

13. Jacobijeva metoda za računanje singularnega razcepa

Bor Plestenjak

NLA

17. maj 2011

Enostranska Jacobijeva metoda

Za matriko $A \in \mathbb{R}^{m \times n}$ bi radi izračunali singularni razcep $A = U\Sigma V^T$.

Enostranska Jacobijeva metoda

Za matriko $A \in \mathbb{R}^{m \times n}$ bi radi izračunali singularni razcep $A = U\Sigma V^T$.

Delamo s polno matriko A , ideja pa je podobna kot pri enostranski QR iteraciji.

Mislmo si, da delamo Jacobijevo metodo za računanje lastnih vrednosti matrike $A^T A$, pri čemer spet ne računamo $A^T A$, temveč le posodobimo A .

Enostranska Jacobijeva metoda

Za matriko $A \in \mathbb{R}^{m \times n}$ bi radi izračunali singularni razcep $A = U\Sigma V^T$.

Delamo s polno matriko A , ideja pa je podobna kot pri enostranski QR iteraciji.

Mislimo si, da delamo Jacobijevo metodo za računanje lastnih vrednosti matrike $A^T A$, pri čemer spet ne računamo $A^T A$, temveč le posodobimo A .

En korak Jacobijeve iteracije za $A^T A$ je:

$$\text{iz } B = A^T A \text{ dobimo } \tilde{B} = R_{pq}^T B R_{pq}.$$

Namesto tega pri enostranski Jacobijevi metodi

$$A \text{ posodobimo v } \tilde{A} = A R_{pq},$$

da velja $\tilde{A}^T \tilde{A} = \tilde{B}$.

Algoritem

$A = \text{oneside_jac}(A, p, q)$

izračunaj $b_{pp} = (A^T A)_{pp}$, $b_{pq} = (A^T A)_{pq}$, $b_{qq} = (A^T A)_{qq}$

če velja $|b_{pq}| > \epsilon \sqrt{b_{pp} b_{qq}}$, potem

$$\tau = \frac{b_{pp} - b_{qq}}{2b_{pq}}$$

$$t = \frac{\text{sign}(\tau)}{|\tau| + \sqrt{1 + \tau^2}}, \quad c = \frac{1}{1 + t^2}, \quad s = ct$$

$$A = AR_{pq}(\varphi)$$

$$J = JR_{pq}(\varphi) \text{ (če potrebujemo tudi singularne vektorje).}$$

Algoritem

$A = \text{oneside_jac}(A, p, q)$

izračunaj $b_{pp} = (A^T A)_{pp}$, $b_{pq} = (A^T A)_{pq}$, $b_{qq} = (A^T A)_{qq}$

če velja $|b_{pq}| > \epsilon \sqrt{b_{pp} b_{qq}}$, potem

$$\tau = \frac{b_{pp} - b_{qq}}{2b_{pq}}$$

$$t = \frac{\text{sign}(\tau)}{|\tau| + \sqrt{1 + \tau^2}}, \quad c = \frac{1}{1 + t^2}, \quad s = ct$$

$$A = AR_{pq}(\varphi)$$

$$J = JR_{pq}(\varphi) \text{ (če potrebujemo tudi singularne vektorje).}$$

Pri rotaciji (p, q) se v A spremenita stolpca p in q .

Algoritem

$A = \text{oneside_jac}(A, p, q)$

izračunaj $b_{pp} = (A^T A)_{pp}$, $b_{pq} = (A^T A)_{pq}$, $b_{qq} = (A^T A)_{qq}$

če velja $|b_{pq}| > \epsilon \sqrt{b_{pp} b_{qq}}$, potem

$$\tau = \frac{b_{pp} - b_{qq}}{2b_{pq}}$$

$$t = \frac{\text{sign}(\tau)}{|\tau| + \sqrt{1 + \tau^2}}, \quad c = \frac{1}{1 + t^2}, \quad s = ct$$

$$A = AR_{pq}(\varphi)$$

$$J = JR_{pq}(\varphi) \text{ (če potrebujemo tudi singularne vektorje).}$$

Pri rotaciji (p, q) se v A spremenita stolpca p in q .

V poštev pride le pragovna varianta, saj bi morali pri klasični izračunati vse elemente $A^T A$, da bi lahko poiskali maksimalni element in ustrezno rotacijo.

