VII.5. Positivität.

VII.5.1. Definition (Positivität). Eine *selbstadjungierte* lineare Abbildung $\varphi\colon V \to V$ auf einem Euklidischen oder untären Vektorraum V heißt *semipositiv*, falls

$$\langle v, \varphi(v)\rangle \geq 0,$$

für alle $v \in V$. Wir schreiben in diesem Fall $\varphi \geq 0$. Gilt sogar

$$\langle v, \varphi(v)\rangle > 0,$$

für alle $0 \neq v \in V$, so wird φ *positiv* genannt und wir schreiben $\varphi > 0$.

VII.5.2. Lemma. *Für eine lineare Abbildung $\varphi\colon V \to V$ auf einem endlich dimensionalen Euklidischen Vektorraum V sind äquivalent:*

(a) φ ist (semi)positiv.
(b) $\langle v, \varphi(w)\rangle$ ist eine positiv (semi)definite symmetrische Bilinearform auf V.
(c) Die Matrix $[\varphi]_{BB}$ ist positiv (semi)definit, für eine (und dann jede) Orthonormalbasis B von V.

Für eine lineare Abbildung $\varphi\colon V \to V$ auf einem endlich dimensionalen unitären Vektorraum V sind äquivalent:

(a) φ ist (semi)positiv.
(b) $\langle v, \varphi(w)\rangle$ ist eine positiv (semi)definite Hermitesche Form auf V.
(c) Die Matrix $[\varphi]_{BB}$ ist positiv (semi)definit, für eine (und dann jede) Orthonormalbasis B von V.

Beweis. Wir werden dies nur für unitäre Vektorräume V zeigen, der Euklidische Fall lässt sich völlig analog behandeln. Eine lineare Abbildung $\varphi\colon V \to V$ ist genau dann selbstadjungiert, wenn $h(v, w) = \langle v, \varphi(w)\rangle$ eine Hermitesche Form auf V definiert, denn offensichtlich ist $\langle v, \varphi(w)\rangle$ linear in w und es gilt $\overline{h(w, v)} = \overline{\langle w, \varphi(v)\rangle} = \langle \varphi(v), w\rangle = \langle v, \varphi^*(w)\rangle$. Daraus erhalten wir sofort die Äquivalenz (a)$\Leftrightarrow$(b). Ist $B = (b_1, \dots, b_n)$ eine Orthonormalbasis von V, dann gilt $[h]_B = [\varphi]_{BB}$, denn nach Proposition VII.2.22 haben wir

$$\varphi(b_j) = \sum_{i=1}^n \langle b_i, \varphi(b_j)\rangle b_i = \sum_{i=1}^n h(b_i, b_j) b_i,$$

und daher $\big([\varphi]_{BB}\big)_{ij} = h(b_i, b_j) = \big([h]_B\big)_{ij}$. Dies zeigt die Äquivalenz (b)$\Leftrightarrow$(c), vgl. Definition VII.1.29. $\square$

VII.5.3. Beispiel. Für $a \in M_{1\times 1}(\mathbb{C}) = \mathbb{C}$ betrachte die lineare Abbildung $\psi_a\colon \mathbb{C} \to \mathbb{C}$, $\psi_a(x) = ax$. Diese Abbildung ist genau dann selbstadjungiert, wenn $a \in \mathbb{R}$. Sie ist genau dann (semi)positiv, wenn $a > 0$ bzw. $a \geq 0$ gilt. Die Abbildung ψ_a ist genau dann unitär, wenn $|a| = 1$.

VII.5.4. Beispiel. Die lineare Abbildung

$$\psi\colon \mathbb{R}^3 \to \mathbb{R}^3, \qquad \psi(x) = Ax, \qquad \text{wobei} \qquad A = \begin{pmatrix} 3 & 1 & 1 \\ 1 & 4 & 1 \\ 1 & 1 & 5 \end{pmatrix}$$

ist positiv (bezüglich des standard inneren Produkts auf $\mathbb{R}^3$), denn die Matrix A ist symmetrisch und positiv definit, da

$$3 > 0, \qquad \begin{vmatrix} 3 & 1 \\ 1 & 4 \end{vmatrix} = 11 > 0, \qquad \begin{vmatrix} 3 & 1 & 1 \\ 1 & 4 & 1 \\ 1 & 1 & 5 \end{vmatrix} = 50 > 0,$$

vgl. das Sylvesterkriterium in Satz VII.1.42.

VII.5.5. Lemma. *Seien φ, φ_1 und φ_2 lineare Abbildungen auf einem endlich dimensionalen Euklidischen oder unitären Vektorraum V und $\lambda \in \mathbb{R}$. Dann gilt:*

(a) Ist $\varphi_1 \geq 0$ und $\varphi_2 \geq 0$, dann auch $\varphi_1 + \varphi_2 \geq 0$.
(b) Ist $\varphi_1 \geq 0$ und $\varphi_2 > 0$, dann auch $\varphi_1 + \varphi_2 > 0$.
(c) Ist $\lambda \geq 0$ und $\varphi \geq 0$, dann auch $\lambda\varphi \geq 0$.
(d) Ist $\lambda > 0$ und $\varphi > 0$, dann auch $\lambda\varphi > 0$.
(e) Ist $\varphi \geq 0$ und $\psi\colon V \to W$ linear, dann auch $\psi\varphi\psi^* \geq 0$.
(f) Ist $\varphi > 0$ und $\psi\colon V \to W$ surjektiv und linear, dann auch $\psi\varphi\psi^* > 0$.
(g) Ist $\varphi \geq 0$ und $-\varphi \geq 0$, dann gilt schon $\varphi = 0$.
(h) Es gilt $\varphi > 0$ genau dann wenn $\varphi \geq 0$ und φ injektiv (invertierbar) ist.

Insbesondere ist die Menge der (semi)positiven Abbildungen auf V konvex.

Beweis. Ist $\varphi_1 \geq 0$ und $\varphi_2 \geq 0$, dann ist $\varphi_1 + \varphi_2$ als Summe zweier selbstadjungierter Abbildungen wieder selbstadjungiert und es gilt $\langle v, (\varphi_1 + \varphi_2)(v)\rangle = \langle v, \varphi_1(v)\rangle + \langle v, \varphi_2(v)\rangle \geq 0$, also $\varphi_1 + \varphi_2 \geq 0$. Ist $\varphi_2 > 0$, dann erhalten wir analog $\varphi_1 + \varphi_2 > 0$. Dies zeigt die ersten beiden Aussagen. Sei nun $\lambda \geq 0$ und $\varphi \geq 0$. Dann ist auch $\lambda\varphi$ selbstadjungiert und wir erhalten $\langle v, (\lambda\varphi)(v)\rangle = \lambda\langle v, \varphi(v)\rangle \geq 0$, also $\lambda\varphi \geq 0$. Ist $\lambda > 0$ und $\varphi > 0$ dann erhalten wir mit dem gleichen Argument $\lambda\varphi > 0$. Behauptung (e) folgt aus $\big(\psi\varphi\psi^*\big)^* = \psi^{**}\varphi^*\psi^* = \psi\varphi\psi^*$ und $\langle(\psi\varphi\psi^*)(w), w\rangle = \langle\varphi(\psi^*(w)), \psi^*(w)\rangle \geq 0$, für alle $w \in W$. Für surjektives ψ is ψ^* injektiv und wir erhalten wir daraus auch (f). Um auch (g) einzusehen, sei nun $\varphi \geq 0$ und $-\varphi \geq 0$, d.h. $\langle v, \varphi(v)\rangle = 0$, für alle $v \in V$. Mit der Polarisierungsidentität folgt $\langle w, \varphi(v)\rangle = 0$, für alle $v, w \in V$, und daher $\varphi = 0$. Ad (h): Ist $\varphi > 0$, dann auch $\ker(\varphi) = \{0\}$, also $\varphi\colon V \to V$ injektiv und daher invertierbar. □

