DIE SINGULÄRWERTZERLEGUNG

THOMAS WILLWACHER

1. Definition

Theorem/Definition 1.1 (Singulärwertzerlegung). *Sei $\mathbb{K} = \mathbb{R}$ oder $\mathbb{C}$ und $A \in M(m \times n, \mathbb{K})$ eine reelle oder komplexe Matrix. Dann gibt es Matrizen $U \in O(m)$ (für $\mathbb{K} = \mathbb{R}$) bzw. $U \in U(m)$ (für $\mathbb{K} = \mathbb{C}$) und $V \in O(n)$ bzw. $V \in U(n)$ und eine "diagonale" Matrix $D \in M(m \times n, \mathbb{R})$ der Form*

$$D = \begin{pmatrix} \sigma_1 & 0 & \dots & 0 \\ 0 & \sigma_2 & \dots & 0 \\ \vdots & & \ddots & \vdots \end{pmatrix}$$

mit Diagonaleinträgen $\sigma_j \in \mathbb{R}$, $\sigma_j \geq 0$ so dass

$$A = U D \bar{V}^T.$$

Diese Zerlegung heisst Singulärwertzerlegung *von A, und die nichtnegativen reellen Zahlen $\sigma_1, \sigma_2, \dots, \sigma_{\min\{m,n\}}$ heissen* Singulärwerte *von A. In der Regel ordnet man die Singulärwerte so, dass $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_{\min\{m,n\}}$.*

Bemerkung 1.1.

- Man beachte, dass D hier eine rechteckige $m \times n$ Matrix ist, nicht unbedingt quadratisch. Dennoch wolllen wir sie als Diagonalmatrix bezeichnen.
- Man könnte äquivalent auch V surch $V^{-1} = \bar{V}^T$ ersetzen und schreiben $A = UDV$. Allerdings ist die obige Form symmetrischer unter $A \leftrightarrows \bar{A}^T$, und oft praktischer.
- Der Rang von A ist gleich dem Rang von D, da sich der Rang nicht ändert, wenn man mit invertierbaren Matrizen von links oder rechts multipliziert. Also ist der Rang von A die Anzahl der nicht verschwindenden Singulärwerte von A.
- Man findet die Singulärwertzerlegung in der Literatur auch oft in der folgenden alternativen Form:

  Zu $A \in M(m \times n, \mathbb{K})$ vom Rang r gibt es Matrizen $\tilde{U} \in M(m \times r, \mathbb{K})$ und $\tilde{V} \in M(n \times r, \mathbb{K})$ mit orthonormalen Spalten und Zahlen $\sigma_1 \geq \sigma_2, \cdots, \sigma_r > 0$ so dass

  $$A = \tilde{U}\tilde{D}\overline{(\tilde{V})}^T = \tilde{U}\tilde{D}\tilde{V}^\dagger \tag{1}$$

  mit

  $$\tilde{D} = \begin{pmatrix} \sigma_1 & 0 & \dots & 0 \\ 0 & \sigma_2 & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & & \sigma_r \end{pmatrix} \in M(r \times r, \mathbb{R})$$

  und der Abkürzung $(-)^\dagger := \overline{(-)}^T$.

  Um diese Zerlegung aus obigem Theorem zu erhalten ordnet man zunächst die Singulärwerte wie angegeben. Dann definiert man $\tilde{U}$ als die ersten r Spalten von U und $\tilde{V}$ als die ersten r Spalten von V. Umgekehrt erhält man U aus $\tilde{U}$ einfach, indem man die durch die Spalten von $\tilde{U}$ gegebene orthonormale Familie zu einer Orthonormalbasis des $\mathbb{K}^m$ fortsetzt. (Analog erhält man V aus $\tilde{V}$.)
- Man beachte, dass die Singulärwertzerlegung immer existiert, ohne zusätzliche Bedingungen an A.
- Die ersten $\min\{m, n\}$ Spalten von U (bzw. V) sind linke bzw. rechte Singulärvektoren von A. Genauer nennt man Einheitsvektoren $u \in \mathbb{K}^m$, $v \in \mathbb{K}^n$ linke bzw. rechte Singulärvektoren zum Singulärwert $\sigma \in \mathbb{R}_{\geq 0}$ falls gilt

