

# Proseminar-Vortrag: Pseudoinverse, Singulärwertzerlegung

von Sven Heins

gehalten am 08.06.2005

## 1 Problemstellung

### 1.1 Lineare Ausgleichsrechnung

Wir befinden uns im Bereich der **linearen Ausgleichsrechnung**. Dabei lautet die Problemstellung:

<p><b>Lineare Ausgleichsaufgabe</b> (Numerik Skript Formel (9.4)) Gegeben seien eine <math>m \times n</math> – <i>Matrix</i> <math>\mathbf{A}</math> mit <math>m &gt; n</math> und <math>\mathbf{b} \in \mathbb{R}^m</math>. Gesucht wird ein <math>\hat{\mathbf{x}} \in \mathbb{R}^n</math> mit</p> $\ \mathbf{A}\hat{\mathbf{x}} - \mathbf{b}\ _2 \leq \ \mathbf{A}\mathbf{x} - \mathbf{b}\ _2 \quad \forall \mathbf{x} \in \mathbb{R}^n.$	(1)
------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-----

### 1.2 Normalgleichungen

Eine Lösung dieses Problems liefern die Normalgleichungen. Man gewinnt sie wie folgt:

$$\begin{aligned} & \hat{\mathbf{x}} \text{ löst die Ausgleichsaufgabe: } \|\mathbf{A}\hat{\mathbf{x}} - \mathbf{b}\|_2 \stackrel{!}{=} \min. \\ \Leftrightarrow & \hat{\mathbf{x}} \text{ löst das Approximationsproblem:} \\ & \|\mathbf{A}\hat{\mathbf{x}} - \mathbf{b}\|_2 \leq \|\mathbf{y} - \mathbf{b}\|_2 \quad \forall \mathbf{y} \in \mathbf{A}(\mathbb{R}^n) \\ \Leftrightarrow & (\mathbf{A}\hat{\mathbf{x}} - \mathbf{b}, \mathbf{A}\mathbf{x}) = 0 \quad \forall \mathbf{x} \in \mathbb{R}^n \text{ (gem. Projektionssatz)} \\ \Leftrightarrow & \dots \\ \Leftrightarrow & \boxed{\mathbf{A}^T \mathbf{A} \hat{\mathbf{x}} = \mathbf{A}^T \mathbf{b} \quad \textit{Normalgleichungen der Ausgleichsrechnung}} \end{aligned}$$

## 2 Die Pseudoinverse

### 2.1 Umformulierung des Ausgleichsproblems

Aus der Numerik wissen wir, dass die Normalgleichungen zwar die Lösung der Ausgleichsrechnung liefern, jedoch numerisch ungünstig. Als erstes wollen wir das Problem

$$\|\mathbf{A}\hat{\mathbf{x}} - \mathbf{b}\|_2 \stackrel{!}{=} \min$$

formal bezeichnen als:

$$\boxed{\mathbf{x} = \mathbf{A}^+\mathbf{b}}$$

Aus den Normalgleichungen ergibt sich dann folgende Gleichung für  $\mathbf{A}^+$ :

$$\mathbf{A}^+ = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T.$$

Es gilt  $\mathbf{A}^+\mathbf{A} = \mathbf{I}$ . Aus diesem Grund nennt man  $\mathbf{A}^+$  auch *Pseudoinverse* von  $\mathbf{A}$ .

Man möchte diese *Pseudoinverse* auch für beliebige Matrizen  $\mathbf{A} \in M(m \times n, \mathbb{R})$  erweitern.

**Hierbei gibt es also keine Bedingung an den Rang und m, n wie in der Linearen Ausgleichsaufgabe (1).** Man hat also nicht zwangsläufig Eindeutigkeit für die Lösung dieses Problems.

