Gruppe

Unter einer Gruppe $(G, \diamond)$ versteht man eine Menge G, auf der eine binäre Operation $\diamond$ definiert ist:

$$\diamond : G \times G \longmapsto G\,,$$

d.h. jedem Elementepaar $(a, b) :\; a, b \in G$ ist ein Element $a \diamond b \in G$ zugeordnet. Ferner müssen folgende Eigenschaften gelten:

- Assoziativität: $(a \diamond b) \diamond c = a \diamond (b \diamond c) \quad \forall a, b, c \in G$
- Neutrales Element: Es existiert ein eindeutig bestimmtes neutrales Element $e \in G$, d.h.

$$e \diamond a = a \diamond e = a \quad \forall a \in G$$

- Inverses Element: Zu jedem Element $a \in G$ existiert ein eindeutig bestimmtes inverses Element $a^{-1} \in G$ mit

$$a \diamond a^{-1} = a^{-1} \diamond a = e$$

Man nennt eine Gruppe eine kommutative oder abelsche Gruppe, wenn die Operation $\diamond$ kommutativ ist:

$$a \diamond b = b \diamond a \qquad \forall a, b \in G$$

Wenn aus dem Zusammenhang ersichtlich ist, welche Operation verwendet wird, schreibt man häufig statt $(G, \diamond)$ nur G.

Beispiel

Die bijektiven reellen Funktionen $f: \mathbb{R} \to \mathbb{R}$ bilden bezüglich der Hintereinanderschaltung $\circ$ eine Gruppe.

Verifizierung der Gruppenaxiome

- Assoziativität:

$$((f \circ g) \circ h)(x) = f(g(h(x))) = (f \circ (g \circ h))(x)$$

- Neutrales Element: Identität

$$e: x \mapsto x$$

- Inverses Element: Umkehrfunktion

$$f^{-1}: f(x) \mapsto x$$

Gegenbeispiel zur Kommutativität: $f \circ g \neq g \circ f$ für

$$f: x \mapsto 2x, \quad g: x \mapsto x + 1$$

$f(g(x)) = 2(x + 1) \neq 2x + 1 = g(f(x))$

Beispiel

Abelsche Gruppe

$$\mathbb{Z}_n = \mathbb{Z} \bmod n$$

der Restklassen $\{0, 1, \ldots, n-1\}$ mit der Addition modulo n als Gruppenoperation

Illustration der Gruppenaxiome für $\mathbb{Z}_4 = \{0, 1, 2, 3\}$

- Assoziativität

$$\begin{aligned}(1 + 2 \bmod 4) + 3 \bmod 4 &= 3 + 3 \bmod 4 = 6 \bmod 4 = 2\\ 1 + (2 + 3 \bmod 4) \bmod 4 &= 1 + (5 \bmod 4) \bmod 4 = 1 + 1 \bmod 4 = 2\end{aligned}$$

- Neutrales Element 0

$$3 + 0 \bmod 4 = 0 + 3 \bmod 4 = 3$$

- Inverses Element

$$0 = 1 + 3 \bmod 4 = 2 + 2 \bmod 4 = 3 + 1 \bmod 4$$

Die Multiplikation modulo 4 definiert keine Gruppenstruktur, denn beispielsweise gilt

$$\begin{aligned} 2 &= 2 \cdot 1 \text{ mod } 4 = 2 \bmod 4 \\ &= 2 \cdot 3 \text{ mod } 4 = 6 \bmod 4 \end{aligned}$$

$\Longrightarrow$ keine Eindeutigkeit des neutralen Elements $(1 \neq 3)$

Gruppentafel

Die Operation $\diamond$ auf einer endlichen Gruppe $G = \{g_1, \ldots, g_n\}$ kann durch die Verknüpfungsmatrix A definiert werden:

$$A: \quad a_{j,k} = g_j \diamond g_k$$

Für eine abelsche (kommutative) Gruppe ist A symmetrisch:

$$g_j \diamond g_k = g_k \diamond g_j \quad \Longleftrightarrow \quad A = A^{\mathrm{t}}\,.$$

Die erste nicht-abelsche Gruppe hat 6 Elemente und kann mit den Permutationen von $\{1, 2, 3\}$ identifiziert werden.

Beispiel

Gruppentafeln für Gruppen mit $\leq$ 4 Elementen

$\diamond$	e	a
e	e	a
a	a	e

$\diamond$	e	a	b
e	e	a	b
a	a	b	e
b	b	e	a

$\diamond$	e	a	b	c
e	e	a	b	c
a	a	e	c	b
b	b	c	e	a
c	c	b	a	e

$\diamond$	e	a	b	c
e	e	a	b	c
a	a	e	c	b
b	b	c	a	e
c	c	b	e	a

Untergruppe

Für eine Gruppe $(G, \diamond)$ bezeichnet man $(U, \diamond)$ als Untergruppe, wenn U eine Teilmenge von G ist, die selbst eine Gruppe bildet.
Um zu testen, ob $(U, \diamond)$ die Gruppenaxiome erfüllt, genügt es zu überprüfen, dass U bezüglich der Verknüpfung $\diamond$ und der Bildung von Inversen abgeschlossen ist:

$$a, b \in U \implies a \diamond b \in U \quad \wedge \quad a \in U \implies a^{-1} \in U\,.$$

Beispiel

Gruppe der Kongruenzabbildungen eines Quadrates

$$ABCD \mapsto BCDA,\ ABCD \mapsto ADCB,\ \ldots$$

Untergruppe der Drehungen D_α: $\alpha = 0°, 90°, 180°, 270°$

$$\begin{array}{rccc} 0^\circ: & ABCD & \mapsto & ABCD \\ 90^\circ: & ABCD & \mapsto & BCDA \\ 180^\circ: & ABCD & \mapsto & CDAB \\ 270^\circ: & ABCD & \mapsto & DABC \end{array}$$

abgeschlossen bzgl. Verknüpfung und Invertierung

$$D_\alpha \circ D_\beta = D_{\alpha+\beta}, \quad (D_\alpha)^{-1} = D_{-\alpha}$$

Spiegelungen bilden keine Untergruppe:
Verknüpfung $\rightsquigarrow$ Drehung, z.B. Komposition von Spiegelung an der Diagonale durch A und C,

$$S_1 : ABCD \mapsto ADCB\,,$$

und Spiegelung an der Mittelsenkrechten auf der Strecke von A nach B,

$$S_2 : ABCD \mapsto BADC\,,$$

$\rightsquigarrow$

$$D = S_2 \circ S_1 : ABCD \mapsto BCDA$$

$\widehat{=}$ Drehung um 90°

Für Spiegelungen existiert ebenfalls kein neutrales Element.

Permutationen

Für eine endliche Menge M bilden die bijektiven Abbildungen $p: M \to M$, versehen mit der Komposition von Abbildungen als Operation, die symmetrische Gruppe von M.

Für $M = \{1, 2, \ldots, n\}$ bezeichnet man diese Gruppe mit S_n und die $n!$ Elemente p als Permutationen. Die Permutationsgruppe ist nur für $n = 2$ kommutativ.

Man benutzt die Schreibweise

$$p = \begin{pmatrix} 1 & 2 & 3 & \ldots & n \\ p(1) & p(2) & p(3) & \ldots & p(n) \end{pmatrix}$$

zur Beschreibung einer Permutation. Ebenfalls gebräuchlich ist die Zyklendarstellung. Ein Zyklus besteht aus einem Element und seinen Bildern bei wiederholter Ausführung der Permutation, bis wieder das ursprüngliche Element erreicht wird. Aus den Elementen, die im ersten Zyklus nicht vorkommen, werden weitere Zyklen gebildet, bis alle Elemente auftreten. Die Zyklen werden nach der Anzahl der Elemente absteigend sortiert und jeweils in runden Klammern hintereinander geschrieben. Zyklen der Länge 1 werden meist weggelassen.

Beispielsweise ist

$$p = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 4 & 3 & 2 & 6 & 5 & 1 \end{pmatrix} \equiv (1\,4\,6)\,(2\,3)\,(5) \quad \text{bzw.} \quad p = (1\,4\,6)\,(2\,3)\,.$$

Transposition und Signum einer Permutation

Eine Transposition

$$\tau = (j\ k)$$

ist eine Vertauschung von j und k. Durch Verknüpfung dieser elementaren Permutationen lässt sich jede Permutation p darstellen:

$$p = \tau_1 \circ \cdots \circ \tau_m$$

Die Parität (gerades oder ungerades m) ist eindeutig bestimmt, und man definiert

$$\sigma(p) = (-1)^m$$

als Vorzeichen oder Signum der Permutation p.

Für eine zyklische Permutation p ist $\sigma(p) = n - 1$ mit n der Länge des Zyklus. Der Exponent m kann damit aus der Zyklendarstellung einer Permutation als Summe der jeweils um 1 verminderten Zyklenlängen bestimmt werden.

Beweis

(i) Wohldefiniertheit von σ:
Beweis durch Induktion

- Induktionsanfang ($n = 2$):
  $S_2 = \{p, q\}$ mit $p = (1)(2)$ (Identität) und $q = (1\,2)$ (Transposition) in Zykelschreibweise
  $p = q \circ q \circ \cdots \circ q$ (gerade Anzahl) und $q = q \circ q \circ \cdots \circ q$ (ungerade Anzahl)
  $\implies$ eindeutig bestimmte Vorzeichen $\sigma(p) = 1$, $\sigma(q) = -1$
- Induktionsschritt ($(n-1) \to n$):
  betrachte
  $$\tau_1 \circ \cdots \circ \tau_m = p \in S_n$$
  definiere
  $$q = (k\;n) \circ p \quad \text{mit } k = p(n)$$
  $\rightsquigarrow$ $q(n) = n$ und damit Identifikation von q mit einer Permutation $\tilde{q}$ in S_{n-1} mit
  $$\sigma(\tilde{q}) = (-1)(-1)^m$$
  Induktionsvoraussetzung $\implies$ $\sigma(\tilde{q})$ eindeutig bestimmt
  $\implies$ Exponent m eindeutig bestimmt

(ii) Vorzeichen von Zykeln:

$$p = (p_1\, p_2\, \ldots\, p_n) = \underbrace{(p_1\, p_2\, \ldots\, p_{n-1}) \circ (p_{n-1}\, p_n)}_{q},$$

denn $q(p_{n-1}) = p_n$, $q(p_n) = p(p_{n-1}) = p_1$
wiederholte Anwendung der Aufspaltung $\quad\rightsquigarrow\quad$ Darstellung von p als Komposition von $n-1$ Transpositionen

$$(p_1\, p_2) \circ (p_2\, p_3) \circ \ldots \circ (p_{n-1}\, p_n)$$

$\Longrightarrow \quad \sigma(p) = n - 1$

Beispiel

Bestimmung des Signums der Permutation

$$p = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 6 & 5 & 3 & 1 & 2 & 4 \end{pmatrix}$$

(i) Darstellung als Komposition von Transpositionen:
Überführung von

$$(p(1), \ldots, p(6)) = (6, 5, 3, 1, 2, 4)$$

durch Transpositionen sukzessive in die kanonische Reihenfolge:

$$\begin{array}{ccc} (1\,6) & : & (1,5,3,6,2,4) \\ (2\,5) & : & (1,2,3,6,5,4) \\ (4\,6) & : & (1,2,3,4,5,6) \end{array}$$

$\Longrightarrow$

$$\mathrm{id} = (4,6) \circ (2,5) \circ (1,6) \circ p, \quad p = (1,6) \circ (2,5) \circ (4,6)$$

und $\sigma(p) = (-1)^3 = -1$

(ii) Alternative Bestimmung von σ mit Hilfe der Zyklenschreibweise:

$$p = (1\,6\,4)\,(2\,5)\,(3)$$

$\sigma(\tau) = (-1)^{k-1}$ für einen Zyklus τ der Länge k, denn

$$(a\,b\,c\,\ldots\,f) = (a\,b) \circ (b\,c) \circ \ldots \circ (e\,f)$$

Anwendung auf das Beispiel $\rightsquigarrow$

$$\sigma(p) = (-1)^{(3-1)+(2-1)+(1-1)} = (-1)^3 = -1$$

Körper

Eine Menge K, auf der eine Addition „$+$“ und eine Multiplikation „$\cdot$“ definiert sind, nennt man einen Körper, wenn folgendes gilt:

- Additive Gruppenstruktur: $(K, +)$ ist eine abelsche Gruppe mit neutralem Element 0 (Nullelement), d.h. für alle $a, b \in K$ gilt

$$
\begin{aligned}
a + b &= b + a \\
(a + b) + c &= a + (b + c) \\
a + 0 &= a \\
a + (-a) &= 0\,,
\end{aligned}
$$

  wobei $(-a)$ das inverse Element zu a bezeichnet.

- Multiplikative Gruppenstruktur: $(K\backslash\{0\},\cdot)$ ist eine abelsche Gruppe mit neutralem Element 1 (Einselement), d.h. für alle $a,b,c \in K\backslash\{0\}$ gilt

$$\begin{aligned} a\cdot b &= b\cdot a \\ (a\cdot b)\cdot c &= a\cdot(b\cdot c) \\ a\cdot 1 &= a \\ a\cdot a^{-1} &= 1\,, \end{aligned}$$

  wobei a^{-1} das inverse Element zu a bezeichnet.
- Distributivgesetz:

$$a\cdot(b+c) = a\cdot b + a\cdot c$$

  für alle $a,b,c \in K$.

Beispiel

Körper der rationalen, reellen und komplexen Zahlen

$$\mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$$

Nullelement: 0, Einselement: 1

Inverses Element bezüglich der Multiplikation einer komplexen Zahl $z = x + \mathrm{i}y \neq 0$:

$$w = \frac{x}{x^2 + y^2} + \mathrm{i}\,\frac{-y}{x^2 + y^2}\ ,$$

denn

$$\begin{aligned} z \cdot w &= (x + \mathrm{i}y)\left(\frac{x}{x^2 + y^2} + \mathrm{i}\,\frac{-y}{x^2 + y^2}\right) \\ &= \frac{x^2 - \mathrm{i}^2 y^2}{x^2 + y^2} + \mathrm{i}\,\frac{-xy + yx}{x^2 + y^2} = 1 + \mathrm{i}0 = 1 \end{aligned}$$

Beispiel

Gruppentafeln von Addition und Multiplikation für den Galois-Körper GF$[2^2]$ mit 4 Elementen

$+$	0	1	a	b
0	0	1	a	b
1	1	0	b	a
a	a	b	0	1
b	b	a	1	0

$\cdot$	0	1	a	b
0	0	0	0	0
1	0	1	a	b
a	0	a	b	1
b	0	b	1	a

z.B: $(a+1)\cdot b = b\cdot b = 1$, $-b = b$ $(\Longleftrightarrow b + b = 0)$, $1/a = b$

Die Konstruktion von Körpern mit p^ℓ Elementen, $\ell \in \mathbb{N}$, ist für beliebige Primzahlen p durchführbar.

$\rightsquigarrow$ alle endlichen Körper

Primkörper

Für jede Primzahl p ist die Menge

$$\mathbb{Z}_p = \{0, 1, \ldots, p-1\}$$

ein Körper unter der Addition und Multiplikation modulo p.
Allgemeiner existieren endliche Körper mit p^k Elementen für jedes $k \in \mathbb{N}$, die sogenannten Galois-Körper. Dies sind die einzigen Körper mit endlich vielen Elementen.

Beweis

Rechenregeln für Addition und Multiplikation in Körpern gelten in den ganzen Zahlen

$\rightsquigarrow$ Gültigkeit der Rechenregeln für $\mathbb{Z}_p$

$\rightsquigarrow$ noch zu zeigen: Existenz eines inversen Elementes a^{-1} für $a \in \{2, \ldots, p-1\}$

betrachte dazu die Folge

$$a^k \bmod p, \quad k = 0, \ldots, p-1$$

$a^k \neq 0 \bmod p \, \forall k \in \mathbb{N}$, denn

$$a^k = np \implies p \operatorname{teilt} a^k \underset{p \,\text{Primzahl}}{\implies} p \operatorname{teilt} a \implies \text{Widerspruch zu } a < p$$

$\implies$ mindestens ein Rest tritt zweimal auf:

$$a^{k_1} = a^{k_2} \bmod p, \quad k_1 < k_2$$

$a^{k_1} a^{k_2 - k_1} = a^{k_1} \implies$

$$1 = a^{k_2 - k_1} \bmod p = a^{k_2 - k_1 - 1} a \bmod p \implies a^{-1} = a^{k_2 - k_1 - 1} \bmod p$$

Beispiel

Inverse Elemente im Primkörper $\mathbb{Z}_5$

$$2^{-1} \bmod 5 = 3, \quad 3^{-1} \bmod 5 = 2, \quad 4^{-1} \bmod 5 = 4$$

Überprüfung durch Multiplikation, z.B.

$$2 \cdot 2^{-1} \bmod 5 = 2 \cdot 3 \bmod 5 = 6 \bmod 5 = 1 \quad \checkmark$$

Illustration anhand des Distributivgesetzes

$$\begin{aligned}
(2+4) \cdot 3^{-1} \bmod 5 &= 6 \cdot 2 \bmod 5 \\
&= 12 \bmod 5 = 2 \\
2 \cdot 3^{-1} + 4 \cdot 3^{-1} \bmod 5 &= 2 \cdot 2 + 4 \cdot 2 \bmod 5 \\
&= 4 + 8 \bmod 5 = 12 \bmod 5 = 2
\end{aligned}$$

Beispiel

Paarungstabellen für Sportturniere

In Stuttgart, München und Berlin soll an 4 Terminen ein Turnier unter 9 Mannschaften ausgetragen werden. Dabei soll „jeder gegen jeden" spielen. Es sind also an jedem Termin 3 Gruppen aus je 3 Mannschaften zu bilden, die jeweils in einer der Städte ihre Spiele untereinander austragen.

Mathematische Formulierung:

$$\begin{aligned} S_{0,k} \cup S_{1,k} \cup S_{2,k} &= \{1, \ldots, 9\}, \qquad k = 0, \ldots, 3\,, \\ \mid S_{j,k} \cap S_{j',k'} \mid &\leq 1, \end{aligned}$$

mit drei-elementigen Mengen $S_{j,k}$, die jeweils der Dreiergruppe in der Stadt j am Termin k entsprechen. Die Bedingung an den Durchschnitt besagt, dass kein Mannschaftspaar doppelt vorkommt.

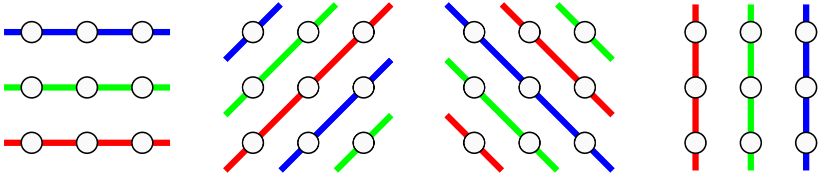
Konstruktion mit Hilfe des Primkörpers $\mathbb{Z}_3 = \{0, 1, 2\}$:
Identifikation der Mannschaften mit Punkten der Ebene $\mathbb{Z}_3^2$, d.h.

$$\{1, 2, ..., 9\} \quad \leftrightarrow \quad \{(x, y) : \; x, y \in \mathbb{Z}_3\}$$

und der Mengen $S_{j,k}$ mit den Geraden in $\mathbb{Z}_3$

$$\begin{aligned} S_{j,k} &= \{(x, k \cdot x + j \bmod 3) : \; x = 0, 1, 2\}, \quad \text{(Steigung } k = 0, 1, 2) \\ S_{j,3} &= \{(j, y) : \; y = 0, 1, 2\} \quad \text{(senkrechte Geraden)} \end{aligned}$$

Durchschnittsbedingung trivialerweise erfüllt: Geraden schneiden sich in höchstens einem Punkt



Paarungstabelle für 16 Mannschaften, 4 Städte und 5 Termine basierend auf 4-elementigen Galois-Körper GF[2^2]

	Spielort 1	Spielort 2	Spielort 3	Spielort 4
1. Spieltag	1,2,3,4	5,6,7,8	9,10,11,12	13,14,15,16
2. Spieltag	1,6,11,16	5,2,15,12	9,14,3,8	13,10,7,4
3. Spieltag	1,10,15,8	5,14,11,4	9,2,7,16	13,6,3,12
4. Spieltag	1,14,7,12	5,10,3,16	9,6,15,4	13,2,11,8
5. Spieltag	1,5,9,13	2,6,10,14	3,7,11,15	4,8,12,16

Galois-Körper GF[q] mit q einer Primzahlpotenz

$\rightsquigarrow$ Paarungstabelle für q^2 Mannschaften und q Städte

Euklidischer Algorithmus

Der größte gemeinsame Teiler zweier natürlicher Zahlen $n_1 > n_2$ kann durch sukzessive Division bestimmt werden. Man berechnet n_{k+1}, $k = 2, 3, \ldots$, als Rest bei der Division $n_{k-1} : n_k$, d.h.

$$n_{k-1} = q_k n_k + n_{k+1}, \quad n_{k+1} < n_k \,,$$

und erhält $\mathrm{GGT}(n_1, n_2) = n_K$ für den Index K mit $n_{K+1} = 0$ (Abbruch der Divisionskette: $n_{K-1} = q_K n_K + 0$).

Beweis

(i) n_K teilt n_1 und n_2:

$n_{K-1} = q_k n_K \quad \implies \quad n_K$ teilt n_{K-1}

rekursive Definition, $n_{k-1} = q_k n_k + n_{k+1}$, für $k = K - 1$

$\implies \quad n_K$ teilt n_{K-2}

Iteration des Arguments für $k = K - 2, \ldots, 3, 2$

$\implies \quad n_K$ teilt $n_{K-3}, \ldots, n_2, n_1$

(ii) n_K ist größter Teiler:

t teilt n_1 und n_2, d.h. $n_1 = s_1 t$, $n_2 = s_2 t$, $\quad \implies$

$$\underbrace{s_1 t}_{n_1} = q_2 \underbrace{s_2 t}_{n_2} + n_3$$

und folglich teilt t ebenfalls n_3, d.h. $n_3 = s_3 t$

Wiederholung des Arguments:

$$n_{k-1} = q_k n_k + n_{k+1} \quad \implies \quad n_{k+1} = s_{k+1} t$$

für $k = 3, 4, \ldots, K - 1$

$\implies \quad n_K = s_K t \geq t$, d.h. n_K ist nicht kleiner als ein Teiler von n_1 und n_2

Beispiel

Größter gemeinsamer Teiler von $n_1 = 156$ und $n_2 = 42$

Euklidischer Algorithmus $\quad\rightsquigarrow\quad$ Divisionskette

$$n_{k-1} = q_k n_k + n_{k+1},\ n_{k+1} < n_k \quad\Longleftrightarrow\quad n_{k-1} : n_k = q_k \text{ Rest } n_{k+1}$$

im konkreten Fall:

$$\begin{aligned}
156 &= 3 \cdot 42 + 30 \\
42 &= 1 \cdot 30 + 12 \\
30 &= 2 \cdot 12 + 6 \\
12 &= 2 \cdot \underbrace{6}_{n_K} + 0 \quad (K = 5)
\end{aligned}$$

$\Longrightarrow \quad \mathrm{GGT}(156, 42) = n_5 = 6$

Beispiel

Berechnung der Inversen x von $m = 81$ modulo $p = 64$, d.h.

$$1 = x \cdot m \bmod p \quad \Longleftrightarrow \quad 1 = x \cdot m + y \cdot p$$

(m, p teilerfremd)

(i) Allgemeine Vorgehensweise:
Euklidischer Algorithmus

$$n_{k-1} = q_k n_k + n_{k+1}, \quad n_{k+1} < n_k$$

mit $n_1 = m$, $n_2 = p$ $\quad \rightsquigarrow \quad$ $n_K = 1$, da $\mathrm{GGT}(n_1, n_2) = 1$, und die vorletzte Gleichung der Divisionskette hat somit die Form

$$n_{K-2} = q_{K-1} n_{K-1} + 1 \quad \text{bzw.} \quad \underbrace{1}_{n_K} = n_{K-2} - q_{K-1} n_{K-1}$$

Rückwärtseinsetzen von

$$n_{K-1} = n_{K-3} - q_{K-2} n_{K-2},\ n_{K-2} = n_{K-4} - q_{K-3} n_{K-3},\ \ldots$$

$\rightsquigarrow \quad 1 = x n_1 + y n_2$, d.h. $x \bmod n_2$ ist die gesuchte Inverse von n_1

(ii) Konkrete Werte $n_1 = m = 81$, $n_2 = p = 64$:
Euklidische Divisionskette

$$\begin{aligned} 81 &= 1 \cdot 64 + \underbrace{17}_{n_3} \\ 64 &= 3 \cdot 17 + 13 \\ 17 &= 1 \cdot 13 + 4 \\ 13 &= 3 \cdot 4 + \underbrace{1}_{n_K} \end{aligned}$$

Aufösen der Gleichungen jeweils nach der Zahl n_j mit dem höchsten Index und Rückwärtseinsetzen

$$\begin{aligned} 13 = 3 \cdot 4 + 1 \quad &\Longrightarrow \quad 1 = 13 - 3 \cdot 4 \\ 4 = 17 - 1 \cdot 13 \quad &\Longrightarrow \quad 1 = 13 - 3 \cdot (17 - 1 \cdot 13) = 4 \cdot 13 - 3 \cdot 17 \\ 13 = 64 - 3 \cdot 17 \quad &\Longrightarrow \quad 1 = 4 \cdot (64 - 3 \cdot 17) - 3 \cdot 17 = 4 \cdot 64 - 15 \cdot 17 \\ 17 = 81 - 1 \cdot 64 \quad &\Longrightarrow \quad 1 = 4 \cdot 64 - 15 \cdot (81 - 1 \cdot 64) = 19 \cdot 64 - 15 \cdot 81 \end{aligned}$$

$\Longrightarrow \quad 1 = -15 \cdot 81 \bmod 64 = 49 \cdot 81 \bmod 64$, d.h. $x = 49$ ist die gesuchte Inverse modulo 64 zu 81

Chinesischer Restsatz

Für teilerfremde natürliche Zahlen $p_1, \dots, p_n$ besitzen die Kongruenzen

$$\begin{aligned} x &= a_1 \bmod p_1 \\ &\dots \\ x &= a_n \bmod p_n \end{aligned}$$

genau eine Lösung $x \in \{0, \dots, P-1\}$, $P = p_1 \cdots p_n$.
Bezeichnet Q_k eine zu $P_k = P/p_k$ inverse ganze Zahl modulo p_k, d.h. ist

$$Q_k P_k = 1 \bmod p_k \,,$$

so gilt

$$x = \sum_{k=1}^{n} a_k Q_k P_k \bmod P \,.$$

Beweis

(i) Existenz:

Darstellung

$$x = \sum_{k=1}^{n} a_k Q_k P_k \bmod P$$

$\Longrightarrow$

$$x \bmod p_\ell = a_\ell Q_\ell P_\ell \bmod p_\ell\,,$$

da p_ℓ Teiler von P_k für $k \neq \ell$

Definition einer zu P_ℓ inversen Zahl Q_ℓ modulo p_ℓ: $Q_\ell P_\ell = 1 \bmod p_\ell$

$\Longrightarrow$

$$x = a_\ell \cdot 1 \bmod p_\ell = a_\ell \bmod p_\ell$$

(ii) Eindeutigkeit:
zu zeigen:

$$x = x' \bmod p_k \text{ für } k = 1, \ldots, n \quad \Longrightarrow \quad x - x' = mP$$

sukzessives Betrachten der Kongruenzen
$x = x' \bmod p_1 \implies$

$$x - x' = m_1 p_1$$

$x = x' \bmod p_2 \quad \implies$

$$m_1 p_1 = 0 \bmod p_2 \quad \Longleftrightarrow \quad m_1 p_1 = s p_2$$

p_1, p_2 teilerfremd $\quad \implies \quad$ p_2 teilt m_1, d.h.

$$m_1 = m_2 p_2, \quad x - x' = m_2 p_1 p_2$$

weitere Kongruenzen $\quad \rightsquigarrow$

$$x - x' = m_3 p_1 p_2 p_3, \ \ldots, \ x - x' = m_n p_1 \cdots p_n$$

Beispiel

Bestimmung einer Lösung x der Kongruenzen

$$\begin{aligned} x &= 6 \bmod 9 \\ x &= 5 \bmod 10 \\ x &= 4 \bmod 13 \end{aligned}$$

Chinesischer Restsatz $\quad \Longrightarrow$

$$x = 6\,Q_1 P_1 + 5\,Q_2 P_2 + 4\,Q_3 P_3 \bmod P, \quad P = 9 \cdot 10 \cdot 13 = 1170$$

mit

$$P_1 = 10 \cdot 13 = 130, \quad P_2 = 9 \cdot 13 = 117, \quad P_3 = 9 \cdot 10 = 90$$

und Q_k der zu P_k inversen natürlichen Zahl modulo p_k, d.h.

$$Q_k P_k + y p_k = 1$$

Bestimmung der Modulo-Inversen

- Q_1:

$$\begin{aligned} \underline{130} &= 14 \cdot \underline{9} + \underline{4} \\ \underline{9} &= 2 \cdot \underline{4} + 1 \end{aligned}$$

Rückwärtseinsetzen $\rightsquigarrow$

$$\begin{aligned} 1 &= 9 - 2 \cdot 4 \\ &= 9 - 2 \cdot (130 - 14 \cdot 9) = 29 \cdot 9 + (-2)130 \end{aligned}$$

$\Longrightarrow \quad Q_1 = (-2) \bmod 9 = 7$

- Q_2:

$$\begin{aligned} \underline{117} &= 11 \cdot \underline{10} + \underline{7} \\ \underline{10} &= 1 \cdot \underline{7} + \underline{3} \\ \underline{7} &= 2 \cdot \underline{3} + 1 \end{aligned}$$

Rückwärtseinsetzen $\rightsquigarrow$

$$\begin{aligned} 1 &= 7 - 2 \cdot 3 \\ &= 7 - 2 \cdot (10 - 1 \cdot 7) = 3 \cdot 7 - 2 \cdot 10 \\ &= 3 \cdot (117 - 11 \cdot 10) - 2 \cdot 10 = 3 \cdot 117 - 35 \cdot 10 \end{aligned}$$

$\Longrightarrow \quad Q_2 = 3 \bmod 10 = 3$

- Q_3:

$$\begin{aligned}\underline{90} &= 6 \cdot \underline{13} + \underline{12} \\ \underline{13} &= 1 \cdot \underline{12} + 1\end{aligned}$$

Rückwärtseinsetzen $\rightsquigarrow$

$$\begin{aligned}1 &= 13 - 1 \cdot 12 \\ &= 13 - 1 \cdot (90 - 6 \cdot 13) = 7 \cdot 13 + (-1) \cdot 90\end{aligned}$$

$$\Longrightarrow \qquad Q_3 = (-1) \bmod 13 = 12$$

Einsetzen in die Darstellung der Lösung

$$\begin{aligned}x &= 6Q_1P_2 + 5Q_2P_2 + 4Q_3P_3 \bmod 1170 \\ &= 6 \cdot 7 \cdot 130 + 5 \cdot 3 \cdot 117 + 4 \cdot 12 \cdot 90 \bmod 1170 \\ &= 11535 \bmod 1170 = 1005\end{aligned}$$

Vektorraum

Ein Vektorraum über einem Körper K (K-Vektorraum) ist eine kommutative Gruppe $(V, +)$, auf der zusätzlich zu der Gruppenoperation „$+$“ eine Skalarmultiplikation „$\cdot$“ definiert ist,

$$K \times V \ni (s, v) \mapsto s \cdot v \in K\,,$$

die folgende Eigenschaften besitzt:

$$\begin{aligned}(s_1 + s_2) \cdot v &= s_1 \cdot v + s_2 \cdot v \\ s \cdot (v_1 + v_2) &= s \cdot v_1 + s \cdot v_2 \\ (s_1 \cdot s_2) \cdot v &= s_1 \cdot (s_2 \cdot v) \\ 1 \cdot v &= v\end{aligned}$$

für alle Skalare $s, s_1, s_2 \in K$, Elemente v, v_1, $v_2 \in V$ und das Einselement $1 \in K$.

Der Einfachheit halber wird das Pluszeichen sowohl für die Addition in V als auch für die Addition in K verwendet. Ebenso wird der Malpunkt für die Skalarmultiplikation meist weggelassen.

Für $K = \mathbb{R}$ bzw. $K = \mathbb{C}$ spricht man von einem reellen bzw. komplexen Vektorraum.

Beispiel

Reeller (komplexer) Vektorraum der Polynome p vom Grad $\leq n$

$$p(x) = a_0 + a_1 x + \cdots + a_n x^n, \quad a_k \in \mathbb{R} \quad (a_k \in \mathbb{C})$$

Definition der Addition und Skalarmultiplikation in der nahe liegenden Weise:

$$(p+q)(x) = p(x) + q(x), \quad (sp)(x) = sp(x)$$

Polynome mit Grad n ($a_n \neq 0$) bilden auf Grund eventueller Gradreduktion bei Addition keinen Vektorraum,
z.B.

$$(\underbrace{x^2 - 1}_{\text{Grad } 2}) + (\underbrace{3 - x^2}_{\text{Grad } 2}) = \underbrace{2}_{\text{Grad } 0}$$

Beispiel

Reeller Vektorraum der Folgen (a_n), $a_n \in \mathbb{R}$

- Addition: $(a_n) + (b_n) = (a_n + b_n)$
- Skalarmultiplikation: $s(a_n) = (sa_n)$

Vektorräume spezieller Folgen:
beschränkte Folgen, konvergente Folen, konvergente Folen, komplexe Folgen

Monotone Folgen bilden keinen Vektorraum, denn Summen monotoner Folgen sind nicht notwendig monoton; z.B.

$$(n^2) + (-2^n) : \; -1, 0, 1, 0, -7, -28, \ldots .$$

Für einen Körper K bilden die n-Tupel oder n-Vektoren

$$a = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}, \quad a_k \in K$$

den K-Vektorraum K^n mit der komponentenweise definierten Addition und Skalarmultiplikation, d.h.

$$\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} + \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} a_1 + b_1 \\ \vdots \\ a_n + b_n \end{pmatrix}, \quad s \cdot \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} s \cdot a_1 \\ \vdots \\ s \cdot a_n \end{pmatrix}$$

für $a_k, b_k, s \in K$.

Oft schreibt man n-Tupel als Zeilenvektor

$$a^{\mathrm{t}} = (a_1, \ldots, a_n) \quad \text{bzw.} \quad a = (a_1, \ldots, a_n)^{\mathrm{t}}\,.$$

Durch das Symbol „t" der Transposition wird von der Standardkonvention als Spaltenvektor unterschieden.

Für $K = \mathbb{R}$ bzw. $K = \mathbb{C}$ erhält man die Vektorräume der n-Tupel reeller und komplexer Zahlen $\mathbb{R}^n$ und $\mathbb{C}^n$.

Beispiel

Addition und Skalarmultiplikation in $\mathbb{R}^3$ und $\mathbb{C}^2$

(i) Reelle (Spalten-) Vektoren im Raum:

$$\begin{pmatrix} 1 \\ -2 \\ 3 \end{pmatrix} + \begin{pmatrix} -4 \\ 5 \\ -6 \end{pmatrix} = \begin{pmatrix} 1+(-4) \\ -2+5 \\ 3+(-6) \end{pmatrix} = \begin{pmatrix} -3 \\ 3 \\ -3 \end{pmatrix}$$

$$2\begin{pmatrix} 3 \\ -1 \\ 4 \end{pmatrix} = \begin{pmatrix} 2\cdot(3) \\ 2\cdot(-1) \\ 2\cdot 4 \end{pmatrix} = \begin{pmatrix} 6 \\ -2 \\ 8 \end{pmatrix}$$

(ii) (Zeilen-) Vektoren in der komplexen Ebene:

$$\begin{pmatrix} 1, \mathrm{i} \end{pmatrix} + \begin{pmatrix} 2-\mathrm{i}, 1+3\mathrm{i} \end{pmatrix} = \begin{pmatrix} 3-\mathrm{i}, 1+4\mathrm{i} \end{pmatrix}$$

$$(2-3\mathrm{i})\begin{pmatrix} 1+\mathrm{i}, 1-\mathrm{i} \end{pmatrix} = \begin{pmatrix} 5-\mathrm{i}, -1-5\mathrm{i} \end{pmatrix}$$

Unterraum

Ein Unterraum U eines K-Vektorraums V besteht aus Elementen $u \in U \subseteq V$, die mit der in V definierten Addition und Skalarmultiplikation selbst einen Vektorraum bilden.

Um zu prüfen, ob $U \subset V$ ein Unterraum ist, genügt es zu zeigen, dass U bzgl. der Addition und Skalarmultiplikation abgeschlossen ist:

$$
\begin{aligned}
u, v \in U \quad &\Longrightarrow \quad u + v \in U \\
s \in K, u \in U \quad &\Longrightarrow \quad s \cdot u \in U\,.
\end{aligned}
$$

Unterräume U werden oft durch Bedingungen an die Elemente von V definiert:

$$U = \{u \in V : \, A(u)\},$$

mit einer Aussage A, die für Elemente von U erfüllt sein muss.

Beispiel

Unterräume des Vektorraums der reellen Funktionen

$$f : \mathbb{R} \to \mathbb{R},$$

definiert durch zusätzliche Eigenschaften

Eigenschaft	Unterraum
(un)gerade	ja
beschränkt	ja
monoton	nein
stetig	ja
positiv	nein
linear	ja

exemplarische Begründungen

- beschränkt:

$$
\begin{aligned}
|f_k| \le c_k \quad &\Longrightarrow \quad |f_1(x) + f_2(x)| \le |f_1(x)| + |f_2(x)| \le c_1 + c_2 \\
|f| \le c \quad &\Longrightarrow \quad |sf(x)| \le sc
\end{aligned}
$$

  jeweils $\forall x$

  $\rightsquigarrow$ $f_1 + f_2$ und sf sind beschränkt

  $\rightsquigarrow$ Unterraumkriterium erfüllt

- monoton:

$$
h(x) = \exp(x) - x \quad \text{nicht monoton wegen} \quad \lim_{x \to \pm\infty} h(x) = +\infty
$$

  trotz Monotonie von $f(x) = \exp(x)$ und $g(x) = x$

Beispiel

Gerade in einem K-Vektorraum V:

$$U:\ u = a + tb, \quad t \in K$$

mit fest gewählten Elementen $a, b \in V$

(i) $a = 0 \quad \rightsquigarrow \quad$ Unterraum:

$$u_1, u_2 \in U \quad \Longrightarrow \quad u_1 + u_2 = t_1 b + t_2 b = (\underbrace{t_1 + t_2}_{t}) b \in U$$

$$s \in K,\ u \in U \quad \Longrightarrow \quad su = s(tb) = (st)b \in U$$

(ii) $0 \notin U$, d.h. $a \neq 0$ und $b \neq sa \quad \rightsquigarrow \quad$ kein Unterraum:

$$u \in U \quad \Longrightarrow \quad su = s(a + tb) = 0 \notin U \quad \text{für } s = 0$$

Linearkombination

Für Elemente $v_1, v_2, \ldots, v_m$ eines K-Vektorraums V bezeichnet man

$$s_1 v_1 + s_2 v_2 + \cdots + s_m v_m = \sum_{k=1}^{m} s_k v_k$$

mit Skalaren $s_k \in K$ als Linearkombination der Elemente v_k.

Die Menge aller solchen Linearkombinationen nennt man die lineare Hülle der v_k:

$$\operatorname{span}(v_1, \ldots, v_m) = \left\{ \sum_{k=0}^{m} s_k v_k : \, s_k \in K \right\} .$$

Allgemeiner definiert man für eine Menge $U \subseteq V$

$$\operatorname{span}(U) = \left\{ \sum_{k} s_k u_k : \, s_k \in K, \, u_k \in U \right\}$$

als die Menge aller Linearkombinationen von Elementen aus der Menge U. Die so gewonnene lineare Hülle von U ist ein Unterraum von V.

Beispiel

Linearkombinationen in $\mathbb{R}^n$

(i) Der Vektor $v = (1, 2, 3)^{\mathrm{t}}$ ist eine Linearkombination der Vektoren

$$v_1 = (3, 4, 5)^{\mathrm{t}},\ v_2 = (1, 1, 1)^{\mathrm{t}}\,,$$

denn $v = v_1 - 2v_2$.

(ii) Der Vektor $v = (1, 0)^{\mathrm{t}}$ ist keine Linearkombination der Vektoren

$$v_1 = (0, 1)^{\mathrm{t}},\ v_2 = (0, 2)^{\mathrm{t}}\,,$$

denn jede Linearkombination von v_1 und v_2 hat die Form $(0, x)^{\mathrm{t}}$.

(iii) Der Vektor $v = (0, 0, 0, 0)^{\mathrm{t}}$ ist auf verschiedene Art als Linearkombination von

$$v_1 = (1, 1, 0, 0)^{\mathrm{t}},\ v_2 = (0, 2, 2, 0)^{\mathrm{t}},\ v_3 = (0, 0, 3, 3)^{\mathrm{t}},\ v_4 = (4, 0, 0, 4)^{\mathrm{t}}$$

darstellbar:

$$v = s(12v_1 - 6v_2 + 4v_3 - 3v_4), \quad s \in \mathbb{R}$$

Beispiel

Lineare Hüllen in $\mathbb{R}^3$

(i) $v_1 = (1, -1, 0)^t$, $v_2 = (0, 1, -1)^t$:

$$\begin{aligned} \operatorname{span}(v_1, v_2) &= \{s_1(1, -1, 0)^t + s_2(0, 1, -1)^t : s_k \in \mathbb{R}\} \\ &= \{(s_1, s_2 - s_1, -s_2)^t : s_k \in \mathbb{R}\} \end{aligned}$$

Ebene, orthogonal zu $(1, 1, 1)^t$
(ii) $v_1 = (1, -1, 0)^t$, $v_2 = (0, 1, -1)^t$, $v_3 = (-1, 0, 1)^t$:
gleiche lineare Hülle, da

$$v_3 = -v_1 - v_2$$

(Darstellung von v_3 als Linearkombination von v_1 und v_2)
keine eindeutige Darstellung der Vektoren in $\operatorname{span}(v_1, v_2, v_3)$, z.B.:

$$(0, 0, 0)^t = s(v_1 + v_2 + v_3), \quad s \in \mathbb{R}$$

Konvexkombination

Eine Konvexkombination ist eine Linearkombination

$$s_1 v_1 + s_2 v_2 + \cdots + s_m v_m$$

von Elementen v_k eines reellen Vektorraums mit

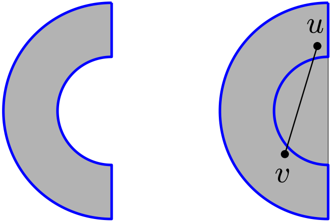
$$s_k \geq 0, \quad \sum_k s_k = 1\,.$$

Die Menge aller Konvexkombinationen von Elementen aus einer Teilmenge $M \subseteq V$ wird als konvexe Hülle von M, $\operatorname{conv}(M)$, bezeichnet.

Geometrisch ist $\operatorname{conv}(M)$ die kleinste M enthaltende Menge, die für je zwei Elemente u, v auch deren Verbindungsstrecke

$$(1-s)u + sv, \quad 0 \leq s \leq 1\,,$$

enthält.



Beispiel

Parametrisierung einer Gerade und eines konvexen Vierecks durch lineare Interpolation

(i) Gerade G:

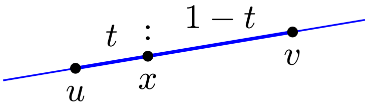
$u, v \in G$, $u \neq v \quad \Longrightarrow$

$$G:\ x = (1-t)u + tv, \quad t \in \mathbb{R}$$

Konvexkombination für $0 \leq t \leq 1$

$\rightsquigarrow$ Geradensegment $\operatorname{conv}(\{u, v\})$ zwischen u und v

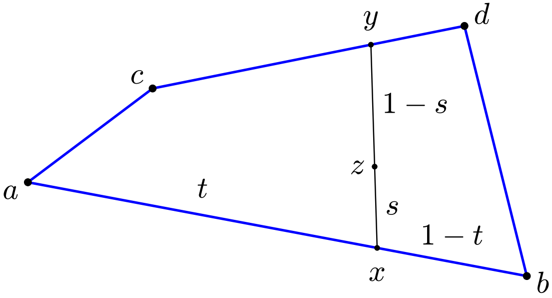
x teilt $\operatorname{conv}(\{u, v\})$ im Verhältnis $t : (1-t)$



(ii) Konvexes Viereck $\square(a,b,c,d)$:
$z \in \square(a,b,c,d) \quad \Longleftrightarrow$

$$\begin{aligned} z &= (1-s)[\underbrace{(1-t)a+tb}_{x}] + s[\underbrace{(1-t)c+td}_{y}] \\ &= (1-s)(1-t)a + (1-s)tb + s(1-t)c + std \end{aligned}$$

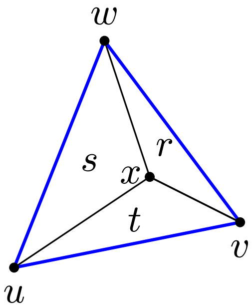
mit $0 \leq s, t \leq 1$
Konvexkombination von a, b, c, d: Koeffizienten ≥ 0 mit Summe 1



Die Konvexkombinationen von $u, v, w \in \mathbb{R}^2$,

$$x = ru + sv + tw, \quad r, s, t \geq 0,\ r + s + t = 1$$

bilden ein Dreieck mit Eckpunkten u, v, w. Die Koeffizienten r, s, t werden als baryzentrische Koordinaten bezeichnet. Sie können durch Lösen des definierenden linearen Gleichungssystems bestimmt werden.



$$r = \frac{\operatorname{area}(x, v, w)}{\operatorname{area}(u, v, w)}$$

$$s = \frac{\operatorname{area}(u, x, w)}{\operatorname{area}(u, v, w)}$$

$$t = \frac{\operatorname{area}(u, v, x)}{\operatorname{area}(u, v, w)}$$

Geometrisch entsprechen die baryzentrischen Koordinaten den Verhältnissen der Flächeninhalte der durch x definierten Teildreiecke zum Flächeninhalt des Dreiecks $\Delta(u, v, w)$.
Analog können die baryzentrischen Koordinaten der Punkte in einem Tetraeder im $\mathbb{R}^3$ definiert werden.

Lineare Unabhängigkeit

Elemente $v_1, \dots, v_m$ eines Vektorraums V sind linear abhängig, wenn ein Element v_k als Linearkombination der anderen darstellbar ist, d.h.

$$s_1 v_1 + \cdots + s_m v_m = 0_V \qquad (*)$$

mit mindestens einem der Skalare $s_k \neq 0$.

Andernfalls bezeichnet man $v_1, \dots, v_m$ als linear unabhängig. In diesem Fall folgt aus $\sum_k s_k v_k = 0_V$ dass $s_1 = \cdots = s_m = 0$.

Für Vektoren $v_k \in \mathbb{R}^n$ ist die Bedingung (*) ein homogenes lineares Gleichungssystem, das bei linearer Unabhängigkeit von $v_1, \dots, v_m$ nur die triviale Lösung $s_1 = \cdots = s_m = 0$ besitzt.

