Jacobi-ähnliche-Verfahren zur Berechnung der Singulärwertzerlegung (SVD) und der verallgemeinerten Singulärwertzerlegung (GSVD)

Udo Schendel Alke El-Fares−Ziegler

Forschungsprojekt *Parallele Numerik*
Fachbereich Mathematik und Informatik
Freie Universität Berlin 1999

Inhaltsverzeichnis

Vorwort

Für jede Matrix $A \in I\!K^{m \times n}$ läßt sich eine Singulärwertzerlegung (SVD[1]) finden: $A = U\Sigma V^H$. $U \in I\!K^{m \times m}$ und $V \in I\!K^{n \times n}$ sind unitäre Matrizen und $\Sigma \in I\!K^{m \times n}$ eine Diagonalmatrix mit den Singulärwerte von A als Diagonalelemente.
Die Singulärwerte sind die positiven Quadratwurzeln der Eigenwerte von $A^H A$.

Die Theorie der Matrixzerlegungen ist schon lange bekannt und bekam durch die Numerische Lineare Algebra eine immer größere Bedeutung, da sie eine sichere Plattform zum Lösen vollkommen unterschiedlicher Probleme liefert.
Eine der wichtigsten Zerlegungen ist die Singulärwertzerlegung, sie spielt in vielen Gebieten eine große Rolle:

- Da die SVD eine Zerlegung mit unitären Matrizen ist, ist sie zur Diskussion von n-dimensionalen Räumen gut geeignet.
- Es ist eine sehr stabile Zerlegung, kleine Veränderungen an der Matrix A erzeugen auch nur kleine Veränderungen an Σ und umgekehrt.
- Da Σ eine Diagonalmatrix ist, ist die Zerlegung auch für eine fast reguläre Matrix A leicht zu bestimmen. Bei einer regulären Matrix sorgt die Zerlegung für eine optimale Rang-Approximation.
- Dank der Pionierarbeiten von Gene Golub existieren effiziente und stabile Algorithmen.

Die verallgemeinerte Singulärwertzerlegung (GSVD[2]) $A^T Ax = \sigma^2 B^T Bx$ wird unter anderem bei der Lösung folgender Probleme verwendet:

- Gedämpfte kleinste Quadrate, d.h. Minimierung von $\|Ax - b\,|^2 + \lambda^2 \|Bx - d\|^2, \quad \lambda > 0$.
- Berechnung kleinster Quadrate unter Nebenbedingungen $\min\limits_{Bx=d} \|Ax - b\|$.
- Lösung gewichteter kleinster Quadrate. [vanL76]

Diese Arbeit entstand im Rahmen des Forschungsprojekts „Parallele Numerik“.

---

[1] singular value decomposition

[2] generalized singular value decomposition

Bezeichnungen

- $\mathbb{R}$ ist der Körper der reellen Zahlen.
- $\mathbb{C}$ ist der Körper der komplexen Zahlen.
- $\mathbb{K}$ bezeichnet einen beliebigen Körper aus $\{\mathbb{R}, \mathbb{C}\}$.
- $\mathbb{N} = \{1, 2, 3, ...\}$ sind die natürlichen Zahlen.
- $\mathbb{K}^n$ bezeichnet die Menge der Vektoren mit $n \in \mathbb{N}$ Koeffizienten aus $\mathbb{K}$.
- $\mathbb{K}^{m\times n}$ bezeichnet die Menge der Matrizen, deren $m \times n$ Koeffizienten in $m \in \mathbb{N}$ Zeilen und $n \in \mathbb{N}$ Spalten angeordnet sind.
- v_i ist je nach Zusammenhang entweder
  - die i-te Komponente des Vektors $v \in \mathbb{K}^n$ mit $v_i \in \mathbb{K}$ oder
  - ein Vektor als Element einer Folge von Vektoren $(v_i)_{i\in\mathbb{N}}$ mit $v_i \in \mathbb{K}^n$.
- $A = [a_{ij}]_{\substack{i=1,\dots,m\\ j=1,\dots,n}} \in \mathbb{K}^{m\times n}$ sei eine *Matrix* mit a_{ij} als Koeffizient in der i-ten Zeile und j-ten Spalte.
- a_{ij}^k oder $(A_k)_{ij}$ sei der Koeffizient in der i-ten Zeile und j-ten Spalte der k-ten Iteration A_k.
- $e_i \in \mathbb{K}^n$ ist der i-te *Einheitsvektor*.
- $I = I_n \in \mathbb{K}^{n\times n}$ ist die *Einheitsmatrix*.
- $A^{-1} \in \mathbb{K}^{n\times n}$ heißt *inverse Matrix* von $A \in \mathbb{K}^{n\times n}$, und es gilt $A^{-1}A = AA^{-1} = I$.
- $A^T =: B = [b_{ij}] \in \mathbb{K}^{n\times m}$ ist die *transponierte Matrix* von $A = [a_{ij}] \in \mathbb{K}^{m\times n}$ und es gilt $b_{ij} = a_{ji}$ für alle $i = 1, ..., n$ und $j = 1, ..., m$.
- $\bar{A} =: B = [b_{ij}] \in \mathbb{C}^{m\times n}$ ist die *konjugiert-komplexe Matrix* von $A = [a_{ij}] \in \mathbb{C}^{m\times n}$ mit $b_{ij} = \bar{a}_{ij}$ für alle $i = 1, ..., n$ und $j = 1, ..., m$.
- $A^H := \bar{A}^T \in \mathbb{C}^{n\times m}$ heißt die zu $A \in \mathbb{C}^{m\times n}$ *konjugiert-transponierte Matrix*.
- $A \in \mathbb{K}^{m\times n}$ heißt *nichtnegativ* $(A \geq 0)$, wenn $a_{ij} \geq 0$ für alle $i = 1, ..., m$ und $j = 1, ..., n$.
- $A \in \mathbb{K}^{m\times n}$ heißt *positiv* $(A > 0)$, wenn $a_{ij} > 0$ für alle $i = 1, ..., m$ und $j = 1, ..., n$.
- $A \in \mathbb{R}^{n\times n}$ heißt *symmetrisch*, falls $A^T = A$.
- $A \in \mathbb{C}^{n\times n}$ heißt *hermitesch*, falls $A^H = \bar{A}^T = A$.
- $A \in \mathbb{R}^{n\times n}$ heißt *orthogonal*, falls $A^T A = I$.
- $A \in \mathbb{R}^{n\times n}$ heißt *normal*, falls $A^T A = AA^T$.
- $A \in \mathbb{C}^{n\times n}$ heißt *unitär*, falls $A^H A = I$.

1 Grundlagen

Es werden zunächst einige klassische Begriffe und Eigenschaften der numerischen Linearen Algebra für die folgenden Betrachtungen zusammengefaßt.

1.1 Matrizen

Definition 1.1.1 (positiv definit) *Eine hermitesche Matrix* $A \in \mathbb{C}^{n\times n}$ *heißt positiv definit, wenn für alle* $x \in \mathbb{C}^n \backslash \{0\}$ *gilt:* $x^H Ax > 0$.

Lemma 1.1.2 *Sei* $A \in \mathbb{K}^{m\times n}$ *und* $Q \in \mathbb{K}^{m\times m}$ *eine unitäre Matrix, dann gilt für die Euklidische Norm* $\|QA\|_2 = \|A\|_2 := \sqrt{\lambda_1}$, *dabei ist* λ_1 *ist der betragsgrößte Eigenwert von* $A^H A$. [Maess85]

Definition 1.1.3 (Äquivalenz von Matrizen)

1. *Zwei Matrizen* $A, B \in \mathbb{K}^{m\times n}$ *sind äquivalent* ($A \sim B$), *wenn es zwei reguläre (orthogonale/unitäre) Matrizen* $U \in \mathbb{K}^{m\times m}$ *und* $V \in \mathbb{K}^{n\times n}$ *mit* $B = U^H AV$ *gibt.*
2. *Die Transformation* $A \mapsto B := U^H AV$ *heißt (orthogonale/unitäre) Äquivalenztransformation.*

Definition 1.1.4 (ähnliche/kongruente Matrizen)

1. *Zwei Matrizen* $A, B \in \mathbb{R}^{n\times n}$, *die mittels einer regulären (orthogonalen) Matrix* $X \in \mathbb{R}^{n\times n}$ *durch* $B = X^{-1}AX$ *verknüpft sind, heißen ähnlich.*
2. *Zwei Matrizen* $A, B \in \mathbb{C}^{n\times n}$, *die mittels einer regulären (unitären) Matrix* $X \in \mathbb{C}^{n\times n}$ *durch* $B = X^H AX$ *verknüpft sind, heißen kongruent.*
3. *Die Transformation* $A \mapsto B := X^{-1}AX$ *heißt (orthogonale) Ähnlichkeitstransformation und* $A \mapsto B := X^H AX$ *(unitäre) Kongruenztransformation.*

Definition 1.1.5 (zeilenproportional) *Zwei Matrizen* $A, B \in \mathbb{R}^{n\times n}$ *werden als zeilenproportional bezeiechnet, wenn eine Diagonalmatrix* D *existiert, so daß* $A = DB$ *gilt.*

Satz 1.1.6 (QR-Zerlegung) *Für jede Matrix* $A \in \mathbb{K}^{m\times n}$ *existiert eine unitäre Matrix* $Q \in \mathbb{K}^{m\times m}$ *und eine obere Dreiecksmatrix* $R \in \mathbb{K}^{m\times n}$, *so daß gilt* $Q^H A = R = \begin{bmatrix} R_0 \\ 0 \end{bmatrix}$ *mit* $R_0 \in \mathbb{K}^{n\times n}$ *und* $0 \in \mathbb{K}^{(m-n)\times(m-n)}$ *ist Nullmatrix.* [BuGe85]

Definition 1.1.7 (Givens-Rotations-Matrix) *Als Givens-Rotations-Matrix wird eine Matrix*

$$R(i,k) = \begin{bmatrix} 1 & & & & & & & & 0 \\ & \ddots & & & & & & & \\ & & 1 & & & & & & \\ & & & c_{ii} & \cdots & s_{ik} & & & \\ & & & & 1 & & & & \\ & & & & & \ddots & & & \\ & & & & & & 1 & & \\ & & & -s_{ki} & \cdots & c_{kk} & & & \\ & & & & & & & 1 & \\ & & & & & & & & \ddots \\ 0 & & & & & & & & 1 \end{bmatrix}$$

mit $c_{ii} = c_{kk} = c$, $s_{ik} = s_{ki} = s$ *und* $s^2 + c^2 = 1$ *bezeichnet.*

Bemerkung 1.1.8 $R(i,k)$ ist eine orthogonale Matrix, da $R(i,k)R^T(i,k) = I$ gilt.