  $$Av = \sigma u \quad \text{und} \quad A^\dagger u = \sigma v.$$

- Man kann die Singulärwertzerlegung auch nochmals äquivalent "matrixfrei" umformulieren: Sei $F \in \mathrm{Hom}(V, W)$ eine lineare Abbildung zwischen den endlich dimensionalen unitären bzw. euklidischen Vektorräumen V und W. Dann gibt es Orthonormalbasen $\mathcal{A}$ von V und $\mathcal{B}$ von W so dass
$$M_{\mathcal{B}}^{\mathcal{A}}(F) = D$$
Diagnalform hat mit nicht-negativen reellen Diagonaleinträgen. (Letztere sind die Singulärwerte.) Vergleiche die Aussage in dieser Form insbesondere mit Korollar 2.4.3 im Fischer.

Wir brauchen das folende Resultat, dass wir evtl. auch schon gesehen haben:

Lemma 1.2. *Sei $A \in M(m \times n, \mathbb{K})$ eine beliebige Matrix vom Rang r. Dann sind die hermiteschen (bzw. symmetrischen) Matrizen $\bar{A}^T A$ und $A\bar{A}^T$ positiv semidefinit und haben Rang r. Genauer gilt, dass*
$$\mathrm{Ker}\bar{A}^T A = \mathrm{Ker}A$$
und dass
$$\mathrm{Im}A\bar{A}^T = \mathrm{Im}A$$

Beweis. Wir betrachten zunächst den Fall $\bar{A}^T A$. Zunächst ist für alle $x \in \mathbb{K}^n$:
$$x^T \bar{A}^T A \bar{x} = (\bar{A})^T \overline{(\bar{A}x)}^T = \langle \bar{A}x, \bar{A}x \rangle \geq 0.$$
Also ist $\bar{A}^T A$ positiv semidefinit. Es gilt offensichtlich $x \in \mathrm{Ker}A \Rightarrow x \in \mathrm{Ker}\bar{A}^T A$. Umgekehrt gilt aber auch
$$x \in \mathrm{Ker}\bar{A}^T A \Rightarrow 0 = \langle \bar{A}^T Ax, x \rangle = \langle Ax, Ax \rangle \Leftrightarrow Ax = 0.$$
Also ist $\mathrm{Ker}\bar{A}^T A = \mathrm{Ker}A$. Damit sind die Ränge der beiden Matrizen nach der Dimensionsformel gleich.

Die Aussagen für $A\bar{A}^T$ folgen aus den schon bewiesenen durch Vertauschung von A und $\bar{A}^T$, und mit der aus der Vorlesung bekannten Formel (Satz 6.2.4 im Fischer und der analoge Satz für komplexe Vektorräume)
$$\mathrm{Im}(A\bar{A}^T) = \left(\mathrm{Ker}(A\bar{A}^T)\right)^{\perp} = \left(\mathrm{Ker}\bar{A}^T\right)^{\perp} = \mathrm{Im}A.$$
□

