Lineární algebra pro fyziky

Zápisky z přednášek, 21/22

Verze z 28. září 2021

Najdete-li chybu, napište mi mail! Díky!

Dalibor Šmíd

Obsah

KAPITOLA 1

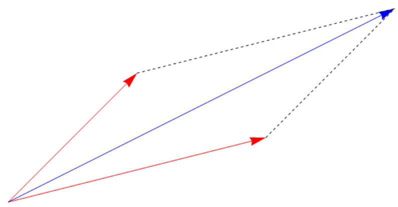
Vektory a zobrazení v $\mathbb{R}^n$

1. Vektory v $\mathbb{R}^n$

Vektor v rovině můžeme zapsat jako uspořádanou dvojici reálných čísel

$$\mathbf{x} := \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

Takové vektory tradičně znázorňujeme jako úsečky s počátkem ve zvoleném počátku kartézských souřadnic v rovině a šipkou v druhém krajním bodě o souřadnicích (x_1, x_2). Sčítání vektorů odpovídá konstrukci úhlopříčky rovnoběžníka jimi svíraného:



Z obrázku je vidět, že souřadnice modré šipky jsou součtem souřadnic šipek červených, součet vektorů tedy můžeme definovat i algebraicky

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} := \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \end{pmatrix}$$

Podobně *po složkách* definujeme i násobek vektoru skalárem (číslem $r \in \mathbb{R}$):

$$r \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} := \begin{pmatrix} rx_1 \\ rx_2 \end{pmatrix}$$

I tato definice odpovídá tradiční geometrické konstrukci: škálování vektoru uvnitř jím přímky, ve které leží.

Algebraickou definici operací snadno zobecníme na vektory v prostoru (uspořádané trojice) nebo rovnou na uspořádané n-tice, tedy vektory z $\mathbb{R}^n \equiv \underbrace{\mathbb{R} \times \ldots \times \mathbb{R}}_{n}$:

$$\mathbf{x} + \mathbf{y} \equiv \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} := \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix}$$

$$r\mathbf{x} \equiv r \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} := \begin{pmatrix} rx_1 \\ \vdots \\ rx_n \end{pmatrix}$$

Čísla x_i jsou *složky* vektoru $\mathbf{x}$. Vektor, jehož všechny složky jsou nulové, označujeme $\mathbf{o}$ (*nulový vektor*), vektor $-\mathbf{x} := (-1)\mathbf{x}$ je *opačný vektor*. Množině $\mathbb{R}^n$ budeme říkat *n-rozměrný prostor*.

2. Skalární součin

Skalární součin dvojice vektorů v rovině je definován předpisem

$$\mathbf{x}.\mathbf{y} := x_1 y_1 + x_2 y_2$$

Pomocí tohoto předpisu lze zapsat *velikost (normu)* vektoru

$$\|\mathbf{x}\| := \sqrt{\mathbf{x}.\mathbf{x}} \equiv \sqrt{x_1^2 + x_2^2}$$

Tvrzení 1. *Je-li* $\phi \in \langle 0, \pi \rangle$ *úhel mezi vektory* $\mathbf{x}, \mathbf{y} \in \mathbb{R}^2$, *pak*

$$\mathbf{x}.\mathbf{y} = \|\mathbf{x}\| \|\mathbf{y}\| \cos \phi$$

Důkaz. Označme orientovaný úhel mezi nezápornou částí první souřadné osy a polopřímkou sestávající z nezáporných násobků vektoru $\mathbf{x}$ symbolem $\alpha \in \langle 0, 2\pi)$ a podobně zaveďme $\beta \in \langle 0, 2\pi)$ pro vektor $\mathbf{y}$. Definujme $\phi' := \beta - \alpha$. Pro složky vektoru $\mathbf{x}$ platí $x_1 = \|\mathbf{x}\| \cos \alpha$, $x_2 = \|\mathbf{x}\| \sin \alpha$, podobně pro složky $\mathbf{y}$.

Dosaďte do definice skalárního součinu:

$$\mathbf{x}.\mathbf{y} = \|\mathbf{x}\| \|\mathbf{y}\| (\cos \alpha \cos \beta + \sin \alpha \sin \beta) = \|\mathbf{x}\| \|\mathbf{y}\| \cos \phi'$$

Pokud $\phi' = \beta - \alpha \in \langle 0, \pi \rangle$, pak je úhel mezi vektory ϕ roven ϕ', tedy tvrzení platí. Pokud $\phi' \in (\pi, 2\pi)$, pak je $\phi = 2\pi - \phi'$, tedy $\cos \phi = \cos \phi'$. Pokud $\phi' \in (-2\pi, 0)$, pak úhel $-\phi' \in (0, 2\pi)$ má stejný cosinus a podle předchozí úvahy buď $-\phi'$ nebo $2\pi + \phi'$ je rovno ϕ, stále se stejnou hodnotou cosinu. □

Zavedeme skalární součin a normu vektoru analogicky i na $\mathbb{R}^n$:

$$\mathbf{x}.\mathbf{y} := \sum_{i=1}^{n} x_i y_i, \quad \|\mathbf{x}\| := \sqrt{\mathbf{x}.\mathbf{x}}$$

V třírozměrném prostoru platí stejné tvrzení o souvislosti skalárního součinu a úhlu mezi vektory (viz cvičení). V obecném $\mathbb{R}^n$ nevíme, co to úhel mezi vektory je, ale můžeme jej (pro nenulové vektory $\mathbf{x}$, $\mathbf{y}$) zadefinovat jako $\phi := \arccos \frac{\mathbf{x}.\mathbf{y}}{\|\mathbf{x}\| \|\mathbf{y}\|}$. Aby tato definice byla korektní, potřebujeme vědět, že

$$-1 \leq \frac{\mathbf{x}.\mathbf{y}}{\|\mathbf{x}\| \|\mathbf{y}\|} \leq 1.$$

To je důsledkem následující věty:

Věta 1 (Schwarzova nerovnost). *Pro každé dva vektory* $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ *platí*

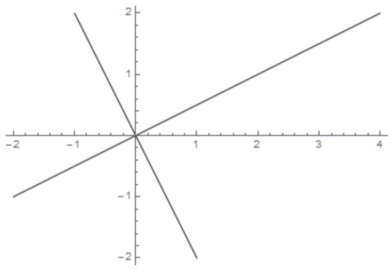
$$|\mathbf{x}.\mathbf{y}| \leq \|\mathbf{x}\| \|\mathbf{y}\|,$$

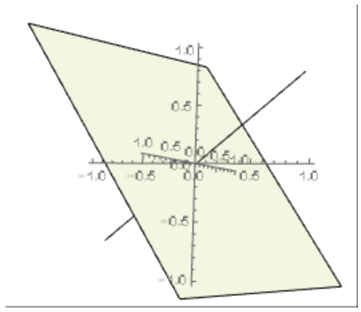
přičemž rovnost je splněna právě když je jeden z vektorů násobkem druhého.

Než větu dokážeme, je užitečné zavést některé další pojmy. Dva vektory prohlásíme za *kolmé* (značíme $\mathbf{x} \perp \mathbf{y}$), právě když $\mathbf{x}.\mathbf{y} = 0$, tedy když je jeden z nich nulový nebo když svírají úhel $\phi = \frac{\pi}{2}$. Množina

$$\mathbf{x}^{\perp} := \{\mathbf{y} \in \mathbb{R}^n | \mathbf{x}.\mathbf{y} = 0\}$$

všech vektorů $\mathbf{y}$ kolmých na vektor $\mathbf{x}$ se nazývá *ortogonální doplněk vektoru* $\mathbf{x}$. Pokud $\mathbf{x} \neq \mathbf{o}$, v $\mathbb{R}^2$ je to přímka, v $\mathbb{R}^3$ rovina, obě procházejí počátkem:





Množinám tvaru $\{\mathbf{y} \in \mathbb{R}^n | \mathbf{x}.\mathbf{y} = c\}$, kde $\mathbf{x} \neq \mathbf{o}$, říkáme obecně *nadroviny*. Nadrovina obsahuje vektor $\mathbf{o}$, právě když $c = 0$.

3. Vektorový součin

Vedle skalárního součinu známe ze středoškolské matematiky ještě součin vektorový. Pro $\mathbf{x}, \mathbf{y} \in \mathbb{R}^3$ je $\mathbf{x} \times \mathbf{y}$ vektor definovaný po složkách jako

$$(\mathbf{x} \times \mathbf{y})_k = \sum_{i=1}^{3} \sum_{j=1}^{3} \varepsilon_{ijk} x_i y_j,$$

kde $\varepsilon_{123} = \varepsilon_{231} = \varepsilon_{312} = 1$, $\varepsilon_{213} = \varepsilon_{321} = \varepsilon_{132} = -1$ a $\varepsilon_{ijk} = 0$, jsou-li dva indexy stejné. Díky tomu z devíti členů dvojité sumy přežijí jen dva, tedy

$$\begin{aligned} \mathbf{x} \times \mathbf{y} &= (\varepsilon_{231} x_2 y_3 + \varepsilon_{321} x_3 y_2, \varepsilon_{132} x_1 y_3 + \varepsilon_{312} x_3 y_1, \varepsilon_{123} x_1 y_2 + \varepsilon_{213} x_2 y_1) \\ &= (x_2 y_3 - x_3 y_2, -x_1 y_3 + x_3 y_1, x_1 y_2 - x_2 y_1) \end{aligned}$$

TVRZENÍ 2. *Nechť* $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^3$, $r, s \in \mathbb{R}$, $\mathbf{x}$ *a* $\mathbf{y}$ *svírají úhel* ϕ. *Pak*

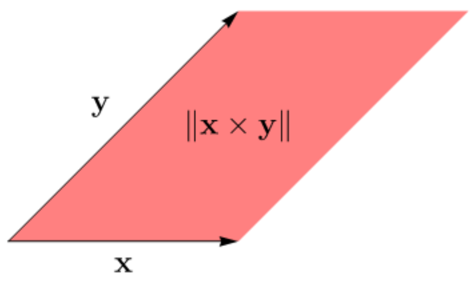
(1) $\mathbf{x} \times \mathbf{y} = -\mathbf{y} \times \mathbf{x}$, *speciálně* $\mathbf{x} \times \mathbf{x} = \mathbf{o}$
(2) $\mathbf{x} \times (r\mathbf{y} + s\mathbf{z}) = r(\mathbf{x} \times \mathbf{y}) + s(\mathbf{x} \times \mathbf{z})$
(3) $\|\mathbf{x} \times \mathbf{y}\| = \|\mathbf{x}\| \|\mathbf{y}\| |\sin \phi|$
(4) $|(\mathbf{x} \times \mathbf{y}).\mathbf{z}|$ *je objem rovnoběžnostěnu definovaného vektory* $\mathbf{x}, \mathbf{y}, \mathbf{z}$, *speciálně* 0 *pro* $z = r\mathbf{x} + s\mathbf{y}$.

Důkaz nyní nebudeme provádět, návod k němu je v dodatku na konci této kapitoly. Z druhé části posledního bodu plyne, že $\mathbf{x} \times \mathbf{y}$ je kolmý na rovinu definovanou $\mathbf{x}$ a $\mathbf{y}$.

Oproti skalárnímu součinu se nenabízí žádné přímočaré zobecnění vektorového součinu do libovolné dimenze n. V $\mathbb{R}^2$ lze vektoru $\mathbf{x}$ přiřadit vektor

$$\tilde{\mathbf{x}} := \begin{pmatrix} \mathbf{x} \\ 0 \end{pmatrix} \in \mathbb{R}^3$$

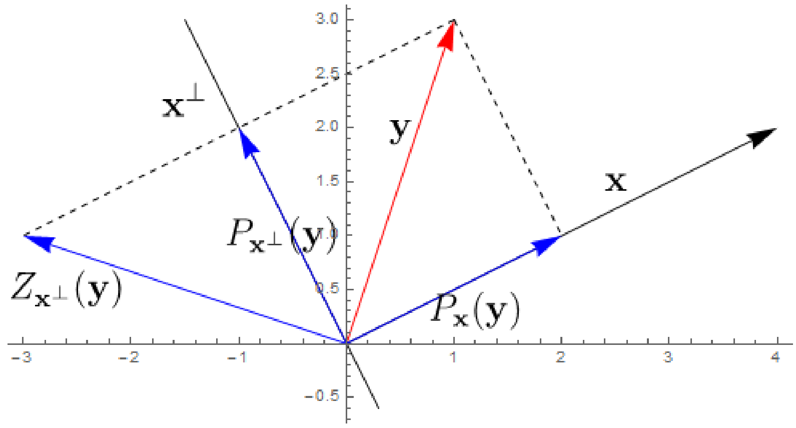
a definovat $\mathbf{x} \times \mathbf{y} := (\tilde{\mathbf{x}} \times \tilde{\mathbf{y}})_3 \in \mathbb{R}$. Pak je $\|\mathbf{x} \times \mathbf{y}\|$ obsah rovnoběžníka definovaného $\mathbf{x}$ a $\mathbf{y}$:



Plyne to ze třetí části tvrzení 2, protože $\|\mathbf{y}\| |\sin \phi|$ je výška rovnoběžníka.

4. Ortogonální projekce a zrcadlení

Pomocí skalárního součinu můžeme vyjádřit kolmý průmět (*ortogonální projekci*) vektoru $\mathbf{y}$ do směru jiného (nenulového) vektoru $\mathbf{x}$, kolmý průmět do směru jeho ortogonálního doplňku a vektor zrcadlený podle nadroviny $\mathbf{x}^\perp$.



$$P_{\mathbf{x}}(\mathbf{y}) = \|\mathbf{y}\| \cos\phi \frac{\mathbf{x}}{\|\mathbf{x}\|} = \frac{\mathbf{x}.\mathbf{y}}{\|\mathbf{x}\|^2}\mathbf{x}$$
$$P_{\mathbf{x}^\perp}(\mathbf{y}) = \mathbf{y} - P_{\mathbf{x}}(\mathbf{y}) = \mathbf{y} - \frac{\mathbf{x}.\mathbf{y}}{\|\mathbf{x}\|^2}\mathbf{x}$$
$$Z_{\mathbf{x}^\perp}(\mathbf{y}) = \mathbf{y} - 2P_{\mathbf{x}}(\mathbf{y}) = \mathbf{y} - 2\frac{\mathbf{x}.\mathbf{y}}{\|\mathbf{x}\|^2}\mathbf{x}$$

Stejné vzorce lze použít jako definice projekce a zrcadlení i v obecné dimenzi n. Pak platí

Tvrzení 3. *Nechť* $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, $\mathbf{x} \neq \mathbf{o}$. *Pak*

(1) $P_{\mathbf{x}}(\mathbf{x}) = \mathbf{x}$
(2) $P_{\mathbf{x}}(\mathbf{y}) = \mathbf{o}$ *právě když* $\mathbf{x} \perp \mathbf{y}$
(3) $P_{\mathbf{x}^\perp}(\mathbf{y}) \perp \mathbf{x}$
(4) $\forall r, s \in \mathbb{R}, \forall \mathbf{y}, \mathbf{z} \in \mathbb{R}^n : P_{\mathbf{x}}(r\mathbf{y} + s\mathbf{z}) = rP_{\mathbf{x}}(\mathbf{y}) + sP_{\mathbf{x}}(\mathbf{z})$
(5) *Vektor* $\mathbf{y}$ *lze jednoznačně zapsat jako* $\mathbf{y}_\| + \mathbf{y}_\perp$, *kde* $\mathbf{y}_\|$ *je násobkem vektoru* $\mathbf{x}$ *a* $\mathbf{y}_\perp$ *je kolmý na* $\mathbf{x}$. *Navíc pak* $\mathbf{y}_\| = P_{\mathbf{x}}(\mathbf{y})$ *a* $\mathbf{y}_\perp = P_{\mathbf{x}^\perp}(\mathbf{y})$.

Důkaz. První čtyři tvrzení plynou přímo z definic skalárního součinu, kolmosti a normy. Například třetí plyne z

$$\mathbf{x}.(\mathbf{y} - P_{\mathbf{x}}(\mathbf{y})) = \mathbf{x}.\mathbf{y} - \frac{\mathbf{x}.\mathbf{y}}{\|\mathbf{x}\|^2}\mathbf{x}.\mathbf{x} = 0$$

Z definic $P_{\mathbf{x}}$ a $P_{\mathbf{x}^\perp}$ a třetí části tvrzení plyne, že každý vektor $\mathbf{y}$ se dá zapsat jako

$$\mathbf{y} = P_{\mathbf{x}}(\mathbf{y}) + P_{\mathbf{x}^\perp}(\mathbf{y}) \equiv r\mathbf{x} + \mathbf{z},$$

kde r je nějaký skalár a $\mathbf{z} \in \mathbf{x}^\perp$. Kdyby bylo možné zapsat vektor $\mathbf{y}$ jako $s\mathbf{x}+\mathbf{z}'$, kde $\mathbf{z}' \in \mathbf{x}^\perp$, pak by $(r-s)\mathbf{x} = \mathbf{z}' - \mathbf{z}$. Vektor $\mathbf{z}' - \mathbf{z}$ je kolmý na $\mathbf{x}$, tedy $(r-s)\mathbf{x}.\mathbf{x} = 0$. Protože $\mathbf{x} \neq \mathbf{o}$, musí být $r = s$ a tedy i $\mathbf{z} = \mathbf{z}'$. Rozklad vektoru $\mathbf{y}$ na paralelní a kolmou část je tedy jednoznačný. □

$P_{\mathbf{x}}$ lze chápat jako *zobrazení*, tedy jakousi „mašinku“, do které se vloží libovolný prvek $\mathbf{y} \in \mathbb{R}^n$ a ona mu přiřadí jednoznačně určený prvek $P_{\mathbf{x}}(\mathbf{y}) \in \mathbb{R}^n$. Čtvrtá část tvrzení znamená, že toto zobrazení zachovává součty a násobení skalárem, geometricky tedy převede každý rovnoběžník opět na rovnoběžník. O takovém zobrazení říkáme, že je *lineární*. Snadno se lze přesvědčit, že i $P_{\mathbf{x}^\perp}, Z_{\mathbf{x}^\perp} : \mathbb{R}^n \to \mathbb{R}^n$ jsou lineární zobrazení, realizující projekci, resp. zrcadlení podle nadroviny $\mathbf{x}^\perp$.

Nyní můžeme konečně dokázat Schwarzovu nerovnost:

Důkaz. Předpokládejme, že alespoň jeden z vektorů je nenulový, jinak je tvrzení triviální. Nechť je to $\mathbf{x}$. Pak lze díky pátému bodu tvrzení 3 zapsat $\mathbf{y}$ jako součet $\mathbf{y}_{\|} + \mathbf{y}_{\perp}$. Protože $\mathbf{y}_{\|} \perp \mathbf{y}_{\perp}$, platí

$$\|\mathbf{y}\|^2 = (\mathbf{y}_{\|} + \mathbf{y}_{\perp}).(\mathbf{y}_{\|} + \mathbf{y}_{\perp}) = \|\mathbf{y}_{\|}\|^2 + \|\mathbf{y}_{\perp}\|^2 \geq \|\mathbf{y}_{\|}\|^2$$

Z pátého bodu tvrzení 3 rovněž víme, že $\mathbf{y}_{\|} = P_{\mathbf{x}}(\mathbf{y})$. Dosazením z definice ortogonální projekce a algebraickou úpravou dostaneme, že

$$\|\mathbf{y}\|^2 \|\mathbf{x}\|^2 \geq (\mathbf{x}.\mathbf{y})^2$$

Protože normy jsou nezáporné, stačí po odmocnění přidat absolutní hodnotu jen na pravou stranu nerovnosti. Rovnost nastává právě když $\|\mathbf{y}_{\perp}\|^2 = 0$, tedy právě když jsou všechny složky $\mathbf{y}_{\perp}$ nulové. To nastane právě když $\mathbf{y} = \mathbf{y}_{\|}$, tedy je-li $\mathbf{y}$ násobkem $\mathbf{x}$. □

5. Rotace v $\mathbb{R}^2$ a $\mathbb{R}^3$, dilatace a posunutí

Otočení (*rotaci*) vektoru $\mathbf{y} \in \mathbb{R}^2$ okolo počátku o úhel ϕ proti směru hodinových ručiček lze zapsat jako zobrazení

$$\mathbf{y} := \begin{pmatrix} \|y\| \cos(\alpha) \\ \|y\| \sin(\alpha) \end{pmatrix} \mapsto R_\phi(\mathbf{y}) := \begin{pmatrix} \|y\| \cos(\alpha + \phi) \\ \|y\| \sin(\alpha + \phi) \end{pmatrix}$$

Po aplikaci součtových vzorců máme vyjádření ve složkách:

$$R_\phi(\mathbf{y}) := \begin{pmatrix} y_1 \cos\phi - y_2 \sin\phi \\ y_1 \sin\phi + y_2 \cos\phi \end{pmatrix} = \cos\phi \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} + \sin\phi \begin{pmatrix} -y_2 \\ y_1 \end{pmatrix}$$

Několik pozorování:

- R_ϕ je lineární zobrazení.
- $R_0(\mathbf{y}) = \mathbf{y}$, tedy $R_0 = \mathrm{Id}$ (*identita*).
- Složení dvou rotací $R_\phi \circ R_\psi$ je rotace $R_{\phi+\psi}$. Na pořadí složení zde nezáleží, říkáme, že R_ϕ a R_ψ *komutují.*
- $R_{-\phi} \circ R_\phi = R_\phi \circ R_{-\phi} = R_0 = \mathrm{Id}$, tedy zobrazení $R_{-\phi}$ je *inverzní* k R_ϕ, $R_{-\phi} = (R_\phi)^{-1}$.

Rotaci vektoru $\mathbf{y} \in \mathbb{R}^3$ okolo osy dané vektorem $\mathbf{v}$ ($\|\mathbf{v}\| = 1$) můžeme zapsat například pomocí Rodriguesovy formule:

$$R_{\mathbf{v},\phi}(\mathbf{y}) = \cos\phi \ \mathbf{y} + (1 - \cos\phi)(\mathbf{v}.\mathbf{y})\mathbf{v} + \sin\phi \ \mathbf{v} \times \mathbf{y}.$$

Návod k jejímu odvození naleznete ve cvičeních.

Rotace v $\mathbb{R}^3$ komutují, pokud mají společnou osu, jindy komutovat nemusí. Vektory

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \mathbf{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

a jejich přímočará zobecnění do $\mathbb{R}^n$ se nazývají *prvky kanonické báze.* Nekomutativita rotací se projeví třeba na příkladu

$$R_{\mathbf{e}_2, \frac{\pi}{2}} \circ R_{\mathbf{e}_3, \frac{\pi}{2}}(\mathbf{e}_3) = R_{\mathbf{e}_2, \frac{\pi}{2}}(\mathbf{e}_3) = \mathbf{e}_1$$
$$R_{\mathbf{e}_3, \frac{\pi}{2}} \circ R_{\mathbf{e}_2, \frac{\pi}{2}}(\mathbf{e}_3) = R_{\mathbf{e}_3, \frac{\pi}{2}}(\mathbf{e}_1) = \mathbf{e}_2$$

Poznámka 1. Rotaci v $\mathbb{R}^n$ je možné definovat podobně jako v $\mathbb{R}^3$, pouze „osa“ již nebude přímka, ale bude mít dimenzi $n-2$. Příkladem rotace v $\mathbb{R}^4$ fixující rovinu vektorů $\mathbf{e_1}$ a $\mathbf{e_4}$ (značíme $\langle \mathbf{e}_1, \mathbf{e}_4 \rangle$) je

$$R_{\langle \mathbf{e}_1, \mathbf{e}_4 \rangle, \phi}(\mathbf{y}) := \begin{pmatrix} y_1 \\ y_2 \cos\phi - y_3 \sin\phi \\ y_2 \sin\phi + y_3 \cos\phi \\ y_4 \end{pmatrix}$$

Dalším příkladem lineárního zobrazení z $\mathbb{R}^n$ do $\mathbb{R}^n$ je roztažení(*dilatace*) faktorem $\lambda \in \mathbb{R}$:

$$D_\lambda(\mathbf{y}) := \lambda \mathbf{y}$$

Naopak posunutí (*translace*) o nenulový vektor $\mathbf{b} \in \mathbb{R}^n$

$$T_\mathbf{b}(\mathbf{y}) := \mathbf{y} + \mathbf{b}$$

lineárním zobrazením není. Lineární zobrazení F totiž musí splňovat $F(\mathbf{o}) = F(0.\mathbf{x} + 0.\mathbf{y}) = 0F(\mathbf{x}) + 0F(\mathbf{y}) = \mathbf{o}$, pro translaci ale $T_\mathbf{b}(\mathbf{o}) = \mathbf{b}$.

6. Matice lineárního zobrazení

Jak tedy poznáme, že je nějaké zobrazení $F : \mathbb{R}^n \to \mathbb{R}^n$ lineární?

Každý vektor $\mathbf{x} \in \mathbb{R}^n$ lze zapsat jako

$$\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \ldots + x_n\mathbf{e}_n,$$

neboli jako *lineární kombinaci* prvků kanonické báze. Z linearity zobrazení F plyne, že

$$\begin{aligned} F(\mathbf{x}) &= F(x_1\mathbf{e}_1 + \ldots + x_n\mathbf{e}_n) = x_1F(\mathbf{e}_1) + \ldots + x_nF(\mathbf{e}_n) \\ &= x_1\mathbf{a}_1 + \ldots + x_n\mathbf{a}_n, \end{aligned}$$

kde jsme jako $\mathbf{a}_j \in \mathbb{R}^n$ označili obraz $F(\mathbf{e}_j)$ vektoru $\mathbf{e}_j$. Označíme-li i-tou složku vektoru $\mathbf{a}_j$ jako a_{ij}, je to dohromady n^2 reálných čísel, která společně zobrazení F jednoznačně určují. Tato čísla souhrnně označujeme jako *matici lineárního zobrazení* a značíme $A \equiv (a_{ij})$.

Zobrazení určené maticí A označujeme symbolem F_A. Pak tedy můžeme zavést zápis lineárního zobrazení pomocí jeho *matice*:

$$\begin{aligned} F_A(\mathbf{x}) &:= x_1 \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{pmatrix} + x_2 \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{pmatrix} + \ldots + x_n \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{nn} \end{pmatrix} \\ &=: \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \equiv A\mathbf{x} \end{aligned}$$

Druhou rovností jsme jednak zavedli standardní zápis matice jako tabulky čísel, v níž první index čísluje řádky a druhý sloupce. Zároveň jsme také zadefinovali *součin matice a vektoru*.

Součin matice A a vektoru $\mathbf{x}$ je vlastně lineární kombinací sloupců matice A s koeficienty rovnými složkám vektoru $\mathbf{x}$. Užívá se proto také *sloupcový zápis matice*

$A = (\mathbf{a}_1|\mathbf{a}_2|\ldots|\mathbf{a}_n)$. Zobrazení s maticí

$$E := (\mathbf{e}_1|\ldots|\mathbf{e}_n) \equiv \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 1 \end{pmatrix}$$

přiřazuje každému vektoru $\mathbf{x}$ vektor

$$F_E(\mathbf{x}) = E\mathbf{x} = x_1\mathbf{e}_1 + \ldots + x_n\mathbf{e}_n = \mathbf{x},$$

je to tedy identita $F_E = \mathrm{Id}$. Matici E se říká *jednotková matice.*

Označíme-li symbolem

$$\tilde{\mathbf{a}}_i := \begin{pmatrix} a_{i1} \\ a_{i2} \\ \vdots \\ a_{in} \end{pmatrix}$$

i-tý řádek matice A, zapsaný ovšem jako sloupcový vektor, pak i-tá složka obrazu vektoru $\mathbf{x}$ v zobrazení F_A je

$$F_A(\mathbf{x})_i \equiv (A\mathbf{x})_i = a_{i1}x_1 + a_{i2}x_2 + \ldots a_{in}x_n \equiv \sum_{j=1}^{n} a_{ij}x_j,$$

což lze interpretovat jako skalární součin vektorů $\tilde{\mathbf{a}}_i$ a $\mathbf{x}$. Pokud tedy $F_A(\mathbf{x})_i = 0$, je vektor $\mathbf{x}$ kolmý na $\tilde{\mathbf{a}}_i$. Podobně pokud $F_A(\mathbf{x})_i = c$, leží $\mathbf{x}$ v nadrovině s rovnicí $\tilde{\mathbf{a}}_i.\mathbf{x} = c$.

7. Dodatek k vektorovému součinu

Při zobecnění vektorového součinu do dimenze $n > 3$ chceme, aby se pomocí něj dál počítaly objemy rovnoběžnostěnů, což lze splnit definicí

$$(\mathbf{x}_1 \times \mathbf{x}_2 \times \ldots \times \mathbf{x}_{n-1})_{k_n} := \sum_{k_1=1}^{n} \sum_{k_2=1}^{n} \ldots \sum_{k_{n-1}=1}^{n} \varepsilon_{k_1 k_2 \ldots k_n} x_{1k_1} x_{2k_2} \ldots x_{n-1,k_{n-1}},$$

kde symbol ε je opět definován tak, aby měnil znaménko při výměně libovolné dvojice indexů a aby $\varepsilon_{12\ldots n} = 1$. Zobecněný vektorový součin tedy již není funkce dvou, ale $n-1$ vektorových argumentů, výsledkem je opět vektor. Jsou ale splněny analogie vlastností $1, 2$ a 4 z tvrzení 2. Například objem rovnoběžnostěnu určeného vektory $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d} \in \mathbb{R}^4$ je roven

$$|(\mathbf{a} \times \mathbf{b} \times \mathbf{c}).\mathbf{d}| = \left| \sum_{i=1}^{4} \sum_{j=1}^{4} \sum_{k=1}^{4} \sum_{l=1}^{4} \varepsilon_{ijkl} a_i b_j c_k d_l \right|$$

V kapitole 9 uvidíme, že podobně se definuje determinant a také jeho geometrický význam je podobný.

Ještě slíbený návod k důkazu tvrzení 2:

(1) Dokažte body 1 a 2 přímým dosazením z definice.
(2) Definujme veličinu

$$V(\mathbf{x}, \mathbf{y}, \mathbf{z}) := \sum_{i=1}^{3} \sum_{j=1}^{3} \sum_{k=1}^{3} \varepsilon_{ijk} x_i y_j z_k$$

S pomocí bodu 2 ukažte, že $V(\mathbf{x}, \mathbf{y}, \mathbf{z}) = V(\mathbf{x} + s\mathbf{y}, \mathbf{y}, \mathbf{z})$. Rozmyslete si, že rovnoběžnostěn definovaný vektory $\mathbf{x}, \mathbf{y}, \mathbf{z}$ má stejný objem jako rovnoběžnostěn definovaný vektory $\mathbf{x} + s\mathbf{y}, \mathbf{y}, \mathbf{z}$.

(3) Buď $x_1 = y_1 = z_1 = 0$, nebo je alespoň jedno z těchto čísel nenulové. Ukažte, že

$$V(\mathbf{x}, \mathbf{y}, \mathbf{z}) = -V(\mathbf{y}, \mathbf{x}, \mathbf{z}) = V(\mathbf{z}, \mathbf{x}, \mathbf{y})$$

a odvoďte z toho, že v druhém případě bez újmy na obecnosti $x_1 \neq 0$. Pak lze pomocí předchozího bodu převést vektory $\mathbf{x}, \mathbf{y}, \mathbf{z}$ na takové, že $y_1' = z_1' = 0$, aniž by se změnila hodnota V nebo objem rovnoběžnostěnu. Rozeberte případ $x_1 = y_1 = z_1 = 0$.

(4) Ukažte, že vektory lze dále převést na trojici $\mathbf{x}, \mathbf{y}', \mathbf{z}''$, kde $z_1'' = z_2'' = 0$ a V ani objem rovnoběžnostěnu se nezměnily. Dále ukažte, že $V(\mathbf{x}, \mathbf{y}', \mathbf{z}'') = x_1 y_2' z_3''$ a že $|x_1 y_2' z_3''|$ je právě objem rovnoběžnostěnu definovaného vektory $\mathbf{x}, \mathbf{y}', \mathbf{z}''$. Tím je dokázán bod 4.

(5) Z bodu 4 ukažte, že $\|\mathbf{x} \times \mathbf{y}\|$ je obsah podstavy rovnoběžnostěnu definované vektory $\mathbf{x}, \mathbf{y}$. Odtud odvoďte bod 3.

KAPITOLA 2

Matice

V této kapitole se podíváme na matice jako na samostatné objekty a zavedeme na nich operace, jejichž vlastnosti vyplývají z vlastností zobrazení. Začneme ale tím, že se podíváme, jak hledání vzoru určitého vektoru v daném lineárním zobrazení vede na soustavu lineárních rovnic.

1. Lineární zobrazení a soustavy rovnic

Příklad 1. Uvažujme matici a vektor

$$A = \begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix}, \mathbf{b} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

Zajímá nás, na jaký vektor zobrazí F_A vektor $\mathbf{b}$ a které vektory se naopak zobrazí na něj. V prvním případě snadno spočteme, že

$$F_A(\mathbf{b}) = \begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1.(-1) + 2.1 \\ 1.(-1) + 3.1 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

V druhém hledáme obecný tvar vektoru $\mathbf{x}$, pro který

$$F_A(\mathbf{x}) \equiv \begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \equiv \begin{pmatrix} x_1 + 2x_2 \\ x_1 + 3x_2 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix},$$

což je vlastně soustava dvou rovnic pro dvě neznámé x_1 a x_2. Každá rovnice je rovnicí přímky v $\mathbb{R}^2$, jsou různoběžné, jediným řešením (a tedy i jediným vzorem vektoru $\mathbf{b}$ v zobrazení F_A) je jejich průsečík, vektor

$$\mathbf{x} = \begin{pmatrix} -5 \\ 2 \end{pmatrix}$$

Co kdybychom hledali průsečnici dvou rovin v $\mathbb{R}^3$? Soustavu

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 &= b_2 \end{aligned}$$

lze opět interpretovat jako rovnost vektorů

$$x_1 \begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix} + x_2 \begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix} + x_3 \begin{pmatrix} a_{13} \\ a_{23} \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$$

Lineární kombinace na levé straně odpovídá definici součinu matice a vektoru, až na to, že matice A nemá stejný počet řádků jako sloupců. Pojďme tedy tuto definici rozšířit.

Definice 1. Nechť $m, n \in \mathbb{N}$ a $\mathbf{a}_1, \ldots, \mathbf{a}_n$ jsou vektory z $\mathbb{R}^m$, $\mathbf{x} \in \mathbb{R}^n$. Pak definujme matici $A = (\mathbf{a}_1 | \ldots | \mathbf{a}_n)$ a její součin s vektorem $\mathbf{x}$

$$A\mathbf{x} := x_1\mathbf{a}_1 + \ldots + x_n\mathbf{a}_n \in \mathbb{R}^m$$

Matice s m řádky a n sloupci se označuje jako matice $m \times n$. Pokud $m = n$, říká se jí *čtvercová*. Pokud $\mathbf{a}$ je nějaký vektor

$$\mathbf{a} := \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \in \mathbb{R}^n,$$

pak rovnici nadroviny $\mathbf{a}.\mathbf{x} = c$ můžeme zapsat pomocí matice $1 \times n$ jako

$$\begin{pmatrix} a_1 & \dots & a_n \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = c$$

Matice vzniklá z matice A výměnou řádků za sloupce se nazývá *matice transponovaná k matici A*:

$$A^T \equiv \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}^T := \begin{pmatrix} a_{11} & \dots & a_{m1} \\ a_{12} & \ddots & a_{m2} \\ \vdots & & \vdots \\ a_{1n} & \dots & a_{mn} \end{pmatrix}$$

S pomocí operace transponování můžeme rovnici nadroviny napsat také jako $\mathbf{a}^T\mathbf{x} = c$, přičemž sloupcový vektor $\mathbf{a}$ interpretujeme jako matici $n \times 1$.

PŘÍKLAD 2. Vraťme se k hledání průsečnice rovin a vezměme nějakou konkrétní matici A a konkrétní *vektor pravých stran* $\mathbf{b}$:

$$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

Řešíme tedy *soustavu lineárních rovnic* (*SLR*)

$$\begin{aligned} x_1 + 2x_2 + 3x_3 &= 1 \\ x_1 + 3x_2 + 4x_3 &= 1 \end{aligned}$$

Nahradíme-li druhou rovnici soustavy rozdílem druhé a první rovnice, množina řešení se nezmění (cvičení 1). Novou soustavu

$$\begin{aligned} x_1 + 2x_2 + 3x_3 &= 1 \\ x_2 + x_3 &= 0 \end{aligned}$$

vyřešíme položením $x_3 := t \in \mathbb{R}$ a dosazením do druhé a poté do první rovnice:

$$\mathbf{x} \equiv \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 - t \\ -t \\ t \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + t \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}.$$

To je v souladu s očekáváním přímka procházející bodem $(1, 0, 0)^T$ a mající směrový vektor $(-1, -1, 1)^T$.

PŘÍKLAD 3. Podobně můžeme najít i *parametrické vyjádření* roviny $x_1 + 2x_2 + 3x_3 = 1$. Položíme neznámé $x_2 := t$, $x_3 := s$ rovny dvěma libovolným parametrům a dopočteme $x_1 = 1 - 2t - 3s$. Vektorově lze množinu řešení zapsat jako

$$\left\{ \mathbf{x} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + t \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} + s \begin{pmatrix} -3 \\ 0 \\ 1 \end{pmatrix} \middle| t, s \in \mathbb{R} \right\},$$

tedy jako množinu „bod plus lineární kombinace směrových vektorů“. Takový tvar má množina řešení každé SLR (cvičení 2).

Cvičení 1

Nechť $(x_1, \ldots, x_n)^T$ je n-tice čísel splňujících SLR1

$$\begin{array}{ccccccccc} a_1x_1 & + & a_2x_2 & + & \ldots & + & a_nx_n & = & c \\ b_1x_1 & + & b_2x_2 & + & \ldots & + & b_nx_n & = & d \end{array}$$

Vyvoďte z toho, že $(x_1, \ldots, x_n)^T$ splňuje také SLR2

$$\begin{array}{ccccccccc} a_1x_1 & + & a_2x_2 & + & \ldots & + & a_nx_n & = & c \\ (b_1 + ra_1)x_1 & + & (b_2 + ra_2)x_2 & + & \ldots & + & (b_n + ra_n)x_n & = & d + rc \end{array}$$

pro libovolné $r \in \mathbb{R}$. Obdobně dokažte, že každé řešení SLR2 je i řešením SLR1. Celkově tedy přičtení násobku první rovnice do druhé rovnice nemění množinu řešení.

Cvičení 2

Nechť $\mathbf{x}_P = (x_{P1}, \ldots, x_{Pn})^T$ je řešením SLR1

$$\begin{array}{ccccccccc} a_1x_1 & + & a_2x_2 & + & \ldots & + & a_nx_n & = & c \\ b_1x_1 & + & b_2x_2 & + & \ldots & + & b_nx_n & = & d \end{array}$$

a $\mathbf{x}_H = (x_{H1}, \ldots, x_{Hn})^T, \mathbf{y} = (y_{H1}, \ldots, y_{Hn})^T$ je řešením SLR2

$$\begin{array}{ccccccccc} a_1x_1 & + & a_2x_2 & + & \ldots & + & a_nx_n & = & 0 \\ b_1x_1 & + & b_2x_2 & + & \ldots & + & b_nx_n & = & 0\text{'} \end{array}$$

která vznikne ze SLR1 nahrazením pravých stran nulami. Ukažte, že $\mathbf{x}_P + \mathbf{x}_H$ je řešením SLR1 a že libovolná lineární kombinace $r\mathbf{x}_H + s\mathbf{y}_H$ je řešením SLR2. K důkazu tvrzení, že množina řešení každé SLR je „bod plus lineární kombinace směrových vektorů“ se potřebujeme ještě seznámit s pojmem báze, viz definice 16.

2. Skládání zobrazení a součin matic

Uvažujme dvě zobrazení $F_A : \mathbb{R}^n \to \mathbb{R}^m$ a $F_B : \mathbb{R}^m \to \mathbb{R}^p$. Složené zobrazení $F_B \circ F_A : \mathbb{R}^n \to \mathbb{R}^p$ je opět lineární (ověřte sami) a jeho matici můžeme určit tak, že spočítáme obrazy vektorů $\mathbf{e}_1, \ldots, \mathbf{e}_n \in \mathbb{R}^n$:

$$(F_B \circ F_A)(\mathbf{e}_i) = F_B(A\mathbf{e}_i) = F_B(\mathbf{a}_i) = B\mathbf{a}_i,$$

tedy pro obecný vektor $\mathbf{x} \in \mathbb{R}^n$

$$(F_B \circ F_A)(\mathbf{x}) = (B\mathbf{a}_1 | \ldots | B\mathbf{a}_n)\mathbf{x}$$

To nám dává návod, jak definovat *součin matic*:

Definice 2. Nechť A je matice $m \times n$ a B je matice $p \times m$. Pak definujme

$$BA := (B\mathbf{a}_1 | \ldots | B\mathbf{a}_n)$$

Z definice automaticky plyne, že $F_B \circ F_A = F_{BA}$. BA je matice $p \times n$, tedy má stejně sloupců jako A a řádků jako B.

Výhoda této definice je, že je konzistentní s dříve definovaným součinem matice a vektoru, chápeme-li vektor jako matici o jednom sloupci. Element matice BA na pozici ij je

$$b_{i1}a_{1j} + b_{i2}a_{2j} + \ldots + b_{im}a_{mj} \equiv \sum_{k=1}^{m} b_{ik}a_{kj},$$

což je vlastně skalární součin $\tilde{\mathbf{b}}_i^T.\mathbf{a}_j$ i-tého řádku matice B a j-tého sloupce matice A. Odtud je jednoduše vidět, že

$$E_mA = AE_n = A,$$

kde E_m, E_n značí jednotkovou matici $m \times m$, resp. $n \times n$.

Označme množinu všech $m \times n$ matic symbolem $\mathbb{R}^{m \times n}$.

Tvrzení 4. *Nechť $m, n, p, q \in \mathbb{N}$, $A \in \mathbb{R}^{m\times n}$, $B \in \mathbb{R}^{n\times p}$, $C \in \mathbb{R}^{p\times q}$. Pak*

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

tedy součin matic je asociativní*.*

Důkaz. Pro libovolný $\mathbf{x} \in \mathbb{R}^q$ platí

$$\begin{aligned}((AB)C)\mathbf{x} &= (F_{AB} \circ F_C)(\mathbf{x}) = F_{AB}(F_C(\mathbf{x})) = \\ &= F_A(F_B(F_C(\mathbf{x}))) = (F_A \circ F_{BC})(\mathbf{x}) = (A(BC))\mathbf{x},\end{aligned}$$

Pokud za $\mathbf{x}$ dosadíme prvek $\mathbf{e}_i$, získáme rovnost mezi i-tým sloupcem matice $(AB)C$ a matice $A(BC)$. Obě matice jsou typu $m \times q$, tedy se musejí rovnat. □

Definice 3. Nechť $A, B \in \mathbb{R}^{m\times n}$. Pak *součet matic* $A + B$ definujeme jako

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} + \begin{pmatrix} b_{11} & \dots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{m1} & \dots & b_{mn} \end{pmatrix} = \begin{pmatrix} a_{11}+b_{11} & \dots & a_{1n}+b_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1}+b_{m1} & \dots & a_{mn}+b_{mn} \end{pmatrix}$$

a násobek matice skalárem $r \in \mathbb{R}$ jako

$$r\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} = \begin{pmatrix} ra_{11} & \dots & ra_{1n} \\ \vdots & \ddots & \vdots \\ ra_{m1} & \dots & ra_{mn} \end{pmatrix}$$

Je to přirozená definice, protože je v případě matic s jediným sloupcem v souladu s definicí součtu vektorů a jejich násobení skalárem. Pojmem *nulová matice* označujeme matici, jejíž všechny elementy jsou 0, obvykle ji stejným symbolem i značíme.

Další důležitou vlastností součinu matic je jeho *distributivita*.

Tvrzení 5. *Nechť $m, n, p \in \mathbb{N}$, $A, B \in \mathbb{R}^{m\times n}$, $C, D \in \mathbb{R}^{n\times p}$. Pak*

$$\begin{aligned}(A+B)C &= AC + BC \\ A(C+D) &= AC + AD,\end{aligned}$$

Důkaz. Nejprve si uvědomme, že $\forall \mathbf{x} \in \mathbb{R}^n$

$$(F_A + F_B)(\mathbf{x}) = F_A(\mathbf{x}) + F_B(\mathbf{x}) = A\mathbf{x} + B\mathbf{x}$$

Odtud ověříme, že $F_A + F_B$ je lineární (dosazením $\mathbf{x} = r\mathbf{y} + s\mathbf{z}$ a rozepsáním pravé strany) a že má matici $A + B$ (dosazením $\mathbf{x} = \mathbf{e}_i$ jako při hledání matice součinu a dokazování asociativity). Tedy $F_A + F_B = F_{A+B}$. Pak, opět jako při dokazování asociativity, ukážeme, že

$$\begin{aligned}((A+B)C)\mathbf{x} &= (F_{A+B} \circ F_C)(\mathbf{x}) = F_{A+B}(F_C(\mathbf{x})) = \\ &= (F_A + F_B)(F_C(\mathbf{x})) = F_A(F_C(\mathbf{x})) + F_B(F_C(\mathbf{x})) = AC\mathbf{x} + BC\mathbf{x}\end{aligned}$$

Rozmyslete si, z jakých přesných důvodů platí jednotlivé rovnosti! Jako v důkazu asociativity vede dosazení $\mathbf{x} = \mathbf{e}_i$ k rovnosti $(A+B)C = AC + BC$. Druhá rovnost v tvrzení se dokáže analogicky. □

Maticový součin ale určitě není komutativní. Je-li definován součin BA, součin AB definován být nemusí, případně nemusí být stejného typu:

$$\begin{pmatrix} b_1 & \dots & b_n \end{pmatrix} \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = \sum_{i=1}^{n} b_i a_i$$

$$\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \begin{pmatrix} b_1 & \dots & b_n \end{pmatrix} = \begin{pmatrix} a_1 b_1 & \dots & a_1 b_n \\ \vdots & \ddots & \vdots \\ a_n b_1 & \dots & a_n b_n \end{pmatrix}$$

I v případě, že jsou A i B čtvercové matice $n \times n$ a tudíž jsou stejného typu i jejich součiny AB a BA, obvykle se tyto součiny nerovnají. Příklad už jsme viděli u rotací v prostoru, ale každý si snadno může vyrobit vlastní vygenerováním nějakých dvou „náhodných“ matic, stačí i 2×2.

Tvrzení 6. *Je-li A matice $m \times n$ a B matice $n \times p$, pak $(AB)^T = B^T A^T$.*

Důkaz. Označme $C = AB$, pak $c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$. Pak

$$c^T_{ji} = c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj} = \sum_{k=1}^{n} a^T_{ki} b^T_{jk} = \sum_{k=1}^{n} b^T_{jk} a^T_{ki}$$

Poslední suma je právě ji-tý element matice $B^T A^T$. □

Cvičení 3

Součin matic umožňuje definovat také mocninu čtvercové matice A^k jako k-násobný součin matice A se sebou samou. Dále definujeme $A^0 = E$. Spočtěte všechny přirozené mocniny matice

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

Na základě předchozího výpočtu se pokuste najít příklady matic, pro které $A^k = 0$, ale $A^{k-1} \neq 0$ pro libovolné $k \in \mathbb{N}$.

3. Inverzní matice

Chceme-li nalézt inverzní zobrazení k $F_A : \mathbb{R}^n \to \mathbb{R}^n$, můžeme se pokusit jej hledat ve tvaru F_B pro nějakou matici B. Podmínka

$$\forall \mathbf{x} \in \mathbb{R}^n : (F_B \circ F_A)(\mathbf{x}) = \mathbf{x} \ \wedge \ (F_A \circ F_B)(\mathbf{x}) = \mathbf{x}$$

znamená, že $BA = AB = E$ (dosaďte opět $\mathbf{x} := \mathbf{e}_i$). Naopak pokud taková matice B existuje, pak $F_B \circ F_A = F_A \circ F_B = \mathrm{Id}$, takže F_B je inverzní zobrazení k F_A.

Definice 4. Nechť $A \in \mathbb{R}^{n \times n}$. Matice A^{-1}, která splňuje $AA^{-1} = A^{-1}A = E$, se nazývá *matice inverzní k A*. Pokud matice A inverzní matici má, nazývá se *regulární*, jinak *singulární*.

Tvrzení 7. *Nechť $A, B, C \in \mathbb{R}^{n \times n}$ splňují $BA = AC = E$. Pak $B = C$.*

Důkaz. Plyne z asociativity součinu:

$$B = BE = B(AC) = (BA)C = EC = C$$

. □

Z tohoto tvrzení plyne, že inverzní matice k A může existovat nejvýše jedna.

Příklad 4. Zkusme si inverzní matici spočítat pro konkrétní volbu matice

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

Hledejme 3×3 matici B s vlastností $AB = E$. Pak její i-tý sloupec splňuje rovnici $A\mathbf{b}_i = \mathbf{e}_i$. Pro $i = 1$ to znamená

$$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 4 \\ 1 & 0 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix},$$

kde jsme označili $\mathbf{x} := \mathbf{b}_1$.

Soustavu lineárních rovnic je zvykem zapisovat ve tvaru *rozšířené matice soustavy*

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 1 & 3 & 4 & 0 \\ 1 & 0 & 2 & 0 \end{array}\right)$$

Každý řádek matice odpovídá rovnici roviny v $\mathbb{R}^3$, celá soustava pak hledání průniku tří rovin. Podobně jako dříve při hledání průsečnice můžeme první rovnici odečíst od druhé. V rozšířené matici se to projeví jako přičtení (-1)-násobku prvního řádku do druhého. Následně provedeme stejnou úpravu i se třetím řádkem.

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 0 & 1 & 1 & -1 \\ 1 & 0 & 2 & 0 \end{array}\right) \sim \left(\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 0 & 1 & 1 & -1 \\ 0 & -2 & -1 & -1 \end{array}\right)$$

Symbol $\sim$ se čte „se převede ekvivalentní úpravou na“. *Ekvivalentní úpravy* jsou ty, které zachovávají množinu řešení.

Po přičtení dvojnásobku druhého řádku do třetího dostáváme soustavu

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 0 & 1 & 1 & -1 \\ 0 & 0 & 1 & -3 \end{array}\right) \quad \leftrightarrow \quad \begin{aligned} x_1 + 2x_2 + 3x_3 &= 1 \\ x_2 + x_3 &= -1 \\ x_3 &= -3 \end{aligned}$$

Z ní postupným dosazením získáme $x_3 = -3$, $x_2 = 2$, $x_1 = 6$. Můžeme také dosazování nahradit dalšími ekvivalentními úpravami:

$$\sim \left(\begin{array}{ccc|c} 1 & 2 & 0 & 10 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & -3 \end{array}\right) \sim \left(\begin{array}{ccc|c} 1 & 0 & 0 & 6 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & -3 \end{array}\right)$$

Získali jsme tak $\mathbf{x} = \mathbf{b}_1$, první sloupec inverzní matice. Stejný výpočet s pravými stranami $\mathbf{e}_2$ a $\mathbf{e}_3$ dá zbývající dva. Pravé strany nemají vliv na to, které ekvivalentní úpravy volíme, můžeme tedy řešit všechny tři soustavy naráz, symbolicky

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 1 & 3 & 4 & 0 & 1 & 0 \\ 1 & 0 & 2 & 0 & 0 & 1 \end{array}\right) \sim \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 6 & -4 & -1 \\ 0 & 1 & 0 & 2 & -1 & -1 \\ 0 & 0 & 1 & -3 & 2 & 1 \end{array}\right)$$

Zjistili jsme tedy, že matice za svislou čarou je jediná, která splňuje rovnost $AB = E$. Vynásobením snadno ověříme, že i $BA = E$, tedy $B = A^{-1}$.

Víme ale, že tímto postupem získaná matice bude vždy dávat jednotkovou matici i při vynásobení z druhé strany? A že lze vždy buď najít ekvivalentní úpravy, kterými lze A^{-1} najít, nebo ukázat, že neexistuje?

Cvičení 4

Dokažte, že pokud $A, B \in \mathbb{R}^{n\times n}$ jsou regulární matice, pak matice $B^{-1}A^{-1}$ je inverzní maticí k matici AB, a tedy i AB je regulární matice.

Cvičení 5

Dokažte, že pokud A je regulární matice, pak A^T je také regulární matice, a její inverzní matice $(A^T)^{-1}$ je $(A^{-1})^T$, tedy matice transponovaná k A^{-1}.

KAPITOLA 3

Soustavy lineárních rovnic

V minulé kapitole jsme se setkali s několika úlohami vedoucími na soustavy lineárních rovnic:

- hledání průsečnice rovin (či obecně průniku nadrovin)
- hledání lineární kombinace vektorů $\mathbf{a}_1, \dots, \mathbf{a}_n$, která dává vektor $\mathbf{b}$
- hledání vzoru vektoru $\mathbf{b}$ v zobrazení F_A.
- hledání sloupců inverzní matice

Další příklady jsou řešení rovnic vyplývajících z Kirchhoffových zákonů pro elektrické obvody, chemických rovnic, numerické aproximace řešení diferenciálních rovnic, proložení bodů polynomem nebo jinou funkcí či křivkou a mnoho dalších.

Řešení soustav lineárních rovnic je nejdůležitější úloha lineární algebry. Pokusíme se jí proto co nejlépe porozumět.

1. Odstupňovaný tvar matice

Definice 5. *Elementární řádkovou úpravou (EŘÚ)* matice rozumíme buď

(1) prohození dvou řádků
(2) přičtení libovolného násobku nějakého řádku do jiného řádku
(3) vynásobení nějakého řádku nenulovým číslem

Elementární maticí rozumíme matici, která vznikne z jednotkové matice elementární řádkovou úpravou.

Příklad 5. Příklady 4×4 elementárních matic jednotlivých úprav jsou

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ r & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & s & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

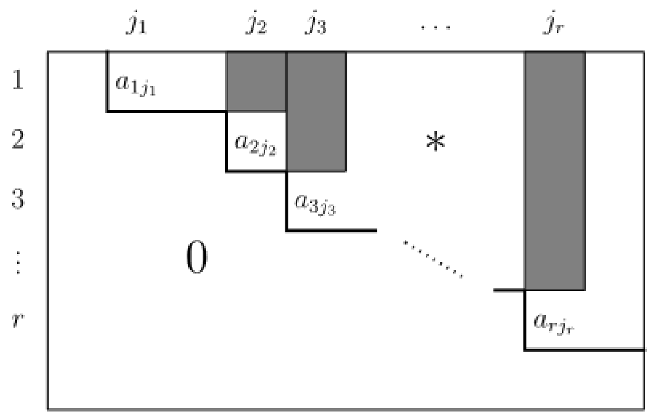
kde $s \neq 0$. Vyzkoušejte si, že pokud libovolnou matici A vynásobíme některou z těchto matic zleva, má to stejný efekt jako provedení příslušné elementární řádkové úpravy přímo na matici A.

Tvrzení 8. *Elementární řádková úprava rozšířené matice soustavy $(A|\mathbf{b})$ nemění množinu řešení. Pokud B je matice této EŘÚ, pak upravená matice je $(BA|B\mathbf{b})$.*

Důkaz. Každou EŘÚ lze vrátit zpět opět pomocí EŘÚ, například přičtení r-násobku i-tého řádku do j-tého se odstraní přičtením $-r$-násobku i-tého řádku do j-tého. Stačí tedy pro jednotlivé typy EŘÚ ukázat, že každé řešení $(A|\mathbf{b})$ je také řešením upravené soustavy. Tvar upravené matice vyplývá z toho, že součin matic je definován po sloupcích. Pokud tedy vektor $\mathbf{x}$ splňuje rovnost $A\mathbf{x} = \mathbf{b}$, pak splňuje i $BA\mathbf{x} = B\mathbf{b}$, tedy soustavu s maticí $(BA|B\mathbf{b})$. □

Řešení soustavy lineárních rovnic spočívá v převedení její rozšířené matice posloupností elementárních řádkových úprav do jednoduššího, tzv. odstupňovaného tvaru. Množina řešení se tím nezmění, ale je možné ji z odstupňovaného tvaru snáze získat.

Definice 6. První nenulový prvek každého řádku matice se nazývá *pivot.* Matice A je v *odstupňovaném tvaru*, pakliže je buď nulová, nebo v ní všechny nulové řádky následují až po všech nenulových a zároveň každý pivot nenulového řádku kromě prvního má vyšší sloupcový index než pivot předchozího řádku. Sloupce, v nichž jsou v odstupňovaném tvaru pivoty, se nazývají *pivotní sloupce.* Matice je v *redukovaném odstupňovaném tvaru*, pokud jsou navíc všechny pivoty rovny 1 a všechny ostatní elementy pivotních sloupců jsou rovny 0.



Na obrázku jsou a_{1j_1} až a_{rj_r} pivoty, nalevo od nich a pod nimi jsou v matici nuly, napravo od nich a nad nimi může být obecně cokoliv. V redukovaném odstupňovaném tvaru jsou na místě tmavých obdélníků nuly.

Tvrzení 9. *Je-li rozšířená matice soustavy $(A|\mathbf{b})$ v odstupňovaném tvaru, pak má soustava rovnic řešení, právě když sloupec pravých stran není pivotní.*

Důkaz. Předpokládejme nejprve, že matice je v redukovaném odstupňovaném tvaru, že rozšířená matice má r nenulových řádků a že pivoty mají sloupcové indexy $j_1 < j_2 < \ldots < j_r$. Pokud je sloupec pravých stran pivotní, pak je součástí soustavy také rovnice, která má na levé straně 0 a na pravé nenulový pivot b_r. Taková rovnice nemá řešení a tedy ani celá soustava. Pokud sloupec pravých stran pivotní není, pak je jedním z řešení vektor $\mathbf{x}$, v němž $\forall k \in \{1, \ldots, r\}$ je $x_{j_k} = b_k$ a ostatní složky $\mathbf{x}$ jsou nulové.

Je-li matice pouze v odstupňovaném tvaru, lze ji do redukovaného odstupňovaného tvaru převést pomocí EŘÚ, podle tvrzení 8 se tím množina řešení nezmění. Konkrétní posloupnost EŘÚ sestává z vynásobení r-tého řádku číslem $(a_{rj_r})^{-1}$, přičtení $(-a_{ij_r})$ násobku r-tého řádku do i-tého řádku pro všechna $i < r$, a následně opakování stejné procedury pro r nahrazené postupně $r-1$, $r-2$ až 1. $\square$

Obecné řešení SLR v odstupňovaném tvaru získáme, pokud položíme každou nepivotní složku rovnu nějakému reálnému parametru a zpětným dosazením dopočítáme složky pivotní. Nejlépe je to vidět na maticích v redukovaném odstupňovaném tvaru:

$$\left(\begin{array}{cccccc|c} 1 & 0 & -2 & -2 & 0 & -3 & 9 \\ 0 & 1 & 3 & -1 & 0 & -3 & 2 \\ 0 & 0 & 0 & 0 & 1 & -2 & 4 \end{array}\right)$$

Zvolíme $x_6 = t$, $x_4 = s$ a $x_3 = r$ a dopočteme

$$\begin{aligned} x_5 &= 4 + 2t \\ x_2 &= 2 - 3r + s + 3t \\ x_1 &= 9 + 2r + 2s + 3t \end{aligned}$$

Přepsáno do vektorového tvaru

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{pmatrix} = \begin{pmatrix} 9 \\ 2 \\ 0 \\ 0 \\ 4 \\ 0 \end{pmatrix} + r \begin{pmatrix} 2 \\ -3 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + s \begin{pmatrix} 2 \\ 1 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} + t \begin{pmatrix} 3 \\ 3 \\ 0 \\ 0 \\ 2 \\ 1 \end{pmatrix}$$

Všimněme si, že vektor $\mathbf{b}$ pravých stran nemá žádný vliv na koeficienty u parametrů r, s, t. Řešení SLR s maticí $(A|\mathbf{o})$ (*příslušné homogenní soustavy*) je tedy

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{pmatrix} = r \begin{pmatrix} 2 \\ -3 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + s \begin{pmatrix} 2 \\ 1 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} + t \begin{pmatrix} 3 \\ 3 \\ 0 \\ 0 \\ 2 \\ 1 \end{pmatrix}$$

2. Řešení soustav a podprostory

Vztah mezi řešením nehomogenní soustavy rovnic a příslušné soustavy homogenní se netýká jen soustav v odstupňovaném tvaru:

Tvrzení 10. *Nechť $\mathbf{x}_P$ je nějaké řešení SLR $(A|\mathbf{b})$. Pak vektor $\mathbf{x}_P + \mathbf{x}_H$ je řešením soustavy $(A|\mathbf{b})$ právě tehdy, když je $\mathbf{x}_H$ řešením příslušné homogenní soustavy.*

Důkaz. Pokud $\mathbf{x}_P + \mathbf{x}_H$ je řešením soustavy, pak $A(\mathbf{x}_P + \mathbf{x}_H) = \mathbf{b}$. Z distributivity maticového násobení plyne

$$A(\mathbf{x}_P + \mathbf{x}_H) = A\mathbf{x}_P + A\mathbf{x}_H = \mathbf{b} + A\mathbf{x}_H$$

Tedy $A\mathbf{x}_H = \mathbf{o}$. Pokud naopak $A\mathbf{x}_H = \mathbf{o}$, pak dosazením do rovnosti výše dostáváme $A(\mathbf{x}_P + \mathbf{x}_H) = \mathbf{b}$. □

Vektoru $\mathbf{x}_P$ říkáme *partikulární řešení nehomogenní soustavy*. Množině všech řešení homogenní soustavy $(A|\mathbf{o})$ pak *jádro matice* A, značíme $\operatorname{Ker} A$. Množina řešení nehomogenní soustavy se v duchu tvrzení zapisuje ve tvaru $\mathbf{x}_P + \operatorname{Ker} A$.

Jádro matice má význačné vlastnosti, které si zaslouží zavedení nového pojmu:

Definice 7. Neprázdná množina $W \subset \mathbb{R}^n$, která s každými vektory $\mathbf{x}, \mathbf{y} \in W$ obsahuje i všechny jejich lineární kombinace $r\mathbf{x}+s\mathbf{y}$, se nazývá *podprostor*. Množina tvaru $\mathbf{u} + W$, kde $\mathbf{u} \in \mathbb{R}^n$ a W je podprostor v $\mathbb{R}^n$, se nazývá *afinní podprostor*.

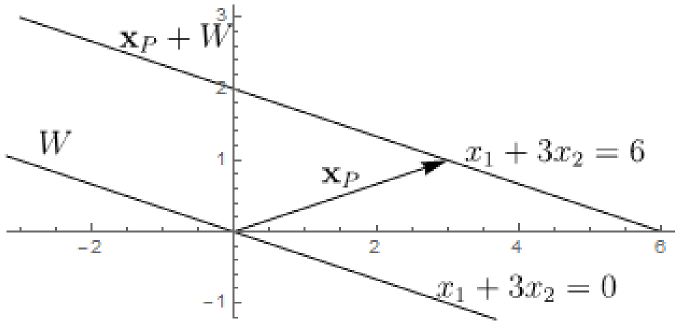
Tvrzení 11. *Jádro každé matice A typu $m \times n$ je podprostor v $\mathbb{R}^n$.*

Důkaz. Pokud $\mathbf{x}, \mathbf{y} \in \operatorname{Ker} A$, pak $A\mathbf{x} = A\mathbf{y} = \mathbf{o}$. Pak ale

$$A(r\mathbf{x} + s\mathbf{y}) = rA\mathbf{x} + sA\mathbf{y} = r\mathbf{o} + s\mathbf{o} = \mathbf{o},$$

tedy i $r\mathbf{x} + s\mathbf{y} \in \operatorname{Ker} A$. □

Každý podprostor musí obsahovat nulový vektor, stačí zvolit $r = s = 0$. Opačná implikace, tedy že podmnožina $\mathbb{R}^n$ obsahující nulový vektor je nutně podprostorem, neplatí (najděte protipříklad!). Afinní podprostor $\mathbf{x} + W$ obsahující $\mathbf{o}$ ovšem už podprostorem být musí. To nastává právě tehdy, když $\mathbf{x} \in W$. Homogenní a nehomogenní soustavy se tedy liší právě tím, že řešením prvních je vždy podprostor, řešením druhých nikdy není podprostor, pouze afinní podprostor:



3. Gaussova eliminace

Každou matici A typu $m \times n$ lze převést do (redukovaného) odstupňovaného tvaru postupem zvaným *Gaussova eliminace*:

(1) Najdeme první nenulový sloupec, jeho index označíme j. Pokud neexistuje, je matice nulová, tedy v redukovaném odstupňovaném tvaru.
(2) Pokud je $a_{1j} = 0$, pak prohodíme 1. řádek s libovolným řádkem i, v němž $a_{ij} \neq 0$.
(3) Pro každé $i = 2, 3, \ldots, m$ přičteme $(-a_{ij}/a_{1j})$-násobek prvního řádku do i-tého řádku.
(4) Opakujeme postup pro matici bez prvního řádku. Proces se zastaví buď v bodě 1, nebo tím, že dojdou nenulové řádky. Pokud v každém průchodu bodu 1 zaznamenáme index j do proměnné j_k, dostaneme posloupnost $j_1 < j_2 < \ldots j_r$ indexů pivotních sloupců, přičemž řádky $r+1$ až m jsou nulové.
(5) Pro každé $i = 1, \ldots, r$ vynásobíme i-tý řádek číslem $1/a_{ij_i}$ a poté pro každé $k < i$ přičteme jeho $(-a_{kj_i})$-násobek do k-tého řádku.