Definiere dann:

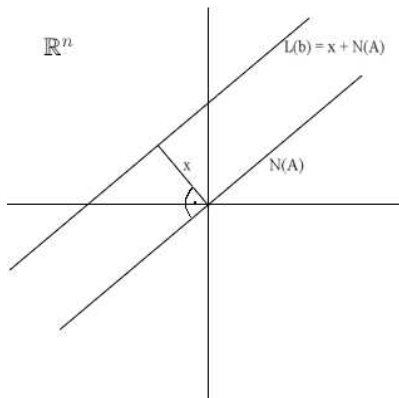
$$\bar{P} : \mathbb{R}^m \longrightarrow \text{Im}(\mathbf{A}) \subset \mathbb{R}^m$$

sei die orthogonale Projektion des  $\mathbb{R}^m$  auf den Bildraum von  $\mathbf{A}$ :  $\text{Im}(\mathbf{A})$

Dann bilden die Lösungen folgenden affinen Unterraum:

$$L(b) := \{\mathbf{x} \in \mathbb{R}^n \mid \|b - \mathbf{A}\mathbf{x}\| = \min\} = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{A}\mathbf{x} = \bar{P}b\}$$

Nun hat man also einen ganzen Raum, der die Lösungen des Problems  $\|\mathbf{A}\mathbf{x} - \mathbf{b}\| \stackrel{!}{=} \min$  beschreibt. Damit gibt man sich jedoch nicht zufrieden, da man eine eindeutige Lösung der Aufgabe benötigt. Also nimmt man sich das  $x \in L(b)$  mit der kleinsten euklidischen Norm. Folgendes Bild zeigt, wie man diese Lösung erhält:



Wie man sieht ist  $\mathbf{x}$  die orthogonale Projektion des Ursprungs auf den affinen Unterraum  $L(b)$ . Somit hat man ein eindeutiges  $\mathbf{x}$  als Lösung für das Problem  $\|\mathbf{Ax} - \mathbf{b}\| \stackrel{!}{=} \min$

## 2.2 Definition und Eigenschaften

### Definition 2.1:

Die *Pseudoinverse* einer Matrix  $\mathbf{A} \in M(m \times n, \mathbb{R})$  wird bezeichnet als  $\mathbf{A}^+$  mit der Eigenschaft, dass für alle  $b \in \mathbb{R}^m$  der Vektor  $\mathbf{x} = \mathbf{A}^+b$  die kleinste Lösung von  $\|\mathbf{Ax} - \mathbf{b}\|_2 \stackrel{!}{=} \min$  ist, d.h.

$$\mathbf{A}^+b \in N(\mathbf{A})^\perp \text{ und } \|\mathbf{AA}^+b - b\|_2 \stackrel{!}{=} \min$$

Folgendes Kommutatives Bild verdeutlicht den Zusammenhang:

	$\mathbf{A}$				
	$\mathbb{R}^n$	$\longleftrightarrow$		$\mathbb{R}^m$	
		$\mathbf{A}^+$			
	$P = \mathbf{A}^+\mathbf{A} \downarrow \uparrow i$			$i \uparrow \downarrow \bar{P} = \mathbf{AA}^+$	
	$\text{Im}(\mathbf{A}^+) = N(\mathbf{A})^\perp$	$\cong$		$\text{Im}(\mathbf{A})$	

(Hierbei ist  $i$  die jeweilige Inklusion, also die allgemeine (Teil-)Mengenbeziehung zwischen den entsprechenden Mengen)

Hieran sieht man, dass gilt

$$\bar{P} = \mathbf{AA}^+ \text{ und } P = \mathbf{A}^+\mathbf{A},$$

denn  $\bar{P}$  ist ja definiert als  $\bar{P} : \mathbb{R}^m \longrightarrow \text{Im}(\mathbf{A}) \subset \mathbb{R}^m$ . Somit ist auch klar, dass  $P$  die Projektion des  $\mathbb{R}^n$  auf das orthogonale Komplement  $N(\mathbf{A})^\perp$  des Nullraums beschreibt. Außerdem stellen wir noch folgende Eigenschaften der Pseudoinversen fest, durch die wir ihre Eindeutigkeit erhalten:

**Satz 2.2:**

Die *Pseudoinverse*  $\mathbf{A}^+ \in M(n \times m, \mathbb{R})$  einer Matrix  $\mathbf{A} \in M(m \times n, \mathbb{R})$  ist eindeutig charakterisiert durch folgende vier Eigenschaften (*Penrose – Axiome*):

i)  $(\mathbf{A}^+\mathbf{A})^T = \mathbf{A}^+\mathbf{A}$   
 ii)  $(\mathbf{A}\mathbf{A}^+)^T = \mathbf{A}\mathbf{A}^+$   
 iii)  $(\mathbf{A}^+\mathbf{A}\mathbf{A}^+) = \mathbf{A}^+$   
 iv)  $(\mathbf{A}\mathbf{A}^+\mathbf{A}) = \mathbf{A}$