Beweis vom Theorem. Wir betrachten zunächst den komplexen Fall $\mathbb{K} = \mathbb{C}$, und wir zeigen die Singulärwertzerlegung in der Version (1). Die Matrix $AA^\dagger \in M(m \times m, \mathbb{C})$ ist hermitesch und (wie wir schon gesehen haben) positiv semidefinit. Es gibt also dank des Spektralsatzes ein $U \in U(m)$ so dass
$$AA^\dagger = UD_1U^\dagger$$
mit D_1 diagonal mit nicht negativen rellen Diagonaleinträgen
$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r > \lambda_{r+1} = \cdots = \lambda_m = 0,$$
die wir o.B.d.A. der Grösse nach ordnen können. Man beachte, dass hier die Anzahl r, also die Zahl der nicht verschwindenden Eigenwerte von $AA^\dagger$, gleichzeitig der Rang von $AA^\dagger$ und nach dem Lemma auch der Rang von A ist. Wir zerlegen nun
$$U = \begin{pmatrix} \tilde{U} & U' \end{pmatrix}$$
in zwei Blöcke, wobei $\tilde{U} \in M(m \times r, \mathbb{C})$ aus den ersten r Spalten von U besteht und entsprechend $U' \in M(m \times (m-r), \mathbb{C})$ aus den restlichen. Man beachte, dass
$$\mathrm{Im}\,\tilde{U} = \mathrm{Im}(AA^\dagger) = \mathrm{Im}\,A,$$
wobei die erste Gleichung aus der Definition von $\tilde{U}$ folgt und die zweite aus dem Lemma. Insbesondere gilt ferner für alle $x = \tilde{U}y \in \mathrm{Im}\,\tilde{U}$:
$$\tilde{U}\tilde{U}^\dagger x = \tilde{U}\tilde{U}^\dagger\tilde{U}y = \tilde{U}E_r y = \tilde{U}y = x.$$
Also gilt auch

(2)
$$\tilde{U}\tilde{U}^\dagger A = A.$$

Wir setzen nun $\sigma_j := \sqrt{\lambda_j}$ für $j = 1, \ldots, r$ und
$$\tilde{D} = \begin{pmatrix} \sigma_1 & 0 & \ldots & 0 \\ 0 & \sigma_2 & \ldots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & & \sigma_r \end{pmatrix} \in M(r \times r, \mathbb{R}).$$

Man beachte, dass dann insbesondere gilt:

$$AA^\dagger = UD_1U^\dagger = \tilde{U}\tilde{D}^2\tilde{U}^\dagger$$

Ferner definieren wir

$$\tilde{V} = A^\dagger\tilde{U}\tilde{D}^{-1}$$

Dann gilt offensichtlich

$$\tilde{U}\tilde{D}\tilde{V}^\dagger = \tilde{U}\tilde{D}\tilde{D}^{-1}\tilde{U}^\dagger A = \tilde{U}U^\dagger A = A.$$

Hier haben wir zum einen benutzt, dass $\tilde{D}$ reell und symmetrisch ist und ausserdem Gleichung (2) im letzten Schritt.

Wir müssen also nur noch zeigen, dass die Spalten von $\tilde{V}$ tatsächlich orthonormal sind:

$$\begin{aligned}\tilde{V}^\dagger\tilde{V} &= \tilde{D}^{-1}\tilde{U}^\dagger \underbrace{AA^\dagger}_{=UD_1U^\dagger=\tilde{U}\tilde{D}^2\tilde{U}^\dagger} \tilde{U}\tilde{D}^{-1} \\ &= \tilde{D}^{-1}\tilde{U}^\dagger\tilde{U}\tilde{D}^2\tilde{U}^\dagger\tilde{U}\tilde{D}^{-1} \\ &= \tilde{D}^{-1}E_r\tilde{D}^2E_r\tilde{D}^{-1} = E_r.\end{aligned}$$

Den rellen Fall $\mathbb{K} = \mathbb{R}$ beweist man genau gleich, ausser dass dann $U \in O(m) \subset U(m)$ ist. □

Bemerkung 1.3. Die Singulärwerte von A sind durch A eindeutig bestimmt. Sei nämlich

$$A = UDV^\dagger$$

eine Singulärwertzerlegung von A wie im Theorem. Dann gilt

$$AA^\dagger = UDD^TU^\dagger.$$

Also sind die Singulärwerte von A gerade die Wurzeln der grössten $\min\{m,n\}$ Eigenwerte von $AA^\dagger$ (und $A^\dagger A$), wie man auch in obigem Beweis sieht, und damit eindeutig durch A festgelegt. Man redet also auch von "den Singulärwerten" von A. Die Matrizen U und V sind allerdings nicht (ganz) eindeutig definiert, trotzdem reden wir von *der* Singulärwertzerlegung.

