Maticový počet – Prednáška č. I

Stručný obsah prednášky 20. 9. 2022

Úvod:

- kontakty
- web-stránka kurzu, Google Classroom
- sylabus
- knihy, prerekvizity (zhrnutie LAG I a LAG II)
- domáce úlohy, hodnotenie
- alternatíva: 1–PMA–215 Maticová algebra pre štatistikov

Aplikácie LinAlg v praxi (Kapitola 1.4):

- ODR a PDR $\longrightarrow$ nasekanie intervalu/domény, aproximácie derivácií $\longrightarrow$ lineárne systémy s veľkými a riedkymi maticami
- príklady: počítanie meteorologických modelov, dynamika objektov v slnečnej sústave (gravitational slingshot)
- súvis PageRank (Google) so spektrálnou teóriou nezáporných matíc (Perron – Frobenius)
- Ako počítať? (existencia algoritmov)
- Ako počítať presne? (analýza typu chýb, ich veľkostí a ich postupného "nabaľovania")
- Ako počítať rýchlo? (počet operácií, zjednodušenie algoritmov, úpravy dát, konvergencia výpočtu, triky a pod.)

Presnosť výpočtov, aritmetika pohyblivej čiarky (Kapitola 1.5):

Čísla s pohyblivou čiarkou

Pod t-číslicovým číslom v artimetike *pohyblivej rádovej čiarky* so *základom* β rozumieme

$$f = \pm .d_1 d_2 \dots d_t \times \beta^{\epsilon}, \qquad \text{pričom} \qquad d_1 \neq 0.$$

Základ (alebo báza) β, *exponent* $L \leq \epsilon \leq U$ a číslice (cifry) $0 \leq d_i \leq \beta - 1$ sú celé čísla. V elektronických zariadeniach sa spravidla používa $\beta = 2$ (binárna reprezentácia), pre výpočty s ceruzkou/kriedou na papieri/tabuli je ľudskejšie používať $\beta = 10$. Hodnota t sa nazýva *presnosť*, povolené hranice (dolná L/horná U) pre exponent ϵ môžu závisieť od voľby hardvéru, softvéru, konvencie, technického štandardu a pod.

- príklad rôzneho priebehu eliminácie v presnej a 3-číslicovej aritmetike

Čiastočné pivotovanie

V každom kroku eliminácie hľadať zložku na a pod pozíciou pivota s najväčšou veľkosťou. Ak je to potrebné, vykonať príslušnú výmenu riadkov a dostať tento maximálny koeficient na pozíciu pivota. Ilustrácia tretieho kroku eliminácie s čiastočným pivotovaním v typickom prípade:

$$\left(\begin{array}{ccccc|c} * & * & * & * & * & * \\ 0 & * & * & * & * & * \\ 0 & 0 & \boxed{s} & * & * & * \\ 0 & 0 & s & * & * & * \\ 0 & 0 & s & * & * & * \end{array}\right).$$

Prehladať pozície v treťom stĺpci označené písmenom s a nájsť koeficient s najväčšou absolútnou hodnotou. Ak je to potrebné, vymeniť riadky tak, aby sa tento koeficient ocitol na zakrúžkovanej pozícii tretieho pivota. Zjednodušene povedané, stratégia je maximalizovať veľkosť pivotov v každom kroku iba za použitia riadkových výmen.

- príklad rôzneho priebehu eliminácie v 3-číslicovej aritmetike s pivotovaním a bez pivotovania

Praktická stratégia pre škálovanie

- Zvoliť jednotky (zodpovedajúce stĺpcom), ktoré sú prirodzené pre daný problém a nedeformovať vzťahy medzi “veľkosťou” vecí. Tieto prirodzené jednotky bývajú spravidla zjavné a ďalšie škálovanie stĺpcov sa spravidla nerobí.

- Riadky systému $(A|b)$ škálovať tak, aby koeficient s najväčšou veľkosťou v každom riadku matice A bol rovný 1. T.j. predeliť každú z rovníc koeficientom s najväčšou veľkosťou.

Čiastočné pivotovanie s takto zvolenou škálovacou stratégiou robí z Gaussovej eliminácie (spolu so spätnou substitúciou) extrémne efektívny nástroj. V priebehu času sa táto technika ukázala byť spoľahlivou pre riešenie väčšiny lineárnych systémov, ktoré priniesla prax.

Kompletné pivotovanie

V k-tom kroku Gaussovej eliminácie sa v rozšírenej matici systému $(A|b)$ nájde koeficient s najväčšou veľkosťou na pozícii pivota, ako aj na všetkých pozíciách v A, ktoré sa nachádzajú nadol alebo napravo od nej. Ak je to nutné, vykoná sa príslušná výmena riadkov (rovníc) a stĺpcov (permutovanie neznámych) tak, aby sa najväčší koeficient ocitol na pozícii pivota. Ilustrácia tretieho kroku eliminácie s kompletným pivotovaním v typickom prípade:

$$\left(\begin{array}{ccccc|c} * & * & * & * & * & * \\ 0 & * & * & * & * & * \\ 0 & 0 & s & s & s & * \\ 0 & 0 & s & s & s & * \\ 0 & 0 & s & s & s & * \end{array}\right).$$

Prehladať pozície označené “s” a nájsť koeficient s najväčšou absolútnou hodnotou. Ak je to potrebné, vymeniť riadky a stĺpce tak, aby sa tento koeficient ocitol na zakrúžkovanej pozícii tretieho pivota. Poznámka: výmena stĺpcov A zodpovedá permutácii (alebo premenovaniu) príslušných neznámych.

- kompletné pivotovanie je aspoň tak dobré ako čiastočné pivotovanie
- pre príklad, kde úplné pivotovanie vyhráva, pozri príklad č. 6 z DÚ 1

Maticový počet – Prednáška č. II

Stručný obsah prednášky 27. 9. 2022

Zle podmienené systémy (Kapitola 1.6):

- Príklad zle podmieneného systému, geometrické vysvetlenie.

Zle podmienené lineárne systémy

Systém lineárnych rovníc sa nazýva *zle podmienený*, ak malá perturbácia systému vedie k relatívne veľkej zmene v presnom riešení. Inak sa systém nazýva *dobre podmienený*.

- Geometria – zmena polohy priesečníka skoro rovnobežných priamok pri ich malom posunutí.
- Algebra – skoro singulárna matica A, A^{-1} "naťahuje" malý vektor $b - \hat{b}$ na veľký vektor $x - \hat{x}$.
- Prejavy: veľké zložky v A^{-1}, resp. veľké vlastné/singulárne hodnoty A^{-1}.
- Potreba identifikovať a vyhnúť sa zle podmieneným systémom v praxi.
- *Rezíduá (resp. reziduálne vektory):* pre nepresné riešenie $x_c = (\xi_1, \xi_2, \dots, \xi_n)$ vyhodnotiť veľkosť výrazov $r_i = a_{i1}\xi_1 + a_{i2}\xi_2 + \cdots + a_{in}\xi_n - b_i$. Maticovo: $r = Ax_c - b$.
- Pre dobre podmienené systémy sú rezíduá užitočné, pre zle podmienené ani veľmi nie.

Inverzné matice súčtu a citlivosť na perturbácie (Kapitola 3.8):

- Zovšeobecnenie formuly pre $(I + cd^T)^{-1}$.

Sherman-Morrisonova formula

- Ak $A_{n\times n}$ je regulárna a ak c a d sú $n \times 1$ stĺpcové vektory spĺňajúce $1 + d^T A^{-1} c \neq 0$, potom je $A + cd^T$ regulárna a
$$\left(A + cd^T\right)^{-1} = A^{-1} - \frac{A^{-1}cd^TA^{-1}}{1 + d^TA^{-1}c}.$$
- Zovšeobecnením je *Sherman-Morrison-Woodburyho* formula. Ak C a D sú typu $n \times k$ také, že $\left(I + D^TA^{-1}C\right)^{-1}$ existuje, potom
$$\left(A + CD^T\right)^{-1} = A^{-1} - A^{-1}C\left(I + D^TA^{-1}C\right)^{-1}D^TA^{-1}.$$

- Aplikácia S-M formuly pre výpočet inverzu aktualizovanej matice (pozmenená jedna zložka).

Neumannov rad

Ak je $\lim_{n\to\infty} A^n = 0$, potom je $I - A$ regulárna a platí
$$(I - A)^{-1} = I + A + A^2 + \cdots = \sum_{k=0}^{\infty} A^k.$$
Takýto nekonečný rad sa nazýva *Neumannov rad*; poskytuje aproximácie pre $(I - A)^{-1}$ ak má matica A zložky malej veľkosti. Napr. pre priblíženie prvého rádu máme $(I - A)^{-1} \approx I + A$.

- Odvodenie približnej rovnosti $\left(A + B\right)^{-1} \approx A^{-1} - A^{-1}BA^{-1}$, ak $\|A\| \gg \|B\|$.
- Potreba zavedenia maticovej normy (zíde sa hlavne vzťah $\|XY\| \leq \|X\|\|Y\|$).
- Odvodenie (približných) ohraničení
$$\frac{\left\|A^{-1} - (A+B)^{-1}\right\|}{\|A^{-1}\|} \lesssim \left\|A^{-1}\right\| \|A\| \left(\frac{\|B\|}{\|A\|}\right),$$
$$\frac{\|x - \tilde{x}\|}{\|x\|} \lesssim \left\|A^{-1}\right\| \|A\| \left(\frac{\|B\|}{\|A\|}\right).$$

Citlivosť a podmienenosť

- Regulárna matica A sa nazýva *zle podmienená*, ak malá relatívna zmena v A vedie k veľkej relatívnej zmene v A^{-1}. Miera zlej podmienenosti je udaná *číslom podmienenosti* $\kappa = \|A\| \left\|A^{-1}\right\|$, kde $\|\cdot\|$ je (niektorá) maticová norma.
- Citlivosť riešenia systému $Ax = b$ vzhľadom na perturbácie (alebo chyby) v A zodpovedá tomu, ako je A zle podmienená.

- Pre t-číslicovú artimetiku a κ rádu p je riešenie $Ax = b$ presné na cca. $t - p$ číslic.
- *Otázka:* ako zistiť podmienenosť matice bez výpočtu A^{-1}? *Odpoveď:* SVD – singulárny rozklad

Stručný obsah prednášky 4. 10. 2022

LU rozklad (Kapitola 3.10):

- Tri typy elementárnych riadkových operácií a ich realizácia ako ľavé násobenie elementárnou maticou.
- Gaussova eliminácia bez výmeny riadkov (iba operácie typu III) $\longrightarrow LU$ rozklad.
- Zložky matice L vyzerajú zaujímavo ...
- *Elementárne dolné trojuholníkové* matice $T_k = I - c_k e_k^T$ pre $c_k = (0, \ldots, 0, \mu_{k+1}, \ldots, \mu_n)^T$,

$$T_k = \begin{pmatrix} 1 & 0 & \ldots & 0 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & 1 & 0 & \ldots & 0 \\ 0 & 0 & \ldots & -\mu_{k+1} & 1 & \ldots & 0 \\ \vdots & \vdots & & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & -\mu_n & 0 & \ldots & 1 \end{pmatrix}.$$

Inverz:

$$T_k^{-1} = I + c_k e_k^T = \begin{pmatrix} 1 & 0 & \ldots & 0 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & 1 & 0 & \ldots & 0 \\ 0 & 0 & \ldots & \mu_{k+1} & 1 & \ldots & 0 \\ \vdots & \vdots & & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \mu_n & 0 & \ldots & 1 \end{pmatrix}.$$

- Vyjadrenie $L = T_1^{-1} T_2^{-1} \ldots T_{k-1}^{-1} = I + c_1 e_1^T + c_2 e_2^T + \cdots + c_{k-1} e_{k-1}^T$.

