47 Singulärwertzerlegung

47.1 Motivation

Wir haben gesehen, dass symmetrische Matrizen vollständig mithilfe ihrer Eigenwerte und Eigenvektoren beschrieben werden können. Diese Darstellung kann unmittelbar zur geometrischen Beschreibung ihrer Wirkung auf Vektoren benutzt werden.

Wir wollen diese Beschreibung auf beliebige Matrizen erweitern. Dies leistet die Singulärwertzerlegung. Singulärwerte können ähnlich gut interpretiert werden wie Eigenwerte symmetrischer Matrizen.

Vorteile der Singulärwertzerlegung gegenüber Eigenwerten und Eigenvektoren:

- Sie ist nicht auf quadratische Matrizen beschränkt.
- In der Singulärwertzerlegung einer reellen Matrix treten nur reelle Matrizen auf (kein Rückgriff auf komplexe Zahlen).

47.2 Definition und Satz: Singulärwertzerlegung

Gegeben sei eine Matrix $A \in \mathbb{R}^{m\times n}$. Dann gibt es orthogonale Matrizen $U \in \mathrm{O}(m)$ und $V \in \mathrm{O}(n)$ sowie eine Matrix $\Sigma = (s_{ij}) \in \mathbb{R}^{m\times n}$ mit $s_{ij} = 0$ für alle $i \neq j$ und nichtnegativen Diagonaleinträgen $s_{11} \geq s_{22} \geq \ldots$, für die

$$A = U\Sigma V^{\mathrm{T}} \tag{$*$}$$

gilt.

Die Darstellung $(*)$ heißt **Singulärwertzerlegung** von A. Die Werte $\sigma_i := s_{ii}$ heißen **Singulärwerte** (singuläre Werte) von A.

47.3 Beispiele

$$A_1 = \begin{pmatrix} 4 & 12 \\ 12 & 11 \end{pmatrix} = \begin{pmatrix} 3/5 & 4/5 \\ 4/5 & -3/5 \end{pmatrix} \begin{pmatrix} 20 & 0 \\ 0 & 5 \end{pmatrix} \begin{pmatrix} 3/5 & 4/5 \\ -4/5 & 3/5 \end{pmatrix}$$

$$A_2 = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$A_3 = \begin{pmatrix} 0{,}36 & 1{,}60 & 0{,}48 \\ 0{,}48 & -1{,}20 & 0{,}64 \end{pmatrix} = \begin{pmatrix} 0{,}8 & 0{,}6 \\ -0{,}6 & 0{,}8 \end{pmatrix} \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 0{,}6 & 0 & 0{,}8 \\ -0{,}8 & 0 & 0{,}6 \end{pmatrix}$$

47.4 Bemerkungen und Beweis

Bemerkungen: 1. Die Anzahl von 0 verschiedener Singulärwerte ist gleich dem Rang r von A.

2. Gleichung $(*)$ kann auch geschrieben werden als

$$A = \sum_{i=1}^{r} \sigma_i u_i v_i^{\mathrm{T}} \ ,$$

wobei u_i der i-te Spaltenvektor von U und v_i der i-te Spaltenvektor von V ist.

Beweis von Satz 47.2: Wir konstruieren eine Singulärwertzerlegung von A.

Schritt 1: Setze $B := A^{\mathrm{T}}A$. Dies ist eine symmetrische $n \times n$-Matrix.

Wir bestimmen die Eigenwerte λ_i und orthonormale Eigenvektoren v_i von B; dabei sei $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$.

Da einerseits

$$v_i^{\mathrm{T}} \cdot Bv_i = \lambda_i v_i^{\mathrm{T}} v_i = \lambda_i$$

und andererseits

$$v_i^{\mathrm{T}} \cdot Bv_i = v_i^{\mathrm{T}} A^{\mathrm{T}} A v_i = (Av_i) \cdot (Av_i) \geq 0$$

gilt, sind alle λ_i nichtnegativ.

Da der Rang von B gleich dem Rang von A ist, sind genau die ersten r Eigenwerte $\lambda_1, \ldots, \lambda_r$ positiv.

Schritt 2: Für $i = 1, \ldots, r$ setze

$$u_i := \frac{1}{\sqrt{\lambda_i}} A v_i \ .$$

Schritt 3: Bestimme $m-r$ orthonormale Vektoren $u_{r+1}, \ldots, u_m$, die zu $u_1, \ldots, u_r$ orthogonal sind.

Schritt 4: Bilde die Matrizen

$$\begin{aligned} U &:= (u_1 \;\; u_2 \;\; \ldots \;\; u_m) \\ V &:= (v_1 \;\; v_2 \;\; \ldots \;\; v_n) \end{aligned}$$

aus den Vektoren $u_1, \ldots, u_m$ bzw. $v_1, \ldots, v_n$ als Spaltenvektoren. Bilde die Matrix $\Sigma = (s_{ij}) \in \mathbb{R}^{m \times n}$ nach der Vorschrift

$$s_{ij} = \begin{cases} \sqrt{\lambda_i}\,, & i = j \leq r\,, \\ 0 & \text{sonst.} \end{cases}$$

Wir zeigen, dass dann $A = U\Sigma V^{\mathrm{T}}$ eine Singulärwertzerlegung von A ist.

