

Miniaturen zur Einführung in die Mathematik

Vertiefungen, Ergänzungen und zusätzliche interessante Aspekte
für Hörsaalaneleitungen oder schlicht als Lektüreangebot

2019

Pseudoinverse Matrizen

Moore-Penrose-Inverse

Robert Labus, Universität Kassel, 2019

Stand 17. März 2019

Inhaltsverzeichnis

1	Pseudoinverse Matrizen	1
1.1	Motivation und Definition von verallgemeinerten Inversen	1
1.2	Die Moore-Penrose-Inverse oder Pseudoinverse	3
1.3	Berechnung von Pseudoinversen für besondere Fälle	11
1.4	Pseudoinverse und Projektionen	16
1.5	Anwendungen zur Pseudoinversen	19

Kapitel 1

Pseudoinverse Matrizen

1.1 Motivation und Definition von verallgemeinerten Inversen

Betrachten wir lineare Abbildungen wie z.B. Streckungen, Drehungen, Spiegelungen im \mathbb{R}^n oder Kombinationen solcher Abbildungen, dann stellt sich mitunter die Frage 'Gibt es Urbilder zu einem bestimmten Element der Zielmenge und wenn ja, gibt es genau ein Urbild oder gibt es mehrere und wie lassen sie sich berechnen?'. Die gerade genannten linearen Abbildungen sind typische Fälle von umkehrbaren linearen Abbildungen eines endlichdimensionalen Vektorraums in sich selbst, Beispiele für sogenannte bijektive Endomorphismen bzw. kurz Automorphismen. Und wegen der Bijektivität oder Umkehrbarkeit ist die gestellte Frage für solche Abbildungen leicht zu beantworten, es gibt für sie stets ein einseitig bestimmtes Urbild.

Wie läßt sich erkennen, ob ein Endomorphismus $f : V \rightarrow V$ eines endlichdimensionalen \mathbb{K} -Vektorraums V bijektiv, also umkehrbar ist?

Ist $\dim V = n$ für ein $n \in \mathbb{N}$, dann entspricht der Abbildung f genau eine Matrix $\underline{A}_f \in \mathbb{K}^{n \times n}$. Die Abbildung f ist bijektiv, wenn die Determinante von \underline{A}_f ungleich Null ist. Die Frage nach dem Urbild hat somit die Antwort, dass es im Fall $\det(\underline{A}_f) \neq 0$ ein Urbild gibt und zwar genau ein Urbild. Berechnen läßt sich das Urbild \underline{x} mit dem linearen Gleichungssystem

$$\underline{A}_f \underline{x} = \underline{b}$$

Wegen $\det(\underline{A}_f) \neq 0$ ist die Matrix \underline{A}_f regulär und besitzt eine inverse Matrix \underline{A}_f^{-1} . Mit dieser Inversen läßt sich nun auch das Urbild \underline{x} berechnen:

$$\begin{aligned} \underline{A}_f \underline{x} = \underline{b} & \Leftrightarrow \underline{A}_f^{-1}(\underline{A}_f \underline{x}) = \underline{A}_f^{-1} \underline{b} & (1.1) \\ & \Leftrightarrow (\underline{A}_f^{-1} \underline{A}_f) \underline{x} = \underline{A}_f^{-1} \underline{b} \\ & \Leftrightarrow \underline{I} \underline{x} = \underline{A}_f^{-1} \underline{b} \\ & \Leftrightarrow \underline{x} = \underline{A}_f^{-1} \underline{b} \end{aligned}$$

Im Ergebnis bedeutet dies, dass für eine reguläre Matrix $\underline{A} \in \mathbb{K}^{n \times n}$ und jedes beliebige $\underline{b} \in \mathbb{K}^n$ gilt:

$$\underline{A}\underline{A}^{-1}\underline{b} = \underline{b}$$

Insbesondere lassen sich für \underline{b} der Reihe nach die Spalten der Matrix \underline{A} einsetzen. Dies heißt dann zusammengefasst

$$\underline{A}\underline{A}^{-1}\underline{A} = \underline{A} \tag{1.2}$$

Was ist aber, wenn die Matrix $\underline{A} \in \mathbb{K}^{n \times n}$ nicht regulär sondern singular ist, nicht den maximalen Rang hat? Dann ist das Bild der zugehörigen linearen Abbildung f ein echter Unterraum von V . Das heißt, dass einige Vektoren aus V gar kein Urbild besitzen und andere unendlich viele. Ähnliches gilt, wenn die Matrix nicht quadratisch und die zugehörige Abbildung kein Endomorphismus ist.

Gibt es für diese Fälle von linearen Abbildungen $f : V \rightarrow W$, deren Matrizen singular oder nicht quadratisch sind, etwas Ähnliches wie die inverse Matrix? Zum Beispiel eine Matrix, die mit \underline{b} multipliziert wenigstens eine von unendlichen vielen Lösungen der Gleichung produziert oder im Fall der Unlösbarkeit wenigstens ein optimales Element $\underline{x} \in V$ für das $f(\underline{x})$ z.B. den kürzesten Abstand von \underline{b} hat? Kurz: Lässt sich der Begriff der Inversen einer Matrix auf beliebige Matrizen verallgemeinern?

Die Antwort ist, das so etwas möglich ist. Aber es ist nicht nur auf eine einzige Art möglich. Die entscheidende Eigenschaft der Inversen, die wir für die Verallgemeinerung nutzen ist die Gleichung (1.2).

Definition 1.1. *verallgemeinerte Inverse*

Für $\underline{A} \in \mathbb{K}^{m \times n}$ $m, n \in \mathbb{N}$ nennen wir eine Matrix $\underline{B} \in \mathbb{K}^{n \times m}$ eine verallgemeinerte Inverse, wenn sie der Gleichung

$$\underline{A}\underline{B}\underline{A} = \underline{A} \tag{1.3}$$

genügt.

Bemerkung 1.2.

(1) Im Fall $m = n$ mit einer regulären Matrix \underline{A} ist $\underline{B} = \underline{A}^{-1}$ die einzige verallgemeinerte Inverse.

Beweis:

Sei \underline{B} eine verallgemeinerte Inverse, dann gilt

$$\begin{aligned} \underline{A}\underline{B}\underline{A} = \underline{A} & \Leftrightarrow \underline{A}^{-1}\underline{A}\underline{B}\underline{A} = \underline{A}^{-1}\underline{A} \\ & \Leftrightarrow \underline{B}\underline{A} = \underline{E} \\ & \Leftrightarrow \underline{B}\underline{A}\underline{A}^{-1} = \underline{E}\underline{A}^{-1} \\ & \Leftrightarrow \underline{B} = \underline{A}^{-1} \quad \square \end{aligned}$$

(2) Im Allgemeinen ist die verallgemeinerte Inverse nicht eindeutig bestimmt, wie folgendes Beispiel zeigt.

Für die singuläre Matrix $\underline{A} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ erfüllen die Matrizen $\underline{B}_1 = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$ und $\underline{B}_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ die Bedingung (1.3). Es gilt $\underline{A}\underline{B}_1\underline{A} = \underline{A}$ und $\underline{A}\underline{B}_2\underline{A} = \underline{A}$.

Daher sind sowohl \underline{B}_1 als auch \underline{B}_2 verallgemeinerte Inverse von \underline{A} .

1.2 Die Moore-Penrose-Inverse oder Pseudoinverse

Wir werden in diesem Abschnitt die Bedingung (1.3) durch weitere Bedingungen ergänzen, die sicherstellen, dass es für jede beliebige $\underline{A} \in \mathbb{C}^{m \times n}$ genau eine verallgemeinerte Inverse \underline{A}^+ mit diesen Zusatzbedingungen existiert. Zur Vorbereitung benötigen wir den Begriff der adjungierten Matrix und den Satz über die Singulärwertzerlegung.

Definition 1.3. *adjungierte Matrix, hermitesche Matrix*

(1) Für $\underline{A} = (a_{i,j}) \in \mathbb{K}^{m \times n}$ $m, n \in \mathbb{N}$ nennen wir die Matrix $\underline{A}^* \in \mathbb{K}^{n \times m}$ mit

$$\underline{A}^* = (a_{i,j}^*) = (\overline{a_{j,i}})$$

die adjungierte Matrix zu \underline{A} .

(Der Querstrich bedeutet hier die konjugiert komplexe Zahl der übertrichenen Zahl.)

(2) Eine quadratische Matrix heißt hermitesch, wenn sie mit ihrer Adjungierten übereinstimmt.

$$\underline{A}^* = \underline{A}$$

Bemerkung 1.4.

Im Fall einer Matrix mit ausschließlich rein reellen Komponenten

- stimmt die adjungierte Matrix mit der transponierten Matrix überein.

- ist die Matrix genau dann hermitesch, wenn sie symmetrisch ist.

Also für $\underline{A} \in \mathbb{R}^{m \times n}$ gilt

$$\underline{A}^* = \underline{A}^T \quad \text{und} \quad \underline{A}^* = \underline{A} \Leftrightarrow \underline{A}^T = \underline{A}$$

Beispiel:

Zur Matrix $\underline{A} = \begin{bmatrix} 3 & 1+i & -2i \\ 2-i & 1 & 5 \end{bmatrix}$ lautet die adjungierte Matrix

$$\underline{A}^* = \begin{bmatrix} \overline{3} & \overline{2-i} \\ \overline{1+i} & \overline{1} \\ \overline{-2i} & \overline{5} \end{bmatrix} = \begin{bmatrix} 3 & 2+i \\ 1-i & 1 \\ 2i & 5 \end{bmatrix}$$

Satz 1.5. *Rechenregeln für adjungierte Matrizen*

Für $\underline{A} \in \mathbb{K}^{m \times n}$ $m, n \in \mathbb{N}$, $\underline{B} \in \mathbb{K}^{n \times m}$ und $z \in \mathbb{C}$ gilt:

$$(1) \quad (\underline{A}\underline{B})^* = (\underline{B}^*\underline{A}^*)$$

$$(2) \quad (z\underline{A})^* = \overline{z}\underline{A}^*$$

$$(3) \quad (\underline{A}^*)^* = \underline{A}$$

$$(4) \quad (\underline{A}^*)^{-1} = (\underline{A}^{-1})^* \quad \text{falls } \underline{A} \text{ regulär ist.}$$

Satz 1.6. *Singulärwertzerlegung (SVD – singular value decomposition) und reduzierte Singulärwertzerlegung (RSVD)*

Zu einer beliebigen Matrix $\underline{\underline{A}} \in \mathbb{C}^{m \times n}$ gibt es stets Matrizen

$\underline{\underline{U}} \in \mathbb{C}^{m \times m}, \underline{\underline{V}} \in \mathbb{C}^{n \times n}$ mit $\underline{\underline{U}}^* \underline{\underline{U}} = \underline{\underline{E}}_m$ bzw. $\underline{\underline{V}}^* \underline{\underline{V}} = \underline{\underline{E}}_n$ und eine Diagonalmatrix $\underline{\underline{\Sigma}} \in \mathbb{R}^{m \times n}$ mit nichtnegativen Komponenten, für die gilt:

$$\underline{\underline{A}} = \underline{\underline{U}} \underline{\underline{\Sigma}} \underline{\underline{V}}^*. \quad (\text{SVD})$$

Daraus läßt sich für $\text{Rang } \underline{\underline{A}} = r > 0$ eine kompaktere reduzierte Singulärwertzerlegung (RSVD) gewinnen, die den Rang berücksichtigt und eine quadratische reguläre Diagonalmatrix $\underline{\underline{\Sigma}}$ verwendet.