LU rozklad

Ak sa počas Gaussovej eliminácie $n \times n$ matice A operáciami typu III nevyskytne na diagonálnej pozícii pivota nula, potom sa A dá vyjariť ako súčin $A = LU$ a platí:

- L je dolná trojuholníková a U je horná trojuholníková.
- Pre diagonálne zložky platí $l_{ii} = 1$ a $u_{ii} \neq 0$ pre $i = 1, 2, \ldots, n$.
- Zložka l_{ij} pod diagonálou matice L zodpovedá násobku riadku j, ktorý bol odpočítaný od riadku i pri nulovaní pozície (i, j) počas eliminácie.
- U je konečný výsledok Gaussovej eliminácie použitej na maticu A.
- Matice L a U sú jednoznačne určené prvými dvoma vlastnosťami (trojuholníkové matice, jednotky na diagonále v L).

Rozklad matice A na súčin $A = LU$ sa nazýva *LU rozklad matice A*. Matice L a U sa nazývajú *LU faktory pre A*.

- Dôkaz jednoznačnosti LU rozkladu ($L_2^{-1} L_1 = U_2 U_1^{-1}$ musí byť diagonálna, resp. identita).
- Využitie pamäte pri výpočte LU rozkladu.

Riešenie systému pomocou LU rozkladu

- Na vyriešenie regulárneho systému $Ax = b$ pomocou LU rozkladu $A = LU$ je potrebné najprv riešiť doprednou substitúciou (dolný) trojuholníkový systém $Ly = b$ pre neznámu y, a potom riešiť (horný) trojuholníkový systém $Ux = y$ spätnou substitúciou.
- Výhodou tohto prístupu je to, že ak sú LU faktory pre maticu A známe, ľubovoľný systém $Ax = \tilde{b}$ sa dá riešiť pomocou n^2 násobení/delení a $n^2 - n$ sčítaní/odčítaní.

- Porovnanie: na výpočet $x = A^{-1}b$ treba tiež n^2 násobení/delení a $n^2 - n$ sčítaní/odčítaní.
- Existencia LU rozkladu a potreba výmeny riadkov počas eliminácie.

Existencia LU faktorov

Nasledujúce tvrdenia sú ekvivalentné tomu, že pre regulárnu $n \times n$ maticu A existuje LU rozklad.

- Počas Gaussovej eliminácie na horný trojuholníkový tvar iba pomocou operácií typu III sa na diagonále neobjavia nulové pivoty.
- Každá hlavná podmatica A_k matice A je regulárna.

- Blokový rozklad ľavej hornej $(k+1) \times (k+1)$ podmatice
$$A_{k+1} = \begin{pmatrix} A_k & b \\ c^T & \alpha_{k+1} \end{pmatrix} = \begin{pmatrix} L_k & 0 \\ c^T U_k^{-1} & 1 \end{pmatrix} \begin{pmatrix} U_k & L_k^{-1} b \\ 0 & \alpha_{k+1} - c^T A_k^{-1} b \end{pmatrix}.$$
- LDU rozklad ako (symetrické) vylepšenie LU rozkladu. Tiež jednoznačný.
- Pre symetrickú maticu ($A = A^T$) máme $A = LDL^T$.
- Pre kladne definitnú symetrickú máme $A = LD^{1/2}D^{1/2}L^T = R^T R$, kde $R = D^{1/2}L^T$ je horná trojuholníková s kladnou diagonálou – *Choleského rozklad*; viac pri QR rozklade.

$PA = LU$ – samoštúdium (Kapitola 3.10, strany 150 - 153):

- Ako postupovať keď je nevyhnutná výmena riadkov (nulové pivoty, resp. potreba čiastočného pivotovania)? $PA = LU$
- výmena riadkov (s poradím $i, j > k$) a elementárne dolné trojuholníkové matice
$$P_{ij}T_k \dots T_2 T_1 = \tilde{T}_k \dots \tilde{T}_2 \tilde{T}_1 P_{ij},$$
kde $T_i = I - c_k e_k^T$ a $\tilde{T}_i = I - \tilde{c}_k e_k^T$, resp. $\tilde{T}_i^{-1} = I + \tilde{c}_k e_k^T$, pričom $\tilde{c}_k = P_{ij} c_k$. To znamená, že v LU rozklade matice PA vychádzajú tie isté koeficienty, len sú v inom poradí – permutované maticou P.

LU rozklad s výmenou riadkov (prax)

- Pre každú regulárnu maticu A existuje permutačná matica P taká, že matica PA má LU rozklad $PA = LU$.
- Pre výpočet L, U a P postupne prepisujeme tabuľku obsahujúcu zložky matice A nasledovne: každá zložka, ktorá sa nuluje, je nahradená násobkom, ktorý bol na toto nulovanie použitý (t.j. $a_{ij} \mapsto l_{ij}$ pre zložky pod diagonálou). Ak sa použije výmena riadkov (napr. pri čiastočnom pivotovaní), násobky sa vymenia korektným spôsobom.
- Zvyčajne sa nevyžaduje pamätať si celú permutačnú maticu P, ale ak to je potrebné, dá sa získať z matice I postupnou výmenou tých riadkov, ktoré sa vymieňajú počas eliminácie. Namiesto $n \times n$ matice spravidla stačí ukladať "stĺpec počítania permutácií" p, v ktorom sa začína so zložkami zoradenými v štandardnom poradí $(1, 2, \dots, n)^T$ a postupne sa vykonávajú príslušné výmeny.
- Na vyriešenie regulárneho systému $Ax = b$ pomocou LU rozkladu s čiastočným pivotovaním je potrebné permutovať zložky b na $\tilde{b}$ tak, aby zodpovedali vykonaným výmenám riadkov – t.j. podľa p – a potom riešiť dopřednou substitúciou systém $Ly = \tilde{b}$ a následne riešiť systém $Ux = y$ spätnou substitúciou.

Stručný obsah prednášky 11. 10. 2022

Štyri základné priestory a ich vlastnosti (Kapitoly 4.2, 4.4):

- Opakovanie: Stĺpcový, nulový, riadkový a ľavý nulový priestor:
  - stĺpcový priestor $\mathcal{R}(A) = \{Ax \mid x \in \mathbb{R}^n\} \subseteq \mathbb{R}^m$.
  - riadkový priestor $\mathcal{R}(A^T) = \{A^T y \mid y \in \mathbb{R}^m\} \subseteq \mathbb{R}^n$.
  - nulový priestor $\mathcal{N}(A) = \{x \mid Ax = 0\} \subseteq \mathbb{R}^n$.
  - ľavý nulový priestor $\mathcal{N}(A^T) = \{y \mid A^T y = 0\} \subseteq \mathbb{R}^m$.
- Dimenzie: $\dim \mathcal{R}(A) = \dim \mathcal{R}(A^T) = r$, $\dim \mathcal{N}(A) = n - r$ a $\dim \mathcal{N}(A^T) = m - r$.
- Kolmosť: $\mathcal{R}(A^T) \perp \mathcal{N}(A)$ a $\mathcal{R}(A) \perp \mathcal{N}(A^T)$.
- Riešiteľnosť: $Ax = b$ má klasické riešenie $\iff b \in \mathcal{R}(A) \iff b \perp \mathcal{N}(A^T)$.
- Inklúzie: $\mathcal{N}(BA) \supseteq \mathcal{N}(A)$ a $\mathcal{R}(BA) \subseteq \mathcal{R}(B)$.

Súčiny $A^T A$ a AA^T

Pre $A \in M_{m \times n}(\mathbb{R})$ platí:

- $h(A^T A) = h(A) = h(AA^T)$.
- $\mathcal{R}(A^T A) = \mathcal{R}(A^T)$ a $\mathcal{R}(AA^T) = \mathcal{R}(A)$.
- $\mathcal{N}(A^T A) = \mathcal{N}(A)$ a $\mathcal{N}(AA^T) = \mathcal{N}(A^T)$.

Pre komplexné matice $A \in M_{m \times n}(\mathbb{C})$ sa musí operácia transpozície $(\cdot)^T$ nahradiť hermitovským združením $(\cdot)^*$.

- Dôkaz využíva kladnú definitnosť: $\langle Ax \mid Ax \rangle = x^T A^T A x = 0 \iff Ax = 0 \iff x \in \mathcal{N}(A)$.

Normálne rovnice

- Systému $Ax = b$ typu $m \times n$ zodpovedá $n \times n$ systém *normálnych rovníc* $A^T A x = A^T b$.
- Systém $A^T A x = A^T b$ je vždy konzistentný, a to aj v prípade, ak $Ax = b$ nie je.
- Ak je systém $Ax = b$ konzistentný, množina jeho riešení sa zhoduje s množinou riešení systému $A^T A x = A^T b$. Vo všeobecnosti, keď je $Ax = b$ nekonzistentný, riešenia normálnych rovníc zodpovedajú *riešeniam $Ax = b$ v najmenších štvorcoch.*
- $A^T A x = A^T b$ má jednoznačné riešenie práve vtedy, keď $h(A) = n$ a vtedy sa toto riešenie dá vyjadriť ako $x = \left(A^T A\right)^{-1} A^T b$.
- Ak je $Ax = b$ konzistentný a má jednoznačné riešenie, tak to platí aj pre $A^T A x = A^T b$ a jednoznačné riešenie oboch systémov je $x = \left(A^T A\right)^{-1} A^T b$.

Stručný obsah prednášky 18. 10. 2022

Štyri základné priestory a ich vlastnosti – pokračovanie (Kapitoly 4.2, 4.4):

Normálne rovnice

- Systému $Ax = b$ typu $m \times n$ zodpovedá $n \times n$ systém *normálnych rovníc* $A^T Ax = A^T b$.
- Systém $A^T Ax = A^T b$ je vždy konzistentný, a to aj v prípade, ak $Ax = b$ nie je.
- Ak je systém $Ax = b$ konzistentný, množina jeho riešení sa zhoduje s množinou riešení systému $A^T Ax = A^T b$. Vo všeobecnosti, keď je $Ax = b$ nekonzistentný, riešenia normálnych rovníc zodpovedajú *riešeniam $Ax = b$ v najmenších štvorcoch.*
- $A^T Ax = A^T b$ má jednoznačné riešenie práve vtedy, keď $h(A) = n$ a vtedy sa toto riešenie dá vyjadriť ako $x = \left(A^T A\right)^{-1} A^T b$.
- Ak je $Ax = b$ konzistentný a má jednoznačné riešenie, tak to platí aj pre $A^T Ax = A^T b$ a jednoznačné riešenie oboch systémov je $x = \left(A^T A\right)^{-1} A^T b$.