- $\{v_1, \ldots, v_n\}$ ist nach Konstruktion eine Orthonormalbasis von $\mathbb{R}^n$, also ist V eine orthogonale Matrix.
- $\{u_1, \ldots, u_m\}$ ist eine Orthonormalbasis von $\mathbb{R}^m$, denn für $i, j = 1, \ldots, r$ ist

$$u_i^{\mathrm{T}} u_j = \frac{1}{\sqrt{\lambda_i \lambda_j}} v_i^{\mathrm{T}} \underbrace{A^{\mathrm{T}} A v_j}_{=\lambda_j v_j} = \frac{\sqrt{\lambda_j}}{\sqrt{\lambda_i}} v_i^{\mathrm{T}} v_j = \begin{cases} \sqrt{\lambda_i/\lambda_i} = 1\,, i = j\,, \\ 0 & \text{sonst,} \end{cases}$$

  also sind $u_1, \ldots, u_r$ orthonormal. Auf $u_{r+1}, \ldots, u_n$ setzt sich die Orthonormalität nach Konstruktion fort. Also ist die Matrix U orthogonal.
- Es gilt

$$\begin{aligned} \sum_{i=1}^{r} \sqrt{\lambda_i} u_i v_i^{\mathrm{T}} &= \sum_{i=1}^{r} A v_i v_i^{\mathrm{T}} && \text{(Definition von } u_i\text{)} \\ &= \sum_{i=1}^{n} A v_i v_i^{\mathrm{T}} && (v_{r+1} \text{ usw. Eigenvektoren zum Eigenwert } 0) \\ &= A \sum_{i=1}^{n} v_i v_i^{\mathrm{T}} = A \underbrace{V V^{\mathrm{T}}}_{=I} \\ &= AI = A && (\{v_1, \ldots, v_n\} \text{ Orthonormalbasis}) \end{aligned}$$

□

47.5 Beispiel

Man bestimme eine Singulärwertzerlegung der Matrix $A = \frac{1}{9}\begin{pmatrix} -8 & 10 & 14 \\ 2 & 2 & 1 \\ -6 & -6 & -3 \\ -16 & 2 & 10 \end{pmatrix}$.

Schritt 1: Es ist

$$B := A^{\mathrm{T}}A = \frac{1}{81}\begin{pmatrix} 360 & -72 & -252 \\ -72 & 144 & 180 \\ -252 & 180 & 306 \end{pmatrix} = \frac{1}{9}\begin{pmatrix} 40 & -8 & -28 \\ -8 & 16 & 20 \\ -28 & 20 & 34 \end{pmatrix}.$$

Die charakteristische Gleichung

$$\begin{aligned} 0 &= -\det(B-\lambda I) \\ &= -\left(\frac{40}{9}-\lambda\right)\left(\frac{16}{9}-\lambda\right)\left(\frac{34}{9}-\lambda\right) - 2\cdot\frac{8}{9}\cdot\frac{20}{9}\cdot\frac{28}{9} \\ &\quad + \frac{8}{9}\cdot\frac{8}{9}\cdot\left(\frac{34}{9}-\lambda\right) + \frac{28}{9}\cdot\frac{28}{9}\cdot\left(\frac{16}{9}-\lambda\right) + \frac{20}{9}\cdot\frac{20}{9}\cdot\left(\frac{40}{9}-\lambda\right) \\ &= \lambda^3 - 10\lambda^2 + 16\lambda \end{aligned}$$

hat die drei Lösungen

$$\begin{aligned} \lambda_1 &= 5 + \sqrt{25-16} = 8 \\ \lambda_2 &= 5 - \sqrt{25-16} = 2 \\ \lambda_3 &= 0 \end{aligned}$$

Eigenvektor zu λ_1: Aus der Bedingung

$$(B-8I)\begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0$$

erhält man das Gleichungssystem (man beachte, dass $\frac{40}{9} - 8 = \frac{40-72}{9} = \frac{-32}{9}$ usw.)

$$\begin{aligned} -32x - 8y - 28z &= 0 \\ -8x - 56y + 20z &= 0 \\ -28x + 20y - 38z &= 0 \end{aligned}$$

mit den Lösungen $x = -2s$, $y = s$, $z = 2s$ für $s \in \mathbb{R}$. Ein normierter Eigenvektor ist daher $v_1 = \begin{pmatrix} -2/3 \\ 1/3 \\ 2/3 \end{pmatrix}$.

