

Singularni razcep - 2. del

5. Primeri singularnega razcepa

Primer

Izračunajmo singularni razcep matrike

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

Rešitev: Najprej izračunamo lastne vrednosti in lastne vektorje matrike

$$A^T A = \begin{bmatrix} 2 & 1 & -2 & -1 \\ 1 & 2 & -1 & -2 \\ -2 & -1 & 2 & 1 \\ -1 & -2 & 1 & 2 \end{bmatrix}$$

Lastni polinom matrike A je $x^2(x-2)(x-6)$. Neničelni singularni vrednosti matrike A sta $\sigma_1 = \sqrt{6}$ in $\sigma_2 = \sqrt{2}$.

Lastni vektorji, ki pripadajo lastnim vrednostim 6, 2 in 0 so

$$w_1 = \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix}, w_2 = \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix}, w_3 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, w_4 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}$$

Vektorja w_3 in w_4 tvorita bazo za lastni podprostor lastne vrednosti 0. Na srečo sta že ortogonalna, zato ne rabimo Gram-Schmidtove ortogonalizacije. Odtod dobimo ortonormirano bazo za $\mathbb{R}^4$

$$v_1 = \frac{w_1}{\|w_1\|}, \quad v_2 = \frac{w_2}{\|w_2\|}, \quad v_3 = \frac{w_3}{\|w_3\|}, \quad v_4 = \frac{w_4}{\|w_4\|}.$$

Tvorimo matriki $Q = [v_1 \ v_2 \ v_3 \ v_4]$ in D

$$Q_2 = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{\sqrt{2}} & 0 \\ -\frac{1}{2} & -\frac{1}{2} & 0 & \frac{1}{\sqrt{2}} \\ -\frac{1}{2} & \frac{1}{2} & \frac{1}{\sqrt{2}} & 0 \end{bmatrix}, \quad D = \begin{bmatrix} \sqrt{6} & 0 & 0 & 0 \\ 0 & \sqrt{2} & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Ostane nam še konstrukcija matrike Q_1 . Najprej izračunajmo vektorja

$$u_1 = \frac{1}{\sigma_1} A v_1 = \frac{1}{\sqrt{6}} \begin{bmatrix} 2 \\ -1 \\ -1 \end{bmatrix}, \quad u_2 = \frac{1}{\sigma_2} A v_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}$$

Poiščimo sedaj tak vektor u_3 , ki bo u_1 in u_2 dopolnil do ortonormirane baze za $\mathbb{R}^3$. Običajno se to naredi z Gram-Schmidtom, ampak pogledjmo si še eno metodo. Rešimo enačbo $AA^T u_3 = 0$ in rešitev normirajmo. Dobimo $u_3 = \frac{1}{\sqrt{3}} [1 \ 1 \ 1]^T$. Ker so u_1, u_2 in u_3 lastni vektorji matrike AA^T , ki pripadajo različnim lastnim vrednostim 6, 2 in 0, so paroma ortogonalni. Sedaj lahko tvorimo

$$Q_1 = [u_1 \ u_2 \ u_3] = \begin{bmatrix} \frac{2}{\sqrt{6}} & 0 & \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \end{bmatrix}$$

Konstruirali smo razcep $A = Q_1 D Q_2^T$, kjer sta Q_1 in Q_2 ortogonalni matriki, D pa diagonalna matrika.

Primer.

Pojasni kako preprosto izračunamo singularni razcep normalne matrike.

Rešitev. Naj bo A $n \times n$ matrika, ki zadošča $A^*A = AA^*$.

Najprej izračunamo singularne vrednosti matrike A . Vemo, da ima A n linearno neodvisnih lastnih vektorjev $v_1, \dots, v_n$. Pripadajoče lastne vrednosti označimo z $\lambda_1, \dots, \lambda_n$. Ker za normalne matrike iz $Av_i = \lambda_i v_i$ sledi $A^*v_i = \bar{\lambda}_i v_i$, imamo $A^*Av_i = \bar{\lambda}_i \lambda_i v_i$ za vsak $i = 1, \dots, n$. Torej so $\sigma_i := \sqrt{\bar{\lambda}_i \lambda_i} = |\lambda_i|$ vse singularne vrednosti A .