Příklad 6. Ukažme si Gaussovu eliminaci na příkladě rozšířené matice SLR:

$$\left(\begin{array}{rrrrr|r} 0 & 3 & -6 & 6 & 4 & -5 \\ 3 & -7 & 8 & -5 & 8 & 1 \\ 3 & -9 & 12 & -9 & 6 & 3 \end{array}\right) \sim \left(\begin{array}{rrrrr|r} 3 & -9 & 12 & -9 & 6 & 3 \\ 3 & -7 & 8 & -5 & 8 & 1 \\ 0 & 3 & -6 & 6 & 4 & -5 \end{array}\right) \sim$$
$$\left(\begin{array}{rrrrr|r} 3 & -9 & 12 & -9 & 6 & 3 \\ 0 & 2 & -4 & 4 & 2 & -2 \\ 0 & 3 & -6 & 6 & 4 & -5 \end{array}\right) \sim \left(\begin{array}{rrrrr|r} 3 & -9 & 12 & -9 & 6 & 3 \\ 0 & 2 & -4 & 4 & 2 & -2 \\ 0 & 0 & 0 & 0 & 1 & -2 \end{array}\right)$$

První úpravě, tedy hledání vhodného pivota, se říká *pivotace*. Po dvou průchodech body 1,2,3,4 jsme dospěli zpět do 1 a nemáme už žádné nenulové řádky. Bod 5, tedy přechod do redukovaného tvaru, je efektivnější provádět odspodu:

$$\left(\begin{array}{rrrrr|r} 3 & -9 & 12 & -9 & 0 & 15 \\ 0 & 2 & -4 & 4 & 0 & 2 \\ 0 & 0 & 0 & 0 & 1 & -2 \end{array}\right) \sim \left(\begin{array}{rrrrr|r} 1 & 0 & -2 & 3 & 0 & 8 \\ 0 & 1 & -2 & 2 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & -2 \end{array}\right)$$

Množina řešení soustavy má tvar

$$\mathbf{x}_P + \operatorname{Ker} A = \begin{pmatrix} 8 \\ 1 \\ 0 \\ 0 \\ -2 \end{pmatrix} + \left\{ s \begin{pmatrix} 2 \\ 2 \\ 1 \\ 0 \\ 0 \end{pmatrix} + t \begin{pmatrix} -3 \\ -2 \\ 0 \\ 1 \\ 0 \end{pmatrix} \middle| s, t \in \mathbb{R} \right\}$$

Množina všech lineárních kombinací nějaké množiny vektorů $M \subset \mathbb{R}^n$ se značí $\langle M \rangle$ a říká se jí *lineární obal*. S tímto značením můžeme řešení zapsat poněkud úsporněji jako

$$(8,1,0,0,-2)^T + \langle (2,2,1,0,0)^T, (-3,-2,0,1,0)^T \rangle$$

Často se v zápise vynechává i symbol transponování.

První 4 kroky Gaussovy eliminace převedou do odstupňovaného tvaru libovolnou matici. To plyne jednoduše matematickou indukcí podle počtu nenulových řádků v matici. Je-li jich nula, pak je matice zjevně v odstupňovaném tvaru. Pokud je jich $k > 0$, pak po provedení prvních 3 kroků získáme sloupcový index j_1 prvního pivotu, který je menší než sloupcový index všech nenulových prvků všech ostatních řádků. Z indukčního předpokladu plyne, že matice vzniklá vynecháním prvního řádku se prvními 4 kroky Gaussovy eliminace převede do odstupňovaného tvaru, přičemž sloupcové indexy pivotů $j_2 < \ldots < j_r$ musí být větší než j_1. Výsledná matice je tedy v odstupňovaném tvaru a bod 5 ji převede do redukovaného odstupňovaného tvaru.

4. Izomorfismus

Následující pojmosloví se užívá pro libovolná zobrazení, nikoli nutně lineární. Zobrazení $f : X \to Y$ se nazývá *prosté* (*injektivní*), pokud má každý prvek z Y nejvýše jeden vzor, tedy pokud z $f(a) = f(b)$ plyne $a = b$. Nazývá se *zobrazení na* (nebo též *surjektivní zobrazení*), pokud má každý prvek z Y alespoň jeden vzor. Je-li zobrazení zároveň prosté i na, říká se mu *vzájemně jednoznačné* (*bijektivní*). To znamená, že každý prvek z Y má v zobrazení f právě jeden vzor, existuje tedy inverzní zobrazení $f^{-1} : Y \to X$.

V případě, že f je navíc lineární, používají se pojmy *monomorfismus*, *epimorfismus* a *izomorfismus*. Z lineárních zobrazení, kterými jsme se zabývali, jsou zrcadlení $Z_{\mathbf{x}^\perp}$, dilatace D_λ pro $\lambda \neq 0$ i rotace izomorfismy (jak vypadají inverzní zobrazení?. Projekce $P_{\mathbf{x}} : \mathbb{R}^n \to \mathbb{R}^n$ není prosté zobrazení, protože libovolný vektor z nadroviny $\mathbf{x}^\perp$ se zobrazí na $\mathbf{o}$, a není ani zobrazení na, protože $P_{\mathbf{x}}(\mathbf{y})$ leží vždy v $\langle \mathbf{x} \rangle$. Podobně pro $P_{\mathbf{x}^\perp}$.

Definice 8. Nechť $f : X \to Y$ je zobrazení. Pak množina všech prvků $p \in Y$, pro které existuje $a \in X$ takové, že $f(a) = p$, se nazývá *obraz* zobrazení f a značí se $\operatorname{Im} f$.

Obrazu se někdy říká také *obor hodnot*. Zobrazení f je na, právě když $\operatorname{Im} f = Y$.

Zobrazení $F_A : \mathbb{R}^n \to \mathbb{R}^m$ je prosté právě tehdy, když rovnice $A\mathbf{x} = \mathbf{b}$ má pro všechna $\mathbf{b}$ nejvýše jedno řešení, tedy když $\operatorname{Ker} A = \{\mathbf{o}\}$. To nastává právě tehdy, když jsou po převodu A na odstupňovaný tvar všechny sloupce pivotní, neboli přejde-li A v redukovaném odstupňovaném tvaru na jednotkovou matici, případně doplněnou o nějaké nulové řádky. Soustava $A\mathbf{x} = \mathbf{b}$ má řešení, pakliže $\mathbf{b}$ lze zapsat jako

$$(\mathbf{a}_1 | \ldots | \mathbf{a}_n)\mathbf{x} = x_1\mathbf{a}_1 + \ldots x_n\mathbf{a}_n,$$

neboli $\operatorname{Im} A = \langle \mathbf{a}_1, \ldots, \mathbf{a}_n \rangle$. Zobrazení F_A je tedy na, právě když $\langle \mathbf{a}_1, \ldots, \mathbf{a}_n \rangle = \mathbb{R}^m$. Lineárnímu obalu sloupců se říká *sloupcový prostor* matice A (analogicky definujeme *řádkový prostor*, ten je roven $\operatorname{Im} A^T$). Rovnost $\langle \mathbf{a}_1, \ldots, \mathbf{a}_n \rangle = \mathbb{R}^m$ formulujeme také tak, že sloupce matice A prostor $\mathbb{R}^m$ *generují*.

Tvrzení 12. *Zobrazení $f_A : \mathbb{R}^n \to \mathbb{R}^m$ je bijektivní, právě když A je čtvercová a existuje matice X, pro kterou $AX = E$.*

Důkaz. Je-li f_A bijektivní, pak pro každou pravou stranu $\mathbf{b}$ má SLR $A\mathbf{x} = \mathbf{b}$ právě jedno řešení. Nemůže tedy být $m < n$, protože pak by po převodu na odstupňovaný tvar nemohly být všechny sloupce pivotní a tedy $\operatorname{Ker} A \neq \{\mathbf{o}\}$. Nemůže být ani $m > n$, protože pak by bylo možné najít pravou stranu $\mathbf{b}$ takovou, že po eliminaci matice $(A|\mathbf{b})$ bude sloupec pravých stran pivotní. Tedy A je čtvercová. Pro pravou stranu $\mathbf{e}_i$ označme $\mathbf{x}_i$ řešení soustavy $A\mathbf{x} = \mathbf{e}_i$. Pak $X = (\mathbf{x}_1|\dots|\mathbf{x}_n)$ splňuje $AX = E$. Tím je dokázána první implikace.

Dokažme nyní druhou. Pokud existuje matice X splňující $AX = E$, pak $\mathbf{x} := X\mathbf{b}$ je řešením soustavy rovnic $A\mathbf{x} = \mathbf{b}$. Tedy f_A je na. Uvažujme eliminaci rozšířené matice $(A|E)$ elementárními úpravami s maticemi $B_1, \dots, B_k$. Pak

$$(A|E) \sim (B_k \dots B_1 A | B_k \dots B_1 E) = (C | B_k \dots B_1),$$

kde matice C je v redukovaném odstupňovaném tvaru. Pokud by měla nulový řádek, neplatilo by, že $(A|\mathbf{b})$ má řešení pro každé $\mathbf{b}$. Protože je čtvercová, musí tedy mít n pivotních sloupců, neboli být jednotkovou maticí. Pak ale j-tý sloupec součinu $B_k \dots B_1$ je jediným řešením soustavy rovnic $A\mathbf{x} = \mathbf{e}_j$, jejímž řešením je i j-tý sloupec X. Tedy $B_k \dots B_1 = X$ a $XA = E$. Tedy X je inverzní matice k A a tudíž i f_X inverzní zobrazení k f_A. □

Připomeňme, že čtvercovou matici A, která má inverzní matici, nazýváme *regulární*. Z tvrzení a jeho důkazu plyne, že regulární jsou právě ty matice, pro které je f_A bijektivní, a že stačí ověřovat jen podmínku $AX = E$, tedy inverznost zprava.

KAPITOLA 4

Vektorové prostory

V prvních třech přednáškách jsme se zabývali vektory v $\mathbb{R}^n$ a setkali se při tom mimo jiné s pojmy lineární kombinace, (afinního) podprostoru, lineárního zobrazení a kanonické báze. Tyto pojmy je možné a užitečné zobecnit na další typy objektů, které lze mezi sebou sčítat a násobit je skalárem. Například levá strana rovnosti

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

je lineární kombinací 2×2 matic, levá strana rovnosti

$$(3x+1) + 2(-x^3 + x^2 - x + 2) + (2x^3 - 1) = 2x^2 + x + 4$$

je lineární kombinací polynomů. V obou případech rovnosti říkají, že matice/polynom na pravé straně je v podprostoru generovaném maticemi/polynomy na levé straně. Označíme-li

$$K_M := \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}, K_P := \{1, x, x^2, x^3\},$$

můžeme každou matici, resp. polynom stupně nejvýše 3 jednoznačně zapsat jako lineární kombinaci prvků množiny K_M, resp. K_P. Tyto množiny tedy pro matice/polynomy hrají podobnou roli jako kanonická báze $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ pro $\mathbb{R}^n$. Rovnost výše pak intuitivně vyjadřuje totéž jako vztah

$$\begin{pmatrix} 0 \\ 0 \\ 3 \\ 1 \end{pmatrix} + 2\begin{pmatrix} -1 \\ 1 \\ -1 \\ 2 \end{pmatrix} + \begin{pmatrix} 2 \\ 0 \\ 0 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \\ 1 \\ 4 \end{pmatrix},$$

pro vektory z $\mathbb{R}^4$.

Podobně můžeme zacházet s mnoha dalšími objekty, například komplexními čísly, maticemi $m \times n$, spojitými funkcemi na daném intervalu, nekonečnými posloupnostmi čísel a mnoha dalšími. Tyto podobnosti vybízejí k tomu, abychom potřebné vlastnosti vektorů abstrahovali a všechny související pojmy zaváděli naráz a obecně pro tyto abstraktní vektory. Dokázaná tvrzení pak budou mít také obecnou platnost. Ještě předtím musíme ale zavést pojem tělesa, což je množina, jejíž prvky jsou abstrakcí pojmu skaláru, a pojem grupy, který se v definici tělesa i vektorového prostoru vyskytuje a který má i sám o sobě pro matematiku a fyziku velkou důležitost.

1. Grupa a těleso

Definice 9 (Grupa). Nechť $\circ : G \times G \to G$ je binární operace na množině G. Pak *grupou* nazýváme dvojici $(G, \circ)$, která splňuje

(1) $\forall a, b, c \in G : (a \circ b) \circ c = a \circ (b \circ c)$ (asociativita)
(2) $\exists e \in G : \forall a \in G : a \circ e = e \circ a = a$ (neutrální prvek)
(3) $\forall a \in G : \exists a^{-1} \in G : a \circ a^{-1} = a^{-1} \circ a = e$ (inverzní prvky)

Pokud navíc $\forall a, b \in G : a \circ b = b \circ a$, pak je $(G, \circ)$ *komutativní grupa*.

Pod označením $\circ$ se může skrývat jakákoli binární operace, nemusí jít nutně o skládání zobrazení. Zároveň mnoho příkladů grup jsou grupy zobrazení s operací jejich skládání. Například množina všech rotací okolo počátku v rovině $\mathbb{R}^2$

$$(\{R_\phi : \mathbb{R}^2 \to \mathbb{R}^2 | \phi \in \langle 0, 2\pi)\}, \circ)$$

je uzavřená na skládání (tedy složení dvou rotací je zase rotace), obsahuje neutrální prvek $R_0 \equiv \mathrm{Id}$ i inverzní prvek $R_{-\phi}$ ke každému R_ϕ. Asociativita a v tomto případě i komutativita jsou splněny rovněž. Dalším typickým příkladem, který můžeme formulovat pomocí značení zavedeného v minulých kapitolách, je grupa

$$(\{\mathrm{Id}, R_{2\pi/3}, R_{-2\pi/3}, Z_{(1,0)^\perp}, Z_{(1,\sqrt{3})^\perp}, Z_{(-1,\sqrt{3})^\perp}\}, \circ)$$

všech zobrazení, která zachovávají rovnostranný trojúhelník s těžištěm v počátku a výškou orientovanou ve směru druhé souřadné osy. Podobné grupy symetrií geometrických útvarů v rovině a prostoru jsou základem disciplín jako krystalografie, spektroskopie, kvantová mechanika a dalších.

Pro nás jsou nyní důležitější grupy založené na množině čísel s nějakou aritmetickou operací. V případě sčítání to mohou být třeba grupy $(\mathbb{Z}, +)$, $(\mathbb{R}, +)$, $(\mathbb{C}, +)$ (proč ne $(\mathbb{N}, +)$?) a grupami jsou také množiny vektorů $(\mathbb{R}^n, +)$ a matic $(\mathbb{R}^{m\times n}, +)$ se sčítáním zavedeným v minulých kapitolách. Neutrálním prvkem je zde 0, inverzním prvkem k x je opačné číslo $-x$.

Příklady grup s operací násobení jsou třeba $(\mathbb{Q} \setminus \{0\}, \cdot)$, $(\mathbb{R} \setminus \{0\}, \cdot)$ nebo $(\{z \in \mathbb{C}|\ |z| = 1\}, \cdot)$. Neutrálním prvkem je zjevně 1, inverzním prvkem k a je $\frac{1}{a}$. Proto také v žádné z množin nemůže být 0.

Množiny matic typu 2×2 s operací násobení matic nám dávají mnoho příkladů nekomutativních grup. Například grupa

$$\left(\left\{\begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix} | \phi \in \langle 0, 2\pi) \right\}, \circ\right)$$

odpovídající maticím rotací komutativní je, ale pokud zkonstruujeme analogicky grupu matic symetrií rovnostranného trojúhelníka, zjistíme, že matice rotací a zrcadlení spolu nekomutují.

Definice 10 (Komutativní těleso). Množina T se nazývá *komutativní těleso*, pokud jsou na ní definovány dvě binární operace $+$ a $\cdot$, splňující

(1) $(T, +)$ je komutativní grupa.
(2) Označíme-li neutrální prvek $(T, +)$ symbolem 0, pak $(T \setminus \{0\}, \cdot)$ je komutativní grupa.
(3) $\forall a, b, c \in T : a \cdot (b + c) = a \cdot b + a \cdot c$ (distributivita)

Název „těleso“ nijak nesouvisí s tělesy geometrickými, je jen překladem slova Zahlenkörper z němčiny. Klasickými příklady těles jsou množiny $\mathbb{Q}$, $\mathbb{R}$ nebo $\mathbb{C}$ s operacemi sčítání a násobení. V algebře a informatice nacházejí uplatnění *konečná tělesa*, jejichž základním příkladem je $\mathbb{Z}_p = \{0, 1, \ldots, p - 1\}$ s operacemi sčítání a násobení modulo prvočíslo p. Tělesa jsou také klíčová pro teorii řešení algebraických rovnic vyšších stupňů, s níž přišel Évariste Galois v roce 1832. Většina jeho matematického dědictví pochází z dopisu, který napsal dva dny před tím, než zahynul v souboji.

Pokud nepožadujeme, aby bylo násobení komutativní, zachováme ostatní axiomy a doplníme pravostrannou distributivitu $(b + c) \cdot a = b \cdot a + c \cdot a$, pak definici splňují např. tzv. kvaterniony, používané pro elegantní popis rotací v $\mathbb{R}^3$.

Prvky T hrají v lineární algebře roli skalárů. Většina textů buduje lineární algebru nad obecnými tělesy, což znamená strávit nějaký čas odvozováním důsledků definice tělesa a zaváděním některých pojmů, zejména tzv. charakteristiky tělesa, která zhruba řečeno pomáhá od sebe odlišit tělesa s nekonečným počtem prvků jako

$\mathbb{R}$ nebo $\mathbb{C}$ a tělesa konečná. Protože počítání s konečnými tělesy není v matematické fyzice příliš často k vidění, omezíme na T rovné buď $\mathbb{R}$ nebo $\mathbb{C}$ a budeme místo pojmu (komutativní) těleso používat pojem *množina skalárů*. Standardně budeme množinu skalárů značit $\mathbb{F}$, což vychází z výrazu *field*, který se pro komutativní těleso používá v angličtině.

2. Vektorový prostor

DEFINICE 11 (Vektorový prostor). Nechť $\mathbb{F}$ je množina skalárů. Množinu V nazveme *vektorovým prostorem nad* $\mathbb{F}$, pokud je na ní definována operace *sčítání vektorů*

$$+ : V \times V \to V$$

taková, že $(V, +)$ je komutativní grupa, a operace *násobení skalárem*

$$\cdot : \mathbb{F} \times V \to V,$$

jejíž výsledek pro $r \in \mathbb{F}$ a $v \in V$ označíme rv, a tyto operace splňují $\forall u, v \in V$ a $\forall r, s \in \mathbb{F}$

(1) $1v = v$ (násobení jednotkou)
(2) $r(sv) = (rs)v$ (asociativita)
(3) $(r + s)v = rv + sv$ (distributivita sčítání skalárů)
(4) $r(u + v) = ru + rv$ (distributivita sčítání vektorů)

Neutrální prvek ve $(V, +)$ značíme o (*nulový vektor*) a inverzní prvek k $v \in V$ značíme $-v$ (*opačný vektor*).

TVRZENÍ 13. *Nechť V je vektorový prostor nad $\mathbb{F}$. Pak $\forall v \in V$, $\forall r \in \mathbb{F}$ platí*

(1) $0v = o$
(2) $(-1)v = -v$
(3) $ro = o$

DŮKAZ. Pro libovolný vektor $v \in V$ platí

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

kde jsme použili nejprve vlastnost násobení jednotkou z definice vektorového prostoru, pak vlastnost nuly jako neutrálního prvku v množině skalárů (tělese), pak distributivitu sčítání skalárů a nakonec znovu násobení jednotkou. Protože $(V, +)$ je grupa, musí mít v opačný vektor $-v$ a s užitím asociativity sčítání v grupě plyne, že

$$o = (-v) + v = (-v) + (v + 0.v) = ((-v) + v) + 0.v = o + 0.v = 0.v,$$

čímž je dokázáno první tvrzení. Zbylá dvě necháváme čtenáři za cvičení. □

Základním příkladem vektorového prostoru je $\mathbb{F}^n$ nad $\mathbb{F}$, tedy množina všech uspořádaných n-tic prvků z $\mathbb{F}$, ve které se sčítá a násobí skalárem po složkách. Doposud jsme vektorem rozuměli prvek $\mathbb{R}^n$ a zapisovali ho jako sloupec čísel. Od nynějška pro nás slovo vektor znamená prvek nějakého vektorového prostoru. Vektorům v původním smyslu budeme od nynějška říkat *aritmetické vektory* a vektorovému prostoru $\mathbb{F}^n$ *aritmetický vektorový prostor*.

Symbolem $\mathbb{F}^{m \times n}$ označujeme vektorový prostor všech $m \times n$ matic s elementy z $\mathbb{F}$. Podobně jako u aritmetického vektoru se snadno ověří, že je to vektorový prostor nad $\mathbb{F}$ s operacemi sčítání matic a násobení matice skalárem, tak jak jsme je dříve zavedli.

Množina $F(X, \mathbb{F})$ všech funkcí z množiny X do $\mathbb{F}$ s operacemi

$$(f + g)(x) := f(x) + g(x)$$
$$(rf)(x) := rf(x),$$

je vektorový prostor nad $\mathbb{F}$. Jaký je vlastně význam tohoto předpisu? První řádka definuje funkci $f+g$ její hodnotou v bodě x, a to tak, že pro libovolné $x \in X$ je tato hodnota součtem hodnot funkcí f a g v x. Druhá obdobně definuje násobek funkce. Pro $X = \mathbb{R}$ to dává prostor všech funkcí jedné reálné proměnné, pro $X = \mathbb{N}$ prostor všech nekonečných posloupností, pro $X = \{1, \dots, n\}$ je to prostor všech n-prvkových posloupností, které se sčítají po složkách, což vlastně není nic jiného než aritmetický vektorový prostor $\mathbb{F}^n$.

Množinu $\mathbb{C}$ můžeme chápat jako vektorový prostor nad sebou samou, nebo jako vektorový prostor nad $\mathbb{R}$. Je v tom rozdíl, protože v prvním případě je každý nenulový vektor násobkem každého jiného nenulového vektoru, v druhém ale 1 a i nejsou skalárním násobkem jeden druhého. Intuitivně se tedy první verze podobá více přímce a druhá více rovině. Analogicky má každý vektorový prostor V nad $\mathbb{C}$ svého „dvojníka“, ve kterém je dovoleno násobit jen reálnými čísly. Pokud ale v dalším napíšeme $\mathbb{C}^n$, rozumíme tím vždy vektorový prostor nad $\mathbb{C}$.

Většina dalších zajímavých příkladů vektorových prostorů vzniká jako podprostory jiných vektorových prostorů. Obecná definice podprostoru se prakticky neliší od té, kterou už jsme měli v $\mathbb{R}^n$:

Definice 12. Nechť V je vektorový prostor nad $\mathbb{F}$ a W je neprázdná podmnožina V taková, že $\forall v, w \in W$, $\forall r, s \in \mathbb{F}$ platí $rv + sw \in W$. Pak nazýváme W *podprostorem* vektorového prostoru V, značíme $W \leq V$.

Tvrzení 14. *Neprázdná podmnožina $W \subset V$ je podprostorem ve V, právě když je uzavřená na součty a na násobení libovolným skalárem.*

Důkaz. Pokud je W uzavřená na součty a na násobení skalárem, pak s každými vektory $v, w \in W$ jsou ve W i rv, sw a $rv + sw$, pro libovolné $r, s \in \mathbb{F}$. Pro důkaz opačné implikace stačí volit $r = s = 1$, resp. r libovolné a $s = 0$. $\square$

Každý vektorový prostor V má dva tzv. *triviální podprostory*, $0 := \{o\}$ (*nulový podprostor*) a sebe sama. Symbol 0 se užívá i pro vektorový prostor obsahující jediný prvek, který samozřejmě musí být totožný s neutrálním prvkem v tomto prostoru, tedy nulovým vektorem.

Podprostor W nějakého vektorového prostoru V je opět vektorový prostor. Tím myslíme, že splňuje axiomy z definice vektorového prostoru. Pro ověření je klíčové, že nám definice podprostoru zaručuje uzavřenost operací, tedy že všechny součty a násobky skalárem, které se v axiomech vyskytují, jsou opět prvky W, a to včetně neutrálního prvku a opačných prvků, které díky prvnímu a druhému bodu tvrzení 13 vzniknou také jako násobky prvku z W skalárem.

Definice 13. Nechť V je vektorový prostor nad $\mathbb{F}$ a $r_1, \dots, r_n \in \mathbb{F}$, $v_1, \dots, v_n \in V$. Výraz $\sum_{i=1}^n r_i v_i$ se nazývá *lineární kombinace* vektorů v_i s koeficienty r_i. Pro neprázdnou $M \subset V$ značí $\langle M \rangle$ množinu všech lineárních kombinací prvků M, neboli její *lineární obal*. Pro $M = \emptyset \subset V$ definujeme jako její lineární obal nulový podprostor $0 \leq V$.

Tvrzení 15. *Nechť V je vektorový prostor nad $\mathbb{F}$ a $M \subset V$. Pak $\langle M \rangle \leq V$.*

Důkaz. Stačí si uvědomit, že součet lineárních kombinací prvků M je opět lineární kombinace prvků M a podobně i pro násobek skalárem. $\square$

Definice 14. Nechť V je vektorový prostor nad $\mathbb{F}$, $W \leq V$, $v \in V$. Množinu $v + W$, jejímiž prvky jsou všechny součty vektoru v s nějakým prvkem podprostoru W, nazýváme *afinní podprostor* ve V, podprostor W nazýváme *zaměřením* tohoto afinního podprostoru.

Afinní podprostor je podprostorem, právě když $v \in W$. Důkaz ponecháváme opět za cvičení.

3. Generování, lineární (ne)závislost, báze

Nechť W je vektorový prostor (který může a nemusí být podprostorem jiného vektorového prostoru). O množině $M \subset W$, pro kterou $\langle M \rangle = W$, říkáme, že prostor W *generuje*, případně, že M je *množinou generátorů* prostoru W.

Typicky existuje mnoho různých množin generujících stejný vektorový prostor:

$$\left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}, \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 3 \end{pmatrix} \right\}, \mathbb{R}^2, \mathbb{R}^2 \setminus \left\{ \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 3 \\ 4 \end{pmatrix} \right\}$$

jsou všechno množiny generátorů prostoru $\mathbb{R}^2$. Přirozenou otázkou je, jak najít pro daný podprostor množinu generátorů co nejmenší.

Definice 15. Lineární kombinace se nazývá *netriviální*, pokud má alespoň jeden koeficient nenulový. Podmožina M vektorového prostoru V nad $\mathbb{F}$ je *lineárně závislá*, pokud existují vektory $v_1, \dots, v_n \in M$ a jejich netriviální lineární kombinace $\sum_{i=1}^n r_i v_i$ se rovná nulovému vektoru o. V opačném případě je M *lineárně nezávislá*. Speciálně prázdnou množinu $\emptyset$ považujeme za lineárně nezávislou.

Příklad 7. Množina

$$\left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 3 \end{pmatrix} \right\}$$

je lineárně závislá, protože lze z vektorů vytvořit lineární kombinaci s koeficienty $-2, -3, 1$, která je netriviální a dává nulový vektor v $\mathbb{R}^2$. Množina

$$\left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$$

je lineárně nezávislá, protože lineární kombinace

$$r \begin{pmatrix} 1 \\ 0 \end{pmatrix} + s \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

může být rovna nulovému vektoru pouze pro $r = s = 0$.

Množina $\{1, i\} \subset \mathbb{C}$ je lineárně nezávislá v $\mathbb{C}$ nad $\mathbb{R}$, ale lineárně závislá v $\mathbb{C}$ nad $\mathbb{C}$.

Tvrzení 16. *Nechť V je vektorový prostor nad $\mathbb{F}$. Pak platí:*

(1) *$M \subset V$ mající alespoň dva prvky je lineárně závislá, právě když existuje $v \in M$, který lze vyjádřit jako lineární kombinaci ostatních prvků M.*
(2) *Nechť M generuje V. Pak M je lineárně závislá, právě když existuje vlastní podmnožina $N \subset M$, která generuje V.*

Důkaz. Nechť existuje netriviální lineární kombinace $\sum_{i=1}^n r_i v_i = o$, kde v_i jsou z M. Vektory v_i si můžeme libovolně přečíslovat, předpokládejme tedy, že $r_1 \neq 0$. Pak $v_1 = \sum_{i=2}^n \frac{-r_i}{r_1} v_i$, tedy v_1 lze vyjádřit jako lineární kombinaci ostatních vektorů. Naopak, pokud lze nějaký $v_1 \in M$ zapsat jako lineární kombinaci ostatních vektorů $\sum_{i=2}^n s_i v_i$, stačí položit $s_1 = -1$ a máme lineární kombinaci $\sum_{i=1}^n s_i v_i = o$ rovnou nulovému vektoru. Tím je dokázán první bod.

Pro druhý bod vyberme v M vektor v, který je dle bodu 1 lineární kombinací ostatních $v = \sum_{i=1}^k r_i v_i$. Definujme $N := M \setminus \{v\}$. Pak můžeme každý vektor u, který je lineární kombinací $u = \sum_{j=1}^n s_j u_j$ vektorů $u_j \in M$, zapsat jako lineární kombinaci vektorů z N. Pokud žádný z vektorů u_j není roven v, je to zřejmé. Pokud některý je roven v, stačí jej nahradit v lineární kombinaci sumou $\sum_{i=1}^k r_i v_i$ a u je pak vyjádřen jen pomocí vektorů z N. Naopak, pokud $N \subset M$, N i M obě generují V a $v \in M \setminus N$, pak lze v zapsat jako lineární kombinaci prvků N a tedy podle bodu 1 je M lineárně závislá. $\square$

Definice 16. Nechť V je vektorový prostor nad $\mathbb{F}$. Množina $M \subset V$, která generuje V a zároveň je lineárně nezávislá, se nazývá *báze vektorového prostoru V*.

Základním příkladem báze je kanonická báze $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ v $\mathbb{F}^n$. Je hned vidět, že je lineárně nezávislá a generuje $\mathbb{F}^n$.

Podle předchozího tvrzení je báze V taková množina generátorů V, která po odebrání libovolného vektoru už množinou generátorů V není.

Tvrzení 17. *Nechť V je vektorový prostor nad $\mathbb{F}$. Množina $M \subset V$ je bází V, právě když lze každý vektor z V vyjádřit jako lineární kombinací prvků M právě jedním způsobem.*

Důkaz. Pokud je M bází V, pak V i generuje, a tedy každý vektor z V lze vyjádřit jako lineární kombinaci prvků M alespoň jedním způsobem. Pokud by nějaký $v \in V$ měl dvojí vyjádření $\sum_{i=1}^{n} r_i v_i = \sum_{i=1}^{n} s_i v_i$, kde alespoň jeden koeficient $r_i \neq s_i$, pak by rozdíl těchto vyjádření byl netriviální lineární kombinací $\sum_{i=1}^{n} (r_i - s_i) v_i$ rovnou nulovému vektoru.

Je-li naopak každý vektor vyjádřen nejvýše jedním způsobem, platí to i pro nulový vektor, jenž je pomocí vektorů z M vyjádřen triviální lineární kombinací (či kombinacemi). Pak ale nemůže být vyjádřen zároveň i netriviální lineární kombinací, a tedy je M lineárně nezávislá. Zároveň je každý vektor vyjádřen alespoň jedním způsobem, tedy M generuje V. M je tedy bází V. $\square$

Tvrzení funguje i pro vektorový prostor 0, jehož jedinými dvěma podmnožinami jsou on sám, tedy $\{o\}$ a prázdná množina $\emptyset$ (je tedy třeba označení 0 a $\emptyset$ důsledně odlišovat). Množina $\{o\}$ obsahuje nulový vektor a je tedy lineárně závislá, bází tedy může být jedině prázdná množina. Jediný prvek prostoru 0, nulový vektor o, lze vyjádřit jako lineární kombinaci prázdné množiny právě jedním způsobem, a sice prázdnou sumou, která je definitoricky rovna nulovému prvku množiny, na níž sčítáme. Prázdná suma nebo prázdná lineární kombinace může být na první pohled trochu neintuitivní, ale jiná konvence by způsobovala řadu nepohodlných výjimek v dalších definicích, tvrzeních a důkazech.

Předpokládejme, že $M = \{v_1, \dots, v_n\}$ je báze vektorového prostoru V nad $\mathbb{F}$. Každému vektoru $v \in V$ lze podle předchozího tvrzení přiřadit jednoznačně n-tici $(r_1, \dots, r_n)$ prvků $\mathbb{F}$. Když tedy určíme nějaké pořadí prvků M, přiřazujeme vlastně vektoru v z V aritmetický vektor $\mathbf{r}$ z $\mathbb{F}^n$. Tomuto vektoru se říká *vektor souřadnic vektoru v vzhledem k bázi M* nebo též *reprezentace vektoru v vzhledem k bázi M*. Pokud například

$$M = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}$$

je báze $\mathbb{R}^{2\times 2}$, pak je matice $\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ reprezentována vzhledem k M vektorem $(1, 2, 3, 4)^T$. Počítání v rozličných vektorových prostorech můžeme takto převést na výpočty s vektory aritmetickými. Ale jak tyto výpočty ovlivňuje to, jakou bázi ve V si zvolíme? Budou mít vždy příslušné aritmetické vektory stejný počet složek? A dá se vždy najít báze s konečně mnoha prvky? Odpovědi nalezneme v příští kapitole.

KAPITOLA 5

Báze a dimenze

V minulé kapitole jsme definovali bázi jako takovou podmnožinu vektorového prostoru V, která je zároveň lineárně nezávislá a zároveň tento prostor generuje. Báze nemusí být nutně konečná množina.

Příklad 8. Množina $M := \{1, x, x^2, x^3, \dots, x^i, \dots\}$ ve vektorovém prostoru $P(x, \mathbb{R})$ všech reálných polynomů v proměnné x je jeho bází, neboť každý polynom $p(x) = \sum_{i=0}^{k} a_i x^i$ lze zapsat jednoznačně jako lineární kombinaci monomů x^i. Nějakou jinou konečnou bázi prostor $P(x, \mathbb{R})$ mít nemůže, protože v lineárním obalu konečné množiny polynomů N mohou být jen polynomy, jejichž stupeň není vyšší než nejvyšší stupeň v N zastoupený.

Definice 17. Vektorový prostor V nad $\mathbb{F}$ se nazývá *prostorem konečné dimenze*, pokud v něm existuje konečná báze.

Tvrzení 18. *Nechť V je vektorový prostor nad $\mathbb{F}$. Pak jsou následující tvrzení ekvivalentní:*

(1) *V je konečné dimenze.*
(2) *Ve V existuje konečná množina generátorů.*
(3) *Z každé množiny generátorů V lze vybrat konečnou bázi V.*

Důkaz. Implikace $1 \Rightarrow 2$ a $3 \Rightarrow 1$ jsou zřejmé, stačí tedy dokázat $2 \Rightarrow 3$. Máme-li ve V konečnou množinu generátorů M a nějakou (ne nutně konečnou) množinu generátorů N, pak lze každý prvek M vyjádřit jako lineární kombinaci prvků N. Lineární kombinace je suma obsahující pouze konečně mnoho sčítanců, množina M je také konečná. Existuje tedy konečná množina N', sestávající ze všech prvků N, u kterých je ve vyjádření některého prvku M nenulový koeficient. Libovolný prvek V je lineární kombinací prvků množiny M a lze jej tedy vyjádřit jako lineární kombinaci prvků N'. Podle druhé části tvrzení 16 je N' buď lineárně nezávislá, nebo z ní lze odebrat prvky, aniž by se změnil lineární obal. Tedy odebráním konečného počtu prvků lze nalézt podmnožinu $N'' \subset N'$, která je bází V. $\square$

Ze třetího bodu tvrzení 18 a druhého bodu tvrzení 16 plyne, že v prostoru V konečné dimenze musejí být všechny báze konečné. Dokážeme-li ještě, že všechny báze V musejí mít stejný počet prvků, budeme moci zadefinovat pojem dimenze V právě jako tento počet. Ještě před tím ale mírně redefinujme pojem báze, aby součástí definice bylo i uspořádání:

Definice 18. Nechť $V \neq 0$ je vektorový prostor nad $\mathbb{F}$ konečné dimenze, pak posloupnost $B = (v_1, \dots, v_n)$ nazveme *bází* V, pokud pro každý vektor $v \in V$ existuje jednoznačně určený aritmetický vektor $[v]^B := (r_1, r_2, \dots, r_n)^T$ takový, že $v = \sum_{i=1}^{n} r_i v_i$. Bází prostoru 0 je prázdná posloupnost. Vektor $[v]^B$ nazýváme *reprezentace vektoru v vzhledem k bázi B.*

Poznámka 2. Zavést bázi jako posloupnost bychom mohli i u některých prostorů nekonečné dimenze, jako je $P(x, \mathbb{R})$. Reprezentací polynomu $p(x) = \sum_{i=0}^{k} a_i x^i$ by pak nebyl aritmetický vektor, ale nekonečná posloupnost $(a_0, a_1, \dots, a_k, 0, 0, \dots)^T$.

Protože lineární kombinace jsou vždy konečné součty, bude mít každá taková reprezentace pouze konečný počet nenulových složek.

Zatím hovoříme pouze o „prostorech konečné dimenze“ a „prostorech nekonečné dimenze“, ale samotný pojem dimenze definován nemáme. K tomu potřebujeme následující velmi důležitou větu:

Věta 2 (O počtu prvků báze). *Všechny báze vektorového prostoru V konečné dimenze mají stejný počet prvků.*

Definice 19. Nechť V je vektorový prostor nad $\mathbb{F}$. Je-li V prostor konečné dimenze, pak definujeme nezáporné celé číslo $\dim V$ jako počet prvků libovolné báze V. Pokud není, píšeme $\dim V = \infty$. V obou případech nazýváme tuto hodnotu *dimenzí* vektorového prostoru V.

Příklady 1.

- $\dim \mathbb{F}^n = n$
- V $\mathbb{F}^{m\times n}$ je báze např. $(E_{ij}, i \in \{1,\ldots,m\}, j \in \{1,\ldots,n\})$, kde E_{ij} označuje matici s jednou jedničkou na pozici ij a nulami všude jinde. Tedy $\dim \mathbb{F}^{m\times n} = mn$.
- Posloupnost
$$\left(\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}i\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}, \begin{pmatrix}0\\i\end{pmatrix}\right)$$
je báze prostoru $\mathbb{C}^2$ nad $\mathbb{R}$. Obecně je dimenze $\mathbb{C}^n$ nad $\mathbb{R}$ rovna $2n$.
- Označme $U_n(\mathbb{F}) = \{A \in \mathbb{F}^{n\times n} | \forall i,j, i > j : a_{ij} = 0\}$ podprostor všech horních trojúhelníkových $n \times n$ matic. Pak $(E_{ij} | i,j \in \{1,\ldots,n\}, j \geq i)$ je bází $U_n(\mathbb{F})$. Tedy $\dim U_n(\mathbb{F}) = 1 + 2 + \ldots + n = \frac{n(n+1)}{2}$
- Označme $P^n(x,\mathbb{F})$ podprostor všech polynomů stupně nejvýše n v $P(x,\mathbb{F})$. Jeho bází je třeba $(1, x, x^2, \ldots, x^n)$, tedy $\dim P^n(x,\mathbb{F}) = n + 1$.

Poznámka 3. Z teorie množin je známo, že nekonečné množiny mohou být spočetné nebo nespočetné. Má-li vektorový prostor spočetnou bázi, pak je možné tuto bázi seřadit do posloupnosti a zacházet s bázemi a reprezentacemi podobně jako u prostoru $P(x,\mathbb{R})$ výše. Prostory, které spočetnou bázi nemají, typicky prostory funkcí, se studují jinými metodami. Obvykle se na nich nějak zavede pojem normy vektoru, který podobně jako na $\mathbb{R}^n$ umožňuje definovat vzdálenost vektorů. Se vzdáleností máme definovanou i konvergenci posloupnosti vektorů a tedy i (některé) nekonečné lineární kombinace. Například k popisu prostoru všech spojitých reálných funkcí na uzavřeném intervalu $[0, 2\pi]$ je možné použít množinu funkcí $\{1\} \cup \{\cos nx | n \in \mathbb{N}\} \cup \{\sin nx | n \in \mathbb{N}\}$, jejíž konvergentní nekonečné lineární kombinace se nazývají Fourierovými řadami a samotná množina pak Fourierovou bází, ačkoli to dle naší definice báze daného vektorového prostoru není.

V tomto kurzu se od nynějška budeme zabývat pouze vektorovými prostory konečné dimenze, aniž bychom to nadále explicitně vypisovali. Tyto prostory ale mohou být zadány i jako podprostory nějakého prostoru dimenze nekonečné, jako například když je prostor $P^n(x,\mathbb{R})$ všech polynomů stupně nejvýše n podprostorem $P(x,\mathbb{R})$.

Věta o počtu prvků báze plyne z tzv. Steinitzova lemmatu o výměně:

Lemma 1 (Steinitz). *Nechť V je vektorový prostor nad $\mathbb{F}$, M jeho n-prvková množina generátorů a $N = \{v_1,\ldots,v_k\}$ lineárně nezávislá množina ve V. Pak $k \leq n$ a prvky M lze uspořádat do posloupnosti $(u_1, u_2, \ldots, u_n)$ tak, že množina $\{v_1,\ldots,v_k,u_{k+1},\ldots u_n\}$ generuje V.*

Nejprve ukažme, jak z lemmatu plyne věta:

Důkaz věty. Označme $|X|$ počet prvků množiny nebo posloupnosti X. Aplikujeme-li lemma na dvě báze B a C ve V, $|B| = p$, $|C| = q$, pak $p \leq q$, protože B je lineárně nezávislá a C generuje V. Zároveň i $q \leq p$, protože C je lineárně nezávislá a B generuje V. Tedy $p = q$. □

Důkaz lemmatu. Budeme postupovat indukcí podle k. Pokud $k = 1$, pak $N = \{v_1\}$, kde $v_1 \neq 0$. Kdyby $n = 0$, je M prázdná množina, tedy $\langle M \rangle = 0$. Protože V obsahuje nenulový vektor v_1, není $M = \emptyset$ množinou generátorů V. Musí tedy být $n \geq 1 = k$, čímž je dokázána první část tvrzení.

Seřaďme M do posloupnosti $(u'_1, u'_2, \ldots, u'_n)$. Protože M generuje V, existují čísla r'_i taková, že $v_1 = \sum_{i=1}^n r'_i u'_i$, a protože $v_1 \neq 0$, musí pro některý index j být $r'_j \neq 0$. Vyměňme v posloupnosti $(u'_1, u'_2, \ldots, u'_n)$ prvek s indexem 1 a prvek s indexem j a označme nově uspořádanou posloupnost $(u_1, u_2, \ldots, u_n)$, analogicky převeďme posloupnost koeficientů $(r'_1, \ldots, r'_n)$ na posloupnost $(r_1, \ldots, r_n)$ výměnou prvního a j-tého členu. Lze tedy psát

$$u_1 = -\frac{1}{r_1}\left(-v_1 + \sum_{i=2}^n r_i u_i\right)$$

Každý vektor, který je lineární kombinací vektorů z $\{u_1, \ldots, u_n\}$, je tudíž také lineární kombinací vektorů z $\{v_1, u_2, \ldots, u_n\}$, čili $\{v_1, u_2, \ldots, u_n\}$ generuje V.

Nyní proveďme indukční krok, tedy předpokládejme, že $k > 1$ a tvrzení platí pro všechna přirozená čísla menší než k. Pokud $N = \{v_1, \ldots, v_k\}$ je lineárně nezávislá, pak $\{v_1, \ldots, v_{k-1}\}$ je lineárně nezávislá, a tudíž podle indukčního předpokladu $n \geq k-1$ a M lze uspořádat do posloupnosti $(u'_1, \ldots, u'_n)$ tak, že $\{v_1, \ldots, v_{k-1}, u'_k, \ldots, u'_n\}$ generuje V. Kdyby $n = k - 1$, šlo by v_k zapsat jako lineární kombinaci vektorů $v_1, \ldots, v_{k-1}$, což je spor s lineární nezávislostí N. Tedy $n \geq k$.

Vektor v_k lze pak zapsat jako

$$v_k = \sum_{i=1}^{k-1} r'_i v_i + \sum_{i=k}^n r'_i u'_i,$$

kde pro nějaké $j \geq k$ musí být $r'_j \neq 0$, jinak bychom opět dostali spor s lineární nezávislostí množiny $\{v_1, \ldots, v_k\}$. Vyměňme v posloupnosti $(u'_1, \ldots, u'_n)$ vektory k-tý a j-tý a označme toto nové uspořádání $(u_1, \ldots, u_n)$, totéž s koeficienty lineární kombinace. Pak

$$u_k = -\frac{1}{r_k}\left(\sum_{i=1}^{k-1} r_i v_i - v_k + \sum_{i=k+1}^n r_i u_i\right)$$

Potom ale každý $v \in V$ je lineární kombinací vektorů z $\{v_1, \ldots, v_k, u_{k+1}, \ldots u_n\}$, a tedy tato množina generuje V. □

Příklad 9. Najděme bázi $W = \langle (1,1,1,1), (2,0,1,-1), (1,3,-3,1) \rangle$, která obsahuje vektory $(0,1,-2,0)$ a $(2,1,-1,-1)$. Nejprve ověříme, že

$$M := \{(1,1,1,1), (2,0,1,-1), (1,3,-3,1)\}$$

je báze W a vektory z $N := \{(0,1,-2,0), (2,1,-1,-1)\}$ jsou lineární kombinace jejích prvků (proveďte sami). Dle Steinitzova lemmatu lze v M nahradit dva vektory prvky lineárně nezávislé množiny N. Zkusme to konkrétně. Protože

$$(0,1,-2,0) = -\frac{1}{2}(1,1,1,1) + \frac{1}{2}(1,3,-3,1),$$

můžeme třeba nahradit $(1,3,-3,1)$ za $(0,1,-2,0)$. V druhém kole pak z rovnosti $(2,1,-1,-1) = (0,1,-2,0) + (2,0,1,-1)$ plyne, že za $(2,1,-1,-1)$ musíme vyjmout vektor $(2,0,1,-1)$. Získáváme tedy

$$W = \langle (0,1,-2,0), (2,1,-1,-1), (1,1,1,1) \rangle$$

Proč je posloupnost $((0,1,-2,0),(2,1,-1,-1),(1,1,1,1))$ báze W? Víme, že generuje, mohli bychom ověřit lineární nezávislost. Lepší ale bude opřít se o další důsledek Steinitzova lemmatu:

Důsledek 1. *Nechť V je vektorový prostor nad $\mathbb{F}$ dimenze n a $M \subset V$.*

(1) *Je-li M lineárně nezávislá, pak $|M| \leq n$.*
(2) *Pokud M generuje V, pak $|M| \geq n$*

Navíc platí-li v některém z případů $|M| = n$, pak je M bází V.

Důkaz. Pro první část stačí uvažovat nějakou bázi N prostoru V a použít ji v lemmatu jako množinu generátorů. Pokud navíc $|M| = |N|$, plyne z lemmatu, že $\langle M \rangle = \langle N \rangle = V$, a tedy že M je bází V. Pro druhou část naopak bude N hrát v lemmatu roli lineárně nezávislé množiny. Pokud $|M| = n$ a M by byla lineárně závislá, mohli bychom z ní vybrat bázi, která by musela mít méně než n prvků. To je ale v rozporu s Větou 2. □

Lemma 2. *Nechť V je vektorový prostor dimenze n a W jeho podprostor. Pak W je prostorem konečné dimenze a* $\dim W \leq n$.

Důkaz. Všechny lineárně nezávislé podmnožiny W jsou zároveň lineárně nezávislé podmnožiny V a mají tedy podle Důsledku 1 nanejvýš n prvků. Zvolme z nich nějakou $N = \{v_1, \dots, v_k\}$ s maximálním počtem prvků. Pokud by N negenerovala celé W, pak by existoval vektor $u \in W$, pro nějž $u \notin \langle N \rangle$. Každá netriviální lineární kombinace $ru + \sum_1^k s_i v_i = 0$ je v rozporu buď s $u \notin \langle N \rangle$, nebo s lineární nezávislostí množiny N, tedy $N \cup \{u\}$ je lineárně nezávislá podmnožina W. Má ale $k+1$ prvků, což je ve sporu s předpokládanou maximalitou N. Tedy N generuje podprostor W, který má konečnou dimenzi $k \leq n = \dim V$. □

Důkaz se může zdát složitý na to, že tvrzení samotné říká na první pohled „zřejmé“ věci. Ve skutečnosti ale z žádných předchozích tvrzení neplyne, že podprostor prostoru konečné dimenze musí mít také konečnou dimenzi, a většina důkazu je tudíž ověřením této skutečnosti.

Podle dalšího důsledku lze bázi podprostoru vždy doplnit na bázi celého prostoru

Důsledek 2. *Nechť V je vektorový prostor dimenze n a W jeho podprostor, N báze W. Pak existuje množina $M \supset N$, která je bází V.*

Důkaz. Podle předchozího lemmatu je W prostor konečné dimenze a má tudíž bázi $N = \{v_1, \dots, v_k\}$. Zvolme ve V libovolnou bázi M', pak druhá část Steinitzova lemmatu říká, že množina $M := \{v_1, \dots, v_k, u_{k+1}, \dots, u_n\}$, kde $u_{k+1}, \dots, u_n$ jsou nějaké prvky M, generuje V. Protože $|M| = n$, musí to být dle Důsledku 1 báze V. □

Bude se nám hodit rozšířit pojem elementární úpravy (EÚ) z posloupnosti řádků nějaké matice na libovolnou posloupnost jakýchkoli vektorů. Řekneme, že posloupnost vznikne *elementární úpravou* posloupnosti vektorů $M = (v_1, \dots, v_m)$ z vektorového prostoru V nad $\mathbb{F}$, pokud má jeden z následujících tvarů:

$$M_1 = (v_1, \dots, v_{k-1}, v_j, v_{k+1}, \dots, v_{j-1}, v_k, v_{j+1}, \dots, v_m)$$
$$M_2 = (v_1, \dots, v_{k-1}, v_k + s v_j, v_{k+1}, \dots, v_m)$$
$$M_3 = (v_1, \dots, v_{k-1}, r v_k, v_{k+1}, \dots, v_m)$$

kde $r, s \in \mathbb{F}$, $r \neq 0$, $j, k \in \{1, \dots, m\}$, $j \neq k$.

Tvrzení 19. *Nechť C je posloupnost vektorů ve vektorovém prostoru V nad $\mathbb{F}$ a posloupnost D z ní vznikne elementární úpravou. Pak platí*

(1) *lineární obaly C a D jsou stejné*
(2) *C je lineárně nezávislá, právě když D je lineárně nezávislá*
(3) *C je bází V, právě když D je bází V*

Důkaz. Protože EÚ jsou vratné opět pomocí EÚ, stačí dokázat jen to, že každý prvek V, který lze vyjádřit jako lineární kombinaci prvků C, lze vyjádřit i jako lineární kombinaci prvků D. Uvažujme elementární úpravu typu 2 a $v = \sum_{i=1}^{m} r_i v_i \in \langle C \rangle$. Pro pohodlnější zápis předpokládejme $k < j$. Pak lze sumu přepsat jako

$$v = \sum_{i=1}^{k-1} r_i v_i + r_k(v_k + s v_j) + \sum_{i=k+1}^{j-1} r_i v_i + (r_j - s r_k) v_j + \sum_{i=j+1}^{m} r_i v_i, \tag{1}$$

tedy v je lineární kombinací vektorů z D. Pro ostatní typy EÚ je postup analogický.

Dále ukážeme, že je-li D lineárně nezávislá, pak musí být i C. Uvažme nějaké vyjádření nulového vektoru ve tvaru lineární kombinace prvků C: $o = \sum_{i=1}^{n} r_i v_i$. Pravou stranu přepišme do stejného tvaru jako v rovnici 1. Z lineární nezávislosti D plyne, že všechny koeficienty v této lineární kombinaci jsou nulové. To ale znamená $r_i = 0$, pokud i není k ani j, dále $r_k = 0$, a díky tomu $r_j = r_j - s r_k = 0$. Tedy lineární kombinace $\sum_{i=1}^{m} r_i v_i$ musí být triviální, a tedy C je lineárně nezávislá. Ostatní typy EÚ opět přenecháváme čtenáři. Třetí bod tvrzení plyne z předchozích dvou a definice báze. □

Příklad 10. Určeme dimenzi lineárního obalu množiny $M = \{(3, -6, 1, -1), (1, -2, 3, 1), (-2, 4, 0, 1), (0, 0, 2, 1)\}$ v $\mathbb{R}^4$. Sestavíme matici, která má tyto vektory jako řádky, a provádějme EŘÚ:

$$\begin{pmatrix} 1 & -2 & 3 & 1 \\ 3 & -6 & 1 & -1 \\ -2 & 4 & 0 & 1 \\ 0 & 0 & 2 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & -2 & 3 & 1 \\ 0 & 0 & -8 & -4 \\ 0 & 0 & 6 & 3 \\ 0 & 0 & 2 & 1 \end{pmatrix} \sim$$
$$\begin{pmatrix} 1 & -2 & 3 & 1 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 6 & 3 \\ 0 & 0 & -8 & -4 \end{pmatrix} \sim \begin{pmatrix} 1 & -2 & 3 & 1 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Řádky poslední matice jsou LZ, tedy i M je LZ. Lineární obal řádků se také zachovává, tedy $\langle M \rangle$ je roven lineárnímu obalu LN množiny $\{(1, -2, 3, 1), (0, 0, 2, 1)\}$. Proto $\dim\langle M \rangle = 2$.

KAPITOLA 6

Hodnost matice

V této kapitole se vrátíme k maticím a soustavám lineárních rovnic. Elementy matic a koeficienty soustav lineárních rovnic budou prvky množiny skalárů $\mathbb{F}$. Všechny definice z kapitol 2 a 3 zůstávají stejné i pro matice z $\mathbb{F}^{m\times n}$, máme tedy definován součet a součin matic, nulovou matici, jednotkovou matici, pojem matice transponované, inverzní, čtvercové a regulární, matice elementární řádkové úpravy 5. Snadno se zobecní i všechna tvrzení, která jsme o těchto pojmech dokázali, tedy asociativita 4 a distributivita 5 maticových operací, nutná a postačující podmínka existence inverzní matice 12 i fakt, že Gaussova eliminace dovede každou matici do redukovaného odstupňovaného tvaru.

Stejně se definují i jádro a obraz matice:

Definice 20. Nechť $A = (\mathbf{a}_1|\dots|\mathbf{a}_n) \in \mathbb{F}^{m\times n}$, kde $\mathbf{a}_i \in \mathbb{F}^m$. Pak *jádrem matice* A rozumíme podprostor

$$\operatorname{Ker} A = \{\mathbf{x} \in \mathbb{F}^n | A\mathbf{x} = \mathbf{o}\} \le \mathbb{F}^n$$

a jejím *obrazem* nebo též *sloupcovým prostorem* pak

$$\operatorname{Im} A = \langle \mathbf{a}_1, \dots, \mathbf{a}_n \rangle \le \mathbb{F}^m$$

Důkaz, že $\operatorname{Ker} A$ je podprostorem $\mathbb{F}^n$ je stejný jako v tvrzení 11. $\operatorname{Im} A$ lze definovat i jako obor hodnot zobrazení F_A. Protože $F_A(\mathbf{x}) = A\mathbf{x} = \sum_{i=1}^n x_i\mathbf{a}_i$, je vektor $\mathbf{y} \in \mathbb{F}^m$ v oboru hodnot F_A, právě když je v lineárním obalu sloupců matice A.

Transponování převádí řádky na sloupce, můžeme proto označit podprostor $\operatorname{Im} A^T \le \mathbb{F}^n$, čili lineární obal řádků matice A, jako její *řádkový prostor*. Čtveřici význačných podprostorů, jejichž vlastnostmi se v této kapitole budeme zabývat, doplňuje *jádro transponované matice* $\operatorname{Ker} A^T \le \mathbb{F}^m$.

Elementární řádkovou úpravu matice lze realizovat násobením elementární maticí zleva. Elementární matice jsou regulární a stejně tak jsou i součiny elementárních matic mezi sebou. Následující tvrzení tedy jen zobecňuje již dokázané vlastnosti elementárních úprav, totiž že zachovávají množinu řešení homogenní soustavy rovnic (tvrzení 8 to říká dokonce i pro nehomogenní) a řádkový prostor (první bod tvrzení 19):

Tvrzení 20. *Nechť* $A = (\mathbf{a}_1|\dots|\mathbf{a}_n) \in \mathbb{F}^{m\times n}$, $R \in \mathbb{F}^{m\times m}$ *regulární. Pak*

(1) $\operatorname{Ker}(RA) = \operatorname{Ker} A$
(2) $\operatorname{Im}(RA)^T = \operatorname{Im} A^T$

Důkaz. Pokud $A\mathbf{x} = \mathbf{o}$, pak $RA\mathbf{x} = R\mathbf{o} = \mathbf{o}$, tedy $\operatorname{Ker} A \subset \operatorname{Ker} RA$. Protože R je regulární, existuje k ní inverzní matice R^{-1}, která je také regulární. Pak $\operatorname{Ker} RA \subset \operatorname{Ker} R^{-1}(RA) = \operatorname{Ker} A$. Tím je dokázána i opačná implikace.

Každý řádek matice RA je lineární kombinací řádků matice A s koeficienty v řádcích matice R, tedy $\operatorname{Im}(RA)^T \subset \operatorname{Im} A^T$. Druhá inkluze plyne opět z $A = R^{-1}(RA)$. $\square$

Elementární sloupcové úpravy (ESÚ) lze definovat jako ty, které lze realizovat násobením elementární maticí zprava. Opět je jednodušší vzít místo elementární matice libovolnou regulární a dokázat, že násobení zprava nemění sloupcový prostor a jádro transponované matice:

Tvrzení 21. *Nechť* $A = (\mathbf{a}_1|\dots|\mathbf{a}_n) \in \mathbb{F}^{m\times n}$, $Q \in \mathbb{F}^{n\times n}$ *regulární. Pak*

(1) $\operatorname{Ker}(AQ)^T = \operatorname{Ker} A^T$
(2) $\operatorname{Im}(AQ) = \operatorname{Im} A$

Na příkladu

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

je vidět, že $\operatorname{Im}(RA)$ a $\operatorname{Im} A$ se rovnat nemusejí a stejně tak $\operatorname{Ker}(RA)^T$ a $\operatorname{Ker} A^T$. Řádkové úpravy tedy sloupcový prostor nezachovávají. Ukážeme ale, že zachovávají jeho dimenzi. První část tvrzení 20 můžeme přeformulovat jako

Lemma 3. *Nechť* $A = (\mathbf{a}_1|\dots|\mathbf{a}_n) \in \mathbb{F}^{m\times n}$, $R \in \mathbb{F}^{m\times m}$ *regulární. Označme* $RA =: A' = (\mathbf{a}'_1|\dots|\mathbf{a}'_n)$. *Je-li* $(s_1,\dots,s_n) \in \mathbb{F}^n$, *pak* $\sum_{i=1}^n s_i\mathbf{a}_i = \mathbf{o}$, *právě když* $\sum_{i=1}^n s_i\mathbf{a}'_i = \mathbf{o}$.

Důsledek 3. *Nechť* A, A' *jsou dvě matice,* $A \sim A'$. *Pak* $\dim\operatorname{Im} A = \dim\operatorname{Im} A'$.

Důkaz. Z lemmatu 3 plyne, že nějaká podposloupnost $(\mathbf{a}_{i_1},\mathbf{a}_{i_2},\dots,\mathbf{a}_{i_k})$ sloupců matice A, kde $i_1 < \dots < i_k$, je lineárně nezávislá, právě když je lineárně nezávislá odpovídající podposloupnost $(\mathbf{a}'_{i_1},\mathbf{a}'_{i_2},\dots,\mathbf{a}'_{i_k})$ sloupců A'. Maximální takové lineárně nezávislé podposloupnosti jsou bázemi $\operatorname{Im} A$, resp. $\operatorname{Im} A'$, a počet prvků takových bází je roven $\dim\operatorname{Im} A = \dim\operatorname{Im} A'$. □

Věta 3. *Pro každou matici* $A \in \mathbb{F}^{m\times n}$ *platí* $\dim\operatorname{Im} A = \dim\operatorname{Im} A^T$.

Důkaz. Matici A lze převést posloupností EŘÚ na redukovaný odstupňovaný tvar A'. Posloupnost všech nenulových řádků A' je lineárně nezávislá, a je tedy bází $\operatorname{Im} A'^T$. Množina všech pivotních sloupců A' je lineárně nezávislá a nepivotní sloupce jsou lineárními kombinacemi sloupců pivotních. Tedy posloupnost všech pivotních sloupců je bází $\operatorname{Im} A'$. Pivotních sloupců i nenulových řádků A' je stejně, tedy $\dim\operatorname{Im} A' = \dim\operatorname{Im} A'^T$. Protože řádkové úpravy $A \sim A'$ zachovávají řádkový prostor, platí $\operatorname{Im} A^T = \operatorname{Im} A'^T$. Z důsledku 3 máme $\dim\operatorname{Im} A = \dim\operatorname{Im} A'$, celkově tedy $\dim\operatorname{Im} A = \dim\operatorname{Im} A^T$. □

Na tvrzení je zajímavé, že nám dává rovnost dimenzí dvou podprostorů ve dvou obecně různých aritmetických vektorových prostorech, protože $\operatorname{Im} A \leq \mathbb{F}^m$, ale $\operatorname{Im} A^T \leq \mathbb{F}^n$. Je to vidět i na příkladu

$$A = \begin{pmatrix}1 & 0 & 1\\ 1 & 0 & 1\end{pmatrix},$$

kde je bází $\operatorname{Im} A$ jednoprvková posloupnost $((1,1))$ a bází $\operatorname{Im} A^T$ posloupnost $((1,0,1))$. Jednoduchý důsledek věty 3 je tedy , že pokud jsou v matici všechny řádky násobkem jednoho z nich, pak jsou i všechny sloupce násobkem jednoho z nich.

Definice 21. Nechť $A \in \mathbb{F}^{m\times n}$. Číslo $\dim\operatorname{Im} A \equiv \dim\operatorname{Im} A^T$ nazýváme *hodnost* matice A, značíme $\operatorname{rank}(A)$.

Tvrzení 22. *Nechť* $A \in \mathbb{F}^{m\times n}$, $R \in \mathbb{F}^{p\times m}$, $Q \in \mathbb{F}^{n\times q}$. *Pak*

(1) $\operatorname{rank}(RA) \leq \operatorname{rank}(A)$ *a pokud* $p = m$ *a* R *je regulární,* $\operatorname{rank}(RA) = \operatorname{rank}(A)$
(2) $\operatorname{rank}(AQ) \leq \operatorname{rank}(A)$ *a pokud* $n = q$ *a* Q *je regulární,* $\operatorname{rank}(AQ) = \operatorname{rank}(A)$
(3) $\operatorname{rank}(A) = \operatorname{rank}(A^T)$

DŮKAZ. Protože řádky RA patři do lineárního obalu řádků A, musí být $\operatorname{rank}(RA) \leq \operatorname{rank}(A)$. Protože R je regulární, existuje inverzní matice R^{-1}, která je rovněž regulární. Platí pak $\operatorname{rank}(R^{-1}(RA)) \leq \operatorname{rank}(RA)$, z čehož plyne i opačná nerovnost. Druhý bod se dokáže analogicky s pomocí sloupcového prostoru matice A, třetí bod plyne z věty 3. □

Důsledkem tvrzení je nerovnost

$$\operatorname{rank}(AB) \leq \min\{\operatorname{rank}(A), \operatorname{rank}(B)\}$$

a rovnost

$$\operatorname{rank}(ABC) = \operatorname{rank}(B), \text{ pro } A, C \text{ regulární.}$$

Hodnost se tedy dá určovat tak, že kombinací řádkových a sloupcových úprav převedeme matici do tvaru, v němž je možné ji určit snadněji, typicky do odstupňovaného. Hodnost nese o matici a o příslušném zobrazení mnoho informací:

VĚTA 4. *Nechť $A \in \mathbb{F}^{n \times n}$. Pak jsou následující tvrzení ekvivalentní*

(1) *A je regulární*
(2) $\operatorname{rank}(A) = n$
(3) *množina všech řádků A je lineárně nezávislá*
(4) *množina všech řádků A generuje $\mathbb{F}^n$*
(5) *posloupnost všech řádků A je bází $\mathbb{F}^n$*
(6) *množina všech sloupců A je lineárně nezávislá*
(7) *množina všech sloupců A generuje $\mathbb{F}^n$*
(8) *posloupnost všech sloupců A je bází $\mathbb{F}^n$*
(9) $\operatorname{Im} A = \mathbb{F}^n$
(10) $\operatorname{Ker} A = 0$
(11) *zobrazení F_A je prosté*
(12) *zobrazení F_A je na*
(13) *zobrazení F_A je bijektivní*
(14) *existuje $X \in \mathbb{F}^{n \times n}$, pro kterou $AX = E$.*
(15) *existuje $X \in \mathbb{F}^{n \times n}$, pro kterou $XA = E$.*

DŮKAZ. Dokážeme řetězec implikací, který propojí všechna tvrzení oběma směry.

$1 \Rightarrow 2$ Je-li A regulární, pak existuje regulární matice A^{-1} splňující $A^{-1}A = E$. Násobení regulární maticí nemění hodnost a $\operatorname{rank}(E) = n$, tedy $\operatorname{rank}(A) = n$.

$2 \Rightarrow 3$ Protože $\dim \operatorname{Im} A^T = \operatorname{rank}(A) = n$, tvoří řádky A n-prvkovou množinu generátorů vektorového prostoru dimenze n. Ta je podle důsledku 1 jeho bází a tudíž je lineárně nezávislá.

$3 \Rightarrow 5$ Množina řádků A je n-prvková lineárně nezávislá podmnožina v $\mathbb{F}^n$, tedy je dle důsledku 1 i jeho bází.

$4 \Rightarrow 5$ Množina řádků A je n-prvková množina generátorů $\mathbb{F}^n$, tedy je dle důsledku 1 i jeho bází.

$2 \Rightarrow 6$ se dokáže podobně jako $2 \Rightarrow 3$. Analogicky máme i $6 \Rightarrow 8$ a $7 \Rightarrow 8$. Implikace $5 \Rightarrow 3$, $5 \Rightarrow 4$, $4 \Rightarrow 2$, $8 \Rightarrow 6$, $8 \Rightarrow 7$, $7 \Rightarrow 2$ plynou z definic. Také je $7 \Leftrightarrow 9 \Leftrightarrow 12$, $6 \Leftrightarrow 10 \Leftrightarrow 11$. Odtud plyne ekvivalence výroků 6 až 12 s výrokem 13.

$13 \Rightarrow 14$ Plyne z tvrzení 12.

$14 \Rightarrow 15$ Je dokázáno v druhé části důkazu 12. Z definice regularity matice pak $14 \Rightarrow 1$.

$15 \Rightarrow 4$ Řádky XA jsou lineárními kombinacemi řádků matice A, lze tedy v lineárním obalu řádků A najít libovolný vektor kanonické báze $\mathbb{F}^n$ a tedy i libovolný vektor $\mathbb{F}^n$.

