8. Skalární součin

Cíl. *Velikost prvků a úhly mezi nimi počítáme v lineárním prostoru pomocí skalárního součinu. Skalární součin definujeme pouze v lineárních prostorech nad reálnými nebo nad komplexními čísly. Dále se budeme věnovat studiu a aplikacím pojmu ortogonalita.*

V abstraktním lineárním prostoru nemáme metrické pojmy jako délka prvku (vektoru) nebo úhel mezi dvěma prvky (vektory). Tyto pojmy zavedeme přidáním skalárního součinu.

8.1. Standardní skalární součin v $\mathbb{R}^n$ a $\mathbb{C}^n$.

8.1.1. *Aritmetický prostor* $\mathbb{R}^n$. Podíváme se nejprve na standardní skalární součin v aritmetickém vektorovém prostoru $\mathbb{R}^n$, někdy se mu také říká *bodový součin.*

Definice 8.1. Pro dva n-složkové aritmetické vektory $\mathbf{u} = (x_1, x_2, \dots, x_n)^T$, $\mathbf{v} = (y_1, y_2, \dots, y_n) \in \mathbb{R}^n$ definujeme jejich *standardní skalární součin* jako reálné číslo

$$\mathbf{u} \cdot \mathbf{v} = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n \ .$$

Všimněme si, že standardní skalární součin $\mathbf{u} \cdot \mathbf{v} = \mathbf{u}^T \mathbf{v}$. Pomocí standardního skalárního součinu definujeme eukleidovskou délku (též zvanou normu) vektoru $\mathbf{u} \in \mathbb{R}^n$.

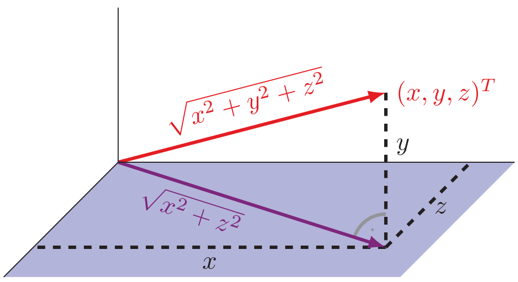
Definice 8.2. *Eukleidovská norma* nebo také *eukleidovská délka* vektoru $\mathbf{u} \in \mathbb{R}^n$ je číslo

$$\|\mathbf{u}\| = \sqrt{\mathbf{u} \cdot \mathbf{u}} = \sqrt{\mathbf{u^T}\,\mathbf{u}} \ .$$

Eukleidovskou normu vektoru $\mathbf{u} = (x_1, x_2, \dots, x_n)^T$ tak spočítáme jako

$$\|\mathbf{u}\| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2} \ .$$

Pro $n = 2$ a $n = 3$ jde o stejný vzorec, který v elementární geometrii plyne z Pythagorovy věty. Pro $n = 1$ dostáváme $\|\mathbf{u}\| = \|(x_1)\| = |x_1|$.



Obrázek 70. Eukleidovská norma v $\mathbb{R}^3$

Geometrický význam skalárního součinu dvojice vektorů $\mathbf{u} = (x_1, x_2)^T, \mathbf{v} = (y_1, y_2)^T$ v rovině $\mathbb{R}^2$ můžeme pochopit pomocí kosinové věty. Je-li posloupnost vektorů $(\mathbf{u}, \mathbf{v})$ lineárně nezávislá, vyjdeme z trojúhelníku o stranách $\mathbf{u}, \mathbf{v}$ a $\mathbf{u} - \mathbf{v}$.

OBRAZEK N1 - kosinova veta

Označíme α úhel mezi vektory $\mathbf{u}$ a $\mathbf{v}$. Podle kosinové věty platí

$$\|\mathbf{u}-\mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - 2\,\|\mathbf{u}\|\,\|\mathbf{v}\|\,\cos\alpha \ .$$

Protože $\mathbf{u}-\mathbf{v} = (x_1-y_1, x_2-y_2)^T$, můžeme normu na levé straně vyjádřit pomocí standardního skalárního součinu ve tvaru

$$\begin{aligned}\|\mathbf{u}-\mathbf{v}\|^2 &= (x_1-y_1)^2 + (x_2-y_2)^2 = x_1^2 + x_2^2 + y_1^2 + y_2^2 - 2(x_1y_1 + x_2y_2) \\ &= \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - 2\,\mathbf{u}\cdot\mathbf{v} \ .\end{aligned}$$

Po dosazení do kosinové věty a úpravě dostaneme

$$\mathbf{u}\cdot\mathbf{v} = \|\mathbf{u}\|\,\|\mathbf{v}\|\,\cos\alpha \ .$$

Výraz $\|\mathbf{u}\|\,\|\mathbf{v}\|\,\cos\alpha$ můžeme chápat jako součin délky vektoru $\mathbf{u}$ a orientované délky $\|\mathbf{v}\|\,\cos\alpha$ kolmé projekce vektoru $\mathbf{v}$ na přímku $\langle\mathbf{u}\rangle$, přičemž znaménko je kladné, pokud vektory $\mathbf{u}$ a $\mathbf{v}$ svírají ostrý úhel, a je záporné pokud svírají tupý úhel. Je-li součin $\|\mathbf{u}\|\,(\|\mathbf{v}\|\,\cos\alpha) = 0$, pak je buď $\|\mathbf{u}\| = 0$, tj. $\mathbf{u}$ je nulový vektor, nebo je $\mathbf{v} = \mathbf{o}$, a nebo jsou oba vektory $\mathbf{u}, \mathbf{v}$ nenulové a $\cos\alpha = 0$, tj. vektory $\mathbf{u}$ a $\mathbf{v}$ svírají pravý úhel.

OBRAZEK N2 - projekce

Symetricky se na standardní skalární součin $\mathbf{u}\cdot\mathbf{v}$ můžeme také dívat jako na součin délky vektoru $\mathbf{v}$ a orientované délky ortogonální projekce vektoru $\mathbf{u}$ na přímku $\langle\mathbf{v}\rangle$. Rozmyslete si také, že geometrický význam skalárního součinu zůstává v platnosti i v případě, kdy je jeden z vektorů $\mathbf{u}, \mathbf{v}$ skalárním násobkem druhého, tj. když je posloupnost vektorů $(\mathbf{u}, \mathbf{v})$ lineárně závislá.

Standardní skalární součin můžem také využít k novému náhledu na rovnici přímky v rovině. Máme-li dánu rovnici

$$a_1x_1 + a_2x_2 = b \ ,$$

ve které je aspoň jeden z koeficientů a_1, a_2 nenulový, množina všech řešení $(x_1, x_2)^T \in \mathbb{R}^2$ tvoří nějakou přímku v rovině. Označíme-li $\mathbf{a} = (a_1, a_2)^T$ a $\mathbf{x} = (x_1, x_2)^T$, můžeme rovnici přímky pomocí standardního skalárního součinu přepsat do tvaru

$$\mathbf{a}\cdot\mathbf{x} = b \ .$$

Použijeme rovnost $\mathbf{a}\cdot\mathbf{x} = \|\mathbf{a}\|\,\|\mathbf{x}\|\,\cos\alpha$. Protože $\mathbf{a} \neq \mathbf{o}$, platí $\|\mathbf{a}\| \neq 0$ a poslední rovnici můžeme přepsat do tvaru

$$\|\mathbf{x}\|\,\cos\alpha = \frac{b}{\|\mathbf{a}\|} \ .$$

Na levé straně dostáváme orientovanou délku projekce proměnného vektoru $\mathbf{u}$ do směru vektoru $\mathbf{a}$, zatímco pravá strana závisí pouze na koeficientech rovnice $a_1x_1 + a_2x_2 = b$ a je tedy konstatntní. Množinu všech řešení rovnice tak tvoří vektory, které mají daný průmět do směru vektoru $\mathbf{a}$. Tak jsme geometrickou úvahou založenou na skalárním součinu dostali ještě jednou, že množina všech řešení rovnice $a_1x_1 + a_2x_2 = b$ tvoří přímku v rovině, pokud je aspoň jeden z koeficientů a_1, a_2 nenulový.

OBRAZEK N3 - přímka s normálovým vektorem

Vektor koeficientů $\mathbf{a} = (a_1, a_2)^T$ u proměnných $\mathbf{x} = (x_1, x_2)^T$ se nazývá *normálový vektor* přímky s rovnicí $a_1x_1 + a_2x_2 = b$, protože tato přímka je kolmá na vektor $\mathbf{a}$ (neboli na přímku $\langle \mathbf{a} \rangle$). Rovnice $a_1x_1 + a_2x_2 = b$ a $a_1x_1 + a_2x_2 = c$ proto určují rovnoběžné přímky.

Pomocí standarního skalárního součinu můžeme také snadno najít rovnici přímky v rovině procházející dvěma různými body.

Příklad 8.3. Najdeme rovnici $a_1x_1 + a_2x_2 = b$ přímky l v rovině procházející body $P = (1, 3)$ a $Q = (2, 1)$. Vektor $\mathbf{u}$ s počátečním bodem P a koncovým bodem Q je rovnoběžný s přímkou l, má souřadnice $(2, 1)^T - (1, 3)^T = (1, -2)^T$ a je kolmý na libovolný normálový vektor přímky l. Za normálový vektor tedy můžeme zvolit například $\mathbf{a} = (2, 1)^T$. Tím jsme našli koeficienty $a_1 = 2$ a $a_2 = 1$ rovnice přímky l. Ta se tedy rovná $2x_1 + 1x_2 = b$ a souřadnice $(1, 3)^T$ bodu P ji musí splňovat. Číslo b se proto rovná $b = 2 \cdot 1 + 1 \cdot 3 = 5$. Jedna z možných rovnic přímky l je tedy $2x_1 + x_2 = 5$.

V první kapitole jsme rovnici přímky procházející dvěma různými body hledali pomocí parametrického tvaru přímky. Přímý postup využívající skalární součin je mnohem rychlejší.

Standardní skalární součin vektorů v rovině je symetrický, pro každé dva vektory $\mathbf{u}, \mathbf{v} \in \mathbb{R}^2$ platí $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$. Symetrie je v souladu s geometrickým významem $\mathbf{u} \cdot \mathbf{v} = \|\mathbf{u}\| \, \|\mathbf{v}\| \cos\alpha$, protože kosinus je sudá funkce a nezáleží tedy na tom, měříme-li úhel „od $\mathbf{u}$ k $\mathbf{v}$“ nebo „od $\mathbf{v}$ k $\mathbf{u}$“.

Pokud zvolíme dva vektory $\mathbf{u} = (x_1, x_2, x_3)^T, \mathbf{v} = (y_1, y_2, y_3)^T \in \mathbb{R}^3$ takové, že posloupnost $(\mathbf{u}, \mathbf{v})$ je lineárně nezávislá, pak z kosinové věty použité na trojúhelník se stranami $\mathbf{u}, \mathbf{v}, \mathbf{u} - \mathbf{v}$ dostaneme stejně jako v případě vektorů v rovině, že

$$\|\mathbf{u} - \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{u}\|^2 + \|\mathbf{u}\| \, \|\mathbf{v}\| \cos\alpha \ ,$$

a po úpravě analogické úpravám, které jsem prováděli s vektory v rovině, dostaneme

$$\mathbf{u} \cdot \mathbf{v} = x_1y_1 + x_2y_2 + x_3y_3 = \|\mathbf{u}\| \, \|\mathbf{v}\| \cos\alpha \ .$$

Také v dimenzi $n = 3$ geometrický význam skalárního součinu $\mathbf{u} \cdot \mathbf{v}$ coby součinu délky vektoru $\mathbf{u}$ s orientovanou délkou ortogonální projekce vektoru $\mathbf{v}$do přímky generované vektorem $\mathbf{u}$ zůstává v platnosti.

Tak jako v rovině vyjádříme množinu všech řešení rovnice $a_1x_1 + a_2x_2 + a_3x_3 = b$ pomocí skalárního součinu ve tvaru

$$\mathbf{a} \cdot \mathbf{x} = b \ .$$

Je-li aspoň jeden z koeficientů a_1, a_2, a_3 nenulový, je vektor $\mathbf{a} = (a_1, a_2, a_3)^T \neq \mathbf{o}$ a vektor neznámých $\mathbf{x} = (x_1, x_2, x_3)^T$ splňuje rovnici

$$\|\mathbf{x}\| \cos\alpha = \frac{b}{\|\mathbf{a}\|} \ .$$

Vektor $\mathbf{x} = (x_1, x_2, x_3)^T$ je tedy řešením rovnice $a_1x_1 + a_2x_2 + a_3x_3 = b$ právě když se pravoúhlá projekce $\mathbf{x}$ do přímky $\langle \mathbf{a} \rangle$ rovná vektoru

$$\frac{b}{\|\mathbf{a}\|} \, \frac{\mathbf{a}}{\|\mathbf{a}\|} = \frac{b}{\|\mathbf{a}\|^2} \mathbf{a}$$

a je to tedy rovina kolmá na přímku $\langle \mathbf{a} \rangle$ procházející bodem $b\mathbf{a} / \|\mathbf{a}\|^2$. V poslední rovnosti jsme orientovanou délku projekce násobili vektorem $\mathbf{a}/\|\mathbf{a}\|$, který má délku 1 a stejný směr jako vektor $\mathbf{a}$. Tím jsme orientovanou délku projekce $\mathbf{x}$ do směru vektoru $\mathbf{a}$ nezměnili. Vektor $\mathbf{a} = (a_1, a_2, a_3)^T$ je *normálový vektor* roviny dané rovnicí $a_1x_1 + a_2x_2 + a_3x_3 = b$.

V prostorech dimenze $n > 3$ postupujeme na základě analogie a považujeme normu

$$\|\mathbf{u}\| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$$

vektoru $\mathbf{u} = (x_1, x_2, \ldots, x_n)^T \in \mathbb{R}^n$ za délku vektoru $\mathbf{u}$. Je-li posloupnost $(\mathbf{u}, \mathbf{v})$ aritmetických vektorů $\mathbf{u}$ a $\mathbf{v} = (y_1, y_2, \ldots, y_n)^T$ lineárně nezávislá v $\mathbb{R}^n$, generuje rovinu $\langle \mathbf{u}, \mathbf{v} \rangle$ v $\mathbb{R}^n$. V této rovině platí pro trojúhelník určený vektory $\mathbf{u}, \mathbf{v}$ kosinová věta a z ní opět plyne geometrický význam standardního skalárního součinu

$$\mathbf{u} \cdot \mathbf{v} = x_1y_1 + x_2y_2 + \cdots + x_ny_n = \|\mathbf{u}\| \; \|\mathbf{v}\| \cos\alpha \ ,$$

kde α znovu značí úhel, který svírají vektory $\mathbf{u}$ a $\mathbf{v}$.

Základní algebraické vlastnosti standardního skalárního součinu shrnuje následující tvrzení.

Tvrzení 8.4. *Jsou-li* $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^n$ *libovolné reálné aritmetické vektory a* $a \in \mathbb{R}$ *skalár, pak platí*

(1) $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$,
(2) $\mathbf{u} \cdot (\mathbf{v} + \mathbf{w}) = \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w}$,
(3) $\mathbf{u} \cdot (a\mathbf{v}) = a(\mathbf{u} \cdot \mathbf{v})$,
(4) $\mathbf{u} \cdot \mathbf{u} \geq 0$ *a* $\mathbf{u} \cdot \mathbf{u} = 0$ *právě když* $\mathbf{u} = \mathbf{o}$.

Důkaz. Všechna tvrzení plynou okamžitě z definice standardního skalárního součinu. Například druhé z nich ověříme výpočtem

$$\mathbf{u} \cdot (\mathbf{v} + \mathbf{w}) = \mathbf{u}^T(\mathbf{v} + \mathbf{w}) = \mathbf{u}^T\mathbf{v} + \mathbf{u}^T\mathbf{w} = \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w} \ ,$$

použili jsme pouze distributivitu násobení matic vzhledem k jejich sčítání.

Poslední tvrzení plyne z toho, že pro $\mathbf{u} = (x_1, x_2, \ldots, x_n)^T$ platí

$$\mathbf{u} \cdot \mathbf{u} = \mathbf{u}^T\mathbf{u} = x_1^2 + x_2^2 + \cdots + x_n^2 \ .$$

□

První rovnost se nazývá *symetrie*, další dvě rovnosti říkají, že standardní skalární součin je *lineární vzhledem ke druhé složce*. Ze symetrie pak plyne, že je také *lineární vzhledem k první složce*. Poslední čtvrtá vlastnost se nazývá *pozitivní definitnost* standardního skalárního součinu. Význam tohoto názvu si ujasníme později.

Linearita standardního skalárního součinu má názorný geometrický význam, který je vidět na následujícím obrázku. Stačí vyjít z toho, že standardní skalární součin $\mathbf{u} \cdot \mathbf{v}$ se rovná součinu délky vektoru $\mathbf{u}$ s orientovanou délkou projekce vektoru $\mathbf{v}$ do přímky $\langle \mathbf{u} \rangle$.

OBRAZEK N4 - linearita

Z linearity samotné už naopak skoro plyne vzorec pro standardní skalární součin vektorů $\mathbf{u} = (x_1, x_2, \ldots, x_n)^T$ a $\mathbf{v} = (y_1, y_2, \ldots, y_n)^T \in \mathbb{R}^n$. Pro přehlednost uvedeme nejprve odvození v případě $n = 2$.

$$\begin{aligned}
\mathbf{u} \cdot \mathbf{v} &= (x_1\mathbf{e}_1 + x_2\mathbf{e}_2) \cdot (y_1\mathbf{e}_1 + y_2\mathbf{e}_2) \\
&= (x_1\mathbf{e}_1) \cdot (y_1\mathbf{e}_1 + y_2\mathbf{e}_2) + (x_2\mathbf{e}_2) \cdot (y_1\mathbf{e}_1 + y_2\mathbf{e}_2) \\
&= (x_1\mathbf{e}_1) \cdot (y_1\mathbf{e}_1) + (x_1\mathbf{e}_1) \cdot (y_2\mathbf{e}_2) + (x_2\mathbf{e}_2) \cdot (y_1\mathbf{e}_1) + (x_2\mathbf{e}_2) \cdot (y_2\mathbf{e}_2) \\
&= x_1y_1(\mathbf{e}_1 \cdot \mathbf{e}_1) + x_1y_2(\mathbf{e}_1 \cdot \mathbf{e}_2) + x_2y_1(\mathbf{e}_2 \cdot \mathbf{e}_1) + x_2y_2(\mathbf{e}_2 \cdot \mathbf{e}_2) \ .
\end{aligned}$$

Pokud ještě stanovíme, že oba vektory $\mathbf{e}_1, \mathbf{e}_2$ mají délku 1, tj. $\mathbf{e}_1 \cdot \mathbf{e}_1 = \mathbf{e}_2 \cdot \mathbf{e}_2 = 1$, a jsou navzájem kolmé, tj. $\mathbf{e}_1 \cdot \mathbf{e}_2$, plyne odtud

$$\mathbf{u} \cdot \mathbf{v} = x_1 y_1 + x_2 y_2 \ .$$

Analogicky z linearity a předpokladu, že vektory kanonické báze $(\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n)$ prostoru $\mathbb{R}^n$ splňují

$$\mathbf{e}_i \cdot \mathbf{e}_j = 0 \ \text{ pokud } \ i \neq j, \quad \mathbf{e}_i \cdot \mathbf{e}_i = 1 \ ,$$

dostaneme v případě libovolné dimenze n, že

$$\begin{aligned} \mathbf{u} \cdot \mathbf{v} &= \left(\sum_{i=1}^{n} x_i \mathbf{e}_i \right) \cdot \left(\sum_{i=1}^{n} y_i \mathbf{e}_i \right) = \sum_{i=1}^{n} \sum_{j=1}^{n} (x_i \mathbf{e}_i) \cdot (y_j \mathbf{e}_j) \\ &= \sum_{i=1}^{n} \sum_{j=1}^{n} x_i y_j (\mathbf{e}_i \cdot \mathbf{e}_j) = \sum_{i=1}^{n} x_i y_i \ . \end{aligned}$$

Všimněte si, že odvození probíhalo podobně jako odvození vzorce pro determinant – předpokládali jsme linearitu ve všech proměnných a řekli jsme si, jak skalární součin (determinant) vypadá na kanonické bázi.

Příklad 8.5. V některých oblastech matematiky bývá zvykem označovat symbolem $\mathbf{1}_n$ n-složkový vektor, který má všechny složky rovné 1. Protože počet složek bývá obvykle jasný z kontextu, index n je vynecháván.

Pro každý vektor $\mathbf{x} \in \mathbb{R}^n$ platí

$$\mathbf{1} \cdot \mathbf{x} = x_1 + x_2 + \cdots + x_n \ .$$

Podobně můžeme vyjádřit aritmetický průměr čísel $x_1, x_2, \dots, x_n$ jako standardní skalární součin

$$\left(\frac{1}{n} \mathbf{1} \right) \cdot \mathbf{x} = \frac{x_1 + x_2 + \cdots + x_n}{n} \ .$$

Obecněji můžeme každé složce x_i vektoru $\mathbf{x}$ přiřadit nějakou váhu $w_i \geq 0$, označit $\mathbf{w} = (w_1, w_2, \dots, w_n)^T$ vektor vah a spočítat *vážený součet* složek vektoru $\mathbf{x}$ s váhami $\mathbf{w}$ jako standardní skalární součin

$$\mathbf{w} \cdot \mathbf{x} = w_1 x_1 + w_2 x_2 + \cdot + w_n x_n \ .$$

Pokud o váhovém vektoru $\mathbf{w}$ navíc předpokládáme, že $w_1 + w_2 + \cdots + w_n = 1$, dostáváme *vážený průměr* prvků vektoru $\mathbf{x}$.

Vážený součet je někdy používán při prohlížení databází informací o dokumentech. Zajímá nás výskyt vybraného tisíce slov v nějakých dokumentech. Každý dokument si zaznamenáme jako vektor $\mathbf{x} = (x_1, x_2, \dots, x_{1000})^T$, kde x_i udává, kolikrát se i-té slovo vyskytne v příslušném dokumentu. Informace o všech dokumentech tak máme uložené jako nějakou množinu aritmetických vektorů $\mathbf{x} \in \mathbb{R}^{1000}$.

Nyní chceme uspořádat dokumenty v databázi podle počtu výskytů nějakých vybraných slov z celkového tisíce. Označíme si $J \subseteq \{1, 2, \dots, 1000\}$ množinu indexů slov, která nás zajímají. Označíme $\mathbf{w}_J = (w_1, w_2, \dots, w_{1000}) \in \mathbb{R}^{1000}$ váhový vektor, jehož složky jsou definované jako

$$w_i = \begin{cases} 1, & \text{pokud } i \in J \\ 0, & \text{pokud } i \notin J \ . \end{cases}$$

Složka $w_i = 1$, pokud nás zajímá výskyt i-tého slova, a $w_i = 0$, pokud nás nezajímá.

Jedotlivé dokumenty v databázi můžeme nyní uspořádat podle hodnoty skalárního součinu $\mathbf{w} \cdot \mathbf{x}$, která udává součet počtu výskytů slov, jejichž indexy leží v J, v dokumentu se záznamem $\mathbf{x}$.

8.1.2. *Aritmetický prostor* $\mathbb{C}^n$. Nad komplexními čísly je standardní skalární součin aritmetických vektorů definován trochu jiným způsobem.

Definice 8.6. Pro dva komplexní aritmetické vektory $\mathbf{u} = (x_1, x_2, \dots, x_n)^T$ a $\mathbf{v} = (y_1, y_2, \dots, y_n)^T$ definujeme *standardní skalární součin* $\mathbf{u} \cdot \mathbf{v}$ předpisem

$$\mathbf{u} \cdot \mathbf{v} = \overline{x_1}y_1 + \overline{x_2}y_2 + \cdots + \overline{x_n}y_n \ ,$$

kde $\overline{x}$ značí číslo komplexně sdružené k x, tj. $\overline{a+bi} = a - bi$.

Pro reálné vektory tato definice souhlasí s předchozí, protože komplexní sdružování s reálnými čísly nic nedělá. Fyzikální motivace pochází z kvantové mechaniky. Z matematického hlediska má tato definice výhodu v tom, že skalární součin $\mathbf{u} \cdot \mathbf{u}$ je vždy kladné reálné číslo, neboť je součtem druhých mocnin absolutních hodnot $\overline{x_i}x_i = |x_i|^2$ složek vektoru $\mathbf{u}$. Také komplexní aritmetické vektory mají nezápornou reálnou délku ve smyslu následující definice.

Definice 8.7. *Eukleidovskou délku* nebo také *eukleidovskou normu* aritmetického vektoru $\mathbf{u} = (x_1, x_2, \dots, x_n)^T \in \mathbb{C}^n$ definujeme jako

$$\|\mathbf{u}\| = \sqrt{\mathbf{u} \cdot \mathbf{u}} = \sqrt{\overline{x_1}x_1 + \overline{x_2}x_2 + \cdots + \overline{x_n}x_n} = \sqrt{|x_1|^2 + |x_2|^2 + \cdots + |x_n|^2} \ .$$

Délka $\|\mathbf{u}\|$ je reálné číslo, které je nulové právě tehdy, když $\mathbf{u} = \mathbf{o}$. Pokud bychom definovali skalární součin bez komplexního sdružování, výraz $\mathbf{u} \cdot \mathbf{u}$ by nebyl vždy reálný a byl by roven nule i pro některé nenulové vektory.

V reálném případě můžeme standardní skalární součin definovat maticovým součinem $\mathbf{u}^T\mathbf{v}$. Abychom mohli maticově zapsat standardní skalární součin nad komplexními čísly, zavedeme pojem hermitovsky sdružené matice.

Definice 8.8. *Hermitovsky sdružená matice* k matici $A = (a_{ij})_{m\times n}$ je matice $A^* = (b_{ji})_{n\times m}$, kde $b_{ji} = \overline{a_{ij}}$ pro libovolné indexy $i \in \{1, 2, \dots, m\}$ a $j \in \{1, 2, \dots, n\}$.

Hermitovsky sdruženou matici k A tedy dostaneme transponováním a následným nahrazením všech prvků prvky komplexně sdruženými. Hermitovské sdružování se chová k ostatním operacím podobně jako transponování, viz cvičení. Speciálně, pokud je definován součin komplexních matic AB, platí

$$(AB)^* = B^*A^* \ .$$

Stejně tak $(A^*)^* = A$ pro každou komplexní matici A. Všimněme si také, že pokud jsou všechny prvky matice A reálné, platí $A^* = A^T$.

Příklad 8.9.

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

Pomocí hermitovského sdružování můžeme také standardní skalární součin komplexních vektorů zapsat pomocí součinu

$$\mathbf{u} \cdot \mathbf{v} = \mathbf{u}^*\mathbf{v} \ .$$

Následující jednoduché tvrzení shrnuje základní vlastnosti standardního skalárního součinu v aritmetickém prostoru $\mathbb{C}^n$.

Tvrzení 8.10. *Pro libovolné tři vektory* $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{C}^n$ *a komplexní číslo* a *platí*

(1) $\mathbf{u} \cdot \mathbf{v} = \overline{\mathbf{v} \cdot \mathbf{u}}$,
(2) $\mathbf{u} \cdot (\mathbf{v} + \mathbf{w}) = \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w}$,
(3) $\mathbf{u} \cdot (a\,\mathbf{v}) = a\,(\mathbf{u} \cdot \mathbf{v})$,
(4) $\mathbf{u} \cdot \mathbf{u}$ *je nezáporné reálné číslo, a* $\mathbf{u} \cdot \mathbf{u} = 0$ *právě když* $\mathbf{u} = \mathbf{o}$.

Důkaz. Tentokrát dokážeme první rovnost, důkaz ostatních vlastnosti ponecháme jako cvičení. Než se do toho pustíme, zavedeme úmluvu, že čtvercovou matici (a) řádu 1 obsahující jediný prvek a budeme v případě potřeby ztotožňovat s prvkem a. Při tomto ztotožnění platí $(a)^* = \bar{a}$.

Při ověření použijeme známou vlastnost, že pro každé komplexní číslo z platí $\bar{\bar{z}} = z$. Pak platí

$$\mathbf{u} \cdot \mathbf{v} = \mathbf{u}^*\mathbf{v} = \overline{\overline{\mathbf{u}^*\mathbf{v}}} = \overline{(\mathbf{u}^*\mathbf{v})^*} = \overline{\mathbf{v}^*\mathbf{u}} = \overline{\mathbf{v} \cdot \mathbf{u}}\ .$$

□

Druhá a třetí rovnost říkají, že standardní skalární součin nad komplexními čísly je rovněž lineární v druhé proměnné. V první proměnné ale lineární není. Platí pouze následující dvě rovnosti.

Pozorování 8.11. *Pro libovolné tři vektory* $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{C}^n$ *a komplexní číslo* a *platí*

(1) $(\mathbf{u} + \mathbf{v}) \cdot \mathbf{w} = \mathbf{u} \cdot \mathbf{w} + \mathbf{v} \cdot \mathbf{w}$,
(2) $(a\mathbf{u}) \cdot \mathbf{v} = \overline{a}\ (\mathbf{u} \cdot \mathbf{v})$.

Druhou rovnost ověříme přímým výpočtem

$$(a\mathbf{u}) \cdot \mathbf{v} = \overline{\mathbf{v} \cdot (a\mathbf{u})} = \overline{\mathbf{v}^*(a\mathbf{u})} = (\mathbf{v}^*(a\mathbf{u}))^* = (a\mathbf{u})^*\mathbf{v} = \mathbf{u}^*\bar{a}\,\mathbf{v} = \bar{a}\,\mathbf{u}^*\,\mathbf{v} = \bar{a}\,(\mathbf{u} \cdot \mathbf{v})\ .$$

8.2. **Obecný skalární součin.** Obecně definujeme skalární součin jako zobrazení přiřazující dvojici prvků nějakého lineárního prostoru skalár, které má podobné vlastnosti jako standardní skalární součin.

Skalární součin prvků $\mathbf{u}$ a $\mathbf{v}$ budeme značit $\langle\, \mathbf{u}, \mathbf{v}\rangle$, značení $\mathbf{u} \cdot \mathbf{v}$ budeme používat pouze pro standardní skalární součin v $\mathbb{R}^n$ nebo $\mathbb{C}^n$. Zdůrazněme ještě jednou, že **skalární součin definujeme pouze pro lineární prostory nad tělesem** $\mathbb{R}$ **nebo** $\mathbb{C}$.

Za základ definice obecného skalárního součinu vezmeme vlastnosti standardního skalárního součinu shrnuté v tvrzeních 8.4 a 8.10. Všimněme si také, že standardní skalární součin na prostoru $\mathbb{R}^n$ má všechny vlastnosti standardního skalárního součinu na prostoru $\mathbb{C}^n$, protože standardní skalární součin dvou reálných vektorů je reálné číslo a pro každé reálné číslo a platí $\bar{a} = a$.

Definice 8.12. Je-li $\mathbf{V}$ lineární prostor nad $\mathbb{R}$ (resp. nad $\mathbb{C}$), pak se zobrazení $\langle\, ,\rangle$ z $V \times V$ do $\mathbb{R}$ (resp do $\mathbb{C}$), které dvojici $\mathbf{u}, \mathbf{v}$ přiřadí skalár $\langle\, \mathbf{u}, \mathbf{v}\rangle$, nazývá *skalární součin* na $\mathbf{V}$, pokud pro libovolné $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ a $a \in \mathbb{R}$ (resp. $a \in \mathbb{C}$) platí

(SCS) $\langle\, \mathbf{u}, \mathbf{v}\rangle = \overline{\langle\, \mathbf{v}, \mathbf{u}\rangle}$,
(SL1) $\langle\, \mathbf{u}, a\mathbf{v}\rangle = a\,\langle\, \mathbf{u}, \mathbf{v}\rangle$,
(SL2) $\langle\, \mathbf{u}, \mathbf{v} + \mathbf{w}\rangle = \langle\, \mathbf{u}, \mathbf{v}\rangle + \langle\, \mathbf{u}, \mathbf{w}\rangle$,
(SP) $\langle\, \mathbf{u}, \mathbf{u}\rangle$ je nezáporné reálné číslo, které je nulové právě tehdy, když $\mathbf{u} = \mathbf{o}$.

První axion je „skorosymetrie“, další dva axiomy říkají, že i obecný skalární součin je lineární vzhledem ke druhé složce, poslední je pozitivní definitnost. Všimněme si také, že pokud je $\mathbf{V}$ reálný lineární prostor se skalárním součinem, splňuje všechny axiomy komplexního lineárního prostoru se skalárním součinem. Odtud plyne, že cokoliv dokážeme pro komplexní lineární prostory se skalárním součinem, platí i pro reálné. Proto budeme většinu důkazů v této kapitole dělat pouze pro komplexní prostory.

Začneme jednoduchými důsledky axiomů v předcházející definici.

Pozorování 8.13. *Je-li* $\mathbf{V}$ *lineární prostor nad* $\mathbb{R}$ *(resp. nad* $\mathbb{C}$*) se skalárním součinem* $\langle\, , \rangle$*, pak pro prvky libovolné* $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ *a skalár a platí*

(1) $\langle\, \mathbf{u}, \mathbf{o} \rangle = 0 = \langle\, \mathbf{o}, \mathbf{u} \rangle$
(2) $\langle\, a\mathbf{u}, \mathbf{v} \rangle = \overline{a}\, \langle\, \mathbf{u}, \mathbf{v} \rangle$
(3) $\langle\, \mathbf{u} + \mathbf{v}, \mathbf{w} \rangle = \langle\, \mathbf{u}, \mathbf{w} \rangle + \langle\, \mathbf{v}, \mathbf{w} \rangle$

Důkaz. Druhou vlastnost jsme dokázali už v pozorování 8.11, neboť jsme i tam použili pouze axiomy z definice 8.12. Třetí vlastnost dokážeme podobně dvojím použitím „skorosymetrie". Zbývá první vlastnost:

$$\langle\, \mathbf{u}, \mathbf{o} \rangle = \langle\, \mathbf{u}, 0\mathbf{o} \rangle = 0\, \langle\, \mathbf{u}, \mathbf{o} \rangle = 0,$$

zbytek plyne z antisymetrie. □

8.2.1. *Příklady.*

Příklad 8.14. Standardní skalární součin v $\mathbb{R}^n$ (resp. $\mathbb{C}^n$) je skalárním součinem v $\mathbb{R}^n$ (resp. $\mathbb{C}^n$) ve smyslu definice 8.12. Všechny axiomy jsme už ověřili v tvrzeních 8.4 a 8.10.

Příklad 8.15. Představme si $\mathbb{R}^2$ jako (nekonečný) list papíru a podívejme se na papír z jiné vzdálenosti a z jiného úhlu. Tím se nám změní vnímané délky vektorů a úhly mezi nimi. Uvažujme například situaci, kdy délka vektoru $\mathbf{e}_1$ zůstane 1, délka vektoru $\mathbf{e}_2$ bude 2 a vektory $\mathbf{e}_1$ a $\mathbf{e}_2$ budou svírat úhel $\pi/3$.

V části o standardním skalárním součinu na $\mathbb{R}^n$ jsme viděli, že formulka pro tento součin plyne z linearity a předpokladu kolmosti různých vektorů kanonické báze a toho, že mají délku 1. Zkusíme podobným způsobem zavést skalární součin pro „šikmý pohled" na rovinu. To znamená, že bude platit

$$\langle\, \mathbf{e}_i, \mathbf{e}_j \rangle = (\text{„délka“ } \mathbf{e}_i)(\text{„délka“ } \mathbf{e}_j)\, \cos\alpha,$$

kde α je úhel, který svírají vektory $\mathbf{e}_i$ a $\mathbf{e}_j$. To znamená, že

$$\begin{aligned}
\langle\, \mathbf{e}_1, \mathbf{e}_1 \rangle &= 1 \cdot 1 \cdot \cos 0 = 1\\
\langle\, \mathbf{e}_2, \mathbf{e}_2 \rangle &= 2 \cdot 2 \cdot \cos 0 = 4\\
\langle\, \mathbf{e}_1, \mathbf{e}_2 \rangle &= 1 \cdot 2 \cdot \cos(\pi/3) = 1 = \langle\, \mathbf{e}_2, \mathbf{e}_1 \rangle
\end{aligned}$$

Podobným výpočtem jako v případě standardního skalárního součinu získáme vzorec

$$\begin{aligned}
\left\langle \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \right\rangle &= \langle\, x_1\mathbf{e}_1 + x_2\mathbf{e}_2, y_1\mathbf{e}_1 + y_2\mathbf{e}_2 \rangle\\
&= x_1y_1\, \langle\, \mathbf{e}_1, \mathbf{e}_1 \rangle + x_1y_2\, \langle\, \mathbf{e}_1, \mathbf{e}_2 \rangle + x_2y_1\, \langle\, \mathbf{e}_2, \mathbf{e}_1 \rangle + x_2y_2\, \langle\, \mathbf{e}_2, \mathbf{e}_2 \rangle\\
&= x_1y_1 + x_1y_2 + x_2y_1 + 4x_2y_2\ .
\end{aligned}$$

Tento vztah lze maticově zapsat

$$\left\langle \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \right\rangle = (x_1\ x_2) \begin{pmatrix} 1 & 1 \\ 1 & 4 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\ .$$

Snadno ověříme, že tento vzorec splňuje všechny axiomy skalárního součinu. Ukážeme si například axiom (SP). Je-li $\mathbf{u} = (x_1, x_2)^T$, pak platí

$$\langle \mathbf{u}, \mathbf{u} \rangle = x_1^2 + 2x_1x_2 + 4x_2^2 = (x_1 + x_2)^2 + 3x_2^2 \ .$$

Odtud plyne, že $\langle \mathbf{u}, \mathbf{u} \rangle \geq 0$ pro každý vektor $\mathbf{u} \in \mathbb{R}^2$ a $\langle \mathbf{u}, \mathbf{u} \rangle = 0$ právě když $x_2 = 0$ a $x_1 + x_2 = 0$, tj. právě když $\mathbf{u} = (x_1, x_2)^T = (0, 0)^T$.

Příklad 8.16. Obecněji, je-li A čtvercová matice nad $\mathbb{R}$ (resp. $\mathbb{C}$), pak zobrazení z $\mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ (resp. $\mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}$) definované vztahem

$$\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u}^* A \mathbf{v}$$

vždy splňuje (SL1) a (SL2) (cvičení). Stejně snadno lze zjistit, pro které matice A platí „skorosymetrie“ (SCS).

Pozorování 8.17. *Matice $A = (a_{ij}) \in \mathbb{C}^{n \times n}$ splňuje rovnost $\mathbf{u}^* A \mathbf{v} = \overline{\mathbf{v}^* A \mathbf{u}}$ pro každé dva vektory $\mathbf{u}, \mathbf{v} \in \mathbb{C}^n$ právě když $A^* = A$.*

Pokud platí $A^* = A$, pak $\overline{\mathbf{v}^* A \mathbf{u}} = (\mathbf{v}^* A \mathbf{u})^* = \mathbf{u}^* A^* \mathbf{v} = \mathbf{u}^* A \mathbf{v}$. Pokud naopak rovnost $A^* = A$ neplatí, existují indexy $i, j \in \{1, 2, \ldots, n\}$ takové, že $a_{ij} \neq \overline{a_{ji}}$. Pak pro prvky $\mathbf{e}_i, \mathbf{e}_j$ kanonické báze v $\mathbb{C}^n$ platí

$$\langle \mathbf{e}_i, \mathbf{e}_j \rangle = \mathbf{e}_i^* A \mathbf{e}_j = a_{ij} \neq \overline{a_{ji}} = \overline{\mathbf{e}_j^* A \mathbf{e}_i} = \overline{\langle \mathbf{e}_i, \mathbf{e}_j \rangle} \ .$$

Komplexním maticím splňujícím rovnost $A^* = A$ říkáme *hermitovské.* V případě reálné matice to znamená, že A je symetrická.

Zdaleka ne pro všechny hermitovské (symetrické) matice splňuje zobrazení $\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u}^* A \mathbf{v}$ podmínku (SP). Má-li matice $A \in \mathbb{C}^{n \times n}$ splňovat podmínku (SP), musí být regulární. Pro singulární matici A totiž existuje nenulový vektor $\mathbf{u} \in \mathbb{C}^n$, pro který platí $A\mathbf{u} = \mathbf{o}$. Pro tento vektor $\mathbf{u}$ pak platí $\mathbf{u}^* A \mathbf{u} = 0$ a matice A tak podmínku (SP) nesplňuje. Regularita matice A ale není postačující pro splnění podmínky (SP). Lze si to snadno ověřit na matici

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \ .$$

Definice 8.18. Hermitovská matice A řádu n se nazývá *pozitivně definitní*, pokud $\mathbf{u}^* A \mathbf{u}$ je nezáporné reálné číslo pro libovolné $\mathbf{u} \in \mathbb{C}^n$ a rovná se 0 právě když $\mathbf{u} = \mathbf{o}$.

Pozitivně definitní matice hrají v lineární algebře a jejích aplikacích důležitou roli. Není ale jednoduché poznat, kdy je nějaké hermitovská matice pozitivně definitní. Příkladem pozitivně definitních matic jsou matice typu $A = B^* B$, kde B je regulární matice řádu n nad $\mathbb{R}$ (resp. nad $\mathbb{C}$). Snadno totiž spočteme, že v takovém případě pro každý vektor $\mathbf{u} \in \mathbb{C}^n$ platí

$$\mathbf{u}^* A \mathbf{u} = \mathbf{u}^* B^* B \mathbf{u} = (\mathbf{u}^* B^*)(B\mathbf{u}) = (B\mathbf{u})^*(B\mathbf{u}) = (B\mathbf{u})^* \cdot (B\mathbf{u}) \ .$$

V posledním výrazu používáme standardní skalární součin na $\mathbb{C}^n$, který podmínku (SP) splňuje. Takže $\mathbf{u}^* A \mathbf{u} = \|B\mathbf{u}\|^2 \geq 0$ pro každý vektor $\mathbf{u} \in \mathbb{C}^n$, přičemž rovnost nule nastává právě když $B\mathbf{u} = \mathbf{o}$. A poslední rovnost vzhledem k regularitě matice B nastává právě když $\mathbf{u} = \mathbf{o}$.

Později ukážeme, že platí i opak, tj. každá pozitivně definitní matice A je tvaru $A = B^* B$ pro nějakou regulární matici B. Dokonce každý skalární součin na $\mathbb{R}^n$ (a na $\mathbb{C}^n$) můžeme vyjádřit ve tvaru

$$\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u}^* B^* B \mathbf{v},$$

kde B je regulární matice řádu n. Navíc můžeme vždy najít jednoznačně určenou horní trojúhelníkovou matici B.

Shrnutí: Je-li $A = B^*B$, pak zobrazení definované $\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u}^* A \mathbf{v}$ je skalární součin na $\mathbb{C}^n$ (nebo na $\mathbb{R}^n$). Pro $A = I_n$ dostáváme standardní skalární součin. Jako ukázku jiného konkrétního příkladu vezmeme

$$B = \begin{pmatrix} \sqrt{3}/4 & 0 \\ 1/2 & 2 \end{pmatrix} ,$$

tedy

$$A = B^*B = B^TB = \begin{pmatrix} \sqrt{3}/4 & 1/2 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} \sqrt{3}/4 & 0 \\ 1/2 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 4 \end{pmatrix} .$$

Příslušný skalární součin v $\mathbb{C}^2$ je dán vztahem

$$\langle \mathbf{u}, \mathbf{v} \rangle = (\overline{x_1}, \overline{x_2}) \begin{pmatrix} 1 & 1 \\ 1 & 4 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \overline{x_1} y_1 + \overline{x_1} y_2 + \overline{x_2} y_1 + 4\overline{x_2} y_2 ,$$

kde $\mathbf{u} = (x_1, x_2)^T$ a $\mathbf{v} = (y_1, y_2)^T$. Stejný vztah (kde nemusíme komplexně sdružovat) definuje skalární součin v $\mathbb{R}^2$, tentýž jako v předchozím příkladu.

Příklad 8.19. Na prostoru spojitých reálných funkcí na intervalu $\langle 0, 2\pi \rangle$ je

$$\langle \mathbf{u}, \mathbf{v} \rangle = \int_0^{2\pi} \mathbf{u}\,\mathbf{v}$$

skalární součin.

Obecnější příklad skalárního součinu na prostoru všech spojitých reálných funkcí na intervalu $\langle 0, 2\pi \rangle$ je definovaný předpisem

$$\langle \mathbf{u}, \mathbf{v} \rangle = \int_0^{2\pi} \mathbf{u}\,\mathbf{v}\,\mathbf{w} ,$$

kde $\mathbf{w}$ je nějaká spojitá kladná *váhová* funkce na intervalu $\langle 0, 2\pi \rangle$.

Příklad 8.20. Prostor ℓ_2 je tvořen posloupnostmi $(a_n)_{n=1}^{\infty}$ komplexních čísel splňujícími

$$\sum_{n=1}^{\infty} |a_n|^2 < \infty .$$

(Je třeba si rozmyslet, že tato množina tvoří spolu s přirozenými operacemi sčítání a násobení skalárem vektorový prostor. Jediný obtížnější krok je uzavřenost na sčítání.) Na tomto prostoru je

$$\langle (a_n)_{n=1}^{\infty}, (b_n)_{n=1}^{\infty} \rangle = \sum_{n=1}^{\infty} \overline{a_n} b_n .$$

skalární součin.

Příklad 8.21. Důležité příklady skalárního součinu pochází z teorie pravděpodobnosti. Vektorový prostor tvoří náhodné veličiny na nějakém pevně zvoleném pravděpodobnostím prostoru. Tzv. kovariance, která, zhruba řečeno, měří míru závislosti jedné veličiny na druhé, splňuje všechny vlastnosti skalárního součinu až na implikaci zleva doprava v podmínce (SP) – $\langle \mathbf{u}, \mathbf{u} \rangle$ může být nula i pro nenulovou veličinu $\mathbf{u}$. (Tento drobný technický nedostatek lze odstranit ztotožněním veličin, jejichž rozdíl má nulový rozptyl.)

8.2.2. *Norma.* Normu vektoru v prostoru se skalárním součinem zavedeme stejným vztahem jakým jsme vyjádřili eukleidovskou normu (délku) pomocí standardního skalárního součinu.

Definice 8.22. Nechť $\mathbf{V}$ je lineární prostor se skalárním součinem $\langle\, ,\rangle$. *Normou* vektoru $\mathbf{v} \in V$ rozumíme reálné číslo

$$\|\mathbf{u}\| = \sqrt{\langle\, \mathbf{u}, \mathbf{u}\rangle}\ .$$

Vektor $\mathbf{u}$ se nazývá *jednotkový*, pokud $\|\mathbf{u}\| = 1$.

Definice dává smysl, protože výraz pod odmocninou je podle (SP) vždy nezáporné reálné číslo. Norma závisí na skalárním součinu, takže když používáme symbol normy, musí být z kontextu jasné, se kterým skalárním součinem pracujeme. Podobně i pro další pojmy jako úhel nebo kolmost, které budou zavedeny později.

Příklad 8.23. Norma vektoru $\mathbf{u} = (1,1)^T$ v prostoru $\mathbb{R}^2$ se standardním skalárním součinem je $\|\mathbf{u}\| = \sqrt{\mathbf{u}^T\mathbf{u}} = \sqrt{2}$. Norma vektoru $\mathbf{u}$ v prostoru $\mathbb{R}^2$ se skalárním součinem

$$\left\langle\, (x_1, x_2)^T, (y_1, y_2)^T \right\rangle = x_1y_1 + x_1y_2 + x_2y_1 + 4x_2y_2$$

je ale $\|\mathbf{u}\| = \sqrt{\langle\, \mathbf{u}, \mathbf{u}\rangle} = \sqrt{7}$.

Příklad 8.24. Norma vektoru $(1+i, 2, 3-2i)^T$ v prostoru $\mathbb{C}^3$ se standardním skalárním součinem je

$$\left\| \begin{pmatrix} 1+i \\ 2 \\ 3-2i \end{pmatrix} \right\| = \sqrt{\begin{pmatrix} 1+i \\ 2 \\ 3-2i \end{pmatrix} \cdot \begin{pmatrix} 1+i \\ 2 \\ 3-2i \end{pmatrix}} = \sqrt{|1+i|^2 + |2|^2 + |3-2i|^2} = \sqrt{19}\ .$$

Norma určená skalárním součinem má následující vlastnosti.

Tvrzení 8.25. *Nechť* $\mathbf{V}$ *je lineární prostor nad* $\mathbb{R}$ *(resp.* $\mathbb{C}$*) se skalárním součinem* $\langle\, ,\rangle$, $\mathbf{u}, \mathbf{v} \in V$ *a* $t \in \mathbb{R}$ *(resp.* $t \in \mathbb{C}$*). Pak platí*

(1) $\|\mathbf{u}\| \geq 0$, *přičemž* $\|\mathbf{u}\| = 0$ *právě tehdy, když* $\mathbf{u} = \mathbf{o}$,
(2) $\|t\mathbf{u}\| = |t|\, \|\mathbf{u}\|$,
(3) $\|\mathbf{u}+\mathbf{v}\|^2 + \|\mathbf{u}-\mathbf{v}\|^2 = 2\,\|\mathbf{u}\|^2 + 2\,\|\mathbf{v}\|^2$, *(rovnoběžníkové pravidlo),*
(4) $\mathrm{Re}\,(\langle\, \mathbf{u}, \mathbf{v}\rangle) = \frac{1}{2}(\|\mathbf{u}+\mathbf{v}\|^2 - \|\mathbf{u}\|^2 - \|\mathbf{v}\|^2)$, *(polarizační identita),*

kde $\mathrm{Re}\,(x)$ *značí reálnou část* x.

Důkaz.

(1) Snadný důsledek (SP).
(2) Použitím (SL1) dostáváme

$$\|t\mathbf{u}\| = \sqrt{\langle\, t\mathbf{u}, t\mathbf{u}\rangle} = \sqrt{t\bar{t}\,\langle\, \mathbf{u}, \mathbf{u}\rangle} = \sqrt{|t|^2\,\langle\, \mathbf{u}, \mathbf{u}\rangle} = |t|\sqrt{\langle\, \mathbf{u}, \mathbf{u}\rangle} = |t|\,\|\mathbf{u}\|\ .$$

(3) Ve výpočtu stačí použít (SL2).

$$\begin{aligned}\|\mathbf{u}+\mathbf{v}\|^2 + \|\mathbf{u}-\mathbf{v}\|^2 &= \langle\, \mathbf{u}+\mathbf{v}, \mathbf{u}+\mathbf{v}\rangle + \langle\, \mathbf{u}-\mathbf{v}, \mathbf{u}-\mathbf{v}\rangle \\ &= \langle\, \mathbf{u}, \mathbf{u}\rangle + \langle\, \mathbf{u}, \mathbf{v}\rangle + \langle\, \mathbf{v}, \mathbf{u}\rangle + \langle\, \mathbf{v}, \mathbf{v}\rangle \\ &\quad + \langle\, \mathbf{u}, \mathbf{u}\rangle - \langle\, \mathbf{u}, \mathbf{v}\rangle - \langle\, \mathbf{v}, \mathbf{u}\rangle + \langle\, \mathbf{v}, \mathbf{v}\rangle \\ &= 2\,\langle\, \mathbf{u}, \mathbf{u}\rangle + 2\,\langle\, \mathbf{v}, \mathbf{v}\rangle = 2\,\|\mathbf{u}\|^2 + 2\,\|\mathbf{v}\|^2\ .\end{aligned}$$

(4) Ze (SL2) a (SCS) vypočteme

$$\|\mathbf{u}+\mathbf{v}\|^2 = \langle\, \mathbf{u},\mathbf{u}\rangle + \langle\, \mathbf{u},\mathbf{v}\rangle + \langle\, \mathbf{v},\mathbf{u}\rangle + \langle\, \mathbf{v},\mathbf{v}\rangle = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 + \langle\, \mathbf{u},\mathbf{v}\rangle + \overline{\langle\, \mathbf{u},\mathbf{v}\rangle}\ .$$

Protože $x + \overline{x} = 2\,\mathrm{Re}\,(x)$, dostáváme

$$2\,\mathrm{Re}\,(\langle\, \mathbf{u},\mathbf{v}\rangle) = \|\mathbf{u}+\mathbf{v}\|^2 - \|\mathbf{u}\|^2 - \|\mathbf{v}\|^2\ .$$

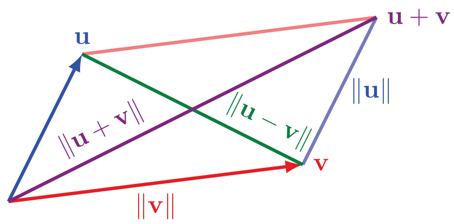
□

Důsledkem (1) a (2) je, že pro nenulový vektor **u** je jeho násobek

$$\frac{\mathbf{u}}{\|\mathbf{u}\|}$$

jednotkový vektor. Říkáme, že vektor $\mathbf{u}/\|\mathbf{u}\|$ vznikl z **u** *znormováním.*

Rovnoběžníkové pravidlo je ilustrováno na obrázku.



Obrázek 71. Rovnoběžníkové pravidlo

Polarizační identita vyjadřuje reálnou část skalárního součinu pouze pomocí normy. Podobný vztah jde napsat i pro imaginární část (pokud pracujeme v prostoru nad $\mathbb{C}$), viz cvičení. Skalární součin je tedy určen normou. Různé další varianty polarizační identity jsou ve cvičeních.

8.2.3. *Cauchyho-Schwarzova nerovnost, úhel.* Pro vektory $\mathbf{u},\mathbf{v} \in \mathbb{R}^3$ jsme nahlédli, že $\mathbf{u}\cdot\mathbf{v} = \|\mathbf{u}\|\,\|\mathbf{v}\|\cos\alpha$. Z toho také vyplývá, že absolutní hodnota $|\mathbf{u}\cdot\mathbf{v}|$ nemůže být větší než součin norem $\|\mathbf{u}\|\,\|\mathbf{v}\|$, protože kosinus úhlu je vždy v intervalu $\langle -1, 1\rangle$.

Vztah $\langle\, \mathbf{u},\mathbf{v}\rangle = \|\mathbf{u}\|\,\|\mathbf{v}\|\cos\alpha$ jde naopak použít pro zavedení úhlu mezi dvěma prvky v libovolném reálném lineárním prostoru se skalárním součinem. Aby byl úhel dobře definován, musíme dokázat, že obecně platí $|\langle\, \mathbf{u},\mathbf{v}\rangle| \leq \|\mathbf{u}\|\,\|\mathbf{v}\|$. Tato nerovnost se nazývá Cauchyho-Schwarzova nerovnost (též Bunjakovského nerovnost, nebo Cauchyho-Schwarzova-Bunjakovského nerovnost, apod.) a je asi jednou z nejdůležitějších nerovností v matematice.

Věta 8.26 (Cauchyho-Schwarzova nerovnost)**.** *Nechť* **V** *je lineární prostor se skalárním součinem* $\langle\, ,\rangle$ *a* $\mathbf{u},\mathbf{v} \in V$. *Pak platí*

$$|\langle\, \mathbf{u},\mathbf{v}\rangle| \leq \|\mathbf{u}\|\,\|\mathbf{v}\|\ ,$$

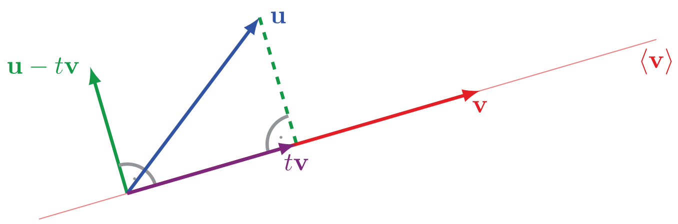
přičemž rovnost nastává právě tehdy, když $(\mathbf{u},\mathbf{v})$ *je lineárně závislá posloupnost.*

Důkaz. Pokud je posloupnost $(\mathbf{u},\mathbf{v})$ lineárně závislá, pak $\mathbf{u} = t\mathbf{v}$ nebo $\mathbf{v} = t\mathbf{u}$ pro nějaké $t \in \mathbb{C}$. V prvním případě je

$$|\langle\, \mathbf{u},\mathbf{v}\rangle| = |\langle\, t\mathbf{v},\mathbf{v}\rangle| = |\overline{t}\,\langle\, \mathbf{v},\mathbf{v}\rangle| = |t|\,\|\mathbf{v}\|^2$$

a
$$\|\mathbf{u}\|\,\|\mathbf{v}\| = \|t\mathbf{v}\|\,\|\mathbf{v}\| = |t|\,\|\mathbf{v}\|^2 \quad .$$
V případě $\mathbf{v} = t\mathbf{u}$ se rovnost $|\langle \mathbf{u}, \mathbf{v}\rangle| = \|\mathbf{u}\|\,\|\mathbf{v}\|$ odvodí podobně.

Nyní předpokládáne, že $(\mathbf{u}, \mathbf{v})$ je lineárně nezávislá posloupnost a odvodíme ostrou nerovnost. Díky lineární nezávislosti pro libovolné $t \in \mathbb{C}$ platí
$$0 < \|\mathbf{u} - t\mathbf{v}\|^2 \quad .$$
Zvolíme $t \in \mathbb{C}$ tak, aby platilo $\langle \mathbf{v}, \mathbf{u} - t\mathbf{v}\rangle = 0$. Geometrický význam v případě standardního skalárního součinu v $\mathbb{R}^n$ je vyznačen na obrázku: vektor $t\mathbf{v}$ je ortogonální projekcí vektoru $\mathbf{u}$ na $\langle \mathbf{v}\rangle$. Později dáme této intuici přesný význam pro obecný skalární součin.



Obrázek 72. K důkazu Cauchy-Schwarzovy nerovnosti

Vztah $\langle \mathbf{v}, \mathbf{u} - t\mathbf{v}\rangle = 0$ je ekvivalentní $\langle \mathbf{v}, \mathbf{u}\rangle - t\,\langle \mathbf{v}, \mathbf{v}\rangle = 0$, což je ekvivalentní
$$t = \frac{\langle \mathbf{v}, \mathbf{u}\rangle}{\|\mathbf{v}\|^2} \quad .$$
(Nulou nedělíme, protože prvek $\mathbf{v}$ je nenulový podle předpokladu o lineární nezávislosti $(\mathbf{u}, \mathbf{v})$.)

Při této volbě t dostáváme
$$\begin{aligned} 0 < \|\mathbf{u} - t\mathbf{v}\|^2 &= \langle \mathbf{u} - t\mathbf{v}, \mathbf{u} - t\mathbf{v}\rangle = \langle \mathbf{u}, \mathbf{u} - t\mathbf{v}\rangle - \bar{t}\,\langle \mathbf{v}, \mathbf{u} - t\mathbf{v}\rangle \\ &= \langle \mathbf{u}, \mathbf{u} - t\mathbf{v}\rangle = \langle \mathbf{u}, \mathbf{u}\rangle - t\,\langle \mathbf{u}, \mathbf{v}\rangle = \|\mathbf{u}\|^2 - \frac{\langle \mathbf{v}, \mathbf{u}\rangle}{\|\mathbf{v}\|^2}\,\langle \mathbf{u}, \mathbf{v}\rangle \\ &= \|\mathbf{u}\|^2 - \frac{\overline{\langle \mathbf{u}, \mathbf{v}\rangle}\,\langle \mathbf{u}, \mathbf{v}\rangle}{\|\mathbf{v}\|^2} = \|\mathbf{u}\|^2 - \frac{|\langle \mathbf{u}, \mathbf{v}\rangle|^2}{\|\mathbf{v}\|^2} \quad . \end{aligned}$$
Po vynásobení $\|\mathbf{v}\|^2$, drobné úpravě a odmocnění (oba výrazy, z nichž se počítá druhá mocnina jsou kladné) vyjde dokazovaná nerovnost:
$$\begin{aligned} 0 &< \|\mathbf{u}\|^2 - \frac{|\langle \mathbf{u}, \mathbf{v}\rangle|^2}{\|\mathbf{v}\|^2} \\ 0 &< \|\mathbf{u}\|^2\,\|\mathbf{v}\|^2 - |\langle \mathbf{u}, \mathbf{v}\rangle|^2 \\ |\langle \mathbf{u}, \mathbf{v}\rangle|^2 &< \|\mathbf{u}\|^2\,\|\mathbf{v}\|^2 \\ |\langle \mathbf{u}, \mathbf{v}\rangle| &< \|\mathbf{u}\|\,\|\mathbf{v}\| \quad . \end{aligned}$$

□

Příklad 8.27. Pro standardní skalární součin v $\mathbb{C}^n$ říká Cauchyho-Schwarzova nerovnost

$$|\overline{x_1}y_1+\overline{x_2}y_2+\cdots+\overline{x_n}y_n| \leq \sqrt{|x_1|^2+|x_2|^2+\cdots+|x_n|^2}\,\sqrt{|y_1|^2+|y_2|^2+\cdots+|y_n|^2}\ .$$

V případě skalárního součinu na $\mathbb{C}^2$ daného vzorcem

$$\left\langle \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \right\rangle = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}^* \begin{pmatrix} 5 & -2 \\ -2 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$

dostáváme

$$\begin{aligned} &|5\overline{x_1}y_1 - 2\overline{x_1}y_2 - 2\overline{x_2}y_1 + \overline{x_2}y_2| \\ &\qquad \leq \sqrt{5|x_1|^2 - 4\mathrm{Re}\,(\overline{x_1}x_2) + |x_2|^2}\,\sqrt{5|y_1|^2 - 4\mathrm{Re}\,(\overline{y_1}y_2) + |y_2|^2}\ . \end{aligned}$$

Pro prostor spojitých reálných funkcí na intervalu $\langle 0, 2\pi\rangle$ se skalárním součinem $\langle f, g\rangle = \int_0^{2\pi} fg$ Cauchyho-Schwartzova nerovnost znamená

$$\left|\int_0^{2\pi} fg\right| \leq \sqrt{\int_0^{2\pi} f^2}\,\sqrt{\int_0^{2\pi} g^2}\ .$$

Důležitým důsledkem Cauchyho-Schwarzovy nerovnosti je trojúhelníková nerovnost.

Důsledek 8.28 (Trojúhelníková nerovnost)**.** *Nechť* $\mathbf{V}$ *je lineární prostor se skalárním součinem* $\langle\,,\rangle$ *a* $\mathbf{u}, \mathbf{v} \in V$*. Pak platí*

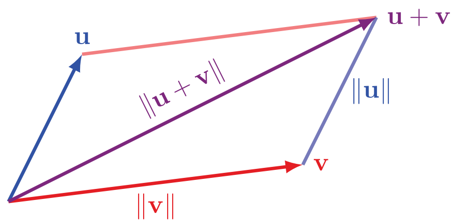
$$\|\mathbf{u}+\mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|\ .$$

Důkaz.

$$\begin{aligned} \|\mathbf{u}+\mathbf{v}\|^2 &= \langle \mathbf{u}+\mathbf{v}, \mathbf{u}+\mathbf{v}\rangle = \langle \mathbf{u}, \mathbf{u}\rangle + \langle \mathbf{u}, \mathbf{v}\rangle + \overline{\langle \mathbf{u}, \mathbf{v}\rangle} + \langle \mathbf{v}, \mathbf{v}\rangle \\ &= \|\mathbf{u}\|^2 + 2\,\mathrm{Re}\,(\langle \mathbf{u}, \mathbf{v}\rangle) + \|\mathbf{v}\|^2 \leq \|\mathbf{u}\|^2 + 2\,|\langle \mathbf{u}, \mathbf{v}\rangle| + \|\mathbf{v}\|^2 \\ &\leq \|\mathbf{u}\|^2 + 2\,\|\mathbf{u}\|\,\|\mathbf{v}\| + \|\mathbf{v}\|^2 = (\|\mathbf{u}\| + \|\mathbf{v}\|)^2 \end{aligned}$$

Cauchyho-Schwarzovu nerovnost jsme použili v předposlední úpravě. Výrazy pod druhými mocninami jsou kladné, takže nerovnost plyne odmocněním. □

Geometrický význam je patrný z obrázku.



Obrázek 73. Trojúhelníková nerovnost

Cauchyho-Schwarzova nerovnost nám umožňuje definovat úhel mezi prvky reálného lineárního prostoru se skalárním součinem.

Definice 8.29. Nechť $\mathbf{V}$ je lineární prostor nad $\mathbb{R}$ se skalárním součinem $\langle\, , \rangle$ a $\mathbf{o} \neq \mathbf{u}, \mathbf{v} \in V$. *Úhlem mezi prvky* $\mathbf{u}$ *a* $\mathbf{v}$ rozumíme reálné číslo $\alpha \in \langle 0, \pi \rangle$ splňující

$$\cos\alpha = \frac{\langle\, \mathbf{u}, \mathbf{v} \rangle}{\|\mathbf{u}\|\,\|\mathbf{v}\|}$$

Úhel mezi dvěma prvky existuje a je určen jednoznačně, protože zlomek je v intervalu $\langle -1, 1 \rangle$ podle Cauchy-Schwarzovo nerovnosti a funkce cos je bijekcí $\langle 0, \pi \rangle$ na interval $\langle -1, 1 \rangle$.

Pro libovolný skalární součin v lineárním prostoru nad reálnými čísly tedy máme vztah

$$\langle\, \mathbf{u}, \mathbf{v} \rangle = \|\mathbf{u}\|\,\|\mathbf{v}\| \cos\alpha \ .$$

Z tohoto vztahu snadno odvodíme kosinovou větu.

Tvrzení 8.30 (Kosinová věta)**.** *Nechť* $\mathbf{V}$ *je lineární prostor nad* $\mathbb{R}$ *se skalárním součinem* $\langle\, , \rangle$ *a* $\mathbf{o} \neq \mathbf{u}, \mathbf{v} \in V$*. Pak platí*

$$\|\mathbf{u} - \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - 2\,\|\mathbf{u}\|\,\|\mathbf{v}\| \cos\alpha \ ,$$

kde α *je úhel mezi vektory* $\mathbf{u}$ *a* $\mathbf{v}$*.*

Důkaz.

$$\begin{aligned}\|\mathbf{u} - \mathbf{v}\|^2 &= \langle\, \mathbf{u} - \mathbf{v}, \mathbf{u} - \mathbf{v} \rangle = \langle\, \mathbf{u}, \mathbf{u} \rangle - 2\,\langle\, \mathbf{u}, \mathbf{v} \rangle + \langle\, \mathbf{v}, \mathbf{v} \rangle \\ &= \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - 2\,\|\mathbf{u}\|\,\|\mathbf{v}\| \cos\alpha\end{aligned}$$

□

8.2.4. *Obecné normy.* Někdy bývá přirozenější měřit délku prvků v lineárním prostoru jiným způsobem, než pomocí normy definované skalárním součinem.

Definice 8.31. Je-li $\mathbf{V}$ lineární prostor nad $\mathbb{C}$ (nebo nad $\mathbb{R}$), pak zobrazení $\|\cdot\|$, které přiřazuje každému prvku $\mathbf{u}$ reálné číslo $\|\mathbf{u}\|$ nazýváme *norma* na prostoru $\mathbf{V}$, pokud platí pro kažé dva prvky $\mathbf{u}, \mathbf{v} \in V$ a každý skalár t

(1) $\|\mathbf{u}\| \geq 0$, přičemž $\|\mathbf{u}\| = 0$ právě tehdy, když $\mathbf{u} = \mathbf{o}$,
(2) $\|t\mathbf{u}\| = |t|\,\|\mathbf{u}\|$,
(3) $\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$.

Existuje mnoho norem, které nepochází ze skalárního součinu, například v $\mathbb{R}^n$ máme normu

$$\|(x_1, x_2, \ldots, x_n)\| = |x_1| + |x_2| + \cdots + |x_n| \ ,$$

která měří vzdálenost, když se můžeme pohybovat pouze pravoúhlým směrem (proto se jí někdy říká manhattanská norma). Norma pochází ze skalárního součinu právě tehdy, když splňuje rovnoběžníkové pravidlo, viz cvičení.

Jiným příkladem normy na $\mathbb{R}^n$ nepocházející ze skalárního součinu je norma

$$\|(x_1, x_2, \ldots, x_n)\| = \max\{|x_i| : i = 1, 2 \ldots, n\} \ .$$

OBRAZEK N5 - l_1-norma

8.3. **Kolmost.**

Ze vztahu $\langle\, \mathbf{u}, \mathbf{v}\rangle = \|\mathbf{u}\|\, \|\mathbf{v}\| \cos\alpha$ vidíme, že (nenulové) prvky lineárního prostoru svírají úhel $\pi/2$ právě tehdy, když je jejich skalární součin nula. To vede k přirozené definici kolmosti prvků lineárního prostoru se skalárním součinem.

Definice 8.32. Nechť $\mathbf{V}$ je lineární prostor se skalárním součinem $\langle\, , \rangle$. Prvky $\mathbf{u}, \mathbf{v} \in V$ nazýváme *kolmé* (nebo *ortogonální*) a píšeme $\mathbf{u} \perp \mathbf{v}$, pokud $\langle\, \mathbf{u}, \mathbf{v}\rangle = 0$.

Množina, nebo posloupnost, M prvků V se nazývá *ortogonální*, pokud $\mathbf{u} \perp \mathbf{v}$ pro libovolné dva různé prvky množiny (nebo posloupnosti) M.

Množina (posloupnost) M se nazývá *ortonormální*, pokud je ortogonální a každý vektor v M je jednotkový.

Z vlastnosti (SCS) plyne, že ortogonalita dvou prvků nezávisí na jejich pořadí. Z vlastnosti (SL1) vidíme, že jsou-li dva prvky kolmé, pak jsou kolmé i jejich libovolné násobky. Máme-li ortogonální množinu nenulových prvků $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$, můžeme z ní vytvořit ortonormální množinu znormováním, tj.

$$\left\{ \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|}, \frac{\mathbf{v}_2}{\|\mathbf{v}_2\|}, \dots, \frac{\mathbf{v}_k}{\|\mathbf{v}_k\|} \right\}$$

je ortonormální.

Z geometrického náhledu v $\mathbb{R}^3$ vidíme, že ortogonální posloupnost nenulových vektorů je lineárně nezávislá. Platí to zcela obecně.

Tvrzení 8.33. *Je-li* $\mathbf{V}$ *lineární prostor se skalárním součinem* $\langle\, , \rangle$*, pak každá ortogonální posloupnost nenulových prvků* V *je lineárně nezávislá.*

Důkaz. Je-li $(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k)$ ortogonální posloupnost prvků V a platí-li

$$a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_k\mathbf{v}_k = \mathbf{o}\ ,$$

pak skalárním vynásobením obou stran zleva vektorem $\mathbf{v}_i$ ($i \in \{1, 2, \dots, k\}$) a využitím (SL1), (SL2) a kolmosti dostáváme

$$\begin{aligned} \langle\, \mathbf{v}_i, a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_k\mathbf{v}_k\rangle &= \langle\, \mathbf{o}, \mathbf{v}\rangle \\ a_1 \langle\, \mathbf{v}_i, \mathbf{v}_1\rangle + a_2 \langle\, \mathbf{v}_i, \mathbf{v}_2\rangle + \cdots + a_k \langle\, \mathbf{v}_i, \mathbf{v}_k\rangle &= 0 \\ a_i \langle\, \mathbf{v}_i, \mathbf{v}_i\rangle &= 0\ . \end{aligned}$$

Protože vektor $\mathbf{v}_i$ je nenulový, platí podle (SP) vztah $\langle\, \mathbf{v}_i, \mathbf{v}_i\rangle = \|\mathbf{v}_i\|^2 > 0$, takže z odvozeného vztahu vyplývá $a_i = 0$. Ukázali jsme tak, že jediná lineární kombinace prvků $\mathbf{v}_i$, která dává nulový vektor, je triviální takže posloupnost $(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k)$ je lineárně nezávislá (viz bod (3) tvrzení 5.30). □

Z tvrzení vyplývá, že ortogonální posloupnost n nenulových vektorů v prostoru dimenze n je ortogonální báze, protože je lineárně nezávislá a lineárně nezávislá posloupnost n vektorů je báze podle bodu (4) v pozorování 5.61

Příklad 8.34. V prostoru $\mathbb{R}^n$ (nebo $\mathbb{C}^n$) se standardním skalárním součinem je kanonická báze ortonormální.

Posloupnost vektorů $((1, 2, 2)^T, (-2, -1, 2)^T)$ v $\mathbb{R}^3$ (nebo $\mathbb{C}^3$) se standardním skalárním součinem je ortogonální, ale není ortonormální. Znormováním dostaneme ortonormální posloupnost

$$\left(\frac{1}{3}(1, 2, 2)^T, \frac{1}{3}(-2, -1, 2)^T \right)\ .$$

Tuto posloupnost lze doplnit na ortonormální bázi – posloupnost

$$\left(\frac{1}{3}(1,2,2)^T, \frac{1}{3}(-2,-1,2)^T, \frac{1}{3}(2,-2,1)\right)$$

je ortonormální, takže je to podle poznámky za předchozím tvrzením ortonormální báze. Později budeme pomocí Gram-Schmidtova ortogonalizačního procesu umět každou ortogonální (resp. ortonormální) posloupnost nenulových vektorů v konečně generovaném prostoru doplnit do ortogonální (resp. ortonormální) báze.

Příklad 8.35. V prostoru $\mathbb{R}^2$ se skalárním součinem daným

$$\left\langle (x_1,x_2)^T, (y_1,y_2)\right\rangle = (x_1,x_2)\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = 2x_1y_1 + x_1y_2 + x_2y_1 + x_2y_2$$

(ověřte, že je to skutečně skalární součin) je posloupnost

$$\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 2 \end{pmatrix}\right)$$

ortogonální, protože

$$\left\langle (1,0)^T, (-1,2)^T\right\rangle = (1,0)\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} -1 \\ 2 \end{pmatrix} = (2,1)\begin{pmatrix} -1 \\ 2 \end{pmatrix} = 0\ ,$$

tedy tvoří ortogonální bázi. Spočítáme normy vektorů a vytvoříme ortonormální bázi.

$$\left\|\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right\| = \sqrt{(1,0)\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix}} = \sqrt{(2,1)\begin{pmatrix} 1 \\ 0 \end{pmatrix}} = \sqrt{2}$$

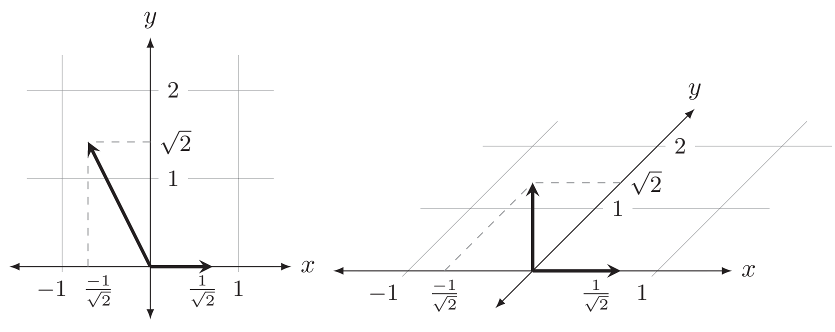
$$\left\|\begin{pmatrix} -1 \\ 2 \end{pmatrix}\right\| = \sqrt{(-1,2)\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} -1 \\ 2 \end{pmatrix}} = \sqrt{(0,1)\begin{pmatrix} -1 \\ 2 \end{pmatrix}} = \sqrt{2}$$

Posloupnost

$$\left(\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \frac{1}{\sqrt{2}}\begin{pmatrix} -1 \\ 2 \end{pmatrix}\right)$$

je tedy ortonormální báze.

Pokud si nakreslíme tyto dva vektory jako kolmé vektory jednotkové velikosti a ostatní vektory kreslíme v tomto souřadném systému, pak délky a úhly při daném skalárním součinu jsou běžné eukleidovské délky a úhly na obrázku. Tento fakt dokážeme v tvrzení 8.40.



Příklad 8.36. V prostoru spojitých funkcí na intervalu $[0, 2\pi]$ se skalárním součinem

$$\langle f, g \rangle = \int_0^{2\pi} fg$$

je množina $\{1, \sin x, \cos x, \sin(2x), \cos(2x), \dots\}$ ortogonální. Toto je základní fakt Fourierovy analýzy, jedné z velmi důležitých oblastí matematiky.

Jednoduchým důsledkem definice kolmosti je zobecnění Pythagorovy věty pro libovolný skalární součin.

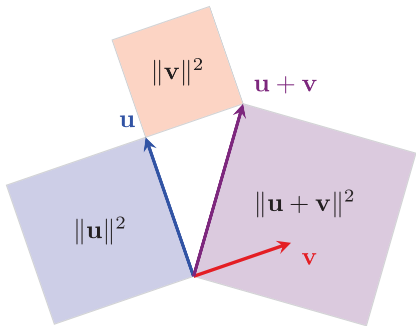
Tvrzení 8.37 (Pythagorova věta)**.** *Je-li* $\mathbf{V}$ *lineární prostor se skalárním součinem* $\langle \, , \rangle$ *a jsou-li vektory* $\mathbf{u}, \mathbf{v} \in V$ *kolmé, pak platí*

$$\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 \quad .$$

Důkaz.

$$\|\mathbf{u} + \mathbf{v}\|^2 = \langle \mathbf{u} + \mathbf{v}, \mathbf{u} + \mathbf{v} \rangle = \langle \mathbf{u}, \mathbf{u} \rangle + \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{v}, \mathbf{u} \rangle + \langle \mathbf{v}, \mathbf{v} \rangle$$

Díky kolmosti jsou prostřední dva členy nulové, takže výraz je roven $\|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$. □



Indukcí lze Pythagorovu větu zobecnit na libovolný konečný počet prvků. Je-li $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ ortogonální množina, pak

$$\|\mathbf{v}_1 + \mathbf{v}_2 + \cdots + \mathbf{v}_k\|^2 = \|\mathbf{v}_1\|^2 + \|\mathbf{v}_2\|^2 + \cdots + \|\mathbf{v}_k\|^2 \quad .$$

Zobecnění této rovnosti na nekonečné ortogonální množiny prvků lineárního prostoru se skalárním součinem se někdy říká *Parsevalova identita.*

8.3.1. *Souřadnice prvku vzhledem k ortonormální bázi.* Souřadnice prvků lineárního prostoru vzhledem k ortonormální bázi se počítají velmi snadno.

Tvrzení 8.38. *Je-li* $\mathbf{V}$ *lineární prostor se skalárním součinem* $\langle \, , \rangle$, $B = (\mathbf{v}_1, \dots, \mathbf{v}_n)$ *nějaká ortonormální báze ve* $\mathbf{V}$ *a* $\mathbf{u} \in V$*, pak platí*

$$\mathbf{u} = \langle \mathbf{v}_1, \mathbf{u} \rangle \mathbf{v}_1 + \langle \mathbf{v}_2, \mathbf{u} \rangle \mathbf{v}_2 + \cdots + \langle \mathbf{v}_n, \mathbf{u} \rangle \mathbf{v}_n \quad .$$

Jinými slovy,

$$[\mathbf{u}]_B = (\langle \mathbf{v}_1, \mathbf{u} \rangle, \langle \mathbf{v}_2, \mathbf{u} \rangle, \dots, \langle \mathbf{v}_n, \mathbf{u} \rangle)^T \quad .$$

Důkaz. Označme $[\mathbf{u}]_B = (a_1, a_2, \dots, a_n)^T$, neboli

$$\mathbf{u} = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_n\mathbf{v}_n \ .$$

Podobně jako v důkazu lineární nezávislosti ortogonální množiny nenulových vektorů skalárně vynásobíme obě strany zleva vektorem $\mathbf{v}_i$ a dostaneme

$$\begin{aligned}
\langle \mathbf{v}_i, \mathbf{u} \rangle &= \langle \mathbf{v}_i, a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_k\mathbf{v}_k \rangle \\
\langle \mathbf{v}_i, \mathbf{u} \rangle &= a_1 \langle \mathbf{v}_i, \mathbf{v}_1 \rangle + a_2 \langle \mathbf{v}_i, \mathbf{v}_2 \rangle + \cdots + a_k \langle \mathbf{v}_i, \mathbf{v}_k \rangle \\
\langle \mathbf{v}_i, \mathbf{u} \rangle &= a_i \langle \mathbf{v}_i, \mathbf{v}_i \rangle = a_i \ ,
\end{aligned}$$

takže $a_i = \langle \mathbf{v}_i, \mathbf{u} \rangle$. □

Souřadnicím vzhledem k ortonormální bázi se někdy říká *Fourierovy koeficienty* vzhledem k této bázi. Obecněji z důkazu vidíme, že pro ortogonální bázi B platí

$$[\mathbf{u}]_B = \left(\frac{\langle \mathbf{v}_1, \mathbf{u} \rangle}{\|\mathbf{v}_1\|^2}, \frac{\langle \mathbf{v}_2, \mathbf{u} \rangle}{\|\mathbf{v}_2\|^2}, \dots, \frac{\langle \mathbf{v}_n, \mathbf{u} \rangle}{\|\mathbf{v}_n\|^2} \right)^T \ .$$

Příklad 8.39. Určíme souřadnice vektoru $\mathbf{u} = (3+i, 2, i)^T \in \mathbb{C}^3$ vzhledem k ortonormální bázi

$$B = (\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3) = \left(\frac{1}{3} \begin{pmatrix} i \\ 2i \\ 2i \end{pmatrix}, \frac{1}{3} \begin{pmatrix} -2 \\ -1 \\ 2 \end{pmatrix}, \frac{1}{3} \begin{pmatrix} 2 \\ -2 \\ 1 \end{pmatrix} \right)$$

prostoru $\mathbb{C}^3$ se standardním skalárním skalárním součinem.

$$\begin{aligned}
[\mathbf{u}]_B &= (\mathbf{v}_1^*\mathbf{u}, \mathbf{v}_2^*\mathbf{u}, \mathbf{v}_3^*\mathbf{u})^T \\
&= \left(\frac{1}{3}(-i, -2i, -2i) \begin{pmatrix} 3+i \\ 2 \\ i \end{pmatrix}, \frac{1}{3}(-2, -1, 2) \begin{pmatrix} 3+i \\ 2 \\ i \end{pmatrix}, \right. \\
&\qquad \left. \frac{1}{3}(2, -2, 1) \begin{pmatrix} 3+i \\ 2 \\ i \end{pmatrix} \right)^T \\
&= \left(\frac{1}{3}(3-7i), -\frac{8}{3}, \frac{1}{3}(2+3i) \right)^T \ .
\end{aligned}$$

Skutečně

$$\begin{pmatrix} 3+i \\ 2 \\ i \end{pmatrix} = \frac{1}{3}(3-7i) \cdot \frac{1}{3} \begin{pmatrix} i \\ 2i \\ 2i \end{pmatrix} - \frac{8}{3} \cdot \frac{1}{3} \begin{pmatrix} -2 \\ -1 \\ 2 \end{pmatrix} + \frac{1}{3}(2+3i) \cdot \frac{1}{3} \begin{pmatrix} 2 \\ -2 \\ 1 \end{pmatrix} \ .$$

Vzhledem k ortonormální bázi přechází skalární součin na standardní. Přesněji řečeno, skalární součin dvou vektorů je roven standardnímu skalárnímu součinu souřadnic těchto vektorů vzhledem k ortonormální bázi.

Tvrzení 8.40. *Je-li* $\mathbf{V}$ *lineární prostor se skalárním součinem* $\langle \ , \ \rangle$, $B = (\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n)$ *jeho ortonormální báze, a* $\mathbf{u}, \mathbf{w} \in V$, *pak*

$$\langle \mathbf{u}, \mathbf{w} \rangle = [\mathbf{u}]_B^* [\mathbf{w}]_B \ .$$

Důkaz. Označme $[\mathbf{u}]_B = (a_1, a_2, \dots, a_n)^T$, $[\mathbf{w}]_B = (b_1, b_2, \dots, b_n)^T$, tedy

$$\mathbf{u} = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_n\mathbf{v}_n, \quad \mathbf{w} = b_1\mathbf{v}_1 + b_2\mathbf{v}_2 + \cdots + b_n\mathbf{v}_n \ .$$

Pomocí (SL2), (SL1) a ortonormality postupně dostáváme

$$\begin{aligned}\langle\, \mathbf{u}, \mathbf{w} \rangle &= \left\langle \sum_{i=1}^{n} a_i\mathbf{v}_i, \sum_{j=1}^{n} b_i\mathbf{v}_i \right\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n} \langle\, a_i\mathbf{v}_i, b_j\mathbf{v}_j \rangle \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} \overline{a_i} b_j \,\langle\, \mathbf{v}_i, \mathbf{v}_j \rangle = \sum_{i=1}^{n} \overline{a_i} b_i = [\mathbf{u}]_B^*[\mathbf{w}]_B \ .\end{aligned}$$

□

Tvrzení ospravedlňuje poznámku z příkladu 8.35 – pokud si nakreslíme vektory ortonormální báze jako jednotkové navzájem kolmé vektory a ostatní vektory kreslíme v tomto souřadném systému, pak délky a úhly při daném skalárním součinu jsou běžné eukleidovské délky a úhly na obrázku.

Příklad 8.41. V prostoru $\mathbb{R}^2$ se skalárním součinem

$$\left\langle\, (x_1, x_2)^T, (y_1, y_2) \right\rangle^T = (x_1, x_2) \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = 2x_1y_1 + x_1y_2 + x_2y_1 + x_2y_2$$

je posloupnost

$$B = \left(\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 2 \end{pmatrix} \right)$$

ortonormální báze (viz příklad 8.35. Uvažujme vektory $\mathbf{u} = (2,3)^T$ a $\mathbf{v} = (1,1)^T$. Z tvrzení 8.38 spočteme jejich souřadnice vzhledem k B a pak vypočítáme skalární součin podle tvrzení 8.40.

$$\begin{aligned}[\mathbf{u}]_B &= \left(\left\langle \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \mathbf{u} \right\rangle, \left\langle \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 2 \end{pmatrix}, \mathbf{u} \right\rangle \right)^T = \frac{1}{\sqrt{2}} \begin{pmatrix} 7 \\ 3 \end{pmatrix} \\ [\mathbf{v}]_B &= \left(\left\langle \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \mathbf{v} \right\rangle, \left\langle \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 2 \end{pmatrix}, \mathbf{v} \right\rangle \right)^T = \frac{1}{\sqrt{2}} \begin{pmatrix} 3 \\ 1 \end{pmatrix} \\ \langle\, \mathbf{u}, \mathbf{v} \rangle = [\mathbf{u}]_B \cdot [\mathbf{v}]_B &= \frac{1}{\sqrt{2}}(7,3)\ \frac{1}{\sqrt{2}} \begin{pmatrix} 3 \\ 1 \end{pmatrix} = 12 \ .\end{aligned}$$

Stejný výsledek dostaneme přímo z definice skalárního součinu v $\mathbb{R}^2$.

Závěrem této části zobecníme kolmost mezi prvky lineárního prostoru se skalárním součinem na podmnožiny tohoto prostoru.

Definice 8.42. Je-li $\mathbf{V}$ lineární prostor se skalárním součinem $\langle\, , \rangle$ a $\mathbf{v} \in V$, $M, N \subseteq V$, pak říkáme, že prvek $\mathbf{v}$ *je kolmý na* M, pokud $\mathbf{v}$ je kolmý na každý prvek z množiny M, což zapisujeme $\mathbf{v} \perp M$.

Říkáme, že M *je kolmá na* N a zapisujeme $M \perp N$, pokud každý prvek množiny M je kolmý na každý prvek množiny N.

Jednoduchým důsledkem definice je následující pozorování.

Pozorování 8.43. *Je-li* $\mathbf{V}$ *lineární prostor se skalárním součinem* $\langle\, , \rangle$ *a* $M, N \subseteq V$, *pak* $M \perp N$ *právě když* $M \perp \langle N \rangle$ *což je právě když* $\langle M \rangle \perp \langle N \rangle$.

Důkaz. Dokážeme ekvivalenci prvních dvou tvrzení. Předpokládáme tedy $M \perp N$. Je-li $\mathbf{x} \in M$ a $\mathbf{y} \in \langle N \rangle$, existuje vyjádření $\mathbf{y} = a_1\mathbf{u}_1 + a_2\mathbf{u}_2 + \cdots + a_k\mathbf{u}_k$ pro nějaké prvky $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k \in N$ a skaláry $a_1, a_2, \ldots, a_k$. Potom

$$\begin{aligned}\langle \mathbf{x}, \mathbf{y} \rangle &= \langle \mathbf{x}, a_1\mathbf{u}_1 + a_2\mathbf{u}_2 + \cdots + a_k\mathbf{u}_k \rangle \\ &= a_1 \langle \mathbf{x}, \mathbf{u}_1 \rangle + a_2 \langle \mathbf{x}, \mathbf{u}_2 \rangle + \cdots + a_k \langle \mathbf{x}, \mathbf{u}_k \rangle = 0 \ .\end{aligned}$$

Proto $M \perp \langle N \rangle$. Opačná implikace je zřejmá. Ekvivalence druhých dvou tvrzení plyne z ekvivalence prvních dvou. □

8.4. **Gramova-Schmidtova ortogonalizace, QR-rozklad.**

V této části se seznámíme s jedním z nejdůležitějších algoritmů v lineární algebře. Jeho důležitost je srovnatelná s Gaussovo eliminací. Základem algoritmu je pojem projekce vektoru na podprostor.

Definice 8.44. Je-li $\mathbf{V}$ lineární prostor se skalárním součinem $\langle\, , \rangle$, $\mathbf{v} \in V$ a $\mathbf{W}$ podprostor $\mathbf{V}$, pak prvek $\mathbf{w} \in W$ nazýváme *ortogonální projekce* $\mathbf{v}$ *na podprostor* $\mathbf{W}$, pokud platí $(\mathbf{v} - \mathbf{w}) \perp W$.

Z definice snadno odvodíme, že pokud $\mathbf{v} \in W$, pak ortogonální projekcí $\mathbf{w}$ na W je vektor $\mathbf{w} = \mathbf{v}$. Platí totiž $\mathbf{v} - \mathbf{w} = \mathbf{o} \perp W$.

Následující jednoduché tvrzení dokazuje intuitivně zřejmý fakt, že ortogonální projekce vektoru na podprostor, pokud existuje, je určená jednoznačně a minimalizuje vzdálenost prvku $\mathbf{v}$ od prvků podprostoru $\mathbf{W}$.

Tvrzení 8.45. *Je-li* $\mathbf{W}$ *podprostor lineárního prostoru* $\mathbf{V}$ *se skalárním součinem* $\langle\, , \rangle$, $\mathbf{v} \in V$ *a* $\mathbf{w}$ *ortogonální projekce prvku* $\mathbf{v}$ *na podprostor* $\mathbf{W}$*, pak pro každý prvek* $\mathbf{w} \neq \mathbf{u} \in W$ *platí*

$$\|\mathbf{v} - \mathbf{w}\| < \|\mathbf{v} - \mathbf{u}\| \ .$$

Ortogonální projekce $\mathbf{v}$ *na podprostor* $\mathbf{W}$ *je určena jednoznačně, pokud existuje.*

Důkaz. Napřed budeme předpokládat, že $\mathbf{v} \notin W$. Protože $\mathbf{u} \neq \mathbf{w}$, je $\mathbf{o} \neq \mathbf{w} - \mathbf{u} \in W$. Protože $\mathbf{w} \neq \mathbf{v}$, je $\mathbf{v} - \mathbf{w} \neq \mathbf{o}$.

Protože $\mathbf{w}$ je ortogonální projekce $\mathbf{v}$ na W, je $(\mathbf{v} - \mathbf{w}) \perp W$, speciálně platí $(\mathbf{v} - \mathbf{w}) \perp (\mathbf{w} - \mathbf{u})$. Protože jsou oba vektory nenulové, můžeme použít Pythagorovu větu a z té plyne

$$\|\mathbf{v} - \mathbf{u}\|^2 = \|(\mathbf{v} - \mathbf{w}) + (\mathbf{w} - \mathbf{u})\|^2 = \|\mathbf{v} - \mathbf{w}\|^2 + \|\mathbf{w} - \mathbf{u}\|^2 > \|\mathbf{v} - \mathbf{w}\|^2 \ .$$

Kdyby oba vektory $\mathbf{w}, \mathbf{u} \in W$ byly ortogonálními projekcemi $\mathbf{v}$ na W, platilo by podle právě dokázaného

$$\|\mathbf{v} - \mathbf{w}\| < \|\mathbf{v} - \mathbf{u}\| < \|\mathbf{v} - \mathbf{w}\| \ ,$$

což nelze. Ortognální projekce $\mathbf{v}$ na W je proto určena jednoznačně.

Případ $\mathbf{v} \in W$ ponecháme jako cvičení. □

Zbývá dokázat, kdy ortogonální projekce $\mathbf{w}$ prvku $\mathbf{v} \in \mathbf{V}$ na podprostor $\mathbf{W} \leq \mathbf{V}$ existuje. Víme už, že existuje, pokud $\mathbf{v} \in W$, a v tom případě $\mathbf{w} = \mathbf{v}$.

Tvrzení 8.46. *Je-li* $\mathbf{V}$ *lineární prostor se skalárním součinem* $\langle\, , \rangle$, $\mathbf{v} \in V$*, a* $\mathbf{W}$ *konečně generovaný podprostor* $\mathbf{V}$ *s ortonormální bází* $(\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k)^T$*, pak prvek*

$$\mathbf{w} = \langle \mathbf{u}_1, \mathbf{v} \rangle \mathbf{u}_1 + \langle \mathbf{u}_2, \mathbf{v} \rangle \mathbf{u}_2 + \cdots + \langle \mathbf{u}_k, \mathbf{v} \rangle \mathbf{u}_k$$

je ortogonální projekcí vektoru $\mathbf{v}$ *na podprostor* $\mathbf{W}$.

Důkaz. Libovolný prvek $\mathbf{w}$ podprostoru $\mathbf{W}$ můžeme vyjádřit jako lineární kombinaci prvků ortonormální báze $(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k)$. Prvek

$$\mathbf{w} = a_1\mathbf{u}_1 + a_2\mathbf{u}_2 + \cdots + a_k\mathbf{u}_k$$

je ortogonální projekcí prvku $\mathbf{v}$ na podprostor $W = \langle \mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k \rangle$ právě když je prvek $\mathbf{v} - \mathbf{w}$ ortogonální ke každému prvku podprostoru $\mathbf{W}$ a to podle pozorování 8.43 nastává právě když je ortogonální ke každému prvku $\mathbf{u}_i$. Pro každé $i = 1, 2, \dots, k$ je prvek $\mathbf{u}_i$ kolmý k prvku $\mathbf{v} - \mathbf{w}$ právě když

$$\begin{aligned}
0 &= \langle \mathbf{u}_i, \mathbf{v} - \mathbf{w} \rangle = \langle \mathbf{u}_i, \mathbf{v} \rangle - \langle \mathbf{u}_i, \mathbf{w} \rangle \\
&= \langle \mathbf{u}_i, \mathbf{v} \rangle - \langle \mathbf{u}_i, a_1\mathbf{u}_1 + a_2\mathbf{u}_2 + \cdots + a_k\mathbf{u}_k \rangle \\
&= \langle \mathbf{u}_i, \mathbf{v} \rangle - a_1 \langle \mathbf{u}_i, \mathbf{u}_1 \rangle - a_2 \langle \mathbf{u}_i, \mathbf{u}_2 \rangle - \cdots - a_k \langle \mathbf{u}_i, \mathbf{u}_k \rangle \\
&= \langle \mathbf{u}_i, \mathbf{v} \rangle - a_i \langle \mathbf{u}_i, \mathbf{u_i} \rangle = \langle \mathbf{u}_i, \mathbf{v} \rangle - a_i \ ,
\end{aligned}$$

tj. právě když $a_i = \langle \mathbf{u}_i, \mathbf{v} \rangle$. □

Všimněme si, že vzorec pro výpočet ortogonální projekce $\mathbf{v}$ na podprostor $\mathbf{W}$ s ortonormální bází je stejný jako vyjádření libovolného prvku lineárního prostoru $\mathbf{V}$ se skalárním součinem jako lineární kombinace prvků ortonormální báze ve $\mathbf{V}$ v tvrzení 8.38. Ve skutečnosti je tvrzení 8.38 speciálním případem předešlého tvrzení, kdy $\mathbf{v} \in \mathbf{W}$.

Pokud báze $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k$ v podprostoru $\mathbf{W}$ není ortonormální, ale pouze ortogonální, napřed ji normalizujeme

$$\frac{\mathbf{u}_1}{\|\mathbf{u}_1\|}, \frac{\mathbf{u}_2}{\|\mathbf{u}_2\|}, \dots, \frac{\mathbf{u}_k}{\|\mathbf{u}_k\|}$$

a pak použijeme předchozí tvrzení. Dostaneme tak vyjádření ortogonální projekce $\mathbf{w}$ prvku $\mathbf{v}$ na podprostor $\mathbf{W}$ ve tvaru

$$\begin{aligned}
\mathbf{w} &= \left\langle \frac{\mathbf{u}_1}{\|\mathbf{u}_1\|}, \mathbf{v} \right\rangle \frac{\mathbf{u}_1}{\|\mathbf{u}_1\|} + \left\langle \frac{\mathbf{u}_2}{\|\mathbf{u}_2\|}, \mathbf{v} \right\rangle \frac{\mathbf{u}_2}{\|\mathbf{u}_2\|} + \cdots + \left\langle \frac{\mathbf{u}_k}{\|\mathbf{u}_k\|}, \mathbf{v} \right\rangle \frac{\mathbf{u}_k}{\|\mathbf{u}_k\|} \\
&= \frac{\langle \mathbf{u}_1, \mathbf{v} \rangle}{\|\mathbf{u}_1\|^2} \mathbf{u}_1 + \frac{\langle \mathbf{u}_2, \mathbf{v} \rangle}{\|\mathbf{u}_2\|^2} \mathbf{u}_2 + \cdots + \frac{\langle \mathbf{u}_k, \mathbf{v} \rangle}{\|\mathbf{u}_k\|^2} \mathbf{u}_k \ .
\end{aligned}$$

Dokázali jsme tak následující důsledek.

Důsledek 8.47. *Je-li* $\mathbf{V}$ *lineární prostor se skalárním součinem* $\langle\, , \rangle$, $\mathbf{v} \in V$, *a* $\mathbf{W}$ *konečně generovaný podprostor* $\mathbf{V}$ *s ortogonální bází* $B = (\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k)^T$, *pak prvek*

$$\mathbf{w} = \frac{\langle \mathbf{u}_1, \mathbf{v} \rangle}{\|\mathbf{u}_1\|^2} \mathbf{u}_1 + \frac{\langle \mathbf{u}_2, \mathbf{v} \rangle}{\|\mathbf{u}_2\|^2} \mathbf{u}_2 + \cdots + \frac{\langle \mathbf{u}_k, \mathbf{v} \rangle}{\|\mathbf{u}_k\|^2} \mathbf{u}_k$$

je ortogonální projekcí vektoru $\mathbf{v}$ *na podprostor* $\mathbf{W}$.

Pomocí souřadnic vzhledem k ortogonální bázi B poslední důsledek zapíšeme ve tvaru

$$[\mathbf{w}]_B = \left(\frac{\langle \mathbf{u}_1, \mathbf{v} \rangle}{\|\mathbf{u}_1\|^2}, \frac{\langle \mathbf{u}_2, \mathbf{v} \rangle}{\|\mathbf{u}_2\|^2}, \dots, \frac{\langle \mathbf{u}_k, \mathbf{v} \rangle}{\|\mathbf{u}_k\|^2} \right) \ .$$

V případě podprostoru $\langle \mathbf{a} \rangle$ dimenze 1 dostáváme projekci $\mathbf{w}$ libovolného prvku $\mathbf{v} \in \mathbf{V}$ na podprostor $\langle \mathbf{a}_1 \rangle$ jako

$$\mathbf{w} = \frac{\langle \mathbf{a}, \mathbf{v} \rangle}{\|\mathbf{a}\|^2} \mathbf{a} \ .$$

Příklad 8.48. V aritmetickém prostoru $\mathbb{R}^3$ se standardním skalárním součinem je $((1,1,2)^T, (2,0,-1)^T)$ ortogonální posloupnost. Ortogonální projekce $\mathbf{w}$ vektoru $\mathbf{v} = (1,2,3)^T$ na rovinu $W = \left\langle (1,1,2)^T, (2,0,-1)^T \right\rangle$ je tedy

$$\begin{aligned}
\mathbf{w} &= \frac{(1,1,2)(1,2,3)^T}{(1,1,2)(1,1,2)^T} \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} + \frac{(2,0,-1)(1,2,3)^T}{(2,0,-1)(2,0,-1)^T} \begin{pmatrix} 2 \\ 0 \\ -1 \end{pmatrix} \\
&= \frac{9}{6} \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} - \frac{1}{5} \begin{pmatrix} 2 \\ 0 \\ -1 \end{pmatrix} = \frac{1}{10} \begin{pmatrix} 11 \\ 15 \\ 32 \end{pmatrix} \quad .
\end{aligned}$$

8.4.1. *Gramova-Schmidtova ortogonalizace.* Nyní dokážeme, že v každém konečně generovaném lineárním prostoru se skalárním součinem existuje ortonormální báze. Existenci ortonormální báze dokážeme pomocí algoritmu, kterému se říká *Gramova-Schmidtova ortogonalizace.*

Tento algoritmus dostane *na vstupu* nějakou lineárně nezávislou posloupnost

$$(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k)$$

prvků lineárního prostoru $\mathbf{V}$ se skalárním součinem. *Na výstupu* vydá ortonormální posloupnost

$$(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k)$$

prvků prostoru $\mathbf{V}$, která splňuje podmínku

$$\langle \mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_i \rangle = \langle \mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_i \rangle$$

pro každé $i = 1, 2, \dots, k$.

První krok je jednoduchý - normalizujeme vektor $\mathbf{v}_1$, tj. položíme

$$\mathbf{u}_i = \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|^2} \quad ,$$

pak platí také $\langle \mathbf{u}_1 \rangle = \langle \mathbf{v}_1 \rangle$.

V druhém kroku najdeme ortogonální projekci $\mathbf{w}_1$ vektoru $\mathbf{v}_2$ na podprostor $\langle \mathbf{u}_1 \rangle = \langle \mathbf{v}_1 \rangle$, podle tvrzení8.46 platí

$$\mathbf{w}_1 = \langle\, \mathbf{u}_1, \mathbf{v}_2 \rangle\ \mathbf{u}_1 \quad ,$$

což podle definice 8.42 znamená $(\mathbf{v}_2 - \mathbf{w}_1) \perp \langle \mathbf{u}_1 \rangle$ a tedy $(\mathbf{v}_2 - \mathbf{w}_1) \perp \mathbf{u}_1$. Položíme

$$\mathbf{u}_2 = \frac{\mathbf{v}_2 - \mathbf{w}_1}{\|\mathbf{v}_2 - \mathbf{w}_1\|} \quad .$$

Potom platí, že $\mathbf{u}_2 \perp \mathbf{u}_1$ a $(\mathbf{u}_1, \mathbf{u}_2)$ je tedy ortonormální posloupnost, pro kterou platí rovnost lineárních obalů

$$\langle \mathbf{u}_1, \mathbf{u}_2 \rangle = \langle \mathbf{u}_1, \mathbf{v}_2 \rangle = \langle \mathbf{v}_1, \mathbf{v}_2 \rangle \quad .$$

Konstrukce prvku $\mathbf{u}_2$ je speciálním případem indukčního kroku. Předpokládáme, že po nějaké $i \leq k$ jsme již sestrojili ortonormální posloupnost $(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_{i-1})$ takovou, že $\langle \mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_{i-1} \rangle = \langle \mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{i-1} \rangle$. Najdeme ortogonální projekci $\mathbf{w}_{i-1}$ prvku $\mathbf{v}_i$ na podprostor $\langle \mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{i-1} \rangle$. Podle indukčního předpokladu je $(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_{i-1})$ ortonormální báze podprostoru $\langle \mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{i-1} \rangle$. Podle tvrzení 8.46 platí

$$\mathbf{w}_{i-1} = \langle\, \mathbf{u}_1, \mathbf{v}_i \rangle\ \mathbf{u}_1 + \langle\, \mathbf{u}_2, \mathbf{v}_i \rangle\ \mathbf{u}_2 + \dots + \langle\, \mathbf{u}_1, \mathbf{v}_{i-1} \rangle\ \mathbf{u}_{i-1} \quad .$$

Z definice ortogonální projekce prvku na podprostor plyne

$$(\mathbf{v}_i - \mathbf{w}_{i-1}) \perp \langle \mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_{i-1} \rangle \supseteq \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_{i-1}\}$$

OBRAZEK N6 - indukcni krok v GS

Dále platí $\mathbf{w}_{i-1} \in \langle \mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_{i-1} \rangle$. Vzhledem k tomu, že původní posloupnost $(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k)$ je lineárně nezávislá, platí $\mathbf{v}_i \notin \langle \mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{i-1} \rangle$. A protože $\langle \mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{i-1} \rangle = \langle \mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_{i-1} \rangle$ podle indukčního předpokladu, plyne odtud také $\mathbf{v}_i \notin \langle \mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_{i-1} \rangle$. Proto $\mathbf{v}_i \neq \mathbf{w}_{i-1}$ a $\|\mathbf{v}_i - \mathbf{w}_{i-1}\| \neq 0$. Můžeme proto položit

$$\mathbf{u}_i = \frac{\mathbf{v}_i - \mathbf{w}_{i-1}}{\|\mathbf{v}_i - \mathbf{w}_{i-1}\|} \quad .$$

Posloupnost $(\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_{i-1}, \mathbf{u}_i)$ je proto ortonormální.

Zbývá dokázat rovnost

$$\langle \mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_i \rangle = \langle \mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_i \rangle \quad .$$

Z indukčního předpokladu $\langle \mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_{i-1} \rangle = \langle \mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{i-1} \rangle$ plyne rovněž

$$\langle \mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_{i-1}, \mathbf{v}_i \rangle = \langle \mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{i-1}, \mathbf{v}_i \rangle \quad .$$

Protože $\mathbf{w}_{i-1} \in \langle \mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_i \rangle$, platí $\mathbf{v}_i - \mathbf{w}_{i-1} \in \langle \mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_{i-1}, \mathbf{v}_i \rangle$ a tedy také $\mathbf{u}_i \in \langle \mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_{i-1}, \mathbf{v}_i \rangle$, což dokazuje inkluzi

$$\langle \mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_{i-1}, \mathbf{u}_i \rangle \subseteq \langle \mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_{i-1}, \mathbf{v}_i \rangle = \langle \mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{i-1}, \mathbf{v}_i \rangle \quad .$$

K důkazu opačné inkluze si stačí uvědomit, že $\mathbf{v}_i = \|\mathbf{v}_i - \mathbf{w}_{i-1}\| \, \mathbf{u}_i + \mathbf{w}_{i-1}$, a proto opětovným použitím indukčního předpokladu dostáváme

$$\mathbf{v}_i \in \langle \mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_{i-1}, \mathbf{u}_i \rangle \quad .$$

Odtud plyne opačná inkluze

$$\langle \mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{i-1}, \mathbf{v}_i \rangle = \langle \mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_{i-1}, \mathbf{v}_i \rangle \subseteq \langle \mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_{i-1}, \mathbf{u}_i \rangle \quad .$$

Celá Gramova-Schmidtova ortogonalizace spočívá v k-násobném iterování cyklu, jehož i-tý průběh sestává ze dvou kroků

(ia) **ortogonalizace**: najdeme prvek

$$\mathbf{v}_i - \mathbf{w}_{i-1} = \mathbf{v}_i - \langle\, \mathbf{u}_1, \mathbf{v}_i \rangle \, \mathbf{u}_1 - \langle\, \mathbf{u}_2, \mathbf{v}_i \rangle \, \mathbf{u}_2 - \cdots - \langle\, \mathbf{u}_1, \mathbf{v}_{i-1} \rangle \, \mathbf{u}_{i-1},$$

(ib) **normalizace**: položíme

$$\mathbf{u}_i = \frac{\mathbf{v}_i - \mathbf{w}_{i-1}}{\|\mathbf{v}_i - \mathbf{w}_{i-1}\|} \quad .$$

V prvním cyklu můžeme krok (1a) vynechat a začít přímo normalizačním krokem (1b), neboť v tom případě $\mathbf{w}_0$ je ortogonální projekce prvku $\mathbf{v}_1$ na podprostor $\langle \emptyset \rangle = \{\mathbf{o}\}$, a tedy $\mathbf{w}_0 = \mathbf{o}$ a $\mathbf{v}_1 - \mathbf{w}_0 = \mathbf{v}_1$. Od druhého cyklu už musíme provést oba kroky - ortogonalizační a normalizační.