∘

Bemerkung 1.1.9 (Eigenschaften der Givens-Rotation) Für $x \in \mathbb{R}^m$ und $y := R(i,k)x$ gilt

$$y_j = \begin{cases} cx_i + sx_k & \text{für } j = i, \\ -sx_i + cx_k & \text{für } j = k, \\ x_j & \text{für } j \neq i,k. \end{cases}$$

1. D.h. der Vektor x wird bei der Multiplikation mit $R(i,k)$ nur in der i-ten und k-ten Komponente verändert.
2. Die Komponenten werden um den Winkel θ gedreht, wobei $c = cos\theta$ und $s = sin\theta$.
3. Multipliziert man also eine Matrix $A \in \mathbb{R}^{n\times n}$ von links mit der Givens-Rotation $R(i,k)$, so bewirkt dies lediglich eine Veränderung der i-ten und der k-ten Zeile.

∘

Bemerkung 1.1.10 (Elimination mit Givens-Rotationen) Um die Komponenten $x_k \neq 0$ mit $R(i,k)$ in $y_k := R(i,k)x_k = 0$ zu überführen setzt man $s := \frac{x_k}{\sqrt{x_i^2+x_k^2}}$ und $c := \frac{x_i}{\sqrt{x_i^2+x_k^2}}$ und erfüllt damit $s^2 + c^2 = 1$, sofern $x_i \neq 0 \wedge x_k \neq 0$.

∘

Im Rahmen dieser Arbeit gilt (abweichend von der Definition für Pivot-Elemente z.B. bei der Gauß-Elimination [Schwarz93]) folgende Definition:

Definition 1.1.11 (Pivot) *Die Elemente a_{ik} einer Matrix, die zu Nullen transformiert werden, nennt man Pivots.*

Bemerkung 1.1.12

- Die Auswahl der Reihenfolge der Pivots wird als *Pivotstrategie* bezeichnet. Dabei unterscheidet man z.B.
- *spaltenzyklisch* bei folgender Eliminationsreihenfolge: (1,2),(1,3),(2,3),...,(1,n),(2,n),...,(n-1,n)
- *zeilenzyklisch* bei der Eliminationsreihenfolge: (1,2),(1,3),...,(1,n),(2,3),(2,4),...,(2,n),...,(n-1,n).

∘

Satz 1.1.13 (CS-Zerlegung) *Wenn*

$$Q = \begin{array}{cc} \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix} & \begin{array}{c} k \\ j \end{array} \\ \begin{array}{cc} k & \quad j \end{array} & \end{array}, \qquad k \geq j$$

eine orthogonale Matrix ist, exisitieren vier orthogonale Matrizen $U_1, V_1 \in \mathbb{R}^{k\times k}$ und $U_2, V_2 \in \mathbb{R}^{j\times j}$, so daß die CS-Zerlegung gilt

$$\begin{bmatrix} U_1 & 0 \\ 0 & U_2 \end{bmatrix}^T \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix} \begin{bmatrix} V_1 & 0 \\ 0 & V_2 \end{bmatrix} = \begin{array}{cc} \begin{bmatrix} I & 0 & 0 \\ 0 & C & S \\ 0 & -S & C \end{bmatrix} & \begin{array}{c} k-j \\ j \\ j \end{array} \\ \begin{array}{ccc} k-j & j & j \end{array} & \end{array}$$

mit	$C = \mathrm{diag}(c_1, \cdots, c_j)$	$c_i = \cos\varphi_i$
	$S = \mathrm{diag}(s_1, \cdots, s_j)$	$s_i = \sin\varphi_i$
und	$0 \leq \varphi_1 \leq \varphi_2 \leq \cdots \leq \varphi_j \leq \frac{\pi}{2}$.	[GoLo96]

Bemerkung 1.1.14 Die CS-Zerlegung diagonalisiert die einzelnen Blöcke einer orthogonalen Matrix gleichzeitig.

∘

Definition 1.1.15 (hierarchische Blockdreiecksmatrix) *$A \in I\!K^{n\times n}$ besitzt eine hierarchische Blockdreiecksgestalt, falls A (oder A^T) die folgende Struktur hat:*

$$A := \begin{bmatrix} A_{11}^r & A_{12}^r \\ A_{21}^r & A_{22}^r \end{bmatrix}$$

mit $A_{11}^r \in I\!K^{k_1\times k_1}, A_{12}^r \in I\!K^{k_1\times k_2}, A_{21}^r \in I\!K^{k_2\times k_1}, A_{22}^r \in I\!K^{k_2\times k_2}$. Für die Blöcke A_{ij}^r gilt:

(i) Wenn A_{22}^r eine obere Dreiecksmatrix ist, dann ist

A_{12}^r eine beliebige Matrix und
A_{21}^r eine Nullmatrix.

(ii) Wenn A_{22}^r eine untere Dreiecksmatrix ist, dann ist

A_{12}^r eine Nullmatrix und
A_{21}^r eine beliebige Matrix.

A_{11}^r (oder $(A_{11}^r)^T$) läßt sich darstellen als $A_{11}^r := \begin{bmatrix} A_{11}^{r-1} & A_{12}^{r-1} \\ A_{21}^{r-1} & A_{22}^{r-1} \end{bmatrix}$.
Die Blöcke von A_{11}^r besitzen wieder die gleichen Eigenschaften wie unter (i) und (ii) definiert.

Satz 1.1.16 *Werden auf eine obere Dreiecksmatrix zur Elimination der Elemente an den Stellen (p,q) und (q,p) spalten- bzw. zeilensyklisch orthogonale Transformationsmatrizen angewandt, so behält die Matrix immer eine hierarische Blockgestalt bei.* [Deich91]

1.2 Eigenwerte

Definition 1.2.1 (Eigenwerte, Eigenvektoren) *$\lambda \in \mathbb{K}$ heißt Eigenwert (EW) der Matrix $A \in \mathbb{K}^{n\times n}$, wenn ein Vektor $z = [z_1, \ldots, z_n]^T \in \mathbb{K}^n \backslash \{0\}$ existiert, für den gilt*

$$Az = \lambda z. \tag{1.2.2}$$

z heißt der zu λ gehörige (Rechts-)Eigenvektor (EV).

Bemerkung 1.2.3

1. Sei $\lambda \in \mathbb{K}$ der zu $z \in \mathbb{K}^n$ gehörende Eigenwert, dann ergibt sich das Gleichungssystem
$$Az - \lambda z = (A - \lambda I)z = 0.$$
2. Die Menge $\mathcal{L}(\lambda) := \{x | \, x \in \mathbb{K}^n : \; (A - \lambda I)x = 0\}$ bildet einen linearen Teilraum des $\mathbb{K}^n$ mit der Dimension $\dim(\mathcal{L}(\lambda)) = n - \operatorname{rang}(A - \lambda I)$.
3. λ ist genau dann Eigenwert von A, wenn $\mathcal{L}(\lambda) \neq \emptyset$, d.h. wenn $\dim(\mathcal{L}(\lambda)) > 0$ und damit $A - \lambda I$ singulär ist, also $\det(A - \lambda I) = 0$.

$\circ$

Definition 1.2.4 (Charakteristische Polynom) *Die Determinante $\det(A - \lambda I)$ definiert das charakteristische Polynom $\varphi_A(\lambda)$ der Matrix A:*

$$\varphi_A(t) := \det(A - tI) = (-1)^n(t^n + c_{n-1}t^{n-1} + \cdots + c_1 t + c_0), \qquad c_i \in \mathbb{K}, \; i = 0, ..., n-1.$$

Definition 1.2.5 (Spektrum, Spektralradius einer Matrix)

1. *Die Menge der Nullstellen $\lambda_i \in \mathbb{C}$ des charakteristischen Polynoms $\lambda(A) := \{\lambda_1, ..., \lambda_n\}$, d.h. die Eigenwerte der Matrix A, bezeichnet man als Spektrum von A.*
2. *Der Spektralradius von A ist definiert als $\rho(A) := \max\{|\lambda| : \lambda \in \lambda(A)\}$.*

Satz 1.2.6 *Sind A und B zwei ähnliche Matrizen, so haben sie das gleiche charakteristische Polynom, $\varphi_A(\lambda) \equiv \varphi_B(\lambda)$.* [HoJo85]

Satz 1.2.7 *Sei $A \in \mathbb{K}^{n\times n}$ symmetrisch/hermitesch. Dann gilt:*

- *Jeder Eigenwert von A ist reell, bei symmetrischen Matrizen sind auch die EV reell.*
- *Die Eigenvektoren zu verschiedenen Eigenwerten sind othogonal/unitär.*
- *Ist A zusätzlich positiv definit, dann sind alle Eigenwerte > 0.*

1.3 Allgemeines Eigenwertproblem

Das allgemeine Eigenwertproblem $Ax = \lambda Bx$ tritt u.a. in der Schwingungsmechanik und bei Differentialgleichungen auf.

Definition 1.3.1 (allgemeines Eigenwertproblem) *Seien $A, B \in \mathbb{K}^{n\times n}$. Dann heißt λ Eigenwert des Matrixpaares (A, B), wenn ein Vektor $z \in \mathbb{K}^n$, $z \neq 0$, existiert, so daß gilt*

$$Az = \lambda Bz \tag{1.3.2}$$

Der Vektor z heißt Eigenvektor zum Eigenwert λ und die Lösungen (λ, z) werden als Eigenpaar bezeichnet.

