Numerische Mathematik für Elektrotechniker

Singulärwertzerlegung/Nichtlineare Gleichungssysteme I

Benjamin Berkels

Karl-Heinz Brakhage, Thomas Jankuhn, Christian Löbbert

Institut für Geometrie und Praktische Mathematik
RWTH Aachen

Wintersemester 2019/2020

Zusammenfassung der letzten Vorlesung

Definition

Zu gegebenen $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ und $\boldsymbol{b} \in \mathbb{R}^{m}$, bestimme $\boldsymbol{x}^* \in \mathbb{R}^{n}$, für dass

$$\|\boldsymbol{A}\,\boldsymbol{x}^* - \boldsymbol{b}\|_2 = \min_{\boldsymbol{x} \in \mathbb{R}^n} \|\boldsymbol{A}\,\boldsymbol{x} - \boldsymbol{b}\|_2$$

gilt. Diese Problemstellung heißt das lineare Ausgleichsproblem.

Normalgleichungen:

$$\boldsymbol{A}^T \boldsymbol{A}\,\boldsymbol{x} = \boldsymbol{A}^T \boldsymbol{b}$$

Satz 4.5.

$\boldsymbol{x}^* \in \mathbb{R}^n$ ist genau dann Lösung des linearen Ausgleichsproblems, wenn $\boldsymbol{x}^*$ Lösung der Normalgleichungen ist.
Das System der Normalgleichungen hat stets mindestens eine Lösung. Sie ist genau dann eindeutig, wenn $\mathbf{Rang}(\boldsymbol{A}) = \boldsymbol{n}$ gilt.

Zusammenfassung der letzten Vorlesung

Lösung x^* über QR-Zerlegung

- Bestimme die QR-Zerlegung von A

$$QA = R = \begin{pmatrix} \diagdown \\ 0 \\ \hline \mathbf{0} \end{pmatrix} = \begin{pmatrix} \widetilde{R} \\ \mathbf{0} \end{pmatrix} \begin{matrix} \}n \\ \}m-n \end{matrix} \qquad (\widetilde{R} \in \mathbb{R}^{n \times n}),$$

und berechne $b_1 \in \mathbb{R}^n$ und $b_2 \in \mathbb{R}^{m-n}$ aus

$$Qb = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}.$$

$\|Qx\|_2 = \|x\|_2$

- Löse $\widetilde{R}\, x^* = b_1$ mittels Rückwärtseinsetzen.
- Die Norm des Residuums $\|A\, x^* - b\|_2$ ist $\|b_2\|_2$.

Zusammenfassung der letzten Vorlesung

Kondition des linearen Ausgleichsproblems

Für Störungen in b gilt:
$$\frac{\|\tilde{x} - x^*\|_2}{\|x^*\|_2} \leq \frac{\kappa_2(A)}{\cos\Theta} \frac{\|\tilde{b} - b\|_2}{\|b\|_2}.$$

wobei

$\langle A x^*, b - A x^* + A x^* \rangle$

$$\cos\Theta = \frac{\langle A\,x^*, b\rangle}{\|A\,x^*\|_2\,\|b\|_2} \overset{b - Ax^* \perp Ax^*}{=} \frac{\|A\,x^*\|_2}{\|b\|_2}$$

Problem schlecht konditioniert für $\Theta \approx \frac{\pi}{2}$ oder $\kappa_2(A) \gg 1$.

Stabilität der Lösungsverfahren

- Die Lösung der Normalgleichung über die Cholesky-Zerlegung von $A^T A$ ist instabil selbst für $\Theta \approx 0$ falls $\kappa_2(A)^2$ der Ordnung **eps**$^{-1}$ ist (Grund $\kappa_2(A^T A) = \kappa_2(A)^2$).
- Die Lösung über die QR-Zerlegung ist stabil.

Heute in der Vorlesung

Themen: Dahmen & Reusken Kap 4.1-4.4,4.7,5.1-5.2

- Singulärwertzerlegung (SVD)

Kapitel 5: Nichtlineare Gleichungssysteme

- Einleitung: Problemstellung

Was Sie mitnehmen sollten:

- Definition, Eigenschaften und Anwendungen der SVD

Motivation

Lösung des linearen Ausgleichsproblems (für $\mathbf{Rang}(A) = n$):

$$x^* = (A^T A)^{-1} A^T b.$$

Problem:
- $(A^T A)^{-1}$ existiert nicht für $\mathbf{Rang}(A) < n$
- Lösung des linearen Ausgleichsproblems nicht eindeutig
  $\Rightarrow$ zusätzliche Auswahlbedingung

$\mathrm{Rang}(A) < n \quad \Rightarrow \exists y \neq 0 \quad A^T A y = 0$

$\Rightarrow A^T A(x + \alpha y) = A^T A x + 0 \qquad \alpha \in \mathbb{R}$

$\min_{x \in L} \|x\|_2$ wobei $L := \operatorname{argmin}_{x \in \mathbb{R}^n} \|Ax - b\|_2$

Motivation

Lösung des linearen Ausgleichsproblems (für $\mathbf{Rang}(A) = n$):

$$x^* = (A^T A)^{-1} A^T b.$$

Problem: – $(A^T A)^{-1}$ existiert nicht für $\mathbf{Rang}(A) < n$
– Lösung des linearen Ausgleichsproblems nicht eindeutig
$\Rightarrow$ zusätzliche Auswahlbedingung