Během popisu Gramovy-Schmidtovy ortogonalizace jsem současně dokázali následující větu.

Věta 8.49. *Gramova-Schmidtova ortogonalizace převede libovolnou lineárně nezávislou posloupnost* $(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k)$ *prvků lineárního prostoru se skalárním součinem na ortonormální posloupnost* $(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k)$*, pro kterou platí*

$$\langle \mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_i \rangle = \langle \mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_i \rangle$$

pro každé $i = 1, 2, \dots, k$.

Pokud chceme najít pouze ortogonální bázi, stačí vynechat v algoritmu kroky (ib). V takovém případě hledáme ortogonální projekci $\mathbf{w}_{i-1}$ vektoru $\mathbf{i}$ na podprostor $\langle \mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_{i-1} \rangle$ s ortogonální bází $(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_{i-1})$ a k jejímu výpočtu musíme použít důsledek 8.47 misto tvrzení 8.46. Výpočtu norem vektorů $\mathbf{v}_i - \mathbf{w}_{i-1}$ se tím ale nevyhneme (s výjimkou toho posledního).

Příklad 8.50. V podprostoru

$$W = \langle \mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3 \rangle = \left\langle (1,2,0,1)^T, (1,-1,1,0)^T, (0,1,1,3)^T \right\rangle$$

prostoru $\mathbb{R}^4$ se standardním skalárním součinem najdeme ortonormální bázi $\mathbf{u}_1$, $\mathbf{u}_2$, $\mathbf{u}_3$. Použijeme Gramovu-Schmidtovou ortogonalizaci na posloupnost vektorů $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$. Postupně počítáme

$$\begin{aligned}
\mathbf{u}_1 &= \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|} = \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ 2 \\ 0 \\ 1 \end{pmatrix} , \\
\mathbf{w}_1 &= \langle \mathbf{u}_1, \mathbf{v}_2 \rangle \, \mathbf{u}_1 = \frac{1}{\sqrt{6}} (1,2,0,1) \begin{pmatrix} 1 \\ -1 \\ 1 \\ 0 \end{pmatrix} \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ 2 \\ 0 \\ 1 \end{pmatrix} = \frac{1}{6} \begin{pmatrix} 1 \\ 2 \\ 0 \\ 1 \end{pmatrix} , \\
\mathbf{v}_2 - \mathbf{w}_1 &= \begin{pmatrix} 1 \\ -1 \\ 1 \\ 0 \end{pmatrix} + \frac{1}{6} \begin{pmatrix} 1 \\ 2 \\ 0 \\ 1 \end{pmatrix} = \frac{1}{6} \begin{pmatrix} 7 \\ -4 \\ 6 \\ 1 \end{pmatrix} , \\
\mathbf{u}_2 &= \frac{\mathbf{v}_2 - \mathbf{w}_1}{\|\mathbf{v}_2 - \mathbf{w}_1\|} = \frac{1}{\sqrt{102}} \begin{pmatrix} 7 \\ -4 \\ 6 \\ 1 \end{pmatrix} .
\end{aligned}$$

$$\mathbf{w}_2 = \langle\, \mathbf{u}_1, \mathbf{v}_3 \rangle\, \mathbf{u}_1 + \langle\, \mathbf{u}_2, \mathbf{v}_3 \rangle\, \mathbf{u}_2$$

$$= \frac{1}{\sqrt{6}}(1,2,0,1) \begin{pmatrix} 0 \\ 1 \\ 1 \\ 3 \end{pmatrix} \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ 2 \\ 0 \\ 1 \end{pmatrix} + \frac{1}{\sqrt{102}}(7,-4,6,1) \begin{pmatrix} 0 \\ 1 \\ 1 \\ 3 \end{pmatrix} \frac{1}{\sqrt{102}} \begin{pmatrix} 7 \\ -4 \\ 6 \\ 1 \end{pmatrix}$$

$$= \frac{5}{6} \begin{pmatrix} 1 \\ 2 \\ 0 \\ 1 \end{pmatrix} + \frac{5}{102} \begin{pmatrix} 7 \\ -4 \\ 6 \\ 1 \end{pmatrix} = \frac{1}{102} \begin{pmatrix} 120 \\ 150 \\ 30 \\ 90 \end{pmatrix} = \frac{5}{17} \begin{pmatrix} 4 \\ 5 \\ 1 \\ 3 \end{pmatrix} ,$$

$$\mathbf{v}_3 - \mathbf{w}_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 3 \end{pmatrix} - \frac{5}{17} \begin{pmatrix} 4 \\ 5 \\ 1 \\ 3 \end{pmatrix} = \frac{4}{17} \begin{pmatrix} -5 \\ -2 \\ 3 \\ 9 \end{pmatrix} ,$$

$$\mathbf{u}_3 = \frac{\mathbf{v}_3 - \mathbf{w}_2}{\|\mathbf{v}_3 - \mathbf{w}_2\|} = \frac{1}{\sqrt{119}} \begin{pmatrix} -5 \\ -2 \\ 3 \\ 9 \end{pmatrix} .$$

Získali jsme tak ortonormální posloupnost

$$\left(\frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ 2 \\ 0 \\ 1 \end{pmatrix}, \frac{1}{\sqrt{102}} \begin{pmatrix} 7 \\ -4 \\ 6 \\ 1 \end{pmatrix}, \frac{1}{\sqrt{119}} \begin{pmatrix} -5 \\ -2 \\ 3 \\ 9 \end{pmatrix} \right) ,$$

která je ortonormální bází podprostoru $W = \left\langle (1,2,0,1)^T, (1,-1,1,0)^T, (0,1,1,3)^T \right\rangle$.

Při řešení předchozího příkladu jsem neověřovali předpoklad, že daná posloupnost $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ je lineárně nezávislá. Není to nutné, protože algoritmus pro Gramovu-Schmidtovu ortogonalizaci sám pozná, je-li některý z daných prvků $\mathbf{v}_i$ lineárně závislý na předchozích. Pokud by takový prvek $\mathbf{v}_i$ existoval, pro první z nich by platilo

$$\mathbf{v}_i \in \langle \mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{i-1} \rangle = \langle \mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_{i-1} \rangle \ .$$

Ortogonální projekce $\mathbf{w}_{i-1}$ prvku $\mathbf{v}_i$ na podprostor $\langle \mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_{i-1} \rangle$ by se rovnala $\mathbf{v}_i$ a rozdíl $\mathbf{v}_i - \mathbf{w}_{i-1}$ by byl nulový vektor. V kroku (ib) by se algoritmus ozval, že má dělit číslem 0, a zastavil by se. Gramovu-Schmidtovu ortogonalizaci tak můžeme použít ke zjištění, je-li nějaká posloupnost $(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k)$ prvků lineárního prostoru se skalárním součinem lineárně závislá nebo nezávislá.

Z věty 8.49 dostáváme ihned dva důležité důsledky.

Věta 8.51. *Je-li* $\mathbf{W}$ *podprostor konečně generovaného lineárního prostoru* $\mathbf{V}$ *se skalárním součinem, pak každou ortonormální (ortogonální) bázi v podprostoru* $\mathbf{W}$ *lze doplnit na ortonormální (ortogonální) bázi celého prostoru* $\mathbf{V}$.

Speciálně, v každém konečně generovaném lineárním prostoru se skalárním součinem existuje ortonormální báze.

Důkaz. Nechť $(\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k)$ ortonormální báze $\mathbf{W}$, je lineárně nezávislá. Doplníme ji vektory $\mathbf{v}_{k+1}, \ldots, \mathbf{v}_n$ na bázi $\mathbf{V}$ (viz důsledek 5.58). Gramova-Schmidtova ortogonalizace z posloupnosti $(\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k, \mathbf{v}_{k+1}, \ldots, \mathbf{v}_n)$ vytvoří ortonormální posloupnost, přičemž prvních k prvků nezmění. (Můžeme ji také „spustit“ až od $(k+1)$-ního cyklu).

Totáž uděláme v případě pouhé ortogonální báze $(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k)$ podprostoru $\mathbf{W}$, Gramovu-Schmidtovu ortogonalizaci použijeme bez normalizačních kroků (ib). $\square$

Druhým důsledkem je následující věta, která formalizuje dříve uváděné intuitivní tvrzení, že každý konečně generovaný lineární prostor se skalárním součinem je „v podstatě stejný“ jako aritmetický vektorový prostor se standardním skalárním součinem.

Věta 8.52. *Je-li* $\mathbf{V}$ *lineární prostor dimenze* n *nad* $\mathbb{R}$ *(nebo nad* $\mathbb{C}$*) se skalárním součinem* $\langle\,,\rangle$*, pak existuje izomorfismus* $f : \mathbf{V} \to \mathbb{R}^n$ *(nebo* $f : \mathbf{V} \to \mathbb{C}^n$*), pro který platí*

$$\langle\, \mathbf{u}, \mathbf{v} \rangle = f(\mathbf{u}) \cdot f(\mathbf{v})$$

pro každé dva prvky $\mathbf{u}, \mathbf{v} \in \mathbf{V}$*.*

Důkaz. V prostoru $\mathbf{V}$ zvolíme ortonormální bázi B a definujeme f předpisem $f(\mathbf{u}) = [\mathbf{u}]_B$. Podle tvrzení 6.29 je f izomorfizmus mezi $\mathbf{V}$ a $\mathbb{R}^n$ (nebo $\mathbb{C}^n$). Podle tvrzení 8.40 platí

$$\langle\, \mathbf{u}, \mathbf{v} \rangle = [\mathbf{u}]_B^* \,[\mathbf{v}]_B = f(\mathbf{u}) \cdot f(\mathbf{v}) \ .$$

$\square$

8.4.2. *QR-rozklad.* QR-rozklad je maticová formulace Gramovy-Schmidtovy ortogonalizace v aritmetických vektorových prostorech se skalárním součinem. Ze vzorce pro Gramovu-Schmidtovu ortogonalizaci vidíme, že původní vektory $\mathbf{v}_i$ lze vyjádřit jako lineární kombinaci vektorů $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k$, které jsou navzájem ortogonální a jednotkové. Použijeme-li tento fakt na lineárně nezávislou posloupnost $(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k)$ reálných (nebo komplexních) n-složkových aritmetických vektorů, získáme vyjádření matice $A = (\mathbf{v}_1|\mathbf{v}_2|\dots|\mathbf{v}_k)$ jako součinu matice $(\mathbf{u}_1|\mathbf{u}_2|\dots|\mathbf{u}_k)$ a horní trojúhelníkové matice řádu k. Tomuto vyjádření říkáme QR-rozklad.

Tvrzení 8.53 (o QR-rozkladu)**.** *Je-li* A *reálná nebo komplexní matice typu* $n \times k$ *s lineárně nezávislými sloupci, pak existuje matice* Q *typu* $n \times k$ *nad stejným tělesem s ortonormálními sloupci a horní trojúhelníková matice* R *řádu* k *s kladnými reálnými prvky na hlavní diagonále taková, že platí* $A = QR$*.*

Důkaz. Označíme $(\mathbf{v}_1, \dots, \mathbf{v}_k)$ posloupnost sloupcových vektorů matice A. Na tuto posloupnost provedeme Gramovu-Schmidtovu ortogonalizaci, tj. označíme prro každé $i = 1, 2, \dots k$

$$\begin{aligned} \mathbf{w}_{i-1} &= \langle\, \mathbf{u}_1, \mathbf{v}_i \rangle\, \mathbf{u}_1 + \langle\, \mathbf{u}_2, \mathbf{v}_i \rangle\, \mathbf{u}_2 + \dots + \langle\, \mathbf{u}_{i-1}, \mathbf{v}_i \rangle\, \mathbf{u}_{i-1} \ , \\ \mathbf{u}_i &= \frac{\mathbf{v}_i - \mathbf{w}_{i-1}}{\|\mathbf{v}_i - \mathbf{w}_{i-1}\|} \ . \end{aligned}$$

Z toho získáme vyjádření

$$\begin{aligned} \mathbf{v}_i &= \|\mathbf{v}_i - \mathbf{w}_{i-1}\|\, \mathbf{u}_i + \mathbf{w}_{i-1} \\ &= \|\mathbf{v}_i - \mathbf{w}_{i-1}\|\, \mathbf{u}_i + \langle\, \mathbf{u}_1, \mathbf{v}_i \rangle\, \mathbf{u}_1 + \langle\, \mathbf{u}_2, \mathbf{v}_i \rangle\, \mathbf{u}_2 + \dots + \langle\, \mathbf{u}_{i-1}, \mathbf{v}_i \rangle\, \mathbf{u}_{i-1} \\ &= \langle\, \mathbf{u}_1, \mathbf{v}_i \rangle\, \mathbf{u}_1 + \langle\, \mathbf{u}_2, \mathbf{v}_i \rangle\, \mathbf{u}_2 + \dots + \langle\, \mathbf{u}_{i-1}, \mathbf{v}_i \rangle\, \mathbf{u}_{i-1} + \|\mathbf{v}_i - \mathbf{w}_{i-1}\|\, \mathbf{u}_i \ , \end{aligned}$$

což můžeme maticově zapsat ve tvaru

$$(\mathbf{v}_1|\mathbf{v}_2|\dots|\mathbf{v}_k) = (\mathbf{u}_1|\dots|\mathbf{u}_k)\begin{pmatrix} \|\mathbf{v}_1-\mathbf{w}_0\| & \langle\,\mathbf{u}_1,\mathbf{v}_2\rangle & \dots & \langle\,\mathbf{u}_1,\mathbf{v}_k\rangle \\ 0 & \|\mathbf{v}_2-\mathbf{w}_1\| & \dots & \langle\,\mathbf{u}_2,\mathbf{v}_k\rangle \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \|\mathbf{v}_k-\mathbf{w}_{k-1}\| \end{pmatrix}.$$

Pokud v předchozí rovnosti nahradíme obecný skalární součin stadardním, dostaneme

$$(\mathbf{v}_1|\mathbf{v}_2|\dots|\mathbf{v}_k) = (\mathbf{u}_1|\dots|\mathbf{u}_k)\begin{pmatrix} \|\mathbf{v}_1-\mathbf{w}_0\| & \mathbf{u}_1^*\mathbf{v}_2 & \dots & \mathbf{u}_1^*\mathbf{v}_k \\ 0 & \|\mathbf{v}_2-\mathbf{w}_1\| & \dots & \mathbf{u}_2^*\mathbf{v}_k \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \|\mathbf{v}_k-\mathbf{w}_{k-1}\| \end{pmatrix}.$$

□

Příklad 8.54. Vypočítáme QR-rozklad reálné matice

$$A = (\mathbf{v}_1|\mathbf{v}_2|\mathbf{v}_3) = \begin{pmatrix} 1 & 1 & 0 \\ 2 & -1 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 3 \end{pmatrix}.$$

Gramovu-Schmidtovu ortogonalizaci posloupnosti sloupcových vektorů $(\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3)$ v aritmetickém prostoru se standardním skalárním součinem jsme spočítali už v příkladu 8.50. Našli jsme tam ortonormální posloupnost

$$(\mathbf{u}_1,\mathbf{u}_2,\mathbf{u}_3) = \left(\frac{1}{\sqrt{6}}\begin{pmatrix}1\\2\\0\\1\end{pmatrix}, \frac{1}{\sqrt{102}}\begin{pmatrix}7\\-4\\6\\1\end{pmatrix}, \frac{1}{\sqrt{119}}\begin{pmatrix}-5\\-2\\3\\9\end{pmatrix}\right).$$

Během výpočtu jsme také našli vektory

$$(\mathbf{v}_1-\mathbf{w}_0,\mathbf{v}_2-\mathbf{w}_1,\mathbf{v}_3-\mathbf{w}_2) = \left(\begin{pmatrix}1\\2\\0\\1\end{pmatrix}, \frac{1}{6}\begin{pmatrix}7\\-4\\6\\1\end{pmatrix}, \frac{4}{17}\begin{pmatrix}-5\\-2\\3\\9\end{pmatrix}\right).$$

Nyní stačí pouze spočítat příslušné normy a standardní skalární součiny a dosadit je do závěrečné formulky z důkazu předchozí věty. Dostaneme

$$\begin{pmatrix} 1 & 1 & 0 \\ 2 & -1 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 3 \end{pmatrix} = \begin{pmatrix} 1/\sqrt{6} & 7/\sqrt{102} & -5/\sqrt{119} \\ 2/\sqrt{6} & -4/\sqrt{102} & -2/\sqrt{119} \\ 0 & 6/\sqrt{102} & 3/\sqrt{119} \\ 1/\sqrt{6} & 1/\sqrt{102} & 9/\sqrt{119} \end{pmatrix}\begin{pmatrix} \sqrt{6} & -1/\sqrt{6} & 5/\sqrt{6} \\ 0 & \sqrt{102}/6 & 5/\sqrt{102} \\ 0 & 0 & 4\sqrt{119}/17 \end{pmatrix}.$$

Gramova-Schmidtova ortogonalizace obecně není numericky stabilní. Její stabilitu lze vylepšit tak, že jednotlivé algebraické operace při výpočtu děláme v jiném pořadí, ale tak, aby se výsledek nezměnil. Tomu se říká *modifikovaná Gramova-Schmidtova ortogonalizace* a více o ní bude v přednášce z numerické lineární algebry ve druhém ročníku.

8.5. **Unitární a ortogonální matice.**

Dalším pojmem, kterým se budeme stručně zabývat, jsou ortogonální a unitární matice. Pokud v tvrzení 8.53 o QR-rozkladu uvažujeme aritmetický prostor $\mathbb{R}^n$ (nebo $\mathbb{C}^n$) se standardním skalárním součinem, můžeme podmínku, že posloupnost sloupcových vektorů matice Q typu $n \times k$ je ortonormální, zapsat elegantně pomocí rovnosti $Q^T Q = I_k$ (nebo $Q^* Q = I - n$). Čtvercové matice s touto vlastností si vysloužily samostatné jméno.

Definice 8.55. Čtvercová reálná matice A řádu n se nazývá *ortogonální*, platí-li $A^T A = I_n$, čtvercová komplexní matice U řádu n se nazývá *unitární*, platí-li $U^* U = I_n$.

Každá reálná (resp. komplexní) čtvercová matice Q řádu n určuje zobrazení $f_Q : \mathbb{R}^n \to \mathbb{R}^n$ (resp. $f_Q : \mathbb{C}^n \to \mathbb{C}^n$). Následující tvrzení shrnuje řadu různých ekvivalentních definic ortogonálních (unitárních) matic.

Tvrzení 8.56. *Je-li Q reálná (resp. komplexní) čtvercová matice řádu n, pak jsou následující tvrzení ekvivalentní.*

(1) *Q je ortogonální (resp. unitární),*
(2) *zobrazení $f_Q(\mathbf{q}) = Q\,\mathbf{u}$ zachovává standardní skalární součin, tj. pro libovolné $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ (resp. $\mathbb{C}^n$) platí $Q\mathbf{u} \cdot Q\mathbf{v} = \mathbf{u} \cdot \mathbf{v}$,*
(3) *f_Q zachovává eukleidovskou normu, tj. pro libovolný vektor $\mathbf{v} \in \mathbb{R}^n$ (resp. $\mathbb{C}^n$) platí $\|Q\mathbf{v}\| = \|\mathbf{v}\|$,*
(4) *f_Q zobrazuje každou ortonormální bázi na ortonormální bázi,*
(5) *$Q^{-1} = Q^T$, (resp. $Q^{-1} = Q^*$),*
(6) *posloupnost řádkových vektorů matice Q tvoří ortonormální bázi v $\mathbb{R}^n$ (resp. $\mathbb{C}^n$) se standardním skalárním součinem,*
(7) *posloupnost sloupcových vektorů matice Q tvoří ortonormální bázi v $\mathbb{R}^n$ (resp. $\mathbb{C}^n$) se standardním skalárním součinem.*

Důkaz. Označíme $Q = (\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_n)$ a $Q^T = (\tilde{\mathbf{q}}_1, \tilde{\mathbf{q}}_2, \dots, \tilde{\mathbf{q}}_n)$. Ekvivalence (1) a (5) plyne z toho, že každá matice inverzní zleva ke čtvercové matici Q je už inverzní z obou stran.

Rovnost $Q^T Q = I_n$ (resp. $Q^* Q = I_n$) je ekvivalentní tomu, že $\mathbf{q}_i \cdot \mathbf{q}_j = \delta_{ij}$ pro každé $i, j = 1, 2, \dots, n$, což je ekvivalentní tomu, že posloupnost sloupcových vektorů $(\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_n)$ je ortonormální vzhledem ke standardnímu skalárnímu součinu v $\mathbb{R}^n$ (resp. $\mathbb{C}^n$). To dokazuje ekvivlenci (1) a (6). Analogicky je rovnost $QQ^T = I_n$ ekvivalentní tomu, že $\tilde{\mathbf{q}}_i \cdot \tilde{\mathbf{q}}_j^T = \delta_{ij}$, což znamená, že posloupnost řádkových vektorů $(\tilde{\mathbf{q}}_1^T, \tilde{\mathbf{q}}_2^T, \dots, \tilde{\mathbf{q}}_n^T)$ je ortonormální v $\mathbb{R}^n$ (resp. $\mathbb{C}^n$) se standardním skalárním součinem. Tvrzení (1), (5), (6) a (7) jsou tedy navzájem ekvivalentní.

Důkaz, že z (1) plyne (2) je přímočarý. Pro každé dva vektory $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ (resp. $\mathbb{C}^n$) platí

$$f_Q(\mathbf{u}) \cdot f_Q(\mathbf{v}) = (Q\,\mathbf{u}) \cdot (Q\,\mathbf{v}) = (Q\,\mathbf{u})^T Q\,\mathbf{v} = \mathbf{u}^T Q^T Q\,\mathbf{v} = \mathbf{u}^T \mathbf{v} = \mathbf{u} \cdot \mathbf{v}$$

(resp. $(Q\,\mathbf{u}) \cdot (Q\,\mathbf{v}) = (Q\,\mathbf{u})^* Q\,\mathbf{v} = \mathbf{u}^* Q^* Q\,\mathbf{v} = \mathbf{u}^* \mathbf{v} = \mathbf{u} \cdot \mathbf{v}$).

Tvrzení (4) plyne z (2) přímo z definic. Z (4) plyne ihned (6), protože $\mathbf{q}_i = Q\mathbf{e}_i = f_Q(\mathbf{e}_i)$ pro každé $i = 1, 2, \dots, n$ a posloupnost $(\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n)$ prvků kanonické báze v $\mathbb{R}^n$ (resp. $\mathbb{C}^n$) je ortonormální. Tím jsme dokázali, že jsou ekvivalentní všechna tvrzení (1), (2), (4), (5), (6) a (7).

Stejně přímo ze (2) plyne (3). Platí $\|f_Q(\mathbf{u})\|^2 = (Q\mathbf{u}) \cdot (Q\mathbf{u}) = \mathbf{u} \cdot \mathbf{u} = \|\mathbf{u}\|^2$, a protože norma každého vektoru je nezáporná, plyne odtud $\|f_Q(\mathbf{u})\| = \|\mathbf{u}\|$ pro každé

$\mathbf{u} \in \mathbb{R}^n$ (resp. $\mathbf{u} \in \mathbb{C}^n$). Naopak z (3) plyne (2) pomocí polarizačních identit, které říkají, že skalární součin je jednoznačně určen normou, kterou definuje. Podrobněji, protože f_Q zachovává normu, dostaneme z bodu (4) tvrzení 8.25

$$\begin{aligned}\mathrm{Re}\,(Q\mathbf{u} \cdot Q\mathbf{v}) &= \frac{1}{2}(\|Q\mathbf{u} + Q\mathbf{v}\|^2 - \|Q\mathbf{u}\|^2 - \|Q\mathbf{v}\|^2) \\ &= \frac{1}{2}(\|Q(\mathbf{u} + \mathbf{v})\|^2 - \|Q\mathbf{u}\|^2 - \|Q\mathbf{v}\|^2) = \frac{1}{2}(\|\mathbf{u} + \mathbf{v}\|^2 - \|\mathbf{u}\|^2 - \|\mathbf{v}\|^2) \\ &= \mathrm{Re}\,(\mathbf{u} \cdot \mathbf{v}) \ .\end{aligned}$$

Tím je dokázáno, že z (3) plyne (2) v případě $\mathbb{R}^n$. Rovnost imaginárních částí skalárního součinu v případě $\mathbb{C}^n$ dostaneme analogicky z polarizační identity ve cvičeních. □

Standardní pojmenování ortogonální matice je poněkud matoucí, smysluplnější by bylo nazývat ji ortonormální. Hezkou vlastností těchto matic je snadné určení inverzní matice – stačí vzít podle bodu (5) matici hermitovsky sdruženou. Příklady ortogonálních matic jsou matice rotací a zrcadlení podle podprostorů vzhledem ke kanonickým bázím.

Důsledek 8.57. *Součin dvou ortogonálních (resp. unitárních) matic téhož řádu je opět ortogonální (resp. unitární) matice.*

Důkaz. Dokážeme pouze ortogonální případ. Jsou-li A, B ortogonální matice, platí $A^{-1} = A^T$ a $B^{-1} = B^T$, a proto $(AB)^{-1} = B^{-1}A^{-1} = B^T A^T = (AB)^T$, což dokazuje, že AB je také ortogonální. □

Rovnost $Q^T Q = I_n$ pro ortogonální matice nyní použijeme k důkazu, že QR-rozklad regulární matice A je určený jednoznačně.

Tvrzení 8.58. *Je-li A regulární (reálná nebo kompexní) matice řádu a a $A = Q_1R_1 = Q_2R_2$ jsou dva QR-rozklady matice A, pak platí $Q_1 = Q_2$ a $R_1 = R_2$.*

Důkaz. Připomňme, že v QR-rozkladu je posloupnost sloupcových vektorů matice Q ortonormální, což v případě čtvercové matice A znamená, že Q je také čtvercová a tedy ortogonální (resp. unitární). Matice R je horní trojúhelníková s kladnými prvky na hlavní diagonále.

Jsou-li $A = Q_1R_1 = Q_2R_2$ dva QR-rozklady regulární matice A, plyne odtud

$$Q_2^{-1}Q_1 = R_2R_1^{-1} \ .$$

Označíme si tento součin $C = (\mathbf{c}_1|\mathbf{c}_2|\cdots|\mathbf{c}_n) = (c_{ij})$. Součin $Q_2^{-1}Q_1$ je ortogonální a součin $R_2R_1^{-1}$ je horní trojúhelníková matice s kladnými prvky na hlavní diagonále. Platí proto $c_{i1} = 0$ pro $i > 1$ a tedy $1 = \|\mathbf{c}_1\| = \mathbf{c}_1^T\mathbf{c}_1 = c_{11}^2$ a tedy $c_{11} = 1$, neboť $c_{11} > 0$. To znamená, že $\mathbf{c}_1 = \mathbf{e}_1$.

Dále postupujeme indukcí podle indexu sloupců $\mathbf{c}_j$ a dokážeme $\mathbf{c}_j = \mathbf{e}_j$ pro každé $j = 1, 2, \ldots, n$. Pro $j = 1$ jsme to právě ověřili. Předpokládejme nyní, že pro nějaké $j > 1$ a $j \leq n$ platí $\mathbf{c}_i = \mathbf{e}_i$ pro každé $i = 1, 2, \ldots, j-1$. Z rovností $\mathbf{c}_i \cdot \mathbf{c}_j = 0$ pro $i < j$ (Q je ortogonální) a indukčního předpokladu $\mathbf{c}_i = \mathbf{e}_i$ plyne $0 = \mathbf{e}_i \cdot \mathbf{c}_j = c_{ij}$ pro každé $i < j$. Matice $(c_{ij}) = R_2R_1^{-1}$ je horní trojúhelníková, proto také $c_{ij} = 0$ pro každé $i > j$. Z rovnosti $\|\mathbf{c}_j\| = 1$ (opět ortogonalita matice C) pak plyne $c_{jj}^2 = 1$ a tedy $c_{jj} = 1$, neboli $\mathbf{c}_j = \mathbf{e}_j$.

Indukcí jsme tak dokázali, že $Q_2^{-1}Q_1 = R_2R_1^{-1} = I_n$ a tedy $Q_1 = Q_2$ a $R_1 = R_2$. □

QR-rozklad lze použít při řešení reálných (komplexních) soustav lineárních rovnic $A\mathbf{x} = \mathbf{b}$ s danou regulární maticí A a různými vektory pravých stran $\mathbf{b}$ stejným způsobem, jakým lze použít LU-rozklad. Spočteme QR-rozklad matice $A = QR$. Rovnici $A\mathbf{x} = \mathbf{b}$ přepíšeme do tvaru $QR\mathbf{x} = \mathbf{b}$ a využijeme toho, že matice Q je ortogonální (unitární). Inverzní matici $Q^{-1} = Q^T$ nemusíme počítat a soustavu

$$R\mathbf{x} = Q^T\mathbf{b}$$

vyřešíme zpětnou substitucí. Algoritmus pro QR-rozklad je numericky stabilnější než Gaussova eliminace, která vede na LU-rozklad. Vyžaduje ale zhruba n^3 aritmetických operací, což je třikrát více než výpočet LU-rozkladu.

Následující příklad ukazuje, že Gramova-Schmidtova ortogonalizace skutečně není numericky stabilní.

Příklad 8.59. V aritmetice se zaokrouhlováním na tři platná místa použijeme Gramovu-Schmidtovu ortogonalizaci na posloupnost sloupcových vektorů matice

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 10^{-3} & 10^{-3} & 0 \\ 10^{-3} & 0 & 10^{-3} \end{pmatrix} .$$

Všimněme si, že všechny sloupcové vektory jsou „skoro“ rovnoběžné. Výsledek zapíšeme jako sloupcové vektory matice

$$Q = \begin{pmatrix} 1 & 0 & 0 \\ 10^{-3} & 0 & -0{,}709 \\ 10^{-3} & -1 & -0{,}709 \end{pmatrix} .$$

Vidíme, že druhý a třetí sloupec příliš kolmé nevyšly.

8.5.1. *Gramova matice.* Pokud známe v lineárním prostoru $\mathbf{V}$ se skalárním součinem obecnou bázi $B = (\mathbf{u}_1, \mathbf{u}_1, \ldots, \mathbf{u}_k)$ nějakého konečně-generovaného podprostoru $\mathbf{W}$, můžeme spočítat projekci $\mathbf{w}$ libovolného prvku $\mathbf{v} \in \mathbf{V}$ na podprostor $\mathbf{W}$ přímo, bez toho abychom napřed hledali ortonormální bázi. Následující tvrzení ukazuje, že vektor souřadnic projekce $\mathbf{w}$ vzhledem k bázi B spočteme jako řešení jisté soustavy lineárních rovnic.

Tvrzení 8.60. *Je-li* $\mathbf{W}$ *konečně generovaný podprostor lineárního prostoru* $\mathbf{V}$ *se skalárním součinem* $\langle\,,\rangle$, $\mathbf{v} \in V$ *a* $B = (\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k)$ *báze ve* $\mathbf{W}$, *pak vektor souřadnic* $[\mathbf{w}]_B = (a_1, a_2, \ldots, a_k)^T$ *ortogonální projekce prvku* $\mathbf{v}$ *na podprostor* $\mathbf{W}$ *spočteme jako řešení soustavy lineárních rovnic*

$$\left(\begin{array}{cccc|c} \langle \mathbf{u}_1, \mathbf{u}_1\rangle & \langle \mathbf{u}_1, \mathbf{u}_2\rangle & \cdots & \langle \mathbf{u}_1, \mathbf{u}_k\rangle & \langle \mathbf{u}_1, \mathbf{v}\rangle \\ \langle \mathbf{u}_2, \mathbf{u}_1\rangle & \langle \mathbf{u}_2, \mathbf{u}_2\rangle & \cdots & \langle \mathbf{u}_2, \mathbf{u}_k\rangle & \langle \mathbf{u}_2, \mathbf{v}\rangle \\ \vdots & & \ddots & \vdots & \vdots \\ \langle \mathbf{u}_k, \mathbf{u}_1\rangle & \langle \mathbf{u}_k, \mathbf{u}_2\rangle & \cdots & \langle \mathbf{u}_k, \mathbf{u}_k\rangle & \langle \mathbf{u}_k, \mathbf{v}\rangle \end{array}\right) .$$

Důkaz. Definice ortogonální projekce $\mathbf{w}$ prvku $\mathbf{v}$ na podprostor $\mathbf{W}$ říká, že musí platit $(\mathbf{v} - \mathbf{w}) \perp \mathbf{W}$, což platí právě když $(\mathbf{v} - \mathbf{w}) \perp \mathbf{u}_i$ pro každé $i = 1, 2, \ldots, k$, neboli právě když $\langle \mathbf{u}_i, \mathbf{w}\rangle = \langle \mathbf{u}_i, \mathbf{v}\rangle$.

Souřadnice $[\mathbf{w}]_B = (a_1, a_2, \ldots, a_k)^T$ projekce $\mathbf{w}$ splňují rovnici

$$\mathbf{w} = a_1\mathbf{u}_1 + a_2\mathbf{u}_2 + \cdots + a_k\mathbf{u}_k \ ,$$

ze které s využitím lienarity skalárního součinu vzhledem ke druhé složce dostáváme

$$\langle \mathbf{u}_i, \mathbf{w}\rangle = a_1 \langle \mathbf{u}_i, \mathbf{u}_1\rangle + a_2 \langle \mathbf{u}_i, \mathbf{u}_2\rangle + \cdots + a_k \langle \mathbf{u}_i, \mathbf{u}_k\rangle \ .$$

Rovnost $\langle \mathbf{u}_i, \mathbf{w}\rangle = \langle \mathbf{u}_i, \mathbf{v}\rangle$ platí tedy právě když souřadnice $a_1, a_2, \ldots, a_k$ splňují rovnici

$$a_1 \langle \mathbf{u}_i, \mathbf{u}_1\rangle + a_2 \langle \mathbf{u}_i, \mathbf{u}_2\rangle + \cdots + a_k \langle \mathbf{u}_i, \mathbf{u}_k\rangle = \langle \mathbf{u}_i, \mathbf{v}\rangle$$

pro každé $i = 1, 2, \ldots, k$. □

Příklad 8.61. V prostoru reálných polynomů se skalárním součinem $\langle f, g\rangle = \int_0^1 fg$ najdeme ortogonální projekci $\mathbf{w}$ polynomu $\mathbf{v} = x^2$ na podprostor $\mathbf{W} = \langle \mathbf{u}_1, \mathbf{u}_2\rangle = \langle 1, x\rangle$ polynomů stupně nejvýše 1. Projekce $\mathbf{w} = a + bx$ s neznámými koeficienty $a, b \in \mathbb{R}$, které podle předchozího tvrzení najdeme jako řešení soustavy

$$\left(\begin{array}{cc|c} \langle \mathbf{u}_1, \mathbf{u}_1\rangle & \langle \mathbf{u}_1, \mathbf{u}_2\rangle & \langle \mathbf{u}_1, \mathbf{v}\rangle \\ \langle \mathbf{u}_2, \mathbf{u}_1\rangle & \langle \mathbf{u}_2, \mathbf{u}_2\rangle & \langle \mathbf{u}_2, \mathbf{v}\rangle \end{array}\right) = \left(\begin{array}{cc|c} \int_0^1 1 & \int_0^1 x & \int_0^1 x^2 \\ \int_0^1 x & \int_0^1 x^2 & \int_0^1 x^3 \end{array}\right) = \left(\begin{array}{cc|c} 1 & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \end{array}\right) .$$

Řešením soustavy dostaneme vektor $(a, b)^T = (-\frac{1}{6}, 1)^T$. Ortogonální projekce polynomu $\mathbf{v} = x^2$ je tedy

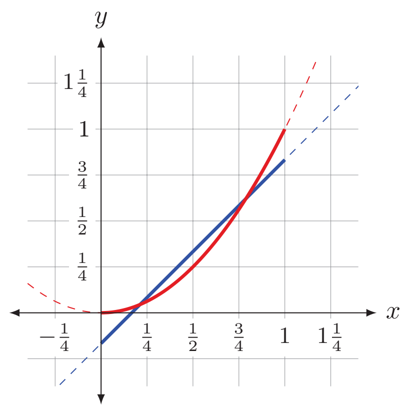
$$\mathbf{w} = a\mathbf{u}_1 + b\mathbf{u}_2 = -\frac{1}{6} + x \ .$$

Vzdálenost polynomu $\mathbf{v} = x^2$ od podprostoru $\mathbf{W}$ se rovná normě vektoru

$$\mathbf{v} - \mathbf{w} = x^2 - x + \frac{1}{6} \ ,$$

kterou spočteme jako

$$\|\mathbf{v} - \mathbf{w}\| = \sqrt{\int_0^1 \left(x^2 - x + \frac{1}{6}\right)^2} = \sqrt{\int_0^1 x^4 - 2x^3 + \frac{4}{3}x^2 - \frac{1}{3}x + \frac{1}{36}}$$
$$= \sqrt{\frac{1}{5} - \frac{1}{2} + \frac{4}{9} - \frac{1}{6} + \frac{1}{36}} = \sqrt{\frac{1}{30}}$$



Matice soustavy z tvrzení 8.60 si také vysloužila vlastní jméno.

Definice 8.62. Jsou-li $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k$ prvky lineárního prostoru se skalárním součinem, pak čtvercovou matici

$$\begin{pmatrix} \langle \mathbf{u}_1, \mathbf{u}_1\rangle & \langle \mathbf{u}_1, \mathbf{u}_2\rangle & \cdots & \langle \mathbf{u}_1, \mathbf{u}_k\rangle \\ \langle \mathbf{u}_2, \mathbf{u}_1\rangle & \langle \mathbf{u}_2, \mathbf{u}_2\rangle & \cdots & \langle \mathbf{u}_2, \mathbf{u}_k\rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle \mathbf{u}_k, \mathbf{u}_1\rangle & \langle \mathbf{u}_k, \mathbf{u}_2\rangle & \cdots & \langle \mathbf{u}_k, \mathbf{u}_k\rangle \end{pmatrix}$$

řádu k nazýváme *Gramova matice* posloupnosti prvků $(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k)$.

Všimněme si, že v případě standardního skalárního součinu na aritmetickém prostoru $\mathbb{R}^n$ můžeme Gramovu matici spočítat rychle tak, že si vektory $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k$ zapíšeme jako sloupce matice

$$A = (\mathbf{u}_1 | \mathbf{u}_2 | \cdots | \mathbf{u}_k)$$

a spočteme

$$A^T A = \begin{pmatrix} \mathbf{u}_1^T \mathbf{u}_1 & \mathbf{u}_1^T \mathbf{u}_2 & \cdots & \mathbf{u}_1^T \mathbf{u}_k \\ \mathbf{u}_2^T \mathbf{u}_1 & \mathbf{u}_2^T \mathbf{u}_2 & \cdots & \mathbf{u}_2^T \mathbf{u}_k \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{u}_k^T \mathbf{u}_1 & \mathbf{u}_k^T \mathbf{u}_2 & \cdots & \mathbf{u}_k^T \mathbf{u}_k \end{pmatrix} = \begin{pmatrix} \mathbf{u}_1 \cdot \mathbf{u}_1 & \mathbf{u}_1 \cdot \mathbf{u}_2 & \cdots & \mathbf{u}_1 \cdot \mathbf{u}_k \\ \mathbf{u}_2 \cdot \mathbf{u}_1 & \mathbf{u}_2 \cdot \mathbf{u}_2 & \cdots & \mathbf{u}_2 \cdot \mathbf{u}_k \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{u}_k \cdot \mathbf{u}_1 & \mathbf{u}_k \cdot \mathbf{u}_2 & \cdots & \mathbf{u}_k \cdot \mathbf{u}_k \end{pmatrix} .$$

Gramova matice posloupnosti vektorů $(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k)$ tedy v tomto případě není nic jiného než součin matic $A^T A$. V případě standardního skalárního součinu na $\mathbb{C}^n$ je to matice $A^* A$.

Základní vlastnosti Gramovy matice shrnuje následující tvrzení.

Tvrzení 8.63. *Pro Gramovu matici $B = (\langle\, \mathbf{u}_i, \mathbf{u}_j \rangle)$ posloupnosti prvků $(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k)$ nějakého lineárního prostoru $\mathbf{V}$ se skalárním součinem platí*

(1) *matice B je regulární právě když je posloupnost $(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k)$ lineárně nezávislá,*
(2) *matice B je symetrická (hermitovská v komplexním případě),*
(3) *je-li posloupnost $(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k)$ lineárně nezávislá, pak je Gramova matice B pozitivně definitní.*

Důkaz. Gramova matice B má řád k. K důkazu (1) zvolíme libovolné skaláry $a_1, a_2, \dots, a_k \in \mathbb{C}$ a pro $i = 1, 2, \dots, k$ spočteme skalární součiny

$$\langle\, \mathbf{u}_i, a_1\mathbf{u}_1 + a_2\mathbf{u}_2 + \cdots + a_k\mathbf{u}_k \rangle = a_1 \langle\, \mathbf{u}_i, \mathbf{u}_1 \rangle + a_2 \langle\, \mathbf{u}_i, \mathbf{u}_2 \rangle + \cdots + a_k \langle\, \mathbf{u}_i, \mathbf{u}_k \rangle \ .$$

Označíme-li $\mathbf{w} = a_1\mathbf{u}_1 + a_2\mathbf{u}_2 + \cdots + a_k\mathbf{u}_k$, dostáváme tak, že součin $B(a_1, a_2, \dots, a_k)^T$ se rovná

$$B(a_1, a_2, \dots, a_k)^T = \begin{pmatrix} \langle\, \mathbf{u}_1, \mathbf{w} \rangle \\ \langle\, \mathbf{u}_2, \mathbf{w} \rangle \\ \vdots \\ \langle\, \mathbf{u}_k, \mathbf{w} \rangle \end{pmatrix} .$$

Je-li nyní Gramova matice B singulární, existuje nenulový vektor $\mathbf{a} = (a_1, a_2, \dots, a_k)^T$ takový, že $B\mathbf{a} = \mathbf{o}$. Odtud plyne, že pro vektor $\mathbf{w} = a_1\mathbf{u}_1 + a_2\mathbf{u}_2 + \cdots + a_k\mathbf{u}_k$ platí $\langle\, \mathbf{u}_i, \mathbf{w} \rangle = 0$ pro každé $i = 1, 2, \dots, k$. Potom ale

$$\|\mathbf{w}\|^2 = \langle\, \mathbf{w}, \mathbf{w} \rangle = \langle\, w, a_1\mathbf{u}_1 + a_2\mathbf{u}_2 + \cdots + a_k\mathbf{u}_k \rangle = a_1 \langle\, \mathbf{w}, \mathbf{u}_1 \rangle + \cdots + a_k \langle\, \mathbf{w}, \mathbf{u}_k \rangle = 0 \ ,$$

což podle axiomu (SP) pro skalární součin znamená, že $\mathbf{w} = \mathbf{o}$ a posloupnost $(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k)$ je lineárně závislá.

Je-li naopak $(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k)$ lineárně závislá posloupnost, existují skaláry $a_1, a_2, \dots, a_k$, které nejsou všechny nulové, a pro které platí $a_1\mathbf{u}_1 + a_2\mathbf{u}_2 + \cdots + a_k\mathbf{u}_k = \mathbf{o}$. Pak platí také

$$0 = \langle\, \mathbf{u}_i, \mathbf{o} \rangle = \langle\, \mathbf{u}_i, a_1\mathbf{u}_1 + a_2\mathbf{u}_2 + \cdots + a_k\mathbf{u}_k \rangle = a_1 \langle\, \mathbf{u}_i, \mathbf{u}_1 \rangle + \cdots + a_k \langle\, \mathbf{u}_i, \mathbf{u}_k \rangle$$

pro každé $i = 1, 2, \ldots, k$, a proto

$$B(a_1, a_2, \ldots, a_k)^T = \begin{pmatrix} \langle \mathbf{u}_1, \mathbf{o}\rangle \\ \langle \mathbf{u}_2, \mathbf{o}\rangle \\ \vdots \\ \langle \mathbf{u}_k, \mathbf{o}\rangle \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} .$$

Protože vektor $(a_1, a_2, \ldots, a_k)^T$ je nenulový, je Gramova matice B singulární.

Vlastnost (2) plyne ze skorosymetrie (SCS) skalárního součinu.

Pokud jde o vlastnost (3), potřebujeme dokázat, že pro každý aritmetický vektor $\mathbf{a} = (a_1, a_2, \ldots, a_k)^T \in \mathbb{C}^k$ platí $\mathbf{a}^* B\mathbf{a} \geq 0$, přičemž rovnost nastává právě když je $\mathbf{a} = \mathbf{o}$. Označíme $\mathbf{w} = a_1\mathbf{u}_1 + a_2\mathbf{u}_2 + \cdots + a_k\mathbf{u}_k$. Z důkazu (1) víme, že

$$B\mathbf{a} = B(a_1, a_2, \ldots, a_k)^T = \begin{pmatrix} \langle \mathbf{u}_1, \mathbf{w}\rangle \\ \langle \mathbf{u}_2, \mathbf{w}\rangle \\ \vdots \\ \langle \mathbf{u}_k, \mathbf{w}\rangle \end{pmatrix}$$

a tedy

$$\begin{aligned} \mathbf{a}^* B\mathbf{a} &= \overline{a_1}\,\langle \mathbf{u}_1, \mathbf{w}\rangle + \overline{a_2}\,\langle \mathbf{u}_2, \mathbf{w}\rangle + \cdots + \overline{a_k}\,\langle \mathbf{u}_k, \mathbf{w}\rangle \\ &= \langle a_1\mathbf{u}_1 + a_2\mathbf{u}_2 + \cdots + a_k\mathbf{u}_k, \mathbf{w}\rangle = \|\mathbf{w}\|^2 \geq 0 \end{aligned}$$

pro každý vektor $\mathbf{a} \in \mathbb{C}^k$, přičemž rovnost nastává právě když $\mathbf{w} = \mathbf{0}$ podle axiomu (SP) z definice skalárního součinu. Protože předpokládáme, že $(\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k)$ je lineárně nezávislá posloupnost, rovnost $\mathbf{o} = \mathbf{w} = a_1\mathbf{u}_1 + a_2\mathbf{u}_2 + \cdots + a_k\mathbf{u}_k$ nastává právě když $a_1 = a_2 = \cdots = a_k = 0$. Odtod plyne, že $\mathbf{a}^* B\mathbf{a} = 0$ právě když $\mathbf{a} = \mathbf{o}$, což znamená, že B je pozitivně definitní matice. □

Všimněme si ještě, že předchozí důkaz lze výrazně zjednodušit, pokud uvažujeme aritmetický vektorový prostor $\mathbb{C}^n$ se standardním skalárním součinem. V tom případě se Gramova matice B posloupnosti prvků $(\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k)$ rovná rovná součinu A^*A, kde $A = (\mathbf{u}_1|\mathbf{u}_2|\cdots|\mathbf{u}_k)$. K důkazu (1) pak stačí spočítat, že v případě že $B\mathbf{a} = \mathbf{o}$ pro nějaký vektor $\mathbf{a} \in \mathbb{C}^k$, plyne odtud $A^*A\mathbf{a} = \mathbf{o}$ a tedy také $\mathbf{a}^*A^*A\mathbf{a} = \mathbf{o}$, neboli $\|A\mathbf{a}\| = \mathbf{o}$, což znamená, že $A\mathbf{a} = \mathbf{o}$ a z předpokladu lineární nezávislosti posloupnosti sloupcových vektorů matice A dostáváme $\mathbf{a} = \mathbf{o}$. Gramova matice B posloupnosti $(\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k)$ je tedy regulární, pokud je posloupnost vektorů $(\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k)$ lineárně nezávislá. Opačná implikace a pozitivní definitnost matice B, neboli vlastnost (3), se dokážou analogicky.

Předchozí tvrzení ukazuje, že Gramova matice jakékoliv báze lineárního prostoru se skalárním součinem je vždy pozitivně definitní. Později ukážeme, že každá pozitivně definitní matice je naopak Gramovo maticí nějaké báze aritmetického prostoru se skalárním součinem.

8.6. **Ortogonální doplněk.**

Největší množina prvků kolmá na danou množinu $M \subseteq V$ v lineárním prostoru se skalárním součinem se nazývá ortogonální doplněk.

Definice 8.64. Je-li $\mathbf{V}$ prostor se skalárním součinem $\langle\,,\rangle$ a $M \subseteq V$, pak *ortogonální doplněk* $M^\perp$ množiny M je množina všech prvků $\mathbf{V}$ kolmých na každý prvek M, tj.

$$M^\perp = \{\mathbf{v} \in \mathbf{V} : \mathbf{v} \perp M\} .$$

Podle definice M je kolmá na $M^\perp$ a $M^\perp$ je největší taková množina. Další jednoduché vlastnosti ortogonálního doplňku jsou:

Pozorování 8.65. *Je-li* $\mathbf{V}$ *prostor se skalárním součinem* $\langle\, ,\rangle$ *a* $M, N \subseteq V$*, pak platí*

(1) $M^\perp = \langle M\rangle^\perp$,
(2) $M^\perp$ *je podprostor* $\mathbf{V}$,
(3) *je-li* $M \subseteq N$*, pak* $N^\perp \subseteq M^\perp$.

Důkaz. Platí $\mathbf{v} \in M^\perp$ právě když $\mathbf{v} \perp M$ což je právě když $\mathbf{v} \perp \langle M\rangle$ podle pozorování 8.43, a to je právě když $\mathbf{v} \in \langle M\rangle^\perp$.

K důkazu (2) stačí ověřit, že ortogonální doplněk $M^\perp$ je uzavřený na sčítání a násobení skalárem. Je-li $\mathbf{u}, \mathbf{v} \in M^\perp$, platí pro každé $\mathbf{w} \in M$, že $\mathbf{u} \perp \mathbf{w}$ a $\mathbf{v} \perp \mathbf{w}$, a tedy $\langle\, \mathbf{w}, \mathbf{u}+\mathbf{v}\rangle = \langle\, \mathbf{w}, \mathbf{u}\rangle + \langle\, \mathbf{w}, \mathbf{v}\rangle = 0$, což dokazuje $\mathbf{u}+\mathbf{v} \in M^\perp$. Podobně lze dokázat uzavřenost $M^\perp$ na skalární násobky.

Stejně snadné je ověřit (3). Je-li $\mathbf{v} \in N^\perp$, platí $\mathbf{v} \perp N \supseteq M$, tj. také $\mathbf{v} \perp M$ a tedy $\mathbf{v} \in M^\perp$. □

V $\mathbb{R}^3$ se standardním skalárním součinem je ortogonální doplněk množiny $M = \{\mathbf{u}, \mathbf{v}\}$ dvou lineárně nezávislých vektorů přímka kolmá na rovinu $\langle \mathbf{u}, \mathbf{v}\rangle$. Ortogonálním doplňkem nenulového vektoru (nebo jeho lineárního obalu) je rovina.

Příklad 8.66. Určíme ortogonální doplněk roviny $U = \left\langle (1,2,5)^T, (0,1,1)^T)\right\rangle$ v prostoru $\mathbb{R}^3$ se standardním skalárním součinem. Podle (3) je $U^\perp$ rovná množině všech vektorů $\mathbf{x}$ kolmých na oba generátory, tj. množině vektorů, pro které $(1,2,5)\mathbf{x} = 0$ a $(0,1,1)\mathbf{x} = 0$. Maticově

$$\begin{pmatrix} 1 & 2 & 5 \\ 0 & 1 & 1 \end{pmatrix} \mathbf{x} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \ .$$

Hledáme tedy řešení homogenní soustavy s maticí, jejíž řádkové vektory jsou generátory U,

$$U^\perp = \operatorname{Ker} \begin{pmatrix} 1 & 2 & 5 \\ 0 & 1 & 1 \end{pmatrix} = \left\langle \begin{pmatrix} -3 \\ -1 \\ 1 \end{pmatrix} \right\rangle \ .$$

V příkladu jsme viděli, že k určení ortogonálního doplňku množiny vektorů $M = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ (nebo podprostoru $\langle M\rangle$) v aritmetickém vektorovém prostoru $\mathbb{R}^n$ se standardním skalárním součinem stačí napsat vektory $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ do řádků matice a vyřešit příslušnou homogenní soustavu. Při standardním skalárním součinu tedy platí

$$(\operatorname{Im} A^T)^\perp = \operatorname{Ker} A \ .$$

To nám dává nad $\mathbb{R}$ další interpretaci řešení homogenní soustavy rovnic $A\mathbf{x} = \mathbf{o}$ – určujeme ortogonální doplněk řádků matice A. V $\mathbb{C}^n$ se standardním skalárním součinem je ještě třeba přidat komplexní sdružování,

$$(\operatorname{Im} A^*)^\perp = \operatorname{Ker} A \ .$$

Obecněji, počítáme-li vzhledem k ortonormální bázi, pak skalární součin se chová jako standardní (viz tvrzení 8.40), takže ortogonální doplněk množiny vektorů můžeme spočítat podobně.

Pozorování 8.67. *Nechť* $\mathbf{V}$ *je konečně generovaný prostor se skalárním součinem* $\langle\,,\rangle$*,* B *jeho ortonormální báze,* $M = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$*. Označíme* A *matici s řádky* $[\mathbf{v}_1]_B^*$*,* $[\mathbf{v}_2]_B^*$*,* $\ldots$*,* $[\mathbf{v}_k]_B^*$*. Pak*

$$[M^\perp]_B = \operatorname{Ker} A \ .$$

Důkaz.

$$\begin{aligned}[M^\perp]_B &= \{[\mathbf{u}]_B : \mathbf{u} \perp M\} = \{[\mathbf{u}]_B : \langle\, \mathbf{v}_1, \mathbf{u}\rangle = \langle\, \mathbf{v}_2, \mathbf{u}\rangle = \cdots = \langle\, \mathbf{v}_k, \mathbf{u}\rangle = 0\} \\ &= \{[\mathbf{u}]_B : [\mathbf{v}_1]_B^*[\mathbf{u}]_B = [\mathbf{v}_2]_B^*[\mathbf{u}]_B = \cdots = [\mathbf{v}_k]_B^*[\mathbf{u}]_B = 0\} \\ &= \{\mathbf{x} : A\mathbf{x} = \mathbf{o}\} = \operatorname{Ker} A \ .\end{aligned}$$

□

Důležité netriviální vlastnosti ortogonálního doplňku jsou shrnuty v následující větě o ortogonálním doplňku.

Věta 8.68. *Nechť* $\mathbf{V}$ *je konečně generovaný prostor dimenze* n *se skalárním součinem* $\langle\,,\rangle$ *a* $\mathbf{W}$ *je podprostor* $\mathbf{V}$*. Pak platí*

(1) $\dim(\mathbf{W}^\perp) = n - \dim(\mathbf{W})$,
(2) $\mathbf{V} = \mathbf{W} \oplus \mathbf{W}^\perp$,
(3) $(\mathbf{W}^\perp)^\perp = \mathbf{W}$.

Důkaz. V důkazu použijeme skutečnost dokázanou ve větě 8.51, a to, že každý prostor konečné dimenze má nějakou ortonormální bázi B.

Zvolme nějakou bázi $(\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_k)$ prostoru W, tj. $\dim(W) = k$.

(1) Podle věty 8.51 můžeme zvolit nějakou ortonormální bázi $(\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_k)$ prostoru W, platí tedy $\dim \mathbf{W} = k$. Podle téže věty ji doplníme na ortonormální bázi $(\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_k, \mathbf{w}_{k+1}, \ldots, \mathbf{w}_n)$ celého prostoru $\mathbf{V}$. Platí tedy $\mathbf{w}_j \perp \mathbf{W}$ pro každé $j = k+1, \ldots, n$ a tedy $\mathbf{W} + \mathbf{W}^\perp = \mathbf{V}$ a tedy $\dim(\mathbf{W} + \mathbf{W}^\perp) = n$.

Dokážeme dále, že $\mathbf{W} \cap \mathbf{W}^\perp = \{\mathbf{o}\}$. Je-li totiž $\mathbf{v} \in \mathbf{W} \cap \mathbf{W}^\perp$, platí $\mathbf{v} \perp \mathbf{v}$ a tedy $\|\mathbf{v}\|^2 = \langle\, \mathbf{v}, \mathbf{v}\rangle = 0$, což podle axiomu (SP) z definice skalárního součinu znamená $\mathbf{v} = \mathbf{o}$. Takže $\dim(\mathbf{W} \cap \mathbf{W}^\perp) = 0$. Z věty o dimenzi součtu a průniku podprostorů pak dostáváme, že

$$\dim \mathbf{W}^\perp = \dim(\mathbf{W} + \mathbf{W}^\perp) + \dim(\mathbf{W} \cap \mathbf{W}^\perp) - \dim W = n - k \ .$$

(2) Z důkazu (1) víme, že $\mathbf{W} \cap \mathbf{W}^\perp = \{\mathbf{o}\}$ a $\mathbf{W} + \mathbf{W}^\perp = \mathbf{V}$, což zapisujeme jako $\mathbf{V} = \mathbf{W} \oplus \mathbf{W}^\perp$.

(3) Podprostor $\mathbf{W}$ je kolmý na $\mathbf{W}^\perp$, takže $\mathbf{W}$ je podprostorem $(\mathbf{W}^\perp)^\perp$. Podle (1) máme $\dim(\mathbf{W}^\perp) = n - k$ a $\dim((\mathbf{W}^\perp)^\perp) = n - (n-k) = k$. Takže $\mathbf{W} = (\mathbf{W}^\perp)^\perp$ opět podle tvrzení 5.63.

□

Předpoklad konečné generovanosti V v bodech (2), (3) věty 8.68 můžeme nahradit slabším předpokladem, že W je konečně generovaný. To získáme jako důsledek Gramovy-Schmidtovy ortogonalizace, viz cvičení.

8.6.1. *Prostory určené maticí a kolmost.* Podíváme se ještě krátce na vztahy prostorů určených maticí z hlediska kolmosti.

Uvažujeme standardní skalární součin nad reálnými čísly a reálnou matici A typu $m \times n$.

Všimli jsme si, že pro standardní skalární součin nad $\mathbb{R}$ máme $(\operatorname{Im} A^T)^\perp = \operatorname{Ker} A$. Podle bodů (3) a (2) z věty 8.68 také platí

$$(\operatorname{Ker} A)^\perp = \operatorname{Im} A^T \quad , \quad \operatorname{Ker} A \oplus \operatorname{Im} A^T = \mathbb{R}^n \quad ,$$

kde n je počet sloupců matice A.

Obdobně pro prostory $\operatorname{Im} A$ a $\operatorname{Ker} A^T$ máme vztahy

$$(\operatorname{Im} A)^\perp = \operatorname{Ker} A^T, \quad (\operatorname{Ker} A^T)^\perp = \operatorname{Im} A, \quad \operatorname{Ker} A^T \oplus \operatorname{Im} A = \mathbb{R}^m \quad ,$$

kde m je počet řádků matice A. Dostaneme je z předchozích rovností $(\operatorname{Im} A^T)^\perp = \operatorname{Ker} A$ a $(\operatorname{Ker} A)^\perp = \operatorname{Im} A^T$ nahrazením matice A transponovanou maticí A^T.

Nad komplexními čísly vychází stejné vztahy, jen je potřeba transponování nahradit komplexním sdružováním a reálné prostory $\mathbb{R}^n, \mathbb{R}^m$ komplexními prostory $\mathbb{C}^n, \mathbb{C}^m$.

Příklad 8.69. Pro matici

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

máme

$$\operatorname{Ker} A = \left\langle (-1,5,3)^T \right\rangle, \quad \operatorname{Im} A^T = \left\langle (1,2,-3)^T, (1,-1,2)^T \right\rangle \quad .$$

Skutečně $\operatorname{Ker} A \perp \operatorname{Im} A^T$ a $\operatorname{Ker} A \oplus \operatorname{Im} A^T = \mathbb{R}^3$.

8.7. **Aplikace a zajímavosti.**

8.7.1. *Metoda nejmenších čtverců.* Při řešení praktických problémů se často stává, že vedou na soustavu rovnic $A\mathbf{x} = \mathbf{b}$ s reálnými nebo komplexními koeficienty, která nemá řešení. Řekněme, že $A = (\mathbf{a}_1|\mathbf{a}_2|\cdots|\mathbf{a}_n)$ je matice typu $m \times n$ nad $\mathbb{R}$ nebo $\mathbb{C}$, typicky $m >> n$. Taková soustava může například vzniknout sestavením rovnic z velkého množství měření, která jsou zatížená chybami. Chceme nalézt „co nejlepší“ přibližné řešení $\hat{\mathbf{x}}$ v tom smyslu, aby vektor $A\hat{\mathbf{x}}$ byl co nejblíže pravé straně soustavy $\mathbf{b}$, tj. aby norma $\|\mathbf{b} - A\hat{\mathbf{x}}\|$ byla co nejmenší možná.

Metoda nejmenších čtverců je založená na měření normy vektorů pomocí eukleidovské normy (odtud její název), která je určená standardním skalárním součinem v $\mathbb{R}^n$ (nebo v $\mathbb{C}^n$). Zapíšeme-li součin $A\mathbf{x}$ jako lineární kombinaci sloupců matice A, můžeme se na tento problém podívat tak, že hledáme vektor $\hat{\mathbf{x}} = (x_1, \dots, x_n)^T$ tak, aby vektor $x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n$ byl co nejblíže vektoru $\mathbf{b}$. Pro libovolný vektor $\mathbf{x} \in \mathbb{R}^n$ (nebo $\mathbf{x} \in \mathbb{C}^n$) leží součin $A\mathbf{x}$ v podprostoru $\operatorname{Im}(A)$. Podle tvrzení 8.45 je mezi všemi vektory $A\mathbf{x} \in \operatorname{Im} A$ nejblíže k $\mathbf{b}$ ortogonální projekce $\mathbf{w}$ vektoru $\mathbf{b}$ na $\operatorname{Im} A$. Podle definice ortogonální projekce je vektor $\mathbf{w} \in \operatorname{Im} A$ ortogonální projekcí vektoru $\mathbf{b}$ na $\operatorname{Im} A$ právě když $(\mathbf{b} - \mathbf{w}) \perp (\operatorname{Im} A)$.

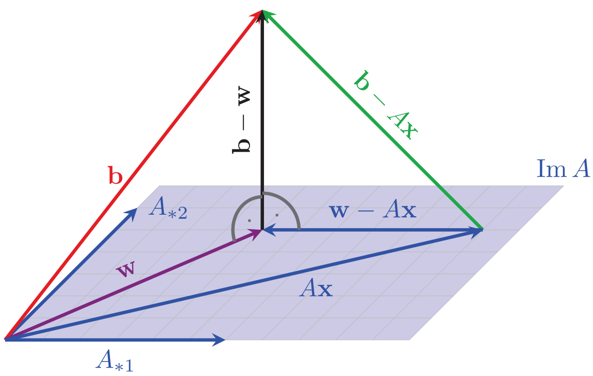
Definice 8.70. Je-li $A\mathbf{x} = \mathbf{b}$ soustava lineárních rovnic s reálnými (nebo komplexními) koeficienty. Každý vektor $\hat{\mathbf{x}} \in \mathbb{R}^n$ (nebo $\hat{\mathbf{x}} \in \mathbb{C}^n$) takový, že $A\hat{\mathbf{x}}$ se rovná ortogonální projekci vektoru pravých stran $\mathbf{b}$ na sloupcový prostor $\operatorname{Im} A$ matice A se nazývá *přibližné řešení* (nebo *aproximace řešení*) *soustavy* $A\mathbf{x} = \mathbf{b}$ *metodou nejmenších čtverců*.

Vektor $\hat{\mathbf{x}}$ je tedy aproximace řešení soustavy $A\mathbf{x} = \mathbf{b}$ metodou nejmenších čtverců právě když $(\mathbf{b} - A\hat{\mathbf{x}}) \perp (\operatorname{Im} A)$ neboli právě když $\mathbf{b} - A\hat{\mathbf{x}} \in (\operatorname{Im} A)^{\perp} = \operatorname{Ker}(A^*)$, což je právě když $A^*(\mathbf{b} - A\hat{\mathbf{x}}) = 0$, a to je právě když $A^*A\hat{\mathbf{x}} = A^*\mathbf{b}$. Dokázali jsme tak následující tvrzení.

Tvrzení 8.71. *Je-li A matice typu $m \times n$ nad $\mathbb{R}$ nebo $\mathbb{C}$ a $\mathbf{b} \in \mathbb{R}^m$ (resp. $\mathbb{C}^m$), pak množina všech přibližných řešení soustavy $A\mathbf{x} = \mathbf{b}$ metodou nejmenších čtverců je rovna množině všech (přesných) řešení soustavy*

$$A^*A\hat{\mathbf{x}} = A^*\mathbf{b} \quad .$$

Definice 8.72. Soustavu $A^*A\mathbf{x} = A^*\mathbf{b}$ nazýváme *soustava normálních rovnic* příslušná k soustavě $A\mathbf{x} = \mathbf{b}$.



Přeformulujeme si tvrzení 8.60 na tento důležitý speciální případ. Matice soustavy z tohoto tvrzení, tj. Gramova matice vektorů $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$, má na místě (i, j) číslo $\mathbf{a}_i \cdot \mathbf{a}_j = \mathbf{a}_i^* \mathbf{a}_j$. Je tedy rovná matici A^*A. Z tvrzení 8.63.1 také plyne, že matice A^*A je regulární právě když je posloupnost sloupcových vektorů $(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n)$ lineárně nezávislá. V takovém případě proto existuje inverzní matice $(A^*A)^{-1}$ a jednoznačně určenou aproximaci $\hat{\mathbf{x}}$ řešení soustavy $A\mathbf{x} = \mathbf{b}$ metodou nejmenších čtverců můžeme vyjádřit ve tvaru

$$\hat{\mathbf{x}} = (A^*A)^{-1} A^*\mathbf{b} \quad .$$

Všimněme si také, že v případě, že posloupnost sloupcových vektorů matice A je lineárně nezávislá, platí

$$(A^*A)^{-1} A^*A = I_n \quad ,$$

což znamená, že matice $(A^*A)^{-1} A^*$ je inverzní zleva k matici A. Nazývá se *Moore-Penroseova pseudoinverze* matice A. V případě, že matice A je regulární, je pseudoinverze $(A^*A)^{-1} A^*$ inverzní maticí k matici A a rovná se tedy A^{-1}.

A poslední poznámka. Je-li soustava $A\mathbf{x} = \mathbf{b}$ řešitelná, je $\mathbf{b} \in \operatorname{Im} A$ a tedy ortogonální projekce $\mathbf{w}$ vektoru $\mathbf{b}$ na podprostor $\operatorname{Im} A$ se rovná $\mathbf{b}$. Vektor $\hat{\mathbf{x}}$ je přibližným řešením soustavy $A\mathbf{x} = \mathbf{b}$ metodou nejmenších čtverců právě když platí $A\hat{\mathbf{x}} = \mathbf{w} = \mathbf{b}$, tj. právě když je (pravým) řešením soustavy $A\mathbf{x} = \mathbf{b}$.

Příklad 8.73. Řešení reálné soustavy $(A|\mathbf{b})$, kde

$$(A|\mathbf{b}) = \left(\begin{array}{cc|c} 2 & 0 & 3 \\ 1 & 1 & 5 \\ -2 & -1 & -2 \end{array} \right) ,$$

metodou nejmenších čtverců je řešení soustavy

$$A^T A\hat{\mathbf{x}} = A^T\mathbf{b}$$

$$\begin{pmatrix} 2 & 1 & -2 \\ 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 1 & 1 \\ -2 & -1 \end{pmatrix} \hat{\mathbf{x}} = \begin{pmatrix} 2 & 1 & -2 \\ 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} 3 \\ 5 \\ -2 \end{pmatrix}$$

$$\begin{pmatrix} 9 & 3 \\ 3 & 2 \end{pmatrix} \hat{\mathbf{x}} = \begin{pmatrix} 15 \\ 7 \end{pmatrix} .$$

Eliminací dostaneme $\hat{\mathbf{x}} = (x_1, x_2)^T = (1, 2)^T$.

Součin $A(1,2)^T = (2,3,-4)$ je ortogonální projekcí $\mathbf{w}$ vektoru $\mathbf{b}$ na podprostor $\operatorname{Im} A$. Rozdíl mezi vektorem pravých stran $\mathbf{b}$ a vektorem $\mathbf{w} = A\hat{\mathbf{x}}$ je potom

$$\mathbf{b} - \mathbf{w} = (3,5,-2)^T - (2,3,-4)^T = (1,2,2)^T$$

a vzdálenost $\mathbf{b}$ od $\operatorname{Im} A$ je tedy

$$\|\mathbf{b} - \mathbf{w}\| = \sqrt{1^2 + 2^2 + 2^2} = 3 \ .$$

8.7.2. *(Lineární) regrese.* Jedním z často využívaných příkladů použití metody nejmenších čtverců je *lineární regrese.* V této úloze chceme „co nejlépe“ proložit přímku $y = ax + b$ danými naměřenými hodnotami (x_1, y_1), (x_2, y_2), …, (x_n, y_n). Přesněji řečeno, hledáme aproximaci „řešení“ soustavy

$$\left(\begin{array}{cc|c} x_1 & 1 & y_1 \\ x_2 & 1 & y_2 \\ \vdots & \vdots & \vdots \\ x_n & 1 & y_n \end{array} \right) .$$

Vidíme, že posloupnost sloupcových vektorů matice soustavy je lineárně nezávislá, pokud se liší aspoň dvě hodnoty x_i pro $i = 1, 2, \ldots, n$.

Příklad 8.74. Metodou nejmenších čtverců proložíme body $(0,1)$, $(1,1)$, $(2,2)$, $(3,4)$, $(4,5)$ v $\mathbb{R}^2$ přímku $y = ax+b$. Dvojice koeficientů (a,b) je přibližným řešením soustavy lineárních rovnic

$$\left(\begin{array}{cc|c} 0 & 1 & 1 \\ 1 & 1 & 1 \\ 2 & 1 & 2 \\ 3 & 1 & 4 \\ 4 & 1 & 5 \end{array} \right)$$

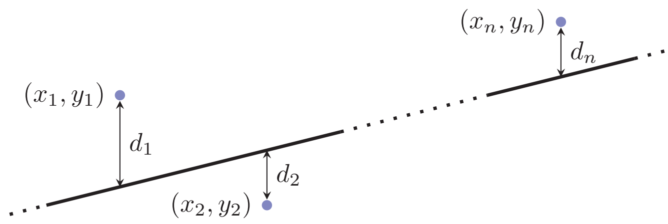
metodou nejmenších čtverců. Příslušná soustava normálních rovnic je

$$\begin{pmatrix} 0 & 1 & 2 & 3 & 4 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 1 \\ 2 & 1 \\ 3 & 1 \\ 4 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 0 & 1 & 2 & 3 & 4 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 2 \\ 4 \\ 5 \end{pmatrix} ,$$

$$\begin{pmatrix} 30 & 10 \\ 10 & 5 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 37 \\ 13 \end{pmatrix} .$$

Řešením vyjde $(a,b)^T = (11/10, 2/5)$ takže hledaná přímka je

$$y = \frac{11}{10}x + \frac{2}{5} \ .$$



OBRÁZEK 74. Lineární regrese – minimalizujeme $\sum d_i^2$.

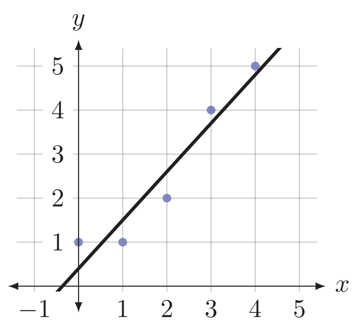
Daty můžeme prokládat složitější útvary, jako paraboly, polynomy vyššího stupně, elipsy (např. při hledání dráhy planety), apod. Také takové úlohy vedou na hledání řešení soustavy metodou nejmenších čtverců.

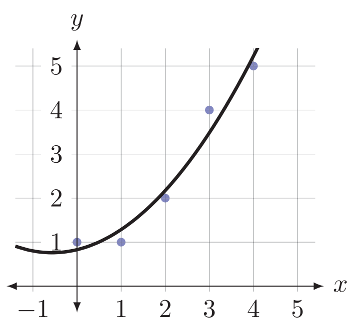
Příklad 8.75. Stejnými body proložíme „co nejlépe“ parabolu $y = ax^2 + bx + c$. Koeficienty jsou řešením soustavy

$$\left(\begin{array}{ccc|c} 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 4 & 2 & 1 & 2 \\ 9 & 3 & 1 & 4 \\ 16 & 4 & 1 & 5 \end{array}\right)$$

metodou nejmenších čtverců. Vyjde $(a,b,c)^T = 1/70(15, 17, 58)^T$,

$$y = \frac{3}{14}x^2 + \frac{17}{70}x + \frac{29}{35}$$





Metodu řešení posledních dvou příkladů můžeme zobecnit následujícím způsobem. Chceme danými daty (x_1, y_1), (x_2, y_2), $\dots$, (x_n, y_n) „co nejlépe“ proložit funkci, která je lineární kombinací předem zvolených reálných funkcí $f_1(x), f_2(x), \dots, f_q(x)$, kterým se v některých oborech říká *regresory*. V prvním z příkladů těmito funkcemi byla konstantní funkce $f_1(x) = 1$ a lineární funkce $f_2(x) = x$. Ve druhém z příkladů jsme si k těmto dvěma funkcím přidali ještě kvadratickou funkci $f_3(x) = x^2$.

Hledáme reálná čísla $c_1, c_2, \ldots, c_q$ taková, že lineární kombinace

$$\hat{f}(x) = c_1 f_1(x) + c_2 f_2(x) + \cdots + c_q f_q(x)$$

minimalizuje eukleidovskou vzdálenost mezi vektorem $\mathbf{y} = (y_1, y_2, \ldots, y_n)^T$ a vektorem $\mathbf{w} = (\hat{f}(x_1), \hat{f}(x_2), \ldots, \hat{f}(x_n))^T$, tj. eukleidovskou normu vektoru $\mathbf{y} - \mathbf{w}$.

Označíme $a_{ij} = f_j(x_i)$ pro $i = 1, 2, \ldots, n$ a $j = 1, 2, \ldots, q$. Potom pro každé $i = 1, 2, \ldots, n$ platí

$$\hat{f}(x_i) = c_1 f_1(x_i) + c_2 f_2(x_i) + \cdots + c_q f_q(x_i) = a_{i1}c_1 + a_{i2}c_2 + \cdots + a_{iq}c_q \ ,$$

což můžeme zapsat maticově ve tvaru $\mathbf{w} = A\hat{\mathbf{c}}$, kde $\hat{\mathbf{c}} = (c_1, c_2, \ldots, c_q)^T$. Ve sloupcovém prostoru $\operatorname{Im} A$ matice A tak hledáme vektor $A\hat{\mathbf{c}}$, který minimalizuje vzdálenost od vektoru $\mathbf{y}$. Pro hledaný vektor $\hat{\mathbf{c}}$ musí platit, že $A\hat{\mathbf{c}}$ je ortogonální projekcí vektoru $\mathbf{y}$ na podprostor $\operatorname{Im} A$. Vektor $\hat{\mathbf{c}}$ tedy najdeme jako přibližné řešení soustavy $\mathbf{y} = A\mathbf{c}$ metodou nejmenších čtverců, tj. jako (pravé) řešení soustavy

$$A^T A\hat{\mathbf{c}} = A^T\mathbf{y} \ .$$

A nakonec závěrečné varování před zneužíváním lineární regrese. Lineární regrese je exaktní metoda, kterou lze použít na jakákoliv data (x_1, y_1), (x_2, y_2), $\ldots$, (x_n, y_n), pokud všechny body neleží na jedné přímce kolmé k první souřadné ose x. Smysluplné je takové použití pouze v případě, kdy máme nějaký dobrý důvod si myslet, že mezi proměnnými x a y existuje nějaká lineární závislost tvaru $y = ax+b$ a naše naměřená data na přímce neleží pouze kvůli nepřesnostem v měření.

Dobrým důvodem rozhodně není tvrzení firmy OCIS, že výsledky v testu matematiky od OCIS lineárně závisí na výsledcích testu obecných studijních předpokladů od téže firmy. Takový předpoklad není ničím podložený, z podstaty věci je nesmyslný, naměřená data jej nepodporují, a jakékoliv závěry plynoucí z použití lineární regrese na takto získaná data nelze nazvat jinak než blábolem.

8.7.3. *Optimalizační úlohy (matematické programování).*

Budu průběžně doplňovat.

8.7.4. *Klasifikační úlohy.*

8.7.5. *Atd.*

Cvičení

1. Jsou-li A, B matice nad tělesem $\mathbb{C}$ typu $m \times n$, C je matice typu $n \times p$ nad $\mathbb{C}$ a $a \in \mathbb{C}$, pak

(1) $(A+B)^* = A^* + B^*$,
(2) $(aA)^* = \overline{a}A^*$,
(3) $(A^*)^* = A$.
(4) $(BC)^* = C^*B^*$.

Dokažte.

2. Nechť A je čtvercová matice nad $\mathbb{C}$. Dokažte, že $\det(A^*) = (\det(A))^*$.

3. Nechť A je regulární matice nad $\mathbb{C}$. Dokažte, že $(A^*)^{-1} = (A^{-1})^*$.

4. Nechť A je čtvercová matice řádu n nad $\mathbb{C}$. Dokažte, že zobrazení $\mathbb{C} \times \mathbb{C} \to \mathbb{C}$ definované vztahem $\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u}^* A \mathbf{v}$ splňuje podmínky (SL1) a (SL2).

5. Nechť A je čtvercová matice řádu n nad $\mathbb{C}$. Dokažte, že zobrazení $\mathbb{C} \times \mathbb{C} \to \mathbb{C}$ definované vztahem $\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u}^* A \mathbf{v}$ splňuje podmínku (SCS) právě tehdy, když A je hermitovská (tj. $A^* = A$).

6. Nechť B je regulární matice řádu n nad $\mathbb{C}$ a $A = B^*B$. Dokažte, že zobrazení $\mathbb{C} \times \mathbb{C} \to \mathbb{C}$ definované vztahem $\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u}^* A \mathbf{v}$ je skalární součin.

7. Dokažte, že v libovolném vektorovém prostoru se skalárním součinem $\langle\, , \rangle$ platí

- $\mathrm{Re}\,(\langle \mathbf{u}, \mathbf{v} \rangle) = \frac{1}{2}(\|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - \|\mathbf{u} - \mathbf{v}\|^2)$
- $\mathrm{Re}\,(\langle \mathbf{u}, \mathbf{v} \rangle) = \frac{1}{4}(\|\mathbf{u} + \mathbf{v}\|^2 - \|\mathbf{u} - \mathbf{v}\|^2)$
- $\mathrm{Im}\,(\langle \mathbf{u}, \mathbf{v} \rangle) = \frac{1}{2}(\|\mathbf{u} - i\mathbf{v}\|^2 - \|\mathbf{u}\|^2 - \left\|\mathbf{v}^2\right\|^2)$
- $\mathrm{Im}\,(\langle \mathbf{u}, \mathbf{v} \rangle) = \frac{1}{2}(\|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - \|\mathbf{u} + i\mathbf{v}\|^2)$
- $\mathrm{Im}\,(\langle \mathbf{u}, \mathbf{v} \rangle) = \frac{1}{4}(\|\mathbf{u} - i\mathbf{v}\|^2 - \|\mathbf{u} + i\mathbf{v}\|^2)$

$\mathrm{Im}\,(x)$ značí imaginární část čísla $x \in \mathbb{C}$.

8. Nad reálnými čísly lze Cauchy-Schwarzovu nerovnost dokázat také následujícím způsobem: Výraz $\|\mathbf{u} + t\mathbf{v}\|^2$ definuje kvadratickou funkci. Protože musí být nezáporná, její diskriminant je nekladný a to dává C-S nerovnost. Doplňte detaily.

9. Kdy nastává v trojúhelníkové nerovnosti rovnost?

10. Dokažte, že norma pochází ze skalárního součinu právě tehdy, když splňuje rovnoběžníkové pravidlo.

11. Dokažte, že platí-li $M \perp N$, pak $M \cap N \subseteq \{\mathbf{o}\}$.

12. Dokažte pozorování 8.65.

13. Dokažte, že prostorech nad $\mathbb{R}$ se skalárním součinem platí opačná implikace v Pythagorově větě, tj. pokud $\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$, pak $\mathbf{u} \perp \mathbf{v}$. Platí opačná implikace v prostorech nad $\mathbb{C}$?

14. Nechť $f : \mathbf{V} \to \mathbf{W}$ je lineární zobrazení a $\mathbf{U} \leq \mathbf{V}$ je doplněk $\mathrm{Ker}\, f$, tj. $\mathrm{Ker}\, f \oplus U = V$. Dokažte, že zúžení f na $\mathbf{U}$ je izomorfismus z $\mathbf{U}$ na obraz f.

15. Dokažte, že determinant Gramovy matice vektorů $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n \in \mathbb{R}^n$ je rovný druhé mocnině determinantu matice

$$(\mathbf{w}_1|\mathbf{w}_2|\ldots|\mathbf{w}_n)\ .$$

Interpretujte geometricky.

16. Pomocí Gramovy-Schmidtovy ortogonalizace dokažte body (2) a (3) věty 8.68 za předpokladu, že W je konečně generovaný (prostor V konečně generovaný být nemusí).

17. Využijte QR-rozklad na důkaz následující nerovnosti pro komplexní matici $A = (\mathbf{a}_1|\ldots|\mathbf{a}_n)$ typu $m \times n$ a standardní skalární součin:

$$\det(A^*A) \leq \|\mathbf{a}_1\|^2\, \|\mathbf{a}_2\|^2 \ldots \|\mathbf{a}_n\|^2$$

Připomeňme si geometrický význam determinantu $\det(A^* A)$ a interpretujte nerovnost geoemetricky.

18. Dokažte, že součinem unitárních matic stejných řádů je unitární matice.

Shrnutí osmé kapitoly

(1) Pro dva n-složkové aritmetické vektory $\mathbf{u} = (x_1, x_2, \ldots, x_n)^T$, $\mathbf{v} = (y_1, y_2, \ldots, y_n) \in \mathbb{R}^n$ definujeme jejich *standardní skalární součin* jako reálné číslo
$$\mathbf{u} \cdot \mathbf{v} = x_1y_1 + x_2y_2 + \cdots + x_ny_n \ .$$
(2) *Eukleidovská norma* nebo také *eukleidovská délka* vektoru $\mathbf{u} \in \mathbb{R}^n$ je číslo
$$\|\mathbf{u}\| = \sqrt{\mathbf{u} \cdot \mathbf{u}} = \sqrt{\mathbf{u}^\mathbf{T}\,\mathbf{u}} \ .$$
Eukleidovskou normu vektoru $\mathbf{u} = (x_1, x_2, \ldots, x_n)^T$ tak spočítáme jako
$$\|\mathbf{u}\| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2} \ .$$
(3) Geometrický význam standardního skalárního součinu je vyjádřen vztahem
$$\mathbf{u} \cdot \mathbf{v} = \|\mathbf{u}\|\ \|\mathbf{v}\| \cos\alpha \ ,$$
kde α je úhel, který vektory $\mathbf{u}$ a $\mathbf{v}$ svírají.

(4) Jiný geometrický význam spočívá v tom, že absolutní hodnota součinu $\|\mathbf{v}\| \cos\alpha$ je délkou ortogonální projekce vektoru $\mathbf{v}$ do přímky určené vektorem $\mathbf{u} \neq \mathbf{o}$. Projekce má stejný směr jako vektor $\mathbf{u}$ v případě, že oba vektory $\mathbf{u}, \mathbf{v}$ svírají úhel menší než $\pi/2$, a má opačný směr, pokud oba vektory svírají úhel větší než $\pi/2$.

(5) Rovnici přímky v rovině $a_1x_1 + a_2x_2 = b$ můžeme přepsat pomocí standardního skalární součinu do tvaru
$$\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = b \ ,$$
a protože v případě rovnice přímky je vektor $(a_1, a_2)^T \neq \mathbf{o}$, jde o množinu všech bodů (x_1, x_2) v rovině, jejichž polohové vektory $\mathbf{x} = (x_1, x_2)^T$ mají stejnou ortogonální projekci do přímky $\langle \mathbf{a} \rangle$. Vektor $\mathbf{a}$ nazýváme *normálovým vektorem* přímky $a_1x_1 + a_2x_2 = b$.

(6) Jsou-li $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^n$ libovolné reálné aritmetické vektory a $a \in \mathbb{R}$ skalár, pak platí
 (a) $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$,
 (b) $\mathbf{u} \cdot (\mathbf{v} + \mathbf{w}) = \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w}$,
 (c) $\mathbf{u} \cdot (a\mathbf{v}) = a(\mathbf{u} \cdot \mathbf{v})$,
 (d) $\mathbf{u} \cdot \mathbf{u} \geq 0$ a $\mathbf{u} \cdot \mathbf{u} = 0$ právě když $\mathbf{u} = \mathbf{o}$.

(7) Pro dva komplexní aritmetické vektory $\mathbf{u} = (x_1,\ x_2,\ \ldots,\ x_n)^T$ a $\mathbf{v} = (y_1, y_2, \ldots, y_n)^T$ definujeme *standardní skalární součin* $\mathbf{u} \cdot \mathbf{v}$ předpisem
$$\mathbf{u} \cdot \mathbf{v} = \overline{x_1}y_1 + \overline{x_2}y_2 + \cdots + \overline{x_n}y_n \ ,$$
kde $\overline{x}$ značí číslo komplexně sdružené k x, tj. $\overline{a + bi} = a - bi$.

(8) *Eukleidovskou délku* nebo také *eukleidovskou normu* aritmetického vektoru $\mathbf{u} = (x_1,\ x_2,\ \ldots,\ x_n)^T \in \mathbb{C}^n$ definujeme jako
$$\|\mathbf{u}\| = \sqrt{\mathbf{u} \cdot \mathbf{u}} = \sqrt{\overline{x_1}x_1 + \overline{x_2}x_2 + \cdots + \overline{x_n}x_n} = \sqrt{|x_1|^2 + |x_2|^2 + \cdots + |x_n|^2} \ .$$
(9) *Hermitovsky sdružená matice* k matici $A = (a_{ij})_{m\times n}$ je matice $A^* = (b_{ji})_{n\times m}$, kde $b_{ji} = \overline{a_{ij}}$ pro libovolné indexy $i \in \{1, 2, \ldots, m\}$ a $j \in \{1, 2, \ldots, n\}$.

(10) Pro libovolné tři vektory $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{C}^n$ a komplexní číslo a platí
 (a) $\mathbf{u} \cdot \mathbf{v} = \overline{\mathbf{v} \cdot \mathbf{u}}$,

   (b) $\mathbf{u} \cdot (\mathbf{v} + \mathbf{w}) = \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w}$,
   (c) $\mathbf{u} \cdot (a\,\mathbf{v}) = a\,(\mathbf{u} \cdot \mathbf{v})$,
   (d) $\mathbf{u} \cdot \mathbf{u}$ je nezáporné reálné číslo, a $\mathbf{u} \cdot \mathbf{u} = 0$ právě když $\mathbf{u} = \mathbf{o}$.

(11) Pro libovolné tři vektory $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{C}^n$ a komplexní číslo a platí
   (a) $(\mathbf{u} + \mathbf{v}) \cdot \mathbf{w} = \mathbf{u} \cdot \mathbf{w} + \mathbf{v} \cdot \mathbf{w}$,
   (b) $(a\mathbf{u}) \cdot \mathbf{v} = \overline{a}\ (\mathbf{u} \cdot \mathbf{v})$.

(12) Je-li $\mathbf{V}$ lineární prostor nad $\mathbb{R}$ (resp. nad $\mathbb{C}$), pak se zobrazení $\langle\,,\rangle$ z $V \times V$ do $\mathbb{R}$ (resp do $\mathbb{C}$), které dvojici $\mathbf{u}, \mathbf{v}$ přiřadí skalár $\langle\, \mathbf{u}, \mathbf{v}\rangle$, nazývá *skalární součin* na $\mathbf{V}$, pokud pro libovolné $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ a $a \in \mathbb{R}$ (resp. $a \in \mathbb{C}$) platí
   (SCS) $\langle\, \mathbf{u}, \mathbf{v}\rangle = \overline{\langle\, \mathbf{v}, \mathbf{u}\rangle}$,
   (SL1) $\langle\, \mathbf{u}, a\mathbf{v}\rangle = a\,\langle\, \mathbf{u}, \mathbf{v}\rangle$,
   (SL2) $\langle\, \mathbf{u}, \mathbf{v} + \mathbf{w}\rangle = \langle\, \mathbf{u}, \mathbf{v}\rangle + \langle\, \mathbf{u}, \mathbf{w}\rangle$,
   (SP) $\langle\, \mathbf{u}, \mathbf{u}\rangle$ je nezáporné reálné číslo, které je nulové právě tehdy, když $\mathbf{u} = \mathbf{o}$.

(13) Je-li $\mathbf{V}$ lineární prostor nad $\mathbb{R}$ (resp. nad $\mathbb{C}$) se skalárním součinem $\langle\,,\rangle$, pak pro prvky libovolné $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ a skalár a platí
   (a) $\langle\, \mathbf{u}, \mathbf{o}\rangle = 0 = \langle\, \mathbf{o}, \mathbf{u}\rangle$
   (b) $\langle\, a\mathbf{u}, \mathbf{v}\rangle = \overline{a}\,\langle\, \mathbf{u}, \mathbf{v}\rangle$
   (c) $\langle\, \mathbf{u} + \mathbf{v}, \mathbf{w}\rangle = \langle\, \mathbf{u}, \mathbf{w}\rangle + \langle\, \mathbf{v}, \mathbf{w}\rangle$.

(14) Komplexním maticím A, které splňují rovnost $A^* = A$ říkáme *hermitovské.*

(15) Hermitovská matice A řádu n se nazývá *pozitivně definitní*, pokud $\mathbf{u}^* A\mathbf{u}$ je nezáporné reálné číslo pro libovolné $\mathbf{u} \in \mathbb{C}^n$ a rovná se 0 právě když $\mathbf{u} = \mathbf{o}$.

(16) Je-li $A = B^* B$, pak zobrazení definované $\langle\, \mathbf{u}, \mathbf{v}\rangle = \mathbf{u}^* A\mathbf{v}$ je skalární součin na $\mathbb{C}^n$ (nebo na $\mathbb{R}^n$).

(17) Nechť $\mathbf{V}$ je lineární prostor se skalárním součinem $\langle\,,\rangle$. *Normou* vektoru $\mathbf{v} \in V$ rozumíme reálné číslo

$$\|\mathbf{u}\| = \sqrt{\langle\, \mathbf{u}, \mathbf{u}\rangle}\ .$$

Vektor $\mathbf{u}$ se nazývá *jednotkový*, pokud $\|\mathbf{u}\| = 1$.

(18) Nechť $\mathbf{V}$ je lineární prostor nad $\mathbb{R}$ (resp. $\mathbb{C}$) se skalárním součinem $\langle\,,\rangle$, $\mathbf{u}, \mathbf{v} \in V$ a $t \in \mathbb{R}$ (resp. $t \in \mathbb{C}$). Pak platí
   (a) $\|\mathbf{u}\| \geq 0$, přičemž $\|\mathbf{u}\| = 0$ právě tehdy, když $\mathbf{u} = \mathbf{o}$,
   (b) $\|t\mathbf{u}\| = |t|\,\|\mathbf{u}\|$,
   (c) $\|\mathbf{u} + \mathbf{v}\|^2 + \|\mathbf{u} - \mathbf{v}\|^2 = 2\,\|\mathbf{u}\|^2 + 2\,\|\mathbf{v}\|^2$, (rovnoběžníkové pravidlo),
   (d) $\mathrm{Re}\,(\langle\, \mathbf{u}, \mathbf{v}\rangle) = \frac{1}{2}(\|\mathbf{u} + \mathbf{v}\|^2 - \|\mathbf{u}\|^2 - \|\mathbf{v}\|^2)$, (polarizační identita),
   kde $\mathrm{Re}\,(x)$ značí reálnou část x.

(19) Cauchyho-Schwarzova nerovnost. Nechť $\mathbf{V}$ je lineární prostor se skalárním součinem $\langle\,,\rangle$ a $\mathbf{u}, \mathbf{v} \in V$. Pak platí

$$|\,\langle\, \mathbf{u}, \mathbf{v}\rangle\,| \leq \|\mathbf{u}\|\,\|\mathbf{v}\|\ ,$$

přičemž rovnost nastává právě tehdy, když $(\mathbf{u}, \mathbf{v})$ je lineárně závislá posloupnost.

(20) Trojúhelníková nerovnost. Nechť $\mathbf{V}$ je lineární prostor se skalárním součinem $\langle\,,\rangle$ a $\mathbf{u}, \mathbf{v} \in V$. Pak platí

$$\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|\ .$$

(21) Nechť $\mathbf{V}$ je lineární prostor nad $\mathbb{R}$ se skalárním součinem $\langle\,,\rangle$ a $\mathbf{o} \neq \mathbf{u}, \mathbf{v} \in V$. *Úhlem mezi prvky* $\mathbf{u}$ *a* $\mathbf{v}$ rozumíme reálné číslo $\alpha \in \langle 0, \pi\rangle$ splňující

$$\cos\alpha = \frac{\langle\, \mathbf{u}, \mathbf{v}\rangle}{\|\mathbf{u}\|\,\|\mathbf{v}\|} \quad .$$

(22) Kosinová věta. Nechť $\mathbf{V}$ je lineární prostor nad $\mathbb{R}$ se skalárním součinem $\langle\,,\rangle$ a $\mathbf{o} \neq \mathbf{u}, \mathbf{v} \in V$. Pak platí

$$\|\mathbf{u} - \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - 2\,\|\mathbf{u}\|\,\|\mathbf{v}\| \cos\alpha \quad ,$$

kde α je úhel mezi vektory $\mathbf{u}$ a $\mathbf{v}$.

(23) Je-li $\mathbf{V}$ lineární prostor nad $\mathbb{C}$ (nebo nad $\mathbb{R}$), pak zobrazení $\|\cdot\|$, které přiřazuje každému prvku $\mathbf{u}$ reálné číslo $\|\mathbf{u}\|$ nazýváme *norma* na prostoru $\mathbf{V}$, pokud platí pro kažé dva prvky $\mathbf{u}, \mathbf{v} \in V$ a každý skalár t
- (a) $\|\mathbf{u}\| \geq 0$, přičemž $\|\mathbf{u}\| = 0$ právě tehdy, když $\mathbf{u} = \mathbf{o}$,
- (b) $\|t\mathbf{u}\| = |t|\,\|\mathbf{u}\|$,
- (c) $\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$.

(24) Nechť $\mathbf{V}$ je lineární prostor se skalárním součinem $\langle\,,\rangle$. Prvky $\mathbf{u}, \mathbf{v} \in V$ nazýváme *kolmé* (nebo *ortogonální*) a píšeme $\mathbf{u} \perp \mathbf{v}$, pokud $\langle\, \mathbf{u}, \mathbf{v}\rangle = 0$.

Množina, nebo posloupnost, M prvků V se nazývá *ortogonální*, pokud $\mathbf{u} \perp \mathbf{v}$ pro libovolné dva různé prvky množiny (nebo posloupnosti) M.

Množina (posloupnost) M se nazývá *ortonormální*, pokud je ortogonální a každý vektor v M je jednotkový.

(25) Je-li $\mathbf{V}$ lineární prostor se skalárním součinem $\langle\,,\rangle$, pak každá ortogonální posloupnost nenulových prvků V je lineárně nezávislá.

(26) Pythagorova věta. Je-li $\mathbf{V}$ lineární prostor se skalárním součinem $\langle\,,\rangle$ a jsou-li vektory $\mathbf{u}, \mathbf{v} \in V$ kolmé, pak platí

$$\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 \quad .$$

(27) Indukcí lze Pythagorovu větu zobecnit na libovolný konečný počet prvků: je-li $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ ortogonální množina, pak

$$\|\mathbf{v}_1 + \mathbf{v}_2 + \cdots + \mathbf{v}_k\|^2 = \|\mathbf{v}_1\|^2 + \|\mathbf{v}_2\|^2 + \cdots + \|\mathbf{v}_k\|^2 \quad .$$

(28) Je-li $\mathbf{V}$ lineární prostor se skalárním součinem $\langle\,,\rangle$, $B = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$ nějaká ortonormální báze ve $\mathbf{V}$ a $\mathbf{u} \in V$, pak platí

$$\mathbf{u} = \langle\, \mathbf{v}_1, \mathbf{u}\rangle\, \mathbf{v}_1 + \langle\, \mathbf{v}_2, \mathbf{u}\rangle\, \mathbf{v}_2 + \cdots + \langle\, \mathbf{v}_n, \mathbf{u}\rangle\, \mathbf{v}_n \quad .$$

Jinými slovy,

$$[\mathbf{u}]_B = (\langle\, \mathbf{v}_1, \mathbf{u}\rangle, \langle\, \mathbf{v}_2, \mathbf{u}\rangle, \ldots, \langle\, \mathbf{v}_n, \mathbf{u}\rangle)^T \quad .$$

(29) Souřadnicím vzhledem k ortonormální bázi se někdy říká *Fourierovy koeficienty* vzhledem k této bázi.

(30) Je-li $\mathbf{V}$ lineární prostor se skalárním součinem $\langle\,,\rangle$, $B = (\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n)$ jeho ortonormální báze, a $\mathbf{u}, \mathbf{w} \in V$, pak

$$\langle\, \mathbf{u}, \mathbf{w}\rangle = [\mathbf{u}]_B^*[\mathbf{w}]_B \quad .$$

(31) Je-li $\mathbf{V}$ lineární prostor se skalárním součinem $\langle\,,\rangle$ a $\mathbf{v} \in V$, $M, N \subseteq V$, pak říkáme, že prvek $\mathbf{v}$ *je kolmý na* M, pokud $\mathbf{v}$ je kolmý na každý prvek z množiny M, což zapisujeme $\mathbf{v} \perp M$.

Říkáme, že M *je kolmá na* N a zapisujeme $M \perp N$, pokud každý prvek množiny M je kolmý na každý prvek množiny N.

(32) Je-li $\mathbf{V}$ lineární prostor se skalárním součinem $\langle\,,\rangle$ a $M, N \subseteq V$, pak $M \perp N$ právě když $M \perp \langle N\rangle$ což je právě když $\langle M\rangle \perp \langle N\rangle$.

(33) Je-li $\mathbf{V}$ lineární prostor se skalárním součinem $\langle\,,\rangle$, $\mathbf{v} \in V$ a $\mathbf{W}$ podprostor $\mathbf{U}$, pak prvek $\mathbf{w} \in W$ nazýváme *ortogonální projekce* $\mathbf{v}$ *na podprostor* $\mathbf{W}$, pokud platí $(\mathbf{v} - \mathbf{w}) \perp W$.

(34) Je-li $\mathbf{W}$ podprostor lineárního prostoru $\mathbf{V}$ se skalárním součinem $\langle\,,\rangle$, $\mathbf{v} \in V$ a $\mathbf{w}$ ortogonální projekce prvku $\mathbf{v}$ na podprostor $\mathbf{W}$, pak pro každý prvek $\mathbf{w} \neq \mathbf{u} \in W$ platí

$$\|\mathbf{v} - \mathbf{w}\| < \|\mathbf{v} - \mathbf{u}\| \quad .$$

Ortogonální projekce $\mathbf{v}$ na podprostor $\mathbf{W}$ je určena jednoznačně, pokud existuje.

(35) Je-li $\mathbf{V}$ lineární prostor se skalárním součinem $\langle\,,\rangle$, $\mathbf{v} \in V$, a $\mathbf{W}$ konečně generovaný podprostor $\mathbf{V}$ s ortonormální bází $(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k)^T$, pak prvek

$$\mathbf{w} = \langle\, \mathbf{u}_1, \mathbf{v}\rangle\, \mathbf{u}_1 + \langle\, \mathbf{u}_2, \mathbf{v}\rangle\, \mathbf{u}_2 + \cdots + \langle\, \mathbf{u}_k, \mathbf{v}\rangle\, \mathbf{u}_k$$

je ortogonální projekcí vektoru $\mathbf{v}$ na podprostor $\mathbf{W}$.

(36) Je-li $\mathbf{V}$ lineární prostor se skalárním součinem $\langle\,,\rangle$, $\mathbf{v} \in V$, a $\mathbf{W}$ konečně generovaný podprostor $\mathbf{V}$ s ortogonální bází $B = (\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k)^T$, pak prvek

$$\mathbf{w} = \frac{\langle\, \mathbf{u}_1, \mathbf{v}\rangle}{\|\mathbf{u}_1\|^2}\, \mathbf{u}_1 + \frac{\langle\, \mathbf{u}_2, \mathbf{v}\rangle}{\|\mathbf{u}_2\|^2}\, \mathbf{u}_2 + \cdots + \frac{\langle\, \mathbf{u}_k, \mathbf{v}\rangle}{\|\mathbf{u}_k\|^2}\, \mathbf{u}_k$$

je ortogonální projekcí vektoru $\mathbf{v}$ na podprostor $\mathbf{W}$.

(37) *Gramova-Schmidtova ortogonalizace* je algoritmus, který dostane *na vstupu* nějakou lineárně nezávislou posloupnost

$$(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k)$$

prvků lineárního prostoru $\mathbf{V}$ se skalárním součinem. *Na výstupu* vydá ortonormální posloupnost

$$(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k)$$

prvků prostoru $\mathbf{V}$, která splňuje podmínku

$$\langle \mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_i\rangle = \langle \mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_i\rangle$$

pro každé $i = 1, 2, \dots, k$.

(38) Gramova-Schmidtova ortogonalizace spočívá v k-násobném iterování cyklu, jehož i-tý průběh sestává ze dvou kroků

(ia) **ortogonalizace**: najdeme prvek

$$\mathbf{v}_i - \mathbf{w}_{i-1} = \mathbf{v}_i - \langle\, \mathbf{u}_1, \mathbf{v}_i\rangle\ \mathbf{u}_1 - \langle\, \mathbf{u}_2, \mathbf{v}_i\rangle\ \mathbf{u}_2 - \cdots - \langle\, \mathbf{u}_1, \mathbf{v}_{i-1}\rangle\ \mathbf{u}_{i-1},$$

(ib) **normalizace**: položíme

$$\mathbf{u}_i = \frac{\mathbf{v}_i - \mathbf{w}_{i-1}}{\|\mathbf{v}_i - \mathbf{w}_{i-1}\|} \quad .$$

(39) Je-li $\mathbf{W}$ podprostor konečně generovaného lineárního prostoru $\mathbf{V}$ se skalárním součinem, pak každou ortonormální (ortogonální) bázi v podprostoru $\mathbf{W}$ lze doplnit na ortonormální (ortogonální) bázi celého prostoru $\mathbf{V}$.

Speciálně, v každém konečně generovaném lineárním prostoru se skalárním součinem existuje ortonormální báze.

(40) Je-li $\mathbf{V}$ lineární prostor dimenze n nad $\mathbb{R}$ (nebo nad $\mathbb{C}$) se skalárním součinem $\langle\,,\rangle$, pak existuje izomorfismus $f : \mathbf{V} \to \mathbb{R}^n$ (nebo $f : \mathbf{V} \to \mathbb{C}^n$), pro který platí

$$\langle\, \mathbf{u}, \mathbf{v} \rangle = f(\mathbf{u}) \cdot f(\mathbf{v})$$

pro každé dva prvky $\mathbf{u}, \mathbf{v} \in \mathbf{V}$.

(41) QR-rozkladu. Je-li A reálná nebo komplexní matice typu $n \times k$ s lineárně nezávislými sloupci, pak existuje matice Q typu $n \times k$ nad stejným tělesem s ortonormálními sloupci a horní trojúhelníková matice R řádu k s kladnými reálnými prvky na hlavní diagonále taková, že platí $A = QR$.

(42) Čtvercová reálná matice A řádu n se nazývá *ortogonální*, platí-li $A^T A = I_n$, čtvercová komplexní matice U řádu n se nazývá *unitární*, platí-li $U^* U = I_n$.

(43) Je-li Q reálná (resp. komplexní) čtvercová matice řádu n, pak jsou následující tvrzení ekvivalentní.
   (a) Q je ortogonální (resp. unitární),
   (b) zobrazení $f_Q(\mathbf{q}) = Q\,\mathbf{u}$ zachovává standardní skalární součin, tj. pro libovolné $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ (resp. $\mathbb{C}^n$) platí $Q\mathbf{u} \cdot Q\mathbf{v} = \mathbf{u} \cdot \mathbf{v}$,
   (c) f_Q zachovává eukleidovskou normu, tj. pro libovolný vektor $\mathbf{v} \in \mathbb{R}^n$ (resp. $\mathbb{C}^n$) platí $\|Q\mathbf{v}\| = \|\mathbf{v}\|$,
   (d) f_Q zobrazuje každou ortonormální bázi na ortonormální bázi,
   (e) $Q^{-1} = Q^T$, (resp. $Q^{-1} = Q^*$),
   (f) posloupnost řádkových vektorů matice Q tvoří ortonormální bázi v $\mathbb{R}^n$ (resp. $\mathbb{C}^n$) se standardním skalárním součinem,
   (g) posloupnost sloupcových vektorů matice Q tvoří ortonormální bázi v $\mathbb{R}^n$ (resp. $\mathbb{C}^n$) se standardním skalárním součinem.

(44) Součin dvou ortogonálních (resp. unitárních) matic téhož řádu je opět ortogonální (resp. unitární) matice.

(45) Je-li A regulární (reálná nebo kompexní) matice řádu a a $A = Q_1 R_1 = Q_2 R_2$ jsou dva QR-rozklady matice A, pak platí $Q_1 = Q_2$ a $R_1 = R_2$.

