2. Singulärwertzerlegung

Motivation & Ziel des Vortrages:

- $\forall A \in M_{m,n}(\mathbb{C})\ \exists U \in M_n(\mathbb{C})$ unitär $\wedge\ \exists V \in M_m(\mathbb{C})$ unitär : $A = V\Sigma U^H$ (Fall $\mathbb{K} = \mathbb{C}$)
- $\forall A \in M_{m,n}(\mathbb{R})\ \exists U \in M_n(\mathbb{R})$ orth. $\wedge\ \exists V \in M_m(\mathbb{R})$ orth. : $A = V\Sigma U^T$ (Fall $\mathbb{K} = \mathbb{R}$)

2.1. Singulärwertzerlegung:

Bezeichnung 1: Für $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$, $m, n \in \mathbb{N}$ definiere

$$M_{m,n}(\mathbb{K}) := \{A : A \in \mathbb{K}^{m,n}\}$$ Menge aller $\mathbb{K}$-wertigen $m\times n$-Matrizen

$$M_n(\mathbb{K}) := M_{n,n}(\mathbb{K})$$ Menge aller $\mathbb{K}$-wertigen (quadratischen) $n\times n$-Matrizen

Lemma 2: Sei $A \in M_{m,n}(\mathbb{C})$, dann gilt:

① $A^H A$ ist hermitesch, $\quad (A^H := \bar{A}^T)$

② Die Eigenwerte von $A^H A$ sind reell und nichtnegativ.

Beweis: (Lemma 2)

zu ①: zz: $B^H = B$ für $B := A^H A \in M_{n,n}(\mathbb{C})$: Es gilt

$$B^H \underset{\text{Def } B}{=} (A^H A)^H \underset{\text{Def } \cdot^H}{=} \overline{(\bar{A}^T A)}^T \underset{\overline{z\cdot w} = \bar{z}\cdot\bar{w}}{=} (A^T \bar{A})^T \underset{(AB)^T = B^T A^T}{=} (\bar{A}^T A) \underset{\text{Def } \cdot^H}{=} A^H A \underset{\text{Def } B}{=} B.$$

zu ②: ⓐ: zz: Eigenwerte von $A^H A$ sind reell:

Sei $0 \neq u \in \mathbb{C}^n$ ein Eigenvektor zum Eigenwert $\lambda \in \mathbb{C}$ der Matrix $B := A^H A$, d.h.

(1) $\quad Bu = A^H A u = \lambda u.$

Definiere

$$\langle \cdot, \cdot \rangle : \mathbb{C}^n \times \mathbb{C}^n \longrightarrow \mathbb{C}, \quad \langle u, v \rangle := u^H v, \quad (\text{Sesquilinearform auf } \mathbb{C}^n \times \mathbb{C}^n)$$

$$\|.\| : \mathbb{C}^n \longrightarrow \mathbb{R}_+, \quad \|u\| := \sqrt{\langle u, u \rangle}. \quad (\text{Norm auf } \mathbb{C}^n)$$

Multiplikation von (1) von links mit u^H liefert

$$u^H B u = \langle u, Bu \rangle = \langle u, \lambda u \rangle = \lambda \langle u, u \rangle = \lambda \|u\|^2$$

$$\| $$

$$(\underbrace{B^H}_{\overset{①}{=} B} u)^H u = \langle Bu, u \rangle = \langle \lambda u, u \rangle = \bar{\lambda} \langle u, u \rangle = \bar{\lambda} \|u\|^2$$

Damit gilt $0 = (\lambda - \bar{\lambda})\|u\|^2$, also (wegen $u \neq 0$ gilt $\|u\| \neq 0$) erhalten wir (nach Division durch $2i\|u\|^2$) $0 = \frac{1}{2i}(\lambda - \bar{\lambda}) = \operatorname{Im}\lambda$,

~~Schreiben wir $\lambda = \lambda_1 + i\lambda_2$, $\lambda_1, \lambda_2 \in \mathbb{R}$, so bedeutet dies $0 = (\lambda_1 + i\lambda_2) - (\lambda_1 - i\lambda_2) = 2i\lambda_2$, also $\lambda_2 = 0$, also ist $\lambda \in \mathbb{R}$~~ d.h. $\lambda \in \mathbb{R}$.

