Institutionen för matematik

SF1681
Linjär algebra, fk
HT21

MATERIAL OM SINGULÄRVÄRDESUPPDELNING OCH PSEUDOINVERSER

1. SINGULÄRVÄRDESUPPDELNING

1.1. **Normalform för linjär avbildning.** När vi arbetar med egenvärden, egenvektorer och diagonalisering handlar det om att välja en bas så att matrisen för en operator L blir på så enkel form som möjligt. Om vi ser på linjära avbildningar $L\colon V \longrightarrow W$ har vi större frihet i och med att vi kan göra valen av bas för V och W oberoende av varandra. Det visar sig då att det alltid går att få matrisen på en mycket enkel form med bara ettor och nollor.

Sats 1.1. *Om $L\colon V \longrightarrow W$ är en linjär avbildning mellan ändligdimensionella vektorrum kan vi välja baser för V och W så att matrisen för L blir en blockmatris på formen*

$$\begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}$$

där I är en identitetsmatris av storlek $r \times r$ där $r = \operatorname{rk} L$.

Bevis. Låt $V'' = \ker L$ och välj ett komplement V' till V'', dvs $V = V' \oplus V''$. Detta kan (till skillnad från ett ortogonalt komplement) göras på många olika sätt. Konkret kan vi börja med en bas $\mathbf{x}_{r+1}, \mathbf{x}_{r+2}, \dots, \mathbf{x}_n$ för V'', sedan utvidga till en bas $\mathscr{B} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$ för hela V och slutligen ta $V' = \operatorname{Span}\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_r\}$.

Låt $W' = \operatorname{im} L$ och välj ett komplement W'' till W', dvs $W = W' \oplus W''$. Återigen kan vi välja ett komplement på många olika vis.

På grund av definitionen av V' är restriktionen av L till V' en isomorfi mellan V' och W'. Vi skriver $L' = L|_{V'}\colon V' \longrightarrow W'$ där alltså $L'(\mathbf{v}) = L(\mathbf{v})$ för alla $\mathbf{v} \in V'$.

Om vi nu låter $\mathbf{y}_i = L(\mathbf{x}_i)$ för $i = 1, 2, \dots, r$ så får vi en bas för W' sådan att matrisen för L' blir identitetsmatrisen. Slutligen utökar vi som tidigare till en bas $\mathscr{B}' = \{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_m\}$ för hela W där alltså $W'' = \operatorname{Span}\{\mathbf{y}_{r+1}, \dots, \mathbf{y}_m\}$.

Matrisen för L blir då på den önskade formen med avseende på baserna $\mathscr{B}$ och $\mathscr{B}'$. $\square$

Notera 1.2. Satsen kan omformuleras som: det finns baser $\mathscr{B} = \{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ och $\mathscr{B}' = \{\mathbf{y}_1, \dots, \mathbf{y}_m\}$ för V och W så att

$$L(\mathbf{x}_i) = \begin{cases} \mathbf{y}_i & \text{om } 1 \le i \le r \\ \mathbf{0} & \text{om } r+1 \le i \le n \end{cases}$$

Datum: 30 november 2021.

1.2. **Adjungerad avbildning.** Tidigare har vi sett på adjungerade *operatorer*. Vi kan göra motsvarande även för linjära avbildningar mellan olika vektorrum. För en linjär avbildning $L\colon V \longrightarrow W$ kan vi definiera en adjungerad avbildning $L^\dagger\colon W \longrightarrow V$. Liksom tidigare existerar $L^\dagger$ inte alltid i det oändligt-dimensionella fallet.

Definition 1.3. Om $L\colon V \longrightarrow W$ är en linjär avbildning mellan inre produktrum ges den *adjungerade avbildningen*, $L^\dagger\colon W \longrightarrow V$ av

$$\langle L^\dagger(\mathbf{y})|\mathbf{x}\rangle = \langle \mathbf{y}|L(\mathbf{x})\rangle, \qquad \forall \mathbf{x} \in V, \forall \mathbf{y} \in W.$$

Notera 1.4. Av definitionen följer att:

$$\ker L^\dagger = (\operatorname{im} L)^\perp \qquad \text{och} \qquad \ker L = (\operatorname{im} L^\dagger)^\perp$$

Sats 1.5. *Om vi väljer ortonormala baser för V och W och matrisen för L är A ges matrisen för $L^\dagger$ av $A^\dagger$.*

Bevis. Om $\mathscr{B} = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ och $\mathscr{B}' = \{\mathbf{f}_1, \ldots, \mathbf{f}_m\}$ är ON-baser för V och W så är alltså $A = [L]_{\mathscr{B}',\mathscr{B}}$ och

$$A_{ij} = \langle \mathbf{f}_i|L(\mathbf{e}_j)\rangle = \langle L^\dagger(\mathbf{f}_i)|\mathbf{e}_j\rangle = \overline{\langle \mathbf{e}_j|L^\dagger(\mathbf{f}_i)\rangle} = \overline{B_{ji}}$$

där B är matrisen för $L^\dagger$, dvs $B = [L^\dagger]_{\mathscr{B},\mathscr{B}'}$. □

1.3. **Singulärvärdesuppdelning.** Vi såg tidigare att vi kunde välja baserna för V och W oberoende av varandra så att matrisen för en linjär avbildning $L\colon V \longrightarrow W$ blir på en mycket enkel form. Om vi har en avbildning mellan inre-produktrum är det naturligt att bara tillåta oss att välja ortonormala baser för V och W och då har vi inte lika stor frihet. Det visar sig dock att vi kan komma ganska nära.

Sats 1.6. *Om $L\colon V \longrightarrow W$ är en linjär avbildning mellan ändligdimensionella inre produktrum kan vi välja ortonormala baser för V och W så att matrisen för L med avseende på dessa baser blir en blockmatris Σ på formen*

$$\begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix}$$

där D är en positiv reell diagonalmatris.

