,$$

ki velja za vsak $i = 1, 2, \ldots, m$ in za vsak $j = 1, 2, \ldots, n$. Poleg tega za poljuben vektor $v = (\alpha_1, \alpha_2, \ldots, \alpha_n) \in \mathbb{F}^n$ z uporabo Cauchy-Schwarzeve neenakosti dobimo oceno

$$\begin{aligned} |Av|^2 = \sum_{i=1}^m |(Av)_{i1}|^2 &= \sum_{i=1}^m \Big| \sum_{j=1}^n A_{ij}\alpha_j \Big|^2 = \sum_{i=1}^m \big| \big\langle (A_{i1}, \ldots, A_{in}), (\bar{\alpha}_1, \ldots, \bar{\alpha}_n) \big\rangle \big|^2 \\ &\leq \sum_{i=1}^m \big| (A_{i1}, \ldots, A_{in}) \big|^2 \big| (\bar{\alpha}_1, \ldots, \bar{\alpha}_n) \big|^2 = \sum_{i=1}^m \sum_{j=1}^n |A_{ij}|^2 |v|^2 = \|A\|_2^2 |v|^2, \end{aligned}$$

odtod pa po točki (iii) sledi še drugi del točke (v).

(5) Naj bo $\mathcal{M}$ metrični prostor z metriko d in naj bo $\mathcal{N} \subset \mathcal{M}$ podmnožica metričnega prostora $\mathcal{M}$. Tedaj je zožitev $d|_{\mathcal{N} \times \mathcal{N}}$ metrika na množici $\mathcal{N}$. Podmnožica $\mathcal{N}$, opremljena z zoženo metriko, je torej spet metrični prostor, ki ga imenujemo tudi *metrični podprostor* metričnega prostora $\mathcal{M}$.

Posebej je torej vsaka podmnožica vektorskega prostora $\mathbb{F}^n$ metrični prostor glede na metriko, ki jo dobimo kot zožitev standardne metrike vektorskega prostora $\mathbb{F}^n$. Tako sta *enotska n-dimenzionalna krogla*

$$\mathbb{B}^n = \{v \in \mathbb{R}^n \ ; \ |v| \leq 1\} \subset \mathbb{R}^n$$

in *enotska* $(n-1)$*-dimenzionalna sfera*

$$\mathbb{S}^{n-1} = \{v \in \mathbb{R}^n \ ; \ |v| = 1\} \subset \mathbb{R}^n$$

metrična podprostora metričnega prostora $\mathbb{R}^n$ s standardno metriko. Vektorski prostor matrik $\mathrm{Mat}(m \times n, \mathbb{F})$ je opremljen z operatorsko normo in s pridruženo operatorsko metriko, zato so grupe $\mathrm{GL}(n, \mathbb{F})$ in $\mathrm{O}(n)$ oziroma $\mathrm{U}(n)$ metrični podprostori metričnega prostora $\mathrm{Mat}(m \times n, \mathbb{F})$.

TRDITEV 4.4. *Naj bo $A \in \mathrm{GL}(n, \mathbb{F})$ obrnljiva matrika velikosti $n \times n$ in naj bo $B \in \mathrm{Mat}(n \times n, \mathbb{F})$ takšna matrika, da velja $\|A - B\| < 1/\|A^{-1}\|$. Tedaj je tudi matrika B obrnljiva.*

DOKAZ. Označimo $\alpha = 1/\|A^{-1}\|$ in $\beta = \|A - B\|$. Velja torej $0 \leq \beta < \alpha$. Za vsak vektor $v \in \mathbb{F}^n$ je

$$\begin{aligned} \alpha|v| &= \alpha|A^{-1}Av| \leq \alpha\|A^{-1}\||Av| = |Av| = |(A-B)v + Bv| \\ &\leq |(A-B)v| + |Bv| \leq \|A-B\||v| + |Bv| = \beta|v| + |Bv| \end{aligned}$$

in zato $(\alpha - \beta)|v| \leq |Bv|$. Če je $Bv = 0$, sledi torej $(\alpha - \beta)|v| = 0$ in zato $v = 0$. S tem smo dokazali, da ima kvadratna matrika B trivialno jedro, odtod pa sledi, da je obrnljiva. □

4.1.1. Odprte krogle, odprte množice in okolice. Naj bo $\mathcal{M}$ metrični prostor z metriko d. Za poljubno točko $x \in \mathcal{M}$ in poljubno pozitivno realno število ϵ množico

$$\mathcal{K}(x, \epsilon) = \{y \in \mathcal{M} \ ; \ d(x, y) < \epsilon\}$$

imenujemo *odprta krogla* v metričnem prostoru $\mathcal{M}$ s središčem v x in s polmerom ϵ. Podmnožica $\mathcal{U} \subset \mathcal{M}$ je *okolica* točke $x \in \mathcal{M}$, če obstaja takšno pozitivno realno število ϵ, da velja $\mathcal{K}(x, \epsilon) \subset \mathcal{U}$. Točka $x \in \mathcal{M}$ je *notranja točka* podmnožice $\mathcal{U} \subset \mathcal{M}$, če je $\mathcal{U}$ okolica točke x. Podmnožica $\mathcal{U} \subset \mathcal{M}$ je *odprta* v $\mathcal{M}$, če je okolica vsake svoje točke. Podmnožica $\mathcal{W} \subset \mathcal{M}$ je *zaprta* v $\mathcal{M}$, če je njen komplement $\mathcal{M} \setminus \mathcal{W}$ odprta podmnožica v $\mathcal{M}$.

VAJA 4.5. (1) Naj bo $\mathcal{M}$ metrični prostor. Pokaži, da je za poljubno točko $x \in \mathcal{M}$ in poljubno pozitivno realno število ϵ odprta krogla $\mathcal{K}(x, \epsilon)$ odprta podmnožica metričnega prostora $\mathcal{M}$.

(2) Naj bo $\mathcal{M}$ metrični prostor. Pokaži, da velja:

- (i) podmnožici $\emptyset$ in $\mathcal{M}$ sta odprti v $\mathcal{M}$,
- (ii) presek vsakih dveh odprtih podmnožic metričnega prostora $\mathcal{M}$ je odprta podmnožica v $\mathcal{M}$, in
- (iii) unija vsake družine odprtih podmnožic metričnega prostora $\mathcal{M}$ je odprta podmnožica v $\mathcal{M}$.

Ker veljajo lastnosti (i-iii), družini vseh odprtih podmnožic metričnega prostora $\mathcal{M}$ pravimo *topologija* metričnega prostora $\mathcal{M}$.

(3) Naj bosta d in d' metriki na isti množici $\mathcal{M}$. Če so v metričnem prostoru $(\mathcal{M}, d)$ odprte natanko tiste podmnožice množice $\mathcal{M}$, ki so odprte v metričnem prostoru $(\mathcal{M}, d')$, potem pravimo, da sta si metriki d in d' *ekvivalentni.* Drugače povedano, metriki d in d' sta si ekvivalentni, če imata metrična prostora $(\mathcal{M}, d)$ in $(\mathcal{M}, d')$ isto topologijo.

Naj bo V vektorski prostor nad $\mathbb{F}$. Dve normi $\| \bullet \|$ in $\| \bullet \|'$ na V sta si *ekvivalentni*, če obstajata takšni pozitivni realni konstanti κ in κ', da za vsak vektor $v \in \mathsf{V}$

velja $\|v\|' \le \kappa\|v\|$ in $\|v\| \le \kappa'\|v\|'$. Pokaži, da velja naslednje: Če sta si normi $\|\bullet\|$ in $\|\bullet\|'$ na V ekvivalentni, potem sta si pripadajoči metriki ekvivalentni.

4.1.2. Zvezne preslikave med metričnimi prostori. Naj bosta $\mathcal{M}$ in $\mathcal{M}'$ metrična prostora. Preslikava $f : \mathcal{M} \to \mathcal{M}'$ je *zvezna v točki* $a \in \mathcal{M}$, če za vsako pozitivno realno število ϵ obstaja takšno pozitivno realno število δ, da velja

$$f(\mathcal{K}(a,\delta)) \subset \mathcal{K}(f(a),\epsilon).$$

Preslikava $f : \mathcal{M} \to \mathcal{M}'$ je *zvezna*, če je zvezna v vsaki točki $a \in \mathcal{M}$.

Zgled 4.6. (1) Identična preslikava in konstantna preslikava sta zvezni preslikavi.

(2) Naj bo $\mathcal{D}$ podmnožica prostora $\mathbb{R}^n$. Vektorska funkcija n spremenljivk $g : \mathcal{D} \to \mathbb{R}^m$ je zvezna (v točki $a \in \mathcal{D}$), če, in samo če, je zvezna kot preslikava med metričnima prostoroma (v točki a), pri čemer prostora $\mathbb{R}^n$ in $\mathbb{R}^m$ opremimo s standardno metriko. Spomnimo se, da takšno vektorsko funkcijo $g : \mathcal{D} \to \mathbb{R}^m$ zapišemo kot urejeno m-terico $g = (g_1, \ldots, g_m)$ njenih komponent, ki so skalarne (realne) funkcije $g_1, \ldots, g_m : \mathcal{D} \to \mathbb{R}$ n spremenljivk, torej

$$g(x_1, \ldots, x_n) = (g_1(x_1, \ldots, x_n), \ldots, g_m(x_1, \ldots, x_n))$$

za vse točke $(x_1, \ldots, x_n) \in \mathcal{D}$. Funkcija g je zvezna, če, in samo če, so vse njene komponente $g_1, \ldots, g_m$ zvezne.

Trditev 4.7. *Naj bosta $\mathcal{M}$ in $\mathcal{M}'$ metrična prostora. Preslikava $f : \mathcal{M} \to \mathcal{M}'$ je zvezna, če, in samo če, je množica $f^{-1}(\mathcal{V})$ odprta v $\mathcal{M}$ za vsako odprto podmnožico $\mathcal{V}$ metričnega prostora $\mathcal{M}'$.*

Dokaz. Najprej predpostavimo, da je preslikava f zvezna. Naj bo $\mathcal{V}$ poljubna odprta podmnožica metričnega prostora $\mathcal{M}'$. Pokazati moramo, da je praslika $f^{-1}(\mathcal{V})$ odprta v $\mathcal{M}$. V ta namen izberimo poljubno točko $a \in f^{-1}(\mathcal{V})$. Ker je podmnožica $\mathcal{V}$ odprta v $\mathcal{M}$, obstaja tako pozitivno realno število ϵ, da velja $\mathcal{K}(f(a),\epsilon) \subset \mathcal{V}$. Ker je preslikava f zvezna v točki a, obstaja takšno pozitivno realno število δ, da je $f(\mathcal{K}(a,\delta)) \subset \mathcal{K}(f(a),\epsilon)$. Posebej odtod sledi, da je $f(\mathcal{K}(a,\delta)) \subset \mathcal{V}$ oziroma $\mathcal{K}(a,\delta) \subset f^{-1}(\mathcal{V})$. Ker to velja za vsako točko $a \in f^{-1}(\mathcal{V})$, je množica $f^{-1}(\mathcal{V})$ odprta v $\mathcal{M}$.