Es gibt Matrizen $\underline{\underline{U}} \in \mathbb{C}^{m \times r}, \underline{\underline{V}} \in \mathbb{C}^{n \times r}$ mit $\underline{\underline{U}}^* \underline{\underline{U}} = \underline{\underline{E}}_r$ bzw. $\underline{\underline{V}}^* \underline{\underline{V}} = \underline{\underline{E}}_r$ und eine Diagonalmatrix $\underline{\underline{\Sigma}} \in \mathbb{R}^{r \times r}$ mit nichtnegativen Komponenten, für die gilt:

$$\underline{\underline{A}} = \underline{\underline{U}} \underline{\underline{\Sigma}} \underline{\underline{V}}^*. \quad (\text{RSVD})$$

Beispiel 1.7

Wir bestimmen die SVD der Matrix

$$\underline{\underline{A}} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

Die $(m \times m)$ -Matrix $\underline{\underline{U}}$ bilden wir aus normierten Eigenvektoren der Matrix $\underline{\underline{A}}; \underline{\underline{A}}^*$ und die $(n \times n)$ -Matrix $\underline{\underline{V}}$ bilden wir aus normierten Eigenvektoren der Matrix $\underline{\underline{A}}^* \underline{\underline{A}}$.

Die $(m \times n)$ -Matrix $\underline{\underline{\Sigma}}$ enthält die Quadratwurzeln der entsprechenden Eigenwerte auf der Hauptdiagonalen.

Berechnung von $\underline{\underline{U}}$ und $\underline{\underline{\Sigma}}$:

$$\underline{\underline{A}} \underline{\underline{A}}^* = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

Das charakteristische Polynom lautet

$$\rho(\lambda) = \det(\underline{\underline{A}} \underline{\underline{A}}^* - \lambda \underline{\underline{E}}) = (2 - \lambda)^2 - 1 = \lambda^2 - 4\lambda + 3$$

Somit haben wir die Eigenwerte $\lambda_1 = 3, \lambda_2 = 1$

Damit erhalten wir die Matrix

$$\underline{\underline{\Sigma}} = \begin{bmatrix} \sqrt{3} & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

Die Bestimmung von normierten Eigenvektoren ergibt

$$\underline{\underline{u}}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \underline{\underline{u}}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Damit erhalten wir die Matrix

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

Berechnung von $\underline{\underline{V}}$:

$$\underline{\underline{A}}^* \underline{\underline{A}} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 1 \end{bmatrix}$$

hat das charakteristische Polynom $p(\lambda) = -\lambda(\lambda-1)(\lambda-3)$ und damit die Eigenwerte $\lambda_1 = 3, \lambda_2 = 1, \lambda_3 = 0$

Die Bestimmung von normierten Eigenvektoren ergibt

$$\underline{v}_1 = \frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \quad \underline{v}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \quad \underline{v}_3 = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$$

Damit erhalten wir die Matrix

$$\underline{\underline{V}} = \frac{1}{\sqrt{6}} \begin{bmatrix} 1 & \sqrt{3} & \sqrt{2} \\ 2 & 0 & -\sqrt{2} \\ 1 & -\sqrt{3} & \sqrt{2} \end{bmatrix}$$

Es folgt

$$\underline{\underline{A}} = \frac{1}{\sqrt{12}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} \sqrt{3} & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 & 1 \\ \sqrt{3} & 0 & -\sqrt{3} \\ \sqrt{2} & -\sqrt{2} & \sqrt{2} \end{bmatrix} \quad (SVD)$$

$$= \frac{1}{2\sqrt{3}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} \sqrt{3} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 1 \\ \sqrt{3} & 0 & -\sqrt{3} \end{bmatrix} \quad (RSVD)$$

Satz und Definition 1.8. Moore-Penrose-Inverse, Pseudoinverse

Ist \mathbb{K} der Körper der reellen oder der komplexen Zahlen ($\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$) und sind $m, n \in \mathbb{N}$ zwei beliebige natürliche Zahlen, dann gibt es zu jeder Matrix $\underline{\underline{A}} \in \mathbb{K}^{m \times n}$ genau eine Matrix $\underline{\underline{P}} \in \mathbb{K}^{n \times m}$ mit den Eigenschaften

- (1) $\underline{\underline{A}} \underline{\underline{P}} \underline{\underline{A}} = \underline{\underline{A}}$
- (2) $\underline{\underline{P}} \underline{\underline{A}} \underline{\underline{P}} = \underline{\underline{P}}$
- (3) $(\underline{\underline{A}} \underline{\underline{P}})^* = \underline{\underline{A}} \underline{\underline{P}}$
- (4) $(\underline{\underline{P}} \underline{\underline{A}})^* = \underline{\underline{P}} \underline{\underline{A}}$

Die vier Gleichungen (1)-(4) werden Penrose-Gleichungen (oder auch Moore-Penrose-Bedingungen) genannt. Die eindeutig bestimmte Matrix $\underline{\underline{P}}$ wird mit $\underline{\underline{A}}^+$ bezeichnet und Moore-Penrose-Inverse oder kurz Pseudoinverse genannt.

Bemerkung 1.9.

Wegen der Bedingung (1) ist eine Pseudoinverse stets auch eine verallgemeinerte Inverse. Damit ist die Existenz und Eindeutigkeit der Pseudoinversen zu einer regulären Matrix mit Bemerkung 1.2 schon bewiesen.

Beweis:

Zu zeigen ist hier zweierlei, erstens, dass es keine zwei verschiedenen Pseudoinversen zu einer Matrix geben kann (Eindeutigkeit) und zweitens, dass zu jeder Matrix mindestens eine Pseudoinverse gibt (Existenz).

Zur Eindeutigkeit:

Es seien $\underline{P}_1, \underline{P}_2$ zwei Pseudoinverse zu einer Matrix \underline{A} .

Wir zeigen zunächst, dass

$$\underline{A}\underline{P}_2 = \underline{A}\underline{P}_1 \tag{1.4}$$

und

$$\underline{P}_2\underline{A} = \underline{P}_1\underline{A} \tag{1.5}$$

gilt.

Nach der Penrose-Gleichung (1) gilt $\underline{A} = \underline{A}\underline{P}_1\underline{A}$.

Damit folgt durch Multiplikation mit \underline{P}_2 von links

$$\begin{aligned} \underline{P}_2\underline{A} &= \underline{P}_2\underline{A}\underline{P}_1\underline{A} \\ &= (\underline{P}_2\underline{A})^* (\underline{P}_1\underline{A})^* && \text{nach (4)} \\ &= [(\underline{P}_1\underline{A})(\underline{P}_2\underline{A})]^* && \text{nach Satz 1.5(1)} \\ &= [\underline{P}_1(\underline{A}\underline{P}_2\underline{A})]^* && \text{Assoziativität der Matrixmultiplikation} \\ &= [\underline{P}_1\underline{A}]^* && \text{nach (1)} \\ &= \underline{P}_1\underline{A} && \text{nach (4)} \end{aligned}$$

Durch Multiplikation mit \underline{P}_2 von rechts folgt

$$\begin{aligned} \underline{A}\underline{P}_2 &= \underline{A}\underline{P}_1\underline{A}\underline{P}_2 \\ &= (\underline{A}\underline{P}_1)^* (\underline{A}\underline{P}_2)^* && \text{nach (3)} \\ &= [(\underline{A}\underline{P}_2)(\underline{A}\underline{P}_1)]^* && \text{nach Satz 1.5(1)} \\ &= [(\underline{A}\underline{P}_2\underline{A})\underline{P}_1]^* && \text{Assoziativität der Matrixmultiplikation} \\ &= (\underline{A}\underline{P}_1)^* && \text{nach(1)} \\ &= \underline{A}\underline{P}_1 && \text{nach(3)} \end{aligned}$$

Mit diesem Ergebnis können wir nun zeigen, dass die beiden Pseudoinversen übereinstimmen müssen.

$$\begin{aligned}
 \underline{\underline{P}}_1 &= \underline{\underline{P}}_1 \underline{\underline{A}} \underline{\underline{P}}_1 && \text{nach (1)} \\
 &= \underline{\underline{P}}_1 (\underline{\underline{A}} \underline{\underline{P}}_1) \\
 &= \underline{\underline{P}}_1 (\underline{\underline{A}} \underline{\underline{P}}_2) && \text{nach (1.4)} \\
 &= (\underline{\underline{P}}_1 \underline{\underline{A}}) \underline{\underline{P}}_2 \\
 &= (\underline{\underline{P}}_2 \underline{\underline{A}}) \underline{\underline{P}}_2 && \text{nach (1.5)} \\
 &= \underline{\underline{P}}_2 \underline{\underline{A}} \underline{\underline{P}}_2 \\
 &= \underline{\underline{P}}_2 && \text{nach (1)}
 \end{aligned}$$

Zur Existenz:

Wir verwenden hier die Singulärwertzerlegung von Matrizen $\underline{\underline{A}} \in \mathbb{C}^{m \times n}$, um ganz allgemein die Existenz von Pseudoinversen zu zeigen. Wir werden im Anschluß an den Beweis sehen, dass wir diese in vielen Fällen nicht benötigen, um die Pseudoinverse zu berechnen.