Positivität erlaubt es auf der Menge der selbstadjungierten Abbildungen eine Ordnungsrelation zu definieren. Sind $\varphi_1, \varphi_2\colon V \to V$ zwei selbstadjungierte Abbildungen dann schreiben wir $\varphi_2 \geq \varphi_1$ falls $\varphi_2 - \varphi_1 \geq 0$. Mit Hilfe des vorangehenden Lemmas erhalten wir sofort:

(a) $\varphi \geq \varphi$.
(b) Ist $\varphi_1 \geq \varphi_2$ und $\varphi_2 \geq \varphi_1$, dann gilt schon $\varphi_1 = \varphi_2$.

(c) Ist $\varphi_1 \geq \varphi_2$ und $\varphi_2 \geq \varphi_3$, dann auch $\varphi_1 \geq \varphi_3$.
(d) Ist $\varphi_1 \geq \tilde{\varphi}_1$ und $\varphi_2 \geq \tilde{\varphi}_2$, dann auch $\varphi_1 + \varphi_2 \geq \tilde{\varphi}_1 + \tilde{\varphi}_2$.
(e) Ist $\lambda \geq 0$ ein Skalar und $\varphi \geq \psi$, dann auch $\lambda\varphi \geq \lambda\psi$.
(f) Ist ψ linear und $\varphi_2 \geq \varphi_1$, dann auch $\psi\varphi_2\psi^* \geq \psi\varphi_1\psi^*$.

Dadurch wird die Menge der selbstadjungierten Abbildungen nur teilweise geordnet, i.A. sind zwei selbstadjungierte Abbildungen φ und ψ nicht vergleichbar, etwa gilt für $A = \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$ weder $A \geq 0$ noch $0 \geq A$. Auch ist diese partielle Ordnung nicht mit der Multiplikation verträglich, d.h. aus $\varphi \geq 0$ und $\psi \geq 0$ lässt sich *nicht* $\varphi\psi \geq 0$ schließen, letztere Abbildung wird ja i.A. nicht einmal selbstadjungiert sein. Allerdings sind die semipositiven selbstadjungierten Abbildungen genau jene, die sich als Quadrate schreiben lassen. Genauer haben wir:

VII.5.6. Satz. *Für eine lineare Abildung $\varphi\colon V \to V$ auf einem endlich dimensionalen Euklidischen oder unitären Vektorraum V sind äquivalent:*

(a) φ ist semipositiv.
(b) φ ist selbstadjungiert und alle Eigenwerte sind größer oder gleich Null.
(c) Es existiert eine Orthonormalbasis B von V, sodass $[\varphi]_{BB}$ eine Diagonalmatrix ist, deren Diagonaleinträge alle größer oder gleich Null sind.
(d) Es existiert eine semipositive Abbildung $\rho\colon V \to V$, sodass $\rho^2 = \varphi$.
(e) Es existiert eine selbstadjungierte Abbildung $\rho\colon V \to V$ mit $\rho^2 = \varphi$.
(f) Es existiert eine lineare Abbildung $\rho\colon V \to V$ mit $\rho^*\rho = \varphi$.

Für $\varphi \geq 0$ ist die semipositive Abbildung $\rho^* = \rho \geq 0$ in (d) eindeutig bestimmt und wird mit $\sqrt{\varphi}$ oder $\varphi^{1/2}$ bezeichnet. Ist $\varphi \geq 0$ dann gilt:

(g) Ist $\varphi = \lambda_1\pi_1 + \cdots + \lambda_k\pi_k$ die Spektralzerlegung von φ, so ist die Spektralzerlegung von $\sqrt{\varphi}$ durch $\sqrt{\varphi} = \sqrt{\lambda_1}\pi_1 + \cdots + \sqrt{\lambda_k}\pi_k$ gegeben. Insbesondere ist λ genau dann Eigenwert von φ wenn $\sqrt{\lambda}$ Eigenwert von $\sqrt{\varphi}$ ist und die entsprechenden Eigenräume stimmen überein.
(h) $\operatorname{img}(\sqrt{\varphi}) = \operatorname{img}(\varphi)$
(i) $\ker(\sqrt{\varphi}) = \ker(\varphi)$.
(j) $\varphi > 0 \Leftrightarrow \sqrt{\varphi} > 0$.
(k) Ist $\psi\colon V \to W$ eine Isometrie, so gilt $\sqrt{\psi\varphi\psi^*} = \psi\sqrt{\varphi}\psi^*$.
(l) Ist $\psi\colon V \to V$ linear und $\psi\varphi = \varphi\psi$, dann auch $\psi\sqrt{\varphi} = \sqrt{\varphi}\psi$.

Beweis. Ad (a)⇒(b): Sei also $\varphi \geq 0$ und λ Eigenwert von φ. Es existiert daher $0 \neq v \in V$ mit $\varphi(v) = \lambda v$, wir erhalten

$$\lambda\|v\|^2 = \lambda\langle v, v\rangle = \langle v, \lambda v\rangle = \langle v, \varphi(v)\rangle \geq 0,$$

und somit $\lambda \geq 0$. Die Implikation (b)⇒(c) folgt aus dem Spektralsatz, siehe Satz VII.3.11. Ad (c)⇒(d): Sei also B eine Orthonormalbasis von V, sodass

$$[\varphi]_{BB} = \begin{pmatrix} a_1 & & \\ & \ddots & \\ & & a_n \end{pmatrix} \qquad \text{und} \qquad a_i \geq 0.$$

Definieren wir $\rho\colon V \to V$ durch

$$[\rho]_{BB} := \begin{pmatrix} \sqrt{a_1} & & \\ & \ddots & \\ & & \sqrt{a_n} \end{pmatrix},$$

dann gilt offensichtlich $\rho^* = \rho \geq 0$ und $\rho^2 = \varphi$. Die Implikationen (d)$\Rightarrow$(e) und (e)$\Rightarrow$(f) sind trivial. Nun zu (f)$\Rightarrow$(a): Jede Abbildung der Form $\varphi = \rho^*\rho$ ist selbstadjungiert, $\varphi^* = (\rho^*\rho)^* = \rho^*\rho^{**} = \rho^*\rho = \varphi$, und semipositiv, denn $\langle v, \varphi(v)\rangle = \langle v, (\rho^*\rho)(v)\rangle = \langle \rho(v), \rho(v)\rangle \geq 0$. Damit ist die Äquivalenz der Eigenschaften (a) bis (f) gezeigt.

