

5.6.3: SVD av lin. avb.

Teorem:

La U og V være endeligdim. indreproduktrom og la $T \in \mathcal{L}(U, V)$. Da finnes basiser $B = (u_1, \dots, u_n)$ og $C = (v_1, \dots, v_m)$ for hhv. U og V slike at $[T]_C^B = \text{diag}_{m \times n}(\sigma_1, \dots, \sigma_r)$, eller ekvivalent,

$$Tu = \sum_{k=1}^r \sigma_k \langle u, u_k \rangle v_k \quad \forall u \in U.$$

Det nevnte er singulærværdidekomposisjonen av T .

Beris:

Om $T \equiv 0$ lar vi B og C være vilkårlige orton. basiser. Ellers har T minst én, og høyst n , positive sing. verd.:

$$\sigma_1 \geq \dots \geq \sigma_r > 0, \quad \sigma_{r+1}, \dots, \sigma_n = 0.$$

Av Spektralteoremet på T^*T , har U en orton. basis $B = (u_1, \dots, u_n)$ slike at

$$T^*T u_k = \sigma_k^2 u_k \quad \forall k = 1, \dots, n.$$

Definer $v_k := \frac{1}{\sigma_k} T u_k$, $k = 1, \dots, r$.

Da er

$$\begin{aligned} \langle v_j, v_k \rangle &= \frac{1}{\sigma_j \sigma_k} \langle T u_j, T u_k \rangle = \frac{1}{\sigma_j \sigma_k} \langle T^* T u_j, u_k \rangle \\ &= \frac{\sigma_j^2}{\sigma_j \sigma_k} \langle u_j, u_k \rangle = \frac{\sigma_j}{\sigma_k} \langle u_j, u_k \rangle = \begin{cases} 0 & j \neq k \\ 1 & j = k, \end{cases} \end{aligned}$$

så $(v_1, \dots, v_r)$ er ortonormal. Om $r < m$ lar vi

$v_{r+1}, \dots, v_m$ være slike at $(v_1, \dots, v_m)$ er ortonormal.

For $u \in U$ får vi

$$\begin{aligned} Tu &= T\left(\sum_{k=1}^n \langle u, u_k \rangle u_k\right) = \sum_{k=1}^n \langle u, u_k \rangle Tu_k \\ &= \sum_{k=1}^n \langle u, u_k \rangle \sigma_k v_k = \sum_{k=1}^r \sigma_k \langle u, u_k \rangle v_k. \end{aligned}$$

Dermed er $[Tu]_{\mathcal{C}} = \text{diag}_{m \times n}(\sigma_1, \dots, \sigma_r) [u]_{\mathcal{B}}$. ▀

Algoritme

La $T \in \mathcal{L}(U, V)$, der $\dim U = n$, $\dim V = m$.

1. Find T^*T

2. Find egenverdierne $\lambda_1, \dots, \lambda_n$ til T^*T

3. Sæt $\sigma_k := \sqrt{\lambda_k}$, $k = 1, \dots, n$.

4. Find en o.n. liste $(u_1, \dots, u_n)$ af egenvektorer tilhørende $\lambda_1, \dots, \lambda_n$.

(a) Find egenvektorer $\tilde{u}_1, \dots, \tilde{u}_n$ tilh. $\lambda_1, \dots, \lambda_n$.

Om $\lambda_j \neq \lambda_k$ vil automatisk $\tilde{u}_j \perp \tilde{u}_k$.

Om en egenverdi har multiplicitet 2 eller mere, velg egenvektorerne til at være ortonormale i egenrummet.

(b) La $u_k := \frac{1}{\|\tilde{u}_k\|} \tilde{u}_k$, $k = 1, \dots, n$.

5. Definér $v_k := \frac{1}{\sigma_k} Tu_k$, $k = 1, \dots, r$, og velg $v_{r+1}, \dots, v_m$

s.d. $(v_1, \dots, v_m)$ er o.n.

Teorem (SVD for matriser)

La $K = \mathbb{R}$ eller $K = \mathbb{C}$, og la $A \in M_{m \times n}(K)$. Da finnes unitære matriser $B \in M_n(K)$ og $C \in M_m(K)$ og $\sigma_1, \dots, \sigma_r \in [0, \infty)$ slike at om $D = \text{diag}_{m \times n}(\sigma_1, \dots, \sigma_r)$, er

$$A = C D B^*.$$

Algoritme: 1. La $\lambda_1, \dots, \lambda_n \geq$ egenverd. til A^*A , med ~~egen~~ orton. egenvekt. $u_1, \dots, u_n$. La $\sigma_k := \sqrt{\lambda_k}$, $k = 1, \dots, n$.