Ist $\underline{\underline{A}}$ eine $(m \times n)$ Nullmatrix, dann hat die $(m \times n)$ Nullmatrix $\underline{\underline{P}}$ die gewünschten Eigenschaften.
Ist $\text{Rang } \underline{\underline{A}} = r > 0$, dann existiert eine reduzierte Singulärwertzerlegung (RSVD)

$$\underline{\underline{A}} = \underline{\underline{U}} \underline{\underline{\Sigma}} \underline{\underline{V}}^* . \quad (RSVD)$$

mit $\underline{\underline{U}} \in \mathbb{C}^{m \times r}, \underline{\underline{V}} \in \mathbb{C}^{n \times r}$ mit $\underline{\underline{U}}^* \underline{\underline{U}} = \underline{\underline{E}}_r$ bzw. $\underline{\underline{V}}^* \underline{\underline{V}} = \underline{\underline{E}}_r$ und einer Diagonalmatrix $\underline{\underline{\Sigma}} \in \mathbb{C}^{r \times r}$.

Die Matrix $\underline{\underline{P}} = \underline{\underline{V}} \underline{\underline{\Sigma}}^{-1} \underline{\underline{U}}^*$ mit $\underline{\underline{U}}, \underline{\underline{V}}, \underline{\underline{\Sigma}}$ ist pseudoinvers zu $\underline{\underline{A}}$.

Dies läßt sich durch Nachrechnen der 4 Penrose-Gleichungen überprüfen.

Da die Pseudoinverse nach dem ersten Teil des Beweises eindeutig ist, folgt dass $\underline{\underline{A}}^+$ entweder eine Nullmatrix ist oder

$$\underline{\underline{A}}^+ = \underline{\underline{V}} \underline{\underline{\Sigma}}^{-1} \underline{\underline{U}}^*$$

mit $\underline{\underline{U}}, \underline{\underline{V}}, \underline{\underline{\Sigma}}$ aus der RSVD zu $\underline{\underline{A}}$. □

Beispiel 1.10

Wir bestimmen die Pseudoinverse der Matrix

$$\underline{\underline{A}} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

Aus Beispiel 1.7 kennen wir die RSVD der Matrix $\underline{\underline{A}}$:

$$\underline{\underline{A}} = \frac{1}{2\sqrt{3}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} \sqrt{3} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 1 \\ \sqrt{3} & 0 & -\sqrt{3} \end{bmatrix} \quad (RSVD)$$

Es folgt

$$\begin{aligned}
 \underline{\underline{A}}^+ &= \frac{1}{2\sqrt{3}} \begin{bmatrix} 1 & 2 & 1 \\ \sqrt{3} & 0 & -\sqrt{3} \end{bmatrix}^* \begin{bmatrix} \sqrt{3} & 0 \\ 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}^* \\
 &= \frac{1}{2\sqrt{3}} \begin{bmatrix} 1 & 2 & 1 \\ \sqrt{3} & 0 & -\sqrt{3} \end{bmatrix}^\top \begin{bmatrix} \sqrt{3} & 0 \\ 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}^\top \\
 &= \frac{1}{2\sqrt{3}} \begin{bmatrix} 1 & \sqrt{3} \\ 2 & 0 \\ 1 & -\sqrt{3} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{3}} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \\
 &= \frac{1}{2\sqrt{3}} \begin{bmatrix} 1 & \sqrt{3} \\ 2 & 0 \\ 1 & -\sqrt{3} \end{bmatrix} \frac{1}{\sqrt{3}} \begin{bmatrix} 1 & 0 \\ 0 & \sqrt{3} \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \\
 &= \frac{1}{6} \begin{bmatrix} 1 & \sqrt{3} \\ 2 & 0 \\ 1 & -\sqrt{3} \end{bmatrix} \begin{bmatrix} 1 & 1 \\ \sqrt{3} & -\sqrt{3} \end{bmatrix} \\
 &= \frac{1}{6} \begin{bmatrix} 4 & -2 \\ 2 & 2 \\ -2 & 4 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 2 & -1 \\ 1 & 1 \\ -1 & 2 \end{bmatrix}
 \end{aligned}$$

Lemma 1.11.

Hat eine reelle ($m \times n$) Matrix $\underline{\underline{A}}$ vollen Spaltenrang (alle Spalten sind linear unabhängig), dann ist $\underline{\underline{A}}^\top \underline{\underline{A}}$ regulär.

Bei vollem Zeilenrang ist entsprechend $\underline{\underline{A}} \underline{\underline{A}}^\top$ regulär.

Für komplexe Matrizen ist bei vollem Spaltenrang $\underline{\underline{A}}^* \underline{\underline{A}}$ und bei vollem Zeilenrang $\underline{\underline{A}} \underline{\underline{A}}^*$ regulär.

Beweis:

Da Spalten- und Zeilenrang durch transponieren bzw. adjungieren ineinander übergehen, brauchen wir die Aussage nur für den Spaltenrang zu zeigen.

Die Spalten $\underline{a}_1, \dots, \underline{a}_n$ der Matrix $\underline{\underline{A}}$ sind nach Voraussetzung linear unabhängig, d.h.

die Gleichung $\underline{\underline{A}} \underline{x} = x_1 \underline{a}_1 + \dots + x_n \underline{a}_n = \underline{0}$ hat nur die triviale Lösung $x_1 = \dots = x_n = 0$.

Somit gilt Kern $\underline{\underline{A}} = \{\underline{0}\}$.

Wir zeigen jetzt, dass auch Kern $\underline{\underline{A}}^\top \underline{\underline{A}} = \{\underline{0}\}$ gilt. Daraus folgt dann mit dem Rangsatz oder Dimensionssatz, dass $\underline{\underline{A}}^\top \underline{\underline{A}}$ regulär ist.

Sei $x \in \text{Kern } \underline{\underline{A}}^\top \underline{\underline{A}}$ also $\underline{\underline{0}} = (\underline{\underline{A}}^\top \underline{\underline{A}})x$, dann folgt durch Multiplikation mit x^\top von links

$$\begin{aligned} \underline{\underline{0}} &= x^\top (\underline{\underline{A}}^\top \underline{\underline{A}}) x = (x^\top \underline{\underline{A}}^\top) (\underline{\underline{A}} x) \\ &= (\underline{\underline{A}} x)^\top (\underline{\underline{A}} x) = \langle \underline{\underline{A}} x, \underline{\underline{A}} x \rangle \\ &= \|\underline{\underline{A}} x\|^2 \\ \Rightarrow \underline{\underline{A}} x &= \underline{\underline{0}} \\ \Rightarrow x &\in \text{Kern } \underline{\underline{A}} = \{\underline{\underline{0}}\} \\ \Rightarrow x &= \underline{\underline{0}} \\ \Rightarrow \text{Kern } \underline{\underline{A}}^\top \underline{\underline{A}} &= \{\underline{\underline{0}}\} \end{aligned}$$

Ersetzen wir überall das Transponieren durch Adjungieren, dann ergibt sich der Beweis für den komplexen Fall. □

Satz 1.12. Rechenregeln für Pseudoinverse

Für beliebiges $\underline{\underline{A}} \in \mathbb{C}^{m \times n}$ gilt:

- (1) $(\underline{\underline{A}}^+)^+ = \underline{\underline{A}}$
- (2) $(\underline{\underline{A}} \underline{\underline{A}}^*)^+ = (\underline{\underline{A}}^*)^+ \underline{\underline{A}}^+$
- (3) $(\lambda \underline{\underline{A}})^+ = \frac{1}{\lambda} \underline{\underline{A}}^+ \quad \text{für } \lambda \in \mathbb{C}^{\neq 0}$
- (4) Das Bilden der Pseudoinversen ist mit Transponieren, Konjugieren und Adjungieren vertauschbar:

$$\left(\underline{\underline{A}}^\top\right)^+ = \left(\underline{\underline{A}}^+\right)^\top, \quad \left(\overline{\underline{\underline{A}}}\right)^+ = \overline{\left(\underline{\underline{A}}^+\right)}, \quad \left(\underline{\underline{A}}^*\right)^+ = \left(\underline{\underline{A}}^+\right)^*$$

- (5) Tychonov-Regularisierung

$$\lim_{\delta \rightarrow 0} (\underline{\underline{A}}^* \underline{\underline{A}} + \delta \underline{\underline{E}})^{-1} \underline{\underline{A}}^* = \underline{\underline{A}}^+ = \lim_{\delta \rightarrow 0} \underline{\underline{A}}^* (\underline{\underline{A}} \underline{\underline{A}}^* + \delta \underline{\underline{E}})^{-1}$$

Beweis:

- (1) Hat $\underline{\underline{A}}$ die (RSVD) $\underline{\underline{A}} = \underline{\underline{U}} \underline{\underline{\Sigma}} \underline{\underline{V}}^*$, dann gilt

$$\underline{\underline{A}}^+ = \underline{\underline{V}} \underline{\underline{\Sigma}}^{-1} \underline{\underline{U}}^*$$

und

$$\left(\underline{\underline{A}}^+\right)^+ = \underline{\underline{U}} \left(\underline{\underline{\Sigma}}^{-1}\right)^{-1} \underline{\underline{V}}^* = \underline{\underline{U}} \underline{\underline{\Sigma}} \underline{\underline{V}}^* = \underline{\underline{A}}$$

(2) Hat $\underline{\underline{A}}$ die (RSVD) $\underline{\underline{A}} = \underline{\underline{U}} \underline{\underline{\Sigma}} \underline{\underline{V}}^*$, dann gilt

$$\begin{aligned} \underline{\underline{A}}^* &= \left(\underline{\underline{U}} \underline{\underline{\Sigma}} \underline{\underline{V}}^* \right)^* = \underline{\underline{V}} \underline{\underline{\Sigma}}^* \underline{\underline{U}}^* \\ &= \underline{\underline{V}} \underline{\underline{\Sigma}} \underline{\underline{U}}^* \end{aligned}$$

Da $\underline{\underline{\Sigma}}$ eine reelle Diagonalmatrix ist, stimmt sie mit $\underline{\underline{\Sigma}}^*$ überein.