Permutirajmo vektorje $v_1, \dots, v_n$ tako, da velja

$$|\lambda_1| \geq \dots \geq |\lambda_r| > 0, \quad \lambda_{r+1} = \dots = \lambda_n = 0$$

in označimo

$$Q_2 = \begin{bmatrix} v_1 & \dots & v_n \end{bmatrix}$$

Upoštevamo, da za $i = 1, \dots, r$ velja $\sigma_i = |\lambda_i|$ in dobimo

$$u_i = \frac{1}{\sigma_i} Av_i = \frac{\lambda_i}{|\lambda_i|} v_i$$

Za dopolnitev do baze $\mathbb{C}^n$ lahko vzamemo kar $v_{r+1}, \dots, v_n$. Torej je

$$Q_1 = \begin{bmatrix} \frac{\lambda_1}{|\lambda_1|} v_1 & \dots & \frac{\lambda_r}{|\lambda_r|} v_r & v_{r+1} & \dots & v_n \end{bmatrix}$$

Naj bo $|D|$ matrika, ki ima po diagonalni $|\lambda_1|, \dots, |\lambda_n|$. Singularni razcep je

$$A = Q_1 |D| Q_2^*.$$

Dodatek: naredimo primerjavo singularnega razcepa z lastnim razcepom

$$A = Q_2 D Q_2^*$$

Označimo s $\text{sgn}(D)$ matriko, ki ima po diagonalni $\frac{\lambda_1}{|\lambda_1|}, \dots, \frac{\lambda_r}{|\lambda_r|}, 1, \dots, 1$.

Opazimo, da je $D = \text{sgn}(D) |D|$ in $Q_1 = Q_2 \text{sgn}(D)$, torej je

$$A = Q_2 D Q_2^* = Q_2 \text{sgn}(D) \text{sgn}(D)^* |D| Q_2^* = Q_1 |D| Q_2^*$$

Za pozitivno semidefinitne matrike se oba razcepa ujemata.

Izrek o singularnem razcepu lahko formuliramo tudi za linearne preslikave.

Posledica

Za vsako linearno preslikavo $L: U \rightarrow V$ med končnorazsežnima vektorskima prostoroma s skalarnim produktom obstaja taka ortonormirana baza $\mathcal{B}$ za U in taka ortonormirana baza $\mathcal{C}$ za V , da je matrika $[L]_{\mathcal{C} \leftarrow \mathcal{B}}$ diagonalna z elementi ≥ 0 .

Dokaz: Najprej izberemo poljubni ortonormirani bazi $\mathcal{B}' = \{b'_1, \dots, b'_m\}$ za U in $\mathcal{C}'$ za V . Potem naredimo singularni razcep $[L]_{\mathcal{C}' \leftarrow \mathcal{B}'} = Q_1 D Q_2^{-1}$. Če je $Q_2 = [r_{i,j}]$, potem z $b_i = \sum_{k=1}^n r_{k,i} b'_k$ definiramo novo bazo $\mathcal{B}$ za U . Ker so stolpci matrike Q_2 ortonormirani, je tudi $\mathcal{B}$ ortonormirana baza.

Podobno definiramo tudi novo bazo $\mathcal{C}$ za V in dokažemo ortonormiranost. Velja $[L]_{\mathcal{C} \leftarrow \mathcal{B}} = P_{\mathcal{C} \leftarrow \mathcal{C}'}([L]_{\mathcal{C}' \leftarrow \mathcal{B}'} P_{\mathcal{B}' \leftarrow \mathcal{B}}) = Q_1^{-1}(Q_1 D Q_2^{-1}) Q_2 = D$. $\square$

Opomba: Če ne zahtevamo ortonormiranosti baz $\mathcal{B}$ in $\mathcal{C}$, potem ju lahko izberemo tako, da ima $[L]_{\mathcal{C} \leftarrow \mathcal{B}}$ po diagonalni enke in ničle, drugod pa ničle.