Pokažimo še obratno. Predpostavimo, da pogoj iz trditve velja, in dokažimo, da je preslikava f zvezna. Naj bo a poljubna točka iz $\mathcal{M}$ in naj bo ϵ poljubno pozitivno realno število. Ker je odprta krogla $\mathcal{K}(f(a),\epsilon)$ odprta podmnožica prostora $\mathcal{M}'$, je po predpostavki praslika $f^{-1}(\mathcal{K}(f(a),\epsilon))$ odprta podmnožica prostora $\mathcal{M}$. To posebej pomeni, da obstaja takšno pozitivno realno število δ, da je $\mathcal{K}(a,\delta) \subset f^{-1}(\mathcal{K}(f(a),\epsilon))$, torej $f(\mathcal{K}(a,\delta)) \subset \mathcal{K}(f(a),\epsilon)$. S tem smo dokazali, da je preslikava f zvezna v točki a. Ker to velja za vsako točko $a \in \mathcal{M}$, je preslikava f torej zvezna. $\square$

Komentar 4.8. Zadnja trditev pove, da je zveznost preslikav med metričnimi prostori odvisna le od topologij danih metričnih prostorov. Drugače povedano, če sta d in d'' ekvivalentni metriki na množici $\mathcal{M}$ ter sta d' in d''' ekvivalentni metriki na množici $\mathcal{M}'$, potem je preslikava $f : \mathcal{M} \to \mathcal{M}'$ zvezna kot preslikava iz $(\mathcal{M}, d)$ v $(\mathcal{M}', d')$, če, in samo če, je zvezna kot preslikava iz $(\mathcal{M}, d'')$ v $(\mathcal{M}', d''')$.

Posledica 4.9. *Kompozicija zveznih preslikav med metričnimi prostori je zvezna preslikava.*

Zaporedje v metričnem prostoru $\mathcal{M}$ je preslikava $\mathbb{N} \to \mathcal{M}$, $k \mapsto a_k$, ki jo običajno zapišemo kot $(a_k)_{k\in\mathbb{N}} = (a_k)_k = (a_k) = (a_1, a_2, \ldots)$. Za poljuben $j \in \mathbb{N}$ je j-ti člen zaporedja (a_k) točka $a_j \in \mathcal{M}$.

Naj bo (a_k) zaporedje v metričnem prostoru $\mathcal{M}$. Točka $a \in \mathcal{M}$ je *stekališče* zaporedja (a_k), če v vsaki okolici točke a v $\mathcal{M}$ leži neskončno členov zaporedja (a_k), oziroma natančneje, če je za vsako okolico $\mathcal{U}$ točke a v $\mathcal{M}$ množica $\{j \in \mathbb{N}\ ;\ a_j \in \mathcal{U}\}$ neskončna.

Točka $a \in \mathcal{M}$ je *limita* zaporedja (a_k), če v vsaki okolici točke a v $\mathcal{M}$ ležijo vsi členi zaporedja (a_k) z izjemo končno mnogih, oziroma natančneje, če je za vsako okolico $\mathcal{U}$ točke a v $\mathcal{M}$ množica $\{j \in \mathbb{N}\ ;\ a_j \notin \mathcal{U}\}$ končna. V tem primeru pravimo, da je zaporedje (a_k) *konvergentno* in konvergira k limiti a, ki jo označimo tudi

$$a = \lim_{k\to\infty} a_k.$$

Če zaporedje ni konvergentno, je *divergentno*.

Podzaporedje zaporedja (a_k) v $\mathcal{M}$ je zaporedje oblike

$$(a_{k_j})_j,$$

kjer je $(k_j)_j$ poljubno strogo naraščajoče zaporedje naravnih števil.

Zaporedje (a_k) v metričnem prostoru $(\mathcal{M}, d)$ je *Cauchyjevo*, če za vsako pozitivno realno število ϵ obstaja tako veliko naravno število N, da za vsaki dve naravni števili $k, j \geq N$ velja $d(a_k, a_j) < \epsilon$. Metrični prostor $\mathcal{M}$ je *poln*, če je vsako Cauchyjevo zaporedje v $\mathcal{M}$ konvergentno.

Vaja 4.10. (1) Pri Matematiki 1 smo spoznali nekaj osnovnih lastnosti zaporedij realnih števil. Z enakimi argumenti in dokazi se lahko prepričamo, da večina teh lastnosti velja tudi za zaporedja v poljubnem metričnem prostoru. Posebej tako za zaporedja v poljubnem metričnem prostoru velja:

(i) Limita zaporedja je tudi stekališče tega zaporedja, obratno pa ni nujno res.

(ii) Konvergentno zaporedje ima natanko eno limito, ki je tudi njegovo edino stekališče.

(iii) Vsako stekališče podzaporedja zaporedja (a_k) je stekališče zaporedja (a_k).

(iv) Vsako podzaporedje konvergentnega zaporedja z limito a je konvergentno z limito a.

(v) Točka a je stekališče zaporedja (a_k), če, in samo če, obstaja konvergentno podzaporedje zaporedja (a_k) z limito a.

(vi) Vsako konvergentno zaporedje je Cauchyjevo. Obratna trditev v splošnem ne velja, primer metričnega prostora, ki ni poln, je že vsak neprazen odprt omejen interval v $\mathbb{R}$. Pač pa iz Matematike 1 vemo, da je metrični prostor $\mathbb{R}$ poln, podobno pa tudi vidimo, da je metrični prostor $\mathbb{R}^n$ poln.

(2) Na enak način kot za funkcije ene realne spremenljivke se lahko prepričamo, da je preslikava $f : \mathcal{M} \to \mathcal{M}'$ med metričnima prostoroma zvezna v točki $a \in \mathcal{M}$, če, in samo če, za vsako zaporedje (a_k) v $\mathcal{M}$, ki konvergira k točki a, zaporedje $(f(a_k))_k$ v $\mathcal{M}'$ konvergira k točki $f(a)$.

(3) Naj bo (a_k) zaporedje vektorjev v $\mathbb{R}^n$. Za vsak $k \in \mathbb{N}$ zapišemo vektor a_k s komponentami kot

$$a_k = (\alpha_{k,1}, \alpha_{k,2}, \ldots, \alpha_{k,n}).$$

Posebej so torej $(\alpha_{k,1})_k, (\alpha_{k,2})_k, \ldots, (\alpha_{k,n})_k$ zaporedja realnih števil. Prepričaj se, da je zaporedje (a_k) konvergentno, če, in samo če, so zaporedja komponent

$(\alpha_{k,1})_k, (\alpha_{k,2})_k, \ldots, (\alpha_{k,n})_k$ vsa konvergentna, in da v tem primeru velja

$$\lim_{k\to\infty} a_k = \big(\lim_{k\to\infty} \alpha_{k,1}, \lim_{k\to\infty} \alpha_{k,2}, \ldots, \lim_{k\to\infty} \alpha_{k,n}\big).$$

4.1.3. Kompaktni metrični prostori. Metrični prostor $\mathcal{M}$ je *kompakten*, če ima vsako zaporedje v $\mathcal{M}$ vsaj eno stekališče v $\mathcal{M}$.

ZGLED 4.11. Prvi primeri kompaktnih metričnih prostorov so končni metrični prostori in zaprti omejeni intervali $[\![a, b]\!] \subset \mathbb{R}$ glede na standardno metriko.

Metrični prostor $\mathcal{M}$ je *omejen*, če je metrika d na $\mathcal{M}$ omejena funkcija, kar pomeni, da obstaja takšna realna konstanta κ, da je $d(x, y) \le \kappa$ za vse $x, y \in \mathcal{M}$.

Spomnimo se, da je podmnožica $\mathcal{N}$ metričnega prostora $\mathcal{M}$ tudi metrični prostor glede na zoženo metriko. Pravimo, da je podmnožica $\mathcal{N}$ metričnega prostora $\mathcal{M}$ kompaktna, če je metrični prostor $\mathcal{N}$ z zoženo metriko kompakten, in da je podmnožica $\mathcal{N}$ metričnega prostora $\mathcal{M}$ omejena, če je če je metrični prostor $\mathcal{N}$ z zoženo metriko omejen.

TRDITEV 4.12. *Naj bo $\mathcal{N}$ podmnožica metričnega prostora $\mathcal{M}$.*
(i) Če je podmnožica $\mathcal{N}$ kompaktna, potem je zaprta v $\mathcal{M}$.
(ii) Če je $\mathcal{M}$ kompakten in je $\mathcal{N}$ zaprta v $\mathcal{M}$, potem je $\mathcal{N}$ kompaktna.
(iii) Če je prostor $\mathcal{M}$ kompakten, potem je omejen.

DOKAZ. (i) Trditev bomo dokazali s protislovjem. Predpostavimo torej, da množica $\mathcal{N}$ ni zaprta. To pomeni, da obstaja točka $a \in \mathcal{M} \setminus \mathcal{N}$, ki ni notranja točka množice $\mathcal{M} \setminus \mathcal{N}$. Za vsak $k \in \mathbb{N}$ torej odprta krogla $\mathcal{K}(a, 1/k)$ seka množico $\mathcal{N}$ in lahko izberemo točko $a_k \in \mathcal{K}(a, 1/k) \cap \mathcal{N}$. Tako je (a_k) zaporedje točk iz $\mathcal{N}$, ki konvergira k točki $a \in \mathcal{M}$. Posebej je a edino stekališče zaporedja (a_k), a ker točka a ni v množici $\mathcal{N}$, zaporedje (a_k) torej nima stekališč v $\mathcal{N}$, kar pa je v protislovju s kompaktnostjo množice $\mathcal{N}$.

(ii) Naj bo (a_k) poljubno zaporedje točk iz $\mathcal{N}$. Ker je prostor $\mathcal{M}$ kompakten, ima zaporedje (a_k) stekališče a v $\mathcal{M}$. Pokazati moramo, da je točka a v množici $\mathcal{N}$. Pa predpostavimo obratno, da točka a ne leži v množici $\mathcal{N}$. V tem primeru bi bila zaradi zaprtosti množice $\mathcal{N}$ množica $\mathcal{M} \setminus \mathcal{N}$ okolica točke a, kar pa ne more biti res, saj je a stekališče zaporedja (a_k) in bi zato moralo biti v okolici $\mathcal{M} \setminus \mathcal{N}$ točke a neskončno členov tega zaporedja.

