Bonner Zentrum für Lehrerbildung (BZL)

Singulärwertzerlegung. Theorie und Anwendung

Bachelorarbeit im Fach: Mathematik

vorgelegt von Noah Pferdekamp
Matrikelnummer: 123456

Erstbetreuer: Prof. Dr. Philipp Hieronymi
Zweitgutachter: Prof. Dr. Thoralf Räsch
Mathematisches Institut

Wintersemester 2024/2025
Bonn, den 31. Januar 2025

Selbstständigkeitserklärung

Ich versichere hiermit, dass die Bachelorarbeit mit dem Titel „Singulärwertzerlegung. Theorie und Anwendung“ von mir selbst und ohne jede unerlaubte Hilfe selbstständig angefertigt wurde, dass sie noch an keiner anderen Hochschule zur Prüfung vorgelegen hat und dass sie weder ganz noch in Auszügen veröffentlicht worden ist. Die Stellen der Arbeit — einschließlich Tabellen, Karten, Abbildungen usw. —, die anderen Werken dem Wortlaut oder dem Sinn nach entnommen sind, habe ich in jedem einzelnen Fall kenntlich gemacht.

Bonn, den 31. Januar 2025
Signatur

Inhaltsverzeichnis

Abbildungsverzeichnis

Tabellenverzeichnis

Notation

$\overline{z}$	Komplexe Konjugation
$\mathfrak{Re}$	Realteil
$\mathfrak{Im}$	Imaginärteil
$\mathcal{L}$	Lineare Hülle
$\lVert v \rVert$	Norm
$\langle v, w \rangle$	Skalarprodukt
$\mathbf{0}$	Nullmatrix
$\operatorname{diag}(\lambda_1, \dots, \lambda_n)$	Diagonaleinträge
$\operatorname{rg}(X)$	Rang
$\operatorname{df}(X)$	Defekt
$\boldsymbol{I}$	Einheitsmatrix

1 Einleitung

Mathematische Theorie 2

In diesem Kapitel wird zunächst unter Verwendung vorher eingeführter Sätze und Definitionen die Existenz und eine fundamentale Eigenschaft der Singulärwertzerlegung formal bewiesen. Anschließend wird an einem konkreten Beispiel die Berechnung durchgeführt und die SVD visualisiert. Um das Kapitel abzuschließen, erfolgt eine Beschreibung der wichtigsten beiden Varianten der Singulärwertzerlegung.

2.1. Beweisführung

Es wird davon ausgegangen, dass der Leser[1] mit den Grundlagen der linearen Algebra vertraut ist, insbesondere mit Matrizen und ihren Eigenschaften. Bekannte Definitionen werden nicht erneut aufgeführt, die einzige Ausnahme bildet Definition 2.1, da diese für jegliche Beweisführung und für das Verständnis in diesem Kapitel unerlässlich ist und deswegen eine Auffrischung sinnvoll erscheint.

Definition 2.1.
Sei $n \in \mathbb{N}$ und $A \in \mathbb{R}^{n \times n}$. Für

$$Av = \lambda v$$

heißen die Lösungen $\mathbb{R}^n \ni v \neq 0$ *Eigenvektoren* und die zugehörigen λ *Eigenwerte*.

Vorausgesetzte Sätze werden ohne weiteren Beweis verwendet, aber in wichtigen Fällen dennoch vor Verwendung kurz rekapituliert, wie in Wiederholung 2.2 verdeutlicht.

[1] Aus Gründen der besseren Lesbarkeit wird das generische Maskulinum verwendet, wobei alle Geschlechter mit eingeschlossen sind.

Wiederholung 2.2 (Basisergänzungssatz).
Sei V ein beliebiger Vektorraum, $L \subseteq V$ linear unabhängig und $E \subseteq V$ ein Erzeugendensystem von V. Dann kann L durch Elemente aus E zu einer Basis von V ergänzt werden.