Beispiel

Lineare Unabhängigkeit von Vektoren in der Ebene

(i) Zwei Vektoren $u, v \in \mathbb{R}^2$ sind genau dann linear unabhängig, wenn keiner der beiden ein Vielfaches des anderen ist.

konkrete Beispiele:

- $(1,0)^t$, $(1,1)^t$ sind linear unabhängig, denn

$$s(1,0)^t + t(1,1)^t = (0,0)^t \quad \Longrightarrow \quad s = t = 0$$

- $(2,3)^t$, $(-4,-6)^t$ sind linear abhängig, denn

$$(-2)(2,3)^t = (-4,-6)^t \quad \Longrightarrow \quad 2(2,3)^t + (-4,6)^t = (0,0)^t\,,$$

d.h. $\exists$ eine nichttriviale Darstellung des Nullvektors

(ii) Drei Vektoren $u, v, w \in \mathbb{R}^2$ sind immer linear abhängig.

$$ru + sv + tw = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$\rightsquigarrow$ unterbestimmtes, homogenes lineares Gleichungssystem

$$\begin{array}{ccccccc} ru_1 & + & sv_1 & + & tw_1 & = & 0 \\ ru_2 & + & sv_2 & + & tw_2 & = & 0 \end{array}$$

für r, s, t, das immer eine nichttriviale Lösung besitzt.

konkretes Beispiel:

$$r \begin{pmatrix} 0 \\ 1 \end{pmatrix} + s \begin{pmatrix} 1 \\ 2 \end{pmatrix} + t \begin{pmatrix} 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

mit der Lösung $r = \lambda$, $s = -2\lambda$ und $t = \lambda$ für beliebiges $\lambda \in \mathbb{R}$

Beispiel

Lineare Unabhängigkeit von Vektoren im Raum

(i) Zwei Vektoren $u, v \in \mathbb{R}^3$ sind linear abhängig, wenn sie parallel sind, d.h. wenn ein Vektor ein Vielfaches des anderen ist.

(ii) Drei Vektoren $u, v, w \in \mathbb{R}^3$ sind linear abhängig, wenn zwei Vektoren parallel sind oder wenn ein Vektor in der von den beiden anderen Vektoren aufgespannten Ebene liegt.

(iii) Vier und mehr Vektoren im $\mathbb{R}^3$ sind immer linear abhängig.

Test für lineare Unabhängigkeit:
$s_1 = \cdots = s_n = 0$ ist die einzige Lösung des homogenes linearen Gleichungssystems

$$s_1 \begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} + \cdots + s_n \begin{pmatrix} x_n \\ y_n \\ z_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Konkrete Fälle:

- $u = (0,1,2)^{\mathrm{t}}$, $v = (1,2,3)^{\mathrm{t}}$ sind linear unabhängig, da $u \nparallel v$
- (ii) $u = (1,0,0)^{\mathrm{t}}$, $v = (2,3,0)^{\mathrm{t}}$, $w = (4,5,6)^{\mathrm{t}}$ sind linear unabhängig, da

$$r \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + s \begin{pmatrix} 2 \\ 3 \\ 0 \end{pmatrix} + t \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad \Longrightarrow \quad r = s = t = 0$$

- (iii) $(1,0,0)^{\mathrm{t}}$, $(0,2,0)^{\mathrm{t}}$, $(0,0,3)^{\mathrm{t}}$, $(4,5,6)^{\mathrm{t}}$ sind linear abhängig, da

$$5 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + 2 \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix} + 2 \begin{pmatrix} 0 \\ 0 \\ 3 \end{pmatrix} + (-1) \begin{pmatrix} 5 \\ 4 \\ 6 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Basis

Eine Teilmenge B eines Vektorraumes V ist eine Basis von V, wenn die Vektoren in B linear unabhängig sind und sich jeder Vektor $v \in V$ eindeutig als Linearkombination

$$v = \sum_k c_k b_k, \quad b_k \in B,$$

darstellen lässt. Die Koeffizienten c_k werden als Koordinaten von v bezüglich der Basis B bezeichnet:

$$v \leftrightarrow v_B = (c_1, c_2, \ldots)^{\mathrm{t}}.$$

Besitzt ein Vektorraum V eine endliche Basis $B = \{b_1, \ldots, b_n\}$, so ist die Anzahl der Basisvektoren eindeutig bestimmt und wird die Dimension von V genannt:

$$n = \dim V.$$

Man setzt $\dim V = 0$ für $V = \{0\}$ und $\dim V = \infty$ für einen Vektorraum ohne endliche Basis.

Für $V = \mathbb{R}^n$ oder $V = \mathbb{C}^n$ besteht eine Basis aus n Vektoren b_k, und die Matrix $B = (b_1, \ldots, b_n)$ mit den Spalten b_k ist invertierbar, d.h. $\det B \neq 0$.

Beweis

Eindeutigkeit der Dimension im endlichen Fall

hinreichend zu zeigen:
Hat ein Vektorraum eine n-elementige Basis

$$b_1, \ldots, b_n\,,$$

so sind $n+1$ Vektoren $v_1, \ldots, v_{n+1}$ (und damit auch mehr als $n+1$ Vektoren) linear abhängig.
($\implies$ Widerspruch zur linearen Unabhängigkeit bei Basen mit unterschiedlich vielen Vektoren)

Beweis durch Induktion:
$(n-1) \to n$: betrachte Basisdarstellung der Vektoren v_k:

$$v_k = \sum_{\ell=1}^{n} \gamma_{k,\ell} b_\ell, \quad k = 1, \ldots, n+1$$

trivialer Fall:

$$\gamma_{n+1,1} = \cdots = \gamma_{n+1,n} = 0$$

$$\implies \quad v_{n+1} = 0 \quad \implies \quad \text{lineare Abhängigkeit der Vektoren } v_k$$

andernfalls, nach geeigneter Nummerierung, $\gamma_{n+1,n} \neq 0$
definiere Vektoren v'_k, $k = 1, \ldots, n$, die sich als Linearkombination der $n-1$ Vektoren $b_1, \ldots, b_{n-1}$ darstellen lassen:

$$\begin{aligned} v'_k &= v_k - \frac{\gamma_{k,n}}{\gamma_{n+1,n}} v_{n+1} \\ &= (\gamma_{k,1} b_1 + \cdots + \gamma_{k,n} b_n) - \frac{\gamma_{k,n}}{\gamma_{n+1,n}} (\gamma_{n+1,1} b_1 + \cdots + \gamma_{n+1,n} b_n) \end{aligned}$$

Koeffizient von $b_n = 0 \quad \Longrightarrow$

$$v'_1, \ldots, v'_n \in V' = \operatorname{span}\{b_1, \ldots, b_{n-1}\}$$

Induktionsvoraussetzung, angewandt auf die Basis $\{b_1, \ldots, b_{n-1}\}$ von V'
$\Longrightarrow \quad \exists$ nichttriviale Linearkombination

$$\lambda_1 v'_1 + \cdots + \lambda_n v'_n = 0$$

Einsetzen der Definition von $v'_k \quad \rightsquigarrow$
Linearkombination der v_k, also die behauptete lineare Abhängigkeit

Reelles Skalarprodukt

Ein Skalarprodukt $\langle \cdot, \cdot \rangle$ auf einem reellen Vektorraum V ist eine Abbildung

$$V \times V \ni (u, v) \mapsto \langle u, v \rangle \in \mathbb{R}$$

mit folgenden Eigenschaften:

- Positivität:
$$\langle v, v \rangle > 0 \quad \text{für } v \neq 0$$
- Symmetrie:
$$\langle u, v \rangle = \langle v, u \rangle$$
- Linearität:
$$\langle su + tv, w \rangle = s\langle u, w \rangle + t\langle v, w \rangle$$

Diese Identitäten gelten für alle $u, v, w \in V$ und $s, t \in \mathbb{R}$.

Aufgrund der Symmetrie ist ein reelles Skalarprodukt auch bezüglich des zweiten Argumentes linear, also eine Bilinearform auf V.

Für $V = \mathbb{R}^n$ ist das kanonische Skalarprodukt

$$\langle u, v\rangle = u^{\mathrm{t}} v = \begin{pmatrix} u_1 & \cdots & u_n \end{pmatrix} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = u_1 v_1 + \cdots + u_n v_n \,,$$

und

$$|u| = \sqrt{\langle u, u\rangle} = \sqrt{u_1^2 + \cdots + u_n^2}$$

ist die assoziierte Norm.

Beispiel

Skalarprodukt-Eigenschaften für Abbildungen

$$\mathbb{R}^2 \ni (x, y) \mapsto f(x, y) \in \mathbb{R}$$

f(x,y)	Positivität	Symmetrie	Linearität
$10x_1y_1 + x_2y_2$	✓	✓	✓
$\lvert x_1y_1\rvert + \lvert x_2y_2\rvert$	✓	✓	
$x_1y_2 + x_2y_1$		✓	✓
$4x_1y_1 + x_1y_2 + 2x_2y_1 + 3x_2y_2$	✓		✓
$x_1y_1^3 + x_2^3y_2$	✓		
$x_1x_2 + y_1y_2$		✓	
$x_1y_2 + 2x_2y_1$			✓
x_1^3			

Beispiel

Skalarprodukt auf dem Vektorraum der auf $[0,1]$ definierten, reellwertigen stetigen Funktionen:

$$\langle f, g\rangle = \int\limits_0^1 f(x)g(x)\,\mathrm{d}x$$

Positivität, Linearität, Symmetrie $\checkmark$

Verallgemeinerung mit einer positiven Gewichtsfunktion w:

$$\langle f, g\rangle_w = \int\limits_0^1 fg\,w$$

z.B: gewichtete Skalarprodukte für radialsymmetrische Funktionen auf der Kreisscheibe oder Kugel:

$$\int_0^1 f(r)g(r)r\,\mathrm{d}r, \quad \int_0^1 f(r)g(r)r^2\,\mathrm{d}r$$

Komplexes Skalarprodukt

Ein Skalarprodukt $\langle \cdot, \cdot \rangle$ auf einem komplexen Vektorraum V ist eine Abbildung

$$V \times V \ni (u, v) \mapsto \langle u, v \rangle \in \mathbb{C}$$

mit folgenden Eigenschaften:

- Positivität: $\langle v, v \rangle > 0 \quad \text{für} \ \ v \neq 0$
- Schiefsymmetrie: $\langle u, v \rangle = \overline{\langle v, u \rangle}$
- Linearität: $\langle u, sv + tw \rangle = s\langle u, v \rangle + t\langle u, w \rangle$

Diese Identitäten gelten für alle $u, v, w \in V$ und $s, t \in \mathbb{C}$.

Aufgrund der Schiefsymmetrie ist ein komplexes Skalarprodukt bezüglich der ersten Variablen nicht linear:

$$\langle su + tv, w \rangle = \bar{s}\langle u, v \rangle + \bar{t}\langle u, w \rangle \, .$$

Lediglich für reelle Skalare s, t ist die komplexe Konjugation ohne Bedeutung.

Für $V = \mathbb{C}^n$ ist das kanonische Skalarprodukt

$$\langle u, v\rangle = u^* v = \bar{u}^{\mathrm{t}} v = \begin{pmatrix} \bar{u}_1 & \cdots & \bar{u}_n \end{pmatrix} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \bar{u}_1 v_1 + \cdots + \bar{u}_n v_n \,,$$

und

$$|u| = \sqrt{\langle u, u\rangle} = \sqrt{|u_1|^2 + \cdots + |u_n|^2}, \quad |u_k|^2 = \bar{u}_k u_k \,,$$

ist die assoziierte Norm.

Gebräuchlich ist auch die komplexe Konjugation des zweiten Arguments des Skalarprodukts,

$$\langle u, v\rangle = \sum_k u_k \bar{v}_k \,.$$

Diese andere Definition des Skalarprodukts komplexer Vektoren ist jedoch weniger konsistent mit den Regeln des Matrix/Vektor-Kalküls - den „liegenden“ adjungierten Vektor u^* erhält man durch Transposition und Konjugation des „stehenden“ Vektors u.

Beispiel

Skalarprodukt von Vektoren in $\mathbb{C}^2$

$$x = \begin{pmatrix} 1+2\mathrm{i} \\ -2-\mathrm{i} \end{pmatrix}, \quad y = \begin{pmatrix} 2 \\ 2\mathrm{i} \end{pmatrix}$$

komplexes Skalarprodukt $\langle x, y\rangle = \bar{x}_1 y_1 + \bar{x}_2 y_2$

$$(\overline{1+2\mathrm{i}})\cdot 2 + (\overline{-2-\mathrm{i}})\cdot 2\mathrm{i} = (1-2\mathrm{i})\cdot 2 + (-2+\mathrm{i})\cdot 2\mathrm{i} = 2-4\mathrm{i}-4\mathrm{i}-2 = -8\mathrm{i}$$

Das Konjugieren ist notwendig für die Positivität der assoziierten Norm.
keine Konjugation $\quad\rightsquigarrow\quad$ falsche Definition der Längen

$$\begin{aligned}
\sqrt{x_1^2+x_2^2} &= \sqrt{(1+2\mathrm{i})^2+(-2-\mathrm{i})^2} \\
&= \sqrt{1+4\mathrm{i}-4+4+4\mathrm{i}-1} = \sqrt{8\mathrm{i}} \notin \mathbb{R} \\
\sqrt{y_1^2+y_2^2} &= \sqrt{2^2+(2\mathrm{i})^2} = \sqrt{4-4} = 0 \not> 0
\end{aligned}$$

richtige Berechnung

$$
\begin{aligned}
|x| &= \sqrt{\bar{x}_1 x_1 + \bar{x}_2 x_2} \\
&= \sqrt{(1+4)+(4+1)} = \sqrt{10} \\
|y| &= \sqrt{2 \cdot 2 + (2\mathrm{i}) \cdot (-2\mathrm{i})} = \sqrt{8}
\end{aligned}
$$

Cauchy-Schwarz-Ungleichung

Ein Skalarprodukt lässt sich mit Hilfe der assoziierten Norm abschätzen:

$$|\langle u, v\rangle| \le |u||v|, \quad |w| = \sqrt{\langle w, w\rangle}\,.$$

Gleichheit gilt genau dann, wenn $u \parallel v$.
Für ein reelles Skalarprodukt kann durch

$$\cos\varphi = \frac{\langle u, v\rangle}{|u|\ |v|}$$

ein Winkel $\varphi \in [0, \pi]$ zwischen u und v definiert werden.

Beweis

$v = su \quad \rightsquigarrow \quad$ Gleichheit

Die Ungleichung bleibt bei Multiplikation von u bzw. v mit einem Skalar unverändert.

$\rightsquigarrow \quad$ o.B.d.A. $|u|^2 = |v|^2 = 1$, $u \not\parallel v$

betrachte ein komplexes Skalarprodukt (Argumentation schließt den reellen Fall ein)

$$\langle v, u\rangle = r \underbrace{\exp(\mathrm{i}\varphi)}_{=:\lambda}, \quad r > 0\,,$$

u, v nicht parallel, $\bar{\lambda}\lambda = 1 \quad \Longrightarrow$

$$\begin{aligned} 0 &< \langle u - \lambda v, u - \lambda v\rangle \\ &= |u|^2 + \bar{\lambda}\lambda|v|^2 - \bar{\lambda}\underbrace{r\exp(\mathrm{i}\varphi)}_{\langle v,u\rangle} - \lambda\underbrace{\overline{r\exp(\mathrm{i}\varphi)}}_{\langle u,v\rangle} = 1 + 1 - r - r\,, \end{aligned}$$

d.h.

$$|\langle u, v\rangle| = r < 1 = |u|\,|v| \quad \checkmark$$

Beispiel

Illustration der Ungleichung von Cauchy-Schwarz für das Skalarprodukt

$$\langle f, g \rangle = \int_0^1 f(x)g(x)\,\mathrm{d}x$$

$f_k(x) = x^k \quad \rightsquigarrow$

$$\begin{aligned} |f_k| &= \left(\int_0^1 (x^k)^2\,\mathrm{d}x\right)^{1/2} = (2k+1)^{-1/2} \\ \langle f_k, f_\ell \rangle &= \int_0^1 x^k x^\ell\,\mathrm{d}x = (k+\ell+1)^{-1} \end{aligned}$$

Ungleichung von Cauchy-Schwarz

$$(k+\ell+1)^{-1} = |\langle f_k, f_\ell \rangle| \leq |f_k||f_\ell| = (2k+1)^{-1/2}(2\ell+1)^{-1/2}$$

$\checkmark$, denn Quadrieren und Kehrwertbildung $\quad \rightsquigarrow$

$$k^2 + \ell^2 + 1 + 2k\ell + 2k + 2\ell \geq 4k\ell + 2k + 2\ell + 1 \quad \iff \quad (k-\ell)^2 \geq 0$$

Norm

Eine Norm auf einem reellen oder komplexen Vektorraum V ist eine Abbildung

$$V \ni v \mapsto \|v\| \in \mathbb{R}$$

mit den folgenden Eigenschaften.

- Positivität:

$$\|v\| > 0 \quad \text{für } v \neq 0$$

- Homogenität:

$$\|sv\| = |s|\|v\|$$

- Dreiecksungleichung:

$$\|u + v\| \leq \|u\| + \|v\|$$

Diese Identitäten gelten für alle $u, v \in V$ und $s \in \mathbb{R}$ bzw. $s \in \mathbb{C}$.

Mit einem Skalarprodukt ist die Norm

$$|v| = \sqrt{\langle v, v \rangle}, \quad v \in V$$

assoziiert. Für diese spezielle Norm werden oft einfache Betragsstriche verwendet. Insbesondere ist für $V = \mathbb{R}^n$ und $V = \mathbb{C}^n$

$$|x| = \sqrt{|x_1|^2 + \cdots + |x_n|^2} \, .$$

Mit Hilfe einer Norm kann durch

$$d(u, v) = \|u - v\|$$

ein Abstand zwischen zwei Vektoren definiert werden.

Beweis

Überprüfung der Normeigenschaften für die einem Skalarprodukt zugeordnete Norm

- Positivität und Homogenität $\checkmark$
- Die Dreiecksungleichung folgt aus

$$\begin{aligned}
|u+v|^2 &= \langle u+v, u+v\rangle \\
&= \langle u,u\rangle + \underbrace{\langle u,v\rangle + \overline{\langle u,v\rangle}}_{\in\mathbb{R}} + \langle v,v\rangle \\
&\leq |u|^2 + 2|\langle u,v\rangle| + |v|^2 \\
&\underset{\text{Cauchy-Schwarz}}{\leq} |u|^2 + 2|u||v| + |v|^2 \\
&= (|u|+|v|)^2
\end{aligned}$$

durch Wurzelziehen.

Beispiel

Normen für die Vektorräume $\mathbb{R}^n$ und $\mathbb{C}^n$

- 2-Norm oder Euklidische Norm

$$\|z\|_2 = \sqrt{|z_1|^2 + \cdots + |z_n|^2}$$

z.B.

$$\|(2, -1, 2)^{\mathrm{t}}\|_2 = \sqrt{4 + 1 + 4} = 9$$

Spezielle Notation für die, dem kanonischen Skalarprodukt zugeordnete Norm: $\|z\|_2 = |z|$

- Maximum-Norm

$$\|z\|_\infty = \max_k |z_k|$$

z.B:

$$\begin{aligned} \|(2+3\mathrm{i},\, 3-4\mathrm{i})^{\mathrm{t}}\|_\infty &= \max(\sqrt{2^2+3^2},\, \sqrt{3^2+4^2}) \\ &= \max(\sqrt{13},\, \sqrt{25}) = 5 \end{aligned}$$

Verallgemeinerung: $\|z\|_{\infty,w} = \max_k(w_k|z_k|)$ mit Gewichten $w_k > 0$

- 1-Norm

$$\|z\|_1 = \sum_{k=1}^{n} |z_k|$$

z.B.

$$\|(1,\, -2,\, 3)^{\mathrm{t}}\|_1 = |1| + |-2| + |3| = 1+2+3 = 6$$

Orthogonale Basis

Eine Basis $B = \{u_1, \ldots, u_n\}$ eines Vektorraums V ist orthogonal, wenn

$$\langle u_j, u_k \rangle = 0, \quad j \neq k\,.$$

Eine normierte orthogonale Basis, d.h. $|u_k| = 1\,\forall\, k$, wird als Orthonormalsystem oder Orthonormalbasis bezeichnet.
Die Elemente v des Vektorraums besitzen die Darstellung

$$v = \sum_{k=1}^{n} c_k u_k\,, \quad c_k = \frac{\langle u_k, v \rangle}{|u_k|^2}\,.$$

Für die Koeffizienten c_k gilt

$$|c_1|^2 |u_1|^2 + \cdots + |c_n|^2 |u_n|^2 = |v|^2\,.$$

Ist die Basis normiert, so fallen die Terme $|u_k|^2$ weg, d.h. es gelten die einfacheren Formeln

$$c_k = \langle u_k, v \rangle, \quad |c_1|^2 + \cdots + |c_n|^2 = |v|^2\,.$$

Für einen komplexen Vektorraum ist die Reihenfolge der Argumente im Skalarprodukt von Bedeutung. Bei der Berechnung der Koeffizienten muss $\langle \cdot, \cdot \rangle$ bezüglich des Argumentes v linear sein.

Beweis

$u_1, \dots, u_n$ Basis $\quad \Longrightarrow \quad$ Existenz von Skalaren $c_1, \dots, c_n$ mit

$$v = \sum_{j=1}^{n} c_j u_j$$

Skalarprodukt mit u_k, Orthogonalität der Basis ($\langle u_k, u_j \rangle = 0$, $j \neq k$) und Linearität $\quad \rightsquigarrow$

$$\langle u_k, v \rangle = \langle u_k, c_k u_k \rangle = c_k \langle u_k, u_k \rangle \quad \text{bzw.} \quad c_k = \frac{\langle v, u_k \rangle}{\langle u_k, u_k \rangle}$$

Identität für die Koeffizienten $\quad \widehat{=} \quad$ Verallgemeinerung des Satzes von Pythagoras

Beweis durch Ausmultiplizieren von

$$|v|^2 = \langle c_1 u_1 + \cdots + c_n u_n, c_1 u_1 + \cdots + c_n u_n \rangle$$

unter Berücksichtigung von

$$\langle c_j u_j, c_k u_k \rangle = 0, \quad j \neq k$$

Beispiel

Darstellung des Vektors $v = \begin{pmatrix} 8 & 4 \end{pmatrix}^t$ bezüglich der orthogonalen Basis

$$u_1 = \begin{pmatrix} 1 & 3 \end{pmatrix}^t, \quad u_2 = \begin{pmatrix} -6 & 2 \end{pmatrix}^t$$

(i) Koeffizienten und Basisdarstellung:

$$\begin{aligned} c_1 &= \frac{\langle u_1, v\rangle}{\langle u_1, u_1\rangle} = \frac{1\cdot 8 + 3\cdot 4}{1^2+3^2} = \frac{20}{10} = 2 \\ c_2 &= \frac{\langle u_2, v\rangle}{\langle u_2, u_2\rangle} = \frac{-6\cdot 8 + 2\cdot 4}{6^2+2^2} = \frac{-40}{40} = -1 \end{aligned}$$

$\rightsquigarrow$ Linearkombination $v = c_1 u_1 + c_2 u_2$

$$\begin{pmatrix} 8 \\ 4 \end{pmatrix} = 2\begin{pmatrix} 1 \\ 3 \end{pmatrix} - \begin{pmatrix} -6 \\ 2 \end{pmatrix}$$

(iii) Quadratsumme der Koeffizienten:
Überprüfung der Identität $c_1^2|u_1|^2 + c_2^2|u_2|^2 = |v|^2$

$$2^2\cdot 10 + (-1)^2\cdot 40 = 80 = 8^2 + 4^2 = 80 \quad \checkmark$$

Beispiel

Darstellung des Vektors $v = \begin{pmatrix} -1 & 1 & 3 \end{pmatrix}^t$ bezüglich der orthonormalen Basis

$$u_1 = \frac{1}{9}\begin{pmatrix} 1 & 4 & 8 \end{pmatrix}^t, \quad u_2 = \frac{1}{9}\begin{pmatrix} 4 & 7 & -4 \end{pmatrix}^t, \quad u_3 = \frac{1}{9}\begin{pmatrix} 8 & -4 & 1 \end{pmatrix}^t$$

$|u_1| = |u_2| = |u_3| = 1 \quad \rightsquigarrow \quad$ Koeffizienten $c_k = \langle u_k, v\rangle$, d.h.

$$\begin{aligned} v &= \langle u_1, v\rangle u_1 + \langle u_2, v\rangle u_2 + \langle u_1, v\rangle u_3 \\ &= \frac{-1\cdot 1 + 4\cdot 1 + 8\cdot 3}{9} u_1 + \frac{-4+7-12}{9} u_2 + \frac{-9}{9} u_2 \\ &= 3\begin{pmatrix} 1/9 \\ 4/9 \\ 8/9 \end{pmatrix} - \begin{pmatrix} 4/9 \\ 7/9 \\ -4/9 \end{pmatrix} - \begin{pmatrix} 8/9 \\ -4/9 \\ 1/9 \end{pmatrix} \end{aligned}$$

Kontrolle der Quadratsumme der Koeffizienten:

$$|3|^2 + |-1|^2 + |-1|^2 \stackrel{!}{=} |v|^2 = (-1)^2 + 1^2 + 3^2 \quad \checkmark$$

Beispiel

Darstellung $w = su + tv$ des Vektors $w = \begin{pmatrix} 2+3\mathrm{i} & 1 \end{pmatrix}^{\mathrm{t}}$ bezüglich der orthogonalen Basis

$$u = \begin{pmatrix} 1 & \mathrm{i} \end{pmatrix}, \quad v = \begin{pmatrix} \mathrm{i} & 1 \end{pmatrix}$$

Überprüfung der Orthogonalität der Basisvektoren

$$\langle u, v \rangle = \overline{1} \cdot \mathrm{i} + \overline{\mathrm{i}} \cdot 1 = 1 \cdot \mathrm{i} + (-\mathrm{i}) \cdot 1 = 0 \quad \checkmark$$

Norm der Basisvektoren

$$|u|^2 = \langle u, u \rangle = \overline{u}_1 u_1 + \overline{u}_2 u_2 = 1 \cdot 1 + (-\mathrm{i}) \cdot \mathrm{i} = 2, \quad |v|^2 = 2$$

Koeffizienten

$$\begin{aligned} s &= \frac{\langle u, w \rangle}{\langle u, u \rangle} = \frac{\overline{1} \cdot (2+3\mathrm{i}) + \overline{\mathrm{i}} \cdot 1}{2} \\ &= \frac{1 \cdot (2+3\mathrm{i}) + (-\mathrm{i}) \cdot 1}{2} = \frac{2+3\mathrm{i}-\mathrm{i}}{2} = 1+\mathrm{i} \\ t &= \frac{(-\mathrm{i}) \cdot (2+3\mathrm{i}) + 1 \cdot 1}{2} = 2-\mathrm{i} \end{aligned}$$

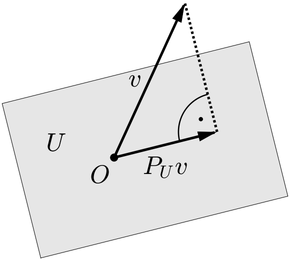
$\rightsquigarrow$ Basisdarstellung

$$\begin{pmatrix} 2+3\mathrm{i} \\ 1 \end{pmatrix} = (1+\mathrm{i}) \begin{pmatrix} 1 \\ \mathrm{i} \end{pmatrix} + (2-\mathrm{i}) \begin{pmatrix} \mathrm{i} \\ 1 \end{pmatrix}$$

Orthogonale Projektion

Die orthogonale Projektion

$$V \ni v \mapsto P_U(v) \in U$$



auf einen Unterraum U eines Vektorraums V ist durch die Orthogonalitätsbedingung

$$\langle u, v - P_U(v) \rangle = 0, \quad \forall u \in U$$

charakterisiert.

Ist $\{u_1, \ldots, u_m\}$ eine orthogonale Basis von U, so besitzt P_U die Darstellung

$$P_U(v) = \sum_{k=1}^{m} \frac{\langle u_k, v \rangle}{\langle u_k, u_k \rangle} u_k \, .$$

Insbesondere gilt für $V = \mathbb{R}^n$

$$P_U v = \sum_{k=1}^{m} |u_k|^{-2}\, u_k\, (u_k^{\mathrm{t}} v)$$

mit der Projektionsmatrix $\sum_k |u_k|^{-2}\, u_k u_k^{\mathrm{t}}$. Für $v = \mathbb{C}^n$ ist der transponierte Vektor u^{t} durch den adjungierten Vektor u^* (Transposition und komplexe Konjugation) zu ersetzen.

Beispiel

Projektion $P_H w$ von $w = (-7, 2, 1, -6)^{\mathrm{t}}$ auf den von den orthogonalen Basisvektoren

$$u = (-2, 1, 2, 0)^{\mathrm{t}},\ v = (0, 2, -1, 2)^{\mathrm{t}}$$

aufgespannten Unterraum von $\mathbb{R}^4$

$u \perp v \quad \implies$

$$P_U w = \frac{u^{\mathrm{t}} w}{|u|^2} u + \frac{v^{\mathrm{t}} w}{|v|^2} v$$

Koeffizient von u

$$\frac{(-2, 1, 2, 0)^{\mathrm{t}},\ (-7, 2, 1, -6)^{\mathrm{t}}}{|(-2, 1, 2, 0)^{\mathrm{t}}|^2} = \frac{14 + 2 + 2 + 0}{4 + 1 + 4 + 0} = 2$$

Koeffizient von v

$$\frac{(0, 2, -1, 2)^{\mathrm{t}},\ (-7, 2, 1, -6)^{\mathrm{t}}}{|(0, 2, -1, 2)^{\mathrm{t}}|^2} = \frac{0 + 4 - 1 - 12}{0 + 4 + 1 + 4} = -1$$

$\rightsquigarrow$ Projektion

$$P_U w = 2 \cdot \begin{pmatrix} -2 \\ 1 \\ 2 \\ 0 \end{pmatrix} + (-1) \cdot \begin{pmatrix} 0 \\ 2 \\ -1 \\ 2 \end{pmatrix} = \begin{pmatrix} -4 \\ 0 \\ 5 \\ -2 \end{pmatrix}$$

Verfahren von Gram-Schmidt

Aus einer Basis $\{b_1, \dots, b_n\}$ kann wie folgt eine orthogonale Basis $\{u_1, \dots, u_n\}$ konstruiert werden. Man definiert sukzessive

$$u_j = b_j - \sum_{k<j} \frac{\langle u_k, b_j \rangle}{\langle u_k, u_k \rangle} u_k,$$

für $j = 1, \dots, n$. Dabei ist die Summe die orthogonale Projektion von b_j auf $\operatorname{span}(u_1, \dots, u_{j-1})$.

Die Rekursion vereinfacht sich, wenn man die Basisvektoren nach jedem Schritt normiert:

$$u_j \leftarrow \frac{u_j}{|u_j|}.$$

In diesem Fall ist $\langle u_k, u_k \rangle = 1$.

Beweis

Zeige induktiv: $u_1, \ldots, u_\ell$ bilden eine orthogonale Basis für span $(b_1, \ldots, b_\ell)$.

- $\ell = 1: \quad u_1 = b_1 \quad \checkmark$
- Induktionsschritt $(\ell \to \ell + 1)$:
  für $j \leq \ell$

$$\begin{aligned}\langle u_j, u_{\ell+1}\rangle &= \left\langle u_j, b_{\ell+1} - \sum_{k\leq\ell} \frac{\langle u_k, b_{\ell+1}\rangle}{\langle u_k, u_k\rangle} u_k \right\rangle \\ &= \langle u_j, b_{\ell+1}\rangle - \sum_{k\leq\ell} \frac{\langle b_{\ell+1}, u_k\rangle}{\langle u_k, u_k\rangle} \underbrace{\langle u_j, u_k\rangle}_{\delta_{k,j}|u_j|^2} = 0\end{aligned}$$

$\Longrightarrow \quad u_1, \ldots, u_{\ell+1}$ orthogonal, da $u_j \perp u_k$ für $j, k \leq \ell$ nach Induktionsvoraussetzung

$$b_{\ell+1} - u_{\ell+1} \in \operatorname{span}(u_1, \ldots, u_\ell) = \operatorname{span}(b_1, \ldots, b_\ell)$$

$\Longrightarrow \quad$ Beide Basen spannen denselben Unterraum auf.

Beispiel

Orthonormale Basis $\{u_1, u_2, u_3\}$ für den von den Vektoren

$$b_1 = (1,1,1,1)^{\mathrm{t}},\ b_2 = (1,2,1,0)^{\mathrm{t}},\ b_3 = (2,2,0,0)^{\mathrm{t}}$$

aufgespannten Unterraum von $\mathbb{R}^4$

$|b_1| = \sqrt{1+1+1+1} = 2 \quad \rightsquigarrow \quad u_1 = (1,1,1,1)^{\mathrm{t}}/2$

zwei Schritte des Gram-Schmidt-Verfahrens

$$u_2 \quad \| \quad b_2 - \langle u_1, b_2\rangle u_1 = \begin{pmatrix} 1 \\ 2 \\ 1 \\ 0 \end{pmatrix} - \left\langle \frac{1}{2}\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 1 \\ 0 \end{pmatrix} \right\rangle \frac{1}{2}\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

$$= \begin{pmatrix} 1 \\ 2 \\ 1 \\ 0 \end{pmatrix} - 2\cdot\frac{1}{2}\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ -1 \end{pmatrix}$$

Normierung $\quad \rightsquigarrow \quad u_2 = (0,1,0,-1)^{\mathrm{t}}/\sqrt{2}$

$$u_3 \parallel b_3 - \langle u_1, b_3\rangle u_1 - \langle u_2, b_3\rangle u_2 =$$

$$\begin{pmatrix} 2 \\ 2 \\ 0 \\ 0 \end{pmatrix} - 2 \cdot \frac{1}{2} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} - \sqrt{2} \cdot \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ 0 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix}$$

Normierung $\quad \rightsquigarrow \quad u_3 = (1, 0, -1, 0)^{\mathrm{t}}/\sqrt{2}$

Kontrolle

$$\begin{aligned} \underbrace{(1,2,1,0)^{\mathrm{t}}}_{b_2} &= 2u_1 + \sqrt{2}u_2 = (1,1,1,1)^{\mathrm{t}} + (0,1,0,-1)^{\mathrm{t}} \\ \underbrace{(2,2,0,0)^{\mathrm{t}}}_{b_3} &= 2u_1 + \sqrt{2}u_2 + \sqrt{2}u_3 \\ &= (1,1,1,1)^{\mathrm{t}} + (0,1,0,-1)^{\mathrm{t}} + (1,0,-1,0)^{\mathrm{t}} \end{aligned}$$

$\Longrightarrow$ Übereinstimmung der aufgespannten Unterräume

Beispiel

Konstruktion einer bzgl. des Skalarproduktes

$$\langle f, g\rangle = \int_{-1}^{1} f(x)g(x)\,\mathrm{d}x$$

orthogonalen Folge von Polynomen $p_0, p_1, \ldots$ aus den Monomen $q_j(x) = x^j$

Verfahren von Gram-Schmidt $\rightsquigarrow$

$$\begin{aligned}
p_0 &= q_0 = 1 \\
p_1 &= q_1 - \frac{\langle p_0, q_1\rangle}{\langle p_0, p_0\rangle} p_0 = q_1 - \frac{\int_{-1}^{1} 1\cdot x\,\mathrm{d}x}{\int_{-1}^{1} 1\cdot 1\,\mathrm{d}x}\cdot 1 = x - 0 \\
p_2 &= q_2 - \frac{\int_{-1}^{1} x\cdot x^2\,\mathrm{d}x}{\int_{-1}^{1} x\cdot x\,\mathrm{d}x}\cdot x - \frac{\int_{-1}^{1} 1\cdot x^2\,\mathrm{d}x}{\int_{-1}^{1} 1\cdot 1\,\mathrm{d}x}\cdot 1 = x^2 - 0\cdot x - \frac{1}{3}
\end{aligned}$$

allgemeine Formel: $p_{n+1} = q_{n+1} - \sum_{j=0}^{n} \frac{\langle p_j, q_{n+1}\rangle}{\langle p_j, p_j\rangle}\, p_j$

Vereinfachung durch Verwendung von $\tilde{q}_{n+1}(x) = xp_n(x)$ anstelle von $q_{n+1}(x) = x^{n+1}$
Legitim, da $xp_n(x) = x^{n+1} + \sum_{j \le n} c_j x^j \quad \Longrightarrow$

$$\operatorname{span}\{p_1, \ldots, p_n, q_{n+1}\} = \operatorname{span}\{p_1, \ldots, p_n, \tilde{q}_{n+1}\}$$

$\rightsquigarrow$ alle Summanden bis auf $j = n - 1$ Null
$j = n$: aufgrund der Antisymmetrie des Integranden

$$\langle p_n, \tilde{q}_{n+1} \rangle = \int_{-1}^{1} p_n(x)\, x p_n(x)\, \mathrm{d}x = 0$$

$j < n - 1$: $[x\, p_j(x)]$ Linearkombination von $p_k(x)$, $k < n \quad \Longrightarrow$

$$\langle \tilde{p}_j, q_{n+1} \rangle = \int_{-1}^{1} p_n(x)\, [x\, p_j(x)]\, \mathrm{d}x = 0\,,$$

da $p_n \perp p_k$

andere Normierung $\rightsquigarrow$ Legendre-Polynome

$$p_n(x) = \alpha_n x^n + \dots, \quad \alpha_n = \frac{(2n)!}{2^n (n!)^2}$$

Gram-Schmidt-Prozess $\rightsquigarrow$ 3-Term-Rekursion

$$(n+1)p_{n+1}(x) = (2n+1)xp_n(x) - np_{n-1}(x)$$

mit $p_0(x) = 1$ und $p_1(x) = x$

Lineare Abbildung

Eine lineare Abbildung $L : V \longmapsto W$ zwischen K-Vektorräumen V und W besitzt folgende Eigenschaften.

- Additivität: $L(u+v) = L(u) + L(v)$
- Homogenität: $L(sv) = sL(v)$

Dabei sind u, v beliebige Elemente von V und s beliebige Skalare in K. Insbesondere gilt $L(0_V) = 0_W$ und $L(-v) = -L(v)$.

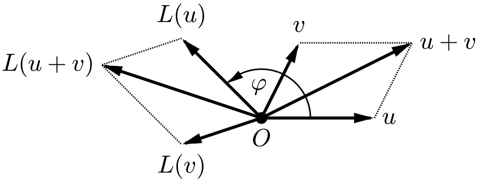
Aufgrund der beiden definierenden Eigenschaften ist eine lineare Abbildung L eindeutig durch die Bilder $L(b_k)$ einer Basis $\{b_1, b_2, \ldots\}$ von V bestimmt.

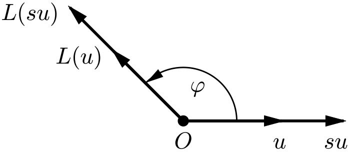
Oft werden bei linearen Abbildungen die Klammern weggelassen, d.h. man schreibt Lu anstelle von $L(u)$.

Beispiel

Elementare lineare Abbildungen der Ebene

(i) Drehung um einen Winkel φ:

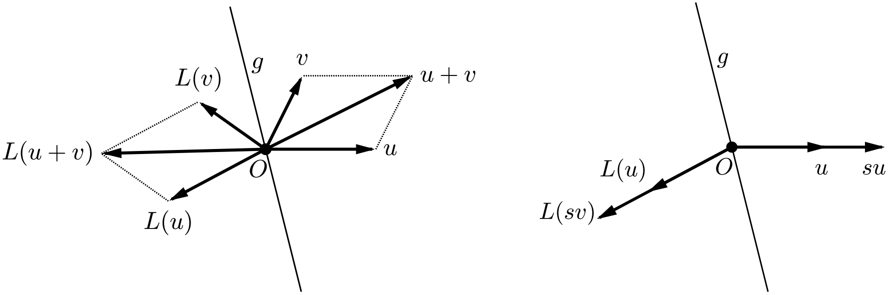




Vertauschbarkeit von Addition und skalarer Multiplikation ✓

- gedrehte Vektorsumme $L(u+v)$ = Summe der gedrehten Vektoren $L(u)$, $L(v)$
- gedrehtes Vektorvielfaches $L(su)$ = Vielfaches des gedrehten Vektors $sL(u)$

(ii) Spiegelung:



Vertauschbarkeit von Addition und skalarer Multiplikation ✓

$$L(u+v) = L(u) + L(v), \quad L(su) = sL(u)$$

(iii) Verschiebung:
nicht linear, denn beispielsweise für

$$T : \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \begin{pmatrix} x_1 + 1 \\ x_2 \end{pmatrix}$$

und

$$u = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad v = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad s = 2$$

gilt

$$T(u+v) = T\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \neq (3,1) = T(u) + T(v) = \begin{pmatrix} 2 \\ 0 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$T(su) = T\begin{pmatrix} 2 \\ 0 \end{pmatrix} = (3,0) \neq (4,0) = sT(u) = s\begin{pmatrix} 2 \\ 0 \end{pmatrix}$$

d.h. T ist weder additiv noch homogen

Beispiel

Abbildungen reeller Funktionen

Abbildung	additiv	homogen
$f \mapsto f'$	X	X
$f \mapsto \lvert f \rvert$	–	–
$f \mapsto \int_0^1 f$	X	X
$f \mapsto \max f$	–	–
$f \mapsto f(0)$	X	X
$f \mapsto (\max f + \min f)/2$	–	X

additive aber nicht homogene Abbildung:

$$T\,:\; f \mapsto \operatorname{Re} f, \quad f \text{ komplexwertig}$$

nicht homogen, da

$$T(\mathrm{i}f) \neq \mathrm{i}\,T(f)$$

Komposition linearer Abbildungen

Für zwei lineare Abbildungen

$$P:\ U \to V, \quad Q:\ V \to W$$

ist die Komposition oder Hintereinanderausführung

$$Q \circ P:\ U \ni u \mapsto Q(P(u)) \in W$$

ebenfalls linear.

Matrix

Eine $m \times n$-Matrix besteht aus $m \cdot n$ Elementen, die in m Zeilen und n Spalten angeordnet sind:

$$A = (a_{j,k}) = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ \vdots & \vdots & & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{pmatrix} .$$

Man bezeichnet

$$u_j^{\mathrm{t}} = (a_{j,1}, \ldots, a_{j,n}), \quad v_k = \begin{pmatrix} a_{1,k} \\ \vdots \\ a_{m,k} \end{pmatrix}$$

als j-ten Zeilen- bzw. k-ten Spaltenvektor und schreibt ebenfalls

$$A = \begin{pmatrix} u_1^{\mathrm{t}} \\ \vdots \\ u_m^{\mathrm{t}} \end{pmatrix} = (v_1, \ldots, v_n) .$$

Speziell ist eine $1 \times n$-Matrix ein Zeilen- und eine $m \times 1$-Matrix ein Spaltenvektor.

Für Elemente in $\mathbb{Q}$, $\mathbb{R}$ oder $\mathbb{C}$ bilden die $m \times n$-Matrizen einen Vektorraum mit elementweise definierter Addition und skalarer Multiplikation:

$$C = A + B \iff c_{j,k} = a_{j,k} + b_{j,k}, \quad C = sA \iff c_{j,k} = sa_{j,k}\,.$$

Beispiel

Verschiedene Matrixdimensionen

- Zeilen- und Spaltenvektor

$$(11, 12, 13), \quad \begin{pmatrix} 11 \\ 21 \end{pmatrix}$$

- Komplexe 2×2-Matrix

$$\begin{pmatrix} 1+\mathrm{i} & 1+2\mathrm{i} \\ 2+\mathrm{i} & 2+2\mathrm{i} \end{pmatrix}$$

- Rationale 2×3-Matrix

$$\begin{pmatrix} 1/1 & 1/2 & 1/3 \\ 2/1 & 2/2 & 2/3 \end{pmatrix}$$

Matrix einer linearen Abbildung

Eine lineare Abbildung $L: V \to W$ zwischen zwei K-Vektorräumen mit den Basen $E = \{e_1, \ldots, e_n\}$ und $F = \{f_1, \ldots, f_m\}$ ist durch die Bilder der Basisvektoren

$$L(e_k) = a_{1,k} f_1 + \cdots + a_{m,k} f_m, \quad k = 1, \ldots, n\,,$$

eindeutig bestimmt. Sie besitzt die Matrixdarstellung

$$w = L(v) \quad \Longleftrightarrow \quad w_j = \sum_{k=1}^{n} a_{j,k} v_k, \quad j = 1, \ldots, m,$$

wobei v_k und w_j die Koordinaten von v und $w = L(v)$ bzgl. der Basen E und F bezeichnen. In der k-ten Spalte der Matrix A stehen also die Koordinaten von $L(e_k)$ bezüglich der Basis F.