Notera 1.7. Precis som för diagonalisering är baserna inte unika, men de positiva diagonalelementen, *singulärvärdena*, är unikt bestämda (inklusive deras multiplicitet). Oftast brukar man ordna singulärvärdena i fallande ordning och då blir Σ unikt bestämt. Ibland räknas även nollor på diagonalen av Σ, utanför D, som singulärvärden.

Bevis. Vi skriver $V = V' \oplus V''$ och $W = W' \oplus W''$ där

- $V' = (\ker L)^\perp$
- $V'' = \ker L$
- $W' = \operatorname{im} L$
- $W'' = (\operatorname{im} L)^\perp$

Observera att det *ortogonala* komplementet (som är meningsfullt eftersom vi har inre produktrum) är unikt till skillnad från de komplement vi valde i beviset till Sats 1.1. Som i det beviset får vi en isomorfi $L' = L|_{V'}\colon V' \xrightarrow{\cong} W'$.

Om vi väljer godtyckliga ON-baser för V'' och W'' så återstår det att välja ON-baser för V' och W' så att $[L'] = D$. Observera att vi inte kan välja en ON-bas $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_r$ för V' och sedan sätta $\mathbf{y}_i = L(\mathbf{x}_i)$ för $i = 1, 2, \ldots, r$. Detta skulle ge en bas för W' men inte nödvändigtvis en ON-bas (såvida inte L' är unitär). Genom att betrakta $L'\colon V' \longrightarrow W'$ istället för $L\colon V \longrightarrow W$ har vi alltså reducerat till fallet då L är inverterbar.

Bilda nu operatorn $H\colon V\oplus W \longrightarrow V\oplus W$ genom $H(\mathbf{x},\mathbf{y}) = \big(L^\dagger(\mathbf{y}), L(\mathbf{x})\big)$. Med avseende på ortonormala baser för V och W får H den Hermiteska matrisen[1]

$$B = \begin{bmatrix} 0 & A^\dagger \\ A & 0 \end{bmatrix}$$

om A är matrisen för L. Alltså är H självadjungerad och eftersom H kommuterar med H^2 kan vi därför välja en gemensam ortogonal bas av egenvektorer. Dessa är på formen $\begin{bmatrix} \mathbf{x}_i \\ \mathbf{y}_i \end{bmatrix}$ med

$$\begin{bmatrix} 0 & A^\dagger \\ A & 0 \end{bmatrix}\begin{bmatrix} \mathbf{x}_i \\ \mathbf{y}_i \end{bmatrix} = \lambda_i \begin{bmatrix} \mathbf{x}_i \\ \mathbf{y}_i \end{bmatrix} \quad \text{och} \quad \begin{bmatrix} A^\dagger A & 0 \\ 0 & AA^\dagger \end{bmatrix}\begin{bmatrix} \mathbf{x}_i \\ \mathbf{y}_i \end{bmatrix} = \mu_i \begin{bmatrix} \mathbf{x}_i \\ \mathbf{y}_i \end{bmatrix}$$

det vill säga

$$\left\{ \begin{array}{rcl} A^\dagger\mathbf{y}_i & = & \lambda_i\mathbf{x}_i \\ A\mathbf{x}_i & = & \lambda_i\mathbf{y}_i \end{array} \right. \quad \text{och} \quad \left\{ \begin{array}{rcl} A^\dagger A\mathbf{x}_i & = & \mu_i\mathbf{x}_i \\ AA^\dagger\mathbf{y}_i & = & \mu_i\mathbf{y}_i \end{array} \right.$$

och därmed är $\mu_i = \lambda_i^2$.

Eftersom H är självadjungerad och inverterbar är λ_i (och μ_i) reella och noll-skilda. Det visar sig att hälften av de $2n$ egenvärdena är positiva och hälften är negativa: givet en egenvektor med egenvärde λ_i som ovan så är $\begin{bmatrix} \mathbf{x}_i \\ -\mathbf{y}_i \end{bmatrix}$ en egenvektor med egenvärde $-\lambda_i$ eftersom

$$\begin{bmatrix} 0 & A^\dagger \\ A & 0 \end{bmatrix}\begin{bmatrix} \mathbf{x}_i \\ -\mathbf{y}_i \end{bmatrix} = \begin{bmatrix} -A^\dagger\mathbf{y}_i \\ A\mathbf{x}_i \end{bmatrix} = \begin{bmatrix} -\lambda_i\mathbf{x}_i \\ \lambda_i\mathbf{y}_i \end{bmatrix} = -\lambda_i \begin{bmatrix} \mathbf{x}_i \\ -\mathbf{y}_i \end{bmatrix}.$$

Vi kan nu anta att egenvärdena är ordnade så att de positiva egenvärdena kommer först och då kommer $\{\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_n\}$ och $\{\mathbf{y}_1,\mathbf{y}_2,\ldots,\mathbf{y}_n\}$ utgöra ortogonala baser för V respektive W. Efter att normerat dessa har vi ortonormala baser och matrisen för L med avseende på dessa baser är en diagonalmatris med positiva diagonalelement $\lambda_1,\lambda_2,\ldots,\lambda_n$. □

Definition 1.8. En *singulärvärdesuppdelning* av den linjära avbildningen $L\colon V \longrightarrow W$ mellan två inreproduktrum V och W är val av ortonormala baser $\mathscr{B} = \{\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_n\}$ och $\mathscr{B}' = \{\mathbf{y}_1,\mathbf{y}_2,\ldots,\mathbf{y}_m\}$ för V och W så att

$$[L]_{\mathscr{B}',\mathscr{B}} = \Sigma = \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix}$$

där $D = \operatorname{diag}(\sigma_1,\sigma_2,\ldots,\sigma_r)$ är en positiv reell diagonalmatris.

