Drsná matematika I – 10
Maticový počet nad komplexními čísly

Jan Slovák

Masarykova univerzita

18–23.11. 2019

Obsah přednášky

1. Unitární prostory
2. Spektrální rozklad
3. Jordanův rozklad

Duální prostory, skalární součin

Množina všech lineárních forem na daném prostoru V je opět vektorový prostor, značíme jej V^*. Pokud je V konečněrozměrný, je V^* izomorfní prostoru V. Realizace takového izomorfismu je dána např. volbou tzv. **duální báze** k zvolené bázi na V, jejímiž prvky α_i jsou právě formy zadávající i-tou souřadnici.
Sesquilineární formy jsou takové formy α na komplexním vektorovém prostoru, které jsou lineární nad $\mathbb{R}$, zatímco pro komplexní a platí $\alpha(av) = \bar{a}\alpha(v)$.

Definition (Unitární vektorový prostor)

Skalární součin na (komplexním) vektorovém prostoru V je forma $\langle\ ,\ \rangle : V \times V \to \mathbb{R}$, která je lineární v prvním argumentu, $\langle u, v\rangle = \overline{\langle v, u\rangle}$, $\langle v, v\rangle \geq 0$ a nulová hodnota nastává pouze při $v = 0$.

Pro skalární součin se často používá také obvyklé tečky, tj. $\langle u, v\rangle = u \cdot v$.

Ortonormální báze

Definition

Vektory v a $w \in V$ se nazývají **ortogonální**, jestliže $\langle v, w \rangle = 0$. Vektor v se nazývá **normovaný**, jestliže $\|v\| = 1$. Báze prostoru V složená z ortogonálních vektorů se nazývá **ortogonální báze**. Jsou-li bázové vektory navíc i normované, je to **ortonormální báze**.

Theorem

Skalární součin je v každé ortonormální bázi dán výrazem

$$\langle x, y \rangle = x^T \cdot \bar{y}.$$

V obecné bázi V existuje symetrická matice S taková, že

$$\langle x, y \rangle = x^T \cdot S \cdot \bar{y}.$$

Grammova-Schmidtova ortogonalizace vždy dá ortonormální báze.

Duální prostory a zobrazení

Pevně zvolený vektor $v \in V$ dosazený za druhý argument ve skalárním součinu dá zobrazení $V \to V^* = \mathrm{Hom}(V, \mathbb{R})$

$$V \ni v \mapsto (w \mapsto \langle v, w \rangle \in \mathbb{R}).$$

Nedegenerovanosti skalárního součinu zaručuje, že je to bijekce. Je vidět, že vektory ortonormální báze jsou zobrazeny na formy tvořící bázi duální.

Lieární $f : V \to W$ zadává tzv. duální zobrazení $f^* : W^* \to V^*$ mezi formami, definované pro všechny $w^* \in W^*$, $v \in V$

$$f^*(w^*)(v) = w^*(f(v)).$$

V libovolných bazích na V a W a jejich duálních bazích na V^* a W^* pak tentýž definiční vztah má tvar (píšeme $\tilde{A}$ pro matici zobrazení f^*, x^T souřadnice formy w^*, y souřadnice v)

$$(\tilde{A} x^T) \cdot y = x^T \cdot (A \cdot y)$$

a proto duální zobrazení má v duálních bazích matici A^T.

Adjungovaná a samoadjungovaná zobrazení

V případě vektorových prostorů se skalárním součinem (reálné nebo komplexní), převádí výše uvedené bijekce duální zobrazení f^* na tobrazení $f^* : W \to V$ zadané vztahem

$$\langle f(u), v\rangle = \langle u, f^*(v)\rangle$$

a tomuto zobrazení se říká **adjungované zobrazení** k f. Předchozí výpočet v souřadnicích pro symetrická zobrazení nám ve skutečnosti sdělil, že je-li A matice zobrazení f v ortonormální bázi, pak matice adjungovaného zobrazení f^* je matice konjugovaná a transponovaná $A^* := \bar{A}^T$.