Še ena reformulacija singularnega razcepa.

Posledica

Za vsako matriko $A \in \mathbb{C}^{m \times n}$ ranga r obstaja taka ortonormirana množica $u_1, \dots, u_r \in \mathbb{C}^m$ in taka ortonormirana množica $v_1, \dots, v_r \in \mathbb{C}^n$, da velja

$$A = \sigma_1 u_1 v_1^* + \dots + \sigma_r u_r v_r^*,$$

kjer so $\sigma_1, \dots, \sigma_r$ neničelne singularne vrednosti matrike A .

Dokaz. Sledi iz singularnega razcepa $A = Q_1 D Q_2^*$ in formule

$$\begin{bmatrix} u_1 & \dots & u_r & \dots & u_m \end{bmatrix} \begin{bmatrix} \sigma_1 & \dots & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & & \vdots \\ 0 & \dots & \sigma_r & \dots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & \dots & 0 & \dots & 0 \end{bmatrix} \begin{bmatrix} v_1^* \\ \vdots \\ v_r^* \\ \vdots \\ v_n^* \end{bmatrix} = \sum_{i=1}^r \sigma_i u_i v_i^*$$

6. Aproximacija z matrikami nizkega ranga

Naj bo $\mathbb{C}^{m \times n}$ vektorski prostor vseh kompleksnih $m \times n$ matrik. Na tem vektorskem prostoru lahko definiramo skalarni produkt z

$$\langle B, C \rangle := \text{Sled}(BC^*) = \sum_{i=1}^m \sum_{j=1}^n b_{i,j} \overline{c_{i,j}}.$$

Pripadajoča norma je $\|B\| := \sqrt{\langle B, B \rangle} = \sqrt{\sum_{i=1}^m \sum_{j=1}^n |b_{i,j}|^2}$. Tej normi pravimo **Frobeniusova norma**. Pogosto se označuje z $\|B\|_F$.

Trditev

Če je Q_1 unitarna $m \times m$ matrika in Q_2 unitarna $n \times n$ matrika, potem za vsako $m \times n$ matriko A velja $\|Q_1^* A Q_2\| = \|A\|$.

Dokaz: Upoštevamo definicijo norme in lastnost $\text{Sled}(XY) = \text{Sled}(YX)$. Dobimo $\|Q_1^* A Q_2\|^2 = \text{Sled}(Q_1^* A Q_2 Q_2^* A^* Q_1) = \text{Sled}(Q_1^* A A^* Q_1) = \text{Sled}(A A^* Q_1 Q_1^*) = \text{Sled}(A A^*) = \|A\|^2$. □

Problem: Naj bo $A \in \mathbb{C}^{m \times n}$ dana matrika in $k \in \mathbb{N}$ dano število. Iščemo tako matriko A_k ranga $\leq k$, da za vsako drugo matriko B ranga $\leq k$ velja

$$\|A - A_k\| \leq \|A - B\|.$$

To pomeni, da se med vsemi matrikami v $\mathbb{C}^{m \times n}$, ki so ranga $\leq k$, matrika A_k najboljše prilega matriki A . Odgovor nam daje naslednji izrek.

Izrek (Eckart & Young 1936)

Naj bo $A \in \mathbb{C}^{m \times n}$ matrika ranga r in naj bo $k < r$. Naj bo $A = Q_1 D Q_2^*$ singularni razcep matrike A , kjer $\sigma_1 \geq \dots \geq \sigma_r > 0$. Definirajmo matriko D_k tako, da v D zamenjamo $\sigma_{k+1}, \dots, \sigma_r$ z nič. Med vsemi matrikami v $\mathbb{C}^{m \times n}$, ki so ranga $\leq k$, se $A_k := Q_1 D_k Q_2^*$ najboljše prilega A .