Att en singulärvärdesuppdelning existerar är Sats 1.6.

Notera 1.9. I en singulärvärdesuppdelning har vi alltså att

$$L(\mathbf{x}_i) = \begin{cases} \sigma_i\mathbf{y}_i & \text{om } 1 \le i \le r \\ \mathbf{0} & \text{om } r+1 \le i \le n \end{cases}$$

Eftersom $\{\mathbf{x}_i\}$ och $\{\mathbf{y}_i\}$ är ortonormala baser gäller även att

$$L^\dagger(\mathbf{y}_i) = \begin{cases} \sigma_i\mathbf{x}_i & \text{om } 1 \le i \le r \\ \mathbf{0} & \text{om } r+1 \le i \le m \end{cases}$$

[1]Detta svarar mot att ge $V\oplus W$ den inre produkten $\langle(\mathbf{x}_1,\mathbf{y}_1)|(\mathbf{x}_2,\mathbf{y}_2)\rangle = \langle\mathbf{x}_1|\mathbf{x}_2\rangle + \langle\mathbf{y}_1|\mathbf{y}_2\rangle$.

Definition 1.10. Ett *singulärvärde* till $L\colon V \longrightarrow W$ är ett positivt reellt tal $\sigma > 0$ så att det finns nollskilda vektorer $\mathbf{x} \in V$ och $\mathbf{y} \in W$ med

$$L(\mathbf{x}) = \sigma \mathbf{y} \quad \text{och} \quad L^\dagger(\mathbf{y}) = \sigma \mathbf{x}.$$

Vektorn $\mathbf{x}$ kallas en *högersingulär* vektor och $\mathbf{y}$ en *vänstersingulär* vektor.

Notera 1.11. Notera att varje högersingulär vektor $\mathbf{x}$ med singulärvärde σ har en *motsvarande* vänstersingulär vektor $\mathbf{y} = \frac{L(\mathbf{x})}{\sigma}$ och omvänt $\mathbf{x} = \frac{L^\dagger(\mathbf{y})}{\sigma}$.

Notera 1.12. Om $\mathbf{x}$ och $\mathbf{y}$ är höger- och vänstersingulära vektorer till σ så är $|\mathbf{x}| = |\mathbf{y}|$ ty $\sigma^2|\mathbf{y}|^2 = \langle L\mathbf{x}|L\mathbf{x}\rangle = \langle \sigma^2\mathbf{x}|\mathbf{x}\rangle = \sigma^2|\mathbf{x}|^2$. Vi kan alltså skala om $\mathbf{x}$ och $\mathbf{y}$ så att de har norm 1.

Notera 1.13 (Singulära vektorer till singulärvärdet 0). Även 0 betraktas ibland som ett singulärvärde och vi kallar även nollskilda vektorer $\mathbf{x} \in \ker L$ för högersingulära och nollskilda vektorer $\mathbf{y} \in \ker L^\dagger$ för vänstersingulära. Observera dock att dessa inte svarar mot varandra på samma sätt som för positiva singulärvärden.

Notera 1.14. I en singulärvärdesuppdelning för $L\colon V \longrightarrow W$ är alltså $\mathscr{B}$ en ortonormal bas av högersingulära vektorer $\mathbf{x}_i$ och $\mathscr{B}'$ är en ortonormal bas av vänstersingulära vektorer $\mathbf{y}_j$ och $D = \operatorname{diag}(\sigma_1, \ldots, \sigma_r)$ är de nollskilda singulärvärdena. Dessutom korresponderar de första r vänster- och högersingulära vektorerna mot varandra.

I termer av matriser har vi:

Definition 1.15 (Singulärvärdesuppdelning för matriser)**.** Om $L\colon \mathbb{C}^n \longrightarrow \mathbb{C}^m$ är en linjär avbildning med motsvarande $m \times n$-matris A så är en singulärvärdesuppdelning för A en faktorisering

$$A = Y\Sigma X^\dagger$$

där

(a) kolonnerna i X är en ortonormal bas av högersingulära vektorer,
(b) kolonnerna i Y är en ortonormal bas av vänstersingulära vektorerna,
(c) Σ är en $m \times n$-matris som är noll utanför diagonalen och har singulärvärdena i fallande storlek längs diagonalen.
(d) de första $r = \operatorname{rk} L$ kolonnerna i X och Y, dvs som har noll-skilda singulärvärden, korresponderar mot varandra.

Notera 1.16. Faktoriseringen

$$A = Y\Sigma X^\dagger$$

är ett ortonormalt basbyte:

$$[L]_{\mathscr{E}',\mathscr{E}} = P_{\mathscr{E}',\mathscr{B}'}[L]_{\mathscr{B}',\mathscr{B}}P_{\mathscr{B},\mathscr{E}}$$

där $P_{\mathscr{E}',\mathscr{B}'} = Y$, $[L]_{\mathscr{B}',\mathscr{B}} = \Sigma$ och $P_{\mathscr{B},\mathscr{E}} = X^{-1} = X^\dagger$ eftersom X är unitär.

Notera 1.17. Det ortogonala komplementet till kärnan, $\ker(L)^\perp$, har en ortonormal bas av högersingulära vektorer $\mathbf{x}_1, \ldots, \mathbf{x}_r$ som svarar mot nollskilda singulärvärden och bildrummet, $\operatorname{im}(L)$, har en ortonormal bas av vänstersingulära vektorer $\mathbf{y}_1, \ldots, \mathbf{y}_r$ som svarar mot nollskilda singulärvärden.

Genom konstruktionen har vi att matriserna X och Y är *ortogonala* om vi arbetar över de reella talen och *unitära* om vi arbetar över de komplexa talen.

