§ 32 Singulärwertzerlegung und Pseudoinverse

Die Abbildung

$$\mathbb{R}^{m\times n} \longrightarrow \mathbb{Z}$$
$$A \longrightarrow \operatorname{rank} A$$

ist nicht stetig (zum Beispiel weil die Wertemenge diskret ist)

Die Singulärwertzerlegung ist in gewisser Weise eine stetige Approximation dieser Abbildung.

<u>Satz</u> (Singulärwertzerlegung)
$A \in \mathbb{R}^{m\times n}$ $\quad p = \min(m,n)$
Dann existieren durch A eindeutig bestimmte Werte $\sigma_1, \dots, \sigma_p \in \mathbb{R}$ mit $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_p \geq 0$ und orthogonale Matrizen $V \in O(n)$, $U \in O(m)$, so dass

$${}^tUAV = \Sigma$$

wobei $\Sigma = (\sigma_{ij})$ Einträge

$$\sigma_{ij} = \begin{cases} \sigma_i & \text{Fall } i=j \\ 0 & \text{sonst} \end{cases}$$

Also

$${}^tUAV = \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_p \\ \hline & 0 & \end{pmatrix} \Big\}n \quad \text{im Fall } m \geq n$$

bzw

$${}^tUAV = \underbrace{\begin{pmatrix} \sigma_1 & & & \\ & \ddots & & 0 \\ & & \sigma_p & \end{pmatrix}}_{m} \quad \text{im Fall } m \leq n$$

Bem. Vergleichen soll man diesen Satz mit dem Normalformensatz für lineare Abbildungen.

Zu $A \in \mathbb{R}^{m\times n}$ $\exists\, T \in GL(m,\mathbb{R}),\ S \in GL(n,\mathbb{R})$

$$T^{-1}AS = \left(\begin{array}{c|c} \begin{matrix} 1 & & \\ & \ddots & \\ & & 1 \end{matrix} & 0 \\ \hline 0 & 0 \end{array}\right)$$

$r = \operatorname{rank} A$

Anwendungen:

1) Ist A eine Matrix von der wir vermuten, dass sie nicht vollen Rang hat und $\tilde{A}$ eine Approximation, etwa eine floating-point-Darstellung. So gibt es die Singulärwertzerlegung von $\tilde{A}$

$${}^tU\tilde{A}V = \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_p \\ \hline & 0 & \end{pmatrix}$$

Ist σ_p deutlich größer als 0 und $1 \gg \sigma_{r+1} \geq \dots \geq \sigma_p \geq 0$ so lässt sich die Hypothese $\operatorname{rank} A = r$ plausibel aufstellen.

2) $A \in \mathbb{R}^{m\times n}$ beliebig, r vorgegeben

$${}^tUAV = \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_p \\ \hline & 0 & \end{pmatrix} = \Sigma \Rightarrow A = U\Sigma\,{}^tV$$

$$\Sigma_r = \left(\begin{array}{c|c} \begin{matrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{matrix} & \\ \hline & 0 \end{array}\right) \rightsquigarrow A_r := U\Sigma_r\,{}^tV$$

Dann gilt A_r ist die "beste Approximation" von A durch eine Matrix vom Rang $\leq r$.

Beweis: Wir betrachten den Fall $m \geq n$ also $p = n$ (Im Fall $m < n$ argumentieren wir mit ${}^tA$)

Die Matrix $B = {}^tAA$ ist symmetrisch und hat nur nicht negative EW:

Sei ~~…~~ v ein normierter EV von B zum EW λ.