(46) Je-li $\mathbf{W}$ konečně generovaný podprostor lineárního prostoru $\mathbf{V}$ se skalárním součinem $\langle\,,\rangle$, $\mathbf{v} \in V$ a $B = (\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k)$ báze ve $\mathbf{W}$, pak vektor souřadnic $[\mathbf{w}]_B = (a_1, a_2, \ldots, a_k)^T$ ortogonální projekce prvku $\mathbf{v}$ na podprostor $\mathbf{W}$ spočteme jako řešení soustavy lineárních rovnic

$$\left(\begin{array}{cccc|c} \langle\, \mathbf{u}_1, \mathbf{u}_1 \rangle & \langle\, \mathbf{u}_1, \mathbf{u}_2 \rangle & \cdots & \langle\, \mathbf{u}_1, \mathbf{u}_k \rangle & \langle\, \mathbf{u}_1, \mathbf{v} \rangle \\ \langle\, \mathbf{u}_2, \mathbf{u}_1 \rangle & \langle\, \mathbf{u}_2, \mathbf{u}_2 \rangle & \cdots & \langle\, \mathbf{u}_2, \mathbf{u}_k \rangle & \langle\, \mathbf{u}_2, \mathbf{v} \rangle \\ \vdots & & \ddots & \vdots & \vdots \\ \langle\, \mathbf{u}_k, \mathbf{u}_1 \rangle & \langle\, \mathbf{u}_k, \mathbf{u}_2 \rangle & \cdots & \langle\, \mathbf{u}_k, \mathbf{u}_k \rangle & \langle\, \mathbf{u}_1, \mathbf{v} \rangle \end{array}\right).$$

(47) Jsou-li $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k$ prvky lineárního prostoru se skalárním součinem, pak čtvercovou matici

$$\begin{pmatrix} \langle\, \mathbf{u}_1, \mathbf{u}_1 \rangle & \langle\, \mathbf{u}_1, \mathbf{u}_2 \rangle & \cdots & \langle\, \mathbf{u}_1, \mathbf{u}_k \rangle \\ \langle\, \mathbf{u}_2, \mathbf{u}_1 \rangle & \langle\, \mathbf{u}_2, \mathbf{u}_2 \rangle & \cdots & \langle\, \mathbf{u}_2, \mathbf{u}_k \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle\, \mathbf{u}_k, \mathbf{u}_1 \rangle & \langle\, \mathbf{u}_k, \mathbf{u}_2 \rangle & \cdots & \langle\, \mathbf{u}_k, \mathbf{u}_k \rangle \end{pmatrix}$$

řádu k nazýváme *Gramova matice* posloupnosti prvků $(\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k)$.

(48) V případě standardního skalárního součinu na aritmetickém prostoru $\mathbb{R}^n$ můžeme Gramovu matici spočítat rychle tak, že si vektory $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k$

zapíšeme jako sloupce matice

$$A = (\mathbf{u}_1|\mathbf{u}_2|\cdots|\mathbf{u}_k)$$

a spočteme $A^T A$. V případě standardního skalárního součinu na $\mathbb{C}^n$ je to matice $A^* A$.

(49) Pro Gramovu matici $B = (\langle \mathbf{u}_i, \mathbf{u}_j \rangle)$ posloupnosti prvků $(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k)$ nějakého lineárního prostoru $\mathbf{V}$ se skalárním součinem platí
   (a) matice B je regulární právě když je posloupnost $(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k)$ lineárně nezávislá,
   (b) matice B je symetrická (hermitovská v komplexním případě),
   (c) je-li posloupnost $(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k)$ lineárně nezávislá, pak je Gramova matice B pozitivně definitní.

(50) Je-li $\mathbf{V}$ prostor se skalárním součinem $\langle\, , \rangle$ a $M \subseteq V$, pak *ortogonální doplněk* $M^\perp$ množiny M je množina všech prvků $\mathbf{V}$ kolmých na každý prvek M, tj.

$$M^\perp = \{\mathbf{v} \in \mathbf{V} : \mathbf{v} \perp M\} \ .$$

(51) Je-li $\mathbf{V}$ prostor se skalárním součinem $\langle\, , \rangle$ a $M, N \subseteq V$, pak platí
   (a) $M^\perp = \langle M \rangle^\perp$,
   (b) $M^\perp$ je podprostor V,
   (c) je-li $M \subseteq N$, pak $N^\perp \subseteq M^\perp$.

(52) Nechť V je konečně generovaný prostor dimenze n se skalárním součinem $\langle\, , \rangle$ a W je podprostor V. Pak platí
   (a) $\dim(W^\perp) = n - \dim(W)$,
   (b) $V = W \oplus W^\perp$,
   (c) $(W^\perp)^\perp = W$.

(53) Pro každou komplexní matici A platí

$$(\operatorname{Im} A^*)^\perp = \operatorname{Ker} A \quad \text{a} \quad (\operatorname{Im} A)^\perp = \operatorname{Ker} A^* \ .$$

(54) Nechť V je konečně generovaný prostor se skalárním součinem $\langle\, , \rangle$, B jeho ortonormální báze, $M = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$. Označíme A matici s řádky $[\mathbf{v}_1]_B^*$, $[\mathbf{v}_2]_B^*$, ..., $[\mathbf{v}_k]_B^*$. Pak

$$[M^\perp]_B = \operatorname{Ker} A \ .$$

(55) Je-li $A\mathbf{x} = \mathbf{b}$ soustava lineárních rovnic s reálnými (nebo komplexními) koeficienty. Každý vektor $\hat{\mathbf{x}} \in \mathbb{R}^n$ (nebo $\hat{\mathbf{x}} \in \mathbb{C}^n$) takový, že $A\hat{\mathbf{x}}$ se rovná ortogonální projekci vektoru pravých stran $\mathbf{b}$ na sloupcový prostor $\operatorname{Im} A$ matice A se nazývá *přibližné řešení* (nebo *aproximace řešení*) *soustavy* $A\mathbf{x} = \mathbf{b}$ *metodou nejmenších čtverců.*

(56) Je-li A je matice typu $m \times n$ nad $\mathbb{R}$ nebo $\mathbb{C}$ a $\mathbf{b} \in \mathbb{R}^m$ (resp. $\mathbb{C}^m$), pak množina všech přibližných řešení soustavy $A\mathbf{x} = \mathbf{b}$ metodou nejmenších čtverců je rovna množině všech (přesných) řešení soustavy

$$A^* A\hat{\mathbf{x}} = A^*\mathbf{b} \ .$$

(57) Soustavu $A^*A\mathbf{x} = A^*\mathbf{b}$ nazýváme *soustava normálních rovnic* příslušná k soustavě $A\mathbf{x} = \mathbf{b}$.

(58) Lineární regrese je přibližné řešení soustavy lineárních rovnic

$$a_1 x_i = b, \quad i = 1, 2, \dots, n$$

s neznámými a, b metodou nejmenších čtverců. Používá se k prokládání přímky množinou bodů (x_i, y_i), $i = 1, 2, \ldots, n$, získanou obvykle nějakým měřením.

Klíčové znalosti z osmé kapitoly nezbytné pro průběžné sledování přednášek s pochopením

(1) Definice standardního skalárního součinu v $\mathbb{R}^n$ a $\mathbb{C}^n$ a jeho geometrické významy v $\mathbb{R}^n$. Eukleidovská norma v $\mathbb{R}^n$ a $\mathbb{C}^n$.
(2) Rovnice přímky v $\mathbb{R}^2$ (obecně nadroviny v $\mathbb{R}^n$) pomocí skalárního součinu.
(3) Definice obecného skalárního součinu a normy určené skalárním součinem.
(4) Cauchyho-Schwartzova nerovnost a její důsledky (např. trojúhelníková nerovnost).
(5) Kolmost, ortogonální a ortonormální množiny a posloupnosti. Lineární nezávislost ortogonální posloupnosti nenulových vektorů.
(6) Souřadnice prvku vzhledem k ortonormální bázi. Skalární součin dvou prvků pomocí jejich souřanic vzhledem k ortonormální bázi.
(7) Ortogonální projekce prvku $\mathbf{v}$ na podprostor a fakt, že minimalizuje vzdálenost prvků podprostoru od $\mathbf{v}$.
(8) Výpočet ortogonální projekce prvku na podprostor, ve kterém je dána ortonormální báze.
(9) Gramova-Schmidtova ortogonalizace e její zápis pomocí QR-rozkladu matice.
(10) Ortogonální a unitární matice a jejich různé ekvivalentní definice.
(11) Ortogonální projekce prvku na podprostor zadaný libovolnou bází.
(12) Gramova matice posloupnosti prvků a její vlastnosti.
(13) Ortogonální doplněk.
(14) Ortogonální doplňky základních podprostorů určených maticí.
(15) Metoda nejmenších čtverců včetně lineární regrese.

9. Vlastní čísla a vlastní vektory

Cíl. *Vlastní čísla a vlastní vektory jsou základní nástroj pro zkoumání lineárních operátorů. Poznatky o vlastních číslech a vektorech použijeme ke studiu rozsáhlé třídy problémů shrnutých pod společný název lineární dynamické systémy. Ukážeme si řadu aplikací.*

V této kapitole pronikneme hlouběji do struktury matic a lineárních operátorů, hlavně na konečně generovaných prostorech. Vyvinutá teorie nám umožní mimo jiné počítat iterace daného operárou $f : \mathbf{V} \to \mathbf{V}$, tj. výrazy tvaru

$$f^n = \underbrace{f \circ f \circ \cdots \circ f}_{n\times} \quad .$$

Spočítat n-tou mocninu lineárního zobrazení $f_A : \mathbf{T}^n \to \mathbf{T}^n$ určeného čtvercovou maticí řádu n nad $\mathbf{T}$ znamená spočítat mocninu A^n. Nahlédnout to můžeme například tak, že si vzpomeneme na rovnost $A = [f_A]^K_K$ vyjadřující matici A jako matici zobrazení f_A vzhledem ke kanonickým bázím K a K. V řeči matic, naučíme se počítat n-tou mocninu čtvercové matice A.

9.1. **Lineární dynamické systémy.** Začneme několika motivujícími příklady.

9.1.1. *Úročení.* Jednoduchým příkladem lineárního dynamického systému je úročení vkladu na účtu. Na účet s úrokem 1%, který banka připisuje jednou za rok, vložíme počáteční vklad $x_0 = 1000$ Kč. Po roce budeme mít na účtě částku

$$x_1 = 1000 + 10 = (1 + 0{,}01)1000 = (1 + 0,01) \quad .$$

Po dvou letech to bude částka

$$x_2 = (1 + 0{,}01)x_1 = (1 + 0{,}01)^2 x_0$$

a po k letech budeme mít

$$x_k = (1 + 0{,}01)x_{k-1} = (1 + 0{,}01)^2 x_{k-2} = \cdots = (1 + 0{,}01)^k x_0 \quad .$$

Jiná banka nám nabídne účet, na kterém každé čtvrtletí připíše úrok 0,25%. U takové banky budeme mít při počátečním vkladu y_0 po jednom čtvrtletí částku

$$y_1 = 1000 + 2{,}50 = (1 + 0{,}0025)y_0 \ ,$$

po jednom roce to bude $y_4 = (1 + 0{,}0025)^4 y_0$ a po k čtvrtletích to bude $y_k = (1 + 0{,}0025)^k y_0$.

Je dobré si všimnout, že u druhé banky budeme mít po roce částku y_4, která je nepatrně větší než částka x_1 u první banky. Rozdíl je nepatrný kvůli malému úroku a malému počátečnímu vkladu. Pokud bychom si ale u banky půjčili milion korun na roční úrok 5%, rozdíl by už byl znatelný.

Vrátíme se k původnímu příkladu s úrokem 1% a vybereme si banku, která každý den připisuje úrok (1/365)%. Při počátečním vkladu z_0 budeme mít po roce částku

$$z_{365} = \left(1 + \frac{0{,}01}{365} z_0\right)^{365} \quad .$$

V rámci konkurenčního boje se banky začnou předhánět v tom, kolikrát za rok úrok připisují. U banky, která připisuje úrok $(1/n)\%$ n-krát ročně, bude po roce na účtu částka

$$z_n = \left(1 + \frac{0{,}01}{n}\right)^n z_0 \ .$$

Přechodem k limitě pro $n \to \infty$ nakonec konkurenční boj vyhraje banka **FURT** s reklamním sloganem *„V bance FURT úročíme furt“*, a poté zkrachuje. V jakémkoliv čase t, přičemž jednotkou času je 1 rok, u ní bude při počátečním vkladu z_0 na účtě částka

$$z(t) = \lim_{n\to\infty} \left(1 + \frac{0{,}01t}{n}\right)^n z(0) = e^{0{,}01t} z_0 \ .$$

Po jednom roce tedy na účtě bude $e^{0{,}01} z_0$ korun.

9.1.2. *Lineární rekurentní posloupnosti.* Fibonacciho posloupnost je příkladem *rekurentní posloupnosti* $(a_k)_{k=0}^{\infty}$ definované *rekurentním vztahem*

$$a_{k+2} = a_{k+1} + a_k \quad \text{pro každé } k = 0, 1, 2, \dots \ .$$

Celou posloupnost jednoznačně určují první dva prvky a_0 a a_1. Fibonacciho posloupnost dostaneme volbou $a_0 = 0$ a $a_1 = 1$.

V části 4.3.2 jsme nahlédli, že platí

$$\begin{pmatrix} a_{k+1} \\ a_{k+2} \end{pmatrix} = C \begin{pmatrix} a_k \\ a_{k+1} \end{pmatrix}, \text{ kde } C = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \ .$$

V terminologii lineárních dynamických systémů nazýváme vektor $\mathbf{x}_{k+1} = (a_{k+1}, a_{k+2})^T$ *stav systému*, tomto případě posloupnosti, v čase $k+1$. *Vývoj systému* je pak určen vztahem

$$\mathbf{x}_{k+1} = C\,\mathbf{x}_k \quad \text{pro každé } k \in \mathbb{N}$$

a *počáteční podmínkou* $\mathbf{x}_0 = (a_0, a_1)^T$.

Stav $\mathbf{x}_k$ v čase k se potom rovná

$$\mathbf{x}_k = \begin{pmatrix} a_k \\ a_{k+1} \end{pmatrix} = C^k \begin{pmatrix} a_0 \\ a_1 \end{pmatrix} = C^k\,\mathbf{x}_0 \ .$$

K určení k-tého členu posloupnosti nám tedy stačí umět vypočítat k-tou mocninu matice C pro libovolné $k \in \mathbb{N}$.

9.1.3. *Diskrétní lineární dynamické systémy.* Jak běžné připisování úroků za daný časový interval tak lineární rekurentní posloupnosti jsou příklady *diskrétních lineárních dynamických systémů.* Tento systém je zadán lineárním zobrazením $f : \mathbf{V} \to \mathbf{V}$ na lineárním prostoru $\mathbf{V}$ nad tělesem $\mathbf{T}$ a *počátečním stavem* $\mathbf{x}_0 \in \mathbf{V}$. Vývoj tohoto dynamického systému je dán předpisem

$$\mathbf{x}_{k+1} = f(\mathbf{x}_k) \quad \text{pro každé } k = 0, 1, 2, \dots \ .$$

To znamená, že např. $\mathbf{x}_2 = f(\mathbf{x}_1) = f(f(\mathbf{x}_0)) = f^2(\mathbf{x}_0)$. Jednoduchou indukcí podle k ověříme, že pro každé $k \geq 0$ platí

$$\mathbf{x}_k = f^k(\mathbf{x}_0) \ .$$

Je-li lineární prostor $\mathbf{V} = \mathbf{T}^n$ a A matice řádu n nad $\mathbf{T}$, pak lineární zobrazení $f_A : \mathbf{T}^n \to \mathbf{T}^n$ určené maticí A definuje diskrétní lineární systém předpisem

$$\mathbf{x}_{k+1} = f_A(\mathbf{x}_k) = A\,\mathbf{x}_k \quad \text{pro každé } k = 0, 1, 2, \dots \ .$$

Pro každé $k \geq 0$ pak platí

$$\mathbf{x}_k = f_A^k(\mathbf{x}_0) = A^k \mathbf{x}_0 \ .$$

Zkoumáme-li nějaký diskrétní lineární dynamický systém, zajímá nás průběh posloupnosti stavů $\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \ldots$ v závislosti na počátečním stavu $\mathbf{x}_0$. Nejsnazší to je, pokud se nám podaří najít explicitní vzorec pro k-tý prvek posloupnosti. Ale i bez explicitního vyjádření stavu $\mathbf{x}_k$ si můžeme klást otázky, jaké je limitní chování posloupnosti $(\mathbf{x}_k)_{k=0}^{\infty}$ pro $k \to \infty$, konverguje-li k nějakému limitnímu stavu $\mathbf{x}_\infty$, pokud ano, jak rychle k němu konverguje, atd. Pro různé počáteční stavy může být limitní chování různé. Jedním počátečním stavem není třeba se zabývat. Pokud je $\mathbf{x}_0 = \mathbf{o}$, pak $\mathbf{x}_k = \mathbf{o}$ pro každé k (důsledek linearity operátoru f).

Vývoj diskrétního lineárního dynamického systému v dimenzi 1, tj. v případě, kdy má stavový prostor $\mathbf{V}$ dimenzi 1, je průhledný. V prostoru $\mathbf{V}$ zvolíme jakýkoliv nenulový prvek $\mathbf{u}$ a jednotlivé stavy $\mathbf{x}_k$ budeme vyjadřovat pomocí jejich souřadnic vzhledem k bázi $B = (\mathbf{u})$ ve $\mathbf{V}$, tj. pomocí koeficientu $x_k \in \mathbf{T}$ ve vyjádření $\mathbf{x}_k = x_k \mathbf{u}$.

Lineární operátor $f : \mathbf{V} \to \mathbf{V}$ je jednoznačně určený svojí hodnotou na bázi B, tj. hodnotou $f(\mathbf{u}) = a\,\mathbf{u}$. Takže

$$\mathbf{x}_k = f^k(\mathbf{x}_0) = f^k(x_0 \mathbf{u}) = x_0\, f^k(\mathbf{u}) = x_0\, a^k \mathbf{u} \quad \text{pro každé } k = 0, 1, 2, \ldots \ .$$

Pomocí souřadnic vzhledem k bázi $B = (\mathbf{u})$ vyjádříme $\mathbf{x}_k = x_k \mathbf{u}$ a dostaneme rovnost

$$x_k = a^k x_0 \quad \text{pro každé } k = 0, 1, 2, \ldots \ .$$

Jde tedy o geometrickou posloupnost.

Její limitní chování závisí na tom, v jakém tělese $\mathbf{T}$ počítáme. Rozebereme si jej v případě tělesa reálných čísel. V tom případě

- je-li $|a| < 1$, posloupnost $(x_k)_{k=0}^{\infty}$ konverguje k 0 pro jakýkoliv počáteční stav x_0.
- je-li $a = 1$, je posloupnost $(x_k)_{k=0}^{\infty}$ konstatntní rovná x_0.
- je-li $a = -1$, posloupnost $(x_k)_{k=0}^{\infty}$ osciluje mezi hodnotami $\pm x_0$.
- je-li $a > 1$, posloupnost konverguje k $\pm\infty$ v závislosti na znaménku x_0.
- je-li $a < -1$, posloupnost $(|x_k|)_{k=0}^{\infty}$ konverguje do $+\infty$ a znaménka čísel x_k se střídají.

Jako cvičení si můžete udělat podobný rozbor v případě tělesa $\mathbb{C}$. Nejzajímavější je případ $|a| = 1$. Posloupnost $(x_k)_{k=0}^{\infty}$ může být v takovém případě konstantní, periodicky nabývat konečně mnoha hodnot a nebo můžou být její prvky po dvou různé, všechny ale leží na kružnici o poloměru $|x_0|$.

9.1.4. *Spojité lineární dynamické systémy.* V takovém případě sledujeme vývoj stavu „spojitě“, nikoliv po diskrétních časových intervalech. Stav spojitého dynamického systému v čase $t \in \mathbb{R}$ zapíšeme jako prvek $\mathbf{x}(t) \in \mathbf{V}$, kde $\mathbf{V}$ je nějaký lineární prostor. V případě spojitých dynamických systémů budeme předpokládat, že $\mathbf{V}$ je aritmetický prostor nad $\mathbb{R}$ nebo nad $\mathbb{C}$. V případě, že $\mathbf{x}(t) \in \mathbb{R}^n$, můžeme stav systému v čase t zapsat jako

$$\mathbf{x}(t) = (x_1(t), x_2(t), \ldots, x_n(t))^T \ ,$$

tj. jako uspořádanou n-tici reálných funkcí jedné reálné proměnné. *Derivací* stavového vektoru $\mathbf{x}(t)$ budeme nazývat vektor

$$\mathbf{x}'(t) = (x_1'(t), x_2'(t), \ldots, x'_n(t))^T \ .$$

Vývoj *spojitého lineární dynamického systému* nad $\mathbb{R}$ je definován rovností

$$\mathbf{x}'(t) = f(\mathbf{x}(t)) \quad \text{pro každé } t \in \mathbb{R} \ ,$$

kde $f : \mathbb{R}^n \to \mathbb{R}^n$ je lineární zobrazení, a *počátečním stavem* $\mathbf{x}(0) \in \mathbb{R}^n$. Protože každé lineární zobrazení $f : \mathbb{R}^n \to \mathbb{R}^n$ je tvaru f_A, kde $A = [f]^K_K$ je matice f vzhledem ke kanonickým bázím K a K, můžeme vývoj spojitého lineárního dynamického systému nad $\mathbb{R}$ zapsat také rovnicí

$$\mathbf{x}'(t) = A\,\mathbf{x}(t) \quad \text{pro každé } t \in \mathbb{R} \ ,$$

kde A je reálná matice řádu n.

Je-li $\mathbf{V} = \mathbb{C}^n$, můžeme každý stavový vektor $\mathbf{x}(t)$ zapsat jako

$$\mathbf{x}(t) = (x_1(t) + i\,y_1(t), x_2(t) + i\,y_2(t), \ldots, x_n(t) + i\,y_n(t))^T,$$

kde i je imaginární jednotka a funkce $x_j(t), y_j(t)$ jsou reálné funkce jedné reálné proměnné. Derivací stavového vektoru $\mathbf{x}(t)$ pak rozumíme vektor

$$\mathbf{x}'(t) = (x_1'(t) + i\,y_1'(t), x_2'(t) + i\,y_2'(t), \ldots, x_n'(t) + i\,y_n'(t))^T \ .$$

Vývoj *spojitého lineární dynamického systému* nad $\mathbb{C}$ je definován rovností

$$\mathbf{x}'(t) = f(\mathbf{x}(t)) \quad \text{pro každé } t \in \mathbb{R} \ ,$$

kde $f : \mathbb{C}^n \to \mathbb{C}^n$ je lineární zobrazení a $\mathbf{x}(0) \in \mathbb{C}^n$ je počáteční stav. Také v případě komplexních skalárů můžeme pomocí komplexní matice $A = [f]^K_K$ systém zapsat jako

$$\mathbf{x}'(t) = A\,\mathbf{x}(t) \quad \text{pro každé } t \in \mathbb{R} \ .$$

Poznamenejme ještě, že v mnoha učebnicích bývá derivace stavového vektoru $\mathbf{x}(t)$ označována jako $\dot{\mathbf{x}}(t)$.

9.1.5. *Rozpad atomů radioaktivní látky.* Míra radioaktivity jaderného materiálu se měří pomocí *rozpadové konstanty* $k \in (0,1)$, která udává pravděpodobnost, s jakou se jádro rozpadne během jedné vteřiny. Čím větší rozpadová konstanta, tím vyšší pravděpodobnost a tím vyšší radioaktivita materiálu.

Označíme $x(t)$ počet radioaktivních jader v čase t. Po krátkém časovém intevralu $\delta > 0$ bude toto množství přibližně

$$x(t+\delta) \approx x(t) - k\delta x(t) \ ,$$

rovnost je pouze přibližná, protože množství rozpadlých radioaktivních jader se snižuje i během intervalu $(t, t+\delta)$, zatímco v naší přibližné rovnosti je považujeme za konstantní a rovné $k\delta x(t)$ po celý interval. Odhad je tím přesnější, čím kratší je interval, tj. čím menší je δ. Přibližnou rovnost si přepíšeme do tvaru

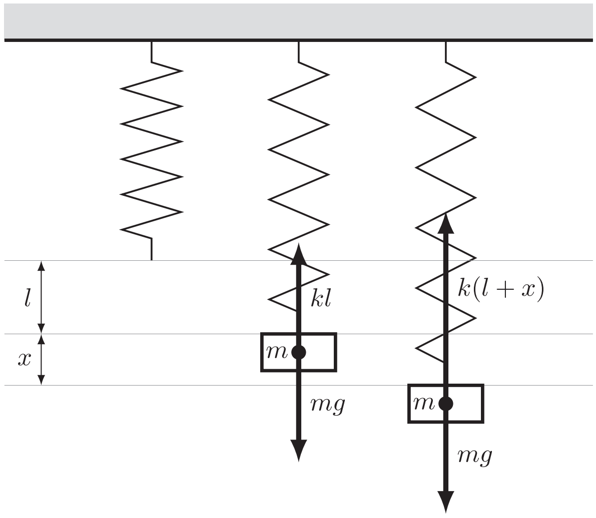
$$\frac{x(t+\delta) - x(t)}{\delta} \approx -k\,x(t) \ .$$

Přechodem k limitě pro $\delta \to 0$ dostaneme rovnost

$$x'(t) = -k\,x(t) \quad \text{pro každé } t \in \mathbb{R} \ .$$

Vývoj počtu radioaktivních jader $x(t)$ v čase je tedy popsán rovnicí obsahující derivaci neznámé funkce $x(t)$, a počátečním stavem (podmínkou) $x(0)$. Všimněme si, že počet radioaktivních jader se v každém časovém okamžiku mění rychlostí přímo úměrnou jejich počtu $x(t)$ v čase t. Koeficient přímé úměrnosti je v tomto případě $-k < 0$.

9.1.6. *Vlastní kmity pružiny.* Na pružinu s koeficientem pružnosti k zavěsíme závaží o hmotnosti m. Pružina se protáhne o délku l. Jak l spočítáme?



V rovnovážném stavu se vyrovnává gravitační síla mg, která táhne závaží směrem dolů, se silou pružiny, která táhne závaží směrem nahoru. Tato síla je podle Hookeova zákona přímo úměrná prodloužení pružiny a koeficient přímé úměrnosti je koeficient pružnosti k. Síla působící směrem vzhůru má proto velikost kl. V rovnovážném stavu pak platí rovnost

$$mg = kl \ ,$$

ze které plyne velikost prodloužení $l = mg/k$.

Když závaží vychýlíme z rovnovážného stavu směrem dolů nebo nahoru o $x_1(0) = b$ a pustíme je, začne se pohybovat. Jeho pohyb je popsán fyzikálními zákony. Stav závaží v čase t zapíšeme pomocí dvojice čísel $(x_1(t), x_2(t))^T$, kde $x_1(t)$ je odchylka od rovnovážného stavu v čase t a $x_2(t)$ je okamžitá rychlost závaží v čase t.

Směrem dolů na závaží působí konstantní gravitační síla mg, směrem vzhůru síla pružnosti $k(l + x_1(t))$. Celková síla působící na závaží v čase t je potom

$$F(t) = mg - k(l + x_1(t)) = (mg - kl) - kx_1(t) = -kx_1(t) \ ,$$

neboť pro prodloužení l v rovnovážném stavu platí rovnost $mg = kl$. Podle Newtonova zákona síla $F(t)$ uděluje závaží v čase t okamžité zrychlení $a(t)$, které vypočteme ze vztahu $F(t) = a(t)\, m$, neboli

$$a(t) = \frac{F(t)}{m} = -\frac{k}{m}\, x_1(t) \ .$$

K výpočtu derivace $\mathbf{x}'(t)$ stavového vektoru využijeme toho, že „derivace dráhy podle času je okamžitá rychlost“. Průměrná rychlost během krátkého časového intervalu $(t, t + \delta)$ se rovná

$$\frac{x_1(t + \delta) - x_1(t)}{\delta}$$

a okamžitou rychlost v čase t pak získáme jako limitu pro $\delta \to 0$, tj. $x_1'(t) = x_2(t)$. Podobně ze vzorce pro průměrné zrychlení během intervalu $(t, t+\delta)$

$$\frac{x_2(t+\delta) - x_2(t)}{\delta}$$

dostaneme v limitě pro $\delta \to 0$, že „derivace rychlosti podle času je okamžité zrychlení“, tj.

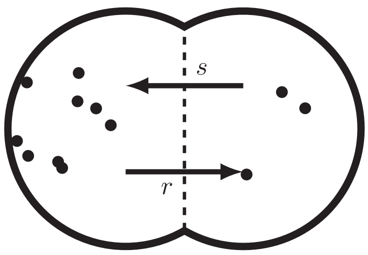
$$x_2'(t) = a(t) = -\frac{k}{m}\, x_1(t) \ .$$

Pohyb závaží na pružině je tedy spojitý lineární dynamický systém

$$\begin{aligned} \mathbf{x}'(t) &= \begin{pmatrix} x_1'(t) \\ x_2'(t) \end{pmatrix} = \begin{pmatrix} x_2(t) \\ -\frac{k}{m}\, x_1(t) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -\frac{k}{m} & 0 \end{pmatrix} \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} \\ &= \begin{pmatrix} 0 & 1 \\ -\frac{k}{m} & 0 \end{pmatrix} \mathbf{x}(t) \end{aligned}$$

s počáteční podmínkou $(x_1(0), x_2(0))^T = (b, 0)^T$.

9.1.7. *Přechod substance přes buněčnou blánu.* Přes buněčnou blánu mezi dvěma buňkami se šíří nějaká substance, např. vápník, alkohol, vitamín C, apod. Na počátku v čase $t = 0$ je do jedné buňky injektováno jednotkové množství substance. Víme, že rychlost šíření substance přes buněčnou blánu z jedné buňky do druhé je přímo úměrná množství substance v buňce, ze které se substance šíří, koeficient rychlosti šíření z buňky 1 do buňky 2 je $r > 0$, a z buňky 2 do buňky 1 je koeficient rovný $s > 0$. Máme určit množství substance v obou buňkách v čase t.



Označíme si $x_1(t)$, resp. $x_2(t)$, množství substance v buňce 1, resp. 2, v čase t. Rychlost změny množství substance v buňce 1 je $sx_2(t)$ (šíření z buňky 2) minus $rx_1(t)$ (šíření do buňky 2). Podobně pro druhou buňku. Dostáváme rovnice

$$\begin{aligned} x_1'(t) &= -rx_1(t) + sx_2(t) \ , \\ x_2'(t) &= rx_1(t) - sx_2(t) \ . \end{aligned}$$

Označíme-li $\mathbf{x}(t) = (x_1(t), x_2(t))^T$ a $\mathbf{x}'(t) = (x_1'(t), x_2'(t))^T$, můžeme proces šíření substance mezi buňkami popsat jako spojitý lineární dynamický systém

$$\mathbf{x}'(t) = \begin{pmatrix} -r & s \\ r & -s \end{pmatrix} \mathbf{x}(t)$$

s počáteční podmínkou že $\mathbf{x}(0) = (1, 0)^T$.

Poznamenejme ještě, že obvyklý název pro spojitý lineární dynamický systém

$$\mathbf{x}'(t) = A\mathbf{x}(t),\ \mathbf{x}(0) = \mathbf{b}$$

definovaný maticí A řádu n nad tělesem $\mathbb{R}$ (nebo $\mathbb{C}$) a počátečním stavem $\mathbf{b} \in \mathbb{R}^n$ (nebo $\mathbb{C}^n$), je *soustava lineárních diferenciálních rovnic s konstantními koeficienty* a s *počáteční podmínkou* $\mathbf{x}(0) = \mathbf{b}$.

9.1.8. *Vývoj spojitého lineárního dynamického systému v dimenzi* 1. V případě reálného spojitého lineárního dynamického systému v dimenzi 1 hledáme reálnou funkci $f(t)$ jedné reálné proměnné t, která splňuje rovnici

$$f'(t) = \lambda f(t)$$

pro nějaké reálné číslo λ a splňující počáteční podmínku $f(0) = s$.

Jednu takovou funkci vidíme hned, a to

$$f(t) = s\,e^{\lambda t},\quad t \in \mathbb{R}\ .$$

Je-li $g : \mathbb{R} \to \mathbb{R}$ jakákoliv funkce splňující $g'(t) = \lambda\, g(t)$ a $g(0) = s$, spočteme

$$(g(t)e^{-\lambda t})' = g'(t)e^{-\lambda t} - g(t)\lambda e^{-\lambda t} = g(t)(e^{-\lambda t} - e^{-\lambda t}) = 0\ ,$$

což znamená, že funkce $g(t)e^{-\lambda t}$ je konstantní na $\mathbb{R}$. Její hodnotu získáme volbou $t = 0$, tj.

$$g(0)e^{-\lambda 0} = g(0) = s\ .$$

Platí tedy $g(t)e^{-\lambda t} = s$, neboli $g(t) = s\,e^{\lambda t}$. Dokázali jsme tak následující tvrzení.

Tvrzení 9.1. *Je-li λ reálné číslo, pak pro funkci $f : \mathbb{R} \to \mathbb{R}$ splňující podmínky $f'(t) = \lambda f(t)$ pro každé $t \in \mathbb{R}$ a $f(0) = s$ platí*

$$f(t) = s\,e^{\lambda t}\ .$$

Příklad 9.2. Využijeme právě nalezeného průběhu reálného spojitého lineárního dynamického systému v dimenzi 1 k porovnání rozpadové konstanty $k > 0$ radioaktivní látky s jinou běžně používanou mírou radioaktivity, a to *poločasem rozpadu* T. Ten je definovaný jako doba, za kterou se množství radioaktivních jader sníží na polovinu. Vývoj počtu radioktivních jader $f(t)$ v čase t je dán rovnicí $f'(t) = -k\,f(t)$ a počátečním stavem $f(0) = s$. Podle předchozího tvrzení tedy platí $f(t) = s\,e^{-kt}$ pro každé t. Pro poločas rozpadu T potom platí $f(T) = s/2$, neboli

$$s\,e^{-kT} = s/2\ ,$$

což po zkrácení s a přirozeném logaritmování vede na rovnost

$$k\,T = \ln 2\ .$$

Reálný spojitý lineární dynamický systém $f(t) = \lambda f(t)$ s počáteční podmínkou $f(0) = s \neq 0$ se může vyvíjet v čase třemi různými způsoby v závislosti na hodnotě λ:

- je-li $\lambda < 0$, platí $\lim_{t\to\infty} f(t) = 0$,
- je-li $\lambda = 0$, platí $f(t) = s$ pro každé $t \in \mathbb{R}$,
- je-li $\lambda > 0$, platí $\lim_{t\to\infty} = \pm\infty$ v závislosti na znaménku s.

Na závěr úvodní motivační části si ještě ukážeme řešení komplexního spojitého dynamického systému v dimenzi 1. Při něm hledáme pro komplexní číslo $\lambda = \mu + i\nu$ komplexní funkci $z(t) = f(t) + ig(t)$ reálné proměnné splňující podmínku

$$z'(t) = \lambda z(t) = (\mu + i\nu)z(t) \quad \text{pro každé} \quad t \in \mathbb{R}$$

spolu s počáteční podmínkou $z(0) = f(0) + ig(0) = p + iq$. Na základě analogie s reálným případem můžeme zkusit řešení

$$z(t) = (p + iq)e^{\lambda t} \ .$$

Abychom ověřili, že takto definnovaná funkce $z(t)$ skutečně splňuje rovnici $z'(t) = \lambda f(t)$, připomeneme si Eulerovu formuli z části 1.2.11:

$$e^{ix} = \cos x + i \sin x \quad \text{pro každé} \quad x \in \mathbb{R} \ .$$

Spočteme derivaci

$$\begin{aligned}(e^{ix})' &= (\cos x + i \sin x)' = (\cos x)' + i(\sin x)' = -\sin x + i \cos x = i(\cos x + i \sin x) \\ &= i\, e^{ix}\end{aligned}$$

a použijeme tuto rovnost při výpočtu

$$\begin{aligned} z'(t) = &\left((p + iq)e^{(\mu+i\nu)t}\right)' = (p + iq)\left(e^{\mu t}e^{i\nu t}\right)' \\ =&(p + iq)\left(\mu e^{\mu t}e^{i\nu t} + i\nu\mu e^{\mu t}e^{i\nu t}\right) = (p + iq)(\mu + i\nu)e^{(\mu+i\nu)t} \\ =&(\mu + i\nu)z(t) = \lambda\, e^{\lambda t} \ . \end{aligned}$$

A protože rovněž $z(0) = p+iq$, popisuje funkce $z(t) = (p+iq)e^{\lambda t}$ vývoj komplexního spojitého dynamického systému $z'(t) = \lambda z(t)$ s počáteční podmínkou $z(0) = p+iq$.

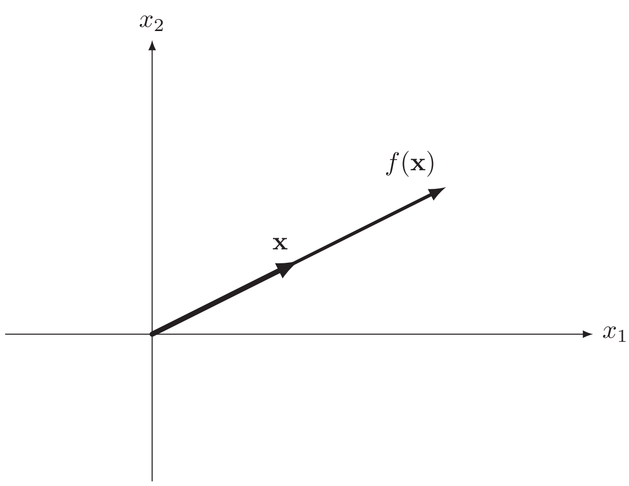
Vývoj komplexního dynamického systému $z'(t) = \lambda\, z(t)$ pro $\lambda = \mu + i\nu$ a počáteční podmínku $z(0) = p + iq$ je následující. Protože

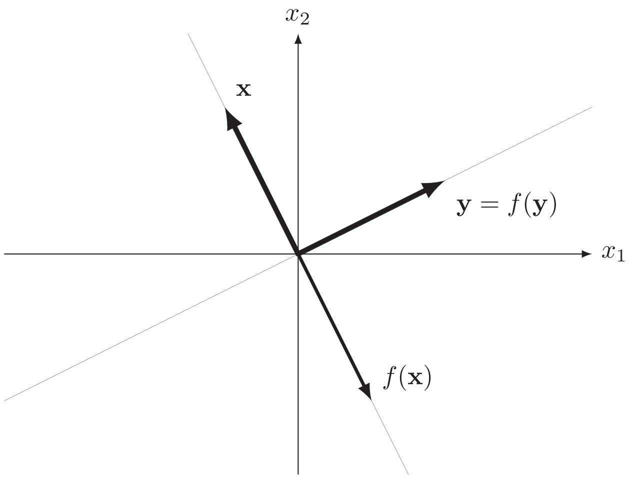
$$\begin{aligned} |z(t)| =&|(p + iq)\lambda e^{\lambda t}| = |p + iq|\,|e^{\mu t+i\nu t}| = |p + iq|\,|e^{\mu t}|\,|e^{i\nu t}| \\ =&|p + iq|\,|e^{\mu t}|\,|\cos(\nu t) + i \sin(\nu t)| = |p + iq|\,|e^{\mu t}| \ , \end{aligned}$$

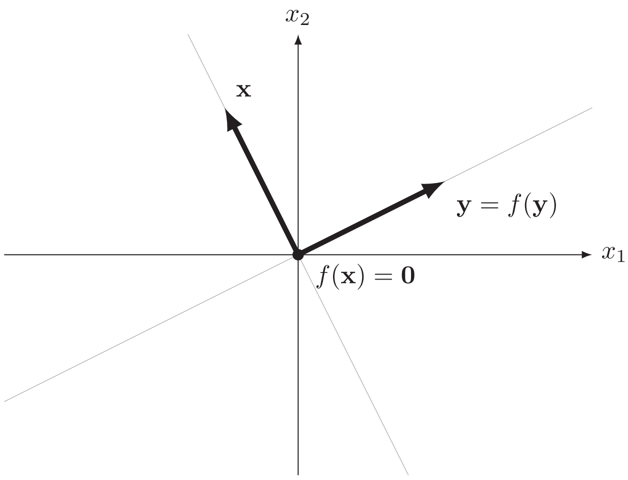
mohou nastat následující možnosti

- $|z(t)| \to \infty$ v případě, že $\mu = \operatorname{Re} \lambda > 0$,
- $|z(t)| = \sqrt{p^2 + q^2}$ pro každé $t \in \mathbb{R}$, pokud $\mu = \operatorname{Re} \lambda = 0$,
- $|z(t)| \to 0$, pokud $\mu = \operatorname{Re} \lambda < 0$.

V případě, že číslo $\nu = \operatorname{Im} \lambda \neq 0$, udává frekvenci „kroužení“ bodu $z(t)$ kolem počátku souřadnic, přičemž jeden „oběh“ trvá $2\pi\nu^{-1}$.







9.2. **Vlastní čísla a vlastní vektory.** V dimenzi 1 umíme předpovědět vývoj jak diskrétních tak spojitých lineárních dynamických systémů pro jakýkoliv počáteční stav. Podobným způsobem umíme předpovědět vývoj i v dimenzi větší než 1 aspoň pro některé počáteční stavy.

Příklad 9.3. Uvažujme lineární operátor f_A na $\mathbb{R}^2$ určený maticí

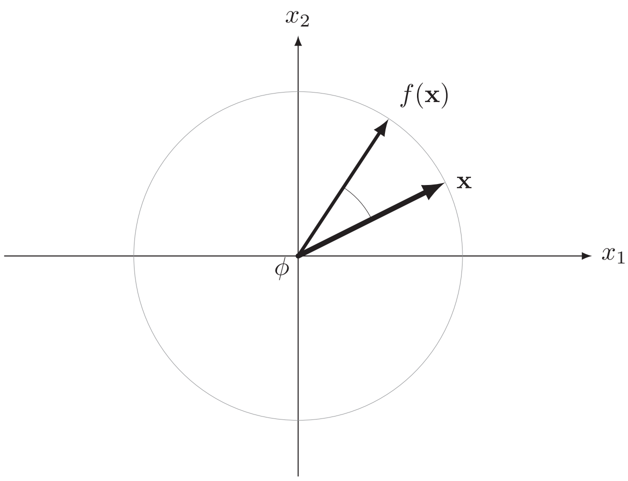
$$A = \begin{pmatrix} 3 & 0 \\ 1 & 2 \end{pmatrix}$$

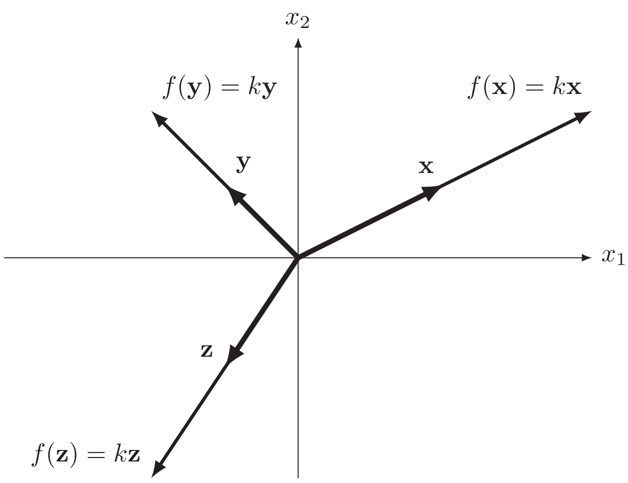
a diskrétní dynamický systém

$$\mathbf{x}_{k+1} = f_A(\mathbf{x}_k) = A\,\mathbf{x}_k \ .$$

Zvolíme-li počáteční stav $\mathbf{x}_0 = (1,1)^T$, dostáváme

$$\mathbf{x}_1 = f_A \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 & 0 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = 3 \begin{pmatrix} 1 \\ 1 \end{pmatrix} = 3\mathbf{x}_0 \ .$$





Dále

$$\mathbf{x}_2 = f_A(\mathbf{x}_1) = f_A(3\mathbf{x}_0) = 3\, f_A(\mathbf{x}_0) = 3^2\mathbf{x}_0$$

a podobně

$$\mathbf{x}_k = f_A^k(\mathbf{x}_0) = 3^k\mathbf{x}_0 \ .$$

Formálně bychom poslední vztah dokázali indukcí podle k.

Řešení příkladu pro počáteční vektor $\mathbf{x}_0 = (1,1)^T$ nám umožnila skutečnost, že $f(\mathbf{x}_0)$ je skalárním násobkem vektoru $\mathbf{x}_0$. To vede k velmi důležité definici vlastních čísel a vektorů.

Definice 9.4. Je-li $f : \mathbf{V} \to \mathbf{V}$ lineární operátor na lineárním prostoru $\mathbf{V}$ nad tělesem $\mathbf{T}$, pak skalár $\lambda \in T$ nazýváme *vlastní číslo* operátoru f, pokud existuje nenulový vektor $\mathbf{x} \in V$, pro který platí

$$f(\mathbf{x}) = \lambda\mathbf{x} \ .$$

Je-li λ vlastní číslo operátoru f, pak libovolný prvek $\mathbf{x} \in V$, pro který platí $f(\mathbf{x}) = \lambda\mathbf{x}$, nazýváme *vlastní vektor* operátoru f *příslušný vlastnímu číslu* λ.

Vlastní čísla a vektory pro čtvercovou matici řádu n nad $\mathbf{T}$ definujeme jako vlastní čísla a vektory příslušného operátoru $f_A : \mathbf{T}^n \to \mathbf{T}^n$. Podobně tomu bude i pro další pojmy v této kapitole. Přeložíme si poslední definici pro případ matic.

Definice 9.5. Je-li A čtvercová matice řádu n nad tělesem $\mathbf{T}$, pak skalár $\lambda \in \mathbf{T}$ nazýváme *vlastní číslo* matice A, pokud existuje nenulový vektor $\mathbf{x} \in T^n$ takový, že

$$A\mathbf{x} = \lambda\mathbf{x} \ .$$

Je-li λ vlastní číslo matice A, pak libovolný vektor $\mathbf{x} \in T^n$, pro který platí $A\mathbf{x} = \lambda\mathbf{x}$, nazýváme *vlastní vektor* matice A *příslušný vlastnímu číslu* λ.

Je důležité uvědomit si geometrický význam definice vlastního čísla operátoru. Číslo $\lambda \in \mathbf{T}$ je vlastní číslo operátoru f, pokud existuje nenulový prvek $\mathbf{x} \in V$, který operátor f zobrazí na λ-násobek $\lambda\mathbf{x}$ prvku $\mathbf{x}$, tj. do směru vektoru $\mathbf{x}$. V případě prostoru nad reálnými čísly tak operátor f vektor $\mathbf{x}$ buď „natahuje“ (pokud $\lambda > 1$) nebo „zkracuje“ (pokud $0 < \lambda < 1$), případně navíc „obrací“ (pokud $\lambda < 0$).

OBRAZEK - vlastni vektor

Příklad 9.6. Na obrázku 75 je nakresleno zobrazení $f_A : \mathbb{R}^2 \to \mathbb{R}^2$ určené maticí

$$A = \begin{pmatrix} 1{,}035 & 0{,}09 \\ 0{,}135 & 0{,}99 \end{pmatrix}$$

tak, že pro některé body $\mathbf{x}$ je zobrazena šipka z $\mathbf{x}$ do $f_A(\mathbf{x}) = A\,\mathbf{x}$.

Zobrazení f_A (stejně jako matice A) má dvě vlastní čísla 1,125 a 0,9. Vlastní vektory příslušné 1,125 jsou vektory z $\left\langle (1,1)^T \right\rangle$, což na obrázku vidíme tak, že tyto vektory zobrazení f_A natáhne na 1,125-násobek. Vlastní vektory příslušné 0,9 jsou vektory z $\left\langle (-2,3)^T \right\rangle$, tyto vektory zobrazení zkrátí na f_A 0,9-násobek.

Na obrázku je také pěkně kvalitativně vidět chování posloupnosti $(f^k(\mathbf{x}_0))_{k=1}^{\infty}$ pro různé počáteční vektory. Výsledek v příští části odůvodníme algebraicky.

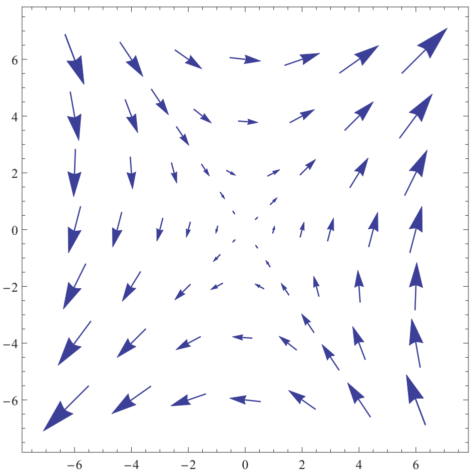
Pro každé číslo $\lambda \in \mathbf{T}$ platí, že $f(\mathbf{o}) = \mathbf{o} = \lambda\mathbf{o}$. To ale neznamená, že λ je vlastní číslo f. K tomu, aby λ bylo vlastní číslo f, je nutná existence **nenulového** prvku $\mathbf{x}$, pro který platí $f(\mathbf{x}) = \lambda\mathbf{x}$. V takovém případě pak i nulový vektor je vlastním vektorem příslušným λ.

Číslo 0 může být vlastním číslem operátoru f, k tomu je ale nutná (a stačí) existence vektoru $\mathbf{x} \neq \mathbf{o}$, pro který platí $f(\mathbf{x}) = 0\mathbf{x} = \mathbf{o}$, což nastává právě když $\mathrm{Ker}\,(f) \neq \{\mathbf{o}\}$, neboli právě když operátor f není prostý (viz tvrzení 6.24). Zformulujeme učiněné pozorování.

Pozorování 9.7. *Operátor* $f : \mathbf{V} \to \mathbf{V}$ *má vlastní číslo* 0 *právě tehdy, když* f *není prostý.*

Pro čtvercovou matici A je operátor f_A prostý právě tehdy, když je A regulární, takže maticová verze předchozího pozorování dává další kriterium regularity.

Pozorování 9.8. *Čtvercová matice* A *má vlastní číslo* 0 *právě tehdy, když* A *je singulární.*



OBRÁZEK 75. Obrázek zobrazení f_A. Šipka vede z bodu $\mathbf{x}$ do bodu $f_A(\mathbf{x})$.

Příklad 9.9. Identické zobrazení $f : \mathbf{V} \to \mathbf{V}$ má jediné vlastní číslo 1. Každý vektor z $\mathbf{V}$ je vlastním vektorem příslušným vlastnímu číslu 1. Speciálně, identická matice řádu n nad $\mathbf{T}$ má jediné vlastní číslo 1 a každý vektor z $\mathbf{T}^n$ je vlastním vektorem příslušným vlastnímu číslu 1.

Nulové zobrazení $0 : \mathbf{V} \to \mathbf{V}$ má jediné vlastní číslo 0. Každý vektor z $\mathbf{V}$ je vlastním vektorem příslušným vlastnímu číslu 0. Speciálně, nulová matice řádu n nad $\mathbf{T}$ má jediné vlastní číslo 0 a každý vektor z $\mathbf{T}^n$ je vlastním vektorem příslušným vlastnímu číslu 0.

Příklad 9.10. V příkladu 9.3 jsem využili toho, že matice

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

(a lineární operátor f_A) má vlastní číslo 3 a každý vektor z $\left\langle (1,1)^T \right\rangle$ je vlastním vektorem matice A příslušným vlastnímu číslu 3.

Příklad 9.11. Osová symetrie $f : \mathbb{R}^2 \to \mathbb{R}^2$ určená přímkou generovanou nenulovým vektorem $(a,b)^T$ má jedno vlastní číslo 1, neboť všechny vektory na ose symetrie se zobrazí samy do sebe a jsou to tedy vlastní vektory příslušné vlastnímu číslu 1. Vektory na přímce kolmé na osu symetrie (generované např. vektorem $(-b,a)^T$) se zobrazují do vektorů opačných, jsou to tedy vlastní vektory příslušné vlastnímu číslu -1.

OBRAZEK - vlastni cisla osove symetrie

Příklad 9.12. Ortogonální projekce $g : \mathbb{R}^2 \to \mathbb{R}^2$ na přímku generovanou $(a,b)^T$ má také dvě vlastní čísla. Jedno je opět 1, protože vektory přímky, na kterou projektujeme, se zobrazují na sebe. Druhé vlastní číslo je 0, protože všechny vektory z přímky kolmé na přímku projekce se zobrazují do nulového vektoru.

OBRAZEK - vlastni cisla ortogonalni projekce na primku

Příklad 9.13. Rotace kolem počátku souřadnic o úhel φ nemá žádné reálné vlastní číslo, pokud φ není násobkem π, neboť v takovém případě se žádný nenulový vektor nezobrazí na svůj násobek.

OBRAZEK - vlastni cisla rotace

Příklad 9.14. Stejnolehlost v rovině $\mathbb{R}^2$ s koeficientem k, která zobrazuje každý vektor $\mathbf{x}$ do jeho k-násobku $k\mathbf{x}$, má jediné vlastní číslo k, každý vektor $\mathbb{R}^2$ je vlastním vektorem příslušným vlastnímu číslu k. Mezi stejnolehlosti řadíme i mezní případ $k = 0$ (konstantní zobrazení do nulového vektoru), $k = 1$, což je identické zobrazení, a také rotace o úhel 0, a $k = -1$ neboli středová symetrie (a také rotace o úhel π.

OBRAZEK - vlastni cisla stejnolehlosti

V definici vlastních čísel a vektorů nepředpokládáme, že prostor $\mathbf{V}$ má konečnou dimenzi. Důležitým „nekonečnědimenzionálním" příkladem je operátor derivace.

Příklad 9.15. Označíme D lineární operátor definovaný předpisem $D(f) = f'$ na prostoru $\mathbf{V}$ všech reálných funkcí reálné proměnné, které mají spojité derivace všech řádů.

Reálné číslo λ je vlastním číslem operátoru D právě když existuje nenulová funkce $f(t) \in \mathbf{V}$, pro kterou platí $D(f) = \lambda\, f$, neboli

$$f'(t) = \lambda f(t) \quad \text{pro každé} \quad t \in \mathbb{R}\ .$$

V části 9.1.8 jsme ukázali, že pro každé $\lambda \in \mathbb{R}$ jsou všechny funkce splňující poslední rovnost tvaru $f(t) = s\, e^{\lambda t}$, kde s je libovolné reálné číslo – tvrzení 9.1.

9.2.1. *Výpočet vlastních čísel a vlastních vektorů.* Na příkladech jste si mohli všimnout, že množina vlastních vektorů příslušných vlastnímu číslu λ vždy tvořila podprostor. To není náhoda.

K důkazu použijeme obrat, který v této kapitole budeme často používat. Uvažujeme lineární operátor $f : \mathbf{V} \to \mathbf{V}$, vektor $\mathbf{x} \in V$ a skalár λ. Vztah $f(\mathbf{x}) = \lambda\mathbf{x}$ lze ekvivalentně upravit

$$\begin{aligned} f(\mathbf{x}) &= \lambda\mathbf{x} \\ f(\mathbf{x}) &= (\lambda\, \mathrm{id}_V)(\mathbf{x}) \\ f(\mathbf{x}) - (\lambda\, \mathrm{id}_V)(\mathbf{x}) &= \mathbf{o} \\ (f - \lambda\, \mathrm{id}_V)(\mathbf{x}) &= \mathbf{o} \\ \mathbf{x} &\in \mathrm{Ker}\,(f - \lambda\, \mathrm{id}_V) \end{aligned}$$

Odtud plyne následující tvrzení.

Tvrzení 9.16. *Nechť f je lineární operátor na prostoru $\mathbf{V}$ nad tělesem $\mathbf{T}$. Pak $\lambda \in T$ je vlastním číslem operátoru f právě tehdy, když operátor $(f - \lambda\, \mathrm{id}_V)$ není prostý.*

Je-li λ vlastním číslem operátoru f, pak množina M_λ všech vlastních vektorů operátoru f příslušných vlastnímu číslu λ je podprostorem $\mathbf{V}$ a platí

$$M_\lambda = \mathrm{Ker}\,(f - \lambda\, \mathrm{id}_V)\ .$$

Důkaz. Podle předchozích úprav, $f(\mathbf{x}) = \lambda\mathbf{x}$ platí právě tehdy, když $\mathbf{x} \in \mathrm{Ker}\,(f - \lambda\,\mathrm{id}_V)$, takže nenulový vektor $\mathbf{x}$ splňující $f(\mathbf{x}) = \lambda\mathbf{x}$ existuje právě tehdy, když je $\mathrm{Ker}\,(f - \lambda\,\mathrm{id}_V)$ netriviální, tedy (viz tvrzení 6.24) právě tehdy, když f není prostý.

Druhá část tvrzení pak plyne ze stejného výpočtu a z toho, že jádro je vždy podprostorem (viz 6.25). □

Explicitně zformulujeme také maticovou verzi. Odvození má v tomto případě podobu

$$\begin{aligned} A\mathbf{x} &= \lambda\mathbf{x} \\ A\mathbf{x} &= \lambda I_n\mathbf{x} \\ A\mathbf{x} - \lambda I_n\mathbf{x} &= \mathbf{o} \\ (A - \lambda I_n)(\mathbf{x}) &= \mathbf{o} \\ \mathbf{x} &\in \mathrm{Ker}\,(A - \lambda I_n) \end{aligned}$$

Tvrzení 9.17. *Je-li A čtvercová matice řádu n nad tělesem $\mathbf{T}$, pak $\lambda \in T$ je vlastním číslem matice A právě tehdy, když je matice $A - \lambda I_n$ singulární.*

Je-li λ vlastním číslem matice f, pak množina M_λ všech vlastních vektorů matice A příslušných vlastnímu číslu λ je podprostorem $\mathbf{T}^n$ a platí

$$M_\lambda = \mathrm{Ker}\,(A - \lambda I_n) \ .$$

Důkaz. Plyne opět z výpočtu před formulací tvrzení spolu s tím, že $\mathrm{Ker}\,(A - \lambda I_n)$ je netriviální právě tehdy, když $A - \lambda I_n$ je singulární (viz větu 4.59 charakterizující regulární matice). □

K výpočtu vlastních čísel matice A si uvědomíme, že matice $A - \lambda I_n$ je singulární právě tehdy, když je její determinant nulový.

Pozorování 9.18. *Je-li A čtvercová matice řádu n nad tělesem $\mathbf{T}$, pak $\lambda \in T$ je vlastním číslem matice A právě tehdy, když* $\det\,(A - \lambda I_n) = 0$.

Obecněji, k výpočtu vlastních čísel operátoru f na konečně generovaném prostoru $\mathbf{V}$ zvolíme bázi B prostoru $\mathbf{V}$. Pro každý skalár $\lambda \in T$ je vyjádření jádra operátoru $f - \lambda\,\mathrm{id}_V$ vzhledem k B podle tvrzení 6.25 a 6.37 rovno

$$[\mathrm{Ker}\,(f - \lambda\,\mathrm{id}_V)]_B = \mathrm{Ker}\,[f - \lambda\,\mathrm{id}_V]^B_B = \mathrm{Ker}\,([f]^B_B - \lambda[\mathrm{id}_V]^B_B) = \mathrm{Ker}\,([f]^B_B - \lambda I_n) \ .$$

Dostáváme obecnější verzi předchozího pozorování.

Pozorování 9.19. *Je-li f lineární operátor na konečně generovaném prostoru $\mathbf{V}$ dimenze n nad tělesem $\mathbf{T}$ a B je báze $\mathbf{V}$, pak $\lambda \in T$ je vlastním číslem operátoru f právě když je λ vlastní číslo matice $[f]^B_B$ vzhledem k bázím B a B, což nastává právě když* $\det\left([f]^B_B - \lambda I_n\right) = 0$.

Příklad 9.20. Určíme vlastní čísla a vlastní vektory matice A (= operátoru f_A) z příkladu 9.3.

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

Vypočteme pro obecný skalár $\lambda \in \mathbb{R}$ determinant

$$\det\,(A - \lambda I_2) = \det \begin{pmatrix} 3-\lambda & 0 \\ 1 & 2-\lambda \end{pmatrix} = (3-\lambda)(2-\lambda) - 0 \cdot 1 = (3-\lambda)(2-\lambda) \ .$$

Podle pozorování **??** má matice A dvě vlastní čísla 2 a 3. Podle tvrzení 9.17 tvoří vlastní vektory příslušné vlastnímu číslu 2 podprostor

$$M_2 = \operatorname{Ker}(A - 2I_2) = \operatorname{Ker}\begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} = \left\langle \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\rangle$$

a vlastní vektory příslušné vlastnímu číslu 3 tvoří podprostor

$$M_3 = \operatorname{Ker}(A - 3I_2) = \operatorname{Ker}\begin{pmatrix} 0 & 0 \\ 1 & -1 \end{pmatrix} = \left\langle \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\rangle \ .$$

Příklad 9.21. Známe-li lineárně nezávislou dvojici vlastních vektorů matice A řádu 2, můžeme snadno spočítat jakýkoliv prvek $\mathbf{x}_k$ diskrétního dynamického systému

$$\mathbf{x}_{k+1} = A\,\mathbf{x}_k$$

určeného maticí A pro každý počáteční stav $\mathbf{x}_0 = (a,b)^T$. Ukážeme si jak na matici

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

z předchozího příkladu.

Tato matice má dvě vlastní čísla $2, 3$, k vlastnímu číslu 2 je příslušný vlastní vektor např. $\mathbf{u}_2 = (0,1)^T$, k vlastnímu číslu 3 je vlastní vektor například $\mathbf{u}_3 = (1,1)^T$. Posloupnost B=$(\mathbf{u}_2, \mathbf{u}_3)$ je zjevně lineárně nezávislá a tedy báze v $\mathbb{R}^2$.

Protože $\mathbf{u}_2$ je vlastní vektor příslušný vlastnímu číslu 2, platí

$$A^k\mathbf{u}_2 = 2^k\mathbf{u}_2$$

a analogicky

$$A^k\mathbf{u}_3 = 3^k\mathbf{u}_3 \ .$$

Libovolný počáteční stav $\mathbf{x}_0 \in \mathbb{R}^2$ vyjádříme jako lineární kombinaci $\mathbf{x}_0 = a\mathbf{u}_2 + b\mathbf{u}_3$ vlastních vektorů matice A a spočítáme

$$\begin{aligned} \mathbf{x}_k =& A^k\mathbf{x}_0 = A^k(a\mathbf{u}_2 + b\mathbf{u}_3) = a\,A^k\mathbf{u}_2 + b\,A^k\mathbf{u}_3 = a\,2^k\mathbf{u}_2 + b\,3^k\mathbf{u}_3 \\ =& a\,2^k\begin{pmatrix} 0 \\ 1 \end{pmatrix} + b\,3^k\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} b3^k \\ a2^k + b3^k \end{pmatrix} \ . \end{aligned}$$

Koeficienty a, b najdeme jako řešení soustavy lineárních rovnic

$$\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = \mathbf{x}_0 \ .$$

Všimněme si, že matice této soustavy se rovná matici přechodu $[\mathrm{id}]^B_K$ od báze B složené z vlastních vektorů matice A ke kanonické bázi K. Koeficienty $(a,b)^T$ tedy najdeme jako

$$\begin{pmatrix} a \\ b \end{pmatrix} = \left([\mathrm{id}]^B_K\right)^{-1}\mathbf{x}_0 = [\mathrm{id}]^K_B\,\mathbf{x}_0 \ .$$

Příklad 9.22. Najdeme vlastní čísla a vektory ortogonální projekce v $\mathbb{R}^2$ na přímku určenou vektorem $(1,2)^T$. Označme tento operátor f.

Jedna možnost je najít matici A operátoru f vzhledem ke kanonickým bázím. Pak $f = f_A$ a vlastní čísla a vektory f se vypočítají jako v předchozím příkladu (jako vlastní čísla a vektory matice A). Ukážeme nejprve tento, méně efektivní, postup. V příkladu **??** jsme odvodili, že f má vzhledem ke kanonické bázi matici

$$A = \frac{\begin{pmatrix} 1 \\ 2 \end{pmatrix}(1,2)}{\|(1,2)^T\|^2} = \frac{1}{5}\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} .$$

Determinant matice $A - \lambda I_2$ je roven

$$\det(A - \lambda I_2) = \det\begin{pmatrix} 1/5 - \lambda & 2/5 \\ 2/5 & 4/5 - \lambda \end{pmatrix} = \lambda^2 - \lambda .$$

Matice A má tedy dvě vlastní čísla 1 a 0. Vlastní vektory příslušné vlastnímu číslu 1 tvoří podprostor

$$M_1 = \mathrm{Ker}\,(A - I_2) = \mathrm{Ker}\begin{pmatrix} -4/5 & 2/5 \\ 2/5 & -1/5 \end{pmatrix} = \left\langle \begin{pmatrix} 1 \\ 2 \end{pmatrix} \right\rangle$$

a vlastní vektory příslušné vlastnímu číslu 0 tvoří podprostor

$$M_0 = \mathrm{Ker}\,(A) = \left\langle \begin{pmatrix} -2 \\ 1 \end{pmatrix} \right\rangle .$$

Výsledek je v souladu z geometrickým náhledem z příkladu 9.12

Početně jednodušší postup je pracovat s maticí vzhledem f vzhledem k jiné bázi. Protože zřejmě $f((1,2)^T) = (1,2)^T$ a $f((-2,1)) = 0$ je matice f vzhledem k bázi $B = ((1,2)^T, (-2,1)^T)$ a B rovná

$$C = [f]^B_B = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

Determinant matice $C - \lambda I_2$ je $(1-\lambda)(-\lambda) = \lambda^2 - \lambda$ a vlastní čísla jsou 0 a 1 podle pozorování **??**. Podprostory M_1 a M_0 vypočítáme nejprve vzhledem k bázi B:

$$\begin{aligned} [M_1]_B =& [\mathrm{Ker}\,(f - \mathrm{id}_V)]_B = \mathrm{Ker}\,([f]^B_B - I_n) = \mathrm{Ker}\begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix} = \left\langle \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\rangle , \\ [M_0]_B =& \mathrm{Ker}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \left\langle \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\rangle . \end{aligned}$$

Převodem do kanonické báze dostaneme stejný výsledek jako prvním postupem:

$$\begin{aligned} M_1 =& [M_1]_K = \left\langle [\mathrm{id}]^B_K \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\rangle = \left\langle 1\begin{pmatrix} 1 \\ 2 \end{pmatrix} + 0\begin{pmatrix} -2 \\ 1 \end{pmatrix} \right\rangle = \left\langle \begin{pmatrix} 1 \\ 2 \end{pmatrix} \right\rangle , \\ M_0 =& [M_2]_K = \left\langle [\mathrm{id}]^B_K \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\rangle = \left\langle 0\begin{pmatrix} 1 \\ 2 \end{pmatrix} + 1\begin{pmatrix} -2 \\ 1 \end{pmatrix} \right\rangle = \left\langle \begin{pmatrix} -2 \\ 1 \end{pmatrix} \right\rangle . \end{aligned}$$

Příklad 9.23. Spočítáme vlastní čísla rotace v $\mathbb{R}^2$ o úhel $\pi/2$ v kladném směru. Matice této rotace vzhledem ke kanonické bázi se rovná

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

příslušný determinant je $\det(A - \lambda I_2) = \lambda^2 + 1$. Vidíme, že matice A nemá žádné reálné vlastní číslo a tedy ani žádný vlastní vektor v $\mathbb{R}^2$.

Považujeme-li matici A za matici nad komplexními čísly, má dvě vlastní čísla i a $-i$. Vlastní vektory příslušné vlastnímu číslu i jsou všechny komplexní násobky vektoru $(i,1)^T$ a vlastní vektory příslušné vlastnímu číslu $-i$ jsou všechny komplexní násobky vektoru $(i,-1)^T$.

9.2.2. *Charakteristický polynom, podobnost.* V předchozích příkladech jsme našli vlastní čísla matice A řádu 2 jako kořeny kvadratického polynomu $\det(A - \lambda I_2)$. Obecně víme podle pozorování 9.18 a 9.19, že $\lambda \in \mathbf{T}$ je vlastní číslo matice nebo operátoru právě když platí $\det(A - \lambda I_n)$ pro nějakou matici A. V následujícím tvrzení si ukážeme, že $\det(A - \lambda I_n)$ je polynom stupně n pro každou matici A řádu n a uvedeme nějaké jeho koeficienty.

Tvrzení 9.24. *Pro každou matici $A = (a_{ij})$ řádu n nad libovolným tělesem $\mathbf{T}$ platí*

(1) $\det(A - \lambda I_n)$ *je polynom stupně n s koeficienty v $\mathbf{T}$,*
(2) *koeficient u λ^n se rovná $(-1)^n$,*
(3) *koeficient u λ^{n-1} se rovná $(-1)^{n-1}(a_{11} + a_{22} + \cdots + a_{nn})$,*
(4) *absolutní člen se rovná* $\det A$.

Důkaz. První tři body ukážeme najednou. Označíme si prvky matice $A - \lambda I_n = (b_{ij})$. Z definice determinantu

$$\det(A - \lambda I_n) = \sum_{\pi \in S_n} \operatorname{sgn}(\pi) b_{1\pi(1)} b_{2\pi(2)} \cdots b_{n\pi(n)}$$

vidíme, že každý součin je součinem n prvků matice $A - \lambda I_n$, přičemž $b_{ij} = a_{ij}$ pokud $i \neq j$, a $b_{ii} = a_{ii} - \lambda$ pro každé $i = 1, 2, \ldots, n$. Součin odpovídající permutaci π obsahuje tolik činitelů $(a_{ii} - \lambda)$, kolik je prvků $i \in \{1, 2, \ldots, n\}$, pro které platí $\pi(i) = i$, tj. kolik je pevných bodů permutace π. Těch je nejvýše n, proto po roznásobení každého součinu dostaneme nějaký polynom stupně nejvýše n v proměnné λ. Jejich součtem je tedy také polynom stupně nejvýše n v proměnné λ.

Mocnina λ^n se může vyskytnout pouze v součinu, ve kterém je všech n činitelů rovných $a_{ii} - \lambda$, tj. v součinu definovaném identickou permutací na množině $\{1, 2, \ldots, n\}$. Platí

$$\begin{aligned} &\operatorname{sgn}(\mathrm{id}_n)(a_{11} - \lambda)(a_{22} - \lambda) \cdots (a_{nn} - \lambda) \\ &\qquad = (-1)^n \lambda^n + (a_{11} + a_{22} + \cdots + a_{nn})(-1)^{n-1} \lambda^{n-1} + \cdots \ . \end{aligned}$$

V jakémkoliv součinu, který je definován neidentickou permutací $\pi \in S_n$, se musí vyskytnout aspoň dva činitelé b_{ij}, které neleží na hlavní diagonále matice B, a proto se v něm vyskytuje nejvýše $n - 2$ činitelů obsahujícíh λ, po roznásobení můžeme dostat nenulové koeficienty poue u mocnin λ^{n-2} nebo nižších. Koeficienty u λ^n a λ^{n-1} pocházejí proto pouze ze součinu určeného identickou permutací. Tím jsou dokázány body (1), (2) a (3).

Hodnotu absolutního členu c_0 jakéhokoliv polynomu

$$p(\lambda) = c_n \lambda^n + c_{n-1} \lambda^{n-1} + \cdots + c_1 \lambda + c_0$$

získáme tak, že do něho dosadíme hodnotu proměnné $\lambda = 0$. Proto se absolutní člen polynomu $\det(A - \lambda I_n) = \det(A - 0\, I_n) = \det A$. □

Z definice determinantu vyplývá, že $\det(A - \lambda I_n)$ je polynom nejvýše n-tého stupně (kde n je řád matice A) v proměnné λ. Nazýváme jej charakteristický polynom matice A. Jeho kořeny jsou podle pozorování **??** vlastní čísla matice A.

Definice 9.25. Je-li A čtvercová matice řádu n nad tělesem $\mathbf{T}$, pak *charakteristický polynom matice* A je polynom

$$p_A(\lambda) = \det(A - \lambda I_n) \ .$$

Tvrzení 9.24 ukazuje, že charakteristický polynom jakékoliv matice A řádu n má stupeň právě n a pozorování 9.18 a 9.19 říkají, že vlastní čísla matice A nebo operátoru f na lineárním prostoru dimanze n, jsou jeho kořeny.

Například charakteristický polynom matice A z příkladů 9.3 a 9.20 je

$$\det(A - \lambda I_2) = \det \begin{pmatrix} 3-\lambda & 0 \\ 1 & 2-\lambda \end{pmatrix} = (3-\lambda)(2-\lambda) = \lambda^2 - 5\lambda + 6 \ .$$

Charakteristický polynom lineárního operátoru f na konečně dimenzionálním prostoru $\mathbf{V}$ definujeme jako charakteristický polynom matice $[f]_B^B$, kde B je nějaká báze prostoru $\mathbf{V}$. Musíme ovšem ověřit, že polynom nezávisí na volbě báze, jak jsme v konkrétní situaci viděli v příkladu 9.22.

Uvažujme tedy dvě různé báze B, C prostoru $\mathbf{V}$. Podle tvrzení 6.20 je

$$[f]_B^B = R^{-1}[f]_C^C R \ ,$$

kde R je matice přechodu od B k C. Matice svázané takovým vztahem nazýváme podobné.

Definice 9.26. Dvě čtvercové matice X, Y téhož řádu nad tělesem $\mathbf{T}$ se nazývají *podobné*, pokud existuje regulární matice R taková, že $Y = R^{-1}XR$.

Relace podobnosti matic je ekvivalence na množině všech čtvercových matic téhož řádu n nad tělesem $\mathbf{T}$, důkaz ponecháme jako cvičení. Podle diskuze předcházející definici 9.26 jsou matice $[f]_B^B$ a $[f]_C^C$ podobné.

Tvrzení 9.27. *Podobné matice mají stejný charakteristický polynom.*

Důkaz. Jsou-li X, Y dvě podobné matice téhož řádu nad tělesem $\mathbf{T}$, pak existuje regulární matice R taková, že $Y = R^{-1}XR$. Potom platí

$$\begin{aligned} \det(Y - \lambda I_n) &= \det(R^{-1}XR - \lambda I_n) = \det(R^{-1}XR - R^{-1}\lambda I_n R) \\ &= \det(R^{-1}(X - \lambda I_n)R) = \det(R)^{-1}\det(X - \lambda I_n)\det(R) \\ &= \det(X - \lambda I_n) \end{aligned}$$

podle věty o násobení determinantů a jejím důsledku pro determinant inverzní matice. $\square$

To opravňuje definici charakteristického polynomu lineární operátoru na konečně generovaném prostoru.

Definice 9.28. Je-li $f : \mathbf{V} \to \mathbf{V}$ lineární operátor na konečně generovaném prostoru $\mathbf{V}$ dimenze n, pak *charakteristický polynom operátoru* f je polynom

$$p_f(\lambda) = \det\left([f]_B^B - \lambda I_n\right) \ ,$$

kde B je libovolná báze prostoru $\mathbf{V}$.

Charakteristický polynom operátoru na (konečně generovaném prostoru) $\mathbf{V}$ je tedy roven charakteristickému polynomu matice $[f]_B^B$ pro jakoukoliv bázi B prostoru $\mathbf{V}$.

Příklad 9.29. V příkladu 9.12 jsme spočetli, že charakteristický polynom ortogonální projekce f na přímku $\left\langle (1,2)^T \right\rangle$ je $p_f(\lambda) = \lambda^2 - \lambda$. Z druhého postupu nahlédněte, že stejně vyjde charakteristický polynom pro ortogonální projekci na libovolnou přímku, dokonce projekci na libovolnou přímku p ve směru přímky q, pokud $p \neq q$.

Příklad 9.30. Charakteristický polynom reálné matice

$$A = \begin{pmatrix} 3 & 7 \\ 4 & 5 \end{pmatrix}$$

je podle tvrzení roven

$$p_A(\lambda) = (-1)^2\lambda^2 - (3+5)\lambda + (3 \cdot 5 - 7 \cdot 4) = \lambda^2 - 8\lambda - 13 \ .$$

9.2.3. *Kořeny polynomů, algebraická násobnost.* K určení vlastních čísel matice nebo operátoru na konečně generovaném prostoru potřebujeme najít kořeny charakteristického polynomu. Uvedeme několik pojmů a tvrzení (bez důkazů), které budeme o kořenech polynomů potřebovat. Důkazy se dozvíte v kurzu obecné algebry ve druhém ročníku.

Připomeňme, že polynom stupně n nad tělesem $\mathbf{T}$ je výraz

$$p(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n, \text{ kde } a_0, \ldots, a_n \in T, \ a_n \neq 0 \ .$$

Kořenem takového polynomu je prvek $t \in T$, pro který $p(t) = a_0 + a_1t + a_2t^2 + \cdots + a_nt^n = 0$. Nulovému polynomu $p(x) = 0$ předchozí definice nepřidělila stupeň, někdy se říká, že je stupně -1, jindy se stupeň nedefinuje.

Nyní směřujeme k pojmu násobnosti kořene. Řekneme, že polynom $p(x)$ dělí polynom $s(x)$ (oba polynomy jsou nad tělesem $\mathbf{T}$), pokud existuje polynom $q(x)$ (nad $\mathbf{T}$) takový, že $p(x)q(x) = s(x)$. Například reálný polynom $x - 1$ dělí polynom $x^2 - 1$, protože $(x-1)(x+1) = x^2 - 1$.

Tvrzení 9.31. *Nechť $p(x)$ je polynom nad $\mathbf{T}$. Prvek $t \in T$ je kořenem polynomu $p(x)$ právě tehdy, když polynom $x - t$ dělí polynom $p(x)$.*

Největší číslo l takové, že $(x-t)^l$ stále dělí polynom $p(x)$ nazýváme násobnost kořene t.

Definice 9.32. Nechť $p(x)$ je polynom nad $\mathbf{T}$ a $t \in T$ je jeho kořen. *Násobnost kořene t* polynomu $p(x)$ definujeme jako největší přirozené číslo l takové, že polynom $(x-t)^l$ dělí polynom $p(x)$.

Tvrzení 9.33. *Nechť $p(x)$ je polynom nad $\mathbf{T}$, $t_1, \ldots, t_k \in T$ po dvou různé a $l_1, \ldots, l_k \in \mathbb{N}$. Pak následující tvrzení jsou ekvivalentní.*

(1) *Pro každé $i \in \{1, \ldots, k\}$ je t_i kořen násobnosti l_i.*
(2) *$p(x) = (x-t_1)^{l_1} \ldots (x-t_k)^{l_k} q(x)$ pro nějaký polynom $q(x)$ takový, že ani jeden z prvků $t_1, \ldots, t_k$ není kořen.*

Implikaci (2) $\Rightarrow$ (1) můžeme použít k hledání násobnosti kořenů. Uhádneme nějaký kořen t polynomu $p(x)$ a polynom $p(x)$ vydělíme polynomem $x - t$. Dále pokračujeme stejným způsobem s výsledným polynomem. Proces ukončíme, když získáme polynom, který již žádný kořen nemá. Nakonec získáme rozklad jako v části (2) a tvrzení nám dá informace o násobnostech.

Příklad 9.34. Určíme kořeny a jejich násobnosti pro reálný polynom $p(x) = 2x^3 - 8x^2 + 10x - 4$.

Uhádneme kořen $t = 1$ a vydělíme polynom $p(x)$ polynomem $x - 1$.

$$q(x) = (2x^3 - 8x^2 + 10x - 4) : (x-1) = 2x^2 - 6x + 4 \ .$$

Hledat kořeny reálného polynomu druhé stupně umíme, polynom $q(x)$ má kořeny 1 a 2. Polynom $q(x)$ lze proto zapsat

$$q(x) = 2(x-1)(x-2) \ .$$

Pro původní polynom máme

$$p(x) = (x-1)q(x) = 2(x-1)^2(x-2) \ .$$

Tedy $p(x)$ má kořen 1 násobnosti 2 a kořen 2 násobnosti 1.

Příklad 9.35. Určíme kořeny a násobnosti kořenů reálného polynomu $p(x) = x^4 + x^2$.

Ihned vidíme, že x^2 dělí polynom $p(x)$. Máme

$$p(x) = x^2(x^2+1) \ ,$$

Polynom x^2+1 již žádný reálný kořen nemá, tedy $p(x)$ má jediný kořen 0 násobnosti 2.

Chápeme-li ovšem $p(x)$ jako polynom nad komplexními čísly, pak x^2+1 má dva kořeny i a $-i$ a polynom $p(x)$ lze psát

$$p(x) = x^2(x+i)(x-i)$$

Nad komplexními čísly tedy máme kořen 0 násobnosti 2 a kořeny i a $-i$ násobnosti 1.

Příklad 9.36. Určíme kořeny polynomu $p(x) = x^4 + 2x^3 + x^2 + 2x$ nad tělesem $\mathbb{Z}_3$.

Vidíme, že $t = 0$ je kořen. Dosazením zjistíme, že $t = 1$ je kořenem a $t = 2$ kořenem není. K určení násobností vydělíme polynom $p(x)$ polynomem $x(x-1) = x^2 - x = x^2 + 2x$.

$$(x^4 + 2x^3 + x^2 + 2x) : (x^2 + 2x) = x^2 + 1 \ .$$

Takže

$$p(x) = x(x+2)(x^2+1) \ .$$

Dosazením prvků tělesa $\mathbb{Z}_3$ do polynomu x^2+1 zjistíme, že x^2+1 žádné kořeny v $\mathbb{Z}_3$ nemá, takže oba kořeny 0 i 1 mají násobnost 1.

Důsledkem implikace (1) $\Rightarrow$ (2) v tvrzení 9.24 je, že polynom stupně $n \geq 0$ má nejvýše n kořenů, i když počítáme každé tolikrát, kolik je jeho násobnost.

Tvrzení 9.37. *Polymom stupně n nad libovolným tělesem má nejvýše n kořenů včetně násobností.*

Obrat „včetně násobností“ budeme používat pro stručnost vyjadřování. Přesný význam je vysvětlený nad tvrzením, tj. pokud každý kořen počítáme tolikrát, kolik je jeho násobnost, vyjde nejvýše n.

Ze základní věty algebry (věta **??**) lze odvodit, že nad komplexními čísly je kořenů vždy maximální možný počet (opět musíme počítat i s násobnostmi).

Věta 9.38. *Každý polynom stupně $n \geq 1$ nad tělesem $\mathbb{C}$ lze napsat jako součin lineárních polynomů (tj. polynomů stupně 1).*

Speciálně, každý polynom stupně $n \geq 0$ nad tělesem $\mathbb{C}$ má právě n kořenů včetně násobností.

(Na okraj poznamenejme, že každé těleso lze rozšířit do většího tělesa, kde platí obdoba předchozí věty.)

Ještě uvedeme jeden pozitivní výsledek pro polynomy nad reálnými čísly, tentokrát výjimečně i s náznakem důkazem.

Tvrzení 9.39. *Polynom lichého stupně nad tělesem $\mathbb{R}$ má alespoň jeden kořen.*

Důkaz. Připomeňme, že je-li komplexní číslo z kořenem reálného polynomu $p(x)$, pak je jeho kořenem také číslo $\overline{z}$ komplexně sdružené se z (viz věta 1.9). Kořeny polynomu $p(x)$ tak můžeme uspořádat do dvojic komplexně sdružených kořenů. Protože ale všech kořenů (spolu s násobnostmi) je lichý počet, existuje aspoň jeden kořen z, pro který platí $z = \overline{z}$, tj. aspoň jeden reálný kořen.

Alternativně lze tvrzení dokázat analyticky, bez použití komplexních čísel. Je-li koeficent u x^n kladný, pak pro $x \to \infty$ je $p(x) \to \infty$ a pro $x \to -\infty$ je $p(x) \to -\infty$. Ze spojitosti funkce $p(x)$ pak vidíme, že nutně existuje číslo z, splňující $p(z) = 0$. (Je-li koefcient u x^n záporný, důkaz je obdobný.) □

Nyní se vrátíme k vlastním číslům. Ty lze spočítat jako kořeny charakteristického polynomu, jejich násobnosti jsou důležitou informací, proto si zaslouží samostatný pojem.

Definice 9.40. Nechť f je lineární operátor na konečně generovaném prostoru a λ je jeho vlastní číslo. *Algebraickou násobností* vlastního čísla λ rozumíme jeho násobnost jako kořene charakteristického polynomu operátoru f.

Definice 9.41. Nechť A je čtvercová matice a λ je její vlastní číslo. *Algebraickou násobností* vlastního čísla λ rozumíme jeho násobnost jako kořene charakteristického polynomu matice A.

Příklad 9.42. Najdeme vlastní čísla a jejich algebraické násobnosti pro lineární operátor $f : \mathbb{R}^3 \to \mathbb{R}^3$ definovaný předpisem

$$f\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -y+z \\ -3x-2y+3z \\ -2x-2y+3z \end{pmatrix} .$$

Matice operátoru f vzhledem ke kanonické bázi je

$$A = [f]_K^K = \begin{pmatrix} 0 & -1 & 1 \\ -3 & -2 & 3 \\ -2 & -2 & 3 \end{pmatrix} .$$

Charakteristický polynom operátoru f se rovná

$$\det(A-\lambda I_3) = \det\begin{pmatrix} 0-\lambda & -1 & 1 \\ -3 & -2-\lambda & 3 \\ -2 & -2 & 3-\lambda \end{pmatrix} = -\lambda^3+\lambda^2+\lambda-1 = -(\lambda-1)^2(\lambda+1) .$$

Operátor má tedy 2 různá vlastní čísla: vlastní číslo 1 algebraické násobnosti 2 a vlastní číslo -1 algebraické násobnosti 1. (Takže dohromady máme 3 vlastní čísla včetně násobností.)

Zformulujeme důsledky tvrzení 9.37, věty 9.38 a tvrzení 9.39 pro vlastní čísla operátorů.

Důsledek 9.43.

- *Každý lineární operátor $f : \mathbf{V} \to \mathbf{V}$ na konečně generovaném prostoru dimenze n nad tělesem $\mathbf{T}$ má nejvýše n vlastních čísel včetně násobností.*
- *Lineární operátor $f : \mathbf{V} \to \mathbf{V}$ má právě n vlastních čísel včetně násobností právě tehdy, když je jeho charakteristický polynom součinem lineárních polynomů.*
- *Každý lineární operátor $f : \mathbf{V} \to \mathbf{V}$ na konečně generovaném prostoru dimenze n nad tělesem $\mathbb{C}$ má právě n vlastních čísel včetně násobností.*

- *Každý lineární operátor $f : \mathbf{V} \to \mathbf{V}$ na konečně generovaném vektorovém prostoru liché dimenze nad $\mathbb{R}$ má aspoň jedno (reálné) vlastní číslo.*

V řeči matic.

Důsledek 9.44.

- *Každá čtvercová matice řádu n nad tělesem $\mathbf{T}$ má nejvýše n vlastních čísel včetně algebraických násobností.*
- *Čtvecová matice řádu n nad tělesem $\mathbf{T}$ má právě n vlastních čísel včetně násobností právě tehdy, když je její charakteristický polynom součinem lineárních polynomů.*
- *Každá čtvercová matice řádu n nad tělesem $\mathbb{C}$ má právě n vlastních čísel včetně algebraických násobností.*
- *Každá čtvercová matice lichého řádu nad tělesem $\mathbb{R}$ má alespoň jedno reálné vlastní číslo.*

9.3. **Diagonalizovatelné operátory.**

Má-li operátor f (resp. matice A) „dostatek“ vlastních vektorů, můžeme diskrétní dynamický systém $\mathbf{x}_k = f(\mathbf{x}_{k-1})$ (resp. $\mathbf{x}_k = A\mathbf{x}_{k-1}$) snadno vyřešit. K ilustraci poslouží opět operátor z příkladů 9.3 a 9.20.

V kapitole budeme často používat matici operátoru vzhledem k bázi B a B, tj. matici $[f]^B_B$. Budeme ji proto jednoduše nazývat *matice f vzhledem k B*.

Důležitou roli v této části budou také hrát diagonální matice, proto si pro ně zavedeme speciální označení. Diagonální matici $D = (d_{ij})$ řádu n budeme zapisovat $\mathrm{diag}(d_{11}, d_{22}, \cdots, d_{nn})$. Diagonální matice umíme snadno umocnit.

$$\mathrm{diag}(t_1, t_2, \ldots, t_n)^k = \mathrm{diag}(t_1^k, t_2^k, \ldots, t_n^k)$$

pro každý exponent $k = 0, 1, 2, \ldots$.

Příklad 9.45. Uvažujme operátor f_A na $\mathbb{R}^2$ určený maticí

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

V příkladu 9.20 jsme vypočítali, že vlastní čísla tohoto operátoru jsou 2 a 3 a příslušné podprostory vlastních vektorů jsou

$$M_2 = \left\langle \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\rangle, \quad M_3 = \left\langle \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\rangle .$$

Vektory $\mathbf{v}_1 = (0,1)^T$, $\mathbf{v}_2 = (1,1)^T$ tvoří lineárně nezávislou posloupnost $B = (\mathbf{v}_1, \mathbf{v}_2)$, tedy bázi prostoru $\mathbb{R}^2$. (To, že posloupnost je lineárně nezávislá, není náhoda – viz věta 9.54.)

Platí $f_A(\mathbf{v}_1) = 2\mathbf{v}_1$ a $f_A(\mathbf{v}_2) = 3\mathbf{v}_2$. Z toho vidíme, že matice operátoru f_A vzhledem k bázi B je

$$[f_A]^B_B = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} .$$

Diagonální matice ale mocnit umíme! Pro libovolné přirozené k proto umíme vypočítat matici operátoru $(f_A)^k$ vzhledem k bázi B:

$$[(f_A)^k]^B_B = ([f_A]^B_B)^k = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}^k = \begin{pmatrix} 2^k & 0 \\ 0 & 3^k \end{pmatrix} .$$

Nyní můžeme odpovědět na řadu otázek o operátoru f_A a matici A.

• **Řešení diskrétního dynamického systému $\mathbf{x}_k = f_A(\mathbf{x}_{k-1})$ „v bázi B“.** Platí

$$[\mathbf{x}_k]_B = [(f_A)^k(\mathbf{x}_0)]_B = [(f_A)^k]^B_B\,[\mathbf{x}_0]_B = \begin{pmatrix} 2^k & 0 \\ 0 & 3^k \end{pmatrix}[\mathbf{x}_0]_B \ .$$

Jsou-li tedy souřadnice $[\mathbf{x}_0]_B$ počátečního stavu $\mathbf{x}_0$ vzhledem k bázi B rovny

$$[\mathbf{x}_0]_B = \begin{pmatrix} r \\ s \end{pmatrix} \ ,$$

pak souřadnice stavu $\mathbf{x}_k$ v bázi B jsou

$$[\mathbf{x}_k]_B = \begin{pmatrix} 2^k & 0 \\ 0 & 3^k \end{pmatrix}\begin{pmatrix} r \\ s \end{pmatrix} = \begin{pmatrix} r2^k \\ s3^k \end{pmatrix} \ .$$

• **Kvalitativní chování diskrétního dynamického systému $\mathbf{x}_k = f_A(\mathbf{x}_{k-1})$.** Pokud $r \neq 0$ a $s \neq 0$, pak se pro $k \to \infty$ budou obě souřadnice vzhledem k B v absolutní hodnotě blížit nekonečnu. Přitom první složka bude pro velká k zanedbatelná vzhledem ke složce druhé.

• **Řešení diskrétního dynamického systému $\mathbf{x}_k = f_A(\mathbf{x}_{k-1})$ „v kanonické bázi“.** Z matice $[(f_A)^k]^B_B$ můžeme určit matici $(f_A)^k$ vzhledem ke kanonickým bázím pomocí matic přechodu (opakujeme výpočet v tvrzení 6.20):

$$\begin{aligned}[(f_A)^k]^K_K = [\mathrm{id}]^B_K\,[(f_A)^k]^B_B\,[\mathrm{id}]^K_B &= \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 2^k & 0 \\ 0 & 3^k \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}^{-1} \\ &= \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 2^k & 0 \\ 0 & 3^k \end{pmatrix}\begin{pmatrix} -1 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 3^k & 0 \\ 3^k - 2^k & 2^k \end{pmatrix}\end{aligned}$$

Z toho dostáváme

$$\mathbf{x}_k = (f_A)^k(\mathbf{x}_0) = [(f_A)^k]^K_K\,\mathbf{x}_0 = \begin{pmatrix} 3^k & 0 \\ 3^k - 2^k & 2^k \end{pmatrix}\mathbf{x}_0 \ .$$

Tedy pokud $\mathbf{x}_0 = (a,b)^T$, pak

$$\mathbf{x}_k = (f_A)^k(\mathbf{x}_0) = \begin{pmatrix} 3^k a \\ (3^k - 2^k)a + 2^k b \end{pmatrix} \ .$$

• **Výpočet k-té mocniny A^k matice A pro $k \geq 1$.** Protože $[(f_A)^k]^K_K = [f_{A^k}]^K_K = A^k$ máme z předchozího bodu

$$A^k = \begin{pmatrix} 3^k & 0 \\ 3^k - 2^k & 2^k \end{pmatrix} \ .$$

Příklad nás vede k definici diagonalizovatelného operátoru.

Definice 9.46. Lineární operátor $f : \mathbf{V} \to \mathbf{V}$ na konečně generovaném prostoru $\mathbf{V}$ nazýváme *diagonalizovatelný*, pokud má vzhledem k nějaké bázi diagonální matici.

Tvrzení 9.47. *Je-li $f : \mathbf{V} \to \mathbf{V}$ lineární operátor na konečně generovaném prostoru $\mathbf{V}$ a je-li $B = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$ báze prostoru $\mathbf{V}$, pak $[f]^B_B = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ platí právě tehdy, když pro každé $i \in \{1, 2, \ldots, n\}$ je $\mathbf{v}_i$ vlastní vektor příslušný vlastnímu číslu λ_i.*

Důkaz. Rovnost $[f]^B_B = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ platí právě když pro každé $i \in \{1, \ldots, n\}$ se i-tý sloupec matice $[f]^B_B$ rovná i-tému sloupci matice $\mathrm{diag}(\lambda_1, \ldots, \lambda_n)$, tj. právě když $[f(\mathbf{v}_i)]_B = \lambda_i \mathbf{e}_i$ (použili jsme definici matice lineárního zobrazení).

Rovnost $[f(\mathbf{v}_i)]_B = \lambda_i \mathbf{e}_i$ je ekvivalentní vztahu $f(\mathbf{v}_i) = \lambda_i \mathbf{v}_i$ podle definice souřadnic prvku vzhledem k bázi. Protože $\mathbf{v}_i \neq \mathbf{o}$, vztah $f(\mathbf{v}_i) = \lambda_i \mathbf{v}_i$ je ekvivalentní tomu, že $\mathbf{v}_i$ je vlastní vektor příslušný vlastnímu číslu λ_i. □

Důsledek 9.48. *Lineární operátor $f : \mathbf{V} \to \mathbf{V}$ na konečně generovaném prostoru $\mathbf{V}$ je diagonalizovatelný právě tehdy, když existuje báze prostoru $\mathbf{V}$ tvořená vlastními vektory operátoru f.*

Ekvivalentně můžeme diagonalizovatelnost charakterizovat pomocí podobnosti matice operátoru (vzhledem k libovolné bázi) s diagonální maticí.

Tvrzení 9.49. *Je-li $f : \mathbf{V} \to \mathbf{V}$ lineární operátor na konečně generovaném prostoru $\mathbf{V}$ dimenze n nad tělesem $\mathbf{T}$ a C báze prostoru $\mathbf{V}$, pak operátor f je diagonalizovatelný právě tehdy, když je matice $[f]_C^C$ podobná diagonální matici.*

Důkaz. Je-li operátor f diagonalizovatelný, pak existuje báze B prostoru $\mathbf{V}$ taková, že $[f]_B^B$ je diagonální matice. Podle tvrzení 6.20) jsou matice $[f]_C^C$ a $[f]_B^B$ jsou podobné.

Naopak, je-li matice $[f]_C^C$ podobná diagonální matici D, pak existuje regulární matice $R = (\mathbf{r}_1|\mathbf{r}_2|\cdots|\mathbf{r}_n)$ taková, že $D = R^{-1}[f]_C^C\, R$. Protože je matice R regulární, je posloupnost sloupcových vektorů $(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_n)$ lineárně nezávislá a proto báze v $\mathbf{T}^n$. Zvolíme vektory $\mathbf{v}_i \in \mathbf{V}$ tak, aby platilo $[\mathbf{v}_i] = \mathbf{r}_i$ pro každé $i = 1, 2, \ldots, n$. Posloupnost $B = (\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n)$ je pak báze prostoru $\mathbf{V}$ a platí pro ni $[\mathrm{id}]_C^B = R$. Opět podle tvrzení 6.20) pak platí $[f]_B^B = R^{-1}[f]_C^C\, R = D$. □

Přeformulujeme si definici a tvrzení o diagonalizovatelnosti operátorů pro matice.

Definice 9.50. Čtvercová matice A řádu n nad tělesem $\mathbf{T}$ se nazývá *diagonalizovatelná*, pokud je operátor $f_A : \mathbf{T}^n \to \mathbf{T}^n$ diagonalizovatelný.

Pro maticovou formulaci tvrzení 9.51 si opět připomeneme, že pro čtvercovou matici A řádu n nad $\mathbf{T}$ a bázi $B = (\mathbf{v}_1, \ldots\ \mathbf{v}_n)$ prostoru $\mathbf{T}^n$ platí

$$[f_A]_B^B = [\mathrm{id}]_B^K\,[f_A]_K^K\,[\mathrm{id}]_K^B = R^{-1}AR, \quad \text{kde } R = [\mathrm{id}]_K^B = (\mathbf{v}_1|\ldots|\mathbf{v}_n)\ .$$

Tvrzení 9.51. *Je-li A čtvercová matice řádu n nad tělesem $\mathbf{T}$, $B = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$ báze prostoru $\mathbf{T}^n$ a $R = [\mathrm{id}]_K^B = (\mathbf{v}_1|\ldots|\mathbf{v}_n)$, pak matice $[f_A]_B^B = R^{-1}AR$ se rovná diagonální matici $\mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ právě tehdy, když pro každé $i \in \{1, 2, \ldots, n\}$ je $\mathbf{v}_i$ vlastní vektor příslušný vlastnímu číslu λ_i.*

Důsledek 9.52. *Čtvercová matice A řádu n nad tělesem $\mathbf{T}$ je diagonalizovatelná právě tehdy, když existuje báze prostoru $\mathbf{T}^n$ tvořená vlastními vektory matice A.*

V situaci, kdy $[f]_B^B = [\mathrm{id}]_B^K\,[f_A]_K^K\,[\mathrm{id}]_K^B = R^{-1}AR = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$, umíme matici A umocnit stejně jako v příkladu 9.45:

$$\begin{aligned} A^k &= [(f_A)^k]_K^K = [\mathrm{id}]_K^B\,[(f_A)^k]_B^B\,[\mathrm{id}]_B^K = R\,([f_A]_B^B)^k\,R^{-1} \\ &= R\ \mathrm{diag}(\lambda_1, \ldots, \lambda_n)^k\,R^{-1} = R\ \mathrm{diag}(\lambda_1^k, \ldots, \lambda_n^k)\,R^{-1}\ . \end{aligned}$$

Výpočet mocniny A^k můžeme nahlédnout také algebraicky. Označíme $D = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ a vztah $R^{-1}AR = D$ přepíšeme na $A = RDR^{-1}$. Pak

$$\begin{aligned} A^k &= \underbrace{(RDR^{-1})(RDR^{-1})\ldots(RDR^{-1})}_{k\times} = R\,\underbrace{DD\cdots D}_{k\times}\,R^{-1} = RD^kR^{-1} \\ &= R\ \mathrm{diag}(\lambda_1^k, \ldots, \lambda_n^k)\,R^{-1}\ . \end{aligned}$$

Použitím tvrzení 9.49 na kanonickou bázi $C = K_n$ dostaneme následující maticovou verzi.

Tvrzení 9.53. *Čtvercová matice A řádu n nad tělesem $\mathbf{T}$ je diagonalizovatelná právě tehdy, když je podobná diagonální matici.*

9.3.1. *Lineární nezávislost vlastních vektorů.* Chceme nalézt nutné a postačující podmínky pro to, aby byl lineární operátor $f : \mathbf{V} \to \mathbf{V}$ na konečně generovaném prostoru $\mathbf{V}$ diagonalizovatelný. Základem je následující věta, která platí zcela obecně, není nutné předpokládat, že prostor $\mathbf{V}$ má konečnou dimenzi.

Věta 9.54. *Je-li $f : \mathbf{V} \to \mathbf{V}$ lineární operátor a $(\mathbf{v_1}, \mathbf{v}_2, \ldots, \mathbf{v}_k)$ posloupnost nenulových vlastních vektorů operátoru f příslušných navzájem různým vlastním číslům $\lambda_1, \ldots, \lambda_k$, pak je posloupnost $(\mathbf{v_1}, \mathbf{v}_2, \ldots, \mathbf{v}_k)$ lineárně nezávislá.*

Důkaz. Použijeme indukci podle k. Je-li $k = 1$, tvrzení platí, protože $\mathbf{v}_1 \neq \mathbf{o}$. Předpokládejme, že $k > 1$ a tvrzení platí pro $k-1$, tj. že posloupnost $(\mathbf{v_1}, \ldots, \mathbf{v}_{k-1})$ je lineárně nezávislá. Uvažujme skaláry $a_1, \ldots, a_k \in \mathbf{T}$ takové, že platí

$$\mathbf{o} = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_{k-1}\mathbf{v}_{k-1} + a_k\mathbf{v}_k \ .$$

Aplikujeme na obě strany operátor $(f - \lambda_k \,\mathrm{id}_V)$ a upravíme. V prvních dvou úpravách používáme linearitu operátoru $(f - \lambda_k \,\mathrm{id}_V)$.

$$(f - \lambda_k \,\mathrm{id}_V)(a_1\mathbf{v}_1 + \cdots + a_{k-1}\mathbf{v}_{k-1} + a_k\mathbf{v}_k) = (f - \lambda_k \,\mathrm{id}_V)(\mathbf{o})$$
$$(f - \lambda_k \,\mathrm{id}_V)(a_1\mathbf{v}_1) + \cdots + (f - \lambda_k \,\mathrm{id}_V)(a_{k-1}\mathbf{v}_{k-1}) + (f - \lambda_k \,\mathrm{id}_V)(a_k\mathbf{v}_k) = \mathbf{o}$$
$$a_1(f - \lambda_k \,\mathrm{id}_V)(\mathbf{v}_1) + \cdots + a_{k-1}(f - \lambda_k \,\mathrm{id}_V)(\mathbf{v}_{k-1}) + a_k(f - \lambda_k \,\mathrm{id}_V)(\mathbf{v}_k) = \mathbf{o} \ .$$

K poslední úpravě využijeme toho, že pro každé $i = 1, 2, \ldots, k$ platí

$$(f - \lambda_k \,\mathrm{id}_V)(\mathbf{v}_i) = f(\mathbf{v}_i) - (\lambda_k \,\mathrm{id}_V)(\mathbf{v}_i) = \lambda_i\mathbf{v}_i - \lambda_k\mathbf{v}_i = (\lambda_i - \lambda_k)\mathbf{v}_i \ ,$$

a dostaneme

$$\mathbf{o} = a_1(\lambda_1 - \lambda_k)\mathbf{v}_1 + \cdots + a_{k-1}(\lambda_{k-1} - \lambda_k)\mathbf{v}_{k-1} \ .$$

Posloupnost vektorů $(\mathbf{v_1}, \ldots, \mathbf{v}_{k-1})$ je lineárně nezávislá podle indukčního předpokladu. Odtud plyne

$$a_1(\lambda_1 - \lambda_k) = a_2(\lambda_2 - \lambda_k) = \cdots = a_{k-1}(\lambda_{k-1} - \lambda_k) = 0 \ .$$

Protože vlastní čísla $\lambda_1, \ldots, \lambda_{k-1}, \lambda_k$ jsou navzájem různá, vyplývá odtud, že $a_1 = a_2 = \cdots = a_{k-1} = 0$. Po dosazení do rovnosti

$$a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_{k-1}\mathbf{v}_{k-1} + a_k\mathbf{v}_k = \mathbf{o}$$

dostaneme $a_k\mathbf{v}_k = \mathbf{o}$ a tedy také $a_k = 0$, protože $\mathbf{v}_k \neq \mathbf{o}$. Tím je dokázáno, že posloupnost $(\mathbf{v_1}, \ldots, \mathbf{v}_{k-1}, \mathbf{v}_k)$ je lineárně nezávislá. □

Důsledek 9.55. *Má-li lineární operátor $f : \mathbf{V} \to \mathbf{V}$ na vektorovém prostoru $\mathbf{V}$ dimenze n nad tělesem $\mathbf{T}$ celkem n navzájem různých vlastních čísel, pak je diagonalizovatelný.*

Důkaz. Má-li operátor f celkem n navzájem různých vlastních čísel $\lambda_1, \ldots, \lambda_n$, existuje pro každé $i = 1, \ldots, n$ nenulový vlastní vektor $\mathbf{v}_i$ příslušný λ_i. Podle předchozí věty je posloupnost vlastních vektorů $(\mathbf{v}_1, \ldots, \mathbf{v}_n)$ lineárně nezávislá a tedy je to báze prostoru $\mathbf{V}$. Operátor f má tak bázi složenou z vlastních vektorů operátoru f, je proto diagonalizovatelný podle důsledku 9.48. □

Důsledek 9.56. *Má-li matice A řádu n nad tělesem $\mathbf{T}$ celkem n navzájem různých vlastních čísel, pak je diagonalizovatelná.*

Operátor v příkladu 9.45 a operátor v motivačním příkladu o Fibonacciho posloupnosti v části 9.1.2 splňují předpoklad důsledku 9.55.

Příklad 9.57. Ještě jednou spočítáme, jak vypadá k-tý prvek Fibonacciho posloupnosti. Najdeme vlastní čísla a vlastní vektory matice C. Charakteristický polynom matice C je (podle tvrzení 9.24 o koeficientech charakteristického polynomu) roven

$$p_C(\lambda) = \lambda^2 - \lambda - 1$$

Vlastní čísla matice C, neboli kořeny rovnice $p_C(\lambda) = 0$, jsou

$$\lambda_1 = \frac{1+\sqrt{5}}{2}, \quad \lambda_2 = \frac{1-\sqrt{5}}{2} = 1 - \lambda_1 \ .$$

Všechny vlastní vektory příslušné vlastnímu číslu λ_1 jsou právě všechna řešení homogenní soustavy s maticí

$$\begin{pmatrix} 0-\lambda_1 & 1 \\ 1 & 1-\lambda_1 \end{pmatrix} ,$$

což jsou všechny vektory v $M_{\lambda_1} = \langle (1, 1/2 + \sqrt{5}/2)^T \rangle = \langle (1, \lambda_1)^T \rangle$. Podobně jsou všechny vlastní vektory příslušné vlastnímu číslu $\lambda_2 = 1 - \lambda_1$ právě vektory z lineárního obalu vektoru $(1, \lambda_2)^T$.

Posloupnost

$$B = (\mathbf{v}_1, \mathbf{v}_2) = \left(\begin{pmatrix} 1 \\ \lambda_1 \end{pmatrix}, \begin{pmatrix} 1 \\ \lambda_2 \end{pmatrix} \right)$$

je (i podle věty 9.54) lineárně nezávislá, takže je bází $\mathbb{R}^2$, a platí

$$[f_C]^B_B = \mathrm{diag}(\lambda_1, \lambda_2), \quad [(f_C)^k]^B_B = \mathrm{diag}(\lambda_1^k, \lambda_2^k) \ .$$

Chceme znát $C^k(a_0, a_1)^T = (f_C)^k(a_0, a_1)^T$. Teď máme více možností, jak výpočet dokončit. Můžeme přejít ke kanonické bázi (tj. spočítat $[(f_C)^k]^K_K = C^k$), nebo pracovat přímo v bázi B. Zvolíme druhý přístup.