Dokaz: Najprej izračunajmo razdaljo med A in A_k :

$$\|A - A_k\|^2 = \|Q_1(D - D_k)Q_2^*\|^2 = \|D - D_k\|^2 = \sigma_{k+1}^2 + \dots + \sigma_r^2.$$

Naj bo B $m \times n$ matrika ranga $\leq k$, ki se najboljše prilega matriki A . Radi bi dokazali, da je $\|A - B\|^2 = \sigma_{k+1}^2 + \dots + \sigma_r^2$.

Začnimo tokrat s singularnim razcepom matrike B :

$$B = U \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} V^*,$$

kjer je D diagonalna $k \times k$ matrika in U, V unitarni matriki. Pišimo

$$A = U(U^*AV)V^* = U \begin{bmatrix} E & F \\ G & H \end{bmatrix} V^*,$$

kjer je E $k \times k$ matrika. Dokažimo, da je $F = G = 0$ in $E = D$. Naj bo

$$C = U \begin{bmatrix} E & F \\ 0 & 0 \end{bmatrix} V^* \quad \text{in} \quad C' = U \begin{bmatrix} E & 0 \\ G & 0 \end{bmatrix} V^*.$$

Ker sta $\begin{bmatrix} E & F \\ 0 & 0 \end{bmatrix}$ in $\begin{bmatrix} E & 0 \\ G & 0 \end{bmatrix}$ ranga $\leq k$, sta tudi C in C' ranga $\leq k$.

Iz definicije norme in lastnosti $\|UXV^*\| = \|X\|$ sledi

$$\|A - B\|^2 = \left\| \begin{bmatrix} E - D & F \\ G & H \end{bmatrix} \right\|^2 = \|E - D\|^2 + \|F\|^2 + \|G\|^2 + \|H\|^2$$

$$\|A - C\|^2 = \left\| \begin{bmatrix} 0 & 0 \\ G & H \end{bmatrix} \right\|^2 = \|G\|^2 + \|H\|^2$$

$$\|A - C'\|^2 = \left\| \begin{bmatrix} 0 & F \\ 0 & H \end{bmatrix} \right\|^2 = \|F\|^2 + \|H\|^2$$

Odtod sledi, da je

$$\begin{aligned} \|A - B\|^2 &= \|A - C\|^2 + \|E - D\|^2 + \|F\|^2 \\ &= \|A - C'\|^2 + \|E - D\|^2 + \|G\|^2 \end{aligned}$$

Ker se matrika B med vsemi matrikami ranga $\leq k$ najbolj prilega k , velja

$$\|A - B\|^2 \leq \|A - C\|^2 \quad \text{in} \quad \|A - B\|^2 \leq \|A - C'\|^2.$$

Odtod sledi, da je

$$\|E - D\|^2 + \|F\|^2 = 0 \quad \text{in} \quad \|E - D\|^2 + \|G\|^2 = 0.$$

Torej je $\|E - D\| = \|F\| = \|G\| = 0$; se pravi $E = D$ in $F = G = 0$.

Naj bo $H = WJZ^*$ singularni razcep matrike H . Odtod sledi, da je

$$A = U \begin{bmatrix} D & 0 \\ 0 & H \end{bmatrix} V^* = (U \begin{bmatrix} I & 0 \\ 0 & W \end{bmatrix}) \begin{bmatrix} D & 0 \\ 0 & J \end{bmatrix} (V \begin{bmatrix} I & 0 \\ 0 & Z \end{bmatrix})^* \quad (1)$$

Singularni razcep matrike B popravimo v

$$B = U \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} V^* = (U \begin{bmatrix} I & 0 \\ 0 & W \end{bmatrix}) \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} (V \begin{bmatrix} I & 0 \\ 0 & Z \end{bmatrix})^*. \quad (2)$$

Ker se B najboljše prilega A , so vsi elementi J pod vsemi elementi D .

