Milan Hladík

LINEÁRNÍ ALGEBRA

(NEJEN)

PRO INFORMATIKY

29. listopadu 2017

Co?

Právě čtete učební text k přednáškám Lineární algebra I a II pro první ročník studia informatiky na Matematicko-fyzikální fakultě Univerzity Karlovy. Nicméně věřím, že může dobře posloužit i na jiných školách.

Proč?

Tento text vznikl mj. i proto, že žádná současná učebnice nekopíruje přesně sylabus přednášky. Nejblíže je [Tůma, 2003], pěkná skripta jsou [Rohn, 2004], kniha [Bečvář, 2005], nebo matematičtěji laděná kniha [Bican, 2009]. Anglicky psané knihy, které stojí za doporučení, jsou [Gareth, 2001], [Meyer, 2000], [Strang, 1988] nebo stručná, ale pěkně motivující knížka [Chartier, 2015]. Na stránkách J. Matouška [Matoušek, 2010] je možno nalézt přibližný podrobný sylabus přednášky, stejně jako „Šestnáct miniatur“, šestnáct aplikací lineární algebry.

Jak?

Text poskytuje poměrně ucelený výklad na sebe navazujících témat základní lineární algebry. Několik složitých důkazů z didaktických důvodů vynecháváme – konkrétně důkaz věty 4.33 o velikosti konečných těles, věty 10.39 o Jordanově normální formě, Perronovy věty 10.54 a věty 13.14 o SVD rozkladu.

Formát textu je tradiční se členěním do definic, vět, důkazů apod. Navzdory jiným moderním trendům mi připadá tento styl přehlednější a srozumitelnější. Snažil jsem, nicméně, obohatit text vysvětlujícími komentáři a ilustrativními obrázky. Na konci každé kapitoly je vždy uvedeno několik problémů na zamyšlení a často poukazující na pokročilejší souvislosti. Dále je na konci každé kapitoly její krátké neformální shrnutí a zdůraznění hlavních myšlenek a výsledků.

Jistým specifikem textu je něco, co nazývám „nahlédnutí pod pokličku“ jiných oborů. Bez detailů, na které není prostor a mnohdy ani matematické zázemí, ukazujeme netriviální aplikace a souvislosti s jinými matematickými obory. Toto se týká mj. aplikací o diskrétní a rychlé Fourierově transformaci (příklady 3.54 a 10.36), samoopravných kódech (příklad 4.40), Stewartově–Goughově platformě v robotice (příklad 7.20), vyhledávači od Google (příklad 10.64), určování struktury bílkovin (příklad 11.21), kuželoseček (sekce 12.2) či komprese dat (sekce 13.5).

Některé partie a věty dále nerozvíjíme a v případě nutnosti je lze při výkladu vynechat. Jedná se například o sekci o jednoznačnosti RREF (sekce 3.4), o dodatcích k soustavám rovnic (sekce 3.5), o prostoru lineárních zobrazení (sekce 6.4), o afinních prostorech (kapitola 7), o teorii nezáporných matic (sekce 10.6), o výpočtu vlastních čísel (sekce 10.7) a o maticových rozkladech (kapitola 13). Z vět je možno přeskočit například Shermanovu–Morrisonovu formuli (věta 3.41), Besselovu nerovnost a Parsevalovu rovnost (věta 8.25), Gramovu matici (věta 8.39), ortogonální matici a lineární zobrazení (věta 8.55 a tvrzení 8.56), nebo Courantovu–Fischerovu formuli (věta 10.53).

Chyby?

Jakékoli připomínky a případné chyby neváhejte prosím zasílat na adresu hladik@kam.mff.cuni.cz.

Poděkování!

Základ tohoto textu jsem začal psát během zimního semestru roku 2010. Děkuji všem, kteří nějakým způsobem pomohli k vylepšení textu. Mnoho drobných chyb a překlepů mi pomohla opravit řada mých studentů z let 2010–2017. Za podnětné připomínky a diskuse děkuji kolegům Jaroslavu Horáčkovi, Jiřímu Šejnohovi a Pavlu Klavíkovi, a za konstruktivní poznámky a podrobné čtení prof. Jiřímu Matouškovi.

„Pokud jsem dohlédl dále než jiní, bylo to proto, že jsem stál na ramenech obrů.“

[Isaac Newton, 1675]

Obsah

Obsah **5**

1 Úvod **7**
1.1 O knize a lineární algebře . . . 7
1.2 Pojmy a značení . . . 8
1.3 Reprezentace čísel . . . 11
1.4 Komplexní čísla . . . 11
1.5 Polynomy . . . 12
1.6 Analytická geometrie . . . 13
1.7 Optimalizace . . . 15

2 Soustavy lineárních rovnic **17**
2.1 Základní pojmy . . . 17
2.2 Gaussova eliminace . . . 19
2.3 Gaussova–Jordanova eliminace . . . 23
2.4 Aplikace . . . 25

3 Matice **29**
3.1 Základní operace s maticemi . . . 29
3.2 Regulární matice . . . 36
3.3 Inverzní matice . . . 38
3.4 Jednoznačnost RREF . . . 41
3.5 Ještě k soustavám rovnic . . . 41
3.5.1 Numerická stabilita při řešení soustav . . . 41
3.5.2 LU rozklad . . . 43
3.5.3 Iterativní metody . . . 44
3.6 Aplikace . . . 45

4 Grupy a tělesa **51**
4.1 Grupy . . . 51
4.2 Permutace . . . 53
4.3 Tělesa . . . 56
4.4 Aplikace . . . 59

5 Vektorové prostory **63**
5.1 Základní pojmy . . . 63
5.2 Podprostory a lineární kombinace . . . 64
5.3 Lineární nezávislost . . . 68
5.4 Báze . . . 69
5.5 Dimenze . . . 72
5.6 Maticové prostory . . . 75
5.7 Aplikace . . . 79

6 Lineární zobrazení **83**
6.1 Lineární zobrazení mezi obecnými prostory . . . 83
6.2 Maticová reprezentace lineárního zobrazení . . . 86
6.3 Isomorfismus . . . 91
6.4 Prostor lineárních zobrazení . . . 94

7 Afinní podprostory **97**
7.1 Základní pojmy . . . 97
7.2 Aplikace . . . 101

8 Skalární součin **107**
8.1 Skalární součin a norma . . . 107
8.2 Ortonormální báze, Gramova–Schmidtova ortogonalizace . . . 112
8.3 Ortogonální doplněk a projekce . . . 115
8.4 Ortogonální doplněk a projekce v $\mathbb{R}^n$ . . . 120
8.5 Metoda nejmenších čtverců . . . 121
8.6 Ortogonální matice . . . 123

9 Determinanty **129**
9.1 Determinant a elementární úpravy . . . 130
9.2 Další vlastnosti determinantu . . . 131
9.3 Adjungovaná matice . . . 134
9.4 Aplikace . . . 135

10 Vlastní čísla **141**
10.1 Charakteristický polynom . . . 142
10.2 Cayleyho–Hamiltonova věta . . . 145
10.3 Diagonalizovatelnost . . . 147
10.4 Jordanova normální forma . . . 152
10.5 Symetrické matice . . . 156
10.6 Teorie nezáporných matic . . . 158
10.7 Výpočet vlastních čísel . . . 159

11 Positivně (semi-)definitní matice **165**
11.1 Metody na testování positivní definitnosti . . . 166
11.2 Aplikace . . . 169

12 Kvadratické formy **173**
12.1 Bilineární a kvadratické formy . . . 173
12.2 Sylvestrův zákon setrvačnosti . . . 176

13 Maticové rozklady **183**
13.1 Householderova transformace . . . 184
13.2 QR rozklad . . . 185
13.3 Aplikace QR rozkladu . . . 187
13.4 SVD rozklad . . . 189
13.5 Aplikace SVD rozkladu . . . 191
13.6 Pseudoinverzní matice . . . 194
13.7 Maticová norma . . . 196

14 Závěrem **201**

Značení **203**

Literatura **205**

Kapitola 1

Úvod

V této kapitole nejprve nastíníme obsah a cíl celého textu a v dalších sekcích připomeneme některé základní pojmy a poznatky, které by čtenář měl znát, anebo jsou mimo hlavní proud tohoto textu, ale nějak se ho dotýkají. Vlastní látka začne až v následující kapitole.

1.1 O knize a lineární algebře

Co je lineární algebra?

Lineární algebra je jeden za základních matematických oborů a je nedílnou součástí matematických, informatických a technických studijních odvětví. Mezi hlavní nástroje, se kterými lineární algebra pracuje, patří vektory a matice. Vektory žijí v nějakém prostoru, a právě lineární algebra se zabývá lineárními objekty v prostoru, jako jsou body, přímky, roviny a podobně. Podobně studuje lineární zobrazení, jako například překlopení či rotace. Vidíme tedy, že navzdory svému názvu lineární algebra velmi úzce souvisí s geometrií a poskytuje efektivní techniky na řešení geometrických problémů.

Matice představuje další klíčový pojem v lineární algebře. Vstupní data nejrozličnějších problémů jsou v maticové formě, proto umět pracovat s maticemi a rozuměti jim je zásadní schopnost nejen pro teoretický výzkum, ale i ryze z praktického hlediska. Matice tak může reprezentovat počítačový obrázek nebo třeba data nějaké velké databáze (katalog knih, databáze zákazníků,...). Vlastnosti matic tedy přímo odráží vlastnosti objektů, které popisují.Matice dále odpovídají lineárním transformacím v eukleidovském prostoru, tudíž na matice můžeme opět nahlížet geometricky. Tento dvojí pohled – algebraický a geometrický – je nesmírně užitečný pro pochopení a pro práci s maticemi.

O čem je tento text?

První téma, které probíráme, jsou *Soustavy lineárních rovnic*. Ty tvoří nejzákladnější problém lineární algebry a jsou stále předmětem výzkumu (některé praktické problémy vedou na obrovské soustavy rovnic, které se musí počítačově vyřešit). Ukážeme algoritmus, nazývaný Gaussova eliminace, který umí vyřešit principiálně každou soustavu rovnic. Rozhodne, za je soustava řešitelná, a pokud, ano, tak vydá popis celé množiny řešení. Množina řešení je vždy lineární útvar typu přímky, roviny atp. Později, v kapitole 7 – Afinní podprostory uvidíme, že naopak každý takovýto útvar lze popsat pomocí soustavy rovnic.

Jak jsme již nastínili, *Matice* představují fundamentální nástroj lineární algebry. V kapitole 3 ukážeme základní operace s maticemi, typy matic a jejich vztah k soustavám lineárních rovnic.

Kapitola 4 – *Grupy a tělesa* ukazuje, že řada výsledků, ke kterým jsme dospěli a které později rozvineme, platí nejenom pro reálná čísla, ale jdou přirozeně rozšířit i na jiné číselné obory. Lze tudíž analogicky pracovat s komplexními čísly, nebo pro informatiky důležitou množinou bitů 0 a 1.

Vektory a vektorové prostory jsou probírané v kapitole 5. Prostory zavádíme axiomaticky pomocí vlastností, které mají splňovat. Výhodou jest, že veškerou teorii, kterou vybudujeme, můžeme použít nejen pro standardní eukleidovský prostor, ale i pro jiné, abstraktní prostory. Ukážeme, že v těchto prostorech jde zavést souřadný systém, pojem dimenze a řada dalších známých konceptů. Nakonec se ukáže, že všechny n-dimenzionální prostory jsou isomorfní, tedy z pohledu lineární algebry se chovají stejně.

Lineární zobrazení jsou transformace v prostoru, které zobrazují přímky zase na přímky a počátek na počátek. Mezi takovéto zobrazení patří otočení, zkosení, natažení, zrcadlení či projekce. Důležitou vlastností je, že každé lineární zobrazení lze popsat maticově a naopak každá matice reprezentuje nějaké lineární zobrazení. Jde tedy o dva různé pohledy na tutéž věc. A právě tento dvojí pohled je velmi užitečný, protože někdy je lepší použít algebraický přístup, a jindy zase geometrický pomocí lineárních zobrazení.

Kapitola 8 – *Skalární součin* se věnuje více geometrii. Budeme zkoumat kolmost, vzdálenosti, projekce a pod. Dosažené výsledky poslouží nejen k efektivnímu řešení všelijakých geometrických úloh, ale dají nám také geometrický pohled na algebraické problémy.

Determinanty a *Vlastní čísla* jsou důležité charakteristiky matic, a tím pádem ukazují některé vlastnosti objektu, který matice reprezentuje (ať už se jedná o soustavu rovnic nebo lineární zobrazení). Konkrétněji, determinant například umožňuje spočítat objem určitých geometrických útvarů nebo dává explicitní vzoreček na řešení soustavy rovnic. Vlastní čísla jsou ještě jemnější charakteristikou a proto dávají mnohem detailnější informaci o chování matice. Použití vlastních čísel je vskutku široké. Jeden příklad za všechny: Pokud vyjádříme maticově internetovou síť webových stránek, pak vlastní vektor odpovídající největšímu vlastnímu číslu můžeme chápat jako vektor, jehož složky udávají „důležitost“ jednotlivých stránek. Na tomto pozorování je mj. založen PageRank vyhledávače Google.

Navzdory svému názvu, lineární algebra studuje i vybrané nelineární problémy. *Positivně definitní matice* a *Kvadratické formy* (kapitola 11 a 12) se zabývají kvadratickými výrazy. Positivně definitní matice, které jsou definovány pomocí kvadratického výrazu (tzv. kvadratické formy), představují důležitou třídu matic, které se objevují při projekcích, v popisu elipsoidů, v optimalizaci či statistice. Kvadratické formy pak klasifikujeme na základní typy a vyvineme účinný nástroj na jejich rozlišování.

Maticové rozklady představují zlatý hřeb celého textu. Maticovým rozkladem rozumíme vyjádření matice jako součin několika (většinou dvou nebo tří) speciálních matic. I když se to takto z obecného popisu nezdá, maticové rozklady jako QR nebo SVD mají velkou teoretickou a výpočetní sílu. V zásadě skoro všechny problémy, kterými se tento text zabývá (a ještě mnoho dalších), jdou elegantně vyřešit pomocí QR či SVD rozkladu.

A co čtenář?

Co by si měl čtenář odnést? Čtenář by si měl odnést nejenom vlastní výsledky, ale měl by jim rozumět a chápat souvislosti. Myslím, že člověk něčemu pořádně rozumí jen tehdy, když to dokáže aplikovat a adaptovat v novém prostředí.

A jakého čtenáře si přeji? Přeji si především zvídavého čtenáře. Učením se nazpaměť se matematika nedá pochopit. Řešením úloh se sice získá patřičná dovednost, ale ani to ještě úplně nestačí. Podle mého názoru je pro pochopení matematiky důležité, když si člověk klade otázky „Proč?“ a „Co by kdyby?“. Proto si přeji si zvídavého čtenáře, který si klade otázky: Proč je tento pojem definován zrovna takto? Co by se stalo, kdybychom změnili definici takto? Co by pak ještě platilo a co už ne? Proč platí ve větě jenom implikace a ne ekvivalence? Proč jsou předpoklady takové a takové a jak by se mělo za důsledek, kdybychom je změnili? Zvídavého čtenáře pak leckdy nějaké téma zaujme tak, že sám pátrá po dalších výsledcích a souvislostech, a tím si rozšiřuje svůj obzor nad rámec, který může dát tento text.

1.2 Pojmy a značení

Nyní zavedeme některé standardní značení, které budeme používat.

Číselné obory

Připomeňme základní číselné obory:

- $\mathbb{N}$ ... množina přirozených čísel,
- $\mathbb{Z}$ ... množina celých čísel,
- $\mathbb{Q}$ ... množina racionálních čísel,
- $\mathbb{R}$ ... množina reálných čísel,

- $\mathbb{C}$ ... množina komplexních čísel (podrobněji v sekci 1.4).

Suma

Symbol sumy „$\sum$“ reprezentuje součet všech instancí výrazu za sumou pro všechny přípustné hodnoty indexu (resp. indexů) pod sumou. Konkrétně:

$$\begin{aligned}
\sum_{i=1}^{n} a_i &\quad \text{je zkratka za} \quad a_1 + \ldots + a_n, \\
\sum_{1 \le i \le 8 \text{ a liché}} a_i &\quad \text{je zkratka za} \quad a_1 + a_3 + a_5 + a_7, \\
\sum_{i,j \in \{1,2\}} a_{ij} &\quad \text{je zkratka za} \quad a_{11} + a_{12} + a_{21} + a_{22}.
\end{aligned}$$

Z praktických důvodů součet přes prázdnou množinu definujeme jako 0, protože přičtením nuly se hodnota nezmění.

Produkt

Analogicky symbol „$\prod$“ se používá pro součin, tedy např.

$$\prod_{i=1}^{n} a_i \quad \text{je zkratka za} \quad a_1 \cdot a_2 \cdot \ldots \cdot a_n.$$

Kvantifikátory

Symbol „$\forall$“ je zkratka pro sousloví „pro všechna“, a symbol „$\exists$“ znamená „existuje“. Např.

$$\forall x \in \mathbb{R} : x + 1 \le e^x$$

se čte

pro všechna reálná čísla x platí nerovnost $x + 1 \le e^x$.

Podobně,

$$\forall x, y \in \mathbb{R}, x < y, \exists z \in \mathbb{Q} : x < z < y$$

můžeme číst

pro jakékoliv dvě různá reálná čísla existuje racionální číslo ležící mezi nimi.

Modulo

Modulo, zapisované výrazem „$a \bmod n$“, udává zbytek při dělení celého čísla a přirozeným číslem n. Například

$$\begin{aligned}
17 \bmod 7 &= 3, \\
-17 \bmod 7 &= 4.
\end{aligned}$$

Faktoriál

Faktoriál přirozeného čísla n je součin přirozených čísel $1, 2, \ldots, n$ a značí se $n!$. Tedy

$$n! = 1 \cdot 2 \cdot 3 \cdot \ldots \cdot n.$$

Body a vektory

Pojem bod se v základním kurzu z lineární algebry moc nepoužívá. Bod, jakožto jako n-tice reálných čísel $(v_1, \ldots, v_n)$, se algebraicky chová stejně jako aritmetický vektor (definice 2.3). Oba pojmy se interpretují různě až z geometrického hlediska, kdy nenulovému vektoru odpovídá určitý směr (mimochodem, slovo „vektor“, stejně jako „vehikl“, pochází z latinského *vehere*, tedy „vézt“).

Zobrazení

Zobrazení f z množiny $\mathcal{A}$ do množiny $\mathcal{B}$ se značí $f\colon \mathcal{A} \to \mathcal{B}$. Pro každé $a \in \mathcal{A}$ je tedy $f(a)$ definováno a náleží do $\mathcal{B}$. Některé důležité typy zobrazení, se kterými budeme pracovat, jsou následující:

- Zobrazení $f\colon \mathcal{A} \to \mathcal{B}$ je *prosté*, pokud pro každé $a_1, a_2 \in \mathcal{A}$, $a_1 \neq a_2$, je $f(a_1) \neq f(a_2)$. Prosté zobrazení tedy nezobrazí dva různé prvky z množiny $\mathcal{A}$ na jeden prvek množiny $\mathcal{B}$.
- Zobrazení $f\colon \mathcal{A} \to \mathcal{B}$ je „*na*“, pokud pro každé $b \in \mathcal{B}$ existuje $a \in \mathcal{A}$ takové, že $f(a) = b$. Jinými slovy, zobrazení f pokryje celou množinu $\mathcal{B}$.
- Zobrazení $f\colon \mathcal{A} \to \mathcal{B}$ je *vzájemně jednoznačné (bijekce)*, pokud je prosté a „na“.
- Je-li $f\colon \mathcal{A} \to \mathcal{B}$ vzájemně jednoznačné, pak *inverzní zobrazení* k f je zobrazení $f^{-1}\colon \mathcal{B} \to \mathcal{A}$ definované $f^{-1}(b) = a$ pokud $f(a) = b$. Vzájemná jednoznačnost zobrazení f zajistí, že inverzní zobrazení f^{-1} je dobře definované v každém bodě množiny $\mathcal{B}$ a je také vzájemně jednoznačné (ověřte!).
- Jsou-li $f\colon \mathcal{A} \to \mathcal{B}$, $g\colon \mathcal{B} \to \mathcal{C}$, pak *složené zobrazení* f a g je zobrazení $g \circ f\colon \mathcal{A} \to \mathcal{C}$ definované předpisem $(g \circ f)(a) = g(f(a))$. Složením vzájemně jednoznačných zobrazení dostaneme vzájemně jednoznačné zobrazení (ověřte!).

Relace

Binární relace R na množině M je libovolná podmnožina kartézského součinu $M \times M$, tedy $R \subseteq M \times M$. Relace R je

- *reflexivní*, pokud $(x, x) \in R$ pro každé $x \in M$,
- *symetrická*, pokud pro každé $x, y \in M$ platí, že je-li $(x, y) \in R$, potom $(y, x) \in R$,
- *anti-symetrická*, pokud pro každé $x, y \in M$ platí, že je-li $(x, y), (y, x) \in R$, potom $x = y$,
- *transitivní*, pokud pro každé $x, y, z \in M$ platí, že je-li $(x, y), (y, z) \in R$, potom $(y, z) \in R$,
- *ekvivalence*, pokud je reflexivní, symetrická a transitivní.
- *(částečné) uspořádání*, pokud je reflexivní, anti-symetrická a transitivní.

Příkladem relace ekvivalence je

- rovnost čísel,
- rovnost zbytků přírozených čísel při dělení číslem n (například při dělení číslem 5 jsou v relaci čísla 2 a 7, ale nikoliv čísla 1 a 8),
- shodnost (či podobnost) geometrických objektů,
- shoda barev různých předmětů.

Příkladem částečného uspořádání je

- nerovnost $\leq$ mezi čísly,
- inkluze $\subseteq$ mezi množinami,
- dělitelnost na přirozených číslech (například 2 dělí 6, 6 dělí 30 a 2 dělí 30, ale 30 nedělí 2).

1.3 Reprezentace čísel

Vzhledem k omezené operační paměti nemůžeme v počítači reprezentovat všechna reálná čísla. Proto se čísla standardně reprezentují v tzv. tvaru s pohyblivou řádovou čárkou

$$m \times b^t,$$

kde b je daný základ vyjadřující, jakou číselnou soustavu používáme; většinou je $b = 2$. Hodnota mantisy m a exponentu t pak určují konkrétní hodnotu reprezentovaného čísla. Navíc vyjádření normujeme tak, aby mantisa m splňovala podmínku $1 \leq m < b$. Protože pro mantisu a exponent bývá na počítači omezená velikost zápisu, můžeme reprezentovat pouze konečně mnoho reálných čísel. Ty ostatní pak alespoň chceme vyjádřit nejbližším reprezentovatelným číslem. Tím přirozeně dochází ke ztrátě informace.

Například, číslo 1/3 se v dekadickém zápisu při přesnosti na čtyři plané cifry reprezentuje jako 3.333×10^{-1}, čili odpovídá číslu 0.3333. Takovéto zaokrouhlovací chyby se pak mohou dále akumulovat při vyhodnocování aritmetických operací. Například, součet čísel 1/3 a 1/3 vede na součet reprezentací 0.3333 a 0.3333 s výsledkem 0.6666. Nicméně, nejbližší reprezentovatelné číslo ke skutečnému součtu $1/3+1/3 = 2/3$ je 0.6667.

Vliv zaokrouhlovacích chyb a omezené reprezentace čísel proto musíme vzít v úvahu při návrhu algoritmů. Podrobněji se touto problematikou zabývá obor *numerická analýza*, ale i my se numerického hlediska dotkneme např. v sekci 3.5.1 a sekci 13.5.

1.4 Komplexní čísla

Komplexní číslo z zavádíme jako výraz $a + bi$, kde $a, b \in \mathbb{R}$ a imaginární jednotka i splňuje $i^2 = -1$. Zde a je *reálná část* komplexního čísla z a značí se $\mathsf{Re}(z)$, a b je *imaginární část* komplexního čísla z a značí se $\mathsf{Im}(z)$. Díky vlastnosti imaginární jednotky definujeme základní operace sčítání a násobení pro dvě komplexní čísla $z_1 = a + bi$, $z_2 = c + di$ takto:

$$\begin{aligned} z_1 + z_2 &= (a + c) + (b + d)i, \\ z_1 z_2 &= (ac - bd) + (cb + ad)i. \end{aligned}$$

Podobně pro odečítání. Pro $z_2 \neq 0$ pak vychází podíl

$$\frac{z_1}{z_2} = \frac{a + bi}{c + di} = \frac{a + bi}{c + di} \cdot \frac{c - di}{c - di} = \frac{ac + bd}{c^2 + d^2} + \frac{cb - ad}{c^2 + d^2} i.$$

Komplexní čísla mají také geometrickou interpretaci. Číslo $a + bi$ si lze představit jako bod (a, b) v rovině. Tato rovina se nazývá komplexní rovina (též Gaussova rovina). Množinu komplexních čísel značíme $\mathbb{C}$.

Pro komplexní číslo $z = a + bi$ pak zavádíme následující operace:

$$\text{komplexně sdružené číslo:} \quad \overline{z} := a - bi,$$

$$\text{absolutní hodnota:} \quad |z| := \sqrt{z\overline{z}} = \sqrt{a^2 + b^2}.$$

Absolutní hodnota čísla $z = a + bi$ vlastně určuje v komplexní rovině eukleidovskou vzdálenost bodu (a, b) od počátku. Komplexně sdružené číslo k číslu z pak představuje překlopení v komplexní rovině podle reálné osy.

Je jednoduchým cvičením ukázat, že komplexně sdružená čísla splňují vlastnosti:

$$\overline{z_1 + z_2} = \overline{z_1} + \overline{z_1},$$
$$\overline{z_1 \cdot z_2} = \overline{z_1} \cdot \overline{z_1}.$$

Na druhou stranu poznamenejme, že obecně $|z|^2 \neq z^2$.

Sčítání a násobení komplexních čísel mají také geometrický význam. Buď $v \in \mathbb{C}$ pevné komplexní číslo a uvažujme zobrazení $z \mapsto z + v$ nad komplexními čísly. V komplexní rovině toto zobrazení představuje posun ve směru vektoru $(\mathsf{Re}(v), \mathsf{Im}(v))$.

Buď $v \in \mathbb{C}$ opět pevné komplexní číslo a uvažujme zobrazení $z \mapsto zv$ nad komplexními čísly. Je-li v reálné číslo (tedy $\mathsf{Im}(v) = 0$), pak zobrazení představuje škálování s násobkem v. Je-li v komplexní a $|v| = 1$, pak zobrazení představuje otočení o úhel α, který v komplexní rovině svírá v s reálnou osou. V obecném případě se kombinují obě vlastnosti dohromady, tedy zobrazení $z \mapsto zv$ otáčí o úhel α a škáluje s násobkem v.

1.5 Polynomy

Reálným polynomem stupně n je funkce $p(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0$, kde $a_0, \ldots, a_n \in \mathbb{R}$ a $a_n \neq 0$. Kromě reálných polynomů můžeme uvažovat polynomy s komplexními koeficienty, popř. nad jinými číselnými obory (o tom až později).

Polynomy můžeme sčítat, odčítat, násobit a dělit se zbytkem. Buďte $p(x) = a_n x^n + \ldots + a_1 x + a_0$, $q(x) = b_m x^m + \ldots + b_1 x + b_0$ dva polynomy a nechť bez újmy na obecnosti $n \geq m$. Pak máme operace

- Sčítání:
$$p(x) + q(x) = a_n x^n + \ldots + a_{m+1} x^{m+1} + (a_m + b_m) x^m + \ldots + (a_1 + b_1) x + (a_0 + b_0).$$
- Násobení:
$$p(x)q(x) = a_n b_m x^{n+m} + \ldots + (a_k b_0 + a_{k-1} b_1 + \ldots + a_0 b_k) x^k + \ldots + a_0 b_0,$$
kde a_k pro $k > n$ a b_k pro $k > m$ definujeme jako 0.
- Dělení se zbytkem: Existují jednoznačně polynom $r(x)$ stupně $n-m$ a polynom $s(x)$ stupně menšího než m tak, že $p(x) = r(x)q(x) + s(x)$. Zde $r(x)$ představuje podíl, a $s(x)$ představuje zbytek. Například, dělíme-li polynom $p(x) = x^3$ polynomem $q(x) = x^2 - x$, dostaneme polynom $r(x) = x+1$ a zbytek $s(x) = x$, tedy
$$p(x) = (x+1)q(x) + x.$$

Kořeny

Kořen polynomu $p(x)$ je taková hodnota $x^* \in \mathbb{R}$, že $p(x^*) = 0$. Například, $p(x) = x^2 - 1$ má kořeny 1 a -1. Polynom $p(x) = x^2 + 1$ nemá reálný kořen, ale má dva komplexní, i a $-i$. Gaussova základní věta algebry z roku 1799 říká, že aspoň jeden kořen, byť komplexní, vždy existuje.

Věta 1.1 (Základní věta algebry)**.** *Každý polynom s komplexními koeficienty má alespoň jeden komplexní kořen.*

Důkaz. Důkazů existuje celá řada a žádný není zcela elementární. Myšlenkově snadno uchopitelný je důkaz autorů Melane & Birkhof a základní idea je následující. Uvažujme obraz kružnice v komplexní rovině se středem v počátku a poloměrem r při zobrazení $x \mapsto p(x)$. Je-li r hodně blízko nuly, je obrazem uzavřená křivka kolem bodu a_0. Naopak, je-li r dost velké, pak $p(x) \approx a_n x^n$ a obrazem křivka probíhající přibližně kolem kružnice se středem v počátku a poloměrem $a_n r^n$. Postupným spojitým zvětšováním r od nuly nakonec musí někde obraz protnout počátek, což odpovídá kořenu. □

Je-li x_1 kořen polynomu $p(x)$, pak $p(x)$ je dělitelný členem $(x - x_1)$ beze zbytku a podíl je polynom stupně $n-1$. Ten má podle základní věty algebry kořen x_2, opět můžeme beze zbytku dělit členem $(x-x_2)$ atd. až snížíme stupeň polynomu na nulu. Každý polynom tudíž lze zapsat jako $p(x) = a_n(x-x_1)\ldots(x-x_n)$, kde $x_1, \ldots, x_n$ jsou jeho kořeny. Dalším důsledkem je, že polynom stupně n má právě n kořenů, pokud započítáváme i násobnosti.

Poznámka 1.2. Nyní víme, že každý polynom má kořen, ale zatím není jasné, jak ho určit. Kořeny polynomů druhého stupně snadno najdeme podle známého vzorečku $x_{1,2} = \frac{1}{2a_2}(-a_1 \pm \sqrt{a_1^2 - 4a_2a_0})$. Pro kořeny polynomu třetího stupně existují také vzorce, tzv. Cardanovy, ale již mnohem komplikovanější. Důležitým zjištěním bylo, když roku 1824 přišel Abel na veřejnost s tím, že pro polynomy stupňů vyšších než 4 obecně žádný vzoreček na výpočet kořenů nemůže existovat. Veškeré praktické metody jsou tedy pouze iterační, kdy kořeny pouze aproximujeme, ale v jistém iteračním procesu aproximaci vylepšujeme na libovolnou přesnost.

1.6 Analytická geometrie

Přímka v rovině

Rovnicový popis přímky v rovině je
$$a_1 x_1 + a_2 x_2 = b,$$
kde a_1, a_2, b jsou daná reálná čísla taková, že alespoň jedno z a_1, a_2 je nenulové; toto vyjadřujeme vztahem $(a_1, a_2) \neq (0, 0)$. Potom všechny body (x_1, x_2) splňující rovnici představují přímku v rovině, a naopak každá přímka se dá vyjádřit tímto způsobem pro vhodné koeficienty $a_1, a_2, b \in \mathbb{R}$. Vektoru (a_1, a_2) se říká *normálový vektor* a je kolmý na přímku.

Parametrický popis přímky v rovině je

$$(x_1, x_2) = (b_1, b_2) + t \cdot (v_1, v_2), \quad t \in \mathbb{R},$$

kde (b_1, b_2) je daný bod přímky a $(v_1, v_2) \neq (0, 0)$ směrový vektor přímky. Libovolný bod přímky lze potom vyjádřit jako $(b_1, b_2) + t \cdot (v_1, v_2)$ pro vhodné reálné číslo t. Na obrázku dole máme znázorněny body odpovídající $t = 1$, $t = 0$ a $t = -2$.

Přímka v prostoru

Parametrický popis přímky v prostoru je

$$(x_1, x_2, x_3) = (b_1, b_2, b_3) + t \cdot (v_1, v_2, v_3), \quad t \in \mathbb{R},$$

kde (b_1, b_2, b_2) je opět daný bod přímky a $(v_1, v_2, v_3) \neq (0, 0, 0)$ její směrový vektor. Tím, jak parametr t prochází všechna reálná čísla postupně projdeme všechny body přímky.

Takovéto parametrické vyjádření přímky funguje i ve vyšších dimenzích. V n-dimenzionálním prostoru tedy libovolnou přímku můžeme vyjádřit

$$(x_1, \ldots, x_n) = (b_1, \ldots, b_n) + t \cdot (v_1, \ldots, v_n), \quad t \in \mathbb{R},$$

kde $(b_1, \ldots, b_n)$ je daný bod na přímce a $(v_1, \ldots, v_3) \neq (0, \ldots, 0)$ její směrový vektor. Směrový vektor je až na násobek určený jednoznačně, zatímco bod $(b_1, \ldots, b_n)$ lze zvolit na přímce libovolně.

Rovina v prostoru

Jedna rovnice

$$a_1x_1 + a_2x_2 + a_3x_3 = b,$$

kde $(a_1, a_2, a_3) \neq (0, 0, 0)$, tentokrát v prostoru nepopisuje již přímku, ale rovinu. Vektor (a_1, a_2, a_3) je její normálový vektor, tedy vektor kolmý na tuto rovinu. Tento normálový vektor je až na násobek určený jednoznačně.

Abychom popsali rovnicově přímku v prostoru, potřebujeme k tomu již dvě rovnice

$$a_1x_1 + a_2x_2 + a_3x_3 = b_1,$$
$$a_1'x_1 + a_2'x_2 + a_3'x_3 = b_2.$$

Nyní musíme předpokládat nejenom to, že normály jsou nenulové, to jest $(a_1, a_2, a_3) \neq (0, 0, 0)$ a $(a_1', a_2', a_3') \neq (0, 0, 0)$, ale navíc nesmí udávat stejný směr. To nám zaručí, že roviny, které dané rovnice popisují, nejsou rovnoběžné, a tudíž jejich průnikem je přímka.

1.7 Optimalizace

Buď $f\colon \mathbb{R}^n \to \mathbb{R}$ reálná funkce, která přiřazuje každému bodu $x = (x_1, \ldots, x_n)$ reálné číslo $f(x)$. Buď $M \subseteq \mathbb{R}^n$ množina bodů. Pak úloha *optimalizace* je

$$\min f(x) \text{ za podmínky } x \in M. \tag{1.1}$$

Chceme tedy najít minimální hodnotu funkce $f(x)$ na množině M. Jinými slovy, hledáme takový bod $x \in M$, že $f(x) \leq f(y)$ pro všechna $y \in M$.

Příklad 1.3. Uvažujme funkci $f(x_1, x_2) = x_1 - x_2$ a množinu bodů M v rovině zadanou omezeními

$$x_1 + 2x_2 \leq 4,\ x_1 \geq 0,\ x_2 \geq 0.$$

Množina M představuje trojúhelník s vrcholy $(0, 0)$, $(4, 0)$ a $(0, 2)$. Minimální hodnota funkce se nabyde v bodě $(0, 2)$, což je hledané optimální řešení optimalizační úlohy (1.1).

□

Kapitola 2

Soustavy lineárních rovnic

2.1 Základní pojmy

Soustavy lineárních rovnic patří mezi základní (algebraické) úlohy a setkáme se s nimi skoro všude – pokud nějaký problém nevede na soustavu rovnic přímo, tak se soustavy rovnic často objeví jako jeho podproblém.

Příklad 2.1 ([Meyer, 2000])**.** Nejstarší zaznamenaná úloha na soustavy rovnic z čínské knihy Chiu-chang Suan-shu (ca 200 př.n.l.):

> *Tři snopy dobrého obilí, dva snopy průměrného a jeden podřadného se prodávají celkem za 39 dou. Dva snopy dobrého obilí, tři průměrného a jeden podřadného se prodávají za 34 dou. Jeden snop dobrého obilí, dva průměrného a tři podřadného se prodávají za 26 dou. Jaká je cena za jeden snop dobrého / průměrného / podřadného obilí?*

Zapsáno dnešní matematikou, dostáváme soustavu rovnic

$$\begin{aligned} 3x + 2y + z &= 39, \\ 2x + 3y + z &= 34, \\ x + 2y + 3z &= 26, \end{aligned}$$

kde x, y, z jsou neznámé pro ceny za jeden snop dobrého / průměrného / podřadného obilí. □

Soustavy rovnic budeme zapisovat maticově, proto nejprve zavedeme pojem *matice*.

Definice 2.2 (Matice)**.** Reálná *matice* typu $m \times n$ je obdélníkové schema (tabulka)

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}.$$

Prvek na pozici (i, j) matice A (tj. v i-tém řádku a j-tém sloupci) značíme a_{ij} nebo A_{ij}. Množinu všech reálných matic typu $m \times n$ značíme $\mathbb{R}^{m\times n}$; podobně pro komplexní, racionální, atd. Je-li $m = n$, potom matici nazýváme *čtvercovou*.

Definice 2.3 (Vektor)**.** Reálný n-rozměrný aritmetický sloupcový *vektor* je matice typu $n \times 1$

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

a řádkový vektor je matice typu $1 \times n$

$$x = (x_1, \dots, x_n).$$

Standardně, pokud není řečeno jinak, uvažujeme vektory sloupcové. Množina všech n-rozměrných vektorů se značí $\mathbb{R}^n$ (namísto $\mathbb{R}^{n\times 1}$). Pro odlišení značíme obecné matice velkými písmeny a vektory malými písmeny.

Definice 2.4 ($*$ notace)**.**
i-tý řádek matice A se značí: $A_{i*} = (a_{i1}, a_{i2}, \ldots, a_{in})$.
j-tý sloupec matice A se značí: $A_{*j} = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{nj} \end{pmatrix}$.

Matici $A \in \mathbb{R}^{m\times n}$ tudíž můžeme rozepsat po sloupcích a po řádcích následovně

$$A = \begin{pmatrix} | & | & & | \\ A_{*1} & A_{*2} & \ldots & A_{*n} \\ | & | & & | \end{pmatrix} = \begin{pmatrix} — & A_{1*} & — \\ — & A_{2*} & — \\ & \vdots & \\ — & A_{m*} & — \end{pmatrix}.$$

Definice 2.5 (Soustava lineárních rovnic)**.** Mějme soustavu m lineárních rovnic o n neznámých

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= b_2, \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n &= b_m, \end{aligned} \tag{2.1}$$

kde a_{ij}, b_i jsou dané koeficienty a $x_1, \ldots, x_n$ jsou neznámé. *Řešením* rozumíme každý vektor $x \in \mathbb{R}^n$ vyhovující všem rovnicím.

Zápis (2.1) v zásadě obsahuje trochu nadbytečné informace. Podstatné jsou jen rozměry soustavy a jednotlivé koeficienty. To nás vede na maticovou formu zápisu soustavy, obsahující jen nezbytné informace.

Definice 2.6 (Matice soustavy)**.** *Matice soustavy* je matice

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix}$$

a *rozšířená matice soustavy* je

$$(A \mid b) = \left(\begin{array}{cccc|c} a_{11} & a_{12} & \ldots & a_{1n} & b_1 \\ a_{21} & a_{22} & \ldots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} & b_m \end{array}\right).$$

Svislá čára v rozšířené matice soustavy symbolizuje rovnost mezi levou a pravou stranou soustavy, Z formálního hlediska sice není nutné ji zobrazovat, ale pomáhá mnemotechnicky chápat význam této matice.

Poznamenejme, že rozšířená matice soustavy plně popisuje soustavu rovnic; řádky odpovídají rovnicím, sloupce nalevo postupně proměnným $x_1, \ldots, x_n$ a poslední sloupec hodnotám na pravé straně soustavy. Tudíž můžeme soustavu rovnic zadávat i maticovém tvaru.

Kupříkladu soustava z příkladu 2.1 by se maticově zapsala jako

$$\left(\begin{array}{ccc|c} 3 & 2 & 1 & 39 \\ 2 & 3 & 1 & 34 \\ 1 & 2 & 3 & 26 \end{array}\right).$$

Poznámka 2.7 (Geometrický význam soustavy rovnic)**.** Pro jednoduchost uvažujme nejprve případ $m = n = 2$, tedy dvě rovnice o dvou neznámých

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= b_1, \\ a_{21}x_1 + a_{22}x_2 &= b_2. \end{aligned}$$

Za obecných předpokladů ($a_{11} \neq 0$ nebo $a_{12} \neq 0$) popisuje první rovnice přímku v rovině $\mathbb{R}^2$, a stejně tak druhá rovnice popisuje nějakou přímku. Řešení soustavy leží tedy v průniku obou přímek.

Podobně pro $n = 3$, každá rovnice s aspoň jedním nenulovým koeficientem popisuje rovinu v prostoru $\mathbb{R}^3$ a řešení představuje průnik těchto rovin. Obecně pro libovolné n, rovnice určují tzv. nadroviny (srov. kapitola 7) a řešení soustavy hledáme v jejich průniku.

Definice 2.8 (Elementární řádkové úpravy)**.** Elementární řádkové úpravy matice jsou

1. vynásobení i-tého řádku reálným číslem $\alpha \neq 0$ (tj. vynásobí se všechny prvky řádku),
2. přičtení α-násobku j-tého řádku k i-tému, přičemž $i \neq j$ a $\alpha \in \mathbb{R}$,
3. výměna i-tého a j-tého řádku.

Poznámka 2.9. Ve skutečnosti výše zmíněné úpravy nejsou zas tak elementární. U druhé řádkové úpravy vystačíme jen s $\alpha = 1$ a třetí úpravu lze simulovat pomocí předchozích dvou. Schematicky

$$\begin{pmatrix} \dots \\ A_{i*} \\ \dots \\ A_{j*} \\ \dots \end{pmatrix} \sim \begin{pmatrix} \dots \\ A_{i*} \\ \dots \\ A_{j*} - A_{i*} \\ \dots \end{pmatrix} \sim \begin{pmatrix} \dots \\ A_{j*} \\ \dots \\ A_{j*} - A_{i*} \\ \dots \end{pmatrix} \sim \begin{pmatrix} \dots \\ A_{j*} \\ \dots \\ -A_{i*} \\ \dots \end{pmatrix} \sim \begin{pmatrix} \dots \\ A_{j*} \\ \dots \\ A_{i*} \\ \dots \end{pmatrix}.$$

Tvrzení 2.10. *Elementární řádkové operace zachovávají množinu řešení soustavy.*

Důkaz. Základní myšlenkou je ukázat, že elementární úpravou se množina řešení nemění. Elementární úpravou neztratíme žádné řešení, protože pokud je x řešením před úpravou, je i po úpravě. A naopak, úpravou žádné řešení nepřibyde. To se nahlédne ze symetrie, protože každá úprava má svoji inverzní úpravu – vhodnou elementární úpravou můžu dojít zpět k původnímu tvaru soustavy. Detailní důkaz jednotlivých úprav ponecháváme za cvičení. □

2.2 Gaussova eliminace

Nyní se přesuneme k tomu, jak soustavy rovnic řešit. Než představíme obecný algoritmus a jeho maticový zápis, ukážeme nejprve myšlenku Gaussovy eliminace na příkladu.

Příklad 2.11. Uvažujme soustavu lineárních rovnic

$$\begin{array}{rcrcrcl} x_1 & + & 2x_2 & + & 3x_3 & = & 32, \\ x_1 & + & x_2 & + & 2x_3 & = & 21, \\ 3x_1 & + & x_2 & + & 3x_3 & = & 35. \end{array}$$

Nejprve eliminujeme proměnnou x_1 z druhé a třetí rovnice. Odečtením první rovnice od druhé dostaneme

$$\begin{array}{rcrcrcr} x_1 & + & 2x_2 & + & 3x_3 & = & 32, \\ & - & x_2 & - & x_3 & = & -11, \\ 3x_1 & + & x_2 & + & 3x_3 & = & 35, \end{array}$$

a odečtením trojnásobku první rovnice od třetí získáme

$$\begin{array}{rcrcrcr} x_1 & + & 2x_2 & + & 3x_3 & = & 32, \\ & - & x_2 & - & x_3 & = & -11, \\ & - & 5x_2 & - & 6x_3 & = & -61. \end{array}$$

Nyní eliminujeme proměnnou x_2 z třetí rovnice. Odečtením pětinásobku druhé rovnice od třetí dostane soustava výsledný tvar

$$\begin{array}{rcrcrcr} x_1 & + & 2x_2 & + & 3x_3 & = & 32, \\ & - & x_2 & - & x_3 & = & -11, \\ & & & - & x_3 & = & -6. \end{array}$$

Tím končí první fáze (tzv. dopředná eliminace), kdy soustavu upravujeme do jednoduššího tvaru eliminací proměnných. Následuje druhá fáze, nazývaná zpětná substituce. Ze třetí rovnice okamžitě vidíme hodnotu třetí proměnné, $x_3 = 6$. Dosadíme tuto hodnotu do druhé rovnice

$$-x_2 - 6 = -11,$$

ze které určíme hodnotu druhé proměnné, $x_2 = 5$. Nakonec hodnoty x_1, x_2 dosadíme do první rovnice

$$x_1 + 2 \cdot 5 + 3 \cdot 6 = 32,$$

a dopočítáme hodnotu první proměnné, $x_1 = 4$. Výsledné řešení soustavy je tedy $(x_1, x_2, x_3) = (4, 5, 6)$. □

Základní myšlenka dvou metod, které popíšeme, je transformace rozšířené matice soustavy pomocí elementárních úprav na jednodušší matici, ze které řešení snadno vyčteme. Ten jednodušší tvar matice se nazývá *odstupňovaný tvar matice*, v angličtině „row echelon form“ (REF).

Definice 2.12 (Odstupňovaný tvar matice)**.** Matice $A \in \mathbb{R}^{m \times n}$ je v řádkově odstupňovaném tvaru, pokud existuje r takové, že platí

- řádky $1, \dots, r$ jsou nenulové (tj. každý obsahuje aspoň jednu nenulovou hodnotu),
- řádky $r+1, \dots, m$ jsou nulové,

a navíc označíme-li $p_i = \min\{j;\ a_{ij} \neq 0\}$, tak platí

- $p_1 < p_2 < \dots < p_r$.

K odstupňovanému tvaru se vztahují některé zásadní pojmy: Pozice $(1, p_1)$, $(2, p_2), \dots,$ (r, p_r) se nazývají *pivoty*, sloupce p_1, $p_2, \dots,$ p_r se nazývají *bázické* a ostatní sloupce *nebázické* (význam bude zřejmý později).

Schematické znázornění odstupňovaného tvaru. Pivoty jsou pozice černých teček.

Definice 2.13. *Hodností* matice A rozumíme počet nenulových řádků po převodu do odstupňovaného tvaru a značíme $\operatorname{rank}(A)$.

Hodnost matice je tedy rovna počtu pivotů (tj. číslu r) po převedení do odstupňovaného tvaru. I když odstupňovaný tvar není jednoznačný, pozice pivotů jednoznačné jsou (ukážeme si později ve větě 3.43). Proto je pojem hodnosti[1)] dobře definován, i když to v tuto chvíli ještě není zřejmé.

Každou matici lze převést elementárními řádkovými úpravami do odstupňovaného tvaru. Například následující algoritmus převádí matici do odstupňovaného tvaru po nejvýše $\min(m,n)$ iteracích hlavního cyklu; důkaz správnosti se udělá matematickou indukcí podle počtu sloupců a rozborem. Pochopitelně, existují i různé varianty.

Algoritmus 2.14 (REF(A))**.** Buď $A \in \mathbb{R}^{m\times n}$.

1: $i := 1$, $j := 1$,

2: **if** $a_{k\ell} = 0$ pro všechna $k \geq i$ a $\ell \geq j$ **then** konec,

3: $j := \min\{\ell;\ l \geq j,\ a_{k\ell} \neq 0$ pro nějaké $k \geq i\}$, //přeskočíme nulové podsloupečky

4: urči k takové, že $a_{kj} \neq 0$, $k \geq i$ a vyměň řádky A_{i*} a A_{k*},

//nyní je na pozici pivota hodnota $a_{ij} \neq 0$

5: pro všechna $k > i$ polož $A_{k*} := A_{k*} - \frac{a_{kj}}{a_{ij}} A_{i*}$, //2. elementární úprava

6: polož $i := i + 1$, $j := j + 1$, a jdi na krok 2.

Krok 2 je ukončovací: Skončíme, když podmatice vpravo dole od aktuální pozice (i,j) je nulová. Speciálně, konec nastane také když $i = m$ nebo $j = n$. V kroku 3 se posouváme indexem j na nejbližší pozici napravo tak, aby v j-tém podsloupečku o prvcích $a_{i,j}, a_{i+1,j}, \ldots, a_{m,j}$ existoval nenulový prvek. Celkem neurčitě jsme definovali index k v kroku 4. Teoreticky si můžeme zvolit libovolně, v praxi se doporučuje kandidát a_{kj} s maximální absolutní hodnotou, tzv. *parciální pivotizace*, protože má lepší numerické vlastnosti při řešení soustav rovnic na počítačích. V kroku 5 zkráceně popisujeme druhou elementární úpravu, kdy od k-tého řádku odečítáme $\frac{a_{kj}}{a_{ij}}$-násobek i-tého řádku.

Důkaz správnosti algoritmu uvádět nebudeme, ale povšimneme si alespoň, že v kroku 5 vynulujeme všechny prvky pod pivotem (i,j). Pro každé $k > i$ je hodnota a_{kj} po úpravě rovna j-tému prvku řádku $A_{k*} := A_{k*} - \frac{a_{kj}}{a_{ij}} A_{i*}$, a ten má hodnotu $a_{kj} - \frac{a_{kj}}{a_{ij}} a_{ij} = 0$.

Příklad 2.15. Ukázka převodu matice na odstupňovaný tvar, zakroužkované hodnoty ukazují aktuální pozice i, j, často to jsou pozice pivotů. U každého kroku označujeme, který krok algoritmu 2.14 se zrovna použije:

$$\begin{pmatrix} 2 & 2 & -1 & 5 \\ 4 & 5 & 0 & 9 \\ 0 & 1 & 2 & 2 \\ 2 & 4 & 3 & 7 \end{pmatrix} \overset{5}{\sim} \begin{pmatrix} 2 & 2 & -1 & 5 \\ 0 & 1 & 2 & -1 \\ 0 & 1 & 2 & 2 \\ 2 & 4 & 3 & 7 \end{pmatrix} \overset{5}{\sim} \begin{pmatrix} 2 & 2 & -1 & 5 \\ 0 & 1 & 2 & -1 \\ 0 & 1 & 2 & 2 \\ 0 & 2 & 4 & 2 \end{pmatrix} \sim$$

$$\overset{6}{\sim} \begin{pmatrix} 2 & 2 & -1 & 5 \\ 0 & 1 & 2 & -1 \\ 0 & 1 & 2 & 2 \\ 0 & 2 & 4 & 2 \end{pmatrix} \overset{5}{\sim} \begin{pmatrix} 2 & 2 & -1 & 5 \\ 0 & 1 & 2 & -1 \\ 0 & 0 & 0 & 3 \\ 0 & 2 & 4 & 2 \end{pmatrix} \sim$$

$$\overset{5}{\sim} \begin{pmatrix} 2 & 2 & -1 & 5 \\ 0 & 1 & 2 & -1 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 4 \end{pmatrix} \overset{6}{\sim} \begin{pmatrix} 2 & 2 & -1 & 5 \\ 0 & 1 & 2 & -1 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 4 \end{pmatrix} \sim$$

$$\overset{3}{\sim} \begin{pmatrix} 2 & 2 & -1 & 5 \\ 0 & 1 & 2 & -1 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 4 \end{pmatrix} \overset{5}{\sim} \begin{pmatrix} 2 & 2 & -1 & 5 \\ 0 & 1 & 2 & -1 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

□

[1)] Pojem hodnosti zavedl německý matematik Ferdinand Georg Frobenius, ale podobný koncept už byl používán dříve např. anglickým matematikem Sylvestrem.

Teď už máme vše připraveno na Gaussovu eliminaci pro řešení soustav rovnic.

Algoritmus 2.16 (Gaussova eliminace[2])**.** Buď dána soustava rovnic $(A \mid b)$, kde $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$. Převedeme rozšířenou matici soustavy $(A \mid b)$ na odstupňovaný tvar $(A' \mid b')$ a označíme $r = \operatorname{rank}(A \mid b)$. Nyní nastala právě jediná z následujících tří situací:

(A) *Soustava nemá řešení.*

Tato situace nastane v případě, že poslední sloupec je bázický, čili v posledním sloupci je pivot. Ekvivalentně vyjádřeno, $\operatorname{rank}(A) < \operatorname{rank}(A \mid b)$, nebo ještě jinak $\operatorname{rank}(A') < \operatorname{rank}(A' \mid b')$.

Důkaz. r-tý řádek soustavy (v REF tvaru) má tvar

$$0x_1 + 0x_2 + \ldots + 0x_n = b'_r.$$

Vzhledem k tomu, že $b'_r \neq 0$, neboť je to pivot, soustava nemůže mít řešení. □

(B) *Soustava má alespoň jedno řešení.*

Tato situace naopak nastane, pokud poslední sloupec je nebázický, čili neobsahuje pivota. Jinými slovy to znamená, že $\operatorname{rank}(A) = \operatorname{rank}(A \mid b)$, neboli $\operatorname{rank}(A') = \operatorname{rank}(A' \mid b')$. Rozlišíme dva podppřípady, odpovídající tomu, jestli má soustava jediné řešení nebo nekonečně mnoho řešení.

(B1) *Soustava má jediné řešení.*

Jediné řešení existuje pokud $r = n$, to znamená, že počet proměnných je roven počtu pivotů. V každém sloupci, kromě nejpravějšího, se tudíž nachází pivot. Řešení nyní najdeme tzv. *zpětnou substitucí*: Postupně pro $k = n, n-1, \ldots, 1$ v tomto pořadí dosadíme

$$x_k := \frac{b'_k - \sum_{j=k+1}^{n} a'_{kj} x_j}{a'_{kk}}.$$

Důkaz. Soustava v odstupňovaném tvaru má podobu

$$\left(\begin{array}{ccc|c} a'_{11} & \ldots & a'_{1n} & b'_1 \\ & \ddots & \vdots & \vdots \\ 0 & & a'_{nn} & b'_n \\ 0 & \ldots & 0 & 0 \end{array}\right).$$

V rovnicovém podobě má soustava tvar

$$\begin{aligned} a'_{11}x_1 + a'_{12}x_2 + \ldots + a'_{1n}x_n &= b'_1, \\ a'_{22}x_2 + \ldots + a'_{2n}x_n &= b'_2, \\ &\vdots \\ a'_{kk}x_k + \ldots + a'_{kn}x_n &= b'_k, \\ &\vdots \\ a'_{nn}x_n &= b'_n \end{aligned}$$

(popřípadě ještě nějaké rovnice typu $0x_1 + \ldots + 0x_n = 0$, které lze vynechat). Hodnotu neznámé x_n spočítáme z poslední rovnice, tu dosadíme do předposlední a dopočítáme x_{n-1} atd. až nakonec spočítáme hodnotu x_1. Máme tedy jediné řešení, které je dobře definované, neboť $a'_{kk} \neq 0$. □

[2] Carl Friedrich Gauss, z r. 1810.

(B2) *Soustava má nekonečně mnoho řešení.*
Tento případ nastane, pokud $r < n$. To znamená, že kromě nejpravějšího sloupečku je v matici alespoň jeden další nebázický sloupec. Množinu všech nekonečně mnoho řešení[3)] popíšeme parametricky. Jako *bázické proměnné* označíme ty, které odpovídají bázickým sloupcům, tj. x_{p_1}, x_{p_2}, $\ldots$,x_{p_r}, a jako *nebázické proměnné* ty zbývající. Potom nebázické proměnné budou parametry, které mohou nabývat libovolných reálných hodnot a pomocí nichž dopočítáme bázické proměnné opět zpětnou substitucí: Postupně pro $k = r, r-1, \ldots, 1$ v tomto pořadí dosadíme

$$x_{p_k} := \frac{b'_k - \sum_{j=p_k+1}^{n} a'_{kj} x_j}{a'_{kp_k}}.$$

Příklad 2.17. Vyřešme Gaussovou eliminací následující soustavu. Nejprve převedeme rozšířenou matici soustavy na odstupňovaný tvar

$$\left(\begin{array}{cccc|c} 2 & 2 & -1 & 5 & 1 \\ 4 & 5 & 0 & 9 & 3 \\ 0 & 1 & 2 & 2 & 4 \\ 2 & 4 & 3 & 7 & 7 \end{array}\right) \overset{REF}{\sim} \left(\begin{array}{cccc|c} 2 & 2 & -1 & 5 & 1 \\ 0 & 1 & 2 & -1 & 1 \\ 0 & 0 & 0 & 3 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

Zpětná substituce probíhá takto:

1. $x_4 = 1$
2. x_3 je volná (nebázická) proměnná
3. $x_2 = 1 + x_4 - 2x_3 = 2 - 2x_3$
4. $x_1 = \frac{1}{2}(1 - 5x_4 + x_3 - 2x_2) = -4 + \frac{5}{2}x_3$

Všechna řešení jsou tvaru (zapsáno v řádku)

$$(-4 + \tfrac{5}{2}x_3, 2 - 2x_3, x_3, 1), \text{ kde } x_3 \in \mathbb{R}.$$

Tato řešení můžeme vyjádřit ekvivalentně ve tvaru

$$(-4, 2, 0, 1) + x_3(\tfrac{5}{2}, -2, 1, 0), \text{ kde } x_3 \in \mathbb{R},$$

jehož význam vyplyne v následující kapitole. Z tohoto vyjádření rovněž vyplyne, že množina řešení představuje přímku v $\mathbb{R}^4$ se směrnicí $(\frac{5}{2}, -2, 1, 0)$ a procházející bodem $(-4, 2, 0, 1)$. □

2.3 Gaussova–Jordanova eliminace

Druhý algoritmus, který si uvedeme, je Gaussova–Jordanova eliminace. Namísto odstupňovaného tvaru používá ještě specifičtější redukovaný odstupňovaný tvar, v angličtině „reduced row echelon form“ (RREF).

Definice 2.18 (Redukovaný odstupňovaný tvar matice)**.** Matice $A \in \mathbb{R}^{m \times n}$ je v redukovaném řádkově odstupňovaném tvaru, pokud je v REF tvaru a navíc platí

- $a_{1p_1} = a_{2p_2} = \ldots = a_{rp_r} = 1$,
- pro každé $i = 1, \ldots, r$ je $a_{1p_i} = a_{2p_i} = \ldots = a_{i-1,p_i} = 0$.

Schematické znázornění redukovaného odstupňovaného tvaru. Narozdíl od REF jsou pivoty navíc znormovány na jedničku a nad nimi jsou samé nuly.

Algoritmus, který převede libovolnou matici do RREF tvaru je podobný tomu pro REF. Jediná změna je v kroku 5, který nahradíme dvěma novými.

[3)] Toto platí, pokud pracujeme nad reálnými čísly. Nad konečnými tělesy (sekce 4.3) se postupuje stejně, ale počet řešení bude konečný.

Algoritmus 2.19 (RREF(A)). Buď $A \in \mathbb{R}^{m\times n}$.

1: $i := 1$, $j := 1$,
2: **if** $a_{k\ell} = 0$ pro všechna $k \geq i$ a $\ell \geq j$ **then** konec,
3: $j := \min\{l;\ l \geq j,\ a_{kl} \neq 0$ pro nějaké $k \geq i\}$, //přeskočíme nulové podsloupečky
4: urči $a_{kj} \neq 0$, $k \geq i$ a vyměň řádky A_{i*} a A_{k*}, //nyní je na pozici pivota hodnota $a_{ij} \neq 0$
5: polož $A_{i*} := \frac{1}{a_{ij}} A_{i*}$, //nyní je na pozici pivota hodnota $a_{ij} = 1$
6: pro všechna $k \neq i$ polož $A_{k*} := A_{k*} - a_{kj} A_{i*}$, //2. elementární úprava
7: polož $i := i + 1$, $j := j + 1$, a jdi na krok 2.

Příklad 2.20. Ukázka převodu matice na redukovaný odstupňovaný tvar, zakroužkované hodnoty ukazují aktuální pivoty:

$$\begin{pmatrix} 2 & 2 & -1 & 5 \\ 4 & 5 & 0 & 9 \\ 0 & 1 & 2 & 2 \\ 2 & 4 & 3 & 7 \end{pmatrix} \sim \begin{pmatrix} 1 & 1 & -0.5 & 2.5 \\ 4 & 5 & 0 & 9 \\ 0 & 1 & 2 & 2 \\ 2 & 4 & 3 & 7 \end{pmatrix} \sim \begin{pmatrix} 1 & 1 & -0.5 & 2.5 \\ 0 & 1 & 2 & -1 \\ 0 & 1 & 2 & 2 \\ 2 & 4 & 3 & 7 \end{pmatrix} \sim$$

$$\sim \begin{pmatrix} 1 & 1 & -0.5 & 2.5 \\ 0 & 1 & 2 & -1 \\ 0 & 1 & 2 & 2 \\ 0 & 2 & 4 & 2 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & -2.5 & 3.5 \\ 0 & 1 & 2 & -1 \\ 0 & 1 & 2 & 2 \\ 0 & 2 & 4 & 2 \end{pmatrix} \sim$$

$$\sim \begin{pmatrix} 1 & 0 & -2.5 & 3.5 \\ 0 & 1 & 2 & -1 \\ 0 & 0 & 0 & 3 \\ 0 & 2 & 4 & 2 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & -2.5 & 3.5 \\ 0 & 1 & 2 & -1 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 4 \end{pmatrix} \sim$$

$$\sim \begin{pmatrix} 1 & 0 & -2.5 & 3.5 \\ 0 & 1 & 2 & -1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 4 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & -2.5 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

□

Gaussova–Jordanova eliminace pak funguje následujícím způsobem.

Algoritmus 2.21 (Gaussova–Jordanova eliminace[4]). Buď dána soustava rovnic $(A \mid b)$, kde $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$. Převedeme rozšířenou matici soustavy na redukovaný odstupňovaný tvar $(A' \mid b')$ a označíme $r = \operatorname{rank}(A \mid b)$. Rozlišíme tři situace:

(A) *Soustava nemá řešení.*

Tento případ nastane, pokud je poslední sloupec bázický. Jinými slovy, poslední sloupec obsahuje pivota, čili $\operatorname{rank}(A) < \operatorname{rank}(A \mid b)$. Důkaz analogický jako pro Gaussovu eliminaci.

(B1) *Soustava má právě jedno řešení.*

Toto nastane, pokud je poslední sloupec je nebázický (neobsahuje pivota) a zároveň $r = n$. Tudíž sloupce $1, \ldots, n$ jsou bázické a poslední je nebázický. Tvrdíme, že jednoznačné řešení má tvar $(x_1, \ldots, x_n) = (b'_1, \ldots, b'_n)$.

Důkaz. Soustava v RREF tvaru má rovnicovou podobu

$$\begin{aligned} x_1 &= b'_1, \\ x_2 &= b'_2, \\ &\vdots \\ x_n &= b'_n \end{aligned}$$

[4] Wilhelm Jordan, z r. 1887.

(opět tam mohou být ještě nějaké rovnice typu $0x_1 + \ldots + 0x_n = 0$, které nemusíme uvažovat). Rovnice zde přímo určují řešení. □

(B2) *Soustava má nekonečně mnoho řešení.*

Pro tento případ musí být poslední sloupec nebázický a $r < n$. Množinu řešení popíšeme parametricky. Označíme nebázické proměnné x_i, $i \in N$, kde $N = \{1, \ldots, n\} \setminus \{p_1, \ldots, p_r\}$. Opět, nebázické proměnné budou parametry, které mohou nabývat libovolných reálných hodnot a pomocí nich dopočítáme bázické proměnné opět zpětnou substitucí (zde lze i dopřednou): Postupně pro $k = r, r-1, \ldots, 1$ v tomto pořadí dosadíme

$$x_{p_k} := b'_k - \sum_{j \in N,\ j > p_k} a'_{kj} x_j.$$

Poznámka 2.22 (Frobeniova věta)**.** Při odvozování Gaussovy a Gaussovy–Jordanovy eliminace jsme jako důsledek dostali důležitou větu lineární algebry nazývanou Frobeniova věta,[5)] která charakterizuje řešitelnost soustav rovnic pomocí hodností matice a rozšířené matice soustavy:

Soustava $(A \mid b)$ má (aspoň jedno) řešení právě tehdy, když $\operatorname{rank}(A) = \operatorname{rank}(A \mid b)$.

Později (poznámka 5.65), v kapitole věnované vektorovým prostorům k tomu získáme ještě jiný náhled.

Poznámka 2.23 (Gaussova versus Gaussova–Jordanova eliminace)**.** Čtenář se může ptát, proč jsme uváděli dvě poměrně podobné metody na řešení soustav rovnic. Gaussova eliminace má výhodu v tom, že je ca o 50% rychlejší než ta druhá, na druhou stranu Gaussovu–Jordanovu eliminaci (či spíše RREF tvar) budeme potřebovat při invertování matic (sekce 3.3).

2.4 Aplikace

Soustavy lineárních rovnic se v praktických problémech objevují velice často, zejména jako podúloha většího problému. Uvedeme jednu aplikaci, vedoucí přímo na soustavu rovnic. Další jsou zmíněné například v sekci 3.6.

Příklad 2.24. Uvažujme elektrický obvod jak je vyznačený na obrázku

Chceme-li určit hodnoty elektrických proudů I_1, I_2, I_3, využijeme fyzikálních zákonů:

1. *Ohmův zákon:* Napětí je rovno součinu proudu a odporu: $U = IR$,
2. *Kirchhoffův zákon o proudu:* Součet proudů vstupujících do uzlu se rovná součtu proudů z uzlu vystupujících.
3. *Kirchhoffův zákon o napětí:* Součet napětí ve smyčce je roven nule.

[5)] V angličtině větu nazývají Rouchého–Capelliho věta, v Rusku Kroneckerova–Capelliho věta a ve Španělsku Rouchého–Frobeniova věta. Můžete se setkat i s názvem Rouché-Fonteného věta. Inu, jiný kraj, jiný mrav.

Konkrétně, Kirchhoffův zákon o proudu dává rovnici $I_1 + I_2 - I_3 = 0$. Kirchhoffův zákon o napětí (spolu s Ohmovým zákonem) pro smyčku DABE dává rovnici $10I_1 - 10I_2 = 10$, a pro smyčku EBCF dává $10I_2 + 20I_3 = 5$. (Smyčku DABCFE již uvažovat nemusíme, neboť vyplývá z předchozích dvou.) Tím dostáváme soustavu rovnic

$$\left(\begin{array}{rrr|r} 1 & 1 & -1 & 0 \\ 10 & -10 & 0 & 10 \\ 0 & 10 & 20 & 5 \end{array}\right).$$

Vyřešením máme $I_1 = 0.7A$, $I_2 = -0.3A$, $I_3 = 0.4A$. □

Problémy

2.1. Jaké maximální absolutní hodnoty mohou nabývat složky řešení soustavy $(A \mid b)$ pokud nenulové prvky matice $A \in \mathbb{R}^{n\times n}$ jsou omezeny $q^{-1} \leq |a_{ij}| \leq q$ pro nějaké dané q?

2.2. Určete maximální počet aritmetických operací Gaussovy a Gaussovy–Jordanovy eliminace (tj. kolikrát sčítáme / odčítáme, a násobíme / dělíme reálná čísla) pro soustavu $n \times n$.

Shrnutí ke kapitole 2. Soustavy lineárních rovnic

Soustava lineárních rovnic je základní úlohou nejen lineární algebry, a umět vyřešit soustavu je znalost potřebná skoro pro každý složitější problém. Jedna lineární rovnice popisuje v prostoru proměnných tzv. nadrovinu. Množina řešení soustavy rovnic pak tedy geometricky představuje průnik několika nadrovin, čímž může být bod, přímka, rovina,... Řešit soustavu pak znamená rozhodnout, zda je soustava řešitelná, a pokud ano, vydat řešení. Pokud je řešení nekonečně mnoho, popisujeme je pomocí parametrů.

Předvedli jsme Gaussovu(–Jordanovu) eliminaci jako metodu, která vyřeší jakoukoli soustavu lineárních rovnic. Pozná, zda je soustava řešitelná, a pokud ano tak v parametrickém zápisu popíše všechna řešení. Metoda je založená na elementárních úpravách, které sice mění soustavu a matici, která ji reprezentuje, ale nemění množinu řešení. Elementárními úpravami převedeme matici do jednoduššího (odstupňovaného) tvaru, ze kterého se řešení snadno vyčte.

Kapitola 3

Matice

Matice jsme zavedli v minulé kapitole ke kompaktnímu zápisu soustav lineárních rovnic a popisu metod na jejich řešení. Matice však mají mnohem širší využití, a proto se na ně v této kapitole podíváme podrobněji. Zavedeme několik důležitých typů matic a základní operace, které umožní s maticemi lépe a jednodušeji zacházet.

3.1 Základní operace s maticemi

Definice 3.1 (Rovnost)**.** Dvě matice se rovnají, $A = B$, pokud mají stejné rozměry $m \times n$ a $A_{ij} = B_{ij}$ pro $i = 1, \dots, m$, $j = 1, \dots, n$.

Definice 3.2 (Součet)**.** Buď $A, B \in \mathbb{R}^{m\times n}$. Pak $A+B$ je matice typu $m\times n$ s prvky $(A+B)_{ij} = A_{ij}+B_{ij}$, $i = 1, \dots, m$, $j = 1, \dots, n$.

Definice 3.3 (Násobek)**.** Buď $\alpha \in \mathbb{R}$ a $A \in \mathbb{R}^{m\times n}$. Pak αA je matice typu $m \times n$ s prvky $(\alpha A)_{ij} = \alpha A_{ij}$, $i = 1, \dots, m$, $j = 1, \dots, n$.

Příklad 3.4 (Součet a násobek matic)**.**

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} + \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix} = \begin{pmatrix} 6 & 8 \\ 10 & 12 \end{pmatrix}, \quad 5\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 5 & 10 \\ 15 & 20 \end{pmatrix}. \qquad \square$$

Výše zmíněné operace umožňují zavést přirozeně i odčítání jako $A - B \coloneqq A + (-1)B$.

Speciální maticí je nulová matice, jejíž všechny prvky jsou nuly. Značíme ji 0 či $0_{m\times n}$ pro zdůraznění rozměru.

Tvrzení 3.5 (Vlastnosti součtu a násobků matic)**.** *Pro reálná čísla α, β a matice $A, B, C \in \mathbb{R}^{m\times n}$ platí*

(1) $A + B = B + A$ *(komutativita)*

(2) $(A + B) + C = A + (B + C)$ *(asociativita)*

(3) $A + 0 = A$

(4) $A + (-1)A = 0$

(5) $\alpha(\beta A) = (\alpha\beta)A$

(6) $1A = A$

(7) $\alpha(A + B) = \alpha A + \alpha B$ *(distributivita)*

(8) $(\alpha + \beta)A = \alpha A + \beta A$ *(distributivita)*

Důkaz. Důkaz vlastností je vesměs triviální, ale je vhodný pro procvičení formálního přístupu. Základní idea důkazů je redukce dané vlastnosti na odpovídající vlastnost reálných čísel. Dokážeme vlastnost (1), zbytek necháváme čtenáři.

(1) Nejprve ověříme, že $A + B$ i $B + A$ mají stejný typ. Pak ukážeme, že odpovídající si prvky jsou shodné: $(A + B)_{ij} = A_{ij} + B_{ij} = B_{ij} + A_{ij} = (B + A)_{ij}$, kde jsme v druhé rovnici využili komutativitu sčítání reálných čísel. $\square$

Příklad 3.6. Jednoduchým použitím maticových operací je manipulace s obrázky. Nechť obrázek je reprezentován maticí $A \in \mathbb{R}^{m\times n}$ tak, že pixel obrázku na pozici (i, j) má barvu s číslem a_{ij}. Násobení matice A skalárem pak mění odstíny barev. Na obrázcích listovnice dole je ilustrována matice αA pro různé volby α. Pro $\alpha = 1$ dostaneme původní matici, pro $\alpha > 1$ se obrázek zesvětlí a naopak pro $\alpha < 1$ se obrázek ztmaví.

originál ($\alpha = 1$)

ztmavení ($\alpha = 0.5$)

originál ($\alpha = 1$)

zesvětlení ($\alpha = 1.5$) □

Nyní zavedeme násobení matic; to je definováno na první pohled trochu neobvykle, jeho význam vyplyne později v sekci 6.2.

Definice 3.7 (Součin)**.** Buď $A \in \mathbb{R}^{m\times p}$ a $B \in \mathbb{R}^{p\times n}$. Pak AB je matice typu $m \times n$ s prvky $(AB)_{ij} = \sum_{k=1}^{p} A_{ik}B_{kj}$.

Příklad 3.8 (Mnemotechnika násobení matic)**.** Buď

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 0 & 1 \\ 2 & 2 & 2 & 2 \end{pmatrix}, \; B = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 0 & 2 & 2 \\ 1 & 2 & 1 & 3 \\ 1 & 2 & 1 & 0 \end{pmatrix}.$$

Mnemotechnická pomůcka pro násobení matic AB, prvek na pozici (i, j) spočítáme jako skalární součin

i-tého řádku matice A a j-tého sloupce matice B:

$$\begin{array}{c|c} & \begin{pmatrix} 1 & 1 & \mathbf{1} & 1 \\ 1 & 0 & \mathbf{2} & 2 \\ 1 & 2 & \mathbf{1} & 3 \\ 1 & 2 & \mathbf{1} & 0 \end{pmatrix} \\ \hline \begin{pmatrix} 1 & 2 & 3 & 4 \\ \mathbf{0} & \mathbf{1} & \mathbf{0} & \mathbf{1} \\ 2 & 2 & 2 & 2 \end{pmatrix} & \begin{pmatrix} 10 & 15 & 12 & 14 \\ 2 & 2 & \mathbf{3} & 2 \\ 8 & 10 & 10 & 12 \end{pmatrix} \end{array}$$

□

Důležitou maticí je *jednotková matice*. Značí se I či I_n a je to čtvercová matice řádu n s prvky $I_{ij} = 1$ pro $i = j$ a $I_{ij} = 0$ jinak. Je to tedy matice s jedničkami na diagonále a s nulami jinde, přičemž *diagonálou* se rozumí prvky na pozicích $(1,1)$, $(2,2)$, … Schematicky

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

Související pojem *jednotkový vektor* e_i je pak i-tý sloupec jednotkové matice, tj. $e_i = I_{*i}$.

Tvrzení 3.9 (Vlastnosti součinu matic)**.** *Platí následující vlastnosti; α je číslo a A, B, C matice vhodných rozměrů.*

(1) $(AB)C = A(BC)$ *(asociativita)*

(2) $A(B + C) = AB + AC$ *(distributivita zleva)*

(3) $(A + B)C = AC + BC$ *(distributivita zprava)*

(4) $\alpha(AB) = (\alpha A)B = A(\alpha B)$

(5) $0A = A0 = 0$

(6) $I_m A = AI_n = A$, *kde* $A \in \mathbb{R}^{m\times n}$

Důkaz. Opět dokážeme jen vlastnost (1), ostatní necháváme za cvičení.

(1) Buď $A \in \mathbb{R}^{m\times p}$, $B \in \mathbb{R}^{p\times r}$ a $C \in \mathbb{R}^{r\times n}$. Pak AB má typ $m \times r$, BC má typ $p \times n$ a oba součiny $(AB)C$, $A(BC)$ mají typ $m \times n$. Nyní ukážeme, že odpovídající si prvky jsou shodné. Na pozici (i, j) jest

$$((AB)C)_{ij} = \sum_{k=1}^{r} (AB)_{ik} C_{kj} = \sum_{k=1}^{r} \left(\sum_{\ell=1}^{p} A_{i\ell} B_{\ell k} \right) C_{kj} = \sum_{k=1}^{r} \sum_{\ell=1}^{p} A_{i\ell} B_{\ell k} C_{kj},$$

$$(A(BC))_{ij} = \sum_{\ell=1}^{p} A_{i\ell} (BC)_{\ell j} = \sum_{\ell=1}^{p} A_{i\ell} \left(\sum_{k=1}^{r} B_{\ell k} C_{kj} \right) = \sum_{\ell=1}^{p} \sum_{k=1}^{r} A_{i\ell} B_{\ell k} C_{kj}.$$

Vidíme, že oba výrazy jsou shodné až na pořadí sčítanců. Ale protože sčítání reálných čísel je komutativní, tak jsou hodnoty stejné. □

Poznámka 3.10. Součin matic obecně není komutativní, pro mnoho matic je $AB \neq BA$. Navíc, může se stát, že součin AB má smysl, a přitom násobit matice v pořadí BA nelze. Najděte takové příklady!

Definice 3.11 (Transpozice)**.** Buď $A \in \mathbb{R}^{m\times n}$. Pak *transponovaná matice* má typ $n \times m$, značí se A^T a je definovaná $(A^T)_{ij} := a_{ji}$.

Příklad 3.12. Transpozice vlastně znamená překlopení dle hlavní diagonály, např.

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}, \quad A^T = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}.$$

Díky transpozici můžeme sloupcové vektory $x \in \mathbb{R}^n$ zapisovat do řádků takto: $x = (x_1, \dots, x_n)^T$.

Tvrzení 3.13 (Vlastnosti transpozice)**.** *Platí následující vlastnosti; α je číslo a A, B matice vhodných rozměrů.*

(1) $(A^T)^T = A$

(2) $(A+B)^T = A^T + B^T$

(3) $(\alpha A)^T = \alpha A^T$

(4) $(AB)^T = B^T A^T$

Důkaz. Pro ilustraci dokážeme jen vlastnost (1), zbytek ponecháme za cvičení.

(1) Buď $A \in \mathbb{R}^{m\times n}$. Pak A^T má rozměr $n \times m$ a $(A^T)^T$ má tedy rozměr $m \times n$, shodný s A. Porovnáním odpovídajících si prvků $((A^T)^T)_{ij} = (A^T)_{ji} = A_{ij}$ konstatujeme rovnost, tudíž $(A^T)^T = A$. □

Vlastnost (4) lze matematickou indukcí snadno rozšířit na součin k matic: $(A_1 A_2 \dots A_k)^T = A_k^T \dots A_2^T A_1^T$.

Definice 3.14 (Symetrická matice)**.** Matice $A \in \mathbb{R}^{n\times n}$ je *symetrická*, pokud $A = A^T$.

Symetrické matice jsou tedy takové matice, které jsou invariantní vůči operaci transpozice. Vizuálně jsou symetrické dle hlavní diagonály. Příkladem symetrických matic jsou jednotková matice I_n, čtvercová nulová matice 0_n nebo $\left(\begin{smallmatrix} 1 & 2 \\ 2 & 3 \end{smallmatrix}\right)$. Součet symetrických matic stejného řádu je zase symetrická matice (dokažte!), ale pro součin už to obecně neplatí (najděte protipříklad!).

Existuje celá řada dalších speciálních typů matic. Mezi ty nejpoužívanější patří např.:

- *Diagonální matice.* Matice $A \in \mathbb{R}^{n\times n}$ je diagonální, pokud $a_{ij} = 0$ pro všechna $i \neq j$. Tedy diagonální matice má na diagonále libovolné prvky a mimo ni jsou nuly.

  Příkladem diagonální matice je I_n nebo 0_n. Diagonální matici s diagonálními prvky $v_1, \dots, v_n$ značíme $\operatorname{diag}(v)$, tedy

$$\operatorname{diag}(v) = \begin{pmatrix} v_1 & 0 & \dots & 0 \\ 0 & v_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & v_n \end{pmatrix}.$$

- *Horní trojúhelníková matice.* Matice $A \in \mathbb{R}^{m\times n}$ je horní trojúhelníková, pokud $a_{ij} = 0$ pro všechna $i > j$. Horní trojúhelníková matice má tedy pod diagonálou nuly. Může mít libovolné rozměry, i když nejčastější případ je čtvercová matice. Schematicky pro $m \leq n$

$$\begin{pmatrix} a_{11} & \dots & \dots & a_{1m} & \dots & a_{1n} \\ 0 & a_{22} & & \vdots & & \vdots \\ \vdots & \ddots & \ddots & \vdots & & \vdots \\ 0 & \dots & 0 & a_{mm} & \dots & a_{mn} \end{pmatrix}.$$

  Podobně se zavádí i *dolní trojúhelníková matice* jako matice s nulami nad diagonálou.

  Příkladem horní trojúhelníkové matice je jakákoli matice v REF tvaru, protože pivoty musí být na nebo nad diagonálou. Obráceně to ovšem neplatí, horní trojúhelníková matice není automaticky v REF tvaru.

Příklad 3.15. Ukažte, že součin dvou horních trojúhelníkových matic $A, B \in \mathbb{R}^{n\times n}$ je opět horní trojúhelníková matice. □

Vraťme se na chvíli k aritmetickým vektorům. Transpozice a součin vektorů jakožto matic o jednom sloupci umožňují zavést součin vektorů $x, y \in \mathbb{R}^n$ jako $x^T y$ nebo jako xy^T, jiné kombinace nejsou možné. První z výrazů definuje *standardní skalární součin*, a má tvar

$$x^T y = \sum_{i=1}^{n} x_i y_i$$

(formálně je to matice 1×1, ale ztotožníme ji s reálným číslem). Standardní eukleidovskou normu vektoru $x \in \mathbb{R}^n$ lze pak zavést jako

$$\|x\| = \sqrt{x^T x} = \sqrt{\textstyle\sum_{i=1}^n x_i^2}$$

Obecnější definice skalárního součinu a normy přijde později (kapitola 8).

Vnější součin vektorů x, y je čtvercová matice řádu n

$$xy^T = \begin{pmatrix} x_1y_1 & x_1y_2 & \dots & x_1y_n \\ x_2y_1 & x_2y_2 & \dots & x_2y_n \\ \vdots & \vdots & & \vdots \\ x_ny_1 & x_ny_2 & \dots & x_ny_n \end{pmatrix} = \begin{pmatrix} — & x_1y^T & — \\ — & x_2y^T & — \\ & \vdots & \\ — & x_ny^T & — \end{pmatrix} =$$

$$= \begin{pmatrix} | & | & & | \\ xy_1 & xy_2 & \dots & xy_n \\ | & | & & | \end{pmatrix}.$$

Například $e_ie_j^T$ je matice s jedničkou na pozici (i, j) a jinde s nulami. Protože v matici xy^T jsou všechny řádky násobkem vektoru y^T (a všechny sloupce násobkem vektoru x), tak má matice xy^T hodnost nanejvýš 1. Nulová hodnost nastane pouze pokud jeden z vektorů x, y je nulový, takže pro x, y nenulové má matice xy^T hodnost přesně 1.

V následující větě zmíníme ještě nějaké vlastnosti maticového násobení, které občas budeme využívat. První dvě vlastnosti říkají, co je výsledkem násobení matice s jednotkovým vektorem, další dvě ukazují, jak snadno získat řádek či sloupec součinu matic a poslední dvě dávají jiný pohled na násobení matice s vektorem.

Tvrzení 3.16. *Buď $A \in \mathbb{R}^{m\times p}$, $B \in \mathbb{R}^{p\times n}$, $x \in \mathbb{R}^p$ a $y \in \mathbb{R}^m$. Pak platí:*

(1) $Ae_j = A_{*j}$

(2) $e_i^T A = A_{i*}$

(3) $(AB)_{*j} = AB_{*j}$

(4) $(AB)_{i*} = A_{i*}B$

(5) $Ax = \sum_{j=1}^p x_jA_{*j}$

(6) $y^T A = \sum_{i=1}^m y_iA_{i*}$

Důkaz. Jednoduché z definice. Pro ilustraci si uvedeme jen některá tvrzení:

(1) $(Ae_j)_i = \sum_{k=1}^n a_{ik}(e_j)_k = \sum_{k\neq j}^n a_{ik} \cdot 0 + a_{ij} \cdot 1 = a_{ij}$.

(3) S využitím první vlastnosti, $(AB)_{*j} = (AB)e_j = A(Be_j) = AB_{*j}$.

(5) Levá strana: $(Ax)_i = \sum_{j=1}^p a_{ij}x_j$, pravá strana: $(\sum_{j=1}^p x_jA_{*j})_i = \sum_{j=1}^p x_j(A_{*j})_i = \sum_{j=1}^p x_ja_{ij}$. □

Schematické vyjádření vlastnosti (1)

$$Ae_j = \begin{pmatrix} | & & | & & | \\ A_{*1} & \dots & A_{*j} & \dots & A_{*n} \\ | & & | & & | \end{pmatrix} \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} | \\ A_{*j} \\ | \end{pmatrix},$$

vlastnosti (3)

$$\begin{array}{c|c} & \begin{pmatrix} | & & | \\ B_{*1} & \dots & B_{*n} \\ | & & | \end{pmatrix} \\ \hline \begin{pmatrix} & & \\ & A & \\ & & \end{pmatrix} & \begin{pmatrix} | & & | \\ AB_{*1} & \dots & AB_{*n} \\ | & & | \end{pmatrix} \end{array}, \tag{3.1}$$

a vlastnosti (3)

$$\sum_{j=1}^{p} x_j A_{*j} = \begin{pmatrix} | \\ A_{*1} \\ | \end{pmatrix} x_1 + \begin{pmatrix} | \\ A_{*2} \\ | \end{pmatrix} x_2 + \ldots + \begin{pmatrix} | \\ A_{*n} \\ | \end{pmatrix} x_n.$$

Poznámka 3.17 (Zápis a interpretace soustavy rovnic)**.** Soustavu lineárních rovnic (2.1) můžeme maticově zapsat také takto: $Ax = b$, kde $x = (x_1, \ldots, x_n)^T$ je vektor proměnných, $b \in \mathbb{R}^m$ vektor pravých stran a $A \in \mathbb{R}^{m \times n}$ matice soustavy. Toto je běžně používaný zápis pro soustavu lineárních rovnic.

Na soustavu rovnic $Ax = b$ můžeme nahlížet řádkově a sloupcově. Její i-tá rovnice má tvar $A_{i*}x = b_i$, neboli $a_{i1}x_1 + \ldots + a_{in}x_n = b_i$, a popisuje nadrovinu v prostoru $\mathbb{R}^n$ (srov. poznámka 2.7). Hledáme-li řešení soustavy, hledáme vektory x, které splňují všechny rovnice, neboli body v průniku všech nadrovin.

Oproti tomu sloupcová interpretace soustavy rovnic vychází z bodu (5) tvrzení 3.16. Soustavu $Ax = b$ můžeme vyjádřit jako $\sum_{j=1}^{p} x_j A_{*j} = b$, nebo ještě názorněji jako

$$\begin{pmatrix} | \\ A_{*1} \\ | \end{pmatrix} x_1 + \begin{pmatrix} | \\ A_{*2} \\ | \end{pmatrix} x_2 + \ldots + \begin{pmatrix} | \\ A_{*n} \\ | \end{pmatrix} x_n = \begin{pmatrix} | \\ b \\ | \end{pmatrix}.$$

Při řešení soustavy tedy chceme vektor b vyjádřit jako kombinaci sloupečků matice A, přičemž sloupečky můžeme pronásobovat jednotlivými proměnnými. Více viz poznámka 5.18.

Poznámka 3.18 (Matice a zobrazení)**.** Často bude užitečné se na matici dívat jako na určité zobrazení (více v kapitole 6). Buď $A \in \mathbb{R}^{m \times n}$ a uvažujme zobrazení z $\mathbb{R}^n$ do $\mathbb{R}^m$ definované předpisem $x \mapsto Ax$. Řešit soustavu rovnic $Ax = b$ pak z tohoto pohledu znamená najít všechny vektory x, který se zobrazí na vektor b.

Ještě zajímavější je dívat se na složení dvou takovýchto zobrazení. Mějme dvě zobrazení $x \mapsto Ax$, $y \mapsto By$, kde $A \in \mathbb{R}^{p \times n}, B \in \mathbb{R}^{m \times p}$. Jak vypadá složené zobrazení? Vektor x se při prvním zobrazení zobrazí na Ax a při druhém na $B(Ax) = (BA)x$. Složené zobrazení jde tedy opět reprezentovat maticově, a to maticí BA. Skládání zobrazení tudíž odpovídá násobení matic! To není náhoda, protože maticové násobení bylo původně vyvinuto právě k těmto účelům.

Asociativitu maticového násobení můžeme také alternativně nahlédnout pomocí lineárních zobrazení. Protože skládání zobrazení je vždy asociativní, musí být asociativní i součin matic, který reprezentuje skládání lineárních zobrazení.

Příklad 3.19 (Matice a zobrazení)**.** Uvažujme konkrétní zobrazení $x \mapsto Ax$ v rovině $\mathbb{R}^2$ s maticí $A = \left(\begin{smallmatrix} -1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$. Vektor $x = (x_1, x_2)^T \in \mathbb{R}^2$ se pak zobrazí na vektor $Ax = (-x_1, x_2)^T$ a zobrazení tudíž představuje překlopení podle osy x_2, jak ilustruje následující obrázek:

Uvažujme nyní jiné zobrazení $x \mapsto Ax$ v rovině $\mathbb{R}^2$ s maticí $A = \left(\begin{smallmatrix} 2.5 & 0 \\ 0 & 1 \end{smallmatrix}\right)$. Vektor $x \in \mathbb{R}^2$ se zobrazí na vektor $Ax = (2.5x_1, x_2)^T$ a zobrazení představuje roztáhnutí ve směru osy x_1:

Jako poslední zobrazení uvažujme otočení v rovině o úhel α proti směru hodinových ručiček. I toto zobrazení lze reprezentovat maticově jako $x \mapsto Ax$ s maticí

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

Zde vektor $x \in \mathbb{R}^2$ se zobrazí na vektor $Ax = (x_1\cos(\alpha) - x_2\sin(\alpha),\ x_1\sin(\alpha) + x_2\cos(\alpha))^T$.