- Podmienenosť $\kappa_{A^T A} \approx (\kappa_A)^2 \longrightarrow$ vyhnúť sa normálnym rovniciam pri numerických výpočtoch.
- Lineárna regresia dát $\mathcal{D} = \{(t_1, b_1), (t_2, b_2), \ldots, (t_m, b_m)\}$ ako príklad použitia metódy najmenších štvorcov pre “preurčený” (overdetermined) $m \times 2$ systém $Ax = b$ pre

$$x = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}, \quad A = \begin{pmatrix} 1 & t_1 \\ 1 & t_2 \\ \vdots & \vdots \\ 1 & t_m \end{pmatrix} \quad \text{a} \quad b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}.$$

- *Chybový vektor* $\varepsilon = Ax - b$.
- Minimalizácia $\|\varepsilon\|^2 = \varepsilon^T \varepsilon = (Ax - b)^T (Ax - b) \longrightarrow$ riešenie normálneho systému $A^T Ax = A^T b$.

Najmenšie štvorce – klasický pohľad (Kapitola 4.5):

- Minimalizácia chyby $f(x) = (Ax - b)^T (Ax - b)$ pre $x \in \mathbb{R}^n$.
- Nutná podmienka minima: $\frac{\partial f}{\partial x_i} = \frac{\partial}{\partial x_i}(x^T A^T Ax - 2x^T A^T b + b^T b) = e_i^T A^T Ax + x^T A^T Ae_i - 2e_i^T A^T b = 0 \longrightarrow$ normálny systém $A^T Ax = A^T b$.
- Ak z rieši $A^T Az = A^T b$ a $y = z + u$, potom $f(y) = f(z) + \|Au\|^2 \geq f(z) \longrightarrow$ v $z + \mathcal{N}(A)$ sa nadobúda globálne minimum f.

Všeobecný problém najmenších štvorcov

Pre $A \in M_{m\times n}(\mathbb{R})$ a $b \in \mathbb{R}^m$ označme $\varepsilon = \varepsilon(x) = Ax - b$. Všeobecný problém najmenších štvorcov je nájsť vektor $x \in \mathbb{R}^n$, ktorý minimalizuje hodnotu

$$\sum_{i=1}^{m} \varepsilon_i^2 = \varepsilon^T \varepsilon = (Ax - b)^T (Ax - b).$$

Ľubovoľný vektor, v ktorom sa takéto minimum nadobúda sa nazýva *riešením $Ax = b$ v najmenších štvorcoch.*

- Množina riešení v najmenších štvorcoch je rovnaká ako množina riešení systému normálnych rovníc $A^T Ax = A^T b$.
- Riešenie v najmenších švorcoch je jednoznačné práve vtedy, keď $h(A) = n$ a v tom prípade je dané ako $x = \left(A^T A\right)^{-1} A^T b$.
- Ak je $Ax = b$ konzistentný, potom je množina riešení $Ax = b$ rovnaká ako množina riešení v najmenších štvorcoch.

Komplementárne podpriestory – šikmé projekcie (Kapitola 5.9):

Komplementárne podpriestory

Podpriestory $\mathcal{X}$, $\mathcal{Y}$ priestoru $\mathcal{V}$ sa nazývajú *komplementárne*, ak

$$\mathcal{V} = \mathcal{X} + \mathcal{Y} \qquad \text{a} \qquad \mathcal{X} \cap \mathcal{Y} = \left\{\vec{0}\right\}.$$

V tom prípade sa $\mathcal{V}$ nazýva *priamy súčet* $\mathcal{X}$ a $\mathcal{Y}$; značenie $\mathcal{V} = \mathcal{X} \oplus \mathcal{Y}$.

Pre vektorový priestor $\mathcal{V}$ a jeho podpriestory $\mathcal{X}$, $\mathcal{Y}$ s bázami $\mathcal{B}_{\mathcal{X}}$ a $\mathcal{B}_{\mathcal{Y}}$ sú nasledujúce tvrdenia ekvivalentné

- $\mathcal{V} = \mathcal{X} \oplus \mathcal{Y}$.
- Pre každé $v \in \mathcal{V}$ existuje jediná dvojica vektorov $x \in \mathcal{X}$ a $y \in \mathcal{Y}$ spĺňajúca $v = x + y$.
- $\mathcal{B}_{\mathcal{X}} \cap \mathcal{B}_{\mathcal{Y}} = \emptyset$ a $\mathcal{B}_{\mathcal{X}} \cup \mathcal{B}_{\mathcal{Y}}$ je báza $\mathcal{V}$.

Projekcie

Nech $\mathcal{V} = \mathcal{X} \oplus \mathcal{Y}$, teda pre každý vektor $v \in \mathcal{V}$ existujú jednoznačné vektory $x \in \mathcal{X}$ a $y \in \mathcal{Y}$ spĺňajúce $v = x + y$.

- Vektor x sa nazýva projekcia v na $\mathcal{X}$ v smere (pozdĺž) podpriestoru $\mathcal{Y}$.
- Vektor y sa nazýva projekcia v na $\mathcal{Y}$ v smere (pozdĺž) podpriestoru $\mathcal{X}$.

- Jednoznačný rozklad $v = x + y$ pre $x \in \mathcal{X}$ a $y \in \mathcal{Y}$ určuje lineárny operátor P daný predpisom $Pv = x$. Nazýva sa *projektor na* $\mathcal{X}$ *v smere* $\mathcal{Y}$.
- $P^2 = P \qquad (P$ je idempotentný).

Maticový počet – Prednáška č. VI

Stručný obsah prednášky 25. 10. 2022

Komplementárne podpriestory – šikmé projekcie – pokračovanie (Kapitola 5.9):

Projekčné operátory

Nech $\mathcal{X}$ a $\mathcal{Y}$ sú komplementárne podpriestory priestoru $\mathcal{V}$, t.j. každé $v \in \mathcal{V}$ sa dá jednoznačne rozložiť ako $v = x + y$ pre $x \in \mathcal{X}$ a $y \in \mathcal{Y}$. Jednoznačne daný lineárny operátor P daný predpisom $Pv = x$ sa nazýva *projektor na $\mathcal{X}$ v smere $\mathcal{Y}$* a platí preň:

- $P^2 = P$ (P je idempotentný).
- $I - P$ je komplementárny projektor na $\mathcal{Y}$ v smere $\mathcal{X}$.
- $\mathcal{R}(P) = \{x \mid Px = x\}$ (množina fixných bodov P)
- $\mathcal{R}(P) = \mathcal{N}(I-P) = \mathcal{X}$ a $\mathcal{R}(I-P) = \mathcal{N}(P) = \mathcal{Y}$.
- Ak $(V) = \mathbb{R}^n$ alebo $\mathbb{C}^n$, potom P sa dá maticovo vyjadriť ako

$$P = (X|0)\,(X|Y)^{-1} = (X|Y)\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}(X|Y)^{-1}$$

  a komplementárny projektor Q ako

$$Q = I - P = (0|Y)\,(X|Y)^{-1} = (X|Y)\begin{pmatrix} 0 & 0 \\ 0 & I_{n-r} \end{pmatrix}(X|Y)^{-1},$$

  kde stĺpce $X_{n\times r}$ a $Y_{n\times n-r}$ tvoria bázy $\mathcal{X}$ a $\mathcal{Y}$.

Projekcie a idempotentné zobrazenia

Lineárny operátor P na vektorovom priestore $\mathcal{V}$ je projekčný operátor práve vtedy, keď $P^2 = P$.

Kolmé projekcie (Kapitola 5.13):

Kolmé (ortogonálne) projekcie

Pre $v \in \mathcal{V}$ majme rozklad $v = m + n$, pričom $m \in \mathcal{M}$ a $n \in \mathcal{M}^\perp$.

- Vektor m sa nazýva *kolmá (ortogonálna) projekcia v* na $\mathcal{M}$.
- Projekčný operátor $P_\mathcal{M}$ na $\mathcal{M}$ v smere $\mathcal{M}^\perp$ sa nazýva *ortogonálny projektor* na $\mathcal{M}$.
- $P_\mathcal{M}$ je (jediný) lineárny operátor spĺňajúci $P_\mathcal{M} v = m$.

- Ak stĺpce $M_{n\times r}$ tvoria bázu $\mathcal{M}$ a stĺpce $N_{n\times n-r}$ tvoria ortonormálnu bázu $\mathcal{M}^\perp$, potom

$$N^T M = 0, \quad M^T N = 0, \quad N^T N = I_{n-r} \quad \text{a} \quad (M|N)^{-1} = \begin{pmatrix} \dfrac{\left(M^T M\right)^{-1} M^T}{N^T} \end{pmatrix}.$$

Konštrukcie matíc kolmej projekcie

Nech $\mathcal{M}$ je r-rozmerný podpriestor $\mathbb{R}^n$ a stĺpce $M_{n\times r}$, $N_{n\times n-r}$ tvoria bázy $\mathcal{M}$, resp. $\mathcal{M}^\perp$. Ortogonálne projektory na $\mathcal{M}$ a $\mathcal{M}^\perp$ sú

- $P_{\mathcal{M}} = M\left(M^T M\right)^{-1} M^T \quad$ a $\quad P_{\mathcal{M}^\perp} = N\left(N^T N\right)^{-1} N^T$;
- navyše platí $P_{\mathcal{M}^\perp} = I - P_{\mathcal{M}}$.

Ak sú stĺpce matíc M a N *ortonormálne*, potom

- $P_{\mathcal{M}} = MM^T \quad$ a $\quad P_{\mathcal{M}^\perp} = NN^T$.
- $P_{\mathcal{M}} = U\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}U^T$, pre ortogonálnu maticu $U = (M|N)$.

Maticový počet – Prednáška č. VII

Stručný obsah prednášky 8. 11. 2022

Kolmé projekcie – dokončenie (Kapitola 5.13):

Charakterizácia projekčných matíc

Nech $P \in M_{n\times n}(\mathbb{R})$ je matica projekčného operátora, t.j. $P^2 = P$. Nasledujúce tvrdenia sú ekvivalentné tomu, že P je matica *kolmej* projekcie:

- $\mathcal{R}(P) \perp \mathcal{N}(P)$.
- $P = P^T$ (t.j. P reprezentuje kolmú projekciu $\Longleftrightarrow PP^T = P^2 = P = P^T$).
- $\|P\|_2 = 1$ pre maticovú 2-normu. (*toto sme neukázali*)

Najbližší bod podpriestoru

Nech $\mathcal{M}$ je podpriestor euklidovského vektorového priestoru $\mathcal{V}$ a b je vektor (bod) vo $\mathcal{V}$. Potom v $\mathcal{M}$ existuje jediný najbližší vektor (bod) p k b a je to práve kolmá projekcia $p = P_{\mathcal{M}}b$. Teda

$$\min_{m\in\mathcal{M}} \|b-m\| = \|b - P_{\mathcal{M}}b\| = dist(b, \mathcal{M}).$$

Táto vzdialenosť sa nazýva *kolmou vzdialenosťou* medzi b a $\mathcal{M}$.

- Problém najmenších štvrocov – minimalizácia $\min_{x\in\mathbb{R}^n} \|Ax-b\|^2$ ako minimalizácia vzdialenosti $\min_{m\in\mathcal{R}(A)} \|b-m\|$, resp. hľadanie projekcie $P_{\mathcal{R}(A)}b$.
- $Ax = P_{\mathcal{R}(A)}b \quad \Longleftrightarrow \quad A^TAx = A^Tb$ (normálne rovnice).