Allgemeines Lineares Ausgleichsproblem

Für $b \in \mathbb{R}^m$ und $A \in \mathbb{R}^{m \times n}$ hat die Aufgabe

Bestimme x^* mit minimaler Euklidischer Norm, für das $\|Ax^* - b\|_2 = \min_{x \in \mathbb{R}^n} \|Ax - b\|_2$ gilt

eine eindeutige Lösung. Diese kann bestimmt werden mittels

- QR-Zerlegung mit Pivotisierung $A = QR = Q \begin{pmatrix} \nabla \\ 0 \end{pmatrix}$

Motivation

Lösung des linearen Ausgleichsproblems (für $\mathbf{Rang}(A) = n$):

$$x^* = (A^T A)^{-1} A^T b.$$

Problem:
- $(A^T A)^{-1}$ existiert nicht für $\mathbf{Rang}(A) < n$
- Lösung des linearen Ausgleichsproblems nicht eindeutig
  $\Rightarrow$ zusätzliche Auswahlbedingung

Allgemeines Lineares Ausgleichsproblem

Für $b \in \mathbb{R}^m$ und $A \in \mathbb{R}^{m \times n}$ hat die Aufgabe

Bestimme x^* mit minimaler Euklidischer Norm, für das $\|Ax^* - b\|_2 = \min_{x \in \mathbb{R}^n} \|Ax - b\|_2$ gilt

eine eindeutige Lösung. Diese kann bestimmt werden mittels

- QR-Zerlegung mit Pivotisierung
- Singulärwertzerlegung (SVD = Singular Value Decomposition)

Singulärwertzerlegung

Satz 4.27.

Zu jeder Matrix $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ existieren orthogonale Matrizen $\boldsymbol{U} \in \mathbb{R}^{m \times m}$, $\boldsymbol{V} \in \mathbb{R}^{n \times n}$ und eine Diagonalmatrix

$$\boldsymbol{\Sigma} := \mathbf{diag}_{mn}(\boldsymbol{\sigma}_1, \ldots, \boldsymbol{\sigma}_p) \in \mathbb{R}^{m \times n}, \quad \boldsymbol{p} = \min(\boldsymbol{m}, \boldsymbol{n}),$$

mit

$$\boldsymbol{\sigma}_1 \geq \boldsymbol{\sigma}_2 \geq \cdots \geq \boldsymbol{\sigma}_p \geq 0,$$

so dass

$$\boldsymbol{A} = \boldsymbol{U}\,\boldsymbol{\Sigma}\,\boldsymbol{V}^T.$$

$A = A^T \Rightarrow \exists U$ orth. $A = U D U^T$

Singulärwertzerlegung

Satz 4.27.

Zu jeder Matrix $A \in \mathbb{R}^{m\times n}$ existieren orthogonale Matrizen $U \in \mathbb{R}^{m\times m}$, $V \in \mathbb{R}^{n\times n}$ und eine Diagonalmatrix

$$\Sigma := \mathrm{diag}_{mn}(\sigma_1, \ldots, \sigma_p) \in \mathbb{R}^{m\times n}, \quad p = \min(m, n),$$

mit

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_p \geq 0,$$

$m > n$

$$\begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \\ & 0 & \end{pmatrix}$$

so dass

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

Hierbei heißen:

- Singulärwerte von A : $\sigma_i,\ i = 1, \ldots, p$
- Linkssingulärvektoren : Spalten von $U = [u_1\, u_2 \cdots u_m]$
- Rechtssingulärvektoren: Spalten von $V = [v_1\, v_2 \cdots v_n]$

Singulärwertzerlegung – Geometrische Interpretation

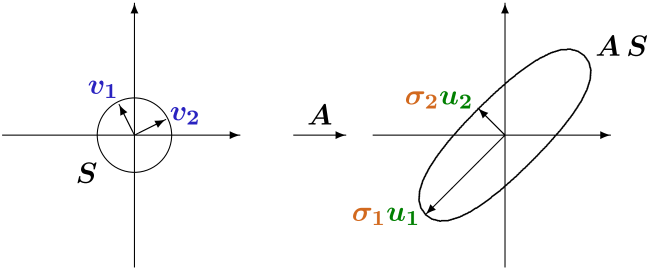
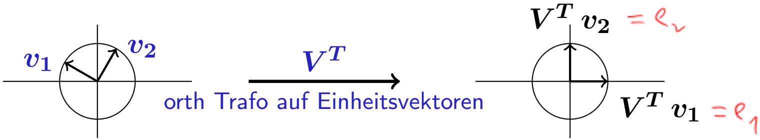