Podrobněji toto zobrazení probereme v příkladu 6.25, ale čtenář může již nyní nahlédnout platnost pro speciální volbu $\alpha = 90°$ nebo $\alpha = 180°$. □

Příklad 3.20 (Matice a skládání zobrazení)**.** Uvažujme zobrazení $x \mapsto Ax$ v rovině $\mathbb{R}^2$ s maticí $A = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$. Vektor $x \in \mathbb{R}^2$ se pak zobrazí na vektor $Ax = (-x_2, x_1)^T$ a není těžké nahlédnout, že toto zobrazení představuje otočení o 90° proti směru hodinových ručiček. Uvažujme ještě jinou matici $B = \left(\begin{smallmatrix} -1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$, která přestavuje překlopení podle osy x_2. Pokud složíme obě zobrazení, dostaneme zobrazení

$$x \mapsto B(Ax) = (BA)x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} x = (x_2, x_1)^T,$$

které představuje překlopení podle osy $x_2 = x_1$. Pokud bychom zobrazení skládali v opačném pořadí, výsledné zobrazení

$$x \mapsto A(Bx) = (AB)x = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} x = (-x_2, -x_1)^T$$

by zase představovalo překlopení podle osy $x_2 = -x_1$. Dostali bychom tedy jiné zobrazení, což odpovídá tomu, že součin matic není komutativní. Skládání zobrazení totiž obecně často nekomutuje. □

Poznámka 3.21 (Blokové násobení matic)**.** Občas je výhodné použít tzv. blokové násobení matic, tj. matice rozdělíme do několika bloků (podmatic) a pak se matice násobí jako by byly podmatice obyčejná čísla. Např. pokud matice A, B rozdělíme na 4 podmatice, pak

$$AB = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} = \begin{pmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \end{pmatrix}.$$

Zde je nutné si dát pozor, aby podmatice $A_{11}, \dots, B_{22}$ měly vhodné rozměry a součiny v pravé části rovnosti dávaly smysl.

Poznámka 3.22 (Rychlé násobení matic)**.** Jestliže násobíme dvě čtvercové matice řádu n, tak nás výpočet stojí řádově n^3 aritmetických operací. Na první pohled se může zdát, že výpočet nelze výrazně urychlit. Nicméně německý matematik Volker Strassen přišel r. 1969 s algoritmem, který potřebuje řádově pouze $n^{\log_2 7} \approx n^{2.807}$ operací. Strassenův algoritmus využívá právě rozdělení matic do bloků a chytrého přeuspořádání. Vývoj šel dál, Coppersmithův–Winogradův algoritmus z r. 1990 snížil výpočetní složitost na řádově $n^{2.376}$. Poznamenejme ale, že tyto rychlé algoritmy se uplatní pouze pro velké n, protože skryté koeficienty u řádových odhadů jsou poměrně vysoké. Jaká je nejmenší možná asymptotická složitost je stále otevřený problém. Zajímavé také je, že násobení matic má stejnou asymptotickou složitost jako maticová inverze probíraná v sekci 3.3, tedy oba problémy jsou na sebe efektivně převoditelné.

3.2 Regulární matice

Regulární matice představují jeden z nejdůležitějších typů matic a budou nás provázet celou knihou. Uvažujme čtvercovou soustavu n lineárních rovnic o n neznámých $Ax = b$. Víme z poznámky 2.7, že geometricky se jedná o průnik jakýchsi nadrovin v prostoru $\mathbb{R}^n$. Může se stát, že rovnice jsou takzvaně lineárně závislé (více o tomto pojmu v sekci 5.3) – například když se rovnice opakují nebo je jedna součtem několika jiných. Potom žádné řešení existovat nemusí nebo jich je nekonečně mnoho, např. celá přímka. Pokud jsou ale rovnice nezávislé a tedy odpovídající nadroviny jsou v prostoru v obecné poloze, pak je řešení (= průnik nadrovin) jednoznačné. Právě tento případ je charakterizován tím, že matice soustavy je regulární. Pro účely definice se nejprve omezíme na soustavy s nulovou pravou stranou, a pak ukážeme souvislost s obecnou čtvercovou soustavou.

Definice 3.23 (Regulární matice)**.** Buď $A \in \mathbb{R}^{n\times n}$. Matice A je *regulární*, pokud soustava $Ax = 0$ má jediné řešení $x = 0$. V opačném případě se matice A nazývá *singulární*.

Soustava $Ax = 0$ s nulovou pravou stranou se nazývá *homogenní*. Evidentně, nulový vektor je vždy jejím řešením. Pro A regulární ale žádné jiné řešení existovat nesmí. Jiný ekvivalentní pohled na regulární matice je, že $Ax \neq 0$ pro všechna $x \neq 0$. Typickým příkladem regulární matice je I_n a singulární matice 0_n.

Tvrzení 3.24. *Buď $A \in \mathbb{R}^{n\times n}$. Pak následující jsou ekvivalentní:*

(1) A je regulární,

(2) $\mathrm{RREF}(A) = I$,

(3) $\mathrm{rank}(A) = n$.

Důkaz. Plyne z rozboru Gaussovy–Jordanovy eliminace. □

Nyní ukážeme, že nulová pravá strana soustavy z definice regulární matice není tak podstatná.

Tvrzení 3.25. *Buď $A \in \mathbb{R}^{n\times n}$. Pak následující jsou ekvivalentní:*

(1) A je regulární,

(2) pro nějaké $b \in \mathbb{R}^n$ má soustava $Ax = b$ jediné řešení,

(3) pro každé $b \in \mathbb{R}^n$ má soustava $Ax = b$ jediné řešení.

Důkaz. Plyne z rozboru Gaussovy–Jordanovy eliminace a tvrzení 3.24. □

Podívejme se na základní vlastnosti regulárních matic. Součet regulárních matic nemusí být regulární matice, vezměme např. $I + (-I) = 0$. Součin matic ale regularitu zachovává.

Tvrzení 3.26. *Buďte $A, B \in \mathbb{R}^{n\times n}$ regulární matice. Pak AB je také regulární.*

Důkaz. Buď x řešení soustavy $ABx = 0$. Chceme ukázat, že x musí být nulový vektor. Označme $y := Bx$. Pak soustava lze přepsat na $Ay = 0$ a z regularity A je $y = 0$, neboli $Bx = 0$. Z regularity B je $x = 0$. □

Tvrzení 3.27. *Je-li aspoň jedna z matic $A, B \in \mathbb{R}^{n\times n}$ singulární, pak AB je také singulární.*

Důkaz. Uvažme dva případy. Je-li B singulární, pak $Bx = 0$ pro nějaké $x \neq 0$. Z toho ale plyne $(AB)x = A(Bx) = A0 = 0$, tedy i AB je singulární.

Nyní předpokládejme, že B je regulární, tedy A je singulární a existuje $y \neq 0$ takové, že $Ay = 0$. Z regularity B dostáváme, že existuje $x \neq 0$ tak, že $Bx = y$. Celkem dostáváme $(AB)x = A(Bx) = Ay = 0$, tedy AB je singulární. □

Vraťme se k elementárním řádkovým úpravám. Ukážeme si, že jdou reprezentovat maticově, a že tyto matice jsou regulární. To, že jdou reprezentovat maticově, znamená, že výsledek úpravy na matici A se dá vyjádřit jako EA pro nějakou matici E. Jak najít tuto matici? Pomůže nám uvědomíme-li si, že aplikací dané úpravy na jednotkovou matici I_n dostaneme $EI = E$, tedy matici reprezentující danou úpravy dostaneme tak, že tuto úpravy provedeme na jednotkovou matici. Ale pozor! To je pouze nutná podmínka, jak by taková matice měla vypadat. To, že to skutečně funguje, se musí dokázat pro každou úpravu zvlášť (důkaz necháváme na cvičení):

1. Vynásobení i-tého řádku číslem $\alpha \neq 0$ lze reprezentovat vynásobením zleva maticí

$$E_i(\alpha) = I + (\alpha - 1)e_i e_i^T = \begin{pmatrix} 1 & 0 & \dots & \dots & 0 \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \alpha & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \dots & \dots & 0 & 1 \end{pmatrix}.$$

2. Přičtení α-násobku j-tého řádku k i-tému, přičemž $i \neq j$, lze reprezentovat vynásobením zleva maticí

$$E_{ij}(\alpha) = I + \alpha e_i e_j^T = \begin{array}{c} \\ \\ \\ i \\ \\ \\ \end{array} \begin{pmatrix} 1 & 0 & \dots & \dots & 0 \\ & \ddots & \ddots & & \vdots \\ & & 1 & \ddots & \vdots \\ & \alpha & & \ddots & 0 \\ & & & & 1 \end{pmatrix}.$$
$$\begin{array}{ccccc} & & j & & \end{array}$$

3. Výměna i-tého a j-tého řádku jde reprezentovat vynásobením zleva maticí

$$E_{ij} = I + (e_j - e_i)(e_i - e_j)^T = \begin{array}{c} i \\ j \end{array} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$
$$\begin{array}{cc} i & j \end{array}$$

Snadno se ukáže, že matice elementárních operací jsou regulární. Získali jsme je totiž aplikací elementární úpravy na jednotkovou matici, tudíž inverzní úpravou převedeme matici zpět na jednotkovou.

Maticově jdou reprezentovat nejenom elementární řádkové úpravy, ale také celá transformace převodu matice na odstupňovaný tvar.

Věta 3.28. *Buď $A \in \mathbb{R}^{m\times n}$. Pak* $\mathrm{RREF}(A) = QA$ *pro nějakou regulární matici $Q \in \mathbb{R}^{m\times m}$.*

Důkaz. $\mathrm{RREF}(A)$ získáme aplikací konečně mnoha elementárních řádkových úprav. Nechť jdou reprezentovat maticemi $E_1, E_2, \dots, E_k$. Pak $\mathrm{RREF}(A) = E_k \dots E_2 E_1 A = QA$, kde $Q = E_k \dots E_2 E_1$. Protože matice $E_1, E_2, \dots, E_k$ jsou regulární, i jejich součin Q je regulární. □

Tvrzení 3.29. *Každá regulární matice $A \in \mathbb{R}^{n\times n}$ se dá vyjádřit jako součin konečně mnoha elementárních matic.*

Důkaz. Pokud k elementárními úpravami dokážu dovést matici A na jednotkovou I_n, pak jistými k elementárními úpravami mohu převést naopak I_n na A. Je to tím, že každá elementární úprava má svojí inverzní, která vykonává opačnou úpravu. Tudíž existují matice $E_1, \dots, E_k$ elementárních úprav tak, že $A = E_k \dots E_2 E_1 I_n = E_k \dots E_2 E_1$. □

3.3 Inverzní matice

Motivace pro inverzní matice: Matice umíme sčítat, odečítat, násobit, tak nešly by i dělit? Ukážeme si, že něco jako dělení lze zavést, ale jen pro regulární matice.

Definice 3.30. Buď $A \in \mathbb{R}^{n\times n}$. Pak A^{-1} je *inverzní* maticí k A, pokud splňuje $AA^{-1} = A^{-1}A = I_n$.

Například, matice inverzní k I_n je opět I_n. Inverzní matice k nulové 0_n evidentně neexistuje. Které matice tedy mají inverzi? Ukážeme, že pouze a jen ty regulární.

Věta 3.31 (O existenci inverzní matice)**.** *Buď $A \in \mathbb{R}^{n\times n}$. Je-li A regulární, pak k ní existuje inverzní matice, a je určená jednoznačně. Naopak, existuje-li k A inverzní, pak A musí být regulární.*

Důkaz. „Existence.“ Z předpokladu regularity matice A plyne, že soustava $Ax = e_j$ má (jediné) řešení pro každé $j = 1, \dots, n$, označme je x_j, $j = 1, \dots, n$. Vytvořme matici A^{-1} tak, aby její sloupce byly vektory $x_1, \dots, x_n$, to jest, $A^{-1} = (x_1|x_2|\dots|x_n)$. Ukážeme, že tato matice je hledaná inverze. Rovnost $AA^{-1} = I$ ukážeme po sloupcích. Buď $j \in \{1, \dots, n\}$, pak

$$(AA^{-1})_{*j} = A(A^{-1})_{*j} = Ax_j = e_j = I_{*j}.$$

Druhou rovnost dokážeme trochu trikem. Uvažme výraz

$$A(A^{-1}A - I) = AA^{-1}A - A = IA - A = 0.$$

Matice $A(A^{-1}A - I)$ je tedy nulová a její j-tý sloupec je nulový vektor: $A(A^{-1}A - I)_{*j} = 0$. Z regularity matice A dostáváme, že $(A^{-1}A - I)_{*j} = 0$. Protože to platí pro každé $j \in \{1, \dots, n\}$, je $A^{-1}A - I = 0$, neboli $A^{-1}A = I$.

„Jednoznačnost.“ Nechť pro nějakou matici B platí $AB = BA = I$. Pak

$$B = BI = B(AA^{-1}) = (BA)A^{-1} = IA^{-1} = A^{-1},$$

tedy B už musí být automaticky rovno naší zkonstruované matici A^{-1}.

„Naopak.“ Nechť pro A existuje inverzní matice. Buď x řešení soustavy $Ax = 0$. Pak $x = Ix = (A^{-1}A)x = A^{-1}(Ax) = A^{-1}0 = 0$. Tedy A je regulární. □

Pomocí inverzních matic snadno dokážeme následující větu, kterou bychom přímo z definice bez vybudovaného aparátu dokazovali jen těžko.

Tvrzení 3.32. *Je-li A regulární, pak A^T je regulární.*

Důkaz. Je-li A regulární, pak existuje inverzní matice a platí $AA^{-1} = A^{-1}A = I_n$. Po transponování všech stran rovností dostaneme $(AA^{-1})^T = (A^{-1}A)^T = I_n^T$, neboli $(A^{-1})^T A^T = A^T (A^{-1})^T = I_n$. Vidíme, že matice A^T má inverzní (rovnou $(A^{-1})^T$) a je tudíž regulární. □

Z důkazu vyplynul pěkný vztah $(A^T)^{-1} = (A^{-1})^T$, kterážto matice se někdy značívá zkráceně A^{-T}.

Nyní ukážeme, že dvě rovnosti $AA^{-1} = I_n$, $A^{-1}A = I_n$ z definice inverzní matice jsou zbytečný přepych a k jejímu určení stačí jen jedna z nich.

Věta 3.33 (Jedna rovnost stačí)**.** *Buďte $A, B \in \mathbb{R}^{n\times n}$. Je-li $BA = I$, pak obě matice A, B jsou regulární a navzájem k sobě inverzní, to jest $B = A^{-1}$ a $A = B^{-1}$.*

Důkaz. Regularita vyplývá z tvrzení 3.27 vzhledem k regularitě jednotkové matice I_n. Nyní víme, že A, B jsou regulární a tudíž mají inverze A^{-1}, B^{-1}. Odvodíme

$$B = BI = B(AA^{-1}) = (BA)A^{-1} = IA^{-1} = A^{-1},$$

a podobně

$$A = AI = A(BB^{-1}) = (AB)B^{-1} = IB^{-1} = B^{-1}. \qquad \square$$

Jak vypočítat inverzní matici? První část důkazu věty 3.31 ukázala návod, jak inverzní matici spočítat pomocí n soustav lineárních rovnic. Návod říká, že j-tý sloupec inverzní matice A^{-1} je řešením soustavy $Ax = e_j$. Není však třeba řešit tyto soustavy zvlášť. Protože všechny soustavy mají stejnou matici A, vyřešíme je naráz tak, že na pravou stranu umístíme vedle sebe všechny jednotkové vektory $e_1, \dots, e_n$. Pravá strana je tak tvořena jednotkovou maticí I_n a soustava má tvar tzv. maticové soustavy $Ax = I_n$. Formální opodstatnění postupu shrnuje následující věta.

Věta 3.34 (Výpočet inverzní matice)**.** *Buď $A \in \mathbb{R}^{n\times n}$. Nechť matice $(A \,|\, I_n)$ typu $n \times 2n$ má RREF tvar $(I_n \,|\, B)$. Pak $B = A^{-1}$. Netvoří-li první část RREF tvaru jednotkovou matici, pak A je singulární.*

Důkaz. Je-li $\mathrm{RREF}(A \,|\, I_n) = (I_n \,|\, B)$, potom dle věty 3.28 existuje regulární matice Q taková, že $(I_n \,|\, B) = Q(A \,|\, I_n)$, neboli po roztržení na dvě části $I_n = QA$ a $B = QI_n$. První rovnost říká $Q = A^{-1}$ a druhá $B = Q = A^{-1}$.

Netvoří-li první část RREF tvaru jednotkovou matici, pak $\mathrm{RREF}(A) \neq I_n$ a tudíž A není regulární. $\square$

Příklad 3.35. Buď $A = \begin{pmatrix} 1 & 1 & 3 \\ 0 & 2 & -1 \\ 3 & 5 & 7 \end{pmatrix}$. Inverzní matici spočítáme takto:

$$(A \,|\, I_3) = \left(\begin{array}{ccc|ccc} 1 & 1 & 3 & 1 & 0 & 0 \\ 0 & 2 & -1 & 0 & 1 & 0 \\ 3 & 5 & 7 & 0 & 0 & 1 \end{array}\right) \sim \left(\begin{array}{ccc|ccc} 1 & 1 & 3 & 1 & 0 & 0 \\ 0 & 2 & -1 & 0 & 1 & 0 \\ 0 & 2 & -2 & -3 & 0 & 1 \end{array}\right) \sim \left(\begin{array}{ccc|ccc} 1 & 1 & 3 & 1 & 0 & 0 \\ 0 & 1 & -0.5 & 0 & 0.5 & 0 \\ 0 & 2 & -2 & -3 & 0 & 1 \end{array}\right) \sim$$

$$\sim \left(\begin{array}{ccc|ccc} 1 & 0 & 3.5 & 1 & -0.5 & 0 \\ 0 & 1 & -0.5 & 0 & 0.5 & 0 \\ 0 & 0 & -1 & -3 & -1 & 1 \end{array}\right) \sim \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -9.5 & -4 & 3.5 \\ 0 & 1 & 0 & 1.5 & 1 & -0.5 \\ 0 & 0 & 1 & 3 & 1 & -1 \end{array}\right) = \left(I_3 \,|\, A^{-1}\right)$$

Tedy máme $A^{-1} = \begin{pmatrix} -9.5 & -4 & 3.5 \\ 1.5 & 1 & -0.5 \\ 3 & 1 & -1 \end{pmatrix}$. $\square$

Shrňme základní vlastnosti inverzních matic. Poznamenejme, že pro inverzi součtu dvou matic není znám žádný jednoduchý vzoreček.

Tvrzení 3.36 (Vlastnosti inverzní matice)**.** *Buďte $A, B \in \mathbb{R}^{n\times n}$ regulární. Pak:*

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

(2) $(A^{-1})^T = (A^T)^{-1}$,

(3) $(\alpha A)^{-1} = \frac{1}{\alpha}A^{-1}$ *pro* $\alpha \neq 0$,

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

Důkaz.

(1) Z rovnosti $A^{-1}A = I$ máme, že inverzní matice k A^{-1} je právě A.

(2) Bylo ukázáno v důkazu věty 3.32.
(3) Plyne z $(\alpha A)(\frac{1}{\alpha}A^{-1}) = \frac{\alpha}{\alpha}AA^{-1} = I$.
(4) Plyne z $(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AIA^{-1} = AA^{-1} = I$. □

Pomocí inverzních matic můžeme elegantně vyjádřit řešení soustavy rovnic s regulární maticí. V praxi se ovšem tento výpočet nepoužívá, neboť je časově dražší než Gaussova eliminace.

Věta 3.37 (Soustava rovnic a inverzní matice)**.** *Buď $A \in \mathbb{R}^{n\times n}$ regulární. Pak řešení soustavy $Ax = b$ je dáno vzorcem $x = A^{-1}b$.*

Důkaz. Protože A je regulární, má soustava jediné řešení x. Platí $x = Ix = (A^{-1}A)x = A^{-1}(Ax) = A^{-1}b$. □

Poznámka 3.38. Význam věty 3.37 spočívá spíš v tom, že udává explicitní vyjádření řešení soustav lineárních rovnic, než že by dávala nejlepší způsob na řešení. Popravdě, použití může být spíše opačné: chceme-li vypočítat výsledek maticového součinu $BA^{-1}b$ pro matice $A, B \in \mathbb{R}^{n\times n}$ a vektor $b \in \mathbb{R}^n$, pak dobrý způsob je vyřešit soustavu $Ax = b$ a toto řešení x pak vynásobit s maticí B jako Bx.

Příklad 3.39. Jak se změní množina řešení soustavy $Ax = b$, když obě strany vynásobíme maticí Q, tj. přejdeme k soustavě $(QA)x = Qb$? Jak se změní množina řešení, když Q je regulární?

Pokud je vektor x řešením soustavy $Ax = b$, pak je řešením i soustavy $(QA)x = Qb$, protože levá a pravá strana se rovnají. Tudíž přenásobením obou stran rovnice maticí zůstane množina řešení stejná nebo se zvětší. Jednoduchý příklad, kdy se množina řešení zvětší, je třeba při přenásobení nulovou maticí.

Pokud je matice Q regulární, tak se množina řešení nezmění. Od soustavy $(QA)x = Qb$ totiž můžeme přejít k původní soustavě přenásobením inverzní maticí Q^{-1}. Tudíž množina řešení soustavy $(QA)x = Qb$ je částí množiny řešení soustavy $Ax = b$, což spolu s předchozím dává rovnost. Rovnost množin řešení lze rovněž nahlédnout z tvrzení 3.29: Každou regulární matici lze složit z elementárních matic a elementární úpravy množinu řešení soustavy nemění. □

Poznámka 3.40. Uvažujme opět zobrazení $x \mapsto Ax$ z množiny $\mathbb{R}^n$ do množiny $\mathbb{R}^n$, kde $A \in \mathbb{R}^{n\times n}$ (viz poznámka 3.18). Je-li matice A regulární, pak pro každé $y \in \mathbb{R}^n$ existuje $x \in \mathbb{R}^n$ takové, že $Ax = y$. Jinými slovy, každý vektor z $\mathbb{R}^n$ má svůj jediný vzor, který se na něj zobrazí. Zobrazení je tedy bijekcí a existuje k němu inverzní zobrazení, jež má zjevně předpis $y \mapsto A^{-1}y$. Vidíme tedy, že inverzní zobrazení jde také vyjádřit maticově, a to s inverzní maticí.

Víme, že regulární matice jsou uzavřené na součin. Lineární zobrazení nám dávají na tuto vlastnost další náhled: Regulární matici odpovídá lineární zobrazení, které je bijekcí a skládání bijekcí je opět bijekce. Proto součin regulárních matic je opět regulární matice.

Geometricky pak zobrazení $x \mapsto Ax$ s regulární maticí $A \in \mathbb{R}^{n\times n}$ představuje takové zobrazení, které zobrazí celý prostor $\mathbb{R}^n$ opět na $\mathbb{R}^n$. Příkladem takových zobrazení je otočení, převrácení, natažení atp. Pokud je ale matice $A \in \mathbb{R}^{n\times n}$ singulární, tak dojde k deformaci prostoru a $\mathbb{R}^n$ se zobrazí jen na část prostoru $\mathbb{R}^n$. Typickým příkladem takového zobrazení je projekce. Například zobrazení $x \mapsto Ax$ v rovině $\mathbb{R}^2$ s maticí $A = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right)$ představuje projekci na osu x_1, protože vektor $x = (x_1, x_2)^T$ se zobrazí na vektor $Ax = (x_1, 0)^T$. Tím pádem obor hodnot tohoto zobrazení, kam se zobrazení rovina $\mathbb{R}^2$, je pouze osa x_1. Na tomto příkladu také vidíme, proč zobrazení (a tedy i matice A) nemá inverzi: z obrazu $(x_1, 0)^T$ nejsme schopni jednoznačně zrekonstruovat, který vektor se na něj zobrazil.

Přestože jsme zmínili, že pro inverzi součtu matic není znám žádný jednoduchý vzoreček, pro určitou třídu matic můžeme inverzi součtu matic vyjádřit explicitně. Tím speciálním případem je tzv. *rank-one update*, tedy situace, kdy jedna matice má hodnost 1. Tato formule tedy umožňuje rychle přepočítat inverzní matici, pokud původní matici „málo“ změníme, např. změny jsou jen v jednom řádku či jednom sloupci (pak b resp. c je jednotkový vektor).

Věta 3.41 (Shermanova–Morrisonova formule[1])**.** *Buď $A \in \mathbb{R}^{n\times n}$ regulární a $b, c \in \mathbb{R}^n$. Pokud $c^TA^{-1}b = -1$, tak $A + bc^T$ je singulární, jinak*

$$(A + bc^T)^{-1} = A^{-1} - \frac{1}{1 + c^TA^{-1}b}A^{-1}bc^TA^{-1}.$$

[1] Nazývána podle amerických statistiků se jmény Jack Sherman a Winifred J. Morrison, kteří ji odvodili v letech 1949–50. Nezávisle na nich byla objevena řadou jiných osobností, např. obecnější tvar odvodil Max Woodbury v r. 1950.

Důkaz. V případě $c^TA^{-1}b = -1$ máme $(A + bc^T)A^{-1}b = AA^{-1}b + bc^TA^{-1}b = b(1 + c^TA^{-1}b) = 0$. Protože $b \neq 0$ a vzhledem k regularitě A je $A^{-1}b \neq 0$, musí matice $(A + bc^T)$ být singulární.

Pokud $c^TA^{-1}b \neq -1$, dostáváme:

$$\begin{aligned}(A + bc^T)\left(A^{-1} - \frac{1}{1 + c^TA^{-1}b}A^{-1}bc^TA^{-1}\right) &= \\ = I_n + bc^TA^{-1} - \frac{1}{1 + c^TA^{-1}b}bc^TA^{-1} - \frac{1}{1 + c^TA^{-1}b}b(c^TA^{-1}b)c^TA^{-1} &= \\ = I_n + \Big(1 - \frac{1}{1 + c^TA^{-1}b} - \frac{c^TA^{-1}b}{1 + c^TA^{-1}b}\Big)bc^TA^{-1} = I_n + 0 \cdot bc^TA^{-1} = I_n.\end{aligned}$$

□

3.4 Jednoznačnost RREF

V této sekci ukážeme, že RREF tvar matice je jednoznačný. Sekci je možno při výkladu přeskočit, na zde uvedené výsledky již dále nenavazujeme.

Abychom dokázali jednoznačnost RREF tvaru, zformulujeme nejprve pomocné tvrzení.

Lemma 3.42. *Buďte $A, B \in \mathbb{R}^{m\times n}$ matice v RREF tvaru, a nechť platí $A = QB$ pro nějakou regulární matici Q. Potom $A = B$.*

Důkaz. Matematickou indukcí podle n, tj. počtu sloupců.

Je-li $n = 1$, pak nastane jedna z možností: buď $B = 0$ nebo $B = e_1$. V prvním případě je $A = QB = Q0 = 0 = B$. V druhém případě je $QB \neq 0$ díky regularitě matice Q, tedy $A = e_1$.

Indukční krok $n \leftarrow n - 1$. Rozdělme matice A, B na podmatice $m \times (n - 1)$ a poslední sloupec takto: $A = (A' \,|\, a)$, $B = (B' \,|\, b)$. Zřejmě A', B' jsou také v RREF tvaru. Rovnost $A = QB$ lze rozepsat jako $A' = QB'$ a $a = Qb$. První z nich díky indukčnímu předpokladu dává $A' = B'$, takže zbývá dokázat $a = b$.

Označme r počet pivotů matice A' a nechť jsou ve sloupcích $p_1, \ldots, p_r$. Pak pro každé $i = 1, \ldots, r$ je $e_i = A'_{*p_i} = QB'_{*p_i} = Qe_i = Q_{*i}$. Tedy prvních r sloupců matice Q je tvořeno jednotkovými sloupci. Nyní uvažme dvě možnosti: buď je poslední sloupec b matice B bázický nebo ne. Jestliže není, pak $b_{r+1} = \ldots = b_m = 0$ a tedy $b = \sum_{j=1}^m b_je_j = \sum_{j=1}^r b_je_j = \sum_{j=1}^r b_jQ_{*j} = \sum_{j=1}^m b_jQ_{*j} = Qb = a$. Naopak, je-li poslední sloupec B bázický, pak $b = e_{r+1}$. Tvrdíme, že v tomto případě je také $a = e_{r+1}$. Jinak, pokud a je nebázický, použijeme předchozí úvahy symetricky na vztah $B = Q^{-1}A$ a dostaneme, že b je také nebázický, což je spor. □

Věta 3.43 (Jednoznačnost RREF)**.** *RREF tvar matice je jednoznačně určen.*

Důkaz. Buď $A \in \mathbb{R}^{m\times n}$ a nechť má dva různé RREF tvary A_1 a A_2. Podle věty 3.28 existují regulární matice Q_1, Q_2 takové, že $Q_1A = A_1$, $Q_2A = A_2$. Přenásobením první rovnice zleva Q_1^{-1} a druhé Q_2^{-1} pak máme $A = Q_1^{-1}A_1 = Q_2^{-1}A_2$, a tedy $A_1 = Q_1Q_2^{-1}A_2$. Protože $Q_1Q_2^{-1}$ je regulární, podle lemmatu 3.42 dostáváme $A_1 = A_2$, což je spor. □

Poznámka 3.44. Předchozí věta je důležitá mj. z toho důvodu, že ospravedlňuje definici hodnosti matice jako počet pivotů RREF tvaru matice (viz definice 2.13). Nyní je tedy pojem hodnosti dobře definován. Korektnost definice hodnosti vyplyne rovněž z teorie maticových prostorů (věta 5.62).

Věta říká také to, že ačkoli REF tvar jednoznačný není, tak pozice pivotů a tvar samotný (až na numerické hodnoty) jednoznačný jest.

3.5 Ještě k soustavám rovnic

3.5.1 Numerická stabilita při řešení soustav

Na závěr kapitol věnovaných soustavám rovnic a maticím zmíníme, jak je to s numerickým řešením soustav lineárních rovnic (srov. sekce 1.3). Při numerickém řešení na počítačích dochází k zaokrouhlovacím chybám a vypočtený výsledek se může diametrálně lišit od správného řešení. Pokud je matice tzv. *špatně podmíněná*, tak je v jistém smyslu blízko k singulární matici, a při počítání s maticí dochází k amplifikaci

zaokrouhlovacích chyb. Toto je vnitřní vlastnost matice, a žádná numerická metoda si s tím v principu nedokáže zcela poradit. Základní algoritmus Gaussovy eliminace tak, jak jsme ho představili v sekci 2.2, je zrovna na zaokrouhlovací chyby citlivý. Existují však jiné metody nebo varianty Gaussovy eliminace, které se dokáží s případnou nestabilitou vypořádat trochu lépe.

Příklad 3.45. Uvažujme dvě soustavy, které se liší v jediném koeficientu podle toho, zda jsme zaokrouhlili číslo $\frac{2}{30}$ nahoru či dolů (na tři desetinná místa).

$$\begin{aligned} 0.835x_1 + 0.667x_2 &= 0.168 \\ 0.333x_1 + 0.266x_2 &= \mathbf{0.067} \end{aligned} \qquad \begin{aligned} 0.835x_1 + 0.667x_2 &= 0.168 \\ 0.333x_1 + 0.266x_2 &= \mathbf{0.066} \end{aligned}$$

Zatímco první soustava má řešení $(x_1, x_2) = (1, -1)$, ta druhá má řešení $(x_1, x_2) = (-666, 834)$.

Geometrická představa je průsečík dvou téměř identických přímek, takže malá změna v datech znamená potencionálně velkou změnu v průsečíku.

□

Příklad 3.46. Jiným, typickým příkladem špatně podmíněných matic jsou tzv. Hilbertovy matice. Hilbertova matice H_n řádu n je definována $(H_n)_{ij} = \dfrac{1}{i+j-1}$, $i, j = 1, \ldots, n$. Např.

$$H_3 = \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \end{pmatrix}.$$

Uvažujme soustavu $H_n x = b$, kde $b = H_n e$ a $e = (1, \ldots, 1)^T$. Řešení soustavy je evidentně $x = e$, a protože H_n je regulární, je to jediné řešení. Jak se se soustavou vypořádá počítač? Výpočty v Matlabu (R 2008b), double precision 52 bitů, což odpovídá přesnosti $\approx 10^{-16}$, v roce 2009 ukázaly následující složky numericky spočítaného řešení:

n	řešení
8	$x_i = 1$
10	$x_i \in [0.9995,\ 1.0003]$
12	$x_i \in [0.8246,\ 1.1500]$
14	$x_i \in [-45.4628,\ 53.3428]$

Tedy již při $n \approx 14$ mají numerické chyby enormní vliv na přesnost řešení. □

Příklad 3.47 (Parciální pivotizace)**.** Nyní ukážeme, že parciální pivotizace často vede k přesnějšímu řešení, i když ani ta samozřejmě není všelék. Řešíme soustavu v aritmetice s přesností na 3 číslice:

$$\begin{aligned} 10^{-3}x_1 - x_2 &= 1, \\ 2x_1 + x_2 &= 0. \end{aligned}$$

Bez pivotizace:

$$\left(\begin{array}{cc|c} \mathbf{10^{-3}} & -1 & 1 \\ 2 & 1 & 0 \end{array}\right) \sim \left(\begin{array}{cc|c} \mathbf{1} & -1000 & 1000 \\ 2 & 1 & 0 \end{array}\right) \sim \left(\begin{array}{cc|c} 1 & -1000 & 1000 \\ 0 & \mathbf{2000} & -2000 \end{array}\right) \sim$$
$$\sim \left(\begin{array}{cc|c} 1 & -1000 & 1000 \\ 0 & \mathbf{1} & -1 \end{array}\right) \sim \left(\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 1 & -1 \end{array}\right).$$

Parciální pivotizace:

$$\left(\begin{array}{cc|c} 10^{-3} & -1 & 1 \\ \mathbf{2} & 1 & 0 \end{array}\right) \sim \left(\begin{array}{cc|c} \mathbf{2} & 1 & 0 \\ 10^{-3} & -1 & 1 \end{array}\right) \sim \cdots \sim \left(\begin{array}{cc|c} 1 & 0 & \frac{1}{2} \\ 0 & 1 & -1 \end{array}\right).$$

Pro porovnání, skutečné řešení je $(\frac{1000}{2001}, -\frac{2000}{2001})^T$. □

3.5.2 LU rozklad

LU rozklad je rozklad čtvercové matice $A \in \mathbb{R}^{n\times n}$ na součin $A = LU$, kde L je dolní trojúhelníková matice s jedničkami na diagonále a U horní trojúhelníková matice.

LU rozklad úzce souvisí s odstupňovaným tvarem matice. V zasadě, U je odstupňovaný tvar matice A a matici L můžeme získat z elementárních úprav, protože v zásadě představuje akumulované elementární úpravy. Pokud z elementárních úprav používáme pouze přičtení násobku řádku k nějakému pod ním (tedy bez prohazování řádků), tak matice $E_{ij}(\alpha)$ takovýchto úprav jsou dolní trojúhelníkové a jejich inverze $E_{ij}(\alpha)^{-1} = E_{ij}(-\alpha)$ taky. Protože součin dolních trojúhelníkových matic je opět dolní trojúhelníková matice, tak nám dají dohromady hledanou matici L. Základní algoritmus tedy může být:

Převeď A na odstupňovaný tvar U: tedy $E_k \dots E_1 A = U$, z čehož $A = \underbrace{E_1^{-1} \dots E_k^{-1}}_{L} U$.

Uvědomíme-li si jak se invertuje matice elementární úpravy, lze L konstruovat velice efektivně a dokonce obě matice L a U můžeme udržovat v jedné matici.

Příklad 3.48. Upravme matici A tak, že namísto nul pod diagonálou budeme zapisovat koeficienty $-\alpha$ z elementárních úprav s maticí $E_{ij}(\alpha)$, pro zdůraznění jsou zakroužkovány:

$$A = \begin{pmatrix} 2 & 1 & 3 \\ 4 & 1 & 7 \\ -6 & -2 & -12 \end{pmatrix} \sim \begin{pmatrix} 2 & 1 & 3 \\ 2 & -1 & 1 \\ -6 & -2 & -12 \end{pmatrix} \sim \begin{pmatrix} 2 & 1 & 3 \\ 2 & -1 & 1 \\ -3 & 1 & -3 \end{pmatrix} \sim \begin{pmatrix} 2 & 1 & 3 \\ 2 & -1 & 1 \\ -3 & -1 & -2 \end{pmatrix}.$$

Tedy

$$A = \begin{pmatrix} 2 & 1 & 3 \\ 4 & 1 & 7 \\ -6 & -2 & -12 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -3 & -1 & 1 \end{pmatrix} \begin{pmatrix} 2 & 1 & 3 \\ 0 & -1 & 1 \\ 0 & 0 & -2 \end{pmatrix} = LU.$$

□

Algoritmus se dá adaptovat i na případ, když při elementárních úpravách musíme někde prohodit řádky. Pak dostaneme LU rozklad matice vniklé z A prohozením nějakých řádků. Obecně totiž LU rozklad neexistuje pro každou matici (např. pro $A = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$), ale po vhodném proházení řádků už ano.

LU rozklad má široké uplatnění. Například pro invertování matic: $A^{-1} = U^{-1}L^{-1}$, pro počítání determinantu $\det(A) = \det(L)\det(U)$ (viz kapitola 9) nebo pro řešení soustav rovnic. V zásadě umožňuje matici A předzpracovat tak, aby další výpočty na ní byly pak jednodušší.

Příklad 3.49. Použití LU rozkladu pro řešení soustavy $Ax = b$ (tedy $LUx = b$):

1. Najdi LU rozklad matice A, tj. $A = LU$,
2. vyřeš soustavu $Ly = b$ dopřednou substitucí,
3. vyřeš soustavu $Ux = y$ zpětnou substitucí.

Druhý krok v podstatě odpovídá aplikaci jednotlivých kroků Gaussovy eliminace na pravou stranu b, zatímco třetí krok je zpětná substituce tak, jak ji známe z Gaussovy eliminace. Například pro soustavu s maticí z příkladu 3.48

$$(A \mid b) = \left(\begin{array}{rrr|r} 2 & 1 & 3 & -1 \\ 4 & 1 & 7 & 5 \\ -6 & -2 & -12 & -2 \end{array}\right).$$

Krok 2.

$$(L \mid b) = \left(\begin{array}{rrr|r} 1 & 0 & 0 & -1 \\ 2 & 1 & 0 & 5 \\ -3 & -1 & 1 & -2 \end{array}\right) \to y = (-1, 7, 2)^T.$$

Krok 3.

$$(U \mid y) = \left(\begin{array}{rrr|r} 2 & 1 & 3 & -1 \\ 0 & -1 & 1 & 7 \\ 0 & 0 & -2 & 2 \end{array}\right) \to x = (5, -8, -1)^T.$$

Řešením soustavy $Ax = b$ je tedy vektor $x = (5, -8, -1)^T$. □

Podívejme se nyní na to, co se stane, když při Gaussově eliminaci musíme někde prohodit dva řádky. Nechť matice A je upravena na tvar $E_\ell \dots E_1 A$ pomocí ℓ elementárních úprav přičtení násobku řádku k nějakému pod ním, a nechť nyní potřebuje prohodit řádky i, j, což je reprezentováno elementární maticí E_{ij}, $i < j$. Dále, nechť předposlední úprava s maticí E_ℓ je tvaru $E_{p,q}(\alpha)$. Uvědomme si, že z charakteru Gaussovy eliminace je $q < i$.

Jsou-li indexové množiny $\{i, j\}$ a $\{p, q\}$ disjunktní, pak $E_{ij}E_{p,q}(\alpha) = E_{p,q}(\alpha)E_{ij}$. V opačném případě je $p \in \{i, j\}$. Pro $p = i$ dostaneme $E_{ij}E_{p,q}(\alpha) = E_{j,q}(\alpha)E_{ij}$ a pro $p = j$ analogicky $E_{ij}E_{p,q}(\alpha) = E_{i,q}(\alpha)E_{ij}$. Tudíž můžeme souhrnně psát $E_{ij}E_\ell = E'_\ell E_{ij}$, kde E'_ℓ je matice elementární úpravy přičtení násobku řádku k nějakému pod ním. Podobně můžeme pokračovat dál, a nakonec přepíšeme $E_{ij}E_\ell \dots E_1 A = E'_\ell \dots E'_1 E_{ij} A$, kde $E'_\ell, \dots, E'_1$ jsou matice elementární úpravy přičtení násobku řádku k řádku pod ním.

Tento postup lze aplikovat pokaždé, když musíme vyměnit dva řádky. Takže po upravení na odstupňovaný tvar U máme $E'_k \dots E'_1 E_{i_1 j_1} \dots E_{i_r j_r} A = U$, kde $E'_k, \dots, E'_1$ jsou matice elementární úpravy přičtení násobku řádku k řádku pod ním a $E_{i_1 j_1}, \dots, E_{i_r j_r}$ jsou matice úpravy prohození dvou řádků. Po úpravě dostaneme $PA = LU$, kde $L = (E'_k \dots E'_1)^{-1}$ je hledaná dolní dolní trojúhelníková matice s jedničkami na diagonále a $P = E_{i_1 j_1} \cdots E_{i_r j_r}$ je takzvaná *permutační matice*, neboť způsobuje permutaci řádků matice A (srov. problém 4.2).

Důsledek 3.50. *Každá matice $A \in \mathbb{R}^{n \times n}$ jde rozložit na tvar $PA = LU$, kde $P \in \mathbb{R}^{n \times n}$ je permutační matice, $L \in \mathbb{R}^{n \times n}$ je dolní trojúhelníková matice s jedničkami na diagonále a $U \in \mathbb{R}^{n \times n}$ horní trojúhelníková matice.*

3.5.3 Iterativní metody

Kromě přímých metod typu Gaussovy eliminace na řešení soustav lineárních rovnic existují i iterativní metody, které od počátečního vektoru postupně konvergují k řešení soustavy. Výhoda iterativních metod je menší citlivost k zaokrouhlovacím chybám, a menší časové a paměťové nároky pro velké a řídké soustavy (řídké soustavy mají jen malý zlomek hodnot nenulových, většina jsou nuly).

Jedna ze základních iterativních metod je Gaussova–Seidelova metoda, ukážeme si ji na konkrétním příkladu.

Příklad 3.51. Uvažujme soustavu

$$\left.\begin{array}{r} 6x + 2y - z = 4 \\ x + 5y + z = 3 \\ 2x + y + 4z = 27 \end{array}\right\} \quad \begin{array}{l} x = \frac{1}{6}(4 - 2y + z) \\ y = \frac{1}{5}(3 - x - z) \\ z = \frac{1}{4}(27 - 2x - y) \end{array}$$

Zvolme počáteční hodnoty $x^{(0)} = y^{(0)} = z^{(0)} = 1$. Iterační krok je pak:

$$\begin{aligned} x^{(i)} &= \tfrac{1}{6}(4 - 2y^{(i-1)} + z^{(i-1)}), \\ y^{(i)} &= \tfrac{1}{5}(3 - x^{(i)} - z^{(i-1)}), \\ z^{(i)} &= \tfrac{1}{4}(27 - 2x^{(i)} - y^{(i)}). \end{aligned}$$

Průběh prvních šesti iterací:

iterace	$x^{(i)}$	$y^{(i)}$	$z^{(i)}$
0	1	1	1
1	0.5	0.3	6.425
2	1.6375	−1.0125	6.184375
3	2.034896	−1.043854	5.993516
4	2.013537	−1.001411	5.993584
5	1.999401	−0.998597	5.999949
6	1.999624	−0.999895	6.000212

Vidíme, že již po několika iteracích máme aproximaci skutečného řešení $(2, -1, 6)^T$. □

3.6 Aplikace

Příklad 3.52 (Leontiefův model ekonomiky)**.** Leontiefův[2)] vstupně-výstupní model ekonomiky popisuje vztahy mezi různými odvětvími ekonomiky.

Uvažujme ekonomiku s n sektory (např. zemědělství, průmysl, dopravu, atp.). Sektor i vyrábí jednu komoditu o množství x_i. Předpokládejme, že výroba jednotky j-té komodity potřebuje a_{ij} jednotek i-té komodity. Označme d_i výsledný požadavek na výrobu sektoru i. Nyní nás model vypadá

$$x_i = a_{i1}x_1 + \ldots + a_{in}x_n + d_i, \quad i = 1, \ldots, n.$$

V maticové formě

$$x = Ax + d,$$

neboli

$$(I_n - A)x = d.$$

Problém tedy vede na řešení soustavy lineárních rovnic, kde $x = (x_1, \ldots, x_n)^T$ je vektor neznámých a $A \in \mathbb{R}^n$, $d \in \mathbb{R}^n$ jsou dány. Řešení má explicitní vyjádření $x = (I_n - A)^{-1}d$ a za mírných předpokladů na matici A se dá (s pokročilými znalostmi maticové teorie) ukázat, že řešení je nezáporné, což odpovídá našemu očekávání.

Leontief model aplikoval na ekonomiku USA ve 40. letech 20. století. Ekonomiku rozdělil na 500 sektorů, což v tehdejší době to byla příliš velká soustava na vyřešení, a tak zredukoval model na 42 sektorů. Z praktických důvodů ukážeme zjednodušený model se třemi sektory zemědělství, zpracovatelský průmysl a služby. Reálná data z roku 1947 jsou

$$x = Ax + d = \begin{pmatrix} 0.4102 & 0.0301 & 0.0257 \\ 0.0624 & 0.3783 & 0.1050 \\ 0.1236 & 0.1588 & 0.1919 \end{pmatrix} x + \begin{pmatrix} 39.24 \\ 60.02 \\ 130.65 \end{pmatrix},$$

kde hodnoty složek x a d jsou v mld. \$. Výsledné řešení je $x = (82.4, 138.85, 201.57)^T$, tedy celková produkce v zemědělství musí být 82.4 mld. \$, ve zpracovatelském průmyslu 138.85 mld. \$ a ve službách 201.57 mld. \$. □

Příklad 3.53 (Interpolace polynomem)**.** Uvažujme problém interpolace bodů polynomem. Mějme v rovině $n+1$ bodů (x_0, y_0), (x_1, y_1),…,(x_n, y_n), kde $x_i \neq x_j$ pro $i \neq j$. Cílem je najít polynom $p(x)$ procházející těmito body.

[2)] Wassily Leontief (1906–1999) byl rusko-americký ekonom. Jeho model ekonomiky pochází z 30. let 20. století, a byl za něj roku 1973 oceněn Nobelovou cenou za ekonomii.

Hledejme interpolační polynom ve tvaru $p(x) = a_n x^n + \ldots + a_1 x + a_0$. Dosadíme-li dané body, dostáváme soustavu rovnic

$$\begin{aligned} a_n x_0^n + \ldots + a_1 x_0 + a_0 &= y_0, \\ a_n x_1^n + \ldots + a_1 x_1 + a_0 &= y_1, \\ &\vdots \\ a_n x_n^n + \ldots + a_1 x_n + a_0 &= y_n. \end{aligned}$$

Rovnic je $n+1$ a proměnných také, jsou to koeficienty $a_n, \ldots, a_0$. Máme tedy soustavu rovnic se čtvercovou maticí, nazývanou Vandermondova[3]. Proto jsme také hledali polynom $p(x)$ stupně n; polynom menšího stupně by nemusel existovat (více rovnic než proměnných) a naopak polynom vyššího stupně by nemusel být jednoznačný (více proměnných než rovnic). Jak ukážeme dole, naše čtvercová matice je regulární a proto polynom stupně n vždy existuje a je určený jednoznačně. Polynom pak získáme vyřešením soustavy rovnic.

Regularita matice soustavy se ukáže následujícími elementárními úpravami. Protože se regularita zachovává maticovou transpozicí (tvrzení 3.32), můžeme provádět elementární úpravy i na sloupce. Nejprve od každého sloupce kromě posledního odečteme x_n-násobek sloupce těsně napravo a pak pro $i = 1, \ldots, n$ vydělíme i-tý řádek číslem $x_{i-1} - x_n$. Až na poslední řádek a sloupec dostaneme Vandermondovu matici menšího řádu, kterou upravujeme rekurzivně dále.

$$\begin{pmatrix} x_0^n & \ldots & x_0^2 & x_0 & 1 \\ x_1^n & \ldots & x_1^2 & x_1 & 1 \\ \vdots & \vdots & & & \vdots \\ x_n^n & \ldots & x_n^2 & x_n & 1 \end{pmatrix} \sim \begin{pmatrix} (x_0 - x_n)x_0^{n-1} & \ldots & (x_0 - x_n)x_0 & x_0 - x_n & 1 \\ (x_1 - x_n)x_1^{n-1} & \ldots & (x_1 - x_n)x_1 & x_1 - x_n & 1 \\ \vdots & \vdots & & & \vdots \\ (x_n - x_n)x_n^{n-1} & \ldots & (x_n - x_n)x_n & x_n - x_n & 1 \end{pmatrix} \sim$$

$$\sim \begin{pmatrix} x_0^{n-1} & \ldots & x_0 & 1 & \frac{1}{x_0 - x_n} \\ x_1^{n-1} & \ldots & x_1 & 1 & \frac{1}{x_1 - x_n} \\ \vdots & \vdots & & & \vdots \\ x_{n-1}^{n-1} & \ldots & x_{n-1} & 1 & \frac{1}{x_{n-1} - x_n} \\ 0 & \ldots & 0 & 0 & 1 \end{pmatrix} \sim \cdots \sim \begin{pmatrix} 1 & & & \\ 0 & \ddots & & \\ \vdots & \ddots & \ddots & \\ 0 & \ldots & 0 & 1 \end{pmatrix}.$$

Interpolaci polynomem z pohledu vektorových prostorů nastíníme později v příkladu 5.71. □

Příklad 3.54 (Diskrétní a rychlá Fourierova transformace)**.** Z předchozího příkladu máme v zásadě dvě možné reprezentace polynomu $p(x)$, první je základní tvar $p(x) = a_n x^n + \ldots + a_1 x + a_0$ a druhá je seznamem funkčních hodnot v $n+1$ různých bodech. Mezi těmito reprezentacemi můžeme snadno přecházet. Z první na druhou stačí zvolit libovolných $n+1$ bodů a spočítat funkční hodnoty, a druhý směr jsme rozebírali nahoře.

Která reprezentace je výhodnější? Každá na něco jiného. Dejme tomu, že chceme umět efektivně sčítat a násobit polynomy. V první reprezentaci je sčítání jednoduché, stojí nás to řádově n aritmetických operací,

[3] Francouzský matematik Alexandre–Théophile Vandermonde (1735–1796) byl i jedním ze zakladatelů matematické teorie uzlů.

ale vynásobit polynomy dá trochu zabrat, to už stojí řádově n^2 aritmetických operací. V druhé reprezentaci stojí sčítání i násobení řádově n, pokud známe funkční hodnoty polynomů ve stejných bodech. Toto by nás mohlo inspirovat k tomu násobit polynomy v základním tvaru tak, že spočítáme funkční hodnoty ve vhodných bodech, vynásobíme a převedeme zpět. A skutečně, pokud se zvolí vhodné body, lze transformaci mezi reprezentacemi implementovat tak, že stojí řádově $n\log(n)$ aritmetických operací, a tolik řádově stojí i výsledné násobení polynomů. Těmto transformacím se říká Fourierova transformace, viz příklad 10.36 a [Tůma, 2003, kap. 15].

Umět rychle násobit polynomy je velmi důležité. Například i obyčejné násobení reálných čísel v desetinném rozvoji si lze představit jako násobení polynomů. □

Příklad 3.55 (Komprese obrázku)**.** Ukážeme jednoduchou kompresi obrázku pomocí tzv. Haarovy transformace (jinou kompresi zmíníme v příkladu 13.21). Předpokládejme, že matice $M \in \mathbb{R}^{m\times n}$ reprezentuje obrázek, ve kterém pixel na pozici (i,j) má barvu s číslem m_{ij}. Rozdělíme matici M na jednotlivé podmatice A velikosti 8×8, které budeme postupně komprimovat (pro jednoduchost předpokládáme, že čísla m,n jsou dělitelná osmi).

Základní myšlenka komprese spočívá v takové transformaci matice A, abychom dostali v matici co nejvíce nul – protože pak stačí ukládat pouze nenulová čísla. Kompresi rozdělíme do tří kroků. Nejprve sdružíme prvky matice do dvojic $a_{ij}, a_{i,j+1}$, $i \in \{1,\dots,8\}$, $j \in \{1,3,5,7\}$. Jejich průměr $\frac{1}{2}(a_{ij}+a_{i,j+1})$ uložíme do matice 8×4 a zprava doplníme tuto matici o jejich odchylky $\frac{1}{2}(a_{ij}-a_{i,j+1})$. Výslednou matici z matice A dostaneme alternativně tak, že matici A vynásobíme zprava regulární maticí

$$H_8 = \begin{pmatrix} \frac12 & 0 & 0 & 0 & \frac12 & 0 & 0 & 0 \\ \frac12 & 0 & 0 & 0 & -\frac12 & 0 & 0 & 0 \\ 0 & \frac12 & 0 & 0 & 0 & \frac12 & 0 & 0 \\ 0 & \frac12 & 0 & 0 & 0 & -\frac12 & 0 & 0 \\ 0 & 0 & \frac12 & 0 & 0 & 0 & \frac12 & 0 \\ 0 & 0 & \frac12 & 0 & 0 & 0 & -\frac12 & 0 \\ 0 & 0 & 0 & \frac12 & 0 & 0 & 0 & \frac12 \\ 0 & 0 & 0 & \frac12 & 0 & 0 & 0 & -\frac12 \end{pmatrix}.$$

V druhém kroku aplikujeme analogický postup na matici AH_8, ale tentokrát jen na první čtyři sloupce. Výsledek tedy dostaneme vynásobením matice AH_8 zprava maticí

$$\begin{pmatrix} H_4 & 0_{4,4} \\ 0_{4,4} & I_4 \end{pmatrix}.$$

Ve třetím kroku pak stejný postup použijeme pouze na první dva sloupce výsledné matice, což vyjádříme vynásobením zprava maticí

$$\begin{pmatrix} H_2 & 0_{2,6} \\ 0_{6,2} & I_6 \end{pmatrix}.$$

V souhrnu můžeme všechny tři kroky vyjádřit jako jediné maticové násobení AH, kde

$$H = H_8 \begin{pmatrix} H_4 & 0_{4,4} \\ 0_{4,4} & I_4 \end{pmatrix} \begin{pmatrix} H_2 & 0_{2,6} \\ 0_{6,2} & I_6 \end{pmatrix} = \begin{pmatrix} \frac18 & \frac18 & \frac14 & 0 & \frac12 & 0 & 0 & 0 \\ \frac18 & \frac18 & \frac14 & 0 & -\frac12 & 0 & 0 & 0 \\ \frac18 & \frac18 & -\frac14 & 0 & 0 & \frac12 & 0 & 0 \\ \frac18 & \frac18 & -\frac14 & 0 & 0 & -\frac12 & 0 & 0 \\ \frac18 & -\frac18 & 0 & \frac14 & 0 & 0 & \frac12 & 0 \\ \frac18 & -\frac18 & 0 & \frac14 & 0 & 0 & -\frac12 & 0 \\ \frac18 & -\frac18 & 0 & -\frac14 & 0 & 0 & 0 & \frac12 \\ \frac18 & -\frac18 & 0 & -\frac14 & 0 & 0 & 0 & -\frac12 \end{pmatrix}.$$

Tuto sérii úprav budeme aplikovat i na řádky, tudíž výsledná matice A' po transformaci se dá vyjádřit jako $A' = H^TAH$. Protože matice H je regulární, mohu se vrátit k původní matici A úpravou $A = H^{-T}A'H^{-1}$.

Protože často v obrázku mají sousední pixely stejnou či podobnou barvu, takovýmto průměrováním hodnot příslušné matice můžeme dostat po transformaci hodně nul nebo nule blízkých čísel.

Výše zmíněná komprese je bezeztrátová. Vyšší účinnosti komprese můžeme dosáhnout pokud v matici A' vynulujeme všechny hodnoty, které jsou v absolutní hodnotě menší než pevná hranice $\varepsilon > 0$. Tento přístup už vede ke ztrátě určité informace. Poměr nenulových čísel v matici A' před a po vynulování malých hodnot se pak nazývá *kompresní poměr*.

Následující obrázky ilustrují kompresi pro různou volbu kompresního poměru k. Ve skutečnosti matici H ještě před použitím upravíme tak, že každý sloupeček vydělíme jeho eukleidovskou normou – výsledná matice je tzv. ortogonální (viz sekce 8.6) a má lepší vlastnosti.

originál ($k = 1$)

$k = 10$

$k = 50$

$k = 100$ □

Problémy

3.1. Buď $A \in \mathbb{R}^{n\times n}$, $b \in \mathbb{R}^n$. Navrhněte co nejefektivnější způsob výpočtu $A^k b$ v závislosti na k, n, měříme-li efektivitu počtem aritmetických operací s čísly (pro jednoduchost počítejte jen součiny čísel).

3.2. Dokažte, že soustava $Ax = b$ má řešení právě tehdy když $A^T y = 0,\ b^T y = 1$ nemá řešení.

3.3. Na řešení soustavy rovnic $Ax = b$ s regulární maticí se můžeme dívat jako na funkci $(A \mid b) \mapsto x = A^{-1}b$, která vstupním hodnotám, reprezentovaným prvky matice A a vektoru b, přiřadí řešení soustavy. Znáte-li derivace, spočítejte derivaci této funkce podle a_{ij}.

3.4. *(Souboj o regularitu)* René a Simona hrají hru s maticí řádu $n \geq 2$. René přiřadí nějakému políčku libovolné reálné číslo, pak Simona přiřadí jinému políčku číslo atd. dokud se nezaplní celá matice. René vyhraje, pokud je výsledná matice regulární, a Simona vyhraje, pokud je singulární. Má někdo vítěznou strategii? A jaká bude situace, když začne Simona?

Shrnutí ke kapitole 3. Matice

Matice je jeden ze základních objektů ve vědeckém počítání. Matice kompaktním způsobem reprezentuje data, zobrazení, vztahy aj., a proto různé charakteristiky matic odráží zajímavé vlastnosti původního problému.

Abychom s maticemi mohli dobře pracovat, zavedli jsme základní operace sčítání, násobku, součinu a transpozice. Existuje celá řada speciálních typů matic. Důležitým typem jsou regulární matice, které odpovídají soustavám lineárních rovnic s jediným řešením pro libovolnou pravou stranu. Navíc jsou to přesně ty matice, pro které existuje inverzní matice pro operaci maticového součinu. Dají se také poskládat jako součin několika matic elementárních úprav. Regulární matice tedy odpovídají ekvivalentním úpravám – přenásobením soustavy regulární maticí se množina řešení nezmění.

Pokud maticově analyzujeme Gaussovu eliminaci, dospějeme k tomu, že (za určitých předpokladů) lze matice soustavy zapsat jako součin dolní a horní trojúhelníkové matice. Tento LU rozklad se hodí nejen pro některé výpočty s maticemi, ale i pro teoretický rozbor Gaussovy eliminace.

Kapitola 4

Grupy a tělesa

Tato kapitola je věnovaná základním algebraickým strukturám jako jsou grupy a tělesa. Jsou to abstraktní pojmy zobecňující dobře známé obory reálných (racionálních, komplexních aj.) čísel s operacemi sčítání a násobení.

Pro hlubší informace a souvislosti doplněné mnoha názornými vizualizacemi doporučuji knihu Carter [2009].

4.1 Grupy

Pojem grupy zavedl francouzský matematik Èvariste Galois (1811–1832) při budování teorie řešitelnosti hledání kořenů polynomů. Pro kořeny polynomů stupně alespoň 5 neexistuje obecně žádný vzoreček, ale Galoisova teorie dává návod, jak to otestovat pro konkrétní polynom, tj. jestli kořeny daného polynomu jdou vyjádřit pomocí základních aritmetických operací a odmocnin. Příkladem situace, kdy to nelze, je polynom $x^5 - 2x - 1$.

Grupy mají však mnohem širší použití. Díky jejich obecnosti a abstraktnosti je můžeme najít v různých oborech: fyzika (Lieovy grupy), architektura (Friezovy grupy), geometrie a molekulární chemie (symetrické grupy) aj.

Definice 4.1 (Grupa)**.** Buď $\circ\colon G^2 \to G$ binární operace na množině G. Pak *grupa* je dvojice $(G, \circ)$ splňující:

(1) $\forall a, b, c \in G : a \circ (b \circ c) = (a \circ b) \circ c$ (asociativita),

(2) $\exists e \in G\ \forall a \in G : e \circ a = a \circ e = a$ (existence neutrálního prvku),

(3) $\forall a \in G\ \exists b \in G : a \circ b = b \circ a = e$ (existence inverzního prvku).

Abelova (komutativní) grupa je taková grupa, která navíc splňuje:

(4) $\forall a, b \in G\ a \circ b = b \circ a$ (komutativita).

Výše zmíněné podmínky se také občas nazývají axiomy. Poznamenejme, že implicitně je v definici grupy schovaná podmínka uzavřenosti, aby výsledek operace nevypadl ven z množiny G, tedy aby $\forall a, b \in G : a \circ b \in G$. Pokud je operací $\circ$ sčítání, většinou se značí neutrální prvek 0 a inverzní $-a$, pokud jde o násobení, neutrální prvek se označuje 1 a inverzní a^{-1}.

Poznámka 4.2 (Definice konstrukcí vs. axiomy)**.** Matematický objekt lze zavést buď konstrukcí z nějakých již vytvořených objektů, nebo specifikací vlastností (axiomů), které má splňovat. Definice grupy padá do druhé skupiny, podobně jako definice tělesa v sekci 4.3 či vektorových prostorů v kapitole 5. Grupou pak je jakýkoli objekt, který splňuje dané vlastnosti. Axiomatická definice má tu výhodu, že nás nesvazuje s jedním konkrétním objektem – jakoukoli vlastnost, kterou odvodíme pro axiomaticky definovaný objekt potom automaticky platí pro každý konkrétní případ. Mnoho konkrétních příkladů grup ukazujeme v následujícím odstavci.

Příklad 4.3. Příklady grup:

- Dobře známé $(\mathbb{Z}, +)$, $(\mathbb{Q}, +)$, $(\mathbb{R}, +)$, $(\mathbb{C}, +)$.

- Grupy matic $(\mathbb{R}^{m\times n}, +)$.
- Množina $\mathbb{Z}_n := \{0, 1, \dots, n-1\}$ a sčítání modulo n, značení $(\mathbb{Z}_n, +)$.
- Známé obory s násobením, např. $(\mathbb{Q} \setminus \{0\}, \cdot)$, $(\mathbb{R} \setminus \{0\}, \cdot)$.
- Množina reálných polynomů proměnné x se sčítáním.

Výše zmíněné grupy jsou Abelovy. Z určitého pohledu jsou důležitější neabelovské grupy. Dva důležité příklady neabelovských grup jsou:

- Zobrazení na množině s operací skládání, např. rotace v $\mathbb{R}^n$ podle počátku nebo později probírané permutace (sekce 4.2),
- Regulární matice s násobením (tzv. maticová grupa).

Příklady negrup:

- $(\mathbb{N}, +)$, $(\mathbb{Z}, -)$, $(\mathbb{R} \setminus \{0\}, :)$, ... □

Tvrzení 4.4 (Základní vlastnosti v grupě)**.** *Pro prvky grupy $(G, \circ)$ platí následující vlastnosti.*

(1) $a \circ c = b \circ c$ *implikuje* $a = b$ *(tzv. krácení),*

(2) neutrální prvek e je určen jednoznačně,

(3) pro každé $a \in G$ je jeho inverzní prvek určen jednoznačně,

(4) rovnice $a \circ x = b$ má právě jedno řešení pro každé $a, b \in G$,

(5) $(a^{-1})^{-1} = a$,

(6) $(a \circ b)^{-1} = b^{-1} \circ a^{-1}$.

Důkaz. Ukážeme jen několik vlastností.

(1)

$$\begin{aligned} a \circ c &= b \circ c && / \circ c^{-1} \text{ zprava} \\ a \circ (c \circ c^{-1}) &= b \circ (c \circ c^{-1}) \\ a \circ e &= b \circ e \\ a &= b \end{aligned}$$

(2) Existují-li dva různé neutrální prvky e_1, e_2, pak $e_1 = e_1 \circ e_2 = e_2$, což je spor.

(3) Existují-li k $a \in G$ dva různé inverzní prvky a_1, a_2, pak $a \circ a_1 = e = a \circ a_2$ a z vlastnosti krácení dostáváme $a_1 = a_2$, což je spor.

(4) Vynásobme rovnost $a \circ x = b$ zleva prvkem a^{-1} a dostaneme jediného kandidáta $x = a^{-1} \circ b$. Dosazením ověříme, že rovnost splňuje. □

Tak jako množiny doprovází pojem podmnožina, tak nelze mluvit o grupách a nezmínit podgrupy.

Definice 4.5 (Podgrupa)**.** *Podgrupa* grupy $(G, \circ)$ je grupa $(H, \diamond)$ taková, že $H \subseteq G$ a pro všechna $a, b \in H$ platí $a \circ b = a \diamond b$. Značení: $(H, \diamond) \leq (G, \circ)$.

Jinými slovy, se stejně definovanou operací splňuje H vlastnosti uzavřenost a existence neutrálního a inverzního prvku. To jest, pro každé $a, b \in H$ je $a \circ b \in H$, dále $e \in H$, a pro každé $a \in H$ je $a^{-1} \in H$.

Příklad 4.6.

- Každá grupa $(G, \circ)$ má dvě triviální podgrupy: sama sebe $(G, \circ)$ a $(\{e\}, \circ)$.
- $(\mathbb{N}, +) \not\leq (\mathbb{Z}, +) \leq (\mathbb{Q}, +) \leq (\mathbb{R}, +) \leq (\mathbb{C}, +)$. □

Příklad 4.7. Ukažte, že podgrupy jsou uzavřené na průnik, ale ne na sjednocení. Jinými slovy, ukažte, že průnik dvou podgrup grupy $(G, \circ)$ je opět její podgrupa a najděte příklad, kdy sjednocení podgrup již podgrupa není. □

4.2 Permutace

Důležitým příkladem grup je takzvaná symetrická grupa permutací, proto si povíme něco více o permutacích. Připomeňme, že vzájemně jednoznačné zobrazení $f\colon X \to Y$ je zobrazení, které je prosté (žádné dva různé prvky se nezobrazí na jeden) a „na“ (pokryje celou množinu Y).

Definice 4.8 (Permutace)**.** *Permutace* na konečné množině X je vzájemně jednoznačné zobrazení $p\colon X \to X$.

Většinou budeme uvažovat $X = \{1, \dots, n\}$. Množina všech permutací na množině $\{1, \dots, n\}$ se pak značí S_n.

Zadání permutace je možné například:

- Tabulkou, kde nahoře jsou vzory a dole jejich obrazy

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

- Grafem vyznačujícím kam se který prvek zobrazí

- Rozložením na cykly

$$p = (1, 2)(3)(4, 5, 6),$$

kde každá závorka uvádí seznam prvků v jednom cyklu. Tedy $(a_1, \dots, a_k)$ znamená, že a_1 se zobrazí na a_2, a_2 se zobrazí na a_3, atd. až a_{k-1} se zobrazí na a_k. Z definice je patrné, že každou permutaci lze rozložit na disjunktní cykly. V následujícím textu budeme nejčastěji používat redukovaný zápis

$$p = (1, 2)(4, 5, 6),$$

ve kterém vynecháváme cykly délky 1.

Příkladem jednoduché, ale důležité, permutace je *transpozice*, což je permutace $= (i, j)$ s jedním cyklem délky 2 prohazující dva prvky. Jednodušší už je jenom identita id zobrazující každý prvek na sebe.

Inverzní permutace a skládání permutací je definováno stejně jako pro jiná zobrazení:

Definice 4.9 (Inverzní permutace)**.** Buď $p \in S_n$. *Inverzní permutace* k p je permutace p^{-1} definovaná $p^{-1}(i) = j$, pokud $p(j) = i$.

Příklad 4.10. $(i, j)^{-1} = (i, j)$, $(i, j, k)^{-1} = (k, j, i)$, ... □

Definice 4.11 (Skládání permutací)**.** Buďte $p, q \in S_n$. *Složená permutace* $p \circ q$ je permutace definovaná $(p \circ q)(i) = p(q(i))$.

Příklad 4.12. $id \circ p = p \circ id = p$, $p \circ p^{-1} = p^{-1} \circ p = id$, ... □

Skládání permutací je asociativní (jako každé zobrazení), ale komutativní obecně není. Např. pro $p = (1, 2)$, $q = (1, 3, 2)$ máme $p \circ q = (1, 3)$, ale $q \circ p = (2, 3)$.

Významná charakteristika permutace je tzv. znaménko.

Definice 4.13 (Znaménko permutace)**.** Nechť se permutace $p \in S_n$ skládá z k cyklů. Pak *znaménko permutace* je číslo $\operatorname{sgn}(p) = (-1)^{n-k}$.

Příklad 4.14. $\operatorname{sgn}(id) = 1$, $\operatorname{sgn}((i, j)) = -1$, ... □

Znaménko je vždy 1 nebo -1. Podle toho se též rozdělují permutace na *sudé* (ty, co mají znaménko 1) a na *liché* (ty se znaménkem -1).

Věta 4.15 (O znaménku složení permutace a transpozice)**.** *Buď $p \in S_n$ a buď $t = (i, j)$ transpozice. Pak* $\operatorname{sgn}(p) = -\operatorname{sgn}(t \circ p) = -\operatorname{sgn}(p \circ t)$.

Důkaz. Dokážeme $\operatorname{sgn}(p) = -\operatorname{sgn}(t \circ p)$, druhá rovnost je analogická. Permutace p se skládá z několika cyklů. Rozlišme dva případy:

Nechť i, j jsou částí stejného cyklu, označme jej $(i, u_1, \ldots, u_r, j, v_1, \ldots, v_s)$. Pak

$$(i, j) \circ (i, u_1, \ldots, u_r, j, v_1, \ldots, v_s) = (i, u_1, \ldots, u_r)(j, v_1, \ldots, v_s),$$

tedy počet cyklů se zvýší o jedna. Viz obrázek, kde černé šipky znázorňují původní cyklus a plné šipky nové dva cykly:

Nechť i, j náleží do dvou různých cyklů, např. $(i, u_1, \ldots, u_r)(j, v_1, \ldots, v_s)$. Pak

$$(i, j) \circ (i, u_1, \ldots, u_r)(j, v_1, \ldots, v_s) = (i, u_1, \ldots, u_r, j, v_1, \ldots, v_s),$$

tedy počet cyklů se sníží o jedna.

V každém případě se počet cyklů změní o jedna, a tudíž i výsledné znaménko. □

Věta 4.16. *Každou permutaci lze rozložit na složení transpozic.*

Důkaz. Rozložíme na transpozice postupně všechny cykly permutace. Libovolný cyklus $(u_1, \ldots, u_r)$ se rozloží

$$(u_1, \ldots, u_r) = (u_1, u_2) \circ (u_2, u_3) \circ (u_3, u_4) \circ \ldots \circ (u_{r-1}, u_r).$$ □

Poznamenejme, že rozklad na transpozice není jednoznačný, dokonce ani počet transpozic ne. Pouze jejich parita zůstane stejná.

Výše zmíněné vlastnosti mají řadu pěkných důsledků.

Důsledek 4.17. *Platí* $\operatorname{sgn}(p) = (-1)^r$*, kde r je počet transpozic při rozkladu p na transpozice.*

Důkaz. Je to důsledek věty 4.15. Vyjdeme z identity, která je sudá. Každá transpozice mění znaménko, tedy výsledné znaménko bude $(-1)^r$. □

Důsledek 4.18. *Buď $p, q \in S_n$. Pak* $\operatorname{sgn}(p \circ q) = \operatorname{sgn}(p)\operatorname{sgn}(q)$.

Důkaz. Nechť p se dá rozložit na r_1 transpozic a q na r_2 transpozic. Tedy $p \circ q$ lze složit z $r_1 + r_2$ transpozic. Pak $\operatorname{sgn}(p \circ q) = (-1)^{r_1+r_2} = (-1)^{r_1}(-1)^{r_2} = \operatorname{sgn}(p)\operatorname{sgn}(q)$. □

Důsledek 4.19. *Buď* $p \in S_n$. *Pak* $\operatorname{sgn}(p) = \operatorname{sgn}(p^{-1})$.

Důkaz. Platí $1 = \operatorname{sgn}(id) = \operatorname{sgn}(p \circ p^{-1}) = \operatorname{sgn}(p)\operatorname{sgn}(p^{-1})$, tedy p, p^{-1} musí mít stejné znaménko. □

Poznámka 4.20. Kromě počtu cyklů a počtu transpozic jde znaménko permutace p zavést také např. pomocí počtu inverzí. Inverzí zde rozumíme uspořádanou dvojici (i, j) takovou, že $i < j$ a $p(i) > p(j)$. Označíme-li počet inverzí permutace p jako $I(p)$, pak platí $\operatorname{sgn}(p) = (-1)^{I(p)}$.

Poznámka 4.21 (Symetrická grupa)**.** Vraťme se zpět ke grupám. Množina permutací S_n s operací skládání $\circ$ tvoří nekomutativní grupu $(S_n, \circ)$, tzv. *symetrickou grupu.* Ta hraje důležitou roli v algebře, protože se dá ukázat, že každá grupa je isomorfní nějaké symetrické grupě či její podgrupě (tzv. Cayleyova reprezentace, dokonce platí i zobecnění na nekonečné grupy). Podobnou roli hrají maticové grupy, protože každá konečná grupa je isomorfní nějaké maticové podgrupě (lineární reprezentace) s tím, že těleso, nad kterým s maticemi pracujeme, si můžeme dopředu zvolit.

Grupa $(S_n, \circ)$ se nazývá symetrická, protože ona a její podgrupy popisují symetrie různých objektů. Kupříkladu rovnoramenný trojúhelník dole na obrázku je symetrický podle svislé osy, a této symetrii odpovídá permutace $(2, 3)$. Uvažujeme-li ještě základní symetrii danou podobností trojúhelníku se sebou samým, které odpovídá permutace id, potom symetrie trojúhelníku odpovídají podgrupě $\{id,\ (2, 3)\}$.

Symetrie následujícího objektu jsou rotace o 0°, 120° a o 240°. Tyto symetrie odpovídají permutacím id, $(1, 2, 3)$ a $(1, 3, 2)$. Tudíž symetrie tohoto objektu odpovídají podgrupě $\{id,\ (1, 2, 3),\ (1, 3, 2)\}$.

Symetrie rovnostranného trojúhelníku jsou souměrnosti podle těžnic, které odpovídají permutacím $(1, 2)$, $(2, 3)$ a $(1, 3)$, a dále otočení o 0°, 120° a o 240°. Všechny symetrie tedy představují celou grupu $(S_3, \circ)$.

Naopak nesymetrickému trojúhelníku jako na obrázku dole přísluší pouze identita, a proto jeho symetrie představují grupu $\{id\}$.

Loydova patnáctka

Symetrické grupy a znaménko permutace se využijí také při analýze hlavolamů jako je Loydova patnáctka nebo Rubikova kostka. Rubikova kostka vyžaduje trochu hlubší rozbor, proto nahlédneme pod pokličku pouze Loydova patnáctky.

Loydova patnáctka. Cílový stav. [zdroj: Wikipedia]

Loydova patnáctka je hra, která sestává z pole 4×4 a kachlíky očíslovanými 1 až 15. Jedno pole je prázdné a přesouváním sousedních kachlíků na prázdné pole měníme rozložení kachlíků. Cílem je dospět pomocí těchto přesunů k vzestupnému uspořádání kachlíků tak, jak je uvedeno na obrázku.

Otázka zní, které počáteční konfigurace jsou řešitelné a které ne. Jestliže očíslujeme 1 až 16 jednotlivá políčka, pak konfigurace kachlíků odpovídá nějaké permutaci $p \in S_{16}$ a přesun kachlíku odpovídá složení p s transpozicí. Označíme-li (r, s) pozici prázdného pole, pak hodnota $h = (-1)^{r+s} \operatorname{sgn}(p)$ zůstává po celou hru stejná, protože každý posun kachlíku změní o jedničku buď r nebo s, ale zároveň posun kachlíku odpovídá složení p s odpovídající transpozicí, čili i $\operatorname{sgn}(p)$ změní znaménko.

Cílová konfigurace má hodnotu $h = 1$, tedy počáteční konfigurace s $h = -1$ řešitelné být nemohou. Detailnější analýza [Výborný a Zahradník, 2002] ukáže, že všechny počáteční konfigurace s $h = 1$ už řešitelné jsou.

4.3 Tělesa

Algebraická tělesa zobecňují třídu tradičních číselných oborů jako je třeba množina reálných čísel na abstraktní množinu se dvěma operacemi a řadou vlastností. To nám umožní pracovat s maticemi (sčítat, násobit, invertovat, řešit soustavy rovnic, ...) nad jinými obory než jen nad $\mathbb{R}$.

Definice 4.22 (Těleso)**.** *Těleso* je množina $\mathbb{T}$ spolu se dvěma komutativními binárními operacemi $+$ a $\cdot$ splňující

(1) $(\mathbb{T}, +)$ je Abelova grupa, neutrální prvek značíme 0 a inverzní k a pak $-a$,

(2) $(\mathbb{T} \setminus \{0\}, \cdot)$ je Abelova grupa, neutrální prvek značíme 1 a inverzní k a pak a^{-1},

(3) $\forall a, b, c \in \mathbb{T}$: $a(b + c) = ab + ac$ (distributivita).

Z definice tělesa nutně vyplývá, že $0 \neq 1$.

Těleso se občas zavádí bez komutativity násobení a tělesu s komutativním násobením se pak říká komutativní těleso nebo pole, ale pro naše účely budeme komutativitu automaticky předpokládat. Podobně jako podgrupy můžeme zavést i pojem podtěleso jako podmnožinu tělesa, která se stejně definovanými operacemi tvoří těleso.

Příklad 4.23. Příkladem nekonečných těles je např. $\mathbb{Q}$, $\mathbb{R}$ nebo $\mathbb{C}$. Množina celých čísel $\mathbb{Z}$ ale těleso netvoří, protože chybí inverzní prvky pro násobení, (např. když invertujeme hezkou celočíselnou matici, tak často vycházejí zlomky a tím pádem se dostáváme mimo obor $\mathbb{Z}$). Těleso netvoří ani čísla reprezentovaná na počítači v aritmetice s pohyblivou desetinnou čárkou – jednak nejsou operace sčítání a násobení uzavřené (pokud by výsledkem bylo hodně velké či hodně malé číslo), a jednak nejsou ani asociativní (díky zaokrouhlování). Konečná tělesa prozkoumáme později. □

Příklad 4.24 (Kvaterniony)**.** Oblíbeným příkladem těles jsou *kvaterniony.* Jedná se o zobecnění komplexních čísel přidáním dalších dvou imaginárních jednotek j a k, jejichž druhá mocnina je -1. Zatímco sčítání se definuje přirozeně, násobení je trochu komplikovanější a neplatí už pro něj komutativita. Kvaterniony pak tudíž tvoří nekomutativní těleso. Pomocí kvaternionů se dobře popisují rotace ve třírozměrném prostoru a našly využití i ve fyzikální kvantové teorii. □

Řadu pěkných vlastností zdědí těleso z vlastností příslušných grup $(\mathbb{T}, +)$ a $(\mathbb{T} \setminus \{0\}, \cdot)$. Např. distributivita zprava $(b + c)a = ba + ca$ plyne z levé distributivity a komutativity násobení. Některé specifické vlastnosti uvádíme v následující větě.

Tvrzení 4.25 (Základní vlastnosti v tělese)**.** *Pro prvky tělesa platí následující vlastnosti.*

(1) $0a = 0$,

(2) $ab = 0$ *implikuje, že* $a = 0$ *nebo* $b = 0$,

(3) $-a = (-1)a$.

Důkaz.

(1) Odvodíme

$$\begin{aligned} 0a = (0+0)a &= 0a + 0a && / + (-0a) \\ (-0a) + 0a = (0+0)a &= (-0a) + 0a + 0a \\ 0 &= 0 + 0a \\ 0 &= 0a \end{aligned}$$

(2) Je-li $a = 0$, pak věta platí. Je-li $a \neq 0$, pak existuje a^{-1}. Pronásobením obou stran rovnice $ab = 0$ zleva prvkem a^{-1} dostaneme $a^{-1}ab = a^{-1}0$, neboli $1b = 0$.

(3) Máme $0 = 0a = (1-1)a = 1a + (-1)a = a + (-1)a$, tedy $-a = (-1)a$. □

Druhá vlastnost (a její důkaz) předchozí věty mj. říkají, že při rozhodování, zda nějaká struktura tvoří těleso, nemusíme ověřovat uzavřenost násobení na množině $\mathbb{T} \setminus \{0\}$ (žádné dva nenulové prvky se nevynásobí na nulu), tato vlastnost vyplývá z ostatních. Stačí tedy jen uzavřenost na $\mathbb{T}$, což bývá snáze vidět.

Podívejme se teď na konečná tělesa. Zaveďme množinu $\mathbb{Z}_n = \{0, 1, \ldots, n-1\}$ a operace $+$ a $\cdot$ modulo n. Snadno nahlédneme, že $\mathbb{Z}_2$ a $\mathbb{Z}_3$ jsou tělesy, ale $\mathbb{Z}_4$ už není, neboť prvek 2 nemá inverzi 2^{-1}. Tento výsledek můžeme zobecnit.

Lemma 4.26. *Buď n prvočíslo a buď $0 \neq a \in \mathbb{Z}_n$. Pak*

$$\{0, 1, \ldots, n-1\} = \{0a, 1a, \ldots, (n-1)a\},$$

kde prvky v množině napravo se počítají modulo n.

Důkaz. Sporem předpokládejme, že $ak = a\ell$ pro nějaké $k, \ell \in \mathbb{Z}_n$, $k \neq \ell$. Pak dostáváme $a(k - \ell) = 0$, tudíž buď a nebo $k - \ell$ je dělitelné n. To znamená buď $a = 0$ nebo $k - \ell = 0$. Ani jedna možnost ale nastat nemůže, což je spor. □

Věta 4.27. *$\mathbb{Z}_n$ je těleso právě tehdy, když n je prvočíslo.*

Důkaz. Je-li n složené, pak $n = pq$, kde $1 < p, q < n$. Kdyby $\mathbb{Z}_n$ bylo těleso, pak $pq = 0$ implikuje podle tvrzení 4.25 buď $p = 0$ nebo $q = 0$, ale ani jedno neplatí.

Je-li n prvočíslo, pak se snadno ověří všechny axiomy z definice tělesa. Jediný pracnější může být existence inverze a^{-1} pro libovolné $a \neq 0$. To ale nahlédneme snadno z lemmatu 4.26. Protože $\{0, 1, \ldots, n-1\} = \{0a, 1a, \ldots, (n-1)a\}$, musí být v množině napravo prvek 1, a tudíž existuje $b \in \mathbb{Z}_n \setminus \{0\}$ takové, že $ba = 1$. Proto $b = a^{-1}$. □

Příklad 4.28 (Těleso $\mathbb{Z}_5$)**.** Pro ilustraci uvádíme v tabulkách dole explicitní vyjádření obou operací nad tělesem $\mathbb{Z}_5$:

$+$	0	1	2	3	4
0	0	1	2	3	4
1	1	2	3	4	0
2	2	3	4	0	1
3	3	4	0	1	2
4	4	0	1	2	3

$\cdot$	0	1	2	3	4
0	0	0	0	0	0
1	0	1	2	3	4
2	0	2	4	1	3
3	0	3	1	4	2
4	0	4	3	2	1

V tabulkách se odráží některé základní vlastnosti těles: komutativita se projevuje jako symetrie tabulek, neutrální prvek kopíruje záhlaví tabulky do příslušného řádku a sloupce, a násobení nulou dává nulu. Vlastnost inverzního prvku se pak projevuje tak, že v každém řádku a sloupci (kromě násobení nulou) je uveden každý prvek tělesa právě jednou.

Inverzní prvky tělesa $\mathbb{Z}_5$ jsou pak:

x	0	1	2	3	4
$-x$	0	4	3	2	1

x	0	1	2	3	4
x^{-1}	–	1	3	2	4

□

Příklad 4.29 (Těleso $\mathbb{Z}_2$)**.** Těleso $\mathbb{Z}_2$ má pro informatiky obzvláště velký význam, protože pracuje se dvěma prvky 0 a 1, na které můžeme nahlížet jako na počítačové bity. Je snadné pak nahlédnout, že operace sčítání v $\mathbb{Z}_2$ odpovídá počítačové operaci XOR a násobení odpovídá operaci AND. □

Poznámka 4.30. Soustavy rovnic a operace s maticemi jsme zaváděli nad tělesem reálných čísel. Nicméně nic nám nebrání rozšířit tyto pojmy a pracovat nad jakýmkoli jiným tělesem. Je-li $\mathbb{T}$ těleso, pak $\mathbb{T}^{m\times n}$ bude značit matici řádu $m \times n$ s prvky v tělese $\mathbb{T}$. Jediné vlastnosti reálných čísel, který jsme používali, jsou přesně ty, které se vyskytují v definici tělesa; nepotřebovali jsme čísla odmocňovat ani mezi sebou porovnávat ani nic podobného. Proto veškeré postupy a teorie vybudovaná v předchozích kapitolách 2 a 3 zůstane v platnosti. Můžeme tak například řešit soustavy lineárních rovnic nad libovolným tělesem pomocí Gaussovy eliminace, hovořit o regularitě matice z $\mathbb{T}^{n\times n}$ či hledat její inverzi.

Příklad 4.31 (Výpočet inverzní matice nad $\mathbb{Z}_5$)**.**

$$(A\,|\,I_3) = \left(\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 2 & 0 & 4 & 0 & 1 & 0 \\ 3 & 3 & 4 & 0 & 0 & 1 \end{array}\right) \sim \left(\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & 3 & 3 & 1 & 0 \\ 0 & 2 & 0 & 2 & 0 & 1 \end{array}\right) \sim \left(\begin{array}{ccc|ccc} 1 & 0 & 2 & 0 & 3 & 0 \\ 0 & 1 & 3 & 3 & 1 & 0 \\ 0 & 0 & 4 & 1 & 3 & 1 \end{array}\right) \sim$$

$$\sim \left(\begin{array}{ccc|ccc} 1 & 0 & 2 & 0 & 3 & 0 \\ 0 & 1 & 3 & 3 & 1 & 0 \\ 0 & 0 & 1 & 4 & 2 & 4 \end{array}\right) \sim \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 2 & 4 & 2 \\ 0 & 1 & 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & 4 & 2 & 4 \end{array}\right) = \left(I_3\,|\,A^{-1}\right)$$

□

Poznámka 4.32 (Jak najít inverzi)**.** Přirozená otázka při počítání nad tělesem $\mathbb{Z}_p$ zní, jak najít inverzní prvek k $x \in \mathbb{Z}_p \setminus \{0\}$. Pro malé hodnoty p mohu zkusit postupně $1, 2, \ldots, p-1$ dokud nenarazím na inverzní prvek k x. Pokud p je hodně velké prvočíslo, tento postup už není efektivní a postupuje se tzv. *rozšířeným Eukleidovým algoritmem*, který najde $a, b \in \mathbb{Z}$ taková, že $ax + bp = 1$, z čehož vidíme, že hledanou inverzí x^{-1} je prvek a (resp. jeho zbytek po dělení p).

Nyní víme, že existují tělesa o velikostech odpovídajícím prvočíslům. Existují však tělesa jiných velikostí?

Věta 4.33 (O velikosti konečných těles)**.** *Existují konečná tělesa právě o velikostech p^n, kde p je prvočíslo a $n \geq 1$.*

Důkaz vynecháme, ale ukážeme základní myšlenku, jak sestrojit těleso o velikosti p^n. Takové těleso se značí[1)] symbolem $\mathrm{GF}(p^n)$ a jeho prvky jsou polynomy stupně nanejvýš $n-1$ s koeficienty v tělese $\mathbb{Z}_p$. Sčítání je definováno analogicky jako pro reálné polynomy. Násobí se modulo ireducibilní polynom stupně n, kde ireducibilní znamená nerozložitelný na součin dvou polynomů stupně aspoň jedna (takový polynom vždy existuje).

Další zajímavá vlastnost je, že každé konečné těleso velikosti p^n je isomorfní s $\mathrm{GF}(p^n)$, to znamená, že taková tělesa jsou v zásadě stejná až na jiné označení prvků.

[1)] GF = Galois field, tedy Galoisovo těleso.

Příklad 4.34 (Těleso GF(8))**.** Množina má za prvky polynomy stupňů nanejvýš dva s koeficienty v $\mathbb{Z}_2$

$$\mathrm{GF}(8) = \{0, 1, x, x+1, x^2, x^2+1, x^2+x, x^2+x+1\}.$$

Sčítání je definované

$$(a_2x^2 + a_1x + a_0) + (b_2x^2 + b_1x + b_0) = (a_2 + b_2)x^2 + (a_1 + b_1)x + (a_0 + b_0),$$

např. $(x+1) + (x^2+x) = x^2+1$. Uvažme ireducibilní polynom, např. x^3+x+1. Pak násobíme modulo tento polynom, např. $x^2 \cdot x = -x-1 = x+1$, nebo $x^2 \cdot (x^2+1) = -x = x$. □

Definice 4.35 (Charakteristika tělesa)**.** *Charakteristika tělesa* $\mathbb{T}$ je nejmenší n takové, že

$$\underbrace{1+1+\ldots+1}_{n} = 0.$$

Pokud takové n neexistuje pak ji definujeme jako 0.

Tvrzení 4.36. *Charakteristika tělesa je buď nula, nebo prvočíslo.*

Důkaz. Protože $0 \neq 1$, charakteristika nemůže být 1. Pokud by byla charakteristika složené číslo $n = pq$, pak

$$0 = \underbrace{1+1+\ldots+1}_{n=pq} = (\underbrace{1+\ldots+1}_{p})(\underbrace{1+\ldots+1}_{q}),$$

tedy součet p nebo q jedniček dá nulu, což je spor s minimalitou n. □

Příklad 4.37. Jestliže charakteristika tělesa $\mathbb{T}$ není 2, tak můžeme zavést něco jako průměr. Označme symbolem 2 hodnotu $1+1$ a pak pro libovolné $a, b \in \mathbb{T}$ má číslo $p = \frac{1}{2}(a+b)$ má vlastnost $a - p = p - b$, je tedy stejně „vzdálené“ od a jako od b. (Viz důkaz věty 7.4 či příklad 12.10.)