Riešenie v najmenších štvorcoch

Nasledujúce štyri tvrdenia sú ekvivalentné tomu, že $\hat{x}$ je riešením systému $Ax = b$ v najmenších štvorcoch:

- $\|A\hat{x} - b\| = \min_{x\in\mathbb{R}^n} \|Ax-b\|$.
- $A\hat{x} = P_{\mathcal{R}(A)}b$.
- $A^TA\hat{x} = A^Tb$ (resp. $A^*A\hat{x} = A^*b$ ak $A \in M_{m\times n}(\mathbb{C})$).
- $\hat{x} \in A^\dagger b + \mathcal{N}(A)$ ($A^\dagger b$ je riešením v najmenších štvorcoch s najmenšou 2-normou).

Pozor! Tento popis je významný z pohľadu teórie, ale v aritmetike pohyblivej čiarky zvyknú byť výpočty nestabilné (podmienenosť: $\kappa_{A^TA} \gg \kappa_A$; výpočet $A^\dagger$ môže byť numericky nestabilný).

- Pre $a \neq \vec{0} \in \mathbb{R}^n$ je ortogonálny projektor na priamku $\mathcal{L} = span(a)$:

$$P_{\mathcal{L}} = a\left(a^Ta\right)^{-1}a^T = \frac{aa^T}{a^Ta}.$$

- Ak $\|a\| = 1$, tak $P_{\mathcal{L}} = aa^T$ a $P_{\mathcal{L}}(x) = \langle a\,|\,x\rangle a$.

Gram-Schmidtov proces (Kapitola 5.5):

Gram–Schmidtov ortogonalizačný proces

Ak $\mathcal{B} = \{x_1, x_2, \ldots, x_n\}$ je báza vektorového priestoru $\mathcal{S}$ so skalárnym súčinom, potom *Gram-Schmidtova postupnosť* definovaná ako

$$u_1 = \frac{x_1}{\|x_1\|} \qquad \text{a} \qquad u_k = \frac{x_k - \sum_{i=1}^{k-1} \langle u_i \mid x_k \rangle u_i}{\left\| x_k - \sum_{i=1}^{k-1} \langle u_i \mid x_k \rangle u_i \right\|} \qquad \text{pre} \qquad k = 2, \ldots, n$$

tvorí ortonormálnu bázu $\mathcal{S}$.
Ak $\mathcal{S}$ je n-rozmerný podpriestor priestoru $\mathbb{C}^m$, Gram-Schmidtova postupnosť sa dá vyjadriť maticovo ako

$$u_{k+1} = \frac{(I - U_k U_k^*) x_{k+1}}{\|(I - U_k U_k^*) x_{k+1}\|} \qquad \text{pre} \qquad k = 0, 1, \ldots, n-1,$$

kde $U_0 = 0_{m \times 1}$ a $U_k = (u_1 | u_2 | \ldots | u_k)_{m \times k}$ pre $k \geq 1$.

QR rozklad

Každá matica $A_{m\times n}$ s lineárne nezávislými stĺpcami sa dá jednoznačne vyjadriť ako súčin $A = QR$, kde $Q_{m\times n}$ má ortonormálne stĺpce tvoriace bázu stĺpcového priestoru $\mathcal{R}(A)$ a $R_{n\times n}$ je horná trojuholníková matica s kladnými zložkami na diagonále.

- QR rozklad presne popisuje Gram-Schmidtovu ortogonalizáciu lebo stĺpce matice $Q = (q_1|q_2|\ldots|q_n)$ tvoria Gram-Schmidtovu postupnosť pre stĺpce matice $A = (a_1|a_2|\ldots|a_n)$ a R je daná skalárnymi súčinmi

$$R = \begin{pmatrix} \nu_1 & q_1^* a_2 & q_1^* a_3 & \ldots & q_1^* a_n \\ 0 & \nu_2 & q_2^* a_3 & \ldots & q_2^* a_n \\ 0 & 0 & \nu_3 & \ldots & q_3^* a_n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & \nu_n \end{pmatrix},$$

kde $\nu_1 = \|a_1\|$ a $\nu_k = \|a_k - \sum_{i=1}^{k-1} \langle q_i \mid a_k \rangle q_i\|$ pre $k > 1$.

Lineárne systémy a QR rozklad

Ak $h(A_{m\times n}) = n$ a $A = QR$ je QR rozklad matice A, potom riešenie $x = R^{-1}Q^T b$ regulárneho horného trojuholníkového systému

$$Rx = Q^T b$$

zodpovedá klasickému riešeniu $Ax = b$ alebo riešeniu $Ax = b$ v najmenších štvorcoch v závislosti od toho, či je $Ax = b$ konzistentný.

Maticový počet – Prednáška č. VIII

Stručný obsah prednášky 15. 11. 2022

Gram-Schmidtov proces – dokončenie (Kapitola 5.5):

- Systém $Rx = Q^T b$ je numericky vhodnejší (rádovo polovičná podmienenosť) ako systém normálnych rovníc $A^T Ax = A^T b$.
- Choleského faktorizácia ($LD^{1/2}D^{1/2}L^T$) symetrickej, kladne definitnej matice $A^T A$ je $R^T R$, t.j. Gram-Schmidtova procedúra pre A a Gaussova eliminácia Gramovej matice $A^T A$ je v zásade ten istý výpočet.

Vektorové normy (Kapitola 5.1):

Euklidovská norma

Pre vektor $x_{n\times 1}$ sa jeho *euklidovská norma* definuje ako
- $\|x\| = \left(\sum_{i=1}^n x_i^2\right)^{1/2} = \sqrt{x^T x}$ pre $x \in \mathbb{R}^n$,
- $\|x\| = \left(\sum_{i=1}^n |x_i|^2\right)^{1/2} = \sqrt{x^* x}$ pre $x \in \mathbb{C}^n$.

Štandardný skalárny súčin

Skalárne výrazy definované vzťahmi

$$x^T y = \sum_{i=1}^n x_i y_i \in \mathbb{R} \qquad \text{a} \qquad x^* y = \sum_{i=1}^n \overline{x}_i y_i \in \mathbb{C}$$

sa naývajú *štandardný skalárny súčin vektorov x a y* v $\mathbb{R}^n$, resp. v $\mathbb{C}^n$.

Cauchy–Schwarzova nerovnosť

$$|x^* y| \le \|x\|\,\|y\| \qquad \text{pre všetky } x, y \in \mathbb{C}^n.$$

Rovnosť nastáva práve vtedy, keď $y = \alpha x$ pre $\alpha = x^* y / x^* x$.

Trojuholníková nerovnosť

$$\|x + y\| \le \|x\| + \|y\| \qquad \text{pre všetky } x, y \in \mathbb{C}^n.$$

p-normy

Pre $p \ge 1$ je *p-norma* vektora $x \in \mathbb{C}^n$ definovaná ako $\|x\|_p = \left(\sum_{i=1}^n |x_i|^p\right)^{1/p}$.

- Hölderova nerovnosť $\longrightarrow$ trojuholníková nerovnosť pre p-normy.
- Monotónnosť p-noriem, ∞-norma ako limita.

Všeobecné normy na vektorových priestoroch

Norma na reálnom alebo komplexnom vektorovom priestore V je funkcia $\|\cdot\|$ zobrazujúca V do $\mathbb{R}$ spĺňajúca:

$$\begin{aligned}
&\|x\| \geq 0 \qquad \text{a} \qquad \|x\| = 0 \Longleftrightarrow x = 0,\\
&\|\alpha x\| = |\alpha|\|x\| \qquad \text{pre všetky skaláry } \alpha,\\
&\|x+y\| \leq \|x\| + \|y\|.
\end{aligned}$$

- Ekvivalencia vektorových noriem v konečnorozmernom vektorovom priestore.

Maticové normy (Kapitola 5.2):

Frobeniova norma pre matice

Frobeniova maticová norma je pre $A \in M_{m\times n}(\mathbb{C})$ definovaná vzťahmi

$$\|A\|_F^2 = \sum_{i,j} |a_{ij}|^2 = \sum_i \|A_{i*}\|^2 = \sum_j \|A_{*j}\|^2 = \mathrm{Tr}(A^*A) = \mathrm{Tr}(AA^*).$$

- Nerovnosti $\|Ax\|_2^2 \leq \|A\|_F^2 \|x\|_2^2$ a $\|AB\|_F^2 \leq \|A\|_F^2 \|B\|_F^2$.

Všeobecné normy pre matice

Maticová norma je funkcia $\|\cdot\|$ zobrazujúca množinu všetkých komplexných matíc (všetkých konečných typov) do $\mathbb{R}$ spĺňajúca:

$$\begin{aligned}
&\|A\| \geq 0 \qquad \text{a} \qquad \|A\| = 0 \Longleftrightarrow A = 0.\\
&\|\alpha A\| = |\alpha|\|A\| \qquad \text{pre všetky skaláry } \alpha.\\
&\|A+B\| \leq \|A\| + \|B\| \qquad \text{pre matice rovnakého typu.}\\
&\|AB\| \leq \|A\|\,\|B\| \qquad \text{pre všetky matice kompatibilné vzhľadom na násobenie.}
\end{aligned}$$

Indukované maticové normy

Vektorové normy definované na $\mathbb{C}^m$ a $\mathbb{C}^n$ indukujú normu na priestore matíc $M_{m\times n}(\mathbb{C})$ nasledovne:

$$\|A\| = \max_{\|x\|=1} \|Ax\| \qquad \text{pre } A \in M_{m\times n}(\mathbb{C}),\ x \in \mathbb{C}^n.$$

Vo všeobecnosti je v definícii takejto *operátorovej normy* potrebné použiť $\sup \|Ax\|$, vďaka lokálnej kompaktnosti $\mathbb{C}^n$ sa však toto suprémum niekde na jednotkovej sfére v $\mathbb{C}^n$ nadobúda.

- Indukovaná maticová norma je zjavne kompatibilná s vektorovou normou v zmysle:
$$\|Ax\| \leq \|A\|\,\|x\|.$$
- Ak je A regulárna štvorcová matica, potom $\min_{\|x\|=1} \|Ax\| = \dfrac{1}{\|A^{-1}\|}$.

Maticový počet – Prednáška č. IX

Stručný obsah prednášky 22. 11. 2022

Maticové normy – dokončenie (Kapitola 5.2):

Maticová 2-norma

- Pre maticovú normu indukovanú euklidovskou vektorovou normou platí
$$\|A\|_2 = \max_{\|x\|_2=1} \|Ax\|_2 = \sqrt{\lambda_{\max}},$$
kde $\lambda_{\max}$ je najväčšie (reálne) číslo λ také, že $A^*A - \lambda I$ je singulárna.
- Ak je A regulárna (štvorcová), potom
$$\left\|A^{-1}\right\|_2 = \frac{1}{\min\limits_{\|x\|_2=1} \|Ax\|_2} = \frac{1}{\sqrt{\lambda_{\min}}},$$
kde $\lambda_{\min}$ je je najmenšie (reálne) číslo λ také, že $A^*A - \lambda I$ je singulárna.