- Das allgemeine Eigenwertproblem (1.3.2) läßt sich umformen in die äquivalente Gleichung
$$(A - \lambda B)z = 0.$$
- Die Menge aller Matrizen $(A - \lambda B)$, mit $\lambda \in \mathbb{C}$, heißt *Matrixbüschel.*
- Als Bedingung für die Existenz einer nichttrivialen Lösung z muß für das verallgemeinerte charakteristische Polynom gelten: $\varphi_{A,B}(\lambda) := \det(A - \lambda B) = 0$.
$\lambda \in \mathbb{C}$ ist Eigenwerte des Matrixpaares (A, B).
- Im Unterschied zum speziellen Eigenwertproblem ist das charakteristieche Polynom im allgemeinen kein Polynom n-ten Grades, und daher existieren nicht immer n Eigenwerte.
- Ist B *singulär*, kann es endlich viele, gar keine oder unendliche viele Eigenwerte geben, wie folgende Beispiel zeigen:

  1. $$A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \qquad \varphi(\lambda) \equiv \lambda, \quad \lambda(A, B) = \{0\}$$
  2. $$A = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \qquad \varphi(\lambda) \equiv 3, \quad \lambda(A, B) = \emptyset$$
  3. $$A = \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \qquad \varphi(\lambda) \equiv 0, \quad \lambda(A, B) = \mathbb{C}$$

- Matrixbüschel, deren charakteristische Polynome konstant Null sind, werden als *singuläre Büschel* bezeichnet. Wie in Beispiel 3 sind Büschel oft singulär, weil sie einen gemeinsamen Nullraum haben. Allerdings zeigt Beispiel 1, daß sowohl A als auch B singulär sein können, ohne daß das Büschel $(A - \lambda B)$ singulär ist.

Satz 1.3.3 *1. Das allgemeine Eigenwertproblem $Az = \lambda Bz$ hat genau dann n Eigenwerte, wenn B nichtsingulär ist.*

2. Sei $\lambda \neq 0$ ein Eigenwert von $Az = \lambda Bz$, dann ist $\frac{1}{\lambda}$ Eigenwert des reziproken Eigenwertproblems $Bz = \mu Az$. [GoLo96]

Bemerkung 1.3.4 Für eine nichtsinguläre Matrix B läßt sich allgemeine Eigenwertproblem (1.3.2) auf das spezielle Eigenwertproblem $B^{-1}Az = \lambda z$ reduzieren. Diese Reduktion kann jedoch die Symmetrie zerstören und ist wegen der Inversenbildung oft numerisch instabil.

○

Um das allgemeine Eigenwertproblem numerisch stabil zu lösen, werden beide Matrizen gleichzeitig auf eine einfachere Gestalt transformiert.

Satz 1.3.5 *Seien (A, B) und $(\tilde{A}, \tilde{B})$ äquivalent ($\tilde{A} = SAT$ und $\tilde{B} = SBT$, $S, T \in \mathbb{K}^{n\times n}$ regulär, vgl. Definition 1.1.3) und (λ, z) ist ein Eigenpaar von (A, B), dann ist $(\lambda, T^{-1}z)$ Eigenpaar von $(\tilde{A}, \tilde{B})$.*

Beweis: $Az = \lambda Bz \Longleftrightarrow \tilde{A}y = \lambda \tilde{B}y$ mit $z = Ty$.

□

Mit unitären Äquivalenztransformationen lassen sich beliebige Matrizenpaare gleichzeitig auf Dreiecksgestalt reduzieren, an der die Eigenwerte ablesbar sind:

Satz 1.3.6 (allgemeine Schur-Zerlegung)

1. *Zu $A, B \in \mathbb{K}^{n\times n}$ existieren unitäre Matrizen $U, V \in \mathbb{K}^{n\times n}$, so daß*
$$UAV = \tilde{A} \quad \text{und} \quad UBV = \tilde{B}$$
obere Dreiecksmatrizen sind.

2. *Für die Eigenwerte von (A, B) gilt: sind für ein i, $1 \leq i \leq n$, die Diagonalelemente $\tilde{a}_{ii} = \tilde{b}_{ii} = 0$, dann ist das Spektrum $\lambda(A, B) = \mathbb{C}$. Anderenfalls gilt*
$$\lambda(A, B) = \{\tfrac{\tilde{a}_{ii}}{\tilde{b}_{ii}} \mid \tilde{b}_{ii} \neq 0\}.$$
[GoLo96]

Bemerkung 1.3.7

1. In der praktischen Anwendung treten meist die Eigenwertprobleme für hermitesche Matrizen A und B auf und B ist zusätzlich positiv definit. In diesem Fall hat das Matrixpaar (A, B) reelle Eigenwerte. [BuGe85]

2. Ein Matrixpaar mit hermiteschen Matrizen hat nicht unbedingt die Eigenschaften einer hermiteschen Matrix:
$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}, \quad a, b \in \mathbb{R} \Longrightarrow \varphi_{A,B}(\lambda) = \lambda^2 ab - 1 \Rightarrow \lambda^2 = \frac{1}{ab}.$$
Das Matrixpaar (A, B) mit den symmetrischen Matrizen A und B hat komplexe Eigenwerte, wenn a und b verschiedne Vorzeichen haben.

○

Bei einer Kongruenztransformation (vgl. Definition 1.1.4) bleiben Symmetrie und Definitheit erhalten:

Satz 1.3.8 *Sei $A \in \mathbb{K}^{n\times n}$ hermitesch und $B \in \mathbb{K}^{n\times n}$ hermitesch und positiv definit. Dann existiert eine nichtsinguläre Matrix $F \in \mathbb{K}^{n\times n}$, so daß $F^H AF$ und $F^H BF$ reelle Diagonalmatrizen sind.* [BuGe85]

Bemerkung 1.3.9 Mit der entsprechenden Wahl von F läßt sich also das allgemeine Eigenwertproblem $Az = \lambda Bz$, $A \in \mathbb{K}^{n\times n}$ hermitesch und $B \in \mathbb{K}^{n\times n}$ hermitesch und positiv definit, auf das Standard-Problem $\hat{A}z = \lambda z$ zurückführen. [BuGe85]

○

2 Singulärwertzerlegungen

2.1 Spezielle Singulärwertzerlegung (SVD)

Die Singulärwertzerlegung spielt eine entscheidende Rolle u.a. bei der Behandlung von Ausgleichsproblemen mit und ohne Bedingungen, Pseudoinversen, Integralgleichungen und in der Approximationstheorie. [Björck96]

Definition 2.1.1 (Singulärwerte) *Die Singulärwerte einer Matrix $A \in I\!K^{m\times n}$ sind die positiven Wurzeln der Eigenwerte der Matrix $A^H A \in I\!K^{n\times n}$.*

Auch nicht-quadratische Matrizen können mittels orthogonaler bzw. unitärer Transformationen auf Diagonalgestalt umgeformt werden.

Satz 2.1.2 (Singulärwertzerlegung) *Für jede Matrix $A \in I\!K^{m\times n}$ gibt es zwei unitäre Matrizen*

$$U = [u_1, ..., u_m] \in I\!K^{m\times m} \quad \text{und} \quad V = [v_1, ..., v_n] \in I\!K^{n\times n},$$

mit denen eine Äquivalenztransformation von A auf Diagonalform

$$U^H AV = \Sigma := diag(\sigma_1, \sigma_2, ..., \sigma_k) \in I\!K^{m\times n} \qquad k := \min\{m, n\} \tag{2.1.3}$$

möglich ist, so daß die Diagonalwerte von Σ folgende Ungleichung erfüllen:

$$\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_r > \sigma_{r+1} = \ldots = \sigma_k = 0 \qquad \text{mit } r = \text{rang(A)}.$$

Die Diagonalelemente σ_i sind die Singulärwerte der Matrix A.
Falls $A \in I\!K^{n\times n}$ eine reguläre Matrix ist, dann gilt speziell $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n > 0$.

Beweis: $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_n$ seien die Eigenwerte von $A^H A$ und
$\sigma_1, ..., \sigma_r$ die positiven Singulärwerte von A mit $\sigma_i := +\sqrt{\lambda_i}$ für alle $i = 1, 2, ...n$.
$V = [v_1...v_n] \in I\!K^{n\times n}$ sei eine unitäre Matrix, deren Spalten $v_1, ..., v_n$ ein Orthonormalsystem von Eigenvektoren von $A^H A$ zu den Eigenvektoren $\lambda_1, \lambda_2, ..., \lambda_n$ bilden.
Es gelte $V = [V_1 V_2]$, mit
$V_1 = [v_1...v_r]$ und
$V_2 = [v_{r+1}...v_n]$, wobei $v_{r+1}, ..., v_n$ die zum Eigenwert 0 gehörenden Eigenvektoren sind.
Sei $\Sigma_r := \text{diag}(\sigma_1, ..., \sigma_r)$ und $\Sigma := \begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix} \in I\!R^{m\times n}$.
Dann ist $V^H A^H AV = [V_1 V_2]^H A^H A[V_1 V_2] = \Sigma^H \Sigma = \begin{bmatrix} \Sigma_r^2 & 0 \\ 0 & 0 \end{bmatrix}$,
es gilt also $V_2^H A^H AV_2 = 0$, d.h. $AV_2 = 0$ und $V_1^H A^H AV_1 = \Sigma_r^2$, d.h. $\Sigma_r^{-1} V_1^H A^H AV_1 \Sigma_r = I_r$.
$U_1 := AV_1 \Sigma_r^{-1} \Rightarrow U_1^H U_1 = I_r$.
Die Spalten von U_1 bilden also ein Orthonormalsystem. Diese werden mit $m-r$ weiteren Vektoren $[u_{r+1}...u_m] =: U_2 \in I\!K^{m-r}$ zu einer Orthonormalbasis ergänzt.
$U := [U_1 U_2] \in I\!K^m$ ist also unitär, insbesondere ist $0 = U_2^H U_1 \Sigma_r = U_2^H AV_1$.

$$\Rightarrow U^H AV = \begin{bmatrix} U_1^H \\ U_2^H \end{bmatrix} A[V_1 V_2] = \begin{bmatrix} U_1^H AV_1 & U_1^H AV_2 \\ U_2^H AV_1 & U_2^H AV_2 \end{bmatrix} = \begin{bmatrix} \Sigma_r^{-1} V_1^H A^H AV_1 & 0 \\ U_2^H U_1 \Sigma_r & 0 \end{bmatrix} = \begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix}.$$

$\square$

Definition 2.1.4 (Singulärvektoren)

1. *Die Matrix U besteht aus den orthonormierten Eigenvektoren $u_1, ..., u_m$ von AA^H, die man Linkssingulärvektoren von A nennt, und*

2. *die Matrix V aus den orthonormierten Eigenvektoren $v_1, ..., v_n$ von $A^H A$, den Rechtssingulärvektoren von A.*

Bemerkung 2.1.5 • Es gelten die Relationen $Av_i = \sigma_i u_i$ und $A^H u_i = \sigma_i v_i$ mit $i = 1, ..., r := \min\{m, n\}$.