□

Věta 5 (O dimenzi jádra a obrazu). *Nechť $A \in \mathbb{F}^{m\times n}$. Pak*

$$\dim \operatorname{Ker} A + \dim \operatorname{Im} A = n$$

Důkaz. Pro nulovou matici je tvrzení triviální, předpokládejme tedy $A \neq 0$. Matici A je možné převést pomocí EŘÚ do redukovaného odstupňovaného tvaru, tím se zachovává jak $\operatorname{Ker} A$, tak $\dim \operatorname{Im} A$. Dále vynechejme všechny nulové řádky (ani tím se $\operatorname{Ker} A$ ani $\dim \operatorname{Im} A$ nezmění) a označme výslednou matici A'. Označme $k_1 < k_2 < \ldots < k_r$ sloupcové indexy pivotních sloupců A', tedy $\mathbf{a}'_{k_i} = \mathbf{e}_i \in \mathbb{F}^r$. Označme sloupcové indexy nepivotních sloupců $j_1 < j_2 < \ldots < j_p$, a tyto sloupce samotné $\mathbf{c}_i := \mathbf{a}'_{j_i}$. Uvažujme vektor $\mathbf{x} \in \mathbb{F}^n$, jehož složky s indexy j_1 až j_p jsou zvoleny libovolně. Pokud $A'\mathbf{x} = 0$, pak

$$\sum_{s=1}^{r} x_{k_s}\mathbf{e}_s + \sum_{q=1}^{p} x_{j_q}\mathbf{c}_q = \mathbf{o},$$

neboli

$$\begin{pmatrix} x_{k_1} \\ \vdots \\ x_{k_r} \end{pmatrix} = -\sum_{q=1}^{p} x_{j_q}\mathbf{c}_q$$

Tím jsou určeny všechny zbývající složky vektoru $\mathbf{x}$. Definujme pro každé $i \in \{1,\ldots,p\}$ vektor $\mathbf{u}_i \in \mathbb{F}^n$ tak, že pro každé $q \in \{1,\ldots,p\}$ je jeho j_q-tá složka rovna q-té složce vektoru $\mathbf{e}_i \in \mathbb{F}^p$, a pro každé $s \in \{1,\ldots,r\}$ je jeho k_s-tá složka rovná s-té složce vektoru $-\mathbf{c}_i \in \mathbb{F}^r$. Pak $\mathbf{x} = \sum_{q=1}^{p} x_{j_q}\mathbf{u}_q$, tedy posloupnost $M = (\mathbf{u}_1,\ldots,\mathbf{u}_p)$ generuje $\operatorname{Ker} A'$.

Zároveň, pokud by $\sum_{q=1}^{p} y_q\mathbf{u}_q = \mathbf{o}$, musí být každý koeficient y_q roven nule, neboť se rovná j_q-té složce vektoru $\sum_{q=1}^{p} y_q\mathbf{u}_q$. Tedy M je lineárně nezávislá. Musí být tedy bází $\operatorname{Ker} A$, $\dim \operatorname{Ker} A' = p$. Protože $\dim \operatorname{Im} A' = r$ a $p + r = n$, dostáváme odtud tvrzení věty. □

Konstrukce báze M je názornější, pokud $k_s = s$, $j_q = r + q$, tedy pokud jdou v A' nejprve sloupce pivotní a pak nepivotní:

$$A' = \begin{pmatrix} 1 & 0 & \ldots & 0 & c_{11} & c_{12} & \ldots & c_{1p} \\ 0 & 1 & \ldots & 0 & c_{21} & c_{22} & \ldots & c_{2p} \\ \vdots & & \ddots & \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & \ldots & 1 & c_{r1} & c_{r2} & \ldots & c_{rp} \end{pmatrix}$$

Báze jádra pak má tvar

$$\begin{aligned} M = \{&(-c_{11}, -c_{21}, \ldots, -c_{r1}, 1, 0, \ldots, 0)^T \\ &(-c_{12}, -c_{22}, \ldots, -c_{r2}, 0, 1, \ldots, 0)^T \\ &\vdots \\ &(-c_{1p}, -c_{2p}, \ldots, -c_{rp}, 0, 0, \ldots, 1)^T\} \end{aligned}$$

a tyto bázové vektory odpovídají vektorům $\mathbf{u}_1,\ldots,\mathbf{u}_p$ v důkazu. Vektor $\mathbf{x}$ v důkazu by byl v tomto speciální případě roven

$$\left(-\sum_{q=1}^{p} c_{1q}x_{r+q}, -\sum_{q=1}^{p} c_{2q}x_{r+q}, \ldots, -\sum_{q=1}^{p} c_{rq}x_{r+q}, x_{r+1}, x_{r+2}, \ldots, x_{r+p}\right)^T$$

Dimenze jádra A se nazývá *nulita* (též *defekt*) matice, značí se $n(A)$. Věta o dimenzi jádra a obrazu je proto známá též pod názvem věta o hodnosti a nulitě:

$$\forall A \in \mathbb{F}^{m\times n} : \operatorname{rank}(A) + n(A) = n$$

Plyne z ní například

TVRZENÍ 23. *Nechť* $A \in \mathbb{R}^{m \times n}$. *Pak* $\operatorname{Ker} A^T A = \operatorname{Ker} A$ *a* $\operatorname{rank} A^T A = \operatorname{rank} A$.

DŮKAZ. Pokud $A\mathbf{x} = \mathbf{o}$, pak jistě $A^T A\mathbf{x} = \mathbf{o}$, tedy $\operatorname{Ker} A \leq \operatorname{Ker} A^T A$. Pokud $A^T A\mathbf{x} = \mathbf{o}$, pak $\mathbf{x}^T A^T A\mathbf{x} = 0$. Označme $\mathbf{y} := A\mathbf{x}$, pak $\mathbf{y}^T = \mathbf{x}^T A^T$ a tedy $0 = \mathbf{y}^T\mathbf{y} = \sum_{i=1}^m y_i^2$. Musí tedy být $\mathbf{y} = 0$, neboli $\mathbf{x} \in \operatorname{Ker} A$. Tím je dokázána druhá inkluze. Protože počet sloupců $A^T A$ je stejný jako počet sloupců A, plyne odsud rovnost hodností. $\square$

Tvrzení platí jen pro **reálné** matice. Kde selže důkaz pro $A \in \mathbb{C}^{m \times n}$? Jak by se muselo upravit tvrzení, aby důkaz prošel?

Pomocí hodnosti matice se dá také formulovat kritérium řešitelnosti SLR:

VĚTA 6 (Frobeniova). *Nechť* $A \in \mathbb{F}^{m \times n}$, $\mathbf{b} \in \mathbb{F}^m$. *Soustava* $A\mathbf{x} = \mathbf{b}$ *má řešení, právě když* $\operatorname{rank}(A) = \operatorname{rank}(A|b)$.

DŮKAZ. Soustava má řešení, právě když je $\mathbf{b} \in \operatorname{Im} A$, což nastává právě když $\operatorname{Im}(A|b) = \operatorname{Im}(A)$. Protože $\operatorname{Im} A \leq \operatorname{Im}(A|b)$, nastane toto právě když se rovnají dimenze. $\square$

Pomocí hodnosti se dá formulovat i to, kolik řešení soustava má. Víme, že množina všech řešení soustavy $A\mathbf{x} = \mathbf{b}$ je afinní podprostor $\mathbb{F}^n$ ve tvaru

$$\mathbf{x}_P + \operatorname{Ker} A$$

Z věty o hodnosti a nulitě víme, že dimenze $\operatorname{Ker} A$ je $n - \operatorname{rank}(A)$.

KAPITOLA 7

Reprezentace vektoru a lineárního zobrazení

V definici 18 jsme zavedli pojem báze a reprezentace vektoru vůči bázi. Reprezentaci vzhledem k bázi $B = (v_1, \ldots, v_n)$ prostoru V nad $\mathbb{F}$ dimenze n lze chápat jako zobrazení $[\ ]^B : V \to \mathbb{F}^n$, které vektoru $v = \sum_{i=1}^n r_i v_i \in V$ přiřadí aritmetický vektor

$$[v]^B := \begin{pmatrix} r_1 \\ r_2 \\ \vdots \\ r_n \end{pmatrix} \in \mathbb{F}^n$$

Toto zobrazení splňuje důležitou vlastnost:

Tvrzení 24. *Nechť V je vektorový prostor $\mathbb{F}$ konečné dimenze, B jeho báze, $u, v \in V$, $r, s \in \mathbb{F}$. Pak $[ru + sv]^B = r[u]^B + s[v]^B$.*

Důkaz. Nechť $B = (v_1, \ldots, v_n)$, $u = \sum_{i=1}^n r_i v_i$, $v = \sum_{i=1}^n s_i v_i$. Pak

$$[ru + sv]^B = \left[r \sum_{i=1}^n r_i v_i + s \sum_{i=1}^n s_i v_i \right]^B = \left[\sum_{i=1}^n (rr_i + ss_i) v_i \right]^B =$$

$$= (rr_1 + ss_1, \ldots, rr_n + ss_n)^T = r(r_1, \ldots, r_n)^T + s(s_1, \ldots, s_n)^T = r[u]^B + s[v]^B$$

□

Stejnou vlastnost jsme viděli poprvé v tvrzení 3 v první kapitole, kde se týkala zobrazení ortogonální projekce $P_{\mathbf{x}}$, a posléze i u dalších zobrazení. Definujme ji nyní obecně:

Definice 22. Nechť V, W jsou dva vektorové prostory nad $\mathbb{F}$ a $f : V \to W$ je zobrazení splňující $\forall r, s \in \mathbb{F}$, $\forall u, v \in V$

$$f(ru + sv) = rf(u) + sf(v)$$

Takové f nazýváme *lineární zobrazení*.

Označme o_V nulový vektor ve V a o_W nulový vektor ve W.

Důsledek 4. *Nechť V, W jsou dva vektorové prostory nad $\mathbb{F}$ a $f : V \to W$ je lineární zobrazení. Pak*

(1) $\forall u, v \in V : f(u + v) = f(u) + f(v)$
(2) $\forall u \in V$, $\forall r \in \mathbb{F}$: $f(ru) = rf(u)$.
(3) $f(o_V) = o_W$
(4) $\forall v \in V$: $f(-v) = -f(v)$

Důkaz. Volíme v definici lineárního zobrazení postupně $r = s = 1$; $s = 0$; $r = s = 0$; $r = 0$, $s = -1$ a využíváme základní vlastnosti vektorového prostoru. □

Příklady 2. Kromě $[\ ]^B$ a příkladů uvedených v první kapitole uveďme ještě několik dalších:

- Pokud $A \in \mathbb{F}^{m \times n}$ je matice, pak zobrazení $F_A : \mathbb{F}^n \to \mathbb{F}^m$ zavedené stejně jako v první kapitole předpisem $F_A(\mathbf{x}) = A\mathbf{x}$, je lineární.

- Zobrazení výše můžeme snadno zobecnit na zobrazení $f : \mathbb{F}^{n\times k} \to \mathbb{F}^{m\times k}$, $f(X) = AX$ mezi prostory matic.
- Zobrazení $g : P^n(x, \mathbb{R}) \to P^{n-1}(x, \mathbb{R})$, které přiřazuje polynomu $p(x)$ jeho první derivaci $\frac{d}{dx}p(x)$, je také lineární.
- Zobrazení $h : F(\mathbb{R}, \mathbb{R}) \to \mathbb{R}$, které přiřazuje funkci $\phi \in F(\mathbb{R}, \mathbb{R})$ její hodnotu v nějakém bodě, např $h(\phi) = \phi(7)$, je rovněž lineární.

V úvodní kapitole jsme vyvodili, že lineárnímu zobrazení $f : \mathbb{F}^n \to \mathbb{F}^m$ můžeme přiřadit matici $A = (\mathbf{a}_1|\dots|\mathbf{a}_n)$, jejíž i-tý sloupec je obrazem i-tého vektoru kanonické báze, tedy $f(\mathbf{e}_i) = \mathbf{a}_i$. Tuto úvahu můžeme rozšířit na obecné lineární zobrazení:

Definice 23. Nechť V, W jsou dva vektorové prostory nad $\mathbb{F}$, $f : V \to W$ je lineární zobrazení, $B = (v_1, \dots, v_n)$ je báze V, $C = (w_1, \dots, w_m)$ je báze W. Pak matici

$$[f]^C_B := \big([f(v_1)]^C \big| [f(v_2)]^C \big| \dots \big| [f(v_n)]^C\big)$$

nazýváme *reprezentací lineárního zobrazení f vzhledem k bázím B a C.*

Zvolíme-li za B a C kanonické báze $K_n = (\mathbf{e}_1, \dots, \mathbf{e}_n)$ v $\mathbb{F}^n$, resp. $K_m = (\mathbf{e}_1, \dots, \mathbf{e}_m)$ v $\mathbb{F}^m$, vidíme, že

$$[f]^{K_m}_{K_n} = ([f(\mathbf{e}_1)]^{K_m} | [f(\mathbf{e}_2)]^{K_m} | \dots | [f(\mathbf{e}_n)]^{K_m}) = (f(\mathbf{e}_1)|f(\mathbf{e}_2)| \dots |f(\mathbf{e}_n))$$

Pro zobrazení F_A mezi aritmetickými vektorovými prostory je tedy matice A tohoto zobrazení podle definice v první kapitole rovna reprezentaci $[F_A]^{K_m}_{K_n}$ zobrazení F_A vzhledem ke kanonickým bázím.

Příklad 11. Uvažujme zobrazení $g : P^2(x, \mathbb{R}) \to P^1(x, \mathbb{R})$, $g(p(x)) = \frac{d}{dx}p(x)$, a báze $B = \{1, x, x^2\} \subset P^2(x, \mathbb{R})$, $C = \{1, x\} \subset P^1(x, \mathbb{R})$. Reprezentace obecného prvku prostoru $P^2(x, \mathbb{R})$ vzhledem k bázi B je

$$[ax^2 + bx + c]^B = \begin{pmatrix} c \\ b \\ a \end{pmatrix}$$

Obrazy prvků báze B (čili derivace monomů) jsou

$$g(1) = 0, \qquad g(x) = 1, \qquad g(x^2) = 2x$$

a jejich reprezentace vzhledem k bázi C jsou

$$[g(1)]^C = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \qquad [g(x)]^C = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad [g(x^2)]^C = \begin{pmatrix} 0 \\ 2 \end{pmatrix},$$

tedy reprezentace g vzhledem k B a C je $[g]^C_B = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$.

Příklad 12. Uvažujme lineární zobrazení $F_A : \mathbb{R}^2 \to \mathbb{R}^3$, dané předpisem

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ -2 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \equiv A\mathbf{x}$$

Pokud K_n označuje kanonickou bázi v $\mathbb{F}^n$, pak $[F_A]^{K_3}_{K_2} = A$. Zvolme jiné báze než kanonické: $B = ((1, 2)^T, (1, 3)^T)$, $C = ((1, 1, 0)^T, (1, 0, 0)^T, (1, 0, 1)^T)$. Určíme

$$F_A\left(\begin{pmatrix} 1 \\ 2 \end{pmatrix}\right) = \begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix}, F_A\left(\begin{pmatrix} 1 \\ 3 \end{pmatrix}\right) = \begin{pmatrix} 4 \\ 1 \\ 1 \end{pmatrix}$$

Zbývá najít reprezentaci obou vektorů vzhledem k C

$$\left(\begin{array}{ccc|cc} 1 & 1 & 1 & 3 & 4 \\ 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 \end{array}\right) \sim \left(\begin{array}{ccc|cc} 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 2 & 2 \\ 0 & 0 & 1 & 0 & 1 \end{array}\right) \Rightarrow [F_A]_B^C = \begin{pmatrix} 1 & 1 \\ 2 & 2 \\ 0 & 1 \end{pmatrix}$$

TVRZENÍ 25. *Nechť V, W jsou dva vektorové prostory nad $\mathbb{F}$, $f : V \to W$ lineární zobrazení, B báze V, C báze W, $v \in V$. Pak*

$$[f(v)]^C = [f]_B^C [v]^B$$

.

DŮKAZ. Označme $B = (v_1, \ldots, v_n)$ a zapišme $v = \sum_{i=1}^n r_i v_i$, kde $(r_1, \ldots, r_n)^T = [v]^B$, pak

$$[f(v)]^C = \left[\sum_{i=1}^n r_i f(v_i)\right]^C = \sum_{i=1}^n r_i [f(v_i)]^C = [f]_B^C [v]^B$$

Nejprve jsme využili linearitu f, pak linearitu $[\ ]^C$ a nakonec definici součinu matice a vektoru. □

Bijektivní lineární zobrazení se nazývá *izomorfismus.* S tímto pojmem jsme se setkali už v kapitole 3, a ve větě 4 jsme ukázali, že F_A je izomorfismus, právě když A je regulární matice. To se dá snadno zobecnit. Nejprve ale potřebujeme dokázat některé vlastnosti izomorfismů:

TVRZENÍ 26. *Nechť V, W jsou vektorové prostory nad $\mathbb{F}$, $f : V \to W$ je izomorfismus a $(v_1, \ldots, v_n)$ je posloupnost vektorů ve V. Pak*

(1) *$(v_1, \ldots, v_n)$ je lineárně nezávislá, právě když $(f(v_1), \ldots, f(v_n))$ je lineárně nezávislá.*
(2) *$(v_1, \ldots, v_n)$ generuje V, právě když $(f(v_1), \ldots, f(v_n))$ generuje W.*
(3) *$(v_1, \ldots, v_n)$ je báze V, právě když $(f(v_1), \ldots, f(v_n))$ je báze W.*

DŮKAZ. Označme o_V nulový vektor ve V a o_W nulový vektor ve W. Pokud $\sum_{i=1}^n r_i v_i = o_V$, pak i $\sum_{i=1}^n r_i f(v_i) = o_W$. Je-li tedy $(f(v_1), \ldots, f(v_k))$ lineárně nezávislá, musí být všechna $r_i = 0$, tedy $(v_1, \ldots, v_k)$ je také lineárně nezávislá. Nechť naopak $\sum_{i=1}^n r_i f(v_i) = o_W$, pak z linearity zobrazení f plyne $f(\sum_{i=1}^n r_i v_i) = o_W$. Protože také $f(o_V) = o_W$ a f je prosté, musí být $\sum_{i=1}^n r_i v_i = o_V$. Je-li $(v_1, \ldots, v_n)$ lineárně nezávislá, znamená to, že všechna r_i jsou nulová, tedy i $(f(v_1), \ldots, f(v_n))$ je lineárně nezávislá. Tím jsme ukázali obě implikace, které dohromady dávají tvrzení v prvním bodě.

Protože f je na, existuje ke každému $w \in W$ vektor $v \in V$ takový, že $f(v) = w$. Pokud $(v_1, \ldots, v_n)$ generuje V, pak existují r_i taková, že $v = \sum_{i=1}^n r_i v_i$. Pak ale z linearity dostáváme, že $w = \sum_{i=1}^n r_i f(v_i)$, tedy $(f(v_1), \ldots, f(v_n))$ generuje W. Naopak, je-li $v \in V$, $w = f(v)$ a $w = \sum_{i=1}^n r_i f(v_i)$, pak z linearity $f(v - \sum_{i=1}^n r_i v_i) = o_W$. Protože f je prosté, znamená to, že $v = \sum_{i=1}^n r_i v_i$, tedy $(v_1, \ldots, v_n)$ generuje V. Tím je dokázán druhý bod.

Třetí bod plyne z prvního a druhého. □

DŮSLEDEK 5. *Nechť V, W jsou dva vektorové prostory nad $\mathbb{F}$ a $f : V \to W$ je izomorfismus. Pak* $\dim V = \dim W$.

DŮKAZ. Ze třetího bodu tvrzení 26 plyne, že pokud je ve V nebo ve W konečná báze, pak je v druhém z prostorů také konečná báze o stejném počtu prvků. Pokud ani v jednom z prostorů konečná báze není, pak $\dim V = \dim W = \infty$. □

TVRZENÍ 27. *Nechť V, W jsou dva vektorové prostory nad $\mathbb{F}$ konečné dimenze, $f : V \to W$ lineární zobrazení. Pak jsou následující výroky ekvivalentní:*

(1) f *je izomorfismus*
(2) *Pro všechny dvojice bází* $B \subset V$, $C \subset W$ *je* $[f]_B^C$ *regulární*
(3) *Pro některou dvojici bází je* $B \subset V$, $C \subset W$ *je* $[f]_B^C$ *regulární*

Důkaz. Pokud je f izomorfismus, pak je to i prosté zobrazení, platí tedy $f(v) = o_W$ pouze pro $v = o_V$. Zvolme bázi $B \in V$ a $C \in W$, pak z důsledku 5 plyne, že počet prvků B a C je stejný, označme jej n. Platí $f(v) = o_W$ právě když $[f(v)]^C = \mathbf{o} \in \mathbb{F}^m$ a $v = o_V$ právě když $[v]^B = \mathbf{o} \in \mathbb{F}^n$. Tedy $[f(v)]^C \equiv [f]_B^C[v]^B = \mathbf{o}$, právě když $[v]^B = \mathbf{o}$, neboli $\mathrm{Ker}[f]_B^C = 0$. Protože matice $[f]_B^C$ je čtvercová, plyne z toho na základě tvrzení 4, že je i regulární. Tím je dokázána implikace $1 \Rightarrow 2$. Implikace $2 \Rightarrow 3$ je zřejmá, zbývá dokázat $3 \Rightarrow 1$. Pokud $[f]_B^C$ je regulární matice z $\mathbb{F}^{n\times n}$, pak z tvrzení 4 vyplývá $\mathrm{Ker}[f]_B^C = 0$ a $\mathrm{Im}[f]_B^C = \mathbb{F}^n$. Z první vlastnosti plyne, že $[f(v)]^C = \mathbf{o}$, právě když $[v]^B = \mathbf{o}$, tedy $f(v) = o_W$ právě když $v = o_V$, tedy f je prosté. Z druhé víme, že ke každému $\mathbf{y} \in \mathbb{F}^n$ existuje $\mathbf{x} \in \mathbb{F}^n$ takový, že $\mathbf{y} = [f]_B^C\mathbf{x}$. Protože ke každému $\mathbf{x}$ existuje $v \in V$ takový, že $\mathbf{x} = [v]^B$, plyne z toho, že pro každý $\mathbf{y} \in \mathbb{F}^n$ existuje $v \in V$ takové, že $[f(v)]^C = \mathbf{y}$. Pokud $\mathbf{y} = [w]^C$ pro nějaký $w \in W$, pak $[f(v)]^C = [w]^C$ a tedy i $f(v) = w$. Zobrazení f je tudíž nejen prosté, ale také na, je to tedy izomorfismus. □

Aplikujme tvrzení 26 na $[\ ]^B : V \to \mathbb{F}^n$:

Důsledek 6. *Nechť* $(v_1, \dots, v_k)$ *je posloupnost vektorů ve vektorovém prostoru* V *dimenze* n *a* B *je báze* V. *Pak*

(1) $(v_1, \dots, v_k)$ *je LN, právě když* $([v_1]^B, \dots, [v_k]^B)$ *je LN.*
(2) $(v_1, \dots, v_k)$ *generuje* V, *právě když* $([v_1]^B, \dots, [v_k]^B)$ *generuje* $\mathbb{F}^n$.
(3) $(v_1, \dots, v_k)$ *je báze* V, *právě když* $([v_1]^B, \dots, [v_k]^B)$ *báze* $\mathbb{F}^n$.

Tento důsledek nám umožňuje převést mnoho výpočtů na hledání hodnosti nějaké matice. Například

$$\{x^2 + x + 1, x^2 + 2x - 1, 2x^2 - 5x + 3\} \text{ je LN} \Leftrightarrow$$

$$\left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}, \begin{pmatrix} 2 \\ -5 \\ 3 \end{pmatrix} \right\} \text{ je LN} \Leftrightarrow \mathrm{rank} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & -1 \\ 2 & -5 & 3 \end{pmatrix} = 3$$

Definice 24. Nechť $f : V \to W$ je lineární zobrazení. Pak množina

$$\mathrm{Ker}\, f = \{v \in V | f(v) = o\}$$

se nazývá *jádro zobrazení* f a množina

$$\mathrm{Im}\, f = \{w \in W | \exists v \in V : f(v) = w\}$$

se nazývá *obraz zobrazení* f.

Definice je stejná, jako jsme měli pro jádro a obraz zobrazení mezi aritmetickými vektorovými prostory. Víme už, že tam je $\mathrm{Ker}\, F_A = \mathrm{Ker}\, A$, tedy množina všech řešení homogenní SLR s maticí A, a $\mathrm{Im}\, F_A = \mathrm{Im}\, A$, tedy sloupcový prostor matice A.

Pokud M je množina vektorů z V, zavedeme označení $f(M)$ pro množinu všech obrazů množiny M v zobrazení f. Speciálně $[M]^B$ je množina reprezentací všech vektorů z M vzhledem k bázi B.

Tvrzení 28. *Nechť* $f : V \to W$ *je lineární zobrazení,* $B = (v_1, \dots, v_n)$ *báze* V, $C = (w_1, \dots, w_m)$ *báze* W. *Pak*

$$[\mathrm{Ker}\, f]^B = \mathrm{Ker}[f]_B^C, \qquad [\mathrm{Im}\, f]^C = \mathrm{Im}[f]_B^C$$

Důkaz. Platí $\mathbf{x} \in [\operatorname{Ker} f]^B$, právě když $v := \sum_{i=1}^n x_i v_i \in \operatorname{Ker} f$. To zase nastává právě když $f(v) = o_W$, což je ekvivalentní s $[f(v)]^C = \mathbf{o}$ neboli $[f]_B^C [v]^B = \mathbf{o}$. Protože $\mathbf{x} = [v]^B$, je tím dokázána první část tvrzení.

Vektor $\mathbf{y}$ patří do $[\operatorname{Im} f]^C$, právě když existuje $v \in V$ takový, že $[f(v)]^C = \mathbf{y}$. Protože $[f(v)]^C = [f]_B^C [v]^B$, znamená to rovněž, že $\mathbf{y} \in \operatorname{Im}[f]_B^C$. Tím je dokázána druhá část tvrzení. □

Bude-li zobrazení f ve vztahu $[f(v)]^C = [f]_B^C [v]^B$ identita $\mathrm{Id} : V \to V$, dostáváme z něj rovnost

$$[v]^C = [\mathrm{Id}]_B^C [v]^B,$$

která vyjadřuje reprezentaci vektoru v vzhledem k bázi C pomocí reprezentace téhož vektoru vzhledem k bázi B.

Definice 25. Nechť V je vektorový prostor nad $\mathbb{F}$ dimenze n, B, C jsou báze V. Matice $[\mathrm{Id}]_B^C \in \mathbb{F}^{n \times n}$ se nazývá *matice přechodu od báze C k bázi B*.

Protože Id je izomorfismus, je podle tvrzení 27 $[\mathrm{Id}]_B^C$ regulární matice. Ze vztahu $[v]^B = [\mathrm{Id}]_C^B [v]^C$ vzniklého výměnou bází dostáváme, že $\forall v \in V$ platí

$$[v]^B = [\mathrm{Id}]_C^B [\mathrm{Id}]_B^C [v]^B$$

Pokud do první rovnosti dosadíme za vektor v i-tý prvek báze B, máme na levé straně vektor $\mathbf{e}_i$ a na pravé i-tý sloupec matice $[\mathrm{Id}]_C^B [\mathrm{Id}]_B^C$. Ta je tudíž rovna jednotkové matici E. Z posledního bodu tvrzení 4 pak plyne, že $[\mathrm{Id}]_B^C$ je regulární matice a $[\mathrm{Id}]_C^B$ je k ní matice inverzní.

Příklad 13. Spočtěme matici přechodu od báze $C = \{(0,1), (-1,1)\}$ k bázi $B = \{(1,2), (1,3)\}$ v $\mathbb{R}^2$. Zobrazíme pomocí Id prvky B a hledáme jejich reprezentaci vzhledem k C, tj. řešíme

$$\left(\begin{array}{rr|rr} 0 & -1 & 1 & 1 \\ 1 & 1 & 2 & 3 \end{array}\right) \sim \left(\begin{array}{rr|rr} 1 & 0 & 3 & 4 \\ 0 & 1 & -1 & -1 \end{array}\right)$$

Hledaná matice přechodu je tedy $[\mathrm{Id}]_B^C = \begin{pmatrix} 3 & 4 \\ -1 & -1 \end{pmatrix}$. Otestujme příkladem: reprezentace vektoru $v = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ jsou

$$[v]^C = \begin{pmatrix} 3 \\ -1 \end{pmatrix}, \quad [v]^B = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

a skutečně platí

$$[\mathrm{Id}]_B^C [v]^B = \begin{pmatrix} 3 & 4 \\ -1 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 3 \\ -1 \end{pmatrix} = [v]^C$$

KAPITOLA 8

Lineární zobrazení

Nechť V, W jsou vektorové prostory nad $\mathbb{F}$. V předchozí kapitole jsme definovali lineární zobrazení $f : V \to W$ a dva podprostory s ním spojené, jádro $\operatorname{Ker} f \leq V$ a obraz $\operatorname{Im} f \leq W$. Lineární zobrazení mezi dvěma vektorovými prostory konečné dimenze je možné reprezentovat maticí $[f]_B^C$, kde B je nějaká báze V a C je nějaká báze W. Jádro a obraz zobrazení f je možné vyjádřit pomocí jádra a obrazu matice $[f]_B^C$.

Uvažujme nyní další lineární zobrazení $g : V \to W$. Součet zobrazení f a g je zobrazení $f + g : V \to W$ definované předpisem

$$\forall v \in V : (f + g)(v) := f(v) + g(v)$$

Pro $r \in \mathbb{F}$ je r-násobek f zobrazení $rf : V \to W$ definované

$$\forall v \in V : (rf)(v) := rf(v)$$

Snadno se ověří, že jak $f + g$, tak rf jsou také lineární zobrazení.

Tvrzení 29. *Nechť V, W jsou vektorové prostory konečné dimenze nad $\mathbb{F}$, B je báze V, C je báze W, $f, g : V \to W$ lineární zobrazení, $r \in \mathbb{F}$. Pak*

$$[f + g]_B^C = [f]_B^C + [g]_B^C, \qquad [rf]_B^C = r[f]_B^C$$

Důkaz. Nechť $B = (v_1, \ldots, v_n)$. Pak z definic plyne

$$\begin{aligned}
[f + g]_B^C &= ([(f + g)(v_1)]^C \mid \ldots \mid [(f + g)(v_n)]^C) \\
&= ([f(v_1) + g(v_1)]^C \mid \ldots \mid [f(v_n) + g(v_n)]^C) \\
&= ([f(v_1)]^C + [g(v_1)]^C \mid \ldots \mid [f(v_n)]^C + [g(v_n)]^C) = [f]_B^C + [g]_B^C
\end{aligned}$$

Podobně

$$[rf]_B^C = ([rf(v_1)]^C \mid \ldots \mid [rf(v_n)]^C) = r[f]_B^C$$

□

Je-li U další vektorový prostor nad $\mathbb{F}$ a $h : U \to V$ lineární zobrazení, pak lze opět jednoduše ověřit, že složené zobrazení $f \circ h : U \to W$ je také lineární.

Tvrzení 30. *Nechť U, V, W jsou vektorové prostory konečné dimenze nad $\mathbb{F}$, D je báze U, B je báze V, C je báze W, $f : V \to W$, $h : U \to V$ lineární zobrazení. Pak*

$$[f \circ h]_D^C = [f]_B^C \; [h]_D^B$$

Důkaz. Nechť $D = (u_1, \ldots, u_p)$, $u \in U$. Pak

$$[(f \circ h)(u)]^C = [f(h(u))]^C = [f]_B^C \; [h(u)]^B = [f]_B^C \; [h]_D^B \; [u]^D$$

Dosadíme-li za vektor u i-tý prvek báze D, dostáváme, že i-tý sloupec matice $[f]_B^C \; [h]_D^B$ je roven $[(f \circ h)(u_i)]^C$, tedy i-tému sloupci matice $[f \circ h]_D^C$. □

Jinými slovy reprezentace složeného zobrazení $f \circ h$ je rovna součinu matic reprezentujících f a h. Speciálně odtud plyne *transformační formule pro reprezentaci lineárního zobrazení*:

Tvrzení 31. *Nechť V, W jsou vektorové prostory konečné dimenze nad $\mathbb{F}$, B, B' jsou báze V, C, C' jsou báze W, $f : V \to W$. Pak*

$$[f]^{C'}_{B'} = [\mathrm{Id}]^{C'}_{C}\,[f]^{C}_{B}\,[\mathrm{Id}]^{B}_{B'}$$

Důkaz. Stačí použít tvrzení 30 na složení tří lineárních zobrazení $\mathrm{Id} \circ f \circ \mathrm{Id}$:

$$[f]^{C'}_{B'} = [\mathrm{Id} \circ f \circ \mathrm{Id}]^{C'}_{B'} = [\mathrm{Id}]^{C'}_{C}\,[f]^{C}_{B}\,[\mathrm{Id}]^{B}_{B'},$$

□

Příklad 14. V minulé kapitole jsme měli zobrazení $F : \mathbb{R}^2 \to \mathbb{R}^3$ s reprezentacemi

$$[F]^{K_3}_{K_2} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ -2 & 1 \end{pmatrix}, \quad [F]^{C}_{B} = \begin{pmatrix} 1 & 1 \\ 2 & 2 \\ 0 & 1 \end{pmatrix},$$

kde $B = ((1,2)^T, (1,3)^T)$, $C = ((1,1,0)^T, (1,0,0)^T, (1,0,1)^T)$. Protože

$$[\mathrm{Id}]^{K_3}_{C} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad [\mathrm{Id}]^{B}_{K_2} = \left([\mathrm{Id}]^{K_2}_{B}\right)^{-1} = \begin{pmatrix} 3 & -1 \\ -2 & 1 \end{pmatrix},$$

můžeme ověřit $[\mathrm{Id}]^{K_3}_{C}\,[F]^{C}_{B}\,[\mathrm{Id}]^{B}_{K_2} = [F]^{K_3}_{K_2}$ dosazením:

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

Lineárnímu zobrazení $f : V \to W$ se říká také *homomorfismus vektorových prostorů V a W*. Zobrazení, pro které $V = W$, tedy se stejným zdrojovým a cílovým prostorem, se nazývá *endomorfismus vektorového prostoru V*. Pokud B, B' jsou dvě báze V, označme

- $R := [\mathrm{Id}]^{B}_{B'}$ matici přechodu od B k B'
- $A := [f]^{B}_{B}$ *matici endomorfismu vzhledem k bázi B.*
- $A' := [f]^{B'}_{B'}$ matici endomorfismu vzhledem k bázi B'.

Pak dostáváme *transformační formuli pro matici endomorfismu*:

$$A' = [f]^{B'}_{B'} = [\mathrm{Id}]^{B'}_{B}[f]^{B}_{B}[\mathrm{Id}]^{B}_{B'} = R^{-1}AR$$

Dvě matice A', A, pro něž existuje regulární matice R taková, že $A' = R^{-1}AR$, se nazývají *podobné*. Podobnost je relace ekvivalence. Charakterizovat třídy této ekvivalence, a tedy umět poznat, zda jsou dvě matice podobné, se naučíme v kapitole o Jordanově tvaru.

Připomeňme z minula, že izomorfismus je bijektivní lineární zobrazení. Složením dvou izomorfismů je opět izomorfismus. Inverzní zobrazení k izomorfismu $f : V \to W$ díky bijektivitě existuje a je také bijektivní. Abychom ukázali jeho linearitu, definujme pro nějaké $w_1, w_2 \in W$ vektory $v_1 = f^{-1}(w_1)$ a $v_2 = f^{-1}(w_2)$. Pak pro $r, s \in \mathbb{F}$

$$f(rv_1 + sv_2) = rf(v_1) + sf(v_2) = rw_1 + sw_2, \text{ tedy}$$
$$f^{-1}(rw_1 + sw_2) = rv_1 + sv_2 = rf^{-1}(w_1) + sf^{-1}(w_2).$$

Dva vektorové prostory, mezi nimiž existuje izomorfismus, se nazývají *izomorfní*. Izomorfnost je relace ekvivalence vektorových prostorů. Už v důsledku 5 jsme ukázali, že izomorfní vektorové prostory mají stejnou dimenzi. V následující větě ukážeme, že pro vektorové prostory konečné dimenze platí i opačná implikace.

Věta 7. *Dva vektorové prostory V, W nad $\mathbb{F}$ konečné dimenze jsou izomorfní, právě když mají stejnou dimenzi.*

Důkaz. Pokud $\dim V = \dim W =: n$, pak zvolme bázi B ve V, bázi C ve W a označme

$$g := [\ ]^B : V \to \mathbb{F}^n$$
$$h := [\ ]^C : W \to \mathbb{F}^n$$

Protože g i h jsou izomorfismy, jsou izomorfismem i zobrazení

$$h^{-1} : \mathbb{F}^n \to W$$
$$h^{-1} \circ g : V \to W$$

Tedy V a W jsou izomorfní. □

Nechť V, W jsou dva vektorové prostory nad $\mathbb{F}$. Množina všech homomorfismů z V do W je vektorový prostor s operacemi součtu homomorfismů a násobení homomorfismu skalárem z $\mathbb{F}$, značí se $\mathrm{Hom}(V, W)$. Ověření, že je splněna definice 11, necháváme na čtenáři.

Věta 8. *Pokud* $\dim V = n$ *a* $\dim W = m$, *pak* $\dim \mathrm{Hom}(V, W) = mn$.

Důkaz. Zvolme ve V bázi B a v W bázi C. Zobrazení

$$[\ ]_B^C : \mathrm{Hom}(V, W) \to \mathbb{F}^{m\times n},$$

které přiřazuje homomorfismu f jeho reprezentaci $[f]_B^C$, je lineární a bijektivní, tedy izomorfismus. Protože prostor $\mathbb{F}^{m\times n}$ má bázi například $\{E_{ij} | i \in \{1, \dots, m\}, j \in \{1, \dots, n\}\}$ o mn prvcích, je $\dim \mathbb{F}^{m\times n} = mn$ a tedy také $\dim \mathrm{Hom}(V, W) = mn$. □

Vektorový prostor všech endomorfismů prostoru V se místo $\mathrm{Hom}(V, V)$ častěji značí symbolem $\mathrm{End}(V)$, $\dim \mathrm{End}(V) = n^2$. Prostý homomorfismus se označuje slovem *monomorfismus*, pokud je na, pak mu říkáme *epimorfismus*.

Věta 9 (O zadání homomorfismu hodnotami na bázi). *Nechť* V, W *jsou vektorové prostory nad* $\mathbb{F}$, $B = (v_1, \dots, v_n)$ *je báze* V *a* $M = (w_1, \dots, w_n)$ *posloupnost vektorů ve* W. *Pak existuje právě jeden* $f \in \mathrm{Hom}(V, W)$ *takový, že*

$$f(v_1) = w_1, \quad f(v_2) = w_2, \quad \dots, \quad f(v_n) = w_n$$

Navíc f *je monomorfismus, právě když* M *je lineárně nezávislá a* f *je epimorfismus, právě když* M *generuje* W.

Důkaz. Je-li $v \in V$, $\mathbf{x} := [v]^B$, pak definujme $f(v) := \sum_{i=1}^n x_i w_i$. Uvažujme další vektor $u \in V$, $\mathbf{y} := [u]^B$ a skaláry $r, s \in \mathbb{F}$. Pak z tvrzení 24 plyne

$$f(ru + sv) = \sum_{i=1}^n (ry_i + sx_i) w_i = r\left(\sum_{i=1}^n y_i w_i\right) + s\left(\sum_{i=1}^n x_i w_i\right) = rf(u) + sf(v),$$

tedy takto definované f je homomorfismus. Tím je dokázána existence požadovaného zobrazení. Jednoznačnost plyne z toho, že podle tvrzení 17 lze každý vektor $v \in V$ zapsat jako $v = \sum_{i=1}^n x_i v_i$ právě jedním způsobem, a z linearity pak dostáváme, že f mající předepsané hodnoty na bázi B splňuje

$$f\left(\sum_{i=1}^n x_i v_i\right) = \sum_{i=1}^n x_i f(v_i) = \sum_{i=1}^n x_i w_i$$

Je-li M lineárně závislá, pak lze vyjádřit nulový vektor ve W jako netriviální lineární kombinaci $\sum_{i=1}^n x_i w_i$, ale jeho vzorem v zobrazení f je nenulový vektor $\sum_{i=1}^n x_i v_i$. Tedy $\mathrm{Ker}\, f$ je nenulové a f není monomorfismus. Naopak, je-li nenulový prvek $\sum_{i=1}^n x_i v_i$ v jádru f, je $\sum_{i=1}^n x_i w_i$ netriviální lineární kombinace posloupnosti M dávající nulový vektor.

Analogický důkaz ekvivalentní podmínky epimorfismu ponecháváme čtenáři za cvičení. □

V následující větě nemusíme předpokládat, že vektorové prostory mají konečnou dimenzi.

VĚTA 10. *Nechť U, V, W jsou vektorové prostory nad $\mathbb{F}$, $f \in \mathrm{Hom}(V, W)$, $g \in \mathrm{Hom}(U, V)$. Platí*

(1) *f je monomorfismus, právě když $\mathrm{Ker}\, f = 0$*
(2) *f je epimorfismus, právě když $\mathrm{Im}\, f = W$*
(3) *Jsou-li f, g monomorfismy, pak je $f \circ g$ monomorfismus.*
(4) *Jsou-li f, g epimorfismy, pak je $f \circ g$ epimorfismus.*
(5) *Je-li $f \circ g$ monomorfismus, je g monomorfismus.*
(6) *Je-li $f \circ g$ epimorfismus, je f epimorfismus.*

DŮKAZ. Pokud pro $v_1, v_2 \in V$ platí $f(v_1) = f(v_2)$, je $f(v_1 - v_2) = o_W$ a tedy $v_1 - v_2 \in \mathrm{Ker}\, f$. Tedy $\mathrm{Ker}\, f = 0$ znamená, že $v_1 = v_2$, čili f je monomorfismus. Naopak pokud f je monomorfismus, nemůže být vzorem o_W jiný vektor než o_V, tedy $\mathrm{Ker}\, f = 0$. Tím jsme ověřili první bod, druhý plyne z definic.

Pokud f je monomorfismus a $f(g(v)) = o_W$, pak $g(v) = o_V$, a je-li i g monomorfismus, musí být $v = o_U$. S využitím prvního bodu odtud plyne třetí bod. Podobně, je-li $w \in W$ a f je epimorfismus, pak existuje $v \in V$ takový, že $f(v) = w$. Je-li g epimorfismus, pak existuje $u \in U$ takový, že $g(u) = v$. Celkově tedy pro každý $w \in W$ existuje $u \in U$ takový, že $(f \circ g)(u) = w$, čímž je dokázán čtvrtý bod.

Pro pátý bod si stačí uvědomit, že pokud g není monomorfismus, tedy dle prvního bodu existuje nenulový $u \in U$ takový, že $g(u) = o_V$, pak i $(f \circ g)(u) = o_W$, tedy i $f \circ g$ má nenulové jádro. Podobně šestý bod vyplývá z toho, že vektor, který není v obrazu f, nemůže být ani v obrazu $f \circ g$. □

VĚTA 11 (O dimenzi jádra a obrazu). *Nechť V, W jsou vektorové prostory nad $\mathbb{F}$, $\dim V = n$, $f \in \mathrm{Hom}(V, W)$. Pak*

$$\dim \mathrm{Ker}\, f + \dim \mathrm{Im}\, f = n$$

DŮKAZ. Argumentace je jednodušší, pokud předpokládáme, že $\dim W < \infty$. Zvolíme-li bázi B ve V a bázi C ve W, pak

$$\dim \mathrm{Ker}\, f = \dim[\mathrm{Ker}\, f]^B = \dim \mathrm{Ker}[f]^C_B$$
$$\dim \mathrm{Im}\, f = \dim[\mathrm{Im}\, f]^C = \dim \mathrm{Im}[f]^C_B$$

Protože n je rovno počtu sloupců matice $[f]^C_B$, plyne tvrzení z věty o hodnosti a nulitě pro tuto matici. □

Dimenzi $\mathrm{Ker}\, f$ nazýváme *nulitou homomorfismu* f, značíme $n(f)$. Dimenzi $\mathrm{Im}\, f$ nazýváme *hodností homomorfismu* f, značíme $\mathrm{rank}(f)$. Předchozí věta se tedy dá nazývat i větou o hodnosti a nulitě homomorfismu. Plyne z ní

TVRZENÍ 32. *Nechť V je vektorový prostor nad $\mathbb{F}$ konečné dimenze, $f \in \mathrm{End}(V)$. Je-li f mono- nebo epimorfismus, pak už musí být i izomorfismem.*

DŮKAZ. Pokud f je monomorfismus, je $\mathrm{Ker}\, f = 0$, tedy $n(f) = 0$. Pak ale $\mathrm{rank}(f) = n$, tedy $\dim \mathrm{Im}\, f = \dim V$. Musí být proto $\mathrm{Im}\, f = V$, neboli f je epimorfismus a tudíž i izomorfismus. Pokud f je epimorfismus, má důkaz tytéž kroky, jen v opačném pořadí. □

Izomorfismus $f \in \mathrm{End}(V)$ se nazývá *automorfismem* prostoru V.

KAPITOLA 9

Determinant

V první kapitole jsme se dotkli souvislosti vektorového součinu a určení objemu rovnoběžníka a rovnoběžnostěnu. Ve cvičení **??** jsme zavedli veličinu

$$V(\mathbf{x},\mathbf{y},\mathbf{z}) := (\mathbf{x}\times\mathbf{y}).\mathbf{z} \equiv \sum_{i=1}^{3}\sum_{j=1}^{3}\sum_{k=1}^{3}\varepsilon_{ijk}x_i y_j z_k,$$

určenou trojicí vektorů $\mathbf{x},\mathbf{y},\mathbf{z}\in\mathbb{R}^3$. Ukázali jsme, že její absolutní hodnota udává objem rovnoběžnostěnu určeného vektory $\mathbf{x},\mathbf{y},\mathbf{z}\in\mathbb{R}^3$. Důkaz byl založen na tom, že zobrazení $V:\mathbb{R}^3\times\mathbb{R}^3\times\mathbb{R}^3\to\mathbb{R}$ má následující vlastnosti:

(1) $V(r\mathbf{x}_1+s\mathbf{x}_2,\mathbf{y},\mathbf{z}) = rV(\mathbf{x}_1,\mathbf{y},\mathbf{z})+sV(\mathbf{x}_2,\mathbf{y},\mathbf{z})$
(2) $V(\mathbf{x},\mathbf{y},\mathbf{z}) = -V(\mathbf{y},\mathbf{x},\mathbf{z}) = V(\mathbf{y},\mathbf{z},\mathbf{x})$
(3) $V(\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3)=1$

První vlastnost, linearita, je hned vidět z definice V. Druhá vyplývá z toho, že $\varepsilon_{ijk} = -\varepsilon_{jik} = \varepsilon_{jki}$ pro všechny možné trojice indexů $i,j,k\in\{1,2,3\}$. Pokud zobrazení při výměně proměnných v nějaké dvojici argumentů změní znaménko, říkáme, že je v této dvojici argumentů *antisymetrické.* Je-li antisymetrické v každé dvojici argumentů, říkáme, že je *úplně antisymetrické.* Poslední vlastnost platí, protože $\varepsilon_{123}=1$. Z druhé vlastnosti plyne linearita i ve 2. a 3. argumentu a také, že $V(\mathbf{x},\mathbf{y},\mathbf{z})=0$, kdykoli se některé dva vektory z $\mathbf{x},\mathbf{y},\mathbf{z}$ rovnají.

Není těžké si rozmyslet, že zobrazení V je těmito třemi vlastnostmi určeno jednoznačně. Chvíli předstírejme, že definici V neznáme, a používejme jen vlastnosti a jejich důsledky. Pak

$$\begin{aligned}
V(\mathbf{x},\mathbf{y},\mathbf{z}) &= V(x_1\mathbf{e}_1+x_2\mathbf{e}_2+x_3\mathbf{e}_3, y_1\mathbf{e}_1+y_2\mathbf{e}_2+y_3\mathbf{e}_3, z_1\mathbf{e}_1+z_2\mathbf{e}_2+z_3\mathbf{e}_3)\\
&= \sum_{i=1}^{3}\sum_{j=1}^{3}\sum_{k=1}^{3} x_i y_j z_k V(\mathbf{e}_i,\mathbf{e}_j,\mathbf{e}_k)\\
&= x_1y_2z_3V(\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3)+x_1y_3z_2V(\mathbf{e}_1,\mathbf{e}_3,\mathbf{e}_2)+x_2y_1z_3V(\mathbf{e}_2,\mathbf{e}_1,\mathbf{e}_3)\\
&\quad + x_2y_3z_1V(\mathbf{e}_2,\mathbf{e}_3,\mathbf{e}_1)+x_3y_1z_2V(\mathbf{e}_3,\mathbf{e}_1,\mathbf{e}_2)+x_3y_2z_1V(\mathbf{e}_3,\mathbf{e}_2,\mathbf{e}_1)\\
&= x_1y_2z_3 - x_1y_3z_2 - x_2y_1z_3 + x_2y_3z_1 + x_3y_1z_2 - x_3y_2z_1
\end{aligned}$$

Druhá rovnost je důsledkem linearity ve všech třech argumentech, ve třetí jsme z $3.3.3=27$ sčítanců vynechali 21, které jsou nulové, protože se do nějaké dvojice argumentů dosazují dva stejné vektory. Nakonec využíváme druhé a třetí vlastnosti. Výsledné vyjádření $V(\mathbf{x},\mathbf{y},\mathbf{z})$ se shoduje s jeho definicí pomocí vektorového a skalárního součinu.

V této kapitole zobecníme V do libovolné dimenze n, tedy na zobrazení

$$V:\underbrace{\mathbb{R}^n\times\ldots\times\mathbb{R}^n}_{n}\to\mathbb{R}$$

Jednou z cest by bylo požadovat opět linearitu v každém argumentu, antisymetrii v každé dvojici argumentů a $V(\mathbf{e}_1,\ldots,\mathbf{e}_n)=1$. Tím bychom získali jednoznačné vyjádření $V(\mathbf{x}_1,\ldots,\mathbf{x}_n)$ pomocí složek vektorů $\mathbf{x}_i$, a znaménka v tomto vyjádření by definovala hodnoty $\varepsilon_{i_1\ldots i_n}$. Zobecněné vlastnosti (1),(2) a (3) pak zaručují, že je V

rozumným vyjádřením objemu rovnoběžnostěnu definovaného vektory $\mathbf{x}_1, \dots, \mathbf{x}_n$. My se vydáme opačnou cestou: nejprve zobecníme veličinu ε_{ijk} a vlastnosti (1),(2) a (3) a s nimi i souvislost s objemem získáme jako důsledky.

1. Permutace

Uvažujme nějakou konečnou množinu, například $M = \{1, 2, \dots, n\}$, a množinu všech bijektivních zobrazení z M do M. Pak tato množina s operací skládání zobrazení splňuje definici 9 a je to tedy grupa. Neutrální prvek této grupy, tedy zobrazení identita na M, značíme id.

Definice 26. Grupa všech bijektivních zobrazení množiny $\{1, 2, \dots, n\}$ do sebe s operací skládání se nazývá *symetrická grupa* a značí se S_n. Její prvky se nazývají *permutace*.

Dva základní způsoby zápisu permutace si můžeme ukázat na příkladu

$$\pi := 1 \quad 2 \quad 3 \quad 4 \quad 5 \quad 6 \quad 7 \;\equiv \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 3 & 5 & 7 & 1 & 2 & 6 & 4 \end{pmatrix}$$

V tabulkovém zápisu je každý sloupec tvořen dvojicí vzor, obraz, tedy například $\pi(4) = 1$. Šipkový zápis můžeme přepsat jako

$$\begin{tikzcd} 1 \arrow[r] & 3 \arrow[d] & 2 \arrow[d, bend left] & 6 \arrow[loop below] \\ 4 \arrow[u] & 7 \arrow[l] & 5 \arrow[u, bend left] & \end{tikzcd}$$

Definice 27. Nechť $\pi \in S_n$. Prvek $i \in \{1, \dots, n\}$ nazýváme *samodružný*, pakliže $\pi(i) = i$. Permutaci π nazýváme *cyklus délky k*, pokud existuje posloupnost $(i_1, \dots, i_k)$ z $\{1, \dots, n\}$ taková, že pro všechna $j \in \{1, \dots, k-1\}$ je $\pi(i_j) = i_{j+1}$ a $\pi(i_k) = i_1$, a že všechny prvky $\{1, \dots, n\}$ mimo tuto posloupnost jsou samodružné. Cyklus délky 2 nazýváme *transpozice*.

V pojmosloví i v zápise bývá běžné ztotožňovat cyklus délky alespoň 2 s příslušnou posloupností nesamodružných prvků. Cykly jsou *nezávislé*, pokud tyto posloupnosti neobsahují žádný společný prvek. Permutaci π z příkladu pak můžeme zapsat jako (1374)(25), tedy jako rozklad na nezávislé cykly délky alespoň 2, nebo jako (1374)(25)(6), pokud vypíšeme i samodružné prvky coby cykly délky 1. Je snadné si rozmyslet, že takový rozklad má každá permutace a že je určen jednoznačně. Můžeme ho chápat buď jako množinu cyklů v permutaci se vyskytujících, nebo jako složení příslušných permutací, v příkladu $(1374) \circ (25)$.

Permutace obecně nekomutují, to je vidět třeba na příkladu transpozic z S_3

$$(12) \circ (13) = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} = (132)$$
$$(13) \circ (12) = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} = (123)$$

Pokud ale skládáme nezávislé cykly, na pořadí nezáleží, permutaci π tedy lze zapsat i jako $(25) \circ (1374)$.

Podobně jako jsme zapsali cyklus (i_1, i_2, i_3) délky 3 jako složení dvou transpozic $(i_1, i_3) \circ (i_1, i_2)$, můžeme obecný cyklus délky k zapsat jako

$$(i_1, i_2, \dots, i_k) = (i_1, i_k) \circ (i_1, i_2, \dots, i_{k-1}) = \dots = (i_1, i_k) \circ (i_1, i_{k-1}) \circ \dots \circ (i_1, i_2),$$

tedy jako složení $k - 1$ transpozic. Takových rozkladů je mnoho, například

$$(1374) = (14) \circ (17) \circ (13) = (13) \circ (37) \circ (74) = (47) \circ (43) \circ (41) = \dots$$

Každou permutaci tedy umíme zapsat jako složení transpozic, např $\pi = (14)\circ(17)\circ(13)\circ(25)$. Budeme chtít na základě takového rozkladu přiřadit permutaci hodnotu $+1$ nebo -1, které bude odpovídat hodnotě ϵ_{ijk} z úvodu kapitoly.

Lemma 4. *Je-li $\pi \in S_n$ permutace a $(ij) \in S_n$ transpozice, pak označme p, resp. q počet cyklů sudé délky v rozkladu permutace π, resp. $\pi \circ (ij)$ na nezávislé cykly. Pak $|p-q| = 1$.*

Důkaz. Rozebereme dva případy. Pokud jsou i, j součástí dvou různých nezávislých cyklů v rozkladu π, zapišme tyto cykly jako $(i, i_2, \dots, i_r)$ a $(j, j_2, \dots, j_s)$, $r, s \in \mathbb{N}$. Pak

$$(i, i_2, \dots, i_r) \circ (j, j_2, \dots, j_s) \circ (ij) = (i, j_2, \dots, j_s, j, i_2, \dots, i_r)$$

Tedy ze dvou cyklů délek r, s vznikl jeden cyklus délky $r+s$. Ostatní cykly v rozkladu π se složením s (ij) nemění. Jsou-li r, s obě sudá, je i $r+s$ sudé. Je-li právě jedno z r, s liché, je $r+s$ liché. V obou případech počet cyklů sudé délky klesl o 1. Jsou-li r, s obě lichá, je $r+s$ sudé a počet cyklů sudé délky vzrostl o 1. Tedy p a q se liší vždy o 1.

Jsou-li naopak i, j součástí téhož cyklu v π, pak pro tento cyklus platí

$$(i, j_2, \dots, j_s, j, i_2, \dots, i_r) \circ (ij) = (i, i_2, \dots, i_r) \circ (j, j_2, \dots, j_s)$$

a můžeme použít tytéž úvahy jako v předchozím případě. □

Definice 28. Nechť $\pi \in S_n$ je permutace a $\pi_1, \dots, \pi_k \in S_n$ jsou transpozice takové, že $\pi = \pi_1 \circ \pi_2 \circ \dots \circ \pi_k$. Pak definujeme *znaménko permutace* π jako $\operatorname{sgn}\pi := (-1)^k$. Pokud $\operatorname{sgn}\pi = +1$, mluvíme o *sudé permutaci*, pokud $\operatorname{sgn}\pi = -1$, je π *lichá permutace*.

Permutaci π lze zapsat jako složení transpozic mnoha způsoby. To, že všechny vedou ke stejné hodnotě znaménka, a tedy že je definice korektní, lze ukázat s pomocí lemmatu 4. Protože π_1 obsahuje právě jeden cyklus sudé délky a složení s transpozicí podle lemmatu mění počet cyklů sudé délky právě o 1, nahoru, či dolů, je parita (sudost či lichost) počtu cyklů sudé délky π stejná jako parita čísla k. To ale bude platit pro jakýkoli jiný zápis π jako složení transpozic $\pi'_1 \circ \dots \circ \pi'_l$, tedy parita čísla l je stejná jako parita k.

Důsledek 7. *Nechť $\pi, \rho \in S_n$ a p je počet cyklů sudé délky v rozkladu permutace π na nezávislé cykly.*

(1) $\operatorname{sgn}\pi = (-1)^p$.
(2) $\operatorname{sgn}(\pi \circ \rho) = \operatorname{sgn}\pi \operatorname{sgn}\rho$
(3) $\operatorname{sgn}\pi^{-1} = \operatorname{sgn}\pi$

Příklad 15. Grupa S_3 obsahuje $3! = 6$ prvků: identitu, tři transpozice a dva cykly délky 3. Jejich znaménka jsou

$$\begin{array}{ccc} \operatorname{sgn}\mathrm{id} = +1 & \operatorname{sgn}(12) = -1 & \operatorname{sgn}(123) = 1 \\ \operatorname{sgn}(13) = -1 & \operatorname{sgn}(23) = -1 & \operatorname{sgn}(132) = 1 \end{array}$$

Pokud definujeme $\varepsilon_{\pi(1)\pi(2)\pi(3)} := \operatorname{sgn}\pi$ a $\varepsilon_{ijk} = 0$ jindy, dostáváme přesně veličinu ε z definice zobrazení V.

2. Determinant a jeho výpočet

Definice 29. Nechť $A \in \mathbb{F}^{n\times n}$. *Determinantem* matice A rozumíme číslo

$$\det A := \sum_{\pi \in S_n} \operatorname{sgn}\pi \; a_{1\pi(1)} a_{2\pi(2)} \dots a_{n\pi(n)}$$

Determinant $n \times n$ matice je tedy součet $n!$ sčítanců, z nichž každý je až na znaménko součinem n-tice elementů matice, z nichž žádné dva neleží ve stejném řádku ani sloupci. Pro 2×2 a 3×3 matice lze definici použít přímo k výpočtu:

$$\begin{aligned}
\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} &= \operatorname{sgn} \mathrm{id}\, a_{11}a_{22} + \operatorname{sgn}(12) a_{12}a_{21} \\
&= a_{11}a_{22} - a_{12}a_{21} \\
\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} &= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\
&\quad - a_{11}a_{23}a_{32} - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33}
\end{aligned}$$

Pro $n > 3$ je už počet potřebných operací neprakticky vysoký. Důležitý speciální případ je dolní trojúhelníková matice, v níž $a_{ij} = 0$ pro $i < j$. Pro každou permutaci $\pi \neq \mathrm{id}$ platí, že $\pi(i) > i$ pro nejmenší prvek $\{1, \ldots, n\}$, který není vůči π samodružný. Pak ale ze sčítanců v definici determinantu zbyde jen ten první $a_{11}a_{22} \ldots a_{nn}$, ostatní obsahují vždy alespoň jeden činitel $a_{i\pi(i)}$ rovný nule. Stejně tak i determinant horní trojúhelníkové matice je roven součinu jejích diagonálních prvků. Plyne to i z následující věty:

Věta 12. *Nechť $A \in \mathbb{F}^{n \times n}$. Pak $\det A = \det A^T$.*

Důkaz. Každá permutace π má k sobě jednoznačně přiřazenou permutaci inverzní $\rho := \pi^{-1}$, v definici determinantu tedy můžeme sčítat i přes ni:

$$\begin{aligned}
\det A &= \sum_{\pi \in S_n} \operatorname{sgn}(\pi) a_{1\pi(1)} a_{2\pi(2)} \ldots a_{n\pi(n)} \\
&= \sum_{\rho \in S_n} \operatorname{sgn}(\rho^{-1}) a_{1\rho^{-1}(1)} a_{2\rho^{-1}(2)} \ldots a_{n\rho^{-1}(n)}
\end{aligned}$$

Z důsledku 7 víme, že $\operatorname{sgn} \rho = \operatorname{sgn} \rho^{-1}$. Množina uspořádaných dvojic $\{(1, \rho^{-1}(1)), (2, \rho^{-1}(2)), \ldots, (n, \rho^{-1}(n))\}$ obsahuje tytéž prvky jako množina $\{(\rho(1), 1), (\rho(2), 2), \ldots, (\rho(n), n)\}$, jen v jiném pořadí. Tedy

$$\det A = \sum_{\rho \in S_n} \operatorname{sgn}(\rho) a_{\rho(1)1} a_{\rho(2)2} \ldots a_{\rho(n)n} = \sum_{\rho \in S_n} \operatorname{sgn}(\rho) a^T_{1\rho(1)} a^T_{2\rho(2)} \ldots a^T_{n\rho(n)},$$

což je z definice rovno $\det A^T$. □

Ukažme nyní souvislost determinantu se zobrazením V z úvodu kapitoly. Pro $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^3$ definujme $A := (\mathbf{x}|\mathbf{y}|\mathbf{z})$. Pak z věty plyne, že

$$\begin{aligned}
\det A = \det A^T &= \sum_{\rho \in S_3} \operatorname{sgn}(\rho) a_{\rho(1)1} a_{\rho(2)2} a_{\rho(3)3} = \sum_{\rho \in S_3} \operatorname{sgn}(\rho) x_{\rho(1)} y_{\rho(2)} z_{\rho(3)} \\
&= \sum_{i=1}^{3} \sum_{j=1}^{3} \sum_{k=1}^{3} \varepsilon_{ijk} x_i y_j z_k = V(\mathbf{x}, \mathbf{y}, \mathbf{z})
\end{aligned}$$

Determinant je tedy skutečně zobecněním zobrazení V a vyplatí se na něj nahlížet jako na funkci, která přiřazuje n-prvkové posloupnosti vektorů $\mathbf{a}_1, \ldots, \mathbf{a}_n$ z $\mathbb{F}^n$ číslo $\det(\mathbf{a}_1| \ldots |\mathbf{a}_n)$. Ze tří vlastností úplně charakterizujících zobrazení V i jeho zobecnění do vyšší dimenze vidíme tu třetí ihned:

$$V(e_1, \ldots, e_n) = \det(\mathbf{e}_1| \ldots |\mathbf{e}_n) = 1$$

Úplnou antisymetrii a linearitu v každém argumentu ukážeme v následujícím tvrzení:

Tvrzení 33. *Nechť $i \in \{1, \ldots, n\}$, $\mathbf{a}_1, \ldots, \mathbf{a}_n, \mathbf{a}'_i \in \mathbb{F}^n$, $r, r' \in \mathbb{F}$, $\rho \in S_n$. Pak*

(1) $\det(\mathbf{a}_1|\dots|r\mathbf{a}_i{+}r'\mathbf{a}'_i|\dots|\mathbf{a}_n) = r\det(\mathbf{a}_1|\dots|\mathbf{a}_i|\dots|\mathbf{a}_n){+}r'\det(\mathbf{a}_1|\dots|\mathbf{a}'_i|\dots|\mathbf{a}_n)$
(2) $\det(\mathbf{a}_{\rho(1)}|\dots|\mathbf{a}_{\rho(n)}) = \operatorname{sgn}(\rho)\det(\mathbf{a}_1|\dots|\mathbf{a}_n)$

Důkaz. Pro první tvrzení stačí jen roznásobit každý člen sumy

$$\sum_{\rho\in S_n}\operatorname{sgn}(\rho)a_{\rho(1)1}\dots(ra_{\rho(i)i}+r'a'_{\rho(i)i})\dots a_{\rho(n)n}$$

Druhé plyne z úprav využívajících důsledek 7 a skutečnost, že množiny $\{\pi|\pi\in S_n\}$ a $\{\rho\circ\pi|\pi\in S_n\}$ jsou totožné:

$$\begin{aligned}
\det(\mathbf{a}_{\rho(1)}|\dots|\mathbf{a}_{\rho(n)}) &= \sum_{\pi\in S_n}\operatorname{sgn}\pi\ a_{1,\rho(\pi(1))}a_{2,\rho(\pi(2))}\dots a_{n,\rho(\pi(n))}\\
&= \operatorname{sgn}\rho^{-1}\sum_{\pi\in S_n}\operatorname{sgn}(\rho\circ\pi)\ a_{1,(\rho\circ\pi)(1)}a_{2,(\rho\circ\pi)(2)}\dots a_{n,(\rho\circ\pi)(n)}\\
&= \operatorname{sgn}\rho\sum_{\tau\in S_n}\operatorname{sgn}\tau\ a_{1,\tau(1)}a_{2,\tau(2)}\dots a_{n,\tau(n)}
\end{aligned}$$

□

Důsledek 8. *Nechť $A\in\mathbb{F}^{n\times n}$, $r\in\mathbb{F}$. Pak*

(1) *Má-li A dva sloupce stejné nebo jeden ze sloupců nulový, pak* $\det A = 0$.
(2) *ESÚ typu přičtení r-násobku sloupce do jiného sloupce nemění determinant.*
(3) *ESÚ typu násobení sloupce číslem r násobí celý determinant číslem r.*
(4) *ESÚ typu prohození dvou sloupců obrací znaménko determinantu.*
(5) *Platí i analogická tvrzení pro řádky a EŘÚ.*

Důkaz. První a čtvrté tvrzení plynou z druhé části tvrzení 33, třetí z jeho první části. Druhé dostaneme z rovnosti

$$\det(\mathbf{a}_1|\dots|\mathbf{a}_i+r\mathbf{a}_j|\dots|\mathbf{a}_n) = \det(\mathbf{a}_1|\dots|\mathbf{a}_i|\dots|\mathbf{a}_n) + r\det(\mathbf{a}_1|\dots|\mathbf{a}_j|\dots|\mathbf{a}_n),$$

v níž je druhý člen na pravé straně nula z první části tohoto tvrzení. Pátý bod plyne z věty 12. □

Determinant matice můžeme tedy efektivně vypočítat pomocí posloupnosti EŘÚ a ESÚ, které matici převedou na jednodušší, nejlépe horní či dolní trojúhelníkovou matici. Zkusme si výpočet determinantu na příkladě, v rámci nějž rovněž začneme používat obvyklé značení determinantu matice nahrazením závorky svislou čárou.

$$\begin{aligned}
&\begin{vmatrix}7&2&1&4\\3&6&-9&3\\2&4&0&3\\5&-2&3&0\end{vmatrix} = 3\begin{vmatrix}7&2&1&4\\1&2&-3&1\\2&4&0&3\\5&-2&3&0\end{vmatrix} = 6\begin{vmatrix}7&1&1&4\\1&1&-3&1\\2&2&0&3\\5&-1&3&0\end{vmatrix} =\\
&= -6\begin{vmatrix}1&7&1&4\\1&1&-3&1\\2&2&0&3\\-1&5&3&0\end{vmatrix} = -6\begin{vmatrix}1&7&1&4\\0&-6&-4&-3\\0&-12&-2&-5\\0&12&4&4\end{vmatrix} = 24\begin{vmatrix}1&1&7&4\\0&4&6&3\\0&2&12&5\\0&1&3&1\end{vmatrix} =\\
&= -24\begin{vmatrix}1&1&7&4\\0&1&3&1\\0&0&6&3\\0&0&-6&-1\end{vmatrix} = -72\begin{vmatrix}1&1&7&4\\0&1&3&1\\0&0&2&1\\0&0&0&2\end{vmatrix} = -288
\end{aligned}$$

Věta 13. *Matice $A\in\mathbb{F}^{n\times n}$ je regulární, právě když* $\det(A)\neq 0$.