Beweis

L ist aufgrund der Bedingungen für Linearität durch die Bilder einer Basis eindeutig bestimmt:

$$w = L(v) = L\left(\sum_k v_k e_k\right) = \sum_k v_k L(e_k) = \sum_k \sum_j v_k a_{j,k} f_j$$

mit den Basiskoeffizienten v_k von v.
Basisdarstellung von w

$$w = \sum_j w_j f_j$$

Vergleich der Koordinaten der Basisvektoren $\quad\rightsquigarrow\quad$ Matrixdarstellung

$$w_j = \sum_{k=1}^{n} a_{j,k} v_k$$

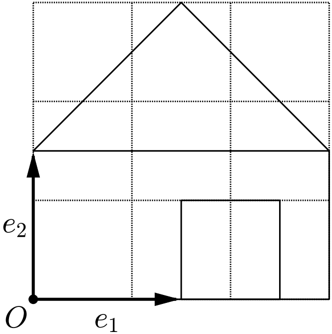
Beispiel

Lineare Abbildungen

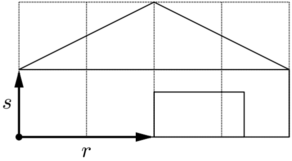
$$L:\ \mathbb{R}^2 \to \mathbb{R}^2$$

der Ebene festgelegt durch die Bilder der Einheitsvektoren

$$e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$



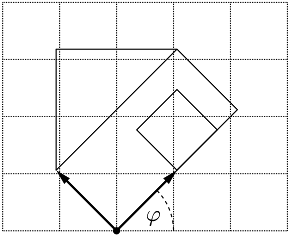
(i) Skalierung mit Faktoren r und s:



$$L(e_1) = \begin{pmatrix} r \\ 0 \end{pmatrix},\ L(e_2) = \begin{pmatrix} 0 \\ s \end{pmatrix}$$

$$\Longrightarrow \quad A = \begin{pmatrix} r & 0 \\ 0 & s \end{pmatrix}$$

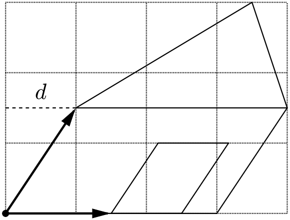
(ii) Drehung um einen Winkel φ:



$$L(e_1) = \begin{pmatrix} \cos\varphi \\ \sin\varphi \end{pmatrix},\ L(e_2) = \begin{pmatrix} -\sin\varphi \\ \cos\varphi \end{pmatrix}$$

$$\Longrightarrow \quad A = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}$$

(iii) Scherung um d in horizontaler Richtung:



$$L(e_1) = \begin{pmatrix} 1 \\ 0 \end{pmatrix},\ L(e_2) = \begin{pmatrix} d \\ 1 \end{pmatrix}$$

$$\Longrightarrow \quad A = \begin{pmatrix} 1 & d \\ 0 & 1 \end{pmatrix}$$

Beispiel

Interpolationsmatrizen

(i) Auswertung einer linearen Funktion p an den Punkten $x = 0, 1$:

$$L:\ p \mapsto \begin{pmatrix} p(0) \\ p(1) \end{pmatrix}, \quad p(x) = a_0 + a_1 x$$

Matrix bzgl. der Monombasis $p_1 : x \mapsto 1$, $p_2 : x \mapsto x$

$$\big(L(p_1), L(p_2)\big) = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$$

Matrix bzgl. der Basis $p_1 : x \mapsto 1 - x$, $p_2 : x \mapsto x$

$$\begin{pmatrix} p_1(0) & p_2(0) \\ p_1(1) & p_2(1) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

(ii) Auswertung eines Polynoms vom Grad $\le n$ an m Stützstellen $x = x_1, \ldots, x_m$:
Monombasis, $q_k : x \mapsto x^k$, $k = 0, \ldots, n$ $\quad \rightsquigarrow \quad$ Vandermonde-Matrix

$$V = \begin{pmatrix} x_1^0 & x_1^1 & \cdots & x_1^n \\ x_2^0 & x_2^1 & \cdots & x_2^n \\ \vdots & \vdots & \ddots & \vdots \\ x_m^0 & x_m^1 & \cdots & x_m^n \end{pmatrix}$$

Spalte k: Auswertung des Monoms q_k an den Punkten $x_1, \ldots, x_m$
Zeile j: Auswertung der Monome q_k, $k = 0, \ldots, n$, am Punkt x_j

$p(x) = a_0 + a_1 x + \cdots + a_n x^n \quad \Longrightarrow$

$$p(x_j) = \sum_{k=0}^{n} v_{j,k} a_k$$

Affine Abbildung

Eine affine Abbildung $f : \mathbb{R}^n \to \mathbb{R}^m$ setzt sich aus einer linearen Abbildung und einer Verschiebung zusammen:

$$f(x) = Ax + v$$

mit einer $m \times n$-Matrix A und einem m-Vektor v.
Der Verschiebungsvektor v ist das Bild des Nullvektors und für die Spalten von A gilt

$$\begin{pmatrix} a_{1,k} \\ \vdots \\ a_{m,k} \end{pmatrix} = f(e_k) - v$$

mit $e_k \in \mathbb{R}^n$ dem k-ten Einheitsvektor.

Koordinatentransformation bei Basiswechsel

Für zwei Basen $\{e_1, \ldots, e_n\}$ und $\{f_1, \ldots, f_n\}$ eines Vektorraums V kann die Umrechnung der Koordinaten x und y eines Elements $v \in V$,

$$v = \sum_k x_k e_k \quad \to \quad v = \sum_j y_j f_j$$

durch eine quadratische Matrix A beschrieben werden:

$$y = Ax \quad \Longleftrightarrow \quad y_j = \sum_{k=1}^{n} a_{j,k} x_k, \quad j = 1, \ldots, n\,.$$

Die k-te Spalte der Matrix A enthält die Koeffizienten von e_k bezüglich der Basis $\{f_1, \ldots, f_n\}$, d.h.

$$e_k = \sum_{j=1}^{n} a_{j,k} f_j\,.$$

Für $V = K^n$ ($K = \mathbb{R}^n$, $K = \mathbb{C}^n$, ...) können die Basisvektoren jeweils spaltenweise in einer Matrix zusammengefasst werden:

$$E = (e_1, \ldots, e_n), \quad F = (f_1, \ldots, f_n)\,.$$

Die Transformationsmatrix A lässt sich in diesem Fall als Matrixprodukt schreiben

$$E = FA \quad \Longleftrightarrow \quad A = EF^{-1}\,.$$

Beweis

Darstellung von v bezüglich der Basen E und F:

$$\sum_k x_k e_k = v = \sum_j y_j e_j$$

Darstellung von e_k bezüglich der Basis F $\quad\Longrightarrow$

$$\begin{aligned} v &= \sum_k x_k e_k = \sum_k x_k \left(\sum_j a_{j,k} f_j \right) \\ &= \sum_j \left(\sum_k a_{j,k} x_k \right) f_j = \sum_j y_j f_j \end{aligned}$$

Koordinatenvergleich $\quad\Longrightarrow\quad y = Ax$

Beispiel

Koordinatentransformation für die Basen

$$\{e_1, e_2\} = \left\{ \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 3 \end{pmatrix} \right\}, \quad \{f_1, f_2\} = \left\{ \begin{pmatrix} 3 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \end{pmatrix} \right\}$$

Darstellung von e_k als Linearkombination von f_j: $e_k = a_{1,k} f_1 + a_{2,k} f_2$

$$\begin{pmatrix} 0 \\ 1 \end{pmatrix} = 1 \cdot \begin{pmatrix} 3 \\ 1 \end{pmatrix} - \frac{3}{2} \cdot \begin{pmatrix} 2 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 2 \\ 3 \end{pmatrix} = 3 \cdot \begin{pmatrix} 3 \\ 1 \end{pmatrix} - \frac{7}{2} \cdot \begin{pmatrix} 2 \\ 0 \end{pmatrix}$$

$\rightsquigarrow$ Transformationsmatrix

$$A = \begin{pmatrix} 1 & 3 \\ -3/2 & -7/2 \end{pmatrix}$$

alternative Berechnung mit den Matrizen $E = (e_1, e_2)$, $F = (f_1, f_2)$

$$A = EF^{-1} = \begin{pmatrix} 0 & 2 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} 3 & 2 \\ 1 & 0 \end{pmatrix}^{-1} = \begin{pmatrix} 0 & 2 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1/2 & -3/2 \end{pmatrix} \quad \checkmark$$

Koordinaten x von $v = (2,5)^{\mathrm{t}}$ bezüglich $\{e_1, e_2\}$

$$\begin{pmatrix} 2 \\ 5 \end{pmatrix} = 2 \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} + 1 \cdot \begin{pmatrix} 2 \\ 3 \end{pmatrix} \quad \Longrightarrow \quad x = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

Umrechnung von x in Koordinaten bezüglich $\{f_1, f_2\}$ $\quad \rightsquigarrow$

$$y = \begin{pmatrix} 1 & 3 \\ -3/2 & -7/2 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 5 \\ -13/2 \end{pmatrix}$$

Probe

$$\begin{pmatrix} 2 \\ 5 \end{pmatrix} v \overset{!}{=} y_1 f_1 + y_2 f_2 = 5 \begin{pmatrix} 3 \\ 1 \end{pmatrix} - \frac{13}{2} \begin{pmatrix} 2 \\ 0 \end{pmatrix} \quad \checkmark$$

Bild und Kern

Für eine lineare Abbildung $L : V \to W$ bezeichnet man mit

$$\operatorname{Kern} L = \{v \in V : \; L(v) = 0_W\} \; \subseteq V$$

den Kern und mit

$$\operatorname{Bild} L = \{w \in W : \; \exists v \in V \text{ mit } L(v) = w\} \; \subseteq W$$

das Bild von L.

Beide Mengen sind Unterräume und

$$\dim V = \dim \operatorname{Kern} L + \dim \operatorname{Bild} L \,,$$

falls $\dim V < \infty$.

Beweis

(i) Kern L ist Unterraum von V:

zu zeigen: Abgeschlossenheit bezüglich Addition und skalarer Multiplikation

Für $u, v \in$ Kern L, $s \in K$ folgt aus der Linearität von L

$$L(s\,u) = s\,L(u) = s\,0 = 0$$

und

$$L(u + v) = L(u) + L(v) = 0 + 0 = 0\,,$$

d.h. $su,\ u + v \in \mathrm{Kern}(L)$

(ii) Bild L ist Unterraum von W:

Für $u, v \in$ Bild L mit $u = L(x)$, $v = L(y)$ und $s \in K$ folgt analog

$$s\,u = s\,L(x) = L(s\,x)$$

und

$$u + v = L(x) + L(y) = L(x + y)\,,$$

d.h. die Existenz entsprechender Urbilder

$\Longrightarrow \quad su,\ u + v \in$ Bild L

(iii) Dimensionsformel:
Wähle eine Basis $E' = \{e_1, \ldots, e_m\}$ für Kern L und ergänze diese durch $\{e_{m+1}, \ldots, e_n\}$ zu einer Basis E von V.

$$L\left(\sum_{k=1}^{n} v_k e_k\right) = \sum_{k=m+1}^{n} v_k L(e_k)$$

$\implies$ Bild $L = \operatorname{span} F,\ F = \{L(e_{m+1}), \ldots, L(e_n)\}$
F linear unabhängig, also eine Basis von Bild L, denn

$$0 = \sum_{k=m+1}^{n} v_k L(e_k) = L\left(\sum_{k=m+1}^{n} v_k e_k\right) \quad \implies \quad \sum_{k=m+1}^{n} v_k e_k \in \text{Kern } L$$

und somit $v_k = 0$
Vergleich der Anzahlen der Basisvektoren $\implies$ Dimensionsformel

Beispiel

Kern und Bild der linearen Abbildung

$$\mathbb{R}^3 \ni x \mapsto Lx = \begin{pmatrix} x_1 - x_2 \\ x_2 - x_3 \\ x_3 - x_1 \end{pmatrix} \in \mathbb{R}^3$$

(i) Kern:

$$\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = Lx = \begin{pmatrix} x_1 - x_2 \\ x_2 - x_3 \\ x_3 - x_1 \end{pmatrix} \quad \Longrightarrow \quad x_1 = x_2 = x_3\,,$$

d.h. Kern $L = \{(t,\ t,\ t)^{\mathrm{t}} :\ t \in \mathbb{R}\} = \mathrm{span}((1,\ 1,\ 1)^{\mathrm{t}})$

(ii) Bild:
Bild L wird durch die Bilder der Einheitsvektoren $e_1 = (1,\, 0,\, 0)^{\mathrm{t}}$, $e_2 = (0,\, 1,\, 0)^{\mathrm{t}}$, $e_3 = (0,\, 0,\, 1)^{\mathrm{t}}$ aufgespannt.

$$\operatorname{Bild} L = \operatorname{span}(Le_1, Le_2, Le_3) = \operatorname{span}\left(\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}\right)$$

$Le_3 = -Le_1 - Le_2$ und $Le_1 \not\parallel Le_2$, d.h. Le_1, Le_2 linear unabhängig $\quad\Longrightarrow$

$$\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \quad \text{Basis für Bild } L$$

(iii) Dimensionsformel:

$$3 = \dim \mathbb{R}^3 \stackrel{!}{=} \dim \operatorname{Kern} L + \dim \operatorname{Bild} L = 1 + 2 \quad \checkmark$$

Beispiel

k-te Ableitung D^k auf dem Raum der Polynome $\mathbb{P}_n$ vom Grad $\leq n$

(i) $k = 1$ und $n = 2$: $\{x \mapsto 1, x \mapsto x, x \mapsto x^2\}$ Basis für $\mathbb{P}_2$

$$p(x) = a_0 + a_1 x + a_2 x^2 \quad \implies \quad (Dp)(x) = a_1 + 2a_2 x \in \mathbb{P}_1$$

Die Ableitung D annuliert Konstanten, einen eindimensionalen Unterraum.

$\implies \quad \dim \mathbb{P}_2 = 3, \quad \dim \operatorname{Kern} D = 1, \quad \dim \operatorname{Bild} = 2$

(ii) Allgemeiner Fall:

$$p(x) = \sum_{\ell=0}^{n} a_\ell x^\ell \quad \implies \quad (D^k p)(x) = \sum_{\ell=k}^{n} \frac{\ell!}{(\ell - k)!} a_\ell x^{\ell - k} \in \mathbb{P}_{n-k}$$

D^k annullient Polynome vom Grad $< k$

$\rightsquigarrow$ Dimensionsformel

$$n + 1 = \underbrace{((n-k)+1)}_{\dim \operatorname{Bild}} + \underbrace{k}_{\dim \operatorname{Kern}} = \dim \mathbb{P}_{n-k} + \dim \mathbb{P}_{k-1} \quad \checkmark$$

Inverse Abbildung

Eine lineare Abbildung $L : V \to W$ ist genau dann injektiv, wenn Kern $L = 0_V$, d.h. nur das Nullelement von V wird auf das Nullelement von W abgebildet:

$$Lv = 0_W \quad \Longrightarrow \quad v = 0_V \, .$$

In diesem Fall kann durch

$$w \mapsto v, \quad w = L(v) \, ,$$

eine inverse Abbildung L^{-1} : Bild $L \to V$ definiert werden, die ebenfalls linear ist. Insbesondere gilt

$$(L^{-1} \circ L)v = v, \quad (L \circ L^{-1})w = w$$

für alle $v \in V$ und $w \in$ Bild L.

Beweis

(i) Injektivität von L:

- L injektiv $\implies$ 0_V ist einziges Urbild von 0_W, d.h. $\text{Kern}\, L = 0_V$
- $\text{Kern}\, L = 0_V \wedge L(v_1) = L(v_2)$
  $\implies 0_W = L(v_1 - v_2) \implies v_1 - v_2 \in \text{Kern}\, L \implies v_1 - v_2 = 0_V$
  $\implies$ L injektiv

(ii) Linearität derUmkehrabbildung:

$$\begin{aligned} L^{-1}(L(v_1) + L(v_2)) &= L^{-1}(L(v_1 + v_2)) \\ &= v_1 + v_2 = L^{-1}(L(v_1)) + L^{-1}(L(v_2)) \\ L^{-1}(sL(v_1)) &= L^{-1}(L(sv_1)) \\ &= sv_1 = sL^{-1}(L(v_1)) \end{aligned}$$

Beispiel

Konstruktion der zu

$$L : \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto y = \begin{pmatrix} 2x_1 + x_2 \\ x_1 - x_2 \\ x_1 - 2x_2 \end{pmatrix}$$

inversen Abbildung

(i) Überprüfung der Injektivität:
zu zeigen: Kern $L = (0,0)^{\mathrm{t}}$, d.h. $y = Lx = (0,0,0)^{\mathrm{t}} \implies x = (0,0)^{\mathrm{t}}$
sukzessives Betrachten der Gleichungen

$$0 = y_1 = 2x_1 + x_2, \quad 0 = y_2 = x_1 - x_2, \quad 0 = y_3 = x_1 - 2x_2$$

$$\implies x_2 = -2x_1,\ 0 = x_1 + 2x_1 \iff x_1 = 0,\ 0 = 0 - 2x_2 \iff x_2 = 0 \quad \checkmark$$

(ii) Inverse Abbildung:
Auflösen von $y = Lx$ nach x

$y_1 + y_2 = 3x_1 \quad \Longrightarrow \qquad x_1 = (y_1 + y_2)/3 \quad$ und $\quad y_2 - y_3 = x_2 \quad \rightsquigarrow$

$$L^{-1} : \, y \mapsto x = \begin{pmatrix} (y_1 + y_2)/3 \\ y_2 - y_3 \end{pmatrix}$$

(iii) Eindeutigkeit von L^{-1}:
anderes Auflösen der Gleichungen $y = Lx \quad \rightsquigarrow$

$$\tilde{L}^{-1} y = \begin{pmatrix} 2y_2 - y_3 \\ (y_1 - 2y_2)/3 \end{pmatrix}$$

kein Widerspruch, denn $L^{-1}y = \tilde{L}^{-1}y$ für $y \in \operatorname{Bild} L$, dem Definitionsbereich der inversen Abbildung
Überprüfung der Übereinstimmung für die kanonische Basis von Bild L,

$$\{(L(1,0)^{\mathrm{t}},\ (L(0,1)^{\mathrm{t}}\} = \{(2,1,1)^{\mathrm{t}},\ (1,-1,-2)^{\mathrm{t}}\}$$

$$L^{-1}\begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} (2+1)/3 \\ 1-1 \end{pmatrix} = \begin{pmatrix} 2\cdot 1 - 1 \\ (2-2\cdot 1)/3 \end{pmatrix} = \tilde{L}^{-1}\begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} \quad \checkmark$$

Matrix-Multiplikation

Das Produkt einer $\ell \times m$-Matrix A und einer $m \times n$-Matrix B ist die $\ell \times n$-Matrix

$$C = AB, \quad c_{i,k} = \sum_{j=1}^{m} a_{i,j} b_{j,k} \,,$$

d.h. zur Definition von $c_{i,k}$ werden die Produkte der Elemente aus Zeile i von A und Spalte k von B summiert:

$$\begin{pmatrix} c_{1,1} \cdots c_{1,k} \cdots c_{1,n} \\ \vdots \qquad \vdots \qquad \vdots \\ c_{i,1} \cdots \mathbf{c_{i,k}} \cdots c_{i,n} \\ \vdots \qquad \vdots \qquad \vdots \\ c_{\ell,1} \cdots c_{\ell,k} \cdots c_{\ell,n} \end{pmatrix} = \begin{pmatrix} a_{1,1} \cdots a_{1,m} \\ \vdots \qquad \vdots \\ \mathbf{a_{i,1}} \cdots \mathbf{a_{i,m}} \\ \vdots \qquad \vdots \\ c_{\ell,1} \cdots c_{\ell,m} \end{pmatrix} \begin{pmatrix} b_{1,1} \cdots \mathbf{b_{1,k}} \cdots b_{1,n} \\ \vdots \qquad \vdots \qquad \vdots \\ b_{m,1} \cdots \mathbf{b_{m,k}} \cdots b_{m,n} \end{pmatrix}$$

Dafür ist essentiell, dass die Spaltenanzahl m von A mit der Zeilenanzahl von B (ebenfalls m) übereinstimmt.

Die Matrix-Multiplikation entspricht der Komposition der linearen Abbildungen

$$P:\ u \mapsto v = Bu, \quad Q:\ v \mapsto w = Av\,,$$

d.h. $Q \circ P:\ u \mapsto ABu$.

Die Multiplikation von Matrizen ist im allgemeinen nicht kommutativ.

Ein Spezialfall der Matrix-Multiplikation ist die Multiplikation mit einem Vektor.

- Für einen Spaltenvektor b ($n = 1$) ist der Spaltenvektor $c = Ab$ eine Linearkombination der Spalten von A:

$$c_i = \sum_{j=1}^{m} a_{i,j} b_j \iff \begin{pmatrix} c_1 \\ \vdots \\ c_\ell \end{pmatrix} = b_1 \begin{pmatrix} a_{1,1} \\ \vdots \\ a_{\ell,1} \end{pmatrix} + \cdots + b_m \begin{pmatrix} a_{1,m} \\ \vdots \\ a_{\ell,m} \end{pmatrix}.$$

  Insbesondere erhält man bei Multiplikation mit dem j-ten Einheitsvektor die j-te Spalte von A.

- Entsprechend gilt für einen Zeilenvektor a ($\ell = 1$)

$$c = aB, \quad c_k = \sum_{j=1}^{m} a_j b_{j,k}$$

bzw.

$$(c_1, \ldots, c_n) = a_1(b_{1,1}, \ldots, b_{1,n}) + \cdots + a_m(b_{m,1}, \ldots, b_{m,n}),$$

und Multiplikation mit dem k-ten Einheitsvektor ergibt die k-te Zeile von B.

Beweis

Komposition der durch

$$w_i = \sum_{j=1}^{m} a_{i,j} v_j, \quad v_j = \sum_{k=1}^{n} b_{j,k} u_k$$

definierten Abbildungen $\quad \rightsquigarrow$

$$w_i = \sum_j \sum_k a_{i,j} b_{j,k} u_k = \sum_k \left[\sum_j a_{i,j} b_{j,k} \right] u_k$$

mit $[\ldots] = c_{i,k}$ den Elementen der Matrix $C = AB$

Beispiel

Matrix-Produkte verschiedener Dimensionen

$$\begin{pmatrix} 1 & 10 & 100 \\ 100 & 10 & 1 \end{pmatrix} \begin{pmatrix} 1 & 3 & 2 \\ 2 & 1 & 3 \\ 3 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 321 & 213 & 132 \\ 123 & 312 & 231 \end{pmatrix}$$

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

$$\begin{pmatrix} 3 & 4 \end{pmatrix} \begin{pmatrix} 3 \\ 4 \end{pmatrix} = 25$$

$$\begin{pmatrix} 1 & \mathrm{i} \\ \mathrm{i} & 1 \end{pmatrix} \begin{pmatrix} 1 \\ \mathrm{i} \end{pmatrix} = \begin{pmatrix} 0 \\ 2\mathrm{i} \end{pmatrix}$$

Beispiel

Strassens Multiplikationen sparende Matrix-Multiplikation
[Gaussian elimination is not optimal, Numer. Math. 13 (1969), 357–361]

Berechnung des Produkts $C = AB$ zweier 2×2-Matrizen mit 7 anstelle der üblichen 8 Multiplikationen

$$\begin{aligned}
p_1 &= (a_{1,2} - a_{2,2})(b_{2,1} + b_{2,2}) \\
p_2 &= (a_{1,1} + a_{2,2})(b_{1,1} + b_{2,2}) \\
p_3 &= (a_{1,1} - a_{2,1})(b_{1,1} + b_{1,2}) \\
p_4 &= (a_{1,1} + a_{1,2})b_{2,2} \\
p_5 &= a_{1,1}(b_{1,2} - b_{2,2}) \\
p_6 &= a_{2,2}(b_{2,1} - b_{1,1}) \\
p_7 &= (a_{2,1} + a_{2,2})b_{1,1}
\end{aligned}$$

$$\rightsquigarrow \quad \begin{aligned}
c_{1,1} &= a_{1,1}b_{1,1} + a_{1,2}b_{2,1} &&= p_1 + p_2 - p_4 + p_6 \\
c_{1,2} &= a_{1,1}b_{1,2} + a_{1,2}b_{2,2} &&= p_4 + p_5 \\
c_{2,1} &= a_{2,1}b_{1,1} + a_{2,2}b_{2,1} &&= p_6 + p_7 \\
c_{2,2} &= a_{2,1}b_{1,2} + a_{2,2}b_{2,2} &&= p_2 - p_3 + p_5 - p_7
\end{aligned}$$

Rekursive Anwendung der Konstruktion in Blockform auf Matrizen der Dimension $n = 2^k$

$\rightsquigarrow$ Reduktion der Multiplikationen von n^3 auf $O(n^{\log_2 7})$

Verbesserung: $O(n^{2.376\ldots})$

[D. Coppersmith, S. Winograd: Matrix Multiplication via arithmetic progressions, J. Symb. Comp. 9 (1990), 251–280]

Vermutung: $O(n^2)$ möglich (optimaler Exponent noch unbekannt)

Die Methode hat heute keine praktische Bedeutung mehr, da auf modernen Computern Multiplikation und Addition fast gleich schnell sind.

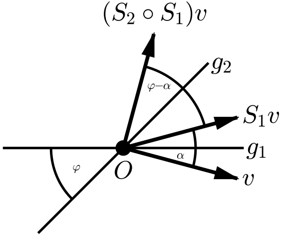
Beispiel

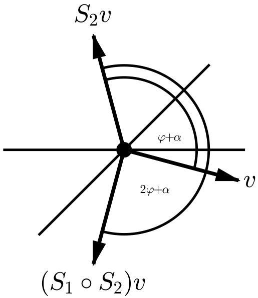
Spiegelungen S_1 und S_2 an Geraden g_1 und g_2

$$g_1 \parallel d_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad g_2 \parallel d_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \varphi = \sphericalangle(g_1, g_2) = \pi/4$$

Bei Hintereinanderausführung der Spiegelungen ist die Reihenfolge relevant: $(S_2 \circ S_1)v \neq (S_1 \circ S_2)v$

$$(S_2 \circ S_1)v : \text{ Drehung um } 2\varphi, \quad (S_1 \circ S_2)v : \text{ Drehung um } -2\varphi$$





Matrixdarstellung ($S \to A$, Spalten: Bilder der Einheitsvektoren)

$$\begin{aligned}
& S_1: \begin{pmatrix} 1 \\ 0 \end{pmatrix} \mapsto \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 1 \end{pmatrix} \mapsto \begin{pmatrix} 0 \\ -1 \end{pmatrix} \\
& \Longrightarrow \quad A_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \\
& S_2: \begin{pmatrix} 1 \\ 0 \end{pmatrix} \mapsto \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 1 \end{pmatrix} \mapsto \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\
& \Longrightarrow \quad A_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}
\end{aligned}$$

Produktmatrizen der beiden verschiedenen Doppelspiegelungen

$$\begin{aligned}
S_2 \circ S_1 : A_2 A_1 &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \\
S_1 \circ S_2 : A_1 A_2 &= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}
\end{aligned}$$

Inverse Matrix

Die mit einer quadratischen Matrix A assoziierte lineare Abbildung $x \mapsto Ax$ ist invertierbar, wenn

$$Ax = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \implies x = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \quad \text{bzw.} \quad \det A \neq 0\,.$$

In diesem Fall bezeichnet A^{-1} die Matrix der inversen linearen Abildung, d.h.

$$AA^{-1} = A^{-1}A = E$$

mit

$$E = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

der Einheitsmatrix.

Für die Multiplikation, Transposition und Bildung adjungierter inverser Matrizen gilt

- $(AB)^{-1} = B^{-1}A^{-1}$,
- $(A^{\mathrm{t}})^{-1} = (A^{-1})^{\mathrm{t}}$,
- $(A^*)^{-1} = (A^{-1})^*$, $A^* = \bar{A}^{\mathrm{t}}$.

Beweis

(i) Kriteria für Invertierbarkeit:
Matrix-Form der Bedingung Kern $L = 0_V$ für allgemeine lineare Abbildungen $L: V \to W \quad \rightsquigarrow \quad$ erstes Kriterium
Die Äquivalenz zum Determinanten-Kriterium folgt aus der Theorie linearer Gleichungssysteme:
$\det A \neq 0 \quad \Longrightarrow \quad Ax = (0, \ldots, 0)^{\mathrm{t}}$ hat nur die triviale Lösung $x = (0, \ldots, 0)^{\mathrm{t}}$

(ii) Regeln für inverse Matrizen:

- $(AB)^{-1} = B^{-1}A^{-1}$ folgt aus

$$(AB)\,(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AEA^{-1} = AA^{-1} = E$$

- Zum Beweis von $A^{\mathrm{t}}(A^{-1})^{\mathrm{t}} = E$ setze $B = A^{-1}$, $C(j,k) = c_{j,k}$ und berechne das Element (i,k) der linken Seite:

$$\begin{aligned}(A^{\mathrm{t}}B^{\mathrm{t}})(i,k) &= \sum_j A^{\mathrm{t}}(i,j)B^{\mathrm{t}}(j,k) = \sum_j A(j,i)B(k,j) \\ &= (BA)(k,i) = E(k,i) = E(i,k)\end{aligned}$$

- Analog erhält man für das Element (i,k) von A^*B^*

$$\sum_j A^*(i,j)B^*(j,k) = \sum_j \overline{A(j,i)}\,\overline{B(k,j)}$$

$$\overline{\sum_j B(k,j)A(j,i)} = \overline{(BA)(k,i)} = \overline{E(k,i)} = E(i,k)$$

Beispiel

Inverse spezieller Matrizen

- Diagonalmatrix:

$$A = \begin{pmatrix} 2 & 0 \\ 0 & 1/2 \end{pmatrix}, \quad A^{-1} = \begin{pmatrix} 1/2 & 0 \\ 0 & 2 \end{pmatrix}$$

  Bilden der Kehrwerte der Diagonalelemente

- Dreiecks-Matrix:

$$B = \begin{pmatrix} 2 & 0 \\ 1 & 1 \end{pmatrix}, \quad B^{-1} = \begin{pmatrix} 1/2 & 0 \\ -1/2 & 1 \end{pmatrix}$$

  Erhaltung der Struktur

- Voll besetzte Matrix:

$$C = \begin{pmatrix} -1 & 2 \\ -1 & 1 \end{pmatrix}, \quad C^{-1} = \begin{pmatrix} 1 & -2 \\ 1 & -1 \end{pmatrix}$$

  Berechnung durch Lösen eines linearen Gleichungssystems

Transponierte und adjungierte Matrix

Durch Vertauschen der Indizes (Spiegelung an der Diagonalen) erhält man aus einer $m \times n$-Matrix A die transponierte $n \times m$-Matrix A^{t}, d.h.

$$A = \begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & & \vdots \\ a_{m,1} & \cdots & a_{m,n} \end{pmatrix}, \quad A^{\mathrm{t}} = \begin{pmatrix} a_{1,1} & \cdots & a_{1,m} \\ \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,m} \end{pmatrix}.$$

Konjugiert man bei einer (komplexen) Matrix A zusätzlich alle Einträge, so ergibt sich die adjungierte Matrix $C = A^* = \bar{A}^{\mathrm{t}}$, d.h.

$$a_{j,k} = u_{j,k} + \mathrm{i}v_{j,k}, \quad c_{j,k} = \bar{a}_{k,j} = u_{k,j} - \mathrm{i}v_{k,j}\,.$$

Stimmt eine Matrix A mit ihrer Transponierten überein, $A = A^{\mathrm{t}}$, bezeichnet man A als symmetrisch, eine Matrix mit $A = A^*$ heißt selbst-adjungiert oder hermitesch. Für reelle Matrizen bedeuten die Begriffe symmetrisch und hermitesch dasselbe.

Es gelten folgende Regeln:

$$
\begin{aligned}
(AB)^{\mathrm{t}} &= B^{\mathrm{t}}A^{\mathrm{t}}, & (AB)^* &= B^*A^*, \\
(A^{\mathrm{t}})^{-1} &= (A^{-1})^{\mathrm{t}}, & (A^*)^{-1} &= (A^{-1})^*.
\end{aligned}
$$

Beweis

(i) Elemente von $C = AB$ und $D = (AB)^{\mathrm{t}}$

$$c_{i,k} = \sum_j a_{i,j} b_{j,k}, \quad d_{i,k} = c_{k,i} = \sum_j a_{k,j} b_{j,i} = \sum_j b_{j,i} a_{k,j}$$

$\implies$ $d_{i,k}$: Element (i,k) von $B^{\mathrm{t}}A^{\mathrm{t}}$, d.h. $D = (AB)^{\mathrm{t}} = B^{\mathrm{t}}A^{\mathrm{t}}$

$\overline{a+b} = \bar{a} + \bar{b}$ und $\overline{ab} = \bar{a}\bar{b}$ $\implies$ gleiche Formel für die Adjungierte

(ii) $\left(A^{-1}\right)^{\mathrm{t}} = (A^{\mathrm{t}})^{-1}$ folgt aus

$$E = E^{\mathrm{t}} = \left(AA^{-1}\right)^{\mathrm{t}} = \left(A^{-1}\right)^{\mathrm{t}} A^{\mathrm{t}}$$

Entsprechendes gilt für die Adjungierte.

Beispiel

Illustration der Regeln für transponierte und adjungierte Matrizen für

$$A = \begin{pmatrix} 0 & \mathrm{i} \\ 1 & 2 \end{pmatrix}, \quad B = \begin{pmatrix} \mathrm{i} & 1 & 0 \\ 2 & 0 & 1+2\mathrm{i} \end{pmatrix}$$

(i) Matrix-Produkt:

$$\begin{aligned}
(AB)^* &= \begin{pmatrix} 2\mathrm{i} & 0 & -2+\mathrm{i} \\ 4+\mathrm{i} & 1 & 2+4\mathrm{i} \end{pmatrix}^* = \begin{pmatrix} \overline{2\mathrm{i}} & \overline{0} & \overline{-2+\mathrm{i}} \\ \overline{4+\mathrm{i}} & \overline{1} & \overline{2+4\mathrm{i}} \end{pmatrix}^{\mathrm{t}} \\
&= \begin{pmatrix} -2\mathrm{i} & 4-\mathrm{i} \\ 0 & 1 \\ -2-\mathrm{i} & 2-4\mathrm{i} \end{pmatrix} \\
B^*A^* &= \begin{pmatrix} -\mathrm{i} & 2 \\ 1 & 0 \\ 0 & 1-2\mathrm{i} \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -\mathrm{i} & 2 \end{pmatrix} = \begin{pmatrix} -2\mathrm{i} & 4-\mathrm{i} \\ 0 & 1 \\ -2-\mathrm{i} & 2-4\mathrm{i} \end{pmatrix}
\end{aligned}$$

(ii) Matrix-Inverse:

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \underbrace{\begin{pmatrix} 0 & \mathrm{i} \\ 1 & 2 \end{pmatrix}}_{A} \begin{pmatrix} 2\mathrm{i} & 1 \\ -\mathrm{i} & 0 \end{pmatrix} = \underbrace{\begin{pmatrix} 0 & 1 \\ -\mathrm{i} & 2 \end{pmatrix}}_{A^*} \begin{pmatrix} -2\mathrm{i} & \mathrm{i} \\ 1 & 0 \end{pmatrix}$$

$\Longrightarrow$

$$(A^{-1})^* = \begin{pmatrix} 2\mathrm{i} & 1 \\ -\mathrm{i} & 0 \end{pmatrix}^* = \begin{pmatrix} -2\mathrm{i} & \mathrm{i} \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -\mathrm{i} & 2 \end{pmatrix}^{-1} = (A^*)^{-1}$$

Spur

Die Spur einer $n \times n$-Matrix A ist die Summe ihrer Diagonalenelemente:

$$\operatorname{Spur} A = \sum_{k=1}^{n} a_{k,k}\,.$$

Für beliebige quadratische Matrizen A, B und invertierbare Matrizen Q gilt

$$\operatorname{Spur}(AB) = \operatorname{Spur}(BA), \quad \operatorname{Spur}(Q^{-1}AQ) = \operatorname{Spur} A\,.$$

Aufgrund letzterer Eigenschaft, der Invarianz der Spur unter Koordinatentransformationen, insbesondere auch unter Transformationen auf Diagonal- oder Dreiecksform, kann die Spur einer Matrix auch als Summe der Eigenwerte (entsprechend ihrer Vielfachheit gezählt) berechnet werden:

$$\operatorname{Spur} A = \sum_{k=1}^{n} \lambda_k\,.$$

Beweis

(i) $\operatorname{Spur}(AB) = \operatorname{Spur}(BA)$:

Definition der Spur und des Matrix-Produkts mit $C = AB$, $D = BA$ $\quad\rightsquigarrow$

$$\operatorname{Spur} C = \sum_j c_{j,j} = \sum_j \sum_k a_{j,k} b_{k,j} = \sum_k \underbrace{\sum_j b_{k,j} a_{j,k}}_{d_{k,k}} = \operatorname{Spur} D$$

(ii) $\operatorname{Spur}(Q^{-1}AQ) = \operatorname{Spur} A$:

Anwendung von (i) auf das Produkt von Q^{-1} und $(AQ)Q$ $\quad\rightsquigarrow$

$$\operatorname{Spur}(Q^{-1}(AQ)) = \operatorname{Spur}((AQ)Q^{-1}) \underset{QQ^{-1}=E}{=} \operatorname{Spur} A$$

(iii) $\operatorname{Spur} A = \sum_k \lambda_k$:

(ii) $\implies$ Invarianz der Spur bei Ähnlichkeitstransformation auf Jordan-Form ($A \to J = Q^{-1}AQ$)

Die Diagonale von J enthält die Eigenwerte von A und folglich ist $\operatorname{Spur} A = \operatorname{Spur} J = \sum_k \lambda_k$.

Beispiel

Illustration der Eigenschaften des Spur-Operators für die Matrizen

$$A = \begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix},\ B = \begin{pmatrix} 2 & 1 \\ 4 & 3 \end{pmatrix},\ Q = \begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix}$$

(i) Matrixprodukt:

$$AB = \begin{pmatrix} 18 & 13 \\ 16 & 11 \end{pmatrix}, \quad BA = \begin{pmatrix} 4 & 11 \\ 10 & 25 \end{pmatrix}$$

und $\mathrm{Spur}(AB) = 18 + 11 = 29 = 4 + 25 = \mathrm{Spur}(BA)$

(ii) Koordinatentransformation:

$$\begin{aligned} Q^{-1}AQ &= \begin{pmatrix} 3 & -2 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix} \\ &= \begin{pmatrix} -1 & 6 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix} = \begin{pmatrix} 5 & 0 \\ 16 & -1 \end{pmatrix} \end{aligned}$$

und $\mathrm{Spur}(Q^{-1}AQ) = 5 + (-1) = 4 = 1 + 3 = \mathrm{Spur}\, A$

(iii) Eigenwerte:

$$\begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = 5 \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} 2 \\ -1 \end{pmatrix} = (-1) \begin{pmatrix} 2 \\ -1 \end{pmatrix}$$

$\rightsquigarrow$ Eigenwerte $\lambda_1 = 5$, $\lambda_2 = -1$ und

$$\operatorname{Spur} A = 5 + (-1) = 4 \quad \checkmark$$

Für eine $m \times n$-Matrix A stimmt die maximale Anzahl linear unabhängiger Zeilen u_j^{t} mit der maximalen Anzahl linear unabhängiger Spalten v_k überein und wird Rang der Matrix genannt:

$$\operatorname{Rang} A = \dim \underbrace{\operatorname{span}(u_1, \ldots, u_m)}_{U} = \dim \underbrace{\operatorname{span}(v_1, \ldots, v_n)}_{V} .$$

Bezeichnet $L : \mathbb{R}^n \ni x \mapsto Ax \in \mathbb{R}^m$ die durch A definierte lineare Abbildung, so ist

$$U = (\operatorname{Kern} L)^\perp, \quad V = \operatorname{Bild} L$$

mit $W^\perp$ dem orthogonalen Komplement eines Unterraums W ($w^\perp \in W^\perp \iff \langle w^\perp, w \rangle = 0 \quad \forall w \in W$).

Offensichtlich gilt

$$\operatorname{Rang} A \leq \min(m, n) ,$$

und für eine invertierbare $n \times n$-Matrix A ist $\operatorname{Rang} A = n$.

Der Rang von A lässt sich bestimmen, indem man mit dem Gram-Schmidt-Verfahren eine orthogonale Basis für U oder V konstruiert. Alternativ kann man A auf Zeilenstufen-Form (Echelon-Form) transformieren. Die dazu nötigen Operationen, Addition von Zeilenvielfachen und Permutationen von Zeilen oder Spalten, lassen den Rang unverändert. $\operatorname{Rang} A$ ist dann die Anzahl der Pivots.

Beweis

zu zeigen:

$$\dim U = \dim V \quad (= \operatorname{Rang} A)$$

Satz über die Dimension von Bild und Kern der mit A assoziierten linearen Abbildung $L:\ \mathbb{R}^n \to \mathbb{R}^m \qquad \Longrightarrow$

$$\dim V = \dim \operatorname{Bild} A = n - \dim \operatorname{Kern} A$$

$$w \in \operatorname{Kern} A \quad \Longleftrightarrow \quad \begin{pmatrix} u_1^{\mathrm{t}} \\ \vdots \\ u_m^{\mathrm{t}} \end{pmatrix} \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \text{d.h. } u_j^{\mathrm{t}}\, w = 0\ \forall j$$

$\Longrightarrow \qquad U = (\operatorname{Kern} A)^\perp$ (orthogonales Komplement)

$\Longrightarrow \qquad$ Dimensionen addieren zu n:

$$n = \dim U + \dim \operatorname{Kern} A$$

Einsetzen in die Dimensionsformel $\quad \Longrightarrow \quad$ Behauptung

Beispiel

Rang von $m \times n$-Matrizen mit $m, n \leq 3$

(i) 2×2-Matrix $A = (v_1, v_2)$:

$$\operatorname{Rang} A = \begin{cases} 0, & \text{falls } v_1 = v_2 = (0,0)^{\mathrm{t}} \\ 1, & \text{falls } v_1 \parallel v_2 \\ 2, & \text{falls } v_1 \nparallel v_2 \end{cases}$$

z.B.

$$\operatorname{Rang} \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} = 1, \quad \operatorname{Rang} \begin{pmatrix} 1 & 2 \\ 4 & 2 \end{pmatrix} = 2$$

(ii) 2×3-Matrix $A = \begin{pmatrix} u_1^{\mathrm{t}} \\ u_2^{\mathrm{t}} \end{pmatrix}$, 3×2-Matrix $A = (v_1, v_2)$:

gleiches Kriterium für die beiden, die Matrix definierenden Vektoren, wie im Fall (i), z.B.

$$\operatorname{Rang} \begin{pmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \end{pmatrix} = 1, \quad \operatorname{Rang} \begin{pmatrix} 2 & 1 \\ 2 & 2 \\ 2 & 4 \end{pmatrix} = 2$$

(iii) 3×3-Matrix $A = (v_1, v_2, v_3)$:

$$\operatorname{Rang} A = \begin{cases} 0, & \text{falls } v_1 = v_2 = v_3 = (0,0,0)^{\mathrm{t}} \\ 1, & \text{falls } v_1 \parallel v_2 \parallel v_3 \\ 2, & \text{falls } v_k \text{ linear abhängig, nicht alle parallel} \\ 3, & \text{falls } v_k \text{ linear unabhängig} \end{cases}$$

z.B.

$$\operatorname{Rang} \begin{pmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \\ 3 & 3 & 3 \end{pmatrix} = 1, \ \operatorname{Rang} \begin{pmatrix} 1 & 1 & 2 \\ 2 & 2 & 1 \\ 3 & 3 & 3 \end{pmatrix} \underset{(1)}{=} 2, \ \operatorname{Rang} \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \\ 2 & 3 & 1 \end{pmatrix} \underset{(2)}{=} 3$$

(1) Zeilen linear abhängig:

$$(1,1,2) + (2,2,1) - (3,3,3) = (0,0,0)$$

(2) Berechnung des Rangs durch Transformation auf Dreiecksform durch Addition von Zeilenvielfachen:

$$\begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \\ 2 & 3 & 1 \end{pmatrix} \underset{(a)}{\rightarrow} \begin{pmatrix} 1 & 2 & 3 \\ 0 & -5 & -7 \\ 0 & -1 & -5 \end{pmatrix} \underset{(b)}{\rightarrow} \begin{pmatrix} 1 & 2 & 3 \\ 0 & -1 & -5 \\ 0 & 0 & 18 \end{pmatrix}$$

(a) (Zeile 2)-3·(Zeile 1), (Zeile 3)-2·(Zeile 1)
(b) Vertauschen von Zeile 2 und Zeile 3, (Zeile 3)-5·(Zeile 2)
Diagonalelemente der Dreiecksform ungleich 0 $\quad\Longrightarrow\quad$ Rang $A = 3$

Beispiel

Rang der Matrix

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

(i) Gram-Schmidt-Verfahren:
Konstruktion einer orthogonalen Basis $\{b_1, b_2, \ldots\}$ für den durch die Spalten $v_1, \ldots, v_5$ von A aufgespannten Bildraum V der linearen Abbildung $x \mapsto Ax$ (Rang $A = \dim V$)

Schritt 1:
$b_1 = v_1 = (1, 1, 2, 2)^{\mathrm{t}}$

Schritt 2:
$$b_2 = v_2 - \frac{b_1^{\mathrm{t}} v_2}{b_1^{\mathrm{t}} b_1} b_1 = (1, 1, 1, 1)^{\mathrm{t}} - \frac{(1,1,2,2)(1,1,1,1)^{\mathrm{t}}}{(1,1,2,2)(1,1,2,2)^{\mathrm{t}}}(1, 1, 2, 2)^{\mathrm{t}} =$$
$$(1, 1, 1, 1)^{\mathrm{t}} - \frac{6}{10}(1, 1, 2, 2)^{\mathrm{t}} = (2/5, 2/5, -1/5, -1/5)^{\mathrm{t}}$$

Schritt 3:
$b_1^t v_3 = (1,1,2,2)(1,3,1,3)^t = 12, \quad b_1^t b_1 = 10,$
$b_2^t v_3 = (2/5, 2/5, -1/5, -1/5)(1,3,1,3)^t = 4/5, \quad b_2^t b_2 = 10/25$
$\rightsquigarrow$
$$b_3 = v_3 - \frac{b_1^t v_3}{b_1^t b_1} b_1 - \frac{b_2^t v_3}{b_2^t b_2} b_2 =$$
$$(1,3,1,3)^t - \frac{12}{10}(1,1,2,2)^t - \frac{4/5}{10/25}(2/5,2/5,-1/5,-1/5)^t =$$
$$(-1,1,-1,1)^t$$
Schritte 4 und 5:
$$v_4 - \sum_{k=1}^{3} \frac{b_k^t v_4}{b_k^t b_k} b_k = (0,0,0,0)^t \quad \Longrightarrow \quad v_4 \in \operatorname{span}(b_1, b_2, b_3),$$
d.h. kein weiterer Basisvektor, ebenso wie für v_5
$\Longrightarrow \quad \text{Rang}\, A = \dim V = 3$

(ii) Transformation auf Zeilenstufenform:
Gauß-Operationen

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 2 \\ 1 & 1 & 3 & 2 & 3 \\ 2 & 1 & 1 & 2 & 2 \\ 2 & 1 & 3 & 3 & 3 \end{pmatrix} \underset{(1)}{\longrightarrow} \begin{pmatrix} 1 & 1 & 1 & 1 & 2 \\ 0 & 0 & 2 & 1 & 1 \\ 0 & -1 & -1 & 0 & -2 \\ 0 & -1 & 1 & 1 & -1 \end{pmatrix} \underset{(2)}{\longrightarrow}$$

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 2 \\ 0 & -1 & -1 & 0 & -2 \\ 0 & 0 & 2 & 1 & 1 \\ 0 & 0 & 2 & 1 & 1 \end{pmatrix} \underset{(3)}{\longrightarrow} \begin{pmatrix} 1 & 1 & 1 & 1 & 2 \\ 0 & -1 & -1 & 0 & -2 \\ 0 & 0 & 2 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

(1) (Zeile 2) $-$ (Zeile 1), (Zeile 3) $-$ 2$\cdot$(Zeile 1), (Zeile 4) $-$ 2$\cdot$(Zeile 1)
(2) Vertauschen von (Zeile 2) und (Zeile 3), (Zeile 4) $-$ (Zeile 2)
(3) (Zeile 4) $-$ (Zeile 3)

3 Pivots,1, -1 und 2 $\quad\Longrightarrow\quad$ Rang $A = 3$

Norm einer Matrix

Einer Vektornorm ist die Matrixnorm

$$\|A\| = \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|} = \max_{\|x\|=1} \|Ax\|$$

zugeordnet. Zusätzlich zu den Normeigenschaften (Positivität, Homogenität, Dreiecksungleichung) gilt

$$\|AB\| \leq \|A\| \|B\|$$

für Produkte von Matrizen, d.h. die zugeordnete Matrixnorm ist submultiplikativ. Diese Eigenschaft müssen andere Matrixnormen, die ebenfalls verwendet werden können, nicht besitzen. Es wird jedoch in Verbindung mit dem Matrix/Vektor-Kalkül gefordert, dass eine Matrixnorm kompatibel mit der Vektornorm ist, d.h.

$$\|Ax\| \leq \|A\| \|x\| \, .$$

Diese schwächere Bedingung als Submultiplikativität ist für die zugeordnete Matrixnorm eine unmittelbare Folgerung aus der Definition.

