7.4 Singulærverdi dekomposisjon

Def

Hvis A er en $m \times n$ matrise, vi sier at

$$A = U \Sigma V^T$$

er en singulærverdi dekomposisjon (SVD) hvis

- U er en $m \times m$ ortogonal matrise,
- V er en $n \times n$ ortogonal matrise,
- $\Sigma = \left(\begin{array}{c|c} D & 0 \\ \hline 0 & 0 \end{array} \right)^{r}_{m-r}$, hvor

D er en $r \times r$ diagonal matrise $\begin{pmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_r \end{pmatrix}$ med strengt positive tall langs diagonalen.

Vanligvis trenger vi også $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_r$.

Anta at $\lambda_1, \ldots, \lambda_n$ er egenverdiene til $A^T A$. Da singulærverdiene, eller s-tallene, til A er $\sigma_1 = \lambda_1^{1/2}, \ldots, \sigma_n = \lambda_n^{1/2}$. Vanligvis skrives vi dem i avtagende rekkefølge:

$$\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n \geq 0.$$

Teorem

Anta at $A = U\Sigma V^T$ er en SVD av en $m\times n$ matrise A,

$$\Sigma = \left(\begin{array}{c|c} D & 0 \\ \hline 0 & 0 \end{array}\right), \quad D = \begin{pmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_r \end{pmatrix}, \quad \sigma_i > 0.$$

Da gjelder:

1) $\sigma_1, \dots, \sigma_r$ er singulærverdiene til A som er strengt positive (men ikke nødvendigvis i avtagende rekkefølge);

2) hvis $U = \{u_1, \dots, u_m\}$, er $\{u_1, \dots, u_r\}$ en ortonormal basis for $\operatorname{Col} A$;

3) $\{u_{r+1}, \dots, u_m\}$ er en ortonormal basis for $(\operatorname{Col} A)^\perp = \operatorname{Nul} A^T$;

4) hvis $V = \{v_1, \dots, v_n\}$, er $\{v_1, \dots, v_r\}$ en ortonormal basis for $\operatorname{Row}(A)$;

5) $\{v_{r+1}, \dots, v_n\}$ er en ortonormal basis for $\operatorname{Row}(A)^\perp = \operatorname{Nul} A$.

Bevis

1)
$$\begin{aligned} A^TA &= (U\Sigma V^T)^T U\Sigma V^T \\ &= V\Sigma^T U^T U \Sigma V^T \\ &= V \Sigma^T \Sigma V^T \\ &= V \begin{pmatrix} \sigma_1^2 & & & & 0 \\ & \ddots & & & \\ & & \sigma_r^2 & & \\ & & & 0 & \\ 0 & & & & \ddots \; 0 \end{pmatrix} V^T \end{aligned}$$

$\boxed{U^TU = I \text{ fordi } U \text{ er ortogonal}}$

Vi ser at egenverdiene til A^TA er $\sigma_1^2, \dots, \sigma_r^2, \underbrace{0, \dots, 0}_{n-r}$. Det viser at $\sigma_1, \dots, \sigma_r$ er singulærverdiene til A som er strengt positive.

2) $A = U\Sigma V^T$

Det følger at $AV = U\Sigma = [\sigma_1 u_1, \dots, \sigma_r u_r, 0, \dots 0]$.

Vi konkluderer

$\mathrm{span}\{u_1, \dots, u_r\} = \mathrm{Im}\, AV = \mathrm{Im}\, A = \mathrm{Col}(A)$.

fordi V er invertibel

3) Siden $\{u_1, \dots, u_m\}$ er en ortonormal basis for $\mathbb{R}^m$ og $\{u_1, \dots, u_r\}$ er en ortonormal basis for $\mathrm{Col}(A)$, blir $\{u_{r+1}, \dots, u_m\}$ en ortonormal basis for $\mathrm{Col}(A)^\perp = \mathrm{Nul}\, A^T$.

4), 5) følger fra 2), 3), fordi

$$A^T = V \Sigma^T U^T$$

er en SVD for A^T. $\blacksquare$

Hvordan finner vi SVD?

<u>Oppsummering</u>:

1) A^TA er en symmetrisk matrise, vi finner strengt positive egenverdiene $\lambda_1, \dots, \lambda_r$ til A^TA, og la

$$\sigma_1 = \lambda_1^{1/2}, \dots, \sigma_r = \lambda_r^{1/2};$$

2) Finner en ortonormal basis $\{v_1, \dots, v_n\}$ som består av egenvektorer til A^TA:

$A^TAv_1 = \lambda_1 v_1, \dots, A^TAv_r = \lambda_r v_r$, $A^TAv_{r+1} = 0, \dots,$

$A^TAv_n = 0$.

3) Definerer
$$u_i = \frac{1}{\sigma_i} A v_i \quad , \; i = 1, \dots, r .$$
Etterpå finner $u_{r+1}, \dots, u_m$ slik at $\{u_1, \dots, u_m\}$ er en ortonormal basis for $\mathbb{R}^m$.