Abbildung der Einheitskugel $S = \{x : \|x\|_2 = 1\}$ durch $A = U\Sigma V^T$

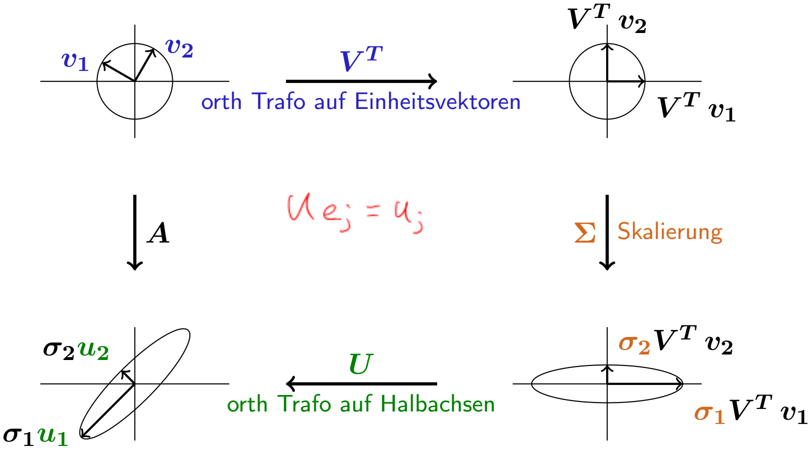
- Singulärwerte: Längen σ_1, σ_2 der Halbachsen von AS.
- Linkssingulärvektoren: Einheitsvektoren $\{u_1, u_2\}$ in Richtung der Halbachsen von AS
- Rechtssingulärvektoren: Einheitsvektoren $\{v_1, v_2\} \in S$, Urbild der Halbachsen von AS, so dass $Av_j = \sigma_j u_j$.

Singulärwertzerlegung – Geometrische Interpretation



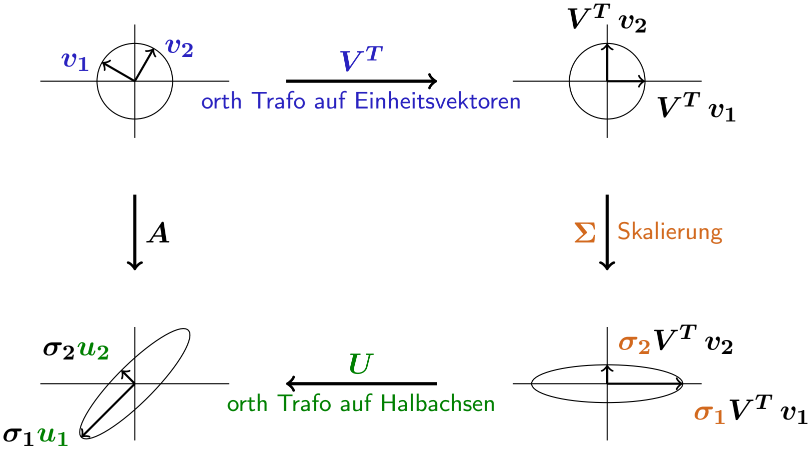
$$V^T v_j = e_j$$

Singulärwertzerlegung – Geometrische Interpretation



$Ue_j = u_j$

Singulärwertzerlegung – Geometrische Interpretation



Singulärwertzerlegung

Fall: $m = 4,\ n = 2,\ \mathrm{Rang}(A) = 2$

- (Volle) SVD: $U \in \mathbb{R}^{m\times m}$, $\Sigma \in \mathbb{R}^{m\times n}$, $V \in \mathbb{R}^{n\times n}$

$$\underbrace{\begin{pmatrix} * & * \\ * & * \\ * & * \\ * & * \end{pmatrix}}_{A} = \underbrace{\begin{pmatrix} * & * & * & * \\ * & * & * & * \\ * & * & * & * \\ * & * & * & * \end{pmatrix}}_{U} \underbrace{\begin{pmatrix} * & 0 \\ 0 & * \\ 0 & 0 \\ 0 & 0 \end{pmatrix}}_{\Sigma} \underbrace{\begin{pmatrix} * & * \\ * & * \end{pmatrix}}_{V^T}$$

- Reduzierte SVD $m \geq n$: $\hat{U} \in \mathbb{R}^{m\times n}$, $\hat{\Sigma} \in \mathbb{R}^{n\times n}$

$$\underbrace{\begin{pmatrix} * & * \\ * & * \\ * & * \\ * & * \end{pmatrix}}_{A} = \underbrace{\begin{pmatrix} * & * \\ * & * \\ * & * \\ * & * \end{pmatrix}}_{\hat{U}} \underbrace{\begin{pmatrix} * & 0 \\ 0 & * \end{pmatrix}}_{\hat{\Sigma}} \underbrace{\begin{pmatrix} * & * \\ * & * \end{pmatrix}}_{V^T}$$

Singulärwertzerlegung

- Annahme: $m \geq n$

  Falls $m < n$, betrachte SVD von A^T

$$A^T = U\,\Sigma\,V^T \quad \Rightarrow \quad A = V\,\Sigma^T\,U^T$$

- Singulärwerte sind reell und nichtnegativ
- Konvention
  - $\sigma_{\max} = \sigma_1$ größter Singulärwert
  - $\sigma_{\min} = \sigma_n$ kleinster Singulärwert
  - der Größe nach geordnet

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0$$

- SVD ist numerisch rückwärtsstabil berechenbar

$$A + \Delta A = \tilde{U}\,\tilde{\Sigma}\,\tilde{V}^T \quad \text{mit} \quad \|\Delta A\|_2 = \mathcal{O}(\text{eps})$$

Eigenschaften

Sei $A = U\,\Sigma\,V^T$ eine Singulärwertzerlegung von $A \in \mathbb{R}^{m\times n}$ mit Singulärwerten

$$\sigma_1 \geq \cdots \geq \sigma_r > \sigma_{r+1} = \cdots = \sigma_n = 0.$$

Dann gilt:

- für $i = 1, \ldots, n$.