Um die Existenz der Singulärwertzerlegung für beliebige Matrizen zu beweisen, bedarf es der Hilfe eines anderen Satzes, des sogenannten Spektralsatzes. Dieser hat keine eindeutige Ausführung, sondern beschreibt vielmehr mehrere verwandte Aussagen der Mathematik, wobei sich in dieser Arbeit auf seine Folgerungen für symmetrische Matrizen beschränkt wird. Es ist ebenfalls wichtig zu betonen, dass der Spektralsatz zwar hier „nur“ für den Beweis der Singulärwertzerlegung verwendet wird, eine Bezeichnung als Hilfssatz jedoch irreführend wäre, da der Satz für sich genommen bereits eine bedeutende Aussage der linearen Algebra und Funktionalanalysis darstellt. Um den Spektralsatz für symmetrische Matrizen einführen und anschließend beweisen zu können, benötigen wir zunächst Lemma 2.3 und Wiederholung 2.4.

Lemma 2.3.
Sei $z \in \mathbb{C}$ und $a, b \in \mathbb{R}$ mit $z = a + bi$. $z = \overline{z}$ gilt genau dann, wenn $z \in \mathbb{R}$.

Beweis. „$\Rightarrow$“ Durch $z = \overline{z}$ gilt

$$\begin{aligned} a + bi &= a - bi \\ \Leftrightarrow \quad 2bi &= 0. \end{aligned}$$

Da $i \neq 0$ muss $b = 0$, womit $\mathfrak{Im}(z) = 0$. Also ist $z = \mathfrak{Re}(z) \in \mathbb{R}$.
„$\Leftarrow$“ Folgt direkt aus der Definition. □

Wiederholung 2.4 (Gram-Schmidtsches Orthonormalisierungsverfahren).
Sei V ein euklidischer Vektorraum und $\{u_1, \dots, u_n\}$ eine Menge von linear unabhängigen Vektoren in V. Dann kann eine Menge $\{v_1, \dots, v_n\}$ aus Vektoren in V konstruiert werden, sodass $\{v_1, \dots, v_n\}$ orthonormal ist und

$$\mathcal{L}\{v_1, \dots, v_n\} = \mathcal{L}\{u_1, \dots, u_n\}.$$

Satz 2.5 (Spektralsatz).
Sei $n \in \mathbb{N}$ und $A \in \mathbb{R}^{n \times n}$ quadratisch und symmetrisch. Dann gilt:

(i) A hat reelle Eigenwerte.

(ii) Es existiert eine orthogonale Matrix $R \in \mathbb{R}^{n \times n}$, sodass $R^{-1}AR = \Lambda$ diagonal ist.

Beweis. Die Behauptungen werden nacheinander bewiesen (vgl. [Cra22]).

Zu (i): Sei $\lambda \in \mathbb{C}$ ein Eigenwert von A mit zugehörigem Eigenvektor $v \in \mathbb{C}^n$. Dann ist mit Definition 2.1

$$\begin{aligned} & Av = \lambda v \\ \Leftrightarrow \quad & A\overline{v} = \overline{\lambda}\overline{v}, \end{aligned} \tag{2.1}$$

da $A \in \mathbb{R}$ und somit $A = \overline{A}$ nach Lemma 2.3. Nun gilt zum einen

$$(Av)^T\overline{v} = (\lambda v)^T\overline{v} = \lambda v^T\overline{v} \tag{2.2}$$

und zum anderen

$$(Av)^T\overline{v} = v^TA^T\overline{v} \overset{A \text{ sym.}}{=} v^TA\overline{v} \overset{(2.1)}{=} v^T\overline{\lambda}\overline{v} = \overline{\lambda}v^T\overline{v}. \tag{2.3}$$

Mit (2.2)=(2.3) ergibt sich

$$\lambda v^T\overline{v} = \overline{\lambda}v^T\overline{v}.$$

Da $v \neq 0$ ist erhalten wir

$$\lambda = \overline{\lambda}.$$

Nach Lemma 2.3 ist dann $\lambda \in \mathbb{R}$, wodurch auch $v \in \mathbb{R}^n$ sein muss. □