Um die Eindeutigkeit der Wurzel zu zeigen, sei $\psi\colon V \to V$ mit $\psi \geq 0$ und $\psi^2 = \varphi$. Dann gilt $\psi\varphi = \psi\psi^2 = \psi^2\psi = \varphi\psi$, also bewahrt ψ jeden Eigenraum von φ. Bezeichne die Einschränkung von ψ auf einen solchen Eigenraum mit $\psi_i := \psi|_{E^\varphi_{\lambda_i}} \colon E^\varphi_{\lambda_i} \to E^\varphi_{\lambda_i}$. Beachte $\psi_i^* = \psi_i \geq 0$ und $\psi_i^2 = \lambda_i \operatorname{id}_{E^\varphi_{\lambda_i}}$. Da ψ_i diagonalisierbar ist, folgt daraus $\psi_i = \sqrt{\lambda_i} \operatorname{id}_{E^\varphi_{\lambda_i}}$, also $\psi|_{E^\varphi_{\lambda_i}} = \sqrt{\varphi}|_{E^\varphi_{\lambda_i}}$. Da dies für jeden Eigenraum gilt, folgt $\psi = \sqrt{\varphi}$, die Wurzel ist daher eindeutig.

Die verbleibenden Eigenschaften der Wurzel sind nun offensichtlich. □

Wir definieren die Wurzel einer positiven Matrix mit Hilfe der damit assoziierten linearen Abbildung.

VII.5.7. Beispiel. Für Diagonalmatrizen mit positiven Einträgen gilt:

$$\begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}^{1/2} = \begin{pmatrix} \sqrt{\lambda_1} & & \\ & \ddots & \\ & & \sqrt{\lambda_n} \end{pmatrix}, \qquad \lambda_i \geq 0.$$

VII.5.8. Beispiel. Wir wollen die Wurzel der positiven Matrix

$$A = \begin{pmatrix} 10 & 6 \\ 6 & 10 \end{pmatrix}$$

berechnen. Wir bestimmen zunächst eine orthogonale Matrix $U \in \mathrm{O}_2$ und eine Diagonalmatrix D, sodass $A = UDU^{-1}$:

$$U = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}, \qquad D = \begin{pmatrix} 16 & 0 \\ 0 & 4 \end{pmatrix}, \qquad A = UDU^{-1} = UDU^t.$$

Es gilt daher

$$\sqrt{A} = \sqrt{UDU^{-1}} = U\sqrt{D}U^{-1} = U\sqrt{D}U^t = U\begin{pmatrix} 4 & 0 \\ 0 & 2 \end{pmatrix}U^t = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}.$$

VII.5.9. Beispiel. Wir wollen die Wurzel der positiven Matrix

$$A = \begin{pmatrix} 3 & 1 & -2 \\ 1 & 3 & -2 \\ -2 & -2 & 6 \end{pmatrix}$$

bestimmen. Wir berechnen zunächst eine orthogonale Matrix $U \in \mathrm{O}_3$ und eine Diagonalmatrix D, sodass $A = UDU^{-1} = UDU^t$:

$$U = \begin{pmatrix} 1/\sqrt{3} & -1/\sqrt{2} & 1/\sqrt{6} \\ 1/\sqrt{3} & 1/\sqrt{2} & 1/\sqrt{6} \\ 1/\sqrt{3} & 0 & -2/\sqrt{6} \end{pmatrix}, \qquad D = \begin{pmatrix} 2 & & \\ & 2 & \\ & & 8 \end{pmatrix}.$$

Für die Wurzel von A erhalten wir

$$\sqrt{A} = U\sqrt{D}U^t = U \begin{pmatrix} \sqrt{2} & & \\ & \sqrt{2} & \\ & & 2\sqrt{2} \end{pmatrix} U^t = \frac{\sqrt{2}}{6} \begin{pmatrix} 7 & 1 & -2 \\ 1 & 7 & -2 \\ -2 & -2 & 10 \end{pmatrix}.$$

VII.5.10. Satz (Polarzerlegung). *Sei $\varphi\colon V \to W$ eine lineare Abbildung zwischen endlich dimensionalen Euklidischen oder unitären Vektorräumen.*

(a) Ist $\dim(V) \leq \dim(W)$, dann lässt sich φ in der Form $\varphi = \psi\rho$ schreiben, wobei $\rho\colon V \to V$ semipositiv und $\psi\colon V \to W$ eine Isometrie ist. Dabei ist ρ eindeutig bestimmt, $\rho = \sqrt{\varphi^*\varphi}$. Ist φ injektiv, dann gilt $\rho > 0$ und $\psi = \varphi\rho^{-1}$, in diesem Fall ist also auch ψ eindeutig bestimmt.

(b) Ist $\dim(V) \geq \dim(W)$, dann lässt sich φ in der Form $\varphi = \rho\psi^*$ schreiben, wobei $\rho\colon W \to W$ semipositiv und $\psi\colon W \to V$ eine Isometrie ist. Dabei ist ρ eindeutig bestimmt, $\rho = \sqrt{\varphi\varphi^*}$. Ist φ surjektiv, dann gilt $\rho > 0$ und $\psi = \varphi^*\rho^{-1}$, in diesem Fall ist also auch ψ eindeutig bestimmt.

Beweis. Ad (a): Ist $\varphi = \psi\rho$ eine Darstellung wie in (a), dann folgt

$$\varphi^*\varphi = (\psi\rho)^*(\psi\rho) = \rho^*\psi^*\psi\rho = \rho^*\rho = \rho^2,$$

also $\rho = \sqrt{\varphi^*\varphi}$. Ist φ injektiv, dann gilt $\ker(\rho) = \ker(\varphi^*\varphi) = \ker(\varphi) = 0$, also ist ρ invertierbar und $\psi = \varphi\rho^{-1}$. Damit sind die Eindeutigkeitsaussagen gezeigt.

Um die Existenz der Darstellung zu zeigen, sei zunächst $\varphi\colon V \to W$ injektiv. Dann ist $\rho := \sqrt{\varphi^*\varphi}$ invertierbar und wir definieren $\psi := \varphi\rho^{-1}$. Da ρ selbstadjungiert ist, gilt auch $(\rho^{-1})^* = \rho^{-1}$ und somit

$$\psi^*\psi = (\varphi\rho^{-1})^*(\varphi\rho^{-1}) = \rho^{-1}\varphi^*\varphi\rho^{-1} = \rho^{-1}\rho^2\rho^{-1} = \mathrm{id}_V,$$

also ist ψ eine Isometrie. Damit ist die Existenz der gesuchten Zerlegung für injektive φ gezeigt.