ⓑ: zz: Eigenwerte von $A^H A$ sind nichtnegativ: Seien (u, λ) wie in (1), dann gilt:

$$\lambda = \lambda \cdot \frac{\langle u, u \rangle}{\langle u, u \rangle} = \frac{\langle u, \lambda u \rangle}{\|u\|^2} = \frac{\langle u, A^H A u \rangle}{\|u\|^2} = \frac{\langle Au, Au \rangle}{\|u\|^2} = \frac{\|Au\|^2}{\|u\|^2} \underset{\geq 0}{\in \mathbb{R}}. \quad \blacksquare$$

Definition 3: Sei $A \in M_{m,n}(\mathbb{C})$ und seien $\lambda_1, \dots, \lambda_n$ die Eigenwerte von $A^H A$. Dann heißen

$$\sigma_j := \sqrt{\lambda_j} \underset{\geq 0}{\in \mathbb{R}}, \quad j = 1, \dots, n$$

die **Singulärwerte von A**.

Das folgende Lemma besagt, dass es für die Definition der Singulärwerte von A egal ist, ob wir diese über die Eigenwerte λ_j von $A^H A$ oder von AA^H definieren.

Lemma 4: Seien $A \in M_{m,n}(\mathbb{C})$ und $B \in M_{n,m}(\mathbb{C})$. Dann sind für $0 \neq \lambda \in \mathbb{C}$ die folgenden Aussagen äquivalent:

①: λ ist Eigenwert von $AB \in M_m(\mathbb{C})$

②: λ ist Eigenwert von $BA \in M_n(\mathbb{C})$.

Beweis: (Lemma 4)

① ⇒ ②: Sei λ ein Eigenwert von AB, d.h.

(2) $\exists\, 0 \neq u \in \mathbb{C}^m: \; ABu = \lambda u$

Angenommen $u \in Ker(B)$, d.h. $Bu = 0$, dann gilt

$$0 = A\underbrace{Bu}_{=0} = \lambda u \overset{u\neq 0}{\Rightarrow} \lambda = 0 \; \lightning_{\lambda \neq 0}$$

Also sei $u \notin Ker(B)$, d.h. $Bu \neq 0$, dann gilt (Multiplikation von (2) von links mit B)

$$BAv = BA\underbrace{(Bu)}_{=:v} = \lambda\underbrace{(Bu)}_{=:v} = \lambda v \quad , \; 0 \neq v = Bu \in \mathbb{C}^n,$$

d.h. λ ist ein Eigenwert von BA.

② ⇒ ①: Analog. ∎

Bemerkung 5: Sei $A \in M_n(\mathbb{C})$ mit Singulärwerten $\sigma_1, \dots, \sigma_n \geq 0$. Dann gilt

$$A \text{ invertierbar} \iff \sigma_j > 0 \quad \forall j = 1, \dots, n$$

Proposition 6: Sei $A \in M_n(\mathbb{C})$ normal (d.h. $A^H A = A A^H$) mit Eigenwerten $\lambda_1, \dots, \lambda_n$ und Singulärwerten $\sigma_1, \dots, \sigma_n$. Dann gilt:

$$\sigma_j = |\lambda_j|, \quad j = 1, \dots, n.$$

Insbesondere gilt: $\max\limits_{j=1,\dots,n} \sigma_j = \max\limits_{j=1,\dots,n} |\lambda_j| = \max\limits_{\lambda \in \sigma(A)} |\lambda| =: \varrho(A)$ Spektralradius von A

Beweis: (Proposition 6)

Da $A \in M_n(\mathbb{C})$ normal, gilt (vgl. Thm. 3.1, Vortrag 1)

$$\exists\, U \in M_n(\mathbb{C}) \text{ unitär}: \; U^H A U = D := \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \quad \left[\overset{U \text{ unitär}}{\iff} A = UDU^H.\right]$$

Damit gilt

$$A^H A = (UDU^H)^H (UDU^H) = UD^H \underbrace{U^H U}_{=I_n} DU^H = U(D^H D)U^H$$

d.h. die Matrizen $A^H A$ und $D^H D = \begin{pmatrix} |\lambda_1|^2 & & 0 \\ & \ddots & \\ 0 & & |\lambda_n|^2 \end{pmatrix}$ sind ähnlich, haben also die gleichen Eigenwerte.