Pozn. V reči vlastných a singulárnych hodnôt toto znamená, že $\lambda_{\max}$ a $\lambda_{\min}$ sú najväčšia a najmenšia vlastná hodnota matice A^*A a $\sigma_1 = (\lambda_{\max})^{1/2}$, $\sigma_n = (\lambda_{\min})^{1/2}$ sú najväčšia a najmenšia singulárna hodnota matice A.

Vlastnosti 2-normy

Okrem vlastností indukovanej normy maticová 2-norma má aj nasledujúce vlastnosti

- $\|A\|_2 = \max\limits_{\|x\|_2=1} \max\limits_{\|y\|_2=1} |y^*Ax|$.
- $\|A\|_2 = \|A^*\|_2$.
- $\|A^*A\|_2 = \|A\|_2^2$.
- $\left\|\left(\begin{smallmatrix} A & 0 \\ 0 & B \end{smallmatrix}\right)\right\|_2 = \max\{\|A\|_2, \|B\|_2\}$.
- $\|U^*AV\|_2 = \|A\|_2$ ak $UU^* = I$ a $V^*V = I$.

Maticová 1-norma a ∞-norma

Maticové normy indukované vektorovou 1-normou a ∞-normou sú:

- $\|A\|_1 = \max\limits_{\|x\|_1=1} \|Ax\|_1 = \max\limits_j \sum\limits_i |a_{ij}|$
$=$ najväčší súčet absolútnych hodnôt po stĺpcoch.
- $\|A\|_\infty = \max\limits_{\|x\|_\infty=1} \|Ax\|_\infty = \max\limits_i \sum\limits_j |a_{ij}|$
$=$ najväčší súčet absolútnych hodnôt po riadkoch.

- Dôkazy tvrdení o indukovaných maticových normách $\|\cdot\|_1$ a $\|\cdot\|_\infty$.
- Indukované maticové normy majú vlastnosť submultiplikativity, lebo $\|(AB)x\| = \|A(Bx)\| \leq \|A\|\|Bx\| \leq \|A\|(\|B\|\|x\|)$, t.j. $\max_{\|x\|=1} \|ABx\| \leq \|A\|\|B\|$.

Priestory so skalárnym súčinom – opakovanie (Kapitola 5.3):

Všeobecný skalárny súčin

Skalárny súčin na reálnom (alebo komplexnom) vektorovom priestore V je funkcia, ktorá priradí každej usporiadanej dvojici vektorov x, y reálny (alebo komplexný) skalár $\langle x \,|\, y \rangle$, pričom platí

- $\langle x \,|\, x \rangle$ je reálne číslo, $\langle x \,|\, x \rangle \geq 0$ a $\langle x \,|\, x \rangle = 0$ práve vtedy, keď $x = 0$,
- $\langle x \,|\, \alpha y \rangle = \alpha \langle x \,|\, y \rangle$ pre všetky skaláry α,
- $\langle x \,|\, y + z \rangle = \langle x \,|\, y \rangle + \langle x \,|\, z \rangle$,
- $\langle x \,|\, y \rangle = \overline{\langle y \,|\, x \rangle}$ (v reálnom priestore to znamená $\langle x \,|\, y \rangle = \langle y \,|\, x \rangle$).

Treba si všimnúť, že pre fixné x druhá a tretia vlastnosť hovoria, že $\langle x \,|\, \cdot \rangle : y \mapsto \langle x \,|\, y \rangle$ je lineárnou funkciou v premennej y.
Reálny (alebo komplexný) vektorový priestor so skalárnym súčinom sa nazýva *euklidovský priestor* (resp. hermitovský priestor).

Norma v priestore so skalárnym súčinom

Ak V je euklidovský (hermitovský) priestor so skalárnym súčinom $\langle x \,|\, y \rangle$, potom predpis

$$\| \cdot \| = \sqrt{\langle \cdot \,|\, \cdot \rangle}$$

definuje vektorovú normu na V.

Rovnobežníková rovnosť

Pre danú normu $\| \cdot \|$ na vektorovom priestore V k nej existuje skalárny súčin na V, spĺňajúci $\langle \cdot \,|\, \cdot \rangle = \| \cdot \|^2$, práve vtedy, keď pre všetky $x, y \in V$ platí *rovnobežníková rovnosť*:

$$\|x + y\|^2 + \|x - y\|^2 = 2\left(\|x\|^2 + \|y\|^2\right).$$

Unitárne a ortogonálne matice, ortogonálna redukcia (Kapitoly 5.6 a 5.7):

Unitárne a ortogonálne matice

- *Unitárna matica* je taká komplexná matica $U_{n \times n}$, ktorej stĺpce (alebo riadky) tvoria ortonormálnu bázu priestoru $\mathbb{C}^n$.
- *Ortogonálna matica* je taká reálna matica $Q_{n \times n}$, ktorej stĺpce (alebo riadky) tvoria ortonormálnu bázu priestoru $\mathbb{R}^n$.

Charakterizácia unitárnych a ortogonálnych matíc

- Nasledujúce podmienky sú ekvivalentné tomu, že komplexná matica $U_{n \times n}$ je unitárna:
  - ▷ U má ortonormálne stĺpce.
  - ▷ U má ortonormálne riadky.
  - ▷ $U^{-1} = U^*$.
  - ▷ $\|Ux\|_2 = \|x\|_2$ pre všetky $x \in \mathbb{C}^n$.
- Nasledujúce podmienky sú ekvivalentné tomu, že reálna matica $Q_{n \times n}$ je ortogonálna:
  - ▷ Q má ortonormálne stĺpce.
  - ▷ Q má ortonormálne riadky.
  - ▷ $Q^{-1} = Q^T$.
  - ▷ $\|Qx\|_2 = \|x\|_2$ pre všetky $x \in \mathbb{R}^n$.

Elementárne kolmé projekcie

Pre vektor $u \in \mathbb{C}^n$ s normou $\|u\| = 1$ sa matica tvaru

$$P = I - uu^*$$

nazýva *elementárny kolmý projektor.*

Geometria elementárnych projekcií

Pre vektory $u, x \in \mathbb{C}^n$, pričom $\|u\| = 1$, platí

- $(I - uu^*)x$ je kolmá projekcia vektora x do ortogonálneho doplnku $u^\perp$, t.j. do priestoru zloženého zo všetkých vektorov kolmých na vektor u;
- u^*ux je kolmá projekcia vektora x na jednorozmerný priestor $span(u)$;
- $|u^*x|$ predstavuje dĺžku ortogonálnej projekcie vektora x na jednorozmerný priestor $span(u)$.

Elementárne reflexie

Pre nenulový vektor $u_{n\times 1}$ je *elementárna reflexia* cez $u^\perp$ (tiež *Householderova transformácia*) definovaná maticou

$$R = I - 2\frac{uu^*}{u^*u}$$

alebo, ekvivalentne,

$$R = I - 2uu^* \qquad \text{ak} \qquad \|u\| = 1.$$

Vlastnosti elementárnych reflexií

- Všetky elementárne reflexie R sú unitárne, hermitovské a involutórne ($R^2 = I$). T.j.
$$R = R^* = R^{-1}.$$
- Ak $x_{n\times 1}$ je vektor, ktorého prvá zložka $x_1 \neq 0$ a ak sa vektor
$$u = x \pm \mu\|x\|e_1, \qquad \text{kde} \quad \mu = \begin{cases} 1 & \text{ak } x_1 \text{ je reálna,} \\ x_1/|x_1| & \text{ak } x_1 \text{ nie je reálna} \end{cases}$$
použije na vytvorenie elementárnej reflexie R, potom
$$Rx = \mp\mu\|x\|e_1.$$
Inými slovami, R "zrkadlí" vektor x na prvú súradnicovú os.
Výpočtová poznámka: Aby sa vyhlo chybám/vynulovaniu pri výpočtoch v aritmetike pohyblivej rádovej čiarky, pre reálne matice sa volí $u = x + \operatorname{sign}(x_1)\|x\|e_1$.

Maticový počet – Prednáška č. X

Stručný obsah prednášky 30. 11. 2022

Ortogonálna redukcia - dokončenie (Kapitola 5.7):

- *Výpočtová poznámka:* Aby sa vyhlo chybám/vynulovaniu pri výpočtoch v aritmetike pohyblivej rádovej čiarky, pre reálne matice sa volí $u = x + \operatorname{sign}(x_1)\|x\|e_1$.
- Elementárna reflexia zobrazujúca stĺpec $A_{*1} = (a_{11}, a_{21}, \ldots, a_{m1})^T$ na $t_{11}e_1 = (t_{11}, 0, \ldots, 0)^T$ sa dá využiť na redukciu A na horný lichobežníkový tvar $T \longrightarrow$ Householderova redukcia.

Ortogonálna redukcia

- Pre každú maticu $A \in M_{m\times n}(\mathbb{C})$ existuje unitárna matica $P_{m\times m}$ taká, že
$$PA = T$$
má horný lichobežníkový tvar. Ak sa P získa ako súčin elementárnych reflektorov, tento proces sa nazýva *Householderova redukcia.*
- Ak je A štvorcová matica, potom T je horná trojuholníková štvorcová matica.
- Ak je A reálna, potom sa aj P dá zvoliť s reálnymi zložkami – ide o ortogonálnu maticu.

- Pre regulárnu štvorcovú maticu A máme QR rozklad $A = QR$ aj ortogonálnu redukciu $A = P^*T$. Vďaka jednoznačnosti QR-rozkladu sa rovnajú.

QR rozklad

- Pre každú regulárnu maticu $A \in M_{n\times n}(\mathbb{R})$ existujú jednoznačne určená ortogonálna matica Q a jednoznačne určená horná trojuholníková matica R s kladnými zložkami na diagonále také, že
$$A = QR.$$
Štvorcový QR-rozklad je špeciálnym prípadom obdĺžnikového QR-rozkladu pre maticu A typu $m \times n$ s $h(A) = n$.

- Algoritmus Householderovej redukcie je numericky stabilný, preto je niekedy vhodnejší pre výpočet QR-rozkladu ako Gram-Schmidtova ortogonalizácia. Podobne, je stabilnejši ako algoritmus pre LU rozklad (bez pivotovania, resp. s čiastočným pivotovaním).

Rozklady matíc BT, URV (Kapitola 5.11):

- Príklady pravej ($AB = I$) a ľavej ($CA = I$) inverznej matice pre $A_{m\times n}$ plnej hodnosti:
$$C = (A^TA)^{-1}A^T, \qquad B = A^T(AA^T)^{-1}.$$
- Vzorce fungujú vďaka $h(A) = h(A^TA)$, resp. $h(A^T) = h(AA^T)$ v reálnom prípade a $h(A) = h(A^*A)$, resp. $h(A^*) = h(AA^*)$ v komplexnom.
- Pre $\alpha : \mathbb{R}^n \to \mathbb{R}^m$ dané predpisom $\alpha(x) = Ax$, kde matica A typu $m\times n$ má hodnosť $h(A) = r$ je zúžené zobrazenie $\alpha\big|_{\mathcal{R}(A^T)} : \mathcal{R}(A^T) \to \mathcal{R}(A)$ izomorfizmom r-rozmerných podpriestorov $\mathcal{R}(A^T)$ a $\mathcal{R}(A)$, lebo $\mathcal{N}(A) \cap \mathcal{R}(A^T) = \{0\}$, t.j. $\ker \alpha\big|_{\mathcal{R}(A^T)} = \{0\}$.
- Ako rozumne definovať *zovšeobecnené inverzné zobrazenie* $\alpha^{GI} : \mathbb{R}^m \to \mathbb{R}^n$, aby $\alpha^{GI}\big|_{\mathcal{R}(A)} = \left(\alpha\big|_{\mathcal{R}(A^T)}\right)^{-1}$? Dať $\ker(\alpha^{GI}) = \mathcal{N}(A^T)$?