1.4. **Att bestämma singulärvärdesuppdelning.** Låt A vara en komplex $m \times n$-matris. Datorprogram som MatLab bestämmer singulärvärdesuppdelning till A på mer effektiva sätt men vi kan bestämma den på följande vis:

(1) Bestäm egenvärdena $\mu_1 \geq \mu_2 \geq \cdots \geq \mu_n$ till den Hermiteska $n \times n$-matrisen $A^\dagger A$. De första $r = \text{rk} A$ är de (noll-skilda) singulärvärdena i kvadrat $\mu_i = \sigma_i^2$ och de sista $n - r = \dim(\ker A)$ är noll: $\mu_{r+1} = \cdots = \mu_n = 0$.
(2) Bestäm en ortonormal bas $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ av egenvektorer till $A^\dagger A$. Detta är en ortonormal bas av högersingulära vektorer, inklusive de med singulärvärde 0 (dvs de i $\ker A$).
(3) Låt $\mathbf{y}_i = \frac{1}{\sigma_i} A\mathbf{x}_i$ för $i = 1, 2, \ldots, r$. Detta är en ortonormal uppsättning av vänstersingulära vektorer för noll-skilda singulärvärden.
(4) Bestäm en ortonormal bas $\mathbf{y}_{r+1}, \ldots, \mathbf{y}_m$ för $\ker A^\dagger = (\text{im} A)^\perp$.
(5) Då har vi ortonormala baser $\mathbf{x}_1, \ldots, \mathbf{x}_n$ och $\mathbf{y}_1, \ldots, \mathbf{y}_m$ som ger unitära matriser $X = [\mathbf{x}_1 \ldots \mathbf{x}_n]$ och $Y = [\mathbf{y}_1 \ldots \mathbf{y}_m]$ så att vi får singulärvärdesuppdelningen

$$A = Y \Sigma X^\dagger$$

där Σ är en $m \times n$-matris där alla element är noll förutom de r första diagonalelementen som är $\sigma_1 \geq \cdots \geq \sigma_r$.

Vi kan förstås byta ut steg (1) mot att istället ta egenvärden och egenvektorer till $AA^\dagger$. Då får vi vänstersingulära vektorer och kan bestämma de högersingulära genom $\mathbf{x}_i = \frac{1}{\sigma_i} A^\dagger \mathbf{y}_i$ för $i = 1, \ldots, r$ och sedan fylla på med en bas för $\ker A$.

Exempel 1.18 (Tentamen 2018-01-10 #3)**.** Bestäm en singulärvärdesuppdelning till matrisen

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \end{bmatrix}$$

(Se handskrivna anteckningar till föreläsning 11.)

1.5. **Projektion.** En av poängerna med singulärvärdesuppdelning är att vi kan bortse helt från singulärvärden som är noll och motsvarande höger- och vänstersingulära vektorer. Vi kan ta steget vidare och bortse från små singulärvärden. Genom att sortera singulärvärdena i fallande ordning $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$ kan vi se att de första termerna i summan

$$A = Y \Sigma X^\dagger = \sum_{i=1}^{r} \sigma_i \mathbf{y}_i \mathbf{x}_i^\dagger$$

innehåller den *viktigaste* informationen om A. Vi kan approximera A med delsummorna

$$A_s = \sum_{i=1}^{s} \sigma_i \mathbf{y}_i \mathbf{x}_i^\dagger, \qquad s = 1, 2, \ldots, r.$$

Sats 1.19. *A_s är den matris av rang s som bäst approximerar A i minsta kvadratmening.*

Exempel 1.20. Om vi har matrisen

$$A = \begin{bmatrix} 2 & 4 & 6 \\ 3 & 6 & 10 \end{bmatrix}$$

så får vi singulärvärdena till kvadratrötterna till egenvärdena av

$$AA^T = \begin{bmatrix} 56 & 90 \\ 90 & 145 \end{bmatrix}$$

som är $\lambda_1 \approx 200{,}9$ och $\lambda_2 \approx 0{,}09956$. Därmed är $\sigma_1 \approx 14{,}17$ och $\sigma_2 \approx 0{,}316$. Om vi bara tar med det största singulärvärdet får vi approximationen

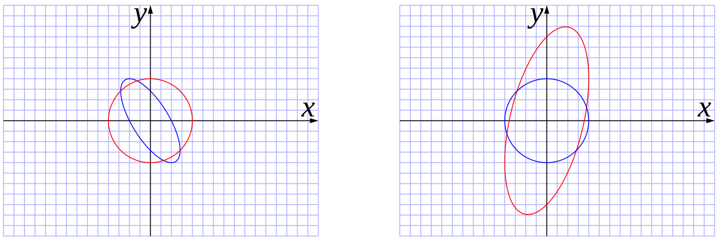
$$A_1 = Y\Sigma_1 X^T = \begin{bmatrix} 1{,}90139 & 3{,}80278 & 6{,}15238 \\ 3{,}06125 & 6{,}1225 & 9{,}90536 \end{bmatrix}$$

Om vi vill spara utrymme och singulärvärdena avtar relativt snabbt kan vi använda en approximation av relativt låg rang. Detta har bland annat använts för bildkompression.

1.6. **Geometrisk innebörd för operatorer på $\mathbb{R}^2$.** Om vi har en inverterbar 2×2-matris A kan vi se på hur enhetscirkeln avbildas genom multiplikation med A. Bilden kommer att bli en ellips och ekvationen för ellipsen ges av $\langle A^{-1}\mathbf{y}|A^{-1}\mathbf{y}\rangle = 1$. Vi kan också se på inversbilden av enhetscirkeln som också blir en ellips. Denna gång blir ekvationen $\langle A\mathbf{x}|A\mathbf{x}\rangle = 1$. Ellipsernas axlar ges av egenvektorer till matriserna $(A^{-1})^T A^{-1}$ respektive $A^T A$. Egenvektorerna till $B = (A^{-1})^T A^{-1}$ är de samma som för dess invers, $B^{-1} = AA^T$.