Zu (ii): Induktion über $n \in \mathbb{N}$:
Induktionsanfang. Für $n = 1$ sind A und R Skalare. Setze $R = 1$. Damit ist R orthogonal, da $R^{-1} = R^T$ und $\mathbb{R} \ni A = R^{-1}AR$ trivialerweise diagonal.
Induktionshypothese. Die Behauptung (ii) gelte für festes, beliebiges $\mathbb{N} \ni n - 1$. Es soll gezeigt werden, dass sie dann auch für n gilt.
Induktionsschritt. Sei λ_1 ein beliebiger Eigenwert von A mit zugehörigem normalisiertem Eigenvektor v_1, also $\|v_1\| = 1$. Nach (i) gilt $\lambda_1 \in \mathbb{R}$ und $v_1 \in \mathbb{R}^n$. Mit dem Basisergänzungssatz (Wiederholung 2.2) kann v_1 durch Vektoren $u_2, \dots, u_n$ zu einer Basis von $\mathbb{R}^n$ ergänzt werden. Nun kann das Gram-Schmidt'sche Orthonormalisierungsverfahren (Wiederholung 2.4) angewendet

werden, wodurch eine orthonormale Basis $\{v_1, \dots, v_n\}$ von $\mathbb{R}^n$ konstruiert wird. Der Leser wird daran erinnert, dass orthonormale Vektoren normiert und orthogonal sind. Sei

$$P = \begin{bmatrix} | & | & | & | \\ v_1 & v_2 & \cdots & v_n \\ | & | & | & | \end{bmatrix} \in \mathbb{R}^{n \times n}$$

mit $v_1, \dots, v_n$ als Spaltenvektoren und setze $\mathbb{R}^{n\times n} \ni B = P^{-1}AP = P^TAP$.

Das Ziel ist, die Induktionshypothese auf eine symmetrische Untermatrix $C \in \mathbb{R}^{(n-1)\times(n-1)}$ von B anzuwenden.

Schritt 1. Dafür wird zunächst die Symmetrie von B gezeigt:

$$B^T = \left(P^TAP\right)^T = (AP)^TP = P^TA^TP \overset{A \text{ sym.}}{=} P^TAP = B.$$

Schritt 2. Betrachte jetzt die erste Spalte von B. Die erste Spalte einer beliebigen Matrix erhält man durch Multiplikation mit dem kanonischen Einheitsvektor $\mathbf{e}_1$:

$$\begin{aligned}
B\mathbf{e}_1 &= P^TAP\mathbf{e}_1 \\
&= P^TAv_1 && \big(v_1 \text{ ist die erste Spalte von P}\big) \\
&= P^T\lambda_1 v_1 && \big(\lambda_1 \text{ ist der Eigenwert zu } v_1\big) \\
&= P^Tv_1\lambda_1 \\
&= \begin{bmatrix} - \, v_1 \, - \\ - \, v_2 \, - \\ \vdots \\ - \, v_n \, - \end{bmatrix} v_1\lambda_1 \\
&= \begin{bmatrix} \langle v_1, v_1\rangle \\ \langle v_2, v_1\rangle \\ \cdots \\ \langle v_n, v_1\rangle \end{bmatrix} \lambda_1 \\
&= \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \lambda_1. && \big(\|v_1\| = 1 \text{ und bel. } v_i, v_j \in \{v_1, \dots, v_n\} \text{ orthogonal}\big)
\end{aligned}$$

Mit der Darstellung als Blockmatrix und durch Symmetrie von B gilt somit

$B = \begin{bmatrix} \lambda_1 & \mathbf{0} \\ \mathbf{0} & C \end{bmatrix}$ mit $C \in \mathbb{R}^{(n-1)\times(n-1)}$ symmetrisch.

Nach Induktionshypothese gibt es ein orthogonales $Q \in \mathbb{R}^{(n-1)\times(n-1)}$ mit $Q^T C Q = D$ diagonal. Damit gilt

$$\begin{aligned} P^T A P &= B \\ &= \begin{bmatrix} \lambda_1 & \mathbf{0} \\ \mathbf{0} & C \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1 & \mathbf{0} \\ \mathbf{0} & Q D Q^T \end{bmatrix} \\ &= \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & Q \end{bmatrix} \begin{bmatrix} \lambda_1 & \mathbf{0} \\ \mathbf{0} & D \end{bmatrix} \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & Q^T \end{bmatrix}. \end{aligned}$$

Also ist

$$\begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & Q^T \end{bmatrix} P^T A P \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & Q \end{bmatrix} = \begin{bmatrix} \lambda_1 & \mathbf{0} \\ \mathbf{0} & D \end{bmatrix}.$$

Definiere

$$R = P \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & Q \end{bmatrix}.$$