Algoritem

$A = \text{oneside_jac}(A, p, q)$

izračunaj $b_{pp} = (A^T A)_{pp}$, $b_{pq} = (A^T A)_{pq}$, $b_{qq} = (A^T A)_{qq}$

če velja $|b_{pq}| > \epsilon \sqrt{b_{pp} b_{qq}}$, potem

$$\tau = \frac{b_{pp} - b_{qq}}{2b_{pq}}$$

$$t = \frac{\text{sign}(\tau)}{|\tau| + \sqrt{1 + \tau^2}}, \quad c = \frac{1}{1 + t^2}, \quad s = ct$$

$$A = AR_{pq}(\varphi)$$

$$J = JR_{pq}(\varphi) \text{ (če potrebujemo tudi singularne vektorje).}$$

Pri rotaciji (p, q) se v A spremenita stolpca p in q .

V poštev pride le pragovna varianta, saj bi morali pri klasični izračunati vse elemente $A^T A$, da bi lahko poiskali maksimalni element in ustrezno rotacijo.

Končamo, ko velja $|b_{pq}| \leq \epsilon \sqrt{b_{pp} b_{qq}}$ za vse $p \neq q$.

Kako razberemo singularni razcep

Na koncu matrika A skonvergira proti taki matriki, da je $A^T A$ diagonalna matrika, ki ima na diagonali kvadrate singularnih vrednosti.

Kako razberemo singularni razcep

Na koncu matrika A skonvergira proti taki matriki, da je $A^T A$ diagonalna matrika, ki ima na diagonali kvadrate singularnih vrednosti.

Na koncu dobimo:

- $\sigma_i = \|A(:, i)\|_2$, (to je v bistvu $\sqrt{b_{ii}}$),
- $U = [u_1 \ \cdots \ u_n]$, kjer je $u_i = \frac{1}{\sigma_i} A(:, i)$,
- $V = J$ (če smo shranjevali produkte).

Izrek

Naj bo R točna Givensova rotacija (Householderjevo zrcaljenje) in $\tilde{R}$ njena aproksimacija v plavajoči vejici. Potem je

$$\begin{aligned} fl(\tilde{R}A) &= R(A + E), & \|E\|_2 &= \mathcal{O}(u)\|A\|_2 \\ fl(A\tilde{R}) &= (A + F)R, & \|F\|_2 &= \mathcal{O}(u)\|A\|_2. \end{aligned}$$

Izrek

Naj bo R točna Givensova rotacija (Householderjevo zrcaljenje) in $\tilde{R}$ njena aproksimacija v plavajoči vejici. Potem je

$$\begin{aligned} fl(\tilde{R}A) &= R(A + E), \quad \|E\|_2 = \mathcal{O}(u)\|A\|_2 \\ fl(A\tilde{R}) &= (A + F)R, \quad \|F\|_2 = \mathcal{O}(u)\|A\|_2. \end{aligned}$$

Izrek

Vzemimo zaporedje ortogonalnih transformacij (rotacij ali zrcaljenj) $P_1, \dots, P_j$ in $Q_1, \dots, Q_j$. Za numerično izračunano matriko $A_j = P_j \cdots P_1 A Q_1 \cdots Q_j$ velja

$$\hat{A}_j = P_j \cdots P_1 (A + E) Q_1 \cdots Q_j,$$

kjer je $\|E\|_2 = j\mathcal{O}(u)\|A\|_2$.

Izrek (za lastne vrednosti)

Naj bo A simetrična matrika z lastnimi vrednostmi $\lambda_n \leq \dots \leq \lambda_1$ in naj bodo $\tilde{\lambda}_n \leq \dots \leq \tilde{\lambda}_1$ lastne vrednosti matrike $X^T A X$. Potem za $i = 1, \dots, n$ velja

$$|\tilde{\lambda}_i - \lambda_i| \leq |\lambda_i| \epsilon,$$

kjer je $\epsilon = \|X^T X - I\|_2$.

Relativni Weylov izrek

Izrek (za lastne vrednosti)

Naj bo A simetrična matrika z lastnimi vrednostmi $\lambda_n \leq \dots \leq \lambda_1$ in naj bodo $\tilde{\lambda}_n \leq \dots \leq \tilde{\lambda}_1$ lastne vrednosti matrike $X^T A X$. Potem za $i = 1, \dots, n$ velja

$$|\tilde{\lambda}_i - \lambda_i| \leq |\lambda_i| \epsilon,$$

kjer je $\epsilon = \|X^T X - I\|_2$.

Posledica (za singularne vrednosti)

Naj bodo $\sigma_1 \geq \dots \geq \sigma_n$ singularne vrednosti matrike A in $\tilde{\sigma}_1 \geq \dots \geq \tilde{\sigma}_n$ singularne vrednosti $Y^T A X$. Potem za $i = 1, \dots, n$ velja

$$|\tilde{\sigma}_i - \sigma_i| \leq |\sigma_i| \epsilon,$$

kjer je $\epsilon = \max(\|X^T X - I\|_2, \|Y^T Y - I\|_2)$.