**Beweis:**

Die Eigenschaften i) bis iv) werden durch  $\mathbf{A}^+$  erfüllt, denn  $\mathbf{A}^+\mathbf{A}$  und  $\mathbf{A}\mathbf{A}^+$  lassen sich auffassen als orthogonale Projektionen auf  $N(\mathbf{A})^\perp = \text{Im}(\mathbf{A}^+)$  bzw. auf  $\text{Im}(\mathbf{A})$ . Andersherum definieren i) bis iv) orthogonale Projektionen  $P := \mathbf{A}^+\mathbf{A}$  und  $\bar{P} := \mathbf{A}\mathbf{A}^+$ , weil

$$P^T = P = P^2 \text{ und } \bar{P}^T = \bar{P} = \bar{P}^2.$$

Aus iv) und  $\bar{P} = \mathbf{A}\mathbf{A}^+$  folgt außerdem, dass  $\bar{P}$  die orthogonale Projektion auf  $\text{Im}(\mathbf{A})$ , d.h.  $\text{Im}(\mathbf{A}) = \text{Im}(\bar{P})$ . Analog folgt aus iii) und  $P = \mathbf{A}^+\mathbf{A}$ , dass  $N(P) = N(\mathbf{A})$  ist.

Also sind die Projektoren  $P$  und  $\bar{P}$  unabhängig von  $\mathbf{A}^+$  eindeutig durch die Eigenschaften i) bis iv) bestimmt. Und daraus kann man auch schon die Eindeutigkeit von  $\mathbf{A}^+$  schließen: Erfüllen  $\mathbf{A}_1^+$  und  $\mathbf{A}_2^+$  die Bedingungen i) bis iv), so gilt

$$P = \mathbf{A}_1^+A = \mathbf{A}_2^+A \text{ und } \bar{P} = \mathbf{A}\mathbf{A}_1^+ = \mathbf{A}\mathbf{A}_2^+$$

und daher gilt

$$\mathbf{A}_1^+ \stackrel{iii)}{=} \mathbf{A}_1^+\mathbf{A}\mathbf{A}_1^+ = \mathbf{A}_2^+\mathbf{A}\mathbf{A}_2^+ \stackrel{iii)}{=} \mathbf{A}_2^+.$$

□

**Bemerkung (2.3):**

Gilt nur ein Teil der Penrose-Axiome, so spricht man von einer *verallgemeinerten Inversen*.

**2.3 Berechnung der kleinsten Lösung**

Die kleinste Lösung  $x = \mathbf{A}^+b$  für  $\mathbf{A} \in M(m \times n, \mathbb{R})$  und  $b \in \mathbb{R}^m$  wird über die QR-Zerlegung berechnet.

Sei  $p := \text{Rang}(\mathbf{A}) \leq \min(m, n)$  der Rang der Matrix  $\mathbf{A}$ . Durch orthogonale Transformation (z.B. Householder-Transformation) bringt man  $\mathbf{A}$  auf obere Dreiecksgestalt:

$$\mathbf{QA} = \left( \begin{array}{c|c} R & S \\ \hline 0 & 0 \end{array} \right)$$

hierbei ist  $\mathbf{R} \in M(p \times p, \mathbb{R})$  eine invertierbare obere Dreiecksmatrix und  $S \in M(p \times (n - p), \mathbb{R})$ . Analog zerlegt man

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \text{ mit } x_1 \in \mathbb{R}^p \text{ und } x_2 \in \mathbb{R}^{n-p}.$$

$$Qb = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \text{ mit } b_1 \in \mathbb{R}^p \text{ und } b_2 \in \mathbb{R}^{m-p}.$$