$$Av_i = \sigma_i u_i,$$
$$A^T u_i = \sigma_i v_i,$$

$A v_i = U \Sigma V^T v_i = U \Sigma e_i = U \sigma_i e_i = \sigma_i u_i$

Eigenschaften

Sei $A = U\,\Sigma\,V^T$ eine Singulärwertzerlegung von $A \in \mathbb{R}^{m\times n}$ mit Singulärwerten

$$\sigma_1 \geq \cdots \geq \sigma_r > \sigma_{r+1} = \cdots = \sigma_n = 0.$$

Dann gilt:

- für $i = 1, \ldots, n$.

$$Av_i = \sigma_i u_i,$$
$$A^T u_i = \sigma_i v_i,$$

- Der **Rang** A ist die Anzahl der von Null verschiedenen Singulärwerte.

$$\mathrm{Rang}(A) = r$$

$\Rightarrow \mathrm{Rang}(A) = \mathrm{Rang}(\Sigma)$

$\mathrm{Rang}(A) = \mathrm{Rang}(BA) = \mathrm{Rang}(AC)$ B, C invertierbar

Eigenschaften

Sei $A = U\,\Sigma\,V^T$ eine Singulärwertzerlegung von $A \in \mathbb{R}^{m\times n}$ mit Singulärwerten

$$\sigma_1 \geq \cdots \geq \sigma_r > \sigma_{r+1} = \cdots = \sigma_n = 0.$$

Dann gilt:

- $\|A\|_2 = \sigma_1 = \sigma_{\max}$

$$\|A\|_2 = \max_{x\neq 0} \frac{\|Ax\|_2}{\|x\|_2} \overset{x=Vy}{=} \max_{y\neq 0} \frac{\|U\Sigma y\|_2}{\|Vy\|_2} = \max_{y\neq 0} \frac{\|\Sigma y\|_2}{\|y\|_2} = \sigma_1$$

Eigenschaften

Sei $A = U\,\Sigma\,V^T$ eine Singulärwertzerlegung von $A \in \mathbb{R}^{m\times n}$ mit Singulärwerten

$$\sigma_1 \geq \cdots \geq \sigma_r > \sigma_{r+1} = \cdots = \sigma_n = 0.$$

Dann gilt:

- $\|A\|_2 = \sigma_1 = \sigma_{\max}$
- $\|A^{-1}\|_2 = \dfrac{1}{\sigma_n} = \dfrac{1}{\sigma_{\min}}$,
  falls A eine reguläre $n \times n$ Matrix ist
- $\kappa_2(A) = \|A\|_2\|A^{-1}\|_2 = \dfrac{\sigma_1}{\sigma_n}$,
  falls A eine reguläre $n \times n$ Matrix ist
- $\kappa_2(A) = \|A\|_2\|A^{+}\|_2 = \dfrac{\sigma_1}{\sigma_r} = \dfrac{\sigma_{\max}}{\sigma_{\min}}$,
  falls $\mathbf{Rang}(A) = n$. Hier ist A^+ die Pseudoinverse.

$$\kappa_2(A) = \max_{x\neq 0} \frac{\|Ax\|_2}{\|x\|_2} \Big/ \min_{x\neq 0} \frac{\|Ax\|_2}{\|x\|_2}$$

Eigenschaften

Sei $A = U\,\Sigma\,V^T$ eine Singulärwertzerlegung von $A \in \mathbb{R}^{m\times n}$ mit Singulärwerten

$$\sigma_1 \geq \cdots \geq \sigma_r > \sigma_{r+1} = \cdots = \sigma_n = 0.$$

Dann gilt:

- Die Pseudoinverse $A^+ \in \mathbb{R}^{n\times m}$ ist definiert durch
$$A^+ = V\,\Sigma^+\,U^T$$
mit $\Sigma^+ = \mathrm{diag}_{nm}(\sigma_1^{-1}, \ldots, \sigma_r^{-1}, 0, \ldots, 0) \in \mathbb{R}^{n\times m}$.

Es gilt

$$AA^+ = U\Sigma\Sigma^+U^T = U\mathrm{diag}(\underbrace{1, \ldots, 1}_{r \text{ mal}}, 0, \ldots, 0)U^T \in \mathbb{R}^{m\times m},$$

$$A^+A = V\Sigma^+\Sigma V^T = V\mathrm{diag}(\underbrace{1, \ldots, 1}_{r \text{ mal}}, 0, \ldots, 0)V^T \in \mathbb{R}^{n\times n}.$$

Eigenschaften

Sei $A = U\,\Sigma\,V^T$ eine Singulärwertzerlegung von $A \in \mathbb{R}^{m\times n}$ mit Singulärwerten

$$\sigma_1 \geq \cdots \geq \sigma_r > \sigma_{r+1} = \cdots = \sigma_n = 0.$$

Dann gilt:

- Die strikt positiven Singulärwerte sind die Wurzeln der strikt positiven Eigenwerte von $A^T A$:

$$\{\sigma_i \mid i = 1,\ldots,r\} = \left\{ \sqrt{\lambda_i(A^T A)} \,\middle|\, i = 1,\ldots,r \right\}$$

$$A^T A = V \Sigma^T \cancel{U^T U} \Sigma V^T = V \underbrace{\Sigma^T}_{n\times m} \underbrace{\Sigma}_{m\times n} V^T = V \begin{pmatrix} \sigma_1^2 & & \\ & \ddots & \\ & & \sigma_n^2 \end{pmatrix} V^T$$