Těleso s charakteristikou 2 je $\mathbb{Z}_2$ nebo obecněji jakékoliv těleso $\mathrm{GF}(2^n)$, kde $n \in \mathbb{N}$. V těchto tělesech tedy průměr 0 a 1 nelze zadefinovat, zatímco například v tělese $\mathbb{Z}_5$ je průměr 0 a 1 číslo 3. □

Další užitečný výsledek je *Malá Fermatova věta*[2], používá se např. pro pravděpodobnostní test prvočíselnosti velkých čísel. Často se uvádí ve znění, že $a^{p-1} \equiv 1 \bmod p$, tedy, že čísla a^{p-1} a 1 mají stejný zbytek při dělení číslem p. V jazyce konečných těles větu formulujeme takto:

Věta 4.38 (Malá Fermatova věta)**.** *Buď p prvočíslo a buď $0 \neq a \in \mathbb{Z}_p$. Pak $a^{p-1} = 1$ v tělese $\mathbb{Z}_p$.*

Důkaz. Podle lemmatu 4.26 je $\{0, 1, \ldots, p-1\} = \{0a, 1a, \ldots, (p-1)a\}$. Protože $0 = 0a$, tak dostáváme $\{1, \ldots, p-1\} = \{1a, \ldots, (p-1)a\}$. Tudíž $1 \cdot 2 \cdot 3 \cdot \ldots \cdot (p-1) = (1a) \cdot (2a) \cdot (3a) \cdot \ldots \cdot (p-1)a$. Zkrácením obou stran čísly $1, 2, \ldots, p-1$ získáme požadovanou rovnost $1 = a \cdots a$. □

Příklad 4.39. Jaká je hodnota 2^{111} v tělese $\mathbb{Z}_{11}$? Podle Malé Fermatovy věty je $2^{10} = 1$, tudíž i $2^{110} = 1$. Proto $2^{111} = 2^{110+1} = 2^{110}2^1 = 2$. □

4.4 Aplikace

Konečná tělesa se používají např. v kódování a šifrování. Na závěr této kapitoly ukážeme praktické využití těles právě v kódování, viz [Tůma, 2003, kap. 11]. Jinou ukázkou použití je tzv. „Secret sharing“, viz [Tůma, 2003, kap. 4].

[2] Malá Fermatova věta byla formulována francouzským právníkem a amatérským matematikem Pierre de Fermatem r. 1640. Pro srovnání, Velká Fermatova věta z r. 1637 pak říká, že neexistují přirozená čísla x, y, z splňující rovnici $x^n + y^n = z^n$ pro $n > 2$. Tato věta zůstávala dlouho jako otevřený problém bez důkazu až ji r. 1993 dokázal britský matematik Andrew Wiles.

Příklad 4.40 (Samoopravné kódy – Hammingův kód $(7,4,3)$)**.** Uvažujme problém přenosu dat, která jsou tvořena posloupností nul a jedniček. Zatímco úlohou šifrování je transformovat data tak, aby je nikdo nepovolaný nepřečetl, úlohou kódování je zlepšit jejich přenosové vlastnosti. Tím myslíme zejména umět detekovat a opravit chyby, které při přenosu přirozeně vznikají.

Kódování vesměs funguje tak, že odesílatel rozdělí binární posloupnost na úseky o délce k. Každý úsek pak určitou metodou přetransformuje na úsek délky k', který pak odešle. Příjemce dat pak transformuje každý úsek na původní hodnoty. Podle zvolené metody je pak schopen detekovat nebo i rovnou opravit určitý počet chyb, které v úseku vznikly.

Jednoduchý příklad. Pokud kódujeme tak, že zdvojíme každý bit, tedy např. úsek $v = 010$ zakódujeme na $v' = 001100$, tak jsme schopni detekovat maximálně jednu chybu v každém úseku. Nicméně, neumíme data opravit. Pokud budeme kódovat tak, že každý bit ztrojíme, tedy např. úsek $v = 010$ zakódujeme na $v' = 000111000$, tak už jsme schopni nejen detekovat, ale i opravit jednu chybu. Pokud příjemce dostane úsek 000111010, ví, že původní úsek byl 000111000, anebo došlo aspoň ke dvěma přenosovým chybám. Tento způsob kódování je značně neefektivní, ukážeme šikovnější způsob.

Hammingův kód $(7,4,3)$ spočívá v rozdělení přenosových dat na úseky o čtyřech bitech, které zakódujeme na sedm bitů. Tento kód umí detekovat a opravit jednu přenosovou chybu. Kódování a dekódování jde elegantně reprezentovat maticovým násobením. Úsek čtyř bitů si představíme jako aritmetický vektor a nad tělesem $\mathbb{Z}_2$. Kódování probíhá vynásobením vektoru a takzvanou generující maticí $H \in \mathbb{Z}_2^{7\times 4}$,

$$\text{např.: } Ha = \begin{pmatrix} 1&1&0&1\\ 1&0&1&1\\ 1&0&0&0\\ 0&1&1&1\\ 0&1&0&0\\ 0&0&1&0\\ 0&0&0&1 \end{pmatrix} \begin{pmatrix} \mathbf{0}\\ \mathbf{1}\\ \mathbf{1}\\ \mathbf{0} \end{pmatrix} = \begin{pmatrix} 1\\ 1\\ \mathbf{0}\\ 0\\ \mathbf{1}\\ \mathbf{1}\\ \mathbf{0} \end{pmatrix} = b.$$

Příjemce obdrží úsek reprezentovaný vektorem b. Bity původních dat jsou na zvýrazněných pozicích b_3, b_5, b_6, b_7, ostatní bity b_1, b_2, b_4 jsou kontrolní. K detekci a opravě chyb používá příjemce detekční matici $D \in \mathbb{Z}_2^{3\times 7}$. Pokud $Db = 0$, nedošlo k žádné chybě v přenosu (nebo nastaly více než dvě chyby). V opačném případě nastala přenosová chyba a chybný bit je na pozici Db, bereme-li tento vektor jako binární zápis přirozeného čísla.

$$\text{např.: } Db = \begin{pmatrix} 0&0&0&1&1&1&1\\ 0&1&1&0&0&1&1\\ 1&0&1&0&1&0&1 \end{pmatrix} \begin{pmatrix} 1\\ 1\\ 0\\ 0\\ 1\\ 1\\ 0 \end{pmatrix} = \begin{pmatrix} 0\\ 0\\ 0 \end{pmatrix} \ \ldots \text{v pořádku.}$$

$$\text{např.: } Db = \begin{pmatrix} 0&0&0&1&1&1&1\\ 0&1&1&0&0&1&1\\ 1&0&1&0&1&0&1 \end{pmatrix} \begin{pmatrix} 1\\ 1\\ 0\\ 0\\ 1\\ \mathbf{0}\\ 0 \end{pmatrix} = \begin{pmatrix} 1\\ 1\\ 0 \end{pmatrix} \ \ldots \text{chyba na pozici } 110_2 = 6.$$

Jak to, že chybný bit najdeme tak snadno? Protože vektor $Db = (1,1,0)^T$ je obsažen v matici D v šestém sloupečku, stačí změnit šestý bit vektoru b a už bude platit rovnost $Db = 0$. Všimněme si, že matice D obsahuje ve sloupcích všechny nenulové vektory, tedy všechny možné výsledky součinu Db jsou pokryty. Navíc sloupce matice D vyjadřují v binárním zápisu čísla 1 až 7, proto určíme index pokaženého bitu pomocí dvojkového vyjádření vektoru Db.

Popsat obecnou konstrukci matic H a D by bylo na pokročilou přednášku šifrování. Nicméně ještě několik podrobností k matici D zmíníme v příkladu 5.69. □

Problémy

4.1. Buď G konečná grupa a H její podgrupa. Dokažte, že velikost G je dělitelná velikostí H. Při důkazu možná využijete následující mezikroky:

(a) označme $aH := \{ah;\ h \in H\}$[3],

(b) pro každé $a, b \in G$ platí buď $aH = bH$, anebo $aH \cap bH = \emptyset$,

(c) pro každé $a \in G$ platí, že velikost aH je stejná jako velikost H.

4.2. Pro permutaci $p \in S_n$ definujme matici $P \in \mathbb{R}^n$ tak, že $P_{ij} = 1$ pokud $p(i) = j$ a nula jinak. Ukažte, že tato definice je ekvivalentní definici permutační matice ze sekce 3.5.2. Dále zjistěte, jak vypadá permutační matice inverzní permutace a permutační matice složení dvou permutací.

4.3. Dokažte vlastnost z poznámky 4.20.

4.4. Spočítejte průměrný počet cyklů v n-prvkové permutaci.

4.5. Určete pravděpodobnost, že náhodně zvolená permutace $p \in S_n$ má cyklus obsahující prvek 1 dlouhý přesně k.

[3] Tato množina se nazývá levý koset.

Shrnutí ke kapitole 4. Grupy a tělesa

Grupy představují první abstraktní pojem, se kterým jsme se setkali. Grupa je jakákoli množina, na které máme zavedenou operaci splňující několik základních vlastností (asociativita, neutrální a inverzní prvek, případně komutativita). Právě tato abstraktní definice umožňuje obsáhnout velkou řadu objektů a tak rozšiřuje pole působnosti. Jako význačný příklad nekomutativní grupy jsme probírali permutace s operací skládání.

Algebraická tělesa jsou oproti grupám bohatší o další operaci. Tělesem je tedy množina se dvěma operacemi, splňujícími určité vlastnosti. Maticové operace probírané v minulých kapitolách tak směle lze rozšířit a pracovat nad libovolným tělesem, nikoliv jen nad $\mathbb{R}$; veškeré výsledky zůstanou v zásadě v platnosti. Známe nekonečná tělesa jako například $\mathbb{R}$ či $\mathbb{C}$, a konečná tělesa jako například informatikům blízké dvojprvkové těleso $\mathbb{Z}_2$. Konečná tělesa existují právě o velikosti p^n, kde p je prvočíslo.

Kapitola 5

Vektorové prostory

Vektorové prostory zobecňují dobře známý prostor aritmetických vektorů $\mathbb{R}^n$. Stejně jako u grup a těles je zadefinujeme pomocí abstraktních axiomů.

První myšlenky na zavedení vektorů pochází od Gottfrieda Wilhelma Leibnize (1646–1716). Solidní teorii však vybudoval až německý učitel, lingvista a filosof Hermann Grassmann (1809–1877) poté, co jej roku 1845 o tomto problému (a o ceně za rozvinutí Leibnizových myšlenek) informoval August Ferdinand Möbius (1790–1868). Pojmy, se kterými se seznámíme, jako např. podprostor či lineární závislost, stejně jako axiomatizaci pomocí komutativity, asociativity a distributivity, zavedl právě Grassmann. Vybudovaná teorie však byla těžká na pochopení a ve své době se nesetkala moc s pochopením. S výjimkou Williama Rowana Hamiltona (1805–1865), který též přispěl k budování teorie vektorových prostorů a považoval Grassmanna za génia, upadl Grassmann trochu v pozapomění, než ho „znovuobjevil“ italský matematik Giuseppe Peano (1858–1932). Současná verze definice vektorových prostorů pochází od německého matematika Hermanna Weyla (1885–1955).

Poznamenejme, že v některých odvětvích se vektorové prostory označují také jako *lineární prostory*.

5.1 Základní pojmy

Definice 5.1 (Vektorový prostor)**.** Buď $\mathbb{T}$ těleso s neutrálními prvky 0 pro sčítání a 1 pro násobení. *Vektorovým prostorem nad tělesem* $\mathbb{T}$ rozumíme množinu V s operacemi sčítání vektorů $+\colon V^2 \to V$, a násobení vektoru skalárem $\cdot\colon \mathbb{T} \times V \to V$ splňující pro každé $\alpha, \beta \in \mathbb{T}$ a $u, v \in V$:

(1) $(V, +)$ je Abelova grupa, neutrální prvek značíme o a inverzní k v pak $-v$,

(2) $\alpha(\beta v) = (\alpha\beta)v$ (asociativita),

(3) $1v = v$,

(4) $(\alpha + \beta)v = \alpha v + \beta v$ (distributivita),

(5) $\alpha(u + v) = \alpha u + \alpha v$ (distributivita).

Prvkům vektorového prostoru V říkáme *vektory* a budeme je značit latinkou. Vektory píšeme bez šipek, tedy v a ne $\vec{v}$. Prvkům tělesa $\mathbb{T}$ pak říkáme *skaláry*, a pro odlišení je budeme značit řeckými písmeny.

Příklad 5.2. Příklady vektorových prostorů:

- Aritmetický prostor $\mathbb{R}^n$ nad $\mathbb{R}$, či obecněji $\mathbb{T}^n$ nad $\mathbb{T}$, kde $\mathbb{T}$ je libovolné těleso; n-tice prvků z tělesa $\mathbb{T}$ sčítáme a násobíme skalárem po složkách podobně jako u $\mathbb{R}^n$. Axiomy z definice vektorového prostoru pak vyplývají z vlastností tělesa.
- Prostor $\mathbb{R}^n$ nad $\mathbb{Q}$, opět s analogicky definovanými operacemi.
- Prostor matic $\mathbb{R}^{m\times n}$ nad $\mathbb{R}$, či obecněji $\mathbb{T}^{m\times n}$ nad $\mathbb{T}$. Axiomy z definice vektorového prostoru se snadno nahlédnou z vlastností matic a těles.
- Prostor všech reálných polynomů proměnné x, který značíme $\mathcal{P}$.
- Prostor všech reálných polynomů proměnné x stupně nanejvýš n, který značíme $\mathcal{P}^n$. Operace jsou definovány známým způsobem.

– Sčítání:

$$(a_nx^n + a_{n-1}x^{n-1} + \ldots + a_1x + a_0) + (b_nx^n + b_{n-1}x^{n-1} + \ldots + b_1x + b_0) = \\ = (a_n + b_n)x^n + (a_{n-1} + b_{n-1})x^{n-1} + \ldots + (a_1 + b_1)x + (a_0 + b_0).$$

– Násobení skalárem $\alpha \in \mathbb{R}$:

$$\alpha(a_nx^n + a_{n-1}x^{n-1} + \ldots + a_1x + a_0) = \\ = (\alpha a_n)x^n + (\alpha a_{n-1})x^{n-1} + \ldots + (\alpha a_1)x + (\alpha a_0)$$

– Nulový vektor: 0.

– Opačný vektor:

$$-(a_nx^n + a_{n-1}x^{n-1} + \ldots + a_1x + a_0) = \\ = (-a_n)x^n + (-a_{n-1})x^{n-1} + \ldots + (-a_1)x + (-a_0).$$

- Prostor všech reálných funkcí $f\colon \mathbb{R} \to \mathbb{R}$, který značíme $\mathcal{F}$. Funkce $f, g\colon \mathbb{R} \to \mathbb{R}$ sčítáme tak, že sečteme příslušné funkční hodnoty, tedy $(f + g)(x) = f(x) + g(x)$. Podobně funkci $f\colon \mathbb{R} \to \mathbb{R}$ násobíme skalárem $\alpha \in \mathbb{R}$ tak, že vynásobíme všechny funkční hodnoty, tj. $(\alpha f)(x) = \alpha f(x)$.

Součet vektorů.

Vynásobení vektoru skalárem.

- Prostor všech spojitých funkcí $f\colon \mathbb{R} \to \mathbb{R}$, který značíme $\mathcal{C}$. Prostor všech spojitých funkcí $f\colon [a, b] \to \mathbb{R}$ na intervalu $[a, b]$ pak značíme $\mathcal{C}_{[a,b]}$. Operace jsou definovány analogicky jako pro $\mathcal{F}$.

Tvrzení 5.3 (Základní vlastnosti vektorů)**.** *V prostoru V nad tělesem $\mathbb{T}$ platí:*

(1) $\forall v \in V : 0v = o$,

(2) $\forall \alpha \in \mathbb{T} : \alpha o = o$,

(3) $\forall v \in V\ \forall \alpha \in \mathbb{T} : \alpha v = o$ *implikuje, že* $\alpha = 0$ *nebo* $v = o$,

(4) $\forall v \in V : (-1)v = -v$.

Důkaz. Analogicky jako u vlastností v tělese. □

5.2 Podprostory a lineární kombinace

Definice 5.4 (Podprostor)**.** Buď V vektorový prostor nad $\mathbb{T}$. Pak $U \subseteq V$ je *podprostorem* V, pokud tvoří vektorový prostor nad $\mathbb{T}$ se stejně definovanými operacemi. Značení: $U \Subset V$.

Jak ukazuje následující tvrzení, ekvivalentní definice podprostoru je, že musí obsahovat nulový vektor a být uzavřený na obě operace.

Tvrzení 5.5. *Buď U podmnožina vektorového prostoru V nad $\mathbb{T}$. Pak U je podprostorem V právě tehdy, když platí:*

(1) $o \in U$,

(2) $\forall u, v \in U : u + v \in U$,

(3) $\forall \alpha \in \mathbb{T}\ \forall u \in U : \alpha u \in U$.

Důkaz. Pokud je U podprostorem V, pak musí splňovat požadované tři vlastnosti z definice vektorového prostoru.

Předpokládejme naopak, že U splňuje zadané tři vlastnosti. Ostatní vlastnosti z definice vektorového prostoru (jako je komutativita, asociativita, distributivita) pak platí také, protože platí pro množinu V, a tudíž automaticky platí i pro každou její podmnožinu. To, že je množina U uzavřená na opačné vektory, vyplývá z uzavřenosti na násobky, neboť podle tvrzení 5.3 je $-v = (-1)v$. □

Příklad 5.6. Příklady vektorových podprostorů:

- Dva triviální podprostory prostoru V jsou: V a $\{o\}$.
- Libovolná přímka v rovině procházející počátkem je podprostorem $\mathbb{R}^2$ nad $\mathbb{R}$, jiná ne.
- $\mathcal{P}^n \Subset \mathcal{P} \Subset \mathcal{C} \Subset \mathcal{F}$.
- Množina symetrických matic v $\mathbb{R}^{n\times n}$.
- $\mathbb{Q}^n$ nad $\mathbb{Q}$ je podprostorem prostoru $\mathbb{R}^n$ nad $\mathbb{Q}$, ale není podprostorem prostoru $\mathbb{R}^n$ nad $\mathbb{R}$, protože pracuje nad jiným tělesem.
- Jsou-li U, V podprostory W a platí-li $U \subseteq V$, pak $U \Subset V$.
- Pro vlastnost „býti podprostorem“ platí transitivita, čili $U \Subset V \Subset W$ implikuje $U \Subset W$. □

Nyní ukážeme, že průnik libovolného systému (i nekonečného nespočetného) podprostorů je zase podprostor. Pro sjednocení tato vlastnost obecně neplatí (najděte protipříklad).

Tvrzení 5.7 (Průnik podprostorů)**.** *Buď V vektorový prostor nad $\mathbb{T}$, a mějme V_i, $i \in I$, libovolný systém podprostorů V. Pak $\bigcap_{i\in I} V_i$ je opět podprostor V.*

Důkaz. Podle tvrzení 5.5 stačí ověřit tři vlastnosti: Protože $o \in V_i$ pro každé $i \in I$, musí být i v jejich průniku. Uzavřenost na sčítání: Buď $u, v \in \bigcap_{i\in I} V_i$, tj. pro každé $i \in I$ je $u, v \in V_i$, tedy i $u + v \in V_i$. Proto $u + v \in \bigcap_{i\in I} V_i$. Analogicky uzavřenost na násobky: Buď $\alpha \in \mathbb{T}$ a $v \in \bigcap_{i\in I} V_i$, tj. pro každé $i \in I$ je $v \in V_i$, tedy i $\alpha v \in V_i$. Proto $\alpha v \in \bigcap_{i\in I} V_i$. □

Tato vlastnost nás opravňuje k následující definici.

Definice 5.8 (Lineární obal)**.** Buď V vektorový prostor nad $\mathbb{T}$, a $W \subseteq V$. Pak *lineární obal* W, značený $\operatorname{span}(W)$, je průnik všech podprostorů V obsahujících W, to jest $\operatorname{span}(W) = \bigcap_{U:W\subseteq U\Subset V} U$.

Z jiného pohledu je lineární obal množiny W nejmenší prostor obsahující W v tom smyslu, že jakýkoli jiný prostor obsahující W je jeho nadmnožinou.

Ještě k terminologii: Nechť prostor U je lineárním obalem množiny vektorů W, tedy $U = \operatorname{span}(W)$. Pak říkáme, že W *generuje* prostor U, a prvky množiny W jsou *generátory* prostoru U. Prostor U se nazývá *konečně generovaný*, jestliže je generovaný nějakou konečnou množinou vektorů.

Příklad 5.9. Příklady lineárních obalů ve vektorovém prostoru $\mathbb{R}^2$:

- $\operatorname{span}\{(1,0)^T\}$ je přímka, konkrétně osa x_1. Je to tedy konečně generovaný podprostor $\mathbb{R}^2$.
- $\operatorname{span}\{(1,0)^T, (2,0)^T\}$ je totéž.
- $\operatorname{span}\{(1,1)^T, (1,2)^T\}$ je celá rovina $\mathbb{R}^2$.
- $\operatorname{span}\{\} = \{o\}$. □

Příklad 5.10. Uvažujme vektorový prostor $\mathbb{R}^2$ a jeho podprostor U reprezentovaný osou x_1. Tento podprostor lze vygenerovat vektorem $(1,0)^T$, nebo lze vygenerovat vektorem $(-3,0)^T$, anebo jakýmkoli jiným tvaru $(a,0)^T$, kde $a \neq 0$. Nicméně, U lze vygenerovat i množinou vektorů $\{(2,0)^T, (5,0)^T\}$. Vidíme, že tato množina není minimální – jeden vektor lze odstranit a zbylý vektor stále generuje podprostor U. Tato snaha o minimální reprezentaci a odstranění redundancí povede později k pojmu báze (sekce 5.4). □

Proč jsme označili obal jako lineární? Odpověď se nazývá vyjádření lineárního obalu pomocí lineárních kombinací.

Definice 5.11 (Lineární kombinace)**.** Buď V vektorový prostor nad $\mathbb{T}$ a $v_1, \ldots, v_n \in V$. Pak *lineární kombinací* vektorů $v_1, \ldots, v_n$ rozumíme výraz typu $\sum_{i=1}^n \alpha_i v_i = \alpha_1 v_1 + \ldots + \alpha_n v_n$, kde $\alpha_1, \ldots, \alpha_n \in \mathbb{T}$.

Poznámka 5.12. Zde je potřeba zdůraznit, že uvažujeme pouze lineární kombinace konečně mnoha vektorů. To pro naše účely plně postačuje, protože vesměs budeme pracovat s konečně generovanými vektorovými prostory. Nekonečné lineární kombinace je možné také zavést, ale potřebovali bychom silnější aparát (limity, konvergenci, ...)

Poznámka 5.13. Lineární kombinaci lze chápat dvěma způsoby. První způsob je chápat ji jako výraz $\sum_{i=1}^n \alpha_i v_i$ a druhý způsob je uvažovat její konkrétní hodnotu, tedy výsledný vektor. Budeme používat oba tyto pohledy.

Poznámka 5.14. Označení typu $v_1, \ldots, v_n$ jsme doposud používali výhradně pro jednotlivé složky aritmetického vektoru $v = (v_1, \ldots, v_n)$. Nicméně, nyní ho budeme používat spíše pro n nějakých vektorů. Význam by však měl být vždy jasný z kontextu.

Pomocí lineárních kombinací můžeme vygenerovat celý lineární obal konečné množiny vektorů.

Věta 5.15. *Buď V vektorový prostor nad $\mathbb{T}$, a mějme $v_1, \ldots, v_n \in V$. Pak*

$$\operatorname{span}\{v_1, \ldots, v_n\} = \{\textstyle\sum_{i=1}^n \alpha_i v_i;\ \alpha_1, \ldots, \alpha_n \in \mathbb{T}\}\,. \tag{5.1}$$

Důkaz. Inkluze „$\supseteq$“. Lineární obal $\operatorname{span}\{v_1, \ldots, v_n\}$ je podprostor V obsahující vektory $v_1, \ldots, v_n$, tedy musí být uzavřený na násobky a součty. Tudíž obsahuje i násobky $\alpha_i v_i$, $i = 1, \ldots, n$, a také jejich součet $\sum_{i=1}^n \alpha_i v_i$.

Inkluze „$\subseteq$“. Stačí ukázat, že množina lineárních kombinací

$$M \coloneqq \{\textstyle\sum_{i=1}^n \alpha_i v_i;\ \alpha_1, \ldots, \alpha_n \in \mathbb{T}\}$$

je vektorový podprostor V obsahující vektory $v_1, \ldots, v_n$, a proto je jednou z množin, jejichž průnikem $\operatorname{span}\{v_1, \ldots, v_n\}$ vzniklo. Pro každé i je vektor v_i v množině M obsažen, stačí vzít lineární kombinaci s $\alpha_i = 1$ a $\alpha_j = 0$, $j \neq i$. Nulový vektor rovněž obsahuje, vezměme lineární kombinaci s nulovými koeficienty. Uzavřenost na součty: vezměme libovolné dva vektory $u = \sum_{i=1}^n \beta_i v_i$, $u' = \sum_{i=1}^n \beta_i' v_i$ z množiny M. Pak $u + u' = \sum_{i=1}^n \beta_i v_i + \sum_{i=1}^n \beta_i' v_i = \sum_{i=1}^n (\beta_i + \beta_i') v_i$, což je prvek množiny. Podobně pro násobky, buď $\alpha \in \mathbb{T}$, pak $\alpha u = \alpha \sum_{i=1}^n \beta_i v v_i = \sum_{i=1}^n (\alpha \beta_i) v_i$, což opět náleží do množiny M. □

Příklad 5.16. Lineární obal jednoho vektoru v je dán množinou všech jeho lineárních kombinací, tedy jeho násobků:

Lineární obal dvou vektorů u, v (s různými směry) v prostoru $\mathbb{R}^3$ představuje rovinu:

□

Překvapivě můžeme pomocí (konečných) lineárních kombinací vygenerovat lineární obal i nekonečné množiny vektorů.

Tvrzení 5.17. *Buď V vektorový prostor nad $\mathbb{T}$ a buď $M \subseteq V$. Pak* $\operatorname{span}(M)$ *je tvořen všemi lineárními kombinacemi všech konečně mnoha vektorů z M.*

Důkaz. Analogický důkazu věty 5.15, necháváme na cvičení. □

Poznámka 5.18 (Trochu jiný pohled na soustavu rovnic $Ax = b$)**.** Výraz $Ax = \sum_j x_j A_{*j}$ je vlastně lineární kombinace sloupců matice A (srov. poznámka 3.17), takže řešit soustavu $Ax = b$ znamená hledat lineární kombinaci sloupců, která se rovná b. Řešení tedy existuje právě tehdy, když b náleží do podprostoru generovaného sloupci matice A.

Poznámka 5.19 (Trochu jiný pohled na součin matic AB)**.** Předchozí úvahu můžeme použít i pro maticové násobení. Uvažujme $A \in \mathbb{T}^{m\times p}, B \in \mathbb{T}^{p\times n}$. Zaměříme se nejprve na sloupce výsledné matice AB. Libovolný j-tý sloupec vyjádříme $(AB)_{*j} = AB_{*j} = \sum_{k=1}^{p} b_{kj} A_{*k}$, je tedy lineární kombinací sloupců matice A. Schematicky:

$$\begin{array}{c|c}
 & \begin{pmatrix} \dots & b_{1j} & \dots \\ & b_{2j} & \\ & \vdots & \\ \dots & b_{pj} & \dots \end{pmatrix} \\
\hline
\begin{pmatrix} | & | & & | \\ A_{*1} & A_{*2} & \dots & A_{*p} \\ | & | & & | \end{pmatrix} & \begin{pmatrix} & | & \\ \dots & (AB)_{*j} & \dots \\ & | & \end{pmatrix}
\end{array}$$

Podobně lze interpretovat maticové násobení jako vytváření lineárních kombinací řádků. Libovolný i-tý řádek výsledné matice AB vyjádříme jako $(AB)_{i*} = A_{i*}B = \sum_{k=1}^{p} a_{ik} B_{k*}$, a tedy představuje lineární kombinaci řádků matice B.

Poznámka 5.20 (Ještě jiný pohled na součin matic AB)**.** Součin $A \in \mathbb{T}^{m\times p}, B \in \mathbb{T}^{p\times n}$ lze vyjádřit ještě jiným způsobem jako $AB = \sum_{k=1}^{p} A_{*k} B_{k*}$. Každý člen sumy představuje vnější součin dvou vektorů, což vytvoří matici hodnosti nanejvýš 1. Tímto předpisem jsme tedy rozepsali matici na součet maximálně k matic hodnosti 1. Nahlédnout toto vyjádření součinu matic je snadné, neboť porovnáním prvků na pozici i, j máme

$$\begin{aligned}
(AB)_{ij} &= \textstyle\sum_{k=1}^{p} A_{ik} B_{kj}, \\
\left(\textstyle\sum_{k=1}^{p} A_{*k} B_{k*}\right)_{ij} &= \textstyle\sum_{k=1}^{p} (A_{*k} B_{k*})_{ij} = \sum_{k=1}^{p} A_{ik} B_{kj}.
\end{aligned}$$

Která z uvedených forem maticového součinu se používá v praktických implementacích?[1)] Na to není jednoduchá odpověď. Pokud jsou matice A, B velké, rozdělují se na menší bloky a na nejvyšší úrovni

[1)] Specifikace rozhraní pro základní maticové operace jsou dané v knihovně BLAS (Basic Linear Algebra Subprograms). Konkrétních implementací je pak celá řada a používají vnitřně různé algoritmy mj. v závislosti na cílové architektuře.

se někdy používá forma součinu z této poznámky. Pro součin jednotlivých bloků se pak používá forma z poznámky 5.19. Oproti standardní definici 3.7 maticového součinu má výhodu v tom, že lépe hospodaří s načítáním hodnot matice do různých úrovní paměti. Více podrobností viz Goto and Geijn [2008].

5.3 Lineární nezávislost

Konečně generovaný prostor typicky může být generován různými množinami vektorů. Motivací pro tuto podkapitolu je snaha najít množinu generátorů, která bude minimální co do počtu i co do inkluze (tedy žádná ostrá podmnožina už prostor negeneruje), srov. příklad 5.10. To pak povede i k pojmům jako báze, souřadnice a dimenze.

Definice 5.21 (Lineární nezávislost)**.** Buď V vektorový prostor nad $\mathbb{T}$ a mějme vektory $v_1, \dots, v_n \in V$. Pak vektory $v_1, \dots, v_n$ se nazývají *lineárně nezávislé*, pokud rovnost $\sum_{i=1}^n \alpha_i v_i = o$ nastane pouze pro $\alpha_1 = \dots = \alpha_n = 0$. V opačném případě jsou vektory *lineárně závislé*.

Tedy vektory $v_1, \dots, v_n$ jsou lineárně závislé, pokud existují $\alpha_1, \dots, \alpha_n \in \mathbb{T}$, ne všechna nulová a taková, že $\sum_{i=1}^n \alpha_i v_i = o$.

Pojem lineární nezávislosti zobecníme i na nekonečné množiny vektorů, nicméně s nekonečny bývá trochu potíž (např. co by se myslelo nekonečnou lineární kombinací?), proto se to definuje takto:

Definice 5.22 (Lineární nezávislost nekonečné množiny)**.** Buď V vektorový prostor nad $\mathbb{T}$ a buď $M \subseteq V$ nekonečná množina vektorů. Pak M je *lineárně nezávislá*, pokud každá konečná podmnožina M je lineárně nezávislá. V opačném případě je M *lineárně závislá*.

Příklad 5.23. Příklady lineárně (ne)závislých vektorů v $\mathbb{R}^2$:

- $(1,0)^T$ je lineárně nezávislý,
- $(1,0)^T$, $(2,0)^T$ jsou lineárně závislé,
- $(1,1)^T$, $(1,2)^T$ jsou lineárně nezávislé,
- $(1,0)^T$, $(0,1)^T$, $(1,1)^T$ jsou lineárně závislé,
- $(0,0)^T$ je lineárně závislý,
- prázdná množina je lineárně nezávislá. □

Příklad 5.24. Není těžké nahlédnout, že dva vektory tvoří lineárně závislý systém pokud jeden z nich je násobkem druhého. Pro více vektorů však lineární závislost není tak snadno vidět. Jak prakticky zjistit, zda dané aritmetické vektory, např. $(1,3,2)^T$, $(2,5,3)^T$, $(2,3,1)^T$, jsou lineárně závislé či nezávislé? Podle definice hledejme, kdy lineární kombinace vektorů dá nulový vektor:

$$\alpha \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix} + \beta \begin{pmatrix} 2 \\ 5 \\ 3 \end{pmatrix} + \gamma \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Toto vyjádříme ekvivalentně jako soustavu rovnic s neznámými α, β, γ, kdy první rovnice odpovídá rovnosti vektorů v první složce, a podobně pro další dvě:

$$\begin{aligned} 1\alpha + 2\beta + 2\gamma &= 0, \\ 3\alpha + 5\beta + 3\gamma &= 0, \\ 2\alpha + 3\beta + 1\gamma &= 0. \end{aligned}$$

Soustavu vyřešíme upravením matice soustavy na odstupňovaný tvar

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

Soustava má nekonečně mnoho řešení a určitě najdeme nějaké nenulové, např. $\alpha = 4$, $\beta = -3$, $\gamma = 1$. To znamená, že dané vektory jsou lineárně závislé. (Ještě jiné postupy uvedeme později v sekci 5.6.) □

Příklad 5.25. Definice lineární nezávislosti trochu připomíná definici regularity (definice 3.23). Není to náhoda, sloupce regulární matice (a potažmo i řádky) představují další příklad lineárně nezávislých vektorů. Podle definice je čtvercová matice A regulární, pokud rovnost $\sum_j A_{*j}x_j = o$ nastane pouze pro $x = o$, a toto přesně odpovídá lineární nezávislosti sloupců matice A. □

Věta 5.26. *Buď V vektorový prostor nad $\mathbb{T}$, a mějme $v_1, \dots, v_n \in V$. Pak vektory $v_1, \dots, v_n$ jsou lineárně závislé právě tehdy, když existuje $k \in \{1, \dots, n\}$ takové, že $v_k = \sum_{i \neq k} \alpha_i v_i$ pro nějaké $\alpha_1, \dots, \alpha_n \in \mathbb{T}$, to jest $v_k \in \operatorname{span}\{v_1, \dots, v_{k-1}, v_{k+1}, \dots, v_n\}$.*

Důkaz. Implikace „⇒“. Jsou-li vektory $v_1, \dots, v_n$ lineárně závislé, pak existuje jejich netriviální lineární kombinace rovna nule, tj. $\sum_{i=1}^n \beta_i v_i = o$ pro $\beta_1, \dots, \beta_n \in \mathbb{T}$ a $\beta_k \neq 0$ pro nějaké $k \in \{1, \dots, n\}$. Zde můžeme zvolit libovolné k takové, že $\beta_k \neq 0$. Vyjádříme k-tý člen $\beta_k v_k = -\sum_{i \neq k} \beta_i v_i$ a po zkrácení dostáváme požadovaný předpis $v_k = \sum_{i \neq k} (-\beta_k^{-1} \beta_i) v_i$.

Implikace „⇐“. Je-li $v_k = \sum_{i \neq k} \alpha_i v_i$, pak $v_k - \sum_{i \neq k} \alpha_i v_i = o$, což je požadovaná netriviální kombinace rovna nule, neboť koeficient u v_k je $1 \neq 0$. □

Důsledkem je ještě jiná charakterizace lineární závislosti. Ta mj. říká, že vektory jsou lineárně závislé právě tehdy, když odebráním nějakého (ale ne libovolného, viz příklad 5.28) z nich se jejich lineární obal nezmenší. Tudíž mezi nimi je nějaký nadbytečný. U lineárně nezávislého systému je tomu naopak: Odebráním libovolného z nich se jejich lineární obal ostře zmenší, není mezi nimi tedy žádný nadbytečný.

Důsledek 5.27. *Buď V vektorový prostor nad $\mathbb{T}$, a mějme $v_1, \dots, v_n \in V$. Pak vektory $v_1, \dots, v_n$ jsou lineárně závislé právě tehdy, když existuje $k \in \{1, \dots, n\}$ takové, že*

$$\operatorname{span}\{v_1, \dots, v_n\} = \operatorname{span}\{v_1, \dots, v_{k-1}, v_{k+1}, \dots, v_n\}. \tag{5.2}$$

Důkaz. Implikace „⇒“. Jsou-li vektory $v_1, \dots, v_n$ lineárně závislé, pak podle věty 5.26 existuje $k \in \{1, \dots, n\}$ takové, že $v_k = \sum_{i \neq k} \alpha_i v_i$ pro nějaké $\alpha_1, \dots, \alpha_n \in \mathbb{T}$. Inkluze $\supseteq$ v (5.2) je splněna triviálně, zaměříme se na tu opačnou. Libovolný vektor $u \in \operatorname{span}\{v_1, \dots, v_n\}$ se dá vyjádřit

$$u = \sum_{i=1}^n \beta_i v_i = \beta_k v_k + \sum_{i \neq k} \beta_i v_i = \beta_k \Big(\sum_{i \neq k} \alpha_i v_i \Big) + \sum_{i \neq k} \beta_i v_i = \sum_{i \neq k} (\beta_k \alpha_i + \beta_i) v_i.$$

Tedy $u \in \operatorname{span}\{v_1, \dots, v_{k-1}, v_{k+1}, \dots, v_n\}$ a máme dokázanou inkluzi „⊆“ v (5.2)

Implikace „⇐“. Pokud platí rovnost (5.2), tak

$$v_k \in \operatorname{span}\{v_1, \dots, v_n\} = \operatorname{span}\{v_1, \dots, v_{k-1}, v_{k+1}, \dots, v_n\}$$

a podle věty 5.26 jsou vektory $v_1, \dots, v_n$ lineárně závislé. □

Příklad 5.28. Vektory $(2,3)^T, (2,1)^T, (4,2)^T \in \mathbb{R}^2$ jsou lineárně závislé, tudíž jejich lineární obal lze vygenerovat i z vlastní podmnožiny těchto vektorů. Můžeme odstranit například druhý, anebo třetí vektor (ale ne oba zároveň) a výsledné dva vektory pořád budou generovat stejný prostor $\mathbb{R}^2$. Nicméně, první vektor odebrat nelze, zbylé dva vektory už $\mathbb{R}^2$ nevygenerují! □

5.4 Báze

Definice 5.29 (Báze)**.** Buď V vektorový prostor nad $\mathbb{T}$. Pak *bází* rozumíme jakýkoli lineárně nezávislý systém generátorů V.

V definici pod pojmem systém rozumíme uspořádanou množinu, časem uvidíme, proč je uspořádání důležité (pro souřadnice atp.). Nicméně pro jednoduchost značení budeme bázi, skládající se z konečně mnoha vektorů $v_1, \dots, v_n$, značit $\{v_1, \dots, v_n\}$.

Báze je tedy podle definice takový systém generátorů prostoru V, který je minimální ve smyslu inkluze. Každý z generátorů má svůj smysl, nemůžeme žádný vynechat, jinak bychom nevygenerovali celý prostor V.

Příklad 5.30. Příklady bází:

- V $\mathbb{R}^2$ máme bázi např. $e_1 = (1,0)^T$, $e_2 = (0,1)^T$. Jiná báze je $(7,5)^T$, $(2,3)^T$.
- V $\mathbb{R}^n$ máme např. bázi $e_1, \dots, e_n$, říká se jí *kanonická* a značí se kan.
- V $\mathcal{P}^n$ je bází např. $1, x, x^2, \dots, x^n$. Je to na první pohled nejjednodušší báze, nikoliv však jediná možná. Užitečná je i Bernsteinova báze skládající se z vektorů $\binom{n}{i} x^i (1-x)^{n-i}$, používá se pro různé aproximace, např. ve výpočetní geometrii pro aproximaci křivek procházejících nebo kontrolovaných danými body (tzv. Bézierovy křivky, používají se třeba v typografických fontech). □
- V $\mathcal{P}$ je bází např. nekonečný ale spočetný systém polynomů $1, x, x^2, \dots$
- V prostoru $\mathcal{C}_{[a,b]}$ také existuje báze, ale není jednoduché žádnou explicitně vyjádřit.

Věta 5.31. *Nechť $v_1, \dots, v_n$ je báze prostoru V. Pak pro každý vektor $u \in V$ existují jednoznačně určené koeficienty $\alpha_1, \dots, \alpha_n \in \mathbb{T}$ takové, že $u = \sum_{i=1}^n \alpha_i v_i$.*

Důkaz. Vektory $v_1, \dots, v_n$ tvoří bázi V, tedy každé $u \in V$ se dá vyjádřit jako $u = \sum_{i=1}^n \alpha_i v_i$ pro vhodné skaláry $\alpha_1, \dots, \alpha_n \in \mathbb{T}$. Jednoznačnost ukážeme sporem. Nechť existuje i jiné vyjádření $u = \sum_{i=1}^n \beta_i v_i$. Potom $\sum_{i=1}^n \alpha_i v_i - \sum_{i=1}^n \beta_i v_i = u - u = o$, neboli $\sum_{i=1}^n (\alpha_i - \beta_i) v_i = o$. Protože $v_1, \dots, v_n$ jsou lineárně nezávislé, musí $\alpha_i = \beta_i$ pro každé $i = 1, \dots, n$. To je spor s tím, že vyjádření jsou různá. □

Díky zmíněné jednoznačnosti můžeme zavést pojem souřadnice.

Definice 5.32 (Souřadnice)**.** Nechť $B = \{v_1, \dots, v_n\}$ je báze prostoru V a nechť vektor $u \in V$ má vyjádření $u = \sum_{i=1}^n \alpha_i v_i$. Pak *souřadnicemi* vektoru $u \in V$ vzhledem k bázi B rozumíme koeficienty $\alpha_1, \dots, \alpha_n$ a vektor souřadnic značíme $[u]_B := (\alpha_1, \dots, \alpha_n)^T$.

Pojem souřadnic je důležitější, než se na první pohled zdá. Umožňuje totiž reprezentovat těžko uchopitelné vektory a (konečně generované) prostory pomocí souřadnic, tedy aritmetických vektorů. Každý vektor má určité souřadnice a naopak každá n-tice skalárů dává souřadnici nějakého vektoru. Existuje tedy vzájemně jednoznačný vztah mezi vektory a souřadnicemi, který později (sekce 6.3) využijeme k tomu, abychom řadu, např. početních, problémů z prostoru V převedli do aritmetického prostoru, kde se pracuje snadněji.

Příklad 5.33. Souřadnice vektoru vzhledem k bázi v prostoru $\mathbb{R}^2$.

Souřadnice vektoru $(-2,3)^T$ vzhledem ke kanonické bázi: $[(-2,3)^T]_{\text{kan}} = (-2,3)^T$.

Souřadnice vektoru $(-2,3)^T$ vzhledem k bázi $B = \{(-3,1^T), (1,1)^T\}$: $[(-2,3^T)]_B = (\frac{5}{4}, \frac{7}{4})^T$. □

Příklad 5.34. Pro každé $v \in \mathbb{R}^n$ je $[v]_{\text{kan}} = v$. □

Příklad 5.35. Uvažujme bázi $B = \{1, x, x^2\}$ prostoru $\mathcal{P}^2$. Pak $[3x^2 - 5]_B = (-5, 0, 3)^T$. □

Příklad 5.36. Nahlédněte, že pro bázi B prostoru V nad $\mathbb{T}$, vektory $u, v \in V$ a skalár $\alpha \in \mathbb{T}$ platí

$$[u+v]_B = [u]_B + [v]_B,$$
$$[\alpha v]_B = \alpha [v]_B.$$

Později v kapitole 6 uvidíme, že díky této vlastnosti dokážeme efektivně vyjadřovat souřadnice. □

Příklad 5.37. Nahlédněte následující pozorování pro vektorový prostor V:

- Je-li $v_1, \dots, v_n \in V$ systém generátorů V, pak každý vektor $u \in V$ lze vyjádřit jako lineární kombinaci vektorů $v_1, \dots, v_n$ alespoň jedním způsobem.
- Jsou-li $v_1, \dots, v_n \in V$ lineárně nezávislé, pak každý vektor $u \in V$ lze vyjádřit jako lineární kombinaci vektorů $v_1, \dots, v_n$ nejvýše jedním způsobem.
- Je-li $v_1, \dots, v_n \in V$ báze V, pak každý vektor $u \in V$ lze vyjádřit jako lineární kombinaci vektorů $v_1, \dots, v_n$ právě jedním způsobem. □

Věta 5.38 (O existenci báze)**.** *Každý vektorový prostor má bázi.*

Důkaz. Důkaz provedeme pouze pro konečně generovaný prostor V. Pro ty ostatní je důkaz složitější a je potřeba určité poznatky z teorie množin (tzv. Zornovo lemma).

Buď $v_1, \dots, v_n$ systém generátorů V. Jsou-li lineárně nezávislé, tak už tvoří bázi. Jinak podle důsledku 5.27 existuje index k tak, že

$$\operatorname{span}\{v_1, \dots, v_n\} = \operatorname{span}\{v_1, \dots, v_{k-1}, v_{k+1}, \dots, v_n\}.$$

Tedy odstraněním v_k bude systém vektorů stále generovat V. Je-li nyní systém vektorů lineárně nezávislý, tvoří bázi. Jinak postup opakujeme dokud nenajdeme bázi. Postup je konečný, protože máme konečnou množinu generátorů, tudíž bázi najít musíme. □

Nyní směřujeme k tomu, že pro daný konečně generovaný prostor jsou všechny jeho báze stejně velké, což povede k zavedení pojmu dimenze. K tomuto účelu nejprve ukážeme pomocné tvrzení, které říká, kdy mohu v systému generátorů vektorového prostoru vyměnit nějaký z generátorů za úplně jiný vektor.

Lemma 5.39 (O výměně)**.** *Buď $y_1, \dots, y_n$ systém generátorů vektorového prostoru V a nechť vektor $x \in V$ má vyjádření $x = \sum_{i=1}^n \alpha_i y_i$. Pak pro libovolné k takové, že $\alpha_k \neq 0$, je $y_1, \dots, y_{k-1}, x, y_{k+1}, \dots, y_n$ systém generátorů prostoru V .*

Důkaz. Ze vztahu $x = \sum_{i=1}^n \alpha_i y_i$ vyjádříme y_k

$$y_k = \frac{1}{\alpha_k}\Big(x - \sum_{i \neq k} \alpha_i y_i\Big).$$

Chceme dokázat, že vektory $y_1, \dots, y_{k-1}, x, y_{k+1}, \dots, y_n$ generují prostor V. Vezměme libovolný vektor $z \in V$. Pro vhodné koeficienty β_i můžeme vektor z vyjádřit jako

$$\begin{aligned} z &= \sum_{i=1}^n \beta_i y_i = \beta_k y_k + \sum_{i \neq k} \beta_i y_i = \frac{\beta_k}{\alpha_k}\Big(x - \sum_{i \neq k} \alpha_i y_i\Big) + \sum_{i \neq k} \beta_i y_i = \\ &= \frac{\beta_k}{\alpha_k} x + \sum_{i \neq k} \Big(\beta_i - \frac{\beta_k}{\alpha_k}\alpha_i\Big) y_i. \end{aligned}$$ □

Věta 5.40 (Steinitzova věta o výměně[2])**.** *Buď V vektorový prostor, buď $x_1, \dots, x_m$ lineárně nezávislý systém ve V, a nechť $y_1, \dots, y_n$ je systém generátorů V. Pak platí*

(1) $m \leq n$,

(2) existují navzájem různé indexy $k_1, \dots, k_{n-m}$ takové, že $x_1, \dots, x_m, y_{k_1}, \dots, y_{k_{n-m}}$ tvoří systém generátorů V.

Důkaz. Důkaz provedeme matematickou indukcí podle m. Je-li $m = 0$, pak tvrzení platí triviálně. Přejděme k indukčnímu kroku. Předpokládejme, že tvrzení platí pro $m - 1$ a ukážeme, že platí i pro m.

Uvažujme vektory $x_1, \dots, x_{m-1}$. Ty jsou lineárně nezávislé, a podle indukčního předpokladu je $m-1 \leq n$ a existují navzájem různé indexy $\ell_1, \dots, \ell_{n-m+1}$ takové, že $x_1, \dots, x_{m-1}, y_{\ell_1}, \dots, y_{\ell_{n-m+1}}$ generují V.

[2] V angličtině *replacement theorem*, autorem je matematik Ernst Steinitz (1871–1928) z dříve německého, dnes polského Slezska.

Kdyby $m-1=n$, pak vektory $x_1,\dots,x_{m-1}$ jsou generátory prostoru V, a dostáváme $x_m \in V = \operatorname{span}\{x_1,\dots,x_{m-1}\}$, což je spor s lineární nezávislostí $x_1,\dots,x_m$. Tudíž jsme dokázali první tvrzení $m \le n$.

Pro důkaz druhé části uvažujme lineární kombinaci $x_m = \sum_{i=1}^{m-1} \alpha_i x_i + \sum_{j=1}^{n-m+1} \beta_j y_{\ell_j}$, což si můžeme dovolit díky tomu, že vektory v sumě generují V. Kdyby $\beta_1 = \dots = \beta_{n-m+1} = 0$, pak dostáváme spor s lineární nezávislostí $x_1,\dots,x_m$. Proto existuje k takové, že $\beta_k \neq 0$. Podle lemmatu 5.39 lze vyměnit y_{ℓ_k} za x_m a výsledné vektory $x_1,\dots,x_m$, $y_{\ell_1},\dots,y_{\ell_{k-1}}$, $y_{\ell_{k+1}},\dots,y_{\ell_{n-m+1}}$ budou opět generovat prostor V. □

Důsledek 5.41. *Všechny báze konečně generovaného vektorového prostoru V jsou stejně velké.*

Důkaz. Buďte $x_1,\dots,x_m$ a $y_1,\dots,y_n$ dvě báze prostoru V. Speciálně, $x_1,\dots,x_m$ jsou lineárně nezávislé a $y_1,\dots,y_n$ jsou generátory V, tedy $m \le n$. Analogicky naopak, $y_1,\dots,y_n$ jsou lineárně nezávislé a $x_1,\dots,x_m$ generují V, tedy $n \le m$. Dohromady dostáváme $m = n$. □

Tvrzení se dá zobecnit na nekonečně generované prostory s tím, že všechny báze mají stejnou mohutnost.

5.5 Dimenze

Každý konečně generovaný prostor má bázi (věta 5.38) a všechny báze jsou stejně velké (důsledek 5.41), což ospravedlňuje zavedení dimenze prostoru jako velikosti (libovolné) báze.

Definice 5.42 (Dimenze)**.** *Dimenze* konečně generovaného vektorového prostoru V je velikost nějaké jeho báze a dimenze nekonečně generovaného prostoru je ∞. Značíme ji $\dim V$.

Příklad 5.43. Příklady dimenzí:

- $\dim \mathbb{R}^n = n$, $\dim \mathbb{R}^{m\times n} = mn$, $\dim\{o\} = 0$, $\dim \mathcal{P}^n = n+1$,
- reálné prostory $\mathcal{P}$, $\mathcal{F}$, a prostor $\mathbb{R}$ nad $\mathbb{Q}$ nejsou konečně generované, mají dimenzi ∞ (viz problém 5.1). □

Nadále budeme uvažovat pouze konečně generované vektorové prostory.

Věta 5.44 (Vztah počtu prvků systému k dimenzi)**.** *Pro vektorový prostor V platí*

(1) Nechť $x_1,\dots,x_m$ jsou lineárně nezávislé. Pak $m \le \dim V$. Pokud $m = \dim V$, potom $x_1,\dots,x_m$ je báze.

(2) Nechť $y_1,\dots,y_n$ jsou generátory V. Pak $n \ge \dim V$. Pokud $n = \dim V$, potom $y_1,\dots,y_n$ je báze.

Důkaz. Označme $d = \dim V$ a nechť $z_1,\dots,z_d$ je báze V.

(1) Protože $x_1,\dots,x_m$ jsou lineárně nezávislé, podle Steinitzovy věty 5.40 je $m \le d$. Pokud $m = d$, pak podle stejné věty lze systém $x_1,\dots,x_m$ doplnit o $d-m=0$ vektorů na systém generátorů prostoru V. Tedy jsou to nutně i generátory a tím i báze.

(2) Protože $y_1,\dots,y_n$ jsou generátory V, podle Steinitzovy věty 5.40 je $n \ge d$. Nechť $n = d$. Jsou-li $y_1,\dots,y_n$ lineárně nezávislé, pak tvoří bázi. Pokud jsou lineárně závislé, pak lze jeden vynechat a získat systém generátorů o velikosti $n-1$ (důsledek 5.27). Podle Steinitzovy věty by pak ale platilo $d \le n-1$, což vede ke sporu. □

První část věty 5.44 mj. říká, že na bázi se dá nahlížet jako na maximální lineárně nezávislý systém. Druhá část věty pak říká, že báze je minimální systém generátorů (co do inkluze i co do počtu).

Věta 5.45 (Rozšíření lineárně nezávislého systému na bázi)**.** *Každý lineárně nezávislý systém vektorového prostoru V lze rozšířit na bázi V.*

Důkaz. Nechť $x_1,\dots,x_m$ jsou lineárně nezávislé a $z_1,\dots,z_d$ je báze prostoru V. Podle Steinitzovy věty 5.40 existují indexy $k_1,\dots,k_{d-m}$ takové, že $x_1,\dots,x_m,z_{k_1},\dots,z_{k_{d-m}}$ jsou generátory V. Jejich počet je d, tedy podle věty 5.44 je to báze V. □

Věta 5.46 (Dimenze podprostoru)**.** *Je-li $W \Subset V$, pak $\dim W \le \dim V$. Pokud navíc $\dim W = \dim V$, tak $W = V$.*

Důkaz. Definujme množinu $M := \emptyset$. Pokud $\operatorname{span}(M) = W$, jsme hotovi. V opačném případě existuje vektor $v \in W \setminus \operatorname{span}(M)$. Přidáme vektor v do množiny M a celý postup opakujeme. Protože M je lineárně nezávislá množina vektorů, podle věty 5.44 je velikost M shora omezena dimenzí prostoru V. Proces je tedy konečný. Protože $\operatorname{span}(M) = W$, množina M tvoří bázi prostoru M, a proto $\dim W \leq \dim V$.

Pokud $\dim W = \dim V$, tak množina M musí podle věty 5.44 tvořit bázi V, a proto $W = V$. □

Příklad 5.47. Najděme všechny podprostory prostoru $\mathbb{R}^2$:

- dimenze 2: to je pouze $\mathbb{R}^2$ (z věty 5.46),
- dimenze 1: ty jsou generovány jedním vektorem, tedy jsou to všechny přímky procházející počátkem,
- dimenze 0: to je pouze $\{o\}$. □

Příklad 5.48 (Struktura podprostorů)**.** K tomu, abychom ilustrovali strukturu podprostorů, nejprve uvažujme všechny podmnožiny množiny $\{1, \ldots, n\}$ a relaci „býti podmnožinou“, neboli inkluzi $\subseteq$. Některé podmnožiny jsou neporovnatelné co do inkluze a jiné zase jsou. Inkluze je tedy částečné uspořádání a můžeme ji znázornit tzv. *Hasseovým diagramem*, kde spojnice značí „sousední“ podmnožiny v inkluzi:

Podobným způsobem můžeme znázornit i strukturu podprostorů prostoru V dimenze n, protože relace „býti podprostorem“ je také částečné uspořádání. Diagram bude mít $n + 1$ hladin, přičemž na i-té hladině budou podprostory dimenze i. Ty jsou mezi sebou neporovnatelné ve smyslu inkluze či ve smyslu „býti podprostorem“, nicméně některé vektory mohou sdílet. Mezi jednotlivými hladinami pak opět vede spojnice mezi podprostory, z nichž jeden je podprostorem druhého. Následující obrázek ilustruje strukturu podprostorů prostoru $\mathbb{R}^3$; narozdíl od předchozího případu se v prostředních hladinách vyskytuje nekonečně mnoho objektů.

□

Víme, že sjednocení podprostorů obecně podprostor netvoří. Nicméně můžeme sestrojit lineární obal sjednocení, tomu se říká spojení podprostorů a má následující ekvivalentní předpis.

Definice 5.49 (Spojení podprostorů)**.** Buďte U, V podprostory vektorového prostoru W. Pak *spojení podprostorů* U, V je definováno jako $U + V := \{u + v;\ u \in U, v \in V\}$.

Věta 5.50 (Spojení podprostorů)**.** *Buďte U, V podprostory vektorového prostoru W. Pak*

$$U + V = \operatorname{span}(U \cup V).$$

Důkaz. Inkluze „$\subseteq$“: je triviální, neboť prostor $\operatorname{span}(U \cup V)$ je uzavřený na součty.

Inkluze „$\supseteq$“: Stačí ukázat, že $U + V$ obsahuje prostory U, V a že je podprostorem W. První část je zřejmá, pro druhou uvažujme $x_1, x_2 \in U + V$. Vektory se dají vyjádřit jako $x_1 = u_1 + v_1$, $u_1 \in U$, $v_1 \in V$, a $x_2 = u_2 + v_2$, $u_2 \in U$, $v_2 \in V$. Potom $x_1 + x_2 = u_1 + v_1 + u_2 + v_2 = (u_1 + u_2) + (v_1 + v_2) \in U + V$, což dokazuje uzavřenost na sčítání. Pro uzavřenost na násobky uvažujme $x = u + v \in U + V$, $u \in U$, $v \in V$ a skalár α. Pak $\alpha x = \alpha(u + v) = (\alpha u) + (\alpha v) \in U + V$. □

Příklad 5.51.

- $\mathbb{R}^2 = \operatorname{span}\{e_1\} + \operatorname{span}\{e_2\}$,
- $\mathbb{R}^3 = \operatorname{span}\{e_1\} + \operatorname{span}\{e_2\} + \operatorname{span}\{e_3\}$,
- $\mathbb{R}^3 = \operatorname{span}\{e_1, e_2\} + \operatorname{span}\{e_3\}$,
- $\mathbb{R}^2 = \operatorname{span}\{(1,2)^T\} + \operatorname{span}\{(3,4)^T\}$,
- ale i $\mathbb{R}^2 = \operatorname{span}\{(1,2)^T\} + \operatorname{span}\{(3,4)^T\} + \operatorname{span}\{(5,6)^T\}$. □

Pro dimenzi podprostorů a jejich spojení a průniku platí podobný vztah jako známý vztah pro velikost konečných množin a jejich sjednocení a průniku (princip inkluze a exkluze).

Věta 5.52 (Dimenze spojení a průniku)**.** *Buďte U, V podprostory vektorového prostoru W. Pak platí*

$$\dim(U + V) + \dim(U \cap V) = \dim U + \dim V. \tag{5.3}$$

Důkaz. $U \cap V$ je podprostor prostoru W, tedy má konečnou bázi $z_1, \dots, z_p$. Podle věty 5.45 ji můžeme rozšířit na bázi U tvaru $z_1, \dots, z_p, x_1, \dots, x_m$. Podobně ji můžeme rozšířit na bázi V tvaru $z_1, \dots, z_p, y_1, \dots, y_n$. Stačí, když ukážeme, že vektory $z_1, \dots, z_p$, $x_1, \dots, x_m$, $y_1, \dots, y_n$ dohromady tvoří bázi $U + V$, a rovnost (5.3) už bude platit. Nejprve ukážeme, že to jsou generátory, a pak, že jsou lineárně nezávislé.

„Generujícnost.“ Buď $z \in U + V$, pak $z = u + v$, kde $u \in U, v \in V$. Vektor u lze vyjádřit $u = \sum_{i=1}^{p} \alpha_i z_i + \sum_{j=1}^{m} \beta_j x_j$ a podobně $v = \sum_{i=1}^{p} \gamma_i z_i + \sum_{k=1}^{n} \delta_k y_k$. Potom $z = u + v = \sum_{i=1}^{p} (\alpha_i + \gamma_i) z_i + \sum_{j=1}^{m} \beta_j x_j + \sum_{k=1}^{n} \delta_k y_k$, tedy z je lineární kombinací našich vektorů.

„Lineární nezávislost.“ Buď $\sum_{i=1}^{p} \alpha_i z_i + \sum_{j=1}^{m} \beta_j x_j + \sum_{k=1}^{n} \gamma_k y_k = o$, chceme ukázat, že všechny koeficienty musí být nulové. Označme $z := \sum_{i=1}^{p} \alpha_i z_i + \sum_{j=1}^{m} \beta_j x_j = -\sum_{k=1}^{n} \gamma_k y_k$. Zřejmě $z \in U \cap V$, tedy lze vyjádřit $z = \sum_{i=1}^{p} \delta_i z_i$. Tím dostáváme $z = \sum_{i=1}^{p} \delta_i z_i = -\sum_{k=1}^{n} \gamma_k y_k$, neboli $\sum_{i=1}^{p} \delta_i z_i + \sum_{k=1}^{n} \gamma_k y_k = o$. Jediná lineární kombinace lineárně nezávislých vektorů, která dá nulový vektor, je triviální, proto $\delta_i = 0$ pro všechna i a $\gamma_k = 0$ pro všechna k. Dosazením do původní rovnosti dostaneme $\sum_{i=1}^{p} \alpha_i z_i + \sum_{j=1}^{m} \beta_j x_j = o$, a tudíž z lineární nezávislosti máme $\alpha_i = 0$ pro všechna i a $\beta_j = 0$ pro všechna j. □

Příklad 5.53. Uvažujme následující podprostory prostoru $\mathbb{R}^4$

$$U = \operatorname{span}\{(1,0,1,0)^T,\ (1,2,3,4)^T\},$$
$$V = \operatorname{span}\{(2,2,1,1,)^T,\ (3,2,1,0)^T\}.$$

Jaká je dimenze jejich průniku? Výsledek spočítáme snadno podle vzorečku (5.3). Protože generátory podprostoru U i podprostoru V jsou zřejmě lineárně nezávislé, tak $\dim U = \dim V = 2$. Dimenzi spojení

$U + V$ spočítáme jako dimenzi prostoru generovaného generátory U i V dohromady, tedy jako hodnost matice

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

Výpočtem určíme, že hodnost je 3, tudíž dimenze průniku podprostorů je $\dim(U \cap V) = \dim U + \dim V - \dim(U + V) = 2 + 2 - 3 = 1$. □

Poznámka 5.54 (Direktní součet podprostorů)**.** Je-li $U \cap V = \{o\}$, pak spojení podprostorů $W = U + V$ se nazývá *direktní součet* podprostorů U, V a značí se $W = U \oplus V$. Podle věty 5.52 je $\dim(U \oplus V) = \dim U + \dim V$. Podmínka $U \cap V = \{o\}$ pak navíc způsobí, že každý vektor $w \in W$ lze zapsat jediným způsobem ve tvaru $w = u + v$, kde $u \in U$ a $v \in V$. Nyní jsou např. $\mathbb{R}^2 = \operatorname{span}\{e_1\} \oplus \operatorname{span}\{e_2\}$, $\mathbb{R}^2 = \operatorname{span}\{(1,2)^T\} \oplus \operatorname{span}\{(3,4)^T\}$ nebo $\mathbb{R}^3 = \operatorname{span}\{e_1\} \oplus \operatorname{span}\{e_2\} \oplus \operatorname{span}\{e_3\}$ direktními součty, ale $\mathbb{R}^2 = \operatorname{span}\{(1,2)^T\} \oplus \operatorname{span}\{(3,4)^T\} \oplus \operatorname{span}\{(5,6)^T\}$ není.

5.6 Maticové prostory

Nyní se vrátíme zpátky k tématu teorie matic, které jsme zdánlivě opustili, a skloubíme ji s vektorovými prostory. Oba obory se vzájemně obohatí: Vektorově prostorový pohled nám umožní jednoduše odvodit další vlastnosti matic, a naopak, postupy z maticové teorie nám poskytnou nástroje na testování lineární nezávislosti, určování dimenze atp.

Definice 5.55 (Maticové prostory)**.** Buď $A \in \mathbb{T}^{m \times n}$. Pak definujeme

(1) sloupcový prostor $\mathcal{S}(A) := \operatorname{span}\{A_{*1}, \dots, A_{*n}\}$,

(2) řádkový prostor $\mathcal{R}(A) := \mathcal{S}(A^T)$,

(3) jádro $\operatorname{Ker}(A) := \{x \in \mathbb{T}^n;\ Ax = o\}$.

Sloupcový prostor je tedy prostor generovaný sloupci matice A, a je to podprostor $\mathbb{T}^m$. Podobně, řádkový prostor je prostor generovaný řádky matice A, a je to podprostor $\mathbb{T}^n$. Jádro pak je tvořeno všemi řešeními soustavy $Ax = o$ a ponecháváme na rozmyšlení, že je také jedná o podprostor $\mathbb{T}^n$.

Příklad 5.56. Uvažme matici

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

Pak její sloupcový prostor je $\mathcal{S}(A) = \mathbb{T}^2$ a její řádkový prostor je $\mathcal{S}(A) = \operatorname{span}\{(1,1,1)^T, (0,1,0)^T\}$. Jádro matice A určíme vyřešením soustavy $Ax = o$. Matice A již je v odstupňovaném tvaru, proto pomocí volné proměnné x_3 popíšeme množinu řešení jako $\{(x_3, 0, x_3)^T;\ x_3 \in \mathbb{R}\}$. Jádro má tedy tvar $\operatorname{Ker}(A) = \operatorname{span}\{(1, 0, -1)^T\}$. □

Díky větě 5.15 můžeme maticové prostory ekvivalentně charakterizovat pomocí lineárních kombinací, což vede na následující tvrzení.

Tvrzení 5.57. *Buď $A \in \mathbb{T}^{m \times n}$. Pak*

(1) $\mathcal{S}(A) = \{Ax;\ x \in \mathbb{T}^n\}$,

(2) $\mathcal{R}(A) = \{A^T y;\ y \in \mathbb{T}^m\}$.

Důkaz. Zřejmý z toho, že $Ax = \sum_{j=1}^n x_j A_{*j}$ představuje lineární kombinaci sloupců matice A. V druhé části analogicky $A^T y$ představuje lineární kombinaci řádků matice A. □

Je užitečné si uvědomit, že maticově můžeme reprezentovat libovolný podprostor V prostoru $\mathbb{T}^n$. Stačí vzít nějaké jeho generátory $v_1, \dots, v_m$ a sestavit matici $A \in \mathbb{T}^{m \times n}$, jejíž řádky tvoří právě vektory $v_1, \dots, v_m$. Pak $V = \mathcal{R}(A)$. Podobně V můžeme vyjádřit jako sloupcový prostor vhodné matice z $\mathbb{T}^{n \times m}$, a dokonce i jako jádro vhodné matice z $\mathbb{T}^{m \times n}$ (to již není zřejmé). Pokud tedy dokážeme dobře manipulovat s maticovými prostory, umožní nám to zacházet i s podprostory $\mathbb{T}^n$. Jak ukážeme později v sekci 6.3, můžeme takto pracovat s libovolnými konečně generovanými prostory.

Poznámka 5.58 (Geometrický pohled na maticové prostory)**.** Uvažujme zobrazení $x \mapsto Ax$ s maticí $A \in \mathbb{T}^{m\times n}$. Jádro matice A je tedy tvořeno všemi vektory z $\mathbb{T}^n$, které se zobrazí na nulový vektor. Sloupcový prostor matice A pak zase představuje množinu všech obrazů, neboli obraz prostoru $\mathbb{T}^n$ při tomto zobrazení. Jak později ukážeme, tyto prostory hrají klíčovou roli pro analýzu geometrické struktury tohoto zobrazení.

Podívejme se, jak se mění maticové prostory, když matici násobíme zleva nějakou jinou maticí (to vlastně dělá Gaussova eliminace).

Tvrzení 5.59 (Prostory a násobení maticí zleva)**.** *Buď $A \in \mathbb{T}^{m\times n}$, $Q \in \mathbb{T}^{p\times m}$. Pak*

(1) $\mathcal{R}(QA) \Subset \mathcal{R}(A)$,

(2) Pokud $A_{*k} = \sum_{j\neq k} \alpha_j A_{*j}$ pro nějaké $k \in \{1,\dots,n\}$ a nějaká $\alpha_j \in \mathbb{T}$, $j \neq k$, pak $(QA)_{*k} = \sum_{j\neq k} \alpha_j (QA)_{*j}$.

Důkaz.

(1) Stačí ukázat $\mathcal{R}(QA) \subseteq \mathcal{R}(A)$. Buď $x \in \mathcal{R}(QA)$, pak existuje $y \in \mathbb{T}^p$ takové, že $x = (QA)^T y = A^T Q^T y = A^T (Q^T y) \in \mathcal{R}(A)$.

(2) $(QA)_{*k} = QA_{*k} = Q(\sum_{j\neq k} \alpha_j A_{*j}) = \sum_{j\neq k} \alpha_j QA_{*j} = \sum_{j\neq k} \alpha_j (QA)_{*j}$, kde jsme využili tvrzení 3.16(3). □

Věta říká, že řádkové prostory jsou porovnatelné přímo – po pronásobení libovolnou maticí zleva dostaneme podprostor. To se snadno nahlédne i z toho, že každý řádek matice QA je vlastně lineární kombinací řádků matice A (viz poznámka 5.19), a vybranými lineárními kombinacemi lze vygenerovat pouze podprostor. Konkrétně, i-tý řádek matice QA má vyjádření $(QA)_{i*} = \sum_{j=1}^m q_{ij} A_{j*}$, schematicky

$$\begin{array}{c|c} & A \\ \hline Q & QA \end{array} \quad \text{odpovídá} \quad \begin{array}{c|c} & \begin{pmatrix} — & A_{1*} & — \\ — & A_{2*} & — \\ & \vdots & \\ — & A_{m*} & — \end{pmatrix} \\ \hline \begin{pmatrix} \cdots & & & \cdots \\ q_{i1} & q_{i2} & \cdots & q_{im} \\ \cdots & & & \cdots \end{pmatrix} & \begin{pmatrix} & \cdots & \\ — & \sum_{j=1}^m q_{ij} A_{j*} & — \\ & \cdots & \end{pmatrix} \end{array}$$

Sloupcové prostory z principu porovnávat nelze, protože jsou to podprostory různých prostorů ($\mathbb{T}^m$ a $\mathbb{T}^p$). Nicméně, jak říká bod (2) tvrzení 5.59, mezi sloupci se zachovává jakási lineárně závislostní vazba: Je-li i-tý sloupec matice A závislý na ostatních, potom i-tý sloupec matice QA je závislý na ostatních se stejnou lineární kombinací (pozor, lineární nezávislost se nemusí zachovávat). Tuto vlastnost můžeme nahlédnout i geometricky. Uvažujme lineární zobrazení $x \mapsto Qx$. Pak sloupce matice A se zobrazí na sloupce matice QA, neboť podle (3.1) je

$$QA = \begin{pmatrix} | & & | \\ QA_{*1} & \dots & QA_{*n} \\ | & & | \end{pmatrix}.$$

Tudíž stejná geometrická transformace $x \mapsto Qx$ se aplikuje na všechny sloupce matice A, a proto závislosti mezi sloupci zůstanou zachovány i pro výslednou matici QA.

Jestliže násobíme zleva regulární maticí, což je typický případ, tak můžeme odvodit silnější tvrzení.

Tvrzení 5.60 (Prostory a násobení regulární maticí zleva)**.** *Buď $Q \in \mathbb{T}^{m\times m}$ regulární a $A \in \mathbb{T}^{m\times n}$. Pak*

(1) $\mathcal{R}(QA) = \mathcal{R}(A)$,

(2) Rovnost $A_{*k} = \sum_{j\neq k} \alpha_j A_{*j}$ platí právě tehdy, když $(QA)_{*k} = \sum_{j\neq k} \alpha_j (QA)_{*j}$, kde $k \in \{1,\dots,n\}$ a $\alpha_j \in \mathbb{T}$, $j \neq k$.

Důkaz.

(1) Podle tvrzení 5.59 je $\mathcal{R}(QA) \subseteq \mathcal{R}(A)$. Aplikujeme-li tvrzení 5.59 na matici (QA) násobenou zleva Q^{-1}, tak dostaneme $\mathcal{R}(Q^{-1}QA) \subseteq \mathcal{R}(QA)$, tedy $\mathcal{R}(A) \subseteq \mathcal{R}(QA)$. Dohromady máme $\mathcal{R}(QA) = \mathcal{R}(A)$.

(2) Implikaci zleva doprava dostaneme z tvrzení 5.59. Obrácenou implikaci dostaneme z tvrzení 5.59 aplikované na matici (QA) násobenou zleva Q^{-1}. □

Důsledkem předchozí věty je, že pokud sloupce matice A jsou lineárně nezávislé, tak zůstanou i po vynásobení regulární maticí zleva.

Příklad 5.61. Jak se změní prostory $\mathcal{R}(A)$ a $\mathcal{S}(A)$ pokud matici A násobíme maticí Q zprava namísto zleva? A jak se změní jádro matice? Konkrétně, pro matice $A \in \mathbb{T}^{m\times p}$, $B \in \mathbb{T}^{p\times n}$, jaký je vztah prostorů $\mathrm{Ker}(A)$, $\mathrm{Ker}(AB)$ a $\mathrm{Ker}(B)$? □

Tvrzení 5.60 nám také usnadní dokázat nejdůležitější výsledek o maticových prostorech.

Věta 5.62 (Maticové prostory a RREF)**.** *Buď $A \in \mathbb{T}^{m\times n}$ a buď A^R její RREF tvar s pivoty na pozicích $(1, p_1), \ldots, (r, p_r)$, kde $r = \mathrm{rank}(A)$. Pak*

(1) nenulové řádky A^R, tedy vektory $A^R_{1*}, \ldots, A^R_{r*}$, tvoří bázi $\mathcal{R}(A)$,

(2) sloupce $A_{*p_1}, \ldots, A_{*p_r}$ tvoří bázi $\mathcal{S}(A)$,

(3) $\dim \mathcal{R}(A) = \dim \mathcal{S}(A) = r$.

Důkaz. Víme z věty 3.28, že $A^R = QA$ pro nějakou regulární matici Q.

(1) Podle tvrzení 5.60 je $\mathcal{R}(A) = \mathcal{R}(QA) = \mathcal{R}(A^R)$. Nenulové řádky A^R jsou lineárně nezávislé, tedy tvoří bázi $\mathcal{R}(A^R)$ i $\mathcal{R}(A)$.

(2) Nejprve ukážeme, že sloupce $A^R_{*p_1}, \ldots, A^R_{*p_r}$ tvoří bázi $\mathcal{S}(A^R)$. Tyto vektory jsou jistě lineárně nezávislé (jsou to jednotkové vektory). Generují $\mathcal{S}(A^R)$, neboť libovolný nebázický sloupec se dá vyjádřit jako lineární kombinace těch bázických:

$$A^R_{*j} = \sum_{i=1}^{m} a^R_{ij} e_i = \sum_{i=1}^{r} a^R_{ij} e_i = \sum_{i=1}^{r} a^R_{ij} A^R_{*p_i}.$$

Nyní použijeme tvrzení 5.60, která zaručí, že i $A_{*p_1}, \ldots, A_{*p_r}$ jsou lineárně nezávislé a generují ostatní sloupce, tedy tvoří bázi $\mathcal{S}(A)$.

(3) Hodnota $\dim \mathcal{R}(A)$ je velikost báze $\mathcal{R}(A)$, tedy r, a podobně $\dim \mathcal{S}(A)$ je velikost báze $\mathcal{S}(A)$, také r. Navíc $r = \mathrm{rank}(A)$. □

Třetí vlastnost věty 5.62 dává důležitý a netriviální důsledek pro hodnost matice a její transpozice, neboť

$$\mathrm{rank}(A) = \dim \mathcal{R}(A) = \dim \mathcal{S}(A) = \dim \mathcal{R}(A^T) = \mathrm{rank}(A^T).$$

Dostáváme tedy následující větu.

Věta 5.63. *Pro každou matici $A \in \mathbb{T}^{m\times n}$ platí $\mathrm{rank}(A) = \mathrm{rank}(A^T)$.*

Tuto větu jsme nezmiňovali v kapitole 3, protože k jejímu dokázání potřebujeme netriviální poznatky z vektorových prostorů. A naopak, řadu charakteristik vektorových prostorů jako je určování dimenze, hledání báze atp. můžeme testovat pomocí známých postupů teorie matic. Spojují se tady dvě teorie a společně produkují zajímavé výsledky.

Věta 5.62 rovněž nabízí ekvivalentní charakterizaci hodnosti matice jako dimenzi řádkového nebo sloupcového prostoru. Tím potvrzuje korektnost definice hodnosti z definice 2.13, alternativně k větě 3.43 o jednoznačnosti RREF tvaru matice.

Věta 5.62 dále dává návod, jak zjistit určité charakteristiky prostorů pomocí RREF tvaru matice. Stačí dát aritmetické vektory do matice, převést do RREF tvaru a z něj pak vyčíst danou informaci. Jestliže vektory nejsou z aritmetického prostoru $\mathbb{T}^n$, pak je potřeba na to jít oklikou, pomocí tzv. isomorfismu (sekce 6.3).

Příklad 5.64. Uvažujme prostor

$$V = \operatorname{span}\{(1,2,3,4,5)^T,\ (1,1,1,1,1)^T,\ (1,3,5,7,9)^T,\ (2,1,1,0,0)^T\} \Subset \mathbb{R}^5.$$

Nejprve sestavme matici, jejíž sloupce jsou rovny daným generátorům V, tedy $V = \mathcal{S}(A)$:

$$\begin{pmatrix} 1 & 1 & 1 & 2 \\ 2 & 1 & 3 & 1 \\ 3 & 1 & 5 & 1 \\ 4 & 1 & 7 & 0 \\ 5 & 1 & 9 & 0 \end{pmatrix} \overset{\text{RREF}}{\sim} \begin{pmatrix} 1 & 0 & 2 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Z RREF tvaru vidíme, že $\dim(V) = 3$ a báze V je například $(1,2,3,4,5)^T$, $(1,1,1,1,1)^T$, $(2,1,1,0,0)^T$. Třetí z generátorů je závislý na ostatních, konkrétně je roven dvojnásobku prvního minus druhý (koeficienty vidíme ve třetím sloupci matice v RREF tvaru).

Nyní dejme generující vektory do řádků matice, tedy $V = \mathcal{R}(A)$:

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 3 & 5 & 7 & 9 \\ 2 & 1 & 1 & 0 & 0 \end{pmatrix} \overset{\text{RREF}}{\sim} \begin{pmatrix} 1 & 0 & 0 & -1 & -1 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Opět z RREF tvaru vyčteme, že $\dim(V) = 3$, dostaneme ale jinou bázi: $(1,0,0,-1,-1)^T$, $(0,1,0,1,0)^T$, $(0,0,1,1,2)^T$. □

Poznámka 5.65. Uvažujme soustavu lineárních rovnic $Ax = b$. Řešitelnost soustavy vlastně znamená, že vektor pravých stran b se dá vyjádřit jako lineární kombinace sloupců matice A (srov. poznámka 5.18). Tudíž soustava je řešitelná právě tehdy, když $b \in \mathcal{S}(A)$, neboli $\mathcal{S}(A) = \mathcal{S}(A \mid b)$. Věta 5.62 pak přímo dává znění Frobeniovy věty z poznámky 2.22.

Věta 5.66 (O dimenzi jádra a hodnosti matice)**.** *Pro každou matici $A \in \mathbb{T}^{m\times n}$ platí*

$$\dim \operatorname{Ker}(A) + \operatorname{rank}(A) = n. \tag{5.4}$$

Důkaz. Buď $\dim \operatorname{Ker}(A) = k$. Nechť vektory $v_1, \dots, v_k$ tvoří bázi $\operatorname{Ker}(A)$, což mj. znamená, že $Av_1 = \dots = Av_k = o$. Rozšiřme vektory $v_1, \dots, v_k$ na bázi celého prostoru $\mathbb{T}^n$ doplněním o vektory $v_{k+1}, \dots, v_n$. Stačí ukázat, že vektory $Av_{k+1}, \dots, Av_n$ tvoří bázi $\mathcal{S}(A)$, protože pak $\operatorname{rank}(A) = \dim \mathcal{S}(A) = n - k$ a rovnost z věty je splněna.

„Generujícnost.“ Buď $y \in \mathcal{S}(A)$, pak $y = Ax$ pro nějaké $x \in \mathbb{T}^n$. Toto x lze vyjádřit $x = \sum_{i=1}^n \alpha_i v_i$. Dosazením

$$y = Ax = A\left(\sum_{i=1}^n \alpha_i v_i\right) = \sum_{i=1}^n \alpha_i A v_i = \sum_{i=k+1}^n \alpha_i (Av_i).$$

„Lineární nezávislost.“ Buď $\sum_{i=k+1}^n \alpha_i A v_i = o$. Pak platí $A\big(\sum_{i=k+1}^n \alpha_i v_i\big) = o$, čili $\sum_{i=k+1}^n \alpha_i v_i$ patří do jádra matice A. Proto $\sum_{i=k+1}^n \alpha_i v_i = \sum_{i=1}^k \beta_i v_i$ pro nějaké skaláry $\beta_1, \dots, \beta_k$. Přepsáním rovnice dostáváme $\sum_{i=k+1}^n \alpha_i v_i + \sum_{i=1}^k (-\beta_i) v_i = o$ a vzhledem k lineární nezávislosti $v_1, \dots, v_n$ je $\alpha_{k+1} = \dots = \alpha_n = \beta_1 = \dots = \beta_k = 0$. □

Poznámka 5.67 (Geometrický pohled na větu 5.66)**.** Uvažujme zobrazení $x \mapsto Ax$ s maticí $A \in \mathbb{T}^{m\times n}$. Prostor $\mathbb{T}^n$ se zobrazí na prostor $\mathcal{S}(A)$, jehož dimenze je $r = \operatorname{rank}(A)$. Tudíž zobrazení zobrazuje n-dimenzionální prostor na r-dimenzionální prostor. Právě ten deficit $n - r \geq 0$ je podle vzorečku (5.4) roven dimenzi jádra matice A. Pro regulární matici je jádro triviální ($\operatorname{Ker}(A) = \{o\}$), a proto zobrazuje $\mathbb{T}^n$ na celé $\mathbb{T}^n$. Čím je však jádro větší, tím menší je obraz prostoru $\mathbb{T}^n$. Dimenze jádra tedy popisuje míru „degenerace“ zobrazení. Nicméně i jádro samotné popisuje způsob této degenerace, protože $\operatorname{Ker}(A)$ obsahuje právě ty vektory, které se zobrazí na o.

Příklad 5.68. Uvažujme matici a její RREF tvar

$$A = \begin{pmatrix} 2 & 4 & 4 & 4 \\ -3 & -4 & 2 & 0 \\ 5 & 7 & -2 & 1 \end{pmatrix} \overset{\text{RREF}}{\sim} \begin{pmatrix} 1 & 0 & -6 & -4 \\ 0 & 1 & 4 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Tedy $\dim \operatorname{Ker}(A) = 4 - 2 = 2$. Prostor $\operatorname{Ker}(A)$ představuje všechna řešení soustavy $Ax = o$ a ta jsou tvaru

$$(6x_3 + 4x_4, -4x_3 - 3x_4, x_3, x_4)^T, \ x_3, x_4 \in \mathbb{R},$$

neboli

$$x_3(6, -4, 1, 0)^T + x_4(4, -3, 0, 1)^T, \ x_3, x_4 \in \mathbb{R}.$$

Tudíž vektory $(6, -4, 1, 0)$, $(4, -3, 0, 1)$ tvoří bázi $\operatorname{Ker}(A)$. Vektory získané tímto postupem představují vždy bázi $\operatorname{Ker}(A)$, protože jsou to generátory jádra a je jich stejný počet jako je $\dim \operatorname{Ker}(A)$, tedy $n - \operatorname{rank}(A)$ (srov. věta 5.44). □

Další vlastnosti maticových prostorů ukážeme v důsledku 8.42.

5.7 Aplikace

Příklad 5.69 (Ještě ke kódování)**.** Navažme na příklad 4.40 o Hammingově kódu $(7, 4, 3)$. Ke kódování jsme používali generující matici H rozměru 7×4 jednoduše tak, že vstupní úsek a délky 4 se zakóduje na úsek $b := Ha$ délky 7. Všechny zakódované úseky tak představují sloupcový prostor matice H. Protože H má lineárně nezávislé sloupce, jedná se o podprostor dimenze 4 v prostoru $\mathbb{Z}_2^7$.

Detekce chyb přijatého úseku b probíhá pomocí detekční matice D rozměru 3×7. Pokud $Db = o$, nenastala chyba (nebo nastaly alespoň dvě). Po detekční matici tedy chceme, aby (pouze) vektory ze sloupcového prostoru matice H zobrazovala na nulový vektor. Tudíž musí $\mathcal{S}(H) = \operatorname{Ker}(D)$. Nyní již vidíme, proč má matice D dané rozměry – aby její jádro byl čtyř-dimenzionální podprostor, musí mít podle věty 5.66 hodnost 3, a proto 3 lineárně nezávislé řádky postačují.

Příklad 5.70 (Rozpoznávání obličejů [Turk and Pentland, 1991])**.** Detekce a rozpoznávání obličejů z digitálního obrazu je moderní úloha počítačové grafiky. Je to příliš složitá úloha, abychom mohli vysvětlit všechny detaily úspěšných algoritmů, ale zkusíme objasnit jejich podstatu z hlediska vektorových prostorů.

Digitální obraz reprezentujeme jako matici $A \in \mathbb{R}^{m \times n}$, kde a_{ij} udává barvu pixelu na pozici i, j. Množinu obrázků s obličeji si můžeme s jistou mírou zjednodušení představit jako podprostor prostoru všech obrázků $\mathbb{R}^{m \times n}$. Báze tohoto podprostoru jsou tzv. *eigenfaces*, čili určité základní typy nebo rysy obličejů, ze kterých skládáme ostatní obličeje.

Pokud chceme rozhodnout, zda obrázek odpovídá obličeji, tak spočítáme, zda odpovídající vektor leží v podprostoru obličejů nebo v jejich blízkosti. Podobně postupujeme, pokud chceme rozpoznat zda daný obrázek odpovídá nějakému známému obličeji: Ve vektorovém prostoru $\mathbb{R}^{m \times n}$ zjistíme, který z vektorů odpovídajících známým tvářím je nejblíže vektoru našeho obrázku.

Několikrát jsme použili pojem „vzdálenost“ vektorů. Eukleidovskou vzdálenost čtenář patrně zná, podrobněji a obecněji však tento termín rozebíráme v kapitole 8. □

Příklad 5.71 (Lagrangeův interpolační polynom)**.** Vraťme se nyní k problému interpolace bodů polynomem. Mějme v rovině $n+1$ bodů $(x_0, y_0), (x_1, y_1), \ldots, (x_n, y_n)$ v rovině, kde $x_i \neq x_j$ pro $i \neq j$. Úkolem je najít polynom $p(x)$ procházející těmito body. V příkladu 3.53 jsme ukázali, jak najít interpolační polynom v základním tvaru $p(x) = a_n x^n + \ldots + a_1 x + a_0$ vyřešením soustavy rovnic s Vandermondovou maticí. Na tento problém můžeme nahlížet pohledem vektorových prostorů. Polynomy $1, x, x^2, \ldots, x^n$ tvoří standardní bázi vektorového prostoru $\mathcal{P}^n$, a naším cílem vlastně je najít souřadnice $a_0, a_1, \ldots, a_n$ hledaného polynomu $p(x)$ vzhledem k této bázi.

Nyní se nabízí otázka, jestli bychom nenašli polynom snadněji, kdybychom zvolili jinou bázi prostoru $\mathcal{P}^n$? Odpověď zní „ano“. Zvolíme následující bázi prostoru $\mathcal{P}^n$. Pro $i = 0, 1, \ldots, n$ definujeme polynom

$$p_i(x) = \prod_{j=0,\, j \neq i}^{n} \frac{1}{x_i - x_j}(x - x_j).$$

Tento polynom má v bodě x_i hodnotu 1 a v ostatních bodech x_j, $j \neq i$, hodnotu 0. Je snadné nahlédnout, že tyto polynomy jsou lineárně nezávislé: žádný polynom $p_i(x)$ není lineární kombinací ostatních, protože ostatní polynomy mají v bodě x_i hodnotu 0. Tudíž polynomy $p_0(x), \ldots, p_n(x)$ tvoří bázi prostoru $\mathcal{P}^n$. Interpolační polynom $p(x)$ se tak dá jednoznačně vyjádřit jako lineární kombinace těchto polynomů a souřadnice tvoří právě funkční hodnoty $y_0, \ldots, y_n$. Tím dostáváme explicitní vyjádření interpolačního polynomu v tzv. Lagrangeově tvaru

$$p(x) = \sum_{i=0}^{n} y_j p_i(x).$$

Tento výsledek dává rovněž alternativní zdůvodnění, že interpolační polynom je určen jednoznačně. □

Problémy

5.1. Ukažte, že prostor reálných polynomů $\mathcal{P}$, prostor reálných funkcí $\mathcal{F}$ a prostor $\mathbb{R}$ nad $\mathbb{Q}$ nejsou konečně generované.

5.2. Dokažte, že hodnost matice $A \in \mathbb{T}^{m \times n}$ se dá ekvivalentně definovat jako:

(a) velikost největší regulární podmatice (podmatice vznikne odstraněním určitého počtu řádků a sloupců, klidně i nulového).

(b) nejmenší z rozměrů matic B, C ze všech možných rozkladů $A = BC$.

5.3. Pro matice $A \in \mathbb{R}^{m \times n}$, $B \in \mathbb{R}^{n \times p}$ zdůvodněte následující odhady pro hodnost jejich součinu

$$\operatorname{rank}(A) + \operatorname{rank}(B) - n \leq \operatorname{rank}(AB) \leq \min\{\operatorname{rank}(A), \operatorname{rank}(B)\}.$$

Shrnutí ke kapitole 5. Vektorové prostory

Vektorové prostory představují další abstraktní pojem. Vektory v prostoru umíme sčítat a každý jednotlivě násobit skalárem (nikoli mezi sebou!). Aplikací obou operací na n vektorů získáme lineární kombinaci těchto vektorů. Množina všech lineárních kombinací zadaných vektorů vytvoří vektorový podprostor. Pokud stejný podprostor nevygeneruje žádná ostře menší podmnožina vektorů, jsou tyto vektory lineárně nezávislé, jinak jsou lineárně závislé. Alternativně, vektory jsou závislé pokud mezi nimi je aspoň jeden, který je lineární kombinací ostatních. Lineárně nezávislé generátory prostoru se nazývají báze tohoto prostoru. Každý prostor má nějakou bázi a pokud jich je více, tak mají všechny stejnou velikost (Steinitzova větě o výměně). To nás opravňuje zavést dimenzi prostoru jako počet vektorů v bázi. Báze prostoru pak představuje jakýsi souřadný systém v tomto prostoru, protože každý vektor prostoru se dá jednoznačně vyjádřit jako lineární kombinace bázických vektorů; příslušné koeficienty se nazývají souřadnice.

