

SS 04, Universität Augsburg

Mitschrift aus der Vorlesung von Prof. Dr. Pukelsheim

Lineare Algebra II

Tobias Wichtrey, tobias@tarphos.de

21. September 2004

Inhaltsverzeichnis

Einführung	5
8 Normaldarstellungen für Matrizen von Vektorraum-Endomorphismen	8
8.1 Basis-Konvertierungslemma	8
8.2 f -Invarianz	10
8.3 Nilpotenz	13
8.4 Darstellungssatz für nilpotente Endomorphismen (Jordan-Darstellung)	13
8.5 Eigenvektoren und Eigenwerte	18
8.6 Algebraische Abgeschlossenheit	20
8.7 Darstellungssatz für beliebige Endomorphismen (Jordan-Normalform)	21
8.8 Eindeutigkeit der Jordan-Darstellung	25
8.9 Matrizenformulierung	27
8.10 Satz über Ähnlichkeitstransformationen	27
8.11 Satz über diagonale Darstellungen (Diagonalisierbarkeit)	28
8.12 Vorsatz zu Kap. 8.13	30
8.13 Satz über simultane Diagonalisierbarkeit	30
9 Transformationsgruppen als elementarste Bausteine der Mathematik	33
9.1 Gruppen	33
9.2 Untergruppen, Kern, Bild	35
9.3 Normalteiler	36
9.4 Homomorphiesatz für Gruppen	37
10 Euklidische Vektorräume	38
10.1 Bilinearformen	38
10.2 Gram-Matrizen	39
10.3 Orthogonalität	41
10.3.1 Exkurs über Normen	44
10.3.2 Orthonormalbasen in euklidischen Vektorräumen	46
10.3.3 Orthogonale Abbildungen (Isometrien)	49
11 Projektionen und verallgemeinerte Matrixinverse	57
11.1 Projektionen	57
11.2 Orthogonale Projektionen	58
11.3 Generalisiertinverse Matrizen	60

Inhaltsverzeichnis

11.4 Pseudoinverse Matrizen	60
11.5 Vektorhalbordnungen	64
12 Tensorprodukte oder Multilineare Algebra	70
12.1 Tensorprodukte von Vektorräumen	70
12.1.1 Reduktionssatz	71
12.1.2 Satz zur Minimaleigenschaft von T	71
12.1.3 Tensorprodukte von Teilräumen	72
12.1.4 Eindeutigkeit von Tensorprodukten	72
12.1.5 Existenz von Tensorprodukten	73
12.2 Tensorprodukte von linearen Abbildungen	74
12.2.1 Rangformel	74
12.2.2 Tensorprodukte von Produkten linearer Abbildungen	75
12.3 Das Kroneckerprodukt (Tensorprodukt für euklidische Vektorräume)	75
12.3.1 Kroneckerprodukt von zwei Matrizen	76
12.3.2 Der Vektorisierungsoperator	78

Einführung

21. 4. 2004

Sichtwechsel: Bisher Basen fest, lineare Abbildungen beliebig, jetzt lineare Abbildungen fest, Basen beliebig.

Satz (Basen-Konvertierungs-Lemma). Seien

K Körper,

V, W endlichdimensionale K -Vektorräume,

$f \in \text{Hom}_K(V, W)$,

$B = \{v_1, \dots, v_n\}$ Basis in V ,

$\tilde{B} = \{\tilde{v}_1, \dots, \tilde{v}_n\}$ Basis in V ,

$C = \{w_1, \dots, w_m\}$ Basis in W ,

$\tilde{C} = \{\tilde{w}_1, \dots, \tilde{w}_m\}$ Basis in W ,

$M(f)$ Darstellungsmatrix von f bezüglich B und C ,

$\tilde{M}(f)$ Darstellungsmatrix von f bezüglich \tilde{B} und \tilde{C} ,

$S \in \text{GL}(n)$ bestimmt durch $\tilde{v}_j = \sum_{i=1}^n s_{ij} v_i$ für $j = 1, \dots, n$,

$T \in \text{GL}(m)$ bestimmt durch $w_k = \sum_{l=1}^m s_{lk} \tilde{w}_l$ für $k = 1, \dots, m$.

Dann gilt:

$$\tilde{M}(f) = TM(f)S$$

Beweis.

I. Die lineare Abbildung $s : V \rightarrow V$, die definiert ist durch $s(v_j) := \tilde{v}_j$ hat Darstellungsmatrix S . Da s bijektiv ist (Basistransformationssatz I. 5.1), ist S invertierbar. Somit $S \in \text{GL}(n)$.

II. Analog ist T die invertierbare Darstellungsmatrix der linearen Abbildung $t(\tilde{w}_k) = w_k$.

Einführung

III. Mit $\tilde{M}(f) = (\tilde{a}_{ij})$ und $M(f) = (a_{ij})$ gilt:

$$\begin{aligned}
 \sum_{l=1}^m \tilde{a}_{lj} \tilde{w}_l &= f(\tilde{v}_j) && \text{(wg. Def. der Darstellungsmatrix)} \\
 &= f\left(\sum_{i=1}^m s_{ij} v_i\right) && \text{(wg. Definition von } S) \\
 &= \sum_i s_{ij} f(v_i) && \text{(da } f \text{ linear)} \\
 &= \sum_i s_{ij} \sum_k a_{ki} w_k && \text{(wg. Definition der Darstellungsmatrix)} \\
 &= \sum_k \sum_i a_{ki} s_{ij} w_k \\
 &= \sum_k \sum_i a_{ki} s_{ij} \sum_l t_{lk} \tilde{w}_l && \text{(wg. Definition von } T) \\
 &= \sum_l \left(\sum_k \sum_i t_{lk} a_{ki} s_{ij} \right) \tilde{w}_l
 \end{aligned}$$

Koeffizientenvergleich (da \tilde{B} Basis) erzwingt

$$a_{lj} = \sum_k \sum_i t_{lk} a_{ki} s_{ij},$$

d. h. $\tilde{M}(f) = TM(f)S$.

□

Satz (Normaldarstellung für Matrizen von Vektorraum-Homomorphismen).

Sei $f \in \text{Hom}_K(V, W)$. Dann gilt:

$$\begin{aligned}
 \exists \tilde{B} &= \{\tilde{v}_1, \dots, \tilde{v}_n\} \text{ Basis von } V \\
 \exists \tilde{C} &= \{\tilde{w}_1, \dots, \tilde{w}_m\} \text{ Basis von } W :
 \end{aligned}$$

$$M(f) = \begin{pmatrix} 1 & & & & \\ & \ddots & & & 0 \\ & & 1 & & \\ & & & 0 & \\ 0 & & & & \ddots \\ & & & & & 0 \end{pmatrix}$$

mit genau $r := \text{Rang}(f)$ Einsen.

Andere Schreibweise: $\begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}$

Beweis. Sei $B = \{v_1, \dots, v_n\}$ Basis von V und $C = \{w_1, \dots, w_m\}$ Basis von W und $M(f)$ die Darstellungsmatrix bezüglich B und C .

Bekanntlich kann $M(f)$ durch elementare Zeilen- und Spaltentransformationen auf die Gestalt $\begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}$ gebracht werden. Die Komposition der elementaren Zeilenumformungen definiert eine Matrix $T \in \text{GL}(m)$, siehe Beweis des Multiplikationssatzes für Determinanten. Analog ergibt die Komposition der elementaren Spaltenumformungen eine Matrix $S \in \text{GL}(n)$. Also

$$TM(f)S = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}.$$

Setze nun $\tilde{v}_j := \sum_{i=1}^n s_{ij}v_i$ und $\tilde{w}_l := \sum_l \tilde{u}_{lj}w_i$ mit $(\tilde{u}_{kl}) =: T^{-1}$. Gemäß Basis-Konvertierungs-Lemma ist $\tilde{M}(f) = TM(f)S$ die Darstellungsmatrix bezüglich \tilde{B} und \tilde{C} .

□

Speziell im folgenden Kapitel $V = W$.

Definition. $\text{End}_K(V) := \text{Hom}_K(V, V)$ Menge der „**Vektorraum-Endomorphismen**“.

Dann interessiert ein Basiswechsel von $B = C$ zu $\tilde{B} = \tilde{C}$. Also statt vier Basen nur noch zwei Basen.

8 Normaldarstellungen für Matrizen von Vektorraum-Endomorphismen

8.1 Basis-Konvertierungslemma

Satz. Seien

$f \in \text{End}_K(V)$,

$B := \{v_1, \dots, v_n\}$ Basis von V ,

$\tilde{B} := \{\tilde{v}_1, \dots, \tilde{v}_n\}$ Basis von V ,

$M(f)$ Darstellungsmatrix von f bezüglich B ,

$\tilde{M}(f)$ Darstellungsmatrix von f bezüglich \tilde{B} .

Dann gilt

$$\tilde{M}(f) = S^{-1}M(f)S,$$

wobei $S \in \text{GL}(n)$ bestimmt ist durch

$$\tilde{v}_j = \sum_{i=1}^n s_{ij}v_i \quad \forall j.$$

Beweis. Im Konvertierungslemma für Vektorraum-Homomorphismen galt $\tilde{M}(f) = TM(f)S$, wobei zu S die lineare Abbildung $s(v_j) = \tilde{v}_j$ gehört und zu T die lineare Abbildung $t(\tilde{w}_k) = w_k$. Jetzt $B = C$ und $\tilde{B} = \tilde{C}$, woraus folgt $(t \circ s)(v_j) = t(s(v_j)) = t(\tilde{v}_j) = v_j$, somit $t \circ s = \text{id}_V$, $t = s^{-1}$ und $T = S^{-1}$. \square

Bemerkungen.

(i) Wegen

$$\det(\tilde{M}(f)) = \det(S^{-1}M(f)S) = \frac{1}{\det S} \cdot \det M(f) \cdot \det S = \det M(f) =: \det f$$

ist $\det f$ wohldefiniert.

(ii) *Definition.*

$$\forall A \in K^{n \times n} : \text{Spur } A := \sum_{i=1}^n a_{ii}$$

Die „**Spur**“¹ einer Matrix ist die Summe der Hauptdiagonalelemente.

¹engl. *trace*

Es gilt

$$\det(AB) = (\det A)(\det B) = (\det B)(\det A) = \det(BA)$$

und

$$\text{Spur}(AB) = \sum_i \sum_j a_{ij} b_{ji} = \sum_j \sum_i b_{ji} a_{ij} = \text{Spur}(BA).$$

Zudem analog zu (i):

Wegen

$$\text{Spur } \tilde{M}(f) = \text{Spur} \left(\underbrace{S^{-1}}_A \underbrace{M(f)S}_B \right) = \text{Spur}(M(f)SS^{-1}) = \text{Spur } M(f) =: \text{Spur } f$$

ist Spur f wohldefiniert.

(iii) Die Rechenregeln für \det und Spur übertragen sich somit von Matrizen auf lineare Abbildungen, z. B.

$$\begin{aligned} \det \text{id}_V &= \det E_n = 1 \\ \text{Spur } \text{id}_V &= \text{Spur } E_n = n \\ \det(kf) &= \det(kM(f)) = k^n \det M(f) = k^n \det f \\ \text{Spur}(kf) &= \dots = k \text{ Spur } f. \end{aligned}$$

(iv) Besonders hilfreich:

$$\begin{aligned} f \text{ Isomorphismus} &\Leftrightarrow f \text{ bijektiv} \\ &\Leftrightarrow M(f) \text{ invertierbar} \\ &\Leftrightarrow \det M(f) \neq 0 \\ &\Leftrightarrow \det f \neq 0 \end{aligned}$$

(v) Welche „Normaldarstellung“ für $M(f)$ ist erreichbar?