Somit gilt

$$\sigma_j = \sqrt{|\lambda_j|^2} = |\lambda_j|, \quad j = 1, \dots, n. \; ∎$$

Theorem 7: (Singulärwertzerlegung)

Sei $A \in M_{m,n}(\mathbb{C})$ mit positiven Singulärwerten $\sigma_1, \dots, \sigma_r > 0$ für ein $1 \leq r \leq n$. Dann gilt:

$$\exists\, U \in M_n(\mathbb{C}) \text{ unitär} \wedge \exists\, V \in M_m(\mathbb{C}) \text{ unitär}: \; A = V \Sigma U^H,$$

wobei

$$\Sigma = \begin{pmatrix} \Sigma_r & 0 \\ 0 & 0 \end{pmatrix} \in M_{m,n}(\mathbb{R}) \quad , \; \Sigma_r = \begin{pmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_r \end{pmatrix} \in M_r(\mathbb{R}).$$

Bemerkungen 8:

①: Falls $n = m$ und $U = V \Rightarrow A = V \Sigma V^H$ (Komplexe Schursche Normalform)
und wegen $V^{-1} = V^H$ (da V unitär) $A = V \Sigma V^{-1}$ (Diagonalform)

②: Es gilt $\operatorname{rang}(A) = r \leq \min\{m, n\}$

③: Es gilt

$$A^H A = (V\Sigma U^H)^H (V \Sigma U^H) = U \Sigma^H \overbrace{V^H V}^{=I_m} \Sigma U^H = U \Sigma^T \Sigma U^H$$

$$A A^H = (V \Sigma U^H)(V \Sigma U^H)^H = V \Sigma \underbrace{U^H U}_{=I_n} \Sigma^H V^H = V \Sigma \Sigma^T V^H$$

d.h.

Spalten von $U \in M_n(\mathbb{C})$ sind Eigenvektoren der hermiteschen Matrix $A^H A \in M_n(\mathbb{C})$
" " $V \in M_m(\mathbb{C})$ " " " " " $AA^H \in M_m(\mathbb{C})$

Diagonaleinträge von $\Sigma^T \Sigma \in M_n(\mathbb{R})$ sind Eigenwerte von $A^H A$
" " " $\Sigma \Sigma^T \in M_m(\mathbb{R})$ " " " AA^H

④: Theorem 7 gilt für $\mathbb{K} = \mathbb{R}$: $\forall A \in M_{m,n}(\mathbb{R}) \; \exists\, U \in M_n(\mathbb{R})$ orth. $\wedge \; \exists\, V \in M_m(\mathbb{R})$ orth.: $A = V \Sigma U^T$

⑤: Die Spaltenvektoren v_i von V (bzw. u_j von U) heißen Links- (bzw. Rechts-) Singulärvektoren von A, $\substack{i=1,\dots,m \\ j=1,\dots,n}$

Beweis: (Theorem 7) Sei o.B.d.A $m \geq n$.

Sei $0 \neq u_j \in \mathbb{C}^n$ ein normierter Eigenvektor zum Eigenwert $\sigma_j^2 := \lambda_j \geq 0$ von $A^H A$ (vgl. Lemma 2 ②), d.h.

$$A^H A u_j = \sigma_j^2 u_j \quad , \; j = 1, \ldots, n.$$

Definiere

$$U := (u_1 | u_2 | \cdots | u_n) \in M_n(\mathbb{C}),$$

dann ist U unitär (da $\{u_j : j = 1, \ldots, n\}$ Orthonormalbasis von $\mathbb{C}^n$)

$$(U^H U)_{ij} = u_i^H u_j = \langle u_i, u_j \rangle = \delta_{ij} := \begin{cases} 1 \,, & i = j \quad (\text{da } |u_i|^2 = 1) \\ 0 \,, & i \neq j \quad (\text{da } u_i \perp u_j) \end{cases} \Rightarrow U^H U = I_n$$

und es gilt

$$A^H A U = (A^H A u_1 | \cdots | A^H A u_n) = (\sigma_1^2 u_1 | \cdots | \sigma_n^2 u_n) = U \begin{pmatrix} \sigma_1^2 & & 0 \\ & \ddots & \\ 0 & & \sigma_n^2 \end{pmatrix}$$

$$\Rightarrow U^H A^H A U = U^H U \begin{pmatrix} \sigma_1^2 & & 0 \\ & \ddots & \\ 0 & & \sigma_n^2 \end{pmatrix} \underset{U \text{ unitär}}{=} \begin{pmatrix} \sigma_1^2 & & 0 \\ & \ddots & \\ 0 & & \sigma_n^2 \end{pmatrix} \underset{\text{Bemerkung 8 ③}}{=} \Sigma^T \Sigma$$