Prostory úzce souvisí s maticemi, a to dvojím způsobem. S každou maticí A je spjato několik vektorových prostorů: ten, generovaný sloupci, ten, generovaný řádky, a pak jádro, čili prostor řešení soustavy $Ax = o$. Tím, že jsme prozkoumali, jak elementární aj. maticové úpravy mění tyto prostory pak na druhou stranu dokážeme pomocí matic snadno řešit spoustu úloh: zjistit zda dané vektory jsou lineárně nezávislé, určit dimenzi prostoru, který generují, vybrat z nich vhodnou bázi, spočítat souřadnice vektoru v dané bázi atp.

Kapitola 6

Lineární zobrazení

S lineárními zobrazeními jsme se již letmo setkali v poznámkách 3.18, 3.40, 5.58 a 5.67 jako se zobrazeními typu $x \mapsto Ax$, kde $A \in \mathbb{T}^{m\times n}$. Nahlédli jsme, že zobrazení je bijekcí právě pro regulární matice a inverzní zobrazení má popis $y \mapsto A^{-1}y$. Dále víme, že prostor $\mathbb{T}^n$ se zobrazí na prostor $\mathcal{S}(A)$, jehož dimenze je $r = \operatorname{rank}(A)$. Rozdíl dimenzí $n - r$ vzoru a obrazu pak odpovídá dimenzi jádra matice A.

Pro lineární zobrazení $x \mapsto Ax$ zřejmě také platí

$$\begin{aligned}(x+y) &\mapsto A(x+y) = Ax + Ay,\\ (\alpha x) &\mapsto A(\alpha x) = \alpha(Ax).\end{aligned}$$

Právě tuto vlastnost použijeme jako definici lineárního zobrazení pro obecné prostory. Jinými slovy tato vlastnost říká, že obraz součtu dvou vektorů je roven součtu jejich obrazů a analogicky pro násobky. Lineární zobrazení tedy zachovává vztah mezi vektory: lineárně závislé vektory se zobrazí na lineárně závislé obrazy (ale ne naopak!); vektor, který je závislý na jiných vektorech se zobrazí na vektor závislý na jejich obrazech při stejné lineární kombinaci atp.

6.1 Lineární zobrazení mezi obecnými prostory

Definice 6.1 (Lineární zobrazení)**.** Buďte U, V vektorové prostory nad tělesem $\mathbb{T}$. Zobrazení $f\colon U \to V$ je *lineární*, pokud pro každé $x, y \in U$ a $\alpha \in \mathbb{T}$ platí:

- $f(x+y) = f(x) + f(y)$,
- $f(\alpha x) = \alpha f(x)$.

Lineární zobrazení se též nazývá *homomorfismus*[1]. Pro toho, koho by zajímala zoologie latinských názvů druhů zobrazení, poznamenejme, že prosté zobrazení je *injektivní*, zobrazení „na“ je *surjektivní*, injektivní homomorfismus je *monomorfismus*, surjektivní homomorfismus je *epimorfismus*, homomorfismus množiny do sebe sama je *endomorfismus*, surjektivní a injektivní homomorfismus je *isomorfismus*, a isomorfní endomorfismus se nazývá *automorfismus*.[2]

Příklad 6.2 (Příklady lineárních zobrazení)**.**

- $f\colon U \to V$ definované $f(x) = o$,
- identita je zobrazení $id\colon U \to U$ definované $id(x) = x$,
- $f\colon \mathbb{R}^n \to \mathbb{R}^m$ definované $f(x) = Ax$, kde $A \in \mathbb{R}^{m\times n}$ je pevná matice,
- derivace z prostoru reálných diferencovatelných funkcí do prostoru reálných funkcí $\mathcal{F}$.

[1] Homomorfismus je obecně zobrazení, které zachovává nějakou podstatnou strukturu. V teorii vektorových prostorů jsou důležité operace součet vektorů a jejich násobek, a potažmo je tedy důležitou strukturou vyjádření vektoru jako lineární kombinace jiných vektorů. Proto homomorfismus v teorii vektorových prostorů znamená zachování lineárně-závislostní vazby, což se dá elementárně vyjádřit jako zachování struktury operací součtu a násobků vektorů.

[2] V teorii kategorií mají pojmy jako monomorfismus a další ještě trochu obecnější význam.

- zobrazení $V \to \mathbb{T}^n$, které vektoru $x \in V$ přiřadí vektor souřadnic vzhledem k dané bázi B, tedy $x \mapsto [x]_B \in \mathbb{T}^n$, kde $n = \dim V < \infty$. □

Příklad 6.3 (Příklady lineárních zobrazení v rovině)**.**

Škálování: $(x, y) \mapsto (\alpha x, \alpha y)$

Projekce do osy x: $(x, y) \mapsto (x, 0)$

Překlopení dle osy y: $(x, y) \mapsto (-x, y)$

Obecný tvar lineárního zobrazení:
$(x, y) \mapsto (a_{11}x + a_{12}y, a_{21}x + a_{22}y)$ □

Tvrzení 6.4 (Vlastnosti lineárních zobrazení)**.** *Buď $f\colon U \to V$ lineární zobrazení. Pak*

(1) $f(\sum_{i=1}^n \alpha_i x_i) = \sum_{i=1}^n \alpha_i f(x_i)$ *pro každé* $\alpha_i \in \mathbb{T},\ x_i \in U,\ i = 1, \dots, n,$

(2) $f(o) = o$.

Důkaz.

(1) Z definice lineárního zobrazení máme $f(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 f(x_1) + \alpha_2 f(x_2)$ a zbytek dostaneme rozšířením matematickou indukcí pro libovolné přirozené n.

(2) $f(o) = f(0 \cdot o) = 0 \cdot f(o) = o$. □

Lineární zobrazení tedy zobrazuje lineární kombinace vzorů na lineární kombinace obrazů. To v důsledku znamená, že je-li vektor y lineárně závislý na vektorech $x_1, \dots, x_n$, pak jeho obraz $f(y)$ je lineárně závislý na obrazech $f(x_1), \dots, f(x_n)$. Speciálně, je-li $y = \sum_{i=1}^n \alpha_i x_i$, pak $f(y) = \sum_{i=1}^n \alpha_i f(x_i)$. Lineární zobrazení tudíž zachovává lineární závislost (včetně koeficientů).

Poznámka 6.5. Jedna z geometrických vlastností lineárních zobrazení je ta, že zobrazují přímku na přímku (nebo na bod). Přímka (viz str. 99) určená dvěma různými vektory v_1, v_2 je množina vektorů tvaru $\lambda v_1 + (1 - \lambda) v_2$, kde $\lambda \in \mathbb{T}$. Obrazem této množiny při lineárním zobrazení f je množina popsaná $f(\lambda v_1 + (1 - \lambda) v_2) = \lambda f(v_1) + (1 - \lambda) f(v_2)$, což je opět přímka nebo bod (je-li $f(v_1) = f(v_2)$). Pozor, opačným směrem tvrzení neplatí, ne každé zobrazení zachovávající přímky je lineární. Například posunutí je nelineární zobrazení, ale zobrazuje přímky na přímky.

Ke každému lineárnímu zobrazení se vztahují dva důležité vektorové prostory, obraz a jádro (též nazývané nulátor).

Definice 6.6 (Obraz a jádro)**.** Buď $f\colon U \to V$ lineární zobrazení. Pak definujeme

- obraz $f(U) := \{f(x);\ x \in U\}$,
- jádro $\operatorname{Ker}(f) := \{x \in U;\ f(x) = o\}$.

Zatímco obraz má přirozený význam jako obor hodnot zobrazení, jádro popisuje chování tohoto zobrazení. Jak uvidíme, triviální jádro (tj. $\operatorname{Ker}(f) = \{o\}$) značí, že zobrazení je prosté a tím pádem dimenze vzoru U i obrazu $f(U)$ jsou stejné. Naopak, čím větší je jádro, tím více zobrazení degeneruje, více vektorů se zobrazí na tu samou hodnotu, a tím menší má obraz $f(U)$ dimenzi vzhledem k dimenzi vzoru U.

Schematické znázornění obrazu a jádra.

Poznámka 6.7. Jádro matice a jádro lineárního zobrazení spolu úzce souvisí. Definujeme-li zobrazení f předpisem $f(x) = Ax$, potom $\mathrm{Ker}\,(f) = \mathrm{Ker}\,(A)$ a $f(U) = \mathcal{S}(A)$.

Příklad 6.8 (Obraz a jádro)**.** V příkladu 6.3:

- překlopení má obraz $f(\mathbb{R}^2) = \mathbb{R}^2$, a jádro $\mathrm{Ker}(f) = \{o\}$,
- projekce na osu x má obraz $f(\mathbb{R}^2) =$ osa x, a jádro $\mathrm{Ker}(f) =$ osa y. □

Snadno nahlédneme, že obraz představuje podprostor prostoru V a jádro podprostor prostoru U.

Tvrzení 6.9. *Buď $f\colon U \to V$ lineární zobrazení. Pak:*

(1) $f(U) \Subset V$,

(2) $\mathrm{Ker}(f) \Subset U$,

(3) pro každé $x_1, \dots, x_n \in U$ platí: $f(\mathrm{span}\{x_1, \dots, x_n\}) = \mathrm{span}\{f(x_1), \dots, f(x_n)\}$.

Důkaz. Pro ilustraci dokážeme jen první tvrzení, zbytek necháme za cvičení.

(1) Stačí ověřit, že $f(U)$ obsahuje nulový vektor a je uzavřený na součty a násobky vektorů. Protože $f(o) = o$, máme $o \in V$. Pokud $v_1, v_2 \in f(U)$, tak existují $u_1, u_2 \in U$ takové, že $f(u_1) = v_1$ a $f(u_2) = v_2$. Potom $f(u_1 + u_2) = f(u_1) + f(u_2) = v_1 + v_2$, tudíž i $v_1 + v_2 \in f(U)$. Konečně pokud $v \in f(U)$, tak existuje $u \in U : f(u) = v$. Pak pro libovolné $\alpha \in \mathbb{T}$ je $f(\alpha u) = \alpha f(u) = \alpha v \in f(U)$, z čehož je $\alpha v \in f(U)$. □

Bod (3) věty 6.9 zároveň dává návod jak určovat obraz podprostoru U' prostoru U: určím obrazy báze (nebo obecně generátorů U'), a ty tvoří generátory obrazu $f(U')$.

Připomeňme, že zobrazení $f\colon U \to V$ je prosté, pokud $f(x) = f(y)$ nastane jenom pro $x = y$. Následující věta charakterizuje, kdy je lineární zobrazení prosté. Za cvičení aplikujte podmínky na lineární zobrazení z příkladu 6.3.

Věta 6.10 (Prosté lineární zobrazení)**.** *Buď $f\colon U \to V$ lineární zobrazení. Pak následující jsou ekvivalentní:*

(1) f je prosté,

(2) $\mathrm{Ker}(f) = \{o\}$,

(3) obraz libovolné lineárně nezávislé množiny je lineárně nezávislá množina.

Důkaz. Dokážeme implikace (1) ⇒ (2) ⇒ (3) ⇒ (1).

- Implikace „(1) ⇒ (2)“. Protože $f(o) = o$, tak $o \in \mathrm{Ker}(f)$. Ale vzhledem k tomu, že f je prosté zobrazení, tak jádro už jiný prvek neobsahuje.
- Implikace „(2) ⇒ (3)“. Buďte $x_1, \dots, x_n \in U$ lineárně nezávislé a nechť $\sum_{i=1}^n \alpha_i f(x_i) = o$. Pak $f\big(\sum_{i=1}^n \alpha_i x_i\big) = o$, čili $\sum_{i=1}^n \alpha_i x_i$ náleží do jádra $\mathrm{Ker}(f) = \{o\}$. Tudíž musí $\sum_{i=1}^n \alpha_i x_i = o$ a z lineární nezávislosti vektorů máme $\alpha_i = 0$ pro všechna i.

- Implikace „(3) $\Rightarrow$ (1)“. Sporem předpokládejme, že existují dva různé vektory $x, y \in U$ takové, že $f(x) = f(y)$. Potom $o = f(x) - f(y) = f(x - y)$. Vektor o představuje lineárně závislou množinu vektorů, tedy $x - y$ musí být podle předpokladu (3) také lineárně závislá množina, a tudíž $x - y = o$, neboli $x = y$. To je spor. □

Speciálně, prosté lineární zobrazení $f\colon U \to V$ zobrazuje bázi prostoru U na bázi $f(U)$. Nicméně, ani prosté lineární zobrazení nemusí být „na“, o čemž svědčí kupříkladu zobrazení vnoření $\mathbb{R}^n$ do $\mathbb{R}^{n+1}$ definované předpisem $(v_1, \ldots, v_n)^T \mapsto (v_1, \ldots, v_n, 0)^T$.

U vektorových prostorů víme, že je každý (konečně generovaný) podprostor jednoznačně určený nějakou bází. Takovouto minimální reprezentaci bychom chtěli i pro lineární zobrazení. Jak uvidíme, jistá analogie platí i u lineárních zobrazení, protože každé lineární zobrazení je jednoznačně určeno tím, kam se zobrazí vektory z báze.

Příklad 6.11. Uvažujme lineární zobrazení $f\colon \mathbb{R}^2 \to V$. Pokud známe pouze obraz vektoru $x \neq o$, pak můžeme určit obrazy všech jeho násobků, tj. vektorů na přímce $\operatorname{span}\{x\}$, jednoduše ze vztahu $f(\alpha x) = \alpha f(x)$. Nedokážeme však zrekonstruovat celé zobrazení. K tomu potřebuje znát ještě obraz nějakého jiného (lineárně nezávislého) vektoru y. Potom umíme dopočítat obrazy nejen všech násobků vektorů x a y, ale i jejich součtu a všech lineárních kombinací, tedy všech vektorů prostoru $\mathbb{R}^2$ ze vztahu $f(\alpha x + \beta y) = \alpha f(x) + \beta f(y)$. Tudíž lineární zobrazení f je charakterizováno pouze obrazy dvou lineárně nezávislých vektorů, tedy báze. □

Věta 6.12 (Lineární zobrazení a jednoznačnost vzhledem k obrazům báze)**.** *Buďte U, V prostory nad $\mathbb{T}$ a $x_1, \ldots, x_n$ báze U. Pak pro libovolné vektory $y_1, \ldots, y_n \in V$ existuje právě jedno lineární zobrazení takové, že $f(x_i) = y_i$, $i = 1, \ldots, n$.*

Důkaz. „Existence“. Buď $x \in U$ libovolné. Pak $x = \sum_{i=1}^n \alpha_i x_i$ pro nějaké skaláry $\alpha_1, \ldots, \alpha_n \in \mathbb{T}$. Definujme obraz x jako $f(x) = \sum_{i=1}^n \alpha_i y_i$, protože lineární zobrazení musí splňovat

$$f(x) = f\left(\sum_{i=1}^n \alpha_i x_i\right) = \sum_{i=1}^n \alpha_i f(x_i) = \sum_{i=1}^n \alpha_i y_i.$$

To, že takto definované zobrazení je lineární, se ověří už snadno.

„Jednoznačnost.“ Mějme dvě různá lineární zobrazení f a g splňující $f(x_i) = g(x_i) = y_i$ pro všechna $i = 1, \ldots, n$. Pak pro libovolné $x \in U$ je

$$f(x) = f\left(\sum_{i=1}^n \alpha_i x_i\right) = \sum_{i=1}^n \alpha_i f(x_i) = \sum_{i=1}^n \alpha_i y_i = \sum_{i=1}^n \alpha_i g(x_i) = g\left(\sum_{i=1}^n \alpha_i x_i\right) = g(x).$$

Tedy $f(x) = g(x)$ $\forall x \in U$, což je ve sporu s tím, že jsou to různá zobrazení. □

6.2 Maticová reprezentace lineárního zobrazení

Každé lineární zobrazení jde reprezentovat maticově. Protože vektory mohou být všelijaké podivné objekty, je nutno je popisovat v řeči souřadnic. Potom s nimi můžeme operovat jako s aritmetickými vektory, což je často pohodlnější. Než přejdeme k vlastní definici, ukážeme několik motivačních příkladů.

Příklad 6.13 (Úvod k matici lineárního zobrazení)**.** Uvažujme lineární zobrazení $f\colon \mathbb{T}^n \to \mathbb{T}^m$. Potom pro libovolné $x \in \mathbb{T}^n$ platí

$$f(x) = f\left(\sum_{i=1}^n x_i e_i\right) = \sum_{i=1}^n x_i f(e_i).$$

Označíme-li matici se sloupci $f(e_1), \ldots, f(e_n)$ jako

$$A = \begin{pmatrix} | & & | \\ f(e_1) & \ldots & f(e_n) \\ | & & | \end{pmatrix},$$

pak zřejmě $f(x) = Ax$. Každé lineární zobrazení $f\colon \mathbb{T}^n \to \mathbb{T}^m$ lze tedy reprezentovat maticově jako $f(x) = Ax$. Toto pozorování nahlédneme ještě jednou jiným způsobem v důsledku 6.17.

Uvažujme nyní lineární zobrazení $f\colon U \to \mathbb{T}^m$ a bázi $B = \{v_1, \dots, v_n\}$ prostoru U. Nechť vektor $x \in U$ má vyjádření $x = \sum_{i=1}^n \alpha_i v_i$, tedy $[x]_B = (\alpha_1, \dots, \alpha_n)^T$. Potom

$$f(x) = f\left(\sum_{i=1}^n \alpha_i v_i\right) = \sum_{i=1}^n \alpha_i f(v_i).$$

Označíme-li matici se sloupci $f(v_1), \dots, f(v_n)$ jako

$$A = \begin{pmatrix} | & & | \\ f(v_1) & \dots & f(v_n) \\ | & & | \end{pmatrix},$$

pak zřejmě $f(x) = A \cdot [x]_B$. Narozdíl od předchozího případu násobíme matici vektorem souřadnic $[x]_B$ vektoru x, a ne vektorem samotným. Pokud změníme i druhý prostor $\mathbb{T}^m$, budeme muset pracovat v souřadnicích i z hlediska matice A a obrazu $f(x)$. □

Definice 6.14 (Matice lineárního zobrazení)**.** Buď $f\colon U \to V$ lineární zobrazení, $B_1 = \{x_1, \dots, x_n\}$ báze prostoru U nad $\mathbb{T}$, a $B_2 = \{y_1, \dots, y_m\}$ báze prostoru V nad $\mathbb{T}$. Nechť $f(x_j) = \sum_{i=1}^m a_{ij} y_i$. Potom matice $A \in \mathbb{T}^{m \times n}$ s prvky a_{ij}, $i = 1, \dots, m$, $j = 1, \dots, n$, se nazývá *matice lineárního zobrazení f vzhledem k bázím B_1, B_2* a značí se ${}_{B_2}[f]_{B_1}$.

Jinými slovy, matice lineárního zobrazení vypadá tak, že její j-tý sloupec je tvořen souřadnicemi obrazu vektoru x_j vzhledem k bázi B_2, to jest $A_{*j} = [f(x_j)]_{B_2}$.

Příklad 6.15 (Matice lineárního zobrazení)**.** Uvažujme lineární zobrazení $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ s předpisem $f(x) = Ax$, kde

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

Zvolme báze $B_1 = \{(1,2)^T, (2,1)^T\}$, $B_2 = \{(1,-1)^T, (0,1)^T\}$ a najděme matici zobrazení f vzhledem k bázím B_1, B_2.

Obraz prvního vektoru báze B_1 je $f(1,2) = (5,-5)^T$, a jeho souřadnice vzhledem k bázi B_2 jsou $[f(1,2)]_{B_2} = (5,0)^T$. Podobně, obraz druhého vektoru báze B_1 je $f(2,1) = (4,2)^T$, a jeho souřadnice vzhledem k bázi B_2 jsou $[f(2,1)]_{B_2} = (4,6)^T$. Tudíž

$${}_{B_2}[f]_{B_1} = \begin{pmatrix} 5 & 4 \\ 0 & 6 \end{pmatrix}.$$ □

K čemu je matice lineárního zobrazení užitečná říká následující věta.

Věta 6.16 (Maticová reprezentace lineárního zobrazení)**.** *Buď $f\colon U \to V$ lineární zobrazení, $B_1 = \{x_1, \dots, x_n\}$ báze prostoru U, a $B_2 = \{y_1, \dots, y_m\}$ báze prostoru V. Pak pro každé $x \in U$ je*

$$[f(x)]_{B_2} = {}_{B_2}[f]_{B_1} \cdot [x]_{B_1}. \tag{6.1}$$

Důkaz. Označme $A := {}_{B_2}[f]_{B_1}$. Buď $x \in U$, tedy $x = \sum_{i=1}^n \alpha_i x_i$, neboli $[x]_{B_1} = (\alpha_1, \dots, \alpha_n)^T$. Pak

$$\begin{aligned} f(x) = f\left(\sum_{j=1}^n \alpha_j x_j\right) = \sum_{j=1}^n \alpha_j f(x_j) = \sum_{j=1}^n \alpha_j \left(\sum_{i=1}^m a_{ij} y_i\right) = \\ = \sum_{j=1}^n \sum_{i=1}^m \alpha_j a_{ij} y_i = \sum_{i=1}^m \left(\sum_{j=1}^n \alpha_j a_{ij}\right) y_i. \end{aligned}$$

Tedy výraz $\sum_{j=1}^n \alpha_j a_{ij}$ reprezentuje i-tou souřadnici vektoru $[f(x)]_{B_2}$, ale jeho hodnota je $\sum_{j=1}^n \alpha_j a_{ij} = (A[x]_{B_1})_i$, což je i-tá složka vektoru ${}_{B_2}[f]_{B_1}\, [x]_{B_1}$. □

Matice lineárního zobrazení tedy převádí souřadnice vektoru vzhledem k dané bázi na souřadnice jeho obrazu. Plně tak popisuje lineární zobrazení a navíc obraz libovolného vektoru můžeme vyjádřit jednoduchým způsobem jako násobení maticí.

Připomeňme, že symbol kan značí kanonickou bázi, tj. tu skládající se z jednotkových vektorů.

Důsledek 6.17. *Každé lineární zobrazení $f\colon \mathbb{R}^n \to \mathbb{R}^m$ se dá vyjádřit jako $f(x) = Ax$ pro nějakou matici $A \in \mathbb{R}^{m\times n}$.*

Důkaz.

$$f(x) = [f(x)]_{\text{kan}} = {}_{\text{kan}}[f]_{\text{kan}} \cdot [x]_{\text{kan}} = {}_{\text{kan}}[f]_{\text{kan}} \cdot x$$

Tedy $f(x) = Ax$, kde $A = {}_{\text{kan}}[f]_{\text{kan}}$. □

Mějme lineární zobrazení $f\colon U \to V$ a báze B_1, B_2 prostorů. Víme, že matice $A = {}_{B_2}[f]_{B_1}$ splňuje

$$[f(x)]_{B_2} = A \cdot [x]_{B_1} \quad \forall x \in U.$$

Ukážeme, že žádná jiná matice tuto vlastnost nemá.

Věta 6.18 (Jednoznačnost matice lineárního zobrazení)**.** *Buď $f\colon U \to V$ lineární zobrazení, B_1 báze prostoru U, a B_2 báze prostoru V. Pak jediná matice A splňující* (6.1) *je $A = {}_{B_2}[f]_{B_1}$.*

Důkaz. Nechť báze B_1 sestává z vektorů $x_1, \dots, x_n$. Pro spor předpokládejme, že lineární zobrazení f má dvě maticové reprezentace (6.1) pomocí matic $A \neq A'$. Tudíž existuje vektor $s \in \mathbb{T}^n$ takový, že $As \neq A's$; takový vektor lze volit např. jako jednotkový s jedničkou na takové pozici, ve kterém sloupci se matice A, A' liší. Definujme vektor $x := \sum_{i=1}^n s_i x_i$. Pak $[f(x)]_{B_2} = As \neq A's = [f(x)]_{B_2}$, což je spor s jednoznačností souřadnic (věta 5.31). □

Poznámka 6.19. Nejen že každé lineární zobrazení jde reprezentovat maticově, ale i naopak každá matice představuje matici nějakého lineárního zobrazení. Buďte B_1, B_2 báze prostorů U, V dimenzí n, m a mějme $A \in \mathbb{T}^{m\times n}$. Pak existuje jediné lineární zobrazení $f\colon U \to V$ takové, že $A = {}_{B_2}[f]_{B_1}$; ve sloupcích matice A vyčteme souřadnice obrazů vektorů báze B_1, což plně určuje zobrazení f dle věty 6.12.

Speciálním případem zobrazení je identita, jeho matici pak nazýváme *maticí přechodu*, protože umožňuje přecházet od jednoho souřadného systému k jinému.

Definice 6.20 (Matice přechodu)**.** Buď V vektorový prostor a B_1, B_2 dvě jeho báze. Pak *maticí přechodu* od B_1 k B_2 nazveme matici ${}_{B_2}[id]_{B_1}$.

Matice přechodu má pak podle maticové reprezentace tento význam; buď $x \in U$, pak

$$[x]_{B_2} = {}_{B_2}[id]_{B_1} \cdot [x]_{B_1},$$

tedy prostým maticovým násobením získáváme souřadnice vzhledem k jiné bázi.

Příklad 6.21. Najděte matici přechodu v $\mathbb{R}^3$ od báze

$$B_1 = \{(1,1,-1)^T, (3,-2,0)^T, (2,-1,1)^T\}$$

k bázi

$$B_2 = \{(8,-4,1)^T, (-8,5,-2)^T, (3,-2,1)^T\}.$$

Řešení: spočítáme

$$\begin{aligned}
[(1,1,-1)^T]_{B_2} &= (2,3,3)^T,\\
[(3,-2,0)^T]_{B_2} &= (-1,-4,-7)^T,\\
[(2,-1,1)^T]_{B_2} &= (1,3,6)^T.
\end{aligned}$$

Tedy

$$ {}_{B_2}[id]_{B_1} = \begin{pmatrix} 2 & -1 & 1 \\ 3 & -4 & 3 \\ 3 & -7 & 6 \end{pmatrix}. $$

Víme-li např., že souřadnice vektoru $(4,-1,-1)^T$ vzhledem k bázi B_1 jsou $(1,1,0)^T$, pak souřadnice vzhledem k B_2 získáme

$$ [(4,-1,-1)^T]_{B_2} = {}_{B_2}[id]_{B_1} \cdot [(4,-1,-1)^T]_{B_1} = {}_{B_2}[id]_{B_1} \cdot (1,1,0)^T = (1,-1,-4)^T. \qquad \square $$

Příklad 6.22. Buď B báze prostoru $\mathbb{T}^n$. Podle maticové reprezentace lineárního zobrazení pak speciálně dostaneme pak

$$ [x]_B = {}_B[id]_{\text{kan}} \cdot [x]_{\text{kan}} = {}_B[id]_{\text{kan}} \cdot x. $$

Souřadnice libovolného vektoru tudíž získáme jednoduše vynásobením matice přechodu s vektorem x. $\square$

Důležitou roli v teorii lineárních zobrazení hraje jejich vzájemné skládání. Připomeňme, že pro zobrazení $f\colon U \to V$ a $g\colon V \to W$ je složené zobrazení $g \circ f$ je definované předpisem $(g \circ f)(x) := g(f(x))$, $x \in U$.

Tvrzení 6.23 (Složené lineární zobrazení)**.** *Buďte $f\colon U \to V$, $g\colon V \to W$ lineární zobrazení. Pak složené zobrazení $g \circ f$ je zase lineární zobrazení.*

Důkaz. Podle definice ověříme pro libovolné $x, y \in U$ a $\alpha \in \mathbb{T}$:

$$ \begin{aligned} (g \circ f)(x+y) &= g(f(x+y)) = g(f(x)+f(y)) = \\ &= g(f(x)) + g(f(y)) = (g \circ f)(x) + (g \circ f)(y), \\ (g \circ f)(\alpha x) &= g(f(\alpha x)) = g(\alpha f(x)) = \alpha g(f(x)) = \alpha (g \circ f)(x). \end{aligned} \qquad \square $$

Věta 6.24 (Matice složeného lineárního zobrazení)**.** *Buďte $f\colon U \to V$ a $g\colon V \to W$ lineární zobrazení, buď B_1 báze U, B_2 báze V a B_3 báze W. Pak*

$$ {}_{B_3}[g \circ f]_{B_1} = {}_{B_3}[g]_{B_2} \cdot {}_{B_2}[f]_{B_1}. $$

Důkaz. Pro každé $x \in U$ je

$$ [(g \circ f)(x)]_{B_3} = [g(f(x))]_{B_3} = {}_{B_3}[g]_{B_2} \cdot [f(x)]_{B_2} = {}_{B_3}[g]_{B_2} \cdot {}_{B_2}[f]_{B_1} \cdot [x]_{B_1}. $$

Díky jednoznačnosti matice lineárního zobrazení (věta 6.18) je ${}_{B_3}[g]_{B_2} \cdot {}_{B_2}[f]_{B_1}$ hledaná matice složeného zobrazení. $\square$

Vidíme, že skládání zobrazení se v maticové reprezentaci projeví jako součin příslušných matic. To není náhoda, neboť zakladatelé teorie matic, jako např. A. Cayley, definovali (kolem roku 1855) násobení matic právě tak, aby mělo požadované vlastnosti pro skládání zobrazení. Takže i když význam násobení matic daleko přesáhl původní myšlenky, jeho kořeny je třeba hledat zde.

Příklad 6.25 (Skládání otočení a součtové vzorce pro sin a cos)**.** Otočení v rovině o úhel α proti směru hodinových ručiček má vzhledem ke kanonické bázi matici

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

Podobně otočení o úhel β. Matici otočení o úhel $\alpha+\beta$ můžeme získat přímo nebo složením otočení o α a pak o β. Porovnáním získáme součtové vzorce pro sin a cos:

$$\begin{pmatrix} \cos(\alpha+\beta) & -\sin(\alpha+\beta) \\ \sin(\alpha+\beta) & \cos(\alpha+\beta) \end{pmatrix} = \begin{pmatrix} \cos\beta & -\sin\beta \\ \sin\beta & \cos\beta \end{pmatrix} \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} =$$
$$= \begin{pmatrix} \cos\alpha\cos\beta - \sin\alpha\sin\beta & -\sin\alpha\cos\beta - \sin\beta\cos\alpha \\ \cos\alpha\sin\beta + \cos\beta\sin\alpha & -\sin\alpha\sin\beta + \cos\alpha\cos\beta \end{pmatrix}. \qquad \square$$

Příklad 6.26 (Otočení v $\mathbb{R}^n$)**.** Matici otočení z předchozího příkladu jednoduše zobecníme na matici otočení o úhel α v rovině os x_i, x_j. Schematicky (prázdné místo odpovídá nulám):

$$\begin{pmatrix} I & & & & \\ & \cos\alpha & & -\sin\alpha & \\ & & I & & \\ & \sin\alpha & & \cos\alpha & \\ & & & & I \end{pmatrix}$$

S touto maticí se setkáme ještě později v příkladu 8.53. $\square$

Příklad 6.27. Nechť máme dánu matici lineárního zobrazení f vzhledem k bázím B_1, B_2, tj. ${}_{B_2}[f]_{B_1}$. Jak určit matici vzhledem k bázím B_3, B_4, tj. ${}_{B_4}[f]_{B_3}$? Podle věty o matici složeného zobrazení aplikované na $f = id \circ f \circ id$ a příslušné báze máme

$${}_{B_4}[f]_{B_3} = {}_{B_4}[id]_{B_2} \cdot {}_{B_2}[f]_{B_1} \cdot {}_{B_1}[id]_{B_3}.$$

Tedy veškerou práci vykonají matice přechodu mezi bázemi. $\square$

Příklad 6.28 (Posunutí jako lineární zobrazení?)**.** Buď $v \in \mathbb{R}^n$ pevné a uvažujme zobrazení $f\colon \mathbb{R}^n \to \mathbb{R}^n$ dané předpisem $f(x) = x + v$. Toto zobrazení není lineární! Nicméně můžeme ho jako lineární simulovat. Vnoříme se do prostoru o jednu dimenzi většího tak, aby pro určité lineární zobrazení $g\colon \mathbb{R}^{n+1} \to \mathbb{R}^{n+1}$ platilo

$$g(x_1, \dots, x_n, 1) = (x_1 + v_1, \dots, x_n + v_n, 1).$$

Dodefinujeme g pro ostatní body tak, aby tvořilo lineární zobrazení

$$g(x_1, \dots, x_n, x_{n+1}) = (x_1 + v_1 x_{n+1}, \dots, x_n + v_n x_{n+1}, x_{n+1}).$$

Ilustrace vnoření:

Matice zobrazení:

$$\begin{pmatrix} 1 & 0 & \dots & 0 & v_1 \\ 0 & 1 & \dots & 0 & v_2 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & \dots & 1 & v_n \\ 0 & 0 & \dots & 0 & 1 \end{pmatrix}$$

$\square$

Věta o matici složeného zobrazení má ještě řadu pěkných důsledků týkajících se isomorfismů.

6.3 Isomorfismus

Definice 6.29 (Isomorfismus)**.** *Isomorfismus* mezi prostory U, V je vzájemně jednoznačné lineární zobrazení $f\colon U \to V$. Pokud mezi prostory U, V existuje isomorfismus, pak říkáme, že U, V jsou *isomorfní.*

Příklad 6.30. Příkladem isomorfismu je třeba škálování, překlápění v $\mathbb{R}^2$ (příklad 6.3) nebo otáčení (příklad 6.25). Příkladem lineárního zobrazení, které není isomorfismem, je projekce (příklad 6.3).

Přikladem isomorfních prostorů je například $\mathcal{P}^n$ a $\mathbb{R}^{n+1}$, kdy vhodným (a nikoliv jediným) isomorfismem je

$$a_n x^n + \ldots + a_1 x + a_0 \mapsto (a_n, \ldots, a_1, a_0),$$

Jiným příkladem isomorfních prostorů je $\mathbb{R}^{m\times n}$ a $\mathbb{R}^{mn}$, kdy vhodným isomorfismem je například

$$A \mapsto (a_{11}, \ldots, a_{1n}, a_{21}, \ldots, a_{2n}, \ldots, a_{m1}, \ldots, a_{mn}).$$

Isomorfismus najdeme i mezi některými nekonečně generovanými prostory, například mezi prostorem polynomů $\mathcal{P}$ a prostorem reálných posloupností s konečně mnoha nenulovými prvky. Isomorfismem je pak třeba zobrazení $a_n x^n + \ldots + a_1 x + a_0 \mapsto (a_0, a_1, \ldots, a_n, 0, \ldots)$. □

Tvrzení 6.31 (Vlastnosti isomorfismu)**.**

(1) Je-li $f\colon U \to V$ isomorfismus, pak $f^{-1}\colon V \to U$ existuje a je to také isomorfismus.

(2) Jsou-li $f\colon U \to V$ a $g\colon V \to W$ isomorfismy, pak $g \circ f\colon U \to W$ je také isomorfismus.

(3) Je-li $f\colon U \to V$ isomorfismus, pak libovolná báze prostoru U se zobrazuje na bázi prostoru V.

(4) Je-li $f\colon U \to V$ isomorfismus, pak $\dim U = \dim V$.

Důkaz.

(1) Zobrazení f je vzájemně jednoznačné, tedy f^{-1} existuje a je také vzájemně jednoznačné. Zbývá dokázat linearitu. Buď $v_1, v_2 \in V$ a nechť $f^{-1}(v_1) = u_1$ a $f^{-1}(v_2) = u_2$. Pak $f(u_1 + u_2) = f(u_1) + f(u_2) = v_1 + v_2$, tedy $f^{-1}(v_1 + v_2) = u_1 + u_2 = f^{-1}(v_1) + f^{-1}(v_2)$. Podobně pro násobky: Nechť $v \in V$ a $f^{-1}(v) = u$, pak $f(\alpha u) = \alpha f(u) = \alpha v$, tedy $f^{-1}(\alpha v) = \alpha u = \alpha f^{-1}(v)$.

(2) Snadné z tvrzení 6.23.

(3) Buď $x_1, \ldots, x_n$ báze U. Protože f je prosté, dle věty 6.10(3) jsou $f(x_1), \ldots, f(x_n)$ lineárně nezávislé. Protože f je na, generují vektory $f(x_1), \ldots, f(x_n)$ dle věty 6.9(3) prostor $f(U) = V$. Tedy vektory $f(x_1), \ldots, f(x_n)$ tvoří bázi V.

(4) Plyne z předchozího bodu. □

Nyní přichází na řadu slíbené důsledky věty o matici složeného lineárního zobrazení.

Tvrzení 6.32. *Buď $f\colon U \to V$ isomorfismus, B_1 báze U a B_2 báze V. Pak*

$${}_{B_1}[f^{-1}]_{B_2} = {}_{B_2}[f]_{B_1}^{-1}.$$

Důkaz. Protože $f^{-1} \circ f = id$, dostáváme

$${}_{B_1}[f^{-1}]_{B_2} \cdot {}_{B_2}[f]_{B_1} = {}_{B_1}[f^{-1} \circ f]_{B_1} = {}_{B_1}[id]_{B_1} = I.$$

Jelikož ${}_{B_2}[f]_{B_1}$ je podle věty 6.31(4) čtvercová, je ${}_{B_1}[f^{-1}]_{B_2}$ její inverzní matice. □

Matice isomorfismu má matici inverzní, tedy musí být regulární. Toto tvrzení platí i naopak: Je-li matice lineárního zobrazení f regulární, pak je f isomorfismem, protože inverzní matice dává předpis pro inverzní zobrazení f^{-1}.

Další důsledek tvrzení 6.32 dostaneme speciálně pro matici přechodu mezi bázemi B_1 a B_2, a to

$${}_{B_1}[id]_{B_2} = {}_{B_2}[id]_{B_1}^{-1}.$$

Příklad 6.33 (Mnemotechnika počítání matic přechodu)**.** Pro počítání matice přechodu v $\mathbb{R}^n$ od báze B_1 do báze B_2, tj. ${}_{B_2}[id]_{B_1}$, lze použít následující mnemotechniku:

$$\big(B_2 \mid B_1\big) \overset{\text{RREF}}{\sim} \big(I_n \mid {}_{B_2}[id]_{B_1}\big).$$

První matice ve sloupcích obsahuje bázi B_2 a pak bázi B_1, což jou vlastně matice ${}_{\text{kan}}[id]_{B_2}$ a ${}_{\text{kan}}[id]_{B_1}$. Převedením na RREF tvar dostaneme napravo hledanou matici přechodu. Důvod pramení ze vztahu ${}_{B_2}[id]_{B_1} = {}_{B_2}[id]_{\text{kan}} \cdot {}_{\text{kan}}[id]_{B_1} = {}_{\text{kan}}[id]_{B_2}^{-1} \cdot {}_{\text{kan}}[id]_{B_1}$.

Konkrétně, pro příklad 6.21, dostaneme

$$\left(\begin{array}{rrr|rrr} 8 & -8 & 3 & 1 & 3 & 2 \\ -4 & 5 & -2 & 1 & -2 & -1 \\ 1 & -2 & 1 & -1 & 0 & 1 \end{array}\right) \overset{\text{RREF}}{\sim} \left(\begin{array}{rrr|rrr} 1 & 0 & 0 & 2 & -1 & 1 \\ 0 & 1 & 0 & 3 & -4 & 3 \\ 0 & 0 & 1 & 3 & -7 & 6 \end{array}\right). \qquad \square$$

Nyní se obraťme od matic zpět k prostorům mezi nimiž existuje isomorfismus. Schyluje se k důležitým výsledkům, které vytváří spojitost mezi dimenzí a isomorfismem prostorů. Již víme z věty 6.31(4), že isomorfní prostory mají stejnou dimenzi. Platí to i naopak?

Tvrzení 6.34. *Buď V vektorový prostor nad tělesem $\mathbb{T}$ dimenze n s bází B. Pak zobrazení $x \mapsto [x]_B$ je isomorfismus mezi prostory V a $\mathbb{T}^n$ nad $\mathbb{T}$.*

Důkaz. Nechť báze B sestává z vektorů $v_1, \dots, v_n$. Snadno se nahlédne, že zobrazení $x \mapsto [x]_B$ je lineární, že je na a prosté: Prostota plyne z jednoznačnosti souřadnic, věta 5.31. Dále, zobrazení je „na“, protože každá n-tice $(\alpha_1, \dots, \alpha_n) \in \mathbb{T}^n$ představuje souřadnice nějakého vektoru, konkrétně vektoru $\sum_{i=1}^n \alpha_i v_i$. Linearitu zobrazení dokážeme takto. Buďte $u, v \in V$ a nechť $u = \sum_{i=1}^n \alpha_i v_i$ a $v = \sum_{i=1}^n \beta_i v_i$, potom $u + v = \sum_{i=1}^n (\alpha_i + \beta_i) v_i$. Tedy

$$[u]_B + [v]_B = (\alpha_1, \dots, \alpha_n)^T + (\beta_1, \dots, \beta_n)^T = (\alpha_1 + \beta_1, \dots, \alpha_n + \beta_n)^T = [u + v]_B.$$

Podobně, pro každé $\gamma \in \mathbb{T}$ je $\gamma[u]_B = \gamma(\alpha_1, \dots, \alpha_n)^T = (\gamma\alpha_1, \dots, \gamma\alpha_n)^T = [\gamma u]_B$. $\square$

Věta 6.35 (Isomorfismus n-dimenzionálních prostorů)**.** *Všechny n-dimenzionální vektorové prostory nad tělesem $\mathbb{T}$ jsou navzájem isomorfní.*

Důkaz. Podle tvrzení 6.34 jsou všechny n-dimenzionální vektorové prostory nad tělesem $\mathbb{T}$ isomorfní s $\mathbb{T}^n$ nad $\mathbb{T}$, a tím pádem i navzájem mezi sebou, neboť složení isomorfismů je zase isomorfismus. $\square$

Věta říká, že všechny n-dimenzionální prostory nad stejným tělesem jsou navzájem isomorfní. To znamená, že jsou z lineárněalgebraického pohledu stejné, přestože každý má svá specifika, zvláštní operace atp. Z hlediska lineárních prostorů se ale chovají stejně. Tudíž při hledání dimenze, ověřování lineární nezávislosti atp. stačí přejít isomorfismem do prostoru $\mathbb{T}^n$ nad $\mathbb{T}$, kde se pracuje mnohem lépe. Isomorfismus totiž zachovává lineární nezávislost vektorů, zachovává lineární závislost vektorů, a také zachovává dimenzi obrazu podprostoru.

Příklad 6.36. Uvažujme polynomy

$$2x^3 + x^2 + x + 3, \quad x^3 + 2x^2 + 3x + 1, \quad x^3 - x^2 - 2x + 2, \quad 4x^3 - x^2 - 3x + 7.$$

Jsou lineárně nezávislé? Jakou dimenzi má prostor jimi generovaný? Jaká je jeho báze? Na tyto otázky snadno odpovíme při použití isomorfismu $a_3x^3 + a_2x^2 + a_1x + a_0 \mapsto (a_3, a_2, a_1, a_0)$. Takto se polynomy zobrazí na vektory

$$(2, 1, 1, 3)^T, \quad (1, 2, 3, 1)^T, \quad (1, -1, -2, 2)^T, \quad (4, -1, -3, 7)^T.$$

Nyní již standardním způsobem (příklad 5.64) zjistíme, že vektory (a tedy i polynomy) jsou lineárně závislé, generují dvou-dimenzionální podprostor a bázi tvoří například první dva. $\square$

Pro lineární zobrazení $f\colon \mathbb{R}^n \to \mathbb{R}^m$ definované předpisem $f(x) = Ax$ platí $\operatorname{Ker}(f) = \operatorname{Ker}(A)$ a $f(\mathbb{R}^n) = \mathcal{S}(A)$. Pro lineární zobrazení mezi jinými prostory tato vlastnost pochopitelně neplatí, ale dá se ukázat aspoň shoda dimenzí.

Věta 6.37 (O dimenzi jádra a obrazu)**.** *Buď $f\colon U \to V$ lineární zobrazení, U, V prostory nad $\mathbb{T}$, B_1 báze prostoru U a B_2 báze prostoru V. Označme $A = {}_{B_2}[f]_{B_1}$. Pak:*

(1) $\dim \operatorname{Ker}(f) = \dim \operatorname{Ker}(A)$,

(2) $\dim f(U) = \dim \mathcal{S}(A) = \operatorname{rank}(A)$.

Důkaz. (1) Podle věty 6.31(4) stačí sestrojit isomorfismus mezi prostory $\operatorname{Ker}(f)$ a $\operatorname{Ker}(A)$. Isomorfismem může být např. zobrazení $x \in \operatorname{Ker}(f) \mapsto [x]_{B_1}$. Z tvrzení 6.34 víme, že je lineární a prosté. Zbývá ukázat, že $[x]_{B_1} \in \operatorname{Ker}(A)$ a že zobrazení je „na“. Buď $x \in \operatorname{Ker}(f)$, pak

$$o = [o]_{B_2} = [f(x)]_{B_2} = {}_{B_2}[f]_{B_1} \cdot [x]_{B_1},$$

tedy $[x]_{B_1}$ náleží do jádra matice A. Také naopak, pro každé $[x]_{B_1} \in \operatorname{Ker}(A)$ je $f(x) = o$.

(2) Označme $\dim U = n$, $\dim V = m$. Opět sestrojíme isomorfismus, nyní mezi $f(U)$ a $\mathcal{S}(A)$, a to takto $y \in f(U) \mapsto [y]_{B_2}$. A opět, zobrazení je lineární a prosté. Dále, pro $y \in f(U)$ existuje $x \in U$ takové, že $f(x) = y$. Nyní $[y]_{B_2} = [f(x)]_{B_2} = A \cdot [x]_{B_1}$, tedy $[y]_{B_2}$ náleží do sloupcového prostoru $\mathcal{S}(A)$. A naopak, pro každé $b \in \mathcal{S}(A)$ existuje $a \in \mathbb{T}^n$ takové, že $b = Aa$. Čili pro vektor $x \in U$ takový, že $[x]_{B_1} = a$, platí $y := f(x) \in f(U)$ a zároveň $[y]_{B_2} = [f(x)]_{B_2} = A \cdot [x]_{B_1} = Aa = b \in \mathcal{S}(A)$. □

Poznámka 6.38. Důkaz věty 6.37 je konstruktivní – říká nejen jak spočítat dimenzi jádra a obrazu f, ale také jak najít jejich báze. Je-li $x_1, \ldots, x_k$ báze $\operatorname{Ker}(A)$, pak tyto vektory tvoří souřadnice (vzhledem k bázi B_1) báze $\operatorname{Ker}(f)$. Podobně, je-li $y_1, \ldots, y_r$ báze prostoru $\mathcal{S}(A)$, pak tyto vektory představují souřadnice báze prostoru $f(U)$ vzhledem k B_2.

Jako důsledek věty 6.37 dostáváme následující zobecnění rovnosti z věty 6.31(4), neboť pro isomorfismus máme $\dim \operatorname{Ker}(f) = 0$.

Důsledek 6.39. *Buď $f\colon U \to V$ lineární zobrazení, pak* $\dim U = \dim \operatorname{Ker}(f) + \dim f(U)$.

Důkaz. Podle věty 5.66 platí pro matici A typu $m \times n$ rovnost $n = \dim \operatorname{Ker}(A) + \operatorname{rank}(A)$. Speciálně, pro $A = {}_{B_2}[f]_{B_1}$ dostáváme hledanou identitu, neboť $n = \dim U$, $\dim \operatorname{Ker}(f) = \dim \operatorname{Ker}(A)$ a $\dim f(U) = \operatorname{rank}(A)$. □

S ohledem na věty 6.10 a 6.37 pak dostáváme, že lineární zobrazení $f\colon U \to V$ je prosté právě tehdy, když $\dim U = \dim f(U)$, neboli $\dim U = \operatorname{rank}({}_{B_2}[f]_{B_1})$. Nutná a postačující podmínka pro to, aby f bylo prosté, tedy je, aby matice zobrazení f vzhledem k libovolným bázím měla lineárně nezávislé sloupce.

Jak poznáme, že lineární zobrazení $f\colon U \to V$ je „na“? Tuto situaci můžeme vyjádřit podmínkou $\dim V = \dim f(U)$, neboli $\dim V = \operatorname{rank}({}_{B_2}[f]_{B_1})$. Ekvivalentně tedy matice f vzhledem k libovolným bázím musí mít lineárně nezávislé řádky.

Příklad 6.40. Mějme lineární zobrazení $f\colon \mathbb{R}^3 \to \mathcal{P}^2$ dané

$${}_{B_2}[f]_{B_1} = A = \begin{pmatrix} 1 & 1 & 1 \\ 3 & 2 & 0 \\ 0 & 1 & 3 \end{pmatrix},$$

kde

$$B_1 = \{(1,2,1)^T,\ (0,1,1)^T,\ (1,2,4)^T\},$$
$$B_2 = \{x^2 - 2x + 3,\ x - 1,\ 2x^2 + x\}.$$

Protože $\operatorname{rank}(A) = 2$, dostáváme ihned, že $\dim \operatorname{Ker}(f) = 3 - \operatorname{rank}(A) = 1$ a $\dim f(\mathbb{R}^3) = \operatorname{rank}(A) = 2$. Jelikož má jádro kladnou dimenzi a je tudíž netriviální, podle věty 6.10 to znamená, že zobrazení f není prosté. Jelikož má obraz dimenzi 2, ale prostor $\mathcal{P}^2$ má dimenzi 3, tak zobrazení f není „na“.

Báze $\mathrm{Ker}(A)$ je $(2,-3,1)^T$, což reprezentuje souřadnice hledaného vektoru v bázi B_1. Tedy báze $\mathrm{Ker}(f)$ je tvořena vektorem

$$2(1,2,1)^T - 3(0,1,1)^T + 1(1,2,4)^T = (3,3,3)^T.$$

Báze $\mathcal{S}(A)$ je $(1,3,0)^T$, $(1,2,1)^T$, což opět reprezentuje souřadnice hledaných vektorů. Tudíž bázi obrazu $f(\mathbb{R}^3)$ tvoří dva vektory

$$\begin{aligned} 1(x^2-2x+3)+3(x-1)+0(2x^2+x) &= x^2+x,\\ 1(x^2-2x+3)+2(x-1)+1(2x^2+x) &= 3x^2+x+1. \end{aligned}$$

□

6.4 Prostor lineárních zobrazení

Není těžké nahlédnout, že množina lineárních zobrazení z prostoru U nad $\mathbb{T}$ dimenze n do prostoru V nad $\mathbb{T}$ dimenze m tvoří vektorový prostor: Součet lineárních zobrazení $f,g\colon U\to V$ je opět lineární zobrazení $(f+g)\colon U\to V$, a násobek αf lineárního zobrazení $f\colon U\to V$ je také lineární zobrazení. Nulovým vektorem je zobrazení $u\mapsto o_V\ \forall u\in U$.

Navíc, protože každé lineární zobrazení je jednoznačně určeno maticí vzhledem k daným bázím, je tento prostor lineárních zobrazení isomorfní s prostorem matic $\mathbb{T}^{m\times n}$ a má tedy dimenzi mn. Příslušným isomorfismem pak může být zobrazení $f\mapsto {}_{B_2}[f]_{B_1}$, kde B_1 je pevná báze prostoru U a B_2 je pevná báze prostoru V. Linearita tohoto zobrazení plyne jednoduše (díky linearitě souřadnic) z vlastností

$$\begin{aligned} {}_{B_2}[f+g]_{B_1} &= {}_{B_2}[f]_{B_1} + {}_{B_2}[g]_{B_1},\\ {}_{B_2}[\alpha f]_{B_1} &= \alpha\, {}_{B_2}[f]_{B_1}. \end{aligned}$$

Nejzajímavější případ prostoru lineárních zobrazení je pro $V=\mathbb{T}$.

Definice 6.41. Buď V vektorový prostor nad $\mathbb{T}$. Pak *lineární funkcionál* (nebo též lineární forma) je libovolné lineární zobrazení z V do $\mathbb{T}$. *Duální prostor*, značený V^*, je vektorový prostor všech lineárních funkcionálů.

Je-li $\dim V = n$, pak také $\dim V^* = n$. Je-li $v_1,\dots,v_n$ báze V, pak duální prostor má např. bázi $f_1,\dots,f_n$, kde f_i je určeno obrazy báze $f_i(v_i)=1$ a $f_i(v_j)=0$ pro $i\neq j$. Tato báze se nazývá duální k bázi $v_1,\dots,v_n$.

Pro konečně generovaný prostor je tedy V isomorfní s duálním prostorem V^*, s duálem k duálnímu prostoru V^{**} atd. Pro nekonečně generované prostory už to pravda obecně není. Nicméně vždy existuje kanonické vnoření V do V^{**}. Pokud navíc platí, že V a V^{**} jsou isomorfní, tak V má určité pěkné vlastnosti.

Další podrobnosti odkrývá obor zvaný funkcionální analýza. I když se matematika za tím může zdát hodně náročná, pomáhá řešit tak složité problémy jako je nalezení rovnovážné pozice membrány vychýlené nějakou překážkou, problémy ve zpracování a analýze signálů či obrázků, ve strojovém učení, nemluvě o fyzikálních problémech (třeba kvantové mechaniky) a mnoha jiných.

Problémy

6.1. Uvažujme lineární zobrazení $x\mapsto Dx$, kde $D\in\mathbb{R}^{n\times n}$ je regulární. Buď $A\in\mathbb{R}^{m\times n}$ a $b\in\mathbb{R}^m$.

(a) Co je obrazem množiny $\{x\in\mathbb{R}^n;\ Ax=b\}$?

(b) Co se zobrazí na množinu $\{x\in\mathbb{R}^n;\ Ax=b\}$?

Jak se změní předchozí výsledky když D nebude regulární?

6.2. Dokažte, že lineární zobrazení $f\colon U\to V$ je

(a) prosté právě tehdy, když existuje lineární zobrazení $g\colon V\to U$ takové, že $g\circ f = id$ (na U).

(b) „na“ právě tehdy, když existuje lineární zobrazení $g\colon V\to U$ takové, že $f\circ g = id$ (na V).

Shrnutí ke kapitole 6. Lineární zobrazení

Lineární zobrazení mezi vektorovými prostory zachovává strukturu lineárních kombinací: lineární kombinaci vektorů zobrazuje na tutéž lineární kombinaci jejich obrazů. K zadání lineárního zobrazení stačí uvést, kam se zobrazí vektory nějaké báze, a to plně určuje i obrazy ostatních vektorů i tím celého prostoru.

U aritmetických prostorů (typu $\mathbb{T}^n$) se lineární zobrazení dá vyjádřit maticově jako $x \mapsto Ax$. Každá matice tedy odpovídá nějakému lineárnímu zobrazení a naopak každé lineární zobrazení má maticové vyjádření. Tato dvojakost je velmi důležitá, protože mnoho problémů lze nahlížet algebraicky (operace s maticí A) nebo geometricky (pomocí lineárního zobrazení $x \mapsto Ax$). Řada vlastností lineárního zobrazení $x \mapsto Ax$ pak opět souvisí s vlastnostmi matice A:

- skládání zobrazení odpovídá maticovému součinu,
- zobrazení je prosté právě tehdy, když v jádru matice je pouze o,
- hodnost matice udává dimenzi obrazu,
- atp.

U obecných prostorů je situace trochu složitější, ale věci tam fungují podobně. Jenom se musí pracovat se souřadnicemi namísto vektorů samotných. Souřadnice jsou vektory z prostoru $\mathbb{T}^n$, tudíž výše zmíněné postřehy lze přizpůsobit na obecný případ. Při práci v souřadnicích se pak hojně využije matice přechodu, která převádí souřadnice v jedné bázi na souřadnice v jiné bázi; tedy změnu souřadného systému lze opět efektivně reprezentovat maticově.

Lineární zobrazení, které je bijekcí, se nazývá isomorfismus. Isomorfismy odpovídají regulárním maticím; proto také k nim (isomorfismům i maticím) existují inverze. Prostory, mezi kterými existuje isomorfismus, pak nazýváme isomorfní. Isomorfní prostory jsou z pohledu lineární algebry vlastně skoro stejné – mají jiné prvky, ale chovají se stejně. Mají také stejnou dimenzi, ale toto pozorování platí i naopak: Všechny n-dimenzionální prostory nad stejným tělesem jsou vzájemně isomorfní. To nám umožňuje nad každým prostorem snadno pracovat jako kdyby to byl prostor $\mathbb{T}^n$.

Kapitola 7

Afinní podprostory

Toto je velmi letmý úvod do afinních podprostorů. Podrobněji o afinních prostorech pojednává například Bican [2009].

Vektorové prostory a podprostory jsou omezeny tím, že musí obsahovat nulový vektor. Motivací pro afinní podprostory je takové zobecnění, abychom se vyhnuli této restrikci a přiblížili se reálným situacím. Afinním podprostorem v $\mathbb{R}^3$ tak může být jakákoli přímka či rovina, ne jenom ta procházející počátkem.

7.1 Základní pojmy

Afinní podprostor je možno definovat axiomaticky podobně jako vektorový prostor, ale pro naše účely se omezíme na speciální typ.

Definice 7.1 (Afinní podprostor)**.** Buď V vektorový prostor nad $\mathbb{T}$. Pak *afinní podprostor* je jakákoli množina $M \subseteq V$ tvaru

$$M = U + a = \{v + a;\ v \in U\},$$

kde $a \in V$ a U je vektorový podprostor V.

Afinní podprostor (používá se i pojem afinní prostor či afinní množina se stejným významem) je tedy jakýkoli podprostor U „posunutý“ nějakým vektorem a.

Protože $o \in U$, je $a \in M$. Tento reprezentant a není jednoznačný, můžeme zvolit libovolný vektor z M. Naopak, U je u každého afinního podprostoru určený jednoznačně (problém 7.1).

Příklad 7.2. Buď V vektorový prostor. Každý jeho vektorový podprostor U je zároveň jeho afinním podprostorem, neboť lze volit $a = o$, čímž $U = U + o$ je tvaru afinního podprostoru.

Dále, pro každý vektor $a \in V$ je množina $\{a\}$ jednoprvkový afinní podprostor ve V. Ten dostaneme volbou $U := \{o\}$, protože potom $U + a = \{a\}$. □

Poznámka 7.3 (Dláždění afinními podprostory)**.** Buď V vektorový prostor a U jeho podprostor. Pak afinní podprostory tvaru $U + a$, $U + a'$ jsou buď shodné či disjunktní. Navíc každý vektor $v \in V$ leží v nějakém afinním podprostoru tohoto tvaru, například v $U + v$. Tudíž prostor V lze rozložit na disjunktní sjednocení afinních podprostorů tvaru $U + a$ pro vhodné volby vektorů a.

Nad tělesem reálných čísel, nebo obecněji nad jakýmkoliv tělesem charakteristiky různé od 2, můžeme afinní podprostory charakterizovat i jinak. Zatímco vektorové podprostory jsou takové množiny vektorů, které jsou uzavřené na lineární kombinace, afinní podprostory jsou takové množiny vektorů, které jsou uzavřené na tzv. afinní kombinace. *Afinní kombinace* vektorů x, y je výraz (vektor) $\alpha x + (1-\alpha)y$, kde α je libovolný skalár. Afinní kombinaci lze přepsat do tvaru $\alpha x + (1-\alpha)y = y + \alpha(x-y)$, což je parametrický popis přímky s bodem y a směrnicí $x - y$. Jinými slovy, afinní podprostor s každými dvěma body musí obsahovat i přímku, která jimi prochází:

Věta 7.4 (Charakterizace afinního podprostoru)**.** *Buď V vektorový prostor nad tělesem $\mathbb{T}$ charakteristiky různé od 2, a buď $\emptyset \neq M \subseteq V$. Pak M je afinní, tj. je tvaru $M = U + a$ právě tehdy, když pro každé $x, y \in M$ a $\alpha \in \mathbb{T}$ platí $\alpha x + (1-\alpha)y \in M$.*

Důkaz. Implikace „$\Rightarrow$“. Buď $x, y \in M$, tedy jsou tvaru $x = u + a$, $y = v + a$, kde $u, v \in U$. Potom $\alpha x + (1-\alpha)y = \alpha(u+a) + (1-\alpha)(v+a) = \alpha u + (1-\alpha)v + a \in U + a = M$.

Implikace „$\Leftarrow$“. Ukážeme, že stačí zvolit $a \in M$ libovolně pevně a $U := M - M = \{x - y;\ x, y \in M\}$. Tedy ukážeme, že $M = (M - M) + a$.

„$\subseteq$“: Buď $x \in M$, pak $x = x - a + a \in (M - M) + a = U + a$.

„$\supseteq$“: Buď $x - y + a \in (M - M) + a$. Protože $a, x, y \in M$ dostáváme, že afinní kombinace $\frac{1}{2}a + \frac{1}{2}x \in M$ a také $2\left(\frac{1}{2}a + \frac{1}{2}x\right) + (1-2)y = x - y + a \in M$. □

Nad tělesem charakteristiky 2 tato charakterizace obecně nefunguje. Stačí vzít za příklad prostor $\mathbb{Z}_2^n$ nad $\mathbb{Z}_2$, v němž je každá množina vektorů uzavřená na afinní kombinace.

Věta 7.5. *Množina řešení soustavy rovnic $Ax = b$ je prázdná nebo afinní. Je-li neprázdná, můžeme tuto množinu řešení vyjádřit ve tvaru $\mathrm{Ker}(A) + x_0$, kde x_0 je jedno libovolné řešení soustavy.*

Důkaz. Pokud x_1 je řešením, pak lze psát $x_1 = x_1 - x_0 + x_0$. Stačí ukázat, že $x_1 - x_0 \in \mathrm{Ker}(A)$. Dosazením $A(x_1 - x_0) = Ax_1 - Ax_0 = b - b = 0$. Tedy $x_1 \in \mathrm{Ker}(A) + x_0$. Naopak, je-li $x_2 \in \mathrm{Ker}(A)$, pak $x_2 + x_0$ je řešením soustavy, neboť $A(x_2 + x_0) = Ax_2 + Ax_0 = 0 + b = b$. □

Poznámka. Platí i obrácená implikace, tedy každý afinní podprostor lze popsat pomocí soustavy rovnic (v souřadnicích), viz Bican [2009].

Příklad 7.6. Věta 7.5 dává následující geometrický pohled na soustavy rovnic při perturbaci pravé strany. Nechť $Ax = b$ je řešitelná, tedy popisuje afinní podprostor $\mathrm{Ker}(A)+x_0$. Změníme-li pravou stranu soustavy b na b', pak buďto soustava přestane mít řešení (což se nestává často), nebo se afinní podprostor posune na $\mathrm{Ker}(A) + x_0'$, kde x_0' je jedno vybrané řešení. Typicky se tedy množina řešení při změně pravé strany pouze posouvá nějakým směrem. □

Dimenze afinního podprostoru

Definice 7.7 (Dimenze afinního podprostoru)**.** *Dimenze* afinního podprostoru $M = U + a$ je definována jako $\dim(M) := \dim(U)$.

Protože každý vektorový podprostor prostoru V je zároveň jeho afinním podprostorem, definice tedy zobecňuje pojem dimenze, zavedený pro vektorové prostory. Definice přirozeně zavádí dimenzi bodu jako nula, dimenzi přímky v $\mathbb{R}^n$ jako jedna a dimenzi roviny jako dva.

Také umožňuje definovat *přímku* p v libovolném vektorovém prostoru W nad $\mathbb{T}$ jakožto afinní množinu dimenze jedna. Jinými slovy, $p = \mathrm{span}\{v\} + a$, kde $a, v \in U$ a $v \neq o$. Odsud dostáváme i známý parametrický popis přímky $p = \{\alpha v + a;\ \alpha \in \mathbb{T}\}$.

Nadrovinou v prostoru dimenze n rozumíme pak libovolný afinní podprostor dimenze $n-1$. Tedy např. v $\mathbb{R}^2$ jsou to přímky, v $\mathbb{R}^3$ roviny, atd.

Příklad 7.8. Množina $\{e^x + \alpha \sin x;\ \alpha \in \mathbb{R}\}$ je přímka v prostoru funkcí $\mathcal{F}$. □

Příklad 7.9. Pro jakékoli $a \in \mathbb{R}^n \setminus \{o\}$ a $b \in \mathbb{R}$ je množina popsaná rovnicí $a^T x = b$ nadrovinou v $\mathbb{R}^n$. A naopak, každá nadrovina v $\mathbb{R}^n$ se dá popsat rovnicí $a^T x = b$ pro určité $a \in \mathbb{R}^n \setminus \{o\}$ a $b \in \mathbb{R}$. □

Afinní nezávislost

Lineární nezávislost vektorů $x_1, \ldots, x_n$ znamenala, že podprostor, který tyto vektory generují, se nedá vygenerovat nějakou vlastní podmnožinou těchto vektorů. Žádný z nich tedy není v jistém smyslu zbytečný. Chtěli bychom analogickou vlastnost mít i pro afinní podprostory, tedy umět charakterizovat nejmenší množinu vektorů, jenž jednoznačně určují daný afinní podprostor. Toto vede na pojem afinní nezávislost. Afinní podprostor jsme definovali jako $U + a$, tedy vektorový podprostor U posunutý o vektor a. Nabízí se tedy definovat afinní nezávislost vektorů tak, že je posuneme zpět a po výsledných vektorech budeme požadovat lineární nezávislost.

Definice 7.10 (Afinní nezávislost)**.** Vektory $x_0, x_1, \ldots, x_n$ vektorového prostoru jsou *afinně nezávislé*, pokud $x_1 - x_0, \ldots, x_n - x_0$ jsou lineárně nezávislé. V opačném případě vektory nazýváme *afinně závislé*.

Příklad 7.11. Vektory $(1,1)^T, (2,2)^T, (1,2)^T \in \mathbb{R}^2$ jsou sice lineárně závislé, ale afinně nezávislé. □

Příklad 7.12. Dva různé body na přímce v $\mathbb{R}^n$ jsou afinně nezávislé. Nicméně, tři body na přímce už jsou afinně závislé, protože přímka je jednoznačně určena již dvěma body. □

Není těžké nahlédnout, že afinní nezávislost nezávisí na pořadí vektorů, a tedy ani na volbě x_0 (dokažte!).

Afinní nezávislost navíc umožňuje jednoduše formalizovat to, co známe pod pojmem „Mějme body v obecné poloze“: Množina bodů $x_1, \ldots, x_m \in \mathbb{R}^n$ je v obecné poloze, když každá její podmnožina velikosti nanejvýš $n + 1$ je afinně nezávislá. Tedy například v rovině $\mathbb{R}^2$ jsou dané body v obecné poloze pokud žádné tři neleží na společné přímce.

Souřadnice v afinním podprostoru

Buď $M = U + a$ afinní podprostor a $B = \{v_1, \ldots, v_n\}$ báze U. Pak každé $x \in M$ se dá jednoznačně zapsat ve tvaru $x = a + \sum_{i=1}^n \alpha_i v_i$. Tedy $S = \{a, v_1, \ldots, v_n\}$ lze považovat za *souřadný systém* a vektor $[x]_S := [x - a]_B = (\alpha_1, \ldots, \alpha_n)^T$ za příslušné *souřadnice*.

Příklad 7.13. Buď $v = (2,1)^T$, $a = (1,2)^T$ a uvažujme přímku $\operatorname{span}\{v\} + a$. Uvažujme dále souřadný systém $S = \{a, v\}$. Potom bod $x = (5,4)^T$ lze vyjádřit jako $x = a + 2v$, a proto jeho souřadnice jsou $[x]_S = (2)$.

Uvažujme nyní jiný vektor $a' = (-1,1)^T$. Potom se vyjádření vektoru x změní na $x = a' + 3v$, a proto jeho souřadnice v systému $S' = \{a', v\}$ budou $[x]_{S'} = (3)$. □

K přechodu mezi souřadnými systémy pak můžeme použít naši známou matici přechodu přesně tak, jak jsme zvyklí. V případě, že měníme i vektor a, objeví se i konstantní aditivní člen.

Tvrzení 7.14. *Buď $M = U + a$ afinní podprostor a $B = \{v_1, \dots, v_n\}$, $B' = \{v'_1, \dots, v'_n\}$ dvě báze U.*

(1) Pro dva dané souřadné systémy $S = \{a, v_1, \dots, v_n\}$ a $S' = \{a, v'_1, \dots, v'_n\}$ máme

$$[x]_{S'} = {}_{B'}[id]_B \cdot [x]_S, \quad \forall x \in U + a.$$

(2) Pro dva dané souřadné systémy $S = \{a, v_1, \dots, v_n\}$ a $S' = \{a', v'_1, \dots, v'_n\}$ máme

$$[x]_{S'} = [a - a']_{B'} + {}_{B'}[id]_B \cdot [x]_S, \quad \forall x \in U + a.$$

Důkaz.

(1) Nechť $x \in U + a$ je tvaru $x = u + a$. Pak

$$[x]_{S'} = [u]_{B'} = {}_{B'}[id]_B \cdot [u]_B = {}_{B'}[id]_B \cdot [x]_S.$$

(2) Nechť $x \in U + a$ je tvaru $x = u + a$, tj. $x = (a - a') + u + a'$. Pak

$$[x]_{S'} = [(a - a') + u]_{B'} = [(a - a')]_{B'} + [u]_{B'} = [(a - a')]_{B'} + {}_{B'}[id]_B \cdot [x]_S.$$ □

Vztah afinních podprostorů

Afinní podprostory $U + a$ a $W + b$ jsou *rovnoběžné*, pokud $U \Subset W$ nebo $W \Subset U$; *různoběžné*, pokud nejsou rovnoběžné a mají neprázdný průnik; a *mimoběžné*, pokud nejsou rovnoběžné a mají prázdný průnik.

Příklad 7.15. Buď v libovolný vektor vektorového prostoru V. Pak afinní podprostor $\{v\}$ je rovnoběžný s každým afinním podprostorem V. Stejnou vlastnost má i celý prostor V. □

Afinní zobrazení

Buď $g\colon U \to V$ lineární zobrazení a mějme pevný vektor $b \in V$. Potom *afinní zobrazení* má tvar $f(u) = g(u) + b$. Jednoduchým příkladem afinního zobrazení je posunutí, tedy zobrazení $g\colon V \to V$ s popisem $f(x) = x + b$, kde $b \in V$ je pevné.

Afinní zobrazení nemusí zobrazovat nulový vektor v U na nulový vektor ve V, protože jsou obrazy posunuté o aditivní člen b.

Příklad 7.16. Buď $U + a$ afinní podprostor dimenze k v prostoru V, a buď S souřadný systém v $U + a$. Pak zobrazení $f(v) = [v]_S$ je afinní zobrazení zobrazující isomorfně $U + a$ na $\mathbb{T}^k$. □

Tvrzení 7.17.

(1) Obraz afinního podprostoru při afinním zobrazení je afinní podprostor.

(2) Složením dvou afinních zobrazení dostaneme opět afinní zobrazení.

(3) Buď $f\colon U \to V$ lineární zobrazení a $v \in V$. Pak úplný vzor vektoru v

$$f^{-1}(v) := \{u \in U;\ f(u) = v\}$$

je buďto prázdná množina, nebo afinní podprostor v U.

Tvrzení nedokazujeme a necháme je čtenáři na rozmyšlení, nicméně obohatíme je o několik poznámek. Bod (1) tvrzení se nahlédne snadno pokud si uvědomíme, že obraz afinního podprostoru $U + a$ se dá vyjádřit jako $f(U + a) = f(U) + f(a)$.

Bod (3) tvrzení 7.17 má analogii s řešením soustav lineárních rovnic. Uvažujme lineární zobrazení $f\colon \mathbb{R}^n \to \mathbb{R}^m$ vyjádřené maticově $f(x) = Ax$ a mějme dané $b \in \mathbb{R}^m$. Potom hledat všechna řešení soustavy $Ax = b$ vlastně znamená najít úplný vzor vektoru b, tedy množinu

$$\begin{aligned} f^{-1}(b) &= \{x \in \mathbb{R}^m;\ f(x) = b\} \\ &= \{x \in \mathbb{R}^m;\ Ax = b\}. \end{aligned}$$

Z věty 7.5 pak víme, že množina řešení soustavy $Ax = b$ je prázdná nebo afinní podprostor.

Ještě jiný pohled na bod (3) tvrzení 7.17 je pomocí jádra lineárního zobrazení. Podobně jako ve větě 7.5 můžeme úplný vzor vektoru v vyjádřit jako afinní podprostor $\mathrm{Ker}(f) + u$, kde $u \in U$ je jeden pevný vzor vektoru v.

7.2 Aplikace

Příklad 7.18 (Rovnice ano, ale nerovnice?)**.** V minulých kapitolách jsme studovali soustavy lineárních rovnic $Ax = b$, tudíž je přirozená otázka zabývat se i soustavou nerovnic $Ax \leq b$. Nerovnost mezi vektory znamená nerovnost v každé složce, tedy $(Ax)_i \leq b_i$ pro všechna i. Dále, omezíme se jen na těleso $\mathbb{R}$, kde máme definováno uspořádání.

Zatímco jedna rovnice vytyčuje v prostoru nadrovinu a soustava rovnic pak nějaký afinní podprostor, tak jedna nerovnice vymezuje v prostoru poloprostor a soustava nerovnic pak tudíž průnik poloprostorů, což je *konvexní polyedr (mnohostěn)*.

Příklad v $\mathbb{R}^2$.

Příklad v $\mathbb{R}^3$.

Pro konkrétnost uvažujme soustavu lineárních nerovnic

$$\begin{aligned} x_1 + 2x_2 &\leq 9, \\ x_1 - x_2 &\leq 3, \\ -x_1 + 2x_2 &\leq 3, \\ -x_1 - x_2 &\leq -3. \end{aligned}$$

Tato soustava popisuje čtyřúhelník vykreslený na obrázku

Vrchol $v_1 = (5, 2)^T$ leží v průniku nadrovin $x_1 + 2x_2 = 9$ a $x_1 - x_2 = 3$. Je tedy řešením odpovídající soustavy rovnic a podle věty 7.5 je to afinní podprostor dimenze nula. Hrana spojující vrcholy v_1, v_2 leží na přímce $x_1 + 2x_2 = 9$. Je tedy určena jedno-dimenzionálním afinním podprostorem $x_1 + 2x_2 = 9$.

Podobným způsobem pokračujeme ve vyšších dimenzích. Například tří-dimenzionální polyedry, jako je krychle, osmistěn atp. mají následující strukturu. Vrcholy jsou nula-dimenzionální afinní podprostory určené průnikem tří nadrovin, tj. popsané třemi rovnicemi. Hrany leží v jedno-dimenzionálním afinním podprostoru popsaným soustavou dvou rovnic. A konečně stěny leží v nadrovině, tedy ve dvou-dimenzionálním afinním podprostoru, který je určen jednou rovnicí.

Na závěr podotkněme, že konvexními polyedry se více zabývá např. obor *lineární programování*. Ten zkoumá nejen konvexní polyedry, ale také nad nimi řeší optimalizační úlohy typu

$$\min\ c^T x \text{ za podmínek } Ax \leq b,$$

kde $c \in \mathbb{R}^n$, $A \in \mathbb{R}^{m \times n}$ a $b \in \mathbb{R}^m$ jsou dané a $x \in \mathbb{R}^n$ je vektor proměnných. □

Příklad 7.19 (Afinní zobrazení a fraktály [Gareth, 2001, sekce 4.4])**.** Fraktál je soběpodobný geometrický útvar na první pohled složitého tvaru. Ukážeme si na příkladu, že i poměrně složitý fraktál může mít jednoduchý popis pomocí afinních zobrazení.

Pomocí čtyř afinních zobrazení dokážeme v rovině vykreslit složitý fraktál. Začněme v počátku a

s danými pravděpodobnostmi uvažujme přechod podle příslušného afinního zobrazení.

$$T_1(x,y) = \begin{pmatrix} 0.86 & 0.03 \\ -0.03 & 0.86 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 0 \\ 1.5 \end{pmatrix} \quad \text{s pravděpodobností } 0.83$$

$$T_2(x,y) = \begin{pmatrix} 0.2 & -0.25 \\ 0.21 & 0.23 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 0 \\ 1.5 \end{pmatrix} \quad \text{s pravděpodobností } 0.08$$

$$T_3(x,y) = \begin{pmatrix} -0.15 & 0.27 \\ 0.25 & 0.26 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 0 \\ 0.45 \end{pmatrix} \quad \text{s pravděpodobností } 0.08$$

$$T_4(x,y) = \begin{pmatrix} 0 & 0 \\ 0 & 0.17 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \text{s pravděpodobností } 0.01$$

Navštívené body postupně vykreslí tzv. Barnsleyho fraktál ve tvaru listu kapradiny.

2500 iterací.

10000 iterací. □

Příklad 7.20 (Stewartova–Goughova platforma v robotice)**.** Stewartova–Goughova platforma je tzv. paralelní manipulátor v oboru kinematické robotiky. Pevná základna je připevněna několika (většinou šesti) pohyblivými rameny k mobilní plošině. Tyto platformy se využívají jako manipulátory, v simulacích (např. letů), nebo třeba v biomechanice kloubů k ověřování implantátů mimo lidské tělo.

Stewartova–Goughova platforma. [zdroj: Wikipedia]

Základna i mobilní plošina mají své vlastní souřadné systémy, mezi kterými můžeme přecházet pomocí afinního zobrazení. Např., jsou-li $x = (x_1, x_2, x_3)^T$ souřadnice bodu v systému plošiny, pak souřadnice vůči základně získáme jako $x' = Px + c$, kde P matice reprezentující naklonění a c je nějaký pevný vektor reprezentující posun. Pochopitelně, P a c nejsou pevné, ale závisí na míře natažení pohyblivých ramen. Navíc se dá ukázat, že matice P závisí pouze na třech parametrech, protože systém plošiny vzhledem k základně je pouze natočený a není nijak deformovaný (natáhnutý, zkosený atp.).

Uvažujme problém určení délek pohyblivých ramen. Označme $x^{(1)}, \ldots, x^{(6)}$ koncové body ramen u základny v souřadném systému základny a $y^{(1)}, \ldots, y^{(6)}$ koncové body na plošině v souřadném systému

plošiny. Ty druhé převedeme výše zmíněnou transformací do souřadného systému základny: $y'^{(1)} = Px^{(1)} + c, \ldots, y'^{(6)} = Px^{(6)} + c$. Nyní můžeme jednoduše spočítat délku ramen jako vzdálenosti bodů $x^{(i)}$ a $y'^{(i)}$ pro $i = 1, \ldots, 6$.

Typicky ale problém zní opačně: délky ramen známe, protože ty ovládáme, a je potřeba spočítat pozici plošiny, tj. koncových bodů. Jiným problémem pak je třeba zjistit všechny pozice nebo omezit hranice, ve kterých se může plošina nacházet. To už jsou úlohy nad rámec úvodního kurzu lineární algebry. □

Příklad 7.21 (Lineární klasifikátor a neuronové sítě)**.** Mějme data, reprezentovaná vektory $v_1, \ldots, v_m \in \mathbb{R}^n$, a pro každou hodnotu víme, zda patří do skupiny A či B. Nechť v_i patři do skupiny A pro $i \in \mathcal{A}$ a do skupiny B pro $i \in \mathcal{B}$. Chceme sestrojit klasifikátor, který bude umět pro novou hodnotu $v \in \mathbb{R}^n$ automaticky rozhodnout do které skupiny patří. Jednoduchý klasifikátor můžeme sestrojit na základě lineárního separátoru. Jeho podstata spočívá v tom sestrojit nadrovinu $a^T x = b$ takovou, aby vektory skupiny A byly v jedné polorovině a vektory skupiny B byly ve druhé, viz obrázek:

Matematicky popsáno, hledáme $a \in \mathbb{R}^n$, $b \in \mathbb{R}$ tak, aby

$$a^T v_i < b \quad \forall i \in \mathcal{A},$$
$$a^T v_i > b \quad \forall i \in \mathcal{B}.$$

Jestliže takovou nadrovinu najdeme, klasifikace nového údaje $v \in \mathbb{R}^n$ funguje jednoduše. Pokud $a^T v < b$, pak hodnotu v považujeme za člena sklupiny A a v opačném případě za člena skupiny B. Pokud lineární separátor nenajdeme, je situace trochu složitější. Jedna z možností, jak se s tím vypořádat, je vnořit data do prostoru vyšší dimenze, kde je větší šance na úspěch.

Poznamenejme, že lineární klasifikátory se využívají v neuronových sítích. Je to jeden ze způsobů, jaký perceptronové algoritmy učení využívají k hledání váhových koeficientů vazeb neuronů. Více informací viz Šíma and Neruda [1996]. □

Problémy

7.1. Buď $M = U + a$ afinní podprostor. Dokažte, že prostor U je dán jednoznačně.

7.2. Ukažte, že průnik afinních podprostorů je zase afinní podprostor nebo prázdná množina.

7.3. Buďte afinní podprostory $U + a$ a $W + b$ rovnoběžné. Ukažte, že pak jsou disjunktní právě tehdy, když $a - b \notin U \cup W$.

7.4. (Analogie vět o isomorfismu vektorových prostorů.) Ukažte, že afinní prostory $U + a$, $W + b$ mají stejnou dimenzi právě tehdy, když existuje prosté afinní zobrazení $U + a$ na $W + b$.

Shrnutí ke kapitole 7. Afinní podprostory

Podprostor vektorového prostoru musí procházet počátkem. Pokud však podprostor posuneme ve směru nějakého vektoru, dostáváme nový objekt – afinní podprostor. Zatímco vektorové podprostory jsou uzavřené na lineární kombinace, afinní podprostory jsou (za obecmých předpokladů) uzavřené na afinní kombinace. Důležitý je vztah se soustavami lineárních rovnic – množina řešení soustavy $Ax = b$ je afinním podprostorem, a naopak každý afinní podprostor se dá vyjádřit jako řešení určité soustavy.

Řada pojmů a vlastností z vektorových prostorů se přirozeně přenese na afinní podprostory. Tudíž snadno zavedeme pojmy jako (afinní) nezávislost, báze, souřadnice či dimenze. Afinní zobrazení pak je zobrazení, které má tvar lineárního zobrazení s konstantním aditivním členem; u prostorů typu $\mathbb{T}^n$ má tvar $x \mapsto Ax + b$. Toto zobrazení má opět podobné vlastnosti jako lineární zobrazení, jenom přeložené do světa afinních podprostorů.

Kapitola 8

Skalární součin

Vektorové prostory byly definovány velice obecně, takže pokryjí velkou třídu objektů. Na druhou stanu, jestliže přidáme další požadavky na to co mají prostory splňovat, tak budeme moci odvodit hlubší výsledky. Konkrétně, skalární součin dává možnost přirozeně zavést pojem kolmosti, velikost a vzdálenost vektorů (a tím i limity) atd.

8.1 Skalární součin a norma

Skalární součin (stejně jako grupu, vektorové prostory aj.) zavádíme obecně výčtem vlastností, které má splňovat.

Definice 8.1 (Skalární součin nad $\mathbb{R}$)**.** Buď V vektorový prostor nad $\mathbb{R}$. Pak *skalární součin* je zobrazení $\langle\cdot,\cdot\rangle\colon V^2 \to \mathbb{R}$, splňující:

(1) $\langle x, x\rangle \geq 0\ \forall x \in V$, a rovnost nastane pouze pro $x = 0$,

(2) $\langle x + y, z\rangle = \langle x, z\rangle + \langle y, z\rangle\ \forall x, y, z \in V$,

(3) $\langle \alpha x, y\rangle = \alpha\langle x, y\rangle\ \forall x, y \in V$, $\forall \alpha \in \mathbb{R}$,

(4) $\langle x, y\rangle = \langle y, x\rangle\ \forall x, y \in V$.

Zobecnění na těleso komplexních čísel je následující. Připomeňme ještě, že komplexně sdružené číslo k $a + bi \in \mathbb{C}$ je definované jako $\overline{a + bi} = a - bi$.

Definice 8.2 (Skalární součin nad $\mathbb{C}$)**.** Buď V vektorový prostor nad $\mathbb{C}$. Pak *skalární součin* je zobrazení $\langle\cdot,\cdot\rangle\colon V^2 \to \mathbb{C}$, splňující:

(1) $\langle x, x\rangle \geq 0\ \forall x \in V$, a rovnost nastane pouze pro $x = 0$,

(2) $\langle x + y, z\rangle = \langle x, z\rangle + \langle y, z\rangle\ \forall x, y, z \in V$,

(3) $\langle \alpha x, y\rangle = \alpha\langle x, y\rangle\ \forall x, y \in V$, $\forall \alpha \in \mathbb{C}$,

(4) $\langle x, y\rangle = \overline{\langle y, x\rangle}\ \forall x, y \in V$.

Kvůli vlastnosti (1), kde je potřeba uspořádání, zavádíme skalární součin pouze nad tělesy $\mathbb{R}$ a $\mathbb{C}$.

Čtvrtá vlastnost u komplexního skalárního součinu zařídí, že $\langle x, x\rangle = \overline{\langle x, x\rangle} \in \mathbb{R}$, tedy $\langle x, x\rangle$ je vždy reálné číslo a lze jej porovnávat s nulou u první vlastnosti.

Vlastnosti (2) a (3) říkají, že skalární součin je lineární funkcí v první složce. Jak je to s druhou?

$$\langle x, y + z\rangle = \overline{\langle y + z, x\rangle} = \overline{\langle y, x\rangle} + \overline{\langle z, x\rangle} = \langle x, y\rangle + \langle x, z\rangle,$$
$$\langle x, \alpha y\rangle = \overline{\langle \alpha y, x\rangle} = \overline{\alpha}\overline{\langle y, x\rangle} = \overline{\alpha}\langle x, y\rangle.$$

Ve druhé složce tedy komplexní skalární součin není lineární, ačkoli reálný skalární součin jest.

Pokud dosadíme $\alpha = 0$, dostáváme $\langle o, x\rangle = \langle x, o\rangle = 0$, tedy násobení jakéhokoli vektoru s nulovým dá nulu.

Příklad 8.3 (Příklady standardních skalárních součinů)**.**

- V prostoru $\mathbb{R}^n$: standardní skalární součin $\langle x, y\rangle = x^T y = \sum_{i=1}^n x_i y_i$.
- V prostoru $\mathbb{C}^n$: standardní skalární součin $\langle x, y\rangle = x^T \overline{y} = \sum_{i=1}^n x_i \overline{y}_i$.
- V prostoru $\mathbb{R}^{m\times n}$: standardní skalární součin $\langle A, B\rangle = \sum_{i=1}^m \sum_{j=1}^n a_{ij} b_{ij}$.
- V $\mathcal{C}_{[a,b]}$, prostoru spojitých funkcí na intervalu $[a, b]$: standardní skalární součin $\langle f, g\rangle = \int_a^b f(x)g(x)dx$.

Výše zmíněné skalární součiny jsou pouze příklady možných zavedení součinů na daných prostorech; jako skalární součin mohou fungovat i jiné operace. Později, ve větě 11.17, popíšeme všechny skalární součiny v prostoru $\mathbb{R}^n$. □

Nadále uvažujme vektorový prostor V nad $\mathbb{R}$ či $\mathbb{C}$ se skalárním součinem. Nejprve ukážeme, že skalární součin umožňuje zavést normu, neboli velikost vektoru.

Definice 8.4 (Norma indukovaná skalárním součinem)**.** *Norma indukovaná skalárním součinem* je definovaná $\|x\| := \sqrt{\langle x, x\rangle}$, kde $x \in V$.

Norma je dobře definovaná díky první vlastnosti z definice skalárního součinu, a je to vždy nezáporná hodnota.

Pro standardní skalární součin v $\mathbb{R}^n$ dostáváme známou eukleidovskou normu $\|x\| = \sqrt{x^T x} = \sqrt{\sum_{i=1}^n x_i^2}$.

Geometrická interpretace standardního skalárního součinu v $\mathbb{R}^n$ je $\langle x, y\rangle = \|x\| \cdot \|y\| \cos(\varphi)$, kde φ je úhel mezi vektory x, y. Speciálně, x, y jsou kolmé právě tehdy, když $\langle x, y\rangle = 0$. V jiných prostorech takováto geometrie chybí, proto kolmost zavedeme právě pomocí vztahu $\langle x, y\rangle = 0$.

Definice 8.5 (Kolmost)**.** Vektory $x, y \in V$ jsou *kolmé*, pokud $\langle x, y\rangle = 0$. Značení: $x \perp y$.

Příklad 8.6 (Příklady kolmých vektorů pro standardní skalární součiny)**.**

- V prostoru $\mathbb{R}^3$: $(1, 2, 3) \perp (1, 1, -1)$.
- V prostoru $\mathcal{C}_{[-\pi,\pi]}$: $\sin x \perp \cos x \perp 1$. □

Věta 8.7 (Pythagorova)**.** *Pokud $x, y \in V$ jsou kolmé, tak $\|x + y\|^2 = \|x\|^2 + \|y\|^2$.*

Důkaz. $\|x + y\|^2 = \langle x + y, x + y\rangle = \langle x, x\rangle + \underbrace{\langle x, y\rangle}_{=0} + \underbrace{\langle y, x\rangle}_{=0} + \langle y, y\rangle = \langle x, x\rangle + \langle y, y\rangle = \|x\|^2 + \|y\|^2$. □

Poznamenejme, že nad $\mathbb{R}$ platí i opačná implikace, ale nad $\mathbb{C}$ obecně nikoli (viz problém 8.2).

Věta 8.8 (Cauchyho–Schwarzova nerovnost[1])**.** *Pro každé $x, y \in V$ platí $|\langle x, y\rangle| \le \|x\| \cdot \|y\|$.*

Důkaz. (Reálná verze) Nejprve ukážeme reálnou verzi, protože má elegantní důkaz. Pro $y = o$ platí nerovnost triviálně, tak předpokládejme $y \neq o$. Uvažujme reálnou funkci $f(t) = \langle x + ty, x + ty\rangle \ge 0$ proměnné $t \in \mathbb{R}$. Pak

$$f(t) = \langle x, x\rangle + t\langle x, y\rangle + t\langle y, x\rangle + t^2\langle y, y\rangle = \langle x, x\rangle + 2t\langle x, y\rangle + t^2\langle y, y\rangle.$$

Jedná se o kvadratickou funkci, která je všude nezáporná, nemůže mít tedy dva různé kořeny. Proto je příslušný diskriminant nekladný:

$$4\langle x, y\rangle^2 - 4\langle x, x\rangle\langle y, y\rangle \le 0.$$

Z toho dostáváme $\langle x, y\rangle^2 \le \langle x, x\rangle\langle y, y\rangle$, odmocněním $|\langle x, y\rangle| \le \|x\| \cdot \|y\|$. □

Důkaz. (Komplexní verze) Je-li $\langle x, y\rangle = 0$, pak tvrzení platí triviálně. Budeme tedy předpokládat, že $\langle x, y\rangle \neq 0$, a zavedeme vektor

$$z := \frac{\langle y, y\rangle}{\langle x, y\rangle} x - y.$$

[1] Nerovnost se také někdy nazývá jen Schwarzova, nebo Cauchyho–Bunjakovského, popř. Cauchyho–Schwarzova–Bunjakovského. Augustin-Louis Cauchy ji dokázal r. 1821 pro prostor $\mathbb{R}^n$ a později ji nezávisle na sobě zobecnili Hermann Amandus Schwarz (1880) a Viktor Jakovlevič Bunjakovskij (1859).