Důkaz. Regularita i nenulovost determinantu se zachovávají pomocí EŘÚ a ESÚ, stačí se tedy dívat jen na matici v odstupňovaném tvaru. Ta je regulární, právě když má všechny prvky na hlavní diagonále nenulové, což nastává právě když je nenulový determinant. $\square$

Minorem nebo též *subdeterminantem* rozumíme determinant nějaké čtvercové podmatice, tj. matice vzniklé vynecháním některých řádků a sloupců.

Tvrzení 34. *Nechť $A \in \mathbb{F}^{m\times n}$. Pak* $\operatorname{rank}(A) = k$, *právě když největší řád nenulového minoru A je k.*

Důkaz. Je-li $\operatorname{rank}(A) = k$, pak lze ze sloupců A vybrat nějakou bázi $(\mathbf{a}_{i_1}, \dots, \mathbf{a}_{i_k})$ prostoru $\operatorname{Im} A$. Protože $A' = (\mathbf{a}_{i_1}|\dots|\mathbf{a}_{i_k})$ má hodnost k, lze vybrat k-prvkovou bázi jejího řádkového prostoru. Výsledná $k\times k$ podmatice je regulární a má tedy nenulový determinant. Kdyby v A existoval nenulový minor vyššího řádu, pak by jeho sloupce byly lineárně nezávislé a tudíž by musely být lineárně nezávislé i příslušné sloupce matice A. To je ale v rozporu s tím, že každá lineárně nezávislá posloupnost $\operatorname{Im} A$ má nejvýše k prvků. $\square$

Označme A_{ij} podmatici matice A vzniklou vyškrtnutím i-tého řádku a j-tého sloupce.

Věta 14 (Laplaceův rozvoj podle sloupce). *Nechť $A \in \mathbb{F}^{n\times n}$, $j \in \{1,\dots,n\}$. Pak*

$$\det A = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det(A_{ij})$$

Důkaz. Protože $\mathbf{a}_j = \sum_{i=1} a_{ij}\mathbf{e}_i$, dostáváme z linearity v j-tém sloupci

$$\det A = \sum_{i=1}^{n} a_{ij} \det(\mathbf{a}_1|\dots|\mathbf{a}_{j-1}|\mathbf{e}_i|\mathbf{a}_{j+1}|\dots|\mathbf{a}_n)$$

U determinantu v i-tém sčítanci je třeba $n-j$ transpozic na přesunutí j-tého sloupce na poslední pozici a $n-i$ transpozic na přesunutí i-tého řádku na poslední pozici, determinant se tím vynásobí faktorem $(-1)^{n-j+n-i} = (-1)^{i+j}$. Matice A', která takto vznikne, má podmatici A'_{nn} rovnu A_{ij}, a poslední sloupec $\mathbf{e}_n$. Její determinant je s využitím věty 12 roven

$$\det A' = \sum_{\pi\in S_n} \operatorname{sgn}\pi\ a'_{\pi(1),1}\dots a'_{\pi(n),n} \quad = \sum_{\substack{\pi\in S_n\\ \pi(n)=n}} \operatorname{sgn}\pi\ a'_{\pi(1),1}\dots a'_{\pi(n-1),n-1}a'_{n,n},$$

kde jsme se v posledním výrazu mohli omezit jen na permutace splňující $\pi(n) = n$, protože pro ostatní je $a'_{\pi(n),n} = 0$. Ale $a'_{n,n} = 1$ a $\operatorname{sgn}\pi = \operatorname{sgn}\pi'$ pro $\pi' \in S_{n-1}$ vzniklou zúžením π na množinu $\{1,\dots,n-1\}$. Tedy

$$\det A' = \sum_{\pi'\in S_{n-1}} \operatorname{sgn}\pi'\ a'_{\pi'(1),1}\dots a'_{\pi'(n-1),n-1} = \det(A'_{nn}) = \det(A_{ij})$$

Tedy determinant v i-tém sčítanci je roven $(-1)^{i+j}\det(A_{ij})$. $\square$

Laplaceův rozvoj je vlastně rekurentní předpis pro determinant. Z věty 12 plyne, že jej lze provést i podle kteréhokoli řádku. Kombinace úprav a rozvoje podle řádku či sloupce s velkým počtem nul je obvykle nejrychlejší cesta k výpočtu determinantu:

$$\begin{vmatrix} 0 & 0 & 2 & 0 & 3 \\ 2 & 0 & 0 & 3 & 0 \\ 3 & 2 & 0 & 1 & 0 \\ 7 & 1 & 5 & 3 & 0 \\ 4 & 2 & 0 & 2 & 1 \end{vmatrix} = 2 \begin{vmatrix} 2 & 0 & 3 & 0 \\ 3 & 2 & 1 & 0 \\ 7 & 1 & 3 & 0 \\ 4 & 2 & 2 & 1 \end{vmatrix} - 5 \begin{vmatrix} 0 & 0 & 0 & 3 \\ 2 & 0 & 3 & 0 \\ 3 & 2 & 1 & 0 \\ 4 & 2 & 2 & 1 \end{vmatrix} =$$

$$2 \begin{vmatrix} 2 & 0 & 3 \\ 3 & 2 & 1 \\ 7 & 1 & 3 \end{vmatrix} + 15 \begin{vmatrix} 2 & 0 & 3 \\ 3 & 2 & 1 \\ 4 & 2 & 2 \end{vmatrix} = 2 \begin{vmatrix} 2 & 0 & 3 \\ -11 & 0 & -5 \\ 7 & 1 & 3 \end{vmatrix} + 15 \begin{vmatrix} 2 & 0 & 3 \\ -1 & 0 & -1 \\ 4 & 2 & 2 \end{vmatrix}$$

$$= -2 \begin{vmatrix} 2 & 3 \\ -11 & -5 \end{vmatrix} - 30 \begin{vmatrix} 2 & 3 \\ -1 & -1 \end{vmatrix} = -2(-10+33) - 30(-2+3) = -76$$

KAPITOLA 10

Aplikace determinantu

Jednou z nejstarších aplikací determinantu je explicitní určení řešení soustavy lineárních rovnic, tzv. Cramerovo pravidlo:

VĚTA 15. *Nechť $A \in \mathbb{F}^{n\times n}$ je regulární matice, $\mathbf{b} \in \mathbb{F}^n$. Označme $A_{i,\mathbf{b}}$ matici, která vznikne nahrazením i-tého sloupce A vektorem $\mathbf{b}$. Pak i-tá složka vektoru $\mathbf{x}$, který je jediným řešením soustavy rovnic $A\mathbf{x} = \mathbf{b}$, je rovna*

$$x_i = \frac{\det A_{i,\mathbf{b}}}{\det A}$$

DŮKAZ. Pro řešení SLR $A\mathbf{x} = \mathbf{b}$ platí $\sum_{j=1}^n x_j \mathbf{a}_j = \mathbf{b}$. Pak

$$\begin{aligned}
\det A_{i,\mathbf{b}} &= \det(\mathbf{a}_1|\dots|\mathbf{a}_{i-1}|\sum_{j=1}^n x_j\mathbf{a}_j|\mathbf{a}_{i+1}|\dots|\mathbf{a}_n) \\
&= \sum_{j=1}^n x_j \det(\mathbf{a}_1|\dots|\mathbf{a}_{i-1}|\mathbf{a}_j|\mathbf{a}_{i+1}|\dots|\mathbf{a}_n) \\
&= x_i \det(\mathbf{a}_1|\dots|\mathbf{a}_{i-1}|\mathbf{a}_i|\mathbf{a}_{i+1}|\dots|\mathbf{a}_n) = x_i \det A
\end{aligned}$$

Využili jsme linearitu v i-tém sloupci a fakt, že determinant matice se dvěma stejnými sloupci je roven 0. Je-li A regulární, pak $\det A \neq 0$, po vydělení $\det A$ dostáváme vyjádření x_i. □

Jednoduchým příkladem užití Cramerova pravidla je

$$\begin{pmatrix} 7 & 8 \\ 2 & 3 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 4 \\ 1 \end{pmatrix} \Rightarrow x_1 = \frac{\begin{vmatrix} 4 & 8 \\ 1 & 3 \end{vmatrix}}{\begin{vmatrix} 7 & 8 \\ 2 & 3 \end{vmatrix}} = \frac{4}{5}, \quad x_2 = \frac{\begin{vmatrix} 7 & 4 \\ 2 & 1 \end{vmatrix}}{\begin{vmatrix} 7 & 8 \\ 2 & 3 \end{vmatrix}} = -\frac{1}{5}$$

Pro větší soustavy je Gaussova eliminace rychlejší, ale i tak Cramerovo pravidlo nalézá využití v teorii, případně pro výpočet konkrétní i-té složky řešení, která nás zajímá.

Ve větě o Laplaceově rozvoji se vyskytly výrazy tvaru $(-1)^{i+j}\det(A_{ij})$, jimž se říká *algebraický doplněk prvku a_{ij}*. Matice, která má na pozici ij algebraický doplněk prvku a_{ji} (**pozor, indexy jsou obráceně!**) matice A, se nazývá *matice adjungovaná k A* a značí $\operatorname{adj}(A)$. Má jednoduchý vztah k matici inverzní:

VĚTA 16. *Nechť $A \in \mathbb{F}^{n\times n}$ je regulární. Pak $A^{-1} = \frac{1}{\det A}\operatorname{adj}(A)$.*

DŮKAZ. Jediným řešením soustavy rovnic $A\mathbf{x} = \mathbf{e}_j$ je j-tý sloupec matice A^{-1}. Podle Cramerova pravidla má x_i, tedy element A^{-1} na pozici ij, hodnotu $\frac{1}{\det A}\det A_{i,\mathbf{e}_j}$. Rozvojem podle i-tého sloupce dostáváme, že $\det A_{i,\mathbf{e}_j}$ je roven

$$\det(\mathbf{a}_1|\dots|\mathbf{a}_{i-1}|\mathbf{e}_j|\mathbf{a}_{i+1}|\dots|\mathbf{a}_n) = (-1)^{j+i}\det A_{ji} = \operatorname{adj}(A)_{ij}$$

□

Příklady 3.

$$A = \begin{pmatrix} 7 & 8 \\ 2 & 3 \end{pmatrix} \Rightarrow \operatorname{adj}(A) = \begin{pmatrix} 3 & -8 \\ -2 & 7 \end{pmatrix} \Rightarrow A^{-1} = \frac{1}{5}\begin{pmatrix} 3 & -8 \\ -2 & 7 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 0 & 5 \\ 1 & 0 & 1 \end{pmatrix}^{-1} = \frac{1}{2}\begin{pmatrix} \begin{vmatrix} 0 & 5 \\ 0 & 1 \end{vmatrix} & -\begin{vmatrix} 2 & 3 \\ 0 & 1 \end{vmatrix} & \begin{vmatrix} 2 & 3 \\ 0 & 5 \end{vmatrix} \\ -\begin{vmatrix} 4 & 5 \\ 1 & 1 \end{vmatrix} & \begin{vmatrix} 1 & 3 \\ 1 & 1 \end{vmatrix} & -\begin{vmatrix} 1 & 3 \\ 4 & 5 \end{vmatrix} \\ \begin{vmatrix} 4 & 0 \\ 1 & 0 \end{vmatrix} & -\begin{vmatrix} 1 & 2 \\ 1 & 0 \end{vmatrix} & \begin{vmatrix} 1 & 2 \\ 4 & 0 \end{vmatrix} \end{pmatrix},$$

což je po vyčíslení $\frac{1}{2}\left(\begin{smallmatrix} 0 & -2 & 10 \\ 1 & -2 & 7 \\ 0 & 2 & -8 \end{smallmatrix}\right)$. Pro větší matice než 3×3 se už vzorec nevyplatí, ale opět je užitečný pro teorii a pro získání konkrétního elementu inverzní matice bez nutnosti spočítat inverzní matici celou.

Je-li A regulární matice, existuje posloupnost elementárních matic $E_1, \dots, E_k$, která převede A na matici jednotkovou:

$$E_k \dots E_2 E_1 A = E$$

Pak ale $A^{-1} = E_k \dots E_2 E_1$ a $A = E_1^{-1} E_2^{-1} \dots E_k^{-1}$, a protože inverzní matice k elementární matici je elementární, plyne z toho, že každou regulární matici lze zapsat jako součin elementárních matic. Determinant elementární matice umíme snadno určit:

- matice přičtení násobku řádku do jiného řádku má determinant 1
- matice vynásobení řádku číslem r má determinant r
- matice prohození dvojice řádků má determinant -1.

Uvažujme nyní součin AB dvou čtvercových matic, přičemž $A = E_1 E_2 \dots E_k$, kde E_i jsou elementární matice. Pak jistě

$$\det(AB) = \det(E_1(E_2 \dots E_k B)) = \det(E_1)\det(E_2 \dots E_k B),$$

protože násobení maticí E_1 je řádková úprava matice $E_2 \dots E_k B$, která změní její determinant přesně stejně, jako kdyby se vynásobil číslem $\det E_1$. Opakováním stejné úvahy zjistíme, že

$$\det(AB) = \det(E_1)\det(E_2)\dots\det(E_k)\det(B)$$
$$\det(A) = \det(E_1)\det(E_2)\dots\det(E_k)$$

Odtud plyne

Věta 17. *Nechť $A, B \in \mathbb{F}^{n\times n}$. Pak $\det AB = \det A \det B$.*

Důkaz. Je-li A regulární, plyne věta z úvah výše. Je-li A singulární, pak je singulární i AB a obě mají determinant 0. □

Důsledek 9. *Nechť $A, R \in \mathbb{F}^{n\times n}$, R regulární. Pak*

(1) $\det(R^{-1}) = \frac{1}{\det R}$
(2) $\det(R^{-1}AR) = \det A$

Důkaz. První bod plyne z

$$\det(R^{-1})\det(R) = \det(R^{-1}R) = \det E = 1,$$

druhý z

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

□

Podobné matice mají tedy stejný determinant. Pokud A je matice $[f]_B^B$ nějakého endomorfismu $f \in \mathrm{End}(V)$ vzhledem k bázi B a $A' = [f]_{B'}^{B'}$ je matice téhož endomorfismu vzhledem k bázi B', plyne z transformační formule, že $\det A = \det A'$. Následující definice je tedy korektní:

Definice 30. Nechť V je vektorový prostor konečné dimenze nad $\mathbb{F}$, B jeho báze a $f \in \mathrm{End}(V)$. Pak *determinant endomorfismu* f je determinant jeho matice $[f]_B^B$.

Uvažujme bázi $B = (v_1, \ldots, v_n)$ reálného vektorového prostoru V a posloupnost $C = (w_1, \ldots, w_n)$ vzniklou z posloupnosti B jednou z elementárních úprav definovaných nad tvrzením 19. Podle věty 9 existuje právě jeden endomorfismus $f_1 \in \mathrm{End}(V)$ takový, že $f_1(B) = C$, tedy že pro všechna $i \in \{1, \ldots, n\}$ je $f_1(v_i) = w_i$. Nazývejme jej *elementární endomorfismus.* Snadno si rozmyslíme, že matice $E_1 := [f_1]_B^B$ je elementární matice, a že $\det E_1$ je až na znaménko roven poměru objemů rovnoběžnostěnů definovaných posloupnostmi $f_1(B)$ a B. Stačí si uvědomit, že při elementární úpravě prvního i druhého typu se zachovává $n-1$ vektorů, které můžeme vzít jako základnu rovnoběžnostěnu, a změna zbývajícího vektoru výšku měřenou od této základny buď nemění (úprava prvního typu), nebo násobí číslem r (úprava druhého typu). Úprava třetího typu rovnoběžnostěn nemění, jen ho zadává posloupností vektorů v jiném pořadí.

Obecný regulární endomorfismus f s maticí $A = [f]_B^B$ můžeme s pomocí rozkladu A na součin $E_1 E_2 \ldots E_k$ elementárních matic chápat jako složení elementárních endomorfismů $f = f_1 \circ \ldots \circ f_k$. Protože $\det A = \det E_1 \ldots \det E_k$ je roven součinu koeficientů, jimiž endomorfismy f_i mění objem rovnoběžnostěnu určeného posloupností B, můžeme $\det A$ interpretovat jako podíl objemů $f(B)$ a B. Protože $\det f = \det A$ nezávisí na volbě báze B, můžeme determinant endomorfismu chápat jako koeficient, jímž se mění při zobrazení f objem libovolného rovnoběžnostěnu, nebo dokonce jakéhokoli útvaru ve V, který lze rovnoběžnostěny (infinitezimálně) pokrýt. Toto pozorování je základem teorie vícerozměrné integrace.

Definice 31. Nechť $B = (\mathbf{b}_1, \ldots, \mathbf{b}_n)$ je báze $\mathbb{R}^n$. Pokud $\det(\mathbf{b}_1 | \ldots | \mathbf{b}_n)$ je kladný, nazveme bázi B *pravotočivou*, pokud je záporný, pak *levotočivou.* Levotočivost nebo pravotočivost báze nazýváme souhrnně její *orientací.*

Z úvah výše vyplývá, že elementární úprava nemění orientaci báze, právě když je determinant příslušné elementární matice kladný. Tedy znaménko $\det f$ vyjadřuje, zda mají báze B a $f(B)$ stejnou orientaci. I v případě bází v reálném vektorovém prostoru V, který není aritmetický, můžeme zavést orientaci jako rozdělení množiny všech bází na dvě třídy ekvivalence vzhledem k relaci „existuje endomorfismus $f \in \mathrm{End}(V)$ s kladným determinantem, který zobrazuje jednu na druhou“ nebo ekvivalentně „matice přechodu mezi těmito bázemi má kladný determinant“.

Další veličinou, která se zachovává při podobnosti je *stopa matice* $A \in \mathbb{F}^{n \times n}$, definovaná jako $\mathrm{Tr}\, A = \sum_{i=1}^n a_{ii}$, tedy součet prvků na diagonále. Pro $A \in \mathbb{F}^{n \times p}$, $B \in \mathbb{F}^{p \times q}$, $C \in \mathbb{F}^{q \times n}$ platí

$$\mathrm{Tr}\, ABC = \sum_{i=1}^{n} \sum_{j=1}^{p} \sum_{k=1}^{q} a_{ij} b_{jk} c_{ki} = \sum_{j=1}^{p} \sum_{k=1}^{q} \sum_{i=1}^{n} b_{jk} c_{ki} a_{ij} = \mathrm{Tr}\, BCA,$$

tzv. *cykličnost stopy.* Odtud pak

$$\mathrm{Tr}\, R^{-1} A R = \mathrm{Tr}\, R R^{-1} A = \mathrm{Tr}\, A$$

Proto i stopu lze definovat pro libovolný endomorfismus $f \in \mathrm{End}(V)$ jako stopu $\mathrm{Tr}[f]_B^B$ jeho libovolné matice. Další maticové veličiny, které lze takto vztáhnout na příslušné endomorfismy (tzv. *invarianty*, protože nezávisí na zvolené reprezentaci) potkáme v kapitole o diagonalizaci.

KAPITOLA 11

Diagonalizace

Diagonalizace $n \times n$ matice A je hledání diagonální matice

$$D = \begin{pmatrix} \lambda_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_n \end{pmatrix} =: \operatorname{diag}(\lambda_1, \dots, \lambda_n),$$

splňující pro nějakou regulární matici R vztah

$$A = RDR^{-1}$$

Protože $D^k = \operatorname{diag}(\lambda_1^k, \dots, \lambda_n^k)$ a

$$A^k = \underbrace{RDR^{-1}RDR^{-1} \dots RDR^{-1}}_{k} = RD^kR^{-1},$$

stačí nám znalost D a R k nalezení libovolné mocniny matice A. Protože $A = [F_A]_K^K$, lze vztah $D = R^{-1}AR$ chápat jako transformační formuli

$$[F_A]_B^B = [\mathrm{Id}]_K^B \; [F_A]_K^K \; [\mathrm{Id}]_B^K,$$

pro nějakou vhodnou bázi B, zvanou *báze z vlastních vektorů.* Budeme se tedy nejprve zabývat tím, jak se taková báze dá nalézt.

1. Vlastní čísla a vlastní vektory

Definice 32. Nechť V je vektorový prostor nad $\mathbb{F}$. Číslo $\lambda \in \mathbb{F}$ je *vlastním číslem* endomorfismu $f \in \operatorname{End}(V)$, pakliže pro nějaký nenulový vektor $v \in V$ platí $f(v) = \lambda v$. Každý vektor, který toto splňuje, se nazývá *vlastním vektorem* f příslušným vlastnímu číslu λ.

Podmínku $f(v) = \lambda v$ lze přepsat jako $(f - \lambda \,\mathrm{Id})(v) = 0$. Vlastní vektory jsou tedy prvky $V_\lambda := \operatorname{Ker}(f - \lambda \,\mathrm{Id})$ neboli *vlastního podprostoru endomorfismu f příslušného vlastnímu číslu λ.*

Dle definice je λ vlastní číslo f, pakliže $\operatorname{Ker}(f - \lambda \,\mathrm{Id}) \neq 0$. Je-li $\dim V = n$ a C báze V, nastává to právě když $[f - \lambda \,\mathrm{Id}]_C^C$ je singulární matice. V označení $[f]_C^C = A$ to znamená, že

$$\det(A - \lambda E) \equiv \begin{vmatrix} a_{11} - \lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - \lambda & & \vdots \\ \vdots & & \ddots & \\ a_{n1} & \dots & & a_{nn} - \lambda \end{vmatrix} = 0$$

Příklad 16. Pro $n = 2$ a $n = 3$ můžeme přímým výpočtem ověřit, že

$$\begin{vmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{vmatrix} = (-\lambda)^2 + (a_{11} + a_{22})(-\lambda) + (a_{11}a_{22} - a_{21}a_{12})$$
$$= \lambda^2 - \operatorname{Tr} A \; \lambda + \det A$$

$$\det(A - \lambda E) = (-\lambda)^3 + (\operatorname{Tr} A)(-\lambda)^2 + (\operatorname{Tr}(\operatorname{adj}(A)))(-\lambda) + \det A$$

Definice 33. *Charakteristickým polynomem matice* $A \in \mathbb{F}^{n\times n}$ nazýváme polynom $p_A(\lambda) := \det(A - \lambda E)$. *Charakteristický polynom endomorfismu* f je $\det(f - \lambda \operatorname{Id})$, čili charakteristický polynom jeho libovolné matice $[f]^C_C$.

Charakteristický polynom $n \times n$ matice má stupeň n a koeficient $(-1)^n$ u λ^n, protože jediný sčítanec v definici determinantu $\det(A - \lambda E)$, který k λ^n přispívá, je ten odpovídající identické permutaci. Protože každý sčítanec odpovídající neidentické permutaci obsahuje alespoň jeden činitel nad diagonálou a alespoň jeden činitel pod ní, pochází i koeficient λ^{n-1} pouze ze součinu diagonálních elementů $(a_{11} - \lambda)\dots(a_{nn} - \lambda)$ a je tedy roven $(-1)^{n-1} \operatorname{Tr} A$ u λ^{n-1}. Absolutní člen $p_A(\lambda)$ je roven $p_A(0)$, tedy $\det A$.

Pro obecné $k > 0$ je koeficient u λ^{n-k} roven

$$(-1)^{n-k} \sum_{\substack{I\subset\{1,\dots,n\}\\ |I|=k}} \det(A_{II}),$$

kde A_{II} je $k \times k$ podmatice A vzniklá vynecháním všech řádků a sloupců, jejichž indexy nejsou v množině I. Speciálně koeficient u λ je roven $-\sum_{i=1}^n \det(A_{ii}) \equiv -\operatorname{Tr}(\operatorname{adj}(A))$, kde A_{ii} je $n-1 \times n-1$ podmatice vzniklá vynecháním i-tého řádku i sloupce. Nejlépe je to vidět, když si představíme každé λ na diagonále $\det(A - \lambda E)$ obarvené jinou barvou a nahlédneme, že koeficient u i-tého barevného λ je právě $\det(A_{ii})$.

Příklady 4.

(1) Zobrazení $f : P^n(x,\mathbb{R}) \to P^n(x,\mathbb{R})$ definované jako derivace polynomu $[f(p)](x) := \frac{d}{dx}p(x)$ má jediné vlastní číslo 0, vlastní podprostor je tvořen konstantními polynomy.
(2) Vlastními čísly $D = \operatorname{diag}(d_1,\dots,d_n)$ jsou $d_1,\dots,d_n$ a příslušné vlastní podprostory jsou (pro případ vzájemně různých d_i) lineární obaly prvků kanonické báze $\langle \mathbf{e}_1\rangle,\dots,\langle \mathbf{e}_n\rangle \in \mathbb{C}^n$.
(3) Pokud $Ax = \lambda x$ a A je regulární, pak $A^{-1}x = \lambda^{-1}x$, čili inverzní matice má stejné vlastní podprostory, ale převrácená vlastní čísla.
(4) Pokud $A = Q^{-1}BQ$, pak $Ax = \lambda x$ znamená $BQx = \lambda Qx$, tedy λ je vlastním číslem B s vlastním vektorem Qx. Podobné matice mají tedy stejná vlastní čísla.
(5) $\det(A^T - \lambda E) = \det(A - \lambda E)^T = \det(A - \lambda E)$, tedy A^T má stejná vlastní čísla jako A.

2. Diagonalizace matice

Vlastní čísla lze tedy nalézt jako kořeny charakteristického polynomu. Teorie se zjednoduší, když budeme moci předpokládat, že A je komplexní matice a vlastní čísla také hledáme v $\mathbb{C}$. Opíráme se přitom o tzv. Základní větu algebry:

Věta 18. *Každý nekonstantní polynom s komplexními koeficienty má alespoň jeden kořen v* $\mathbb{C}$.

Důkaz je nad rámec kurzu. Můžeme ale odvodit jednoduchý důsledek:

Důsledek 10. *Nechť* $p(x)$ *je polynom stupně* $n > 0$ *s komplexními koeficienty. Pak má* n *komplexních kořenů včetně násobností.*

Důkaz indukcí. Pro $n = 1$ zjevně platí. Nechť $x_0 \in \mathbb{C}$ je kořen $p(x) = \sum_{i=0}^n a_i x^i$. Pak $p(x) = p(x) - p(x_0) = \sum_{i=0}^n a_i(x^i - x_0^i)$, což lze zapsat jako$(x - x_0)q(x)$, kde $q(x)$ je polynom stupně $n-1$. Ten má z indukčního předpokladu $n-1$ kořenů včetně násobností. Tedy $p(x)$ má n kořenů včetně násobností.

Definice 34. Množina všech vlastních čísel matice A se nazývá její *spektrum*, značí se $\sigma(A)$. Pro $\lambda_i \in \sigma(A)$ je jeho *algebraická násobnost* definována jako násobnost λ_i coby kořenu $p_A(\lambda)$, a jeho *geometrická násobnost* jako dimenze prostoru $\operatorname{Ker}(A - \lambda_i E)$. Analogicky jsou definovány tytéž pojmy pro endomorfismus.

Lemma 5. *Nechť V je vektorový prostor nad $\mathbb{F}$ dimenze n, $f \in \operatorname{End}(V)$ a pro všechny prvky množiny $M = \{\lambda_1, \dots, \lambda_k\}$ je zvolena nějaká báze B_i vlastního prostoru V_{λ_i}. Pak je množina $B = B_1 \cup \ldots \cup B_k$ lineárně nezávislá.*

Důkaz. Uvažujme množinu $N \subset V$ nenulových vektorů, v níž každý prvek patří jinému vlastnímu prostoru V_{λ_i}. Je-li N lineárně závislá, lze z ní vybrat bázi $N' = \{v_{i_1}, \dots, v_{i_q}\}$ jejího lineárního obalu, kde $v_{i_j} \in V_{\lambda_{i_j}}$, a vyjádřit nějaký $v_j \in N \setminus N'$ jako $v_j = \sum_{p=1}^{q} r_p v_{i_p}$. Je-li $v_j \in V_{\lambda_j}$, pak $f(v_j) = \lambda_j v_j = \sum_{p=1}^{q} r_p \lambda_j v_{i_p}$, ale zároveň

$$f(v_j) = \sum_{p=1}^{q} r_p f(v_{i_p}) = \sum_{p=1}^{q} r_p \lambda_{i_p} v_{i_p}$$

Tedy $\sum_{p=1}^{q} r_p (\lambda_{i_p} - \lambda_j) v_{i_p} = 0$. Protože žádný rozdíl $\lambda_{i_p} - \lambda_j$ není 0 a i některé r_p musí být nenulové, je to ve sporu s lineární nezávislostí N'. Tedy N nemůže být lineárně závislá.

Označme nyní $B_i = (w_{i1}, \dots, w_{in_i})$ a uvažujme lineární kombinaci ve tvaru

$$\sum_{i_1=1}^{n_1} r_{1i_1} w_{1i_1} + \ldots + \sum_{i_k=1}^{n_k} r_{ki_k} w_{ki_k} = o$$

Označme jednotlivé sumy jako $v_1, \dots, v_k$, pak $v_j \in V_{\lambda_j}$. Kdyby byly některé z vektorů $v_1, \dots, v_k$ nenulové, pak bychom měli jejich lineární kombinaci rovnou nulovému vektoru, což podle úvahy z prvního odstavce nemůže nastat. Tedy musí být všechny v_j nulové, a protože posloupnosti B_j jsou lineárně nezávislé, musí být pak nulové i všechny koeficienty r_{ji_j}. Tedy B je lineárně nezávislá. □

Je-li V nad $\mathbb{C}$, $M = \sigma(f)$ a každé vlastní číslo f má stejnou algebraickou i geometrickou násobnost, pak $|B| = n$ a je to tedy na základě lemmatu báze V. Speciálně to musí nastat v případě, že jsou všechny algebraické násobnosti všech $\lambda_i \in \sigma(f)$ rovny jedné. Protože $\dim \operatorname{Ker}(f - \lambda_i \operatorname{Id})$ je vždy alespoň 1, má B alespoň n prvků, a protože je lineárně nezávislá, musí jich mít právě n.

Předpokládejme nyní, že pro $f \in \operatorname{End}(V)$ v prostoru V máme bázi $B = (v_1, \dots, v_n)$ složenou z vlastních vektorů f. Pokud $v_i \in V_\lambda$ pro nějaké $\lambda_i \in \sigma(f)$, pak $f(v_i) = \lambda_i v_i$, neboli $[f(v_i)]^B = \lambda_i \mathbf{e}_i$. Pak ale

$$[f]_B^B = \big([f(v_1)]^B \big| \ldots \big| [f(v_n)]^B\big) = \begin{pmatrix} \lambda_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_n \end{pmatrix}$$

V tomto vztahu na rozdíl od lemmatu 5 nepředpokládáme všechna λ_i vzájemně různá. Speciálně pro $f = F_A$ s bází z vlastních vektorů $B = (\mathbf{v}_1, \dots, \mathbf{v}_n)$ máme

$$\begin{aligned} A \equiv [F_A]_K^K &= [\operatorname{Id}]_B^K \; [F_A]_B^B \; [\operatorname{Id}]_K^B = \\ &= (\mathbf{v}_1 | \ldots | \mathbf{v}_n) \begin{pmatrix} \lambda_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_n \end{pmatrix} (\mathbf{v}_1 | \ldots | \mathbf{v}_n)^{-1} \end{aligned}$$

Příklad 17. Diagonalizujme matici $A = \begin{pmatrix} 1 & 1 \\ -3 & 5 \end{pmatrix}$. Char. polynom je

$$p_A(\lambda) = \lambda^2 - \operatorname{Tr} A \; \lambda + \det A = \lambda^2 - 6\lambda + 8 = (\lambda - 2)(\lambda - 4)$$

Tedy $\sigma(A) = \{2, 4\}$. Vlastní podprostory pak jsou

$$V_4 = \operatorname{Ker}\begin{pmatrix} -3 & 1 \\ -3 & 1 \end{pmatrix} = \left\langle \begin{pmatrix} 1 \\ 3 \end{pmatrix} \right\rangle, \; V_2 = \operatorname{Ker}\begin{pmatrix} -1 & 1 \\ -3 & 3 \end{pmatrix} = \left\langle \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\rangle$$

Tedy

$$A = \begin{pmatrix} 1 & 1 \\ -3 & 5 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} 4 & 0 \\ 0 & 2 \end{pmatrix} \frac{1}{(-2)} \begin{pmatrix} 1 & -1 \\ -3 & 1 \end{pmatrix}$$

Má-li A včetně násobností n vlastních čísel, pak je její charakteristický polynom rozložitelný, tj.

$$\begin{aligned} p_A(\lambda) &\equiv (-\lambda)^n + \operatorname{Tr} A(-\lambda)^{n-1} + \ldots + \det(A) \\ &= (\lambda_1 - \lambda)(\lambda_2 - \lambda) \ldots (\lambda_n - \lambda) \\ &= (-\lambda)^n + (\lambda_1 + \ldots + \lambda_n)(-\lambda)^{n-1} + \ldots + \lambda_1 \lambda_2 \ldots \lambda_n \end{aligned}$$

Z porovnání absolutních členů vidíme, že $\det A$ je roven součinu všech n vlastních čísel a $\operatorname{Tr} A$ je rovna součtu všech vlastních čísel. Charakteristický polynom, vlastní čísla a jejich algebraické i geometrické násobnosti, stopa, determinant i ostatní koeficienty charakteristického polynomu se nemění při podobnostní transformaci $A \to R^{-1}AR$. Jsou to tedy invarianty, lze je zavést i pro endomorfismy, se zachováním vztahů mezi nimi. Jeden takový, který platí pro 2×2 matice a lze jej snadno ověřit přímým výpočtem, je

$$\det A = \frac{1}{2}(\operatorname{Tr}(A)^2 - \operatorname{Tr}(A^2))$$

Obecně platí, že koeficienty $p_A(\lambda)$ lze vždy vyjádřit pomocí $\operatorname{Tr}(A^k)$, $k \leq n$.

Některé endomorfismy a matice diagonalizovatelné nejsou, protože lineárně nezávislá množina $B = B_1 \cup \ldots \cup B_k$ neobsahuje dost vektorů, aby to byla báze. Příkladem je třeba matice

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \;\Rightarrow\; \sigma(A) = \{0\}, \text{ tedy } \; V_0 = \left\langle \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\rangle,$$

nebo endomorfismus D prostoru $V = P^n(x, \mathbb{C})$, který přiřazuje polynomu $p(x)$ jeho první derivaci podle x. Pokud definujeme ve V bázi $C = (1, x, \frac{x^2}{2}, \ldots, \frac{x^n}{n!})$, spojuje ji D do řetízku

$$0 \xleftarrow{D} 1 \xleftarrow{D} x \xleftarrow{D} \tfrac{x^2}{2} \xleftarrow{D} \ldots \xleftarrow{D} \tfrac{x^{n-1}}{(n-1)!} \xleftarrow{D} \tfrac{x^n}{n!} \;,$$

z něhož lze vyčíst, že

$$[D]^C_C = \begin{pmatrix} 0 & 1 & 0 & \ldots & 0 & 0 \\ 0 & 0 & 1 & \ldots & 0 & 0 \\ \vdots & & & \ddots & 0 & 0 \\ 0 & 0 & 0 & \ldots & 1 & 0 \\ 0 & 0 & 0 & \ldots & 0 & 1 \\ 0 & 0 & 0 & \ldots & 0 & 0 \end{pmatrix}$$

Matice tohoto typu se objevují v teorii Jordanova tvaru, která řeší otázku matic, které diagonalizovatelné nejsou.

Příklad 18. Rotace v rovině $\mathbb{R}^2$ o pravý úhel má matici vzhledem ke K

$$[F_A]^K_K \equiv A = \begin{pmatrix} \cos\frac{\pi}{2} & -\sin\frac{\pi}{2} \\ \sin\frac{\pi}{2} & \cos\frac{\pi}{2} \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in \mathbb{R}^{2\times 2}$$

Ta má charakteristický polynom $\lambda^2 + 1$, tedy nemá reálná vlastní čísla, ani nemá v $\mathbb{R}^2$ žádné vlastní vektory. Chápeme-li ale F_A jako endomorfismus prostoru $\mathbb{C}^2$, pak $\sigma(A) = \{i, -i\}$ a příslušné vlastní podprostory jsou $\langle(1, -i)\rangle$, resp. $\langle(1, i)\rangle$. Diagonalizace pak má tvar

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix} \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} \frac{1}{2i} \begin{pmatrix} i & -1 \\ i & 1 \end{pmatrix}$$

Pro obecnou matici rotace se dá analogicky spočítat, že

$$\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix} \begin{pmatrix} e^{i\alpha} & 0 \\ 0 & e^{-i\alpha} \end{pmatrix} \frac{1}{2i} \begin{pmatrix} i & -1 \\ i & 1 \end{pmatrix}$$

3. Diagonalizace symetrické matice

Nechť $A \in \mathbb{C}^{m\times n}$, pak $\bar{A}$ označuje matici stejného typu, jejíž každý element $\bar{a}_{ij}$ je komplexně sdružený k odpovídajícímu elementu a_{ij} matice A. Matice $A^+ := \bar{A}^T$ se nazývá matice *hermitovsky sdružená* k matici A.

Tvrzení 35. *Nechť $A \in \mathbb{C}^{m\times n}$, $B \in \mathbb{C}^{n\times p}$. Pak*

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

Pokud $m = n$ a A je regulární, pak $(A^+)^{-1} = (A^{-1})^+$.

Důkaz. Z vlastností operací na $\mathbb{C}$ plyne, že $\overline{AB} = \bar{A}\bar{B}$. Vztah $(AB)^+ = B^+A^+$ pak plyne z tvrzení 6. Druhá část věty vyplývá z $(A^{-1})^+A^+ = (AA^{-1})^+ = E^+ = E$ a jednoznačnosti inverzní matice. □

Pokud $A^+ = A$, říkáme, že A je *hermitovská matice.*

Věta 19. *Nechť $A \in \mathbb{C}^{n\times n}$ je hermitovská matice. Pak každé její vlastní číslo je reálné.*

Důkaz. Pokud $\lambda \in \mathbb{C}$, $\mathbf{v} \in \mathbb{C}^n$ splňují $A\mathbf{v} = \lambda\mathbf{v}$, máme z tvrzení 35 i $\mathbf{v}^+A^+ = \bar{\lambda}\mathbf{v}^+$. Pak s využitím $A = A^+$

$$\lambda\mathbf{v}^+\mathbf{v} = \mathbf{v}^+A\mathbf{v} = \mathbf{v}^+A^+\mathbf{v} = \bar{\lambda}\mathbf{v}^+\mathbf{v}$$

Protože $\mathbf{v}^+\mathbf{v} = \sum_{i=1}^n |v_i|^2 \neq 0$ pro $\mathbf{v} \neq 0$, musí být $\lambda = \bar{\lambda}$. □

Speciálním případem hermitovské matice je reálná symetrická matice, tedy $A = A^T$. Pokud $\lambda \in \mathbb{R}$ je její vlastní číslo, pak i $A - \lambda E$ je reálná matice a báze jejího jádra sestává z vektorů z $\mathbb{R}^n$.

Tvrzení 36. *Nechť $A \in \mathbb{R}^{n\times n}$ je symetrická matice, λ, μ dvě různá vlastní čísla matice A, $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$ příslušné vlastní vektory, tj. $\mathbf{v} \in V_\lambda$, $\mathbf{w} \in V_\mu$. Pak $\mathbf{v} \perp \mathbf{w}$.*

Důkaz. Z $A\mathbf{v} = \lambda\mathbf{v}$ a $A\mathbf{w} = \mu\mathbf{w}$ plyne, že

$$\mu\mathbf{v}^T\mathbf{w} = \mathbf{v}^TA\mathbf{w} = \mathbf{v}^TA^T\mathbf{w} = (A\mathbf{v})^T\mathbf{w} = \lambda\mathbf{v}^T\mathbf{w}$$

Protože $\lambda \neq \mu$, musí být $\mathbf{v}^T\mathbf{w} = 0$, neboli $\mathbf{v} \perp \mathbf{w}$. □

Má-li matice A všech n vlastních čísel vzájemně různých, je v bázi z vlastních vektorů každý vektor kolmý na každý jiný. Takové bázi se říká *ortogonální.* Později ukážeme, že ortogonální bázi z vlastních vektorů lze najít pro každou symetrickou matici.

Z každé ortogonální báze $(\mathbf{v}_1, \ldots, \mathbf{v}_n)$ v $\mathbb{R}^n$ lze vytvořit bázi $\left(\frac{\mathbf{v}_1}{\|\mathbf{v}_1\|}, \ldots, \frac{\mathbf{v}_n}{\|\mathbf{v}_n\|}\right)$, která je stále ortogonální a navíc má každý vektor normu 1. Taková báze se nazývá *ortonormální.* Protože $\|\mathbf{v}\|^2 = \mathbf{v}^T\mathbf{v}$, plyne odtud

Tvrzení 37. *Nechť $B = (\mathbf{u}_1, \ldots, \mathbf{u}_n)$ je ortonormální báze $\mathbb{R}^n$, $U = [\mathrm{Id}]^K_B$. Pak $U^T = U^{-1}$.*

Důkaz. Platí $U^TU = (\mathbf{u}_1|\ldots|\mathbf{u}_n)^T(\mathbf{u}_1|\ldots|\mathbf{u}_n) = E$. □

Reálná čtvercová matice U s vlastností $U^T U = E$ se nazývá *ortogonální matice*. Pokud tedy A je reálná $n \times n$ symetrická matice, která má n vzájemně různých vlastních čísel, víme, že musí existovat ortogonální matice U taková, že $U^T AU$ je diagonální.

Několik pozorování k ortogonálním maticím:

(1) Endomorfismus $F_R \in \mathrm{End}(\mathbb{R}^n)$, kde R je ortogonální, zachovává normu, neboť $\|R\mathbf{v}\|^2 = \mathbf{v}^T R^T R\mathbf{v} = \|\mathbf{v}\|^2$.

(2) Zachovává i skalární součin, protože ten lze vyjádřit pomocí norem

$$\mathbf{v}^T\mathbf{w} = \frac{1}{4}(\|\mathbf{v}+\mathbf{w}\|^2 - \|\mathbf{v}-\mathbf{w}\|^2)$$

Stačí ověřit, že $\|\mathbf{v} \pm \mathbf{w}\|^2 = \mathbf{v}^T\mathbf{v} \pm 2\mathbf{v}^T\mathbf{w} + \mathbf{w}^T\mathbf{w}$.

(3) Naopak každá čtvercová matice U, která zachovává skalární součin, tedy $(U\mathbf{v})^T U\mathbf{w} = \mathbf{v}^T\mathbf{w}$ pro libovolné $\mathbf{v}, \mathbf{w}$, je ortogonální, protože $(U\mathbf{e}_i)^T U\mathbf{e}_j = \mathbf{e}_i^T U^T U\mathbf{e}_j$ se rovná ij-tému elementu matice $U^T U$ a $\mathbf{e}_i^T\mathbf{e}_j$ je 1 právě když $i = j$, jinak 0. Tedy například matice rotace vzhledem ke kanonické bázi je také ortogonální matice.

(4) Snadno se ověří, že množina všech $n \times n$ ortogonálních matic tvoří grupu, tedy součin dvou ortogonálních matic je ortogonální matice a $U^{-1} = U^T$ je ortogonální matice, pokud U je ortogonální matice. Pak i matice $U^T RU$, kde U, R jsou ortogonální, je ortogonální. Pokud tedy B je ortonormální báze a $U = [\mathrm{Id}]_B^K$, pak $[F_R]_B^B$ je ortogonální matice, neboli zobrazení určené ortogonální maticí vzhledem ke kanonické bázi je reprezentované ortogonální maticí i ke všem ostatním ortonormálním bázím.

(5) Chápeme-li F_R, kde R je reálná ortogonální matice, jako endomorfismus prostoru $\mathbb{C}^n$ a $\mathbf{v} \in \mathbb{C}$ je jeho nenulový vlastní vektor příslušný vlastnímu číslu $\lambda \in \mathbb{C}$, pak

$$\bar{\lambda}\lambda\mathbf{v}^+\mathbf{v} = (R\mathbf{v})^+(R\mathbf{v}) = \mathbf{v}^+R^+R\mathbf{v} = \mathbf{v}^+R^TR\mathbf{v} = \mathbf{v}^+\mathbf{v},$$

tedy musí být $\bar{\lambda}\lambda = |\lambda|^2 = 1$, neboli vlastní čísla F_R musí ležet na jednotkové kružnici v $\mathbb{C}$.

Podrobněji se těmito vlastnostmi budeme zabývat v kapitole o obecném skalárním součinu v letním semestru.

KAPITOLA 12

Direktní součet

Připomeňme z první kapitoly ortogonální projekci $P_{\mathbf{x}} \in \mathrm{End}(\mathbb{R}^n)$ do směru nenulového vektoru $\mathbf{x} \in \mathbb{R}^n$:

$$P_{\mathbf{x}}(\mathbf{y}) = \frac{\mathbf{x}.\mathbf{y}}{\|\mathbf{x}\|^2}\mathbf{x}$$

Zvolme $B_1 = (r\mathbf{x})$ bázi prostoru $\langle \mathbf{x} \rangle$ a $B_2 = (\mathbf{u}_1, \dots, \mathbf{u}_{n-1})$ nějakou bázi ortogonálního doplňku $\mathbf{x}^\perp$. Pro bázi $\mathbb{R}^n$ ve tvaru $B = (r\mathbf{x}, \mathbf{u}_1, \dots, \mathbf{u}_{n-1})$ má reprezentace $P_{\mathbf{x}}$ jednoduchý tvar

$$[P_{\mathbf{x}}]_B^B = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix}$$

Máme tedy dvojici podprostorů $W_1 := \langle \mathbf{x} \rangle$, $W_2 := \mathbf{x}^\perp$ takových, že sjednocením jejich bází B_1, B_2 vznikne báze celého prostoru. Právě taková konfigurace se popisuje pomocí direktního součtu.

1. Součet a direktní součet

Definice 35. Nechť V je vektorový prostor nad $\mathbb{F}$ a W_1, W_2 jeho dva podprostory. Pak jejich *součtem* $W_1 + W_2$ je množina všech vektorů, které lze zapsat jako součet $w_1 + w_2$, kde $w_i \in W_i$. Pokud navíc $W_1 \cap W_2 = 0$, nazýváme $W_1 + W_2$ *direktní součet podprostorů* a označujeme jej $W_1 \oplus W_2$. Pokud $W_1 \oplus W_2 = V$, pak W_2 je *doplňkem podprostoru* W_1 ve V.

Příklady 5. V $\mathbb{R}^3$ jsou co do možných dimenzí tři netriviální příklady součtu podprostorů:

- Dvě různé přímky, např. $\left\langle \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right\rangle \oplus \left\langle \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\rangle = \left\langle \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\rangle$
- Dvě různé roviny, např. $\left\langle \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\rangle + \left\langle \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\rangle = \mathbb{R}^3$, ale součet není direktní, protože průnikem rovin je jejich průsečnice.
- Rovina a přímka v ní neležící, např. $\left\langle \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\rangle \oplus \left\langle \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\rangle = \mathbb{R}^3$.

Zbylé možnosti by zahrnovaly buď triviální podprostory 0 a $\mathbb{R}^3$, nebo situaci, kdy je jeden podprostor rovný druhému podprostoru nebo je jeho podprostorem.

Tvrzení 38. *Každý prvek $W_1 \oplus W_2$ lze zapsat jako součet vektoru z $w_1 \in W_1$ a vektoru $w_2 \in W_2$ právě jedním způsobem.*

Důkaz. Stačí ukázat jednoznačnost. Pokud by existovaly $w_1, w_1' \in W_1$, $w_2, w_2' \in W_2$ takové, že $w_1 + w_2 = w_1' + w_2'$, pak $w_1 - w_1' = w_2' - w_2$. Levá strana rovnosti patří do W_1, pravá do W_2, musí tedy být obě nula. □

Sjednocení dvou podprostorů obecně není podprostor (stačí si představit dvě různé přímky v $\mathbb{R}^2$), ale součet ano:

Tvrzení 39. *Pokud W_1, W_2 jsou podprostory ve V, pak $W_1 + W_2 = \langle W_1 \cup W_2 \rangle$, tedy součet podprostorů je také podprostor. Pokud M_1 je množina generátorů W_1 a M_2 je množina generátorů W_2, pak jejich sjednocení $M_1 \cup M_2$ generuje $W_1 + W_2$.*

Důkaz. Každý součet vektorů $w_1 + w_2$ je lineární kombinací prvků z $W_1 \cup W_2$ a lze jej zapsat jako lineární kombinaci prvků z $M_1 \cup M_2$. Naopak každou lineární kombinaci prvků z $W_1 \cup W_2$ lze zapsat jako součet vektoru z W_1 a z W_2, stejně tak pro lineární kombinaci prvků z $M_1 \cup M_2$. $\square$

Tuto charakterizaci součtu můžeme rozšířit na součet libovolného množství podprostorů:

Definice 36. Nechť $\mathcal{W}$ je množina podprostorů ve vektorovém prostoru V. Pak definujeme jejich *součet* jako

$$\sum_{W \in \mathcal{W}} W := \left\langle \bigcup_{W \in \mathcal{W}} W \right\rangle$$

I zde je zřejmé, že sjednocení množin generátorů jednotlivých podprostorů generuje jejich součet. Je třeba si uvědomit, že ačkoli množina $\mathcal{W}$ může být nekonečná, v $\sum_{W \in \mathcal{W}} W$ se vyskytují jen součty konečně mnoha vektorů.

Věta 20 (O dimenzi součtu a průniku). *Nechť $W_1, W_2 \leq V$, oba konečné dimenze. Pak*

$$\dim(W_1 + W_2) = \dim W_1 + \dim W_2 - \dim(W_1 \cap W_2)$$

Důkaz. Nechť $(u_1, \dots, u_p)$ je báze $W_1 \cap W_2$. Doplňme ji podle důsledku 2 na bázi $(u_1, \dots, u_p, v_1, \dots, v_q)$ prostoru W_1 a na bázi $(u_1, \dots, u_p, w_1, \dots, w_r)$ prostoru W_2. Posloupnost $(u_1, \dots, u_p, v_1, \dots, v_q, w_1, \dots, w_r)$ generuje $W_1 + W_2$, ukážeme, že je lineárně nezávislá. Pokud

$$\sum_{i=1}^{p} a_i u_i + \sum_{i=1}^{q} b_i v_i + \sum_{i=1}^{r} c_i w_i = o$$

pro nějaké skaláry a_i, b_i, c_i, pak $\sum a_i u_i + \sum b_i v_i = -\sum c_i w_i$ je vektor, který je díky tvaru levé strany ve W_1 a zároveň díky tvaru pravé strany ve W_2. Je tedy ve $W_1 \cap W_2$ a lze jej zapsat jako $\sum_{i=1}^{p} d_i u_i$. Pak ale

$$\sum_{i=1}^{p} d_i u_i + \sum_{i=1}^{r} c_i w_i = o,$$

což z lineární nezávislosti posloupnosti $(u_1, \dots, u_p, w_1, \dots, w_r)$ znamená, že všechny koeficienty d_i, c_i jsou rovny nule. Pak ale

$$\sum_{i=1}^{p} a_i u_i + \sum_{i=1}^{q} b_i v_i = o,$$

takže z lineární nezávislosti $(u_1, \dots, u_p, v_1, \dots, v_q)$ dovodíme, že i všechny a_i, b_i jsou rovny nule. Tedy $(u_1, \dots, u_p, v_1, \dots, v_q, w_1, \dots, w_r)$ je báze $W_1 + W_2$ a stačí jen ověřit, že dimenze splňují

$$p + q + r = (p + q) + (p + r) - p$$

$\square$

Pro direktní součet tedy platí $\dim(W_1 \oplus W_2) = \dim W_1 + \dim W_2$. Navíc posloupnost (B_1, B_2), kde B_i je báze W_i, je bází $W_1 \oplus W_2$. Chceme definovat direktní součet více než dvou podprostorů tak, aby pro něj platily analogické vlastnosti. Mohli bychom zkusit požadovat, že součet tří podprostorů W_1, W_2, W_3 je direktní, pokud všechny průniky dvojic i celé trojice jsou nulové, ale to nefunguje: stačí uvažovat tři různé přímky v $\mathbb{R}^3$, z jejich průniků není poznat, zda jejich sjednocení generuje celý prostor nebo jen rovinu. Založíme tedy definici na tvrzení 38.

Definice 37. Nechť $W_1, \ldots, W_k \leq V$. Pak součet $W = W_1 + \ldots + W_k$ označíme za *direktní*, pokud lze každý vektor $w \in W$ zapsat jako součet $w_1 + \ldots + w_k$, kde $w_i \in W_i$, právě jedním způsobem. Označujeme jej pak

$$W = W_1 \oplus \ldots \oplus W_k \equiv \bigoplus_{i=1}^{k} W_i$$

Věta 21. *Nechť $W_1, \ldots, W_k \leq V$ jsou podprostory konečné dimenze. Pak*

$$\dim(W_1 \oplus \ldots \oplus W_k) = \dim W_1 + \ldots + \dim W_k$$

Důkaz. Nechť B_i je báze W_i, pak ukážeme, že $(B_1, \ldots, B_k)$ je báze $W_1 \oplus \ldots \oplus W_k$. Libovolný vektor $w \in W_1 + \ldots + W_k$ lze zapsat jako součet $w_1 + \ldots + w_k$, kde $w_i \in W_i$, a každý w_i lze zapsat jako lineární kombinaci prvků posloupnosti B_i. Tedy w lze zapsat jako lineární kombinaci prvků posloupnosti $(B_1, \ldots, B_k)$. Kdyby existoval druhý takový zápis, pak jej zapišme jako $w_1' + \ldots + w_k'$, kde w_i' je lineární kombinace prvků B_i. Protože součet $W_1 \oplus \ldots \oplus W_k$ je direktní, musí být $w_i = w_i'$ pro všechna i, a protože B_i jsou báze, jsou koeficienty lineární kombinace tvořící w_i' určeny jednoznačně. Tedy koeficienty v zápisu w jako lineární kombinace prvků posloupnosti $(B_1, \ldots, B_k)$ jsou určeny jednoznačně, a tedy dle tvrzení 17 (nebo dle definice 18) je to báze $W_1 \oplus \ldots \oplus W_k$. □

Bázi direktního součtu, která vznikne z bází jednotlivých sčítanců, budeme říkat *báze podřízená direktnímu součtu*, popřípadě *rozkladu*. Direktní součet jakožto podprostor má samozřejmě mnoho dalších bází, například $B = ((1,1),(1,-1))$ je báze prostoru $\langle(1,0)\rangle \oplus \langle(0,1)\rangle = \mathbb{R}^2$, jejíž prvky neleží v žádném z direktních sčítanců. Pokud bychom $\mathbb{R}^2$ zapsali jako direktní součet $\langle(1,1)\rangle \oplus \langle(1,-1)\rangle$, pak by B tomuto rozkladu podřízená byla.

2. Blokový zápis

Uvažujme dva vektorové prostory a jejich rozklady na direktní součet $V = V_1 \oplus V_2$ a $W = W_1 \oplus W_2$ a jejich dimenze $n = n_1 + n_2$ a $m = m_1 + m_2$. Pokud B_j je báze V_j, $B = (B_1, B_2)$ báze V, $v = v_1 + v_2$, kde $v_i \in V_i$ a $\mathbf{x}_i := [v_i]^{B_i}$, můžeme reprezentaci vektoru v zapsat *blokovým zápisem*

$$[v]^B \equiv \mathbf{x} = \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{pmatrix} \equiv \begin{pmatrix} [v_1]^{B_1} \\ [v_2]^{B_2} \end{pmatrix}$$

Aritmetický vektor o n složkách je zde zapsán jako vektor složený ze dvou bloků, jimiž jsou aritmetické vektory o n_1, resp. n_2 složkách.

Označme $\iota_1 : V_1 \to V$ lineární zobrazení *vložení* definované předpisem $\iota_1(v_1) = v_1 \in V$. Jeho reprezentace pak splňuje

$$[\iota_1]^B_{B_1} \mathbf{x}_1 = [\iota_1]^B_{B_1} [v_1]^{B_1} = [v_1]^B = \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{o} \end{pmatrix}$$

Tedy

$$[\iota_1]^B_{B_1} = \left(\begin{array}{c|c|c|c} \mathbf{e}_1 & \mathbf{e}_2 & \multirow{2}{*}{\ldots} & \mathbf{e}_{n_1} \\ \mathbf{o} & \mathbf{o} & & \mathbf{o} \end{array}\right) =: \begin{pmatrix} E_{n_1} \\ 0 \end{pmatrix},$$

kde $0 \in \mathbb{F}^{n_2 \times n_1}$. Podobně lze definovat $\iota_2 : V_2 \to V$.

Dále zaveďme lineární zobrazení *projekce* $\pi_i : W \to W_i$ tak, že pro libovolný $w \in W$, $w = w_1 + w_2$, kde $w_i \in W_i$, je $\pi_i(w) = w_i$. Pokud $C = (C_1, C_2)$ je báze W složená z bází W_1, W_2 a $[w_i]^{C_i} =: \mathbf{y}_i$, pak

$$[\pi_1]^{C_1}_C \begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{pmatrix} = [\pi_1]^{C_1}_C [w]^C = [w_1]^{C_1} = \mathbf{y}_1$$

Tedy $[\pi_1]^{C_1}_C = (\mathbf{e}_1|\dots|\mathbf{e}_{m_1}|\mathbf{o}|\dots|\mathbf{o}) =: (E_{m_1}\ 0)$, kde $0 \in \mathbb{F}^{m_1 \times m_2}$.

Uvažujme nyní $f \in \operatorname{Hom}(V, W)$ s reprezentací $[f]^C_B = A \in \mathbb{F}^{m \times n}$ a čtyři zobrazení $f_{ij} \in \operatorname{Hom}(V_i, W_j)$, $f_{ij} = \pi_i \circ f \circ \iota_j$. Pro f_{11} máme reprezentaci

$$[f_{11}]^{C_1}_{B_1} = [\pi_1]^{C_1}_C [f]^C_B [\iota_1]^B_{B_1} = \begin{pmatrix} E_{m_1} & 0 \end{pmatrix} A \begin{pmatrix} E_{n_1} \\ 0 \end{pmatrix} =: A_{11} \in \mathbb{F}^{m_1 \times n_1},$$

kde A_{11} je matice vzniklá vyškrtnutím posledních n_2 sloupců a posledních m_2 řádků z matice A. Analogicky lze zavést matice

$$\begin{aligned}
[f_{12}]^{C_1}_{B_2} &= [\pi_1]^{C_1}_C [f]^C_B [\iota_2]^B_{B_2} = \begin{pmatrix} E_{m_1} & 0 \end{pmatrix} A \begin{pmatrix} 0 \\ E_{n_2} \end{pmatrix} =: A_{12} \in \mathbb{F}^{m_1 \times n_2} \\
[f_{21}]^{C_2}_{B_1} &= [\pi_2]^{C_2}_C [f]^C_B [\iota_1]^B_{B_1} = \begin{pmatrix} 0 & E_{m_2} \end{pmatrix} A \begin{pmatrix} E_{n_1} \\ 0 \end{pmatrix} =: A_{21} \in \mathbb{F}^{m_2 \times n_1} \\
[f_{22}]^{C_2}_{B_2} &= [\pi_2]^{C_2}_C [f]^C_B [\iota_2]^B_{B_2} = \begin{pmatrix} 0 & E_{m_2} \end{pmatrix} A \begin{pmatrix} 0 \\ E_{n_2} \end{pmatrix} =: A_{22} \in \mathbb{F}^{m_2 \times n_2}
\end{aligned}$$

Ty dávají dohromady *blokový zápis matice* $A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$.

Tvrzení 40. *Nechť $n = n_1 + n_2$, $m = m_1 + m_2$, $p = p_1 + p_2$, $A \in \mathbb{F}^{m \times n}$, $B \in \mathbb{F}^{n \times p}$ s bloky $A_{ij} \in \mathbb{F}^{m_i \times n_j}$, $B_{ij} \in \mathbb{F}^{n_i \times p_j}$. Pak součin $C := AB$ má blokový zápis*

$$\begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix} = \begin{pmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \end{pmatrix}$$

Důkaz lze provést rozepsáním

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj} = \sum_{k=1}^{n_1} a_{ik} b_{kj} + \sum_{k=n_1+1}^{n} a_{ik} b_{kj}$$

a interpretací obou sum jako elementů součinu nějakých bloků matic A, B. Nebo můžeme zavést $g \in \operatorname{Hom}(U, V)$, $U = U_1 \oplus U_2$, $D = (D_1, D_2)$ podřízenou bázi, ι'_i, vložení U_i do U a π'_i, projekci V na V_i, rozmyslet si, že

$$\iota_1 \circ \pi'_1 + \iota_2 \circ \pi'_2 = \operatorname{Id} \in \operatorname{End}(V)$$

a poté odvodit, že pro $q, r \in \{1, 2\}$

$$\begin{aligned}
(f \circ g)_{qr} &= \pi_q \circ f \circ g \circ \iota'_r \\
&= \pi_q \circ f \circ (\iota_1 \circ \pi'_1 + \iota_2 \circ \pi'_2) \circ g \circ \iota'_r \\
&= (\pi_q \circ f \circ \iota_1) \circ (\pi'_1 \circ g \circ \iota'_r) + (\pi_q \circ f \circ \iota_2) \circ (\pi'_2 \circ g \circ \iota'_r) \\
&= f_{q1} \circ g_{1r} + f_{q2} \circ g_{2r},
\end{aligned}$$

odkud dostáváme

$$[(f \circ g)_{qr}]^{C_q}_{D_r} = [f_{q1}]^{C_q}_{B_1} [g_{1r}]^{B_1}_{D_r} + [f_{q2}]^{C_q}_{B_2} [g_{2r}]^{B_2}_{D_r} = A_{q1}B_{1r} + A_{q2}B_{2r}$$

Protože direktní sčítanec v rozkladu $V = V_1 \oplus V_2$ může být sám direktním součtem nějakých svých podprostorů, mohou mít i jednotlivé bloky matice samy blokovou strukturu. Tvrzení se pak snadno zobecní pro matice s libovolným blokovým členěním, pouze musí pro každý vyskytující se součin bloků $A_{ik}B_{kj}$ mít blok A_{ik} stejně sloupců, jako má blok B_{kj} řádků.

Pro $f \in \operatorname{End}(V)$, $V = V_1 \oplus \ldots \oplus V_k$ s bází $B = (B_1, \ldots, B_k)$, kde B_i je báze V_i, jsou diagonální bloky $[f]_B^B$ čtvercové matice $[f_{ii}]_{B_i}^{B_i}$, a pokud jsou zároveň mimodiagonální bloky nulové, mluvíme o *blokově diagonální matici*, jsou-li nulové jen na jedné straně diagonály, o *blokově horní/dolní trojúhelníkové matici.*

Tvrzení 41. *Nechť $A = \operatorname{diag}(A_{11}, A_{22}, \ldots, A_{kk})$ je blokově diagonální matice. Pak*

(1) $\forall p \in \mathbb{N} : A^p = \operatorname{diag}(A_{11}^p, \ldots, A_{kk}^p)$ *a jsou-li všechny A_{ii} regulární, pak i* $A^{-p} = \operatorname{diag}(A_{11}^{-p}, \ldots, A_{kk}^{-p})$
(2) $\operatorname{rank}(A) = \operatorname{rank}(A_{11}) + \operatorname{rank}(A_{22}) + \ldots + \operatorname{rank}(A_{kk})$
(3) $\operatorname{Tr}(A) = \operatorname{Tr}(A_{11}) + \operatorname{Tr}(A_{22}) + \ldots + \operatorname{Tr}(A_{kk})$
(4) $\det(A) = \det(A_{11}) \det(A_{22}) \ldots \det(A_{kk})$
(5) $\sigma(A) = \sigma(A_{11}) \cup \ldots \cup \sigma(A_{kk})$ *včetně algebraických násobností.*

Navíc části 3,4 a 5 platí i pro horní/dolní trojúhelníkovou matici s týmiž diagonálními bloky.