Definition. $M(f)$ heißt „**diagonalisierbar**“, falls gilt:

$$\exists S \in \text{GL}(n) \exists k_1, \dots, k_n \in K : \quad \tilde{M}(f) := \underbrace{S^{-1}M(f)S}_{\text{„Ähnlichkeitstransformation“}} = \underbrace{\begin{pmatrix} k_1 & & 0 \\ & \ddots & \\ 0 & & k_n \end{pmatrix}}_{\text{„Diagonalgestalt“}}$$

Selbst Diagonalisierbarkeit ist nicht immer gegeben. Satz 7 wird ergeben, daß i. a. die

8 Normaldarstellungen für Matrizen von Vektorraum-Endomorphismen

folgende Darstellung erreichbar ist:

$$J := \begin{pmatrix} J_1 & & 0 \\ & J_2 & \\ & & \ddots \\ 0 & & & J_t \\ & & & & 0 \end{pmatrix}$$

mit „quadratischen Blöcken“ $J_i = \begin{pmatrix} x_i & & 0 \\ 1 & \ddots & \\ & \ddots & \ddots \\ 0 & & 1 & x_i \end{pmatrix} \in K^{k_i \times k_i}, x_i \in K$

8.2 f -Invarianz

Definition. Sei $f \in \text{End}_K(V)$ und $U \subseteq V$ Untervektorraum. Der Untervektorraum U heißt „ f -invariant“, falls gilt $f(U) := \{f(u) \mid u \in U\} \subseteq U$.

Bemerkungen.

- (i) In diesem Fall ist die Einschränkung

$$\begin{aligned} f|_U: U &\rightarrow U \\ u &\mapsto f(u) \end{aligned}$$

wohldefiniert und linear. $f|_U \in \text{End}_K(U)$.

- (ii) Wählt man eine Basis b_1, \dots, b_m von U und ergänzt diese mit den Vektoren c_1, \dots, c_n zu einer Basis $b_1, \dots, b_m, c_1, \dots, c_n$ von V , so hat die Darstellungsmatrix von f die „Block-Dreiecksform“

$$\begin{pmatrix} B & D \\ 0 & C \end{pmatrix},$$

wobei B die Darstellungsmatrix von $f|_U$ ist.

- (iii) Sei $V = U_1 + U_2$ Vektorraum-Summe zweier Untervektorräume U_1 und U_2 , die zueinander komplementär sind. (Schreibweise $V = U_1 \oplus U_2$)

Wählt man eine Basis b_1, \dots, b_m von U_1 und eine Basis c_1, \dots, c_n von U_2 , so erhält die Darstellungsmatrix von f bezüglich $B = \{b_1, \dots, b_m, c_1, \dots, c_n\}$ die „Blockdiagonalgestalt“

$$M(f) = \begin{pmatrix} B & 0 \\ 0 & C \end{pmatrix},$$

wenn sowohl U_1 als auch U_2 f -invariant sind.

(iv) Invariante Untervektorräume erhält man mittels der „iterierten Abbildungen“

$$\begin{aligned} f^1 &:= f \\ f^2 &:= f \circ f \\ f^3 &:= f \circ (f \circ f) \\ &\vdots \\ f^k &:= f \circ f^{k-1} = f^{k-1} \circ f. \end{aligned}$$

Satz (von Hans FITTING, 1906–1938). Sei $f \in \text{End}_K(V)$. Dann gilt:

- (i) $V \supseteq \text{Bild } f \supseteq \text{Bild } f^2 \supseteq \dots$
- (ii) $\{0\} \subseteq \text{Kern } f \subseteq \text{Kern } f^2 \subseteq \dots$
- (iii) $\exists k \in \mathbb{N} : \text{Bild } f^k = \text{Bild } f^{k+1}$
- (iv) $\exists l \in \mathbb{N} : \text{Kern } f^l = \text{Kern } f^{l+1}$
- (v) $\forall k \in \mathbb{N} : \text{Bild } f^k = \text{Bild } f^{k+1} \Rightarrow \text{Bild } f^{k+1} = \text{Bild } f^{k+2} = \dots$
- (vi) $\forall l \in \mathbb{N} : \text{Kern } f^l = \text{Kern } f^{l+1} \Rightarrow \text{Kern } f^{l+1} = \text{Kern } f^{l+2} = \dots$
- (vii) $\min\{k \in \mathbb{N} \mid \text{Bild } f^k = \text{Bild } f^{k+1}\} = \min\{l \in \mathbb{N} \mid \text{Kern } f^l = \text{Kern } f^{l+1}\} =: m$
- (viii) $V = \text{Bild } f^m + \text{Kern } f^m$ **und** $\text{Bild } f^m \cap \text{Kern } f^m = \{0\}$.
Schreibweise: $V = \text{Bild } f^m \oplus \text{Kern } f^m$.
- (ix) $\text{Bild } f^m$ f -invariant **und** $\text{Kern } f^m$ f -invariant.

28. 4. 2004

Beweis.

- (i) $\text{Bild } f \subseteq V \Rightarrow \text{Bild } f^2 = f(\text{Bild } f) \subseteq f(V) = \text{Bild } f$
...
- (ii) $x \in \text{Kern } f \Rightarrow f(x) = 0 \Rightarrow f(f(x)) = f(0) = 0 \Rightarrow x \in \text{Kern } f^2$
...
- (iii) Aus (i) folgt, daß die Dimensionen auch fallen:

$$\begin{aligned} n \geq \text{Rang } f \geq \dots \geq 0 \Rightarrow \exists k : \text{Rang } f^k = \text{Rang } f^{k+1} \\ \text{und } \text{Bild } f^k = \text{Bild } f^{k+1} \end{aligned}$$

(iv) Aus (ii) folgt, daß die Dimensionen steigen:

$$\begin{aligned} 0 \leq \dim \text{Kern } f \leq \dim \text{Kern } f^2 \leq \dots \leq n \Rightarrow \exists l : \dim \text{Kern } f^l = \dim \text{Kern } f^{l+1} \\ \text{und } \text{Kern } f^l = \text{Kern } f^{l+1} \end{aligned}$$

8 Normaldarstellungen für Matrizen von Vektorraum-Endomorphismen

$$(v) \text{ Bild } f^{k+1} = f(\text{Bild } f^k) \underset{\text{vor.}}{=} f(\text{Bild } f^{k+1}) = \text{Bild } f^{k+2}$$

(vi)

$$\begin{aligned} x \in \text{Kern } f^{l+2} &\Rightarrow 0 = f^{l+2}(x) = f^{l+1}(f(x)) \\ &\Rightarrow f(x) \in \text{Kern } f^{l+1} \\ &\underset{\text{vor.}}{\Rightarrow} f^l(f(x)) = 0 \\ &\Rightarrow x \in \text{Kern } f^{l+1} \end{aligned}$$

(vii) $\forall k \in \mathbb{N}$:

$$\begin{aligned} \text{Bild } f^k = \text{Bild } f^{k+1} &\Leftrightarrow \text{Rang } f^k = \text{Rang } f^{k+1} \\ &\Leftrightarrow \dim V - \text{Rang } f^k = \dim V - \text{Rang } f^{k+1} \\ &\Leftrightarrow \dim(\text{Kern } f^k) = \dim(\text{Kern } f^{k+1}) \\ &\Leftrightarrow \text{Kern } f^k = \text{Kern } f^{k+1} \end{aligned}$$

Daraus folgt die Behauptung.

(viii) I. „ \supseteq “ \checkmark

II. „ \subseteq “:

Sei $v \in V$. Wegen $f^m(v) \in \text{Bild } f^m = \text{Bild } f^{m+1} = \dots = \text{Bild } f^{2m}$ existiert ein $w \in V$ mit $f^m(v) = f^{2m}(w)$. Dafür: $0 = f^m(v) - f^{2m}(w) = f^m(v - f^m(w))$, also $v - f^m(w) \in \text{Kern } f^m$. Somit hat man $v = f^m(w) + (v - f^m(w)) \in \text{Bild } f^m + \text{Kern } f^m$.

III. Sei $v \in \text{Bild } f^m \wedge \text{Kern } f^m$. Dann existiert ein $w \in V$ mit $v = f^m(w)$, wofür wegen $v \in \text{Kern } f^m$ gilt $f^m(v) = f^{2m}(w) = 0$. Also $w \in \text{Kern } f^{2m} = \dots = \text{Kern } f^m$ und es folgt $v = f^m(w) = 0$.

(ix) Zu zeigen $f(U) \subseteq U$ mit $U := \text{Bild } f^m$. Offenbar gilt sogar Gleichheit, denn

$$f(\text{Bild } f^m) = \text{Bild } f^{m+1} = \text{Bild } f^m.$$

Noch zu zeigen $f(U) \subseteq U$ mit $U := \text{Kern } f^m$. Das folgt aus

$$\begin{aligned} v \in \text{Kern } f^m = \text{Kern } f^{m+1} &\Rightarrow 0 = f^{m+1}(v) = f^m(f(v)) \\ &\Rightarrow f(v) \in \text{Kern } f^m. \end{aligned}$$

□

Bemerkungen.

(i)

$$\begin{aligned} \text{Bild } f^m = V &\Leftrightarrow f^m \text{ surjektiv} \\ &\Leftrightarrow f^m \text{ bijektiv} \\ &\Leftrightarrow f \text{ bijektiv (und } m = 1) \end{aligned}$$

(ii) $\text{Bild } f^m = \{0\} \Leftrightarrow f^m = 0 \Leftrightarrow \exists l : f^l = 0$

8.3 Nilpotenz

Definition. $f \in \text{End}_K(V)$ heißt „**nilpotent**“, falls $l \in \mathbb{N}$ existiert mit $f^l = 0$. Ist m das kleinste solcher l , so heißt f nilpotent „**vom Index m** “.

8.4 Darstellungssatz für nilpotente Endomorphismen (Jordan-Darstellung)

Satz. Sei $f \in \text{End}_K(V)$ nilpotent. Dann gilt

$$\begin{aligned} \exists l \in \mathbb{N} \exists n_1, \dots, n_l \in \mathbb{N} \exists b_1, \dots, b_l \in V : \\ b_1, f(b_1), \dots, f^{n_1-1}(b_1), \\ b_2, f(b_2), \dots, f^{n_2-1}(b_2), \\ \vdots \\ b_l, f(b_l), \dots, f^{n_l-1}(b_l) \\ \text{ist eine Basis von } V \text{ und } f^{n_1}(b_1) = \dots = f^{n_l}(b_l) = 0. \end{aligned}$$

Bezüglich dieser Basis hat f die in Abb. 8.1 auf der nächsten Seite definierte Darstellungsmatrix $M(f)$.