Es gilt

$$R^T = \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & Q^T \end{bmatrix} P^T = \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & Q^{-1} \end{bmatrix} P^{-1} = R^{-1}.$$

Dementsprechend ist R orthogonal und $R^{-1} A R = R^T A R = \begin{bmatrix} \lambda_1 & \mathbf{0} \\ \mathbf{0} & D \end{bmatrix} = \Lambda$ diagonal, da D diagonal ist. □

Korollar 2.6.
Sei $n \in \mathbb{N}$ und $A \in \mathbb{R}^{n\times n}$ quadratisch und symmetrisch. Dann gilt:
Es existiert eine orthogonale Matrix $R \in \mathbb{R}^{n\times n}$, sodass $R^{-1} A R = \Lambda$ diagonal ist, wobei die Diagonalwerte von $\Lambda \in \mathbb{R}^{n\times n}$ die Eigenwerte von A und die Spalten von R die normierten Eigenvektoren von A sind.

Beweis. Nach dem Spektralsatz (Satz 2.5) gibt es ein orthogonales $R = [v_1 \dots v_n]$ mit $v_1, \dots, v_n \in \mathbb{R}^n$ und $\Lambda = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$, sodass

$$R^{-1} A R = \Lambda.$$

Dann ist
$$AR = R\Lambda,$$
oder spaltenweise
$$Av_i = \lambda_i v_i, \quad \text{für } i = 1, \dots, n.$$
Da R orthogonal ist, sind nach Definition die Spaltenvektoren von R, also $v_1, \dots, v_n$, orthonormal. Für beliebiges v_i gilt somit $\|v_i\| = 1$, wodurch $v_i \neq 0$ sein muss. Mit Definition 2.1 sind also $v_1, \dots, v_n$ die (normierten) Eigenvektoren von A mit zugehörigen Eigenwerten $\lambda_1, \dots, \lambda_n$.

Hinweis. Mithilfe von Zeilen- und Spaltenvertauschungen innerhalb von R und Λ kann Λ so geordnet werden, dass $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$. Diese Sortierung wird für den Rest des Kapitels angenommen. □

BEMERKUNG 2.7.
Durch Korollar 2.6 lässt sich direkt die nützliche Aussage treffen, dass bei einer symmetrischen Matrix Eigenvektoren zu verschiedenen Eigenwerten orthogonal zueinander sind.

Nun kann mithilfe der vorangegangenen Sätze die Existenz der Singulärwertzerlegung für beliebige reelle Matrizen bewiesen werden. Der Beweis orientiert sich dabei an [Che20].

SATZ 2.8 (Singulärwertzerlegung).
Sei $m, n \in \mathbb{N}$ und $X \in \mathbb{R}^{m \times n}$.
Dann existieren orthogonale Matrizen $U \in \mathbb{R}^{m \times m}, V \in \mathbb{R}^{n \times n}$ und eine Diagonalmatrix $\Sigma \in \mathbb{R}^{m \times n}$, sodass
$$X = U \Sigma V^T.$$

Beweis. Sei $C = X^T X \in \mathbb{R}^{n \times n}$ und $r = \mathrm{rg}(X) \leq \min(m, n)$. Dann ist C symmetrisch und positiv semidefinit (alle Eigenwerte sind positiv oder gleich 0 bei symmetrischen Matrizen). Nach Korollar 2.6 gibt es ein orthogonales
$$\mathbf{V} = [v_1 \dots v_n] \in \mathbb{R}^{n \times n}$$
und diagonales $\mathbb{R}^{n \times n} \ni \Lambda = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$ mit $\lambda_1 \geq \cdots \geq \lambda_r > 0 = \lambda_{r+1} = \cdots = \lambda_n$, sodass $C = V \Lambda V^T$.