(iii) Pa predpostavimo, da prostor $\mathcal{M}$ z metriko d ni omejen. Posebej to pomeni, da obstaja takšno zaporedje (a_k) v $\mathcal{M}$, da je $d(a_k, a_1) \ge k$ za vsak $k \in \mathbb{N}$. Trikotniška neenakost nam pokaže, da takšno zaporedje (a_k) ni Cauchyjevo, enako pa ni Cauchyjevo nobeno podzaporedje zaporedja (a_k). Odtod sledi, da nobeno podzaporedje zaporedja (a_k) ni konvergentno in da je torej zaporedje (a_k) brez stekališč, kar pa je v protislovju s kompaktnostjo prostora $\mathcal{M}$. $\square$

TRDITEV 4.13. *Podmnožica prostora $\mathbb{R}^n$ je kompaktna, če, in samo če, je omejena in zaprta v $\mathbb{R}^n$.*

DOKAZ. Če je podmnožica prostora $\mathbb{R}^n$ kompaktna, po Trditvi 4.12 sledi, da je omejena in zaprta v $\mathbb{R}^n$. Obratno, predpostavimo da je $\mathcal{N}$ omejena in zaprta podmnožica prostora $\mathbb{R}^n$. Naj bo (a_k) poljubno zaporedje točk iz $\mathcal{N}$, in zapišimo

$$a_k = (\alpha_{k,1}, \alpha_{k,2}, \ldots, \alpha_{k,n}).$$

Ker je množica $\mathcal{N}$ omejena, so zaporedja realnih števil $(\alpha_{k,1})_k, (\alpha_{k,2})_k, \ldots, (\alpha_{k,n})_k$ vsa omejena in ima torej vsako od njih vsaj eno stekališče v $\mathbb{R}$.

Zdaj lahko izberemo konvergentno podzaporedje (b_l) zaporedja (a_k), in sicer takole: Ker ima zaporedje prvih komponent $(\alpha_{k,1})_k$ stekališče, lahko najprej izberemo podzaporedje (b_l) zaporedja (a_k) tako, da je zaporedje prvih komponent $(\beta_{l,1})_l$ zaporedja (b_l) konvergentno. V naslednjem koraku zaporedje (b_l) nadomestimo z njegovim podzaporedjem tako, da tudi zaporedje drugih komponent $(\beta_{l,2})_l$ konvergira. Tako nadaljujemo in nazadnje zaporedje (b_l) nadomestimo z njegovim podzaporedjem tako, da tudi zaporedje zadnjih komponent $(\beta_{l,n})_l$ konvergira. Na ta način smo dobili takšno podzaporedje (b_l) zaporedja (a_k), da vsa njegova zaporedja komponent konvergirajo, takšno zaporedje (b_l) pa je konvergentno v $\mathbb{R}^n$.

Limita $a \in \mathbb{R}^n$ podzaporedja (b_l) je stekališče zaporedja (a_k). Ostane nam še, da se prepričamo, da je $a \in \mathcal{N}$. Predpostavimo obratno, da točka a ne leži v množici $\mathcal{N}$. Zaradi zaprtosti množice $\mathcal{N}$ bi bila tedaj množica $\mathbb{R}^n \setminus \mathcal{N}$ okolica točke a, kar pa ne more biti res, saj je a stekališče zaporedja (a_k) in bi zato moralo biti v okolici $\mathbb{R}^n \setminus \mathcal{N}$ točke a neskončno členov tega zaporedja. □

TRDITEV 4.14. *Naj bo $f : \mathcal{M} \to \mathcal{M}'$ zvezna preslikava med metričnima prostoroma. Če je metrični prostor $\mathcal{M}$ kompakten, potem je slika $f(\mathcal{M})$ kompaktna podmnožica metričnega prostora $\mathcal{M}'$.*

DOKAZ. Naj bo (a'_k) poljubno zaporedje točk iz množice $f(\mathcal{M})$. Za vsak $k \in \mathbb{N}$ torej lahko izberemo takšno točko a_k iz $\mathcal{M}$, da je $f(a_k) = a'_k$. Ker je (a_k) zaporedje točk v kompaktnem metričnem prostoru $\mathcal{M}$, ima to zaporedje stekališče a v $\mathcal{M}$. Izberemo lahko podzaporedje $(a_{k_l})_l$ zaporedja (a_k), ki konvergira k točki a. Ker je preslikava f zvezna, velja

$$\lim_{l\to\infty} a'_{k_l} = \lim_{l\to\infty} f(a_{k_l}) = f(\lim_{l\to\infty} a_{k_l}) = f(a).$$

Odtod sledi, da je $f(a)$ stekališče zaporedja (a'_k). □

POSLEDICA 4.15. *Naj bo $\mathcal{M}$ kompakten metrični prostor in $f : \mathcal{M} \to \mathbb{R}$ zvezna funkcija. Tedaj obstajata takšni točki $a, b \in \mathcal{M}$, da je $f(a)$ najmanjša vrednost funkcije f in da je $f(b)$ največja vrednost funkcije f, torej da velja*

$$f(a) \leq f(x) \leq f(b)$$

za vse $x \in \mathcal{M}$.

VAJA 4.16. (1) Pokaži, da je vsak kompakten metrični prostor poln.

(2) Naj bosta $(\mathcal{M}, d)$ in $(\mathcal{M}', d')$ metrična prostora in $f : \mathcal{M} \to \mathcal{M}'$ preslikava med njima. Preslikava f je *enakomerno zvezna*, če za vsako pozitivno realno število ϵ obstaja takšno pozitivno realno število δ, da za vsaki točki $x, y \in \mathcal{M}$, za kateri je $d(x, y) < \delta$, velja $d'(f(x), f(y)) < \epsilon$. Pokaži:

(i) Če je preslikava f enakomerno zvezna, potem je tudi zvezna.

(ii) Če je preslikava f zvezna in če je prostor $\mathcal{M}$ kompakten, potem je f enakomerno zvezna.

4.1.4. S potmi povezani metrični prostori. *Pot* v metričnem prostoru $\mathcal{M}$ od točke $a \in \mathcal{M}$ do točke $b \in \mathcal{M}$ je zvezna preslikava $\gamma : [\![0, 1]\!] \to \mathcal{M}$, za katero je $\gamma(0) = a$ in $\gamma(1) = b$. Metrični prostor $\mathcal{M}$ je *povezan s potmi*, če za vsaki dve točki $a, b \in \mathcal{M}$ obstaja pot v $\mathcal{M}$ od a do b.

ZGLED 4.17. Interval je podmnožica $\mathcal{J} \subset \mathbb{R}$ z lastnostjo, da za poljubni dve točki $a, b \in \mathcal{J}$ velja $[\![a, b]\!] \subset \mathcal{J}$. Iz osnovnih lastnosti zveznih funkcij ene realne spremenljivke, ki jih poznamo iz Matematike 1, direktno sledi, da so intervali natanko vse s potmi povezane podmnožice prostora $\mathbb{R}$. Posebej so zaprti omejeni intervali natanko vse kompaktne s potmi povezane podmnožice prostora $\mathbb{R}$.

TRDITEV 4.18. *Naj bo $f : \mathcal{M} \to \mathcal{M}'$ zvezna preslikava med metričnima prostoroma in predpostavimo, da je metrični prostor $\mathcal{M}$ povezan s potmi. Tedaj je slika $f(\mathcal{M})$ s potmi povezan metrični prostor.*

DOKAZ. Naj bosta a' in b' poljubni točki iz $f(\mathcal{M})$. Izberemo lahko točki $a, b \in \mathcal{M}$ tako da je $f(a) = a'$ in $f(b) = b'$. Ker je prostor $\mathcal{M}$ povezan s potmi, obstaja pot γ v $\mathcal{M}$ od a do b. Kompozicija $f \circ \gamma$ je tedaj pot v $f(\mathcal{M})$ od a' do b'. □

POSLEDICA 4.19. *Naj bo $\mathcal{M}$ s potmi povezan metrični prostor in naj bo $f : \mathcal{M} \to \mathbb{R}$ zvezna funkcija. Tedaj velja:*

(i) Za poljubni dve števili $t, s \in f(\mathcal{M})$, $t \leq s$, je $[\![t, s]\!] \subset f(\mathcal{M})$.

(ii) Če obstajata takšni točki $a, b \in \mathcal{M}$, da je $f(a)f(b) < 0$, potem obstaja takšna točka $c \in \mathcal{M}$, da velja $f(c) = 0$.

POSLEDICA 4.20. *Če je $\mathcal{M}$ kompakten s potmi povezan metrični prostor in $f : \mathcal{M} \to \mathbb{R}$ zvezna funkcija, potem je slika $f(\mathcal{M})$ zaprt omejen interval.*

4.1.5. Banachov izrek o negibni točki. Naj bo sta $(\mathcal{M}, d)$ in $(\mathcal{M}', d')$ metrična prostora. Preslikava $f : \mathcal{M} \to \mathcal{M}'$ je *skrčitev*, če obstaja takšna realna konstanta q, da je $0 \leq q < 1$ in da velja

$$d'(f(x), f(y)) \leq qd(x, y)$$

za vse točke $x, y \in \mathcal{M}$.

VAJA 4.21. Pokaži, da je vsaka skrčitev zvezna.

IZREK 4.22 (Banachov izrek o negibni točki). *Naj bo $\mathcal{M}$ poln neprazen metrični prostor in naj bo preslikava $f : \mathcal{M} \to \mathcal{M}$ skrčitev. Tedaj obstaja natanko ena točka $b \in \mathcal{M}$, za katero velja $f(b) = b$.*

DOKAZ. Izberimo poljubno točko $a_0 \in \mathcal{M}$ in rekurzivno definirajmo zaporedje (a_k) v $\mathcal{M}$ s predpisom $a_k = f(a_{k-1})$. Velja torej

$$a_k = f^k(a_0) = f(f(\cdots(f(a_0))\cdots)).$$

Najprej bomo pokazali, da je zaporedje (a_k) Cauchyjevo. Naj bo d metrika metričnega prostora $\mathcal{M}$. Ker je preslikava f skrčitev, obstaja takšno realno število q, $0 \leq q < 1$, da velja $d(f(x), f(y)) \leq qd(x, y)$ za vse $x, y \in \mathcal{M}$. Posebej odtod sledi

$$d(a_k, a_{k+1}) \leq qd(a_{k-1}, a_k) \leq q^2 d(a_{k-2}, a_{k-1}) \leq \cdots \leq q^k d(a_0, a_1).$$

Za poljubni naravni števili k in l, $k < l$, zato velja

$$\begin{aligned}
d(a_k, a_l) &= d(f^k(a_0), f^l(a_0)) \leq q^k d(a_0, f^{l-k}(a_0)) \\
&\leq q^k \big(d(a_0, a_1) + d(a_1, a_2) + \cdots + d(a_{l-k-1}, a_{l-k})\big) \\
&\leq q^k (1 + q + \cdots + q^{l-k-1}) d(a_0, a_1) \\
&\leq \frac{q^k}{1-q} d(a_0, a_1).
\end{aligned}$$

Odtod direktno sledi, da je zaporedje (a_k) res Cauchyjevo.