- $A^T A$ und $\Sigma^T \Sigma = \mathrm{diag}(\sigma_1^2, \sigma_2^2, ..., \sigma_r^2, 0, ..., 0)$ sind orthogonalähnlich (Definition 1.1.4), denn es gilt $\Sigma^T \Sigma = (U^T AV)^T (U^T AV) = V^T A^T U U^T AV = V^{-1} A^T AV$.

- Da ähnliche Matrizen das gleiche Spektrum besitzen (vgl. Satz 1.2.6) gilt $\sigma_i = +\sqrt{\lambda_i}$ mit $\lambda_i \in \lambda(A^T A),\ i = 1, 2, ..., n$.

- Die Matrizen $A \in I\!R^{m \times n}$ und $A^T \in I\!R^{n \times m}$ haben dieselben Singulärwerte.

- Die Singulärvektoren hingegen sind nur dann eindeutig, wenn σ_i^2 ein einfacher Eigenwert von $A^T A$ bzw. AA^T ist. Für mehrfache Singulärwerte können die zugeordneten Singulärvektoren als irgendeine orthonormale Basis des entsprechenden Eigenraums gewählt werden.

- Die Euklidische Norm und Frobenius-Norm $\|A\|_F := \sqrt{\sum_{j=1}^{n} \sum_{i=1}^{m} |a_{ij}|^2} = \sqrt{\mathrm{spur}(A^H A)}$ einer Matrix $A \in I\!R^{n \times n}$ stehen in engem Zusammenhang mit den Singulärwerten von A:

$$\begin{aligned} \|A\|_2 &= \sigma_1, \\ \|A\|_F &= \sqrt{\sigma_1^2 + ... + \sigma_k^2}. \end{aligned}$$

- Auch die Determinante $\det(A)$ einer quadratischen Matrix $A \in I\!R^{n \times n}$ hängt mit den Singulärwerten zusammen:

$$|\det(A)| = |\det(U) \cdot \det(S) \cdot \det(V^T)| = \prod_{i=1}^{k} \sigma_i,$$

da die Determinate einer orthogonalen Matrix entweder +1 oder -1 ist.

∘

Lemma 2.1.6 *Sei $A \in I\!R^{m \times n}$, $m \geq n$ und $P = [p_1, ..., p_m] \in I\!R^{m \times m}$, $Q = [q_1, ..., q_n] \in I\!R^{n \times n}$ orthogonale Matrizen und*

$$P^T AQ = \begin{bmatrix} B \\ 0 \end{bmatrix} \begin{matrix} \}n \\ \}m-n \end{matrix}.$$

Dann gilt

1. *$A \in I\!R^{m \times n}$ und $B \in I\!R^{n \times n}$ haben dieselben Singulärwerte.*

2. *Ist $B = U_B \mathrm{diag}(\sigma_1, ..., \sigma_n) V_B^T$ mit den orthogonalen Matrizen $U_B \in I\!R^{n \times n}$, $V_B \in I\!R^{n \times n}$ eine Singulärwertzerlegung von B, dann ist durch*

$$A = U \mathrm{diag}(\sigma_1, ..., \sigma_n) V^T, \qquad \text{mit } U := (p_1, ..., p_n) U_B,\ V := QV_B,$$

eine Singulärwertzerlegung von A gegeben. [Werner92]

2.2 Verallgemeinerte Singulärwertzerlegung (GSVD)

Definition 2.2.1 (verallgemeinerte Singulärwerte) *Seien $A \in \mathbb{K}^{m\times n}$ und $B \in \mathbb{K}^{p\times n}$ mit $m, n, p \in \mathbb{N}$ beliebig. Das nichtnegative Zahlenpaar (α, β) heißt verallgemeinertes singuläres Wertepaar des Matrixpaares (A, B), falls gilt*

$$(\alpha,\beta) = (\lambda^{\frac{1}{2}}, \mu^{\frac{1}{2}}) \quad mit \quad (\lambda,\mu) \in \lambda(A^TA, B^TB) \qquad \lambda,\mu \geq 0,$$

wobei $\lambda(A^TA, B^TB)$ die Menge der verallgemeinerten Eigenwerte des Matrixpaares (A^TA, B^TB) ist.
Die Menge aller verallgemeinertan Singulärwerte wird mit $\sigma(A,B) \equiv \{\frac{\alpha_1}{\beta_1}, \cdots, \frac{\alpha_r}{\beta_r}\}$ bezeichnet.

Bemerkung 2.2.2
1. $(\lambda,\mu) \in \lambda(A^TA, B^TB) \overset{(\lambda,\mu)\neq(0,0)}{\Longleftrightarrow} \mu A^TAx = \lambda B^TBx, \Rightarrow$ $A^TAx = \frac{\lambda}{\mu}B^TBx$. [Deich91]
2. Es gilt: $\sigma(A,B) = (\lambda,\mu) \neq (0,0) | \det(\beta^2 A^TA - \alpha^2 B^TB) = 0, \quad \alpha,\beta \geq 0.$
3. $\sigma \in \sigma(A,B) \Rightarrow \sigma^2 \in \lambda(A^TA, B^TB).$
4. Ist $B \in \mathbb{K}^{n\times n}$ regulär, dann gilt: $U^TAB^{-1}V = \Sigma := \mathrm{diag}(\frac{\alpha_1}{\beta_1}, ..., \frac{\alpha_p}{\beta_p}), \;\; p = \mathrm{rang}(B).$

∘

Satz 2.2.3 (verallgemeinerte Singulärwertzerlegung) *Für die Matrizen $A \in \mathbb{R}^{m\times n}$, $m \geq n$, und $B \in \mathbb{R}^{p\times n}$ existieren die orthogonalen Matrizen $U \in \mathbb{R}^{m\times m}$ und $V \in \mathbb{R}^{p\times p}$ und eine reguläre Matrix $X \in \mathbb{R}^{n\times n}$, so daß gilt:*

$$\begin{array}{llll} & U^TAX = \Sigma_A = \mathrm{diag}(\alpha_1, \cdots, \alpha_n) & \alpha_i \geq 0 & \\ und & V^TBX = \Sigma_B = \mathrm{diag}(\beta_1, \cdots, \beta_q) & \beta_i \geq 0, & q = \min\{p, n\} \\ mit & \beta_1 \geq \cdots \geq \beta_r > \beta_{r+1} = \cdots = \beta_q = 0 & & r = \mathrm{rang}(B). \end{array}$$

Beweis: O.B.d.A. sei $m = n = p$ und $\mathcal{N}(A) \cap \mathcal{N}(B) = 0$.

$$\begin{bmatrix} A \\ B \end{bmatrix} = \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix} \begin{bmatrix} R \\ 0 \end{bmatrix} \tag{2.2.4}$$

ist eine QR-Zerlgung mit einer orthogonaler Matrix $Q = \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix}$ und einer oberen Dreiecksmatrix $R \in \mathbb{R}^{n\times n}$. Da $\mathcal{N}(A) \cap \mathcal{N}(B) = 0$ gilt ist R nichtsingulär.

$$\begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix} = \begin{bmatrix} U & 0 \\ 0 & V \end{bmatrix} \begin{bmatrix} C & S \\ -S & C \end{bmatrix} \begin{bmatrix} W & 0 \\ 0 & Y \end{bmatrix}^T \tag{2.2.5}$$

ist eine CS-Zerlegung (vgl. Satz 1.1.13) von Q mit den orthogonalen Matrizen U, V, W und Y, $C = \mathrm{diag}(c_i)$, $c_i \geq 0$ und $S = \mathrm{diag}(s_i)$, $s_i \geq 0$ und $C^2 + S^2 = I_n$.
Durch Vergleichen der Blöcke in (2.2.4) und anwenden der Gleichung (2.2.5) ergibt sich:

$$\begin{aligned} A &= Q_{11}R = UC(W^TR) \\ B &= Q_{21}R = -VS(W^TR) \end{aligned}$$

Mit $\Sigma_A = C$, $\Sigma_B = -S$ und $X = (W^TR)^{-1}$ gilt Satz 2.2.3.

□

Bemerkung 2.2.6 Der Satz ist eine Verallgemeinerung der speziellen Singulärwertzerlegung (Satz 2.1.2). Ist $B = I_n$, dann gilt $\sigma(A,B) = \sigma(A)$.

∘

3 Numerische Verfahren der Singulärwertzerlegungen

3.1 Das Verfahren von Hari und Veselić zur Berechnung der SVD

Das Verfahren von Hari und Veselić berechnet die Singulärwerte einer bereits auf obere Dreiecksgestalt transformierte Matrix $A \in I\!R^{n\times n}$, z.B. mit Housholder-Matrizen. Dies bedeutet keine Einschränkung, denn wenn $A = QR$ mit orthogonaler Matrix Q und oberer Dreiecksmatrix R eine QR-Zerlegung der Matrix A ist, so ist das Problem $A^TAx = \lambda^2 x$ äquivalent zu dem Problem $R^TRx = \lambda^2 x$.
Zur Berechnung der Singulärwertzerlegung $U^TAV = \Sigma$ werden die Matrizen U und V als Produkt von Rotationsmatrizen U_k und V_k gebildet. Daraus ergibt sich, in Anlehnung an das Jacobi-Verfahren, das folgende Iterationsverfahren zur Berechnung der Singulärwerte:

$$\begin{aligned} &A_0 := A \\ &A_k := U_k^T A_{k-1} V_k \qquad \text{für } k \geq 1. \end{aligned} \tag{3.1.1}$$

Eigenschaften der Transformationsmatrizen

- U_k und V_k sind die k-ten Transformationen, die sich nur an den Stellen (p,p), (p,q), (q,p) und (q,q) von der Einheitsmatrix unterscheiden, d.h. $U_k = U_k(p,q)$ und $V_k = V_k(p,q)$ sind Givens-Rotations-Matrizen (siehe Definition 1.1.7).
  φ_k und ψ_k seien die zur k-ten Transformation gehörenden Winkel. Die 2×2-Pivotrestriktionen der Matrizen U_k, V_k und A_k haben also die Form:

$$\tilde{U}_k := \begin{bmatrix} u_{pp} & u_{pq} \\ u_{qp} & u_{qq} \end{bmatrix} = \begin{bmatrix} \cos\varphi_k & \sin\varphi_k \\ -\sin\varphi_k & \cos\varphi_k \end{bmatrix} \tag{3.1.2}$$

$$\tilde{V}_k := \begin{bmatrix} v_{pp} & v_{pq} \\ v_{qp} & v_{qq} \end{bmatrix} = \begin{bmatrix} \cos\psi_k & \sin\psi_k \\ -\sin\psi_k & \cos\psi_k \end{bmatrix} \tag{3.1.3}$$

$$\tilde{A}_k := \begin{bmatrix} a^k_{pp} & a^k_{pq} \\ a^k_{qp} & a^k_{qq} \end{bmatrix} \tag{3.1.4}$$

- Die Pivotpaare (p,q) sind von k abhängig, d.h. als Funktionen von k aufzufassen:
  $p := p(k)$ und $q := q(k)$ für alle $(p,q) \in \{1,...,n\} \times \{1,...,n\},\ p < q$.