Om vi till exempel tar matrisen $A = \begin{bmatrix} 0 & 1 \\ 2 & 1 \end{bmatrix}$ får vi bilderna enligt Figur 1 där den röda enhetscirkeln till vänster avbildas på den röda ellipsen till höger och den blå ellipsen till vänster avbildas på den blå enhetscirkeln till höger.

Det ortogonala basbytet som gör ellipsen till vänster axelparallell svarar mot valet av vänstersingulära vektorer och det ortogonala basbytet som gör ellipsen till höger axelparallell svarar mot de högersingulära vektorerna.



FIGUR 1. En bild av hur singulärvärdesuppdelningen ser ut för matrisen en 2×2-matris.

2. PSEUDOINVERS

Det finns flera olika sätt att generalisera begreppet inversmatris till situationer där matriserna inte är inverterbara i vanlig mening. En sådan generalisering ges av *Moore–Penrose-inversen* [Dre20, Pen55].

Definition 2.1 (Moore–Penrose invers)**.** Om A är en komplex $m \times n$-matris ges *Moore–Penrose-inversen* A^+ till A av

$$A^+ = Q \begin{bmatrix} D^{-1} & 0 \\ 0 & 0 \end{bmatrix} P^\dagger, \quad \text{där } A = P \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} Q^\dagger$$

är singulärvärdesuppdelningen av A.

Sats 2.2 (Penrose[Pen55])**.** *A^+ är den unika lösningen till de fyra ekvationerna*

$$AXA = A, \quad XAX = X, \quad (AX)^\dagger = AX \quad \textit{och} \quad (XA)^\dagger = XA.$$

Bevis. Vi använder singulärvärdesuppdelning av A och kan därmed anta att

$$A = \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} \quad \text{och} \quad X = \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix}.$$

Då blir

$$AX = \begin{bmatrix} DX_{11} & DX_{12} \\ 0 & 0 \end{bmatrix} \quad \text{och} \quad XA = \begin{bmatrix} X_{11}D & 0 \\ X_{21}D & 0 \end{bmatrix}$$

och de två sista ekvationerna ger $DX_{12} = 0$, $X_{21}D = 0$ och $D_{11}X = XD_{11}$. Eftersom D är inverterbar är $X_{21} = 0$ och $X_{12} = 0$.

Nu blir

$$AXA = \begin{bmatrix} DX_{11}D & 0 \\ 0 & 0 \end{bmatrix} \quad \text{och} \quad XAX = \begin{bmatrix} X_{11}DX_{11} & 0 \\ 0 & 0 \end{bmatrix}$$

och den första ekvationen ger $DX_{11}D = D$, vilket ger $X_{11} = D^{-1}$ och den andra ger $X_{22} = 0$. Den unika lösningen är därmed

$$X = \begin{bmatrix} D^{-1} & 0 \\ 0 & 0 \end{bmatrix}$$

med avseende på de givna ortonormala baserna för V och W. □

Notera 2.3. I och med att den vanliga inversen har dessa egenskaper sammanfaller Moore–Penrose-inversen med den vanliga inversen när den existerar.

Moore–Penrose-inversen används bland annat vid minsta-kvadratproblem. Minsta kvadratmetoden ger att vi ska välja en lösning till $A^\dagger A\mathbf{x} = A^\dagger \mathbf{b}$. Om $\ker(A) \neq \{0\}$ finns inte en unik lösning och vi kan bestämma oss för att välja den vektor i lösningsmängden som ligger närmast origo, dvs som har minst norm.

Vi har följande sats.

Sats 2.4. *Till ekvationen $A\mathbf{x} = \mathbf{b}$ är $\mathbf{x} = A^+\mathbf{b}$ den minsta-kvadratlösning som har minst norm $|\mathbf{x}|$, dvs den vektor som minimerar $|\mathbf{x}|$ bland de vektorer där $|A\mathbf{x} - \mathbf{b}|$ är minimal.*

Bevisidé. Vi vill hitta en minsta-kvadrat lösning till $AX = I$ som minimerar normen av kolonnerna i X. Vi börjar med att se att $X = A^+$ löser normalekvationen $A^\dagger AX = A^\dagger$ eftersom $AA^+ = (A^+)^\dagger A^\dagger$, vilket ger $A^\dagger (A^+)^\dagger A^\dagger = A^\dagger$ som är ekvivalent med $AA^+A = A$.

Om vi har en annan lösning kan den skrivas $A^+ + Y$ där $A^\dagger AY = 0$. Enligt konstruktionen är kolonnerna i A^+ ortogonala mot kärnan till $A^\dagger A$, vilket gör att normen av kolonnerna i $A^+ + Y$ minimeras så $Y = 0$. □

Läs mer om singulärvärdesuppdelning i exempelvis *Linear Algebra Done Right* [Axl15, §7.D] eller *Numerical Analysis* [Sau14]. Det finns många andra källor på nätet, exempelvis *Wikipedia* [Wik17] och dokumentationen för Matlab [Mat]. Det kan också vara intressant att se på Penroses artikel [Pen55] från 1955.

Referenser

[Axl15] Sheldon Axler. *Linear algebra done right*. Undergraduate Texts in Mathematics. Springer, Cham, third edition, 2015.

[Dre20] Arnold Dresden. The fourteenth western meeting of the american mathematical society. *Bull. Amer. Math. Soc.*, 26(9):385–396, 06 1920.

[Mat] MathWorks. svd — documentation.

[Pen55] R. Penrose. A generalized inverse for matrices. *Proc. Cambridge Philos. Soc.*, 51:406–413, 1955.

[Sau14] Timothy Sauer. *Numerical Analysis*. Pearson, 2014.

[Wik17] Wikipedia. Singular-value decomposition — wikipedia, the free encyclopedia, 2017.