Beweis

(i) Positivität: ✓

(ii) Homogenität:

$$\begin{aligned}\|sA\| &= \max_{\|x\|=1} \|sAx\| = \max_{\|x\|=1} |s|\|Ax\| \\ &= |s| \max_{\|x\|=1} \|Ax\| = |s|\|A\|\end{aligned}$$

(iii) Dreiecksungleichung:

$$\begin{aligned}\|A+B\| &= \max_{\|x\|=1} \|(A+B)x\| = \max_{\|x\|=1} \|Ax+Bx\| \\ &\leq \max_{\|x\|=1} (\|Ax\| + \|Bx\|) \\ &\leq \max_{\|x\|=1} \|Ax\| + \max_{\|x\|=1} \|Bx\| = \|A\| + \|B\|\end{aligned}$$

(iv) Submultiplikativität:

Für eine Nullmatrix B ($\|B\| = 0$) ist die Aussage trivial, da in diesem Fall AB ebenfalls die Nullmatrix ist.

Andernfalls gilt Folgendes:

$$\begin{aligned} \|AB\| &= \sup_{x\neq 0} \frac{\|ABx\|}{\|x\|} \underset{(*)}{=} \sup_{x:Bx\neq 0} \frac{\|A(Bx)\|}{\|Bx\|} \frac{\|Bx\|}{\|x\|} \\ &\leq \sup_{y\neq 0} \frac{\|Ay\|}{\|y\|} \sup_{x\neq 0} \frac{\|Bx\|}{\|x\|} = \|A\|\|B\| \end{aligned}$$

Die Einschränkung bei der Umformung $(*)$ ist unproblematisch, da $Bx = 0 \implies ABx = 0$.

Maximumnorm

Der Maximumnorm für Vektoren,

$$\|v\|_\infty = \max_k |v_k| ,$$

ist für Matrizen die Zeilensummennorm zugeordnet:

$$\|A\|_\infty = \sup_{x \neq 0} \frac{\|Ax\|_\infty}{\|x\|_\infty} = \max_j \sum_k |a_{j,k}| .$$

Dabei ist zu beachten, dass im allgemeinen $\|A\|_\infty \neq \|A^t\|_\infty$. Insbesondere gilt für einen Zeilenvektor $(x_1, \ldots, x_n)$ (d.h. eine $1 \times n$-Matrix)

$$\|x\|_\infty = \sum_{j=1}^{n} |x_j| .$$

Beweis

Definition der zugeorneten Norm $\implies$

$$\|A\|_\infty = \max_{\|x\|_\infty=1} \max_j \Big|\sum_k a_{j,k} x_k\Big| \le \max_j \sum_k |a_{j,k}|$$

Gleichheit gilt für

$$x_k = \operatorname{sign}\, a_{j,k}$$

mit j einem Index, für den die Zeilensumme maximal ist.

Euklidische Matrixnorm

Der Euklidischen Norm (2-Norm) für Vektoren,

$$|v| = \left(\sum_k |v_k|^2 \right)^{1/2} ,$$

ist die Matrixnorm

$$\|A\|_2 = \max\{\sqrt{\lambda} : \lambda \,\text{ist Eigenwert von}\, A^*A\}$$

zugeordnet. Die Wurzeln der Eigenwerte der symmetrischen, positiv definiten Matrix A^*A sind die Singulärwerte von A.

Kompatibel mit der Euklidischen Norm ist ebenfalls die Frobenius-Norm

$$\|A\|_F = \left(\sum_{j,k} |a_{j,k}|^2 \right)^{1/2} ,$$

d.h. es gilt

$$|Ax| \le \|A\|_F |x| .$$

Diese Norm ist jedoch nicht submultiplikativ.

Beweis

(i) Zugeordnete Norm:
benutze die Singulärwertzerlegung

$$A = USV^*, \quad U, V \text{ unitär}, \ S \text{ diagonal}$$

Invarianz der Euklidischen Norm bei unitären Transformationen $\implies$

$$\|A\|_2^2 = \sup_{x \neq 0} \frac{|Ax|^2}{|x|^2} = \sup_{x \neq 0} \frac{|USV^*x|^2}{|x|^2} = \sup_{x \neq 0} \frac{|SV^*x|^2}{|V^*x|^2} = \sup_{y \neq 0} \frac{|Sy|^2}{|y|^2}$$

$|Sy|^2 = \sum_{k=1}^{r} |s_k y_y|^2$ mit r dem Rang und $s_1 \geq \cdots \geq s_r$ den Singulärwerten von A

$$\implies \quad \|A\|_2^2 = \sup_{y} \frac{\sum_k |s_k y_k|^2}{\sum_k |y_k|^2} \leq \max_k s_k = s_1$$

mit Gleichheit für $y = (1, 0, \ldots)^{\mathrm{t}}$

(ii) Frobenius-Norm:

Kompatibilität $\Longleftarrow$ Ungleichung von Cauchy-Schwarz

$$
\begin{aligned}
|Ax|^2 &= \sum_j \left(\sum_k a_{j,k} x_k \right)^2 \le \sum_j \left(\sqrt{\sum_k |a_{j,k}|^2} \sqrt{\sum_k |x_k|^2} \right)^2 \\
&= \left(\sum_j \sum_k |a_{j,k}|^2 \right) |x|^2 = \|A\|_F^2 |x|^2
\end{aligned}
$$

Beispiel

verschiedene Normen der Matrix

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

(i) Zeilensummennorm:

$$\|A\|_\infty = \max_j \sum_k |a_{j,k}| = \max\{|-1|, |1|, |-2| + |2|\} = \max\{1, 1, 4\} = 4$$

Gleichheit in der Ungleichung $\|Ax\|_\infty \leq \|A\|_\infty \|x\|_\infty$ für $x_k = \operatorname{sign} a_{3,k}$, d.h. $x = (-1, 1)^{\mathrm{t}}$:

$$\|Ax\|_\infty = \left\| \begin{pmatrix} -1 \\ -1 \\ 4 \end{pmatrix} \right\|_\infty = 4, \quad \|x\|_\infty = 1$$

(ii) Euklidische Norm:
$\|A\|_2 = \max\{\sqrt{\lambda} : \lambda \text{ ist Eigenwert von } A^{\mathrm{t}}A\}$

$$A^{\mathrm{t}}A = \begin{pmatrix} 0 & 1 & -2 \\ -1 & 0 & 2 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \\ -2 & 2 \end{pmatrix} = \begin{pmatrix} 5 & -4 \\ -4 & 5 \end{pmatrix}$$

Eigenwerte $\lambda_1 = 9$, $\lambda_2 = 1$, denn

$$A^{\mathrm{t}}A \begin{pmatrix} 1 \\ -1 \end{pmatrix} = 9 \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad A^{\mathrm{t}}A \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

Gleichheit in der Ungleichung $|Ax| \leq \|A\|_2 |x|$ für $x = (1, -1)^{\mathrm{t}}$:

$$|Ax| = \left| \begin{pmatrix} 1 \\ 1 \\ -4 \end{pmatrix} \right| = \sqrt{18}, \quad |x| = \sqrt{2}$$

und $\sqrt{18} = 9\sqrt{2}$ $\quad \checkmark$

(iii) Frobenius-Norm:
$\|A\|_F = \sqrt{1 + 1 + (-2)^2 + 2^2} = \sqrt{10}$

Eine komplexe $n \times n$-Matrix A ist unitär, falls

$$A^{-1} = \bar{A}^{\mathrm{t}} = A^{*}\,,$$

d.h. falls die Spalten von A eine orthonormale Basis von $\mathbb{C}^n$ bilden.

Für relle Matrizen entfällt (wie auch beim Skalarprodukt) die komplexe Konjugation,

$$A^{-1} = A^{\mathrm{t}}\,,$$

und man bezeichnet A als orthogonal.

Eine reelle (komplexe) $n \times n$-Matrix A ist genau dann orthogonal (unitär), wenn sie die euklidische Norm jedes Vektors invariant lässt:

$$|Av| = |v| \quad \forall v \in \mathbb{R}^n \, (\in \mathbb{C}^n)\,.$$

Beweis

Es genügt, den allgemeineren komplexen Fall zu betrachten.

(i) A unitär $\implies$ $(Ay)^*(Ax) = y^*A^*Ax = y^*x$

$\rightsquigarrow$ Invarianz des komplexen Skalarprodukts und damit auch der Norm

(ii) Normtreue $\implies$ Normierung der Spalten v_j von A, die Bilder der Einheitsvektoren e_j sind:

$$|v_j| = |Ae_j| = 1$$

zum Beweis der Orthogonalität wähle $\lambda = \exp(\mathrm{i}\vartheta)$, so dass

$$s = v_j^*(\lambda v_k) \in \mathbb{R}$$

Normtreue und Definition der euklidischen Norm $\rightsquigarrow$

$$\begin{aligned} 2 &= |e_j + \lambda e_k|^2 = |v_j + \lambda v_k|^2 \\ &= |v_j|^2 + |\lambda|^2 |v_k|^2 + v_j^*(\lambda v_k) + \underbrace{(\lambda v_k)^* v_j}_{=\bar{s}=s} = 2 + 2s \end{aligned}$$

$\implies$ $s = 0$ und somit ebenfalls $v_j^* v_k = 0$

Einige elementare orthogonale und unitäre Matrizen

(i) Drehung (orthogonal):

$$A = \begin{pmatrix} c & -s \\ s & c \end{pmatrix}, \quad c = \cos\varphi,\ s = \sin\varphi$$

$A^{\mathrm{t}} A \overset{!}{=} E$, d.h. $A^{\mathrm{t}} = A^{-1}$

$$\begin{pmatrix} c & s \\ -s & c \end{pmatrix} \begin{pmatrix} c & -s \\ s & c \end{pmatrix} = \begin{pmatrix} c^2 + s^2 & -cs + sc \\ -sc + cs & (-s)^2 + c^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \checkmark$$

(ii) Spiegelung (symmetrisch, orthogonal):
d: normierter Normalenvektor der Spiegelungsebene $\rightsquigarrow$

$$A = E - 2dd^{\mathrm{t}}$$

$A^{\mathrm{t}} \stackrel{!}{=} A$

$$(E - 2dd^{\mathrm{t}})^{\mathrm{t}} = E^{\mathrm{t}} - 2(dd^{\mathrm{t}})^{\mathrm{t}} = E - 2(d^{\mathrm{t}})^{\mathrm{t}} d^{\mathrm{t}} = E - 2dd^{\mathrm{t}} = A \quad \checkmark$$

$A^{\mathrm{t}}A = AA \stackrel{!}{=} E$
geometrisch offensichtlich (zweifache Spiegelung $\rightsquigarrow$ Identität)
rechnerisch:

$$(E - 2dd^{\mathrm{t}})(E - 2dd^{\mathrm{t}}) = E - 2dd^{\mathrm{t}} - 2dd^{\mathrm{t}} + 4d \underbrace{d^{\mathrm{t}}d}_{=1} d^{\mathrm{t}} = E \quad \checkmark$$

konkrete Spiegelung: $d = (1,\, 1,\, 1)^{\mathrm{t}}/\sqrt{3}$ $\rightsquigarrow$

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

(iii) Fourier-Matrix (unitär):

$$A = \frac{1}{2} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & \mathrm{i} & -1 & -\mathrm{i} \\ 1 & -1 & 1 & -1 \\ 1 & -\mathrm{i} & -1 & \mathrm{i} \end{pmatrix}$$

$A^* A \stackrel{!}{=} E$, d.h. $A^* = A^{-1}$

$$\frac{1}{4} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -\mathrm{i} & -1 & \mathrm{i} \\ 1 & -1 & 1 & -1 \\ 1 & \mathrm{i} & -1 & -\mathrm{i} \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & \mathrm{i} & -1 & -\mathrm{i} \\ 1 & -1 & 1 & -1 \\ 1 & -\mathrm{i} & -1 & \mathrm{i} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \checkmark$$

Beispiel

Hadamard- Matrizen H: nach Normierung orthogonal, $h_{j,k} \in \{-1, 1\}$

Konstuktion für $n = 2^k$

$$H_1 = \begin{pmatrix} 1 \end{pmatrix}, \quad H_2 = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \quad H_4 = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}$$

Rekursion

$$H_{2k} = \begin{pmatrix} H_k & H_k \\ H_k & -H_k \end{pmatrix}$$

Normierung $H_k/\sqrt{k} \quad \rightsquigarrow \quad$ orthogonale Matrizen

Fourier-Matrix

Die Fourier-Matrix besteht aus Potenzen der Einheitswurzel $w_n = \mathrm{e}^{2\pi \mathrm{i}/n}$:

$$W_n = \begin{pmatrix} w_n^{0\cdot 0} & \cdots & w_n^{0\cdot(n-1)} \\ \vdots & & \vdots \\ w_n^{(n-1)\cdot 0} & \cdots & w_n^{(n-1)\cdot(n-1)} \end{pmatrix} .$$

Sie ist nach Normierung ($W_n \to W_n/\sqrt{n}$) unitär, d.h. $W_n^* W_n / n$ ist die Einheitsmatrix.

Die Orthogonalität der Spalten v_j folgt aus der geometrischen Summenformel

$$v_j^* v_k = \sum_{\ell=0}^{n-1} \overline{w_n^{\ell j}} w_n^{\ell k} = \sum_{\ell=0}^{n-1} w_n^{(k-j)\ell} = \frac{w_n^{(k-j)n} - 1}{w_n^{k-j} - 1} = 0, \quad j \neq k, \, ,$$

da $w_n^n = 1$.

Normale Matrizen

Eine komplexe Matrix A ist normal, falls

$$AA^* = A^*A, \quad A^* = \bar{A}^{\mathrm{t}}\,.$$

Insbesondere sind unitäre und hermitesche Matrizen normal.

Für reelle normale Matrizen entfällt die komplexe Konjugation, d.h.,

$$AA^{\mathrm{t}} = A^{\mathrm{t}}A\,.$$

Dies ist für orthogonale und symmetrische Matrizen erfüllt.

Beweis

A unitär, d.h. $A^* = A^{-1} \quad \Longrightarrow$

$$AA^* = AA^{-1} = E = A^{-1}A = A^*A$$

A hermitesch, d.h. $A^* = A \quad \Longrightarrow$

$$AA^* = AA = A^*A$$

Beide Eigenschaften sind demnach hinreichend für Normalität.

Für reelle Matrizen A ist $A^* = A^{\mathrm{t}}$. Die obigen Argumente gelten also auch für diesen Fall.

Beispiel

Konstruktion einer normalen 2×2-Matrix

$$A = \begin{pmatrix} 1 & a \\ b & c \end{pmatrix}$$

$$AA^{\mathrm{t}} = \begin{pmatrix} 1 & a \\ b & c \end{pmatrix} \begin{pmatrix} 1 & b \\ a & c \end{pmatrix} = \begin{pmatrix} 1+a^2 & b+ac \\ b+ac & b^2+c^2 \end{pmatrix}$$
$$A^{\mathrm{t}}A = \begin{pmatrix} 1 & b \\ a & c \end{pmatrix} \begin{pmatrix} 1 & a \\ b & c \end{pmatrix} = \begin{pmatrix} 1+b^2 & a+bc \\ a+bc & a^2+c^2 \end{pmatrix}$$

normal, falls $AA^{\mathrm{t}} = A^{\mathrm{t}}A$, d.h. falls

$$a = \pm b \quad \wedge \quad b + ac = a + bc$$

$\iff$

$$a = b \quad \vee \quad (a = -b,\ c = 1)$$

Zyklische Matrizen

Bei einer zyklischen $n \times n$-Matrix werden die Spalten durch zyklisches Verschieben der ersten Spalte gebildet:

$$C = \begin{pmatrix} a_0 & a_{n-1} & a_{n-2} & \ldots \\ a_1 & a_0 & a_{n-1} & \ldots \\ a_2 & a_1 & a_0 & \ldots \\ \vdots & \vdots & \vdots & \end{pmatrix}, \quad \text{d.h. } c_{j,k} = a_{j-k \bmod n}\,.$$

Die zyklische Struktur bleibt bei Transposition, Multiplikation und Invertierung erhalten.

Beweis

(i) Transposition: $\checkmark$

(ii) Produkt:

$Q = CD, \quad c_{i,j} = a_{i-j \bmod n},\ d_{i,j} = b_{i-j \bmod n} \quad \Longrightarrow$

$$q_{i,k} = \sum_{j=1}^{n} a_{i-j \bmod n}\, b_{j-k \bmod n}$$

Substitution $j = j' + k \quad \rightsquigarrow$

$$q_{i,k} = \sum_{j'=1-k}^{n-k} a_{(i-k)-j' \bmod n}\, b_{j' \bmod n}$$

Modulo-Arithmetik $\Longrightarrow$ Invarianz der Summanden bei Ersetzen von j' durch $j' \pm n$

$\rightsquigarrow$ Änderung des Summationsbereiches

$$\begin{aligned}\{1-k,\ldots,0,1,\ldots,n-k\} &\to \{n+1-k,\ldots,n,1,\ldots,n-k\}\\ &= \{1,\ldots,n\}\end{aligned}$$

$\Longrightarrow$ $q_{i,k}$ hängt nur von $(i-k) \bmod n$ ab, d.h. Q ist zyklisch: $q_{i,k} = p_{i-k \bmod n}$

(iii) Inverse:

$D = C^{-1}$, $c_{j,k} = a_{j-k \bmod n}$

Für die erste Zeile $(b_0, b_{n-1}, \ldots, b_1)$ von D gilt wegen $E = DC$

$$\delta_{1,\ell} = \sum_{k=1}^{n} b_{1-k \bmod n}\, a_{k-\ell \bmod n}$$

zeige $d_{i,k} = b_{i-k \bmod n}$:

wähle dazu $\ell = j - i + 1 \bmod n$, substituiere $k' = k + i - 1$ und erhalte

$$\begin{aligned}\delta_{i,j} = \delta_{1,j-i+1 \bmod n} &= \sum_{k=1}^{n} b_{1-k \bmod n}\, a_{k-j+i-1 \bmod n} \\ &= \sum_{k'=i}^{i+n-1} b_{i-k' \bmod n} a_{k'-j \bmod n}\end{aligned}$$

Verschiebung des Summationsbereiches, $\{i, \ldots, i+n-1\} \to \{1, \ldots, n\}$

$\Longrightarrow \quad d_{i,k} = b_{i-k \bmod n}$ erfüllt $E = DC$ und ist somit inverse Matrix

Beispiel

Produkt CD und Inverse C^{-1} für die zyklischen Matrizen

$$C = \begin{pmatrix} 0 & 2 & -1 \\ -1 & 0 & 2 \\ 2 & -1 & 0 \end{pmatrix}, \quad D = \begin{pmatrix} 3 & 5 & 4 \\ 4 & 3 & 5 \\ 5 & 4 & 3 \end{pmatrix}$$

- Produkt:

$$CD = \begin{pmatrix} 3 & 2 & 7 \\ 7 & 3 & 2 \\ 2 & 7 & 3 \end{pmatrix}$$

- Inverse:

$$C^{-1} = \frac{1}{7} \begin{pmatrix} 2 & 1 & 4 \\ 4 & 2 & 1 \\ 1 & 4 & 2 \end{pmatrix}$$

Positiv definite Matrizen

Eine quadratische Matrix A ist positiv definit, falls

$$v^* A v > 0 \quad \forall v \neq (0, \ldots, 0)^{\mathrm{t}}$$

($v^* = v^{\mathrm{t}}$ für reelle Vektoren).

Ist $v^* A v$ lediglich nicht-negativ, so bezeichnet man A als positiv semidefinit.

Eine positiv definite Matrix A hat ausschließlich positive Diagonalelemente und Eigenwerte. Insbesondere ist A invertierbar und die Inverse ist ebenfalls positiv definit.

Beweis

(i) Diagonalelemente:

$$g_{j,k} = e_j^{\mathrm{t}} \underbrace{Ae_k}_{\text{Spalte } k} > 0$$

(ii) Eigenwerte:

$$\begin{aligned} Av = \lambda v \qquad &\Longrightarrow \qquad 0 < v^* A v = v^*(\lambda v) = \lambda |v|^2 \\ &\Longrightarrow \qquad \lambda > 0 \end{aligned}$$

(iii) Inverse:

Eigenwerte $> 0 \quad \Longrightarrow \quad$ Existenz von A^{-1} und

$$v^* A^{-1} v = (AA^{-1}v)^* A^{-1} (A \underbrace{A^{-1}v}_{w}) = w^* A^* w = (\underbrace{w^* A w}_{>0})^* > 0$$

Beispiel

Gramsche Matrix G einer Basis $\{v_1, \ldots, v_n\}$:

$$g_{j,k} = \langle v_j, v_k \rangle$$

G ist positiv definit, da

$$x^* G x = \sum_{j,k} \overline{x}_j g_{j,k} x_k = \Big\langle \underbrace{\sum_j x_j v_j}_{u}, \sum_k x_k v_k \Big\rangle = \langle u, u \rangle > 0$$

für $u \neq 0 \quad \Longleftrightarrow \quad x \neq (0, \ldots, 0)^{\mathrm{t}}$

Hilbert Matrix:

Basis $p_k : x \mapsto x^k$, $k = 0, \ldots, n$, für Polynome vom Grad $\leq n$,

Skalarprodukt $\langle f, g \rangle = \int_0^1 f(x) g(x)\, \mathrm{d}x$

$$\rightsquigarrow \quad G : g_{j,k} = \langle p_j, p_k \rangle = \int_0^1 x^j x^k \, \mathrm{d}x = \frac{1}{1 + j + k}$$

Die Determinante

$$\det A = \det(a_1, \ldots, a_n)$$

einer quadratischen Matrix A mit Spalten a_j kann durch folgende Eigenschaften definiert werden.

- Multilinearität:

$$\det(\ldots, \alpha a_j + \beta b_j, \ldots) = \alpha \det(\ldots, a_j, \ldots) + \beta \det(\ldots, b_j, \ldots)$$

- Antisymmetrie:

$$\det(\ldots, a_j, \ldots, a_k, \ldots) = -\det(\ldots, a_k, \ldots, a_j, \ldots)\,,$$

  insbesondere $\det A = 0$ bei zwei gleichen Spalten

- Normierung:

$$\det(e_1, \ldots, e_n) = 1, \quad (e_k)_\ell = \delta_{k\ell}\,.$$

Mit den definierenden Regeln lässt sich eine Determinante als Summe n-facher Produkte entwickeln:

$$\det A = \sum_{p \in S_n} \sigma(p)\, a_{p(1),1} \cdots a_{p(n),n}\,,$$

wobei über alle Permutationen p von $(1, \ldots, n)$ summiert wird, und $\sigma(p)$ das Vorzeichen von p bezeichnet.

Man benutzt ebenfalls die Schreibweisen

$$\det A = |A| = \begin{vmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,n} \end{vmatrix}.$$

Wegen der hohen Anzahl der Summanden (es exisitieren n! Permutationen) ist die explizite Darstellung der Determinante für die praktische Berechnung schlecht geeignet. Sie ist jedoch unmittelbar mit den definierenden Eigenschaften verknüpft und wird zum Beweis sowie zur Herleitung einiger anderer Eigenschaften benötigt.

Beweis

(i) Eigenschaften $\implies$ Entwicklung der Determinante via Permutationen: Darstellung der Spalten von A als Linearkombinationen der Einheitsvektoren e_j,

$$a_k = \sum_{j=1}^{n} a_{j,k} e_j \,,$$

und Multilinearität $\implies$

$$\det A = \sum_{k_1=1}^{n} \cdots \sum_{k_n=1}^{n} a_{k_1,1} \cdots a_{k_n,n} \underbrace{\det(e_{k_1}, \ldots, e_{k_n})}_{=d_k}$$

Antisymmetrie $\implies$

- $\det(e_{k_1}, \ldots, e_{k_n}) = 0$, falls nicht alle e_{k_ν} verschieden sind, d.h. nur Permutationen sind zu berücksichtigen, $k = (p(1), \ldots, p(n))$, $p \in S_n$
- $d_k = (-1)^{\tau(k)} \det(e_1, \ldots, e_n)$, wobei $\tau(k)$ die modulo 2 eindeutig bestimmte Anzahl von Vertauschungen bezeichnet, um die Einheitsvektoren in die kanonische Reihenfolge zu bringen

Definition des Vorzeichens einer Permutation und Normierung $\implies$

$$d_k = \det(e_{p(1)}, \ldots, e_{p(n)}) = \sigma(p) \det(e_1, \ldots, e_n) = \sigma(p)$$

(ii) Entwicklung $\implies$ Eigenschaften:
Multilinearität $\impliedby$ Produkte

$$a_{k_1,1} \cdots a_{k_n,n}$$

enthalten aus jeder Spalte jeweils genau ein Element.
Antisymmetrie $\impliedby$ Vertauschung von Spalten ändert Vorzeichen der Permutation.
Normierung $\impliedby$ Für die Einheitsmatrix existiert nur ein nichttrivialer Summand:

$$a_{1,1} \cdots a_{n,n} = 1 \cdots 1 = 1\,.$$

Beispiel

Determinante einer (2×2)-Matrix:

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

(i) Berechnung mit Hilfe der definierenden Eigenschaften:

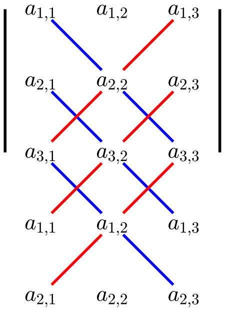
$$\begin{aligned} \det(ae_1 + ce_2, be_1 + de_2) &= a\det(e_1, be_1 + de_2) + c\det(e_2, be_1 + de_2) \\ &= ab\det(e_1, e_1) + ad\det(e_1, e_2) + cb\det(e_2, e_1) + cd\det(e_2, e_2) \\ &= ab \cdot 0 + ad \cdot 1 + cb \cdot (-1) + cd \cdot 0 = ad - bc \end{aligned}$$

(ii) Entwicklung nach Permutationen:
Dimension $n = 2 \quad \rightsquigarrow \quad n! = 2$ Permuationen $p = (1, 2)$ und $q = (2, 1)$

$$\begin{aligned} \begin{vmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{vmatrix} &= \sigma(p) a_{p(1),1} a_{p(2),2} + \sigma(\ell) a_{q(1),1} a_{q(2),2} \\ &= 1 \cdot a_{1,1} a_{2,2} + (-1) \cdot a_{2,1} a_{1,2} = ad - cb \end{aligned}$$

Beispiel

Die Determinante einer 3×3-Matrix ist eine Summe von Produkten, die den verschiedenen Diagonalen entsprechen. Dieses sogenannte Sarrus-Schema ist in der Abbildung illustriert.



$|A| =$

$$+a_{1,1}a_{2,2}a_{3,3} + a_{2,1}a_{3,2}a_{1,3} + a_{3,1}a_{1,2}a_{2,3}$$
$$-a_{3,1}a_{2,2}a_{1,3} - a_{1,1}a_{3,2}a_{2,3} - a_{2,1}a_{1,2}a_{3,3}$$

konkrete Matrix:

$$\begin{vmatrix} 8 & 1 & 6 \\ 3 & 5 & 7 \\ 4 & 9 & 2 \end{vmatrix} = (8 \cdot 5 \cdot 2 + 3 \cdot 9 \cdot 6 + 4 \cdot 1 \cdot 7)$$
$$-(4 \cdot 5 \cdot 6 + 8 \cdot 9 \cdot 7 + 3 \cdot 1 \cdot 2) = 360$$

Überprüfung des Sarrus-Schemas durch Entwicklung nach Permutationen

$$|A| = \sum_{p \in S_3} \sigma(p)\, a_{p(1),1} a_{p(2),2} a_{p(3),3}$$

$\rightsquigarrow$ 6 Summanden

p	Vertauschungen	$\sigma(p)$	Produkt
$(1,2,3)$		$+$	$a_{1,1}a_{2,2}a_{3,3}$
$(2,3,1)$	$\to (3,2,1) \to (1,2,3)$	$+$	$a_{2,1}a_{3,2}a_{1,3}$
$(3,1,2)$	$\to (2,1,3) \to (1,2,3)$	$+$	$a_{3,1}a_{1,2}a_{2,3}$
$(3,2,1)$	$\to (1,2,3)$	$-$	$a_{3,1}a_{2,2}a_{1,3}$
$(1,3,2)$	$\to (1,2,3)$	$-$	$a_{1,1}a_{3,2}a_{2,3}$
$(2,1,3)$	$\to (1,2,3)$	$-$	$a_{2,1}a_{1,2}a_{3,3}$

Die Determinante einer Matrix bleibt bei Addition eines Vielfachen einer Spalte oder Zeile zu einer anderen Spalte oder Zeile unverändert. Sie ändert ihr Vorzeichen bei Vertauschung von Spalten oder Zeilen. Analog zum Gauß-Algorithmus für die Lösung linearer Gleichungssysteme kann man mit diesen Operationen die Determinante auf Dreiecksform transformieren,

$$\begin{vmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,n} \end{vmatrix} \rightarrow \begin{vmatrix} d_{1,1} & \cdots & d_{1,n} \\ O & & \vdots \\ O & O & d_{n,n} \end{vmatrix}, \quad d_{j,k} = 0 \, \text{für} \, j > k \,,$$

und als Produkt der Diagonalelemente berechnen:

$$(-1)^{\ell} \det A = \det D = d_{1,1} \cdots d_{n,n}$$

mit ℓ der Anzahl der Spalten- bzw. Zeilen-Permutationen.

Eine Determinante ist genau dann nicht Null, wenn die Spalten (Zeilen) eine Basis bilden. Insbesondere ist die Determinante einer Matrix mit zwei gleichen Spalten oder Zeilen Null.

Es gelten die folgenden Regeln:

- $\det(AB) = (\det A)(\det B)$
- $\det A = \det A^{\mathrm{t}}$
- $\det(A^{-1}) = (\det A)^{-1}$

Beweis

(i) Transposition:

Entwicklung nach Permutationen

$$\det A = \sum_{p \in S_n} \sigma(p)\, a_{p(1),1} \cdots a_{p(n),1}$$

Umordnung der Faktoren (Spaltenindex $1, \ldots, n \to$ Spaltenindex $p^{-1}(1), \ldots, p^{-1}(n)$) $\quad \rightsquigarrow$

$$a_{p(1),1} \cdots a_{p(n),1} = a_{1,p^{-1}(1)} \cdots a_{n,p^{-1}(n)}$$

mit p^{-1} der inversen Permutation zu p

$\sigma(p) = \sigma(p^{-1}) \quad \rightsquigarrow$

$$\sum_{p^{-1}} \sigma(p^{-1})\, a^{\mathrm{t}}_{p^{-1}(1),1} \cdots a^{\mathrm{t}}_{p^{-1}(n),1} = \det A^{\mathrm{t}}\,,$$

d.h. $\det A = \det A^{\mathrm{t}}$

$\implies$ Aussagen für Spalten gelten entsprechend für Zeilen.

(ii) Addition von Spaltenvielfachen:

$$\det(\ldots, a_j + sa_k, \ldots, a_k, \ldots) = \\ \det(\ldots, a_j, \ldots, a_k, \ldots) + s\,\underbrace{\det(\ldots, a_k, \ldots, a_k, \ldots)}_{=0}$$

aufgrund der Linearität und Antisymmetrie der Determinante

(iii) Basis-Test:
Die lineare Hülle der Spalten einer Matrix A bleibt bei Addition von Spaltenvielfachen und Spaltenvertauschungen unverändert, damit auch der Rang, ebenso wie der Betrag der Determinante.
Die Spalten a_k bilden genau dann eine Basis, wenn Rang A maximal ist, d.h. mit der Dimension n von A übereinstimmt.
Damit gilt

$$\begin{aligned} \{a_1, \ldots, a_n\}\ \text{Basis} \quad &\Longleftrightarrow \quad \text{Rang}\, A = \text{Rang}\, D = n \\ &\Longleftrightarrow \quad d_{k,k} \neq 0,\ k = 1, \ldots, n \\ &\Longleftrightarrow \quad |d_{1,1} \cdots d_{n,n}| = |\det D| = |\det A| \neq 0 \end{aligned}$$

(v) Produkte von Matrizen:
$\det A = 0 \quad \Longleftrightarrow \quad a_1, \ldots, a_n$ linear abhängig
$\Longrightarrow \quad Ab_1, \ldots, Ab_n$ linear abhängig, da die Spalten Ab_k von AB Linearkombinationen der Spalten a_ℓ von A sind
$\Longrightarrow \quad \det(AB) = \det(Ab_1, \ldots, Ab_n) = 0 = \det A \det B \quad \checkmark$
für $\det A \neq 0$ definiere

$$d(b_1, \ldots, b_n) = \det(AB)/\det A$$

und verifiziere die definierenden Eigenschaften der Determinante
Antisymmetrie und Normierung unmittelbar ersichtlich
zum Beweis der Linearität bemerke für $b_k = \alpha u + \beta v$

$$\begin{aligned} d(\ldots, \alpha u + \beta v, \ldots) &= \det(\ldots, \alpha Au + \beta Av, \ldots)/\det A \\ &= \alpha d(\ldots, u, \ldots) + \beta d(\ldots, v, \ldots) \end{aligned}$$

(v) Inverse:
$AA^{-1} = E \quad \Longrightarrow$

$$1 = \det E = \det(AA^{-1}) = \det(A)\det(A^{-1})$$

Beispiel

Berechnung der Determinante von

$$A = \begin{pmatrix} 1 & 4 & 0 & 2 \\ 0 & 8 & 0 & 4 \\ 3 & 3 & 3 & 2 \\ 0 & 6 & 0 & 4 \end{pmatrix}$$

durch Transformation auf Dreiecksform mit Gauß-Operationen (Skalierung, Addition von Zeilenvielfachen und Zeilenvertauschungen)

Zeile 3 - 3 $\times$ Zeile 1 $\quad \rightsquigarrow$

$$\det A = \det \begin{pmatrix} 1 & 4 & 0 & 2 \\ 0 & 8 & 0 & 4 \\ 0 & -9 & 3 & -4 \\ 0 & 6 & 0 & 4 \end{pmatrix}$$

Vertauschen von Spalte 2 und Spalte 4 $\rightsquigarrow$

$$\det A = -\det \begin{pmatrix} 1 & 2 & 0 & 4 \\ 0 & 4 & 0 & 8 \\ 0 & -4 & 3 & -9 \\ 0 & 4 & 0 & 6 \end{pmatrix}$$

Zeile 3 + Zeile 2, Zeile 4 - Zeile 2 $\rightsquigarrow$

$$\det A = -\det \begin{pmatrix} 1 & 2 & 0 & 4 \\ 0 & 4 & 0 & 8 \\ 0 & 0 & 3 & -1 \\ 0 & 0 & 0 & -2 \end{pmatrix} = -1 \cdot 4 \cdot 3 \cdot (-2) = 24$$

Volumen eines Parallelepipeds

Der Betrag der Determinante einer reellen Matrix $A = (a_1, \ldots, a_n)$ stimmt mit dem n-dimensionalen Volumen des von den Spalten a_k von A aufgespannten Parallelepipeds überein:

$$|\det A| = \operatorname{vol}\left\{\sum_{k=1}^{n} s_k a_k :\ 0 \le s_k \le 1\right\} = \operatorname{vol}\left(A[0,1]^n\right) .$$

Beweis

Das orientierte Volumen

$$\operatorname{vol}_* A = \operatorname{sign}(\det A) \operatorname{vol}(A[0,1]^n)$$

besitzt alle drei definierenden Eigenschaften der Determinante (Multilinearität, Antisymmetrie, Normierung).

$\Longrightarrow$

$$\operatorname{vol}_* A = \det A$$

Beispiel

Alternative Berechnung der Determinante einer 3×3-Matrix $A = (u, v, w)$ mit den Spalten

$$u^{\mathrm{t}} = (2, 1, 1),\ v^{\mathrm{t}} = (1, 2, 1),\ w^{\mathrm{t}} = (1, 1, 2)$$

(i) Spatprodukt:
$\det(u, v, w) = [u, v, w] = u^{\mathrm{t}}\,(b \times c) \quad \rightsquigarrow$

$$(2, 1, 1) \left(\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} \times \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} \right) = (2, 1, 1) \begin{pmatrix} 3 \\ -1 \\ -1 \end{pmatrix} = 6 - 1 - 1 = 4$$

(ii) ε-Tensor:
$\det(u, v, w) = \sum_j \sum_k \sum_\ell \varepsilon_{j,k,\ell} u_j v_k w_\ell \quad \rightsquigarrow$

$$\begin{aligned} \det(u, v, w) &= \varepsilon_{1,2,3} \cdot 2 \cdot 2 \cdot 2 + \varepsilon_{1,3,2} \cdot 2 \cdot 1 \cdot 1 + \varepsilon_{2,1,3} \cdot 1 \cdot 1 \cdot 2 \\ &\quad + \varepsilon_{2,3,1} \cdot 1 \cdot 1 \cdot 1 + \varepsilon_{3,1,2} \cdot 1 \cdot 1 \cdot 1 + \varepsilon_{3,2,1} \cdot 1 \cdot 2 \cdot 1 \\ &= 8 - 2 - 2 + 1 + 1 - 2 = 4 \end{aligned}$$

Determinanten spezieller Matrizen

Die Determinante einiger spezieller $n \times n$-Matrizen A läßt sich unmittelbar angeben.

- Dreiecksmatrix ($a_{j,k} = 0$ für $j < k$ oder $j > k$):

$$\det A = a_{1,1} \cdots a_{n,n}$$

- Blockdiagonalmatrix (Null bis auf quadratische Diagonalblöcke $A_{k,k}$):

$$\det A = \prod_k \det A_{k,k}$$

- Unitäre und orthogonale Matrix: ($A^{-1} = A^*$):

$$|\det A| = 1$$

  bzw. $\det A \in \{-1, 1\}$ für $a_{j,k} \in \mathbb{R}$ und $A^{-1} = A^{\mathrm{t}}$

Beweis

(i) Obere Dreiecksmatrix:

Wegen $\det A = \det A^{\mathrm{t}}$ genügt es, eine obere Dreiecksmatrix zu betrachten.

Permutation p von $(1,\ldots,n)$ ungleich der Identität

$$\begin{aligned} &\implies \quad \exists k \,\mathrm{mit}\, p(k) > k \,(k = \min\{\ell :\ p(\ell \neq \ell)\}) \implies a_{p(k),k} = 0 \\ &\implies \quad \det A = \sum_{p \in S_n} \sigma(p)\, a_{p(1),1} \cdots a_{p(n),n} = a_{1,1} \cdots a_{n,n} \end{aligned}$$

(ii) Blockdiagonalmatrix:

betrachte zunächst eine Aufteilung in zwei Diagonalblöcke der Größe n_1 und n_2

$$\begin{aligned} &a_{p(1),1} \cdots a_{p(n),n} \neq 0 \quad \text{nur für} \\ &\quad p(1), \cdots, p(n_1) \leq n_1 \,\wedge\, p(n_1+1), \cdots, p(n) \geq n_1 + 1 \end{aligned}$$

$\Longrightarrow$ Produktform der Determinante:

$$\det A = \sum_{p \in S_n} \sigma(p)\,(a_{p(1),1} \cdots a_{p(n_1),n_1})(a_{p(n_1+1),n_1+1} \cdots a_{p(n),n}) =$$
$$= \Big(\sum_{q \in S_{n_1}} \sigma(q)\, a_{q(1),1} \cdots a_{q(n_1),n_1}\Big)\Big(\sum_{r \in S_{n_2}} \sigma(r)\, a_{n_1+r(1),n_1+1} \cdots a_{n_1+r(n_2),n}\Big)$$
$$= \det A_{1,1} \det A_{2,2}$$

Induktion mit Abspaltung von jeweils einem Diagonalblock $\Longrightarrow$ Aussage für eine Aufteilung in mehr als zwei Blöcke

(iii) Orthogonale und unitäre Matrix:
Vertauschbarkeit von komplexer Konjugation mit Summen und Produkten, $\det A = \det A^{\mathrm{t}}$ und $\det(AB) = \det A \det B$ $\Longrightarrow$

$$\begin{aligned} 1 &= \det E = \det(A^* A) = \det A^* \det A \\ &= \overline{\det A^{\mathrm{t}}}\det A = \overline{\det A}\det A = |\det A|^2 \end{aligned}$$

A reell $\Longrightarrow$ $\det A \in \{-1, 1\}$

Beispiel

Determinanten der Matrizen

$$A = \begin{pmatrix} 1 & 4 & 6 \\ 0 & -2 & 5 \\ 0 & 0 & 3 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 2 & 0 \\ 3 & 4 & 0 \\ 0 & 0 & 5 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & -2 & 2 \\ -2 & 1 & 2 \\ 2 & 2 & 1 \end{pmatrix}$$

A: Dreiecksmatrix $\quad\Longrightarrow\quad \det A = a_{1,1}a_{2,2}a_{3,3} = 1 \cdot (-2) \cdot 3 = -6$

B: Blockdiagonalmatrix

$$\begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} = 1 \cdot 4 - 2 \cdot 3 = -2, \quad\Longrightarrow\quad \det B = (-2) \cdot 5 = -10$$

$C/3$: orthogonal $\quad\Longrightarrow\quad |\det C| = 3^3\,|\det(C/3)| = 27$

Sarrus-Regel $\quad\rightsquigarrow$

$$\begin{vmatrix} 1 & -2 & 2 \\ -2 & 1 & 2 \\ 2 & 2 & 1 \end{vmatrix} = 1 \cdot 1 \cdot 1 + (-2) \cdot 2 \cdot 2 + 2 \cdot (-2) \cdot 2$$

$$-2 \cdot 1 \cdot 2 - 2 \cdot 2 \cdot 1 - 1 \cdot (-2) \cdot (-2) = -27$$

Entwicklung von Determinanten

Die Determinante einer $n \times n$-Matrix A lässt sich nach einer beliebigen Zeile oder Spalte entwickeln:

$$\begin{aligned} \det A &= \sum_{k=1}^{n} (-1)^{j+k} a_{j,k} \det \tilde{A}_{j,k} \quad \text{(Entwicklung nach Zeile } j\text{)} \\ &= \sum_{j=1}^{n} (-1)^{j+k} a_{j,k} \det \tilde{A}_{j,k} \quad \text{(Entwicklung nach Spalte } k\text{)}\,, \end{aligned}$$

wobei $\tilde{A}_{j,k}$ die Matrix bezeichnet, die durch Streichen der j-ten Zeile und k-ten Spalte von A entsteht,

$$A : \begin{pmatrix} a_{1,1} & \cdots & a_{1,k} & \cdots & a_{1,n} \\ \vdots & & \vdots & & \vdots \\ a_{j,1} & \cdots & \mathbf{a_{j,k}} & \cdots & a_{j,n} \\ \vdots & & \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,k} & \cdots & a_{n,n} \end{pmatrix}, \quad \tilde{A}_{j,k} : \begin{pmatrix} a_{1,1} & \vdots & \vdots & a_{1,n} \\ \cdots & a_{j-1,k-1} & a_{j-1,k+1} & \cdots \\ \cdots & a_{j+1,k-1} & a_{j+1,k+1} & \cdots \\ a_{n,1} & \vdots & \vdots & a_{n,n} \end{pmatrix}.$$

Durch wiederholte Anwendung der Prozedur kann die Dimension der Determinanten sukzessive reduziert werden.

Beweis

$\det A = \det A^{\mathrm{t}} \quad \rightsquigarrow \quad$ Entwicklung nach einer Spalte

Darstellung der k-ten Spalte als Linearkombination von Einheitsvektoren,

$$a_k = \sum_{j=1}^{n} a_{j,k} e_j \,,$$

und Multilinearität der Determinante $\quad \Longrightarrow$

$$\det A = \sum_{j=1}^{n} a_{j,k} \det \underbrace{(a_1, \ldots, a_{k-1}, e_j, a_{k+1}, \ldots, a_n)}_{=B_j}$$

$n-k$ Spaltenvertauschungen und $n-j$ Zeilenvertauschungen

$\rightsquigarrow \quad$ 1 von e_j in rechter unterer Ecke:

$$\det B_j = \underbrace{(-1)^{n-k}(-1)^{n-j}}_{(-1)^{j+k}} \det \begin{pmatrix} & & & & & & 0 \\ & & & \tilde{A}_{j,k} & & & \vdots \\ & & & & & & 0 \\ a_{j,1} & \cdots & a_{j,k-1} & a_{j,k+1} & \cdots & a_{j,n} & 1 \end{pmatrix}$$

Entwicklung nach Permutationen $\rightsquigarrow$ nichttriviale Beiträge nur von Permutationen p mit $p(n) = n$, d.h. $(p(1), \ldots, p(n-1)) \in S_{n-1}$ und

$$\det \begin{pmatrix} & & & & & & 0 \\ & & \tilde{A}_{j,k} & & & & \vdots \\ & & & & & & 0 \\ a_{j,1} & \cdots & a_{j,k-1} & a_{j,k+1} & \cdots & a_{j,n} & 1 \end{pmatrix} = \det \tilde{A}_{i,j}$$

$\rightsquigarrow$ angegebene Entwicklungsformel

Beispiel

Entwicklung der Determinante

$$d = \begin{vmatrix} 1 & 3 & 0 & 4 \\ 2 & 0 & 0 & 3 \\ 0 & 6 & 1 & 5 \\ 0 & 0 & 2 & 9 \end{vmatrix}$$

Entwicklung nach der zweiten Zeile $\rightsquigarrow$

$$(-1)^{2+1} \cdot 2 \begin{vmatrix} 3 & 0 & 4 \\ 6 & 1 & 5 \\ 0 & 2 & 9 \end{vmatrix} + 0 + 0 + (-1)^{2+4} \cdot 3 \begin{vmatrix} 1 & 3 & 0 \\ 0 & 6 & 1 \\ 0 & 0 & 2 \end{vmatrix}$$

Entwicklung der ersten Determinante nach der ersten Spalte $\rightsquigarrow$

$$(-1)^{1+1} \cdot 3 \begin{vmatrix} 1 & 5 \\ 2 & 9 \end{vmatrix} + (-1)^{2+1} \cdot 6 \begin{vmatrix} 0 & 4 \\ 2 & 9 \end{vmatrix} = -3 + 48 = 45$$

Determinante der Dreiecksmatrix: $1 \cdot 6 \cdot 2 = 12$

insgesamt: $d = -2 \cdot 45 + 3 \cdot 12 = -54$

Beispiel

Ebene durch die Punkte

$$P_k = (x_k, y_k, z_k), \quad k = 1, 2, 3$$

Ebenengleichung in Determinantenform

$$E : \begin{vmatrix} x & y & z & 1 \\ x_1 & y_1 & z_1 & 1 \\ x_2 & y_2 & z_2 & 1 \\ x_3 & y_3 & z_3 & 1 \end{vmatrix} = 0$$

Entwicklung nach der ersten Zeile $\quad \rightsquigarrow$

$$E : \, ax + by + cz = d$$

mit

$$a = \begin{vmatrix} y_1 & z_1 & 1 \\ y_2 & z_2 & 1 \\ y_3 & z_3 & 1 \end{vmatrix}, \; b = -\begin{vmatrix} x_1 & z_1 & 1 \\ x_2 & z_2 & 1 \\ x_3 & z_3 & 1 \end{vmatrix}, \; c = \begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix}, \; d = \begin{vmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{vmatrix}$$

Vandermonde-Determinante

Die Determinante der an n Punkten x_j ausgewerteten Monome $1, x, \ldots x^{n-1}$ ist

$$\begin{vmatrix} 1 & x_1 & \ldots & x_1^{n-1} \\ 1 & x_2 & \ldots & x_2^{n-1} \\ \vdots & \vdots & & \vdots \\ 1 & x_n & \ldots & x_n^{n-1} \end{vmatrix} = \prod_{j>k} (x_j - x_k)\,.$$

Beweis

Induktion

(i) $n = 2$:

$$\det \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \end{pmatrix} = x_2 - x_1$$

(ii) $(n-1) \to n$:

Subtraktion der ersten Zeile von allen weiteren Zeilen

$$\begin{vmatrix} 1 & x_1 & \dots & x_1^{n-1} \\ 1 & x_2 & \dots & x_2^{n-1} \\ \vdots & \vdots & & \vdots \\ 1 & x_n & \dots & x_n^{n-1} \end{vmatrix} = \begin{vmatrix} 1 & x_1 & \dots & x_1^{n-1} \\ 0 & x_2 - x_1 & \dots & x_2^{n-1} - x_1^{n-1} \\ \vdots & \vdots & & \vdots \\ 0 & x_n - x_1 & \dots & x_n^{n-1} - x_1^{n-1} \end{vmatrix}$$

Entwicklung nach der ersten Spalte

$$\begin{vmatrix} x_2 - x_1 & \dots & x_2^{n-1} - x_1^{n-1} \\ \vdots & & \vdots \\ x_n - x_1 & \dots & x_n^{n-1} - x_1^{n-1} \end{vmatrix}$$

Herausziehen des Faktors $x_k - x_1$ aus der k-ten Zeile

$$\left(\prod_{k=2}^{n}(x_k - x_1)\right)\begin{vmatrix} 1 & x_2 + x_1 & x_2^2 + x_2x_1 + x_1^2 & \dots & \sum\limits_{i=0}^{n-2} x_2^{n-2-i}x_1^i \\ 1 & x_3 + x_1 & x_3^2 + x_3x_1 + x_1^2 & \dots & \sum\limits_{i=0}^{n-2} x_3^{n-2-i}x_1^i \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n - x_1 & x_n^2 + x_nx_1 + x_1^2 & \dots & \sum\limits_{i=0}^{n-2} x_n^{n-2-i}x_1^i \end{vmatrix}$$

Subtraktion des x_1-fachen der vorhergehenden Spalte von jeder Spalte

$$\left(\prod_{k=2}^{n}(x_k - x_1)\right)\begin{vmatrix} 1 & x_2 & \dots & x_2^{n-2} \\ 1 & x_3 & \dots & x_3^{n-2} \\ \vdots & \vdots & & \vdots \\ 1 & x_n & \dots & x_n^{n-2} \end{vmatrix}$$

Induktionsvoraussetzung $\quad\Longrightarrow\quad |\cdots| = \prod_{j>k\geq 2}(x_j - x_k)$

Lineares Gleichungssystem

Ein lineares Gleichungssystem hat die Form

$$\begin{array}{ccccccc} a_{1,1}x_1 & + & \cdots & + & a_{1,n}x_n & = & b_1 \\ \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots \\ a_{m,1}x_1 & + & \cdots & + & a_{m,n}x_n & = & b_m \end{array} \quad \Longleftrightarrow \quad Ax = b$$

mit einer Koeffizientenmatrix $A = (a_{j,k})$, zu bestimmenden Unbekannten x_k und einer rechten Seite b.