Potom platí

$$\langle z, y\rangle = \left\langle \frac{\langle y, y\rangle}{\langle x, y\rangle} x - y, y \right\rangle = \frac{\langle y, y\rangle}{\langle x, y\rangle}\langle x, y\rangle - \langle y, y\rangle = 0.$$

Vektory z, y jsou tudíž na sebe kolmé a podle Pythagorovy věty

$$\|z + y\|^2 = \|z\|^2 + \|y\|^2,$$

neboli

$$\frac{\langle y, y\rangle^2}{|\langle x, y\rangle|^2}\|x\|^2 = \|z\|^2 + \|y\|^2 \geq \|y\|^2.$$

Zkrácením $\|y\|^2$ a vynásobením $|\langle x, y\rangle|^2$ dostaneme hledanou nerovnost $\|x\|^2\|y\|^2 \geq |\langle x, y\rangle|^2$ v druhých mocninách. □

Občas se Cauchyho–Schwarzova nerovnost uvádí v ekvivalentní podobě

$$|\langle x, y\rangle|^2 \leq \langle x, x\rangle\langle y, y\rangle.$$

Cauchyho–Schwarzova nerovnost je užitečná a často používaná pro odvozování dalších výsledků na obecné bázi, nebo i pro konkrétní algebraické výrazy. Např. pro standardní skalární součin v $\mathbb{R}^n$ dostaneme nerovnost

$$\left(\sum_{i=1}^{n} x_i y_i\right)^2 \leq \left(\sum_{i=1}^{n} x_i^2\right)\left(\sum_{i=1}^{n} y_i^2\right).$$

Další využití viz např. [Krisl, 2008]. My použijeme Cauchyho–Schwarzovu nerovnost hned v následujícím k odvození trojúhelníkové nerovnosti.

Důsledek 8.9 (Trojúhelníková nerovnost)**.** *Pro každé $x, y \in V$ platí $\|x + y\| \leq \|x\| + \|y\|$.*

Důkaz. Nejprve připomeňme, že pro každé komplexní číslo $z = a + bi$ platí: $z + \overline{z} = 2a = 2\,\mathsf{Re}(z)$, a dále $a \leq |z|$. Nyní můžeme odvodit:

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

kde poslední nerovnost plyne z Cauchyho–Schwarzovy nerovnosti. Tedy máme $\|x + y\|^2 \leq (\|x\| + \|y\|)^2$ a odmocněním získáme hledaný vztah. □

Norma indukovaná skalárním součinem je jen jedním typem normy, pojem normy je ale definován obecněji. My budeme vesměs pracovat s normou indukovanou skalárním součinem, takže následující je pouze malou odbočkou.

Definice 8.10 (Norma)**.** Buď V vektorový prostor nad $\mathbb{R}$ nebo $\mathbb{C}$. Pak *norma* je zobrazení $\|\cdot\|\colon V \to \mathbb{R}$, splňující:

(1) $\|x\| \geq 0$ pro všechna $x \in V$, a rovnost nastane pouze pro $x = 0$,

(2) $\|\alpha x\| = |\alpha| \cdot \|x\|$ pro všechna $x \in V$, a pro všechna $\alpha \in \mathbb{R}$ resp. $\alpha \in \mathbb{C}$,

(3) $\|x+y\| \leq \|x\| + \|y\|$.

Tvrzení 8.11. *Norma indukovaná skalárním součinem je normou.*

Důkaz. Vlastnost (1) je splněna díky definici normy indukované skalárním součinem. Vlastnost (3) je ukázána v důsledku 8.9. Zbývá vlastnost (2):

$$\|\alpha x\| = \sqrt{\langle \alpha x, \alpha x\rangle} = \sqrt{\alpha\overline{\alpha}\langle x, x\rangle} = \sqrt{\alpha\overline{\alpha}}\sqrt{\langle x, x\rangle} = |\alpha| \cdot \|x\|. \qquad \square$$

Příklad 8.12 (Příklady norem v $\mathbb{R}^n$)**.** Užitečná třída norem jsou tzv. p-normy. Pro $p = 1, 2, \ldots$ definujeme p-normu vektoru $x \in \mathbb{R}^n$ jako

$$\|x\|_p = \left(\sum_{i=1}^n |x_i|^p\right)^{\frac{1}{p}}.$$

Speciální volby p vedou ke známým normám:

- pro $p = 2$: eukleidovská norma $\|x\|_2 = \sqrt{\sum_{i=1}^n x_i^2}$, což je norma indukovaná standardním skalárním součinem,
- pro $p = 1$: součtová norma $\|x\|_1 = \sum_{i=1}^n |x_i|$; nazývá se manhattanská norma, protože odpovídá reálným vzdálenostem při procházení pravoúhlé sítě ulic v městě,
- pro $p = \infty$ (limitním přechodem): maximová (Čebyševova) norma $\|x\|_\infty = \max_{i=1,\ldots,n} |x_i|$. $\square$

Příklad 8.13 (Jednotková koule)**.** Jednotková koule je množina vektorů, které mají normu nanejvýš 1 a tedy jsou od počátku vzdáleny maximálně 1. Formálně definujeme jednotkovou kouli jako

$$\{x \in V;\ \|x\| \leq 1\}.$$

Tento pojem je nejenom užitečný nástroj, ale zároveň umožňuje normu geometricky vizualizovat. Uvažujme pro konkrétnost rovinu $\mathbb{R}^2$ a tři základní normy z předchozího přikladu 8.12. Jednotková koule má tvar:

eukleidovská norma součtová norma maximová norma

Jiné normy mají za jednotkovou kouli jiný geometrický objekt. Každá jednotková koule v $\mathbb{R}^n$ ale musí být množina, která je uzavřená, omezená, symetrická dle počátku, konvexní (tj., s každými dvěma body obsahuje i jejich spojnici) a počátek leží v jejím vnitřku. Platí i opačné tvrzení – každá množina splňující tyto vlastnosti představuje jednotkovou kouli nějaké normy. $\square$

Příklad 8.14 (Příklady norem v $\mathcal{C}_{[a,b]}$)**.** Normu spojité funkce $f\colon [a,b] \to \mathbb{R}$ lze zavést analogicky jako pro Eukleidovský prostor:

- analogie eukleidovské normy: $\|f\|_2 = \sqrt{\int_a^b f(x)^2 dx}$,
- analogie součtové normy: $\|f\|_1 = \int_a^b |f(x)| dx$,
- analogie maximové normy: $\|f\|_\infty = \max_{x\in[a,b]} |f(x)|$,
- analogie p-normy: $\|f\|_p = \left(\int_a^b |f(x)|^p dx\right)^{\frac{1}{p}}$. $\square$

Poznámka 8.15 (Rovnoběžníkové pravidlo)**.** Pro normu indukovanou skalárním součinem platí tzv. *rovnoběžníkové pravidlo:*

$$\|x-y\|^2 + \|x+y\|^2 = 2\|x\|^2 + 2\|y\|^2.$$

Důkaz. $\|x-y\|^2 + \|x+y\|^2 = \langle x-y, x-y\rangle + \langle x+y, x+y\rangle = 2\langle x,x\rangle + 2\langle y,y\rangle = 2\|x\|^2 + 2\|y\|^2.$ □

Díky tomu snadno nahlédneme, že součtová a maximová norma nejsou indukované žádným skalárním součinem. Např. pro $x = (1,0)$ a $y = (0,1)$ nesplňují rovnoběžníkové pravidlo.

Platí dokonce silnější tvrzení: Pokud pro normu platí rovnoběžníkové pravidlo, pak je indukovaná nějakým skalárním součinem; viz Horn and Johnson [1985].

Norma umožňuje zavést vzdálenost (neboli metriku) mezi vektory x, y jako $\|x-y\|$. A pokud máme vzdálenost, můžeme zavést limity, etc. Čtenáře by již nemělo překvapit, že i metriku lze zavést axiomaticky. Navíc k definici metriky nepotřebujeme ani vektorový prostor, stačí libovolná množina.

Poznámka 8.16 (Metrika)**.** Metriku na množině M definujeme jako je zobrazení $d\colon M^2 \to \mathbb{R}$, splňující:

(1) $d(x,y) \geq 0$ pro všechna $x, y \in M$, a rovnost nastane pouze pro $x = y$,

(2) $d(x,y) = d(y,x)$ pro všechna $x, y \in M$,

(3) $d(x,z) \leq d(x,y) + d(y,z)$ pro všechna $x, y, z \in M$.

Každá norma určuje metriku předpisem $d(x,y) := \|x-y\|$, tedy vzdálenost vektorů x, y zavádí jako velikost jejich rozdílu. Na druhou stranu to ale obecně neplatí. Existují prostory s metrikou, která není indukována žádnou normou, např. diskrétní metrika $d(x,y) := \lceil \|x-y\|_2 \rceil$, nebo diskrétní metrika definovaná $d(x,y) := 1$ pro $x \neq y$ a $d(x,y) := 0$ pro $x = y$.

Příklad 8.17 (Klasifikace psaných číslic)**.** Ukážeme použití vzdálenosti na vytvoření jednoduchého klasifikátoru pro automatickou identifikaci ručně psaných číslic. Tento problém je velmi důležitý mj. pro efektivitu poštovní služby. Uvádí se [Chartier, 2015], že jen v roce 2012 bylo v USA doručeno 160 miliard dopisů. Automatizace pak znamená zrychlení a zlevnění celého procesu doručování dopisů. První program pro detekci číslic byl spuštěn v roce 1997 a přestože měl úspěšnost pouze 10%, znamenal výrazný posun kupředu.

Předpokládáme, že každá číslice je zadaná jako obrázek, reprezentovaný maticí $A \in \mathbb{R}^{m\times n}$, tedy pixel obrázku na pozici (i,j) má barvu s číslem a_{ij}. Jako vzory použijeme zprůměrované hodnoty z databáze MNIST (`http://yann.lecun.com/exdb/mnist/`):

Nyní uvažujme následující obrázek, který chceme klasifikovat jako číslici:

Na prostoru matic proto musíme zavést metriku. K tomuto účelu adaptujeme klasickou eukleidovskou vzdálenost a vzdálenost matic $A, B \in \mathbb{R}^{m\times n}$ definujeme jako

$$\|A - B\| := \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n}(a_{ij} - b_{ij})^2}.$$

Pokud spočítáme vzdálenost mezi maticí reprezentující klasifikovaný obrázek a jednotlivými vzory, pak dostaneme hodnoty:

0: 1957.44	1: 2237.30	2: 2015.79	3: 1816.23	4: 1868.78
5: 1771.64	6: 2038.57	7: 2090.51	8: 1843.22	9: 1900.81

Vidíme, že nejmenší vzdálenost je k obrázku číslo 5, proto klasifikujeme daný obrázek správně jako číslo 5. Takovýto jednoduchý klasifikátor se ale snadno může splést. Náš klasifikátor nedokáže rozpoznat tvary čar (znaménko křivosti atp.), a tudíž je malá vzdálenost je i třeba k obrázku číslo 3. □

8.2 Ortonormální báze, Gramova–Schmidtova ortogonalizace

Každý vektorový prostor má bázi. U prostoru se skalárním součinem je přirozené se ptát, zda existuje báze složená z navzájem kolmých vektorů. V této sekci ukážeme, že je to pravda, že taková báze má řadu pozoruhodných vlastností a také odvodíme algoritmus na její nalezení.

Definice 8.18 (Ortogonální a ortonormální systém)**.** Systém vektorů $z_1, \dots, z_n$ je *ortogonální*, pokud $\langle z_i, z_j\rangle = 0$ pro všechna $i \neq j$. Systém vektorů $z_1, \dots, z_n$ je *ortonormální*, pokud je ortogonální a $\|z_i\| = 1$ pro všechna $i = 1, \dots, n$.

Je-li systém $z_1, \dots, z_n$ ortonormální, pak je také ortogonální. Naopak to obecně neplatí, ale není problém ortogonální systém zortonormalizovat. Jsou-li $z_1, \dots, z_n$ nenulové a ortogonální, pak $\frac{1}{\|z_1\|}z_1, \dots, \frac{1}{\|z_n\|}z_n$ je ortonormální. Důkaz: $\|\frac{1}{\|z_i\|}z_i\| = \frac{1}{\|z_i\|}\|z_i\| = 1$.

Příklad 8.19. V prostoru $\mathbb{R}^n$ se standardním skalárním součinem je ortonormálním systémem například kanonická báze $e_1, \dots, e_n$. Speciálně v rovině $\mathbb{R}^2$ tvoří normální bázi vektory $(1, 0)^T$, $(0, 1)^T$. Jiný příklad normální báze v $\mathbb{R}^2$ je například: $\frac{\sqrt{2}}{2}(1, 1)^T$, $\frac{\sqrt{2}}{2}(1, -1)^T$. □

Tvrzení 8.20. *Je-li systém vektorů $z_1, \dots, z_n$ ortonormální, pak je lineárně nezávislý.*

Důkaz. Uvažujme lineární kombinaci $\sum_{i=1}^{n} \alpha_i z_i = o$. Pak pro každé $k = 1, \dots, n$ platí:

$$0 = \langle o, z_k\rangle = \left\langle \sum_{i=1}^{n} \alpha_i z_i, z_k \right\rangle = \sum_{i=1}^{n} \alpha_i \langle z_i, z_k\rangle = \alpha_k \langle z_k, z_k\rangle = \alpha_k. \qquad \square$$

Ortonormalita vektorů tedy znamená jejich lineární nezávislost plus něco navíc – jejich kolmost. A právě tato vlastnost umožní některé problémy řešit efektivně. Zrovna následující věta říká, jak jednoduše spočítat souřadnice vůči bázi, která je ortonormální.

Věta 8.21 (Fourierovy koeficienty)**.** *Buď $z_1, \dots, z_n$ ortonormální báze prostoru V. Pak pro každé $x \in V$ platí $x = \sum_{i=1}^{n} \langle x, z_i\rangle z_i$.*

Důkaz. Víme, že $x = \sum_{i=1}^{n} \alpha_i z_i$ a souřadnice $\alpha_1, \dots, \alpha_n$ jsou jednoznačné (věta 5.31). Nyní pro každé $k = 1, \dots, n$ platí:

$$\langle x, z_k\rangle = \left\langle \sum_{i=1}^{n} \alpha_i z_i, z_k \right\rangle = \sum_{i=1}^{n} \alpha_i \langle z_i, z_k\rangle = \alpha_k \langle z_k, z_k\rangle = \alpha_k. \qquad \square$$

Vyjádření $x = \sum_{i=1}^{n} \langle x, z_i \rangle z_i$ se nazývá *Fourierův rozvoj*, a skaláry $\langle x, z_i \rangle$, $i = 1, \ldots, n$ se nazývají *Fourierovy koeficienty*[2]. Geometrický význam Fourierova koeficientu $\langle x, z_i \rangle$ je ten, že $\langle x, z_i \rangle z_i$ udává projekci vektoru x na přímku $\operatorname{span}\{z_i\}$. Jinými slovy, $\langle x, z_i \rangle z_i$ je vektor na přímce se směrnicí z_i, který je nejblíže vektoru x. Potom vektor x lze složit z těchto dílčích projekcí jednoduchým součtem $x = \sum_{i=1}^{n} \langle x, z_i \rangle z_i$ (více o projekcích budeme hovořit v sekci 8.3). Jak ilustruje obrázek dole, pokud by báze $z_1, \ldots, z_n$ nebyla ortonormální, pak by tato vlastnost už obecně neplatila.

Vektory z_1, z_2 ortonormální.

Vektory z_1, z_2 délky 1, ale ne ortogonální.

Jak sestrojit ortonormální bázi nějakého prostoru? Následující procedura, Gramova–Schmidtova ortogonalizační metoda, začne s libovolnou bází a postupným nakolmováním vektorů vytvoří bázi, která je ortonormální. Nakolmování v kroku 2 funguje tak, že od vektoru x_k odečteme jeho projekci do prostoru generovaného vektory $x_1, \ldots, x_{k-1}$; tak bude kolmý na všechny předchozí. O projekci budeme pojednávat více v sekci 8.3.

Nakolmení druhého vektoru.

Nakolmení třetího vektoru.

Algoritmus 8.22 (Gramova–Schmidtova ortogonalizace[3])**.** Buďte $x_1, \ldots, x_n \in V$ lineárně nezávislé.

1: **for** $k := 1$ **to** n **do**

2: $\quad y_k := x_k - \sum_{j=1}^{k-1} \langle x_k, z_j \rangle z_j,$ //vypočítáme kolmici

3: $\quad z_k := \frac{1}{\|y_k\|} y_k,$ //normalizujeme délku na 1

4: **end for**

Výstup: $z_1, \ldots, z_n$ ortonormální báze prostoru $\operatorname{span}\{x_1, \ldots, x_n\}$.

Důkaz. (Správnost Gramovy–Schmidtovy ortogonalizace.) Matematickou indukcí podle n dokážeme, že $z_1, \ldots, z_n$ je ortonormální báze prostoru $\operatorname{span}\{x_1, \ldots, x_n\}$. Pro $n = 1$ je $y_1 = x_1 \neq o$ a $z_1 = \frac{1}{\|x_1\|} x_1$ je dobře definované a $\operatorname{span}\{x_1\} = \operatorname{span}\{z_1\}$.

[2] Jean Baptiste Joseph Fourier (1768–1830), francouzský matematik a fyzik. Rozvoj použil kolem r. 1807 pro řešení problému vedení tepla v pevných látkách.

[3] Metoda pochází od dánského finančního matematika Jørgen Pedersen Grama z r. 1883, explicitní vzorec publikoval r. 1907 německý matematik Erhard Schmidt. Jak už to bývá, nezávisle na nich a dříve objevili postup také P.S. Laplace (1816) nebo A.L. Cauchy (1836).

Indukční krok $n \leftarrow n-1$. Předpokládejme, že $z_1, \dots, z_{n-1}$ je ortonormální báze prostoru $\operatorname{span}\{x_1, \dots, x_{n-1}\}$. Kdyby bylo $y_n = o$, tak $x_n = \sum_{j=1}^{n-1} \langle x_n, z_j \rangle z_j$ a $x_n \in \operatorname{span}\{z_1, \dots, z_{n-1}\} = \operatorname{span}\{x_1, \dots, x_{n-1}\}$, což by byl spor s lineární nezávislostí vektorů $x_1, \dots, x_n$. Proto $y_n \neq o$ a $z_n = \frac{1}{\|y_n\|} y_n$ je dobře definovaný a má jednotkovou normu.

Nyní dokážeme, že $z_1, \dots, z_n$ je ortonormální systém. Z indukčního předpokladu je $z_1, \dots, z_{n-1}$ ortonormální systém a proto $\langle z_i, z_j \rangle$ je rovno 0 pro $i \neq j$ a rovno 1 pro $i = j$. Stačí ukázat, že z_n je kolmé na ostatní z_i pro $i < n$:

$$\begin{aligned}\langle z_n, z_i \rangle &= \frac{1}{\|y_n\|} \langle y_n, z_i \rangle = \frac{1}{\|y_n\|} \left\langle x_n - \sum_{j=1}^{n-1} \langle x_n, z_j \rangle z_j, z_i \right\rangle = \\ &= \frac{1}{\|y_n\|} \langle x_n, z_i \rangle - \frac{1}{\|y_n\|} \sum_{j=1}^{n-1} \langle x_n, z_j \rangle \langle z_j, z_i \rangle = \frac{1}{\|y_n\|} \langle x_n, z_i \rangle - \frac{1}{\|y_n\|} \langle x_n, z_i \rangle = 0.\end{aligned}$$

Zbývá ověřit $\operatorname{span}\{z_1, \dots, z_n\} = \operatorname{span}\{x_1, \dots, x_n\}$. Z algoritmu je vidět, že $z_n \in \operatorname{span}\{z_1, \dots, z_{n-1}, x_n\} \subseteq \operatorname{span}\{x_1, \dots, x_n\}$, a tedy $\operatorname{span}\{z_1, \dots, z_n\} \subseteq \operatorname{span}\{x_1, \dots, x_n\}$. Protože oba prostory mají stejnou dimenzi, nastane rovnost (věta 5.46). □

Gramova–Schmidtova ortogonalizace má tu přednost, že je použitelná v každém prostoru se skalárním součinem. Speciálně při standardním skalárním součinu v $\mathbb{R}^n$ můžeme ortogonalizaci vyjádřit maticově (viz poznámka 13.9), ale na druhou stranu v tomto případě existují i jiné metody, které mají lepší numerické vlastnosti; srov. sekce 13.3.

Důsledek 8.23 (Existence ortonormální báze)**.** *Každý konečně generovaný prostor (se skalárním součinem) má ortonormální bázi.*

Důkaz. Víme (věta 5.38), že každý konečně generovaný prostor má bázi, a tu můžeme Gramovou–Schmidtovou metodou zortogonalizovat. □

Poznamenejme, že pro nekonečně generované prostory tvrzení věty neplatí – existují prostory se skalárním součinem, které nemají ortonormální bázi; viz Bečvář [2005].

Důsledek 8.24 (Rozšíření ortonormálního systému na ortonormální bázi)**.** *Každý ortonormální systém vektorů v konečně generovaném prostoru lze rozšířit na ortonormální bázi.*

Důkaz. Víme (věta 5.45), že každý ortonormální systém vektorů $z_1, \dots, z_m$ lze rozšířit na bázi $z_1, \dots, z_m$, $x_{m+1}, \dots, x_n$, a tu můžeme Gramovou–Schmidtovou metodou zortogonalizovat na $z_1, \dots, z_m, z_{m+1}, \dots, z_n$. Povšimněme si, že ortogonalizací se prvních m vektorů nezmění. □

Další užitečný vztah je Besselova nerovnost a Parsevalova rovnost.

Věta 8.25. *Buď $z_1, \dots, z_n$ ortonormální systém ve V a buď $x \in V$. Pak platí:*

(1) Besselova nerovnost: $\|x\|^2 \geq \sum_{j=1}^n |\langle x, z_j \rangle|^2$,

(2) Parsevalova rovnost: $\|x\|^2 = \sum_{j=1}^n |\langle x, z_j \rangle|^2$ právě tehdy, když $x \in \operatorname{span}\{z_1, \dots, z_n\}$.

Důkaz.

(1) Vyplývá z úpravy

$$\begin{aligned}0 &\leq \left\langle x - \sum_{j=1}^n \langle x, z_j \rangle z_j, x - \sum_{j=1}^n \langle x, z_j \rangle z_j \right\rangle = \\ &= \langle x, x \rangle - \sum_{j=1}^n \overline{\langle x, z_j \rangle} \langle x, z_j \rangle - \sum_{j=1}^n \langle x, z_j \rangle \langle z_j, x \rangle + \sum_{j=1}^n \langle x, z_j \rangle \overline{\langle x, z_j \rangle} = \\ &= \|x\|^2 - \sum_{j=1}^n |\langle x, z_j \rangle|^2.\end{aligned}$$

(2) Vyplývá z předchozího, neboť rovnost nastane právě tehdy, když $x = \sum_{j=1}^{n} \langle x, z_j \rangle z_j$. □

Parsevalova rovnost ukazuje, že pro vektory blízké počátku musí i jejich souřadnice být dostatečně malé. Dále, rovnost se dá zobecnit i pro nekonečně generované prostory jako je $\mathcal{C}_{[-\pi,\pi]}$, což mj. znamená, že Fourierovy koeficienty v nekonečném rozvoji musí konvergovat k nule.

Parsevalova rovnost také jinými slovy říká, že v jakémkoli konečně generovaném prostoru V se norma libovolného $x \in V$ dá vyjádřit jako standardní eukleidovská norma jeho vektoru souřadnic: $\|x\| = \sqrt{[x]_B^T [x]_B}$, kde B je ortonormální báze V. Jak ukážeme dole v tvrzení 8.26, tato vlastnost analogicky platí i pro skalární součin: $\langle x, y \rangle = [x]_B^T [y]_B$ pro reálný a $\langle x, y \rangle = [x]_B^T \overline{[y]}_B$ pro komplexní prostor.

Tvrzení 8.26. *Buď B ortonormální báze prostoru V a buď $x, y \in V$. Pak $\langle x, y \rangle = [x]_B^T \overline{[y]}_B$.*

Důkaz. Buď $B = \{z_1, \ldots, z_n\}$. Podle věty 8.21 je $[x]_B = (\langle x, z_1 \rangle, \ldots, \langle x, z_n \rangle)^T$. Nyní

$$\begin{aligned} \langle x, y \rangle &= \langle \textstyle\sum_{j=1}^{n} \langle x, z_j \rangle z_j, y \rangle = \sum_{j=1}^{n} \langle x, z_j \rangle \langle z_j, y \rangle \\ &= \textstyle\sum_{j=1}^{n} \langle x, z_j \rangle \overline{\langle y, z_j \rangle} = [x]_B^T \overline{[y]}_B. \end{aligned}$$ □

Není těžké nahlédnout, že tato věta platí i naopak. Čili dostáváme, že zobrazení $\langle \cdot, \cdot \rangle$ je skalárním součinem na prostoru V právě tehdy, když se dá vyjádřit jako $\langle x, y \rangle = [x]_B^T \overline{[y]}_B$ pro nějakou (či pro libovolnou) ortonormální bázi B. Každý skalární součin je tedy standardním skalárním součinem při pohledu z libovolné ortonormální báze.

8.3 Ortogonální doplněk a projekce

Ortogonální doplněk je užitečný pojem s názornou geometrickou interpretací. Ortogonální projekce je pak navíc ohromně důležitý nástroj, jehož použití v mnoha různých oborech překonává jeho základní geometrický význam.

Definice 8.27 (Ortogonální doplněk)**.** Buď V vektorový prostor a $M \subseteq V$. Pak *ortogonální doplněk* množiny M je $M^\perp := \{x \in V;\ \langle x, y \rangle = 0\ \forall y \in M\}$.

Ortogonální doplněk $M^\perp$ tedy obsahuje takové vektory x, které jsou kolmé na všechny vektory z M (někdy zkráceně říkáme, že x je kolmé na M).

Příklad 8.28. Ortogonální doplněk k vektoru $(2, 5)^T$ je přímka $\operatorname{span}\{(5, -2)^T\}$. Ortogonální doplněk k celé přímce $\operatorname{span}\{(2, 5)^T\}$ je rovněž přímka $\operatorname{span}\{(5, -2)^T\}$. □

Tvrzení 8.29 (Vlastnosti ortogonálního doplňku množiny)**.** *Buď V vektorový prostor a $M, N \subseteq V$. Pak*

(1) $M^\perp$ je podprostor V,

(2) je-li $M \subseteq N$ pak $M^\perp \supseteq N^\perp$,

(3) $M^\perp = \operatorname{span}(M)^\perp$.

Důkaz.

(1) Ověříme vlastnosti podprostoru: $o \in M^\perp$ triviálně. Nyní buďte $x_1, x_2 \in M^\perp$. Pak $\langle x_1, y \rangle = \langle x_2, y \rangle = 0\ \forall y \in M$, tedy i $\langle x_1 + x_2, y \rangle = \langle x_1, y \rangle + \langle x_2, y \rangle = 0$. Nakonec, buď $x \in M^\perp$, tedy $\langle x, y \rangle = 0\ \forall y \in M$. Pak pro každý skalár α je $\langle \alpha x, y \rangle = \alpha \langle x, y \rangle = 0$.

(2) Buď $x \in N^\perp$, tedy $\langle x, y \rangle = 0\ \forall y \in N$. Tím spíš $\langle x, y \rangle = 0\ \forall y \in M \subseteq N$, a proto $x \in M^\perp$.

(3) $M \subseteq \operatorname{span}(M)$, tedy dle předchozího je $M^\perp \supseteq \operatorname{span}(M)^\perp$. Důkaz druhé inkluze spočívá v tom, že je-li vektor x kolmý na určité vektory, pak je kolmý na jejich lineární kombinace, a tím pádem na jejich lineární obal. Formálně: buď $x \in M^\perp$ tedy $\langle x, y \rangle = 0\ \forall y \in M$. Speciálně, $\langle x, y_i \rangle = 0$, kde $y_1, \ldots, y_n \in M$ je báze $\operatorname{span}(M)$. Pak pro libovolné $y = \sum_{i=1}^{n} \alpha_i y_i \in \operatorname{span}(M)$ jest $\langle x, y \rangle = \langle x, \sum_{i=1}^{n} \alpha_i y_i \rangle = \sum_{i=1}^{n} \overline{\alpha}_i \langle x, y_i \rangle = 0$. □

Vlastnost (3) říká, že ortogonální doplněk prostoru nebo jeho báze je ten samý. To ulehčí práci pro praktické hledání ortogonálního doplňku, protože stačí ověřit kolmost jen na bázické vektory.

Zatímco předchozí věta se týkala ortogonálního doplňku libovolné množiny vektorů, nyní se zaměříme na ortogonální doplněk podprostoru. Povšimněme si, že důkaz první části je poměrně konstruktivní a dává návod jak ortogonální doplněk (resp. jeho bázi) spočítat.

Věta 8.30 (Vlastnosti ortogonálního doplňku podprostoru)**.** *Buď V vektorový prostor a $U \Subset V$. Potom platí:*

(1) Je-li $z_1, \dots, z_m$ ortonormální báze U, a je-li $z_1, \dots, z_m, z_{m+1}, \dots, z_n$ její rozšíření na ortonormální bázi V, pak $z_{m+1}, \dots, z_n$ je ortonormální báze $U^\perp$.

(2) $\dim V = \dim U + \dim U^\perp$,

(3) $V = U + U^\perp$,

(4) $(U^\perp)^\perp = U$,

(5) $U \cap U^\perp = \{o\}$.

Důkaz.

(1) Zřejmě $z_{m+1}, \dots, z_n$ je ortonormální systém v V, a tudíž stačí dokázat $\operatorname{span}\{z_{m+1}, \dots, z_n\} = U^\perp$.

Inkluze „$\supseteq$“. Každý $x \in V$ má Fourierův rozvoj $x = \sum_{i=1}^n \langle x, z_i\rangle z_i$. Je-li $x \in U^\perp$, pak $\langle x, z_i\rangle = 0$, $i = 1, \dots, m$, a tudíž $x = \sum_{i=m+1}^n \langle x, z_i\rangle z_i \in \operatorname{span}\{z_{m+1}, \dots, z_n\}$.

Inkluze „$\subseteq$“. Buď $x \in \operatorname{span}\{z_{m+1}, \dots, z_n\}$, pak $x = \sum_{i=m+1}^n \langle x, z_i\rangle z_i = \sum_{i=1}^m 0 z_i + \sum_{i=m+1}^n \langle x, z_i\rangle z_i$. Z jednoznačnosti souřadnic dostáváme $\langle x, z_i\rangle = 0$, $i = 1, \dots, m$, a tím $x \in U^\perp$.

(2) Z první vlastnosti máme $\dim V = n$, $\dim U = m$, $\dim U^\perp = n - m$.

(3) Z první vlastnosti máme $x = \underbrace{\sum_{i=1}^m \langle x, z_i\rangle z_i}_{\in U} + \underbrace{\sum_{i=m+1}^n \langle x, z_i\rangle z_i}_{\in U^\perp} \in U + U^\perp$.

(4) Z první vlastnosti je $z_{m+1}, \dots, z_n$ ortonormální báze $U^\perp$, tedy $z_1, \dots, z_m$ je ortonormální báze $(U^\perp)^\perp$.

(5) Z předchozího a podle věty 5.52 o dimenzi spojení a průniku je $\dim(U \cap U^\perp) = \dim V - \dim U - \dim U^\perp = 0$. □

Příklad 8.31. Ilustrace podprostoru U a jeho ortogonálního doplňku $U^\perp$:

□

Další z pěkných vlastností ortonormálních systémů je, že umožňují jednoduše spočítat projekci x_U vektoru x do podprostoru U, což je ten vektor z U, který je nejbližší k x.[4)] Jak ilustruje obrázek dole, a jak ukážeme formálně v následujících větách, projekce x_U je jednoznačná a je to takový vektor z U, pro který $x - x_U$ je kolmé na U, tedy $x - x_U \in U^\perp$. Díky této kolmosti se mluví o *ortogonální* projekci. Vzhledem k tomu, že jiné typy projekcí uvažovat nebudeme, nadále používáme pouze pojem *projekce*.

[4)] Pro lepší názornost projekce v $\mathbb{R}^n$ je možná lepší interpretovat vektory z $\mathbb{R}^n$ jako body. Potom představa projekce bodu $x \in \mathbb{R}^n$ do podprostoru U jako nejbližšího bodu $x_U \in U$ k bodu x je intuitivně pochopitelnější.

Definice 8.32 (Ortogonální projekce)**.** Buď V vektorový prostor a U jeho podprostor. Pak *projekcí* vektoru $x \in V$ rozumíme takový vektor $x_U \in U$, který splňuje

$$\|x - x_U\| = \min_{y \in U} \|x - y\|.$$

Následující věta ukazuje, že každý vektor má jednoznačnou projekci, a tím věta také opravňuje k zavedení projekce jakožto zobrazení $V \to U$ definované $x \mapsto x_U$. Pojem projekce tedy budeme používat jak pro zobrazení $x \mapsto x_U$, tak pro obraz x_U konkrétního vektoru x.

Věta 8.33 (O ortogonální projekci)**.** *Buď V vektorový prostor a $U \Subset V$. Pak pro každé $x \in V$ existuje právě jedna projekce $x_U \in U$ do podprostoru U. Navíc, je-li $z_1, \ldots, z_m$ ortonormální báze U, pak*

$$x_U = \sum_{i=1}^{m} \langle x, z_i \rangle z_i. \tag{8.1}$$

Důkaz. Buď $z_1, \ldots, z_m, z_{m+1}, \ldots, z_n$ rozšíření na ortonormální bázi V. Zadefinujme $x_U := \sum_{i=1}^{m} \langle x, z_i \rangle z_i \in U$ a ukážeme, že je to hledaný vektor. Nyní

$$x - x_U = \textstyle\sum_{i=1}^{n} \langle x, z_i \rangle z_i - \sum_{i=1}^{m} \langle x, z_i \rangle z_i = \sum_{i=m+1}^{n} \langle x, z_i \rangle z_i \in U^\perp. \tag{8.2}$$

Buď $y \in U$ libovolné. Nyní máme $x - x_U \in U^\perp$ a $x_U - y \in U$, viz obrázek:

Tudíž $(x - x_U) \perp (x_U - y)$ a můžeme použít Pythagorovu větu, která dává

$$\|x - y\|^2 = \|(x - x_U) + (x_U - y)\|^2 = \|x - x_U\|^2 + \|x_U - y\|^2 \geq \|x - x_U\|^2,$$

neboli $\|x - y\| \geq \|x - x_U\|$, což dokazuje minimalitu. Abychom dokázali jednoznačnost, uvědomíme si, že rovnost nastane pouze tehdy, když $\|x_U - y\|^2 = 0$, čili když $x_U = y$. □

Pokud vektor x náleží do podprostoru U, pak jeho projekcí do U je on sám, a vzoreček (8.1) odpovídá Fourierově rozvoji z věty 8.21. Rovněž je snadné nahlédnout, že náleží-li vektor x do podprostoru $U^\perp$, pak jeho projekcí do U je o. Předpis (8.1) navíc ukazuje, že zobrazení $x \mapsto x_U$, které vektor $x \in V$ zobrazuje na jeho projekci do podprostoru U, je lineárním zobrazením.

Příklad 8.34 (Projekce na přímku)**.** Buď $a \in \mathbb{R}^n$ nenulový vektor a uvažujme projekci vektoru $x \in \mathbb{R}^n$ na přímku se směrnicí a, čili projekci do podprostoru $U = \operatorname{span}\{a\}$. Ortonormální báze prostoru U je vektor $z = \frac{1}{\|a\|}a$ a podle vzorce (8.1) má projekce vektoru x tvar

$$x_U = \langle x, z\rangle z = \frac{1}{\|a\|^2}\langle x, a\rangle a = \frac{x^T a}{a^T a}a. \qquad \square$$

Poznámka 8.35. Vzhledem k vlastnostem (3) a (5) věty 8.30 se dá prostor V vyjádřit jako direktní součet podprostorů U a $U^\perp$ (poznámka 5.54). To mj. znamená, že každý vektor $v \in V$ má jednoznačné vyjádření $v = u + u'$, kde $u \in U$ a $u' \in U^\perp$. Podle věty 8.33 je navíc vektor u projekcí vektoru v do U, a vektor u' projekcí v do $U^\perp$.

Víme ze vztahu (8.2) důkazu věty 8.33, že $x - x_U \in U^\perp$. Nyní nahlédneme, že tato vlastnost je nejenom nutnou, ale i postačující podmínkou pro to, aby x_U byla projekce vektoru x.

Tvrzení 8.36. *Při značení z věty 8.33, pokud nějaké $y \in U$ splňuje $x - y \in U^\perp$, pak $y = x_U$.*

Důkaz. Protože $(x - y) \perp (y - x_U)$, použijeme Pythagorovu větu, která říká

$$\|x - x_U\|^2 = \|x - y\|^2 + \|y - x_U\|^2 \geq \|x - y\|^2.$$

Dostáváme $\|x - x_U\| \geq \|x - y\|$, tedy z vlastnosti a jednoznačnosti projekce musí $y = x_U$. $\square$

Příklad 8.37 (Legendreovy polynomy)**.** Uvažujme prostor polynomů $\mathcal{P}^n$. Jaký pro něj zavést skalární součin? První nápad je využít isomorfismu s $\mathbb{R}^{n+1}$ a pro polynomy $p(x) = a_n x^n + \ldots + a_0$, $q(x) = b_n x^n + \ldots + b_0$ zavést skalární součin jako

$$\langle p, q\rangle = \sum_{i=1}^{n} a_i b_i$$

a vektory $1, x, x^2, \ldots, x^n$ pak budou tvořit ortonormální systém. To je sice v pořádku, ale není to jediný možný způsob. Pokud si uvědomíme, že $\mathcal{P}^n$ je podprostorem prostoru spojitých funkcí $\mathcal{C}_{[a,b]}$, tak můžeme na $\mathcal{P}^n$ použít standardní skalární součin prostoru $\mathcal{C}_{[a,b]}$. Pokud zortogonalizujeme Gramovou–Schmidtovou metodou vektory $1, x, x^2, \ldots$ speciálně na $\mathcal{C}_{[-1,1]}$, pak dostaneme tzv. *Legendreovy polynomy*

$$p_0(x) = 1,\ p_1(x) = x,\ p_2(x) = \frac{1}{2}(3x^2 - 1),\ p_3(x) = \frac{1}{5}(5x^3 - 3x),\ \ldots$$

Technické detaily výpočtu přeskočíme. Ani explicitní vyjádření není příliš jednoduché, neboť n-tý člen má tvar

$$p_n(x) = 2^{-n}\sum_{k=0}^{n}\binom{n}{k}^2 (x-1)^{n-k}(x+1)^k.$$

Tyto polynomy jsou na sebe kolmé, ale z důvodu určitých aplikací jsou znormovány tak, že n-tý polynom má normu $2/(2n+1)$.

Legendreovy polynomy můžeme použít třeba k aproximaci funkce polynomem, srov. metodu v sekci 3.6. Pokud funkci f chceme aproximovat polynomem n-tého stupně, tak spočítáme projekci f do podprostoru $\mathcal{P}^n$ v tomto skalárním součinu. Projekci spočítáme podle věty 8.33 a za ortonormální bázi $\mathcal{P}^n$ použijeme Legendreovy polynomy. Výsledná projekce má třeba tu vlastnost, že ze všech polynomů stupně n je nejblíže k f v normě indukované daným skalárním součinem, což zhruba odpovídá snaze minimalizovat plochu mezi f a polynomem. Historicky byly Legendreovy polynomy poprvé použity ve fyzice k vyjádření určitých operátorů a řešení diferenciálních rovnic. $\square$

Příklad 8.38 (Ortonormální systém v prostoru funkcí)**.** V prostoru $\mathcal{C}_{[-\pi,\pi]}$ existuje spočetný ortonormální systém $z_1, z_2, \ldots$ sestávající z vektorů

$$\frac{1}{\sqrt{2\pi}},\ \frac{1}{\sqrt{\pi}}\cos x,\ \frac{1}{\sqrt{\pi}}\sin x,\ \frac{1}{\sqrt{\pi}}\cos 2x,\ \frac{1}{\sqrt{\pi}}\sin 2x,\ \frac{1}{\sqrt{\pi}}\cos 3x, \frac{1}{\sqrt{\pi}}\sin 3x,\ \ldots$$

I když to není báze v pravém slova smyslu, každou funkci $f \in \mathcal{C}_{[-\pi,\pi]}$ lze vyjádřit jako nekonečnou řadu $f(x) = \sum_{i=1}^{\infty} \langle f, z_i \rangle z_i$. Poznámka: zde trochu zjednodušujeme a odbýváme pojem nekonečného součtu, ale pro intuitivní pochopení to snad postačuje.

Vyjádření několika prvních členů $f(x) \approx \sum_{i=1}^{k} \langle f, z_i \rangle z_i$, což je vlastně projekce do prostoru $\operatorname{span}\{z_1, \dots, z_k\}$ dimenze k, dává dobrou aproximaci funkce $f(x)$. Takovéto aproximace se používá hojně v teorii zpracování signálů (např. zvuku).

Konkrétně, spočítejme Fourierův rozvoj funkce $f(x) = x$ na intervalu $[-\pi, \pi]$

$$x = a_0 + \sum_{k=1}^{\infty} (a_k \sin(kx) + b_k \cos(kx)),$$

kde

$$\begin{aligned} a_0 &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) \cdot 1 \, dx = \frac{1}{2\pi} \int_{-\pi}^{\pi} x \, dx = 0, \\ a_k &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin(kx) \, dx = \frac{1}{\pi} \int_{-\pi}^{\pi} x \sin(kx) \, dx = (-1)^{k+1} \frac{2}{k}, \\ b_k &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos(kx) \, dx = \frac{1}{\pi} \int_{-\pi}^{\pi} x \cos(kx) \, dx = 0. \end{aligned}$$

Tedy $x = \sum_{k=1}^{\infty} (-1)^{k+1} \frac{2}{k} \sin(kx)$.

□

Je možné spočítat projekci bez nutnosti mít ortonormální bázi podprostoru? Ano, pomocí řešení soustavy s tzv. Gramovou maticí.

Věta 8.39 (Gramova matice)**.** *Buď V reálný vektorový prostor a $U \Subset V$. Nechť U má bázi $B = \{w_1, \dots, w_m\}$. Označme jako Gramovu matici $G \in \mathbb{R}^{m \times m}$ matici s prvky $G_{ij} = \langle w_i, w_j \rangle$. Pak G je regulární maticí a vektor souřadnic $s = [x_U]_B$ projekce x_U libovolného vektoru $x \in V$ do podprostoru U je řešením soustavy*

$$Gs = (\langle w_1, x \rangle, \dots, \langle w_m, x \rangle)^T. \tag{8.3}$$

Důkaz. Pro důkaz regularity G buď $s \in \mathbb{R}^m$ řešení soustavy $Gs = o$. Pak i-tý řádek soustavy rovnic má tvar $\sum_{j=1}^{m} G_{ij} s_j = \langle w_i, \sum_{j=1}^{m} s_j w_j \rangle = 0$, čili $\sum_{j=1}^{m} s_j w_j \in U^{\perp} \cap U = \{o\}$. Z lineární nezávislosti $w_1, \dots, w_m$ nutně $s = o$.

Víme, že x_U existuje a je jednoznačná a lze psát ve tvaru $x_U = \sum_{j=1}^{m} \alpha_j w_j$ pro vhodné skaláry α_j. Protože $x - x_U \in U^{\perp}$, dostáváme speciálně $\langle w_i, x - x_U \rangle = 0$, pro všechna $i = 1, \dots, m$. Dosazením za x_U získáme $\langle w_i, x - \sum_{j=1}^{m} \alpha_j w_j \rangle = 0$, neboli

$$\sum_{j=1}^{m} \alpha_j \langle w_i, w_j \rangle = \langle w_i, x \rangle, \quad i = 1, \dots, m.$$

Tedy $s := [x_U]_B = (\alpha_1, \dots, \alpha_m)^T$ řeší soustavu (8.3). Z regularity G pak existuje pouze jediné řešení soustavy a odpovídá dané projekci. □

V důkazu jsme nahlédli, že Gramova matice G je regulární. Není těžké nahlédnout, že pokud by generátory $w_1, \dots, w_m$ podprostoru U byly lineárně závislé, pak by matice G byla singulární. Tudíž G je regulární právě tehdy, když vektory $w_1, \dots, w_m$ jsou lineárně nezávislé.

8.4 Ortogonální doplněk a projekce v $\mathbb{R}^n$

Z minulé sekce víme, jak počítat ortogonální doplněk a projekci pro libovolný konečně generovaný vektorový prostor se skalárním součinem, a to pomocí ortonormální báze. Nyní ukážeme, že v $\mathbb{R}^n$ pro standardní skalární součin tyto transformace lze vyjádřit explicitně a přímo bez počítání ortonormální báze.

Následující věta říká, jak spočítat ortogonální doplněk libovolného podprostoru $\mathbb{R}^n$, známe-li jeho bázi nebo konečný systém generátorů (představují řádky matice A).

Věta 8.40 (Ortogonální doplněk v $\mathbb{R}^n$)**.** *Buď $A \in \mathbb{R}^{m\times n}$. Pak $\mathcal{R}(A)^\perp = \mathrm{Ker}(A)$.*

Důkaz. Z vlastností ortogonálního doplňku (tvrzení 8.29(3)) víme $\mathcal{R}(A)^\perp = \{A_{1*}, \dots, A_{m*}\}^\perp$. Tedy $x \in \mathcal{R}(A)^\perp$ právě tehdy, když x je kolmé na řádky matice A, neboli $A_{i*}x = 0$ pro všechna $i = 1, \dots, m$. Ekvivalentně, $Ax = o$, to jest $x \in \mathrm{Ker}(A)$. □

Příklad 8.41. Buď V prostor generovaný vektory $(1, 2, 3)^T$ a $(1, -1, 0)^T$. Chceme-li určit $V^\perp$, tak sestavíme matici

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

protože $V = \mathcal{R}(A)$. Nyní již stačí nalézt bázi $V^\perp = \mathrm{Ker}(A)$, kterou tvoří např. vektor $(1, 1, -1)^T$. □

Charakterizace ortogonálního doplňku má i teoretické důsledky, např. vztah matice A a matice A^TA. Pozor, pro sloupcové prostory analogie neplatí!

Důsledek 8.42. *Buď $A \in \mathbb{R}^{m\times n}$. Pak*

(1) $\mathrm{Ker}(A^TA) = \mathrm{Ker}(A)$,

(2) $\mathcal{R}(A^TA) = \mathcal{R}(A)$,

(3) $\mathrm{rank}(A^TA) = \mathrm{rank}(A)$.

Důkaz.

(1) Je-li $x \in \mathrm{Ker}(A)$, pak $Ax = o$, a tedy také $A^TAx = A^To = o$, čímž $x \in \mathrm{Ker}(A^TA)$. Naopak, je-li $x \in \mathrm{Ker}(A^TA)$, pak $A^TAx = o$. Pronásobením x^T dostaneme $x^TA^TAx = o$, neboli $\|Ax\|^2 = o$. Z vlastnosti normy musí $Ax = o$ a tudíž $x \in \mathrm{Ker}(A)$.

(2) $\mathcal{R}(A^TA) = \mathrm{Ker}(A^TA)^\perp = \mathrm{Ker}(A)^\perp = \mathcal{R}(A)$.

(3) Triviálně z předchozího bodu. □

Nyní se podíváme na projekci vektoru x do podprostoru V, pro kterou odvodíme explicitní vzorec. Pokud vektory báze prostoru V dáme do sloupců matice A, tak projekci vektoru x do V lze formulovat jako projekci x do $\mathcal{S}(A)$.

Věta 8.43 (Ortogonální projekce v $\mathbb{R}^m$)**.** *Buď $A \in \mathbb{R}^{m\times n}$ hodnosti n. Pak projekce vektoru $x \in \mathbb{R}^m$ do sloupcového prostoru $\mathcal{S}(A)$ je $x' = A(A^TA)^{-1}A^Tx$.*

Důkaz. Nejprve si uvědomíme, že x' je dobře definované. Matice A^TA má dimenzi n (důsledek 8.42(3)), tedy je regulární a má inverzi. Podle tvrzení 8.36 stačí nyní ukázat, že $x' \in \mathcal{S}(A)$ a $x - x' \in \mathcal{S}(A)^\perp$. První vlastnost platí, neboť $x' = Az$ pro $z = (A^TA)^{-1}A^Tx$. Pro druhou vlastnost stačí ověřit, že $x - x' \in \mathcal{S}(A)^\perp = \mathcal{R}(A^T)^\perp = \mathrm{Ker}(A^T)$, a to plyne z vyjádření

$$A^T(x - x') = A^T(x - A(A^TA)^{-1}A^Tx) = A^Tx - A^TA(A^TA)^{-1}A^Tx = A^Tx - A^Tx = o. \qquad \square$$

Poznamenejme, že projekce je lineární zobrazení a podle předchozí věty je $P := A(A^TA)^{-1}A^T$ jeho matice (vzhledem ke kanonické bázi). Navíc tato matice má pozoruhodné vlastnosti. Např. je symetrická, $P^2 = P$ a regulární pouze tehdy, když $m = n$.

Speciálně, matice projekce na jednodimenzionální podprostor (přímku) má tvar $P = a(a^Ta)^{-1}a^T$, kde $a \in \mathbb{R}^n$ je směrnice přímky. Projekce vektoru x na přímku je pak vektor $Px = a(a^Ta)^{-1}a^Tx = \frac{a^Tx}{a^Ta}a$ (srov. příklad 8.34). Pokud navíc směrnici normujeme tak, aby $\|a\|_2 = 1$, potom $a^Ta = 1$ a tudíž matice projekce získá jednoduchý tvar $P = aa^T$.

Věta 8.44 (Ortogonální projekce do doplňku)**.** *Buď $P \in \mathbb{R}^{n\times n}$ matice projekce do podprostoru $V \Subset \mathbb{R}^n$. Pak $I - P$ je maticí projekce do $V^\perp$.*

Důkaz. Podle věty 8.30 lze každý vektor $x \in \mathbb{R}^n$ jednoznačně rozložit na součet $x = y + z$, kde $y \in V$ a $z \in V^\perp$. Z pohledu věty 8.33 je y projekce x do V a z projekce x do $V^\perp$. Tedy $z = x - y = x - Px = (I - P)x$. □

Příklad 8.45 (Matice projekce do $\mathrm{Ker}(A)$)**.** Věta 8.44 umožňuje elegantně vyjádřit projekci do jádra matice $A \in \mathbb{R}^{m\times n}$. Předpokládejme, že $\mathrm{rank}(A) = m$. Protože $\mathrm{Ker}(A)^\perp = \mathcal{R}(A) = \mathcal{S}(A^T)$, tak matice projekce do $\mathrm{Ker}(A)$ je dána předpisem $I - A^T(AA^T)^{-1}A$, kde $A^T(AA^T)^{-1}A$ je matice projekce do $\mathcal{S}(A^T)$. □

Poznámka 8.46 (Vzdálenosti podprostorů)**.** Jedním z elegantních využití projekcí v geometrii je určení vzdálenosti afinních podprostorů – vzdálenost bodu od přímky, vzdálenost dvou přímek, vzdálenost bodu od roviny atp. Vzdáleností dvou afinních podprostorů $U + a$, $V + b$ pak rozumíme nejmenší vzdálenost $\|x - y\|$, kde $x \in U + a$, $y \in V + b$. Bez důkazu poznamenáváme, že nejmenší vzdálenost se vždy nabyde.

Univerzální postup je následující. Mějme $U + a$, $V + b$ dva afinní podprostory prostoru $\mathbb{R}^n$, kde $U = \mathrm{span}\{u_1, \dots, u_m\}$ a $V = \mathrm{span}\{v_1, \dots, v_n\}$. Nechť nejmenší vzdálenost se nabyde pro body $x \in U + a$, $y \in V + b$; tyto body jdou vyjádřit jako $x = a + \sum_{i=1}^m \alpha_i u_i$, $y = b + \sum_{j=1}^n \beta_j v_j$. Vzdálenost těchto dvou bodů je stejná jako vzdálenost bodu a od bodu $b + \sum_{j=1}^n \beta_j v_j - \sum_{i=1}^m \alpha_i u_i$. Čili hledanou vzdálenost můžeme ekvivalentně vyjádřit jako vzdálenost bodu a od afinního podprostoru $U + V + b$. Posunutím ve směru $-b$ pak vzdálenost spočítáme jako vzdálenost bodu $a - b$ od podprostoru $U + V = \mathrm{span}\{u_1, \dots, u_m, v_1, \dots, v_n\}$. To už je standardní úloha, kterou vyřešíme pomocí věty 8.33 resp. věty 8.43 jakožto vzdálenost bodu $a - b$ od své projekce do prostoru $U + V$.

Uvažujme pro konkrétnost dvě přímky $\mathrm{span}\{u_1\} + a$, $\mathrm{span}\{v_1\} + b$, kde $u_1 = (1, 1, 4)^T$ $a = (3, 3, 3)^T$, $v_1 = (1, 0, 2)^T$, $b = (1, 2, 6)^T$. Jejich vzdálenost je tedy stejná jako vzdálenost bodu $a - b = (2, 1, -3)^T$ od $\mathrm{span}\{u_1, v_1\}$. Podle věty 8.43 je projekce bodu $a - b$ na podprostor $\mathrm{span}\{u_1, v_1\}$ rovna $(0, -1, -2)^T$. Tudíž hledaná vzdálenost je $\|(2, 1, -3)^T - (0, -1, -2)^T\| = \|(2, 2, -1)^T\| = 3$.

8.5 Metoda nejmenších čtverců

Věta o projekci má široké použití nejenom v geometrii. Uvažujme soustavu $Ax = b$, která nemá řešení (typicky, když $m \gg n$). V tom případě bychom chtěli nějakou dobrou aproximaci, tj. takový vektor x, že levá a pravá strana jsou si co nejblíže. Formálně,

$$\min_{x\in\mathbb{R}^n} \|Ax - b\|$$

Tento přístup se studuje pro různé normy, ale pro eukleidovskou dostáváme

$$\min_{x\in\mathbb{R}^n} \|Ax - b\|_2,$$

což je vzhledem k monotonii druhé mocniny ekvivalentní s úlohou

$$\min_{x\in\mathbb{R}^n} \|Ax - b\|_2^2 = \min_{x\in\mathbb{R}^n} \sum_{j=1}^n (A_{j*}x - b_j)^2.$$

Odtud název *metoda nejmenších čtverců*. S využitím věty o projekci najdeme řešení snadno. Následující věta říká, že řešení metodou nejmenších čtverců jsou zároveň řešeními soustavy rovnic

$$A^T Ax = A^T b. \tag{8.4}$$

Tato soustava se nazývá *soustava normálních rovnic*. Zajímavé je, že tuto soustavu dostaneme z původní soustavy $Ax = b$ pouhým přenásobením maticí A^T.

Věta 8.47 (Množina řešení metodou nejmenších čtverců)**.** *Buď $A \in \mathbb{R}^{m\times n}$. Pak množina přibližných řešení soustavy $Ax = b$ metodou nejmenších čtverců je neprázdná a rovna množině řešení normálních rovnic* (8.4).

Důkaz. Hledáme vlastně projekci vektoru b do podprostoru $\mathcal{S}(A)$, a tato projekce je vektor tvaru Ax, kde $x \in \mathbb{R}^n$. Podle tvrzení 8.36 je Ax projekcí právě tehdy, když $Ax - b \in \mathcal{S}(A)^\perp = \mathrm{Ker}(A^T)$. Jinými slovy, musí platit $A^T(Ax - b) = 0$, neboli $A^TAx = A^Tb$. Tato soustava má řešení, protože projekce musí existovat. □

Poznamenejme, že pokud soustava $Ax = b$ je řešitelná, potom každé její řešení je zároveň řešením metodou nejmenších čtverců, a naopak, každé řešení metodou nejmenších čtverců je skutečným řešením soustavy.

Jednoznačnost řešení nejmenších čtverců nastane, má-li matice A lineárně nezávislé sloupce. Pak totiž podle důsledku 8.42(3)) je matice A^TA regulární a soustava normálních rovnic tak má právě jedno řešení. To je typický případ.

Důsledek 8.48. *Buď $A \in \mathbb{R}^{m\times n}$ hodnosti n. Pak přibližné řešení soustavy $Ax = b$ metodou nejmenších čtverců je $x = (A^TA)^{-1}A^Tb$, a je jednoznačné.*

Metoda nejmenších čtverců[5)] má uplatnění v řadě oborů, zejména ve statistice při lineární regresi. Ta studuje chování a odhaduje budoucí vývoj různých veličin, např. globální teploty, HDP, ceny akcií či ropy v čase. Setkáme se s ní ale skoro ve všech vědních oborech. Např. ve fyzice tzv. Hookeův zákon říká, že natažení materiálu je přímo úměrné působící síle. Chceme-li odhadnout konstantu úměrnosti, provedeme velké množství pozorování a z nich vypočítáme hodnotu metodou nejmenších čtverců.

Příklad 8.49 (Lineární regrese: vývoj světové populace)**.** Data vývoje světové populace jsou následující:

rok	1950	1960	1970	1980	1990	2000
populace (mld.)	2,519	2,982	3,692	4,435	5,263	6,070

Chceme najít závislost velikosti populace na čase. Předpokládejme, že závislost je lineární. (To není vůbec samozřejmé, třeba Fibonacciho zjednodušený model růstu populace králíků byl exponenciální.) Obrázek dole, ilustrující data, v zásadě lineární závislost naznačuje. Lineární vztah popíšeme přímkou $y = px + q$, kde x je čas a y velikost populace. Neznámé parametry p, q vypočítáme. Po dosazení dat do rovnic by parametry p, q měly splňovat podmínky

$$\begin{aligned} 2,519 &= p \cdot 1950 + q \\ &\vdots \\ 6,070 &= p \cdot 2000 + q \end{aligned}$$

Přesné řešení neexistuje ale řešení metodou nejmenších čtverců je $p = 0,0724$, $q = -138,84$. Grafické znázornění závislosti a zdrojový kód pro Matlab / Octave:

```
A = [1950 1960 1970
     1980 1990 2000;
     1 1 1 1 1 1 ]';
b = [2.519 2.982 3.692
     4.435 5.263 6.070]';
x=(A'*A)\(A'*b),
x'*[2009; 1]
```

Výslednou závislost lze využít pro predikce na další rok. Odhad pro rok 2010 je 6,6943 mld. obyvatel, ve skutečnost jich bylo 6,853 mld. Ovšem pozor, má smysl vytvářet pouze krátkodobé odhady – v roce 1900 určitě nebyla velikost populace záporná. □

[5)] Metoda nejmenších čtverců byla vyvinuta Gaussem kolem roku 1801 pro astronomická pozorování. Tehdy se objevil asteroid Ceres, aby vzápětí zase zmizel. Gauss metodou popsal jeho dráhu a předpověděl kdy se znovu objeví.

8.6 Ortogonální matice

Uvažujme lineární zobrazení v prostoru $\mathbb{R}^n$. Jaké toto zobrazení (potažmo jeho matice) musí být, aby nijak nedeformovalo geometrické objekty? Otočení kolem počátku či překlopení podle osy jsou příklady takových zobrazení, ale chtěli bychom je analyzovat podrobně. Ukážeme, že tato vlastnost souvisí s tzv. ortogonálními maticemi. Ty ale mají dalekosáhlejší význam. Protože mají dobré numerické vlastnosti (viz sekce 1.3 a 3.5.1), setkáváme se s nimi často v nejrůznějších numerických algoritmech.

I v této sekci uvažujeme standardní skalární součin v $\mathbb{R}^n$ a eukleidovskou normu.

Definice 8.50 (Ortogonální a unitární matice)**.** Matice $Q \in \mathbb{R}^{n\times n}$ je *ortogonální*, pokud $Q^TQ = I_n$. Matice $Q \in \mathbb{C}^{n\times n}$ je *unitární*, pokud $\overline{Q}^TQ = I_n$.

Pojem unitární matice je zobecnění ortogonálních matic pro komplexní čísla. Nadále ale budeme vesměs pracovat jen s ortogonálními maticemi.

Věta 8.51 (Charakterizace ortogonálních matic)**.** *Buď $Q \in \mathbb{R}^{n\times n}$. Pak následující jsou ekvivalentní:*

(1) Q je ortogonální,

(2) Q je regulární a $Q^{-1} = Q^T$,

(3) $QQ^T = I_n$,

(4) Q^T je ortogonální,

(5) Q^{-1} existuje a je ortogonální,

(6) sloupce Q tvoří ortonormální bázi $\mathbb{R}^n$,

(7) řádky Q tvoří ortonormální bázi $\mathbb{R}^n$.

Důkaz. Stručně. (1)–(5): Je-li Q ortogonální, pak $Q^TQ = I$ a tedy $Q^{-1} = Q^T$; podobně naopak. Dle vlastnosti inverze máme i $QQ^T = I$, neboli $(Q^T)^TQ^T = I$, tedy Q^T je ortogonální.

(6): Z rovnosti $Q^TQ = I$ dostáváme porovnáním prvků na pozici i, j, že $\langle Q_{*i}, Q_{*j}\rangle = 1$, pokud $i = j$, a $\langle Q_{*i}, Q_{*j}\rangle = 0$, pokud $i \neq j$. Tedy sloupce Q tvoří ortonormální systém. Analogicky naopak. □

Vzhledem k vlastnosti (6) by se spíš slušelo říkat „ortonormální matice“, ale termín ortogonální matice je již zažitý.

Tvrzení 8.52 (Součin ortogonálních matic)**.** *Jsou-li $Q_1, Q_2 \in \mathbb{R}^{n\times n}$ ortogonální, pak Q_1Q_2 je ortogonální.*

Důkaz. $(Q_1Q_2)^TQ_1Q_2 = Q_2^TQ_1^TQ_1Q_2 = Q_2^TQ_2 = I_n$. □

Příklad 8.53 (Příklady ortogonálních matic)**.**

- Jednotková matice I_n, nebo k ní opačná $-I_n$.
- *Householderova matice*: $H(a) := I_n - \frac{2}{a^Ta}aa^T$, kde $o \neq a \in \mathbb{R}^n$. Její geometrický význam je následující. Nechť x' je projekce bodu x na přímku $\operatorname{span}\{a\}$, a uvažujme lineární zobrazení otočení bodu x dle přímky $\operatorname{span}\{a\}$ o úhel 180°. Pomocí věty 8.43 o projekci dostáváme, že bod x se zobrazí na vektor

$$x + 2(x' - x) = 2x' - x = 2a(a^Ta)^{-1}a^Tx - x = \left(2\frac{aa^T}{a^Ta} - I\right)x.$$

  Tedy matice otočení je $\frac{2}{a^Ta}aa^T - I$. Uvažujme nyní zrcadlení dle nadroviny s normálou a. To můžeme reprezentovat jako otočení o 180° dle a, a pak překlopení dle počátku. Tedy matice tohoto zobrazení je $I - 2\frac{aa^T}{a^Ta} = H(a)$.

Otočení kolem přímky a o 180°.

Zrcadlení dle nadroviny s normálou a.

Navíc se dá ukázat, že každou ortogonální matici řádu n lze rozložit jako součin nanejvýš n vhodných Householderových matic. Tudíž lineární zobrazení s ortogonální maticí geometricky reprezentuje složení nanejvýš n zrcadlení. Další vlastnosti Householderových matic poodkryjeme v sekci 13.1.

- *Givensova matice*[6]: Pro $n = 2$ je to matice otočení o úhel α proti směru hodinových ručiček

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

Je to tedy matice tvaru $\left(\begin{smallmatrix} c & -s \\ s & c \end{smallmatrix}\right)$, kde $c^2 + s^2 = 1$ a každá takováto matice odpovídá nějaké matici otočení. Obecně pro dimenzi n je to matice reprezentující otočení o úhel α v rovině os x_i, x_j, tedy schematicky

$$G_{i,j}(c,s) = \begin{pmatrix} I & & & & \\ & c & & -s & \\ & & I & & \\ & s & & c & \\ & & & & I \end{pmatrix}$$

Také z Givensových matic lze složit každou ortogonální matici, ale je jich potřeba v součinu až $\binom{n}{2}$ a případně navíc jedna diagonální matice s ± 1 na diagonále. Geometricky to znamená, že každé lineární zobrazení s ortogonální maticí reprezentuje složení nanejvýš $\binom{n}{2}$ jednoduchých otočení a případně jedno zrcadlení ve směru souřadných os. □

Věta 8.54 (Vlastnosti ortogonálních matic)**.** *Buď $Q \in \mathbb{R}^{n\times n}$ ortogonální. Pak:*

(1) $\langle Qx, Qy\rangle = \langle x, y\rangle$ *pro každé* $x, y \in \mathbb{R}^n$,

(2) $\|Qx\| = \|x\|$ *pro každé* $x \in \mathbb{R}^n$,

(3) $|Q_{ij}| \le 1$ *a* $|Q^{-1}_{ij}| \le 1$ *pro každé* $i, j = 1, \dots, n$,

(4) $\begin{pmatrix} 1 & o^T \\ o & Q \end{pmatrix}$ *je ortogonální matice.*

Důkaz.

(1) $\langle Qx, Qy\rangle = (Qx)^T Qy = x^T Q^T Qy = x^T I y = x^T y = \langle x, y\rangle$.

(2) $\|Qx\| = \sqrt{\langle Qx, Qx\rangle} = \sqrt{\langle x, x\rangle} = \|x\|$.

(3) Vzhledem k vlastnosti (6) z věty 8.51 je $\|Q_{*j}\| = 1$ pro každé $j = 1, \dots, n$. Tedy $1 = \|Q_{*j}\|^2 = \sum_{i=1}^n q_{ij}^2$, z čehož $q_{ij}^2 \le 1$, a proto $|q_{ij}| \le 1$. Matice Q^{-1} je ortogonální, takže pro ni tvrzení platí také.

(4) Z definice $\begin{pmatrix} 1 & o^T \\ o & Q \end{pmatrix}^T \begin{pmatrix} 1 & o^T \\ o & Q \end{pmatrix} = \begin{pmatrix} 1 & o^T \\ o & Q^T Q \end{pmatrix} = I_{n+1}$. □

[6] James Wallace Givens (1910–1993), Jr., americký matematik.

Díváme-li se na Q jako na matici příslušného lineárního zobrazení $x \mapsto Qx$, pak vlastnost (1) věty 8.54 říká, že při tomto zobrazení se zachovávají úhly, a vlastnost (2) zase říká, že se zachovávají délky. Tvrzení platí i naopak: matice zobrazení zachovávající skalární součin musí být nutně ortogonální (srov. věta 8.55) a dokonce matice zobrazení zachovávající eukleidovskou normu musí být ortogonální [Horn and Johnson, 1985]. Vlastnost (3) je zase ceněná v numerické matematice, protože Q a Q^{-1} mají omezené velikosti složek. Nejdůležitější vlastností pro numerické počítání je (2), protože při násobení s ortogonální maticí prvky (a tedy i zaokrouhlovací chyby) nemají tendenci se zvětšovat.

Na závěr ukažme některá zobecnění výše zmíněných vlastností na libovolný skalární součin a libovolné lineární zobrazení.

Věta 8.55 (Ortogonální matice a lineární zobrazení)**.** *Buďte U, V prostory nad $\mathbb{R}$ s libovolným skalárním součinem a $f\colon U \to V$ lineární zobrazení. Nechť B_1 resp. B_2 je ortonormální báze U resp. V. Pak matice zobrazení ${}_{B_2}[f]_{B_1}$ je ortogonální právě tehdy, když $\langle f(x), f(y)\rangle = \langle x, y\rangle$ pro každé $x, y \in U$.*

Důkaz. Podle tvrzení 8.26 a vlastností matice zobrazení je

$$\begin{aligned}
\langle x, y\rangle &= [x]_{B_1}^T \cdot [y]_{B_1},\\
\langle f(x), f(y)\rangle &= [f(x)]_{B_2}^T \cdot [f(y)]_{B_2} = \left({}_{B_2}[f]_{B_1} \cdot [x]_{B_1}\right)^T \cdot {}_{B_2}[f]_{B_1} \cdot [y]_{B_1} =\\
&= [x]_{B_1}^T \cdot {}_{B_2}[f]_{B_1}^T \cdot {}_{B_2}[f]_{B_1} \cdot [y]_{B_1}.
\end{aligned}$$

Tudíž, je-li ${}_{B_2}[f]_{B_1}$ ortogonální, pak $\langle f(x), f(y)\rangle = \langle x, y\rangle$. Naopak, pokud $\langle f(x), f(y)\rangle = \langle x, y\rangle$ platí pro každé $x, y \in U$, platí rovnost speciálně pro vektory jejichž souřadnice jsou jednotkové vektory. Dosadíme-li za x konkrétně i-tý vektor báze B_1 a za y pak j-tý vektor báze B_2, máme $[x]_{B_1} = e_i$, $[y]_{B_2} = e_j$ a proto

$$\begin{aligned}
(I_n)_{ij} = e_i^T e_j = [x]_{B_1}^T \cdot [y]_{B_1} = \langle x, y\rangle = \langle f(x), f(y)\rangle = [x]_{B_1}^T \cdot {}_{B_2}[f]_{B_1}^T \cdot {}_{B_2}[f]_{B_1} \cdot [y]_{B_1} =\\
= e_i^T \cdot {}_{B_2}[f]_{B_1}^T \cdot {}_{B_2}[f]_{B_1} \cdot e_j = \left({}_{B_2}[f]_{B_1}^T \cdot {}_{B_2}[f]_{B_1}\right)_{ij}.
\end{aligned}$$

Tímto po složkách dostáváme rovnost $I_n = {}_{B_2}[f]_{B_1}^T \cdot {}_{B_2}[f]_{B_1}$. □

Tvrzení 8.56 (Ortogonální matice a matice přechodu)**.** *Buď V prostor nad $\mathbb{R}$ s libovolným skalárním součinem a B_1, B_2 dvě jeho báze. Jakékoli dvě z následujících vlastností implikují tu třetí:*

(1) B_1 je ortonormální báze,

(2) B_2 je ortonormální báze,

(3) ${}_{B_2}[id]_{B_1}$ je ortogonální matice.

Důkaz.

Implikace „(1), (2) ⇒ (3)“. Plyne z věty 8.55, neboť identita zachovává skalární součin.

Implikace „(2), (3) ⇒ (1)“. Buď $B_1 = \{x_1, \dots, x_n\}$. Z definice pak sloupce ${}_{B_2}[id]_{B_1}$ tvoří vektory $[x_i]_{B_2}$, které jsou (díky ortogonalitě matice přechodu) ortonormální při standardním skalárním součinu v $\mathbb{R}^n$. Podle tvrzení 8.26 pak $\langle x_i, x_j\rangle = [x_i]_{B_2}^T [x_j]_{B_2}$, což je 0 pro $i = j$ a 1 jinak.

Implikace „(3), (1) ⇒ (2)“. Platí z předchozího ze symetrie, neboť ${}_{B_1}[id]_{B_2} = {}_{B_2}[id]_{B_1}^{-1}$. □

Problémy

8.1. *Stopa* matice $A \in \mathbb{R}^{n\times n}$ je číslo $\operatorname{trace}(A) = \sum_{i=1}^n a_{ii}$. Ukažte, že $\langle A, B\rangle := \operatorname{trace}(A^T B)$ je skalární součin na prostoru matic $\mathbb{R}^{m\times n}$.

8.2. Dokažte, že následující tvrzení platí pro reálný vektorový prostor, ale pro komplexní už obecně ne: Vektory $x, y \in V$ jsou kolmé právě tehdy, když $\|x + y\|^2 = \|x\|^2 + \|y\|^2$.

8.3. Rozhodněte, pro které vektory se Cauchyho–Schwarzova nerovnost nabyde jako rovnost a pro které jako ostrá nerovnost.

8.4. Porovnejte velikost (normu) vektoru $x \in V$ s velikostí jeho projekce do podprostoru $U \Subset V$.

8.5. Určete vzdálenost $c \in \mathbb{R}^n$ od

(a) nadroviny $a^T x = b$, kde $o \neq a \in \mathbb{R}^n$ a $b \in \mathbb{R}$.

(b) přímky $p = b + \operatorname{span}\{a\}$, $a \neq o$.

8.6. Ukažte, že projekce vektoru x do podprostoru U se dá vyjádřit jako jednoznačný bod v průniku podprostoru U a afinního podprostoru $U^\perp + x$.

8.7. Buď P matice projekce do $U \Subset \mathbb{R}^n$. Dokažte, že $\operatorname{rank}(P) = \operatorname{trace}(P)$.

8.8. Odvoďte vzorečky pro ortogonální projekci v $\mathbb{R}^n$ (věta 8.43) a pro metodu nejmenších čtverců (důsledek 8.48) pomocí Gramovy matice (věta 8.39).

8.9. Může být součet ortogonálních matic zase ortogonální matice?

8.10. Buďte U, V prostory nad $\mathbb{R}$ se skalárním součinem a $f\colon U \to V$ zobrazení zachovávající skalární součin, tj. $\langle f(x), f(y)\rangle = \langle x, y\rangle$ pro všechna $x, y \in U$. Ukažte, že f je lineární a prosté.

Shrnutí ke kapitole 8. Skalární součin

Skalární součin zavádí speciální součin dvou vektorů, kdy výsledkem je skalár. Máme-li vektorový prostor vybaven skalárním součinem, pak tento skalární součin přirozeně na prostoru definuje také normu, tedy velikost vektoru. A norma pak definuje vzdálenost vektorů jako normu jejich rozdílu. Oba pojmy potřebujeme k tomu, abychom byli schopni měřit v prostoru, ale taky třeba vyjádřit, že posloupnost vektorů konverguje.

Skalární součin dále přirozeně zavádí kolmost vektorů. Ortonormální báze je báze složená z vektorů velikosti 1 a navzájem kolmých. S takovouto bází se pak jednoduše počítají souřadnice, projekce aj. Ortonormální bázi umíme sestrojit Gramovou–Schmidtovou ortogonalizační metodou. Ačkoli jsme pojem skalárního součinu definovali abstraktně, ukázalo se, že každý skalární součin má podobu standardního skalárního součinu v souřadném systému (libovolné) ortonormální báze.

Ortogonální projekce je zobrazení, které vektor zobrazí na jemu nejbližší v daném podprostoru. Přímka, vedená od vektoru k jeho projekci, musí být kolmá na podprostoru (odtud „ortogonální“ projekce). Projekce se snadno spočítá, pokud známe ortonormální bázi podprostoru. V opačném případě použijeme maticový vzoreček. Projekce je velmi užitečný nástroj nejen v geometrii, kde nám umožňuje elegantně vyjádřit vzdálenosti různých objektů. Jako negeometrickou aplikaci jsme uvedli metodu nejmenších čtverců, počítající nejlepší přibližné řešení přeurčené soustavy rovnic.

Algebraicky jsou ortogonální matice takové matice, jejichž inverzní matice se jednoduše vyjádří jako transpozice. Ortogonální matice pak geometricky reprezentují lineární zobrazení, které nedeformují objekty – zachovávají úhly i vzdálenosti. Jedná se tedy o rotace a zrcadlení a jejich složení. Tato geometrická podstata se odráží i v numerických vlastnostech – počítání s ortogonálními maticemi je výhodné, protože zaokrouhlovací chyby se tolik neamplifikují.

Kapitola 9

Determinanty

Determinanty byly vyvinuty pro účely řešení čtvercové soustavy lineárních rovnic a dávají explicitní vzorec pro jejich řešení (viz věta 9.13). Za autora determinantu se považuje Gottfried Wilhelm Leibniz a nezávisle na něm jej objevil stejného roku 1683 japonský matematik Seki Kōwa. Jejich přístup (v trochu jiné podobě než uvádíme v definici) upadl trochu v zapomnění a determinanty se staly populární pro řešení soustav rovnic až tak v letech 1750–1900, pak je vystřídaly jiné metody. Nicméně se ukázalo, že determinant sám o sobě je důležitá charakteristika čtvercové matice s řadou uplatnění. Samotný pojem „determinant“ pochází od Gausse (*Disquisitiones arithmeticae*, 1801), i když jej používal v trochu jiném smyslu. Významnou měrou do teorie determinantů přispěli také mj. A.L. Cauchy či C.G.J. Jacobi.

Připomeňme, že S_n značí množinu všech permutací na množině $\{1, \dots, n\}$, viz sekce 4.2.

Definice 9.1 (Determinant)**.** Buď $A \in \mathbb{T}^{n\times n}$. Pak *determinant* matice A je číslo

$$\det(A) = \sum_{p\in S_n} \operatorname{sgn}(p) \prod_{i=1}^{n} a_{i,p(i)} = \sum_{p\in S_n} \operatorname{sgn}(p) a_{1,p(1)} \dots a_{n,p(n)}.$$

Značení: $\det(A)$ nebo $|A|$.

Co vlastně říká vzoreček z definice determinantu? Každý sčítanec má tvar $\operatorname{sgn}(p)a_{1,p(1)} \dots a_{n,p(n)}$, což odpovídá tomu, že v matici A vybereme n prvků tak, že z každého řádku a sloupce máme právě jeden. Tyto prvky pak mezi sebou vynásobíme a ještě sčítanci přiřadíme kladné či záporné znaménko podle toho, jaké bylo znaménko permutace, která tyto prvky určovala.

Příklad 9.2 (Příklady determinantů)**.** Matice řádu 2 má determinant

$$\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11}a_{22} - a_{21}a_{12}.$$

Matice řádu 3 má determinant

$$\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{cases} a_{11}a_{22}a_{33} + a_{21}a_{32}a_{13} + a_{31}a_{12}a_{23} \\ -a_{31}a_{22}a_{13} - a_{11}a_{32}a_{23} - a_{21}a_{12}a_{33}. \end{cases}$$

Počítat determinanty z definice pro větší matice je obecně značně neefektivní, protože vyžaduje zpracovat $n!$ sčítanců. Výpočet je jednodušší jen pro speciální matice. Například, determinant trojúhelníkové matice je součin diagonálních prvků, protože v definici determinantu je jediný potenciálně nenulový člen sumy pro permutaci $p = id$, která odpovídá výběru diagonálních prvků. Speciálně, $\det(I_n) = 1$. □

Tvrzení 9.3 (Determinant transpozice)**.** *Buď $A \in \mathbb{T}^{n\times n}$. Pak* $\det(A^T) = \det(A)$.

Důkaz.

$$\det(A^T) = \sum_{p\in S_n} \operatorname{sgn}(p) \prod_{i=1}^{n} A^T_{i,p(i)} = \sum_{p\in S_n} \operatorname{sgn}(p) \prod_{i=1}^{n} a_{p(i),i} = \sum_{p\in S_n} \operatorname{sgn}(p^{-1}) \prod_{i=1}^{n} a_{i,p^{-1}(i)} =$$

$$= \sum_{p^{-1}\in S_n} \operatorname{sgn}(p^{-1}) \prod_{i=1}^{n} a_{i,p^{-1}(i)} = \sum_{q\in S_n} \operatorname{sgn}(q) \prod_{i=1}^{n} a_{i,q(i)} = \det(A).$$ □

Pro determinanty obecně $\det(A+B) \neq \det(A) + \det(B)$, ani není znám jednoduchý vzoreček na determinant součtu matic. Výjimkou je následující speciální případ řádkové linearity.

Věta 9.4 (Řádková linearita determinantu)**.** *Buď $A \in \mathbb{T}^{n\times n}$ a $b \in \mathbb{T}^n$. Pak pro libovolné $i = 1, \dots, n$ platí:*

$$\det(A + e_i b^T) = \det(A) + \det(A + e_i(b^T - A_{i*})).$$

Jinými slovy,

$$\det\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{i1}+b_1 & \dots & a_{in}+b_n \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} = \det\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{i1} & \dots & a_{in} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} + \det\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ b_1 & \dots & b_n \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix}.$$

Důkaz.

$$\begin{aligned}\det(A + e_i b^T) &= \sum_{p\in S_n} \operatorname{sgn}(p) a_{1,p(1)} \dots (a_{i,p(i)} + b_{p(i)}) \dots a_{n,p(n)} = \\ &= \sum_{p\in S_n} \operatorname{sgn}(p) a_{1,p(1)} \dots a_{i,p(i)} \dots a_{n,p(n)} + \sum_{p\in S_n} \operatorname{sgn}(p) a_{1,p(1)} \dots b_{p(i)} \dots a_{n,p(n)} = \\ &= \det(A) + \det(A + e_i(b^T - A_{i*}))\end{aligned}$$

□

Vzhledem k tvrzení 9.3 je determinant nejen řádkově, ale i sloupcově lineární.

9.1 Determinant a elementární úpravy

Naším plánem je k výpočtu determinantu využít Gaussovu eliminaci. K tomu musíme nejprve umět spočítat determinant matice v odstupňovaném tvaru, a vědět, jak hodnotu determinantu ovlivňují elementární řádkové úpravy. Na první otázku je jednoduchá odpověď, protože matice v odstupňovaném tvaru je zároveň horní trojúhelníková, a tudíž je její determinant roven součinu diagonálních prvků. Druhou otázku zodpovíme rozborem jednotlivých elementárních úprav. Nechť matice A' vznikne z A nějakou elementární úpravou:

1. Vynásobení i-tého řádku číslem $\alpha \in \mathbb{T}$: $\det(A') = \alpha \det(A)$.

   Důkaz.

$$\begin{aligned}\det(A') &= \sum_{p\in S_n} \operatorname{sgn}(p) a'_{1,p(1)} \dots a'_{i,p(i)} \dots a'_{n,p(n)} = \\ &= \sum_{p\in S_n} \operatorname{sgn}(p) a_{1,p(1)} \dots (\alpha a_{i,p(i)}) \dots a_{n,p(n)} = \\ &= \alpha \sum_{p\in S_n} \operatorname{sgn}(p) a_{1,p(1)} \dots a_{i,p(i)} \dots a_{n,p(n)} = \alpha \det(A).\end{aligned}$$

   □

2. Výměna i-tého a j-tého řádku: $\det(A') = -\det(A)$.

   Důkaz. Označme transpozici $t = (i, j)$, pak

$$\det(A') = \sum_{p\in S_n} \operatorname{sgn}(p) a'_{1,p(1)} \dots a'_{i,p(i)} \dots a'_{j,p(j)} \dots a'_{n,p(n)},$$

kde $a'_{1,p(1)} = a_{1,p(1)} = a_{1,p\circ t(1)}$, $a'_{i,p(i)} = a_{j,p(i)} = a_{j,p\circ t(j)}$, atd. Tedy

$$\begin{aligned}\det(A') &= \sum_{p\in S_n} \operatorname{sgn}(p) a_{1,p\circ t(1)} \dots a_{j,p\circ t(j)} \dots a_{i,p\circ t(i)} \dots a_{n,p\circ t(n)} = \\ &= \sum_{p\in S_n} \operatorname{sgn}(p) \prod_{i=1}^{n} a_{i,p\circ t(i)} = -\sum_{p\circ t\in S_n} \operatorname{sgn}(p\circ t) \prod_{i=1}^{n} a_{i,p\circ t(i)} = \\ &= -\sum_{q\in S_n} \operatorname{sgn}(q) \prod_{i=1}^{n} a_{i,q(i)} = -\det(A).\end{aligned}$$

□

Důsledek 9.5. *Pokud má matice $A \in \mathbb{T}^{n\times n}$ dva stejné řádky, pak* $\det(A) = 0$.

Důkaz. Prohozením těchto dvou řádků dostaneme $\det(A) = -\det(A)$, a tedy $\det(A) = 0$. □

Poznamenejme, že toto tvrzení platí nad jakýmkoli tělesem, ale pro tělesa charakteristiky 2 se musí použít jiný důkaz, protože např. v $\mathbb{Z}_2$ je $-1 = 1$.

Důkaz č. 2. Definujme transpozici $t := (i, j)$, kde i, j jsou indexy stejných řádků. Nechť S'_n je množina sudých permutací z S_n. Pak S_n lze disjunktně rozložit na sjednocení S'_n a $\{p \circ t;\ p \in S'_n\}$. Tudíž

$$\begin{aligned}\det(A) &= \sum_{p\in S'_n} \operatorname{sgn}(p) \prod_{i=1}^{n} a_{i,p(i)} + \sum_{p\in S'_n} \operatorname{sgn}(p\circ t) \prod_{i=1}^{n} a_{i,p\circ t(i)} = \\ &= \sum_{p\in S'_n} \operatorname{sgn}(p) \prod_{i=1}^{n} a_{i,p(i)} - \sum_{p\in S'_n} \operatorname{sgn}(p) \prod_{i=1}^{n} a_{i,p(i)} = 0.\end{aligned}$$

□

3. Přičtení α-násobku j-tého řádku k i-tému, přičemž $i \neq j$: $\det(A') = \det(A)$.

Důkaz. Z řádkové linearity determinantu, důsledku 9.5 a první elementární úpravy dostáváme

$$\det(A') = \det\begin{pmatrix} A_{1*} \\ \vdots \\ A_{i*} + \alpha A_{j*} \\ \vdots \\ A_{n*} \end{pmatrix} = \det(A) + \det\begin{pmatrix} A_{1*} \\ \vdots \\ \alpha A_{j*} \\ \vdots \\ A_{n*} \end{pmatrix} = \det(A) + \alpha 0 = \det(A).$$

□

Výše zmíněná pozorování mají několik důsledků: Pro libovolnou matici $A \in \mathbb{T}^{n\times n}$ je $\det(\alpha A) = \alpha^n \det(A)$. Dále, obsahuje-li A nulový řádek nebo sloupec, tak $\det(A) = 0$.

Hlavní význam vlivu elementárních úprav na determinant je, že determinanty můžeme počítat pomocí Gaussovy eliminace:

Algoritmus 9.6 (Výpočet determinantu pomocí REF)**.** Převeď matici A do odstupňovaného tvaru A' a pamatuj si případné změny determinantu v koeficientu c; pak $\det(A)$ je roven součinu c^{-1} a diagonálních prvků matice A'.

9.2 Další vlastnosti determinantu

Věta 9.7 (Kriterium regularity)**.** *Matice $A \in \mathbb{T}^{n\times n}$ je regulární právě tehdy, když* $\det(A) \neq 0$.

Důkaz. Převedeme matici A elementárními úpravami na odstupňovaný tvar A', ty mohou měnit hodnotu determinantu, ale nikoli jeho (ne)nulovost. Pak A je regulární právě tehdy, když A' má na diagonále nenulová čísla. □

Poznámka 9.8 (Míra regularity)**.** Věta 9.7 umožňuje zavést jakousi míru regularity. Čím je $\det(A)$ blíže k 0, tím je matice A blíž k nějaké singulární matici. Příkladem je Hilbertova matice H_n (viz příklad 3.46), která je špatně podmíněná, protože je „skoro“ singulární. Skutečně, jak ukazuje tabulka, determinant matice je velmi blízko nule.

n	$\det(H_n)$
4	$\approx 10^{-7}$
6	$\approx 10^{-18}$
8	$\approx 10^{-33}$
10	$\approx 10^{-53}$

Tato míra není ale ideální (lepší je např. pomocí vlastních nebo singulárních čísel, viz sekce 13.5), protože je hodně citlivá ke škálování. Uvažujme např. matici $0.1I_n$, pro níž $\det(0.1I_n) = 10^{-n}$. Přestože 10^{-n} může být libovolně malé číslo, samotná matice má k singulární relativně daleko.

Věta 9.9 (Multiplikativnost determinantu)**.** *Pro každé $A, B \in \mathbb{T}^{n\times n}$ platí* $\det(AB) = \det(A)\det(B)$.

Důkaz. (1) Nejprve uvažujme speciální případ, když A je matice elementární úpravy:

1. $A = E_i(\alpha)$, vynásobení i-tého řádku číslem α. Potom $\det(AB) = \alpha\det(B)$ a $\det(A)\det(B) = \alpha\det(B)$.
2. $A = E_{ij}$, prohození i-tého a j-tého řádku. Pak $\det(AB) = -\det(B)$ a $\det(A)\det(B) = -1\det(B)$.
3. $A = E_{ij}(\alpha)$, přičtení α-násobku j-tého řádku k i-tému. Pak $\det(AB) = \det(B)$ a $\det(A)\det(B) = 1\det(B)$.

Tedy rovnost platí ve všech případech.

(2) Nyní uvažme obecný případ. Je-li A singulární, pak i AB je singulární (tvrzení 3.27) a tedy podle věty 9.7 je $\det(AB) = 0 = 0\det(B) = \det(A)\det(B)$. Je-li A regulární, pak jde rozložit na součin elementárních matic $A = E_1 \dots E_k$. Nyní postupujme matematickou indukcí, případ $k = 1$ máme vyřešený v bodě (1), takže se věnujme indukčnímu kroku. Podle indukčního předpokladu a z bodu (1) dostáváme

$$\begin{aligned}\det(AB) &= \det(E_1(E_2\dots E_kB)) = \det(E_1)\det((E_2\dots E_k)B) = \\ &= \det(E_1)\det(E_2\dots E_k)\det(B) = \det(E_1E_2\dots E_k)\det(B) = \\ &= \det(A)\det(B).\end{aligned}$$

□

Důsledek 9.10. *Buď $A \in \mathbb{T}^{n\times n}$ regulární, pak* $\det(A^{-1}) = \det(A)^{-1}$.

Důkaz. $1 = \det(I_n) = \det(AA^{-1}) = \det(A)\det(A^{-1})$. □

Nyní ukážeme rekurentní vzoreček na výpočet determinantu.

Věta 9.11 (Laplaceův rozvoj podle i-tého řádku)**.** *Buď $A \in \mathbb{T}^{n\times n}$, $n \geq 2$. Pak pro každé $i = 1, \dots, n$ platí*

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

kde A^{ij} je matice vzniklá z A vyškrtnutím i-tého řádku a j-tého sloupce.

Poznámka. Podobně jako podle řádku můžeme rozvíjet podle libovolného sloupce.