Zobrazením $f : V \to V$, která jsou rovna svým adjungovaným zobrazením f^*, se říká **samoadjungovaná**.

Lineární zobrazení $f : V \to V$ se nazývá **projekce**, jestliže platí

$$f \circ f = f.$$

V takovém případě je pro každý vektor $v \in V$

$$v = f(v) + (v - f(v)) \in \mathrm{Im}(f) + \mathrm{Ker}(f) = V$$

a je-li $v \in \mathrm{Im}(f)$ a $f(v) = 0$, pak je i $v = 0$. Je tedy přechozí součet podprostorů přímý. Říkáme, že f je projekce na podprostor $W = \mathrm{Im}(f)$ podél podprostoru $U = \mathrm{Ker}(f)$. Slovy se dá projekce popsat přirozeně takto: rozložíme daný vektor na komponentu ve W a v U a tu druhou zapomeneme.

Předpokládejme nyní, že na V je definován skalární součin. Pro každý pevně zvolený podprostor $W \subset V$ definujeme jeho **ortogonální doplněk**

$$W^\perp = \{u \in V;\ \langle u, v\rangle = 0 \text{ pro všechny } v \in W\}.$$

Přímo z definice je zjevné, že $W^\perp$ je vektorový podprostor. Jestliže $W \subset V$ má bázi $(u_1, \dots, u_k)$ je podmínka pro $W^\perp$ dána jako k homogenních rovnic pro n proměnných. Bude tedy mít $W^\perp$ dimenzi alespoň $n - k$. Zároveň ale $u \in W \cap W^\perp$ znamená $\langle u, u\rangle = 0$ a tedy i $u = 0$ podle definice skalárního součinu. Zřejmě je tedy vždy

$$V = W \oplus W^\perp.$$

Každý podprostor $W \neq V$ definuje **kolmou projekci** na W. Je to projekce na W podél $W^\perp$.

Pokud máme zadán podprostor $W \subset V$ a jeho ortonormální bázi $(e_1, \dots, e_k)$, jde ji jistě doplnit na ortonormální bázi $(e_1, \dots, e_n)$ celého V. Kolmá projekce obecného vektoru $v \in V$ do W pak bude dána vztahem

$$v \mapsto \langle e_1, v \rangle e_1 + \cdots + \langle e_n, v \rangle e_k.$$

Pro kolmou projekci nám tedy stačí znát jen ortonormání bázi podprostoru W, na nejž promítáme.
Povšimněme si také, že obecně jsou projekce f na podprostor W podél U a projekce g na U podél W svázány vztahem $g = \mathrm{id}_V - f$. Je tedy u kolmých projekcí na daný podprostor W vždy výhodnější počítat ortonormální bázi toho z dvojice W, $W^\perp$, který má menší dimenzi.

Ortogonální a unitární zobrazení

Zobrazení $f : V \to W$, které zachovává velikosti pro všechny vektory $u \in V$, se nazývá **ortogonální zobrazení**. Požadujeme tedy

$$\langle f(u), f(u)\rangle = \langle u, u\rangle.$$

Z linearity f a symetrie skalárního součinu plyne

$$\langle f(u+v), f(u+v)\rangle = \langle f(u), f(u)\rangle + \langle f(v), f(v)\rangle + 2\langle f(u), f(v)\rangle,$$

je tedy ekvivalentní podmínkou i zdánlivě silnější požadavek, aby

$$\langle f(u), f(v)\rangle = \langle u, v\rangle,$$

pro všechny vektory $u, v \in V$.

Obecně, ortogonální zobrazení musí vždycky být injektivní, protože podmínka
$$\langle f(u), f(u)\rangle = 0$$
znamená i $\langle u, u\rangle = 0$ a tedy $u = 0$. Je tedy vždy v takovém případě dimenze oboru hodnot alespoň taková, jako je dimenze definičního oboru f.
Bez újmy na obecnost proto můžeme rovnou předpokládat, že jsou stejné a $f : V \to V$ (pokud by nebyly, doplníme ortonormální bázi na oboru hodnot).
Matice ortogonálního zobrazení v ortonormální bázi splňuje pro všechny vektory x a y v prostoru $\mathbb{K}^n$:
$$(A \cdot x)^T \cdot (\bar{A} \cdot \bar{y}) = x^T \cdot (A^T \cdot \bar{A}) \cdot \bar{y} = x^T \cdot \bar{y}.$$
Speciálními volbami vektorů standardní báze za x a y dostaneme přímo, že $\bar{A}^T \cdot A = E$, tedy tentýž výsledek jako v reálné dimenzi 2!