Důkaz. Stačí uvažovat $k = 2$ a $A = \operatorname{diag}(A', A'')$, kde $A' \in \mathbb{F}^{p \times p}$, $A'' \in \mathbb{F}^{q \times q}$, $p + q = n$. Body 1 a 3 jsou jasné, v bodě 4 lze v definici determinantu uvažovat pouze permutace $\pi \in S_n$, pro něž $\pi(i) \leq p$ pro všechna $i \in \{1, \ldots, p\}$. Rozložme π na nezávislé cykly, pak $\pi = \pi'\pi''$, kde π' je složení cyklů na množině $\{1, \ldots, p\}$ a π'' složení cyklů na množině $\{p+1, \ldots, n\}$. Označme $\rho' \in S_p$ permutaci, která vznikne zúžením π' na množinu $\{1, \ldots, p\}$ a $\rho'' \in S_q$ permutaci definovanou $\rho''(i) = \pi''(i) - p$. Pak je $\det(A) = \sum_{\pi \in S_n} \operatorname{sgn} \pi a_{1\pi(1)} \ldots a_{n\pi(n)}$

$$= \sum_{\rho' \in S_p} \sum_{\rho'' \in S_q} \operatorname{sgn}(\pi'\pi'') a_{1\rho'(1)} \ldots a_{p\rho'(p)} a_{p+1,p+\rho''(1)} \ldots a_{p+q,p+\rho''(q)}$$
$$= \sum_{\rho' \in S_p} \operatorname{sgn}(\rho') a'_{1\rho'(1)} \ldots a'_{p\rho'(p)} \sum_{\rho'' \in S_q} \operatorname{sgn}(\rho'') a''_{1\rho''(1)} \ldots a''_{q\rho''(q)}$$

Z bodu 4 plyne 5, protože vlastní čísla jsou kořeny polynomu $\det(A - \lambda E_n) = \det(A' - \lambda E_p) \det(A'' - \lambda E_q)$. Bod 2 plyne z Gaussovy eliminace provedené zvlášť na prvních p řádků a zvlášť na zbylých q řádků, kterou se A', A'' převedou do odstupňovaného tvaru, takže jejich hodnost je rovna počtu nenulových řádků z prvních p, resp ze zbylých q. Celá matice A se tím rovněž dostala do odstupňovaného tvaru, takže její hodnost je rovna počtu nenulových řádků v celé upravené matici. Zobecnění 3 horní/dolní trojúhelníkovou matici je zřejmé, pro zobecnění 4 a 5 si stačí uvědomit, že každá permutace, kterou nelze napsat jako $\pi'\pi''$ dle prvního odstavce, zobrazuje alespoň jeden index z množiny $\{1, \ldots, p\}$ do množiny $\{p+1, \ldots, n\}$ a alespoň jeden index naopak. Tedy sčítanec v definici determinantu příslušný této permutaci bude obsahovat alespoň jeden činitel z bloku nad diagonálou a alespoň jeden činitel z bloku pod diagonálou, přičemž alespoň jeden z těchto činitelů je nulový. $\square$

Příklad 19. Nechť $\mathbf{x} \in \mathbb{R}^n$, $\|\mathbf{x}\| = 1$, $P_{\mathbf{x}}(\mathbf{y}) = (\mathbf{x}.\mathbf{y})\mathbf{x}$ a $B = (B_1, B_2)$ je báze složená z bází $\langle \mathbf{x} \rangle$ a $\mathbf{x}^\perp$. Označme 0_{n-1} nulovou matici z $\mathbb{F}^{n-1 \times n-1}$ a $\mathbf{o}_{n-1}$ nulový vektor z $\mathbb{F}^{n-1}$. Pak

$$[P_{\mathbf{x}}]_B^B = \begin{pmatrix} 1 & \mathbf{o}_{n-1}^T \\ \mathbf{o}_{n-1} & 0_{n-1} \end{pmatrix}, \quad [P_{\mathbf{x}^\perp}]_B^B = \begin{pmatrix} 0 & \mathbf{o}_{n-1}^T \\ \mathbf{o}_{n-1} & E_{n-1} \end{pmatrix}$$

a matice zrcadlení $Z_{\mathbf{x}^\perp} = \operatorname{Id} - 2P_{\mathbf{x}}$ má blokový tvar

$$[Z_{\mathbf{x}^\perp}]_B^B = \begin{pmatrix} -1 & \mathbf{o}_{n-1}^T \\ \mathbf{o}_{n-1} & E_{n-1} \end{pmatrix}$$

Pokud $n = 3$ a B_2 je ortonormální báze $\mathbf{x}^\perp$, pak matice rotace okolo osy $\langle \mathbf{x} \rangle$ o úhel ϕ vzhledem k bázi B je

$$[R_{\mathbf{x},\phi}]_B^B = \begin{pmatrix} 1 & \mathbf{o}_2^T \\ \mathbf{o}_2 & [R_\phi]_{B_2}^{B_2} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\phi & -\sin\phi \\ 0 & \sin\phi & \cos\phi \end{pmatrix}$$

Příklad 20. Uvažujme blokově horní trojúhelníkovou matici $D := \begin{pmatrix} A & B \\ 0 & C \end{pmatrix}$, kde $A \in \mathbb{F}^{p\times p}$, $C \in \mathbb{F}^{q\times q}$ jsou regulární. Pak platí

$$\begin{pmatrix} A^{-1} & -A^{-1}BC^{-1} \\ 0 & C^{-1} \end{pmatrix} \begin{pmatrix} A & B \\ 0 & C \end{pmatrix} = \begin{pmatrix} A^{-1}A & A^{-1}B - A^{-1}B \\ 0 & C^{-1}C \end{pmatrix}$$

Matice napravo je jednotková, tedy D je regulární a první matice nalevo je rovna D^{-1}. Jak jsme mohli inverzní matici uhodnout? Stačí se podívat na výpočet inverzní matice k 2×2 matici s elementy a, b, c z $\mathbb{F}$, $a, c \neq 0$:

$$\begin{pmatrix} a & b \\ 0 & c \end{pmatrix}^{-1} = \frac{1}{ac}\begin{pmatrix} c & -b \\ 0 & a \end{pmatrix} = \begin{pmatrix} \frac{1}{a} & -\frac{b}{ac} \\ 0 & \frac{1}{c} \end{pmatrix}$$

Mezi tvary blokového a neblokového inverzu je zřejmá analogie, až na to, že matice A^{-1}, B, C^{-1} na rozdíl od prvků a^{-1}, b, c^{-1} nekomutují a nemusí jít ani v jiném pořadí vynásobit.

Jako ilustraci použití blokových matic v důkazech uveďme

Tvrzení 42. *Nechť $A, B \in \mathbb{F}^{n\times n}$. Pak $\sigma(AB) = \sigma(BA)$ včetně násobností.*

Důkaz. Výpočtem ověříme, že

$$\begin{pmatrix} E_n & -A \\ 0_n & E_n \end{pmatrix} \begin{pmatrix} AB & 0_n \\ B & 0_n \end{pmatrix} \begin{pmatrix} E_n & A \\ 0_n & E_n \end{pmatrix} = \begin{pmatrix} 0_n & 0_n \\ B & BA \end{pmatrix}$$

Levá strana má tvar RCR^{-1}, tedy matice C a matice C' na pravé straně jsou podobné a mají stejná vlastní čísla včetně násobností. Protože C i C' jsou obě blokově dolní trojúhelníkové, platí pro ně $\sigma(C) = \sigma(AB) \cup \sigma(0_n)$, $\sigma(C') = \sigma(BA) \cup \sigma(0_n)$ včetně násobností. $\square$

Tvrzení i důkaz se snadno zobecní na obdélníkové matice $A \in \mathbb{C}^{m\times n}$, $B \in \mathbb{C}^{n\times m}$, pouze se bude u AB a BA lišit algebraická násobnost vlastního čísla nula.

KAPITOLA 13

Skalární součin

Skalární součin umožňuje definovat na vektorovém prostoru geometrické vlastnosti jako vzdálenost, kolmost nebo úhel. To se hodí nejen v geometrii samotné, ale také třeba k zavedení konvergence posloupností funkcí nebo popisu chyb při numerickém počítání s maticemi. S komplexním skalárním součinem se setkáte v kvantové mechanice při popisu pravděpodobnosti přechodu jednoho stavu do druhého. Mnoho oblastí matematické fyziky se neobejde bez různých variant ortogonálních polynomů a jiných funkcí, příkladem mohou být Fourierovy řady nebo multipólový rozvoj. Metody jako lineární regrese ukazují využití skalárního součinu v analýze dat.

Doposud jsme měli skalární součin definovaný pouze na $\mathbb{R}^n$ vůči ortonormální bázi předpisem

$$\mathbf{x}.\mathbf{y} = \sum_{i=1}^{n} x_i y_i \equiv \mathbf{x}^T \mathbf{y}$$

Pomocí něj jsme pak zavedli velikost vektoru, vzdálenost vektorů, ortogonální doplněk, ortogonální projekci, ortogonální matici a další pojmy.

Tuto definici budeme chtít zobecnit na libovolný vektorový prostor V. Uvažujme jeho libovolné báze B, B' a vektory $u, v \in V$ vůči nim vyjádřené jako $\mathbf{x} := [u]^B$, $\mathbf{x}' := [u]^{B'}$, $\mathbf{y} := [v]^B$, $\mathbf{y}' := [v]^{B'}$. Jelikož $\mathbf{x} = R\mathbf{x}'$, kde R je matice přechodu od B k B'. Pak

$$\mathbf{x}^T \mathbf{y} = \mathbf{x}'^T R^T R \mathbf{y}'$$

Protože $R^T R$ není obecně jednotková matice, výrazy $\mathbf{x}^T\mathbf{y}$ a $\mathbf{x}'^T\mathbf{y}'$ se liší. Skalární součin vektorů $u, v \in V$ nelze zadefinovat jednoznačně pomocí jejich reprezentací, což znamená, že skalárních součinů na V musí být víc. Právě proto je nutné zadefinovat nějaké základní vlastnosti, které mají splňovat.

Definice 38. Nechť V je vektorový prostor nad $\mathbb{R}$. Zobrazení $g : V \times V \to \mathbb{R}$ je *skalární součin* na V, pakliže

(1) $\forall u, v, w \in V$, $r, s \in \mathbb{R}$:

$$g(ru + sv, w) = rg(u, w) + sg(v, w)$$
$$g(u, rv + sw) = rg(u, v) + sg(u, w)$$

(2) $\forall u, v \in V : g(u, v) = g(v, u)$
(3) $\forall u \in V, u \neq 0 : g(u, u) > 0$

Zobrazení $g : V \times V \to \mathbb{R}$ splňující pouze vlastnost (1) nazýváme *bilineární forma*, splňuje-li navíc (2), pak je *symetrická*, splňuje-li i (3), pak je *pozitivně definitní*. Ověřte sami, že zobrazení $g : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ definovaná $g(\mathbf{x}, \mathbf{y}) := \mathbf{x}^T\mathbf{y}$, které se nazývá *standardní skalární součin* na vektorovém prostoru $\mathbb{R}^n$, tyto vlastnosti splňuje. Úvaha nad definicí ukazuje, že na jiných vektorových prostorech žádný „standardní“ skalární součin zavést nemůžeme.

Další potíže jsou s rozšířením definice na komplexní vektorové prostory. Zobrazení $g : \mathbb{C}^2 \times \mathbb{C}^2 \to \mathbb{C}$ stejným předpisem $g(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T \mathbf{y}$ sice bude splňovat vlastnosti (1) a (2), ale s vlastností (3) vzniknou potíže, jak ukazují příklady

$$\begin{pmatrix} 1 & i \end{pmatrix} \begin{pmatrix} 1 \\ i \end{pmatrix} = 0, \ \begin{pmatrix} 0 & i \end{pmatrix} \begin{pmatrix} 0 \\ i \end{pmatrix} = -1, \ \begin{pmatrix} 0 & 1+i \end{pmatrix} \begin{pmatrix} 0 \\ 1+i \end{pmatrix} = 2i,$$

.

Definice 39. Nechť V je vektorový prostor nad $\mathbb{C}$. Zobrazení $g : V \times V \to \mathbb{C}$ je *skalární součin* na V, pakliže

(1) $\forall u, v, w \in V,\ r, s \in \mathbb{C}$:

$$g(ru + sv, w) = \bar{r} g(u, w) + \bar{s} g(v, w)$$
$$g(u, rv + sw) = r g(u, v) + s g(u, w)$$

(2) $\forall u, v \in V : g(u, v) = \overline{g(v, u)}$

(3) $\forall u \in V, u \neq 0 : g(u, u) > 0$

Zobrazení splňující vlastnost (1) se nazývá *seskvilineární forma*, termín vychází z latinského výrazu pro „jeden a půl“, protože zobrazení je v druhém argumentu lineární a v prvním *antilineární*. Někdy se můžete setkat i s opačnou konvencí, kdy je skalární součin definován jako antilineární v druhém a lineární v prvním argumentu. Standardní skalární součin na $\mathbb{C}^n$ je definován předpisem $g(\mathbf{x}, \mathbf{y}) := \mathbf{x}^+ \mathbf{y}$. Ověřte opět sami, že splňuje vlastnosti (1), (2) a (3).

Dvojice (V, g), kde V je vektorový prostor a g skalární součin na něm, se nazývá *prostor se skalárním součinem* nebo též *unitární prostor*.

Definice 40. Nechť V je vektorový prostor a g bilineární/seskvilineární forma na něm. Je-li $\dim V = n$ a $B = (u_i)_1^n$ báze V, pak matice $[g]_B \in \mathbb{C}^{n \times n}$, jejíž ij-tý element je $g(u_i, u_j)$, se nazývá *matice bilineární/seskvilineární formy vzhledem k bázi B*. Je-li forma skalárním součinem, mluvíme o *matici skalárního součinu vzhledem k bázi B*.

Díky vlastnosti (1) matice $G \equiv [g]_B$ skalárního součinu jednoznačně definuje jeho hodnotu na libovolné dvojici vektorů:

$$g(u, v) = g\left(\sum_{i=1}^{n} x_i u_i, \sum_{j=1}^{n} y_j u_j \right) = \sum_{i,j=1}^{n} \bar{x}_i y_j g(u_i, u_j) = \mathbf{x}^+ G \mathbf{y}$$

Vlastnost (2) znamená, že $g_{ij} = \bar{g}_{ji}$, tedy, že G je hermitovská matice, proto mluvíme o *hermitovské seskvilineární formě*. Z pozitivní definitnosti (3) pro matici G plyne, že $\mathbf{x}^+ G \mathbf{x} > 0$ pro libovolný nenulový $\mathbf{x} \in \mathbb{C}^n$. Kdyby byla matice diagonální $G = \operatorname{diag}(g_{11}, \ldots, g_{nn})$, pak by

$$\mathbf{x}^+ G \mathbf{x} = g_{11} \bar{x_1} x_1 + \ldots + g_{nn} \bar{x_n} x_n > 0,$$

což nastává právě když $g_{ii} > 0$ pro všechna $i \in \{1, \ldots, n\}$.

Definice 41. Je-li V vektorový prostor a g symetrická bilineární, resp. hermitovská seskvilineární forma na něm. Bázi $B' = (u_i')_1^n$ ve V nazýváme *polární báze formy g*, pakliže je matice $[g]_{B'}$ diagonální.

Je-li $R = [\mathrm{Id}]_{B'}^{B}$ matice přechodu od B k B', platí tedy $u_p' = \sum_{i=1}^{n} r_{ip} u_i$. Pak

$$g(u_p', u_q') = g\left(\sum_{i=1}^{n} r_{ip} u_i, \sum_{j=1}^{n} r_{jq} u_j \right) = \sum_{i,j=1}^{n} r_{pi}^{+} g_{ij} r_{jq},$$

neboli

$$[g]_{B'} \equiv G' = R^+ G R = \left([\mathrm{Id}]_{B'}^{B} \right)^+ [g]_B [\mathrm{Id}]_{B'}^{B}$$

Tedy matice bilineární, resp. seskvilineární formy se při změně báze transformují podle vztahů

$$G' = R^T GR, \text{ resp. } G' = R^+ GR$$

Uvědomte si rozdíl oproti transformační formuli pro matici endomorfismu $A' = R^{-1}AR$.

Pokud tedy dokážeme nalézt matici R takovou, že $R^T GR$, resp. $R^+ GR$ je diagonální, našli jsme i polární bázi a můžeme ověřit, zda je forma pozitivně definitní. Že to vždy jde nám zaručuje následující věta:

VĚTA 22. *Každá symetrická bilineární forma na vektorovém prostoru V konečné dimenze má polární bázi, stejně tak každá hermitovská seskvilineární.*

V důkazu se nám bude hodit tzv. polarizační identita:

TVRZENÍ 43. *Je-li g hermitovská seskvilineární forma na V nad $\mathbb{C}$, pak $\forall u, v \in V$ platí, že*

$$\operatorname{Re} g(u,v) = \frac{1}{2}(g(u+v, u+v) - g(u,u) - g(v,v))$$
$$\operatorname{Im} g(u,v) = \frac{1}{2}(g(iu+v, iu+v) - g(u,u) - g(v,v))$$

Pro symetrickou bilineární formu na V nad $\mathbb{R}$

$$g(u,v) = \frac{1}{2}(g(u+v, u+v) - g(u,u) - g(v,v))$$

DŮKAZ TVRZENÍ. Stačí rozepsat dle vlastností (1) a (2)

$$g(u+v, u+v) = g(u,u) + g(u,v) + g(v,u) + g(v,v) = g(u,u) + 2\operatorname{Re} g(u,v) + g(v,v)$$

a podobně pro imaginární část. Reálná verze se dokáže stejně. □

DŮKAZ VĚTY. Větu dokážeme pro komplexní případ indukcí podle dimenze V. Pro $\dim V = 1$ není co dokazovat, protože každá 1×1 matice je diagonální. Předpokládejme nyní, že tvrzení platí pro všechny prostory dimenze menší než n, že V má dimenzi n a že g je nenulová. Z polarizační identity plyne, že existuje vektor $u_1 \in V$, pro který $g(u_1, u_1) \neq 0$. Doplňme jej na bázi $B = (u_i)_1^n$ prostoru V a označme $G := [g]_B$. Chceme najít R takové, aby $G' = R^+ GR$ byla diagonální.

Volme nejprve matici přechodu ve speciálním tvaru a pišme blokově

$$\begin{aligned} G_1 := R_1^+ G R_1 &= \begin{pmatrix} 1 & 0 \\ \mathbf{r} & E \end{pmatrix} \begin{pmatrix} g_{11} & \mathbf{h}^+ \\ \mathbf{h} & \tilde{G} \end{pmatrix} \begin{pmatrix} 1 & \mathbf{r}^+ \\ 0 & E \end{pmatrix} \\ &= \begin{pmatrix} g_{11} & \mathbf{r}^+ g_{11} + \mathbf{h}^+ \\ g_{11}\mathbf{r} + \mathbf{h} & \tilde{G}' \end{pmatrix} \end{aligned}$$

kde $\tilde{G}' := \tilde{G} + \mathbf{r}\mathbf{h}^+ + \mathbf{h}\mathbf{r}^+ + g_{11}\mathbf{r}\mathbf{r}^+$ je hermitovská matice. Protože $g_{11} = g(u_1, u_1) \neq 0$, můžeme volbou $\mathbf{r} = -\frac{1}{g_{11}}\mathbf{h}$ zajistit, že G_1 bude blokově diagonální. Podle indukčního předpokladu existuje matice $\tilde{R}$, že $\tilde{R}^+ \tilde{G}' \tilde{R}$ je diagonální. Pak ale

$$\begin{pmatrix} 1 & 0 \\ 0 & \tilde{R}^+ \end{pmatrix} R_1^+ G R_1 \begin{pmatrix} 1 & 0 \\ 0 & \tilde{R} \end{pmatrix}$$

je hledaná diagonální G'.

Důkaz pro reálný případ se provede obdobně. □

PŘÍKLAD 21. Uvažujme třeba $g : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ s předpisem

$$g(\mathbf{x}, \mathbf{y}) := \mathbf{x}^T \begin{pmatrix} 1 & 1 \\ 1 & 5 \end{pmatrix} \mathbf{y} \equiv x_1 y_1 + x_1 y_2 + x_2 y_1 + 5 x_2 y_2$$

Použijeme-li jako R_1^T matici vhodné řádkové úpravy, získáme

$$R_1^T G R_1 = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 5 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix} =: [g]_B,$$

kde polární bázi $B = ((1,0),(-1,1))$ čteme jako sloupce matice přechodu $R_1 = [\mathrm{Id}]_B^K$. Násobení maticí R zprava se chová jako sloupcová úprava odpovídající řádkové úpravě R_1^T. Vzhledem k asociativitě násobení matic je jedno, která z nich se provede dřív. Úpravám typu $G \to R^T G R$ se říká *symetrické úpravy*, jejich komplexní verzi $G \to R^+ G R$ *hermitovské úpravy*. Protože 1 a 4 jsou kladná čísla, je g pozitivně definitní a B je polární báze. Další symetrickou úpravou, sestávající z vydělení druhého řádku a druhého sloupce číslem 2, převedeme matici $[g]_B$ na jednotkovou. Celý výpočet se obvykle zaznamenává pomocí rozšířené matice

$$\left(\begin{array}{cc|cc} 1 & 1 & 1 & 0 \\ 1 & 5 & 0 & 1 \end{array}\right) \sim \left(\begin{array}{cc|cc} 1 & 0 & 1 & 0 \\ 0 & 4 & -1 & 1 \end{array}\right) \sim \left(\begin{array}{cc|cc} 1 & 0 & 1 & 0 \\ 0 & 1 & -\frac{1}{2} & \frac{1}{2} \end{array}\right),$$

kde v první polovině matice provádíme symetrické úpravy a v pravé jen odpovídající úpravy řádkové. Pak je na konci výpočtu matice vpravo rovna matici R^T transponované k matici přechodu do polární báze. V našem případě tedy pro $\tilde{B} = ((1,0), \frac{1}{2}(-1,1))$ platí $[g]_{\tilde{B}} = E$.

Definice 42. Pokud g je skalární součin na vektorovém prostoru V konečné dimenze a B je jeho polární báze, říkáme, že B je *ortogonální báze* unitárního prostoru (V, g). Je-li dokonce $[g]_B = E$, je B báze *ortonormální*.

Skalární součin budeme od nynějška značit místo $g(u,v)$ pouze $\langle u, v\rangle$.

Definice 43. Nechť $(V, \langle , \rangle)$ je unitární prostor. Dva vektory $u, v \in V$ jsou *kolmé*, pakliže $\langle u, v\rangle = 0$, značíme $u \perp v$. Je-li $M \subset V$, pak její *ortogonální doplněk* $M^\perp$ je množina všech vektorů V kolmých na všechny prvky M. *Norma vektoru* $u \in V$ je číslo $\|u\| = \sqrt{\langle u, u\rangle}$, *vzdálenost vektorů* $u, v \in V$ definujeme pomocí ní jako $\rho(u,v) := \|u - v\|$. Jsou-li u, v nenulové, pak jejich *úhel* definujeme jako $\arccos \frac{|\langle u,v\rangle|}{\|u\|\|v\|}$ v komplexním a jako $\arccos \frac{\langle u,v\rangle}{\|u\|\|v\|}$ v reálném případě.

Věta 23 (Pythagorova). *Pokud $u \perp v$, pak $\|u + v\|^2 = \|u\|^2 + \|v\|^2$.*

Důkaz. Stačí dosadit do polarizační identity a použít definici kolmosti. □

Ortogonální projekce na nenulový vektor $v \in V$ je zobrazení $P_v : V \to V$ definované předpisem

$$P_v(u) = \frac{\langle v, u\rangle}{\|v\|^2} v$$

Ortogonální projekce na ortogonální doplněk $v^\perp$ je pak definována předpisem $P_{v^\perp} = \mathrm{Id} - P_v$, tedy

$$P_{v^\perp}(u) = u - \frac{\langle v, u\rangle}{\|v\|^2} v$$

Tvrzení 44. *Nechť V je unitární prostor nad $\mathbb{F}$, $u, v \in V$, $v \neq o$. Pak*

(1) *$P_v, P_{v^\perp}$ jsou lineární zobrazení*
(2) *$P_v(v) = v$, $\operatorname{Ker} P_v = v^\perp$, $\operatorname{Im} P_v = \langle v\rangle$,*
(3) *$\operatorname{Ker} P_{v^\perp} = \langle v\rangle$, $\operatorname{Im} P_{v^\perp} = v^\perp$*
(4) *$P_v \circ P_v = P_v$, $P_{v^\perp} \circ P_{v^\perp} = P_{v^\perp}$*

Důkaz. První bod plyne z vlastnosti (1) skalárního součinu. Protože $\frac{\langle v,u\rangle}{\|v\|^2}$ je nulový vektor právě když $u \perp v$, vidíme, že $\operatorname{Ker} P_v = v^\perp$. Dosazením $u = v$ do podílu $\frac{\langle v,u\rangle}{\|v\|^2}$ plyne z definice normy $P_v(v) = v$. Vektor $P_v(u)$ je vždy násobkem v a volbou $u = \lambda v$ dokážeme takto získat libovolný násobek v, tedy platí i třetí tvrzení druhého bodu. Obdobně se dokáže třetí bod. Z linearity P_v máme $\forall u \in V$

$$(P_v \circ P_v)(u) = P_v\left(\frac{\langle v, u\rangle}{\|v\|^2} v\right) = \frac{\langle v, u\rangle}{\|v\|^2}\frac{\langle v, v\rangle}{\|v\|^2} v = P_v(u),$$

a podobně plyne i druhá část čtvrtého bodu. $\square$

Čtvrtý bod tvrzení odpovídá obecné matematické definici projekce jako lineárního zobrazení, pro které $P^2 \equiv P \circ P = P$. Je dobré si také uvědomit, že první bod by v komplexním případě neplatil pro zobrazení $u \mapsto \frac{\langle u,v\rangle}{\|v\|^2} v$, které je antilineární. Při definici projekce P_v tedy na pořadí vektorů v čitateli záleží.

Věta 24 (Schwarzova nerovnost). *Nechť $(V, \langle , \rangle)$ je unitární prostor, $u, v \in V$. Pak*

$$|\langle u, v\rangle| \leq \|u\|\|v\|,$$

přičemž rovnost platí, právě když je jeden z vektorů násobkem druhého.

Důkaz. Pro $v = 0$ tvrzení platí, předpokládejme tedy $v \neq o$. Z definice ortogonální projekce a jejích vlastností platí, že $u = P_v(u) + P_{v^\perp}(u)$, přičemž $P_v(u) \perp P_{v^\perp}(u)$. Z Pythagorovy věty pak

$$\|u\|^2 = \|P_v(u)\|^2 + \|P_{v^\perp}(u)\|^2 = \frac{|\langle u, v\rangle|^2}{\|v\|^4}\|v\|^2 + \|P_{v^\perp}(u)\|^2 \geq \frac{|\langle u, v\rangle|^2}{\|v\|^2}$$

V posledním kroku platí rovnost, právě když $P_{v^\perp}(u) = o$, tj. právě když u je násobkem v. Násobením číslem $\|v\|^2$ dostáváme Schwarzovu nerovnost. $\square$

Věta 25 (Trojúhelníková nerovnost). *Nechť $(V, \langle , \rangle)$ je unitární prostor, $u, v \in V$. Pak*

$$\|u + v\| \leq \|u\| + \|v\|$$

Důkaz.

$$\begin{aligned}\|u + v\|^2 &= \|u\|^2 + \langle u, v\rangle + \langle v, u\rangle + \|v\|^2 = \|u\|^2 + 2\operatorname{Re}\langle u, v\rangle + \|v\|^2 \\ &\leq \|u\|^2 + 2|\langle u, v\rangle| + \|v\|^2 \leq \|u\|^2 + 2\|u\|\|v\| + \|v\|^2 \\ &= (\|u\| + \|v\|)^2\end{aligned}$$

Ve třetím kroku jsme využili nerovnost $Re(a+ib) \leq \sqrt{a^2 + b^2} \equiv |a+ib|$, která platí pro každé komplexní číslo, ve čtvrtém pak Schwarzovu nerovnost. $\square$

Posloupnost $(w_1, \ldots, w_k)$ ve V, pro kterou platí $\langle w_i, w_j\rangle = \delta_{ij}$ (*Kroneckerovo delta*, tj. 1 pro $i = j$ a 0 jindy) se nazývá *ortonormální*.

Tvrzení 45. *Nechť $(V, \langle , \rangle)$ je unitární prostor, $(w_i)_1^k$ ortonormální posloupnost v něm, W její lineární obal. Definujme $P_W : V \to V$ předpisem*

$$P_W(u) := \langle w_1, u\rangle w_1 + \ldots + \langle w_k, u\rangle w_k \equiv \sum_{i=1}^{k} P_{w_i}(u)$$

Pak

(1) $u \in W$ *právě když* $P_W(u) = u$; $\operatorname{Im} P_W = W$.
(2) $P_W \circ P_W = P_W$
(3) $u - P_W(u) \in W^\perp$
(4) $\forall u \in V : \|P_W(u)\| \leq \|u\|$, *přičemž rovnost nastává právě když* $u \in W$
(5) $\forall v \in W : \|u - P_W(u)\| \leq \|u - v\|$, *přičemž rovnost nastává právě když* $v = P_W(u)$.

Podle druhého bodu je tedy P_W projekce a podle prvního promítá vektor ortogonálně na W. Pátý bod říká, že $P_W(u)$ je vektor W, který je nejblíže vektoru $u \in V$, a že je tento vektor určen jednoznačně. Nezáleží tedy na tom, pomocí které ortonormální báze W zobrazení P_W zadefinujeme. Snadno se ale ukáže, že pokud by $(w_i)_1^k$ nebyla ortogonální bází, pak $\sum_{i=1}^k P_{w_i}$ nebude splňovat bod (1) a nebude tedy projekcí.

Důkaz. Pokud $P_W(u) = u$, pak je na pravé straně u a na levé lineární kombinace vektorů z $(w_i)_1^k$, tedy $u \in W$. Naopak pokud $u = \sum_{j=1}^k r_j w_j \in W$, pak

$$\begin{aligned} P_W(u) &= \sum_{i=1}^k \left\langle w_i, \sum_{j=1}^k r_j w_j \right\rangle w_i = \sum_{i,j=1}^k r_j \langle w_i, w_j \rangle w_i \\ &= \sum_{i,j=1}^k r_j \delta_{ij} w_i = \sum_{i=1}^k r_i w_i = u \end{aligned}$$

Odtud plyne i $\operatorname{Im} P_W = W$ a $(P_W)^2 = P_W$. Pro důkaz třetího tvrzení spočtěme

$$\langle w_j, u - P_W(u) \rangle = \langle w_j, u \rangle - \sum_{i=1}^k \langle w_i, w_j \rangle \langle w_j, u \rangle = \langle w_j, u \rangle - \sum_{i=1}^k \delta_{ij} \langle w_j, u \rangle = 0$$

Tedy $u - P_W(u)$ je kolmé na všechny vektory z $(w_i)_1^k$, tedy i na každou jejich lineární kombinaci, patří tedy do $W^\perp$. Díky prvnímu bodu je pak $u - P_W(u) \perp P_W(u)$. Podle Pythagorovy věty potom

$$\|u\|^2 = \|P_W(u)\|^2 + \|u - P_W(u)\|^2 \geq \|P_W(u)\|^2,$$

přičemž rovnost nastává právě když $P_W(u) = u$, což je podle prvního bodu právě když $u \in W$. Tím je dokázán čtvrtý bod.

Podobně v posledním bodě víme, že vektor $P_W(u) - v \in W$ je kolmý na $u - P_W(u)$. Z Pythagorovy věty aplikované na součet těchto vektorů dostáváme, že

$$\|u - v\|^2 = \|u - P_W(u)\|^2 + \|P_W(u) - v\|^2 \geq \|u - P_W(u)\|^2,$$

s rovností právě v případě $v = P_W(u)$. □

Tvrzení 46. *Nechť $(V, \langle , \rangle)$ je unitární prostor, $W \leq V$ je podprostor konečné dimenze. Pak*

(1) $W \oplus W^\perp = V$
(2) $\operatorname{Ker} P_W = W^\perp$
(3) $(W^\perp)^\perp = W$
(4) *Pokud V je konečné dimenze, pak* $\dim W = \dim V - \dim W^\perp$

Důkaz. Je-li $u \in V$, pak se dá zapsat jako součet $P_W(u) + (u - P_W(u))$, ze třetího bodu tvrzení 45 a definice součtu podprostorů tedy $V = W + W^\perp$. Je-li $w \in W \cap W^\perp$, pak $\langle w, w \rangle = 0$, z vlastnosti (3) skalárního součinu tedy $w = o$, součet je tedy direktní.

Díky prvnímu bodu má každý vektor $v \in V$ jednoznačný rozklad na součet vektoru z W a vektoru z $W^\perp$, a to $v = P_W(v) + (v - P_W(v))$. První složka $P_W(v)$ rozkladu je nulová, právě když je v roven druhé složce $v - P_W(v) \in W^\perp$. Tím je dokázáno druhé tvrzení.

Pokud $w \in W$, pak je jistě kolmý na všechny vektory $W^\perp$, tedy $w \in (W^\perp)^\perp$. Nechť naopak pak je $w \in (W^\perp)^\perp$. Z prvního bodu tvrzení 45 víme, že $P_W(w) \in W$, a ze třetího bodu, že $w - P_W(w) \in W^\perp$. Pak ale

$$\|w - P_W(w)\|^2 = \langle w, w - P_W(w) \rangle - \langle P_W(w), w - P_W(w) \rangle = 0 + 0,$$

neboli $w = P_W(w) \in W$. Poslední bod tvrzení je aplikace věty o dimenzi direktního součtu. $\square$

KAPITOLA 14

Ortogonalizace

V minulé kapitole jsme definovali ortogonální projekci P_W na podprostor W unitárního prostoru V pomocí ortonormální báze ve W. Zatím ale nevíme, že taková báze v každém unitárním prostoru existuje. Následující věta říká, že ano, a dává návod na její konstrukci ortogonalizací obecně neortogonální báze. Libovolně dlouhý začátek zkonstruované báze navíc bude generovat stejný podprostor jako odpovídající začátek báze původní.

Věta 26 (Gramova-Schmidtova ortogonalizace). *Nechť $(V, \langle , \rangle)$ je unitární prostor a $B = (u_1, \ldots, u_n)$ jeho báze. Pak existuje ortonormální báze $C = (w_1, \ldots, w_n)$ prostoru V taková, že $\forall k \in \{1, \ldots, n\}$ platí $\langle u_1, \ldots, u_k\rangle = \langle w_1, \ldots, w_k\rangle$.*

Důkaz. Budeme postupovat indukcí podle dimenze n. Pro $n = 1$ definujeme $v_1 := u_1$ a $w_1 := \frac{v_1}{\|v_1\|}$ a věta platí triviálně. Nechť $n > 1$ a předpokládejme platnost věty pro všechny prostory dimenze menší než n, tedy i pro prostor $W_{n-1} := \langle u_1, \ldots, u_{n-1}\rangle$. Podle indukčního předpokladu v něm máme ortonormální bázi $(w_1, \ldots, w_{n-1})$, kde pro $\forall k \in \{1, \ldots, n-1\}$ platí $\langle u_1, \ldots, u_k\rangle = \langle w_1, \ldots, w_k\rangle$. Nyní ortogonalizujeme vektor u_n:

$$v_n := u_n - P_{W_{n-1}}(u_n) = u_n - \sum_{i=1}^{n-1} \langle w_i, u_n\rangle w_i$$

Že je v_n skutečně kolmý na podprostor W_{n-1}, se můžeme přesvědčit jeho skalárním vynásobením se všemi předchozími vektory. Pro všechna $j \in \{1, \ldots, n-1\}$ tak dostáváme

$$\langle w_j, v_n\rangle = \langle w_j, u_n\rangle - \sum_{i=1}^{n-1} \langle w_i, u_n\rangle\langle w_j, w_i\rangle = 0,$$

kde poslední rovnost plyne z ortonormality posloupnosti $(w_i)_1^{n-1}$, tedy $\langle w_j, w_i\rangle$ je rovno 1 pro $i = j$ a 0 jinak. Kdyby $v_n = o$, pak by $u_n \in \langle (w_i)_1^{n-1}\rangle = W_{n-1}$. Z indukčního předpokladu $W_{n-1} = \langle (u_i)_1^{n-1}\rangle$, tedy by $(u_1, \ldots, u_n)$ byla lineárně závislá, což je v rozporu s předpokladem, že $(u_i)_1^n$ je báze. Můžeme proto definovat $w_n := \frac{v_n}{\|v_n\|}$. To je vektor s normou 1, kolmý na všechny vektory $w_1, \ldots, w_{n-1}$. Z indukčního předpokladu víme, že $\langle w_1, \ldots, w_k\rangle = \langle u_1, \ldots, u_k\rangle$ pro všechna $k < n$, a pro $k = n$ to plyne z definice vektorů v_n a w_n. □

Součástí důkazu je i algoritmus výpočtu. Při ručním počítání je výhodnější normalizaci provést až na závěr se všemi vektory a k-tý krok Gramovy-Schmidtovy ortogonalizace zapsat jako

$$v_k := u_k - \sum_{i=1}^{k-1} \frac{\langle v_i, u_k\rangle}{\|v_i\|^2} v_i$$

Příklad 22. V $\mathbb{R}^4$ se standardním skalárním součinem ortogonalizujme posloupnost

$$(u_1, u_2, u_3) := ((1, 2, 2, -1), (1, 1, -5, 3), (3, 2, 8, -7))$$

Začneme vektorem $v_1 := u_1$ a dopočteme vektor

$$v_2 := u_2 - \frac{\langle u_2, v_1\rangle}{\|v_1\|^2} v_1 = (1,1,-5,3) - \frac{-10}{10}(1,2,2,-1) = (2,3,-3,2)$$

Vidíme, že skutečně $v_2 \perp v_1$. Spočteme $\|v_2\|^2 = 26$, $\langle u_3, v_1\rangle = 30$ a $\langle u_3, v_2\rangle = -26$, takže

$$v_3 := (3,2,8,-7) - \frac{30}{10}(1,2,2,-1) - \frac{-26}{26}(2,3,-3,2) = (2,-1,-1,-2),$$

což je opět vektor kolmý na v_2 i v_1. Ortonormální báze pak bude

$$C = \left(\frac{1}{\sqrt{10}}(1,2,2,-1), \frac{1}{\sqrt{26}}(2,3,-3,2), \frac{1}{\sqrt{10}}(2,-1,-1,-2)\right)$$

TVRZENÍ 47 (Fourierovy koeficienty). *Nechť $C = (w_i)_1^n$ je ortonormální báze unitárního prostoru $(V, \langle,\rangle)$. Pak lze každý vektor $u \in V$ reprezentovat vůči C jako*

$$[u]^C = \begin{pmatrix} \langle w_1, u\rangle \\ \vdots \\ \langle w_n, u\rangle \end{pmatrix} \in \mathbb{C}^n$$

DŮKAZ. Stačí si uvědomit, že ortogonální projekce na podprostor V ve V je

$$u = P_V(u) = \sum_{i=1}^n \langle w_i, u\rangle w_i,$$

□

Pojmenování Fourierovy koeficienty pro takto vyjádřené souřadnice vektoru vůči ortonormální bázi vychází ze souvislosti s Fourierovými řadami. Máme-li na vektorovém prostoru všech reálných funkcí na intervalu $[-\pi, \pi]$, které jsou lineární kombinací funkcí $1, \cos x, \cos 2x, \dots, \cos nx, \sin x, \sin 2x, \dots, \sin nx,$ definován skalární součin předpisem

$$\langle f, g\rangle := \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)g(x)dx,$$

je posloupnost generátorů $\left(\frac{1}{\sqrt{2}}, \cos x, \dots, \cos nx, \sin x, \dots, \sin nx\right)$ ortonormální báze a koeficienty a_i, b_i v rozvoji

$$f(x) = \frac{a_0}{2} + \sum_{k=1}^n a_k \cos kx + \sum_{k=1}^n b_k \sin kx$$

mají vyjádření

$$a_k = \langle \cos kx, f(x)\rangle = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos kx dx$$
$$b_k = \langle \sin kx, f(x)\rangle = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin kx dx$$

V teorii Fourierových řad se pak tato myšlenka zobecňuje na situaci, kdy má prostor funkcí nekonečnou dimenzi a $f(x)$ může být libovolná funkce, pro kterou její vyjádření nekonečnou řadou s koeficienty $(a_i)_0^\infty$, $(b_i)_1^\infty$ konverguje. Platí pak vztah

$$\sum_{k=0}^\infty |a_k|^2 + \sum_{k=1}^\infty |b_k|^2 = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)^2 dx,$$

který se nazývá *Parsevalova rovnost.* V naší konečně dimenzionální verzi má následující tvar:

TVRZENÍ 48 (Parsevalova rovnost). *Nechť $(V, \langle , \rangle)$ je unitární prostor konečné dimenze, $C = (w_1, \ldots, w_n)$ jeho ortonormální báze a $u \in V$. Pak*

$$\sum_{i=1}^{n} |\langle w_i, u\rangle|^2 = \|u\|^2$$

DŮKAZ. Protože $u = \sum_{i=1}^{n} \langle w_i, u\rangle w_i$, dostáváme z definice normy, tvrzení 47 a vlastností skalárního součinu

$$\|u\|^2 = \langle u, u\rangle = \left\langle \sum_{i=1}^{n} \langle w_i, u\rangle w_i, \sum_{j=1}^{n} \langle w_j, u\rangle w_j \right\rangle =$$
$$= \sum_{i=1}^{n} \sum_{j=1}^{n} \overline{\langle w_i, u\rangle} \langle w_j, u\rangle \langle w_i, w_j\rangle = \sum_{i=1}^{n} |\langle w_i, u\rangle|^2$$

□

Zapíšeme-li $u = \sum_{i=1}^{n} x_i w_i$, dá se Parsevalova rovnost přepsat jako

$$\|u\| = \sqrt{|x_1|^2 + \ldots + |x_n|^2},$$

je to tedy vyjádření normy na V pomocí standardní normy v prostoru souřadnic $\mathbb{C}^n$. Obdobný vztah je mezi skalárním součinem na V a standardním skalárním součinem na $\mathbb{C}^n$:

TVRZENÍ 49. *Nechť $(V, \langle , \rangle)$ je unitární prostor konečné dimenze, $C = (w_1, \ldots, w_n)$ jeho ortonormální báze, $u, v \in V$, $\mathbf{x} = [u]^C$, $\mathbf{y} = [v]^C$. Pak*

$$\langle u, v\rangle = \sum_{i=1}^{n} \bar{x}_i y_i$$

DŮKAZ.

$$\langle u, v\rangle = \left\langle \sum_{i=1}^{n} \langle w_i, u\rangle w_i, \sum_{j=1}^{n} \langle w_j, v\rangle w_j \right\rangle = \sum_{i=1}^{n} \sum_{j=1}^{n} \overline{\langle w_i, u\rangle} \langle w_j, v\rangle \langle w_i, w_j\rangle =$$
$$= \sum_{i=1}^{n} \overline{\langle w_i, u\rangle} \langle w_i, v\rangle = \sum_{i=1}^{n} \bar{x}_i y_i$$

□

Tvrzení se dá rovněž dokázat pomocí polarizační identity a Parsevalovy rovnosti (Tvrzení 43 a 48). Skalární součin na V i na $\mathbb{C}^n$ se dají vyjádřit pomocí norem, a ty se rovnají.

PŘÍKLAD 23. Chceme-li vektor $u := u_3 = (3, 2, 8, -7)$ vyjádřit vůči C, počítáme místo nehomogenní soustavy rovnic jen tři Fourierovy koeficienty:

$$\langle w_1, u\rangle = \frac{1}{\sqrt{10}} \langle (1, 2, 2, -1), (3, 2, 8, -7)\rangle = \frac{1}{\sqrt{10}} 30 = 3\sqrt{10},$$
$$\langle w_2, u\rangle = \frac{1}{\sqrt{26}} \langle (2, 3, -3, 2), (3, 2, 8, -7)\rangle = \frac{1}{\sqrt{26}} (-26) = -\sqrt{26},$$
$$\langle w_3, u\rangle = \frac{1}{\sqrt{10}} \langle (2, -1, -1, -2), (3, 2, 8, -7)\rangle = \frac{1}{\sqrt{10}} 10 = \sqrt{10}$$

Otestujme, že platí i Parsevalova rovnost:

$$\|u\|^2 = 9 + 4 + 64 + 49 = 126$$

$$\sum_{i=1}^{3} |\langle w_i, u\rangle|^2 = 90 + 26 + 10 = 126$$

S Gramovou-Schmidtovou ortogonalizací úzce souvisí maticový rozklad, který patří k nejdůležitějším nástrojům numerické lineární algebry.

VĚTA 27 (QR rozklad matice). *Nechť $A \in \mathbb{C}^{m\times n}$ je matice s lineárně nezávislými sloupci. Pak existuje právě jedna dvojice matic $Q \in \mathbb{C}^{m\times n}$, $R \in \mathbb{C}^{n\times n}$ taková, že*

(1) $A = QR$
(2) *sloupce Q tvoří ortonormální množinu vzhledem ke standardnímu skalárnímu součinu na $\mathbb{C}^m$*
(3) *R je horní trojúhelníková matice s kladnými čísly na diagonále.*

DŮKAZ. Označme $A = (\mathbf{a}_1|\dots|\mathbf{a}_n)$ a budiž $(\mathbf{q}_1,\dots,\mathbf{q}_n)$ ortonormální posloupnost vzniklá Gramovou-Schmidtovou ortogonalizací posloupnosti $(\mathbf{a}_1,\dots,\mathbf{a}_n)$. Položme $Q := (\mathbf{q}_1|\dots|\mathbf{q}_n)$. Vyjádření $\mathbf{a}_k$ pomocí Fourierových koeficientů

$$\mathbf{a}_k = \sum_{i=1}^{k} \langle \mathbf{q}_i, \mathbf{a}_k\rangle \mathbf{q}_i$$

zapišme maticově jako

$$(\mathbf{a}_1|\dots|\mathbf{a}_n) = (\mathbf{q}_1|\dots|\mathbf{q}_n)\begin{pmatrix} \langle\mathbf{q}_1,\mathbf{a}_1\rangle & \langle\mathbf{q}_1,\mathbf{a}_2\rangle & \dots & \langle\mathbf{q}_1,\mathbf{a}_n\rangle \\ 0 & \langle\mathbf{q}_2,\mathbf{a}_2\rangle & \dots & \langle\mathbf{q}_2,\mathbf{a}_n\rangle \\ \vdots & & \ddots & \vdots \\ 0 & \dots & 0 & \langle\mathbf{q}_n,\mathbf{a}_n\rangle \end{pmatrix}$$

Označme $\mathbf{q}'_k := \mathbf{a}_k - \sum_{i=1}^{k-1}\langle\mathbf{q}_i,\mathbf{a}_k\rangle\mathbf{q}_i$ vektor ortogonální báze, která vzniká v Gramově-Schmidtově algoritmu před normalizací. Pak

$$\langle\mathbf{q}_k,\mathbf{a}_k\rangle = \left\langle \mathbf{q}_k, \mathbf{q}'_k + \sum_{i=1}^{k-1}\langle\mathbf{q}_i,\mathbf{a}_k\rangle\mathbf{q}_i, \right\rangle = \langle\mathbf{q}_k,\mathbf{q}'_k\rangle = \left\langle \frac{\mathbf{q}'_k}{\|\mathbf{q}'_k\|}, \mathbf{q}'_k \right\rangle = \|\mathbf{q}'_k\| > 0$$

Druhá rovnost platí díky linearitě skalárního součinu ve druhém argumentu a ortogonalitě posloupnosti $(\mathbf{q}_1,\dots,\mathbf{q}_k)$. Tím je dokázána existence Q a R s danými vlastnostmi.

Jednoznačnost dokážeme sporem. Předpokládejme že existují dva rozklady, tedy $A = Q_1R_1 = Q_2R_2$. Z ortonormality sloupců Q_2 plyne $Q_2^+Q_2 = E$, R_1 je regulární, máme tedy rovnost $Q_2^+Q_1 = R_2R_1^{-1}$. Inverzní matice k horní trojúhelníkové matici je horní trojúhelníková, součin dvou horních trojúhelníkových matic je rovněž horní trojúhelníková, zachovává se i kladnost prvků na diagonále, z toho plyne, že první sloupec matice $R_2R_1^{-1}$ je kladný násobek vektoru $\mathbf{e}_1$. Tedy druhý až n-tý sloupec Q_2 jsou všechny kolmé na první sloupec Q_1. Protože první sloupec Q_1 leží v lineárním obalu sloupců matice Q_2 a zároveň v ortogonálním doplňku jejího druhého až n-tého sloupce, musí být násobkem prvního sloupce Q_2. Oba první sloupce mají normu 1 a jejich skalární součin je kladný, musí se tedy rovnat. Analogickým způsobem ukážeme rovnost ostatních odpovídajících si sloupců matic Q_1 a Q_2. Pak je ovšem $Q_2^+Q_1 = E$ a tedy $R_1 = R_2$. □

Důkaz jednoznačnosti se trochu zjednoduší, když je A čtvercová, a tedy je čtvercová i Q. Vztah $Q^+Q = E$, který je maticovým vyjádřením ortonormality sloupců, znamená, že Q^+ je inverzní maticí ke Q. Ta je ale inverzní i zprava, tedy $QQ^+ = E$, neboli ortonormální bázi $\mathbb{C}^n$ tvoří i řádky matice Q.

DEFINICE 44. Čtvercová matice $U \in \mathbb{C}^{n\times n}$ s vlastností $U^+U = UU^+ = E$ se nazývá *unitární*. Reálná unitární matice se nazývá *ortogonální*.

TVRZENÍ 50. *Jsou-li $U, V \in \mathbb{C}^{n\times n}$ unitární matice, pak i $U^+ \equiv U^{-1}$ a UV jsou unitární matice.*

DŮKAZ. Přímým dosazením do definice. $\square$

TVRZENÍ 51. *Nechť $(V, \langle , \rangle)$ je unitární prostor, $B = (v_1, \dots, v_n)$, $C = (w_1, \dots, w_n)$ jeho ortonormální báze. Pak je matice přechodu $U := [\mathrm{Id}]^B_C$ unitární.*

DŮKAZ. Z definice matice přechodu platí $w_k = \sum_{i=1}^n u_{ik} v_i$ a z vlastností skalárního součinu a ortonormality báze B pak plyne, že

$$\delta_{kj} = \langle w_k, w_j \rangle = \sum_{i=1}^n \sum_{l=1}^n \overline{u}_{ik} u_{lj} \langle v_i, v_l \rangle = \sum_{i=1}^n \sum_{l=1}^n \overline{u}_{ik} u_{lj} \delta_{il} = \sum_{i=1}^n u^+_{ki} u_{ij},$$

tedy $U^+ U = E$. Protože U je čtvercová, tak i $UU^+ = E$. $\square$

Matice přechodu v opačném směru je $U^{-1} \equiv [\mathrm{Id}]^C_B = U^+$. Transformační formule pro matici $A := [f]^B_B$ endomorfismu $f \in \mathrm{End}(V)$ vzhledem k ortonormální bázi B pak je

$$A' := [f]^C_C = [\mathrm{Id}]^C_B [f]^B_B [\mathrm{Id}]^B_C = U^+ A U$$

nebo v reálném případě $A' = U^T A U$. Zjistili jsme pozoruhodnou vlastnost, totiž že v ortonormálních bázích transformační formule pro endomorfismy a pro bilineární resp. seskvilineární formy splývají. V dalších kapitolách se dozvíme, jak toho využít například při klasifikaci kuželoseček.

POZNÁMKA 4. K čemu se může QR rozklad hodit? Uvažujme soustavu rovnic $A\mathbf{x} = \mathbf{b}$, kde A je regulární matice. Máme-li k dispozici QR rozklad A, můžeme soustavu upravit na tvar $R\mathbf{x} = Q^+\mathbf{b}$. Protože R je horní trojúhelníková matice, lze tuto soustavu snadno vyřešit postupným dosazováním odspoda. To je numericky stabilnější než Gaussova eliminace, lze totiž ukázat, že násobení unitární maticí Q^+ nemění velikost chyby ve vstupních datech (matici A a vektoru $\mathbf{b}$). Gramova-Schmidtova ortogonalizace, která vedla k nalezení QR rozkladu, sice také není numericky stabilní, ale existují jiné metody jeho nalezení, pomocí Householderových nebo Givensových transformací. Ty jsou realizovány pomocí unitárních matic a tedy opět nezvětšují chyby ve vstupních datech.

KAPITOLA 15

Ortogonální diagonalizace

V některých oblastech matematiky a fyziky (funkcionální analýza, kvantová teorie) se používá místo pojmů (lineární) zobrazení mezi vektorovými prostory pojem *(lineární) operátor*. Operátor mezi vektorovými prostory V a W budeme značit velkým tučným či zdvojeným písmenem

$$\mathbb{A} : V \to W$$

a $\mathbb{A}v \in W$ označuje obraz vektoru $v \in V$ pomocí operátoru $\mathbb{A}$. Budeme předpokládat, že V a W jsou unitární prostory nad $\mathbb{C}$ a pojmem operátor rozumět vždy lineární operátor. Pokud $V = W$, mluvíme o *operátoru na* V. Operátor $\mathbb{A}^* : W \to V$, který splňuje $\forall v \in V$, $w \in W$

$$\langle w, \mathbb{A}v\rangle_W = \langle \mathbb{A}^* w, v\rangle_V,$$

nazýváme operátor *adjungovaný* (či *sdružený*) k $\mathbb{A}$. Indexy W, V rozlišují, ve kterém z prostorů skalární součin počítáme. To je obvykle zřejmé z kontextu, proto dále značíme oba skalární součiny jen $\langle , \rangle$. Pokud oba prostory popisujeme vůči ortonormálním bázím, existuje jednoduchý vztah mezi operátorem a jeho operátorem k němu adjungovaným:

TVRZENÍ 52. *Nechť $B = (v_i)_1^n$, $C = (w_i)_1^m$ jsou ortonormální báze V, resp. W, $A = [\mathbb{A}]_B^C \in \mathbb{C}^{m\times n}$. Pak $\mathbb{A}^*$ existuje a $[\mathbb{A}^*]_C^B = A^+ \in \mathbb{C}^{n\times m}$.*

DŮKAZ. Z definice matice lineárního zobrazení víme, že $\mathbb{A}v_j = \sum_{i=1}^m a_{ij} w_i$. Definujme operátor $\mathbb{A}^*$ hodnotami na bázi C jako $\mathbb{A}^* w_k := \sum_{i=1}^n \overline{a}_{ki} v_i$. Ze seskvilinearity skalárního součinu a ortonormality bází plyne

$$\langle w_k, \mathbb{A}v_j\rangle = \left\langle w_k, \sum_{i=1}^m a_{ij} w_i \right\rangle = \sum_{i=1}^m a_{ij} \langle w_k, w_i\rangle = a_{kj}$$

$$\langle \mathbb{A}^* w_k, v_j\rangle = \left\langle \sum_{i=1}^n \overline{a}_{ki} v_i, v_j \right\rangle = \sum_{i=1}^n a_{ki} \langle v_i, v_j\rangle = a_{kj}.$$

Pro libovolné vektory $w = \sum_{k=1}^m r_k w_k \in W$ a $v = \sum_{j=1}^n s_j v_j \in V$ pak platí

$$\langle w, \mathbb{A}v\rangle = \sum_{k=1}^m \sum_{j=1}^n \overline{r}_k s_j \langle w_k, \mathbb{A}v_j\rangle = \sum_{k=1}^m \sum_{j=1}^n \overline{r}_k s_j \langle \mathbb{A}^* w_k, v_j\rangle = \langle \mathbb{A}^* w, v\rangle$$

Z toho plyne, že takto definovaný operátor $\mathbb{A}^*$ splňuje definici operátoru adjungovaného k $\mathbb{A}$. Z jeho definice je jk-tý element jeho matice $[\mathbb{A}^*]_C^B$ roven $\overline{a}_{kj} \equiv a_{jk}^+$. Pokud by existoval jiný operátor $\mathbb{B} : W \to V$ splňující definici adjungovaného operátoru k $\mathbb{A}$, pak by speciálně musel být pro všechna $k \in \{1, \ldots, n\}$ a $j \in \{1, \ldots, m\}$ splněn vztah $\langle \mathbb{B}w_k, v_j\rangle = a_{kj}$. Na levé straně vztahu jsou komplexně sdružené Fourierovy koeficienty vektorů $\mathbb{B}w_k$ vzhledem k bázi B. Tyto koeficienty vektor $\mathbb{B}w_k$ jednoznačně určují a posloupnost $(\mathbb{B}w_k)_1^m$ zase jednoznačně určuje operátor $\mathbb{B}$. $\square$

Operace $*$ sdružení operátoru je tedy v ortonormální bázích vyjádřena operací $+$ hermitovského sdružení matic.

Tvrzení 53. *Nechť V, W, U jsou unitární prostory, $\mathbb{A}, \mathbb{B} : V \to W$, $\mathbb{G} : U \to V$ operátory, k nimž existují operátory adjungované, $\alpha, \beta \in \mathbb{C}$. Pak platí*

(1) $(\alpha\mathbb{A} + \beta\mathbb{B})^* = \bar{\alpha}\mathbb{A}^* + \bar{\beta}\mathbb{B}^*$
(2) $(\mathbb{A}\mathbb{G})^* = \mathbb{G}^*\mathbb{A}^*$
(3) $(\mathbb{A}^*)^* = \mathbb{A}$
(4) $\operatorname{Ker}\mathbb{A} = (\operatorname{Im}\mathbb{A}^*)^\perp, \operatorname{Ker}\mathbb{A}^* = (\operatorname{Im}\mathbb{A})^\perp$
(5) $\operatorname{Ker}\mathbb{A} = \operatorname{Ker}\mathbb{A}^*\mathbb{A}$
(6) *Pokud je navíc* $\dim V = n \in \mathbb{N}$, *pak* $\operatorname{rank}(\mathbb{A}) = \operatorname{rank}(\mathbb{A}^*)$ *a* $\operatorname{rank}(\mathbb{A}) = \operatorname{rank}(\mathbb{A}^*\mathbb{A})$

Důkaz. První tři tvrzení plynou okamžitě z definice. Pro vektor $v \in V$ platí $\mathbb{A}v = 0$, právě když $\forall w \in W$ je $\langle \mathbb{A}v, w\rangle = 0$, tedy právě když $v \perp \mathbb{A}^*w$. Analogicky se dokáže i druhá půlka čtvrtého tvrzení. Pro důkaz pátého tvrzení nejprve předpokládejme, že $v \in \operatorname{Ker}\mathbb{A}^*\mathbb{A}$. Pak

$$0 = \langle \mathbb{A}^*\mathbb{A}v, v\rangle = \langle \mathbb{A}v, \mathbb{A}v\rangle = \|\mathbb{A}v\|^2,$$

což nastává právě tehdy, když $v \in \operatorname{Ker}\mathbb{A}$. Naopak pokud $v \in \operatorname{Ker}\mathbb{A}$, musí $\mathbb{A}v = 0$ a tedy $\mathbb{A}^*(\mathbb{A}v) = 0$. Šesté tvrzení odtud plyne aplikací věty o hodnosti a nulitě na čtvrté:

$$\operatorname{rank}(\mathbb{A}) = n - \dim\operatorname{Ker}(\mathbb{A}) = n - \dim(\operatorname{Im}\mathbb{A}^*)^\perp = \dim\operatorname{Im}\mathbb{A}^* \equiv \operatorname{rank}(\mathbb{A}^*)$$

a obdobně i na páté tvrzení. □

Definice 45. Operátor $\mathbb{A} : V \to V$ se nazývá

- *samosdružený*, pokud $\mathbb{A} = \mathbb{A}^*$.
- *unitární*, pokud $\mathbb{A}\mathbb{A}^* = \mathbb{A}^*\mathbb{A} = \mathrm{Id}$.
- *normální*, pokud $\mathbb{A}\mathbb{A}^* = \mathbb{A}^*\mathbb{A}$.

Z definice plyne, že samosdružený nebo unitární operátor je nutně také normální. Z tvrzení 52 má samosdružený operátor vzhledem k ortonormální bázi hermitovskou matici, matice unitárního je unitární. Matice normálního operátoru komutuje se svou hermitovsky sdruženou, tj. $A^+A = AA^+$. Taková matice se nazývá *normální.*

Hlavní výsledek této kapitoly je Věta 29 o ortogonální diagonalizaci, která říká, že jsou to právě normální operátory a matice, které lze převést na diagonální tvar pomocí unitární matice přechodu. Nejprve ukážeme, že každý operátor lze takto převést na horní trojúhelníkový tvar.

Lemma 6. *Nechť V je unitární prostor dimenze n a $(w_1, \dots, w_k)$ ortonormální posloupnost ve V. Pak lze tuto posloupnost doplnit na ortonormální bázi $(w_1, \dots, w_n)$ prostoru V.*

Důkaz. Stačí doplnit posloupnost na bázi V a použít na doplněnou posloupnost Gramovu-Schmidtovu ortogonalizaci. □

Věta 28 (Schurova reprezentace operátoru). *Nechť V je unitární prostor konečné dimenze, $\mathbb{A}$ operátor na něm. Pak existuje ortonormální báze B ve V taková, že matice $[\mathbb{A}]_B^B$ je horní trojúhelníková.*

Důsledek 11 (Schurův rozklad matice). *Nechť $A \in \mathbb{C}^{n\times n}$. Pak existují $U \in \mathbb{C}^{n\times n}$ unitární matice a $R \in \mathbb{C}^{n\times n}$ horní trojúhelníková matice takové, že $A = URU^+$.*

Důkaz. Důsledek plyne z věty díky transformační formuli a tvrzení 51:

$$A = [\mathbb{A}]_K^K = [\mathrm{Id}]_B^K[\mathbb{A}]_B^B[\mathrm{Id}]_K^B = URU^+$$

Pro důkaz věty potřebujeme Důsledek 10 Základní věty algebry, podle nějž má operátor $\mathbb{A}$ alespoň jedno vlastní číslo $\lambda \in \mathbb{C}$. Mezi vlastními vektory k tomuto vlastnímu číslu vyberme nějaký normalizovaný v, tj. $\|v\| = 1$. Na základě lemmatu 6 zvolme nějakou ortonormální bázi $C = (v, C')$ ve V a reprezentujme operátor blokovou maticí

$$[\mathbb{A}]_C^C = \begin{pmatrix} \lambda & \mathbf{x}^T \\ \mathbf{o} & A' \end{pmatrix},$$

kde $\mathbf{x} \in \mathbb{C}^{n-1}$ je nějaký vektor a $A' = [\mathbb{A}']_{C'}^{C'} \in \mathbb{C}^{n-1\times n-1}$ definuje operátor na podprostoru $v^\perp$. Z indukčního předpokladu má A' Schurův rozklad $U'R'U'^+$, kde $R' = [\mathbb{A}']_{B'}^{B'}$ je horní trojúhelníková matice a B' je nějaká ortonormální báze $v^\perp$. Báze $B := (v, B')$ je ortonormální a

$$[\mathbb{A}]_B^B = \begin{pmatrix} 1 & \mathbf{o}^T \\ \mathbf{o} & U'^+ \end{pmatrix} \begin{pmatrix} \lambda & \mathbf{x}^T \\ \mathbf{o} & A' \end{pmatrix} \begin{pmatrix} 1 & \mathbf{o}^T \\ \mathbf{o} & U' \end{pmatrix} = \begin{pmatrix} \lambda & \mathbf{x}^T U' \\ \mathbf{o} & R' \end{pmatrix}$$

Poslední matice je horní trojúhelníková, čímž je důkaz dokončen. □

Věta 29 (Ortogonální diagonalizace normálního operátoru). *Nechť V je unitární prostor konečné dimenze. Pro operátor $\mathbb{A} : V \to V$ existuje ortonormální báze prostoru V, vůči níž je jeho matice diagonální, právě když je $\mathbb{A}$ normální.*

Důkaz. Existuje-li ortonormální báze B taková, že $[\mathbb{A}]_B^B =: D = \operatorname{diag}(d_1, \ldots, d_n)$ je diagonální, pak

$$[\mathbb{A}^*\mathbb{A}]_B^B = [\mathbb{A}^*]_B^B[\mathbb{A}]_B^B = D^+D = \operatorname{diag}(|d_1|^2, \ldots, |d_n|^2) = DD^+ = [\mathbb{A}]_B^B[\mathbb{A}^*]_B^B = [\mathbb{A}\mathbb{A}^*]_B^B$$

Tím je dokázána implikace zleva doprava. Předpokládejme tedy naopak, že $\mathbb{A}$ je normální operátor. Dokážeme indukcí, že každá normální horní trojúhelníková matice je diagonální, pro 1×1 matice je to triviální. Nechť B je ortonormální báze V z Věty 28. Pak $A := [\mathbb{A}]_B^B$ je horní trojúhelníková a zároveň normální. Matice A' v blokových zápisech

$$A = \begin{pmatrix} \lambda & \mathbf{x}^T \\ \mathbf{o} & A' \end{pmatrix}, \quad A^+ = \begin{pmatrix} \bar\lambda & \mathbf{o}^T \\ \bar{\mathbf{x}} & A'^+ \end{pmatrix}$$

je tedy také horní trojúhelníková. Podmínka $A^+A = AA^+$ dává blokově

$$\begin{pmatrix} |\lambda|^2 & \bar\lambda\mathbf{x}^T \\ \lambda\bar{\mathbf{x}} & \bar{\mathbf{x}}\mathbf{x}^T + A'^+A' \end{pmatrix} = \begin{pmatrix} |\lambda|^2 + \mathbf{x}^T\bar{\mathbf{x}} & \mathbf{x}^T A'^+ \\ A'\bar{\mathbf{x}} & A'A'^+ \end{pmatrix}$$

Protože $\mathbf{x}^T\bar{\mathbf{x}} = \sum_{i=1}^{n-1} |x_i|^2$, dostáváme z porovnání obou levých horních rohů, že $\mathbf{x} = 0$. Z pravého dolního rohu pak plyne, že A' je i normální matice. Z indukčního předpokladu je A' diagonální a tedy je i A. Schurova reprezentace normálního operátoru je tedy diagonální matice. □

Velmi důležitým důsledkem věty je, že pro každou čtvercovou matici A, která je hermitovská nebo unitární, existuje unitární matice U taková, že U^+AU je diagonální.