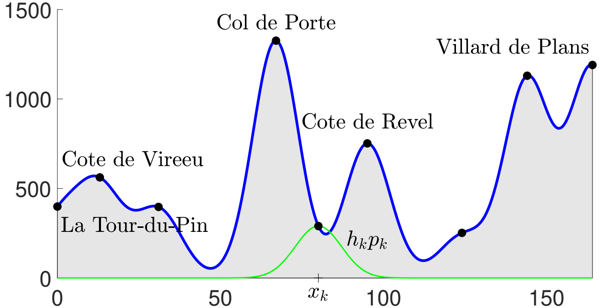
Das lineare Gleichungssystem ist homogen, wenn $b = (0, \ldots, 0)^{\mathrm{t}}$, andernfalls bezeichnet man es als inhomogen.

Besitzt das lineare Gleichungssystem keine Lösung (im Allgemeinen für $m > n$), so nennt man es überbestimmt. Man spricht in diesem Fall auch von einem Ausgleichsproblem. Ein lineares Gleichungssystem mit keiner eindeutigen Lösung (im Allgemeinen für $m < n$) wird als unterbestimmt bezeichnet.

Beispiel

Approximation eines Höhenprofils einer Etappe der Tour-de-France aus den Daten

x_j	0	13	31	67	80	95	124	144	164
h_j	399	562	397	1326	290	752	252	1130	1190



Ansatz

$$h(x) \approx p(x) = \sum_{k=1}^{n} c_k p_k(x)$$

mit Exponentialfunktionen

$$p_k(x) = \exp(-(x - x_k)^2/100)$$

als Basisfunktionen

Vorteil der Basiswahl: starkes Abklingen von p_k für $|x - x_k| \to \infty$ $\quad \rightsquigarrow$ keine globale Auswirkung von Änderungen

Interpolationsbedingungen

$$h_j = p(x_j) = \sum_{k=1}^{n} c_k p_k(x_j), \quad j = 1, \ldots, n$$

$\iff$ lineares Gleichungssystem

$$Ac = b$$

mit $a_{j,k} = p_k(x_j)$ und $b_j = h_j$

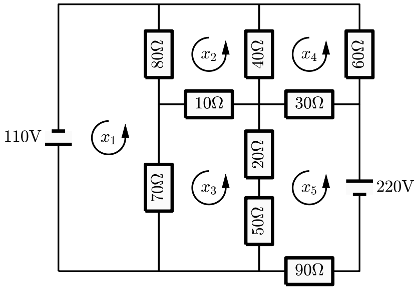
Beispiel

Elektrischer Schaltkreis

x_k: Kreisströme mit Fließrichtung entgegen dem Uhrzeigersinn

$R_{j,k}$: gemeinsamer Widerstand der j-ten und k-ten Schleife

U_j: angelegte Spannungen



Ohmsches und Kirchhoffsches Gesetz $\rightsquigarrow$ lineares Gleichungssystem

$$\sum_{j\sim 0} x_j R_{j,0} + \sum_{j\sim k} (x_j - x_k) R_{j,k} = U_j$$

$j \sim k$: j-te und k-te Schleife haben einen gemeinsamen Widerstand durchflossen vom Strom $x_j - x_k$

$R_{j,0}$, $j \sim 0$: Widerstände, die nur in der j-ten Schleife liegen

Daten in der Abbildung $\rightsquigarrow$

$$\begin{pmatrix} 150 & -70 & -80 & 0 & 0 \\ -70 & 120 & -10 & -40 & 0 \\ -80 & -10 & 160 & 0 & -70 \\ 0 & -40 & 0 & 130 & -30 \\ 0 & 0 & -70 & -30 & 190 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 110 \\ 0 \\ 0 \\ 0 \\ -220 \end{pmatrix}$$

Diagonale: Summe der zu einer Schleife gehörigen Widerstände

$a_{j,k}$: negativer gemeinsamer Widerstand der Schleifen j und k

Lösung

$$x = \begin{pmatrix} 1.0157 \\ 0.5641 \\ 0.0358 \\ -0.0940 \\ -1.1595 \end{pmatrix}$$

Lösungsmenge eines linearen Gleichungssystems

Die Lösungsmenge eines homogenen linearen Gleichungssystems

$$Ax = (0, \ldots, 0)^{\mathrm{t}},$$

mit einer $m \times n$-Koeffizientenmatrix A ist ein Unterraum U des Vektorraums der n-Tupel $(x_1, \cdots, x_n)^{\mathrm{t}}$, $U = \operatorname{Kern} A$.

Besitzt das inhomogene lineare Gleichungssystem

$$Ax = b$$

eine Lösung v, so gilt für die allgemeine Lösung

$$x \in v + U,$$

d.h. die Lösungsmenge ist ein affiner Unterraum.

Insbesondere kann also ein inhomogenes lineares Gleichungssystem entweder keine ($\nexists v$), eine ($U = \{0\}$) oder unendlich viele ($\dim U > 0$) Lösungen besitzen.

Beweis

(i) Für Lösungen x, y des homogenen linearen Gleichungssystems gilt

$$A(x+y) = Ax + Ay = (0,\dots,0)^{\mathrm{t}} + (0,\dots,0)^{\mathrm{t}} = (0,\dots,0)^{\mathrm{t}}$$
$$A(sx) = sAx = s(0,\dots,0)^{\mathrm{t}} = (0,\dots,0)^{\mathrm{t}},$$

d.h. die Lösungsmenge des homogenen linearen Gleichungssystems bildet einen Unterraum U.

(ii) Für eine Lösung v des inhomogenen linearen Gleichungssystems $Ax = b$ gilt

$$A(x-v) = Ax - Av = b - b = (0,\dots,0)^{\mathrm{t}},$$

d.h. $u = x - v \in U$.

$\rightsquigarrow$ Lösungsmenge $v + U$ (affiner Unterraum)

Beispiel

Verschiedene Typen linearer Gleichungssysteme

(i) Eindeutige Lösung:

$$\begin{pmatrix} 0 & 2 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 4 \\ 5 \end{pmatrix} \quad \Longleftrightarrow \quad \begin{array}{ccccc} & + & 2x_2 & = & 4 \\ x_1 & + & 3x_2 & = & 5 \end{array}$$

$\rightsquigarrow \quad x_2 = 2,\ x_1 = -1$

(ii) Unendlich viele Lösungen:

$$\begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \Longleftrightarrow \quad \begin{array}{ccccccc} & & x_2 & + & x_3 & = & 1 \\ x_1 & + & x_2 & & & = & 1 \end{array}$$

$x_2 = t \implies x_3 = 1 - t$, $x_1 = 1 - t$, d.h. $x = \underbrace{(1, 0, 1)^{\mathrm{t}}}_{v} + \underbrace{t(-1, 1, -1)^{\mathrm{t}}}_{\in U}$

U: Lösungsraum des homogenen Gleichungssytems (dim $U = 1$, Gerade mit Richtung $(-1, 1, -1)^{\mathrm{t}}$)

(iii) Keine Lösung:

$$\begin{pmatrix} 0 & 1 \\ 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad \Longleftrightarrow \quad \begin{array}{ccccc} & & x_2 & = & 0 \\ x_1 & + & x_2 & = & 1 \\ x_1 & & & = & 0 \end{array}$$

Gleichungen 1 und 3, $x_2 = 0$ und $x_1 = 0$, sind inkonsistent zu Gleichung 2.

Cramersche Regel

Für ein quadratisches lineares Gleichungssystem $Ax = b$ ist

$$x_j \det A = \det(a_1, \ldots, a_{j-1}, b, a_{j+1}, \ldots, a_n)\,,$$

wobei $a_1, \ldots, a_n$ die Spalten der Koeffizientenmatrix A bezeichnen.

Ist $\det A \neq 0$, so existiert eine eindeutige Lösung x, deren Komponenten x_j als Quotienten der Determinanten in obiger Gleichung bestimmt werden können.

Wählt man b als den k-ten Einheitsvektor e_k, so erhält man die k-te Spalte $(c_{1,k}, \ldots, c_{n,k})^{\mathrm{t}}$ der Inversen $C = A^{-1}$:

$$c_{j,k} = \frac{\det(a_1, \ldots, a_{j-1}, e_k, a_{j+1}, \ldots, a_n)}{\det A}\,.$$

Nach dem Entwicklungsatz für Determinanten ist

$$\det(a_1, \dots, a_{j-1}, e_k, a_{j+1}, \dots, a_n) = (-1)^{j+k} \det \tilde{A}_{j,k}$$

mit $\tilde{A}_{j,k}$ der Matrix, die durch Streichen der Spalte j und der Zeile k von A entsteht.

Beweis

Multilinearität und Antisymmetrie der Determinante $\quad\Longrightarrow$

$$\begin{aligned}
\det(a_1,\ldots,a_{k-1},b,a_{k+1},\ldots,a_n) &= \det(a_1,\ldots,a_{k-1},\sum_{j=1}^{n} a_j x_j, a_{k+1},..,a_n) \\
&= \sum_{j=1}^{n} x_j \det(a_1,\ldots,a_{k-1},a_j,a_{k+1},..,a_n) \\
&= x_k \det(a_1,\ldots,a_{k-1},a_k,a_{k+1},\ldots,a_n)
\end{aligned}$$

$b = e_k \quad \rightsquigarrow \quad k$-te Spalte $x = (c_{1,k},\ldots,c_{n,k})^{\mathrm{t}}$ der Inversen, denn

$$AC = E = (e_1,\ldots,e_n) \quad \Longrightarrow \quad Ax = e_k$$

Beispiel

Berechnung der Inversen C der (2×2)-Matrix

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

Cramersche Regel $\quad \rightsquigarrow$

$$c_{1,1} = \frac{\begin{vmatrix} 1 & b \\ 0 & d \end{vmatrix}}{\det A} = \frac{d}{\det A}, \qquad c_{1,2} = \frac{\begin{vmatrix} 0 & b \\ 1 & d \end{vmatrix}}{\det A} = \frac{-b}{\det A},$$

$$c_{2,1} = \frac{\begin{vmatrix} a & 1 \\ c & 0 \end{vmatrix}}{\det A} = \frac{-c}{\det A}, \qquad c_{2,2} = \frac{\begin{vmatrix} a & 0 \\ c & 1 \end{vmatrix}}{\det A} = \frac{a}{\det A}$$

bzw. mit $\det A = ad - cb$

$$A^{-1} = \frac{1}{ad - cb} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

Beispiel

Lineares Gleichungssystem

$$\underbrace{\begin{pmatrix} 4 & -1 & -1 \\ -3 & -4 & 4 \\ -2 & 2 & -1 \end{pmatrix}}_{A} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \underbrace{\begin{pmatrix} 8 \\ 5 \\ -7 \end{pmatrix}}_{b}$$

eindeutige Lösung, da

$$\begin{aligned} \det A &= \begin{vmatrix} 4 & -1 & -1 \\ -3 & -4 & 4 \\ -2 & 2 & -1 \end{vmatrix} \\ &\underset{\substack{\text{Entwickeln} \\ \text{nach Zeile 1}}}{=} 4\begin{vmatrix} -4 & 4 \\ 2 & -1 \end{vmatrix} + \begin{vmatrix} -3 & 4 \\ -2 & -1 \end{vmatrix} - \begin{vmatrix} -3 & -4 \\ -2 & 2 \end{vmatrix} \\ &= -16 + 11 + 14 = 9 \neq 0 \end{aligned}$$

Cramersche Regel $\quad\Longrightarrow$

$$x_1 = \det(b, a_2, a_3) / \det A \text{ mit } a_k \text{ den Spalten von } A$$

Einsetzen $\quad\rightsquigarrow$

$$\begin{aligned}
x_1 &= \frac{1}{9}\begin{vmatrix} 8 & -1 & -1 \\ 5 & -4 & 4 \\ -7 & 2 & -1 \end{vmatrix} \\
&= \frac{1}{9}\left(8\begin{vmatrix} -4 & 4 \\ 2 & -1 \end{vmatrix} - (-1)\begin{vmatrix} 5 & 4 \\ -7 & -1 \end{vmatrix} + (-1)\begin{vmatrix} 5 & -4 \\ -7 & -2 \end{vmatrix}\right) \\
&= \frac{-32 + 23 + 18}{9} = 1
\end{aligned}$$

analog, $x_2 = \det(a_1, b, a_3) / \det A$ und $x_3 = \det(a_1, a_2, b) / \det A$, d.h.

$$x_2 = \frac{1}{9}\begin{vmatrix} 4 & 8 & -1 \\ -3 & 5 & 4 \\ -2 & -7 & -1 \end{vmatrix} = -3, \quad x_3 = \frac{1}{9}\begin{vmatrix} 4 & -1 & 8 \\ -3 & -4 & 5 \\ -2 & 2 & -7 \end{vmatrix} = -1$$

Rückwärtseinsetzen

Bei einem linearen Gleichungssystem in oberer Dreiecksform,

$$\underbrace{\begin{pmatrix} r_{1,1} & \cdots & r_{1,n} \\ & \ddots & \vdots \\ 0 & & r_{n,n} \end{pmatrix}}_{R} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix},$$

mit $\det R = r_{1,1} \cdots r_{n,n} \neq 0$ können die Unbekannten $x_n, \ldots, x_1$ nacheinander bestimmt werden:

$$r_{n,n} x_n = b_n \implies x_n = b_n / r_{n,n}$$

und, für $\ell = n-1, \ldots 1$,

$$r_{\ell,\ell} x_\ell + \cdots + r_{\ell,n} x_n = b_\ell \implies x_\ell = \left(b_\ell - r_{\ell,\ell+1} x_{\ell+1} - \cdots - r_{\ell,n} x_n\right) / r_{\ell,\ell}\,.$$

Dabei werden jeweils die schon berechneten Werte $x_{\ell+1}, \ldots, x_n$ verwendet.

Bei einem linearen Gleichungssystem mit einer unteren Dreiecksmatrix kann man analog nacheinander $x_1, \ldots, x_n$ bestimmen.

Die Berechnungen können mit Hilfe eines Tableaus $R|b$ erfolgen, in dem die Koeffizientenmatrix und die rechte Seite zusammengefasst sind. Für $\ell = n, \ldots, 1$ dividiert man die Zeile ℓ durch $r_{\ell,\ell}$ ($\rightsquigarrow$ Diagonalelement gleich 1) und zieht für $k = \ell - 1, \ldots, 1$ das $r_{k,\ell}$-fache der ℓ-ten Zeile von der k-ten Zeile ab. Dadurch werden oberhalb von $r_{\ell,\ell} = 1$ Nullen erzeugt. Als Resultat enthält das Tableau die Einheitsmatrix und in der letzten Spalte (modifizierte rechte Seite b) die Lösung x.
Bei der praktischen Durchführung werden nur die sich ändernden Zeilen untereinander notiert.

Beispiel

Lösen des Gleichungssystems

$$\begin{array}{rcrcrcl} 4x_1 & + & 3x_2 & + & x_3 & = & 6 \\ & & 2x_2 & + & 2x_3 & = & 0 \\ & & & & 7x_3 & = & 7 \end{array}$$

Rückwärtseinsetzen $\quad \rightsquigarrow$

$$\begin{aligned} x_3 &= 7/7 = 1 \\ x_2 &= (0 - 2x_3)/2 = (0 - 2 \cdot 1)/2 = -1 \\ x_1 &= (6 - 3x_2 - x_3)/4 = (6 - 3(-1) - 1)/4 = 2 \end{aligned}$$

alternative Lösung mit Hilfe des Tableaus $R|b$

$$
\begin{array}{ccc|c||l}
4 & 3 & 1 & 6 & Z1 \\
0 & 2 & 2 & 0 & Z2 \\
0 & 0 & 7 & 7 & Z3 \\
\hline
0 & 0 & \boxed{1} & 1 & Z3' = Z3/2 \\
4 & 3 & 0 & 5 & Z1' = Z1 - Z3' \\
0 & 2 & 0 & -2 & Z2' = Z2 - 2 * Z3' \\
\hline
0 & \boxed{1} & 0 & -1 & Z2'' = Z2'/2 \\
4 & 0 & 0 & 8 & Z1'' = Z1' - 3 * Z2'' \\
\hline
\boxed{1} & 0 & 0 & 2 & Z1''' = Z1''/4
\end{array}
$$

markierte Zeilen $Z1'''$, $Z2''$, $Z3'$ (Diagonalelement gleich 1) $\rightsquigarrow$ Lösung

$$x_3 = 1,\ x_2 = -1,\ x_1 = 2$$

Durch Gauß-Transformationen lässt sich ein lineares Gleichungssystem mit invertierbarer $n \times n$-Koeffizientenmatrix A in maximal $n-1$ Schritten auf obere Dreiecksform bringen.
Dazu werden sukzessive die Koeffizienten unterhalb der Diagonalen annulliert, d.h. nach $\ell - 1$ Schritten hat das lineare Gleichungssystem die Form

$$
\begin{array}{ccccccccccccc}
a_{1,1}x_1 & + & a_{1,2}x_2 & + & \dots & + & a_{1,\ell}x_\ell & + & \dots & + & a_{1,n}x_n & = & b_1 \\
 & & a_{2,2}x_2 & + & \dots & + & a_{2,\ell}x_\ell & + & \dots & + & a_{2,n}x_n & = & b_2 \\
 & & & & \ddots & & & \vdots & & \vdots & & \vdots & \\
 & & & & & & a_{\ell,\ell}x_\ell & + & \dots & + & a_{\ell,n}x_n & = & b_\ell \\
 & & & & & & a_{\ell+1,\ell}x_\ell & + & \dots & + & a_{\ell+1,n}x_n & = & b_{\ell+1} \\
 & & & & & & & \vdots & & \vdots & & \vdots & \\
 & & & & & & a_{n,\ell}x_\ell & + & \dots & + & a_{n,n}x_n & = & b_n
\end{array}
$$

Im einzelnen verläuft der ℓ-te Eliminationsschritt wie folgt.

- Aus den Koeffizienten $a_{\ell,\ell}, \ldots, a_{n,\ell}$ wird ein von Null verschiedener Koeffizient $a_{i,\ell}$, das sogenannte Pivot-Element, ausgewählt, und die ℓ-te mit der i-ten Gleichung vertauscht.
- Für $j > \ell$ wird die j-te Gleichung durch eine Linearkombination der j-ten und ℓ-ten Gleichung ersetzt, um den Term mit der Unbekannten x_ℓ zu eliminieren ($a_{j,\ell} \to 0$):

  $$a_{j,k} \leftarrow \alpha_j a_{j,k} + \beta_j a_{\ell,k},\ k = \ell, \ldots, n, \quad b_j \leftarrow \alpha_j b_j + \beta_j b_\ell$$

  Dabei sind α_j und β_j so gewählt, dass $\alpha_j a_{j,\ell} + \beta_j a_{\ell,\ell} = 0$ ($\to a_{j,k} = 0$ für $k = \ell$, Nullen unterhalb des Pivot-Elements $a_{\ell,\ell}$)

Eine kanonische Wahl bei den Eliminationsschritten ist $\alpha_j = 1$, $\beta_j = -a_{j,\ell}/a_{\ell,\ell}$; jedoch kann, um gegebenenfalls Brüche zu vermeiden, auch eine andere Wahl getroffen werden, z.B. $\alpha_j = a_{\ell,\ell}$, $\beta_j = -a_{j,\ell}$.

Üblicherweise werden für die Eliminationsschritte die Matrix A und die rechte Seite b zu einem Tableau $A|b$ zusammengefasst. Die modifizierten Zeilen werden dann untereinander aufgelistet und die Pivot-Elemente markiert. Zur besseren Erläuterung können die Faktoren α_j und β_j in einer zusätzlichen Spalte notiert werden. Die resultierende Dreiecksform besteht dann aus den markierten Pivot-Zeilen.

Das Schema kann zur simultanen Behandlung mehrerer rechter Seiten benutzt werden. Insbesondere kann so mit $b = (e_1, \ldots, e_n)$ die Inverse der Matrix A bestimmt werden.

Beispiel

Gauß-Elimination für das lineare Gleichungssystem

$$\begin{array}{rcrcrcrcr} & & 2x_2 & + & x_3 & - & x_4 & = & -1 \\ 3x_1 & + & 2x_2 & & & + & x_4 & = & 5 \\ 3x_1 & + & x_2 & - & 2x_3 & + & x_4 & = & 3 \\ 6x_1 & + & 4x_2 & - & x_3 & + & x_4 & = & 7 \end{array}$$

Matrix und rechte Seite

$$A = \begin{pmatrix} 0 & 2 & 1 & -1 \\ 3 & 2 & 0 & 1 \\ 3 & 1 & -2 & 1 \\ 6 & 4 & -1 & 1 \end{pmatrix}, \quad b = \begin{pmatrix} -1 \\ 5 \\ 3 \\ 7 \end{pmatrix}$$

Ablauf des Gauß-Algorithmus mit Hilfe des Tableaus $A|b$ (erste vier Zeilen des folgenden Schemas)

A				b	α, β		
0	2	1	−1	−1	1		
$\boxed{\mathbf{3}}$	**2**	**0**	**1**	**5**	0	−1	−2
3	1	−2	1	3		1	
6	4	−1	1	7			1
0	2	1	−1	−1	1		
0	$\boxed{\mathbf{-1}}$	**−2**	**0**	**−2**	2	0	
0	0	−1	−1	−3		1	
0	0	−3	−1	−5	1		
0	0	$\boxed{\mathbf{-1}}$	**−1**	**−3**	−3		
0	0	0	$\boxed{\mathbf{2}}$	**4**			

Schritt 1:

Wahl des Pivot-Elements $\boxed{3}$ in Spalte 1

$\rightarrow$ Zeile 1 der Dreiecksform

Modifikation der (nicht-pivot) Zeilen 1,3,4, z.B.

Zeile 4 $\leftarrow$ 1 * Zeile 4 + (-2) * Zeile 2

(Koeffizienten $\alpha = 1$, $\beta = -2$, notiert im rechten Block des Schemas)

drei modifizierte Zeilen $\rightarrow$ nächstes Tableau

Schritt 2:

Wahl des Pivot-Elements $\boxed{-1}$

$\rightarrow$ Zeile 2 der Dreiecksform

Modifikation der (nicht-pivot) Zeilen, z.B.

Zeile 1 $\leftarrow$ 1 * Zeile 1 + 2 * Zeile 2

(Koeffizienten notiert im rechten Block des Schemas)

zwei modifizierte Zeilen $\rightarrow$ nächstes Tableau

Schritt 3:

Pivot-Element -1, nur eine Zeile zu modifizieren, Linearkombination mit der Pivot-Zeile mit den Koeffizienten 1 und -3

Die vier Pivot-Zeilen bilden die Matrix und rechte Seite der resultierenden Dreiecksform $Rx = b$.

Rückwärtseinsetzen beginnend mit dem Tableau $R|b$ (erste vier Zeilen)

$$
\begin{array}{cccc|l}
3 & 2 & 0 & 1 & 5 \\
0 & -1 & -2 & 0 & -2 \\
0 & 0 & -1 & -1 & -3 \\
0 & 0 & 0 & 2 & 4 \\
\hline
0 & 0 & 0 & \boxed{1} & 2 = x_4 \\
3 & 2 & 0 & 0 & 3 \\
0 & -1 & -2 & 0 & -2 \\
0 & 0 & -1 & 0 & -1 \\
\hline
0 & 0 & \boxed{1} & 0 & 1 = x_3 \\
3 & 2 & 0 & 0 & 3 \\
0 & -1 & 0 & 0 & 0 \\
\hline
0 & \boxed{1} & 0 & 0 & 0 = x_2 \\
3 & 0 & 0 & 0 & 3 \\
\hline
\boxed{1} & 0 & 0 & 0 & 1 = x_1
\end{array}
$$

Bei jeder der vier Tableau-Modifikationen wird zunächst die jeweils letzte Zeile durch das Diagonalelement (Pivot-Element bei dem entsprechenden Gauß-Eliminations-Schritt) dividiert ($\rightarrow$ eingerahmte 1). Die resultierende Zeile enthält als letzten Eintrag eine Komponente der Lösung und wird nun von den anderen Zeilen abgezogen, z.B.

$$(3\,2\,0\,1|5) - (0\,0\,0\,\boxed{1}|2) = (3\,2\,0\,0|3)\,.$$

Zeilenstufenform

Die Zeilenstufenform oder Echelon-Form D einer $m \times n$-Matrix A ist eine Verallgemeinerung der Dreiecksform quadratischer Matrizen:

$$D = \begin{pmatrix} 0\ldots0 & \boxed{d_{1,i_1}} & *\ldots* & & & & * \\ & & \ddots & & & & \\ & & 0\ldots0 & \boxed{d_{1,i_r}} & *\ldots* & & \\ & & & & 0\ldots0 & & \\ & & & & \ddots & & \\ 0 & & & & & & 0\ldots0 \end{pmatrix}.$$

Für die ersten von Null verschiedenen Elemente $p_j = d_{j,i_j}$ (Pivots) der Zeilen muss $i_{j+1} > i_j$ gelten, d.h. die Anzahl der führenden Nullen nimmt sukzessive um mindestens 1 zu. Insbesondere bilden die $m - r$ Nullzeilen (wenn vorhanden, d.h., wenn $r < m$) den unteren Block der Matrix D. Die Anzahl r der Pivots ist der Rang von D.

Bei der reduzierten Zeilenstufenform wird zusätzlich verlangt, dass

$$d_{j,i_j} = 1,\ d_{k,i_j} = 0 \,\text{für}\, k \neq j, \quad j = 1, \ldots, r\,,$$

d.h. außer der 1 in Position (j, i_j) enthält die Pivot-Spalte i_j nur Nullen.

Analog zur Gauß-Elimination für quadratische Systeme lässt sich ein lineares Gleichungssystem $Ax = b$ mit einer $m \times n$-Koeffizientenmatrix durch Vertauschung von Gleichungen und Addition von Gleichungsvielfachen auf Zeilenstufenform transformieren:

$$Ax = b \quad \to \quad Dx = c\,.$$

Der ℓ-te Transformationsschritt verläuft wie folgt:

- Ein von Null verschiedenes Element der Koeffizientenmatrix mit Zeilenindex $k \geq \ell$ und kleinstem Spaltenindex i_ℓ wird als Pivot gewählt und die Gleichung k mit der Gleichung ℓ vertauscht.
- Existiert kein Pivot-Element, so ist die Zeilenstufenform erreicht.
- Andernfalls werden in den Gleichungen $\ell + 1, \ldots, m$ durch Bilden von Linearkombinationen mit der ℓ-ten Gleichung,

  $$\text{Gleichung}\, j \leftarrow \alpha_j * \text{Gleichung}\, j \,+\, \beta_j * \text{Gleichung}\, \ell, \quad j = \ell + 1, \ldots, m\,,$$

  die Terme mit der i_ℓ-ten Unbekannten annulliert (Koeffizient von x_{i_ℓ} in Gleichung $j \to 0$ für $j > \ell$).

Bei Transformation auf reduzierte Zeilenstufenform werden die Linearkombinationen auch für $j = 1, \ldots, \ell - 1$ gebildet, nachdem zuvor die Gleichung ℓ durch das Pivot-Element dividiert wurde (Koeffizient von x_{i_ℓ} $\to 1$).

Analog zur Gauß-Elimination werden die Modifikationen des Gleichungssystems mit Hilfe eines Tableaus durchgeführt,

$$A|b \to D|c\,,$$

wobei die sich ändernden Koeffizienten und Elemente der rechten Seite zeilenweise aufgelistet werden.

Beispiel

Transformation des Gleichungssystems

$$\begin{array}{rrrrrrrcr}
 & & x_3 & +5x_4 & -3x_5 & & +9x_7 & = & -1 \\
2x_1 & +3x_2 & & -x_4 & +4x_5 & +x_6 & -2x_7 & = & 2 \\
4x_1 & +6x_2 & +x_3 & +3x_4 & +5x_5 & +2x_6 & +5x_7 & = & 3 \\
-2x_1 & -3x_2 & +x_3 & +6x_4 & -7x_5 & -x_6 & +6x_7 & = & 0 \\
 & & x_3 & +5x_4 & -3x_5 & & +9x_7 & = & 4
\end{array}$$

auf Zeilenstufenform

Modifikation des Tableaus $A|b$ aus Koeffizientenmatrix und rechter Seite (erste 5 Zeilen des nachfolgenden Schemas) mit Gauß-Operationen (Vertauschung von Zeilen und Addition von Zeilenvielfachen)

2 modifierte Tableaus mit 4 und 3 Zeilen im Schema für das betrachtete Beispiel

			A				b	α, β			
0	0	1	5	−3	0	9	−1	1			
$\boxed{\mathbf{2}}$	**3**	**0**	**−1**	**4**	**1**	**−2**	**2**	0	−2	1	0
4	6	1	3	5	2	5	3		1		
−2	−3	1	6	−7	−1	6	0			1	
0	0	1	5	−3	0	9	4				1
0	**0**	$\boxed{\mathbf{1}}$	**5**	**−3**	**0**	**9**	**−1**	−1	−1	−1	
0	0	1	5	−3	0	9	−1	1			
0	0	1	5	−3	0	4	2		1		
0	0	1	5	−3	0	9	4			1	
0	0	0	0	0	0	0	0				
0	**0**	**0**	**0**	**0**	**0**	$\boxed{\mathbf{-5}}$	**3**				
0	0	0	0	0	0	0	5				

Schritt 1:
Wahl des Pivot-Elements (ungleich Null, kleinster Spaltenindex)
$\rightsquigarrow$ fett gedruckte Zeile 2
Ersetzen der Zeilen 1,3,4,5 jeweils durch eine Linearkombination mit der Pivot-Zeile mit Koeffizienten α, β in den entsprechenden 4 Spalten rechts im Schema, z.B.

$$\text{Zeile}\,3 \leftarrow (-2) * \text{Zeile } 2 + (1) * \text{Zeile}\,3$$

$\rightsquigarrow$ nächster Block (4 Zeilen) des Schemas (Pivot-Zeile 2 wird nicht nochmals gelisted)
Schritt 2:
Zeile 1 $\rightsquigarrow$ Pivot-Zeile (fett)
Abziehen der Pivot-Zeile von den anderen 3 Zeilen (Koeffizienten $\alpha = -1, \beta = 1$)
$\rightsquigarrow$ letzter Block (3 Zeilen)
Schritt 3:
keine weiteren Additionen von Zeilenvielfachen, da außer der Pivot-Zeilen nur Null-Zeilen vorhanden sind
$\rightsquigarrow$ Zeilenstufenform bestehend aus den Pivot- und Nullzeilen

resultierendes Gleichungssystem

$$\begin{pmatrix} \boxed{2} & 3 & 0 & -1 & 4 & 1 & -2 \\ 0 & 0 & \boxed{1} & 5 & -3 & 0 & 9 \\ 0 & 0 & 0 & 0 & 0 & 0 & \boxed{\text{-5}} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \end{pmatrix} = \begin{pmatrix} 2 \\ -1 \\ 3 \\ 0 \\ 5 \end{pmatrix}$$

Koeffizientenmatrix mit Rang gleich 3 (Anzahl der Pivots)
keine Lösung, da fünftes Element der rechten Seite ungleich Null

Lösungsmenge eines linearen Gleichungssystems in Zeilenstufenform

Ein lineares Gleichungssystem $Dx = c$ in Zeilenstufenform,

$$\begin{pmatrix} 0\ldots 0 & \boxed{p_1} & *\ldots * & & & & * \\ & & \ddots & & & & \\ & & 0\ldots 0 & \boxed{p_r} & *\ldots * & & \\ & & & & 0\ldots 0 & & \\ & & & & \ddots & & \\ 0 & & & & & & 0\ldots 0 \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} c_1 \\ \vdots \\ c_r \\ c_{r+1} \\ \vdots \\ c_m \end{pmatrix},$$

mit Pivots $p_j = d_{j,i_j}$ ist genau dann lösbar, wenn $c_{r+1} = \cdots = c_m = 0$. Die Lösung x kann durch Rückwärtseinsetzen bestimmt werden. Sie ist eindeutig, falls $r = n$, d.h., wenn $r = \operatorname{Rang} D$ maximal ist.

Für $r < n$ gibt es $n - r$ linear unabhängige Lösungen u_k des homogenen linearen Gleichungssystems ($c_1 = \cdots = c_m = 0$), d.h. $\dim \operatorname{Kern} D = n - r$. Die Unbekannten, die den Spalten ohne Pivots entsprechen, können frei gewählt werden.

Die allgemeine Lösung des inhomogenen Systems hat die Form

$$x = x_* + \sum_{k=1}^{n-r} s_k u_k$$

mit x_* einer beliebigen speziellen Lösung von $Dx = c$.

Beispiel

Typische Fälle von Gleichungssystemen $Dx = c$ in Zeilenstufenform

(i) $m = n = 2$:

$$\begin{pmatrix} \boxed{1} & 3 \\ 0 & \boxed{2} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 5 \\ 4 \end{pmatrix}$$

$r = 2$ $\boxed{\text{Pivots}}$ $\implies$ Rang $D = 2$ maximal, eindeutige Lösung
Rückwärtseinsetzen $\rightsquigarrow$ $x = (-1, 2)^{\mathrm{t}}$

(ii) $m = 3 < n = 4$:

$$\begin{pmatrix} \boxed{1} & 2 & 0 & 3 \\ 0 & 0 & \boxed{1} & 4 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 5 \end{pmatrix}$$

reduzierte Zeilenstufenform (Pivot-Spalten $=$ Einheitsvektoren)
$r = 2 < m$ $\boxed{\text{Pivots}}$, $c_{r+1} = 5 \neq 0$ (inkompatibel mit der Nullzeile) $\implies$ keine Lösung

(iii) $m = 4 < n = 5$:

$$\begin{pmatrix} \boxed{1} & 1 & 0 & 3 & 0 \\ 0 & \boxed{2} & 3 & 0 & 5 \\ 0 & 0 & 0 & \boxed{2} & 4 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 5 \\ 0 \\ 4 \\ 0 \end{pmatrix}$$

$r = 3 < n$ $\boxed{\text{Pivots}}$, $c_{r+1} = 0 \quad \Longrightarrow \quad$ lösbar
kanonische Lösung: nur die Unbekannten x_1, x_2, x_4 zu den Pivot-Spalten ungleich Null
$\rightsquigarrow \quad x_* = (-1, 0, 0, 2, 0)^{\mathrm{t}}$
Rang $D = 3 = n - 2 \quad \Longrightarrow$
zwei linear unabhängige Lösungen u_k des homogenen Systems
$\rightsquigarrow \quad$ allgemeine Lösung:

$$x = x_* + s_1 u_1 + s_2 u_2, \quad s_k \in \mathbb{R}$$

Bestimmung von u_k durch kanonische Wahl der nicht zu den Pivot-Spalten gehörigen Unbekannten x_3 und x_5 und Berechnung der restlichen Unbekannten als Lösungen des homogenen Systems $Dx = (0, \dots, 0)^{\mathrm{t}}$:

$$\begin{aligned} x_3 = 1, x_5 = 0 \quad &\rightsquigarrow \quad u_1 = (3/2, -3/2, 1, 0, 0)^{\mathrm{t}} \\ x_3 = 0, x_5 = 1 \quad &\rightsquigarrow \quad u_2 = (-17/2, -5/2, 0, -2, 1)^{\mathrm{t}} \end{aligned}$$

Eigenwert und Eigenvektor

Gilt für eine quadratische Matrix

$$Av = \lambda v, \quad v \neq (0, \ldots, 0)^{\mathrm{t}},$$

so bezeichnet man den Skalar λ als Eigenwert von A und den Vektor v als einen Eigenvektor zu λ.

Die Eigenvektoren zu einem Eigenwert bilden zusammen mit dem Nullvektor einen Unterraum, den so genannten Eigenraum

$$V_\lambda = \mathrm{Kern}(A - \lambda E)$$

von λ (E bezeichnet die Einheitsmatrix). Die mögliche Eigenvektoren v zu λ kann man somit als nichttriviale Lösungen des linearen Gleichungssystems

$$(A - \lambda E)v = (0, \ldots, 0)^{\mathrm{t}}$$

bestimmen.

Beispiel

Eigenwerte und -Vektoren der Matrix

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

(i) Eigenwert $\lambda = 0$:

$$\begin{pmatrix} 1 & 2 & -1 \\ 0 & 2 & 0 \\ -1 & 2 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = 0 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$$

$\Longrightarrow \quad v = (1,0,1)^{\mathrm{t}}$ ist ein Eigenvektor zu λ

$\operatorname{Rang}(A - 0E) = 2 \quad \Longrightarrow \quad \dim \underbrace{\operatorname{Kern}(A - 0E)}_{V_\lambda} = 3 - 2 = 1$, d.h.

$$V_\lambda = \{sv : s \in \mathbb{R}\},$$

alle Eigenvektoren zu $\lambda = 0$ sind parallel zu v

(ii) Eigenwert $\lambda = 2$:

$$\begin{pmatrix} 1 & 2 & -1 \\ 0 & 2 & 0 \\ -1 & 2 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix} = 2 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

$\implies \quad (1,1,1)^{\mathrm{t}}$ ist ein Eigenvektor zu λ

Bestimmung aller Eigenvektoren durch Lösen des linearen Gleichungssystems $(A - 2E)v = (0,0,0)^{\mathrm{t}}$:

$$\begin{pmatrix} -1 & 2 & -1 \\ 0 & 0 & 0 \\ -1 & 2 & -1 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\implies \quad v = \begin{pmatrix} s \\ (s+t)/2 \\ t \end{pmatrix} = s \underbrace{\begin{pmatrix} 1 \\ 1/2 \\ 0 \end{pmatrix}}_{u} + t \underbrace{\begin{pmatrix} 0 \\ 1/2 \\ 1 \end{pmatrix}}_{w}$$

mit $s, t \in \mathbb{R}$, d.h. der Eigenraum V_λ ist zweidimensional mit einer möglichen Basis $\{u, w\}$

Mögliche Fälle des Eigenwertproblems für reelle (2×2)-Matrizen

- Zwei reelle Eigenwerte

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}, \quad \lambda = -1,\, v \parallel \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \quad \tilde{\lambda} = 3,\, \tilde{v} \parallel \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

- Ein reeller Eigenwert, zwei linear unabhängige Eigenvektoren

$$A = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}, \quad \lambda = 2,\, v \parallel \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \tilde{v} \parallel \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad V_\lambda = \mathbb{R}^2$$

- Ein reeller Eigenwert, ein bis auf Vielfache eindeutiger Eigenvektor

$$A = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}, \quad \lambda = 1,\, v \parallel \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

- Zwei konjugiert komplexe Eigenwerte und Eigenvektoren

$$A = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}, \quad \lambda_\pm = 1 \pm \mathrm{i}\,, v_{1,2} \parallel \begin{pmatrix} 1 \\ \mp \mathrm{i} \end{pmatrix}$$

Ähnlichkeitstransformation

Bei einer Ähnlichkeitstransformation einer quadratischen Matrix A,

$$A \to B = Q^{-1} A Q\,,$$

bleiben die Eigenwerte erhalten.

Die Eigenvektoren werden gemäß dem durch die invertierbare Matrix Q beschriebenen Basiswechsel transformiert, d.h. einem Eigenvektor v von A zum Eigenwert λ entspricht der Eigenvektor $w = Q^{-1} v$ von B zum gleichen Eigenwert.

Ebenfalls invariant unter Ähnlichkeitstransformationen sind die Determinante und die Spur einer Matrix, da sie aus den Eigenwerten berechnet werden können:

$$\det A = \prod \lambda_k, \quad \operatorname{Spur} A = \sum \lambda_k\,.$$

Beweis

$$Av = \lambda v \quad \wedge \quad w = Q^{-1}v$$

$\Longrightarrow$

$$\begin{aligned} Bw &= Q^{-1}AQw = Q^{-1}AQQ^{-1}v \\ &= Q^{-1}Av = Q^{-1}\lambda v = \lambda Q^{-1}v \\ &= \lambda w \end{aligned}$$

Beispiel

Ähnlichkeitstransformation $A \to B = Q^{-1}AQ$ mit

$$A = \begin{pmatrix} 1 & -2 \\ 1 & 4 \end{pmatrix}, \quad Q = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$$

(i) Berechnung von B:

$$\begin{aligned} B = Q^{-1}AQ &= \begin{pmatrix} -1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & -2 \\ 1 & 4 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \\ &= \begin{pmatrix} -1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} -2 & -1 \\ 4 & 5 \end{pmatrix} = \begin{pmatrix} 6 & 6 \\ -2 & -1 \end{pmatrix} \end{aligned}$$

(ii) Eigenwerte und -Vektoren von A:

$$\lambda = 2,\ v = \begin{pmatrix} -2 \\ 1 \end{pmatrix}, \quad \tilde{\lambda} = 3,\ \tilde{v} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

(iii) Transformation der Eigenvektoren:

$\rightsquigarrow$ Eigenvektoren von B

$$w = Q^{-1}v = \begin{pmatrix} -1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} -2 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ -2 \end{pmatrix}$$

und

$$\tilde{w} = Q^{-1}\tilde{v} = \begin{pmatrix} -1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$$

Probe

$$\begin{aligned} Bw &= \begin{pmatrix} 6 & 6 \\ -2 & -1 \end{pmatrix} \begin{pmatrix} 3 \\ -2 \end{pmatrix} = \begin{pmatrix} 6 \\ -4 \end{pmatrix} = 2w \\ B\tilde{w} &= \begin{pmatrix} 6 & 6 \\ -2 & -1 \end{pmatrix} \begin{pmatrix} 2 \\ -1 \end{pmatrix} = \begin{pmatrix} 6 \\ -3 \end{pmatrix} = 3\tilde{w} \end{aligned}$$

gleiche Eigenwerte $\lambda = 2$, $\tilde{\lambda} = 3$ $\quad\checkmark$

(iv) Invarianz von Determinante und Spur:

$$\begin{aligned} \det A = 4 + 2 = 6 = -6 + 12 = \det B \\ \operatorname{Spur} A = 1 + 4 = 5 = -1 + 6 = \operatorname{Spur} B \end{aligned}$$

Die Eigenwerte einer $n \times n$-Matrix A sind die Nullstellen des charakteristischen Polynoms

$$p_A(\lambda) = \det(A - \lambda E) = \begin{vmatrix} a_{1,1} - \lambda & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} - \lambda & \dots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \dots & a_{n,n} - \lambda \end{vmatrix} .$$

Aufgrund der Eigenschaften von Determinanten sind die Eigenwerte invariant unter Transposition und Ähnlichkeitstransformation der Matrix A.

Für jeden Eigenwert λ sind Eigenvektoren v nichttriviale Lösungen des homogenen linearen Gleichungssystems

$$(A - \lambda E)v = (0, \dots, 0)^{\mathrm{t}} .$$

Um den Eigenraum V_λ, d.h. alle Eigenvektoren zum Eigenwert λ, zu bestimmen, kann man dieses System auf Zeilenstufenform transformieren. Dann sind

$$\dim V_\lambda = n - \underbrace{\mathrm{Rang}(A - \lambda E)}_{\text{Anzahl der Pivots}}$$

Komponenten von v frei wählbar. Bei kleineren Matrizen ist die Bestimmung der Eigenvektoren nach Vorgabe geeigneter Komponenten auch unmittelbar durchführbar.