Izrek

Naj bo $A = DX$, kjer je D nesingularna diagonalna in X nesingularna matrika. Če $\hat{A}_m$ dobimo po m korakih enostranske Jacobijeve metode in so $\sigma_1 \geq \dots \geq \sigma_n$ singularne vrednosti A , $\hat{\sigma}_1 \geq \dots \geq \hat{\sigma}_n$ pa singularne vrednosti $\hat{A}_m$, potem velja

$$\frac{|\hat{\sigma}_i - \sigma_i|}{\sigma_i} \leq O(mu)\kappa(X).$$

Izrek

Naj bo $A = DX$, kjer je D nesingularna diagonalna in X nesingularna matrika. Če $\hat{A}_m$ dobimo po m korakih enostranske Jacobijeve metode in so $\sigma_1 \geq \dots \geq \sigma_n$ singularne vrednosti A , $\hat{\sigma}_1 \geq \dots \geq \hat{\sigma}_n$ pa singularne vrednosti $\hat{A}_m$, potem velja

$$\frac{|\hat{\sigma}_i - \sigma_i|}{\sigma_i} \leq O(mu)\kappa(X).$$

Izrek

Če Jacobijevo enostransko metodo zaustavimo, ko za vse $j \neq k$ velja

$$|b_{jk}| \leq \epsilon \sqrt{b_{jj}b_{kk}},$$

kjer je $B = A^T A$, in so $\sigma_1 \geq \dots \geq \sigma_n$ singularne vrednosti A , $\alpha_1^2 \geq \dots \geq \alpha_n^2$ pa diagonalni elementi B , potem je

$$|\sigma_i - \alpha_i| \leq n\epsilon|\alpha_i|.$$

Visoka relativna natančnost

Jacobijeva metoda izračuna singularne vrednosti (in vektorje) z visoko relativno natančnostjo, če lahko zapišemo $A = DX$, kjer je X dobro pogojena matrika, D pa nesingularna diagonalna matrika.

Jacobijeva metoda ponavadi deluje bolje od drugih v primeru, ko ima D elemente, ki se po velikosti bistveno razlikujejo.

Jacobijeva metoda izračuna singularne vrednosti (in vektorje) z visoko relativno natančnostjo, če lahko zapišemo $A = DX$, kjer je X dobro pogojena matrika, D pa nesingularna diagonalna matrika.

Jacobijeva metoda ponavadi deluje bolje od drugih v primeru, ko ima D elemente, ki se po velikosti bistveno razlikujejo.

Za nesingularno bidiagonalno matriko je dqds direktno stabilen algoritem, problem pa je, če imamo na začetku polno matriko. Stabilnost lahko izgubimo pri redukciji na bidiagonalno obliko, kar se pri Jacobijevi metodi, kjer ni začetne redukcije, ne zgodi.

Naj bo

$$A = \begin{bmatrix} \delta & 1 & 1 & 1 \\ \delta & \delta & 0 & 0 \\ \delta & 0 & \delta & 0 \\ \delta & 0 & 0 & \delta \end{bmatrix},$$

kjer je $\delta = 10^{-20}$. Velja

$$A = \begin{bmatrix} 1 & & & \\ & \delta & & \\ & & \delta & \\ & & & \delta \end{bmatrix} \cdot \begin{bmatrix} \delta & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}.$$

Po prvem koraku bidiagonalizacije dobimo

$$\begin{bmatrix} -2\delta & -1/2 - \delta/2 & -1/2 - \delta/2 & -1/2 - \delta/2 \\ 0 & -1/2 + 5\delta/6 & -1/2 - \delta/6 & -1/2 - \delta/6 \\ 0 & -1/2 - \delta/6 & -1/2 + 5\delta/6 & -1/2 - \delta/6 \\ 0 & -1/2 - \delta/6 & -1/2 - \delta/6 & -1/2 + 5\delta/6 \end{bmatrix},$$