Vyjádříme vektor $(a_0, a_1)^T = (0, 1)^T$ vzhledem k bázi B. Vyjde

$$\left[\begin{pmatrix} 0 \\ 1 \end{pmatrix} \right]_B = \frac{1}{\sqrt{5}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \ .$$

Odtud dostáváme pro každé celé číslo $k \geq 0$

$$\left[\begin{pmatrix} a_k \\ a_{k+1} \end{pmatrix} \right]_B = [(f_C)^k]^B_B \left[\begin{pmatrix} 0 \\ 1 \end{pmatrix} \right]_B = \begin{pmatrix} \lambda_1^k & 0 \\ 0 & \lambda_2^k \end{pmatrix} \frac{1}{\sqrt{5}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \frac{1}{\sqrt{5}} \begin{pmatrix} \lambda_1^k \\ -\lambda_2^k \end{pmatrix} \ .$$

Z toho vyplývá

$$\begin{pmatrix} a_k \\ a_{k+1} \end{pmatrix} = \frac{1}{\sqrt{5}} \left(\lambda_1^k \begin{pmatrix} 1 \\ \lambda_1 \end{pmatrix} - \lambda_2^k \begin{pmatrix} 1 \\ \lambda_2 \end{pmatrix} \right) = \frac{1}{\sqrt{5}} \begin{pmatrix} \lambda_1^k - \lambda_2^k \\ \lambda_1^{k+1} - \lambda_2^{k+1} \end{pmatrix} \ .$$

Srovnáním prvních složek dostáváme

$$a_k = \frac{\lambda_1^k}{\sqrt{5}} - \frac{\lambda_2^k}{\sqrt{5}}$$

pro každé $k > 0$ (přičemž si můžeme všimnout, že vzorec platí i pro $k = 0$).

Všimněme si také, že $|\lambda_2^k| < 1$, druhý sčítanec se proto pro $k \to \infty$ blíží 0, takže

$$a_k \approx \lambda_1^k/\sqrt{5} = \frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^k \quad .$$

9.3.2. *Geometrická násobnost, charakterizace diagonalizovatelných operátorů.* Pokud chceme operátor f na konečně generovaném prostoru dimenze n diagonalizovat, musíme najít n-prvkovou lineárně nezávislou posloupnost B složenou z vlastních vektorů operátoru f. Každý z vektorů v B musí ležet v podprostoru $\mathbf{M}_\lambda$ vlastních vektorů příslušných nějakému vlastnímu číslu λ. Z něho může báze B obsahovat nanejvýš $\dim \mathbf{M}_\lambda$ prvků. Této dimenzi říkáme geometrická násobnost vlastního čísla λ.

Definice 9.58. *Geometrickou násobností* vlastního čísla λ operátoru f na konečně generovaném prostoru (nebo čtvercové matice A) rozumíme dimenzi podprostoru $\mathbf{M}_\lambda$ vlastních vektorů operátoru f (nebo matice A) příslušných vlastnímu číslu λ.

Geometrická násobnost každého vlastního čísla λ operátoru f je aspoň 1 (jinak by λ nebylo vlastní číslo). V následujícím tvrzení dokážeme, že je menší nebo rovná algebraické násobnosti λ. K důkazu budeme potřebovat následující tvrzení o determinantech.

Tvrzení 9.59. *Pro čtvercovou blokovou matici*

$$A = \begin{pmatrix} B & C \\ 0 & D \end{pmatrix}$$

se čtvercovými diagonálními bloky B, D platí

$$\det A = (\det B)\,(\det D) \quad .$$

Důkaz. Označíme n řád matice $A = (a_{ij})$ a $k < n$ řád matice B. Důkaz uděláme indukcí podle k. Je-li $k = 1$, je $B = (a_{11})$ a všechny ostatní prvky v prvním sloupci jsou 0. Rozvineme $\det A$ podle prvního sloupce a dostaneme

$$\det A = a_{11} \det A_{11} = (\det B)\,(\det D) \quad .$$

Nyní předpokládáme, že $k > 1$. Indukční předpoklad je, že determinant blokově horní trojúhelníkové matice se dvěma diagonálními bloky se rovná součinu determinantů diagonálních bloků kdykoliv má diagonální blok vlevo nahoře řád $k-1$. Opět rozvineme $\det A$ podle prvního sloupce, tentokrát $a_{i1} = 0$ pro každé $i > k$. Dostaneme

$$\det A = (-1)^{1+1} a_{11} \det A_{11} + (-1)^{2+1} a_{21} \det A_{21} + \cdots + (-1)^{k+1} a_{k1} \det A_{k1} \quad ,$$

kde A_{i1} je minor matice A vzniklý vynecháním prvního sloupce a i-tého řádku. Pro každé $i = 1, 2, \ldots, k$ je minor

$$A_{i1} = \begin{pmatrix} B_{i1} & C_i \\ 0 & D \end{pmatrix} \quad ,$$

kde B_{i1} je minor v diagonálním bloku B vzniklý vynechnáním i-tého řádku a prvního sloupce, a C_i je matice, kterou dostaneme z bloku C vynecháním i-tého řádku. Minor A_{i1} je blokově horní trojúhelníková matice a podle indukčního předpokladu platí

$$\det A_{i1} = (\det B_{i1})\,(\det D) \quad .$$

Dostáváme tak

$$\begin{aligned}\det A =&(-1)^{1+1}a_{11}\det A_{11} + (-1)^{2+1}a_{21}\det A_{21} + \cdots + (-1)^{k+1}a_{k1}\det A_{k1}\\ =&(-1)^{1+1}a_{11}(\det B_{11})\,(\det D) + \cdots + (-1)^{k+1}a_{k1}(\det B_{k1})\,(\det D)\\ =&\big((-1)^{1+1}a_{11}(\det B_{11}) + \cdots + (-1)^{k+1}a_{k1}(\det B_{k1})\big)\,(\det D)\\ =&(\det B)\,(\det D)\ ,\end{aligned}$$

podle věty o rozvoji determinantu podle sloupce použité na diagonální blok B. □

Tvrzení 9.60. *Pro každé vlastní číslo μ lineárního operátoru $f : \mathbf{V} \to \mathbf{V}$ na konečně generovaném prostoru $\mathbf{V}$ (čtvercové matice A) nad tělesem $\mathbf{T}$ platí, že geometrická násobnost μ je menší nebo rovná algebraické násobnosti λ.*

Důkaz. Buď k geometrická násobnost vlastního čísla μ operátoru f. Zvolíme nějakou bázi $(\mathbf{v}_1, \ldots, \mathbf{v}_k)$ podprostoru $\mathbf{M}_\mu$ vlastních vektorů příslušných μ a doplníme ji vektory $\mathbf{v}_{k+1}, \ldots, \mathbf{v}_n$ na bázi $B = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$ celého prostoru $\mathbf{V}$.

Protože pro každé $i = 1, \ldots, k$ platí $[f(\mathbf{v}_i)]_B = [\mu\mathbf{v}_i]_B = \mu\mathbf{e}_i$, matice $[f]_B^B$ má blokově diagonální tvar

$$\begin{pmatrix} \mu I_k & C \\ 0 & D \end{pmatrix}\ ,$$

kde C je vhodný blok typu $k \times (n-k)$ a D vhodný čtvercový blok řádu $n-k$. Charakteristický polynom $p_f(\lambda)$ operátoru f se tedy rovná determinantu matice

$$[f]_B^B - \lambda I_n = \begin{pmatrix} (\mu-\lambda)I_k & C \\ 0 & D-\lambda I_{n-k} \end{pmatrix}\ .$$

Determinant této blokově horní trojúhelníkové matice se podle předchozího tvrzení rovná součinu determinantů diagonálních bloků. Proto

$$p_f(\lambda) = \det\left((\mu-\lambda)I_k\right)\det(D-\lambda I_{n-k}) = (\mu-\lambda)^k\det(D-\lambda I_{n-k})\ .$$

Číslo μ je tedy aspoň k-násobným kořenem charakteristického polynomu operátoru f, jeho algebraická násobnost je proto aspoň k. □

Příkladem kdy nastane rovnost, je jednotková matice I_2, která má jediné vlastní číslo 1, které má geometrickou i algebraickou násobnost 2. Dále samozřejmě rovnost algebraické a geometrické násobnosti platí pro každé vlastní číslo s algebraickou násobností 1. V jistém smyslu typický případ, kdy je nerovnost ostrá, ukazuje následující příklad.

Příklad 9.61. Reálná matice

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

má charakteristický polynom $\det A - \lambda I_2 = (\lambda - 3)^2$ a tedy jediné vlastní číslo $\lambda = 3$ algebraické násobnosti 2. Podprostor vlastních vektorů příslušných vlastnímu číslu $\lambda = 3$ je

$$M_3 = \operatorname{Ker}(A - 3I_2) = \operatorname{Ker}\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \left\langle \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\rangle\ .$$

geometrická násobnost vlastního čísla $\lambda = 3$ je proto 1.

Tato matice není diagonalizovatelná, protože z M_3 zřejmě nelze vybrat dvoučlennou lineárně nezávislou posloupnost.

Nediagonalizovatelnost operátoru nebo matice může mít dvě příčiny – nedostatek vlastních čísel (jako rotace o $\pi/2$ v příkladu 9.23) nebo nedostatek vlastních vektorů (jako v předchozím příkladu). Tím se dostáváme se k charakterizaci diagonalizovatelných operátorů.

Věta 9.62. *Buď $f : \mathbf{V} \to \mathbf{V}$ lineární operátor na konečně generovaném vektorovém prostoru $\mathbf{V}$ dimenze n (resp. buď A je čtvercová matice řádu n) nad tělesem $\mathbf{T}$. Pak jsou následující tvrzení jsou ekvivalentní.*

(1) *Operátor f je diagonalizovatelný (resp. matice A je diagonalizovatelná).*
(2) *Operátor f (resp. matice A) má*
  - *n vlastních čísel včetně algebraických násobností a*
  - *geometrická násobnost každého vlastního čísla operátoru f (resp. matice A) je rovná jeho algebraické násobnosti.*

Důkaz. (1) $\Rightarrow$ (2). Předpokládáme, že je f diagonalizovatelný. Existuje tedy báze $B = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$ prostoru $\mathbf{V}$ složená z vlastních vektorů operátoru f. Označme $\lambda_1, \ldots, \lambda_k$ všechna navzájem různá vlastní čísla operátoru $\mathbf{V}$, $l_1, \ldots, l_k$ jejich algebraické násobnosti a $m_1, \ldots, m_k$ jejich geometrické násobnosti. Každý z vektorů v B leží v jednom z podprostorů $\mathbf{M}_{\lambda_1}, \ldots, \mathbf{M}_{\lambda_k}$, přičemž z každého podprostoru $\mathbf{M}_{\lambda_i}$ může v bázi B ležet nejvýše $\dim(\mathbf{M}_{\lambda_i}) = m_i$ vektorů (protože pouze nejvýše tolik vektorů může tvořit lineárně nezávislou posloupnost v $\mathbf{M}_{\lambda_i}$). Z toho vyplývá nerovnost $n \leq m_1 + \cdots + m_k$. Podle tvrzení 9.60 je geometrická násobnost menší nebo rovná algebraické násobnosti, tedy $m_i \leq l_i$ (pro každé $i \in \{1, \ldots, k\}$). Součet algebraických násobností je přitom nejvýše n (viz první bod důsledku 9.43). Dohromady máme

$$n \leq m_1 + \cdots + m_k \leq l_1 + \cdots + l_k \leq n \quad \text{a} \quad m_i \leq l_i \quad \text{pro každé} \quad i = 1, 2, \ldots, k \ .$$

To znamená, že $l_1 + \ldots l_k = n$ a současně $m_i = l_i$ pro každé $i \in \{1, 2 \ldots, k\}$, jak jsme chtěli dokázat.

(2) $\Rightarrow$ (1). Předpokládejme naopak, že podmínky na násobnosti jsou splněné. Označíme $\lambda_1, \ldots, \lambda_k$ vlastní čísla operátoru a $l_1, \ldots, l_k$ jejich algebraické (= geometrické) násobnosti. Pro každé $i \in \{1, \ldots, k\}$ má $\mathbf{M}_{\lambda_i}$ dimenzi l_i, vezmeme jeho libovolnou bázi

$$B_i = (\mathbf{v}^i_1, \mathbf{v}^i_2, \ldots, \mathbf{v}^i_{l_i}) \ .$$

Ukážeme, že posloupnost

$$B = (\mathbf{v}^1_1, \mathbf{v}^1_2, \ldots, \mathbf{v}^1_{l_1}, \mathbf{v}^2_1, \mathbf{v}^2_2, \ldots, \mathbf{v}^2_{l_1}, \ldots, \mathbf{v}^k_1, \mathbf{v}^k_2, \ldots, \mathbf{v}^k_{l_k})$$

vytvořená ze všech vektorů všech bází podprostorů $\mathbf{M}_{\lambda_i}$ tvoří bázi prostoru $\mathbf{V}$.

Počet prvků této posloupnosti je $l_1 + l_2 + \cdots + l_k = n = \dim \mathbf{V}$, stačí proto ukázat, že B je lineárně nezávislá posloupnost. Předpokládejme tedy, že pro nějaké skaláry a^i_j, $i = 1, \ldots, k$ a $j = 1, \ldots, l_i$, platí

$$a^1_1\mathbf{v}^1_1 + a^1_2\mathbf{v}^1_2 + \cdots + a^1_{l_1}\mathbf{v}^1_{l_1} + \cdots + a^k_1\mathbf{v}^k_1 + a^k_2\mathbf{v}^k_2 + \cdots + a^k_{l_k}\mathbf{v}^k_{l_k} = \mathbf{o} \ .$$

Pro každé $i = 1, 2, \ldots, k$ je vektor

$$\mathbf{w}_i = a^i_1\mathbf{v}^i_1 + a^i_2\mathbf{v}^i_2 + \cdots + a^i_{l_i}\mathbf{v}^i_{l_i}$$

vlastní vektor operátoru f příslušný vlastnímu číslu λ_i. Dále platí

$$\mathbf{w}_1 + \mathbf{w}_2 + \cdots + \mathbf{w}_k = \mathbf{o}.$$

Pokud by některý z vektorů $\mathbf{w}_i$ byl nenulový, vynechali bychom z poslední rovnosti všechny nulové vektory na levé straně a zůstal by nám součet, tj. lineární kombinace s koeficienty 1,

$$\mathbf{w}_{i_1} + \mathbf{w}_{i_2} + \cdots + \mathbf{w}_{i_l} = \mathbf{o}$$

nenulových vlastních vektorů příslušných různým vlastním číslům $\lambda_{i_1}, \lambda_{i_2}, \ldots, \lambda_{i_l}$. To ale není možné podle věty 9.54 o lineární nezávislosti posloupnosti nenulových vlastních vektorů příslušných různým vlastním číslům.

Pro každé $i = 1, 2, \ldots, k$ tedy platí $\mathbf{w}_i = \mathbf{o}$, čili

$$\mathbf{o} = a_1^i \mathbf{v}_1^i + a_2^i \mathbf{v}_2^i + \cdots + a_{l_i}^i \mathbf{v}_{l_i}^i \quad .$$

Posloupnost vektorů $B_i = (\mathbf{v}_1^i, \ldots, \mathbf{v}_{l_i}^i)$ je ale lineárně nezávislá neboť tvoří bázi $\mathbf{M}_{\lambda_i}$. Dostáváme tak, že $a_1^i = a_2^i = \cdots = a_{l_i}^i = 0$ pro každé $i = 1, 2, \ldots, k$. Posloupnost B je tedy lineárně nezávislá a tvoří proto bázi prostoru $\mathbf{V}$ složenou z vlastních vektorů operátoru f. □

Z důkazu vidíme, že v případě diagonalizovatelných operátorů bude v nalezené bázi B počet vlastních vektorů příslušných vlastnímu číslu λ roven algebraické (= geometrické) násobnosti λ. Proto $[f]_B^B$ bude mít na diagonále každé vlastní číslo tolikrát, kolik je jeho algebraická násobnost.

Příklad 9.63. V příkladu 9.42 jsme spočítali vlastní čísla a jejich algebraické násobnosti pro operátor $f : \mathbb{R}^3 \to \mathbb{R}^3$ definovaný předpisem

$$f \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -y + z \\ -3x - 2y + 3z \\ -2x - 2y + 3z \end{pmatrix} \quad .$$

Zjistíme je-li diagonalizovatelný, a pokud ano, najdeme bázi v $\mathbb{R}^3$ složenou z vlastních vektorů operátoru f.

Operátor f je roven f_A pro matici

$$A = [f]_K^K = \begin{pmatrix} 0 & -1 & 1 \\ -3 & -2 & 3 \\ -2 & -2 & 3 \end{pmatrix} \quad .$$

Charakteristický polynom operátoru f vyšel $-(\lambda - 1)^2(\lambda + 1)$, takže operátor f má vlastní číslo $\lambda_1 = 1$ algebraické násobnosti 1 a vlastní číslo $\lambda_2 = -1$ algebraické násobnosti 2. Splňuje tedy první podmínku pro diagonalizovatelnost. Zbývá ověřit rovnost algebraické a geometrické násobnosti obou vlastních čísel $\lambda_1 = 1$ a $\lambda_2 = -1$.

Algebraická násobnost vlastního čísla $\lambda_1 = 1$ je 2. Jeho geometrická násobnost se rovná dimenzi jádra matice

$$A - \lambda_1 I_3 = A - I_3 = \begin{pmatrix} -1 & -1 & 1 \\ -3 & -3 & 3 \\ -2 & -2 & 2 \end{pmatrix} \quad .$$

Hodnost této matice se rovná 1, dimenze jádra je proto 2. Geometrická násobnost vlastního čísla $\lambda_1 = 1$ je rovná jeho algebraické násobnosti.

Algebraická násobnost vlastního čísla $\lambda_2 = -1$ je 1 a rovná se tak jeho geometrické násobnosti, protože ta je aspoň 1 pro jakékoliv vlastní číslo. Operátor f je tedy diagonalizovatelný.

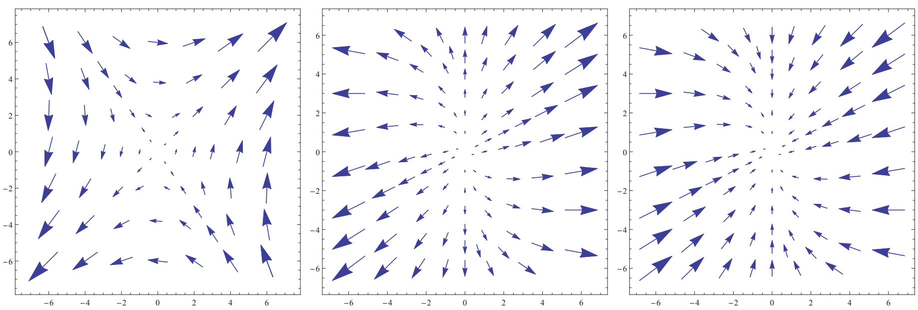
Najdeme ještě bázi $\mathbb{R}^3$, vzhledem ke které je matice operátoru f diagonální. Bázi jádra matice $A - \lambda_1 I_3 = A - I_3$, které má dimenzi 2, můžeme zvolit například $(1,0,1)^T$, $(0,1,1)^T$. Bázi jádra matice

$$A - \lambda_2 I_3 = A + I_3 = \begin{pmatrix} 1 & -1 & 1 \\ -3 & -1 & 3 \\ -2 & -2 & 4 \end{pmatrix} \sim \begin{pmatrix} 1 & -1 & 1 \\ 0 & -4 & 6 \end{pmatrix}$$

můžeme zvolit například $(1,3,2)^T$. Posloupnost $B = ((1,0,1)^T, (0,1,1)^T, (1,3,2)^T)$ tak tvoří bázi prostoru $\mathbb{R}^3$ tvořenou vlastními vektory operátoru f a platí $[f]_B^B = \operatorname{diag}(1,1,-1)$.

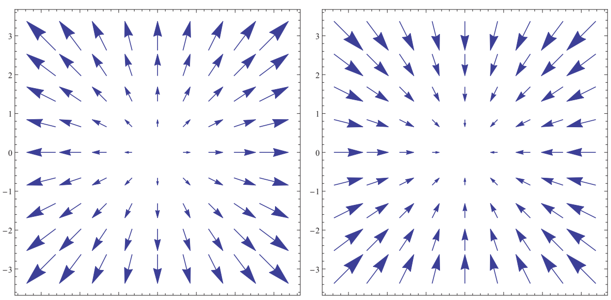
9.3.3. *Diagonalizovatelné lineární operátory na reálném vektorovém prostoru dimenze* 2. Probereme možnosti, které mohou nastat pro diagonalizovatelný lineární operátor f na vektorovém prostoru $\mathbf{V}$ dimenze 2 nad $\mathbb{R}$.

• Operátor f má dvě různá vlastní čísla λ_1, λ_2. Pak je diagonalizovatelný. Pro představu o vývoji diskrétního dynamického systému $\mathbf{x}_k = f(\mathbf{x}_{k-1})$ se podívejte na následující obrázek. Stejně jako v příkladu 9.6 vedou šipky z bodu $\mathbf{x}$ do bodu $f(\mathbf{x})$.



Obrázek 76. Dvě různá vlastní čísla. Vlevo $\lambda_1 > 1 > \lambda_2 > 0$, uprostřed $\lambda_1, \lambda_2 > 1$, vpravo $0 < \lambda_1, \lambda_2 < 1$

• Operátor f má jedno vlastní číslo algebraické násobnosti 2 a geometrické násobnosti 2. Pak je diagonalizovatelný, dokonce $\mathbf{M}_\lambda$ se rovná $\mathbf{V}$, tj. $f(\mathbf{x}) = \lambda\mathbf{x}$ pro každé $\mathbf{x}$. Vývoj diskrétního dynamického systému $\mathbf{x}_k = f(\mathbf{x}_{k-1})$ si můžeme představit pomocí dalšího obrázku.



Obrázek 77. Jedno vlastní číslo algebraické násobnosti 2 a geometrické násobnosti 2. Vlevo $\lambda > 1$, vpravo $1 > \lambda > 0$.

9.3.4. *Operátory na prostorech nad $\mathbb{R}$, které jsou „diagonalizovatelné nad $\mathbb{C}$“.* Rozebereme situaci, kdy reálná matice A řádu n není diagonalizovatelná, ale stejná matice, chápaná jako matice nad $\mathbb{C}$, již diagonalizovatelná je. (Diskuzi budeme provádět pouze pro matice (nebo operátory f_A), abychom nemuseli používat pojem komplexního rozšíření reálného lineárního prostoru a operátoru na něm.) Charakteristický polynom $p_A(\lambda)$ má v tom případě reálné koeficienty a spolu s každým komplexním vlastním číslem λ má matice A také komplexně sdružené vlastní číslo $\overline{\lambda}$.

Následující jednoduché tvrzení ukazuje, že také komplexní vlastní vektory matice A můžeme sdružit do párů. Pro komplexní matici $A = (a_{ij})$ označíme $\overline{A} = (\overline{a_{ij}})$ matici, ve které každý prvek a_{ij} matice A nahradíme číslem $\overline{a_{ij}}$ komplexně sdruženým k a_{ij}. Matice A má všechny prvky reálné právě když $\overline{A} = A$.

Tvrzení 9.64. *Je-li $\mathbf{x} \in \mathbb{C}^n$ vlastní vektor reálné matice $A = (a_{ij})$ příslušný vlastnímu číslu λ, pak $\overline{\mathbf{x}}$ je vlastní vektor matice A příslušný vlastnímu číslu $\overline{\lambda}$.*

Důkaz. Je-li $\mathbf{x} = (x_1, x_2, \ldots, x_n)^T$ vlastní vektor reálné matice A příslušný vlastnímu číslu λ, platí $A\mathbf{x} = \lambda\mathbf{x}$. To znamená, že pro každé $i = 1, 2, \ldots, n$ platí

$$a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n = \lambda x_i \ .$$

Rovnají se tedy také čísla komplexně sdružená k oběma stranám, tj.

$$\overline{a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n} = \overline{\lambda x_i}$$

a tedy také

$$\overline{a_{i1}}\ \overline{x_1} + \overline{a_{i2}}\ \overline{x_2} + \cdots + \overline{a_{in}}\ \overline{x_n} = \overline{\lambda}\ \overline{x_i} \ .$$

Protože A je reálná matice, platí $\overline{a_{ij}} = a_{ij}$ pro každé $i, j = 1, 2, \ldots, n$ a tedy

$$a_{i1}\overline{x_1} + a_{i2}\overline{x_2} + \cdots + a_{in}\overline{x_n} = \overline{\lambda}\overline{x_i} \ .$$

pro každé $i = 1, 2, \ldots, n$, což dokazuje rovnost $A\overline{\mathbf{x}} = \overline{\lambda}\overline{\mathbf{x}}$. □

Začneme tím, že na konkrétním příkladu ukážeme přímočarý, ale nepříliš efektivní postup pro mocnění reálné matice, která nemá žádné reálné vlastní číslo.

Příklad 9.65. Najdeme vzorec pro k-tou mocninu reálné matice

$$A = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \ .$$

Charakteristický polynom matice A je

$$p_A(\lambda) = \lambda^2 - 2\lambda + 2 \ .$$

Tento polynom nemá reálné kořeny. Budeme proto považovat A za matici nad $\mathbb{C}$. Nyní má $p_A(\lambda)$ dva komplexně sdružené kořeny $\lambda = 1 + i$, $\overline{\lambda} = 1 - i$, matice je tedy nad $\mathbb{C}$ diagonalizovatelná.

Vlastní vektory příslušné vlastnímu číslu λ tvoří podprostor

$$M_{1+i} = \mathrm{Ker}\,(A - (1+i)I_2) = \mathrm{Ker} \begin{pmatrix} -i & -1 \\ 1 & -i \end{pmatrix} = \left\langle \begin{pmatrix} 1 \\ -i \end{pmatrix} \right\rangle \ .$$

Protože A je reálná matice, platí podle předchozího tvrzení 9.64, že $\mathbf{v}$ je vlastní vektor matice A příslušný komplexnímu λ právě tehdy, když je $\overline{\mathbf{v}}$ vlastní vektor matice A příslušný vlastnímu číslu $\overline{\lambda}$. Proto bez počítání víme, že

$$M_{1-i} = \left\langle \begin{pmatrix} 1 \\ i \end{pmatrix} \right\rangle \ .$$

Bázi v $\mathbb{C}^2$ z vlastních vektorů matice A zvolíme například

$$B = \left(\begin{pmatrix} 1 \\ -i \end{pmatrix}, \begin{pmatrix} 1 \\ i \end{pmatrix} \right) .$$

Pak

$$\begin{aligned} A^k =& [(f_A)^k]_K^K = [\mathrm{id}]_K^B \, [(f_A)^k]_B^B \, [\mathrm{id}]_B^K \\ =& \begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix} \begin{pmatrix} (1+i)^k & 0 \\ 0 & (1-i)^k \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix}^{-1} . \end{aligned}$$

Vyjde

$$A^k = \frac{1}{2} \begin{pmatrix} (1+i)^k + (1-i)^k & i(1+i)^k - i(1-i)^k \\ -i(1+i)^k + i(1-i)^k & (1+i)^k + (1-i)^k \end{pmatrix} .$$

To je vcelku komplikovaný výraz, navíc obsahuje imaginární čísla, i když výsledek musí zřejmě být reálná matice.

Proto je v tomto případě lepší počítat s goniometrickým tvarem komplexních čísel. Je $1 + i = \sqrt{2}e^{i\pi/4} = \sqrt{2}(\cos(\pi/4) + i\sin(\pi/4))$ a $1 - i = \sqrt{2}e^{-i\pi/4} = \sqrt{2}(\cos(\pi/4) - i\sin(\pi/4))$. Dosazením a využitím Moivreovy věty vyjde daleko přijatelnější výsledek

$$\begin{aligned} A^k =& \begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix} \begin{pmatrix} (1+i)^k & 0 \\ 0 & (1-i)^k \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix}^{-1} \\ =& \begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix} \begin{pmatrix} (\sqrt{2})^k e^{ik\pi/4} & 0 \\ 0 & (\sqrt{2})^k e^{-ik\pi/4} \end{pmatrix} \frac{1}{2i} \begin{pmatrix} i & -1 \\ i & -i \end{pmatrix} \\ =& \sqrt{2}^k \begin{pmatrix} \cos(k\pi/4) & -\sin(k\pi/4) \\ \sin(k\pi/4) & \cos(k\pi/4) \end{pmatrix} . \end{aligned}$$

Stejného výsledku bychom docílili, kdybychom si hned na začátku všimli, že

$$A = \sqrt{2} \begin{pmatrix} \cos(\pi/4) & -\sin(\pi/4) \\ \sin(\pi/4) & \cos(\pi/4) \end{pmatrix} .$$

Tedy že A je matice složení rotace o $\pi/4$ a stejnolehlosti s koeficientem $\sqrt{2}$. Mocninu A^k pak vidíme geometricky okamžitě.

Nyní budeme uvažovat obecnou reálnou matici A řádu 2, která nemá řádné reálné vlastní číslo. Charakteristický polynom matice A má tedy dva různé komplexně sdružené kořeny λ, $\overline{\lambda}$, takže operátor f_A má dvě různá komplexní vlastní čísla λ, $\overline{\lambda}$. Jako operátor $f_A : \mathbb{R}^2 \to \mathbb{R}^2$ diagonalizovatelný není (nemá žádné reálné vlastní číslo), jako operátor $f_A : \mathbb{C}^2 \to \mathbb{C}^2$ diagonalizovatelný je (dvě různá vlastní čísla). Vlastní čísla si vyjádříme v goniometrickém tvaru:

$$\lambda = re^{i\varphi} = r(\cos\varphi + i\sin\varphi), \quad \overline{\lambda} = re^{-i\varphi} = r(\cos\varphi - i\sin\varphi) .$$

Ukážeme, že existuje báze B prostoru $\mathbb{R}^2$ (!!) taková, že

$$[f_A]_B^B = r \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix} ,$$

tj. „vzhledem k bázi B“ se operátor $f_A : \mathbb{R}^2 \to \mathbb{R}^2$ rovná rotaci o φ složené se stejnolehlostí s koeficientem r. (Ve skutečnosti můžeme dokonce nalézt ortogonální bázi B, vzhledem ke které má operátor uvedenou matici, nemůžeme ale obecně požadovat ortonormální bázi. Výpočet je složitější, nebudeme jej provádět nyní, ale až v příští kapitole.)

Označme $\mathbf{v} \in \mathbb{C}^2$ nějaký nenulový vlastní vektor operátoru $f_A : \mathbb{C}^2 \to \mathbb{C}^2$ příslušný vlastnímu číslu λ. Podle tvrzení 9.64 víme, že $\overline{\mathbf{v}}$ je vlastní vektor operátoru $f_A : \mathbb{C}^2 \to \mathbb{C}^2$ příslušný vlastnímu číslu $\overline{\lambda}$.

Posloupnost vektorů $C = (\mathbf{v}, \overline{\mathbf{v}})$ je báze $\mathbb{C}^2$ a vzhledem k bázi C máme

$$[f_A]_C^C = \mathrm{diag}(\lambda, \overline{\lambda}) \ .$$

Označíme

$$\mathbf{w}_1 = \mathbf{v} + \overline{\mathbf{v}}, \quad \mathbf{w}_2 = i(\mathbf{v} - \overline{\mathbf{v}}) \ .$$

Vektory $\mathbf{w}_1, \mathbf{w}_2$ jsou reálné vektory, vektor $\mathbf{w}_1$ je dvojnásobek reálné části vektoru $\mathbf{v}$, vektor $\mathbf{w}_2$ je (-2)-násobek jeho imaginární části. Z definice vektorů $\mathbf{w}_1, \mathbf{w}_2$ plyne $[\mathbf{w}_1]_C = (1,1)^T$, $[\mathbf{w}_2]_C = (i,-i)^T$. Z toho vidíme, že $B = (\mathbf{w}_1, \mathbf{w}_2)$ je lineárně nezávislá posloupnost (v $\mathbb{C}^2$ i v $\mathbb{R}^2$) a tvoří tedy bázi prostoru $\mathbb{C}^2$ i $\mathbb{R}^2$. Vzhledem k této bázi máme

$$\begin{aligned}[f_A]_B^B &= [\mathrm{id}]_B^C \, [f_A]_C^C \, [\mathrm{id}]_C^B \\ &= \begin{pmatrix} 1 & i \\ 1 & -i \end{pmatrix}^{-1} \begin{pmatrix} r(\cos\varphi + i\sin\varphi) & 0 \\ 0 & r(\cos\varphi - i\sin\varphi) \end{pmatrix} \begin{pmatrix} 1 & i \\ 1 & -i \end{pmatrix}\end{aligned}$$

Výpočtem získáme

$$[f]_B^B = r \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix} \ ,$$

což jsme chtěli ukázat. Dokázali jsme tak následující tvrzení.

Tvrzení 9.66. *Je-li A reálná matice řádu 2, která nemá reálná vlastní čísla, a $\lambda = r(\cos\varphi + i\sin\varphi)$ je komplexní vlastní číslo s nenulovým vlastním vektorem $\mathbf{v}$, pak platí*

(1) *vektory $\mathbf{w}_1 = \mathbf{v} + \overline{\mathbf{v}}$ a $\mathbf{w}_2 = i(\mathbf{v} - \overline{\mathbf{v}})$ tvoří bázi $B = (\mathbf{w}_1, \mathbf{w}_2)$ prostoru $\mathbb{R}^2$,*
(2) *lineární zobrazení $f_A : \mathbb{R}^2 \to \mathbb{R}^2$ určené maticí A má vzhledem k bázi B matici*

$$[f_A]_B^B = r \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}$$

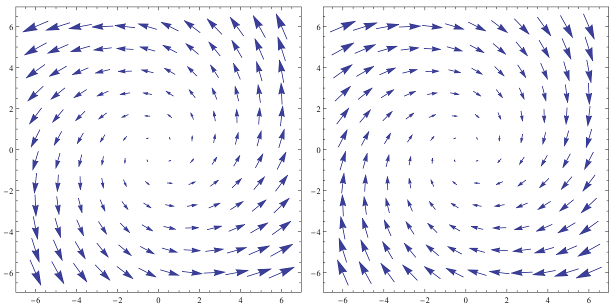
a je tedy složením rotace o úhel φ se stejnolehlostí s koeficientem $r > 0$

Vývoj diskrétního reálného dynamického systému $\mathbf{x}_{k+1} = f_A(\mathbf{x}_k)$ pak můžeme také intuitivně nahlédnout geometricky.

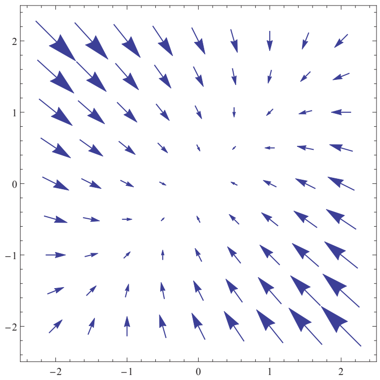
• Jestliže reálná matice A řádu 2 nemá žádné reálné vlastní číslo, pak ani operátor f_A nemá žádná reálná vlastní čísla. Není proto diagonalizovatelný nad $\mathbb{R}$. Operátor f_A má dvě různá komplexně sdružená vlastní čísla a vývoj dynamického systému $\mathbf{x}_{k+1} = f_A(\mathbf{x}_k)$ „vzhledem k bázi B“ můžeme intuitivně nahlédnout pomocí následujícího obrázku.

9.3.5. *Vývoj reálného spojitého dynamického systému s diagonalizovatelnou maticí.* Uvažujme spojitý dynamický systém $\mathbf{x}'(t) = A\mathbf{x}(t)$, kde A je reálná matice řádu n. Vektor $\mathbf{x}(t) \in \mathbb{R}^n$ může například udávat polohu pohybujícího se objektu v čase t a rovnice $\mathbf{x}'(t) = A\mathbf{x}(t)$ pak říká, že vektor rychlosti tohoto objektu je v bodě $\mathbf{x}(t)$ roven $A\mathbf{x}(t)$. Dobrou představu o řešení si můžeme udělat tak, že si do několika bodů $\mathbf{x}$ nakreslíme šipku z bodu $\mathbf{x}$ do bodu $\mathbf{x} + A\mathbf{x}$, neboli vektor $A\mathbf{x}$ s počátečním bodem $\mathbf{x}$, jako na následujícím obrázku.

Pokud je matice A diagonalizovatelná, můžeme soustavu vyřešit následujícím způsobem. V prostoru $\mathbb{R}^n$ existuje báze $B = (\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n)$ tvořená vlastními



Obrázek 78. Dvě komplexně sdružená vlastní čísla. Vlevo $|\lambda| > 1$, vpravo $1 > |\lambda|$.



Obrázek 79. K příkladu 9.67. Šipky jsou pro přehlednost zkráceny.

.

vektory matice A, přičemž vektor $\mathbf{u}_i$ je příslušný vlastnímu číslu λ_i matice A. Matice $R = (\mathbf{u}_1|\mathbf{u}_2|\cdots|\mathbf{u}_n)$ je regulární a rovná se matici přechodu $[\mathrm{id}]^B_K$ od báze B ke kanonické bázi K. Platí pro ni

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

Úpravou dostaneme $A = RDR^{-1}$ a dosadíme do rovnice $\mathbf{x}'(t) = A\mathbf{x}(t)$. Dostaneme tak

$$\begin{aligned} \mathbf{x}'(t) &= RDR^{-1}\mathbf{x}(t) \\ R^{-1}\mathbf{x}'(t) &= DR^{-1}\mathbf{x}(t) \end{aligned}$$

Položíme $\mathbf{y}(t) = R^{-1}\mathbf{x}(t)$. Snadno se pak ověří, že $\mathbf{y}'(t) = R^{-1}\mathbf{x}'(t)$. Stačí si uvědomit, že každá složka $y_j(t)$ vektoru $\mathbf{y}(t)$ je lineární kombinací funkcí $x_i(t)$ a tedy derivace $y_j'(t)$ je lineární kombinací derivací $x_i'(t)$ se stejnými koeficienty (podrobněji je tento krok vysvětlen v příkladu). Dosazením dostáváme

$$\mathbf{y}'(t) = D\mathbf{y}(t)$$

Tuto soustavu vyřešíme užitím příkladu 9.15 o řešení diferenciální rovnice $f'(t) = \lambda f(t)$. Původní funkce $\mathbf{x}(t)$ pak dopočteme ze vztahu $\mathbf{x}(t) = R\mathbf{y}(t)$.

Příklad 9.67. Vyřešíme soustavu

$$\begin{pmatrix} x_1'(t) \\ x_2'(t) \end{pmatrix} = A \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} = \begin{pmatrix} -2 & 1 \\ 1 & -2 \end{pmatrix} \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix}$$

s počáteční podmínkou $x_1(0) = 5, x_2(0) = 7$. Charakteristický polynom matice A je $p_A(\lambda) = \lambda^2 + 4\lambda + 3$, který má dva různé reálné kořeny $\lambda_1 = -1$ a $\lambda_2 = -3$. Vlastní vektory příslušné vlastnímu číslu $\lambda_1 = -1$ tvoří lineární obal $\langle (1,1)^T \rangle$. Vlastní vektory příslušné vlastnímu číslu $\lambda_2 = -3$ tvoří linerání obal $\langle (1,-1)^T \rangle$. Položíme $B = ((1,1)^T, (1,-1)^T)$. Pak

$$R = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} .$$

a

$$D = R^{-1}AR = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} .$$

Původní soustavu si přepíšeme do tvaru

$$R^{-1} \begin{pmatrix} x_1'(t) \\ x_2'(t) \end{pmatrix} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} R^{-1} \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} .$$

Označíme

$$\begin{pmatrix} y_1(t) \\ y_2(t) \end{pmatrix} = R^{-1} \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} .$$

Obě funkce $y_i(t)$ jsou lineární kombinace funkcí $x_1(t), x_2(t)$ s konstantními koeficienty v i-tém řádku matice R^{-1}. Platí proto

$$\begin{pmatrix} y_1'(t) \\ y_2'(t) \end{pmatrix} = R^{-1} \begin{pmatrix} x_1'(t) \\ x_2'(t) \end{pmatrix} .$$

Dvojice funkcí $y_1(t), y_2(t)$ tak splňuje soustavu lineárních diferenciálních rovnic

$$\begin{pmatrix} y_1'(t) \\ y_2'(t) \end{pmatrix} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \begin{pmatrix} y_1(t) \\ y_2(t) \end{pmatrix} = \begin{pmatrix} \lambda_1 y_1(t) \\ \lambda_2 y_2(t) \end{pmatrix} .$$

Tu už umíme řešit: $y_1(t) = y_1(0)e^{-2t}$ a $y_2(t) = y_2(0)e^{-3t}$, kde

$$\begin{pmatrix} y_1(0) \\ y_2(0) \end{pmatrix} = R^{-1} \begin{pmatrix} x_1(0) \\ x_2(0) \end{pmatrix} .$$

Spočítáme původní funkce:

$$\begin{aligned} \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} &= R \begin{pmatrix} y_1(t) \\ y_2(t) \end{pmatrix} = R \begin{pmatrix} y_1(0)e^{\lambda_1 t} \\ y_2(0)e^{\lambda_2 t} \end{pmatrix} = R \begin{pmatrix} e^{\lambda_1 t} & 0 \\ 0 & e^{\lambda_2 t} \end{pmatrix} \begin{pmatrix} y_1(0) \\ y_2(0) \end{pmatrix} \\ &= R \begin{pmatrix} e^{\lambda_1 t} & 0 \\ 0 & e^{\lambda_2 t} \end{pmatrix} R^{-1} \begin{pmatrix} x_1(0) \\ x_2(0) \end{pmatrix} \\ &= \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} e^{\lambda_1 t} & 0 \\ 0 & e^{\lambda_2 t} \end{pmatrix} \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 5 \\ 7 \end{pmatrix} \\ &= \begin{pmatrix} 6e^{-t} - e^{-3t} \\ 6e^{-t} + e^{-3t} \end{pmatrix} . \end{aligned}$$

Z postupu při řešení předchozího příkladu můžeme nahlédnout, že pro obecnou diagonalizovatelnou matici A můžeme řešení spojitého dynamického systému $\mathbf{x}'(t) = A\mathbf{x}(t)$ vyjádřit ve tvaru řešením soustavy

$$\mathbf{x}(t) = R\,\mathrm{diag}(e^{\lambda_1 t}, \dots, e^{\lambda_n t})R^{-1}\mathbf{x}(0) ,$$

kde $A = R\,\mathrm{diag}(\lambda_1, \dots, \lambda_n)\, R^{-1}$.

Pro zajímavost uveďme, že matice $R\,\mathrm{diag}(e^{\lambda_1 t}, \dots, e^{\lambda_n t}) R^{-1}$ je rovna tzv. exponenciále matice tA a značí se e^{tA}. To dává smysl, protože „dosadíme-li“ matici tA do Taylorovy řady funkce e^x máme

$$
\begin{aligned}
& I_n + tA + t^2\frac{A^2}{2!} + t^3\frac{A^3}{3!} + \cdots + t^k\frac{A^k}{k!} + \cdots \\
= \;& I_n + tRDR^{-1} + t^2\frac{RD^2R^{-1}}{2!} + t^3\frac{RD^3R^{-1}}{3!} + \cdots + t^k\frac{RD^kR^{-1}}{k!} + \cdots \\
= \;& R\left(I_n + tD + \frac{(tD)^2}{2!} + \frac{(tD)^3}{3!} + \cdots + \frac{(tD)^k}{k!} + \cdots \right) R^{-1} \\
= \;& R\,\mathrm{diag}(e^{\lambda_1 t}, \dots, e^{\lambda_n t}) R^{-1} \;.
\end{aligned}
$$

Řešení soustavy $\mathbf{x}'(t) = A\mathbf{x}(t)$ pak můžeme psát ve tvaru

$$\mathbf{x}(t) = e^{tA}\mathbf{x}(0) \;,$$

zcela analogicky k vyjádření $f(t) = e^{t\lambda} f(0)$ jako řešení diferenciální rovnice $f' = \lambda f$.

Příklad 9.68. Vyřešíme ještě spojitý dynamický systém z části 9.1.7 popisující přechod substance přes buněčnou blánu. Ten vede k soustavě

$$
\begin{aligned}
u_1'(t) &= -ru_1(t) + su_2(t) \;, \\
u_2'(t) &= ru_1(t) - su_2(t) \;.
\end{aligned}
$$

s počáteční podmínkou $u_1(0) = 1$, $u_2(0) = 0$. Matice soustavy

$$A = \begin{pmatrix} -r & s \\ r & -s \end{pmatrix}$$

má charakteristický polynom $p(\lambda) = \lambda^2 + (r+s)\lambda$ a tudíž dvě různá vlastní čísla $\lambda_1 = 0$ a $\lambda_2 = -(r+s)$, odtud plyne její diagonalizovatelnost.

Vlastní vektory příslušné vlastnímu číslu $\lambda_1 = 0$ tvoří jádro matice

$$A - 0I_2 = \begin{pmatrix} -r & s \\ r & -s \end{pmatrix} \;,$$

který se rovná lineárnímu obalu $\langle (s,r)^T \rangle$.

Vlastní vektory příslušné vlastnímu číslu $\lambda_2 = -(r+s)$ tvoří jádro matice

$$A + (r+s)I_2 = \begin{pmatrix} s & s \\ r & r \end{pmatrix} \;,$$

který se rovná lineárnímu obalu $\langle (1,-1)^T \rangle$. Zvolíme bázi $B = ((s,r)^T, (1,-1)^T))$. Pak

$$R = [\mathrm{id}]^B_K = \begin{pmatrix} s & 1 \\ r & -1 \end{pmatrix}, \quad R^{-1} = \frac{-1}{r+s}\begin{pmatrix} -1 & -1 \\ -r & s \end{pmatrix} \;.$$

Soustava má řešení

$$
\begin{aligned}
\begin{pmatrix} u_1(t) \\ u_2(t) \end{pmatrix} &= \begin{pmatrix} s & 1 \\ r & -1 \end{pmatrix}\begin{pmatrix} e^{0t} & 0 \\ 0 & e^{-(r+s)t} \end{pmatrix}\frac{-1}{r+s}\begin{pmatrix} -1 & -1 \\ -r & s \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} \\
&= \frac{1}{r+s}\begin{pmatrix} s + re^{-(r+s)t} \\ r - re^{-(r+s)t} \end{pmatrix} \;.
\end{aligned}
$$

Vidíme, že pro $t \to \infty$ hodnota $u_1(t)$ konverguje k $\frac{s}{r+s}$ a hodnota $u_2(t)$ konverguje k $\frac{r}{r+s}$.

V následujícím příkladu si ukážeme, jak se vyvíjí spojitý dynamický systém $\mathbf{x}'(t) = A\mathbf{x}(t)$ v dimenzi 2 v případě, že reálná matice A nemá žádné reálné vlastní číslo.

Příklad 9.69. Spočítáme vlastní kmity pružiny se závažím. V části 9.1.6 jsme odvodili, že vlastní kmity jsou popsané spojitým dynamickým systémem

$$\mathbf{x}'(t) = \begin{pmatrix} x_1'(t) \\ x_2'(t) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -\frac{k}{m} & 0 \end{pmatrix} \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -\frac{k}{m} & 0 \end{pmatrix} \mathbf{x}(t) \ ,$$

kde k označuje koeficient pružnosti pružiny a m hmotnost závaží. Proto $k/m > 0$ existuje jednoznačně určené reálné číslo $\omega > 0$, pro které platí $k/m = \omega^2$. Řešíme tedy spojitý dynamický systém s maticí

$$A = \begin{pmatrix} 0 & 1 \\ -\omega^2 & 0 \end{pmatrix}$$

s libovolným počátečním stavem $\mathbf{x}(0) = (p_1, p_2)^T$. Charakteristický polynom matice A je $p_A\lambda = \lambda^2 + \omega^2$ a má imaginární kořeny $\lambda_{1,2} = \pm\, i\,\omega$. Matice A je tedy diagonalizovatelná pouze nad $\mathbb{C}$, nikoliv nad $\mathbb{R}$. Množina všech vlastních vektorů matice A příslušných vlastnímu číslu $\lambda = i\omega$ se rovná

$$M_{i\omega} = \mathrm{Ker} \begin{pmatrix} -i\omega & 1 \\ -\omega^2 & -i\omega \end{pmatrix} = \left\langle \begin{pmatrix} 1 \\ i\omega \end{pmatrix} \right\rangle$$

a podle tvrzení 9.64 je

$$M_{-i\omega} = \left\langle \begin{pmatrix} 1 \\ -i\omega \end{pmatrix} \right\rangle \ .$$

Pro matici

$$R = \begin{pmatrix} 1\omega & 1 \\ i\omega & -i\omega \end{pmatrix}$$

pak platí

$$R^{-1}AR = \begin{pmatrix} i\omega & 0 \\ 0 & -i\omega \end{pmatrix} \ .$$

Dynamický systém $\mathbf{x}'(t) = A\mathbf{x}(t)$ upravíme do tvaru

$$R^{-1}\mathbf{x}'(t) = \begin{pmatrix} i\omega & 0 \\ 0 & -i\omega \end{pmatrix} R^{-1}\mathbf{x}(t) \ .$$

a vyřešíme pro nový neznámý vektor $\mathbf{y}(t) = (y_1(t), y_2(t))^T = R^{-1}\mathbf{x}(t)$ s novým počátečním stavem $\mathbf{y}_0 = (y_1(0), y_2(0))^T = R^{-1}\mathbf{x}_0$. Spojitý dynamický systém

$$\begin{pmatrix} y_1'(t) \\ y_2'(t) \end{pmatrix} = \begin{pmatrix} i\omega & 0 \\ 0 & -i\omega \end{pmatrix} \begin{pmatrix} y_1(t) \\ y_2(t) \end{pmatrix}$$

se rozpadá na dva komplexní spojité dynamické systémy

$$y_1'(t) = (i\omega)y_1 t$$
$$y_2'(t) = (-i\omega)y_2 t$$

v dimenzi 1 a v závěru úvodní motivační části 9.1 jsme si ukázali, že má řešení

$$y_1(t) = y_1(0)e^{i\omega t}$$
$$y_2(t) = y_2(0)e^{-i\omega t} \ ,$$

což můžeme přepsat do tvaru

$$\begin{pmatrix} y_1(t) \\ y_2(t) \end{pmatrix} = \begin{pmatrix} e^{i\omega t} & 0 \\ 0 & e^{-i\omega t} \end{pmatrix} \begin{pmatrix} y_1(0) \\ y_2(0) \end{pmatrix} .$$

Spočteme ještě matici

$$R^{-1} = \frac{1}{2\omega} \begin{pmatrix} \omega & -i \\ \omega & i \end{pmatrix}$$

a vyjádříme řešení původního spojitého dynamického systému $\mathbf{x}'(t) = A\mathbf{x}(t)$ s počáteční podmínkou $\mathbf{x}(0) = (p_1, p_2)^T$ ve tvaru

$$\mathbf{x}(t) = R\mathbf{y}(t) = \begin{pmatrix} 1\omega & 1 \\ i\omega & -i\omega \end{pmatrix} \begin{pmatrix} e^{i\omega t} & 0 \\ 0 & e^{-i\omega t} \end{pmatrix} \frac{1}{2\omega} \begin{pmatrix} \omega & -i \\ \omega & i \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \end{pmatrix} .$$

Po delším počítání s využitím Eulerovy formule $e^{i\omega t} = \cos(\omega t) + i\sin(\omega t)$ dostaneme výsledek

$$\begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} = \begin{pmatrix} p_1 \cos(\omega t) + (1/\omega) p_2 \sin(\omega t) \\ -\omega p_1 \sin(\omega t) + p_2 \cos(\omega t) \end{pmatrix} .$$

Abychom pochopili trajektorii bodu $(x_1(t), x_2(t))^T$ v rovině $\mathbb{R}^2$, vynásobíme první složku $x_1(t)$ číslem ω a spočteme součet čtverců

$$\begin{aligned}
(\omega x_1(t))^2 + (x_2(t))^2 &= (\omega p_1 \cos(\omega t) + p_2 \sin(\omega t))^2 + (-\omega p_1 \sin(\omega t) + p_2 \cos(\omega t))^2 \\
&= \omega^2 p_1^2 \cos^2(\omega t) + p_2^2 \sin^2(\omega t) + 2\omega p_1 p_2 \cos(\omega t) \sin(\omega t) \\
&+ \omega^2 p_1^2 \sin^2(\omega t) + p_2^2 \cos^2(\omega t) - 2\omega p_1 p_2 \cos(\omega t) \sin(\omega t) \\
&= \omega^2 p_1^2 + p_2^2 \ .
\end{aligned}$$

Vidíme tedy, že bod $(x_1(t), x_2(t))^T$ se pohybuje po elipse se středem v počátku a poloosami délky $\sqrt{\omega^2 p_1^2 + p_2^2}/\omega$ ve směru osy x_1 a délky $\sqrt{\omega^2 p_1^2 + p_2^2}$ ve směru osy x_2.

Ve speciálním případě, kdy $\omega = 1$, se bod $(x_1, x_2)^T$ pohybuje po kružnici se středem v počátku o poloměru $\sqrt{p_1^2 + p_2^2}$ s konstantní úhlovou rychlostí $\omega = -1$, tj. po směru hodinových ručiček. Směr pohybu a fázový posun zjistíme nejsnáze pomocí komplexních čísel. Použijeme goniometrický tvar $p_1 + ip_2 = re^{i\varphi}$ pro počáteční podmínku a spočteme

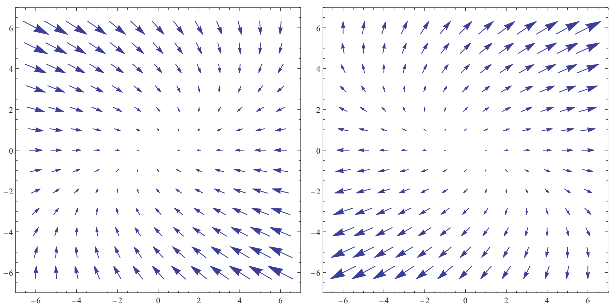
$$\begin{aligned}
x_1(t) + ix_2(t) &= p_1 \cos t + p_2 \sin t - ip_1 \sin t + ip_2 \cos t = (p_1 + ip_2)(\cos t - i \sin t) \\
&= r\, e^{i\varphi} e^{-it} = r\, e^{i(-t+\varphi)} \ .
\end{aligned}$$

Spočítali jsme tak, že fázový posun se rovná φ.

9.4. **Jordanův kanonický tvar.** Jak jsme dokázali ve větě 9.62, lineární operátor na prostoru dimenze n může být nediagonalizovatelný ze dvou důvodů – součet algebraických násobností vlastních čísel je menší než n nebo geometrická násobnost nějakého vlastního čísla je menší než jeho algebraická násobnost. První příčinu lze obejít tím, že pracujeme ve větším tělese (například místo $\mathbb{R}$ v $\mathbb{C}$). Druhá příčina takto obejít nejde, musíme slevit z požadavku diagonalizovatelnosti. Naštěstí lze v případě splnění první podmínky na algebraické násobnosti vždy najít bázi, vzhledem ke které má operátor „téměř“ diagonální matici, přesněji tzv. matici v Jordanově tvaru. Mocninu takové matice stále lze explicitně vypočítat, tedy lze také spočítat libovolnou mocninu příslušného operátoru.

9.4.1. *Nediagonalizovatelné operátory v dimenzi* 2. V odstavci 9.3.3 jsme probrali možnosti, jaké mohou nastat pro operátory na reálném vektorovém prostoru dimenze 2. Zbyl jediný případ, který nyní podrobně rozebereme. Diskuze také snad poslouží k orientaci v obecnějších pojmech a tvrzeních.

• Operátor f má jedno vlastní číslo λ algebraické násobnosti 2 a geometrické násobnosti 1. Takovými operátory se budeme nyní zabývat v části o Jordanovu tvaru. Ilustrace je na následujícím obrázku.



OBRÁZEK 80. Jedno vlastní číslo algebrické násobnosti 2 a geometrické násobnosti 1. Vlevo $\lambda > 1$, vpravo $1 > \lambda > 0$.

Ukážeme, že v takovém případě existuje báze $B = (\mathbf{u}, \mathbf{v})$ prostoru $\mathbf{V}$, vzhledem ke které má operátor f matici

$$[f]^B_B = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} .$$

To je nejjednodušší příklad matice v tzv. Jordanově tvaru, která není diagonální. Takou matici umíme umocnit, platí totiž

$$\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}^m = \begin{pmatrix} \lambda^m & m\lambda^{m-1} \\ 0 & \lambda^m \end{pmatrix} .$$

Když takovou bázi B najdeme, budeme umět spočítat $[f^m]^B_B = ([f]^B_B)^m$ a tím pádem například vyřešit diskrétní dynamický systém $\mathbf{x}_m = f(\mathbf{x}_{m-1})$. Obecně je Jordanův tvar a mocnění rozebráno v odstavci 9.4.2.

Hledání báze B začneme přeformulováním podmínky na matici $[f]^B_B$. Podle definice matice f vzhledem k B a B potřebujeme, aby platilo

$$[f(\mathbf{u})]_B = \begin{pmatrix} \lambda \\ 0 \end{pmatrix}, \quad [f(\mathbf{v})]_B = \begin{pmatrix} 1 \\ \lambda \end{pmatrix} .$$

Podle definice souřadnic vzhledem k bázi tedy chceme, aby

$$f(\mathbf{u}) = \lambda\mathbf{u}, \quad f(\mathbf{v}) = \mathbf{u} + \lambda\mathbf{v} .$$

Označíme-li g operátor $f - \lambda\,\mathrm{id}_V$ můžeme tyto podmínky zapsat

$$g(\mathbf{u}) = 0, \quad g(\mathbf{v}) = \mathbf{u} ,$$

schematicky

$$\mathbf{v} \xmapsto{\ g\ } \mathbf{u} \xmapsto{\ g\ } \mathbf{o} .$$

Potřebujeme tedy, aby $\mathbf{u}$ byl vlastní vektor operátoru f příslušný vlastnímu číslu λ a aby $g(\mathbf{v}) = \mathbf{u}$. Obecně je toto přeformulování provedeno v odstavci 9.4.3.

Dále ukážeme, že kdykoliv máme vektory $\mathbf{u}, \mathbf{v}$ splňující podmínky $g(\mathbf{v}) = \mathbf{u}$, $g(\mathbf{u}) = \mathbf{o}$ a vektor $\mathbf{u}$ je nenulový, pak $B = (\mathbf{u}, \mathbf{v})$ je lineárně nezávislá posloupnost, tedy báze $\mathbf{V}$. Skutečně, je-li

$$a\mathbf{v} + b\mathbf{u} = \mathbf{o} \ ,$$

pak aplikací operátoru g na obě strany dostaneme

$$\begin{aligned} g(a\mathbf{v} + b\mathbf{u}) &= g(\mathbf{o}) \\ ag(\mathbf{v}) + bg(\mathbf{u}) &= \mathbf{o} \\ a\mathbf{u} &= \mathbf{o} \ . \end{aligned}$$

Protože je $\mathbf{u}$ nenulový vektor, plyne odsud $a = 0$. Dosazením do původního vztahu získáme $b\mathbf{u} = \mathbf{o}$, takže i $b = 0$. Obecné tvrzení je dokázáno v odstavci 9.4.4

V odstavci 9.4.5 probereme, jak takové vektory obecně hledat za předpokladu, že existují (v odstavcích 9.4.6, 9.4.7, 9.4.8, 9.4.9 ukážeme postup na řadě příkladů). Důkaz, že skutečně existují, využívá pojem invariantního podprostoru diskutovaného v odstavci 9.4.10, samotný důkaz je pak obsažen v odstavci 9.4.11.

V našem případě označme $\mathbf{W} = \operatorname{Im} g$. Tento prostor je invariantní ve smyslu, že $g(\mathbf{x}) \in W$ kdykoliv $\mathbf{x} \in W$. Skutečně, je-li $\mathbf{x} \in W$, pak z definice $\mathbf{W}$ existuje vektor $\mathbf{y} \in V$ takový, že $g(\mathbf{y}) = \mathbf{x}$. Pak $g(\mathbf{x}) = g(g(\mathbf{y})) \in \operatorname{Im} g$.

Podprostor $\operatorname{Ker} g \leq \mathbf{V}$ je tvořen všemi vlastními vektory operátoru f příslušnými vlastnímu číslu λ. Protože geometrická násobnost vlastního čísla λ je podle předpokladu 1, je $\dim \operatorname{Ker} g = 1$. Podle věty o dimenzi jádra a obrazu je $\dim(\operatorname{Im} g) = \dim \mathbf{W} = 1$. Vezmeme libovolný nenulový vektor $\mathbf{u} \in W$. Protože $\mathbf{W}$ má dimenzi 1, je vektor $g(\mathbf{u}) \in W$ násobkem vektoru $\mathbf{u}$, tj. $g(\mathbf{u}) = a\mathbf{u}$ pro nějaký skalár $a \in T$. Pak ale $(f - \lambda \operatorname{id}_V)\mathbf{u} = a\mathbf{u}$, takže $f(\mathbf{u}) = (\lambda + a)\mathbf{u}$. Protože f má jediné vlastní číslo λ, musí být nutně $a = 0$, platí tedy $g(\mathbf{u}) = \mathbf{o}$. Vektor $\mathbf{u}$ leží ve $W = \operatorname{Im} g$, existuje proto vektor $\mathbf{v} \in V$ takový, že $g(\mathbf{v}) = \mathbf{u}$, a důkaz je hotov – nalezli jsme vektory $\mathbf{u}, \mathbf{v} \in V$ takové, že $g(\mathbf{v}) = \mathbf{u}$, $g(\mathbf{u}) = \mathbf{o}$ a $\mathbf{u} \neq \mathbf{o}$, což znamená, že pro bázi $B = (\mathbf{u}, \mathbf{v})$ v prostoru $\mathbf{V}$ platí

$$[f]^B_B = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} \ .$$

9.4.2. *Matice v Jordanově tvaru.* Matice v Jordanově tvaru je blokově diagonální matice, jejíž bloky tvoří Jordanovy buňky. Jordanova buňka je matice, která má všechny diagonální prvky rovny nějakému $\lambda \in T$ a všechny prvky o jednu pozici nad diagonálou rovny 1.

Definice 9.70. *Jordanova buňka* řádu $k \geq 1$ nad tělesem $\mathbf{T}$ příslušná prvku $\lambda \in T$ je čtvercová matice

$$J_{\lambda,k} = \begin{pmatrix} \lambda & 1 & 0 & & 0 & 0 \\ 0 & \lambda & 1 & \dots & 0 & 0 \\ 0 & 0 & \lambda & & 0 & 0 \\ & \vdots & & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \lambda & 1 \\ 0 & 0 & 0 & \dots & 0 & \lambda \end{pmatrix} \ .$$

Příklad 9.71. Reálné matice

$$\begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 3 & 1 & 0 \\ 0 & 3 & 1 \\ 0 & 0 & 3 \end{pmatrix}, (4)$$

jsou Jordanovy buňky $J_{2,2}$, $J_{0,2}$, $J_{3,3}$, $J_{4,1}$ (příslušné pořadě číslům $2,0,3,4$).

Definice 9.72. Matice J nad tělesem $\mathbf{T}$ je v *Jordanově kanonickém tvaru* (nebo stručněji v *Jordanově tvaru*), pokud J je blokově diagonální matice, jejíž každý diagonální blok je Jordanova buňka (nějakého řádu příslušná nějakému číslu), tj.

$$J = \mathrm{diag}(J_{\lambda_1,k_1}, \dots, J_{\lambda_s,k_s}) = \begin{pmatrix} J_{\lambda_1,k_1} & 0 & \dots & 0 \\ 0 & J_{\lambda_2,k_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & J_{\lambda_s,k_s} \end{pmatrix},$$

kde $\lambda_1, \dots, \lambda_s \in T$ a $k_1, \dots, k_s$ jsou kladná celá čásla. (Nuly v matici v tomto případě značí nulové matice vhodných typů.)

Příklad 9.73. Diagonální matice $\mathrm{diag}(\lambda_1, \dots, \lambda_n)$ je v Jordanově tvaru. Je složená z Jordanových buněk $J_{\lambda_1,1}$, $\dots$, $J_{\lambda_n,1}$ řádu 1.

Příklad 9.74. Matice

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

je v Jordanově tvaru. Je složená z Jordanových buněk $J_{0,2}$, $J_{0,1}$, $J_{2,3}$.

Nyní najdeme vzorec pro mocninu Jordanovy matice. Matici

$$J = \begin{pmatrix} J_1 & 0 & \dots & 0 \\ 0 & J_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & J_s \end{pmatrix}$$

v blokově diagonálním tvaru můžeme mocnit po diagonálních blocích:

$$J^m = \begin{pmatrix} J_1 & 0 & \dots & 0 \\ 0 & J_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & J_s \end{pmatrix}^m = \begin{pmatrix} J_1^m & 0 & \dots & 0 \\ 0 & J_2^m & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & J_s^m \end{pmatrix}.$$

Stačí se proto zaměřit pouze na mocnění Jordanových buněk.

Jednoduchý je speciální případ Jordanových buněk příslušných prvku 0.

Tvrzení 9.75. *Pro libovolná přirozená čísla $m < k$ platí*

$$J_{0,k}^m = (\underbrace{\mathbf{o}|\cdots|\mathbf{o}}_{m\times}|\mathbf{e}_1|\mathbf{e}_2|\cdots|\mathbf{e}_{k-m})$$

Pro $m \geq k$ je $J_{0,k}^m = 0$.

Důkaz. Indukcí podle $m < k$, případ $m = 1$ je zjevný, neboť $J_{0,k} = (\mathbf{o}|\mathbf{e}_1|\mathbf{e}_2|\cdots|\mathbf{e}_{k-1})$. Platí-li tvrzení pro nějaké m menší než k, máme ze sloupcového pohledu na násobení

$$\begin{aligned} J_{0,k}^{m+1} &= J_{0,k}^m J_{0,k} = (\underbrace{\mathbf{o}|\ldots|\mathbf{o}}_{m\times}|\mathbf{e}_1|\mathbf{e}_2|\ldots|\mathbf{e}_{k-m})(\mathbf{o}|\mathbf{e}_1|\mathbf{e}_2|\ldots|\mathbf{e}_{k-1}) \\ &= (\underbrace{\mathbf{o}|\ldots|\mathbf{o}}_{(m+1)\times}|\mathbf{e}_1|\ldots|\mathbf{e}_{k-(m+1)}) \ . \end{aligned}$$

Pro $m \geq k$ je indukční krok zřejmý, neboť $J_{0,k}^m = 0$. □

Příklad 9.76.

$$J_{0,4}^2 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \ J_{0,4}^3 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} .$$

Jordanovu buňku $J_{\lambda,k}$ můžeme rozepsat

$$J_{\lambda,k} = \lambda I_k + J_{0,k}$$

Pokud dvě čtvercové matice A, B komutují, tj. platí-li $AB = BA$, pak pro ně platí obdoba binomické věty (cvičení)

$$(A+B)^m = \binom{m}{0}A^m + \binom{m}{1}A^{m-1}B + \binom{m}{2}A^{m-2}B^2 + \cdots + \binom{m}{m}B^m \ .$$

Použitím na matice λI_k a $J_{0,k}$ dostaneme vzorec v následujícím tvrzení. Používáme konvenci, že binomické číslo $\binom{m}{j} = 0$ pokud $m < j$. Dále pro $i \in \{0, 1, \ldots\}$ a prvek t v tělese $\mathbf{T}$ definujeme it jako $\underbrace{t + t + \cdots + t}_{i\times}$.

Tvrzení 9.77. *Je-li $J = J_{\lambda,k}$ Jordanova buňka, pak pro každé kladné m platí*

$$J_{\lambda,k}^m = \begin{pmatrix} \lambda^m & \binom{m}{1}\lambda^{m-1} & \binom{m}{2}\lambda^{m-2} & \ldots & \binom{m}{k-1}\lambda^{m-k+1} \\ 0 & \lambda^m & \binom{m}{1}\lambda^{m-1} & \ldots & \binom{m}{k-2}\lambda^{m-k+2} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda^m & \binom{m}{1}\lambda^{m-1} \\ 0 & 0 & \ldots & 0 & \lambda^m \end{pmatrix} .$$

Důkaz. Jeden z možných výpočtů byl naznačen před tvrzením, ukážeme alternativní důkaz.

Prvek na místě (i, j) v mocnině $J_{\lambda,k}^m$ zapsat jako $\binom{m}{j-i}\lambda^{m-(j-i)}$. K důkazu lze použít indukci podle m, případ $m = 1$ je zjevný. Pokud formulka platí pro $m \geq 1$, spočítáme $J_{\lambda,k}^{m+1} = J_{\lambda,k}J_{\lambda,k}^m$. Obě matice, které násobíme, jsou horní trojúhelníkové, součin je proto také horní trojúhelníkový. Zbývá spočítat prvky na místě (i, j) v součinu $J_{\lambda,k}J_{\lambda,k}^m$ pro $i \leq j$. Prvek na místě (i, j) v matici $J_{\lambda,k}^m$ se podle indukčního předpokladu rovná $\binom{m}{j-i}\lambda^{m-(j-i)}$. Prvek na místě (i, j) v matici $J_{\lambda,k}J_{\lambda,k}^m$ se pak rovná

$$\begin{aligned} &\lambda\binom{m}{j-i}\lambda^{m-(j-i)} + 1\binom{m}{j-(i+1)}\lambda^{m-(j-i-1)} \\ = \ &\binom{m}{j-i}\lambda^{m+1-(j-i)} + \binom{m}{j-i-1}\lambda^{m+1-(j-i)} \\ = \ &\binom{m+1}{j-i}\lambda^{m+1-(j-i)} \ , \end{aligned}$$

použili jsme vztah mezi kombinačními čísly $\binom{m}{l} + \binom{m}{l-1} = \binom{m+1}{l}$. □

9.4.3. *Operátory s Jordanovým tvarem.* Chceme zjistit, zda daný operátor f na konečně generovaném prostoru má vzhledem k nějaké bázi B matici v Jordanově tvaru, jak takovou bázi najít, a z jakých buněk se matice $[f]_B^B$ skládá.

Formulujeme obdobu definice 9.46 diagonalizovatelnosti. Pojem „jordanizovatelnost“ se nepoužívá, raději říkáme, že pro operátor existuje Jordanův kanonický tvar.

Definice 9.78. Říkáme, že pro lineární operátor $f : \mathbf{V} \to \mathbf{V}$ na konečně generovaném prostoru $\mathbf{V}$ *existuje Jordanův kanonický tvar*, pokud má vzhledem k nějaké bázi matici v Jordanově kanonickém tvaru.

Odvodíme obdobu tvrzení 9.47. Nejprve pro samotné buňky. Kdy má operátor $f : \mathbf{V} \to \mathbf{V}$ vzhledem k nějaké bázi $B = (\mathbf{v}_1, \ldots, \mathbf{v}_k)$ matici $[f]_B^B = J_{\lambda,k}$? Podle definice matice operátoru musí platit (a stačí)

$$[f(\mathbf{v}_1)]_B = \begin{pmatrix} \lambda \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad [f(\mathbf{v}_2)]_B = \begin{pmatrix} 1 \\ \lambda \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ldots, [f(\mathbf{v}_k)]_B = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ \lambda \end{pmatrix} \ ,$$

neboli

$$f(\mathbf{v}_1) = \lambda\mathbf{v}_1,\ f(\mathbf{v}_2) = \lambda\mathbf{v}_2 + \mathbf{v}_1,\ f(\mathbf{v}_3) = \lambda\mathbf{v}_3 + \mathbf{v}_2,\ \ldots, f(\mathbf{v}_k) = \lambda\mathbf{v}_k + \mathbf{v}_{k-1} \ .$$

Úpravou (podobně jako v části 9.2.1) dostaneme ekvivalentně

$$\begin{aligned} &(f - \lambda\,\mathrm{id}_V)(\mathbf{v}_1) = \mathbf{o},\ (f - \lambda\,\mathrm{id}_V)(\mathbf{v}_2) = \mathbf{v}_1,\ (f - \lambda\,\mathrm{id}_V)(\mathbf{v}_3) = \mathbf{v}_2,\ \ldots, \\ &(f - \lambda\,\mathrm{id}_V)(\mathbf{v}_k) = \mathbf{v}_{k-1} \ , \end{aligned}$$

schematicky

$$\mathbf{v}_k \xmapsto{f-\lambda\,\mathrm{id}_V} \mathbf{v}_{k-1} \xmapsto{f-\lambda\,\mathrm{id}_V} \ldots \xmapsto{f-\lambda\,\mathrm{id}_V} \mathbf{v}_3 \xmapsto{f-\lambda\,\mathrm{id}_V} \mathbf{v}_2 \xmapsto{f-\lambda\,\mathrm{id}_V} \mathbf{v}_1 \xmapsto{f-\lambda\,\mathrm{id}_V} \mathbf{o} \ .$$

Vidíme, že v tom případě je λ vlastní číslo operátoru f, a že $\mathbf{v}_1$ je vlastní vektor příslušný λ. Posloupnosti $(\mathbf{v}_1, \ldots, \mathbf{v}_k)$ budeme říkat Jordanův řetízek, vektorům $\mathbf{v}_2, \ldots, \mathbf{v}_k$ se někdy říká *zobecněné vlastní vektory* příslušné vlastnímu číslu λ.

Definice 9.79. Je-li f lineární operátor na vektorovém prostoru $\mathbf{V}$ nad tělesem $\mathbf{T}$ a λ vlastní číslo operátoru f, pak posloupnost $(\mathbf{v}_1, \ldots, \mathbf{v}_k)$ vektorů z $\mathbf{V}$ nazýváme *Jordanův řetízek operátoru* f *délky* k *příslušný vlastnímu číslu* λ *s počátkem* $\mathbf{v}_1$, pokud platí

$$\begin{aligned} &(f - \lambda\,\mathrm{id}_V)(\mathbf{v}_1) = \mathbf{o},\ (f - \lambda\,\mathrm{id}_V)(\mathbf{v}_2) = \mathbf{v}_1,\ (f - \lambda\,\mathrm{id}_V)(\mathbf{v}_3) = \mathbf{v}_2,\ \ldots, \\ &(f - \lambda\,\mathrm{id}_V)(\mathbf{v}_k) = \mathbf{v}_{k-1} \ . \end{aligned}$$

Před definicí jsme odvodili následující tvrzení.

Tvrzení 9.80. *Je-li $f : \mathbf{V} \to \mathbf{V}$ lineární operátor na konečně generovaném prostoru $\mathbf{V}$ a $B = (\mathbf{v}_1, \dots, \mathbf{v}_k)$ báze prostoru $\mathbf{V}$, pak $[f]^B_B = J_{\lambda,k}$ právě tehdy, když $(\mathbf{v}_1, \dots, \mathbf{v}_k)$ je Jordanův řetízek operátoru f příslušný vlastnímu číslu λ s počátkem $\mathbf{v}_1$.*

Snadno se tvrzení zobecní na obecné matice v Jordanově tvaru. Budeme říkat, že posloupnost vektorů B je *spojením* posloupností

$$B_1 = (\mathbf{v}^1_1, \dots, \mathbf{v}^1_{k_1}), B_2 = (\mathbf{v}^2_1, \dots, \mathbf{v}^2_{k_2}), \dots, B_s = (\mathbf{v}^s_1, \dots, \mathbf{v}^s_{k_s}) \ ,$$

pokud

$$B = (\mathbf{v}^1_1, \dots, \mathbf{v}^1_{k_1}, \mathbf{v}^2_1, \dots, \mathbf{v}^2_{k_2}, \dots, \mathbf{v}^s_1, \dots, \mathbf{v}^s_{k_s}) \ .$$

Budeme také používat zápis

$$B = B_1, \dots, B_s \ .$$

Tvrzení 9.81. *Je-li $f : \mathbf{V} \to \mathbf{V}$ lineární operátor na konečně generovaném prostoru $\mathbf{V}$ a B báze prostoru $\mathbf{V}$, pak $[f]^B_B = \mathrm{diag}(J_{\lambda_1,k_1}, \dots, J_{\lambda_s,k_s})$ platí právě tehdy, když B je spojením posloupností $B_1, \dots, B_s$, kde pro každé $i \in \{1, \dots, s\}$ je B_i Jordanův řetízek operátoru f délky k_i příslušný vlastnímu číslu λ_i s počátkem $\mathbf{v}^i_1$.*

Důsledek 9.82. *Pro lineární operátor $f : \mathbf{V} \to \mathbf{V}$ na konečně generovaném prostoru $\mathbf{V}$ existuje Jordanův tvar právě tehdy, když existuje báze prostoru $\mathbf{V}$ vzniklá spojením Jordanových řetízků operátoru f.*

Nakonec formulujeme obdobu tvrzení 9.49. Důkaz je stejný.

Tvrzení 9.83. *Nechť $f : \mathbf{V} \to \mathbf{V}$ je lineární operátor na konečně generovaném prostoru $\mathbf{V}$ a C je báze prostoru $\mathbf{V}$. Pak pro operátor f existuje Jordanův tvar právě tehdy, když je matice $[f]^C_C$ podobná matici v Jordanově tvaru.*

Maticové verze definic a tvrzení přenecháme k rozmyšlení čtenáři.

9.4.4. *Lineární nezávislost zobecněných vlastních vektorů.* Chceme-li najít bázi, vzhledem ke které má operátor na prostoru dimenze n matici v Jordanově tvaru, musíme najít Jordanovy řetízky celkové délky n, tak aby jejich spojení byla lineárně nezávislá posloupnost. Následující věta, která zobecňuje větu 9.54 o lineární nezávislosti vlastních vektorů příslušných různým vlastním číslům, říká že stačí zaručit, aby pro každé vlastní číslo λ tvořily počáteční vektory řetízků příslušných vlastnímu číslu λ lineárně nezávislou posloupnost.

Věta 9.84. *Předpokládáme, že $f : \mathbf{V} \to \mathbf{V}$ je lineární operátor a $B_1, \dots, B_s$ jsou Jordanovy řetízky operátoru f příslušné vlastním číslům $\lambda_1, \dots, \lambda_s$. Předpokládejme dále, že pro každé $\lambda \in \{\lambda_1, \dots, \lambda_s\}$ je posloupnost počátečních vektorů těch řetízků z $B_1, \dots, B_s$, které přísluší vlastnímu číslu λ, lineárně nezávislá. Pak spojení $B = B_1, \dots, B_s$ je lineárně nezávislá posloupnost.*

Důkaz. Použijeme indukci podle celkového počtu k vektorů v řetízcích $B_1, \dots, B_s$. Pro $k = 1$ je tvrzení zřejmé, neboť v tom případě máme jediný řetízek délky 1 a jeho počáteční vektor je nenulový.

Předpokládáme nyní, že součet délek řetízků $B_1, B_2, \dots, B_s$ je $k > 1$. Indukční předpoklad je, že kdykoliv máme nějaké Jordanovy řetízky $C_1, C_2, \dots, C_t$ operátoru f o celkové délce menší než k a takové, že lineárně nezávislou posloupnost tvoří počáteční vektory těch řetízků mezi $C_1, C_2, \ldots, C_t$, které přísluší stejnému vlastnímu číslu, pak je spojení řetízků $C_1, C_2, \ldots, C_t$ lineárně nezávislá posloupnost.

Označíme r počet řetízků příslušných vlastnímu číslu λ_1 a uspořádáme si řetízky tak, že všechny řetízky příslušející vlastnímu číslu λ_1 jsou na začátku, tj. řetízky $B_1, \ldots, B_r$ přísluší vlastnímu číslu λ_1 a zbylé přísluší jiným vlastním číslům. Označme pro $i = 1, \ldots, s$

$$B_i = (\mathbf{v}_1^i, \mathbf{v}_2^i, \ldots, \mathbf{v}_{k_i}^i) \ .$$

Uvažujme skaláry $a_j^i \in T$ $(i \in \{1, \ldots, s\}, j \in \{1, \ldots, k_i\})$ takové, že

$$\begin{aligned}\mathbf{o} =& a_1^1\mathbf{v}_1^1 + a_2^1\mathbf{v}_2^1 + \cdots + a_{k_1}^1\mathbf{v}_{k_1}^1 \\ &+a_1^2\mathbf{v}_1^2 + a_2^2\mathbf{v}_2^2 + \vdots + a_{k_2}^2\mathbf{v}_{k_2}^2 \\ &\vdots \\ &+a_1^s\mathbf{v}_1^s + a_2^s\mathbf{v}_2^s + \cdots + a_{k_s}^s\mathbf{v}_{k_s}^s \ .\end{aligned}$$

Potřebujeme ukázat, že všechny skaláry a_j^i jsou nulové.

Aplikujeme na obě strany operátor $f - \lambda_1 \operatorname{id}_V$. Využitím linearity, podobně jako v důkazu věty 9.54, získáme

$$\begin{aligned}\mathbf{o} =& a_1^1(f-\lambda_1 \operatorname{id}_V)(\mathbf{v}_1^1) + a_2^1(f-\lambda_1 \operatorname{id}_V)(\mathbf{v}_2^1) + \cdots + a_{k_1}^1(f-\lambda_1 \operatorname{id}_V)(\mathbf{v}_{k_1}^1) \\ &+a_1^2(f-\lambda_1 \operatorname{id}_V)(\mathbf{v}_1^2) + a_2^2(f-\lambda_1 \operatorname{id}_V)(\mathbf{v}_2^2) + \cdots + a_{k_2}^2(f-\lambda_1 \operatorname{id}_V)(\mathbf{v}_{k_2}^2) \\ &\vdots \\ &+a_1^s(f-\lambda_1 \operatorname{id}_V)(\mathbf{v}_1^s) + a_2^s(f-\lambda_1 \operatorname{id}_V)(\mathbf{v}_2^s) + \cdots + a_{k_s}^s(f-\lambda_1 \operatorname{id})(\mathbf{v}_{k_s}^s) \ .\end{aligned}$$

Rozebereme výraz po jednotlivých řetízcích B_i. Pro $i \in \{1, \ldots, r\}$ je z definice Jordanova řetízku příslušného vlastnímu číslu λ_1

$$\begin{aligned}&a_1^i(f-\lambda_1 \operatorname{id}_V)(\mathbf{v}_1^i) + a_2^i(f-\lambda_1 \operatorname{id}_V)(\mathbf{v}_2^i) + \cdots + a_{k_i}^i(f-\lambda_1 \operatorname{id}_V)(\mathbf{v}_{k_i}^i) \\ &= a_2^i\mathbf{v}_1^i + a_3^i\mathbf{v}_2^i + \cdots + a_{k_i}^i\mathbf{v}_{k_i-1}^i \ .\end{aligned}$$

Pro $i > r$ využijeme úpravy

$$\begin{aligned}(f-\lambda_1 \operatorname{id}_V)(\mathbf{v}_j^i) &= f(\mathbf{v}_j^i) - \lambda_1\mathbf{v}_j^i \\ &= f(\mathbf{v}_j^i) - \lambda_i\mathbf{v}_j^i + \lambda_i\mathbf{v}_j^i - \lambda_1\mathbf{v}_j^i \\ &= (f-\lambda_i \operatorname{id}_V)(\mathbf{v}_j^i) + (\lambda_i - \lambda_1)\mathbf{v}_j^i \\ &= \mathbf{v}_{j-1}^i + (\lambda_i - \lambda_1)\mathbf{v}_j^i \ .\end{aligned}$$

Dosadíme do následujícího součtu a upravíme

$$\begin{aligned}&a_1^i(f-\lambda_1 \operatorname{id}_V)(\mathbf{v}_1^i) + a_2^i(f-\lambda_1 \operatorname{id}_V)(\mathbf{v}_2^i) + \cdots + \\ &\qquad + a_{k_i-1}^i(f-\lambda_1 \operatorname{id}_V)(\mathbf{v}_{k_i-1}^i) + a_{k_i}^i(f-\lambda_1 \operatorname{id}_V)(\mathbf{v}_{k_i}^i) \\ &= a_1^i\mathbf{o} + a_1^i(\lambda_i-\lambda_1)\mathbf{v}_1^i + a_2^i\mathbf{v}_1^i + a_2^i(\lambda_i-\lambda_1)\mathbf{v}_2^i + \cdots + a_{k_i-1}^i\mathbf{v}_{k_i-2}^i + \\ &\qquad + a_{k_i-1}^i(\lambda_i-\lambda_1)\mathbf{v}_{k_i-1}^i + a_{k_i}^i\mathbf{v}_{k_i-1}^i + a_{k_i}^i(\lambda_i-\lambda_1)\mathbf{v}_{k_i}^i \\ &= (a_1^i(\lambda_i-\lambda_1) + a_2^i)\mathbf{v}_1^i + (a_2^i(\lambda_i-\lambda_1) + a_3^i)\mathbf{v}_2^i + \cdots + \\ &\qquad + (a_{k_i-1}^i(\lambda_i-\lambda_1) + a_{k_i}^i)\mathbf{v}_{k_i-1}^i + a_{k_i}^i(\lambda_i-\lambda_1)\mathbf{v}_{k_i}^i \ .\end{aligned}$$

Dosadíme do hodnoty operátoru $f-\lambda_1 \,\mathrm{id}_V$ na původní lineární kombinaci prvků spojení řetízků $B_1, B_2, \dots, B_s$ a dostaneme

$$\begin{aligned}
\mathbf{o} &= a_2^1\mathbf{v}_1^1 + a_3^1\mathbf{v}_2^1 + \cdots + a_{k_1}^1\mathbf{v}_{k_1-1}^1\\
&\ \vdots\\
&+ a_2^r\mathbf{v}_1^r + a_3^r\mathbf{v}_2^r + \cdots + a_{k_1}^r\mathbf{v}_{k_r-1}^r\\
&+ (a_1^{r+1}(\lambda_{r+1}-\lambda_1) + a_2^{r+1})\mathbf{v}_1^{r+1} + (a_2^{r+1}(\lambda_{r+1}-\lambda_1) + a_3^{r+1})\mathbf{v}_2^{r+1} + \cdots +\\
&\qquad + (a_{k_{r+1}-1}^{r+1}(\lambda_{r+1}-\lambda_1) + a_{k_{r+1}}^{r+1})\mathbf{v}_{k_{r+1}-1}^{r+1} + a_{k_{r+1}}^{r+1}(\lambda_{r+1}-\lambda_1)\mathbf{v}_{k_{r+1}}^{r+1}\\
&\quad (a_1^s(\lambda_s-\lambda_1) + a_2^s)\mathbf{v}_1^s + (a_2^s(\lambda_s-\lambda_1) + a_3^s)\mathbf{v}_2^s + \cdots +\\
&\qquad + (a_{k_s-1}^s(\lambda_s-\lambda_1) + a_{k_s}^s)\mathbf{v}_{k_s-1}^s + a_{k_s}^s(\lambda_s-\lambda_1)\mathbf{v}_{k_s}^s \ .
\end{aligned}$$

Tento výraz je lineární kombinací vektorů v řetízcích $B_1', B_2' \dots, B_r', B_{r+1}, \dots, B_s$, kde řetízek B_i' vznikne z B_i odebráním posledního vektoru v řetízku B_i, může tak vzniknout i prázdný řetízek, pokud měl některý z řetízků B_i pro $i = 1, 2, \dots, r$ délku 1. Počáteční vektory v řetízcích $B_1', B_2' \dots, B_r', B_{r+1}, \dots, B_s$ jsou podposloupností počátečních vektorů v řetízcích $B_1, B_2, \dots, B_s$ a ty z nich, které jsou příslušné stejnému vlastnímu číslu λ, proto tvoří lineárně nezávislou posloupnost podle předpokladu věty. Celková délka řetízků $B_1', B_2' \dots, B_r', B_{r+1}, \dots, B_s$ je o $r \geq 1$ menší než celková délka řetízků $B_1, B_2, \dots, B_s$, a tedy je menší než k. Z indukčního předpokladu proto vyplývá, že všechny koeficienty v poslední lineární kombinaci jsou nutně nulové, tj.

$$a_2^1 = a_3^1 \cdots = a_{k_1}^1 = a_2^2 = \cdots a_{k_2}^2 = \cdots = a_2^r = \cdots = a_{k_r}^r = 0$$

a také

$$a_1^i(\lambda_i-\lambda_1) + a_2^i = a_2^i(\lambda_i-\lambda_1) + a_3^i = \cdots = a_{k_i-1}^i(\lambda_i-\lambda_1) + a_{k_i}^i = a_{k_i}^i(\lambda_i-\lambda_1) = 0$$

pro každé $i = r+1, \dots, n$. Protože pro $i > r$ platí $\lambda_i \neq \lambda_1$, plyne z poslední rovnosti $a_{k_i}^i = 0$, po dosazení $a_{k_i}^i = 0$ do předposledního členu dostáváme $a_{k_i-1}^i$ a tak pokračujeme zpětně až nakonec dostaneme také $a_2^i = a_1^i = 0$. Tím jsme získali rovněž

$$a_1^{r+1} = a_2^{r+1} = \cdots = a_{k_{r+1}}^{r+1} = \cdots = a_1^s = a_2^s = \cdots = a_{k_s}^s = 0 \ .$$

Po dosazení 0 za všechny tyto koeficienty do původní lineární kombinace prvků řetízků $B_1, B_2, \dots, B_s$ zůstane

$$\mathbf{o} = a_1^1\mathbf{v}_1^1 + a_1^2\mathbf{v}_1^2 + \cdots + a_1^r\mathbf{v}_1^r$$

a z předpokladu o lineární nezávislosti počátků řetízků $B_1, B_2, \dots, B_r$ příslušných vlastnímu číslu λ_1 získáme konečně také

$$a_1^1 = a_1^2 = \cdots = a_1^r = 0 \ .$$

□

9.4.5. *Výpočet řetízků.* Uvažujme operátor $f : \mathbf{V} \to \mathbf{V}$ na konečně generovaném prostoru $\mathbf{V}$ dimenze n a bázi $B = B_1, \dots, B_s$ složenou ze Jordanových řetízků $B_1, \dots, B_s$ délek $k_1, \dots, k_s$ příslušných vlastním číslům $\lambda_1, \dots, \lambda_s$. Pro přehlednost si je uspořádáme tak, aby řetízky příslušné stejným číslům byly pohromadě. Řekněme, že prvních r odpovídá stejnému vlastnímu číslu λ, tj. $\lambda = \lambda_1 = \lambda_2 = \cdots = \lambda_r$

a $\lambda_i \neq \lambda$ pro $i > r$. Označme $B_i = (\mathbf{v}^i_1, \mathbf{v}^i_2, \dots, \mathbf{v}^i_{k_i})$ pro $i \in \{1, \dots, s\}$. Schematicky:

$$\begin{array}{c}
\mathbf{v}^1_{k_1} \xmapsto{f-\lambda\,\mathrm{id}_V} \mathbf{v}^1_{k_1-1} \longmapsto \dots \longmapsto \mathbf{v}^1_2 \xmapsto{f-\lambda\,\mathrm{id}_V} \mathbf{v}^1_1 \xmapsto{f-\lambda\,\mathrm{id}_V} \mathbf{o} \\
\vdots \\
\mathbf{v}^r_{k_1} \xmapsto{f-\lambda\,\mathrm{id}_V} \mathbf{v}^r_{k_1-1} \longmapsto \dots \longmapsto \mathbf{v}^r_2 \xmapsto{f-\lambda\,\mathrm{id}_V} \mathbf{v}^r_1 \xmapsto{f-\lambda\,\mathrm{id}_V} \mathbf{o} \\
\vdots \\
\mathbf{v}^s_{k_s} \xmapsto{f-\lambda_s\,\mathrm{id}_V} \mathbf{v}^s_{k_s-1} \longmapsto \dots \longmapsto \mathbf{v}^s_2 \xmapsto{f-\lambda_s\,\mathrm{id}_V} \mathbf{v}^s_1 \xmapsto{f-\lambda_s\,\mathrm{id}_V} \mathbf{o}
\end{array}$$

a

$$[f]^B_B = J = \mathrm{diag}(J_{\lambda_1,k_1}, J_{\lambda_2,k_2}, \dots, J_{\lambda_s,k_s}) \ .$$

Za této situace spočítáme charakteristický polynom operátoru f, vlastní čísla a vektory, geometrické násobnosti a navíc jádra a obrazy operátorů $(f - \lambda_i\,\mathrm{id}_V)^l$ pro $l = 1, 2, \dots$. (Zaměříme se na vlastní číslo $\lambda_1 = \dots = \lambda_r$, přičemž výsledky přirozeně budou platit pro všechna další vlastní čísla.) Tyto poznatky nám pak umožní hledat Jordanovy řetízky i v situaci, kdy je předem neznáme.

Charakteristický polynom $p_f(\lambda)$ operátoru f je roven determinantu matice $J - \lambda I_n$. Tato matice je horní trojúhelníková a na diagonále má postupně k_1-krát výraz $(\lambda_1 - \lambda)$, k_2-krát výraz $(\lambda_2 - \lambda)$, atd. Charakteristický polynom je proto roven

$$(\lambda_1 - \lambda)^{k_1} (\lambda_2 - \lambda)^{k_2} \dots (\lambda_s - \lambda)^{k_s} \ .$$

Důsledkem je, že operátor f má n vlastních čísel včetně násobností a algebraická násobnost vlastního čísla λ_1 je rovna součtu délek Jordanových řetízků příslušných vlastnímu číslu λ_1 (to jest $k_1 + k_2 + \dots + k_r$).

Dále vypočítáme jádro a obraz operátoru $f - \lambda_1\,\mathrm{id}_V$. (Pro představu je dobré sledovat výpočet na konkrétní situaci, viz např. příklad 9.85). Jeho matice $[f - \lambda_1\,\mathrm{id}_V]^B_B$ vzhledem k bázi B je

$$J - \lambda_1 I_n = \mathrm{diag}(J_{0,k_1}, \dots, J_{0,k_r}, J_{\lambda_{r+1}-\lambda_1,k_{r+1}}, \dots, J_{\lambda_s-\lambda_1,k_s})$$

Tato matice má nulové řádky, které odpovídají pozici koncových vektorů řetízků $B_1, \dots, B_r$ v bázi B. Vynecháme-li je, dostaneme matici v řádkově odstupňovaném tvaru s $(n - r)$ nenulovými řádky. Dimenze jádra matice $J - \lambda_1 I_n$ je r a také vidíme, že množina všech řešení homogenní soustavy rovnic s maticí $J - \lambda_1 I_n$ je $\mathrm{Ker}\,(J - \lambda_1 I_n) = \langle \mathbf{e}_{i_1}, \dots, \mathbf{e}_{i_r} \rangle$, kde indexy $i_1, \dots, i_r$ odpovídají pozicím počátečních vektorů řetízků $B_1, \dots, B_r$ v bázi B, tj. $i_1 = 1$, $i_2 = 1 + k_1$, $i_3 = 1 + k_1 + k_2$, …, $i_r = 1 + k_1 + \dots + k_{r-1}$. Je tedy

$$[\mathrm{Ker}\,(f - \lambda_1\,\mathrm{id}_V)]_B = \mathrm{Ker}\,(J - \lambda_1 I_n) = \langle \mathbf{e}_{i_1}, \dots, \mathbf{e}_{i_r} \rangle \ ,$$

takže

$$\mathrm{Ker}\,(f - \lambda_1\,\mathrm{id}_V) = \left\langle \mathbf{v}^1_1, \mathbf{v}^2_1, \dots, \mathbf{v}^r_1 \right\rangle \ .$$

Geometrická násobnost r vlastního čísla λ_1 je tedy rovná počtu řetízků příslušných vlastnímu číslu λ_1 a jádro operátoru $(f - \lambda_1\,\mathrm{id}_V)$ je rovno lineárnímu obalu počátečních vektorů těchto řetízků.

Přejdeme k výpočtu obrazu (tj. oboru hodnot) $\mathrm{Im}\,(f - \lambda_1\,\mathrm{id}_V)$ operátoru $f - \lambda_1\,\mathrm{id}_V$. Obor hodnot matice $J - \lambda_1 I_n = [f - \lambda_1\,\mathrm{id}_V]^B_B$ se rovná lineárnímu obalu sloupcových vektorů. Sloupce odpovídající pozicím počátečních vektorů řetízků $B_1, \dots, B_r$ jsou nulové a zbylé sloupce příslušné kterékoliv z buněk J_{0,k_i} pro $i =$

$1, 2, \dots, r$ obsahují vektory kanonické báze. Ostatní buňky (příslušné vlastním číslům různým od λ_1) jsou horní trojúhelníkové matice s nenulovými prvky na hlavní diagonále, můžeme je tedy elementárními sloupcovými úpravami (které nemění obraz) převést na jednotkové matice. Obraz matice $J - \lambda_1 I_n$ je tedy roven lineárnímu obalu těch vektorů kanonické báze, které neodpovídají pozicím koncových vektorů řetízků $B_1, \dots, B_r$. Protože $[\mathrm{Im}\,(f - \lambda_1 \,\mathrm{id}_V)]_B = \mathrm{Im}\,(J - \lambda_1 I_n)$, je obraz operátoru $f - \lambda_1 \,\mathrm{id}_V$ roven lineárnímu obalu všech vektorů v B kromě koncových vektorů řetízků $B_1, \dots, B_r$. Můžeme si představovat, že umažeme jeden vektor z konce každého řetízku příslušného vlastnímu číslu λ_1.

Příklad 9.85. Pro $\lambda = \lambda_1 = \lambda_2 = \lambda_3 = 7$, $\lambda_4 = 9$, $k_1 = 1$, $k_2 = 2$, $k_3 = 3$, $k_4 = 2$ máme řetízky

$$\begin{array}{rrrr}
 & & \mathbf{v}_1^1 \xmapsto{f-7\,\mathrm{id}_V} & \mathbf{o} \\
 & \mathbf{v}_2^2 \xmapsto{f-7\,\mathrm{id}_V} & \mathbf{v}_1^2 \xmapsto{f-7\,\mathrm{id}_V} & \mathbf{o} \\
\mathbf{v}_3^3 \xmapsto{f-7\,\mathrm{id}_V} & \mathbf{v}_2^3 \xmapsto{f-7\,\mathrm{id}_V} & \mathbf{v}_1^3 \xmapsto{f-7\,\mathrm{id}_V} & \mathbf{o} \\
 & \mathbf{v}_2^4 \xmapsto{f-9\,\mathrm{id}_V} & \mathbf{v}_1^4 \xmapsto{f-9\,\mathrm{id}_V} & \mathbf{o}
\end{array}$$

Operátor f má charakteristický polynom $p_f(\lambda) = (7 - \lambda)^6 (9 - \lambda)^2$, vlastní číslo 7 algebraické násobnosti 6 a vlastní číslo 9 algebraické násobnosti 2.

Matice operátoru $f - 7\,\mathrm{id}_V$ vzhledem k B je

$$J - 7I_8 = \begin{pmatrix} 7&0&0&0&0&0&0&0\\ 0&7&1&0&0&0&0&0\\ 0&0&7&0&0&0&0&0\\ 0&0&0&7&1&0&0&0\\ 0&0&0&0&7&1&0&0\\ 0&0&0&0&0&7&0&0\\ 0&0&0&0&0&0&9&1\\ 0&0&0&0&0&0&0&9 \end{pmatrix} - 7I_8 = \begin{pmatrix} 0&0&0&0&0&0&0&0\\ 0&0&1&0&0&0&0&0\\ 0&0&0&0&0&0&0&0\\ 0&0&0&0&1&0&0&0\\ 0&0&0&0&0&1&0&0\\ 0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&2&1\\ 0&0&0&0&0&0&0&2 \end{pmatrix}.$$

Jádro matice $J - 7I_8$ je

$$\mathrm{Ker} \begin{pmatrix} 0&0&1&0&0&0&0&0\\ 0&0&0&0&1&0&0&0\\ 0&0&0&0&0&1&0&0\\ 0&0&0&0&0&0&2&1\\ 0&0&0&0&0&0&0&2 \end{pmatrix} = \langle \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_4 \rangle$$

a tudíž jádro operátoru $f - 7\,\mathrm{id}_V$ je $\mathrm{Ker}\,(f - 7\,\mathrm{id}_V) = \left\langle \mathbf{v}_1^1, \mathbf{v}_1^2, \mathbf{v}_1^3 \right\rangle$ (lineární obal počátečních vektorů příslušných vlastnímu číslu 7), jeho dimenze je 3 a je rovná geometrické násobnost vlastního čísla 7, která udává počet řetízků příslušných tomuto vlastnímu číslu.

Obraz matice $J - 7I_8$ je

$$\mathrm{Im}\begin{pmatrix} 0&0&0&0&0\\ 1&0&0&0&0\\ 0&0&0&0&0\\ 0&1&0&0&0\\ 0&0&1&0&0\\ 0&0&0&0&0\\ 0&0&0&2&1\\ 0&0&0&0&2 \end{pmatrix} = \mathrm{Im}\begin{pmatrix} 0&0&0&0&0\\ 1&0&0&0&0\\ 0&0&0&0&0\\ 0&1&0&0&0\\ 0&0&1&0&0\\ 0&0&0&0&0\\ 0&0&0&1&0\\ 0&0&0&0&1 \end{pmatrix} = \langle \mathbf{e}_2, \mathbf{e}_4, \mathbf{e}_5, \mathbf{e}_7, \mathbf{e}_8 \rangle$$

a tudíž obraz operátoru $f - 7\,\mathrm{id}_V$ je $\mathrm{Im}\,(f - 7\,\mathrm{id}_V) = \left\langle \mathbf{v}_1^2, \mathbf{v}_1^3, \mathbf{v}_2^3, \mathbf{v}_1^4, \mathbf{v}_2^4 \right\rangle$ (lineární obal všech vektorů v řetízcích kromě koncových vektorů příslušných vlastnímu číslu 7).

Nakonec obecněji vypočteme jádro a obraz operátoru $(f - \lambda_1\,\mathrm{id}_V)^l$ pro $l \geq 2$. Jeho matice vzhledem k B je

$$(J - \lambda_1 I_n)^l = \mathrm{diag}(J^l_{0,k_1}, \ldots, J^l_{0,k_r}, J^l_{\lambda_{r+1}-\lambda_1,k_{r+1}}, \ldots, J^l_{\lambda_s-\lambda_1,k_s})$$

Tvar prvních r diagonálních buněk jsme spočítali v tvrzení 9.75, tvar ostatních v tvrzení 9.77, ten ale teď nebudeme potřebovat, stačí vědět, že v případě buněk příslušných vlastním číslům různým od λ_1 vyjdou regulární matice (horní trojúhelníkové s nenulovými prvky na hlavní diagonále).

Matice $(J - \lambda_1 I_n)^l$ má nulové řádky odpovídající pozici l koncových prvků řetízků $B_1, \ldots, B_r$ v bázi B (pokud má některý z těchto řetízků délku nejvýše l, pak uvažujeme všechny jeho prvky). Vynecháme-li je, dostaneme matici v řádkově odstupňovaném tvaru a jádro $\mathrm{Ker}\,(J - \lambda_1 I_n)^l$ matice $(J - \lambda_1 I_n)^l$ je rovno lineárnímu obalu $\langle \mathbf{e}_{i_1}, \ldots, \rangle$, kde indexy $i_1, \ldots$, odpovídají pozicím l počátečních vektorů řetízků $B_1, \ldots, B_r$ v bázi B. To znamená, že

$$\mathrm{Ker}\,(f - \lambda_1\,\mathrm{id}_V)^l = \left\langle \mathbf{v}_1^1, \ldots, \mathbf{v}_l^1, \mathbf{v}_1^2, \ldots, \mathbf{v}_l^2, \ldots, \mathbf{v}_1^r, \ldots, \mathbf{v}_l^r \right\rangle \quad .$$

(Pro řetízky B_i délky menší než l nejsou vektory $\mathbf{v}^i_{k_i+1}, \ldots, \mathbf{v}^i_l$ definované.) Jádro operátoru $(f - \lambda_1\,\mathrm{id}_V)^l$ se rovná lineárnímu obalu l počátečních vektorů z každého řetízku příslušného vlastnímu číslu λ_1 (z řetízků délky menší než l bereme všechny vektory.)

Z toho také vyplývá důležité pozorování – počet řetízků příslušných vlastnímu číslu λ, které mají délku aspoň l, se rovná

$$\dim \mathrm{Ker}\,(f - \lambda_1\,\mathrm{id}_V)^l - \dim \mathrm{Ker}\,(f - \lambda_1\,\mathrm{id}_V)^{l-1} \quad .$$

Obraz operátoru $(f - \lambda_1\,\mathrm{id}_V)^l$ také spočteme obdobně jako v případu $l = 1$.

Příklad 9.86. Vrátíme se k příkladu 9.85. Operátory $(f - 7\,\mathrm{id})^2$, $(f - 7\,\mathrm{id})^3$ mají vzhledem k B matici

$$(J - 7I_8)^2 = \begin{pmatrix} 0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0\\ 0&0&0&0&0&1&0&0\\ 0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&4&4\\ 0&0&0&0&0&0&0&4 \end{pmatrix}, (J - 7I_8)^3 = \begin{pmatrix} 0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&16&12\\ 0&0&0&0&0&0&0&16 \end{pmatrix} .$$

Jádra jsou

$$\mathrm{Ker}\,(J-7I_8)^2 = \langle \mathbf{e}_1,\dots,\mathbf{e}_5\rangle\,,\ \mathrm{Ker}\,(J-7I_8)^3 = \mathrm{Ker}\,(J-7I_8)^4 = \dots = \langle \mathbf{e}_1,\dots,\mathbf{e}_6\rangle$$

a proto

$$\mathrm{Ker}\,(f-7\,\mathrm{id})^2 = \left\langle \mathbf{v}_1^1,\mathbf{v}_1^2,\mathbf{v}_2^2,\mathbf{v}_1^3,\mathbf{v}_2^3\right\rangle,\ \mathrm{Ker}\,(f-7\,\mathrm{id})^3 = \left\langle \mathbf{v}_1^1,\mathbf{v}_1^2,\mathbf{v}_2^2,\mathbf{v}_1^3,\mathbf{v}_2^3,\mathbf{v}_3^3\right\rangle \ .$$

Obrazy jsou

$$\mathrm{Im}\,(J-7I_8)^2 = \langle \mathbf{e}_4,\mathbf{e}_7,\mathbf{e}_8\rangle\,,\ \mathrm{Im}\,(J-7I_8)^3 = \mathrm{Im}\,(J-7I_8)^4 = \dots = \langle \mathbf{e}_7,\mathbf{e}_8\rangle$$

a proto

$$\mathrm{Im}\,(f-7\,\mathrm{id})^2 = \left\langle \mathbf{v}_1^3,\mathbf{v}_1^4,\mathbf{v}_2^4\right\rangle,\ \mathrm{Im}\,(f-7\,\mathrm{id})^3 = \left\langle \mathbf{v}_1^4,\mathbf{v}_2^4\right\rangle \ .$$

Shrneme získané poznatky.

Tvrzení 9.87. *Je-li $f : \mathbf{V}\to\mathbf{V}$ lineární operátor na prostoru $\mathbf{V}$ dimenze n a B báze vzniklá spojením Jordanových řetízků operátoru f, pak platí*

(1) *operátor f má n vlastních čísel včetně násobností,*
(2) *pro libovolné vlastní číslo λ operátoru f je jeho algebraická násobnost rovna součtu délek Jordanových řetízků v B příslušných vlastnímu číslu λ,*
(3) *pro libovolné vlastní číslo λ operátoru f a libovolné $l\in\mathbb{N}$ je jádro operátoru $(f-\lambda\,\mathrm{id}_V)^l$ rovno lineárnímu obalu l počátečních vektorů z každého řetízku v B příslušného vlastnímu číslu λ (z řetízků délky menší než l bereme všechny vektory),*
(4) *pro libovolné vlastní číslo λ operátoru f a libovolné $l\in\mathbb{N}$ je obraz operátoru $(f-\lambda\,\mathrm{id}_V)^l$, roven lineárnímu obalu všech vektorů v B kromě l koncových vektorů z řetízků příslušných vlastnímu číslu λ (z řetízků příslušných vlastnímu číslu λ délky menší než l nebereme žádný vektor).*

Speciálně pro libovolné vlastní číslo λ operátoru f platí

(5) *geometrická násobnost vlastního čísla λ se rovná počtu řetízků v B příslušných vlastnímu číslu λ a prostor $M_\lambda = \mathrm{Ker}\,(f-\lambda\,\mathrm{id}_V)$ je roven lineárnímu obalu počátečních vektorů těchto řetízků,*
(6) *počet řetízků příslušných vlastnímu číslu λ délky alespoň l je roven*

$$z_l = \dim\mathrm{Ker}\,(f-\lambda\,\mathrm{id}_V)^l - \dim\mathrm{Ker}\,(f-\lambda\,\mathrm{id}_V)^{l-1}\ ,$$

(aby měl výraz smysl i pro $l=1$ definujeme $(f-\lambda\,\mathrm{id}_V)^0 = \mathrm{id}_V$),
(7) *počet řetízků příslušných vlastnímu číslu λ délky právě l je $z_l - z_{l+1}$.*

Z prvního bodu vyplývá nutná podmínka pro existenci Jordanova kanonického tvaru – operátor musí mít dostatek vlastních čísel (včetně násobností). Tato podmínka je i dostačující podle následující veledůležité věty.

