



LUDWIG-  
MAXIMILIANS-  
UNIVERSITÄT  
MÜNCHEN

INSTITUT FÜR STATISTIK



Leo Knüsel

# Singulärwert-Zerlegung und Methode der kleinsten Quadrate

Technical Report Number 031, 2008  
Department of Statistics  
University of Munich

<http://www.stat.uni-muenchen.de>



## Vorwort

a) *Zur 2. Version (September 2008)*

Es wurden einige Verbesserungen und Ergänzungen vorgenommen (schwache Multikollinearität, Rang-k-Approximation mit Frobenius-Norm).

b) *Zur 3. Version (Oktober 2008)*

Die Rang-k-Approximation wurde ergänzt um einen nützlichen Satz, der eine Beziehung herstellt zwischen dem betragsmäßig größten Matrixelement und der Spektralnorm einer beliebigen reellen Matrix  $\mathbf{A} = (m \times n) = (a_{ij})$ :

$$\max_{i,j} |a_{ij}| \leq \|\mathbf{A}\| = \sigma_{\max}.$$

Die Grundlagen für diesen Satz finden sich im Buch *Matrix Algebra and Its Applications to Statistics and Econometrics* von Rao-Rao (1998).

c) *Zur 4. Version (Juli 2009)*

Die Herleitung der Singulärwert-Zerlegung nach Stewart (1973) wurde im Abschnitt 7 hinzugefügt.

# Singulärwert-Zerlegung und Methode der kleinsten Quadrate

Leo Knüsel

Institut für Statistik  
Ludwig-Maximilians-Universität München  
80539 München  
knuesel@stat.uni-muenchen.de

**Abstract:** The method of least squares is an important instrument to determine the optimal linear estimators in regression models. By means of the singular value decomposition we can find the least squares estimators without differentiation, without solving the normal equations and without assumptions on the rank of the data matrix. Even in case of multicollinearity we can find the simple and natural solutions. The results in the paper are not new, they have been developed mainly in numerical publications, but they are hardly to be found in statistical textbooks.

**Keywords:** least squares, multicollinearity, singular value decomposition, generalized inverse, Moore-Penrose-inverse, pseudo-inverse, rank-k-approximation, ridge estimation, shrinkage estimation.

## 1. Einführung und Überblick

Die Methode der kleinsten Quadrate ist in der Regressionsrechnung ein wichtiges Instrument zur Bestimmung von optimalen linearen Schätzfunktionen. Die Singulärwert-Zerlegung einer beliebigen Daten-Matrix  $\mathbf{X} = (n \times p)$  ( $n =$  Anzahl der Beobachtungen,  $p =$  Anzahl der Variablen) erlaubt es, die Kleinste-Quadrat-Schätzungen zu finden ohne Differentiation, ohne das Lösung von Normalgleichungen und ohne Voraussetzungen über den Rang der Matrix  $X$ . Auch im Falle von Multikollinearität liefert diese Methode die einfachen natürlichen Lösungen. Die Ergebnisse, die hier vorgestellt werden, sind vor allem in der numerischen Fachliteratur entwickelt worden (vgl. Golub, 1965; Golub-Van Loan, 1996; Stör, 2005), sie haben aber in unseren Statistik-Lehrbüchern noch kaum ihren Niederschlag gefunden.

### *Überblick*

- Abschnitt 2: Lineares Modell und Singulärwert-Zerlegung in der deskriptiven Statistik
- Abschnitt 3: Zur Multikollinearität
- Abschnitt 4: Vergleich mit klassischer Lösung im regulären Fall
- Abschnitt 5: Elimination unnötiger Variablen
- Abschnitt 6: Stochastisches Modell
- Abschnitt 7: Singulärwert-Zerlegung und Eigenwert-Zerlegung
- Abschnitt 8: Singulärwert-Zerlegung und verallgemeinerte Inverse
- Abschnitt 9: Rang-k-Approximation mit Singulärwert-Zerlegung
- Abschnitt 10: Singuläre Werte von orthogonalen und symmetrischen Matrizen

## 2. Lineares Modell und Singulärwert-Zerlegung in der deskriptiven Statistik

Wir betrachten das lineare Modell (bzw. den linearen Ansatz) der beschreibenden Statistik in Matrix-Notation

$$(1) \quad \mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e},$$

wobei

$$\mathbf{X} = (x_{ij}) = (n \times p) = \text{Datenmatrix (Design-Matrix),}$$

( $n = \text{Anzahl der Beobachtungen, } p = \text{Anzahl der Variablen, } n > p$ ),

$$\mathbf{y} = (n \times 1) = \text{Beobachtungen der abhängigen Variablen,}$$

$$\boldsymbol{\beta} = (p \times 1) = \text{unbekannter Parametervektor.}$$

Es sei

$$(2) \quad \mathbf{e} = \mathbf{y} - \mathbf{X}\boldsymbol{\beta}, \text{ d.h. } e_i = y_i - \sum_{j=1}^p x_{ij}\beta_j.$$

Bestimmung von  $\boldsymbol{\beta}$  mit KQ-Methode (Methode der kleinsten Quadrate):

$$(3) \quad \boldsymbol{\beta} \text{ so wählen, dass } \sum_{i=1}^n e_i^2 = \mathbf{e}^T \mathbf{e} = \|\mathbf{X}\boldsymbol{\beta} - \mathbf{y}\|^2 = \min.$$

*Transformation des Modells mit Singulärwert-Zerlegung (SVD = Singular value decomposition)*

Die Matrix  $\mathbf{X}$  kann geschrieben werden als (vgl. Abschnitt 7)

$$(4) \quad \mathbf{X} = \mathbf{U}\mathbf{D}\mathbf{V}^T \text{ (Singulärwert-Zerlegung),}$$

wobei

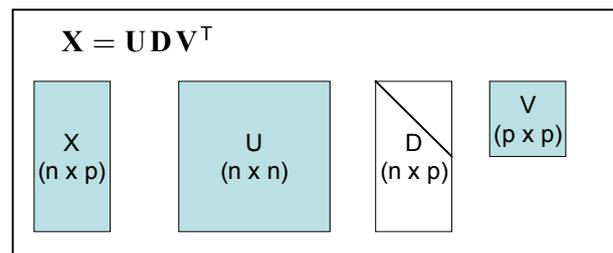
$$\mathbf{U} = (n \times n) \text{ orthogonal,}$$

$$\mathbf{V} = (p \times p) \text{ orthogonal,}$$

$$\mathbf{D} = (n \times p) = \text{diag}(d_1, \dots, d_p)$$

$$d_1 \geq d_2 \geq \dots \geq d_r > 0, \quad d_{r+1} = \dots = d_p = 0$$

$$d_1, \dots, d_p \text{ heißen } \textit{singuläre Werte} \text{ von } \mathbf{X}, \quad r = \text{rk}(\mathbf{X}) = \text{Rang von } \mathbf{X}.$$



Somit kann das lineare Modell (1) geschrieben werden als

$$(5) \quad \mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e} = \mathbf{U}\mathbf{D}\mathbf{V}^T\boldsymbol{\beta} + \mathbf{e}.$$

Wenn wir (5) von links mit  $\mathbf{U}^T$  multiplizieren, erhalten wir

$$(6) \quad \underbrace{\mathbf{U}^T \mathbf{y}}_{=\tilde{\mathbf{y}}} = \underbrace{\mathbf{U}^T \mathbf{U}}_{=\mathbf{I}_n} \underbrace{\mathbf{D}\mathbf{V}^T \boldsymbol{\beta}}_{=\tilde{\boldsymbol{\beta}}} + \underbrace{\mathbf{U}^T \mathbf{e}}_{=\tilde{\mathbf{e}}}$$

d.h.

$$(7) \quad \tilde{\mathbf{y}} = \mathbf{D}\tilde{\boldsymbol{\beta}} + \tilde{\mathbf{e}}, \text{ wobei } \tilde{\mathbf{y}} = \mathbf{U}^T \mathbf{y}, \quad \tilde{\boldsymbol{\beta}} = \mathbf{V}^T \boldsymbol{\beta}, \quad \tilde{\mathbf{e}} = \mathbf{U}^T \mathbf{e}.$$

Die Darstellung (7) nennen wir die *kanonische Form* des linearen Modells (1). Es gilt jetzt

$$(8) \quad \begin{aligned} \tilde{y}_1 &= d_1 \tilde{\beta}_1 + \tilde{e}_1 \\ \dots & \\ \tilde{y}_r &= d_r \tilde{\beta}_r + \tilde{e}_r \\ \tilde{y}_{r+1} &= \tilde{e}_{r+1} && \text{(unabhängig von } \tilde{\beta}_{r+1}) \\ \dots & \\ \tilde{y}_p &= \tilde{e}_p && \text{(unabhängig von } \tilde{\beta}_p) \\ \tilde{y}_{p+1} &= \tilde{e}_{p+1} \\ \dots & \\ \tilde{y}_n &= \tilde{e}_n. \end{aligned}$$

Es sind also nur die Parameter  $\tilde{\beta}_1, \dots, \tilde{\beta}_r$  durch das Modell (7) festgelegt, die übrigen  $p-r$  Parameter können beliebig gewählt werden (sind nicht identifizierbar). Nun wollen wir die KQ-Schätzung für den Parametervektor  $\tilde{\boldsymbol{\beta}}$  bestimmen. Es gilt

$$\tilde{\mathbf{e}}^\top \tilde{\mathbf{e}} = \mathbf{e}^\top \underbrace{\mathbf{U}\mathbf{U}^\top}_{=\mathbf{I}_n} \mathbf{e} = \mathbf{e}^\top \mathbf{e}, \quad \text{d.h.} \quad \sum_{i=1}^n \tilde{e}_i^2 = \sum_{i=1}^n e_i^2.$$

Das KQ-Problem lautet jetzt

$$\tilde{\boldsymbol{\beta}} \quad (= \mathbf{V}^\top \boldsymbol{\beta}) \quad \text{in (7) so wählen, dass} \quad \sum_{i=1}^n \tilde{e}_i^2 = \tilde{\mathbf{e}}^\top \tilde{\mathbf{e}} = \min.$$

Gemäß (7) und (8) gilt  $\tilde{\mathbf{e}} = \tilde{\mathbf{y}} - \mathbf{D}\tilde{\boldsymbol{\beta}}$ , d.h.

$$\tilde{e}_i = \begin{cases} \tilde{y}_i - d_i \tilde{\beta}_i & \text{für } i = 1, \dots, r \\ \tilde{y}_i & \text{für } i = r+1, \dots, n. \end{cases}$$

Nach der KQ-Methode sollen wir

$$\tilde{\beta}_1, \dots, \tilde{\beta}_p \quad \text{so wählen, dass} \quad \sum_{i=1}^n \tilde{e}_i^2 = \min.$$

Die Quadratsumme  $\sum \tilde{e}_i^2$  ( $= \sum e_i^2$ ) wird offensichtlich minimal falls

$$\begin{aligned} \tilde{\beta}_1, \dots, \tilde{\beta}_r & \text{ so gewählt werden, dass } \tilde{e}_1 = \dots = \tilde{e}_r = 0 \\ \tilde{\beta}_{r+1}, \dots, \tilde{\beta}_p & \text{ können beliebig gewählt werden (ohne Einfluss auf Quadratsumme).} \end{aligned}$$

Somit ergibt sich als KQ-Lösung

$$(9) \quad \tilde{\beta}_i = \hat{\beta}_i = \begin{cases} \tilde{y}_i / d_i & \text{für } i = 1, \dots, r \\ \text{beliebig} & \text{für } i = r+1, \dots, p. \end{cases}$$

*Minimale Quadratsumme:*

$$Q_{\min} = \sum_{i=1}^n e_i^2 = \sum_{i=1}^n \tilde{e}_i^2 = \sum_{i=r+1}^n \tilde{e}_i^2 = \sum_{i=r+1}^n \tilde{y}_i^2.$$

*Zurückspielen auf Originalgrößen:*

Da  $\tilde{\boldsymbol{\beta}} = \mathbf{V}^\top \boldsymbol{\beta}$  so gilt  $\boldsymbol{\beta} = \mathbf{V}\tilde{\boldsymbol{\beta}}$ , wobei  $\tilde{\beta}_1, \dots, \tilde{\beta}_r$  eindeutig festgelegt (identifizierbar) und  $\tilde{\beta}_{r+1}, \dots, \tilde{\beta}_p$  beliebig wählbar sind (Minimum  $Q_{\min}$  eindeutig, Minimalstelle nicht eindeutig). Die Gesamtheit aller Lösungsvektoren  $\boldsymbol{\beta}$  ist ein  $(p-r)$ -dimensionaler Raum.

*Beachte:*

- Keine Voraussetzungen, Rang  $r$  von  $\mathbf{X}$  kann kleiner als  $p$  sein (Multikollinearität);
- Keine Differentiation zur Bestimmung der Extremalstelle, kein Lösen Normalgleichungen;
- Das Vorgehen mit Hilfe der SVD ist ähnlich wie beim Nachweis der *Minimumseigenschaft des arithmetischen Mittels* mit Hilfe des Verschiebungssatzes von Steiner:

$x_1, \dots, x_n$  beliebige reelle Zahlen;

$a$  so wählen, dass  $\sum (x_i - a)^2 = \min$ ;

$$\text{Satz von Steiner: } \sum (x_i - a)^2 = \underbrace{\sum (x_i - \bar{x})^2}_{(1)} + \underbrace{n(\bar{x} - a)^2}_{(2)}$$

(1) unabhängig von  $a$ ;

(2) minimal für  $a = \bar{x}$ .