Sei nun $\varphi\colon V \to W$ beliebig. Da die Einschränkung $\varphi|_{\ker(\varphi)^\perp}\colon \ker(\varphi)^\perp \to W$ injektiv ist, erhalten wir aus dem eben Bewiesenen eine (semi)positive Abbildung $\tilde\rho\colon \ker(\varphi)^\perp \to \ker(\varphi)^\perp$ und eine Isometrie $\tilde\psi\colon \ker(\varphi)^\perp \to W$, sodass

$$\tilde\psi\tilde\rho = \varphi|_{\ker(\varphi)^\perp}. \tag{VII.26}$$

Aus bekannten Dimensionsformeln folgt

$$\begin{aligned} \dim\ker(\varphi) &= \dim V - \dim\operatorname{img}(\varphi) \\ &= \dim V - \dim W + \dim\operatorname{img}(\varphi)^\perp \leq \dim\operatorname{img}(\varphi)^\perp, \end{aligned}$$

denn nach Voraussetzung ist $\dim V - \dim W \leq 0$. Es existiert daher eine Orthonormalbasis $v_1, \dots, v_k$ von $\ker(\varphi)$ und ein Orthonormalsystem $w_1, \dots, w_k$ von $\mathrm{img}(\varphi)^\perp$. Die Abbildung $\tilde\psi\colon \ker(\varphi)^\perp \to W$ lässt sich daher zu einer Isometrie $\psi\colon V \to W$ fortsetzen, $\psi|_{\ker(\varphi)^\perp} = \tilde\psi$, $\psi(v_i) = w_i$, $i = 1, \dots, k$. Aus (VII.26) folgt:

$$\psi\iota\tilde\rho = \varphi|_{\ker(\varphi)^\perp},$$

wobei $\iota\colon \ker(\varphi)^\perp \to V$ die kanonische Inklusion bezeichnet. Daraus erhalten wir

$$\psi\iota\tilde\rho\pi = \varphi,$$

wobei $\pi\colon V \to \ker(\varphi)^\perp$ die Orthogonalprojektion bezeichnet. Es bleibt nur noch zu zeigen, dass die Abbildung $\rho := \iota\tilde\rho\pi\colon V \to V$ semipositiv ist. Da jedoch $\pi = \iota^*$, folgt dies aus Lemma VII.5.5(e).

Wir erhalten (b) indem wir (a) auf $\varphi^*\colon W \to V$ anwenden. □

VII.5.11. Korollar (Polarzerlegung für Matrizen).

a) Jede Matrix $A \in M_{n\times n}(\mathbb{C})$ lässt sich in der Form $A = UR$ schreiben, wobei $R^* = R \geq 0$ und $U \in \mathrm{U}_n$. Dabei ist R eindeutig bestimmt, $R = \sqrt{A^*A}$. Für invertierbares A gilt $R > 0$ und $U = AR^{-1}$, in diesem Fall ist also auch U eindeutig bestimmt.

b) Jede Metraix $A \in M_{n\times n}(\mathbb{R})$ lässt sich in der Form $A = UR$ schreiben, wobei $R = R^t \geq 0$ und $U \in \mathrm{O}_n$. Dabei ist R eindeutig bestimmt, $R = \sqrt{A^tA}$. Für invertierbares A gilt $R > 0$ und $U = AR^{-1}$, in diesem Fall ist also auch U eindeutig bestimmt.

VII.5.12. Beispiel. Ist $0 \neq a \in \mathbb{C}$, dann hat die Polarzerlegung der Matrix $A = (a) \in M_{1\times 1}(\mathbb{C})$ die Form:

$$A = UR, \qquad U = \left(\tfrac{a}{|a|}\right), \qquad R = (|a|).$$

Die Polarzerlegung von Matrizen kann daher als Verallgemeinerung der klassischen Polardarstellung verstanden werden.

VII.5.13. Beispiel. Die Polarzerlegung einer Diagonalmatrix

$$A = \begin{pmatrix} a_1 & & \\ & \ddots & \\ & & a_n \end{pmatrix}, \qquad a_i \in \mathbb{C},$$

ist $A = UR$, wobei

$$U = \begin{pmatrix} \frac{a_1}{|a_1|} & & \\ & \ddots & \\ & & \frac{a_n}{|a_n|} \end{pmatrix} \qquad \text{und} \qquad R = \begin{pmatrix} |a_1| & & \\ & \ddots & \\ & & |a_n| \end{pmatrix}.$$

Ist $a_j = 0$, dann ist $a_j/|a_j|$ nicht definiert und der entsprechende Diagonaleintrag in U kann beliebig (mit Absolutbetrag Eins) gewählt werden.

VII.5.14. Beispiel. Wir wollen die Polarzerlegung der invertierbaren Matrix

$$A = \begin{pmatrix} 13 & -15 \\ 9 & 5 \end{pmatrix}$$

bestimmen. Wir berechnen

$$A^*A = \begin{pmatrix} 13 & 9 \\ -15 & 5 \end{pmatrix} \begin{pmatrix} 13 & -15 \\ 9 & 5 \end{pmatrix} = 50 \begin{pmatrix} 5 & -3 \\ -3 & 5 \end{pmatrix}.$$

Durch Bestimmen einer Orthonormalbasis von Eigenvektoren der Matrix A^*A erhalten wir eine orthogonale Matrix $V \in O_2$ und eine Diagonalmatrix D, sodass:

$$A^*A = VDV^{-1} = VDV^t, \qquad V = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \qquad D = 100 \begin{pmatrix} 1 & \\ & 4 \end{pmatrix}.$$

Für die Wurzel folgt

$$\begin{aligned} R = (A^*A)^{1/2} &= VD^{1/2}V^t \\ &= \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} 10 \begin{pmatrix} 1 & \\ & 2 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = 5 \begin{pmatrix} 3 & -1 \\ -1 & 3 \end{pmatrix}. \end{aligned}$$

Nach Konstruktion ist

$$U := AR^{-1} = \begin{pmatrix} 13 & -15 \\ 9 & 5 \end{pmatrix} \frac{1}{40} \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix} = \frac{1}{5} \begin{pmatrix} 3 & -4 \\ 4 & 3 \end{pmatrix}$$

eine orthogonale Matrix, und daher $A = UR$ die gesuchte Polarzerlegung von A.

VII.5.15. Satz (Singulärwertzerlegung). *Sei $\varphi\colon V \to W$ eine lineare Abbildung zwischen Euklidischen oder unitären Vektorräumen endlicher Dimension. Dann existieren eine Orthonormalbasis B von V, eine Orthonormalbasis C von W und $\sigma_1, \dots, \sigma_k > 0$, sodass*

$$[\varphi]_{CB} = \begin{pmatrix} \sigma_1 & & & 0 & \cdots & 0 \\ & \ddots & & \vdots & & \vdots \\ & & \sigma_k & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 \end{pmatrix}. \tag{VII.27}$$

Dabei sind $\sigma_1, \dots, \sigma_k$, die sogenannten Singulärwerte *von φ, bis auf ihre Reihenfolge eindeutig bestimmt. Es gilt nämlich $k = \operatorname{rank}(\varphi^*\varphi) = \operatorname{rank}(\varphi\varphi^*)$ und $\sigma_1^2, \dots, \sigma_k^2$ sind die nicht-trivialen Eigenwerte von $\varphi^*\varphi$ und gleichzeitig auch die nicht-trivialen Eigenwerte von $\varphi\varphi^*$.*

Darüber hinaus ist $\sigma > 0$ genau dann Singulärwert von φ, wenn Einheitsvektoren $v \in V$ und $w \in W$ existieren, $\|v\| = 1 = \|w\|$, sodass $\varphi(v) = \sigma w$ und $\varphi^*(w) = \sigma v$.