Matrika

$$\begin{bmatrix} -2\delta & -1/2 - \delta/2 & -1/2 - \delta/2 & -1/2 - \delta/2 \\ 0 & -1/2 + 5\delta/6 & -1/2 - \delta/6 & -1/2 - \delta/6 \\ 0 & -1/2 - \delta/6 & -1/2 + 5\delta/6 & -1/2 - \delta/6 \\ 0 & -1/2 - \delta/6 & -1/2 - \delta/6 & -1/2 + 5\delta/6 \end{bmatrix},$$

se zaokroži v

$$\begin{bmatrix} -2\delta & -1/2 & -1/2 & -1/2 \\ 0 & -1/2 & -1/2 & -1/2 \\ 0 & -1/2 & -1/2 & -1/2 \\ 0 & -1/2 & -1/2 & -1/2 \end{bmatrix}$$

in v naslednjem koraku dobimo

$$\begin{bmatrix} -2\delta & \sqrt{3}/2 & 0 & 0 \\ 0 & 3/2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Singularne vrednosti dobljene bidiagonalne matrike so $\sqrt{3}, \sqrt{3}\delta, 0, 0$, točne singularne vrednosti pa so $\sqrt{3}, \sqrt{3}\delta, \delta, \delta$.

Računanje lastnih vrednosti s.p.d. matrike

Če je matrika A s.p.d., lahko lastni razcep izračunamo na naslednji način:

a) izračunamo razcep Choleskega $A = LL^T$,

Računanje lastnih vrednosti s.p.d. matrike

Če je matrika A s.p.d., lahko lastni razcep izračunamo na naslednji način:

- a) izračunamo razcep Choleskega $A = LL^T$,
- b) izračunamo singularni razcep $L = U\Sigma V^T$,

Računanje lastnih vrednosti s.p.d. matrike

Če je matrika A s.p.d., lahko lastni razcep izračunamo na naslednji način:

- a) izračunamo razcep Choleskega $A = LL^T$,
- b) izračunamo singularni razcep $L = U\Sigma V^T$,
- c) $A = U\Sigma^2 U^T$.

Računanje lastnih vrednosti s.p.d. matrike

Če je matrika A s.p.d., lahko lastni razcep izračunamo na naslednji način:

- a) izračunamo razcep Choleskega $A = LL^T$,
- b) izračunamo singularni razcep $L = U\Sigma V^T$,
- c) $A = U\Sigma^2 U^T$.

Če za b) uporabimo enostransko Jacobijevo metodo in je $L = DX$, kjer je D diagonalna matrika in X dobro pogojena, imajo izračunane singularne vrednosti relativno napako omejeno z $\mathcal{O}(u)\kappa(X)$.

Računanje lastnih vrednosti s.p.d. matrike

Če je matrika A s.p.d., lahko lastni razcep izračunamo na naslednji način:

- a) izračunamo razcep Choleskega $A = LL^T$,
- b) izračunamo singularni razcep $L = U\Sigma V^T$,
- c) $A = U\Sigma^2 U^T$.

Če za b) uporabimo enostransko Jacobijevo metodo in je $L = DX$, kjer je D diagonalna matrika in X dobro pogojena, imajo izračunane singularne vrednosti relativno napako omejeno z $\mathcal{O}(u)\kappa(X)$.

Če upoštevamo še obratno stabilnost razcepa Choleskega v točki a), celotni postopek vrne singularne vrednosti z relativno napako $\mathcal{O}(u)\kappa(X)^2$.

Če je X dobro pogojena, lahko vse lastne vrednosti s.p.d. matrike izračunamo z visoko relativno natančnostjo.

Vzemimo

$$A = \begin{bmatrix} 1 & \sqrt{\delta} & \sqrt{\delta} \\ \sqrt{\delta} & 1 & 10\delta \\ \sqrt{\delta} & 10\delta & 100\delta \end{bmatrix},$$

kjer je $\delta = 10^{-20}$. Če A reduciramo na tridiagonalno matriko, je točen rezultat

$$T = \begin{bmatrix} 1 & \sqrt{2\delta} & \\ \sqrt{2\delta} & 1/2 + 60\delta & 1/2 - 50\delta \\ & 1/2 - 50\delta & 1/2 + 40\delta \end{bmatrix},$$

pri numeričnem računanju v dvojni natančnosti pa dobimo

$$\tilde{T} = \begin{bmatrix} 1 & \sqrt{2\delta} & \\ \sqrt{2\delta} & 1/2 & 1/2 \\ & 1/2 & 1/2 \end{bmatrix},$$

ki ni s.p.d., zato smo izgubili relativno točnost najmanjše lastne vrednosti.

Razcep Choleskega in nato enostranski Jacobi izračunata vse lastne vrednosti s polno natančnostjo: $\lambda_1 = 1 + \sqrt{\delta}$, $\lambda_2 = 1 - \sqrt{\delta}$, $\lambda_3 = 99\delta$.