Beweis

(i) Charakteristisches Polynom:

λ ist genau dann Eigenwert der $n \times n$-Matrix A, wenn

$$Av = \lambda v\,, \quad v \neq (0, \ldots, 0)^{\mathrm{t}}\,.$$

$\iff \quad \exists$ eine nicht-triviale Lösung v des homogenen linearen Gleichungssystem $(A - \lambda E)v = (0, \ldots, 0)^{\mathrm{t}}$

$\iff \quad (A - \lambda E)$ ist singulär $\quad \iff \quad \det(A - \lambda E) = 0$

(ii) Transposition und Ähnlichkeitstransformation:

- $\det B = \det B^{\mathrm{t}} \quad \implies$

$$\det(A^{\mathrm{t}} - \lambda E) = \det(A^{\mathrm{t}} - \lambda E)^{\mathrm{t}} = \det(A - \lambda E)$$

- $\det Q^{-1} = 1/\det Q$, $\det(CD) = (\det C)(\det D) \quad \implies$

$$\begin{aligned} \det(Q^{-1}AQ - \lambda E) &= \det(Q^{-1}(A - \lambda E)Q) \\ &= \det Q^{-1} \det(A - \lambda E) \det Q = \det(A - \lambda E) \end{aligned}$$

Beispiel

Eigenwerte und Eigenvektoren der Matrix

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

(i) Charakteristisches Polynom $p_A(\lambda)$:

$$\begin{vmatrix} -\lambda & -2 & -1 \\ 2 & 2-\lambda & 1 \\ 0 & 2 & 2-\lambda \end{vmatrix} \underset{\text{Sarrus}}{=} -\lambda(2-\lambda)^2 - 4 + 2\lambda + 4(2-\lambda)$$

$$= -\lambda^3 + 4\lambda^2 - 6\lambda + 4$$

Nullstellen $\rightsquigarrow$ Eigenwerte
Raten (Teiler des Absolutgliedes) $\rightsquigarrow$ $\lambda_1 = 2$
Abdividieren des Linearfaktors $\lambda - 2$,

$$(-\lambda^3 + 4\lambda^2 - 6\lambda + 4)/(\lambda - 2) = -\lambda^2 + 2\lambda - 2\,,$$

und Lösungsformel für quadratische Gleichungen $\rightsquigarrow$ $\lambda_{2,3} = 1 \pm \mathrm{i}$

(ii) Eigenvektoren:

- Ein Eigenvektor v zum Eigenwert 2 löst das homogene System

$$(A-2E)v = \begin{pmatrix} -2 & -2 & -1 \\ 2 & 0 & 1 \\ 0 & 2 & 0 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Rang $A > 1$ und $\leq 3 \quad \Longrightarrow \quad$ Rang $A = 2 \quad \Longrightarrow \quad$ eine Komponente von v beliebig wählbar
$v_1 = s \quad \underset{\text{Gleichung 2,3}}{\Longrightarrow} \quad v_3 = -2s,\ v_2 = 0$, d.h.

$$v \parallel (1,0,-2)^{\mathrm{t}}, \quad V_2 = \{(s,0,-2s)^{\mathrm{t}} :\ s \in \mathbb{R}\}$$

Ebenfalls möglich (notwendig in komplizierteren Fällen): Transformation von $A - \lambda E$ auf Zeilenstufenform

$$\begin{pmatrix} -2 & -2 & -1 \\ 2 & 0 & 1 \\ 0 & 2 & 0 \end{pmatrix} \to \begin{pmatrix} -2 & -2 & -1 \\ 0 & -2 & 0 \\ 0 & 2 & 0 \end{pmatrix} \to \begin{pmatrix} -2 & -2 & -1 \\ 0 & -2 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

- Ein Eigenvektor w zum Eigenwert $1+\mathrm{i}$ löst

$$(A-(1+\mathrm{i})E)w_1 = \begin{pmatrix} -1-\mathrm{i} & -2 & -1 \\ 2 & 1-\mathrm{i} & 1 \\ 0 & 2 & 1-\mathrm{i} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

$\dim V_{1+\mathrm{i}} = 1 \quad \rightsquigarrow \quad$ wähle (z.B.) $w_3 = s$
Gleichung 3 $\quad \Longrightarrow \quad w_2 = s(-1+\mathrm{i})/2$
Gleichung 2 $\quad \Longrightarrow \quad w_1 = (-(1-\mathrm{i})w_2 - w_3) = s(-1-\mathrm{i})/2$
konsistent (notwendigerweise!) mit Gleichung 1
$\rightsquigarrow$

$$w \parallel \left(\frac{-1-\mathrm{i}}{2}, \frac{-1+\mathrm{i}}{2}, 1\right)^{\mathrm{t}}, \quad V_{1+\mathrm{i}} = \left\{ s\left(\frac{-1-\mathrm{i}}{2}, \frac{-1+\mathrm{i}}{2}, 1\right)^{\mathrm{t}} : s \in \mathbb{R} \right\}$$

- A reell $\quad \Longrightarrow \quad$ Eigenvektoren zu $\lambda = 1 \pm \mathrm{i}$ komplex konjugiert, d.h. die zu $1-\mathrm{i}$ gehörigen Eigenvektoren sind

$$\overline{w} \parallel \left(\frac{-1+\mathrm{i}}{2}, \frac{-1-\mathrm{i}}{2}, 1\right)^{\mathrm{t}}$$

Algebraische und geometrische Vielfachheit

Die Vielfachheit eines Eigenwerts λ einer $n \times n$-Matrix A als Nullstelle des charakteristischen Polynoms

$$p_A(\lambda) = \det(A - \lambda E)$$

wird als algebraische Vielfachheit m_λ bezeichnet, die Dimension d_λ des Eigenraums V_λ als geometrische Vielfachheit.

Es gilt

$$d_\lambda \leq m_\lambda, \quad \sum_\lambda m_\lambda = n$$

sowie

$$d_\lambda = n - \operatorname{Rang}(A - \lambda E)\,.$$

Beispiel

Eigenvektorstruktur der Matrizen

$$A = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{pmatrix}, \quad B = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 5 & 1 \\ 0 & -1 & 3 \end{pmatrix}, \quad C = \begin{pmatrix} 4 & 4 & 0 \\ 3 & 4 & 6 \\ 0 & -2 & 4 \end{pmatrix}$$

mit identischen charakteristischen Polynomen

(i) Charakteristisches Polynom:

$$\begin{aligned} p(\lambda) &= \det(\text{Matrix} - \lambda\,\text{Einheitsmatrix}) \\ &\underset{A}{=} (4-\lambda)^3 \\ &\underset{B}{=} (4-\lambda)(5-\lambda)(3-\lambda) + 4 - \lambda \\ &\underset{C}{=} (4-\lambda)^3 + 12(4-\lambda) - 12(4-\lambda) \end{aligned}$$

$\Longrightarrow$ algebraische Vielfachheit $m_4 = 3$ für alle Matrizen

(ii) Eigenräume:

- A:

$$A - 4E = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \mathrm{Rang}(A - 4E) = 0, \, \dim V_4 = 3 - 0 = 3$$

$$\Longrightarrow \qquad V_4 = \mathbb{R}^3, \text{ jeder Vektor ungleich } (0,0,0)^t \text{ ist Eigenvektor}$$

- B:

$$B - 4E = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & -1 & -1 \end{pmatrix}, \quad \mathrm{Rang}(B - 4E) = 1, \, \dim V_4 = 2$$

Eigenvektoren $\perp (0,1,1)^t$,
mögliche Basis für V_4: $\{(1,0,0)^t, (0,1,-1)^t\}$

- C:

$$C - 4E = \begin{pmatrix} 0 & 4 & 0 \\ 3 & 0 & 6 \\ 0 & -2 & 0 \end{pmatrix}, \quad \mathrm{Rang}(C - 4E) = 2,\ \dim V_4 = 1$$

Eigenvektoren $\parallel (2,0,-1)^{\mathrm{t}},\ V_4 = \{(2s,0,-s)^{\mathrm{t}} :\ s \in \mathbb{R}\}$

Summe und Produkt von Eigenwerten

Für die Eigenwerte λ_k einer $n \times n$-Matrix A gilt

$$\sum_{k=1}^{n} \lambda_k = \operatorname{Spur} A, \quad \prod_{k=1}^{n} \lambda_k = \det A\,,$$

wobei mehrfache Eigenwerte entsprechend ihrer algebraischen Vielfachheit gezählt werden.

Diese Identitäten können unter anderem zur Kontrolle bei Eigenwertberechnungen verwendet werden.

Beweis

Die Eigenwerte λ_k sind die Nullstellen des charakteristischen Polynoms

$$p_A(\lambda) = \det(A - \lambda E) = \det(a_1 - \lambda e_1, \ldots, a_n - \lambda e_n)\,,$$

wobei a_k die Spalten von A und e_k die Einheitsvektoren bezeichnen.
Multilinearität der Determinante $\quad\leadsto\quad$ Summe von 2^n Determinanten, die man nach Potenzen von λ zusammenfassen kann:

$$p_A(\lambda) = (-\lambda)^n \det(e_1, \ldots, e_n) + (-\lambda)^{n-1} c + \cdots + (-\lambda)^0 \det(a_1, \ldots, a_n)$$

mit

$$c = \det(a_1, e_2, \ldots, e_n) + \det(e_1, a_2, \ldots, e_n) + \cdots + \det(e_1, e_2, \ldots, a_n)$$

$a_1 = \sum_{k=1}^n a_{k,1} e_k$ und $\det(e_j, \ldots, e_j, \ldots, e_n) = 0$ (Antisymmetrie der Determinante)
$\implies \quad \det(a_1, e_2, \ldots, e_n) = a_{1,1} \det(e_1, \ldots, e_n)$

gleiches Argument für die anderen Summanden in der Darstellung von c, $\det(e_1, \ldots, e_n) = 1 \quad \rightsquigarrow$

$$p_A(\lambda) = (-\lambda)^n + \underbrace{(a_{1,1} + \cdots + a_{n,n})}_{\text{Spur } A}(-\lambda)^{n-1} + \cdots + \det(a_1, \ldots, a_n)$$

Vergleich mit der Faktorisierung von p_A in Linearfaktoren,

$$p_A(\lambda) = \prod_{k=1}^{n}(\lambda_k - \lambda) = (-\lambda)^n + (\sum_k \lambda_k)(-\lambda)^{n-1} + \cdots + \prod_k \lambda_k$$

$\Longrightarrow \quad$ behauptete Identitäten

Beispiel

Summe und Produkt der Eigenwerte einer 2×2-Matrix

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

charakteristisches Polynom

$$p_A(\lambda) = \begin{vmatrix} a - \lambda & b \\ c - \lambda & d \end{vmatrix} = \lambda^2 - (a+d)\lambda + (ad - bc)$$

mit den Nullstellen

$$\lambda_{1,2} = \frac{(a+d) \pm \sqrt{(a+d)^2 - 4(ad-bc)}}{2}$$

Summe der Eigenwerte

$$\lambda_1 + \lambda_2 = \frac{(a+d)+(a+d)}{2} = a + d = \operatorname{Spur} A$$

Produkt der Eigenwerte (via dritter Binomischer Formel)

$$\lambda_1 \lambda_2 = \frac{(a+d)^2 - (a+d)^2 + 4(ad-bc)}{4} = ad - bc = \det A$$

Basis aus Eigenvektoren

Existiert zu einer $n \times n$-Matrix A eine Basis $B = \{v_1, \ldots, v_n\}$ aus Eigenvektoren, so hat die lineare Abbildung $L : x \mapsto Ax$ bzgl. dieser Basis Diagonalform,

$$V^{-1}AV = \operatorname{diag}(\lambda_1, \ldots, \lambda_n), \quad V = (v_1, \ldots, v_n)\,,$$

mit λ_k dem Eigenwert zu v_k. Mit anderen Worten: Sind y_k die Koordinaten eines Vektors x bezüglich B, so sind $\lambda_k y_k$ die Koordinaten von $L(x)$,

$$x = \sum_{k=1}^{n} y_k v_k \quad \Longrightarrow \quad L(x) = \sum_{k=1}^{n} (\lambda_k y_k) v_k\,.$$

Beweis

(i) Diagonalisierung:
Zusammenfassen von $Av_k = \lambda_k v_k$, $k = 1, \ldots, n$ als Spalten einer Matrix
$\rightsquigarrow$

$$AV = A(v_1, \ldots, v_n) = (v_1\lambda_1, \ldots, v_n\lambda_n) = VD, \quad D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)\,,$$

d.h, $V^{-1}AV = D$

(ii) Basiswechsel:
$x = \sum_k v_k y_k = Vy$, $L(x) = Ax = \sum_k v_k \tilde{y}_k = V\tilde{y} \quad \Longrightarrow$

$$\tilde{y} = V^{-1}Ax = V^{-1}AVy\,,$$

d.h. $D = V^{-1}AV$ ist die Matrixdarstellung von L bzgl. der Basis $B = \{v_1, \ldots, v_n\}$

Beispiel

Diagonalisierung der Matrix

$$A = \begin{pmatrix} 4 & 6 \\ -3 & -5 \end{pmatrix}$$

(i) Eigenwerte:
Nullstellen des charakteristischen Polynoms

$$p_A(\lambda) = \begin{vmatrix} 4-\lambda & 6 \\ -3 & -5-\lambda \end{vmatrix} = (4-\lambda)(-5-\lambda) + 18 = \lambda^2 + \lambda - 2$$

Lösungsformel für quadratische Gleichungen
$\rightsquigarrow \quad -1/2 \pm \sqrt{1/4 + 2} = -1/2 \pm 3/2$
$\rightsquigarrow \quad$ Eigenwerte $\lambda_+ = 1$, $\lambda_- = -2$

(ii) Eigenvektoren:
Lösungen des linearen Gleichungssystems $(A - \lambda E)v = (0,0)^{\mathrm{t}}$

- $\lambda_+ = 1$
$$\begin{pmatrix} 4-1 & 6 \\ -3 & -5-1 \end{pmatrix} v_+ = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \Longrightarrow \quad v_+ \parallel \begin{pmatrix} 2 \\ -1 \end{pmatrix}$$
- $\lambda_- = -2$
$$\begin{pmatrix} 4+2 & 6 \\ -3 & -5+2 \end{pmatrix} v_- = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \Longrightarrow \quad v_- \parallel \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

(iii) Ähnlichkeitstransformation:
Transformmationsmatrix mit Eigenvektoren als Spalten
$$V = (v_+, v_-) = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}$$
Cramersche Regel $\quad\rightsquigarrow\quad$ Inverse
$$V^{-1} = \frac{1}{\det V} \begin{pmatrix} v_{2,2} & -v_{1,2} \\ -v_{2,1} & v_{1,1} \end{pmatrix} \underset{\det V = 1}{=} \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$$
$$\Longrightarrow \qquad V^{-1}AV = D,\ VDV^{-1} \text{ mit } D = \operatorname{diag}(1,-2)$$

Probe Ausmultiplizieren der Matrizen

$$\begin{aligned} D = V^{-1}AV &= \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 4 & 6 \\ -3 & -5 \end{pmatrix} \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 2 & 2 \\ -1 & -2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -2 \end{pmatrix} \quad \checkmark \end{aligned}$$

Diagonalisierung zyklischer Matrizen

Eine zyklische $n \times n$-Matrix A kann mit Hilfe der Fourier-Matrix

$$W = (w^{jk})_{j,k=0,\ldots,n-1}, \quad w = \exp(2\pi \mathrm{i}/n)\,,$$

diagonalisiert werden:

$$\frac{1}{n}\overline{W}\begin{pmatrix} a_0 & a_{n-1} & \cdots & a_1 \\ a_1 & a_0 & \cdots & a_2 \\ \vdots & & \ddots & \vdots \\ a_{n-1} & a_{n-2} & \cdots & a_0 \end{pmatrix} W = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix},$$

d.h. die Spalten von W, $(1, w^k, w^{2k}, \ldots, w^{(n-1)k})^{\mathrm{t}}$, $k = 0, \ldots, n-1$, sind die Eigenvektoren von A. Die Eigenwerte $\lambda_0, \ldots, \lambda_{n-1}$ lassen sich auch unmittelbar aus dem erzeugenden Vektor der zyklischen Matrix berechnen:

$$\begin{pmatrix} \lambda_0 \\ \vdots \\ \lambda_{n-1} \end{pmatrix} = \overline{W}\begin{pmatrix} a_0 \\ \vdots \\ a_{n-1} \end{pmatrix} \quad \Longleftrightarrow \quad \lambda_\ell = \sum_{k=0}^{n-1} a_k w^{-k\ell},\ \ell = 0, \ldots, n-1\,.$$

Beweis

$A: \ (a_{j-k \bmod n})_{j,k=0,\ldots,n-1}$
j-te Komponente des Vektors $A\,v_\ell$, $v_\ell = (w^{0\ell}, \cdots, w^{(n-1)\ell})^{\mathrm{t}}$

$$\sum_{k=0}^{n-1} a_{j-k \bmod n} w^{(k-j)\ell} w^{j\ell} \underset{k'=j-k}{=} w^{j\ell} \sum_{k'=j-n+1}^{j} a_{k' \bmod n} w^{-k'\ell}$$

Indizes modulo n
$\rightsquigarrow \quad k' \in \{j-n+1, \ldots, -1\} \mathrel{\widehat{=}} k' \in \{j+1, \ldots, n-1\}$
$\implies \quad \sum_{k'} \cdots = \lambda_\ell$, d.h. $Av_\ell = \lambda_\ell v_\ell$
schreibe die Identität für v_ℓ, $\ell = 0, \ldots, n-1$, in Matrixform

$$A \underbrace{(v_0, \ldots, v_{n-1})}_{W} = (v_0, \ldots, v_{n-1}) D, \quad D = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix}$$

$W/\sqrt{n}$ unitär und $W = W^{\mathrm{t}} \implies (W/\sqrt{n})^{-1} = W^*/\sqrt{n}$ bzw.
$W^{-1} = \overline{W}/n$
$\rightsquigarrow \quad (\overline{W}/n)AW = D$, d.h. die behauptete Diagonalisierung

Beispiel

Diagonalform der zyklischen Matrix

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

Fourier-Matrix: $W = (\mathrm{i}^{jk})_{j,k=0,1,2,3} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & \mathrm{i} & -1 & -\mathrm{i} \\ 1 & -1 & 1 & -1 \\ 1 & -\mathrm{i} & -1 & \mathrm{i} \end{pmatrix}$

Eigenwerte: $\begin{pmatrix} \lambda_0 \\ \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -\mathrm{i} & -1 & \mathrm{i} \\ 1 & -1 & 1 & -1 \\ 1 & \mathrm{i} & -1 & -\mathrm{i} \end{pmatrix}}_{\overline{W}} \begin{pmatrix} 2 \\ 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 2 \\ 0 \\ 2 \end{pmatrix}$

$\rightsquigarrow$ Diagonalisierung: $\frac{1}{4}\overline{W}AW = \mathrm{diag}(4,2,0,2)$

A reell $\quad\rightsquigarrow\quad$ reelle Basis für den Eigenraum zum Eigenwert 2 durch Linearkombinationen der komplexen Eigenvektoren

$$\begin{pmatrix} 1 \\ \mathrm{i} \\ -1 \\ -\mathrm{i} \end{pmatrix}, \quad \begin{pmatrix} 1 \\ -\mathrm{i} \\ -1 \\ \mathrm{i} \end{pmatrix} \qquad \text{(zweite und vierte Spalte von } W\text{)}$$

$\rightsquigarrow\quad$ reelle Eigenvektoren

$$\begin{pmatrix} 1 \\ \mathrm{i} \\ -1 \\ -\mathrm{i} \end{pmatrix} + \begin{pmatrix} 1 \\ -\mathrm{i} \\ -1 \\ \mathrm{i} \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \\ -2 \\ 0 \end{pmatrix}, \quad \mathrm{i}\begin{pmatrix} 1 \\ \mathrm{i} \\ -1 \\ -\mathrm{i} \end{pmatrix} - \mathrm{i}\begin{pmatrix} 1 \\ -\mathrm{i} \\ -1 \\ \mathrm{i} \end{pmatrix} = \begin{pmatrix} 0 \\ -2 \\ 0 \\ 2 \end{pmatrix}$$

Unitäre Diagonalisierung

Eine $n \times n$-Matrix A ist normal, wenn sie mit ihrer Adjungierten $A^* = \bar{A}^\mathrm{t}$ ($A^* = A^\mathrm{t}$ für reelle Matrizen) kommutiert:

$$A^*A = AA^* \,.$$

Normalität von A ist äquivalent zu unitärer Diagonalisierbarkeit, d.h. es existiert eine orthonormale Basis aus Eigenvektoren $u_1, \ldots, u_n$ und

$$U^*AU = \operatorname{diag}(\lambda_1, \ldots, \lambda_n), \quad U = (u_1, \ldots, u_n),\ U^* = U^{-1}\,,$$

mit λ_k den Eigenwerten von A.

Beweis

(i) Orthogonale Diagonalisierbarkeit $\implies$ Normalität:

$$U^{-1} \underset{U\,\text{unitär}}{=} U^*, \quad U^{-1}AU = D \,\text{bzw.}\, UDU^{-1} = UDU^*$$

mit $D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n) \implies$

$$\begin{aligned}
AA^* &= (UDU^*)(UDU^*)^* \\
&\underset{(BC)^*=C^*B^*}{=} UD(U^* \underbrace{U^{**}}_{U})D^*U^* \\
&\underset{U^*U=E}{=} U(D^*D)U^* \\
&\underset{U^*U=E}{=} (UD^*U^*)(UDU^*) \\
&= A^*A
\end{aligned}$$

$DD^* = D^*D$ wegen $\lambda\bar{\lambda} = \bar{\lambda}\lambda$

(ii) Normalität $\implies$ orthogonale Diagonalisierbarkeit:
konstruiere zunächst eine orthogonale Matrix V mit

$$V^*AV = \left(\begin{array}{c|c} \lambda & \underline{0}^{\mathrm{t}} \\ \hline \underline{0} & B \end{array}\right), \quad B \,\text{normal},\ \underline{0} = (0,\ldots,0)^{\mathrm{t}} \in \mathbb{R}^{n-1}$$

Ansatz: $V = (v|\tilde{V})$, wobei die Spalten von $\tilde{V}$ einen normierten Eigenvektor v zum Eigenwert λ zu einer Orthonormalbasis ergänzen

$$\rightsquigarrow \quad V^*AV = \left(\begin{array}{c} v^* \\ \hline \tilde{V}^* \end{array}\right)\left(\begin{array}{c|c} \lambda v & A\tilde{V} \end{array}\right) = \left(\begin{array}{c|c} \lambda & c^* \\ \hline \underline{0} & B \end{array}\right) = D$$

noch zu zeigen: $BB^* = B^*B \ \wedge\ c = (0,\ldots,0)^{\mathrm{t}}$
Ähnlichkeitstransformationen erhalten Normalität $\rightsquigarrow$ $DD^* = D^*D$, d.h.

$$\left(\begin{array}{c|c} |\lambda|^2 + |c|^2 & c^*B^* \\ \hline Bc & BB^* \end{array}\right) = \left(\begin{array}{c|c} |\lambda|^2 & \overline{\lambda}c^* \\ \hline \lambda c & B^*B \end{array}\right)$$

Vergleich der Diagonalblöcke $\implies$ gewünschte Eigenschaften nehme induktiv an, dass B mit einer unitären Matrix W diagonalisierbar ist, d.h.

$$W^*BW = \tilde{D}$$

$\Longrightarrow$ A wird durch

$$U = V\left(\begin{array}{c|c} 1 & \underline{0}^t \\ \hline \underline{0} & W \end{array}\right),$$

diagonalisiert, denn

$$\begin{aligned} U^*AU &= \left(\begin{array}{c|c} 1 & \underline{0}^t \\ \hline \underline{0} & W^* \end{array}\right)\left(\begin{array}{c|c} \lambda & \underline{0}^t \\ \hline \underline{0} & B \end{array}\right)\left(\begin{array}{c|c} 1 & \underline{0}^t \\ \hline \underline{0} & W \end{array}\right) \\ &= \left(\begin{array}{c|c} \lambda & \underline{0} \\ \hline \underline{0}^t & W^*BW \end{array}\right) = \left(\begin{array}{c|c} \lambda & \underline{0} \\ \hline \underline{0}^t & \tilde{D} \end{array}\right) \end{aligned}$$

Beispiel

Diagonalisierung der normalen Matrix

$$A = \begin{pmatrix} 1+2\mathrm{i} & 1-2\mathrm{i} \\ 1-2\mathrm{i} & 1+2\mathrm{i} \end{pmatrix}$$

(i) Überprüfung der Normalität:

$$\begin{aligned} A^*A &= \begin{pmatrix} 1-2\mathrm{i} & 1+2\mathrm{i} \\ 1+2\mathrm{i} & 1-2\mathrm{i} \end{pmatrix} \begin{pmatrix} 1+2\mathrm{i} & 1-2\mathrm{i} \\ 1-2\mathrm{i} & 1+2\mathrm{i} \end{pmatrix} = \begin{pmatrix} 10 & -6 \\ -6 & 10 \end{pmatrix} \\ AA^* &= \begin{pmatrix} 1+2\mathrm{i} & 1-2\mathrm{i} \\ 1-2\mathrm{i} & 1+2\mathrm{i} \end{pmatrix} \begin{pmatrix} 1-2\mathrm{i} & 1+2\mathrm{i} \\ 1+2\mathrm{i} & 1-2\mathrm{i} \end{pmatrix} = \text{gleiches Resultat} \quad \checkmark \end{aligned}$$

(ii) Eigenwerte- und Vektoren:
beide Zeilensummen von A gleich 2 $\implies$
2 ist Eigenwert mit Eigenvektor $(1,1)^{\mathrm{t}}$

Orthogonalität der Eigenvektoren $\implies$
$(-1,1)^{\mathrm{t}}$ ist ein zweiter Eigenvektor

$$\begin{pmatrix} 1+2\mathrm{i} & 1-2\mathrm{i} \\ 1-2\mathrm{i} & 1+2\mathrm{i} \end{pmatrix} \begin{pmatrix} -1 \\ 1 \end{pmatrix} = 4\mathrm{i} \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

$\implies$ 4i ist der zugehörige Eigenwert

(iii) Diagonalisierung:
Normierung der Eigenvektoren $\rightsquigarrow$ Transformationsmatrix

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

und $U^*AU = D$, $D = \operatorname{diag}(2, 4\mathrm{i})$
Überprüfung durch Ausmultiplizieren

$$\begin{aligned} U^*AU &= \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1+2\mathrm{i} & 1-2\mathrm{i} \\ 1-2\mathrm{i} & 1+2\mathrm{i} \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \\ &= \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 2 & -4\mathrm{i} \\ 2 & 4\mathrm{i} \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 4 & 0 \\ 0 & 8\mathrm{i} \end{pmatrix} = D \quad \checkmark \end{aligned}$$

Diagonalform hermitescher Matrizen

Die Eigenwerte λ_k einer hermiteschen $n \times n$-Matrix A ($A = A^* = \bar{A}^{\mathrm{t}}$ bzw. $A = A^{\mathrm{t}}$ für reelle Matrizen) sind reell, und es existiert eine Orthonormalbasis aus Eigenvektoren u_k. Folglich ist

$$U^* A U = \operatorname{diag}(\lambda_1, \ldots, \lambda_n), \quad U^* = U^{-1}\,,$$

mit der unitären Matrix $U = (u_1, \ldots, u_n)$ und λ_k den Eigenwerten von A.

Beweis

hermitesch $\implies$ normal $\implies$ A ist unitär diagonalisierbar:

$$U^*AU = D, \quad U^* = U^{-1},\ D = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$$

zu zeigen: $\lambda \in \mathbb{R}$ für alle Eigenwerte λ

$Av = \lambda \quad \implies$

$$\begin{aligned} \lambda v^* v &= v^*(\lambda v) = v^*(Av) = (A^* v)^* v = (Av)^* v = (\lambda v)^* v \\ &= \overline{\lambda} v^* v \end{aligned}$$

Division durch $|v|^2 = v^* v \quad \implies \quad \lambda = \overline{\lambda}$, d.h. $\lambda \in \mathbb{R}$

Beispiel

Diagonalisierung der hermiteschen Matrix

$$A = \begin{pmatrix} 2 & -1-\mathrm{i} \\ -1+\mathrm{i} & 1 \end{pmatrix} = A^* = \overline{A}^{\mathrm{t}}$$

$\det A = 0 \quad \implies \quad \exists$ Eigenwert $\lambda = 0$
zugehöriger Eigenvektor

$$v = \begin{pmatrix} 1 \\ 1-\mathrm{i} \end{pmatrix}$$

$\operatorname{Spur} A = 3 = \lambda + \varrho \quad \rightsquigarrow \quad$ zweiter Eigenwert $\varrho = 3$
zugehöriger Eigenvektor $w \perp v$, d.h.

$$w = \begin{pmatrix} 1+\mathrm{i} \\ -1 \end{pmatrix}$$

$(w^* v = \overline{w}^{\mathrm{t}} v = (1-\mathrm{i}, -1)(1, 1-\mathrm{i})^{\mathrm{t}} = 0)$

Normierung $\quad\rightsquigarrow\quad$ unitäre Matrix

$$U = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 1+\mathrm{i} \\ 1-\mathrm{i} & -1 \end{pmatrix}$$

Diagonalisierung

$$A = U \begin{pmatrix} 0 & 0 \\ 0 & 3 \end{pmatrix} U^*$$

Überprüfung durch Ausmultiplizieren

$$\underbrace{\begin{pmatrix} 2 & -1-\mathrm{i} \\ -1+\mathrm{i} & 1 \end{pmatrix}}_{A} \overset{!}{=} \frac{1}{3} \begin{pmatrix} 1 & 1+\mathrm{i} \\ 1-\mathrm{i} & -1 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} 1 & 1+\mathrm{i} \\ 1-\mathrm{i} & -1 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 1+\mathrm{i} \\ 1-\mathrm{i} & -1 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 1-\mathrm{i} & -1 \end{pmatrix} \quad \checkmark$$

Rayleigh-Quotient

Für eine hermitesche positiv definite Matrix S ($S^* = S$, $x^* S x > 0$ für $x \neq (0, \ldots, 0)^{\mathrm{t}}$) sind die Extremwerte des sogenannten Rayleigh-Quotienten

$$r_S(x) = \frac{x^* S x}{x^* x}, \quad x \neq (0, \ldots, 0)^{\mathrm{t}},$$

der kleinste und größte Eigenwert von S.
Insbesondere gilt für die der euklidischen Norm zugeordneten Matrix-Norm

$$\|S\|_2 = \max_x r_S(x) = \lambda_{\max}, \quad \|S^{-1}\|_2 = 1 / \min_x r_S(x) = 1/\lambda_{\min}.$$

Beweis

Transformation auf Diagonalform mit unitärer Matrix $U = (u_1, \cdots, u_n)$ aus Eigenvektoren $\rightsquigarrow$

$$U^* S U = D, \quad D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n), \quad S u_k = \lambda_k u_k \quad \text{mit } \lambda_k \in \mathbb{R}$$

S positiv definit $\implies$

$$\lambda_k = \frac{u_k^* S u_k}{u_k^* u_k} = \frac{u_k^* \lambda u_k}{u_k^* u_k} > 0$$

und o.B.d.A. $\lambda_1 \geq \cdots \geq \lambda_n$
Substitution von $S = UDU^*$ und $x = Uy$ $\rightsquigarrow$

$$r_S(x) = \frac{x^* S x}{x^* x} = \frac{(Uy)^*\, UDU^*\, Uy}{y^*(U^* U) y} \underset{U^* U = E}{=} \frac{\sum_k \lambda_k |y_k|^2}{\sum_k |y_k|^2}$$

$\lambda_k > 0 \implies$

$$\lambda_{\min} \leq r_S(x) \leq \lambda_{\max}$$

mit Gleichheit für $y = e_n$ und $y = e_1$

Beispiel

Rayleigh-Quotient der Matrix

$$S = \begin{pmatrix} 3 & 2 \\ 2 & 3 \end{pmatrix}$$

Invarianz des Rayleigh-Quotienten $r_s(x)$ unter Skalierung $(x \to \lambda x)$
$\rightsquigarrow \quad$ o.B.d.A. $x = (\cos t, \sin t)^{\mathrm{t}}$
$x^* x = |x|^2 = 1 \qquad \Longrightarrow$

$$r(x) = \frac{x^* S x}{x^* x} = 3\cos^2 t + 4\cos t \sin t + 3\sin^2 t = 3 + 2\sin(2t)$$

Extrema

$$0 \stackrel{!}{=} 4\cos(2t) \qquad \Longrightarrow \qquad t \in \{\pm\pi/4 + k\pi/2 : \; k \in \mathbb{Z}\}$$

kleinster und größter Eigenwert

$$\lambda_+ = \max_x r_S(x) = 3 + 2\sin(\frac{\pi}{2}) = 5, \; \lambda_- = \min_x r_S(x) = 3 + 2\sin(-\frac{\pi}{2}) = 1$$

Kontrolle mit Hilfe der Identitäten für Determinante und Spur

$$\begin{aligned}
9 - 4 &= \det A = \lambda_+ \lambda_- = 5 \cdot 1, \\
3 + 3 &= \operatorname{Spur} A = \lambda_+ + \lambda_- = 5 + 1
\end{aligned}$$

Dreiecksform

Jede $n \times n$-Matrix A lässt sich durch eine unitäre Ähnlichkeitstransformation U auf obere Dreiecksform bringen:

$$R = \begin{pmatrix} \lambda_1 & r_{1,2} & \cdots & r_{1,n} \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & r_{n-1,n} \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix} = U^* A U, \quad U^* = \bar{U}^{\mathrm{t}} = U^{-1}\,.$$

Die Diagonale von R enthält die Eigenwerte λ_k von A und die Spalten von U bilden eine orthonormale Basis.

Beweis

Induktion über die Dimension n:

- $n = 1$: 1×1-Matrix $\quad \checkmark$
- $n - 1 \to n$:
  Eine $n \times n$-Matrix A hat (mindestens) einen Eigenvektor v.
  Normierung und orthogonale Ergänzung $\quad \rightsquigarrow \quad$ orthonormale Basis $\{v, w_1, \ldots, w_{n-1}\}$
  Ähnlichkeitstransformation mit $V = (v, w_1, \ldots, w_{n-1}) \quad \rightsquigarrow$

$$\begin{aligned}
\tilde{A} \;&=\; V^* A V = \begin{pmatrix} v^* \\ w_1^* \\ \vdots \\ w_{n-1}^* \end{pmatrix} \left(\lambda v \;\middle|\; A w_1, \ldots, A w_{n-1} \right) \\
&=\; \begin{pmatrix} \lambda & x^* \\ \underline{0} & B \end{pmatrix},
\end{aligned}$$

da $w_k^* v = 0$ und mit $\bar{x}_k = A w_k$, $\underline{0} = (0, \ldots, 0)^{\mathrm{t}}$

Induktionsvoraussetzung $\Longrightarrow$
$\exists$ unitäre Ähnlichkeitstransformation auf Dreiecksform für die $n-1 \times n-1$-Matrix B:

$$\tilde{U}^* B \tilde{U} = R_{n-1}$$

Mit

$$U = V \begin{pmatrix} 1 & \underline{0}^* \\ \underline{0} & \tilde{U} \end{pmatrix}, \quad \underline{0} = (0, \ldots, 0)^{\mathrm{t}},\ \underline{0}^* = \underline{0}^{\mathrm{t}}$$

folgt

$$\begin{aligned} U^* A U &= \begin{pmatrix} 1 & \underline{0}^* \\ \underline{0} & \tilde{U}^* \end{pmatrix} V^* A V \begin{pmatrix} 1 & \underline{0}^* \\ \underline{0} & \tilde{U} \end{pmatrix} \\ &= \begin{pmatrix} 1 & \underline{0}^* \\ \underline{0} & \tilde{U}^* \end{pmatrix} \begin{pmatrix} \lambda & x^* \\ \underline{0} & B \end{pmatrix} \begin{pmatrix} 1 & \underline{0}^* \\ \underline{0} & \tilde{U} \end{pmatrix} \\ &= \begin{pmatrix} 1 & \underline{0}^* \\ \underline{0} & \tilde{U}^* \end{pmatrix} \begin{pmatrix} \lambda & x^* \tilde{U} \\ \underline{0} & B \tilde{U} \end{pmatrix} \\ &= \begin{pmatrix} \lambda & x^* \tilde{U} \\ \underline{0} & \tilde{U}^* B \tilde{U} \end{pmatrix} = \begin{pmatrix} \lambda & x^* \tilde{U} \\ \underline{0} & R_{n-1} \end{pmatrix} = R \end{aligned}$$

Beispiel

Unitäre Ähnlichkeitstransformation der Matrix

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

auf Dreiecksform

charakteristisches Polynom

$$p_A(\lambda) = \begin{vmatrix} -1-\lambda & -4 \\ 1 & 3-\lambda \end{vmatrix} = (-1-\lambda)(3-\lambda)+4 = (\lambda-1)^2$$

doppelte Nullstelle $\lambda = 1 \quad \Longrightarrow \quad \lambda = 1$ ist einziger Eigenwert
Eigenvektor

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix} = (A-\lambda E)v = \begin{pmatrix} -2 & -4 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \quad \Longrightarrow \quad v \parallel \begin{pmatrix} 2 \\ -1 \end{pmatrix}$$

Normierung und Ergänzung zu einer orthonormalen Basis $\{v, w\}$ $\rightsquigarrow$ unitäre Transformationsmatrix

$$U = (v, w) = \frac{1}{\sqrt{5}} \begin{pmatrix} 2 & 1 \\ -1 & 2 \end{pmatrix}$$

$U^* = U^{\mathrm{t}}$ für die reelle Matrix U $\rightsquigarrow$ Dreiecksform

$$\begin{aligned} U^{\mathrm{t}} A U &= \frac{1}{\sqrt{5}} \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} -1 & -4 \\ 1 & 3 \end{pmatrix} \frac{1}{\sqrt{5}} \begin{pmatrix} 2 & 1 \\ -1 & 2 \end{pmatrix} \\ &= \frac{1}{5} \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 2 & -9 \\ -1 & 7 \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & -5 \\ 0 & 1 \end{pmatrix}}_{R} \end{aligned}$$

Jordan-Form

Jede $n \times n$-Matrix A lässt sich durch eine Ähnlichkeitstransformation auf die Blockdiagonalform

$$J = \begin{pmatrix} J_1 & & 0 \\ & \ddots & \\ 0 & & J_k \end{pmatrix} = Q^{-1} A Q$$

bringen. Der ℓ-te Jordanblock ist eine $n_\ell \times n_\ell$-Matrix der Form

$$J_\ell = \begin{pmatrix} j_{i_\ell, i_\ell} & \cdots \\ \vdots & \ddots \end{pmatrix} = \begin{pmatrix} \lambda_\ell & 1 & & & 0 \\ 0 & \lambda_\ell & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda_\ell & 1 \\ 0 & & & & \lambda_\ell \end{pmatrix},$$

mit einem Eigenwert λ_ℓ von A und der i_ℓ-ten Spalte von Q einem zugehörigen Eigenvektor.

Bis auf Permutationen der Blöcke ist die Jordan-Form eindeutig.

Ist die Matrix A diagonalisierbar, d.h. existiert eine Basis aus Eigenvektoren, so treten keine Einsen auf der Nebendiagonale auf. Alle Jordanblöcke sind in diesem Fall 1×1-Matrizen, $J_\ell = \lambda_\ell$.

Nicht-triviale Jordanblöcke zu einem Eigenwert λ_ℓ existieren genau dann wenn die algebraische Vielfachheit von λ_ℓ größer als die geometrische Vielfachheit ist. Die Anzahl dieser Jordanblöcke entspricht der Dimension des Eigenraums V_{λ_ℓ}, d.h. der Anzahl der linear unabhängigen Eigenvektoren zu λ_ℓ.

Beispiel

verschiedene Jordan-Formen $J = Q^{-1}AQ$ für eine 2×2-Matrix A

(i) Zwei einfache Eigenwerte λ und ϱ:

$$J = \begin{pmatrix} \lambda & 0 \\ 0 & \varrho \end{pmatrix} = Q^{-1}AQ, \quad Q = (v, w)$$

mit v und w Eigenvektoren zu λ und ϱ

(ii) Doppelter Eigenwert λ mit zwei linear unabhängigen Eigenvektoren v und w:

$$J = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} = A\,,$$

denn mit $Q = (v, w)$ ist

$$A = QJQ^{-1} = Q(\lambda E)Q^{-1} = \lambda E\,,$$

jeder Vektor ist Eigenvektor zu λ

(iii) Doppelter Eigenwert λ mit nur einem, bis auf Vielfache eindeutig bestimmten Eigenvektor v:

$$J = Q^{-1}AQ = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}, \quad Q = (v, w)\,,$$

mit einem Hauptvektor w

$\mathrm{Rang}(A - \lambda E) = 1 \quad \Longrightarrow$

Ein Eigenvektor v ist mit einer nicht-trivialen Gleichung des homogenen Systems $(A - \lambda E)v = 0$ bestimmbar, die andere Gleichung ist redundant.

Bestimmung von w aus der zweiten Spalte der Identität

$$QJ = AQ \quad \Longleftrightarrow \quad (\lambda v, v + \lambda w) = (Av, Aw)\,,$$

d.h. $v = (A - \lambda E)w$

Beispiel

Jordan-Form der Matrix

$$A = \begin{pmatrix} -4 & 4 \\ -9 & 8 \end{pmatrix}$$

charakteristisches Polynom

$$p_A(\lambda) = \underbrace{\begin{vmatrix} -4-\lambda & 4 \\ -9 & 8-\lambda \end{vmatrix}}_{A-\lambda E} = (-4-\lambda)(8-\lambda)+36 = \lambda^2-4\lambda+4 = (\lambda-2)^2$$

doppelte Nullstelle $\rightsquigarrow$ Eigenwert $\lambda = 2$ mit algebraischer Vielfachheit 2
lineares Gleichungssystem für einen Eigenvektor v

$$(A-2E)v = \begin{pmatrix} -6 & 4 \\ -9 & 6 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$\rightsquigarrow \quad v = (2,3)^{\mathrm{t}}$

kein zweiter linear unabhängiger Eigenvektor (geometrische Vielfachheit von λ gleich 1) $\quad\rightsquigarrow\quad$ Jordanform

$$J = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix} = Q^{-1}AQ, \quad Q = (v, w)$$

Bestimmung eines Hauptvektors w aus der zweiten Spalte der Identität

$$QJ = AQ \quad \Longleftrightarrow \quad (2v, v + 2w) = (Av, Aw)$$

d.h.

$$\underbrace{\begin{pmatrix} 2 \\ 3 \end{pmatrix}}_{v} = \underbrace{\begin{pmatrix} -6 & 4 \\ -9 & 6 \end{pmatrix}}_{A-2E} \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}$$

$\rightsquigarrow \quad w = (1, 2)^{\mathrm{t}}$

Probe

$$\begin{pmatrix} -4 & 4 \\ -9 & 8 \end{pmatrix} = A = QJQ^{-1} = \begin{pmatrix} 2 & 1 \\ 3 & 2 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 2 & -1 \\ -3 & 2 \end{pmatrix} \quad \checkmark$$

Beispiel

verschiedene Jordan-Formen

$$J = Q^{-1}AQ, \quad Q = (u, v, w)$$

für eine 3×3-Matrix A

(i) Paarweise verschiedene Eigenwerte λ, ϱ, σ:

$$J = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \varrho & 0 \\ 0 & 0 & \sigma \end{pmatrix}$$

(ii) Doppelter Eigenwert λ:

$$\begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \varrho \end{pmatrix}, \quad \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \varrho \end{pmatrix}$$

Zweiter Fall:

$$\mathrm{Rang}(A - \lambda E) = 2$$

Bestimmung von Eigenvektoren u und w und einem Hauptvektor v durch Lösen der Gleichungssysteme

$$(A - \lambda E)u = \underline{0}, \quad u = (A - \lambda E)v, \quad (A - \varrho E)w = \underline{0}, \quad \underline{0} = (0,0,0)^{\mathrm{t}}$$

(iii) Dreifacher Eigenwert λ:

$$\begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix}, \quad \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix}, \quad \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}$$

Zweiter Fall:

$$\mathrm{Rang}(A - \lambda E) = 1$$

Bestimmung eines Hauptvektors v durch Lösen von

$$(A - \lambda E)^2 v = \underline{0}, \quad (A - \lambda E)v \neq \underline{0}$$

zugehöriger Eigenvektor: $u = (A - \lambda E)v$
Ein weiterer linear unabhängiger Eigenvektor w erfüllt

$$(A - \lambda E)w = \underline{0}, \quad w \perp u$$

Dritter Fall:

$$\text{Rang}(A - \lambda E) = 2$$

Bestimmung eines Eigenvektors u und von zwei Hauptvektoren v und w durch Lösen von

$$(A - \lambda E)u = \underline{0}, \quad u = (A - \lambda E)v, \quad v = (A - \lambda E)w$$

Beispiel

Jordan-Form der Matrix

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

(i) Eigenwerte:
Entwickeln von $\det(A - \lambda E)$ nach der ersten Zeile $\quad \rightsquigarrow$
charakteristisches Polynom

$$\begin{aligned} p_A(\lambda) &= (-1-\lambda)^3 + 4(4 - 3(-1-\lambda)) = -\lambda^3 - 3\lambda^2 + 9\lambda + 27 \\ &= (\lambda+3)^2(3-\lambda) \end{aligned}$$

Nullstellen: $\lambda_1 = -3$ (doppelt), $\lambda_2 = 3$ (einfach)

$$\mathrm{Rang}(A - \lambda_1 E) = \mathrm{Rang}\begin{pmatrix} 2 & 0 & 4 \\ 2 & 2 & 0 \\ 3 & 2 & 2 \end{pmatrix} = 2$$

$\implies$ bis auf Vielfache nur ein Eigenvektor u zu $\lambda = -3$

$\Longrightarrow$ ein nicht-trivialer Jordanblock, d.h.

$$J = \begin{pmatrix} -3 & 1 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 3 \end{pmatrix} = Q^{-1}AQ, \quad Q = (u, v, w)$$

mit u und w Eigenvektoren zu -3 und 3

(ii) Eigenvektoren und Hauptvektor:

- $\lambda_1 = -3$:

$$\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = (A - (-3)E)u = \begin{pmatrix} 2 & 0 & 4 \\ 2 & 2 & 0 \\ 3 & 2 & 2 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}$$

$\Longrightarrow$ $u = (-2, 2, 1)^{\mathrm{t}}$ ist ein Eigenvektor

- $\lambda_2 = 3$:
  analog $(0, 0, 0)^{\mathrm{t}} = (A - 3E)w \quad \rightsquigarrow \quad w = (2, 1, 2)^{\mathrm{t}}$

Bestimmung eines Hauptvektors v aus der zweiten Spalte der Identität

$$QJ = AQ \quad \Longleftrightarrow \quad (-3u, u - 3v, 3w) = (Au, Av, Aw)$$

d.h.

$$\begin{pmatrix} -2 \\ 2 \\ 1 \end{pmatrix} = u = Av+3v = \begin{pmatrix} 2 & 0 & 4 \\ 2 & 2 & 0 \\ 3 & 2 & 2 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \quad \Longrightarrow \quad v = \begin{pmatrix} -3 \\ 4 \\ 1 \end{pmatrix}$$

$\rightsquigarrow$ Transformationsmatrix

$$Q = (u, v, w) = \begin{pmatrix} -2 & -3 & 2 \\ 2 & 4 & 1 \\ 1 & 1 & 2 \end{pmatrix}$$

Beispiel

14 mögliche Belegunsstrukturen für Jordan-Normalformen von 4×4-Matrizen

(i) Vier verschiedene Eigenwerte $\lambda_1, \lambda_2, \lambda_3, \lambda_4$:

$$\begin{pmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ 0 & 0 & \lambda_3 & 0 \\ 0 & 0 & 0 & \lambda_4 \end{pmatrix}$$

(ii) Ein zweifacher Eigenwert λ_1 und zwei einfache Eigenwerte λ_2, λ_3:

$$\begin{pmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 \\ 0 & 0 & \lambda_2 & 0 \\ 0 & 0 & 0 & \lambda_3 \end{pmatrix}, \begin{pmatrix} \lambda_1 & 1 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 \\ 0 & 0 & \lambda_2 & 0 \\ 0 & 0 & 0 & \lambda_3 \end{pmatrix}$$

(iii) Zwei zweifache Eigenwerte λ_1, λ_2:

$$\begin{pmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 \\ 0 & 0 & \lambda_2 & 0 \\ 0 & 0 & 0 & \lambda_2 \end{pmatrix}, \begin{pmatrix} \lambda_1 & 1 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 \\ 0 & 0 & \lambda_2 & 0 \\ 0 & 0 & 0 & \lambda_2 \end{pmatrix}, \begin{pmatrix} \lambda_1 & 1 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 \\ 0 & 0 & \lambda_2 & 1 \\ 0 & 0 & 0 & \lambda_2 \end{pmatrix}$$

(iv) Ein dreifacher Eigenwert λ_1 und ein einfacher Eigenwert λ_2:

$$\begin{pmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 \\ 0 & 0 & \lambda_1 & 0 \\ 0 & 0 & 0 & \lambda_2 \end{pmatrix}, \begin{pmatrix} \lambda_1 & 1 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 \\ 0 & 0 & \lambda_1 & 0 \\ 0 & 0 & 0 & \lambda_2 \end{pmatrix}, \begin{pmatrix} \lambda_1 & 1 & 0 & 0 \\ 0 & \lambda_1 & 1 & 0 \\ 0 & 0 & \lambda_1 & 0 \\ 0 & 0 & 0 & \lambda_2 \end{pmatrix}$$

(v) Ein vierfacher Eigenwert λ_1:

$$\begin{pmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 \\ 0 & 0 & \lambda_1 & 0 \\ 0 & 0 & 0 & \lambda_1 \end{pmatrix}, \begin{pmatrix} \lambda_1 & 1 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 \\ 0 & 0 & \lambda_1 & 0 \\ 0 & 0 & 0 & \lambda_1 \end{pmatrix}, \begin{pmatrix} \lambda_1 & 1 & 0 & 0 \\ 0 & \lambda_1 & 1 & 0 \\ 0 & 0 & \lambda_1 & 0 \\ 0 & 0 & 0 & \lambda_1 \end{pmatrix}$$

$$\begin{pmatrix} \lambda_1 & 1 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 \\ 0 & 0 & \lambda_1 & 1 \\ 0 & 0 & 0 & \lambda_1 \end{pmatrix}, \begin{pmatrix} \lambda_1 & 1 & 0 & 0 \\ 0 & \lambda_1 & 1 & 0 \\ 0 & 0 & \lambda_1 & 1 \\ 0 & 0 & 0 & \lambda_1 \end{pmatrix}$$

Dominanz des größten Eigenwerts

Besitzt A einen betragsmäßig größten Eigenwert λ und ist die Dimension des Eigenraums V_λ maximal, d.h. gleich der algebraischen Vielfachheit von λ, so gilt für einen Vektor x mit einer nicht-trivialen Komponente $v \in V_\lambda$

$$A^n x = \lambda^n (v + o(1)), \quad n \to \infty$$

bzw. nach Normierung

$$\lim_{n\to\infty} \frac{A^n x}{\|A^n x\|} = \frac{v}{\|v\|},$$

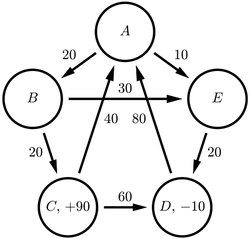
d.h. die normierten Potenzen streben gegen einen normierten Eigenvektor.