Torej je (1) singularni razcep matrike A . Iz (1) in (2) sedaj sledi

$$\|A - B\|^2 = \sigma_{k+1}^2 + \dots + \sigma_r^2. \quad \square$$

To se uporablja pri zgoščevanju slik. Recimo, da imamo sliko velikosti $m \times n$. Za vsako točko si moramo zapomniti njeno barvo, se pravi neko realno število. (Odtенок sive ali kombinacija odtenkov modre, zelene in rdeče.) Dobimo torej realno $m \times n$ matriko A .

Naredimo singularni razcep in si zapomnimo prvih k singularnih vrednosti, ter vektorje $u_1, \dots, u_k$ in $v_1, \dots, v_k$ za katere velja

$$A_k = \sigma_1 u_1 v_1^* + \dots + \sigma_k u_k v_k^*$$

Vektorji u_i so dolžine m , vektorji v_i pa dolžine n , torej je A_k podana z $k(m + n + 1)$ podatki. Razmerje z začetno količino podatkov je

$$\frac{k(m + n + 1)}{mn} = k\left(\frac{1}{m} + \frac{1}{n} + \frac{1}{mn}\right),$$

kjer lahko zadnji člen zanemarimo. Če je recimo $m = n = 1000$ in $k = 50$, si moramo zapomniti samo 10% od začetne količine podatkov.

7. Pseudoinverz

Kot ponavadi, se omejimo na realna in kompleksna števila.

Za diagonalno $m \times n$ matriko D z neničelnimi elementi $\sigma_1, \dots, \sigma_r$ definiramo njen pseudoinverz D^+ kot diagonalno $n \times m$ matriko z neničelnimi elementi $\frac{1}{\sigma_1}, \dots, \frac{1}{\sigma_r}$.

Primer

$$\begin{bmatrix} \sigma_1 & 0 & 0 & 0 \\ 0 & \sigma_2 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}^+ = \begin{bmatrix} \frac{1}{\sigma_1} & 0 & 0 \\ 0 & \frac{1}{\sigma_2} & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Opomba: Očitno velja $(D^+)^+ = D$ za vsako diagonalno matriko D .

Opomba: Če je D obrnljiva, potem velja $D^+ = D^{-1}$.

Za splošno matriko A najprej izračunamo singularni razcep

$$A = Q_1 D Q_2^*$$

in potem definiramo

$$A^+ = Q_2 D^+ Q_1^*.$$

Opomba: Za vsako matriko A je $(A^+)^+ = Q_1 (D^+)^+ Q_2^* = Q_1 D Q_2^* = A$.

Opomba: Če je matrika A obrnljiva, potem velja $A^+ = A^{-1}$.

Primer

Izračunajmo psevdoinverz matrike

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

Rešitev: Najprej izračunamo singularni razcep matrike A .

$$A = \begin{bmatrix} \frac{2}{\sqrt{6}} & 0 & \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \end{bmatrix} \begin{bmatrix} \sqrt{6} & 0 & 0 & 0 \\ 0 & \sqrt{2} & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{\sqrt{2}} & 0 \\ -\frac{1}{2} & -\frac{1}{2} & 0 & \frac{1}{\sqrt{2}} \\ -\frac{1}{2} & \frac{1}{2} & \frac{1}{\sqrt{2}} & 0 \end{bmatrix}^T$$

Potem je

$$A^+ = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{\sqrt{2}} & 0 \\ -\frac{1}{2} & -\frac{1}{2} & 0 & \frac{1}{\sqrt{2}} \\ -\frac{1}{2} & \frac{1}{2} & \frac{1}{\sqrt{2}} & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{6}} & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{2}{\sqrt{6}} & 0 & \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \end{bmatrix}^T$$

$$= \frac{1}{6} \begin{bmatrix} 1 & -2 & 1 \\ 1 & 1 & -2 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix}$$