- Die Parameter φ_k und ψ_k der Matrizen U_k und V_k werden so gewählt, daß

$$a^k_{pq} = a^k_{qp} = 0$$

  gilt. Für die Restriktion folgt daraus:

$$\begin{bmatrix} \cos\varphi_k & \sin\varphi_k \\ -\sin\varphi_k & \cos\varphi_k \end{bmatrix} \begin{bmatrix} a^{k-1}_{pp} & a^{k-1}_{pq} \\ a^{k-1}_{qp} & a^{k-1}_{qq} \end{bmatrix} \begin{bmatrix} \cos\psi_k & \sin\psi_k \\ -\sin\psi_k & \cos\psi_k \end{bmatrix} = \begin{bmatrix} a^k_{pp} & 0 \\ 0 & a^k_{qq} \end{bmatrix} \tag{3.1.5}$$

  Als Bestimmungsgleichung für die Parameter φ_k und ψ_k ergibt sich dann:

$$
\begin{aligned}
(\cos\varphi_k a_{pp}^{k-1} - \sin\varphi_k a_{qp}^{k-1})\sin\psi_k \quad &+ \quad (\cos\varphi_k a_{pq}^{k-1} - \sin\varphi_k a_{qq}^{k-1})\cos\psi_k = 0 \\
(\sin\varphi_k a_{pp}^{k-1} - \cos\varphi_k a_{qp}^{k-1})\cos\psi_k \quad &+ \quad (\sin\varphi_k a_{pq}^{k-1} - \cos\varphi_k a_{qq}^{k-1})\sin\psi_k = 0
\end{aligned}
\tag{3.1.6}
$$

und die Parameter φ_k und ψ_k berechnen sich wie folgt:

$$
\begin{aligned}
\tan 2\varphi_k &= \frac{2(a_{qp}^{k-1}a_{pp}^{k-1}+a_{qq}^{k-1}a_{pq}^{k-1})}{(a_{qq}^{k-1})^2+(a_{qp}^{k-1})^2-(a_{pp}^{k-1})^2-(a_{pq}^{k-1})^2} \\
\tan\psi_k &= \frac{-a_{pq}^{k-1}+a_{qq}^{k-1}\tan\varphi_k}{a_{pp}^{k-1}-a_{qp}^{k-1}\tan\varphi_k} = \frac{a_{qp}^{k-1}+a_{pp}^{k-1}\tan\varphi_k}{a_{qq}^{k-1}+a_{pq}^{k-1}\tan\varphi_k}
\end{aligned}
\tag{3.1.7}
$$

und äquivalent dazu

$$
\begin{aligned}
\tan 2\psi_k &= \frac{2(a_{pp}^{k-1}a_{pq}^{k-1}+a_{qp}^{k-1}a_{qq}^{k-1})}{(a_{qq}^{k-1})^2+(a_{pq}^{k-1})^2-(a_{pp}^{k-1})^2-(a_{qp}^{k-1})^2} \\
\tan\varphi_k &= \frac{-a_{qp}^{k-1}+a_{qq}^{k-1}\tan\psi_k}{a_{pp}^{k-1}-a_{pq}^{k-1}\tan\psi_k} = \frac{a_{pq}^{k-1}+a_{pp}^{k-1}\tan\psi_k}{a_{qq}^{k-1}+a_{qp}^{k-1}\tan\psi_k}
\end{aligned}
\tag{3.1.8}
$$

- Der Algorithmus startet mit einer oberen Dreiecksmatrix A. Nach dem ersten Zyklus entsteht im Laufe der Transformationen eine untere Dreiecksmatrix, nach dem zweiten wieder eine obere Dreiecksmatrix usw., d.h. also im ungeraden Zyklus entsteht eine obere, im geraden Zyklus eine untere Dreiecksmatrix. Mit jeder 2×2-SVD wird jeweils nur ein Element eliminiert, die im unteren Schema eingekreisten Elemente werden im nächsten Schritt eliminiert. Eliminationsschritte des geraden Zyklus:

$$
\begin{array}{cccc}
\times & \otimes & \times & \times \\
 & \times & \times & \times \\
 & & \times & \times \\
 & & & \times
\end{array}
\longrightarrow
\begin{array}{cccc}
\times & & \otimes & \times \\
 & \times & \times & \times \\
 & & \times & \times \\
 & & & \times
\end{array}
\longrightarrow
\begin{array}{cccc}
\times & & & \otimes \\
\times & \times & \times & \times \\
 & & \times & \times \\
 & & & \times
\end{array}
\longrightarrow
\begin{array}{cccc}
\times & & & \\
\times & \times & \otimes & \times \\
\times & & \times & \times \\
 & & & \times
\end{array}
$$

$$
\longrightarrow
\begin{array}{cccc}
\times & & & \\
\times & \times & & \otimes \\
\times & & \times & \times \\
 & & & \times
\end{array}
\longrightarrow
\begin{array}{cccc}
\times & & & \\
\times & \times & & \\
\times & \times & \times & \otimes \\
\times & & & \times
\end{array}
\longrightarrow
\begin{array}{cccc}
\times & & & \\
\otimes & \times & & \\
\times & \times & \times & \\
\times & \times & & \times
\end{array}
$$

- Zur Berechnung der Winkel wird für gerade Zyklen die Formel (3.1.7) und für ungerade Zyklen (3.1.8) verwendet.
  Winkel für die geraden Zyklen:

$$
\begin{aligned}
\tan 2\varphi_k &= \frac{2(a_{pp}^{k-1}a_{qp}^{k-1})}{(a_{qq}^{k-1})^2-(a_{pp}^{k-1})^2+(a_{qp}^{k-1})^2} \\
\tan\psi_k &= \frac{a_{qq}^{k-1}\tan\varphi_k}{a_{pp}^{k-1}-a_{qp}^{k-1}\tan\varphi_k} = \frac{a_{qp}^{k-1}+a_{pp}^{k-1}\tan\varphi_k}{a_{qq}^{k-1}}
\end{aligned}
\tag{3.1.9}
$$

  Winkel für die ungeraden Zyklen:

$$
\begin{aligned}
\tan 2\psi_k &= \frac{2(a_{pp}^{k-1}a_{pq}^{k-1})}{(a_{qq}^{k-1})^2+(a_{pq}^{k-1})^2-(a_{pp}^{k-1})^2} \\
\tan\varphi_k &= \frac{a_{qq}^{k-1}\tan\psi_k}{a_{pp}^{k-1}-a_{pq}^{k-1}\tan\psi_k} = \frac{a_{pq}^{k-1}+a_{pp}^{k-1}\tan\psi_k}{a_{qq}^{k-1}}
\end{aligned}
\tag{3.1.10}
$$

- Diese Wahl der Winkel sichert die allgemeine Konvergenz des Verfahrens.

3.2 Berechnung der GSVD

Um die verallgemeinerten Singulärwerte von zwei oberen Dreiecksmatrizen $A = [a_{ij}] \in \mathbb{R}^{n \times n}$ und $B = [b_{ij}] \in \mathbb{R}^{n \times n}$, B nichtsingulär, zu berechnen, läßt sich die verallgemeinerte Singulärwertzerlegung (Bemerkung 2.2.2 1.)

$$\mu A^T Ax = \lambda B^T Bx$$

als Singulärwertzerlegung des Matrixproduktes AB^{-1} auffassen, d.h. es müssen die orthogonalen Matrizen U und V berechnet werden mit

$$U^T(AB^{-1})V = D \tag{3.2.1}$$

D ist eine Diagonalmatrix. Da bei schlecht konditionierter Matrix B die numerische Berechnung des Produktes AB^{-1} zu großen Fehlern führen kann wird, um die Inversenbildung zu vermeiden, folgende Darstellung betrachtet:

$$U^T A = DV^T B. \tag{3.2.2}$$

Die Matrizen $U^T A$ und $V^T B$ sind zeilenproportional (vgl. Definition 1.1.5).
Mit der der nichtsingulären Matrix X ergibt sich das folgende Iterationsverfahren (vgl. Satz 2.2.3):

$$\begin{array}{lll} A_0 := A & B_0 := B & \text{Startmatrizen} \\ A_k := U_k^T A_{k-1} X_k & B_k := V_k^T B_{k-1} X_k & \text{für } k = 1, 2, ... \end{array} \tag{3.2.3}$$

Die Transformationematrizen U und V werden gebildet als Produkt der Givens-Rotationsmatrizen $U_k = U_k(p,q)$ und $V_k = V_k(p,q)$ (siehe Definition 1.1.7).

Bemerkung 3.2.4 (Konvergenz)

1. Das Iterationsverfahren (3.2.3) erzeugt zeilenproportionale Matritzen A_k und B_k.
2. Die Matrizen A_k, B_k konvergieren quadratisch gegen Zeilenproportionalität.
3. Die Zeilenproportionalitätsfaktoren stellen die Lösung des verallgemeinerten Singulärwertproblems $\mu A^T Ax = \lambda B^T Bx$ dar.[Deich91]

○

Eigenschaften der Transformationsmatrizen

- Um die Elemente a_{pq}, a_{qp}, b_{pq} und b_{qp} der Matrizen A und B zu eliminieren unterscheiden sich die Transformationsmatrizen U_k, V_k und X_k nur an den Stellen (p,p), (p,q), (q,p) und (q,q) von der Einheitsmatrix. Die 2×2-Restriktionen von U_k, V_k und X_k haben also die Form:

$$\tilde{U}_k := \begin{bmatrix} u_{pp} & u_{pq} \\ u_{qp} & u_{qq} \end{bmatrix} = \begin{bmatrix} \cos\varphi_k & \sin\varphi_k \\ -\sin\varphi_k & \cos\varphi_k \end{bmatrix} \tag{3.2.5}$$

$$\tilde{V}_k := \begin{bmatrix} v_{pp} & v_{pq} \\ v_{qp} & v_{qq} \end{bmatrix} = \begin{bmatrix} \cos\psi_k & \sin\psi_k \\ -\sin\psi_k & \cos\psi_k \end{bmatrix} \tag{3.2.6}$$

$$\tilde{X}_k := \begin{bmatrix} 1 & \alpha_k \\ -\beta_k & 1 \end{bmatrix} \tag{3.2.7}$$

- Die Parameter $\varphi_k, \psi_k, \alpha_k$ und β_k werden so bestimmt, daß sie die folgenden Gleichungen erfüllen:

$$\tilde{U}_k^T \tilde{A}_{k-1} \tilde{X}_k = \begin{bmatrix} a_{pp}^k & 0 \\ 0 & a_{qq}^k \end{bmatrix} \quad \text{und} \quad \tilde{V}_k^T \tilde{B}_{k-1} \tilde{X}_k = \begin{bmatrix} b_{pp}^k & 0 \\ 0 & b_{qq}^k \end{bmatrix}. \tag{3.2.8}$$

- Die Indizes (p,q) sind als Funktionen von k aufzufassen: $p := p(k)$ und $q := q(k)$.
- Es wird eine Normierung von B_k gefordert:

$$(B_k^T B_k)_{ii} = 1, \tag{3.2.9}$$

$(B_k^T B_k)_{ii}$ ist das i-te Diagonalelement der Matrix $B_k^T B_k$.