Věta 9.88 (o Jordanově kanonickém tvaru)**.** *Je-li $f : \mathbf{V}\to\mathbf{V}$ lineární operátor na konečně generovaném vektorovém prostoru $\mathbf{V}$ dimenze n, pak jsou následující tvrzení ekvivalentní.*

(1) *Pro operátor f existuje Jordanův kanonický tvar.*
(2) *Operátor f (resp. matice A) má n vlastních čísel včetně algebraických násobností.*

Důsledek 9.89. *Pro každý operátor $f : \mathbf{V}\to\mathbf{V}$ na konečně dimenzionálním prostoru $\mathbf{V}$ nad tělesem komplexních čísel $\mathbb{C}$ existuje Jordanův kanonický tvar.*

Důkaz chybějící implikace odložíme na později. Poznamenejme, že Jordanův tvar operátoru je určen jednoznačně v tom smyslu, že jsou-li matice $[f]^B_B$ a $[f]^C_C$ obě v Jordanově tvaru, pak se mohou lišit pouze pořadím Jordanových buněk. To vyplývá z tvrzení 9.87, protože matice je určena vlastními čísly λ operátoru, jejich algebraickou násobností, a dimenzemi podprostorů $\mathrm{Ker}\,(f - \lambda\,\mathrm{id}_V)^l$. Těmto číselným charakteristikám operátoru f říkáme *algebraické invarianty operátoru f*.

Algoritmus pro hledání Jordanova tvaru je možné odvodit z tvrzení 9.87. Obecná diskuze by byla dost nepřehledná, proto ukážeme postup na konktétních příkladech.

9.4.6. *Jordanův tvar v dimenzi* 2. Jediný případ, kdy má operátor f na prostoru dimenze 2 dvě vlastní čísla včetně násobností a není diagonalizovatelný, je případ kdy f má vlastní číslo λ algebraické násobnosti 2 a geometrické násobnosti 1. V tom případě máme jeden Jordanův řetízek délky 2.

Příklad 9.90. Uvažujme operátor f_A na $\mathbb{R}^2$ určený maticí

$$A = \begin{pmatrix} 1 & -1 \\ 4 & -3 \end{pmatrix} .$$

Charakteristický polynom je $p_A(\lambda) = \lambda^2 + 2\lambda + 1 = (\lambda+1)^2$. Operátor f_A má vlastní číslo -1 algebraické násobnosti 2, existuje pro něj proto Jordanův kanonický tvar.

Spočítáme $M_{-1} = \mathrm{Ker}\,(f_A + \mathrm{id})$.

$$M_{-1} = \mathrm{Ker} \begin{pmatrix} 2 & -1 \\ 4 & -2 \end{pmatrix} = \left\langle \begin{pmatrix} 1 \\ 2 \end{pmatrix} \right\rangle .$$

Geometrická násobnost vlastního čísla -1 je 1. Operátor není diagonalizovatelný a budeme proto hledat Jordanův řetízek délky 2

$$\mathbf{v}_2 \xmapsto{f_A+\mathrm{id}} \mathbf{v}_1 \xmapsto{f_A+\mathrm{id}} \mathbf{o} .$$

Vektor $\mathbf{v}_1$ zvolíme jako libovolný nenulový vlastní vektor, například $\mathbf{v}_1 = (1,2)^T$.

Podle tvrzení 9.87 je $\mathrm{Im}\,(f_A + \mathrm{id}) = \mathrm{Ker}\,(f_A + \mathrm{id}) = \langle \mathbf{v}_1 \rangle$, takže speciálně $\mathbf{v}_1 \in \mathrm{Im}\,(f_A + \mathrm{id})$, proto můžeme vždy počáteční vektor $\mathbf{v}_1$ doplnit vektorem $\mathbf{v}_2$ tak, aby platilo $(f_A + \mathrm{id})(\mathbf{v}_2) = \mathbf{v}_1$. Pro takový vektor $\mathbf{v}_2$ platí

$$(A + \mathrm{id})\mathbf{v}_2 = \mathbf{v}_1, \quad \begin{pmatrix} 2 & -1 \\ 4 & -2 \end{pmatrix} \mathbf{v}_2 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \quad \text{zvolíme např. } \mathbf{v}_2 = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$$

Podle věty 9.84 o lineární nezávislosti zobecněných vlastních vektorů je $B = (\mathbf{v}_1, \mathbf{v}_2)$ báze (v takto malém případě to vidíme okamžitě, ve větších dimenzích už to tak zřejmé není). Matice operátoru f_A vzhledem k bázi B je

$$[f_A]^B_B = \begin{pmatrix} -1 & 1 \\ 0 & -1 \end{pmatrix}$$

9.4.7. *Jordanův tvar v dimenzi* 3. Pro nediagonalizovatelný operátor na prostoru dimenze 3 s třemi vlastními čísly včetně násobností můžou nastat následující možnosti.

(1) Operátor f má dvě různá vlastní čísla λ_1, λ_2, kde λ_1 má algebraickou (i geometrickou) násobnost 1 a λ_2 má algebraickou násobnost 2, zatímco geometrickou násobnost 1. V tom případě máme jeden řetízek délky 1 příslušný

λ_1 a jeden řetízek délky 2 příslušný λ_2:

$$\mathbf{v}_1^1 \xmapsto{f-\lambda_1 \mathrm{id}_V} \mathbf{o}$$

$$\mathbf{v}_2^2 \xmapsto{f-\lambda_2 \mathrm{id}_V} \mathbf{v}_1^2 \xmapsto{f-\lambda_2 \mathrm{id}_V} \mathbf{o}\ .$$

(2) Operátor f má vlastní číslo λ algebraické násobnosti 3 a geometrické násobnosti 2. Pak máme dva řetízky příslušné vlastnímu číslu λ a tím pádem nutně jeden z nich má délku 1 a druhý má délku 2:

$$\mathbf{v}_1^1 \xmapsto{f-\lambda \mathrm{id}_V} \mathbf{o}$$

$$\mathbf{v}_2^2 \xmapsto{f-\lambda \mathrm{id}_V} \mathbf{v}_1^2 \xmapsto{f-\lambda \mathrm{id}_V} \mathbf{o}\ .$$

(3) Operátor f má vlastní číslo λ algebraické násobnosti 3 a geometrické násobnosti 1. Pak máme jede řetízek délky 3:

$$\mathbf{v}_3 \xmapsto{f-\lambda \mathrm{id}_V} \mathbf{v}_3 \xmapsto{f-\lambda \mathrm{id}_V} \mathbf{v}_3 \xmapsto{f-\lambda \mathrm{id}_V} \mathbf{o}\ .$$

Příklad 9.91. Uvažujme operátor $f_A : \mathbb{R}^3 \to \mathbb{R}^3$ určený maticí

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

Charakteristický polynom operátoru f_A je $p_A(\lambda) = -\lambda^3 + \lambda^2 + \lambda - 1 = -(\lambda - 1)^2(\lambda + 1)$. Vlastní čísla operátoru A jsou 1 (algebraická násobnost je 2) a -1 (s algebraickou násobností 1), existuje pro něj Jordanův tvar. Příslušné prostory vlastních vektorů jsou

$$M_1 = \mathrm{Ker}\,(f_A - \mathrm{id}) = \mathrm{Ker} \begin{pmatrix} -2 & 0 & 1 \\ 0 & -1 & 0 \\ -4 & 0 & 2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}, M_{-1} = \mathrm{Ker}\,(f_A + \mathrm{id}) = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}\ .$$

Geometrická násobnost vlastního čísla 1 je 1, takže operátor není diagonalizovatelný a Jordanovy řetízky budou tvaru

$$\mathbf{v}_1^1 \xmapsto{f_A + \mathrm{id}} \mathbf{o}$$

$$\mathbf{v}_2^2 \xmapsto{f_A - \mathrm{id}} \mathbf{v}_1^2 \xmapsto{f_A - \mathrm{id}} \mathbf{o}$$

Za vektor $\mathbf{v}_1^1$ zvolíme libovolný nenulový vektor z M_{-1}, např. $\mathbf{v}_1^1 = (0,1,0)^T$. Za vektor $\mathbf{v}_1^2$ zvolíme libovolný nenulový vektor z M_1, např. $\mathbf{v}_1^2 = (1,0,2)^T$, protože (podobně jako v příkladu 9.90) z tvrzení 9.87 plyne, že $\mathbf{v}_1^2 \in \mathrm{Im}\,(f_A - \mathrm{id})$, takže řetízek můžeme doplnit. Řešením soustavy

$$(A - I_3)\mathbf{v}_2^2 = \mathbf{v}_1^1, \quad \text{tj.} \begin{pmatrix} -2 & 0 & 1 \\ 0 & -1 & 0 \\ -4 & 0 & 2 \end{pmatrix} \mathbf{v}_2^2 = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}$$

najdeme například vektor $\mathbf{v}_2^2 = (0,0,1)^T$. Podle věty 9.84 je $B = (\mathbf{v}_1^1, \mathbf{v}_1^2, \mathbf{v}_2^2)$ bází prostoru $\mathbb{R}^3$. Matice operátoru f_A vzhledem k B je

$$[f_A]_B^B = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}\ .$$

Příklad 9.92. Uvažujme operátor $f_A : \mathbb{R}^3 \to \mathbb{R}^3$ určený maticí

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

Charakteristický polynom operátoru f_A vyjde $p_A(\lambda) = -\lambda^3$. Operátor f_A má vlastní číslo 0 algebraické násobnosti 3, existuje pro něj Jordanův tvar. Prostor vlastních vektorů příslušných nule je

$$M_0 = \mathrm{Ker}\,(f_A - 0\,\mathrm{id}) = \mathrm{Ker}\, f_A = \mathrm{Ker}\, A = \left\langle \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \right\rangle .$$

Geometrická násobnost vlastního čísla 0 je 2, proto operátor není diagonalizovatelný, budeme mít dva Jordanovy řetízky tvaru

$$\begin{array}{rl} & \mathbf{v}_1^1 \xmapsto{f_A} \mathbf{o} \\ \mathbf{v}_2^2 \xmapsto{f_A} & \mathbf{v}_1^2 \xmapsto{f_A} \mathbf{o} \,. \end{array}$$

Podle tvrzení 9.87 je $\mathrm{Im}\, f_A = \langle \mathbf{v}_1^2 \rangle$ a $\mathrm{Ker}\, f_A = \langle \mathbf{v}_1^1, \mathbf{v}_1^2 \rangle$. Při hledání vektorů v řetízku postupujeme od počátku nejdelšího řetízku. Vektor $\mathbf{v}_1^2$ zvolíme v $\mathrm{Im}\, f_A$, např. $\mathbf{v}_1^2 = (2,0,1)^T$. Pak je $\mathbf{v}_1^2 \in \mathrm{Ker}\, f_A$. Doplníme $\mathbf{v}_1^2$ na bázi $(\mathbf{v}_1^2, \mathbf{v}_1^1)$ prostoru $\mathrm{Ker}\, f_A$, třeba vektorem $\mathbf{v}_1^1 = (-1,1,0)^T$. Nakonec najdeme $\mathbf{v}_2^2$ tak, aby $f_A(\mathbf{v}_2^2) = \mathbf{v}_1^2$. To musí jít, protože $\mathbf{v}_1^2 \in \mathrm{Im}\, f_A$. Řešením rovnice $A\mathbf{v}_2^2 = \mathbf{v}_1^2$ je například vektor $\mathbf{v}_2^2 = (1,0,0)^T$.

Počátky řetízků tvoří lineárně nezávislou posloupnost (tak jsme je zvolili), takže podle věty 9.84 je $B = (\mathbf{v}_1^1, \mathbf{v}_1^2, \mathbf{v}_2^2)$ báze prostoru $\mathbb{R}^3$. Matice operátoru f_A vzhledem k B je

$$[f_A]_B^B = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} .$$

Příklad 9.93. Uvažujme operátor $f_A : \mathbb{R}^3 \to \mathbb{R}^3$ určený maticí

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

Charakteristický polynom operátoru f_A je $p_A(\lambda) = -(\lambda+1)^3$, máme jedno vlastní číslo -1 algebraické násobnosti 3.

$$M_{-1} = \mathrm{Ker}\,(f_A + \mathrm{id}) = \mathrm{Ker}\,(A + I_3) = \left\langle \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix} \right\rangle .$$

Geometrická násobnost vlastního čísla -1 je 1. Operátor f_A není diagonalizovatelný, existuje pro něj Jordanův tvar a příslušná báze B bude obsahovat jeden řetízek

$$\mathbf{v}_3 \xmapsto{f_A+\mathrm{id}} \mathbf{v}_2 \xmapsto{f_A+\mathrm{id}} \mathbf{v}_1 \xmapsto{f_A+\mathrm{id}} \mathbf{o} \,.$$

Prodle tvrzení 9.87 je $\mathrm{Ker}\,(f_A+\mathrm{id}) = \mathrm{Im}\,(f_A+\mathrm{id})^2 = \langle \mathbf{v}_1 \rangle$, takže za počátek můžeme zvolit libovolný nenulový vektor v tomto prostoru, například $\mathbf{v}_1 = (0,2,1)^T$. Vektor

$\mathbf{v}_2$ musíme zvolit tak, aby $(f_A + \mathrm{id})(\mathbf{v}_2) = \mathbf{v}_1$ a aby ležel v $\mathrm{Im}\,(f_A + \mathrm{id})$, abychom pak mohli nalézt vektor $\mathbf{v}_1$. První podmínka je

$$(A + \mathrm{id})\mathbf{v}_2 = \mathbf{v}_1, \qquad \begin{pmatrix} 0 & 0 & 0 \\ 1 & 2 & -4 \\ 1 & 1 & -2 \end{pmatrix} \mathbf{v}_2 = \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix}$$

Řešením soustavy je $(0,1,0)^T + \left\langle (0,2,1)^T \right\rangle = (0,1,0)^T + \mathrm{Ker}\,(A + \mathrm{id})$. Některý z takových vektorů leží v $\mathrm{Im}\,(f_A + \mathrm{id})$, protože ale $\mathrm{Ker}\,(f_A + \mathrm{id}) \subseteq \mathrm{Im}\,(f_A + \mathrm{id})$ (viz opět tvrzení 9.87), každý z těchto vektorů leží v $\mathrm{Im}\,(f_A + \mathrm{id})$. Druhá podmínka je splněná v tomto případě automaticky a můžeme zvolit třeba $\mathbf{v}_2 = (0,1,0)^T$. (Pokud bychom měli více řetízků, neplatilo by $\mathrm{Ker}\,(f - \lambda\,\mathrm{id}) \subseteq \mathrm{Im}\,(f - \lambda\,\mathrm{id})$, takže volba by nemohla být libovolná.) Nakonec najdeme vektor $\mathbf{v}_1$, aby $(f_A + \mathrm{id})\mathbf{v}_1 = \mathbf{v}_2$. Můžeme vzít například $\mathbf{v}_1 = (-1,1,0)^T$.

Podle věty 9.84 je $B = (\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ báze prostoru $\mathbb{R}^3$. Matice operátoru f_A vzhledem k B je

$$[f_A]_B^B = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{pmatrix} .$$

9.4.8. *Jordanův tvar ve vyšších dimenzích.* Do dimenze 3 je možné o počtu a délkách řetízků rozhodnout pouze z algebraických a geometrických násobností. Stejně je tomu v dimenzi 4 kromě případu, že má operátor vlastní číslo λ algebraické násobnosti 4 a geometrické násobnosti 2. Pak má bázi ze dvou Jordanových řetízků, nevíme ale, jsou-li oba délky 2, nebo jeden z nich délky 1 a druhý délky 3.

Příklad 9.94. Uvažujme operátor $f_A : \mathbb{R}^4 \to \mathbb{R}^4$ pro matici

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

Výpočtem charakteristického polynomu zjistíme, že f_A má jediné vlastní číslo 0 algebraické násobnosti 4. Prostor vlastních vektorů příslušných vlastnímu číslu 0 je

$$M_0 = \left\langle \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} \right\rangle .$$

Geometrická násobnost nuly je 2, takže hledaná báze B, vzhledem ke které je $[f]_B^B$ v Jordanově tvaru, bude spojením dvou řetízků příslušných vlastnímu číslu 0. Nevíme, ale budou-li jejich délky $1,3$ nebo $2,2$. Vypočteme proto ještě jádro operátoru $(f_A - 0\,\mathrm{id})^2$.

$$\mathrm{Ker}\,(f_A - 0\,\mathrm{id})^2 = \mathrm{Ker}\,A^2 = \mathrm{Ker}\,0_{4\times 4} = \mathbb{R}^4 \ .$$

Dimenze jádra operátoru $(f_A + 0\,\mathrm{id})^2$ je o 2 vyšší než dimenze jádra operátoru $(f_A + 0\,\mathrm{id})$, takže podle tvrzení 9.87 budou v B právě 2 řetízky délky alespoň 2.

Tím pádem je B spojením řetízků

$$
\begin{array}{l}
\mathbf{v}_2^1 \overset{f_A}{\longmapsto} \mathbf{v}_1^1 \overset{f_A}{\longmapsto} \mathbf{o} \\
\mathbf{v}_2^2 \overset{f_A}{\longmapsto} \mathbf{v}_1^2 \overset{f_A}{\longmapsto} \mathbf{o} \ .
\end{array}
$$

Protože (opět podle tvrzení 9.87) je $\mathrm{Im}\, f_A = \mathrm{Ker}\, f_A = \langle \mathbf{v}_1^1, \mathbf{v}_1^2 \rangle$, můžeme za $\mathbf{v}_1^1, \mathbf{v}_1^2$ zvolit libovolnou bázi $\mathrm{Ker}\, f_A$, například $\mathbf{v}_1^1 = (0,1,0,1)$, $\mathbf{v}_1^2 = (1,0,1,0)$, a pak nalézt vektory $\mathbf{v}_2^1$ a $\mathbf{v}_2^2$ tak, aby $f_A(\mathbf{v}_2^1) = \mathbf{v}_1^1$ a $f_A(\mathbf{v}_2^2) = \mathbf{v}_1^2$, třeba $\mathbf{v}_2^1 = (0,1,0,0)^T$ a $\mathbf{v}_2^2 = (1,0,0,0)^T$. Pak $B = (\mathbf{v}_1^1, \mathbf{v}_2^1, \mathbf{v}_1^2, \mathbf{v}_2^2)$ je podle věty 9.84 báze prostoru $\mathbf{R}^4$ a platí

$$
[f]_B^B = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \ .
$$

Příklad 9.95. Uvažujme operátor $f_A : \mathbb{R}^4 \to \mathbb{R}^4$ pro matici

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

Charakteristický polynom vyjde $p_A(\lambda) = (\lambda - 4)^4$, operátor f_A má vlastní číslo 4 algebraické násobnosti 4.

$$
M_4 = \mathrm{Ker}\,(f_A - 4\,\mathrm{id}) = \mathrm{Ker}\,(A - 4I_4) = \mathrm{Ker} \begin{pmatrix} 1 & 0 & 0 & -1 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & -1 & 0 \\ 0 & 1 & -1 & 0 \end{pmatrix} = \left\langle \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} \right\rangle \ .
$$

Geometrická násobnost vlastního čísla 4 je 2. Operátor bude mít dva řetízky. Abychom zjistili jejich délky, spočítáme $\mathrm{Ker}\,(f_A - 4\,\mathrm{id})^2$.

$$
\mathrm{Ker}\,(f_A - 4\,\mathrm{id})^2 = \mathrm{Ker} \begin{pmatrix} 1 & -1 & 1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & 1 & -1 \end{pmatrix} = \left\langle \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} \right\rangle \ .
$$

Dimenze je o 1 vyšší než dimenze $\mathrm{Ker}\,(A - 4I_4)$, takže počet řetízků délky alespoň 2 je 1. Hledaná báze B je tedy spojením řetízku délky 1 a řetízku délky 3.

$$
\begin{array}{r}
\mathbf{v}_1^1 \overset{f_A - 4\,\mathrm{id}}{\longmapsto} \mathbf{o} \\
\mathbf{v}_3^2 \overset{f_A - 4\,\mathrm{id}}{\longmapsto} \mathbf{v}_2^2 \overset{f_A - 4\,\mathrm{id}}{\longmapsto} \mathbf{v}_1^2 \overset{f_A - 4\,\mathrm{id}}{\longmapsto} \mathbf{o} \ .
\end{array}
$$

Protože $\mathrm{Im}\,(f - 4\,\mathrm{id})^2 = \langle \mathbf{v}_1^2 \rangle$, zvolíme vektor $\mathbf{v}_1^2$ v tomto prostoru, např. $\mathbf{v}_1^2 = (1,1,1,1)^T$. Je $\mathrm{Ker}\,(f - 4\,\mathrm{id}) = \langle \mathbf{v}_1^1, \mathbf{v}_1^2 \rangle$, doplníme vektor $\mathbf{v}_1^2$ na bázi prostoru $\mathrm{Ker}\,(f - 4\,\mathrm{id})$, například vektorem $\mathbf{v}_1^1 = (1,0,0,1)^T$.

Vektor $\mathbf{v}_2^2$ musíme zvolit tak, aby $(f_A - 4\,\mathrm{id})(\mathbf{v}_2^2) = \mathbf{v}_1^2$ a aby $\mathbf{v}_2^2 \in \mathrm{Im}\,(f_A - 4\,\mathrm{id})$ (druhou podmínku již nemůžeme ignorovat jako v předcházejícím příkladu).

Množina všech řešení soustavy $(A - 4I_4)\mathbf{v}_2^2 = (1,1,1,1)^T$ je

$$\begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \left\langle \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} \right\rangle .$$

Obraz operátoru $f_A - 4\,\mathrm{id}$ je lineární obal sloupců matice $A - 4I_4$, který se rovná $\left\langle (1,1,0,0)^T, (0,0,1,1)^T \right\rangle$. Oběma podmínkám vyhovuje například vektor $\mathbf{v}_2^2 = (1,1,0,0)^T$. Nyní už stačí vzít libovolný vektor $\mathbf{v}_3^2$ tak, aby platilo $(f_A - 4\,\mathrm{id})\mathbf{v}_3^2 = \mathbf{v}_2^2$, např. $\mathbf{v}_3^2 = (1,0,0,0)^T$.

Podle věty 9.84 je $B = (\mathbf{v}_1^1, \mathbf{v}_1^2, \mathbf{v}_2^2, \mathbf{v}_3^2)$ báze. Vzhledem k B má operátor f matici

$$[f]_B^B = \begin{pmatrix} 4 & 0 & 0 & 0 \\ 0 & 4 & 1 & 0 \\ 0 & 0 & 4 & 1 \\ 0 & 0 & 0 & 4 \end{pmatrix}$$

Na příkladu si rozmyslíme postup v ještě vyšší dimenzi.

Příklad 9.96. Operátor $f : \mathbf{V} \to \mathbf{V}$ na prostoru $\mathbf{V}$ dimenze 15 splňuje následující podmínky.

(1) f má vlastní číslo 31 algebraické násobnosti 11 a vlastní číslo 47 algebraické násobnosti 4,
(2) $\dim \mathrm{Ker}\,(f - 31\,\mathrm{id}) = 4$, $\dim \mathrm{Ker}\,(f - 47\,\mathrm{id}) = 2$,
(3) $\dim \mathrm{Ker}\,(f - 31\,\mathrm{id})^2 = 7$, $\dim \mathrm{Ker}\,(f - 47\,\mathrm{id})^2 = 3$,
(4) $\dim \mathrm{Ker}\,(f - 31\,\mathrm{id})^3 = 9$,
(5) $\dim \mathrm{Ker}\,(f - 31\,\mathrm{id})^4 = 11$.

Rozmyslíme si, že existuje báze B, pro které je $[f]_B^B$ v Jordanově tvaru, z jakých blolů se $[f]_B^B$ skládá a jak bychom takovou bázi hledali.

Z první podmínky vidíme, že pro f existuje Jordanův kanonický tvar, celkový počet vektorů v řetízcích příslušným vlastnímu číslu 31 je 11 a celkový počet vektorů v řetízcích příslušných vlastnímu číslu 47 je 4.

Zaměříme se nejprve na vlastní číslo 47 a příslušné řetízky. Z druhé podmínky vyplývá, že řetízky jsou 2, takže zbývají dvě možnosti: délky 1, 3, nebo délky 2,2. Ze třetí podmínky máme $\dim \mathrm{Ker}\,(f - 47\,\mathrm{id})^2 - \dim \mathrm{Ker}\,(f - 47\,\mathrm{id}) = 1$, takže máme právě jeden řetízek délky alespoň 2, což vylučuje první možnost. Řetízky příslušné vlastnímu číslu 47 tedy budou

$$\begin{array}{ccccccc} & & & & \mathbf{v}_1^1 & \xmapsto{f-47\,\mathrm{id}} & \mathbf{o} \\ \mathbf{v}_3^2 & \xmapsto{f-47\,\mathrm{id}} & \mathbf{v}_2^2 & \xmapsto{f-47\,\mathrm{id}} & \mathbf{v}_1^2 & \xmapsto{f-47\,\mathrm{id}} & \mathbf{o} \end{array}$$

Počet řetízků pro vlastní číslo 31 je podle druhé podmínky 4. Ze třetí podmínky máme $\dim \mathrm{Ker}\,(f - 31\,\mathrm{id})^2 - \dim \mathrm{Ker}\,(f - 31\,\mathrm{id}) = 3$, takže tři řetízky mají délku alespoň 2. Možnosti délek jsou $1+2+2+6$, $1+2+3+5$, $1+2+4+4$, $1+3+3+4$. Z předposlední podmínky víme, že počet řetízků délky alespoň 3 je 2 (protože $\dim \mathrm{Ker}\,(f - 31\,\mathrm{id})^3 - \dim \mathrm{Ker}\,(f - 31\,\mathrm{id})^2 = 2$). Zbývají dvě možnosti $1+2+3+5$ a $2+4+4$, první možnost vylučuje poslední podmínka. Řetízky příslušné vlastnímu číslu 31 jsou

$$\begin{array}{ccccccccc}
 & & & & & & \mathbf{v}_1^3 & \xmapsto{f-31\,\mathrm{id}} & \mathbf{o} \\
 & & & & \mathbf{v}_2^4 & \xmapsto{f-31\,\mathrm{id}} & \mathbf{v}_1^4 & \xmapsto{f-31\,\mathrm{id}} & \mathbf{o} \\
\mathbf{v}_4^5 & \xmapsto{f-31\,\mathrm{id}} & \mathbf{v}_3^5 & \xmapsto{f-31\,\mathrm{id}} & \mathbf{v}_2^5 & \xmapsto{f-31\,\mathrm{id}} & \mathbf{v}_1^5 & \xmapsto{f-31\,\mathrm{id}} & \mathbf{o} \\
\mathbf{v}_4^6 & \xmapsto{f-31\,\mathrm{id}} & \mathbf{v}_3^6 & \xmapsto{f-31\,\mathrm{id}} & \mathbf{v}_2^6 & \xmapsto{f-31\,\mathrm{id}} & \mathbf{v}_1^6 & \xmapsto{f-31\,\mathrm{id}} & \mathbf{o}
\end{array}$$

Vzhledem k bázi B složené z těchto řetízků bude

$$[f]_B^B = \mathrm{diag}(J_{47,1}, J_{47,3}, J_{31,1}, J_{31,2}, J_{31,4}, J_{31,4}) \ .$$

Řetízky bychom opět hledali od počátků. Zaměříme se na vlastní číslo 31. Protože $\mathrm{Im}\,(f-31\,\mathrm{id})^3 \cap \mathrm{Ker}\,(f-31\,\mathrm{id}) = \langle \mathbf{v}_1^5, \mathbf{v}_1^6 \rangle$, vektory $\mathbf{v}_1^5, \mathbf{v}_1^6$ bychom zvolili tak, aby tvořili bázi tohoto průniku. Tyto dva vektory bychom doplnili vektorem $\mathbf{v}_1^4$ do báze prostoru $\mathrm{Ker}\,(f-31\,\mathrm{id}) \cap \mathrm{Im}\,(f-31\,\mathrm{id})$. A vektory $\mathbf{v}_1^4, \mathbf{v}_1^5, \mathbf{v}_1^6$ doplnili vektorem $\mathbf{v}_1^3$ do báze prostoru $\mathrm{Ker}\,(f-31\,\mathrm{id})$. Pak bychom pro každý počáteční vektor postupně doplnili zbylé vektory do řetízku. Rozmyslíme si třetí z řetízků. Vektor $\mathbf{v}_2^5$ bychom zvolili tak, aby $(f-31\,\mathrm{id})(\mathbf{v}_2^5) = \mathbf{v}_1^5$ a zároveň $\mathbf{v}_2^5 \in \mathrm{Im}\,(f-31\,\mathrm{id})^2$ (druhá podmínka je nutná, abychom mohli pokračovat). Dále vektor $\mathbf{v}_3^5$ bychom zvolili tak, aby $(f-31\,\mathrm{id})(\mathbf{v}_3^5) = \mathbf{v}_2^5$ a $\mathbf{v}_3^5 \in \mathrm{Im}\,(f-31\,\mathrm{id})$. Konečně vektor $\mathbf{v}_4^5$ bychom zvolili tak, aby $(f-31\,\mathrm{id})(\mathbf{v}_4^5) = \mathbf{v}_3^5$.

Podobně bychom posupovali pro další řetízky a druhé vlastní číslo 47.

9.4.9. *Řešení spojitého dynamického systému.* V odstavci 9.3.5 jsme ukázali, jak vyřešit spojitý dynamický systém $\mathbf{x}'(t) = A\mathbf{x}(t)$ v případě, že A je diagonalizovatelná matice. Řešení spočívalo v tom, že jsme původní soustavu převedli na soustavu $\mathbf{y}'(t) = D\mathbf{y}(t)$, kde D je diagonální matice, a takovou soustavu již umíme řešit. Stejný postup lze použít pro matici podobnou matici v Jordanově tvaru, získáme soustavu $\mathbf{y}'(t) = J\mathbf{y}(t)$, kde J je v Jordanově tvaru.

Zopakujeme tento postup. Předpokládejme, že pro f_A existuje Jordanův kanonický tvar. Pak umíme najít matici J v Jordanově tvaru a regulární matici R takovou, že $J = R^{-1}AR$. (Připomeňme, že R je matice přechodu od báze B tvořené spojením Jordanových řetízků matice A ke kanonické bázi a $J = [f_A]_B^B$.) Úpravou dostaneme $A = RJR^{-1}$ a rovnici můžeme ekvivalentně přepsat

$$\begin{aligned}
\mathbf{x}'(t) &= RJR^{-1}\mathbf{x}(t) \\
R^{-1}\mathbf{x}'(t) &= JR^{-1}\mathbf{x}(t) \ .
\end{aligned}$$

Definujeme $\mathbf{y}(t) = R^{-1}\mathbf{x}(t)$ a dostaneme

$$\mathbf{y}'(t) = J\mathbf{y}(t) \ .$$

Původní funkce $\mathbf{x}(t)$ dopočteme ze vztahu $\mathbf{x}(t) = R\mathbf{y}(t)$.

Stačí tedy umět řešit spojité dynamické systémy tvaru $\mathbf{y}'(t) = J\mathbf{y}(t)$, kde J je matice v Jordanově kanonickém tvaru. Pro diagonální matici J jsme v 9.3.5 ukázali, že řešením

$$\begin{pmatrix} y_1'(t) \\ y_2'(t) \\ \vdots \\ y_n'(t) \end{pmatrix} = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ & & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix} \begin{pmatrix} y_1(t) \\ y_2(t) \\ \vdots \\ y_n(t) \end{pmatrix} .$$

funkce

$$\begin{pmatrix} y_1(t) \\ y_2(t) \\ \vdots \\ y_n(t) \end{pmatrix} = \begin{pmatrix} y_1(0)e^{\lambda_1 t} \\ y_2(0)e^{\lambda_2 t} \\ \vdots \\ y_n(0)e^{\lambda_n t} \end{pmatrix} = \begin{pmatrix} e^{\lambda_1 t} & 0 & \dots & 0 \\ 0 & e^{\lambda_2 t} & \dots & 0 \\ & & \ddots & \vdots \\ 0 & 0 & \dots & e^{\lambda_n t} \end{pmatrix} \begin{pmatrix} y_1(0) \\ y_2(0) \\ \vdots \\ y_n(0) \end{pmatrix} .$$

Ukážeme si řešení v případě Jordanovy buňky řádu 2

$$\begin{pmatrix} y_1'(t) \\ y_2'(t) \end{pmatrix} = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} \begin{pmatrix} y_1(t) \\ y_2(t) \end{pmatrix} .$$

Řešíme tedy spojitý dynamický systém

$$\begin{aligned} y_1'(t) &= \lambda y_1(t) + y_2(t) \\ y_2'(t) &= \lambda y_2(t) \ . \end{aligned}$$

(V řeči operátorů, $y_2(t)$ je vlastním vektorem operátoru derivování, $y_1(t)$ je zobecněným vlastním vektorem.) Z druhé rovnice máme $y_2(t) = y_2(0)e^{\lambda t}$. Dosazením do první rovnice dostáváme

$$y_1'(t) = \lambda y_1(t) + y_2(0)e^{\lambda t} \ ,$$

což přepíšeme do tvaru

$$y_1'(t) - \lambda y_1(t) = y_2(0)e^{\lambda t} \ .$$

Funkci $y_1(t)$ najdeme pomocí jednoduchého triku. Napíšeme si ji jako součin $y_1(t) = u(t)v(t)$ dvou jiných funkcí. Po dosazení do předchozí rovnosti a použití vzorečku pro derivaci součinu dvou funkcí dostaneme

$$u'(t)v(t) + u(t)v'(t) - \lambda u(t)v(t) = y_2(0)e^{\lambda t}$$

neboli

$$u'(t)v(t) + u(t)(v'(t) - \lambda v(t)) = y_2(0)e^{\lambda t} \ .$$

Funkci $v(t)$ zvolíme tak, aby byla závorka na levé straně rovná 0, tj. tak aby platilo $v'(t) = \lambda v(t)$ pro každé $t \in \mathbb{R}$. Jednou z možností je zvolit $v(t) = e^{\lambda t}$. Tím se poslední rovnice redukuje na

$$u'(t)e^{\lambda t} = y_2(0)e^{\lambda t}$$

a tedy

$$u'(t) = y_2(0) \ .$$

Pro libovolnou konstantu d funkce $u(t) = y_2(0)t + d$ splňuje poslední rovnici, takže

$$y_1(t) = u(t)v(t) = (y_2(0)t + d)e^{\lambda t} = y_2(0)te^{\lambda t} + de^{\lambda t}.$$

Dosazením $t = 0$ vyjde $y_1(0) = d$. Dostali jsem tak, že musí platit

$$y_1(t) = y_2(0)te^{\lambda t} + y_1(0)e^{\lambda t} \ .$$

Řešením spojitého dynamického systému

$$\begin{aligned} y_1'(t) &= \lambda y_1(t) + y_2(t) \\ y_2'(t) &= \lambda y_2(t) \end{aligned}$$

jsou tedy funkce

$$\begin{pmatrix} y_1(t) \\ y_2(t) \end{pmatrix} = \begin{pmatrix} y_2(0)te^{\lambda t} + y_1(0)e^{\lambda t} \\ y_2(0)e^{\lambda t} \end{pmatrix} = \begin{pmatrix} e^{\lambda t} & te^{\lambda t} \\ 0 & e^{\lambda t} \end{pmatrix} \begin{pmatrix} y_1(0) \\ y_2(0) \end{pmatrix} .$$

Uvedeným postupem jsme našli všechna možná řešení $(y_1(t), y_2(t))^T$. Jednoznačnost funkce $y_2(t) = y_2(0)e^{\lambda t}$ jsme si ukázali už v tvrzení 9.1. Pokud jde o jednoznačnost funkce $y_1(t)$, můžeme se o ní přesvědčit také přímo podobně jako v tvrzení 9.1. Stačí spočítat, že pro každou funkci $f(t)$ splňující rovnost $f'(t) = \lambda f(t) + y_2(0)e^{\lambda t}$ a $f(0) = y_1(0)$ je derivace funkce $(f(t) - y_2(0)te^{\lambda t})e^{-\lambda t}$ rovná 0.

Příklad 9.97. Vyřešíme spojitý dynamický systém

$$\begin{pmatrix} x_1'(t) \\ x_2'(t) \end{pmatrix} = A \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 4 & -3 \end{pmatrix} \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix}$$

s počáteční podmínkou $x_1(0) = 3, x_2(0) = 4$. V příkladu 9.90 jsme vypočetli, že vzhledem k bázi $B = ((1,2)^T, (0,-1)^T)$ je $[f]_B^B = J_{-1,2}$, tj. platí

$$J = RAR^{-1}, \text{ kde } J = \begin{pmatrix} -1 & 1 \\ 0 & -1 \end{pmatrix} \text{ a } R = [\mathrm{id}]_K^B = \begin{pmatrix} 1 & 0 \\ 2 & -1 \end{pmatrix} .$$

Původní soustavu si připíšeme do tvaru

$$R^{-1} \begin{pmatrix} x_1'(t) \\ x_2'(t) \end{pmatrix} = \begin{pmatrix} -1 & 1 \\ 0 & -1 \end{pmatrix} R^{-1} \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} .$$

Označíme-li

$$\begin{pmatrix} y_1(t) \\ y_2(t) \end{pmatrix} = R^{-1} \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} ,$$

platí

$$\begin{pmatrix} y_1'(t) \\ y_2'(t) \end{pmatrix} = \begin{pmatrix} -1 & 1 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} y_1(t) \\ y_2(t) \end{pmatrix}.$$

Řešením je

$$\begin{pmatrix} y_1(t) \\ y_2(t) \end{pmatrix} = \begin{pmatrix} e^{-t} & te^{-t} \\ 0 & e^{-t} \end{pmatrix} \begin{pmatrix} y_1(0) \\ y_2(0) \end{pmatrix} ,$$

takže

$$\begin{aligned} \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} &= R \begin{pmatrix} y_1(t) \\ y_2(t) \end{pmatrix} = R \begin{pmatrix} e^{-t} & te^{-t} \\ 0 & e^{-t} \end{pmatrix} \begin{pmatrix} y_1(0) \\ y_2(0) \end{pmatrix} \\ &= R \begin{pmatrix} e^{-t} & te^{-t} \\ 0 & e^{-t} \end{pmatrix} R^{-1} \begin{pmatrix} x_1(0) \\ x_2(0) \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} e^{-t} & te^{-t} \\ 0 & e^{-t} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} 3 \\ 4 \end{pmatrix} = \\ &= \begin{pmatrix} 3e^{-t} + 2te^{-t} \\ 4e^{-t} + 4te^{-t} \end{pmatrix} . \end{aligned}$$

Příklad 9.98. Tři chemikálie E, F, G spolu reagují podle schématu

$$E \longrightarrow F \longrightarrow G \ .$$

To znamená, že E se při reakci mění na F a F se mění na G. Rychlost přeměny je přímo úměrná koncentraci, pro jednoduchost bude v naší reakci koeficient úměrnosti rovný 1. Na začátku, v čase $t = 0$, bude přítomná pouze chemikálie E. Zajímá nás, jak se budou koncentrace všech tří chemikálií vyvíjet v čase.

Označme $\mathbf{x}(t) = (x_E(t), x_F(t), x_G(t))^T$ vektor koncentrací v čase t. Z popisu reakce vyplývá, že koncentrace splňují

$$\begin{aligned} x'_E(t) &= -x_E(t) \\ x'_F(t) &= x_E(t) - x_F(t) \\ x'_G(t) &= x_F(t) \ . \end{aligned}$$

Navíc víme, že $\mathbf{x}(0) = (1, 0, 0)^T$. Maticově zapsáno máme

$$\mathbf{x}'(t) = A\mathbf{x}(t), \quad \text{kde } A = \begin{pmatrix} -1 & 0 & 0 \\ 1 & -1 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \ \mathbf{x}(0) = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \ .$$

Nyní již stačí aplikovat probraný postup. Zjistíme, že matice A se rovná

$$A = RJR^{-1}, \ \text{kde } R = \begin{pmatrix} 0 & -1 & 0 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}, \ J = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \ .$$

(R je matice přechodu od báze B tvořené spojením Jordanových řetízků ke kanonické bázi a $J = [f_A]^B_B$.) Označíme $\mathbf{y}(t) = R^{-1}\mathbf{x}(t)$ a původní soustavu přepíšeme do tvaru

$$\mathbf{y}'(t) = J\mathbf{y}(t) = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \mathbf{y}(t) \ .$$

Podle předchozího příkladu dostáváme řešení

$$\mathbf{y}(t) = \begin{pmatrix} e^{-t} & te^{-t} & 0 \\ 0 & e^{-t} & 0 \\ 0 & 0 & e^{0t} \end{pmatrix} \mathbf{y}(0) \ .$$

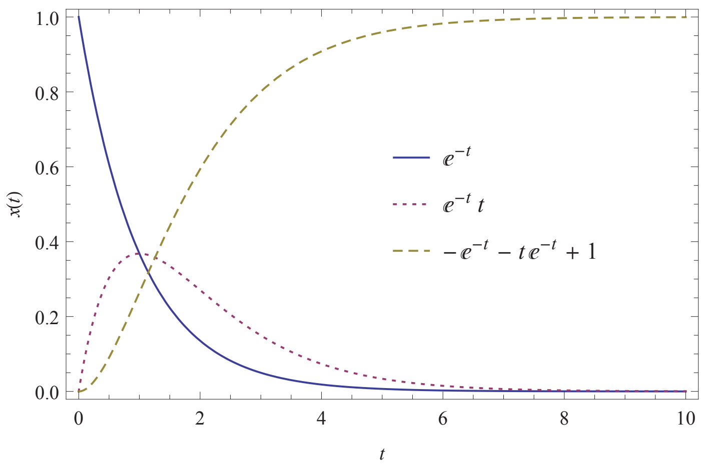
Z toho

$$\mathbf{x}(t) = R \begin{pmatrix} e^{-t} & te^{-t} & 0 \\ 0 & e^{-t} & 0 \\ 0 & 0 & e^{0t} \end{pmatrix} R^{-1}\mathbf{x}(0) = \begin{pmatrix} e^{-t} \\ te^{-t} \\ -e^{-t} - te^{-t} + 1 \end{pmatrix} \ .$$

Koncentrace chemikálií E, F, G v čase t tedy bude $x_E(t) = e^{-t}$, $x_F(t) = te^{-t}$, $x_G(t) = -e^{-t} - te^{-t} + 1$.

Poznamejme, že obecněji pro Jordanovu buňku $J_{\lambda,n}$ jsou řešením spojitého dynamického systému $\mathbf{y}'(t) = J_{\lambda,n}\mathbf{y}(t)$ funkce

$$\begin{pmatrix} y_1(t) \\ y_2(t) \\ \vdots \\ y_{n-1}(t) \\ y_n(t) \end{pmatrix} = \begin{pmatrix} e^{\lambda t} & \frac{t}{1!}e^{\lambda t} & \frac{t^2}{2!}e^{\lambda_t} & \dots & \frac{t^n}{n!}e^{\lambda t} \\ 0 & e^{\lambda t} & \frac{t}{1!}e^{\lambda t} & \dots & \frac{t^{n-1}}{(n-1)!}e^{\lambda t} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & e^{\lambda t} & \frac{t}{1!}e^{\lambda t} \\ 0 & 0 & \dots & 0 & e^{\lambda t} \end{pmatrix} \begin{pmatrix} y_1(0) \\ y_2(0) \\ \vdots \\ y_{n-1}(0) \\ y_n(0) \end{pmatrix} \ .$$



Obrázek 81. Grafy průběhu koncentrací jednotlivých chemikálií.

Stačí k tomu použít indukci podle n a v indukčním kroku stejný trik jako v případě Jordanovy buňky řádu 2.

9.4.10. *Invariantní podprostory.* Invariantní podprostory operátoru f jsou podprostory, které operátor f zachovává v následujícím smyslu.

Definice 9.99. Je-li $f : \mathbf{V} \to \mathbf{V}$ lineární operátor na vektorovém prostoru $\mathbf{V}$, pak podprostor $\mathbf{W} \leq \mathbf{V}$ nazýváme *invariantní podprostor operátoru* f, pokud platí pro každý vektor $\mathbf{x} \in W$, že také $f(\mathbf{x}) \in W$.

Invariantní podprostor čtvercové matice A definujeme jako invariantní podprostor operátoru f_A určeného maticí A.

Příklad 9.100. Každý operátor má dva triviální invariantní podprostory $\{\mathbf{o}\}$ a $\mathbf{V}$.

Z geometrického náhledu vidíme, že rotace v $\mathbb{R}^2$ má pouze triviální invariantní podprostory.

Osová souměrnost v $\mathbb{R}^2$ podle přímky $\langle \mathbf{v} \rangle$ má kromě triviálních podprostorů ještě dva invariantní podprostory: $\langle \mathbf{v} \rangle$ a $\mathbf{v}^\perp$ (ortogonální doplněk je vzhledem ke standardnímu skalárnímu součinu.)

Pro rotaci v $\mathbb{R}^3$ kolem přímky $\langle \mathbf{p} \rangle$ jsou $\langle \mathbf{p} \rangle$ a $\mathbf{p}^\perp$ invariantní podprostory. Rotace o π má ještě další invariantní podprostory.

Každý podprostor prostoru $\mathbf{V}$ je invariantním podprostorem operátoru id a také operátoru $\lambda\,\mathrm{id}$ pro libolný skalár λ.

Tvrzení 9.101. *Pro každý lineární operátor $f : \mathbf{V} \to \mathbf{V}$ jsou následující podprostory $\mathbf{V}$ invariantní podprostory operátoru f:*

(1) $\mathrm{Ker}\,(f)$,
(2) $\mathrm{Im}\,(f)$,
(3) *podprostor $\langle \mathbf{u} \rangle$ generovaný libovolným nenulovým vlastním vektorem $\mathbf{u}$ operátoru f,*
(4) *obecněji, podprostor $\langle \mathbf{v}_1, \dots, \mathbf{v}_k \rangle$ generovaný Jordanovým řetízkem $(\mathbf{v}_1, \dots, \mathbf{v}_k)$ operátoru f příslušným vlastnímu číslu λ.*

Důkaz. Bod (1) je triviální.

Pro důkaz (2) uvažujme libovolný vektor $\mathbf{x} \in \mathbf{V}$. Pak existuje vektor $\mathbf{y} \in \mathrm{Im}\,(f)$ takový, že $f(\mathbf{y}) = \mathbf{x}$. Obrazem vektoru $\mathbf{x}$ je vektor $f(\mathbf{x}) = f(f(\mathbf{y}))$, takže $f(\mathbf{x}) \in \mathrm{Im}\, f$.

Bod (3) je speciálním případem bodu (4).

Pro důkaz (4) uvažujme libovolný vektor $\mathbf{x} = a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k$. Jeho obraz je po úpravě

$$f(\mathbf{x}) = a_1 f(\mathbf{v}_1) + a_2 f(\mathbf{v}_2) + \cdots + a_k f(\mathbf{v}_k) = a_1\lambda\mathbf{v}_1 + a_2(\lambda\mathbf{v}_2 + \mathbf{v}_1) + \cdots + a_k(\lambda\mathbf{v}_k + \mathbf{v}_{k-1}) \ .$$

Výraz na pravé straně jde vyjádřit jako lineární kombinaci vektorů $\mathbf{v}_1, \ldots, \mathbf{v}_k$, takže skutečně $f(\mathbf{x}) \in \langle \mathbf{v}_1, \ldots, \mathbf{v}_k \rangle$. □

Další invariantní podprostory můžeme získat průniky a součty invariantních podprostorů.

Tvrzení 9.102. *Jsou-li $\mathbf{U}$ a $\mathbf{W}$ dva invariantní podprostory operátoru $f : \mathbf{V} \to \mathbf{V}$, pak jsou podprostory $\mathbf{U} \cap \mathbf{W}$ a $\mathbf{U} + \mathbf{W}$ rovněž invariantními podprostory operátoru f.*

Důkaz. Je-li $\mathbf{x} \in U \cap W$, pak $f(\mathbf{x}) \in U$, protože $\mathbf{U}$ je invariatní, a $f(\mathbf{x}) \in W$, protože $\mathbf{W}$ je invariantní. Z toho plyne $f(\mathbf{x}) \in U \cap W$.

Je-li $\mathbf{x} \in U + W$, pak existují vektory $\mathbf{u} \in U$, $\mathbf{w} \in W$ takové, že $\mathbf{x} = \mathbf{u} + \mathbf{w}$. Z invariance U a W víme, že $f(\mathbf{u}) \in U$ a $f(\mathbf{w}) \in W$, proto $f(\mathbf{x}) = f(\mathbf{u} + \mathbf{w}) = f(\mathbf{u}) + f(\mathbf{w}) \in U + W$. □

Z předchozích dvou tvrzení vyplývá, že lineární obal spojení libovolného počtu Jordanových řetízků nějakého operátoru je jeho invariantním podprostorem.

Je-li $\mathbf{W}$ invariantní podprostor operátoru f, pak zúžení $g = f|_W$ operátoru f na podprostor $\mathbf{W}$ je lineární operátor na prostoru $\mathbf{W}$. Je zřejmé, že každé vlastní číslo operátoru $g = f|_W$ je vlastním číslem operátoru f a každý vlastní vektor operátoru g je také vlastním vektorem operátoru f (příslušný stejnému vlastnímu číslu). Dokážeme silnější tvrzení. Metodu důkazu jsme použili už v důkazu věty o tom, že geometrická násobnost libovolného vlastního čísla operátoru f je nejvýše rovná jeho algebraické násobnosti.

Tvrzení 9.103. *Buď $f : \mathbf{V} \to \mathbf{V}$ lineární operátor na konečně dimenzionálním prostoru $\mathbf{V}$ nad tělesem $\mathbf{T}$ a $\mathbf{W} \leq \mathbf{V}$ invariantní podprostor operátoru f. Potom charakteristický polynom zúžení $g = f|_W$ operátoru f na podprostor $\mathbf{W}$ dělí charakteristický polynom operátoru f.*

Důkaz. Zvolme nějakou bázi $C = (\mathbf{v}_1, \ldots, \mathbf{v}_k)$ podprostoru $\mathbf{W}$ a doplňme ji na bázi $B = (\mathbf{v}_1, \ldots, \mathbf{v}_k, \mathbf{v}_{k+1}, \ldots, \mathbf{v}_n)$ prostoru $\mathbf{V}$. Pro každý vektor $\mathbf{v}_j$, $j = 1, \ldots, k$ platí $f(\mathbf{v}_j) \in W$, neboť $\mathbf{W}$ je invariantní podprostor operátoru f. Vyjádření $[f(\mathbf{v}_j)]_B$ vektoru $f(\mathbf{v}_j)$ v bázi B proto bude mít složky $k+1, \ldots, n$ nulové a vektor tvořený prvními k složkami bude rovný $[g(\mathbf{v}_j)]_C$. Matice $[f]_B^B$ operátoru f vzhledem k bázi B má potom blokový tvar

$$[f]_B^B = \begin{pmatrix} A & E \\ 0 & F \end{pmatrix} \ ,$$

kde $A = [g]_C^C$, F je nějaká čtvercová matice řádu $n-k$ a E je matice typu $k \times (n-k)$. Potom

$$[f]_B^B - \lambda I_n = \begin{pmatrix} A - \lambda I_k & E \\ 0 & F - \lambda I_{n-k} \end{pmatrix} \ ,$$

$p_f(\lambda) = \det([f]_B^B - \lambda I_n) = \det(A - \lambda I_k)\det(F - \lambda I_{n-k})$ a $p_g(\lambda) = \det([f]_C^C - \lambda I_n) = \det(A - \lambda I_k)$. Takže $p_g(\lambda)$ skutečně dělí $p_f(\lambda)$. □

Formulujeme důležitý důsledek.

Důsledek 9.104. *Nechť $f : \mathbf{V} \to \mathbf{V}$ je operátor na prostoru $\mathbf{V}$ dimenze n a $\mathbf{W}$ je invariantní podprostor operátoru f dimenze k. Pokud má operátor f právě n vlastních čísel včetně násobností, pak má operátor $g = f|_W : \mathbf{W} \to \mathbf{W}$ právě k vlastních čísel včetně násobností.*

Důkaz. Bez důkazu použijeme tvrzení, které dokážete v kurzu algebry – pokud se polynom rozkládá na lineární faktory, pak se na lineární faktory rozkládá i libovolný jeho dělitel.

Pokud má operátor f právě n vlastních čísel včetně násobností, pak se jeho charakteristický polynom $p_f(\lambda)$ rozkládá na lineární faktory. Polynom $p_g(\lambda)$ podle předchozího tvrzení dělí polynom $p_f(\lambda)$, z toho vyplývá, že se $p_g(\lambda)$ rovněž rozkládá na lineární faktory, operátor g má tedy k vlastních čísel včetně násobností. □

Příklad 9.105. Uvažujme operátor $f = f_A : \mathbb{R}^3 \to \mathbb{R}^3$ určený maticí

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

Ukážeme, že $\mathbf{W} = \langle \mathbf{u}, \mathbf{v} \rangle = \left\langle (0,1,0)^T, (1,1,2)^T \right\rangle$ je jeho invariatní podprostor. Platí $f(\mathbf{u}) = (0,-1,0)^T$ a $f(\mathbf{v}) = (1,-1,2)^T$. Obrazy obou generátorů jsou lineární kombinace vektorů $\mathbf{u}, \mathbf{v}$:

$$f(\mathbf{u}) = -\mathbf{u}, \quad f(\mathbf{v}) = -2\mathbf{u} + \mathbf{v}\ .$$

Z toho vyplývá, že každý vektor z W se zobrazí do W: Je-li totiž $\mathbf{x} = a\mathbf{u} + b\mathbf{v}$, pak $f(\mathbf{x}) = af(\mathbf{u}) + bf(\mathbf{v})$. Podprostor $\mathbf{W}$ je tedy invariantní podprostor operátoru f. (Operátor f_A je shodný s operátorem v příkladu 9.91, podprostor $\mathbf{W}$ je rovný lineárnímu obalu vlastních vektorů.)

Podívejme se ještě na operátor $g = f|_W$. Jeho matice vzhledem k bázi $C = (\mathbf{u}, \mathbf{v})$ je

$$[g]_C^C = \begin{pmatrix} -1 & -2 \\ 0 & 1 \end{pmatrix}$$

Charakteristický polynom operátoru g je $p_g(\lambda) = (\lambda - 1)(\lambda + 1)$ a příslušné vlastní podprostory jsou

$$[M_1]_C = \left\langle \begin{pmatrix} -1 \\ 1 \end{pmatrix} \right\rangle, \quad [M_{-1}]_C = \left\langle \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\rangle\ ,$$

neboli

$$M_1 = \langle -u + v \rangle = \left\langle (1,0,2)^T \right\rangle, \quad M_{-1} = \langle u \rangle = \left\langle (0,1,0)^T \right\rangle$$

Matice operátoru g vzhledem k bázi $D = ((1,0,2)^T, (0,1,0)^T)$ je

$$[g]_D^D = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\ .$$

Geometricky, operátor g je reflexe podle přímky $\left\langle (1,0,2)^T \right\rangle$ ve směru přímky $\left\langle (0,1,0)^T \right\rangle$. To nám dává představu, jak operátor f „vypadá“ v rovině $\mathbf{W}$.

Pro ilustraci předchozího tvrzení ještě uveďme, že $p_f(\lambda) = -(\lambda - 1)^2(\lambda + 1)$. Polynom $p_g(\lambda)$ skutečně tělí polynom $p_f(\lambda)$.

Na závěr si ještě všimneme, že množina operátorů, pro které je daný podprostor $\mathbf{W}$ prostoru $\mathbf{V}$ invariantní, je uzavřená na sčítání a násobení skalárem.

Tvrzení 9.106. *Nechť* $\mathbf{V}$ *je vektorový prostor,* $\mathbf{W}$ *jeho podprostor,* f, g *lineární operátory na* $\mathbf{V}$ *a* $t \in T$. *Pak platí:*

(1) *Je-li* $\mathbf{W}$ *invariantní podprostor operátorů* f *i* g, *pak je* $\mathbf{W}$ *invariantní podprostor operátoru* $f + g$.
(2) *Je-li* $\mathbf{W}$ *invariantní podprostor operátoru* f, *pak je* $\mathbf{W}$ *invariantní podprostor operátoru* tf.

Důkaz. (1). Je-li $\mathbf{x} \in W$ a $f(\mathbf{x}) \in W$, $g(\mathbf{x}) \in W$, pak $(f+g)(\mathbf{x}) = f(\mathbf{x}) + g(\mathbf{x}) \in W$.
(2). Je-li $\mathbf{x} \in W$ a $f(\mathbf{x}) \in W$, pak $(tf)(\mathbf{x}) = t(f(\mathbf{x})) \in W$. □

Například, je-li $\mathbf{W}$ invariatní podprostor operátoru f, pak je také invariantním podprostorem operátoru $f - \lambda\,\mathrm{id}$ pro libovolné $\lambda \in T$.

9.4.11. *Důkaz věty o Jordanově kanonickém tvaru.* Nyní dokážeme chybějící implikaci ve větě 9.88 o Jordanově kanonickém tvaru. Předpokládejme, že $\mathbf{V}$ je konečně generovaný prostor dimenze n a $f : \mathbf{V} \to \mathbf{V}$ je lineární operátor, který má n vlastních čísel včetně algebraických násobností. Chceme dokázat, že pro operátor f existuje Jordanův kanonický tvar.

Větu dokážeme tak, že najdeme bázi $\mathbf{V}$, která je spojením Jordanových řetízků operátoru $\mathbf{V}$. Postupovat budeme indukcí podle dimenze n.

Je-li $n = 1$, matice f vzhledem k jakékoliv bázi B prostoru $\mathbf{V}$ má řád 1 a je tedy Jordanovo buňkou a báze B je tvořena jedním Jordanovým řetízkem délky 1. Předpokládejme, že $n > 1$ a že tvrzení platí pro všechna menší n.

Označme λ libovolné vlastní číslo operátoru f a pro přehlednost označme $g = f - \lambda\,\mathrm{id}$. Pak $\dim(\mathrm{Ker}\, g) > 0$ (protože prostor $\mathrm{Ker}\, g$ je tvořen vlastními vektory operátoru f příslušnými vlastními číslu λ) a podle věty o dimenzi jádra a obrazu je $\dim(\mathrm{Im}\, g) = n - \dim(\mathrm{Ker}\, g) < n$.

Podprostor $\mathrm{Im}\, g$ je podle tvrzení 9.101 invariantním podprostorem operátoru g, takže také operátoru $f = g + \lambda\,\mathrm{id}$ (viz tvrzení 9.106). Charakteristický polynom zúžení h operátoru f na $\mathrm{Im}\, g$ dělí charakteristický polynom operátoru f, a ten má n vlastních čísel včetně násobností. Podle důsledku 9.104 má operátor h $\dim \mathrm{Im}\, g$ vlastních čísel čísel včetně násobností, takže na prostor $\mathrm{Im}\, g$ můžeme použít indukční předpoklad. Existuje tedy báze C prostoru $\mathrm{Im}\, g$, která je složením Jordanových řetízků operátoru h (ty jsou samořejmě rovněž Jordanovými řetízky operátoru f). Jordanovy řetízky příslušné vlastnímu číslu λ označíme podle schématu

$$\begin{array}{c}
\mathbf{v}^1_{k_1} \overset{g}{\longmapsto} \dots \overset{g}{\longmapsto} \mathbf{v}^1_2 \overset{g}{\longmapsto} \mathbf{v}^1_1 \overset{g}{\longmapsto} \mathbf{o} \\
\vdots \\
\mathbf{v}^r_{k_r} \overset{g}{\longmapsto} \dots \overset{g}{\longmapsto} \mathbf{v}^r_2 \overset{g}{\longmapsto} \mathbf{v}^r_1 \overset{g}{\longmapsto} \mathbf{o}
\end{array}$$

(V bázi C mohou být ještě řetízky příslušné jiným vlastním číslům.) Počáteční vektory $\mathbf{v}^1_1, \dots, \mathbf{v}^r_1$ tvoří lineárně nezávislou posloupnost v $\mathrm{Ker}\, g$, doplníme tyto vektory na bázi $(\mathbf{v}^1_1, \dots, \mathbf{v}^s_1)$ prostoru $\mathrm{Ker}\, g$. Pro každé $i = 1, \dots, r$ leží koncový vektor $\mathbf{v}^i_{k_i}$ v prostoru $\mathrm{Im}\, g$, existují proto vektory $\mathbf{v}^i_{k_i+1}$ takové, že $g(\mathbf{v}^i_{k_i+1}) = \mathbf{v}^i_{k_i}$.

Tím nám vznikne soubor řetízků

$$\begin{array}{ccccccccccc}
\mathbf{v}^1_{k_1+1} & \stackrel{g}{\longmapsto} & \mathbf{v}^1_{k_1} & \stackrel{g}{\longmapsto} & \dots & \stackrel{g}{\longmapsto} & \mathbf{v}^1_2 & \stackrel{g}{\longmapsto} & \mathbf{v}^1_1 & \stackrel{g}{\longmapsto} & \mathbf{o} \\
 & & & & & & & & & & \vdots \\
\mathbf{v}^r_{k_r+1} & \stackrel{g}{\longmapsto} & \mathbf{v}^r_{k_r} & \stackrel{g}{\longmapsto} & \dots & \stackrel{g}{\longmapsto} & \mathbf{v}^r_2 & \stackrel{g}{\longmapsto} & \mathbf{v}^r_1 & \stackrel{g}{\longmapsto} & \mathbf{o} \\
 & & & & & & & & \mathbf{v}^{r+1}_1 & \stackrel{g}{\longmapsto} & \mathbf{o} \\
 & & & & & & & & & & \vdots \\
 & & & & & & & & \mathbf{v}^s_1 & \stackrel{g}{\longmapsto} & \mathbf{o}
\end{array}$$

plus řetízky v bázi C, které přísluší jiným vlastním číslům. Zkonstruovali jsme novou posloupnost B, která je spojením Jordanových řetízků operátoru f. Zbývá ukázat, že B je báze.

Podle věty 9.84 je B lineárně nezávislá, protože počáteční vektory příslušné vlastnímu číslu λ tvoří z konstrukce lineárně nezávislou posloupnost a pro jiná vlastní čísla jsme nic nezměnili. Nyní stačí spočítat, že počet vektorů v B je n. V bázi C je $\dim \operatorname{Im} g$ vektorů k nim jsme doplnili $\dim \operatorname{Ker} g - r$ vektorů z $\operatorname{Ker} g$ a poté jsme k existujícím řetízkům doplnili r vektorů, ke každému z r řetízků jeden. Dohromady je v B $\dim \operatorname{Im} g + (\dim \operatorname{Ker} g - r) + r = \dim \operatorname{Im} g + \dim \operatorname{Ker} g = n$ vektorů. Tím je důkaz ukončen.

9.4.12. *Cayleyho-Hamiltonova věta.* Uvažujme čtvercovou matici A řádu n nad tělesem $\mathbf{T}$ (nebo operátor f na prostoru $\mathbf{V}$ dimenze n). Posloupnost matic

$$(I_n, A, A^2, A^3, \dots, A^{n^2})$$

(resp. operátorů $(\mathrm{id}, f, f^2, \dots, f^{n^2})$) je lineárně závislá posloupnost v prostoru $\mathbf{T}^{n\times n}$ (resp. $\operatorname{Hom}(\mathbf{V},\mathbf{V})$), protože tento prostor má dimenzi n^2. Existují proto skaláry $a_0, a_1, \dots$ takové, že

$$a_0 I_n + a_1 A + a_2 A^2 + \dots + a_{n^2} A^{n^2} = 0_{n\times n}$$

(resp. $a_0 \,\mathrm{id} + a_1 f + \dots + a_{n^2} f^{n^2} = 0$). Cayleyho-Hamiltonova věta říká, že taková závislost existuje mnohem dříve – stačí prvních $n+1$ členů posloupnosti, přičemž za koeficienty lze vzít koeficenty charakteristického polynomu matice A (resp. polynomu f). Zhruba řečeno, každá matice (resp. každý operátor) je „kořenem“ svého charakteristického polynomu.

Definujeme dosazení matice (operátoru) do polynomu.

Definice 9.107. Nechť $\mathbf{T}$ je těleso, $p(t) = a_0 + a_1 t + a_2 t^2 + \dots + a_n t^n$ polynom s koeficienty $a_0, \dots, a_n$ v $\mathbf{T}$, A čtvercová matice řádu k nad $\mathbf{T}$ a f lineární operátor na prostoru $\mathbf{V}$ nad tělesem $\mathbf{T}$. *Dosazením matice A do polynomu $p(t)$* rozumíme matici

$$p(A) = a_0 I_k + a_1 A + a_2 A^2 + \dots + a_n A^n \ .$$

Doszením operátoru f do polynomu $p(t)$ rozumíme operátor

$$p(f) = a_0 \,\mathrm{id}_V + a_1 f + a_2 f^2 + \dots + a_n f^n \ .$$

Příklad 9.108. Je-li f operátor na $\mathbf{V}$, pak dosazením operátoru f do polynomu $p(t) = t - 3$ je operátor $p(f) = f - 3\,\mathrm{id}$.

Příklad 9.109. Uvažujme reálnou matici

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

Její charakteristický polynom je

$$p_A(\lambda) = \lambda^2 - 5\lambda - 2 \ .$$

Dosazením matice A do tohoto polynomu získáme matici

$$p_A(A) = A^2 - 5A - 2I_2 = \begin{pmatrix} 7 & 15 \\ 10 & 22 \end{pmatrix} + \begin{pmatrix} -5 & -15 \\ -10 & -20 \end{pmatrix} + \begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix} = 0_{2\times 2} \ .$$

Před důkazem Cayleyho-Hamiltonovy věty si všimneme, že dosazování do součinu polynomů lze provádět po jednotlivých činitelích. Je-li $p(t) = p_1(t)p_2(t)\dots p_i(t)$, pak $p(A) = p_1(A)p_2(A)\dots p_i(A)$. Důvodem je, že při roznásobení maticového výrazu $p_1(A)\dots p_i(A)$ je koeficient u A^j stejný jako koeficient u t^j při roznásobování výrazu $p_1(t)p_2(t)\dots p_i(t)$ (pro každé $j \in \{0,\dots,i\}$). Podobně pro operátory $p(f) = p_1(f)p_2(f)\dots p_i(f)$.

Věta 9.110 (Cayleyho-Hamiltonova věta)**.** *Je-li f lineární operátor na konečně generovaném prostoru $\mathbf{V}$ dimenze n nad tělesem $\mathbf{T}$ (resp. je-li A čtvercová matice řádu n nad $\mathbf{T}$), pak $p_f(f) = 0$ (resp. $p_A(A) = 0$).*

Důkaz. Dokážeme si operátorovou verzi, maticovou přenecháme čtenáři. Větu dokážeme pouze v případě, že f má n vlastních čísel včetně násobností. V případě, že tomu tak není, je nutné napřed rozšířit těleso $\mathbf{T}$ do většího tělesa tak, aby v tom větším tělese měl charakteristický polynom dostatek kořenů. To lze udělat vždy a bude to v kursu algebry ve druhém ročníku.

Označme $\lambda_1,\dots,\lambda_m$ vlastní čísla operátoru f a $l_1,\dots,l_m$ jejich násobnosti. Podle předpokladu je $l_1 + \dots + l_m = n$ a charakteristický polynom je proto

$$p_f(\lambda) = (-1)^n(\lambda-\lambda_1)^{l_1}(\lambda-\lambda_2)^{l_2}\cdots(\lambda-\lambda_m)^{l_m} \ .$$

Podle věty 9.88 o Jordanově kanonickém tvaru existuje báze B taková, že $J = [f]^B_B$ je v Jordanově tvaru. Podle pozorování nad větou platí

$$p_f(f) = (-1)^n(f-\lambda_1\,\mathrm{id})^{l_1}(f-\lambda_2\,\mathrm{id})^{l_2}\cdots(f-\lambda_m\,\mathrm{id})^{l_m}$$

a tedy

$$\begin{aligned}
[p_f(f)]^B_B &= [(-1)^n(f-\lambda_1\,\mathrm{id})^{l_1}(f-\lambda_2\,\mathrm{id})\cdots(f-\lambda_m\,\mathrm{id})^{l_m}]^B_B \\
&= (-1)^n([f-\lambda_1\,\mathrm{id}]^B_B)^{l_1}\cdots([f-\lambda_m\,\mathrm{id}]^B_B)^{l_m} \\
&= (-1)^n(J-\lambda_1 I_n)^{l_1}\cdots(J-\lambda_m I_n)^{l_m} \ .
\end{aligned}$$

Matice v součinu jsou blokově diagonální (bloky odpovídají Jordanovým buňkám matice J), můžeme je násobit po blocích. Uvažujme libovolný blok K. Ten odpovídá nějaké Jordanově buňce $J_{\lambda_i,k}$, přičemž k je nejvýše l_i, protože velikost žádné buňky příslušné vlastnímu číslu λ_i nemůže být větší než jeho algebraická násobnost (viz tvrzení 9.87). Pak je ale $(J-\lambda_i I_n)^{l_i} = J^{l_i}_{0,k_i}$ nulová matice podle tvrzení 9.75, takže v celém součinu bude blok K nulový. Dokázali jsme, že $[p_f(f)]^B_B = 0_{n\times n}$, takže $p_f(f) = 0$. □

Příklad 9.111. Ukážeme si použití Cayleyho-Hamiltonovy věty v teorii řízení. Diskrétní dynamický systém $\mathbf{x}_{k+1} = A\mathbf{x}_k$ dimenze n nad tělesem $\mathbf{T}$ s počáteční podmínkou $\mathbf{x}_0 = \mathbf{o}$ má řešení $\mathbf{x}_k = \mathbf{o}$ pro každé k, systém zůstává stále v počátečním stavu.

Přidáme si k němu možnost lineárního „řízení“

$$\mathbf{x}_{k+1} = A\mathbf{x}_k + B\mathbf{u}_k \ ,$$

kde $B = (\mathbf{b}_1|\mathbf{b}_2|\cdots,\mathbf{b}_n)$ je matice stejného řádu jako A. Můžeme si ji představit jako „joystick“ nebo „knipl“, kterým systém uvedeme do pohybu a pak jej řídíme volbou vstupů $\mathbf{u}_0, \mathbf{u}_1, \dots, \mathbf{u}_k, \dots$. Chceme vědět, jakých stavů $\mathbf{x}_k$ můžeme dosáhnout v čase k.

Pro $k = 1$ dostáváme

$$\mathbf{x}_1 = A\mathbf{x}_0 + B\mathbf{u}_0 = B\mathbf{u}_0 \in \operatorname{Im} B$$

a protože vstup $\mathbf{u}_0$ můžeme zvolit libovolně, tvoří možné stavy v čase $k = 1$ sloupcový prostor $\operatorname{Im} B$ „řídící“ matice B.

Pro $k = 2$ dostáváme

$$\mathbf{x}_2 = A\mathbf{x}_1 + B\mathbf{u}_1 = AB\mathbf{u}_0 + B\mathbf{u}_1 \in \operatorname{Im}(AB|B)$$

a protože vstupy $\mathbf{u}_0, \mathbf{u}_1 \in \mathbf{T}^n$ můžeme volit libovolně, tvoří možné stavy v čase $k = 2$ celý sloupcový prostor $\operatorname{Im}(AB|B)$ matice $(AB|B)$ typu $n \times (2n)$.

Jednoduchou indukcí podle k odvodíme, že v čase k tvoří možné stavy $\mathbf{x}_k$ sloupcový prostor

$$\operatorname{Im}(A^{k-1}B|A^{k-2}B|\cdots|AB|B)$$

matice $(A^{k-1}B|A^{k-2}B|\cdots|AB|B)$ typu $n \times (nk)$.

Porovnáme množiny možných stavů v časech $k = n$ a $k = n + 1$, tj. sloupcové prostory

$$\operatorname{Im}(A^{n-1}B|A^{n-2}B|\cdots|AB|B), \quad \operatorname{Im}(A^n|A^{n-1}B|A^{n-2}B|\cdots|AB|B) \ .$$

Zřejmě platí

$$\operatorname{Im}(A^{n-1}B|A^{n-2}B|\cdots|AB|B) \subseteq \operatorname{Im}(A^nB|A^{n-1}B|\cdots|AB|B) \ ,$$

protože každý sloupec matice vlevo je mezi sloupci matice vpravo. Podle Cayleyho-Hamiltonovy věty můžeme matici A^n vyjádřit jako lineární kombinaci

$$A^n = c_{n-1}A^{n-1} + \cdots + c_1A + c_0I_n$$

pro nějaké koeficienty c_i, které se rovnají $+/-$ koeficientům charakteristického polynomu matice A, protože koeficient u λ^n v charakteristickém polynomu $p_A(\lambda)$ je $(-1)^n$. V každém případě koeficienty c_i leží v $\mathbf{T}$. Poslední rovnost přenásobíme zprava maticí B a dostaneme

$$A^nB = c_{n-1}A^{n-1}B + \cdots + c_1AB + c_0B \ .$$

Každý sloupec $A^n\mathbf{b}_j$ matice A^nB, tj. každý nový sloupec matice $(A^nB|A^{n-1}B|A^{n-2}B|\cdots|AB|B)$ se tedy rovná lineární kombinaci

$$A^n\mathbf{b}_j = c_{n-1}A^{n-1}\mathbf{b}_j + \cdots + c_1A\mathbf{b}_j + c_0\mathbf{b}_j$$

nějakých sloupců v matici $(A^{n-1}B|A^{n-2}B|\cdots|AB|B)$. Každý nový sloupec matice $(A^nB|A^{n-1}B|\cdots|AB|B)$ proto už leží ve sloupcovém prostoru matice $\operatorname{Im}(A^{n-1}B|A^{n-2}B|\cdots|AB|B)$.

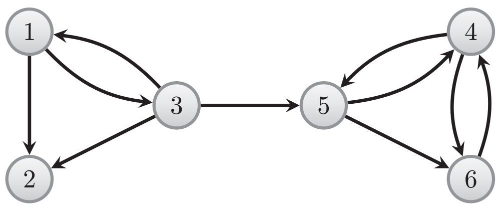
Dostali jsme tak, že každého možného stavu $\mathbf{x}_{n+1}$ už můžeme dosáhnout po $k = n$ krocích, rovná se nějakému z možných stavů $\mathbf{x}_n$.

9.5. **Google.** Ukážeme si jednu moderní aplikaci vlastních čísel a vlastních vektorů. Myšlenku uspořádání webových stránek podle důležitosti si napřed předvedeme na jednoduchém příkladu. Poté odvodíme obecnou formulaci problému.

Představme si malou síť šesti webových stránek, které na sebe odkazují. Odkazy si zapíšeme do matice $A = (a_{ij})$, kde $a_{ij} = 1$ právě když stránka j odkazuje na stránku i. Naše síť je zadána maticí

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

Protože $a_{21} = 1$, stránka 1 odkazuje na stránku 2. Dále $a_{23} = 1$, také stránka 3 odkazuje na stránku 2. Žádná jiná stránka na stránku 2 neodkazuje. Takto si můžeme nakreslit graf sítě.



Obrázek 82. Google

Z vrcholu j vede šipka do i právě když stránka j odkazuje na stránku i. Matice A je tak maticí incidence grafu sítě. Z prvního semestru víme, že prvek na místě (i, j) v mocnině A^k říká, kolik orientovaných cest délky k vede z vrcholu j do vrcholu i.

Základní myšlenka vyhledávače Google spočívá v tom, že měří důležitost stránky pravděpodobností, s jakou se na stránku dostaneme náhodným klikáním. Důležitosti stránky se dopracujeme tak, že na začátku přiřadíme všem stránkám stejnou důležitost 1/6. Počáteční aproximací vektoru důležitosti stránek tak bude vektor $\mathbf{r}_0 = (1/6, 1/6, 1/6, 1/6, 1/6, 1/6)^T$, i-tá složka je důležitost i-té stránky.

Nyní musíme matici incidence webu upravit tak, aby její hodnoty říkali, s jakou pravděpodobností klikneme na link ze stránky j na stránku i. Pokud ze stránky j vede více odkazů, řekněme k, pak na každý z nich klikneme s pravděpodobností $1/k$. Matici A si upravíme tak, že každou jednotku v j-tém sloupci nahradíme číslem $1/k$, kde k je počet prvků rovných 1 v j-tém sloupci matice A. Dostaneme tak matici

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

Všechny prvky matice H jsou nezáporné a součet každého sloupce se rovná buď 1 nebo 0. Druhá možnost nastane v případě, že z příslušné stránky nevede žádný odkaz. Jako třeba ze stránky s pdf souborem těchto přednášek.

První iteraci vektoru důležitosti stránek v naší síti pak získáme jako $\mathbf{r}_1 = H\mathbf{r}_0$. Složka i tohoto vektoru říká, s jakou pravděpodobností se na stránku i dostaneme z náhodně vybrané stránky po jednom kliknutí. Platí

$$\mathbf{r}_1 = H\mathbf{r}_0 = \begin{pmatrix} 0 & 0 & 1/3 & 0 & 0 & 0 \\ 1/2 & 0 & 1/3 & 0 & 0 & 0 \\ 1/2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1/2 & 1 \\ 0 & 0 & 1/3 & 1/2 & 0 & 0 \\ 0 & 0 & 0 & 1/2 & 1/2 & 0 \end{pmatrix} \begin{pmatrix} 1/6 \\ 1/6 \\ 1/6 \\ 1/6 \\ 1/6 \\ 1/6 \end{pmatrix} = \begin{pmatrix} 1/18 \\ 5/36 \\ 1/12 \\ 1/4 \\ 5/36 \\ 1/6 \end{pmatrix}.$$

Druhou iteraci vektoru důležitosti $\mathbf{r}_2$ dostaneme jako $H\mathbf{r}_1$. Můžeme ji slovně popsat tak, že uvádí, s jakou pravděpodobností se na i-tou stránku dostaneme jedním kliknutím z nějaké stránky, přičemž počáteční stránky volíme s pravděpodobnostmi danými vektorem $\mathbf{r}_1$. Stránka je tedy tím „důležitější“, čím „důležitější“ stránky na ni odkazují. Vyjde

$$\mathbf{r}_2 = H\mathbf{r}_1 = \begin{pmatrix} 0 & 0 & 1/3 & 0 & 0 & 0 \\ 1/2 & 0 & 1/3 & 0 & 0 & 0 \\ 1/2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1/2 & 1 \\ 0 & 0 & 1/3 & 1/2 & 0 & 0 \\ 0 & 0 & 0 & 1/2 & 1/2 & 0 \end{pmatrix} \begin{pmatrix} 1/18 \\ 5/36 \\ 1/12 \\ 1/4 \\ 5/36 \\ 1/6 \end{pmatrix} = \begin{pmatrix} 1/36 \\ 1/18 \\ 1/36 \\ 17/72 \\ 11/72 \\ 14/72 \end{pmatrix}.$$

Hledání vektoru důležitosti jednotlivých stránek tak vede na diferenční rovnici $\mathbf{r}_k = H\mathbf{r}_{k-1}$, která jak víme má řešení $\mathbf{r}_k = H^k\mathbf{r}_0$. Tento vektor můžeme interpretovat tak, že udává, s jakou pravděpodobností se dostaneme na danou stránku po k náhodných kliknutích.

Pro porovnávání důležitosti všech webových stránek bychom museli uvažovat matici celého webu, tedy matici řádu n, kde n je číslo v současnosti větší než třicet miliard. Každá iterace navíc vyžaduje spočítat součin matice tohoto řádu s jedním n-složkovým vektorem, počet aritmetických operací je tak řádu n^2. To všechno se zdá být zhola nemožné. Nicméně matice H je velmi řídká, naprostá většina jejích prvků se rovná 0. Pro ty jsou vypracované efektivní metody ukládání. Dále v každém sloupci matice H je v průměru 10 odkazů na jiné stránky, aspoň tak je jejich počet odhadován. Takže součin matice s vektorem vyžaduje pouze $10n$ operací. A to už je v současnosti výpočetně zvládnutelné.

Popsaná diferenční rovnice vyvolává řadu důležitých otázek:

- Konverguje posloupnost vektorů $\mathbf{r}_k$ k nějakému vektoru nebo je celý proces nestabilní?
- Může se stát, že posloupnost vektorů osciluje kolem několika různých limitních vektorů?
- Za jakých podmínek na matici H proces konverguje k jedinému vektoru?
- Pokud konverguje, dává výsledný limitní vektor dobrou míru důležitosti jednotlivých webových stránek?
- Závisí konvergence na počáteční aproximaci $\mathbf{r}_0$?
- Pokud proces konverguje, kolik iterací musíme provést, abychom dostali dobrou aproximaci limitního vektoru?

Už při prvním hraní si s naším malých příkladem zjistíme jeden problém tohoto přístupu. Díky tomu, že v našem příkladu ze stránky 2 nevede žádný odkaz, důležitost této stránky se nijak neprojeví na důležitosti jiných stránek. Na druhou stranu při každé iteraci do sebe nasaje něco z důležitosti jiných stránek a celková suma důležitostí všech stránek se postupně snižuje. Stránkou 2 tak důležitost „odtéká“. Mnohem závažnější je skutečnost, že klastr stránek 4,5,6 odkazuje pouze na stránky 4,5,6, a žádná z nich neodkazuje na žádnou ze stránek 1,2,3, zatímco stránka 3 odkazuje na stránku 5 z tohoto klastru. Klastr stránek 4,5,6 tak bude akumulovat důležitost stránek z celé sítě. Skutečně, již třináctá iterace $\mathbf{r}_{13}$ má první tři složky zanedbatelně malé a zbylé tři složky v poměru $(2/3) : (1/3) : (1/5)$.

Problém se stránkami, ze kterých nevede žádný odkaz, vyřešíme předpokladem, že z takové stránky můžeme náhodně přeskočit na jakoukoliv jinou stránku, na všechny se stejnou pravděpodobností. V našem malém příkladu je takovou stránkou stránka 2, nulový sloupec v matici H nahradíme sloupcem ze samých hodnot $1/6$. Dostaneme tak matici

$$S = \begin{pmatrix} 0 & 1/6 & 1/3 & 0 & 0 & 0 \\ 1/2 & 1/6 & 1/3 & 0 & 0 & 0 \\ 1/2 & 1/6 & 0 & 0 & 0 & 0 \\ 0 & 1/6 & 0 & 0 & 1/2 & 1 \\ 0 & 1/6 & 1/3 & 1/2 & 0 & 0 \\ 0 & 1/6 & 0 & 1/2 & 1/2 & 0 \end{pmatrix} .$$

V obecném případě bychom matici H nahradili maticí

$$S = H + \frac{1}{n}\mathbf{e}\mathbf{a}^T,$$

kde $\mathbf{e}$ je sloupcový vektor se všemi složkami rovnými 1 a $\mathbf{a}$ je vektor, jehož j-tá složka je rovna 1, pokud z j-té stránky nevede žádný odkaz, a rovná se 0, pokud z j-té stránky nějaký odkaz na jinou stránku vede. Matice S je markovovská matice, to znamená, že její prvky jsou nezáporné a každý sloupec má součet rovný 1. O takových maticích už víme, že číslo 1 je jejich vlastním číslem.

Problém klastru stránek, které akumulují důležitost všech ostatních stránek, touto úpravou nevyřešíme. V našem příkladu bude pořád platit, že mezi klastrem stránek 1,2,3 a klastrem stránek 4,5,6 vedou odkazy pouze jednosměrně, ze stránek 1,2,3 na stránky 4,5,6. Naše brouzdání po webu upravíme ještě jedním způsobem. Zvolíme si nějaké číslo $\alpha \in (1/2, 1)$. Toto číslo je pravděpodobnost, se kterou volíme následující krok při prohlížení webu tak, že klikneme na nějaký odkaz. Pravděpodobnost $1 - \alpha$ je pak pravděpodobnost, že skočíme náhodně na jakoukoliv jinou stránku webu. Dostaneme tak další matici

$$G = \alpha S + \frac{1}{n}(1-\alpha)\mathbf{e}\mathbf{e}^T \ .$$

Tato **Google matice** je matice, kterou zakladatelé firmy Google Larry Page a Sergey Brin uvedli ve svém prvním článku o jejich algoritmu PageRank na porovnávání důležitosti webových stránek. Všimněme si, že všechny prvky matice G jsou kladné a součet prvků v každém sloupci zůstává rovný 1.

Náš malý příklad vede při volbě $\alpha = 0,9$ na matici

$$G = 0,9 \cdot S + 0,1 \cdot \frac{1}{6}\mathbf{e}\mathbf{e}^T = \begin{pmatrix} 1/60 & 1/6 & 19/60 & 1/60 & 1/60 & 1/60 \\ 7/15 & 1/6 & 19/60 & 1/60 & 1/60 & 1/60 \\ 7/15 & 1/6 & 1/60 & 1/60 & 1/60 & 1/60 \\ 1/60 & 1/6 & 1/60 & 1/60 & 7/15 & 11/12 \\ 1/60 & 1/6 & 19/60 & 7/15 & 1/60 & 1/60 \\ 1/60 & 1/6 & 1/60 & 7/15 & 7/15 & 1/60 \end{pmatrix} .$$

Diferenční rovnice $\mathbf{r}_k = G\mathbf{r}_{k-1}$ s počátečním vektorem $\mathbf{r}_0$ má pak řešení $\mathbf{r}_k = G^k\mathbf{r}_0$, které konverguje k jednoznačně určenému vektoru

$$\mathbf{r} = \begin{pmatrix} 0,03721 \\ 0,05396 \\ 0,04151 \\ 0,3751 \\ 0,206 \\ 0,2862 \end{pmatrix} .$$

Tento limitní vektor interpretujeme tak, že náhodný brouzdal po webu řídící se našimi pravidly stráví v průměru 3,721% času na stránce 1, 5,396% času na stránce 2, 37,51% času na stránce 4, atd.

Vlastnosti vlastních čísel matice G plynou z **Perronovy věty**, kterou dokázal již v roce 1907 německý matematik Oskar Perron. Uvedeme si bez důkazu její důsledky pro Google matici G.

Věta 9.112. *Pro Google matici G platí*

(1) *Číslo* 1 *je vlastním číslem matice G,*
(2) *geometrická i algebraická násobnost vlastního čísla* 1 *se rovná jedné,*
(3) *existuje vlastní vektor* $\mathbf{r}$ *příslušný vlastnímu číslu* 1*, který má všechny složky kladné,*
(4) *pro jakékoliv jiné vlastní číslo* λ *matice G platí* $|\lambda| < 1$.

Pokud kladný vlastní vektor $\mathbf{r}$ splňuje navíc podmínku $\|\mathbf{r}\| = 1$, nazývá se *Perronův vektor* matice G. První vlastnost jsme si už ukázali dříve, protože matice G je markovovská (tj. nezáporná a součet každého sloupce se rovná 1) a 1 je proto vlastní číslo G. Můžeme si také ověřit, že z dalších uvedených vlastností matice G plyne konvergence vektorů $\mathbf{r}_k = G^k\mathbf{r}_0$. Pokud si matici G převedeme do Jordanova kanonického tvaru $J = R^{-1}GR$ pomocí nějaké regulární matice R, můžeme předpokládat, že první Jordanova buňka $J_1 = J_{1,1}$ odpovídá vlastnímu číslu 1 a Perronův vektor $\mathbf{r}$ je prvním sloupcem matice R, jejíž sloupce tvoří bázi $B = (\mathbf{r} = \mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n)$ aritmetického prostoru $\mathbb{R}^n$ složenou ze Jordanových řetízků. Potom pro matici $J = \operatorname{diag}(J_1, J_2, \dots, J_s)$ platí

$$\mathbf{r}_k = RJ^kR^{-1}\mathbf{r}_0 = R\operatorname{diag}(J_1^k, J_2^k, \dots, J_s^k)R^{-1}\mathbf{r}_0 \ .$$

Protože $|\lambda| < 1$ pro jakékoliv vlastní číslo matice G různé od 1, platí $J_i^k \to O$ pro jakoukoliv Jordanovu buňku různou od J_1. Matice J^k tak konverguje k matici, která má na místě $(1,1)$ prvek 1 a všechny ostatní prvky nulové. Odtud plyne, že posloupnost vektorů

$$\mathbf{r}_k = RJ^kR^{-1}\mathbf{r}_0 = R\operatorname{diag}(J_1^k, J_2^k, \dots, J_s^k)R^{-1}\mathbf{r}_0$$

konverguje k nějakému skalárnímu násobku vektoru $\mathbf{r}$. Protože začínáme s vektorem $\mathbf{r}_0$, který má součet složek rovný 1, a násobíme jej markovovskou maticí, každý vektor $\mathbf{r}_k$ má součet složek rovný 1 a tedy jej má rovný 1 i limita posloupnosti vektorů $\mathbf{r}_k$. Posloupnost vektorů $\mathbf{r}_k$ tak komverguje k nějakému kladnému násobku Perronova vektoru $\mathbf{r}$, který má všechny složky kladné.

Tento výpočet ukazuje, že vhodný násobek Perronova vektoru odpovídá na všechny otázky spojené s řešením diferenční rovnice $\mathbf{r}_k = G^k\mathbf{r}_{k-1}$ s výjimkou rychlosti konvergence. Rychlost konvergence posloupnosti $\mathbf{r}_k$ závisí na tom, jak rychle konvergují k O mocniny Jordanovy buňky příslušné vlastním číslům $\lambda \neq 1$. Nejpomaleji z nich konvergují buňky odpovídající vlastnímu číslu $\lambda \neq 1$, který má co největší absolutní hodnotu $|\lambda|$. Rychlost konvergence tak závisí nejvíce na $|\lambda_2|$, kde λ_2 je druhé největší (pokud jde o absolutní hodnotu) vlastní číslo matice G.

Pokud jde o volbu parametru α, autoři algoritmu uvádějí $\alpha = 0,85$. Na volbě α závisí rychlost konvergence a numerická stabilita výpočtů. Z odhadů absolutní hodnoty druhého největšího vlastního čísla matice G vyplývá, že při této volbě α stačí k přesnosti na tři desetinná místa zhruba 50 iterací, tj. stačí spočítat vektor $\mathbf{r}_{50}$. Rychlost konvergence výpočtu také závisí na volbě počátečního vektoru $\mathbf{r}_0$. Otázka volby $\mathbf{r}_0$ je teoreticky podrobně zkoumána, žádné definitivní výsledky zatím nejsou. Firma Google uvádí, že každý výpočet začíná vždy od stejného počátečního vektoru $\mathbf{r}_0 = (1/n)\mathbf{e}$. Zatím se nepodařilo najít způsob, jak využít předchozích masivních výpočtů při výpočtu nové aktualizace vektoru důležitosti stránek.

Uvedené použití Jordanova kanonického tvaru pro důkaz konvergence posloupnosti vektorů $\mathbf{r}_k$ dobře ilustruje význam teoretických výsledků. Při vlastním výpočtu iterací $\mathbf{r}_k = G\mathbf{r}_{k-1}$ jej nepotřebujeme, součin počítáme přímo. Jordanův kanonický tvar nám umožňuje dokázat, že uvedený numerický postup vede k očekávanému výsledku.

Poslední poznámka se týká rychlosti násobení matice s vektorem. Matice G už není řídká, všechny její prvky jsou nenulové. Její tvar je

$$G = \alpha S + (1-\alpha)\frac{1}{n}\mathbf{e}\mathbf{e}^T = H + \alpha\mathbf{e}\mathbf{a}^T + (1-\alpha)\frac{1}{n}\mathbf{e}\mathbf{e}^T \ .$$

Matice H je řídká, s naprostou většinou prvků rovných 0. Matice G se od ní liší přičtením dvou matic s hodností rovnou 1. Násobíme-li maticí G libovolný vektor $\mathbf{x}$, počítáme

$$G\mathbf{x} = (\alpha S + (1-\alpha)\frac{1}{n}\mathbf{e}\mathbf{e}^T)\mathbf{x} = H\mathbf{x} + \alpha\mathbf{e}\mathbf{a}^T\mathbf{x} + (1-\alpha)\frac{1}{n}\mathbf{e}\mathbf{e}^T\mathbf{x} \ .$$

Člen $\alpha\mathbf{e}\mathbf{a}^T\mathbf{x}$ vyžaduje pouze výpočet standardního skalárního součinu $\mathbf{a}^T\mathbf{x}$, což je n násobení, doplněného o jedno další násobení $\alpha(\mathbf{a}^T\mathbf{x})$. Stejný počet násobení vyžaduje výpočet třetího členu. Celá složitost výpočtu $G\mathbf{x}$ tak závisí na složitosti výpočtu součinu velmi řídké matice H s vektorem $\mathbf{x}$.

Tento tvar matice G tak stále umožňuje řadu optimalizací výpočtů vytvořených pro počítání s řídkými maticemi.

Označíme-li $E = \frac{1}{n}\mathbf{e}\mathbf{e}^T$ matici, jejíž všechny prvky jsou rovné $1/n$, můžeme rovnici definující vektor $\mathbf{r}$ napsat ve tvaru

$$(\alpha S + (1-\alpha)E)\mathbf{r} = \mathbf{r} \ .$$

Její jednoduchost a elegance vede některé autory k názoru, že by měla být zařazena do příštího vydání knihy *It Must Be Beautiful: Great Equations of Modern Science*, jejíž první vydání vyšlo v roce 2002.

Cvičení

1. Dokažte, že relace podobnosti matic je ekvivalence na množině všech čtvercových matic téhož řádu n nad tělesem $\mathbf{T}$.

2. Nechť $\mathbf{V}$ je vektorový prostor dimenze n nad tělesem T, $B = (\mathbf{u}, \mathbf{u}_2, \dots, \mathbf{u}_n)$ je báze $\mathbf{V}$, $f : \mathbf{V} \to \mathbf{V}$ lineární operátor na $\mathbf{V}$ a $R = (r_{ij})$ regulární matice řádu n nad $\mathbf{T}$. Najděte bázi $C = (\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n)$ ve $\mathbf{V}$, pro kterou platí

$$[f]_C^C = R^{-1}[f]_B^B R.$$

3. Dokažte, že číslo 0 je vlastní číslo lineárního operátoru $f : \mathbf{V} \to \mathbf{V}$ právě když $\mathrm{Ker}\,(f) \neq \{\mathbf{o}\}$.

4. Dokažte, že jediné vlastní číslo jednotkové matice I_n je 1.

5. Dokaže, že je-li λ vlastní číslo matice A, je λ^2 vlastní číslo matice A^2.

6. Dokažte, že je-li A regulární matice a λ vlastní číslo A, pak λ^{-1} je vlastní číslo inverzní matice A^{-1}.

7. Derivace komplexni funkce realne promenne a reseni rovnice $f' = \lambda f$.

8. Nechť $\mathbf{V}$ je konečně generovaný prostor nad $\mathbf{T}$, $C = (\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n)$ je báze $\mathbf{V}$ a matice $[f]_C^C$ je podobná matici A. Dokažte, že existuje báze D ve $\mathbf{V}$, pro kterou platí $A = [f]_D^D$.

9. Spočítejte vlastní čísla a vlastní vektory matice ortogonální projekce na přímku určenou nenulovým vektorem $(a, b)^T \in \mathbb{R}^2$.

10. Najděte matici osové souměrnosti určené přímkou – lineárním obalem nenulového vektoru $(a, b)^T \in \mathbb{R}^2$ vzhledem ke kanonické bázi $\mathbb{R}^2$ a spočítejte její vlastní čísla a vlastní vektory.

11. Dokažte, že rotace kolem počátku souřadnic o úhel φ má reálná vlastní čísla právě když φ je násobkem π.

12. Spočítejte vlastní čísla a vlastní vektory projekce na obecnou přímku v $\mathbb{R}^3$. Napřed odhadněte výsledek z geometrického významu vlastních čísel a vektorů.

13. Jaké kořeny má v tělese $\mathbb{Z}_2$ polynom $x^2 + 1$? Jaké kořeny má v tělesech $\mathbb{Z}_3$ a $\mathbb{Z}_5$?

Shrnutí deváté kapitoly

(1) Je-li $\mathbf{V}$ lineární prostor nad tělesem $\mathbf{T}$, $f : \mathbf{V} \to \mathbf{V}$ lineární zobrazení, a $\mathbf{x}_0 \in \mathbf{V}$, pak *diskrétní lineární dynamický systém* je definovaný rovností $\mathbf{x}_k = f(\mathbf{x}_{k-1})$ pro každé $k \in \mathbb{N}$ a *počátečním stavem* $\mathbf{x}_0$. Prvek $\mathbf{x}_k \in \mathbf{V}$ nazýváme *stav systému* v čase k. Má-li prostor $\mathbf{V}$ konečnou dimenzi n, říkáme také, že dynamický systém má dimenzi n.

(2) Vývoj diskrétního lineárního dynamického systému je popsán vztahem $\mathbf{x}_k = f^k(\mathbf{x}_0)$ pro každé $k \geq 0$. Speciálně, je-li $\dim \mathbf{V} = 1$ a $f(x) = ax$ pro každé $x \in \mathbf{T}$, pak stav diskrétního lineárního dynamického systému $\mathbf{x}_k = f(\mathbf{x}_{k-1})$ v čase k se rovná $x_k = a^k\, x_0$ pro každé $k \geq 0$.

(3) Je-li $\mathbf{V} = \mathbb{R}^n$ (nebo $\mathbf{V} = \mathbb{C}^n$) a $\mathbf{x}(t) = (x_1(t), x_2(t), \ldots, x_n(t))^T \in \mathbf{V}$ pro každé $t \in \mathbb{R}$, pak definujeme derivaci $\mathbf{x}'(t)$ stavového vektoru $\mathbf{x}(t)$ jako vektor $\mathbf{x}'(t) = (x_1'(t), x_2'(t), \ldots, x_n'(t))^T$.

(4) Spojitý lineární dynamický systém je definovaný rovností $\mathbf{x}'(t) = f(\mathbf{x}(\mathbf{t}))$ pro každé $t \in \mathbb{R}$ a počátečním stavem $\mathbf{x}(0) \in \mathbb{R}^n$ (nebo $\mathbf{x}(0) \in \mathbb{C}^n$).

(5) Je-li $n = 1$, pak v čase t se stav spojitého lineárního dynamického systému $x'(t) = ax(t)$ s počátečním stavem $x(0) \in \mathbb{R}$ (nebo $x(0) \in \mathbb{C}^n$) rovná $x(t) = x(0)\, e^{at}$.

(6) Příklady diskrétních lineárních dynamických systémů - úročení, Fibonacciho posloupnost. Příklady spojitých lineárních dynamických systémů - rozpad radioaktivních jader, vlastní kmity pružiny, přechod substance přes buněčnou bránu.

(7) Je-li $f : \mathbf{V} \to \mathbf{V}$ lineární operátor na lineárním prostoru $\mathbf{V}$ nad tělesem $\mathbf{T}$, pak skalár $\lambda \in T$ nazýváme *vlastní číslo* operátoru f, pokud existuje nenulový vektor $\mathbf{x} \in V$, pro který platí
$$f(\mathbf{x}) = \lambda \mathbf{x} \ .$$
Je-li λ vlastní číslo operátoru f, pak libovolný prvek $\mathbf{x} \in V$, pro který platí $f(\mathbf{x}) = \lambda\mathbf{x}$, nazýváme *vlastní vektor* operátoru f *příslušný vlastnímu číslu* λ.

(8) Je-li A čtvercová matice řádu n nad tělesem $\mathbf{T}$, pak skalár $\lambda \in \mathbf{T}$ nazýváme *vlastní číslo* matice A, pokud existuje nenulový vektor $\mathbf{x} \in T^n$ takový, že
$$A\mathbf{x} = \lambda \mathbf{x} \ .$$
Je-li λ vlastní číslo matice A, pak libovolný vektor $\mathbf{x} \in T^n$, pro který platí $A\mathbf{x} = \lambda\mathbf{x}$, nazýváme *vlastní vektor* matice A *příslušný vlastnímu číslu* λ.

(9) Operátor $f : \mathbf{V} \to \mathbf{V}$ má vlastní číslo 0 právě tehdy, když f není prostý. Čtvercová matice A má vlastní číslo 0 právě tehdy, když A je singulární.

(10) Geometrický význam vlastních čísel a vektorů v případě jednoducgých geometrických zobrazení v $\mathbb{R}^2$.

(11) Nechť f je lineární operátor na prostoru $\mathbf{V}$ nad tělesem $\mathbf{T}$. Pak $\lambda \in T$ je vlastním číslem operátoru f právě tehdy, když operátor $(f - \lambda\, \mathrm{id}_V)$ není prostý.

Je-li λ vlastním číslem operátoru f, pak množina M_λ všech vlastních vektorů operátoru f příslušných vlastnímu číslu λ je podprostorem $\mathbf{V}$ a platí
$$M_\lambda = \mathrm{Ker}\,(f - \lambda\, \mathrm{id}_V) \ .$$

(12) Je-li A čtvercová matice řádu n nad tělesem $\mathbf{T}$, pak $\lambda \in T$ je vlastním číslem matice A právě tehdy, když je matice $A - \lambda I_n$ singulární.

Je-li λ vlastním číslem matice f, pak množina M_λ všech vlastních vektorů matice A příslušných vlastnímu číslu λ je podprostorem $\mathbf{T}^n$ a platí
$$M_\lambda = \mathrm{Ker}\,(A - \lambda I_n) \ .$$

(13) Je-li A čtvercová matice řádu n nad tělesem $\mathbf{T}$, pak $\lambda \in T$ je vlastním číslem matice A právě tehdy, když $\det\,(A - \lambda I_n) = 0$.

Je-li f lineární operátor na konečně generovaném prostoru $\mathbf{V}$ dimenze n nad tělesem $\mathbf{T}$ a B je báze $\mathbf{V}$, pak $\lambda \in T$ je vlastním číslem operátoru f právě když je λ vlastní číslo matice $[f]_B^B$ vzhledem k bázím B a B, což nastává právě když $\det\left([f]_B^B - \lambda I_n\right) = 0$.

(14) Pro každou matici $A = (a_{ij})$ řádu n nad libovolným tělesem $\mathbf{T}$ platí
 (a) $\det(A - \lambda I_n)$ je polynom stupně n s koeficienty v $\mathbf{T}$,
 (b) koeficient u λ^n se rovná $(-1)^n$,
 (c) koeficient u λ^{n-1} se rovná $(-1)^{n-1}(a_{11} + a_{22} + \cdots + a_{nn})$,
 (d) absolutní člen se rovná $\det A$.

(15) Je-li A čtvercová matice řádu n nad tělesem $\mathbf{T}$, pak *charakteristický polynom matice* A je polynom
$$p_A(\lambda) = \det\left(A - \lambda I_n\right) \quad .$$

(16) Dvě čtvercové matice X, Y téhož řádu nad tělesem $\mathbf{T}$ se nazývají *podobné*, pokud existuje regulární matice R taková, že $Y = R^{-1}XR$.

(17) Podobné matice mají stejný charakteristický polynom.

(18) Je-li $f : \mathbf{V} \to \mathbf{V}$ lineární operátor na konečně generovaném prostoru $\mathbf{V}$ dimenze n, pak *charakteristický polynom operátoru* f je polynom
$$p_f(\lambda) = \det\left([f]_B^B - \lambda I_n\right) \quad ,$$
kde B je libovolná báze prostoru $\mathbf{V}$.

(19) Nechť $p(x)$ je polynom nad $\mathbf{T}$. Prvek $t \in T$ je kořenem polynomu $p(x)$ právě tehdy, když polynom $x - t$ dělí polynom $p(x)$.

(20) Nechť $p(x)$ je polynom nad $\mathbf{T}$ a $t \in T$ je jeho kořen. *Násobnost kořene* t polynomu $p(x)$ definujeme jako největší přirozené číslo l takové, že polynom $(x - t)^l$ dělí polynom $p(x)$.

(21) Nechť $p(x)$ je polynom nad $\mathbf{T}$, $t_1, \ldots, t_k \in T$ po dvou různé a $l_1, \ldots, l_k \in \mathbb{N}$. Pak následující tvrzení jsou ekvivalentní.
 (1) Pro každé $i \in \{1, \ldots, k\}$ je t_i kořen násobnosti l_i.
 (2) $p(x) = (x - t_1)^{l_1} \ldots (x - t_k)^{l_k} q(x)$ pro nějaký polynom $q(x)$ takový, že ani jeden z prvků $t_1, \ldots, t_k$ není kořen.

(22) Polymom stupně n nad libovolným tělesem má nejvýše n kořenů včetně násobností.

(23) Každý polynom stupně $n \geq 1$ nad tělesem $\mathbb{C}$ lze napsat jako součin lineárních polynomů (tj. polynomů stupně 1).

Speciálně, každý polynom stupně $n \geq 0$ nad tělesem $\mathbb{C}$ má právě n kořenů včetně násobností.

Polynom lichého stupně nad tělesem $\mathbb{R}$ má alespoň jeden kořen.

(24) Nechť f je lineární operátor na konečně generovaném prostoru a λ je jeho vlastní číslo. *Algebraickou násobností* vlastního čísla λ rozumíme jeho násobnost jako kořene charakteristického polynomu operátoru f. Nechť A je čtvercová matice a λ je její vlastní číslo.

Algebraickou násobností vlastního čísla λ rozumíme jeho násobnost jako kořene charakteristického polynomu matice A.

(25)
- Každý lineární operátor $f : \mathbf{V} \to \mathbf{V}$ na konečně generovaném prostoru dimenze n nad tělesem $\mathbf{T}$ má nejvýše n vlastních čísel včetně násobností.
- Lineární operátor $f : \mathbf{V} \to \mathbf{V}$ má právě n vlastních čísel včetně násobností právě tehdy, když je jeho charakteristický polynom součinem lineárních polynomů.
- Každý lineární operátor $f : \mathbf{V} \to \mathbf{V}$ na konečně generovaném prostoru dimenze n nad tělesem $\mathbb{C}$ má právě n vlastních čísel včetně násobností.
- Každý lineární operátor $f : \mathbf{V} \to \mathbf{V}$ na konečně generovaném vektorovém prostoru liché dimenze nad $\mathbb{R}$ má aspoň jedno (reálné) vlastní číslo.

(26)
- Každá čtvercová matice řádu n nad tělesem $\mathbf{T}$ má nejvýše n vlastních čísel včetně algebraických násobností.

- Čtvecová matice řádu n nad tělesem $\mathbf{T}$ má právě n vlastních čísel včetně násobností právě tehdy, když je její charakteristický polynom součinem lineárních polynomů.
- Každá čtvercová matice řádu n nad tělesem $\mathbb{C}$ má právě n vlastních čísel včetně algebraických násobností.
- Každá čtvercová matice lichého řádu nad tělesem $\mathbb{R}$ má alespoň jedno reálné vlastní číslo.