Ker je prostor $\mathcal{M}$ poln, je Cauchyjevo zaporedje (a_k) konvergentno. Naj bo $b \in \mathcal{M}$ limita tega zaporedja. Ker je preslikava f skrčitev, je tudi zvezna, zato velja

$$f(b) = f(\lim_{k\to\infty} a_k) = \lim_{k\to\infty} f(a_k) = \lim_{k\to\infty} a_{k+1} = b.$$

S tem smo dokazali obstoj negibne točke b. Enoličnost takšne točke sledi direktno z definicije skrčitve. □

4.2. Funkcije več spremenljivk

Najprej se spomnimo nekaj osnovnih lastnosti realnih funkcij več spremenljivk, ki smo jih spoznali pri Matematiki 1. Naj bo $f : \mathcal{D} \to \mathbb{R}$ skalarna funkcija n spremenljivk $x_1, x_2, \ldots, x_n$, definirana na podmnožici $\mathcal{D} \subset \mathbb{R}^n$, in naj bo a notranja točka množice $\mathcal{D}$. Za poljuben $j \in \{1, 2, \ldots, n\}$ je (prvi) parcialni odvod funkcije f na spremenljivko x_j v točki a definiran kot limita funkcije realnega parametra h

$$\frac{\partial f}{\partial x_j}(a) = \frac{\partial}{\partial x_j}(f)(a) = \lim_{h\to 0} \frac{f(a + he_j) - f(a)}{h},$$

kadar ta limita obstaja. Če vsi prvi parcialni odvodi funkcije f v točki a obstajajo, vektor

$$(\nabla f)(a) = \Big(\frac{\partial f}{\partial x_1}(a), \frac{\partial f}{\partial x_2}(a), \ldots, \frac{\partial f}{\partial x_n}(a)\Big) \in \mathbb{R}^n$$

imenujemo gradient funkcije f v točki a.

Funkcija f je totalno odvedljiva (ali diferenciabilna) v točki a, če obstaja tak vektor $v \in \mathbb{R}^n$, da velja

$$\lim_{w\to 0} \frac{f(a + w) - f(a) - v \cdot w}{|w|} = 0,$$

pri čemer je w vektorski parameter. V tem primeru je funkcija f zvezna v točki a, vsi prvi parcialni odvodi funkcije f v točki a obstajajo in velja $v = (\nabla f)(a)$.

Naj bo $f : \mathcal{D} \to \mathbb{R}$ funkcija spremenljivk $x_1, x_2, \ldots, x_n$, definirana na odprti podmnožici $\mathcal{D} \subset \mathbb{R}^n$. Funkcija f je totalno odvedljiva, če je totalno odvedljiva v vsaki točki iz $\mathcal{D}$. Funkcija f je zvezno odvedljiva, če je zvezna in če vsi prvi parcialni odvodi funkcije f obstajajo v vseh točkah iz množice $\mathcal{D}$ ter so zvezni kot funkcije

$$\frac{\partial f}{\partial x_j} : \mathcal{D} \to \mathbb{R},$$

za $j = 1, 2, \ldots, n$. V tem primeru je torej gradient

$$\nabla f : \mathcal{D} \to \mathbb{R}^n$$

zvezna funkcija. Če je funkcija f zvezno odvedljiva, potem je tudi totalno odvedljiva.

Za poljubno naravno število k in poljubne indekse $j_1, j_2, \ldots, j_k \in \{1, 2, \ldots, n\}$ k-ti parcialni odvod funkcije f na spremenljivke $x_{j_1}, x_{j_2}, \ldots, x_{j_k}$, če ta obstaja, označimo z

$$\frac{\partial^k f}{\partial x_{j_k} \cdots \partial x_{j_2} \partial x_{j_1}} = \frac{\partial}{\partial x_{j_k}} \left(\cdots \frac{\partial}{\partial x_{j_2}} \left(\frac{\partial}{\partial x_{j_1}}(f) \right) \cdots \right).$$

Funkcija f je k-krat zvezno odvedljiva, če je zvezna in če vsi p-ti parcialni odvodi funkcije f obstajajo v vseh točkah iz $\mathcal{D}$ ter so vsi zvezni kot funkcije $\mathcal{D} \to \mathbb{R}$, za vse $p = 1, 2, \ldots, n$. Množico vseh k-krat zvezno odvedljivih funkcij $\mathcal{D} \to \mathbb{R}$ označimo s $C^k(\mathcal{D})$. Funkcije v preseku $C^\infty(\mathcal{D}) = \cap_{k\in\mathbb{N}} C^k(\mathcal{D})$ so gladke funkcije.

Vektorsko funkcijo $g : \mathcal{E} \to \mathbb{R}^m$ spremenljivk $x_1, x_2, \ldots, x_n$, definirano na podmnožici $\mathcal{E} \subset \mathbb{R}^n$, zapišemo kot urejeno m-terico $g = (g_1, \ldots, g_m)$, kjer so $g_1, \ldots, g_m$ komponente funkcije g, ki so skalarne funkcije n spremenljivk. Za vse točke $(x_1, \ldots, x_n) \in \mathcal{E}$ torej velja

$$g(x_1, \ldots, x_n) = (g_1(x_1, \ldots, x_n), \ldots, g_m(x_1, \ldots, x_n)).$$

Če za nek indeks $j \in \{1, 2, \ldots, n\}$ v neki notranji točki a množice $\mathcal{E}$ obstajajo parcialni odvodi vseh komponent funkcije g na spremenljivko x_j, potem vektorju

$$\frac{\partial g}{\partial x_j}(a) = \Big(\frac{\partial g_1}{\partial x_j}(a), \frac{\partial g_2}{\partial x_j}(a), \ldots, \frac{\partial g_m}{\partial x_j}(a)\Big) \in \mathbb{R}^m$$

pravimo parcialni odvod funkcije g na spremenljivko x_j v točki a. Funkcija g je totalno odvedljiva (ali diferenciabilna) v notranji točki a množice $\mathcal{E}$, če so vse njene komponente totalno odvedljive v točki a.

Naj bo vektorska funkcija $g : \mathcal{E} \to \mathbb{R}^m$ definirana na odprti podmnožici $\mathcal{E} \subset \mathbb{R}^n$. Funkcija g je totalno odvedljiva (ali diferenciabilna), če so vse njene komponente totalno odvedljive. Funkcija g je k-krat zvezno odvedljiva (oziroma gladka), če so vse njene komponente $g_1, \ldots, g_m$ k-krat zvezno odvedljive (oziroma gladke). Množico vseh k-krat zvezno odvedljivih (oziroma gladkih) funkcij $\mathcal{E} \to \mathbb{R}^m$ označimo s $C^k(\mathcal{E}, \mathbb{R}^m)$ (oziroma s $C^\infty(\mathcal{E}, \mathbb{R}^m)$).

4.2.1. Totalni odvod vektorske funkcije več spremenljivk. Naj bo $g : \mathcal{E} \to \mathbb{R}^m$ vektorska funkcija n spremenljivk $x_1, \ldots, x_n$, definirana na podmnožici $\mathcal{E} \subset \mathbb{R}^n$, in zapišimo $g = (g_1, g_2, \ldots, g_m)$. Če je funkcija g totalno odvedljiva v notranji točki a množice $\mathcal{E}$, potem matriki parcialnih odvodov

$$(\mathrm{D}g)(a) = (\mathrm{J}g)(a) = \begin{bmatrix} \frac{\partial g_1}{\partial x_1}(a) & \frac{\partial g_1}{\partial x_2}(a) & \cdots & \frac{\partial g_1}{\partial x_n}(a) \\ \frac{\partial g_2}{\partial x_1}(a) & \frac{\partial g_2}{\partial x_2}(a) & \cdots & \frac{\partial g_2}{\partial x_n}(a) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial g_m}{\partial x_1}(a) & \frac{\partial g_m}{\partial x_2}(a) & \cdots & \frac{\partial g_m}{\partial x_n}(a) \end{bmatrix} \in \mathrm{Mat}(m \times n, \mathbb{R})$$

pravimo *totalni odvod* (ali *Jacobijeva matrika*) funkcije g v točki a.

Naj bo funkcija g totalno odvedljiva v notranji točki a množice $\mathcal{E}$. Totalni odvod $(\mathrm{D}g)(a)$ lahko zapišemo tudi bločno v oblikah:

$$(\mathrm{D}g)(a) = \begin{bmatrix} (\nabla g_1)(a)^\mathsf{t} \\ (\nabla g_2)(a)^\mathsf{t} \\ \vdots \\ (\nabla g_m)(a)^\mathsf{t} \end{bmatrix} = \begin{bmatrix} \frac{\partial g}{\partial x_1}(a) & \frac{\partial g}{\partial x_2}(a) & \cdots & \frac{\partial g}{\partial x_n}(a) \end{bmatrix}$$

Za poljuben vektor $w \in \mathbb{R}^n$ torej velja:

$$(\mathrm{D}g)(a)w = \begin{bmatrix} (\nabla g_1)(a)^\mathsf{t} w \\ (\nabla g_2)(a)^\mathsf{t} w \\ \vdots \\ (\nabla g_m)(a)^\mathsf{t} w \end{bmatrix} = \begin{bmatrix} (\nabla g_1)(a) \cdot w \\ (\nabla g_2)(a) \cdot w \\ \vdots \\ (\nabla g_m)(a) \cdot w \end{bmatrix}$$

Ker je za vsak $i = 1, 2, \ldots, m$

$$\lim_{w \to 0} \frac{g_i(a + w) - g_i(a) - (\nabla g_i)(a) \cdot w}{|w|} = 0,$$

odtod dobimo

$$\lim_{w\to 0}\frac{g(a+w)-g(a)-(\mathrm{D}g)(a)w}{|w|}=0.$$

Z zadnjo enakostjo je totalni odvod v resnici enolično določen. Še več, če je $g:\mathcal{E}\to\mathbb{R}^m$ poljubna vektorska funkcija, definirana na podmnožici $\mathcal{E}\subset\mathbb{R}^n$, če je a notranja točka množice $\mathcal{E}$ in če je A takšna realna matrika velikosti $m\times n$, da velja

$$\lim_{w\to 0}\frac{g(a+w)-g(a)-Aw}{|w|}=0,$$

potem je funkcija g totalno odvedljiva v točki a in velja $A=(\mathrm{D}g)(a)$.

Če je $g:\mathcal{E}\to\mathbb{R}^m$ totalno odvedljiva vektorska funkcija, definirana na odprti podmnožici $\mathcal{E}\subset\mathbb{R}^n$, je totalni odvod funkcija

$$\mathrm{D}g:\mathcal{E}\to\mathrm{Mat}(m\times n,\mathbb{R}).$$

Ta funkcija je zvezna, če, in samo če, je funkcija g zvezno odvedljiva.

Zgled 4.23. (1) Naj bo $f:\mathcal{D}\to\mathbb{R}$ skalarna funkcija n spremenljivk, definirana na odprti podmnožici $\mathcal{D}\subset\mathbb{R}^n$. Če je f totalno odvedljiva v točki $a\in\mathcal{D}$, potem velja

$$(\mathrm{D}f)(a)=(\nabla f)(a)^{\mathrm{t}}\in\mathrm{Mat}(1\times n,\mathbb{R}).$$

(2) Totalna odvedljivost realne funkcije ene spremenljivke je ekvivalentna običajni odvedljivosti. Vektorska funkcija ene spremenljivke je torej totalno odvedljiva, če, in samo če, so vse njene komponente odvedljive funkcije ene spremenljvke. Totalni odvod vektorske funkcije ene spremenljivke je stolpec, torej vektor.