$$\begin{aligned} \underline{\underline{A}} \underline{\underline{A}}^* &= \underline{\underline{U}} \underline{\underline{\Sigma}} \underbrace{\underline{\underline{V}}^* \underline{\underline{V}}}_{=\underline{\underline{E}}} \underline{\underline{\Sigma}} \underline{\underline{U}}^* \\ &= \underline{\underline{U}} \underline{\underline{\Sigma}} \underline{\underline{\Sigma}} \underline{\underline{U}}^* \\ &= \underline{\underline{U}} \underline{\underline{\Sigma}}^2 \underline{\underline{U}}^* \\ \left(\underline{\underline{A}} \underline{\underline{A}}^* \right)^+ &= \left(\underline{\underline{U}} \underline{\underline{\Sigma}}^2 \underline{\underline{U}}^* \right)^+ \\ &= \underline{\underline{U}} \left(\underline{\underline{\Sigma}}^2 \right)^{-1} \underline{\underline{U}}^* \\ \left(\underline{\underline{A}}^* \right)^+ \underline{\underline{A}}^+ &= \left(\underline{\underline{V}} \underline{\underline{\Sigma}} \underline{\underline{U}}^* \right)^+ \left(\underline{\underline{U}} \underline{\underline{\Sigma}} \underline{\underline{V}}^* \right)^+ \\ &= \underline{\underline{U}} \underline{\underline{\Sigma}}^{-1} \underbrace{\underline{\underline{V}}^* \underline{\underline{V}}}_{=\underline{\underline{E}}} \underline{\underline{\Sigma}}^{-1} \underline{\underline{U}}^* \\ &= \underline{\underline{U}} \underline{\underline{\Sigma}}^{-1} \underline{\underline{\Sigma}}^{-1} \underline{\underline{U}}^* \\ &= \underline{\underline{U}} \left(\underline{\underline{\Sigma}}^{-1} \right)^2 \underline{\underline{U}}^* \end{aligned}$$

Wegen $\underline{\underline{\Sigma}} = \text{diag}(\sigma_1, \dots, \sigma_r)$ gilt

$$\underline{\underline{\Sigma}}^{-1} = \text{diag}\left(\frac{1}{\sigma_1}, \dots, \frac{1}{\sigma_r}\right) \quad \text{und}$$

$$\left(\underline{\underline{\Sigma}}^{-1}\right)^2 = \text{diag}\left(\frac{1}{\sigma_1^2}, \dots, \frac{1}{\sigma_r^2}\right) = \left(\underline{\underline{\Sigma}}^2\right)^{-1}$$

Damit folgt

$$\left(\underline{\underline{A}} \underline{\underline{A}}^* \right)^+ = \underline{\underline{U}} \left(\underline{\underline{\Sigma}}^2 \right)^{-1} \underline{\underline{U}}^* = \underline{\underline{U}} \left(\underline{\underline{\Sigma}}^{-1} \right)^2 \underline{\underline{U}}^* = \left(\underline{\underline{A}}^* \right)^+ \underline{\underline{A}}^+$$

(3) ...als Übung empfohlen

(4) ...als Übung empfohlen

(5) Es sei $\underline{\underline{A}} = \underline{\underline{U}} \underline{\underline{\Sigma}} \underline{\underline{V}}^*$ die SVD zu $\underline{\underline{A}}$, dann gilt analog zu Teil (2)

$$\underline{\underline{A}}^* \underline{\underline{A}} = \underline{\underline{V}} \underline{\underline{\Sigma}}^2 \underline{\underline{V}}^*$$

Außerdem gilt

$$\delta \underline{\underline{E}} = \delta \underline{\underline{V}} \underline{\underline{V}}^* = \delta \underline{\underline{V}} \underline{\underline{E}} \underline{\underline{V}}^* = \underline{\underline{V}} \delta \underline{\underline{E}} \underline{\underline{V}}^*$$

Damit erhalten wir

$$\begin{aligned} \underline{\underline{A}}^* \underline{\underline{A}} + \delta \underline{\underline{E}} &= \underline{\underline{V}} \underline{\underline{\Sigma}}^2 \underline{\underline{V}}^* + \underline{\underline{V}} \delta \underline{\underline{E}} \underline{\underline{V}}^* \\ &= \underline{\underline{V}} (\underline{\underline{\Sigma}}^2 + \delta \underline{\underline{E}}) \underline{\underline{V}}^* \\ &= \underline{\underline{V}} \begin{bmatrix} \sigma_1^2 + \delta & 0 & \dots & 0 \\ & \vdots & & \\ 0 & 0 & \dots & \sigma_n^2 + \delta \end{bmatrix} \underline{\underline{V}}^* \end{aligned}$$

und damit

$$\begin{aligned} (\underline{\underline{A}}^* \underline{\underline{A}} + \delta \underline{\underline{E}})^{-1} \underline{\underline{A}}^* &= \underline{\underline{V}} \begin{bmatrix} \frac{1}{\sigma_1^2 + \delta} & 0 & \dots & 0 \\ & \vdots & & \\ 0 & 0 & \dots & \frac{1}{\sigma_n^2 + \delta} \end{bmatrix} \underline{\underline{V}}^* \underline{\underline{V}} \underline{\underline{\Sigma}} \underline{\underline{U}}^* \\ &= \underline{\underline{V}} \begin{bmatrix} \frac{\sigma_1}{\sigma_1^2 + \delta} & 0 & \dots & 0 \\ & \vdots & & \\ 0 & 0 & \dots & \frac{\sigma_n}{\sigma_n^2 + \delta} \end{bmatrix} \underline{\underline{U}}^* && \text{An Stellen mit } \sigma_k = 0 \text{ steht hier eine Null} \\ \longrightarrow \underline{\underline{V}} \begin{bmatrix} \frac{1}{\sigma_1} & 0 & \dots & 0 \\ & \vdots & & \\ 0 & 0 & \dots & \frac{1}{\sigma_n} \end{bmatrix} \underline{\underline{U}}^* && \text{An Stellen mit } \sigma_k = 0 \text{ steht hier eine Null} \\ &= \underline{\underline{V}} \underline{\underline{\Sigma}}^+ \underline{\underline{U}}^* = \underline{\underline{A}}^+ \end{aligned}$$

Damit haben wir

$$\lim_{\delta \rightarrow 0} (\underline{\underline{A}}^* \underline{\underline{A}} + \delta \underline{\underline{E}})^{-1} \underline{\underline{A}}^* = \underline{\underline{A}}^+$$

gezeigt. Die zweite Grenzwertaussage folgt analog.

□

1.3 Berechnung von Pseudoinversen für besondere Fälle

Diagonalmatrizen

Wenn $\underline{\underline{A}}$ eine Diagonalmatrix ist, dann ist auch $\underline{\underline{A}}^+$ eine Diagonalmatrix. Die Pseudoinverse entsteht aus $\underline{\underline{A}}$ durch Ersetzen der Diagonalelemente $a_{kk} \neq 0$ durch ihren reziproken Wert.

$$\underline{\underline{A}} = \text{diag}(a_{11}, \dots, a_{nn}) \quad \Rightarrow \quad \underline{\underline{A}}^+ = \text{diag}(\hat{a}_{11}, \dots, \hat{a}_{nn}) \quad \text{mit} \quad \hat{a}_{kk} = \begin{cases} 0, & \text{wenn } a_{kk} = 0 \\ \frac{1}{a_{kk}}, & \text{sonst.} \end{cases}$$

Matrizen mit vollem Spaltenrang (d.h. mit linear unabhängigen Spaltenvektoren)

Nach Lemma 1.11 ist $\underline{\underline{A}}^* \underline{\underline{A}}$ unter dieser Voraussetzung regulär und besitzt somit eine Inverse mit

$$\underline{\underline{E}}_n = \left(\underline{\underline{A}}^* \underline{\underline{A}} \right)^{-1} \left(\underline{\underline{A}}^* \underline{\underline{A}} \right) = \left(\underbrace{\left(\underline{\underline{A}}^* \underline{\underline{A}} \right)^{-1} \underline{\underline{A}}^*}_{=\underline{\underline{A}}^+} \right) \underline{\underline{A}}$$

Damit erhalten wir für Matrizen $\underline{\underline{A}} \in \mathbb{C}^{m \times n}$ mit $\text{Rang } \underline{\underline{A}} = n$ die Pseudoinverse

$$\underline{\underline{A}}^+ = \left(\underline{\underline{A}}^* \underline{\underline{A}} \right)^{-1} \underline{\underline{A}}^* \tag{1.6}$$

Matrizen mit vollem Zeilenrang (d.h. mit linear unabhängigen Zeilenvektoren)

In diesem Fall ist $\underline{\underline{B}} := \underline{\underline{A}}^*$ eine Matrix mit vollem Spaltenrang und es gilt

$$\underline{\underline{B}}^+ = \left(\underline{\underline{B}}^* \underline{\underline{B}} \right)^{-1} \underline{\underline{B}}^* .$$

Durch Adjungieren folgt

$$\left(\underline{\underline{B}}^+ \right)^* = \left(\left(\underline{\underline{B}}^* \underline{\underline{B}} \right)^{-1} \underline{\underline{B}}^* \right)^* = \left(\underline{\underline{B}}^* \right)^* \left(\left(\underline{\underline{B}}^* \underline{\underline{B}} \right)^{-1} \right)^*$$

und nach den Rechenregeln 1.5, 1.12 gilt

$$\underline{\underline{B}}^* = \left(\underline{\underline{A}}^* \right)^* = \underline{\underline{A}} \quad \left(\underline{\underline{B}}^+ \right)^* = \left(\underline{\underline{B}}^* \right)^+ = \underline{\underline{A}}^+$$

Es folgt

$$\begin{aligned} \underline{\underline{A}}^+ &= \left(\left(\underline{\underline{B}}^* \underline{\underline{B}} \right)^{-1} \underline{\underline{B}}^* \right)^* \\ &= \left(\underline{\underline{B}}^* \right)^* \left(\left(\underline{\underline{B}}^* \underline{\underline{B}} \right)^{-1} \right)^* \\ &= \underline{\underline{B}} \left(\left(\underline{\underline{B}}^* \underline{\underline{B}} \right)^* \right)^{-1} \\ &= \underline{\underline{B}} \left(\underline{\underline{B}}^* \underline{\underline{B}} \right)^{-1} \\ &= \underline{\underline{A}}^* \left(\underline{\underline{A}} \underline{\underline{A}}^* \right)^{-1} \end{aligned}$$

Also gilt für Matrizen mit vollem Zeilenrang

$$\underline{\underline{A}}^+ = \underline{\underline{A}}^* \left(\underline{\underline{A}} \underline{\underline{A}}^* \right)^{-1} \tag{1.7}$$

Matrizen, nicht diagonal und nicht mit maximalem Rang

(d.h. $(m \times n)$ -Matrizen mit $r = \text{Rang } \underline{\underline{A}} < \min\{m, n\}$ und mindestens ein $a_{ij} \neq 0$ für $i \neq j$)

Mit den bisherigen Hilfsmitteln bleibt hier die SVD bzw RSVD als Methode übrig.