Damit ergibt sich folgendes

**Lemma (2.4):**

Mit den obigen Bezeichnungen ist  $x$  genau dann eine Lösung von  $\|\mathbf{Ax} - \mathbf{b}\|_2 \stackrel{!}{=} \min$ , falls

$$x_1 = R^{-1}b_1 - R^{-1}Sx_2.$$

**Beweis:**

Bezüglich orthogonalen Transformationen ist die Euklidische Norm invariant, also gilt

$$\|\mathbf{Ax} - \mathbf{b}\|^2 = \|\mathbf{QAx} - \mathbf{Qb}\|^2 = \|\mathbf{Rx}_1 + \mathbf{Sx}_2 - \mathbf{b}_1\|^2 + \|\mathbf{b}_2\|^2.$$

Diese Gleichung wird genau dann minimal, wenn

$$\mathbf{Rx}_1 + \mathbf{Sx}_2 - \mathbf{b}_1 = 0.$$

□

**Bemerkung:**

Im Fall  $p = \text{Rang}(\mathbf{A}) = n$  liegt ein überbestimmtes Gleichungssystem mit vollem Rang vor, das bereits in der Numerik behandelt wurde. Die Matrix  $S$  verschwindet in diesem Fall, und man erhält die Lösung  $x = x_1 = R^{-1}b_1$ . Im folgenden betrachtet man den Fall des unterbestimmten oder rangdefekten Gleichungssystems ( $p < n$ ).

**Lemma 2.5:**

Sei  $p < n$ ,  $V := R^{-1}S \in M(p \times (n - p), \mathbb{R})$  und  $u := R^{-1}b_1 \in \mathbb{R}^p$ . Dann ist die kleinste Lösung  $x$  von  $\|\mathbf{Ax} - \mathbf{b}\|_2 \stackrel{!}{=} \min$  gegeben durch  $x = (x_1, x_2) \in \mathbb{R}^p \times \mathbb{R}^{n-p}$  mit

$$(I + V^T V)x_2 = V^T u \text{ und } x_1 = u - Vx_2.$$

**Beweis:**

Nach Lemma 2.4 sind die Lösungen von  $\|\mathbf{Ax} - \mathbf{b}\|_2 \stackrel{!}{=} \min$  durch  $x_1 = u - Vx_2$  gekennzeichnet. Eingesetzt in  $\|x\|$  erhalten wir

$$\begin{aligned} \|x\|^2 &= \|x_1\|^2 + \|x_2\|^2 = \|u - Vx_2\|^2 + \|x_2\|^2 \\ &= \|u\|^2 - 2\langle u, Vx_2 \rangle + \langle Vx_2, Vx_2 \rangle + \langle x_2, x_2 \rangle \\ &= \|u\|^2 + \langle x_2, (I + V^T V)x_2 - 2V^T u \rangle =: \varphi(x_2). \end{aligned}$$

Dabei ist  $\varphi'(x_2) = -2V^T u + 2(I + V^T V)x_2$  und  $\varphi''(x_2) = 2(I + V^T V)$ . Da  $I + V^T V$  eine symmetrische, positiv definite (spd) Matrix ist, nimmt  $\varphi(x_2)$  sein Minimum für  $x_2$  mit  $\varphi'(x_2) = 0$  an, d.h.  $(I + V^T V)x_2 = V^T u$ . Das ist die Behauptung.

□

**2.4 Algorithmus:**

Berechnung der kleinsten Lösung  $\|\mathbf{Ax} - \mathbf{b}\|_2 \stackrel{!}{=} \min$  durch QR - Zerlegung:

1. QR-Zerlegung von  $\mathbf{A}$  mit  $p = \text{Rang}(\mathbf{A})$  wobei  $Q \in O(m)$ ,  $R \in M(p \times p, \mathbb{R})$  obere Dreiecksmatrix und  $S \in M(p \times (n - p), \mathbb{R})$ .
2. Berechne  $V \in M(p \times (n - p), \mathbb{R})$  aus  $RV = S$ .
3. Cholesky-Zerlegung von  $I + V^T V$

$$I + V^T V = LL^T,$$

wobei  $L \in M((n - p) \times (n - p), \mathbb{R})$  untere Dreiecksmatrix ist.