2. La $v_k := \frac{1}{\sigma_k} A u_k$, $k = 1, \dots, r$. Velg $v_{r+1}, \dots, v_m$ o.a. ($v_1, \dots, v_m$) er o.n.

3. Velg $B = (u_1 \dots u_n)$, $C = (v_1 \dots v_m)$.

Eksempel:

Finn SVD av $A := \begin{pmatrix} 1 & -1 \\ 2 & 2 \\ 2 & -2 \end{pmatrix}$. Vi har

$$A^*A = A^T A = \begin{pmatrix} 1 & -2 & 2 \\ -1 & 2 & -2 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 2 & 2 \\ 2 & -2 \end{pmatrix} = \begin{pmatrix} 9 & -9 \\ -9 & 9 \end{pmatrix},$$

som har kar. pol. $p(\lambda) = \lambda^2 - 18\lambda$, så $\lambda_1 = 18$, $\lambda_2 = 0$.

Vi får $\sigma_1 = \sqrt{18}$, $\sigma_2 = 0$. Egenvektorene blir:

$$(A^*A - \lambda_1 I_2) \tilde{u}_1 = \begin{pmatrix} 9 & -9 \\ -9 & -9 \end{pmatrix} \tilde{u}_1 = 0 \iff \tilde{u}_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix},$$

$$(A^*A - \lambda_2 I_2) \tilde{u}_2 = \begin{pmatrix} 9 & -9 \\ -9 & 9 \end{pmatrix} \tilde{u}_2 = 0 \iff \tilde{u}_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Normaliser: $u_1 = \begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{pmatrix}$, $u_2 = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}$.

$$\text{Velg } v_1 = \frac{1}{\sqrt{18}} A u_1 = \frac{1}{\sqrt{18}} \cdot \frac{1}{\sqrt{2}} \begin{pmatrix} 2 \\ -4 \\ 4 \end{pmatrix} = \frac{1}{6} \begin{pmatrix} 2 \\ -4 \\ 4 \end{pmatrix} = \begin{pmatrix} 1/3 \\ -2/3 \\ 2/3 \end{pmatrix}$$

Velg nå v_2, v_3 n.a. (v_1, v_2, v_3) er ortonormal:

Om $v_1 \cdot \tilde{v} = 0$ er $\frac{1}{3}(a - 2b + 2c) = 0$ (der $\tilde{v} = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$)

nå $a - 2b + 2c = 0$. Velg f.eks. $a = 0$, nå $b = c$,

og vi får $\tilde{v}_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$. Om $b = -c = 1$ må

$a = 2b - 2c = 4$, nå $\tilde{v}_3 = \begin{pmatrix} 4 \\ 1 \\ -1 \end{pmatrix}$. Vi får til slutt

$$v_1 = \frac{1}{3} \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}, \quad v_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \quad v_3 = \frac{1}{\sqrt{18}} \begin{pmatrix} 4 \\ 1 \\ -1 \end{pmatrix} = \frac{1}{3\sqrt{2}} \begin{pmatrix} 4 \\ 1 \\ -1 \end{pmatrix}$$

ELLER vi kan gjøre Gram-Schmidt på (v_1, e_1, e_2, e_3) :

$$\tilde{v}_2 := e_1 - (v_1 \cdot e_1)v_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} - \frac{1}{3} \begin{pmatrix} 1/3 \\ 2/3 \\ 2/3 \end{pmatrix} = \begin{pmatrix} 8/9 \\ -2/9 \\ -2/9 \end{pmatrix},$$

$$\text{nå } v_2 = \frac{1}{\sqrt{72+4+4}} \tilde{v}_2 = \frac{1}{\sqrt{80}} \tilde{v}_2 = \frac{1}{4\sqrt{5}} \begin{pmatrix} 8 \\ 2 \\ -2 \end{pmatrix},$$

$$\tilde{v}_3 := e_2 - (v_1 \cdot e_2)v_1 - (v_2 \cdot e_2)v_2$$

$$= \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} - \left(-\frac{2}{3}\right) \begin{pmatrix} 1/3 \\ -2/3 \\ 2/3 \end{pmatrix} - \frac{1}{2\sqrt{5}} \cdot \frac{1}{4\sqrt{5}} \begin{pmatrix} 8 \\ 2 \\ -2 \end{pmatrix}$$

$$= \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 2/9 \\ -4/9 \\ 4/9 \end{pmatrix} - \frac{1}{40} \begin{pmatrix} 8 \\ 2 \\ -2 \end{pmatrix} = \begin{pmatrix} 2/9 - 1/5 \\ 5/9 - 1/20 \\ 4/9 + 1/20 \end{pmatrix}$$

$$\text{og } v_3 = \frac{1}{\|\tilde{v}_3\|} \tilde{v}_3.$$

Vi får $A = QDQ^*$, der $B = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}$, $D = \begin{pmatrix} \sqrt{18} & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$,

$$Q = \begin{pmatrix} 1/3 & 0 & 4/3\sqrt{2} \\ -2/3 & 1/\sqrt{2} & 1/3\sqrt{2} \\ 2/3 & 1/\sqrt{2} & -1/3\sqrt{2} \end{pmatrix}.$$

□

Korollar 5.6.11

- (i) $(u_{r+1}, \dots, u_n)$ er en basis for $\ker T$
- (ii) $(v_1, \dots, v_r)$ er en basis for $\operatorname{im} T$
- (iii) T har rang r
- (iv) (σ_j^2, u_j) ($j=1, \dots, n$) er egenparene til T^*T
- (v) (σ_j^2, v_j) ($j=1, \dots, m$) er egenparene til TT^* .