Příklad 24. Proveďme ortogonální diagonalizaci matice

$$A = \begin{pmatrix} 0 & 2 & 2 \\ 2 & 0 & 2 \\ 2 & 2 & 0 \end{pmatrix}$$

Vlastní čísla jsou 4 a -2, druhé z nich je dvojnásobné. Příslušné vlastní podprostory jsou $\langle(1, 1, 1)\rangle$ a $\langle(1, -1, 0), (0, 1, -1)\rangle$, vidíme, že jsou skutečně navzájem kolmé. Ve druhém z nich pomocí Gramovy-Schmidtovy metody najdeme ortogonální bázi $((1, -1, 0), (1, 1, -2))$. V předpisu pro diagonalizaci $A = UDU^+$ je matice U přechodu od kanonické báze do báze z vlastních vektorů unitární, její sloupce tedy tvoří normalizované vlastní vektory. V tomto případě, kdy bylo možné vzít všechny vlastní vektory reálné, je $U^+ = U^T$, tedy U je ortogonální matice. Celkově máme A zapsáno jako součin

$$\begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & -\frac{2}{\sqrt{6}} \end{pmatrix} \begin{pmatrix} 4 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & -\frac{2}{\sqrt{6}} \end{pmatrix}$$

VĚTA 30 (Spektrální rozklad normálního operátoru). *Nechť $\mathbb{A} : V \to V$ je normální operátor, $\mathbb{P}_\lambda : V \to V$ operátor ortogonální projekce na vlastní podprostor příslušný vlastnímu číslu λ. Pak*

$$\mathbb{A} = \sum_{\lambda \in \sigma(\mathbb{A})} \lambda \mathbb{P}_\lambda$$

DŮKAZ. Nechť $(w_i)_1^n$ je ortonormální báze, vzhledem k níž má $\mathbb{A}$ diagonální matici. Vektor w_i je vlastním vektorem $\mathbb{A}$ s vlastním číslem $\lambda \in \sigma(\mathbb{A})$, takže $\mathbb{P}_\lambda w_i = w_i$ a $\mathbb{P}_\mu w_i = 0$, pokud $\mu \neq \lambda$. Pravá i levá strana rovnosti tedy mají na w_i hodnotu λw_i, rovnají se proto na bázi a tudíž i celkově jako operátory. □

PŘÍKLAD 25. Spektrální rozklad pro matici $A = U \operatorname{diag}(4, -2, -2) U^T$ z předchozího příkladu, kde označíme $U = (\mathbf{u}_1 | \mathbf{u}_2 | \mathbf{u}_3)$, vypadá jako

$$A = 4\mathbf{u}_1\mathbf{u}_1^T - 2(\mathbf{u}_2\mathbf{u}_2^T + \mathbf{u}_3\mathbf{u}_3^T)$$

Zde

$$\mathbf{u}_1\mathbf{u}_1^T = \frac{1}{3}\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}, \quad \mathbf{u}_2\mathbf{u}_2^T + \mathbf{u}_3\mathbf{u}_3^T = \frac{1}{3}\begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix}$$

jsou matice $[P_{\mathbf{u}_1}]_K^K$, resp. $[P_{\langle \mathbf{u}_2, \mathbf{u}_3 \rangle}]_K^K$ jsou matice ortogonálních projekcí na jednotlivé vlastní podprostory.

Příkladem spektrálního rozkladu unitární matice může být třeba

$$\begin{pmatrix} \cos\frac{\pi}{2} & -\sin\frac{\pi}{2} \\ \sin\frac{\pi}{2} & \cos\frac{\pi}{2} \end{pmatrix} \equiv \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \lambda_1 \mathbf{w}_1\mathbf{w}_1^+ + \lambda_2 \mathbf{w}_2\mathbf{w}_2^+ =$$

$$= i\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -i \end{pmatrix}\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & i \end{pmatrix} + (-i)\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ i \end{pmatrix}\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -i \end{pmatrix} = i\frac{1}{2}\begin{pmatrix} 1 & i \\ -i & 1 \end{pmatrix} - i\frac{1}{2}\begin{pmatrix} 1 & -i \\ i & 1 \end{pmatrix}$$

Všimněte si, že všechny matice projekcí jsou hermitovské.

Z příkladu vidíme, jak zapsat matici ortogonální projekce na podprostor $W \leq \mathbb{C}^n$, v němž známe ortonormální bázi. Konkrétně ortogonální projekce do sloupcového prostoru $\operatorname{Im}(A)$ matice $A \in \mathbb{C}^{n \times k}$ s lineárně nezávislými sloupci má matici

$$[P_{\operatorname{Im} A}]_K^K = \left[\sum_{i=1}^k P_{q_i}\right]_K^K = \mathbf{q}_1\mathbf{q}_1^+ + \dots + \mathbf{q}_k\mathbf{q}_k^+ = QQ^+,$$

kde $Q = (\mathbf{q}_1 | \dots | \mathbf{q}_k)$ je matice z QR-rozkladu matice A.

VĚTA 31. *Operátor $\mathbb{A} : V \to V$ je samosdružený, právě když je normální a všechna jeho vlastní čísla jsou reálná.*

DŮKAZ. Je-li operátor samosdružený, pak je normální. Pokud $\mathbb{A}v = \lambda v, v \neq 0$ a $\mathbb{A} = \mathbb{A}^*$, pak

$$\lambda\|v\|^2 = \langle v, \lambda v \rangle = \langle v, \mathbb{A}v \rangle = \langle \mathbb{A}^* v, v \rangle = \langle \mathbb{A}v, v \rangle = \langle \lambda v, v \rangle = \bar{\lambda}\|v\|^2$$

Protože $\|v\| \neq 0$, musí být $\bar{\lambda} = \lambda$, neboli $\lambda \in \mathbb{R}$. Naopak je-li $\mathbb{A}$ normální operátor, pak lze vyjádřit vzhledem k ortonormální bázi diagonální maticí s vlastními čísly na diagonále. Jsou-li tato čísla reálná, je matice hermitovská a tedy $\mathbb{A}$ je samosdružený. □

Důsledek 12. *Je-li $A \in \mathbb{R}^{n\times n}$ symetrická matice, existuje ortogonální matice U taková, že matice $U^T AU$ je diagonální.*

Důkaz. Operátor $F_A : \mathbb{C}^n \to \mathbb{C}^n$ je samosdružený a tedy má dle věty 31 reálné spektrum a dle věty 29 je diagonalizovatelný. Existuje tedy báze $\mathbb{C}^n$, která je sjednocením bází jednotlivých vlastních podprostorů matice A. Protože pro každé $\lambda \in \sigma(A)$ je $A-\lambda E$ reálná matice, lze v každém vlastním podprostoru $\mathrm{Ker}(A-\lambda E)$ zvolit bázi složenou pouze z reálných vektorů a tu ortonormalizovat. Spojením všech těchto bází je ortonormální báze $\mathbb{C}^n$ složená pouze z reálných vektorů. Ta je ortonormální bází $\mathbb{R}^n$ a definujeme-li U jako matici přechodu od kanonické báze do této, pak je U ortogonální matice a $U^{-1}AU = U^T AU$ je diagonální. □

Tvrzení 54. *Nechť V nad $\mathbb{C}$ je unitární prostor dimenze n a $\mathbb{U} : V \to V$ je operátor. Pak jsou následující podmínky ekvivalentní*

(1) *$\mathbb{U}$ je unitární.*
(2) *$\forall v \in V$ je $\|\mathbb{U}v\| = \|v\|$*
(3) *$\forall v, w \in V$ je $\langle \mathbb{U}v, \mathbb{U}w\rangle = \langle v, w\rangle$*
(4) *Pokud B je ortonormální báze V, je $U := [\mathbb{U}]^B_B$ unitární matice.*
(5) *Je-li $(w_i)^n_1$ ortonormální báze V, pak je jí i $(\mathbb{U}w_i)^n_1$.*
(6) *$\mathbb{U}$ je normální operátor, jehož vlastní čísla leží na jednotkové kružnici.*

Důkaz. Pokud $\mathbb{U}^*\mathbb{U} = \mathrm{Id}$, pak $\forall v \in V$

$$\|v\|^2 = \langle \mathbb{U}^*\mathbb{U}v, v\rangle = \langle \mathbb{U}v, \mathbb{U}v\rangle = \|\mathbb{U}v\|^2,$$

tedy (1) ⇒ (2). Z polarizační identity 43 máme (2) ⇒ (3). Platí-li (3), pak z $\langle \mathbb{U}^*\mathbb{U}v, w\rangle = \langle v, w\rangle$ po úpravě plyne, že vektor $\mathbb{U}^*\mathbb{U}v - v$ je kolmý na libovolný vektor $w \in V$ a musí být tudíž roven nule. Pokud ale $\mathbb{U}^*\mathbb{U}v = v$ pro libovolný vektor $v \in V$, je $\mathbb{U}^*\mathbb{U} = \mathrm{Id}$. Obdobně ukážeme $\mathbb{U}\mathbb{U}^* = \mathrm{Id}$, dohromady máme implikaci (3) ⇒ (1)

Je-li $(w_i)^n_1$ ortonormální báze V, plyne ze (3), že

$$\langle \mathbb{U}w_i, \mathbb{U}w_j\rangle = \langle w_i, w_j\rangle = \delta_{ij},$$

tedy dostáváme tvrzení (5). Označíme-li $B = (w_i)^n_1$, pak je $[\mathbb{U}]^B_B$ rovna matici přechodu od báze B k bázi $(\mathbb{U}w_i)^n_1$ a díky tvrzení 51 matice je unitární, máme tedy (5) ⇒ (4). Z bodu (4) plyne (1) na základě tvrzení 52.

Unitární operátor je normální a z bodu (2) plyne, že pro $\mathbb{U}v = \lambda v$ musí být $|\lambda|$ rovno jedné. Naopak normální operátor je možné vyjádřit vzhledem k nějaké ortonormální bázi diagonální maticí a leží-li jeho vlastní čísla na jednotkové kružnici, pak tato diagonální matice má tvar $D = \mathrm{diag}(e^{i\alpha_1}, \dots, e^{i\alpha_n})$ pri nějaká $\alpha_i \in \mathbb{R}$. Taková matice je unitární a vzhledem k jakékoli jiné ortonormální bázi má matice operátoru tvar VDV^+, kde V je unitární. Součin VDV^+ je díky tvrzení 50 opět unitární matice, tedy z (6) plyne (4). □

KAPITOLA 16

Singulární rozklad

Říkáme, že soustava rovnic $A\mathbf{x} = \mathbf{b}$ je *přeurčená*, pokud má více rovnic než neznámých a v důsledku toho nemá řešení. Pro takové soustavy rovnic můžeme hledat alespoň řešení aproximativní, pro něž má vektor $A\mathbf{x} - \mathbf{b}$ minimální normu. Uvidíme, že taková úloha je ekvivalentní k hledání ortogonální projekce vektoru $\mathbf{b}$ na podprostor $\operatorname{Im} A$. Podobně pro *podurčené* soustavy, které mají více neznámých než rovnic, a v důsledku toho více řešení, můžeme mezi těmito řešeními hledat to, které má minimální normu $\|x\|$. Obě úlohy vedou na speciální tvary tzv. pseudoinverzní matice. Připomeňme si, že pro A regulární má soustava $A\mathbf{x} = \mathbf{b}$ řešení $\mathbf{x} = A^{-1}\mathbf{b}$. Pseudoinverzní matice $A^\dagger$ umožňuje vyjádřit aproximativní řešení, které má navíc minimální normu, ve tvaru $\mathbf{x} = A^\dagger\mathbf{b}$. Matice $A^\dagger$ se dá spočítat pomocí tzv. singulárního rozkladu (někdy též SVD z anglického singular value decomposition). Ten patří mezi nejdůležitější maticové metody se širokou škálou aplikací především v analýze dat.

1. Přeurčená soustava

Definice 46. Je-li $(u_1, \dots, u_k)$ posloupnost vektorů unitárního prostoru $(V, \langle , \rangle)$ nad $\mathbb{F}$, pak *Gramovou maticí* této posloupnosti rozumíme matici $G \in \mathbb{F}^{k\times k}$, jejíž ij-tý element je $g_{ij} = \langle u_i, u_j\rangle$.

Ze seskvilinearity skalárního součinu plyne, že $g_{ij} = \overline{g}_{ji}$, tedy G je hermitovská matice. Je-li B ortonormální báze V a $A = ([u_1]^B|\dots|[u_k]^B)$, pak $A^+A = G$, neboť ij-tý element matice A^+A je roven

$$\mathbf{a}_i^+\mathbf{a}_j \equiv \langle \mathbf{a}_i, \mathbf{a}_j\rangle = \langle u_i, u_j\rangle$$

Druhá rovnost plyne z tvrzení 49.

Tvrzení 55. *Gramova matice posloupnosti* $(u_1, \dots, u_k)$ *je regulární, právě když je tato posloupnost lineárně nezávislá.*

Důkaz. Plyne z šestého bodu tvrzení 53. □

Ukažme nyní, jak se dá pomocí Gramovy matice formulovat hledání ortogonální projekce na podprostor a najít matici této projekce bez nutnosti hledat v podprostoru ortonormální bázi.

Uvažujme $W := \langle u_1, \dots, u_k\rangle$ podprostor unitárního prostoru V a vektor $v \in V$, který v něm obecně neleží. Jeho projekce $w := P_W(v)$ bude lineární kombinací generátorů W, tedy $w = \sum_{i=1}^k x_i u_i$ pro nějaké skaláry x_i. Vektor w bude ortogonální projekce vektoru v, právě když je jejich rozdíl kolmý na všechny prvky množiny generátorů W, tedy pakliže $\forall i : \langle u_i, v - w\rangle = 0$. Dosadíme-li za w a použijeme linearitu skalárního součinu ve druhém argumentu, dostáváme pro koeficienty $x_1, \dots, x_k$ lineární rovnice tvaru

$$\langle u_i, v\rangle = x_1\langle u_i, u_1\rangle + \dots + x_k\langle u_i, u_k\rangle$$

Ty lze dohromady zapsat jako soustavu $G\mathbf{x} = \mathbf{y} \in \mathbb{F}^k$, jejíž matice je Gramova matice G a složky pravé strany jsou $y_i := \langle u_i, v\rangle$.

Zvolme ve V nějakou ortonormální bázi B a označme $\mathbf{b} = [v]^B \in \mathbb{F}^n$, $A = ([u_1]^B|\dots|[u_k]^B)$. Soustavu $G\mathbf{x} = \mathbf{y}$ lze pak přepsat do tvaru

$$A^+A\mathbf{x} = A^+\mathbf{b}$$

To je tzv. *normální soustava* příslušející soustavě rovnic $A\mathbf{x} = \mathbf{b}$. Můžeme se na to dívat tak, že jsme jakoby nejprve zkoušeli řešit soustavu $A\mathbf{x} = \mathbf{b}$, která je souřadnicovým vyjádřením hledání $\mathbf{x} \in \mathbb{F}^k$ takového, aby $\sum_{i=1}^k x_i u_i = v$. Není-li $v \in W$, tedy nemá-li přeurčená soustava $A\mathbf{x} = \mathbf{b}$ řešení, nalezne její normální soustava vektor minimalizující hodnotu $\|\sum_{i=1}^k x_i u_i - v\|$ nebo ekvivalentně $\|A\mathbf{x} - \mathbf{b}\|$, tzv. *aproximativní řešení*. Najde tedy $\mathbf{x}$ takové, pro něž se levá strana $A\mathbf{x}$ liší od pravé strany $\mathbf{b}$ nejméně, měřeno normou vektoru.

Má-li matice A lineárně nezávislé sloupce, je A^+A regulární a má tedy inverzní matici. Pak můžeme řešení normální soustavy zapsat jako $\mathbf{x} = (A^+A)^{-1}A^+\mathbf{b}$. Matici projekce P_W můžeme pak odvodit z rovnosti

$$[P_W]_B^B[v]^B = [P_W(v)]_B^B = [w]^B = \sum_{i=1}^k x_i[u_i]^B = A\mathbf{x} = A(A^+A)^{-1}A^+[v]^B,$$

neboli $[P_W]_B^B = A(A^+A)^{-1}A^+$. Je-li W definován přímo jako sloupcový prostor nějaké matice A, můžeme tento výsledek zapsat jako

Tvrzení 56. *Je-li $A \in \mathbb{C}^{n\times k}$ matice s lineárně nezávislými sloupci a K kanonická báze v $\mathbb{C}^n$, pak*

$$[P_{\operatorname{Im} A}]_K^K = A(A^+A)^{-1}A^+$$

Klasický příklad užití aproximativního řešení je lineární regrese, tj. prokládání funkce zadané v n bodech lineární funkcí metodou nejmenších čtverců. Minimalizace výrazu

$$\sum_{i=1}^n |ax_i + b - y_i|^2$$

vzhledem k $a, b \in \mathbb{R}$ se zadanými $(x_i, y_i), i \in \{1, \dots, n\}$ vede na aproximativní řešení přeurčené soustavy lineárních rovnic

$$\begin{pmatrix} x_1 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix},$$

jejíž normální soustava je

$$\begin{pmatrix} \sum_{i=1}^n x_i^2 & \sum_{i=1}^n x_i \\ \sum_{i=1}^n x_i & n \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^n x_i y_i \\ \sum_{i=1}^n y_i \end{pmatrix}$$

Tu je možné explicitně vyřešit třeba Cramerovým pravidlem, tj.

$$a = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{n\sum x_i^2 - (\sum x_i)^2}, \quad b = \frac{\sum x_i^2 \sum y_i - \sum x_i \sum x_i y_i}{n\sum x_i^2 - (\sum x_i)^2}$$

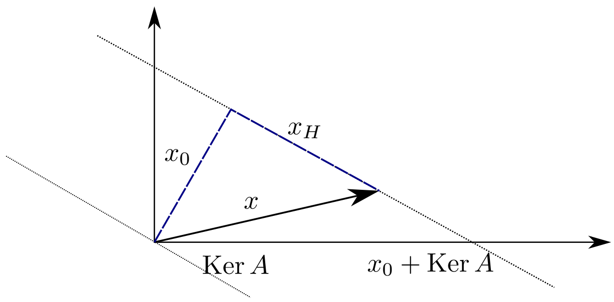
Meze sum jsou pro lepší čitelnost vynechány. Podobně je možné spočítat libovolnou regresní funkci, která se dá vyjádřit jako lineární kombinace konečně mnoha daných funkcí, tedy například proložit data polynomem zvoleného stupně, nebo hledat regresní funkce více proměnných.

2. Podurčená soustava

Uvažujme nyní případ podurčené soustavy rovnic $A\mathbf{x} = \mathbf{b}$, $A \in \mathbb{C}^{m\times n}$, která má množinu řešení tvaru $\mathbf{x}_p + \operatorname{Ker} A$, kde $\operatorname{Ker} A \neq \{\mathbf{o}\}$. Ortogonální doplněk jádra je řádkový prostor komplexně sdružené matice, z čehož spolu s tvrzením 46 plyne, že

$$\mathbb{C}^n = \operatorname{Ker} A \oplus (\operatorname{Ker} A)^\perp = \operatorname{Ker} A \oplus \operatorname{Im} A^+$$

Každý vektor z $\mathbb{C}^n$ a tudíž i každé řešení soustavy $A\mathbf{x} = \mathbf{b}$ můžeme tedy jednoznačně rozložit na součet $\mathbf{x} = \mathbf{x}_0 + \mathbf{x}_H$, kde $\mathbf{x}_0 \in \operatorname{Im} A^+$ a $\mathbf{x}_H \in \operatorname{Ker} A$.



Obrázek 1. Řešení s nejmenší normou

Jak $\mathbf{x}_0$, tak $\mathbf{x}_H$ jsou ortogonální projekcí vektoru $\mathbf{x}$ na příslušný podprostor, speciálně $\mathbf{x}_0 = P_{\operatorname{Im} A^+}(\mathbf{x})$. Ze čtvrtého bodu tvrzení 45 plyne, že $\|\mathbf{x}_0\| \leq \|\mathbf{x}\|$, a tedy $\mathbf{x}_0$ má mezi všemi řešeními soustavy $A\mathbf{x} = \mathbf{b}$ nejmenší normu. Navíc z $\mathbf{x}_0 \in \operatorname{Im} A^+$ plyne existence vektoru $\mathbf{z} \in \mathbb{C}^m$ takového, že $\mathbf{x}_0 = A^+\mathbf{z}$. Platí tedy

$$AA^+\mathbf{z} = A\mathbf{x_0} = \mathbf{b}$$

Vektor $\mathbf{z}$ nemusí být určen jednoznačně, ale splňuje-li soustavu $AA^+\mathbf{z} = \mathbf{b}$, pak je $\mathbf{x}_0 = A^+\mathbf{z}$ řešením soustavy $A\mathbf{x} = \mathbf{b}$ s nejmenší normou. Má-li A lineárně nezávislé řádky, je AA^+ regulární, $\mathbf{z}$ je určen jednoznačně a $\mathbf{x}_0 = A^+\mathbf{z} = A^+(AA^+)^{-1}\mathbf{b}$. To nám umožňuje zapsat vztah mezi obecným řešením $\mathbf{x}$ soustavy $A\mathbf{x} = \mathbf{b}$ a řešením $\mathbf{x}_0$ s nejmenší normou jako

$$\mathbf{x}_0 = A^+(AA^+)^{-1}A\mathbf{x} = [P_{\operatorname{Im} A^+}]_K^K\,\mathbf{x} \equiv P_{\operatorname{Im} A^+}(\mathbf{x}),$$

což nabízí alternativní způsob důkazu tvrzení 56.

3. Pseudoinverzní matice

Definice 47. Nechť $A \in \mathbb{C}^{m\times n}$, pak matice $A^\dagger \in \mathbb{C}^{n\times m}$ splňující

(1) $AA^\dagger A = A$
(2) $A^\dagger AA^\dagger = A^\dagger$
(3) $AA^\dagger$ i $A^\dagger A$ jsou hermitovské matice

se nazývá *Moore-Penroseova pseudoinverze* matice A.

To je axiomatická definice, není z ní jasné, že $A^\dagger$ existuje nebo že je jen jedna. Pro A s lineárně nezávislými sloupci je díky poslednímu bodu tvrzení 53 A^+A regulární matice a definujeme-li

$$A^\dagger := (A^+A)^{-1}A^+,$$

vidíme okamžitě, že první dvě podmínky jsou splněny. U třetí zjevně $A^\dagger A = E_n$ a z $AA^\dagger = [P_{\operatorname{Im} A}]_K^K$, jak plyne z tvrzení 56. Z vět 30 a 31 i z explicitního tvaru matic projekcí plyne, že $[P_{\operatorname{Im} A}]_K^K$ je hermitovská matice, E_n je očividně také.

Podobně pro A s lineárně nezávislými řádky je AA^+ regulární a o

$$A^\dagger := A^+(AA^+)^{-1}$$

z řešení podurčené soustavy víme, že $A^\dagger A = [P_{\operatorname{Im} A^+}]_K^K$. I takto definovaná $A^\dagger$ tedy splňuje podmínky definice 47. Pro A regulární v obou případech vyjde $A^\dagger = A^{-1}$.

Pro obecný případ, kdy A nemá lineárně nezávislé řádky ani sloupce, musíme zkonstruovat $A^\dagger$ jinak. Cestou k tomu bude singulární rozklad, který existuje pro každou matici A a lze pomocí něj jednoduše napsat matici splňující podmínky definice 47. Nejprve ještě dokážeme jednoznačnost:

TVRZENÍ 57. *Nechť $A \in \mathbb{C}^{m\times n}$ a matice $B, C \in \mathbb{C}^{n\times m}$ splňují podmínky definice 47. Pak $B = C$.*

DůKAZ. Vlastnost 1 říká, že $ACA = A$, tedy $AB = ACAB$. Podle vlastnosti 3 jsou AC i AB hermitovské, z vlastností hermitovského sdružení a z rovnosti $ABA = A$ pak plyne

$$ACAB = (AC)^+(AB)^+ = C^+A^+(AB)^+ = C^+(ABA)^+ = C^+A^+ = (AC)^+ = AC$$

Tedy $AB = AC$, podobně ukážeme $BA = CA$. Odtud následně z vlastnosti 2

$$B = BAB = BAC = CAC = C$$

□

Řekneme, že samosdružený operátor $\mathbb{B} : V \to V$ je *pozitivně semidefinitní*, pokud má všechna vlastní čísla nezáporná. Je-li C ortonormální báze V, pak díky větě 29 existuje unitární matice U taková, že

$$[\mathbb{B}]_C^C = U \begin{pmatrix} \lambda_1 & 0 & \\ 0 & \ddots & \\ & & \lambda_n \end{pmatrix} U^+,$$

kde $\lambda_i \geq 0$. Označme D diagonální matici uprostřed a pro vektor $v \in V$ s reprezentací $[v]^C =: \mathbf{x}$ označme $\mathbf{y} = U^+\mathbf{x}$. Pak z tvrzení 49 plyne, že

$$\langle v, \mathbb{B}v \rangle = \mathbf{x}^+ UDU^+ \mathbf{x} = \mathbf{y}^+ D \mathbf{y} = \sum_{i=1}^n \lambda_i |y_i|^2$$

Tedy $\mathbb{B}$ je pozitivně semidefinitní, právě když $\langle v, \mathbb{B}v \rangle \geq 0$ pro každý vektor $v \in V$. Pro pozitivně semidefinitní operátor $\mathbb{B}$ pak můžeme definovat jeho pozitivně semidefinitní odmocninu $\sqrt{\mathbb{B}} : V \to V$ operátoru $\mathbb{B}$ předpisem

$$[\sqrt{\mathbb{B}}]_C^C := U \begin{pmatrix} \sqrt{\lambda_1} & 0 & \\ 0 & \ddots & \\ & & \sqrt{\lambda_n} \end{pmatrix} U^+$$

Nechť nyní $\mathbb{A} : V \to W$ je libovolný operátor. Pak operátor $\mathbb{A}^*\mathbb{A}$ je samosdružený a

$$\langle \mathbf{v}, \mathbb{A}^*\mathbb{A}\mathbf{v} \rangle = \langle \mathbb{A}\mathbf{v}, \mathbb{A}\mathbf{v} \rangle = \|\mathbb{A}\mathbf{v}\|^2 \geq 0,$$

tedy i pozitivně semidefinitní. Jeho odmocninu $|\mathbb{A}| := \sqrt{\mathbb{A}^*\mathbb{A}}$ nazýváme *modul operátoru* $\mathbb{A}$. Modul má díky poslednímu bodu tvrzení 53 stejné jádro i hodnost jako původní operátor. Nenulová vlastní čísla modulu $|\mathbb{A}|$ se označují jako *singulární hodnoty operátoru* $\mathbb{A}$, je jich tedy včetně násobností právě $r := \operatorname{rank}(\mathbb{A})$. Označují se $\sigma_1, \ldots, \sigma_r$ a konvenčně bývají řazeny sestupně podle velikosti. Modul $|\mathbb{A}|$ může mít i vlastní číslo nula, to se mezi singulární hodnoty operátoru $\mathbb{A}$ nepočítá.

Označením $D = \operatorname{diag}_{m\times n}(\sigma_1, \ldots, \sigma_r)$ rozumíme obdélníkovou matici, na jejíž diagonále $d_{11}, d_{22}, \ldots$ jsou postupně čísla $\sigma_1, \sigma_2, \ldots, \sigma_r$ a zbytek diagonály i matice je nulový.

VĚTA 32. *Nechť $\mathbb{A} : V \to W$ je operátor mezi unitárními prostory. Pak existují ortonormální báze $B = (v_i)_1^n$ ve V, $C = (u_i)_1^m$ ve W takové, že $[\mathbb{A}]_B^C = \mathrm{diag}_{m\times n}(\sigma_1, \ldots, \sigma_r)$, kde σ_i jsou singulární hodnoty operátoru $\mathbb{A}$.*

DŮKAZ. Bázi B volme tak, aby vůči ní byl operátor $\mathbb{A}^*\mathbb{A}$ diagonální a byla seřazena tak, aby $v_1, \ldots, v_r$ byly vlastní vektory operátoru $\mathbb{A}^*\mathbb{A}$ s vlastními čísly $\sigma_1^2, \ldots, \sigma_r^2$ a zbylé $v_{r+1}, \ldots, v_n$ příslušely vlastnímu číslu 0. Prvních r prvků báze C pak definujme předpisem $u_i := \frac{1}{\sigma_i}\mathbb{A}v_i$. Potom

$$\langle u_i, u_j\rangle = \left\langle \frac{1}{\sigma_i}\mathbb{A}v_i, \frac{1}{\sigma_j}\mathbb{A}v_j \right\rangle = \frac{1}{\sigma_i\sigma_j}\langle v_i, \mathbb{A}^*\mathbb{A}v_j\rangle = \frac{\sigma_j^2}{\sigma_i\sigma_j}\langle v_i, v_j\rangle = \delta_{ij},$$

tedy $(u_1, \ldots, u_r)$ je ortonormální posloupnost. Tu můžeme na základě Lemmatu 6 doplnit vektory $u_{r+1}, \ldots, u_m$ na ortonormální bázi prostoru W. Pak $\mathbb{A}v_i = \sigma_i u_i$ pro $i \leq r$ a nula jindy, protože z pátého bodu tvrzení 53 a z vlastností báze B plyne, že

$$\mathrm{Ker}\,\mathbb{A} = \mathrm{Ker}\,\mathbb{A}^*\mathbb{A} = \langle v_{r+1}, \ldots, v_n\rangle.$$

Tedy $[\mathbb{A}v_i]^C$ je vektor, který má na pozici i vlastní číslo σ_i a jinak samé nuly, tudíž matice $[\mathbb{A}]_B^C$ má požadovaný tvar. □

Je-li operátorem $\mathbf{A}$ zobrazení F_A určené maticí $A \in \mathbb{C}^{m\times n}$ mezi prostory $\mathbb{C}^n$ a $\mathbb{C}^m$ se standardními skalárními součiny, vyplývá z věty 32, že lze matici A zapsat jako

$$A = U\Sigma V^+ = \sigma_1\mathbf{u}_1\mathbf{v}_1^+ + \ldots + \sigma_r\mathbf{u}_r\mathbf{v}_r^+,$$

kde $U = (\mathbf{u}_1|\ldots|\mathbf{u}_m) \in \mathbb{C}^{m\times m}$ a $V = (\mathbf{v}_1|\ldots|\mathbf{v}_n) \in \mathbb{C}^{n\times n}$ jsou unitární matice a $\Sigma = \mathrm{diag}_{m\times n}(\sigma_1, \ldots, \sigma_r)$. Prvnímu zápisu se říká *singulární rozklad* matice A, druhý z něj vznikne analogicky jako vznikl spektrální rozklad z věty 30 z ortogonální diagonalizace z věty 29. Na rozdíl od spektrálního rozkladu ale matice $\mathbf{u}_i\mathbf{v}_i^+$ hodnosti 1 neodpovídají ortogonálním projekcím, třeba už jen proto, že nejsou nutně čtvercové.

Návod k jeho nalezení nám dává důkaz věty 32:

(1) Ortogonální diagonalizací vyjádříme matici A^+A jako součin $A^+A = V\Lambda V^+$, kde $V = (\mathbf{v}_1|\ldots|\mathbf{v}_n)$ je unitární a $\Lambda = \mathrm{diag}(\lambda_1, ..., \lambda_n)$ má vlastní čísla seřazená od největšího po nejmenší, počínaje λ_{r+1} jsou to nuly.
(2) Singulární hodnoty jsou $\sigma_i = \sqrt{\lambda_i}$, pro $i \leq r$, potom $\Sigma = \mathrm{diag}_{m\times n}(\sigma_1, ..., \sigma_r)$
(3) Prvních r sloupců matice U je dáno jako $\mathbf{u}_i = \frac{1}{\sigma_i}A\mathbf{v}_i$, zbylé nalezneme libovolným doplněním na unitární matici.

Singulární rozklad má celou řadu aplikací. Pokud ze součtu $\sum_{i=1}^r \sigma_i\mathbf{u}_i\mathbf{v}_i^+$ odebereme určitý počet členů od konce, dostaneme tzv. *low-rank aproximaci* matice A. Dá se ukázat, že například matice $\sigma_1\mathbf{u}_1\mathbf{v}_1^+$ je matice hodnosti 1, která je v určitém přesném smyslu nejméně vzdálená od matice A. To se dá využít ke kompresi dat (například je-li A obrázek o $m \times n$ pixelech, dá se jeho aproximace ranku r uložit jen jako $(m+n+1)r$ čísel) nebo k tzv. analýze hlavních komponent (PCA), která pomáhá nalézt veličiny, které nesou nejvíce informací o daném souboru dat.

Pomocí singulárního rozkladu lze také popsat, jak lineární zobrazení natahuje či zkracuje vektory. Uvažujme jednotkovou kouli $K_1 := \{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x}\| = 1\}$ a zobrazme ji pomocí zobrazení $F_A : \mathbb{R}^n \to \mathbb{R}^m$. Matice V v singulárním rozkladu A bude reálná a ortogonální a stejně tak i U. Zobrazení pak můžeme zapsat jako $F_A = F_U \circ F_\Sigma \circ F_{V^T}$. Z tvrzení 54 víme, že

$$\|F_{V^T}(\mathbf{x})\| = \|V^T\mathbf{x}\| = \|\mathbf{x}\|$$

a V^T je regulární, tedy $F_{V^T}(K_1) = K_1$. Násobení maticí Σ převede K_1 na r-rozměrný elipsoid v $\mathbb{R}^m$ s poloosami délek $\sigma_1, \ldots, \sigma_r$. Poslední zobrazení F_U opět zachová normu a tedy i délky poloos elipsoidu. Pokud bychom nahradili A její aproximací hodnosti 1, zobrazila by se jednotková koule na úsečku, která je nejdelší osou elipsoidu, kdybychom zvolili aproximaci hodnosti 2, byla by výsledkem elipsa, jejíž osy by byly dvě nejdelší osy původního elipsoidu. Tento geometrický pohled poskytuje i intuitivní důvody, proč jsou oříznuté sumy nejbližšími aproximacemi daného zobrazení ranku 1, 2, atd.

Naším plánem bylo ale především využít singulární rozklad ke konstrukci Moore-Penroseovy pseudoinverzní matice. Definujeme-li $\Sigma^\dagger := \mathrm{diag}_{n\times m}(\sigma_1^{-1}, \dots, \sigma_r^{-1})$ (pozor, to není nutně inverzní matice k Σ), a

$$A^\dagger := V\Sigma^\dagger U^+,$$

pak

$$AA^\dagger A = U\Sigma V^+ V\Sigma^\dagger U^+ U\Sigma V^+ = U\Sigma\Sigma^\dagger\Sigma V^+ = U\Sigma V^+ = A$$

a podobně ověříme i zbývající podmínky v definici 47. Následující věta osvětlí, proč matici $A^\dagger$ říkáme pseudoinverzní:

VĚTA 33. *Nechť $A \in \mathbb{C}^{m\times n}$, $\mathbf{b} \in \mathbb{C}^m$. Pak vektor $A^\dagger\mathbf{b}$ je aproximativním řešením soustavy rovnic $A\mathbf{x} = \mathbf{b}$, které má navíc nejmenší možnou normu.*

DŮKAZ. Předpokládejme nejprve, že $A = \mathrm{diag}_{m\times n}(\sigma_1, ..., \sigma_r)$, kde $\sigma_1 \geq \dots \geq \sigma_r > 0$. Pak

$$\|A\mathbf{x} - \mathbf{b}\|^2 = \sum_{i=1}^{r} |\sigma_i x_i - b_i|^2 + \sum_{i=r+1}^{m} |b_i|^2$$

bude pro volbu $x_i = \frac{1}{\sigma_i} b_i$ pro $i = 1, \dots, r$ rovno druhé sumě, která sama na $\mathbf{x}$ nezávisí a je minimální možnou hodnotou výrazu $\|A\mathbf{x} - \mathbf{b}\|^2$. Norma vektoru

$$\mathbf{x} = \left(\frac{1}{\sigma_1} b_1, \dots, \frac{1}{\sigma_r} b_r, x_{r+1}, \dots, x_n\right)^T,$$

v němž už máme prvních r složek pevně určených, bude minimální pro $x_{r+1} = \dots = x_n = 0$. Celkově tedy

$$\mathbf{x} = \mathrm{diag}_{n\times m}(\sigma_1^{-1}, \dots, \sigma_r^{-1})\mathbf{b} = A^\dagger\mathbf{b}$$

Uvažujme nyní obecnou matici se singulárním rozkladem $A = U\Sigma V^+$. Norma vektorů $U^+ A\mathbf{x} - U^+\mathbf{b}$ a $A\mathbf{x} - \mathbf{b}$ je stejná. Stejně tak je stejná norma vektorů $\mathbf{x}$ a $\mathbf{x}' := V^+\mathbf{x}$. Tedy

$$\min_{\mathbf{x}} \|U\Sigma V^+\mathbf{x} - \mathbf{b}\| = \min_{\mathbf{x}} \|\Sigma V^+\mathbf{x} - U^+\mathbf{b}\| = \min_{\mathbf{x}'} \|\Sigma\mathbf{x}' - U^+\mathbf{b}\|$$

a poslední výraz nabývá dle diskuse v první části důkazu minima s minimální normou právě pro $\mathbf{x}' = \Sigma^\dagger U^+\mathbf{b}$. Tedy $\mathbf{x} = V\mathbf{x}' = V\Sigma^\dagger U^+\mathbf{b} = A^\dagger\mathbf{b}$. □

Poslední aplikací singulárního rozkladu, kterou zde uvedeme, bude věta, na základě níž je možné každou čtvercovou matici zapsat jakou součin unitární a hermitovské matice. Speciálně pro reálné 3×3 matice to znamená rozklad matice na součin matice rotace a matice symetrické, tedy s reálnými vlastními čísly. To se hodí například v mechanice kontinua při zavedení tenzoru malých deformací, tedy vlastně k rozdělení infinitezimálního pohybu kontinua na část rotační a část deformační.

Upravme singulární rozklad čtvercové matice $A \in \mathbb{C}^{n\times n}$ na

$$A = U\Sigma V^+ = UV^+ V\Sigma V^+ = WP,$$

kde $W := UV^+$ je unitární matice a $P := V\Sigma V^+$ je pozitivně semidefinitní hermitovská matice. V řeči lineárních zobrazení tato úvaha dává větu o *polárním rozkladu operátoru:*

Věta 34. *Nechť* $\mathbb{A} : V \to V$ *je operátor na unitárním prostoru. Pak existují unitární operátor* $\mathbb{W}$ *a pozitivně semidefinitní samosdružený operátor* $\mathbb{P}$ *takové, že* $\mathbb{A} = \mathbb{W}\mathbb{P}$. *Navíc* $\mathbb{P} = |\mathbb{A}|$ *a pokud je* $\mathbb{A}$ *izomorfismus, pak je* $\mathbb{W}$ *určen jednoznačně.*

Důkaz. Je-li C ortonormální báze V, definujme $A := [\mathbb{A}]_C^C$. Z konstrukce modulu víme, že je určen jednoznačně a že $[|\mathbb{A}|]_C^C = V\Sigma V^+$, kde Σ, V jsou matice vyskytující se v singulárním rozkladu $A = U\Sigma V^+$. Definujeme-li $\mathbb{W}$ na bázi jako $[\mathbb{W}]_C^C := W = UV^+$, je to podle tvrzení 54 unitární operátor. Je-li $\mathbb{A}$ izomorfismus, pak je jím i podle posledního bodu tvrzení 53 $\mathbb{A}^*\mathbb{A}$ a $|\mathbb{A}| =: \mathbb{P}$. Pak $\mathbb{W} = \mathbb{A}\mathbb{P}^{-1}$ je rovněž izomorfismus a je tímto předpisem určen jednoznačně. □

V dimenzi 1 odpovídá polární rozklad zápisu komplexního čísla v polárních souřadnicích:

$$z = a + ib = \left(\frac{a}{\sqrt{a^2+b^2}} + i\frac{b}{\sqrt{a^2+b^2}} \right) \sqrt{a^2+b^2} = e^{i\alpha}|z|$$

Zde lze $|z|$ chápat jako pozitivně semidefinitní 1×1 matici a $e^{i\alpha} \equiv \cos\alpha + i\sin\alpha$ jako unitární 1×1 matici.

Poznámka 5. Je-li $A = U\Sigma V^+$ singulární rozklad matice A, v němž $\Sigma = \mathrm{diag}_{m\times n}(\sigma_1, \ldots, \sigma_r)$, pak $A^+ = V\Sigma^+U^+$ je singulární rozklad matice A^+, a protože $\Sigma^+ = \mathrm{diag}_{n\times m}(\sigma_1, \ldots, \sigma_r)$, má A^+ stejné singulární hodnoty jako A. Je možné je najít jako kladné odmocniny z nenulových vlastních čísel matice $(A^+)^+A^+ = AA^+$. Také z tvrzení 42 víme, že A^+A a AA^+ mají stejná nenulová vlastní čísla se stejnými algebraickými násobnostmi. Je-li $A^+A\mathbf{v} = \lambda\mathbf{v}$, platí $\lambda A\mathbf{v} = A(A^+A)\mathbf{v} = AA^+(A\mathbf{v})$, tedy vlastní vektor matice A^+A zobrazí matice A na vlastní vektor matice AA^+. Normalizované vlastní vektory matice A^+A jsou sloupce matice V, jejich A-obrazy jsou buď rovny nulovému vektoru, nebo jsou po normalizaci rovny sloupcům matice U. To vyplývá jak z důkazu věty 32, tak z výše uvedené úvahy, že $A^+ = V\Sigma^+U^+$. Naopak vlastní vektory matice AA^+ zobrazuje matice A^+ na vlastní vektory matice A^+A. Tedy prvních r sloupců matice U tvoří bázi $\mathrm{Im}\, A$ a matice A^+ je až na násobek zobrazuje na prvních r sloupců matice V, zbylých $m - r$ sloupců U je bází prostoru $(\mathrm{Im}\, A)^\perp = \mathrm{Ker}\, A^+$. Podobně prvních r sloupců matice V tvoří bázi $\mathrm{Im}\, A^+$ a matice A je až na násobek zobrazuje na prvních r sloupců matice U, zbylých $n - r$ sloupců V je bází prostoru $(\mathrm{Im}\, A^+)^\perp = \mathrm{Ker}\, A$.

KAPITOLA 17

Kvadratické formy

1. Polární báze kvadratické formy

DEFINICE 48. Nechť g je bilineární forma na V. Zobrazení $Q_g : V \to \mathbb{R}$, které je definované předpisem $Q_g(u) := g(u,u)$, nazýváme *kvadratická forma* na V příslušná bilineární formě g.

PŘÍKLAD 26. Bilineární forma na $\mathbb{R}^2$

$$g(\mathbf{x},\mathbf{y}) = \mathbf{x}^T \begin{pmatrix} 1 & -3 \\ 1 & 1 \end{pmatrix} \mathbf{y} = x_1y_1 - 3x_1y_2 + x_2y_1 + x_2y_2$$

určuje kvadratickou formu

$$Q_g(\mathbf{x}) = \mathbf{x}^T \begin{pmatrix} 1 & -3 \\ 1 & 1 \end{pmatrix} \mathbf{x} = x_1^2 - 2x_1x_2 + x_2^2$$

Stejnou kvadratickou formu ovšem určuje i bilineární forma

$$h(\mathbf{x},\mathbf{y}) = \mathbf{x}^T \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \mathbf{y} = x_1y_1 - x_1y_2 - x_2y_1 + x_2y_2$$

DEFINICE 49. Nechť g je bilineární forma na V. Její *symetrizace* a *antisymetrizace* jsou bilineární formy definované $\forall u,v \in V$ předpisy

$$g_S(u,v) := \frac{1}{2}(g(u,v) + g(v,u))$$
$$g_A(u,v) := \frac{1}{2}(g(u,v) - g(v,u))$$

Hned vidíme, že $g_S(u,v) = g_S(v,u)$, g_S je tedy opravdu symetrická, a podobně $g_A(u,v) = -g_A(v,u)$, takové bilineární formě říkáme *antisymetrická*. Protože

$$g(u,v) = g_S(u,v) + g_A(u,v),$$

dá se libovolná bilineární forma rozložit na součet symetrické a antisymetrické bilineární formy. Pro formu z našeho příkladu je tímto rozkladem

$$\mathbf{x}^T \begin{pmatrix} 1 & -3 \\ 1 & 1 \end{pmatrix} \mathbf{y} = \mathbf{x}^T \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \mathbf{y} + \mathbf{x}^T \begin{pmatrix} 0 & -2 \\ 2 & 0 \end{pmatrix} \mathbf{y}$$

Antisymetrie g_A znamená také, že $g_A(u,u) = 0$, proto

$$Q_g(u) \equiv g(u,u) = g_S(u,u),$$

tedy kvadratická forma je určena pouze symetrickou částí bilineární formy g. Z polarizační identity 43 plyne, že

$$g_S(u,v) = \frac{1}{2}(Q_g(u+v) - Q_g(u) - Q_g(v)),$$

tj. že symetrická bilineární forma g_S je naopak určena kvadratickou formou Q_g. Zjistili jsme tedy, že:

TVRZENÍ 58. *Kvadratická forma Q_g je jednoznačně určena symetrickou částí g_S bilineární formy g.*

Je-li B báze V a $G = [g]_B$ matice bilineární formy, pak pro $u, v \in V$, $\mathbf{x} = [u]^B$, $\mathbf{y} = [v]^B$ je

$$g(u, v) = \mathbf{x}^T G \mathbf{y}$$

Symetrická část g_S má matici $G_S := \frac{1}{2}(G + G^T)$, antisymetrická g_A matici $G_A := \frac{1}{2}(G - G^T)$. Kvadratickou formu je možné psát v souřadnicích jako

$$Q_g(u) = \mathbf{x}^T G \mathbf{x} = \mathbf{x}^T G_S \mathbf{x}$$

To umožňuje *matici kvadratické formy* $[Q_g]_B$ vzhledem k bázi B definovat jako matici $G_S = [g_S]_B$ symetrické části příslušné bilineární formy. V dalším textu budeme předpokládat, že symetrická je už přímo g, a tedy $G = G_S$.

Polární báze kvadratické formy je definována jako polární báze příslušné symetrické bilineární formy g, tedy taková B', že $[g]_{B'}$ je diagonální. Z věty 22 víme, že polární báze existuje a umíme ji nalézt symetrickými úpravami, tedy vlastně hledáním matice přechodu $R = [\mathrm{Id}]^B_{B'}$ takové, že $G' := R^T G R$ je diagonální.

PŘÍKLAD 27. Různými posloupnostmi symetrických úprav můžeme získat rovnosti

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$
$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 2 & 2 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 0 & 2 \end{pmatrix}$$
$$\begin{pmatrix} 4 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix}$$

Sloupce poslední matice na řádku tvoří vždy nějakou polární bázi symetrické bilineární formy h na $\mathbb{R}^2$.

Jinou možnost, jak polární bázi najít, dává důsledek 12. Podle něj k symetrické matici G existuje ortogonální matice U taková, že $G' := U^T G U$ je diagonální matice, jejíž diagonála je tvořena vlastními čísly matice G. Interpretujeme-li matici U jako matici přechodu od báze B k nějaké bázi B', je G' maticí bilineární formy g vzhledem k této bázi B', tedy je to polární báze. Její prvky jsou vzhledem k bázi B reprezentovány vlastními vektory matice G. Byla-li původní báze B ortonormální vůči nějakému skalárnímu součinu ve V, je díky ortogonalitě matice U a tvrzení 49 ortonormální i polární báze B'. Je to tedy vlastně polární báze pro dvě různé symetrické bilineární formy na V: pro g a pro onen skalární součin. Například pokud je B kanonická báze v $\mathbb{R}^n$ se standardním skalárním součinem, pak je i $B' = (\mathbf{u}'_i)_1^n$ ortonormální polární báze g, tedy jak Gramova matice $((\mathbf{u}'_i, \mathbf{u}'_j))$, tak matice $[g]_{B'} = (g(\mathbf{u}'_i, \mathbf{u}'_j))$ jsou obě diagonální, Gramova je přímo rovna jednotkové matici.

PŘÍKLAD 28. Rovnost

$$\begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$$

získáme nalezením vlastních čísel a normalizovaných vlastních vektorů matice vpravo uprostřed. Sloupce poslední matice dávají opět polární bázi, která je navíc ortonormální vzhledem ke standardnímu skalárnímu součinu. V tomto i předchozím příkladě se jednotlivé matice na levé straně, které jsou vyjádřením h vzhledem k různým polárním bázím, liší. Na rozdíl od endomorfismů tedy nelze mluvit o jednom konkrétním diagonálním tvaru symetrické bilineární formy.

Využití ortogonální diagonalizace má výhodu na unitárních prostorech, protože změna souřadnic zachovává velikosti a úhly vektorů. V kapitole o kvadrikách uvidíme, že se pak při přechodu do nových souřadnic zachovají například délky poloos u elipsoidu, zatímco při diagonalizaci symetrickými úpravami přejde obecně elipsoid v jiný elipsoid s jinou délkou poloos. Výhoda symetrických úprav je zase v tom, že nevyžadují výpočet vlastních čísel a dají se tedy vždy provést v přesné aritmetice. Následující věta ukazuje, že ač je polárních bází i způsobů jejich nalezení mnoho, něco mají přece jen všechny společné:

2. Signatura

VĚTA 35 (Sylvesterův zákon setrvačnosti kvadratických forem). *Nechť B, C jsou dvě polární báze kvadratické formy Q_g na V. Pak je počet kladných hodnot na diagonále matic $[g]_B$ a $[g]_C$ stejný, a stejně tak počet záporných hodnot a počet nul.*

DŮKAZ. Matice $[g]_B$, $[g]_C$ vzniknou jedna z druhé symetrickými úpravami a ty zachovávají hodnost, počet nul tedy musí být stejný. Stačí proto dokazovat jen shodnost počtu kladných hodnot, a to sporem.

Nechť $B = (u_i)_1^m$ a $C = (v_i)_1^m$ jsou seřazeny tak, že v maticích

$$[g]_B = \operatorname{diag}(a_1, \ldots, a_p, b_1, \ldots, b_q, 0, \ldots, 0)$$
$$[g]_C = \operatorname{diag}(c_1, \ldots, c_s, d_1, \ldots, d_t, 0, \ldots, 0)$$

jsou všechna čísla a_i, c_i kladná a všechna b_i, d_i záporná. Víme, že $p + q = s + t$, bez újmy na obecnosti předpokládejme $p > s$, jinak můžeme prohodit báze B a C. Definujme podprostory $U := \langle u_1, \ldots, u_p\rangle$, $W := \langle v_{s+1}, \ldots, v_m\rangle$. Pak z věty 20 o dimenzi součtu a průniku dostáváme

$$\dim U \cap W = \dim U + \dim W - \dim(U + W) \geq p + (m - s) - m \geq 1$$

Tedy existuje nenulový $u = \sum_{i=1}^{p} x_i u_i = \sum_{i=s+1}^{m} y_i v_i \in U \cap W$ a

$$Q_g(u) = ([u]^B)^T [g]_B [u]^B = \sum_{i=1}^{p} a_i x_i^2 > 0$$
$$Q_g(u) = ([u]^C)^T [g]_C [u]^C = \sum_{i=s+1}^{s+t} d_{i-s} y_i^2 \leq 0$$

Nerovnosti vyplývají ze znamének čísel a_i, d_j a nenulovosti vektoru $(x_1, \ldots, x_p)$. Jsou ale ve sporu, musí tedy být $p = s$ a $q = t$. □

Sylvesterův zákon umožňuje definovat *signaturu* $\operatorname{sign} Q_g$ kvadratické formy Q_g jakožto trojici nezáporných čísel (p, q, n), která udávají počet kladných, záporných a nulových elementů na diagonále matice kvadratické formy vzhledem k nějaké polární bázi. Víme už, že forma je pozitivně definitní, má-li její signatura tvar $(p, 0, 0)$, kde $p > 0$. Další podobná označení jsou

- $\operatorname{sign} Q_g = (p, 0, n)$, kde $p, n > 0$: Q_g je *pozitivně semidefinitní*
- $\operatorname{sign} Q_g = (0, q, 0)$, kde $q > 0$: Q_g je *negativně definitní*
- $\operatorname{sign} Q_g = (0, q, n)$, kde $q, n > 0$: Q_g je *negativně semidefinitní*
- $\operatorname{sign} Q_g = (p, q, 0)$, kde $p, q > 0$: Q_g je *regulární indefinitní*
- $\operatorname{sign} Q_g = (p, q, n)$, kde $p, q, n > 0$: Q_g je *singulární indefinitní*

Tytéž pojmy se vztahují i na symetrické bilineární formy a symetrické matice. Celý dosavadní výklad včetně Sylvesterova zákona bychom mohli zobecnit i na případ hermitovské seskvilineární formy, pro který už jsme formulovali a dokazovali větu 22. Stejně tak můžeme mluvit o signatuře hermitovské matice a samosdruženého operátoru.

3. Sylvesterovo kritérium

Nechť $G \in \mathbb{R}^{n\times n}$, označme $G_k \in \mathbb{R}^{k\times k}$ její podmatici vzniklou vyškrtnutím posledních $n-k$ řádků a sloupců.

Věta 36 (Jacobi-Sylvesterova). *Nechť g je symetrická bilineární forma na V, $B = (u_i)_1^m$ báze V, $G = [g]_B$ splňuje podmínku $\forall k \in \{1,\dots,m\} : \det G_k \neq 0$. Pak má g polární bázi $C = (v_i)_1^m$ takovou, že $v_k = \sum_{j=1}^k r_{jk}u_j$ pro nějakou regulární horní trojúhelníkovou matici $R \in \mathbb{R}^{m\times m}$ a*

$$[g]_C = \operatorname{diag}\left(\frac{1}{\det G_1}, \frac{\det G_1}{\det G_2}, \frac{\det G_2}{\det G_3} \dots, \frac{\det G_{m-1}}{\det G_m}\right)$$

Tedy signatura g je $(p,q,0)$, kde q je počet znaménkových změn v posloupnosti $(1, \det G_1, \dots, \det G_m)$.

Věta bývá označována také jako *Sylvesterovo kritérium* a je to další způsob určení signatury kvadratické formy, která ale musí být regulární. Důkaz trochu připomíná Gramovu-Schmidtovu ortogonalizaci:

Důkaz. Zvolme $v_1 := \frac{1}{g_{11}}u_1$, pak $g(v_1,v_1) = \frac{1}{g_{11}^2}g(u_1,u_1) = \frac{1}{g_{11}}$, protože $g_{11} = g(u_1,u_1)$. Předpokládejme, že jsme zkonstruovali vektory $(v_1,\dots,v_{k-1})$, tedy i prvních $k-1$ sloupců matice R. Zkonstruujme její k-tý sloupec tak, že nalezneme vektor $v_k = \sum_{j=1}^k r_{jk}u_j$ tak, aby pro $i<k$ platilo

$$g(u_i, v_k) \equiv g\left(u_i, \sum_{j=1}^k r_{jk}u_j\right) \equiv \sum_{j=1}^k g_{ij}r_{jk} = 0$$

Přidáme-li podmínku $g(u_k,v_k) = 1$, máme pro vektor $(r_{1k},\dots,r_{kk})^T$ soustavu rovnic s maticí G_k a pravou stranou $\mathbf{e}_k$. Protože $\det G_k \neq 0$, řešení soustavy existuje a je jednoznačné. Z prvních $k-1$ podmínek této soustavy platí $g(u_j,v_k) = 0$ pro $j<k$, a z konstrukce předchozích $k-1$ sloupců R máme $v_j \in \langle u_1,\dots,u_j\rangle$, tedy $g(v_j,v_k) = 0$ pro $j<k$. Když takto zkonstruujeme celou bázi $C := (v_i)_1^m$, plyne odtud, že $[g]_C$ je horní trojúhelníková matice, ale protože je zároveň symetrická, musí být diagonální. Tedy C je polární báze pro g. Diagonální elementy $[g]_C$ získáme Cramerovým pravidlem:

$$g(v_k,v_k) = g\left(\sum_{j=1}^k r_{jk}u_j, v_k\right) = g(r_{kk}u_k, v_k) = r_{kk} = \frac{\det\begin{pmatrix} G_{k-1} & \mathbf{o} \\ * & 1\end{pmatrix}}{\det G_k} = \frac{\det G_{k-1}}{\det G_k}$$

V druhém kroku jsme využili bilinearitu g a vlastnost $g(u_i,v_k) = 0$ pro $i<k$, ve třetím podmínku $g(u_k,v_k) = 1$, ve čtvrtém Cramerovo pravidlo a v pátém Laplaceův rozvoj. □

Příklad 29. Určeme signaturu kvadratické formy na $\mathbb{R}^3$

$$Q_g(\mathbf{x}) = \mathbf{x}^T\begin{pmatrix}1 & 0 & 1\\ 0 & -3 & 3\\ 1 & 3 & 1\end{pmatrix}\mathbf{x} = x_1^2 + 2x_1x_3 - 3x_2^2 + 6x_2x_3 + x_3^2$$

Spočteme $\det G_1 = 1$, $\det G_2 = -3$, $\det G_3 = -9$. V posloupnosti $(1,1,-3,-9)$ je jedna znaménková změna, tedy $\operatorname{sign} Q_g = (2,1,0)$, forma je indefinitní. Znamená to, že vzhledem k libovolné polární bázi bude Q_g vyjádřena diagonální maticí, na jejíž diagonále budou dvě kladná a jedno záporné číslo. Protože navíc vždy lze symetrickou úpravou udělat z kladné diagonální hodnoty $+1$ a ze záporné -1, znamená to, že lze v $\mathbb{R}^3$ zavést nové (čárkované) souřadnice, pomocí nichž je $Q_g(\mathbf{x})$ vyjádřena jako

$$\mathbf{x}'^T \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \mathbf{x}' = (x_1')^2 + (x_2')^2 - (x_3')^2$$

Definice 50. *Nulová množina* kvadratické formy Q_g je množina

$$N_g := \{v \in V | Q_g(v) = 0\},$$

obsahující všechny vektory, které se zobrazí pomocí Q_g na nulu.

Příklad 30. Pro Q_g na $\mathbb{R}^2$ s předpisem

$$Q_g(\mathbf{x}) = \mathbf{x}^T \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \mathbf{x} = x_1^2 - x_2^2$$

je $N_g = \langle(1,1)\rangle \cup \langle(1,-1)\rangle$, tedy sjednocení dvou různoběžných přímek. Taková množina není podprostorem $\mathbb{R}^2$. Podobná situace nastává u všech indefinitních kvadratických forem. Jedna taková, tzv. prostoročasový interval v $\mathbb{R}^4$ se signaturou $(3,1,0)$, je důležitá pro speciální teorii relativity, a její nulové množině se říká světelný (nebo též nulový) kužel. Jeho trojrozměrnou verzí je nulová množina kvadratické formy z předchozího příkladu, její tvar lze nejlépe nahlédnout v čárkovaných souřadnicích, v nichž je to rotační kuželová plocha s osou x_3'.

Matice $A \in \mathbb{R}^{n \times n}$, která zachovává obecnou kvadratickou formu Q_g s maticí $[g]_K^K = G$, musí splňovat $\forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n$

$$g(A\mathbf{x}, A\mathbf{y}) = \mathbf{x}^T A^T G A \mathbf{y} = \mathbf{x}^T G \mathbf{y} = g(\mathbf{x}, \mathbf{y}),$$

neboli $A^T G A = G$. Pro Q_g prostoročasový interval a matici G v základním tvaru $G = \mathrm{diag}(1,1,1,-1)$ se množina všech takových matic A nazývá Lorentzova grupa, nebo též grupa Lorentzových transformací. Je v ní obsažena grupa všech rotací v prostoru, ale také středová souměrnost, symetrie podle časové proměnné a především tzv. vlastní Lorentzovy transformace (*boosty*), které odpovídají přechodu do inerciální soustavy souřadnic pohybující se vůči původní soustavě nějakou rychlostí.

KAPITOLA 18

Kvadriky

1. Afinní prostor

V zimním semestru jsme zavedli pojem afinního podprostoru vektorového prostoru V jako množinu $a + W$, kde $a \in V$ a $W \leq V$. Podprostor W se nazývá *zaměření* afinního podprostoru. Podmnožina $a + W$ ve V není obecně uzavřená na sčítání vektorů a násobení vektoru skalárem, není tedy podprostorem V. S afinními podprostory jsme se zatím setkali především jako s množinami řešení nehomogenních soustav lineárních rovnic, které jsme si zvykli zapisovat jako $\mathbf{x}_P + \operatorname{Ker} A$, součet partikulárního řešení a jádra matice soustavy.

Definice 51. Afinní podprostor $A_n := \{\tilde{\mathbf{x}} \in \mathbb{F}^{n+1} | \tilde{\mathbf{x}} = (1, x_1, \ldots, x_n)^T\}$ vektorového prostoru $\mathbb{F}^{n+1}$ se zaměřením $V_n := \{\tilde{\mathbf{x}} \in \mathbb{F}^{n+1} | \tilde{\mathbf{x}} = (0, x_1, \ldots, x_n)^T\}$ se nazývá *aritmetický afinní prostor.*

Složky vektoru $\tilde{\mathbf{x}}$ číslujeme od nuly, takže A_n je také možno popsat rovnicí $\tilde{x}_0 = 1$, nebo blokovým zápisem

$$A_n = \left\{ \tilde{\mathbf{x}} = \begin{pmatrix} 1 \\ \mathbf{x} \end{pmatrix} \middle| \mathbf{x} \in \mathbb{F}^n \right\}$$

Aritmetický afinní prostor je vhodná struktura, uvnitř níž se dají studovat přímky, roviny, nadroviny a další lineární útvary neprocházející počátkem, ale také kuželosečky, kvadratické plochy a další. Než se k tomu dostaneme, zavedeme na A_n analogie některých konstrukcí na vektorových prostorech.

Jak A_n, tak V_n lze identifikovat s množinou $\mathbb{F}^n$ pomocí přirozených bijekcí

$$\begin{aligned} \iota_A : \mathbb{F}^n &\to A_n \\ \mathbf{a} &\mapsto \begin{pmatrix} 1 \\ \mathbf{a} \end{pmatrix} \end{aligned} \qquad \qquad \begin{aligned} \iota_V : \mathbb{F}^n &\to V_n \\ \mathbf{x} &\mapsto \begin{pmatrix} 0 \\ \mathbf{x} \end{pmatrix} \end{aligned}$$

Prvky prvního typu označujeme jako *body* a druhého jako *vektory.* Pro lepší odlišení bodů a vektorů budeme u bodů užívat hranaté závorky, netučný font a obvykle i písmena ze začátku abecedy:

$$a \equiv \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} = \iota_A^{-1} \begin{pmatrix} 1 \\ a_1 \\ \vdots \\ a_n \end{pmatrix}, \quad \mathbf{x} \equiv \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \iota_V^{-1} \begin{pmatrix} 0 \\ x_1 \\ \vdots \\ x_n \end{pmatrix}$$

Rozdíl dvou prvků A_n je prvkem V_n a prostřednictvím bijekcí ι_A, ι_V lze tuto operaci přenést i na odpovídající prvky $\mathbb{F}^n$:

$$a - b := \iota_V^{-1}(\iota_A(a) - \iota_A(b)), \text{ tj. } \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} - \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} = \begin{pmatrix} a_1 - b_1 \\ \vdots \\ a_n - b_n \end{pmatrix}$$

Podobně součet bodu a vektoru je

$$a + \mathbf{x} := \iota_A^{-1}(\iota_A(a) + \iota_V(\mathbf{x})), \quad \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} + \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{bmatrix} a_1 + x_1 \\ \vdots \\ a_n + x_n \end{bmatrix}$$

Na vektorech jsou (narozdíl od bodů) samozřejmě definovány i součet a násobek skalárem.

Afinní přímka spojující body a, b se dá zapsat jako množina bodů

$$\{a + \lambda(b - a) | \lambda \in \mathbb{F}\} \subset \mathbb{F}^n$$

Výraz $a + \lambda(b - a)$ lze formálně přepsat i jako $(1 - \lambda)a + \lambda b$. Ačkoli na bodech součet ani násobek skalárem nedefinujeme, výraz $\mu\, \iota_A(a) + \lambda\, \iota_A(b)$ smysl dává a pro $\mu + \lambda = 1$ je ι_A-obrazem nějakého bodu.

Definice 52. Nechť $a_1, \ldots, a_k \in \mathbb{F}^n$ jsou body a $\lambda_1, \ldots, \lambda_k \in \mathbb{F}$ skaláry, které splňují $\sum_1^k \lambda_i = 1$. Pak výraz

$$\lambda_1 a_1 + \lambda_2 a_2 + \ldots + \lambda_k a_k,$$

označujeme jako *afinní lineární kombinaci* daných bodů.

Z příkladu s přímkou je vidět, že množina všech afinních lineárních kombinací bodů $a_1, \ldots, a_k$ je něco jiného než lineární obal v $\mathbb{F}^n$. Dá se ale vyjádřit pomocí lineárního obalu odpovídajících vektorů $\mathbb{F}^{n+1}$ jako

$$\iota_A^{-1}(\, \langle \iota_A(a_1), \ldots, \iota_A(a_k) \rangle \cap A_n \,)$$

V případě afinní přímky procházející body a, b je to průnik dvourozměrného podprostorou $\langle \iota_A(a), \iota_A(b) \rangle$ v $\mathbb{F}^{n+1}$ s nadrovinou A_n, identifikovaný pomocí ι_A^{-1} s afinní přímkou v $\mathbb{F}^n$. Podobně průnik nadroviny $\{\tilde{\mathbf{x}} \in \mathbb{F}^{n+1} | r_0 x_0 + \ldots + r_n x_n = 0\}$ a A_n je ztotožněn s afinní nadrovinou $\{x \in \mathbb{F}^n | r_0 + r_1 x_1 + \ldots + r_n x_n = 0\}$.

Ztotožnění afinních přímek, rovin, nadrovin a dalších lineárních objektů v $\mathbb{F}^n$ s podprostory vektorového prostoru $\mathbb{F}^{n+1}$ by nám spolu s větou o dimenzi spojení a průniku umožnilo zobecnit standardní teorii o vzájemné poloze dvou přímek či přímky a roviny v $\mathbb{R}^3$ do libovolné dimenze. My se analytické geometrii lineárních objektů věnovat nebudeme, přejdeme rovnou k objektům kvadratickým. Afinní prostor využijeme především jako vhodný rámec ke studiu kvadrik, tj. množin ztotožněných pomocí ι_A s průnikem nadroviny A_n s nulovou množinou nějaké kvadratické formy na $\mathbb{F}^{n+1}$.

Ztotožnění mezi prvky $\mathbb{F}^n$ a A_n budeme používat natolik často, že psát všude bijekci ι_A nebo její inverzi by bylo zbytečně nepřehledné, podobně u $\mathbb{F}^n$ a V_n. Obvykle bude z kontextu, ze značení a z použitých operací jasné, s jakým typem objektu máme co do činění. Například v následující definici by bylo přesnější psát $\iota_A(a), \iota_A(b) \in A_n$ a místo $b - a$ vždy $\iota_V(b - a)$:

Definice 53. Nechť $a, b \in A_n$. Pak *vzdálenost* $\rho(a, b)$ těchto bodů definujeme jako $\|b - a\|$, tj. velikost vektoru $b - a \in V_n$ v normě indukované standardním skalárním součinem.

Prostor A_n s takto definovanou funkcí ρ se nazývá *eukleidovský prostor E_n*. Snadno se ověří, že splňuje axiomy *metrického prostoru*:

Definice 54. Metrický prostor je dvojice (X, ρ), kde X je množina a $\rho : X \times X \to \mathbb{R}_0^+$ je funkce splňující

(1) $\rho(x, y) = 0 \Leftrightarrow x = y$
(2) $\forall x, y \in X : \rho(x, y) = \rho(y, x)$
(3) $\forall x, y, z \in X : \rho(x, y) + \rho(y, z) \geq \rho(x, z)$

Tvrzení 59. *Nechť $\tilde{A} \in \mathbb{F}^{n+1 \times n+1}$. Pak lineární zobrazení $F_{\tilde{A}} : \mathbb{F}^{n+1} \to \mathbb{F}^{n+1}$ zachovává A_n a V_n, právě když*

$$\tilde{A} = \begin{pmatrix} 1 & \mathbf{o}^T \\ \mathbf{r} & A \end{pmatrix},$$

kde $A \in \mathbb{F}^{n \times n}$, $\mathbf{r} \in \mathbb{F}^n$ mohou být libovolné. Navíc $F_{\tilde{A}}$ zachovává vzdálenost mezi body A_n, právě když je A unitární matice.