4.  $(b_1, b_2)^T := Qb$  mit  $b_1 \in \mathbb{R}^p$ ,  $b_2 \in \mathbb{R}^{n-p}$ .
5. Berechne  $u \in \mathbb{R}^p$  aus  $Ru = b_1$ .
6. Berechne  $x_2 \in \mathbb{R}^{n-p}$  aus  $LL^T x_2 = V^T u$ .
7. Setze  $x_1 := u - Vx_2$ .  
Dann ist  $x = (x_1, x_2)^T = A^+ b$ .

Man beachte, dass die Schritte 1. bis 3. für verschiedene rechte Seiten  $b$  nur einmal durchgeführt werden müssen.

Mit diesem Verfahren hat man also eine Möglichkeit, rangdefekte, singuläre Gleichungssysteme bestmöglich zu lösen. Eine weitere Möglichkeit erhält man mit der *Singulärwertzerlegung*, auf die wir im nächsten Kapitel eingehen wollen.

### 3 Singulärwertzerlegung singular value decomposition (SVD)

#### 3.1 Motivation

Mithilfe der *Singulärwertzerlegung* haben wir eine weitere Möglichkeit, rangdefekte, singuläre Gleichungssysteme zu lösen. Des Weiteren ist die *Singulärwertzerlegung* ein nützliches Mittel um Matrizen im Allgemeinen zu analysieren.

**Satz 3.1:**

Sei  $\mathbf{A} \in M(m \times n, \mathbb{R})$  eine beliebige reelle Matrix. Dann gibt es orthogonale Matrizen  $U \in \mathbf{O}(m)$  und  $V \in \mathbf{O}(n)$ , so dass gilt

$$U^T AV = \Sigma = \text{diag}(\sigma_1, \dots, \sigma_p) \in M(m \times n, \mathbb{R}),$$

wobei  $p = \min(m, n)$  und  $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_p \geq 0$ .

**Beweis:**

Es genügt zu zeigen, dass es  $U \in \mathbf{O}(m)$  und  $V \in \mathbf{O}(n)$  gibt, so dass

$$U^T AV = \begin{pmatrix} \sigma & 0 \\ 0 & B \end{pmatrix}.$$

Durch Induktion folgt dann die Behauptung. Sei  $\sigma := \|A\|_2 = \max_{\|x\|=1} \|Ax\|$ . Da das Maximum angenommen wird, gibt es  $v \in \mathbb{R}^n, u \in \mathbb{R}^m$ , so dass

$$Av = \sigma u \text{ und } \|u\|_2 = \|v\|_2 = 1.$$

Nun erweitert man  $v$  zu einer Orthonormalbasis  $\{v = V_1, \dots, V_n\}$  des  $\mathbb{R}^n$  bzw.  $u$  zu einer Orthonormalbasis  $\{u = U_1, \dots, U_m\}$  des  $\mathbb{R}^m$ .

Dann sind

$$V := \begin{bmatrix} V_1, \dots, V_n \end{bmatrix} \text{ und } U := \begin{bmatrix} U_1, \dots, U_m \end{bmatrix}$$

orthogonale Matrizen,  $V \in \mathbf{O}(n), U \in \mathbf{O}(m)$ , und  $U^T AV$  von der Form

$$A_1 := U^T AV = \begin{pmatrix} \sigma & w^T \\ 0 & B \end{pmatrix}.$$

mit  $w \in \mathbb{R}^{n-1}$ . Da

$$\left\| A_1 \begin{pmatrix} \sigma \\ w \end{pmatrix} \right\|_2^2 \geq (\sigma^2 + \|w\|_2^2)^2 \text{ und } \left\| \begin{pmatrix} \sigma \\ w \end{pmatrix} \right\|_2^2 = \sigma^2 + \|w\|_2^2,$$

gilt  $\sigma^2 = \|A\|_2^2 = \|A_1\|_2^2 \geq \sigma^2 + \|w\|_2^2$  und daher  $w = 0$ , also

$$U^T AV = \begin{pmatrix} \sigma & 0 \\ 0 & B \end{pmatrix}$$

□

**Definition 3.2:**

Die Zerlegung  $U^T AV = \Sigma$  heißt *Singulärwertzerlegung* von  $A$ , die  $\sigma_i$  sind die *Singulärwerte* von  $A$ .