(27) Lineární operátor $f : \mathbf{V} \to \mathbf{V}$ na konečně generovaném prostoru $\mathbf{V}$ nazýváme *diagonalizovatelný*, pokud má vzhledem k nějaké bázi diagonální matici.

(28) Je-li $f : \mathbf{V} \to \mathbf{V}$ lineární operátor na konečně generovaném prostoru $\mathbf{V}$ a je-li $B = (\mathbf{v}_1, \dots, \mathbf{v}_n)$ báze prostoru $\mathbf{V}$, pak $[f]_B^B = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$ platí právě tehdy, když pro každé $i \in \{1, 2, \dots, n\}$ je $\mathbf{v}_i$ vlastní vektor příslušný vlastnímu číslu λ_i.

(29) Lineární operátor $f : \mathbf{V} \to \mathbf{V}$ na konečně generovaném prostoru $\mathbf{V}$ je diagonalizovatelný právě tehdy, když existuje báze prostoru $\mathbf{V}$ tvořená vlastními vektory operátoru f.

(30) Je-li $f : \mathbf{V} \to \mathbf{V}$ lineární operátor na konečně generovaném prostoru $\mathbf{V}$ dimenze n nad tělesem $\mathbf{T}$ a C báze prostoru $\mathbf{V}$, pak operátor f je diagonalizovatelný právě tehdy, když je matice $[f]_C^C$ podobná diagonální matici.

(31) Čtvercová matice A řádu n nad tělesem $\mathbf{T}$ se nazývá *diagonalizovatelná*, pokud je operátor $f_A : \mathbf{T}^n \to \mathbf{T}^n$ diagonalizovatelný.

(32) Je-li A čtvercová matice řádu n nad tělesem $\mathbf{T}$, $B = (\mathbf{v}_1, \dots, \mathbf{v}_n)$ báze prostoru $\mathbf{T}^n$ a $R = [\mathrm{id}]_K^B = (\mathbf{v}_1 | \dots | \mathbf{v}_n)$, pak matice $[f_A]_B^B = R^{-1}AR$ se rovná diagonální matici $\mathrm{diag}(\lambda_1, \dots, \lambda_n)$ právě tehdy, když pro každé $i \in \{1, 2, \dots, n\}$ je $\mathbf{v}_i$ vlastní vektor příslušný vlastnímu číslu λ_i.

(33) Čtvercová matice A řádu n nad tělesem $\mathbf{T}$ je diagonalizovatelná právě tehdy, když existuje báze prostoru $\mathbf{T}^n$ tvořená vlastními vektory matice A.

(34) Čtvercová matice A řádu n nad tělesem $\mathbf{T}$ je diagonalizovatelná právě tehdy, když je podobná diagonální matici.

(35) Je-li $f : \mathbf{V} \to \mathbf{V}$ lineární operátor a $(\mathbf{v_1}, \mathbf{v}_2, \dots, \mathbf{v}_k)$ posloupnost nenulových vlastních vektorů operátoru f příslušných navzájem různým vlastním číslům $\lambda_1, \dots, \lambda_k$, pak je posloupnost $(\mathbf{v_1}, \mathbf{v}_2, \dots, \mathbf{v}_k)$ lineárně nezávislá.

(36) Má-li lineární operátor $f : \mathbf{V} \to \mathbf{V}$ na vektorovém prostoru $\mathbf{V}$ dimenze n nad tělesem $\mathbf{T}$ celkem n navzájem různých vlastních čísel, pak je diagonalizovatelný.

(37) Má-li matice A řádu n nad tělesem $\mathbf{T}$ celkem n navzájem různých vlastních čísel, pak je diagonalizovatelná.

(38) Celé odvození vztahu pro k-tý člen Fibonacciho posloupnosti.

(39) *Geometrickou násobností* vlastního čísla λ operátoru f na konečně generovaném prostoru (nebo čtvercové matice A) rozumíme dimenzi podprostoru $\mathbf{M}_\lambda$ vlastních vektorů operátoru f (nebo matice A) příslušných vlastnímu číslu λ.

(40) Pro čtvercovou blokovou matici

$$A = \begin{pmatrix} B & C \\ 0 & D \end{pmatrix}$$

se čtvercovými diagonálními bloky B, D platí

$$\det A = (\det B)\,(\det D)\ .$$

(41) Pro každé vlastní číslo μ lineárního operátoru $f : \mathbf{V} \to \mathbf{V}$ na konečně generovaném prostoru $\mathbf{V}$ (čtvercové matice A) nad tělesem $\mathbf{T}$ platí, že geometrická násobnost μ je menší nebo rovná algebraické násobnosti λ.

(42) Buď $f : \mathbf{V} \to \mathbf{V}$ lineární operátor na konečně generovaném vektorovém prostoru $\mathbf{V}$ dimenze n (resp. buď A je čtvercová matice řádu n) nad tělesem $\mathbf{T}$. Pak jsou následující tvrzení jsou ekvivalentní.

(a) Operátor f je diagonalizovatelný (resp. matice A je diagonalizovatelná).
(b) Operátor f (resp. matice A) má
- n vlastních čísel včetně algebraických násobností a
- geometrická násobnost každého vlastního čísla operátoru f (resp. matice A) je rovná jeho algebraické násobnosti.

(43) Je-li $\mathbf{x} \in \mathbb{C}^n$ vlastní vektor reálné matice $A = (a_{ij})$ příslušný vlastnímu číslu λ, pak $\overline{\mathbf{x}}$ je vlastní vektor matice A příslušný vlastnímu číslu $\overline{\lambda}$.

(44) Je-li A reálná matice řádu 2, která nemá reálná vlastní čísla, a $\lambda = r(\cos\varphi + i \sin\varphi)$ je komplexní vlastní číslo s nenulovým vlastním vektorem $\mathbf{v}$, pak platí
(a) vektory $\mathbf{w}_1 = \mathbf{v} + \overline{\mathbf{v}}$ a $\mathbf{w}_2 = i(\mathbf{v} - \overline{\mathbf{v}})$ tvoří bázi $B = (\mathbf{w}_1, \mathbf{w}_2)$ prostoru $\mathbb{R}^2$,
(b) lineární zobrazení $f_A : \mathbb{R}^2 \to \mathbb{R}^2$ určené maticí A má vzhledem k bázi B matici

$$[f_A]_B^B = r \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}$$

a je tedy složením rotace o úhel φ se stejnolehlostí s koeficientem $r > 0$

(45) Vývoj spojitého lineárního dynamického systému s diagonalizovatelnou maticí.

(46) *Jordanova buňka* řádu $k \geq 1$ nad tělesem $\mathbf{T}$ příslušná prvku $\lambda \in T$ je čtvercová matice

$$J_{\lambda,k} = \begin{pmatrix} \lambda & 1 & 0 & & 0 & 0 \\ 0 & \lambda & 1 & \dots & 0 & 0 \\ 0 & 0 & \lambda & & 0 & 0 \\ & \vdots & & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \lambda & 1 \\ 0 & 0 & 0 & \dots & 0 & \lambda \end{pmatrix}.$$

(47) Matice J nad tělesem $\mathbf{T}$ je v *Jordanově kanonickém tvaru* (nebo stručněji v *Jordanově tvaru*), pokud J je blokově diagonální matice, jejíž každý diagonální blok je Jordanova buňka (nějakého řádu příslušná nějakému číslu), tj.

$$J = \operatorname{diag}(J_{\lambda_1,k_1}, \dots, J_{\lambda_s,k_s}) = \begin{pmatrix} J_{\lambda_1,k_1} & 0 & \dots & 0 \\ 0 & J_{\lambda_2,k_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & J_{\lambda_s,k_s} \end{pmatrix},$$

kde $\lambda_1, \dots, \lambda_s \in T$ a $k_1, \dots, k_s$ jsou kladná celá čísla.

(48) Pro libovolná přirozená čísla $m < k$ platí

$$J_{0,k}^m = (\underbrace{\mathbf{o}|\dots|\mathbf{o}}_{m\times}|\mathbf{e}_1|\mathbf{e}_2|\mathbf{e}_{k-m})$$

Pro $m \geq k$ je $J_{0,k}^m = 0$.

(49) Je-li $J = J_{\lambda,k}$ Jordanova buňka, pak pro každé kladné m platí

$$J_{\lambda,k}^m = \begin{pmatrix} \lambda^m & \binom{m}{1}\lambda^{m-1} & \binom{m}{2}\lambda^{m-2} & \dots & \binom{m}{k-1}\lambda^{m-k+1} \\ 0 & \lambda^m & \binom{m}{1}\lambda^{m-1} & \dots & \binom{m}{k-2}\lambda^{m-k+2} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda^m & \binom{m}{1}\lambda^{m-1} \\ 0 & 0 & \dots & 0 & \lambda^m \end{pmatrix}.$$

(50) Říkáme, že pro lineární operátor $f : \mathbf{V} \to \mathbf{V}$ na konečně generovaném prostoru $\mathbf{V}$ *existuje Jordanův kanonický tvar*, pokud má vzhledem k nějaké bázi matici v Jordanově kanonickém tvaru.

(51) Je-li f lineární operátor na vektorovém prostoru $\mathbf{V}$ nad tělesem $\mathbf{T}$ a λ vlastní číslo operátoru f, pak posloupnost $(\mathbf{v}_1, \dots, \mathbf{v}_k)$ vektorů z $\mathbf{V}$ nazýváme *Jordanův řetízek operátoru f délky k příslušný vlastnímu číslu λ s počátkem* $\mathbf{v}_1$, pokud platí

$$(f - \lambda\,\mathrm{id}_V)(\mathbf{v}_1) = \mathbf{o},\ (f - \lambda\,\mathrm{id}_V)(\mathbf{v}_2) = \mathbf{v}_1,\ (f - \lambda\,\mathrm{id}_V)(\mathbf{v}_3) = \mathbf{v}_2,\ \dots,$$
$$(f - \lambda\,\mathrm{id}_V)(\mathbf{v}_k) = \mathbf{v}_{k-1}\ .$$

(52) Je-li $f : \mathbf{V} \to \mathbf{V}$ lineární operátor na konečně generovaném prostoru $\mathbf{V}$ a $B = (\mathbf{v}_1, \dots, \mathbf{v}_k)$ báze prostoru $\mathbf{V}$, pak $[f]^B_B = J_{\lambda,k}$ právě tehdy, když $(\mathbf{v}_1, \dots, \mathbf{v}_k)$ je Jordanův řetízek operátoru f příslušný vlastnímu číslu λ s počátkem $\mathbf{v}_1$.

(53) Je-li $f : \mathbf{V} \to \mathbf{V}$ lineární operátor na konečně generovaném prostoru $\mathbf{V}$ a B báze prostoru $\mathbf{V}$, pak $[f]^B_B = \mathrm{diag}(J_{\lambda_1,k_1}, \dots, J_{\lambda_s,k_s})$ platí právě tehdy, když B je spojením posloupností $B_1, \dots, B_s$, kde pro každé $i \in \{1, \dots, s\}$ je B_i Jordanův řetízek operátoru f délky k_i příslušný vlastnímu číslu λ_i s počátkem $\mathbf{v}^i_1$.

(54) Pro lineární operátor $f : \mathbf{V} \to \mathbf{V}$ na konečně generovaném prostoru $\mathbf{V}$ existuje Jordanův tvar právě tehdy, když existuje báze prostoru $\mathbf{V}$ vzniklá spojením Jordanových řetízků operátoru f.

(55) Nechť $f : \mathbf{V} \to \mathbf{V}$ je lineární operátor na konečně generovaném prostoru $\mathbf{V}$ a C je báze prostoru $\mathbf{V}$. Pak pro operátor f existuje Jordanův tvar právě tehdy, když je matice $[f]^C_C$ podobná matici v Jordanově tvaru.

(56) Předpokládáme, že $f : \mathbf{V} \to \mathbf{V}$ lineární operátor a $B_1, \dots, B_s$ jsou Jordanovy řetízky operátoru f příslušné vlastním číslům $\lambda_1, \dots, \lambda_s$. Předpokládejme dále, že pro každé $\lambda \in \{\lambda_1, \dots, \lambda_s\}$ je posloupnost počátečních vektorů těch řetízků z $B_1, \dots, B_s$, které přísluší vlastnímu číslu λ, lineárně nezávislá. Pak spojení $B = B_1, \dots, B_s$ je lineárně nezávislá posloupnost.

(57) Je-li $f : \mathbf{V} \to \mathbf{V}$ lineární operátor na prostoru $\mathbf{V}$ dimenze n a B báze vzniklá spojením Jordanových řetízků operátoru f, pak platí
   (a) operátor f má n vlastních čísel včetně násobností,
   (b) pro libovolné vlastní číslo λ operátoru f je jeho algebraická násobnost rovna součtu délek Jordanových řetízků v B příslušných vlastnímu číslu λ,
   (c) pro libovolné vlastní číslo λ operátoru f a libovolné $l \in \mathbb{N}$ je jádro operátoru $(f-\lambda\,\mathrm{id}_V)^l$ rovno lineárnímu obalu l počátečních vektorů z každého řetízku v B příslušného vlastnímu číslu λ (z řetízků délky menší než l bereme všechny vektory),
   (d) pro libovolné vlastní číslo λ operátoru f a libovolné $l \in \mathbb{N}$ je obraz operátoru $(f - \lambda\,\mathrm{id}_V)^l$, roven lineárnímu obalu všech vektorů v B kromě l koncových vektorů z řetízků příslušných vlastnímu číslu λ (z řetízků příslušných vlastnímu číslu λ délky menší než l nebereme žádný vektor).

   Speciálně pro libovolné vlastní číslo λ operátoru f platí
   (5) geometrická násobnost vlastního čísla λ se rovná počtu řetízků v B příslušných vlastnímu číslu λ a prostor $M_\lambda = \mathrm{Ker}\,(f - \lambda\,\mathrm{id}_V)$ je roven lineárnímu obalu počátečních vektorů těchto řetízků,
   (6) počet řetízků příslušných vlastnímu číslu λ délky alespoň l je roven

$$z_l = \dim \mathrm{Ker}\,(f - \lambda\,\mathrm{id}_V)^l - \dim \mathrm{Ker}\,(f - \lambda\,\mathrm{id}_V)^{l-1}\ ,$$

   (aby měl výraz smysl i pro $l = 1$ definujeme $(f - \lambda\,\mathrm{id}_V)^0 = \mathrm{id}_V$),
   (7) počet řetízků příslušných vlastnímu číslu λ délky právě l je $z_l - z_{l+1}$.

(58) Je-li $f : \mathbf{V} \to \mathbf{V}$ lineární operátor na konečně generovaném vektorovém prostoru $\mathbf{V}$ dimenze n, pak jsou následující tvrzení ekvivalentní.
   (a) Pro operátor f existuje Jordanův kanonický tvar.
   (b) Operátor f (resp. matice A) má n vlastních čísel včetně algebraických násobností.

(59) Pro každý operátor $f : \mathbf{V} \to \mathbf{V}$ na konečně dimenzionálním prostoru $\mathbf{V}$ nad tělesem komplexních čísel $\mathbb{C}$ existuje Jordanův kanonický tvar.

(60) Řešení spojitého lineárního dynamického systému s maticí rovnou Jordanově buňce řádu 2.

(61) Nechť $\mathbf{T}$ je těleso, $p(t) = a_0 + a_1 t + a_2 t^2 + \cdots + a_n t^n$ polynom s koeficienty $a_0, \ldots, a_n$ v $\mathbf{T}$, A čtvercová matice řádu k nad $\mathbf{T}$ a f lineární operátor na prostoru $\mathbf{V}$ nad tělesem $\mathbf{T}$. *Dosazením matice A do polynomu* $p(t)$ rozumíme matici

$$p(A) = a_0 I_k + a_1 A + a_2 A^2 + \cdots + a_n A^n \ .$$

Dosazením operátoru f do polynomu $p(t)$ rozumíme operátor

$$p(f) = a_0 \,\mathrm{id}_V + a_1 f + a_2 f^2 + \cdots + a_n f^n \ .$$

(62) Cayleyho-Hamiltonova věta. Je-li f lineární operátor na konečně generovaném prostoru $\mathbf{V}$ dimenze n nad tělesem $\mathbf{T}$ (resp. je-li A čtvercová matice řádu n nad $\mathbf{T}$), pak $p_f(f) = 0$ (resp. $p_A(A) = 0$).

Klíčové znalosti z deváté kapitoly nezbytné pro průběžné sledování přednášek s pochopením

(1) Diskrétní a spojité lineární dynamické systémy.
(2) Vlastní čísla a vlastní vektory lineárních operátorů a matic.
(3) Charakteristický polynom matice a operátoru na prostoru konečné dimenze.
(4) Vlastní čísla matice nebo operátoru na prostoru konečné dimenze jsou kořeny charakteristického polynomu.
(5) Algebraická násobnost vlastních čísel.
(6) Diagonalizovatelné operátory a matice.
(7) Lineární nezávislost posloupnosti vlastních vektorů příslušných různým vlastním číslům.
(8) Geometrická násobnost vlastních čísel.
(9) Charakterizace diagonalizovatelných matic a operátorů.
(10) Jordanovy buňky a Jordanův kanonický tvar.
(11) Věta o Jordanově kanonickém tvaru.
(12) Cayleyho-Hamiltonova věta.

10. Ortogonální a unitární diagonalizace

Cíl. *V této kapitole budeme zkoumat vlastní čísla a vlastní vektory reálných a komplexních matic. Bude nás zajímat, kdy existuje ortonormální báze složená z vlastních vektorů matice. Ukážeme si jak poznatky o maticích přenést na operátory na konečně generovaných prostorech se skalárním součinem. Nakonec si ukážeme singulární rozklad reálné nebo komplexní matice. Z hlediska numerických výpočtů je singulární rozklad vhodnější nástroj pro zkoumání matic než Jordanův kanonický tvar.*

10.1. **Unitární diagonalizovatelnost.** Má-li operátor f vzhledem k nějaké bázi diagonální matici, máme docela dobrou představu, co operátor "dělá". Víme-li například, že operátor f na prostoru $\mathbb{R}^2$ má vzhledem k bázi $B = (\mathbf{v}_1, \mathbf{v}_2)$ matici $D = \mathrm{diag}(1, 2)$, víme, že f zachovává vektor $\mathbf{v}_1$ a dvakrát prodlužuje vektor $\mathbf{v}_2$. Tím je díky linearitě operátor f zcela určen.

Informace, že matice f je vzhledem k bázi B prostoru $\mathbb{R}^2$ diagonální, ale není úplně uspokojivá, pokud vezmeme do úvahy standardní skalární součin na $\mathbb{R}^2$. Z obrázku lze sice odhadnout, že obraz jednotkové kružnice v $\mathbb{R}^2$ je nějaká elipsa, z vlastních čísel a obecných vlastních vektorů $\mathbf{v}_1, \mathbf{v}_2$ ale přímo nepoznáme, jaké jsou délky a směry jejích poloos. Z obrázku odhadujeme, že jde o elipsu, směr poloos a jejich velikosti ale nejsou v jednoduchém vztahu s bází B a maticí D.

OBRAZEK - obraz jednotkove kruznice operatorem

Pokud ale najdeme v $\mathbb{R}^2$ ortonormální bázi $B = (\mathbf{v}_1, \mathbf{v}_2)$ takovou, že $[f]_B^B = \mathrm{diag}(1, 2)$, pak hned vidíme, že obraz jednotkové kružnice je elipsa s osami $\langle \mathbf{v}_1 \rangle$, $\langle \mathbf{v}_2 \rangle$, délkou poloosy 1 ve směru $\langle \mathbf{v}_1 \rangle$ a délkou poloosy 2 ve směru $\langle \mathbf{v}_2 \rangle$. S úhly mezi vektory báze B souvisí problém numerické stability výpočtů s maticí přechodu od báze B ke kanonické bázi K. Báze, ve které jsou některé vektory jsou „téměř rovnoběžné“, vedou na numerickou nestabilitu výpočtů (jak jsme viděli u soustav lineárních rovnic), jsou totiž příliš blízko lineárně závislým množinám.

OBRAZEK - obraz jednotkove kruznice operatorem vzhledem k ortonormalni bazi

V této části se budeme zabývat otázkou, kdy pro operátor f na prostoru $\mathbf{V}$ se skalárním součinem existuje ortonormální báze B prostoru $\mathbf{V}$ taková, že $[f]_B^B$ je diagonální, tj. ortonormální báze složená z vlastních vektorů operátoru f. Takovým operátorům říkáme unitárně diagonalizovatelné. Protože na $\mathbf{V}$ je nutný skalární součin, abychom mohli mluvit o úhlech mezi prvky prostoru $\mathbf{V}$, bude $\mathbf{V}$ vždy reálný nebo komplexní vektorový prostor.

Definice 10.1. Je-li $\mathbf{V}$ konečně generovaný lineární prostor nad $\mathbb{C}$ (resp. $\mathbb{R}$) se skalárním součinem $\langle\, , \rangle$ a f lineární operátor na $\mathbf{V}$, pak říkáme, že f je *unitárně diagonalizovatelný* (resp. *ortogonálně diagonalizovatelný*), pokud existuje ortonormální báze B prostoru $\mathbf{V}$ taková, že $[f]_B^B$ je diagonální.

Následující tvrzení je obdobou věty 9.62 charakterizující unitárně diagonalizovatelné operátory. Formulaci uvedeme v operátorové verzi, maticovou verzi přenecháme čtenáři.

Věta 10.2. *Je-li $f : \mathbf{V} \to \mathbf{V}$ lineární operátor na konečně generovaném lineárním prostoru $\mathbf{V}$ dimenze n se skalárním součinem nad tělesem $\mathbb{C}$ (resp. $\mathbb{R}$), pak jsou následující tvrzení ekvivalentní.*

(1) *Operátor f je unitárně diagonalizovatelný (resp. ortogonálně diagonalizovatelný).*
(2) *Operátor f*
- *má n vlastních čísel včetně algebraických násobností,*

- *geometrická násobnost každého vlastního čísla operátoru f se rovná jeho algebraické násobnosti a*
- *pro libovolná dvě různá vlastní čísla λ_i, λ_j operátoru f platí $M_{\lambda_i} \perp M_{\lambda_j}$.*

Důkaz. Důkaz je podobný jako ve větě 9.62. Předpokládáme, že $\lambda_1, \ldots, \lambda_k$ jsou všechna navzájem různá vlastní čísla operátoru f.

Pro důkaz (2) $\Rightarrow$ (1) vybereme v každém z prostorů $\mathbf{M}_{\lambda_i}$ ortonormální bázi B_i a spojení B bází $B_1, B_2, \cdots, B_s$ bude báze ve $\mathbf{V}$. Stejně jako v důkazu věty 9.62 k tomu stačí první dva z předpokladů na operátor f v bodě (2). Báze B je navíc ortonormální. Všechny prvky báze B mají normu 1, protože B_i je ortonormální báze v $\mathbf{M}_{\lambda_i}$, a z téhož důvodu jsou libovolné dva různé prvky B_i kolmé. Je-li $i \neq j$, pak každý prvek $B_i \subseteq M_{\lambda_i}$ je kolmý na každý prvek $B_j \subseteq M_{\lambda_j}$ podle třetího z předpokladů.

Přepodkládejme naopak, že B je ortonormální báze prostoru $\mathbf{V}$ taková, že matice $[f]_B^B$ je diagonální. To znamená, že každý z prvků báze B je nenulový vlastní vektor operátoru f. Stejně jako v důkazu implikace (1) $\Rightarrow$ (2) ve větě 9.62 z toho vyplývá, že báze B je složená z bází prostorů $M_{\lambda_1}, \ldots, M_{\lambda_k}$. Protože všechny vektory v B jsou navzájem kolmé, jsou navzájem kolmé i podprostory $\mathbf{M}_{\lambda_i} = \langle B_i \rangle$ a $\mathbf{M}_{\lambda_j} = \langle B_j \rangle$ (viz pozorování 8.43 o kolmosti lineárního obalu). □

Jinými slovy poslední věta říká, že operátor je unitárně diagonalizovatelný právě tehdy, když je diagonalizovatelný a vlastní vektory příslušné různým vlastním číslům jsou na sebe kolmé.

Je-li f operátor na $\mathbf{V}$, $B = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$ ortonormální báze prostoru $\mathbf{V}$ a $[f]_B^B = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$, pak pro libovolný prvek $\mathbf{x} \in \mathbf{V}$ je podle tvrzení 8.38

$$[\mathbf{x}]_B = (\langle \mathbf{v}_1, \mathbf{x} \rangle, \ldots, \langle \mathbf{v}_n, \mathbf{x} \rangle)^T \ .$$

Souřadnice obrazu $f(\mathbf{x})$ prvku $\mathbf{x}$ vzhledem k bázi B jsou

$$[f(\mathbf{x})]_B = (\lambda_1 \langle \mathbf{v}_1, \mathbf{x} \rangle, \ldots, \lambda_n \langle \mathbf{v}_n, \mathbf{x} \rangle)^T \ ,$$

což znamená, že

$$f(\mathbf{x}) = \lambda_1 \langle \mathbf{v}_1, \mathbf{x} \rangle \mathbf{v}_1 + \cdots + \lambda_n \langle \mathbf{v}_n, \mathbf{x} \rangle \mathbf{v}_n \ .$$

Prvek $\langle \mathbf{v}_i, \mathbf{x} \rangle \mathbf{v}_i$ se rovná ortogonální projekci prvku $\mathbf{x}$ na přímku $\langle \mathbf{v}_i \rangle$, podle tvrzení 8.46. Označíme-li p_i ortogonální projekci na $\langle \mathbf{v}_i \rangle$ chápanou jako lineární zobrazení, tj. $p_i(\mathbf{x}) = \mathbf{x}_{\langle \mathbf{v}_i \rangle}$, můžeme psát

$$f = \lambda_1 p_1 + \ldots \lambda_n p_n \ .$$

Unitárně diagonalizovatelný operátor lze tedy vyjádřit jako lineární kombinaci projekcí na vzájemně kolmé jednodimenzionální podprostory.

Hlavní výsledky v této části jsou, že tzv. hermitovské operátory (resp. symetrické) a unitární operátory (resp. ortogonální) jsou unitárně diagonalizovatelné. Tyto výsledky vyplynou z charakterizace unitárně diagonalizovatelných operátorů jako tzv. operátorů normálních.

Pojmy hermitovský (symetrický), unitární (ortogonální) a normální používáme také pro čtvercové matice, kromě normálních matic jsme již dokonce všechny definovali. Předem poznamejme, že matice A je hermitovská, ... právě tehdy, když je operátor f_A na prostoru $\mathbb{C}^n$ (resp. $\mathbb{R}^n$) se standardním skalárním součinem hermitovský, ...

10.1.1. *Sdružené lineární zobrazení.* Pro komplexní matice (ne nutně čtvercové) jsme v kapitole o skalárním součinu definovali matici hermitovsky sdruženou jako matici komplexně sdruženou k transponované. Pro reálné matice tento pojem splývá s transponovanou maticí. Nyní obecněji definujeme pojem sdruženého lineárního zobrazení mezi prostory se skalárním součinem. Tím také ukážeme geometrický význam hermitovsky sdružené matice.

Reálná matice A typu $m \times n$ a příslušná transponovaná matice A^T splňují pro libovolné vektory $\mathbf{x} \in \mathbb{R}^m$, $\mathbf{y} \in \mathbb{R}^n$ vztah

$$A^T\mathbf{x} \cdot \mathbf{y} = \mathbf{x} \cdot A\mathbf{y} \ .$$

(Na levé straně značí $\cdot$ standardní skalární součin v $\mathbb{R}^n$, na pravé straně v $\mathbb{R}^m$.) Skutečně, $A^T\mathbf{x} \cdot \mathbf{y} = (A^T\mathbf{x})^T\mathbf{y} = \mathbf{x}^T A\mathbf{y} = \mathbf{x} \cdot A\mathbf{y}$. Tento vztah transponovanou matici k A charakterizuje – A^T je jediná taková matice B, pro kterou platí formulka $B\mathbf{x} \cdot \mathbf{y} = \mathbf{x} \cdot A\mathbf{y}$, jak se přesvědčíme dosazením všech dvojic vektorů kanonické báze.

Pro komplexní matici A je obdobně

$$A^*\mathbf{x} \cdot \mathbf{y} = \mathbf{x} \cdot A\mathbf{y} \ ,$$

protože $A^*\mathbf{x} \cdot \mathbf{y} = (A^*\mathbf{x})^*\mathbf{y} = \mathbf{x}^* A\mathbf{y} = \mathbf{x} \cdot A\mathbf{y}$.

Tento pohled na hermitovské sdružování využijeme k definici sdruženého operátoru.

Tvrzení 10.3. *Nechť* $\mathbf{V}$ *a* $\mathbf{W}$ *jsou konečně generované vektorové prostory nad* $\mathbb{C}$ *(nebo* $\mathbb{R}$*) se skalárními součiny (které jsou jako obvykle značeny* $\langle\, , \rangle$*) a* $f : \mathbf{V} \to \mathbf{W}$ *je lineární zobrazení. Pak existuje právě jedno lineární zobrazení* $g : \mathbf{W} \to \mathbf{V}$ *splňující pro každé* $\mathbf{x} \in W$*,* $\mathbf{y} \in V$ *rovnost* $\langle\, g(\mathbf{x}), \mathbf{y}\rangle = \langle\, \mathbf{x}, f(\mathbf{y})\rangle$.

Důkaz. Dokážeme nejprve existenci. Zvolíme libovolnou ortonormální bázi $B = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$ prostoru $\mathbf{V}$ a ortonormální bázi $C = (\mathbf{w}_1, \ldots, \mathbf{w}_m)$ prostoru $\mathbf{W}$. Každé lineární zobrazení z $\mathbf{W}$ do $\mathbf{V}$ můžeme zadat maticí vzhledem k bázím C a B. Definujeme operátor g tak, aby $[g]^C_B = ([f]^B_C)^*$. Ověříme, že g splňuje pro libovolné vektory $\mathbf{x} \in W$, $\mathbf{y} \in V$ vztah $\langle\, g(\mathbf{x}), \mathbf{y}\rangle = \langle\, \mathbf{x}, f(\mathbf{y})\rangle$. Skutečně, užitím tvrzení 8.40 o skalárním součinu vzhledem k ortonormální bázi dostáváme

$$\begin{aligned}\langle\, g(\mathbf{x}), \mathbf{y}\rangle &= [g(\mathbf{x})]^*_B[\mathbf{y}]_B = ([g]^C_B[\mathbf{x}]_C)^*[\mathbf{y}]_B = [\mathbf{x}]^*_C([g]^C_B)^*[\mathbf{y}]_B \\ &= [\mathbf{x}]^*_C[f]^B_C[\mathbf{y}]_B = [\mathbf{x}]^*_C[f(\mathbf{y})]_C = \langle\, \mathbf{x}, f(\mathbf{y})\rangle \ .\end{aligned}$$

Jednoznačnost ukážeme dosazením dvojic vektorů bází B, C do rovnosti $\langle\, g(\mathbf{x}), \mathbf{y}\rangle = \langle\, \mathbf{x}, f(\mathbf{y})\rangle$. Pro libovolné $i \in \{1, \ldots, m\}$, $j \in \{1, \ldots, n\}$ je $\langle\, g(\mathbf{w}_i), \mathbf{v}_j\rangle = \langle\, \mathbf{w}_i, f(\mathbf{v}_j)\rangle$. Úpravou obou stran dostaneme

$$\begin{aligned}[\mathbf{w}_i]^*_C([g]^C_B)^*[\mathbf{v}_j]_B &= [\mathbf{w}_i]^*_B[f]^B_C[\mathbf{v}_j]_B \\ \mathbf{e}_i([g]^C_B)^*\mathbf{e}_j &= \mathbf{e}_i[f]^B_C\mathbf{e}_j\end{aligned}$$

Na pravé straně je prvek na místě (i, j) v matici $[f]^B_C$, na levé straně je prvek na místě (i, j) v matici $([g]^C_B)^*$. Platí tedy $([g]^C_B)^* = [f]^B_C$, neboli $[g]^C_B = ([f]^B_C)^*$. Ukázali jsme, že pro g splňující rovnost $\langle\, g(\mathbf{x}), \mathbf{y}\rangle = \langle\, \mathbf{x}, f(\mathbf{y})\rangle$ musí nutně platit $[g]^C_B = ([f]^B_C)^*$. Protože operátor $\mathbf{W} \to \mathbf{V}$ je jednoznačně určen svou maticí vzhledem k C a B, důkaz je ukončen. □

Alternativně jde tvrzení dokázat použitím věty **??** o reprezentaci lineárních forem skalárním součinem, viz cvičení.

Definice 10.4. V situaci tvrzení 10.3 nazýváme g *sdružené lineární zobrazení k lineárnímu zobrazení* f a značíme $f^* = g$.

Definující vztah pro f^* je tedy

$$\langle\, f^*(\mathbf{x}), \mathbf{y}\rangle = \langle\, \mathbf{x}, f(\mathbf{y})\rangle$$

Příklad 10.5. Platí $\mathrm{id}^* = \mathrm{id}$, $0^* = 0$.

Protože $\langle\, \overline{a}\mathbf{x}, \mathbf{y}\rangle = a\,\langle\, \mathbf{x}, \mathbf{y}\rangle = \langle\, \mathbf{x}, a\mathbf{y}\rangle$, je sdružené zobrazení k $a\,\mathrm{id}$ rovno $(a\,\mathrm{id})^* = \overline{a}\,\mathrm{id}$.

Sdružené zobrazení k rotaci $f : \mathbb{R}^2 \to \mathbb{R}^2$ o úhel α je rotace o úhel $-\alpha$, jak je vidět na obrázku (v prostorech $\mathbb{R}^2$ bereme standardní skalární součin):

OBRAZEK

Příklad 10.6. Na prostoru, který není konečně generovaný, nemusí sdružené lineární zobrazení obecně existovat (lze ukázat, že pokud existuje, je určené jednoznačně). Ukážeme, že operátor derivování na vhodném prostoru sdružený operátor má.

Označme $\mathbf{V}$ vektorový prostor všech hladkých reálných funkcí f na nějakém intervalu, např. $[0,1]$, takových, že $f(0) = f(1) = 0$, se skalárním součinem

$$\langle f, g \rangle = \int_0^1 fg \ .$$

Uvažujme operátor D, který každé funkci f přiřazuje její derivaci $D(f) = f'$. Pomocí integrace per partes vypočítáme, že sdružený operátor k D existuje a je roven $-D$:

$$\langle -D(f), g \rangle = \langle -f', g \rangle = \int_0^1 -f'g = \int_0^1 fg' - [fg]_0^1 = \int_0^1 fg' = \langle f, D(g) \rangle$$

Z důkazu tvrzení 10.3 vyplývá, že matice lineárního zobrazení f^* vzhledem k bázím C a B je hermitovsky sdružená k matici lineárního zobrazení f vzhledem k B a C:

Důsledek 10.7. *Nechť f je lineární zobrazení $\mathbf{V} \to \mathbf{W}$, kde $\mathbf{V}$ a $\mathbf{W}$ jsou konečně generované komplexní (resp. reálné) lineární prostory se skalárním součinem, B je ortonormální báze prostoru $\mathbf{V}$ a C je ortonormální báze prostoru $\mathbf{W}$. Pak*

$$[f^*]_B^C = ([f]_C^B)^* \ .$$

Tímto vztahem by také bylo možné sdružené lineární zobrazení definovat, museli bychom ale ukázat, že f^* nezávisí na volbě ortonormálních bází.

Ujasníme si vztah hermitovsky sdružené matice k matici A a sdruženého lineárního zobrazení k lineárnímu zobrazení f_A.

Pozorování 10.8. *Pro libovolnou komplexní (resp. reálnou) matici A typu $m \times n$ platí*

$$(f_A)^* = f_{A^*} \ ,$$

kde sdružování na levé straně je vzhledem ke standardnímu skalárnímu součinu.

Důkaz. Plyne z důsledku 10.7 volbou kanonickách bází, protože $[(f_A)^*]_K^K = ([f_A]_K^K)^* = A^*$, je $(f_A)^* = f_{A^*}$. □

V reálném případě tedy $(f_A)^* = f_{A^T}$.

Příklad 10.9. Sdružené lineární zobrazení k lineárnímu zobrazení $f_A : \mathbb{C}^3 \to \mathbb{C}^2$ (na obou prostorech uvažujeme standardní skalární součin), kde

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

je lineární zobrazení $(f_A)^* = f_{A^*}$ určené maticí.

$$A^* = \begin{pmatrix} 1-i & 3+2i \\ 3 & 4-4i \\ 2 & 1 \end{pmatrix}$$

Následující tvrzení shrnuje některé jednoduché vlastnosti sdružování, které budeme používat automaticky.

Tvrzení 10.10. *Nechť $\mathbf{V}$, $\mathbf{W}$ jsou konečně generované vektorové prostory se skalárním součinem nad $\mathbb{C}$ (resp. $\mathbb{R}$), f, g jsou lineární zobrazení $\mathbf{V} \to \mathbf{W}$, a $a \in \mathbb{C}$ (resp. $a \in \mathbb{R}$). Pak platí*

(1) $f^{**} = f$,
(2) $(f+g)^* = f^* + g^*$,
(3) $(af)^* = \overline{a} f^*$,
(4) $(fg)^* = g^* f^*$,
(5) *je-li f izomorfismus, pak je f^* izomorfismus a platí* $(f^{-1})^* = (f^*)^{-1}$.

Důkaz. Všechny uvedené vlastnosti můžeme dokázat porovnáním matic operátorů vzhledem k nějakým ortonormálním bázím B, C, využitím důsledku 10.7 a vlastností hermitovského sdružování (resp. transponování) matic. Lepší možnost je vyjít přímo z definice.

Například pro důkaz (2) můžeme počítat

$$[(f+g)^*]^C_B = ([f+g]^B_C)^* = ([f]^B_C+[g]^B_C)^* = ([f]^B_C)^*+([g]^B_C)^* = [f^*]^C_B+[g^*]^C_B = [f^*+g^*]^C_B \ ,$$

tedy $(f+g)^* = f^* + g^*$. Pro ověření z definice spočítáme

$$\begin{aligned}\langle\, (f^*+g^*)(\mathbf{x}), \mathbf{y}\rangle &= \langle\, f^*(\mathbf{x}) + g^*(\mathbf{x}), \mathbf{y}\rangle = \langle\, f^*(\mathbf{x}), \mathbf{y}\rangle + \langle\, g^*(\mathbf{x}), \mathbf{y}\rangle \\ &= \langle\, \mathbf{x}, f(\mathbf{y})\rangle + \langle\, \mathbf{x}, g(\mathbf{y})\rangle = \langle\, \mathbf{x}, f(\mathbf{y}) + g(\mathbf{y})\rangle = \langle\, \mathbf{x}, (f+g)(\mathbf{y})\rangle \ ,\end{aligned}$$

takže platí $(f+g)^* = f^* + g^*$.

Ostatní vlastnosti přenecháme jako cvičení čtenáři. □

Opustíme obecná lineární zobrazení a vrátíme se zpět k operátorům. Důležitou vlastností sdružených operátorů je, že jejich vlastní čísla jsou komplexně sdružená k vlastním číslům původního operátoru.

Tvrzení 10.11. *Nechť* $\mathbf{V}$ *je konečně generovaný komplexní (resp. reálný) vektorový prostor se skalárním součinem a* f *je lineární operátor na* $\mathbf{V}$*. Pak* $\lambda \in \mathbb{C}$ *(resp.* $\lambda \in \mathbb{R}$*) je vlastní číslo operátoru* f *právě tehdy, když je* $\overline{\lambda}$ *(resp.* λ*) vlastní číslo operátoru* f^**.*

Důkaz. Díky vlastnosti $f^{**} = f$ stačí dokázat jednu implikaci. Předpokládejme, že λ je vlastní číslo operátoru f. Pak operátor $f - \lambda\,\mathrm{id}$ není prostý, tedy není to izomorfismus. Pak ani $(f - \lambda\,\mathrm{id})^* = f^* - \overline{\lambda}\,\mathrm{id}^* = f^* - \overline{\lambda}\,\mathrm{id}$ není izomorfismem, takže není prostý (připomeňme, že pro operátory na konečně generovaném platí „prostý ⇔ na ⇔ izo“). Z toho vyplývá, že $\overline{\lambda}$ je vlastní číslo operátoru f^*. □

Pomocí pozorování 10.8 můžeme tvrzení přeformulovat pro matice: λ je vlastní číslo komplexní čtvercové matice A právě tehdy, když je $\overline{\lambda}$ vlastní číslo matice A^*. Speciálně, reálná čtvercová matice A a transponovaná matice A^T mají stejná vlastní čísla.

Jako cvičení dokažte, že charakteristický polynom operátoru f je „komplexně sdružený polynom“ k charakteristickému polynomu operátoru f^*. To dává alternativní důkaz předchozího tvrzení.

Tvrzení nedává žádnou informaci o příslušných vlastních vektorech, žádný jednoduchý vztah totiž obecně neplatí.

Příklad 10.12. Reálná matice

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

má vlastní čísla 1 a -2 a příslušné podprostory vlastních čísel $M_1 = \left\langle(-1,4)^T\right\rangle$, $M_{-2} = \left\langle(-1,1)^T\right\rangle$. Matice transponovaná má stejná vlastní čísla a $M_1 = \left\langle(1,1)^T\right\rangle$, $M_{-2} = \left\langle(4,1)^T\right\rangle$.

10.1.2. *Normální operátory.*

Definice 10.13. Operátor na komplexním (resp. reálném) lineárním prostoru $\mathbf{V}$ se skalárním součinem se nazývá *normální*, pokud $f^*f = ff^*$.

Definice 10.14 (10.13*.)**.** Komplexní (resp. reálná) čtvercová matice A se nazývá *normální*, pokud $A^*A = AA^*$ (v reálném případě můžeme psát $A^TA = AA^T$).

Pojem normální matice je zaveden v souladu s pojmem normální operátor – matice A je normální právě tehdy, když je normální operátor f_A na prostoru $\mathbb{C}^n$ (resp. $\mathbb{R}^n$) se standardním skalárním součinem.

Příklady normálních operátorů zahrnují operátory unitární (v reálném případě ortogonální) a operátory hermitovské, kterými se budeme zabývat v dalším odstavci. Příklady normálních matic jsou tedy unitární matice (ortogonální matice) a hermitovské matice (symetrické matice). Dále také například diagonální matice a antihermitovské (antisymetrické) matice, tj. matice splňucí $-A^* = A$.

Příklad 10.15. Reálná matice

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

je normální, protože

$$A^T A = AA^T = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix} .$$

Matice A není symetrická, antisymetrická, ani ortogonální.

Skalární násobek normálního operátoru je rovněž normální a sdružený operátor k normálnímu operátoru je normální (cvičení). Součet ani složení dvou normálních operátorů ale normální být nemusí, stačí ale, aby operátory komutovaly. Budeme potřebovat pouze speciální případ:

Tvrzení 10.16. *Je-li f normální operátor na komplexním (reálném) vektorovém prostoru $\mathbf{V}$ se skalárním součinem a $t \in \mathbb{C}$ ($t \in \mathbb{R}$), pak je operátor $f - t\,\mathrm{id}_V$ také normální.*

Důkaz. Je $(f - t\,\mathrm{id})^*(f - t\,\mathrm{id}) = (f^* - (t\,\mathrm{id})^*)(f - t\,\mathrm{id}) = (f^* - \overline{t}\,\mathrm{id})(f - t\,\mathrm{id}) = f^*f - tf^* - \overline{t}f + t\overline{t}\,\mathrm{id}$. Stejně vyjde i $(f - t\,\mathrm{id})(f - t\,\mathrm{id})^*$. □

Normální operátory se vyznačují tím, že se normy f-obrazu a f^*-obrazu libovolného vektoru rovnají.

Tvrzení 10.17. *Nechť f je normální operátor na komplexním (resp. reálném) vektorovém prostoru $\mathbf{V}$ se skalárním součinem a $\mathbf{v} \in V$. Pak platí*

$$\|f(\mathbf{v})\| = \|f^*(\mathbf{v})\| \quad .$$

Důkaz. Protože norma je vždy nezáporné reálné číslo, stačí dokázat rovnost druhých mocnic norem.

$$\begin{aligned} \|f(\mathbf{v})\|^2 &= \langle\, f(\mathbf{v}), f(\mathbf{v})\rangle = \langle\, f^*f(\mathbf{v}), \mathbf{v}\rangle = \langle\, ff^*(\mathbf{v}), \mathbf{v}\rangle \\ &= \langle\, f^{**}f^*(\mathbf{v}), \mathbf{v}\rangle = \langle\, f^*(\mathbf{v}), f^*(\mathbf{v})\rangle = \|f^*(\mathbf{v})\|^2 \end{aligned}$$

□

Jako cvičení dokažte, že vlastnost v předchozím tvrzení normální operátory charakterizuje.

Z tvrzení 10.11 víme, že λ je vlastní číslo operátoru f právě tehdy, když je $\overline{\lambda}$ vlastní číslo operátoru f^*. Příslušné vlastní vektory ale nejsou obecně v jednoduchém vztahu. Pro normální operátory je situace daleko přehlednější.

Tvrzení 10.18. *Nechť f je normální operátor na komplexním (resp. reálném) vektorovém prostoru $\mathbf{V}$ se skalárním součinem, $\lambda \in \mathbb{C}$ (resp. $\lambda \in \mathbb{R}$) a $\mathbf{v} \in V$. Pak $\mathbf{v}$ je vlastní vektor operátoru f příslušný vlastnímu číslu λ právě tehdy, když je $\mathbf{v}$ vlastní vektor operátoru f^* příslušný vlastnímu číslu $\overline{\lambda}$.*

Důkaz. Předpokládejme, že $\mathbf{v}$ je vlastní vektor operátoru f příslušný vlastnímu číslu λ. Z tvzení 10.11 již víme, že $\overline{\lambda}$ je vlastní číslo operátoru f^*, zbývá dokázat, že $\mathbf{v}$ je příslušný vlastní vektor. Platí $(f - \lambda\,\mathrm{id})(\mathbf{v}) = \mathbf{o}$, tedy také $\|(f - \lambda\,\mathrm{id})(\mathbf{v})\| = 0$. Protože f je normální, je podle tvrzení 10.16 normální také operátor $f - \lambda\,\mathrm{id}$. Z tvrzení 10.17 o normách vyplývá $\|(f - \lambda\,\mathrm{id})^*(\mathbf{v})\| = \left\|(f^* - \overline{\lambda}\,\mathrm{id})(\mathbf{v})\right\| = 0$. Z toho $(f^* - \overline{\lambda}\,\mathrm{id})(\mathbf{v}) = \mathbf{o}$, tedy $\mathbf{v}$ je skutečně vlastní vektor operátoru f^* příslušný vlastnímu číslu $\overline{\lambda}$.

Pro důkaz druhé implikace stačí připomenout, že f^* je normální operátor a $f^{**} = f$. □

Dostáváme se k hlavní větě o normálních operátorech.

Věta 10.19 (Spektrální věta pro normální operátory)**.** *Nechť* $\mathbf{V}$ *je konečně generovaný vektorový prostor nad* $\mathbb{C}$ *se skalárním součinem a* f *lineární operátor na* $\mathbf{V}$ *(resp. nechť* A *je čtvercová matice nad* $\mathbb{C}$*) . Pak následující tvrzení jsou ekvivalentní.*

(1) *Operátor* f *(resp. matice* A*) je unitárně diagonalizovatelný (-á).*
(2) *Operátor* f *(resp. matice* A*) je normální.*

Důkaz. (1) $\Rightarrow$ (2). Je-li B ortonormální báze taková, že $[f]_B^B = D = \mathrm{diag}(t_1, \dots, t_n)$ je diagonální, pak

$$[f^* f]_B^B = [f^*]_B^B [f]_B^B = ([f]_B^B)^* [f]_B^B = D^* D$$

a podobně $[ff^*]_B^B = DD^*$. Protože $D^*D = DD^* = \mathrm{diag}(|t_1|^2, \dots, |t_n|^2)$, platí $[f^* f]_B^B = [ff^*]_B^B$, tedy také $f^*f = ff^*$.

(2) $\Rightarrow$ (1). Tvrzení dokážeme indukcí podle dimenze n prostoru $\mathbf{V}$. Pro $n = 1$ tvrzení zřejmě platí. Předpokládejme, že f je normální a každý normální operátor na prostoru dimenze $n-1$ je unitárně diagonalizovatelný. Chceme ukázat, že f je unitárně diagonalizovatelný, ekvivalentně, pro f existuje ortonormální báze složená z vlastních vektorů.

Každý operátor na konečně generovaném prostoru nad $\mathbb{C}$ má vlastní číslo λ (viz důsledek 9.43). Vezmeme libovolný nenulový vlastní vektor příslušný vlastnímu číslu λ, znormujeme jej a označíme $\mathbf{v}_n$.

Ukážeme, že $\mathbf{W} = \mathbf{v}_n^\perp$ je invariantní prostor operátoru f. Nechť tedy $\mathbf{x} \in W$ je libovolný vektor. Protože $\mathbf{x}$ je kolmý na $\mathbf{v}_n$, platí $\langle \mathbf{v}_n, \mathbf{x} \rangle = 0$. Pak také

$$\langle \mathbf{v}_n, f(\mathbf{x}) \rangle = \langle f^*(\mathbf{v}_n), \mathbf{x} \rangle = \langle \overline{\lambda} \mathbf{v}_n, \mathbf{x} \rangle = \lambda \langle \mathbf{v}_n, \mathbf{x} \rangle = 0 \ ,$$

kde v druhé úpravě jsme využili tvrzení 10.18 o vlastních vektorech normálního operátoru. Takže skutečně je $f(\mathbf{x}) \in \mathbf{v}_n^\perp = W$.

Prostor $\mathbf{W}$ je ortogonální doplněk jednodimenzionálního prostoru, má tedy dimenzi $n-1$ (viz větu 8.68). Na zúžení $f|_W$ použijeme indukční předpoklad a získáme ortonormální bázi $C = (\mathbf{v}_1, \dots, \mathbf{v}_{n-1})$ prostoru $\mathbf{W}$ tvořenou vlastními vektory operátoru f. Pak je posloupnost $B = (\mathbf{v}_1, \dots, \mathbf{v}_{n-1}, \mathbf{v}_n)$ tvořená vlastními vektory operátoru f, je ortonormální (protože C je ortonormální, $\mathbf{v}_n$ je jednotkový a kolmý na všechny vektory z C), takže B je ortonormální báze prostoru $\mathbf{V}$ tvořená vlastními vektory operátoru f.

Alternativně lze větu dokázat užitím Jordanova kanonického tvaru, viz cvičení. □

Speciálně, normální reálná matice je unitárně diagonalizovatelná, **chápeme-li ji jako matici nad** $\mathbb{C}$. Není pravda, že je nutně ortogonálně diagonalizovatelná nad $\mathbb{R}$! Reálné matice, které jsou ortogonálně diagonalizovatelné, charakterizujeme v důsledku 10.23 – jsou to přesně symetrické matice.

Příklad 10.20. V příkladu 10.15 jsme viděli, že reálná matice

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

je normální. Její charakteristický polynom je

$$p_A(t) = -t^3 + 3t^2 - 3t + 2 = -(t-2)(t^2 - t + 1) \ .$$

Tento polynom má pouze jeden reálný kořen $\lambda = 2$ násobnosti 1, matice A tedy není ortogonálně diagonalizovatelná.

Chápejme nyní A jako matici nad $\mathbb{C}$. Podle spektrální věty je unitárně diagonalizovatelná. Má tři vlastní čísla

$$\lambda_1 = 2,\ \lambda_2 = \frac{1}{2} + \frac{\sqrt{3}}{2} i,\ \lambda_3 = \overline{\lambda_2} = \frac{1}{2} - \frac{\sqrt{3}}{2} i \ ,$$

prostory vlastních vektorů mají tedy dimenzi 1 a stačí v každém z nich zvolit jednotkový vektor.

$$\mathbf{v}_1 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \ \mathbf{v}_2 = \frac{1}{\sqrt{3}} \begin{pmatrix} \frac{-1+\sqrt{3}i}{2} \\ \frac{-1-\sqrt{3}i}{2} \\ 1 \end{pmatrix}, \mathbf{v}_3 = \overline{\mathbf{v}_2} = \frac{1}{\sqrt{3}} \begin{pmatrix} \frac{-1-\sqrt{3}i}{2} \\ \frac{-1-\sqrt{3}i}{2} \\ 1 \end{pmatrix}$$

Vzhledem k bázi $B = (\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ je matice operátoru $f_A : \mathbb{C}^3 \to \mathbb{C}^3$ rovná

$$[f_A]_B^B = \begin{pmatrix} 2 & 0 & 0 \\ 0 & \frac{1+\sqrt{3}i}{2} & 0 \\ 0 & 0 & \frac{1-\sqrt{3}i}{2} \end{pmatrix} .$$

To nám také dává maticový rozklad

$$A = [f_A]_K^K = [\mathrm{id}]_K^B [f_A]_B^B [\mathrm{id}]_B^K = [\mathrm{id}]_K^B [f_A]_B^B ([\mathrm{id}]_K^B)^{-1} = [\mathrm{id}]_K^B [f_A]_B^B ([\mathrm{id}]_K^B)^*$$

$$= \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & \frac{-1+\sqrt{3}i}{2} & \frac{-1-\sqrt{3}i}{2} \\ 1 & \frac{-1-\sqrt{3}i}{2} & \frac{-1+\sqrt{3}i}{2} \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 & 0 \\ 0 & \frac{1+\sqrt{3}i}{2} & 0 \\ 0 & 0 & \frac{1-\sqrt{3}i}{2} \end{pmatrix} \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 1 & 1 \\ \frac{-1-\sqrt{3}i}{2} & \frac{-1+\sqrt{3}i}{2} & 1 \\ \frac{-1+\sqrt{3}i}{2} & \frac{-1-\sqrt{3}i}{2} & 1 \end{pmatrix} .$$

10.1.3. *Hermitovské a symetrické operátory.* Důležitou podtřídou normálních operátorů jsou operátory hermitovské (v reálném případě symetrické).

Definice 10.21. Operátor na komplexním (resp. reálném) lineárním prostoru $\mathbf{V}$ se skalárním součinem se nazývá *hermitovský* (resp. *symetrický*), pokud $f^* = f$.

Pojem hermitovského (symetrického) operátoru je zaveden v souladu s pojmem hermitovské (symetrické) matice – matice A je hermitovská (symetrická) právě tehdy, když je hermitovský (symetrický) operátor f_A na prostoru $\mathbb{C}^n$ (resp. $\mathbb{R}^n$) se standardním skalárním součinem.

Hermitovské operátory jsou přesně ty normální operátory, jejichž vlastní čísla jsou reálná.

Věta 10.22 (Spektrální věta pro hermitovské operátory)**.** *Nechť* $\mathbf{V}$ *je konečně generovaný vektorový prostor nad* $\mathbb{C}$ *se skalárním součinem a* f *je lineární operátor na* $\mathbf{V}$ *(resp. nechť* A *je čtvercová matice nad* $\mathbb{C}$*). Pak následující tvrzení jsou ekvivalentní.*

(1) *Operátor* f *(resp. matice* A*) je unitárně diagonalizovatelný (-á) a všechna jeho (její) vlastní čísla jsou reálná.*
(2) *Operátor* f *(resp. matice* A*) je hermitovský (-á).*

Důkaz. (1) $\Rightarrow$ (2). Je-li f unitárně diagonalizovatelný a všechna jeho vlastní čísla jsou reálná, pak existuje ortonormální báze B prostoru $\mathbf{V}$ taková, že $[f]_B^B = D$ je reálná diagonální matice. Platí tedy

$$[f^*]_B^B = ([f]_B^B)^* = D^* = D = [f]_B^B \ ,$$

neboli $f^* = f$.

(2) $\Rightarrow$ (1). Protože každý hermitovský operátor je normální, stačí podle spektrální věty o normálních operátorech (věta 10.19) ukázat, že všechna vlastní čísla operátoru f jsou reálná. To nahlédneme z tvrzení 10.18 o vlastních vektorech normálních operátorů: Je-li $\lambda \in \mathbb{C}$ vlastní číslo operátoru f a $\mathbf{v}$ nenulový vlastní vektor příslušný λ, pak $\mathbf{v}$ je vlastní vektor operátoru $f^* = f$ příslušný vlastnímu číslu $\overline{\lambda}$. Jeden nenulový vektor nemůže příslušet více vlastním číslům, platí tedy $\lambda = \overline{\lambda}$, neboli $\lambda \in \mathbb{R}$. □

Důsledek 10.23 (Spektrální věta pro symetrické operátory)**.** *Nechť* $\mathbf{V}$ *je konečně generovaný vektorový prostor nad* $\mathbb{R}$ *se skalárním součinem a* f *je lineární operátor na* $\mathbf{V}$ *(resp. nechť* A *je čtvercová matice nad* $\mathbb{R}$*). Pak následující tvrzení jsou ekvivalentní.*

(1) *Operátor* f *(resp. matice* A*) je ortogonálně diagonalizovatelný (-á).*

(2) *Operátor f (resp. matice A) je symetrický (-á).*

Důkaz. Důkaz (1) $\Rightarrow$ (2) se udělá stejně jako v předchozí větě.

Dokážeme implikaci (2) $\Rightarrow$ (1) v maticové verzi. Předpokládejme tedy, že A je reálná symetrická matice. Chápejme nyní A jako matici nad $\mathbb{C}$. Protože je A hermitovská, podle předchozí věty je unitárně diagonalizovatelná a všechna vlastní čísla jsou reálná. Z toho vyplývá (viz větu 10.2), že A má n reálných vlastních čísel včetně násobností, geometrická násobnost každého vlastního čísla je rovná jeho algebraické násobnosti a prostory M_λ jsou navzájem kolmé (vzhledem ke standardnímu skalárnímu součinu). Algebraická (geometrická) násobnost nad $\mathbb{C}$ je rovná algebraické (geometrické) násobnosti nad $\mathbb{R}$, takže chápeme-li A opět jako reálnou matici, bude splňovat podmínky z věty 10.2, a bude proto ortogonálně diagonalizovatelná. □

Příklad 10.24. Jako ilustraci spektrální věty pro reálné symetrické matice najdeme pro operátor f_A určený maticí

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

ortonormální bázi složenou z vlastních vektorů.

Operátor f_A má charakteristický polynom $p(\lambda) = (1-\lambda)(\lambda^2 - 1)$ a tedy vlastní čísla 1 a -1. Příslušné prostory vlastních vektorů jsou

$$\mathbf{M}_1 = \left\langle (1,1,0)^T, (0,0,1)^T \right\rangle, \mathbf{M}_{-1} = \left\langle (1,-1,0)^T \right\rangle$$

V prostoru $\mathbf{M}_1$ je ortonormální báze například $(\mathbf{v}_1, \mathbf{v}_2)$, kde

$$\mathbf{v_1} = \frac{\sqrt{2}}{2} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \ \mathbf{v}_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \ .$$

V prostoru $\mathbf{M}_{-1}$ tvoří ortonromální bázi například vektor

$$\mathbf{v}_3 = \frac{\sqrt{2}}{2} \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \ .$$

Báze $B = (\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ je ortonormální báze prostoru $\mathbf{V}$ a $[f_A]_B^B = \mathrm{diag}(1,1,-1)$.

Zapíšeme ještě výsledek maticově. Označme

$$Q = [\mathrm{id}]_K^B = \begin{pmatrix} \frac{\sqrt{2}}{2} & 0 & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & 0 & -\frac{\sqrt{2}}{2} \\ 0 & 1 & 0 \end{pmatrix} \ .$$

Matice Q je ortogonální, takže $Q^{-1} = Q^T$ a můžeme psát

$$\mathrm{diag}(1,1,-1) = [f_A]_B^B = [\mathrm{id}]_B^K [f_A]_K^K [\mathrm{id}]_K^B = Q^{-1} A Q = Q^T A Q \ ,$$

nebo ve formě rozkladu matice A

$$A = Q \,\mathrm{diag}(1,1,-1) Q^{-1} = Q \,\mathrm{diag}(1,1,-1) Q^T \ .$$

10.1.4. *Pozitivně (semi)definitní operátory.* Je-li f hermitovský operátor na $\mathbf{V}$, pak pro libovolný vektor $\mathbf{x} \in V$ platí

$$\langle \mathbf{x}, f(\mathbf{x}) \rangle = \langle f^*(\mathbf{x}), \mathbf{x} \rangle = \langle f(\mathbf{x}), \mathbf{x} \rangle = \overline{\langle \mathbf{x}, f(\mathbf{x}) \rangle} \ .$$

Z toho plyne, že $\langle \mathbf{x}, f(\mathbf{x}) \rangle$ je vždy reálné číslo. Pro $\mathbf{x} = \mathbf{o}$ je rovno 0. Ty operátory, pro které je jinak toto číslo vždy kladné (resp. nezáporné) nazýváme pozitivně definitní (resp. semidefinitní).

Definice 10.25. Operátor f na konečně generovaném komplexním (resp. reálném) lineárním prostoru $\mathbf{V}$ se skalárním součinem se nazývá

- *pozitivně definitní*, pokud je hermitovský (resp. symetrický) a pro všechna $\mathbf{o} \neq \mathbf{x} \in V$ platí $\langle\, \mathbf{x}, f(\mathbf{x})\rangle > 0$;
- *pozitivně semidefinitní*, pokud je hermitovský (resp. symetrický) a pro všechna $\mathbf{x} \in V$ platí $\langle\, \mathbf{x}, f(\mathbf{x})\rangle \geq 0$.

Pro matice je pojem jako obykle definován pomocí příslušného operátoru a standardního skalárního součinu. Explicitně:

Definice 10.26 (10.25*.)**.** Čtvercová matice A nad $\mathbb{C}$ (resp. $\mathbb{R}$) se nazývá

- *pozitivně definitní*, pokud je hermitovská (resp. symetrická) a pro všechna $\mathbf{o} \neq \mathbf{x} \in V$ platí $\mathbf{x}^* A\mathbf{x} > 0$;
- *pozitivně semidefinitní*, pokud je hermitovská (resp. symetrická) a $\mathbf{x}^* A\mathbf{x} \geq 0$.

Pozitivně definitní operátory nebo matice lze ekvivalentně definovat tak, že jsou hermitovské (symetrické) a všechna vlastní čísla jsou kladná. Podobně pro semidefinitnost.

Tvrzení 10.27. *Nechť f je hermitovský (symetrický) operátor f na komplexním (reálném) vektorovém prostoru $\mathbf{V}$ se skalárním součinem. Pak f je pozitivně definitní (resp. semidefinitní) právě tehdy, když jsou všechna vlastní čísla operátoru f kladná (resp. nezáporná).*

Důkaz. Dokážeme pouze verzi pro pozitivně definitní operátory. Pro semidefinitní je důkaz téměř totožný.

Je-li f pozitivně definitní a λ je vlastní číslo operátoru f (to je nutně reálné), pak pro nenulový vlastní vektor $\mathbf{v}$ příslušný λ platí $0 < \langle\, \mathbf{v}, f(\mathbf{v})\rangle = \langle\, \mathbf{v}, \lambda\mathbf{v}\rangle = \lambda \|\mathbf{v}\|^2$. Protože norma je kladná, plyne odsud $\lambda > 0$.

Jsou-li naopak všechna vlastní čísla operátoru f kladná, existuje podle spektrální věty pro hermitovské matice (věta 10.22) ortonormální báze B prostoru $\mathbf{V}$ taková, že $[f]_B^B = D = \mathrm{diag}(t_1, \dots, t_n)$, kde t_i jsou vlastní čísla operátoru f, tj. $t_i > 0$ pro všechna $i \in \{1, \dots, n\}$. Pak užitím tvrzení o skalárním součinu vzhledem k ortonormální bázi dostáváme pro libovolný vektor $\mathbf{x}$ vztah

$$\langle\, \mathbf{x}, f(\mathbf{x})\rangle = ([\mathbf{x}]_B)^*[f(\mathbf{x})]_B = ([\mathbf{x}]_B)^*[f]_B^B[\mathbf{x}]_B = ([\mathbf{x}]_B)^* D[\mathbf{x}]_B \ .$$

Označíme-li $[\mathbf{x}]_B = (x_1, \dots, x_n)$ je výraz roven $t_1|x_1|^2 + \cdots + t_n|x_n|^2$, což je ostře větší než 0, kdykoliv $\mathbf{x} \neq 0$. □

Všimněte si také, že pozitivně definitní operátor je vždy izomorfismus (a pozitivně definitní matice je regulární), protože nemůže mít vlastní číslo 0.

Příklad 10.28. Reálná matice

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

není pozitivně definitní ani semidefinitní, protože má záporné vlastní číslo -1.

Reálná matice

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

má vlastní čísla 0, 5, je proto pozitivně semidefinitní, ale není pozitivně definitní.

Reálná matice

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

je pozitivně definitní. Pro libovolný nenulový vektor $(a, b) \in \mathbb{R}^2$ je proto číslo

$$(a, b)A \begin{pmatrix} a \\ b \end{pmatrix} = a^2 + 4ab + 5b^2$$

kladné.

Mnoho soustav lineárních rovnic vzniklých přeformulováním úloh z přírodních věd má pozitivně (semi)definitní matici. Je to často způsobeno tím, že matice soustavy je tvaru A^TA pro nějakou reálnou obdélníkovou matici, nebo obecněji A^TDA pro matici A typu $m\times n$ a diagonální matici $D = \mathrm{diag}(d_1,\dots,d_m)$ řádu m s nezápornými prvky na diagonále.

Matice tvaru A^TA jsou skutečně pozitivně semidefinitní, protože jsou symetrické

$$(A^TA)^T = A^TA = A^TA$$

a pro libovolný vektor $\mathbf{x} \in \mathbb{R}^n$ platí

$$\mathbf{x}^TA^TA\mathbf{x} = \mathbf{x}\cdot A^TA\mathbf{x} = A\mathbf{x}\cdot A\mathbf{x} = \|A\mathbf{x}\|^2 \quad .$$

(Stejný výsledek můžeme získat přímo maticovým výpočtem $\mathbf{x}^TA^TA\mathbf{x} = (A\mathbf{x})^TA\mathbf{x} = \|A\mathbf{x}\|^2$.) Také vidíme, že matice A^TA je pozitivně definitní právě tehdy, když má soustava $A\mathbf{x} = \mathbf{o}$ pouze triviální řešení, tzn. právě tehdy, když má matice A lineárně nezávislou posloupnost sloupcových vektorů.

Matice tvaru A^TDA, kde $D = \mathrm{diag}(d_1,\dots,d_m)$, $d_i \geq 0$, lze psát ve tvaru $(\sqrt{D}A)^T(\sqrt{D}A)$, kde $\sqrt{D}$ značí matici $\mathrm{diag}(\sqrt{d_1},\dots,\sqrt{d_m})$, jsou tedy také pozitivně semidefinitní.

Podobně, každá komplexní matice tvaru A^*A je pozitivně semidefinitní. Formulujeme verzi pro lineární zobrazení.

Tvrzení 10.29. *Nechť $\mathbf{V}$, $\mathbf{W}$ jsou vektorové prostory se skalárními součiny a $f : \mathbf{V} \to \mathbf{W}$ je lineární zobrazení. Pak je operátor f^*f pozitivně semidefinitní.*

Důkaz. Operátor f^*f je hermitovský, protože $(f^*f)^* = f^*f^{**} = f^*f$. Pro libovolný vektor $\mathbf{x} \in V$ platí $\langle \mathbf{x}, f^*f(\mathbf{x})\rangle = \langle f(\mathbf{x}), f(\mathbf{x})\rangle = \|f(\mathbf{x})\|^2 \geq 0$, takže operátor f^*f je pozitivně semidefinitní. □

Naopak platí, že každý semidefinitní operátor g lze psát ve tvaru $g = f^*f$ pro nějaký operátor f. Důkaz přecháme čtenáři jako cvičení.

10.1.5. *Unitární operátory.* Další důležitou podtřídou normálních operátorů jsou operátory unitární, v reálném případě ortogonální. Připomeňme, že f je unitární, pokud f zachovává normu. V tvrzení **??** jsme uvedli řadu ekvivalentních definic. Pomocí sdruženého operátoru můžeme unitární operátory charakterizovat pomocí vlastnosti $f^* = f^{-1}$, podobně jako unitární matice:

Tvrzení 10.30. *Operátor f na konečně generovaném prostoru $\mathbf{V}$ nad $\mathbb{C}$ (resp. $\mathbb{R}$) je unitární (resp. ortogonální) právě tehdy, když $f^* = f^{-1}$.*

Důkaz. Je-li zobrazení f unitární, pak je podle pozorování **??** prosté, tedy f je izomorfismus. Inverzní zobrazení f^{-1} proto existuje. Pro libovolné $\mathbf{x}, \mathbf{y} \in V$ platí

$$\left\langle f^{-1}(\mathbf{x}), \mathbf{y}\right\rangle = \left\langle f(f^{-1}(\mathbf{x})), f(\mathbf{y})\right\rangle = \langle \mathbf{x}, f(\mathbf{y})\rangle \quad .$$

(V první úpravě využíváme toho, že unitární zobrazení zachovává skalární součin, viz tvrzení **??**.) Z toho vyplývá $f^* = f^{-1}$.

Platí-li naopak $f^* = f^{-1}$, pak

$$\|f(\mathbf{x})\| = \sqrt{\langle f(\mathbf{x}), f(\mathbf{x})\rangle} = \sqrt{\langle f^*f(\mathbf{x}), \mathbf{x}\rangle} = \sqrt{\langle \mathbf{x}, \mathbf{x}\rangle} = \|\mathbf{x}\| \quad .$$

Zobrazení f je tedy unitární. □

Z tvrzení také vidíme, že unitární (ortogonální) operátory jsou skutečně normální, protože $f^*f = ff^* = \mathrm{id}$.

Věta 10.31 (Spektrální věta pro unitární operátory)**.** *Nechť $\mathbf{V}$ je konečně generovaný vektorový protor nad $\mathbb{C}$ se skalárním součinem a f je lineární operátor na $\mathbf{V}$ (resp. nechť A je čtvercová matice nad $\mathbb{C}$). Pak následující tvrzení jsou ekvivalentní.*

(1) *Operátor f (resp. matice A) je unitárně diagonalizovatelný (-á) a pro všechna vlastní čísla $\lambda \in \mathbb{C}$ platí $|\lambda| = 1$.*

(2) *Operátor f (resp. matice A) je unitární.*

Důkaz. (1) $\Rightarrow$ (2). Jsou-li splněny předpoklady, pak $[f]_B^B = D$, kde B je ortonormální a na diagonále D jsou komplexní čísla s absolutní hodnotou 1. Sloupce matice D tvoří ortonormální posloupnost vzhledem ke standardnímu skalárnímu součinu, takže f je unitární podle bodu (5) tvrzení **??**.

(2) $\Rightarrow$ (1). Protože každý unitární operátor je normální, stačí podle věty 10.19 ukázat, že pro všechna vlastní čísla λ operátoru f platí $|\lambda| = 1$. Vezmeme libovolný nenulový vlastní vektor $\mathbf{v}$ příslušný vlastnímu číslu λ. Protože f zachovává normu, platí $\|\mathbf{v}\| = \|f(\mathbf{v})\| = \|\lambda\mathbf{v}\| = |\lambda|\,\|\mathbf{v}\|$. Z toho plyne $|\lambda| = 1$, jak jsme chtěli. □

10.1.6. *Ortogonální operátory v dimenzi* 2. Jak vypadají ortogonální operátory f na reálném prostoru dimenze 2? Pro zjednodušení zápisu budeme uvažovat prostor $\mathbb{R}^2$ se standardním skalárním součinem.

Nechť tedy $f : \mathbb{R}^2 \to \mathbb{R}^2$ je ortogonální operátor. Obrazy vektorů nějaké ortonormální báze (např. kanonické) jsou jednotkové a na sebe kolmé. Z toho lze geometricky nahlédnout, že f je buď rotace nebo reflexe (osová souměrnost). Toto pozorování teď dokážeme.

Označme si A matici f vzhledem ke kanonické bázi, tj. $f = f_A$. Někdy budeme A chápat jako komplexní matici a $f = f_A$ jako operátor na $\mathbb{C}^2$. Podle charakterizace unitárních operátorů pro všechna vlastní čísla matice A platí $|\lambda| = 1$. Protože má charakteristický polynom matice A reálné koeficienty, jsou obě vlastní čísla buď reálná nebo je tvoří dvojice komplexně sdružených čísel $\cos\varphi + i\sin\varphi = e^{i\varphi}$ a $\cos\varphi - i\sin\varphi = e^{-i\varphi}$. Označme $C = (\mathbf{v}_1, \mathbf{v}_2)$ nějakou ortonormální bázi prostoru $\mathbb{C}^2$ složenou z vlastních vektorů matice A.

Nejdříve probereme případ, kdy vlastní čísla matice A, tj. operátoru f, jsou reálná. Pak můžeme zvolit oba vlastní vektory $\mathbf{v}_1, \mathbf{v}_2$ reálné. Zde máme tři možnosti.

- Obě vlastní čísla se rovnají 1. Matice $[f]_C^C$ se pak rovná I_2 a operátor f se rovná identickému zobrazení.
- Obě vlastní čísla se rovnají -1. Matice $[f]_C^C$ se pak rovná $-I_2$ a operátor f se rovná středové symetrii - stejnolehlosti s koeficientem -1.
- Jedno vlastní číslo se rovná 1 a druhé -1. Matice $[f]_C^C$ se pak rovná
$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} .$$
Zobrazení f je reflexe (osová souměrnost) vzhledem k přímce generované vektorem $\mathbf{v}_1$.

Zbývá případ komplexních vlastních čísel, která nejsou reálná. Označme $\mathbf{v} = (a + bi, c + di)$ vlastní vektor příslušný číslu $\lambda = \cos\varphi + i\sin\varphi$. V části 9.3.4 jsme ukázali, že $\overline{\mathbf{v}}$ je vlastní vektor příslušný vlastnímu číslu $\overline{\lambda}$ a vzhledem k bázi $C' = (\mathbf{w}_1, \mathbf{w}_2) = (\mathbf{v} + \overline{\mathbf{v}}, i(\mathbf{v} - \overline{\mathbf{v}})) = (2Re(\mathbf{v}), -2Im(\mathbf{v}))$ má operátor matici rovnou matici rotace o úhel φ.

Protože $\mathbf{v}$ a $\overline{\mathbf{v}}$ jsou na sebe kolmé, platí
$$\mathbf{v}^*\overline{\mathbf{v}} = (a - ib \,|\, c - id) \begin{pmatrix} a - ib \\ c - id \end{pmatrix} = a^2 - b^2 + c^2 - d^2 - 2i(ab + cd) = 0 \ .$$

Z imaginární části výrazu vidíme, že $ab + cd = 0$ a reálné vektory $\mathbf{w}_1 = 2Re\,\mathbf{v} = (2a, 2c)^T$ a $\mathbf{w}_2 = -2Im\,\mathbf{v} = (-2b, -2d)$ jsou na sebe kolmé. Z reálné části výrazu vidíme, že oba vektory $\mathbf{w}_1, \mathbf{w}_2$ mají stejnou normu $e = \sqrt{4a^2 + 4c^2} = \sqrt{4b^2 + 4d^2}$ Báze
$$B = (\mathbf{w}_1/e, \mathbf{w}_2/e)$$
je ortonormální báze prostoru $\mathbb{R}^2$. Protože vznikla vynásobením vektorů báze C' stejným skalárem je
$$[f]_B^B = [f]_{C'}^{C'} = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}$$
a zobrazení f je tedy rotace.

Pokud vezmeme do úvahy, že středová symetrie je rotace o úhel π a identické zobrazení je rotace o úhel 0, dokázali jsme následující klasifikaci ortogonálních zobrazení v $\mathbb{R}^2$.

Tvrzení 10.32. *Každé ortogonální zobrazení f v lineárním prostoru $\mathbb{R}^2$ se standardním skalárním součinem je buď rotace nebo reflexe. Rotace je to právě když $\det[f]_B^B = 1$ a reflexe je to právě když $\det[f]_B^B = -1$, kde B je libovolná báze $\mathbb{R}^2$.*

Protože složení dvou ortogonálních zobrazení je opět ortogonální zobrazení, dostáváme s použitím věty o součinu determinantů tento důsledek.

Důsledek 10.33. *Složení dvou rotací v $\mathbb{R}^2$ je opět rotace, složení dvou reflexí je rotace a složení rotace s reflexí (v libovolném pořadí) je opět nějaká reflexe.*

10.1.7. *Ortogonální operátory v dimenzi* 3. Rozšíříme klasifikaci ortogonálních zobrazení na reálné prostory dimenze 3.

Nechť $f : \mathbb{R}^3 \to \mathbb{R}^3$ je ortogonální operátor na prostoru $\mathbb{R}^3$ se standardním skalárním součinem a $A = [f]_K^K$ jeho matice vzhledem ke kanonické bázi v $\mathbb{R}^3$. Zobrazení $f = f_A$ budeme také chápat jako unitární operátor na $\mathbb{C}^3$. Charakteristický polynom má všechna vlastní čísla rovná v absolutní hodnotě 1 a existuje ortonormální báze $C = (\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ prostoru $\mathbb{C}^3$ složená z vlastních vektorů matice A. Protože má navíc reálné koeficienty, jsou buď všechna vlastní čísla reálná (rovná ± 1) a nebo je jedno reálné a zbylá dvě jsou komplexně sdružená čísla $e^{i\varphi}$ a $e^{-i\varphi}$ pro nějaký úhel φ.

Předpokládejme, že pouze jedno vlastní číslo je reálné. K němu příslušný vlastní vektor $\mathbf{v}_1$ můžeme proto také zvolit reálný. Podprostor $\langle \mathbf{v}_2, \mathbf{v}_3 \rangle$ prostoru $\mathbb{C}^3$ (ortogonální doplněk vektoru $\mathbf{v}_1$) je invariantní podprostor operátoru f. Na tomto podprostoru dimenze 2 je zúžení f také ortogonální operátor a má komplexní vlastní čísla $e^{i\varphi}$ a $e^{-i\varphi}$. Podle předchozí diskuze je matice zúžení operátoru f na podprostor $\langle \mathbf{v}_2, \mathbf{v}_3 \rangle$ vzhledem k ortonormální bázi $(a\, Re\, \mathbf{v}_2, -a\, Im\, \mathbf{v}_2)$ tohoto podprostoru, kde $a = \| Re\, \mathbf{v}_2 \|^{-1}$, rovná

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

Je-li reálné vlastní číslo rovné 1, má potom f vzhledem k ortonormální bázi $B = (\mathbf{v_1}, a\, Re\, \mathbf{v}_2, -a\, Im\, \mathbf{v}_2)$ prostoru $\mathbb{R}^3$ matici

$$[f]_B^B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\varphi & -\sin\varphi \\ 0 & \sin\varphi & \cos\varphi \end{pmatrix} .$$

a jde tedy o rotaci kolem osy generované vektorem $\mathbf{v}_1$ o úhel φ.

Je-li jediné reálné vlastní číslo operátoru f rovné -1, platí

$$[f]_B^B = \begin{pmatrix} -1 & 0 & 0 \\ 0 & \cos\varphi & -\sin\varphi \\ 0 & \sin\varphi & \cos\varphi \end{pmatrix} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\varphi & -\sin\varphi \\ 0 & \sin\varphi & \cos\varphi \end{pmatrix}$$

a zobrazení f je tedy složením rotace kolem osy generované $\mathbf{v}_1$ o úhel φ s reflexí (zrcadlením) určenou rovinou kolmou na vektor $\mathbf{v}_1$.

Jsou-li všechna vlastní čísla operátoru f reálná, můžeme zvolit ortonormální bázi $\mathbb{C}^3$ složenou z reálných vektorů a matice $[f]_B^B$ má (až na pořadí prvků na hlavní diagonále) jeden ze tvarů

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} .$$

V prvním případě jde o identické zobrazení (tj. rotaci o úhel 0), ve druhém případě jde o zrcadlení vzhledem k rovině $\langle \mathbf{v}_1, \mathbf{v}_2 \rangle = \{\mathbf{v}_3\}^\perp$, ve třetím případě jde o rotaci kolem osy generované $\mathbf{v}_1$ o úhel π a ve čtvrtém případě jde o složení této rotace s reflexí (zrcadlením) určenou rovinou $\langle \mathbf{v}_2, \mathbf{v}_3 \rangle$.

Platí proto následující tvrzení.

Tvrzení 10.34. *Každé ortogonální zobrazení v euklidovském prostoru $\mathbb{R}^3$ je buď rotace kolem nějaké osy, ortogonální reflexe vzhledem k nějaké rovině a nebo složení rotace s ortogonální reflexí. Rotace je to právě tehdy, když determinant matice tohoto zobrazení vzhledem k jakékoliv bázi je rovný* 1.

Důsledek 10.35. *Složení dvou rotací v $\mathbb{R}^3$ je zase rotace v $\mathbb{R}^3$, složení dvou reflexí je rotace (osa rotace je rovná průniku rovin reflexí).*

10.2. **Singulární rozklad.** Najít ortonormální bázi, vzhledem ke které má daný operátor diagonální matici, jde v komplexním případě pro normální operátory a v reálném případě pro symetrické operátory. Když slevíme z požadavku, že báze pro vzory a obrazy jsou stejné, lze „unitárně diagonalizovat" každý lineární operátor, dokonce každé lineární zobrazení (mezi konečně generovanými prostory se skalárním součinem). Navíc na diagonále budou nezáporná reálná čísla.

Začneme ilustrativním příkladem.

Příklad 10.36. Uvažujme „zkosení" $f_A : \mathbb{R}^2 \to \mathbb{R}^2$, kde

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

Budeme umět spočítat, že vzhledem k bázím B a C, kde

$$B = (\mathbf{v}_1, \mathbf{v}_2) \approx \left(\begin{pmatrix} 0{,}526 \\ 0{,}851 \end{pmatrix}, \begin{pmatrix} -0{,}851 \\ 0{,}526 \end{pmatrix} \right) ,$$

$$C = (\mathbf{u}_1, \mathbf{u}_2) \approx \left(\begin{pmatrix} 0{,}851 \\ 0{,}526 \end{pmatrix}, \begin{pmatrix} -0{,}526 \\ 0{,}851 \end{pmatrix} \right) ,$$

má f matici

$$[f_A]^B_C \approx \begin{pmatrix} 1{,}618 & 0 \\ 0 & 0{,}618 \end{pmatrix} .$$

Z toho vidíme, že vektor $\mathbf{v}_1$ se při zobrazení f_A zobrazí na přibližně 1,618-násobek vektoru $\mathbf{u}_1$ a vektor $\mathbf{v}_2$ se zobrazí na přibližně 0,618-násobek vektoru $\mathbf{u}_2$. Obecněji, obraz vektoru $\mathbf{x}$ o souřadnicích $[\mathbf{x}]_B = (x_1, x_2)^T$ vzhledem k bázi B je vektor $A\mathbf{x}$ o souřadnicích $[A\mathbf{x}]_C \approx (1{,}618x_1, 0{,}618x_2)^T$ vzhledem k bázi C.

OBRAZEK

Výhodou matice f_A vzhledem k bázím B a C oproti matici f_A vzhledem ke kanonickým bázím (tj. matici A) je, že snadno určíme obraz jednotkového kruhu $O = \{\mathbf{x} : \|\mathbf{x}\| \leq 1\}$. Protože B je ortonormální báze, platí $\|\mathbf{x}\| = \|[\mathbf{x}]_B\|$. Vektor $\mathbf{x}$ proto leží v O právě tehdy, když jeho souřadnice $[\mathbf{x}]_B = (x_1, x_2)^T$ vzhledem k bázi B splňují $x_1^2 + x_2^2 \leq 1$. Obrazem vektoru $\mathbf{x}$ je vektor $A\mathbf{x}$ o souřadnicích $[A\mathbf{x}]_C = (y_1, y_2)^T \approx (1{,}618x_1, 0{,}618x_2)^T$. V obrazu kruhu O tedy budou právě ty body, jejichž souřadnice vzhledem k C splňují

$$\left(\frac{y_1}{1{,}618}\right)^2 + \left(\frac{y_2}{0{,}618}\right)^2 \leq 1$$

Z toho vidíme, že obrazem je elipsa s délkami poloos (přibližně) 1,618, 0,618. Směry poloos jsou určeny vektory $\mathbf{u}_1, \mathbf{u}_2$ báze C.

Užitečná je také maticová verze. Označíme-li $U = [\mathrm{id}]^C_K$ a $V = [\mathrm{id}]^B_K$, dostaneme z $[f_A]^B_C = D$ vztah $A = UDV^{-1}$. Protože V je ortogonální matice, platí $V^{-1} = V^T$, takže také můžeme psát $A = UDV^T$.

$$A = UDV^T = \begin{pmatrix} 0{,}851 & -0{,}526 \\ 0{,}526 & 0{,}851 \end{pmatrix} \begin{pmatrix} 1{,}618 & 0 \\ 0 & 0{,}618 \end{pmatrix} \begin{pmatrix} 0{,}526 & 0{,}851 \\ -0{,}851 & 0{,}526 \end{pmatrix}$$

Na tento rozklad se také můžeme dívat tak, že zobrazení f_A vyjadřujeme jako složení $f_A = f_U f_D f_{V^T}$, kde f_U a f_{V^T} jsou ortogonální zobrazení a f_D je zobrazení určené diagonální maticí. Zobrazení f_A je tak v našem případě složením zobrazení f_{V^T}, což je rotace o přibližně $-58{,}28^\circ$, zobrazení f_D, které natahuje vektory 1,618-krát ve směru první osy a zkracuje 0,618-krát ve směru druhé osy, a zobrazení f_U, což je rotace o přibližně 31,72°. I z tohoto pohledu vidíme obraz jednotkového kruhu: Zobrazení $f_{V^T} = f_{V^{-1}}$ zobrazí O na O (přičemž vektor $\mathbf{v}_1$ se zobrazí na $\mathbf{e}_1$ a vektor $\mathbf{v}_2$ na $\mathbf{e}_2$). Zobrazení f_D kružnici deformuje ve směru souřadnicových os, čímž vznikne elipsa s osami $\langle \mathbf{e}_1 \rangle$, $\langle \mathbf{e}_2 \rangle$. Tuto elipsu zobrazení f_U otočí.

OBRAZEK

Uvažujme lineární zobrazení $f : \mathbf{V} \to \mathbf{U}$, kde $\mathbf{V}$ a $\mathbf{U}$ jsou konečně generované vektorové prostory nad $\mathbb{C}$ (nebo oba nad $\mathbb{R}$) se skalárními součiny. Chceme najít ortonormální bázi B prostoru $\mathbf{V}$ a ortonormální bázi C prostoru $\mathbf{U}$ tak, že $[f]^B_C$ je "diagonální" s nezápornými reálnými čísly na diagonále.

Tato matice nemusí být čtvercová, pojem diagonální matice proto rozšíříme. Říkáme, že matice $D = (d_{ij})$ typu $m \times n$ je *obdélníková diagonální matice*, pokud $d_{ij} = 0$, kdykoliv $i \neq j$ (kde $i \in \{1, \dots, m\}$, $j \in \{1, \dots, n\}$). Obdélníkovou diagonální matici budeme zapisovat $D = \mathrm{diag}(d_{11}, \dots, d_{rr})$ nebo obšírněji

$$D = \mathrm{diag}_{m\times n}(d_{11}, \dots, d_{rr}) \ ,$$

chceme-li zvýraznit typ matice D. Budeme často vypisovat pouze nenulové prvky, tj. je-li $r < \min(m, n)$, rozumí se, že zbylé diagonální prvky jsou nulové.

Pokud $[f]^B_C = D = \mathrm{diag}_{m\times n}(d_{11}, \dots, d_{rr})$, $d_{ii} \geq 0$, pak podle důsledku 10.7 o hermitovském sdružování vzhledem k ortonormální bázím platí $[f^*]^C_B = D^* = \mathrm{diag}_{n\times m}(d_{11}, \dots, d_{rr})$. Z toho

$$[f^*f]^B_B = [f^*]^C_B [f]^B_C = \mathrm{diag}_{n\times n}(d_{11}^2, \dots, d_{rr}^2) \ .$$

Vidíme, že na diagonále matice $[f]^B_C$ musí nutně být druhé odmocniny vlastních čísel operátoru f^*f a báze B musí sestávat z vlastních vektorů tohoto operátoru. Navíc ze vztahu $[f]^B_C = D$ vidíme, že obraz i-tého vektoru báze B musí být d_{ii}-násobkem i-tého vektoru v bázi C (pro $i \leq \min(m, n)$). Tato pozorování dávají návod k důkazu věty o singulárním rozkladu.

Věta 10.37. *[o singuláním rozkladu] Nechť* $\mathbf{V}$ *a* $\mathbf{U}$ *jsou konečně generované komplexní nebo reálné vektorové prostory a* $f : \mathbf{V} \to \mathbf{U}$ *je lineární zobrazení. Pak existují ortonormální báze* B, C *prostorů* $\mathbf{V}$, $\mathbf{U}$ *takové, že* $[f]^B_C$ *je obdélníková diagonální matice.*

Důkaz. Operátor $f^*f : \mathbf{V} \to \mathbf{V}$ je podle tvrzení 10.29 pozitivně semidefinitní, takže podle spektrální věty pro hermitovské (resp. v reálném případě symetrické) operátory (věta 10.22 nebo důsledek 10.23) existuje ortonormální báze $B = (\mathbf{v}_1, \dots, \mathbf{v}_n)$ prostoru $\mathbf{V}$ složená z vlastních vektorů operátoru f^*f a $[f^*f]^B_B = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$, kde $\lambda_1, \dots, \lambda_n$ jsou nezáporná reálná vlastní čísla operátoru f^*f. Vektory v bázi B uspořádáme tak, aby $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$. Řekněme, že prvních r vektorů je nenulových.

Pro $i \in \{1, \dots, r\}$ označíme $\sigma_i = \sqrt{\lambda_i}$ a $\mathbf{u}_i = \sigma_i^{-1} f(\mathbf{v}_i)$. Pak pro libovolná $i, j \in \{1, \dots, r\}$ platí

$$\begin{aligned} \langle\, \mathbf{u}_i, \mathbf{u}_j \rangle &= \left\langle\, \sigma_i^{-1} f(\mathbf{v}_i), \sigma_j^{-1} f(\mathbf{v}_j) \right\rangle = \sigma_i^{-1}\sigma_j^{-1} \langle\, f(\mathbf{v}_i), f(\mathbf{v}_j) \rangle \\ &= \sigma_i^{-1}\sigma_j^{-1} \langle\, f^*f(\mathbf{v}_i), \mathbf{v}_j \rangle = \sigma_i^{-1}\sigma_j^{-1} \lambda_i \langle\, \mathbf{v}_i, \mathbf{v}_j \rangle \ . \end{aligned}$$

Z toho vyplývá, že pro $i \neq j$ jsou vektory $\mathbf{u}_i, \mathbf{u}_j \in U$ na sebe kolmé a navíc $\langle\, \mathbf{u}_i, \mathbf{u}_i \rangle = \sigma_i^{-2} \lambda_i = 1$, takže každý z vektorů $\mathbf{u}_1, \dots, \mathbf{u}_r$ má jednotkovou normu. Můžeme tedy tuto posloupnost doplnit na ortonormální bázi $C = (\mathbf{u}_1, \dots, \mathbf{u}_m)$ prostoru $\mathbf{U}$.

Nyní pro $i \in \{1, \dots, r\}$ je $f(\mathbf{v}_i) = \sigma_i \mathbf{u}_i$, neboli $[f(\mathbf{v}_i)]_C = \sigma_i \mathbf{e}_i$, a pro $i > k$ je $[f(\mathbf{v}_i)]_C = \mathbf{o}$. Matice lineárního zobrazení f vzhledem k bázím B a C je tedy skutečně

$$[f]^B_C = \mathrm{diag}_{m\times n}(\sigma_1, \dots, \sigma_r) \ .$$

□

Je zvykem vektory v bázích B a C uspořádat tak, že na diagonále matice $[f]_C^B$ jsou prvky uspořádány sestupně podle velikosti. Tyto prvky jsou podle pozorování nad větou rovny druhým odmocninám vlastních čísel operátoru f^*f, nazývají se *singulární hodnoty* lineárního zobrazení f. Z technických důvodů budeme singulárními hodnotami nazývat pouze nenulové prvky na diagonále.

Definice 10.38. Nechť $f : \mathbf{V} \to \mathbf{U}$ je lineární zobrazení mezi konečně generovanými komplexními nebo reálnými vektorovými prostory se skalárním součinem, B, C jsou ortonormální báze $\mathbf{V}, \mathbf{U}$ takové, že $[f]_C^B = \mathrm{diag}_{m\times n}(\sigma_1, \ldots, \sigma_r)$, kde $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$. Pak čísla $\sigma_1, \ldots, \sigma_r$ nazýváme *singulární hodnoty* lineárního zobrazení f.

Z pozorování nad větou 10.37 také vyplývá, že nenulová hodnota σ je na diagonále tolikrát, kolik je násobnost σ^2 jako vlastního čísla operátoru f^*f.

Rozmyslíme si podrobněji, co vztah $[f]_C^B = \mathrm{diag}_{m\times n}(\sigma_1, \ldots, \sigma_r)$ říká o lineárním zobrazení f. Označme $B = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$ a $C = (\mathbf{u}_1, \ldots, \mathbf{u}_m)$. Podle definice matice operátoru je $f(\mathbf{v}_1) = \sigma_1\mathbf{u}_1, \ldots, f(\mathbf{v}_r) = \sigma_r\mathbf{u}_r$. Pro zbylé vektory v bázi B je $f(\mathbf{v}_{r+1}) = \cdots = f(\mathbf{v}_n) = \mathbf{o}$. Obecněji, obraz vektoru $\mathbf{x}$ spočítáme vzorcem $[f(\mathbf{x})]_C = [f]_C^B[\mathbf{x}]_B$, tedy

(1) Vektor $\mathbf{x}$ vyjádříme v bázi B. Protože B je ortonormální, můžeme explicitně psát

$$[\mathbf{x}]_B = (x_1, \ldots, x_n)^T \text{ , kde } x_i = \langle\, \mathbf{v}_i, \mathbf{x} \rangle \quad .$$

(2) Vynásobíme zleva maticí $[f]_C^B$.

$$[f(\mathbf{x})]_C = [f]_C^B[\mathbf{x}]_B = \underbrace{(\sigma_1 x_1, \ldots, \sigma_r x_r, 0, \ldots, 0)^T}_{m \text{ složek}} \quad .$$

Složky vektoru $[\mathbf{x}]_B$ tedy vynásobíme postupně $\sigma_1, \ldots, \sigma_r$ a případně doplníme nulami na m-složkový vektor

(3) Z toho dostáváme vyjádření pro $f(\mathbf{x})$:

$$\begin{aligned} f(\mathbf{x}) &= \sigma_1 x_1 \mathbf{u}_1 + \cdots + \sigma_r x_r \mathbf{u}_r \\ &= \sigma_1 \langle\, \mathbf{v}_1, \mathbf{x} \rangle\, \mathbf{u}_1 + \cdots + \sigma_r \langle\, \mathbf{v}_r, \mathbf{x} \rangle\, \mathbf{u}_r \end{aligned}$$

OBRAZEK

Z vyjádření $[f]_C^B$ také vidíme jádro a obraz operátoru f. Vzhledem k bázi B je $[\mathrm{Ker}\, f]_B = \langle \mathbf{e}_{r+1}, \ldots, \mathbf{e}_n \rangle$, takže $\mathrm{Ker}\, f = \langle \mathbf{v}_{r+1}, \ldots, \mathbf{v}_n \rangle = \langle \mathbf{v}_1, \ldots, \mathbf{v}_r \rangle^\perp$. Pro obraz máme $[\mathrm{Im}\, f]_C = \langle \mathbf{e}_1, \ldots, \mathbf{e}_r \rangle$, takže $\mathrm{Im}\, f = \langle \mathbf{u}_1, \ldots, \mathbf{u}_r \rangle = \langle \mathbf{u}_{r+1}, \ldots, \mathbf{u}_m \rangle^\perp$. Speciálně $\dim \mathrm{Im}\, f = r$.

Příklad 10.39. Předpokládejme $\dim \mathbf{V} = 5$, $\dim \mathbf{U} = 4$, $B = (\mathbf{v}_1, \ldots, \mathbf{v}_5)$, $C = (\mathbf{u}_1, \ldots, \mathbf{u}_4)$ a $[f]_C^B = \mathrm{diag}_{4\times 5}(10,\ 9,\ 0{,}1)$.

Vektor $\mathbf{x}$ s vyjádřením $[\mathbf{x}]_B = (x_1, x_2, x_3, x_4, x_5)^T$ se zobrazí na vektor $f(\mathbf{x})$ s vyjádřením $[f(\mathbf{x})]_C = (10x_1, 9x_2, 0{,}1x_3, 0)$, čili vektor $f(\mathbf{x}) = 10x_1\mathbf{u}_1 + 9x_2\mathbf{u}_2 + 0{,}1x_3\mathbf{u}_3$. To lze interpretovat tak, že největší vliv na $f(\mathbf{x})$ mají první dvě složky x_1, x_2 odpovídající vektorům $\mathbf{v}_1, \mathbf{v}_2$. Tyto složky se přibližně zdesetinásobí, $f(\mathbf{x})$ bude „blízko“ roviny $\langle \mathbf{u}_1, \mathbf{u}_2 \rangle$ a bude mít přibližně desetkrát větší normu (pokud není třetí složka příliš velká). Vliv třetí složky x_3 je malý a další složky nemají na výsledek žádný vliv.

Jádrem f je prostor $\mathrm{Ker}\, f = \langle \mathbf{v}_4, \mathbf{v}_5 \rangle = \langle \mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3 \rangle^\perp$ a obrazem f je prostor $\langle \mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3 \rangle = \mathbf{u}_4^\perp$.

Jaká je souvislost unitární diagonalizace operátoru a singulárního rozkladu? Uvažujme normální operátor f na konečně generovaném prostoru $\mathbf{V}$ a ortonormální bázi $B = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$ prostoru $\mathbf{V}$ takovou, že $[f]_B^B = D = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$. Vektory v B si uspořádáme tak, aby $\lambda_1, \ldots, \lambda_r \neq 0$ a $\lambda_{r+1} = \cdots = \lambda_n = 0$. Singulární hodnoty zobrazení f jsou druhé odmocniny nenulových vlastních čísel operátoru f^*f, které můžeme spočítat jako (nenulová) vlastní čísla matice

$$[f^*f]_B^B = [f^*]_B^B[f]_B^B = D^*D = \mathrm{diag}(|\lambda_1|^2, \ldots, |\lambda_n|^2) \quad .$$

Proto platí:

Pozorování 10.40. *Singulární hodnoty normálního operátoru jsou rovny absolutním hodnotám jeho nenulových vlastních čísel.*

Z tvaru $[f]_B^B = D$ navíc můžeme snadno získat singulární rozklad. Pro $i \leq r$ položíme $\mathbf{u}_i = (\lambda_i/|\lambda_i|)\mathbf{v}_i$. Posloupnost $(\mathbf{u}_1, \dots, \mathbf{u}_r)$ je ortonormální, protože vznikla vynásobením vektorů $\mathbf{v}_1, \dots, \mathbf{v}_r$ komplexními čísly s absolutní hodnotou 1. Tyto vektory doplníme na ortonormální bázi $(\mathbf{u}_1, \dots, \mathbf{u}_n)$. Protože pro $i \leq r$ je $f(\mathbf{v}_i) = \lambda_i \mathbf{v}_i = |\lambda_i|\mathbf{u}_i$ a pro $i > r$ je $f(\mathbf{v}_i) = \mathbf{o} = \lambda_i \mathbf{u}_i$, platí

$$[f]_C^B = \mathrm{diag}(|\lambda_1|, \dots, |\lambda_n|) \ .$$

Všimněte si ještě, že pro pozitivně definitní operátory unitární diagonalizace a singulární rozklad splývají.

10.2.1. *Singulární rozklad matice.* Následující věta je maticovou verzí věty 10.37. Poskytne nám také další geometrickou interpretaci.

Věta 10.41 (10.37*.)**.** *[singulární rozklad matice] Nechť A je komplexní (resp. reálná) matice typu $m \times n$. Pak existují unitární (resp. ortogonální) matice U, V řádů m, n a obdélníková diagonální matice D typu $m \times n$ takové, že*

$$A = UDV^{-1} = UDV^* \ .$$

Důkaz. Důkaz budeme formulovat pro komplexní případ. Použijeme větu 10.37 na lineární zobrazení $f_A : \mathbb{C}^n \to \mathbb{C}^m$ mezi aritmetickými prostory se skalárním součinem. Existuje ortonormální báze B prostoru $\mathbb{C}^n$ a ortonormální báze C prostoru $\mathbb{C}^m$ tak, že $[f_A]_C^B = D$ je obdélníková diagonální matice. Označme U matici přechodu od C ke kanonické bázi prostoru $\mathbb{C}^m$ a V matici přechodu od B ke kanonické bázi prostoru $\mathbb{C}^n$. Matice U, V jsou unitární, takže $U^{-1} = U^*$ (a $V^{-1} = V^*$). Pak

$$[f_A]_K^K = [\mathrm{id}]_K^C [f_A]_C^B [\mathrm{id}]_B^K = [\mathrm{id}]_K^C [f_A]_C^B ([\mathrm{id}]_K^B)^{-1} = UDV^{-1} = UDV^* \ .$$

□

Rozklad $A = UDV^T = UDV^{-1}$ můžeme geometricky interpretovat jako $f_A = f_U f_D f_{V^{-1}}$. V případě reálné čtvercové matice řádu n je tedy $f_A : \mathbb{R}^n \to \mathbb{R}^n$ složením pořadě ortogonálního zobrazení $f_{V^{-1}}$, zobrazení f_D, které natahuje nebo zkracuje souřadnicové osy a ortogonálního zobrazení f_U.

Pro zobrazení $f_A = f_U f_D f_{V^T} : \mathbb{R}^n \to \mathbb{R}^m$ sledujme postupně obraz n-dimenzionální koule $O = \{\mathbf{x} : \|\mathbf{x}\| \leq 1\}$. Zobrazení f_{V^T} je ortogonální, proto zobrazí O opět na O. Zobrazení f_D pak sféru O natáhne nebo smrští ve směru souřadnicových os (případně ještě ubere nebo přidá složky pokud $m \neq n$), tím vznikne tzv. *zobecněný elipsoid* s poloosami $\sigma_1\mathbf{e}_1, \sigma_2\mathbf{e}_2, \dots$. Nakonec se na vzniklou množinu aplikuje zobrazení f_U, které vektor $\mathbf{e}_i$ zobrazí na vektor $\mathbf{u}_i$. Tím vznikne zobecněný elipsoid s polosami $\sigma_1\mathbf{u}_1, \sigma_2\mathbf{u}_2, \dots$ velikostí $\sigma_1, \sigma_2, \dots$.

Příklad 10.42. Spočítáme singulární rozklad zkosení $f_A : \mathbb{R}^2 \to \mathbb{R}^2$ diskutovaného v příkladu 10.36 daného maticí

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

Postupujeme podle důkazu věty 10.37. Najdeme ortonormální bázi $B = (\mathbf{v}_1, \mathbf{v}_2)$ prostoru $\mathbb{R}^2$ se standardním skalárním součinem takovou, že matice operátoru $(f_A)^* f_A = f_{A^T A}$ vzhledem k B je diagonální.

Vlastní čísla matice

$$A^T A = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$$

jsou $\lambda_{1,2} = (3 \pm \sqrt{5})/2$, singulární hodnoty jsou proto

$$\sigma_{1,2} = \sqrt{\frac{3 \pm \sqrt{5}}{2}}, \quad \sigma_1 \approx 1{,}618, \quad \sigma_2 \approx 0{,}618 \ .$$

Příslušné prostory vlastních vektorů matice $A^T A$ jsou jednodimenzionální, vybereme v nich vektory jednotkové velikosti. Vyjde přibližně

$$\mathbf{v}_1 \approx \begin{pmatrix} 0{,}526 \\ 0{,}851 \end{pmatrix} \in M_{\lambda_1}, \quad \mathbf{v}_2 \approx \begin{pmatrix} -0{,}851 \\ 0{,}526 \end{pmatrix} \in M_{\lambda_2} \ .$$

Vektory $\mathbf{u}_1, \mathbf{u}_2$ vypočteme ze vzorce $\mathbf{u}_i = \sigma_i^{-1} A\mathbf{v}_i$.

$$\mathbf{u}_1 \approx \begin{pmatrix} 0{,}851 \\ 0{,}526 \end{pmatrix}, \quad \mathbf{u}_2 \approx \begin{pmatrix} -0{,}526 \\ 0{,}851 \end{pmatrix} \ .$$

Vzhledem k bázím $B = (\mathbf{v}_1, \mathbf{v}_2)$ a $C = (\mathbf{u}_1, \mathbf{u}_2)$ má f_A matici

$$D = [f_A]^B_C \approx \begin{pmatrix} 1{,}618 & 0 \\ 0 & 0{,}618 \end{pmatrix} \ .$$

Příklad 10.43. Spočítáme singulární rozklad pro reálnou čtvercovou matici

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

Tato matice je normální, v příkladu 10.20 jsme ji unitárně diagonalizovali. Pokud by nás zajímali pouze singulární hodnoty, můžeme je spočítat jako absolutní hodnoty vlastních čísel $2, (1+\sqrt{3}i)/2, (1-\sqrt{3}i)/2$, tj. $\sigma_1 = 2$ (násobnost 1), $\sigma_2 = \sigma_3 = 1$ (násobnost 2).

Z unitární diagonalizace lze také určit singulární rozklad, budeme ale postupovat podle důkazu věty 10.37. Najdeme ortonormální bázi $B = (\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ prostoru $\mathbb{R}^3$ se standardním skalárním součinem takovou, že matice operátoru $(f_A)^* f_A = f_{A^T A}$ vzhledem k B je diagonální. Matice

$$A^T A = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$$

má vlastní číslo $\lambda_1 = 4$ násobnosti 1 a vlastní číslo $\lambda_2 = \lambda_3 = 1$ násobnosti 2. Singulární hodnoty matice A jsou $\sigma_1 = 2$ a $\sigma_2 = \sigma_3 = 1$. Příslušné prostory vlastních vektorů jsou

$$M_4 = \left\langle \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\rangle \ , \quad M_1 = \left\langle \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \right\rangle,$$

V nich najdeme ortonormální báze. V prostoru M_1 je ortonormální báze třeba

$$(\mathbf{v}_2, \mathbf{v}_3) = \left(\frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix} \right) \ ,$$

v prostoru M_4

$$\mathbf{v}_1 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \ .$$

Vektory báze $C = (\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3)$ vypočteme ze vztahu $\mathbf{u}_i = \sigma_i^{-1} f_A(\mathbf{v}_i)$.

$$\mathbf{u}_1 = 2^{-1} A\mathbf{v}_3 = 2 \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix} \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \ .$$

$$\mathbf{u}_2 = A\mathbf{v}_2 = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}, \quad \mathbf{u}_3 = A\mathbf{v}_3 = \frac{1}{\sqrt{6}}\begin{pmatrix} 2 \\ -1 \\ -1 \end{pmatrix},$$

Vzhledem k ortonormálním bázím B a C je

$$[f_A]^B_C = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} .$$

To nám dává singulární rozklad matice A.

$$A = [f_A]^K_K = [\mathrm{id}]^C_K [f_A]^B_C [\mathrm{id}]^K_B = \begin{pmatrix} \frac{1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \end{pmatrix} \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & -\frac{2}{\sqrt{6}} \end{pmatrix}$$

Obrazem jednotkové koule při zobrazení f_A je elipsoid s osami $\langle \mathbf{u}_1 \rangle, \langle \mathbf{u}_2 \rangle, \langle \mathbf{u}_3 \rangle$ a velikostí poloos $2, 1, 1$. Protože velikosti druhé a třetí poloosy jsou stejné, je tento elipsoid rotačně symetrický podle osy $\langle \mathbf{u}_1 \rangle$.

