Matematika I – 10b
Něco málo maticového počtu

Jan Slovák

Masarykova univerzita
Fakulta informatiky

19. 11. 2012

Obsah přednášky

Uvažujme opět reálný vektorový prostor V se skalárním součinem. Zobrazení $f : V \to V$ se nazývá **symetrické**, jestliže pro všechny vektory $u, v \in V$ platí

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

V libovolné ortonormální bázi můžeme předchozí vztah v souřadnicích vyjádřit takto:

$$(A \cdot x)^T \cdot y = x^T \cdot (A \cdot y) = x^T \cdot (A^T y).$$

Volbou souřadnic bázových vektorů (tj. jedna jednička a zbytek nuly) se dostaneme ke vztahům $a_{ij} = a_{ji}$ pro jednotlivé komponenty matice A, tzn. ke vztahu $A = A^T$.

Theorem

Zobrazení $f : V \to V$ na vektorovém prostoru se skalárním součinem je symetrické právě tehdy, když v některé (a pak už všech) ortonormální bázi má symetrickou matici.

Adjungovaná zobrazení

Jestliže zvolíme pevně jeden vektor $v \in V$, dosazování vektorů za druhý argument nám dává 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 A^* pro matici zobrazení f^*, x^T souřadnice formy w^*, y souřadnice v)

$$(A^* x^T) \cdot y = x^T \cdot (A \cdot y)$$

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

V případě vektorových prostorů se skalárním součinem, převádí výše uvedené bijekce duální zobrazení f^* na tobrazení $f^* : W \to V$ zadané formulí

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

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 transponovaná A^T. Můžeme proto také přeformulovat definici takto: Symetrické je takové zobrazení $f : V \to V$, které je rovno svému adjungovanému zobrazení f^*. Často se takovým zobrazením také proto říká **samoadjungovaná**.

Theorem (Spektrální rozklad symetrického zobrazení)

Pro každé symetrické 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

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

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.

I při počítání s reálnými čísly užíváme pro zjednodušení rozklady na součiny. Nejjednodušším je vyjádření každého reálného čísla jednoznačně ve tvaru $a = \operatorname{sgn}(a) \cdot |a|$, tj. jako součin znaménka a abolutní hodnoty. Obdobné rozklady se užívají u matic a mají dalekosáhlé důsledky pro numeriku. Zde jen naznačíme.

LU rozklad

Gausova eliminace spočívá v postupném násobení naší matice invertibilními maticemi P_i, které postihovaly přičítání násobků řádků pod právě zpravovávaným. Předpokládejme pro jednoduchost, že nepotřebujeme při eliminaci přehazovat řádky, tj. všechny naše matice P_i mohou být dolní trojúhelníkové s jedničkami na diagonálách. Konečně, inverzní matice k takovýmto P_i jsou opět dolní trojúhelníkové s jedničkami na diagonálách a dostáváme

$$U = P \cdot A = P_k \cdots P_1 \cdot A$$

kde U je horní trojúhelníková matice a tedy

$$A = L \cdot U$$

kde L je dolní trojúhelníková matice s jedničkami na diagonále a U je horní trojúhelníková.

Přímým důsledkem Gausovy eliminace bylo také zjištění, že až na volbu vhodných bází na definičním oboru a oboru hodnot, je každé zobrazení $f : V \to W$ zadáno maticí v blokově diagonálním tvaru s jednotkovou maticí s rozměrem daným dimenzí obrazu f nulovými bloky všude kolem. To lze přeformulovat takto:

Každou matici A typu m/n nad polem skalárů $\mathbb{K}$ lze rozložit na součin

$$A = P \cdot \begin{pmatrix} E & 0 \\ 0 & 0 \end{pmatrix} \cdot Q.$$

Pro čtvercové matice jsme ukázali při diskusi vlastností lineárních zobrazení $f : V \to V$ na komplexních vektorových prostorech, že každou čtvercovou matici A dimenze m umíme rozložit na součin

$$A = P \cdot B \cdot P^{-1}$$

kde B je blokově diagonální s Jordanovými bloky příslušnými k vlastním číslům na diagonále. Všimněme si, že násobení maticí P a její inverzí z opačných stran odpovídá v tomto přípaě právě změně báze na vektorovém prostoru V.