Aufgrund der Invarianz der Euklidischen Norm  $\|\cdot\|_2$  unter den orthogonalen Transformationen  $U$  und  $V$  erhalten wir aus der Singulärwertzerlegung von  $A$  auch eine weitere Darstellung der Pseudoinversen  $A^+$  von  $A$ .

**Folgerung 3.3:**

Sei  $U^T AV = \Sigma$  die *Singulärwertzerlegung* einer Matrix  $A \in M(m \times n, \mathbb{R})$  mit  $p = \text{Rang}(A)$  und

$$\Sigma = \text{diag}(\sigma_1, \dots, \sigma_p, 0, \dots, 0).$$

Dann ist die Pseudoinverse  $A^+ \in M(n \times m, \mathbb{R})$  gerade

$$A^+ = V\Sigma^+U^T \text{ mit } \Sigma^+ = \text{diag}(\sigma_1^{-1}, \dots, \sigma_p^{-1}, 0, \dots, 0).$$

**Beweis:**

Man muss die Penrose-Axiome für die rechte Seite  $B := V\Sigma^+U^T$  zeigen. Dann wäre man fertig.

Da trivialerweise  $\Sigma^+$  die Pseudoinverse der Diagonalmatrix  $\Sigma$  ist, folgen sofort die Penrose-Axiome für  $B$ , da  $V^T V = I$  und  $U^T U = I$ .

□

**Folgerung 3.4:**

Sei  $U^T AV = \Sigma = \text{diag}(\sigma_1, \dots, \sigma_p)$  die Singulärwertzerlegung von  $A$  mit den Singulärwerten  $\sigma_1, \dots, \sigma_p$ , wobei  $p = \min(m, n)$ . Dann gilt:

1. Das lineare Ausgleichsproblem (1)  $\|\mathbf{Ax} - \mathbf{b}\|_2 \stackrel{!}{=} \min$  wird gelöst, indem man für die Matrix  $A$  die entsprechende Singulärwertmatrix  $\Sigma$  benutzt. Man hat es somit geschafft, zu einer Matrix, die singulär sein kann, die beste Lösung unter obiger Voraussetzung zu berechnen.
2. Sind  $U_i$  und  $V_i$  die Spalten von  $U$  bzw.  $V$ , so ist

$$AV_i = \sigma_i U_i \text{ und } A^T U_i = \sigma_i V_i \text{ für } i = 1, \dots, p.$$

3. Falls  $\sigma_1 \geq \dots \geq \sigma_r > \sigma_{r+1} = \dots = \sigma_p = 0$ , so gilt  $\text{Rang}(A) = r$ ,

$$\ker(A) = \text{span}\{V_{r+1}, \dots, V_n\} \text{ und } \text{Im}(A) = \text{span}\{U_1, \dots, U_r\}.$$

4. Die Euklidische Norm von  $A$  ist der größte Singulärwert, d.h.

$$\|A\|_2 = \sigma_1.$$



5. Für die Frobenius-Norm von  $\|A\|_F = (\sum_{i=1}^n \|A_i\|_2^2)^{\frac{1}{2}}$  gilt

$$\|A\|_F^2 = \sigma_1^2 + \dots + \sigma_p^2.$$

6. Die Kondition von  $A$  bzgl. der Euklidischen Norm ist der Quotient von größtem und kleinstem Singulärwert, d.h.

$$\kappa_2(A) = \frac{\sigma_1}{\sigma_p}.$$

7. Die Quadrate  $\sigma_1^2, \dots, \sigma_p^2$  der Singulärwerte sind die Eigenwerte von  $A^T A$  und  $AA^T$  zu den Eigenvektoren  $V_1, \dots, V_p$  bzw.  $U_1, \dots, U_p$

### Bemerkungen:

Des weiteren liefert die Singulärwertzerlegung selbstverständlich eine numerisch günstige Möglichkeit, den Rang einer Matrix zu bestimmen. Das ist z.B. mit dem Gauß-Algorithmus bei schlecht konditionierten Matrizen nicht möglich.

Weitere Eigenschaften (ohne Beweise):

Man kann zeigen, dass der kleinste Singulärwert einer regulären, quadratischen Matrix den Abstand zur nächsten singulären Matrix angibt.

Hat man zwei Matrizen  $A, B \in M(m \times n, \mathbb{R})$ , so kann man mithilfe der SVD überprüfen ob  $A$  durch Rotation von  $B$  entstanden ist.