Beweis. Da $\varphi^*\varphi \geq 0$, existiet eine Orthonormalbasis $B = (b_1, \dots, b_n)$ aus Eigenvektoren, d.h. $(\varphi^*\varphi)(b_i) = \lambda_i b_i$, und für die Eigenwerte gilt $\lambda_i \geq 0$. Durch

Umbenennen der Basiselemente können wir daher $\lambda_1, \dots, \lambda_k > 0$ und $\lambda_{k+1} = \dots = \lambda_n = 0$ erreichen. Wir setzen

$$\sigma_i := \sqrt{\lambda_i} \qquad \text{und} \qquad c_i := \tfrac{1}{\sigma_i}\varphi(b_i), \qquad i = 1, \dots, k.$$

Beachte, dass die Vektoren $c_1, \dots, c_k$ ein Orthonormalsystem in W bilden, denn

$$\begin{aligned}\langle c_i, c_j\rangle = \tfrac{1}{\sigma_i\sigma_j}\langle\varphi(b_i), \varphi(b_j)\rangle = \tfrac{1}{\sigma_i\sigma_j}\langle b_i, (\varphi^*\varphi)(b_j)\rangle \\ = \tfrac{1}{\sigma_i\sigma_j}\langle b_i, \lambda_j b_j\rangle = \tfrac{\lambda_j}{\sigma_i\sigma_j}\langle b_i, b_j\rangle = \tfrac{\lambda_j}{\sigma_i\sigma_j}\delta_{ij} = \delta_{ij}.\end{aligned}$$

Dieses lässt sich zu einer Orthonormalbasis, $C = (c_1, \dots, c_k, c_{k+1}, \dots, c_m)$ von W ergänzen. Nach Konstruktion hat die Matrixdarstellung von φ die Form (VII.27). Die restlichen Behauptungen sind nun offensichtlich. □

Für Matrizen erhalten wir daraus:

VII.5.16. Korollar (Singulärwertzerlegung für Matrizen).
a) Jede reelle Matrix $A \in M_{m\times n}(\mathbb{R})$ lässt sich in der Form $A = U\Sigma V^t$ schreiben, wobei $U \in \mathrm{O}_m$, $V \in \mathrm{O}_n$ und Σ wie in (VII.27).
b) Jede komplexe Matrix $A \in M_{m\times n}(\mathbb{C})$ lässt sich in der Form $A = U\Sigma V^*$ schreiben, wobei $U \in \mathrm{U}_m$, $V \in \mathrm{U}_n$ und Σ wie in (VII.27).

Beweis. Wir zeigen nur b), der reelle Fall lässt sich analog behandeln. Es bezeichne $\psi\colon \mathbb{C}^n \to \mathbb{C}^m$, $\psi(x) = Ax$, die mit A assoziierte lineare Abbildung. Nach Satz VII.5.15 existieren eine Orthonormalbasis $B = (b_1, \dots, b_n)$ von $\mathbb{C}^n$ und eine Orthonormalbasis $C = (c_1, \dots, c_m)$ von $\mathbb{C}^m$, sodass die Matrixdarstellung $\Sigma := [\psi]_{CB}$ die Gestalt (VII.27) hat. Da B und C Orthonormalbasen bilden, sind die Basiswechselmatrizen $V := T_{EB} = (b_1|\cdots|b_n)$ und $U := T_{\tilde{E}C} = (c_1|\cdots|c_m)$ unitär, wobei E die Standardbasis con $\mathbb{C}^n$ und $\tilde{E}$ die Standardbasis von $\mathbb{C}^m$ bezeichnen. Es gilt daher $A = [\psi]_{\tilde{E}E} = T_{\tilde{E}C}[\psi]_{CB}T_{EB}^{-1} = U\Sigma V^{-1} = U\Sigma V^*$. □

Beachte, dass die orthogonalen/unitären Matrizen U und V i.A. nicht eindeutig bestimmt sind.

VII.5.17. Beispiel. Wir wollen eine Singulärwertzerlegung der Matrix

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

bestimmen. Wir gehen genau wie im Beweis des vorangehenden Satzes vor und berechnen zunächst:

$$A^*A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 2 \\ 0 & 2 & 2 \end{pmatrix}.$$

Es ist dann

$$b_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \qquad b_2 = \begin{pmatrix} 0 \\ 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}, \qquad b_3 = \begin{pmatrix} 0 \\ 1/\sqrt{2} \\ -1/\sqrt{2} \end{pmatrix}$$

eine Orthonormalbasis aus Eigenvektoren von A^*A mit Eigenwerten $1, 4, 0$. Somit

$$\sigma_1 = 1, \quad \sigma_2 = 2, \quad c_1 = \tfrac{1}{\sigma_1} Ab_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad c_2 = \tfrac{1}{\sigma_2} Ab_2 = \begin{pmatrix} 0 \\ 1/\sqrt{2} \\ -1/\sqrt{2} \\ 0 \end{pmatrix}.$$

Wir ergänzen dies durch

$$c_3 = \begin{pmatrix} 0 \\ 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{pmatrix}, \qquad c_4 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

zu einer Orthonormalbasis und erhalten die Singulärwertzerlegung

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & -1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1/\sqrt{2} & 1/\sqrt{2} \\ 0 & 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix}^*.$$

Wir wollen diesen Abschnitt mit einigen Bemerkungen zu anderen Funktionen von Matrizen beenden. Wir erläutern dies an einem der wichtigsten Beispiele, der *Exponentialfunktion.* Für jede quadratische Matrix A setzen wir:

$$\exp(A) = e^A = \sum_{n=0}^{\infty} \frac{1}{n!} A^n. \tag{VII.28}$$

Mit Hilfe der weiter oben besprochenen Operatornorm erhalten wir

$$\sum_{n=0}^{\infty} \frac{1}{n!} ||A^n\| \leq \sum_{n=0}^{\infty} \frac{1}{n!} \|A\|^n = e^{\|A\|} < \infty$$

also konvergiert die Reihe (VII.28) für jede quadratische Matrix A und es gilt

$$\|e^A\| \leq e^{\|A\|}.$$

Etwa gilt für Diagonalmatrizen:

$$\exp \begin{pmatrix} a_1 & & \\ & \ddots & \\ & & a_n \end{pmatrix} = \begin{pmatrix} e^{a_1} & & \\ & \ddots & \\ & & e^{a_n} \end{pmatrix}.$$

Für jede invertierbare Matrix U haben wir:

$$e^{UAU^{-1}} = Ue^A U^{-1}. \tag{VII.29}$$

I.A. ist $e^{AB} \neq e^A e^B$, für kommutierende A und B gilt jedoch Gleichheit. Insbesondere ist $t \mapsto e^{tA}$ ein Gruppenhomomorphismus:

$$e^{(s+t)A} = e^{sA}e^{tA}, \qquad (e^A)^{-1} = e^{-A} \qquad \text{und} \qquad \tfrac{d}{dt}e^{tA} = Ae^{tA}.$$

Für jedes v_0 ist daher $v(t) = e^{tA}v_0$ die Lösung des Systems gewöhnlicher Differentialgleichung $v'(t) = Av(t)$ zum Anfangswert $v(0) = v_0$.