ÖVNINGAR PÅ SINGULÄRVÄRDESUPPDELNING

Uppgift 1. Låt L vara en operator på ett inre produktrum V.

(a) Visa att $L^\dagger L$ och $LL^\dagger$ har samma egenvärden.
(b) Visa att alla egenvärden till $L^\dagger L$ och $LL^\dagger$ är icke-negativa reella tal.
(c) Visa att om V är ändligt-dimensionellt så finns det en självadjungerad operator K_1 så att $(K_1)^2 = L^\dagger L$ och en självadjungerad operator K_2 så att $(K_2)^2 = LL^\dagger$.
(d) Om L i en ortonormal bas $\mathscr{B}$ ges av matrisen

$$A = \begin{bmatrix} 2 & 3 \\ 0 & 2 \end{bmatrix}$$

bestäm matriserna för K_1 och K_2 i samma bas.

Uppgift 2. Bestäm en 2×3-matris A så att A har singulärvärdena 2 och 3 med motsvarande höger- och vänstersingulära vektorer $\frac{1}{3}\begin{bmatrix} 1 & 2 & 2 \end{bmatrix}^T$ och $\frac{1}{\sqrt{5}}\begin{bmatrix} 1 & 2 \end{bmatrix}^T$, respektive $\frac{1}{\sqrt{2}}\begin{bmatrix} 0 & 1 & -1 \end{bmatrix}^T$ och $\frac{1}{\sqrt{5}}\begin{bmatrix} 2 & -1 \end{bmatrix}^T$.

Uppgift 3. Bestäm Moore–Penrose inversen A^+ till matrisen

$$A = \begin{bmatrix} 1 & 2 & 4 \\ 2 & 4 & 8 \end{bmatrix}.$$

Uppgift 4. Använd Moore–Penrose inversen för att lösa ekvationssystemet

$$\begin{bmatrix} 1 & 2 & 4 \\ 2 & 4 & 8 \end{bmatrix} \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \mathbf{x}_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$$

i minsta kvadratmening.

Uppgift 5. Visa att om A är en reell 2×2-matris kan vi alltid välja X eller Y som en symmetrisk matris i singulärvärdesuppdelningen $A = Y\Sigma X^T$. När kan vi välja både X och Y som symmetriska matriser?

Uppgift 6 (Uppgift 3 på tentamen 2019-04-17)**.** Bestäm det största singulärvärdet och motsvarande höger- och vänstersingulära vektorer som hör till matrisen

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

Uppgift 7 (Uppgift 3 på tentamen 2018-04-05)**.** Bestäm singulärvärden och singulärvektorer till matriserna

$$A = \begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix} \quad \text{och} \quad B = \begin{bmatrix} 1 & i \\ -i & 1 \end{bmatrix}.$$

LÖSNINGSFÖRSLAG

Lösningsförslag till uppgift 1.

(a) Låt $\mathbf{x}$ vara en egenvektor till $L^\dagger L$ med egenvärde $\lambda \neq 0$. Då är $L\mathbf{x}$ en egenvektor till $LL^\dagger$ med egenvärde λ eftersom

$$LL^\dagger(L\mathbf{x}) = L\big((L^\dagger L)\mathbf{x}\big) = L(\lambda\mathbf{x}) = \lambda(L\mathbf{x}).$$

Notera att $L\mathbf{x} \neq 0$ eftersom $L^\dagger L\mathbf{x} \neq 0$.

Omvänt, om $\mathbf{y}$ är en egenvektor till $LL^\dagger$ med egenvärde $\lambda \neq 0$ så är $L^\dagger\mathbf{y}$ en egenvektor till $L^\dagger L$ med egenvärde λ eftersom

$$L^\dagger L(L^\dagger\mathbf{y}) = L^\dagger\big((LL^\dagger)\mathbf{y}\big) = L^\dagger(\lambda\mathbf{y}) = \lambda(L^\dagger\mathbf{y}).$$

Egenvärdet 0 förekommer hos L precis när $\operatorname{rk} L \neq \dim V$ och analogt för $L^\dagger$. Eftersom $\operatorname{rk} L = \operatorname{rk} L^\dagger$ så förekommer alltså egenvärdet 0 antingen för både L och $L^\dagger$ eller för ingen av dem.

(b) Låt $\mathbf{x}$ vara en egenvektor till $L^\dagger L$ med egenvärde λ. Då är

$$\lambda|\mathbf{x}| = \langle\mathbf{x}|\lambda\mathbf{x}\rangle = \langle\mathbf{x}|L^\dagger L\mathbf{x}\rangle = \langle L\mathbf{x}|L\mathbf{x}\rangle = |L\mathbf{x}| \geq 0.$$

Alltså är $\lambda \geq 0$.

(c) Eftersom $L^\dagger L$ är självadjungerad så finns en ortonormal bas $\mathscr{B} = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ av egenvektorer, dvs $[L^\dagger L]_{\mathscr{B}} = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$. Vi kan nu låta K_1 vara operatorn med matrisen $\operatorname{diag}(\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_n})$ i basen $\mathscr{B}$, dvs $K_1(\mathbf{x}_i) = \sqrt{\lambda_i}\mathbf{x}_i$ för $i = 1, 2, \ldots, n$. Eftersom $[K_1]_{\mathscr{B}}$ är reell och symmetrisk så är K_1 självadjungerad. Analogt för $LL^\dagger$.

Vi hade kunnat välja $-\sqrt{\lambda_i}$ istället för några eller alla i. Vårt val av K_1 (och K_2) har den ytterligare egenskapen att alla egenvärdena hos K_1 är icke-negativa, eller ekvivalent, att K_1 är *positivt semidefinit*, dvs $\langle K_1\mathbf{x}|\mathbf{x}\rangle \geq 0$ för alla $\mathbf{x} \in V$.