## 3.2 Berechnung der Singulärwerte

Um die Singulärwerte numerisch zu berechnen, könnte man nun nach Folgerung 3.4 die Singulärwerte einer Matrix  $A$  ausrechnen, indem man die Wurzeln der Eigenwerte von  $A^T A$  ausrechnet.

$$\sigma_i(A) = \sqrt{\lambda_i(A^T A)} \quad (2)$$

Das Eigenwertproblem von  $A^T A$  ist schließlich gut konditioniert, sofern man eine konkrete Berechnung von  $A^T A$  vermeidet. Folgendes Beispiel zeigt jedoch, dass eine Berechnung nach (2) ungeeignet ist.

### Beispiel 3.5:

Sei die Rechnung bis auf 4 Ziffern genau. (Rundung) Und sei

$$A = A^T = \begin{pmatrix} 1.005 & 0.995 \\ 0.995 & 1.005 \end{pmatrix}, \sigma_1 = \lambda_1(A) = 2, \sigma_2 = \lambda_2(A) = 0.01.$$

Über  $A^T A$  erhalten wir

$$fl(A^T A) = \begin{pmatrix} 2.000 & 2.000 \\ 2.000 & 2.000 \end{pmatrix}, \sigma_1^2 = 4, \sigma_2^2 = 0.$$

(fl berücksichtigt die Rechengenauigkeit, will heißen es wird nicht  $A^T A$  exakt ausgerechnet)

Also versuchen wir, eine Berechnung der Singulärwerte ohne  $A^T A$  zu erreichen. Man untersucht deswegen erst, unter welchen Operationen die Singulärwerte invariant sind.

**Lemma 3.6:**

Sei  $A \in M(m \times n, \mathbb{R})$ , und seinen  $P \in \mathbf{O}(m), Q \in \mathbf{O}(n)$  orthogonale Matrizen. Dann haben  $A$  und  $B := PAQ$  die gleichen Singulärwerte.

**Beweis:**

Wir wissen, dass sich unter den orthogonalen Matrizen die Eigenwerte von  $A^T A$  nicht ändern und somit bleiben auch die Singulärwerte von  $A$  erhalten.

□

Man darf also eine Matrix  $A$  von links und rechts mit orthogonalen Matrizen multiplizieren, ohne dass sich die Singulärwerte ändern. Des weiteren möchte man für den QR-Algorithmus eine Tridiagonalgestalt für  $A^T A$  herstellen. Am einfachsten geht dies, wenn  $A$  auf Bidiagonalgestalt ist. Mit folgendem Lemma kann man dies mithilfe von Householder-Transformationen erreichen:

**Lemma 3.7:**

Für jede Matrix  $A \in M(m \times n, \mathbb{R})$  mit  $m \geq n$  (o.B.d.A) existieren orthogonale Matrizen  $P \in \mathbf{O}(m)$  und  $Q \in \mathbf{O}(n)$ , so dass

$$PAQ = \left[ \begin{array}{cccc} * & * & & \\ & \ddots & \ddots & \\ & & \ddots & * \\ & & & * \\ \hline 0 & \dots & \dots & 0 \\ \vdots & & & \vdots \\ 0 & \dots & \dots & 0 \end{array} \right] = \left[ \begin{array}{c} B \\ 0 \end{array} \right]$$

wobei  $B$  eine (quadratische) Bidiagonalmatrix ist.



eliminiere $z_2$	(Givens von links)	→ erzeugt	$z_3$
eliminiere $z_3$	(Givens von rechts)	→ erzeugt	$z_4$
$\vdots$			
eliminiere $z_{2n-3}$	(Givens von rechts)	→ erzeugt	$z_{2n-2}$
eliminiere $z_{2n-2}$	(Givens von links)		

Man beseitigt also mithilfe von Givens-Rotationen von rechts und links (abwechselnd) die fill-in-Elemente, welche entlang der Diagonalen immer weiter nach rechts unten wandern. Im Englischen nennt man diesen Prozess auch *chasing*  $\sim$  *Verfolgungsjagd*, da man die fill-in-Elemente anschaulich die Diagonale herunterjagt.