Dokázali jsme tak následující tvrzení:

Theorem

Nechť V je reálný nebo komplexní vektorový prostor se skalárním součinem a $f : V \to V$ je lineární zobrazení. Pak f je ortogonální právě, když v některé ortonormální bázi (a pak už všech) má matici A splňující $\bar{A}^T = A^{-1}$.

Skutečně, jestliže zachovává f velikosti, musí mít uvedenou vlastnost v každé ortonormální bázi. Naopak, předchozí výpočet ukazuje, že vlastnost matice v jedné bázi už zaručuje zachovávání velikostí.

Důsledkem této věty je také popis všech matic přechodu S mezi ortonormálními bázemi. Každá totiž musí zadávat zobrazení $\mathbb{K}^n \to \mathbb{K}^n$ zachovávající velikosti a splňují tady také právě podmínku $S^{-1} = \bar{S}^T$.

Při přechodu od jedné ortonormální báze ke druhé se tedy matice ortogonálního/unitárního zobrazení mění podle vztahu (v reálném případě je $S^* = S^T$)

$$A' = S^* A S.$$

Invariantní podprostory

Nechť $f : V \to V$ je lineární a předpokládejme, že pro nějaký podprostor $W \subset V$ platí $f(W) \subset W$. Říkáme, že W je **invariantní podprostor** pro zobrazení f. Jestliže je V konečněrozměrné a vybereme nějakou bázi $(u_1, \dots, u_k)$ podprostoru W, můžeme ji vždy doplnit na bázi $(u_1, \dots, u_n)$ celého V a v každé takové bázi má naše zobrazení matici A tvaru

$$A = \begin{pmatrix} B & C \\ 0 & D \end{pmatrix} \tag{1}$$

kde B je čtvercová matice dimenze k, D je čtvercová matice dimenze $n - k$ a C je matice typu $n/(n - k)$. Naopak, jestliže existuje v nějaké bázi matice zobrazení f tvaru (1), je $W = \langle u_1, \dots, u_k \rangle$ invariantní podprostor zobrazení f.

Extrémní případy jsme viděli při hledání báze z vlastních vektorů. V případě existence n různých vlastních čísel zobrazení f jsme dostali rozklad V na přímý součet n vlastních podprostorů a v bazích z vlastních vektorů má naše zobrazení diagonální tvar s vlastními čísly na diagonále.
Zároveň jsme viděli dva různé příklady důvodů, proč zobrazení diagonální matici mít nemusí. První souvisel s nilpotentními zobrazeními, druhý s rotacemi v dvourozměrných podprostorech.

Rozklad ortogonálního/unitárního zobrazení

Nechť je zobrazení $f : V \to V$ ortogonální, s maticí A v nějaké ortonormální bázi.

Jestliže pro libovolný podprostor $W \subset V$ a ortogonální zobrazení $f : V \to V$ platí $f(W) \subset W$, pak také platí pro všechny $v \in W^\perp$, $w \in W$

$$\langle f(v), w \rangle = \langle f(v), f \circ f^{-1}(w) \rangle = \langle v, f^{-1}(w) \rangle = 0$$

protože i $f^{-1}(w) \in W$. To ale znamená, že také $f(W^\perp) \subset W^\perp$.

Theorem

Ortogonální doplněk k invariantnímu podprostoru je také invariantní.