Für die SVD gilt:

Die Singulärwerte, also die Diagonalelemente von Σ sind die Quadratwurzeln aus den Eigenwerten von $\underline{\underline{A}}\underline{\underline{A}}^*$.

Die Spalten von $\underline{\underline{U}}$ sind die normierten Eigenvektoren zu den Eigenwerten bezüglich $\underline{\underline{A}}\underline{\underline{A}}^*$.

Die Spalten von $\underline{\underline{V}}$ sind die normierten Eigenvektoren zu den Eigenwerten bezüglich $\underline{\underline{A}}^*\underline{\underline{A}}$.

Beispiel 1.13

$$(1) \underline{\underline{A}}_1 = \begin{bmatrix} 7 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -\frac{1}{2} \end{bmatrix}$$

$$\text{Die Pseudoinverse ist } \underline{\underline{A}}_1^+ = \begin{bmatrix} \frac{1}{7} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -2 \end{bmatrix}$$

$$(2) \underline{\underline{A}}_2 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

Die Matrix hat nur eine Spalte. Diese ist vom Nullvektor verschieden und deshalb linear unabhängig.

Die Matrix hat also vollen Spaltenrang. Daher gilt nach (1.6)

$$\underline{\underline{A}}_2^+ = \left(\underline{\underline{A}}_2^* \underline{\underline{A}}_2 \right)^{-1} \underline{\underline{A}}_2^*$$

$$\underline{\underline{A}}_2^* \underline{\underline{A}}_2 = \begin{bmatrix} 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = 1 + 4 + 9 = 14 \quad \Rightarrow \quad \left(\underline{\underline{A}}_2^* \underline{\underline{A}}_2 \right)^{-1} = \frac{1}{14} \text{ und daher}$$

$$\underline{\underline{A}}_2^+ = \frac{1}{14} \begin{bmatrix} 1 & 2 & 3 \end{bmatrix}$$

$$(3) \underline{\underline{A}}_3 = \begin{bmatrix} 1 & 2 & 3 \end{bmatrix}$$

Wegen $\underline{\underline{A}}_3 = \underline{\underline{A}}_2^\top$ gilt

$$\underline{\underline{A}}_3^+ = \left(\underline{\underline{A}}_2^\top \right)^+ = \left(\underline{\underline{A}}_2^+ \right)^\top = \frac{1}{14} \begin{bmatrix} 1 & 2 & 3 \end{bmatrix}^\top = \frac{1}{14} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

$$(4) \underline{\underline{A_4}} = \begin{bmatrix} 2 & 0 & 0 \\ 3 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Hier bestimmen wir die Pseudoinverse über die SVD.

$$\underline{\underline{A_4}} \underline{\underline{A_4^*}} = \begin{bmatrix} 2 & 0 & 0 \\ 3 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 2 & 3 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 4 & 6 & 0 \\ 6 & 13 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Das charakteristische Polynom lautet

$$\rho(\lambda) = \det(\underline{\underline{A_4}} - \lambda \underline{\underline{E}}) = -\lambda(\lambda - 16)(\lambda - 1)$$

Somit haben wir die Eigenwerte $\lambda_1 = 16, \lambda_2 = 1, \lambda_3 = 0$

Damit erhalten wir die Matrix

$$\underline{\underline{\Sigma}} = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Die Bestimmung von normierten Eigenvektoren ergibt

$$\underline{\underline{v_1}} = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \quad \underline{\underline{v_2}} = \frac{1}{\sqrt{5}} \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix}, \quad \underline{\underline{v_3}} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Damit erhalten wir die Matrix

$$\underline{\underline{U}} = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 & 2 & 0 \\ 2 & -1 & 0 \\ 0 & 0 & \sqrt{5} \end{bmatrix}$$

$$\underline{\underline{A_4^*}} \underline{\underline{A_4}} = \begin{bmatrix} 2 & 3 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 2 & 0 & 0 \\ 3 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 13 & 6 & 0 \\ 6 & 4 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

hat die gleichen Eigenwerte.

Die Bestimmung von normierten Eigenvektoren ergibt

$$\underline{\underline{v_1}} = \frac{1}{\sqrt{5}} \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}, \quad \underline{\underline{v_2}} = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix}, \quad \underline{\underline{v_3}} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Damit erhalten wir die Matrix

$$\underline{\underline{V}} = \frac{1}{\sqrt{5}} \begin{bmatrix} 2 & 1 & 0 \\ 1 & -2 & 0 \\ 0 & 0 & \sqrt{5} \end{bmatrix}$$

Es folgt

$$\underline{\underline{A_4}} = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 & 2 & 0 \\ 2 & -1 & 0 \\ 0 & 0 & \sqrt{5} \end{bmatrix} \begin{bmatrix} 4 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \frac{1}{\sqrt{5}} \begin{bmatrix} 2 & 1 & 0 \\ 1 & -2 & 0 \\ 0 & 0 & \sqrt{5} \end{bmatrix} \quad (SVD)$$

$$= \frac{1}{\sqrt{5}} \begin{bmatrix} 1 & 2 \\ 2 & -1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix}, \frac{1}{\sqrt{5}} \begin{bmatrix} 2 & 1 & 0 \\ 1 & -2 & 0 \end{bmatrix} \quad (RSVD)$$

Damit erhalten wir die Pseudoinverse

$$\begin{aligned} \underline{\underline{A_4}}^+ &= \frac{1}{5} \begin{bmatrix} 2 & 1 \\ 1 & -2 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{4} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 0 \\ 2 & -1 & 0 \end{bmatrix} \\ &= \frac{1}{5} \begin{bmatrix} 2 & 1 \\ 1 & -2 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{4} & \frac{1}{2} & 0 \\ 2 & -1 & 0 \end{bmatrix} \\ &= \frac{1}{5} \begin{bmatrix} \frac{5}{2} & 0 & 0 \\ -\frac{15}{4} & \frac{5}{2} & 0 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & 0 & 0 \\ -\frac{3}{4} & \frac{1}{2} & 0 \\ 0 & 0 & 0 \end{bmatrix} \\ &= \frac{1}{4} \begin{bmatrix} 2 & 0 & 0 \\ -3 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix} \end{aligned}$$

Wegen der speziellen Struktur dieser Matrix hätten wir hier alternativ auch zuerst die Pseudoinverse von $\underline{\underline{B}} = \begin{bmatrix} 2 & 0 \\ 3 & 2 \end{bmatrix}$ berechnen können.

Alternative Berechnung von $\underline{\underline{A}}_4$:

$$\underline{\underline{B}}^+ = \left(\underline{\underline{B}}^* \underline{\underline{B}} \right)^{-1} \underline{\underline{B}}^*$$

$$\underline{\underline{B}}^* \underline{\underline{B}} = \begin{bmatrix} 2 & 3 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 3 & 2 \end{bmatrix} = \begin{bmatrix} 13 & 6 \\ 6 & 4 \end{bmatrix}$$

$$\det(\underline{\underline{B}}^* \underline{\underline{B}}) = 52 - 36 = 16$$

$$\left(\underline{\underline{B}}^* \underline{\underline{B}} \right)^{-1} = \frac{1}{16} \begin{bmatrix} 4 & -6 \\ -6 & 13 \end{bmatrix}$$

$$\begin{aligned} \underline{\underline{B}}^+ &= \frac{1}{16} \begin{bmatrix} 4 & -6 \\ -6 & 13 \end{bmatrix} \begin{bmatrix} 2 & 3 \\ 0 & 2 \end{bmatrix} = \frac{1}{16} \begin{bmatrix} 8 & 0 \\ -12 & 8 \end{bmatrix} \\ &= \frac{1}{4} \begin{bmatrix} 2 & 0 \\ -3 & 2 \end{bmatrix} \end{aligned}$$

Wenn wir jetzt die zuvor weggeschnittenen Nullen wieder anfügen, erhalten wir

$$\underline{\underline{A}}_4^+ = \frac{1}{4} \begin{bmatrix} 2 & 0 & 0 \\ -3 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

1.4 Pseudoinverse und Projektionen

Vorbemerkungen über Projektionen:

Eine lineare Abbildung $P : \mathbb{K}^m \rightarrow \mathbb{K}^m$ ist eine Projektion, wenn $P \circ P = P$ bzw. $\underline{\underline{A}}^2 = \underline{\underline{A}}$ für die zugehörige Matrix $\underline{\underline{A}}$ gilt.