Dann gilt

$$\lambda = \lambda\langle v,v\rangle = \langle v,\lambda v\rangle = \langle v,Bv\rangle = \langle v,{}^tAAv\rangle$$
$$= \langle Av,Av\rangle \geq 0$$

Sei $V \in O(n)$ die B in Diagonalgestalt überführt.

$${}^tVBV = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} \quad \text{mit } \lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$$

Dann ~~…~~ setzen wir $\sigma_i = \sqrt{\lambda_i}$ $i = 1,\dots,p = n$

$\sigma_1,\dots,\sigma_n$ sind die Singulärwerte von A.

$${}^t(UAV)\,{}^tUAV = {}^t\Sigma\,\Sigma = \begin{pmatrix} \sigma_1^2 & & 0 \\ & \ddots & \\ 0 & & \sigma_n^2 \end{pmatrix}$$

$${}^tV\,{}^tA\underbrace{U\,{}^tU}_{E}AV = {}^tVBV = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$$

Sei $r = \operatorname{rank} A = \operatorname{rank}\Sigma$, $\sigma_1 \geq \dots \geq \sigma_r > 0$, $\sigma_{r+1} = \dots = \sigma_p = 0$

Betrachten

$$u_i = \frac{1}{\sigma_i} A\cdot v_i \in \mathbb{R}^m \quad i = 1,\dots,r$$

Dann gilt also $\sigma_i u_i = Av_i$ $i = 1,\dots,r$

$u_1, \dots, u_r$ bilden ein Orthonormalsystem

$$\langle u_i, u_j \rangle = \frac{1}{\sigma_i \sigma_j} \langle A v_i, A v_j \rangle$$

$$= \frac{1}{\sigma_i \sigma_j} \langle v_i, \underbrace{{}^tA A}_{B} v_j \rangle$$

$$= \frac{\lambda_j}{\sigma_i \sigma_j} \langle v_i, v_j \rangle = \begin{cases} 1 & i = j \\ 0 & \text{sonst} \end{cases}$$

Wir ergänzen $u_1, \dots, u_r$ zu einer Orthonormalbasis $\{u_1, \dots, u_r, u_{r+1}, \dots, u_m\}$ von $\mathbb{R}^m$

Es sei $U = (u_1, \dots, u_m) \in O(m)$

$V = (v_1, \dots, v_n) \in O(n)$

Dann gilt: $\sigma_i u_i = A v_i$ für $i = 1, \dots, n$

Für $\langle A v_i, A v_i \rangle = \lambda_i \langle v_i, v_i \rangle = 0$ für $i = r+1, \dots, n$

$\Rightarrow A v_i = 0$ für $i = r+1, \dots, n$

$\sigma_i u_i = A v_i$ gilt da $\sigma_i = 0$

In der Basis $u_1, \dots, u_m$ von $\mathbb{R}^m$ und $v_1, \dots, v_n$ von $\mathbb{R}^n$ hat die Abbildung $f_A : \mathbb{R}^n \to \mathbb{R}^m$ die darstellende Matrix

$$\Sigma = \begin{pmatrix} \sigma_1 & & & & \\ & \ddots & & & \\ & & \sigma_r & & \\ & & & 0 = \sigma_{r+1} & \\ & & & & \ddots \\ & & & & 0 = \sigma_p \\ \hline & & 0 & & \end{pmatrix} \Rightarrow {}^tU A V = \Sigma$$

aus der Kommutativität des Diagramms $\square$

$$\begin{array}{ccc} \mathbb{R}^n & \xrightarrow{A} & \mathbb{R}^m \\ V \uparrow & & \uparrow U \\ \mathbb{R}^n & \xrightarrow[\Sigma]{} & \mathbb{R}^m \end{array}$$

Beispiele 1) Die Singulärwerte einer Orthogonalen Matrix $S \in O(n)$ sind alle 1

In der Tat $U = S \quad V = E$

$${}^tU S \cdot V = {}^tS \cdot S \cdot E = E$$

2) Die Singulärwerte einer symmetrischen Matrix $A \in \mathbb{R}^{n \times n}$ sind $|\lambda_1|, \dots, |\lambda_n|$ der EW von A

$${}^tV A V = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$$

$$\underbrace{{}^t({}^tVAV)({}^tVAV)}_{{}^tV \underbrace{{}^tAA}_{B} V} = \begin{pmatrix} \lambda_1^2 & & \\ & \ddots & \\ & & \lambda_n^2 \end{pmatrix}$$