OBRAZEK

Příklad 10.44. Spočítáme singulární rozklad pro reálnou matici

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

Vlastní čísla matice

$$A^T A = \begin{pmatrix} 6 & 12 \\ 12 & 24 \end{pmatrix}$$

jsou 30 a 0, singulární hodnota matice A je tedy $\sigma_1 = \sqrt{30}$. Příslušné normované vlastní vektory jsou

$$\mathbf{v}_1 = \frac{1}{\sqrt{5}}\begin{pmatrix} 1 \\ 2 \end{pmatrix}, \quad \mathbf{v}_2 = \frac{1}{\sqrt{5}}\begin{pmatrix} -2 \\ 1 \end{pmatrix} .$$

Vektor

$$\mathbf{u}_1 = \sigma_1^{-1} A\mathbf{v}_1 = \frac{1}{\sqrt{6}}\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$$

doplníme do ortonormální báze $\mathbb{R}^3$ například vektory

$$\mathbf{u}_2 = \frac{1}{\sqrt{2}}\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, \quad \mathbf{u}_3 = \frac{1}{\sqrt{3}}\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} .$$

Vzhledem k bázím $B = (\mathbf{v}_1, \mathbf{v}_2)$ a $C = (\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3)$ máme

$$[f_A]^B_C = \begin{pmatrix} \sqrt{30} & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} ,$$

což nám dává rozklad

$$A = \begin{pmatrix} \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ \frac{2}{\sqrt{6}} & 0 & -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \end{pmatrix} \begin{pmatrix} \sqrt{30} & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \\ -\frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{pmatrix}$$

Obrazem jednotkového kruhu při zobrazení f_A je „elipsoid“ s osami $\langle \mathbf{u}_1 \rangle, \langle \mathbf{u}_2 \rangle, \langle \mathbf{u}_3 \rangle$ a poloosami velikostí $\sqrt{30}, 0, 0$, takže ve skutečnosti úsečka spojující $-\sqrt{30}\mathbf{u}_1$ a $\sqrt{30}\mathbf{u}_1$.

OBRAZEK

Uvažujme singulární rozklad $A = UDV^*$ matice typu $m \times n$ hodnosti r, kde $U = (\mathbf{u}_1|\dots|\mathbf{u}_m)$, $D = \mathrm{diag}_{m\times n}(\sigma_1, \dots, \sigma_r)$ a $V = (\mathbf{v}_1|\dots|\mathbf{v}_n)$. Formulku $A = UDV^*$ můžeme také psát

$$A = \sigma_1 \mathbf{u}_1 \mathbf{v}_1^* + \cdots + \sigma_r \mathbf{u}_r \mathbf{v}_r^* \enspace ,$$

čímž vyjadřujeme matici A jako součet matic hodnosti 1.

Úspornější forma, tzv. *kompaktní singulární rozklad*, je $A = U'D'(V')^*$, kde $U' = (\mathbf{u}_1|\dots|\mathbf{u}_r)$ je typu $m \times r$, $D' = \mathrm{diag}(\sigma_1, \dots, \sigma_r)$ je čtvercová matice řádu r a $V' = (\mathbf{v}_1|\dots|\mathbf{v}_r)$ je typu $n \times r$ (čili $(V')^*$ je typu $r \times n$). Například pro matici z příkladu 10.44 můžeme psát

$$\begin{pmatrix} 1 & 2 \\ 2 & 4 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{6}} \\ \frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \end{pmatrix} \begin{pmatrix} \sqrt{30} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{pmatrix}$$

Za zmínku v souvislosti se singulárním rozkladem stojí tzv. *polární rozklad* $A = RW$ čtvercové matice A, kde R je pozitivně semidefinitní matice a W je unitární. Ze singulárního rozkladu $A = UDV^T$ jej dostaneme úpravou $A = (UDU^T)(UV^T)$. Matice UDU^T je pozitivně semidefinitní a matice UV^T je unitární, takže můžeme položit $R = UDU^T$ a $W = UV^T$. Polární rozklad lze chápat jako zobecnění rozkladu komplexního čísla na součin nezáporného reálného čísla a komplexního čísla jednotkové velikosti.

10.2.2. *Spektrální norma.* Singulární rozklad lineárního zobrazení f nám umožňuje odpovědět na otázku, jaký nejvýše (nejméně) může být podíl $\|f(\mathbf{x})\| \,/\, \|\mathbf{x}\|$ pro $\mathbf{x} \neq \mathbf{o}$. Jinými slovy, jak nejvíc se může změnit délka vektoru při zobrazení f. Pro které vektory se tohoto maxima (minima) nabývá?

Nejprve si všimneme, že

$$\frac{\|f(\mathbf{x})\|}{\|\mathbf{x}\|} = \left\| f\left(\frac{\mathbf{x}}{\|\mathbf{x}\|}\right) \right\| \enspace ,$$

takže se stačí zabývat otázkou, jaká je největší, nebo nejmenší hodnota $\|f(\mathbf{x})\|$ pro vektory $\mathbf{x}$ jednotkové velikosti (tj. pro vektory na jednotkové sféře). Geometricky je v případě reálných matic odpověď patrná z diskuze o obrazu jednotkové koule. Ukážeme algebraické odvození v obecném případě.

Nechť B, C jsou ortonormální báze takové, že $[f]^B_C = \mathrm{diag}_{m\times n}(\sigma_1, \dots, \sigma_r)$, $\sigma_1 \geq \cdots \geq \sigma_r > 0$. Označme $[\mathbf{x}]_B = (x_1, \dots, x_n)$. Protože $\|\mathbf{x}\| = 1$ a B je ortonormální, je $\|[\mathbf{x}]_B\| = 1$, čili

$$|x_1|^2 + |x_2|^2 + \cdots + |x_n|^2 = 1 \enspace .$$

Norma vektoru $f(\mathbf{x})$ je potom

$$\|f(\mathbf{x})\| = \|[f(\mathbf{x})]_C\| = \left\| (\sigma_1 x_1, \dots, \sigma_r x_r, 0, \dots, 0)^T \right\| = \sqrt{\sigma_1^2 |x_1|^2 + \cdots + \sigma_r^2 |x_r|^2} \enspace .$$

Protože $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$ je tento výraz menší nebo roven

$$\sqrt{\sigma_1^2 (|x_1|^2 + \cdots + |x_n|^2)} = \sigma_1 \enspace ,$$

přičemž rovnost nastává právě tehdy, když $|x_i| = 0$ pro každé i takové, že $\sigma_i < \sigma_1$, neboli právě pro vektory $\mathbf{x}$ v lineárním obalu vektorů báze B příslušných singulární hodnotě σ_1, neboli právě pro vlastní vektory operátoru $f^* f$ příslušné vlastnímu číslu σ_1^2. Odvodili jsme následující tvrzení

Tvrzení 10.45. *Nechť $f : \mathbf{V} \to \mathbf{W}$ je lineární zobrazení mezi reálnými nebo komplexními vektorovými prostory se skalárním součinem. Pak pro libovolný vektor $\mathbf{o} \neq \mathbf{x} \in V$ platí*

$$\frac{\|f(\mathbf{x})\|}{\|\mathbf{x}\|} \leq \sigma_1 \enspace ,$$

kde σ_1 je největší singulární hodnota operátoru f. Rovnost nastává právě tehdy, když $\mathbf{x}$ je nenulový vlastní vektor operátoru $f^* f$ příslušný vlastnímu číslu σ_1^2.

Podobné tvrzení samozřejmě můžeme formulovat pro matice a výraz $\|A\mathbf{x}\| / \|\mathbf{x}\|$. Maximu výrazu $\|f(\mathbf{x})\| / \|\mathbf{x}\|$ (resp. $\|A\mathbf{x}\| / \|\mathbf{x}\|$) se také říká *spektrální norma* operátoru f (resp. matice A). Je rovná největší singulární hodnotě. Budeme ji značit $\|f\|$ (resp. $\|A\|$). Podle definice je

$$\|f(\mathbf{x})\| \leq \|f\| \, \|\mathbf{x}\| , \quad \|A\mathbf{x}\| \leq \|A\| \, \|\mathbf{x}\| \ .$$

Obdobné tvrzení se odvodí pro minima:

Tvrzení 10.46. *Nechť $f : \mathbf{V} \to \mathbf{W}$ je lineární zobrazení mezi reálnými nebo komplexními vektorovými prostory se skalárním součinem,* $\dim V = n$. *Pak pro libovolný vektor $\mathbf{x} \in V$ takový, že $f(\mathbf{x}) \neq \mathbf{o}$ platí*

$$\frac{\|f(\mathbf{x})\|}{\|\mathbf{x}\|} \geq \sigma_r \ ,$$

kde σ_r je nejmenší singulární hodnota operátoru f. Rovnost nastává právě tehdy, když $\mathbf{x}$ je nenulový vlastní vektor operátoru $f^* f$ příslušný vlastnímu číslu σ_r^2.

Příklad 10.47. Matice

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

z příkladu 10.43 má spektrální normu $\|A\| = 2$, tj. pro libovolný vektor $\mathbf{x} \in \mathbb{C}^3$ platí

$$\frac{\|A\mathbf{x}\|}{\|\mathbf{x}\|} \leq 2 \ .$$

Rovnost nastává právě tehdy, když $\mathbf{o} \neq \mathbf{x} \in M_4 = \langle \mathbf{v}_1 \rangle = \left\langle (1,1,1)^T \right\rangle$ (značení přebíráme z příkladu 10.43).

Pro libovolný vektor $\mathbf{x} \in C^3$ platí

$$\frac{\|A\mathbf{x}\|}{\|\mathbf{x}\|} \geq 1 \ ,$$

přičemž rovnost nastává právě tehdy, když $\mathbf{o} \neq \mathbf{x} \in M_1 = \langle \mathbf{v}_2, \mathbf{v}_3 \rangle = \left\langle (-1,1,0)^T, (-1,0,1)^T \right\rangle$.

Příklad 10.48. Pro matici

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

z příkladu 10.44 platí

$$0 \leq \frac{\|A\mathbf{x}\|}{\|\mathbf{x}\|} \leq \sqrt{30} \ .$$

(První nerovnost je triviální.) Rovnost v první nerovnosti nastává pro $\mathbf{o} \neq \mathbf{x} \in \operatorname{Ker} A = \langle \mathbf{v}_2 \rangle = \left\langle (-2,1)^T \right\rangle$, v druhé nerovnosti pro $\mathbf{o} \neq \mathbf{x} \in \langle \mathbf{v}_1 \rangle = \left\langle (1,2)^T \right\rangle$. Spektrální norma matice A je $\|A\| = \sqrt{30}$.

10.2.3. *Numerická stabilita řešení soustavy lineárních rovnic s regulární maticí.* Uvažujme soustavu $A\mathbf{x} = \mathbf{b}$, kde A je reálná regulární matice, jejíž řešení je, jak víme, $\mathbf{x} = A^{-1}\mathbf{b}$.

Řekněme, že vektor $\mathbf{b}$ získáme měřením, které je zatíženo chybou $\delta\mathbf{b}$ (výraz $\delta\mathbf{b}$ chápejte jako označení vektoru, nikoliv jako součin). Ve skutečnosti tedy neznámé hodnoty $\mathbf{x}$ budou zatížené chybou $\delta\mathbf{x}$, kde

$$A(\mathbf{x} + \delta\mathbf{x}) = \mathbf{b} + \delta\mathbf{b} \ , \text{ tj. } \delta\mathbf{x} = A^{-1}\delta\mathbf{b} \ .$$

Velikost chyby bude

$$\|\delta\mathbf{x}\| = \left\|A^{-1}\delta\mathbf{b}\right\| \leq \left\|A^{-1}\right\| \, \|\delta\mathbf{b}\| \ .$$

(Přičemž rovnost může nastat.) Pokud je spektrální norma $\left\|A^{-1}\right\|$ vysoká, např. 10^6, velikost chyby neznámých hodnot může být až 10^6-krát větší než velikost chyby naměřených hodnot. To je neuspokojivé a je nejspíše potřeba změnit model. Všimněte si, že norma $\left\|A^{-1}\right\|$ je rovná převrácené hodnotě nejmenší singulární hodnoty matice A (cvičení).

V praxi nás spíše bude zajímat odhad na velikost relativní chyby $\|\delta\mathbf{x}\| \,/\, \|\mathbf{x}\|$ neznámých hodnot v závislosti na velikosti relativní chyby $\|\delta\mathbf{b}\| \,/\, \|\mathbf{b}\|$ měření. K tomu si všimneme

$$\|\mathbf{b}\| = \|A\mathbf{x}\| \leq \|A\| \, \|\mathbf{x}\| \quad ,$$

takže

$$\frac{\|\delta\mathbf{x}\|}{\|\mathbf{x}\|} \leq \left\|A^{-1}\right\| \|\delta\mathbf{b}\| \frac{1}{\|\mathbf{x}\|} \leq \left\|A^{-1}\right\| \|\delta\mathbf{b}\| \frac{\|A\|}{\|\mathbf{b}\|} = \|A\| \left\|A^{-1}\right\| \frac{\|\delta\mathbf{b}\|}{\|\mathbf{b}\|} \quad .$$

Číslu $\|A\| \left\|A^{-1}\right\|$ se říká *číslo podmíněnosti matice* A, je rovno podílu největší a nejmenší singulární hodnoty. Relativní chybu řešení lze tedy odhadnout relativní chybou měření krát číslo podmíněnosti.

Příklad 10.49. Číslo podmíněnosti matice

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

z příkladu 10.43 je $2/1 = 2$. Relativní chyba řešení soustavy $A\mathbf{x} = \mathbf{b}$ bude tedy nejvýše dvakrát větší než relativní chyba měření pravé strany.

10.2.4. *Aproximace maticí nižší hodnosti.* Uvažujme lineární zobrazení $f : \mathbb{R}^n \to \mathbb{R}^m$ hodnosti r, tj. $r = \dim \operatorname{Im} f$. Chceme najít lineární zobrazení $\hat{f} : \mathbb{R}^n \to \mathbb{R}^m$ dané hodnosti $s < r$, které co nejlépe aproximuje f ve smyslu, že spektrální norma $\left\|f - \hat{f}\right\|$ je co nejmenší. To se nám může hodit při komprimaci dat nebo při zjednodušování matematických modelů.

Nechť $B = (\mathbf{v}_1, \dots, \mathbf{v}_n)$, C jsou ortonormální báze $\mathbb{R}^n$, $\mathbb{R}^m$ takové, že $[f]^B_C = \operatorname{diag}_{m\times n}(\sigma_1, \dots, \sigma_r)$, kde $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$. Ukážeme, že hledaná nejlepší aproximace $\hat{f}$ lineárního zobrazení f je určená vztahem

$$[\hat{f}]^B_C = \operatorname{diag}_{m\times n}(\sigma_1, \dots, \sigma_s) \quad .$$

Při této volbě je

$$[f - \hat{f}]^B_C = \operatorname{diag}_{m\times n}(0, \dots, 0, \sigma_{s+1}, \dots, \sigma_r)$$

největší singulární číslo lineárního zobrazení $f - \hat{f}$ je σ_{s+1}, takže

$$\left\|f - \hat{f}\right\| = \sigma_{s+1} \quad .$$

Zbývá ukázat, že lepší normy nelze dosáhnout. Předpokládejme, že $g : \mathbb{R}^n \to \mathbb{R}^m$ je lineární zobrazení hodnosti nejvýše s. Protože $\dim \langle \mathbf{v}_1, \dots, \mathbf{v}_{s+1} \rangle = s+1$ a $\dim \operatorname{Ker} g = n - \dim \operatorname{Im} g \geq n-s$, plyne z věty o dimenzi součtu a průniku, že se tyto dva prostory protínají. Uvažujme libovolný nenulový vektor $\mathbf{x}$ v jejich průniku a označme $[\mathbf{x}]_B = (x_1, \dots, x_n)$, tj. $g(\mathbf{x}) = \mathbf{o}$ a $x_{s+2} = \cdots = x_n = 0$. Pak

$$\begin{aligned} \|f - g\| &\geq \frac{\|(f-g)(\mathbf{x})\|}{\|\mathbf{x}\|} = \frac{\|f(\mathbf{x})\|}{\|\mathbf{x}\|} = \frac{\|[f(\mathbf{x})]_C\|}{\|[\mathbf{x}]_B\|} = \frac{\left\|(\sigma_1 x_1, \dots, \sigma_{s+1} x_{s+1}, 0, \dots, 0)^T\right\|}{\|(x_1, \dots, x_{s+1}, 0, \dots, 0)^T\|} \\ &\geq \frac{\left\|(\sigma_{s+1} x_1, \dots, \sigma_{s+1} x_{s+1}, 0, \dots, 0)^T\right\|}{\|(x_1, \dots, x_{s+1}, 0, \dots, 0)^T\|} = \sigma_{s+1} \quad , \end{aligned}$$

Tedy norma je skutečně alespoň σ_{s+1}.

V maticovém pohledu můžeme výsledek formulovat následujícím způsobem. Je-li singulární rozklad matice A roven

$$A = UDV^T = U \operatorname{diag}_{m\times n}(\sigma_1, \dots, \sigma_r) V^T = \sigma_1 \mathbf{u}_1 \mathbf{v}_1^* + \cdots + \sigma_r \mathbf{u}_r \mathbf{v}_r^* \quad ,$$

kde $\sigma_1 \geq \cdots \geq \sigma_r$, pak nejlepší aproximace $\hat{A}$ matice A maticí hodnosti s je

$$\hat{A} = U \operatorname{diag}_{m\times n}(\sigma_1, \dots, \sigma_s) V^T = \sigma_1 \mathbf{u}_1 \mathbf{v}_1^* + \cdots + \sigma_s \mathbf{u}_s \mathbf{v}_s^* \quad .$$

K uložení matice A typu $m \times n$ v počítači potřebujeme mn skalárů. K uložení aproximace $\hat{A}$ stačí $s(m + n + 1)$ skalárů (protože máme s sčítanců a každý sčítanec obsahuje skalár σ_i, m-složkový vektor $\mathbf{u}_i$ a n-složkový vektor $\mathbf{v}_i$). Toho lze využít pro komprimaci dat.

Příklad 10.50. Nejlepší aproximace matice

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

z příkladu 10.43 maticí hodnosti 1 je

$$\hat{A} = \sigma_1 \mathbf{u}_1 \mathbf{v}_1^* = 2\ \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}\ \frac{1}{\sqrt{3}}(1\ 1\ 1) = \frac{2}{3} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} .$$

Aproximace je nejlepší v tom smyslu, že $\left\| A - \hat{A} \right\| = \sigma_2 = 1$ a pro žádnou matici B hodnosti 1 neplatí $\|A - B\| < 1$.

Aproximovat A maticí hodnosti 2 se nevyplatí, protože norma rozdílu $A - \hat{A}$ by byla také rovna $\sigma_3 = 1$, takže bychom v tomto smyslu nedosáhli žádného zlepšení.

10.2.5. *Pseudoinverze.* Uvažujme soustavu rovnic $A\mathbf{x} = \mathbf{b}$, kde A je reálná matice typu $m \times n$. Soustava nemusí mít žádné řešení. V tom případě víme, že aproximace řešení metodou nejmenších čtverců jsou právě všechna řešení soustavy rovnic

$$A^T A\mathbf{x} = A^T\mathbf{b} .$$

Tato soustava může mít řešení více, najdeme takové, pro které je norma $\|\mathbf{x}\|$ nejmenší.

Uvažme nejprve speciální případ, kdy $A = D = \mathrm{diag}_{m\times n}(\sigma_1, \ldots, \sigma_r)$, kde $\sigma_1, \ldots, \sigma_r$ jsou nenulová reálná čísla. Chceme najít řešení soustavy

$$D^T D\mathbf{x} = D^T\mathbf{b} ,$$

s nejmenší normou. Označme $\mathbf{x} = (x_1, \ldots, x_n)$ a $\mathbf{b} = (b_1, \ldots, b_m)$. Pak

$$D^T D\mathbf{x} = \mathrm{diag}_{n\times n}(\sigma_1^2, \ldots, \sigma_r^2)(x_1, \ldots, x_n)^T = (\sigma_1^2 x_1, \ldots, \sigma_r^2 x_r, 0, \ldots, 0)^T$$

a

$$D^T\mathbf{b} = \mathrm{diag}_{n\times m}(\sigma_1, \ldots, \sigma_r)(b_1, \ldots, b_m)^T = (\sigma_1 b_1, \ldots, \sigma_r b_r, 0, \ldots, 0)^T .$$

Řešení $\mathbf{x}$ s nejmenší normou bude tedy

$$(x_1, \ldots, x_n)^T = (b_1\sigma_1^{-1}, \ldots, b_r\sigma_r^{-1}, 0, \ldots, 0)^T .$$

Označíme-li

$$D^\dagger = \mathrm{diag}_{n\times m}(\sigma_1^{-1}, \ldots, \sigma_r^{-1}) ,$$

můžeme vztah maticově zapsat

$$\mathbf{x} = D^\dagger\mathbf{b} .$$

Pomocí singulárního rozkladu teď tento výsledek zobecníme na obecnou matici $A = UDV^T$. Hledáme $\mathbf{x}$ s nejmenší normou, aby

$$\begin{aligned} A^T A\mathbf{x} &= A^T\mathbf{b} \\ VD^T U^T\ UDV^T\mathbf{x} &= VD^T U^T\mathbf{b} \\ VD^T DV^T\mathbf{x} &= VD^T U^T\mathbf{b} \\ D^T DV^T\mathbf{x} &= D^T U^T\mathbf{b} . \end{aligned}$$

Protože $\left\|V^T\mathbf{x}\right\| = \|\mathbf{x}\|$ (matice V^T je ortogonální), pro hledané $\mathbf{x}$ podle předchozího výsledku platí

$$V^T\mathbf{x} = D^\dagger U^T\mathbf{b}$$

a vynásobením maticí V zleva získáme

$$\mathbf{x} = VD^\dagger U^T\mathbf{b}$$

Matice $A^\dagger = VD^\dagger U^T$ je tzv. *Moore-Penroseova pseudoinverze* matice A. Při použití zápisu

$$A = \sigma_1 \mathbf{u}_1 \mathbf{v}_1^* + \cdots + \sigma_r \mathbf{u}_r \mathbf{v}_r^*$$

můžeme psát

$$A^\dagger = \sigma_1^{-1} \mathbf{v}_1 \mathbf{u}_1^* + \cdots + \sigma_r^{-1} \mathbf{v}_r \mathbf{u}_r^* \ .$$

Ukázali jsme, že pro soustavu $A\mathbf{x} = \mathbf{b}$ je vektor $\mathbf{x} = A^\dagger \mathbf{b}$ nejkratším vektorem, který je zároveň aproximací řešení metodou nejmenších čtverců. Speciálně, pokud $A\mathbf{x} = \mathbf{b}$ má právě jedno řešení, pak je tímto řešením vektor $\mathbf{x} = A^\dagger \mathbf{b}$ (tj. pro regulární matice A je $A^\dagger = A^{-1}$). Má-li soustava $A\mathbf{x} = \mathbf{b}$ více řešení, pak je $\mathbf{x} = A^\dagger \mathbf{b}$ řešením s nejmenší normou.

Příklad 10.51. Uvažujme soustavu

$$A\mathbf{x} = \mathbf{b}, \quad A = \begin{pmatrix} 1 & 2 \\ 2 & 4 \\ 1 & 2 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 2 \\ 2 \\ 0 \end{pmatrix} \ .$$

Aproximace soustavy metodou nejmenších čtverců jsou řešení soustavy $A^T A\mathbf{x} = A^T \mathbf{b}$:

$$\begin{pmatrix} 6 & 12 \\ 12 & 24 \end{pmatrix} \mathbf{x} = \begin{pmatrix} 6 \\ 12 \end{pmatrix}$$

$$\mathbf{x} \in \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \left\langle \begin{pmatrix} -2 \\ 1 \end{pmatrix} \right\rangle$$

OBRAZEK

Z obrázku je vidět, že aproximace s nejmenší normou má směr $\left\langle (1,2)^T \right\rangle$ a snadno vypočteme $\mathbf{x} = (1/5, 2/5)^T$.

Toto řešení můžeme vypočítat pomocí pseudoinverze. V příkladu 10.44 jsme nalezli singulární rozklad

$$A = \sigma_1 \mathbf{u}_1 \mathbf{v}_1^* = \sqrt{30} \begin{pmatrix} \frac{1}{\sqrt{6}} \\ \frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{pmatrix} \ .$$

Pseudoinverze je

$$A^\dagger = \sigma_1^{-1} \mathbf{v}_1 \mathbf{u}_1^* = \frac{1}{\sqrt{30}} \begin{pmatrix} \frac{1}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \end{pmatrix} = \frac{1}{30} \begin{pmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \end{pmatrix} \ ,$$

takže hledaná aproximace je

$$\mathbf{x} = A^\dagger \mathbf{b} = \frac{1}{30} \begin{pmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \end{pmatrix} \begin{pmatrix} 2 \\ 2 \\ 0 \end{pmatrix} = \frac{1}{5} \begin{pmatrix} 1 \\ 2 \end{pmatrix} \ .$$

Shrnutí desáté kapitoly

(1) Je-li $\mathbf{V}$ konečně generovaný lineární prostor nad $\mathbb{C}$ (resp. $\mathbb{R}$) se skalárním součinem $\langle\, , \rangle$ a f lineární operátor na $\mathbf{V}$, pak říkáme, že f je *unitárně diagonalizovatelný* (resp. *ortogonálně diagonalizovatelný*), pokud existuje ortonormální báze B prostoru $\mathbf{V}$ taková, že $[f]_B^B$ je diagonální.

(2) Je-li $f : \mathbf{V} \to \mathbf{V}$ lineární operátor na konečně generovaném lineárním prostoru $\mathbf{V}$ dimenze n se skalárním součinem nad tělesem $\mathbb{C}$ (resp. $\mathbb{R}$), pak jsou následující tvrzení ekvivalentní.
   (a) Operátor f je unitárně diagonalizovatelný (resp. ortogonálně diagonalizovatelný).
   (b) Operátor f
      - má n vlastních čísel včetně algebraických násobností,
      - geometrická násobnost každého vlastního čísla operátoru f se rovná jeho algebraické násobnosti a
      - pro libovolná dvě různá vlastní čísla λ_i, λ_j operátoru f platí $\mathbf{M}_{\lambda_i} \perp M_{\lambda_j}$.

(3) Je-li A komplexní (nebo reálná) matice řádu n, pak $\lambda \in \mathbb{C}$ (nebo $\lambda \in \mathbb{R}$) je vlastní číslo matice A právě když $\bar{\lambda}$ je vlastní číslo matice A^* (nebo A^T).

(4) Komplexní (nebo reálná) čtvercová matice A se nazývá *normální*, pokud $A^*A = AA^*$ (pro reálné matice můžeme psát $A^TA = AA^T$).

(5) Je-li A normální komplexní (nebo reálná) matice a $\lambda \in \mathbb{C}$ (nebo $\lambda \in \mathbb{R}$), pak matice $A - \lambda I_n$ je také normální.

(6) Je-li A normální komplexní (nebo reálná) matice řádu n, pak pro každý vektor $\mathbf{v} \in \mathbb{C}^n$ (nebo $\mathbf{v} \in \mathbb{R}^n$) platí

$$\|A\mathbf{v}\| = \|A^*\mathbf{v}\| \ ,$$

kde $\|\cdot\|$ je norma určená standardním skalárním součinem na $\mathbb{C}^n$ (resp. $\mathbb{R}^n$).

(7) Je-li A komplexní (nebo reálná) matice řádu n, $\lambda \in \mathbb{C}$ (nebo $\lambda \in \mathbb{R}$) a $\mathbf{v} \in \mathbb{C}^n$ (nebo $\mathbf{v} \in \mathbb{R}^n$), pak $\mathbf{v}$ je vlastní vektor matice A příslušný vlastnímu číslu λ právě tehdy, když je $\mathbf{v}$ vlastní vektor matice A^* příslušný vlastnímu číslu $\overline{\lambda}$.

(8) **Spektrální věta pro normální matice**. Je-li A komplexní matice řádu n, pak jsou následující tvrzení ekvivalentní
   (a) matice A je unitárně (ortogonálně) diagonalizovatelná,
   (b) matice A je normální.

(9) **Spektrální věta pro hermitovské matice**. Pro čtvercovou matici A nad $\mathbb{C}$ jsou následující tvrzení jsou ekvivalentní:
   (a) matice A je unitárně diagonalizovatelná a všechna její vlastní čísla jsou reálná,
   (b) matice A je hermitovská.

(10) **Spektrální věta pro symetrické matice**. Pro reálnou čtvercovou matici A jsou následující tvrzení jsou ekvivalentní:
   (a) matice A je ortogonálně diagonalizovatelná a všechna její vlastní čísla jsou reálná,
   (b) matice A je symetrická.

(11) Komplexní (nebo reálná) matice A řádu n se nazývá
   - *pozitivně definitní*, pokud je hermitovská (nebo symetrická) a platí $\mathbf{x} \cdot A\mathbf{x} \geq 0$ pro každý nenulový vektor $\mathbf{x} \in \mathbb{C}^n$ (nebo $\mathbf{x} \in \mathbb{R}^n$),
   - *pozitivně semidefinitní*, pokud je hermitovská (nebo symetrická) a platí $\mathbf{x} \cdot A\mathbf{x} \geq 0$ pro každé $\mathbf{x} \in \mathbb{C}^n$ (nebo $\mathbf{x} \in \mathbb{R}^n$).

(12) **Spektrální věta pro pozitivně (semi)definitní matice**. Pro hermitovskou (symetrickou) matici A je ekvivalentní
   (a) A je pozitivně definitní (nebo semidefinitní),

(b) všechna vlastní čísla matice A jsou kladná (nebo nezáporná).

(13) Komplexní (nebo reálná) matice A je pozitivně semidefinitní právě když $A = B^*B$ (nebo $A = B^TB$) pro nějakou komplexní (nebo reálnou) matici B.

(14) **Spektrální věta pro unitární (ortogonální) matice**. Pro čtvercovou komplexní matici A jsou následující podmínky ekvivalentní:

(a) matice A je unitárně diagonalizovatelná a pro každé vlastní číslo λ matice A platí $|\lambda| = 1$,

(b) matice A je unitární.

(15) Operátor $f : \mathbf{V} \to \mathbf{V}$ na komplexním (nebo reálném) prostoru $\mathbf{V}$ se skalárním součinem $\langle\, , \rangle$ se nazývá unitární (ortogonální), pokud pro každý prvek $\mathbf{x} \in \mathbf{V}$ platí

$$\|f(\mathbf{x})\| = \|\mathbf{x}\| \ .$$

(16) Pro lineární operátor f na komplexním (reálném) lineárním prostoru $\mathbf{V}$ se skalárním součinem $\langle\, , \rangle$ jsou následující podmínky ekvivalentní

(a) f je unitární (ortogonální)

(b) $\langle f(\mathbf{x}), f(\mathbf{y})\rangle = \langle \mathbf{x}, \mathbf{y}\rangle$ pro každé $\mathbf{x}, \mathbf{y} \in \mathbf{V}$

Má-li $\mathbf{V}$ navíc konečnou dimenzi, pak jsou předchozí podmínky ekvivalentní také s

(c) matice $[f]_B^B$ vzhledem k ortonormální bázi B ve $\mathbf{V}$ je unitární (nebo ortogonální)

(d) f zobrazuje každou ortonormální bázi ve $\mathbf{V}$ opět na ortonormální bázi ve $\mathbf{V}$

(e) f zobrazuje nějakou ortonormální bázi ve $\mathbf{V}$ opět na ortonormální bázi ve $\mathbf{V}$

(17) Každé ortogonální zobrazení $f : \mathbb{R}^2 \to \mathbb{R}^2$ na prostoru $\mathbb{R}^2$ se standardním skalárním součinem je buď reflexe (tj. osová symetrie) nebo rotace.

Zobrazení je reflexe právě když $\det[f]_B^B = -1$ a je rotace právě když $\det[f]_B^B = 1$ pro jakoukoliv bázi B v $\mathbb{R}^2$.

(18)
- Složení dvou reflexí v $\mathbb{R}^2$ je rotace,
- složení rotace s reflexí je reflexe,
- složení dvou rotací je rotace.

(19) Každé ortogonální zobrazení v euklidovském prostoru $\mathbb{R}^3$ je buď rotace kolem nějaké osy, ortogonální reflexe vzhledem k nějaké rovině a nebo složení rotace s ortogonální reflexí. Rotace je to právě tehdy, když determinant matice tohoto zobrazení vzhledem k jakékoliv bázi je rovný 1.

(20) Složení dvou rotací v $\mathbb{R}^3$ je zase rotace v $\mathbb{R}^3$, složení dvou reflexí je rotace (osa rotace je rovná průniku rovin reflexí).

(21) Je-li A matice typu $m \times n$ nad $\mathbb{C}$, pak kladné odmocniny z nenulových vlastních čísel matice A^*A nazýváme *singulární čísla* matice A.

(22) **Věta o singulárním rozkladu, geometrická varianta**. Pro každou matici $A \in \mathbb{C}^{m\times n}$ hodnosti r existují ortonormální báze $B = (\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n)$ v prostoru $\mathbb{C}^n$, $C = (\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m)$ v prostoru $\mathbb{C}^m$, a reálná čísla

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$$

taková, že

$$[f_A]_C^B = \begin{pmatrix} \Sigma_r & 0 \\ 0 & 0 \end{pmatrix}$$

neboli $f_A(\mathbf{v}_j) = A\mathbf{v}_j = \begin{cases} \sigma_j \mathbf{u}_j, & \text{pro } j = 1, 2, \ldots, r \\ \mathbf{o}, & \text{pro } j = 1, 2, \ldots, r \end{cases}$

(23) **Věta o singulárním rozkladu, algebraická varianta**. Pro každou matici $A \in \mathbb{C}^{m\times n}$ hodnosti r existují unitární matice U řádu m, unitární matice V řádu n, a reálná čísla

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$$

taková, že

$$A = U\Sigma V^*$$

kde $\Sigma = \begin{pmatrix} \Sigma_r & 0 \\ 0 & 0 \end{pmatrix} \in \mathbb{R}^{m\times n}, \quad \Sigma_r = \mathrm{diag}(\sigma_1, \sigma_2, \dots, \sigma_r)$.

(24) Singulární rozklad $A = UDV^*$ matice typu $m\times n$ hodnosti r, kde $U = (\mathbf{u}_1|\dots|\mathbf{u}_m)$, $D = \mathrm{diag}_{m\times n}(\sigma_1,\dots,\sigma_r)$ a $V = (\mathbf{v}_1|\dots|\mathbf{v}_n)$ můžeme také zapsat jako

$$A = \sigma_1\mathbf{u}_1\mathbf{v}_1^* + \dots + \sigma_r\mathbf{u}_r\mathbf{v}_r^* \ ,$$

říkáme tomu *dyadický rozvoj* matice A.

(25) Ze singulárního rozkladu $A = UDV^T$ dostaneme úpravou $A = (UDU^T)(UV^T)$. Matice UDU^T je pozitivně semidefinitní a matice UV^T je unitární, takže můžeme položit $R = UDU^T$ a $W = UV^T$ a matici A vyjádřit ve tvaru $A = RW$. Poslednímu vyjádření říkáme *polární rozklad* matice A. Polární rozklad lze chápat jako zobecnění rozkladu komplexního čísla na součin nezáporného reálného čísla a komplexního čísla jednotkové velikosti.

(26) Je-li A komplexní matice typu $m\times n$, pak číslo $\max\{\|A\mathbf{x}\| : \mathbf{x}\in\mathbb{C}^n,\ \|\mathbf{x}\| = 1\}$ nazýváme *spektrální norma* matice A, a značíme jej $\|A\|$.

(27) Pro každou matici $A\in\mathbb{C}^{m\times n}$ a pro libovolný nenulový vektor $\mathbf{x}\in\mathbb{C}^n$ platí

$$\frac{\|A\,\mathbf{x}\|}{\|\mathbf{x}\|} \le \sigma_1 \ ,$$

kde σ_1 je největší singulární hodnota matice A, přičemž rovnost nastává právě tehdy, když $\mathbf{x}$ je vlastní vektor matice A^*A příslušný vlastnímu číslu σ_1^2.

(28) Pro každou matici $A\in\mathbb{C}^{m\times n}$ a pro libovolný nenulový vektor $\mathbf{x}\in\mathbb{C}^n$ platí

$$\frac{\|A\,\mathbf{x}\|}{\|\mathbf{x}\|} \ge \sigma_r \ ,$$

kde σ_r je nejmenší singulární hodnota matice A, přičemž rovnost nastává právě tehdy, když $\mathbf{x}$ je vlastní vektor matice A^*A příslušný vlastnímu číslu σ_r^2.

(29) Je-li $A = U\,\Sigma\,V^*$ singulární rozklad matice A s hodností r, pak matice matice $A^\dagger = V\,\Sigma^\dagger\,U^*$, kde

$$\Sigma^\dagger = \begin{pmatrix} \Sigma_r^{-1} & 0 \\ 0 & 0 \end{pmatrix}$$

se nazývá *Mooreova-Penroseova pseudoinverze* matice A.

(30) Je-li A regulární, platí $A^\dagger = A^{-1}$.

Klíčové znalosti z desáté kapitoly nezbytné pro průběžné sledování přednášek s pochopením

(1) Definice unitární diagonalizovatelnosti pro operátory a matice.
(2) Varianty spektrálních vět pro různé typy matic.
(3) Různé ekvivalentní definice unitárních a ortogonálních operátorů.
(4) Charakterizace ortogonálních operátorů v $\mathbb{R}^2$ a $\mathbb{R}^3$.
(5) Singulární čísla matice a obě varinaty věty o singulárním rozkladu.
(6) Dyadický rozvoj matice, polární rozklad matice, spektrální norma matice, Mooreova-Penroseova pseudoinverze.

Upozornění. Spektrální věty jsou ve skriptech formulovány pro operátory, v přehledu pro matice tak, jak byly probírány na přednášce.

11. Bilineární formy a kvadratické formy

Cíl. *Bilineární formu lze chápat jako zobecnění skalárního součinu. Ponecháme pouze vlastnosti linearity v každé složce a vzdáme se symetrie a pozitivní definitnosti. Taková zobecnění skalárního součinu se používají například ve fyzice, konkrétně ve speciální teorii relativity. Naší hlavní motivací pro studium bilineárních forem je porozumění kvadratickým formám, které určují „kvadratické útvary“. Ukážeme, že kvadratické formy vzájemně jednoznačně odpovídají symetrickým bilineárním formám. Hlavní náplní bude nalezení báze, vzhledem ke které má symetrická bilineární forma, a tím i příslušná kvadratická forma, jednoduchý tvar. To nám umožní analyzovat tvar kvadratických útvarů.*

Bilineární forma je zobrazení přiřazující každé dvojici vektorů prvek tělesa, které je lineární v obou složkách.

Definice 11.1. Nechť $\mathbf{V}$ je vektorový prostor nad tělesem $\mathbf{T}$. *Bilineární forma* na prostoru $\mathbf{V}$ je zobrazení $f : V \times V \to T$, které je lineární v obou složkách, tj. pro libovolné $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w} \in V$, $t \in T$ platí

(1) $f(\mathbf{u}+\mathbf{v},\mathbf{w}) = f(\mathbf{u},\mathbf{w}) + f(\mathbf{v},\mathbf{w})$, $f(\mathbf{w},\mathbf{u}+\mathbf{v}) = f(\mathbf{w},\mathbf{u}) + f(\mathbf{w},\mathbf{v})$ a
(2) $f(t\mathbf{v},\mathbf{w}) = t\,f(\mathbf{v},\mathbf{w})$, $f(\mathbf{v},t\mathbf{w}) = tf(\mathbf{v},\mathbf{w})$.

Příklad 11.2. Bilineární formou na $\mathbb{R}^3$ je například zobrazení

$$\begin{aligned} f((x_1,x_2,x_3)^T,(y_1,y_2,y_3)^T) &= 2x_1y_1 - 3x_1y_2 + 5x_1y_3 + 6x_2y_1 + x_2y_3 + 10x_3y_2 \\ &= (x_1\ x_2\ x_3) \begin{pmatrix} 2 & -3 & 5 \\ 6 & 0 & 1 \\ 0 & -10 & 0 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} . \end{aligned}$$

Uvidíme, že každá bilineární forma na aritmetickém vektorovém prostoru $\mathbf{T}^n$ je tvaru $f(\mathbf{x},\mathbf{y}) = \mathbf{x}^T A\mathbf{y}$, pro nějakou čtvercovou matici řádu n nad $\mathbf{T}$.

Příklad 11.3. Libovolný skalární součin na reálném vektorovém prostoru $\mathbf{V}$ je bilineární forma na $\mathbf{V}$ (která je navíc symetrická a pozitivně definitní). Bilineární formy tedy můžeme chápat jako zobecnění skalárního součinu. Axiomy (1) a (2) jsme využili k odvození formulky pro standardní skalární součin.

Obecněji, pro libovolný operátor g na reálném prostoru $\mathbf{V}$ se skalárním součinem $\langle\,,\rangle$ je $f(\mathbf{x},\mathbf{y}) = \langle\,\mathbf{x}, g(\mathbf{y})\rangle$ bilineární forma. Takové bilineární formy jsme potkali při sdružování lineárních zobrazení.

Pozor! Skalární součin na komplexním vektorovém prostoru bilineární forma není — vlastnost $f(t\mathbf{u},\mathbf{v}) = f(\mathbf{u},\mathbf{v})$ bychom museli nahradit vlastností $f(t\mathbf{u},\mathbf{v}) = \overline{t}f(\mathbf{u},\mathbf{v})$. Takovým formám se říká *seskvilineární* a nebudeme se jimi podrobněji zabývat.

Příklad 11.4. Zobrazení $\mathbf{T}^2 \times \mathbf{T}^2 \to \mathbf{T}$ definované vztahem $f(\mathbf{x},\mathbf{y}) = \det(\mathbf{x},\mathbf{y})$ je bilineární forma na $\mathbf{T}^2$. Axiomy (1) a (2) byly také základní vlastnosti použité při odvození vzorce pro determinant matic 2×2.

Pro matice vyšších řádů je determinant příkladem tzv. *multilineární* formy, tedy zobrazení $V \times V \times \cdots \times V \to T$ lineární v každé složce.

My využijeme bilineární formy hlavně při studiu kvadratických forem.

Definice 11.5. Je-li f bilineární forma na vektorovém prostoru $\mathbf{V}$ nad tělesem $\mathbf{T}$, pak zobrazení $f_2 : V \to T$ definované předpisem

$$f_2(\mathbf{v}) = f(\mathbf{v},\mathbf{v}) \qquad \text{pro každé } \mathbf{v} \in V$$

nazýváme *kvadratickou formou* vytvořenou bilineární formou f.

Rovněž říkáme, že f_2 je kvadratická forma příslušná bilineární formě f.

Příklad 11.6. Pro bilineární formu na $\mathbb{R}^3$ z příkladu 11.2 je

$$\begin{aligned} f_2((x_1,x_2,x_3)^T) &= 2x_1^2 - 3x_1x_2 + 5x_1x_3 + 6x_2x_1 + x_2x_3 + 10x_3x_2 \\ &= 2x_1^2 + 3x_1x_2 + 5x_1x_3 + 11x_2x_3 \end{aligned}$$

Maticově,

$$f_2(\mathbf{x}) = \mathbf{x}^T A\mathbf{x} \ ,$$

kde A je matice řádu 3 ze stejného příkladu.

Příklad 11.7. Je-li f skalární součin na reálném vektorovém prostoru $\mathbf{V}$ (tj. $f(\mathbf{x},\mathbf{y}) = \langle\, \mathbf{x},\mathbf{y}\rangle$), pak $f_2(\mathbf{x}) = \|\mathbf{x}\|^2$.

Kvadratické formy se vyskytují při analýze funkcí více proměnných. Například nás zajímá, jak vypadá daná hladká funkce $h : \mathbb{R}^2 \to \mathbb{R}$ v okolí nějakého bodu $\mathbf{d} \in \mathbb{R}^2$, řekněme $\mathbf{d} = (0,0)^T$. Velmi hrubá aproximace je nahradit funkci její funkční hodnotou $c = h(\mathbf{d})$

$$h(x_1,x_2) \approx c \ .$$

Přesnější je lineární aproximace, kdy nahradíme funkci její tečnou rovinou

$$h(x_1,x_2) \approx c + b_1x_1 + b_2x_2 \ .$$

Nekonstantní část $g(x_1,x_2) = b_1x_1 + b_2x_2$ je lineární forma na $\mathbb{R}^2$, koeficienty b_1, b_2 se vypočtou pomocí parciálních derivací. Ještě přesnější je aproximace polynomem stupně 2:

$$h(x_1,x_2) \approx c + b_1x_1 + b_2x_2 + a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2$$

Kvadratická část $f(x_1,x_2) = a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2$ je kvadratická forma na $\mathbb{R}^2$ (koeficienty se vypočtou z druhých parciálních derivací). Tato aproximace je důležitá například při hledání extrémů.

Proto nás zajímá, jak vypadá graf kvadratické funkce více proměnných. Obecněji nás zajímá, jak vypadá implicitně zadaný kvadratický útvar, například množina bodů v $\mathbb{R}^3$ splňujících rovnici

$$10x_1^2 + 13x_2^2 + 13x_3^2 + 4x_1x_2 + 4x_1x_3 + 8x_2x_3 = 9 \ .$$

Základní myšlenka na řešení takových problémů je stejná jako u lineárních operátorů: najít bázi, vzhledem ke které je bilineární forma přehledná.

11.1. **Matice.**

Podobně jako v úvodu do determinantů spočítáme, že každá bilineární forma je určena obrazy dvojic prvků libovolné báze. To nám dává maticovou reprezentaci bilineárních forem a tzv. analytické vyjádření. Nechť f je bilineární forma na $\mathbf{V}$ a $B = (\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n)$ je báze prostoru $\mathbf{V}$. Vezmeme dva vektory $\mathbf{x}, \mathbf{y} \in V$ a vyjádříme $f(\mathbf{x},\mathbf{y})$ pomocí souřadnic vektorů $\mathbf{x}, \mathbf{y}$ v bázi B a pomocí hodnot $a_{ij} = f(\mathbf{v}_i, \mathbf{v}_j)$:

$$[\mathbf{x}]_B = (x_1, x_2, \ldots, x_n)^T, \quad [\mathbf{y}]_B = (y_1, y_2, \ldots, y_n)^T \ .$$

$$
\begin{aligned}
f(\mathbf{x},\mathbf{y}) &= f(x_1\mathbf{v}_1 + \dots + x_n\mathbf{v}_n, y_1\mathbf{v}_1 + \dots + y_n\mathbf{v}_n) = f\left(\sum_{i=1}^n x_i\mathbf{v}_i, \sum_{i=1}^n y_i\mathbf{v}_i\right) \\
&= \sum_{i=1}^n x_i f\left(\mathbf{v}_i, \sum_{j=1}^n y_j\mathbf{v}_j\right) = \sum_{i=1}^n\sum_{j=1}^n x_i y_j f(\mathbf{v}_i,\mathbf{v}_j) \\
&= \sum_{i=1}^n\sum_{j=1}^n x_i y_j a_{ij} \\
&= (x_1\ x_2\ \dots\ x_n)\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & \dots & \dots & \dots \\ \vdots & & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn}\end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n\end{pmatrix}
\end{aligned}
$$

To vede k pojmu matice bilineární formy vzhledem k bázi.

Definice 11.8. Nechť $B = (\mathbf{v}_1, \dots, \mathbf{v}_n)$ je báze vektorového prostoru $\mathbf{V}$ nad tělesem $\mathbf{T}$ a f je bilineární forma na $\mathbf{V}$. *Maticí bilineární formy f vzhledem k B* rozumíme čtvercovou matici řádu n nad $\mathbf{T}$, která má na pozici (i,j) prvek $f(\mathbf{v}_i,\mathbf{v}_j)$. Tuto matici značíme $[f]_B$.

Tvrzení 11.9. *Je-li B báze konečně generovaného prostoru $\mathbf{V}$ a $\mathbf{x},\mathbf{y} \in V$, pak*

$$f(\mathbf{x},\mathbf{y}) = [\mathbf{x}]_B^T[f]_B[\mathbf{y}]_B \ .$$

Jsou-li souřadnice vektorů $[\mathbf{x}]_B = (x_1, \dots, x_n)^T$, $[\mathbf{y}]_B = (y_1, \dots, y_n)^T$ a $[f]_B = (a_{ij})_{n\times n}$, pak

$$f(\mathbf{x},\mathbf{y}) = \sum_{i=1}^n\sum_{j=1}^n a_{ij}x_iy_j \ .$$

Tomuto vyjádření také říkáme *analytické vyjádření bilineární formy f*.

Naopak, každou bilineární formu na konečně generovaném prostoru můžeme vztahem $f(\mathbf{x},\mathbf{y}) = [\mathbf{x}]_B^T A[\mathbf{y}]_B$ definovat a matice A je tímto určená jednoznačně:

Tvrzení 11.10. *Nechť $\mathbf{V}$ je konečně generovaný prostor nad tělesem $\mathbf{T}$, $B = (\mathbf{v}_1, \dots, \mathbf{v}_n)$ je jeho báze a A je čtvercová matice nad $\mathbf{T}$ řádu n. Pak zobrazení $f : V\times V \to T$ definované vztahem*

$$f(\mathbf{x},\mathbf{y}) = [\mathbf{x}]_B^T A[\mathbf{y}]_B \qquad \textit{pro každé } \mathbf{x},\mathbf{y} \in V$$

je bilineární forma na $\mathbf{V}$ a platí $[f]_B = A$.

Důkaz. Pro libovolné $\mathbf{u},\mathbf{v},\mathbf{w}$ platí

$$
\begin{aligned}
f(\mathbf{u}+\mathbf{v},\mathbf{w}) &= [\mathbf{u}+\mathbf{v}]_B^T A[\mathbf{w}]_B = ([\mathbf{u}]_B^T + [\mathbf{v}]_B^T)A[\mathbf{w}]_B = [\mathbf{u}]_B^T A[\mathbf{w}]_B + [\mathbf{v}]_B^T A[\mathbf{w}]_B \\
&= f(\mathbf{u},\mathbf{w}) + f(\mathbf{v},\mathbf{w}) \ .
\end{aligned}
$$

Ostatní axiomy se ověří podobně.

Dosazením $\mathbf{x} = \mathbf{v}_i$ a $\mathbf{y} = \mathbf{v}_j$ získáme

$$f(\mathbf{v}_i,\mathbf{v}_j) = [\mathbf{v}_i]_B^T A[\mathbf{v}_j] = \mathbf{e}_i^T A\mathbf{e}_j \ ,$$

což je prvek na místě (i,j) v matici A, takže skutečně $[f]_B = A$. □

Při pevně zvolené bázi B tedy takto bilineární formy na $\mathbf{V}$ vzájemně jednoznačně odpovídají čtvercovým maticím nad $\mathbf{T}$ řádu n.

Příklad 11.11. Zobrazení $f : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ definované předpisem

$$f((x_1,x_2)^T,(y_1,y_2)^T) = 2x_1y_1 + 4x_2y_1 = (x_1\ x_2)\begin{pmatrix} 2 & 0 \\ 4 & 0\end{pmatrix}\begin{pmatrix} y_1 \\ y_2\end{pmatrix}$$

je bilineární forma na $\mathbb{R}^2$. Jeho matice vzhledem ke kanonické bázi je

$$[f]_{K_2} = \begin{pmatrix} 2 & 0 \\ 4 & 0 \end{pmatrix}.$$

Vezmeme jinou bázi $\mathbb{R}^2$, například

$$B = \left(\begin{pmatrix} 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \end{pmatrix} \right) .$$

Matice f vzhledem k B je podle definice

$$[f]_B = \begin{pmatrix} f((1,-1)^T,(1,-1)^T) & f((1,-1)^T,(2,0)^T) \\ f((2,0)^T,(1,-1)^T) & f((2,0)^T,(2,0)^T) \end{pmatrix} = \begin{pmatrix} -2 & -4 \\ 4 & 8 \end{pmatrix} ,$$

kde například prvek na místě $(1,2)$ spočteme

$$f((1,-1)^T,(2,0)^T) = (1,-1) \begin{pmatrix} 2 & 0 \\ 4 & 0 \end{pmatrix} \begin{pmatrix} 2 \\ 0 \end{pmatrix} = (1,-1) \begin{pmatrix} 4 \\ 8 \end{pmatrix} = -4 .$$

Matice bilineární formy f vzhledem k B nám umožňuje rychle spočítat $f((x_1,x_2)^T,(y_1,y_2)^T)$ známe-li vyjádření vektorů vzhledem k bázi B:

$$[(x_1,x_2)^T]_B = (x_1',x_2')^T, \quad [(y_1,y_2)^T]_B = (y_1',y_2')^T,$$

$$\begin{aligned} f((x_1,x_2)^T,(y_1,y_2)^T) &= (x_1' \; x_2') \begin{pmatrix} -2 & -4 \\ 4 & 8 \end{pmatrix} \begin{pmatrix} y_1' \\ y_2' \end{pmatrix} \\ &= -2x_1'y_1' - 4x_1'y_2' + 4x_2'y_1' + 8x_2'y_2' . \end{aligned}$$

Matici $[f]_B$ spočítáme ještě jedním způsobem, který nám zároveň ukáže, jak se obecně mění matice bilineární formy při přechodu od báze k bázi. Označme X matici přechodu od B ke kanonické bázi K_2.

$$X = [\mathrm{id}]^B_{K_2} = \begin{pmatrix} 1 & 2 \\ -1 & 0 \end{pmatrix}$$

Pro libovolný vektor $\mathbf{z} \in V$ platí $[\mathbf{z}]_{K_2} = X[\mathbf{z}]_B$ a transponováním získáme $[\mathbf{z}]^T_{K_2} = [\mathbf{z}]^T_B X^T$. Pak

$$\begin{aligned} f(\mathbf{x},\mathbf{y}) = (x_1 \; x_2) \begin{pmatrix} 2 & 0 \\ 4 & 0 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} &= (x_1',x_2') \begin{pmatrix} 1 & -1 \\ 2 & 0 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 4 & 0 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} y_1' \\ y_2' \end{pmatrix} \\ &= (x_1',x_2') \begin{pmatrix} -2 & -4 \\ 4 & 8 \end{pmatrix} \begin{pmatrix} y_1' \\ y_2' \end{pmatrix} \end{aligned}$$

Z jednoznačnosti maticového vyjádření f nyní plyne, že matice f vzhledem k B je stejná jako u předchozího výpočtu.

Zobecněním výpočtu v předchozím příkladu dostáváme vztah o změně matice při přechodu od báze k bázi.

Tvrzení 11.12. *Nechť f je bilineární forma na vektorovém prostoru $\mathbf{V}$, B a C jsou báze $\mathbf{V}$ a $X = [\mathrm{id}]^C_B$ je matice přechodu od C k B. Pak $[f]_C = X^T[f]_B X$.*

Důkaz. Pro libovolné vektory $\mathbf{x},\mathbf{y} \in V$ platí

$$f(\mathbf{x},\mathbf{y}) = [\mathbf{x}]^T_B[f]_B[\mathbf{y}]_B = ([\mathrm{id}]^C_B[\mathbf{x}]_C)^T[f]_B([\mathrm{id}]^C_B[\mathbf{y}]_C) = [\mathbf{x}]^T_C X^T[f]_B X[\mathbf{y}]_C .$$

Z jednoznačnosti matice bilineární formy vzhledem k bázi nyní plyne $[f]_C = X^T[f]_B X$. □

Čtvercová matice A řádu n má teď pro nás dva geometrické významy: lineární operátor f_A na $\mathbf{T}^n$ a bilineární forma $\mathbf{x}^T A\mathbf{y}$ na $\mathbf{T}^n$. Všimněte si rozdílu při změně báze. Je-li R matice přechodu od B ke kanonické bázi, pak matice příslušného lineárního operátoru vzhledem k B je $R^{-1}AR$ zatímco matice příslušné bilineární formy vzhledem k B je R^TAR.

11.2. **Symetrické a antisymetrické formy.** Kvadratická forma může být vytvořena různými bilineárními formami, například bilineární formy

$$f((x_1,x_2)^T,(y_1,y_2)^T) = 2x_1y_1 + 3x_1y_2 + x_2y_1, \quad g((x_1,x_2)^T,(y_1,y_2)^T) = 2x_1y_1 + 4x_2y_1$$

vytváří stejnou kvadratickou formu

$$f_2((x_1,x_2)^T) = g_2((x_1,x_2)^T) = 2x_1^2 + 4x_1x_2$$

V této části si, v případě těles charakteristiky různé od dva, jednoznačně rozložíme každou bilineární formu na součet symetrické a antisymetrické, a ukážeme, že vytvořená kvadratická forma je určena symetrickou částí.

Definice 11.13. Bilineární forma f na vektorovém prostoru $\mathbf{V}$ se nazývá

- *symetrická*, pokud pro libovolné $\mathbf{x},\mathbf{y} \in V$ platí $f(\mathbf{x},\mathbf{y}) = f(\mathbf{y},\mathbf{x})$;
- *antisymetrická*, pokud pro libovolné $\mathbf{x},\mathbf{y} \in V$ platí $f(\mathbf{x},\mathbf{y}) = -f(\mathbf{y},\mathbf{x})$.

Příkladem symetrické formy je skalární součin na reálném vektorovém prostoru.

Zda je forma symetrická (antisymetrická) poznáme snadno z matice vzhledem k libovolné bázi.

Tvrzení 11.14. *Nechť* $\mathbf{V}$ *je konečně generovaný vektorový prostor,* B *je báze* $\mathbf{V}$ *a* f *je bilineární forma na* $\mathbf{V}$. *Pak*

- f *je symetrická právě tehdy, když je* $[f]_B$ *symetrická matice;*
- f *je antisymetrická právě tehdy, když je* $[f]_B$ *antisymetrická matice.*

Důkaz. Dokážeme první ekvivalenci, druhá se dokáže podobně. Označme $B = (\mathbf{v}_1,\ldots,\mathbf{v}_n)$.

Prvek na místě (i,j) v matici $[f]_B$ je podle definice rovný $f(\mathbf{v}_i,\mathbf{v}_j)$. Je-li tedy f symetrický pak prvek na místě (i,j) je stejný jako prvek na místě (j,i), takže $[f]_B$ je symetrická matice.

Je-li naopak $[f]_B$ symetrická matice, pak pro libovolné vektory $\mathbf{x},\mathbf{y} \in V$ platí

$$f(\mathbf{x},\mathbf{y}) = [\mathbf{x}]_B^T[f]_B[\mathbf{y}]_B = [\mathbf{x}]_B^T[f]_B^T[\mathbf{y}]_B = ([\mathbf{x}]_B^T[f]_B^T[\mathbf{y}]_B)^T = [\mathbf{y}]_B^T[f]_B[\mathbf{x}]_B = f(\mathbf{y},\mathbf{x}) \ ,$$

kde ve třetí rovnosti jsme využili, že $(t)^T = t$ pro libovolný skalár $t \in T$. □

Bilineární formy můžeme přirozeným způsobem sčítat a násobit skalárem. Jsou-li f,g dvě bilineární formy na $\mathbf{V}$ a $t \in T$ pak definujeme

$$(f+g)(\mathbf{x},\mathbf{y}) = f(\mathbf{x},\mathbf{y}) + g(\mathbf{x},\mathbf{y}), \quad (tf)(\mathbf{x},\mathbf{y}) = tf(\mathbf{x},\mathbf{y}).$$

S těmito operacemi tvoří množina všech bilineárních forem na $\mathbf{V}$ vektorový prostor. Je-li B konečná báze $\mathbf{V}$, snadno se ověří vztahy

$$[f+g]_B = [f]_B + [g]_B, \quad [tf]_B = t[f]_B \ .$$

Zamyslíme se nyní, jak rozložit danou bilineární formu f na prostoru $\mathbf{V}$ nad tělesem $\mathbf{T}$ na součet symetrické formy f_s a antisymetrické formy f_a. Pro konečně generované prostory je tento úkol ekvivalentní rozkladu čtvercové matice na součet symetrické a antisymetrické. Pro libovolné dva vektory $\mathbf{x},\mathbf{y} \in V$ chceme, aby platilo

$$\begin{aligned} f(\mathbf{x},\mathbf{y}) &= f_s(\mathbf{x},\mathbf{y}) + f_a(\mathbf{x},\mathbf{y}) \\ f(\mathbf{y},\mathbf{x}) &= f_s(\mathbf{y},\mathbf{x}) + f_a(\mathbf{y},\mathbf{x}) = f_s(\mathbf{x},\mathbf{y}) - f_a(\mathbf{x},\mathbf{y}) \end{aligned}$$

Dostali jsme pro $f_s(\mathbf{x},\mathbf{y})$ a $f_a(\mathbf{x},\mathbf{y})$ soustavu dvou rovnic s řešením

$$f_s(\mathbf{x},\mathbf{y})) = \frac{1}{2}(f(\mathbf{x},\mathbf{y}) + f(\mathbf{y},\mathbf{x})), \quad f_a(\mathbf{x},\mathbf{y})) = \frac{1}{2}(f(\mathbf{x},\mathbf{y}) - f(\mathbf{y},\mathbf{x})) \ .$$

Je snadné nahlédnout, že bilineární forma f_s definovaná tímto předpisem je symetrická a f_a je antisymetrická. Problém je pouze v případě, kdy soustava má singulární matici, tj. v případě, že $1 = -1$, ekvivalentně, charakteristika tělesa $\mathbf{T}$ je 2. V opačném případě z postupu vyplývá, že f_s, f_a jsou určeny jednoznačně. Dokázali jsme tak následující tvrzení.

Tvrzení 11.15. *Nechť* $\mathbf{V}$ *je vektorový prostor nad tělesem* $\mathbf{T}$ *charakteristiky různé od* 2. *Pak každou bilineární formu* f *na* $\mathbf{V}$ *lze psát jako součet* $f = f_s + f_a$, *kde* f_s *je symetrická a* f_a *je antisymetrická. Tento rozklad je jednoznačný a platí*

$$f_s(\mathbf{x}, \mathbf{y}) = \frac{1}{2}(f(\mathbf{x}, \mathbf{y}) + f(\mathbf{y}, \mathbf{x})), \quad f_a(\mathbf{x}, \mathbf{y}) = \frac{1}{2}(f(\mathbf{x}, \mathbf{y}) - f(\mathbf{y}, \mathbf{x})) \ .$$

Množina symetrických bilineárních forem na $\mathbf{V}$ i množina antisymetrických bilineárních forem na $\mathbf{V}$ tvoří podprostory prostoru všech bilineárních forem na $\mathbf{V}$ (cvičení). Tvrzení lze formulovat také tak, že vektorový prostor všech bilineárních forem na $\mathbf{V}$ je direktním součtem těchto dvou podprostorů.

Příklad 11.16. Bilineární forma f na $\mathbb{R}^2$

$$f((x_1, x_2)^T, (y_1, y_2)^T) = 2x_1y_1 + 4x_2y_1 + 2x_1y_2 = (x_1 \ x_2) \begin{pmatrix} 2 & 2 \\ 4 & 0 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$

je součtem

$$f_s((x_1, x_2)^T, (y_1, y_2)^T) = 2x_1y_1 + 3x_2y_1 + 3x_1y_2 = (x_1 \ x_2) \begin{pmatrix} 2 & 3 \\ 3 & 0 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$

$$f_a((x_1, x_2)^T, (y_1, y_2)^T) = x_1y_2 - x_2y_1 = (x_1 \ x_2) \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$

To odpovídá maticovému vztahu

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

Pro tělesa charakteristiky dva, například $\mathbf{T} = \mathbb{Z}_2$, je teorie bilineárních forem odlišná, ale tímto případem se nebudeme zvlášť zabývat. Poznamenejme jen, že pojmy symetrická a antisymetrická v tomto případě splývají (cvičení).

Bilineární formy využíváme mimo jiné ke studiu příslušných kvadratických forem. Tato kvadratická forma závisí pouze na symetrické části bilineární formy:

Tvrzení 11.17. *Nechť* f, g *jsou bilineární formy na vektorovém prostoru* $\mathbf{V}$ *nad tělesem charakteristiky různé od* 2. *Pak* $f_2 = g_2$ *právě tehdy, když* $f_s = g_s$. *Navíc*

$$f_s(\mathbf{x}, \mathbf{y}) = \frac{1}{2}(f_2(\mathbf{x} + \mathbf{y}) - f_2(\mathbf{x}) - f_2(\mathbf{y})) \ .$$

Důkaz. Je-li g antisymetrická forma, pak pro libovolný vektor $\mathbf{x} \in V$ platí $g(\mathbf{x}, \mathbf{x}) = -g(\mathbf{x}, \mathbf{x})$. Pokud je charakteristika tělesa různá od dva, vyplývá z tohoto vztahu $g_2(\mathbf{x}) = g(\mathbf{x}, \mathbf{x}) = 0$. Pro libovolnou bilineární formu f pak máme

$$f_2(\mathbf{x}) = f(\mathbf{x}, \mathbf{x}) = f_s(\mathbf{x}, \mathbf{x}) + f_a(\mathbf{x}, \mathbf{x}) = f_s(\mathbf{x}, \mathbf{x}) \ .$$

Vytvořená kvadratická forma tedy závisí jen na symetrické části. Odtud plyne implikace zprava doleva.

Vzorec z tvrzení ověříme přímočarým výpočtem.

$$\begin{aligned} \frac{1}{2}(f_2(\mathbf{x} + \mathbf{y}) - f_2(\mathbf{x}) - f_2(\mathbf{y})) &= \frac{1}{2}(f_s(\mathbf{x} + \mathbf{y}, \mathbf{x} + \mathbf{y}) - f_s(\mathbf{x}, \mathbf{x}) - f_s(\mathbf{y}, \mathbf{y})) \\ &= \frac{1}{2}(f_s(\mathbf{x}, \mathbf{x}) + f_s(\mathbf{x}, \mathbf{y}) + f_s(\mathbf{y}, \mathbf{x}) + f_s(\mathbf{y}, \mathbf{y}) - f_s(\mathbf{x}, \mathbf{x}) - f_s(\mathbf{y}, \mathbf{y})) \\ &= \frac{1}{2}(2f_s(\mathbf{x}, \mathbf{y})) = f_s(\mathbf{x}, \mathbf{y}) \end{aligned}$$

Implikace zleva doprava je nyní zřejmá. □

Vztah v předchozí větě je varianta polarizační identity z tvrzení 8.25. Dává explicitní vzorec na výpočet hodnoty symetrické bilineární formy pomocí příslušné formy kvadratické. Tuto jednoznačně určenou symetrickou bilineární formu také nazýváme *symetrická forma příslušná dané kvadratické formě.*

Příklad 11.18. Uvažujme kvadratickou formu

$$f_2((x_1,x_2)^T) = 2x_1^2 + 7x_1x_2 + 5x_2^2$$

Pro nalezení symetrické formy f_s není třeba používat vzorec z předchozího tvrzení, stačí si uvědomit z jakých členů bilineární formy pochází členy f_2. Je-li

$$f((x_1,x_2)^T,(y_1,y_2)^T) = a_{11}x_1y_1 + a_{12}x_1y_2 + a_{21}x_2y_1 + a_{22}x_2y_2$$

a $f_2((x_1,x_2)^T) = f((x_1,x_2)^T,(x_1,x_2)^T)$, pak koeficient u x_1^2 v kvadratické formě f_2 musí pocházet ze členu $a_{11}x_1y_1$, tedy $a_{11} = 2$. Podobně $a_{22} = 5$. Koeficient u x_1x_2 vznikne součtem $a_{12} + a_{21}$ a kvůli symetrii je $a_{12} = a_{21} = 7/2$. Takže je

$$f_s((x_1,x_2)^T,(y_1,y_2)^T) = 2x_1^2 + 3{,}5x_1y_2 + 3{,}5x_2y_1 + 5x_2y_2 = (x_1\ x_2)\begin{pmatrix} 2 & 3{,}5 \\ 3{,}5 & 5 \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}.$$

11.3. **Ortogonální báze.**

V celém zbytku kapitoly se budeme věnovat pouze symetrickým formám nad tělesy charakteristiky různé od 2. Budeme se snažit najít bázi vzhledem k níž má daná bilineární forma co nejjednodušší matici, ideálně diagonální. Narozdíl od lineárních operátorů to vždy lze provést.

Symetrické bilineární formy vzájemně jednoznačně odpovídají kvadratickým. Všechny pojmy a výsledky pro symetrické bilineární pojmy budeme proto používat i pro příslušné kvadratické formy.

Co pro bilineární formu f znamená, že matice vzhledem k bázi $B = (\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n)$ je diagonální? Podle definice musí pro dva různé vektory $\mathbf{v}_i$, $\mathbf{v}_j$, $i \neq j$ platit $f(\mathbf{v}_i, \mathbf{v}_j) = 0$. To motivuje pojem ortogonality vektorů.

Definice 11.19. Nechť f je symetrická bilineární forma na $\mathbf{V}$ a $\mathbf{x}, \mathbf{y} \in V$. Říkáme, že $\mathbf{x}$ a $\mathbf{y}$ jsou *f-ortogonální*, pokud $f(\mathbf{x}, \mathbf{y}) = 0$. Zapisujeme $\mathbf{x} \perp_f \mathbf{y}$.

Báze $B = (\mathbf{v}_1, \dots \mathbf{v}_n)$ prostoru $\mathbf{V}$ se nazývá *f-ortogonální*, pokud je $[f]_B$ diagonální, tj. pro libovolné $i, j \in \{1, 2, \dots, n\}$, $i \neq j$, jsou vektory $\mathbf{v}_i, \mathbf{v}_j$ f-ortogonální.

(Pokud je f zřejmé z kontextu, říkáme někdy pouze ortogonální.)

V případě, že f je skalární součin na reálném vektorovém prostoru, se pojmy shodují s již zavedenými. Na hledání ortogonální báze v takovém případě můžeme použít například Gram-Schmidtův ortogonalizační proces. Pro obecnou symetrickou bilineární formu lze zavést obdoby dalších pojmů z kapitoly o skalárním součinu (jako například ortogonální doplněk), teorie je ale o něco složitější a nebudeme se jí věnovat.

Má-li f vzhledem k B diagonální matici $\mathrm{diag}(a_1, \dots, a_n)$, pak pro příslušnou kvadratickou formu platí

$$f_2(\mathbf{x}) = a_1x_1^2 + \dots + a_nx_n^2, \quad [\mathbf{x}]_B = (x_1, \dots, x_n)$$

Z takového vyjádření lépe vidíme, jak daná kvadratická forma vypadá. Na obrázku jsou znázorněny grafy několika kvadratických forem na $\mathbb{R}^2$.

OBRAZEK

11.3.1. *Hodnost.* Je-li f bilineární forma na konečně generovaném prostoru $\mathbf{V}$ a B, C jsou báze prostoru $\mathbf{V}$, pak podle tvrzení 11.12 platí $[f]_C = X^T[f]_BX$, kde X je matice přechodu od C k B. Protože X je regulární, podle důsledku 5.87 o hodnosti součinu s regulární maticí platí $r([f]_C) = r([f]_B)$. To nám umožňuje zavést hodnost bilineární formy.

Definice 11.20. *Hodností* bilineární formy f na konečně generovaném prostoru $\mathbf{V}$ rozumíme hodnost její matice vzhledem k libovolné bázi, značíme $r(f)$.

Je-li matice symetrické bilineární formy f vzhledem k B diagonální matice $D = [f]_B$, pak hodnost $r(D)$ je rovná počtu nenulových prvků na diagonále. Počet nul tedy nezávisí na volbě f-ortogonální báze.

11.3.2. *Metoda symetrických úprav.* Předpokládejme, že f je bilineární forma na vektorovém prostoru $\mathbf{V}$ dimenze n nad tělesem $\mathbf{T}$, C je báze $\mathbf{V}$ a $A = [f]_C$. Vytvoříme matici typu $n \times 2n$ tak, že vedle A napíšeme jednotkovou matici, tj. $(A|I_n)$. S touto maticí provádíme tzv. *symetrické úpravy.* Jedna symetrická úprava sestává z elementární řádkové úpravy a následné „stejné“ úpravy na sloupce. Máme tedy tři typy symetrických úprav:

- prohození i-tého a j-tého řádku, následné prohození i-tého a j-tého sloupce,
- vynásobení i-tého řádku nenulovým prvkem $t \in T$, následné vynásobení i-tého sloupce prvkem t,
- přičtení t-násobku i-tého řádku k j-tému, kde $t \in T$ a $i \neq j$, následné přičtení t-násobku i-tého sloupce k j-tému.

Řádkové úpravy provádíme s celými řádky (vektory z $\mathbf{T}^{2n}$), sloupcové úpravy se vždy týkají jen levého bloku matice.

Odvodíme maticový popis symetrické úpravy matice $(X|Y)$ typu $n \times 2n$. Označíme E matici příslušné řádkové úpravy. Po provedení řádkové úpravy vznikne matice $E(X|Y) = (EX|EY)$. Příslušná sloupcová úprava odpovídá násobení maticí E^T zprava, takže po provedení obou úprav máme matici $(EXE^T|EY)$. Začneme-li tedy s maticí $(A|I_n)$ a provedeme několik symetrických úprav, dostaneme posloupnost matic

$$(A|I_n), (E_1AE_1^T|E_1), (E_2E_1AE_1^TE_2^T|E_2E_1), \ldots, (E_k \ldots E_1AE_1^T \ldots E_k^T|E_k \ldots E_1) \ .$$

Z maticového popisu je vidět, že sloupcové úpravy příslušné řádkovým úpravám není nutné provádět okamžitě. Můžeme je provést kdykoliv, musíme ale zachovat pořadí. Rovněž si všimněte, že po každém kroku je levý blok symetrická matice.

Označme $F = (E_k \ldots E_1)^T$, tj. poslední matice je $(F^TAF|F^T)$. Matice F je regulární, protože je transponovaným součinem elementárních matic. Označme B bázi $\mathbf{V}$ takovou, že $[\mathrm{id}]_C^B = F$, tj. vyjádření vektorů báze B vzhledem k bázi C je ve sloupcích matice F, neboli v řádcích pravého bloku výsledné matice $(F^TAF|F^T)$. Podle tvrzení 11.12 o změně matice bilineární formy při změně báze je matice F^TAF v levé bloku matice $(F^TAF|F^T)$ rovná matici f vzhledem k B.

Tyto úvahy vedou na metodu diagonalizace bilneární formy f. Symetrickými úpravami převedeme matici $(A|I_n)$ do tvaru $(D|G)$, kde D je diagonální. V řádcích matice G pak máme vyjádření vektorů jisté báze B v původní bázi C a platí $[f]_B = D$, tj. speciálně B je f-ortogonální. Jak převod do diagálního tvaru provádět ukážeme na příkladě.

Příklad 11.21. Najdeme f-ortogonální bázi pro bilineární formu f na $\mathbb{Z}_5^3$.

$$[f]_{K_3} = A = \begin{pmatrix} 0 & 1 & 2 \\ 1 & 0 & 1 \\ 2 & 1 & 0 \end{pmatrix}$$

Upravujeme matici $(A|I_n)$ symetrickými úpravami do tvaru $(D|G)$, kde D je diagonální.

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

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

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

Komentář k úpravám: V prvním kroku potřebujeme na pivotní pozici $(1,1)$ nenulový prvek, docílíme toho přičtením druhého řádku k prvnímu (a následnou symetrickou úpravou – přičtení druhého sloupce k prvnímu). Všimněte si, že prohozením řádků v tomto případě ničeho nedocílíme. Kdybychom například prohodili první a druhý řádek, a následně symetricky první a druhý sloupec, na pozici $(1,1)$ by byla stále nula. Po této úpravě jsme přičetli 2-násobek prvního řádku ke druhému a první řádek ke třetímu, a symetricky se sloupci (tím se pouze vynulují pozice $(1,2)$ a $(1,3)$). Nakonec jsme přičetli 4-násobek druhého řádku ke třetímu, a symetricky se sloupci.

Z diskuze nad příkladem vyplývá, že $B = ((1,1,0)^T, (2,3,0)^T, (4,3,1)^T)$ je f-ortogonální báze a $[f]_B = \mathrm{diag}(2,2,1)$.

Věta 11.22. *Každá symetrická bilineární forma f na konečně generovaném vektorovém prostoru nad tělesem charakteristiky různé od 2 má f-ortogonální bázi.*

Důkaz. Podle diskuze nad příkladem se zbývá přesvědčit, že každou čtvercovou matici A řádu n nad tělesem $\mathbf{T}$ lze symetrickými úpravami převést na diagonální tvar. Budeme postupně elimovat řádky a sloupce – po provedení i kroků bude mít matice blokově diagonální tvar

$$A' = \begin{pmatrix} D & 0 \\ 0 & X \end{pmatrix} ,$$

kde D je diagonální matice řádu i. Předpokládejme, že jsme již provedli $i-1$ kroků a provededeme i-tý.

Jsou-li všechny prvky v i-tém sloupci nulové (a tím i prvky v i-tém řádku), nemusíme nic dělat.

Je-li pivot, tj. prvek na místě (i,i) v matici A' nulový a nějaký prvek na místě (j,i) nenulový, řekněme $b \in T$, přičteme j-tý řádek k i-tému a následně j-tý sloupec k i-tému. Tím převedeme matici do tvaru, kdy prvek na místě (i,i) je roven $2b$. Tento prvek není nulový díky tomu, že charakteristika tělesa není 2.

Konečně, je-li prvek na místě (i,i) nenulový, přičteme vhodné násobky i-tého řádku k ostatním řádkům, aby prvky na místech (j,i), $j \neq i$, byly nulové. Příslušné sloupcové úpravy pak pouze vynulují prvky na místech (i,j), $j \neq i$. □

11.3.3. *Bez nulových pivotů.* Jak je vidět z důkazu předchozí věty, při převodu symetrické matice A symetrickými úpravami na diagonální tvar si v řadě případů vystačíme jen s jedním typem symetrických úprav, a to

(*) přičtení t-násobku i-tého řádku k j-tému, kde $t \in T$ a $j > i$ (!) (a následná symetrická sloupcová úprava).

Nastane to v případě, že v každém kroku máme nenulový pivot nebo je celý sloupec (a řádek) nulový. V takovém případě vlastně provádíme Gaussovu eliminaci bez prohazování řádků s tím, že po vyeliminování sloupce vynulujeme také nediagonální hodnoty v příslušném řádku.

Po provedení úprav dostaneme diagonální matici

$$D = E_k \dots E_1 A E_1^T \dots E_k^T$$

složenou z pivotů. Matice E_i řádkové úpravy typu (*) je dolní trojúhelníková s jedničkami na diagonále, součinem takových matic je opět dolní trojúhelníková matice s jedničkami na diagonále a rovněž invertování tuto vlastnost zachovává. To nám dává následující rozklad.

Tvrzení 11.23. *Je-li A symetrická matice taková, že při Gaussově eliminaci nemusíme prohazovat řádky, pak existuje dolní trojúhelníková matice L s jedničkami na diagonále a diagonální matice D (složená z pivotů) tak, že*

$$A = LDL^T \ .$$

Důkaz. Stačí položit $L = (E_k \dots E_1)^{-1}$. Podle diskuze výše je L dolní trojúhelníková s jedničkami na diagonále a platí $D = L^{-1}A(L^{-1})^T$, neboli $LDL^T = A$. □

Příklad 11.24. Najdeme rozklad $A = LDL^T$ pro reálnou symetrickou matici

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

Symetrickými úpravami typu (*) převedeme matici $(A|I_3)$ na tvar $(D|G)$.

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

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

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

Nyní platí $D = GAG^T$. Položíme-li

$$L = G^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 2 & -1 & 1 \end{pmatrix} ,$$

platí $A = LDL^T$.

Kvadratickou formu lze také diagonalizovat tzv. Langrangovou metodou doplňování na čtverce. Tato metoda úzce souvisí s metodou symetrických úprav. Ukážeme si princip na příkladu kvadratické formy na $\mathbb{R}^3$, jejíž příslušná symetrická bilineární forma má matici $A = (a_{ij})$, tj.

$$f_2(\mathbf{x}) = f_2(x_1, x_2, x_3) = a_{11}x_1^2 + a_{22}x_2^2 + a_{33}x_3^2 + 2a_{12}x_1x_2 + 2a_{13}x_1x_2 + 2a_{23}x_2x_3 \ .$$

Pokud $a_{11} \neq 0$, smíšených členů x_{12}, x_{13} se můžeme zbavit doplněním na čtverec

$$\begin{aligned} f_2(\mathbf{x}) = a_{11} & \left(x_1 + \frac{a_{12}}{a_{11}}x_2 + \frac{a_{13}}{a_{11}}x_3 \right)^2 \\ & + \left(a_{22} - \frac{a_{12}^2}{a_{11}} \right) x_2^2 + \left(a_{33} - \frac{a_{13}^2}{a_{11}} \right) x_3^2 + \left(2a_{23} - 2\frac{a_{12}a_{13}}{a_{11}} \right) x_2x_3 \ . \end{aligned}$$

Zvolíme-li novou bázi B tak, aby $[\mathbf{x}]_B = (x_1', x_2', x_3')$, kde

$$x_1' = x_1 + \frac{a_{12}}{a_{11}}x_2 + \frac{a_{13}}{a_{11}}x_3, \quad x_2' = x_2, \quad x_3' = x_3 \ ,$$

pak analytické vyjádření f_2 vzhledem k B je

$$f_2(\mathbf{x}) = a_{11}(x_1')^2 + \left(a_{22} - \frac{a_{12}^2}{a_{11}} \right)(x_2')^2 + \left(a_{33} - \frac{a_{13}^2}{a_{11}} \right)(x_3')^2 + \left(2a_{23} - 2\frac{a_{12}a_{13}}{a_{11}} \right)(x_2')(x_3')$$

a matice příslušné symetrické bilineární formy vzhledem k B je

$$\begin{pmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} - \frac{a_{12}^2}{a_{11}} & a_{23} - \frac{a_{12}a_{13}}{a_{11}} \\ 0 & a_{23} - \frac{a_{12}a_{13}}{a_{11}} & a_{33} - \frac{a_{13}^2}{a_{11}} \end{pmatrix} .$$

To je tatáž matice jako po provedení jednoho kroku metodou důkazu věty 11.22. Vyelimování sloupce (a řádku) metodou symetrických úprav můžeme tedy chápat jako maticový zápis doplnění na čtverce. Symetrické úpravy jsou flexibilnější v tom, že máme více možností úprav a snadnou kontrolu změn bází.

11.4. **Ortogonální báze nad $\mathbb{R}$.**

11.4.1. *Setrvačnost, signatura.* Ortogonální báze ani matice vzhledem k této bázi není určená jednoznačně. Uvažme bilineární formu f na prostoru $\mathbf{V}$ nad tělesem $\mathbf{T}$ a f-ortogonální bázi $B = (\mathbf{v}_1, \dots, \mathbf{v}_n)$ prostoru $\mathbf{V}$. Matice $[f]_B$ je diagonální, řekněme $[f]_B = \mathrm{diag}(a_1, \dots, a_n)$. Vynásobíme i-tý vektor báze B prvkem $t_i \in T$. Vzniklá báze $C = (t_1\mathbf{v}_1, \dots, t_n\mathbf{v}_n)$ je stále f-ortogonální (protože $f(t_i\mathbf{v}_i, t_j\mathbf{v}_j) = t_it_jf(\mathbf{v}_i, \mathbf{v}_j) = 0$ pro $i \neq j$) a na diagonále matice $[f]_C$ jsou prvky $f(t_i\mathbf{v}_i, t_i\mathbf{v}_i) = t_i^2 f(\mathbf{v}_i, \mathbf{v}_i) = t_i^2 a_i$, tj. $[f]_C = \mathrm{diag}(a_1t_1^2, \dots, a_nt_n^2)$.

V případě, že $\mathbf{T} = \mathbb{C}$ z provedené úvahy vyplývá, že pro každou bilineární formu na $\mathbf{V}$ můžeme najít bázi takovou, že $[f]_C = \mathrm{diag}(1, 1, \dots, 1, 0, 0, \dots, 0)$, protože zřejmě pro každé nenulové $a_i \in \mathbb{C}$ můžeme najít $t_i \in \mathbb{C}$ tak, že $a_it_i^2 = 1$.

Pro $\mathbf{T} = \mathbb{R}$ můžeme volbou $t_i = \left(\sqrt{|a_i|}\right)^{-1}$ docílit toho, že $[f]_C$ má na diagonále pouze čísla $1, -1, 0$, tj. při vhodném uspořádání bázových vektorů je $[f]_C = \mathrm{diag}(1, 1, \dots, 1, -1, -1, \dots, -1, 0, 0, \dots, 0)$. Počet jedniček je roven počtu kladných prvků na diagonále $[f]_B$, apod.

Víme, že počet nenulových prvků nezávisí na volbě báze, je roven hodnosti bilineární formy f. Na první pohled ale není jasné, že počet jedniček a minus jedniček také na volbě báze nezávisí. Věta 11.25, tzv. *zákon setrvačnosti kvadratických forem* říká, že tomu tak skutečně je.

Věta 11.25 (Zákon setrvačnosti kvadratických forem)**.** *Nechť f je symetrická bilineární forma na reálném vektorovém prostoru $\mathbf{V}$ dimenze n a C, C' báze $\mathbf{V}$ takové, že*

$$[f]_C = \mathrm{diag}(\underbrace{1, 1, \dots, 1}_{k\times}, \underbrace{-1, -1, \dots, -1}_{l\times}, \underbrace{0, 0, \dots, 0}_{m\times})$$
$$[f]_{C'} = \mathrm{diag}(\underbrace{1, 1, \dots, 1}_{k'\times}, \underbrace{-1, -1, \dots, -1}_{l'\times}, \underbrace{0, 0, \dots, 0}_{m'\times})$$

Pak $k = k', l = l', m = m'$.

Důkaz. Již víme, že $m = m' = n - r(f)$.

Předpokládejme pro spor, že $k > k'$. Označme $C = (\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{v}_1, \dots, \mathbf{v}_l, \mathbf{w}_1, \dots, \mathbf{w}_m)$, $U = \langle \mathbf{u}_1, \dots, \mathbf{u}_k \rangle$, $C' = (\mathbf{u}'_1, \dots, \mathbf{u}'_k, \mathbf{v}'_1, \dots, \mathbf{v}'_l, \mathbf{w}'_1, \dots, \mathbf{w}'_m)$ a $W = \langle \mathbf{v}'_1, \dots, \mathbf{v}'_l, \mathbf{w}'_1, \dots, \mathbf{w}'_m \rangle$. Platí $\dim U = k$, $\dim W = l' + m' = n - k'$ a $\dim(U + W) \leq n$. Podle věty o dimenzi součtu a průniku je

$$\dim(U \cap W) = \dim U + \dim W - \dim(U + W) \geq k + n - k' - n > 0 \ ,$$

takže průnik $U \cap W$ obsahuje nenulový vektor $\mathbf{x} \in U \cap W$. Protože $\mathbf{x} \in U$, ve vyjádření $[\mathbf{x}]_C = (a_1, \dots, a_k, b_1, \dots, b_l, c_1, \dots, c_m)$ máme $b_1 = \dots = b_k = c_1 = \dots = c_m = 0$. Platí tedy

$$\begin{aligned} f_2(x) &= [\mathbf{x}]_C^T [f]_C [\mathbf{x}]_C = 1a_1^2 + \dots 1a_k^2 + (-1)b_1^2 + \dots + (-1)b_l^2 + 0c_1^2 + \dots + 0c_m^2 \\ &= a_1^2 + \dots + a_k^2 > 0. \end{aligned}$$

(Nerovonost je ostrá, protože $\mathbf{x} \neq 0$, takže alespoň jedno a_i je nenulové.)

Podobně, z $\mathbf{x} \in W$ plyne, že ve vyjádření $[\mathbf{x}]'_C = (a'_1, \dots, a'_{k'}, b'_1, \dots, b'_{l'}, c'_1, \dots, c'_{m'})$ je $a'_1 = \dots = a'_{k'} = 0$ a proto

$$\begin{aligned} f_2(x) &= 1(a'_1)^2 + \dots 1(a'_{k'})^2 + (-1)(b'_1)^2 + \dots + (-1)(b'_{l'})^2 + 0(c'_1)^2 + \dots + 0(c'_{m'})^2 \\ &= -(b'_1)^2 + \dots - (b'_{l'})^2 \leq 0 \ , \end{aligned}$$

spor.

Obdobně se ukáže, že nemůže platit $k < k'$. Dokázali jsme, že $m = m'$ a $k = k'$, tedy také $l = l'$. □

Definice 11.26. Nechť f je symetrická bilineární forma na reálném konečně generovaném vektorovém prostoru $\mathbf{V}$. Číslo k (resp. l) z předchozí věty nazýváme *pozitivní (resp.*

negativní) index setrvačnosti formy f, značíme $n_+(f)$ (resp. $n_-(f)$). *Signaturou formy* f rozumíme trojici $(n_0(f), n_+(f), n_-(f))$.

Příklad 11.27. Určíme signaturu bilineární formy f na $\mathbb{R}^3$, jejíž matice vzhledem ke kanonické bázi je

$$A = [f]_K = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} .$$

Symetrickými úpravami převedeme matici do diagonálního tvaru.

$$\begin{pmatrix} 2 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \sim \begin{pmatrix} 2 & 1 & 1 \\ 0 & -\frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{2} & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} 2 & 0 & 0 \\ 0 & -\frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{2} & -\frac{1}{2} \end{pmatrix} \sim \begin{pmatrix} 2 & 0 & 0 \\ 0 & -\frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 0 \end{pmatrix} \sim \begin{pmatrix} 2 & 0 & 0 \\ 0 & -\frac{1}{2} & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Vzniklá matice je maticí stejné bilineární formy f vzhledem k nějaké bázi (která nás teď nezajímala). Signatura f je proto $(1, 1, 1)$.

Příklad 11.28. Určíme signaturu kvadratické formy

$$f_2(x_1, x_2) = 4x_1x_2 + x_2^2$$

na prostoru $\mathbb{R}^2$.

Příslušná symetrická bilineární forma má vzhledem ke kanonické bázi matici

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

Symetrickými úpravami získáme

$$\begin{pmatrix} 0 & 2 \\ 2 & 1 \end{pmatrix} \sim \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 \\ 2 & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 \\ 0 & -4 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 \\ 0 & -4 \end{pmatrix} .$$

(V úpravách jsme tentokrát nepostupovali podle důkazu věty 11.22 – v první úpravě jsme pro pohodlí prohodili první a druhý řádek a následně první a druhý sloupec.) Signatura kvadratické formy f_2 je $(0, 1, 1)$.

11.4.2. *Pozitivní definitnost.* Má-li bilineární forma nenulový pouze index $n_+(f) = n$, mluvíme o pozitivně definitní formě. Obdobně se zavádí *pozitivně semidefinitní* a *negativně (semi)definitní* bilineární formy, o těch však mluvit nebudeme.

Definice 11.29. Symetrická bilineární forma f na reálném vektorovém prostoru $\mathbf{V}$ je *pozitivně definitní*, pokud $f_2(\mathbf{x}) > 0$ pro libovolný vektor $\mathbf{o} \neq \mathbf{x} \in V$.

Tvrzení 11.30. *Symetrická bilineární forma f na reálném vektorovém prostoru $\mathbf{V}$ dimenze n je pozitivně definitní právě tehdy, když $n_+(f) = n$.*

Důkaz. Je-li B ortogonální báze a $[f]_B = \mathrm{diag}(a_1, \ldots, a_n)$, pak pro libovolný vektor $\mathbf{x} \in V$ je $f_2(\mathbf{x}) = f(\mathbf{x}, \mathbf{x})$ rovno

$$f_2(\mathbf{x}) = a_1x_1^2 + \cdots + a_nx_n^2, \text{ kde } [\mathbf{x}]_B = (x_1, \ldots, x_n) .$$

Z toho se snadno vidí obě implikace. Je-li $f_2(\mathbf{x}) > 0$ pro libovolné $\mathbf{o} \neq \mathbf{x} \in V$, pak volbou $[\mathbf{x}]_B = \mathbf{e}_i$ získáme $a_i > 0$ pro každé $i \in \{1, \ldots, n\}$, čili $n_+(f) = n$. Naopak, pokud $n_+(f) = n$, neboli $a_1, \ldots, a_n > 0$, pak je zřejmě $f_2(\mathbf{x}) > 0$ pro libovolný nenulový vektor $\mathbf{x}$. □

Pro reálný vektorový prostor $\mathbf{V}$ je pozitivně definitní symetrická bilineární forma totéž jako skalární součin. Vlastnosti (SL1), (SL2) a (SL3) z definice 8.12 skalárního součinu říkají, že skalární součin je symetrická bilineární forma, a vlastnost (SP) je pozitivní definitnost. Názvy se používají podle toho, jak se na bilineární formu díváme.

Pozitivní definitnost je definovaná v souladu se stejným pojmem pro operátory ve smyslu, že operátor g na prostoru $\mathbf{V}$ se skalárním součinem je pozitivně definitní právě tehdy, když je pozitivně definitní bilineární forma $f(\mathbf{x}, \mathbf{y}) = \langle \mathbf{x}, g(\mathbf{y}) \rangle$. Podobně, matice A řádu n je pozitivně definitní ve smyslu definice 10.25* právě tehdy, když je pozitivně definitní bilineární forma $f(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T A \mathbf{y}$ na aritmetickém prostoru $\mathbb{R}^n$. Navíc platí:

Pozorování 11.31. *Symetrická bilineární forma f na reálném konečně generovaném prostoru $\mathbf{V}$ je pozitivně definitní právě tehdy, když je pozitivně definitní její matice vzhledem k libovolné bázi B.*

Důkaz. Vztah $f_2(\mathbf{x}) > 0$ platí právě tehdy, když $[\mathbf{x}]_B^T [f]_B [\mathbf{x}]_B > 0$. Z toho vyplývá, že $f_2(\mathbf{x}) > 0$ pro každý nenulový vektor $\mathbf{x} \in V$ platí právě tehdy, když $\mathbf{y}^T [f]_B \mathbf{y} > 0$ platí pro každý nenulový vektor $\mathbf{y} \in \mathbb{R}^{\dim \mathbf{V}}$. □

Z části o unitární diagonalizovatelnosti víme, že pozitivně definitní matice jsou právě ty symetrické matice, jejichž vlastní čísla jsou kladná. Charakterizaci nyní můžeme doplnit o další kriteria. *Hlavním minorem* matice A řádu n rozumíme matici tvořenou prvními i řádky a i sloupci matice A pro nějaké $i \in \{1, \dots, n\}$.

Věta 11.32. *Nechť A je reálná symetrická matice řádu n. Následující tvrzení jsou ekvivalentní.*

(1) *A je pozitivně definitní.*
(2) *(Sylvestrovo kritérium) Všechny hlavní minory matice A mají kladný determinant.*
(3) *Gaussova eliminace použitá na matici A může proběhnout bez prohazování řádků a všechny pivoty vyjdou kladné.*
(4) *$A = LDL^T$ pro nějakou dolní trojúhelníkovou matici L s jedničkami na diagonále a nějakou diagonální matici D s kladnými čísly na diagonále.*
(5) *(Choleského rozklad) $A = RR^T$ pro nějakou regulární dolní trojúhelníkovou matici R.*

Důkaz. (1) ⇒ (2). Nejprve dokážeme, že každý minor A_i matice A tvořený prvními i řádky a i sloupci je pozitivně definitní. Vezmeme libovolný nenulový vektor $\mathbf{y} \in \mathbb{R}^i$ a doplníme jej nulami na vektor $\mathbf{x} \in \mathbb{R}^n$. Protože A je pozitivně definitní, platí $\mathbf{x} A \mathbf{x} > 0$. Pak ale

$$\mathbf{y}^T A_i \mathbf{y} = \mathbf{x}^T A \mathbf{x} > 0 \ .$$

Matice A_i je podle důskedku 10.23 ortogonálně diagonalizovatelná, má proto i vlastních čísel $\lambda_1, \dots, \lambda_i$ včetně násobností a podle tvrzení 10.27 jsou všechna vlastní čísla kladná. Charakteristický polynom $p_{A_i}(t) = (\lambda_1 - t) \dots (\lambda_i - t)$ má podle tvrzení 9.24 absolutní člen rovný $\det(A_i)$. Roznásobením výrazu pro $p_{A_i}(t)$ ale také vidíme, že absolutní člen je rovný $\lambda_1 \lambda_2 \dots \lambda_i > 0$, takže $\det(A_i) > 0$.

(2) ⇒ (3). Indukcí podle i dokážeme, že před eliminací i-tého sloupce jsou všechny pivoty (prvky na místech $(1,1)$, …, (i,i)) kladné (speciálně, Gaussova eliminace bude používat pouze úpravy typu přičtení násobku řádku k jinému řádku). Pro $i = 1$ není co dokazovat, předpokládejme, že tvrzení platí pro $i - 1$. Před eliminací i-tého sloupce má matice tvar

$$B = \begin{pmatrix} X & Y \\ 0 & Z \end{pmatrix} \ ,$$

kde X je horní trojúhelníková matice řádu $i - 1$ s kladnými prvky na diagonále. Všechny dosud použité úpravy byly typu přičtení násobku řádku k jinému. Takové úpravy nemění determinant žádného minoru, pro i-tý minor B_i matice B tedy platí

$$\det(B_i) = x_{11} \dots x_{i-1,i-1} z_{11} = \det(A_i) > 0 \ .$$

Z toho vyplývá, že $z_{11} > 0$, takže pivot před eliminací i-tého sloupce bude skutečně kladný.

Implikace (3) ⇒ (4) je důsledkem v tvrzení 11.23.

(4) $\Rightarrow$ (5). Je-li $A = LDL^T$, kde $D = \mathrm{diag}(d_1, \dots, d_n)$, $d_1, \dots, d_n > 0$, pak položíme $R = L\sqrt{D}$, kde $\sqrt{D} = \mathrm{diag}(\sqrt{d_1}, \dots, \sqrt{d_n})$. Matice R je regulární a dolní trojúhelníková, protože je součinem dvou regulárních dolních trojúhelníkových matic, a platí

$$RR^T = (L\sqrt{D})(L\sqrt{D})^T = L\sqrt{D}\sqrt{D}^T L^T = L(\sqrt{D}\sqrt{D})L^T = LDL^T = A \ .$$

(5) $\Rightarrow$ (1). Pro libovolný nenulový vektor $\mathbf{x} \in \mathbb{R}^n$ platí $R^T\mathbf{x} \neq \mathbf{o}$, protože R^T je regulární. Potom

$$\mathbf{x}^T A\mathbf{x} = \mathbf{x}^T RR^T\mathbf{x} = (R^T\mathbf{x})^T R^T\mathbf{x} = \left\|R^T\mathbf{x}\right\|^2 > 0 \ .$$

□

11.4.3. *Ortonormální diagonalizace.* Pro geometrické aplikace se hodí najít f-ortogonální bázi symetrické bilineární formy, která je navíc ortonormální vzhledem k nějakému skalárnímu součinu. Takovou bázi můžeme vždy najít (ale nemůžeme vyžadovat, aby koeficienty u kvadratických členů byly z množiny $\{-1, 0, 1\}$).

Tvrzení 11.33. *Nechť* $\mathbf{V}$ *je reálný vektorový prostor dimenze* n *se skalárním součinem* $\langle\, , \rangle$ *a* f *je symetrická bilineární forma na* $\mathbf{V}$. *Pak existuje báze* B *prostoru* $\mathbf{V}$, *která je* f*-ortogonální a zároveň ortonormální vzhledem k* $\langle\, , \rangle$.

Důkaz. Pro skalární součin $\langle\, , \rangle$ existuje podle věty 8.51 ortonormální báze C prostoru $\mathbf{V}$. Označme $A = [f]_C$. V kapitole o unitární diagonalizaci jsme se dozvěděli, že existuje ortonormální báze $(\mathbf{u}_1, \dots, \mathbf{u}_n)$ prostoru $\mathbb{R}^n$ (ortonormalita je zde vzhledem ke standardnímu skalárnímu součinu!) složená z vlastních vektorů matice A. Maticově napsáno, označíme-li $U = (\mathbf{u}_1|\dots|\mathbf{u}_n)$, je U ortogonální matice a $U^{-1}AU = U^TAU = D$ je diagonální. Vezmeme $B = (\mathbf{v}_1, \dots, \mathbf{v}_n)$, aby $[\mathbf{v}_i]_C = \mathbf{u}_i$, tj. báze B je zvolená tak, že U je matice přechodu od B k C. Podle tvrzení 11.12 o změně báze je matice f vzhledem k B rovná $U^TAU = D$, takže B je f-ortogonální báze. Protože vyjádření vektorů $\mathbf{v}_1, \dots, \mathbf{v}_n$ v bázi C tvoří ortonormální bázi vzhledem ke standardnímu skalárnímu součinu a C je ortonormální báze vzhledem k $\langle\, , \rangle$, dostáváme, že $\mathbf{v}_1, \dots, \mathbf{v}_n$ tvoří ortonormální bázi vzhledem ke skalárnímu součinu $\langle\, , \rangle$ (viz tvrzení 8.40). □

Z tvrzení vyplývá, že jsou-li f, g dvě symetrické bilineární formy na reálném konečně generovaném prostoru $\mathbf{V}$, **z nichž alespoň jedna je pozitivně definitní**, pak existuje báze B, která je zároveň f-ortogonální a g-ortogonální. To obecně neplatí, vynecháme-li zvýrazněný požadavek, že alespoň jedna z forem je pozitivně definitní, viz cvičení.

11.5. **Příklady.** Podíváme se na aplikace nabytých poznatků na určení tvaru „kvadratického útvaru“.

Příklad 11.34. Podíváme se na množinu bodů $(x_1, x_2, x_3)^T \in \mathbb{R}^3$ splňujících $x_3 = -x_1^2 + x_1x_2 - 3x_2^2$. Je to graf kvadratické formy $f_2((x_1, x_2)^T) = -x_1^2 + x_1x_2 - 3x_2^2$. Příslušná symetrická bilineární forma f na $\mathbb{R}^2$ je

$$f((x_1, x_2)^T, (y_1, y_2)^T) = -x_1y_1 + 1/2x_1y_2 + 1/2x_2y_1 - 3x_2^2$$

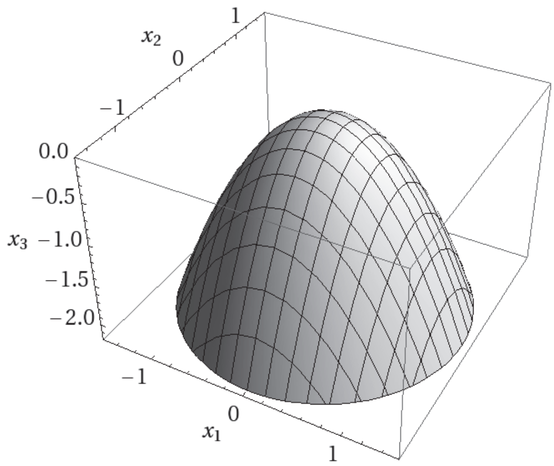
a její matice vzhledem ke kanonické bázi je

$$[f]_{K_2} = A = \begin{pmatrix} -1 & 1/2 \\ 1/2 & -3 \end{pmatrix} \sim \begin{pmatrix} -1 & 0 \\ 0 & -3 + (1/4) \end{pmatrix} \ .$$

Signatura je tedy $(0, 0, 2)$. Analytické vyjádření f_2 vzhledem k jisté bázi B je proto

$$f_2((x_1, x_2)^T) = -(x_1')^2 - (x_2')^2, \ \text{kde } [(x_1, x_2)^T]_B = (x_1', x_2')^T.$$

Grafem $x_3 = -x_1^2 - x_2^2$ je rotační paraboloid otevřený směrem dolů (viz obrázek). Tak vypadá graf vzhledem k bázi B. To nám dává představu, jak vypadá původní útvar – jde o „lineárně zdeformovaný“ rotační paraboloid. Ve skutečnosti je to eliptický paraboloid (ale není to zřejmé).



Abychom přesněji určili tvar útvaru, museli bychom najít B-ortogonální bázi, která je zároveň ortonormální vzhledem ke standardnímu skalárnímu součinu.

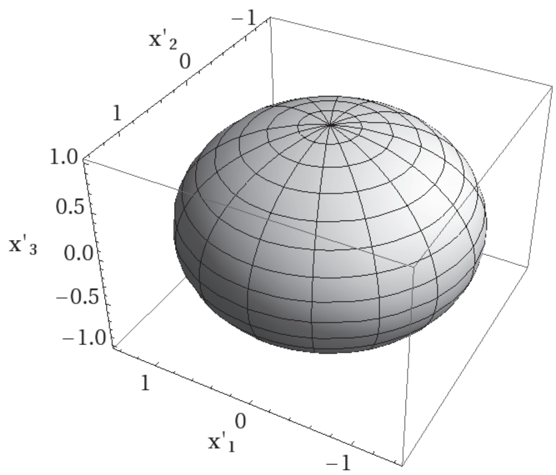
Příklad 11.35. Uvažujme množinu bodů $(x_1, x_2, x_3)^T \in \mathbb{R}^3$ splňujících

$$10x_1^2 + 13x_2^2 + 13x_3^2 + 4x_1x_2 + 4x_1x_3 + 8x_2x_3 = 9 \ .$$

Levá strana je kvadratická forma f_2 na $\mathbb{R}^3$. Příslušná symetrická bilineární forma f má matici

$$[f]_{K_3} = A = \begin{pmatrix} 10 & 2 & 2 \\ 2 & 13 & 4 \\ 2 & 4 & 13 \end{pmatrix}$$

Signatura f je $(0, 3, 0)$. Vzhledem k jisté bázi B má tedy útvar rovnici $(x_1')^2 + (x_2')^2 + (x_3')^2 = 9$, takže jde o „lineárně zdeformovanou" sféru.



OBRAZEK

Ve skutečnosti jde o elipsoid, ale opět to není zřejmé. Abychom určili útvar přesněji, najdeme ortonormální bázi (vzhledem ke standardnímu skalárnímu součinu), která je zároveň f-ortogonální. Jako ortonormální bázi C v tvrzení 11.33 zvolíme kanonickou, tj.

$$[f]_{K_3} = A = \begin{pmatrix} 10 & 2 & 2 \\ 2 & 13 & 4 \\ 2 & 4 & 13 \end{pmatrix}$$

Najdeme ortonormální bázi složenou z vlastních vektorů. Vlastní čísla vyjdou $\lambda_1 = \lambda_2 = 9$ (dvojnásobné) a $\lambda_3 = 18$. V příslušných podprostorech vybereme ortonormální bázi, v M_9 je to např. $(\mathbf{v}_1, \mathbf{v}_2)$ a v M_{18} $(\mathbf{v}_3)$.

$$(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3) = \left(\frac{1}{3}\begin{pmatrix} 2 \\ -2 \\ 1 \end{pmatrix}, \frac{1}{3}\begin{pmatrix} 2 \\ 1 \\ -2 \end{pmatrix}, \frac{1}{3}\begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix} \right)$$

Matice f vzhledem k (ortonormální) bázi $B = (\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ je $[f]_B = \mathrm{diag}(9, 9, 18)$, takže vzhledem k B je rovnice našeho útvaru

$$9(x_1')^2 + 9(x_2')^2 + 18(x_3')^2 = 9$$

a po drobné úpravě

$$\frac{(x_1')^2}{1} + \frac{(x_2')^2}{1} + \frac{(x_3')^2}{\left(\frac{1}{\sqrt{2}}\right)^2} = 1 \ .$$

Vidíme, že jde o elipsoid s poloosami $\mathbf{v}_1, \mathbf{v}_2, \frac{1}{\sqrt{2}}\mathbf{v}_3$, viz obrázek.