(d) $L^\dagger L$ och $LL^\dagger$ ges av matriserna

$$A^\dagger A = \begin{bmatrix} 4 & 6 \\ 6 & 13 \end{bmatrix}, \quad \text{och} \quad AA^\dagger = \begin{bmatrix} 13 & 6 \\ 6 & 4 \end{bmatrix}.$$

Egenvärdena för båda operatorerna är 1 och 16. En egenvektorbas för $A^\dagger A$ är

$$\mathbf{x}_1 = \begin{bmatrix} 2 \\ -1 \end{bmatrix}, \quad \mathbf{x}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

och en egenvektorbas för $AA^\dagger$ är

$$\mathbf{y}_1 = \begin{bmatrix} 1 \\ -2 \end{bmatrix}, \quad \mathbf{y}_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}.$$

Operatorerna K_1 och K_2 har båda egenvärdena 1 och 4 och egenvektorbaserna $\{\mathbf{x}_1, \mathbf{x}_2\}$ respektive $\{\mathbf{y}_1, \mathbf{y}_2\}$. Detta ger

$$[K_1] = \begin{bmatrix} 2 & 1 \\ -1 & 2 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}\begin{bmatrix} 2 & 1 \\ -1 & 2 \end{bmatrix}^{-1} = \frac{1}{5}\begin{bmatrix} 8 & 6 \\ 6 & 17 \end{bmatrix}$$

$$[K_2] = \begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix}^{-1} = \frac{1}{5}\begin{bmatrix} 17 & 6 \\ 6 & 8 \end{bmatrix}.$$

Lösningsförslag till uppgift 2. För $\mathbb{R}^3$ har vi en ortonormal bas $\mathscr{B}$ bestående av de 2 höger-singulära vektorerna

$$\mathbf{x}_1 = \frac{1}{3}\begin{bmatrix}1\\2\\2\end{bmatrix}, \quad \mathbf{x}_2 = \frac{1}{\sqrt{2}}\begin{bmatrix}0\\1\\-1\end{bmatrix},$$

och en tredje godtycklig vektor $\mathbf{x}_3$ i det ortogonala komplementet till $\mathrm{Span}\{\mathbf{x}_1, \mathbf{x}_2\}$.

För $\mathbb{R}^2$ har vi en ortogonal bas $\mathscr{B}'$ bestående av de 2 vänster-singulär vektorerna.

$$\mathbf{y}_1 = \frac{1}{\sqrt{5}}\begin{bmatrix}1\\2\end{bmatrix}, \quad \mathbf{y}_2 = \frac{1}{\sqrt{5}}\begin{bmatrix}2\\-1\end{bmatrix},$$

Med dessa baser har vi:

$$[A]_{\mathscr{B}',\mathscr{B}} = \begin{bmatrix}2 & 0 & 0\\0 & 3 & 0\end{bmatrix}$$

dvs

$$A = \frac{1}{\sqrt{5}}\begin{bmatrix}1 & 2\\2 & -1\end{bmatrix}\begin{bmatrix}2 & 0 & 0\\0 & 3 & 0\end{bmatrix}\begin{bmatrix}\frac{1}{3} & 0 & *\\ \frac{2}{3} & \frac{1}{\sqrt{2}} & *\\ \frac{2}{3} & \frac{-1}{\sqrt{2}} & *\end{bmatrix}^{-1}$$

Eftersom den tredje matrisen är ortogonal så erhåller vi:

$$A = \frac{1}{\sqrt{5}}\begin{bmatrix}1 & 2\\2 & -1\end{bmatrix}\begin{bmatrix}2 & 0 & 0\\0 & 3 & 0\end{bmatrix}\begin{bmatrix}\frac{1}{3} & \frac{2}{3} & \frac{2}{3}\\ 0 & \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}}\\ * & * & *\end{bmatrix} = \frac{1}{6\sqrt{5}}\begin{bmatrix}4 & 8+18\sqrt{2} & 8-18\sqrt{2}\\8 & 16-9\sqrt{2} & 16+9\sqrt{2}\end{bmatrix}$$

Lösningsförslag till uppgift 3. Vi beräknar singulärvärdena och singulärvektorer till A. Vi utgår från $AA^\dagger$ eftersom detta är en 2×2-matris.

$$AA^\dagger = \begin{bmatrix}21 & 42\\42 & 84\end{bmatrix}$$

Denna har egenvärdena 0 och 105. En normerad egenvektor för egenvärdet 105 är

$$\mathbf{y} = \frac{1}{\sqrt{5}}\begin{bmatrix}1\\2\end{bmatrix}$$

vilket ger det enda singulärvärdet $\sqrt{105}$ med vänstersingulär vektor $\mathbf{y}$ och motsvarande högersingulära vektor

$$\mathbf{x} = \frac{1}{\sqrt{105}}A^\dagger\mathbf{y} = \frac{1}{\sqrt{525}}\begin{bmatrix}5\\10\\20\end{bmatrix} = \frac{1}{\sqrt{21}}\begin{bmatrix}1\\2\\4\end{bmatrix}.$$

Detta ger singulärvärdesuppdelningen

$$A = Q\Sigma P^\dagger = \frac{1}{\sqrt{5}}\begin{bmatrix}1 & *\\2 & *\end{bmatrix}\begin{bmatrix}\sqrt{105} & 0 & 0\\0 & 0 & 0\end{bmatrix}\frac{1}{\sqrt{21}}\begin{bmatrix}1 & 2 & 4\\ * & * & *\end{bmatrix}$$

där $*$ väljs så att vi får ortonormala baser men inte behöver bestämmas för att beräkna Moore–Penrose inversen. Denna får vi som:

$$A^+ = P\Sigma^+Q^\dagger = \frac{1}{\sqrt{21}}\begin{bmatrix}1 & *\\2 & *\\4 & *\end{bmatrix}\begin{bmatrix}\frac{1}{\sqrt{105}} & 0\\0 & 0\\0 & 0\end{bmatrix}\frac{1}{\sqrt{5}}\begin{bmatrix}1 & 2\\ * & *\end{bmatrix} = \frac{1}{108}\begin{bmatrix}1 & 2\\2 & 4\\4 & 8\end{bmatrix}$$