Basis

(i): Lidsa $D := (u_{r+1}, \dots, u_n)$ er ortonormal, og

$$T^*T u_k = \sigma_k^2 u_k = 0 \quad \text{for } k = r+1, \dots, n, \text{ n\u00e5 } u_{r+1}, \dots, u_n \in \ker T^*T.$$

Men af Lemma 5.6.1 (iii) er $\ker T^*T = \ker T$.

Videre er (Prop. 5.6.7 (iii'))

$$n - r = (\# \text{ n\u00f8g. verd. lik } 0) = \dim \ker T,$$

n\u00e5 D udsprenger hele $\ker T$.

(ii): $E := (v_1, \dots, v_r)$ er ortonormal og ligger i $\operatorname{im} T$.

Av Prop. 5.6.7. (ii) er

$$r = (\# \text{ positive n\u00f8g. verd. }) = \dim \operatorname{im} T,$$

n\u00e5 E er en basis for $\operatorname{im} T$.

(iii): Opgave.

(iv): Per def.

(v): Mer innviklet; se \u00f8reboka.



5.6.5: Redusert SVD

Eksempel: Vi nå tidligere at $A := \begin{pmatrix} 1 & -1 \\ -2 & 2 \\ 2 & -2 \end{pmatrix}$ har SVD

$$A = (v_1 \quad v_2 \quad v_3) \begin{pmatrix} \sqrt{18} & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} u_1^T \\ u_2^T \end{pmatrix}, \text{ der } u_1 = \dots$$

$$= (\sqrt{18} v_1 \quad 0) \begin{pmatrix} u_1^T \\ u_2^T \end{pmatrix} = \sqrt{18} v_1 u_1^T.$$

Dermed vil kolonnene v_2, v_3 og raden u_2^T sannsynligvis fjernes.

□

Def: La A ha SVD $A = C D B^*$, der $B = (u_1, \dots, u_n)$,

$C = (v_1, \dots, v_m)$ og $D = \text{diag}_{m \times n}(\sigma_1, \dots, \sigma_r)$. Den reduserte

SVD av A er da $A = C_r D_r B_r^*$, der

$B_r = (u_1, \dots, u_r)$, $C = (v_1, \dots, v_r)$ og $D = \text{diag}(\sigma_1, \dots, \sigma_r)$.

5.7 Anvendelser av SVD

5.7.1: Beregne operatornorm

For $A \in M_{m \times n}(\mathbb{K})$ har vi $\|A\|_2 := \sup_{\substack{x \in \mathbb{K}^n \\ x \neq 0}} \frac{\|Ax\|_2}{\|x\|_2}$.

Om $A = CDB^*$ er SVD, er B og C unitære, så

$$\|A\|_2 = \|CDB^*\|_2 = \|DB^*\|_2 = \|D\|_2,$$

og operatornormen av en diagonalmatrise er bare største element, som her er $\|D\|_2 = \max_k |\sigma_k| = \sigma_1$. Altså er

$$\|A\|_2 = \sigma_1.$$

$$\begin{aligned} \text{Om } x = u_1, \text{ er } \frac{\|Ax\|_2}{\|x\|_2} &= \|CDB^*u_1\|_2 = \|CDe_1\|_2 = \|C(\sigma_1 e_1)\|_2 \\ &= \sigma_1 \|e_1\|_2 = \sigma_1, \end{aligned}$$

så supremum oppnås i $x = u_1$.

5.7.2: Over- og underbestemte systemer

Et system $Ax=b$ er overbestemt om $m > n$ (m =antall rader, n =# kolonner) og underbestemt om $m < n$. Et overbestemt system vil generelt ikke ha løsnings, mens et underbestemt system vil kunne ha uendelig mange løsnings.

Om $Ax=b$ er overbestemt, vil det (antagelig) ikke finnes x n.a. $Ax=b$, dvs $Ax-b=0$, men vi kan ønske å finne x n.a. $Ax-b$ er så nærme 0 som mulig.

Tilsvarende om $Ax=b$ er underbestemt, vil vi ønske å finne den løsningen som er minst.