für

$$\Sigma = \begin{pmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_n \\ 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{pmatrix} \in M_{m,n}(\mathbb{R}).$$

wobei $\sigma_1 \geq \cdots \geq \sigma_r > \sigma_{r+1} = \cdots = \sigma_n = 0$ (ansonsten vertausche Spalten in U). Es gilt

$$\langle Au_i, Au_j \rangle = \langle A^H A u_i, u_j \rangle = \langle \sigma_i^2 u_i, u_j \rangle = \sigma_i^2 \langle u_i, u_j \rangle = \sigma_i^2 \delta_{ij} \quad , \quad 1 \leq i,j \leq n$$

und insbesondere (für $i = j$)

$$\|Au_i\|^2 = \langle Au_i, Au_i \rangle = \sigma_i^2 \begin{cases} = 0 & , \; r < i \leq n \\ \neq 0 & , \; 1 \leq i \leq r \end{cases} \Rightarrow Au_i \begin{cases} = 0 \,, & r < i \leq n, \\ \neq 0 \,, & 1 \leq i \leq r. \end{cases}$$

Definiere

$$0 \neq v_i := \frac{Au_i}{\sigma_i} \in \mathbb{C}^m \quad , \; 1 \leq i \leq r,$$

dann vervollständige diese durch $v_{r+1}, \ldots, v_m \in \mathbb{C}^m$ zu einer Orthonormalbasis $\{v_i : i = 1, \ldots, m\}$ von $\mathbb{C}^m$ und definiere

$$V := (v_1 | \cdots | v_m) \in M_m(\mathbb{C}).$$

- $v_i := \frac{Au_i}{\sigma_i} \Leftrightarrow \sigma_i v_i = Au_i \,, \; 1 \leq i \leq r$
- $\sigma_i = 0 \,, \; r < i \leq n$

Dann gilt:

$$V \Sigma U^H = (v_1 | \cdots | v_m) \begin{pmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_n \\ 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{pmatrix} U^H = (\sigma_1 v_1 | \cdots | \sigma_n v_n) U^H = (Au_1 | \cdots | Au_r | 0 | \cdots | 0) U^H$$

$$\underset{Au_i = 0, \, r < i \leq n.}{=} A (u_1 | \cdots | u_n) U^H \underset{\text{Def } U}{=} A U U^H \underset{UU^H = I_n}{=} A. \quad \blacksquare$$

Bemerkung 9:

Geometrische Interpretation: Seien

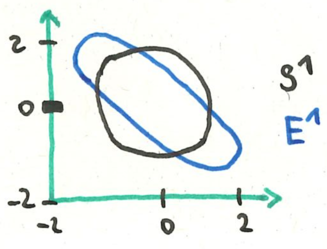
$$S^{n-1} := \left\{ x \in \mathbb{R}^n : |x|^2 := \sum_{i=1}^n x_i^2 = 1 \right\} \qquad \text{Einheitssphäre im } \mathbb{R}^n$$

$$E^{n-1} := \left\{ x \in \mathbb{R}^n : \sum_{i=1}^n \frac{x_i^2}{c_i^2} = 1 \right\} \quad , \; c_i > 0 \qquad \text{Ellipsoid im } \mathbb{R}^n$$

Für $A \in M_n(\mathbb{R})$ invertierbar gilt

$$A S^{n-1} = E^{n-1} \quad \text{für geeignete } c_i > 0.$$

Fall $n = 2$:



2.2. Pseudoinverse:

Definition 10:

Sei $A \in M_{m,n}(\mathbb{C})$ mit positiven Singulärwerten $\sigma_1, \dots, \sigma_r > 0$ für ein $1 \leq r \leq \min\{m,n\}$ und Singulärwertzerlegung $A = V \Sigma U^H$ (vgl. Thm. 7). Dann nennen wir die Matrix

$$A^+ := U \Sigma^+ V^H \in M_{n,m}(\mathbb{C}) \quad \text{mit } \Sigma^+ := \begin{pmatrix} \Sigma_r^{-1} & 0 \\ 0 & 0 \end{pmatrix} \in M_{n,m}(\mathbb{R})$$

die **Pseudoinverse von A**.