“BT” rozklad

Pre maticu A typu $m \times n$ hodnosti $h(A) = r$ existujú matice $B_{m\times r}$ (“báza”) a $T_{r\times n}$ (“horná trojuholníková”), obe s hodnosťou r, spĺňajúce

$$A_{m\times n} = B_{m\times r} T_{r\times n}.$$

Matice B a T sa získajú pomocou Gaussovej eliminácie – stĺpce B tvoria bázu $\mathcal{R}(A)$ (“pivotové” stĺpce), stupňovitá matica T pozostáva z r nenulových riadkov redukovaného stupňovitého tvaru E_A matice A (pozri kap. 2.2).

- Ak $A = BT$ je BT-rozklad, potom k matici $B_{m\times r}$ existuje (nejaká) ľavá inverzná $L_{r\times m}$ a k matici $T_{r\times n}$ (nejaká) pravá inverzná $R_{n\times r}$. Potom matica RL typu $n \times m$ hodnosti r spĺňa rovnice vonkajšej ($YAY = Y$) a vnútornej ($AXA = A$) zovšeobecnenej inverzie. *Toto sme na prednáške nespravili, ale na niektorom z cvík sme skúmali vzorec* $A^\dagger = (BT)^\dagger = T^\dagger B^\dagger = T^T(TT^T)^{-1}(B^TB)^{-1}B^T$ *a ukázali, že spĺňa Penroseove axiómy pre pseudoinverz, ktoré požadujú okrem vnútornej a vonkajšej inverzie aj symetriu súčinov* $AA^\dagger$ *a* $A^\dagger A$.

“URV” rozklad

Pre každú maticu $A \in M_{m\times n}(\mathbb{R})$ hodnosti r existujú ortogonálne matice $U_{m\times m}$ a $V_{n\times n}$ a regulárna matica $C_{r\times r}$ taká, že

$$A = URV^T = U \begin{pmatrix} C_{r\times r} & 0 \\ 0 & 0 \end{pmatrix} V^T.$$

- Prvých r stĺpcov matice U tvorí ortonormálnu bázu $\mathcal{R}(A)$.
- Posledných $m - r$ stĺpcov matice U tvorí ortonormálnu bázu $\mathcal{N}(A^T)$.
- Prvých r stĺpcov matice V tvorí ortonormálnu bázu $\mathcal{R}(A^T)$.
- Posledných $n - r$ stĺpcov matice V tvorí ortonormálnu bázu $\mathcal{N}(A)$.

Každý výber ortonormálnych báz pre štyri základné podpriestory matice A vedie k jej inému URV rozkladu. V komplexnom prípade treba nahradiť “ortogonálne” “unitárnym” a transponovanie $(\cdot)^T$ hermitovským združením $(\cdot)^*$.

- Pomocou ortogonálnej redukcie (Householderova redukcia) sa dá dosiahnuť, že matica C v URV rozklade bude horná trojuholníková.

Singulárny rozklad a pseudoinverzy (Kapitola 5.12):

- Ak sa za stĺpce $U_{m\times m}$ a $V_{n\times n}$ v URV-rozklade vezmú vlastné vektory matíc AA^T, resp. A^TA, dostaneme dokonca diagonálnu maticu R – singulárny rozklad.
- Singulárne vektory u_1 a v_1 v U a V, zodpovedajúce najväčšej singulárnej hodnote σ_1, sú tie, kde sa realizuje indukovaná maticová norma, t.j.

$$\sigma_1 = \|A^T\|_2 = \max_{\|y\|_2=1} \|A^Ty\|_2 = \|A^Tu_1\|_2 \qquad \text{a} \qquad \sigma_1 = \|A\|_2 = \max_{\|x\|_2=1} \|Ax\|_2 = \|Av_1\|_2.$$

- dôkaz pomocou Lagrangeových multiplikátorov - pozri dôkaz rovnosti 5.2.7 na str. 281–282 (plus odkaz na str. 227) v C. Meyer: *Matrix Analysis and Applied Linear Algebra*
- Platí tiež $v_i = \frac{A^Tu_i}{\sigma_i}$ a $u_i = \frac{Av_i}{\sigma_i}$.

Stručný obsah prednášky 6. 12. 2022

Singulárny rozklad a pseudoinverzy – dokončenie (Kapitola 5.12):

Singulárny rozklad

Pre každú maticu $A \in M_{m\times n}(\mathbb{R})$ hodnosti r existujú ortogonálne matice $U_{m\times m}$ a $V_{n\times n}$ a diagonálna matica $D_{r\times r} = \mathrm{diag}(\sigma_1, \sigma_2, \dots, \sigma_r)$ taká, že

$$A = U \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix} V^T, \qquad \text{kde} \quad \sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0.$$

Reálne čísla σ_i sa nazývajú *singulárne hodnoty* matice A. Ak $r < p = \min\{m, n\}$, hovoríme, že A má dodatočných $p-r$ nulových singulárnych hodnôt. Takýto rozklad sa nazýva *singulárny rozklad* matice A (singular value decomposition), stĺpce matíc U a V sú ľavé, resp. pravé *singulárne vektory* matice A.

- Penroseove rovnice pre pseudoinverziu:

$$\begin{aligned} AA^\dagger A &= A, & A^\dagger AA^\dagger &= A^\dagger, \\ (AA^\dagger)^T &= AA^\dagger, & (A^\dagger A)^T &= A^\dagger A. \end{aligned}$$

- Dá sa ukázať, že tento systém rovníc má jednoznačné riešenie.

Moore–Penrosove pseudoinverzné matice

Pomocou URV rozkladu sa dá Mooreova-Penroseova pseudoinverzná matica pre

$$A_{m\times n} = U \begin{pmatrix} C_{r\times r} & 0 \\ 0 & 0 \end{pmatrix} V^T \qquad \text{vyjadriť ako} \quad A^\dagger_{n\times m} = V \begin{pmatrix} C^{-1} & 0 \\ 0 & 0 \end{pmatrix} U^T.$$

- Ak je systém $Ax = b$ konzistentný, $x = A^\dagger b$ je jeho riešenie s minimálnou euklidovskou normou.
- Ak je systém $Ax = b$ nekonzistentný, $x = A^\dagger b$ je jeho riešením v najmenších štvorcoch s minimálnou euklidovskou normou.
- Pri použití SVD-rozkladu s $C = D = \mathrm{diag}(\sigma_1, \sigma_2, \dots, \sigma_r)$ máme

$$A^\dagger = V \begin{pmatrix} C^{-1} & 0 \\ 0 & 0 \end{pmatrix} U^T = \sum_{i=1}^{r} \frac{v_i u_i^T}{\sigma_i} \qquad \text{a} \quad A^\dagger b = \sum_{i=1}^{r} \frac{(u_i^T b)}{\sigma_i} v_i.$$

Obraz jednotkovej sféry

Pre regulárnu $n \times n$ maticu A so singulárnymi hodnotami $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n$ a SVD rozkladom $A = UDV^T$ s $D = \mathrm{diag}(\sigma_1, \sigma_2, \dots, \sigma_n)$ je obraz jednotkovej sféry vzhľadom na 2-normu elipsoid, ktorého k-ta polos je $\sigma_k U_{*k}$ (σ_k-násobok k-teho stĺpca matice U). Navyše, vektor V_{*k} reprezentuje bod na jednotkovej sfére, ktorý sa zobrazí na koncový bod tejto polosi, $AV_{*k} = \sigma_k U_{*k}$. Špeciálne

- $\sigma_1 = \|AV_{*1}\|_2 = \max_{\|x\|_2=1} \|Ax\|_2 = \|A\|_2$,
- $\sigma_n = \|AV_{*n}\|_2 = \min_{\|x\|_2=1} \|Ax\|_2 = 1/\|A^{-1}\|_2$,

Miera deformácie (sploštenia) jednotkovej sféry transformáciou danou A vieme určiť pomocou *čísla podmienenosti* vzhľadom na 2-normu

- $\kappa_2 = \frac{\sigma_1}{\sigma_n} = \|A\|_2 \|A^{-1}\|_2 \geq 1$.

Treba si všimnúť, že $\kappa_2 = 1$ práve vtedy, keď A je ortogonálna matica a ňou daná transformácia je izometria.

- pre systém $Ax = b$ má riešenie $A^\dagger b$ (s použitím pseudoinverzu $A^\dagger$) najmenšiu euklidovskú normu spomedzi riešení (klasických, či v najmenších štvorcoch)

Determinanty (Kapitoly 6.1 a 6.2):

- opakovanie z I. ročníka, str. 459–488 v knihe C. Meyera, “rámčeky” uvedené v Prehľade III.

Základné vlastnosti vlastných hodnôt a vektorov (Kapitola 7.1):

- opakovanie z I. ročníka, str. 489–496 v knihe C. Meyera, “rámčeky” uvedené v Prehľade IV.
- ak $\lambda \in \sigma(A)$, tak aj $\bar{\lambda} \in \sigma(A)$ pre reálnu maticu A
- *symetrické polynómy*, príklad pre $\lambda_1, \lambda_2, \lambda_3, \lambda_4$:

$$
\begin{aligned}
s_1 &= \sum_i \lambda_i = \lambda_1 + \lambda_2 + \lambda_3 + \lambda_4, \\
s_2 &= \sum_{i<j} \lambda_i\lambda_j = \lambda_1\lambda_2 + \lambda_1\lambda_3 + \lambda_1\lambda_4 + \lambda_2\lambda_3 + \lambda_2\lambda_4 + \lambda_3\lambda_4, \\
s_3 &= \sum_{i<j<k} \lambda_i\lambda_j\lambda_k = \lambda_1\lambda_2\lambda_3 + \lambda_1\lambda_2\lambda_4 + \lambda_1\lambda_3\lambda_4 + \lambda_2\lambda_3\lambda_4 \\
s_4 &= \sum_{i<j<k<l} \lambda_i\lambda_j\lambda_k\lambda_l = \lambda_1\lambda_2\lambda_3\lambda_4.
\end{aligned}
$$

Koeficienty charakteristického polynómu

Ak má charakteristická rovnica matice $A_{n\times n}$ tvar $\lambda^n + c_1\lambda^{n-1} + c_2\lambda^{n-2} + \cdots + c_{n-1}\lambda + c_n = 0$ a ak s_k označuje k-ty symetrický polynóm vlastných hodnôt $\lambda_1, \lambda_2, \ldots, \lambda_n$, potom

- $c_k = (-1)^k \sum$(všetky hlavné $k \times k$ minory),
- $s_k = \sum$(všetky hlavné $k \times k$ minory),
- $\mathrm{Tr}(A) = \lambda_1 + \lambda_2 + \cdots + \lambda_n = -c_1$,
- $\det(A) = \lambda_1\lambda_2 \cdots \lambda_n = (-1)^n c_n$.