VII.5.18. Beispiel. Wir wollen e^{tA} bestimmen, wobei

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Durch Berechnung einer Eigenbasis von A erhalten wir die Darstellung:

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ -\mathbf{i} & \mathbf{i} \end{pmatrix} \begin{pmatrix} \mathbf{i} & 0 \\ 0 & -\mathbf{i} \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -\mathbf{i} & \mathbf{i} \end{pmatrix}^{-1}$$

Mit (VII.29) daher

$$\begin{aligned} e^{tA} &= \begin{pmatrix} 1 & 1 \\ -\mathbf{i} & \mathbf{i} \end{pmatrix} \begin{pmatrix} e^{t\mathbf{i}} & 0 \\ 0 & e^{-t\mathbf{i}} \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -\mathbf{i} & \mathbf{i} \end{pmatrix}^{-1} \\ &= \begin{pmatrix} \frac{e^{t\mathbf{i}}+e^{-t\mathbf{i}}}{2} & -\frac{e^{t\mathbf{i}}-e^{-t\mathbf{i}}}{2\mathbf{i}} \\ \frac{e^{t\mathbf{i}}-e^{-t\mathbf{i}}}{2\mathbf{i}} & \frac{e^{t\mathbf{i}}+e^{-t\mathbf{i}}}{2} \end{pmatrix} = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}, \end{aligned}$$

d.h. e^{tA} ist eine Drehung um den Winkel t.

Analog lassen sich konvergente Potenzreihe, $f(z) = \sum_n a_n z^n$, auf quadratische Matrizen anwenden, $f(A) := \sum_n a_n A^n$. So lässt sich etwa der Gleichung

$$e^{\mathbf{i}A} = \cos(A) + \mathbf{i}\sin(A),$$

auch für quadratische Matrizen A Sinn geben. Basisunabhängig, d.h. für Endomorphismen φ eines endlich dimensionalen Vektorraums: $f(\varphi) = \sum_n a_n \varphi^n$.

VII.6. Moore–Penrose Pseudoinverse. Sei $\varphi\colon V \to W$ eine lineare Abbildung zwischen endlich dimensionalen Euklidischen oder unitären Vektorräumen. Dann ist die Einschränkung,

$$\varphi|_{\ker(\varphi)^\perp}\colon \ker(\varphi)^\perp \xrightarrow{\cong} \operatorname{img}(\varphi)$$

invertierbar und wir definieren ihre sogenannte *Moore–Penrose Pseudoinverse*, $\varphi^+\colon W \to V$, durch

$$\varphi^+|_{\operatorname{img}(\varphi)} := \varphi|_{\ker(\varphi)^\perp}^{-1} \qquad \text{und} \qquad \varphi^+|_{\operatorname{img}(\varphi)^\perp} := 0.$$

VII.6.1. Satz (Moore–Penrose Pseudoinverse). *Sei $\varphi\colon V \to W$ eine lineare Abbildung zwischen endlich dimensionalen Euklidischen oder unitären Vektorräumen. Dann hat die Pseudoinverse, $\varphi^+\colon W \to V$, folgende vier Eigenschaften und ist dadurch eindeutig charakterisiert:*

(a) $\varphi\varphi^+\varphi = \varphi$.

(b) $\varphi^+\varphi\varphi^+ = \varphi^+$.
(c) $\varphi\varphi^+$ *ist selbstadjungiert, d.h.* $(\varphi\varphi^+)^* = \varphi\varphi^+$.
(d) $\varphi^+\varphi$ *ist selbstadjungiert, d.h.* $(\varphi^+\varphi)^* = \varphi^+\varphi$.

Darüber hinaus gilt:

(e) $\varphi^*\varphi\varphi^+ = \varphi^*$.
(f) $\varphi^+\varphi\varphi^* = \varphi^*$.
(g) $\varphi\varphi^+$ *ist die Orthogonalprojektion auf* $\mathrm{img}(\varphi) = \ker(\varphi^*)^\perp$.
(h) $\varphi^+\varphi$ *ist die Orthogonalprojektion auf* $\mathrm{img}(\varphi^*) = \ker(\varphi)^\perp$.
(i) $\ker(\varphi^+) = \ker(\varphi^*) = \mathrm{img}(\varphi)^\perp$.
(j) $\mathrm{img}(\varphi^+) = \mathrm{img}(\varphi^*) = \ker(\varphi)^\perp$.
(k) Ist φ *invertierbar, dann gilt* $\varphi^+ = \varphi^{-1}$.
(l) Ist φ *surjektiv, so ist* $\varphi\varphi^+ = \mathrm{id}_W$ *und* $\varphi^+ = \varphi^*(\varphi\varphi^*)^{-1}$.
(m) Ist φ *injektiv, so ist* $\varphi^+\varphi = \mathrm{id}_V$ *und* $\varphi^+ = (\varphi^*\varphi)^{-1}\varphi^*$.
(n) Ist $\psi\colon U \to V$ *surjektiv und* $\varphi\colon V \to W$ *injektiv, dann gilt* $(\varphi\psi)^+ = \psi^+\varphi^+ = \psi^*(\psi\psi^*)^{-1}(\varphi^*\varphi)^{-1}\varphi^* = \psi^*(\varphi^*\varphi\psi\psi^*)^{-1}\varphi^*$.
(o) $(\lambda\varphi)^+ = \lambda^{-1}\varphi^+$, *für alle* $0 \neq \lambda \in \mathbb{K}$.
(p) $(\varphi^*)^+ = (\varphi^+)^*$
(q) Sei $b_1, \dots, b_k$ *eine Basis von* $\ker(\varphi)^\perp$ *und* $b_{k+1}, \dots, b_n$ *eine Basis von* $\ker(\varphi)$. *Weiters sei* $c_1, \dots, c_k$ *eine Basis von* $\mathrm{img}(\varphi)$ *und* $c_{k+1}, \dots, c_m$ *eine Basis von* $\mathrm{img}(\varphi)^\perp$. *Dann gilt*

$$[\varphi]_{CB} = \begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix} \qquad \text{und} \qquad [\varphi^+]_{BC} = \begin{pmatrix} A^{-1} & 0 \\ 0 & 0 \end{pmatrix}.$$

wobei A *eine invertierbare* $(k \times k)$*-Matrix bezeichnet.*

Beweis. Nach Konstruktion ist $\varphi\varphi^+|_{\mathrm{img}(\varphi)} = \mathrm{id}_{\mathrm{img}(\varphi)}$ und $\varphi\varphi^+|_{\mathrm{img}(\varphi)^\perp} = 0$, also stimmt $\varphi\varphi^+$ mit der Orthogonalprojektion auf $\mathrm{img}(\varphi)$ überein. Dies zeigt (a), (c) und (g). Nach Konstruktion gilt $\varphi^+\varphi|_{\ker(\varphi)^\perp} = \mathrm{id}_{\ker(\varphi)^\perp}$ und offensichtlich auch $\varphi^+\varphi|_{\ker(\varphi)} = 0$. Daraus schließen wir, dass $\varphi^+\varphi$ mit der Orthogonalprojektion auf $\ker(\varphi)^\perp$ übereinstimmt. Dies zeigt (b), (d) und (h). Behauptung (e) folgt aus (a) und (c). Analog erhalten wir (f) sofort aus (a) und (d). Auch (i), (j), (k), (o) und (q) folgen sofort aus der Definition. Behauptung (l) folgt aus (g) und (f), denn für surjektives φ ist $\varphi\varphi^*$ invertierbar. Analog erhalten wir aus (h) und (e) sofort (m), denn für injektives φ ist $\varphi^*\varphi$ invertierbar. Ad (n): In diesem Fall gilt $\ker(\varphi\psi) = \ker(\psi)$ und $\mathrm{img}(\varphi\psi) = \mathrm{img}(\varphi)$ und aus unserer Definition folgt sofort $(\varphi\psi)^+ = \psi^+\varphi^+$. Die anderen Formeln für $\psi^+\varphi^+$ folgen dann aus (l) und (m).