Pro unitární prostory už toto tvrzení zaručuje existenci báze z vlastních vektorů! Skutečně, zúžení unitárního zobrazení f na ortogonální doplněk invariantního podprostoru je opět unitární zobrazení, takže můžeme do báze přibírat jeden vlastní vektor za druhým, až dostaneme celý rozklad V.
Kdyby byla vlastní čísla ortogonálního, dostali bychom stejný výsledek také.
Nicméně většinou nejsou vlastní čísla ortogonálních zorbazení reálná. Musíme si proto pomoci opět výletem do komplexních vektorových prostorů.

Theorem

Nechť $f : V \to V$ je ortogonální zobrazení na prostoru se skalárním součinem. Pak všechny kořeny charakteristického polynomu f mají velikost jedna a existuje rozklad V na jednorozměrné vlastní podprostory odpovídající vlastním číslům $\lambda = \pm 1$ a dvourozměrné podprostory $P_{\lambda,\bar{\lambda}}$, na kterých působí f rotací o úhel rovný argumentu komplexního čísla λ. Všchny tyto různé podprostory jsou po dvou ortogonální.

Náznak důkazu

Jestliže považujeme matici A za matici lineárního zobrazení na komplexním prostoru $\mathbb{C}^n$ (která je jen shodou okolností reálná), budeme mít právě n kořenů charakteristického polynomu, včetně jejich algebraické násobnosti. Navíc, protože charakteristický polynom zobrazení bude mít výhradně reálné koeficienty, budou tyto kořeny buď reálné, nebo půjde o dvojice komplexně sdružených kořenů λ a $\bar{\lambda}$. Příslušné vlastní vektory v $\mathbb{C}^n$ k takové dvojici vektorů budou také komplexně sdružené, protože budou řešením dvou komplexně sdružených systémů lineárních rovnic.

Označme v_λ vlastní vektor příslušný k vlastnímu číslu $\lambda = \alpha + i\beta$, $\beta \neq 0$. Reálný vektorový podprostor P_λ generovaný reálnou a imaginární částí $x_\lambda = \operatorname{re} v_\lambda$, $y_\lambda = \operatorname{im} v_\lambda$ je zjevně invariantní vůči násobení maticí A a dostáváme

$$A \cdot x_\lambda = \alpha x_\lambda - \beta y_\lambda,\ A \cdot y_\lambda = \alpha y_\lambda + \beta x_\lambda.$$

To ale neznamená nic jiného, než že zúžení našeho zobrazení na P_λ je dáno složením rotace o argument vlastní honoty λ (úhel $\arccos \frac{\alpha}{\sqrt{\alpha^2+\beta^2}}$) s násobením velikostí vlastní hodnoty λ (skalárem $\sqrt{\alpha^2 + \beta^2}$). Protože naše zobrazení zachovává velikosti, musí být velikost vlastní hodnoty λ rovna jedné.

Example

V dimenzi tři má charakteristický polynom alespoň jeden reálný kořen, kterým musí být buď jednička nebo mínus jednička. Další dva musí být opět ± 1 nebo dva komplexně sdružené nereálné. V posledním případě zadává vlastní vektor odpovídající reálnému vlastnímu číslu osu rotace o argument vlastního čísla druhého. Pokud je reálné vlastní číslo -1, bude navíc ještě uplatněno zrcadlení podle roviny rotace.

Uvažme tedy zobrazení s maticí ve standardní bázi

$$f : \mathbb{R}^3 \to \mathbb{R}^3,\ A = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{pmatrix}.$$

Dostaneme polynom $-\lambda^3 + \lambda^2 - \lambda + 1 = -(\lambda - 1)(\lambda^2 + 1)$ s kořeny $\lambda_1 = 1$, $\lambda = i$ a $\bar{\lambda} = -i$. Pochopitelně matice zadává rotaci o devadesát stupnů podle osy y.

Theorem (Spektrální rozklad symetrického zobrazení)

Pro každé symetrické/samoadjungované zobrazení $f : V \to V$ na vektorovém prostoru se skalárním součinem existuje ortonormální báze z vlastních vektorů. Jsou-li $\lambda_1, \dots, \lambda_k$ všechna různá vlastní čísla f a $P_1, \dots, P_k$ příslušné kolmé a navzájem kolmé projektory na vlastní podprostory, pak

$$f = \lambda_1 P_1 + \dots + \lambda_k P_k.$$

Vlastní čísla jsou přitom i v unitárním případě vždy reálná.