4.2.2. Verižno pravilo. Naj bo $g:\mathcal{E}\to\mathbb{R}^m$ vektorska funkcija n spremenljivk, definirana na podmnožici $\mathcal{E}\subset\mathbb{R}^n$, in naj bo $f:\mathcal{D}\to\mathbb{R}^p$ vektorska funkcija m spremenljivk, definirana na podmnožici $\mathcal{D}\subset\mathbb{R}^m$. Predpostavimo, da je funkcija g totalno odvedljiva v notranji točki a množice $\mathcal{E}$, da je $g(a)$ notranja točka množice $\mathcal{D}$ in da je funkcija f totalno odvedljiva v točki $g(a)$.

Kompozicija funkcij g in f je tedaj vektorska funkcija n spremenljivk

$$f\circ g:g^{-1}(\mathcal{D})\to\mathbb{R}^p.$$

Ker je funkcija g zvezna, je a notranja točka množice $g^{-1}(\mathcal{D})$. Komponente kompozicije $f\circ g$ so kompozicije komponent funkcije $f=(f_1,f_2,\ldots,f_p)$ s funkcijo g, torej

$$f\circ g=(f_1\circ g,f_2\circ g,\ldots,f_p\circ g).$$

Verižno pravilo, ki smo ga dokazali pri Matematiki 1, nam pove, da so tedaj vse komponente funkcije $f\circ g$ totalno odvedljive v točki a in da za vsak $k=1,2,\ldots,p$ in za vsak $j=1,2,\ldots,n$ velja

$$\frac{\partial(f_k\circ g)}{\partial x_j}(a)=(\nabla f_k)(g(a))\cdot\frac{\partial g}{\partial x_j}(a)=(\nabla f_k)(g(a))^{\mathrm{t}}\frac{\partial g}{\partial x_j}(a).$$

Odtod sledi, da je vektorska funkcija $f\circ g$ totalno odvedljiva v točki a in da je totalni odvod funkcije $f\circ g$ v točki a enak produktu totalnega odvoda funkcije g v točki a in totalnega odvoda funkcije f v točki $g(a)$,

$$\mathrm{D}(f\circ g)(a)=(\mathrm{D}f)(g(a))(\mathrm{D}g)(a).$$

Če sta množici $\mathcal{D}$ in $\mathcal{E}$ odprti in če sta funkciji f in g zvezno odvedljivi, potem je množica $g^{-1}(\mathcal{D})$ odprta in funkcija $f\circ g:g^{-1}(\mathcal{D})\to\mathbb{R}^p$ je zvezno odvedljiva.

Podmnožica $\mathcal{E} \subset \mathbb{R}^n$ je *konveksna*, če za vsaki dve točki iz množice $\mathcal{E}$ cela daljica med tema dvema točkama leži znotraj množice $\mathcal{E}$. Primer konveksnih množic so odprte krogle. Kot enostavno posledico verižnega pravila in Lagrangeovega izreka za funkcije ene spremenljivke lahko izpeljemo naslednjo trditev, ki velja za totalno odvedljive funkcije, definirane na odprtih konveksnih podmnožicah prostora $\mathbb{R}^n$:

TRDITEV 4.24. *Naj bo $\mathcal{E}$ konveksna odprta podmnožica prostora $\mathbb{R}^n$ in naj bo $g : \mathcal{E} \to \mathbb{R}^m$ totalno odvedljiva funkcija. Tedaj za poljubni točki $u, v \in \mathcal{E}$ obstaja takšno realno število ϑ, $0 < \vartheta < 1$, da je*

$$|g(v) - g(u)| \leq \|(\mathrm{D}g)(\vartheta v + (1-\vartheta)u)\| |v - u|.$$

DOKAZ. Naj bo funkcija $\rho : \mathbb{R} \to \mathbb{R}^n$ dana s predpisom

$$\rho(t) = tv + (1-t)u.$$

Ta funkcija zaprt interval $[\![0, 1]\!]$ preslika na daljico med točkama u in v. Funkcija ρ je seveda zvezno odvedljiva in velja

$$\dot{\rho}(t) = (\mathrm{D}\rho)(t) = v - u$$

za vsak $t \in \mathbb{R}$. Naj bo $\phi : \rho^{-1}(\mathcal{E}) \to \mathbb{R}^m$ kompozicija funkcij ρ in g, torej

$$\phi(t) = g(\rho(t))$$

za $t \in \rho^{-1}(\mathcal{E})$. Po verižnem pravilu je funkcija ϕ totalno odvedljiva in velja

$$(\mathrm{D}\phi)(t) = (\mathrm{D}g)(\rho(t))\dot{\rho}(t) = (\mathrm{D}g)(tv + (1-t)u)(v - u).$$

Posebej je torej

$$|(\mathrm{D}\phi)(t)| \leq \|(\mathrm{D}g)(tv + (1-t)u)\| |v - u|.$$

Definirajmo še funkcijo $f : \rho^{-1}(\mathcal{E}) \to \mathbb{R}$ s predpisom

$$f(t) = \phi(t) \cdot (g(v) - g(u)).$$

To je torej odvedljiva realna funkcija ene realne spremenljivke. Na intervalu $[\![0, 1]\!] \subset g^{-1}(\mathcal{E})$ zanjo lahko uporabimo Lagrangeov izrek, ki pove, da obstaja takšno realno število ϑ, $0 < \vartheta < 1$, da je

$$f(1) - f(0) = \dot{f}(\vartheta) = (\mathrm{D}f)(\vartheta).$$

Ker je

$$\dot{f}(\vartheta) = (\mathrm{D}\phi)(\vartheta) \cdot (g(v) - g(u)),$$

iz prejšnje enakosti dobimo

$$\phi(1) \cdot (g(v) - g(u)) - \phi(0) \cdot (g(v) - g(u)) = (\mathrm{D}\phi)(\vartheta) \cdot (g(v) - g(u))$$

in odtod

$$\begin{aligned} |g(v) - g(u)|^2 &= g(v) \cdot (g(v) - g(u)) - g(u) \cdot (g(v) - g(u)) \\ &= (\mathrm{D}\phi)(\vartheta) \cdot (g(v) - g(u)) \\ &\leq |(\mathrm{D}\phi)(\vartheta)| |g(v) - g(u)| \\ &\leq \|(\mathrm{D}g)(\vartheta v + (1-\vartheta)u)\| |v - u| |g(v) - g(u)|. \end{aligned}$$

□

4.2.3. Izrek o inverzni funkciji.

Izrek 4.25 (Izrek o inverzni funkciji). *Naj bo $f : \mathcal{D} \to \mathbb{R}^n$ zvezno odvedljiva funkcija n spremenljivk, definirana na odprti podmnožici $\mathcal{D} \subset \mathbb{R}^n$, in na bo $a \in \mathcal{D}$ takšna točka, da velja $\det(\mathrm{D}f)(a) \neq 0$. Tedaj obstaja takšna odprta podmnožica $\mathcal{U} \subset \mathbb{R}^n$, da je $a \in \mathcal{U} \subset \mathcal{D}$ in da velja:*

(i) Zožitev $f|_{\mathcal{U}}$ je injektivna funkcija in podmnožica $f(\mathcal{U})$ je odprta v $\mathbb{R}^n$.

(ii) Inverz $(f|_{\mathcal{U}})^{-1} : f(\mathcal{U}) \to \mathcal{U}$ bijekcije $f|_{\mathcal{U}} : \mathcal{U} \to f(\mathcal{U})$ je zvezno odvedljiva funkcija in velja

$$\mathrm{D}((f|_{\mathcal{U}})^{-1})(f(a)) = ((\mathrm{D}f)(a))^{-1}.$$

Dokaz. Označimo $A = (\mathrm{D}f)(a)$ in $\lambda = 1/(2\|A^{-1}\|)$. Ker je f zvezno odvedljiva, je funkcija $\mathrm{D}f : \mathcal{D} \to \mathrm{Mat}(n \times n, \mathbb{R})$ zvezna, zato obstaja tako majhen $\delta > 0$, da je $\mathcal{K}(a, \delta) \subset \mathcal{D}$ in da za vsak $u \in \mathcal{K}(a, \delta)$ velja

$$\|(\mathrm{D}f)(u) - A\| < \lambda. \tag{4.1}$$

Ker je torej $\|(\mathrm{D}f)(u) - A\| < 1/(2\|A^{-1}\|) < 1/\|A^{-1}\|$, je po trditvi 4.4 matrika $(\mathrm{D}f)(u)$ obrnljiva. Inverz te matrike označimo z $B_u = ((\mathrm{D}f)(u))^{-1}$. Za množico $\mathcal{U}$ iz izreka vzamemo odprto kroglo $\mathcal{K}(a, \delta)$, torej

$$\mathcal{U} = \mathcal{K}(a, \delta).$$

Za poljuben $v \in \mathbb{R}^n$ definirajmo funkcijo $\phi_v : \mathcal{U} \to \mathbb{R}^n$ s predpisom

$$\phi_v(u) = u + A^{-1}(v - f(u)),$$

za vsak $u \in \mathcal{U}$. Opazimo, da je $f(u) = v$, če, in samo če, velja $\phi_v(u) = u$. Poleg tega je

$$(\mathrm{D}\phi_v)(u) = \mathrm{I} - A^{-1}(\mathrm{D}f)(u) = A^{-1}(A - (\mathrm{D}f)(u)),$$

odtod in iz neenakosti (4.1) pa sledi

$$\|(\mathrm{D}\phi_v)(u)\| \leq \|A^{-1}\|\|A - (\mathrm{D}f)(u)\| \leq \|A^{-1}\|\lambda = \frac{1}{2}.$$

Po trditvi 4.24 za poljubna vektorja $u, u_0 \in \mathcal{U}$ zato velja

$$|\phi_v(u) - \phi_v(u_0)| \leq \frac{1}{2}|u - u_0|. \tag{4.2}$$

Preslikava ϕ_v je torej skrčitev in ima zato največ eno negibno točko, s tem pa smo dokazali, da je zožitev $f|_{\mathcal{U}}$ injektivna.

Dokažimo, da je množica $\mathcal{V} = f(\mathcal{U})$ odprta v $\mathbb{R}^n$. Vzemimo poljubno točko $v_0 \in \mathcal{V}$. Naj bo u_0 tista točka iz $\mathcal{U}$, za katero je $f(u_0) = v_0$ in naj bo $\mu > 0$ tako majhen, da je $\mathcal{K}(u_0, 2\mu) \subset \mathcal{U}$. Posebej torej velja, da je tudi zaprta krogla

$$\mathcal{B} = \{u \in \mathbb{R}^n \ ; \ |u - u_0| \leq \mu\}$$

podmnožica množice $\mathcal{U}$. Pokazali bomo, da tedaj velja

$$\mathcal{K}(v_0, \lambda\mu) \subset \mathcal{V},$$

odtod pa sledi, da je množica $\mathcal{V}$ odprta. Res, vzemimo poljuben vektor $v \in \mathcal{K}(v_0, \lambda\mu)$. Opazimo, da velja

$$\begin{aligned}|\phi_v(u_0) - u_0| &= |A^{-1}(v - f(u_0))| = |A^{-1}(v - v_0)| \\ &\leq \|A^{-1}\||v - v_0| < \|A^{-1}\|\lambda\mu = \frac{\mu}{2}.\end{aligned}$$

Za vsak $u \in \mathcal{B}$ je torej

$$|\phi_v(u) - u_0| \le |\phi_v(u) - \phi_v(u_0)| + |\phi_v(u_0) - u_0| < \frac{1}{2}|u - u_0| + \frac{\mu}{2} \le \mu.$$

Pokazali smo torej, da je $\phi_v(\mathcal{B}) \subset \mathcal{B}$. Ker je zaprta krogla $\mathcal{B}$ kompaktna, je posebej tudi poln metrični prostor. Za preslikavo $\phi_v|_{\mathcal{B}} : \mathcal{B} \to \mathcal{B}$, ki je kontrakcija (neenakost (4.2)), lahko torej uporabimo Banachov izrek o negibni točki 4.22. Po tem izreku obstaja natanko ena točka $u \in \mathcal{B}$, za katero je $\phi_v(u) = u$, odtod pa sledi $v = f(u) \in f(\mathcal{B}) \subset f(\mathcal{U}) = \mathcal{V}$. S tem smo končali dokaz točke (i) iz izreka.