Nach diesem Verfahren hat die Matrix wieder Bidiagonalgestalt und wir haben einen Iterationsschritt des QR-Verfahrens für  $B^T B$  nur auf  $B$  ausgeführt.

Nach QR-Zerlegung gilt Konvergenz:

$$B_k^T B_k \rightarrow \Lambda = \text{diag}(\sigma_1^2, \dots, \sigma_n^2) = \Sigma^2 \text{ für } k \rightarrow \infty.$$

Daher konvergiert die Folge  $B_k$  gegen die Diagonalmatrix der Singulärwerte von  $B$ , d.h.

$$B_k \rightarrow \Sigma \text{ für } k \rightarrow \infty.$$

Zusammengefasst erhalten wir folgenden Algorithmus zur Bestimmung der Singulärwerte von  $A$ .

### 3.3 Algorithmus für Singulärwerte

QR-Algorithmus für Singulärwerte.

- a) Bringe  $A \in M(m \times n, \mathbb{R})$  mit Hilfe orthogonaler Transformation,  $P \in \mathbf{O}(m)$  und  $Q \in \mathbf{O}(n)$  (z.B. Householder-Reflexionen), auf Bidiagonalgestalt.

$$PAQ = \begin{pmatrix} B \\ 0 \end{pmatrix}, B \in M(n \times n, \mathbb{R}) \text{ obere Bidiagonalmatrix}$$

- b) Führe den QR-Algorithmus für  $B^T B$  nach dem *chasing*-Verfahren auf  $B$  aus und erhalte so eine Folge von Bidiagonalmatrizen  $\{B_k\}$ , die gegen die Diagonalmatrix  $\Sigma$  der Singulärwerte konvergiert.

Für den Aufwand zählen wir dabei für  $m = n$

- a)  $\sim \frac{4}{3}n^3$  Multiplikationen für die Reduktion auf Bidiagonalgestalt,
- b)  $O(n^2)$  Multiplikationen für den modifizierten QR-Algorithmus.

Zum Vergleich mit Gauß-Eliminationsverfahren:  $\mathbf{O}(\frac{2}{3}n^3) + \mathbf{O}(n^2)$

Das SVD benötigt also ungefähr doppelt so viele Operationen wie das gaußsche Eliminationsverfahren bei einer entsprechend großen Matrix.

## 4 Zusammenfassung

In diesem Vortrag werden weitere Herangehensweisen an das lineare Ausgleichsproblem behandelt.

Durch Umformulieren des linearen Gleichungssystems  $Ax = b$  auf  $x = A^+b$  erhält man die Darstellung der Pseudoinversen. Mithilfe der QR-Zerlegung erhält man dann einen konkreten Algorithmus, mit dem man die beste Lösung unter der Bedingung

$$\|\mathbf{Ax} - \mathbf{b}\|_2 \stackrel{!}{=} \min$$

ausrechnen kann.

Eine allumfassendere Herangehensweise ist die Singulärwertzerlegung, mit deren Hilfe man nicht nur die bestmögliche Lösung des Gleichungssystems erhält, sondern auch viele andere Möglichkeiten zur Analyse von sehr allgemeinen Matrizen bekommt.

Der springende Punkt bei den Verfahren der Pseudoinversen und Singulärwertzerlegung ist, dass man vor einem sehr allgemeinen Gleichungssystem steht

$$\mathbf{Ax} = b \quad \mathbf{A} \in M(m \times n, \mathbb{R}), x \in \mathbb{R}^n, b \in \mathbb{R}^m$$

und hierbei weder Anforderungen an den Rang stellt, noch an  $m$  und  $n$ . Das stellt einen wesentlichen Unterschied zu den bislang behandelten Verfahren (z.B. Householder) dar. Und das ist die Stärke dieser beider Verfahren.

## 5 Literaturhinweise:

**Deuffhard-Hohmann: Numerische Mathematik I, 2.Auflage;**

Kapitel 3.3 (Verallgemeinerte Inverse) und

Kapitel 5.4 (Singulärwertzerlegung)

**Stoer-Bulirsch: Numerische Mathematik II, 3.Auflage;**

Kapitel 6.4 (weitere Eigenschaften der Singulärwertzerlegung)