OBRAZEK - skutecny (tj. otoceny) elipsoid

Příklad 11.36. Budeme analyzovat následující útvar v $\mathbb{R}^2$:

$$U = \{(x_1, x_2)^T \in \mathbb{R}^2 : 3x_1^2 + 2x_1x_2 + 3x_2^2 - 10x_1 - 14x_2 + 7 = 0\}$$

Výraz z definice je součtem kvadratické formy $f_2((x_1, x_2)^T) = 3x_1^2 + 2x_1x_2 + 3x_2^2$, lineární formy $h((x_1, x_2)^T) = -10x_1 - 14x_2$ a konstanty 7.

Najdeme nejprve ortonormální f-ortogonální bázi $\mathbb{R}^2$, kde f je symetrická bilineární příslušná f_2:

$$[f]_{K_2} = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}$$

Vlastní čísla matice $[f]_{K_2}$ jsou 2 a 4 a příslušné znormované vlastní vektory jsou $\frac{\sqrt{2}}{2}(1, -1)$ a $\frac{\sqrt{2}}{2}(1, 1)$. Hledaná báze B a je tedy

$$B = (\mathbf{v}_1, \mathbf{v}_2) = \left(\frac{\sqrt{2}}{2}\begin{pmatrix} 1 \\ -1 \end{pmatrix}, \frac{\sqrt{2}}{2}\begin{pmatrix} 1 \\ 1 \end{pmatrix} \right) \ .$$

Vyjádříme útvar U v bázi B. Matice f vzhledem k bázi B je $\mathrm{diag}(2, 4)$, matice lineární formy h vzhledem k B je

$$[h]_{K_1}^{B} = [h]_{K_1}^{K_2}[\mathrm{id}]_{K_2}^{B} = (-10, -14)\frac{\sqrt{2}}{2}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} = \sqrt{2}(2, -12) \ ,$$

takže U má vzhledem k B vyjádření

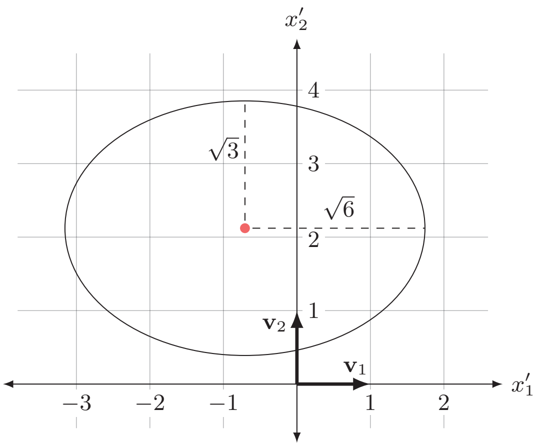
$$[U]_B = \{(x_1', x_2')^T \in \mathbb{R}^2 : 2(x_1')^2 + 4(x_2')^2 + 2\sqrt{2}x_1' - 12\sqrt{2}x_2' + 7 = 0\} \ .$$

Doplněním na čtverce a drobnými úpravami získáme

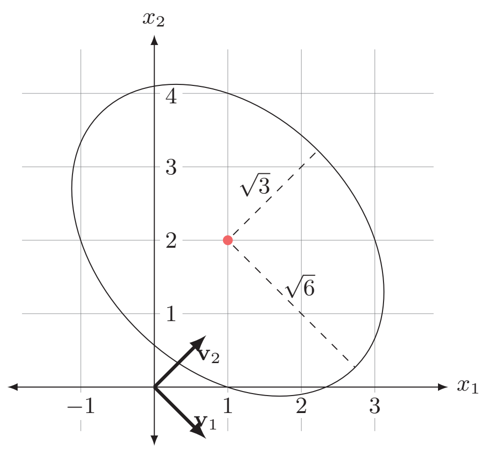
$$[U]_B = \left\{ (x_1', x_2')^T \in \mathbb{R}^2 : 2\left(x_1' + \frac{\sqrt{2}}{2}\right)^2 + 4\left(x_2' - \frac{3\sqrt{2}}{2}\right)^2 = 12 \right\}$$

$$= \left\{ (x_1', x_2')^T \in \mathbb{R}^2 : \left(\frac{x_1' + \frac{\sqrt{2}}{2}}{\sqrt{6}}\right)^2 + \left(\frac{x_2' - \frac{3\sqrt{2}}{2}}{\sqrt{3}}\right)^2 = 1 \right\}$$

Z toho vidíme, že vzhledem k B je útvar elipsa se středem $(-\frac{\sqrt{2}}{2}, \frac{3\sqrt{2}}{2})^T$ a velikostmi poloos $\sqrt{6}$ a $\sqrt{3}$.



Přepočteme střed do původních souřadnic: $-\frac{\sqrt{2}}{2}\mathbf{v}_1 + \frac{3\sqrt{2}}{2}\mathbf{v}_2 = (1,2)^T$. Vidíme, že U je elipsa se středem v bodě $(1,2)^T$, hlavní poloosou ve směru $\langle \mathbf{v}_1 \rangle$ a velikostí $\sqrt{6}$ a vedlejší poloosou ve směru $\langle \mathbf{v}_2 \rangle$ a velikostí $\sqrt{3}$.



Shrnutí jedenácté kapitoly

(1) Je-li $\mathbf{V}$ lineární prostor nad tělesem $\mathbf{T}$, pak *bilineární forma* na prostoru $\mathbf{V}$ je zobrazení $f : V \times V \to T$, které je lineární v obou složkách, tj. pro libovolné $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w} \in V$ a každý skalár $t \in T$ platí
  (1) $f(\mathbf{u}+\mathbf{v},\mathbf{w}) = f(\mathbf{u},\mathbf{w}) + f(\mathbf{v},\mathbf{w})$, $f(\mathbf{w},\mathbf{u}+\mathbf{v}) = f(\mathbf{w},\mathbf{u}) + f(\mathbf{w},\mathbf{v})$ a
  (2) $f(t\mathbf{v},\mathbf{w}) = t\,f(\mathbf{v},\mathbf{w})$, $f(\mathbf{v},t\mathbf{w}) = tf(\mathbf{v},\mathbf{w})$.

(2) Skalární součin na reálném lineárním prostoru je bilineární forma, skalární součin na komplexním lineárním prostoru není bilineární forma.

(3) Je-li f bilineární forma na lineárním prostoru $\mathbf{V}$ nad tělesem $\mathbf{T}$, pak zobrazení $f_2 : V \to T$ definované předpisem
$$f_2(\mathbf{v}) = f(\mathbf{v},\mathbf{v}) \quad \text{pro každé } \mathbf{v} \in V$$
nazýváme *kvadratickou formou* vytvořenou bilineární formou f.
Rovněž říkáme, že f_2 je kvadratická forma příslušná bilineární formě f.

(4) Je-li $B = (\mathbf{v}_1, \dots, \mathbf{v}_n)$ báze lineárního prostoru $\mathbf{V}$ nad tělesem $\mathbf{T}$ a f bilineární forma na $\mathbf{V}$, pak *maticí bilineární formy f vzhledem k B* rozumíme čtvercovou matici řádu n nad $\mathbf{T}$, která má na pozici (i,j) prvek $f(\mathbf{v}_i,\mathbf{v}_j)$. Tuto matici značíme $[f]_B$.

(5) Je-li B báze konečně generovaného lineárního prostoru $\mathbf{V}$ a $\mathbf{x},\mathbf{y} \in V$, pak
$$f(\mathbf{x},\mathbf{y}) = [\mathbf{x}]_B^T [f]_B [\mathbf{y}]_B \ .$$

(6) Jsou-li souřadnice vektorů $[\mathbf{x}]_B = (x_1, \dots, x_n)^T$, $[\mathbf{y}]_B = (y_1, \dots, y_n)^T$ a $[f]_B = (a_{ij})_{n\times n}$, pak
$$f(\mathbf{x},\mathbf{y}) = \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i y_j \ .$$
Tomuto vyjádření také říkáme *analytické vyjádření bilineární formy f*.

(7) Je-li $\mathbf{V}$ konečně generovaný lineární prostor nad tělesem $\mathbf{T}$, $B = (\mathbf{v}_1, \dots, \mathbf{v}_n)$ jeho báze a A čtvercová matice řádu n nad $\mathbf{T}$, pak zobrazení $f : V \times V \to T$ definované vztahem
$$f(\mathbf{x},\mathbf{y}) = [\mathbf{x}]_B^T A [\mathbf{y}]_B \quad \text{pro každé } \mathbf{x},\mathbf{y} \in V$$
je bilineární forma na $\mathbf{V}$ a platí $[f]_B = A$.

(8) Je-li f bilineární forma na lineárním prostoru $\mathbf{V}$, jsou-li B a C báze $\mathbf{V}$ a $X = [\mathrm{id}]_B^C$ matice přechodu od C k B, pak $[f]_C = X^T [f]_B X$.

(9) Bilineární forma f na lineárním prostoru $\mathbf{V}$ se nazývá
  - *symetrická*, pokud pro libovolné $\mathbf{x},\mathbf{y} \in V$ platí $f(\mathbf{x},\mathbf{y}) = f(\mathbf{y},\mathbf{x})$;
  - *antisymetrická*, pokud pro libovolné $\mathbf{x},\mathbf{y} \in V$ platí $f(\mathbf{x},\mathbf{y}) = -f(\mathbf{y},\mathbf{x})$.

(10) Je-li $\mathbf{V}$ konečně generovaný lineární prostor, B báze $\mathbf{V}$ a f bilineární forma na $\mathbf{V}$, pak
  - f je symetrická právě tehdy, když je $[f]_B$ symetrická matice;
  - f je antisymetrická právě tehdy, když je $[f]_B$ antisymetrická matice.

(11) Je-li $\mathbf{V}$ lineární prostor nad tělesem $\mathbf{T}$ charakteristiky různé od 2, pak každou bilineární formu f na $\mathbf{V}$ lze psát jako součet $f = f_s + f_a$, kde f_s je symetrická a f_a je antisymetrická. Tento rozklad je jednoznačný a platí
$$f_s(\mathbf{x},\mathbf{y}) = \frac{1}{2}(f(\mathbf{x},\mathbf{y}) + f(\mathbf{y},\mathbf{x})), \quad f_a(\mathbf{x},\mathbf{y}) = \frac{1}{2}(f(\mathbf{x},\mathbf{y}) - f(\mathbf{y},\mathbf{x})) \ .$$

(12) Jsou-li f,g bilineární formy na lineárním prostoru $\mathbf{V}$ nad tělesem charakteristiky různé od 2, pak $f_2 = g_2$ právě tehdy, když $f_s = g_s$. Navíc
$$f_s(\mathbf{x},\mathbf{y}) = \frac{1}{2}(f_2(\mathbf{x}+\mathbf{y}) - f_2(\mathbf{x}) - f_2(\mathbf{y})) \ .$$

(13) Je-li f symetrická bilineární forma na $\mathbf{V}$ a $\mathbf{x}, \mathbf{y} \in V$, pak říkáme, že $\mathbf{x}$ a $\mathbf{y}$ jsou f-*ortogonální*, pokud $f(\mathbf{x}, \mathbf{y}) = 0$. Zapisujeme $\mathbf{x} \perp_f \mathbf{y}$.

Báze $B = (\mathbf{v}_1, \dots \mathbf{v}_n)$ prostoru $\mathbf{V}$ se nazývá f-*ortogonální*, pokud je $[f]_B$ diagonální, tj. pro libovolné $i, j \in \{1, 2, \dots, n\}$, $i \neq j$, jsou vektory $\mathbf{v}_i, \mathbf{v}_j$ f-ortogonální.

(14) *Hodností* bilineární formy f na konečně generovaném lineárním prostoru $\mathbf{V}$ rozumíme hodnost její matice vzhledem k libovolné bázi, značíme $r(f)$.

(15) Každá symetrická bilineární forma f na konečně generovaném vektorovém prostoru nad tělesem charakteristiky různé od 2 má f-ortogonální bázi.

(16) Je-li A symetrická matice taková, že při Gaussově eliminaci nemusíme prohazovat řádky, pak existuje dolní trojúhelníková matice L s jedničkami na diagonále a diagonální matice D (složená z pivotů) tak, že

$$A = LDL^T \quad .$$

(17) Je-li f symetrická bilineární forma na reálném lineárním prostoru $\mathbf{V}$ dimenze n a C, C' báze $\mathbf{V}$ takové, že

$$[f]_C = \mathrm{diag}(\underbrace{1, 1, \dots, 1}_{k\times}, \underbrace{-1, -1, \dots, -1}_{l\times}, \underbrace{0, 0, \dots, 0}_{m\times})$$
$$[f]_{C'} = \mathrm{diag}(\underbrace{1, 1, \dots, 1}_{k'\times}, \underbrace{-1, -1, \dots, -1}_{l'\times}, \underbrace{0, 0, \dots, 0}_{m'\times}),$$

pak $k = k', l = l', m = m'$.

(18) Je-li f symetrická bilineární forma na reálném konečně generovaném lineárním prostoru $\mathbf{V}$, pak číslo k (resp. l) z předchozí věty nazýváme *pozitivní (resp. negativní) index setrvačnosti formy* f, značíme $n_+(f)$ (resp. $n_-(f)$). *Signaturou formy* f rozumíme trojici $(n_0(f), n_+(f), n_-(f))$.

(19) Symetrická bilineární forma f na reálném lineárním prostoru $\mathbf{V}$ se nazývá *pozitivně definitní*, pokud $f_2(\mathbf{x}) > 0$ pro libovolný nenulový prvek $\mathbf{x} \in V$.

(20) Symetrická bilineární forma f na reálném lineárním prostoru $\mathbf{V}$ dimenze n je pozitivně definitní právě tehdy, když $n_+(f) = n$.

(21) Symetrická bilineární forma f na reálném konečně generovaném prostoru $\mathbf{V}$ je pozitivně definitní právě tehdy, když je pozitivně definitní její matice vzhledem k libovolné bázi B.

(22) Pro reálnou symetrickou matici A řádu n jsou následující podmínky ekvivalentní
  (a) A je pozitivně definitní,
  (b) (Sylvestrovo kritérium) v šechny hlavní minory matice A mají kladný determinant,
  (c) Gaussova eliminace použitá na matici A může proběhnout bez prohazování řádků a všechny pivoty vyjdou kladné,
  (d) $A = LDL^T$ pro nějakou dolní trojúhelníkovou matici L s jedničkami na diagonále a nějakou diagonální matici D s kladnými čísly na diagonále,
  (e) (Choleského rozklad) $A = RR^T$ pro nějakou regulární dolní trojúhelníkovou matici R.

(23) Je-li $\mathbf{V}$ reálný vektorový prostor dimenze n se skalárním součinem $\langle\,,\rangle$ a f symetrická bilineární forma na $\mathbf{V}$, pak existuje báze B prostoru $\mathbf{V}$, která je f-ortogonální a zároveň ortonormální vzhledem k $\langle\,,\rangle$.

Klíčové znalosti z jedenácté kapitoly nezbytné pro průběžné sledování přednášek s pochopením

(1) Definice bilineární formy a její matice vzhledem k bázi.
(2) Definice kvadratické formy vytvořené bilineární formou.
(3) Jak se změní matice bilineární formy změní-li se báze.

(4) Rozklad bilineární formy na součet symetrické a antisymetrické formy.
(5) Věta o setrvačnosti symetrických bilineárních forem, signatura symetrické bilineární formy.
(6) Pozitivně definitní bilineární formy.
(7) Různé ekvivalentní definice pozitivně definitních matic.
(8) Věta o ortonormální diagonalizaci symetrických bilineárních forem na lineárním prostoru se skalárním součinem.

12. Afinní prostory

Cíl. *Až dosud byl pro nás základní pojem lineárního prostoru. V případě aritmetického vektorového prostoru $\mathbb{R}^3$ jsme nějaký aritmetický vektor $\mathbf{x} = (x_1, x_2, x_3)^T$ geometricky interpretovali podle potřeby buď jako bod o souřadnicích (x_1, x_2, x_3) ve třídimenzionálním prostoru s nějakým systémem souřadnic nebo jako (geometrický) vektor se souřadnicemi (x_1, x_2, x_3). V této kapitole se budeme zabývat více geometrií roviny a prostoru.*

V této kapitole se začneme blíže zaobírat geometrií. Zkoumanými objekty jsou množiny *bodů*, například množina bodů v prostoru, a množiny *vektorů*. Vektory si představujeme jako „šipky“ určené dvěma body, přičemž dva vektory považujeme za stejné, pokud se liší jenom umístěním. S vektory můžeme provádět známé operace sčítání a násobením skalárem. Další přirozenou geometrickou operací je přičtení bodu a vektoru. To provedeme umístěním počátku vektoru do daného bodu, výsledkem je koncový bod.

OBRAZEK (pricteni bodu a vektoru)

Tento pohled je přirozenější lidskému vnímání. Prostor se skládá z bodů a bod je tedy základním objektem, vektor je pojem odvozený. Doposud jsme tento nedostatek řešili tak, že jsme si v prostoru zvolili počátek a vektory umísťovali do počátku. Bod jsme pak ztotožňovali s jeho polohovým vektorem. Tento pohled má několik nedostatků. Jedním z nich je, že prostor nemá apriori žádný význačný bod, takže volba nějakého počátku je nepřirozená. Podstatnější nevýhoda vynikne, když si připomeneme, že lineární algebru lze chápat jako studium „rovných“ útvarů (přímky, roviny, atd.) a „rovných“ zobrazení mezi nimi. Odpovídající objekty ve vektorových prostorech jsou podprostory a lineární zobrazení. Podprostory ale nepopisují všechny rovné útvary, pouze rovné útvary *procházející počátkem*, i když jiné rovné útvary se přirozeně objevily, například jako množiny řešení nehomogenní soustavy rovnic. Podobně, lineární zobrazení popisují jen rovná zobrazení *zachovávající počátek*, tedy například žádné posunutí o nenulový vektor nebylo objektem studia.

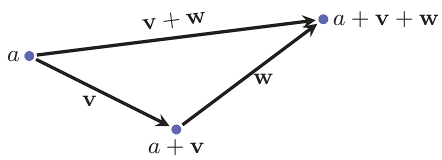
Nyní tedy začneme rozlišovat body a vektory. V další kapitole pak nahlédneme, že body a vektory lze vlastně chápat jako různé instance stejného geometrického objektu, a tím se poněkud paradoxně vrátíme ke studiu rovných útvarů pouze pomocí vektorů. Tento pohled nám přinese řadu výhod.

V celé kapitole budeme pracovat výhradně s prostory konečné dimenze, které jsou bližší geometrickému náhledu. Řada pojmů a tvrzení se přirozeně přenáší na prostory, které nejsou konečně generované.

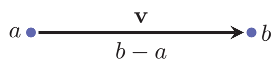
12.1. **Definice afinního prostoru.** Jak jsme předeslali v úvodu, afinní prostor je tvořen množinou bodů a množinou vektorů. Na množině vektorů máme operace sčítání a násobení skalárem, které mají všechny doposud používané vlastnosti, tedy množina vektorů tvoří spolu s těmito operacemi vektorový prostor. Přibude operace sčítání bodu a vektoru. Požadované axiomy jsou opět ve shodě s geometrickou představou.

Definice 12.1. Nechť $\mathbf{T}$ je těleso. Afinním prostorem $\mathbf{A}$ nad $\mathbf{T}$ rozumíme množinu A, jejíž prvky nazýváme *body*, spolu s vektorovým prostorem $\mathbf{V}$ nad $\mathbf{T}$ a operací $+ : A \times V \to A$, která bodu $a \in A$ a vektoru $\mathbf{v} \in V$ přiřadí bod $a + \mathbf{v} \in A$, splňující axiomy:

- (aS2) Pro libovolný bod $a \in A$ a libovolné vektory $\mathbf{v}, \mathbf{w} \in V$ platí $a + (\mathbf{v} + \mathbf{w}) = (a + \mathbf{v}) + \mathbf{w}$.
- (aS1) Pro libovolný bod $a \in A$ platí $a + \mathbf{o} = a$.
- (aM) Ke každé dvojici bodů $a, b \in A$ existuje právě jeden vektor $\mathbf{v} \in V$, pro který $a + \mathbf{v} = b$. Tento vektor značíme $b - a$.



Axiom (aS2)



Axiom (aM)

Sčítat můžeme dva vektory a bod s vektorem. Sčítání dvou bodů nedává (zatím) žádný geometrický smysl. Pro body budeme používat stejně jako v definici malá písmena abecedy. Z axiomu (aS2) vidíme, že ve výrazech tvaru $a + \mathbf{v}_1 + \mathbf{v}_2 + \cdots + \mathbf{v}_n$ nemusíme psát závorky. Při popisu afinního prostoru $\mathbf{A}$ budeme většinou zdůrazňovat jen množinu bodů A s tím, že vektorový prostor a sčítání je zřejmé z kontextu. Vektorový prostor $\mathbf{V}$ budeme někdy nazývat *prostor vektorů* afinního prostoru $\mathbf{A}$.

Pokud v afinním prostoru zvolíme nějaký bod $a \in A$, pak každému bodu $b \in A$ můžeme podle (aM) přiřadit vektor $b - a$ a naopak, každému vektoru $\mathbf{v}$ můžeme přiřadit bod $a + \mathbf{v}$. Jak se snadno ověří (cvičení), tato zobrazení jsou navzájem inverzní bijekce bodů a vektorů (bijekce nejsou přirozené, závisí na volbě bodu a). V tomto smyslu si body a vektory vzájemně jednoznačně odpovídají, proto například dává smysl mluvit o dimenzi afinního prostoru.

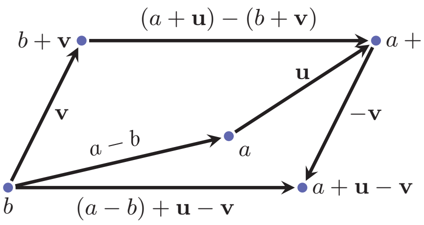
Definice 12.2. *Dimenzí* afinního prostoru $\mathbf{A}$ rozumíme dimenzi jeho prostoru vektorů.

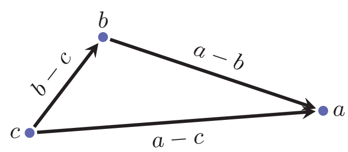
Afinní prostor dimenze 0 tvoří jediný bod $A = \{a\}$. Afinní prostor dimenze 1 nazýváme *afinní přímka*, nebo jen *přímka*, afinní prostor dimenze 2 nazýváme *afinní rovina*, nebo jen *rovina*.

Mechanickým cvičením jsou následující vlastnosti operací, které platí pro libovolné body $a, b, c, d \in A$ a vektory $\mathbf{u}, \mathbf{v} \in V$. Geometrický význam je jasný z obrázku.

- $a - b = -(b - a)$
- $(a + \mathbf{u}) - (b + \mathbf{v}) = (a - b) + \mathbf{u} - \mathbf{v}$
- $(a - b) + (c - d) = (a - d) + (c - b)$
- $(a - b) + (b - c) = a - c$

Tyto a podobné vlastnosti budou podrobněji diskutovány v části o lineárních kombinacích bodů.





Příklady. Pro libovolný vektorový prostor $\mathbf{V}$ tvoří $A = V$ spolu se sčítáním ve $\mathbf{V}$ afinní prostor. Množiny bodů a vektorů jsou tedy stejné, rozdíl je jen v pohledu – na prvky A se díváme jako na body, na prvky V jako na vektory. Rozdílný bude také například pojem podprostoru, jak jsme diskutovali v úvodu. Speciálně pro $\mathbf{V} = \mathbf{T}^n$ dostáváme *aritmetický afinní prostor*. Budeme jej značit stejně jako aritmetický vektorový prostor, tj. $\mathbf{T}^n$, jeho dimenze je n.

Trochu jiným příkladem je

$$A = (1,2,3)^T + \langle (2,3,4)^T, (6,7,8)^T \rangle, \quad V = \langle (2,3,4)^T, (6,7,8)^T \rangle \ .$$

Vektorový prostor $\mathbf{V}$ je podprostor $\mathbb{R}^3$ generovaný vektory $(2,3,4)^T$ a $(6,7,8)^T$ a A je rovina v $\mathbb{R}^3$ se „směrem“ $\mathbf{V}$ procházející bodem $(1,2,3)^T$. (Sčítání bodu a vektoru probíhá po složkách.) V tomto případě A není vektorovým podprostorem $\mathbb{R}^3$. Bod v A můžeme sečíst s vektorem ve V, ale součet dvou bodů, pokud bychom ho počítali jako v $\mathbb{R}^3$, v A obecně neleží. Toto je příklad podprostoru afinního prostoru $\mathbb{R}^3$. Jeho dimenze je 2, je to afinní rovina.

Obecněji, pro libovolný afinní prostor $\mathbf{A}$ s prostorem směrů $\mathbf{V}$ je každá množina bodů tvaru $a + W$, kde $\mathbf{W} \leq \mathbf{V}$ se zděděnými operacemi afinní prostor, jehož prostor směrů je $\mathbf{W}$. Tento prostor je podprostorem $\mathbf{A}$. Takové podprostory aritmetických prostorů vznikají například při řešení soustavy lineárních rovnic. Podrobněji se podprostory budeme zabývat zanedlouho, zatím jsme ani přesně nepopsali, co je podprostor. Vystačíme s intuitivní představou.

Chceme-li ještě pracovat s metrickými vlastnostmi, jako velikosti vektorů, vzdálenosti bodů, atd., potřebujeme na $\mathbf{V}$ mít ještě dán skalární součin. V tomto případě musí být $\mathbf{T} = \mathbb{R}$ nebo $\mathbf{T} = \mathbb{C}$.

Definice 12.3. *Afinním eukleidovským prostorem* (resp. *afinním unitárním prostorem*) rozumíme afinní prostor $\mathbf{A}$ nad tělesem $\mathbb{R}$ (resp. $\mathbb{C}$) spolu se skalárním součinem $\langle\, , \rangle$ na jeho prostoru vektorů.

Nejjednodušším příkladem afinního eukleidovského prostoru je $\mathbb{R}^n$ se standardním skalárním součinem. Nejjednodušším příkladem afinního unitárního prostoru je $\mathbb{C}^n$ se standardním skalárním součinem. V této kapitole budeme uvažovat pouze afinní prostory a afinní eukleidovské prostory. Přímočaré rozšíření na komplexní případ si čtenář může rozmyslet sám.

Již víme, co pro afinním eukleidovský prostor znamená velikost vektoru, úhel dvou vektorů, kolmost, apod. Vzdálenost bodů definujeme opět ve shodě s intuicí.

Definice 12.4. Vzdáleností dvou bodů $a, b \in A$ v afinním eukleidovském prostoru $\mathbf{A}$ rozumíme číslo $\|a - b\|$.

12.1.1. *Soustava souřadnic.* Na bázi vektorového prostoru lze nazírat jako na jeho soustavu souřadnic – zvolíme-li bázi, můžeme vektory vyjadřovat jako n-tice skalárů (prvky $\mathbf{T}^n$) a počítat s nimi jako v $\mathbf{T}^n$ (viz odstavec 5.4.3). Soustava souřadnic v afinním prostoru má podobnou roli. Sestává z bodu, tzv. počátku soustavy souřadnic, a n-tice vektorů, které si představujeme umístěné do počátku. Máme-li zadanou soustavu, můžeme přirozeným způsobem vyjadřovat body i vektory jako n-tice prvků tělesa a počítání pak probíhá jako v aritmetickém afinním prostoru $\mathbf{T}^n$.

Definice 12.5. *Soustavou souřadnic* v afinním prostoru $\mathbf{A}$ dimenze n s prostorem vektorů $\mathbf{V}$ rozumíme $(n+1)$-tici $S = (a, \mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n)$, kde $a \in A$ je bod nazývaný *počátek soustavy souřadnic* a $B = (\mathbf{u}_1, \ldots, \mathbf{u}_n)$ je báze $\mathbf{V}$.

Je-li S soustava souřadnic jako výše, $b \in A$ je bod a $\mathbf{w} \in V$ je vektor, pak *souřadnice vektoru* $\mathbf{w}$ *v soustavě souřadnic* S definujeme jako souřadnice $\mathbf{w}$ vzhledem k bázi B a značíme $[\mathbf{w}]_S$, tj.

$$[\mathbf{w}]_S = [\mathbf{w}]_B$$

a *souřadnice bodu* b *v soustavě souřadnic* S definujeme jako souřadnice vektoru $b - a$ v bázi B, tj.

$$[b]_S = [b-a]_S = [b-a]_B \ .$$

Souřadnice bodu jsou definovány ve shodě s geometrickou intuicí. To je možná ještě lépe vidět s následujícího přeformulování definice: Souřadnice bodu b v soustavě S je rovno té jednoznačně určené n-tici prvků $(t_1, \ldots, t_n) \in T^n$, pro kterou platí

$$b = a + t_1\mathbf{u}_1 + \cdots + t_n\mathbf{u}_n \ .$$

OBRAZEK

Souřadnice počátku a vzhledem k S jsou $[a]_S = (0, 0, \ldots, 0)^T$ a $[a + \mathbf{u}_i]_S = \mathbf{e}_i$.

Příklad 12.6. V aritmetickém afinním prostoru $\mathbb{R}^2$ je

$$S = (a, \mathbf{u}_1, \mathbf{u}_2) = \left(\begin{pmatrix} 3 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} -2 \\ -1 \end{pmatrix} \right)$$

soustava souřadnic, protože $(\mathbf{u}_1, \mathbf{u}_2)$ je bází aritmetického vektorového prostoru $\mathbb{R}^2$. Určíme souřadnice vektoru $\mathbf{w} = (-1,3)^T$ a bodu $b = (-1,3)^T$ v S. K tomu potřebujeme nalézt vyjádření vektoru $(-1,3)^T$ a vektoru $(-1,3)^T - (3,2)^T = (-4,1)^T$ v bázi $(\mathbf{u}_1, \mathbf{u}_2)$. To vede na řešení dvou soustav rovnic se stejnou maticí. Vyřešíme je současně.

$$\left(\begin{array}{cc|cc} 1 & -2 & -1 & -4 \\ 1 & -1 & 3 & 1 \end{array} \right) \sim \left(\begin{array}{cc|cc} 1 & -2 & -1 & -4 \\ 0 & 1 & 4 & 5 \end{array} \right)$$

Z toho dopočteme řešení

$$[\mathbf{w}]_S = \begin{pmatrix} 7 \\ 4 \end{pmatrix}, \quad [b]_S = \begin{pmatrix} 6 \\ 5 \end{pmatrix} \ .$$

Pro kontrolu můžeme ověřit, že skutečně $\mathbf{w} = 7\mathbf{u}_1 + 4\mathbf{u}_2$ a $b = a + 6\mathbf{u}_1 + 5\mathbf{u}_2$.

Příklad 12.7. V aritmetických afinních prostorech máme význačnou soustavu souřadnic, budeme jí nazývat *kanonická*:

$$S = ((0,0,\dots,0)^T, \mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n) \ .$$

Je charakterizovaná tím, že $[a]_S = a$ a $[\mathbf{w}]_S = \mathbf{w}$ pro libovolný bod a a libovolný vektor $\mathbf{w}$.

V afinním eukleidovském prostoru jsou „nejlepší“ soustavy souřadnic kartézské.

Definice 12.8. Soustava souřadnic $S = (a, \mathbf{u}_1, \dots, \mathbf{u}_n)$ v afinním eukleidovském prostoru se nazývá *kartézská*, pokud $(\mathbf{u}_1, \dots, \mathbf{u}_n)$ je ortonormální báze.

V kartézské soustavě souřadnic jsou tedy vektory $\mathbf{u}_1, \dots, \mathbf{u}_n$ jednotkové a navzájem kolmé. V aritmetickém afinním prostoru se standardním skalárním součinem (budeme mu říkat aritmetický afinní eukleidovský prostor) je kanonická soustava souřadnic kartézská.

Volba soustavy souřadnic převádí počítání v afinním prostoru na počítání v aritmetickém vektorovém prostoru, podobně jako báze pro vektorové prostory (viz tvrzení 5.69). Je-li prostor afinní eukleidovský, tak v kartézské soustavě souřadnic se skalární součin převádí na standardní (viz TODO).

Tvrzení 12.9. *Je-li S soustava souřadnic afinního prostoru $\mathbf{A}$ s prostorem vektorů $\mathbf{V}$ nad tělesem $\mathbf{T}$, pak pro libovolné $\mathbf{v}_1, \mathbf{v}_2 \in V$, $b, c \in A$, $t \in T$ platí*

$$[\mathbf{v}_1 + \mathbf{v}_2]_S = [\mathbf{v}_1]_S + [\mathbf{v}_2]_S, \quad [t\mathbf{v}_1]_S = t[\mathbf{v}_1]_S, \quad [b + \mathbf{v}_1]_S = [b]_S + [\mathbf{v}_1]_S, \quad [b - c]_S = [b]_S - [c]_S \ .$$

Je-li navíc $\mathbf{A}$ afinní eukleidovský prostor a soustava S je kartézská, pak

$$\langle \mathbf{v}_1, \mathbf{v}_2 \rangle = [\mathbf{v}_1]_S \cdot [\mathbf{v}_2]_S \ .$$

Důkaz. cviceni □

Nyní spočítáme, jak se změní souřadnice bodů a vektorů při změně soustavy souřadnic. Uvažujme dvě soustavy $S = (a, \mathbf{u}_1, \dots, \mathbf{u}_n)$ a $S' = (a', \mathbf{u}'_1, \dots, \mathbf{u}'_n)$. Označme X matici přechodu od báze $B = (\mathbf{u}_1, \dots, \mathbf{u}_n)$ k bázi $B' = (\mathbf{u}'_1, \dots, \mathbf{u}'_n)$. Přepočítávat souřadnice vektorů už umíme: pro libovolný vektor $\mathbf{v} \in V$ máme

$$[\mathbf{v}]_{S'} = X[\mathbf{v}]_S \ .$$

Pro bod $b \in A$ využijeme vztahu $b - a' = (b - a) + (a - a')$ a dostaneme

$$[b]_{S'} = [b - a']_{S'} = [b - a]_{S'} + [a - a']_{S'} = X[b - a]_S + [a - a']_{S'} = X[b]_S + [a]_{S'} \ .$$

Shrneme výsledek do tvrzení.

Tvrzení 12.10. *Nechť $S = (a, \mathbf{u}_1, \dots, \mathbf{u}_n)$ a $S' = (a', \mathbf{u}'_1, \dots, \mathbf{u}'_n)$ jsou soustavy souřadnic v afinním prostoru $\mathbf{A}$ s prostorem vektorů $\mathbf{V}$ a X je matice přechodu od $(\mathbf{u}_1, \dots, \mathbf{u}_n)$ k $(\mathbf{u}'_1, \dots, \mathbf{u}'_n)$. Pak pro každé $b \in A$, $\mathbf{v} \in V$ platí*

$$[\mathbf{v}]_{S'} = X[\mathbf{v}]_S, \quad [b]_{S'} = X[b]_S + [a]_{S'} \ .$$

Příklad 12.11. Ilustrujeme přechodové vztahy na soustavách souřadnic S, S' aritmetického afinního prostoru $\mathbb{R}^2$.

$$S = (a, \mathbf{u}_1, \mathbf{u}_2) = \left(\begin{pmatrix} -4 \\ 5 \end{pmatrix}, \begin{pmatrix} 5 \\ 3 \end{pmatrix}, \begin{pmatrix} -7 \\ 14 \end{pmatrix} \right) S' = (a', \mathbf{u}'_1, \mathbf{u}'_2) = \left(\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} -2 \\ 3 \end{pmatrix} \right),$$

Najdeme matici přechodu X od báze $B = (\mathbf{u}_1, \mathbf{u}_2)$ k bázi $B' = (\mathbf{u}'_1, \mathbf{u}'_2)$.

$$[\mathrm{id}]^B_{B'} = [\mathrm{id}]^{K_2}_{B'} [\mathrm{id}]^B_{K_2} = \begin{pmatrix} 1 & -2 \\ 2 & 3 \end{pmatrix}^{-1} \begin{pmatrix} 5 & -7 \\ 3 & 14 \end{pmatrix} = \frac{1}{7} \begin{pmatrix} 3 & 2 \\ -2 & 1 \end{pmatrix} \begin{pmatrix} 5 & -7 \\ 3 & 14 \end{pmatrix} = \begin{pmatrix} 3 & 1 \\ -1 & 4 \end{pmatrix}$$

Najdeme ještě $[a]_{S'} = [a - a']_{S'}$.

$$[a - a']_{S'} = \left[\begin{pmatrix} -5 \\ 4 \end{pmatrix} \right]_{S'} = [id]^{K_2}_{S'} \begin{pmatrix} -5 \\ 4 \end{pmatrix} = \frac{1}{7} \begin{pmatrix} 3 & 2 \\ -2 & 1 \end{pmatrix} \begin{pmatrix} -5 \\ 4 \end{pmatrix} = \begin{pmatrix} -1 \\ 2 \end{pmatrix}$$

Pro libovolný bod $b \in A$ nyní máme

$$[b]_{S'} = \begin{pmatrix} 3 & 1 \\ -1 & 4 \end{pmatrix} [b]_S + \begin{pmatrix} -1 \\ 2 \end{pmatrix}$$

Abychom ještě lépe viděli tvar přechodových vztahů, označíme $[b]_S = (x, y)^T$ a $[b]_{S'} = (x', y')^T$ a vztahy přepíšeme.

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} 3 & 1 \\ -1 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} -1 \\ 2 \end{pmatrix} = \begin{pmatrix} 3x + y - 1 \\ -x + 4y + 2 \end{pmatrix}$$

Nové souřadnice jsou tedy lineární výrazy ve starých souřadnicích (tj. výrazy tvaru lineární forma + konstanta). Pro vektory dostaneme stejné výrazy bez konstantních členů.

12.2. **Lineární kombinace bodů.** Tvořit „lineární kombinace" bodů nedává obecně žádný geometrický smysl, i když na některé smysluplné výrazy (např. vektor $b - a$ a bod $a + (b - a) = b$) lze nazírat jako na lineární kombinace.

Abychom nahlédli, že všem výrazům skutečně nelze dát v afinním prostoru geometrický smysl, podívejme se na výraz $a + b$, kde a, b jsou body nějakého afinního prostoru $\mathbf{A}$ s prostorem vektorů $\mathbf{V}$. Přirozenou myšlenkou je zvolit v $\mathbf{A}$ soustavu souřadnic S a definovat $a + b$ jako ten bod, jehož souřadnice vzhledem k S jsou $[a]_S + [b]_S$. Problém je, že výsledný bod závisí na volbě soustavy souřadnic. Například pro $\mathbf{A} = \mathbb{R}^2$, $a = (0, 0)^T$, $b = (1, 0)^T$ by vzhledem ke kanonické soustavě souřadnic vyšlo $a + b = (1, 0)^T$, ale vzhledem k soustavě souřadnic $S = ((2, 3)^T, (1, 0)^T, (0, -1)^T)$ bychom měli

$$[a]_S = \begin{pmatrix} -2 \\ 3 \end{pmatrix}, \quad [b]_S = \begin{pmatrix} -1 \\ 3 \end{pmatrix}, \quad [a + b]_S = \begin{pmatrix} -3 \\ 6 \end{pmatrix},$$

takže $a + b = (2, 3)^T + -3(1, 0)^T + 6(0, -1)^T = (-1, -3)^T$. Ještě by nás mohlo napadnout, že $a + b$ je nějaký vektor, ale ani v tom případě bychom neuspěli – našli bychom dvě soustavy souřadnic, ve které se výsledky liší.

12.2.1. *Afinní kombinace.* Některým lineárním kombinacím ale smysl lze dát. Pokud bychom například počítali $\frac{1}{2}a + \frac{1}{2}b$ stejným postupem vyšel by nám v obou případech stejný bod $(\frac{1}{2}, 0)^T$. Je to proto, že tento bod lze vyjádřit jako $a + \frac{1}{2}(b - a)$ $(= b + \frac{1}{2}(a - b))$ a tento výraz je definován – je to součet bodu a a $\frac{1}{2}$-násobku vektoru $b - a$. Geometricky, je to střed úsečky a, b. Následující tvrzení zodpovídá přesně na otázku, kdy lze definovat bod jako lineární kombinace bodů.

Tvrzení 12.12. *Nechť* $\mathbf{A}$ *je afinní prostor nad* $\mathbf{T}$ *dimenze alespoň* 1, $a_1, \dots, a_k \in A$ *body a* $\lambda_1, \dots, \lambda_k \in \mathbf{T}$ *skaláry. Pak následující tvrzení jsou ekvivalentní.*

(1) *Bod* b *o souřadnicích* $[b]_S = \lambda_1[a_1]_S + \dots + \lambda_k[a_k]_S$ *nezávisí na volbě soustavy souřadnic* S.
(2) $\lambda_1 + \dots + \lambda_k = 1$.

Důkaz. Snazší je dokázat implikaci (2) $\Rightarrow$ (1). Je-li $\lambda_1 + \dots + \lambda_k = 1$, stačí si uvědomit, že v libovolné soustavě souřadnic S díky podmínce této podmínce a tvrzení 12.9 o souřadnicích a operacích máme

$$\begin{aligned}\lambda_1[a_1]_S + \dots + \lambda_k[a_k]_S &= [a_1]_S + \lambda_2([a_2]_S - [a_1]_S) + \dots + \lambda_k([a_k]_S - [a_1]_S) \\ &= [a_1 + \lambda_2(a_2 - a_1) + \dots + \lambda_k(a_k - a_1)]_S\end{aligned}$$

Protože body jsou jednoznačně určené svými souřadnicemi, bod b v (1) je nutně roven (korektně definovanému) bodu $a_1 + \lambda_2(a_2 - a_1) + \dots + \lambda_k(a_k - a_1)$, který samozřejmě na S nezávisí.

(1) $\Rightarrow$ (2). □

To nám umožňuje zavést afinní kombinaci bodů.

Definice 12.13. Nechť $\mathbf{A}$ je afinní prostor nad $\mathbf{T}$, $a_1, \dots, a_k \in A$ body a $\lambda_1, \dots, \lambda_k \in \mathbf{T}$ skaláry takové, že $\lambda_1 + \dots + \lambda_k = 1$. *Afinní kombinací bodů* $a_1, \dots, a_k$ *s koeficienty* $\lambda_1, \dots, \lambda_k$ rozumíme bod $b \in A$ takový, že

$$[b]_S = \lambda_1[a_1]_S + \dots + \lambda_k[a_k]_S$$

kde S je libovolná soustava souřadnic prostoru $\mathbf{A}$. Značíme $b = \lambda_1 a_1 + \dots + \lambda_k a_k$.

Afinní kombinaci jsme zavedli pomocí (libovolně zvolené) soustavy souřadnic, přičemž definice dává smysl díky předchozímu tvrzení. Z důkazu tohoto tvrzení také plyne, že afinní kombinaci lze zavést bez volby soustavy, například vztahem

$$\lambda_1 a_1 + \dots + \lambda_k a_k = a_1 + \lambda_2(a_2 - a_1) + \dots + \lambda_k(a_k - a_1) \ .$$

Tento výraz je ale poněkud nesymetrický.

Alternativní, symetrická definice a geometrický význam asi nejlépe vynikne z fyzikálního pohledu (i když ten můžeme uplatnit pouze pro reálné afinní prostory malých dimenzí a pouze pro afinní kombinace s nezápornými koeficienty). Afinní kombinaci $\lambda_1 a_1 + \dots + \lambda_k a_k$ totiž můžeme chápat jako těžiště soustavy hmotných bodů $a_1, \dots, a_k$ s hmotnostmi $\lambda_1, \dots, \lambda_k$. To je lépe vidět z následující charakterizace.

Tvrzení 12.14. *Nechť* $\mathbf{A}$ *je afinní prostor nad* $\mathbf{T}$, $a_1, \dots, a_k \in A$ *body a* $\lambda_1, \dots, \lambda_k \in T$ *skaláry takové, že* $\lambda_1 + \dots + \lambda_k = 1$. *Pak bod* $\lambda_1 a_1 + \dots + \lambda_k a_k$ *je roven tomu jednoznačně určenému bodu* b, *pro který*

$$\lambda_1(a_1 - b) + \lambda_2(a_2 - b) + \dots + \lambda_k(a_k - b) = \mathbf{o} \ .$$

Důkaz. V $\mathbf{A}$ zvolíme libovolnou soustavu souřadnic S s počátkem $\lambda_1 a_1 + \dots + \lambda_k a_k$. Pak pro libovolný bod b jsou souřadnice vektoru na levé straně vzhledem k S rovny

$$\begin{aligned}[\lambda_1(a_1 - b) + \lambda_2(a_2 - b) + \dots + \lambda_k(a_k - b)]_S &= [\lambda_1 a_1 + \lambda_2 a_2 + \dots + \lambda_k a_k]_S \\ &- [\lambda_1 b + \dots + \lambda_k b]_S = -[b]_S\end{aligned}$$

(Používáme definici afinní kombinace a tvrzení 12.9 o počítání v souřadnicích.)

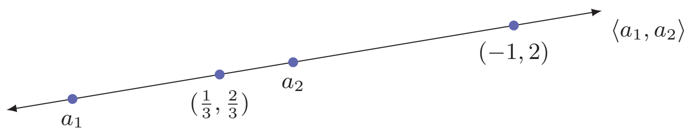
Vidíme, že vektor na levé straně je nulový právě tehdy, když $b = \lambda_1 a_1 + \dots + \lambda_k a_k$, což jsme měli dokázat. □

OBRAZEK (ruzne afin. kombinace dvou bodu, trojuhelnik, 4.bod v rovnobezniku)

12.2.2. *Barycentrické souřadnice.* Podíváme se blíže na afinní kombinace dvou bodů na afinní přímce. Mějme tedy afinní prostor $\mathbf{A}$ s prostorem vektorů $\mathbf{V}$ nad tělesem $\mathbf{T}$, kde $\dim \mathbf{A}(= \dim \mathbf{V}) = 1$. Konkrétně například $\mathbb{R}$ nebo podprostor $\mathbb{R}^2$ nebo $\mathbb{R}^3$ tvaru $A = c + \langle \mathbf{v} \rangle$, $\mathbf{v} \neq \mathbf{o}$.

Jsou-li $a, b \in A$ dva různé body, pak každý bod $c \in A$ lze vyjádřit právě jedním způsobem jako jejich afinní kombinace. Existenci takového vyjádření můžeme zdůvodnit například následujícím způsobem. Protože $b - a$ je nenulový vektor a $\dim \mathbf{V} = 1$, je každý vektor ve $\mathbf{V}$ jeho násobkem. Existuje proto $\lambda \in \mathbf{T}$ takové, že $c - a = \lambda(b - a)$. Nyní můžeme psát $c = a + \lambda(b - a) = (1 - \lambda)a + \lambda b$ (rovnost dokážeme například pomocí souřadnic a tvrzení 12.9). Jednoznačnost se nahlédne například z jednoznačnosti λ ve vyjádření $c - a = \lambda(b - a)$. Důkaz obecnějšího tvrzení provedeme za okamžik.

Bod $c = \lambda_1 a + \lambda_2 b$ „dělí“ body a, b v poměru $\lambda_2 : \lambda_1$. Přesněji, $\lambda_1(c - a) = \lambda_2(b - c)$. Pokud $\mathbf{A}$ je eukleidovský tak tento vztah znamená, že poměr „orientovaných vzdáleností“ c od a a c od b je $\lambda_2 : \lambda_1$, tj. v případě, že c leží na úsečce ab (ekvivalentně $\lambda_1, \lambda_2 \geq 0$) je poměr vzdáleností $\lambda_2 : \lambda_1$, v opačném případě je poměr vzdáleností $|\lambda_1| : |\lambda_2|$.



Obrázek 83. Souřadnice dvou bodů vzhledem k barycentrické soustavě souřadnic (a_1, a_2). Afinní obal $\langle a_1, a_2 \rangle$.

Příklad 12.15. Vyjádříme bod $c = (2, 3)^T \in \mathbb{R}^2$ jako afinní kombinaci bodů $a = (1, 2)^T$ a $b = (5, 6)^T$. Úloha dává smysl, protože všechny tři body leží na afinní přímce $(0, 1)^T + \langle (1, 1)^T \rangle$.

Srovnáním prvních složek ve vztahu $c = \lambda_1 a + \lambda_2 b$ získáme $\lambda_1 + 5\lambda_2 = 2$, což spolu s $\lambda_1 + \lambda_2 = 1$ dává $\lambda_1 = \frac{3}{4}, \lambda_2 = \frac{1}{4}$. Tedy $c = \frac{3}{4}a + \frac{1}{4}b$. Skutečně, bod c dělí body a, b v poměru $\frac{1}{4} : \frac{3}{4} = 1 : 3$. Fyzikální interpretace je taková, že má-li bod a hmotnost $\frac{3}{4}$ a bod b hmotnost $\frac{1}{4}$, pak je jejich těžištěm bod c.

Dvojice (λ_1, λ_2) tvoří tzv. *barycentrické souřadnice* bodu c vzhledem k (a, b). Vyjadřují, jakým způsobem musíme body a, b zatížit, aby jejich těžištěm byl bod c. Podobným způsobem lze definovat barycentrické souřadnice bodu v rovině vzhledem ke třem bodům neležících na jedné přímce, apod.

Tvrzení 12.16. *Nechť $\mathbf{A}$ je afinní prostor dimenze n s prostorem vektorů $\mathbf{V}$ a $a_1, \dots, a_k \in A$ jsou body. Pak následující tvrzení jsou ekvivalentní.*

(1) *Každý bod $b \in A$ lze jednoznačným způsobem zapsat jako afinní kombinaci bodů $a_1, \dots, a_k$.*
(2) *Posloupnost vektorů $(a_2 - a_1, a_3 - a_1, \dots, a_k - a_1)$ tvoří bázi prostoru $\mathbf{V}$ (speciálně $k = n + 1$).*

Důkaz. K důkazu obou implikací si všimneme, že pro libovolný bod $b \in A$ a skaláry $\lambda_1, \dots, \lambda_k$, $\lambda_1 + \dots + \lambda_k = 1$ vztah

$$b = \lambda_1 a_1 + \dots + \lambda_k a_k \quad ,$$

platí právě tehdy, když platí vztah

$$b - a_1 = \lambda_2(a_2 - a_1) + \lambda_3(a_3 - a_1) + \dots + \lambda_k(a_k - a_1) \quad .$$

(1) $\Rightarrow$ (2). Pro libovolný vektor $\mathbf{v}$ najdeme vyjádření bodu $b = a_1 + \mathbf{v}$ jako afinní kombinaci bodů $a_1, \dots, a_k$ a druhá ekvivalentní rovnost nám dává vyjádření vektoru $b -$

$a_1 = \mathbf{v}$ jako lineární kombinaci vektorů $a_2 - a_1, \ldots, a_k - a_1$. To dokazuje, že posloupnost generuje $\mathbf{V}$. Je-li $\mathbf{o} = \lambda_2(a_2 - a_1) + \cdots + \lambda_k(a_k - a_1)$ netriviální lineární kombinace a položíme-li $\lambda_1 = 1 - \lambda_2 - \ldots - \lambda_k$, $b = a_1$ dostáváme z první rovnosti vyjádření bodu $b = a_1$ jako afinní kombinaci bodů $a_1, \ldots, a_k$ rozdílnou od $a_1 = 1a_1 + 0a_2 + \cdots + 0a_k$. Tento spor dokazuje, že posloupnost $(a_2 - a_1, \ldots, a_k - a_1)$ je lineárně nezávislá, takže je to báze.

(2) $\Rightarrow$ (1). Důkaz je rovněž přímočarý užitím výše uvedené ekvivalence. □

První podmínka nezávisí na pořadí bodů $a_1, \ldots, a_k$, tedy lineární nezávislost posloupnosti v druhé části rovněž nezávisí na pořadí těchto bodů. Jako cvičení dokažte toto pozorování přímo.

Jsou-li splněny ekvivalentní podmínky v tvrzení, říkáme, že $Z = (a_1, \ldots, a_{n+1})$ je barycentrická soustava souřadnic a $(n+1)$-tici koeficientů $(\lambda_1, \ldots, \lambda_{n+1})^T$ ve vyjádření bodu $b \in A$ nazýváme barycentrické souřadnice bodu b vzhledem Z.

Definice 12.17. Nechť $\mathbf{A}$ je afinní prostor dimenze n s prostorem vektorů $\mathbf{V}$. *Barycentrická soustava souřadnic* je $(n+1)$-tice bodů $(a_1, \ldots, a_{n+1})$, které splňují ekvivalentní podmínky v tvrzení 12.16.

Je-li $Z = (a_1, \ldots, a_{n+1})$ barycentrická soustava souřadnic afinního prostoru $\mathbf{A}$ a $b \in A$, pak $(n+1)$-tici skalárů $(\lambda_1, \ldots, \lambda_{n+1})^T$ nazýváme *barycentrické souřadnice bodu b vzhledem k Z*, pokud $b = \lambda a_1 + \cdots + \lambda_{n+1} a_{n+1}$.

Podle tvrzení je $Z = (a_1, \ldots, a_{n+1})$ barycentrická soustava souřadnic právě tehdy, když je $S = (a_1, a_2 - a_1, a_3 - a_1, \ldots, a_{n+1} - a_1)$ soustava souřadnic prostoru $\mathbf{A}$. V důkazu jsme si všimli, že pokud známe souřadnice bodu b vzhledem k S, řekněme $[b]_S = (\lambda_2, \ldots, \lambda_{n+1})^T$, pak snadno spočítáme barycentrické souřadnice bodu b: $(1 - \lambda_2 - \ldots - \lambda_k, \lambda_2, \ldots, \lambda_{n+1})$.

Příklad 12.18. V afinním prostoru $\mathbb{R}^2$ vyjádříme b v barycentrické soustavě souřadnic (a_1, a_2, a_3).

$$b = \begin{pmatrix} 0 \\ -1 \end{pmatrix}, \quad a_1 = \begin{pmatrix} 2 \\ 7 \end{pmatrix}, \quad a_2 = \begin{pmatrix} 8 \\ 1 \end{pmatrix}, \quad a_3 = \begin{pmatrix} -6 \\ -5 \end{pmatrix}$$

Protože vektory $a_2 - a_1 = (6, -6)^T$ a $a_3 - a_1 = (-8, -12)^T$ jsou lineárně nezávislé, posloupnost (a_1, a_2, a_3) je skutečně barycentrickou soustavou souřadnic. Hledáme $\lambda_1, \lambda_2, \lambda_3$ takové, že $b = \lambda_1 a_1 + \lambda_2 a_2 + \lambda_3 a_3$ a $\lambda_1 + \lambda_2 + \lambda_3 = 1$. Přepsáním do složek dostaneme soustavu tří rovnic o třech neznámých. Druhou možností je vypočítat $[b]_S = (\lambda_2, \lambda_3)^T$, kde $S = (a_1, a_2 - a_1, a_3 - a_1)$, a dopočítat λ_1. Zvolíme druhou alternativu. Dostáváme soustavu

$$(a_2 - a_1 | a_3 - a_1 | b - a_1) = \left(\begin{array}{cc|c} 6 & -8 & -2 \\ -6 & -12 & -8 \end{array}\right) \sim \left(\begin{array}{cc|c} 6 & -8 & -2 \\ 0 & -20 & -10 \end{array}\right)$$

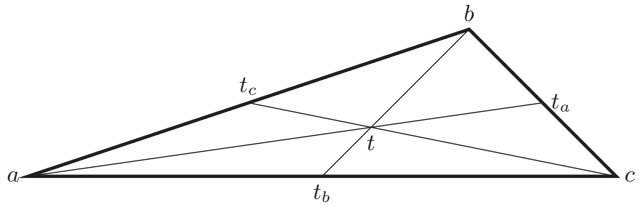
Vychází $\lambda_3 = \frac{1}{2}$, $\lambda_2 = \frac{1}{3}$ a $\lambda_1 = 1 - \lambda_2 - \lambda_3 = \frac{1}{6}$. Barycentrické souřadnice bodu b vzhledem k (a_1, a_2, a_3) jsou tedy $(\frac{1}{6}, \frac{1}{3}, \frac{1}{2})^T$.

12.2.3. *Afinní kombinace pomocí dvojic.* Afinní kombinaci více bodů v afinním prostoru $\mathbf{A}$ nad $\mathbf{T}$ lze, v případě, že charakteristika $\mathbf{T}$ není 2, získat pomocí afinních kombinací dvojic. Například pro $\mathbf{T} = \mathbb{R}$, $\lambda_1, \lambda_2, \lambda_3 \neq 0$, $\lambda_1 + \lambda_2 + \lambda_3 = 1$, $\lambda_1 + \lambda_2 \neq 0$ můžeme psát

$$\lambda_1 a + \lambda_2 b + \lambda_3 c = (\lambda_1 + \lambda_2)\left(\frac{\lambda_1}{\lambda_1 + \lambda_2} a + \frac{\lambda_2}{\lambda_1 + \lambda_2} b\right) + \lambda_3 c \ .$$

Výraz v závorce je afinní kombinací bodů a, b a celkově se jedná o afinní kombinaci této kombinace a bodu c, celý výraz tedy dává smysl. Fyzikální interpretace je taková, že těžiště soustavy hmotných bodů a, b, c s hmotnostmi $\lambda_1, \lambda_2, \lambda_3$ můžeme určit tak, že nejprve určíme těžiště hmotných bodů a, b a pak těžiště výsledného bodu (o hmotnosti $\lambda_1 + \lambda_2$) a bodu c.

Uvažujme nyní konkrétní situaci trojice bodů a, b, c v reálné afinní rovině, které neleží na jedné přímce a položme $\lambda_1 = \lambda_2 = \lambda_3 = \frac{1}{3}$. Bod $t = \frac{1}{3}a + \frac{1}{3}b + \frac{1}{3}c$ je těžištěm trojúhelníka s vrcholy a, b, c. Označíme-li $t_c = \frac{1}{2}a + \frac{1}{2}b$, tj. t_c je střed úsečky ab (co je úsečka jde formálně definovat pomocí konvexních kombinací diskutovaných níže). Podle vyjádření v předchozím odstavci máme $t = \frac{2}{3}t_c + \frac{1}{3}c$, tj. t leží na úsečce ct_c (těžnice) a tuto úsečku dělí v poměru 2 : 1. Podobně se ukáže, že t leží na úsečkách at_a a bt_b (kde t_a a t_b jsou středy stran bc a ac) a dělí tyto úsečku ve stejném poměru 2 : 1. Přirozeným způsobem jsme mimochodem nahlédli, že úsečky spojující vrcholy a středy protilehlých stran se protínají v jednom bodě a tento bod je dělí v poměru 2 : 1! Podobným způsobem lze dokázat řadu podobných geometrických poznatků (viz cvičení).



12.2.4. *Konvexní kombinace.* Krátkou neformální poznámku věnujeme tzv. konvexním kombinacím v reálných afinních prostorech. Afinní kombinace $\lambda_1 a_1 + \cdots + \lambda_k a_k$ se nazývá *konvexní*, pokud jsou všechny koeficienty nezáporné (a tím pádem také menší než 1). Konvexní kombinace souvisí s konvexními útvary. Množinu bodů nazveme *konvexní*, pokud s každými dvěma body obsahuje celou úsečku, která je spojuje. Není těžké ukázat, že každá konvexní množina je uzavřená na konvexní kombinace (cvičení).

Množina všech konvexních kombinací daných bodů $a_1, \dots, a_k$ je proto nejmenším konvexní množinou obsahující tyto body. Této množině říkáme *konvexní obal*. Rozmyslete si, že konvexním obalem dvojice bodů a, b jsou právě body ležící na úsečce ab a že konvexním obal trojice bodů a, b, c je trojúhelník (i se svým vnitřkem) s vrcholy a, b, c. Naopak, tento geometrický názor můžeme využít k formální definici úsečky ab jako konvexního obalu bodů a, b.

Příklad 12.19. Ukážeme, jak lze barycentrické souřadnice použít při zjišťování zda daný bod leží uvnitř daného trojúhelníka.

V příkladu 12.18 jsme zjistili, že barycentrické souřadnice bodu $b = (0, -1)^T$ vzhledem k $(a_1, a_2, a_3) = ((2, 7)^T, (8, 1)^T, (-6, -5)^T)$ jsou $(\frac{1}{6}, \frac{1}{3}, \frac{1}{2})$. Bod b je tedy afinní kombinací bodů (a_1, a_2, a_3) s kladnými koeficienty, proto leží uvnitř trojúhelníka s vrcholy a_1, a_2, a_3.

Konvexní množiny vznikají například při řešení soustavy lineárních nerovnic. Řešení takových soustav se týká řada důležitých teoretických i praktických problémů.

12.2.5. *Lineární kombinace odpovídající vektorům.* V tvrzení 12.12 jsme ukázali, kdy lineární kombinace bodů určuje **bod** nezávisle na volbě soustavy souřadnic, a to nám umožnilo definovat afinní kombinaci bodů. Výraz $b - a$ napovídá, kdy lze lineární kombinaci bodů smysluplně interpretovat jako **vektor**.

Tvrzení 12.20. *Nechť* $\mathbf{A}$ *je afinní prostor nad* $\mathbf{T}$*,* $a_1, \dots, a_k \in A$ *body a* $\lambda_1, \dots, \lambda_k \in \mathbf{T}$ *skaláry. Pak následující tvrzení jsou ekvivalentní.*

(1) *Vektor* $\mathbf{v}$ *o souřadnicích* $[\mathbf{v}]_S = \lambda_1[a_1]_S + \cdots + \lambda_k[a_k]_S$ *nezávisí na volbě soustavy souřadnic* S*.*
(2) $\lambda_1 + \cdots + \lambda_k = 0$.

Důkaz. Důkaz je obdobný jako u tvrzení 12.12 a přenecháme jej do cvičení. □

Podobně jako u afinním kombinací nyní můžeme v případě, že $\lambda_1 + \cdots + \lambda_k = 0$, definovat vektor $\lambda_1 a_1 + \ldots \lambda_k a_k$ předpisem

$$[\lambda_1 a_1 + \cdots + \lambda_k a_k]_S = \lambda_1 [a_1]_S + \cdots + \lambda_k [a_k]_S$$

kde S je libovolná soustava souřadnic prostoru $\mathbf{A}$, nebo například vztahem

$$\lambda_1 a_1 + \cdots + \lambda_k a_k = \lambda_2 (a_2 - a_1) + \cdots + \lambda_k (a_k - a_1) \ .$$

Obecněji, pro libovolný bod $b \in A$ platí

$$\lambda_1 a_1 + \cdots + \lambda_k a_k = \lambda_1 (a_1 - b) + \lambda_2 (a_2 - b) + \cdots + \lambda_k (a_k - b) \ .$$

12.3. **Podprostory.** Podprostory afinních prostorů definujeme analogicky jako podprostory vektorových prostorů.

Definice 12.21. Nechť $\mathbf{A}$ je afinní prostor nad tělesem $\mathbf{T}$ s prostorem vektorů $\mathbf{V}$. Afinní prostor $\mathbf{B}$ nad tělesem $\mathbf{T}$ s prostorem vektorů $\mathbf{W}$ se nazývá *(afinní) podprostor prostoru* $\mathbf{A}$, pokud $B \subseteq A$, $\mathbf{W} \leq \mathbf{V}$ a sčítání bodu a vektoru v $\mathbf{B}$ je zúžením sčítání bodu a vektoru v $\mathbf{A}$.

Je-li $\mathbf{A}$ afinní eukleidovský prostor pak $\mathbf{B}$ nazýváme *(afinním eukleidovským) podprostorem* $\mathbf{A}$, pokud je $\mathbf{B}$ afinním podprostorem $\mathbf{A}$ a navíc je skalární součin v $\mathbf{B}$ zúžením skalárního součinu v $\mathbf{A}$.

Již jsme se setkali s jedním typem podprostorů: Pro libovolný bod $a \in A$ a (vektorový) podprostor $\mathbf{W} \leq \mathbf{V}$ tvoří množina bodů $a + W$ (spolu se sčítáním zděděným z $\mathbf{A}$) afinní podprostor prostoru $\mathbf{A}$, jehož prostor vektorů je $\mathbf{W}$. Následující tvrzení ukazuje, že takto získáme všechny podprostory.

Tvrzení 12.22. *Nechť* $\mathbf{A}$ *je afinní prostor nad tělesem* $\mathbf{T}$ *s prostorem vektorů* $\mathbf{V}$ *a* $\mathbf{B}$ *je jeho podprostor s prostorem vektorů* $\mathbf{W}$. *Pak pro libovolný bod* $b \in B$ *platí* $B = b + W$. *Navíc platí* $W = \{c - b : c \in B\} = \{d - c : c, d \in B\}$.

Poznámka: Sčítání bodu z b a vektoru z W můžeme provádět v libovolném z prostorů $\mathbf{A}$ nebo $\mathbf{B}$, protože se podle definice shodují. Tím pádem se rovněž shoduje odčítání: Jsou-li $c, d \in B$ dva body v $\mathbf{B}$, pak vektor $c - d$ ve $\mathbf{W}$ je definován jako ten jednoznačně určený vektor $\mathbf{w} \in W$, pro který platí $d + \mathbf{w} = c$. Protože sčítání v $\mathbf{A}$ a $\mathbf{B}$ se shodují, vztah $d + \mathbf{w} = c$ platí i v $\mathbf{A}$, takže $d - c = \mathbf{w}$ v $\mathbf{A}$ podle definice odčítání v $\mathbf{A}$. Shodují se také jakékoliv další operace, které jsou odvozené z operací afinního prostoru, například afinní kombinace.

Důkaz. Pro libovolný vektor $\mathbf{w} \in W$ platí $b + \mathbf{w} \in B$, protože B je uzavřená na sčítání bodu a vektoru. Proto platí $b + W \subseteq B$. Naopak, pro libovolný bod $c \in B$ máme $c - b \in W$, takže $c = b + (c - b) \in b + W$, což dokazuje opačnou inkluzi.

Dodatek je rovněž snadný, plyne například z korespondence bodů a vektorů diskutované za definicí afinního prostoru. □

Příklad 12.23. Podprostory afinního prostoru $\mathbb{R}^3$ jsou čtyř typů:

- body, tj. podprostory tvaru $B = b + W$, $\dim(\mathbf{W}) = 0$, čili $W = \{\mathbf{o}\}$ a $B = \{b\}$;
- přímky, tj. podprostory tvaru $B = b + W$, $\dim(\mathbf{W}) = 1$, čili $W = \langle \mathbf{v} \rangle$, kde $\mathbf{v} \neq \mathbf{o}$, a $B = b + \langle \mathbf{v} \rangle$
- roviny, tj. podprostory tvaru $B = b + W$, $\dim(\mathbf{W}) = 2$, čili $W = \langle \mathbf{v}, \mathbf{w} \rangle$, kde $(\mathbf{v}, \mathbf{w})$ je lineárně nezávislá posloupnost, a $B = b + \langle \mathbf{v}, \mathbf{w} \rangle$
- celý prostor $B = \mathbb{R}^3$

Zavedli jsme názvy pro prostory dimenze 0 (body), 1 (přímky) a 2 (roviny). Ještě se používá pojem *nadrovina*, to je podprostor dimenze $n-1$ v prostoru dimenze n. Například nadroviny v $\mathbb{R}^1$ jsou body, nadroviny v $\mathbb{R}^2$ jsou přímky a nadroviny v $\mathbb{R}^3$ jsou roviny.

Podle tvrzení je prostor vektorů $\mathbf{W}$ podprostoru $\mathbf{B}$ prostoru $\mathbf{A}$ jednoznačně určen množinou bodů B, protože W je množina všech rozdílů bodů v B (jeden z bodů můžeme libovolně zafixovat). Proto při zadávání podprostoru často uvádíme jenom množinu bodů B a říkáme, že B je podprostor $\mathbf{A}$.

K tomu, aby neprázdná množina $B \subseteq A$ byla podprostorem afinního prostoru $\mathbf{A}$ je nutné a stačí, aby množina vektorů $W = \{c - b : c \in B\}$ (kde $b \in B$ je libovolný bod) tvořila podprostor vektorového prostoru $\mathbf{V}$. Podprostory lze také charakterizovat jako množiny bodů uzavřené na afinní kombinace.

Tvrzení 12.24. *Nechť $\mathbf{A}$ je afinní prostor a $B \subseteq A$, $B \neq \emptyset$. Pak B je podprostorem $\mathbf{A}$ právě tehdy, když každá afinní kombinace bodů z B leží v B.*

Důkaz. Je-li $\mathbf{B}$ podprostorem afinního prostoru $\mathbf{A}$, pak triviálně každá afinní kombinace bodů z B leží v B.

Předpokládejme naopak, že každá afinní kombinace bodů z B leží v B a zvolme libovolný bod $b \in B$. Je potřeba ukázat, že množina $W = \{c - b : c \in B\}$ je podprostorem prostoru vektorů $\mathbf{V}$ afinního prostoru $\mathbf{A}$. K tomu je potřeba ověřit, že W je uzavřená na sčítání a násobení skalárem. Jsou-li c, c' dva body z B, pak

$$(c - b) + (c' - b) = (c + c' - b) - b \ ,$$

kde $c + c' - b$ je afinní kombinací bodů z B, která v B podle předpokladu leží, takže $(c - b) + (c' - b) \in W$ a množina W je proto uzavřená na sčítání. Je-li $c \in B$ a $t \in T$, pak

$$t(c - b) = (tc + (1 - t)b) - b \ .$$

Závorka na pravé straně je opět afinní kombinace bodů z B a dostáváme uzavřenost W na násobení skalárem. □

Podprostory vektorových prostorů často zadáváme pomocí množiny generátorů. Podobně, podprostory afinního prostoru $\mathbf{A}$ často zadáváme pomocí „generující" množiny bodů X, říkáme například přímka určená body a, b nebo rovina určená body a, b, c, atd.

Definice 12.25. Nechť X je neprázdná podmnožina bodů afinního prostoru $\mathbf{A}$ nad tělesem $\mathbf{T}$. *Afinním obalem* množiny X rozumíme množinu $\langle X \rangle$ všech afinních kombinací bodů z X, tj.

$$\langle X \rangle = \{\lambda_1 a_1 + \cdots + \lambda_k a_k : a_1, \ldots, a_k \in X,\ \lambda_1, \ldots, \lambda_k \in T,\ \lambda_1 + \cdots + \lambda_k = 1\}$$

Pro afinní obal bodů užíváme stejné značení jako pro lineární obal. Musíme si proto vždy uvědomit, zda prvky X jsou body nebo vektory.

Tvrzení 12.26. *Nechť X je neprázdná podmnožina bodů afinního prostoru $\mathbf{A}$ nad tělesem $\mathbf{T}$. Pak $\langle X \rangle$ je podprostor afinního prostoru $\mathbf{A}$ a pro jeho prostor vektorů $\mathbf{W}$ platí*

$$\begin{aligned} W &= \{\lambda_1 a_1 + \cdots + \lambda_k a_k : a_1, \ldots, a_k \in X,\ \lambda_1, \ldots, \lambda_k \in T,\ \lambda_1 + \cdots + \lambda_k = 0\} \\ &= \langle \{c - b : c \in X\} \rangle \ , \end{aligned}$$

kde b je libovolný bod v X.

Důkaz. Protože afinní kombinace afinních kombinací je afinní kombinace, je $\langle X \rangle$ je podprostorem $\mathbf{A}$ podle charakterizace podprostorů pomocí afinních kombinací v tvrzení 12.24. Zvolme $b \in X$ libovolně. Prostor vektorů $\mathbf{W}$ podprostoru $\langle X \rangle$ je roven (viz tvrzení 12.22) $W = \{c - b : c \in \langle X \rangle\}$. Každý bod c v $\langle X \rangle$ je tvaru $c = \lambda_1 a_1 + \cdots + \lambda_k a_k$, kde $\lambda_1 + \cdots + \lambda_k = 1$, takže každý vektor $c - b$ je tvaru $\lambda_1 a_1 + \cdots + \lambda_k a_k + (-1)b$, kde $\lambda_1 + \cdots + \lambda_k + (-1) = 0$. To dokazuje inkluzi $\subseteq$ v první rovnosti. Naopak, každý vektor tvaru $\lambda_1 a_1 + \cdots + \lambda_k a_k$, kde $\lambda_1 + \cdots + \lambda_k = 0$, lze psát ve tvaru $(\lambda_1 a_1 + \cdots + \lambda_k a_k + 1 \cdot b) - b$, kde $\lambda_1 + \cdots + \lambda_k + 1 = 1$, což dokazuje druhou inkluzi.

Druhou část přenecháme do cvičení. □

Každý podprostor je uzavřený na afinní kombinace bodů. Proto každý podprostor afinního prostoru $\mathbf{A}$ obsahující množinu X musí obsahovat také $\langle X \rangle$. V tomto smyslu je $\langle X \rangle$ „nejmenší" podprostor $\mathbf{A}$ obsahující X.

Příklad 12.27. Afinním obalem dvojice bodů $X = \{a, b\}$, $a \neq b$ je přímka

$$\langle X \rangle = \{\lambda_1 a + \lambda_2 b : \lambda_1 + \lambda_2 = 1\} = a + W = b + W \ ,$$

kde

$$W = \{\lambda_1 a + \lambda_2 b : \lambda_1 + \lambda_2 = 0\} = \langle b - a \rangle$$

Konkrétně, pro body $a = (1,2)^T$, $b = (4,6)^T$ v afinním prostoru $\mathbb{R}^2$ je

$$\begin{aligned}\langle X \rangle &= \left\{\lambda_1 \begin{pmatrix} 1 \\ 2 \end{pmatrix} + \lambda_2 \begin{pmatrix} 4 \\ 6 \end{pmatrix} : \lambda_1 + \lambda_2 = 1\right\} \\ &= \begin{pmatrix} 1 \\ 2 \end{pmatrix} + \left\{\lambda_1 \begin{pmatrix} 1 \\ 2 \end{pmatrix} + \lambda_2 \begin{pmatrix} 4 \\ 6 \end{pmatrix} : \lambda_1 + \lambda_2 = 0\right\} \\ &= \begin{pmatrix} 1 \\ 2 \end{pmatrix} + \left\langle \begin{pmatrix} 3 \\ 4 \end{pmatrix} \right\rangle = \begin{pmatrix} 4 \\ 6 \end{pmatrix} + \left\langle \begin{pmatrix} 3 \\ 4 \end{pmatrix} \right\rangle\end{aligned}$$

12.3.1. *Bodový, parametrický a rovnicový popis podprostoru.* Podprostor afinního prostoru $\mathbf{A}$ dimenze n můžeme popsat následujícími způsoby:

- *Bodově,* zadáním množiny bodů $X = \{a_1, \dots, a_l\}$. Množina X určuje podprostor $B = \langle X \rangle$ tvořený všemi afinními kombinacemi bodů z X. Prostor vektorů W je roven lineárnímu obalu $\langle a_2 - a_1, \dots, a_l - a_1 \rangle$, takže na zadání prostoru dimenze k potřebujeme alespoň $k + 1$ bodů. Naopak, máme-li prostor $\mathbf{B}$ dimenze k a zvolíme $a_1, \dots, a_{k+1} \in B$ tak, aby $(a_2 - a_1, \dots, a_{k+1} - a_1)$ byla lineárně nezávislá posloupnost, pak je $(a_1, \dots, a_{k+1})$ barycentrická soustava souřadnic prostoru $\mathbf{B}$, tj. každý bod lze jednoznačným způsobem zapsat jako afinní kombinaci bodů $a_1, \dots, a_{k+1}$ (viz tvrzení 12.16).
- *Parametricky,* zadáním bodu b a množiny vektorů $\{\mathbf{v}_1, \dots, \mathbf{v}_l\}$. Daný bod a dané vektory určují podprostor $B = b + W = b + \langle \mathbf{v}_1, \dots, \mathbf{v}_l \rangle$. Na zadání prostoru dimenze k potřebujeme bod a alespoň k vektorů. Naopak, máme-li prostor $\mathbf{B}$ dimenze k s prostorem vektorů $\mathbf{W}$, zvolíme $b \in B$ libovolně a zvolíme k-tici lineárně nezávislých vektorů z W, pak $B = b + W$ a každý bod lze jednoznačným způsobem vyjádřit ve tvaru $b + t_1\mathbf{v}_1 + \cdots + t_k\mathbf{v}_k$.

Máme-li B zadán parametricky jako $B = b + \langle \mathbf{v}_1, \dots, \mathbf{v}_l \rangle$ a S je soustava souřadnic prostoru $\mathbf{A}$, pak vyjádření B v soustavě souřadnic S je afinní podprostor $[B]_S = [b]_S + \langle [\mathbf{v}_1]_S, \dots, [\mathbf{v}_l]_S \rangle \leq \mathbf{T}^n$. Takové podprostory aritmetických afinních prostorů vznikají při řešení soustav lineárních rovnic. To nám dává další možný popis podprostorů.

- *Rovnicově,* zadáním soustavy souřadnic S prostoru $\mathbf{A}$ a soustavy lineárních rovnic $Rx = c$ o n neznámých. Řešení soustavy je afinní podprostor $[B]_S = \{x \in \mathbf{T}^n : Rx = c\}$ prostoru $\mathbf{T}^n$, ten určuje podprostor $B = b + W$. Souřadnice $[b]_S$ bodu b jsou partikulárním řešením soustavy a $[W]_S = \operatorname{Ker} R$. Máme-li l rovnic, pak jádro matice soustavy má dimenzi alespoň $n - l$, takže $\dim(\mathbf{W}) \geq n - l$. Pokud má matice soustavy plnou hodnost l, pak $\dim(\mathbf{W}) = n - l$. K zadání prostoru dimenze k proto potřebujeme alespoň $n - k$ rovnic.

Přechod od rovnicového popisu k parametrickému spočívá ve vyřešení soustavy lineárních rovnic. Jak z parametrického popisu vytvořit rovnicový popisuje důkaz následujícího tvrzení.

Tvrzení 12.28. *Nechť $b + W$ je podprostor dimenze k aritmetického afinního prostoru $\mathbf{T}^n$. Pak existuje matice R typu $(n-k) \times n$ nad $\mathbf{T}$ a bod $c \in \mathbf{T}^k$ takový, že množina řešení soustavy rovnic $Rx = c$ je rovná $b + W$.*

Důkaz. Označme $\mathbf{v}_1, \dots, \mathbf{v}_k$ nějakou bázi W, tj. $W = \langle \mathbf{v}_1, \dots, \mathbf{v}_k \rangle$ a uvažujme matici $C = (\mathbf{v}_1 | \dots | \mathbf{v}_k)^T$. Podle věty o dimenzi jádra a obrazu je $\dim \operatorname{Ker} C = n - k$. Označme $(\mathbf{w}_1, \dots, \mathbf{w}_{n-k})$ nějakou bázi $\operatorname{Ker} C$, $R = (\mathbf{w}_1 | \dots | \mathbf{w}_{n-k})^T$ a $c = Rb$.

Jádro matice R má dimenzi $n-(n-k)=k$ a obsahuje každý z vektorů $\mathbf{v}_i$, protože pro libovolné $j \in \{1,\dots,n-k\}$ platí $\mathbf{w}_j^T\mathbf{v}_i = \mathbf{v}_i^T\mathbf{w}_j = 0$ z volby vektorů $\mathbf{w}_1,\dots,\mathbf{w}_{n-k}$. Platí proto $\operatorname{Ker} R = W$.

Protože b je podle volby c partikulárním řešením soustavy $Rx=c$, je množina všech řešení soustavy $Rx=c$ rovna $b+\operatorname{Ker} R = b+W$. □

V důkazu máme zároveň návod jak hledat rovnicový popis podprostoru zadaného parametricky. Pokud vzhledem k soustavě souřadnic S je $[B]_S = b+W$, napíšeme nějakou bázi W (nebo množinu generátorů W) do řádků matice a vyřešíme homogenní soustavu rovnic s touto maticí. Bázi množiny řešení napíšeme do řádků matice R a určíme pravou stranu $c=Rb$. Tím získáme rovnicový popis $[B]_S = \{x \in T^n : Rx = c\}$.

Navíc, je-li $\mathbf{A}$ afinní eukleidovský prostor a S jeho kartézská soustava, pak řádky matice R generují prostor $([W]_S)^\perp = [W^\perp]_S$, tj. generují vyjádření ortogonálního doplňku prostoru W vzhledem k S. Prvkům ortogonálního doplňku W říkáme *normálové vektory.*

Příklad 12.29. Určíme parametricky podprostor $\mathbf{B}$ prostoru $\mathbb{R}^5$ daný rovnicovým popisem vzhledem ke kanonické bázi:

$$\begin{pmatrix} 1 & 2 & -1 & 0 & 2 \\ 2 & 4 & 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 1 \\ 4 \end{pmatrix}$$

Na tomto místě si rovněž můžeme uvědomit, že každá netriviální rovnice určuje nadrovinu v $\mathbf{A}$ (v našem případě nadrovinu v $\mathbb{R}^5$), takže rovnicové vyjádření podprostoru můžeme chápat jako vyjádření pomocí průniku nadrovin.

Soustavu vyřešíme Gaussovou eliminační metodou

$$\left(\begin{array}{ccccc|c} 1 & 2 & -1 & 0 & 2 & 1 \\ 2 & 4 & 0 & 1 & -1 & 4 \end{array}\right) \sim \left(\begin{array}{ccccc|c} 1 & 2 & -1 & 0 & 2 & 1 \\ 0 & 0 & 2 & 1 & -5 & 2 \end{array}\right)$$

$$B = b+W = \begin{pmatrix} 2 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \left\langle \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ -1 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 5 \\ 0 \\ 2 \end{pmatrix} \right\rangle$$

Vidíme, že $\mathbf{B}$ je podprostor dimenze 3.

Nyní si představme, že B je zadaný parametricky a zapomeňme na původní rovnicové vyjádření. Chceme nalézt soustavu $(R|c)$, aby jejím řešením byl podprostor $B=b+W$. Napíšeme generátory prostoru W do řádků matice a najdeme její jádro.

$$\operatorname{Ker} \begin{pmatrix} -2 & 1 & 0 & 0 & 0 \\ -1 & 0 & -1 & 2 & 0 \\ 1 & 0 & 5 & 0 & 2 \end{pmatrix} = \operatorname{Ker} \begin{pmatrix} 1 & 0 & 5 & 0 & 2 \\ 0 & 1 & 10 & 0 & 4 \\ 0 & 0 & 4 & 2 & 2 \end{pmatrix} = \left\langle \begin{pmatrix} 5 \\ 10 \\ -1 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ -1 \\ 0 \\ 2 \end{pmatrix} \right\rangle$$

Matici R tedy zvolíme takto:

$$R = \begin{pmatrix} 1 & 2 & -1 & 0 & 2 \\ 5 & 10 & -1 & 2 & 0 \end{pmatrix}.$$

Zbývá zvolit pravou stranu c tak, aby bod b byl partikulárním řešením. Dosazením získáme $c = Ab = (1,9)^T$. Rovnicový popis prostoru B je tedy například

$$\begin{pmatrix} 1 & 2 & -1 & 0 & 2 \\ 5 & 10 & -1 & 2 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 1 \\ 9 \end{pmatrix}$$

Vyšel jiný rovnicový popis než původní. To není překvapivé, podprostor můžeme parametricky i rovnicově zpravidla vyjádřit mnoha způsoby.

Z rovnicového popisu vidíme také normálové vektory – lineární obal řádků matice A tvoří právě vektory kolmé na W vzhledem ke standardímu skalárnímu součinu.

Shrňme různé způsoby vyjádření přímek a rovin v afinního eukleidovském prostoru $\mathbb{R}^3$ se standardním skalárním součinem.

- *Přímku* můžeme popsat jako afinní obal dvojice různých bodů, parametricky ve tvaru $b + \langle \mathbf{v} \rangle, \mathbf{v} \neq \mathbf{o}$, nebo dvěma rovnicemi $a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = c_1, a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = c_2$, přičemž normálové vektory této přímky jsou právě vektory v $\left\langle (a_{11}, a_{12}, a_{13})^T, (a_{21}, a_{22}, a_{23})^T \right\rangle$.
- *Rovinu* můžeme popsat jako afinní obal trojice bodů neležících na jedné přímce, parametricky ve tvaru $b + \langle \mathbf{v}_1, \mathbf{v}_2 \rangle$, kde $(\mathbf{v}_1, \mathbf{v}_2)$ je lineárně nezávislá posloupnost, nebo rovnicí $a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = c_1$, přičemž normálové vektory této roviny jsou právě vektory v $\left\langle (a_{11}, a_{12}, a_{13})^T \right\rangle$.

OBRAZEK

Stejná diskuze platí pro libovolný afinní eukleidovský prostor dimenze 3, kde rovnicový popis bereme vzhledem k nějaké kartézské soustavě souřadnic. Vynecháme-li poznámky o normálových vektorech, pak diskuze platí v libovolném afinním prostoru dimenze 3, kde rovnicový popis bereme vzhledem k libovolné soustavě souřadnic.

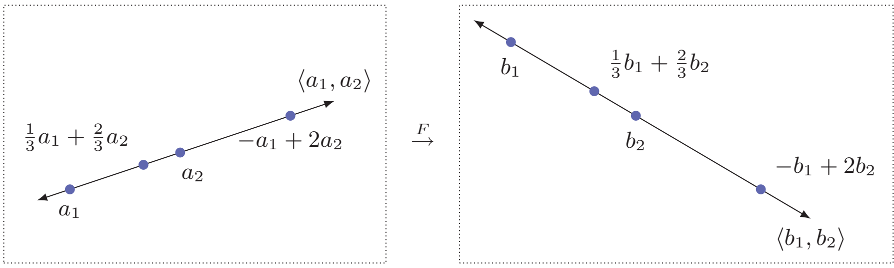
12.4. **Afinní zobrazení.** Lineární zobrazení mezi vektorovými prostory je zobrazení zachovávající součet a násobení skalárem, ekvivalentně, zobrazení zachovávající lineární kombinace. Obdobně zavedeme afinní zobrazení mezi afinními prostory jako zobrazení zachovávající afinní kombinace bodů.

Definice 12.30. Nechť $\mathbf{A}$ a $\mathbf{B}$ jsou afinní prostory nad stejným tělesem $\mathbf{T}$. Zobrazení $F : A \to B$ nazýváme *afinní zobrazení z* $\mathbf{A}$ *do* $\mathbf{B}$, značíme $F : \mathbf{A} \to \mathbf{B}$, pokud zachovává afinní kombinace, tj. pro libovolné $k \in \mathbb{N}$, $a_1, \dots, a_k \in A$, $\lambda_1, \dots, \lambda_k \in T$, $\lambda_1 + \cdots + \lambda_k = 1$ platí

$$F(\lambda_1 a_1 + \cdots + \lambda_k a_k) = \lambda_1 F(a_1) + \cdots + \lambda_k F(a_k) \ .$$

Slovy, obraz afinní kombinace je afinní kombinace obrazů se stejnými koeficienty. Fyzikální interpretace: těžiště soustavy hmotných bodů se musí zobrazit na těžiště obrazů se stejnými hmotnostmi.

Podíváme se podrobněji na případ $k = 2$ v definici. Zvolíme pevně dva různé body $a_1, a_2 \in A$ a označíme $b_1 = F(a_1)$, $b_2 = F(a_2)$. Každý bod c na přímce $\langle a_1, a_2 \rangle$ lze zapsat jako afinní kombinaci $c = \lambda_1 a_1 + \lambda_2 a_2$. Jeho obrazem musí být bod $F(c) = \lambda_1 b_1 + \lambda_2 b_2$. Obrazem je tedy bod v $\langle b_1, b_2 \rangle$, který má stejné poměry „orientovaných vzdáleností“ od bodů b_1, b_2 jako má bod c od bodů a_1, a_2. V degenerovaném případě kdy $b_1 = b_2$ se všechny body přímky $\langle a_1, a_2 \rangle$ zobrazí do b_1. V části 12.2.3 (viz cvičení **??**) jsme diskutovali, že v případě, že těleso má charakteristiku různou od dva, lze každou afinní kombinaci napsat pomocí afinní kombinace dvojic. Rozmyslete si (cvičení), že tím pádem by pro taková tělesa stačilo v definici požadovat zachovávání afinních kombinací dvojic. Jinými slovy, afinní zobrazení je takové zobrazení, které zobrazuje přímky na přímky nebo body a zachovává poměry „orientovaných vzdáleností“ bodů na přímce (opět předpokládáme charakteristiku různou od dva).



OBRÁZEK 84. Afinní zobrazení F, kde $b_i = F(a_i)$.

Dobrou představu o afinních zobrazeních z prostoru **A** dimenze n do **B** (libovolné dimenze) si vytvoříme, uvážíme-li nějakou barycentrickou soustavu souřadnic $(a_1, \dots, a_{n+1})$ v **A** a obrazy $b_i = F(a_i)$. Každý bod $a \in A$ lze zapsat jednoznačně jako afinní kombinaci $a = \lambda_1 a_1 + \cdots + \lambda_{n+1} a_{n+1}$ a obraz je pak nutně $F(a) = \lambda_1 b_1 + \cdots + \lambda_{n+1} b_{k+1}$. Naopak, na barycentrické soustavě souřadnic si můžeme obrazy předepsat libovolně a to jednoznačně určuje afinní zobrazení. Tyto skutečnosti jsou obdobou tvrzení 6.4 o určení lineárního zobrazení na bázi.

OBRAZEK (v R2)

Tvrzení 12.31. *Nechť* **A** *a* **B** *jsou vektorové prostory nad tělesem* **T**, $\dim \mathbf{A} = n$, $(a_1, \dots, a_{n+1})$ *je barycentrická soustava souřadnic prostoru* **A** *a* $b_1, \dots, b_{n+1} \in B$. *Pak existuje právě jedno afinní zobrazení* $F : \mathbf{A} \to \mathbf{B}$ *splňující* $f(a_i) = b_i$ *pro každé* $i \in \{1, 2, \dots, n+1\}$.

Důkaz. Jednoznačnost plyne z definice. Abychom dokázali existenci, definujeme F jak si vynucuje definice, tj. pro bod $a \in A$ položíme $F(a) = \lambda_1 b_1 + \cdots + \lambda_{n+1} b_{n+1}$, kde $(\lambda_1, \dots, \lambda_{n+1})^T$ jsou barycentrické souřadnice bodu a vzhledem k dané barycentrické soustavě. Je potřeba ověřit, že vzniklé zobrazení je afinní, tj. podmínka z definice platí pro libovolné k a libovolné body. To přenecháme do cvičení. □

Konkrétní příklady afinních zobrazení:

- Konstantní zobrazení $F : \mathbf{A} \to \mathbf{B}$, které každému bodu v **A** přiřazuje pevně zvolený bod $b \in B$.
- *Posunutí* o vektor $\mathbf{v}$ (který leží v prostoru směrů prostoru **A**) je afinní zobrazení $F : \mathbf{A} \to \mathbf{A}$. Posunutím o vektor $\mathbf{v}$ přirozeně myslíme zobrazení definované $F(c) = c + \mathbf{v}$.
- Rotace o nějaký úhel, zrcadlení podle přímky, zkosení, projekce na přímku v nějakém směru, posunutí a každé složení těchto zobrazení je afinním zobrazením $F : \mathbb{R}^2 \to \mathbb{R}^2$.
- Zobrazení přiřazující bodu **A** jeho souřadnice vzhledem ke zvolené soustavě souřadnic je afinní zobrazení $F : \mathbf{A} \to \mathbf{T}^n$.

12.4.1. *Afinní a lineární zobrazení.* Afinní zobrazení mezi afinními prostory určuje přirozeným způsobem lineární zobrazení mezi prostory vektorů. Naopak, lineární zobrazení mezi jejich prostory vektorů a obraz jednoho bodu určují jednoznačně afinní zobrazení.

Podrobněji. Uvažujme afinní prostor **A** s prostorem vektorů **V**, afinní prostor **B** s prostorem vektorů **W** (oboje nad tělesem **T**) a afinní zobrazení $F : \mathbf{A} \to \mathbf{B}$. Zvolíme libovolný bod $a \in A$ a definujeme zobrazení $f : V \to W$ vztahem

$$f(\mathbf{v}) = F(a + \mathbf{v}) - F(a) \quad \text{pro každý vektor } \mathbf{v} \in V \ .$$

Alternativně můžeme stejnou definici psát

$$f(c-a) = F(c) - F(a) \quad \text{pro každý bod } c \in A \ .$$

Ukážeme, že takto definované zobrazení f nezávisí na volbě bodu a. Z definice afinního zobrazení dostaneme, že pro libovolný bod $a' \in A$ a vektor $\mathbf{v} \in V$ platí

$$F(a' + \mathbf{v}) = F((a+\mathbf{v}) - a + a') = F(a+\mathbf{v}) - F(a) + F(a') \ ,$$

což po úpravě dává

$$F(a' + \mathbf{v}) - F(a') = F(a+\mathbf{v}) - F(a) \ ,$$

takže f skutečně nezávisí na volbě bodu a. Jednoduchou úpravou definice f zjistíme, že zobrazení F je určené f a obrazem libovolného bodu $a \in A$ vztahem

$$F(c) = F(a) + f(c-a) \qquad \text{pro libovolný bod } c \in A$$

nebo

$$F(a+\mathbf{v}) = F(a) + f(\mathbf{v}) \qquad \text{pro libovolný vektor } \mathbf{v} \in V \ .$$

Jsou-li $a_1, \dots, a_k \in A$ libovolné body a $\lambda_1, \dots, \lambda_k \in T$ skaláry takové, že $\lambda_1 + \dots + \lambda_k = 0$, pak „lineární kombinace“ $\lambda_1 a_1 + \dots + \lambda_k a_k$ odpovídá nějakému vektoru ve $\mathbf{V}$. Podíváme se na jeho obraz při zobrazení f. Podle definice f a definice afinního zobrazení je

$$\begin{aligned} f(\lambda_1 a_1 + \dots + \lambda_k a_k) &= F(a + \lambda_1 a_1 + \dots + \lambda_k a_k) - F(a) \\ &= F(a) + \lambda_1 F(a_1) + \dots + \lambda_k F(a_k) - F(a) \\ &= \lambda_1 F(a_1) + \dots + \lambda_k F(a_k) \ . \end{aligned}$$

Ještě nahlédneme, že f je skutečně lineární zobrazení: Pro libovolné dva vektory $\mathbf{u}, \mathbf{v} \in V$ a skalár $\lambda \in T$ označíme $b = a + \mathbf{u}$, $c = a + \mathbf{u} + \mathbf{v}$ a spočítáme

$$\begin{aligned} f(\mathbf{u}+\mathbf{v}) &= f(c-a) = F(c) - F(a) = (F(c) - F(b)) + (F(b) - F(a)) \\ &= f(c-b) + f(b-a) = f(\mathbf{u}) + f(\mathbf{v}) \\ f(\lambda\mathbf{u}) &= f(\lambda b - \lambda a) = \lambda F(b) - \lambda F(a) = \lambda(F(b) - F(a)) = \lambda f(b-a) = \lambda f(\mathbf{u}) \end{aligned}$$

Naopak, je-li $f : \mathbf{V} \to \mathbf{W}$ lineární zobrazení a $a \in A$, $b \in B$, pak zobrazení $F : A \to B$ definované vztahem

$$F(c) = b + f(c-a) \qquad \text{pro každé } c \in A$$

ekvivalentně

$$F(a+\mathbf{v}) = b + f(\mathbf{v}) \qquad \text{pro každé } \mathbf{v} \in V$$

je afinní zobrazení $F : \mathbf{A} \to \mathbf{B}$ (pro které $F(a) = b$), protože pro libovolnou afinní kombinaci $\lambda_1 a_1 + \dots + \lambda_k a_k$ ($\sum_1^k \lambda_i = 1$) máme

$$\begin{aligned} F(\lambda_1 a_1 + \dots + \lambda_k a_k) &= b + f(\lambda_1 a_1 + \dots + \lambda_k a_k - a) \\ &= b + f(\lambda_1(a_1 - a) + \dots + \lambda_k(a_k - a)) \\ &= b + \lambda_1 f(a_1 - a) + \dots + \lambda_k f(a_k - a) \\ &= \lambda_1(b + f(a_1 - a)) + \dots + \lambda_k(b + f(a_k - a)) \\ &= \lambda_1 F(a_1) + \dots + \lambda_k F(a_k) \end{aligned}$$

Shrneme učiněná pozorování.

Tvrzení 12.32. *Nechť $\mathbf{A}, \mathbf{B}$ jsou afinní prostory nad stejným tělesem $\mathbf{T}$ a $\mathbf{V}, \mathbf{W}$ jsou jejich prostory vektorů. Pak platí:*

(1) *Pro libovolné afinní zobrazení $F : \mathbf{A} \to \mathbf{B}$ zobrazení $f : V \to W$ definované pro $\mathbf{v} \in V$ vztahem $f(\mathbf{v}) = F(a+\mathbf{v}) - F(a)$ nezávisí na volbě bodu a a je lineárním zobrazením $\mathbf{V} \to \mathbf{W}$. Pro libovolné $a, c \in A$ platí $F(c) = F(a) + f(c-a)$ a pro libovolnou kombinaci $\lambda_1 a_1 + \dots \lambda_k a_k$, $\lambda_1 + \dots + \lambda_k = 0$ platí*

$$f(\lambda_1 a_1 + \dots + \lambda_k a_k) = \lambda_1 F(a_1) + \dots + \lambda_k F(a_k) \ .$$

(2) *pro libovolné lineární zobrazení* $f : \mathbf{V} \to \mathbf{W}$ *a body* $a \in A$, $b \in B$ *je zobrazení* $F : A \to B$ *definované vztahem* $F(c) = b + f(c - a)$ *afinní zobrazení* $\mathbf{A} \to \mathbf{B}$.

V situaci předchozího tvrzení říkáme, že afinní zobrazení F vytváří lineární zobrazení f nebo, že f je lineární zobrazení příslušné F, apod. Například afinní zobrazení $F : \mathbf{A} \to \mathbf{A}$ vytvořená identitou jsou právě posunutí, zobrazení vytvořená rotací jsou rotace složené s posunutím.

Následující pozorování shrnuje některé jednoduché, ale důležité vlastnosti afinních zobrazení a příslušných lineárních.

Pozorování 12.33. *Nechť* $F : \mathbf{A} \to \mathbf{B}$ *je afinní zobrazení a* $f : \mathbf{V} \to \mathbf{W}$ *příslušné lineární zobrazení. Pak platí:*

(1) *F je prosté právě tehdy, když f je prosté,*
(2) *F je na právě tehdy, když f je na.*
(3) *Obrazem podprostoru $B = b + U$ prostoru $\mathbf{A}$ při zobrazení F je podprostor $F(B) = F(b) + f(U)$ prostoru $\mathbf{B}$.*
(4) *Je-li $G : \mathbf{B} \to \mathbf{C}$ afinní zobrazení a g příslušné lineární zobrazení, pak složené zobrazení $G \circ F$ je afinním zobrazením $\mathbf{A} \to \mathbf{C}$ a jemu příslušné lineární zobrazení je $g \circ f$.*

Důkaz. Cvičení. □

12.4.2. *Afinní zobrazení v souřadnicích.* Na příkladu ukážeme jak popsat afinní zobrazení mezi konečně dimenzionálními prostory v souřadnicích.

Příklad 12.34. Popíšeme zobrazení, které zobrazuje trojici bodů $a_1, a_2, a_3 \in \mathbb{R}^2$ na trojici bodů $b_1, b_2, b_3 \in \mathbb{R}^3$ (v tomto pořadí).

$$a_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, a_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}, a_3 = \begin{pmatrix} 2 \\ -1 \end{pmatrix}, b_1 = \begin{pmatrix} 5 \\ 3 \\ 2 \end{pmatrix}, b_2 = \begin{pmatrix} 3 \\ -1 \\ 4 \end{pmatrix} b_3 = \begin{pmatrix} 0 \\ 3 \\ -1 \end{pmatrix}$$

Protože $D = (a_2 - a_1, a_3 - a_1) = ((-2,0)^T, (1,-2)^T)$ je báze $\mathbb{R}^2$, tvoří trojice (a_1, a_2, a_3) barycentrickou soustavu souřadnic, takže afinní zobrazení $F : \mathbb{R}^2 \to \mathbb{R}^3$ je podmínkami jednoznačně určené (viz tvrzení 12.31). Určíme příslušné lineární zobrazení $f : \mathbb{R}^2 \to \mathbb{R}^2$. Obrazem vektoru $a_2 - a_1$ je vektor $f(a_2 - a_1) = F(a_2) - F(a_1) = b_2 - b_1 = (-2,-4,2)^T$ a obrazem $a_3 - a_1$ je $f(a_3 - a_1) = b_3 - b_1 = (-5,0,-3)^T$. Matice f vzhledem k D a K_3 je proto

$$[f]^D_{K_3} = \begin{pmatrix} -2 & -5 \\ -4 & 0 \\ 2 & -3 \end{pmatrix}$$

takže vzhledem ke kanonickým bázím je

$$[f]^{K_2}_{K_3} = [f]^D_{K_3}[\mathrm{id}]^{K_2}_D = \begin{pmatrix} -2 & -5 \\ -4 & 0 \\ 2 & -3 \end{pmatrix} \begin{pmatrix} -2 & 1 \\ 0 & -2 \end{pmatrix}^{-1}$$

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

Nyní pro libovolný bod $c \in \mathbb{R}^2$ a vektor $\mathbf{v} \in \mathbb{R}^2$ je $F(c + \mathbf{v}) = F(c) + f(\mathbf{v})$. Použijeme tento vztah pro $c = (0,0)^T$ a $\mathbf{v} = (x_1, x_2)^T$ a dostáváme obraz bodu $(x_1, x_2)^T$:

$$F \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = F \begin{pmatrix} 0 \\ 0 \end{pmatrix} + f \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = F \begin{pmatrix} 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 1 & 3 \\ 2 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

Místo určování $F((0,0)^T)$ přímo, můžeme do vztahu dosadit například bod a_1 a dopočítat.

$$\begin{aligned}
F\begin{pmatrix} 1 \\ 1 \end{pmatrix} &= F\begin{pmatrix} 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 1 & 3 \\ 2 & 1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} \\
\begin{pmatrix} 5 \\ 3 \\ 2 \end{pmatrix} &= F\begin{pmatrix} 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 4 \\ 3 \\ 0 \end{pmatrix} \\
F\begin{pmatrix} 0 \\ 0 \end{pmatrix} &= \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}
\end{aligned}$$

Celkově dostáváme

$$F\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} + \begin{pmatrix} 1 & 3 \\ 2 & 1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 + x_1 + 3x_2 \\ 2x_1 + x_2 \\ 2 - x_1 + x_2 \end{pmatrix}$$

Jako zkoušku ověříme, že skutečně $F(a_i) = b_i$, $i = 1, 2, 3$.

Obecněji, máme-li afinní zobrazení $F : \mathbf{A} \to \mathbf{B}$, soustavu souřadnic $S = (a, \mathbf{v}_1, \dots, \mathbf{v}_n)$ v prostoru $\mathbf{A}$ a soustavu souřadnic $Q = (b, \mathbf{w}_1, \dots, \mathbf{w}_m)$ v prostoru $\mathbf{B}$, pak souřadnice obrazu bodu c, který máme zadaný v soustavě S, vzhledem k Q spočítáme

$$[F(c)]_Q = [F(a)]_Q + [f(c-a)]_Q = [F(a)]_Q + X[c]_S \ ,$$

kde X je matice f vzhledem k bázím $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ a $(\mathbf{w}_1, \dots, \mathbf{w}_m)$. Heslovitě, obraz je tvaru „bod plus matice krát vzor“. Když na okamžik přestaneme rozlišovat body a vektory (zvolíme počátek a bod ztotožňujeme z jeho polohovým vektorem), pak lineární zobrazení jsou „rovná zobrazení“, která zachovávají počátky, a afinní zobrazení jsou všechna rovná zobrazení. Vzniknou z lineárních složením s posunutím.

12.4.3. *Izometrie.* Izometrie mezi afinními eukleidovskými prostory je zobrazení, které zachovává vzdálenosti. Používá se také název *shodnost*, zejména v případě zobrazení mezi stejnými prostory.

Definice 12.35. Nechť $\mathbf{A}$, $\mathbf{B}$ jsou afinní eukleidovské prostory. Zobrazení $F : A \to B$ nazýváme *izometrie*, pokud zachovává vzdálenosti, tzn. pro libovolné $a, c \in A$ platí

$$\|a - c\| = \|F(a) - F(c)\| \ .$$

Intuice napovídá, že izometrie je „rovné“, tj. afinní zobrazení, a příslušné lineární zobrazení mezi prostory vektorů je ortogonální. Intuice se nemýlí, jak ukazuje následující věta.

Věta 12.36. *Nechť $\mathbf{A}$ a $\mathbf{B}$ jsou afinní eukleidovské prostory konečné dimenze a $F : A \to B$ je zobrazení. Pak následující tvrzení jsou ekvivalentní.*

(1) *F je izometrie.*
(2) *F je afinní zobrazení $\mathbf{A} \to \mathbf{B}$ a příslušné lineární zobrazení mezi prostory vektorů je ortogonální.*

Důkaz. Označíme $\mathbf{V}$,$\mathbf{W}$ prostory vektorů afinních prostorů $\mathbf{A}$, $\mathbf{B}$.

Implikace (2) $\Rightarrow$ (1) je jednoduchá: Jsou-li $a, c \in A$ libovolné body, pak

$$\|F(a) - F(c)\| = \|f(a-c)\| = \|a - c\| \ .$$

Zajímavá je opačná implikace (1) $\Rightarrow$ (2). Ukážeme myšlenku důkazu a některé technické detaily přenecháme do cvičení.

- Pro libovolné dva body $a_1, a_2 \in A$ a jejich afinní kombinaci $a = \lambda_1 a_1 + \lambda_2 a_2$ platí $F(a) = \lambda_1 F(a_1) + \lambda_2 F(a_2)$. K důkazu si všimneme, že vztah „bod je afinní kombinací dvojice bodů s koeficienty λ_1, λ_2“ můžeme charakterizovat pomocí jejich vzájemných vzdáleností (cvičení).
- Protože F zachovává afinní kombinace dvojic, je F afinní zobrazení podle cvičení **??**. Označme f příslušné lineární zobrazení mezi prostory vektorů.
- Zobrazení f zachovává normy: Pro libovolný vektor $\mathbf{v} \in V$ a bod $a \in A$ platí

$$\|f(\mathbf{v})\| = \|f((a+\mathbf{v}) - a)\| = \|F(a+\mathbf{v}) - F(a)\| = \|a + \mathbf{v} - a\| = \|\mathbf{v}\|$$

- Protože f zachovává normu, je podle tvrzení **??** ortogonální.

□

V příkladech **??** z kapitoly o vlastních číslech jsme popsali všechny ortogonální zobrazení $\mathbb{R}^2 \to \mathbb{R}^2$ a $\mathbb{R}^3 \to \mathbb{R}^3$. Z dokázané věty tak získáme v těchto případech popis všech izometrií. Izometrie $F : \mathbb{R}^2 \to \mathbb{R}^2$ jsou rotace složené s posunutím a ortogonální reflexe složené z posunutím. Izometrie $F : \mathbb{R}^3 \to \mathbb{R}^3$ jsou rotace kolem osy složené z posunutím a rotace kolem osy složené s ortogonální reflexí vzhledem k rovině a posunutím. Obdobné výsledky samozřejmě platí pro izometrie mezi dvěma libovolnými eukleidovskými prostory dimenze 2 nebo 3, stačí vše převést do $\mathbb{R}^2$ nebo $\mathbb{R}^3$ pomocí kartézských soustav souřadnic.

Obsah