- spojitá závislosť vlastných hodnôt matice od jej zložiek
- spektrálny polomer $\rho(A) = \max_{\lambda\in\sigma(A)} |\lambda|$ a jeho ohraničenie pomocou (ľubovoľnej) maticovej normy $\rho(A) \le \|A\|$.

Geršgorinove kruhy

- Všetky vlastné hodnoty matice $A \in M_{n\times n}(\mathbb{C})$ sa nachádzajú v množine $\mathcal{G}_r$ – zjednotení n *Geršgorinovych kruhov* daných pomocou

$$|z - a_{ii}| \le r_i, \qquad \text{kde} \qquad r_i = \sum_{\substack{j=1 \\ j\neq i}}^{n} \|a_{ij}\| \quad \text{pre } i = 1, 2, \ldots, n.$$

  Inými slovami, vlastné hodnoty sú “uväznené” v sade kruhov so stredmi a_{ii} s polomermi danými súčtami absolútnych hodnôt zložiek v stĺpci A_{*i} okrem diagonálnej zložky a_{ii}.
- Navyše, ak zjednotenie $\mathcal{U}$ k-tich Geršgorinovych kruhov nemá prienik so zvyšnými $n - k$ kruhmi, potom sa v k-kruhovom $\mathcal{U}$ nachádza práve k vlastných hodnôt matice A, počítajúc s násobnostami.
- Keďže $\sigma(A) = \sigma(A^T)$, sčítanie absolútnych hodnôt mimodiagonálnych zložiek po riadkoch sa dá nahradiť súčtom po stĺpcoch, teda vlastné hodnoty sa nachádzajú aj v zjednotení kruhov $\mathcal{G}_c$ daných pomocou

$$|z - a_{ii}| \le c_j, \qquad \text{kde} \qquad c_j = \sum_{\substack{i=1 \\ i\neq j}}^{n} \|a_{ij}\| \quad \text{pre } j = 1, 2, \ldots, n.$$

- Kombináciou riadkového a stĺpcového prístupu dostávame, že vlastné hodnoty matice A sa nachádzajú v prieniku $\mathcal{G}_r \cap \mathcal{G}_c$.

- dôkaz nabudúce

Maticový počet – Prednáška č. XII

Stručný obsah prednášky 13. 12. 2022

Základné vlastnosti vlastných hodnôt a vektorov – dokončenie (Kapitola 7.1):

Geršgorinove kruhy

- Všetky vlastné hodnoty matice $A \in M_{n\times n}(\mathbb{C})$ sa nachádzajú v množine $\mathcal{G}_r$ – zjednotení n *Geršgorinovych kruhov* daných pomocou

$$|z - a_{ii}| \leq r_i, \qquad \text{kde} \qquad r_i = \sum_{\substack{j=1\\ j\neq i}}^{n} \|a_{ij}\| \quad \text{pre } i = 1, 2, \dots, n.$$

  Inými slovami, vlastné hodnoty sú "uväznené" v sade kruhov so stredmi a_{ii} s polomermi danými súčtami absolútnych hodnôt zložiek v stĺpci A_{*i} okrem diagonálnej zložky a_{ii}.
- Navyše, ak zjednotenie $\mathcal{U}$ k-tich Geršgorinovych kruhov nemá prienik so zvyšnými $n-k$ kruhmi, potom sa v k-kruhovom $\mathcal{U}$ nachádza práve k vlastných hodnôt matice A, počítajúc s násobnosťami.
- Keďže $\sigma(A) = \sigma(A^T)$, sčítanie absolútnych hodnôt mimodiagonálnych zložiek po riadkoch sa dá nahradiť súčtom po stĺpcoch, teda vlastné hodnoty sa nachádzajú aj v zjednotení kruhov $\mathcal{G}_c$ daných pomocou

$$|z - a_{ii}| \leq c_j, \qquad \text{kde} \qquad c_j = \sum_{\substack{i=1\\ i\neq j}}^{n} \|a_{ij}\| \quad \text{pre } j = 1, 2, \dots, n.$$

- Kombináciou riadkového a stĺpcového prístupu dostávame, že vlastné hodnoty matice A sa nachádzajú v prieniku $\mathcal{G}_r \cap \mathcal{G}_c$.

- dôkaz

Diagonalizácia pomocou podobnostnej transformácie (Kapitola 7.2):

Schurova veta o triangularizácii

Každá štvorcová matica je unitárne podobná hornej trojuholníkovej matici. To znamená, že pre každú $A_{n\times n}$ existuje unitárna matica U (nie jednoznačná) a horná trojuholníková T (nie jednoznačná) také, že $U^*AU = T$. Diagonálne zložky T sú vlastnými hodnotami A.

- dôkaz indukciou
- vybrať vlastný vektor x matice A, doplniť na ortonormálnu bázu, napr. pomocou elementárnej reflexie posielajúcej $x \leftrightarrow e_1$, $R = (x|V)$, potom $R^{-1}AR = RAR = \begin{pmatrix} \lambda & x^*AV \\ 0 & V^*AV \end{pmatrix}$, pokračovať s $(n-1)\times(n-1)$ maticou V^*AV.
- *Cayley-Hamiltonova veta:* $\chi_A(A) = 0$.
- Dôkaz pomocou Schurovej lemy; najprv ukázať pre trojuholníkové matice $\chi_T(T) = 0$, tvrdenie potom vyplýva z podobnosti – *pozrieť ZS 2020/21, resp. knihu*

Násobnosti

Pre $\lambda \in \sigma(A) = \{\lambda_1, \lambda_2, \ldots, \lambda_s\}$ definujme:

- *Algebraická násobnosť* vlastnej hodnoty λ zodpovedá násobnosti λ ako koreňa charakteristického polynómu $\chi_A(x)$. Inými slovami, $\mathit{alg\ mult}_A(\lambda_i) = a_i$ práve vtedy, keď $(x-\lambda_1)^{a_1} \ldots (x-\lambda_s)^{a_s} = 0$ je charakteristickou rovnicou matice A.
- Ak $\mathit{alg\ mult}_A(\lambda) = 1$, λ sa nazýva *jednoduchou* vlastnou hodnotou matice A.
- *Geometrická násobnosť* vlastnej hodnoty λ je $\dim \mathcal{N}(A - \lambda I)$. Inými slovami, $\mathit{geo\ mult}_A(\lambda)$ zodpovedá najväčšiemu počtu lineárne nezávislých vlastných vektorov prislúchajúcich k vlastnej hodnote λ.
- Vlastné hodnoty, pre ktoré $\mathit{alg\ mult}_A(\lambda) = \mathit{geo\ mult}_A(\lambda)$, sa nazývajú *polo-jednoduché* (semisimple) vlastné hodnoty matice A.

Nerovnosť násobností

Pre každú maticu $A \in M_{n\times n}(\mathbb{C})$ a pre každé $\lambda \in \sigma(A)$ platí:

$$\mathit{geo\ mult}_A(\lambda) \leq \mathit{alg\ mult}_A(\lambda).$$

Diagonalizovateľnosť a násobnosti

Matica A typu $n \times n$ je diagonalizovateľná práve vtedy, keď

$$\mathit{geo\ mult}_A(\lambda) = \mathit{alg\ mult}_A(\lambda)$$

pre každé $\lambda \in \sigma(A)$, teda práve vtedy, keď je každá jej vlastná hodnota polojednoduchá.

Spektrálna veta pre diagonalizovateľné matice

Matica $A_{n\times n}$ so spektrom $\sigma(A) = \{\lambda_1, \lambda_2, \ldots, \lambda_k\}$ je diagonalizovateľná práve vtedy, ak existujú matice $\{G_1, G_2, \ldots, G_k\}$ spĺňajúce

$$A = \lambda_1 G_1 + \lambda_2 G_2 + \cdots + \lambda_k G_k,$$

pričom G_i majú nasledujúce vlastnosti:

- G_i je projekčná matica na $\mathcal{N}(A - \lambda_i I)$ v smere $\mathcal{R}(A - \lambda_i I)$.
- $G_i G_j = 0$ pre $i \neq j$.
- $G_1 + G_2 + \cdots + G_k = I$.

Takýto rozklad sa nazýva *spektrálny rozklad matice* A a G_i sa nazývajú *spektrálne projektory* prislúchajúce A.

Jednoduché vlastné hodnoty a projektory

Ak x a y^* sú pravé a ľavé vlastné vektory prislúchajúce jednoduchej vlastnej hodnote $\lambda \in \sigma(A)$, potom

$$G = \frac{xy^*}{y^*x}$$

je projektor na $\mathcal{N}(A - \lambda I)$ v smere $\mathcal{R}(A - \lambda I)$, teda G je spektrálny projektor prislúchajúci vlastnej hodnote λ.

Súhrn diagonalizovateľnosti

Pre $n \times n$ maticu A so spektrom $\sigma(A) = \{\lambda_1, \lambda_2, \dots, \lambda_k\}$ sú nasledujúce podmienky ekvivalentné:

- A je podobná diagonálnej matici, t.j. $P^{-1}AP = D$.
- A má úplnú sadu lineárne nezávislých vlastných vektorov.
- Každá vlastná hodnota λ_i je polojednoduchá, t.j. $\textit{geo mult}_A(\lambda_i) = \textit{alg mult}_A(\lambda_i)$.
- $A = \lambda_1 G_1 + \lambda_2 G_2 + \cdots + \lambda_k G_k$, kde
  - ▷ G_i je projektor na $\mathcal{N}(A - \lambda_i I)$ v smere $\mathcal{R}(A - \lambda_i I)$,
  - ▷ $G_i G_j = 0$ pre $i \neq j$,
  - ▷ $G_1 + G_2 + \cdots + G_k = I$,
  - ▷ $G_i = \prod_{\substack{j=1 \\ j \neq i}}^{k} (A - \lambda_j I) / \prod_{\substack{j=1 \\ j \neq i}}^{k} (\lambda_i - \lambda_j)$,
  - ▷ Ak λ_i je jednoduchá vlastná hodnota, ktorej prislúchajú pravý a ľavý vlastný vektor x, resp. y^*, potom $G_i = xy^*/y^*x$.

Normálne matice (Kapitola 7.5):

- Kedy sú pravé vlastné vektory (stĺpce P) aj ľavými (riadky P^T)? Tvoria ortonormálnu bázu, resp. $P^{-1} = P^T$ – matica P je ortogonálna (unitárna).

Unitárna diagonalizácia

Matica $A_{n \times n}$ je unitárne podobná diagonálnej matici (t.j. A má úplnú sadu ortonormálnych vlastných vektorov) práve vtedy, keď $A^*A = AA^*$. Matice spĺňajúce túto rovnosť sa nazývajú *normálne*.

- Ak $U^*AU = D$, kde U je unitárna a D diagonálna, stĺpce U tvoria úplnú sadu ortonormálnych vlastných vektorov matice A a diagonálne zložky matice D sú príslušné vlastné hodnoty.

- Dôkaz v knihe používa fakt, že normálne matice sú tzv. RPN matice ("range perpendicular to null-space"), t.j. $\mathcal{R}(A) \perp \mathcal{N}(A)$. Pre ne má URV rozklad tvar URU^*.
- Na prednáške sme ukázali, že unitárna podobnosť zachováva normalitu a povedali si, že normálne trojuholníkové matice musia byť diagonálne. Potom tvrdenie vyplýva zo Schurovej lemmy.