Definiere $\sigma_i := \sqrt{\lambda_i}$ für $i = 1, \dots, r$ und

$$\Sigma_{i,j} = \begin{cases} \sigma_i, & i = j. \\ 0, & i \neq j. \end{cases}$$
$$\Leftrightarrow \quad \mathbf{\Sigma} = \begin{bmatrix} \operatorname{diag}(\sigma_i, \dots, \sigma_r) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \in \mathbb{R}^{m \times n}$$

als Blockmatrix. Definiere außerdem

$$u_i = \frac{1}{\sigma_i} X v_i \in \mathbb{R}^m, \quad \text{für } 1 \leq i \leq r.$$

Dann sind $u_1, \dots, u_r$ orthornomal:

$$\begin{aligned} u_i^T u_j &= \left(\frac{1}{\sigma_i} X v_i\right)^T \left(\frac{1}{\sigma_i} X v_j\right) \\ &= \frac{1}{\sigma_i \sigma_j} v_i^T \underbrace{X^T X}_{=C} v_j \\ &= \frac{1}{\sigma_i \sigma_j} v_i^T (\lambda_j v_j) && (\lambda_j \text{ ist Eigenwert zu } v_j) \\ &= \frac{\sigma_j}{\sigma_i} v_i^T v_j && (\lambda_j = \sigma_j^2) \\ &= \begin{cases} 1, & i = j. \\ 0, & i \neq j. \end{cases} && (v_i, v_j \text{ orthonormal}) \end{aligned}$$

Wie bereits im Beweis des Spektralsatzes können $u_1, \dots, u_r$ mithilfe des Basisergänzungssatzes (Wiederholung 2.2) und des Gram-Schmidt-Verfahrens (Wiederholung 2.4) durch Vektoren $u_{r+1} \dots, u_m \in \mathbb{R}^m$ zu einer orthonormalen Basis von $\mathbb{R}^m$ ergänzt werden. Damit ist

$$\mathbf{U} = [u_1 \dots u_r u_{r+1} \dots u_m] \in \mathbb{R}^{m \times m}$$

orthogonal. Es bleibt zu zeigen, dass $XV = U\Sigma$ ist, also:

$$X[v_1 \dots v_r v_{r+1} \dots v_n] = [u_1 \dots u_r u_{r+1} \dots u_m] \begin{bmatrix} \sigma_1 & & & & \\ & \sigma_2 & & & \mathbf{0} \\ & & \ddots & & \\ & & & \sigma_r & \\ & \mathbf{0} & & & \mathbf{0} \end{bmatrix}.$$

Für $1 \leq i \leq r$ gilt $Xv_i = u_i\sigma_i$ nach Konstruktion.
Für $i > r$ soll gezeigt werden, dass $Xv_i = 0u_i = \mathbf{0}$ ist. Betrachte dafür

$$X^T X v_i = C v_i \overset{(*)}{=} 0 v_i = \mathbf{0}$$

(∗) *Der zugehörige Eigenwert zum Eigenvektor v_i ist 0 für $i > r$.*

Damit muss wie erwünscht $Xv_i = \mathbf{0}$ gelten, oder $X^T = \mathbf{0}$, wodurch ebenfalls $Xv_i = \mathbf{0}$ folgt. Dementsprechend ist $X = U\Sigma V^T$ und die Aussage ist bewiesen.
□

BEMERKUNG 2.9.

- Die Diagonalwerte von Σ heißen **Singulärwerte** von X und werden meist absteigend sortiert.
- Die Spalten von U heißen **linke Singulärvektoren** von X.
- Die Spalten von V heißen **rechte Singulärvektoren** von X.

Nachdem der zentrale Beweis dieses Kapitels geführt und die Existenz der Singulärwertzerlegung für beliebige reelle Matrizen gezeigt wurde, wird nun auf eine wichtige Eigenschaft der SVD eingegangen.

Dafür werden zunächst die vier fundamentalen Unterräume zu einer Matrix nach [Str03, S. 185] definiert:

DEFINITION 2.10.
Sei $A \in \mathbb{R}^{m \times n}$. Definiere folgende Unterräume zu A:

- *Spaltenraum:* $\operatorname{Bild}(A) = \{b \in \mathbb{R}^m \mid \exists x \in \mathbb{R}^n, Ax = b\}$.
- *Zeilenraum:* $\operatorname{Bild}(A^T) = \{z \in \mathbb{R}^n \mid \exists y \in \mathbb{R}^m, A^T y = z\}$.
- *Kern/Nullraum:* $\operatorname{Kern}(A) = \{x \in \mathbb{R}^n \mid Ax = \mathbf{0}\}$.
- *Linkskern:* $\operatorname{Kern}(A^T) = \{y \in \mathbb{R}^m \mid A^T y = \mathbf{0}\}$.