Nå kan vi definere:

$V = \{v_1, \dots, v_n\}$, $U = \{u_1, \dots, u_m\}$,

$$D = \begin{pmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_r \end{pmatrix} \quad , \quad \Sigma = \begin{array}{c c} \left(\begin{array}{c|c} D & 0 \\ \hline 0 & 0 \end{array} \right) & \begin{array}{l} r \\ m-r \end{array} \\ \begin{array}{cc} r & \;\; n-r \end{array} & \end{array}$$

Anta at $A = U \Sigma V^T$ er en SVD,

$$\Sigma = \left(\begin{array}{c|c} D & 0 \\ \hline 0 & 0 \end{array} \right) , \quad D = \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{pmatrix} , \quad \sigma_i > 0 .$$

Definer

$U_r = \{u_1, \dots, u_r\}$ (hvis $U = \{u_1, \dots, u_m\}$),

$V_r = \{v_1, \dots, v_r\}$ (hvis $V = \{v_1, \dots, v_n\}$).

Vi har
$$A = U_r D V_r^T .$$
Dette kalles en <u>redusert SVD</u>.

Matrise $A^+ = V_r D^{-1} U_r^T$ kalles <u>pseudoinvers</u> til A

Egenskaper av pseudoinvers:

1) $AA^+ = \text{Proj}_{Col(A)} : \mathbb{R}^m \to Col(A)$,

2) $A^+A = \text{Proj}_{Row(A)} : \mathbb{R}^n \to Row(A)$,

3) $AA^+A = A$, $A^+AA^+ = A^+$.

Bevis

1)
$$\begin{aligned} AA^+ &= U_r D V_r^T V_r D^{-1} U_r^T \\ &= U_r D D^{-1} U_r^T \\ &= U_r U_r^T = u_1 u_1^T + \dots + u_r u_r^T \\ &= \text{Proj}_{\text{span}\{u_1, \dots, u_r\}} \\ &= \text{Proj}_{Col(A)}. \end{aligned}$$

2)
$$\begin{aligned} A^+A &= V_r D^{-1} U_r^T U_r D V_r^T = V_r V_r^T \\ &= \text{Proj}_{\text{span}\{v_1, \dots, v_r\}} \\ &= \text{Proj}_{Row(A)}. \end{aligned}$$

3)
$$\begin{aligned} AA^+A &= (U_r U_r^T)(U_r D V_r^T) \\ &= U_r D V_r^T = A ; \end{aligned}$$

$$\begin{aligned} A^+AA^+ &= (V_r V_r^T) V_r D^{-1} U_r^T \\ &= V_r D^{-1} U_r^T = A^+. \end{aligned}$$

Som en anvendelse, betrakt et lineært system
$$Ax = b. \qquad (*)$$
La $A = U_r D V_r^T$ være en redusert SVD for A, betrakt $A^+ = V_r D^{-1} U_r^T$.

Da er $x^+ = A^+ b$ er den minste kvadraters løsningen til $(*)$ med minst mulig normen.

Bevis

Vi har
$$Ax^+ = AA^+ b = \mathrm{Proj}_{\mathrm{Col}(A)}(b),$$
Så er x^+ en minste kvadraters løsning til $(*)$.

Merk at vi har
$$x^+ = A^+ b = A^+ A A^+ b = \mathrm{Proj}_{\mathrm{Row}(A)}(A^+ b).$$
Ta en minste kvadraters løsning $\hat{x}$ til $(*)$.

Da $x^+ - \hat{x} \in \mathrm{Nul}\, A = \mathrm{Row}(A)^\perp$. Det følger at
$$\|\hat{x}\|^2 = \|\underbrace{(\hat{x} - x^+)}_{\in \mathrm{Row}(A)^\perp} + \underbrace{x^+}_{\in \mathrm{Row}(A)}\|^2 = \|\hat{x} - x^+\|^2 + \|x^+\|^2$$
$$\geq \|x^+\|^2, \text{ så}$$
$\|\hat{x}\| \geq \|x^+\|$, og $\|\hat{x}\| = \|x^+\|$ hvis og bare hvis $\hat{x} = x^+$. ∎

7.2. Kvadratiske former

Husk at en kvadratisk form på $\mathbb{R}^n$ er en funksjon $Q: \mathbb{R}^n \to \mathbb{R}$,

$$Q(x) = x^T A x = \sum_{i,j=1}^{n} a_{ij} x_i x_j .$$

Matrisen A kalles matrisen til Q.

Vi kan velge A å være symmetrisk, da blir A entydig bestemt av Q.

Hva skjer med Q ved variabelskift?

Anta at S er en inverterbar matrise. Definer

$$Q'(y) = Q(Sy).$$

Vi har

$$Q'(y) = (Sy)^T A S y = y^T S^T A S y.$$

Vi ser at Q' er en kvadratisk form og $S^T A S$ er en matrise til Q'.

Teorem 4

Betrakt en kvadratisk form Q på $\mathbb{R}^n$. Da finnes en ortogonal matrise P slik at formen $Q'(y) = Q(Py)$ har ingen kryssledd.

Bevis

La A være den symmetriske matrisen til Q.
Vi kan finne en ortogonal diagonalisering av A:

$$A = PDP^T, \quad P \text{ er ortogonal}, \quad D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

Vi har $P^T A P = D$.
Dette viser at $Q'(y) = Q(Py)$ har matrise D, så $Q'(y) = \lambda_1 y_1^2 + \dots + \lambda_n y_n^2$. ∎