Ist $\underline{\underline{A}}$ eine Projektionsmatrix, dann gilt:

$$\text{Bild } \underline{\underline{A}} = \text{Kern}(\underline{\underline{E}} - \underline{\underline{A}}) = \text{Kern}(\underline{\underline{A}} - \underline{\underline{E}}) \quad (1.8)$$

$$\text{Kern } \underline{\underline{A}} = \text{Bild}(\underline{\underline{E}} - \underline{\underline{A}}) = \text{Bild}(\underline{\underline{A}} - \underline{\underline{E}}) \quad (1.9)$$

Zum Beweis:

Zu (1.8):

Sei $\underline{y} \in \text{Bild } \underline{\underline{A}}$, dann gibt es ein $\underline{x} \in \mathbb{K}^m$ mit $\underline{y} = \underline{\underline{A}} \underline{x}$.

$$\begin{aligned} (\underline{\underline{E}} - \underline{\underline{A}})\underline{y} &= (\underline{\underline{E}} - \underline{\underline{A}})\underline{\underline{A}} \underline{x} = \underline{\underline{A}} \underline{x} - \underline{\underline{A}}^2 \underline{x} = \underline{\underline{A}} \underline{x} - \underline{\underline{A}} \underline{x} = \underline{0} \\ \Rightarrow \underline{y} &\in \text{Kern}(\underline{\underline{E}} - \underline{\underline{A}}) \end{aligned}$$

$$\underline{x} \in \text{Kern}(\underline{\underline{E}} - \underline{\underline{A}}) \Rightarrow \underline{0} = (\underline{\underline{E}} - \underline{\underline{A}})\underline{x} = \underline{x} - \underline{\underline{A}} \underline{x} \Rightarrow \underline{\underline{A}} \underline{x} = \underline{x} \Rightarrow \underline{x} \in \text{Bild } \underline{\underline{A}}$$

Zu (1.9):

$$\underline{x} \in \text{Kern } \underline{A} \Rightarrow (\underline{E} - \underline{A})\underline{x} = \underline{x} - \underline{A}\underline{x} = \underline{x} - \underline{0} = \underline{x} \Rightarrow \underline{x} \in \text{Bild}(\underline{E} - \underline{A})$$

Sei $\underline{y} \in \text{Bild}(\underline{E} - \underline{A})$, dann gibt es ein $\underline{x} \in \mathbb{K}^m$ mit $\underline{y} = (\underline{E} - \underline{A})\underline{x}$.

$$\begin{aligned} \underline{A}\underline{y} &= \underline{A}(\underline{E} - \underline{A})\underline{x} = (\underline{A} - \underline{A}^2)\underline{x} = (\underline{A} - \underline{A})\underline{x} = \underline{0}\underline{x} = \underline{0} \\ &\Rightarrow \underline{y} \in \text{Kern } \underline{A} \end{aligned}$$

Eine Projektion ist orthogonal, wenn zusätzlich zur Projektionseigenschaft der Vektor $\underline{x} - \underline{A}\underline{x}$ stets das Lot von \underline{x} auf den Unterraum ist, auf den projiziert wird:

$$\text{Bild}(\underline{E} - \underline{A}) \perp \text{Bild } \underline{A} \quad (1.10)$$

Aus 1.10 folgt mit der Gleichung (1.9):

Die Projektion mit der Matrix \underline{A} ist genau dann orthogonal, wenn

$$\text{Kern } \underline{A} \perp \text{Bild } \underline{A}. \quad (1.11)$$

Für lineare Projektionen gibt es zwei Extremfälle: Die Projektion auf den gesamten Raum und die Projektion auf den Nullpunkt. Davon verschiedene Projektionen bezeichnen wir als echte Projektionen.

Ist P eine echte Projektion, dann ist die zugehörige Matrix \underline{A} weder regulär noch die Nullmatrix. Also sind sowohl $\text{Bild } P = \text{Bild } \underline{A}$ als auch $\text{Kern } P = \text{Kern } \underline{A}$ echte Unterräume des \mathbb{K}^m und verschieden vom Nullraum.

Als Folgerung ergibt sich daraus, dass echte Projektionen genau zwei Eigenwerte haben.

$U_0 := \text{Kern } \underline{A} \neq \{0\}$ ist Eigenraum zum Eigenwert $\lambda_0 = 0$, denn für $\underline{x} \in \text{Kern } \underline{A}$ gilt

$$\underline{A}\underline{x} = \underline{0} = 0 \cdot \underline{x} = \lambda_0 \underline{x}$$

$U_1 := \text{Bild } \underline{A} \neq \{0\}$ ist Eigenraum zum Eigenwert $\lambda_1 = 1$:

Ist $\underline{y} \in \text{Bild } \underline{A}$ dann gibt es ein $\underline{x} \in \mathbb{K}^m$ mit $\underline{y} = \underline{A}\underline{x}$.

$$\underline{A}\underline{y} = \underline{A}(\underline{A}\underline{x}) = (\underline{A}\underline{A})\underline{x} = \underline{A}\underline{x} = \underline{y} = \lambda_1 \underline{y}$$

Verschiedene Eigenräume haben nur den Nullvektor gemeinsam. Aus dem Dimensionssatz folgt dann, dass es keine weiteren Eigenräume und damit auch keine weiteren Eigenwerte geben kann.

Ist die Matrix \underline{A} reell und symmetrisch oder komplex und hermitesch, also $\underline{A} = \underline{A}^*$, dann gilt

$$\text{für alle } \underline{x}, \underline{y} \in \mathbb{K}^m : \quad \langle \underline{A}\underline{x}, \underline{y} \rangle = \langle \underline{x}, \underline{A}\underline{y} \rangle \quad (1.12)$$

und die Eigenräume sind paarweise orthogonal.

Nach diesen Vorbemerkungen kommen wir nun zu Projektionen, die Pseudomatrizen verwenden.

Satz 1.14.

Für eine beliebige $(m \times n)$ Matrix \underline{A} gilt:

Die Matrix $\underline{B} := \underline{A} \underline{A}^+$ beschreibt eine orthogonale Projektion auf Bild \underline{A} .

Beweis:

Wir zeigen zunächst, dass \underline{B} die Matrix einer Projektion ist.

$$\begin{aligned} \underline{B}^2 &= \overbrace{\underline{A} \underline{A}^+ \underline{A}} \underline{A}^+ \\ &= \underline{A} \underline{A}^+ && \text{nach der Penrosebedingung (1)} \\ &= \underline{B} \end{aligned}$$

Damit ist \underline{B} die Matrix einer Projektion. Der Unterraum auf den projiziert wird, ist Bild \underline{B} .

Als nächstes zeigen wir, dass durch \underline{B} eine Projektion auf Bild \underline{A} gegeben ist. Dazu müssen wir zeigen, dass Bild \underline{B} mit Bild \underline{A} übereinstimmt.

Da \underline{B} die Verkettung zweier linearer Abbildungen beschreibt, ist das Bild eine Teilmenge des Bildes der zuletzt anzuwendenden Matrix, d.h.

$$\text{Bild } \underline{B} \subseteq \text{Bild } \underline{A}$$

Ist \underline{y} ein beliebiges Element aus Bild \underline{A} , dann gibt es ein $\underline{x} \in \mathbb{K}^m$ mit $\underline{y} = \underline{A} \underline{x}$. Es folgt mit der Penrosegleichung (1)

$$\underline{B} \underline{y} = \underline{A} \underline{A}^+ \underline{y} = \underline{A} \underline{A}^+ \underline{A} \underline{x} = \underline{A} \underline{x} = \underline{y}$$

d.h. $\underline{y} \in \text{Bild } \underline{B}$ für alle $\underline{y} \in \text{Bild } \underline{A}$. Also

$$\text{Bild } \underline{A} \subseteq \text{Bild } \underline{B}$$

Damit ist die Gleichheit der Bilder bewiesen.

Es bleibt noch zu zeigen, dass die Projektion orthogonal ist. Nach der Penrosegleichung (3) gilt:

$$\underline{B}^* = (\underline{A} \underline{A}^+)^* = \underline{A} \underline{A}^+ = \underline{B}$$

Damit folgt aus (1.12) für beliebige $\underline{x}, \underline{y} \in \mathbb{K}^m$

$$\begin{aligned} \langle (\underline{E} - \underline{A} \underline{A}^+) \underline{x}, \underline{A} \underline{A}^+ \underline{y} \rangle &= \langle \underline{A} \underline{A}^+ (\underline{E} - \underline{A} \underline{A}^+) \underline{x}, \underline{y} \rangle \\ &= \langle (\underline{A} \underline{A}^+ - (\underline{A} \underline{A}^+)^2) \underline{x}, \underline{y} \rangle \\ &= \langle (\underline{A} \underline{A}^+ - \underline{A} \underline{A}^+) \underline{x}, \underline{y} \rangle \\ &= \langle \underline{0} \underline{x}, \underline{y} \rangle = 0 \\ \Rightarrow (\underline{E} - \underline{A} \underline{A}^+) \underline{x} &\perp \text{Bild } \underline{A} \underline{A}^+ \end{aligned}$$

Damit haben wir nach (1.10) die Orthogonalität gezeigt. □

Durch Vertauschen der Rollen von \underline{A} und \underline{A}^+ und Verwendung der Rechenregeln 1.12 folgt das

Korollar 1.15.

Für eine beliebige $(m \times n)$ Matrix \underline{A} gilt:

Die Matrix $\underline{B} := \underline{A}^+ \underline{A}$ beschreibt eine orthogonale Projektion auf $\text{Bild } \underline{A}^+$.

1.5 Anwendungen zur Pseudoinversen

Die Pseudoinverse kann erfolgreich verwendet werden, um lineare Gleichungssysteme $\underline{A}x = \underline{b}$ zu lösen, deren Matrix nicht regulär ist. Bei homogenen Systemen kann stets die allgemeine Lösung ermittelt werden. Bei inhomogenen Systemen liefert $\underline{x} := \underline{A}^+ \underline{b}$ eine partikuläre Lösung, falls das LGS lösbar ist.

Ist das LGS nicht lösbar, dann ist $\underline{x} := \underline{A}^+ \underline{b}$ eine optimale Näherungslösung, für die gilt

$$\|\underline{A}\underline{x} - \underline{b}\| = \min \left\{ \|\underline{A}\underline{x} - \underline{b}\| \mid \underline{x} \in \mathbb{C}^n \right\},$$

denn nach (1.10) und Satz 1.14 ist $\underline{A}\underline{x} - \underline{b} = \underline{A}\underline{A}^+ \underline{b} - \underline{b}$ das Lot von \underline{b} auf $\text{Bild } \underline{A}$.

\underline{x} wird in diesem Fall als Minimum-Norm-Lösung bezeichnet. Im Detail:

Satz 1.16.

Für eine beliebige $(m \times n)$ Matrix \underline{A} gilt

$$\text{Kern } \underline{A} = \text{Bild}(\underline{E} - \underline{A}^+ \underline{A})$$

Das heißt, das das homogene LGS $\underline{A}x = \underline{0}$ den Unterraum $\mathbb{L}_{\text{hom}} = \text{Bild}(\underline{E} - \underline{A}^+ \underline{A})$ als Lösungsraum besitzt.