Die vier Unterräume geben umfangreichen Aufschluss über die Wirkung einer Matrix auf verschiedene Vektoren und stehen dabei in Verbindung mit zahlreichen Themen der linearen Algebra, wie beispielsweise dem Lösen von Gleichungssystemen.

Die Art der Beziehung zwischen der Singulärwertzerlegung und den vier fundamentalen Unterräumen wird in Korollar 2.12 zusammengefasst. Der

Beweis wird nach [Joh21, S. 214 f.] geführt, vor der Beweisführung wird allerdings Wiederholung 2.11 benötigt.

Wiederholung 2.11 (Dimensionssatz).
Sei $A \in \mathbb{R}^{m \times n}$. Dann gilt:

$$\mathrm{df}(A) + \mathrm{rg}(A) = n.$$

Korollar 2.12.
Sei $m, n \in \mathbb{N}, X \in \mathbb{R}^{m \times n}$ und $r = rg(X)$.
Dann existieren orthogonale Matrizen $U \in \mathbb{R}^{m \times m}, V \in \mathbb{R}^{n \times n}$ und eine Diagonalmatrix $\Sigma \in \mathbb{R}^{m \times n}$, sodass

$$X = U \Sigma V^T$$

und es gilt:

- Die ersten r Spalten von U sind eine Basis des Spaltenraums von X.
- Die letzten $m - r$ Spalten von U sind eine Basis des Linkskerns von X.
- Die ersten r Spalten von V sind eine Basis des Zeilenraums von X.
- Die letzten $n - r$ Spalten von V sind eine Basis des Kerns von X.

Beweis. Mit der Singulärwertzerlegung (Satz 2.8) erhalten wir orthogonales

$$U = [u_1 \dots u_r u_{r+1} \dots u_m] \in \mathbb{R}^{m \times m},$$
$$V = [v_1 \dots v_r v_{r+1} \dots v_n] \in \mathbb{R}^{n \times n}$$

und diagonales

$$\Sigma = \begin{bmatrix} \mathrm{diag}(\sigma_i, \dots, \sigma_r) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \in \mathbb{R}^{m \times n},$$

sodass $X = U \Sigma V^T$. Sei nun $i \in \{1, \dots, n\}$ und betrachte

$$X v_i = U \Sigma V^T v_i \overset{(**)}{=} U \Sigma \mathbf{e}_i = U \sigma_i \mathbf{e}_i = \sigma_i U \mathbf{e}_i = \sigma_i u_i.$$

(∗∗) *Für $i, j \in \{1, \dots, n\}$ sind v_i, v_j orthonormal.*

Fall 1: $1 \leq i \leq r$.
Damit ist $\sigma_i > 0$ und

$$X\frac{v_i}{\sigma_i} = u_i.$$

Nach Definition 2.10 sind dann $u_1, \dots, u_r \in \text{Bild}(X)$. Nun ist $\dim(\text{Bild}(X)) = \text{rg}(X) = r$ und da $\mathcal{B}_S = \{u_1, \dots, u_r\}$ genau r orthonormale Vektoren enthält, bildet $\mathcal{B}_S$ eine Basis vom $\text{Bild}(X)$, also vom Spaltenraum.

Fall 2: $i \geq r+1$.
Damit ist $\sigma_i = 0$ und

$$Xv_i = \mathbf{0}.$$

Nach Definition 2.10 sind dann $v_{r+1}, \dots, v_n \in \text{Kern}(X)$. Durch Wiederholung 2.11 wissen wir, dass $\dim(\text{Kern}(X)) = \text{df}(X) = n - r$. Mit $\mathcal{B}_K = \{v_{r+1}, \dots, v_n\}$ haben wir $n - r$ orthonormale Vektoren gegeben, also bildet $\mathcal{B}_K$ eine Basis vom $\text{Kern}(X)$.

Die Beweise für die Basen des Linkskerns und des Zeilenraums werden analog gezeigt, indem

$$X^T u_i = V \Sigma U^T u_i$$

betrachtet wird. □

Damit ist die Beweisführung dieser Arbeit abgeschlossen und die Berechnung der Singulärwertzerlegung kann an einem Beispiel veranschaulicht und visualisiert werden.