### 3. Zur Multikollinearität

Falls die Datenmatrix  $\mathbf{X} = (n \times p)$  nicht vollen Rang besitzt, d.h. falls  $r = \text{rk}(\mathbf{X}) < p$ , so spricht man von Multikollinearität. Aus der Singulärwert-Zerlegung ergibt sich

$$\mathbf{X} = \mathbf{U}\mathbf{D}\mathbf{V}^T \Leftrightarrow \mathbf{X}\mathbf{V} = \mathbf{U}\mathbf{D}$$

Nun ist

$$\begin{aligned} \mathbf{U} &= (n \times n) = (\mathbf{u}_1, \dots, \mathbf{u}_n), & \mathbf{U}\mathbf{D} &= (n \times p) = (d_1\mathbf{u}_1, \dots, d_r\mathbf{u}_r, \mathbf{0}, \dots, \mathbf{0}) \\ \mathbf{V} &= (p \times p) = (\mathbf{v}_1, \dots, \mathbf{v}_p), & \mathbf{X}\mathbf{V} &= (n \times p) = (\mathbf{X}\mathbf{v}_1, \dots, \mathbf{X}\mathbf{v}_p) \end{aligned}$$

und da  $\mathbf{X}\mathbf{V} = \mathbf{U}\mathbf{D}$ , so gilt

$$\mathbf{X}\mathbf{v}_{r+1} = \dots = \mathbf{X}\mathbf{v}_p = \mathbf{0}.$$

$\mathbf{X}\mathbf{v}_j$  ist eine Linearkombination der Spalten von  $\mathbf{X}$  mit dem Koeffizientenvektor  $\mathbf{v}_j$ . Die Information  $\mathbf{X}\mathbf{v}_j = \mathbf{0}$  besagt, dass ein (unerwarteter) Zusammenhang zwischen den erklärenden Variablen besteht; eine Spalte von  $\mathbf{X}$  und der zugehörige Parameter sind eliminierbar!

*Beispiel:*

Es sei  $\mathbf{X} = (n \times p) = (\mathbf{x}_1, \dots, \mathbf{x}_p)$ ; dann ist  $\mathbf{X}\boldsymbol{\beta} = \beta_1\mathbf{x}_1 + \dots + \beta_p\mathbf{x}_p$ . Wir wollen annehmen, dass die erste und zweite Spalte von  $\mathbf{X}$  identisch sind, d.h. dass  $\mathbf{x}_1 = \mathbf{x}_2$ . Dann gilt

$$\beta_1\mathbf{x}_1 + \beta_2\mathbf{x}_2 = (\beta_1 + \beta_2)\mathbf{x}_1$$

$$\tilde{\beta}_1 = \beta_1 + \beta_2 \quad \text{ist eindeutig festgelegt, identifizierbar,}$$

$$\tilde{\beta}_2 = \beta_1 - \beta_2 \quad \text{ist nicht identifizierbar, kann bei gegebenem } \tilde{\beta}_1 \text{ beliebige Werte annehmen, ohne dass sich das Modell ändert.}$$

Die zweite Spalte  $\mathbf{x}_2$  kann also eliminiert werden und die erste Spalte erhält dann den neuen Koeffizienten  $\tilde{\beta}_1 = \beta_1 + \beta_2$ .

*KQ-Lösung mit minimaler Länge*

Wir haben oben gesehen dass die Modellparameter  $\tilde{\beta}_{r+1}, \dots, \tilde{\beta}_p$  und die entsprechenden KQ-Schätzungen beliebig gewählt werden können. Wir wollen mit einer einfachen zusätzlichen Annahme alle Modellparameter und die zugehörigen KQ-Schätzungen eindeutig festlegen. Wir verlangen, dass die freien Parameter im kanonischen Modell (7) alle gleich null gesetzt werden, dass also  $\tilde{\beta}_{r+1} = \dots = \tilde{\beta}_p = 0$ . Entsprechend setzen wir dann auch bei der KQ-Schätzung (9) die freien Schätzungen alle null:

$$(10) \quad \hat{\tilde{\beta}}_i = \begin{cases} \tilde{y}_i/d_i & \text{für } i = 1, \dots, r \\ 0 & \text{für } i = r+1, \dots, p. \end{cases}$$

Wie lautet die zusätzliche Annahme  $\tilde{\beta}_{r+1} = \dots = \tilde{\beta}_p = 0$  bei den originalen Parametern? Die Forderung  $\tilde{\beta}_{r+1} = \dots = \tilde{\beta}_p = 0$  ist äquivalent zu

$$\sum_{i=1}^p \tilde{\beta}_i^2 = \tilde{\boldsymbol{\beta}}^T \tilde{\boldsymbol{\beta}} = \min,$$

und da  $\tilde{\boldsymbol{\beta}} = \mathbf{V}^T \boldsymbol{\beta}$  so gilt  $\tilde{\boldsymbol{\beta}}^T \tilde{\boldsymbol{\beta}} = \boldsymbol{\beta}^T \mathbf{V}\mathbf{V}^T \boldsymbol{\beta} = \boldsymbol{\beta}^T \boldsymbol{\beta} = \|\boldsymbol{\beta}\|^2$ , da ja  $\mathbf{V}$  orthogonal ist und somit  $\mathbf{V}\mathbf{V}^T = \mathbf{I}_p$ . Im Falle  $r < p$  sollen also die Originalparameter  $\boldsymbol{\beta}$  so festgelegt werden, dass die Norm (=Länge)  $\|\boldsymbol{\beta}\|$  minimal wird; das Entsprechende gilt für den zugehörigen Schätzvektor  $\hat{\boldsymbol{\beta}} = \mathbf{V}\hat{\tilde{\boldsymbol{\beta}}}$ . Damit werden alle

Parameter und die zugehörigen Schätzungen eindeutig (identifizierbar). Man spricht dann von den *KQ-Schätzungen mit minimaler Länge*.

*KQ-Lösung mit minimaler Länge in Matrix-Notation*

Im kanonischen Modell (7) ist

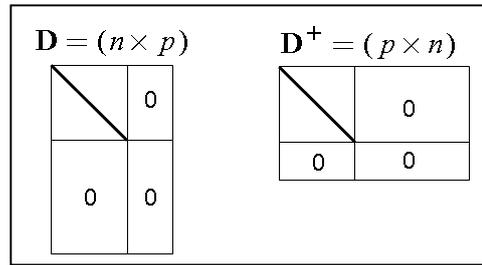
$$\mathbf{D} = (n \times p) = \text{diag}(d_1, \dots, d_r, 0, \dots, 0)$$

und wir setzen

$$\mathbf{D}^+ = (p \times n) = \text{diag}\left(\frac{1}{d_1}, \dots, \frac{1}{d_r}, 0, \dots, 0\right).$$

Es gilt

$$\mathbf{D}^+ \mathbf{D} = (p \times p) = \begin{pmatrix} \mathbf{I}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}, \quad \mathbf{D} \mathbf{D}^+ = (n \times n) = \begin{pmatrix} \mathbf{I}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}.$$



Die Matrix  $\mathbf{D}^+$  heißt *verallgemeinerte Inverse* zu  $\mathbf{D}$  (vgl. Abschnitt 7). Die KQ-Lösung (10) mit minimaler Länge kann jetzt geschrieben werden als  $\hat{\boldsymbol{\beta}} = \mathbf{D}^+ \tilde{\mathbf{y}}$ , und da  $\tilde{\boldsymbol{\beta}} = \mathbf{V}^T \boldsymbol{\beta}$  und  $\tilde{\mathbf{y}} = \mathbf{U}^T \mathbf{y}$  so erhalten wir

$$(11) \quad \hat{\boldsymbol{\beta}} = \mathbf{V} \hat{\tilde{\boldsymbol{\beta}}} = \mathbf{V} \mathbf{D}^+ \tilde{\mathbf{y}} = \mathbf{V} \mathbf{D}^+ \mathbf{U}^T \mathbf{y} = \mathbf{X}^+ \mathbf{y}.$$

Dabei ist  $\mathbf{X}^+ = \mathbf{V} \mathbf{D}^+ \mathbf{U}^T$  die verallgemeinerte Inverse zu  $\mathbf{X} = \mathbf{U} \mathbf{D} \mathbf{V}^T$  (vgl. Abschnitt 7).

*Bemerkung*

Bei *schwacher Multikollinearität* besitzt die Datenmatrix  $\mathbf{X} = (n \times p)$  vollen Rang, aber sie ist nahezu singular. Dann sind alle singulären Werte  $d_1, \dots, d_p$  positiv, d.h.  $d_1 \geq \dots \geq d_p > 0$ , aber  $d_p/d_1$  ist sehr klein ( $< 10^{-6}$ , etwa). Da schwache Multikollinearität in vielen Fällen nur durch Rundungsfehler bei der Datenmatrix  $\mathbf{X}$  verursacht wird, so bietet die Rang-k-Approximation eine nahe liegende Lösung. Dabei wird die Matrix  $\mathbf{X}$  durch eine Matrix  $\tilde{\mathbf{X}}$  ersetzt, welche sich von  $\mathbf{X}$  nur im Bereich der Rundungsfehler unterscheidet, und bei welcher die  $p-k$  kleinsten singulären Werte von  $\mathbf{X}$  null gesetzt werden, sodass  $\tilde{\mathbf{X}}$  den Rang  $k < p$  besitzt (vgl. Abschnitt 9). Man kann dann unnötige Variablen eliminieren oder die Lösungen mit minimaler Länge bestimmen.

#### 4. Vergleich mit klassischer Lösung im regulären Fall

Wir betrachten wieder das lineare Modell  $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e}$  und setzen jetzt voraus, dass die Datenmatrix  $\mathbf{X} = (n \times p)$  vollen Rang hat, dass also  $r = \text{rk}(\mathbf{X}) = p$ . Die klassischen KQ-Schätzer für  $\boldsymbol{\beta}$  sind gegeben durch

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y}.$$

Mit Hilfe der SVD haben wir in (11) als Lösung gefunden  $\hat{\boldsymbol{\beta}} = \mathbf{X}^+ \mathbf{y}$ , wobei

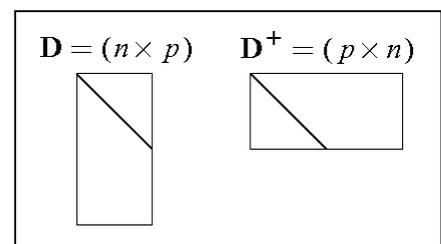
$$\mathbf{X} = (n \times p) = \mathbf{U} \mathbf{D} \mathbf{V}^T \quad \text{Singulärwert-Zerlegung von } \mathbf{X}$$

$$\mathbf{X}^+ = (p \times n) = \mathbf{V} \mathbf{D}^+ \mathbf{U}^T \quad \text{verallgemeinerte Inverse zu } \mathbf{X}$$

$$\mathbf{D} = (n \times p) = \text{diag}(d_1, \dots, d_p) \quad \text{mit } d_1 \geq d_2 \geq \dots \geq d_p > 0$$

$$\mathbf{D}^+ = (p \times n) = \text{diag}(1/d_1, \dots, 1/d_p).$$

Die beiden Lösungen sind identisch, denn



$$\mathbf{X}^T \mathbf{X} = \underbrace{\mathbf{V} \mathbf{D}^T \mathbf{U}^T}_{=\mathbf{X}^T} \underbrace{\mathbf{U} \mathbf{D} \mathbf{V}^T}_{=\mathbf{X}} = \mathbf{V} \mathbf{D}^T \underbrace{\mathbf{U}^T \mathbf{U}}_{=\mathbf{I}_n} \mathbf{D} \mathbf{V}^T = \mathbf{V} \mathbf{D}^T \mathbf{D} \mathbf{V}^T$$

$$\mathbf{D}^T \mathbf{D} = (p \times p) = \text{diag}(d_1^2, \dots, d_p^2) = \mathbf{\Lambda}, \lambda_i = d_i^2 > 0$$

$$\mathbf{X}^T \mathbf{X} = \mathbf{V} \mathbf{\Lambda} \mathbf{V}^T \quad (\text{Spektralzerlegung, Eigenwertdarstellung von } \mathbf{X}^T \mathbf{X})$$

$$(\mathbf{X}^T \mathbf{X})^{-1} = (\mathbf{V} \mathbf{\Lambda} \mathbf{V}^T)^{-1} = \mathbf{V} \mathbf{\Lambda}^{-1} \mathbf{V}^T$$

$$(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T = \mathbf{V} \mathbf{\Lambda}^{-1} \underbrace{\mathbf{V}^T \mathbf{V}}_{=\mathbf{I}_p} \mathbf{D}^T \mathbf{U}^T = \mathbf{V} \underbrace{\mathbf{\Lambda}^{-1} \mathbf{D}^T}_{=\mathbf{D}^+} \mathbf{U}^T = \mathbf{V} \mathbf{D}^+ \mathbf{U}^T = \mathbf{X}^+$$

Somit gilt  $\hat{\boldsymbol{\beta}} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y} = \mathbf{X}^+ \mathbf{y}$ , falls  $\text{rk}(\mathbf{X}) = p$ ; die beiden Lösungen sind dann äquivalent.

