7.4 Singulærverdi-dekomposisjon

La A være en $m \times n$ matrise (m rader og n kolonner). Vi kan tenke på A som en lineærtransformasjon

$$A: \mathbb{R}^n \to \mathbb{R}^m.$$

Vi er inspirert av resultatene for symmetriske matriser og betrakter problemet

$$\begin{cases} \text{maksimere } \|A\vec{x}\|, \\ \text{gitt at } \|\vec{x}\| = 1. \end{cases}$$

Merknad. Hvis A er $n \times n$ og λ er en egenverdi med egenvektor $\vec{v}$ av lengde 1, så er

$$\|A\vec{v}\| = \|\lambda \vec{v}\| = |\lambda| \cdot \|\vec{v}\| = |\lambda|.$$

Eksempel 7.4.1

Løser problemet for $A = \begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix}$. Det er ekvivalent å maksimere $\|A\vec{x}\|^2$. Derfor,

$$\|A\vec{x}\|^2 = (A\vec{x}) \cdot (A\vec{x}) = (A\vec{x})^T A\vec{x}$$

$$= \vec{x}^T A^T A \vec{x} = \vec{x} \cdot (\underbrace{A^T A}_{= B} \vec{x}).$$

Her er

$$B = A^T A = \begin{bmatrix} 4 & 8 \\ 11 & 7 \\ 14 & -2 \end{bmatrix} \begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix} = \begin{bmatrix} 80 & 100 & 40 \\ 100 & 170 & 140 \\ 40 & 140 & 200 \end{bmatrix}.$$

B er symmetrisk og egenverdiene er

$$\lambda_1 = 360, \quad \lambda_2 = 90 \quad \text{og} \quad \lambda_3 = 0.$$

En ortonormal mengde egenvektorer er

$$\vec{v}_1 = \frac{1}{3}\begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}, \quad \vec{v}_2 = \frac{1}{3}\begin{bmatrix} -2 \\ -1 \\ 2 \end{bmatrix} \quad \text{og} \quad \vec{v}_3 = \frac{1}{3}\begin{bmatrix} 2 \\ -2 \\ 1 \end{bmatrix}.$$

Fra Teorem 7.6 får at maksimum av $\|A\vec{x}\|^2$ for $\|\vec{x}\|=1$ er $\lambda_1 = 360$, og at vi kan velge $\vec{x} = \vec{v}_1$ for å oppnå maksimumet. Så for å maksimere $\|A\vec{x}\|$ får vi $\sqrt{\lambda_1} = \sqrt{360} = 6\sqrt{10}$ som oppnås for $\vec{x} = \vec{v}_1$.

Merknad. $A\vec{v}_1 = \begin{bmatrix} 18 \\ 6 \end{bmatrix}$ og $\|A\vec{v}_1\| = 6\sqrt{10}$.

Noen generelle betraktninger.

Hvis A er en $m \times n$ matrise, så $A^T A$ en symmetrisk $n \times n$-matrise, fordi

$$(A^T A)^T = A^T (A^T)^T = A^T A.$$

Siden $\vec{x} \cdot A^TA\vec{x} = \|A\vec{x}\|^2$ følger det at matrisen A^TA positivt semidefinitt, som betyr at egenverdiene til A^TA oppfyller

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n \geq 0.$$

La $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n\}$ være en ortonormal mengde av egenvektorer for A^TA. Da er

$$\|A\vec{v}_j\| = \sqrt{\vec{v}_j \cdot A^TA\vec{v}_j} = \sqrt{\vec{v}_j \cdot \lambda_j\vec{v}_j} = \sqrt{\lambda_j}.$$

Definisjon

La A være en $m \times n$ matrise og la

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$$

være egenverdiene til A^TA. Singulærverdiene til A er tallene $\sigma_j = \sqrt{\lambda_j}$ for $j = 1, 2, \dots, n$.

Eksempel 7.4.2

For matrisen $A = \begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix}$ er singulærverdiene

$$\sigma_1 = \sqrt{360} = 6\sqrt{10}, \quad \sigma_2 = \sqrt{90} = 3\sqrt{10} \quad \text{og} \quad \sigma_3 = \sqrt{0} = 0.$$

Merknad. Det er ikke tilfeldig at $\sigma_3 = 0$. Siden

$$A : \mathbb{R}^3 \to \mathbb{R}^2$$

er det klart at $\text{rank}(A^TA) = \text{rank}(A) \leq 2$, så 3×3 matrisen A^TA må ha 0 som egenverdi.

Teorem 7.9

La A være en $m\times n$ matrise og la

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$$

være egenverdiene til A^TA med ortonormale egenvektorer

$$\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}.$$

Anta $\lambda_j > 0$ hvis og bare hvis $j = 1, 2, \ldots, r$. Da er

- $\{A\vec{v}_1, A\vec{v}_2, \ldots, A\vec{v}_r\}$ er en ortogonal basis for $\text{Col}(A)$ og
- $\text{rank}(A) = r$.

Merknad. Siden A er $m\times n$, så er $r \leq \min(m,n)$.