Beweis. Falls die gegebene Liste eine Basis bildet, ist $M(f)$ die zugehörige Darstellungsmatrix, denn es gilt immer $f(v_j) = \sum_i a_{ij}v_i$. Hier: $v_j = f^p(b_q) \Rightarrow f(v_j) = f^{p+1}(b_q) = v_{j+1}$

$$\begin{aligned} \Rightarrow a_{j+1,j} &= 1, a_{ij} = 0 \text{ sonst, falls } p \neq n_i \\ \text{oder } a_{j+1,j} &= 0, a_{ij} = 0 \text{ sonst, falls } p = n_i. \end{aligned}$$

Zur Hauptsache, daß die gegebene Liste eine Basis bildet: Wir konstruieren Untervektorräume U_1, \dots, U_m von V so, daß gilt

$$\begin{aligned} V &= \text{Kern } f^m = \text{Kern } f^{m-1} \oplus U_1, \\ \text{Kern } f^{m-1} &= (\text{Kern } f^{m-2} + f(U_1)) \oplus U_2, \\ \text{Kern } f^{m-2} &= (\text{Kern } f^{m-3} + f^2(U_1) + f(U_2)) \oplus U_3, \\ &\vdots \\ \text{Kern } f^2 &= (\text{Kern } f + f^{m-2}(U_1) + \dots + f(U_{m-2})) \oplus U_{m-1}, \\ \text{Kern } f &= (\{0\} + f^{m-1}(U_1) + \dots + f(U_{m-1})) \oplus U_m. \end{aligned}$$

8.4 Darstellungssatz für nilpotente Endomorphismen (Jordan-Darstellung)

Wir wählen Basen

$$\begin{array}{ll}
 b_1, \dots, b_{r_1} & \text{von } U_1, \\
 b_{r_1+1}, \dots, b_{r_2} & \text{von } U_2, \\
 & \vdots \\
 b_{r_{m-2}+1}, \dots, b_{r_{m-1}} & \text{von } U_{m-1}, \\
 b_{r_{m-1}+1}, \dots, b_{r_m} & \text{von } U_m.
 \end{array}$$

Nun zeigen wir, daß die folgende Liste von Vektoren eine Basis von V bilden:

$$\begin{array}{ll}
 \left. \begin{array}{l} b_1, f(b_1), \dots, f^{m-1}(b_1), \\ \vdots \\ b_{r_1}, f(b_{r_1}), \dots, f^{m-1}(b_{r_1}), \end{array} \right\} & \text{aus } U_1, \text{ wird } (m-1)\text{-mal iteriert} \\
 \left. \begin{array}{l} b_{r_1+1}, f(b_{r_1+1}), \dots, f^{m-2}(b_{r_1+1}), \\ \vdots \\ b_{r_2}, f(b_{r_2}), \dots, f^{m-2}(b_{r_2}), \end{array} \right\} & \text{aus } U_2, \text{ wird } (m-2)\text{-mal iteriert} \\
 & \vdots \\
 & \vdots \\
 \left. \begin{array}{l} b_{r_{m-2}+1}, f(b_{r_{m-2}+1}), \\ \vdots \\ b_{r_{m-1}}, f(b_{r_{m-1}}), \end{array} \right\} & \text{aus } U_{m-1}, \text{ wird 1-mal iteriert} \\
 \left. \begin{array}{l} b_{r_{m-1}+1}, \\ \vdots \\ b_{r_m} \end{array} \right\} & \text{aus } U_m, \text{ wird 0-mal iteriert}
 \end{array}$$

I. Hilfsaussage: Für $i = 0, 1, 2, \dots$ gilt

$$\text{Kern } f^{m-i} = \text{Kern } f^{m-i-1} \oplus f^i(U_1) \oplus f^{i-1}(U_2) \oplus \dots \oplus f(U_i) \oplus U_{i+1}.$$

Denn für $i = 0$ gilt das per Konstruktion. Mit vollständiger Induktion folgt:

Sei $x \in \text{Kern } f^{m-i-1}$, $u_1 \in U_1, \dots, u_{i+1} \in U_{i+1}$ mit $0 = x + f^i(u_1) + f^{i-1}(u_2) + \dots + f(u_i) + u_{i+1}$. Wir wollen zeigen, daß jeder Summand selber schon gleich 0 ist, denn damit ist die behauptete Zerlegung eine *direkte* Summe.

Per Konstruktion gilt $u_{i+1} = 0$. Weiter $0 = f^{m-i-1}(0) = f^{m-i-1}(x + \dots) = 0 + f^{m-1}(u_1) + f^{m-2}(u_2) + \dots + f^{m-i}(u_i) = f^{m-i}(f^{i-1}(u_1) + \dots + u_i)$, d. h. $f^{i-1}(u_1) + \dots + u_i \in \text{Kern } f^{m-i}$.

Es sind Kern $f^{m-(i-1)-1} \equiv \text{Kern } f^{m-1}$ und $f^{i-1}(U_1), \dots, f(U_i)$ gemäß Induktionsvoraussetzung zueinander komplementär.

Somit folgt $0 = f^{i-1}(u_1) = f^{i-2}(u_2) = \dots = u_i$. Daraus $0 = f^i(u_1) = f^{i-1}(u_2) = \dots = f(u_1)$, was impliziert $x = 0$.

8 Normaldarstellungen für Matrizen von Vektorraum-Endomorphismen

II. Folglich gilt

$$\begin{aligned} V &= U_1 \oplus f(U_1) \oplus f^2(U_1) \oplus \cdots \oplus f^{m-1}(U_1) \\ &\quad \oplus U_2 \oplus f(U_2) \oplus \cdots \oplus f^{m-2}(U_2) \\ &\quad \vdots \\ &\quad \oplus U_{m-1} \oplus f(U_{m-1}) \\ &\quad \oplus U_m \end{aligned}$$

und die aufgelisteten Vektoren bilden ein Erzeugendensystem von V .

III. Die aufgelisteten Vektoren sind auch linear unabhängig, denn eingeschränkt auf U_i sind für $j = 1, \dots, m-1$ die Abbildungen $f^j|_{U_i}$ injektiv, weil ihr Kern entartet:

$$\begin{aligned} x \in U_i \wedge \text{Kern } f^j \Rightarrow 0 &= f^j(x) = f^{j+1}(x) = \cdots = f^{m-i}(x) \\ &\Rightarrow x \in U_i \wedge \text{Kern } f^{m-i} = \{0\} \Rightarrow x = 0. \end{aligned}$$

□

29.4.2004 *Beispiel.* $n_1 = 2 < n_2 = 3$. Sei $f \in \text{End}(\mathbb{R}^5)$ gegeben durch $f(x) = Ax$ mit

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

Dann $f^2(x) = A^2x$ mit

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

Und $f^3(x) = A^3x$ mit $A^3 = 0$. Somit ist f nilpotent mit Index 3. Also $\{0\} \subseteq \text{Kern } A \subseteq \text{Kern } A^2 \subseteq \text{Kern } A^3 = \mathbb{R}^5$.

Jetzt rückwärts rechnen (wie im Beweis).

I.

$$\begin{aligned} x \in \text{Kern } A^2 &\Leftrightarrow \begin{pmatrix} 1 & -2 & 0 & 0 & 1 \\ -1 & 2 & 0 & 0 & -1 \end{pmatrix} x = 0 \\ &\Leftrightarrow x_1 - 2x_2 + x_5 = 0 \end{aligned}$$

Da $\dim(\text{Kern } A^2) = 4$, folgt $\dim U_1 = 1$.

8.4 Darstellungssatz für nilpotente Endomorphismen (Jordan-Darstellung)

Wir wählen

$$b_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

und berechnen

$$f(b_1) = Ab_1 = \begin{pmatrix} 2 \\ 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}$$

und

$$f^2(b_1) = A^2b_1 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \\ 0 \end{pmatrix}.$$

II. Der Untervektorraum Kern A^2 hat Basis

$$\begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}.$$

Jetzt zu Kern A :

$$x \in \text{Kern } A \Leftrightarrow \dots \Leftrightarrow \begin{cases} x_3 = -x_4 \\ x_2 = 0 \\ x_1 = -x_5 \end{cases}$$

Da $\dim(\text{Kern } A) = 2$, suchen wir 2 Basisvektoren dafür, z. B.

$$\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \\ 0 \end{pmatrix}.$$

Zusammen mit

$$f(b_1) = \begin{pmatrix} 2 \\ 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}$$

8 Normaldarstellungen für Matrizen von Vektorraum-Endomorphismen

suchen wir eine Ergänzung b_2 so, daß $\text{Kern } A^2 = (\text{Kern } A + f(U_1)) \oplus \underbrace{b_2}_{U_2}$. Wir wählen

$$b_2 = \begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

und berechnen

$$f(b_2) = Ab_2 = \begin{pmatrix} 1 \\ 0 \\ 2 \\ -2 \\ -1 \end{pmatrix}.$$

III. Finale:

Bezüglich der Basis

$$\underbrace{b_1}_{=v_1}, \underbrace{f(b_1)}_{=v_2=Av_1}, \underbrace{f^2(b_1)}_{=v_3=Av_2}, \underbrace{b_2}_{=v_4}, \underbrace{f(b_2)}_{=v_5=Av_4}$$

besitzt f die Darstellungsmatrix

$$M(f) = \left(\begin{array}{ccc|cc} 0 & 0 & 0 & & \\ 1 & 0 & 0 & & 0 \\ 0 & 1 & 0 & & \\ \hline & & & 0 & 0 \\ & 0 & & 1 & 0 \end{array} \right).$$

Bemerkung. Sei f weder bijektiv (d. h. $\text{Bild } f^m \neq V$) noch nilpotent (d. h. $\text{Bild } f^m \neq \{0\}$). Dann wird V durch $\text{Bild } f^m$ und $\text{Kern } f^m$ in zwei echte Untervektorräume zerlegt. Dabei ist $f|_{\text{Bild } f^m}$ surjektiv (wegen $\text{Bild } f^m = \text{Bild } f^{m+1} = f(\text{Bild } f^m)$) und also auch bijektiv. Zudem ist $f|_{\text{Kern } f^m}$ nilpotent.

Somit ist f „zerlegt“ in eine bijektive Abbildung und eine nilpotente Abbildung.

Jetzt zu bijektiven Endomorphismen.

8.5 Eigenvektoren und Eigenwerte

Definition. Sei $f \in \text{End}(V)$ und $x \in K$.

- (i) x heißt „**Eigenwert von f** “, falls gilt $\exists v \in V \setminus \{0\} : f(v) = xv$ (d. h. $\text{Kern}(f - x \text{id}_V) \neq \{0\}$).
- (ii) $v \in V \setminus \{0\}$ heißt „**Eigenvektor zum Eigenwert x** “, falls gilt $f(v) = xv$ (d. h. $v \in \text{Kern}(f - x \text{id}_V) \setminus \{0\}$).

(iii) $E_x(f) := \text{Kern}(f - x \text{id}_V)$ heißt „**Eigenraum zum Eigenwert x** “.

$H_x(f) := \bigcup_{k \in \mathbb{N}} \text{Kern}((f - x \text{id}_V)^k)$ heißt „**Hauptraum zum Eigenwert x** “.

(iv) $\text{Spec}(f) := \{x \in K \mid x \text{ Eigenwert von } f\}$ heißt „**Spektrum von f** “.

Bemerkungen.

(i)

$$\begin{aligned} v \in E_x(f) &\Leftrightarrow (f - x \text{id}_V)(v) = 0 \\ &\Leftrightarrow f(v) = xv \\ &\Leftrightarrow v = 0 \text{ **oder** } v \text{ Eigenvektor zum Eigenwert } x \end{aligned}$$

(ii) Nach Schachtelungssatz existiert m mit $H_x(f) = \text{Kern}((f - x \text{id}_V)^m) = \text{Kern}((f - x \text{id}_V)^{m+1}) = \dots$ und $H_x(f)$ ist $(f - x \text{id}_V)$ -invariant und $E_x(f) \subseteq H_x(f)$.