Dokažimo zdaj še točko (ii). Inverz zožitve $f|_{\mathcal{U}} : \mathcal{U} \to \mathcal{V}$ označimo z $g : \mathcal{V} \to \mathcal{U}$. Najprej bomo dokazali, da je funkcija g totalno odvedljiva v poljubni točki $v \in \mathcal{V}$. Naj bo $u = g(v)$, naj bo $w \in \mathbb{R}^n$ tak vektor, da je $v + w \in \mathcal{V}$, in označimo $r = r(w) = g(v + w) - u$. Opazimo lahko, da je

$$\phi_v(u + r) - \phi_v(u) = r - A^{-1}(f(u + r) - f(u)) = r - A^{-1}w.$$

Ker iz neenakosti (4.2) sledi $|\phi_v(u + r) - \phi_v(u)| \le \frac{1}{2}|r|$, je torej

$$|r - A^{-1}w| \le \frac{1}{2}|r|.$$

Odtod sledi

$$|r| \le |r - A^{-1}w| + |A^{-1}w| \le \frac{1}{2}|r| + |A^{-1}w|,$$

torej

$$\frac{1}{2}|r| \le |A^{-1}w| \le \|A^{-1}\||w|$$

in zato

$$|r| \le 2\|A^{-1}\||w| = \frac{1}{\lambda}|w|. \tag{4.3}$$

Iz enakosti

$$\begin{aligned} g(v + w) - g(v) - B_u w &= (u + r) - u - B_u w = B_u(B_u^{-1} r - w) \\ &= B_u(B_u^{-1} r - (v + w) + v) \\ &= B_u((\mathrm{D}f)(u)r - f(u + r) + f(u)) \end{aligned}$$

in iz neenakosti (4.3) sledi

$$\frac{|g(v + w) - g(v) - B_u w|}{|w|} \le \frac{\|B_u\||f(u + r) - f(u) - (\mathrm{D}f)(u)r|}{\lambda|r|}.$$

Če se w približuje 0, se po neenakosti (4.3) tudi r približuje 0, zaradi totalne odvedljivosti funkcije f v točki u pa se tedaj tudi izraz $|f(u+r)-f(u)-(\mathrm{D}f)(u)r|/|r|$ približuje 0. Iz zadnje neenakost zato sledi, da je

$$\lim_{w \to 0} \frac{|g(v + w) - g(v) - B_u w|}{|w|} = 0,$$

torej je funkcija g totalno odvedljiva v točki v. Ob tem smo še dokazali, da je $(\mathrm{D}g)(v) = B_u$, torej

$$(\mathrm{D}g)(v) = ((\mathrm{D}f)(g(v)))^{-1}$$

za vsak $v \in \mathcal{V}$. Ker komponente matrike matrike $((\mathrm{D}f)(g(v)))^{-1}$ lahko izračunamo kot kofaktorje matrike $(\mathrm{D}f)(g(v))$, ki jih delimo z determinanto te matrike, in ker so komponente matrike $(\mathrm{D}f)(g(v))$ zvezne funkcije spremenljivke v, so tudi komponente matrike $(\mathrm{D}g)(v)$ zvezne funkcije spremenljivke v. Funkcija g je torej zvezno odvedljiva. □

4.2.4. Izrek o implicitni funkciji.

IZREK 4.26 (Izrek o implicitni funkciji). *Naj bo $g : \mathcal{D} \to \mathbb{R}^m$ zvezno odvedljiva funkcija, definirana na odprti podmnožici $\mathcal{D} \subset \mathbb{R}^{n+m}$, in naj bosta $a \in \mathbb{R}^n$ ter $b \in \mathbb{R}^m$ takšni točki, da je $(a,b) \in \mathcal{D}$, $g(a,b) = 0$ in da so vektorji*

$$\frac{\partial g}{\partial x_{n+1}}(a,b),\ \frac{\partial g}{\partial x_{n+2}}(a,b),\ \ldots,\ \frac{\partial g}{\partial x_{n+m}}(a,b)$$

med seboj linearno neodvisni. Tedaj obstajata takšni odprti podmnožici $\mathcal{U} \subset \mathcal{D}$ in $\mathcal{W} \subset \mathbb{R}^n$, da je $(a,b) \in \mathcal{U}$ ter $a \in \mathcal{W}$ in da velja:

(i) Za vsako točko $x \in \mathcal{W}$ obstaja natanko ena točka $y \in \mathbb{R}^m$, za katero je $(x,y) \in \mathcal{U}$ in

$$g(x,y) = 0.$$

To točko y, ki je odvisna od izbrane točke x, označimo z $f(x)$.

(ii) Tako definirana funkcija $f : \mathcal{W} \to \mathbb{R}^m$ je zvezno odvedljiva, njen graf je množica vseh ničel funkcije $g|_{\mathcal{U}}$, ob tem pa je $f(a) = b$ in

$$(\mathrm{D}f)(a) = -\begin{bmatrix} \frac{\partial g}{\partial x_{n+1}}(a,b) & \cdots & \frac{\partial g}{\partial x_{n+m}}(a,b) \end{bmatrix}^{-1} \begin{bmatrix} \frac{\partial g}{\partial x_1}(a,b) & \cdots & \frac{\partial g}{\partial x_n}(a,b) \end{bmatrix}.$$

DOKAZ. Označimo

$$X = \begin{bmatrix} \frac{\partial g}{\partial x_1}(a,b) & \cdots & \frac{\partial g}{\partial x_n}(a,b) \end{bmatrix} \in \mathrm{Mat}(m \times n, \mathbb{R})$$

in

$$Y = \begin{bmatrix} \frac{\partial g}{\partial x_{n+1}}(a,b) & \cdots & \frac{\partial g}{\partial x_{n+m}}(a,b) \end{bmatrix} \in \mathrm{Mat}(m \times m, \mathbb{R}).$$

Ker so stolpci matrike Y po predpostavki linearno neodvisni, je matrika Y obrnljiva. Totalni odvod funkcije g v točki (a,b) lahko zapišemo v bločni obliki

$$(\mathrm{D}g)(a,b) = \begin{bmatrix} X & Y \end{bmatrix}.$$

Za vse točke $x \in \mathbb{R}^n$ in $y \in \mathbb{R}^m$, za katere je $(x,y) \in \mathcal{D} \subset \mathbb{R}^{n+m} = \mathbb{R}^n \times \mathbb{R}^m$, naj bo

$$G(x,y) = (x, g(x,y)) \in \mathbb{R}^n \times \mathbb{R}^m = \mathbb{R}^{n+m}.$$

S tem smo torej definirali funkcijo $G : \mathcal{D} \to \mathbb{R}^{n+m}$, ki je očitno zvezno odvedljiva, saj je funkcija g zvezno odvedljiva. Ker je totalni odvod funkcije G v točki (a,b) v bločni obliki enak

$$(\mathrm{D}G)(a,b) = \begin{bmatrix} \mathrm{I} & 0 \\ X & Y \end{bmatrix},$$

je $\det(\mathrm{D}G)(a,b) = \det Y \neq 0$. Funkcija G torej zadošča predpostavkam izreka o inverzni funkciji 4.25. Ta izrek nam pove, da obstaja takšna odprta okolica $\mathcal{U} \subset \mathcal{D} \subset \mathbb{R}^{n+m}$ točke (a,b), da je množica $\mathcal{V} = G(\mathcal{U}) \subset \mathbb{R}^{n+m}$ odprta in da je zožitev $G|_{\mathcal{U}} : \mathcal{U} \to \mathcal{V}$ bijekcija z zvezno odvedljivim inverzom $F = (G|_{\mathcal{U}})^{-1} : \mathcal{V} \to \mathcal{U}$.

Naj bo

$$\mathcal{W} = \{x \in \mathbb{R}^n \; ; \; (x,0) \in \mathcal{V}\}.$$

Ker je $(a,b) \in \mathcal{U}$, velja $(a,0) = (a, g(a,b)) = G(a,b) \in \mathcal{V}$, zato je $a \in \mathcal{W}$. Poleg tega iz enakosti $\mathcal{W} \times \{0\}^m = \mathcal{V} \cap (\mathbb{R}^n \times \{0\}^m)$ sledi, da je $\mathcal{W}$ odprta podmnožica prostora $\mathbb{R}^n$. Naj bo $x \in \mathcal{W}$. Za točko $F(x,0) \in \mathcal{U}$ velja $G(F(x,0)) = (x,0)$, zato je

$$F(x,0) = (x,y) \in \mathcal{U}$$

za nek $y \in \mathbb{R}^m$, ob tem pa je $g(x,y) = 0$. Če je $z \in \mathbb{R}^m$ še ena točka, za katero je $(x,z) \in \mathcal{U}$ in $g(x,z) = 0$, velja $G(x,z) = (x,0) = G(x,y)$ in zato $z = y$, saj je G injektivna na $\mathcal{U}$. S tem smo dokazali točko (i) iz izreka.

Dokažimo še točko (ii). Seveda je $f(a) = b$, za vsako ničlo (x, y) funkcije $g|_{\mathcal{U}}$ pa sledi $x \in \mathcal{W}$ in $y = f(x)$. Opazimo še, da je funkcija f enaka kompoziciji funkcij

$$\mathcal{W} \xrightarrow{S} \mathcal{V} \xrightarrow{F} \mathbb{R}^{n+m} \xrightarrow{T} \mathbb{R}^m,$$

kjer je $S(x) = (x, 0)$ in $T(x, y) = y$. Ker so funkcije S, F in T vse zvezno odvedljive, je takšna tudi funkcija f.