Def: La $A \in M_{m \times n}(\mathbb{K})$ ha SVD $A^* := CDB^*$. Den pseudoinverse av A er $A^+ := BD^+C^*$, der $D^+ = \text{diag}_{n \times m}(\frac{1}{\sigma_1}, \dots, \frac{1}{\sigma_r})$.

Teorem:

La $A \in M_{m \times n}(\mathbb{K})$ og $b \in \mathbb{K}^m$.

(i) Om $\nexists$ løsn. av $Ax=b$, vil $x := A^+b$ være vektoren som minimerer $\|Ax-b\|_2^2$.

(ii) Om $\exists$ flere løsn. av $Ax=b$, vil $x := A^+b$ være ~~vektoren~~ den av løsningene med minst norm $\|x\|_2^2$.

Basis (i):

Vi vil minimere $\|Ax-b\|_2^2$ over $x \in \mathbb{K}^n$, dvs. minimere $\|y-b\|_2^2$ over $y \in \text{im } A$, og deretter finne $x \in \mathbb{K}^n$ n.a. $Ax=y$.

Om $x := A^+ b$, vil

$$Ax = CDB^*(BD^+C^*b) = CDD^+C^*b.$$

V_i har $DD^+ = \text{diag}_{m \times m}(\underbrace{1, \dots, 1}_{r \text{ ganger}})$, så

$$CDD^+ = (v_1 \dots v_r \ 0 \dots 0).$$

Altså er

$$Ax = (v_1 \dots v_r \ 0 \dots 0) \begin{pmatrix} \bar{v}_1^T \\ \vdots \\ \bar{v}_r^T \\ \vdots \\ \bar{v}_m^T \end{pmatrix} b$$

$$= \sum_{k=1}^r v_k \bar{v}_k^T b = \sum_{k=1}^r v_k (b \cdot v_k),$$

altså projeksjonen av b ned på $\text{span}(v_1, \dots, v_r) = \text{im } A$.

Men dette er jo nettopp den vektoren som minimerer $\|y - b\|_2$ for $y \in \text{im } A$. 

5.7.3: Lavrangsapprosimasjon

La $T \in \mathcal{L}(U, V)$ ha SVD

$$Tu = \sum_{j=1}^r \sigma_j \langle u, u_j \rangle v_j \quad \forall u \in U.$$

T har rang r fordi $\text{im } T$ spennes av $(v_1, \dots, v_r)$. Om singularverdiene avtar raskt, kan vi ignorere de minste:

$$Tu = \sum_{j=1}^r \sigma_j \langle u, u_j \rangle v_j \approx \sum_{j=1}^s \sigma_j \langle u, u_j \rangle v_j =: T_s u$$

for en $s \in \{1, \dots, r\}$. Den nye lin. avb. T_s har rang s .

Teorem:

T_S har rang s , og om $R \in \mathcal{L}(U, V)$ også har rang s , er

$$\|T - T_S\|_{\mathcal{L}} \leq \|T - R\|_{\mathcal{L}}.$$

Beris:

Vi har $[T]_{\mathcal{B}}^{\mathcal{B}} = \text{diag}_{mn}(\sigma_1, \dots, \sigma_s)$, så

$$\|T - T_S\|_{\mathcal{L}} = \|[T - T_S]_{\mathcal{B}}^{\mathcal{B}}\|_{\mathcal{L}} = \|\text{diag}_{mn}(0, \dots, 0, \sigma_{s+1}, \dots, \sigma_r)\|_{\mathcal{L}} = \sigma_{s+1}.$$

La R ha rang s , så $\dim \ker R = n - s$. Siden $U_s := \text{span}(u_1, \dots, u_{s+1})$ har $\dim. s+1$, og både $\ker R$ og U_s er underrom af U (som har $\dim. n$), må

$$\ker R \cap U_s \neq \{0\}.$$

La derfor $z \in \ker R \cap U_s$, $\|z\| = 1$. Vi får

$$\begin{aligned} \|T - R\|_{\mathcal{L}} &\geq \|(T - R)z\|_{\mathcal{L}} = \|Tz\|_{\mathcal{L}} = \left\| \sum_{j=1}^r \sigma_j \langle z, u_j \rangle u_j \right\|_{\mathcal{L}} \\ &= \left\| \sum_{j=1}^{s+1} \sigma_j \langle z, u_j \rangle u_j \right\|_{\mathcal{L}} \\ &= \sqrt{\sum_{j=1}^{s+1} |\sigma_j \langle z, u_j \rangle|^2} \\ &\geq \min(\sigma_1, \dots, \sigma_{s+1}) \sqrt{\sum_{j=1}^{s+1} |\langle z, u_j \rangle|^2} \\ &= \sigma_{s+1} \|z\| = \sigma_{s+1} = \|T - T_S\|_{\mathcal{L}}. \end{aligned}$$