Důkaz. (1) Nejprve uvažujme případ $A_{i*} = e_j^T$, tj. i-tý řádek matice A je jednotkový vektor. Postupným vyměňováním řádků $(i, i+1)$, $(i+1, i+2)$, $\dots$, $(n-1, n)$ převedeme jednotkový vektor do posledního řádku. Podobně postupujeme pro sloupce a j-tý sloupec převedeme na poslední. Výslednou matici označme

$$A' := \left(\begin{array}{ccc|c} & A^{ij} & & \\ \hline 0 & \dots & 0 & 1 \end{array}\right)$$

a znaménko determinantu se změní koeficientem $(-1)^{(n-i)+(n-j)} = (-1)^{i+j}$. Nyní máme

$$\det(A) = (-1)^{i+j}\det(A') = (-1)^{i+j}\sum_{p\in S_n}\operatorname{sgn}(p)\prod_{i=1}^{n} a'_{i,p(i)} =$$
$$= (-1)^{i+j}\sum_{p;\,p(n)=n}\operatorname{sgn}(p)\prod_{i=1}^{n-1} a'_{i,p(i)} = (-1)^{i+j}\det(A^{ij}).$$

(2) Nyní uvažme obecný případ. Z řádkové linearity determinantu a z předchozího dostáváme

$$\det(A) = \det\begin{pmatrix}\dots & & & \\ a_{i1} & 0 & \dots & 0\\ \dots & & & \end{pmatrix} + \dots + \det\begin{pmatrix}\dots & & & \\ 0 & \dots & 0 & a_{in}\\ \dots & & & \end{pmatrix} =$$
$$= a_{i1}(-1)^{i+1}\det(A^{i1}) + \dots + a_{in}(-1)^{i+n}\det(A^{in}).$$

□

Příklad 9.12 (Laplaceův rozvoj podle 4. řádku)**.**

$$\begin{vmatrix}1&2&3&4\\1&2&1&2\\2&5&5&5\\0&2&-4&-4\end{vmatrix} = (-1)^{4+1}\cdot 0\cdot\begin{vmatrix}2&3&4\\2&1&2\\5&5&5\end{vmatrix} + (-1)^{4+2}\cdot 2\cdot\begin{vmatrix}1&3&4\\1&1&2\\2&5&5\end{vmatrix} +$$
$$+ (-1)^{4+3}\cdot(-4)\cdot\begin{vmatrix}1&2&4\\1&2&2\\2&5&5\end{vmatrix} + (-1)^{4+4}\cdot(-4)\cdot\begin{vmatrix}1&2&3\\1&2&1\\2&5&5\end{vmatrix} =$$
$$= 0 + 2\cdot 4 + 4\cdot 2 - 4\cdot 2 = 8$$

□

Následující věta dává explicitní vzoreček na řešení soustavy s regulární maticí. Matice $A + (b - A_{*i})e_i^T$ ve výrazu představuje matici A, ve které nahradíme i-tý sloupec vektorem b.

Věta 9.13 (Cramerovo pravidlo)**.** *Buď $A \in \mathbb{T}^{n\times n}$ regulární, $b \in \mathbb{T}^n$. Pak řešení soustavy $Ax = b$ je dáno vzorcem*

$$x_i = \frac{\det(A + (b - A_{*i})e_i^T)}{\det(A)}, \quad i = 1, \dots, n.$$

Důkaz. Buď x řešení soustavy $Ax = b$; díky regularitě A řešení existuje a je jednoznačné. Rovnost rozepíšeme $\sum_{j=1}^n A_{*j}x_j = b$. Ze sloupcové linearity determinantu dostaneme

$$\det(A + (b - A_{*i})e_i^T) = \det(A_{*1}|\dots|b|\dots|A_{*n}) = \det(A_{*1}|\dots|\textstyle\sum_{j=1}^n A_{*j}x_j|\dots|A_{*n}) =$$
$$= \sum_{j=1}^n \det(A_{*1}|\dots|A_{*j}|\dots|A_{*n})x_j =$$
$$= \det(A_{*1}|\dots|A_{*i}|\dots|A_{*n})x_i = \det(A)x_i.$$

Nyní stačí obě strany podělit číslem $\det(A) \neq 0$. □

Cramerovo pravidlo z roku 1750 (i když bylo známo už dříve) je pojmenováno po švýcarském matematikovi Gabrielu Cramerovi. Ve své době to byl populární nástroj na řešení soustav lineárních rovnic, ale dnes se pro praktické výpočty již nepoužívá, protože výpočet řešení soustavy pomocí $n + 1$ determinantů není příliš efektivní. Význam determinantu je spíše teoretický, mimo jiné ukazuje a dává:

- Explicitní vyjádření řešení soustavy lineárních rovnic.
- Spojitost řešení vzhledem k prvkům matice A a vektoru b.

  Formálně, zobrazení $(A, b) \mapsto A^{-1}b$ je spojité na definičním oboru regulárních matic A. To je snadné nahlédnout, protože podle Cramerova pravidla počítáme jednotlivé složky řešení pouze pomocí aritmetických operací.

- Odhad velikosti popisu řešení z velikosti popisu vstupních hodnot.

  Potřebujeme-li k zápisu vstupních hodnot soustavy $Ax = b$ (tj., k zápisu a_{ij}, b_j) celkem K bitů, tak její řešení má zápis pomocí polynomiálně mnoha bitů vzhledem ke K. To dá trochu více práce nahlédnout, ale je to důležité pozorování, protože nám zaručuje omezenou velikost zápisu řešení.

Příklad 9.14 (Cramerovo pravidlo)**.** Řešení soustavy rovnic

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

spočítáme po složkách

$$x_1 = \frac{\begin{vmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \\ 4 & 5 & 5 \end{vmatrix}}{\begin{vmatrix} 1 & 2 & 3 \\ 1 & 2 & 1 \\ 2 & 5 & 5 \end{vmatrix}} = \frac{4}{2} = 2,\ x_2 = \frac{\begin{vmatrix} 1 & 1 & 3 \\ 1 & 3 & 1 \\ 2 & 4 & 5 \end{vmatrix}}{\begin{vmatrix} 1 & 2 & 3 \\ 1 & 2 & 1 \\ 2 & 5 & 5 \end{vmatrix}} = \frac{2}{2} = 1,\ x_3 = \frac{\begin{vmatrix} 1 & 2 & 1 \\ 1 & 2 & 3 \\ 2 & 5 & 4 \end{vmatrix}}{\begin{vmatrix} 1 & 2 & 3 \\ 1 & 2 & 1 \\ 2 & 5 & 5 \end{vmatrix}} = \frac{-2}{2} = -1,$$

□

9.3 Adjungovaná matice

Adjungovaná matice[1)] úzce souvisí s determinanty a maticovou inverzí. Využijeme ji při odvozování Cayleyho–Hamiltonovy věty (věta 10.18), ale čtenář se s ní může potkat např. v kryptografii nebo při odvozování vzorečku pro derivaci determinantu.

Definice 9.15 (Adjungovaná matice)**.** Buď $A \in \mathbb{T}^{n\times n}$ a $n \geq 2$. Pak *adjungovaná matice* $\operatorname{adj}(A) \in \mathbb{T}^{n\times n}$ má složky

$$\operatorname{adj}(A)_{ij} = (-1)^{i+j} \det(A^{ji}), \quad i, j = 1, \ldots, n,$$

kde A^{ji} opět značí matici vzniklou z A vyškrtnutím j-tého řádku a i-tého sloupce.

Věta 9.16 (O adjungované matici)**.** *Pro každou matici $A \in \mathbb{T}^{n\times n}$ platí $A \operatorname{adj}(A) = \det(A) I_n$.*

Důkaz. Odvodíme

$$(A \operatorname{adj}(A))_{ij} = \sum_{k=1}^{n} A_{ik} \operatorname{adj}(A)_{kj} = \sum_{k=1}^{n} A_{ik} (-1)^{k+j} \det(A^{jk}) = \begin{cases} \det(A), & \text{pro } i = j, \\ 0, & \text{pro } i \neq j. \end{cases}$$

Zdůvodnění poslední rovnosti je, že pro $i = j$ se jedná o Laplaceův rozvoj $\det(A)$ podle j-tého řádku. Pro $i \neq j$ se zase jedná o rozvoj podle j-tého řádku matice A, v níž ale nejprve j-tý řádek nahradíme i-tým. Tato matice bude mít dva stejné řádky a tím pádem nulový determinant. □

Pro regulární matici A je $\det(A) \neq 0$ a vydělením $\det(A)$ dostaneme explicitní vzoreček pro inverzní matici A^{-1}.

Důsledek 9.17. *Je-li $A \in \mathbb{T}^{n\times n}$ regulární, pak $A^{-1} = \frac{1}{\det(A)} \operatorname{adj}(A)$.*

Příklad 9.18 (Adjungovaná matice)**.** Buď

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

Pak:

$$\operatorname{adj}(A)_{12} = (-1)^{1+2} \begin{vmatrix} 2 & 3 \\ 5 & 5 \end{vmatrix} = 5, \ \ldots$$

[1)] V angličtině *adjugate*, popř. *adjoint*.

Celkem:

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

Tedy:

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

□

Příklad 9.19. Uvažujme determinant jako funkci $\mathbb{R}^{n\times n} \to \mathbb{R}$. Problém nyní zní určit parciální derivaci $\det(A)$ podle a_{ij} a sestavit matici parciálních derivací.

Pro tento účel vyjdeme z Laplaceova rozvoje $\det(A) = \sum_{k=1}^{n}(-1)^{i+k}a_{ik}\det(A^{ik})$ a jednoduše odvodíme $\frac{\partial \det(A)}{\partial a_{ij}} = (-1)^{i+j}\det(A^{ij})$. Tudíž matice parciálních derivací je $\partial \det(A) = \operatorname{adj}(A)^T$.

Použijeme-li konkrétně matici z příkladu 9.18, tak $\det(A) = 2$. Protože $\operatorname{adj}(A)_{33} = 0$, tak determinant matice A se nezmění při změně prvku a_{33}. Na druhou stranu, při malém zvětšení prvku a_{11} se determinant zvětší výrazně (protože $\operatorname{adj}(A)_{11} = 5$) a při zvětšení prvku a_{13} se determinant zvětší méně výrazně (protože $\operatorname{adj}(A)_{31} = 1$). Vyzkoušejte sami! □

9.4 Aplikace

Determinant se používá např. v teorii grafů pro vyjádření počtu koster grafu [Matoušek a Nešetřil, 2009]. Věta o adjungované matici zase dává následující charakterizaci celočíselnosti inverzní matice.

Tvrzení 9.20. *Buď $A \in \mathbb{Z}^{n\times n}$. Pak A^{-1} má celočíselné hodnoty právě tehdy, když* $\det(A) = \pm 1$.

Důkaz. Implikace „$\Rightarrow$“. Víme $1 = \det(A)\det(A^{-1})$. Jsou-li matice A, A^{-1} celočíselné, pak i jejich determinanty jsou celočíselné a tudíž musejí být rovny ± 1.

Implikace „$\Leftarrow$“. Víme $A^{-1}_{ij} = \frac{1}{\det(A)}(-1)^{i+j}\det(A^{ji})$. To je celočíselná hodnota, jestliže $\det(A) = \pm 1$ a $\det(A^{ji})$ je celé číslo. □

Další ukázka použití determinantu je v polynomech. Determinant z následující věty se nazývá *rezultant* a používá se například k řešení nelineárních rovnic.

Tvrzení 9.21. *Polynomy $p(x) = a_n x^n + \ldots + a_1 x + a_0$, $q(x) = b_m x^m + \ldots + b_1 x + b_0$ mají společný kořen právě tehdy, když*

$$\begin{vmatrix} a_n & a_{n-1} & \ldots & a_0 & & & \\ & a_n & a_{n-1} & \ldots & a_0 & & \\ & & \ddots & \ddots & & \ddots & \\ & & & a_n & a_{n-1} & \ldots & a_0 \\ b_m & b_{m-1} & \ldots & b_1 & b_0 & & \\ & \ddots & \ddots & & & \ddots & \\ & & b_m & b_{m-1} & \ldots & b_1 & b_0 \end{vmatrix} = 0.$$

Důkaz. Polynomy $p(x)$, $q(x)$ mají společný kořen právě tehdy, když existují (ne oba triviální) polynomy $r(x)$, $s(x)$ stupňů nanejvýš $m-1, n-1$ takové, že $r(x)p(x) + s(x)q(x) = 0$. Pokud si neznámé polynomy vyjádříme jako $r(x) = c_{m-1}x^{m-1} + \ldots + c_1 x + c_0$, $s(x) = d_{n-1}x^{n-1} + \ldots + d_1 x + d_0$, tak rovnost $r(x)p(x) + s(x)q(x) = 0$ můžeme přepsat jako homogenní soustavu $m+n$ rovnic s $m+n$ neznámými $c_{m-1}, \ldots, c_0, d_{n-1}, \ldots, d_0$. Nenulové řešení existuje právě tehdy, když je matice singulární, neboli její determinant nulový. Ve větě je pak determinant transponované matice. □

Geometrická interpretace determinantu

Determinant má pěkný geometrický význam. Uvažujeme-li lineární zobrazení $x \mapsto Ax$ s maticí $A \in \mathbb{R}^{n\times n}$, pak geometrická tělesa mění v tomto zobrazení svůj objem s koeficientem $|\det(A)|$. Uvažujeme nejprve speciální případ rovnoběžnostěnu. *Rovnoběžnostěn* s lineárně nezávislými hranami $a_1, \dots, a_m$ definujeme jako množinu $\{x \in \mathbb{R}^n;\ x = \sum_{i=1}^m \alpha_i a_i,\ 0 \le \alpha_i \le 1\}$.

Věta 9.22 (Objem rovnoběžnostěnu)**.** *Buď $A \in \mathbb{R}^{m\times n}$ a uvažujme rovnoběžnostěn s hranami danými řádky matice A. Pak jeho objem je $\sqrt{\det(AA^T)}$. Speciálně, pro $m = n$ je objem $|\det(A)|$.*

Poznámka. Svou roli hraje nejen velikost determinantu A, ale také jeho znaménko; to souvisí s pořadím hran rovnoběžnostěnu jako řádků matice A. Speciálně, pro $A \in \mathbb{R}^{3\times 3}$ je $\det(A) > 0$ pokud řádky A tvoří pravotočivou posloupnost vektorů (tzv. pravidlo palce), a $\det(A) < 0$ pokud tvoří levotočivou posloupnost.

Důkaz. Důkaz provedeme matematickou indukcí podle m. Pro $m = 1$ je to zřejmé, postupme k indukčnímu kroku. Označme i-tý řádek matice A jako a_i a definujme matici

$$D := \begin{pmatrix} a_1^T \\ \vdots \\ a_{m-1}^T \end{pmatrix},$$

která vznikne z A odstraněním posledního řádku. Rozložme $a_m = b_m + c_m$, kde $c_m \in \mathcal{R}(D)$ a $b_m \in \mathcal{R}(D)^\perp$ podle poznámky 8.35. Označme

$$A' := \begin{pmatrix} a_1^T \\ \vdots \\ a_{m-1}^T \\ b_m^T \end{pmatrix}.$$

Nyní řádky matice D generují menší rovnoběžnostěn, jenž tvoří podstavu celkového rovnoběžnostěnu a nalézá se podprostoru $\mathcal{R}(D)$. Vektor b_m je kolmý na podstavu a jeho délka odpovídá výšce $\|b_m\|$ rovnoběžnostěnu.

Rovnoběžnostěn $a_1, a_2, \dots, a_m$.

Plocha základny: $\sqrt{\det(DD^T)}$, výška: $\|b_m\|$.

Od A' k A lze přejít pomocí elementárních řádkových úprav, neboť k poslednímu řádku stačí přičíst c_m, což je lineární kombinace $a_1, \dots, a_{m-1}$. Tedy existují elementární matice $E_1, \dots, E_k$ tak, že $A = E_1 \dots E_k A'$; navíc jejich determinant je 1 protože jen přičítají násobek řádku k jinému. Nyní

$$\begin{aligned}\det(AA^T) &= \det(E_1 \dots E_k A' A'^T E_k^T \dots E_1^T) = \\ &= \det(E_k) \dots \det(E_1) \det(A'A'^T) \det(E_k^T) \dots \det(E_1^T) = \det(A'A'^T).\end{aligned}$$

Dále,

$$A'A'^T = \begin{pmatrix} D \\ b_m^T \end{pmatrix} \begin{pmatrix} D^T & b_m \end{pmatrix} = \begin{pmatrix} DD^T & Db_m \\ b_m^T D^T & b_m^T b_m \end{pmatrix} = \begin{pmatrix} DD^T & o \\ o^T & b_m^T b_m \end{pmatrix}$$

Tedy $\det(A'A'^T) = b_m^T b_m \det(DD^T)$ a odmocněním dostaneme $\sqrt{\det(A'A'^T)} = \|b_m\| \sqrt{\det(DD^T)}$. To odpovídá intuitivní představě objemu jako velikosti výšky krát obsah základny. $\square$

Poznámka 9.23 (Objem rovnoběžnostěnu a elementární úpravy)**.** Platnost věty 9.22 lze nahlédnout geometricky rozborem vlivu elementárních úprav. V zásadě důkaz věty jen technicky zpracovává následující myšlenky:

Uvažujme rovnoběžnostěn generovaný řádky matice $A \in \mathbb{R}^{n\times n}$ a chceme ukázat, že jeho objem je $|\det(A)|$. Víme, že determinant se nezmění pokud na matici provádíme třetí elementární úpravu (přičtení násobku jednoho řádku k jinému). Samotný rovnoběžnostěn se ale změní. Pochopit, proč objem zůstává zachován, je snadné z geometrického náhledu: Přičtením násobku řádku k jinému (například poslednímu) znamená, že se rovnoběžnostěn zkosí či narovná, ale jeho základna i výška zůstane stejná.

Představit si ostatní elementární úpravy je ještě snazší: prohození řádků matice A znamená překlopení rovnoběžnostěnu a jeho objem se proto nezmění, a vynásobení řádku matice A číslem α pak protáhne rovnoběžnostěn v jednom směru, a tudíž se objem změní α-krát.

Je-li matice A singulární, pak odpovídající rovnoběžnostěn leží v nějakém podprostoru dimenze menší než n, a tudíž je jeho objem nulový. Je-li matice A regulární, pak ji elementárními úpravami převedeme na jednotkovou matici – odpovídající rovnoběžnostěn je jednotková krychle, a ta má objem 1.

Poznámka 9.24 (Vysvětlení definice determinantu)**.** Předchozí poznámka umožňuje alternativní způsob zavedení determinantu a vysvětlení jeho definice. Kdybychom chtěli zavést determinant matice A jako objem odpovídajícího rovnoběžnostěnu, narazíme na problém znaménka, protože objem je vždy nezáporný. Zavedeme tedy něco jako orientovaný objem, a to pomocí základních vlastností, které by objem měl splňovat:

1. Determinant jednotkové matice I_n je roven 1, což odpovídá objemu jednotkové krychle.
2. Výměna řádků změní znaménko determinantu. To odpovídá vlastnosti, že objem rovnoběžnostěnu se nezmění změnou pořadí hran, tedy překlopením, nicméně změna znaménka zavede právě určitou orientaci do definice determinantu.
3. Vynásobení řádku matice A číslem $\alpha \in \mathbb{R}$ změní determinant s koeficientem α. To odpovídá protažení rovnoběžnostěnu ve směru dané hrany, a tím pádem odpovídající změnu objemu.
4. Řádková linearita determinantu ve smyslu věty 9.4. Důsledek této vlastnosti je například to, že zkosení nezmění objem rovnoběžnostěnu, a proto i determinant zůstane stejný.

Z těchto základních vlastností již jdou odvodit všechny ostatní vlastnosti determinantu a vysvětlit i původní definici

$$\det(A) = \textstyle\sum_{p\in S_n} \operatorname{sgn}(p) a_{1,p(1)} \dots a_{n,p(n)}.$$

Z řádkové linearity determinantu můžeme podobně jako v Laplaceově rozvoji podle prvního řádku psát

$$\det(A) = \begin{vmatrix} a_{11} & 0 & \dots & 0 \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} + \begin{vmatrix} 0 & a_{12} & \dots & 0 \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} + \dots + \begin{vmatrix} 0 & \dots & 0 & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}.$$

Nyní každý z n determinantů napravo rozvineme podle druhého řádku a tak dále až podle posledního řádku. Tím nakonec dostaneme vyjádření $\det(A)$ pomocí součtu n^n determinantů jednoduchých matic:

$$\det(A) = \begin{vmatrix} a_{11} & 0 & \dots & 0 \\ a_{21} & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ a_{n1} & 0 & \dots & 0 \end{vmatrix} + \begin{vmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & \dots & 0 \\ \vdots & \vdots & & \vdots \\ a_{n1} & 0 & \dots & 0 \end{vmatrix} + \dots + \begin{vmatrix} 0 & \dots & 0 & a_{1n} \\ 0 & \dots & 0 & a_{2n} \\ \vdots & & \vdots & \vdots \\ 0 & \dots & 0 & a_{nn} \end{vmatrix}.$$

Každá z těchto jednoduchých matic má v každém řádku nanejvýš jedno nenulové číslo. Pokud tato čísla nejsou v navzájem různých sloupcích, je v matici nulový sloupec a její determinant je nula. Tím pádem z celkového počtu n^n determinantů jednoduchých matic zůstane jen $n!$ jednoduchých matic, jejichž potenciálně nenulové prvky jsou rozmístěny podle nějaké permutace:

$$\det(A) = \begin{vmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & a_{nn} \end{vmatrix} + \begin{vmatrix} 0 & a_{11} & \dots & 0 \\ a_{22} & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & a_{nn} \end{vmatrix} + \dots + \begin{vmatrix} 0 & \dots & 0 & a_{1n} \\ 0 & \dots & a_{2n} & 0 \\ \vdots & & & \vdots \\ a_{nn} & 0 & \dots & 0 \end{vmatrix}.$$

Determinant každé takovéto jednoduché matice je pak součin těch n prvků, a jeho znaménko odpovídá znaménku příslušné permutace p, podle které jsou rozmístěny dané prvky (protože to je parita počtu výměn sloupců, které potřebujeme, abychom matici dovedli na diagonální tvar). Tudíž determinant jednoduché matice je $\operatorname{sgn}(p)a_{1,p(1)} \dots a_{n,p(n)}$, a součtem jich všech dostaneme determinant matice A.

Poznámka 9.25 (Objem jiných geometrických těles)**.** Buď $A \in \mathbb{R}^{n\times n}$. Jak jsme již zmínili, objem geometrických těles se při zobrazení $x \mapsto Ax$ mění s koeficientem $|\det(A)|$. Krychle o hraně 1 se zobrazí na rovnoběžnostěn o hranách, které odpovídají sloupcům matice A, a jeho objem je proto $|\det(A^T)| = |\det(A)|$. Tuto vlastnost můžeme zobecnit na ostatní „rozumná“ geometrická tělesa, protože každé lze aproximovat krychličkami, a jejich obraz je tedy aproximován rovnoběžnostěny a změna objemu je přibližně $|\det(A)|$. Postupným zjemňováním aproximace dostaneme limitním přechodem výsledný poměr.

Příklad 9.26 (Geometrická interpretace determinantu)**.** Obraz jednotkové krychle při zobrazení $x \mapsto Ax$:

objem = 1 objem = $|\det(A)|$

Obraz geometrického tělesa při zobrazení $x \mapsto Ax$:

objem = V objem = $|\det(A)| \cdot V$ □

Determinanty se používají při řešení ještě mnoha dalších geometrických problémů [Berger, 1987; Dattorro, 2011; Ratschek and Rokne, 2003]. Výpočtem determinantu tak například snadno rozhodneme, zda daný bod v rovině leží uvnitř či vně kružnice zadané svými třemi body, a podobně ve vyšších dimenzích.

Zmiňme úlohu, která souvisí s objemem rovnoběžnostěnu, a to určení objemu mnohostěnu s $n+1$ vrcholy v $\mathbb{R}^n$. Bez újmy na obecnosti nechť je jeden vrchol a_0 v počátku a ostatní mají pozice $a_1, \dots, a_n \in \mathbb{R}^n$. Definujme matici $A \in \mathbb{R}^{n\times n}$ tak, že její sloupce jsou vektory $a_1, \dots, a_n$. Pak objem mnohostěnu je $\frac{1}{n!}|\det(A)|$, čili tvoří jen část rovnoběžnostěnu danou faktorem $1 : n!$.

Předchozí způsob výpočtu objemu mnohostěnu předpokládal, že známe pozice jednotlivých vrcholů v prostoru $\mathbb{R}^n$. V některých případech (např. molekulární biologie) jsou ale známy pouze vzdálenosti $d_{ij} = \|a_i - a_j\|$ mezi jednotlivými vrcholy, tj. délky hran mnohostěnu. V tomto případě spočítáme objem pomocí tzv. Cayleyho–Mengerova determinantu jako

$$\frac{(-1)^{n-1}}{2^n (n!)^2} \begin{vmatrix} 0 & 1 & 1 & 1 & \dots & 1 \\ 1 & 0 & d_{01}^2 & d_{02}^2 & \dots & d_{0n}^2 \\ 1 & d_{10}^2 & 0 & d_{12}^2 & \dots & d_{1n}^2 \\ 1 & d_{20}^2 & d_{21}^2 & 0 & \dots & d_{2n}^2 \\ \vdots & \vdots & \vdots & \ddots & & \vdots \\ 1 & d_{n0}^2 & d_{n1}^2 & d_{2n}^2 & \dots & 0 \end{vmatrix}.$$

Problémy

9.1. Dokažte multiplikativnost determinantu přímo ze sloupcové linearity determinantu.

9.2. Dokažte, že v každém kroku Gaussovy–Jordanovy eliminace se každý nenulový prvek matice dá vyjádřit buď jako determinant, nebo podíl dvou determinantů podmatic původní matice.

9.3. Pomocí Cramerova pravidla odvoďte vzorec pro inverzní matici a porovnejte jej s adjungovanou maticí.

9.4. Pomocí věty o adjungované matici odvoďte Cramerovo pravidlo.

9.5. Rozhodněte, zda $\operatorname{adj}(AB) = \operatorname{adj}(BA)$ pro libovolné matice $A, B \in \mathbb{R}^{n\times n}$.

Shrnutí ke kapitole 9. Determinanty

Determinant matice A je číselná charakteristika matice a lze spočítat efektivně pomocí Gaussovy eliminace – stačí jen určit, jak elementární úpravy mění determinant matice. Determinant mj. udává, zda matice je regulární či singulární, a pomocí determinantu můžeme také explicitně vyjádřit řešení soustavy $Ax = b$ s regulární maticí A. Podobně můžeme explicitně vyjádřit inverzi regulární matice, což vede na pojem adjungovaná matice. Geometricky pak determinant reprezentuje koeficient, se kterým se mění objem těles při lineárním zobrazení $x \mapsto Ax$; speciálně udává objem rovnoběžnostěnu jehož hrany jsou dané řádky matice A.

Kapitola 10

Vlastní čísla

Vlastní čísla (dříve též nazývaná „charakteristická čísla“), podobně jako determinant, představují určitou charakteristiku matice. Narozdíl od determinantu je jejich význam ještě dalekosáhlejší.

Definice 10.1 (Vlastní čísla a vlastní vektory)**.** Buď $A \in \mathbb{C}^{n\times n}$. Pak $\lambda \in \mathbb{C}$ je *vlastní číslo* matice A a $x \in \mathbb{C}^n$ jemu příslušný *vlastní vektor*, pokud $Ax = \lambda x$, $x \neq o$.

Poznamenejme, že $x \neq o$ je nezbytná podmínka, protože pro $x = o$ by rovnost byla triviálně splněna pro každé $\lambda \in \mathbb{C}$. Na druhou stranu, $\lambda = 0$ klidně může nastat.

Povšimněme si dále, že vlastní vektor při daném vlastním čísle není určen jednoznačně – každý jeho nenulový násobek je také vlastním vektorem. Někdy se proto vlastní vektor normuje tak, aby $\|x\| = 1$.

Přirozeně, vlastní čísla a vektory lze definovat stejně nad jakýmkoli jiným tělesem. My zůstaneme u $\mathbb{R}$ resp. $\mathbb{C}$. Jak uvidíme později, komplexním číslům se nevyhneme i když matice A je reálná.

Vlastní čísla se dají zavést i obecněji. Buď V vektorový prostor a $f\colon V \to V$ lineární zobrazení. Pak λ je vlastní číslo a $x \neq o$ příslušný vlastní vektor, pokud platí $f(x) = \lambda x$. My se však vesměs budeme zabývat vlastními čísly matic, protože vzhledem k maticové reprezentaci lineárních zobrazení můžeme úlohu hledání vlastních čísel a vektorů lineárních zobrazení redukovat na matice.

Příklad 10.2 (Geometrická interpretace vlastních čísel a vektorů)**.** Vlastní vektor reprezentuje invariantní směr při zobrazení $x \mapsto Ax$, tedy směr, který se zobrazí opět na ten samý směr. Jinými slovy, je-li v vlastní vektor, pak přímka $\operatorname{span}\{v\}$ se zobrazí do sebe sama. Vlastní číslo pak představuje škálování v tomto invariantním směru.

- Překlopení dle přímky $y = -x$, matice zobrazení $A = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}$:

vlastní číslo 1, vlastní vektor $(-1, 1)^T$,

vlastní číslo -1, vlastní vektor $(1, 1)^T$.

- Rotace o úhel 90°, matice zobrazení $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$:

žádná reálná vlastní čísla.

□

Věta 10.3 (Charakterizace vlastních čísel a vektorů)**.** *Buď $A \in \mathbb{C}^{n\times n}$. Pak*

(1) $\lambda \in \mathbb{C}$ je vlastním číslem A právě tehdy, když $\det(A - \lambda I_n) = 0$,

(2) $x \in \mathbb{C}^n$ je vlastním vektorem příslušným k vlastnímu číslu $\lambda \in \mathbb{C}$ právě tehdy, když $o \neq x \in \operatorname{Ker}(A - \lambda I_n)$.

Důkaz.

(1) $\lambda \in \mathbb{C}$ je vlastním číslem A právě tehdy, když $Ax = \lambda I_n x$, $x \neq o$, neboli $(A - \lambda I_n)x = o$, $x \neq o$, což je ekvivalentní singularitě matice $A - \lambda I_n$, a to zase podmínce $\det(A - \lambda I_n) = 0$.

(2) Analogicky, $x \in \mathbb{C}^n$ je vlastním vektorem k vlastnímu číslu $\lambda \in \mathbb{C}$ právě tehdy, když $(A-\lambda I_n)x = o$, $x \neq o$, tedy x je v jádru matice $A - \lambda I_n$. □

Důsledkem věty je, že k danému vlastnímu číslu λ přísluší $\dim \operatorname{Ker}(A - \lambda I_n) = n - \operatorname{rank}(A - \lambda I_n)$ lineárně nezávislých vlastních vektorů.

10.1 Charakteristický polynom

První část věty 10.3 říká, že $\lambda \in \mathbb{C}$ je vlastním číslem matice A právě tehdy, když matice $A - \lambda I_n$ je singulární, neboli $\det(A-\lambda I_n) = 0$. Pokud se na λ díváme jako na komplexní proměnnou, tak najít vlastní číslo je totéž jako najít řešení rovnice $\det(A - \lambda I_n) = 0$. Rozvojem determinantu snadno nahlédneme, že levá strana je polynom stupně n v proměnné λ, což nás vede k tomu zavést následující pojem.

Definice 10.4 (Charakteristický polynom)**.** *Charakteristický polynom* matice $A \in \mathbb{C}^{n\times n}$ vzhledem k proměnné λ je $p_A(\lambda) = \det(A - \lambda I_n)$.

Z definice determinantu je patrné, že charakteristický polynom se dá vyjádřit ve tvaru

$$p_A(\lambda) = \det(A - \lambda I_n) = (-1)^n \lambda^n + a_{n-1}\lambda^{n-1} + \ldots + a_1\lambda + a_0.$$

Tedy je to skutečně polynom a má stupeň n. Snadno nahlédneme, že $a_{n-1} = (-1)^{n-1}(a_{11} + \ldots + a_{nn})$ a po dosazení $\lambda = 0$ získáme $a_0 = \det(A)$.

Podle základní věty algebry (věta 1.1) má tento polynom n komplexních kořenů (včetně násobností), označme je $\lambda_1, \ldots, \lambda_n$. Pak

$$p_A(\lambda) = (-1)^n(\lambda - \lambda_1)\ldots(\lambda - \lambda_n).$$

Podle věty 10.3(1) kořeny $p_A(\lambda)$ odpovídají vlastním číslům matice A. Vlastních čísel je tedy n, započítáme-li i jejich násobnosti. Tímto jsme odvodili následující:[1)]

Věta 10.5. *Vlastní čísla matice $A \in \mathbb{C}^{n\times n}$ jsou právě kořeny jejího charakteristického polynomu $p_A(\lambda)$, a je jich n včetně násobností.*

Příklad 10.6. Mějme matici $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ z příkladu 10.2. Pak

$$p_A(\lambda) = \det(A - \lambda I_n) = \det \begin{pmatrix} -\lambda & -1 \\ 1 & -\lambda \end{pmatrix} = \lambda^2 + 1.$$

Kořeny polynomu, a tedy vlastními čísly matice A, jsou $\pm i$. Vlastní vektor příslušný i je $(1, -i)^T$ a vlastní vektor příslušný $-i$ je $(1, i)^T$. □

[1)] Alternativní odvození existence a počtu vlastních čísel bez použití determinantu je k nalezení například v článku Axler [1995].

Definice 10.7 (Algebraická a geometrická násobnost vlastního čísla)**.** Buď $\lambda \in \mathbb{C}$ vlastní číslo matice $A \in \mathbb{C}^{n\times n}$. *Algebraická násobnost* λ je rovna násobnosti λ jakožto kořene $p_A(\lambda)$. *Geometrická násobnost* λ je rovna $n - \operatorname{rank}(A - \lambda I_n)$, tj. počtu lineárně nezávislých vlastních vektorů, které odpovídají λ.

Algebraická násobnost je vždy větší nebo rovna geometrické násobnosti, což vyplyne v sekci 10.4. Nadále budeme pojem násobnost používat pro algebraickou násobnost.

Definice 10.8 (Spektrum a spektrální poloměr)**.** Nechť $A \in \mathbb{C}^{n\times n}$ má vlastní čísla $\lambda_1, \ldots, \lambda_n$. Pak *spektrum* matice A je množina jejích vlastních čísel $\{\lambda_1, \ldots, \lambda_n\}$ a *spektrální poloměr* je $\rho(A) = \max_{i=1,\ldots,n} |\lambda_i|$.

Počítat vlastní čísla jako kořeny charakteristického polynomu není příliš efektivní. Navíc, jak víme ze sekce 1.5, pro kořeny polynomu neexistuje žádný vzoreček ani konečný postup a počítají se iterativními metodami. Totéž platí i o vlastních číslech (uvidíme ve větě 10.16). Zde nepomůže ani jindy tak oblíbená a všestranná Gaussova eliminace. Nicméně pro některé speciální matice můžeme vlastní čísla určit snadno.

Příklad 10.9 (Vlastní čísla trojúhelníkové matice)**.**

- Nechť $A \in \mathbb{C}^{n\times n}$ je trojúhelníková matice. Pak její vlastní čísla jsou prvky na diagonále, neboť $\det(A - \lambda I_n) = (a_{11} - \lambda)\ldots(a_{nn} - \lambda)$.
- Speciálně, I_n má vlastní číslo 1, které je n-násobné. Množina příslušných vlastních vektorů je $\mathbb{R}^n \backslash \{o\}$.
- Speciálně, 0_n má vlastní číslo 0, které je n-násobné. Množina příslušných vlastních vektorů je $\mathbb{R}^n \backslash \{o\}$.
- Speciálně, matice $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ má vlastní číslo 1, které je dvojnásobné (algebraicky). Odpovídající vlastní vektor je až na násobek pouze $(1, 0)^T$, proto je geometrická násobnost vlastního čísla 1 pouze jedna.

□

Příklad 10.10. Mějme matici $A = \begin{pmatrix} 1.5 & 0.75 \\ 0 & 1 \end{pmatrix}$. Příslušné lineární zobrazení $x \mapsto Ax$ geometricky představuje skosení a protáhnutí v ose x_1 o 50%. Tuto transformaci ilustruje obrázek Stádleckého mostu dole (počátek souřadných os by byl u obrázku vlevo dole)

původní obrázek

obrázek po transformaci

Vlastní čísla matice A jsou 1.5 a 1, a jim příslušející vlastní vektory jsou $(1, 0)^T$ a $(-1.5, 1)^T$. První vlastní číslo a vektor říkají, že se obrázek protáhne o 50% ve směru osy x_1. Druhé vlastní číslo a vektor říkají, že se obrázek ve směru vektoru $(-1.5, 1)^T$ nedeformuje. Tím pádem se obrázek sice neprotahuje ve směru osy x_2, ale dochází ke skosení. □

Pro následující připomeňme, že $\operatorname{trace}(A)$ značí stopu matice A, čili součet diagonálních prvků A.

Tvrzení 10.11 (Součin a součet vlastních čísel)**.** *Buď $A \in \mathbb{C}^{n\times n}$ s vlastními čísly $\lambda_1, \ldots, \lambda_n$. Pak*

(1) $\det(A) = \lambda_1 \ldots \lambda_n$,

(2) $\operatorname{trace}(A) = \lambda_1 + \ldots + \lambda_n$.

Důkaz.

(1) Víme, že $\det(A - \lambda I_n) = (-1)^n(\lambda - \lambda_1)\dots(\lambda - \lambda_n)$. Dosazením $\lambda = 0$ dostáváme $\det(A) = (-1)^n(-\lambda_1)\dots(-\lambda_n) = \lambda_1 \dots \lambda_n$.

(2) Porovnejme koeficienty u λ^{n-1} různých vyjádření charakteristického polynomu. V rozvoji $\det(A - \lambda I_n)$ dostáváme, že koeficient vznikne pouze ze součinu $(a_{11} - \lambda)\dots(a_{nn} - \lambda)$, a má hodnotu $(-1)^{n-1}(a_{11} + \dots + a_{nn})$. Koeficient u λ^{n-1} v rozvoji $(-1)^n(\lambda - \lambda_1)\dots(\lambda - \lambda_n)$ je očividně $(-1)^n(-\lambda_1 - \dots - \lambda_n)$. Porovnáním tedy $(-1)^{n-1}(a_{11} + \dots + a_{nn}) = (-1)^n(-\lambda_1 - \dots - \lambda_n)$. □

Poznamenejme, že porovnáním koeficientů u jiných členů charakteristického polynomu dostaneme další vztahy mezi prvky matice A a vlastními čísly, ale již trochu komplikovanější.

Tvrzení 10.12 (Vlastnosti vlastních čísel)**.** *Nechť $A \in \mathbb{C}^{n\times n}$ má vlastní čísla $\lambda_1, \dots, \lambda_n$ a jim odpovídající vlastní vektory $x_1, \dots, x_n$. Pak:*

(1) A je regulární právě tehdy, když 0 není její vlastní číslo,

(2) je-li A regulární, pak A^{-1} má vlastní čísla $\lambda_1^{-1}, \dots, \lambda_n^{-1}$ a vlastní vektory $x_1, \dots, x_n$,

(3) A^2 má vlastní čísla $\lambda_1^2, \dots, \lambda_n^2$ a vlastní vektory $x_1, \dots, x_n$,

(4) αA má vlastní čísla $\alpha\lambda_1, \dots, \alpha\lambda_n$ a vlastní vektory $x_1, \dots, x_n$,

(5) $A + \alpha I_n$ má vlastní čísla $\lambda_1 + \alpha, \dots, \lambda_n + \alpha$ a vlastní vektory $x_1, \dots, x_n$,

(6) A^T má vlastní čísla $\lambda_1, \dots, \lambda_n$, ale vlastní vektory obecně jiné.

Důkaz. Dokážeme první dvě tvrzení, ostatní necháme čtenáři na rozmyšlení.

(1) A má vlastní číslo 0 právě tehdy, když $0 = \det(A - 0I_n) = \det(A)$, neboli když A je singulární.

(2) Pro každé $i = 1, \dots, n$ je $Ax_i = \lambda_i x_i$. Přenásobením A^{-1} dostaneme $x_i = \lambda_i A^{-1} x_i$ a vydělením $\lambda_i \neq 0$ pak $A^{-1}x_i = \lambda_i^{-1} x_i$. □

V následující větě je důležité, že matice A je reálná, pro komplexní už tvrzení neplatí.

Tvrzení 10.13. *Je-li $\lambda \in \mathbb{C}$ vlastní číslo matice $A \in \mathbb{R}^{n\times n}$, pak i komplexně sdružené $\overline{\lambda}$ je vlastním číslem A.*

Důkaz. Víme, že λ je kořenem $p_A(\lambda) = (-1)^n\lambda^n + a_{n-1}\lambda^{n-1} + \dots + a_1\lambda + a_0 = 0$. Komplexním sdružením obou stran rovnosti máme $(-1)^n\overline{\lambda}^n + a_{n-1}\overline{\lambda}^{n-1} + \dots + a_1\overline{\lambda} + a_0 = 0$, tedy i $\overline{\lambda}$ je kořenem $p_A(\lambda)$. □

Příklad 10.14. Spektrum reálné matice je tedy množina symetrická podle reálné osy. Komplexní matice můžou mít za spektrum jakýchkoli n komplexních čísel.

□

Nyní ukážeme, že výpočet kořenů polynomu lze převést na úlohu hledání vlastních čísel určité matice.

Definice 10.15 (Matice společnice[2]). Buď $p(x) = x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0$. Pak *matice společnice* polynomu $p(x)$ je čtvercová matice řádu n definovaná

$$C(p) := \begin{pmatrix} 0 & \ldots & \ldots & 0 & -a_0 \\ 1 & \ddots & & \vdots & -a_1 \\ 0 & \ddots & \ddots & \vdots & -a_2 \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & \ldots & 0 & 1 & -a_{n-1} \end{pmatrix}.$$

Věta 10.16 (O matici společnici). *Pro charakteristický polynom matice $C(p)$ platí $p_{C(p)}(\lambda) = (-1)^n p(\lambda)$, tedy vlastní čísla $C(p)$ odpovídají kořenům polynomu $p(\lambda)$.*

Důkaz. Upravme

$$p_{C(p)}(\lambda) = \det(C(p) - \lambda I_n) = \det \begin{pmatrix} -\lambda & \ldots & \ldots & 0 & -a_0 \\ 1 & \ddots & & \vdots & -a_1 \\ 0 & \ddots & \ddots & \vdots & -a_2 \\ \vdots & \ddots & \ddots & -\lambda & \vdots \\ 0 & \ldots & 0 & 1 & -a_{n-1} - \lambda \end{pmatrix}.$$

Přičtením λ-násobku posledního řádku k předposlednímu, pak λ-násobku předposledního řádku k předpředposlednímu, atd. dostaneme

$$p_{C(p)}(\lambda) = \det \begin{pmatrix} 0 & \ldots & \ldots & 0 & -p(\lambda) \\ 1 & \ddots & & \vdots & \vdots \\ 0 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & -a_{n-2} - a_{n-1}\lambda - \lambda^2 \\ 0 & \ldots & 0 & 1 & -a_{n-1} - \lambda \end{pmatrix}.$$

Laplaceovým rozvojem podle prvního řádku pak $p_{C(p)}(\lambda) = (-1)^{n+1}(-p(\lambda)) \det(I_{n-1}) = (-1)^n p(\lambda)$. □

Věta má mj. za důsledek, že úloha hledání kořenů reálných polynomů a vlastních čísel matic jsou na sebe navzájem převoditelné: Věta 10.5 redukuje hledání vlastních čísel matice na hledání kořenů polynomu, a věta 10.16 to dělá naopak. S ohledem na poznámku 1.2 to znamená, že vlastní čísla obecně můžeme počítat pouze numericky, žádný konečný postup neexistuje (praktické numerické metody jsou ale efektivní). Zatímco vlastní čísla se přes kořeny charakteristického polynomu v praxi nepočítají, opačný postup použitelný je (i když se využívají spíš specializované metody). Zajímavostí je, že např. Matlab počítá charakteristický polynom $p_A(\lambda)$ matice A tak, že najde vlastní čísla $\lambda_1, \ldots, \lambda_n$ a pak roznásobí dvojčleny ve výrazu $p_A(\lambda) = (-1)^n(\lambda - \lambda_1)\ldots(\lambda - \lambda_n)$.

10.2 Cayleyho–Hamiltonova věta

Příklad 10.17. Abychom lépe porozuměli následující větě, uvádíme příklad polynomiální matice a maticového polynomu

$$\begin{pmatrix} \lambda^2 - \lambda & 2\lambda - 3 \\ 7 & 5\lambda^2 - 4 \end{pmatrix} = \lambda^2 \begin{pmatrix} 1 & 0 \\ 0 & 5 \end{pmatrix} + \lambda \begin{pmatrix} -1 & 2 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & -3 \\ 7 & -4 \end{pmatrix}.$$

Jsou to dva zápisy stejné matice s parametrem λ, a můžeme jednoduše převést jeden zápis na druhý. V různých situacích bývá užitečný jeden či druhý tvar. □

[2] V angličtině *companion matrix*. V češtině se používají různé termíny, např. doprovodná matice polynomu, matice přidružená polynomu atd. Pojem *matice společnice* začal používat nezávisle autor a někteří pracovníci jiných škol.

Následující větu formulujeme pro komplexní matice, ale její platnost je širší – platí nad libovolným tělesem a dokonce ještě obecněji nad tzv. komutativními okruhy [Horn and Johnson, 1985].

Věta 10.18 (Cayleyho–Hamiltonova[3]). *Buď $A \in \mathbb{C}^{n\times n}$ a $p_A(\lambda) = (-1)^n\lambda^n + a_{n-1}\lambda^{n-1} + \ldots + a_1\lambda + a_0$. Pak*

$$(-1)^n A^n + a_{n-1}A^{n-1} + \ldots + a_1 A + a_0 I_n = 0.$$

Důkaz. Víme, že pro adjungované matice platí $(A - \lambda I_n)\operatorname{adj}(A - \lambda I_n) = \det(A - \lambda I_n)I_n$. Každý prvek $\operatorname{adj}(A - \lambda I_n)$ je polynom stupně nanejvýš $n-1$ proměnné λ, takže se dá vyjádřit ve tvaru $\operatorname{adj}(A - \lambda I_n) = \lambda^{n-1}B_{n-1} + \ldots + \lambda B_1 + B_0$ pro určité $B_{n-1}, \ldots, B_0 \in \mathbb{C}^{n\times n}$. Dosazením máme

$$\begin{aligned}(A - \lambda I_n)(\lambda^{n-1}B_{n-1} + \ldots + \lambda B_1 + B_0) &= \det(A - \lambda I_n)I_n = \\ &= ((-1)^n\lambda^n + a_{n-1}\lambda^{n-1} + \ldots + a_1\lambda + a_0)I_n.\end{aligned}$$

Roznásobením

$$\begin{aligned}- B_{n-1}\lambda^n + (AB_{n-1} - B_{n-2})\lambda^{n-1} + \ldots + (AB_1 - B_0)\lambda + AB_0 &= \\ = (-1)^n\lambda^n I_n + a_{n-1}\lambda^{n-1}I_n + \ldots + a_1\lambda I_n + a_0 I_n.\end{aligned}$$

Porovnáním koeficientů

$$\begin{aligned}-B_{n-1} &= (-1)^n I_n, \\ AB_j - B_{j-1} &= a_j I_n, \quad j = 1, \ldots, n-1, \\ AB_0 &= a_0 I_n.\end{aligned}$$

Vynásobme první rovnici A^n, další A^j a poslední $A^0 = I_n$. Sečtením pak získáme

$$\begin{aligned}&-A^nB_{n-1} + (A^nB_{n-1} - A^{n-1}B_{n-2}) + \ldots + (A^2B_1 - AB_0) + AB_0 = \\ &= 0 = (-1)^n A^n + a_{n-1}A^{n-1} + \ldots + a_1 A + a_0 I_n\end{aligned}$$ □

Zkráceně můžeme tvrzení Cayleyho–Hamiltonovy věty vyjádřit jako $p_A(A) = 0$, tj. matice je sama kořenem svého charakteristického polynomu. Přestože věta může působit jenom jako matematická zajímavost, má netriviální důsledky.

Důsledek 10.19. *Buď $A \in \mathbb{C}^{n\times n}$. Pak:*

(1) Pro každé $k \in \mathbb{N}$ je $A^k \in \operatorname{span}\{I_n, A, \ldots, A^{n-1}\}$, tedy A^k je lineární kombinací matic $I_n, A, \ldots, A^{n-1}$.

(2) je-li A regulární, pak $A^{-1} \in \operatorname{span}\{I_n, A, \ldots, A^{n-1}\}$.

Důkaz.

(1) Stačí uvažovat $k \geq n$. Při dělení polynomu λ^k polynomem $p_A(\lambda)$ se zbytkem tak vlastně polynom λ^k rozložíme $\lambda^k = r(\lambda)\,p_A(\lambda) + s(\lambda)$, kde $r(\lambda)$ je polynom stupně $k-n$ a $s(\lambda) = b_{n-1}\lambda^{n-1} + \ldots + b_1\lambda + b_0$ je zbytek. Pak

$$A^k = r(A)\,p_A(A) + s(A) = s(A) = b_{n-1}A^{n-1} + \ldots + b_1 A + b_0 I_n.$$

(2) Víme $p_A(A) = (-1)^n A^n + a_{n-1}A^{n-1} + \ldots + a_1 A + a_0 I_n = 0$ a $a_0 \neq 0$ z regularity A. Tedy

$$\begin{aligned}I &= -\frac{(-1)^n}{a_0}A^n - \frac{a_{n-1}}{a_0}A^{n-1} - \ldots - \frac{a_1}{a_0}A = \\ &= A\left(-\frac{(-1)^n}{a_0}A^{n-1} - \frac{a_{n-1}}{a_0}A^{n-2} - \ldots - \frac{a_1}{a_0}I_n\right).\end{aligned}$$

Tudíž vynásobením A^{-1} dostáváme

$$A^{-1} = -\frac{(-1)^n}{a_0}A^{n-1} - \frac{a_{n-1}}{a_0}A^{n-2} - \ldots - \frac{a_1}{a_0}I_n.$$ □

[3] Arthur Cayley objevil tuto identitu r. 1858, William Rowan Hamilton nezávisle na něm r. 1853, oba pouze pro řád 3. Na vyšší řády ji zobecnil až Ferdinand Georg Frobenius r. 1878.

Podle tohoto důsledku lze velkou mocninu A^k matice A spočítat alternativně tak, že najdeme příslušné koeficienty charakteristického polynomu a vyjádříme A^k jako lineární kombinaci $I_n, A, \ldots, A^{n-1}$.

Podobně můžeme vyjádřit i A^{-1}, a tím pádem vyjádřit řešení soustavy $Ax = b$ s regulární maticí jako $A^{-1}b = \frac{1}{a_0}(-(-1)^n A^{n-1}b - \ldots - a_1 b)$. Podobný přístup se občas používá pro řešení velmi rozměrných soustav rovnic (tzv. Krylovova metoda), ale tam se uvažuje trochu jiný polynom nižšího stupně.

10.3 Diagonalizovatelnost

Při řešení soustav lineárních rovnic pomocí Gaussovy–Jordanovy eliminace jsme používali elementární řádkové úpravy. Ty nemění množinu řešení a upraví matici soustavy na tvar, ze kterého snadno vyčteme řešení. Je přirozené hledat analogické, spektrum neměnící, úpravy i na problém počítání vlastních čísel. Elementární řádkové úpravy použít nelze, protože ty mění spektrum. Vhodnou transformací je tzv. *podobnost*, protože ta spektrum nemění. A pokud se nám podaří touto transformací převést matici na diagonální tvar, máme vyhráno – na diagonále jsou hledaná vlastní čísla.

Definice 10.20 (Podobnost)**.** Matice $A, B \in \mathbb{C}^{n\times n}$ jsou *podobné*, pokud existuje regulární $S \in \mathbb{C}^{n\times n}$ tak, že $A = SBS^{-1}$.

Podobnost jde ekvivalentně definovat vztahem $AS = SB$ pro nějakou regulární matici S. Tímto přepisem můžeme i najít matici S, skrze kterou jsou A, B podobné (více viz Bečvář [2005]).

Příklad 10.21. Matice

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{a} \quad \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

jsou si podobné skrze matici $S = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$.

Věta 10.22 (Vlastní čísla podobných matic)**.** *Podobné matice mají stejná vlastní čísla.*

Důkaz. Z podobnosti matic existuje regulární matice S taková, že $A = SBS^{-1}$. Pak

$$\begin{aligned} p_A(\lambda) &= \det(A - \lambda I_n) = \det(SBS^{-1} - \lambda S I_n S^{-1}) = \det(S(B - \lambda I_n)S^{-1}) = \\ &= \det(S)\det(B - \lambda I_n)\det(S^{-1}) = \det(B - \lambda I_n) = p_B(\lambda). \end{aligned}$$

Obě matice mají stejné charakteristické polynomy, tedy i vlastní čísla. □

Příklad 10.23. Ukažte, že podobnost jako binární relace je reflexivní, symetrická a tranzitivní. Tedy jedná se o relaci ekvivalenci. □

Uvědomme si, že věta neříká nic o vlastních vektorech, ty se měnit mohou. Co ale zůstává neměnné, je jejich počet.

Věta 10.24. *Nechť matice $A, B \in \mathbb{C}^{n\times n}$ jsou podobné a nechť λ je jejich vlastní číslo. Pak počet vlastních vektorů, odpovídajících λ, je stejný u obou matic.*

Důkaz. Buď $A = SBS^{-1}$. Dále víme, že počet (lineárně nezávislých) vlastních vektorů, které odpovídají vlastnímu číslu λ matice A, je roven dimenzi $\operatorname{Ker}(A - \lambda I_n)$, tedy číslu $n - \operatorname{rank}(A - \lambda I_n)$. Protože hodnost matice se nemění přenásobením regulární maticí, tak dostáváme

$$\begin{aligned} \operatorname{rank}(A - \lambda I_n) &= \operatorname{rank}(SBS^{-1} - \lambda I_n) = \operatorname{rank}(SBS^{-1} - \lambda SS^{-1}) = \\ &= \operatorname{rank}(S(B - \lambda I_n)S^{-1}) = \operatorname{rank}(B - \lambda I_n). \end{aligned}$$

Tudíž dimenze jádra obou matic $A - \lambda I_n$ a $B - \lambda I_n$ jsou stejné, a tím pádem i počet vlastních vektorů. □

Vlastní čísla se ale podobnostní transformací nemění, tedy jestliže matici A převedeme podobnostní transformací na diagonální či obecněji trojúhelníkovou, tak na diagonále najdeme její vlastní čísla. Speciálně matice, které jdou převést na diagonální tvar, mají obzvláště pěkné vlastnosti.

Definice 10.25 (Diagonalizovatelnost)**.** Matice $A \in \mathbb{C}^{n\times n}$ je *diagonalizovatelná*, pokud je podobná nějaké diagonální matici.

Diagonalizovatelná matice A jde tedy vyjádřit ve tvaru $A = S\Lambda S^{-1}$, kde S je regulární a Λ diagonální. Tomuto tvaru se říká *spektrální rozklad*, a to proto, že na diagonále matice Λ je spektrum matice A.

Příklad 10.26. Ne každá matice je diagonalizovatelná, např.

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

není. Její vlastní číslo (dvojnásobné) je 0. Pokud by A byla diagonalizovatelná, pak by byla podobná nulové matici, tedy $A = S0S^{-1} = 0$, což je spor. □

Věta 10.27 (Charakterizace diagonalizovatelnosti)**.** *Matice $A \in \mathbb{C}^{n\times n}$ je diagonalizovatelná právě tehdy, když má n lineárně nezávislých vlastních vektorů.*

Důkaz. Implikace „$\Rightarrow$“. Je-li A diagonalizovatelná, pak má spektrální rozklad $A = S\Lambda S^{-1}$, kde S je regulární a Λ diagonální. Rovnost přepíšeme $AS = S\Lambda$ a porovnáním j-tých sloupců dostaneme

$$AS_{*j} = (AS)_{*j} = (S\Lambda)_{*j} = S\Lambda_{*j} = S\Lambda_{jj}e_j = \Lambda_{jj}S_{*j},$$

což můžeme názorně zapsat jako

$$AS = A\begin{pmatrix} | & & | \\ S_{*1} & \dots & S_{*n} \\ | & & | \end{pmatrix} = \begin{pmatrix} | & & | \\ (\Lambda_{11}S_{*1}) & \dots & (\Lambda_{nn}S_{*n}) \\ | & & | \end{pmatrix} = S\Lambda.$$

Tedy Λ_{jj} je vlastní číslo a S_{*j} je příslušný vlastní vektor. Sloupce matice S jsou lineárně nezávislé díky její regularitě.

Implikace „$\Leftarrow$“. Analogicky opačným směrem. Nechť A má vlastní čísla $\lambda_1, \dots, \lambda_n$ a jim přísluší lineárně nezávislé vlastní vektory $x_1, \dots, x_n$. Sestavme regulární matici $S := (x_1 \mid \cdots \mid x_n)$ a diagonální $\Lambda := \operatorname{diag}(\lambda_1, \dots, \lambda_n)$. Pak

$$(AS)_{*j} = AS_{*j} = Ax_j = \lambda_j x_j = \Lambda_{jj}S_{*j} = S(\Lambda_{jj}e_j) = S\Lambda_{*j} = (S\Lambda)_{*j}.$$

Tedy $AS = S\Lambda$, z čehož $A = S\Lambda S^{-1}$. □

Důkaz věty byl konstruktivní, tedy dává návod jak diagonalizovat matici ze znalosti vlastních čísel a vlastních vektorů. Podobně i naopak ze znalosti spektrálního rozkladu $A = S\Lambda S^{-1}$ snadno určíme vlastní vektory (jsou to sloupce matice S v pořadí odpovídajícím vlastním číslům na diagonále Λ).

Nediagonalizovatelné matice jsou ty, pro které nastávají určité patologické situace, ale diagonalizovatelné matice mají celou řadu přirozených vlastností. Je-li matice A diagonalizovatelná, pak:

- Algebraická a geometrická násobnost každého vlastního čísla A je stejná.

  (*Důkaz.* Ze spektrálního rozkladu.)

- Hodnost matice A je rovna počtu nenulových vlastních čísel A.

  (*Důkaz.* Ze spektrálního rozkladu $\operatorname{rank}(A) = \operatorname{rank}(S\Lambda S^{-1}) = \operatorname{rank}(\Lambda) =$ počet nenulových vlastních čísel.)

Příklad 10.28. Pro diagonalizovatelné matice dokažte vlastnosti z tvrzení 10.12 pomocí spektrálního rozkladu. □

Poznámka 10.29 (Geometrická interpretace diagonalizace)**.** Jiný pohled na diagonalizaci je geometrický: víme, že vlastní vektor představuje invariantní směr při zobrazení $x \mapsto Ax$. Nyní si představme, že A představuje matici nějakého lineárního zobrazení $f\colon \mathbb{C}^n \to \mathbb{C}^n$ vzhledem k bázi B. Buď $S = {}_{B'}[id]_B$ matice přechodu od B k jiné bázi B'. Pak $SAS^{-1} = {}_{B'}[id]_B \cdot {}_B[f]_B \cdot {}_B[id]_{B'} = {}_{B'}[f]_{B'}$ je matice zobrazení f vzhledem k nové bázi B'. Nyní diagonalizovatelnost můžeme chápat jako hledání vhodné báze B', aby příslušná matice byla diagonální, a tak jednoduše popisovala chování zobrazení.

Příklad 10.30 (Geometrická interpretace diagonalizace)**.** Buď

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

Vlastní čísla a vlastní vektory matice A jsou:

$$\lambda_1 = 4,\ x_1 = (1,1)^T, \quad \lambda_2 = 2,\ x_2 = (-1,1)^T.$$

Diagonalizace má tvar:

$$A = S\Lambda S^{-1} = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 4 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

Geometrická interpretace: v souřadném systému vlastních vektorů je matice zobrazení diagonální a zobrazení představuje jen škálování na osách. Na obrázku dole je znázorněný jednotkový čtverec a jeho obraz při zobrazení $x \mapsto Ax$. Nalevo je situace, kdy pracujeme v souřadném systému kanonické báze, a napravo v souřadném systému báze z vlastních vektorů.

□

Nyní ukážeme, že různým vlastním číslům odpovídají lineárně nezávislé vlastní vektory.

Tvrzení 10.31 (Vlastní vektory různých vlastních čísel)**.** *Nechť $\lambda_1, \ldots, \lambda_k$ jsou navzájem různá vlastní čísla (ne nutně všechna) matice $A \in \mathbb{C}^{n\times n}$. Pak odpovídající vlastní vektory $x_1, \ldots, x_k$ jsou lineárně nezávislé.*

Důkaz. Matematickou indukcí podle k. Pro $k = 1$ zřejmé, neboť vlastní vektor je nenulový.

Indukční krok $k \leftarrow k - 1$. Uvažme lineární kombinaci

$$\alpha_1 x_1 + \ldots + \alpha_k x_k = o. \tag{10.1}$$

Pak přenásobením maticí A dostaneme

$$A(\alpha_1 x_1 + \ldots + \alpha_k x_k) = \alpha_1 A x_1 + \ldots + \alpha_k A x_k = \alpha_1 \lambda_1 x_1 + \ldots + \alpha_k \lambda_k x_k = o. \tag{10.2}$$

Odečtením λ_k-násobku (10.1) od (10.2) dostaneme

$$\alpha_1(\lambda_1 - \lambda_k)x_1 + \ldots + \alpha_{k-1}(\lambda_{k-1} - \lambda_k)x_{k-1} = o.$$

Z indukčního předpokladu jsou $x_1, \ldots, x_{k-1}$ lineárně nezávislé, tedy $\alpha_1 = \ldots = \alpha_{k-1} = 0$. Dosazením do (10.1) máme $\alpha_k x_k = o$, neboli $\alpha_k = 0$. □

Důsledek 10.32. *Pokud matice $A \in \mathbb{C}^{n\times n}$ má n navzájem různých vlastních čísel, pak je diagonalizovatelná.*

Nyní předvedeme několik teoretických i praktických aplikací diagonalizace.

Tvrzení 10.33. *Buď $A, B \in \mathbb{C}^{n\times n}$. Pak matice AB i BA mají stejná vlastní čísla včetně násobností.*

Důkaz. Matice

$$\begin{pmatrix} AB & 0 \\ B & 0 \end{pmatrix} \quad \text{resp.} \quad \begin{pmatrix} 0 & 0 \\ B & BA \end{pmatrix}$$

jsou blokově trojúhelníkové, proto mají stejná vlastní čísla jako AB resp. BA, plus navíc n-násobné vlastní číslo 0. Nyní stačí ukázat shodu spekter obou blokových matic. Tyto matice jsou podobné skrze matici $S = \left(\begin{smallmatrix} I & A \\ 0 & I \end{smallmatrix}\right)$, neboť

$$\begin{pmatrix} AB & 0 \\ B & 0 \end{pmatrix} \begin{pmatrix} I & A \\ 0 & I \end{pmatrix} = \begin{pmatrix} AB & ABA \\ B & BA \end{pmatrix} = \begin{pmatrix} I & A \\ 0 & I \end{pmatrix} \begin{pmatrix} 0 & 0 \\ B & BA \end{pmatrix}$$ □

Předchozí věta platí i pro obdélníkové matice $A, B^T \in \mathbb{R}^{m\times n}$, ale znění platí pouze pro nenulová vlastní čísla; násobnost nulových vlastních čísel může být (a typicky je) odlišná.

Příklad 10.34 (Mocnina matice)**.** Buď $A = S\Lambda S^{-1}$ spektrální rozklad matice $A \in \mathbb{C}^{n\times n}$. Pak $A^2 = S\Lambda S^{-1}S\Lambda S^{-1} = S\Lambda^2 S^{-1}$. Obecněji:

$$A^k = S\Lambda^k S^{-1} = S \begin{pmatrix} \lambda_1^k & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_n^k \end{pmatrix} S^{-1}.$$

Můžeme studovat i asymptotické chování. Zjednodušeně:

$$\lim_{k\to\infty} A^k = S \begin{pmatrix} \lim_{k\to\infty} \lambda_1^k & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lim_{k\to\infty} \lambda_n^k \end{pmatrix} S^{-1} =$$

$$= \begin{cases} 0, & \text{pokud } \rho(A) < 1, \\ \text{diverguje}, & \text{pokud } \rho(A) > 1, \\ \text{konverguje / diverguje}, & \text{pokud } \rho(A) = 1. \end{cases}$$

Případ $\rho(A) = 1$ se nedá obecně rozhodnout: Pro $A = I_n$ matice konverguje k I_n, pro $A = -I_n$ matice osciluje mezi I_n a $-I_n$.

Zde je opět užitečné sledovat geometrickou paralelu pro lineární zobrazení $x \mapsto Ax$. Mocnění matice A odpovídá skládání zobrazení se sebou samým. Pokud jsou vlastní čísla v absolutní hodnotě menší než 1 (tj., $\rho(A) < 1$), tak lineární zobrazení kontrahuje vzdálenosti a proto konverguje k nule pro $k \to \infty$. Pokud je alespoň jedno vlastní číslo v absolutní hodnotě větší než 1, pak lineární zobrazení ve směru vlastního vektoru roztahuje vzdálenosti, a proto diverguje pro $k \to \infty$. Mezní případ, kdy $\rho(A) = 1$ je nejzajímavější. Tento případ nastává například pro ortogonální matice nebo matice Markovových řetězců, které zmíníme v příkladu 10.55. □

Příklad 10.35 (Rekurentní vzorečky a Fibonacci)**.** Uvažujme posloupnost $a_1, a_2, \ldots$ zadanou rekurentním vztahem

$$a_n = pa_{n-1} + qa_{n-2},$$

kde a_1, a_2 jsou dané první hodnoty posloupnosti a p, q konstanty. Ukážeme, jak „rozbít“ rekurzi a vyjádřit explicitně n-tý prvek posloupnosti. Uvedený postup funguje i na složitější rekurze, kdy a_n závisí na více než dvou předchozích členech.

Rekurenci vyjádříme maticově

$$\begin{pmatrix} a_n \\ a_{n-1} \end{pmatrix} = \begin{pmatrix} p & q \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_{n-1} \\ a_{n-2} \end{pmatrix}.$$

Označíme-li

$$x_n := \begin{pmatrix} a_n \\ a_{n-1} \end{pmatrix}, \quad A = \begin{pmatrix} p & q \\ 1 & 0 \end{pmatrix},$$

pak rekurence má tvar

$$x_n = Ax_{n-1} = A^2 x_{n-2} = \ldots = A^{n-2} x_2.$$

Potřebujeme tedy určit vyšší mocninu matice A. K tomu poslouží postup z příkladu 10.34. Diagonalizujeme $A = S\Lambda S^{-1}$, a pak $x_n = S\Lambda^{n-2}S^{-1}x_2$. Nyní jsme hotovi, explicitní vyjádření a_n je skryto v první složce vektoru $x_n = S\Lambda^{n-2}S^{-1}x_2$.

Ukažme si postup konkrétně pro Fibonacciho posloupnost, která je daná rekurencí $a_n = a_{n-1} + a_{n-2}$ a prvními členy $a_1 = a_2 = 1$. Označme $\varphi := \frac{1}{2}(1+\sqrt{5})$ hodnotu zlatého řezu. Nyní

$$x_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} = S\Lambda S^{-1},$$

kde

$$S = \begin{pmatrix} -1 & \varphi \\ \varphi & 1 \end{pmatrix}, \quad S^{-1} = \frac{\sqrt{5}}{5}\begin{pmatrix} 1-\varphi & 1 \\ 1 & \varphi - 1 \end{pmatrix}, \quad \Lambda = \begin{pmatrix} 1-\varphi & 0 \\ 0 & \varphi \end{pmatrix}.$$

Tudíž

$$a_n = S_{1*}\Lambda^{n-2}S^{-1}x_2 = \frac{5-3\sqrt{5}}{10}(1-\varphi)^{n-2} + \frac{5+3\sqrt{5}}{10}\varphi^{n-2},$$

což se dá snadno přepsat do běžněji používaného tvaru

$$a_n = -\frac{\sqrt{5}}{5}(1-\varphi)^n + \frac{\sqrt{5}}{5}\varphi^n. \qquad \square$$

Příklad 10.36 (Diskrétní a rychlá Fourierova transformace, viz příklad 3.54)**.** Uvažujme dva polynomy

$$\begin{aligned} p(x) &= \textstyle\sum_{k=0}^{m} a_k x^k = a_m x^m + \ldots + a_1 x + a_0, \\ q(x) &= \textstyle\sum_{k=0}^{m} b_k x^k = b_m x^m + \ldots + b_1 x + b_0, \end{aligned}$$

kde ne nutně $a_m \neq 0$, $b_m \neq 0$. Naším cílem je rychle vynásobit tyto polynomy, a tedy získat polynom

$$\begin{aligned} p(x)q(x) &= \textstyle\sum_{k=0}^{2m} \left(\sum_{j=\max\{0,k-m\}}^{\min\{k,m\}} a_j b_{k-j} \right) x^k \\ &= a_m b_m x^{2m} + \ldots + (a_0 b_1 + a_1 b_0)x + a_0 b_0. \end{aligned}$$

Tento problém je důležitý nejenom pro násobení polynomů, ale analogická operace se provádí i při obyčejném násobení čísel nebo při diskrétní konvoluci konečných posloupností.

Součin $p(x)q(x)$ budeme reprezentovat maticově, a sice tak, že polynom $p(x)$ zakódujeme do matice a polynom $g(x)$ do vektoru s příslušnými koeficienty (prázdná políčka jsou nuly):

$$\begin{pmatrix} a_0 & & & \\ a_1 & a_0 & & \\ \vdots & a_1 & \ddots & \\ a_m & \vdots & \ddots & a_0 \\ & a_m & & a_1 \\ & & \ddots & \vdots \\ & & & a_m \end{pmatrix} \begin{pmatrix} b_0 \\ b_1 \\ \vdots \\ b_m \end{pmatrix}. \tag{10.3}$$

Výsledkem součinu bude vektor koeficientů polynomu $p(x)q(x)$. Aby se s maticí a vektory lépe pracovalo, rozšíříme je na velikost $n = 2m+1$ a doplníme nulami. Pokud navíc definujeme $a_{m+1} = \ldots = a_{n-1} = 0$, $b_{m+1} = \ldots = b_{n-1} = 0$, má výraz (10.3) tvar

$$Ab = \begin{pmatrix} a_0 & a_{n-1} & \ldots & a_2 & a_1 \\ a_1 & a_0 & a_{n-1} & & a_2 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ a_{n-2} & & \ddots & \ddots & a_{n-1} \\ a_{n-1} & a_{n-2} & \ldots & a_1 & a_0 \end{pmatrix} \begin{pmatrix} b_0 \\ \vdots \\ b_m \\ \vdots \\ b_n \end{pmatrix}.$$

Matice, které jsou ve tvaru jako má matice A, se nazývají *cirkulanty*. Tyto matice mají pozoruhodné vlastnosti. Jedna z nich je, že její vlastní vektory nezávisí na konkrétních hodnotách $a_0, \dots, a_{n-1}$, ale pouze na struktuře cirkulantu. Vlastní čísla matice A se dají jednoduše dopočítat z hodnot v matici. Navíc je matice A diagonalizovatelná, můžeme ji tedy vyjádřit ve tvaru $A = S\Lambda S^{-1}$, kde $S \in \mathbb{C}^{n\times n}$ se skládá z vlastních vektorů matice A a $\Lambda \in \mathbb{C}^{n\times n}$ je diagonální matice s vlastními čísly matice A na diagonále. Pro matici S navíc platí $S^{-1} = \frac{1}{n}\overline{S}^T$. Součin Ab se dá nyní vyjádřit jako

$$Ab = S\Lambda\tfrac{1}{n}\overline{S}^T b.$$

Součin Ab můžeme tedy vyjádřit pomocí tří operací: vynásobení postupně třemi maticemi $\frac{1}{n}\overline{S}^T$, Λ a S. Pokud si matici S představíme jako matici přechodu z báze vlastních vektorů do kanonické báze, pak první operace převede vektor b do souřadného systému vlastních vektorů (tzv. *diskrétní Fourierova transformace*), druhá provede hlavní operaci a třetí převede výsledný vektor zpět do kanonické báze (tzv. *inverzní Fourierova transformace*). Druhá operace je zřejmě triviální, protože Λ je diagonální.

Na první pohled se může zdát, že jsme celý problém pouze zkomplikovali. Nicméně, za použití vhodných algoritmů (založených např. na principu rozděl a panuj) lze vynásobení vektoru b maticí $\frac{1}{n}\overline{S}^T$ provést v čase, který je úměrný funkci $n\log(n)$. Podobně pro násobení maticí S. To je asymptoticky výrazné vylepšení oproti obyčejnému maticovému součinu Ab, který vyžaduje řádově n^2 aritmetických operací. Takto vylepšená metoda se nazývá *rychlá Fourierova transformace* a je to jeden z nejdůležitějších numerických algoritmů, viz strana 8. Kromě efektivního násobení velkých čísel nebo polynomů má velké využití ve zpracováni signálů a obrazů, pro kompresi dat atp. □

10.4 Jordanova normální forma

Nejjednodušší tvar matice, ke kterému lze dospět pomocí elementárních řádkových úprav, je redukovaný odstupňovaný tvar. Jaký však je nejjednodušší tvar matice, ke kterému lze dospět pomocí podobnosti? Diagonální matice to není, protože již víme, že ne všechny matice jsou diagonalizovatelné. Nicméně každou matici lze podobnostní transformací převést na poměrně jednoduchý tvar, nazývaný Jordanova normální forma.[4)]

Definice 10.37 (Jordanova buňka)**.** Buď $\lambda \in \mathbb{C}$, $k \in \mathbb{N}$. *Jordanova buňka* $J_k(\lambda)$ je čtvercová matice řádu k definovaná

$$J_k(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \dots & \dots & 0 & \lambda \end{pmatrix}.$$

Jordanova buňka má vlastní číslo λ, které je k-násobné, a přísluší mu pouze jeden vlastní vektor, protože matice $J_k(\lambda) - \lambda I_k$ má hodnost $k-1$.

Definice 10.38 (Jordanova normální forma)**.** Matice $J \in \mathbb{C}^{n\times n}$ je v *Jordanově normální formě*, pokud je v blokově diagonálním tvaru

$$J = \begin{pmatrix} J_{k_1}(\lambda_1) & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & J_{k_m}(\lambda_m) \end{pmatrix}$$

a na diagonále jsou Jordanovy buňky $J_{k_1}(\lambda_1), \dots, J_{k_m}(\lambda_m)$.

[4)] Autorem je francouzský matematik Marie Ennemond Camille Jordan. Spolutvůrcem Gaussovy–Jordanovy eliminace je však někdo jiný, německý geodet Wilhelm Jordan.

Hodnoty λ_i a k_i nemusí být navzájem různé. Stejně tak nějaká Jordanova buňka se může vyskytovat vícekrát.

Věta 10.39 (O Jordanově normální formě)**.** *Každá matice $A \in \mathbb{C}^{n\times n}$ je podobná matici v Jordanově normální formě. Tato matice je až na pořadí buněk určena jednoznačně.*

Důkaz. Úplný důkaz viz např. [Bečvář, 2005; Bican, 2009; Horn and Johnson, 1985]. Zde řekneme alespoň pár myšlenek z důkazu a konstrukce Jordanovy normální formy. Hledáme tedy bázi, vůči níž má zobrazení $x \mapsto Ax$ Jordanovu normální formu. Ukážeme, jak vypadají Jordanovy buňky a odpovídající část báze pro vlastní číslo λ. Bez újmu na obecnost buď $\lambda = 0$, jinak použijeme transformaci $A - \lambda I_n$.

Uvažujme podprostory $\mathbb{C}^n$:

$$\operatorname{Ker}(A) \subsetneqq \operatorname{Ker}(A^2) \subsetneqq \cdots \subsetneqq \operatorname{Ker}(A^p) = \operatorname{Ker}(A^{p+1}) = \dots$$

Buď $v \in \operatorname{Ker}(A^p) \setminus \operatorname{Ker}(A^{p-1})$, a uvažujme bázi B tvořenou vektory $A^{p-1}v, \dots, Av, v$. Na podprostoru generovaném těmito vektory pak lineární zobrazení $x \mapsto Ax$ má matici vůči bázi B rovnou Jordanově buňce $J_p(0)$.

Podobně dostaneme i ostatní Jordanovy buňky, jen postup trochu zobecníme. Namísto vektoru v vezmeme systém vektorů $v_1, \dots, v_\ell$ o velikosti $\dim \operatorname{Ker}(A^p) - \dim \operatorname{Ker}(A^{p-1})$, který doplňuje nějakou bázi $\operatorname{Ker}(A^{p-1})$ na bázi $\operatorname{Ker}(A^p)$. Každý z těchto vektorů je zdrojem řetízku vektorů $A^{p-1}v_i, \dots, Av_i, v_i$, který odpovídá Jordanově buňce $J_p(0)$; těchto buněk je tedy ℓ. Pokud $\ell' := \dim \operatorname{Ker}(A^{p-1}) - \dim \operatorname{Ker}(A^{p-2}) - \ell > 0$, tak musíme vektory $Av_1, \dots, Av_\ell$ doplnit o nové vektory $v_{\ell+1}, \dots, v_{\ell+\ell'}$, aby opět doplňovaly nějakou bázi $\operatorname{Ker}(A^{p-2})$ na bázi $\operatorname{Ker}(A^{p-1})$. Tyto nové vektory se stanou základem řetízků $A^{p-2}v_j, \dots, Av_j, v_j$, které odpovídají ℓ' Jordanovým buňkám typu $J_{p-1}(0)$. Tento postup opakujeme až do dimenze 1. □

Příklad 10.40. Matice

$$A = \begin{pmatrix} 5 & -2 & 2 & -2 & 0 \\ 0 & 6 & -1 & 3 & 2 \\ 2 & 2 & 7 & -2 & -2 \\ 2 & 3 & 1 & 2 & -4 \\ -2 & -2 & -2 & 6 & 11 \end{pmatrix}$$

má vlastní čísla 5 (dvojnásobné) a 7 (trojnásobné). Protože $3 = \operatorname{rank}(A - 5I_5) = \operatorname{rank}(A - 5I_5)^2$, tak budeme hledat dva řetízky o délce 1. Najdeme dva lineárně nezávislé vektory $x_1, x_2 \in \operatorname{Ker}(A - 5I_5)$, například $x_1 = (-2, 1, 1, 0, 0)^T$ a $x_2 = (-1, 1, 0, -1, 1)^T$, a ty tvoří základ hledané báze.

Přistupme k vlastnímu číslu 7. Nyní máme $\operatorname{rank}(A - 7I_5) = 3$ a $\operatorname{rank}(A - 7I_5)^2 = \operatorname{rank}(A - 7I_5)^3 = 2$. Zvolíme tedy $x_4 \in \operatorname{Ker}(A - 7I_5)^2 \setminus \operatorname{Ker}(A - 5I_5)$, například $x_4 = (1, 0, 1, 0, 0)^T$ a potom příslušnou část báze tvoří řetízek $x_3 = (A - 7I_5)x_4 = (0, -1, 2, 3, -4)^T$, x_4. Posledním vektorem báze bude vektor z $\operatorname{Ker}(A - 7I_5)$ lineárně nezávislý s vektorem x_3, tedy například $x_5 = (0, 1, 1, 0, 1)^T$.

Hledaná báze je tvořena vektory $x_1, \dots, x_5$. Dáme-li tyto vektory do sloupců matice S, pak

$$J = S^{-1}AS = \begin{pmatrix} 5 & 0 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 & 0 \\ 0 & 0 & 7 & 1 & 0 \\ 0 & 0 & 0 & 7 & 0 \\ 0 & 0 & 0 & 0 & 7 \end{pmatrix}$$

je hledaná Jordanova normální forma matice A. □

Jordanova normální forma je nestabilní v tom smyslu, že i nepatrná změna v původní matici způsobí skokovou změnu Jordanovy formy – není to tedy spojitá funkce vzhledem k prvkům matice A. To je také důvod, proč třeba výpočetní systém Maple odmítl zařadit výpočet Jordanovy formy do své knihovny.

Důkaz věty 10.39 naznačil postup jak spočítat Jordanovu normální formu pro danou matici A, teď se ale zaměříme na určité rysy Jordanovy formy. Z věty 10.24 víme, že počet vlastních vektorů se nemění podobnostní transformací. Vzhledem k tomu, že každé Jordanově buňce odpovídá jeden vlastní vektor, tak jsme právě odvodili:

Tvrzení 10.41. *Počet všech Jordanových buněk odpovídajících λ je roven počtu vlastních vektorů pro λ.*

Jako důsledek dále dostáváme, že (algebraická) násobnost každého vlastního čísla λ je vždy větší nebo rovna počtu vlastních vektorů, které mu přísluší (tedy geometrické násobnosti).

Důsledek 10.42. *Násobnost vlastního čísla je větší nebo rovna počtu vlastních vektorů, které mu přísluší.*

Nicméně ani tyto znalosti ještě nemusí stačit k určení Jordanovy normální formy. Obecně potřebujeme znát ještě nějakou informaci navíc, např. jak to dělá následující vzoreček.

Poznámka 10.43 (Velikosti a počet buněk)**.** Počet buněk $J_k(\lambda)$ matice $A \in \mathbb{C}^{n\times n}$ ve výsledné Jordanově normální formě je roven

$$\operatorname{rank}\left((A-\lambda I_n)^{k-1}\right) - 2\operatorname{rank}\left((A-\lambda I_n)^{k}\right) + \operatorname{rank}\left((A-\lambda I_n)^{k+1}\right).$$

Důkaz viz např. Horn and Johnson [1985]; Meyer [2000]. Vzoreček je možno též odvodit z myšlenky důkazy věty 10.39 o Jordanově normální formě.

Poznámka 10.44 (Zobecněné vlastní vektory)**.** Buď J Jordanova normální forma matice A, tedy $A = SJS^{-1}$ pro nějakou regulární matici S. Pokud je matice J diagonální, pak ve sloupcích matice S jsou vlastní vektory, které v pořadí odpovídají vlastním číslům matice A, umístěným na diagonále matice J. Jak ale interpretovat sloupce matice S (nazývané zobecněné vlastní vektory) pokud J není diagonální? Zaměřme se na speciální případ, kdy $J = J_k(\lambda)$ je jedna Jordanova buňka, obecný případ se vyřeší analogicky rozdělením na jednotlivé Jordanovy buňky. Nyní $A - \lambda I_k = SJ_k(0)S^{-1}$, z čehož

$$(A-\lambda I_k)^k = \left(SJ_k(0)S^{-1}\right)^k = SJ_k(0)^kS^{-1} = S0S^{-1} = 0.$$

Tudíž $(A-\lambda I_k)^kS = 0$, což znamená, že sloupce matice S patří do jádra matice $(A-\lambda I_k)^k$.

Souhrnem, buď $A \in \mathbb{C}^{n\times n}$ a $\lambda \in \mathbb{C}$ vlastní číslo. Prostor vlastních vektorů, příslušných k λ, tvoří jádro matice $A - \lambda I_n$. Prostor zobecněných vlastních vektorů, příslušných k λ, tvoří jádro matice $(A-\lambda I_n)^n$. Vlastních vektorů je plný počet (tedy n) pouze, když matice A je diagonalizovatelná. Na druhou stranu, (lineárně nezávislých) zobecněných vlastních vektorů je vždy n. To přirozeně souvisí s tím, že každá matice je podobná matici v Jordanově normální formě.

V následujících příkladech ukážeme několik aplikací Jordanovy normální formy.

Příklad 10.45 (Mocnění matice)**.** Již v příkladu 10.34 jsme zmínili využití diagonalizace pro počítání mocniny matice. Pomocí Jordanovy normální formy můžeme tvrzení zobecnit pro libovolné $A \in \mathbb{C}^{n\times n}$: Buď $A = SJS^{-1}$, pak

$$A^k = SJ^kS^{-1} = S\begin{pmatrix} J_{k_1}(\lambda_1)^k & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & J_{k_m}(\lambda_m)^k \end{pmatrix} S^{-1}.$$

Zde je třeba si trochu rozmyslet, jak se chovají mocniny Jordanových buněk $J_{k_i}(\lambda_i)^k$, $i = 1, \dots, m$. Asymptoticky pak dostaneme stejně jako pro diagonalizovatelné matice:

$$\lim_{k\to\infty} A^k = \begin{cases} 0, & \text{pokud } \rho(A) < 1, \\ \text{diverguje}, & \text{pokud } \rho(A) > 1, \\ \text{konverguje / diverguje}, & \text{pokud } \rho(A) = 1. \end{cases}$$

Příklad 10.46 (Maticová funkce)**.** Položme si otázku: Jak zavést maticovou funkci jako např. $\cos(A)$, e^A, atp.? Pro reálnou funkci $f\colon \mathbb{R} \to \mathbb{R}$ a matici $A \in \mathbb{R}^{n\times n}$ lze zavést $f(A)$ tak, že aplikujeme funkci na každou složku matice zvlášť,

$$f(A) = \begin{pmatrix} f(a_{11}) & \dots & f(a_{1n}) \\ \vdots & & \vdots \\ f(a_{n1}) & \dots & f(a_{nn}) \end{pmatrix}, \tag{10.4}$$

je sice možné, ale moc pěkných vlastností to mít nebude. Zkusme jiný přístup. Předpokládejme, že funkce $f\colon \mathbb{R} \to \mathbb{R}$ se dá vyjádřit nekonečným rozvojem $f(x) = \sum_{i=0}^{\infty} a_i x^i$; reálné analytické funkce jako např. $\sin(x)$, $\exp(x)$ aj. tento předpoklad splňují. Pak je tedy přirozené zavést $f(A) = \sum_{i=0}^{\infty} a_i A^i$. Mocnit matice již umíme, proto je-li $A = SJS^{-1}$, tak

$$f(A) = \sum_{i=0}^{\infty} a_i S J^i S^{-1} = S f(J) S^{-1}.$$

Dále snadno nahlédneme, že

$$f(J) := \begin{pmatrix} f(J_{k_1}(\lambda_1)) & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & f(J_{k_m}(\lambda_m)) \end{pmatrix}.$$

Chybí tedy ještě zavést obraz Jordanových buněk $J_{k_i}(\lambda_i)$. Pro $k_i = 1$ je to triviální, jde o matici řádu 1. Pro $k_i > 1$ je předpis složitější [Meyer, 2000]:

$$f(J_{k_i}(\lambda_i)) := \begin{pmatrix} f(\lambda_i) & f'(\lambda_i) & \dots & \frac{f^{(k_i-1)}(\lambda_i)}{(k_i-1)!} \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & f'(\lambda_i) \\ 0 & \dots & 0 & f(\lambda_i) \end{pmatrix}.$$

Například funkce $f(x) = x^2$ má maticové rozšíření $f(A) = A^2$, jedná se tedy o klasické maticové mocnění. Na druhou stranu, předpis (10.4) by naznačoval mocnit jednotlivé prvky matice zvlášť, což není to, co bychom chtěli. □

Příklad 10.47 (Soustava lineárních diferenciálních rovnic)**.** Uvažme tzv. soustavu lineárních diferenciálních rovnic:

$$u(t)' = Au(t) \tag{10.5}$$

kde $A \in \mathbb{R}^{n\times n}$. Cílem je nalézt neznámou funkci $u\colon \mathbb{R} \to \mathbb{R}^n$ splňující tuto soustavu pro určitou počáteční podmínku tvaru $u(t_0) = u_0$. Pro případ $n = 1$ je řešením diferenciální rovnice $u(t)' = au(t)$ funkce $u(t) = v \cdot e^{at}$, kde $v \in \mathbb{R}^n$ je libovolné (volí se tak, aby byla splněna počáteční podmínka). To nás motivuje (ale je za tím hlubší teorie) hledat řešení obecného případu ve tvaru

$$u(t) = (u_1(t), \dots, u_n(t)) = (v_1 e^{\lambda t}, \dots, v_n e^{\lambda t}) = e^{\lambda t} v,$$

kde v_i, λ jsou neznámé, $i = 1, \dots, n$. Dosazením $u(t) := e^{\lambda t} v$ do (10.5) dostaneme

$$\lambda e^{\lambda t} v = e^{\lambda t} A v, \quad \text{neboli} \quad \lambda v = Av.$$

To je přímo úloha výpočtu vlastních čísel a vektorů. Nechť matice A má vlastní čísla $\lambda_1, \dots, \lambda_n$ a vlastní vektory $x_1, \dots, x_n$. Pak řešení (10.5) je $u(t) = \sum_{i=1}^{n} \alpha_i e^{\lambda_i t} x_i$, kde $\alpha_i \in \mathbb{R}$ se získá z počátečních podmínek.

Uvažme konkrétní příklad:

$$\begin{aligned} u_1'(t) &= 7u_1(t) - 4u_2(t) \\ u_2'(t) &= 5u_1(t) - 2u_2(t) \end{aligned}$$

Matice $A = \begin{pmatrix} 7 & -4 \\ 5 & -2 \end{pmatrix}$ má vlastní čísla 2 a 3, jim odpovídají vlastní vektory $(4,5)^T$ a $(1,1)^T$. Řešení úlohy jsou tvaru

$$\begin{pmatrix} u_1(t) \\ u_2(t) \end{pmatrix} = \alpha_1 e^{2t} \begin{pmatrix} 4 \\ 5 \end{pmatrix} + \alpha_2 e^{3t} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \alpha_1, \alpha_2 \in \mathbb{R}.$$

Pro řešení lineárních diferenciálních rovnic je tedy znalost vlastních čísel a vlastních vektorů důležitá. Jordanova normální forma se použije, pokud matice A není diagonalizovatelná a deficience počtu vlastních vektorů se projeví tím, že ve vyjádření $u(t)$ nebudou α_i již skaláry, ale polynomy proměnné t stupňů odpovídajících násobnostem λ_i. □

10.5 Symetrické matice

Reálné symetrické matice mají řadu pozoruhodných vlastností týkající se vlastních čísel. Mezi stěžejní vlastnosti patří to, že jsou vždy diagonalizovatelné a jejich vlastní čísla jsou reálná.

Nejprve se podívejme na zobecnění transpozice a symetrických matic pro komplexní matice.

Definice 10.48 (Hermitovská matice a transpozice)**.** *Hermitovská transpozice*[5] matice $A \in \mathbb{C}^{n\times n}$ je matice $A^* := \overline{A}^T$. Matice $A \in \mathbb{C}^{n\times n}$ se nazývá *hermitovská*, pokud $A^* = A$.

Hermitovská transpozice má podobné vlastnosti jako klasická transpozice, např. $(A^*)^* = A$, $(\alpha A)^* = \overline{\alpha} A^*$, $(A+B)^* = A^* + B^*$, $(AB)^* = B^* A^*$.