Eigenschaften

Sei $\boldsymbol{A} = \boldsymbol{U}\,\boldsymbol{\Sigma}\,\boldsymbol{V}^T$ eine Singulärwertzerlegung von $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ mit Singulärwerten

$$\sigma_1 \geq \cdots \geq \sigma_r > \sigma_{r+1} = \cdots = \sigma_n = 0.$$

Dann gilt:

- Die strikt positiven Singulärwerte sind die Wurzeln der strikt positiven Eigenwerte von $\boldsymbol{A}^T\boldsymbol{A}$:

$$\{\sigma_i \,|\, i = 1, \ldots, r\} = \left\{ \sqrt{\lambda_i(\boldsymbol{A}^T\,\boldsymbol{A})} \,\middle|\, i = 1, \ldots, r \right\}$$

- Die Singulärwerte sind gleich dem Absolutbetrag der Eigenwerte von $\boldsymbol{A}$, falls $\boldsymbol{A} = \boldsymbol{A}^T$. $\Rightarrow A^T A = A^2$

EW(A^2) sind die Quadrate der EW von A

Eigenschaften

Sei $A = U\Sigma V^T$ eine Singulärwertzerlegung von $A \in \mathbb{R}^{m\times n}$ mit Singulärwerten

$$\sigma_1 \geq \cdots \geq \sigma_r > \sigma_{r+1} = \cdots = \sigma_n = 0.$$

Dann gilt:

- Die strikt positiven Singulärwerte sind die Wurzeln der strikt positiven Eigenwerte von A^TA:

$$\{\sigma_i \,|\, i = 1,\dots,r\} = \left\{ \sqrt{\lambda_i(A^TA)} \,\middle|\, i = 1,\dots,r \right\}$$

- Die Singulärwerte sind gleich dem Absolutbetrag der Eigenwerte von A, falls $A = A^T$.
- Für $A \in \mathbb{R}^{n\times n}$ gilt

$$|\det(A)| = \prod_{i=1}^{n} \sigma_i$$

$\det(A) = \underbrace{\det(U)}_{=\pm 1} \det(\Sigma) \cdot \underbrace{\det(V^T)}_{=\pm 1}$

Eigenschaften – Niedrigrang-Approximation

Sei $A = U\,\Sigma\,V^T$ eine Singulärwertzerlegung von $A \in \mathbb{R}^{m\times n}$,
mit $r = \mathbf{Rang}(A)$ und Singulärwerten

$$\sigma_1 \geq \cdots \geq \sigma_r > \sigma_{r+1} = \cdots = \sigma_n = 0.$$

Für $k \leq r$ definiere $\quad A_k = U\,\Sigma_k\,V^T$

mit $\Sigma_k = \mathbf{diag}_{mn}(\sigma_1, \ldots, \sigma_k, 0, \ldots, 0) \in \mathbb{R}^{m\times n}$.

Dann gilt:

- $\mathbf{Rang}(A_k) = k$.
- Die Distanz zwischen A_k und A ist

$$\|A - A_k\|_2 = \sigma_{k+1}.$$

$$A - A_k = U \begin{pmatrix} 0 & & & & & \\ & \ddots & & & & \\ & & 0 & & & \\ & & & \sigma_{k+1} & & \\ & & & & \ddots & \\ & & & & & \sigma_r \\ & & & & & & 0 \end{pmatrix} V^T$$

Eigenschaften – Niedrigrang-Approximation

Sei $A = U\,\Sigma\,V^T$ eine Singulärwertzerlegung von $A \in \mathbb{R}^{m\times n}$,
mit $r = \mathbf{Rang}(A)$ und Singulärwerten

$$\sigma_1 \geq \cdots \geq \sigma_r > \sigma_{r+1} = \cdots = \sigma_n = 0.$$

Für $k \leq r$ definiere $\quad A_k = U\,\Sigma_k\,V^T$

mit $\Sigma_k = \mathbf{diag}_{mn}(\sigma_1, \ldots, \sigma_k, 0, \ldots, 0) \in \mathbb{R}^{m\times n}$.

Dann gilt:

- $\mathbf{Rang}(A_k) = k$.
- Die Distanz zwischen A_k und A ist
$$\|A - A_k\|_2 = \sigma_{k+1}.$$
- A_k ist die beste Approximation an A vom Rang $\leq k$
$$\|A - A_k\|_2 = \min_{\mathbf{Rang}(B)\leq k} \|A - B\|_2.$$

Lineares Ausgleichsproblem

Für $A \in \mathbb{R}^{m \times n}$ mit $\mathbf{Rang}(A) = n$ bestimme x^*, so dass

$$x^* = \arg\min_{x \in \mathbb{R}^n} \|A\,x - b\|_2.$$

Lösung mittels SVD: $A = U\,\Sigma\,V^T$

$$\begin{aligned} \|Ax - b\|_2 &= \|U^T\,A\,(V\,V^T)\,x - U^T\,b\|_2 \\ &= \|\Sigma\,(V^T\,x) - (U^T\,b)\|_2 \\ &= \|\Sigma\,y - (U^T\,b)\|_2 \end{aligned}$$

Dann ist $x^* = V\,y$ mit $y = \Sigma^+\,(U^T\,b)$ die eindeutige Lösung.