## 5. Elimination unnötiger Variablen

Wir betrachten wieder das lineare Modell

$$\mathbf{y} = \mathbf{X} \boldsymbol{\beta} + \mathbf{e},$$

und nehmen an, dass die Datenmatrix  $\mathbf{X} = (n \times p)$  den Rang  $r = \text{rk}(\mathbf{X}) < p$  besitzt. Dann sind die Spalten von  $\mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_p)$  linear abhängig, und  $p - r$  Spalten können eliminiert werden (vgl. Abschnitt 3). Wie sieht dann das reduzierte Modell aus?

### 1. Methode: Schrittweise eliminieren

Es sei

$$\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) = \begin{pmatrix} 1 & 1 & 2 \\ 1 & 2 & 3 \\ 1 & 3 & 4 \\ 1 & 4 & 5 \end{pmatrix}.$$

Offensichtlich gilt  $\mathbf{x}_1 + \mathbf{x}_2 = \mathbf{x}_3$ , und damit kann eine der drei Spalten eliminiert werden. Wenn wir die dritte Spalte eliminieren wollen, erhalten wir

$$\begin{aligned} \mathbf{X} \boldsymbol{\beta} &= \beta_1 \mathbf{x}_1 + \beta_2 \mathbf{x}_2 + \beta_3 \mathbf{x}_3 \\ &= \beta_1 \mathbf{x}_1 + \beta_2 \mathbf{x}_2 + \beta_3 (\mathbf{x}_1 + \mathbf{x}_2) \\ &= (\beta_1 + \beta_3) \mathbf{x}_1 + (\beta_2 + \beta_3) \mathbf{x}_2. \end{aligned}$$

Wir setzen nun

$$\mathbf{X}_1 = (\mathbf{x}_1, \mathbf{x}_2) = (n \times 2), \quad \boldsymbol{\beta}_1 = \begin{pmatrix} \beta_1 + \beta_3 \\ \beta_2 + \beta_3 \end{pmatrix},$$

und es gilt  $\mathbf{X}_1 \boldsymbol{\beta}_1 = \mathbf{X} \boldsymbol{\beta}$ , wobei die reduzierte Datenmatrix  $\mathbf{X}_1$  jetzt vollen Rang besitzt. Im allgemeinen Fall wird dieses Verfahren wiederholt, bis die reduzierte Matrix vollen Rang besitzt.

### 2. Methode: Zusammenhang mit Hauptkomponenten

Es sei  $\mathbf{X} = \mathbf{U} \mathbf{D} \mathbf{V}^T$  die Singulärwert-Zerlegung von  $\mathbf{X}$ , wobei  $\mathbf{D} = (n \times p) = \text{diag}(d_1, \dots, d_p)$  mit  $d_1 \geq \dots \geq d_r > 0$ ,  $d_{r+1} = \dots = d_p = 0$ . Nun gilt  $\mathbf{X} \mathbf{V} = \mathbf{U} \mathbf{D} = (d_1 \mathbf{u}_1, \dots, d_r \mathbf{u}_r, 0, \dots, 0)$ , und somit ist  $\mathbf{X} \mathbf{v}_j = 0$  für  $j = r + 1, \dots, p$ . Wir schreiben

$$\mathbf{V} = (\mathbf{V}_1 | \mathbf{V}_2) \quad \text{mit } \mathbf{V}_1 = (p \times r)$$

$$(12) \quad \mathbf{X}_1 = \mathbf{X} \mathbf{V}_1 = (n \times r)$$

$$\boldsymbol{\beta}_1 = \mathbf{V}_1^T \boldsymbol{\beta} = (r \times 1).$$

Da  $\mathbf{X} \mathbf{V}_2 = \mathbf{0}$ , so gilt jetzt

$$\mathbf{X}\boldsymbol{\beta} = \mathbf{X}\mathbf{V}\mathbf{V}^T\boldsymbol{\beta} = (\mathbf{X}\mathbf{V}_1 \mid \mathbf{0}) \begin{pmatrix} \mathbf{V}_1^T\boldsymbol{\beta} \\ \mathbf{V}_2^T\boldsymbol{\beta} \end{pmatrix} = \mathbf{X}\mathbf{V}_1\mathbf{V}_1^T\boldsymbol{\beta} = \mathbf{X}_1\boldsymbol{\beta}_1.$$

Damit haben wir die reduzierte Darstellung gefunden:

$$\mathbf{X}\boldsymbol{\beta} = \mathbf{X}_1\boldsymbol{\beta}_1, \text{ wobei } \mathbf{X}_1 = \mathbf{X}\mathbf{V}_1 = (n \times r) \text{ und } \boldsymbol{\beta}_1 = \mathbf{V}_1^T\boldsymbol{\beta} = (r \times 1).$$

Die Spalten von  $\mathbf{X}_1$  sind unkorreliert, denn es gilt ja

$$\mathbf{X}_1^T\mathbf{X}_1 = \mathbf{V}_1^T\mathbf{X}^T\mathbf{X}\mathbf{V}_1 = \mathbf{V}_1^T\mathbf{V}\boldsymbol{\Lambda}\mathbf{V}^T\mathbf{V}_1 = \text{diag}(\lambda_1, \dots, \lambda_r) \text{ mit } \lambda_i = d_i^2.$$

*Bemerkung:*

Es besteht hier ein enger Zusammenhang mit der Hauptkomponentenmethode (principal components) bei den multivariaten Verfahren. Die Hauptkomponenten zu einer Datenmatrix  $\mathbf{X} = (n \times p)$  sind die Spalten der Matrix  $\mathbf{Y} = \mathbf{X}\mathbf{V} = (n \times p)$ , wobei  $\mathbf{V} = (p \times p)$  gegeben ist durch die Eigenwertzerlegung der symmetrischen Matrix  $\mathbf{X}^T\mathbf{X}$ :

$$\mathbf{X}^T\mathbf{X} = \mathbf{V}\boldsymbol{\Lambda}\mathbf{V}^T, \mathbf{V} \text{ orthogonal, } \boldsymbol{\Lambda} = \text{diag}(\lambda_1, \dots, \lambda_p), \lambda_1 \geq \dots \geq \lambda_p \geq 0.$$

Die Spalten der Matrix  $\mathbf{X}_1 = \mathbf{X}\mathbf{V}_1 = (n \times r)$  bei (12) sind also die die Hauptkomponenten zu den positiven Eigenwerten  $\lambda_1, \dots, \lambda_r$ .

*Beispiel (Berechnungen mit Maple 10):*

Es sei

$$\mathbf{X} = \begin{pmatrix} 1 & 1 & 2 \\ 1 & 2 & 3 \\ 1 & 3 & 4 \\ 1 & 4 & 5 \end{pmatrix}.$$

Für die Singulärwert-Zerlegung  $\mathbf{X} = \mathbf{U}\mathbf{D}\mathbf{V}^T$  gilt  $\mathbf{D} = \text{diag}(d_1, d_2, d_3)$  mit

$$d_1^2 = 44 + 2\sqrt{469} \approx 9.344$$

$$d_2^2 = 44 - 2\sqrt{469} \approx 0.829$$

$$d_3 = 0.$$

Für die orthogonale Matrix  $\mathbf{V}$  gilt

$$\mathbf{V} = (\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3) \approx \begin{pmatrix} 0.202 & 0.791 & 0.577 \\ 0.584 & -0.571 & 0.577 \\ 0.786 & 0.220 & -0.577 \end{pmatrix}.$$

Da  $d_3 = 0$ , so hat  $\mathbf{X}$  den Rang 2, und es besteht eine lineare Beziehung zwischen den drei Spalten. Es muss  $\mathbf{X}\mathbf{v}_3 = \mathbf{0}$  sein, und tatsächlich ist

$$\mathbf{X}\mathbf{v}_3 = \underbrace{\frac{\sqrt{3}}{3}}_{=0.577\dots} (\mathbf{x}_1 + \mathbf{x}_2 - \mathbf{x}_3) = \mathbf{0}.$$

$\mathbf{V}_1 = (\mathbf{v}_1, \mathbf{v}_2)$  besteht aus den ersten beiden Spalten von  $\mathbf{V}$ . Nun gilt

$$\mathbf{X}_1 = \mathbf{X}\mathbf{V}_1 \approx \begin{pmatrix} 2.359 & 0.661 \\ 3.729 & 0.311 \\ 5.099 & -0.039 \\ 6.469 & -0.389 \end{pmatrix} \quad \boldsymbol{\beta}_1 = \mathbf{V}_1^T\boldsymbol{\beta} \approx \begin{pmatrix} 0.202\beta_1 + 0.584\beta_2 + 0.786\beta_3 \\ 0.791\beta_1 - 0.571\beta_2 + 0.220\beta_3 \end{pmatrix},$$

und in der Tat ist

$$\mathbf{X}_1\boldsymbol{\beta}_1 = \mathbf{X}\boldsymbol{\beta}.$$

Überdies sind die Spalten von  $\mathbf{X}_1$  unkorreliert.

## 6. Stochastisches Modell

Wir betrachten nun das stochastische lineare Modell

$$(13) \quad \mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon},$$

wobei

$\mathbf{X} = (n \times p)$  gegebene Datenmatrix (Design-Matrix)

$\boldsymbol{\beta} = (p \times 1)$  unbekannter Parametervektor

$\mathbf{y} = (n \times 1)$  Vektor von beobachtbaren Zufallsgrößen

$\boldsymbol{\varepsilon} = (n \times 1)$  Vektor von nicht-beobachtbaren Zufallsgrößen (Fehlervariablen)

$$E(\boldsymbol{\varepsilon}) = \mathbf{0}, \quad E(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^\top) = \sigma^2 \mathbf{I}_n, \quad \text{d.h. } \boldsymbol{\varepsilon} \sim (\mathbf{0}, \sigma^2 \mathbf{I}_n).$$

Falls man sich für das Modell mit der Normalverteilungsannahme interessiert, so gilt  $\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \sigma^2 \mathbf{I}_n)$ .

Gemäß der KQ-Methode haben wir den unbekannt Parametervektor  $\boldsymbol{\beta}$  wie folgt zu bestimmen:

$$\boldsymbol{\beta} = \hat{\boldsymbol{\beta}} \text{ so, dass } \mathbf{e}^\top \mathbf{e} = \sum e_i^2 = \min, \text{ wobei } \mathbf{e} = \mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}.$$

Mit Hilfe der Singulärwert-Zerlegung  $\mathbf{X} = \mathbf{U}\mathbf{D}\mathbf{V}^\top$  können wir das Modell (13) transformieren zu

$$\tilde{\mathbf{y}} = \mathbf{D}\tilde{\boldsymbol{\beta}} + \tilde{\boldsymbol{\varepsilon}}, \quad \text{wobei } \tilde{\mathbf{y}} = \mathbf{U}^\top \mathbf{y}, \quad \tilde{\boldsymbol{\varepsilon}} = \mathbf{U}^\top \boldsymbol{\varepsilon}, \quad \tilde{\boldsymbol{\beta}} = \mathbf{V}^\top \boldsymbol{\beta}$$

und weiter

$$E(\tilde{\boldsymbol{\varepsilon}}) = \mathbf{U}^\top E(\boldsymbol{\varepsilon}) = \mathbf{0}, \quad E(\tilde{\boldsymbol{\varepsilon}}\tilde{\boldsymbol{\varepsilon}}^\top) = \mathbf{U}^\top E(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^\top)\mathbf{U} = \sigma^2 \mathbf{I}_n.$$

Für die transformierten Fehlervariablen  $\tilde{\boldsymbol{\varepsilon}}$  gelten also die gleichen Verteilungsannahmen wie für die ursprünglichen Variablen  $\boldsymbol{\varepsilon}$ , nämlich  $\tilde{\boldsymbol{\varepsilon}} \sim (\mathbf{0}, \sigma^2 \mathbf{I}_n)$  bzw.  $\tilde{\boldsymbol{\varepsilon}} \sim N(\mathbf{0}, \sigma^2 \mathbf{I}_n)$ . Wir erhalten so das so genannte *kanonische Modell* (in neuer Notation, ohne Schlangen)

$$(14) \quad \begin{aligned} \mathbf{y} &= \mathbf{D}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \\ E(\boldsymbol{\varepsilon}) &= \mathbf{0}, \quad E(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^\top) = \sigma^2 \mathbf{I}_n. \end{aligned}$$

$$\mathbf{D} = (n \times p)$$

/	0
0	0

Wir nehmen an, dass  $\text{rk}(\mathbf{X}) = r \leq p$ . Dann gilt

$$\mathbf{D} = (n \times p) = \left( \begin{array}{c|c} \mathbf{D}_1 & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} \end{array} \right)$$

wobei

$$\mathbf{D}_1 = (r \times r) = \text{diag}(d_1, \dots, d_r), \quad d_1 \geq \dots \geq d_r > 0.$$

Wir schreiben

$$\boldsymbol{\beta} = \begin{pmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \end{pmatrix} = \begin{pmatrix} r \times 1 \\ (p-r) \times 1 \end{pmatrix}, \quad \mathbf{y} = \begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{pmatrix} = \begin{pmatrix} r \times 1 \\ (n-r) \times 1 \end{pmatrix}, \quad \boldsymbol{\varepsilon} = \begin{pmatrix} \boldsymbol{\varepsilon}_1 \\ \boldsymbol{\varepsilon}_2 \end{pmatrix} = \begin{pmatrix} r \times 1 \\ (n-r) \times 1 \end{pmatrix}$$

und erhalten

$$\mathbf{y} = \mathbf{D}\boldsymbol{\beta} + \boldsymbol{\varepsilon} = \begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{pmatrix}, \quad \text{wobei } \mathbf{y}_1 = \mathbf{D}_1\boldsymbol{\beta}_1 + \boldsymbol{\varepsilon}_1, \quad \mathbf{y}_2 = \boldsymbol{\varepsilon}_2.$$

Nur die  $r$  Komponenten von  $\boldsymbol{\beta}_1$  treten im Modell auf, die  $p-r$  Komponenten von  $\boldsymbol{\beta}_2$  können beliebig gewählt werden (Pseudoparameter, nicht identifizierbar). Mit der Zusatzforderung  $\beta_{r+1} = \dots = \beta_p = 0$  werden alle Parameter identifizierbar.