(iii) Daraus folgt offenbar

$$\forall v \in H_x(f) : f(v) = \underbrace{(f - \text{id}_V)(v)}_{\in H_x(f)} + \underbrace{xv}_{\in H_x(f)} \in H_x(f),$$

d. h. $H_x(f)$ ist zudem f -invariant.

(iv) $\forall x \in K :$

$$\begin{aligned} x \in \text{Spec } f &\Leftrightarrow \text{Kern}(f - x \text{id}_V) \neq \{0\} \\ &\Leftrightarrow f - x \text{id}_V \text{ *nicht* bijektiv} \\ &\Leftrightarrow \det(f - x \text{id}_V) = 0 \\ &\Leftrightarrow \det(M(f) - xE_n) = 0, \end{aligned}$$

wobei $M(f)$ die Darstellungsmatrix einer beliebigen Basis ist.

(v) Übertragung auf Matrizen $A \in K^{n \times n}$.

$x \in K$ „Eigenwert von A “, falls $\det(A - xE_n) = 0$.

$v \in V$ „Eigenvektor von A “, falls $v \neq 0$ **und** $Av = xv$.

Ebenso Eigenraum, Hauptraum, Spektrum.

Beispiel. Die Abbildung $f \in \text{End}(\mathbb{R}^3)$ habe die Matrix

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

8 Normaldarstellungen für Matrizen von Vektorraum-Endomorphismen

Dann $x \in \mathbb{R}$,

$$\begin{aligned}\det(A - xE_3) &= \det \begin{pmatrix} 1-x & 1 & 1 \\ 1 & 2-x & 1 \\ 1 & 2 & 3-x \end{pmatrix} = \dots = -x^3 + 6x^2 - 7x + 2 \\ &= -(x-1)(x^2 - 5x + 2) = 0 \\ &\Leftrightarrow x = 1 \text{ oder } x = \frac{5}{2} \pm \sqrt{\frac{25}{4} - \frac{8}{4}} \approx \frac{5}{2} \pm 2 = \begin{cases} \frac{9}{2} \\ \frac{1}{2} \end{cases}\end{aligned}$$

- 5.5.2004 (i) $f(v) = xv \Leftrightarrow \dots \Leftrightarrow v \in \text{Kern}(f - x \text{id}_V) = L(f - x \text{id}_V, 0) \equiv$ Lösen homogener linearer Gleichungen \equiv kennen wir schon!
- (ii) Aber: Die Berechnung der Eigenwerte selber ist das Auffinden von Nullstellen des „**charakteristischen Polynoms**“

$$\det(A - xE_n) = (-1)^n x^n + a_1 x^{n-1} + \dots + a_n = 0.$$

8.6 Algebraische Abgeschlossenheit

Definition. Ein Körper K heißt „**algebraisch abgeschlossen**“, falls gilt

$$\forall n \in \mathbb{N} \forall a_1, \dots, a_n \in K \exists x \in K : x^n + a_1 x^{n-1} + \dots + a_n = 0.$$

In Worten: Jedes Polynom mit Koeffizienten in K hat eine Nullstelle in K .

Bemerkung.

- (i) \mathbb{R} ist nicht algebraisch abgeschlossen, denn $x^2 + 1 = 0$ hat keine Nullstelle x in \mathbb{R} .
- (ii) \mathbb{C} ist algebraisch abgeschlossen, was die Aussage des „**Fundamentalsatzes der Algebra**“ ist und in der Algebra-Vorlesung bewiesen wird.
- (iii) Endliche Körper sind nie algebraisch abgeschlossen, denn im Fall $K = \{x_1, \dots, x_n\}$ ist $(x_1 - x) \cdots (x_n - x) + 1 \neq 0$ für alle $x \in K$. Das heißt, jeder algebraisch abgeschlossene Körper ist unendlich.
- (iv) Ist K algebraisch abgeschlossen, dann hat jedes charakteristische Polynom

$$\det(A - xE_n) = 0$$

eine Nullstelle x . Somit gilt

$$\forall f \in \text{End}_K(V) : \text{Spec } f \neq \emptyset;$$

$$\forall A \in K^{n \times n} : \text{Spec } A \neq \emptyset.$$

8.7 Darstellungssatz für beliebige Endomorphismen (Jordan²-Normalform)

Satz. Sei K algebraisch abgeschlossen, V ein endlich-dimensionaler K -Vektorraum und $f \in \text{End}_K(V)$. Dann existiert eine Basis $B \subseteq V$, bzgl. der die Darstellungsmatrix von f Blockdiagonalgestalt hat:

$$M(f) = \begin{pmatrix} J_1 & & 0 \\ & J_2 & \\ & & \ddots \\ 0 & & & J_t \end{pmatrix}$$

mit „Jordan-Blöcken“ der Form

$$J_i = \begin{pmatrix} x_i & & & 0 \\ 1 & x_i & & \\ & \ddots & \ddots & \\ & & \ddots & x_i \\ 0 & & & 1 & x_i \end{pmatrix}_{n_i \times n_i}$$

für $n_i \in \mathbb{N}, x_i \in K, i = 1, \dots, t$.

Beweis. Induktion nach $n := \dim V$:

Induktionsanfang. $n = 1$:

$\forall v \in V \cong K : f(v) = vf(1) = x_1v$ mit $x_1 := f(1)$. Also $M(f) = (x_1) \equiv (J_1)$ mit simplem Jordan-Block $J_1 = (x_1) \in K^{1 \times 1}$.

Induktionsschritt von $n - 1$ auf n :

Da K algebraisch abgeschlossen ist, existieren Eigenwerte:

$$\exists x \in \text{Spec } f \exists m \in \mathbb{N} : \{0\} \neq E_x(f) \subseteq H_x(f) = \text{Kern}((f - x \text{id}_V)^m)$$

I. Zunächst studieren wir die eingeschränkte Abbildung

$$g := (f - x \text{id}_V)|_{H_x(f)} : H_x(f) \rightarrow H_x(f)$$

(wohldefiniert gemäß Bemerkung (ii) aus Kapitel 8.5 auf Seite 19). Aus der Bemerkung in Kapitel 8.4 auf Seite 18 wissen wir, daß g nilpotent ist. Der Darstellungssatz für nilpotente Endomorphismen liefert eine Basis von $H_x(f)$, bzgl. der g die Darstellungsmatrix

$$\begin{pmatrix} M_1 & & 0 \\ & \ddots & \\ 0 & & M_l \end{pmatrix}$$

²Camille JORDAN, 1838–1922

8 Normaldarstellungen für Matrizen von Vektorraum-Endomorphismen

mit Blöcken

$$M_i = \begin{pmatrix} 0 & & & 0 \\ 1 & 0 & & \\ & \ddots & \ddots & \\ 0 & & 1 & 0 \end{pmatrix}$$

hat.

II. Betrachte die eingeschränkte Abbildung

$$f|_{H_x(f)} : H_x(f) \rightarrow H_x(f)$$

(wohldefiniert gemäß Bemerkung (iii) aus Kapitel 8.5 auf Seite 19). Offenbar $f|_{H_x(f)} = g + x \operatorname{id}_{h_x(f)}$. Somit die Darstellungsmatrix

$$M(f|_{H_x(f)}) = \begin{pmatrix} J_1 & & & 0 \\ & J_2 & & \\ & & \ddots & \\ 0 & & & J_l \end{pmatrix}$$

mit Blöcken

$$J_i = M_i + x E_{n_i} = \begin{pmatrix} x_i & & & & 0 \\ 1 & x_i & & & \\ & \ddots & \ddots & & \\ & & \ddots & x_i & \\ 0 & & & 1 & x_i \end{pmatrix}.$$

III. Nach Satz von Fitting gilt $V = H_x(f) \oplus U$ mit $U := \operatorname{Bild}((f - x \operatorname{id}_V)^m)$. Wegen $H_x(f) \neq \{0\}$ ist $\dim U < n$. Nach Induktionsvoraussetzung existiert eine Basis von U , bzgl. der $f|_U : U \rightarrow U$ eine Darstellungsmatrix der Form

$$M(f|_U) = \begin{pmatrix} J_{l+1} & & & 0 \\ & J_{l+2} & & \\ & & \ddots & \\ 0 & & & J_{l+t} \end{pmatrix}$$

mit Jordan-Blöcken J_i hat.

Zusammen ergibt sich nach der Bemerkung (iii) in Kapitel 8.2 auf Seite 10 die Darstellungsmatrix

$$M(f) = \begin{pmatrix} M(f|_{h_x(f)}) & & & 0 \\ & \ddots & & \\ 0 & & & M(f|_U) \end{pmatrix}$$

mit Jordan-Gestalt. □

8.7 Darstellungssatz für beliebige Endomorphismen (Jordan-Normalform)

Bemerkung.

- (i) Die Reihenfolge der Jordan-Blöcke ist nicht eindeutig, sondern ändert sich mit Umordnung der Basis.
- (ii) Sofortige Auswirkung auf das charakteristische Polynom von f :

$$\det(f - x \operatorname{id}_V) = \det(M(f) - xE_n) \stackrel{\text{Dreiecks-Mx.}}{=} (x_1 - x)^{n_1} \cdots (x_t - x)^{n_t} = 0$$

$$\Leftrightarrow x \in \{x_1, \dots, x_t\}$$

In Worten: Auf der Hauptdiagonalen der Jordan-Darstellung stehen nur Eigenwerte von f .

Warnung. Es muß nicht gelten, daß x_1, \dots, x_t paarweise verschieden sind. Gegenbeispiel:

$$E_n = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}.$$

- (iii) Da K als algebraisch abgeschlossen angenommen wird, hat *jeder* Endomorphismus f eine Jordan-Darstellung. Wenn dagegen nur ein (spezieller) Endomorphismus interessiert, dann reicht es, daß nur *sein* charakteristisches Polynom zerfällt gemäß $\det(f - x \operatorname{id}_V) = (x_1 - x)^{n_1} \cdots (x_t - x)^{n_t}$.

Beispiel. Sei $f \in \operatorname{End}_{\mathbb{R}}(\mathbb{R}^3)$ gegeben durch $f(v) = Av$ mit

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

I. Das charakteristische Polynom ist

$$\det(A - xE_3) = \det \begin{pmatrix} 2-x & -1 & 0 \\ 1 & 1-x & -1 \\ 1 & -1 & 1-x \end{pmatrix} = \cdots = (1-x)^2(2-x)$$

und $\operatorname{Spec} A = \{1, 2\}$.