Also B hat EW $\lambda_1^2, \dots, \lambda_n^2$, A Singulärwerte

$|\lambda_1| = \sqrt{\lambda_1^2}, \dots, |\lambda_n| = \sqrt{\lambda_n^2}$

3)

$$A = \begin{pmatrix} 0.36 & 1.60 & 0.48 \\ 0.48 & -1.20 & 0.64 \end{pmatrix} =$$

$$= \underbrace{\begin{pmatrix} 0.8 & 0.6 \\ -0.6 & 0.8 \end{pmatrix}}_{\in SO(3)} \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \underbrace{\begin{pmatrix} 0 & 1 & 0 \\ 0.6 & 0 & 0.8 \\ -0.8 & 0 & 0.6 \end{pmatrix}}_{\in O(3) \setminus SO(3)}$$

Singulärwerte von A sind 2 und 1

4)

$A = \begin{pmatrix}0&1\\0&0\end{pmatrix}$ hat Singulärwerte $\sigma_1 = 1, \sigma_2 = 0$

$$\begin{pmatrix}1&0\\0&1\end{pmatrix}\underbrace{\begin{pmatrix}1&0\\0&0\end{pmatrix}}_{\Sigma}\begin{pmatrix}0&1\\1&0\end{pmatrix} = \begin{pmatrix}0&1\\0&0\end{pmatrix} = A$$

$$\Sigma = \begin{pmatrix}1&\\&0\end{pmatrix} = \begin{pmatrix}1&0\\0&1\end{pmatrix}\cdot A\cdot\begin{pmatrix}0&1\\1&0\end{pmatrix}$$

$$A_\varepsilon = \begin{pmatrix}0&1\\ \varepsilon&0\end{pmatrix}$$

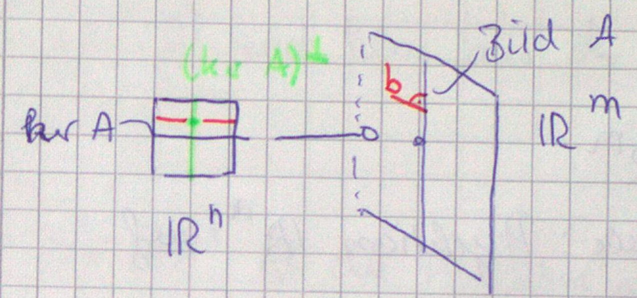
$$\underbrace{\underbrace{\begin{pmatrix}1&0\\0&\text{sign}\,\varepsilon\end{pmatrix}}_{{}^tU}\underbrace{A_\varepsilon\overbrace{\begin{pmatrix}0&1\\1&0\end{pmatrix}}^{V}}_{\begin{pmatrix}1&0\\0&\varepsilon\end{pmatrix}}}_{\Sigma = \begin{pmatrix}1&0\\0&|\varepsilon|\end{pmatrix}} =$$

Mit $\varepsilon \to 0$ geht der 2-te Singulärwert $|\varepsilon| \to 0$

Pseudoinverse

Sei $A \in \mathbb{R}^{m\times n}$ eine Matrix und $f_A : \mathbb{R}^n \to \mathbb{R}^m$ die zugehörige lineare Abbildung.

Im Allgemeinen ist f_A weder injektiv noch surjektiv. Stattdessen haben wir ausgezeichnete Unterräume $\ker A \subset \mathbb{R}^n$, $\text{Bild } A \subset \mathbb{R}^n$



Die Pseudoinverse $A^+ : \mathbb{R}^m \to \mathbb{R}^n$ ist eine Matrix die Folgendes erfüllt.

1) zu $b \in \mathbb{R}^m$ ist $x = A^+ b$ ein $x \in \mathbb{R}^n$ so dass $\|Ax - b\|$ minimal wird

2) Unter allen x mit $\|Ax - b\|$ minimal ist $x = A^+ b \in \mathbb{R}^n$ der Punkt der zum Nullpunkt den kleinsten Abstand hat.