Die KQ-Methode mit der Zusatzforderung liefert die Schätzfunktionen

$$(15) \quad \hat{\beta}_i = \begin{cases} y_i/d_i & \text{für } i = 1, \dots, r \\ 0 & \text{für } i = r+1, \dots, p. \end{cases}$$

Jetzt wollen wir die Varianz-Kovarianz-Matrix des Schätzvektors  $\hat{\boldsymbol{\beta}}$  bestimmen. Es gilt gemäß (15)

$$\hat{\boldsymbol{\beta}} = \begin{pmatrix} \hat{\boldsymbol{\beta}}_1 \\ \hat{\boldsymbol{\beta}}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{D}_1^{-1} \mathbf{y}_1 \\ \mathbf{0} \end{pmatrix} = \mathbf{D}^+ \mathbf{y}.$$

Wir haben

$$\mathbf{E}(\mathbf{y}) = \begin{pmatrix} \mathbf{E}(\mathbf{y}_1) \\ \mathbf{E}(\mathbf{y}_2) \end{pmatrix} = \begin{pmatrix} \mathbf{D}_1 \boldsymbol{\beta}_1 \\ \mathbf{0} \end{pmatrix} = \mathbf{D} \boldsymbol{\beta}$$

$$\mathbf{E}(\hat{\boldsymbol{\beta}}) = \mathbf{E}(\mathbf{D}^+ \mathbf{y}) = \mathbf{D}^+ \mathbf{E}(\mathbf{y}) = \mathbf{D}^+ \mathbf{D} \boldsymbol{\beta} = \begin{pmatrix} \boldsymbol{\beta}_1 \\ \mathbf{0} \end{pmatrix} = \boldsymbol{\beta},$$

$$\mathbf{D} = (n \times p)$$

/	0
0	0

$$\mathbf{D}^+ = (p \times n)$$

/	0
0	0

d.h. die Schätzung  $\hat{\boldsymbol{\beta}} = \mathbf{D}^+ \mathbf{y}$  ist erwartungstreu. Weiter gilt

$$\text{var}(\hat{\boldsymbol{\beta}}) = \mathbf{D}^+ \text{var}(\mathbf{y}) (\mathbf{D}^+)^{\top} = \sigma^2 \mathbf{D}^+ (\mathbf{D}^+)^{\top} = \sigma^2 (\mathbf{D}^{\top} \mathbf{D})^+ = \sigma^2 \begin{pmatrix} \mathbf{D}_1^{-2} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} = (p \times p)$$

da ja

$$\text{var}(\hat{\boldsymbol{\beta}}_1) = \mathbf{D}_1^{-1} \underbrace{\text{var}(\mathbf{y}_1)}_{=\sigma^2 \mathbf{I}_r} \mathbf{D}_1^{-1} = \sigma^2 \mathbf{D}_1^{-2}, \quad \text{var}(\hat{\boldsymbol{\beta}}_2) = \mathbf{0}, \quad \text{cov}(\hat{\boldsymbol{\beta}}_1, \hat{\boldsymbol{\beta}}_2) = \mathbf{0}.$$

Da die Schätzung  $\hat{\boldsymbol{\beta}}_1$  die beste unter den linearen erwartungstreuen Schätzungen ist (BLUE = best linear unbiased estimator, Gauss-Markoff-Theorem), und da wir  $\boldsymbol{\beta}_2 = \hat{\boldsymbol{\beta}}_2 = \mathbf{0}$  gesetzt haben, so besitzt auch  $\hat{\boldsymbol{\beta}} = \mathbf{D}^+ \mathbf{y}$  die BLUE-Eigenschaften. Weiter gilt für  $\mathbf{e} = \mathbf{y} - \mathbf{D} \hat{\boldsymbol{\beta}}$

$$\mathbf{e}^{\top} \mathbf{e} = \sum_{i=1}^n e_i^2 = \sum_{i=r+1}^n y_i^2 = \sum_{i=r+1}^n \varepsilon_i^2$$

$$\Rightarrow \mathbf{E}(\mathbf{e}^{\top} \mathbf{e}) = (n-r) \sigma^2$$

$$\Rightarrow \hat{\sigma}^2 = \frac{1}{n-r} \sum_{i=1}^n e_i^2 \text{ ist ein erwartungstreuer Schätzer für } \sigma^2.$$

Durch Rücktransformation erhalten wir die entsprechenden Ergebnisse für das Originalmodell (13).

### Bemerkung

Bei *schwacher Multikollinearität* besitzt die Datenmatrix  $\mathbf{X} = (n \times p)$  vollen Rang, aber sie ist nahezu singular. Dann sind alle singulären Werte  $d_1, \dots, d_p$  positiv, d.h.  $d_1 \geq \dots \geq d_p > 0$ , aber  $d_p/d_1$  ist sehr klein ( $< 10^{-6}$ , etwa). Nun gilt

$$\hat{\beta}_i = y_i/d_i, \quad \mathbf{E}(\hat{\beta}_i) = \beta_i, \quad \text{var}(\hat{\beta}_i) = \sigma^2/d_i^2, \quad i=1, \dots, p.$$

Die Varianz von  $\hat{\beta}_p$  wird also sehr groß im Vergleich zu den Varianzen der übrigen Schätzer. Im Grenzfall  $d_p = 0$  ist ja der Parameter  $\beta_p$  im Modell gar nicht mehr identifizierbar und kann beliebige Werte annehmen. Da schwache Multikollinearität in vielen Fällen nur durch Rundungsfehler bei der Datenmatrix  $\mathbf{X}$  verursacht wird, so bietet die Rang- $k$ -Approximation eine nahe liegende Lösung. Dabei wird die Matrix  $\mathbf{X}$  durch eine Matrix  $\tilde{\mathbf{X}}$  ersetzt, welche sich von  $\mathbf{X}$  nur im Bereich der Rundungsfehler unterscheidet, und bei welcher die  $p-k$  kleinsten singulären Werte von  $\mathbf{X}$  null gesetzt werden, sodass  $\tilde{\mathbf{X}}$  den Rang  $k < p$  besitzt (vgl. Abschnitt 9). Man kann dann unnötige Variablen eliminieren oder die Lösungen mit minimaler Länge bestimmen. In der statistischen Literatur werden bei schwacher Multikollinearität so genannte Ridge-Schätzer oder Shrinkage-Schätzer vorgeschlagen, bei denen die große Varianz der klassischen KQ-Schätzer reduziert werden kann, welche jedoch nicht mehr er

wartungstreu sondern verzerrt (biased) sind (vgl. Rao et al., 2008). Die Rang-k-Approximation scheint mir die sinnvollere Methode zu sein.

### Zusammenfassung

Wir betrachten das lineare Modell:

$$(16) \quad \mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \quad E(\boldsymbol{\varepsilon}) = 0, \quad E(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^\top) = \sigma^2 \mathbf{I}_n.$$

Mit der Singulärwert-Zerlegung  $\mathbf{X} = \mathbf{U}\mathbf{D}\mathbf{V}^\top$  erhalten wir aus (16) das kanonische Modell

$$(17) \quad \tilde{\mathbf{y}} = \mathbf{D}\tilde{\boldsymbol{\beta}} + \tilde{\boldsymbol{\varepsilon}} \quad \text{wobei} \quad \tilde{\mathbf{y}} = \mathbf{U}^\top \mathbf{y}, \quad \tilde{\boldsymbol{\varepsilon}} = \mathbf{U}^\top \boldsymbol{\varepsilon}, \quad \tilde{\boldsymbol{\beta}} = \mathbf{V}^\top \boldsymbol{\beta}, \quad E(\tilde{\boldsymbol{\varepsilon}}) = 0, \quad E(\tilde{\boldsymbol{\varepsilon}}\tilde{\boldsymbol{\varepsilon}}^\top) = \sigma^2 \mathbf{I}_n.$$

*Identifizierbarkeit der Parameter:*

Falls  $\text{rk}(\mathbf{X}) = \text{rk}(\mathbf{D}) = r < p$ , so sind die Parameter  $\tilde{\beta}_{r+1}, \dots, \tilde{\beta}_p$  Pseudoparameter, die im einfachen Modell (17) gar nicht auftreten; wir setzen diese Parameter gleich null. Wir verlangen also, dass die freien Modellparameter so festgelegt werden, dass  $\tilde{\boldsymbol{\beta}}^\top \tilde{\boldsymbol{\beta}} = \min$ , denn daraus ergibt  $\tilde{\beta}_{r+1} = \dots = \tilde{\beta}_p = 0$ . Da  $\tilde{\boldsymbol{\beta}}^\top \tilde{\boldsymbol{\beta}} = \boldsymbol{\beta}^\top \boldsymbol{\beta}$ , so lautet die entsprechende Forderung im Originalmodell: Parameter so festlegen, dass  $\boldsymbol{\beta}^\top \boldsymbol{\beta} = \min$  (Parametervektor mit minimaler Länge). Mit dieser Zusatzforderung werden alle Modellparameter eindeutig festgelegt auf  $\boldsymbol{\beta} = \mathbf{V}\tilde{\boldsymbol{\beta}}$ .

*KQ-Methode im kanonischen Modell:*  $\tilde{\boldsymbol{\beta}} = \hat{\tilde{\boldsymbol{\beta}}}$  so wählen, dass

- 1)  $\tilde{\mathbf{e}}^\top \tilde{\mathbf{e}} = \min$ , wobei  $\tilde{\mathbf{e}} = \tilde{\mathbf{y}} - \mathbf{X}\hat{\tilde{\boldsymbol{\beta}}}$
- 2)  $\hat{\tilde{\boldsymbol{\beta}}}^\top \hat{\tilde{\boldsymbol{\beta}}} = \min$ , d.h.  $\hat{\tilde{\beta}}_{r+1} = \dots = \hat{\tilde{\beta}}_p = 0$ .

Dann sind auch die Schätzfunktionen eindeutig festgelegt:  $\hat{\boldsymbol{\beta}} = \mathbf{D}^+ \tilde{\mathbf{y}}$ . Weiter gilt

$$E(\hat{\boldsymbol{\beta}}) = \mathbf{D}^+ E(\tilde{\mathbf{y}}) = \mathbf{D}^+ \mathbf{D}\tilde{\boldsymbol{\beta}} = \tilde{\boldsymbol{\beta}}$$

$$\text{var}(\hat{\boldsymbol{\beta}}) = \sigma^2 (\mathbf{D}^\top \mathbf{D})^+ = \sigma^2 \left( \begin{array}{c|c} \mathbf{D}_1^{-2} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} \end{array} \right) = (p \times p)$$

$\hat{\boldsymbol{\beta}}$  ist der BLUE-Schätzer für  $\tilde{\boldsymbol{\beta}}$

$$\hat{\sigma}^2 = \frac{1}{n-r} \sum_{i=1}^n \tilde{e}_i^2 \quad \text{ist ein erwartungstreuer Schätzer für } \sigma^2.$$

*Rücktransformation zum Originalmodell:*

Da  $\mathbf{e}^\top \mathbf{e} = \tilde{\mathbf{e}}^\top \tilde{\mathbf{e}}$ , so erhalten wir die KQ-Schätzungen im Originalmodell durch einfache Rücktransformation. Es gilt

$$\boldsymbol{\beta} = \mathbf{V}\tilde{\boldsymbol{\beta}} \quad \text{und somit} \quad \hat{\boldsymbol{\beta}} = \mathbf{V}\hat{\tilde{\boldsymbol{\beta}}} = \mathbf{V}\mathbf{D}^+ \tilde{\mathbf{y}} = \mathbf{V}\mathbf{D}^+ \mathbf{U}^\top \mathbf{y} = \mathbf{X}^+ \mathbf{y}$$

$$E(\hat{\boldsymbol{\beta}}) = \mathbf{V} E(\hat{\tilde{\boldsymbol{\beta}}}) = \mathbf{V}\tilde{\boldsymbol{\beta}} = \boldsymbol{\beta}$$

$$(18) \quad \text{var}(\hat{\boldsymbol{\beta}}) = \mathbf{X}^+ \text{var}(\mathbf{y})(\mathbf{X}^+)^{\top} = \sigma^2 \mathbf{X}^+ (\mathbf{X}^+)^{\top} = \sigma^2 (\mathbf{X}^\top \mathbf{X})^+ = \mathbf{V}(\mathbf{D}^\top \mathbf{D})^+ \mathbf{V}^\top = (p \times p)$$

$\hat{\boldsymbol{\beta}}$  ist der BLUE-Schätzer für  $\boldsymbol{\beta}$

$$\hat{\sigma}^2 = \frac{1}{n-r} \sum_{i=1}^n e_i^2 \quad \text{ist ein erwartungstreuer Schätzer für } \sigma^2.$$

Falls die Matrix  $\mathbf{X}$  vollen Rang besitzt, d.h. falls  $r = \text{rk}(\mathbf{X}) = p$ , so gilt

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y}$$

$$(19) \quad \text{var}(\hat{\boldsymbol{\beta}}) = \sigma^2 (\mathbf{X}^T \mathbf{X})^{-1}$$

$$\hat{\sigma}^2 = \frac{1}{n-p} \mathbf{e}^T \mathbf{e} \text{ ist ein erwartungstreuer Schätzer für } \sigma^2, \text{ wobei } \mathbf{e} = \mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}$$

(vgl. Toutenburg, 2003, S. 107 ff.), und dies ist äquivalent zu

$$\hat{\boldsymbol{\beta}} = \mathbf{X}^+ \mathbf{y}$$

$$(20) \quad \text{var}(\hat{\boldsymbol{\beta}}) = \sigma^2 (\mathbf{X}^T \mathbf{X})^+$$

$$\hat{\sigma}^2 = \frac{1}{n-r} \mathbf{e}^T \mathbf{e} \text{ ist ein erwartungstreuer Schätzer für } \sigma^2, \text{ wobei } \mathbf{e} = \mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}.$$

Die Beziehung (20) ist jedoch auch gültig im Falle  $r = \text{rk}(\mathbf{X}) < p$ . Numerische Beispiele zum einfachen und allgemeinen (mit beliebiger Kovarianzmatrix) linearen Modell finden sich in Knüsel (2009).