Insbesondere gelten beide Identitäten für Matrizen mit einem betragsmäßig größten einfachen Eigenwert.

Beispiel

Jährliche Veränderung der Marktanteile x_k konkurrierender Firmen

In dem illustrierten Fall gewinnt die Firma A jährlich 80% der Marktanteile der Firma D, und die Firma C vergrößert ihre Marktanteile um 90%, verliert jedoch gleichzeitig Marktanteile an die Firmen A und D.



Veränderung der Marktanteile

$$
\begin{aligned}
A_{\text{neu}} &= 0.7A + 0.4C + 0.8D \\
B_{\text{neu}} &= 0.5B + 0.2A \\
C_{\text{neu}} &= 0.9C + 0.2B \\
D_{\text{neu}} &= 0.1D + 0.6C + 0.2E \\
E_{\text{neu}} &= 0.8E + 0.1A + 0.3B
\end{aligned}
$$

Beschreibung durch eine Matrixmultiplikation

$$x_{\text{neu}} = \underbrace{\begin{pmatrix} 0.7 & 0 & 0.4 & 0.8 & 0 \\ 0.2 & 0.5 & 0 & 0 & 0 \\ 0 & 0.2 & 0.9 & 0 & 0 \\ 0 & 0 & 0.6 & 0.1 & 0.2 \\ 0.1 & 0.3 & 0 & 0 & 0.8 \end{pmatrix}}_{P} x_{\text{alt}}, \quad x = (A, B, C, D, E)^{\text{t}}$$

Multiplikation mit der n-ten Potenz von P

$\rightsquigarrow$ Marktanteile nach n Jahren

Die normierten Vektoren

$$x^\circ = x/\|x\|_1, \quad \|x\|_1 = \sum_k |x_k|$$

konvergieren gegen einen Eigenvektor zum betragsmäßig größten Eigenwert:

$$\lambda_{\max} = 1.1, \quad v_{\max} = (3, 1, 1, 1, 2)^{\text{t}}/8$$

$\rightsquigarrow$ prozentuale Anteile $A:$ 37.5%, $B, C, D:$ 12.5%, $E:$ 25%

Beispiel

Rekursion der Fibonacci-Zahlen

$$a_0 = 0,\ a_1 = 1, \quad a_{n+1} = a_n + a_{n-1},\ n > 0$$

$\rightsquigarrow$ Folge
0, 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, 233, 377, 610, 987, ...

Matrixform der Rekursion:

$$x_{n+1} = Ax_n, \quad x_n = (a_{n-1}, a_n)^{\mathrm{t}}, \quad A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$$

Eigenwerte und Eigenvektoren von A:

$$\lambda_\pm = \frac{1}{2} \pm \frac{\sqrt{5}}{2}, \quad v_\pm = \begin{pmatrix} 1 \\ \lambda_\pm \end{pmatrix}$$

Darstellung des Startvektors als Linearkombination von v_+ und v_-,

$$\begin{pmatrix} a_0 \\ a_1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{5}} v_+ - \frac{1}{\sqrt{5}} v_-$$

$\rightsquigarrow$ asymptotisches Verhalten

$$\begin{pmatrix} a_{n-1} \\ a_n \end{pmatrix} = A^{n-1} \begin{pmatrix} a_0 \\ a_1 \end{pmatrix} = \frac{\lambda_+^{n-1}}{\sqrt{5}} v_+ - \frac{\lambda_-^{n-1}}{\sqrt{5}} v_-$$

d.h.

$$a_n = \frac{\lambda_+^{n-1}}{\sqrt{5}} \lambda_+ - \frac{\lambda_-^{n-1}}{\sqrt{5}} \lambda_- = \frac{1}{\sqrt{5}} \left(\frac{1}{2} + \frac{\sqrt{5}}{2} \right)^n (1 - \underbrace{(\lambda_- / \lambda_+)^n}_{o(1)})$$

andere Startwerte $a_1 = 1$ und $a_2 = 3$ $\rightsquigarrow$ Lucas-Zahlen:

$$1, 3, 4, 7, 11, 18, 29, 47, 76, 123, 199, 322, 521, 843, 1364, 2207, 3571, \ldots$$

gleiches Wachstumsverhalten: $a_{n+1}/a_n \to (1 + \sqrt{5})/2$

Konvergenz von Matrix-Potenzen

Die Potenzen A^n, $n = 0, 1, \ldots$, einer Matrix konvergieren genau dann gegen die Nullmatrix, wenn der Betrag aller Eigenwerte λ kleiner als 1 ist.

Die Folge (A^n) bleibt beschränkt, wenn $|\lambda| \leq 1$ für alle Eigenwerte und für Eigenwerte mit Betrag 1 die algebraische gleich der geometrischen Vielfachheit ist, d.h. eine Basis aus Eigenvektoren für den Eigenraum V_λ existiert.

Andernfalls divergiert die Folge, insbesondere dann, wenn ein Eigenwert mit Betrag größer als 1 existiert.

Beweis

Darstellung der Potenzen mit der Jordan-Form $J = Q^{-1}AQ \quad \rightsquigarrow$

$$A^n = (QJQ^{-1})(QJQ^{-1})\cdots(QJQ^{-1}) = QJ^nQ^{-1}$$

$\implies$ Es ist ausreichend die Konvergenz der Potenzen der Blockdiagonalmatrix J zu zeigen.
betrachte dazu die Blöcke

$$J_k = (\lambda_k E) + D$$

mit λ_k einem Eigenwert von A und D einer Matrix mit einer Nebendiagonale aus Einsen oder (für einen diagonalen Jordanblock) der Nullmatrix
$D^m = 0$ für einen Block der Dimension $m \quad \implies$

$$(J_k)^n = \lambda_k^n E + \binom{n}{1}\lambda_k^{n-1}D + \cdots + \binom{n}{m-1}\lambda_k^{n-m+1}D^{m-1}$$

Definition des Binomialkoeffizienten, $\binom{n}{j} = \frac{n\cdot(n-1)\cdots(n-j+1)}{j\cdot(j-1)\cdots 1} \quad \implies$ Konvergenzeigenschaften

$|\lambda_k| < 1$: $\quad \lim_{n\to\infty} \binom{n}{j} |\lambda_k|^{n-k} \le |\lambda_k^{-k}| \lim_{n\to\infty} n^j |\lambda_k|^n = 0$
$\implies \quad (J_k)^n \to$ Nullmatrix

$|\lambda_k| = 1$: $\quad (J_k)^n$ nur beschränkt, wenn D die Nullmatrix ist

$|\lambda_k| > 1$: $\quad$ Divergenz, da $\lambda_k^n \to \infty$

Beispiel

Potenzen der 3×3-Matrizen

$$A = \begin{pmatrix} \frac{1}{2} & 1 & 0 \\ 0 & \frac{1}{2} & 1 \\ 0 & 0 & \frac{1}{2} \end{pmatrix}, \; B = \begin{pmatrix} \frac{1}{2} & 1 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}, \; C = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{2} \end{pmatrix}$$

(i) $A^n \to$ Nullmatrix, da der Betrag aller Eigenwerte ($1/2$ mit Vielfachheit 3) kleiner als 1 ist

$$A^2 = \begin{pmatrix} \frac{1}{4} & 1 & 1 \\ 0 & \frac{1}{4} & 1 \\ 0 & 0 & \frac{1}{4} \end{pmatrix}, \ldots, A^n = \begin{pmatrix} 2^{-n} & n2^{1-n} & \frac{n(n-1)}{2}2^{2-n} \\ 0 & 2^{-n} & n2^{1-n} \\ 0 & 0 & 2^{-n} \end{pmatrix}$$

(ii) $\|B^n\| \leq \text{const}$, da der Betrag aller Eigenwerte kleiner oder gleich 1 ist, und der Eigenwert 1 einfach ist

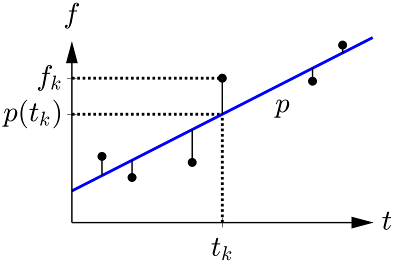
$$B^2 = \begin{pmatrix} \frac{1}{4} & 1 & 0 \\ 0 & \frac{1}{4} & 0 \\ 0 & 0 & 1 \end{pmatrix}, \ldots, B^n = \begin{pmatrix} 2^{-n} & n2^{1-n} & 0 \\ 0 & 2^{-n} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

(iii) C^n divergiert, da ein Eigenwert mit Betrag 1 und einem nicht-diagonalen Jordanblock existiert

$$C^2 = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{4} \end{pmatrix}, \ldots, C^n = \begin{pmatrix} 1 & n & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2^{-n} \end{pmatrix}$$

Ausgleichsgerade

Eine Gerade $g: p(t) = u + vt$, die Daten (t_k, f_k), $k = 1, \ldots, n$, bestmöglichst approximiert ($p(t_k) \approx f_k$), kann durch Minimierung der Fehlerquadratsumme



$$\sum_{k=1}^{n} (p(t_k) - f_k)^2$$

ermittelt werden.
Der Achsenabschnitt u und die Steigung v berechnen sich gemäß

$$u = \frac{(\sum t_k^2)(\sum f_k) - (\sum t_k)(\sum t_k f_k)}{n(\sum t_k^2) - (\sum t_k)^2}, \quad v = \frac{n(\sum t_k f_k) - (\sum t_k)(\sum f_k)}{n(\sum t_k^2) - (\sum t_k)^2},$$

wobei (sinnvollerweise) angenommen wird, dass mindestens zwei Abszissen t_k verschieden sind.

Beweis

(u, v) minimal $\quad\Longrightarrow$

Ableitungen der Fehlerquadratsumme $\sum_k (u + vt_k - f_k)^2$ nach u und v Null:

$$\begin{aligned} 0 &= 2\sum_k (u + vt_k - f_k) \\ 0 &= 2\sum_k t_k(u + vt_k - f_k) \end{aligned}$$

bzw. in Matrixform

$$\underbrace{\begin{pmatrix} n & \sum t_k \\ \sum t_k & \sum t_k^2 \end{pmatrix}}_{A} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} \sum f_k \\ \sum t_k f_k \end{pmatrix}$$

mindestens zwei t_k verschieden, Cauchy-Schwarz-Ungleichung $\quad\Longrightarrow$

$$\det A = \underbrace{|(1, \ldots, 1)|^2}_{=n} |(t_1, \ldots, t_n)|_2^2 - \left(\sum_k 1 \cdot t_k\right)^2 > 0\,,$$

da $(1, \ldots, 1)$ und $(t_1, \ldots, t_n)$ nicht parallel sind

Cramersche Regel $\quad\rightsquigarrow\quad$ angegebene Formeln für u und v

Beispiel

Bestimmung der Ausgleichsgerade $g: t \mapsto p(t) = u + tv$ zu den Daten

t_k	-1	0	2
f_k	-3	1	4

$$\sum t_k = 1,\ \sum f_k = 2,\ \sum t_k^2 = 5,\ \sum t_k f_k = 11$$

Einsetzen in die Formeln für u und v $\rightsquigarrow$

$$u = \frac{(\sum t_k^2)(\sum f_k) - (\sum t_k)(\sum t_k f_k)}{n(\sum t_k^2) - (\sum t_k)^2} = \frac{5 \cdot 2 - 1 \cdot 11}{3 \cdot 5 - 1^2} = -\frac{1}{14}$$

$$v = \frac{n(\sum t_k f_k) - (\sum t_k)(\sum f_k)}{n(\sum t_k^2) - (\sum t_k)^2} = \frac{3 \cdot 11 - 1 \cdot 2}{3 \cdot 5 - 1^2} = \frac{31}{14}$$

Normalengleichungen

Eine Näherungslösung für ein überbestimmtes lineares Gleichungssystem $Ax - b$ kann man durch Minimierung der Norm des Residuums $r = Ax - b$ erhalten:

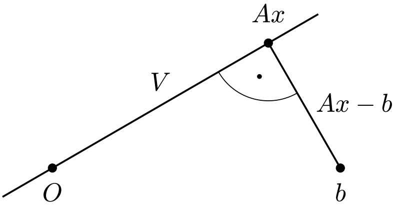
$$|Ax - b|^2 = \sum_{j=1}^{m} \left| \sum_{k=1}^{n} a_{j,k} x_k \, - b_j \right|^2 \to \min \, .$$

Man bestimmt eine beste Approximation $\sum_{k=1}^{n} x_k a_k$ zu b in dem von den Spalten $a_1, \ldots, a_n$ von A aufgespannten Unterraum $V = \operatorname{Bild} A$. Für eine reelle $m \times n$-Matrix A erfüllt jede Lösung x dieses sogenannten Ausgleichsproblems die Normalengleichungen

$$A^{\mathrm{t}} A \, x = A^{\mathrm{t}} b \, .$$

Dies bedeutet, dass das Residuum r (Fehler der Approximation) orthogonal zu dem von den Vektoren $a_1, \ldots, a_n$ aufgespannten Unterraum ist:

$$a_k^{\mathrm{t}} r = 0, \, \forall k \iff A^{\mathrm{t}} x - b \perp V \, .$$



Die Matrix $A^{\mathrm{t}}A$ ist quadratisch und hat Dimension n. Sie ist genau dann invertierbar, wenn $\operatorname{Rang} A = n$, d.h. wenn die Spalten a_k von A linear unabhängig sind. In diesem Fall besitzt das Ausgleichsproblem eine eindeutige Lösung.

Die Normalengleichungen sind auch im singulären Fall lösbar. Das Residuum r ist auch dann eindeutig, die Lösung x jedoch nicht.

Alle Aussagen gelten, allgemeiner, ebenfalls für komplexe Matrizen; A^{t} ist dabei durch die adjungierte Matrix $A^* = \bar{A}^{\mathrm{t}}$ zu ersetzen.

Beweis

Minimalität von $x \quad \Longleftrightarrow \quad \forall t \in \mathbb{R},\, y \in \mathbb{R}^n$:

$$|A(x+ty)-b|^2 \geq |Ax-b|^2, \quad \forall t \in \mathbb{R}, \forall y \in \mathbb{R}^n$$

Einsetzen der Definition der Norm,

$$\begin{aligned} |A(x+ty)-b|^2 &= ((Ax-b)^{\mathrm{t}} + t(Ay)^{\mathrm{t}})(Ax-b+tAy) \\ |Ax-b|^2 &= (Ax-b)^{\mathrm{t}}(Ax-b)\,, \end{aligned}$$

und Vereinfachung mit $r = Ax-b \quad \rightsquigarrow$

$$p = \underbrace{tr^{\mathrm{t}}Ay + ty^{\mathrm{t}}A^{\mathrm{t}}r}_{2ty^{\mathrm{t}}A^{\mathrm{t}}r} + t^2 y^{\mathrm{t}}A^{\mathrm{t}}Ay \geq 0$$

$(r^{\mathrm{t}}Ay = y^{\mathrm{t}}A^{\mathrm{t}}r)$
p: nicht-negative Parabel in t
$p \geq 0$ genau dann, wenn $y^{\mathrm{t}}(A^{\mathrm{t}}r) = 0$
y beliebig $\quad \Longrightarrow \quad A^{\mathrm{t}}r = (0,\ldots,0)^{\mathrm{t}}$

Beispiel

Reguläres und singuläres Ausgleichsproblem

(i) Rang A maximal:

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

Normalengleichungen

$$\underbrace{\begin{pmatrix} 5 & 1 \\ 1 & 5 \end{pmatrix}}_{A^{\mathrm{t}}A} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \underbrace{\begin{pmatrix} 3 \\ 3 \end{pmatrix}}_{A^{\mathrm{t}}b}$$

eindeutige Lösung $x = \begin{pmatrix} 1/2 & 1/2 \end{pmatrix}^{\mathrm{t}}$ mit Residuum

$$r = Ax - b = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$$

(ii) Rang A nicht maximal:

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

linear abhängige Spalten $\quad\rightsquigarrow\quad$ singuläre Normalengleichungen

$$\begin{pmatrix} 5 & 10 \\ 10 & 20 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 3 \\ 6 \end{pmatrix}$$

Lösung

$$x = \begin{pmatrix} 3/5 \\ 0 \end{pmatrix} + t \begin{pmatrix} -2 \\ 1 \end{pmatrix}, \quad t \in \mathbb{R}$$

nicht eindeutig, aber eindeutiges Residuum

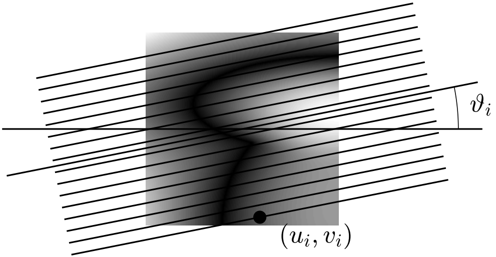
$$r = Ax - b = \begin{pmatrix} 2 & 4 \\ 1 & 2 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 3/5 - 2t \\ t \end{pmatrix} - \begin{pmatrix} 0 \\ 3 \\ 0 \end{pmatrix} = \begin{pmatrix} 6/5 \\ -12/5 \\ 0 \end{pmatrix}$$

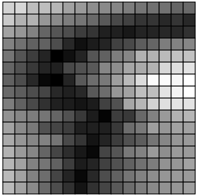
Beispiel

Computer-Tomographie:

Rekonstruktion einer Dichte $x(u,v)$ aus dem Intensitätsverlust von Röntgenstrahlen entlang von k Bündeln aus ℓ parallelen Geraden

$$\mathcal{R}_i: \ (u_i, v_i) + \mathbb{R}(\cos\vartheta_i, \sin\vartheta_i), \quad i = 1, \ldots, m = k\ell$$





Approximation von x durch eine stückweise konstante Funktion auf einem Raster von Quadraten $\mathcal{Q}_j$ und eine Näherung für die Linienintegrale

$$b_i = \int_{\mathbb{R}} x(u_i + t\cos\vartheta_i, v_i + t\sin\vartheta_i)\,\mathrm{d}t \approx \sum_{j=1}^{n} a_{i,j} x_j$$

mit x_j einer Approximation von $x(u,v)$ auf $\mathcal{Q}_j$ und

$$a_{i,j} = |\mathcal{R}_i \cap \mathcal{Q}_j|$$

der Länge des Durchschnitts der Geraden $\mathcal{R}_i$ mit dem Quadrat $\mathcal{Q}_j$

$m \gg n \;\rightsquigarrow\;$ Ausgleichsproblem zur Bestimmung von x aus den Daten b

Abbildung:
16×16 Raster, Winkel

$$\vartheta = 0, \pi/32, 2\pi/32, \ldots$$

mit 16 parallelen Scan-Richtungen im Abstand der Rasterquadratbreite
$\rightsquigarrow \quad 32 \cdot 16 = 512$ Gleichungen für $16 \cdot 16 = 256$ Unbekannte

Zu jeder $m \times n$-Matrix A existieren orthonormale Basen $\{u_1, \ldots, u_m\}$ und $\{v_1, \ldots, v_n\}$ bezüglich derer die lineare Abbildung $L : x \mapsto Ax$ Diagonalform hat:

$$v_k \mapsto Av_k = s_k u_k, \quad k = 1, \ldots, r = \operatorname{Rang} A\,,$$

mit $s_1 \geq \cdots \geq s_r > 0$ den sogenannten Singulärwerten von A.

Mit den unitären Matrizen $U = (u_1, \ldots, u_m)$ und $V = (v_1, \ldots, v_n)$ besitzt A die Faktorisierung

$$A = USV^*, \quad S = \begin{pmatrix} s_1 & & \mathbf{O} \\ & s_2 & \\ \mathbf{O} & & \ddots \end{pmatrix},$$

mit S einer $m \times n$-Diagonalmatrix und $V^* = \bar{V}^{\mathrm{t}}$ der adjungierten Matrix zu V ($V^* = V^{\mathrm{t}}$ für eine reelle Matrix). Entsprechend hat die lineare Abbildung L die Darstellung

$$x \mapsto Ax = \sum_{k=1}^{r} u_k s_k v_k^* x\,.$$

Daraus folgt insbesondere, dass $Av_k = (0, \ldots, 0)^{\mathrm{t}}$ für $k = r+1, \ldots, n$ und $Ax \in \operatorname{span}\{u_1, \ldots, u_r\}$, d.h.

$$\operatorname{Kern} A = \operatorname{span}\{v_{r+1}, \ldots, v_n\}, \quad \operatorname{Bild} A = \operatorname{span}\{u_1, \ldots, u_r\}\,.$$

Schließlich ist $\|A\|_2 = s_1$ und $\|A\|_F^2 = \sum_{j,k} |a_{j,k}|^2 = s_1^2 + \cdots + s_r^2$.

Die Singulärwertzerlegung lässt sich durch Lösen eines Eigenwertproblems bestimmen. Die Singulärwerte s_k sind die positiven Eigenwerte der symmetrischen positiv semidefiniten $n \times n$-Matrix A^*A und die Vektoren v_k sind die zugehörigen Eigenvektoren. Die Vektoren $u_1, \ldots, u_r$ sind die Eigenvektoren der $m \times m$-Matrix AA^* zu den gleichen Eigenwerten. Die Vektoren $u_{r+1}, \ldots, u_m$ und $v_{r+1}, \ldots, v_n$ sind für die Faktorisierung von A irrelevant. Sie ergänzen $u_1, \ldots, u_r$ und $v_1, \ldots, v_r$ zu orthonormalen Basen und spannen die Eigenräume zum Eigenwert 0 der Matrizen AA^* und A^*A auf.

Alternativ kann man u_k auch durch Normierung der Spalten $1, \ldots, r$ von $AV = US = (u_1 s_1, \ldots, u_r s_r, 0, \ldots, 0)$ bestimmen und durch Vektoren $u_{r+1}, \ldots, u_m$ zu einer orthonormalen Basis ergänzen.

Beweis

(i) Konstruktion:

A^*A hermitesch und positiv semidefinit $\quad\Longrightarrow$

$\exists$ orthonormale Basis $\{v_1, \ldots, v_n\}$ aus Eigenvektoren zu den nicht-negativen Eigenwerten $\lambda_j = s_j^2$, d.h.

$$V^*A^*AV = \operatorname{diag}(s_1^2, \ldots, s_k^2, 0, \ldots, 0) = S^{\mathrm{t}}S, \quad V = (v_1, \ldots, v_n),\ V^*V = E$$

(Eigenwerte s_j^2 absteigend sortiert)

Spalten von AV orthogonal mit Norm $s_j \quad\Longrightarrow$

$$AV = (s_1 u_1, \ldots, s_k u_k, \underbrace{\mathbf{0}}_{\text{Spalten } k+1\ldots m}) = US$$

mit einer unitären Matrix U; die Spalten $u_{k+1}, \ldots, u_m$ ergänzen $u_1, \ldots, u_k$ zu einer orthonormalen Basis

$\rightsquigarrow$ Darstellung

$$A = USV^* \quad\Longleftrightarrow\quad U^*AV = S$$

(ii) Rang:
Invarianz des Rangs einer Matrix unter unitären Transformationen $\implies$

$$\text{Rang } A = \text{Rang } S = k$$

(iii) $Av_j = s_j u_j$:
folgt aus

$$A = USV^*, \quad V^* v_j = e_j,\, Se_j = s_j e_j,\, Ue_j = u_j$$

mit e_j dem j-ten Einheitsvektor

(iv) $y = \sum_j s_j (v_j^* x) u_j$:
folgt aus (iii) und

$$x = \sum_j (v_j^* x) v_j$$

(Darstellung bzgl. einer orthonormalen Basis)

(v) $\|A\|_2 = s_1,\, \|A\|_F^2 = s_1^2 + \cdots + s_k^2$:
folgt aus der Invarianz der Euklidischen Norm und der Frobenius-Norm unter unitären Transformationen

Beispiel

Berechnung der Singulärwert-Zerlegung von

$$A = \begin{pmatrix} 6 & -5 & 8 \\ 6 & 5 & 8 \\ 6 & -5 & 8 \\ 6 & 5 & 8 \end{pmatrix}$$

(i) Bestimmung von V und S:

$$A^{\mathrm{t}}A = \begin{pmatrix} 144 & 0 & 192 \\ 0 & 100 & 0 \\ 192 & 0 & 256 \end{pmatrix}$$

Eigenwert 100 mit Eigenvektor $(0,1,0)^{\mathrm{t}}$
Spalte 3 $= (4/3) \times$ Spalte 1
$\rightsquigarrow$ Eigenwert 0 mit Eigenvektor $(4,0,-3)^{\mathrm{t}}/5$
Spur $A = 500$
$\rightsquigarrow$ Eigenwert $500 - 100 - 0 = 400$ mit orthogonalem Eigenvektor

absteigende Sortierung der singulären Werte s_j (Wurzeln der Eigenwerte von $A^{\mathrm{t}}A$) $\quad\rightsquigarrow$

$$V = \begin{pmatrix} 3/5 & 0 & 4/5 \\ 0 & 1 & 0 \\ 4/5 & 0 & -3/5 \end{pmatrix}, \quad S = \begin{pmatrix} 20 & 0 & 0 \\ 0 & 10 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

(ii) Bestimmung von U:

$$AV = \begin{pmatrix} 10 & -5 & 0 \\ 10 & 5 & 0 \\ 10 & -5 & 0 \\ 10 & 5 & 0 \end{pmatrix}$$

Normierung (Division durch s_1, s_2) $\quad\rightsquigarrow\quad$ Spalten 1 und 2 von U:

$$u_1 = \begin{pmatrix} 1/2 \\ 1/2 \\ 1/2 \\ 1/2 \end{pmatrix}, \quad u_2 = \begin{pmatrix} -1/2 \\ 1/2 \\ -1/2 \\ 1/2 \end{pmatrix}$$

Ergänzung zu einer orthogonalen Basis $\rightsquigarrow$

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

Probe $A \overset{!}{=} USV^{\mathrm{t}}$

$$A = \begin{pmatrix} 6 & -5 & 8 \\ 6 & 5 & 8 \\ 6 & -5 & 8 \\ 6 & 5 & 8 \end{pmatrix}$$

$$USV^{\mathrm{t}} = \frac{1}{2}\begin{pmatrix} 1 & -1 & 1 & -1 \\ 1 & 1 & 1 & 1 \\ 1 & -1 & -1 & 1 \\ 1 & 1 & -1 & -1 \end{pmatrix}\begin{pmatrix} 20 & 0 & 0 \\ 0 & 10 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}\frac{1}{5}\begin{pmatrix} 3 & 0 & 4 \\ 0 & 5 & 0 \\ 4 & 0 & -3 \end{pmatrix}$$

$\checkmark$

Mit der Singulärwertzerlegung einer $m \times n$-Matrix A, $A = USV^*$, lässt sich die Lösung x des Ausgleichsproblems $|Ax - b| \to \min$ mit minimaler Norm in der Form

$$x = A^+ b, \quad A^+ = VS^+ U^*,$$

darstellen mit

$$S^+ = \operatorname{diag}(1/s_1, \ldots, 1/s_r, 0, \ldots, 0), \quad r = \operatorname{Rang} A\,,$$

der $n \times m$-Diagonalmatrix mit den Kehrwerten der Singulärwerte. Für eine reelle Matrix A sind die unitären Matrizen U, V ebenfalls reell und $U^* = \bar{U}^{\mathrm{t}} = U^{\mathrm{t}}$, $V^* = \bar{V}^{\mathrm{t}} = V^{\mathrm{t}}$.

Die Matrix A^+ wird als Pseudo-Inverse bezeichnet. Mit $\|C\|_F = (\sum_{j,k} |c_{j,k}|^2)^{1/2}$ der Frobenius-Norm und $E_{m\times m}$ der $m \times m$-Einheitsmatrix gilt

$$\|E_{m\times m} - AA^+\|_F = \min_X \|E_{m\times m} - AX\|_F\,,$$

was die Namesgebung (Approximation einer Inversen für eine nicht invertierbare Matrix) unterstreicht.

Ist Rang $A = n$, so kann man das Ausgleichsproblem mit Hilfe der Normalengleichungen lösen, $x = (A^*A)^{-1}A^*b$. In diesem Fall ist also

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

insbesondere $A^+ = A^{-1}$ für eine quadratische, invertierbare Matrix A.

Bezeichnen $\{u_1, \ldots, u_m\}$ und $\{v_1, \ldots, v_n\}$ die orthonormalen Basen aus den Spalten der unitären Matrizen U bzw. V, so lässt sich die lineare Abbildung $b \mapsto x = A^+b$ in der faktorisierten Form

$$x = \sum_{\ell=1}^{r} \frac{1}{s_\ell}\,(u_\ell^* b)v_\ell$$

darstellen. Daraus folgt insbesondere, dass

$$\text{Kern}\,A^+ = \text{span}\{u_{r+1}, \ldots, u_m\}, \quad \text{Bild}\,A^+ = \text{span}\{v_1, \ldots, v_r\}$$

und $\|A^+\|_2 = 1/s_r$.

Beweis

(i) Ausgleichsproblem:
Invarianz der Euklidischen Norm unter unitären Transformationen $\Longrightarrow$

$$|Ax - b| = |U^*(Ax - b)|$$

setze

$$c = U^* b, \quad y = V^* x \quad \Longleftrightarrow \quad x = Vy$$

$U^* Ax = U^*(USV^*)Vy = Sy \quad \rightsquigarrow \quad$ äquivalentes Minimierungsproblem

$$|Sy - c| = \left| \begin{pmatrix} s_1 y_1 - c_1 \\ \vdots \\ s_r y_r - c_r \\ -c_{r+1} \\ \vdots \\ -c_m \end{pmatrix} \right| \to \min, \quad S = U^* AV$$

Lösung der ersten r Gleichungen $\quad \rightsquigarrow \quad$ Minimum

$$y_k = c_k / s_k, \; k = 1, \ldots r$$

Normtreue der unitären Matrix V ($|x| = |y|$) $\quad \rightsquigarrow$

$$y_{r+1} = \cdots = y_n = 0$$

für die Lösung minimaler Norm
Insgesamt folgt

$$y = S^+ c, \quad s^+_{j,k} = \begin{cases} 1/s_j & \text{für } j = k \le r \\ 0 & \text{sonst} \end{cases}$$

d.h.

$$x = Vy = VS^+ c = \underbrace{VS^+ U^*}_{A^+} b$$

(ii) Approximation der Inversen:

$$\|E_{m\times m} - AX\|_F^2 = \sum_k |e_k - Ax_k|^2$$

mit e_k dem k-ten Einheitsvektor und x_k der k-tex Spalte von X
minimal für $x_k = A^+ e_k$, d.h. x_k ist die k-te Spalte von A^+
$\Longrightarrow \quad X = A^+$

Beispiel

Lösung des Ausgleichproblems $|Ax - b| \to \min$ für

$$A = \begin{pmatrix} 2 & -4 & 5 \\ 6 & 0 & 3 \\ 2 & -4 & 5 \\ 6 & 0 & 3 \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ 3 \\ -1 \\ 3 \end{pmatrix}$$

(i) Berechnung der Singulärwert-Zerlegung:

$$A^{\mathrm{t}}A = \begin{pmatrix} 80 & -16 & 56 \\ -16 & 32 & -40 \\ 56 & -40 & 68 \end{pmatrix}$$

Eigenwerte und Eigenvektoren

$$s_1^2 = 144,\ s_2^2 = 36,\ s_3^2 = 0, \qquad V = \frac{1}{3} \begin{pmatrix} 2 & 2 & -1 \\ -1 & 2 & 2 \\ 2 & -1 & 2 \end{pmatrix}$$

$$\rightsquigarrow \quad S = \begin{pmatrix} 12 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad AV = \begin{pmatrix} 6 & -3 & 0 \\ 6 & 3 & 0 \\ 6 & -3 & 0 \\ 6 & 3 & 0 \end{pmatrix} = US$$

Spalten 1 und 2 von U: Division der entsprechenden Spalten von AV durch die singulären Werte s_1 und s_2 $\quad \rightsquigarrow \quad$ Einheitsvektoren
restliche Spalten: Ergänzen zu einer orthonormalen Basis

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

(ii) Pseudo-Inverse und Ausgleichslösung:
Zerlegung $A = USV^* \quad \rightsquigarrow$

$$A^+ = V \begin{pmatrix} \frac{1}{12} & 0 & 0 & 0 \\ 0 & \frac{1}{6} & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} U^* = \frac{1}{72} \begin{pmatrix} -2 & 6 & -2 & 6 \\ -5 & 3 & -5 & 3 \\ 4 & 0 & 4 & 0 \end{pmatrix}$$

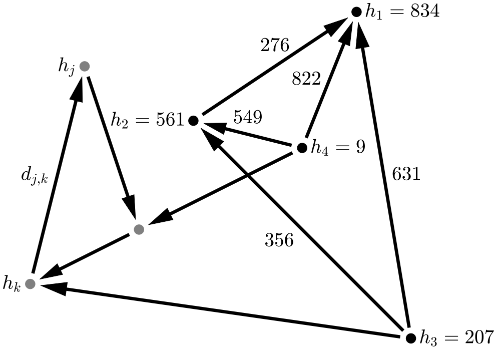
Lösung minimaler Norm: $x = A^+ b = \frac{1}{4}(2, 1, 0)^{\mathrm{t}}$

Beispiel

Kontrolle topographischer Höhendaten h_j durch Messung von Höhendifferenzen $d_{j,k}$

Messfehler $\rightsquigarrow$

$$d_{j,k} \neq h_j - h_k$$



Höhenkorrekturen x_j durch Lösen des Ausgleichsproblems

$$\sum_{(j,k)} \Big(d_{j,k} - \big((h_j + x_j) - (h_k + x_k)\big)\Big)^2 \to \min$$

Höhen und Differenzwerte des abgebildeten Beispiels

$$h = (834, 561, 207, 9)^{\mathrm{t}}, \quad d = (276, 631, 822, 356, 549)^{\mathrm{t}}$$

überbestimmtes System

$$A(h+x) = d, \quad A = \begin{pmatrix} 1 & -1 & 0 & 0 \\ 1 & 0 & -1 & 0 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{pmatrix}, \quad d = \begin{pmatrix} d_{1,2} \\ d_{1,3} \\ d_{1,4} \\ d_{2,3} \\ d_{2,4} \end{pmatrix}$$

Lösung minimaler Norm

$$x = A^+(d - Ah) = \frac{1}{8} \begin{pmatrix} 2 & 2 & 2 & 0 & 0 \\ -2 & 0 & 0 & 2 & 2 \\ 0 & -3 & 1 & -3 & 1 \\ 0 & 1 & -3 & 1 & -3 \end{pmatrix} \begin{pmatrix} 3 \\ 4 \\ -3 \\ 2 \\ -3 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ -3 \\ 3 \end{pmatrix}$$

Minimum-Norm-Lösung $\rightsquigarrow$ kleinstmögliche Korrekturen

Spiegelung

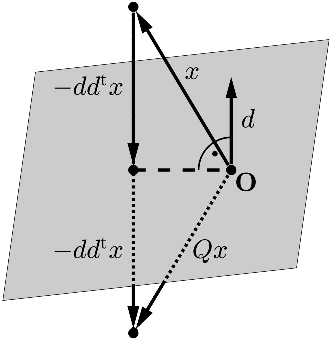
Eine Spiegelung an einer Hyperebene

$$H: \; d^{\mathrm{t}}x = 0\,,$$

mit normiertem Normalenvektor $d \in \mathbb{R}^n$ ($|d| = 1$) wird durch die symmetrische orthogonale Matrix

$$Q = E - 2dd^{\mathrm{t}}$$

mit E der Einheitsmatrix beschrieben.



Die Spiegelung lässt Vektoren in H invariant und ändert bei Vektoren parallel zu d das Vorzeichen:

$$Qx = x,\; x \perp d, \quad Qx = -x,\; x \parallel d\,.$$

Beweis

(i) Symmetrie von Q ✓

(ii) Orthogonalität:

$$QQ^{\mathrm{t}} = Q^2 = (E - 2dd^{\mathrm{t}})^2 = E^2 - 4dd^{\mathrm{t}} + 4d\,\underbrace{d^{\mathrm{t}}d}_{=1}\,d^{\mathrm{t}} = E$$

(iii) Spiegelungseigenschaft:
$x - Qx \parallel d$:

$$x - Qx = x - x + 2dd^{\mathrm{t}}x = 2(d^{\mathrm{t}}x)d$$

Mittelpunkt zwischen x und $Qx \in H$:

$$d^{\mathrm{t}}(x + Qx)/2 = (d^{\mathrm{t}}x + d^{\mathrm{t}}x - 2\,\underbrace{d^{\mathrm{t}}d}_{=1}\,d^{\mathrm{t}}x)/2 = 0$$

$\implies$ Qx ist Spiegelbild von x an der Hyperebene durch den Ursprung, die zu d orthogonal ist

Beispiel

Spiegelung an der Ebene H durch die Punkte

$$(0,0,0),\ (0,1,0),\ (3,0,4)$$

Normale

$$d = (-4,0,3)^{\mathrm{t}}/5, \quad |d| = 1$$

$\rightsquigarrow$ Spiegelungsmatrix

$$\begin{aligned} Q &= E - 2dd^{\mathrm{t}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} - \frac{2}{25} \begin{pmatrix} -4 \\ 0 \\ 3 \end{pmatrix} \begin{pmatrix} -4 & 0 & 3 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} - \frac{2}{25} \begin{pmatrix} 16 & 0 & -12 \\ 0 & 0 & 0 \\ -12 & 0 & 9 \end{pmatrix} = \begin{pmatrix} -0.28 & 0 & .96 \\ 0 & 1 & 0 \\ .96 & 0 & 0.28 \end{pmatrix} \end{aligned}$$

Drehung

Die orthogonale $n \times n$-Matrix

$$
Q: \quad \begin{array}{r}
\\ \\ \text{Zeile } j \ \rightarrow \\ \\ \text{Zeile } k \ \rightarrow \\ \\ \\
\end{array}
\begin{pmatrix}
1 & & & & & & 0 \\
& \ddots & & & & & \\
& & \cos\varphi & & -\sin\varphi & & \\
& & & \ddots & & & \\
& & \sin\varphi & & \cos\varphi & & \\
& & & & & \ddots & \\
0 & & & & & & 1
\end{pmatrix}
$$

beschreibt eine Drehung um den Winkel φ in der $x_j x_k$-Ebene des $\mathbb{R}^n$.

Jede orthogonale Matrix Q mit $\det Q = 1$ ist als Produkt solcher Drehungen in Koordinatenebenen darstellbar:

$$Q = \prod_{j<k} Q_{j,k} \, .$$

Beweis

(i) Orthogonalität:

$c = \cos\varphi,\ s = \sin\varphi, \quad c^2 + s^2 = 1 \quad \Longrightarrow$

$$QQ^{\mathrm{t}} = \begin{pmatrix} 1 & & & & & & 0 \\ & \ddots & & & & & \\ & & c^2+s^2 & & -sc+sc & & \\ & & & \ddots & & & \\ & & -sc+sc & & c^2+s^2 & & \\ & & & & & \ddots & \\ 0 & & & & & & 1 \end{pmatrix} = E$$

$\rightsquigarrow$ Orthogonalität der Drehmatrix Q

(ii) Herleitung der Faktorisierung für $n = 3$:
Bestimmung der Drehungen durch sukzessives Annulieren der Elemente $q_{2,1}, q_{3,1}, q_{3,2}$ von Q:

$$\begin{aligned} Q_{1,2}^{-1} Q &= \begin{pmatrix} \star & \star & \star \\ 0 & \star & \star \\ \star & \star & \star \end{pmatrix} \\ Q_{1,3}^{-1} Q_{1,2}^{-1} Q &= \begin{pmatrix} 1 & \star & \star \\ 0 & \star & \star \\ 0 & \star & \star \end{pmatrix} \\ Q_{2,3}^{-1} Q_{1,3}^{-1} Q_{1,2}^{-1} Q &= \begin{pmatrix} 1 & \star & \star \\ 0 & 1 & \star \\ 0 & 0 & \star \end{pmatrix} = R \end{aligned}$$

Die Drehungen $Q_{2,3}$, $Q_{1,3}$, $Q_{1,2}$ bilden jeweils Vektoren $(v_1, v_2)^{\mathrm{t}}$ auf Vielfache von Einheitsvektoren ab.

$\det Q_{j,k} = 1$, $\det Q = 1 \quad \Longrightarrow \quad \det R = 1 \quad \Longrightarrow \quad r_{3,3} = 1$

Normierung der Spalten $\quad \Longrightarrow \quad R = E$ und $Q = Q_{1,2} Q_{1,3} Q_{2,3}$

Beispiel

Faktorisierung der Drehmatrix

$$Q = \begin{pmatrix} 0 & \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ -\frac{\sqrt{2}}{2} & -\frac{\sqrt{6}}{4} & -\frac{\sqrt{2}}{4} \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{6}}{4} & \frac{\sqrt{2}}{4} \end{pmatrix}$$

Drehung $D_z^{-1} = Q_{1,2}^{-1}$ um $\frac{\pi}{2}$ um die z-Achse:

$$D_z^{-1} Q = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} Q = \begin{pmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{6}}{4} & \frac{\sqrt{2}}{4} \\ 0 & \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{6}}{4} & \frac{\sqrt{2}}{4} \end{pmatrix}$$

Drehung $D_y^{-1} = Q_{1,3}^{-1}$ um $\frac{\pi}{4}$ um die y-Achse:

$$\underbrace{\begin{pmatrix} \frac{\sqrt{2}}{2} & 0 & -\frac{\sqrt{2}}{2} \\ 0 & 1 & 0 \\ \frac{\sqrt{2}}{2} & 0 & \frac{\sqrt{2}}{2} \end{pmatrix}}_{D_y^{-1}} \underbrace{\begin{pmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{6}}{4} & \frac{\sqrt{2}}{4} \\ 0 & \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{6}}{4} & \frac{\sqrt{2}}{4} \end{pmatrix}}_{D_z^{-1}Q} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ 0 & \frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix}$$

Drehung $D_x^{-1} = Q_{2,3}$ um $-\frac{\pi}{3}$ um die x-Achse

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{\sqrt{3}}{2} \\ 0 & -\frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix} D_y^{-1} D_z^{-1} Q = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = E$$

$D_x^{-1} D_y^{-1} D_z^{-1} Q = E,\ D^{-1} = D^{\mathrm{t}} \quad \rightsquigarrow \quad$ Faktorisierung

$$\begin{aligned} Q &= (D_x^{-1} D_y^{-1} D_z^{-1})^{-1} = D_z D_y D_x \\ &= \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{\sqrt{2}}{2} & 0 & \frac{\sqrt{2}}{2} \\ 0 & 1 & 0 \\ -\frac{\sqrt{2}}{2} & 0 & \frac{\sqrt{2}}{2} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ 0 & \frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix} \end{aligned}$$

Drehung im Raum

Eine Drehung im $\mathbb{R}^3$ mit normierter Drehachsenrichtung u und Drehwinkel φ, orientiert wie eine Rechtsschraube, bildet einen Vektor x auf

$$Qx = \cos\varphi\, x + (1-\cos\varphi)uu^{\mathrm{t}}x + \sin\varphi\, u \times x$$

ab, wobei $u \times x$ das Kreuzprodukt von u und x bezeichnet.
Die entsprechende Drehmatrix ist

$$Q: \quad q_{j,\ell} \;=\; \cos\varphi\,\delta_{j,\ell} + (1-\cos\varphi)\,u_j u_\ell + \sin\varphi \sum_{\ell} \varepsilon_{j,k,\ell} u_k$$

mit dem Kroneckersymbol $\delta_{j,\ell}$ und dem ε-Tensor $\varepsilon_{j,k,\ell}$ bzw.

$$Q = \cos\varphi \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + (1-\cos\varphi) \begin{pmatrix} u_1u_1 & u_1u_2 & u_1u_3 \\ u_2u_1 & u_2u_2 & u_2u_3 \\ u_3u_1 & u_3u_2 & u_3u_3 \end{pmatrix} + \sin\varphi \begin{pmatrix} 0 & -u_3 & u_2 \\ u_3 & 0 & -u_1 \\ -u_2 & u_1 & 0 \end{pmatrix}.$$

Beweis

zeige:

$Qu = u$ und Q dreht einen zu u orthogonalen Vektor v um einen Winkel φ um die Achse u.