1.1. **Min-max Eigenschaft.** Analog zum Satz von Courant-Fischer für die Eigenwerte hat man folgendes Resultat für die Singulärwerte.

Korollar 1.4. *Sei $A \in M(m \times n, \mathbb{K})$ eine Matrix und $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_p \geq 0$ die Singulärwerte von A mit $p := \min\{m,n\}$. Dann gilt*

$$\sigma_j = \min_{\substack{U\subset\mathbb{K}^n \\ \dim U = n-j+1}} \max_{\substack{u\in U \\ u\neq 0}} \frac{\|Au\|}{\|u\|} = \min_{\substack{U\subset\mathbb{K}^m \\ \dim U = m-j+1}} \max_{\substack{u\in U \\ u\neq 0}} \frac{\|A^\dagger u\|}{\|u\|}.$$

Insbesondere gilt

$$\sigma_1 = \max_{\substack{u\in\mathbb{K}^n \\ u\neq 0}} \frac{\|Au\|}{\|u\|} = \max_{\substack{u\in\mathbb{K}^n \\ \|u\|=1}} \|Au\|$$

Beweis. Wir zeigen nur die erste Gleichung, die zweite folgt analog. Sei

$$\lambda = \min_{\substack{U\subset\mathbb{K}^m \\ \dim U = m-j+1}} \max_{\substack{u\in U \\ u\neq 0}} \frac{\|Au\|}{\|u\|}.$$

Da $x \mapsto x^2$ auf $\mathbb{R}_{\geq 0}$ eine streng monoton steigende Funktion ist können wir in der Formel oben äquivalent auch die Grösse $\frac{\|Au\|^2}{\|u\|^2}$ extremieren, und erhalten dann einfach λ^2, also

$$\lambda^2 = \min_{\substack{U\subset\mathbb{K}^m \\ \dim U = m-j+1}} \max_{\substack{u\in U \\ u\neq 0}} \frac{\|Au\|^2}{\|u\|^2}.$$

Wegen

$$\|Au\|^2 = \langle u, A^\dagger Au\rangle$$

sehen wir dann aber aus dem Satz von Courant-Fischer, dass λ^2 gerade der j-grösste Eigenwert von $A^\dagger A$ ist. Also ist wegen der Bemerkung oben $\lambda = \sigma_j$. Man beachte, dass man aus der zweiten Gleichung im Satz von Courant-Fischer ähnlich auch eine max-min-Charakterisierung der Singulärwerte erhält, auf die wir hier aber verzichten. □

Eine Anwendung ist die folgende Interpretation des maximalen Singulärwertes.

Definition 1.5. *Die Spektralnorm (oder 2-Norm) von $A \in M(m \times n, \mathbb{K})$ ist der maximale Singulärwert von A,*

$$\|A\|_2 := \sigma_1(A) = \max_{\substack{u \in \mathbb{K}^n \\ u \neq 0}} \frac{\|Au\|}{\|u\|} = \max_{\substack{u \in \mathbb{K}^n \\ \|u\|=1}} \|Au\| .$$

Beweis. Man sollte noch zeigen, dass dies tatsächlich die Axiome einer Norm erfüllt. Das einzige allenfalls nicht offensichtliche Axiom dabei ist die Dreiecksungleichung:

$$\|A+B\|_2 = \max_{\substack{u \in \mathbb{K}^n \\ \|u\|=1}} \|Au + Bu\| \leq \max_{\substack{u \in \mathbb{K}^n \\ \|u\|=1}} (\|Au\| + \|Bu\|) \leq \max_{\substack{u \in \mathbb{K}^n \\ \|u\|=1}} (\|Au\|) + \max_{\substack{u \in \mathbb{K}^n \\ \|u\|=1}} (\|Bu\|) .$$

□

Bemerkung 1.6. Dies ist einfach die von der üblichen 2-Norm auf $\mathbb{K}^m$ bzw. $\mathbb{K}^n$ abgeleitete natürliche Operatornorm. Allgemeiner kann man für $F \in \mathrm{Hom}(V, W)$, mit V, W (sagen wir) endlich dimensionalen normierten (nicht Null-)Vektorräumen die Norm

$$\|F\|_{V,W} := \max_{\substack{v \in V \\ v \neq 0}} \frac{\|F(v)\|_W}{\|v\|_V}$$

erklären.