Věta o singulárním rozkladu

Jestliže se omezíme na ortonormální báze, ale chceme znát více informací o struktuře obecných lineárních zobrazení, musíme postupovat o hodně rafinovaněji, než v případě bazí libovolných:

Theorem

Nechť A je reálná matice typu m/n. Pak existují čtvercové ortogonální matice U a V dimenzí m a n, a reálná diagonální matice s nezápornými prvky D dimenze r, $r \leq \min\{m, n\}$, takové, že

$$A = USV^T, \quad S = \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix},$$

kde r je hodnost matice AA^T. Přitom je S určena jednoznačně až na pořadí prvků a prvky diagonální matice D jsou druhé odmocniny vlastních čísel d_i matice AA^T.

Geometrická interpretace singulárního rozkladu

Diagonálním hodnotám matice D z předchozí věty se říká *singulární hodnoty matice A*. Pro přísluąné zobrazení $\varphi : \mathbb{R}^n \to \mathbb{R}^m$ mají jednoduchý geometrický význam: Nechť $K \subset \mathbb{R}^n$ je jednotková sféra pro standardní skalární součin. Obrazem $\varphi(K)$ pak vždy bude (případně degenerovaný) m-rozměrný elipsoid. Singulární čísla matice A jsou přitom velikosti hlavních poloos a věta navíc říká, že původní sféra vždy připouští ortogonální sdružené průměry, jejichž obrazem budou právě všechny poloosy tohoto elipsoidu.

Pro čtvercové matice je vidět, že A je invertibilní právě, když všechna singulární čísla jsou nenulová. Poměr největšího a nejmenšího singulárního čísla je důležitým parametrem pro robustnost řady numerických výpočtů s maticemi, např. pro výpočet inverzní matice.

Definition

Nechť A je reálná matice typu m/n a nechť$A = USV^T$ je její singulární rozklad, $S = \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix}$. Matici $A^{(-1)} := VS'U^T$ s $S' = \begin{pmatrix} D^{-1} & 0 \\ 0 & 0 \end{pmatrix}$ nazýváme *pseudoinverzní matice* k matici A.

Jak ukazuje následující věta, je pseudoinverze důležité zobecnění pojmu inverzní matice.

Theorem

Nechť A je reálná matice typu m/n. Platí

1. *Je-li A invertibilní (zejména tedy čtvercová), pak*

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

2. *pro pseudoinverzi $A^{(-1)}$ platí, že $A^{(-1)}A$ i $AA^{(-1)}$ jsou symetrické a*

$$AA^{(-1)}A = A, \quad A^{(-1)}AA^{(-1)} = A^{(-1)}.$$

3. *Uvažme pro danou matici A systém lineárních rovnic $Ax = b$, $b \in R^m$. Pak $y = A^{(-1)}b \in \mathbb{R}^n$ minimalizuje vzdálenost $\|Ax - b\|$ pro všechny $x \in \mathbb{R}^n$.*

Lineární regrese

Aproximační vlastnost (3) předchozí věty je velice užitečná v případech, kdy máme najít co nejlepší přiblížení (neexistujícího) řešení přeurčeného systému $Ax = b$, kde A je reálná matice typu m/n a m je větší než n.
Např. máme experimentem dáno mnoho naměřených hodnot b_j a chceme najít lineární kombinaci několika funkcí f_i, která bude co nejlépe aproximovat hodnoty b_j. Skutečné hodnoty zvolených funkcí v bodech $y_j \in \mathbb{R}$ zadají matici $a_{ij} = f_i(y_j)$ a naším úkolem je tedy určit koeficienty $x_j \in \mathbb{R}$ tak, aby $\sum_{i=1}^{m}(b_i - (\sum_{j=1}^{n} x_j a_{ij}))^2$ byla minimální. Jinými slovy, hledáme lineární kombinaci funkcí f_i takovou, abychom "dobře" proložili zadané hodnoty b_i. Díky předchozí větě jsou hledané optimální koeficienty $A^{(-1)}b$.