Eigenvektor zu λ_2: Aus

$$(B - 2I) \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0$$

erhält man

$$\begin{aligned} 22x - 8y - 28z &= 0 \\ -8x - 2y + 20z &= 0 \\ -28x + 20y + 16z &= 0 \end{aligned}$$

mit den Lösungen $x = 2s$, $y = 2s$, $z = s$ für $s \in \mathbb{R}$. Ein normierter Eigenvektor ist daher $v_2 = \begin{pmatrix} 2/3 \\ 2/3 \\ 1/3 \end{pmatrix}$.

Eigenvektor zu λ_3:

$$B \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0$$

erhält man das Gleichungssystem

$$\begin{aligned} 40x - 8y - 28z &= 0 \\ -8x + 16y + 20z &= 0 \\ -28x + 20y + 34z &= 0 \end{aligned}$$

und damit $x = s$, $y = -2s$, $z = 2s$ für $s \in \mathbb{R}$. Ein normierter Eigenvektor ist daher $v_3 = \begin{pmatrix} 1/3 \\ -2/3 \\ 2/3 \end{pmatrix}$.

Schritt 2: Wir berechnen

$$u_1 := \frac{1}{\sqrt{\lambda_1}} A v_1 = \frac{1}{2\sqrt{2}} \cdot \frac{1}{9} \begin{pmatrix} -8 & 10 & 14 \\ 2 & 2 & 1 \\ -6 & -6 & -3 \\ -16 & 2 & 10 \end{pmatrix} \begin{pmatrix} -2/3 \\ 1/3 \\ 2/3 \end{pmatrix}$$

$$= \frac{1}{54\sqrt{2}} \begin{pmatrix} 54 \\ 0 \\ 0 \\ 54 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} ,$$

$$u_2 := \frac{1}{\sqrt{\lambda_2}} A v_2 = \frac{1}{\sqrt{2}} \cdot \frac{1}{9} \begin{pmatrix} -8 & 10 & 14 \\ 2 & 2 & 1 \\ -6 & -6 & -3 \\ -16 & 2 & 10 \end{pmatrix} \begin{pmatrix} 2/3 \\ 2/3 \\ 1/3 \end{pmatrix}$$

$$= \frac{1}{27\sqrt{2}} \begin{pmatrix} 18 \\ 9 \\ -27 \\ -18 \end{pmatrix} = \frac{1}{3\sqrt{2}} \begin{pmatrix} 2 \\ 1 \\ -3 \\ -2 \end{pmatrix} .$$

Schritt 3: Wir bestimmen zwei orthonormale Vektoren u_3, u_4, die $\{u_1, u_2\}$ zu einer Orthonormalbasis des $\mathbb{R}^4$ ergänzen, z. B.

$$u_3 = \frac{1}{3\sqrt{2}} \begin{pmatrix} -2 \\ -1 \\ -3 \\ 2 \end{pmatrix} , \qquad u_4 = \frac{1}{3\sqrt{2}} \begin{pmatrix} 1 \\ -4 \\ 0 \\ -1 \end{pmatrix} .$$

Schritt 4: Die gesuchte Singulärwertzerlegung ergibt sich als

$$A = \underbrace{\frac{1}{3\sqrt{2}} \begin{pmatrix} 3 & 2 & -2 & 1 \\ 0 & 1 & -1 & -4 \\ 0 & -3 & -3 & 0 \\ 3 & -2 & 2 & -1 \end{pmatrix}}_{U} \cdot \underbrace{\begin{pmatrix} 2\sqrt{2} & 0 & 0 \\ 0 & \sqrt{2} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}}_{\Sigma} \cdot \underbrace{\frac{1}{3} \begin{pmatrix} -2 & 1 & 2 \\ 2 & 2 & 1 \\ 1 & -2 & 2 \end{pmatrix}}_{V^{\mathrm{T}}} .$$

47.6 Weitere Bemerkungen

1. Die Singulärwerte von A und damit die „rechteckige Diagonalmatrix“ Σ sind eindeutig bestimmt, die orthogonalen Matrizen U und V dagegen nicht.

2. Es gilt

$$\begin{aligned} \operatorname{Ker} A &= \operatorname{span}\{v_{r+1}, \dots, v_n\} \\ \operatorname{Im} A &= \operatorname{span}\{u_1, \dots, u_r\} \end{aligned}$$

3. Die Bestimmung der Singulärwerte als Quadratwurzeln der Eigenwerte von $A^{\mathrm{T}} A$ kann zu numerischen Ungenauigkeiten führen. Deswegen ist das Verfahren aus dem Beweis 47.4 zur praktischen Berechnung der Singulärwertzerlegung nicht in allen Fällen geeignet. Andere Verfahren nutzen z. B. Umformungen mittels Householder-Matrizen zurück.

4. Für symmetrische Matrizen A sind die Singulärwerte die Beträge der Eigenwerte. Sind alle Eigenwerte nichtnegativ, so ist die Hauptachsentransformation

$$A = Q \Lambda Q^{\mathrm{T}}$$

auch eine Singulärwertzerlegung.