2. Anwendungen der Singulärwertzerlegung

Die Singulärwertzerlung ist in der numerischen Mathematik und für viele praktische Anwendungen von ausserordentlicher Wichtigkeit. Hier seien einige Beispiele gegeben.

2.1. **Regression und Pseudoinverse.**

Wir haben in der Vorlesung das (exakte) Lösen von linearen Gleichungssystemen

$$Ax = b$$

besprochen, mit $A \in M(m \times n, \mathbb{K})$. In der Praxis ist es fast immer der Fall, dass das Gleichungssystem gar keine Lösung hat, z.B. wegen Messfehlern. Man sucht dann stattdessen die bestmögliche approximative Lösung. Was "bestmöglich" heisst hängt dabei von der Anwendung ab, und kann je nachdem das Problem sehr schwierig machen. Wir betrachten hier nur den einfachsten Fall und suchen das (bzw. ein) $x \in \mathbb{K}^n$, das den Fehler

$$\|Ax - b\|$$

minmiert, bzw. äquivalent den quadratischen Fehler

$$E(x) := \|Ax - b\|^2$$

minimiert. Aus der Analysis wissen Sie, dass ein solches x erfüllt:

$$0 = DE(x)$$

mit DE der Ableitung von E. Das heisst, es muss für alle $h \in \mathbb{K}^n$ gelten:

$$\begin{aligned} 0 = DE(x)h &= D((Ax-b)^\dagger (Ax-b))h = D(x^\dagger A^\dagger Ax - b^\dagger Ax - x^\dagger A^\dagger b + b^\dagger b)h \\ &= h^\dagger A^\dagger Ax + x^\dagger A^\dagger Ah - b^\dagger Ah - h^\dagger A^\dagger b = 2\mathrm{Re}(h^\dagger (A^\dagger Ax - A^\dagger b)). \end{aligned}$$

Also muss gelten (ÜA)

$$A^\dagger Ax = A^\dagger b. \tag{3}$$

Sei nun $A = \tilde{U}\tilde{D}\tilde{V}^\dagger$ die Singulärwertzerlegung in der reduzierten Form (1), insbesondere ist also $\tilde{D} \in M(r \times r, \mathbb{R})$ invertierbar. Dann wird unsere Gleichung oben zu

$$\tilde{V}\tilde{D}^2\tilde{V}^\dagger x = \tilde{V}\tilde{D}\tilde{U}^\dagger b.$$

Diese Gleichung hat die (bzw. genauer: eine) Lösung

$$x = \tilde{V}\tilde{D}^{-1}\tilde{U}^\dagger b, \tag{4}$$

denn (setze ein und rechne nach)

$$\tilde{V}\tilde{D}^2\tilde{V}^\dagger\tilde{V}\tilde{D}^{-1}\tilde{U}^\dagger b = \tilde{V}\tilde{D}\tilde{U}^\dagger b.$$

Mann nennt die in (4) auftretende Matrix $\tilde{V}\tilde{D}^{-1}\tilde{U}^\dagger$ auch Pseudoinverse von A, da sie ähnliche Eigenschaften zur Inversen hat und hier die entsprechende Rolle spielt. Ausserdem ist für A quadratisch und invertierbar $\tilde{V}\tilde{D}^{-1}\tilde{U}^\dagger = A^{-1}$.