2.2. Beispiel und Visualisierung

Beispiel 2.13. Sei

$$A = \begin{bmatrix} 1 & -1 & 3 \\ 3 & 1 & 1 \end{bmatrix} \in \mathbb{R}^{2\times 3}.$$

Wir wollen nun die SVD dieser Matrix finden. Dafür muss der Beweis der Singulärwertzerlegung (Satz 2.8) mithilfe unserer konkreten Werte schrittweise nachvollzogen werden. Zuerst wird also

$$A^T A = \begin{bmatrix} 10 & 2 & 6 \\ 2 & 2 & -2 \\ 6 & -2 & 10 \end{bmatrix} \in \mathbb{R}^{3\times 3}$$

bestimmt. Davon sollen die Eigenwerte mit zugehörigen normierten Eigenvektoren berechnet werden. Für die Eigenwerte muss zunächst das charakteristische Polynom

$$\det(A - \lambda \boldsymbol{I}) = 0$$

gesetzt und die Lösungen λ_i für $i = 1, 2, 3$ gefunden werden. Auf die genaue Berechnung wird an dieser Stelle verzichtet, das Ergebnis lautet:

$$\lambda_1 = 16, \quad \lambda_2 = 6, \quad \lambda_3 = 0.$$

Mit $\sigma_j = \sqrt{\lambda_j}$ für $j = 1, 2$ (da $\operatorname{rg}(A) = 2$) erhalten wir

$$\Sigma = \begin{bmatrix} 4 & 0 & 0 \\ 0 & \sqrt{6} & 0 \end{bmatrix} \in \mathbb{R}^{2\times 3}.$$

Durch die Lösungen $\mathbb{R}^3 \ni v_i \neq 0$ von

$$(A - \lambda \boldsymbol{I})v = \mathbf{0}$$

ergeben sich die normierten Eigenvektoren

$$v_1 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{bmatrix}, \quad v_2 = \begin{bmatrix} -\frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix}, \quad v_3 = \begin{bmatrix} -\frac{1}{\sqrt{6}} \\ \frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \end{bmatrix}$$

und damit

$$V^T = \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \end{bmatrix} \in \mathbb{R}^{3\times 3}.$$

Es muss also nur noch $U \in \mathbb{R}^{2\times 2}$ bestimmt werden, welches spaltenweise durch

$$u_j = \frac{1}{\sigma_j} X v_j$$

ausgedrückt wird. Wir erhalten also

$$u_1 = \frac{1}{4} \begin{bmatrix} 1 & -1 & 3 \\ 3 & 1 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$$

und

$$u_2 = \frac{1}{\sqrt{6}} \begin{bmatrix} 1 & -1 & 3 \\ 3 & 1 & 1 \end{bmatrix} \begin{bmatrix} -\frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix}.$$

Dadurch ist

$$U = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}$$

und die Berechnung ist abgeschlossen mit

$$A = \begin{bmatrix} 1 & -1 & 3 \\ 3 & 1 & 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 4 & 0 & 0 \\ 0 & \sqrt{6} & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \end{bmatrix} = U\Sigma V^T.$$

Damit genauer verstanden wird, was genau durch die Singulärwertzerlegung geschieht, betrachten wir die Wirkung der Matrix A aus Beispiel 2.13 auf einen Vektor $v \in \mathbb{R}^3$, indem wir v und Av grafisch darstellen. Da dies an einem einzelnen Vektor schwer visualisiert werden kann, multiplizieren wir A mit allen Punkten auf der Einheitssphäre, also allen

$$v \in \left\{ \begin{bmatrix} \cos(u)\sin(v) \\ \sin(u)\sin(v) \\ \cos(v) \end{bmatrix} \in \mathbb{R}^3 \mid u \in [0, 2\pi],\ v \in [0, \pi] \right\}.$$

Das Ergebnis ist in Abbildung 2.1 dargestellt.