Důkaz. Má-li být nultá složka vektoru

$$\tilde{A} \begin{pmatrix} 1 \\ \mathbf{a} \end{pmatrix} = \begin{pmatrix} \lambda & \mathbf{s}^T \\ \mathbf{r} & A \end{pmatrix} \begin{pmatrix} 1 \\ \mathbf{a} \end{pmatrix} = \begin{pmatrix} \lambda + \mathbf{s}^T \mathbf{a} \\ A\mathbf{a} + \mathbf{r} \end{pmatrix}$$

rovna jedné nezávisle na hodnotě $\mathbf{a}$, musí být $\mathbf{s} = 0$ a $\lambda = 1$. Pro matici $\tilde{A}$ v tomto tvaru pak

$$\tilde{A} \begin{pmatrix} 0 \\ \mathbf{x} \end{pmatrix} = \begin{pmatrix} 1 & \mathbf{o}^T \\ \mathbf{r} & A \end{pmatrix} \begin{pmatrix} 0 \\ \mathbf{x} \end{pmatrix} = \begin{pmatrix} 0 \\ A\mathbf{x} \end{pmatrix}$$

Z definice 53 plyne, že $F_{\tilde{A}}$ zachová vzdálenost, právě když $\forall \mathbf{x} \in \mathbb{F}^n$ bude $\|\mathbf{x}\| = \|A\mathbf{x}\|$. Z tvrzení 54 plyne, že to nastává právě pro A unitární. □

Zavedeme-li součin matice z $\mathbb{F}^{n \times n}$ a bodu z $\mathbb{F}^n$ standardně, můžeme zúžení $F_{\tilde{A}}$ na A_n psát ve tvaru $a \mapsto Aa + \mathbf{r}$. Takové zobrazení na aritmetickém afinním prostoru se nazývá *afinní zobrazení*. Je-li A unitární, nazývá se takové zobrazení *izometrie*.

Definice 55. *Soustava souřadnic* v A_n je dvojice $S = (p, B)$, kde $p \in A_n$ a $B = (\mathbf{v}_i)_1^n$ je báze V_n. Bod p je její *počátek* a soustava souřadnic S se nazývá *ortogonální/ortonormální*, právě když je ortogonální/ortonormální báze B. *Souřadnicemi* bodu $a \in A_n$ vzhledem k soustavě souřadnic S označíme vektor $[a]_S := [a - p]^B \in \mathbb{F}^n$.

Zvolme $S' = (p', B')$, jakožto další soustavu souřadnic, kde $B' = (\mathbf{v}'_i)_1^n$, a pro $a \in A_n$ označme $\mathbf{x} := [a]_S$, $\mathbf{x}' := [a]_{S'} \in \mathbb{F}^n$. Pak

$$a = p + \sum_{j=1}^{n} x_j \mathbf{v}_j = p' + \sum_{i=1}^{n} x'_i \mathbf{v}'_i.$$

Označme dále $U := [\mathrm{Id}]^B_{B'}$ a dosaďme z definice matice přechodu

$$\sum_{j=1}^{n} x_j \mathbf{v}_j = (p' - p) + \sum_{i=1}^{n} x'_i \sum_{j=1}^{n} u_{ji} \mathbf{v}_j.$$

Vektor $p' - p$ lze zapsat jako $\sum_{j=1}^{n} r_j \mathbf{v}_j$, přičemž $\mathbf{r} := [p' - p]^B$. Pak lze poslední rovnost upravit na

$$\sum_{j=1}^{n} x_j \mathbf{v}_j = \sum_{j=1}^{n} r_j \mathbf{v}_j + \sum_{j=1}^{n} \sum_{i=1}^{n} x'_i u_{ji} \mathbf{v}_j.$$

a z jednoznačnosti reprezentace vektoru vůči bázi získat vztah mezi j-tými složkami vektorů $\mathbf{x}$, $\mathbf{x}'$ souřadnic bodu a vzhledem k soustavám S a S':

$$x_j = r_j + \sum_{i=1}^{n} u_{ji} x'_i$$

Ten se dá přepsat maticově jako

$$\begin{pmatrix} 1 \\ \mathbf{x} \end{pmatrix} = \begin{pmatrix} 1 & \mathbf{o}^T \\ \mathbf{r} & U \end{pmatrix} \begin{pmatrix} 1 \\ \mathbf{x}' \end{pmatrix}$$

Podobnost mezi vztahem pro změnu souřadnic a předpisem pro afinní zobrazení je jasně patrná a analogická podobnosti mezi maticí přechodu a maticí lineárního zobrazení ze zimního semestru. Předpokládáme-li, že soustavy souřadnic S a S' jsou ortogonální, je U unitární matice, podobně jako u izometrií.

2. Kvadriky

Kvadrika je podmnožina reálného aritmetického afinního prostoru A_n popsaná (kvadratickou) rovnicí

$$\tilde{\mathbf{x}}^T \tilde{A} \tilde{\mathbf{x}} \equiv \begin{pmatrix} 1 & \mathbf{x}^T \end{pmatrix} \begin{pmatrix} c & \mathbf{b}^T \\ \mathbf{b} & A \end{pmatrix} \begin{pmatrix} 1 \\ \mathbf{x} \end{pmatrix} \equiv \mathbf{x}^T A \mathbf{x} + 2\mathbf{b}^T \mathbf{x} + c = 0,$$

kde $c \in \mathbb{R}$, $\mathbf{b} \in \mathbb{R}^n$ a $A \in \mathbb{R}^{n \times n}$ jsou bloky matice označené $\tilde{A}$. Z předchozí kapitoly víme, že v kvadratické formě $\mathbf{x} \mapsto \mathbf{x}^T A \mathbf{x}$ na $\mathbb{R}^n$ můžeme bez újmy na obecnosti předpokládat, že A je symetrická, a budeme tak činit. Zobrazení $Q_g : \tilde{\mathbf{x}} \mapsto \tilde{\mathbf{x}}^T \tilde{A} \tilde{\mathbf{x}}$ na $\mathbb{R}^{n+1}$ je rovněž kvadratická forma určená symetrickou maticí $\tilde{A} \equiv [g]_K$ a kvadrika je průnikem její nulové množiny N_g s aritmetickým afinním prostorem A_n. Známými příklady kvadrik na A_2 jsou elipsa, hyperbola a parabola. Například rovnici elipsy se středem v 0 a délkou poloos a, b lze přepsat do tvaru

$$\begin{pmatrix} 1 & x_1 & x_2 \end{pmatrix} \begin{pmatrix} -1 & 0 & 0 \\ 0 & a^{-2} & 0 \\ 0 & 0 & b^{-2} \end{pmatrix} \begin{pmatrix} 1 \\ x_1 \\ x_2 \end{pmatrix} = \frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} - 1 = 0$$

Kvadrika se nazývá *regulární*, pokud je regulární matice $\tilde{A}$, a *středová*, pokud je regulární matice A. Vlastnosti kvadriky budeme studovat pomocí přechodu do jiné soustavy souřadnic, vzhledem k níž je její rovnice v *kanonickém tvaru*, jako například výše zmíněná elipsa. Všechny kanonické tvary kvadrik se vyznačují diagonální maticí A, což znamená, že se v rovnici kvadriky nevyskytují smíšené členy $x_i x_j$.

Změna souřadnic změní matici formy Q_g na matici

$$\begin{pmatrix} 1 & \mathbf{r}^T \\ \mathbf{o} & U^T \end{pmatrix} \begin{pmatrix} c & \mathbf{b}^T \\ \mathbf{b} & A \end{pmatrix} \begin{pmatrix} 1 & \mathbf{o}^T \\ \mathbf{r} & U \end{pmatrix} = \begin{pmatrix} c' & \mathbf{b}'^T \\ \mathbf{b}' & U^T A U \end{pmatrix},$$

kde $c' := c + 2\mathbf{b}^T \mathbf{r} + \mathbf{r}^T A \mathbf{r}$ a $\mathbf{b}' := U^T(\mathbf{b} + A\mathbf{r})$. Protože A je symetrická matice, existuje matice U taková, že $U^T A U =: D \equiv \operatorname{diag}(d_1, \ldots, d_n)$. Je-li kvadrika středová, pak existuje A^{-1} a volbou $\mathbf{r} = -A^{-1}\mathbf{b}$ zajistíme, že $\mathbf{b}' = \mathbf{o}$. Zároveň

$$c' = c - 2\mathbf{b}^T A^{-1} \mathbf{b} + \mathbf{b}^T (A^{-1})^T A A^{-1} \mathbf{b} = c - \mathbf{b}^T A^{-1} \mathbf{b}$$

Protože $\tilde{A}$ je regulární a regularita se zachovává při násobení regulárními maticemi, musí být c' nenulové. V nových, čárkovaných souřadnicích, které s původními spojuje vztah

$$\begin{pmatrix} 1 \\ \mathbf{x} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \mathbf{r} & U \end{pmatrix} \begin{pmatrix} 1 \\ \mathbf{x}' \end{pmatrix},$$

má rovnice kvadriky tvar

$$c' + \sum_{i=1}^{n} d_i x_i'^2 = 0$$

Absence lineárních členů znamená, že je množina bodů $\mathbf{x}' \in \mathbb{R}^n$ splňujících rovnici souměrná podle všech nadrovin kolmých na souřadné osy v $\mathbb{R}^n$, dává tedy smysl označit bod se souřadnicemi $\mathbf{x}' = \mathbf{o}$ jako její *střed*. V původní bázi jsou jeho souřadnice rovny $\mathbf{x} = \mathbf{r} + U\mathbf{x}' = -A^{-1}\mathbf{b}$ a dají se tedy snadno zjistit už z původní matice, bez nutnosti hledat U a $\mathbf{r}$.

Na $\mathbb{R}^2$ vidíme, že různá znaménka čísel d_1 a d_2 dávají v čárkovaných souřadnicích rovnici hyperboly a stejná znaménka buď rovnici elipsy, nebo prázdnou množinu, v závislosti na znaménku c'. Znaménka diagonálních elementů matice D

jsou dána signaturou A a znaménko c' je stejné jako znaménko podílu mezi determinanty matice D a matice $\mathrm{diag}(c', d_1, \dots, d_n)$. Tyto determinanty mají ale stejné znaménko jako determinanty matice A a $\tilde{A}$. Platí tedy

VĚTA 37. *Je-li* $\mathbf{x}^T A\mathbf{x} + 2\mathbf{b}^T\mathbf{x} + c = 0$ *regulární středová kvadrika taková, že signatura matice* A *je* $(p, q, 0)$ *a* $C := \frac{\det \tilde{A}}{\det A}$, *pak pro*

- $q = 0 \wedge C < 0$ *nebo* $p = 0 \wedge C > 0$ *je to elipsoid,*
- $q = 0 \wedge C > 0$ *nebo* $p = 0 \wedge C < 0$ *je to prázdná množina,*
- $p \neq 0 \wedge q \neq 0$ *je to hyperboloid.*

PŘÍKLAD 31. Uvažujme kvadriku na $\mathbb{R}^2$ s rovnicí $3x^2 - 10xy + 3y^2 + 14x - 2y + 3 = 0$. Přepišme levou stranu maticově jako

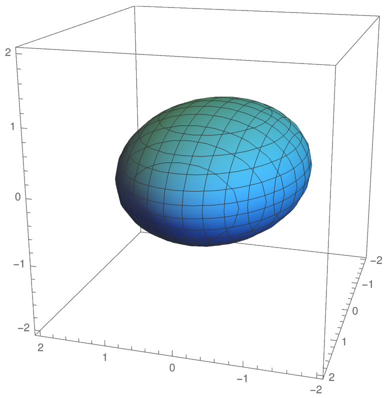
$$\begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} 3 & -5 \\ -5 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + 2\begin{pmatrix} 7 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + 3 \equiv \begin{pmatrix} 1 & x & y \end{pmatrix} \begin{pmatrix} 3 & 7 & -1 \\ 7 & 3 & -5 \\ -1 & -5 & 3 \end{pmatrix} \begin{pmatrix} 1 \\ x \\ y \end{pmatrix}$$

Protože $\operatorname{sign} A = (1, 1, 0)$ a velká 3×3 matice $\tilde{A}$ je regulární, jedná se o hyperbolu.

Věta umožňuje určit typ středové kvadriky z její rovnice v původních souřadnicích. Elipsoid je zobecněním elipsy, je to tedy kvadrika s (kanonickou) rovnicí

$$\sum_{i=1}^{n} \frac{x_i'^2}{a_i^2} = 1$$

Délky jejích poloos jsou rovny $a_i = \sqrt{-\frac{c'}{d_i}}$. Zvolíme-li matici U ortogonální, pak změna souřadnic zachovává vzdálenost a stejné délky poloos má elipsoid i v původních, nečárkovaných souřadnicích. Čísla d_i na diagonále matice D jsou v takovém případě rovna vlastním číslům matice A, tedy i ty lze zjistit již z původní rovnice kvadriky. Směry poloos jsou v čárkovaných souřadnicích dány prvky kanonické báze, v původních souřadnicích to tedy musí být vlastní vektory matice A.



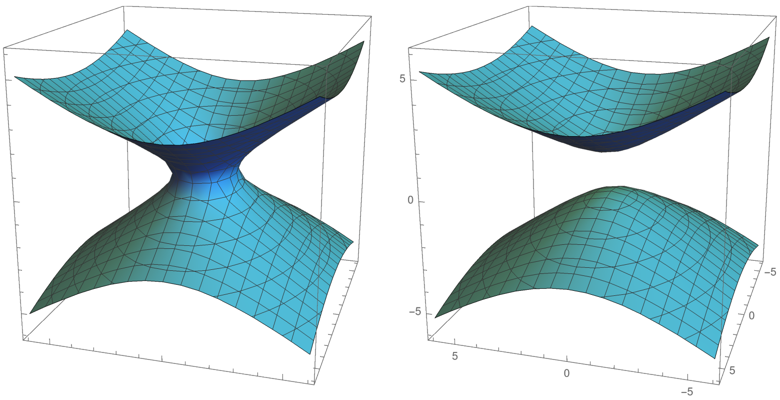
OBRÁZEK 1. *
Elipsoid: $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$

Tuto diskusi je možné vztáhnout i na případ hyperboloidu, tedy kvadriky s kanonickou rovnicí

$$\sum_{i=1}^{n} \pm \frac{x_i'^2}{a_i^2} = 1,$$

kde všechna znaménka nalevo nejsou stejná. Počet kladných a záporných, tedy signatura matice A umožňuje zavést jemnější dělení hyperboloidů, v $\mathbb{R}^3$ jsou dva

- jednodílný a dvojdílný. Odpovídají rotačnímu tělesu, jehož plášť je hyperbola, vzniklému rotací okolo dvou hlavních os této hyperboly.



Obrázek 2. *
Jednodílný hyperboloid: $\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1$ Dvojdílný hyperboloid: $\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = -1$

Rozeberme ještě případ nestředové regulární kvadriky. Stejně jako u středové můžeme najít ortogonální matici U takovou, že $U^T AU$ je diagonální matice D, jejíž alespoň jeden diagonální element je 0. Nechť je to ten poslední, tedy $D = \operatorname{diag}(d_1, \ldots, d_{n-1}, 0)$. Poslední sloupec matice

$$\begin{pmatrix} c' & \mathbf{b}'^T \\ \mathbf{b}' & D \end{pmatrix}$$

tedy musí být násobkem prvního vektoru kanonické báze $\mathbf{e}_1 \in \mathbb{R}^{n+1}$. Musí to být násobek nenulový a žádný jiný sloupec už násobkem vektoru $\mathbf{e}_1$ být nemůže, bylo by to v rozporu z regularitou kvadriky. Tedy čísla $d_1, \ldots, d_{n-1}$ už musí být všechna nenulová a $\operatorname{rank} A = n - 1$. Chceme opět zvolit $\mathbf{r}$, aby $\mathbf{b}' = U^T(\mathbf{b} + A\mathbf{r})$ bylo co nejjednodušší, ale se singulární maticí A už nebude možné, aby to byla nula. Označme $\mathbf{b}'' := U^T\mathbf{b}$, pak

$$\mathbf{b}' = \mathbf{b}'' + U^T AUU^T\mathbf{r} = \mathbf{b}'' + D\mathbf{r}'',$$

kde $\mathbf{r}'' := U^T\mathbf{r}$. Pro $i \in \{1, \ldots, n-1\}$ položme $r''_i := -\frac{1}{d_i} b''_i$, pak má matice tvar

$$\begin{pmatrix} c' & \mathbf{o}^T & b''_n \\ \mathbf{o} & D'' & \mathbf{o} \\ b''_n & \mathbf{o}^T & 0 \end{pmatrix},$$

kde $D'' = \operatorname{diag}(d_1, \ldots, d_{n-1})$ a $c' := c + 2\mathbf{b}^T\mathbf{r} + \mathbf{r}^T A\mathbf{r}$. Po dosazení

$$\begin{aligned} c' &= c + 2\mathbf{b}''^T U^T U\mathbf{r}'' + \mathbf{r}''^T U^T AU\mathbf{r}'' = c + 2\mathbf{b}''^T\mathbf{r}'' + \mathbf{r}''^T D\mathbf{r}'' \\ &= c - 2\sum_{i=1}^{n-1} \frac{1}{d_i}(b''_i)^2 + b''_n r''_n + \sum_{i=1}^{n-1} \frac{1}{d_i}(b''_i)^2 = c - \sum_{i=1}^{n-1} \frac{1}{d_i}(b''_i)^2 + b''_n r''_n \end{aligned}$$

Protože $b''_n \neq 0$, lze zvolit r''_n tak, aby $c' = 0$. Tím jsme zvolili vektor $\mathbf{r}''$ a tedy i $\mathbf{r} = U\mathbf{r}''$ tak, aby v příslušné soustavě souřadnic měla kvadrika rovnici

$$\sum_{i=1}^{n-1} d_i x_i'^2 + 2b''_n x'_n = 0.$$

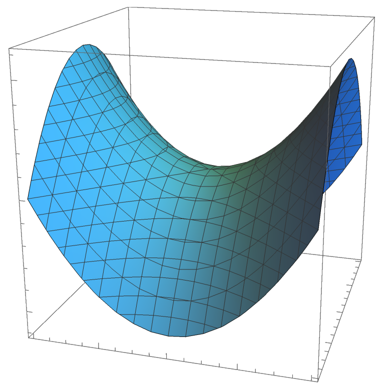
Pro $n = 2$ je to parabola, jejíž rovnici je možné přepsat jako

$$x'_2 = -\frac{d_1}{2b''_n} x_1'^2,$$

obecně se jedná se o paraboloid. Parametry v rovnici je možné určit z původní matice $\tilde{A}$, protože $d_1, \dots, d_{n-1}$ jsou vlastní čísla její podmatice A a

$$\det \tilde{A} = -(b''_n)^2 \prod_{i=1}^{n-1} d_i$$

Na signatuře matice A opět závisí jemnější dělení typů paraboloidů, v $\mathbb{R}$ rozlišujeme eliptický paraboloid pro $\operatorname{sign} A = (2,0,1)$ nebo $(0,2,1)$ a hyperbolický paraboloid pro $\operatorname{sign} A = (1,1,1)$. Z obrázku je vidět, že vodorovné řezy grafu hyperbolického paraboloidu jsou skutečně hyperboly, u eliptického to budou elipsy, případně bod nebo prázdná množina.



Obrázek 3. *
Hyperbolický paraboloid: $\frac{x^2}{a^2} - \frac{y^2}{b^2} - z = 0$

Směr osy paraboloidu je definován jako lineární obal vektoru, jehož čárkované souřadnice jsou $\mathbf{e}_n$, tedy jako vlastní vektor A s vlastním číslem 0, neboli $\operatorname{Ker} A$. *Vrcholem paraboloidu* je bod určený v čárkovaných souřadnicích vektorem $\mathbf{x}' = \mathbf{o}$.

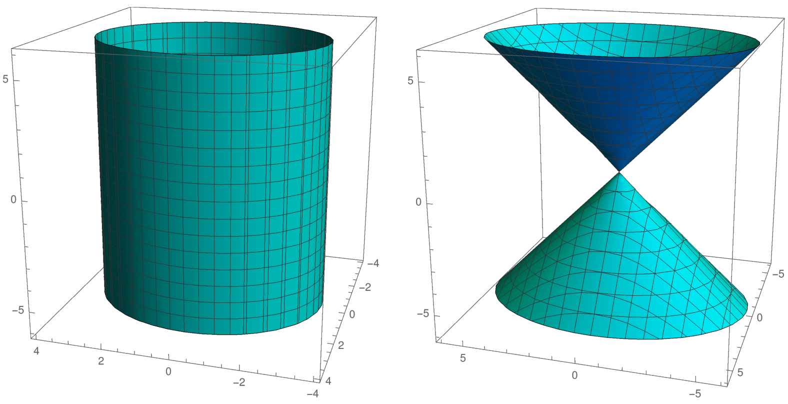
Poznámka 6. V této kapitole jsme klasifikovali všechny regulární kvadriky do tříd ekvivalence. Ty, které mají vzhledem k nějaké ortonormální soustavě souřadnic stejnou rovnici, jsou na sebe převoditelné odpovídající izometrií. Říkáme, že jsou stejného *metrického typu* a seznam těchto typů nazýváme *metrickou klasifikací* kvadrik. Metrických typů kvadrik je nekonečně mnoho, protože například dvě elipsy, které se liší délkami poloos, jsou různého metrického typu.

Afinní klasifikace kvadrik připouští i neortonormální soustavy souřadnic, jinými slovy kvadriky jsou stejného *afinního typu*, pokud se na sebe dají převést afinní transformací. Afinních typů regulárních kvadrik je konečně mnoho, v dimenzi 2 jsou to elipsa, hyperbola, parabola a prázdná množina, v dimenzi 3 je to elipsoid, jednodílný hyperboloid, dvojdílný hyperboloid, hyperbolický paraboloid, eliptický paraboloid a prázdná množina. Afinní typ kvadriky je možné určit jen ze signatury matic A a $\tilde{A}$.

Klasifikací singulárních kvadrik se zabývat nebudeme, podobně jako u regulárních stačí k určení afinního typu nalezení signatur. V $\mathbb{R}^2$ jsou typu afinních singulárních kvadrik

- bod: $x_1^2 + x_2^2 = 0$
- dvě různoběžky: $x_1^2 - x_2^2 = 0$
- dvě rovnoběžky: $x_1^2 - 1 = 0$
- přímka: $x_1^2 = 0$
- rovina: $0 = 0$

Příklady singulárních kvadrik v $\mathbb{R}^3$ jsou například tyto:



Obrázek 4. *

Eliptická válcová plocha: $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ Kuželová plocha: $\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 0$

KAPITOLA 19

Jordanův tvar

Nechť V je vektorový prostor konečné dimenze nad $\mathbb{C}$, $f \in \operatorname{End}(V)$ a $\lambda \in \mathbb{C}$ je vlastní číslo endomorfismu f. Zaveďme označení $f_\lambda := f - \lambda \operatorname{Id}$ pro endomorfismus posunutý oproti f o $-\lambda$-násobek identity. Posloupnost nenulových vektorů $C = (v_i)_1^k$ ve V takovou, že $f_\lambda(v_1) = o$ a $\forall i \in \{2, \ldots, k\}$ je $f_\lambda(v_i) = v_{i-1}$ nazveme *řetízek* příslušný f a λ. Definující vlastnost řetízku je možné ve zkratce zapsat také formou diagramu

$$v_k \xmapsto{f_\lambda} v_{k-1} \xmapsto{f_\lambda} \ldots \xmapsto{f_\lambda} v_2 \xmapsto{f_\lambda} v_1 \xmapsto{f_\lambda} o$$

První vektor v řetízku splňuje $f(v_1) = \lambda v_1$, patří tedy do vlastního podprostoru V_λ endomorfismu f. Ostatní vektory už vlastními vektory nejsou, splňují podmínku

$$f(v_i) = (f_\lambda + \lambda \operatorname{Id})(v_i) = v_{i-1} + \lambda v_i$$

Takovým vektorům se říká *přidružené vlastní vektory* endomorfismu f vůči vlastnímu číslu λ.

Pokud je C navíc i báze V, pak z definice matice endomorfismu plyne, že $[f]_C^C \in \mathbb{C}^{k\times k}$ má tvar

$$J_{\lambda,k} := \begin{pmatrix} \lambda & 1 & 0 & 0 & \ldots & 0 & 0 \\ 0 & \lambda & 1 & 0 & \ldots & 0 & 0 \\ 0 & 0 & \lambda & 1 & \ldots & 0 & 0 \\ \vdots & & & \ddots & \ddots & 0 & 0 \\ 0 & 0 & 0 & 0 & \ldots & 1 & 0 \\ 0 & 0 & 0 & 0 & \ldots & \lambda & 1 \\ 0 & 0 & 0 & 0 & \ldots & 0 & \lambda \end{pmatrix}$$

Této matici se říká *Jordanova buňka.* Hlavním výsledkem této kapitoly je, že pro každý endomorfismus V existuje báze B (*Jordanova báze*) prostoru V složená z řetízků příslušných jeho vlastním číslům. Vůči této bázi je $[f]_B^B$ blokově diagonální matice, jejíž diagonální bloky jsou Jordanovy buňky, říkáme, že je to matice v *Jordanově tvaru.* Mají-li všechny řetízky v Jordanově bázi délku 1, je Jordanův tvar diagonální matice, a tedy že f je diagonalizovatelný endomorfismus. Druhým důsledkem věty o Jordanově tvaru bude, že je určen pro daný endomorfismus jednoznačně až na pořadí Jordanových buněk. Z toho plyne, že matice má netriviální Jordanův tvar (tj. s alespoň jednou Jordanovou buňkou řádu většího než 1), právě když endomorfismus není diagonalizovatelný.

V kapitole 11 jsme viděli, že diagonalizace matice umožňuje efektivně spočítat její mocniny. Nalezení Jordanovy báze a Jordanova tvaru endomorfismu či matice umožňuje rozšířit tuto metodu i na matice nediagonalizovatelné. Nejprve dokážeme, že pro komutující matice platí obdoba binomické věty:

Lemma 7 (Binomická věta). *Nechť $A, B \in \mathbb{F}^{k\times k}$, $AB = BA$. Pak*

$$(A + B)^n = \sum_{i=0}^{n} \binom{n}{i} A^{n-i} B^i$$

Důkaz. Indukcí s využitím relací pro kombinační čísla, podobně jako u binomické věty pro reálná čísla. $\square$

K určení mocniny endomorfismu tedy stačí reprezentovat ho vůči Jordanově bázi a spočítat mocninu Jordanových buněk na blokové diagonále. Protože matice λE komutuje s každou maticí, lze použít binomickou větu:

$$(J_{\lambda,k})^n = (\lambda E + J_{0,k})^n = \sum_{i=0}^{n} \binom{n}{i} \lambda^{n-i} (J_{0,k})^i =$$

$$= \begin{pmatrix} \lambda^n & n\lambda^{n-1} & \binom{n}{2}\lambda^{n-2} & \dots & & \binom{n}{k-1}\lambda^{n-k+1} \\ 0 & \lambda^n & n\lambda^{n-1} & \dots & & \binom{n}{k-2}\lambda^{n-k+2} \\ 0 & 0 & \lambda^n & \ddots & & \binom{n}{k-3}\lambda^{n-k+3} \\ \vdots & & & \ddots & & \vdots \\ 0 & 0 & 0 & \dots & \lambda^n & n\lambda^{n-1} \\ 0 & 0 & 0 & \dots & 0 & \lambda^n \end{pmatrix}$$

Využili jsme jednoduchého tvaru mocnin Jordanových buněk s vlastním číslem 0. Pro dostatečně velké i (konkrétně $i \geq k$) je dokonce $(J_{0,k})^i = 0$, matice i endomorfismy s touto vlastností se nazývají *nilpotentní*.

Definice 56. *Invariantním podprostorem* endomorfismu $f \in \mathrm{End}(V)$ rozumíme podprostor $W \leq V$ takový, že $f(W)$ je podmnožinou W.

Například $W := \mathrm{Ker}\, f$ je invariantním podprostorem, protože $f(W) = 0 \subset W$. Podobně je vidět, že invariantním podprostorem jsou i $\mathrm{Im}\, f$ a lineární obal vektorů libovolného řetízku ve V spojeného f_λ. V definici se nepožaduje, aby šlo o vlastní podmnožinu, může platit i rovnost.

Lemma 8. *Jsou-li $U, W \leq V$ invariantní podprostory pro $f \in \mathrm{End}(V)$, pak jsou jimi i $U \cap W$ a $U + W$.*

Důkaz. Z definic. $\square$

Lemma 9. *Je-li $W \leq V$ invariantní podprostor endomorfismů $f, g \in \mathrm{End}(V)$ a $r, s \in \mathbb{F}$, pak je invariantní i pro endomorfismus $rf + sg$.*

Důkaz. Z definic. $\square$

Protože pro Id je invariantní každý podprostor, má $f_\lambda = f - \lambda\,\mathrm{Id}$ stejné invariantní podprostory jako f. Zvolíme-li bázi $B = (C, B')$ ve V tak, že C je báze invariantního podprostoru W, pak prvních $\dim W \equiv |C|$ sloupců matice $[f]_B^B$ má posledních $\dim V - \dim W \equiv |B'|$ složek nulových. Je to tedy blokově horní trojúhelníková matice

$$[f]_B^B = \begin{pmatrix} A & X \\ 0 & G \end{pmatrix},$$

kde $|C| \times |C|$ matici A jde s pomocí zobrazení $\iota : W \to V$, $\pi : V \to W$ zavedených v kapitole o direktním součtu zapsat jako $A = [\pi \circ f \circ \iota]_C^C$. Je to tedy matice $f_{|W} := \pi \circ f \circ \iota \in \mathrm{End}(W)$ *zúžení endomorfismu f na podprostor W*. Charakteristický polynom původního endomorfismu $p_f(\lambda) = \det(f - \lambda\,\mathrm{Id})$ je díky tvrzení 41 roven

$$\det(A - \lambda E)\det(G - \lambda E) = \det(f_{|W} - \lambda\,\mathrm{Id})\det(G - \lambda E) = p_{f_{|W}}(\lambda) q(\lambda),$$

tudíž každé vlastní číslo endomorfismu $f_{|W}$ je i vlastním číslem endomorfismu f. Pro speciální volbu W odtud plyne nerovnost mezi algebraickou a geometrickou násobností vlastního čísla:

Tvrzení 60. *Nechť $f \in \mathrm{End}(V)$. Pak algebraická násobnost vlastního čísla μ endomorfismu f je vždy větší nebo rovna jeho násobnosti geometrické.*

Důkaz. Je-li v předchozí úvaze W vlastní podprostor f s vlastním číslem μ, pak $A = \mu E$, tedy $p_{f_{|W}}(\lambda) = (\mu - \lambda)^{\dim W}$. Odtud vidíme, že μ je kořenem $p_f(\lambda)$ alespoň násobnosti $\dim W$. $\square$

Existenci Jordanovy báze budeme dokazovat ve dvou krocích. Nejprve ukážeme, že množina vektorů v řetízcích je lineárně nezávislá, právě když je lineárně nezávislá posloupnost jejich počátečních vektorů, tedy těch, které se v řetízcích zobrazují na nulový vektor. Druhým krokem bude ukázat, že lze sestrojit množinu řetízků splňujících tuto podmínku, která je dostatečně velká, aby generovala celé V.

Lemma 10. *Nechť $B = (B_1, \dots, B_p)$ je posloupnost řetízků pro $f \in \operatorname{End}(V)$. B je lineárně nezávislá, právě když je lineárně nezávislá posloupnost $(v_1^1, \dots, v_1^p)$ jejich počátečních vektorů.*

Důkaz. Implikace $\Rightarrow$ je zjevná, implikaci $\Leftarrow$ dokážeme indukcí podle celkového počtu vektorů v B. Nechť $B_1, \dots, B_q$ jsou řetízky spojené f_λ, tedy jejich počáteční vektory $v_1^1, \dots, v_1^q$ splňují $(f - \lambda \operatorname{Id})(v_1^i) = o$, čili jsou to vlastní vektory f s vlastním číslem λ. Podívejme se nejprve, na co se zobrazí lineární kombinace prvků některého z prvních q řetízků v zobrazení f_λ. Je-li $i \le q$, pak

$$f_\lambda\left(\sum_{j=1}^{k_i} r_j^i v_j^i\right) = \sum_{j=1}^{k_i} r_j^i f_\lambda(v_j^i) = \sum_{j=2}^{k_i} r_j^i v_{j-1}^i = \sum_{j=1}^{k_i-1} r_{j+1}^i v_j^i \in \langle B_i' \rangle,$$

kde $B_i' := B_i \setminus \{v_{k_i}^i\}$ je řetízek B_i po odebrání posledního vektoru. Pokud $i > q$, tedy pokud B_i je řetízek příslušný jinému vlastnímu číslu μ, pak je $\langle B_i \rangle$ díky předchozí úvaze invariantní podprostor pro f_μ, tedy i pro f a tedy i pro f_λ. Tedy f_λ-obrazem lineární kombinace prvků B_i bude opět nějaká lineární kombinace prvků B_i.

Lineární kombinaci všech vektorů ve všech řetízcích $B_1, \dots, B_p$ rozdělme na dvě dvojité sumy: první běží přes řetízky příslušející vlastnímu číslu λ, druhá přes řetízky příslušející ostatním vlastním číslům:

$$\sum_{i=1}^{q} \sum_{j=1}^{k_i} r_j^i v_j^i + \sum_{i=q+1}^{p} \sum_{j=1}^{k_i} r_j^i v_j^i$$

Endomorfismus f_λ zobrazí tuto lineární kombinaci na lineární kombinaci vektorů z posloupnosti $(B_1', \dots, B_q', B_{q+1}, \dots, B_p)$, kde koeficient u v_j^i pro $i \le q$ je r_{j+1}^i. Posloupnost počátečních vektorů řetízků $B_1', \dots, B_q', B_{q+1}, \dots, B_p$ je buď stejná jako posloupnost počátečních vektorů řetízků $B_1, \dots, B_q, B_{q+1}, \dots, B_p$, nebo je to její podposloupnost, je tedy lineárně nezávislá. Protože posloupnost vektorů $(B_1', \dots, B_q', B_{q+1}, \dots, B_p)$ obsahuje méně prvků než posloupnost $(B_1, \dots, B_p)$, plyne z indukčního předpokladu, že první z posloupností je lineárně nezávislá.

Položme tedy původní lineární kombinaci rovnu o, pak je roven nulovému vektoru i její f_λ-obraz. Z indukčního předpokladu a tvaru koeficientů f_λ-obrazu vektorů z prvních q řetízků plyne, že $r_j^i = 0$ pro všechna $i \le q$ a $j \ge 2$. Opakujeme-li stejnou úvahu s ostatními vlastními čísly a odpovídajícími zobrazeními f_μ dostáváme, že $r_j^i = 0$ i pro $i > q$ a $j \ge 2$. Pak má ale původní lineární kombinace tvar $\sum_{i=1}^p r_1^i v_1^i = o$ a tedy i z lineární nezávislosti počátečních vektorů plyne $r_1^i = 0$ pro všechna i. Tedy všechny koeficienty původní lineární kombinace musí být nulové a posloupnost $(B_1, \dots, B_p)$ je tudíž lineárně nezávislá. $\square$

Věta 38 (O existenci Jordanova tvaru). *Nechť V je komplexní vektorový prostor konečné dimenze, $f \in \operatorname{End} V$. Pak existuje báze B prostoru V taková, že $[f]_B^B$ je matice v Jordanově tvaru.*

Důkaz. Indukcí podle dimenze V. Pro $\dim V = 1$ je tvrzení zřejmé, předpokládejme tedy, že $n = \dim V > 1$ a pro všechny prostory nižší dimenze tvrzení platí. Pro $\lambda \in \sigma(f)$ jsou $\operatorname{Ker} f_\lambda$ a $\operatorname{Im} f_\lambda$ f-invariantní podprostory a protože $\dim \operatorname{Ker} f_\lambda > 0$, plyne z věty o dimenzi jádra a obrazu, že $\dim \operatorname{Im} f_\lambda < n$. Protože $W := \operatorname{Im} f_\lambda$ je invariantní podprostor f, lze z indukčního předpokladu najít Jordanovu bázi C endomorfismu $f|_W$ v prostoru W. Nechť je v C právě r řetízků s vlastním číslem λ. Protože $C \subset \operatorname{Im} f_\lambda$, lze ke každému koncovému vektoru $v^i_{k_i}$ takového řetízku najít nějaký jeho vzor $v^i_{k_i+1}$ v zobrazení f_λ. Lze také doplnit r počátečních vektorů těchto řetízků na bázi $\operatorname{Ker} f_\lambda$, tím získáme dalších $\dim \operatorname{Ker} f_\lambda - r$ vektorů tvořících řetízky délky 1 spojené f_λ. Tím jsme z C získali novou posloupnost řetízků B, v níž je o

$$r + (\dim \operatorname{Ker} f_\lambda - r) = \dim \operatorname{Ker} f_\lambda = n - \dim \operatorname{Im} f_\lambda = n - |C|$$

vektorů více, má tedy celkem n prvků. Podobnou úvahou jako v důkazu lemmatu 5 odvodíme, že sjednocení lineárně nezávislých množin, z nichž každá leží v jiném vlastním podprostoru, je lineárně nezávislá množina. Množina C_1 všech počátečních vektorů v C je z indukčního předpokladu lineárně nezávislá, tedy je lineárně nezávislý i její průnik s každým vlastním podprostorem f, speciálně i s $\operatorname{Ker} f_\lambda$, který označme C_1'. Množina B_1 všech počátečních vektorů v B vznikla doplněním C_1' na bázi $\operatorname{Ker} f_\lambda$ a přidáním všech ostatních prvků C_1, je tedy lineárně nezávislá. Podle lemmatu 10 je B lineárně nezávislá a jelikož má n prvků, je to báze V. □

Příklad 32. Zkonstruujme Jordanovu bázi pro endomorfismus $F_A : \mathbb{C}^4 \to \mathbb{C}^4$ určený maticí

$$A = \begin{pmatrix} 3 & 1 & 0 & -1 \\ 0 & 2 & 0 & 1 \\ 0 & 0 & 3 & 0 \\ 1 & 0 & 0 & 4 \end{pmatrix}$$

Spočteme postupně $\sigma(A) = \{3\}$, $\dim \operatorname{Ker}(A-3E) = 2$ a $(A-3E)^3 = 0$, $(A-3E)^2 \neq 0$. Zvolíme vektor $\mathbf{v}^1_3$ tak, aby se nenacházel v jádru matice

$$(A - 3E)^2 = \begin{pmatrix} -1 & -1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{pmatrix},$$

například $\mathbf{v}^1_3 := (1,0,0,0)^T$. Násobením maticí $A - 3E$ zleva získáme řetízek

$$\overset{\mathbf{v}^1_3=}{(1,0,0,0)^T} \overset{A-3E}{\longmapsto} \overset{\mathbf{v}^1_2:=}{(0,0,0,1)^T} \overset{A-3E}{\longmapsto} \overset{\mathbf{v}^1_1:=}{(-1,1,0,1)^T} \overset{A-3E}{\longmapsto} \mathbf{o}$$

Vektor $\mathbf{v}^1_1$ doplníme na bázi $\operatorname{Ker}(A - 3E)$ např. vektorem $\mathbf{v}^2_1 := (0,0,1,0)^T$. Z lemmatu 10 je $(\mathbf{v}^1_1, \mathbf{v}^1_2, \mathbf{v}^1_3, \mathbf{v}^2_1)$ lineárně nezávislá, je to tedy hledaná Jordanova báze. Jordanův tvar matice A je $\operatorname{diag}(J_{3,3}, J_{3,1})$.

Zapišme $[f]^B_B$ blokově jako $\operatorname{diag}(J, K)$, kde blok J zahrnuje všechny Jordanovy buňky příslušející vlastnímu číslu λ a K všechny ostatní Jordanovy buňky. Pak existuje $k \in \mathbb{N}$, že k-tá mocnina matice $[f_\lambda]^B_B$ je

$$\begin{pmatrix} (J - \lambda E)^k & 0 \\ 0 & (K - \lambda E)^k \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & R \end{pmatrix},$$

kde R je regulární matice. Nejmenší takové k je rovno délce nejdelšího řetízku z řetízků $B_1, \dots B_q$ příslušejících vlastnímu číslu λ. Protože R je regulární matice, je podprostor $Z_\lambda := \langle B_1, \dots, B_q \rangle$ vůči bázi B vyjádřen jako $\operatorname{Ker}([f_\lambda]^B_B)^k = \operatorname{Ker}([(f_\lambda)^k]^B_B)$. To znamená, že samotný Z_λ je roven $\operatorname{Ker}(f_\lambda)^k$. Poslední vyjádření závisí pouze na endomorfismu f a jeho vlastním čísle λ, ale už ne na volbě Jordanovy báze B.

Definice 57. Nechť $f \in \mathrm{End}(V)$, λ je vlastní číslo endomorfismu f a $k \in \mathbb{N}$ je nejmenší takové, že $\mathrm{Ker}(f_\lambda)^k = \mathrm{Ker}(f_\lambda)^{k+1}$. Podprostor $Z_\lambda := \mathrm{Ker}(f_\lambda)^k$ nazýváme *zobecněný vlastní podprostor* f příslušný vlastnímu číslu λ.

Rozmyslete si, že tato definice Z_λ souhlasí s definicí v předchozím odstavci. Ideu nezávislosti na volbě báze vytěžíme v následujícím tvrzení, jehož důsledkem bude, že počty a délky řetízků musí být v každé Jordanově bázi daného endomorfismu stejné.

Tvrzení 61. *Nechť V je vektorový prostor konečné dimenze n nad $\mathbb{C}$, $f \in \mathrm{End}(V)$ a $\lambda \in \sigma(f)$. Pak*

(1) $Z_\lambda \oplus \mathrm{Im}(f_\lambda)^k = V$, *$\dim Z_\lambda$ je rovno algebraické násobnosti λ*
(2) *Je-li $\sigma(f) = \{\lambda_1, \dots, \lambda_m\}$, pak $V = Z_{\lambda_1} \oplus \dots \oplus Z_{\lambda_m}$*
(3) *Počet řetízků příslušných λ s délkou alespoň j je roven*

$$\dim \mathrm{Ker}(f_\lambda)^j - \dim \mathrm{Ker}(f_\lambda)^{j-1}$$

(4) *Počet řetízků příslušných λ s délkou právě j je roven*

$$-\dim \mathrm{Ker}(f_\lambda)^{j+1} + 2\dim \mathrm{Ker}(f_\lambda)^j - \dim \mathrm{Ker}(f_\lambda)^{j-1}$$

Důkaz. Nechť B je Jordanova báze f a matice J, K jsou zvoleny stejně jako výše. Matice

$$[(f_\lambda)^k]_B^B = \begin{pmatrix} (J-\lambda E)^k & 0 \\ 0 & (K-\lambda E)^k \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & R \end{pmatrix}$$

má blokovou strukturu, v níž je rozměr nulové matice v levém horním rohu roven rozměru J, neboli počtu výskytů λ na diagonále horní trojúhelníkové matice $[f]_B^B$, jinými slovy algebraické násobnosti λ. Označme rozměr regulární matice R jako $n-m$. Zjevně $\mathrm{Im}[(f_\lambda)^k]_B^B = \langle \mathbf{e}_{m+1}, \dots, \mathbf{e}_n \rangle$ a $\mathrm{Ker}[(f_\lambda)^k]_B^B = \langle \mathbf{e}_1, \dots, \mathbf{e}_m \rangle$. Jádro a obraz matice $[(f_\lambda)^k]_B^B$ tedy tvoří direktní součet rovný $\mathbb{C}^n$, z čehož

$$\mathrm{Ker}(f_\lambda)^k \oplus \mathrm{Im}(f_\lambda)^k = V$$

Druhé tvrzení získáme indukcí z prvního a z pozorování, že $Z_\mu \subset \mathrm{Im}(f_\lambda)^k$ pro $\mu \neq \lambda$. To je opět nejlépe vidět z maticového vyjádření, protože matice K obsahuje na diagonále všechny Jordanovy buňky s vlastním číslem μ. Pro ověření třetího tvrzení si stačí rozmyslet, že v $\mathrm{Ker}(f_\lambda)^j$ leží z báze B právě ty vektory, které patří do nějakého řetízku spojeného f_λ a jsou vzdáleny od počátečního vektoru nejvýše o $j-1$ šipek. Čtvrté tvrzení plyne ihned ze třetího. □

Čtvrtý bod vyjadřuje počet řetízků délky j spojených f_λ nezávisle na zvolené Jordanově bázi. Každému takovému řetízku odpovídá v matici $[f]_B^B$ Jordanova buňka $J_{\lambda,j}$. Z toho plyne

Věta 39. *Jordanův tvar endomorfismu $f \in \mathrm{End}(V)$ je určen jednoznačně až na změnu pořadí Jordanových buněk.*

Při praktickém hledání nějaké Jordanovy báze postupujeme po jednotlivých zobecněných vlastních podprostorech Z_λ, v každém nejprve zjistíme strukturu řetízků z dimenzí $\mathrm{Ker}(f_\lambda)^j$. Konkrétní vektory v řetízcích nejsou určeny jednoznačně, nějakou vhodnou volbu můžeme získat následujícím postupem:

Pro k nejmenší takové, že $\mathrm{Ker}(f_\lambda)^k = \mathrm{Ker}(f_\lambda)^{k+1}$ najděme nejprve nějakou bázi $\mathrm{Ker}\, f_\lambda \cap \mathrm{Im}(f_\lambda)^{k-1}$, označme ji jako „nové počáteční vektory“ a položme $j = k-1$. Následně:

(1) Prohlásíme nové počáteční vektory za staré počáteční vektory a prodloužíme je na řetízky délky $j+1$ opakovaným hledáním vzorů v zobrazení f_λ.
(2) Doplníme staré počáteční vektory novými počátečními vektory na bázi $\operatorname{Ker} f_\lambda \cap \operatorname{Im}(f_\lambda)^{j-1}$
(3) Je-li $j=1$, skončíme, jinak snížíme j o 1 a jdeme na bod 1.

Tento postup zaručuje, že vzniknou řetízky s lineárně nezávislými počátečními vektory a správnými délkami. Pro malé matice je někdy výhodnější konstruovat řetízky od konce, protože hledání obrazu v f_λ je jednodušší než hledání vzoru v něm. Pak ale nemusí počáteční vektory vyjít lineárně nezávislé a je nutné to dodatečně ověřit.

Příklad 33. Matice

$$A = \begin{pmatrix} 0 & 1 & 1 \\ -3 & 4 & 3 \\ 2 & -1 & 1 \end{pmatrix}$$

má spektrum $\sigma(A) = \{1, 2\}$ a vlastní podprostory

$$V_1 = \operatorname{Ker}(A - E) = \langle (1, 2, -1)^T \rangle, V_2 = \operatorname{Ker}(A - 2E) = \langle (1, 3, -1)^T \rangle.$$

Protože 1 má algebraickou násobnost 1, je V_1 už roven Z_1. Prostor $\operatorname{Ker}(A-2E)^2 = \langle (0,0,1), (1,3,0) \rangle$ je roven Z_2, protože už má dimenzi rovnou algebraické násobnosti vlastního čísla 2. Dimenze V_2 je 1, bázi Z_2 tedy bude tvořit jeden řetízek délky 2. Vektory v něm můžeme najít buď jako (např.) $(1, 3, -1)$ a nějaký jeho vzor v F_{A-2E}, nebo jako (např.) $(0, 0, 1)$ a nějaký jeho obraz v F_{A-2E}. Obraz se spočte snáz a je roven $(1, 3, -1)$. Nalezená Jordanova báze a Jordanův tvar jsou pak

$$B = \left(\begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right), \quad J_A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$$

Mocniny matice A lze spočíst jako $A^n = R J_A^n R^{-1}$, $R := [\mathrm{Id}]^K_B$.

Poznámka 7. Při numerickém hledání Jordanova tvaru narážíme na problém, že libovolně malá změna vstupních dat může změnit strukturu řetízků. Například matice

$$N = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

není diagonalizovatelná a je sama svým Jordanovým tvarem. K ní blízká matice

$$A = \begin{pmatrix} \varepsilon & 1 \\ 0 & -\varepsilon \end{pmatrix}$$

diagonalizovatelná je, protože má prosté spektrum $\{\varepsilon, -\varepsilon\}$. Její Jordanův tvar tedy je

$$J_A = \begin{pmatrix} \varepsilon & 0 \\ 0 & -\varepsilon \end{pmatrix}$$

Vlastní podprostory $\langle (1, 0)^T \rangle$ a $\langle (1, -2\varepsilon)^T \rangle$ matice A pro $\varepsilon = 0$ splynou v jediný vlastní podprostor matice N.

KAPITOLA 20

Exponenciála

Označení 1. Standardně značíme maticové elementy matice A symbolem a_{ij}. V některých případech, například u inverzní matice A^{-1}, by označení elementů (např. a_{ij}^{-1}) bylo zavádějící. Proto budeme značit ij-tý element matice A v případě potřeby také $(A)_{ij}$. Naopak (a_{ij}) bude označovat matici, jejíž elementy jsou a_{ij}.

Pomocí Jordanova tvaru umíme spočítat libovolnou mocninu libovolné komplexní matice a tedy i dosadit matici do polynomu. Je možné matici dosadit i do mocninné řady?

Definice 58. Je-li $(B_q)_{q=0}^{\infty}$ posloupnost matic z $\mathbb{C}^{m\times n}$, pak matici $B \in \mathbb{C}^{m\times n}$ nazveme limitou $\lim_{q\to\infty} B_q$ posloupnosti, pokud $\forall i \in \{1,\dots,m\}$, $\forall j \in \{1,\dots,n\}$ platí $\lim_{q\to\infty}(B_q)_{ij} = (B)_{ij}$.

Součet nekonečné řady matic je pak definován jako limita jejích částečných součtů. Definici limity můžeme přeformulovat i pomocí zavedení *normy* na množině matic vztahem $\|B\| := \max_{i,j} |(B)_{ij}|$. Platí

Lemma 11. *Nechť $(B_q)_{q=0}^{\infty}$ je posloupnost matic z $\mathbb{C}^{m\times n}$. Pak $B \in \mathbb{C}^{m\times n}$ je její limita, právě když $\lim_{q\to\infty} \|B_q - B\| = 0$.*

Důkaz. Jednoduché cvičení. □

Definice 59. Pro libovolnou čtvercovou matici A definujme její *exponenciálu*

$$\exp A = \sum_{q=0}^{\infty} \frac{1}{q!} A^q$$

Lemma 12. *Pro libovolnou čtvercovou matici je* $\exp A$ *absolutně konvergentní řada.*

Důkaz. Nechť $A \in \mathbb{C}^{n\times n}$. Pro všechna $q \in \mathbb{N}$ platí nerovnost $\|A^q\| \le n^{q-1}\|A\|^q$, která se snadno dokáže indukcí. Pro $q = 1$ je tvrzení triviální. Předpokládejme, že platí pro nějaké $q \in \mathbb{N}$. Pak

$$\begin{aligned} \|A^{q+1}\| = \|A^q A\| = \max_{ij} |\sum_k (A^q)_{ik}(A)_{kj}| \le \max_{ij} \sum_k |(A^q)_{ik}||(A)_{kj}| \le \\ \le n\|A\| \max_{ik} |(A^q)_{ik}| = n\|A\|\|A^q\| \le n^q \|A\|^{q+1}, \end{aligned}$$

kde jsme v posledním kroku použili indukční předpoklad.

Pak ale pro libovolné $Q \in \mathbb{N}$

$$\sum_{q=0}^{Q} \frac{1}{q!}|(A^q)_{ij}| \le \sum_{q=0}^{Q} \frac{1}{q!}\|A^q\| \le \sum_{q=0}^{Q} \frac{n^q}{q!}\|A\|^q \le e^{n\|A\|}$$

□

PŘÍKLAD 34. Pokud je nějaká mocnina matice A rovna nule, říkáme, že je A *nilpotentní matice*. Pak je řada $\exp A$ dána konečným součtem. Příkladem je Jordanova buňka $J_{0,k}$:

$$\exp J_{0,k} = E + J_{0,k} + \frac{1}{2}(J_{0,k})^2 + \ldots + \frac{1}{(k-1)!}(J_{0,k})^{k-1}$$

U některých speciálních matic je možné spočítat obecnou mocninu přímo a i řada se dá pak sečíst. To je případ diagonální matice

$$\exp\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} = \sum_{q=0}^{\infty} \frac{1}{q!}\begin{pmatrix} a^q & 0 \\ 0 & b^q \end{pmatrix} = \begin{pmatrix} e^a & 0 \\ 0 & e^b \end{pmatrix},$$

ale také například matic

$$\exp\begin{pmatrix} 0 & t \\ t & 0 \end{pmatrix} = \begin{pmatrix} \cosh t & \sinh t \\ \sinh t & \cosh t \end{pmatrix}, \quad \exp\begin{pmatrix} 0 & -t \\ t & 0 \end{pmatrix} = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}$$

VĚTA 40. *Nechť $A, B, Q \in \mathbb{C}^{n\times n}$, Q regulární. Pak*

(1) *Pokud $AB = BA$, pak $\exp(A+B) = \exp A \exp B$*
(2) *$\exp A$ je regulární a $(\exp A)^{-1} = \exp(-A)$*
(3) *$\exp(Q^{-1}AQ) = Q^{-1}\exp(A)Q$*
(4) *$\det \exp A = \exp \operatorname{Tr} A$*

DŮKAZ. $\exp A$ a $\exp B$ jsou dvě absolutně konvergentní řady, jejich součin $\exp A \exp B$ je tedy roven řadě

$$\sum_{q=0}^{\infty}\frac{1}{q!}(A)^q \sum_{r=0}^{\infty}\frac{1}{r!}(B)^r = \sum_{p=0}^{\infty}\sum_{q=0}^{p}\frac{1}{q!(p-q)!}A^q B^{p-q} =$$
$$= \sum_{p=0}^{\infty}\frac{1}{p!}\sum_{q=0}^{p}\binom{p}{q}A^q B^{p-q} = \sum_{p=0}^{\infty}\frac{1}{p!}(A+B)^p,$$

v poslední rovnosti jsme využili lemma 7. Tím dostáváme první tvrzení. Druhé plyne z prvního, protože A a $-A$ komutují a $\exp 0 = E$ na základě přímého dosazení. Třetí je důsledkem vlastnosti $QA^qQ^{-1} = (QAQ^{-1})^q$ a věty o aritmetice limit.

Poslední tvrzení dokážeme pomocí věty o Jordanově tvaru. Pokud $A = QJ_AQ^{-1}$, kde J_A je Jordanův tvar A, pak platí $\det\exp A = \det\exp J_A$ ze třetího bodu a vlastností determinantu a $\exp \operatorname{Tr} A = \exp \operatorname{Tr} J_A$ z vlastností stopy. Z tvaru mocnin Jordanových buněk vidíme, že $\exp J_A$ je horní trojúhelníková matice, která má na diagonále čísla e^{λ_i}, $i \in \{1,\ldots,n\}$, kde λ_i jsou vlastní čísla A, každé tolikrát, kolik je jeho algebraická násobnost (viz též níže explicitní tvar exponenciály Jordanovy buňky). Pak ale

$$\det\exp J_A = \prod_{i=1}^{n} e^{\lambda_i} = e^{\sum_{i=1}^{n}\lambda_i} = e^{\operatorname{Tr} J_A}$$

$\square$

Jordanův tvar umožňuje explicitně spočítat exponenciálu libovolné matice. Pro jednoduchost se věnujme případu $A = QJQ^{-1}$, kde J je Jordanova buňka stupně n s vlastním číslem λ. Pak $J = \lambda E + N$ je rozklad na součet dvou komutujících matic, takže díky prvnímu bodu předchozí

$$\begin{aligned}
\exp A &= Q\exp(J)Q^{-1} = Q\exp(\lambda E + N)Q^{-1} \\
&= Q\exp(\lambda E)\exp NQ^{-1} = Qe^{\lambda}\exp NQ^{-1} \\
&= Qe^{\lambda}\begin{pmatrix} 1 & \frac{1}{1!} & \frac{1}{2!} & \dots & \frac{1}{(n-1)!} \\ 0 & 1 & \frac{1}{1!} & \dots & \frac{1}{(n-2)!} \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & 1 & \frac{1}{1!} \\ 0 & \dots & 0 & 0 & 1 \end{pmatrix} Q^{-1}
\end{aligned}$$

Pokud $t \in \mathbb{C}$, lehkým zobecněním tohoto postupu plyne

$$\exp(tA) = Qe^{t\lambda}\begin{pmatrix} 1 & t & \frac{t^2}{2!} & \dots & \frac{t^{n-1}}{(n-1)!} \\ 0 & 1 & t & \dots & \frac{t^{n-2}}{(n-2)!} \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & 1 & t \\ 0 & \dots & 0 & 0 & 1 \end{pmatrix} Q^{-1},$$

což, jak za chvíli uvidíme, je výsledek klíčový pro řešení soustav lineárních diferenciálních rovnic.

Příklad 35. Hledáme dvě funkce $x_1, x_2 : \mathbb{R} \to \mathbb{C}$ splňující rovnice

$$\begin{aligned}
\frac{dx_1}{dt}(t) &= -x_1(t) - x_2(t) \\
\frac{dx_2}{dt}(t) &= -4x_1(t) - x_2(t)
\end{aligned}$$

s počáteční podmínkou $x_1(0) = 2$, $x_2(0) = 1$. Zadání můžeme kompaktně přepsat jako

$$\dot{\mathbf{x}} = A\mathbf{x}, \text{ kde } A = \begin{pmatrix} -1 & -1 \\ -4 & -1 \end{pmatrix}, \mathbf{x}(0) = (2,1)^T$$

Taková úloha se nazývá homogenní soustavou lineárních diferenciálních rovnic prvního řádu s konstantními koeficienty. Ukážeme, že její obecné řešení má tvar $\mathbf{x} = \exp(tA) \cdot \mathbf{c}$.

Definice 60. Derivací maticové funkce $Q : \mathbb{R} \to \mathbb{F}^{m\times n}$ v bodě $t \in \mathbb{R}$ rozumíme

$$\lim_{h\to 0} \frac{Q(t+h) - Q(t)}{h} =: \dot{Q}(t) \equiv \frac{dQ(t)}{dt}$$

Je zřejmé, že derivací maticové funkce $Q(t)$ je matice $(\dot{Q}(t))_{ij}$, jejíž elementy jsou derivacemi maticových elementů $(Q(t))_{ij}$.

Lemma 13. *Nechť $P, Q : \mathbb{R} \to \mathbb{F}^{m\times n}, A \in \mathbb{F}^{n\times p}, B \in \mathbb{F}^{n\times m}, R : \mathbb{R} \to \mathbb{F}^{n\times p}$, $f : \mathbb{R} \to \mathbb{F}$ a funkce P, Q, R, f mají derivaci v bodě t. Pak*

(1) $\frac{d}{dt}(P(t) + Q(t)) = \frac{dP(t)}{dt} + \frac{dQ(t)}{dt}$
(2) $\frac{d}{dt}(Q(t)A) = \frac{dQ(t)}{dt}A$
(3) $\frac{d}{dt}(BQ(t)) = B\frac{dQ(t)}{dt}$
(4) $\frac{d}{dt}(Q(t)R(t)) = \frac{dQ(t)}{dt}R(t) + Q(t)\frac{dR(t)}{dt}$
(5) $\frac{d}{dt}(f(t)Q(t)) = \frac{df(t)}{dt}Q(t) + f(t)\frac{dQ(t)}{dt}$
(6) *Pro A čtvercovou platí*

$$\frac{d}{dt}\exp(tA) = A\exp(tA)$$

Důkaz. Prvních pět tvrzení plyne okamžitě z vlastností derivace číselných funkcí a definic maticového sčítání a násobení. Pro poslední potřebujeme větu z matematické analýzy, která říká, že mocninná řada, která vznikne derivací mocninné řady člen po členu, má stejný poloměr konvergence a uvnitř kruhu konvergence je rovna derivaci součtu původní řady. Mocninná řada $(\exp tA)_{ij}$ konverguje pro všechna $t \in \mathbb{R}$, tedy její derivace v t je rovna

$$\frac{d}{dt}(\exp(tA))_{ij} = \frac{d}{dt}\sum_{q=0}^{\infty}\frac{t^q(A^q)_{ij}}{q!} = \sum_{q=1}^{\infty}\frac{t^{q-1}(A^q)_{ij}}{(q-1)!} = (A\exp(tA))_{ij}$$

□

Poznámka 8. Poslední tvrzení se dá také dokázat čistě algebraicky přechodem k Jordanovu tvaru $\exp(tA) = U\exp(tJ)U^{-1} = U\exp(tD)\exp(tN)U^{-1}$, kde matice $\exp(tD)$ a $\exp(tN)$ explicitně známe.

Věta 41. *Nechť $A \in \mathbb{C}^{n\times n}$, $\mathbf{c} \in \mathbb{C}^n$. Pak jediná funkce $\mathbf{x} : \mathbb{R} \to \mathbb{C}^n$, splňující rovnici $\dot{\mathbf{x}}(t) = A\mathbf{x}(t)$ s počáteční podmínkou $\mathbf{x}(0) = \mathbf{c}$ je $\mathbf{x}(t) = \exp(tA)\mathbf{c}$.*

Důkaz. Rovnosti $\mathbf{x}(0) = \exp(0)\mathbf{c} = \mathbf{c}$ a $\dot{\mathbf{x}}(t) = A\exp(tA)\mathbf{c} = A\mathbf{x}(t)$ jsou pro $\mathbf{x}(t)$ z věty splněny. Pokud by $\mathbf{y}(t)$ splňovalo tyto rovnosti také, pak platí

$$\begin{aligned}\frac{d}{dt}(\exp(-tA)\mathbf{y}(t)) &= \exp(-tA)\dot{\mathbf{y}}(t) + (-A)\exp(-tA)\mathbf{y}(t)\\ &= \exp(-tA)\,[A\mathbf{y}(t) - A\mathbf{y}(t)] = 0,\end{aligned}$$

tedy $\exp(-tA)\mathbf{y}(t) =: \mathbf{d}$ je konstantní vektor. Protože $\mathbf{y}(0) = \mathbf{c}$, platí $\mathbf{d} = \mathbf{c}$. Tedy $\mathbf{y}(t) = \exp(tA)\mathbf{c} \equiv \mathbf{x}(t)$. □

Příklad 36. Matice A z úvodu tohoto oddílu je diagonalizovatelná, platí

$$\begin{pmatrix} -1 & -1 \\ -4 & -1 \end{pmatrix} = \frac{1}{4}\begin{pmatrix} 1 & 1 \\ -2 & 2 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -3 \end{pmatrix}\begin{pmatrix} 2 & -1 \\ 2 & 1 \end{pmatrix}$$

Řešení soustavy rovnic je pak

$$\begin{aligned}\begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} &= \frac{1}{4}\begin{pmatrix} 1 & 1 \\ -2 & 2 \end{pmatrix}\begin{pmatrix} e^t & 0 \\ 0 & e^{-3t} \end{pmatrix}\begin{pmatrix} 2 & -1 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} 2 \\ 1 \end{pmatrix}\\ &= \frac{1}{4}\begin{pmatrix} 3e^t + 5e^{-3t} \\ -6e^t + 10e^{-3t} \end{pmatrix}\end{aligned}$$

Budeme-li interpretovat vektorovou funkci $(x_1(t), x_2(t))$ jako pohyb bodu v rovině, pak v souřadnicích vzhledem k Jordanově bázi má řešení tvar $x_1'(t) = c_1'e^t$, $x_2'(t) = c_2'e^{-3t}$. Bod $(s, 0)$ na vodorovné ose se tedy z počátečního stavu vzdaluje s rostoucím časem do nekonečna, bod $(0, s)$ na svislé ose konverguje k nule. Body mimo osy se pohybují po drahách trochu připomínajících hyperboly směrem k bodům $(\pm\infty, 0)$. Tyto dráhy jsou v každém bodě $\mathbf{x} \in \mathbb{R}^2$ tečné k vektoru $V_A(\mathbf{x}) := A\mathbf{x}$ a nazýváme je integrálními křivkami vektorového pole $V_A : \mathbb{R}^2 \to \mathbb{R}^2$.

Každá 2×2 reálná matice se dvěma různými vlastními čísly je diagonalizovatelná a její vlastní čísla jsou buď obě reálná, nebo jsou komplexně sdružená. V druhém případě označme $\lambda = k + i\omega$ a vytkněme společný faktor e^{kt}:

$$\exp(tA) = Q\begin{pmatrix} e^{t\lambda} & 0 \\ 0 & e^{-t\bar{\lambda}} \end{pmatrix}Q^{-1} = e^{tk}Q\begin{pmatrix} e^{it\omega} & 0 \\ 0 & e^{-it\omega} \end{pmatrix}Q^{-1}$$

Unitární matici uprostřed je možné unitární transformací převést na matici rotace a obě matice přechodu sloučit do jedné:

$$e^{tk}Q\frac{1}{2}\begin{pmatrix} i & 1 \\ -i & 1 \end{pmatrix}\begin{pmatrix} \cos(t\omega) & \sin(t\omega) \\ -\sin(t\omega) & \cos(t\omega) \end{pmatrix}\begin{pmatrix} -i & i \\ 1 & 1 \end{pmatrix}Q^{-1} =$$
$$= Pe^{tk}\begin{pmatrix} \cos(t\omega) & \sin(t\omega) \\ -\sin(t\omega) & \cos(t\omega) \end{pmatrix}P^{-1}.$$

Je vidět, že P musí být reálná matice. Integrální křivka v souřadnicích vzhledem k bázi ze sloupců P má parametrické vyjádření

$$\gamma_c(t) = e^{tk}\begin{pmatrix} \cos(t\omega) & \sin(t\omega) \\ -\sin(t\omega) & \cos(t\omega) \end{pmatrix}\begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$$

Pro $k = 0$ je to kružnice, pro $k < 0$ spirála směřující k $(0,0)$, pro $k > 0$ spirála směřující k nekonečnu.