Bemerkung 2.1. Man kann natürlich (3) auch direkt und ohne Singulärwertzerlegung lösen. Falls $A^\dagger A$ invertierbar ist findet man z.B. direkt $x = (A^\dagger A)^{-1} A^\dagger b$ finden. In der Numerik macht man das allerdings nicht, da $(A^\dagger A)^{-1} A^\dagger$ dann viel schlechter konditioniert wäre.

2.2. **Approximation von Matrizen durch Matrizen kleinen Ranges.**

Definition 2.2. *Sei $A = (a_{ij}) \in M(m \times n, \mathbb{K})$ eine Matrix. Dann definieren wir die Frobenius-Norm von A als*

$$\|A\|_F := \sqrt{\sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^2} = \sqrt{\operatorname{tr}(A^\dagger A)},$$

wobei hier tr *die Spur ist, das heisst die Summe der Diagonalelemente. (Das heisst, für $B = (b_{ij})$ quadratisch ist $\operatorname{tr} B = \sum_i b_{ii}$.)*

Dies ist einfach die übliche 2-Norm auf $\mathbb{K}^{mn} \cong M(m \times n, \mathbb{K})$. Diese kommt von folgendem Skalarprodukt auf dem Raum $M(m \times n, \mathbb{K})$.

$$\langle A, B \rangle = \operatorname{tr}(A^T \bar{B}) = \operatorname{tr}(B^\dagger A).$$

Diese Norm hängt mit den Singulärwerten wie folgt zusammen.

Lemma 2.3. *Sei $A \in M(m \times n, \mathbb{K})$ eine Matrix und $U \in M(m \times m, \mathbb{K})$ und $V \in M(n \times n, \mathbb{K})$ unitär bzw. orthogonal. Dann gilt*

$$\|A\|_F = \|UAV\|_F.$$

Sind $\sigma_1, \dots, \sigma_p$ (mit $p = \min\{m, n\}$) die Singulärwerte von A so gilt ausserdem

$$\|A\|_F = \sqrt{\sigma_1^2 + \cdots + \sigma_p^2}.$$

Beweis. Es gilt

$$\|UAV\|_F^2 = \operatorname{tr}((UAV)^\dagger UAV) = \operatorname{tr}(V^\dagger A^\dagger U^\dagger UAV) = \operatorname{tr}(V^\dagger A^\dagger AV) = \operatorname{tr}(A^\dagger AVV^\dagger) = \operatorname{tr}(A^\dagger A) = \|A\|_F^2.$$

Bemerkung: Hier haben wir die Zyklizität der Spur ausgenutzt, also

$$\operatorname{tr}(XY) = \operatorname{tr}(YX) = \sum_{i,j} x_{ij} y_{ji}$$

für Matrizen $X = (x_{ij})$, $Y = (y_{ij})$ so dass XY quadratisch ist.

Sei nun $A = UDV^\dagger$ die Singulärwertzerlegung. Dann gilt also wegen des ersten Resultats

$$\|A\|_F = \|U^\dagger AV\|_F = \|D\|_F = \sqrt{\sigma_1^2 + \cdots + \sigma_p^2}.$$

□

In vielen "data science" Anwendungen hat man nun folgendes Problem: Gegeben eine Matrix $A \in M(m \times n, \mathbb{K})$, und eine Zahl $k \le p := \min\{m, n\}$, finde eine Matrix $A_k \in M(m \times n, \mathbb{K})$ vom Rang höchstens k, so dass

$$\|A - A_k\|_F$$

minimal ist.

Theorem 2.4 (Eckart–Young–Mirsky). *Sei $A = UDV^\dagger$ die Singulärwertzerlegung von A und seien die Singulärwerte $\sigma_1 \ge \sigma_2 \ge \cdots$ wie oben der Grösse nach sortiert. Seien $u_1, u_2, \dots, u_m \in \mathbb{K}^m$ die Spalten von U und $v_1, \dots, v_n \in \mathbb{K}^n$ die Spalten von V und seien*