Df. $A^+ : \mathbb{R}^m \to \mathbb{R}^n$ ist die Abbildung die 1) und 2) erfüllt.

Also $A^+ : \mathbb{R}^m \to \text{Bild } A$

A induziert einen Isomorphismus $\underset{\mathbb{R}^n}{\overset{\cap}{(\ker A)^\perp}} \xrightarrow{A} \text{Bild } A$

Zunächst gilt $f_A|_{\ker A^\perp}$ ist injektiv, da $(\ker A)^\perp \cap \ker A = 0$

basis definiert ist.

Ferner

dim Bild $A = n$ - dim ker A = dim $(\ker A)^\perp$

$f_A|_{\ker A^\perp} \to$ Bild A ist nicht nur injektiv sondern auch surjektiv

Satz $A \in \mathbb{R}^{m \times n}$, $A^+ \in \mathbb{R}^{n \times m}$ die Pseudoinverse

Dann gilt

1) $P = AA^+ : \mathbb{R}^m \to \mathbb{R}^m$

definiert die orthogonale Projektion $\mathbb{R}^m$ auf Bild $A \subset \mathbb{R}^m$

Insbesondere gilt $P^2 = P$

2) $Q = A^+A : \mathbb{R}^n \to \mathbb{R}^n$ ist die orthogonale Projektion auf $(\ker A)^\perp \subset \mathbb{R}^n$, insbesondere $Q^2 = Q$

3) A und A^+ induzieren zueinander Inverse Isomorphismen

$$\begin{tikzcd}
(\ker A)^\perp \arrow[r, shift left, "A"] \arrow[d, phantom, "\cap" rotate=90] & \text{Bild } A \arrow[l, shift left, "A^+"] \arrow[d, phantom, "\cup"] \\
\mathbb{R}^n \arrow[r, "A"] & \mathbb{R}^m
\end{tikzcd}$$

4) $A^+AA^+ = A^+$, $AA^+A = A$

$(A^+)^+ = A$

5) Ist A invertierbar (also $n = m$) dann gilt $A^+ = A^{-1}$.

6) Ist ${}^tUAV = \Sigma$ die Singulärwertzerlegung von A, dann gilt $A^+ = U\Sigma^+ {}^tV$

wobei zu

$$\Sigma = \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r & \\ & & & 0 \end{pmatrix}; \sigma_r > 0 \quad \Sigma^+ = \left(\begin{array}{ccc|c} \frac{1}{\sigma_1} & & & \\ & \ddots & & 0 \\ & & \frac{1}{\sigma_r} & \\ \hline & 0 & & 0 \end{array}\right)$$

Beweis:

1) 2) 3) sind klar aus der Konstruktion

4) AA^+ ist die Projektion auf Bild A.

$\Rightarrow A^+AA^+ = A^+$ denn der erste Schritt von A^+ ist die Projektion auf das Bild

A^+A ist die Projektion auf $(\ker A)^\perp$

$\Rightarrow AA^+A = A$, da ~~...~~ alle Punkte von $x + \ker A$ unter A das gleiche Bild haben

$\ker(A^+) = (\text{Bild } A)^\perp$ da A^+ mit der orthogonalen Projektion beginnt

$\text{Bild}(A^+) = (\ker A)^\perp$

Also $(A^+)^+$ beginnt mit der orthogonalen Projektion auf $(\ker A)^\perp$

$(\ker A)^\perp \to (\ker A^+)^\perp = (\text{Bild } A)^{\perp\perp} = \text{Bild } A \subset \mathbb{R}^m$

$\Rightarrow A^{++} = A$

5) klar

6) Unsere geometrische Definition von A^+ macht klar, dass die Abbildung unabhängig von der Wahl eines Orthonormalbasis definiert ist.

Also ${}^tU A V = B \Rightarrow {}^tU A^+ V = B^+$, Σ^+ hat die behauptete Gestalt