Na soustavu lineárních diferenciálních rovnic prvního řádu s konstantními koeficienty lze převést i soustavy řádů vyšších. Jednoduchým příkladem je rovnice harmonického oscilátoru

$$\ddot{y}(t) = -\omega^2 y(t)$$

Definicí $x_1(t) := \dot{y}(t)$ a $x_2(t) := \omega y(t)$, lze rovnici přepsat jako

$$\dot{\mathbf{x}}(t) \equiv \begin{pmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{pmatrix} = \begin{pmatrix} 0 & -\omega \\ \omega & 0 \end{pmatrix}\begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix}$$

a tuto soustavu pak vyřešit:

$$\mathbf{x}(t) = \exp t\begin{pmatrix} 0 & -\omega \\ \omega & 0 \end{pmatrix}\mathbf{x}(0) = \begin{pmatrix} \cos\omega t & -\sin\omega t \\ \sin\omega t & \cos\omega t \end{pmatrix}\begin{pmatrix} x_1(0) \\ x_2(0) \end{pmatrix}$$

Tím dostáváme standardní řešení harmonického oscilátoru ve tvaru

$$y(t) = \frac{1}{\omega}x_2(t) = \frac{1}{\omega}\dot{y}(0)\sin\omega t + y(0)\cos\omega t$$

Soustavy s pravou stranou lze řešit pomocí tzv. variace konstant:

Věta 42. *Nechť $A \in \mathbb{C}^{n\times n}$, $\mathbf{b} : \mathbb{R} \to \mathbb{C}^n$. Pak všechna řešení nehomogenní soustavy diferenciálních rovnic*

$$\dot{\mathbf{x}}(t) = A\mathbf{x}(t) + \mathbf{b}(t)$$

jsou tvaru $\mathbf{x}(t) = \exp(tA)\mathbf{c}(t)$, kde $\mathbf{c} : \mathbb{R} \to \mathbb{C}^n$ je funkce splňující $\dot{\mathbf{c}}(t) = \exp(-tA)\mathbf{b}(t)$.

Důkaz. Pokud $\mathbf{x}(t), \tilde{\mathbf{x}}(t)$ jsou dvě řešení soustavy $\dot{\mathbf{x}}(t) = A\mathbf{x}(t) + \mathbf{b}(t)$, pak jejich rozdíl $\mathbf{x}(t) - \tilde{\mathbf{x}}(t)$ je řešením homogenní soustavy $\dot{\mathbf{x}}(t) = A\mathbf{x}(t)$. Stačí tedy najít partikulární řešení nehomogenní soustavy. Pokud ho budeme hledat ve tvaru $\mathbf{x}_P(t) = \exp(tA)\mathbf{c}(t)$, pak s využitím lemmatu 13 dostáváme

$$\dot{\mathbf{x}}_P(t) = \frac{d}{dt}(\exp(tA)\mathbf{c}(t)) = \exp(tA)\dot{\mathbf{c}}(t) + A\exp(tA)\mathbf{c}(t) = \exp(tA)\dot{\mathbf{c}}(t) + A\mathbf{x}_P(t),$$

a po porovnání s rovnicí dostáváme $\dot{\mathbf{c}}(t) = \exp(-tA)\mathbf{b}(t)$. Obecné řešení nehomogenní soustavy je

$$\mathbf{x}(t) = \exp(tA)\mathbf{c}(t) + \exp(tA)\mathbf{d} = \exp(tA)(\mathbf{c}(t) + \mathbf{d}),$$

a protože $\frac{d}{dt}(\mathbf{c}(t) + \mathbf{d}) = \dot{\mathbf{c}}(t)$, mají všechna řešení požadovaný tvar. □

Příklad 37. Zajímavá speciální situace nastává, když A má vlastní vektor $\mathbf{v}$ s vlastním číslem λ a pravá strana je ve tvaru $\mathbf{b}(t) = e^{\mu t}\mathbf{v}$. Pak

$$\dot{\mathbf{c}}(t) = \exp(-tA)\mathbf{b}(t) = e^{(\mu-\lambda)t}\mathbf{v},$$

čili

$$\mathbf{y}(t) = \exp(tA)\mathbf{c}(t) = \frac{1}{\mu - \lambda} e^{(\mu-\lambda)t} \exp(tA)\mathbf{v} = \frac{1}{\mu - \lambda} e^{\mu t}\mathbf{v}$$

Pokud λ i μ jsou ryze imaginární, popisuje výsledek periodický pohyb s úhlovou frekvencí μ a amplitudou úměrnou $\frac{1}{\mu-\lambda}$. Čím jsou si λ a μ blíže, tím více roste amplituda. Dostáváme tedy jednoduchý model rezonance při působení vnější síly s frekvencí blízkou frekvenci vlastních kmitů systému.

KAPITOLA 21

Tenzory

Nechť V je vektorový prostor nad $\mathbb{F}$. Lineární zobrazení $\phi : V \to \mathbb{F}$ se nazývá *lineární forma* (někdy též *lineární funkcionál* či *kovektor*).

PŘÍKLADY 6. Lineárních zobrazení už jsme viděli mnoho, lineární formy jsou pouze ty z nich, jejichž hodnoty jsou skaláry, např.:

- $\phi : \mathbb{R}^2 \to \mathbb{R}$, $\phi((x_1, x_2)^T) = 2x_1 - 5x_2$
- $\phi : \mathbb{R}^{n\times n} \to \mathbb{R}$, $\phi(A) = \operatorname{Tr} A$
- $\phi : P(x, \mathbb{R}) \to \mathbb{R}$, $\phi(p) = p(0)$
- $\phi : C^0([0, 1], \mathbb{R}) \to \mathbb{R}$, $\phi(f) = \int_0^1 f(x)dx$

Je-li V prostor konečné dimenze a $B = (e_i)_1^n$ jeho báze, pak $1 \times n$ matice $[\phi]_B^{K_1} = (\phi(e_1), \ldots, \phi(e_n))$ reprezentuje ϕ vůči bázím B ve V a $K_1 = (1)$ v $\mathbb{F}^1$. Je to tedy řádkový vektor, který budeme od nynějška značit $[\phi]_B$. Pro $v \in V$ platí

$$\phi(v) = [\phi]_B^{K_1}[v]^B = [\phi]_B[v]^B$$

DEFINICE 61. Prostor $\operatorname{Hom}(V, \mathbb{F})$ všech lineárních forem na V se nazývá *duální prostor k V* a značí se V^*. Je-li $\dim V = n$ a $B = (e_i)_1^n$ báze V, pak posloupnost $B^* := (e^i)_1^n \subset V^*$, kde $[e^i]_B = \mathbf{e}_i^T$ (neboli $e^i(e_j) = \delta_j^i$), se nazývá *duální báze* k B.

Zdůrazněme, že v definici

- $\mathbf{e}_i$ značí i-tý vektor kanonické báze v $\mathbb{F}^n$ (tedy aritmetický vektor)
- e_i značí i-tý vektor báze B ve V (tedy *abstraktní* vektor)
- e^i značí i-tý vektor duální báze B^* ve V^* (tedy lineární formu).

Rozlišování typu objektu pomocí pozice indexu je jedním z hlavních rysů tenzorové notace. Dalším je tzv. sumační konvence, kdy výskyt dvojice stejných indexů v horní a dolní poloze znamená, že se přes tento index sčítá. Je-li například $v \in V$ s reprezentací $[v]^B = (v^1, \ldots, v^n)^T$, můžeme psát

$$v = \sum_{i=1}^n v^i e_i =: v^i e_i$$

Proto také souřadnice vektorů budeme již od teď psát vždy s horním indexem. Pro prvky B^* platí

$$e^i(v) = e^i\left(\sum_{j=1}^n v^j e_j\right) = \sum_{j=1}^n v^j e^i(e_j) = \sum_{j=1}^n v^j \delta_j^i = v^i$$

čili i-tý prvek duální báze B^* přiřadí vektoru z V jeho i-tou souřadnici vzhledem k bázi B. Pro lepší pochopení jsme ve výrazech zachovali sumy, se sumační konvencí by stejné odvození bylo zapsáno

$$e^i(v) = e^i(v^j e_j) = v^j e^i(e_j) = v^j \delta_j^i = v^i$$

LEMMA 14. *Je-li V vektorový prostor nad $\mathbb{F}$ dimenze n a $B = (e_i)_1^n$ jeho báze, je posloupnost $B^* := (e^i)_1^n$ bází duálního prostoru V^* a pro libovolný prvek $\alpha \in V^*$ platí, že*

$$[\alpha]^{B^*} = (\alpha(e_1), \ldots, \alpha(e_n))^T = ([\alpha]_B)^T =: (\alpha_1, \ldots, \alpha_n)^T$$

Lemma říká, že pro libovolnou bázi B ve V je báze B^* k ní duální skutečně bází V^* a že souřadnice lineární formy α vůči B^* se rovnají hodnotám α na prvcích B, tedy prvkům jednořádkové matice $[\alpha]_B$ lineárního zobrazení α vzhledem k bázím B a K_1. Zavádí také konvenci, že souřadnice α_i kovektoru α budeme psát s dolními indexy.

Důkaz. Pro libovolný $v \in V$ platí

$$\alpha(v) = [\alpha]_B[v]^B = \alpha_i v^i = \alpha_i e^i(v) = (\alpha_i e^i)(v)$$

Prvky α a $\alpha_i e^i$ z $\mathrm{Hom}(V, \mathbb{F})$ jsou si tedy rovny a kovektor α je tudíž v lineárním obalu posloupnosti kovektorů $(e^1, \ldots, e^n)$. Tedy B^* generuje V^* a protože $\dim V^* = \dim \mathrm{Hom}(V, \mathbb{F}) = n$ z věty 8, musí to být podle důsledku 1 i báze prostoru V^*. □

Pokud $V = \mathbb{F}^n$, označme K jeho kanonickou bázi, její prvky $\varepsilon_i := \mathbf{e}_i$ a

$$K^* := \{\varepsilon^1, \ldots, \varepsilon^n\}$$

bázi ke K duální. Zobrazení ε^i tedy přiřazuje vektoru $x \in \mathbb{F}^n$ jeho i-tou složku.

Příklad 38. Nechť $V = \mathbb{R}^2$, $M = \{(3, 2), (4, 3)\} \equiv \{e_1, e_2\}$. Prvky duální báze $M^* = \{e^1, e^2\}$ lze zapsat jako $e^1 = a\varepsilon^1 + b\varepsilon^2$, $e^2 = c\varepsilon^1 + d\varepsilon^2$, kde a, b, c, d jsou nějaká čísla. Podmínky na duální bázi říkají, že

$$\begin{aligned}
e^1(e_1) &= a\varepsilon^1(e_1) + b\varepsilon^2(e_1) = 3a + 2b = 1\\
e^1(e_2) &= a\varepsilon^1(e_2) + b\varepsilon^2(e_2) = 4a + 3b = 0\\
e^2(e_1) &= c\varepsilon^1(e_1) + d\varepsilon^2(e_1) = 3c + 2d = 0\\
e^2(e_2) &= c\varepsilon^1(e_2) + d\varepsilon^2(e_2) = 4c + 3d = 1
\end{aligned}$$

To je vlastně maticová rovnice

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 3 & 4 \\ 2 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

čili úloha na inverzní matici. Jejím výpočtem dostáváme duální bázi $\{3\varepsilon^1 - 4\varepsilon^2, -2\varepsilon^1 + 3\varepsilon^2\}$.

Je-li $B' = \{e'_1, \ldots, e'_n\}$ jiná báze V a $R := [\mathrm{Id}]^B_{B'}$ je matice přechodu od B k B', pak

$$e'_j = \sum_{i=1}^n e_i (R)_{ij} =: R^i_j e_i$$

První vyjádření vychází z definice 25 matice přechodu, druhé opět rozšiřuje naši sumační a indexovou konvenci. Sčítá se přes řádkový index matice přechodu a dolní index bázového vektoru e_i, budeme tedy řádkový index matice přechodu psát nahoře a sumu přes i vynecháme. Sloupcový index matice přechodu ponecháme v dolní poloze, protože to odpovídá poloze indexu bázového vektoru e'_j na levé straně. Rovnost $[v]^B = [\mathrm{Id}]^B_{B'}[v]^{B'}$, nejprve bez sumační konvence a poté s ní, nabývá tvaru

$$v^j = \sum_{i=1}^n (R)_{ji} v'^i \equiv R^j_i v'^i$$

nebo též

$$v'^j = (R^{-1})^j_i v^i$$

Lemma 15. *Nechť $B^* = \{e^1, \ldots, e^n\}$, $B'^* = \{e'^1, \ldots, e'^n\}$, jsou báze V^* duální k bázím B a B', R je matice přechodu od B k B', $\alpha \in V^*$. Pak*

$$e'^i = (R^{-1})^i_j e^j$$

$$\alpha'_i = R^j_i \alpha_j$$

Důkaz. Pokud $v \in V$, pak $v'^i = (R^{-1})^i_j v^j$. Protože $v^j = e^j(v)$, $v'^i = e'^i(v)$, plyne odtud

$$e'^i(v) = \left((R^{-1})^i_j e^j\right)(v)$$

neboli rovnost zobrazení $e'^i = (R^{-1})^i_j e^j$. Podobně

$$\alpha'_i = \alpha(e'_i) = R^j_i \alpha(e_j) = R^j_i \alpha_j$$

□

Doposud odvozené transformační vztahy můžeme shrnout do tabulky:

Objekt	**maticově**	**tenzorově**
prvky báze V	$e'_i = \sum_j e_j R_{ji}$	$e'_i = R^j_i e_j$
prvky báze V^*	$e'^i = \sum_j (R^{-1})_{ij} e^j$	$e'^i = (R^{-1})^i_j e^j$.
souř. vektoru	$([v]^{B'})^T = ([v]^B)^T (R^{-1})^T$	$v'^i = (R^{-1})^i_j v^j$
souř. kovektoru	$[\alpha]_{B'} = [\alpha]_B R$	$\alpha'_i = R^j_i \alpha_j$

Říkáme, že prvky báze V a souřadnice kovektoru se transformují *kovariantně*, zbylé dva objekty *kontravariantně*. Rovnost ve třetím řádku sloupce „maticově“ je možné jednodušeji zapsat ve tvaru $[v]^{B'} = R^{-1}[v]^B$, transponovaná verze ale dává lepší srovnání s rovností o řádek níž. V obou případech má vztah tvar „řádkový vektor čárkovaných souřadnice je roven řádkovému vektoru nečárkovaných souřadnic násobenému maticí“, ale matice se od sebe liší o složení operací transponování a invertování. Matice $(R^{-1})^T$ (někdy ve zkratce psáno R^{-T} se nazývá matice *kontragredientní* k matici R.

Z tabulky vidíme, že v tenzorovém zápise je typ transformace objektu zakódován polohou indexu. Objekty s indexem dole se transformují pomocí matice R, tzv. *kovariantně*. Objekty s indexem nahoře se transformují pomocí R^{-1}, tzv. *kontravariantně*.

Transformační vztahy pro vektor a kovektor je možné také ověřit pomocí nezávislosti výpočtu hodnoty lineární formy α na vektoru v na použitých souřadnicích. Pokud by transformační vztah pro vektory byl $[v]^{B'} = Q[v]^B$ a pro kovektory $[\alpha]_{B'} = [\alpha]_B P$, pak je ze vztahu

$$\alpha(v) = [\alpha]_{B'}[v]^{B'} = [\alpha]_B PQ[v]^B = [\alpha]_B[v]^B$$

možné vyvodit, že $PQ = E$. Tedy pokud $[v]^{B'} = R^{-1}[v]^B$, pak $[\alpha]_{B'} = [\alpha]_B R$.

Definice 62. Nechť X, Y jsou množiny a $f : X \to \mathbb{F}$, $g : Y \to \mathbb{F}$ dvě funkce na těchto množinách. Jejich *tenzorovým součinem* rozumíme funkci

$$\begin{aligned} f \otimes g : X \times Y &\to \mathbb{F} \\ : (x, y) &\mapsto f(x)g(y) \end{aligned}$$

Tenzorový součin není komutativní ($f(x)g(y) \neq f(y)g(x)$, dokonce to ani nemusí být definováno, pokud $X \neq Y$), ale je asociativní:

$$((f \otimes g) \otimes h)(x, y, z) = f(x)g(y)h(z) = (f \otimes (g \otimes h))(x, y, z)$$

Má tedy smysl psát prostě $f \otimes g \otimes h$. Platí také

$$\begin{aligned} ((r_1 f_1 + r_2 f_2) \otimes g)(x, y) = r_1 f_1(x)g(y) + r_2 f_2(x)g(y) = \\ = r_1(f_1 \otimes g)(x, y) + r_2(f_2 \otimes g)(x, y) \end{aligned}$$

a podobně v druhé složce, je tedy bilineární, nebo při rozšíření na více než dva činitele multilineární.

Tenzorový součin dvou lineárních forem $\alpha : V \to \mathbb{F}$ a $\beta : V \to \mathbb{F}$ je bilineární forma $\alpha \otimes \beta$, která splňuje

$$(\alpha \otimes \beta)(v, w) = \alpha(v)\beta(w)$$

Jsou-li $\alpha = e^i$, $\beta = e^j$ prvky báze B^*, pak

$$(e^i \otimes e^j)(v, w) = v^i w^j$$

Je-li $H \in \mathbb{F}^{n \times n}$ matice, pak

$$h := (h_{ij} e^i \otimes e^j)(v, w) = h_{ij} v^i w^j = ([v]^B)^T H [w]^B$$

je bilineární forma, jejíž matice $[h]_B = (h(e_i, e_j))$ vzhledem k bázi B dle definice 40 je rovna H. Poprvé vidíme výraz, kde jsou dvě dvojice indexů, přes které se sčítá. Narozdíl od matice přechodu, u níž nás konvence „donutila" psát řádkový index jako horní a sloupcový jako dolní, u matice bilineární formy musíme psát oba indexy dole. Bilineární forma typu $\alpha \otimes \beta$ má matici hodnosti 1:

$$[\alpha \otimes \beta]_B = ((\alpha \otimes \beta)(e_i, e_j))_{i,j=1}^n = (\alpha(e_i)\beta(e_j))_{i,j=1}^n = [\alpha]_B^T [\beta]_B$$

Podobně bychom mohli zavést trilineární formu $\alpha \otimes \beta \otimes \gamma$, bázové trilineární formy $e^i \otimes e^j \otimes e^k$ a jejich lineární kombinace $T_{ijk} e^i \otimes e^j \otimes e^k$. Přejděme rovnou k obecné definici:

Definice 63. Nechť V, W jsou dva vektorové prostory nad $\mathbb{F}$, $k \in \mathbb{N}$. Pokud pro zobrazení

$$T : \underbrace{V \times \ldots \times V}_{k} \to W$$

platí, že $\forall u, v, v_1, \ldots, v_k \in V$, $\forall \lambda, \mu \in \mathbb{F}$, $\forall i \in \{1, \ldots, k\}$

$$T(v_1, \ldots, v_{i-1}, \lambda u + \mu v, v_{i+1}, \ldots, v_k) =$$
$$= \lambda T(v_1, \ldots, u, \ldots, v_k) + \mu T(v_1, \ldots, v, \ldots, v_k),$$

nazýváme ho *multilineárním* nebo specifičtěji *k-lineárním zobrazením.* Pokud $W = \mathbb{F}$, mluvíme o *multilineární formě.*

Tenzorový součin k lineárních forem je k-lineární forma. Množinu všech k-lineárních forem na vektorovém prostoru V označme symbolem $T_k(V)$. Pak tenzorový součin definuje také zobrazení

$$\otimes : T_p(V) \times T_q(V) \to T_{p+q}(V)$$

Lemma 16. *Nechť $B = \{e_1, \ldots, e_n\}$ je báze V. Označme*

$$e^{i_1 \ldots i_q} := \underbrace{e^{i_1} \otimes \ldots \otimes e^{i_q}}_{q} \in T_q(V)$$

Množina

$$(B^*)^q := \{e^{i_1 \ldots i_q} | i_1, \ldots, i_q \in \{1, \ldots, n\}\}$$

tvoří bázi prostoru $T_q(V)$ a $\forall T \in T_q(V)$ platí

$$T = T_{i_1 \ldots i_q} e^{i_1 \ldots i_q},$$

kde

$$T_{i_1 \ldots i_q} = T(e_{i_1}, \ldots, e_{i_q})$$

jsou souřadnice T vzhledem k $(B^*)^q$. Pokud $B' = \{e'_1, \ldots, e'_n\}$ a R je matice přechodu od B k B', pak

$$e'^{i_1 \ldots i_q} = (R^{-1})^{i_1}_{j_1} \ldots (R^{-1})^{i_q}_{j_q} e^{j_1 \ldots j_q}$$
$$T'_{i_1 \ldots i_q} = R^{j_1}_{i_1} \ldots R^{j_q}_{i_q} T_{j_1 \ldots j_q}$$

Důkaz. Podle definice tenzorového součinu

$$e^{i_1\dots i_q}(v,\dots,w) = v^{i_1}\dots w^{i_q}$$

Pak ale

$$T(v,\dots,w) = T(v^{i_1}e_{i_1},\dots,w^{i_q}e_{i_q}) = T_{i_1\dots i_q}v^{i_1}\dots w^{i_q} = (T_{i_1\dots i_q}e^{i_1\dots i_q})(v,\dots,w)$$

Vztah platí pro libovolnou q-tici vektorů $v,\dots,w$, takže $T = T_{i_1\dots i_q}e^{i_1\dots i_q}$. Odtud zároveň plyne, že $(B^*)^q$ generuje $T_q(V)$. Pokud by existovala čísla $S_{i_1\dots i_q}$, pro něž $S_{i_1\dots i_q}e^{i_1\dots i_q} = 0$, pak po dosazení vektorů $e_{j_1},\dots,e_{j_q}$ do levé strany plyne

$$(S_{i_1\dots i_q}e^{i_1\dots i_q})(e_{j_1},\dots,e_{j_q}) = S_{i_1\dots i_q}\delta^{i_1}_{j_1}\dots\delta^{i_q}_{j_q} = S_{j_1\dots j_q} = 0,$$

čili všechny koeficienty musí být nulové a $(B^*)^q$ je také lineárně nezávislá.

Multilinearita tenzorového součinu dává

$$\begin{aligned} e'^{i_1\dots i_q} &\equiv e'^{i_1}\otimes\dots\otimes e'^{i_q} = \\ &= (R^{-1})^{i_1}_{j_1}e^{j_1}\otimes\dots\otimes(R^{-1})^{i_q}_{j_q}e^{j_q} = (R^{-1})^{i_1}_{j_1}\dots(R^{-1})^{i_q}_{j_q}e^{j_1\dots j_q}\end{aligned}$$

a poslední tvrzení plyne z

$$T'_{i_1\dots i_q} = T(e'_{i_1},\dots,e'_{i_q}) = T(R^{j_1}_{i_1}e_{j_1},\dots,R^{j_q}_{i_q}e_{j_q}) = R^{j_1}_{i_1}\dots R^{j_q}_{i_q}T_{j_1\dots j_q}$$

□

Na základě tvrzení budeme sadu n^q čísel $(T(e_{i_1},\dots,e_{i_q}))$ označovat jako *reprezentaci multilineární formy* $[T]_B$. Je to vlastně zobecněná matice s více než dvěma indexy, q-rozměrná mřížka čísel.

Pokud $T_{ab\dots k}$ a $S_{li\dots t}$ jsou souřadnice $T\in T_p(V)$ a $S\in T_q(V)$ vůči M, pak

$$(T\otimes S)_{ab\dots t} = T_{ab\dots k}S_{li\dots t}$$

jsou souřadnice $T\otimes S\in T_{p+q}(V)$ vůči stejné bázi, jak je lehce vidět z lemmatu.

Příklad 39. Souřadnice lineární formy α se transformují podle vztahu $\alpha'_a = R^r_a\alpha_r$, maticově $[\alpha]_{B'} = [\alpha]_B R$. Pokud g je bilineární forma, má transformační vztah

$$g'_{ab} = R^r_a R^s_b g_{rs}$$

maticovou podobu $G' = R^TGR$, kde interpretujeme souřadnice $G = (g_{ab})$ jako matici bilineární formy vzhledem k B. Pro trilineární formu T můžeme interpretovat souřadnice T_{abc} buď jako $n\times n\times n$ krychličku čísel, nebo jako řádkový vektor matic

$$(T_{1bc},T_{2bc},\dots,T_{nbc}) =: (((T_1)_{bc}),((T_2)_{bc}),\dots,((T_n)_{bc}))$$

Transformační vztah

$$T'_{abc} = R^r_aR^s_bR^t_cT_{rst}$$

se pak dá přepsat jako

$$(T'_1,\dots,T'_n) = \left(\sum_{i=1}^n r_{i1}R^TT_iR, \sum_{i=1}^n r_{i2}R^TT_iR,\dots,\sum_{i=1}^n r_{in}R^TT_iR\right)$$

Samozřejmě nás nic nenutí vzít jako „maticové“ indexy zrovna ty dva poslední a jako „vektorový“ ten první. Zavedení matic T_i je jenom početní a notační pomůcka, což je zdůrazněno i tím, že jsme v posledním vztahu nepoužili sumační konvenci a zapsali elementy r_{ij} matice R tak, jak jsme zvyklí z dřívějška.

Definice 64 (Kanonický izomorfismus). Nechť V je vektorový prostor konečné dimenze nad $\mathbb{F}$. Pro každý $v \in V$ definujme homomorfismus $f_v : V^* \to \mathbb{F}$ předpisem $f_v(\alpha) := \alpha(v)$. Zobrazení

$$\begin{aligned} \Phi : V &\to (V^*)^* \\ &: v \mapsto f_v \end{aligned}$$

nazýváme *kanonický izomorfismus* vektorového prostoru V a jeho druhého duálu.

Ověříme korektnost definice. Protože $\forall \alpha, \beta \in V^*$, $\forall r, s \in \mathbb{F}$ platí

$$f_v(r\alpha + s\beta) = (r\alpha + s\beta)(v) = r\alpha(v) + s\beta(v) = rf_v(\alpha) + sf_v(\beta),$$

je f_v je lineární forma na V^*, tedy prvek prostoru $\mathrm{Hom}(V^*, \mathbb{F}) \equiv (V^*)^*$.

Zobrazení Φ tedy skutečně zobrazuje do prostoru specifikovaného v definici, ověřme dále, že je lineární. Protože $\forall v, w \in V$, $\forall r, s \in \mathbb{F}$, $\forall \alpha \in V^*$ platí

$$\begin{aligned} [\Phi(rv + sw)](\alpha) &= f_{rv+sw}(\alpha) = \alpha(rv + sw) \\ &= r\alpha(v) + s\alpha(w) = rf_v(\alpha) + sf_w(\alpha) \\ &= r[\Phi(v)](\alpha) + s[\Phi(w)](\alpha) \end{aligned}$$

Zobrazení $\Phi(rv + sw)$ a $r\Phi(v) + s\Phi(w)$ se rovnají pro všechna $\alpha \in V^*$, jsou tedy totožná.

Dále ověříme, že Φ je prosté. Podle definic

$$\begin{aligned} \mathrm{Ker}\,\Phi &= \{v \in V | \Phi(v) = 0\} = \{v \in V | \forall \alpha \in V^*, f_v(\alpha) = 0\} \\ &= \{v \in V | \forall \alpha \in V^*, \alpha(v) = 0\} \end{aligned}$$

Pro každý nenulový vektor v ale existuje lineární forma α, pro kterou $\alpha(v) \neq 0$. Stačí například zvolit doplněk W prostoru $\langle v \rangle$ ve V a definovat $\alpha(u + rv) = r$ pro libovolné $u \in W$ a $r \in \mathbb{F}$ (ověřte sami, že takto definované α je lineární forma). Tedy $\mathrm{Ker}\,\Phi$ musí být nulový podprostor.

Protože $\dim V = \dim V^* = \dim(V^*)^*$, musí být (díky větě o dimenzi jádra a obrazu) Φ izomorfismus.

Přívlastek „kanonický“ se v matematice užívá ve významu nejjednodušší, nejpřirozenější, závislý na co nejméně rovnocenných volbách. Skutečně je těžké si představit mezi izomorfismy prostorů V a $(V^*)^*$ nějaký, který by šel vyjádřit jednodušším předpisem. Pro obecnou dvojici vektorových prostorů V a W stejné dimenzi bychom našli zhruba tolik izomorfismů, kolik je dvojic (B, C), kde B je báze V a C je báze W, ale žádný z nich nemůžeme nazvat kanonickým, protože žádná dvojice bází není lepší nebo přirozenější než všechny ostatní. V případě prostorů V a jeho duálu V^* bychom mohli zkusit vektoru $v \in V$ přiřadit kovektor $\alpha \in V^*$, jehož souřadnice $[\alpha]^{B^*}$ jsou rovny $[v]^B$. Takový izomorfismus je jednodušší, než kdybychom volili ve V a ve V^* bázi nezávisle, pořád ale na volbě báze B závisí a mezi všemi možnostmi, jak B zvolit, žádná kanonická obecně není. Až v případě V a jeho druhého duálu $(V^*)^*$ máme možnost zvolit izomorfismus skutečně kanonicky, tj. v tomto případě nezávisle na volbě báze ve V. Jistě je pořád možné se přít, jestli by zobrazení Φ nešlo nahradit například zobrazením -2Φ nebo $137{,}035999084\Phi$, ale většina z nás patrně bude považovat za nejjednodušší původní variantu. V nekonečné dimenzi je zobrazení Φ obecně pouze monomorfismus, jemuž se říká *kanonické vnoření* V do $(V^*)^*$.

Kanonický izomorfismus nám umožňuje interpretovat vektory jako lineární formy na kovektorech a na základě toho definovat prostor $T^q(V)$ všech q-lineárních forem na kovektorech, kterým budeme říkat *multivektory.* Zcela analogicky jako v lemmatu 16 můžeme formulovat a dokázat

LEMMA 17. *Nechť $B = \{e_1, \ldots, e_n\}$ je báze V. Označme*

$$e_{i_1 \ldots i_q} := \underbrace{e_{i_1} \otimes \ldots \otimes e_{i_q}}_{q} \in T^q(V)$$

Množina

$$B^q := \{e_{i_1 \ldots i_q} | i_1, \ldots, i_q \in \{1, \ldots, n\}\}$$

tvoří bázi prostoru $T^q(V)$ a $\forall T \in T^q(V)$ platí

$$T = T^{i_1 \ldots i_q} e_{i_1 \ldots i_q},$$

kde

$$T^{i_1 \ldots i_q} = T(e^{i_1}, \ldots, e^{i_q})$$

jsou souřadnice T vzhledem k B^q. Pokud $B' = \{e'_1, \ldots, e'_n\}$ a R je matice přechodu od B k B', pak

$$e'_{i_1 \ldots i_q} = R^{j_1}_{i_1} \ldots R^{j_q}_{i_q} e_{j_1 \ldots j_q}$$
$$T'^{i_1 \ldots i_q} = ((R)^{-1})^{i_1}_{j_1} \ldots ((R)^{-1})^{i_q}_{j_q} T^{j_1 \ldots j_q}$$

Sadu n^q čísel $T(e^{i_1}, \ldots, e^{i_q})$ budeme opět značit $[T]_B$ a nazývat reprezentací multivektoru.

PŘÍKLAD 40. Souřadnice bivektoru $g = g^{ab} e_a \otimes e_b \in T^2(V)$ se transformují podle předpisu

$$g'^{ij} = (R^{-1})^i_a (R^{-1})^j_b g^{ab}$$

Definujeme-li matici bivektoru jako $(G)_{ij} := g^{ij}$, pak se dá transformační vztah zapsat jako

$$G' = R^{-1} G (R^{-1})^T$$

Souřadnice trivektoru $T = T^{abc} e_a \otimes e_b \otimes e_c$lze chápat jako $n \times n \times n$ krychličku, nebo jako řádek matic $(T_1, \ldots, T_n)$, kde $(T_i)_{jk} := T^{ijk}$. Pak je $(T'_1, \ldots, T'_n)$ rovno $\left(\sum_{i=1}^n (R^{-1})_{1i} R^{-1} T_i (R^{-1})^T, \ldots, \sum_{i=1}^n (R^{-1})_{ni} R^{-1} T_i (R^{-1})^T\right)$

Multilineární formy označujeme také jako *kovariantní tenzory* a multivektory jako *kontravariatní tenzory.* Nejobecnější typ tenzoru je kombinací těchto dvou typů:

DEFINICE 65. Nechť V je vektorový prostor nad $\mathbb{F}$. Symbolem $T^p_q(V)$ označíme množinu všech zobrazení

$$T : \underbrace{V^* \times \ldots \times V^*}_{p} \times \underbrace{V \times \ldots \times V}_{q} \to \mathbb{F},$$

která jsou lineární v každém argumentu. Mluvíme o množině *p-krát kontravariantních a q-krát kovariantních tenzorů* na V nebo též o *tenzorech typu (p, q)*. Číslo $p + q$ se označuje jako *stupeň tenzoru. Souřadnicemi* či *reprezentací* $[T]_B$ tenzoru $T \in T^p_q(V)$ vzhledem k bázi $B = (e_i)^n_1$ ve V rozumíme soubor čísel

$$T^{i_1 \ldots i_p}_{j_1 \ldots j_q} := T(e^{i_1}, \ldots, e^{i_p}, e_{j_1}, \ldots, e_{j_q})$$

O dolních indexech hovoříme jako o *indexech kovariantních* a horní se nazývají *indexy kontravariantními.*

Na smíšených tenzorech máme definován tenzorový součin

$$\otimes : T^p_q(V) \times T^k_l(V) \to T^{p+k}_{q+l}(V)$$

a reprezentace vzhledem k B je tvořena ve skutečnosti souřadnicemi vzhledem k bázi $B^p \times (B^*)^q$ v $T^p_q(V)$:

$$T = T^{i_1 \ldots i_p}_{j_1 \ldots j_q} e_{i_1} \otimes \ldots \otimes e_{i_p} \otimes e^{j_1} \otimes \ldots \otimes e^{j_q} \equiv T^{i_1 \ldots i_p}_{j_1 \ldots j_q} e^{j_1 \ldots j_q}_{i_1 \ldots i_p}$$

$$T^p_q(V) \equiv \underbrace{V \otimes \ldots \otimes V}_{p} \otimes \underbrace{V^* \otimes \ldots \otimes V^*}_{q},$$

které vyjadřuje, že libovolný tenzor typu (p, q) lze zkonstruovat jako lineární kombinaci tenzorového součinu p vektorů a q kovektorů.

Příklad 41. Prostor $T^1_1(V) \equiv V \otimes V^*$ je přirozeně ztotožněn s prostorem $\mathrm{End}(V)$ všech endomorfismů na V. Prvek $v \otimes \alpha$, kde $v \in V$ a $\alpha \in V^*$, definuje lineární zobrazení $f_{v\otimes\alpha} : V \to V$ předpisem

$$f_{v\otimes\alpha}(u) := \alpha(u)v$$

Odpovídá to vlastně částečnému dosazení do druhého argumentu zobrazení

$$v \otimes \alpha : V^* \times V \to \mathbb{F}$$

Báze $\{e_i \otimes e^j | i, j \in \{1, \ldots, n\}\}$ v $T^1_1(V)$ je v tomto ztotožnění složena z endomorfismů majících na bázi B hodnoty

$$(e_i \otimes e^j)(e_k) \equiv f_{e_i \otimes e^j}(e_k) = e_i \delta^j_k,$$

neboli matice endomorfismu $[e_i \otimes e^j]^B_B$ má na pozici ij jedničku a všude jinde nuly. Endomorfismus $T := T^i_j e_i \otimes e^j$ má na bázi hodnoty

$$T(e_k) = (T^i_j e_i \otimes e^j)(e_k) = T^i_j e_i \delta^j_k = T^i_k e_i,$$

jeho matice vzhledem k B je proto T^i_k, kde horní index chápeme jako řádkový a dolní index sloupcový. Všimněte si, že to je stejná konvence jako pro matice přechodu. To není překvapivé, vzhledem k tomu, že matice přechodu se dá definovat jako speciální případ matice endomorfismu. Odtud je také jasně vidět, že přiřazení $T^1_1(V)$ a $\mathrm{End}(V)$ je opravdu izomorfismus. Kroneckerovo delta δ^i_k je pak matice identického endomorfismu, která je vůči kterékoli bázi stejná, rovná jednotkové matici:

$$(1_V)(e_k) = \delta^i_k e_i = e_k$$

Věta 43. *Nechť V je vektorový prostor, $B = \{e_1, \ldots, e_n\}, B' = \{e'_1, \ldots, e'_n\}$ dvě báze v něm, $T \in T^p_q(V)$. Pak*

$$T'^{i_1 \ldots i_p}_{j_1 \ldots j_q} = (R^{-1})^{i_1}_{a_1} \ldots (R^{-1})^{i_p}_{a_p} R^{b_1}_{j_1} \ldots R^{b_q}_{j_q} T^{a_1 \ldots a_p}_{b_1 \ldots b_q}$$

Důkaz spočívá jen v dosazení vztahů $e'_j = R^b_j e_b$ a $e'^i = (R^{-1})^i_a e^a$ do definice souřadnic.

Příklad 42. Smíšený tenzor f typu $(1, 1)$ se transformuje předpisem $f'^i_j = (R^{-1})^i_a R^b_j f^a_b$. Pro matici F s elementy $(F)_{ab} := f^a_b$ z toho dostáváme

$$F' = R^{-1} F R$$

tedy vlastně standardní vztah pro změnu matice endomorfismu při změně báze.

Souřadnice smíšeného tenzoru $T = T^{ij}_k e_i \otimes e_j \otimes e^k$ typu $(2, 1)$ můžeme zapsat například jako n-tici matic s elementy $(T_i)_{jk} := T^{ij}_k$. Pak

$$T'_a = \sum_{i=1}^{n} (R^{-1})_{ai} (R^{-1} T_i R)$$

Pro tenzor typu $(1, 2)$ se souřadnicemi zapsanými maticemi $(T_i)_{jk} := T^j_{ik}$ máme

$$T'_a = \sum_{i=1}^{n} (R^T)_{ai} (R^{-1} T_i R)$$

KAPITOLA 22

Tenzory podruhé

Pokud do k-tého kovektorového argumentu tenzoru

$$T : \underbrace{V^* \times \ldots \times V^*}_{p} \times \underbrace{V \times \ldots \times V}_{q} \mapsto \mathbb{R},$$

z $T^p_q(V)$ dosadíme pevný kovektor α, vzniklé zobrazení

$$\tilde{T} := T(\cdot, \ldots, \cdot, \alpha, \cdot, \ldots, \cdot)$$

je lineární ve všech zbylých argumentech označených tečkami, je to tedy tenzor typu $(p-1, q)$. Podobně dosazením vektoru $v \in V$ do některého z posledních q argumentů T dostaneme tenzor typu $(p, q-1)$.

Vezmeme-li nyní v roli T bilineární formu $g \in T^0_2(V)$, pak dosazením $v \in V$ do druhého argumentu získáme lineární formu

$$\begin{aligned} g(\cdot, v) :& V \to \mathbb{F} \\ & u \mapsto g(u, v) \end{aligned}$$

Dosazením stejného vektoru do prvního argumentu dostaneme obecně jinou lineární formu $g(v, \cdot)$. Vzhledem k bázi je $g(u, v) = g_{ij} u^i v^j$, souřadnice obou lineárních forem tedy budou

$$\begin{aligned} [g(\cdot, v)]_B &= (g_{1j} v^j, \ldots, g_{nj} v^j) \\ [g(v, \cdot)]_B &= (g_{i1} v^i, \ldots, g_{in} v^i) \end{aligned}$$

Podobně z tenzoru $T := T^i_j e_i \otimes e^j \in T^1_1(V)$ dostaneme dosazením $v \in V$ vektor $T(\cdot, v) := T^i_j e_i e^j(v) = T^i_j v^j e_i$ se souřadnicemi

$$[T(\cdot, v)]^B := (T^1_j v^j, \ldots, T^n_j v^j)^T$$

Každému $T \in T^1_1(V)$ tedy lze přiřadit operátor $\mathbb{T}$ předpisem

$$\mathbb{T}v := T(\cdot, v),$$

který je vzhledem k bázi B reprezentován maticí $[\mathbb{T}]^B_B = (T^i_j)^n_{i,j=1} \equiv [T]_B$. Toto přiřazení jsme v předchozí kapitole nazývali přirozený izomorfismus mezi $T^1_1(V)$ a $\operatorname{End} V$.

Tvrzení 62. *Nechť g je regulární bilineární forma na V. Pak zobrazení*

$$\begin{aligned} \flat_g :& V \to V^* \\ & v \mapsto g(\cdot, v) \end{aligned}$$

je izomorfizmus prostorů V a V^*.

Důkaz. Vzhledem tomu, že V a V^* mají stejnou dimenzi, stačí ukázat, že $\flat_g$ je prosté. Nechť $g(\cdot, v)$ je nulový kovektor, pak pro všechna $u \in V$ je $g(u, v) = 0$. Vyjádříme g vzhledem k bázi $B = (e_1, \ldots, e_n)$ a dosadíme za $u = e_i$, pak

$$g(u, v) = g(e_i, v^j e_j) = v^j g_{ij} = 0$$

Protože (g_{ij}) je regulární matice, musí být všechna v^j nulová a tedy i $v = 0$. $\square$

Nejčastěji se s tímto izomorfizmem setkáváme, když V je reálný vektorový prostor a $g \equiv g_{ab}e^a \otimes e^b$ symetrická bilineární forma na něm, pak jsou kovektory $g(\cdot, v)$ a $g(v, \cdot)$ totožné. V souřadnicích máme

$$\flat_g v \equiv (\flat_g v)_i e^i = (g_{ij} e^i \otimes e^j)(\cdot, v) = g_{ij} v^j e^i,$$

tedy kovektor $\flat_g v$ má souřadnice $(\flat_g v)_i = g_{ij} v^j$. Proto se zobrazení $\flat_g$ nazývá *spuštění indexu* a označuje odpovídajícím hudebním symbolem.

Pokud g je skalární součin a tedy (V, g) bereme jako unitární prostor, pak se předpokládá, že spuštění indexu přísluší skalárnímu součinu (*metrickému tenzoru*) g a souřadnice kovektoru $\flat_g v$ bývají označeny místo $(\flat_g v)_i = g_{ij} v^j$ pouze v_i. Ze symetrie g a interpretace v jako prvku $(V^*)^*$ pak

$$(\flat_g v)(u) = g(u, v) = g(v, u) = v(g(\cdot, u)) = (v \circ \flat_g)(u)$$

Zobecněním rovnosti $\flat_g v = v \circ \flat_g$ můžeme definovat spouštění indexů pro libovolný tenzor $T^p_q(V)$ s $p > 0$. Například pro $T \in T^3_1(V)$ je $T(\flat_g \cdot, \cdot, \flat_g \cdot, \cdot)$ tenzor typu $(1, 3)$ vzniklý spuštěním prvního a třetího indexu. Jeho souřadnice jsou

$$T_a{}^b{}_{cd} := T(\flat_g e_a, e^b, \flat_g e_c, e_d) = T(g_{ai} e^i, e^b, g_{cj} e^j, e_d) = g_{ai} g_{cj} T^{ibj}{}_d$$

Píšeme $T_a{}^b{}_{cd}$ místo T^b_{acd}, abychom dokázali rozlišit, které indexy byly spuštěny a které ne.

Definujme bivektor $g^{-1} := g^{ab} e_a \otimes e_b$, jehož souřadnice splňují $g^{ac} g_{cb} = \delta^a_b$, čili matice skalárního součinu g a bivektoru g^{-1} jsou navzájem inverzní. Dosazením transformačních vztahů pro bilineární formy a bivektory snadno odvodíme, že tato definice nezávisí na volbě báze. Vůči ortonormální bázi $B = (e_a)^n_1$ je i g^{ab} jednotková matice a tedy g^{-1} je pozitivně definitní symetrická bilineární forma na V^*, tzv. *duální metrický tenzor*. Zobrazení

$$\begin{aligned} \sharp_g :& V^* \to V \\ & \alpha \mapsto g^{-1}(\cdot, \alpha), \end{aligned}$$

které v souřadnicích nabývá tvaru

$$\sharp_g \alpha \equiv (\sharp_g \alpha)^i e_i = (g^{ij} e_i \otimes e_j)(\cdot, \alpha) = g^{ij} \alpha_j e_i,$$

tj. $(\sharp_g \alpha)^i = g^{ij} \alpha_j$, nazýváme *zdvižení indexu*. Protože

$$(\sharp_g \flat_g v)^i = g^{ij} g_{jk} v^k = \delta^i_k v^k = v^i,$$

je $\sharp_g$ inverzní izomorfizmus k $\flat_g$. Podobně jako u $\flat_g$ pak $\sharp_g \alpha = \alpha \circ \sharp_g$ a zdvižení indexu lze aplikovat i na tenzory vyššího stupně.

Můžete se divit, proč se vůbec zaobíráme situací, kdy (g_{ab}) není jednotková matice, když ortonormální bázi lze zavést pro každý skalární součin. Existují tři oblasti, kvůli nimž je důležité naučit se počítat s obecným metrickým tenzorem. Jsou to křivočaré systémy souřadnic, kde se metrický tenzor mění v prostoru (je to vlastně tenzorové pole), speciální teorie relativity, kde není pozitivně definitní (mluvíme o *pseudometrickém tenzoru*), a analytická mechanika, která se popisuje pomocí pojmů symplektické geometrie, v níž je „skalární součin“ antisymetrický. Naše formulace se dá snadno adaptovat i na tyto tři odlišné situace.

Zpět k operátorům. Stopa $\mathbb{T} \in \mathrm{End}(V)$ je rovna

$$\mathrm{Tr}\, \mathbb{T} = \sum_{i=1}^{n} ([\mathbb{T}]^B_B)_{ii} \equiv T^i_i = (T^a_b e_a \otimes e^b)(e^i, e_i) \equiv T(e^i, e_i)$$

Na volbě báze B nezáleží, protože

$$T(e'^a, e'_a) = T((R^{-1})^a_c e^c, R^b_a e_b) = \delta^b_c T(e^c, e_b) = T(e^b, e_b)$$

a Tr můžeme chápat jako zobrazení z $T^1_1(V)$ do skalárů $T^0_0(V) := \mathbb{F}$. Konstrukci lze zobecnit:

Definice 66. Nechť V je vektorový prostor, $(e_i)_1^n$ je báze V, $p > 0$, $q > 0$, $a \in \{1, \ldots, p\}$, $b \in \{1, \ldots, q\}$. Zobrazení

$$\begin{aligned} C_{ab} : T_q^p(V) &\to T_{q-1}^{p-1}(V) \\ T &\mapsto T(\ldots, \underset{a}{e^i}, \ldots, \underset{b}{e_i}, \ldots) \end{aligned}$$

se nazývá *kontrakce* tenzoru (někdy též *zúžení* nebo *stopa*) přes a-tý kontravariantní a b-tý kovariantní index.

Na unitárním prostoru lze zkombinovat kontrakci a zdvihání/spouštění indexu. Pokud g je metrický tenzor a např. $T \in T_3^0(V)$ trilineární forma, pak

$$T_{ij}{}^i e^j = g^{ik} T_{ijk} e^j = C_{11} T(\cdot, \cdot, \sharp_g \cdot) \in T_1^0(V)$$

Ze symetrie g plyne, že $T^i{}_{ji} = T_{ij}{}^i$, tedy kovektor výše lze zapsat i jako $C_{12}T(\sharp_g \cdot, \cdot, \cdot)$. Kovektory

$$T_i{}^i{}_j e^j \equiv C_{11} T(\cdot, \sharp_g \cdot, \cdot) \text{ nebo } T_j{}^i{}_i e^j \equiv C_{12} T(\cdot, \sharp_g \cdot, \cdot)$$

se ale od něj obecně liší.

Pokud $v \in V$, $\alpha \in V^*$, lze hodnotu $\alpha(v) = \alpha_i v^i$ chápat jako kontrakci $C_{11}(v \otimes \alpha)$ tenzorového součinu. Analogicky lze pomocí tenzorového součinu a kontrakcí zapsat hodnotu libovolného tenzoru stupně $p + q$ na (ko-)vektorech $\alpha, \ldots, \beta, v, \ldots, w$:

$$T(v, \ldots, w) = \underbrace{(C \circ \ldots \circ C)}_{p+q}(T \otimes v \otimes \ldots \otimes w),$$

Definice 67. Nechť $T \in T_q^0$. Pak definujeme *úplnou symetrizaci*, resp. *úplnou antisymetrizaci* tenzoru T předpisem $\forall v_1, \ldots, v_q \in V$

$$\begin{aligned} [\pi_S(T)](v_1, \ldots, v_q) &:= \frac{1}{q!} \sum_{\rho \in S_q} T(v_{\rho(1)}, \ldots, v_{\rho(q)}), \text{ resp.} \\ [\pi_A(T)](v_1, \ldots, v_q) &:= \frac{1}{q!} \sum_{\rho \in S_q} \operatorname{sgn}(\rho) T(v_{\rho(1)}, \ldots, v_{\rho(q)}) \end{aligned}$$

Snadno se ověří, že

$$\begin{aligned} \pi_A \circ \pi_A &= \pi_A \\ \pi_S \circ \pi_S &= \pi_S \\ \pi_A \circ \pi_S &= \pi_S \circ \pi_A = 0 \end{aligned}$$

Obě zobrazení jsou tedy projekce na dva podprostory, podprostor úplně symetrických kovariantních tenzorů, který se obvykle značí $S_q(V)$ nebo $S^q(V^*)$, a podprostor úplně antisymetrických tenzorů značený $\Lambda_q(V) \equiv \Lambda^q(V^*)$, v jejichž průniku je pouze nulový tenzor.

Pro $q = 2$ definice říká, že

$$\begin{aligned} [\pi_S(T)]_{ab} &= \frac{1}{2}(T_{ab} + T_{ba}) =: T_{(ab)} \\ [\pi_A(T)]_{ab} &= \frac{1}{2}(T_{ab} - T_{ba}) =: T_{[ab]} \end{aligned}$$

Zde jsme zavedli tradiční závorkovou notaci pro symetrizaci a antisymetrizaci indexů. Zjevně platí $T_{(ab)} = T_{(ba)}$ a $T_{[ab]} = -T_{[ba]}$, tedy matice $(T_{(ab)})$ je symetrická a matice $(T_{[ab]})$ antisymetrická. Vlastnost $T_{ab} = T_{(ab)} + T_{[ab]}$ znamená, že každý tenzor typu $(0, 2)$ je možné (jednoznačně) rozložit na jeho symetrickou a antisymetrickou část

$$
\begin{aligned}
T = T_{ab}e^a \otimes e^b &= T_{(ab)}e^a \otimes e^b + T_{[ab]}e^a \otimes e^b \\
&= T_{(ab)}\frac{1}{2}(e^a \otimes e^b + e^b \otimes e^a) + T_{[ab]}\frac{1}{2}(e^a \otimes e^b - e^b \otimes e^a) \\
&= T_{(ab)}e^{(ab)} + T_{[ab]}e^{[ab]}
\end{aligned}
$$

Lineárně nezávislých tenzorů $e^{[ab]}$ je právě tolik, kolik je dvouprvkových množin čísel z $\{1,\dots,n\}$, tedy $\binom{n}{2}$. Počet tenzorů $e^{(ab)}$ se zase rovná počtu dvouprvkových kombinací s opakováním z $\{1,\dots,n\}$, tedy $\binom{n+1}{2}$. Součet obou čísel je n^2, což je rovno dimenzi $T^0_2(V)$, jak jsme mohli předpokládat na základě věty o dimenzi spojení a průniku.

(Anti)symetrizaci indexů můžeme zavést i pro tenzory vyšších stupňů:

$$
\begin{aligned}
T_{(a\dots b)} &:= \frac{1}{q!}\sum_{\rho\in S_q} T(\rho(e_a,\dots,e_b)) \\
T_{[a\dots b]} &:= \frac{1}{q!}\sum_{\rho\in S_q} \operatorname{sgn}(\rho) T(\rho(e_a,\dots,e_b)),
\end{aligned}
$$

kde $\rho(e_a,\dots,e_b)$ znamená přeuspořádání posloupnosti vektorů $e_a,\dots,e_b$ permutací ρ. (Anti)symetrizací prvků báze $\{e^{a\dots b}\}$ v $T^0_q(V)$ dostáváme prvky $e^{(a\dots b)}$ a $e^{[a\dots b]}$, které tvoří bázi v $S^q(V^*)$ a $\Lambda^q(V^*)$. Je vidět, že

$$
\begin{aligned}
\dim S^q(V^*) &= \binom{n+q-1}{q}, \\
\dim \Lambda^q(V^*) &= \binom{n}{q}.
\end{aligned}
$$

Pro $q > n$ tedy už žádné netriviální úplně antisymetrické tenzory nejsou, zatímco prostor úplně symetrických tenzorů je nenulový pro všechna q. Pro $q > 2$ stále platí, že $S^q(V^*) \cap \Lambda^q(V^*) = 0$, ale narozdíl od případu $q = 2$ je

$$\dim S^q(V^*) + \dim \Lambda^q(V^*) < \dim T^0_q(V),$$

tedy existují tenzory, z nichž po odečtení úplně symetrické a úplně antisymetrické části ještě něco zbyde.

(Anti)symetrizaci můžeme úplně stejně zavést i pro kontravariantní tenzory, případně pro tenzory smíšené. U smíšených má ale smysl provádět ji přes indexy stejného typu.

Pomocí zdvihu a spuštění indexu je možné definovat kontrakci přes libovolné dva indexy: tenzor

$$\tilde{T} := C_{11}T(\flat_g\cdot,\cdot,\cdot)$$

má souřadnice

$$T_a{}^{ac} = g_{ai}T^{iac} = g_{ia}T^{iac} = T^a{}_a{}^c.$$

Pokud je T úplně symetrický, pak se všechny jeho kontrakce rovnají,

$$T_a{}^{ac} = g_{ai}T^{iac} = g_{ai}T^{cia} = T^c{}_a{}^a, \text{ apod.}$$

Je-li T úplně antisymetrický, jsou všechny jeho kontrakce nulové, neboť díky symetrii g_{ab} máme

$$T_a{}^{ac} = g_{ai}T^{iac} = g_{ai}(-T^{aic}) = -g_{ia}T^{aic} = -T_i{}^{ic}$$

KAPITOLA 23

Pozdní sběr

Tato kapitola obsahuje zajímavá a užitečná tvrzení, která bychom bývali zvládli formulovat a dokázat už dříve, ale zavedlo by nás to stranou od hlavního výkladu.

1. Cayley-Hamiltonova věta

Dosazením 2×2 matice do jejího charakteristického polynomu získáme pozoruhodnou rovnost:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^2 - (a+d)\begin{pmatrix} a & b \\ c & d \end{pmatrix} + (ad-bc)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Věta 44 (Cayley-Hamilton). *Nechť $A \in \mathbb{C}^{n\times n}$ a p_A její charakteristický polynom. Pak $p_A(A) = 0$.*

Důkaz. Nechť $\sigma(A) = \{\lambda_1, \ldots, \lambda_m\}$, $p_A(\lambda) = \pm\prod_{i=1}^m(\lambda-\lambda_i)^{k_i}$, $J = R^{-1}AR = \operatorname{diag}(J_{k_1}, \ldots, J_{k_m})$ Jordanův tvar A, kde J_{k_i} je $k_i \times k_i$ blok obsahující buňky s vlastním číslem λ_i. Protože

$$p_A(\lambda) = \det(A - \lambda E) = \det(RJR^{-1} - \lambda RR^{-1}) = \det(J - \lambda E) = p_J(\lambda),$$

platí $p_A(A) = p_J(A)$. Zároveň $A^q = RJ^qR^{-1}$ pro všechna $q \geq 0$, tedy i $p_J(A) = Rp_J(J)R^{-1}$. Protože $J_{k_i} - \lambda_i E$ má na blokové diagonále Jordanovy buňky s vlastním číslem nula velikosti menší nebo rovné k_i, platí $(J_{k_i} - \lambda_i E)^{k_i} = 0$. Charakteristický polynom má tvar $p_A(\lambda) = \prod_{i=1}^m(\lambda_i - \lambda)^{k_i}$, tedy

$$p_J(J) = (-1)^n \prod_{i=1}^m (J - \lambda_i E)^{k_i} = \begin{pmatrix} (J_{k_1} - \lambda_1 E)^{k_1} K & 0 \\ 0 & L \end{pmatrix} = 0,$$

kde K je matice $\prod_{i=2}^m(J_{k_1} - \lambda_i E)^{k_i}$. První blok na diagonále je tedy nulový, ale obdobnou úvahou musí i všechny ostatní bloky skryté v matici L být nulové, protože obsahují nulový činitel. $\square$

Z Cayley-Hamiltonovy věty plyne, že A^n je lineární kombinace matic $E, A, \ldots, A^{n-1}$, například pro 2×2 matice

$$A^2 = (\operatorname{Tr} A)A - (\det A)E$$

Lze odtud např. spočítat rekurentně libovolnou mocninu matice, aniž bychom potřebovali znát její spektrum. Pro A regulární lze takto vyjádřit i inverzní matici, například pro 2×2 matice

$$A^{-1} = \frac{1}{\det A}((\operatorname{Tr} A)E - A)$$

2. Současná diagonalizovatelnost

Nechť $\mathbb{A}, \mathbb{B} \in \operatorname{End}(V)$, $\mathbb{A}\mathbb{B} = \mathbb{B}\mathbb{A}$, $\lambda \in \sigma(\mathbb{A})$, $W \leq V$ vlastní podprostor $\mathbb{A}$ příslušný λ. Pak W je invariantní podprostor $\mathbb{A}$ a pro $v \in W$ platí

$$\mathbb{A}(\mathbb{B}v) = \mathbb{B}(\mathbb{A}v) = \mathbb{B}(\lambda v) = \lambda\mathbb{B}v \quad \Rightarrow \quad \mathbb{B}v \in W$$

Zúžený operátor $\mathbb{B}|_W$ má vlastní číslo μ a jemu příslušný vlastní vektor $w \in W$ je společný vlastní vektor komutujících operátorů $\mathbb{A}, \mathbb{B}$. Kdybychom takových vektorů našli $\dim V$, měli bychom bázi, vzhledem k níž jsou oba operátory reprezentovány diagonální maticí. Říkáme pak, že množina operátorů $\mathcal{A} = \{\mathbb{A}, \mathbb{B}\}$ je *současně diagonalizovatelná.* Obecně platí

VĚTA 45. *Nechť $\mathcal{A} \subset \operatorname{End} V$ je množina diagonalizovatelných operátorů. Pak $\mathcal{A}$ je současně diagonalizovatelná, právě když každé dva operátory v ní komutují.*

DŮKAZ. Implikace $\Rightarrow$ je snazší. Pokud $\mathbb{A}, \mathbb{B} \in \mathcal{A}$ a existuje báze C ve V taková, že $D_A := [\mathbb{A}]^C_C$ a $D_B := [\mathbb{B}]^C_C$ jsou diagonální, pak

$$[\mathbb{A}\mathbb{B} - \mathbb{B}\mathbb{A}]^C_C = [\mathbb{A}]^C_C[\mathbb{B}]^C_C - [\mathbb{B}]^C_C[\mathbb{A}]^C_C = D_A D_B - D_B D_A = 0$$

tedy $\mathbb{A}$ a $\mathbb{B}$ nutně komutují.

Druhou implikaci dokážeme indukcí podle $n = \dim V$. Operátory na prostoru V dimenze 1 jsou reprezentovány maticemi 1×1, které spolu komutují všechny a všechny jsou automaticky diagonální, takže není co dokazovat. Uvažujme $\mathcal{A}$ množinu komutujících diagonalizovatelných operátorů na V dimenze $n > 1$, zvolme ve V bázi B a označme $\mathcal{A}'$ množinu všech matic operátorů z $\mathcal{A}$ vzhledem k bázi B. Předpokládejme, že tyto $n \times n$ nejsou všechny násobkem jednotkové matice, jinak by nebylo co dokazovat. Vyberme $A \in \mathcal{A}'$, která není násobkem jednotkové matice a regulární matici R takovou, že $D := R^{-1}AR$ je diagonální. Pro každou matici $X \in \mathcal{A}'$ platí, že $Y := R^{-1}XR$ komutuje s D. Matici D můžeme psát v blokově diagonálním tvaru

$$\begin{pmatrix} \lambda_1 E_{n_1} & 0 & \dots & 0 \\ 0 & \lambda_2 E_{n_2} & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \dots & 0 & \lambda_m E_{n_m} \end{pmatrix},$$

kde λ_i jsou navzájem různá. Zapíšeme-li Y ve stejné blokové struktuře jako má D a porovnáme odpovídající bloky součinů DY a YD, je snadno vidět, že Y musí být také blokově diagonální

$$Y = \begin{pmatrix} Y_1 & 0 & \dots & 0 \\ 0 & Y_2 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \dots & 0 & Y_m \end{pmatrix},$$

kde $Y_i \in \mathbb{F}^{n_i \times n_i}$. Podle indukčního předpokladu existují regulární matice $S_i \in \mathbb{F}^{n_i \times n_i}$ takové, že pro všechny takto získané matice Y a

$$S := \begin{pmatrix} S_1 & 0 & \dots & 0 \\ 0 & S_2 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \dots & 0 & S_m \end{pmatrix}$$

je $S^{-1}YS$ diagonální matice. Pak pro všechna $X \in \mathcal{A}'$ a $Q := RS$ je $Q^{-1}XQ$ diagonální matice. $\square$

Větu můžeme formulovat i tak, že pokud množina matic komutuje, pak lze najít společnou bázi z vlastních vektorů. Je-li spektrum prosté, stačí spočítat vlastní vektory jedné matice a ty budou vlastními vektory i pro ostatní. Věta má význam v kvantové mechanice a kdekoli, kde jsou důležité symetrie.

3. Geršgorinovy kruhy

VĚTA 46 (Geršgorinovy kruhy). *Každé vlastní číslo $A \in \mathbb{C}^{n\times n}$ leží pro nějaké $i \in \{1, \dots, n\}$ v kruhu o středu a_{ii} a poloměru $\sum_{j\neq i} |a_{ij}|$.*

DŮKAZ. Pro $\lambda \in \sigma(A)$ zvolme $\mathbf{x} \in V_\lambda$ a i tak, že $x_i = 1$ a $\forall j \neq i$, $1 \geq |x_j|$. Pak $((A - \lambda E)\mathbf{x})_i = 0$ dává

$$|\lambda - a_{ii}| = \left|\sum_{j\neq i} a_{ij}x_j\right| \leq \sum_{j\neq i} |a_{ij}||x_j| \leq \sum_{j\neq i} |a_{ij}|$$

□

Matice je *diagonálně dominantní*, pokud $|a_{ii}| > \sum_{j\neq i} |a_{ij}|$. Geršgorinovy kruhy pak neobsahují počátek, tedy A musí být regulární. Další aplikací je odhad chyby a test ukončení iterativních algoritmů pro výpočet vlastních čísel.

Příkladem takového algoritmu je *QR-algoritmus*. Je-li $A =: A_0$ reálná čtvercová matice s QR-rozkladem $A_0 = QR$, definujme

$$A_1 := RQ = Q^T QRQ = Q^T A_0 Q = Q^{-1} A_0 Q$$

Protože A_0 a A_1 jsou podobné, mají stejné spektrum. Bez důkazu uveďme, že pokud je A symetrická, dá se iterací tohoto postupu získat posloupnost, která konverguje k diagonální matici, jejíž diagonální prvky jsou vlastní čísla A. Pro většinu ostatních matic konverguje algoritmus k matici blokově diagonální s bloky velikosti 1 a 2 (ty odpovídají dvojici komplexně sdružených kořenů). Příklad:

$$A = \begin{pmatrix} 3 & 1 & -1 \\ 1 & 3 & -1 \\ -1 & -1 & 5 \end{pmatrix} \overset{5 \text{ iterací}}{\longrightarrow} \begin{pmatrix} 5.9940 & 0.0284 & -0.1323 \\ 0.0284 & 2.0002 & -0.0009 \\ -0.1323 & -0.0009 & 3.0058 \end{pmatrix}$$

Z věty o Geršgorinových kruzích plyne odhad spektra

$$\sigma(A) = \{5.9940 \pm 0.1607, 2.0002 \pm 0.0293, 3.0058 \pm 0.1332\}$$

4. Gaussův integrál

O *Gaussově integrálu* $\int_{-\infty}^{\infty} e^{-x^2} dx$ je známo, že je roven $\sqrt{\pi}$. Odtud snadno pro $\lambda_1, \dots, \lambda_n > 0$

$$\int_{\mathbb{R}^n} e^{-\lambda_1 x_1^2 - \dots - \lambda_n x_n^2} dx_1 \dots dx_n = \sqrt{\frac{\pi^n}{\lambda_1 \dots \lambda_n}}$$

Pokud $A \in \mathbb{R}^{n\times n}$ je pozitivně definitní symetrická matice a U ortogonální matice, pro kterou $U^T AU = D = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$, pak

$$\int_{\mathbb{R}^n} e^{-\mathbf{x}^T A\mathbf{x}} dx_1 \dots dx_n = \int_{\mathbb{R}^n} e^{-\mathbf{x}^T UDU^T \mathbf{x}} dx_1 \dots dx_n$$

Protože $|\det U| = 1$, substituce $\mathbf{x}' = U^T\mathbf{x}$ zachovává objem, tedy $dx'_1 \dots dx'_n = dx_1 \dots dx_n =: d\mathbf{x}$. Protože $\det A = \det D$, dostáváme

$$\int_{\mathbb{R}^n} e^{-\mathbf{x}^T A\mathbf{x}} d\mathbf{x} = \int_{\mathbb{R}^n} e^{-\mathbf{x}'^T D\mathbf{x}'} d\mathbf{x}' = \sqrt{\frac{\pi^n}{\det A}}$$

5. Adjungovaný operátor jako polynom

TVRZENÍ 63. *Nechť $\mathbb{N}$ je normální operátor na unitárním prostoru V dimenze n. Pak existuje komplexní polynom p stupně nejvýše n takový, že $\mathbb{N}^* = p(\mathbb{N})$.*

Důkaz. Označme N matici $\mathbb{N}$ vzhledem k libovolné ortonormální bázi B a D jeho diagonální matici vzhledem nějaké ortonormální bázi C zaručené větou 29. Pro tyto matice a unitární matici přechodu U pak platí $N = UDU^+$. Protože $N^+ = UD^+U^+$ je matice $\mathbb{N}^*$, stačí jen dokázat, že D^+ je možné vyjádřit jako $p(D)$ pro nějaký polynom p. Je-li $D = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$, pak $D = \mathrm{diag}(\bar{\lambda}_1, \ldots, \bar{\lambda}_n)$. Polynom, pro který platí $p(\lambda_i) = \bar{\lambda}_i$ pro všechna i lze zkonstruovat jako Lagrangeův interpolační polynom pro k vzájemně různých vlastních čísel operátoru $\mathbb{N}$. $\square$