- Um die Forderung (3.2.9) von Beginn an sicherzustellen, wird das Verfahren mit der Matrix $\hat{B}_0 := B_0 D_0$ gestartet. D_0 ist wie folgt definiert:

$$D_0 := \operatorname{diag}\left(\frac{1}{\sqrt{(B_0^T B_0)_{11}}}, \frac{1}{\sqrt{(B_0^T B_0)_{22}}}, \cdots, \frac{1}{\sqrt{(B_0^T B_0)_{nn}}}\right).$$

Analog wird $\hat{A}_0 := A_0 D_0$ definiert.

Bemerkung 3.2.10 1. Die Matrixpaare $(A_0^T A_0, B_0^T B_0)$ und $(\hat{A}_0^T \hat{A}_0, \hat{B}_0^T \hat{B}_0)$ besitzen dieselben verallgemeinerten Singulärwerte.

2. Die Normierung wird eingeführt, um zu garantieren, daß die Normen von $\hat{A}_0$ und $\hat{B}_0$ beschränkt sind: Es gilt

$$\|\hat{B}_0\|_F = \sqrt{\operatorname{spur}(\hat{B}_0^T \hat{B}_0)} = \sqrt{\sum_{i=1}^{n} (\hat{B}_0^T \hat{B}_0)_{ii}} = \sqrt{n}, \tag{3.2.11}$$

und für die Norm von $\hat{A}_0$ folgt dann aus

$$(\hat{A}_0^T \hat{A}_0)_{ii} = e_i^T \hat{A}_0^T \hat{A}_0 e_i = \frac{e_i^T \hat{A}_0^T \hat{A}_0 e_i}{e_i^T \hat{B}_0^T \hat{B}_0 e_i} e_i^T \hat{B}_0^T \hat{B}_0 e_i \leq \sigma_1^2 \tag{3.2.12}$$

die folgende Ungleichung:

$$\|\hat{A}_0\|_F = \sqrt{\operatorname{spur}(\hat{A}_0^T \hat{A}_0)} = \sqrt{\sum_{i=1}^{n} (\hat{A}_0^T \hat{A}_0)_{ii}} \leq \sqrt{n\sigma_1^2}. \tag{3.2.13}$$

∘

- Um zu garantieren, daß auch die Normen der Matrizen A_k und B_k nach jeder Transformation beschränkt bleiben wird die Diagonalmatrix D_k definiert, die sich nur an den Stellen (p,p) und (q,q) von der Einheitsmatrix unterscheidet:

$$\tilde{D}_k := \begin{bmatrix} \frac{1}{\sqrt{(B_k^T B_k)_{pp}}} & 0 \\ 0 & \frac{1}{\sqrt{(B_k^T B_k)_{qq}}} \end{bmatrix} \tag{3.2.14}$$

$\tilde{D}_k$ bezeichnet die 2×2-Restriktion von D_k.

- Die k-te Transformation hat folgende Gestalt:

$$\begin{array}{ll} \hat{A}_k := U_k^T A_{k-1} X_k & \hat{B}_k := V_k^T B_{k-1} X_k \\ A_k := \hat{A}_k D_k & B_k := \hat{B}_k D_k \end{array} \tag{3.2.15}$$

Bemerkung 3.2.16 Wie aus (3.2.5) und (3.2.6) ersichtlich ist, handelt es sich bei den Zeilentransformationen um unitärte Transformationen, so daß sich bei diesen Transformationen die Norm der Spalentenvektoren von A_k und B_k nicht ändern (vgl. Lemma 1.1.2). Bei der rechtseitigen Multiplikation mit X_k werden die p-te und q-te Spalte kombiniert, daher bietet sich Modifikation der Matrix X_k an. D.h. die k-te Transformation (3.2.15) kann auch geschrieben werde als:

$$\begin{array}{lll} A_k := U_k^T A_{k-1} \hat{X}_k & B_k := V_k^T B_{k-1} \hat{X}_k & k = 1, 2, ... \\ \hat{X}_k := X_k D_k. & & \end{array}$$

∘

Bestimmung der Transformationsparameter bei regulärer Matrix B

- Die Winkel φ_k und ψ_k für die 2×2-Restriktionen von A_{k-1}, B_{k-1}, U_k^T und V_k^T werden entsprechend der Bedingung (3.2.8) bestimmt, so daß gilt:

$$\tilde{U}_k^T \tilde{A}_{k-1} = \Sigma_k \tilde{V}_k^T \tilde{B}_{k-1}, \tag{3.2.17}$$

  wobei Σ_k eine 2×2-Diagonalmatrix ist.

- Für invertierbare $\tilde{B}_{k-1}$ kann (3.2.17) umgeschrieben werden:

$$\tilde{U}_k^T (\tilde{A}_{k-1} \tilde{B}_{k-1}^{-1}) \tilde{V}_k = \Sigma_k \tag{3.2.18}$$

  d.h. es wird eine Singulärwertzerlegung für $\tilde{A}_{k-1} \tilde{B}_{k-1}^{-1}$ ausgeführt.

- Da A_{k-1} und B_{k-1} in jedem Schritt des Verfahrens aufgrund des spalten- bzw. zeilenzyklischen Vorgehens immer hierarische Blockdreiecksgestalt besitzen (vgl. Satz 1.1.16), reicht es wieder die Pivotblöcke zu betrachten:

$$\tilde{B}_{k-1} = \begin{bmatrix} b_{pp}^{k-1} & b_{pq}^{k-1} \\ 0 & b_{qq}^{k-1} \end{bmatrix} \qquad \text{im ungeraden Zyklus und}$$

$$\tilde{B}_{k-1} = \begin{bmatrix} b_{pp}^{k-1} & 0 \\ b_{qp}^{k-1} & b_{qq}^{k-1} \end{bmatrix} \qquad \text{im geraden Zyklus.}$$

- **Bestimmung der Parameter im ungeraden Zyklus:**

$$\begin{aligned} \tilde{A}_{k-1} \tilde{B}_{k-1}^{-1} &= \begin{bmatrix} a_{pp}^{k-1} & a_{pq}^{k-1} \\ 0 & a_{qq}^{k-1} \end{bmatrix} \begin{bmatrix} (b_{pp}^{k-1})^{-1} & -b_{pq}^{k-1}(b_{pp}^{k-1} b_{qq}^{k-1})^{-1} \\ 0 & (b_{qq}^{k-1})^{-1} \end{bmatrix} \\ &= \begin{bmatrix} a_{pp}^{k-1}(b_{pp}^{k-1})^{-1} & (-a_{pp}^{k-1} b_{pq}^{k-1} + a_{pq}^{k-1} b_{pp}^{k-1})(b_{pp}^{k-1} b_{qq}^{k-1})^{-1} \\ 0 & a_{qq}^{k-1}(b_{qq}^{k-1})^{-1} \end{bmatrix} \end{aligned}$$

– Entsprechend zum Verfahren von Hari-Veselić (vgl. (3.1.9)) ergeben sich für die Winkel die folgende Bestimmungsgleichungen:

$$\tan 2\psi_k = \frac{2a_{pp}^{k-1}b_{qq}^{k-1}(a_{pq}^{k-1}b_{pp}^{k-1} - a_{pp}^{k-1}b_{pq}^{k-1})}{(a_{qq}^{k-1})^2(b_{pp}^{k-1})^2 - (a_{pp}^{k-1})^2(b_{qq}^{k-1})^2 + (a_{pq}^{k-1}b_{pp}^{k-1} - a_{pp}^{k-1}b_{pq}^{k-1})^2} \tag{3.2.19}$$

$$\tan \varphi_k = \frac{b_{pp}^{k-1}a_{qq}^{k-1}\tan\psi_k}{a_{pp}^{k-1}b_{qq}^{k-1} - (a_{pq}^{k-1}b_{pp}^{k-1} - a_{pp}^{k-1}b_{pq}^{k-1})\tan\psi_k.} \tag{3.2.20}$$

– Mit $\tilde{V}_k^T\tilde{B}_{k-1}\tilde{X}_k = \begin{bmatrix} \cos\psi_k & -\sin\psi_k \\ \sin\psi_k & \cos\psi_k \end{bmatrix}\begin{bmatrix} b_{pp}^{k-1} & b_{pq}^{k-1} \\ 0 & b_{qq}^{k-1} \end{bmatrix}\begin{bmatrix} 1 & \alpha_k \\ -\beta_k & 1 \end{bmatrix} = \begin{bmatrix} b_{pp}^k & 0 \\ 0 & b_{qq}^k \end{bmatrix}$ (vgl. (3.2.8)) ergibt sich nun:

$$\alpha_k = \tan\psi_k\frac{b_{qq}^{k-1}}{b_{pp}^{k-1}} - \frac{b_{pq}^{k-1}}{b_{pp}^{k-1}} = \tan\varphi_k\frac{a_{qq}^{k-1}}{a_{pp}^{k-1}} - \frac{a_{pq}^{k-1}}{a_{pp}^{k-1}} \tag{3.2.21}$$

$$\beta_k = \frac{\tan\psi_k b_{pp}^{k-1}}{\tan\psi_k b_{pq}^{k-1} + b_{qq}^{k-1}} = \frac{\tan\varphi_k a_{pp}^{k-1}}{\tan\varphi_k a_{pq}^{k-1} + a_{qq}^{k-1}} \tag{3.2.22}$$