Wir verifizieren nun, dass die Pseudoinverse durch die Eigenschaften (a) bis (d) eindeutig bestimmt ist. Sei dazu $\tilde\varphi^+\colon W \to V$, sodass (a) bis (d) auch für $\tilde\varphi^+$ statt φ^+ gelten. Setzen wir $p := \tilde\varphi^+\varphi$, so folgt $p^2 = p = p^*$ und $\ker(p) = \ker(\varphi)$ aus (a) und (d). Somit ist p die Orthogonalprojektion auf $\ker(\varphi)$, also

$$\tilde\varphi^+\varphi = p = \varphi^+\varphi.$$

Analog erhalten wir für $q := \varphi\tilde{\varphi}^+$ aus (a) und (c) sofort $q^2 = q = q^*$ und $\mathrm{img}(q) = \mathrm{img}(\varphi)$, also ist q die Orthogonalprojektion auf $\mathrm{img}(\varphi)$ und daher

$$\varphi\tilde{\varphi}^+ = q = \varphi\varphi^+.$$

Mit (d) folgt

$$\tilde{\varphi}^+ = \tilde{\varphi}^+\varphi\tilde{\varphi}^+ = \varphi^+\varphi\tilde{\varphi}^+ = \varphi^+\varphi\varphi^+ = \varphi^+,$$

womit die Eindeutigkeitsaussage gezeigt wäre.

Ad (p): Aus (a) bis (d) erhalten wir $\varphi^*(\varphi^+)^*\varphi^* = \varphi^*$, $(\varphi^+)^*\varphi^*(\varphi^+)^* = (\varphi^+)^*$, und $\varphi^*(\varphi^+)^*$ sowie $(\varphi^+)^*\varphi^*$ sind selbstadjungiert. Aus der eben bewiesenen Eindeutigkeit der Pseudoinversen folgt daher $(\varphi^*)^+ = (\varphi^+)^*$. □

Für Matrizen erhalten wir daraus:

VII.6.2. Korollar. *Sei $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$. Für jede Matrix $A \in M_{m\times n}(\mathbb{K})$ existiert eine eindeutig bestimmte Matrix $A^+ \in M_{n\times m}(\mathbb{K})$ mit folgenden vier Eigenschaften:*

(a) $AA^+A = A$
(b) $A^+AA^+ = A^+$
(c) $(AA^+)^* = AA^+$
(d) $(A^+A)^* = A^+A$

Diese Pseudoinverse A^+ hat die zu den in Satz VII.6.1 analogen Eigenschaften.

(e) $A^*AA^+ = A^*$
(f) $A^+AA^* = A^*$
(g) AA^+ *ist die Matrix der Orthogonalprojektion auf* $\mathrm{img}(A) = \ker(A^*)^\perp$.
(h) A^+A *ist die Matrix der Orthogonalprojektion auf* $\mathrm{img}(A^*) = \ker(A)^\perp$.
(i) $\ker(A^+) = \ker(A^*) = \mathrm{img}(A)^\perp$.
(j) $\mathrm{img}(A^+) = \mathrm{img}(A^*) = \ker(A)^\perp$.
(k) Ist A invertierbar, dann gilt $A^+ = A^{-1}$.
(l) Ist $\mathrm{rank}(A) = m$, *so gilt* $AA^+ = I_m$ *und* $A^+ = A^*(AA^*)^{-1}$.
(m) Ist $\mathrm{rank}(A) = n$, *so gilt* $A^+A = I_n$ *und* $A^+ = (A^*A)^{-1}A^*$.
(n) Ist $B \in M_{n\times l}(\mathbb{K})$ *und* $\mathrm{rank}(A) = n = \mathrm{rank}(B)$, *dann gilt* $(AB)^+ = B^+A^+ = B^*(BB^*)^{-1}(A^*A)^{-1}A^* = B^*(A^*ABB^*)^{-1}A^*$.
(o) $(\lambda A)^+ = \lambda^{-1}A^+$, *für alle Skalare* $0 \neq \lambda \in \mathbb{K}$.
(p) $(A^*)^+ = (A^+)^*$

VII.6.3. Beispiel. Ist $A \in \mathrm{GL}_k(\mathbb{K})$, dann gilt:

$$\underbrace{\begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix}^+}_{\in M_{m\times n}(\mathbb{K})} = \underbrace{\begin{pmatrix} A^{-1} & 0 \\ 0 & 0 \end{pmatrix}}_{\in M_{n\times m}(\mathbb{K})}.$$

VII.6.4. Beispiel. Nach Korollar VII.6.2(l) gilt:

$$\underbrace{\begin{pmatrix}1&1&1\\1&2&3\end{pmatrix}}_{A}^{+}=A^*(AA^*)^{-1}=\begin{pmatrix}1&1\\1&2\\1&3\end{pmatrix}\left\{\begin{pmatrix}1&1&1\\1&2&3\end{pmatrix}\begin{pmatrix}1&1\\1&2\\1&3\end{pmatrix}\right\}^{-1}$$

$$=\begin{pmatrix}1&1\\1&2\\1&3\end{pmatrix}\begin{pmatrix}3&6\\6&14\end{pmatrix}^{-1}=\begin{pmatrix}1&1\\1&2\\1&3\end{pmatrix}\frac{1}{6}\begin{pmatrix}14&-6\\-6&3\end{pmatrix}=\begin{pmatrix}4/3&-1/2\\1/3&0\\-2/3&1/2\end{pmatrix}.$$

VII.6.5. Beispiel. Wir wollen die Pseudoinverse von

$$A=\begin{pmatrix}1&0&1&0\\1&2&3&4\\2&2&4&4\end{pmatrix}$$

bestimmen. Zunächst mittels Zeilenumformungen:

$$A=\begin{pmatrix}1&0&1&0\\1&2&3&4\\2&2&4&4\end{pmatrix}=\begin{pmatrix}1&0&0\\-1&1&0\\-1&-1&1\end{pmatrix}^{-1}\begin{pmatrix}1&0&1&0\\0&2&2&4\\0&0&0&0\end{pmatrix}$$

$$=\begin{pmatrix}1&0&0\\1&1&0\\2&1&1\end{pmatrix}\begin{pmatrix}1&0&1&0\\0&2&2&4\\0&0&0&0\end{pmatrix}=\begin{pmatrix}1&0\\1&1\\2&1\end{pmatrix}\begin{pmatrix}1&0&1&0\\0&2&2&4\end{pmatrix}=:LU$$

Mit Korollar VII.6.2(n) daher:

$$A^+=U^+L^+=U^*\big(L^*LUU^*\big)^{-1}L^*$$

$$=U^*\left\{\begin{pmatrix}6&3\\3&2\end{pmatrix}\begin{pmatrix}2&2\\2&24\end{pmatrix}\right\}^{-1}L^*=U^*\begin{pmatrix}18&84\\10&54\end{pmatrix}^{-1}L^*$$

$$=\begin{pmatrix}1&0\\0&2\\1&2\\0&4\end{pmatrix}\frac{1}{66}\begin{pmatrix}27&-42\\-5&9\end{pmatrix}\begin{pmatrix}1&1&2\\0&1&1\end{pmatrix}=\frac{1}{66}\begin{pmatrix}27&-15&12\\-10&8&-2\\17&-7&10\\-20&16&-4\end{pmatrix}.$$

Als Anwendung der Pseudoinversen wollen wir nun lineare Gleichungssysteme, $Ax=y$, betrachten, die sich nicht exakt lösen lassen. In solch einer Situation ist es oft hilfreich x so zu bestimmen, dass der Euklidische Abstand, $d(Ax,y)=\|Ax-y\|$, minimal wird.