## 7. Singulärwert-Zerlegung und Eigenwert-Zerlegung

Es sei  $\mathbf{A} = (m \times n)$  eine beliebige reelle Matrix. Dann existieren orthogonale Matrizen  $\mathbf{U} = (m \times m)$  und  $\mathbf{V} = (n \times n)$  so, dass

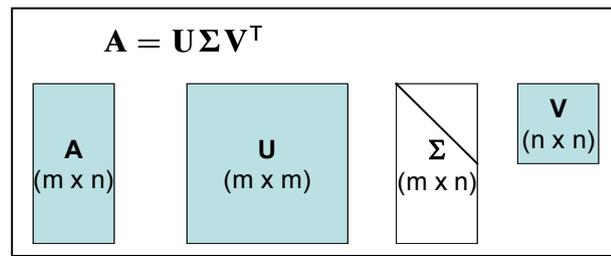
$$(21) \quad \mathbf{A} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T,$$

wobei

$$\boldsymbol{\Sigma} = (m \times n) = \text{diag}(\sigma_1, \dots, \sigma_p)$$

mit

$$p = \min(m, n), \quad \sigma_1 \geq \dots \geq \sigma_r > \sigma_{r+1} = \dots = \sigma_p = 0, \quad r = \text{rk}(\mathbf{A}).$$



Die Werte  $\sigma_1, \dots, \sigma_p$  heißen *singuläre Werte* von  $\mathbf{A}$ .

*Beweis nach Stewart (1973, S. 319)*

Die Matrix  $\mathbf{B} = \mathbf{A}^T \mathbf{A} = (p \times p)$  ist symmetrisch und positiv semidefinit. Es sei  $\mathbf{B} = \mathbf{U}^T \boldsymbol{\Lambda} \mathbf{U}$  die Eigenwertzerlegung von  $\mathbf{B}$ , wobei  $\boldsymbol{\Lambda} = \text{diag}(\lambda_1, \dots, \lambda_p)$  mit  $\lambda_1 \geq \dots \geq \lambda_r > \lambda_{r+1} = \dots = \lambda_p = 0$ . Wir definieren nun  $\sigma_1 \geq \dots \geq \sigma_r > \sigma_{r+1} = \dots = \sigma_p = 0$  so, dass  $\sigma_i^2 = \lambda_i$  für  $i = 1, \dots, p$ . Weiter sei

$$\mathbf{U} = (\mathbf{u}_1, \dots, \mathbf{u}_p) = (p \times p);$$

$$\mathbf{U}_1 = (\mathbf{u}_1, \dots, \mathbf{u}_r) = (p \times r);$$

$$\mathbf{U}_2 = (\mathbf{u}_{r+1}, \dots, \mathbf{u}_p) = (p \times (p-r));$$

$$\boldsymbol{\Sigma} = \text{diag}(\sigma_1, \dots, \sigma_r) = (r \times r).$$

Dann gilt

$$\mathbf{U}_1^T \mathbf{B} \mathbf{U}_1 = \text{diag}(\lambda_1, \dots, \lambda_r) = \boldsymbol{\Sigma}^2;$$

$$\mathbf{U}_2^T \mathbf{B} \mathbf{U}_2 = \mathbf{0}.$$

Daraus ergibt sich

$$(22) \quad \boldsymbol{\Sigma}^{-1} \mathbf{U}_1^T \mathbf{B} \mathbf{U}_1 \boldsymbol{\Sigma}^{-1} = \mathbf{I}_r, \quad \text{und somit } \boldsymbol{\Sigma}^{-1} \mathbf{U}_1^T \mathbf{A}^T \mathbf{A} \mathbf{U}_1 \boldsymbol{\Sigma}^{-1} = \mathbf{I}_r;$$

$$\mathbf{U}_2^T \mathbf{B} \mathbf{U}_2 = \mathbf{U}_2^T \mathbf{A}^T \mathbf{A} \mathbf{U}_2 = (\mathbf{A} \mathbf{U}_2)^T (\mathbf{A} \mathbf{U}_2) = \mathbf{0}, \quad \text{und somit } \mathbf{A} \mathbf{U}_2 = \mathbf{0}.$$

Nun sei  $\mathbf{V}_1 = \mathbf{A}\mathbf{U}_1\boldsymbol{\Sigma}^{-1} = (m \times r)$ . Gemäß (22) ist  $\mathbf{V}_1^T \mathbf{V}_1 = \mathbf{I}_r$ ; die  $r$  Spalten von  $\mathbf{V}_1$  sind also ortho- normiert. Wir wählen  $\mathbf{V}_2$  so, dass die Matrix  $\mathbf{V} = (\mathbf{V}_1 | \mathbf{V}_2) = (m \times m)$  orthogonal wird; es ist dann insbesondere  $\mathbf{V}_2^T \mathbf{V}_1 = \mathbf{0}$ . Und jetzt erhalten wir

$$\mathbf{V}^T \mathbf{A} \mathbf{U} = \begin{pmatrix} \mathbf{V}_1^T \mathbf{A} \mathbf{U}_1 & \mathbf{V}_1^T \mathbf{A} \mathbf{U}_2 \\ \mathbf{V}_2^T \mathbf{A} \mathbf{U}_1 & \mathbf{V}_2^T \mathbf{A} \mathbf{U}_2 \end{pmatrix} = \begin{pmatrix} (\boldsymbol{\Sigma}^{-1} \mathbf{U}_1^T \mathbf{A}^T) \mathbf{A} \mathbf{U}_1 & \mathbf{0} \\ \mathbf{V}_2^T \mathbf{V}_1 \boldsymbol{\Sigma} & \mathbf{0} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\Sigma} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix},$$

was zu beweisen war.

### Bemerkungen

Die Singulärwert-Zerlegung ist in der Fachliteratur bereits bei Eckart-Young (1936, 1939) zu finden. Bisher haben wir die singulären Werte mit  $d_1, d_2, \dots$  bezeichnet, um Verwechslungen mit der Varianz  $\sigma^2$  auszuschließen. Jetzt bevorzugen wir die in der Fachliteratur übliche Bezeichnung  $\sigma_1, \sigma_2, \dots$ .

### Sparversion der Singulärwert-Zerlegung (thin singular value decomposition)

Wir nehmen jetzt an, dass  $m \geq n$  (im Falle  $m \leq n$  ist alles analog); dann ist  $p = \min(m, n) = n$ , und wir schreiben

$$\boldsymbol{\Sigma} = (m \times n) = \begin{pmatrix} \boldsymbol{\Sigma}_1 \\ \mathbf{0} \end{pmatrix} \quad \text{mit } \boldsymbol{\Sigma}_1 = (n \times n) = \text{diag}(\sigma_1, \dots, \sigma_n)$$

$$\mathbf{U} = (\mathbf{U}_1, \mathbf{U}_2) \quad \text{mit } \mathbf{U}_1 = (m \times n)$$

Nun erhalten wir aus (21)

$$(23) \quad \mathbf{A} = \mathbf{U}_1 \boldsymbol{\Sigma}_1 \mathbf{V}^T.$$

Die Darstellung (21) ist bequemer für algebraische Umformungen wegen der Orthogonalität von  $\mathbf{U}$

(vgl. die obigen Abschnitte zum linearen Modell), und die Darstellung (23) ist nützlich bei numerischen Berechnungen, weil Speicherplatz gespart und unnötige Berechnungen vermieden werden können. Die Beziehung (23) wird bei Golub-Van Loan (1996), S. 71, als *Sparversion der Singulärwert-Zerlegung (thin singular value decomposition)* bezeichnet. Die Beziehung (23) kann auch geschrieben werden als

$$\mathbf{A} = \mathbf{U}_1 \boldsymbol{\Sigma}_1 \mathbf{V}^T = \sum_{i=1}^n \sigma_i \mathbf{u}_i \mathbf{v}_i^T, \quad \text{wobei } \mathbf{U}_1 = (\mathbf{u}_1, \dots, \mathbf{u}_n), \quad \mathbf{V} = (\mathbf{v}_1, \dots, \mathbf{v}_n).$$

Diese Darstellung zeigt, dass die Reihenfolge und das Vorzeichen der singulären Werte beliebig geändert werden können durch Umstellen der Spalten von  $\mathbf{U}_1$  und  $\mathbf{V}$ , bzw. durch Ändern der Vorzeichen der Spalten.

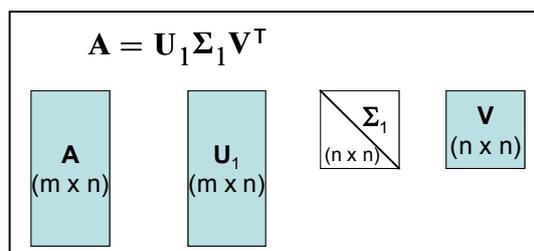
### Zusammenhang mit Eigenwert-Zerlegung

Wir nehmen wieder an, dass  $m \geq n$ . Aus (21) ergibt sich

$\mathbf{A} \mathbf{A}^T = \mathbf{U} \boldsymbol{\Sigma} \mathbf{V}^T \mathbf{V} \boldsymbol{\Sigma}^T \mathbf{U}^T = \mathbf{U} \boldsymbol{\Sigma} \boldsymbol{\Sigma}^T \mathbf{U}^T = \mathbf{U} \boldsymbol{\Lambda} \mathbf{U}^T$  mit  $\boldsymbol{\Lambda} = \boldsymbol{\Sigma} \boldsymbol{\Sigma}^T = (m \times m) = \text{diag}(\lambda_1, \dots, \lambda_n, 0, \dots, 0)$ , wobei  $\lambda_i = \sigma_i^2$ . Die Darstellung  $\mathbf{A} \mathbf{A}^T = \mathbf{U} \boldsymbol{\Lambda} \mathbf{U}^T$  ist die Eigenwert-Zerlegung (Spektraldarstellung, Hauptachsentransformation) der symmetrischen Matrix  $\mathbf{A} \mathbf{A}^T$ . Weiter gilt

$$\mathbf{A}^T \mathbf{A} = \mathbf{V} \boldsymbol{\Sigma}^T \mathbf{U}^T \mathbf{U} \boldsymbol{\Sigma} \mathbf{V}^T = \mathbf{V} \boldsymbol{\Sigma}^T \boldsymbol{\Sigma} \mathbf{V}^T = \mathbf{V} \tilde{\boldsymbol{\Lambda}} \mathbf{V}^T \quad \text{mit } \tilde{\boldsymbol{\Lambda}} = \boldsymbol{\Sigma}^T \boldsymbol{\Sigma} = (n \times n) = \text{diag}(\lambda_1, \dots, \lambda_n).$$

Die drei Bausteine  $\mathbf{U}$ ,  $\boldsymbol{\Sigma}$ ,  $\mathbf{V}$  der Singulärwert-Zerlegung von  $\mathbf{A}$  kann man aber nicht ohne weitere Anpassung aus den Eigenwert-Zerlegungen von  $\mathbf{A} \mathbf{A}^T$  und  $\mathbf{A}^T \mathbf{A}$  übernehmen, da ja die orthogonalen Matrizen  $\mathbf{U}$  und  $\mathbf{V}$  bei mehrfachen Eigenwerten nicht eindeutig festgelegt sind.



### Zur numerischen Berechnung der Singulärwert-Zerlegung

Wie in Stoer-Bulirsch (2005), S. 78 gezeigt wird, gibt es zur numerischen Berechnung der Singulärwert-Zerlegung einer Matrix  $\mathbf{A} = (m \times n) = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$  stabile direkte Verfahren (Methode von Golub-Reinsch). Die Singulärwert-Zerlegung sollte nicht mit Hilfe der Eigenwert-Zerlegungen der symmetrischen Matrizen  $\mathbf{A}^T\mathbf{A}$  und  $\mathbf{A}\mathbf{A}^T$  durchgeführt werden, da dabei ein Verlust an Genauigkeit auftreten kann, wie das folgende Beispiel zeigt.