Pomocí hermitovské transpozice můžeme unitární matice (rozšiřující pojem ortogonální matice pro komplexní matice, viz definice 8.50) definovat jako matice $Q \in \mathbb{C}^{n\times n}$ splňující $Q^*Q = I_n$.

Příklad 10.49. Z matic

$$\begin{pmatrix} 2 & 1+i \\ 1+i & 5 \end{pmatrix}, \quad \begin{pmatrix} 2 & 1+i \\ 1-i & 5 \end{pmatrix}$$

je první symetrická, ale ne hermitovská, a druhá je hermitovská, ale ne symetrická. Pro reálné matice oba pojmy splývají. □

Věta 10.50 (Vlastní čísla symetrických matic)**.** *Vlastní čísla reálných symetrických (resp. obecněji komplexních hermitovských) jsou reálná.*

Důkaz. Buď $A \in \mathbb{C}^{n\times n}$ hermitovská a buď $\lambda \in \mathbb{C}$ její libovolné vlastní číslo a $x \in \mathbb{C}^n$ příslušný vlastní vektor jednotkové velikosti, tj. $\|x\|_2 = 1$. Pak $Ax = \lambda x$, přenásobením x^* máme $x^*Ax = \lambda x^*x = \lambda$. Nyní

$$\lambda = x^*Ax = x^*A^*x = (x^*Ax)^* = \lambda^*.$$

Tedy $\lambda = \lambda^*$ a proto $\lambda \in \mathbb{R}$. □

Poznamenejme, že komplexní symetrické matice mohou mít ryze komplexní vlastní čísla.

Následující věta říká, že symetrické matice jsou diagonalizovatelné[6]. Navíc jsou diagonalizovatelné specifickým způsobem: z vlastních vektorů lze vybrat ortonormální systém, tedy matice podobnosti je ortogonální.

Věta 10.51 (Spektrální rozklad symetrických matic)**.** *Pro každou symetrickou matici $A \in \mathbb{R}^{n\times n}$ existuje ortogonální $Q \in \mathbb{R}^{n\times n}$ a diagonální $\Lambda \in \mathbb{R}^{n\times n}$ tak, že $A = Q\Lambda Q^T$.*

Důkaz. Matematickou indukcí podle n. Případ $n = 1$ je triviální: $\Lambda = A$, $Q = 1$.

Indukční krok $n \leftarrow n-1$. Buď λ vlastní číslo A a x odpovídající vlastní vektor normovaný $\|x\|_2 = 1$. Doplňme x, jakožto ortonormální systém (důsledek 8.24), na ortogonální matici $S := (x \mid \cdots)$. Protože $(A - \lambda I_n)x = o$, máme $(A - \lambda I_n)S = (o \mid \cdots)$, a tudíž $S^T(A - \lambda I_n)S = S^T(o \mid \cdots) = (o \mid \cdots)$. A jelikož je tato matice symetrická, máme

$$S^T(A - \lambda I_n)S = \begin{pmatrix} 0 & o^T \\ o & A' \end{pmatrix},$$

kde A' je nějaká symetrická matice řádu $n-1$. Podle indukčního předpokladu má spektrální rozklad $A' = Q'\Lambda' Q'^T$, kde Λ' je diagonální a Q' ortogonální. Matice a rovnost rozšíříme o jeden řád takto:

$$\begin{pmatrix} 0 & o^T \\ o & A' \end{pmatrix} = \begin{pmatrix} 1 & o^T \\ o & Q' \end{pmatrix} \begin{pmatrix} 0 & o^T \\ o & \Lambda' \end{pmatrix} \begin{pmatrix} 1 & o^T \\ o & Q'^T \end{pmatrix},$$

což snadno můžeme ověřit pronásobením. Označme

$$R := \begin{pmatrix} 1 & o^T \\ o & Q' \end{pmatrix}, \quad \Lambda'' := \begin{pmatrix} 0 & o^T \\ o & \Lambda' \end{pmatrix}.$$

[5] Charles Hermite (1822–1901) byl francouzský matematik. Dokázal mj., že e je transcendentní, to jest není kořenem žádné polynomu s racionálními koeficienty.

[6] Augustin Louis Cauchy, r. 1829.

Matice R je ortogonální (věta 8.54(4)), matice Λ'' diagonální. Nyní můžeme psát

$$S^T(A - \lambda I_n)S = R\Lambda'' R^T,$$

z čehož

$$A = SR\Lambda'' R^T S^T + \lambda I_n = SR\Lambda'' R^T S^T + \lambda SRR^T S^T = SR(\Lambda'' + \lambda I_n)R^T S^T.$$

Nyní máme hledaný rozklad $A = Q\Lambda Q^T$, kde $Q := SR$ je ortogonální matice a $\Lambda := \Lambda'' + \lambda I_n$ je diagonální. □

Podobně můžeme spektrálně rozložit hermitovské matice $A = Q\Lambda Q^*$, kde Q je unitární matice.

Poznámka 10.52 (Jiná forma spektrálního rozkladu)**.** Nechť symetrická $A \in \mathbb{R}^{n\times n}$ má vlastní čísla $\lambda_1, \dots, \lambda_n$ a odpovídající ortonormální vlastní vektory $x_1, \dots, x_n$. Tedy ve spektrálním rozkladu $A = Q\Lambda Q^T$ je $\Lambda_{ii} = \lambda_i$ a $Q_{*i} = x_i$. Pokud rozepíšeme Λ na součet jednodušších diagonálních matic

$$\Lambda = \begin{pmatrix} \lambda_1 & & & \\ & 0 & & \\ & & \ddots & \\ & & & 0 \end{pmatrix} + \begin{pmatrix} 0 & & & \\ & \lambda_2 & & \\ & & 0 & \\ & & & \ddots \end{pmatrix} + \dots + \begin{pmatrix} 0 & & & \\ & \ddots & & \\ & & 0 & \\ & & & \lambda_n \end{pmatrix} = \sum_{i=1}^n \lambda_i e_i e_i^T,$$

tak matici A lze vyjádřit jako

$$A = Q\Lambda Q^T = Q\left(\sum_{i=1}^n \lambda_i e_i e_i^T\right)Q^T = \sum_{i=1}^n \lambda_i Q e_i e_i^T Q^T = \sum_{i=1}^n \lambda_i Q_{*i} Q_{*i}^T = \sum_{i=1}^n \lambda_i x_i x_i^T.$$

Tvar $A = \sum_{i=1}^n \lambda_i x_i x_i^T$ je tak alternativní vyjádření spektrálního rozkladu, ve kterém matici A rozepisujeme na součet n matic hodnosti 0 nebo 1. Navíc, $x_i x_i^T$ je matice projekce na přímku $\operatorname{span}\{x_i\}$, tudíž z geometrického hlediska se na zobrazení $x \mapsto Ax$ můžeme dívat jako součet n zobrazení, kde v každém provádíme projekci na přímku (kolmou na ostatní) a škálování podle hodnoty λ_i.

Jedním z pěkných důsledků spektrálního rozkladu je následující, byť trochu teoretický, vzoreček na výpočet největšího a nejmenšího vlastního čísla[7]. Říká, že největší resp. nejmenší vlastní číslo se dá vyjádřit jako největší resp. nejmenší hodnota kvadratické funkce $f(x) = x^T Ax$ na jednotkové sféře.

Věta 10.53 (Courant–Fischer)**.** *Nechť $\lambda_1 \geq \dots \geq \lambda_n$ jsou vlastní čísla symetrické matice $A \in \mathbb{R}^{n\times n}$. Pak*

$$\lambda_1 = \max_{x:\|x\|_2=1} x^T Ax, \quad \lambda_n = \min_{x:\|x\|_2=1} x^T Ax.$$

Důkaz. Pouze pro λ_1, druhá část je analogická.

Nerovnost „$\leq$“: Buď x_1 vlastní vektor příslušný k λ_1 normovaný $\|x_1\|_2 = 1$. Pak $Ax_1 = \lambda_1 x_1$. Přenásobením x_1^T zleva dostaneme

$$\lambda_1 = \lambda_1 x_1^T x_1 = x_1^T A x_1 \leq \max_{x:\|x\|_2=1} x^T Ax.$$

Nerovnost „$\geq$“: Buď $x \in \mathbb{R}^n$ libovolný vektor takový, že $\|x\|_2 = 1$. Označme $y := Q^T x$, pak $\|y\|_2 = 1$ (věta 8.54(2)). S využitím spektrálního rozkladu $A = Q\Lambda Q^T$ dostaneme:

$$x^T Ax = x^T Q\Lambda Q^T x = y^T \Lambda y = \sum_{i=1}^n \lambda_i y_i^2 \leq \sum_{i=1}^n \lambda_1 y_i^2 = \lambda_1 \|y\|_2^2 = \lambda_1. \qquad \square$$

[7] Ernst Fischer odvodil vzorec r. 1905, Richard Courant ho zobecnil pro nekonečně rozměrné operátory r. 1920. Jejich vzorec zahrnuje i mezilehlá vlastní čísla $\lambda_2, \dots, \lambda_{n-1}$, ale pro ně je to trochu komplikovanější. Tato jednodušší verze se také někdy nazývá Rayleigh-Ritzova věta.

10.6 Teorie nezáporných matic

Perronova–Frobeniova teorie nezáporných matic[8)] je pokročilá teorie kolem vlastních čísel nezáporných matic. Uvedeme jen základní výsledek bez důkazu.

Věta 10.54 (Perronova)**.**

(1) Buď $A \in \mathbb{R}^{n\times n}$ nezáporná matice (tj. $a_{ij} \geq 0$ pro všechna i, j). Pak největší (v absolutní hodnotě) vlastní číslo je reálné nezáporné a příslušný vlastní vektor je nezáporný (ve všech složkách).

(2) Buď $A \in \mathbb{R}^{n\times n}$ kladná matice (tj. $a_{ij} > 0$ pro všechna i, j). Pak největší (v absolutní hodnotě) vlastní číslo je reálné kladné, je jediné, a příslušný vlastní vektor je kladný (ve všech složkách). Navíc žádnému jinému vlastnímu číslu neodpovídá nezáporný vlastní vektor.

Důkaz. Viz např. [Meyer, 2000]. □

Příklad 10.55 (Markovovy řetězce)**.** Jedním z využití mocnin matice (příklad 10.45) a trochu i teorie nezáporných matic jsou Markovovy řetězce.[9)] Buď $x \in \mathbb{R}^n$ stavový vektor, čili x_i udává hodnotu stavu i. Buď $A \in \mathbb{R}^{n\times n}$ matice s hodnotami $a_{ij} \in [0, 1]$ tak, že součet hodnot v každém sloupci je roven 1. Na matici A budeme nahlížet jako na přechodovou matici, to jest, hodnota a_{ij} je pravděpodobnost přechodu ze stavu j do stavu i. Pak Ax udává nový stavový vektor po jednom kroku daného procesu. Zajímá nás, jak se bude stavový vektor vyvíjet v čase a zda se nějak ustálí. K tomu se bude hodit zjistit něco více o matici A. Přímo z definice platí, že $A^T e = e$, kde $e = (1, \ldots, 1)^T$. Tudíž e je vlastní vektor A^T a 1 vlastní číslo A^T (a tedy i A). Navíc se dá ukázat, že žádné jiné vlastní číslo není v absolutní hodnotě větší (viz poznámka 10.58), tudíž 1 je vlastní číslo matice A z Perronovy věty a odpovídající vlastní vektor je nezáporný.

Další analýzu ilustrujeme na konkrétním příkladu: Migrace obyvatel USA v sektorech město–předměstí–venkov probíhá každoročně podle vzorce:

z města:	96% zůstane, 3% do předměstí, 1% na venkov,
z předměstí:	1% do města, 98% zůstane, 1% na venkov,
z venkova:	1.5% do města, 0.5% do předměstí, 98% zůstane.

Počáteční stav: 58 mil. obyvatel ve městě, 142 mil. na předměstí a 60 mil. na venkově. Jak se bude situace vyvíjet v čase? Označme

$$A := \begin{pmatrix} 0.96 & 0.01 & 0.015 \\ 0.03 & 0.98 & 0.005 \\ 0.01 & 0.01 & 0.98 \end{pmatrix}, \quad x_0 = (58, 142, 60)^T.$$

V nultém roce je rozmístění obyvatel dáno vektorem x_0 a v prvním roce vektorem Ax_0, protože Ax_0 má v první složce výsledný počet obyvatel ve městě po jednoroční migraci a podobně v druhé a třetí složce pro předměstí a venkov. Vývoj v čase probíhá tedy takto: x_0, Ax_0, A^2x_0, A^3x_0, $\ldots$,$A^\infty x_0$. Diagonalizací spočítáme

$$A = Q \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0.95 & 0 \\ 0 & 0 & 0.97 \end{pmatrix} Q^{-1}, \quad \text{kde} \quad Q \approx \begin{pmatrix} -0.393 & -0.707 & 0.154 \\ -0.729 & 0.707 & -0.772 \\ -0.561 & 0 & 0.617 \end{pmatrix}.$$

z čehož

$$\lim_{k\to\infty} A^k = Q \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} Q^{-1} = Q_{*1} Q_{1*}^{-1} = \begin{pmatrix} 0.23 & 0.23 & 0.23 \\ 0.43 & 0.43 & 0.43 \\ 0.33 & 0.33 & 0.33 \end{pmatrix}.$$

[8)] Základní věta je od německého matematika Oskara Perrona z r. 1907, a byla rozšířena Ferdinandem Georgem Frobeniem r. 1912.

[9)] Andrej Andrejevič Markov (1856–1922), ruský matematik. Tvrdí se, že první aplikací Markovových řetězců byla analýza distribuce samohlásek a souhlásek v Puškinově Evženu Oněginovi, kdy vytvořil model předpokládající, že pravděpodobnost výskytu samohlásky či souhlásky na dané pozici závisí jen na posledním předchozím znaku, ale na žádném z předešlých už nikoli. Tato jednoduchá Markovova vlastnost se pak stala základem Markovových řetězců. Více o této lingvistické aplikaci se dočtete např. v *Učiteli matematiky*, 1–2(77–78), roč. 19, 2010–2011.

Tedy (bez ohledu na počáteční stav x_0) se rozložení obyvatelstva ustálí na hodnotách: 23% ve městě, 43% předměstí, 33% venkov.

Tento příklad ilustruje ještě jednu věc. Má-li vlastní číslo 1 násobnost jedna a ostatní vlastní čísla jsou v absolutní hodnotě menší (což nastane např. pokud A je kladná), tak $A^\infty = Q_{*1}Q_{1*}^{-1}$, kde Q_{*1} je vlastní vektor A k číslu 1, a $Q_{1*}^{-1} = (1, \dots, 1)$ je vlastní vektor A^T k číslu 1. Tudíž sloupce A^∞ budou opět stejné, a výsledné rozložení odpovídá složkám vlastního vektoru Q_{*1} (ty jsou kladné, podle Perronovy věty). Takže v tomto případě stačí jen spočítat kladný vektor z $\operatorname{Ker}(A - I_n)$, což odpovídá intuici, že ustálené rozložení reprezentované vektorem v má splňovat $Av = v$. □

10.7 Výpočet vlastních čísel

Jak jsme již zmínili (poznámky k větě 10.16), vlastní čísla se počítají veskrze numerickými iteračními metodami, a hledat je jako kořeny charakteristického polynomu není efektivní postup. V této sekci ukážeme jednoduchý odhad na vlastní čísla, a jednoduchou metodu na výpočet největšího vlastního čísla. Další metodu, populární QR algoritmus, probereme v sekci 13.3. Pro velmi přesný výpočet vlastních čísel symetrických (zvláště tzv. positivně definitních) matic se používá také Jacobiho metoda (viz např. [Rohn, 2004]) a pro velké řídké symetrické matice např. Lanczosova metoda[10)] (viz [Meyer, 2000]).

Nejprve uvedeme jednoduchý odhad pro vlastní čísla, tzv. Gerschgorinovy disky[11)].

Věta 10.56 (Gerschgorinovy disky)**.** *Každé vlastní číslo λ matice $A \in \mathbb{C}^{n\times n}$ leží v kruhu o středu a_{ii} a poloměru $\sum_{j\neq i} |a_{ij}|$ pro nějaké $i \in \{1, \dots, n\}$.*

Důkaz. Buď λ vlastní číslo a x odpovídající vlastní vektor, tedy $Ax = \lambda x$. Nechť i-tá složka x má největší absolutní hodnostu, tj. $|x_i| = \max_{k=1,\dots,n} |x_k|$. Protože i-tá rovnice má tvar $\sum_{j=1}^n a_{ij}x_j = \lambda x_i$, vydělením $x_i \neq 0$ dostáváme

$$\lambda = a_{ii} + \sum_{j\neq i} a_{ij}\frac{x_j}{x_i}$$

a tím pádem

$$|\lambda - a_{ii}| = \Big|\sum_{j\neq i} a_{ij}\frac{x_j}{x_i}\Big| \leq \sum_{j\neq i} |a_{ij}|\frac{|x_j|}{|x_i|} \leq \sum_{j\neq i} |a_{ij}|. \qquad \square$$

Příklad 10.57. Mějme

$$A = \begin{pmatrix} 2 & 1 & 0 \\ -2 & 5 & 1 \\ -1 & -2 & -3 \end{pmatrix}.$$

Vlastní čísla matice A jsou $\lambda_1 = -2.78$, $\lambda_2 = 3.39 + 0.6\,i$, $\lambda_3 = 3.39 - 0.6\,i$. Spolu s Gerschgorinovými disky jsou nakresleny dole:

[10)] Algoritmus pochází z r. 1950 od maďarského matematika Cornelia Lanczose.

[11)] Pochází z r. 1931, a autorem je běloruský matematik Semyon Aranovich Gerschgorin. Nicméně tvrzení bylo známo už dříve: L. Lévy (1881), H. Minkowski (1900), J. Hadamard (1903).

Vidíme, že ne každý kruh obsahuje nějaké vlastní číslo. Nicméně platí [Meyer, 2000], že v každé komponentě souvislosti je tolik vlastních čísel, z kolika kruhů daná komponenta vznikla. □

Věta dává jednoduchý ale hrubý odhad na velikost vlastních čísel (existují i vylepšení, např. Cassiniho ovály aj.). Nicméně, v některých aplikacích může takovýto odhad postačovat.

Poznámka 10.58 (Tři použití Gerschgorinových disků)**.**
1) *Kriterium pro zastavení výpočtu iteračních metod.* Např. Jacobiho metoda na výpočet vlastních čísel spočívá v postupném zmenšování nediagonálních prvků symetrické matice, takže matice konverguje k diagonální matici. Gerschgorinovy disky pak dávají horní mez na přesnost vypočtených vlastních čísel. Pokud např. matice $A \in \mathbb{R}^{n\times n}$ je skoro diagonální v tom smyslu, že všechny mimodiagonální prvky jsou menší než 10^{-k} pro nějaké $k \in \mathbb{N}$, pak diagonální prvky aproximují vlastní čísla s přesností $10^{-k}(n-1)$. Například matice

$$A = \begin{pmatrix} 7.0001 & 0.0001 & -0.0002 \\ 0.0001 & 5 & 0.0003 \\ -0.0002 & 0.0003 & 1.999 \end{pmatrix}$$

má vlastní čísla 7.0001 ± 0.0003, 5 ± 0.0004 a 1.999 ± 0.0005.

2) *Diagonálně dominantní matice.* Gerschgorinovy disky dávají také následující postačující podmínku[12)] pro regularitu matice $A \in \mathbb{C}^{n\times n}$: $|a_{ii}| > \sum_{j\neq i} |a_{ij}|$ $\forall i = 1, \dots, n$. V tomto případě totiž disky neobsahují počátek a proto nula není vlastním číslem A. Matice s touto vlastností se nazývají diagonálně dominantní.
3) *Markovovy matice.* Buď A Markovova matice z příkladu 10.55. Všechny Gerschgorinovy disky matice A^T mají střed v bodě v intervalu $[0, 1]$ a pravým krajem protínají hodnotu 1 na reálné ose. To dokazuje, že $\rho(A) \leq 1$, a tudíž číslo 1 je skutečně v absolutní hodnotě největším vlastním číslem matice A.

Nyní ukážeme jednoduchou metodu na výpočet dominantního vlastního čísla.[13)] Přes svou jednoduchost se stala základem některých numerických metod jako je např. inverzní mocninná metoda[14)], nebo metoda Rayleighova podílu[15)] pro symetrické matice.

Algoritmus 10.59 (Mocninná metoda)**.** Buď $A \in \mathbb{C}^{n\times n}$.

1: Zvol $o \neq x_0 \in \mathbb{C}^n$, $i := 1$,
2: **while not** splněna ukončovací podmínka **do**
3: $\quad y_i := Ax_{i-1}$,
4: $\quad x_i := \frac{1}{\|y_i\|_2} y_i$,
5: $\quad i := i + 1$,
6: **end while**

Výstup: $\lambda_1 := x_{i-1}^T y_i$ je odhad vlastního čísla, $v_1 := x_i$ je odhad příslušného vlastního vektoru.

Příklad 10.60. Mějme

$$A = \begin{pmatrix} 2 & 4 & 2 \\ 4 & 2 & 2 \\ 2 & 2 & -1 \end{pmatrix}, \quad x_0 = (1, 0, 1)^T.$$

[12)] Tzv. Lévyho–Desplanquesova věta.
[13)] Mocninou metodu, anglicky *power method*, představil americký matematik a fyzik Richard Edler von Mises r. 1929.
[14)] Autorem je německý matematik Helmut Wielandt, z r. 1944, anglicky *inverse iteration*.
[15)] Autorem anglický fyzik John William Strutt, lord Rayleigh, nositel Nobelovy ceny za fyziku (1904), anglicky *Rayleigh quotient iteration*.

Postup výpočtu a zdrojový kód pro Matlab / Octave:

i	$\frac{1}{\|x_i\|_\infty} x_i$	$x_{i-1}^T y_i$
0	$(1.00, 0.00, 1.00)^T$	–
1	$(0.67, 1.00, 0.17)^T$	5
2	$(1.00, 0.88, 0.56)^T$	6.32
3	$(0.97, 1.00, 0.47)^T$	6.94
4	$(1.00, 1.00, 0.50)^T$	7

```
A=[2 4 2; 4 2 2; 2 2 -1];
x=[1;0;1];
for i=1:4
  y=A*x;
  (y'*x),
  x=y/norm(y);
  x/max(abs(x)),
end
```

□

Metodu ukončíme ve chvíli, když se hodnota $x_{i-1}^T y_i$ resp. vektor x_i ustálí; potom $x_i \approx x_{i-1}$ je odhad vlastního vektoru a $x_{i-1}^T y_i = x_{i-1}^T A x_{i-1} \approx x_{i-1}^T \lambda x_{i-1} \approx \lambda$ odhad odpovídajícího vlastního čísla. Metoda může být pomalá, špatně se odhaduje chyba a míra konvergence a navíc velmi záleží na počáteční volbě x_0. Na druhou stranu je robustní (zaokrouhlovací chyby nemají velký vliv) a snadno aplikovatelná na velké řídké matice. Ne vždy konverguje, ale za určitých předpokladů se to dá zajistit.

Tvrzení 10.61 (Konvergence mocninné metody)**.** *Buď $A \in \mathbb{R}^{n\times n}$ s vlastními čísly $|\lambda_1| > |\lambda_2| \geq \ldots \geq |\lambda_n|$ a odpovídajícími lineárně nezávislými vektory $v_1, \ldots, v_n$ velikosti 1. Nechť x_0 má nenulovou souřadnici ve směru v_1. Pak x_i konverguje (až na znaménko) k vlastnímu vektoru v_1 a $x_{i-1}^T y_i$ konverguje k vlastnímu číslu λ_1.*

Důkaz. Nechť $x_0 = \sum_{j=1}^n \alpha_j v_j$, kde $\alpha_1 \neq 0$ podle předpokladu. Pak $x_i = \frac{1}{\|A^i x_0\|} A^i x_0$ a

$$A^i x_0 = A^i \Big(\sum_{j=1}^n \alpha_j v_j \Big) = \sum_{j=1}^n \alpha_j A^i v_j = \sum_{j=1}^n \alpha_j \lambda_j^i v_j = \lambda_1^i \Big(\alpha_1 v_1 + \sum_{j\neq 1} \alpha_j \Big(\frac{\lambda_j}{\lambda_1}\Big)^i v_j \Big).$$

Protože vektory x_i postupně normujeme, násobek λ_1^i nás nemusí zajímat. Zbylý vektor postupně konverguje k $\alpha_1 v_1$, protože $\left|\frac{\lambda_j}{\lambda_1}\right| < 1$ a tudíž $\left|\frac{\lambda_j}{\lambda_1}\right|^i \to 0$ pro $i \to \infty$.

Nyní předpokládejme, že x_i již dobře aproximuje vlastní vektor v_1. Pak $x_{i-1}^T y_i = x_{i-1}^T A x_{i-1} = x_{i-1}^T \lambda_1 x_{i-1} = \lambda_1 \|x_{i-1}\|_2^2 = \lambda_1$. □

Z důkazu věty vidíme, že rychlost konvergence výrazně závisí na poměru $\left|\frac{\lambda_2}{\lambda_1}\right|$. Poznamenejme, že vzhledem k tomu, že λ_1 je dominantní vlastní číslo, tak pro reálnou matici A musí být reálné a v_1 rovněž tak (tvrzení 10.13).

Mocninná metoda počítá jen dominantní vlastní číslo a vektor. Následující technika ale umožňuje jednoduchou transformací vynulovat jedno vlastní číslo, např. to dominantní, a pak můžeme rekurzivně dopočítat mocninnou metodou i zbylá vlastní čísla. Nejprve uvádíme jednoduchou verzi pro symetrické matice, a potom v poznámce 10.63 jak postupovat obecně.

Tvrzení 10.62 (O deflaci vlastního čísla)**.** *Buď $A \in \mathbb{R}^{n\times n}$ symetrická, $\lambda_1, \ldots, \lambda_n$ její vlastní čísla a $v_1, \ldots, v_n$ odpovídající ortonormální vlastní vektory. Pak matice $A - \lambda_1 v_1 v_1^T$ má vlastní čísla $0, \lambda_2, \ldots, \lambda_n$ a vlastní vektory $v_1, \ldots, v_n$.*

Důkaz. Podle poznámky 10.52 lze psát $A = \sum_{i=1}^n \lambda_i v_i v_i^T$. Pak $A - \lambda_1 v_1 v_1^T = 0 v_1 v_1^T + \sum_{i=2}^n \lambda_i v_i v_i^T$, což je spektrální rozklad matice $A - \lambda_1 v_1 v_1^T$. □

Poznámka 10.63 (K deflaci vlastního čísla obecné matice)**.** Buď λ vlastní číslo a x odpovídající vlastní vektor matice $A \in \mathbb{R}^{n\times n}$. Doplňme x na regulární matici S tak, aby $S_{*1} = x$. Pak

$$S^{-1}AS = S^{-1}A(x \mid \cdots) = S^{-1}(\lambda x \mid \cdots) = (\lambda e_1 \mid \cdots) = \begin{pmatrix} \lambda & \cdots \\ o & A' \end{pmatrix}.$$

Z podobnosti má matice A' stejná vlastní čísla jako A, pouze λ má o jedna menší násobnost. Tudíž zbývající vlastní čísla matice A můžeme najít pomocí A'.

Příklad 10.64 (Vyhledávač Google™ a PageRank[16]). Uvažujme webovou síť s těmito parametry:

$$
\begin{aligned}
&N \text{ webových stránek},\\
&a_{ij} = \begin{cases} 1 & j\text{-tá stránka odkazuje na } i\text{-tou},\\ 0 & \text{jinak},\end{cases}\\
&b_j = \text{počet odkazů z } j\text{-té stránky},\\
&x_i = \text{důležitost } i\text{-té stránky}.
\end{aligned}
$$

Cílem je stanovit důležitosti $x_1, \dots, x_N$ jednotlivých stránek. Základní myšlenka autorů googlovského PageRanku spočívá ve stanovení důležitosti i-té stránky tak, aby byla úměrná součtu důležitosti stránek na ni odkazujících. Řešíme tedy rovnici $x_i = \sum_{j=1}^{n} \frac{a_{ij}}{b_j} x_j$, $i = 1, \dots, N$. Maticově $A'x = x$, kde $a'_{ij} := \frac{a_{ij}}{b_j}$. Tedy x je vlastní vektor matice A příslušný vlastnímu číslu 1. Vlastní číslo 1 je dominantní, což snadno nahlédneme z Gerschgorinových disků pro matici A'^T (součet sloupců matice A' je roven 1, tedy všechny Gerschgorinovy disky mají nejpravější konec v bodě 1). Podle Perronovy věty 10.54 je vlastní vektor x nezáporný. Ten pak normujeme tak, aby $\|x\|_\infty = 10$, to jest, aby složky byly v intervalu $[0, 10]$. Nakonec složky zaokrouhlíme, aby měly hodnoty v rozmezí $0, 1, 2, \dots, 10$.

V praxi je matice A' obrovská, řádově $\approx 10^{10}$, a zároveň řídká (většina hodnot jsou nuly). Proto se na výpočet x hodí mocninná metoda, stačí ≈ 100 iterací. Prakticky se navíc ještě matice A' trochu upravuje, aby byla tzv. stochastická, aperiodická a ireducibilní (vadilo by např. pokud by z nějaké stránky nebyl odkaz na žádnou jinou), více viz [Langville and Meyer, 2006; Tůma, 2003].

Na matici A' se můžeme dívat i z pohledu Markovových řetězců (příklad 10.55). Pokud budeme symbolicky procházet po webové síti a se stejnou pravděpodobností se na každé stránce rozhodneme, kam pokračovat, tak A' odpovídá přechodové matici a hledaný vektor x rovnovážnému stavu.

Příklady Page ranku (k roku 2010):

`www.google.com`	10
`www.cuni.cz`	8
`www.mff.cuni.cz`	7
`kam.mff.cuni.cz`	6
`kam.mff.cuni.cz/~hladik`	4

□

Problémy

10.1. Buď $A \in \mathbb{R}^{n\times n}$ singulární. Dokažte, že existuje posloupnost regulárních matic A_i, $i = 1, \dots,$ konvergující po složkách k A. (Množina regulárních matic je tzv. *hustá* v $\mathbb{R}^{n\times n}$.)

10.2. Buď $A \in \mathbb{R}^{n\times n}$ nediagonalizovatelná. Dokažte, že existuje posloupnost diagonalizovatelných matic A_i, $i = 1, \dots,$ konvergující po složkách k A. (Množina diagonalizovatelných matic je tedy *hustá* v $\mathbb{R}^{n\times n}$.)

10.3. Ukažte přímo (bez Jordanovy normální formy), že algebraická násobnost vlastního čísla je vždy větší nebo rovna geometrické násobnosti.

10.4. Dokažte přímo Cayleyho–Hamiltonovu větu 10.18 pro diagonalizovatelné matice.

10.5. Dokažte Cayleyho–Hamiltonovu větu za použití Jordanovy normální formy.

10.6. Ukažte, že každá matice $A \in \mathbb{R}^{n\times n}$ má invariantní podprostor dimenze 1 nebo 2.

10.7. Dokažte následující odhad pro hodnost druhé mocniny matice $A \in \mathbb{R}^{n\times n}$

$$\operatorname{rank}(A) - \tfrac{n}{2} \le \operatorname{rank}(A^2) \le \operatorname{rank}(A).$$

Hint: Využijte faktu, že A je podobná matici v Jordanově normální formě.

10.8. Domyslete detaily důkazu věty 10.39 o Jordanově normální formě:

[16] Z r. 1997, autory jsou Sergey Brin a Larry Page.

(a) Ukažte, že pro každou matici $A \in \mathbb{C}^{n\times n}$ existuje $p \in \mathbb{N}$ takové, že

$$\mathrm{Ker}(A) \subsetneqq \mathrm{Ker}(A^2) \subsetneqq \cdots \subsetneqq \mathrm{Ker}(A^p) = \mathrm{Ker}(A^{p+1}) = \dots$$

(b) Buď $v \in \mathrm{Ker}(A^p) \setminus \mathrm{Ker}(A^{p-1})$. Ukažte, že vektory $v, Av, \dots, A^{p-1}v$ jsou lineárně nezávislé.

(c) Dokažte vzoreček z poznámky 10.43 o velikostech a počtu Jordanových buněk.

Shrnutí ke kapitole 10. Vlastní čísla

Vlastní čísla a vlastní vektory matice A poskytují o matici a o lineárním zobrazení $x \mapsto Ax$ velmi důležité informace. Vlastní vektory reprezentují invariantní směry, které se zobrazí samy na sebe, a vlastní čísla představují míru škálování v těchto směrech. Vlastní čísla tedy celkem dobře popisují jak moc lineární zobrazení $x \mapsto Ax$ degeneruje objekty a co se děje, když zobrazení iterujeme.

Vlastních čísel matice $A \in \mathbb{C}^{n\times n}$ je právě n včetně násobností, a (lineárně nezávislých) vlastních vektorů je nanejvýš n. Má-li vlastní číslo násobnost k, pak mu odpovídá maximálně k vlastních vektorů. Různá vlastní čísla pak mají lineárně nezávislé vlastní vektory.

Elementární úpravy mění vlastní čísla matice, ale úprava, která je nemění, je podobnost. Geometricky podobnost znamená jen změnu souřadného systému. Každá matice je podobná matici v jednodušším tvaru – Jordanově normálním tvaru. Ten může mít dokonce podobu diagonální matice (pak je původní matice diagonalizovatelná). To nastává právě tehdy, když matice má plný počet vlastních vektorů (počet Jordanových buněk je roven počtu vlastních vektorů).

Důležitou třídou matic jsou symetrické matice. Mají tři podstatné vlastnosti: (1) jsou vždy diagonalizovatelné, (2) mají reálná vlastní čísla, (3) vlastní vektory jdou vybrat tak, aby na sebe byly kolmé. Příslušný spektrální rozklad je pak velmi užitečný nástroj.

Další třídou matic se speciálními vlastnostmi jsou nezáporné matice: největší vlastní číslo v absolutní hodnotě leží na reálné ose vpravo od počátku a odpovídající vlastní vektor je nezáporný.

Problém výpočtu vlastních čísel a výpočtu kořenů polynomu jsou na sebe převoditelné (skrze charakteristický polynom a matici společnici). Na výpočet vlastních čísel nelze jednoduše použít Gaussovu eliminaci – veškeré používané metody jsou iterativní, jako například mocninná metoda. Šikovné jsou i různé odhady jako jsou například Gerschgorinovy disky.

Kapitola 11

Positivně (semi-)definitní matice

Definice 11.1 (Positivně (semi-)definitní matice)**.** Buď $A \in \mathbb{R}^{n\times n}$ symetrická. Pak A je *positivně semidefinitní*, pokud $x^TAx \geq 0$ pro všechna $x \in \mathbb{R}^n$, a A je *positivně definitní*, pokud $x^TAx > 0$ pro všechna $x \neq o$.

Zřejmě, je-li A positivně definitní, pak je i positivně semidefinitní.

Kromě positivně (semi-)definitních matic lze zavést i negativně (semi-)definitní matice pomocí obrácené nerovnosti. Zabývat se jimi nebudeme, protože A je negativně (semi-)definitní právě tehdy, když $-A$ je positivně (semi-)definitní, čili vše se redukuje na základní případ.

Poznámka 11.2. Definice dává smysl i pro nesymetrické matice, ale ty můžeme snadno zesymetrizovat úpravou $\frac{1}{2}(A + A^T)$, neboť

$$x^T\tfrac{1}{2}(A + A^T)x = \tfrac{1}{2}x^TAx + \tfrac{1}{2}x^TA^Tx = \tfrac{1}{2}x^TAx + \left(\tfrac{1}{2}x^TAx\right)^T = x^TAx.$$

Tedy pro testování podmínky lze ekvivalentně použít symetrickou matici $\frac{1}{2}(A+A^T)$. Omezení na symetrické matice je tudíž bez újmy na obecnosti. Důvod, proč se omezujeme na symetrické matice, je, že řada testovacích podmínek funguje pouze pro symetrické matice.

Příklad 11.3. Příkladem positivně semidefinitní matice je 0_n. Příkladem positivně definitní matice je I_n, neboť $x^TAx = x^TI_nx = x^Tx = \|x\|_2^2 > 0$ pro všechna $x \neq o$. □

Poznámka 11.4 (Nutná podmínka pro positivní (semi-)definitnost)**.** Buď $A \in \mathbb{R}^{n\times n}$ symetrická matice. Aby byla positivně semidefinitní, musí podle definice $x^TAx \geq 0$ pro všechna $x \in \mathbb{R}^n$. Postupným dosazením $x = e_i$, $i = 1, \ldots, n$, dostaneme $x^TAx = e_i^TAe_i = a_{ii} \geq 0$. Tím pádem, positivně semidefinitní matice musí mít nezápornou diagonálu a positivně definitní matice dokonce kladnou.

Poznámka 11.5. Matice $A = (a) \in \mathbb{R}^{1\times 1}$ je positivně semidefinitní právě tehdy, když $a \geq 0$, a positivně definitní právě tehdy, když $a > 0$. Tím pádem se můžeme dívat na positivní semidefinitnost jako na zobecnění pojmu nezápornosti z čísel na matice. Proto se také positivní semidefinitnost matice $A \in \mathbb{R}^{n\times n}$ značívá $A \succeq 0$ (narozdíl od $A \geq 0$, což se používá pro nezápornost v každé složce).

Tvrzení 11.6 (Vlastnosti positivně definitních matic)**.**

(1) Jsou-li $A, B \in \mathbb{R}^{n\times n}$ positivně definitní, pak i $A + B$ je positivně definitní,

(2) Je-li $A \in \mathbb{R}^{n\times n}$ positivně definitní a $\alpha > 0$, pak i αA je positivně definitní,

(3) Je-li $A \in \mathbb{R}^{n\times n}$ positivně definitní, pak i A^{-1} je positivně definitní.

Důkaz. První dvě vlastnosti jsou triviální, dokážeme pouze tu třetí.

Nejprve ověříme regularitu matice A. Buď x řešení $Ax = o$. Pak $x^TAx = x^To = 0$. Z předpokladu musí $x = o$.

Nyní ukážeme positivní definitnost. Sporem nechť existuje $x \neq o$ takové, že $x^TA^{-1}x \leq 0$. Pak

$$x^TA^{-1}x = x^TA^{-1}AA^{-1}x = y^TAy \leq 0,$$

kde $y = A^{-1}x \neq o$. To je spor, neboť A je positivně definitní. □

Analogie věty platí i pro positivně semidefinitní matice. Část (1) platí beze změny, část (2) platí pro všechna $\alpha \geq 0$, ale část (3) už obecně neplatí.

Součinem positivně definitních matic se zabýváme v příkladu 12.19.

Následující věta dává důležitou charakterizaci positivně definitních matic jednak pomocí vlastních čísel a jednak pomocí tvaru U^TU; s tímto výrazem jsme se skrytě setkali už např. v metodě nejmenších čtverců (sekce 8.5) u soustavy normálních rovnic $A^TAx = A^Tb$, nebo obecně při explicitním vyjádření ortogonální projekce v $\mathbb{R}^n$ (věta 8.43).

Věta 11.7 (Charakterizace positivní definitnosti)**.** *Buď $A \in \mathbb{R}^{n\times n}$ symetrická. Pak následující podmínky jsou ekvivalentní:*

(1) A je positivně definitní,

(2) vlastní čísla A jsou kladná,

(3) existuje matice $U \in \mathbb{R}^{m\times n}$ hodnosti n taková, že $A = U^TU$.

Důkaz. Implikace (1) $\Rightarrow$ (2): Sporem nechť existuje vlastní číslo $\lambda \leq 0$, a x je příslušný vlastní vektor s eukleidovskou normou 1. Pak $Ax = \lambda x$ implikuje $x^TAx = \lambda x^Tx = \lambda \leq 0$. To je spor s positivní definitností A.

Implikace (2) $\Rightarrow$ (3): Protože A je symetrická, má spektrální rozklad $A = Q\Lambda Q^T$, kde Λ je diagonální matice s prvky $\lambda_1, \dots, \lambda_n > 0$. Definujme matici Λ' jako diagonální s prvky $\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n} > 0$. Pak hledaná matice je $U = \Lambda' Q^T$, neboť $U^TU = Q\Lambda'\Lambda' Q^T = Q\Lambda'^2 Q^T = Q\Lambda Q^T = A$. Uvědomme si, že U má hodnost n a je tudíž regulární, neboť je součinem dvou regulárních matic.

Implikace (3) $\Rightarrow$ (1): Sporem nechť $x^TAx \leq 0$ pro nějaké $x \neq o$. Pak $0 \geq x^TAx = x^TU^TUx = (Ux)^TUx = \langle Ux, Ux\rangle = \|Ux\|_2^2$. Tedy musí $Ux = o$, ale sloupce U jsou lineárně nezávislé, a tak $x = o$, spor. □

Pro positivní semidefinitnost máme následující charakterizaci. Důkaz již neuvádíme, je analogický.

Věta 11.8 (Charakterizace positivní semidefinitnosti)**.** *Buď $A \in \mathbb{R}^{n\times n}$ symetrická. Pak následující podmínky jsou ekvivalentní:*

(1) A je positivně semidefinitní,

(2) vlastní čísla A jsou nezáporná,

(3) existuje matice $U \in \mathbb{R}^{m\times n}$ taková, že $A = U^TU$.

11.1 Metody na testování positivní definitnosti

Nyní se zaměříme na konkrétní metody pro testování positivní definitnosti. Řada z nich vychází z následujícího rekurentního vztahu. Povšimněme si, že pro testování positivní semidefinitnosti použít ani jednoduše přizpůsobit nelze.

Věta 11.9 (Rekurentní vzoreček na testování positivní definitnosti)**.** *Symetrická matice $A = \left(\begin{smallmatrix} \alpha & a^T \\ a & \tilde{A} \end{smallmatrix}\right)$, kde $\alpha \in \mathbb{R}$, $a \in \mathbb{R}^{n-1}$, $\tilde{A} \in \mathbb{R}^{(n-1)\times(n-1)}$ je positivně definitní právě tehdy, když $\alpha > 0$ a $\tilde{A} - \frac{1}{\alpha}aa^T$ je positivně definitní.*

Důkaz. Implikace „$\Rightarrow$“. Buď A positivně definitní. Pak $x^TAx > 0$ pro všechna $x \neq o$, tedy speciálně pro $x = e_1$ dostáváme $\alpha = e_1^TAe_1 > 0$. Dále, buď $\tilde{x} \in \mathbb{R}^{n-1}$, $\tilde{x} \neq o$. Pak

$$\tilde{x}^T\big(\tilde{A} - \tfrac{1}{\alpha}aa^T\big)\tilde{x} = \tilde{x}^T\tilde{A}\tilde{x} - \tfrac{1}{\alpha}(a^T\tilde{x})^2 = \begin{pmatrix} -\frac{1}{\alpha}a^T\tilde{x} & \tilde{x}^T \end{pmatrix} \begin{pmatrix} \alpha & a^T \\ a & \tilde{A} \end{pmatrix} \begin{pmatrix} -\frac{1}{\alpha}a^T\tilde{x} \\ \tilde{x} \end{pmatrix} > 0.$$

Implikace „$\Leftarrow$“. Buď $x = \left(\begin{smallmatrix} \beta \\ \tilde{x} \end{smallmatrix}\right) \in \mathbb{R}^n$. Pak

$$x^TAx = \begin{pmatrix} \beta & \tilde{x}^T \end{pmatrix} \begin{pmatrix} \alpha & a^T \\ a & \tilde{A} \end{pmatrix} \begin{pmatrix} \beta \\ \tilde{x} \end{pmatrix} = \alpha\beta^2 + 2\beta a^T\tilde{x} + \tilde{x}^T\tilde{A}\tilde{x} =$$
$$= \tilde{x}^T\big(\tilde{A} - \tfrac{1}{\alpha}aa^T\big)\tilde{x} + \big(\sqrt{\alpha}\beta + \tfrac{1}{\sqrt{\alpha}}a^T\tilde{x}\big)^2 \geq 0.$$

Rovnost nastane pouze tehdy, když $\tilde{x} = o$ a druhý čtverec je nulový, tj. $\beta = 0$. □

Přestože rekurentní vzoreček je pro testování positivní definitnosti použitelný, větší roli hraje následující Choleského rozklad[1)].

Věta 11.10 (Choleského rozklad)**.** *Pro každou positivně definitní matici $A \in \mathbb{R}^{n\times n}$ existuje jediná dolní trojúhelníková matice $L \in \mathbb{R}^{n\times n}$ s kladnou diagonálou taková, že $A = LL^T$.*

Důkaz. Matematickou indukcí podle n. Pro $n = 1$ máme $A = (a_{11})$ a $L = (\sqrt{a_{11}})$.

Indukční krok $n \leftarrow n-1$. Mějme $A = \begin{pmatrix} \alpha & a^T \\ a & \tilde{A} \end{pmatrix}$. Podle věty 11.9 je $\alpha > 0$ a $\tilde{A} - \frac{1}{\alpha}aa^T$ je positivně definitní. Tedy dle indukčního předpokladu existuje dolní trojúhelníková matice $\tilde{L} \in \mathbb{R}^{(n-1)\times(n-1)}$ s kladnou diagonálou tak, že $\tilde{A} - \frac{1}{\alpha}aa^T = \tilde{L}\tilde{L}^T$. Potom $L = \begin{pmatrix} \sqrt{\alpha} & o^T \\ \frac{1}{\sqrt{\alpha}}a & \tilde{L} \end{pmatrix}$, neboť

$$LL^T = \begin{pmatrix} \sqrt{\alpha} & o^T \\ \frac{1}{\sqrt{\alpha}}a & \tilde{L} \end{pmatrix} \begin{pmatrix} \sqrt{\alpha} & \frac{1}{\sqrt{\alpha}}a^T \\ o & \tilde{L}^T \end{pmatrix} = \begin{pmatrix} \alpha & a^T \\ a & \frac{1}{\alpha}aa^T + \tilde{L}\tilde{L}^T \end{pmatrix} = A.$$

Pro důkaz jednoznačnosti mějme jiný rozklad $A = L'L'^T$, kde $L' = \begin{pmatrix} \beta & o^T \\ b & \tilde{L}' \end{pmatrix}$. Pak

$$\begin{pmatrix} \alpha & a^T \\ a & \tilde{A} \end{pmatrix} = A = L'L'^T = \begin{pmatrix} \beta^2 & \beta b^T \\ \beta b & bb^T + \tilde{L}'\tilde{L}'^T \end{pmatrix}.$$

Porovnáním matic dostaneme: $\beta = \sqrt{\alpha}$, $b = \frac{1}{\sqrt{\alpha}}a$ a $\tilde{A} = bb^T + \tilde{L}'\tilde{L}'^T$, neboli $\tilde{L}'\tilde{L}'^T = \tilde{A} - \frac{1}{\alpha}aa^T$. Jenže podle indukčního předpokladu je $\tilde{A} - \frac{1}{\alpha}aa^T = \tilde{L}\tilde{L}^T$ jednoznačné, tedy $\tilde{L}' = \tilde{L}$, a tudíž i $L' = L$. □

Choleského rozklad existuje i pro positivně semidefinitní matice, ale není už jednoznačný.

Věta byla spíše existenčního charakteru, nicméně sestrojit Choleského rozklad je vcelku jednoduché. Základní idea je postupně porovnávat shora prvky v prvním sloupci matice $A = LL^T$, pak ve druhém atd. Odvození a výsledný postup jsou popsány níže. Jestliže A je positivně definitní, algoritmus najde rozklad, a pokud A není, tak to algoritmus ohlásí.

Algoritmus 11.11 (Choleského rozklad)**.** Buď $A \in \mathbb{R}^{n\times n}$ symetrická.

1: $L := 0_n$,

2: **for** $k := 1$ **to** n **do** //v k-tém cyklu určíme hodnoty L_{*k}

3: **if** $a_{kk} - \sum_{j=1}^{k-1} \ell_{kj}^2 \le 0$ **then return** „A není positivně definitní“,

4: $\ell_{kk} := \sqrt{a_{kk} - \sum_{j=1}^{k-1} \ell_{kj}^2}$,

5: **for** $i := k+1$ **to** n **do**

6: $\ell_{ik} := \frac{1}{\ell_{kk}} \left(a_{ik} - \sum_{j=1}^{k-1} \ell_{ij}\ell_{kj} \right)$,

7: **end for**

8: **end for**

9: **return** $A = LL^T$.

[1)] André-Louis Cholesky, francouzský důstojník (pravděpodobně polského původu), metodu z r. 1910 vyvinul pro účely triangularizace a vytvoření přesnějších map (v podstatě pro řešení soustavy normálních rovnic v metodě nejmenších čtverců), ale publikována byla až po jeho smrti roku 1924.

Odvození Choleského rozkladu. Předpokládejme, že máme spočítaný první až $(k-1)$-ní sloupec matice L. Ze vztahu $A = LL^T$ odvodíme

$$a_{kk} = \sum_{j=1}^{n} L_{kj}(L^T)_{jk} = \sum_{j=1}^{n} \ell_{kj}^2 = \sum_{j=1}^{k} \ell_{kj}^2.$$

Jediná neznámá v tomto vztahu je hodnota ℓ_{kk}, a pokud ji z rovnice vyjádříme, dostaneme výraz z kroku 4.

Nyní předpokládejme, že z matice L máme navíc určeny hodnoty prvních $i-1$ prvků sloupce k. Ze vztahu $A = LL^T$ odvodíme pro $i > k$

$$a_{ik} = \sum_{j=1}^{n} L_{ij}(L^T)_{jk} = \sum_{j=1}^{n} \ell_{ij}\ell_{kj} = \sum_{j=1}^{k} \ell_{ij}\ell_{kj}.$$

V tomto výrazu je jediná neznámá hodnota l_{ik} a jejím vyjádřením dostaneme vzorec z kroku 6. □

Popsaný algoritmus zároveň může posloužit jako alternativní důkaz jednoznačnosti Choleského rozkladu. Prvky matice A jednoznačně určily prvky matice L, nikde jsme neměli na výběr z více možností.

Příklad 11.12. Choleského rozklad matice A:

$$\begin{array}{c|c} & L^T \\ \hline L & A \end{array} \equiv \begin{array}{c|c} & \begin{pmatrix} 2 & -1 & 2 \\ 0 & 3 & 1 \\ 0 & 0 & 1 \end{pmatrix} \\ \hline \begin{pmatrix} 2 & 0 & 0 \\ -1 & 3 & 0 \\ 2 & 1 & 1 \end{pmatrix} & \begin{pmatrix} 4 & -2 & 4 \\ -2 & 10 & 1 \\ 4 & 1 & 6 \end{pmatrix} \end{array}$$

□

Příklad 11.13. Použití Choleského rozkladu pro řešení soustavy $Ax = b$ s positivně definitní maticí A. Pokud máme rozklad $A = LL^T$, pak soustava má tvar $L(L^Tx) = b$. Nejprve vyřešíme soustavu $Ly = b$, potom $L^Tx = y$ a její řešení je to hledané. Postup:

1. Najdi Choleského rozklad $A = LL^T$,
2. Najdi řešení y^* soustavy $Ly = b$ pomocí dopředné substituce,
3. Najdi řešení x^* soustavy $L^Tx = y^*$ pomocí zpětné substituce.

Tento postup je řádově o 50% rychlejší než řešení Gaussovou eliminací.

Choleského rozklad můžeme použít i k invertování positivně definitních matic, protože $A^{-1} = (LL^T)^{-1} = (L^{-1})^T L^{-1}$ a inverze k dolní trojúhelníkové matici L se najde snadno. □

Rekurentní vzoreček má ještě další důsledky, které vyjadřují, jak testovat positivní definitnost pomocí Gaussovy–Jordanovy eliminace a pomocí determinantů.

Tvrzení 11.14 (Gaussova eliminace a positivní definitnost)**.** *Symetrická matice $A \in \mathbb{R}^{n\times n}$ je positivně definitní právě tehdy, když ji Gaussova eliminace převede do odstupňovaného tvaru s kladnou diagonálou za použití pouze elementární úpravy přičtení násobku řádku s pivotem k jinému řádku pod ním.*

Důkaz. Mějme $A = \begin{pmatrix} \alpha & a^T \\ a & \tilde{A} \end{pmatrix}$ positivně definitní. První krok Gaussovy eliminace převede matici na tvar $\begin{pmatrix} \alpha & a^T \\ o & \tilde{A} - \frac{1}{\alpha}aa^T \end{pmatrix}$. Podle věty 11.9 je $\alpha > 0$ a $\tilde{A} - \frac{1}{\alpha}aa^T$ je zase positivně definitní, takže můžeme pokračovat induktivně dál. □

Tvrzení 11.15 (Sylvestrovo[2)] kriterium positivní definitnosti)**.** *Symetrická matice $A \in \mathbb{R}^{n\times n}$ je positivně definitní právě tehdy, když determinanty všech hlavních vedoucích podmatic $A_1, \ldots, A_n$ jsou kladné, přičemž A_i je levá horní podmatice A velikosti i (tj. vznikne z A odstraněním posledních $n-i$ řádků a sloupců).*

[2)] James Joseph Sylvester, anglický matematik, spoluzakladatel teorie matic. Jako první použil pojem „matice“ v práci z r. 1850. Kriterium je z r. 1852.

Důkaz. Implikace „⇒“. Buď $A \in \mathbb{R}^{n\times n}$ positivně definitní. Pak pro každé $i = 1, \dots, n$ je A_i positivně definitní, neboť pokud $x^T A_i x \le 0$ pro jisté $x \ne o$, tak $(x^T\, o^T) A \binom{x}{o} = x^T A_i x \le 0$. Tedy A_i má kladná vlastní čísla a její determinant je také kladný (je roven součinu vlastních čísel).

Implikace „⇐“. Během Gaussovy eliminace matice A jsou všechny pivoty kladné, neboť pokud je i-tý pivot první nekladný, pak $\det(A_i) \le 0$. Podle tvrzení 11.14 je tedy A positivně definitní. □

Z nezápornosti determinantů všech hlavních vedoucích podmatic ještě positivní semidefinitnost nevyplývá (najděte takový příklad!). Analogie Sylvestrovy podmínky pro positivně semidefinitní matice je následující.

Tvrzení 11.16 (Sylvestrovo kriterium positivní semidefinitnosti)**.** *Symetrická matice $A \in \mathbb{R}^{n\times n}$ je positivně semidefinitní právě tehdy, když determinanty všech hlavních podmatic jsou nezáporné, přičemž hlavní podmatice je matice, která vznikne z A odstraněním určitého počtu (i nulového) řádků a sloupců s týmiž indexy.*

Důkaz. Je-li A positivně semidefinitní, pak zřejmě hlavní podmatice jsou také positivně semidefinitní, a tudíž mají nezáporný determinant (= součin vlastních čísel).

Důkaz opačné implikace provedeme matematickou indukcí. Pro $n = 1$ je tvrzení zřejmé.

Indukční krok $n \leftarrow n - 1$. Pro spor buď $\lambda < 0$ vlastní číslo A, a nechť x je odpovídající vlastní vektor znormovaný tak, že $\|x\|_2 = 1$. Jsou-li všechna ostatní vlastní čísla kladná, pak $\det(A) < 0$, a jsme hotovi. V opačném případě buď $\mu \le 0$ dalším vlastním číslem A a buď y, $\|y\|_2 = 1$, odpovídající vlastní vektor. Nyní nalezneme $\alpha \in \mathbb{R}$ takové, že vektor $z := x + \alpha y$ má aspoň jednu složku nulovou; nechť je to i-tá. Protože $x \perp y$, máme

$$\begin{aligned} z^T A z &= (x + \alpha y)^T A (x + \alpha y) = (x + \alpha y)^T (Ax + \alpha Ay) = \\ &= (x + \alpha y)^T (\lambda x + \alpha\mu y) = \lambda x^T x + \alpha^2 \mu y^T y = \lambda + \alpha^2 \mu < 0. \end{aligned}$$

Nechť A' vznikne z A odstraněním i-tého řádku a sloupce, a z' nechť vznikne ze z odstraněním i-té složky. Pak $z'^T A' z' = z^T A z < 0$, tudíž hlavní podmatice A' není positivně semidefinitní a aplikujeme indukční předpoklad. □

Zatímco Sylvestrovo kriterium positivní definitnosti vyžaduje počítání n determinantů, pro positivní semidefinitnost počet vzroste na $2^n - 1$, a proto není moc použitelnou metodou. Lepší způsob ukážeme později v sekci 12.2 (důsledek 12.16).

Přestože jsme uvedli několik metod na testování positivní definitnosti, některé jsou si dost podobné. Důkaz tvrzení 11.14 ukazuje, že rekurentní vzoreček a Gaussova eliminace fungují v podstatě stejně. A pokud počítáme determinanty Gaussovou eliminací, tak i Sylvestrovo pravidlo je varianta prvních dvou. Naproti tomu Choleského rozklad je metoda principiálně odlišná.

11.2 Aplikace

Nejprve ukážeme, že pomocí positivně definitních matic můžeme popsat všechny možné skalární součiny na prostoru $\mathbb{R}^n$.

Věta 11.17 (Skalární součin a positivní definitnost)**.** *Operace $\langle x, y\rangle$ je skalárním součinem v $\mathbb{R}^n$ právě tehdy, když má tvar $\langle x, y\rangle = x^T A y$ pro nějakou positivně definitní matici $A \in \mathbb{R}^{n\times n}$.*

Důkaz. Implikace „⇒“. Definujme matici $A \in \mathbb{R}^{n\times n}$ předpisem $a_{ij} = \langle e_i, e_j\rangle$, kde e_i, e_j jsou standardní jednotkové vektory. Matice A je zjevně symetrická, ale nemusí být jednotková, protože daný skalární součin nemusí být ten standardní. Nyní podle linearity skalárního součinu v první i druhé složce můžeme psát

$$\begin{aligned} \langle x, y\rangle &= \left\langle \sum_{i=1}^n x_i e_i, \sum_{j=1}^n y_j e_j \right\rangle = \sum_{i=1}^n x_i \left\langle e_i, \sum_{j=1}^n y_j e_j \right\rangle = \\ &= \sum_{i=1}^n \sum_{j=1}^n x_i y_j \langle e_i, e_j\rangle = \sum_{i=1}^n \sum_{j=1}^n x_i y_j a_{ij} = x^T A y. \end{aligned}$$

Matice A musí být positivně definitní, neboť z definice skalárního součinu $x^TAx = \langle x, x\rangle \geq 0$ a nulové jen pro $x = o$.

Implikace „$\Leftarrow$“. Nechť A je positivně definitní. Pak $\langle x, y\rangle = x^TAy$ tvoří skalární součin: $\langle x, x\rangle = x^TAx \geq 0$ a nulové jen pro $x = o$, je lineární v první složce a je symetrický neboť

$$\langle x, y\rangle = x^TAy = (x^TAy)^T = y^TA^Tx = y^TAx = \langle y, x\rangle. \qquad \square$$

Víme, že skalární součin indukuje normu (definice 8.4). Norma indukovaná výše zmíněným skalárním součinem je $\|x\| = \sqrt{x^TAx}$. V této normě je jednotková koule elipsoid (viz příklad 12.21). Pro $A = I_n$ dostáváme standardní skalární součin v $\mathbb{R}^n$ a eukleidovskou normu.

Poznámka 11.18. Přestože nestandardní skalární součin $\langle x, y\rangle = x^TAy$ může vypadat podivně, jeho vztah ke standardnímu je velmi blízký. Protože matice A je positivně definitní, lze rozložit jako $A = R^TR$, kde R je regulární. Buď B báze tvořená sloupci matice R^{-1}, tudíž $R = {}_B[id]_{\text{kan}}$ je matice přechodu od kanonické báze do B. Nyní $x^TAy = x^TR^TRy = (Rx)^T(Ry) = [x]_B^T[y]_B$. To ukazuje, že nestandardní skalární součin lze vyjádřit jako standardní skalární součin vzhledem k určité bázi.

Další aplikací je odmocnina z matice. Pro positivně semidefinitní matice můžeme zavést positivně semidefinitní odmocninu $\sqrt{A}$. Odmocnina je dokonce jednoznačná, důkaz viz Rohn [2003].

Tvrzení 11.19 (Odmocnina z matice)**.** *Pro každou positivně semidefinitní matici $A \in \mathbb{R}^{n\times n}$ existuje positivně semidefinitní matice $B \in \mathbb{R}^{n\times n}$ taková, že $B^2 = A$.*

Důkaz. Nechť A má spektrální rozklad $A = Q\Lambda Q^T$, kde $\Lambda = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$, $\lambda_1, \dots, \lambda_n \geq 0$. Definujme diagonální matici $\Lambda' = \operatorname{diag}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n})$ a matici $B = Q\Lambda'Q^T$. Pak $B^2 = Q\Lambda'Q^TQ\Lambda'Q^T = Q\Lambda'^2Q^T = Q\Lambda Q^T = A$. $\square$

Zde je zajímavé porovnat odmocninu z matice s maticovými funkcemi z příkladu 10.46. Odmocninu lze vyjádřit nekonečným rozvojem pouze v malém okolí daného kladného čísla, nicméně tam, kde existuje, se budou obě definice shodovat.

Poznámka 11.20. Positivní (semi-)definitnost je důležitá i v optimalizaci při určování minima funkce. Matice, která zde vystupuje, je tzv. hessián, matice druhých parciálních derivací. Za předpokladu, že jsme v bodě x^* s nulovým gradientem, pak positivní definitnost dává postačující podmínku pro to aby x^* bylo lokální minimum, a naopak positivní semidefinitnost dává nutnou podmínku.

Hessián se podobně používá i při určování konvexity funkce. Positivní definitnost na nějaké otevřené konvexní množině implikuje konvexitu funkce f. Více např. viz [Rohn, 1997].

Positivně definitní matice hrají v optimalizaci ještě jednu důležitou roli. Semidefinitní program je taková optimalizační úloha, při níž hledáme minimum lineární funkce za podmínky na positivní semidefinitnost matice, jejíž prvky jsou lineární funkcí proměnných [Gärtner and Matoušek, 2012]. Formálně, jedná se o úlohu

$$\min\ c^Tx \text{ za podmínky } A_0 + \sum_{i=1}^{m} A_ix_i \text{ je positivně definitní,}$$

kde $c \in \mathbb{R}^m$, $A_0, A_1, \dots, A_m \in \mathbb{R}^{n\times n}$ jsou dány a $x = (x_1, \dots, x_m)^T$ je vektor proměnných. Semidefinitními programy dokážeme nejen modelovat větší třídu úloh než lineárními programy (viz příklad 7.18), ale stále jsou řešitelné efektivně za rozumný čas. Umožnily mj. velký pokrok na poli kombinatorické optimalizace, protože řada výpočetně složitých problémů se dá rychle a těsně aproximovat právě pomocí vhodných semidefinitních programů.

Výskyt positivně (semi-)definitních matic je ještě širší. Například ve statistice se setkáváme s tzv. kovarianční a korelační maticí. Obě dávají jistou informaci o závislosti mezi n náhodnými veličinami a, ne náhodou, jsou vždy positivně semidefinitní.

Na závěr této sekce ukážeme ještě aplikaci z chemie.

Příklad 11.21 (Určování struktury bílkovin)**.** Jedna ze základních úloh modelování bílkovin je určení trojrozměrné struktury bílkovin. Typický postup je určení matice vzdáleností jednotlivých atomů pomocí jaderné magnetické rezonance a pak odvození struktury.

Struktura bílkoviny. [zdroj: Wikipedia]

Buď $X \in \mathbb{R}^{n\times 3}$ matice pozic jednotlivých atomů, to jest, řádek X_{i*} udává souřadnice i-tého atomu v prostoru. Maticí $D \in \mathbb{R}^{n\times n}$ si označíme vzdálenosti jednotlivých atomů, tedy $d_{ij} =$ vzdálenost i-tého a j-tého atomu. Jestliže známe X, tak D spočítáme následujícím způsobem. Posuneme souřadný systém, aby n-tý atom byl v počátku, tj. $X_{n*} = (0,0,0)$ a z matice X odstraníme poslední řádek, který je nyní redundantní. Označme pomocnou matici $D^* := XX^T$ a souřadnice libovolných dvou atomů $u := X_{i*}$, $v := X_{j*}$. Vztah matic D a D^* je tento:

$$d_{ij}^2 = \|u-v\|^2 = \langle u-v, u-v\rangle = \langle u,u\rangle + \langle v,v\rangle - 2\langle u,v\rangle = d_{ii}^* + d_{jj}^* - 2d_{ij}^*. \tag{11.1}$$

Náš problém stojí typicky naopak: ze znalosti naměřených vzdáleností D potřebujeme určit X, čímž získáme představu, kde v prostoru se jaký atom nachází a tím pádem jaká je struktura bílkoviny. Ze vztahu (11.1) odvodíme

$$d_{ij}^* = \frac{1}{2}(d_{ii}^* + d_{jj}^* - d_{ij}^2) = \frac{1}{2}(d_{in}^2 + d_{jn}^2 - d_{ij}^2).$$

Tímto předpisem z matice D spočítáme matici D^*. Protože D^* je symetrická a positivně definitní, ze spektrálního rozkladu $D^* = Q\Lambda Q^T$, kde $\Lambda = \operatorname{diag}(\lambda_1, \lambda_2, \lambda_3)$, sestrojíme hledanou matici $X = Q \cdot \operatorname{diag}\left(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \sqrt{\lambda_3}\right)$. Pak totiž máme $D = XX^T$.

Poznamenejme, že X se dá někdy spočítat i za částečné informace vzdáleností (matice D není známa celá), ale tato úloha může být výpočetně složitá (tzv. NP-úplný problém).

Jiný problém z tohoto oboru je tzv. *Prokrustův problém*,[3] v němž porovnáváme dvě struktury bílkovin jak moc jsou podobné. Označme matice dvou struktur X, Y. Lineární zobrazení s ortogonální maticí Q zachovává úhly a vzdálenosti (věta 8.54), tudíž YQ odpovídá té samé struktuře, jako Y, pouze nějakým způsobem otočené či zrcadlené. Chceme-li určit podobnost obou struktur, hledáme takovou ortogonální matici $Q \in \mathbb{R}^{3\times 3}$, aby matice X a YQ byly „co nejbližší“. Matematická formulace vede na optimalizační úlohu minimalizovat hodnotu maticové normy $\|X - YQ\|$ na množině ortogonálních matic Q. Tato úloha trochu připomíná metodu nejmenších čtverců. □

Problémy

11.1. Porovnejte počet aritmetických operací (nejhorší případ) následujících metod na testování positivní definitnosti matice $A \in \mathbb{R}^{n\times n}$: rekurentní vzoreček, Choleského rozklad, Gaussova eliminace a Sylvestrovo kriterium. Která metoda vychází nejlépe?

11.2. Ukažte, že každá symetrická matice se dá zapsat jako rozdíl dvou positivně definitních matic.

11.3. Buď $A \in \mathbb{R}^{n\times n}$ positivně definitní. Ukažte, že $A_{ii}(A^{-1})_{ii} \geq 1$ pro všechna $i = 1, \ldots, n$.

11.4. Buďte $A, B \in \mathbb{R}^{n\times n}$ symetrické a nechť všechna vlastní čísla A jsou větší než vlastní čísla B. Dokažte, že $A - B$ je positivně definitní.

[3] Prokrustes nebyl matematik, ale řecká mytologická postava, lupič z Attiky, který pocestným usekával končetiny nebo jim je natahoval, aby se přesně vešli na jeho lože. Jeho život ukončil až Theseus.

Shrnutí ke kapitole 11. Positivně (semi-)definitní matice

Positivně definitní matice je speciální typ matice, která se nicméně vyskytuje v rozličných situacích:

- každý skalární součin v prostoru $\mathbb{R}^n$ je tvaru $\langle x, y\rangle = x^T Ay$ pro nějakou positivně definitní matici A,
- funkce $f\colon \mathbb{R}^n \to \mathbb{R}$ je konvexní pokud její hessián je positivně definitní,
- atp.

Positivně definitní matici definujeme jako takovou symetrickou matici A, pro kterou polynom $x^T Ax$ je nezáporný a nulový pouze v počátku. Alternativně ji můžeme charakterizovat jako matici, která má kladná vlastní čísla. Ještě jinak lze positivně definitní matice nahlédnout tak, že mají rozklad $A = U^T U$, kde U je regulární. Dokonce lze po matici U požadovat, aby byla horní trojúhelníková s kladnou diagonálou (pak je i jednoznačná). To dává vzniku i efektivnímu způsobu na testování positivní definitnosti, tzv. Choleskému rozkladu. Navíc rozklad $A = U^T U$ je užitečný při řešení soustav lineárních rovnic či jiných výpočtech s maticí A.

Kapitola 12

Kvadratické formy

Přestože se lineární algebra zabývá převážně lineárními záležitostmi, i kvadratické modely zde mají své opodstatnění. V zásadě jsme se s nimi setkali již u eukleidovské normy (str. 108) a positivní definitnosti (definice 11.1). V této kapitole probereme kvadratické formy podrobněji. Ukážeme souvislosti s positivní (semi-)definitností, znaménky vlastních čísel a popisem určitých geometrických objektů.

Příklad 12.1 (Motivace ke kvadratickým formám)**.** Pro začátek si můžeme kvadratickou formu představit jako polynom n proměnných, kde součet stupňů každého členu je přesně dva. Čili

$$f(x) = 5x^2$$

je kvadratická forma s jednou proměnnou a

$$f(x_1, x_2) = 5x_1^2 - 3x_1x_2 + 12x_2^2$$

je kvadratická forma se dvěma proměnnými. Obecný předpis pro takovýto polynom s n proměnnými $x = (x_1, \dots, x_n)^T$ je

$$f(x) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j = x^T A x$$

kde A je matice $n \times n$ příslušných koeficientů. Kompaktní maticový zápis $x^T A x$ budeme často používat. □

12.1 Bilineární a kvadratické formy

Definice 12.2 (Bilineární a kvadratická forma)**.** Buď V vektorový prostor nad $\mathbb{T}$. *Bilineární forma* je zobrazení $b\colon V^2 \to \mathbb{T}$, které je lineární v první i druhé složce, tj.

$$\begin{aligned} b(\alpha u + \beta v, w) &= \alpha b(u, w) + \beta b(v, w), \quad \forall \alpha, \beta \in \mathbb{T},\ \forall u, v, w \in V, \\ b(w, \alpha u + \beta v) &= \alpha b(w, u) + \beta b(w, v), \quad \forall \alpha, \beta \in \mathbb{T},\ \forall u, v, w \in V. \end{aligned}$$

Bilineární forma se nazývá *symetrická*, pokud $b(u, v) = b(v, u)$ pro všechna $u, v \in V$. Zobrazení $f\colon V \to \mathbb{T}$ je *kvadratická forma*, pokud se dá vyjádřit $f(u) = b(u, u)$ pro nějakou symetrickou bilineární formu b.

Je lehké nahlédnout, že vždy platí $b(o, v) = b(v, o) = 0$, $f(o) = 0$.

Příklad 12.3.

- Každý reálný skalární součin je bilineární formou, tedy pro libovolnou matici $A \in \mathbb{R}^{n \times n}$ je zobrazení $b(x, y) = x^T A y$ bilineární formou. Pro symetrickou matici A je pak kvadratickou formou zobrazení $f(x) = x^T A x$.

  Speciálně, v prostoru $\mathbb{R}^1$ je bilineární formou jakékoli zobrazení $b\colon \mathbb{R}^2 \to \mathbb{R}$ s předpisem $b(x, y) = axy$, kde $a \in \mathbb{R}$ je konstanta. Příslušná kvadratická forma je pak kvadratická funkce jedné proměnné $f(x) = ax^2$. Kvadratické formy na $\mathbb{R}^n$ lze tedy chápat jako zobecnění kvadratické funkce z jedné na n proměnných.

- Komplexní skalární součin není bilineární formou, protože není lineární v druhé složce (viz strana 107). Formy takovéhoto typu se nazývají seskvilineární (více viz Bečvář [2005]).
- Buď $V = \mathbb{R}^2$. Pak
$$b(x, y) = x_1y_1 + 2x_1y_2 + 4x_2y_1 + 10x_2y_2$$
je příkladem bilineární formy,
$$b'(x, y) = x_1y_1 + 3x_1y_2 + 3x_2y_1 + 10x_2y_2$$
je příkladem symetrické bilineární formy a
$$f(x) = b'(x, x) = x_1^2 + 6x_1x_2 + 10x_2^2$$
odpovídající kvadratické formy. □

Příklad 12.4. Ilustrace bilineární formy $b(x, y) = xy$ na intervalu $[-5, 5]^2$ a na intervalu $[0, 5]^2$:

□

V této kapitole se budeme převážně zabývat kvadratickými formami, i když bilineární formy jsou také zajímavé, např. právě vztahem se skalárním součinem.

V analogii s teorií lineárních zobrazení, i bilineární formy jsou jednoznačně určeny obrazy bází a dají se vyjádřit maticově. Čtenáři proto doporučujeme porovnat pojmy a výsledky této sekce se sekcí 6.2 a uvědomit si jistou paralelu.

Poznámka 12.5 (Motivace k maticím forem)**.** Buď $b\colon V^2 \to \mathbb{T}$ bilineární forma a $B = \{w_1, \dots, w_n\}$ báze prostoru V. Mějme libovolné dva vektory $u, v \in V$ a nechť mají v bázi B vyjádření $u = \sum_{i=1}^n x_iw_i$, $v = \sum_{i=1}^n y_iw_i$. Z definice bilineární formy pak obraz vektorů je
$$b(u, v) = b(\textstyle\sum_{i=1}^n x_iw_i, \sum_{j=1}^n y_jw_j) = \sum_{i=1}^n \sum_{j=1}^n x_iy_jb(w_i, w_j).$$
Vidíme, že celá bilineární forma je vlastně určena tím, kam se zobrazí všechny dvojice bázických vektorů. To nás navíc motivuje umístit tyto hodnoty $b(w_i, w_j)$ do matice a pracovat s maticovou reprezentací.

Definice 12.6 (Matice bilineární a kvadratické formy)**.** Buď $b\colon V^2 \to \mathbb{T}$ bilineární forma a $B = \{w_1, \dots, w_n\}$ báze prostoru V. Pak definujme matici $A \in \mathbb{T}^{n\times n}$ bilineární formy vzhledem k bázi B předpisem $a_{ij} = b(w_i, w_j)$. Matice kvadratické formy $f\colon V \to \mathbb{T}$ je definována jako matice libovolné symetrické bilineární formy indukující f.

Pomocí matic forem můžeme elegantně vyjádřit jejich funkční hodnoty explicitním předpisem.

Věta 12.7 (Maticové vyjádření forem)**.** *Buď B báze vektorového prostoru V a buď b bilineární forma na V. Pak A je matice formy b vzhledem k bázi B právě tehdy, když pro každé $u, v \in V$ platí*
$$b(u, v) = [u]_B^T A[v]_B. \tag{12.1}$$
Dále, je-li b symetrická forma, pak odpovídající kvadratická forma f pro každé $u \in V$ splňuje
$$f(u) = [u]_B^T A[u]_B.$$

Důkaz. Označme $x := [u]_B$, $y := [v]_B$, a nechť B se skládá z vektorů $w_1, \dots, w_n$. Je-li A matice formy b, tak

$$b(u,v) = b\Big(\sum_{i=1}^n x_i w_i, \sum_{j=1}^n y_j w_j\Big) = \sum_{i=1}^n \sum_{j=1}^n x_i y_j b(w_i, w_j) = \sum_{i=1}^n \sum_{j=1}^n x_i y_j a_{ij} = x^T A y.$$

Naopak, pokud platí (12.1) pro každé $u, v \in V$, tak dosazením $u := w_i$, $v := w_j$ dostaneme $b(w_i, w_j) = [w_i]_B^T A [w_j]_B = e_i^T A e_j = a_{ij}$ pro všechna $i, j = 1, \dots, n$.

Konečně, $f(u) = b(u,u) = x^T A x$. □

Každá bilineární forma je jednoznačně určená obrazy všech dvojic bázických vektorů.

Důsledek 12.8. *Buď $B = \{w_1, \dots, w_n\}$ báze vektorového prostoru V nad $\mathbb{T}$ a buď $A \in \mathbb{T}^{n\times n}$. Pak existuje jediná bilineární forma $b\colon V^2 \to \mathbb{T}$ taková, že $b(w_i, w_j) = a_{ij}$ pro všechna $i, j = 1, \dots, n$.*

Důkaz. „Existence.“ Stačí ověřit, že zobrazení $b\colon V^2 \to \mathbb{T}$ dané předpisem $b(u,v) = [u]_B^T A [v]_B$ splňuje podmínky bilineární formy. To se nahlédne snadno, neboť zobrazení $u \mapsto [u]_B$ je lineární (srov. tvrzení 6.34).

„Jednoznačnost.“ Z (12.1) plyne, že pro každé $u, v \in V$ je $b(u,v) = [u]_B^T A [v]_B$, tedy obrazy jsou jednoznačně dány. □

Buď B pevná báze prostoru V dimenze n. Každé bilineární formě tedy odpovídá jednoznačná matice $A \in \mathbb{T}^{n\times n}$, a naopak každé matici $A \in \mathbb{T}^{n\times n}$ odpovídá jednoznačně bilineární forma. Existuje tudíž vzájemně jednoznačný vztah mezi množinou bilineárních forem a prostorem matic $\mathbb{T}^{n\times n}$. Jedná se navíc o isomorfismus, protože bilineární formy tvoří vektorový prostor s přirozeně definovaným součtem a násobky (srov. prostor $\mathcal{F}$ ze str. 64).

Ve vektorovém prostoru $\mathbb{R}^n$ mají bilineární formy speciální tvar.

Důsledek 12.9. *Každá bilineární forma na $\mathbb{R}^n$ se dá vyjádřit ve tvaru*

$$b(x,y) = x^T A y$$

pro určitou matici $A \in \mathbb{R}^{n\times n}$, a každá kvadratická forma na $\mathbb{R}^n$ se dá vyjádřit ve tvaru

$$f(x) = x^T A x$$

pro určitou symetrickou matici $A \in \mathbb{R}^{n\times n}$.

Důkaz. Stačí vzít A jako matici formy vzhledem ke kanonické bázi. Pak $b(x,y) = [x]_{\text{kan}}^T A [y]_{\text{kan}} = x^T A y$.

Pro kvadratickou formu pak platí $f(x) = b(x,x) = x^T A x$. Pokud A není symetrická, nahradíme ji maticí $\frac{1}{2}(A + A^T)$ ve smyslu poznámky 11.2. □

Příklad 12.10. Uvažme bilineární formu na $\mathbb{R}^2$

$$b(x,y) = x_1 y_1 + 2x_1 y_2 + 4x_2 y_1 + 10 x_2 y_2.$$

Matice b vzhledem ke kanonické bázi je $A = \begin{pmatrix} 1 & 2 \\ 4 & 10 \end{pmatrix}$, což snadno nahlédneme i z vyjádření

$$b(x,y) = x^T A y = \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 4 & 10 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}.$$

Tato bilineární forma není symetrická, narozdíl od bilineární formy

$$b'(x,y) = x_1 y_1 + 3x_1 y_2 + 3x_2 y_1 + 10 x_2 y_2.$$

Matice b' vzhledem ke kanonické bázi je $A' = \begin{pmatrix} 1 & 3 \\ 3 & 10 \end{pmatrix}$, tedy $b'(x,y) = x^T A' y$. Odpovídající kvadratická forma splňuje

$$f'(x) = b'(x,x) = x^T A' x = \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} 1 & 3 \\ 3 & 10 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$ □

Poznámka 12.11. Mohli bychom si klást otázku, proč kvadratickou formu definujeme pouze pomocí symetrických bilineárních forem. Vždyť nic nám nebrání zavést $f(u) = b(u,u)$ i pro nesymetrickou bilineární formu b. Důvod je podobný jaký jsme uvedli u positivně definitních matic, viz poznámka 11.2. Můžeme zavést bilineární formu $b_s(u,v) := \frac{1}{2}\big(b(u,v) + b(v,u)\big)$, která již bude symetrická. Navíc, jak se snadno nahlédne, obě formy b, b_s indukují stejnou kvadratickou formu f. Zde ovšem těleso $\mathbb{T}$ nesmí mít charakteristiku 2, jinak by zlomek nedával smysl. Restrikcí na symetrický případ pak také v důsledku máme jednoznačnost matice kvadratických forem, opět za předpokladu, že těleso $\mathbb{T}$ nemá charakteristiku 2.

Matice forem závisí na volbě báze. Jak se změní matice, když přejdeme k jiné bázi?

Věta 12.12 (Matice kvadratické formy při změně báze)**.** *Buď $A \in \mathbb{T}^{n\times n}$ matice kvadratické formy f vzhledem k bázi B prostoru V. Buď B' jiná báze a $S = {}_B[id]_{B'}$ matice přechodu od B' k B. Pak matice f vzhledem k bázi B' je S^TAS, a odpovídá stejné symetrické bilineární formě.*

Důkaz. Buď $u, v \in V$ a b symetrická bilineární forma indukující f. Pak

$$b(u,v) = [u]_B^T A[v]_B = ({}_B[id]_{B'} \cdot [u]_{B'})^T A({}_B[id]_{B'} \cdot [v]_{B'}) = [u]_{B'}^T S^T AS[v]_{B'}.$$

Podle věty 12.7 je S^TAS je matice formy b, a tím i f, vzhledem k bázi B'. □

Různou volbou báze prostoru V dosahujeme různé maticové reprezentace. Naším cílem bude najít takovou bázi, vůči níž je matice co nejjednodušší, tedy diagonální.

Zde je jistá paralela z diagonalizací pro vlastní čísla, kde jsme transformovali matici pomocí podobnosti. Nyní transformujeme matici úpravou S^TAS, kde S je regulární. Místo podobnosti nyní máme tzv. *kongruenci.* Jak uvidíme, u kvadratických forem je situace jednodušší – každou matici lze diagonalizovat.

12.2 Sylvestrův zákon setrvačnosti

V této sekci nadále uvažujeme reálný prostor $\mathbb{R}^n$ nad $\mathbb{R}$. Z definice je patrné, že matice kvadratické formy je symetrická. Tato vlastnost budeme velmi potřebovat.

Věta 12.13 (Sylvestrův zákon setrvačnosti[1])**.** *Buď $f(x) = x^TAx$ kvadratická forma. Pak existuje báze, vůči níž má f diagonální matici s prvky $1, -1, 0$. Navíc, tato matice je až na pořadí prvků jednoznačná.*

Důkaz. „Existence“. Protože A je symetrická, tak má spektrální rozklad $A = Q\Lambda Q^T$, kde $\Lambda = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$. Tedy $\Lambda = Q^TAQ$ je diagonalizace formy. Abychom docílili na diagonále ± 1, tak provedeme ještě úpravu $\Lambda' Q^TAQ\Lambda'$, kde Λ' je diagonální matice s prvky $\Lambda'_{ii} = |\lambda_i|^{-\frac{1}{2}}$, pokud $\lambda_i \neq 0$ a $\Lambda'_{ii} = 1$ jinak. Nyní můžeme $Q\Lambda'$ považovat za matici ${}_{\text{kan}}[id]_B$ přechodu od hledané báze B do kanonické báze. Tudíž bázi B vyčteme ve sloupcích matice $Q\Lambda'$.

„Jednoznačnost“. Sporem předpokládejme, že máme dvě různé diagonalizace D, D':

$$D = \begin{array}{c} \begin{matrix} 1 & \dots & p & p+1 & \dots & q & q+1 & \dots & n \end{matrix} \\ \begin{pmatrix} 1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & 1 & & & & & & \\ & & & -1 & & & & & \\ & & & & \ddots & & & & \\ & & & & & -1 & & & \\ & & & & & & 0 & & \\ & & & & & & & \ddots & \\ & & & & & & & & 0 \end{pmatrix} \end{array}, \qquad D' = \begin{array}{c} \begin{matrix} 1 & \dots & s & s+1 & \dots & t & t+1 & \dots & n \end{matrix} \\ \begin{pmatrix} 1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & 1 & & & & & & \\ & & & -1 & & & & & \\ & & & & \ddots & & & & \\ & & & & & -1 & & & \\ & & & & & & 0 & & \\ & & & & & & & \ddots & \\ & & & & & & & & 0 \end{pmatrix} \end{array}.$$

[1] Anglicky *Sylvester's law of inertia*, z r. 1852.

První nechť odpovídá bázi $B = \{w_1, \ldots, w_n\}$ a druhá bázi $B' = \{w'_1, \ldots, w'_n\}$. Buď $u \in \mathbb{R}^n$ libovolné a nechť má souřadnice $y = [u]_B$, $z = [u]_{B'}$. Pak podle věty 12.7

$$f(u) = [u]_B^T D [u]_B = y^T D y = y_1^2 + \ldots + y_p^2 - y_{p+1}^2 - \ldots - y_q^2 + 0y_{q+1}^2 + \ldots + 0y_n^2,$$
$$f(u) = [u]_{B'}^T D [u]_{B'} = z^T D' z = z_1^2 + \ldots + z_s^2 - z_{s+1}^2 - \ldots - z_t^2 + 0z_{t+1}^2 + \ldots + 0z_n^2.$$

Nejprve si povšimneme, že $q = t$. Protože $D = S^T D' S$ pro nějakou regulární S, konkrétně pro $S = {}_{B'}[id]_B$, tak matice D, D' mají stejnou hodnost. Tudíž musí $q = t$. Nyní zbývá ukázat, že nutně $p = s$. Bez újmy na obecnosti předpokládejme $p > s$. Definujme prostory $P = \operatorname{span}\{w_1, \ldots, w_p\}$ a $R = \operatorname{span}\{w'_{s+1}, \ldots, w'_n\}$. Pak

$$\dim P \cap R = \dim P + \dim R - \dim(P + R) \geq p + (n - s) - n = p - s \geq 1.$$

Tedy existuje nenulový vektor $u \in P \cap R$ a pro něj máme $u = \sum_{i=1}^p y_i w_i = \sum_{j=s+1}^n z_j w'_j$, z čehož dostáváme

$$f(u) = \begin{cases} y_1^2 + \ldots + y_p^2 > 0, \\ -z_{s+1}^2 - \ldots - z_t^2 \leq 0. \end{cases}$$

To je spor. □

Poznamenejme, že báze, vůči níž matice kvadratické formy je diagonální, se nazývá *polární báze*. Tedy báze z věty 12.13 je příkladem polární báze, ale typicky existují i další. Dá se také ukázat, že polární báze existuje nejen pro reálné prostory, ale i pro prostory nad libovolným tělesem charakteristiky různé od 2.

Geometrický význam Sylvestrova zákona setrvačnosti spočívá v tom najít vhodný souřadný systém (tj. bázi), ve kterém má kvadratická forma jednoduchý diagonální tvar. Algebraický pohled na věc je ten, že danou symetrickou matici A transformujeme na diagonální tvar pomocí úprav $S^T A S$, kde S je regulární.

Příklad 12.14 (Kvadratické formy v $\mathbb{R}^2$)**.** Podle Sylvestrova zákona mají kvadratické formy v $\mathbb{R}^2$ v podstatě jeden z následujících tvarů v souřadném systému vhodné báze.

$x_1^2 + x_2^2$

$-x_1^2 - x_2^2$

x_1^2

$-x_1^2$

$x_1^2 - x_2^2$

0

□

Význam Sylvestrova zákona nespočívá ani tak v existenci diagonalizace, jako spíš v její jednoznačnosti (odtud název „setrvačnost“). Tato jednoznačnost opravňuje k zavedení pojmu *signatura* jako trojice (p, q, z), kde p je počet jedniček, q počet minus jedniček a z počet nul ve výsledné diagonální matici. Navíc má řadu důsledků týkající se mj. positivní (semi-)definitnosti a vlastních čísel.

Důsledek 12.15. *Buď $A \in \mathbb{R}^{n\times n}$ symetrická a S^TAS převedení na diagonální tvar. Pak počet jedniček resp. minus jedniček resp. nul na diagonále odpovídá počtu kladných resp. záporných resp. nulových vlastních čísel.*

Důkaz. Stačí uvažovat kvadratickou formu $f(x) = x^TAx$, která má matici A. Z důkazu věty je patrné, že jednu diagonalizaci získáme ze spektrálního rozkladu. Ze setrvačnosti pak musí počty vždy souhlasit. □

Diagonalizací matice A tedy nenajdeme vlastní čísla, ale určíme kolik jich je kladných a kolik záporných. Pokud tento postup aplikujeme na matici $A - \alpha I_n$ tak zjistíme, kolik vlastních čísel matice A je větších / menších / rovno číslu α. Takto lze postupným půlením intervalu omezovat potenciální rozsah vlastních čísel a limitně konvergovat k jejich hodnotám. Tento postup se kdysi používal pro výpočet vlastních čísel, ale nyní je již překonaný (mj. to není moc stabilní metoda).

Důsledek 12.16. *Buď $A \in \mathbb{R}^{n\times n}$ symetrická a S^TAS převedení na diagonální tvar. Pak*

(1) A je positivně definitní právě tehdy, když S^TAS má kladnou diagonálu,

(2) A je positivně semidefinitní právě tehdy, když S^TAS má nezápornou diagonálu.

Důkaz. Z důsledku 12.15 a vztahu mezi positivní (semi-)definitností a vlastními čísly (věta 11.7). □

Sylvestrův zákon tedy dává návod, jak jednou metodou rozhodnout o positivní definitnosti resp. positivní semidefinitnosti resp. negativní (semi-)definitnosti v jednom.

Zbývá otázka, jak matici kvadratické formy převést na diagonální tvar. Důkaz věty o Sylvestrově zákonu sice dává návod (přes spektrální rozklad), ale můžeme jednoduše adaptovat elementární maticové úpravy. Co se stane, když symetrickou matici A převedeme na E^TAE, kde E^T je matice elementární řádkové úpravy? Provede se řádková úprava a na sloupce i analogická sloupcová úprava. Budeme na matici tedy aplikovat řádkové úpravy a odpovídající sloupcové úpravy. Tím budeme nulovat prvky pod i nad diagonálou, až matici převedeme na diagonální tvar.

Algoritmus 12.17 (Diagonalizace matice kvadratické formy)**.** Buď $A \in \mathbb{R}^{n\times n}$ symetrická. Aplikováním elementárních řádkových úprav a odpovídajících sloupcových úprav převeď matici na diagonální tvar.

Příklad 12.18. Diagonalizujme matici

$$A = \begin{pmatrix} 1 & 2 & -1 \\ 2 & 5 & -3 \\ -1 & -3 & 2 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & -1 \\ 0 & 1 & -1 \\ -1 & -3 & 2 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ -1 & -1 & 2 \end{pmatrix} \sim$$
$$\sim \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Matice A má dvě kladná vlastní čísla a jedno nulové, je tedy positivně semidefinitní. □

Příklad 12.19. Ukažte, že součin positivně definitních matic $A, B \in \mathbb{R}^{n\times n}$ nemusí být positivně definitní matice, a to ve smyslu původní definice, nejen že AB není nutně symetrická.