Zde o projektoru $P : V \to V$ říkáme, že je **kolmý**, je-li $\operatorname{Im} P \perp \operatorname{Ker} P$. Dva kolmé projektory P, Q jsou **vzájemně kolmé**, je-li $\operatorname{Im} P \perp \mathbf{Im} Q$.

Věta opět vyplývá ze zkoumání invariance ortogonálních doplňků a vlastních čísel.

Jestliže pro libovolný podprostor $W \subset V$ a symetrické zobrazení $f : V \to V$ platí $f(W) \subset W$, pak také platí pro všechny $v \in W^\perp$, $w \in W$

$$\langle f(v), w\rangle = \langle v, f(w)\rangle = 0.$$

To ale znamená, že také $f(W^\perp) \subset W^\perp$.
Představme si dále, že A je matice symetrického zobrazení a $A \cdot x = \lambda x$ pro nějaký komplexní vektor $x \in \mathbb{C}^n$. Rozšíříme si definici skalárního součinu $\langle\ ,\ \rangle$ na $\mathbb{C}^n$ vztahem

$$\langle x, y\rangle = x^T \cdot \bar{y}$$

kde $\bar{y}$ je vektor v $\mathbb{C}^n$ s komplexně konjugovanými souřadnicemi. Zjevně platí i pro rozšířené zobrazení $x \mapsto A \cdot x$ vztah

$$\langle A \cdot x, y\rangle = \langle x, A \cdot y\rangle$$

a pro náš vlastní vektor x tedy dostáváme

$$\lambda\langle x, x\rangle = \bar{\lambda}\langle x, x\rangle.$$

Kladným reálným číslem $\langle x, x\rangle$ můžeme krátit a proto musí být $\bar{\lambda} = \lambda$, tj. vlastní čísla jsou skutečně reálná.

Komplexních kořenů má charakteristický polynom $\det(A - \lambda E)$ tolik, kolik je dimenze čtvercové matice A, a všechny jsou ve skutečnosti reálné.

Lemma

Ortogonální doplněk k invariantnímu podprostoru pro symetrické zobrazení je také invariantní. Navíc jsou všechna vlastní čísla symetrické matice A reálná.

Ze samotné definice je zřejmé, že zúžení symetrického zobrazení na invariantní podprostor je opět symetrické. Předchozí tvrzení nám tedy zaručuje, že bude vždy existovat báze V z vlastních vektorů. Skutečně, zúžení f na ortogonální doplněk invariantního podprostoru je opět ortogonální zobrazení, takže můžeme do báze přibírat jeden vlastní vektor za druhým, až dostaneme celý rozklad V. Vlastní vektory příslušející různým vlastním číslům jsou navíc kolmé, protože z rovností $f(u) = \lambda u$, $f(v) = \mu v$ vyplývá

$$\lambda\langle u, v\rangle = \langle f(u), v\rangle = \langle u, f(v)\rangle = \mu\langle u, v\rangle.$$

Nezáporná zobrazení a odmocniny

Nezáporná reálná čísla jsou právě ta, která umíme psát jako druhé mocniny. Zobecnění takového chování pro matice a zobrazení lze vidět u součinů $B = A^T \cdot A$ (tj. složení zobrazení $f^* \circ f$):

$$\langle B \cdot x, x \rangle = \langle A^T \cdot A \cdot x, x \rangle = \langle A \cdot x, A \cdot x \rangle \geq 0$$

pro všechny vektory x. Navíc zjevně

$$B^T = (A^T \cdot A)^T = A^T \cdot A = B.$$

Symetrickým maticím B s takovou vlastností říkáme **nezáporné** a pokud nastane nulová hodnota pouze pro $x = 0$, pak jim říkáme **kladné**. Obdobně hovoříme o **kladných** a **nezáporných** zobrazeních $f : V \to V$.