Oglejmo si, kako izračunamo psevdoinverz brez singularnega razcepa. Za pozitivno semidefinitno matriko B vemo, da se njen singularni razcep ujema z njenim lastnim razcepom $B = QDQ^*$, kjer je Q unitarna in D diagonalna z nenegativnimi elementi. Torej je $B^+ = QD^+Q^*$. Splošni primer lahko prevedemo na pozitivno semidefinitni primer takole:

Trditev

Za vsako kompleksno matriko A velja

$$A^+ = (A^*A)^+A^* = A^*(AA^*)^+$$

Dokaz: Če je $A = Q_1DQ_2^*$ singularni razcep, potem sta tudi $A^*A = Q_2D^*DQ_2^*$ in $AA^* = Q_1DD^*Q_1^*$ singularna razcepa, torej je $A^* = Q_2D^*Q_1^*$, $(A^*A)^+ = Q_2(D^*D)^+Q_2^*$ in $(AA^*)^+ = Q_1(DD^*)^+Q_1^*$. Trditev sedaj sledi iz $D^+ = (D^*D)^+D^* = D^*(DD^*)^+$. □

Posledica

- Če so stolpci A linearno neodvisni, je $A^+ = (A^*A)^{-1}A^*$.
- Če so vrstice A linearno neodvisne je $A^+ = A^*(AA^*)^{-1}$.

Ker singularni razcep ni enolično določen z matriko A , se zastavi vprašanje ali morda tudi psevdoinverz A^+ ni enolično določen z matriko A .

Izrek o enoličnosti psevdoinverza

Matrika A^+ je enolično določena z matriko A .

Dokaz: Dokažimo najprej nekaj identitet, ki jim zadošča matrika A^+

$$\begin{aligned}AA^+A &= Q_1 D Q_2^* Q_2 D^+ Q_1^* Q_1 D Q_2^* = Q_1 D D^+ D Q_2^* = Q_1 D Q_2^* = A, \\A^+ A A^+ &= Q_2 D^+ Q_1^* Q_1 D Q_2^* Q_2 D^+ Q_1^* = Q_2 D^+ D D^+ Q_1^* = Q_2 D^+ Q_1^* = A^+, \\(AA^+)^* &= (Q_1 D Q_2^* Q_2 D^+ Q_1^*)^* = Q_1 (D D^+)^* Q_1^* = Q_1 D D^+ Q_1^* = AA^+, \\(A^+ A)^* &= (Q_2 D^+ Q_1^* Q_1 D Q_2^*)^* = Q_2 (D^+ D)^* Q_2^* = Q_2 D^+ D Q_2^* = A^+ A\end{aligned}$$

Če bi z dvema singularnima razcepoma dobili različna psevdoinverza B in C , potem bi oba zadoščala tem štirim identitetam. Odtod sledi

$$AB \stackrel{(1)}{=} ACAB \stackrel{(3)}{=} (AC)^*(AB)^* = C^*(ABA)^* \stackrel{(1)}{=} C^*A^* = (AC)^* \stackrel{(3)}{=} AC.$$

Podobno dokažemo, da je tudi $BA = CA$. Oboje uporabimo v računu

$$B \stackrel{(2)}{=} BAB = BAC = CAC \stackrel{(2)}{=} C. \quad \square$$

8. Najkrajša posplošena rešitev sistema linearnih enačb

Psevdo inverz se rabi pri iskanju rešitev sistemov linearnih enačb.

Naj bo A pravokotna matrika. Spomnimo se definicije posplošene rešitve sistema $Ax = b$. To je tak vektor x_0 , da je Ax_0 najbližje vektorju b .