Zajímavé ale je, že AB, přestože ne nutně symetrická, má stále kladná vlastní čísla. Snadno je to vidět z vyjádření $AB = \sqrt{A}\sqrt{A}B$. Vynásobením $\sqrt{A}^{-1}$ zleva a $\sqrt{A}$ zprava dostaneme podobnou matici $\sqrt{A}B\sqrt{A}$, která ze setrvačnosti má stejnou signaturu jako B, což znamená kladná vlastní čísla. □

Poznámka 12.20. Zde je užitečné ukázat vztah s rekurentním vzorečkem na testování positivní definitnosti (věta 11.9). Pokud uvažujeme matici $A = \begin{pmatrix} \alpha & a^T \\ a & \tilde{A} \end{pmatrix}$ jako matici kvadratické formy a elementárními

úpravami vynulujeme prvky pod a napravo od pivota, výsledná blokově diagonální matice se dá maticově vyjádřit jako

$$\begin{pmatrix} \alpha & o^T \\ o & \tilde{A} - \frac{1}{\alpha} a a^T \end{pmatrix}.$$

Tato matice je positivně definitní právě tehdy, když všechny bloky jsou positivně definitní matice, tj. $\alpha > 0$ a $\tilde{A} - \frac{1}{\alpha} a a^T$ je positivně definitní. Tím jsme dostali jiné odvození rekurentního vzorečku.

Kuželosečky a kvadriky

Pomocí kvadratických forem lze popisovat geometrické útvary zvané *kvadriky*. To jsou (stručně řečeno, podrobněji viz např. [Bican, 2009]) množiny popsané rovnicí $x^T A x + b^T x + c = 0$, kde $A \in \mathbb{R}^{n \times n}$ je symetrická, $b \in \mathbb{R}^n$, $c \in \mathbb{R}$. Jak vidíme, tato rovnice již není lineární právě díky kvadratickému členu $x^T A x$. Pomocí různých charakteristik, jako jsou například vlastní čísla resp. signatura matice A, můžeme pak snadno kvadriky klasifikovat na jednotlivé tvary. Těmito tvary jsou elipsoidy, paraboloidy, hyperboloidy aj., viz příklady dole.

Speciálním případem kvadrik v prostoru $\mathbb{R}^2$ jsou pak *kuželosečky*. Mezi ně patří elipsy, paraboly či hyperboly.

Příklad 12.21 (Elipsoidy)**.** Rovnice $\frac{1}{a^2} x_1^2 + \frac{1}{b^2} x_2^2 = 1$ popisuje v rovině $\mathbb{R}^2$ elipsu se středem v počátku, poloosy jsou ve směru souřadných os x_1, x_2 a mají délky a resp. b. Podobně pro vyšší dimenze.

Nyní uvažme rovnici $x^T A x = 1$, kde $A \in \mathbb{R}^{n \times n}$ je positivně definitní a $x = (x_1, \ldots, x_n)^T$ je vektor proměnných. Nahlédneme, že rovnice popisuje elipsoid v prostoru $\mathbb{R}^n$. Buď $A = Q \Lambda Q^T$ spektrální rozklad. Při substituci $y := Q^T x$ dostaneme

$$1 = x^T A x = x^T Q \Lambda Q^T x = y^T \Lambda y = \sum_{i=1}^{n} \lambda_i y_i^2 = \sum_{i=1}^{n} \frac{1}{\left(\lambda_i^{-1/2}\right)^2} y_i^2.$$

Dostáváme tedy popis elipsoidu se středem v počátku, poloosy jsou ve směru souřadnic a mají délky $\frac{1}{\sqrt{\lambda_1}}, \ldots, \frac{1}{\sqrt{\lambda_n}}$. Nicméně, tento popis je v prostoru po transformaci $y = Q^T x$. Vrátíme zpět transformací $x = Qy$. Protože Q je ortogonální matice, dostaneme stejný elipsoid se středem v počátku, jen nějak pootočený. Protože kanonická báze $e_1, \ldots, e_n$ se zobrazí na sloupce matice Q (což jsou vlastní vektory matice A), tak poloosy původního elipsoidu budou ve směrech vlastních vektorů A.

Je-li A symetrická, ale ne positivně definitní, analýza bude stejná. Jenom nedostaneme elipsoid, ale jiný geometrický útvar (hyperboloid aj.)

Příklad 12.22 (Některé kvadriky v $\mathbb{R}^3$)**.** Následující obrázky znázorňují některé další třidimenzionální kvadriky.

$$\frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} + \frac{x_3^2}{c^2} = 1$$

elipsoid

$$\frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} - \frac{x_3^2}{c^2} = 0$$

kuželová plocha

$$\frac{x_1^2}{a^2} - \frac{x_2^2}{b^2} - x_3 = 0$$

hyperbolický paraboloid

$$\frac{x_1^2}{a^2} - \frac{x_2^2}{b^2} = 1$$

hyperbolická válcová plocha

$$\frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} - \frac{x_3^2}{c^2} = 1$$

jednodílný hyperboloid

$$-\frac{x_1^2}{a^2} - \frac{x_2^2}{b^2} + \frac{x_3^2}{c^2} = 1$$

dvojdílný hyperboloid □

Problémy

12.1. Diskutujte jednoznačnost matice kvadratické formy, a jednoznačnost kvadratické formy ze znalosti obrazu nějaké báze.

12.2. Ukažte, že bilineární formy na prostoru V tvoří vektorový prostor a určete jeho dimenzi.

12.3. Ukažte, že kvadratické formy na prostoru V tvoří vektorový prostor a určete jeho dimenzi.

12.4. Buď V vektorový prostor nad tělesem $\mathbb{T}$ charakteristiky různé od 2. Ukažte, že každá symetrická bilineární forma $b\colon V^2 \to \mathbb{T}$ je určena hodnotami $b(v, v)$, $v \in V$. To v důsledku znamená, že ze znalosti kvadratické formy jsme schopni zrekonstruovat jednoznačnou symetrickou bilineární formu, která ji indukuje.

12.5. Platí Sylvestrův zákon setrvačnosti na komplexním prostoru?

12.6. Vyjádřete kvadratickou formu $f(x) = x^T A x$ s positivně definitní maticí $A \in \mathbb{R}^{n\times n}$ jako součet čtverců lineárních forem. (*Poznámka.* Platí dokonce, že každý nezáporný polynom se dá vyjádřit

jako zlomek, jehož čitatel a jmenovatel jsou ve tvaru součtu čtverců polynomů. Toto bylo známé jako sedmnáctý Hilbertův problém a bylo vyřešeno Emilem Artinem roku 1927. Navíc čitatel se dá vyjádřit pomocí nanejvýš 2^n čtverců a jmenovatel pomocí pouze jednoho.)

12.7. Buď $f\colon V \to \mathbb{R}$ kvadratická forma a $h\colon V \to V$ isomorfismus.

(a) Ukažte, že $f \circ h$ je kvadratická forma na V.

(b) Ukažte, že obě kvadratické formy f, $f \circ h$ mají stejnou signaturu.

12.8. Buď f kvadratická forma na reálném prostoru V a buď (p, q, z) její signatura. Ukažte, že v prostoru V existuje podprostor dimenze $z+\min(p, q)$, na kterém je kvadratická forma f nulová, a větší takový podprostor neexistuje.

Shrnutí ke kapitole 12. Kvadratické formy

Kvadratickou formou v prostoru $\mathbb{R}^n$ je polynom $f(x) = x^T Ax$, kde $A \in \mathbb{R}^{n\times n}$ je symetrická. U obecných prostorů pracujeme v souřadnicích podobně jako u lineárních zobrazení. Změníme-li souřadný systém, matice se změní na matici $S^T AS$, kde S je matice přechodu mezi souřadnými systémy. Ústřední věta této kapitoly, Sylvestrův zákon setrvačnosti, pak tvrdí, že vždycky můžeme najít souřadný systém, ve kterém je matice diagonální. Navíc má tato diagonální matice pro pevnou kvadratickou formu vždy stejnou signaturu – stejný počet kladných a záporných čísel (to je hlavní poselství věty!). To nám umožňuje kvadratické formy jednoduše klasifikovat a popisovat. Signaturu pro danou matici určíme jednoduše pomocí elementárních úprav aplikovaných jak na řádky matice, tak symetricky i na sloupce. Tím získáme mimochodem efektivní metodu na testování positivní (semi-)definitnosti aj. Teorie kvadratických forem nám také dává účinný nástroj jak analyzovat polynomiální rovnice typu $x^T Ax + b^T x + c = 0$ a jak charakterizovat různé kuželosečky a kvadriky, jako jsou například elipsoidy.

Kvadratické formy úzce souvisí s bilineárními formami. Reálný skalární součin je vždy bilineární formou. Speciálně v prostoru $\mathbb{R}^n$ mají bilineární formy tvar $b(x, y) = x^T Ay$.

Kapitola 13

Maticové rozklady

Maticové rozklady byly zařazeny do *Top 10* algoritmů 20. století. S několika rozklady jsme se již setkali (Choleského rozklad, spektrální rozklad, ...), ale QR rozklad (bude o něm řeč později) se v seznamu objevuje skrytě ještě jednou, protože je základem QR algoritmu. Jeho důležitost je tedy patrná.

Top 10 algoritmy 20. století podle [Dongarra and Sullivan, 2000; Cipra, 2000] jsou:

1. *Metoda Monte Carlo* (1946, J. von Neumann, S. Ulam, a N. Metropolis)

   Pomocí simulací s náhodnými čísly spočítáme přibližná řešení problémů, které jsou velmi těžké na to spočítat jejich řešení přesně.

2. *Simplexová metoda pro lineární programování* (1946, G. Dantzig)

   Metoda na výpočet optimalizačních úloh s lineárním kriteriem i omezeními, tj. v jakém bodu polyedru se nabyde nejmenší hodnota lineární funkce (viz sekce 1.7 a příklad 7.18).

3. *Iterační metody Krylovových podprostorů* (1950, M. Hestenes, E. Stiefel, C. Lanczos)

   Metody na řešení velkých a řídkých soustav lineárních rovnic.

4. *Dekompozice matic* (1951, A. Householder)

   Maticové rozklady jako je např. Choleského rozklad, LU rozklad, QR rozklad, spektrální rozklad, Schurova triangularizace nebo SVD rozklad.

5. *Překladač Fortranu* (1957, J. Backus)

   Informatikům netřeba představovat překladač programovacího jazyka Fortran, který jako jeden z prvních umožnil náročné numerické výpočty.

6. *QR algoritmus*, (1961, J. Francis)

   Algoritmus pro výpočet vlastních čísel.

7. *Quicksort* (1962, A. Hoare)

   Algoritmus na třídění.

8. *Rychlá Fourierova transformace* (1965, J. Cooley, J. Tukey)

   Pro rychlé násobení polynomů a čísel, zpracování signálu a mnoho dalších věcí, viz příklad 10.36.

9. *Algoritmus „Integer relation detection“* (1977, H. Ferguson, R. Forcade)

   Zobecnění Euklidova algoritmu postupného dělení na problém: Dáno n reálných čísel $x_1, \ldots, x_n$, existuje netriviální celočíselná kombinace $a_1x_1 + \ldots + a_nx_n = 0$?

10. *Metoda více pólů – „Fast multipole algorithm“* (1987, L. Greengard, V. Rokhlin)

    Simulace v problému výpočtu sil dalekého dosahu v úloze n těles. Seskupuje zdroje ležící u sebe a pracuje s nimi jako s jediným.

V této kapitole budeme uvažovat standardní skalární součin v $\mathbb{R}^n$ a eukleidovskou normu (pokud není explicitně řečeno jinak).

13.1 Householderova transformace

Motivace k této sekci je následující. Mějme dané dva vektory $x, y \in \mathbb{R}^n$ a chceme najít lineární zobrazení $x \mapsto Hx$ takové, které zobrazí vektor x na vektor y. Aby mělo lineární zobrazení pěkné vlastnosti, požadujeme navíc, aby matice H byla ortogonální. Protože takové lineární zobrazení zachovává délky (věta 8.54), musí nutně mít oba vektory x, y stejnou normu.

Takovéto zobrazení vždy existuje, a za matici H lze zvolit vhodnou Householderovu matici. Připomeňme (příklad 8.53), že Householderova matice je definována jako $H(x) := I_n - \frac{2}{x^Tx}xx^T$, kde $o \neq x \in \mathbb{R}^n$. Tato matice je ortogonální a symetrická. Vhodnou volbou vektorů x, y pak může nahradit elementární matice při výpočtu odstupňovaného tvaru matice. Tento postup se nazývá Householderova transformace.[1)]

Věta 13.1 (Householderova transformace)**.** *Pro každé $x, y \in \mathbb{R}^n$, $x \neq y$, $\|x\|_2 = \|y\|_2$ platí $y = H(x-y)x$.*

Důkaz. Počítejme

$$\begin{aligned} H(x-y)x &= \left(I_n - \frac{2}{(x-y)^T(x-y)}(x-y)(x-y)^T\right)x = \\ &= x - \frac{2(x-y)^Tx}{(x-y)^T(x-y)}(x-y) = x - \frac{2\|x\|_2^2 - 2y^Tx}{(x-y)^T(x-y)}(x-y) = \\ &= x - \frac{\|x\|_2^2 + \|y\|_2^2 - 2y^Tx}{\|x-y\|_2^2}(x-y) = x - (x-y) = y. \end{aligned}$$

□

Householderova matice tedy převádí jeden vybraný vektor x na jiný y se stejnou normou tím, že vynásobíme vektor x zleva vhodnou Householderovou maticí. Speciálně lze převést každý vektor na vhodný násobek jednotkového vektoru:

Důsledek 13.2. *Buď $x \in \mathbb{R}^n$ a definujme*

$$H := \begin{cases} H(x - \|x\|_2 e_1), & \textit{pokud } x \neq \|x\|_2 e_1, \\ I_n, & \textit{jinak.} \end{cases}$$

Potom $Hx = \|x\|_2 e_1$.

Důkaz. Případ $x = \|x\|_2 e_1$ je jasný. Jinak použijeme větu 13.1, vektory $x, \|x\|_2 e_1$ mají stejnou normu. □

Příklad 13.3. Buď $x = (2, 2, 1)^T$. Pak $\|x\|_2 = 3$ a tedy

$$H = H(x - 3e_1) = \frac{1}{3}\begin{pmatrix} 2 & 2 & 1 \\ 2 & -1 & -2 \\ 1 & -2 & 2 \end{pmatrix}.$$

Nyní $Hx = (3, 0, 0)^T$, tedy lineární zobrazení $x \mapsto Hx$ zobrazuje vektor $x = (2, 2, 1)^T$ na násobek jednotkového vektoru. □

Mějme matici $A \in \mathbb{R}^{m \times n}$ a H sestrojíme podle prvního sloupce A. Vynásobením HA tak vynulujeme prvky v prvním sloupci A až na první z nich. Rekurzivním voláním transformace pak převedeme matici do odstupňovaného tvaru. Tento postup je tedy alternativou k elementárním řádkovým úpravám. Máme tu však ještě něco navíc, a to tzv. QR rozklad, o němž hovoříme podrobněji v následující sekci.

Podobně lze použít i Givensova matice, ale ta nuluje pouze jeden prvek (stejně jako elementární matice), takže ji musíme použít vícekrát (problém 13.2).

Jiná užitečná vlastnost Householderovy matice je uvedena v následujícím příkladu.

[1)] Alston Scott Householder (1904–1993), americký numerický matematik, transformace je z r. 1958. Byl spoluzakladatelem numerické matematiky a prý nikdy necestoval letadlem, neboť si byl vědom, kde všude při výpočtech konstrukce mohlo dojít k chybě.

Příklad 13.4. Buď $u \in \mathbb{R}^n$ vektor jednotkové délky, tj. $\|u\|_2 = 1$. Sestrojte ortogonální matici Q s prvním sloupcem rovným u. Jinými slovy, doplňte u na ortonormální bázi prostoru $\mathbb{R}^n$. Tento problém se vyskytl například v důkazu věty 10.51 o spektrálním rozkladu symetrických matic. Podle důsledku 8.24 víme, že taková ortogonální matice musí existovat, ale doposud jsme neměli pro tento problém účinný nástroj. Householderova matice ale tuto úlohy vyřeší elegantně.

Řešení: Je-li $u = e_1$, pak zřejmě stačí volit $Q = I_n$. V opačném případě můžeme zvolit matici $Q := H(e_1 - u)$, tedy Householderovu matici pro vektor $e_1 - u$. Zdůvodnění je jednoduché, první sloupec této matice je podle Householderovy transformace roven $Q_{1*} = Qe_1 = H(e_1 - u)e_1 = u$. □

13.2 QR rozklad

Věta 13.5 (QR rozklad)**.** *Pro každou matici $A \in \mathbb{R}^{m\times n}$ existuje ortogonální $Q \in \mathbb{R}^{m\times m}$ a horní trojúhelníková matice $R \in \mathbb{R}^{m\times n}$ s nezápornou diagonálou tak, že $A = QR$.*

Důkaz. Matematickou indukcí podle n, tj. počtu sloupců. Je-li $n = 1$, pak $A = a \in \mathbb{R}^m$ a pro matici H sestrojenou podle důsledku 13.2 platí $Ha = \|a\|_2 e_1$. Stačí položit $Q := H^T$ a $R := \|a\|_2 e_1$.

Indukční krok $n \leftarrow n - 1$. Aplikací důsledku 13.2 na první sloupec matice A dostaneme $HA_{*1} = \|A_{*1}\|_2 e_1$. Tedy HA je tvaru

$$HA = \begin{pmatrix} \alpha & b^T \\ o & B \end{pmatrix},$$

kde $B \in \mathbb{R}^{(m-1)\times(n-1)}$ a $\alpha = \|A_{*1}\|_2 \geq 0$. Podle indukčního předpokladu existuje rozklad $B = Q'R'$, kde $Q' \in \mathbb{R}^{(m-1)\times(m-1)}$ je ortogonální a $R' \in \mathbb{R}^{(m-1)\times(n-1)}$ horní trojúhelníková s nezápornou diagonálou. Upravme

$$\begin{pmatrix} 1 & o^T \\ o & Q'^T \end{pmatrix} HA = \begin{pmatrix} 1 & o^T \\ o & Q'^T \end{pmatrix} \begin{pmatrix} \alpha & b^T \\ o & B \end{pmatrix} = \begin{pmatrix} \alpha & b^T \\ o & R' \end{pmatrix}. \tag{13.1}$$

Označme

$$Q := H^T \begin{pmatrix} 1 & o^T \\ o & Q' \end{pmatrix}, \quad R := \begin{pmatrix} \alpha & b^T \\ o & R' \end{pmatrix}.$$

Matice Q je ortogonální a R je horní trojúhelníková s nezápornou diagonálou. Nyní rovnice (13.1) má tvar $Q^T A = R$, neboli $A = QR$ je hledaný rozklad. □

Věta dává i návod na konstrukci QR rozkladu. Poznamenejme, že v průběhu algoritmu platí invariant: $A = QR$ a Q je ortogonální. Podstata tedy spočívá v tom, že postupně R přeměníme na horní trojúhelníkovou matici. Symbol $R(j : m, j)$ značí vektor $(r_{jj}, \ldots, r_{mj})^T$.

Algoritmus 13.6 (QR rozklad)**.** Buď $A \in \mathbb{R}^{m\times n}$.

1: $Q := I_m$, $R := A$,
2: **for** $j := 1$ **to** $\min(m, n)$ **do**
3: $\quad x := R(j : m, j)$,
4: $\quad$ **if** $x \neq \|x\|_2 e_1$ **then**
5: $\quad\quad x := x - \|x\|_2 e_1$,
6: $\quad\quad H(x) := I_{m-j+1} - \frac{2}{x^T x} x x^T$,
7: $\quad\quad H := \begin{pmatrix} I_{j-1} & 0 \\ 0 & H(x) \end{pmatrix}$,
8: $\quad\quad R := HR$, $Q := QH$,
9: $\quad$ **end if**
10: **end for**

Výstup: $A = QR$.

Příklad 13.7 (QR rozklad)**.** Buď

$$A = \begin{pmatrix} 0 & -20 & -14 \\ 3 & 27 & -4 \\ 4 & 11 & -2 \end{pmatrix}.$$

První iterace:

$$x = A_{*1} - \|A_{*1}\| e_1 = (-5, 3, 4)^T,$$
$$Q_1 = I_3 - 2\frac{xx^T}{x^Tx} = \frac{1}{25}\begin{pmatrix} 0 & 15 & 20 \\ 15 & 16 & -12 \\ 20 & -12 & 9 \end{pmatrix}, \quad Q_1 A = \begin{pmatrix} 5 & 25 & -4 \\ 0 & 0 & -10 \\ 0 & -25 & -10 \end{pmatrix}.$$

Druhá iterace:

$$x = (0, -25)^T - 25e_1 = (-25, -25)^T,$$
$$Q_2 = I_2 - 2\frac{xx^T}{x^Tx} = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}, \quad Q_2 \begin{pmatrix} 0 & -10 \\ -25 & -10 \end{pmatrix} = \begin{pmatrix} 25 & 10 \\ 0 & 10 \end{pmatrix}.$$

Výsledek:

$$Q = Q_1 \begin{pmatrix} 1 & 0 \\ 0 & Q_2 \end{pmatrix} = \frac{1}{25}\begin{pmatrix} 0 & -20 & -15 \\ 15 & 12 & -16 \\ 20 & -9 & 12 \end{pmatrix}, \quad R = \begin{pmatrix} 5 & 25 & -4 \\ 0 & 25 & 10 \\ 0 & 0 & 10 \end{pmatrix}.$$

□

QR rozklad je jednoznačný jen za určitých předpokladů. Např. pro nulovou matici $A = 0$ je $R = 0$ a Q libovolná ortogonální matice, tedy jednoznačnost tu nenastává.

Tvrzení 13.8 (Jednoznačnost QR rozkladu)**.** *Pro regulární matici $A \in \mathbb{R}^{n\times n}$ je QR rozklad jednoznačný a R má na diagonále kladné hodnoty.*

Důkaz. Ze vztahu $A = QR$ plyne, že R je regulární, a proto musí mít nenulovou, tudíž kladnou, diagonálu.

Jednoznačnost ukážeme sporem. Nechť A má dva různé rozklady $A = Q_1R_1 = Q_2R_2$. Pak $Q_2^TQ_1 = R_2R_1^{-1}$, a tuto matici označíme jako U. Zřejmě U je ortogonální (je to součin ortogonálních matic Q_2^T a Q_1) a horní trojúhelníková (je to součin horních trojúhelníkových matic R_2 a R_1^{-1}). Speciálně, první sloupec U má tvar $U_{*1} = (u_{11}, 0, \dots, 0)^T$, kde $u_{11} > 0$. Aby měl jednotkovou velikost, musí $u_{11} = 1$ a proto $U_{*1} = e_1$. Druhý sloupec je kolmý na první, proto $u_{21} = 0$, a aby měl jednotkovou velikost, musí $u_{22} = 1$. Tedy $U_{*2} = e_2$. Atd. pokračujeme dále až dostaneme $U = I_n$, z čehož $Q_1 = Q_2$ a $R_1 = R_2$. To je spor. □

Věta se dá zobecnit i na případ kdy $A \in \mathbb{R}^{m\times n}$ má lineárně nezávislé sloupce. Pak matice R a prvních n sloupců Q je jednoznačně určeno, a diagonála matice R je kladná.

Poznámka 13.9 (QR rozklad pomocí Gramovy–Schmidtovy ortogonalizace)**.** QR rozklad matice $A \in \mathbb{R}^{m\times n}$ lze sestrojit i pomocí Gramovy–Schmidtovy ortogonalizace. Prakticky se to sice nedělá, protože právě díky použití ortogonálních matic je Householderova transformace numericky lepší, ale pro pochopení vztahu obou metod je to užitečné.

Základní myšlenka je následující: Zatímco Householderovými transformacemi (tj., sekvencí ortogonálních matic) upravujeme matici A na horní trojúhelníkovou, Gramova–Schmidtova ortogonalizace funguje přesně naopak – pomocí sekvence vhodných horních trojúhelníkových matic upravíme A na matici s ortonormálními sloupci.

Konkrétně popíšeme Gramovu–Schmidtovu ortogonalizaci takto. Mějme matici $A \in \mathbb{R}^{m\times n}$, jejíž sloupce chceme zortonormalizovat. Budeme postupovat podle algoritmu 8.22 s tím, že vektory $x_1, \dots, x_n$ jsou sloupce matice A a v průběhu výpočtu budeme sloupce matice A nahrazovat vektory y_k a z_k. Krok 2 algoritmu má tvar

$$y_k := x_k - \sum_{j=1}^{k-1} \langle x_k, z_j \rangle z_j.$$

Tuto operaci vyjádříme maticově tak, že matici A vynásobíme zprava maticí

$$\begin{pmatrix} 1 & & & \alpha_1 & \\ & \ddots & & \vdots & \\ & & 1 & \alpha_{k-1} & \\ & & & 1 & \\ & & & & 1 \end{pmatrix},$$

kde $\alpha_j = -\langle x_k, z_j \rangle$, $j = 1, \dots, k-1$. Podobně krok 3, který má tvar

$$z_k := \frac{1}{\|y_k\|} y_k,$$

vyjádříme vynásobením matice A zprava diagonální maticí s prvky $1, \dots, 1, \frac{1}{\|y_k\|}, 1, \dots, 1$ na diagonále. Protože obě matice, kterými jsme násobili A zprava, jsou horní trojúhelníkové, můžeme celou ortogonalizaci vyjádřit jako

$$AR_1 \dots R_\ell = Q,$$

kde $R_1, \dots, R_\ell$ jsou horní trojúhelníkové matice a Q má ortonormální sloupce. Hledaný QR nyní dostaneme tak, že vyjádříme $A = QR$, kde $R := (R_1 \dots R_\ell)^{-1}$ je rovněž horní trojúhelníková matice.

Poznámka 13.10 (QR rozklad pomocí Choleského rozkladu)**.** Buď $A \in \mathbb{R}^{m \times n}$ hodnosti n a $A = QR$ její hledaný QR rozklad. Vyjádříme matice Q, R blokově jako

$$A = QR = \begin{pmatrix} \tilde{Q} & \tilde{Q}' \end{pmatrix} \begin{pmatrix} \tilde{R} \\ 0 \end{pmatrix} = \tilde{Q}\tilde{R},$$

kde $\tilde{R}$ je regulární horní trojúhelníková matice s kladnou diagonálou. Pak $A^T A = R^T Q^T Q R = R^T R = \tilde{R}^T \tilde{R}$. Matici $\tilde{R}$ tedy můžeme sestrojit Choleského rozkladem z matice vzniklé součinem $A^T A$. Z rovnice $A = \tilde{Q}\tilde{R}$ pak jednoduše vyjádříme $\tilde{Q} = A\tilde{R}^{-1}$. Zbylé sloupečky matice Q dopočítáme libovolně tak, aby matice Q byla ortogonální (že to lze máme zaručeno důsledkem 8.24).

13.3 Aplikace QR rozkladu

QR rozklad se dá použít na řešení mnoha úloh, se kterými jsme se doposud setkali. Jeho hlavní výhodou je, že pracuje s ortogonální maticí Q. Protože ortogonální matice zachovávají normu (věta 8.54(2)), tak se zaokrouhlovací chyby příliš nezvětšují. To je důvod, proč se ortogonální matice hojně využívají v numerických metodách.

QR rozklad a soustavy rovnic

Uvažujme soustavu lineárních rovnic $Ax = b$, kde $A \in \mathbb{R}^{n \times n}$ je regulární. Řešení vypočítáme následujícím způsobem: Vypočítej QR rozklad $A = QR$. Pak soustava má tvar $QRx = b$, neboli $Rx = Q^T b$. Protože R je horní trojúhelníková matice, řešení dostaneme snadno zpětnou substitucí.

Oproti Gaussově eliminaci je tento způsob přibližně dvakrát pomalejší, na druhou stranu je numericky stabilnější a přesnější.

QR rozklad a ortogonalizace

Pro následující si nejprve zavedeme tzv. *redukovaný QR rozklad.* Nechť $A \in \mathbb{R}^{m \times n}$ má lineárně nezávislé sloupce. Pak QR rozklad rozepíšeme blokově

$$A = QR = \begin{pmatrix} \tilde{Q} & \tilde{Q}' \end{pmatrix} \begin{pmatrix} \tilde{R} \\ 0 \end{pmatrix} = \tilde{Q}\tilde{R},$$

kde $\tilde{Q} \in \mathbb{R}^{m \times n}$ tvoří prvních n sloupců matice Q a $\tilde{R}$ tvoří prvních n řádků matice R. Matice $\tilde{R}$ je regulární.

Nyní se podívejme, jak QR rozklad aplikovat k nalezení ortonormální báze daného prostoru; je to tedy alternativa ke Gramově–Schmidtově ortogonalizaci v $\mathbb{R}^m$. Nechť $A \in \mathbb{R}^{m\times n}$ má lineárně nezávislé sloupce a chceme sestrojit ortonormální bázi sloupcového prostoru $\mathcal{S}(A)$. Z rovnosti $A = \tilde{Q}\tilde{R}$ a regularity $\tilde{R}$ vyplývá (tvrzení 5.60), že $\mathcal{S}(A) = \mathcal{S}(\tilde{Q})$. Tedy ortonormální bázi $\mathcal{S}(A)$ tvoří sloupce $\tilde{Q}$.

Vzhledem k vlastnostem ortogonálních matic pak $\tilde{Q}'$ tvoří ortonormální bázi $\mathrm{Ker}(A^T)$, protože $\mathrm{Ker}(A^T)$ je ortogonální doplněk k $\mathcal{S}(A)$. Z QR rozkladu matice A resp. A^T tedy dokážeme vyčíst ortonormální bázi všech základních maticových prostorů – řádkového, sloupcového a jádra.

QR rozklad a rozšíření na ortonormální bázi

Několikrát jsme v tomto textu použili vlastnost, že ortonormální systém vektorů jde rozšířit na ortonormální bázi. Zaměříme se na speciální případ jak rozšířit vektor jednotkové velikosti na ortonormální bázi. Tento případ se objevil např. v důkazu věty 10.51 o spektrálním rozkladu symetrických matic.

Buď $a \in \mathbb{R}^n$, $\|a\|_2 = 1$, a buď $a = Qr$ jeho QR rozklad. Aby byl vektor r, bráno jako matice s jedním sloupcem, v horním trojúhelníkovém tvaru s nezápornou diagonálou, musí mít tvar $r = (\alpha, 0, \dots, 0)^T$ pro nějaké $\alpha \geq 0$. Protože $\|a\|_2 = 1$ a Q je ortogonální, musí $\|r\|_2 = 1$. Tudíž $r = e_1$, z čehož $a = Q_{*1}$. Vektor a tak leží v prvním sloupci Q a ostatní sloupce představují jeho rozšíření na ortonormální bázi.

QR rozklad a projekce do podprostoru

Nechť $A \in \mathbb{R}^{m\times n}$ má lineárně nezávislé sloupce. Víme (věta 8.43), že projekce vektoru $x \in \mathbb{R}^m$ do sloupcového prostoru $\mathcal{S}(A)$ je tvaru $x' = A(A^TA)^{-1}A^Tx$. Výraz můžeme zjednodušit s použitím redukovaného QR rozkladu $A = \tilde{Q}\tilde{R}$:

$$\begin{aligned} A(A^TA)^{-1}A^T &= \tilde{Q}\tilde{R}(\tilde{R}^T\tilde{Q}^T\tilde{Q}\tilde{R})^{-1}\tilde{R}^T\tilde{Q}^T = \tilde{Q}\tilde{R}(\tilde{R}^T\tilde{R})^{-1}\tilde{R}^T\tilde{Q}^T = \\ &= \tilde{Q}\tilde{R}\tilde{R}^{-1}(\tilde{R}^T)^{-1}\tilde{R}^T\tilde{Q}^T = \tilde{Q}\tilde{Q}^T. \end{aligned}$$

Matice projekce je tedy $\tilde{Q}\tilde{Q}^T$ a vektor x se projektuje na vektor $x' = \tilde{Q}\tilde{Q}^Tx$.

QR rozklad a metoda nejmenších čtverců

Metoda nejmenších čtverců (sekce 8.5) spočívá v přibližném řešení přeurčené soustavy rovnic $Ax = b$, kde $A \in \mathbb{R}^{m\times n}$, $m > n$. Nechť A má hodnost n, pak přibližné řešení metodou nejmenších čtverců je

$$\begin{aligned} x &= (A^TA)^{-1}A^Tb = (\tilde{R}^T\tilde{Q}^T\tilde{Q}\tilde{R})^{-1}\tilde{R}^T\tilde{Q}^Tb = \\ &= \tilde{R}^{-1}(\tilde{R}^T)^{-1}\tilde{R}^T\tilde{Q}^Tb = \tilde{R}^{-1}\tilde{Q}^Tb. \end{aligned}$$

Jinými slovy, x získáme jako řešení regulární soustavy $\tilde{R}x = \tilde{Q}^Tb$, a to zpětnou substitucí protože matice $\tilde{R}$ je horní trojúhelníková. Povšimněme si analogie s řešením regulární soustavy $Ax = b$, které vedlo na $Rx = Q^Tb$; nyní máme oříznutou soustavu $\tilde{R}x = \tilde{Q}^Tb$.

QR algoritmus

QR algoritmus[2)] je metoda na výpočet vlastních čísel matice $A \in \mathbb{R}^{n\times n}$, která se stala základem soudobých efektivních metod.

Algoritmus 13.11 (QR algoritmus)**.** Buď $A \in \mathbb{R}^{n\times n}$.

1: $A_0 := A$, $i := 0$,
2: **while not** splněna ukončovací podmínka **do**
3: sestroj QR rozklad matice A_i, tj. $A_i = QR$,
4: $A_{i+1} := RQ$,

[2)] Autory jsou anglický informatik John G.F. Francis a ruská matematička Vera Nikolaevna Kublanovskaya, kteří jej vyvinuli nezávisle r. 1961.

5: $i := i + 1$,
6: **end while**

Výstup: A_i.

Tvrzení 13.12. *Matice $A_0, A_1, \ldots$ jsou si navzájem podobné.*

Důkaz. $A_{i+1} = RQ = I_n RQ = Q^T QRQ = Q^T A_i Q$. □

Matice A_i na výstupu je podobná s A, a má tím pádem i stejná vlastní čísla. Jak je zjistíme? Algoritmus vesměs konverguje (případy kdy nekonverguje jsou řídké, skoro umělé; dlouho nebyl znám případ kdy by nekonvergoval) k blokově horní trojúhelníkové matici s bloky o velikosti 1 a 2. Bloky o velikosti 1 jsou vlastní čísla, a z bloků o velikosti 2 jednoduše dopočítáme dvojice komplexně sdružených vlastních čísel.

Příklad 13.13. Iterace QR algoritmu pro danou matici A:

$$A = \begin{pmatrix} 2 & 4 & 2 \\ 4 & 2 & 2 \\ 2 & 2 & -1 \end{pmatrix} \to \begin{pmatrix} 6.1667 & -2.4623 & 0.8616 \\ -2.4623 & -1.2576 & -0.2598 \\ 0.8616 & -0.2598 & -1.9091 \end{pmatrix} \to$$
$$\to \begin{pmatrix} 6.9257 & 0.7725 & 0.2586 \\ 0.7725 & -1.9331 & 0.0224 \\ 0.2586 & 0.0224 & -1.9925 \end{pmatrix} \to \begin{pmatrix} 6.9939 & -0.2225 & 0.0742 \\ -0.2225 & -1.9945 & -0.0018 \\ 0.0742 & -0.0018 & -1.9994 \end{pmatrix} \to$$
$$\to \begin{pmatrix} 6.9995 & 0.0636 & 0.0212 \\ 0.0636 & -1.9996 & 0.0001 \\ 0.0212 & 0.0001 & -1.9999 \end{pmatrix} \to \begin{pmatrix} 7 & -0.0182 & 0.0061 \\ -0.0182 & -2 & -10^{-5} \\ 0.0061 & -10^{-5} & -2 \end{pmatrix}$$

Zdrojový kód pro Matlab / Octave:

```
A=[2 4 2; 4 2 2; 2 2 -1];
for i=1:5
  [Q,R]=qr(A);
  A=R*Q,
end
```

Symetrická matice konverguje k diagonální. Přesnost vypočítaných vlastních čísel určuje věta 10.56 o Gerschgorinových discích. □

13.4 SVD rozklad

Stejně jako QR rozklad je SVD rozklad[3)] jednou ze základních technik v numerických výpočtech.

Věta 13.14 (SVD rozklad)**.** *Buď $A \in \mathbb{R}^{m\times n}$, $q := \min\{m, n\}$. Pak existuje diagonální matice $\Sigma \in \mathbb{R}^{m\times n}$ s prvky $\sigma_{11} \geq \ldots \geq \sigma_{qq} \geq 0$ a ortogonální matice $U \in \mathbb{R}^{m\times m}$, $V \in \mathbb{R}^{n\times n}$ tak, že $A = U\Sigma V^T$.*

Ideu důkazu uvádíme za algoritmem 13.17, který konstruuje SVD rozklad. Kladným číslům na diagonále $\sigma_{11}, \ldots, \sigma_{rr}$ říkáme *singulární čísla* matice A a značíme je obvykle $\sigma_1, \ldots, \sigma_r$. Zjevně $r = \operatorname{rank}(A)$.

Věta 13.15 (Vztah singulárních a vlastních čísel)**.** *Buď $A \in \mathbb{R}^{m\times n}$, $r = \operatorname{rank}(A)$, a nechť $A^T A$ má vlastní čísla $\lambda_1 \geq \ldots \geq \lambda_n$. Pak singulární čísla matice A jsou $\sigma_i = \sqrt{\lambda_i}$, $i = 1, \ldots, r$.*

Důkaz. Nechť $A = U\Sigma V^T$ je SVD rozklad A. Pak

$$A^T A = V\Sigma^T U^T U\Sigma V^T = V\Sigma^T \Sigma V^T = V \operatorname{diag}(\sigma_1^2, \ldots, \sigma_q^2, 0, \ldots, 0)V^T,$$

což je spektrální rozklad positivně definitní matice $A^T A$. Tudíž $\lambda_i = \sigma_i^2$. □

[3)] Zkratka za *Singular value decomposition*, rozklad na singulární čísla. Byl objeven r. 1873 nezávisle řadou autorů jako byli např. Ital Eugenio Beltrami, Francouz Marie E. Camille Jordan, Angličan James Sylvester, Němec Erhard Schmidt nebo Švýcar Hermann Weyl.

Příklad 13.16. Buď $Q \in \mathbb{R}^{n\times n}$ ortogonální. Pak $Q^TQ = I_n$ má vlastní čísla samé jedničky. Tedy ortogonální matice Q má singulární čísla také samé jedničky. □

Důkaz věty prozradil navíc, že matice V je ortogonální maticí ze spektrálního rozkladu matice A^TA. Podobně, matice U je ortogonální maticí ze spektrálního rozkladu AA^T:

$$AA^T = U\Sigma V^TV\Sigma^TU^T = U\Sigma\Sigma^TU^T = U\operatorname{diag}(\sigma_1^2,\dots,\sigma_q^2,0,\dots,0)U^T.$$

Bohužel, spektrální rozklady matic A^TA a AA^T nemůžeme použít ke konstrukci SVD rozkladu, protože nejsou jednoznačné. Použít můžeme jen jeden a druhý dopočítat trochu jinak.

Algoritmus 13.17 (SVD rozklad)**.** Buď $A \in \mathbb{R}^{m\times n}$.

1: Sestroj $V\Lambda V^T$ spektrální rozklad matice A^TA;
2: $r := \operatorname{rank}(A)$;
3: $\sigma_i := \sqrt{\lambda_i}$, $i = 1,\dots,r$;
4: $S := \operatorname{diag}(\sigma_1,\dots,\sigma_r)$;
5: V_1 je matice tvořená prvními r sloupci V;
6: $U_1 := AV_1S^{-1}$;
7: doplň U_1 na ortogonální matici $U = (U_1 \mid U_2)$;

Výstup: $A = U\Sigma V^T$ je SVD rozklad A.

Důkaz. Z věty 13.15 víme, že $\sigma_1,\dots,\sigma_r$ jsou hledaná singulární čísla a zjevně V je ortogonální. Musíme dokázat, že U_1 má ortonormální sloupce a $A = U\Sigma V^T$.

Z rovnosti $A^TA = V\Lambda V^T$ odvodíme $\Lambda = V^TA^TAV$ a odříznutím posledních $n-r$ řádků a sloupců dostaneme matici $\operatorname{diag}(\lambda_1,\dots,\lambda_r) = V_1^TA^TAV_1$. Nyní je vidět, že U_1 má ortonormální sloupce, neboť

$$\begin{aligned} U_1^TU_1 &= (S^{-1})^TV_1^TA^TAV_1S^{-1} = (S^{-1})^T\operatorname{diag}(\lambda_1,\dots,\lambda_r)S^{-1} = \\ &= (S^{-1})^TS^2S^{-1} = I_r. \end{aligned}$$

Zbývá ukázat, že $A = U\Sigma V^T$, neboli $\Sigma = U^TAV$. Rozložme $V = (V_1 \mid V_2)$. Odříznutím prvních r řádků a sloupců v matici $\Lambda = V^TA^TAV$ dostaneme $0 = V_2^TA^TAV_2$, z čehož $AV_2 = 0$ (důsledek 8.42). Nyní s využitím rovnosti $AV_1 = U_1S$ máme

$$U^TAV = U^TA(V_1 \mid V_2) = (U^TU_1S \mid U^TAV_2) = \begin{pmatrix} S & 0 \\ 0 & 0 \end{pmatrix} = \Sigma.$$ □

Příklad 13.18 (SVD rozklad)**.** Mějme

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

Spektrální rozklad matice A^TA:

$$A^TA = \begin{pmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \end{pmatrix} \begin{pmatrix} 6 & 0 \\ 0 & 4 \end{pmatrix} \begin{pmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \end{pmatrix} \equiv V\Lambda V^T.$$

Určení S:

$$S = \begin{pmatrix} \sqrt{6} & 0 \\ 0 & 2 \end{pmatrix}.$$

Určení U_1 (v tomto příkladu máme $V_1 = V$):

$$U_1 = AV_1S^{-1} = \begin{pmatrix} \frac{\sqrt{3}}{3} & 0 \\ \frac{\sqrt{3}}{3} & \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{3}}{3} & \frac{\sqrt{2}}{2} \end{pmatrix}.$$

Doplnění U_1 ortogonální matici U:

$$U = \begin{pmatrix} \frac{\sqrt{3}}{3} & 0 & -\frac{\sqrt{6}}{3} \\ \frac{\sqrt{3}}{3} & \frac{\sqrt{2}}{2} & \frac{\sqrt{6}}{6} \\ -\frac{\sqrt{3}}{3} & \frac{\sqrt{2}}{2} & -\frac{\sqrt{6}}{6} \end{pmatrix}.$$

Výsledný SVD rozklad:

$$A = \begin{pmatrix} \frac{\sqrt{3}}{3} & 0 & -\frac{\sqrt{6}}{3} \\ \frac{\sqrt{3}}{3} & \frac{\sqrt{2}}{2} & \frac{\sqrt{6}}{6} \\ -\frac{\sqrt{3}}{3} & \frac{\sqrt{2}}{2} & -\frac{\sqrt{6}}{6} \end{pmatrix} \begin{pmatrix} \sqrt{6} & 0 \\ 0 & 2 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \end{pmatrix} \equiv U\Sigma V^T. \qquad \square$$

Podobně jako u QR rozkladu, tak i pro SVD rozklad existuje redukovaná verze, tzv. *redukovaný* (nebo též *tenký*) SVD rozklad. Buď $A = U\Sigma V^T$ hodnosti $r > 0$. Rozložme $U = (U_1 \mid U_2)$, $V = (V_1 \mid V_2)$ na prvních r sloupců a zbytek, a dále $S := \operatorname{diag}(\sigma_1, \ldots, \sigma_r)$. Pak

$$A = U\Sigma V^T = \begin{pmatrix} U_1 & U_2 \end{pmatrix} \begin{pmatrix} S & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} V_1^T \\ V_2^T \end{pmatrix} = U_1 S V_1^T.$$

Redukovaný SVD používá jen část informace z SVD rozkladu, ale tu podstatnou, ze které můžeme plný SVD rozklad zrekonstruovat (doplněním U_1, V_1 na ortogonální matice). Redukovaný SVD jsme trochu používali už v důkazu algoritmu 13.17.

13.5 Aplikace SVD rozkladu

SVD a ortogonalizace

SVD rozklad lze použít k nalezení ortonormální báze (nejen) sloupcového prostoru $\mathcal{S}(A)$. Na rozdíl od dosavadních přístupů nemusíme předpokládat lineární nezávislost sloupců matice A.

Věta 13.19. *Nechť $A = U\Sigma V^T = U_1 S V_1^T$ je SVD rozklad matice $A \in \mathbb{R}^{m\times n}$. Pak*

(1) Sloupce U_1 tvoří ortonormální bázi prostoru $\mathcal{S}(A)$.

(2) Sloupce V_1 tvoří ortonormální bázi prostoru $\mathcal{R}(A)$.

(3) Sloupce V_2 tvoří ortonormální bázi prostoru $\operatorname{Ker}(A)$.

Důkaz.

(1) Redukovaný SVD rozklad je $A = U_1 S V_1^T$. Přenásobením V_1 zprava dostaneme $AV_1 = U_1 S$. Nyní, $\mathcal{S}(A) \ni \mathcal{S}(AV_1) = \mathcal{S}(U_1 S) = \mathcal{S}(U_1)$ díky regularitě matice S. Protože $\operatorname{rank}(A) = \operatorname{rank}(U_1)$, máme rovnost $\mathcal{S}(A) = \mathcal{S}(U_1)$.

(2) Plyne z předchozího díky $\mathcal{R}(A) = \mathcal{S}(A^T)$.

(3) Z transpozice $A^T = V_1 S U_1^T$ dostáváme redukovaný SVD rozklad matice A^T. Tedy sloupce V_1 tvoří ortonormální bázi prostoru $\mathcal{S}(A^T) = \mathcal{R}(A) = \operatorname{Ker}(A)^\perp$. Proto sloupce V_2, které doplňují sloupce V_1 na ortonormální bázi $\mathbb{R}^n$, představují ortonormální bázi $\operatorname{Ker}(A)$. $\square$

SVD a projekce do podprostoru

Pomocí SVD rozkladu můžeme snadno vyjádřit matici projekce do sloupcového (a řádkového) prostoru dané matice. Dokonce k tomu nepotřebujeme předpoklad na lineární nezávislost sloupců matice, což bylo potřeba ve větě 8.43.

Věta 13.20. *Nechť $A = U\Sigma V^T = U_1 S V_1^T$ je SVD rozklad matice $A \in \mathbb{R}^{m\times n}$. Pak matice projekce do*

(1) sloupcového prostoru $\mathcal{S}(A)$ je $U_1 U_1^T$,

(2) řádkového prostoru $\mathcal{R}(A)$ je $V_1 V_1^T$,

Důkaz.

(1) Z věty 13.19 je $\mathcal{S}(A) = \mathcal{S}(U_1)$. Sloupce U_1 jsou lineárně nezávislé, a proto matice projekce má dle věty 8.43 tvar $U_1(U_1^T U_1)^{-1} U_1^T = U_1 (I_r)^{-1} U_1^T = U_1 U_1^T$.

(2) Plyne z předchozího díky $\mathcal{R}(A) = \mathcal{S}(A^T)$. □

Podobným způsobem lze odvodit i vzoreček pro přibližné řešení soustavy $Ax = b$ metodou nejmenších čtverců. Nicméně, později ve větě 13.28 ukážeme silnější výsledek.

SVD a geometrie lineárního zobrazení

Buď $A \in \mathbb{R}^{n\times n}$ regulární matice a studujme obraz jednotkové koule při zobrazení $x \mapsto Ax$. Z SVD rozkladu $A = U\Sigma V^T$ plyne, že lineární zobrazení lze rozložit na složení tří základních zobrazení: ortogonální zobrazení s maticí V^T, škálování podle Σ a ortogonální zobrazení s maticí U. Konkrétně, zobrazení s maticí V^T zobrazí kouli na sebe sama, Σ ji zdeformuje na elipsoid a U ji otočí/převrátí. Tedy výsledkem bude elipsoid se středem v počátku, poloosy jsou ve směrech sloupců U a délky mají velikost $\sigma_1, \ldots, \sigma_n$.

Hodnota $\frac{\sigma_1}{\sigma_n} \geq 1$ se nazývá *míra deformace* a kvantitativně udává, jak moc zobrazení deformuje geometrické útvary. Je-li hodnota rovna 1, elipsoid bude mít tvar koule, a naopak čím větší bude hodnota, tím protáhlejší bude elipsoid. Význam této hodnoty je ale nejenom geometrický. V numerické matematice se podíl $\frac{\sigma_1}{\sigma_n}$ nazývá číslo podmíněnosti a čím je větší, tím hůře podmíněná je matice A ve smyslu, že vykazuje špatné numerické vlastnosti – zaokrouhlování v počítačové aritmetice s pohyblivou řádkovou čárkou způsobuje chyby (viz sekce 1.3).

Empirické pravidlo říká, že je-li číslo podmíněnosti řádově 10^k, pak při výpočtech s maticí (inverze, řešení soustav, atp.) ztrácíme přesnost o k desetinných míst. Ortogonální matice mají číslo podmíněnosti rovné 1, a proto se v numerické matematice často používají. Naproti tomu např. Hilbertovy matice z příkladu 3.46 mají číslo podmíněnosti velmi vysoké:

n	číslo podmíněnosti H_n
3	≈ 500
5	$\approx 10^5$
10	$\approx 10^{13}$
15	$\approx 10^{17}$

SVD a numerický rank

Hodnost matice A je rovna počtu (kladných) singulárních čísel. Nicméně, pro výpočetní účely se hodně malé kladné číslo považuje za praktickou nulu. Buď $\varepsilon > 0$, pak *numerický rank* matice A je $\max\{s;\ \sigma_s > \varepsilon\}$, tedy počet singulárních čísel větších než ε, ostatní se berou za nulová. Např. Matlab / Octave bere $\varepsilon := \max\{m, n\} \cdot \sigma_1 \cdot eps$, kde $eps \approx 2 \cdot 10^{-16}$ je přesnost počítačové aritmetiky.

SVD a low-rank aproximace

Buď $A \in \mathbb{R}^{m\times n}$ a $A = U\Sigma V^T$ její SVD rozklad. Jestliže ponecháme k největších singulárních čísel a ostatní vynulujeme $\sigma_{k+1} := 0, \ldots, \sigma_r := 0$, tak dostaneme matici

$$A' = U \operatorname{diag}(\sigma_1, \ldots, \sigma_k, 0, \ldots, 0) V^T$$

hodnosti k, která dobře aproximuje A. Navíc tato aproximace je v jistém smyslu nejlepší možná. To jest, v určité normě (viz sekce 13.7) je ze všech matic hodnosti k právě A' nejblíže matici A. Low-rank aproximaci využijeme v následujícím:

SVD a komprese dat

Předpokládejme, že matice $A \in \mathbb{R}^{m\times n}$ reprezentuje data, které chceme zkomprimovat. Pokud $\operatorname{rank}(A) = r$, tak pro redukovaný SVD rozklad $A = U_1 S V_1^T$ si potřebujeme zapamatovat $mr + r + nr = (m + n + 1)r$ hodnot. Při low-rank aproximaci $A \approx U \operatorname{diag}(\sigma_1, \ldots, \sigma_k, 0, \ldots, 0) V^T$ si stačí pamatovat jen $(m + n + 1)k$ hodnot. Tedy kompresní poměr je $k : r$. Čím menší k, tím menší objem dat si stačí pamatovat. Ale na druhou stranu, menší k značí horší aproximaci.

Příklad 13.21. Zmíněný postup ilustrujeme na kompresi obrázku (srov. příklad 3.55). Předpokládáme, že matice $A \in \mathbb{R}^{m\times n}$ reprezentuje obrázek, ve kterém pixel na pozici (i, j) má barvu s číslem a_{ij}. Následující obrázky ilustrují komprimaci pro různou volbu k. Pro $k = 480$ máme originální obrázek, pro 150 asi třetinovou kompresi bez znatelné újmy na kvalitě obrázku, při $k = 50$ už dochází k zrnění a při $k = 5$ je obrázek značně rozmazán (ale na to, že máme zhruba 1% původního objemu dat, je výsledek stále slušný).

originál ($k = 480$)

$k = 150$

$k = 50$

$k = 5$

Obrázek představuje foto z konference o numerické algebře v Gatlinburgu z r. 1964, a zaznamenává největší numerické matematiky své doby, zleva: James H. Wilkinson, Wallace Givens, George Forsythe, Alston Householder, Peter Henrici, a Fritz Bauer. Obrázek se skládá z 480×640 pixelů, SVD rozklad trval cca 5 sec (11.5.2010). Zdrojový kód pro Matlab / Octave:

```
load gatlin,
[X,S,Y] = svd(X);
figure(2), clf,
k = 150;
Xk = X(:,1:k)*S(1:k,1:k)*Y(:,1:k)';
image(Xk),
colormap(map),
axis equal, axis off,
```

□

SVD a míra regularity

Jak jsme uvedli v poznámce 9.8, determinant se jako míra regularity moc nehodí. Zato singulární čísla jsou pro to jako stvořená. Buď $A \in \mathbb{R}^{n\times n}$. Pak σ_n udává vzdálenost (v jisté normě, viz sekce 13.7) k nejbližší singulární matici, podrobněji viz [Rohn, 2004]. Takže je to v souladu s tím, co bychom si pod takovou mírou představovali. Ortogonální matice mají míru 1, naproti tomu Hilbertovy matice mají malou míru regularity, tj. jsou téměř singulární:

n	$\sigma_n(H_n)$
3	≈ 0.0027
5	$\approx 10^{-6}$
10	$\approx 10^{-13}$
15	$\approx 10^{-18}$

13.6 Pseudoinverzní matice

Je přirozenou snahou zobecnit pojem inverze matice i na singulární nebo obdélníkové matice. Takové zobecněné inverzi se říká *pseudoinverze* a existuje několik druhů. Nejčastější je tzv. Mooreova–Penroseova pseudoinverze[4)], která spočívá na SVD rozkladu.

Definice 13.22 (Mooreova–Penroseova pseudoinverze)**.** Buď $A \in \mathbb{R}^{m\times n}$ matice s redukovaným SVD rozkladem $A = U_1 S V_1^T$. Je-li $A \neq 0$, pak její *pseudoinverze* je $A^\dagger = V_1 S^{-1} U_1^T \in \mathbb{R}^{n\times m}$. Pro $A = 0$ definujeme pseudoinverzi předpisem $A^\dagger = A^T$.

Příklad 13.23. Pseudoinverze nenulového vektoru $a \in \mathbb{R}^n$ je $a^\dagger = \frac{1}{a^T a} a^T$, speciálně např. $(1,1,1,1)^\dagger = \frac{1}{4}(1,1,1,1)^T$. □

Tvrzení 13.24 (Vlastnosti pseudoinverze)**.** *Buď $A \in \mathbb{R}^{m\times n}$, pak*

(1) Je-li A regulární, tak $A^{-1} = A^\dagger$,
(2) $(A^\dagger)^\dagger = A$,
(3) $(A^T)^\dagger = (A^\dagger)^T$,
(4) $A = AA^\dagger A$,
(5) $A^\dagger = A^\dagger A A^\dagger$,
(6) $AA^\dagger$ je symetrická,
(7) $A^\dagger A$ je symetrická,
(8) $A^\dagger = (A^T A)^\dagger A^T$,
(9) má-li A lineárně nezávislé sloupce, pak $A^\dagger = (A^T A)^{-1} A^T$,
(10) má-li A lineárně nezávislé řádky, pak $A^\dagger = A^T (AA^T)^{-1}$.

Důkaz. Vlastnosti se dokážou jednoduše z definice. Pro ilustraci ukážeme jen dvě vlastnosti, zbytek necháváme čtenáři.

(4) Z definice $AA^\dagger A = U_1 S V_1^T V_1 S^{-1} U_1^T U_1 S V_1^T = U_1 S S^{-1} S V_1^T = U_1 S V_1^T = A$.

(9) Z předpokladu je V_1 čtvercová, tedy ortogonální. Pak

$$(A^T A)^{-1} = (V_1 S U_1^T U_1 S V_1^T)^{-1} = (V_1 S^2 V_1^T)^{-1} = V_1 S^{-2} V_1^T,$$

z čehož $(A^T A)^{-1} A^T = V_1 S^{-2} V_1^T V_1 S U_1^T = V_1 S^{-1} U_1^T = A^\dagger$. □

První vlastnost říká, že se skutečně jedná o zobecnění klasické inverze. Vlastnosti (4)–(7) jsou zajímavé v tom, že dávají alternativní definici pseudoinverze; ta se totiž ekvivalentně dá definovat jako matice, která splňuje podmínky (4)–(7), a taková matice kupodivu existuje vždy právě jedna.

Poznamenejme, že některé vlastnosti, u kterých bychom očekávali že platí, tak obecně platit nemusí. Např. obecně $AA^\dagger \neq A^\dagger A$ a $(AB)^\dagger \neq B^\dagger A^\dagger$.

Pomocí pseudoinverze elegantně vyjádříme matice projekce do maticových prostorů.

Věta 13.25. *Buď $A \in \mathbb{R}^{m\times n}$. Pak matice projekce do*

(1) sloupcového prostoru $\mathcal{S}(A)$ je $AA^\dagger$,

[4)] Nezávisle ji objevili americký matematik Eliakim Hastings Moore r. 1920 a anglický fyzik, slavný popularizátor, Roger Penrose r. 1955.

(2) řádkového prostoru $\mathcal{R}(A)$ je $A^\dagger A$,

(3) jádra $\mathrm{Ker}(A)$ *je $I_n - A^\dagger A$.*

Důkaz.

(1) S použitím redukovaného SVD rozkladu $A = U_1 S V_1^T$ upravme

$$AA^\dagger = U_1 S V_1^T V_1 S^{-1} U_1^T = U_1 U_1^T.$$

Podle věty 13.20 je to hledaná matice projekce $U_1 U_1^T$.

(2) Analogicky jako v předchozím je $A^\dagger A = V_1 V_1^T$, což je matice projekce do $\mathcal{R}(A)$.

(3) Plyne z věty 8.44 a vlastnosti $\mathrm{Ker}(A) = \mathcal{R}(A)^\perp$ (věta 8.40). □

Zajímavá je interpretace pseudoinverze z hlediska lineárních zobrazení.

Věta 13.26 (Pseudoinverzní matice a lineární zobrazení)**.** *Uvažujme lineární zobrazení $f(x) = Ax$, kde $A \in \mathbb{R}^{m\times n}$.*

(1) Pokud definiční obor $f(x)$ omezíme pouze na prostor $\mathcal{R}(A)$, tak dostaneme isomorfismus mezi $\mathcal{R}(A)$ a $f(\mathbb{R}^n)$.

(2) Inverzní zobrazení k tomuto isomorfismu má tvar $y \mapsto A^\dagger y$.

Důkaz.

(1) Zobrazení s omezeným definičním oborem je „na“, protože podle důsledku 8.42(2) je

$$\begin{aligned} f(\mathbb{R}^n) &= \mathcal{S}(A) = \mathcal{R}(A^T) = \mathcal{R}(AA^T) = \\ &= \mathcal{S}(AA^T) = \{Ay;\, y \in \mathcal{R}(A)\} = f(\mathcal{R}(A)). \end{aligned}$$

Jelikož prostory $f(\mathbb{R}^n)$ a $\mathcal{R}(A)$ mají stejnou dimenzi (věta 5.62), musí být zobrazení isomorfismem.

(2) Podle věty 13.25(2) se každý vektor $x \in \mathcal{R}(A)$ při zobrazení $x \mapsto A^\dagger Ax$ zobrazí na $A^\dagger Ax = x$. Tudíž $A^\dagger$ je matice inverzního zobrazení k zobrazení $x \mapsto Ax$. □

Nejvýznačnější vlastnost pseudoinverze spočívá v popisu množiny řešení řešitelných soustav a množiny přibližných řešení metodou nejmenších čtverců neřešitelných soustav. V obou případech je $A^\dagger b$ v jistém smyslu význačné řešení.

Věta 13.27 (Pseudoinverzní matice a řešení soustav rovnic)**.** *Buď $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$ a X množina řešení soustavy $Ax = b$. Je-li $X \neq \emptyset$, pak*

$$X = A^\dagger b + \mathrm{Ker}(A).$$

kde

$$\mathrm{Ker}(A) = \mathcal{S}(I_n - A^\dagger A).$$

Navíc, ze všech vektorů z množiny X má $A^\dagger b$ nejmenší eukleidovskou normu, a je to jediné řešení s touto vlastností.

Důkaz. „=“ Buď $x \in X$, tj. $Ax = b$. Potom podle tvrzení 13.24(4) je $AA^\dagger b = AA^\dagger Ax = Ax = b$, tedy $A^\dagger b \in X$. Podle věty 7.5 je $X = x_0 + \mathrm{Ker}(A)$, kde x_0 je libovolné řešení. Podle věty 13.25(3) je $\mathrm{Ker}(A) = \mathcal{S}(I_n - A^\dagger A)$ a za x_0 můžeme volit $A^\dagger b$.

„Norma.“ Buď $x \in X$. Podle věty 13.26(2) je $A^\dagger b = A^\dagger Ax \in \mathcal{R}(A)$ a dále platí $\mathcal{R}(A)^\perp = \mathrm{Ker}(A)$. Nyní podle Pythagorovy věty pro každé $y \in \mathcal{S}(I_n - A^\dagger A)$ platí

$$\|A^\dagger b + y\|_2^2 = \|A^\dagger b\|_2^2 + \|y\|_2^2 \geq \|A^\dagger b\|_2^2.$$

Tedy $A^\dagger b$ nejmenší eukleidovskou normu. Každý jiný vektor z X má normu větší, protože $y \neq 0$ implikuje $\|y\|_2 > 0$. □

Věta 13.28 (Pseudoinverzní matice a metoda nejmenších čtverců)**.** *Buď $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$ a X množina přibližných řešení soustavy $Ax = b$ metodou nejmenších čtverců. Pak*

$$X = A^\dagger b + \operatorname{Ker}(A).$$

Navíc, ze všech vektorů z množiny X má $A^\dagger b$ nejmenší eukleidovskou normu, a je to jediné řešení s touto vlastností.

Důkaz. Množina přibližných řešení soustavy $Ax = b$ metodou nejmenších čtverců je popsána soustavou $A^TAx = A^Tb$ a je neprázdná, viz věta 8.47. Podle věty 13.27 máme

$$X = (A^TA)^\dagger(A^Tb) + \operatorname{Ker}(A^TA).$$

Jelikož podle tvrzení 13.24(8) je $(A^TA)^\dagger A^T = A^\dagger$ a podle důsledku 8.42 je $\operatorname{Ker}(A^TA) = \operatorname{Ker}(A)$, množina X má požadovaný popis a $A^\dagger b$ požadovanou vlastnost. □

Předchozí dvě věty tedy mj. říkají, že $A^\dagger b$ je význačný vektor. V případě, že soustava $Ax = b$ má řešení, pak je jejím řešením s minimální normou. A v případě, že soustava $Ax = b$ nemá řešení, pak je jejím přibližným řešením (opět s minimální normou) metodou nejmenších čtverců. Navíc není zapotřebí předpokladu na lineární nezávislost sloupců matice A.

13.7 Maticová norma

Nyní se vrátíme zpět k normám a podíváme se na to, jak zavést normu pro matice. Přestože normy jsme probírali v sekci 8.1, důležitou maticovou normu představuje největší singulární číslo matice, proto zařazujeme tuto část k SVD rozkladu.

V zásadě, matice z $\mathbb{R}^{m\times n}$ tvoří vektorový prostor, proto na matice můžeme pohlížet jako na vektory. Nicméně, pro maticovou normu se uvažuje ještě jedna vlastnost navíc, proto máme speciální definici.

Definice 13.29 (Norma matice)**.** Třída zobrazení $\|\cdot\|\colon \mathbb{R}^{m\times n} \to \mathbb{R}$ je reálná *maticová norma*, pokud je to norma pro libovolné m, n a navíc splňuje:

$$\|AB\| \le \|A\| \cdot \|B\| \quad \text{pro všechna } A \in \mathbb{R}^{m\times p},\ B \in \mathbb{R}^{p\times n}.$$

První příklad maticové normy, se kterým se seznámíme, je *Frobeniova norma*:

$$\|A\|_F := \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}^2}.$$

Je to vlastně eukleidovská norma vektoru tvořeného všemi prvky matice A a využili jsme ji již v příkladu 8.17. Jinou charakterizaci dává následující věta.

Tvrzení 13.30 (Frobeniova norma)**.** *Buď $A \in \mathbb{R}^{m\times n}$ se singulárními čísly $\sigma_1, \dots, \sigma_r$. Pak $\|A\|_F = \sqrt{\sum_{i=1}^r \sigma_i^2}$.*

Důkaz. Podle tvrzení 10.11 a věty 13.15 máme $\|A\|_F = \sqrt{\sum_{i=1}^m \sum_{j=1}^n a_{ij}^2} = \sqrt{\operatorname{trace}(A^TA)} = \sqrt{\sum_{i=1}^r \sigma_i^2}$. □

Druhý příklad maticové normy je *maticová p-norma*:

$$\|A\|_p = \max_{x:\|x\|_p=1} \|Ax\|_p.$$

V této definici používáme vektorovou p-normu. Výslednou normu si můžeme představit takto: Zobrazíme jednotkovou kouli (v p-normě, tedy vektory splňující $\|x\|_p = 1$) lineárním zobrazením $x \mapsto Ax$ a v obrazu vybereme vektor s největší normou. Pro různé hodnoty p dostáváme různé maticové normy. Ty nejdůležitější shrnuje následující věta.

Tvrzení 13.31 (Maticové p-normy). *Buď $A \in \mathbb{R}^{m\times n}$. Pak maticové p-normy pro $p \in \{1, 2, \infty\}$ mají tvar*

(1) $\|A\|_2 = \sigma_1(A)$ *(největší singulární číslo),*

(2) $\|A\|_1 = \max_{j=1,\dots,n} \sum_{i=1}^m |a_{ij}|$,

(3) $\|A\|_\infty = \max_{i=1,\dots,m} \sum_{j=1}^n |a_{ij}| = \|A^T\|_1$.

Důkaz.

(1) Jak jsme již zmínili, $\|A\|_2$ je velikost největšího bodu elipsy, vzniklé obrazem jednotkové koule při zobrazení $x \mapsto Ax$. Ze sekce 13.5 (SVD a geometrie lineárního zobrazení) víme, že tato hodnota je $\sigma_1(A)$.

(2) Označme $c := \max_{j=1,\dots,n} \sum_{i=1}^m |a_{ij}|$. Připomeňme, že $\|x\|_1 = \sum_{k=1}^n |x_k|$. Pro jakékoli x takové, že $\|x\|_1 = 1$, platí

$$\begin{aligned}\|Ax\|_1 &= \sum_{i=1}^m \left|\sum_{j=1}^n a_{ij}x_j\right| \le \sum_{i=1}^m \sum_{j=1}^n |a_{ij}||x_j| = \\ &= \sum_{j=1}^n |x_j| \left(\sum_{i=1}^m |a_{ij}|\right) \le \sum_{j=1}^n |x_j| c = c.\end{aligned}$$

Zároveň se nabyde rovnost $\|Ax\|_1 = c$ vhodnou volbou jednotkového vektoru $x = e_i$. Proto dostáváme $\max_{x:\|x\|_1=1} \|Ax\|_1 = c$.

(3) Označme $c := \max_{i=1,\dots,m} \sum_{j=1}^n |a_{ij}|$. Připomeňme, že $\|x\|_\infty = \max_{k=1,\dots,n} |x_k|$. Pro jakékoli x takové, že $\|x\|_\infty = 1$, platí

$$\begin{aligned}\|Ax\|_\infty &= \max_{i=1,\dots,m} \left|\sum_{j=1}^n a_{ij}x_j\right| \le \max_{i=1,\dots,m} \sum_{j=1}^n |a_{ij}||x_j| \le \\ &\le \max_{i=1,\dots,m} \sum_{j=1}^n |a_{ij}| = c.\end{aligned}$$

Zároveň se nabyde rovnost $\|Ax\|_\infty = c$ vhodnou volbou vektoru $x \in \{\pm 1\}^n$. Proto dostáváme $\max_{x:\|x\|_\infty=1} \|Ax\|_\infty = c$. □

Poznámka 13.32. Mějme vektor $v \in \mathbb{R}^n$ a libovolné p. Nyní výraz $\|v\|_p$ může označovat jak vektorovou, tak maticovou p-normu. To však nevadí, protože obě dávají stejnou hodnotu. Začneme s maticovou normou a ukážeme, že se rovná vektorové:

$$\|v\|_p = \max_{x\in\mathbb{R}:\|x\|=1} \|vx\|_p = \max_{x\in\{\pm 1\}} \|vx\|_p = \|\pm v\|_p = \|v\|_p.$$

Víme z věty 8.54, že přenásobení vektoru ortogonální maticí nemění jeho eukleidovskou normu. Nyní tvrzení zobecníme pro Frobeniovu a maticovou 2-normu.

Tvrzení 13.33. *Buď $A \in \mathbb{R}^{m\times n}$ a buďte $Q \in \mathbb{R}^{m\times m}$, $R \in \mathbb{R}^{n\times n}$ ortogonální. Pak*

(1) $\|QAR\|_F = \|A\|_F$,

(2) $\|QAR\|_2 = \|A\|_2$.

Důkaz.

(1) Analogicky jako v důkazu tvrzení 13.30 máme

$$\begin{aligned}\|QAR\|_F^2 &= \operatorname{trace}((QAR)^T(QAR)) = \operatorname{trace}(R^TA^TQ^TQAR) = \\ &= \operatorname{trace}(R^TA^TAR) = \operatorname{trace}(A^TARR^T) = \operatorname{trace}(A^TA) = \|A\|_F^2,\end{aligned}$$

kde jsme navíc využili fakt, že $\operatorname{trace}(BC) = \operatorname{trace}(CB)$ pro každé $B, C \in \mathbb{R}^{n\times n}$.

(2) Odvodíme

$$\|QAR\|_2 = \max_{x:\|x\|_2=1} \|QARx\|_2 = \max_{x:\|x\|_2=1} \|ARx\|_2 =$$
$$= \max_{y:\|R^Ty\|_2=1} \|Ay\|_2 = \max_{y:\|y\|_2=1} \|Ay\|_2 = \|A\|_2. \qquad \square$$

Maticové normy se objevují v různých souvislostech. Nejprve ukažme, že dávají odhad na velikost vlastních čísel.

Věta 13.34 (Odhad spektrálního poloměru pomocí normy)**.** *Buď $A \in \mathbb{R}^{n\times n}$. Pak pro každou maticovou normu $\rho(A) \le \|A\|$.*

Důkaz. Buď $\lambda \in \mathbb{C}$ libovolné vlastní číslo a x odpovídající vlastní vektor matice A, tedy $Ax = \lambda x$. Definujme matici $X := (x \mid o \mid \cdots \mid o)$. Protože platí $AX = \lambda X$, můžeme odvodit

$$|\lambda| \cdot \|X\| = \|\lambda X\| = \|AX\| \le \|A\| \cdot \|X\|.$$

Vydělením $\|X\| \ne 0$ dostáváme $|\lambda| \le \|A\|$. $\square$

Příklad 13.35. Uvažujme matici

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \\ 3 & 6 & 9 \end{pmatrix}.$$

Její spektrální poloměr a různé typy norem mají hodnoty:

$$\begin{aligned} \rho(A) &= 12, \\ \|A\|_F &= \sqrt{154} \approx 12.4097, \\ \|A\|_2 &= \sqrt{154} \approx 12.4097, \\ \|A\|_1 &= 15, \\ \|A\|_\infty &= 18. \end{aligned} \qquad \square$$

Další velmi zajímavá vlastnost singulárních čísel je, že σ_i udává v 2-normě vzdálenost matice k nejbližší matici hodnosti nanejvýš $i-1$. V důkazu následující věty je schované i to, jak tuto matici sestrojit – ne náhodou je to matice z low-rank aproximace (sekce 13.5).

Věta 13.36 (Interpretace singulárních čísel)**.** *Buď $A \in \mathbb{R}^{m\times n}$ se singulárními čísly $\sigma_1, \dots, \sigma_r$. Pak*

$$\sigma_i = \min\{\|A-B\|_2;\ B \in \mathbb{R}^{m\times n},\ \operatorname{rank}(B) \le i-1\}$$

pro každé $i = 1, \dots, r$.

Důkaz. Nerovnost „$\ge$“. Nechť $A = U\Sigma V^T$ je SVD rozklad matice A. Definujme matici předpisem $B := U \operatorname{diag}(\sigma_1, \dots, \sigma_{i-1}, 0, \dots, 0)V^T$. Pak

$$\|A-B\|_2 = \|U \operatorname{diag}(0, \dots, 0, \sigma_i, \dots, \sigma_n)V^T\|_2 = \|\operatorname{diag}(0, \dots, 0, \sigma_i, \dots, \sigma_n)\|_2 = \sigma_i.$$

Nerovnost „$\le$“. Buď $B \in \mathbb{R}^{n\times n}$ libovolná matice hodnosti nanejvýš $i-1$ a ukážeme, že $\|A-B\|_2 \ge \sigma_i$. Nechť V_1 sestává z prvních i sloupců matice V. Buď $o \ne z \in \operatorname{Ker}(B) \cap \mathcal{S}(V_1)$, to jest $Bz = o$, a navíc normujeme z tak, aby $\|z\|_2 = 1$. Takový vektor existuje, protože $\dim \operatorname{Ker}(B) \ge n-i+1$ a $\dim \mathcal{S}(V_1) = i$. Pak

$$\|A-B\|_2^2 = \max_{x:\|x\|_2=1} \|(A-B)x\|_2^2 \ge \|(A-B)z\|_2^2 = \|Az\|_2^2 = \|U\Sigma V^T z\|_2^2.$$

Protože $z \in \mathcal{S}(V_1)$, lze psát $z = Vy$ pro nějaký vektor $y = (y_1, \dots, y_i, 0, \dots, 0)^T$, přičemž $\|y\|_2 = \|V^T z\|_2 = \|z\|_2 = 1$. Nyní

$$\|U\Sigma V^T z\|_2^2 = \|U\Sigma V^T V y\|_2^2 = \|\Sigma y\|_2^2 = \sum_{j=1}^{i} \sigma_j^2 y_j^2 \ge \sum_{j=1}^{i} \sigma_i^2 y_j^2 = \sigma_i^2 \|y\|_2^2 = \sigma_i^2. \qquad \square$$

Speciálně, nejmenší singulární číslo σ_n matice $A \in \mathbb{R}^{n\times n}$ udává vzdálenost k nejbližší singulární matici, srov. sekce 13.5 (SVD a míra regularity). To znamená, že matice $A + C$ je regulární pro všechny matice $C \in \mathbb{R}^{n\times n}$ splňující $\|C\|_2 < \sigma_n$.

Příklad 13.37. Uvažujme matici

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

Její nejmenší singulární číslo je $\sigma_4 = 2$ (ve skutečnosti jsou všechna singulární čísla stejně velká, protože A je dvojnásobkem ortogonální matice). Tudíž matice zůstane regulární i když k ní přičteme libovolnou matici s 2-normou menší než 2. □

Další zajímavé souvislosti viz Prasolov [1994].

Problémy

13.1. Určete maximální počet aritmetických operací algoritmu 13.6 na QR rozklad a porovnejte jej s počtem operací Gaussovy resp. Gaussovy–Jordanovy eliminace (srov. problém 2.2).

13.2. Navrhněte metodu na QR rozklad za použití Givensových matic.

13.3. Za použití QR rozkladu popište všechna řešení soustavy $Ax = b$, kde A má lineárně nezávislé řádky.

13.4. Ukažte, že Frobeniova a p-norma jsou skutečně maticovými normami.

13.5. Buď $A \in \mathbb{R}^{n\times n}$. Dokažte $\|A\|_2 \le \|A\|_F \le \sqrt{n}\|A\|_2$.

13.6. Buď $A \in \mathbb{R}^{n\times n}$. Dokažte, že ze všech symetrických matic je matice $\frac{1}{2}(A + A^T)$ nejblíže k matici A ve Frobeniově a maticové 2-normě.

Shrnutí ke kapitole 13. Maticové rozklady

Maticové rozklady jsou velmi účinný nástroj vědeckého počítání v moderní době. Rozkladů existuje celá řada, mezi nejdůležitější se řadí QR a SVD rozklad. Ne náhodou oba používají v rozkladu ortogonální matice.

QR rozklad vyjadřuje libovolnou reálnou matici jako součin ortogonální a horní trojúhelníkové. Tento rozklad lze výpočetně jednoduše získat pomocí Gramovy–Schmidtovy ortogonalizace či Householderovou transformací. Využití QR rozkladu je nepřeberné: řešení soustav lineárních rovnic, nalezení ortonormální báze, sestrojení matice ortogonální projekce, řešení metodou nejmenších čtverců, metoda na výpočet vlastních čísel,...

SVD rozklad má analogické vlastnosti. Danou reálnou matici rozkládá na součin ortogonální, diagonální (ale ne nutně čtvercovou!) a ortogonální. Využití je podobné jako pro QR rozklad, ale navíc nám říká něco o geometrii lineárních zobrazení, dává nástroj pro aproximaci a kompresi dat. Umožňuje také přirozeně rozšířit pojem inverzní matice na ne nutně regulární matice.

Singulární čísla (= čísla diagonální matice z SVD rozkladu) pak poskytují zajímavé informace o lineárním zobrazení $x \mapsto Ax$ a o matici A samotné a také o datech, která reprezentuje. Singulární čísla říkají, jak moc lineární zobrazení degeneruje objekty, jaké má matice numerické vlastnosti, jaká je vzdálenost k nejbližší singulární matici a největší singulární číslo reprezentuje často používanou maticovou normu.

Kapitola 14

Závěrem

K sepsání tohoto závěru mě inspiroval kolega, který na zkoušce z lineární algebry položil studentovi otázku: „A k čemu to vlastně je?“ Tato otázka samozřejmě nemá jednoznačnou odpověď a lidé se liší podle priorit, které kladou různým pohledům. Kromě základních znalostí, které jsou potřeba pro další studium informatiky a matematiky, kromě osvojení si práce s matematickými pojmy jako je definice-věta-důkaz a matematické vyjadřování vůbec, kromě tohoto všeho je dobré v lineární algebře hledat i nějaké hlubší myšlenky a souvislosti. Zkusím zde zdůraznit několik z nich:

- *Algebraický vs. geometrický pohled.*

  Na většinu pojmů lineární algebry lze nahlížet algebraicky nebo geometricky. Ani jeden přístup není silnější než druhý, oba mají stejnou dokazovací schopnost. Přesto je dobré se na věci dívat z více úhlů pohledu, protože nám to pomůže porozumět podstatě věci.

  Pro ilustraci: Na soustavy lineárních rovnic se můžeme dívat algebraicky, kdy hledáme vektor splňující dané rovnice, nebo se na ně můžeme dívat geometricky jako na nadroviny a řešení je jejich průsečík. Jiný příklad: interpretaci determinantu jako objem rovnoběžnostěnu jsme dokázali algebraicky ve větě 9.22, ale stejně tak lze myšlenku nahlédnout geometricky, jak jsme učinili v poznámce 9.23. Další příklad: Vlastní čísla a vektory můžeme vyjadřovat algebraicky pomocí rovnice z definice, nebo geometricky jako invariantní směry. V tomto rozlišení je relace podobnosti buďto algebraická úprava vynásobení maticí a její inverzí, anebo změna souřadného systému (báze), ve kterém zkoumáme lineární zobrazení. Poslední příklad: Kvadratické formy a Sylvestrův zákon setrvačnosti. Algebraický pohled je maticová transformace $A \to S^T AS$, kdežto geometrická je opět jako změna báze kvadratické formy.

- *Invarianty a základní tvary.*

  Invariantem rozumím operaci, která nemění nějakou základní vlastnost, o kterou nám jde. Je to důležitý princip vyskytující se i v jiných oblastech matematiky. Například v souvislosti s řešením soustav lineárních rovnic jsou to elementární řádkové úpravy, které mění matici, ale nemění množinu řešení. Lineární zobrazení se nezmění, když změníme bázi, změní se pouze matice, která ho reprezentuje. Pro determinant to byly také elementární řádkové či sloupcové úpravy. Ale pro vlastní čísla už jsme museli hledat jinou úpravu, a byla jí podobnost. A pro kvadratické formy to byla kongruence.

  S invarianty souvisí základní tvary matic, tedy v jistém smyslu nejjednodušší tvary, ke kterým se danou operací můžeme dostat: RREF pro elementární řádkové úpravy, Jordanova normální forma pro podobnost a diagonální matice pro kongruenci.

- *Axiomatizace.*

  Některé pojmy jsme zavedli konstruktivně, například determinanty, ale zvláštní způsob definice je pomocí axiomů, tedy vlastností, které má daný objekt splňovat. Takto jsme definovali pojem grupy, tělesa či vektorového prostoru. Přestože je tento způsob hodně abstraktní, je v matematice často využíván právě pro obecnost definovaných pojmů. Vektorem tedy není nějaký směr v prostoru, ale v různých souvislostech to může být i polynom, matice, nebo třeba funkce.

Problémy

14.1. Jaké souvislosti vidíte v lineární algebře Vy?

Značení

$U + V$	spojení podprostorů, $U + V = \{u + v;\ u \in U, v \in V\}$, str. 73
$U + a$	součet množiny vektorů a vektoru, $U + a = \{v + a;\ v \in U\}$, str. 97
$[v]_B$	souřadnice vektoru v vzhledem k bázi B, str. 70
$\exists$	existenční kvantifikátor („existuje“), str. 9
$\forall$	univerzální kvantifikátor („pro všechna“), str. 9
${}_{B_2}[f]_{B_1}$	matice lineárního zobrazení f vzhledem k bázi B_1, B_2, str. 87
$\|A\|$	norma matice A, str. 196
$\|x\|$	norma vektoru x, str. 109
$\|x\|_2$	eukleidovská norma, $\|x\|_2 = \sqrt{\sum_{i=1}^n x_i^2}$, str. 110
$\Subset$	podprostor, str. 64
$\prod$	produkt (součin), str. 9
$g \circ f$	složené zobrazení, $(g \circ f)(x) = g(f(x))$, str. 10
$\langle x, y \rangle$	skalární součin vektorů x, y, str. 107
$\sum$	suma (součet), str. 9
$H(a)$	Householderova matice vytvořená z a, str. 123
$M^\perp$	ortogonální doplněk množiny M, str. 115
$\operatorname{adj}(A)$	adjungovaná matice k A, str. 134
A_{i*}	i-tý řádek matice A, str. 18
A_{*j}	j-tý sloupec matice A, str. 18
$A^\dagger$	pseudoinverze matice A, str. 194
A^*	Hermitovská transpozice matice, $A^* = \overline{A}^T$, str. 156
A^T	transpozice matice, str. 31
A^{-T}	inverze a transpozice matice, $A^{-T} = (A^T)^{-1} = (A^{-1})^T$, str. 39
$\mathbb{C}$	množina komplexních čísel, str. 11
$C(p)$	matice společnice polynomu $p(x)$, str. 145
$\det(A)$	determinant matice A, str. 129
$\operatorname{diag}(v)$	diagonální matice s diagonálními prvky $v_1, \ldots, v_n$, str. 32
$\dim V$	dimenze prostoru V, str. 72
e_i	jednotkový vektor, str. 31
$E_i(\alpha)$	elementární matice pro vynásobení i-tého řádku číslem $\alpha \neq 0$, str. 37
E_{ij}	elementární matice pro prohození dvou řádků i, j, str. 37
$E_{ij}(\alpha)$	elementární matice pro přičtení α-násobku j-tého řádku k i-tému, str. 37
$f(U)$	obraz množiny U při zobrazení f, str. 84
$\mathsf{Im}(z)$	imaginární část komplexního čísla z, str. 11
I_n, I	jednotková matice, str. 31
$J_k(\lambda)$	Jordanova buňka, str. 152
kan	kanonická báze, str. 70

$\mathrm{Ker}(A)$	jádro matice A, str. 75
$\mathrm{Ker}(f)$	jádro zobrazení f, str. 84
$\mathbb{N}$	množina přirozených čísel
$n!$	faktoriál, $n! = 1 \cdot 2 \cdot 3 \cdot \ldots \cdot n$, str. 9
o	nulový vektor, str. 63
$p_A(\lambda)$	charakteristický polynom matice A, str. 142
$\mathcal{P}^n$	prostor reálných polynomů proměnné x stupně nanejvýš n, str. 63
$\mathbb{Q}$	množina racionálních čísel
$\mathbb{R}$	množina reálných čísel
$\mathrm{rank}(A)$	hodnost matice A, str. 20
$\mathsf{Re}(z)$	reálná část komplexního čísla z, str. 11
REF	odstupňovaný tvar matice, str. 20
$\rho(A)$	spektrální poloměr matice A, str. 143
RREF	redukovaný odstupňovaný tvar matice, str. 23
$\mathcal{R}(A)$	řádkový prostor matice A, str. 75
$\mathcal{S}(A)$	sloupcový prostor matice A, str. 75
$\mathrm{sgn}(p)$	znaménko permutace p, str. 53
S_n	množina všech permutací na $\{1, \ldots, n\}$, str. 53
$\mathrm{span}(W)$	lineární obal množiny vektorů W, str. 65
$\mathbb{T}$	nějaké těleso, str. 56
$\mathrm{trace}(A)$	stopa matice, $\mathrm{trace}(A) = \sum_{i=1}^{n} a_{ii}$, str. 125
$\overline{z}$	komplexně sdružené číslo, str. 11
$\mathbb{Z}$	množina celých čísel
$\mathbb{Z}_n$	množina $\{0, 1, \ldots, n-1\}$, str. 57

Literatura

S. Axler. Down with determinants! *American Mathematical Monthly*, 102(2):139–154, 1995. `https://www.jstor.org/stable/pdf/2975348.pdf`.

J. Bečvář. *Lineární algebra.* Matfyzpress, Praha, třetí vydání, 2005.

M. Berger. *Geometry I.* Springer, Berlin, 1987.

L. Bican. *Lineární algebra a geometrie.* Academia, Praha, druhé vydání, 2009.

N. Carter. *Visual Group Theory.* The Mathematical Association of America, Washington, DC, 2009.

T. Chartier. *When Life is Linear: From Computer Graphics to Bracketology.* The Mathematical Association of America, Washington, DC, 2015.

B. A. Cipra. The best of the 20th century: Editors name top 10 algorithms. *SIAM News*, 33:1–2, 2000. `https://www.siam.org/pdf/news/637.pdf`.

J. Dattorro. *Convex Optimization & Euclidean Distance Geometry.* Meboo Publishing, USA, 2011.

J. Dongarra and F. Sullivan. Guest editors' introduction: The top 10 algorithms. *Computing in Science and Engineering*, 2:22–23, 2000. `https://www.computer.org/csdl/mags/cs/2000/01/c1022.pdf`.

W. Gareth. *Linear Algebra with Applications.* Jones and Bartlett Publishers, Boston, 4th edition, 2001.

B. Gärtner and J. Matoušek. *Approximation Algorithms and Semidefinite Programming.* Springer, Berlin Heidelberg, 2012.

K. Goto and R. A. van de Geijn. Anatomy of high-performance matrix multiplication. *ACM Trans. Math. Softw.*, 34(3):12:1–12:25, 2008. `http://doi.acm.org/10.1145/1356052.1356053`.

R. A. Horn and C. R. Johnson. *Matrix Analysis.* Cambridge University Press, Cambridge, 1985.

T. Krisl. Cauchyova–Schwarzova nerovnost. Diplomová práce, Masarykova Univerzita, Přírodovědecká fakulta, Ústav matematiky a statistiky, Brno, 2008. `http://www.is.muni.cz/th/106635/prif_m/diplomka.pdf`

A. N. Langville and C. D. Meyer. *Google's PageRank and Beyond. The Science of Search Engine Rankings.* Princeton University Press, Princeton, 2006.

J. Matoušek. Informace k přednáškám a cvičením, 2010. `http://kam.mff.cuni.cz/~matousek/vyuka.html`

J. Matoušek a J. Nešetřil. *Kapitoly z diskrétní matematiky.* Karolinum, Praha, čtvrté vydání, 2009.

C. D. Meyer. *Matrix Analysis and Applied Linear Algebra (incl. CD-ROM and solutions manual).* SIAM, Philadelphia, PA, 2000. `http://www.matrixanalysis.com/DownloadChapters.html`.

V. V. Prasolov. *Problems and Theorems in Linear Algebra.* American Mathematical Society, 1994.
`http://www2.math.su.se/~mleites/books/prasolov-1994-problems.pdf`.

H. Ratschek and J. Rokne. *Geometric Computations with Interval and New Robust Methods. Applications in Computer Graphics, GIS and Computational Geometry.* Horwood Publishing, Chichester, 2003.

J. Rohn. Lineární a nelineární programování (Učební text), 1997.
`http://kam.mff.cuni.cz/~hladik/doc/Rohn-LP_NLP-1997.pdf`

J. Rohn. Přehled některých důležitých vět z teorie matic, technická zpráva V-895, ICS AS CR, Praha, 2003.
`http://hdl.handle.net/11104/0125288`

J. Rohn. *Lineární algebra a optimalizace.* Karolinum, Praha, první vydání, 2004.

G. Strang. *Linear Algebra and Its Applications.* Thomson, USA, 3rd edition, 1988.

J. Šíma a R. Neruda. *Teoretické otázky neuronových sítí.* Matfyzpress, Praha, 1996.
`http://www2.cs.cas.cz/~sima/kniha.html`

J. Tůma. Texty k přednášce Lineární algebra, 2003.
`http://www.karlin.mff.cuni.cz/~tuma/NNlinalg.htm`

M. Turk and A. Pentland. Eigenfaces for recognition. *J. Cognitive Neuroscience*, 3(1):71–86, 1991.
`http:/www.face-rec.org/algorithms/PCA/jcn.pdf`.

K. Výborný a M. Zahradník. *Používáme lineární algebru.* Karolinum, Praha, první vydání, 2002.
`http://matematika.cuni.cz/zahradnik-plas.html`