Bemerkung 11:

①: Die Pseudoinverse in Definition 10 verallgemeinert den Begriff „Inverse einer Matrix" für
- rechteckige Matrizen
- quadratische, nicht-invertierbare Matrizen

②: Es gelten die folgenden Eigenschaften:

ⓐ: $$A^+A = U\Sigma^+\underbrace{V^HV}_{=I_m}\Sigma U^H = U(\Sigma^+\Sigma)U^H = U\begin{pmatrix} \Sigma_r^{-1} & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} \Sigma_r & 0 \\ 0 & 0 \end{pmatrix}U^H = U\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}U^H$$
$$= \sum_{i=1}^{r} \underbrace{u_i u_i^H}_{\text{ist eine Matrix!!}}$$

ⓑ: $$AA^+ = V\Sigma\underbrace{U^HU}_{=I_n}\Sigma^+V^H = V(\Sigma\Sigma^+)V^H = V\begin{pmatrix} \Sigma_r & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} \Sigma_r^{-1} & 0 \\ 0 & 0 \end{pmatrix}V^H = V\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}V^H$$
$$= \sum_{i=1}^{r} v_i v_i^H$$

ⓒ: $$A = \sum_{i=1}^{r} \sigma_i v_i u_i^H$$

ⓓ: $$A^+ = \sum_{i=1}^{r} \frac{1}{\sigma_i} u_i v_i^H$$

ⓔ:
- Falls $\operatorname{rang}(A) \underset{\text{Bem. 8②}}{=} r = \min\{n,m\}$ (d.h. A hat maximalen Rang), dann gilt
$$A^+ = (A^HA)^{-1}A^H$$
- Falls $\operatorname{rang}(A) = r = n = m$ (d.h. $A \in M_n(\mathbb{C})$ invertierbar), dann gilt
$$A^+A = AA^+ = I_n ,$$
also
$$A^+ = A^{-1}.$$
Die Pseudoinverse ist also eine Verallgemeinerung des Inversen.

ⓕ: Die Pseudoinverse A^+ ist die einzige Matrix X, die die **Moore-Penrose-Bedingungen** erfüllt:
$$AXA = A, \quad XAX = X, \quad XA = (XA)^H, \quad AX = (AX)^H.$$

2.3. Bildkompression durch Singulärwertzerlegung

Betrachte ein Bild

$$\begin{matrix} a_{11} & \cdots & \cdots & \cdots & a_{1n} \\ \vdots & & & & \vdots \\ \vdots & & a_{ij} & & \vdots \\ a_{m1} & \cdots & \cdots & \cdots & a_{mn} \end{matrix}$$

$\boxed{a_{ij}}$ $\hat{=}$ Farbpixel mit Farbinformation
$a_{ij} \in [0,1]$

(i. Allg.: $a_{ij} = \{$

Die Farbinformationen (eines schwarz/weiß Bildes) speichern wir in einer Matrix

$$A \in M_{m,n}(\mathbb{R}) \quad \text{(reelle!)}$$

Bildkompression: Betrachte die reduzierte Singulärwertzerlegung von A

$$A = V \Sigma U^T = \sum_{i=1}^{r} \sigma_i v_i u_i^T \qquad (v_i \in \mathbb{R}^m \text{ i-te Spalte von } V,\ u_i \in \mathbb{R}^n \text{ i-te Spalte von } U)$$

mit $U \in M_n(\mathbb{R})$, $V \in M_m(\mathbb{R})$, $\Sigma = \begin{pmatrix} \Sigma_r & 0 \\ 0 & 0 \end{pmatrix} \in M_{m,n}(\mathbb{R})$, $\Sigma_r = \begin{pmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_r \end{pmatrix} \in M_r(\mathbb{R})$ und positiven Singulärwerten $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > \sigma_{r+1} = \dots = \sigma_n = 0$ von A für ein $1 \leq r \leq n$.

Wir approximieren A durch

$$A_k := \sum_{i=1}^{k} \sigma_i v_i u_i^T \quad , \quad 1 \leq k \leq r$$

A_k ist die beste Approximation von A unter allen Matrizen vom Rang k.

Speicherbedarf:

$$A: \quad m \cdot n$$
$$A_k: \quad k \cdot (m+n+1)$$

Kompressionsverhältnis:

$$C_R := \frac{k(m+n+1)}{m \cdot n}$$

Speicherreduktion, falls

$$k(m+n+1) < m \cdot n \iff C_R < 1.$$

Hinweis: Zur Bildbearbeitung gibt es effizientere Verfahren als die Singulärwertzerlegung.