(i) Bild von u:

$$Qu \;=\; \cos\varphi\, u \,+\, (1-\cos\varphi) u \underbrace{u^{\mathrm{t}} u}_{=1} + \sin\varphi \; \underbrace{u \times u}_{=(0,0,0)^{\mathrm{t}}} \;=\; u$$

(ii) Bild von v:

$$Qv = \cos\varphi\, v \,+\, (0,0,0)^{\mathrm{t}} \,+\, \sin\varphi\, \underbrace{u \times v}_{=w}, \quad w \perp v,\ |w| = 1$$

$\iff$ Drehung um φ in der von v und $u \times v$ aufgespannten Ebene

Beispiel

Matrix Q einer Drehung um $\varphi = \frac{\pi}{3}$ um die Achse $u = \frac{1}{\sqrt{3}}(1\ 1\ 1)^{\mathrm{t}}$

$\cos(\pi/3) = 1/2$, $\sin(\pi/3) = \sqrt{3}/2 \quad \rightsquigarrow$

$$\cos\varphi\ \delta_{j,\ell} \quad : \quad \begin{pmatrix} \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{2} \end{pmatrix}$$

$$(1-\cos\varphi)u_j u_\ell \quad : \quad \frac{1}{6}\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

$$\sin\varphi \sum_k \varepsilon_{j,k,\ell} u_k \quad : \quad \frac{1}{2}\begin{pmatrix} 0 & \varepsilon_{1,3,2}u_3 & \varepsilon_{1,2,3}u_2 \\ \varepsilon_{2,3,1}u_3 & 0 & \varepsilon_{2,1,3}u_1 \\ \varepsilon_{3,2,1}u_2 & \varepsilon_{3,1,2}u_1 & 0 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 0 & -1 & 1 \\ 1 & 0 & -1 \\ -1 & 1 & 0 \end{pmatrix}$$

Summe der drei Matrizen $\quad \rightsquigarrow$

$$Q = \frac{1}{3}\begin{pmatrix} 2 & -1 & 2 \\ 2 & 2 & -1 \\ -1 & 2 & 2 \end{pmatrix}$$

Drehachse und Drehwinkel

Jede Drehung Q im $\mathbb{R}^3$ besitzt eine Drehachse, d.h. lässt einen Einheitsvektor u invariant, und entspricht einer ebenen Drehung um einen Winkel φ in der zu u orthogonalen Ebene.

Bezüglich eines orthonormalen Rechtssystems u, v, w besitzt Q die Matrixdarstellung

$$\tilde{Q} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\varphi & -\sin\varphi \\ 0 & \sin\varphi & \cos\varphi \end{pmatrix}.$$

Insbesondere gilt für den Drehwinkel

$$\cos\varphi = \frac{1}{2}(\operatorname{Spur} Q - 1).$$

Beweis

Orthogonalität der Drehmatrix $Q \quad \Longrightarrow$

$$Q^{-1} = Q^{\mathrm{t}}, \quad |\det\, Q| = 1$$

$|\lambda_k| = 1$ und $\lambda_1\lambda_2\lambda_3 = \det\, Q = 1 \quad \Longrightarrow \quad \exists$ Eigenwert $\lambda = 1$,
denn bei geeigneter Numerierung gilt

$$\lambda_1 = \overline{\lambda_2}, \quad \lambda_1\lambda_2 = |\lambda_1|^2 = 1$$

oder

$$\lambda_k \in \{-1, 1\}$$

Drehachse u: normierter Eigenvektor u zum Eigenwert 1 orthonormales Rechtssystem $u, v, w \quad \rightsquigarrow$

$$\begin{aligned} Qu &= u \\ Qv &= \alpha v + \beta w \\ Qw &= \gamma v + \delta w \end{aligned}$$

Qv, Qw haben keine u-Komponente, wegen der Winkeltreue orthogonaler Matrizen:

$$x \perp u \quad \Longrightarrow \quad Qx \perp Qu = u$$

Matrixform obiger Gleichungen

$$Q(\underbrace{u, v, w}_{P}) = (u, v, w) \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & \alpha & \gamma \\ 0 & \beta & \delta \end{pmatrix}}_{\tilde{Q}}$$

$\tilde{Q} = P^{-1}QP$ orthogonal mit $\det \tilde{Q} = \det Q = 1 \quad \Longrightarrow$

$$\begin{pmatrix} \alpha & \gamma \\ \beta & \delta \end{pmatrix} = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}$$

Invarianz der Spur unter Ähnlichkeitstransformationen $\quad \Longrightarrow$

$$\operatorname{Spur} Q = \operatorname{Spur} \tilde{Q} = 1 + 2\cos\varphi$$

Beispiel

Bestimmung von Drehachse und Drehwinkel für die Drehmatrix

$$Q = \frac{1}{2} \begin{pmatrix} 1 & -\sqrt{2} & 1 \\ \sqrt{2} & 0 & -\sqrt{2} \\ 1 & \sqrt{2} & 1 \end{pmatrix}$$

(i) Überprüfung der Orthogonalität und der Determinante:

$$Q^{\mathrm{t}} Q = \frac{1}{4} \begin{pmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{pmatrix} = E \quad \Longleftrightarrow \quad Q^{\mathrm{t}} = Q^{-1}$$

und

$$\det Q = \frac{1}{8} \det \begin{pmatrix} 1 & -\sqrt{2} & 1 \\ \sqrt{2} & 0 & -\sqrt{2} \\ 1 & \sqrt{2} & 1 \end{pmatrix} = +1$$

(ii) Drehachse:
Eigenvektor u zum Eigenwert $\lambda = 1$

$$\underbrace{\frac{1}{2}\begin{pmatrix} -1 & -\sqrt{2} & 1 \\ \sqrt{2} & -2 & -\sqrt{2} \\ 1 & \sqrt{2} & -1 \end{pmatrix}}_{Q-E} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad \Longrightarrow \quad u = \begin{pmatrix} \sqrt{2}/2 \\ 0 \\ \sqrt{2}/2 \end{pmatrix}$$

(iii) Drehwinkel:

$$\cos\varphi = \frac{1}{2}(\mathrm{Spur}\, Q - 1) = \frac{1}{2}(1-1) = 0$$

$\Longrightarrow \quad \varphi = \pm\pi/2$

(iv) Orientierung:
Das Vorzeichen von φ hängt von der Orientierung der Drehachsenrichtung u ab und kann mit Hilfe eines Rechtssystems $\{u, v, w\}$ bestimmt werden:

$$w^{\mathrm{t}} Q v = w^{\mathrm{t}}(\cos\varphi v + \sin\varphi w) = \sin\varphi$$

wähle als Rechtssystem

$$u = \begin{pmatrix} \sqrt{2}/2 \\ 0 \\ \sqrt{2}/2 \end{pmatrix}, \quad u \perp v = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad w = u \times v = \begin{pmatrix} -\sqrt{2}/2 \\ 0 \\ \sqrt{2}/2 \end{pmatrix}$$

Bilden des Produktes $w^{\mathrm{t}} Q v \quad \rightsquigarrow$

$$\sin\varphi = (-\sqrt{2}/2,\, 0,\, \sqrt{2}/2) \underbrace{\begin{pmatrix} -\sqrt{2}/2 \\ 0 \\ \sqrt{2}/2 \end{pmatrix}}_{Qv} = 1\,,$$

d.h. $\varphi = \pi/2$

Quadrik

Für eine symmetrische $n \times n$-Matrix A, einen Vektor b und eine Konstante c ist

$$Q:\ x^{\mathrm{t}}Ax + 2b^{\mathrm{t}}x + c = \sum_{j,k} a_{j,k}x_j x_k + \sum_k b_k x_k + c = 0$$

eine Quadrik im $\mathbb{R}^n$. Mit

$$\tilde{A} = \left(\begin{array}{c|c} c & b^{\mathrm{t}} \\ \hline b & A \end{array}\right), \quad \tilde{x} = \left(\begin{array}{c} 1 \\ \hline x \end{array}\right)$$

kann man Q auch in der homogenen Form

$$Q:\ \tilde{x}^{\mathrm{t}}\tilde{A}\tilde{x} = 0$$

darstellen.

Man unterscheidet die folgenden Typen:

- kegelige Quadrik: $\operatorname{Rang}\tilde{A} = \operatorname{Rang}A$,
- Mittelpunktsquadrik: $\operatorname{Rang}\tilde{A} = \operatorname{Rang}A + 1$,
- parabolische Quadrik: $\operatorname{Rang}\tilde{A} = \operatorname{Rang}A + 2$.

Beispiel

Typ der parameterabhängigen Quadrik

$$Q: \quad x_1^2 + \lambda x_2^2 + 2x_2 + \lambda = 0, \quad \lambda \in \mathbb{R}$$

Matrixschreibweise

$$Q: \quad x^{\mathrm{t}}Ax + 2b^{\mathrm{t}}x + c = 0 \quad \Longleftrightarrow \quad \tilde{x}^{\mathrm{t}}\tilde{A}\tilde{x} = 0,\, \tilde{x}^{\mathrm{t}} = (1, x^{\mathrm{t}})$$

mit

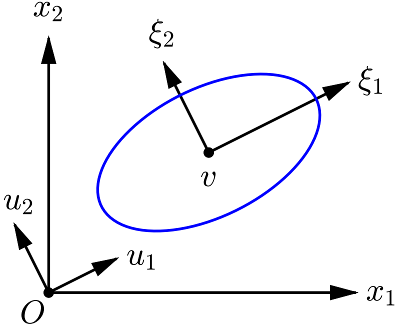
$$A = \begin{pmatrix} 1 & 0 \\ 0 & \lambda \end{pmatrix}, \quad b = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad c = \lambda\,, \quad \tilde{A} = \begin{pmatrix} \lambda & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & \lambda \end{pmatrix}$$

Q ist für $\lambda = 0$ eine parabolische ($\operatorname{Rang} \tilde{A} = \operatorname{Rang} A + 2 = 3$), für $\lambda = \pm 1$ eine kegelige ($\operatorname{Rang} \tilde{A} = \operatorname{Rang} A = 2$) und für alle anderen Werte von λ eine Mittelpunktsquadrik ($\operatorname{Rang} \tilde{A} = \operatorname{Rang} A + 1 = 3$).

Durch eine Drehung U und Verschiebung v, $\xi \mapsto x = U\xi + v$, kann eine Quadrik Q im $\mathbb{R}^n$ auf Diagonalform transformiert werden:

$$Q:\ x^{\mathrm{t}}Ax + 2b^{\mathrm{t}}x + c = 0 \quad \rightarrow \quad \sum_{k=1}^{m} \lambda_k \xi_k^2 + 2\tilde{b}^{\mathrm{t}}\xi + \gamma = 0$$

mit m dem Rang der symmetrischen $n \times n$-Matrix A. Dabei ist $\tilde{b}^{\mathrm{t}} = (0, \ldots, 0)$ für $m = n$ (kein Eigenwert 0) und $2\tilde{b}^{\mathrm{t}}\xi = 2\beta\xi_{m+1}$, $\beta\gamma = 0$ für $m < n$.



Die Spalten der Drehmatrix U enthalten normierte Eigenvektoren u_k zu den Eigenwerten λ_k von A, und die Geraden $g_k : x = v + tu_k$ werden als Hauptachsen bezeichnet. Der Verschiebungsvektor v ist der Mittelpunkt der Quadrik.

Durch Skalierung erhält man eine der Normalformen

$$Q : \sum_{k=1}^{m} \sigma_k \frac{\xi_k^2}{a_k^2} = 1 \qquad (m \leq n)$$

oder

$$Q : \sum_{k=1}^{m} \sigma_k \frac{\xi_k^2}{a_k^2} = 2\xi_{m+1} \qquad (m < n)$$

je nachdem, ob A maximalen Rang hat und β oder γ Null ist. Dabei ist $\sigma_k \in \{-1, 1\}$ und $a_k > 0$ sind die Hauptachsenlängen.

Beweis

Bestimmung von U und v durch sukzessive Variablentransformationen

(i) Drehung:

$A^{\mathrm{t}} = A \quad \Longrightarrow$

$\exists$ orthonormale Basis aus Eigenvektoren u_k mit $\det(u_1, \ldots, u_n) = 1$

Substitution $x = Uy = (u_1, \ldots, u_n)y \quad \rightsquigarrow \quad$ Diagonalform

$$
\begin{aligned}
0 &= y^{\mathrm{t}} \underbrace{U^{\mathrm{t}} A U}_{\text{diagonal}} y + 2b^{\mathrm{t}} U y + c \\
&= y^{\mathrm{t}} \operatorname{diag}(\lambda_1, \ldots, \lambda_m, 0, \ldots, 0) y + 2(\underbrace{U^{\mathrm{t}} b}_{\tilde{b}})^{\mathrm{t}} y + c
\end{aligned}
$$

Ist $m + 1 < n$ (Eigenwert 0 mit Vielfachheit > 1, seltener Fall), so ist die Basis $\{u_{m+1}, \ldots, u_n\}$ für den Eigenraum V_0 so gewählt, dass $\tilde{b}_k = 0$ für $k > m + 1$.

Wähle dazu für eine gegebene Basis $\{\tilde{u}_{m+1}, \ldots, \tilde{u}_n\}$ von V_0 eine $(n-m) \times (n-m)$-Drehmatrix $\tilde{U}$, so dass

$$\underbrace{\tilde{U} \begin{pmatrix} \tilde{u}_{m+1}^{\mathrm{t}} \\ \tilde{u}_{m+2}^{\mathrm{t}} \\ \vdots \\ \tilde{u}_n^{\mathrm{t}} \end{pmatrix}}_{=W} \begin{pmatrix} b_{m+1} \\ b_{m+2} \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} \tilde{b}_{m+1} \\ 0 \\ \vdots \\ 0 \end{pmatrix} .$$

Die Zeilen $u_{m+1}^{\mathrm{t}}, \ldots, u_n^{\mathrm{t}}$ von W bilden dann die gewünschte Basis.

(ii) Verschiebung:

Quadratische Ergänzung für $k \leq m$ ($\lambda_k \neq 0$):

$$y_k = \xi_k - \tilde{b}_k/\lambda_k$$

$$\implies \quad \lambda_k y_k^2 = \lambda_k \xi_k^2 - 2\xi_k \tilde{b}_k + \tilde{b}_k^2/\lambda_k,\ 2\tilde{b}_k y_k = 2\tilde{b}_k \xi_k - 2\tilde{b}_k^2/\lambda_k$$

Summation für $k = 1, \ldots, m \quad \rightsquigarrow \quad$ Elimination der linearen Terme und Änderung der Konstante

$$c \to -\gamma = c - \sum_{k=1}^{m} \tilde{b}_k^2 / \lambda_k$$

- $m = n$ (Rang $A = n$, kein Eigenwert 0):

$$Q : \sum_{k=1}^{n} \lambda_k \xi_k^2 = \gamma$$

$$x = U \underbrace{(\xi - (\tilde{b}_1/\lambda_1, \ldots, \tilde{b}_n/\lambda_n)^{\mathrm{t}})}_{y} \quad \rightsquigarrow \quad v = -U(\tilde{b}_1/\lambda_1, \ldots, \tilde{b}_n/\lambda_n)^{\mathrm{t}})$$

- $m < n$ (Eigenwert 0 mit Vielfachheit $n - m$):

$$Q: \sum_{k=1}^{m} \lambda_k \xi_k^2 + 2\tilde{b}_{m+1} y_{m+1} = \gamma$$

Falls $\tilde{b}_{m+1} = 0$, keine weitere Verschiebung, d.h. $\xi_k = y_k$, $k = m+1, \ldots, n$, und $v = -U(\tilde{b}_1/\lambda_1, \ldots, \tilde{b}_m/\lambda_m, 0, \ldots, 0)^{\mathrm{t}})$
Falls $\tilde{b}_{m+1} \neq 0$, so wird γ durch Setzen von $y_{m+1} = \xi_{m+1} + \gamma/(2\tilde{b}_{m+1})$ eliminiert:

$$Q: \sum_{k=1}^{m} \lambda_k \xi_k^2 + 2\beta\xi_{m+1} = 0, \quad \beta = \tilde{b}_{m+1}\,.$$

$$\rightsquigarrow \quad v = -U(\tilde{b}_1/\lambda_1, \ldots, \tilde{b}_m/\lambda_m, -2\gamma/\tilde{b}_{m+1}, 0, \ldots, 0)^{\mathrm{t}})$$

(iii) Skalierung:
Division durch γ oder $-\beta$ $\quad\rightsquigarrow\quad$ Normalformen

$$Q:\ \sum_{k=1}^{m}\sigma_k\frac{\xi_k^2}{a_k^2}=\gamma,\quad \sigma_k/a_k^2=\lambda_k/\gamma \qquad (m\le n)\,,$$

oder

$$Q:\ \sum_{k=1}^{m}\sigma_k\frac{\xi_k^2}{a_k^2}=2\xi_{m+1},\quad \sigma_k/a_k^2=-\lambda_k/\beta \qquad (m< n)\,.$$

Beispiel

Normalform der Quadrik

$$Q:\ 5x_1^2+5x_2^2+5x_3^2+6x_1x_2+8x_2x_3+2x_2=0$$

(i) Matrixschreibweise:

$$Q:\ x^{\mathrm{t}}Ax+2b^{\mathrm{t}}x+c=0$$

mit

$$A=\begin{pmatrix}5&3&0\\3&5&4\\0&4&5\end{pmatrix},\quad b=\begin{pmatrix}0\\1\\0\end{pmatrix},\quad c=0$$

beachte: $a_{1,2}x_1x_2+a_{2,1}x_2x_1=6x_1x_2 \quad\Longrightarrow\quad a_{1,2}=a_{2,1}=3$, etc.

(i) Eigenwerte von A:
charakteristisches Polynom

$$\begin{vmatrix} 5-\lambda & 3 & 0 \\ 3 & 5-\lambda & 4 \\ 0 & 4 & 5-\lambda \end{vmatrix} = (5-\lambda)^3 - 16(5-\lambda) - 9(5-\lambda) = (5-\lambda)((5-\lambda)^2 - 25$$

$\rightsquigarrow$ Eigenwerte $\lambda_1 = 10$, $\lambda_2 = 5$, $\lambda_3 = 0$

(iii) Eigenvektoren:

$$\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 5-10 & 3 & 0 \\ 3 & 5-10 & 4 \\ 0 & 4 & 5-10 \end{pmatrix} u_1 \quad \Longrightarrow \quad u_1 \parallel \begin{pmatrix} 3 \\ 5 \\ 4 \end{pmatrix}$$

analog

$$u_2 \parallel (4, 0, -3)^{\mathrm{t}}, \quad u_3 \parallel (3, -5, 4)^{\mathrm{t}}$$

Determinante der Matrix aus normierten Eigenvektoren

$$\begin{vmatrix} 3/\sqrt{50} & 4/5 & 3/\sqrt{50} \\ 5/\sqrt{50} & 0 & -5/\sqrt{50} \\ 4/\sqrt{50} & -3/5 & 4/\sqrt{50} \end{vmatrix} \underset{\text{Sarrus}}{=} -\frac{80}{250} - \frac{45}{250} - \frac{45}{250} - \frac{80}{250} = -1$$

Multiplikation eines Eigenvektors (z.B. u_2) mit -1 $\rightsquigarrow$ Drehmatrix (Determinante $= 1$)

$$U = \frac{1}{5} \begin{pmatrix} 3/\sqrt{2} & -4 & 3/\sqrt{2} \\ 5/\sqrt{2} & 0 & -5/\sqrt{2} \\ 4/\sqrt{2} & 3 & 4/\sqrt{2} \end{pmatrix}$$

(iv) Diagonalisierung:

$x = Uy$ $\rightsquigarrow$

$$Q:\; y^{\mathrm{t}} U^{\mathrm{t}} A U y + 2 b^{\mathrm{t}} U y = 0$$

mit

$$U^{\mathrm{t}} A U = \operatorname{diag}(10, 5, 0), \quad b^{\mathrm{t}} U = (0, 1, 0) U = (1/\sqrt{2}, 0, -1/\sqrt{2})\,,$$

d.h.

$$Q:\; 10 y_1^2 + 5 y_2^2 + \sqrt{2} y_1 - \sqrt{2} y_3 = 0$$

(v) Verschiebung, quadratische Terme:
$10y_1^2 + \sqrt{2}y_1 = 10(\underbrace{y_1 + \sqrt{2}/20}_{\xi_1})^2 - \underbrace{10(\sqrt{2}/20)^2}_{1/20}$, $y_2 = \xi_2 \quad \rightsquigarrow$

$$Q:\ 10\xi_1^2 + 5\xi_2^2 - \sqrt{2}y_3 - 1/20 = 0$$

(vi) Verschiebung, linearer Term:
$-\sqrt{2}y_3 - 1/20 = -\sqrt{2}(\underbrace{y_3 + 1/(20\sqrt{2})}_{\xi_3}) \quad \rightsquigarrow$

$$Q:\ 10\xi_1^2 + 5\xi_2^2 0 - \sqrt{2}\xi_3 = 0$$

(vii) Skalierung:
Multiplikation mit $\sqrt{2} \quad \rightsquigarrow \quad$ Normalform

$$10\sqrt{2}\xi_1^2 + 5\sqrt{2}\xi_2^2 = 2\xi_3$$

mit den Halbachsenlängen $a_1 = 1/\sqrt{10\sqrt{2}}$, $a_2 = 1/\sqrt{5\sqrt{2}}$

(viii) Transformation:
$x = Uy = U(\xi - (\sqrt{2}/20, 0, 1/(20\sqrt{2}))^{\mathrm{t}})$ mit dem Verschiebungsvektor

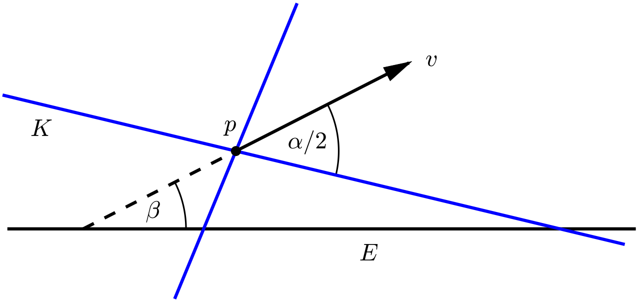
$$
\begin{aligned}
v &= \underbrace{\frac{1}{5}\begin{pmatrix} 3/\sqrt{2} & -4 & 3/\sqrt{2} \\ 5/\sqrt{2} & 0 & -5/\sqrt{2} \\ 4/\sqrt{2} & 3 & 4/\sqrt{2} \end{pmatrix}}_{U} \begin{pmatrix} -\sqrt{2}/20 \\ 0 \\ -1/(20\sqrt{2}) \end{pmatrix} \\
&= \frac{1}{5}\frac{1}{20}\begin{pmatrix} -3-3/2 \\ -5+5/2 \\ -4-4/2 \end{pmatrix} = -\frac{1}{200}\begin{pmatrix} 9 \\ 5 \\ 12 \end{pmatrix}
\end{aligned}
$$

Der Schnitt eines Doppelkegels

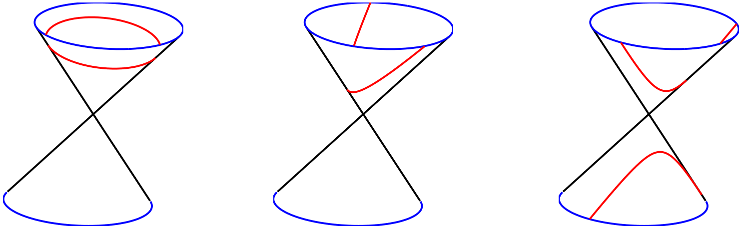
$$K:\ (x-p)^{\mathrm{t}}v = \pm\cos\frac{\alpha}{2}\,|x-p||v|$$

mit Spitze p ($p_3 \neq 0$), Richtung v und Öffnungswinkel α mit der Ebene $E:\ x_3 = 0$ ist eine quadratische Kurve

$$K \cap E:\ a_{1,1}x_1^2 + 2a_{1,2}x_1x_2 + a_{2,2}x_2^2 + 2b_1x_1 + 2b_2x_2 + c = 0\,.$$



Der Typ dieses Kegelschnitts hängt von der Größe des Winkels β zwischen E und der Symmetrieachse $g: p + tv$, $t \in \mathbb{R}$, des Doppelkegels ab.



Ellipse: $\beta > \alpha/2$ Parabel: $\beta = \alpha/2$ Hyperbel: $\beta < \alpha/2$

Beispiel

Schnitt des Doppelkegels mit Spitze $p = (0,0,1)^{\mathrm{t}}$, Achsenrichtung $v = (1,1,1)^{\mathrm{t}}$ und Öffnungswinkel $\alpha = pi/2$ mit der x_1x_2-Ebene

(i) Typ:
halber Öffnungswinkel: $\alpha/2 = \pi/4 = 0.7854$
Winkel β der Achsenrichtung mit der x_1x_2-Ebene:
$\cos(\pi/2 - \beta) = (0,0,1)v/|v| = \frac{1}{\sqrt{3}} \implies$
$\beta = \pi/2 - \arccos(1/\sqrt{3}) = 0.6155$
$\alpha/2 > \beta \quad \implies \quad$ Hyperbel

(ii) Gleichung des Kegelschnitts:
Einsetzen der Daten in die Gleichung $(x-p)^{\mathrm{t}}v = \pm\cos(\alpha/2)\,|x-p|\,|v|$, $\cos((\pi/2)/2) = 1/\sqrt{2}$ und Quadrieren $\quad \rightsquigarrow$

$$\left((x_1, x_2, x_3 - 1)(1,1,1)^{\mathrm{t}}\right)^2 = \frac{1}{2}\,|(x_1, x_2, x_3 - 1)|^2\,|v|^2$$

Schnitt mit der x_1x_2-Ebene $\widehat{=}\ x_3 = 1$, d.h.

$$(x_1 + x_2 - 1)^2 = \frac{1}{2}\,(x_1^2 + x_2^2 + 1) \cdot 3$$

bzw. nach Vereinfachung

$$x_1^2 + x_2^2 - 4x_1x_2 + 4x_1 + 4x_2 + 1 = 0$$

Matrix-Form

$$x^{\mathrm{t}}Ax + 2b^{\mathrm{t}}x + c = 0, \quad A = \begin{pmatrix} 1 & -2 \\ -2 & 1 \end{pmatrix},\ b = \begin{pmatrix} 2 \\ 2 \end{pmatrix},\ c = 1$$

(iii) Transformation auf Normalform:
Eigenwerte λ_k und normierte, ein Rechtssystem bildende Eigenvektoren u_k der Matrix A

$$\lambda_1 = 3,\ \lambda_2 = -1, \quad U = (u_1, u_2) = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$$

Substitution $x = Ay$ (Ähnlichkeitstransformation der Matrix A) $\quad\rightsquigarrow$

$$3y_1^2 - y_2^2 + 2\underbrace{(0, 2\sqrt{2})}_{b^{\mathrm{t}}U}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} + 1 = 0$$

Substitution $y_1 = \xi_1,\ y_2 = \xi_2 + 2\sqrt{2}$ $\quad\rightsquigarrow\quad$ Elimination des linearen Terms

$$3\xi_1^2 - \xi_2^2 \underbrace{-(2\sqrt{2})^2 + 1}_{=9} = 0$$

Division durch (-9) $\quad\rightsquigarrow\quad$ Normalform

$$-\frac{\xi_1^2}{3} + \frac{\xi_2^2}{9} = 1$$

Hyperbel mit Hauptachsenrichtungen u_1, u_2 und Hauptachsenlängen $a_1 = \sqrt{3}$, $a_2 = 3$

Euklidische Normalformen der zweidimensionalen Quadriken

Es existieren 6 (nicht-triviale) Typen ebener Quadriken mit den folgenden Normalformen:

- Kegelige Quadriken

Normalform	Bezeichnung
$\frac{x_1^2}{a_1^2} - \frac{x_2^2}{a_2^2} = 0$	schneidendes Geradenpaar
$\frac{x_1^2}{a_1^2} = 0$	Doppelgerade

- Parabolische Quadriken

Normalform	Bezeichnung
$\frac{x_1^2}{a_1^2} = 2x_2$	Parabel

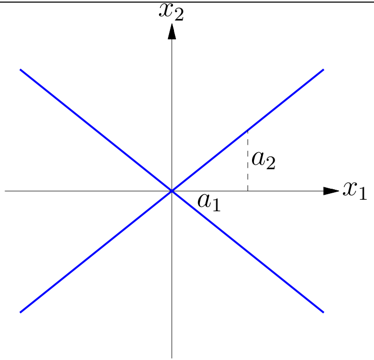
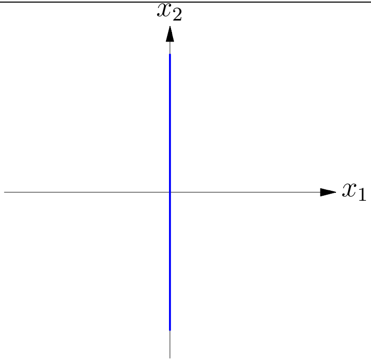
- Mittelpunktsquadriken

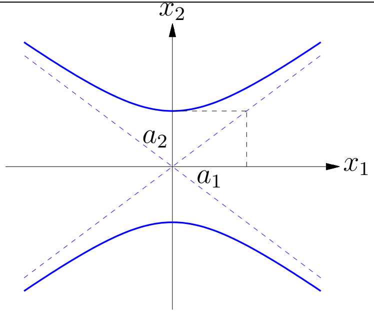
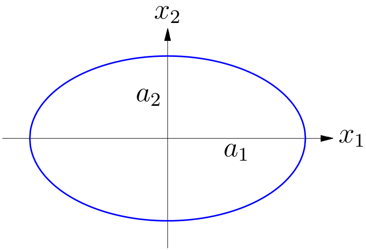
Normalform	Bezeichnung
$-\frac{x_1^2}{a_1^2} + \frac{x_2^2}{a_2^2} = 1$	Hyperbel
$\frac{x_1^2}{a_1^2} + \frac{x_2^2}{a_2^2} = 1$	Ellipse
$\frac{x_1^2}{a_1^2} = 1$	paralleles Geradenpaar

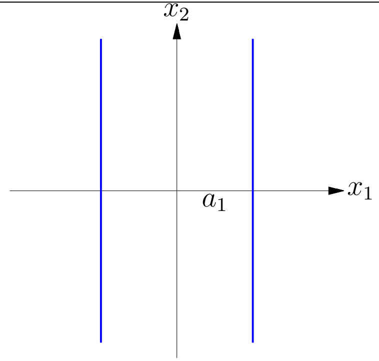
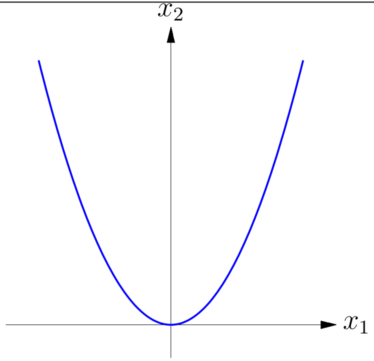
Die Größen $a_k > 0$ sind die Hauptachsenlängen der Quadrik.

Triviale Sonderfälle sind die Normalformen

- $\frac{x_1^2}{a_1^2} + \frac{x_2^2}{a_2^2} = 0$ (Punkt)
- $-\frac{x_1^2}{a_1^2} - \frac{x_2^2}{a_2^2} = 1$ (leere Menge)
- $-\frac{x_1^2}{a_1^2} = 1$ (leere Menge)

schneidendes Geradenpaar	Doppelgerade
	

Hyperbel	Ellipse
	

paralleles Geradenpaar	Parabel
	

Beispiel

Normalform und der Typ der Quadrik

$$Q:\ 3x_1^2 + 3x_2^2 + 10x_1x_2 - 14\sqrt{2}x_1 - 2\sqrt{2}x_2 - 18 = 0$$

(i) Matrixform:

$$0 = x^{\mathrm{t}}Ax + 2b^{\mathrm{t}}x + c = x^{\mathrm{t}}\begin{pmatrix} 3 & 5 \\ 5 & 3 \end{pmatrix} x + 2\sqrt{2}\,(-7,-1)\,x - 18$$

(ii) Eigenwerte:
charakteristisches Polynom

$$\begin{vmatrix} 3-\lambda & 5 \\ 5 & 3-\lambda \end{vmatrix} = (3-\lambda)^2 - 25$$

Nullstellen $\lambda_1 = -2$, $\lambda_2 = 8$

(iii) Eigenvektoren:
homogenes lineares Gleichungssystem

$$\det(A-(-2)E)u_1 = \begin{pmatrix} 5 & 5 \\ 5 & 5 \end{pmatrix} u_1 = 0$$

$\rightsquigarrow$ normierter Eigenvektor $u_1 = (1,-1)^{\mathrm{t}}/\sqrt{2}$ zu $\lambda_1 = -2$
normierter Eigenvektor u_2 zu $\lambda_2 = 8 \perp$ zu $u_1 \implies u_2 = (1,1)^{\mathrm{t}}/\sqrt{2}$,
wobei das Vorzeichen so gewählt ist, dass die Determinante der Transformationsmatrix

$$U = (u_1, u_2) = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$$

positiv ist (Drehmatrix)

(iv) Diagonalisierung:
Substitution $x = Uy \quad \rightsquigarrow$

$$\begin{aligned} 0 &= x^{\mathrm{t}}Ax + 2b^{\mathrm{t}}x + c \\ &= y^{\mathrm{t}}U^{\mathrm{t}}AUy + 2\left(b^{\mathrm{t}}U\right)y + c \\ &= -2y_1^2 + 8y_2^2 - 12y_1 - 16y_2 - 18 \end{aligned}$$

(v) Verschiebung:
quadratische Ergänzung $\xi = y + (3, -1)^{\mathrm{t}} \quad \rightsquigarrow$

$$\begin{aligned} 0 &= -2y_1^2 + 8y_2^2 - 12y_1 - 16y_2 - 18 \\ &= -2(y_1+3)^2 + 18 + 8(y_2-1)^2 - 8 - 18 \\ &= -2\xi_1^2 + 8\xi_2^2 - 8 \end{aligned}$$

(vi) Skalierung:
Division durch 8 $\quad \rightsquigarrow \quad$ Normalform

$$-\frac{\xi_1^2}{2^2} + \frac{\xi_2^2}{1^2} = 1$$

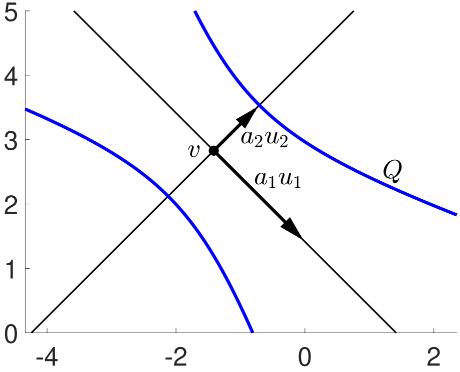
Hyperbel mit Halbachsenlängen 2 und 1

(vii) Transformation:

$x = Uy = U(\xi + (-3, 1)^{\mathrm{t}}) \quad \rightsquigarrow \quad$ Verschiebungsvektor (Mittelpunkt)

$$v = \underbrace{\frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}}_{U} \begin{pmatrix} -3 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} -2 \\ 4 \end{pmatrix}$$

(viii) Skizze:



Euklidische Normalformen der dreidimensionalen Quadriken

Es existieren 10 (nicht-triviale) Typen räumlicher Quadriken mit den folgenden Normalformen:

- Kegelige Quadriken

Normalform	Bezeichnung
$\frac{x_1^2}{a_1^2} + \frac{x_2^2}{a_2^2} - \frac{x_3^2}{a_3^2} = 0$	(Doppel-)Kegel
$\frac{x_1^2}{a_1^2} - \frac{x_2^2}{a_2^2} = 0$	Schneidende Ebenen

- Parabolische Quadriken

Normalform	Bezeichnung
$\frac{x_1^2}{a_1^2} + \frac{x_2^2}{a_2^2} = 2x_3$	elliptisches Paraboloid
$-\frac{x_1^2}{a_1^2} + \frac{x_2^2}{a_2^2} = 2x_3$	hyperbolisches Paraboloid
$\frac{x_1^2}{a_1^2} = 2x_2$	parabolischer Zylinder

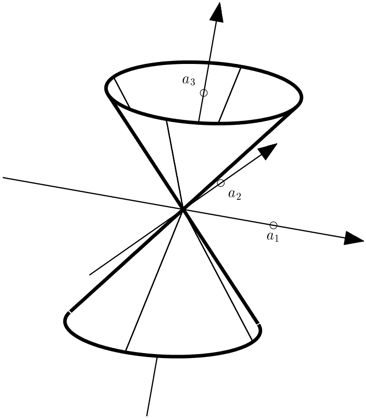
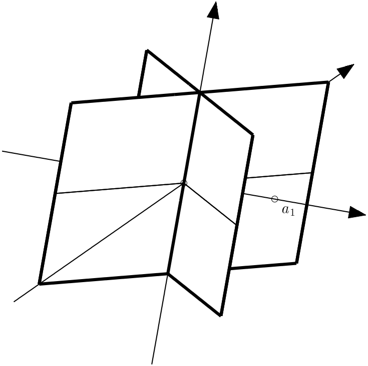
- Mittelpunktsquadriken

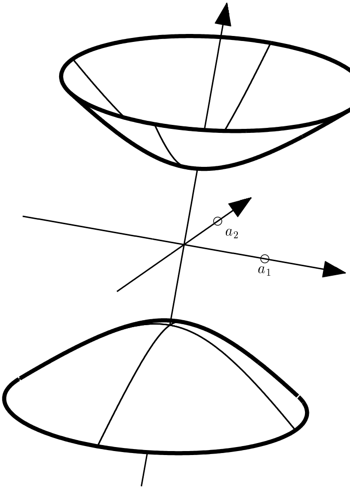
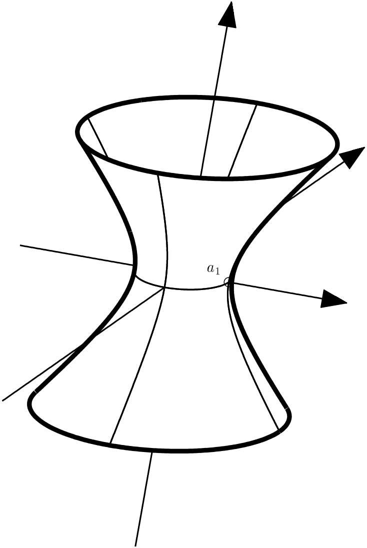
Normalform	Bezeichnung
$-\frac{x_1^2}{a_1^2} - \frac{x_2^2}{a_2^2} + \frac{x_3^2}{a_3^2} = 1$	zweischaliges Hyperboloid
$-\frac{x_1^2}{a_1^2} + \frac{x_2^2}{a_2^2} + \frac{x_3^2}{a_3^2} = 1$	einschaliges Hyperboloid
$\frac{x_1^2}{a_1^2} + \frac{x_2^2}{a_2^2} + \frac{x_3^2}{a_3^2} = 1$	Ellipsoid
$-\frac{x_1^2}{a_1^2} + \frac{x_2^2}{a_2^2} = 1$	hyperbolischer Zylinder
$\frac{x_1^2}{a_1^2} + \frac{x_2^2}{a_2^2} = 1$	elliptischer Zylinder

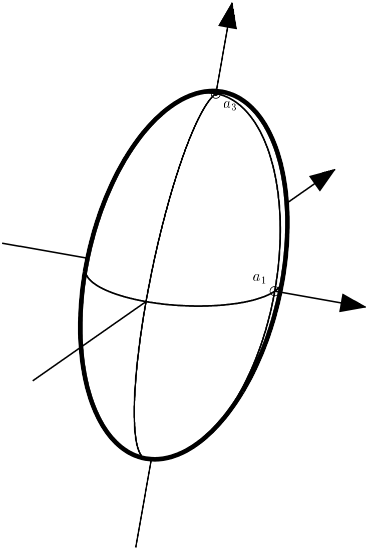
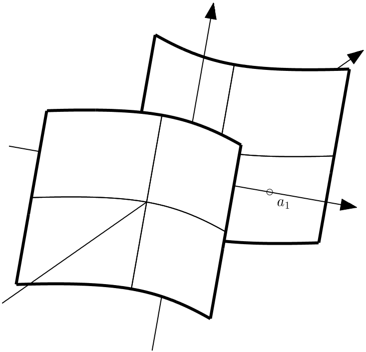
Die Größen a_k sind die Hauptachsenlängen der Quadrik.

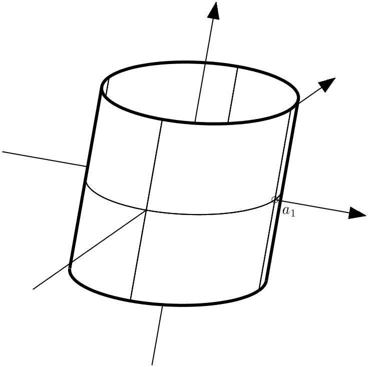
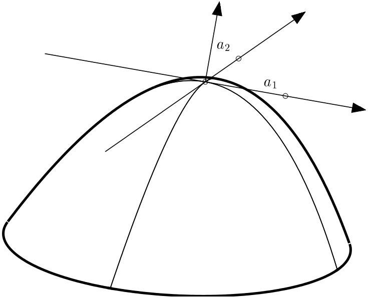
Triviale Sonderfälle sind die Normalformen

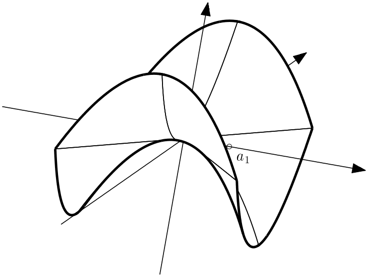
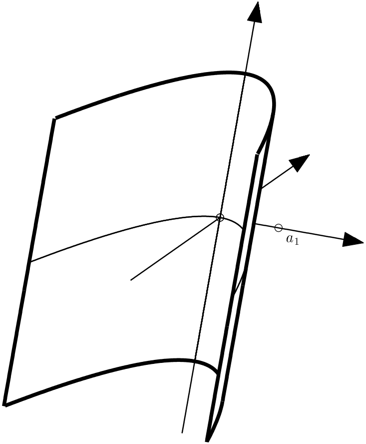
- $-\frac{x_1^2}{a_1^2} - \frac{x_2^2}{a_2^2} - \frac{x_3^2}{a_3^2} = 1,\ -\frac{x_1^2}{a_1^2} - \frac{x_2^2}{a_2^2} = 1,\ -\frac{x_1^2}{a_1^2} = 1$ (leere Mengen)
- $\frac{x_1^2}{a_1^2} + \frac{x_2^2}{a_2^2} + \frac{x_3^2}{a_3^2} = 0,\ \frac{x_1^2}{a_1^2} + \frac{x_2^2}{a_2^2} = 0$ (Punkt und Gerade)
- $\frac{x_1^2}{a_1^2} = 0,\ \frac{x_1^2}{a_1^2} = 1$ (Doppelebene und parallele Ebenen)

(Doppel-)Kegel	schneidende Ebenen
	

zweischaliges Hyperboloid	einschaliges Hyperboloid
	

Ellipsoid	hyperbolischer Zylinder
	

elliptischer Zylinder	elliptisches Paraboloid
	

hyperbolisches Paraboloid	parabolischer Zylinder
	

Beispiel

Normalform und der Typ der Quadrik

$$Q:\ x^{\mathrm{t}}\begin{pmatrix}5&4&0\\4&3&4\\0&4&1\end{pmatrix}x+2\,(-2,1,2)\,x+2=0$$

(i) Eigenwerte:
charakteristisches Polynom

$$\begin{vmatrix}5-\lambda&4&0\\4&3-\lambda&4\\0&4&1-\lambda\end{vmatrix}\underset{\text{Sarrus}}{=}(5-\lambda)(3-\lambda)(1-\lambda)-16(5-\lambda)-16(1-\lambda)$$

$$=\lambda^3-9\lambda^2-9\lambda+81$$

Nullstellen $\rightsquigarrow$ Eigenwerte

$$\lambda_1=9,\quad \lambda_2=3,\quad \lambda_3=-3$$

(ii) Eigenvektoren:
normierte Eigenvektoren zu den Eigenwerten λ_k

$$u_1 = \frac{1}{3}\begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix}, \quad u_2 = \frac{1}{3}\begin{pmatrix} -2 \\ 1 \\ 2 \end{pmatrix}, \quad u_3 = \frac{1}{3}\begin{pmatrix} 1 \\ -2 \\ 2 \end{pmatrix}$$

Vorzeichen so gewählt, dass ein Rechtssystem entsteht $\quad\rightsquigarrow\quad$ Drehmatrix

$$U = (u_1, u_2, u_3) = \frac{1}{3}\begin{pmatrix} 2 & -2 & 1 \\ 2 & 1 & -2 \\ 1 & 2 & 2 \end{pmatrix}$$

(iii) Diagonalisierung:
Substitution $x = Uy \quad \rightsquigarrow$

$$(y_1, y_2, y_3)\begin{pmatrix} 9 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -3 \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} + 2\,\underbrace{(0, 3, 0)}_{(-2,1,2)U}\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} + 2 = 0$$

(iv) Verschiebung:
quadratische Ergänzung $\xi = y + (0,1,0)^{\mathrm{t}}$ $\rightsquigarrow$

$$\begin{aligned} 0 &= 9y_1^2 + 3y_2^2 - 3y_3^2 + 6y_2 + 2 \\ &= 9y_1^2 + 3(y_2+1)^2 - 6y_2 - 3 - 3y_3^2 + 6y_2 + 2 \\ &= 9\xi_1^2 + 3\xi_2^2 - 3\xi_3^2 - 1 \end{aligned}$$

(v) Normalform:

$$-\frac{\xi_3^2}{1/3} + \frac{\xi_1^2}{1/9} + \frac{\xi_2^2}{1/3} = 1$$

Q: einschaliges Hyperboloid mit Hauptachsenlängen

$$a_3 = \sqrt{1/3}, \quad a_1 = 1/3, \quad a_2 = \sqrt{1/3}$$

Transformation $x = Uy = U(\xi - (0,1,0)^{\mathrm{t}})$ $\rightsquigarrow$ Mittelpunkt

$$v = -U \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2/3 \\ -1/3 \\ -2/3 \end{pmatrix}$$