$\Sigma^+ = \mathrm{diag}_{nm}(\sigma_1^{-1}, \ldots, \sigma_n^{-1}) \in \mathbb{R}^{n \times m}$

Lineares Ausgleichsproblem

Für $A \in \mathbb{R}^{m\times n}$ mit $\mathrm{Rang}(A) = n$ bestimme x^*, so dass

$$x^* = \arg\min_{x\in\mathbb{R}^n} \|A\,x - b\|_2.$$

Lösung mittels SVD: $A = U\,\Sigma\,V^T$

$$\begin{aligned} \|Ax - b\|_2 &= \|U^T\,A\,(V\,V^T)\,x - U^T\,b\|_2 \\ &= \|\Sigma\,(V^T\,x) - (U^T\,b)\|_2 \\ &= \|\Sigma\,y - (U^T\,b)\|_2 \end{aligned}$$

Dann ist $x^* = V\,y$ mit $y = \Sigma^+\,(U^T\,b)$ die eindeutige Lösung.
$\Rightarrow x^* = V\,\Sigma^+\,(U^T\,b) = A^+b$
Dies gilt auch falls $\mathrm{Rang}(A) < n$.

Lineares Ausgleichsproblem

Falls $\mathbf{Rang}(A) = r < n$, ist die Pseudoinverse $A^+ \in \mathbb{R}^{n\times m}$

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

mit

$$\Sigma^+ = \mathrm{diag}_{nm}(\sigma_1^{-1}, \ldots, \sigma_r^{-1}, 0, \ldots, 0) \in \mathbb{R}^{n\times m}.$$

Lösung des allgemeinen linearen Ausgleichsproblems:

$$x^* = A^+\,b.$$

Numerischer Rang

Der Rang einer Matrix wird in der Praxis fast immer über die Singulärwerte bestimmt. Falls

$$\sigma_1 \geq \cdots \geq \sigma_r > \sigma_{r+1} = \cdots = \sigma_n = 0$$

so gilt $\textbf{Rang}(A) = r$.
Problem: Rundungsfehler aufgrund der Maschinengenauigkeit

$$A \to \tilde{A}, \quad \sigma_i \to \tilde{\sigma}_i, \; i = 1, \ldots, n.$$

$\Rightarrow \textbf{Rang}(\tilde{A})$ häufig $> r$, Abfrage "$\sigma_k = 0$" nicht sinnvoll.

Numerischer Rang

Der numerische Rang der Matrix $\tilde{A} \in \mathbb{R}^{m \times n}$ ist definiert als

$$\textbf{Rang}_{\textbf{num}}(\tilde{A}) = \min\{1 \leq k \leq n \,|\, \tilde{\sigma}_{k+1} \leq \tilde{\sigma}_1 \sqrt{mn}\, \text{eps}\}.$$

Bildkomprimierung

Singulärwertzerlegung von A

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

kann geschrieben werden als

$$A = \sum_{i=1}^{p} \sigma_i\, u_i\, v_i^T = \sigma_1\, u_1\, v_1^T + \sigma_2\, u_2\, v_2^T + \cdots + \sigma_p\, u_p\, v_p^T$$

mit $p = \min(m, n)$.

Niedrigrang-Approximation der Matrix A vom Rang k mit $k \leq p$, ist

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

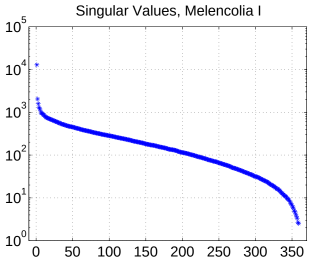
und $\|A - A_k\|_2 = \sigma_{k+1}$.

(Verlustbehaftete) Datenkompression:

$\Rightarrow$ Speicherbedarf für A_k ist $k(m+n+1)$ vs. mn für A.

Bildkomprimierung

Betrachte Schwarz-Weiß-Bild $\boldsymbol{B}(\boldsymbol{x},\boldsymbol{y})$ als Matrix $\boldsymbol{A}$:
Einträge entsprechen Graustufen des jeweiligen Pixels.

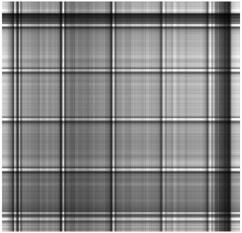


Je nach Anwendung:

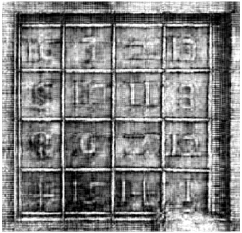
- Principal Component Analysis (PCA),
- Proper Orthogonal Decomposition (POD),
- Karhunen-Loéve Decomposition.

Bildkomprimierung

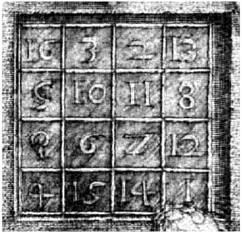
k = 1, compression = 0.00548



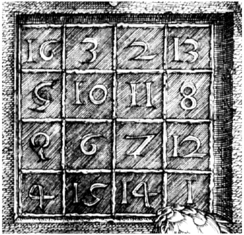
k = 20, compression = 0.11



k = 40, compression = 0.219



Original: A. Duerer, Melencolia I



Zusammenfassung

Singulärwertzerlegung wird eingesetzt für

- Bestimmung des (numerischen) Rangs einer Matrix
- Berechnung von Norm und Konditionszahl
- Niedrigrang-Approximation

Viele Praktische Anwendungen:

- Signalverarbeitung
- Bildverarbeitung
- Statistische Auswertungen
- Modelordnungsreduktion
- Inverse Probleme
- Regularisierung schlecht gestellter Probleme

Verständnisfragen

Für $\boldsymbol{A} \in \mathbb{R}^{\boldsymbol{m}\times\boldsymbol{n}}$ betrachten wir das lineare Ausgleichsproblem: Bestimme $\boldsymbol{x}^*$ mit minimaler 2-Norm so, dass $||\boldsymbol{A}\boldsymbol{x}^* - \boldsymbol{b}||_2 = \min_{\boldsymbol{x}\in\mathbb{R}^n} ||\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}||_2$.