– Die Normierungsfaktoren (vgl. (3.2.14)) berechnen sich aus:

$$\begin{aligned} (B_k^TB_k)_{pp} &= [(1+\alpha\beta)\cos\psi b_{pp}]^2 + \sum_{\substack{i=1\\ i\neq p,q}}^{n}(b_{ip} - \beta b_{iq})^2 \\ (B_k^TB_k)_{qq} &= \left[\frac{b_{qq}}{\cos\psi}\right]^2 + \sum_{\substack{i=1\\ i\neq p,q}}^{n}(\alpha b_{ip} + b_{iq})^2 \end{aligned} \tag{3.2.23}$$

- **Bestimmung der Parameter im geraden Zyklus:**

$$\begin{aligned} \tilde{A}_{k-1}\tilde{B}_{k-1}^{-1} &= \begin{bmatrix} a_{pp}^{k-1} & 0 \\ a_{qp}^{k-1} & a_{qq}^{k-1} \end{bmatrix}\begin{bmatrix} (b_{pp}^{k-1})^{-1} & 0 \\ -b_{qp}^{k-1}(b_{pp}^{k-1}b_{qq}^{k-1})^{-1} & (b_{qq}^{k-1})^{-1} \end{bmatrix} \\ &= \begin{bmatrix} a_{pp}^{k-1}(b_{pp}^{k-1})^{-1} & 0 \\ (-a_{qq}^{k-1}b_{qp}^{k-1} + a_{qp}^{k-1}b_{qq}^{k-1})(b_{pp}^{k-1}b_{qq}^{k-1})^{-1} & a_{qq}^{k-1}(b_{qq}^{k-1})^{-1} \end{bmatrix} \end{aligned}$$

– Entsprechend zum Verfahren von Hari-Veselić (vgl. (3.1.10)) ergeben sich für die Winkel die folgende Bestimmungsgleichungen:

$$\tan 2\varphi_k = \frac{2a_{pp}^{k-1}b_{qq}^{k-1}(a_{qp}^{k-1}b_{qq}^{k-1} - a_{qq}^{k-1}b_{qp}^{k-1})}{(a_{qq}^{k-1})^2(b_{pp}^{k-1})^2 - (a_{pp}^{k-1})^2(b_{qq}^{k-1})^2 + (a_{qp}^{k-1}b_{qq}^{k-1} - a_{qq}^{k-1}b_{qp}^{k-1})^2} \tag{3.2.24}$$

$$\tan \psi_k = \frac{b_{pp}^{k-1}a_{qq}^{k-1}\tan\varphi_k}{a_{pp}^{k-1}b_{qq}^{k-1} - (a_{qp}^{k-1}b_{qq}^{k-1} - a_{qq}^{k-1}b_{qq}^{k-1})\tan\varphi_k.} \tag{3.2.25}$$

– Mit $\tilde{V}_k^T \tilde{B}_{k-1} \tilde{X}_k = \begin{bmatrix} \cos\psi_k & -\sin\psi_k \\ \sin\psi_k & \cos\psi_k \end{bmatrix} \begin{bmatrix} b_{pp}^{k-1} & 0 \\ b_{qp}^{k-1} & b_{qq}^{k-1} \end{bmatrix} \begin{bmatrix} 1 & \alpha_k \\ -\beta_k & 1 \end{bmatrix} = \begin{bmatrix} b_{pp}^k & 0 \\ 0 & b_{qq}^k \end{bmatrix}$
(vgl. (3.2.8)) ergibt sich nun:

$$\alpha_k = \frac{\tan\psi_k b_{qq}^{k-1}}{b_{pp}^{k-1} - \tan\psi_k b_{qp}^{k-1}} = \frac{\tan\varphi_k a_{qq}^{k-1}}{a_{pp}^{k-1} - \tan\varphi_k a_{qp}^{k-1}} \tag{3.2.26}$$

$$\beta_k = \tan\psi_k \frac{b_{pp}^{k-1}}{b_{qq}^{k-1}} + \frac{b_{qp}^{k-1}}{b_{qq}^{k-1}} = \tan\varphi_k \frac{a_{pp}^{k-1}}{a_{qq}^{k-1}} - \frac{a_{qp}^{k-1}}{a_{qq}^{k-1}} \tag{3.2.27}$$

– Die Normierungsfaktoren (vgl. (3.2.14)) berechnen sich aus:

$$\begin{aligned} (B_k^T B_k)_{pp} &= \left[\frac{b_{pp}}{\cos\psi}\right]^2 + \sum_{\substack{i=1 \\ i\neq p,q}}^{n} (b_{ip} + \beta b_{iq})^2 \\ (B_k^T B_k)_{qq} &= [(1+\alpha\beta)\cos\psi b_{qq}]^2 + \sum_{\substack{i=1 \\ i\neq p,q}}^{n} (\alpha b_{ip} - b_{iq})^2 \end{aligned} \tag{3.2.28}$$

Bestimmung der Transformationsparameter bei singulärer Matrix B

Falls B eine singuläre Matrix ist, kann das verallgemeinerte Singulärwertproblem ebenfalls nach diesem Prinzip gelöst werden, jedoch ohne die Normierung.

Satz 3.2.29 *Zu den 2×2-Matrizen A und B werden die unitären Matrizen U und V so gewählt, daß gilt:*

$$U^T A B_{adj} V = \Sigma = \operatorname{diag}(\sigma_1, \sigma_2), \qquad \sigma_1 \geq \sigma_2 \geq 0.$$

Die i-te Zeile von $U^T A$ ist dann parallel zur i-ten Zeile von $V^T B$, $i = 1, 2$. B_{adj} bezeichnet die adjungierte Matrix von B.
Mit $AB_{adj} \neq 0$ und $U = [u_1, u_2]$, $V = [v_1, v_2]$ ergibt sich

$$\begin{aligned} & u_1^T A \neq 0, && u_2^T B \neq 0, \\ & \operatorname{rang}(A) = 1 && \Longrightarrow \; u_2^T A = 0, \\ & \operatorname{rang}(B) = 1 && \Longrightarrow \; v_1^T B = 0. \end{aligned}$$

[Paige86]

Bemerkung 3.2.30 Für singuläre B muß also ein neues, zu (3.2.17) bzw. (3.2.18) äquivalente Problem, gelöst werden:

$$U^T A B_{adj} V = \Sigma. \tag{3.2.31}$$

∘

Satz 3.2.32 *Der Ansatz (3.2.31) führt auf dieselben Bestimmungsgleichungen (3.2.19)-(3.2.22) im ungeraden, bzw. (3.2.24)-(3.2.27) im geraden Zyklus.*

Beweis: Es muß lediglich gezeigt werden, daß die Winkel, die sich in der Gleichung (3.2.31) ergeben, mit denen aus (3.2.18) übereinstimmen:

- **ungerader Zyklus:**

$$\tilde{A}_{k-1}(\tilde{B}_{k-1})_{adj} = \begin{bmatrix} a_{pp}^{k-1} & a_{pq}^{k-1} \\ 0 & a_{qq}^{k-1} \end{bmatrix} \begin{bmatrix} (b_{qq}^{k-1})^{-1} & -b_{pq}^{k-1} \\ 0 & (b_{pp}^{k-1})^{-1} \end{bmatrix} = \begin{bmatrix} a_{pp}^{k-1}b_{qq}^{k-1} & (-a_{pp}^{k-1}b_{pq}^{k-1} + a_{pq}^{k-1}b_{pp}^{k-1}) \\ 0 & a_{qq}^{k-1}(b_{pp}^{k-1}) \end{bmatrix}$$

Entsprechend zum Verfahren von Hari-Veselić (vgl. (3.1.9)) ergeben sich für die Winkel die folgende Bestimmungsgleichungen:

$$\tan 2\psi_k = \frac{2a_{pp}^{k-1}b_{qq}^{k-1}(a_{pq}^{k-1}b_{pp}^{k-1} - a_{pp}^{k-1}b_{pq}^{k-1})}{(a_{qq}^{k-1})^2(b_{pp}^{k-1})^2 - (a_{pp}^{k-1})^2(b_{qq}^{k-1})^2 + (a_{pq}^{k-1}b_{pp}^{k-1} - a_{pp}^{k-1}b_{pq}^{k-1})^2} \tag{3.2.33}$$

$$\tan \varphi_k = \frac{b_{pp}^{k-1}a_{qq}^{k-1}\tan\psi_k}{a_{pp}^{k-1}b_{qq}^{k-1} - (a_{pq}^{k-1}b_{pp}^{k-1} - a_{pp}^{k-1}b_{pq}^{k-1})\tan\psi_k.} \tag{3.2.34}$$

Das sind gerade (3.2.19) und (3.2.20).

- **gerader Zyklus:**

$$\tilde{A}_{k-1}(\tilde{B}_{k-1})_{adj} = \begin{bmatrix} a_{pp}^{k-1} & 0 \\ a_{qp}^{k-1} & a_{qq}^{k-1} \end{bmatrix} \begin{bmatrix} b_{qq}^{k-1} & 0 \\ -b_{qp}^{k-1} & b_{pp}^{k-1} \end{bmatrix} = \begin{bmatrix} a_{pp}^{k-1}b_{qq}^{k-1} & 0 \\ -a_{qq}^{k-1}b_{qp}^{k-1} + a_{qp}^{k-1}b_{qq}^{k-1} & a_{qq}^{k-1}b_{pp}^{k-1} \end{bmatrix}$$

Das Verfahren von Hari-Veselić (vgl. (3.1.10)) liefert wieder für die Winkel genau die Gleichungen (3.2.24) und (3.2.25).

□

3.3 Algorithmus zur Berechnung der GSVD

Der folgende Algorithmus 3.3.7 berechnet die verallgemeinerten Singulärwertzerlegung $U^TAX = \Sigma V^TBX,\ \Sigma = \mathrm{diag}(\sigma_1, \sigma_2, \cdots, \sigma_3)$ des quadratischen Matrixpaares (A, B).
Es können aber auch die Singulärwerte beliebiger Matrizen $\tilde{A} \in I\!R^{m\times n},\ m \le n$, und $\tilde{B} \in I\!R^{p\times n},\ p \le n$, mit dem Algorithmus berechnet werden, indem man sie durch Erweiterung mit entsprechend vielen Null-Zeilen bzw. Null-Spalten in quadratische Matrizen $A, B \in I\!R^{n\times n}$ umwandelt. Dies ist im allgemeinen nur günstig, wenn die Dimensionen der Matrizen sehr groß sind und damit die Anzahl der hinzu gefügten Nullen verhälnismäßig klein ist, so daß Speedup und Performance nicht wesentlich beeinflußt werden.