Mit anderen Worten:

Ist $\underline{s}_1, \dots, \underline{s}_k$ eine maximale Auswahl von linear unabhängigen Spalten der Matrix $\underline{E} - \underline{A}^+ \underline{A}$, dann lautet die allgemeine Lösung des homogenen Systems $\underline{A}x = \underline{0}$

$$\underline{x}_{\text{hom}} = t_1 \underline{s}_1 + \dots + t_k \underline{s}_k = \sum_{v=1}^k t_v \underline{s}_v.$$

Beweis:

Nach der Penrose-Gleichung (1) gilt

$$\begin{aligned} \underline{A} &= \underline{A}\underline{A}^+ \underline{A} && \Leftrightarrow \underline{A}\underline{E} = \underline{A}\underline{A}^+ \underline{A} \\ &&& \Leftrightarrow \underline{A}\underline{E} - \underline{A}\underline{A}^+ \underline{A} = \underline{0} \\ &&& \Leftrightarrow \underline{A} \left(\underline{E} - \underline{A}^+ \underline{A} \right) = \underline{0} \end{aligned}$$

Daraus folgt für alle $\underline{t} \in \mathbb{C}^n$:

$$\underline{A} \left[\left(\underline{E} - \underline{A}^+ \underline{A} \right) \underline{t} \right] = \underline{0}$$

Somit ist Bild $\left(\underline{E} - \underline{A}^+ \underline{A} \right) \subseteq \text{Kern } \underline{A}$
 Andererseits gilt für $\underline{x} \in \text{Kern } \underline{A}$:

$$\left(\underline{E} - \underline{A}^+ \underline{A} \right) \underline{x} = \underline{x} - \underbrace{\underline{A}^+ \underline{A} \underline{x}}_{=0} = \underline{x}$$

Somit gilt $\text{Kern } \underline{A} \subseteq \text{Bild} \left(\underline{E} - \underline{A}^+ \underline{A} \right)$ Das heißt aber nichts anderes als

$$\text{Kern } \underline{A} = \text{Bild} \left(\underline{E} - \underline{A}^+ \underline{A} \right)$$

□

Satz 1.17.

Für das LGS $\underline{A} \underline{x} = \underline{b}$ mit einer beliebigen komplexe $(m \times n)$ Matrix \underline{A} und $\underline{b} \in \mathbb{C}^n$ gilt:

Ist das LGS lösbar, also $\text{Rang } \underline{A} = \text{Rang } \underline{A} | \underline{b}$, dann ist

$\underline{x}_p = \underline{A}^+ \underline{b}$ eine partikuläre und $\underline{x} = \underline{A}^+ \underline{b} + \underline{x}_{\text{hom}}$ die allgemeine Lösung.

Der affine Lösungsraum ist somit durch

$$\mathbb{L}(\underline{b}) = \underline{A}^+ \underline{b} + \text{Bild} \left(\underline{E} - \underline{A}^+ \underline{A} \right)$$

gegeben.

Ist das LGS unlösbar, also $\text{Rang } \underline{A} < \text{Rang } \underline{A} | \underline{b}$, dann ist

$\bar{\underline{x}} = \underline{A}^+ \underline{b}$ eine Minimum-Norm-Lösung.

Beweis:

Im Fall eines lösbaren Systems ist lediglich zu zeigen, dass $\underline{x}_p = \underline{A}^+ \underline{b}$ eine Lösung ist.

Nach Voraussetzung gilt $\text{Rang } \underline{A} = \text{Rang } \underline{A} | \underline{b}$, also liegt \underline{b} in Bild \underline{A} . Damit existieren Skalare $\lambda_k \in \mathbb{C}$ mit

$$\underline{b} = \sum_{k=1}^n \lambda_k \underline{a}_k \tag{1.13}$$

Nach der Penrose-Gleichung (1) gilt

$$\begin{aligned} \underline{A} &= \underline{A} \underline{A}^+ \underline{A} \\ &= \left(\underline{A} \underline{A}^+ \right) \underline{A} \\ &= \left(\underline{A} \underline{A}^+ \right) \left[\underline{a}_1 \quad \dots \quad \underline{a}_n \right] \\ &= \left[\left(\underline{A} \underline{A}^+ \right) \underline{a}_1 \quad \dots \quad \left(\underline{A} \underline{A}^+ \right) \underline{a}_n \right] \end{aligned}$$

Also bildet $\left(\underline{\underline{A}}\underline{\underline{A}}^+\right)$ die Spalten von $\underline{\underline{A}}$ auf sich selbst ab.

$$\left(\underline{\underline{A}}\underline{\underline{A}}^+\right) \underline{\underline{a}}_k = \underline{\underline{a}}_k \quad k = 1, \dots, n$$

Daraus folgt mit (1.13) :

$$\begin{aligned} \underline{\underline{A}} \underline{\underline{x}}_p &= \underline{\underline{A}} \left(\underline{\underline{A}}^+ \underline{\underline{b}}\right) \\ &= \left(\underline{\underline{A}}\underline{\underline{A}}^+\right) \underline{\underline{b}} \\ &= \left(\underline{\underline{A}}\underline{\underline{A}}^+\right) \sum_{k=1}^n \lambda_k \underline{\underline{a}}_k \\ &= \sum_{k=1}^n \lambda_k \left(\underline{\underline{A}}\underline{\underline{A}}^+\right) \underline{\underline{a}}_k \\ &= \sum_{k=1}^n \lambda_k \underline{\underline{a}}_k = \underline{\underline{b}} \end{aligned}$$

Zum unlösbaren System:

Wegen $\underline{\underline{A}}\underline{\underline{x}} \in \text{Bild } \underline{\underline{A}}$ gilt für alle $\underline{\underline{x}} \in \mathbb{C}^n$:

$$\left\| \tilde{\underline{\underline{b}}} - \underline{\underline{b}} \right\| \leq \left\| \underline{\underline{A}}\underline{\underline{x}} - \underline{\underline{b}} \right\| ,$$

wenn $\tilde{\underline{\underline{b}}}$ die orthogonale Projektion von $\underline{\underline{b}}$ auf Bild $\underline{\underline{A}}$ ist.

Nach Satz 1.14 ist die orthogonale Projektion durch die Matrix $\underline{\underline{A}}\underline{\underline{A}}^+$ gegeben.

Aus

$$\underline{\underline{A}}\bar{\underline{\underline{x}}} = \underline{\underline{A}} \left(\underline{\underline{A}}^+ \underline{\underline{b}}\right) = \left(\underline{\underline{A}}\underline{\underline{A}}^+\right) \underline{\underline{b}} = \tilde{\underline{\underline{b}}}$$

folgt die Behauptung, dass $\bar{\underline{\underline{x}}} = \underline{\underline{A}}^+ \underline{\underline{b}}$ eine Minimum-Norm-Lösung ist.

$$\left\| \underline{\underline{A}}\bar{\underline{\underline{x}}} - \underline{\underline{b}} \right\| = \min \left\{ \left\| \underline{\underline{A}}\underline{\underline{x}} - \underline{\underline{b}} \right\| \mid \underline{\underline{x}} \in \mathbb{C}^n \right\} .$$

□

Beispiel 1.18 Unterbestimmte lineare Gleichungssystem

(1) Gesucht sind die Lösungen von

$$x + 2y + 3z = 28$$

Hier ist $\underline{A} = \begin{bmatrix} 1 & 2 & 3 \end{bmatrix}$ und wir erhalten $\underline{x}_p = \underline{A}^+ 28 = \frac{1}{14} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} 28 = \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix}$

$$\underline{E} - \underline{A}^+ \underline{A} = \underline{E} - \frac{1}{14} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \end{bmatrix} = \underline{E} - \frac{1}{14} \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 3 & 6 & 9 \end{bmatrix} = \frac{1}{14} \begin{bmatrix} 13 & -2 & -3 \\ -2 & 10 & -6 \\ -3 & -6 & 5 \end{bmatrix}$$

Es folgt

$$\mathbb{L} = \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix} + \text{Span} \left\{ \begin{bmatrix} 13 \\ -2 \\ -3 \end{bmatrix}, \begin{bmatrix} -2 \\ 10 \\ -6 \end{bmatrix} \right\}$$

(2) Zu lösen ist $\underline{A} \underline{x} = \underline{b}$ mit

$$\underline{A} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}, \quad \underline{b} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$$

Die Matrix \underline{A} hat vollen Zeilenrang. Daher gilt

$$\underline{A}^+ = \underline{A}^* (\underline{A} \underline{A}^*)^{-1}$$

Mit

$$\underline{A} \underline{A}^* = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 3 & 2 \\ 2 & 2 \end{bmatrix}$$

und

$$(\underline{A} \underline{A}^*)^{-1} = \frac{1}{2} \begin{bmatrix} 2 & -2 \\ -2 & 3 \end{bmatrix}$$

folgt

$$\underline{A}^+ = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 0 \end{bmatrix} \frac{1}{2} \begin{bmatrix} 2 & -2 \\ -2 & 3 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 0 & 1 \\ 0 & 1 \\ 2 & -2 \end{bmatrix}$$

$$\underline{E} - \underline{A}^+ \underline{A} = \frac{1}{2} \left(2\underline{E} - \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix} \right) = \frac{1}{2} \begin{bmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Diese Matrix hat den Spaltenrang 1 und wir erhalten

$$\mathbb{L}_{hom} = \text{Kern } \underline{\underline{A}} = \text{Span} \left\{ \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \right\}$$

Eine partikuläre Lösung erhalten wir durch

$$\underline{\underline{x}}_p = \underline{\underline{A}}^+ \underline{\underline{b}} = \frac{1}{2} \begin{bmatrix} 0 & 1 \\ 0 & 1 \\ 2 & -2 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

Die allgemeine Lösung lautet somit

$$\underline{\underline{x}} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + t \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1+t \\ 1-t \\ 1 \end{bmatrix}, \quad t \in \mathbb{C}$$

(3) Beispiel für ein komplexes System

Zu lösen ist $\underline{\underline{A}} \underline{\underline{x}} = \underline{\underline{b}}$ mit

$$\underline{\underline{A}} = \begin{bmatrix} 2 & 1+i & -i \\ 1-i & 1 & -1 \end{bmatrix}, \quad \underline{\underline{b}} = \begin{bmatrix} -1 \\ -1+i \end{bmatrix}$$

Die Matrix $\underline{\underline{A}}$ hat vollen Zeilenrang. Daher gilt

$$\underline{\underline{A}}^+ = \underline{\underline{A}}^* (\underline{\underline{A}} \underline{\underline{A}}^*)^{-1}$$