Maticový počet – Prednáška č. XIII

Stručný obsah náhradnej prednášky, on-line zo záznamu

Normálne matice – dokončenie (Kapitola 7.5):

Symetrické a hermitovské matice

Okrem vlastností, ktoré majú všetky normálne matice,

- reálne symetrické a hermitovské matice majú reálne vlastné hodnoty,
- A je reálna symetrická práve vtedy, keď je ortogonálne podobná reálnej diagonálnej matici D – t.j. $P^T AP = D$ pre nejakú ortogonálnu maticu P,
- reálne antisymetrické a antihermitovské matice majú rýdzo imaginárne vlastné hodnoty.

Funkcie diagonalizovateľných matíc (Kapitola 7.3):

- Ak $P^{-1}AP = B$, potom pre ľuboboľný polynóm $p(x)$ platí $p(B) = P^{-1}p(A)P$.
- Inšpirácia mocninovými radmi, napr. $e^z = \sum_{k=0}^{\infty} \frac{z^k}{k!}$, vedie k definícii *maticovej exponenciály:* $e^A = I + A + \frac{A^2}{2!} + \frac{A^3}{3!} + \dots$
- Pre diagonalizovateľnú maticu $A = PDP^{-1}$ dostaneme rovnosť $e^A = P\,\mathrm{diag}(e^{\lambda_1}, e^{\lambda_2}, \dots, e^{\lambda_n})P^{-1}$.

Funkcie diagonalizovateľných matíc

Nech $A = PDP^{-1}$ je diagonalizovateľná matica, v ktorej sú vlastné hodnoty v $D = \mathrm{diag}(\lambda_1 I, \lambda_2 I, \dots, \lambda_k I)$ zlúčené pri opakovaní. Pre funkciu $f(z)$, ktorá je definovaná pre každé $\lambda_i \in \sigma(A)$, definujme

$$\begin{aligned} f(A) = Pf(D)P^{-1} &= P \begin{pmatrix} f(\lambda_1)I & 0 & \dots & 0 \\ 0 & f(\lambda_2)I & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & f(\lambda_k)I \end{pmatrix} P^{-1} \\ &= f(\lambda_1)G_1 + f(\lambda_2)G_2 + \dots + f(\lambda_k)G_k, \end{aligned}$$

kde G_i je i-ty spektrálny projektor prislúchajúci vlastnej hodnote λ_i.

Nekonečné rady

Ak $f(z) = \sum_{n=0}^{\infty} c_n(z - z_0)^n$ konverguje pre $|z - z_0| < r$ a ak $|\lambda_i - z_0| < r$ pre každú vlastnú hodnotu diagonalizovateľnej matice A, potom

$$f(A) = \sum_{n=0}^{\infty} c_n(A - z_0 I)^n.$$

Navyše sa dá ukázať, že maticový rad na pravej strane konverguje práve vtedy, keď $|\lambda_i - z_0| < r$ pre každé λ_i, bez ohľadu na to, či je alebo nie je matica A diagonalizovateľná. Preto takýto rad slúži ako definícia $f(A)$ pre funkcie, ktoré sa dajú vyjadriť pomocou Taylorovho radu bez ohľadu na diagonalizovateľnosť matice A.

Funkcie Jordanovych blokov

Pre $k \times k$ Jordanov blok J_* s vlastnou hodnotou λ a pre funkciu $f(z)$, pre ktorú sú definované $f(\lambda)$, $f'(\lambda)$, …, $f^{(k-1)}(\lambda)$, je $f(J_*)$ definované ako

$$f(J_*) = f\begin{pmatrix} \lambda & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{pmatrix} = \begin{pmatrix} f(\lambda) & f'(\lambda) & \frac{f''(\lambda)}{2!} & \cdots & \frac{f^{(k-1)}(\lambda)}{(k-1)!} \\ & f(\lambda) & f'(\lambda) & \ddots & \vdots \\ & & \ddots & \ddots & \frac{f''(\lambda)}{2!} \\ & & & f(\lambda) & f'(\lambda) \\ & & & & f(\lambda) \end{pmatrix}.$$

Systémy diferenciálnych rovníc (Kapitola 7.4):

- Pre maticovú exponenciálu e^{At} diagonalizovateľnej matice $A = PDP^{-1}$ platí:
  - $\frac{de^{At}}{dt} = P\,\mathrm{diag}\left(\frac{de^{\lambda_1 t}}{dt}, \frac{de^{\lambda_2 t}}{dt}, \dots, \frac{de^{\lambda_n t}}{dt}\right) P^{-1} = P\,\mathrm{diag}\left(\lambda_1 e^{\lambda_1 t}, \lambda_2 e^{\lambda_2 t}, \dots, \lambda_n e^{\lambda_n t}\right) P^{-1} = Ae^{At}$.
  - $Ae^{At} = e^{At}A$ (resp. $Af(A) = f(A)A$ pre ľubovoľnú maticovú funkciu $f(A)$).
  - $e^{-At}e^{At} = e^{At}e^{-At} = I = e^0$.
- Z prvej rovnosti vyplýva, že $u(t) = e^{At}c$ bude riešením lineárneho systému diferenciálnych rovníc (s konštantnými koeficientami) $u' = Au$ a počiatočnou podmienkou $u(0) = c$.

Diferenciálne rovnice

Ak $A_{n\times n}$ je diagonalizovateľná so spektrom $\sigma(A) = \{\lambda_1, \lambda_2, \dots, \lambda_k\}$, potom riešenie lineárnej diferenciálnej rovnice $u' = Au$ s počiatočnou podmienkou $u(0) = c$ sa dá vyjadriť ako

$$u(t) = e^{At}c = e^{\lambda_1 t}v_1 + e^{\lambda_2 t}v_2 + \dots + e^{\lambda_k t}v_k,$$

kde v_i je vlastný vektor získaný i-tym spektrálnym projektorom G_i: $v_i = G_i c$.

Kladne definitné matice (Kapitola 7.6):

- Nezdá sa, že by obsahovala oveľa viac, ako bolo na LAG II.

Jordanove tvary a maticové funkcie (Kapitola 7.7, 7.8 a 7.9):

- Nestihli sme ... Na skúšku sa očakávajú vedomosti na úrovni LAG II, pozri Prehľad 4.

Kladné matice (Kapitola 8.2):

- Definícia kladnej a nezápornej matice: $A > 0$ ak $a_{ij} > 0$, resp. $A \geq 0$ ak $a_{ij} \geq 0$ pre všetky i, j.
- Aplikácie: výrobné vzťahy (produkčná/technologická matica), pravdepodobnosti, stochastické matice prechodu, atď.
- $A > 0 \Longrightarrow \rho(A) > 0$
- Zopár nerovností:
  $P > 0, x \geq 0, x \neq 0 \Longrightarrow Px > 0,$
  $N \geq 0, u \geq v \geq 0 \Longrightarrow Nu \geq Nv,$
  $N \geq 0, z > 0, Nz = 0 \Longrightarrow N = 0,$
  $N \geq 0, N \neq 0, u > v > 0 \Longrightarrow Nu > Nv.$
- $|Ax| \leq |A||x|$, kde $|x|^T = (|x_1|, |x_2|, \dots, |x_n|)$. Tiež $|A| = A \Longleftrightarrow A \geq 0$.

Kladná vlastná hodnota a kladný vlastný vektor

Ak $A_{n\times n} > 0$, potom platí:

- $\rho(A) \in \sigma(A)$.
- Ak $Ax = \rho(A)x$, potom $A|x| = \rho(A)|x|$ a $|x| > 0$.

Inými slovami, kladná matica A má (reálnu) vlastnú hodnotu $\rho(A)$, ktorej prislúcha kladný vlastný vektor $v > 0$.

Index $\rho(A)$

Ak $A_{n\times n} > 0$, potom platí:

- $\rho(A)$ je jediná vlastná hodnota matice A na jej spektrálnej kružnici.
- $index(\rho(A)) = 1$, t.j. $\rho(A)$ je polo-jednoduchá vlastná hodnota – jej geometrická násobnosť sa rovná algebraickej, a teda zodpovedajúce Jordanove bloky sú veľkosti 1×1.

Násobnosť $\rho(A)$

Ak $A_{n\times n} > 0$, potom $alg\ mult_A(\rho(A)) = 1$. Inými slovami, spektrálny polomer je jednoduchou vlastnou hodnotou matice A, preto $\dim \mathcal{N}(A-\rho(A)I) = geo\ mult_A(\rho(A)) = alg\ mult_A(\rho(A)) = 1$.

- Význačná vlastná hodnota $r = \rho(A)$ pre $A > 0$ sa nazýva *Perronov koreň* a jednoznačný vlastný vektor $p \in \mathcal{N}(A - \rho(A)I)$ s $p > 0$ a $\sum_j p_j = 1$ sa nazýva *Perronov vektor*.
- Podobne existuje aj *ľavý Perronov vektor* (keďže aj A^T je kladná).

Žiadne ďalšie kladné vlastné vektory

Pre kladnú maticu $A_{n\times n} > 0$ neexistuje iný nezáporný vlastný vektor okrem Perronovho vektora a jeho kladných násobkov

Perronova veta

Ak $A_{n\times n} > 0$ a $r = \rho(A)$, potom platí:

- $r > 0$.
- $r \in \sigma(A)$ (r sa nazýva Perronov koreň).
- $alg\ mult_A(\rho(A)) = 1$.
- Existuje vlastný vektor $x > 0$, pre ktorý $Ax = rx$.
- Perronov vektor je jednoznačne určený vektor, spĺňajúci
$$Ap = rp, \quad p > 0, \quad \text{a} \quad \|p\|_1 = 1.$$
Okrem jeho kladných násobkov neexistujú žiadne nezáporné vlastné vektory pre maticu A, bez ohľadu na vlastnú hodnotu.
- r je jediná vlastná hodnota na spektrálnej kružnici matice A.
- $r = \max_{x\in\mathcal{N}} f(x)$ (Collatzova-Wielandtova formula),
kde $f(x) = \min_{\substack{1\le i\le n\\ x_i\neq 0}} \frac{[Ax]_i}{x_i}$ a $\mathcal{N} = \{x \mid x \geq 0 \text{ a } x \neq 0\}$.

Nezáporné matice (Kapitola 8.3):

- Permutačné matice mávajú viacero rôznych vlastných hodnôt na spektrálnej kružnici $\longrightarrow$ Frobeniova teória.

Stochastické matice a markovovské reťazce (Kapitola 8.4):

- Zložky *stochastickej matice* A označujú pravdepodobnosti prechodov medzi jednotlivými stavmi, t.j. sú nezáporné a v závislosti od konvencie je ich súčet po stĺpcoch, resp. riadkoch rovný 1. Potom $\rho(A) = 1$, jej ľavý Perronov vektor je $p_l^T = (1, 1, \ldots, 1)$ a pravý Perronov vektor p_r predstavuje (v ireducibilnom prípade) stabilný stav.
- Aplikácie: Google matrix a Page rank, Markov Chain Monte Carlo (pozri odkazy z web-stránky).