Die Matrizen A und B werden zunächst in die folgenden Startmatrizen A_0 und B_0 orthogonal transformiert.

$$A_0 = \begin{array}{l} \left[\begin{array}{ccc} A_{11} & A_{12} & A_{13} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{array}\right] \begin{array}{l} r \\ \\ \\ \end{array} \end{array} \in I\!R^{n\times n} \quad B_0 = \left[\begin{array}{ccc} B_{11} & B_{12} & B_{13} \\ 0 & B_{22} & B_{23} \\ 0 & 0 & 0 \end{array}\right] \begin{array}{l} r \\ \\ \\ \end{array} \in I\!R^{n\times n}. \tag{3.3.1}$$

- Der Elemente $\neq 0$ von B_0 stehen in den oberen Zeilen, die daher vollen Zeilenrang besitzen.
- A_{11} ist eine untere Dreiecksmatrix und regulär,
- B_{11} ist eine untere Dreiecksmatrix,
- B_{22} ist eine untere Dreiecksmatrix und für $q > 0$ regulär.

Die Iterationsmatrizen $A_k = U_{k-1}^T A_{k-1} X_{k-1}$ und $B_k = V_{k-1} B_{k-1} X_{k-1}$ haben nach einem geraden oder ungeraden Zyklus jeweils eine andere Gliederung (vgl. Seite 11):

$$A_k = \left[\begin{array}{cccc} A_{11}^k & A_{12}^k & A_{13}^k & 0 \\ 0 & A_{22}^k & A_{23}^k & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] \begin{array}{l} r_1 \\ r_2 \\ \\ \\ \end{array}, \quad B_k = \left[\begin{array}{cccc} 0 & 0 & 0 & 0 \\ 0 & B_{22}^k & B_{23}^k & 0 \\ 0 & 0 & B_{33}^k & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] \begin{array}{l} r_1 \\ r_2 \\ \\ \\ \end{array}, \quad k > 0 \ \text{ gerade} \tag{3.3.2}$$

$A_{11}^k,\ A_{22}^k,\ B_{22}^k,\ B_{33}^k$ sind reguläre obere Dreiecksmatrizen, $r = r_1 + r_2$,

$$A_k = \left[\begin{array}{cccc} A_{11}^k & 0 & 0 & 0 \\ A_{21}^k & A_{22}^k & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] \begin{array}{l} r_1 \\ r_2 \\ \\ \\ \end{array}, \quad B_k = \left[\begin{array}{cccc} 0 & 0 & 0 & 0 \\ B_{21}^k & B_{22}^k & 0 & 0 \\ B_{31}^k & B_{32}^k & B_{33}^k & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] \begin{array}{l} r_1 \\ r_2 \\ \\ \\ \end{array}, \quad k > 1 \ \text{ ungerade} \tag{3.3.3}$$

$A_{11}^k,\ A_{22}^k,\ B_{22}^k,\ B_{33}^k$ sind reguläre untere Dreiecksmatrizen, $r = r_1 + r_2$.

Der Algorithmus konvergiert gegen eine reguläre obere Dreiecksmatrix

$$R = \begin{array}{cl} \begin{bmatrix} R_{11} & R_{12} & R_{13} \\ 0 & R_{22} & R_{23} \\ 0 & 0 & R_{33} \end{bmatrix} & \begin{array}{l} r_1 \\ r_2 \\ \end{array} \end{array} \in I\!R^{k\times k}, \qquad r = r_1 + r_2,$$

und die positiv definiten Matrizen $\Sigma_A = \operatorname{diag}(\alpha_{r_1+1}, ..., \alpha_r)$ und $\Sigma_B = \operatorname{diag}(\beta_{r_1+1}, ..., \beta_r)$ mit $\Sigma_A^2 + \Sigma_B^2 = I$.

$$U^H AX = \begin{bmatrix} R_{11} & R_{12} & R_{13} & 0 \\ 0 & \Sigma_A R_{22} & \Sigma_A R_{23} & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \tag{3.3.4}$$

$$V^H AB = \begin{bmatrix} 0_B & 0 & 0 & 0 \\ 0 & \Sigma_B R_{22} & \Sigma_B R_{23} & 0 \\ 0 & 0 & R_{33} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}. \tag{3.3.5}$$

- $0_B \in I\!R^{r_1\times r_1}$ ist eine Nullmatrix, so daß
- $[\Sigma_B R_{22}, \Sigma_B R_{23}]$ in (3.3.5) in derselben Position wie $[\Sigma_A R_{22}, \Sigma_A R_{23}]$ in (3.3.4) steht.
- Diese Zerlegung ist eine GSVD von A und B, vgl. [PaSa81].
- Die von Null verschiedenen Spalten entsprechen $U^H AX$ und $V^H BX$ in der Gleichung (3.2.17).

Bemerkung 3.3.6 Schreibweisen in Algorithmus 3.3.7:

- a_{ij} und b_{ij} sind Elemente der Matrizen A und B mit

  $A_{ij} := \begin{bmatrix} a_{ii} & a_{ij} \\ a_{ji} & a_{jj} \end{bmatrix}, \quad B_{ij} := \begin{bmatrix} b_{ii} & b_{ij} \\ b_{ji} & b_{jj} \end{bmatrix}.$
- U_{ij}, V_{ij} und X_{ij} sind orthogonale 2×2-Matrizen.
- Es gilt also $U_{ij}^T A_{ij} (B_{ij})_{adj} V_{ij} = \operatorname{diag}(\sigma_1, \sigma_2)$, $\sigma_1, \sigma_2 \geq 0$ mit $U_{ij} = V_{ij} = I$.
  Vergleiche Kapitel 3.2.

∘

Zusammenfassend gilt der folgende Algorithmus:

Algorithmus 3.3.7 (GSVD)

BEGIN A, B, n sind vorgegeben wie in (3.3.1) beschrieben
 $U := I_n;\ V := I_n;\ X := I_n;$
 cycle:=0;
 WHILE nonconvergence *AND (cycle* < 10*) DO*
 BEGIN
  cycle:=cycle + 1;
  FOR i:=1 TO n-1 DO
   FOR j:= i+1 TO n DO
   BEGIN
    1. Transformation der Elemente in Position (i,j):
     Wähle X_{ij} *so, daß* $U_{ij}^T A_{ij} X_{ij}$ *und* $V_{ij}^T B_{ij} X_{ij}$
      untere Dreiecksmatrizen in ungeraden Zyklen und
      obere Dreiecksmatrizen in geraden Zyklen sind.
    2. Definiere die Matrizen $U_n, V_n, X_n \in \mathbb{R}^{n\times n}$ *als Einheitsmatrizen, die in den Positionen (i,i),(i,j),(j,i) und (j,j) die entsprechenden Einräge der Matrizen* U_{ij}, V_{ij} *und* X_{ij} *haben.*
    3. $A := U_n^T A X_n;$
     $B := V_n^T B X_n;$
     $U := U U_n;\quad V := V V_n;\quad X := X X_n;$
   END
  END
 END

Bemerkung 3.3.8

- Die Partitionierung der Matrizen A und B in (3.3.1) als Blockmatrizen bildet den Grundstein dafür den Algorithmus 3.3.7 parallel zu berechnen.
- Da in jedem Transformationsschritt nur je zwei Spalten und Zeilen verändert werden, läßt sich, wie für das spezielle Singulärwertproblem, parallele Algorithmen finden.
  Vergleiche [SchZi99].

$\circ$

Literatur

[Björck96] Å. Björck, *Least Squares Problems,* siam, Philadelphia 1996.

[BuGe85] W. Bunse, A. Bunse-Gerstner, *Numerische lineare Algebra,* Teubner Studienbücher Mathematik, Stuttgart 1985.

[Deich91] A. Deichmöller, *Über die Berechnung verallgemeinerter Singulär Werte mittels Jacobi-ähnlicher Verfahren,* Dissertation, Fernuniversität-Gesamthochschule-Hagen, 1991.

[GoLo96] G.H. Golub, C.F. van Loan, *Matrix Computations,* The Johns Hopkins University Press, Baltimore 1996.

[HoJo85] R.A. Horn, C.R. Johnson, *Matrix Analysis,* Cambridge University Press, Cambrdge 1985.

[Maess85] G. Maess, *Lineare Algebra, Vorlesungen über numerische Mathematik* Birkhäuser Verlag, Basel 1985.

[Paige86] C.C. Paige, *Computing the Generalized Singular Value Decomposition,* SIAM J. Sci. Stat. Comput., Vol. 7, No. 4, pp.1126-1146, October 1986.

[PaSa81] C.C. Paige, M.A. Sauders, *Towards a Generalized Singular Value Decomosition,* SIAM J. Numer. Anal., Vol. 18, No. 3, pp.398-405, June 1981.

[SchZi99] U. Schendel, A. El-Fares−Ziegler, *Die Singulärwertzerlegung (SVD) und parallele Verfahren zur Berechnung der SVD,* Preprint Nr.A/12/99, Fachbereich Mathematik und Informatik, Freie Universität Berlin, 1999.

[SchZi97] U. Schendel, A. Ziegler, *Numerische Algorithmen der Linearen Algebra,* Vorlesungsskript, Fachbereich Mathematik und Informatik, Freie Universität Berlin, 1997.

[Schwarz93] H.R. Schwarz, *Numerische Mathematik,* B.G. Teubner Stuttgart 1993.

[Überh95] C. Überhuber, *Computer-Numerik 2,* Springer-Verlag, Heidelberg 1995.

[vanL76] Ch.F. van Loan, *Generalizing the Singular Value Decomposition,* SIAM J. Numer. Anal. Vol.13, No. 1, March 1976.

[Werner92] J. Werner, *Numerische Mathematik 1 und 2,* Vieweg & Sohn, Braunschweig/Wiesbaden 1992.

[Wilkin92] J.H. Wilkinson, *The Algebraic Eigenvalue Problem,* Clarendon Press, Oxford, 1992.

[WiRe71] J.H. Wilkinson, C. Reinisch, *Lineare Algebra,* Die Grundlehren der mathematischen Wissenschaften Band 186, Springer-Verlag, Berlin 1971.