$y = \Sigma^+(u^T b) = \begin{pmatrix} \frac{3}{200} \\ 0 \end{pmatrix}$

$x_1^* = V y = \left(\frac{9}{500} \right)$

f — Es seien $\boldsymbol{\sigma_1}$ der größte und $\boldsymbol{\sigma_r}$ der kleinste (positive) Singulärwert von $\boldsymbol{A}$. Dann gilt: $||\boldsymbol{A}||_2 = \frac{\boldsymbol{\sigma_1}}{\boldsymbol{\sigma_r}}$.

w — Die Lösung $\boldsymbol{x}^*$ des linearen Ausgleichsproblem mit minimaler 2-Norm ist immer eindeutig.

Es sei $\boldsymbol{A} = \boldsymbol{U}\,\boldsymbol{\Sigma}\,\boldsymbol{V}^{\boldsymbol{T}}$ mit $\boldsymbol{U} = \begin{pmatrix} \frac{3}{5} & \frac{4}{5} \\ \frac{4}{5} & -\frac{3}{5} \end{pmatrix}$, $\boldsymbol{V} = \begin{pmatrix} \frac{3}{5} & -\frac{4}{5} \\ \frac{4}{5} & \frac{3}{5} \end{pmatrix}$ und

$\boldsymbol{\Sigma} = \begin{pmatrix} 20 & 0 \\ 0 & 0 \end{pmatrix}$ eine Singulärwertzerlegung von $\boldsymbol{A}$ und es sei

$\boldsymbol{b} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$. Bestimmen Sie $\boldsymbol{x}_1^*$.

Verständnisfragen

Für $A \in \mathbb{R}^{m\times n}$ betrachten wir das lineare Ausgleichsproblem: Bestimme x^* mit minimaler 2-Norm so, dass $||Ax^* - b||_2 = \min_{x\in\mathbb{R}^n} ||Ax - b||_2$.

f Es seien σ_1 der größte und σ_r der kleinste (positive) Singulärwert von A. Dann gilt: $||A||_2 = \frac{\sigma_1}{\sigma_r}$.

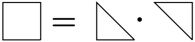
w Die Lösung x^* des linearen Ausgleichsproblem mit minimaler 2-Norm ist immer eindeutig.

Es sei $A = U\,\Sigma\,V^T$ mit $U = \begin{pmatrix} \frac{3}{5} & \frac{4}{5} \\ \frac{4}{5} & -\frac{3}{5} \end{pmatrix}$, $V = \begin{pmatrix} \frac{3}{5} & -\frac{4}{5} \\ \frac{4}{5} & \frac{3}{5} \end{pmatrix}$ und $\Sigma = \begin{pmatrix} 20 & 0 \\ 0 & 0 \end{pmatrix}$ eine Singulärwertzerlegung von A und es sei $b = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$. Bestimmen Sie x_1^*. 0.018

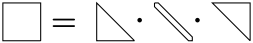
Matrix-Zerlegungen

$\boldsymbol{A}$ reguläre quadratische Matrizen

- $\boldsymbol{LR}$-Zerlegung (Gauß-Elimination)

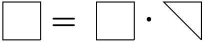


- Falls $\boldsymbol{A}$ s.p.d.: Cholesky $\boldsymbol{L}\,\boldsymbol{D}\,\boldsymbol{L}^{\boldsymbol{T}}$

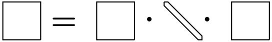


$\boldsymbol{A}$ beliebig (nicht notwendig quadratisch oder voller Rang)

- $\boldsymbol{QR}$-Zerlegung (Givens-Rotation, Householder-Spiegelung)



- Singulärwert-Zerlegung



1) $Ax = b \iff \underbrace{Ax - b}_{=: f(x)} = 0$

linear Gl.

2) A $m \times n$ $\quad m > n$

$\|Ax - b\|_2 \to \min$ $\quad$ lin. Ausgleich

3) $f : \mathbb{R}^n \to \mathbb{R}^n$ $\quad f(x) = 0$

4) $f : \mathbb{R}^n \to \mathbb{R}^m$ $\quad m > n$

$\|f(x)\|_2 \to \min$

Nichtlineare Gleichungssysteme

Kapitel 5:

Nichtlineare Gleichungssysteme

Motivation

1. Viele praktische Probleme führen auf nichtlineare Gleichungssysteme
2. Je realistischer ein (mathematisches) Modell, desto eher ist es nichtlinear:
   - Pendelschwingung: Auslenkungswinkel φ beschrieben durch

     $$\ddot{\varphi}(t) + \frac{g}{\ell}\varphi(t) = 0 \quad \text{vs.} \quad \ddot{\varphi}(t) + \frac{g}{\ell}\sin\big(\varphi(t)\big) = 0$$

     für kleine vs. große Auslenkungen.
   - Lineare vs. nichtlineare Diffusion: Temperatur u lößt

     $$u_t = \operatorname{div}(k\,\nabla u) \quad \text{vs.} \quad u_t = \operatorname{div}\big(k(u)\,\nabla u\big)$$

     mit Wärmeleitfähigkeit $k(u) = c_1 + c_2 u + c_3 u^3$.
   - Strömungsprobleme, Netzwerkanalyse, ...