*Beispiel (vgl. Stoer-Bulirsch, 2005, S. 78)*

Es sei

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

Die singulären Werte von  $\mathbf{A}$  sind  $\sigma_1 = \sqrt{2 + \varepsilon^2}$ ,  $\sigma_2 = \varepsilon$ , wie man mit Maple nachprüfen kann. Weiter ist

$$\mathbf{A}^T\mathbf{A} = \begin{pmatrix} 1 + \varepsilon^2 & 1 \\ 1 & 1 + \varepsilon^2 \end{pmatrix},$$

und die Eigenwerte von  $\mathbf{A}^T\mathbf{A}$  sind  $\lambda_1 = 2 + \varepsilon^2 (= \sigma_1^2)$  und  $\lambda_2 = \varepsilon^2 (= \sigma_2^2)$ . Wenn nun die numerische Rechengenauigkeit z.B. 10 signifikante Ziffern beträgt, und wenn  $\varepsilon < 10^{-5}$  ist, so ist  $1 + \varepsilon^2 = 1$  (numerisch), und die numerische Darstellung von  $\mathbf{X}^T\mathbf{X}$  ist

$$\mathbf{X}^T\mathbf{X} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

mit den Eigenwerten  $\lambda_1 = 2$  und  $\lambda_2 = 0$ , und daraus würde man erhalten  $\sigma_1 = \sqrt{\lambda_1} = 2$  und  $\sigma_2 = \sqrt{\lambda_2} = 0$ . Während  $\sigma_1$  numerisch korrekt ist, ist der Wert  $\sigma_2 = 0$  deutlich verschieden vom korrekten Wert  $\sigma_2 = \varepsilon$  (relativer Fehler 100%).

Die Programme Maple, Matlab und R berechnen die Singulärwert-Zerlegung auf direktem Wege ohne den Umweg über die Eigenwert-Zerlegung, und sie bieten die Möglichkeit, nur einzelne Komponenten der Zerlegung zu berechnen, um Speicherplatz zu sparen, z.B. nur die Sparversion. Auf einem Standard-PC kann die vollständige Zerlegung mit  $m = 1000$  und  $n = 100$  in wenigen Sekunden berechnet werden.

## 8. Singulärwert-Zerlegung und verallgemeinerte Inverse

Definition der verallgemeinerten Inversen (vgl. Toutenburg, 2003, S. 504)

Es sei  $\mathbf{A} = (m \times n)$  eine beliebige reelle Matrix. Eine Matrix  $\mathbf{A}^+ = (n \times m)$  mit

- 1)  $\mathbf{A}\mathbf{A}^+\mathbf{A} = \mathbf{A}$
- 2)  $\mathbf{A}^+\mathbf{A}\mathbf{A}^+ = \mathbf{A}^+$
- 3)  $(\mathbf{A}\mathbf{A}^+)^T = \mathbf{A}\mathbf{A}^+ = (m \times m)$  (Symmetrie von  $\mathbf{A}\mathbf{A}^+$ )
- 4)  $(\mathbf{A}^+\mathbf{A})^T = \mathbf{A}^+\mathbf{A} = (n \times n)$  (Symmetrie von  $\mathbf{A}^+\mathbf{A}$ )

heißt *verallgemeinerte Inverse (Moore-Penrose-Inverse, Pseudo-Inverse)* von  $\mathbf{A}$  (vgl. Penrose, 1955).

**Satz:** Zu jeder reellen Matrix  $\mathbf{A} = (m \times n)$  existiert genau eine verallgemeinerte Inverse  $\mathbf{A}^+$ .

1° *Beweis der Eindeutigkeit* (vgl. Stör, 1999, S. 248)

Es seien  $\mathbf{U}$  und  $\mathbf{V}$  zwei Exemplare von verallgemeinerten Inversen  $\mathbf{A}^+$  mit den Eigenschaften

1) bis 4). Wir wollen zeigen, dass die beiden Matrizen identisch sein müssen. Es gilt

$$\begin{aligned} \mathbf{U} &= \underbrace{\mathbf{U}\mathbf{A}\mathbf{U}}_{\text{wegen 2)}} = \mathbf{U}(\underbrace{\mathbf{A}\mathbf{V}\mathbf{A}}_{=\mathbf{A}})\mathbf{U} = \mathbf{U}(\underbrace{\mathbf{A}\mathbf{V}\mathbf{A}}_{=\mathbf{A}})\mathbf{V}(\underbrace{\mathbf{A}\mathbf{V}\mathbf{A}}_{=\mathbf{A}})\mathbf{U} \\ &= \underbrace{(\mathbf{U}\mathbf{A})^T}_{=\mathbf{U}\mathbf{A}} \underbrace{(\mathbf{V}\mathbf{A})^T}_{=\mathbf{V}\mathbf{A}} \mathbf{V} \underbrace{(\mathbf{A}\mathbf{V})^T}_{=\mathbf{A}\mathbf{V}} \underbrace{(\mathbf{A}\mathbf{U})^T}_{=\mathbf{A}\mathbf{U}} \\ &= \underbrace{(\mathbf{A}^T\mathbf{U}^T\mathbf{A}^T)}_{=\mathbf{A}^T} \mathbf{V}^T \mathbf{V} \mathbf{V}^T \underbrace{(\mathbf{A}^T\mathbf{U}^T\mathbf{A}^T)}_{=\mathbf{A}^T} \\ &= \underbrace{\mathbf{A}^T\mathbf{V}^T}_{=\mathbf{V}\mathbf{A}} \underbrace{\mathbf{V}\mathbf{V}^T\mathbf{A}^T}_{=\mathbf{A}\mathbf{V}} = \mathbf{V} \underbrace{\mathbf{A}\mathbf{V}\mathbf{A}}_{=\mathbf{A}} \mathbf{V} = \mathbf{V}\mathbf{A}\mathbf{V} = \mathbf{V}. \end{aligned}$$

Somit ist die die Moore-Penrose-Inverse eindeutig bestimmt.

2° *Existenz; Zusammenhang mit Singulärwert-Zerlegung*

Die Singulärwert-Zerlegung von  $\mathbf{A}$  hat die Form  $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$ , wobei  $\mathbf{U}, \mathbf{V}$  orthogonal und

$\mathbf{\Sigma} = (m \times n) = \text{diag}(\sigma_1, \dots, \sigma_p)$  mit  $p = \min(m, n)$  und  $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0, \sigma_{r+1} = \dots = \sigma_p = 0$ . Es sei

$$\begin{aligned} \mathbf{\Sigma}^+ &= (n \times m) = \text{diag}\left(\frac{1}{\sigma_1}, \dots, \frac{1}{\sigma_r}, 0, \dots, 0\right), \\ \mathbf{A}^+ &= \mathbf{V}\mathbf{\Sigma}^+\mathbf{U}^T. \end{aligned}$$

Dann sind  $\mathbf{\Sigma}^+$  und  $\mathbf{A}^+$  die verallgemeinerten Inversen zu  $\mathbf{\Sigma}$  und  $\mathbf{A}$ , denn es gilt

$$\mathbf{\Sigma}\mathbf{\Sigma}^+ = (m \times m) = \begin{pmatrix} \mathbf{I}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \quad \text{und} \quad \mathbf{\Sigma}^+\mathbf{\Sigma} = (n \times n) = \begin{pmatrix} \mathbf{I}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix},$$

und daraus folgen sofort die Eigenschaften 1) bis 4) für  $\mathbf{\Sigma}$  und  $\mathbf{\Sigma}^+$ . Weiter gilt

- 1)  $\mathbf{A}\mathbf{A}^+\mathbf{A} = \underbrace{\mathbf{U}\mathbf{\Sigma}\mathbf{V}^T}_{=\mathbf{A}} \underbrace{\mathbf{V}\mathbf{\Sigma}^+\mathbf{U}^T}_{=\mathbf{A}^+} \underbrace{\mathbf{U}\mathbf{\Sigma}\mathbf{V}^T}_{=\mathbf{A}} = \mathbf{U} \underbrace{\mathbf{\Sigma}\mathbf{\Sigma}^+\mathbf{\Sigma}}_{=\mathbf{\Sigma}} \mathbf{V}^T = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T = \mathbf{A}$
- 2)  $\mathbf{A}^+\mathbf{A}\mathbf{A}^+ = \underbrace{\mathbf{V}\mathbf{\Sigma}^+\mathbf{U}^T}_{=\mathbf{A}^+} \underbrace{\mathbf{U}\mathbf{\Sigma}\mathbf{V}^T}_{=\mathbf{A}} \underbrace{\mathbf{V}\mathbf{\Sigma}^+\mathbf{U}^T}_{=\mathbf{A}^+} = \mathbf{V} \underbrace{\mathbf{\Sigma}^+\mathbf{\Sigma}\mathbf{\Sigma}^+}_{=\mathbf{\Sigma}^+} \mathbf{U}^T = \mathbf{V}\mathbf{\Sigma}^+\mathbf{U}^T = \mathbf{A}^+$
- 3)  $\mathbf{A}\mathbf{A}^+ = \underbrace{\mathbf{U}\mathbf{\Sigma}\mathbf{V}^T}_{=\mathbf{A}} \underbrace{\mathbf{V}\mathbf{\Sigma}^+\mathbf{U}^T}_{=\mathbf{A}^+} = \mathbf{U}\mathbf{\Sigma}\mathbf{\Sigma}^+\mathbf{U}^T$  ist offensichtlich symmetrisch.
- 4)  $\mathbf{A}^+\mathbf{A} = \underbrace{\mathbf{V}\mathbf{\Sigma}^+\mathbf{U}^T}_{=\mathbf{A}^+} \underbrace{\mathbf{U}\mathbf{\Sigma}\mathbf{V}^T}_{=\mathbf{A}} = \mathbf{V}\mathbf{\Sigma}^+\mathbf{\Sigma}\mathbf{V}^T$  ist offensichtlich symmetrisch

Somit ist  $\mathbf{A}^+$  die verallgemeinerte Inverse zu  $\mathbf{A}$ . Die Singulärwert-Zerlegung liefert also eine einfache und übersichtliche Darstellung der Moore-Penrose-Inversen!

## 9. Rang-k-Approximation mit Singulärwert-Zerlegung

### a) Vektor- und Matrix-Normen

Es sei  $\mathbf{x}^T = (x_1, \dots, x_n)$  ein reeller Vektor und  $\mathbf{A} = (m \times n)$  eine beliebige reelle Matrix. Dann ist

$$\|\mathbf{x}\| = \sqrt{x_1^2 + \dots + x_n^2} \quad \text{die euklidische Vektornorm}$$

$$\|\mathbf{A}\| = \max_{\mathbf{x} \neq 0} \frac{\|\mathbf{Ax}\|}{\|\mathbf{x}\|} \quad \text{die Grenznorm zur euklidischen Vektornorm (Spektralnorm).}$$

Wenn  $\|\mathbf{A}\|_b$  eine beliebige Matrixnorm ist mit  $\|\mathbf{Ax}\| \leq \|\mathbf{A}\|_b \|\mathbf{x}\|$  für jeden Vektor  $\mathbf{x}$ , dann gilt offensichtlich

$$\|\mathbf{A}\|_b \geq \max_{\mathbf{x} \neq 0} \frac{\|\mathbf{Ax}\|}{\|\mathbf{x}\|} = \|\mathbf{A}\|.$$

Die oben definierte Grenznorm  $\|\mathbf{A}\|$  ist also eine sehr sinnvolle Matrixnorm zur euklidischen Vektornorm; sie wird auch *Spektralnorm* genannt (vgl. Rao-Rao, 1998, S. 376). Es gilt

$$\|\mathbf{A} + \mathbf{B}\| \leq \|\mathbf{A}\| + \|\mathbf{B}\|, \quad \|\mathbf{AB}\| \leq \|\mathbf{A}\| \times \|\mathbf{B}\|$$

$$\|\mathbf{A}\| = \max_{\mathbf{x} \neq 0} \frac{\|\mathbf{Ax}\|}{\|\mathbf{x}\|} = \sigma_{\max} \quad (\text{größter singulärer Wert von } \mathbf{A})$$

$$\min_{\mathbf{x} \neq 0} \frac{\|\mathbf{Ax}\|}{\|\mathbf{x}\|} = \sigma_{\min} \quad (\text{kleinster singulärer Wert von } \mathbf{A}).$$

Die Kondition einer Matrix  $\mathbf{A}$   $\text{cond}(\mathbf{A}) = \sigma_{\max} / \sigma_{\min}$  ist ein Maß für die Empfindlichkeit des relativen Fehlers  $\|\Delta \mathbf{x}\| / \|\mathbf{x}\|$  der Lösung  $\mathbf{x}$  eines linearen Gleichungssystems  $\mathbf{Ax} = \mathbf{b}$  gegenüber Änderungen der Ausgangsdaten  $\mathbf{A}$  und  $\mathbf{b}$  (vgl. Stoer, 2005, S. 217 ff; Golub-Van Loan, 1996, S. 52 ff).

### b) Ungleichung zur Spektralnorm (vgl. Rao-Rao, 1998, S. 365)

Es sei  $\mathbf{A} = (a_{ij}) = (m \times n)$  eine beliebige reelle Matrix und  $\sigma_{\max}$  der größte singuläre Wert von  $\mathbf{A}$ .