II. Zum Eigenwert 1:

$$\operatorname{Kern}(A - E_3) = \operatorname{Kern} \begin{pmatrix} 1 & -1 & 0 \\ 1 & 0 & -1 \\ 1 & -1 & 0 \end{pmatrix} = \left\{ \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \middle| y_1 = y_2 = y_3 \right\} = \operatorname{span} \left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\}$$

$$\begin{aligned} \operatorname{Kern}((A - E_3)^2) &= \operatorname{Kern} \begin{pmatrix} 0 & -1 & 1 \\ 0 & 0 & 0 \\ 0 & -1 & 1 \end{pmatrix} = \left\{ \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \middle| y_1 \in \mathbb{R}, y_2 = y_3 \right\} \\ &= \operatorname{span} \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \right\} \end{aligned}$$

8 Normaldarstellungen für Matrizen von Vektorraum-Endomorphismen

$$(A - E_3)^3 = (A - E_3)^2 \Rightarrow m = 2$$

Schlaue Basis von $H_1(f) = \text{Kern}((A - E_3)^2)$ also

$$b_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, b_2 = (f - \text{id})(b_1) = (A - E_3)(b_1) = \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} - \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

III.

$$U = \text{Bild}((A - E_3)^2) = \left\{ \begin{pmatrix} y_3 - y_2 \\ 0 \\ y_3 - y_2 \end{pmatrix} \mid y_1, y_2, y_3 \in \mathbb{R} \right\} = \text{span} \left\{ \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \right\}$$

Mit

$$b_3 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$$

gilt

$$f|_U(b_3) = Ab_3 = \begin{pmatrix} 2 \\ 0 \\ 2 \end{pmatrix} = 2 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = 2b_3.$$

Zusammen hat f bzgl. der Basis $\{b_1, b_2, b_3\}$ die Darstellungsmatrix

$$M(f) = \left(\begin{array}{cc|c} 1 & 0 & 0 \\ 1 & 1 & 0 \\ \hline 0 & 0 & 2 \end{array} \right).$$

Probe:

$$f(b_1) = Ab_1 = \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} = b_1 + b_2 \checkmark$$

$$f(b_2) = Ab_2 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = b_2 \checkmark$$

$$f(b_3) = Ab_3 = \begin{pmatrix} 2 \\ 0 \\ 2 \end{pmatrix} = 2b_3 \checkmark$$

6. 5. 2004 Bessere Formulierung (für Anwendungen einschließlich $K = \mathbb{R}$):

Satz. Sei K ein beliebiger Körper und $f \in \text{End}_K(V)$. Dann gilt:
Das charakteristische Polynom von f zerfällt in Linearfaktoren, d. h.

$$\det(f - x \text{id}_V) = (x_1 - x)^{m_1} \cdots (x_s - x)^{m_s}.$$

\Leftrightarrow Es gibt eine Basis, bzgl. der die Darstellungsmatrix in Jordan-Darstellung ist.

Zusatz.

K algebraisch abgeschlossen $\Rightarrow \forall f \in \text{End}_K(V) \exists \text{Basis} : M(f)$ Jordan-Darstellung

8.8 Eindeutigkeit der Jordan-Darstellung**Definition („Defekt“³).**

$$\forall f \in \text{End}(V) : \text{def } f := \dim(\text{Kern } f) = \dim V - \text{Rang } f$$

$$\forall A \in K^{n \times n} : \text{def } A := \dim(\text{Kern } A) = n - \text{Rang } A$$

Bemerkung. (i) Hat A Zeilen-Stufen-Form, dann ist $\text{def } A$ die Anzahl der Nullzeilen.

(ii) *Beispiel.*

$$N_{k \times k} = \begin{pmatrix} 0 & & & 0 \\ 1 & 0 & & \\ & \ddots & \ddots & \\ 0 & & 1 & 0 \end{pmatrix}$$

sind die Blöcke aus Kapitel 8.4. Offenbar:

$$N_k^2 = \begin{pmatrix} 0 & & & & 0 \\ 0 & \ddots & & & \\ 1 & \ddots & \ddots & & \\ \vdots & \ddots & \ddots & \ddots & \\ 0 & \cdots & 1 & 0 & 0 \end{pmatrix} \quad N_k^3 = \begin{pmatrix} 0 & & & & & 0 \\ 0 & \ddots & & & & \\ 0 & \ddots & \ddots & & & \\ 1 & \ddots & \ddots & \ddots & & \\ \vdots & \ddots & \ddots & \ddots & \ddots & \\ 0 & \cdots & 1 & 0 & 0 & 0 \end{pmatrix}$$

$N_k^k = 0$. Daher ist N_k nilpotent von Index k und

$$\text{def } N_k^j = \begin{cases} j & \text{für } j \leq k \\ k & \text{für } j > k. \end{cases}$$

Damit kompakte Notation in Kapitel 8.4:

$$M(f) = \text{Blockdiag}(N_{n_1}, \dots, N_{n_t});$$

und in Kapitel 8.7:

$$M(f) = \text{Blockdiag}(J_1, \dots, J_t) \text{ mit } J_i = N_{n_i} + x_i E_{n_i}$$

³engl. *nullity*

8 Normaldarstellungen für Matrizen von Vektorraum-Endomorphismen

(iii) Hat A Blockdiagonalform:

$$A = \begin{pmatrix} B_1 & & 0 \\ & \ddots & \\ 0 & & B_t \end{pmatrix},$$

dann ist der Defekt „additiv“:

$$\text{def } A = \text{def } B_1 + \dots + \text{def } B_t.$$

Satz (Eindeutigkeit). Sei K ein beliebiger Körper und $f \in \text{End}_K(V)$. Sei außerdem $M(f) = \text{Blockdiag}(J_1, \dots, J_t)$ eine Jordan-Darstellung mit $J_i = N_{n_i} + x_i E_{n_i}$. Dann sind für $j \in \mathbb{N}$ und $x \in K$ die Anzahlen

$$m_j(x) := \#\{i \in \{1, \dots, t\} \mid n_i = j \text{ und } x_i = x\}$$

bestimmt durch die Defekte von $(f - x \text{id}_V)^j$ gemäß

$$m_1(x) = 2 \text{def}(f - x \text{id}_V) - \text{def}((f - x \text{id}_V)^2)$$

und für $j \geq 2$: $m_j(x) = 2 \text{def}((f - x \text{id}_V)^j) - \text{def}((f - x \text{id}_V)^{j+1}) - \text{def}((f - x \text{id}_V)^{j-1})$.

Insbesondere sind die Jordan-Blöcke J_1, \dots, J_t bis auf die Reihenfolge eindeutig.

Bemerkung. Bisher: eine Jordan-Darstellung. Nunmehr: die Jordan-Darstellung.

Beweis. I. Für $j \in \mathbb{N}$ und $x \in K$ gilt

$$\text{def}((f - x \text{id}_V)^j) = \text{def}((M(f) - x E_n)^j).$$

Da $M(f)$ blockdiagonal ist, ist $M(f) - x E_n$ blockdiagonal. Dafür gilt

$$\begin{pmatrix} B_1 & & 0 \\ & \ddots & \\ 0 & & B_t \end{pmatrix}^j = \begin{pmatrix} B_1^j & & 0 \\ & \ddots & \\ 0 & & B_t^j \end{pmatrix}.$$

Somit:

$$\text{def}((f - x \text{id}_V)^j) = \text{def}((J_1 - x E_{n_1})^j) + \dots + \text{def}((J_t - x E_{n_t})^j)$$

II. Für $i = 1, \dots, t$ ist $J_i - x E_{n_i} = N_{n_i} + (x_i - x) E_{n_i}$ eine Dreiecksmatrix mit Defekt $\text{def}(J_i - x E_{n_i}) = (x_i - x)^{n_i}$. Im Fall $x \notin \{x_1, \dots, x_t\}$ ist also $\det \neq 0$ und die Matrix invertierbar und $\text{def} = 0$.

III. Im Fall $x = x_i$, $i = 1, \dots, t$ ist $J_i - x_i E_{n_i} = N_{n_i}$ (siehe Beispiel auf der vorherigen Seite) und es gilt

$$\text{def}((J_i - x_i E_{n_i})^j) = \begin{cases} j & \text{für } j \leq n_i \\ n_i & \text{für } j > n_i. \end{cases}$$

IV. Zusammen ergibt sich

$$\begin{aligned}\operatorname{def}(f - x \operatorname{id}_V) &= m_1(x) + m_2(x) + \dots \\ \operatorname{def}((f - x \operatorname{id}_V)^2) &= m_1(x) + 2m_2(x) + 2m_3(x) + \dots \\ \operatorname{def}((f - x \operatorname{id}_V)^3) &= m_1(x) + 2m_2(x) + 3m_3(x) + 3m_4(x) + \dots\end{aligned}$$

und daraus $m_1(x)$ und $m_j(x)$ für $j > 1$.

□

8.9 Matrizenformulierung

Satz. Sei K ein beliebiger Körper und $A \in K^{n \times n}$. Dann zerfällt das charakteristische Polynom $\det(A - xE_n)$ genau dann in Linearfaktoren, wenn $\exists S \in \operatorname{GL}(n) : S^{-1}AS = \operatorname{Blockdiag}(J_1, \dots, J_t)$ mit Jordan-Gestalt für $J_i \forall i = 1, \dots, t$.

Beweis. Sei

$$v_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \dots, v_n = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

die Standardbasis des K^n . Betrachte den Endomorphismus $f(v) = Av$. Sei b_1, \dots, b_n die Basis gemäß Kapitel 8.7. Sei S bestimmt durch $b_j = \sum_{i=1}^n s_{ij}v_i$. Dann liefert das Basiskonvertierungslemma die neue Darstellungsmatrix $S^{-1}AS = M(f) = \operatorname{Jordan-Darstellung}$. □

In Worten: A ist ähnlich zu einer Jordan-Darstellung.

8.10 Satz über Ähnlichkeitstransformationen

Satz. Seien K ein algebraisch abgeschlossener Körper und $A, B \in K^{n \times n}$. Dann sind folgende Aussagen äquivalent:

- (i) $\exists S \in \operatorname{GL}(n) : S^{-1}AS = B$;
- (ii) $\forall j \in \mathbb{N} \forall x \in K : \operatorname{def}((A - xE_n)^j) = \operatorname{def}((B - xE_n)^j)$;
- (iii) $\forall j = 1, \dots, n \forall x \in \operatorname{Spec} A : \operatorname{def}((A - xE_n)^j) = \operatorname{def}((B - xE_n)^j)$ **und** $\operatorname{Spec} A = \operatorname{Spec} B$.

durch Ringschluß.

(i) \Rightarrow (ii):

Bzgl. der Standardbasis v_1, \dots, v_n im K^n hat der durch $f(v) = Av$ definierte Endomorphismus die Darstellungsmatrix A . Bzgl. der durch S induzierten neuen Basis $\tilde{v}_j = \sum s_{ij}v_j$ hat f die Darstellungsmatrix $S^{-1}AS \underset{(\text{Vor.})}{=} B$. Somit $\operatorname{def}((A - xE_n)^j) = \operatorname{def}((f - x \operatorname{id}_V)^j) = \operatorname{def}((B - xE_n)^j)$.

8 Normaldarstellungen für Matrizen von Vektorraum-Endomorphismen

(ii)⇒(iii):

Aus (ii) folgt für alle $x \in K$:

$$\begin{aligned} x \in \text{Spec } A &\stackrel{(j=1)}{\Leftrightarrow} \text{def}(A - xE_n) \neq 0 \\ &\Leftrightarrow \text{def}(B - xE_n) \neq 0 \\ &\Leftrightarrow x \in \text{Spec } B \end{aligned}$$

(iii)⇒(i)

Bzgl. der Standardbasis v_1, \dots, v_n des K^n definieren wir die Endomorphismen $f(v) := Av$ und $g(v) := Bv$. Diese erben aus (iii)

$$\text{Spec } f = \text{Spec } A = \text{Spec } B = \text{Spec } g$$

und

$$\text{def}((f - x \text{id}_V)^j) = \dots = \text{def}((g - x \text{id}_V)^j).$$

Nach den Kapiteln 8.7 und 8.8 gibt es Basen $B := \{b_1, \dots, b_n\}$ und $\tilde{B} := \{\tilde{b}_1, \dots, \tilde{b}_n\}$ so, daß f bzgl. B Jordan-Gestalt-Darstellungsmatrix besitzt und g bzgl. \tilde{B} Jordan-Gestalt-Darstellungsmatrix besitzt und die beiden Jordan-Darstellungen gleich sind. Seien S und \tilde{S} die Basistransformationsmatrizen in B und \tilde{B} . Somit $S^{-1}AS = M(f) = \tilde{M}(g) = \tilde{S}^{-1}B\tilde{S}$. Mit $T := S\tilde{S}^{-1}$ gilt $T^{-1}AT = B$. \square

12. 5. 2004 *Beispiel.* K algebraisch abgeschlossen, $A \in K^{n \times n} \Rightarrow A, A^T$ ähnlich.