Ostane nam še, da izračunamo totalni odvod funkcije f v točki a. Totalni odvod funkcije $\phi = F \circ S : \mathcal{W} \to \mathbb{R}^{n+m}$, $\phi(x) = (x, f(x))$, je v točki a v bločni obliki enak

$$(\mathrm{D}\phi)(a) = \begin{bmatrix} \mathrm{I} \\ (\mathrm{D}f)(a) \end{bmatrix}.$$

Ker je $g(x, f(x)) = 0$ za vsak $x \in \mathcal{W}$, je kompozicija $g \circ \phi : \mathcal{W} \to \mathbb{R}^m$ ničelna funkcija in zato je tudi njen totalni odvod v točki a enak 0, torej

$$(\mathrm{D}(g \circ \phi))(a) = (\mathrm{D}g)(\phi(a))(\mathrm{D}\phi)(a) = 0.$$

Če zadnjo enakost napišemo v bločni obliki, dobimo

$$\begin{bmatrix} X & Y \end{bmatrix} \begin{bmatrix} \mathrm{I} \\ (\mathrm{D}f)(a) \end{bmatrix} = 0$$

oziroma

$$X + Y(\mathrm{D}f)(a) = 0,$$

odtod pa dobimo $(\mathrm{D}f)(a) = -Y^{-1}X$. $\square$

V primeru $n = m = 1$ nam izrek o implicitni funkciji pove, da lahko iz enačbe

$$g(x, y) = 0$$

lokalno, v okolici neke ničle (a, b) funkcije g, izrazimo spremenljivko y kot zvezno odvedljivo funkcijo spremenljivke x, če je le izpolnjen pogoj $\frac{\partial g}{\partial y}(a, b) \neq 0$. Natančneje povedano, rešitve enačbe $g(x, y) = 0$ v tem primeru lokalno zapišemo kot graf zvezno odvedljive funkcije spremenljivke x. Če je $\frac{\partial g}{\partial x}(a, b) \neq 0$, lahko podobno izrazimo spremenljivko x kot funkcijo spremenljivke y.

V splošnem sta x in y vektorja in izrek o implicitni funkciji govori o tem, kdaj lahko rešitve enačbe $g(x, y) = 0$ lokalno zapišemo kot graf zvezno odvedljive vektorske funkcije vektorja spremenljivk x. Če bi v izreku o implicitni funkciji predpostavko linearne neodvisnosti vektorjev

$$\frac{\partial g}{\partial x_{n+1}}(a, b),\ \frac{\partial g}{\partial x_{n+2}}(a, b),\ \ldots,\ \frac{\partial g}{\partial x_{n+m}}(a, b)$$

zamenjali z linearno neodvisnostjo gradientov

$$(\nabla g_1)(a, b),\ (\nabla g_2)(a, b),\ \ldots,\ (\nabla g_m)(a, b),$$

oziroma s predpostavko, da je $\operatorname{rank}(\mathrm{D}g)(a, b) = m$, bi izrek veljal v zapisani obliki, če bi najprej ustrezno spremenili vrstni red spremenljivk in s tem zagotovili, da je zadnjih m stolpcev v totalnem odvodu funkcije g linearno neodvisnih. Drugače povedano, od tega, katera $(m \times m)$-poddeterminanta totalnega odvoda funkcije g je v dani ničli funkcije g neničelna je odvisno, katerih m spremenljivk lahko izrazimo kot funkcijo preostalih n spremenljivk.

VAJA 4.27. (1) Ugotovi, kaj pravi izrek 4.26 v primeru, ko je $g : \mathbb{R}^{n+m} \to \mathbb{R}^m$ linearna preslikava.

(2) Ugotovi, kaj pove izrek 4.26 v primeru funkcije $g : \mathbb{R}^2 \to \mathbb{R}$, podane s predpisom $g(x, y) = x^2 + y^2 - 1$.

4.2.5. Ekstremi funkcij več spremenljivk. Naj bo $f : \mathcal{D} \to \mathbb{R}$ totalno odvedljiva funkcija n spremenljivk, definirana na odprti podmnožici $\mathcal{D} \subset \mathbb{R}^n$. Točka $a \in \mathcal{D}$ je stacionarna točka funkcije f, če velja

$$(\nabla f)(a) = 0.$$

Če ima funkcija f v točki $a \in \mathcal{D}$ lokalni ekstrem, potem je a stacionarna točka funkcije f.

Da bi formulirali zadostni pogoj za nastop lokalnega ekstrema, si moramo ogledati tudi druge parcialne odvode. Če je funkcija f dvakrat zvezno odvedljiva, njene druge parcialne odvode zložimo v matriko

$$(\mathrm{H}f)(a) = \left[\tfrac{\partial^2 f}{\partial x_i \partial x_j}(a)\right]_{i,j=1,\ldots,n} \in \mathrm{Mat}(n \times n, \mathbb{R}),$$

ki je simetrična in jo imenujemo *Hessejeva matrika* funkcije f v točki a.

TRDITEV 4.28. *Naj bo $f : \mathcal{D} \to \mathbb{R}$ dvakrat zvezno odvedljiva funkcija n spremenljivk, definirana na odprti podmnožici $\mathcal{D} \subset \mathbb{R}^n$, in naj bo $a \in \mathcal{D}$ stacionarna točka funkcije f. Velja:*

(i) Če je Hessejeva matrika $(\mathrm{H}f)(a)$ pozitivno definitna, potem ima funkcija f v točki a strogi lokalni minimum.

(ii) Če je Hessejeva matrika $(\mathrm{H}f)(a)$ negativno definitna, potem ima funkcija f v točki a strogi lokalni maksimum.

(iii) Če Hessejeva matrika $(\mathrm{H}f)(a)$ ni niti pozitivno semidefinitna niti negativno semidefinitna, potem funkcija f v točki a nima lokalnega ekstrema.

DOKAZ. Naj bo $\delta > 0$ tako majhen, da velja $\mathcal{K}(a, \delta) \subset \mathcal{D}$. Po Taylorjevi formuli za vsak $w = (\gamma_1, \gamma_2, \ldots, \gamma_n) \in \mathcal{K}(0, \delta)$ obstaja tako realno število ϑ_w, $0 < \vartheta_w < 1$, da je

$$\begin{aligned} f(a + w) &= f(a) + \frac{1}{2} \sum_{1 \le i,j \le n} \frac{\partial^2 f}{\partial x_i \partial x_j}(a + \vartheta_w w)\gamma_i \gamma_j \\ &= f(a) + \frac{1}{2} w^{\mathsf{t}} (\mathrm{H}f)(a + \vartheta_w w) w. \end{aligned}$$

Za to, ali ima funkcija f v točki a lokalni ekstrem, je tako odločilen predznak izraza $w^{\mathsf{t}}(\mathrm{H}f)(a + \vartheta_w w)w$. Opazimo lahko, da je preslikava $\mathrm{H}f : \mathcal{D} \to \mathrm{Mat}(n \times n, \mathbb{R})$, ki preslika poljuben vektor $u \in \mathcal{D}$ v simetrično matriko $(\mathrm{H}f)(u)$, zvezna, saj je funkcija f dvakrat zvezno odvedljiva.

(i) Po predpostavki je matrika $(\mathrm{H}f)(a)$ pozitivno definitna. Dovolj je dokazati, da obstaja tako majhen $\epsilon > 0$, da je $\epsilon < \delta$ in da je matrika $(\mathrm{H}f)(u)$ pozitivno definitna za vsak $u \in \mathcal{K}(a, \epsilon)$. Če to ne bi bilo res, potem bi za vsako naravno število k lahko našli takšna $u_k \in \mathcal{K}(a, \delta) \cap \mathcal{K}(a, 1/k)$ in $v_k \in \mathbb{R}^n \setminus \{0\}$, da bi veljalo

$$v_k^{\mathsf{t}} (\mathrm{H}f)(u_k) v_k \le 0.$$

Tudi za normirane vektorje $w_k = v_k / |v_k|$ tedaj velja

$$w_k^{\mathsf{t}} (\mathrm{H}f)(u_k) w_k \le 0.$$

Zaporedje (u_k) konvergira k točki a. Zaporedje (w_k) je zaporedje točk iz sfere $\mathbb{S}^{n-1}$, ki je kompaktna, zato lahko izberemo njegovo podzaporedje $(w_{k_l})_l$, ki konvergira k neki točki $w \in \mathbb{S}^{n-1}$. Zaradi zveznosti tedaj velja

$$0 \geq \lim_{l \to \infty} w_{k_l}^{\mathsf{t}} (\mathrm{H}f)(u_{k_l}) w_{k_l} = w^{\mathsf{t}} (\mathrm{H}f)(a) w,$$

kar je protislovje s predpostavko, da je matrika $(\mathrm{H}f)(a)$ pozitivno definitna.

(ii) Dokaz te točke je podoben dokazu točke (i).

(iii) Po predpostavki obstajata takšna vektorja $v, w \in \mathbb{R}^n$, da velja

$$v^{\mathsf{t}}(\mathrm{H}f)(a)v > 0 \quad \text{in} \quad w^{\mathsf{t}}(\mathrm{H}f)(a)w < 0.$$

Zaradi zveznosti preslikave $\mathrm{H}f$ obstaja tako majhen $\epsilon > 0$, da je $\epsilon < \delta$ in je da za vsak $u \in \mathcal{K}(a, \epsilon)$

$$v^{\mathsf{t}}(\mathrm{H}f)(u)v > 0 \quad \text{in} \quad w^{\mathsf{t}}(\mathrm{H}f)(u)w < 0.$$

Za vsak $\mu > 0$, $\mu < \epsilon$, lahko najdemo tako majhen $\lambda > 0$, da velja $\lambda v, \lambda w \in \mathcal{K}(0, \mu)$, za ta dva vektorja pa velja

$$(\lambda v)^{\mathsf{t}}(\mathrm{H}f)(a + \vartheta_{\lambda v} \lambda v)(\lambda v) > 0 \quad \text{in} \quad (\lambda w)^{\mathsf{t}}(\mathrm{H}f)(a + \vartheta_{\lambda v} \lambda w)(\lambda w) < 0,$$

kar pomeni, da funkcija f v točki a nima lokalnega ekstrema. □

4.2.6. Vezani ekstremi in Lagrangeova metoda multiplikatorjev. Naj bo $f : \mathcal{D} \to \mathbb{R}$ skalarna funkcija n spremenljivk, definirana na odprti podmnožici $\mathcal{D} \subset \mathbb{R}^n$. Naj bo $\mathcal{R} \subset \mathbb{R}^n$ poljubna podmnožica in naj bo $a \in \mathcal{R} \cap \mathcal{D}$. Če ima zožitev $f|_{\mathcal{R} \cap \mathcal{D}}$ lokalni ekstrem v točki a, potem pravimo, da ima funkcija f v točki a *vezan ekstrem na podmnožici* $\mathcal{R}$.

Običajno je podmnožica $\mathcal{R}$ podana kot množica ničel neke funkcije. Naj bo $g : \mathcal{E} \to \mathbb{R}^m$ vektorska funkcija n spremenljivk, definirana na odprti podmnožici $\mathcal{E} \subset \mathbb{R}^n$, in za množico $\mathcal{R}$ vzemimo množico ničel te funkcije, torej

$$\mathcal{R} = g^{-1}(\{0\}) = \{v \in \mathcal{E} \; ; \; g(v) = 0\} \subset \mathcal{E} \subset \mathbb{R}^n.$$

Če ima zožitev $f|_{\mathcal{R} \cap \mathcal{D}} = f|_{g^{-1}(\{0\}) \cap \mathcal{D}}$ lokalni ekstrem v točki $a \in g^{-1}(\{0\}) \cap \mathcal{D}$, potem pravimo, da ima funkcija f v točki *a vezan ekstrem pri vezi g*.