Dann gilt

$$\max_{i,j} |a_{ij}| \leq \sigma_{\max} (= \|\mathbf{A}\|).$$

Beweis:

1° Es sei  $\mathbf{I}_{ij} = (m \times n)$  eine Matrix, bei der nur das Element  $(i, j)$  mit einer Eins besetzt, alle übrigen Elemente sind null. Dann gilt  $\|\mathbf{I}_{ij}\| = 1$  für beliebige Paare  $(i, j)$ . Anstelle eines allgemeinen Beweises geben wir ein Beispiel. Es sei

$$\mathbf{I}_{ij} = (3 \times 4) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \mathbf{U} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \mathbf{V} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad \mathbf{D} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Dann gilt offensichtlich  $\mathbf{U} \mathbf{I}_{ij} \mathbf{V} = \mathbf{D} = (3 \times 4) = \text{diag}(1, 0, 0)$ , denn für eine beliebige Matrix  $\mathbf{B} = (3 \times 4)$  werden bei der Multiplikation  $\mathbf{UB}$  die Zeilen 1 und 2 von  $\mathbf{B}$  vertauscht, und bei der Multiplikation  $\mathbf{BV}$  werden die Spalten 1 und 3 vertauscht. Die quadratischen und symmetrischen Matrizen  $\mathbf{U}$  und  $\mathbf{V}$  sind orthogonal, da in jeder Zeile und Spalte genau eine Eins steht. Aus der Beziehung  $\mathbf{U} \mathbf{I}_{ij} \mathbf{V} = \mathbf{D}$  ergibt sich die Singulärwert-Zerlegung von  $\mathbf{I}_{ij}$

$$\mathbf{I}_{ij} = \mathbf{UDV}^T, \quad \text{wobei } \mathbf{D} = (3 \times 4) = \text{diag}(1, 0, 0);$$

der größte singuläre Wert von  $\mathbf{I}_{ij}$  hat also stets den Wert 1 unabhängig von der Position  $(i, j)$ , und damit gilt  $\|\mathbf{I}_{ij}\| = 1$  für beliebige Paare  $(i, j)$ .

2° Es gilt  $\mathbf{I}_{ij}\mathbf{A}\mathbf{I}_{ij} = a_{ij}\mathbf{I}_{ij}$  und  $\|a_{ij}\mathbf{I}_{ij}\| = |a_{ij}| \times \|\mathbf{I}_{ij}\| = |a_{ij}|$ . Somit ist

$$|a_{ij}| = \|a_{ij}\mathbf{I}_{ij}\| = \|\mathbf{I}_{ij}\mathbf{A}\mathbf{I}_{ij}\| \leq \|\mathbf{A}\| \times \|\mathbf{I}_{ij}\|^2 = \|\mathbf{A}\| \text{ für beliebige } i, j,$$

und daher gilt

$$\max_{i,j} |a_{ij}| \leq \|\mathbf{A}\| = \sigma_{\max}, \text{ was zu beweisen war.}$$

c) *Optimale Rang-k-Approximation mit Spektralnorm*

Es sei  $\mathbf{A} = (m \times n)$  eine beliebige reelle Matrix mit  $m \geq n$ . Wir wollen die Matrix  $\mathbf{A}$  approximieren mit einer Matrix  $\mathbf{A}_k = (m \times n)$  mit  $\text{rk}(\mathbf{A}_k) = k < r = \text{rk}(\mathbf{A})$ . Es sei  $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$  die Singulärwert-Zerlegung von  $\mathbf{A}$ , wobei  $\mathbf{\Sigma} = (m \times n) = \text{diag}(\sigma_1, \dots, \sigma_n)$  mit  $\sigma_1 \geq \dots \geq \sigma_r > \sigma_{r+1} = \dots = \sigma_n = 0$ . Wir setzen

$$\mathbf{\Sigma}_k = (m \times n) = \text{diag}(\sigma_1, \dots, \sigma_k, 0, \dots, 0) \text{ und } \mathbf{A}_k = \mathbf{U}\mathbf{\Sigma}_k\mathbf{V}^T.$$

Dann gilt:

$$\min_{\text{rk}(\mathbf{B})=k} \|\mathbf{A} - \mathbf{B}\| = \|\mathbf{A} - \mathbf{A}_k\| = \sigma_{k+1}.$$

Die Matrix  $\mathbf{A}_k$  besitzt also unter allen Matrizen mit Rang  $k$  den kleinsten Abstand von  $\mathbf{A}$ .

Beweis (nach Golub-Van Loan, 1996, S. 73)

1° Es gilt

$$\mathbf{A} - \mathbf{A}_k = \mathbf{U}(\mathbf{\Sigma} - \mathbf{\Sigma}_k)\mathbf{V}^T, \text{ wobei } \mathbf{\Sigma} - \mathbf{\Sigma}_k = \text{diag}(0, \dots, 0, \sigma_{k+1}, \dots, \sigma_n)$$

und somit gilt

$$\|\mathbf{A} - \mathbf{A}_k\| = \sigma_{k+1} = \text{größter singulärer Wert von } \mathbf{A} - \mathbf{A}_k.$$

2° Die Matrizen  $\mathbf{A}$  und  $\mathbf{A}_k$  können geschrieben werden als

$$\mathbf{A} = \sum_{i=1}^n \sigma_i \mathbf{u}_i \mathbf{v}_i^T \text{ und } \mathbf{A}_k = \sum_{i=1}^k \sigma_i \mathbf{u}_i \mathbf{v}_i^T, \text{ wobei } \mathbf{U} = (\mathbf{u}_1, \dots, \mathbf{u}_m), \mathbf{V} = (\mathbf{v}_1, \dots, \mathbf{v}_n).$$

Nun sei  $\mathbf{B} = (m \times n)$  eine Matrix mit  $\text{rk}(\mathbf{B}) = k$ , und  $\text{null}(\mathbf{B}) = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{B}\mathbf{x} = 0\}$  sei der Nullraum zu  $\mathbf{B}$ ; er besitzt die Dimension  $n-k$ . Weiter sei  $\text{span}(\mathbf{v}_1, \dots, \mathbf{v}_{k+1})$  der Vektorraum, der von  $\mathbf{v}_1, \dots, \mathbf{v}_{k+1}$  aufgespannt wird. Die beiden Räume können keinen leeren Durchschnitt haben, d.h.

$\text{null}(\mathbf{B}) \cap \text{span}(\mathbf{v}_1, \dots, \mathbf{v}_{k+1}) \neq \emptyset$ , da ja sonst die Vereinigung die Dimension

$(n-k) + (k+1) = n+1$  haben müsste. Es sei nun  $\mathbf{z}$  ein Vektor aus dem Durchschnitt mit  $\|\mathbf{z}\| = 1$ .

$\mathbf{z}$  kann geschrieben werden als  $\sum_{i=1}^{k+1} \lambda_i \mathbf{v}_i$  mit  $\sum \lambda_i^2 = 1$ . Nun ist  $\mathbf{B}\mathbf{z} = 0$  und

$$\mathbf{A}\mathbf{z} = \left( \sum_{i=1}^n \sigma_i \mathbf{u}_i \mathbf{v}_i^T \right) \left( \sum_{j=1}^{k+1} \lambda_j \mathbf{v}_j \right) = \sum_{i=1}^{k+1} \sigma_i \lambda_i \mathbf{u}_i \mathbf{v}_i^T \mathbf{v}_i = \sum_{i=1}^{k+1} \sigma_i \lambda_i \mathbf{u}_i,$$

da ja

$$\mathbf{v}_i^T \mathbf{v}_j = \begin{cases} 0 & \text{falls } i \neq j \\ 1 & \text{falls } i = j. \end{cases}$$

Somit gilt

$$\|\mathbf{A} - \mathbf{B}\|^2 \geq \|(\mathbf{A} - \mathbf{B})\mathbf{z}\|^2 = \|\mathbf{A}\mathbf{z}\|^2 = \left\| \sum_{i=1}^{k+1} \sigma_i \lambda_i \mathbf{u}_i \right\|^2 = \sum_{i=1}^{k+1} (\sigma_i \lambda_i)^2 \geq \sigma_{k+1}^2 \sum_{i=1}^{k+1} \lambda_i^2 = \sigma_{k+1}^2.$$

Die Matrix  $\mathbf{B}$  kann also nicht näher bei  $\mathbf{A}$  liegen als die Matrix  $\mathbf{A}_k$ .

*Folgerung:* Es sei  $\mathbf{\Delta} = \mathbf{A} - \mathbf{A}_k = (\delta_{ij})$ . Dann gilt gemäß b):  $\max_{i,j} |\delta_{ij}| \leq \|\mathbf{\Delta}\| = \|\mathbf{A} - \mathbf{A}_k\| = \sigma_{k+1}$ .

d) *Optimale Rang-k-Approximation mit Frobenius-Norm*

Es sei  $\mathbf{A} = (m \times n)$  eine beliebige reelle Matrix mit  $m \geq n$ . Wir wollen die Matrix  $\mathbf{A}$  approximieren mit einer Matrix  $\mathbf{A}_k = (m \times n)$  mit  $\text{rk}(\mathbf{A}_k) = k < r = \text{rk}(\mathbf{A})$  und zwar so, dass

$$\|\mathbf{A} - \mathbf{A}_k\|_F = \min_{\text{rk}(\mathbf{B})=k} \|\mathbf{A} - \mathbf{B}\|_F.$$

Dabei verwenden wir hier nicht die Spektralnrm wie im Absatz c) sondern die Frobeniusnorm  $\|\mathbf{A}\|_F$ , welche definiert ist als

$$\|\mathbf{A}\|_F^2 = \sum_{i,j} a_{ij}^2 = \text{tr}(\mathbf{A}^T \mathbf{A}).$$

Es sei  $\mathbf{R} = (m \times m)$  eine orthogonale Matrix und  $\mathbf{B} = \mathbf{R}\mathbf{A}$ . Dann gilt  $\|\mathbf{B}\|_F = \|\mathbf{A}\|_F$ , denn  $\text{tr}(\mathbf{B}^T \mathbf{B}) = \text{tr}(\mathbf{A}^T \mathbf{R}^T \mathbf{R} \mathbf{A}) = \text{tr}(\mathbf{A}^T \mathbf{A})$ . Dasselbe gilt falls  $\mathbf{R} = (n \times n)$  und  $\mathbf{B} = \mathbf{A}\mathbf{R}$ , denn es gilt  $\text{tr}(\mathbf{B}^T \mathbf{B}) = \text{tr}(\mathbf{B} \mathbf{B}^T) = \text{tr}(\mathbf{A} \mathbf{R} \mathbf{R}^T \mathbf{A}^T) = \text{tr}(\mathbf{A} \mathbf{A}^T) = \text{tr}(\mathbf{A}^T \mathbf{A})$ . Die Frobeniusnorm bleibt also (wie die Grenznorm) invariant bei einer orthogonalen Transformation (von links oder von rechts). Es sei nun  $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$  die Singulärwert-Zerlegung von  $\mathbf{A}$ . Dann gilt

$$\|\mathbf{A} - \mathbf{B}\|_F = \|\mathbf{U}\mathbf{\Sigma}\mathbf{V}^T - \mathbf{B}\|_F = \|\mathbf{\Sigma} - \mathbf{C}\|_F, \text{ wobei } \mathbf{C} = \mathbf{U}^T \mathbf{B} \mathbf{V}.$$

Nun soll  $\mathbf{C} = (c_{ij})$  so gewählt werden, dass  $\text{rk}(\mathbf{C}) = k$  und

$$\|\mathbf{\Sigma} - \mathbf{C}\|_F^2 = \sum_i (\sigma_i - c_{ii})^2 + \sum_{i \neq j} c_{ij}^2 = \min.$$

Dies wird offensichtlich erreicht, wenn  $\mathbf{C} = \text{diag}(\sigma_1, \dots, \sigma_k) = \mathbf{\Sigma}_k$ . Durch Rücktransformation finden wir  $\mathbf{A}_k = \mathbf{U}\mathbf{\Sigma}_k\mathbf{V}^T$  wie bei der Approximation mit der Grenznorm, wobei jedoch jetzt

$$\min_{\text{rk}(\mathbf{B})=k} \|\mathbf{A} - \mathbf{B}\|_F = \|\mathbf{A} - \mathbf{A}_k\|_F = \|\mathbf{\Sigma} - \mathbf{\Sigma}_k\|_F = \left( \sum_{i=k+1}^r \sigma_i^2 \right)^{1/2}.$$

Dieser Satz ist bereits zu finden bei Eckart-Young (1936); dabei wird wohl erstmals die Singulärwert-Zerlegung einer rechteckigen Matrix erwähnt. Der obige Beweis findet sich bei Golub-Kahan (1965).