VII.6.6. Satz (Methode der kleinsten Quadrate). *Sei $\mathbb{K}=\mathbb{R}$ oder $\mathbb{K}=\mathbb{C}$. Weiters seien $A\in M_{m\times n}(\mathbb{K})$, $y\in\mathbb{K}^m$ und $x\in\mathbb{K}^n$. Dann sind äquivalent:*

(a) $d(Ax,y)\leq d(A\tilde{x},y)$, für alle $\tilde{x}\in\mathbb{K}^n$. (x liefert beste Approximation)
(b) $Ax-y\perp\operatorname{img}(A)$.
(c) $A^*Ax=A^*y$. (Normalgleichungen)

(d) $x \in A^+y + \ker(A)$, wobei $A^+ \in M_{n\times m}(\mathbb{K})$ die Pseudoinverse von A bezeichnet.
Unter allen x, die diesen Bedingungen genügen, ist $x = A^+y$ das eindeutige mit kleinster Norm.

Beweis. Die Äquivalenz (a)⇔(b) folgt aus Satz VII.2.32. Die Äquivalenz (b)⇔(c) folgt aus $\operatorname{img}(A)^\perp = \ker(A^*)$, siehe Proposition VII.3.2. Ad (d)⇔(c): Ein Vektor $x \in \mathbb{R}^n$ genügt genau dann den Normalgleichungen in (c), wenn

$$A^*AA^+y + A^*A(x - A^+y) = A^*y.$$

Nach Korollar VII.6.2(e) gilt $A^*AA^+ = A^*$, also ist dies zu $A^*A(x - A^+y) = 0$ äquivalent. Dies ist genau dann der Fall, wenn $x - A^+y \in \ker(A^*A) = \ker(A)$ gilt, d.h. genau dann wenn $x \in A^+y + \ker(A)$. Damit ist die Äquivalenz von (a) bis (d) gezeigt. Ist $\xi \in \ker(A)$, dann gilt $\xi \perp A^+y$, siehe Korollar VII.6.2(j), und nach dem Satz von Pythagoras daher

$$\|A^+y + \xi\|^2 = \|A^+y\|^2 + \|\xi\|^2 \geq \|A^+y\|^2,$$

wobei Gleichheit nur für $\xi = 0$ gilt. Dies zeigt, dass $x = A^+y$ die eindeutige Lösung der Normalgleichungen mit kleinster Norm ist. □

VII.6.7. Beispiel (Polynomausgleich). Gegeben seien $x_1, \dots, x_N \in \mathbb{R}$ und $y_1, \dots, y_N \in \mathbb{R}$, sowie $n \in \mathbb{N}$. Wir suchen ein Polynom n-ten Grades,

$$p(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n,$$

sodass $p(x_i)$ möglichst nahe bei y_i liegt. Genauer, die Summe der Fehlerquadrate,

$$\sum_{i=1}^N \bigl|p(x_i) - y_i\bigr|^2,$$

soll minimal sein. Betrachten wir

$$A = \begin{pmatrix} 1 & x_0 & x_0^2 & \cdots & x_0^n \\ 1 & x_1 & x_1^2 & \cdots & x_1^n \\ \vdots & \vdots & & \vdots & \\ 1 & x_N & x_N^2 & \cdots & x_N^n \end{pmatrix}, \qquad a = \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{pmatrix} \qquad \text{und} \qquad y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{pmatrix},$$

dann gilt:

$$Aa = \begin{pmatrix} p(x_0) \\ p(x_1) \\ \vdots \\ p(x_N) \end{pmatrix} \qquad \text{und} \qquad d(Aa, y)^2 = \sum_{i=1}^N \bigl|p(x_i) - y_i\bigr|^2$$

Es ist daher $a \in \mathbb{R}^{n+1}$ so zu bestimmen, sodass $d(Aa, y)$ minimal wird. Nach Satz VII.6.6 ist dies für $a = A^+y$ der Fall. Haben wir wenigstens $n + 1$ verschiedene Werte x_i vorliegen, dann hat A maximalen Rang, $\operatorname{rank}(A) = n + 1$, siehe

Satz IV.6.26. Nach Korollar VII.6.2(m) gilt in diesem Fall $A^+ = (A^*A)^{-1}A^*$, die Koeffizienten des gesuchten Polynoms sind daher durch

$$a = (A^*A)^{-1}A^*y$$

gegeben.

VII.6.8. Beispiel (Linearer Ausgleich). Es seien x_i und y_i wie folgt gegeben:

i	1	2	3	4	5	6	7
x_i	2.5	3.1	3.1	4.2	5.1	6.7	6.7
y_i	6.5	7.5	7.9	9.9	11.8	14.9	14.5

Gesucht sei ein lineares Polynom, $p(x) = a + bx$, sodass die Summe der Fehlerquadrate,

$$\sum_{i=1}^{7} \bigl|p(x_i) - y_i\bigr|^2$$

minimal wird. Nach Beispiel VII.6.7 sind die Koeffizienten durch

$$\begin{pmatrix} a \\ b \end{pmatrix} = (A^*A)^{-1}A^*y = \frac{1}{126.34}\begin{pmatrix} 209.978 \\ 246.91 \end{pmatrix} = \begin{pmatrix} 1.6620073\dots \\ 1.9543296\dots \end{pmatrix}$$

gegeben, wobei

$$A = \begin{pmatrix} 1 & 2.5 \\ 1 & 3.1 \\ 1 & 3.1 \\ 1 & 4.2 \\ 1 & 5.1 \\ 1 & 6.7 \\ 1 & 6.7 \end{pmatrix} \qquad \text{und} \qquad y = \begin{pmatrix} 6.5 \\ 7.5 \\ 7.9 \\ 9.9 \\ 11.8 \\ 14.9 \\ 14.5 \end{pmatrix}.$$

Denn

$$(A^*A)^{-1} = \begin{pmatrix} 7 & 31.4 \\ 31.4 & 158.9 \end{pmatrix}^{-1} = \frac{1}{126.34}\begin{pmatrix} 158.9 & -31.4 \\ -31.4 & 7 \end{pmatrix},$$

und

$$(A^*A)^{-1}A^* = \frac{1}{126.34}\begin{pmatrix} 80.4 & 61.56 & 61.56 & 27.02 & -1.24 & -51.48 & -51.48 \\ -13.9 & -9.7 & -9.7 & -2 & 4.3 & 15.5 & 15.5 \end{pmatrix}.$$

Das gesuchte lineare Polynom ist daher:

$$p(x) = a + bx \sim 1.66 + 1.95x.$$