Um nachzuvollziehen, wie dieses Ergebnis zustande gekommen ist, verwen-

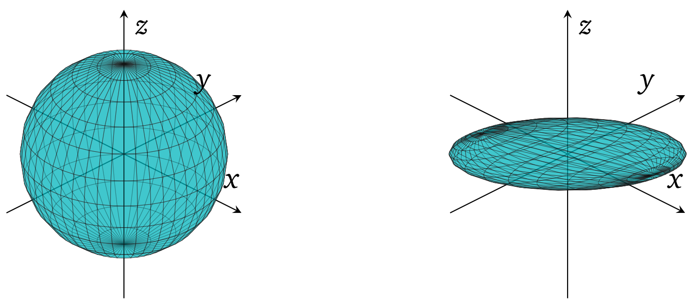

Abb. 2.1. Wirkung von A auf die Einheitssphäre

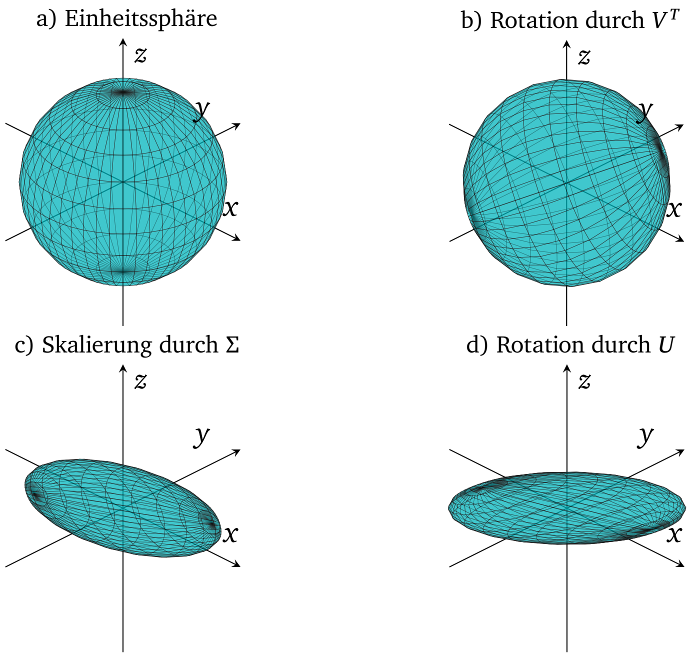

Abb. 2.2. Visualisierung der Singulärwertzerlegung

den wir die Singulärwertzerlegung und veranschaulichen die Zwischenschritte von $Av = U\Sigma V^T v$ anhand von einzelner Plots (siehe Abbildung 2.2).

Grundlage für diese Visualisierung ist das Wissen, dass im euklidischen Raum orthogonale Matrizen Drehungen und Diagonalmatrizen Skalierungen (entlang der Hauptachsen) darstellen. An dieser Stelle ist auch wichtig zu erwähnen, dass in Abbildung 2.2c und Abbildung 2.2d der zu beobachtende Wertebereich vergrößert wurde, also eine größere Streckung durch Σ erfolgt ist, als auf dem Plot zu sehen ist. Außerdem sei angemerkt, dass ab Abbildung 2.2c die Darstellung der z-Achse überflüssig und eigentlich nicht zu empfehlen ist, da sich dort durch die Dimensionsreduktion mittels Σ im zweidimensionalen Raum bewegt wird. Um eine Vergleichbarkeit der Plots zu verbessern, wird die Darstellung dennoch beibehalten.

Zusammenfassend zerlegt also die Singulärwertzerlegung eine Matrix in grundlegende geometrische Transformationen: Drehung, Skalierung und gegebenenfalls Dimensionsreduktion.

Mit dieser Erkenntnis wird sich dem letzten Abschnitt des theoretischen Teils zugewandt, in dem die wichtigsten beiden Varianten der SVD beschrieben werden.

2.3. Varianten der SVD

Test

3 Anwendung

Literaturverzeichnis

[Cra22] T. Crawford. *Oxford Linear Algebra: Spectral Theorem Proof*. 2022. URL: https://tomrocksmaths.com/2022/11/18/oxford-linear-algebra-spectral-theorem-proof/ (Stand: 02. 01. 2025).

[Che20] G. Chen. *Lecture 5: Singular Value Decomposition (SVD)*. 2020. URL: https://www.sjsu.edu/faculty/guangliang.chen/Math253S20/lec5svd.pdf (Stand: 20. 01. 2025).

[Str03] G. Strang. *Lineare Algebra*. Berlin, Heidelberg: Springer, 2003.

[Joh21] N. Johnston. *Advanced Linear and Matrix Algebra*. Cham: Springer, 2021.

A Code