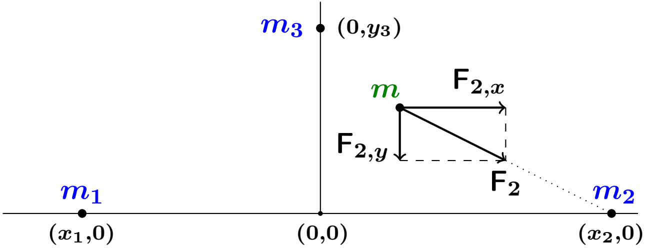
Beispiel 5.1.

Für die Gravitationskraft zwischen zwei Punktmassen M_a und M_b mit gegenseitigem Abstand r gilt (Newtons Gravitationsgesetz):

$$F = G \cdot \frac{M_a M_b}{r^2}$$

wobei $G = 6.67 \cdot 10^{-11}$Nm^2/kg (Gravitationskonstante).

Wie betrachten das folgende Gravitationsfeld:



Beispiel 5.1.

Gesucht: (x, y), so dass für eine Punktmasse m an der Stelle (x, y) die Gravitationskräfte F_i zwischen m und m_i, $i = 1, 2, 3$, im Gleichgewicht sind.

Hilfsgrößen mit $i = 1, 2, 3$ sind

$$r_i := \sqrt{(x - x_i)^2 + (y - y_i)^2}$$

$$F_i := G \cdot \frac{m_i m}{r_i^2}$$

$$F_{i,x} := \frac{F_i(x_i - x)}{r_i}$$

$$F_{i,y} := \frac{F_i(y_i - y)}{r_i}$$

Beispiel 5.1.

Die Gleichgewichtsbedingungen sind wie folgt:

$$F_{1,x} + F_{2,x} + F_{3,x} = 0 \quad \text{und} \quad F_{1,y} + F_{2,y} + F_{3,y} = 0$$

Hieraus ergibt sich das System

$$\begin{aligned} f_1(x,y) &= \sum_{i=1}^{3} \frac{m_i(x_i - x)}{\left((x - x_i)^2 + (y - y_i)^2\right)^{3/2}} = 0 \\ f_2(x,y) &= \sum_{i=1}^{3} \frac{m_i(y_i - y)}{\left((x - x_i)^2 + (y - y_i)^2\right)^{3/2}} = 0. \end{aligned}$$

Beispiel 5.2.

Statt der linearen Integralgleichung im Beispiel 3.3.

$$\boldsymbol{u(x) + 2\int_0^1 \cos(xt)\, u(t)\, dt = 2, \quad x \in [0,1]}$$

ist nun eine nichtlineare Integralgleichung zu lösen:

Gesucht ist eine Funktion $\boldsymbol{u(x) \geq 0}$, die die Integralgleichung

$$\boldsymbol{u(x) + 2\int_0^1 \cos(xt)\, u(t)^3\, dt = 2, \quad x \in [0,1]}$$

erfüllt.

Beispiel 5.2.

Das Problem wird, wie in Beispiel 3.3., auf dem Gitter

$$t_j = \left(j - \frac{1}{2}\right) h, \quad j = 1, \ldots, n, \quad h = \frac{1}{n}.$$

diskretisiert.

Man erhält dann die Gleichungen

$$u_i + 2\,h \sum_{j=1}^{n} \cos(t_i\, t_j)\, u_j^3 = 2, \quad i = 1, \ldots, n,$$

für die Unbekannten $u_i \approx u(t_i)$, $i = 1, \ldots, n$.

Problemstellung

Aufgabe

Zu gegebenem $f = \begin{pmatrix} f_1 \\ \vdots \\ f_n \end{pmatrix} : \mathbb{R}^n \to \mathbb{R}^n$ bestimme $x^* = \begin{pmatrix} x_1^* \\ \vdots \\ x_n^* \end{pmatrix}$,

so dass

$$\begin{aligned} f_1(x_1^*, \ldots, x_n^*) &= 0 \\ \vdots \quad &\vdots \quad \vdots \\ f_n(x_1^*, \ldots, x_n^*) &= 0 \end{aligned}$$

erfüllt ist.

Kompakte Darstellung:

$$f(x^*) = 0$$

Problemstellung

Aufgabe

Gegeben: $f : \mathbb{R}^n \to \mathbb{R}^n$.

Gesucht: $x^* \in \mathbb{R}^n$, so dass $f(x^*) = 0$.

- Lineare Gleichungssysteme: Sonderfall dieser Problemstellung

$$A\,x^* = b \quad \Leftrightarrow \quad f(x^*) := A\,x^* - b = 0.$$

- Der Spezialfall $n = 1$ wird oft als skalare Gleichung in einer Unbekannten bezeichnet.
- Hat man mehr (nichtlineare) Gleichungen als Unbekannte, d.h.

$$f : \mathbb{R}^n \to \mathbb{R}^m \text{ mit } m > n$$

erhält man ein nichtlineares Ausgleichsproblem
$\rightsquigarrow$ siehe nächstes Kapitel.