Drsná matematika I – 11
Rozklady matic

Jan Slovák

Masarykova univerzita

25–29.11. 2019

Obsah přednášky

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 případě 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 příslušné 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.

Polární rozklad

Jestliže přepíšeme vztah ze singulárního rozkladu

$$A = USV* = USU * UV* = PW,$$

pak $P = USU*$ je zjevně samoadjungovaná, pozitivně semidefinitní, zatímco W je unitární.
Tomuto rozkladu se říká **polární rozklad**. Navíc je

$$P = \sqrt{AA^*}$$

zadaná jednoznačně a, pokud je A invertibilní, pak i W je jednoznačně určená.
Všimněte si že jde o pěknou analogii goniometrického tvaru komplexního čísla.

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