Beweis. Mit (ii):

$$\forall j \in \mathbb{N} \forall x \in K :$$

$$\begin{aligned} \text{def}((A^T - xE_n)^j) &= n - \text{Rang}(\underbrace{(A^T - xE_n)^j}_{=((A - xE_n)^j)^T}) \\ &= n - \text{Rang}((A - xE_n)^j) = \text{def}((A - xE_n)^j) \end{aligned}$$

\square

8.11 Satz über diagonale Darstellungen (Diagonalisierbarkeit)

Satz. Sei $f \in \text{End}_K(V)$. Dann sind folgende Aussagen äquivalent:

(i) f diagonalisierbar (d. h. es existiert eine Basis von V , bzgl. der die Darstellungsmatrix von f diagonal ist);

$$(ii) V = \bigoplus_{x \in \text{Spec } f} E_x(f);$$

$$(iii) V = \sum_{x \in \text{Spec } f} E_x(f).$$

8.11 Satz über diagonale Darstellungen (Diagonalisierbarkeit)

Beweis.

(i)⇒(iii)

Wegen (i) existiert eine Basis $B = \{b_1, \dots, b_n\}$ mit

$$M(f) = \text{Diag}(x_1, \dots, x_n) = \begin{pmatrix} x_1 & & 0 \\ & \ddots & \\ 0 & & x_n \end{pmatrix}.$$

Es folgt

$$\forall x \in K : \det(f - x \text{id}_V) = \det(M(f) - xE_n) = \prod_{i=1}^n (x_i - x)$$

und

$$x \in \text{Spec } f \Leftrightarrow \det(f - x \text{id}) = 0 \Leftrightarrow x \in \{x_1, \dots, x_n\}.$$

Also $\text{Spec } f = \{x_1, \dots, x_n\}$. Wegen $f(b_i) = x_i b_i$ ist

$$b_i \in E_{x_i}(f) \subseteq \sum_{x \in \text{Spec } f} E_x(f) \subseteq V.$$

Somit

$$V = \text{span}\{b_1, \dots, b_n\} \subseteq \sum E_x(f) \subseteq V.$$

(iii)⇒(ii)

Sei $x \in \text{Spec } f$ und $v \in E_x(f) \cap \sum_{\substack{y \in \text{Spec } f \\ y \neq x}} E_y(f)$. Zu zeigen $v = 0$.

Jeder Untervektorraum $E_y(f)$ besitzt eine Basis aus Eigenvektoren zum Eigenwert y . Sei also $\{y_1, \dots, y_k\} = \text{Spec}(f) \setminus \{x\}$ und $\{b_1, \dots, b_l\}$ eine Basis von $\sum_{y \neq x} E_y(f) = \sum_{i=1}^k E_{y_i}(f)$. Also existieren $z_1, \dots, z_l \in K$ mit $v = z_1 b_1 + \dots + z_l b_l$. Dann $x z_1 b_1 + \dots + x z_l b_l = x v = f(v) = z_1 f(b_1) + \dots + z_l f(b_l) = z_1 y_{i_1} b_1 + \dots + z_l y_{i_l} b_l$. Koeffizientenvergleich liefert:

$$z_1 \underbrace{(x - y_{i_1})}_{\neq 0} = 0, \dots, z_l \underbrace{(x - y_{i_l})}_{\neq 0} = 0$$

$$\Rightarrow z_1 = 0, \dots, z_l = 0 \Rightarrow v = 0.$$

(ii)⇒(i)

Ist $V = \bigoplus E_x(f)$, dann lassen sich Basen (aus Eigenvektoren) für $E_x(f)$ finden und zusammenfügen zu einer Basis $\{v_1, \dots, v_n\}$ von V . Wegen $f(v_j) = x_j v_j$ hat f die Darstellungsmatrix $M(f) = \text{Diag}(x_1, \dots, x_n)$. □

Bemerkung. Erinnerungswürdig:

$$\forall x, y \in \text{Spec } f : \quad x \neq y \Rightarrow E_x(f) \cap E_y(f) = \{0\}$$

Mit anderen Worten: Eigenvektoren zu unterschiedlichen Eigenwerten sind linear unabhängig.

8.12 Vorsatz zu Kap. 8.13

Satz. Sei $f \in \text{End}_K(V)$, f diagonalisierbar, $U \subseteq V$ Untervektorraum, U f -invariant. Dann ist $g := f|_U \in \text{End}_K(U)$ diagonalisierbar.

Beweis. Zu zeigen: $U \subseteq \sum_{x \in \text{Spec } f} E_x(g)$. Indirekt:

I. Sei $M := U \setminus \sum_{x \in \text{Spec } f} E_x(g) \neq \emptyset$. Für jedes $\tilde{u} \in M$ gilt $\tilde{u} \neq 0$. Wegen $V = \bigoplus E_x(f)$ existiert eine eindeutige „Zerlegung in Eigenvektoren“ gemäß $\tilde{u} = \sum_{x \in \text{Spec } f} u_x$ mit $u_x \in E_x(f)$, $u_x \neq 0$. Sei $m(\tilde{u})$ die Anzahl der Summanden. Wähle $u \in M$ mit $m(u) = \min_{\tilde{u} \in M} m(\tilde{u}) =: m$. Das heißt, u hat eine kürzestmögliche Zerlegung in Eigenvektoren $u = u_1 + \dots + u_m$ mit $u_i \in E_{x_i}(f)$.

II. Damit betrachte den Konkurrenten

$$\begin{aligned} v &:= (x_1 - x_1)u_1 + (x_2 + x_1)u_2 + \dots + (x_m - x_1)u_m \\ &= f(u) - x_1 u \in U \text{ (da } U \text{ } f\text{-invariant)}. \end{aligned}$$

Wegen $m(v) < m$ gilt $v \notin M$ und $v \in U \setminus M \subseteq \sum_{x \in \text{Spec } f} E_x(g)$. Also $v = \sum_{x \in \text{Spec } f} v_x$ mit $v_x \in E_x(g) = E_x(f) \cap U$. Damit notwendigerweise $v_{x_i} = (x_i - x_1)u_i$ für ein $i = 2, \dots, m$ und $v_x = 0$ sonst. Folglich mit $u_i \in E_{x_i}(g)$ für $i = 2, \dots, m$ $u_1 = u - u_2 - \dots - u_m \in U \cap E_{x_1}(f) = E_{x_1}(g)$ und $u \in \sum_{i=1}^m E_{x_i}(g) \subseteq \sum_{x \in \text{Spec } f} E_x(g) \subseteq U \setminus M \uparrow$.

□

8.13 Satz über simultane Diagonalisierbarkeit

Satz. Sei $H \subseteq \text{End}_K(V)$ und $\forall f, g \in H : f \circ g = g \circ f$. In Worten: H sei eine Menge kommutierender Endomorphismen. Dann gilt:

Wenn jeder Endomorphismus in H diagonalisierbar ist, dann existiert eine Basis von V , bzgl. der für jeden Endomorphismus $f \in H$ die Darstellungsmatrix diagonal ist.

Beweis. Induktion nach $n := \dim V$.

Induktionsanfang. $n = 1$ ✓

Induktionsschritt von $n - 1$ auf n .

Fall $\# \text{Spec } f = 1$ für alle $f \in H$:

Dann $f = x_f \text{id}_V$ und $M(f) = x_f E_n$ bezüglich beliebiger Basen. Alles kommutiert. ✓

Fall $\exists f \in H : \text{Spec } f = \{a_1, \dots, a_m\}$ mit $m \geq 2$.

Dann $V = \bigoplus_{i=1}^m E_{a_i}(f)$ mit $\dim E_{a_i}(f) < n$. Dabei $\forall v \in E_{a_i}(f) \forall g \in H : g(v) \in E_{a_i}(f)$, denn $f(g(v)) \underset{\text{vor.}}{=} g(f(v)) = g(a_i v) = a_i g(v)$. Also $E_{a_i}(f)$ invariant unter $g \in H$ und auf $H_i := \{g|_{E_{a_i}(f)} \mid g \in H\}$ läßt sich die Induktionsvoraussetzung anwenden.

Sei also B_i Basis von $E_{a_i}(f)$ so, daß alle Endomorphismen in H_i diagonale Darstellungsmatrizen haben. Dann tut es die Basis $B_1 \cup \dots \cup B_m$ von V für die Behauptung. □

13. 5. 2004 *Beispiel.* Seien $f, g \in \text{End}_{\mathbb{R}}(\mathbb{R}^4)$ wie immer durch ihre Darstellungsmatrizen bzgl. der Standardbasis gegeben gemäß

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

I.

$$AB = \begin{pmatrix} 0 & -3 & 0 & 4 \\ 0 & 1 & 0 & 0 \\ -2 & -3 & 2 & 4 \\ 0 & -3 & 0 & 4 \end{pmatrix} = BA$$

II. Für $x \in \mathbb{R}$ ist $\det(A - xE_4) = \dots = (1 - x)^2(2 - x)^2$. Also $\text{Spec } f = \{1, 2\}$. Man findet

$$E_1(A) = \text{span} \left\{ \underbrace{\begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}}_{b_1}, \underbrace{\begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}}_{b_2} \right\}, \quad E_2(A) = \text{span} \left\{ \underbrace{\begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}}_{b_3}, \underbrace{\begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}}_{b_4} \right\}.$$

Bzgl. $B = \{b_1, \dots, b_4\}$ hat f die Darstellungsmatrix

$$M(f) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix} = \text{Diag}(1, 1, 2, 2).$$

III. Betrachte $g|_{E_1(A)}$:

$$g(b_1) = Bb_1 = 0$$

$$g(b_2) = Bb_2 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = b_1 + b_2$$

$$M(g|_{E_1(A)}) = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} =: C$$

Für $x \in \mathbb{R}$ ist $\det(C - xE_2) = x(x - 1)$. Also $\text{Spec } C = \{0, 1\}$. Man findet:

$$E_0(g|_{E_1(f)}) = \text{span} \left\{ \underbrace{b_1}_{=:c_1} \right\}, \quad E_1(g|_{E_1(f)}) = \text{span} \left\{ \underbrace{b_1 + b_2}_{=:c_2} \right\}.$$

8 Normaldarstellungen für Matrizen von Vektorraum-Endomorphismen

IV. Betrachte $g|_{E_2(A)}$:

$$g(b_3) = \begin{pmatrix} 2 \\ 0 \\ 1 \\ 2 \end{pmatrix} = 2b_3 + b_4$$

$$g(b_4) = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} = b_4$$

$$M(g|_{E_2(A)}) = \begin{pmatrix} 2 & 0 \\ 1 & 1 \end{pmatrix} =: D$$

$$\det(D - xE_2) = (1 - x)(2 - x)$$

$$\text{Spec } D = \{1, 2\}$$

$$E_1(g|_{E_2(A)}) = \text{span}\{\underbrace{b_4}_{=:c_3}\}$$

$$E_2(g|_{E_2(A)}) = \text{span}\{\underbrace{b_3 + b_4}_{=:c_4}\}$$

V. Zusammenfassung:

Bezüglich der Basis $\{c_1, \dots, c_4\}$ hat sowohl f als auch g diagonale Darstellungsmatrix:

$$\tilde{M}(f) = \begin{pmatrix} 1 & & & 0 \\ & 1 & & \\ & & 2 & \\ 0 & & & 2 \end{pmatrix} \quad \tilde{M}(g) = \begin{pmatrix} 0 & & & 0 \\ & 1 & & \\ & & 1 & \\ 0 & & & 2 \end{pmatrix}$$

9 Transformationsgruppen als elementarste Bausteine der Mathematik

9.1 Gruppen

Definition. (G, \circ) heißt „Gruppe“, falls G Menge und $\circ : G \times G \rightarrow G$ erfüllt:

(G1) Assoziativität: $\forall g, h, k \in G : (g \circ h) \circ k = g \circ (h \circ k)$.