Vektorska enačba $g(v) = 0$ je ekvivalentna sistemu m skalarnih enačb

$$g_1(v) = 0, \; g_2(v) = 0, \; \ldots, \; g_m(v) = 0,$$

kjer so $g_1, g_2, \ldots, g_m$ skalarne komponente vektorske vezi g, ki jim pravimo tudi *skalarne vezi.* Vezanim ekstremom funkcije f pri (vektorski) vezi g bi lahko rekli tudi *vezani ekstremi pri m skalarnih vezeh $g_1, g_2, \ldots, g_m$*.

Trditev 4.29 (Lagrangeova metoda multiplikatorjev). *Naj bo $f : \mathcal{D} \to \mathbb{R}$ totalno odvedljiva skalarna funkcija n spremenljivk, definirana na odprti podmnožici $\mathcal{D} \subset \mathbb{R}^n$, in naj bo $g : \mathcal{E} \to \mathbb{R}^m$ zvezno odvedljiva vektorska funkcija n spremenljivk, definirana na odprti podmnožici $\mathcal{E} \subset \mathbb{R}^n$. Naj bo $a \in \mathcal{E} \cap \mathcal{D}$ takšna točka, da je $g(a) = 0$ in da je $\operatorname{rank}(\mathrm{D}g)(a) = m$. Če ima funkcija f v točki a vezan ekstrem pri vezi g, potem obstajajo takšne realne konstante $\lambda_1, \lambda_2, \ldots, \lambda_m$, da velja*

$$(\nabla f)(a) = \sum_{i=1}^{m} \lambda_i (\nabla g_i)(a).$$

Dokaz. Ker je $\operatorname{rank}(\mathrm{D}g)(a) = m$, je $m \leq n$. Če je $m = n$, je trditev trivialna, saj gradienti vezi tedaj sestavljajo bazo prostora $\mathbb{R}^n$. Predpostavimo torej lahko, da je $m < n$, in označimo $p = n - m$.

Totalni odvod $(\mathrm{D}g)(a)$ ima neničelno $(m \times m)$-poddeterminanto. Če je potrebno, lahko spremenimo vrstni red spremenljivk tako, da je $(\mathrm{D}g)(a)$ bločno oblike

$$\begin{bmatrix} X & Y \end{bmatrix},$$

kjer je $X \in \mathrm{Mat}(m \times p, \mathbb{R})$ in $Y \in \mathrm{GL}(m, \mathbb{R})$. Zapišimo še $a = (b, c) \in \mathbb{R}^n = \mathbb{R}^p \times \mathbb{R}^m$, ker je $b \in \mathbb{R}^p$ in $c \in \mathbb{R}^m$. Za funkcijo $g|_{\mathcal{D} \cap \mathcal{E}}$ torej lahko uporabimo izrek o implicitni funkciji 4.26. Ta nam pove, da obstajata takšni odprti podmnožici $\mathcal{U} \subset \mathcal{D} \cap \mathcal{E}$ in $\mathcal{W} \subset \mathbb{R}^p$ ter takšna zvezno odvedljiva funkcija $\phi : \mathcal{W} \to \mathbb{R}^m$, da velja $(b, c) \in \mathcal{U}$, $b \in \mathcal{W}$ in $\phi(b) = c$ ter da je graf funkcije ϕ enak množici tistih ničel funkcije g, ki ležijo v množici $\mathcal{U}$.

Naj bo $\rho : \mathcal{W} \to \mathbb{R}^n$ zvezno odvedljiva funkcija, dana s predpisom

$$\rho(x) = (x, \phi(x)) \in \mathbb{R}^p \times \mathbb{R}^m = \mathbb{R}^n.$$

Slika $\rho(\mathcal{W})$ je torej graf funkcije ϕ, ki je enak množici $g^{-1}(\{0\}) \cap \mathcal{U}$, zato iz predpostavke, da ima funkcija f v točki a vezan ekstrem pri vezi g, sledi, da ima funkcija

$$f \circ \rho : \mathcal{W} \to \mathbb{R}$$

v točki b lokalni ekstrem. Točka b je torej stacionarna točka zvezno odvedljive funkcije $f \circ \rho$, zato velja $(\mathrm{D}(f \circ \rho))(b) = 0$, po verižnem pravilu pa odtod sledi

$$(\mathrm{D}f)(a)(\mathrm{D}\rho)(b) = 0. \tag{4.4}$$

Po izreku o implicitni funkciji dobimo bločni zapis za totalni odvod $(\mathrm{D}\rho)(b)$,

$$(\mathrm{D}\rho)(b) = \begin{bmatrix} \mathrm{I} \\ -Y^{-1}X \end{bmatrix},$$

totalni odvod $(\mathrm{D}f)(a)$ pa bločno zapišimo v obliki

$$(\mathrm{D}f)(a) = \begin{bmatrix} f_x & f_y \end{bmatrix},$$

kjer je

$$f_x = \begin{bmatrix} \frac{\partial f}{\partial x_1}(a) & \cdots & \frac{\partial f}{\partial x_p}(a) \end{bmatrix} \quad \text{in} \quad f_y = \begin{bmatrix} \frac{\partial f}{\partial x_{p+1}}(a) & \cdots & \frac{\partial f}{\partial x_n}(a) \end{bmatrix}.$$

Enakost (4.4) lahko torej zapišemo v bločni obliki

$$\begin{bmatrix} f_x & f_y \end{bmatrix} \begin{bmatrix} \mathrm{I} \\ -Y^{-1}X \end{bmatrix} = 0,$$

odtod pa dobimo

$$f_x - f_y Y^{-1} X = 0.$$

Naj bo

$$\Lambda = \begin{bmatrix} \lambda_1 & \lambda_2 & \cdots & \lambda_m \end{bmatrix} = f_y Y^{-1}.$$

Velja torej

$$f_x = \Lambda X \quad \text{in} \quad f_y = \Lambda Y$$

oziroma

$$(\mathrm{D}f)(a) = \Lambda(\mathrm{D}g)(a).$$

S transponiranjem zadnjo enakost prepišemo v obliko

$$\begin{aligned}(\nabla f)(a) &= ((\mathrm{D}f)(a))^{\mathrm{t}} = ((\mathrm{D}g)(a))^{\mathrm{t}}\Lambda^{\mathrm{t}} \\ &= \begin{bmatrix}(\nabla g_1)(a) & \cdots & (\nabla g_m)(a)\end{bmatrix}\begin{bmatrix}\lambda_1 \\ \vdots \\ \lambda_m\end{bmatrix} \\ &= \sum_{i=1}^{m} \lambda_i (\nabla g_i)(a).\end{aligned}$$

□

Komentar 4.30. Lagrangeova metoda multiplikatorjev nam pomaga pri iskanju vezanih ekstremov, saj nam poda potrebni pogoj za nastop vezanega ekstrema.

Posebej si oglejmo, kaj Lagrangeova metoda multiplikatorjev pove v primeru, ko imamo le eno skalarno vez, torej v primeru $m = 1$. Predpostavimo torej, da je $f : \mathcal{D} \to \mathbb{R}$ totalno odvedljiva skalarna funkcija, definirana na odprti podmnožici $\mathcal{D} \subset \mathbb{R}^n$, in da je $g : \mathcal{E} \to \mathbb{R}$ zvezno odvedljiva skalarna funkcija, definirana na odprti podmnožici $\mathcal{E} \subset \mathbb{R}^n$. Če je tedaj $a \in \mathcal{E} \cap \mathcal{D}$ takšna točka, da je $g(a) = 0$, da je $(\nabla g)(a) \neq 0$ in da ima funkcija f v točki a vezan ekstrem pri vezi g, potem obstaja realna konstanta λ, za katero velja

$$(\nabla f)(a) = \lambda(\nabla g)(a).$$

Zgled 4.31. Poiščimo vezane ekstreme funkcije dveh spremenljivk $f(x, y) = y^2 - x^2$ na krožnici $x^2 + y^2 = 1$. Krožnica

$$\mathcal{R} = \mathbb{S}^1 = \{(x, y) \in \mathbb{R}^2 \; ; \; x^2 + y^2 = 1\}$$

je množica ničel funkcije $g : \mathbb{R}^2 \to \mathbb{R}$,

$$g(x, y) = x^2 + y^2 - 1,$$

zato funkcijo g vzamemo za vez. Funkcija g je zvezno odvedljiva, njen gradient

$$(\nabla g)(x, y) = (2x, 2y)$$

pa je neničelen v vseh točkah iz krožnice $\mathbb{S}^1$. Funkcija f je totalno odvedljiva in velja

$$(\nabla f)(x, y) = (-2x, 2y).$$

Po Lagrangeovi metodi multiplikatorjev ima funkcija f lahko vezan ekstrem pri vezi g le v tistih točkah $(x, y) \in \mathbb{R}^2$, za katere velja

$$g(x, y) = 0$$

in

$$(\nabla f)(x, y) = \lambda(\nabla g)(x, y)$$

za nek $\lambda \in \mathbb{R}$. Kandidate za vezane ekstreme torej dobimo tako, da rešimo sistem treh enačb

$$\begin{aligned} x^2 + y^2 - 1 &= 0 \\ -2x &= \lambda 2x \\ 2y &= \lambda 2y \end{aligned}$$

za tri realne neznanke x, y, λ. Ta sistem ima štiri rešitve: prvi dve sta $(x, y) = (0, \pm 1)$ pri $\lambda = 1$, preostali dve pa sta $(x, y) = (\pm 1, 0)$ pri $\lambda = -1$. Odtod torej vidimo, da ima funkcija f lahko vezan ekstrem pri vezi g le v točkah $(0, \pm 1)$ in $(\pm 1, 0)$.

Lagrangeova metoda multiplikatorjev nam ne pove, ali ima funkcija f v točkah $(0, \pm 1)$ in $(\pm 1, 0)$ res vezane ekstreme pri vezi g, pač pa se moramo o tem prepričati na kakšen drug način. V našem primeru je $\mathbb{S}^1$ kompaktna množica, zato vemo, da mora na njej zvezna funkcija f zavzeti največjo in najmanjšo vrednost. Ker je $f(0, \pm 1) = 1$ in $f(\pm 1, 0) = -1$, lahko torej sklepamo, da ima funkcija f v točkah $(0, \pm 1)$ in $(\pm 1, 0)$ res vezane ekstreme pri vezi g: funkcija f največjo vrednost na krožnici $\mathbb{S}^1$ zavzame v točkah $(0, \pm 1)$, najmanjšo vrednost na krožnici $\mathbb{S}^1$ pa v točkah $(\pm 1, 0)$.

Vaja 4.32. (1) Poišči vezane ekstreme funkcije treh spremenljivk $f(x, y, z) = x - 2y + 2z$ na sferi $x^2 + y^2 + z^2 = 1$.

(2) Poišči točko na ploskvi $z^2 - xy = 1$ v $\mathbb{R}^3$, ki je najbližja izhodišču.