Lösningsförslag till uppgift 4. Vi har att minstakvadratenlösningen med minst norm är:

$$\begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \mathbf{x}_3 \end{bmatrix} = A^+ \begin{bmatrix} 1 \\ 3 \end{bmatrix} = \frac{1}{108} \begin{bmatrix} 1 & 2 \\ 2 & 4 \\ 4 & 8 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \end{bmatrix} = \frac{1}{108} \begin{bmatrix} 7 \\ 14 \\ 28 \end{bmatrix}$$

Lösningsförslag till uppgift 5. Eftersom A är reell så är $A^\dagger A = A^T A$ en reell symmetrisk 2×2-matris och har alltså en ortonormal bas bestående av reella egenvektorer. Detta ger högersingulära vektorer $\mathbf{x}_1, \mathbf{x}_2$ och motsvarande vänstersingulära vektorer $\mathbf{y}_1 = \frac{1}{\sigma_1} L(\mathbf{x}_1), \mathbf{y}_2 = \frac{1}{\sigma_2} L(\mathbf{x}_2)$ så att $X = \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 \end{bmatrix}$ och $Y = \begin{bmatrix} \mathbf{y}_1 & \mathbf{y}_2 \end{bmatrix}$ är reella ortogonala matriser.

Om $\mathbf{x}_1 = \begin{bmatrix} a \\ b \end{bmatrix}$ så finns två möjligheter för $\mathbf{x}_2$: antingen $\begin{bmatrix} b \\ -a \end{bmatrix}$ vilket gör X symmetrisk (determinant -1) eller $\begin{bmatrix} -b \\ a \end{bmatrix}$ vilket gör X anti-symmetrisk (determinant 1). På samma sätt kommer Y antingen vara symmetrisk (determinant -1) eller anti-symmetrisk (determinant 1).

Om $\det(A) > 0$ är $\det(\Sigma) > 0$ och väljer vi X symmetrisk blir även Y symmetrisk. Om $\det(A) < 0$ kommer precis en av X och Y vara symmetrisk och den andra anti-symmetrisk. Om $\det(A) = 0$ har vi högst ett singulärvärde och $\mathbf{x}_2 \in \ker L$ och $\mathbf{y}_2 \in \ker L^\dagger$ och vi kan byta ut $\mathbf{y}_2$ mot $-\mathbf{y}_2$ om det behövs för att göra Y symmetrisk.

Vi kan alltså välja X och Y som symmetriska matriser precis när $\det(A) \geq 0$.

Lösningsförslag till uppgift 6. Singulärvärdena är kvadratrötterna till egenvärdena av $A^T A$ och AA^T. Vi får att

$$A^T A = \begin{bmatrix} 1 & 2 & 2 \\ 2 & 2 & -1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 2 & 2 \\ 2 & -1 \end{bmatrix} = \begin{bmatrix} 9 & 4 \\ 4 & 9 \end{bmatrix}$$

vars karakteristiska ekvation är $t^2 - 18t + 65 = 0$, vilket ger $t = 9 \pm \sqrt{81 - 65} = 9 \pm 4$. Det största egenvärdet är 13. Alltså är det största singulärvärdet $\sigma_1 = \sqrt{13}$. Motsvarande egenvektor till $A^T A$ är $\begin{bmatrix} 1 & 1 \end{bmatrix}^T$. Normerar vi den får vi den höger-singulära vektorn

$$\mathbf{x} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

och när vi multiplicerar med A får vi

$$A\mathbf{x} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 2 \\ 2 & 2 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 3 \\ 4 \\ 1 \end{bmatrix}.$$

När vi normerar denna får vi

$$\mathbf{y} = \frac{1}{\sqrt{26}} \begin{bmatrix} 3 \\ 4 \\ 1 \end{bmatrix}$$

som är motsvarande vänstersingulära vektor.

Lösningsförslag till uppgift 7. Vi beräknar $A^\dagger A$ och $AA^\dagger$ och får

$$A^\dagger A = \begin{bmatrix} 1 & -i \\ -i & 1 \end{bmatrix} \begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \quad \text{och} \quad AA^\dagger = \begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix} \begin{bmatrix} 1 & -i \\ -i & 1 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}.$$

Därmed är båda singulärvärdena $\sqrt{2}$. Vi kan välja standardbasen för de högersingulära vektorerna och får

$$\frac{1}{\sqrt{2}}\begin{bmatrix}1\\ i\end{bmatrix} \quad \text{och} \quad \frac{1}{\sqrt{2}}\begin{bmatrix}i\\ 1\end{bmatrix}$$

som motsvarande vänstersingulära vektorer.

Eftersom B är Hermitesk har den en ortogonal bas av egenvektorer och därmed är dessa egenvektorer singulärvektorer med beloppen av motsvarande egenvärden som singulärvärden. Det karaktäristiska polynomet är $p_B(x) = x^2 - \operatorname{tr}(B)x + \det(B) = x^2 - 2x$. Alltså är egenvärdena 0 och 2 med egenvektorer som ges av

$$\ker\begin{bmatrix}1 & i\\ -i & 1\end{bmatrix} = \operatorname{Span}\left\{\begin{bmatrix}-i\\ 1\end{bmatrix}\right\} \quad \text{respektive} \quad \ker\begin{bmatrix}-1 & i\\ -i & -1\end{bmatrix} = \operatorname{Span}\left\{\begin{bmatrix}i\\ 1\end{bmatrix}\right\}$$