Mit

$$\underline{\underline{A}} \underline{\underline{A}}^* = \begin{bmatrix} 2 & 1+i & -i \\ 1-i & 1 & -1 \end{bmatrix} \begin{bmatrix} 2 & 1+i \\ 1-i & 1 \\ i & -1 \end{bmatrix} = \begin{bmatrix} 7 & 3+4i \\ 3-4i & 4 \end{bmatrix}$$

und

$$(\underline{\underline{A}} \underline{\underline{A}}^*)^{-1} = \frac{1}{3} \begin{bmatrix} 4 & -3-4i \\ -3+4i & 7 \end{bmatrix}$$

folgt

$$\underline{\underline{A}}^+ = \begin{bmatrix} 2 & 1+i \\ 1-i & 1 \\ i & -1 \end{bmatrix} \frac{1}{3} \begin{bmatrix} 4 & -3-4i \\ -3+4i & 7 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 1+i & 1-i \\ 1 & -i \\ 3 & -3-3i \end{bmatrix}$$

Berechnung von $\mathbb{L}_{hom} = \text{Kern } \underline{\underline{A}} = \text{Bild } \underline{\underline{E}} - \underline{\underline{A}}^+ \underline{\underline{A}}$:

$$\underline{\underline{A}}^+ \underline{\underline{A}} = \frac{1}{3} \begin{bmatrix} 1+i & 1-i \\ 1 & -i \\ 3 & -3-3i \end{bmatrix} \begin{bmatrix} 2 & 1+i & -i \\ 1-i & 1 & -1 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 2 & 1+i & 0 \\ 1-i & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

d.h.

$$\underline{\underline{E}} - \underline{\underline{A}}^+ \underline{\underline{A}} = \frac{1}{3} \begin{bmatrix} 1 & -1-i & 0 \\ -1+i & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Die zweite Spalte ist das $(-1-i)$ -fache der ersten Spalte. Diese Matrix hat somit den Spaltenrang 1 und wir erhalten

$$\mathbb{L}_{hom} = \text{Kern } \underline{\underline{A}} = \text{Span} \left\{ \begin{bmatrix} 1 \\ -1+i \\ 0 \end{bmatrix} \right\}$$

Eine partikuläre Lösung erhalten wir durch

$$\underline{x}_p = \underline{\underline{A}}^+ \underline{b} = \frac{1}{3} \begin{bmatrix} 1+i & 1-i \\ 1 & -i \\ 3 & -3-3i \end{bmatrix} \begin{bmatrix} -1 \\ -1+i \end{bmatrix} = \frac{1}{3} \begin{bmatrix} -1+i \\ i \\ 3 \end{bmatrix}$$

Die allgemeine Lösung lautet somit

$$\underline{x} = \frac{1}{3} \begin{bmatrix} -1+i \\ i \\ 3 \end{bmatrix} + t \begin{bmatrix} 1 \\ -1+i \\ 0 \end{bmatrix}, \quad t \in \mathbb{C}$$

Beispiel 1.19 überbestimmte lineare Gleichungssysteme und Regression

Als Regression bezeichnen wir hier das Problem, zu einer gegebenen Punktmenge im $\{(x_1; y_1); \dots; (x_n; y_n)\} \in \mathbb{R}^2$ eine gut approximierende Kurve eines bestimmten Typs zu finden. Genauer suchen wir nach Geraden oder Parabeln (n -ten Grades), d.h. Graphen von Polynomen P mit der Eigenschaft, dass die Summe $(y_1 - P(x_1))^2 + \dots + (y_n - P(x_n))^2$ minimal wird.

Gegeben sind die Punkte $(1; 7), (2; 2), (3; 1), (0; 0)$

- (1) Welche Ursprungsgerade $y = mx$ ist die beste Näherung für die Punkte?
- (2) Welche Gerade $y = mx + b$ ist die beste Näherung für die Punkte inklusive Nullpunkt?
- (3) Welche Parabel 2.ter Ordnung liefert die beste Approximation?

- (1) Welche Ursprungsgerade $y = mx$ ist die beste Näherung für die Punkte?
Die Antwort liefert die Minimum-Norm-Lösung von

$$\underbrace{\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}}_{=\underline{A}} m = \underbrace{\begin{bmatrix} 7 \\ 2 \\ 1 \end{bmatrix}}_{=\underline{b}}.$$

Da \underline{A} vollen Spaltenrang hat, gilt

$$\underline{A}^+ = (\underline{A}^* \underline{A})^{-1} \underline{A}^* = \left(\begin{bmatrix} 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \right)^{-1} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \frac{1}{14} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

Damit erhalten wir

$$\bar{m} = \underline{A}^+ \underline{b} = \frac{1}{14} \begin{bmatrix} 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 7 \\ 2 \\ 1 \end{bmatrix} = 1$$

Die gesuchte Ursprungsgerade ist $y = x$.

- (2) Welche Gerade $y = mx + b$ ist die beste Näherung für die Punkte inklusive Nullpunkt?

Das Gleichungssystem lautet jetzt

$$\begin{bmatrix} 1 \\ 2 \\ 3 \\ 0 \end{bmatrix} m + \begin{bmatrix} b \\ b \\ b \\ b \end{bmatrix} = \begin{bmatrix} 7 \\ 2 \\ 1 \\ 0 \end{bmatrix}$$

bzw

$$\underbrace{\begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \\ 0 & 1 \end{bmatrix}}_{=\underline{A}} \underbrace{\begin{bmatrix} m \\ b \end{bmatrix}}_{=x} = \begin{bmatrix} 7 \\ 2 \\ 1 \\ 0 \end{bmatrix}$$

Die Matrix \underline{A} hat vollen Spaltenrang. Daher gilt

$$\underline{A}^+ = \left(\underline{A}^* \underline{A} \right)^{-1} \underline{A}^*$$

Mit

$$\underline{A}^* \underline{A} = \begin{bmatrix} 1 & 2 & 3 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 14 & 6 \\ 6 & 4 \end{bmatrix}$$

und

$$\left(\underline{A}^* \underline{A} \right)^{-1} = \frac{1}{20} \begin{bmatrix} 2 & -3 \\ -3 & 7 \end{bmatrix}$$

folgt

$$\begin{aligned} \underline{A}^+ &= \frac{1}{20} \begin{bmatrix} 2 & -3 \\ -3 & 7 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix} \\ &= \frac{1}{20} \begin{bmatrix} -1 & 1 & 3 & -3 \\ 4 & 1 & -2 & 7 \end{bmatrix} \end{aligned}$$

Die Minimum-Norm-Lösung lautet:

$$\begin{bmatrix} m \\ b \end{bmatrix} = \underline{A}^+ \begin{bmatrix} 7 \\ 2 \\ 1 \\ 0 \end{bmatrix} = \frac{1}{20} \begin{bmatrix} -1 & 1 & 3 & -3 \\ 4 & 1 & -2 & 7 \end{bmatrix} \begin{bmatrix} 7 \\ 2 \\ 1 \\ 0 \end{bmatrix} = \frac{1}{10} \begin{bmatrix} -1 \\ 14 \end{bmatrix}$$

Die gesuchte Gerade ist somit:

$$y = \frac{14 - x}{10}$$

- (3) Welche Parabel 2.ter Ordnung liefert die beste Approximation?
Gesucht sind die Koeffizienten des Polynoms

$$y(x) = a_2x^2 + a_1x + a_0$$

Dazu bestimmen wir die Minimum-Norm-Lösung von

$$\begin{aligned} a_2 + a_1 + a_0 &= 7 \\ 4a_2 + 2a_1 + a_0 &= 2 \\ 9a_2 + 3a_1 + a_0 &= 1 \\ a_0 &= 0 \end{aligned}$$

Die Matrix $\underline{\underline{A}} = \begin{bmatrix} 1 & 1 & 1 \\ 4 & 2 & 1 \\ 9 & 3 & 1 \\ 0 & 0 & 1 \end{bmatrix}$ hat vollen Spaltenrang. Somit gilt $\underline{\underline{A}}^+ = (\underline{\underline{A}}^* \underline{\underline{A}})^{-1} \underline{\underline{A}}^*$

$$\underline{\underline{A}}^* \underline{\underline{A}} = \begin{bmatrix} 1 & 4 & 9 & 0 \\ 1 & 2 & 3 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 4 & 2 & 1 \\ 9 & 3 & 1 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 98 & 36 & 14 \\ 36 & 14 & 6 \\ 14 & 6 & 4 \end{bmatrix} = 2 \begin{bmatrix} 49 & 18 & 7 \\ 18 & 7 & 3 \\ 7 & 3 & 2 \end{bmatrix}$$

und

$$(\underline{\underline{A}}^* \underline{\underline{A}})^{-1} = \frac{1}{20} \begin{bmatrix} 5 & -15 & 5 \\ -15 & 49 & -21 \\ 5 & -21 & 19 \end{bmatrix}$$

Es folgt

$$\begin{aligned} \underline{\underline{A}}^+ &= \frac{1}{20} \begin{bmatrix} 5 & -15 & 5 \\ -15 & 49 & -21 \\ 5 & -21 & 19 \end{bmatrix} \begin{bmatrix} 1 & 4 & 9 & 0 \\ 1 & 2 & 3 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix} \\ &= \frac{1}{20} \begin{bmatrix} -5 & -5 & 5 & 5 \\ 13 & 17 & -9 & -21 \\ 3 & -3 & 1 & 19 \end{bmatrix} \end{aligned}$$

Die Minimum-Norm-Lösung lautet:

$$\begin{aligned} \begin{bmatrix} a_2 \\ a_1 \\ a_0 \end{bmatrix} &= \underline{\underline{A^+}} \begin{bmatrix} 7 \\ 2 \\ 1 \\ 0 \end{bmatrix} \\ &= \frac{1}{20} \begin{bmatrix} -5 & -5 & 5 & 5 \\ 13 & 17 & -9 & -21 \\ 3 & -3 & 1 & 19 \end{bmatrix} \begin{bmatrix} 7 \\ 2 \\ 1 \\ 0 \end{bmatrix} \\ &= \frac{1}{5} \begin{bmatrix} -10 \\ 29 \\ 4 \end{bmatrix} \end{aligned}$$

Die gesuchte Parabel ist somit der Graph von:

$$y(x) = \frac{-10x^2 + 29x + 4}{5}$$