(G2) Existenz eines neutralen Elements (Identität): $\exists e \in G \forall g \in G : e \circ g = g$.

(G3) Existenz inverser Elemente: $\forall g \in G \exists g^{-1} \in G : g^{-1} \circ g = e$.

Bemerkung.

(i) (G2) ohne Eindeutigkeit und außerdem genauer „linksneutral“; (G3) genauer „linksinvers“.

(ii) Zugabe:

(G4) $\forall g, h \in G : g \circ h = h \circ g$

\leadsto Gruppe heißt „abelsch“¹.

Satz (Längst überfälliges Eindeutigkeitslemma).

(i) *linksinvers = rechtsinvers*

Beweis. Sei $h := g^{-1}$ linksinvers zu g . Dann:

$$g \circ g^{-1} = e \circ g \circ g^{-1} = h^{-1} \circ h \circ g \circ g^{-1} = h^{-1} \circ e \circ g^{-1} = h^{-1} \circ g^{-1} = e$$

□

(ii) *linksneutral = rechtsneutral*

Beweis.

$$g \circ e = g \circ g^{-1} \circ g \underset{(i)}{=} e \circ g = g$$

□

(iii) *Das neutrale Element ist eindeutig.*

¹Niels Henrik ABEL, 1802–1829

9 Transformationsgruppen als elementarste Bausteine der Mathematik

Beweis. Sei $e \circ g = g = \tilde{e} \circ g$ für alle $g \in G$. Dann folgt (mit $g = e$) $\tilde{e} \circ e = e$ und (mit $g = \tilde{e}$ und (ii)) $\tilde{e} \circ e = \tilde{e}$. Zusammen folgt $e = \tilde{e}$. \square

(iv) *Inverse sind eindeutig.*

Beweis.

$$g^{-1} \circ g = e = h \circ g \quad \Rightarrow \quad h \underset{(ii)}{=} h \circ e \underset{(i)}{=} h \circ g \circ g^{-1} = e \circ g^{-1} = g^{-1}$$

\square

Beispiel.

I. Zahlen und Vektoren

- (i)
 - $(\mathbb{Z}, +)$ abelsche Gruppe
 - $(2\mathbb{Z}, +)$ abelsche Gruppe
 - $(3\mathbb{Z}, +)$ abelsche Gruppe
- (ii) $(K, +, \cdot)$ Körper $\Rightarrow (K, +)$ abelsche Gruppe („additive Gruppe von K “) und $(K \setminus \{0\}, \cdot)$ abelsche Gruppe („multiplikative Gruppe von K “).

II. Matrizen-Gruppen

- (i) $GL_K(n) = \{A \in K^{n \times n} \mid \det A \neq 0\}$ generelle lineare Gruppe ($\circ \equiv$ Hintereinanderausführung, nichtabelsch).
- (ii) allgemeiner:
Sei X Menge und $G = \{f : X \rightarrow X \mid f \text{ bijektiv}\}$. Dann ist (G, \circ) Gruppe.
- (iii) $SL_K(n) := \{A \in K^{n \times n} \mid \det A = 1\}$ heißt „**spezielle lineare Gruppe**“. Tatsächlich $\forall A, B \in SL_K(n) : \det(AB) = \det A \cdot \det B = 1$, also ist Matrizenmultiplikation auf $SL_K(n)$ wohldefiniert.

Bemerkung. Manchmal „koordinatenfreie“ Notation:

$$GL_K(V) = \{f : V \rightarrow V \mid f \text{ linear und } f \text{ bijektiv}\}$$
$$SL_K(V) = \dots$$

III. Permutationsgruppen

- (i) $S_n := \{f : \{1, \dots, n\} \rightarrow \{1, \dots, n\} \mid f \text{ bijektiv}\}$ heißt „**Permutationsgruppe auf n Elementen**“ (mit \circ als Hintereinanderausführung von Abbildungen).

Bemerkung.

- a) Permutation = Bijektion einer endlichen Menge auf sich.
- b) Auch genannt „symmetrische Gruppe von n Elementen“.
- (ii) $A_n := \{f \in S_n \mid \text{sig } f = 1\}$ heißt „**alternierende Permutationsgruppe**“. Tatsächlich $\forall f, \tilde{f} \in A_n : \text{sig}(f \circ \tilde{f}) = \text{sig } f \cdot \text{sig } \tilde{f} = 1$.

19. 5. 2004

Definition. Seien $(G, \circ), (H, \diamond)$ Gruppen. Eine Abbildung $f : G \rightarrow H$ heißt „**(Gruppen-)Homomorphismus**“, falls gilt $\forall g, \tilde{g} \in G : f(g \circ \tilde{g}) = f(g) \diamond f(\tilde{g})$.

Beispiel.

(i) Die Determinante

$$\det : \text{GL}_K(n) \rightarrow (K \setminus \{0\}, \cdot)$$

ist Gruppenhomomorphismus, denn $\det(AB) = \det A \det B$. Speziell:

$$\begin{aligned} \det E_n &= 1; \\ \det(A^{-1}) &= \frac{1}{\det A}. \end{aligned}$$

Zudem: $\text{SL}_K(n) = \det^{-1}(\{1\})$ ist das Urbild unter \det von 1.

(ii) Ebenso für Permutationen:

$$\text{sig} : S_n \rightarrow (\{-1, +1\}, \cdot)$$

ist Gruppenhomomorphismus, denn $\forall f, \tilde{f} \in S_n : \text{sig}(f \circ \tilde{f}) = \text{sig } f \cdot \text{sig } \tilde{f}$. Zudem: $A_n = \text{sig}^{-1}(\{1\})$.

9.2 Untergruppen, Kern, Bild

Definition. Ist (G, \circ) und $U \subseteq G$ Teilmenge, für die die Einschränkung $\circ|_U : U \times U \rightarrow U$ erfüllt und die Gruppeneigenschaften (G1), (G2), (G3) hat, dann heißt (U, \circ) „**Untergruppe von G** “.

Bemerkung.

(i) Immer gilt: U Untergruppe $\Rightarrow e \in U \Rightarrow U \neq \emptyset$.

(ii) Untergruppenkriterium:

$$U \text{ Untergruppe} \Leftrightarrow e \in U \text{ und } \forall g, h \in U : g \circ h^{-1} \in U$$

Satz. Sei $f : G \rightarrow H$ ein Gruppenhomomorphismus. Dann gilt:

$$(i) f(e_G) = e_H$$

$$(ii) \forall g \in G : f(g^{-1}) = (f(g))^{-1}$$

Beweis.

(i)

$$f(e_G) = e_H \diamond f(e_G) = (f(e_G))^{-1} \diamond \underbrace{f(e_G) \diamond f(e_G)}_{=f(e_G \circ e_G)} = e_H$$

$$(ii) \quad f(g) \diamond f(g^{-1}) = f(g \circ g^{-1}) = f(e_G) \stackrel{(i)}{=} e_H \Rightarrow f(g^{-1}) = (f(g))^{-1}$$

□

Satz. Sei $f : G \rightarrow H$ ein Gruppenhomomorphismus. Dann gilt:

(i) Bild $f := \{f(g) | g \in G\}$ ist Untergruppe von H .

(ii) Kern $f := \{g \in G | f(g) = e_H\}$ ist Untergruppe von G .

Beweis. Mit Untergruppenkriterium klar.

□

9.3 Normalteiler

Definition. Eine Untergruppe N von G heißt „Normalteiler“ oder „normale Untergruppe“, falls gilt:

$$\forall g \in G : g^{-1}Ng := \{g^{-1} \circ n \circ g | n \in N\} = N$$

Beispiel. Für alle Homomorphismen $f : G \rightarrow H$ ist Kern $f :=: N$ Normalteiler, denn sei $g \in G$ und $n \in N$ so, daß $f(n) = e_H$.

$$\stackrel{„\subseteq“}{f(g^{-1}ng)} = f(g^{-1}) \underbrace{f(n)}_{=e_H} f(g) = e_H \Rightarrow g^{-1}ng \in N.$$

„ \supseteq “
Setze $h := gn g^{-1}$. Dann $f(h) = e_H$ und $h \in N$. Außerdem $n = g^{-1}hg \in g^{-1}Ng$.

Satz (Motivationslemma). Sei $N \subseteq G$ Untergruppe. Dann gilt:

$$N \text{ Normalteiler} \Leftrightarrow \forall g \in G : Ng := \{ng | n \in N\} = \{gn | n \in N\} =: gN$$

In Worten: Rechtsnebenklassen = Linksnebenklassen.

Beweis. Siehe Beutelspacher.

□

Satz (über Faktorgruppen). Sei G Gruppe, $N \subseteq G$ Normalteiler. Dann ist $G/N := \{gN | g \in G\}$ mit der Verknüpfung $(gN) \circ (hN) := (g \circ h)N$ eine Gruppe.

Bemerkung. G/N heißt „Faktorgruppe“.

Beweis.

I. Wohldefiniertheit:

Für $gN = \tilde{g}N$ und $hN = \tilde{h}N$ ist zu zeigen: $(gh)N = (\tilde{g}\tilde{h})N$. Dazu: wegen $\tilde{g}e \in gN$ existiert $n_1 \in N$ mit $\tilde{g} = gn_1$. Also gilt $g^{-1}\tilde{g} = n_1 \in N$. Analog $n_2 := h^{-1}\tilde{h} \in N$.

Damit:

$$\begin{aligned} (gh)^{-1}(\tilde{g}\tilde{h}) &= h^{-1}g^{-1}\tilde{g}\tilde{h} = h^{-1} \underbrace{n_1\tilde{h}}_{\in N\tilde{h}=\tilde{h}N} = h^{-1}\tilde{h}n_3 \quad \text{für passendes } n_3 \in N \\ &= n_2n_3 \end{aligned}$$

Daraus folgt $(gh)N = (\tilde{g}\tilde{h})N$.

II. „ \subseteq “

Sei $ghn \in (gh)N$. Dann mit $n_0 := n_3^{-1}n_2^{-1}n$ gilt:

$$(\tilde{g}\tilde{h}) \underbrace{(n_2n_3)^{-1}n}_{\in N} = \tilde{g}\tilde{h}((gh)^{-1}(\tilde{g}\tilde{h}))^{-1}n = (\tilde{g}\tilde{h})(\tilde{g}\tilde{h})^{-1}ghn = ghn.$$

Also $ghn \in (\tilde{g}\tilde{h})N$.