*Ergänzung:*

Der obige Approximationssatz gilt nicht nur für die Spektralnrm und die Frobeniusnorm, sondern für jede beliebige rotationsinvariante (invariant bei orthogonaler Transformation) Matrixnorm  $\|\cdot\|_{ri}$ :

$$\min_{\text{rk}(\mathbf{B})=k} \|\mathbf{A} - \mathbf{B}\|_{ri} = \|\mathbf{A} - \mathbf{A}_k\|_{ri} = \|\mathbf{\Sigma} - \mathbf{\Sigma}_k\|_{ri}$$

(vgl. Rao-Rao, 1998, S. 392), und wenn die Norm  $\|\cdot\|_{ri}$  kompatibel ist mit der euklidischen Vektornorm, d.h. wenn  $\|\mathbf{A}\mathbf{x}\| \leq \|\mathbf{A}\|_{ri} \|\mathbf{x}\|$  für beliebige Vektoren  $\mathbf{x}$ , dann gilt  $\|\mathbf{\Sigma} - \mathbf{\Sigma}_k\|_{ri} \geq \|\mathbf{\Sigma} - \mathbf{\Sigma}_k\| = \sigma_1$ , wie wir unter a) gesehen haben.

e) *Beispiel:* Es sei

$$\mathbf{X} = (3 \times 2) = \begin{pmatrix} 2 & 2\sqrt{2} \\ 7 & 7\sqrt{2} \\ 9 & 9\sqrt{2} \end{pmatrix} \approx \mathbf{X}_1 = \begin{pmatrix} 2 & 2.828 \\ 7 & 9.899 \\ 9 & 12.728 \end{pmatrix}; \quad \mathbf{X} - \mathbf{X}_1 = \begin{pmatrix} 0 & 0.000427 \\ 0 & 0.000495 \\ 0 & -0.000078 \end{pmatrix}.$$

Offensichtlich besitzt  $\mathbf{X}$  den Rang 1, und die singulären Werte (berechnet mit Maple 10) betragen

bei $\mathbf{X}$ (exakte Arithmetik)	$\sigma_1 = \sqrt{402} = 20.0499\dots$	$\sigma_2 = 0$
bei $\mathbf{X}_1$ :	$\sigma_1 = 20.05$	$\sigma_2 = 0.3346 \times 10^{-3}$
bei $\mathbf{X} - \mathbf{X}_1$ :	$\sigma_1 = 0.000658$	$\sigma_2 = 0.$

Der größte singuläre Wert ist stets größer als das betragsmäßig größte Matrixelement. Die Matrix  $\mathbf{X}_1$  ist regulär, und ihr Abstand zur nächsten Rang-1-Matrix beträgt  $\sigma_2 = 0.3346 \times 10^{-3}$ . Nun berechnen wir die symmetrische Matrix  $\mathbf{X}_1^T \mathbf{X}_1 = (2 \times 2)$  und erhalten

$$\mathbf{X}_1^T \mathbf{X}_1 = \begin{pmatrix} 134 & 189.501 \\ 189.501 & 267.990 \end{pmatrix}, \quad (\mathbf{X}_1^T \mathbf{X}_1)^{-1} = \begin{pmatrix} 5.955 \times 10^6 & -4.211 \times 10^6 \\ -4.211 \times 10^6 & 2.978 \times 10^6 \end{pmatrix}$$

$$\text{Eigenwerte von } \mathbf{X}_1^T \mathbf{X}_1: \lambda_1 = 402.0 (= \sigma_1^2) \quad \lambda_2 = 0.1119 \times 10^{-6} (= \sigma_2^2).$$

Die Matrix  $\mathbf{X}^T \mathbf{X}$  hat den Rang 1 wie  $\mathbf{X}$ ,  $\mathbf{X}_1^T \mathbf{X}_1$  ist jedoch regulär wie  $\mathbf{X}_1$ ; die Eigenwerte erhält man als Quadrat der singulären Werte.

Jetzt berechnen wir die optimale Rang-1-Approximation zu  $\mathbf{X}_1$ . Die Singulärwert-Zerlegung liefert

$\mathbf{X}_1 = \mathbf{U} \mathbf{\Sigma} \mathbf{V}^T$ , wobei  $\mathbf{\Sigma} = (3 \times 2) = \text{diag}(\sigma_1, \sigma_2)$  mit  $\sigma_1 = 20.05$  und  $\sigma_2 = 0.3346 \times 10^{-3}$ . Wir setzen  $\mathbf{D}_1 = \text{diag}(\sigma_1, 0)$  und erhalten

$$\mathbf{X}_2 = \mathbf{U} \mathbf{\Sigma}_1 \mathbf{V}^T \approx \begin{pmatrix} 1.999824 & 2.828124 \\ 6.999856 & 9.899102 \\ 9.000151 & 12.727893 \end{pmatrix}; \quad \mathbf{\Delta} = \mathbf{X}_1 - \mathbf{X}_2 = (\delta_{ij}) \approx \begin{pmatrix} 0.000176 & -0.000124 \\ 0.000144 & -0.000102 \\ -0.000151 & 0.000107 \end{pmatrix}.$$

Die singulären Werte bei  $\mathbf{X}_2$  (Daten 15-stellig) betragen  $\sigma_1 = 20.05$  und  $\sigma_2 = 0.7048 \times 10^{-13}$ .

$\mathbf{X}_2$  ist die nächste Rang-1-Matrix von  $\mathbf{X}_1$ ;  $\mathbf{X}_2$  stimmt jedoch nicht mit der Ausgangsmatrix  $\mathbf{X}$  überein.

Für die Differenz  $\mathbf{\Delta} = \mathbf{X}_1 - \mathbf{X}_2 = (\delta_{ij})$  gilt  $\max |\delta_{ij}| = 0.000176 < \|\mathbf{X}_1 - \mathbf{X}_2\| = \sigma_2 = 0.0003346$ , und damit ist auch hier die Ungleichung b) erfüllt. Der Unterschied zwischen  $\mathbf{X}_1$  und  $\mathbf{X}_2$  bewegt sich im Bereich der Rundungsfehler.

#### e) Rundungsfehler und Rang-k-Approximation (vgl. Golub-Van Loan, 1996, S. 73)

Wenn die Daten einer vorliegenden Datenmatrix  $\mathbf{X} = (n \times p)$  z.B. mit 3 Dezimalstellen gegeben sind, d.h. wenn sie korrekt sind bis auf  $\pm 0.001$ , dann kann man sich fragen, ob eine Matrix  $\tilde{\mathbf{X}} = (n \times p)$  existiert mit  $\|\mathbf{X} - \tilde{\mathbf{X}}\| < 0.001$  und mit  $\text{rk}(\tilde{\mathbf{X}}) < \text{rk}(\mathbf{X})$ . Dazu muss man prüfen, ob die Datenmatrix  $\mathbf{X}$  singuläre Werte kleiner als 0.001 besitzt. Wenn nun  $\sigma_1 \geq \dots \geq \sigma_p \geq 0$  die singulären Werte von  $\mathbf{X}$  bezeichnen, und wenn  $\sigma_k > 0.001$  und  $\sigma_{k+1} < 0.001$ , so liefert die Rang-k-Approximation eine Matrix  $\tilde{\mathbf{X}}$  mit  $\text{rk}(\tilde{\mathbf{X}}) = k$  und mit  $\|\mathbf{X} - \tilde{\mathbf{X}}\| = \sigma_{k+1} < 0.001$ . Der Einfluss von Rundungsfehlern auf den Rang von  $\mathbf{X}$  ist damit eliminiert.

## 10. Singuläre Werte von orthogonalen und symmetrischen Matrizen

### a) Singuläre Werte einer orthogonalen Matrix

Es sei  $\mathbf{R} = (n \times n)$  eine orthogonale Matrix. Dann haben alle singulären Werte von  $\mathbf{R}$  den Wert 1.

*Beweis:*

Es sei  $\mathbf{R} = \mathbf{U} \mathbf{\Sigma} \mathbf{V}^T$  die Singulärwert-Zerlegung von  $\mathbf{R}$  mit  $\mathbf{\Sigma} = \text{diag}(\sigma_1, \dots, \sigma_n)$ ,  $\sigma_i \geq 0$ . Dann gilt

$$\begin{aligned} \mathbf{R} \mathbf{R}^T &= \mathbf{U} \mathbf{\Sigma} \mathbf{\Sigma}^T \mathbf{U}^T = \mathbf{I} \\ \Rightarrow \mathbf{\Sigma} \mathbf{\Sigma}^T &= \mathbf{I}, \text{ d.h. } \sigma_i^2 = 1 \text{ für } i = 1, \dots, n \\ \Rightarrow \sigma_i &= 1 \text{ für } i = 1, \dots, n \text{ da ja } \sigma_i \geq 0. \end{aligned}$$

Die Kondition einer orthogonalen Matrix hat somit den Wert  $\text{cond}(\mathbf{R}) = \sigma_{\max} / \sigma_{\min} = 1$ . Ein lineares Gleichungssystem  $\mathbf{R} \mathbf{x} = \mathbf{b}$  kann also numerisch stabil gelöst werden, da ja bei der Berechnung der inversen Matrix  $\mathbf{R}^{-1} = \mathbf{R}^T$  keinerlei Rundungsfehler auftreten können.

b) *Singuläre Werte einer symmetrischen Matrix*

Es sei  $A = (n \times n)$  eine symmetrische Matrix, und  $\lambda_1, \dots, \lambda_n$  seien die Eigenwerte von  $\mathbf{A}$ . Dann sind  $|\lambda_1|, \dots, |\lambda_n|$  die singulären Werte von  $\mathbf{A}$ .

*Beweis:*

Die Eigenwert-Zerlegung (Spektraldarstellung) von  $\mathbf{A}$  hat die Form

$$\mathbf{A} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^T, \text{ wobei } \mathbf{V} \text{ orthogonal und } \mathbf{\Lambda} = \text{diag}(\lambda_1, \dots, \lambda_n).$$

$\lambda_1, \dots, \lambda_n$  sind die Eigenwerte von  $\mathbf{A}$ ; sie sind reell und können auch negativ sein. Es sei  $\sigma_i = |\lambda_i|$  für  $i = 1, \dots, n$ .  $\mathbf{V} = (\mathbf{v}_1, \dots, \mathbf{v}_n)$ ; dann ist  $\mathbf{V}\mathbf{\Lambda} = (\lambda_1 \mathbf{v}_1, \dots, \lambda_n \mathbf{v}_n)$ . Nun definieren wir

$\mathbf{U} = (n \times n) = (\mathbf{u}_1, \dots, \mathbf{u}_n)$  so, dass  $\sigma_i \mathbf{u}_i = \lambda_i \mathbf{v}_i$ ; die Spalten von  $\mathbf{U}$  sind dann identisch zu den Spalten von  $\mathbf{V}$  bis auf das Vorzeichen, und daher gilt  $\mathbf{U}^T \mathbf{U} = \mathbf{V}^T \mathbf{V} = \mathbf{I}$ . Somit ist auch  $\mathbf{U}$  orthogonal und es gilt  $\mathbf{V}\mathbf{\Lambda} = \mathbf{U}\mathbf{\Sigma}$  und somit

$$\mathbf{A} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^T = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T, \text{ wobei } \mathbf{U} \text{ und } \mathbf{V} \text{ orthogonal, und } \mathbf{\Sigma} = \text{diag}(\sigma_1, \dots, \sigma_n), \sigma_i = |\lambda_i|.$$

Die Darstellung  $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$  entspricht der Singulärwert-Zerlegung von  $\mathbf{A}$ ; die Reihenfolge der singulären Werte kann ja beliebig geändert werden durch Vertauschen der entsprechenden Spalten von  $\mathbf{U}$  und  $\mathbf{V}$ . Falls  $\mathbf{A}$  positiv semidefinit ist, d.h. falls alle Eigenwerte  $\lambda_i \geq 0$  sind, so sind die singulären Werte identisch zu den Eigenwerten.

## Literatur

- Eckart, C. and Young, G. (1936). The Approximation of one Matrix by Another of Lower Rank. *Psychometrika* 1, 211-218.
- Eckart, C. and Young, G. (1939). A Principal Axis Transformation for non-Hermitian Matrices. *Bull. Amer. Math. Soc.* 45, 118-121.
- Golub, G. H. (1965). Numerical Methods for Solving Linear Least Squares Problems. *Numerische Mathematik* 7, 206-216.
- Golub, G. H. and Kahan, W. (1965). Calculating the Singular Values and Pseudoinverse of a Matrix. *SIAM Journal of Numerical Analysis* 2, 205-224.
- Golub, Gene H. and Van Loan, Charles F. (1996). *Matrix Computations*. The Johns Hopkins University Press, Baltimore and London.
- Knüsel, L. (2009). The General Linear Model and the Generalized Singular Value Decomposition. Technical Report Number 48, 2009, Department of Statistics, [www.stat.uni-muenchen.de](http://www.stat.uni-muenchen.de).
- Knüsel, L. (2009). The General Linear Model and the Generalized Singular Value Decomposition; Some Examples. Technical Report Number 49, 2009, Department of Statistics, [www.stat.uni-muenchen.de](http://www.stat.uni-muenchen.de).
- Maple (2006). Waterloo Maple Inc. URL: [www.maplesoft.com](http://www.maplesoft.com).
- Matlab (2008). Technical Software for Engineering and Science. [www.mathworks.com](http://www.mathworks.com).
- Osborne, E. E. (1961). On Least Squares Solutions of Linear Equations. *Journal of Association of Computing Machinery* 8, 628-636.
- Penrose, R. (1955). A Generalized Inverse for Matrices. *Proc. Camb. Phil. Soc.* 51, 406-413.
- R Development Core Team (2006). *R: A language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. ISBN 3-900051-07-0, URL: [www.R-project.org](http://www.R-project.org).
- Rao, C.R. and Rao, M.B. (1998). *Matrix Algebra and Its Applications to Statistics and Econometrics*. World Scientific, Singapore.
- Rao, C.R., Toutenburg, H., Shalabh, Heumann, C. (2008). *Linear Models and Generalizations*. 570 pp. Springer, Berlin.
- Stewart, G.W. (1973). *Introduction to Matrix Computations*. Academic Press. 429 pp.
- Stoer, Josef (1999). *Numerische Mathematik 1*, 8. Auflage. Springer, Berlin.
- Stoer, Josef (2005). *Numerische Mathematik 1*, 9. Auflage. Springer, Berlin
- Stoer, Josef und Bulirsch, Roland (2005). *Numerische Mathematik 2*, 5. Auflage. Springer, Berlin
- Toutenburg, Helge (2003). *Lineare Modelle. Theorie und Anwendungen*, 2. Auflage. Physica-Verlag, Heidelberg.