Posplošena rešitev ni enolično določena. Med vsemi posplošenimi rešitvami sistema $Ax = b$ poiščimo najkrajšo. Izkazuje se, da je to ravno A^+b .

Izrek

Naj bo A kompleksna $m \times n$ matrika in b vektor iz $\mathbb{C}^m$. Najkrajša posplošena rešitev sistema $Ax = b$ je $x_0 = A^+b$.

Dokaz: Naj bo $A = Q_1 D Q_2^*$ singularni razcep. Označimo $b' = Q_1^* b$. Primerjajmo posplošene rešitve sistema $Ax = b$ s posplošenimi rešitvami sistema $Dx' = b'$. Za vsak vektor $x \in \mathbb{C}^m$ velja

$$\|Ax - b\| = \|Q_1 D Q_2^* x - b\| = \|Q_1 (D Q_2^* x - b')\| = \|D Q_2^* x - b'\|$$

in $\|Q_2^* x\| = \|x\|$. Odtod očitno sledi dvoje:

- Če je x_0 posplošena rešitev sistema $Ax = b$, potem je $x'_0 := Q_2^* x_0$ posplošena rešitev sistema $Dx' = b'$. Velja tudi $\|x'_0\| = \|x_0\|$.
- Če je x'_0 posplošena rešitev sistema $Dx' = b'$, potem je $x_0 := Q_2 x'_0$ posplošena rešitev sistema $Ax = b$. Velja tudi $\|x_0\| = \|x'_0\|$.

Torej je x_0 najkrajša posplošena rešitev sistema $Ax = b$ natanko tedaj, ko je $x'_0 = Q_2^* x_0$ najkrajša posplošena rešitev sistema $Dx' = b'$.

Poiščimo sedaj najkrajšo posplošeno rešitev sistema $Dx' = b'$, se pravi najkrajšega izmed vseh vektorjev x' , ki minimizirajo izraz $\|Dx' - b'\|$.

Če vstavimo $x' = (x'_1, \dots, x'_n)$ in $b' = (b'_1, \dots, b'_m)$ dobimo

$$\|Dx' - b'\|^2 = (\sigma_1 x'_1 - b'_1)^2 + \dots + (\sigma_r x'_r - b'_r)^2 + (b'_{r+1})^2 + \dots + (b'_m)^2.$$

Minimum tega izraza dobimo pri

$$x' = \left(\frac{b'_1}{\sigma_1}, \dots, \frac{b'_r}{\sigma_r}, x'_{r+1}, \dots, x'_n \right),$$

kjer so $x_{r+1}, \dots, x_n$ poljubni. Najkrajša od teh posplošenih rešitev je

$$x'_0 = \left(\frac{b'_1}{\sigma_1}, \dots, \frac{b'_r}{\sigma_r}, 0, \dots, 0 \right).$$

Opazimo, da je izraz na desni enak $D^+ b'$. Odtod sledi, da je

$$x_0 := Q_2 x'_0 = Q_2 D^+ b' = Q_2 D^+ Q_1^* b = A^+ b$$

najkrajša posplošena rešitev sistema $Ax = b$.

Primer

Sistem linearnih enačb

$$x_1 + x_2 - x_3 - x_4 = 1$$

$$-x_1 + x_3 = 0$$

$$-x_2 + x_4 = 1$$

ni rešljiv običajnem smislu. Poiščimo najkrajšo posplošeno rešitev.

Rešitev: Sistem najprej zapišemo v matrični obliki

$$\begin{bmatrix} 1 & 1 & -1 & -1 \\ -1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

Najkrajša posplošena rešitev je

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 & 1 & -1 & -1 \\ -1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 \end{bmatrix}^+ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} =$$
$$= \frac{1}{6} \begin{bmatrix} 1 & -2 & 1 \\ 1 & 1 & -2 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{3} \\ -\frac{1}{6} \\ -\frac{1}{3} \\ \frac{1}{6} \end{bmatrix}$$