$$U_k = \begin{pmatrix} u_1 & \cdots & u_k \end{pmatrix} \in M(m \times k, \mathbb{K}) \qquad V_k = \begin{pmatrix} v_1 & \cdots & v_k \end{pmatrix} \in M(n \times k, \mathbb{K})$$

die Matrizen bestehend aus ersten k Spalten von U bzw. V. Sei

$$A_k := U_k D_k V_k^\dagger = \sum_{j=1}^k \sigma_j u_j v_j^\dagger \tag{5}$$

mit

$$D_k = \begin{pmatrix} \sigma_1 & 0 & \dots & 0 \\ 0 & \sigma_2 & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & & \sigma_k \end{pmatrix} \in M(k \times k, \mathbb{R}).$$

Dann gilt für jede Matrix $B \in M(m \times n, \mathbb{K})$ vom Rang höchstens k, dass

$$\|A - B\|_F \ge \|A - A_k\|_F = \sqrt{\sum_{j=k+1}^r \sigma_j^2}.$$

Also ist A_k eine Lösung des obigen Optimierungsproblems.

Bemerkung 2.5. Tatsächlich kann man hier auch die Frobeniusnorm durch die Spektralnorm ersetzen, man hat also für das gleiche A_k und alle B vom Rang $\le k$ auch

$$\|A - B\|_2 \ge \|A - A_k\|_2 = \sigma_{k+1}. \tag{6}$$

Beweis von der Bemerkung, ohne Benutzung des Theorems. Wir nehmen an, dass $k < \min\{m,n\}$, ansonsten ist natürlich $A_k = A$ die eindeutige Lösung des Optimierungsproblems. Da $\mathrm{rank}(B) \le k$ gilt $\dim(\mathrm{Ker}B) \ge n - k$. Also existiert (wegen der Dimensionsformel) ein

$$0 \ne w = \lambda_1 v_1 + \cdots + \lambda_{k+1} v_{k+1} \in \mathrm{span}(v_1, \dots, v_{k+1}) \cap \mathrm{Ker}B,$$

und wir können o.B.d.A. annehmen, dass $\|w\| = 1$. Dann gilt

$$\begin{aligned}\|A - B\|_2^2 &\ge \|(A-B)w\|^2 = \|Aw\|^2 = \|UDV^\dagger w\|^2 = \|\sum_{j=1}^{k+1} \sigma_j \lambda_j u_j\|^2 \\ &= \sum_{j=1}^{k+1} \sigma_j^2 |\lambda_j|^2 \ge \sigma_{k+1}^2 \sum_{j=1}^{k+1} |\lambda_j|^2 = \sigma_{k+1}^2.\end{aligned}$$

Also gilt (6) und die Bemerkung ist bewiesen. □

Als Korollar erhalten wir folgende nützliche Eigenschaft der Singulärwerte, die die Dreiecksungleichung für die Spektralnorm verallgemeinert. Wir schreiben dabei allgemein $\sigma_j(A)$ für den j-ten Singulärwert von einer Matrix A, wobei die Singulärwerte der Grösse nach (absteigend) sortiert werden wie oben.

Korollar 2.6 (Korollar von der Bemerkung)**.** *Seien $A, B \in M(m \times n, \mathbb{K})$ und $i, j = 1, 2, \dots$ beliebig. Dann gilt*

$$\sigma_{i+j-1}(A + B) \le \sigma_i(A) + \sigma_j(B).$$

Falls hierbei ein Index grösser ist als $\min\{m,n\}$ so definiert man den entsprechenden Singulärwert als 0.