Pro každé nezáporné zobrazení $f : V \to V$ umíme najít jeho odmocninu, tj. zobrazení g takové, že $g \circ g = f$. Nejjednodušeji to uvidíme v ortonormální bázi, ve které bude mít f diagonální matici. Taková podle našich předchozích úvah vždy existuje a matice A zobrazení f v ní bude mít na diagonále nezáporná reálná vlastní čísla zobrazení f. Kdyby totiž bylo některé z nich záporné, nebyla by splněna podmínka nezápornosti již pro některý z bázových vektorů. Pak ovšem stačí definovat zobrazení g pomocí matice B s odmocninami příslušných vlastních čísel na diagonále.

Nilpotentní a cyklická zobrazení – připomenutí

Definition

Lineární zobrazení $f : V \to V$ se nazývá **nilpotentní**, jestliže existuje celé číslo $k \geq 1$ takové, že iterované zobrazení f^k je identicky nulové. Nejmenší číslo k s touto vlastností se nazývá *stupněm nilpotentnosti* zobrazení f. Zobrazení $f : V \to V$ se nazývá **cyklické**, jestliže existuje báze $(u_1, \dots, u_n)$ prostoru V taková, že $f(u_1) = 0$ a $f(u_i) = u_{i-1}$ pro všechna $i = 2, \dots, n$.

Jinými slovy, matice f v bázi z definice cyklického zobrazení je tvaru

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

Je-li $f(v) = a \cdot v$, pak pro každé přirozené k je $f^k(v) = a^k \cdot v$. Zejména tedy může spektrum nilpotentního zobrazení obsahovat pouze nulový skalár (a ten tam vždy je).

Přímo z definice plyne, že každé cyklické zobrazení je nilpotentní, navíc je jeho stupeň nilpotentnosti roven dimenzi prostoru V. Operátor derivování na polynomech je příkladem cyklického zobrazení. Kupodivu to platí i naopak a každé nilpotentní zobrazení je přímým součtem cyklických. Navíc pro každé lineární zobrazení $f : V \to V$, pro které je součet algebraických násobností vlastních čísel roven dimenzi (to nastane vždy pro prostory nad komplexními skaláry), existuje jednoznačný rozklad V na invariantní podprostory V_i příslušné k jednotlivým vlastním číslům λ_i, na kterých je zobrazení $f - \lambda_i \,\mathrm{id}_{\mathrm{V_i}}$ nilpotentní.

V Jordanově rozkladu vystupují vektorové (pod)prostory a lineární zobrazení na nich s jediným vlastním číslem λ a maticí

$$J = \begin{pmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \lambda & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & & \vdots \\ 0 & 0 & 0 & \dots & \lambda \end{pmatrix}.$$

Takovýmto maticím (a odpovídajícím invariantním podprostorům) se řídá **Jordanův blok**.

Theorem (Věta o Jordanově kanonickém tvaru)

Nechť V je vektorový prostor dimenze n a $f : V \to V$ je lineární zobrazení s n vlastními čísly včetně algebraických násobností. Pak existuje jednoznačný rozklad prostoru V na přímý součet podprostorů

$$V = V_1 \oplus \cdots \oplus V_k$$

takových, že $f(V_i) \subset V_i$, zúžení f na každé V_i má jediné vlastní číslo λ_i a zúžení $f - \lambda_i \cdot \mathrm{id}$ na V_i je buď cyklické nebo nulové zobrazení.

Věta tedy říká, že ve vhodné bázi má každé lineární zobrazení blokově diagonální tvar s Jordanovými bloky podél diagonály. Celkový počet jedniček nad diagonálou v takovém tvaru je roven rozdílu mezi celkovou algebraickou a geometrickou násobností vlastních čísel.

Všimněme si, že jsme tuto větu plně dokázali v případech, kdy jsou všechna vlastní čísla různá nebo když jsou geometrické a algebraické násobnosti vlastních čísel stejné.

Důkaz fakt komplikovaný – v učebnici je a zde snad v jiné prezentaci (pro náročné).