Bevis. Siden $\vec{v}_j \cdot \vec{v}_k = 0$ for $j \neq k$ får vi

$$A\vec{v}_j \cdot A\vec{v}_k = \vec{v}_j \cdot A^TA\vec{v}_k = \vec{v}_j \cdot \lambda_k \vec{v}_k$$

$$= \lambda_k(\vec{v}_j \cdot \vec{v}_k) = 0.$$

Videre er

$$\|A\vec{v}_j\| = \sqrt{\lambda_j}.$$

Følgelig er

$$A\vec{v}_j \neq \vec{0} \quad \text{for } j = 1,2,\dots,r,$$

$$A\vec{v}_j = \vec{0} \quad \text{for } j = r+1,\dots,n.$$

Altså er $\{A\vec{v}_1, A\vec{v}_2, \dots, A\vec{v}_r\}$ er en ortogonal basis for et underrom av $\mathrm{Col}(A)$, fordi $A\vec{v}_j \in \mathrm{Col}(A)$ for $j = 1,2,\dots,r$. Vi må vise at underrommet faktisk er hele $\mathrm{Col}(A)$.

La $\vec{y} \in \mathrm{Col}(A)$. Da finnes $\vec{x} \in \mathbb{R}^n$ slik at

$$\vec{y} = A\vec{x}.$$

Siden $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n\}$ er en basis for $\mathbb{R}^n$, så er

$$\vec{x} = c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_n\vec{v}_n.$$

Altså blir

$$\vec{y} = A\vec{x} = c_1 A\vec{v}_1 + c_2 A\vec{v}_2 + \cdots + c_n A\vec{v}_n.$$

$$= c_1 A\vec{v}_1 + c_2 A\vec{v}_2 + \cdots + c_r A\vec{v}_r$$

fordi $A\vec{v}_j = \vec{0}$ for $j = r+1,\dots,n$.

Eksempel 7.4.3

For matrisen $A = \begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix}$ har vi

$$\sigma_1 = 6\sqrt{10}, \quad \sigma_2 = 3\sqrt{10} \quad \text{og} \quad \sigma_3 = 0.$$

Videre er

$$\vec{v}_1 = \frac{1}{3}\begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}, \quad \vec{v}_2 = \frac{1}{3}\begin{bmatrix} -2 \\ -1 \\ 2 \end{bmatrix} \quad \text{og} \quad \vec{v}_3 = \frac{1}{3}\begin{bmatrix} 2 \\ -2 \\ 1 \end{bmatrix}.$$

Til slutt

$$A\vec{v}_1 = \begin{bmatrix} 18 \\ 6 \end{bmatrix}, \quad A\vec{v}_2 = \begin{bmatrix} 3 \\ -9 \end{bmatrix} \quad \text{og} \quad A\vec{v}_3 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Definerer

$$\vec{u}_1 = \frac{1}{\sigma_1} A\vec{v}_1 = \frac{1}{\sqrt{10}}\begin{bmatrix} 3 \\ 1 \end{bmatrix},$$

$$\vec{u}_2 = \frac{1}{\sigma_2} A\vec{v}_2 = \frac{1}{\sqrt{10}}\begin{bmatrix} 1 \\ -3 \end{bmatrix}.$$

Da er

$$U = \begin{bmatrix} \vec{u}_1 & \vec{u}_2 \end{bmatrix} = \frac{1}{\sqrt{10}}\begin{bmatrix} 3 & 1 \\ 1 & -3 \end{bmatrix} \quad \text{og}$$

$$V = \begin{bmatrix} \vec{v}_1 & \vec{v}_2 & \vec{v}_3 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 1 & -2 & 2 \\ 2 & -1 & -2 \\ 2 & 2 & 1 \end{bmatrix}$$

begge ortogonale matriser. Vi vil finne en matrise Σ slik at

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

A: 2×3, U: 2×2, V^T: 3×3

så Σ er 2×3. Mirakuløst fungerer dette med

$$\Sigma = \begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \end{bmatrix} = \begin{bmatrix} 6\sqrt{10} & 0 & 0 \\ 0 & 3\sqrt{10} & 0 \end{bmatrix}.$$

Altså:

$$\begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix} = \frac{1}{\sqrt{10}} \begin{bmatrix} 3 & 1 \\ 1 & -3 \end{bmatrix} \cdot \begin{bmatrix} 6\sqrt{10} & 0 & 0 \\ 0 & 3\sqrt{10} & 0 \end{bmatrix} \cdot \frac{1}{3} \begin{bmatrix} 1 & 2 & 2 \\ -2 & -1 & 2 \\ 2 & -2 & 1 \end{bmatrix}.$$

Dodgson-kondensjon (Ikke pensum)

$$\begin{vmatrix} 1 & -1 & 2 \\ 1 & 2 & 4 \\ 3 & 1 & 3 \end{vmatrix} = \frac{1}{2} \begin{vmatrix} 3 & -8 \\ -5 & 2 \end{vmatrix} = -17.$$

$$\begin{vmatrix} 1 & -1 \\ 1 & 2 \end{vmatrix} = 3 \qquad \begin{vmatrix} -1 & 2 \\ 2 & 4 \end{vmatrix} = -8$$

$$\begin{vmatrix} 1 & 2 \\ 3 & 1 \end{vmatrix} = -5 \qquad \begin{vmatrix} 2 & 4 \\ 1 & 3 \end{vmatrix} = 2$$