Beweis. Sei $C = A + B$ und C_{i+j-2}, A_{i-1} und B_{j-1} die optimalen Rang $i + j - 2$, $i - 1$ bzw. $j - 1$ Approximationen von C, A und B wie im Theorem, also wie in (5). (NB: Bis jetzt wissen wir nur, dass diese Approximationen optimal in der 2-Norm sind!) Dann gilt

$$\begin{aligned}\sigma_{i+j-1}(C) &= \sigma_1(C - C_{i+j-2}) \overset{\text{Bemerkung}}{\le} \sigma_1(C - A_{i-1} - B_{j-1}) \\ &= \sigma_1(A - A_{i-1} + B - B_{j-1}) \le \sigma_1(A - A_{i-1}) + \sigma_1(B - B_{j-1}) = \sigma_i(A) + \sigma_j(B)\end{aligned}$$

Die erste Ungleichung ist dabei die Bemerkung ($A_{i-1} + B_{j-1}$ hat Rang $\le i + j - 2$), die zweite Ungleichung ist die Dreiecksungleichung für die Spektralnorm. □

Damit können wir schliesslich das Theorem zeigen.

Beweis vom Theorem. Sei nun B wie im Theorem vom Rang $\le k$. Dann gilt $\sigma_j(B) = 0$ für $j \ge k + 1$. Damit gilt dann mit dem Korollar

$$\sigma_{i+k}(A) = \sigma_{i+k}(A - B + B) \le \sigma_i(A - B) + \sigma_{k+1}(B) = \sigma_i(A - B).$$

Wir haben also, mit Lemma 2.3,

$$\|A - B\|_F^2 = \sum_{j=1}^{p} (\sigma_j(A - B))^2 \ge \sum_{j=1}^{p} (\sigma_{j+k}(A))^2 = \sum_{j=k+1}^{p} (\sigma_j(A))^2 = \|A - A_k\|_F.$$

□

Bemerkung 2.7. Wir haben hier mehrfach benutzt, dass die Singulärwerte von $A - A_k = \sum_{j=k+1}^{p} \sigma_j u_j v_j^\dagger$ gerade

$$\sigma_{k+1}, \sigma_{k+2}, \dots$$

sind, wenn $\sigma_1 \ge \sigma_2 \ge \cdots$ die Singulärwerte von A sind. (ÜA)

2.3. **Approximation durch orthogonale Matrizen.**

Satz 2.8. *Sei $A \in M(n \times n, \mathbb{R})$ eine quadratische reelle Matrix und $A = UDV^T$ die Singulärwertzerlegung. Dann minimiert $R_0 = UV^T$ den Abstand*

$$\|A - R\|_F$$

unter allen orthogonalen Matrizen $R \in O(n)$.

Proof. Zunächst gilt nach Lemma 2.3

$$\|A - R\|_F = \|U^T(A - R)V\|_F = \|D - \underbrace{U^T RV}_{\in O(n)}\|_F.$$

Wir können also direkt annehmen, dass $A = D$ diagonal ist mit nicht ngeativen Diagonaleinträgen. Unser Ziel ist dann zu zeigen, dass $R_0 = E_n$ das Optimierungsproblem löst, also dass für alle $R \in O(n)$ gilt

$$\|D - R\|_F \geq \|D - E_n\|_F.$$

Es gilt

$$\|D - R\|_F^2 = \mathrm{tr}((D - R)^T(D - R)) = \mathrm{tr}(D^2) - 2\mathrm{tr}(DR) + \mathrm{tr}(\underbrace{R^T R}_{=E_n}) = \|D\|_F^2 - 2\mathrm{tr}(DR) + n.$$

Dieser Ausdruck wird also durch solche R minimiert, die $\mathrm{tr}(DR)$ maximieren. Sei $R = (r_{ij})$. Die Einträge einer orthogonalen Matrix sind immer ≤ 1, denn

$$r_{ij} \leq \sqrt{\sum_{i=1}^{n} r_{ij}^2} = 1.$$

Also gilt, da $\sigma_j \geq 0$:

$$\mathrm{tr}(DR) = \sum_{j=1}^{n} r_{jj}\sigma_j \leq \sum_{j=1}^{n} \sigma_j = \mathrm{tr}(D).$$

Eingesetzt in die Gleichung darüber ergibt dies die gewünschte Ungleichung

$$\|D - R\|_F^2 \geq \|D\|_F^2 - 2\mathrm{tr}(D) + n = \|D - E_n\|_F^2.$$

□