III. „ \supseteq “

Analog. Gruppeneigenschaften sind dann klar.

□

9.4 Homomorphiesatz für Gruppen

Satz. Sei $f : F \rightarrow H$ ein Gruppenhomomorphismus. Dann gibt es einen Gruppenisomorphismus φ von $G/\text{Kern } f$ auf $\text{Bild } f$.

Beweis.

I. $\varphi(gN) := f(g) : G/\text{Kern } f \rightarrow \text{Bild } f$ ist wohldefiniert, denn aus $gN = \tilde{g}N$ folgt $n_1 := g^{-1} \circ \tilde{g} \in N := \text{Kern } f$ und $e_H = f(n_1) = f(g^{-1}) \diamond f(\tilde{g})$, woraus $f(g) = f(\tilde{g})$ folgt.

II. φ ist Gruppenhomomorphismus, denn $\varphi(gN \circ hN) = \varphi((gh)N) = f(gh) = f(g)f(h) = \varphi(gN)\varphi(hN)$.

III. φ surjektiv ✓

IV. φ injektiv, denn $\varphi(gN) = \varphi(hN) \Rightarrow f(g) = f(h) \Rightarrow g^{-1}h \in N \Rightarrow gN = hN$

□

10 Euklidische Vektorräume

Ein euklidischer Vektorraum ist ein endlichdimensionaler K -Vektorraum mit einem „Skalarprodukt“ (\equiv „inneres Produkt“ \equiv positiv definite, symmetrische Bilinearform).

10.1 Bilinearformen

Definition.

- (i) Eine Abbildung $B : V \times V \rightarrow K$ heißt „**Bilinearform** (auf V)“, falls gilt

$$\begin{aligned} \forall v \in V : B(v, \cdot) : V \rightarrow K \text{ linear} \quad \text{und} \\ \forall w \in V : B(\cdot, w) : V \rightarrow K \text{ linear.} \end{aligned}$$

- (ii) Eine Bilinearform B auf V heißt „**symmetrisch**“, falls gilt

$$\forall v, w \in V : B(v, w) = B(w, v).$$

- (iii) Eine Bilinearform B auf V heißt „**nicht entartet**“, „**nicht degeneriert**“, falls gilt

$$\forall v \in V : (\forall w \in V : B(v, w) = 0) \Rightarrow v = 0.$$

- (iv) Speziell für $K = \mathbb{R}$:

Eine Bilinearform auf einem \mathbb{R} -Vektorraum V heißt „**positiv definit**“, falls gilt:

$$\forall v \in V \setminus \{0\} : B(v, v) > 0.$$

- (v) Eine Bilinearform B auf einem \mathbb{R} -Vektorraum V heißt ein „**Skalarprodukt**“, falls B positiv definit und symmetrisch ist.

Oft geschrieben als $B(v, w) =: \langle v, w \rangle$

Bemerkung.

(i) $\forall v, w \in V : B(v, 0) = 0 = B(0, w)$

(ii) B ist nicht entartet $\Leftrightarrow \bigcap_{w \in V} \text{Kern}(B(\cdot, w)) = \{0\}$.

(iii) B positiv definit $\Rightarrow B$ nicht entartet

(iv) **Hauptbeispiel:**

Bezüglich der Standardbasis im K^n sei für $v = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ und $w = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ definiert:

$$B(v, w) := \sum_{i=1}^n x_i y_i.$$

Es gilt:

- B ist Bilinearform. ✓
- B ist symmetrisch. ✓
- B ist nichtentartet. ✓

$K = \mathbb{R} \Rightarrow B$ positiv definit und B ist Skalarprodukt. B heißt „**Standardskalarprodukt**“.

(v) „Partielle Abbildungen“ sind auch sonst eine wirksame Idee der Mathematik.

10.2 Gram¹-Matrizen

Sei $\{v_1, \dots, v_n\}$ eine Basis von V und $A \in K^{n \times n}$. Für $v, w \in V$ mit $v = \sum x_i v_i$ und $w = \sum y_j v_j$ wird durch

$$B(v, w) := \sum_{i=1}^n \sum_{j=1}^n x_i a_{ij} y_j$$

eine Bilinearform auf V definiert, denn

$$B(kv, w) = \sum \sum k x_i a_{ij} y_j = k \sum \sum x_i a_{ij} y_j = kB(v, w)$$

und

$$B(v + \tilde{v}, w) = \sum \sum (x_i + \tilde{x}_i) a_{ij} y_j = \dots = B(v, w) + B(\tilde{v}, w).$$

Analog für $B(v, kw) = kB(v, w)$ und $B(v, w + \tilde{w}) = B(v, w) + B(v, \tilde{w})$.

Mit Spaltenvektoren $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ und $y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ läßt sich summenformelfrei schreiben

$B(v, w) = x^T A y$. Man nennt B die von der Matrix A (bzgl. der Basis $\{v_1, \dots, v_n\}$) „**induzierte**“ Bilinearform. Hauptbeispiel hat $A = E_n$. Weitere Beispiele:

$$D = \begin{pmatrix} d_1 & & 0 \\ & \ddots & \\ 0 & & d_n \end{pmatrix}$$

¹J. P. GRAM, 1850–1916

10 Euklidische Vektorräume

induziert die Bilinearform $x^T D y = \sum x_i d_i y_i$ (Summe über die gewichteten gemischten Produkte).

Umgekehrt kann aus B die Koeffizientenmatrix A gewonnen werden mittels $B(v_i, v_j) = e_i^T A e_j = a_{ij}$. Mit anderen Worten: Bei fester Basis wird für eine gegebene Bilinearform B deren Koeffizientenmatrix gegeben durch

$$A = \begin{pmatrix} B(v_1, v_1) & B(v_1, v_2) & \dots & B(v_1, v_n) \\ \dots & \dots & \dots & \dots \\ B(v_n, v_1) & B(v_n, v_2) & \dots & B(v_n, v_n) \end{pmatrix}.$$

A heißt „**Gram-Matrix** von B “.

Saloppes Resümee: Es gibt so viele Bilinearformen wie quadratische Matrizen.

Satz. Die Gram-Matrix ist symmetrisch genau dann, wenn B symmetrisch ist.

Satz. Sei $\{v_1, \dots, v_n\}$ Basis von V , B Bilinearform auf V mit Gram-Matrix A . Dann gilt:

$$\text{Rang } A = n \Leftrightarrow B \text{ nichtentartet.}$$

Beweis. Behauptung ist dasselbe wie

$$\text{Rang } A < n \Leftrightarrow B \text{ entartet}$$

$$\Leftrightarrow \exists v \in V \forall w \in V : B(v, w) = 0 \text{ und } v \neq 0$$

„ \Rightarrow “

Sei $\text{Rang } A < n$. Also Zeilen von A linear abhängig, d. h.

$$\exists x_1, \dots, x_n \in K, \text{ nicht alle } x_i = 0 :$$

$$\sum_{i=1}^n x_i B(v_i, v_j) = 0 \quad \text{für alle } j = 1, \dots, n \quad \text{und} \quad v := \sum_{i=1}^n x_i v_i \neq 0.$$

Offenbar folgt wegen der Linearität im ersten Argument daraus $B(v, v_j) = 0$ für alle j . Mit der Linearität im zweiten Argument dann $\forall w \in V : B(v, w) = 0$.

„ \Leftarrow “

Sei v gegeben mit den angegebenen Eigenschaften. Entwickle v nach der gegebenen Basis zu $v = \sum x_i v_i$. Dann $0 = B(v, v_j) = \sum x_i B(v_i, v_j)$ für alle j . Daher hat A linear abhängige Zeilen und $\text{Rang } A < n$. □

27. 5. 2004 **Satz (Bilinearformen, Endomorphismen und quadratische Matrizen).** Sei B nichtentartete Bilinearform auf V . Dann gilt für alle Abbildungen $\tilde{B} : V \times V \rightarrow K$:

$$\tilde{B} \text{ bilinear} \Leftrightarrow \exists f \in \text{End}(V) \forall v, w \in V : \tilde{B}(v, w) = B(v, f(w)).$$

Beweis. Matrizenfrei siehe Beutelspacher. Hier mit Matrizen:

„ \Rightarrow “

Sei $\{v_1, \dots, v_n\}$ Basis von V und $A = \text{Gram } B$. Also, da B nichtentartet, $A \in \text{GL}(n)$. Sei $\tilde{A} = \text{Gram } \tilde{B} \in K^{n \times n}$. Seien $v, w \in V$ mit Koordinatendarstellung $v = \sum x_i v_i, w = \sum y_j v_j$. Dann: $\tilde{B}(v, w) = x^T \tilde{A} y = x^T A A^{-1} \tilde{A} y = B(v, f(w))$ mit $f(w) := \underbrace{A^{-1} \tilde{A}}_{\in K^{n \times n}} y \in \text{End}(V)$.

„ \Leftarrow “

Klar. □

10.3 Orthogonalität

Definition. Sei B Bilinearform auf V .

- (i) Zwei Vektoren $v, w \in V$ heißen „**orthogonal**“, falls $B(v, w) = 0$.
- (ii) $\forall v \in V : \{v\}^\perp := \{w \in V \mid B(v, w) = 0\}$ heißt „**Orthogonalraum zu v** “.
- (iii) $\forall M \subseteq V : M^\perp := \{w \in V \mid \forall v \in M : B(v, w) = 0\}$ heißt **Orthogonalraum zu M** “.

Bemerkung. M^\perp ist ein Untervektorraum.

Satz. Sei B eine Bilinearform auf V . Dann gilt:

- (i) $\forall M, N \subseteq V : M \subseteq N \Rightarrow M^\perp \supseteq N^\perp$
- (ii) $\forall M \subseteq V : M^\perp = (\text{span } M)^\perp$
- (iii) B symmetrisch $\Rightarrow \forall M \subseteq V : M \subseteq M^{\perp\perp}$ **und** $M^\perp = M^{\perp\perp\perp}$

Beweis.

- (i) ✓
- (ii) Setze $N := \text{span } M$. Aus (i) folgt \supseteq . Umgekehrt $v \in M^\perp \Rightarrow \forall w \in M : B(v, w) = 0 \Rightarrow \forall w \in N : B(v, w) = 0 \Rightarrow v \in N^\perp$.
- (iii) Sei $v \in M$. Zu zeigen $v \in (M^\perp)^\perp$, d. h. $\forall w \in M^\perp : 0 = B(w, v) \stackrel{(\text{sym.})}{=} B(v, w)$ ✓. Und mit (i) $M^\perp \supseteq M^{\perp\perp\perp}$, wie auch für $N := M^\perp$ folgt $N \subseteq N^{\perp\perp}$.

□

Satz (Dimensionsformel). Sei B eine nichtentartete Bilinearform auf V und $U \subseteq V$ ein Untervektorraum. Dann gilt:

- (i) $\dim U + \dim U^\perp = \dim V$
- (ii) B symmetrisch $\Rightarrow U = U^{\perp\perp}$

Beweis.

- (i) Sei $\{v_1, \dots, v_n\}$ Basis von V so, daß $\{v_1, \dots, v_m\}$ Basis des Untervektorraums U ist. Für $w = \sum_{i=1}^n x_i v_i \in U^\perp$ gilt:

$$\forall j \leq m : 0 = B(w, v_j) = \sum_i x_i B(v_i, v_j) = x^T A e_j \quad \text{mit } A := \text{Gram } B$$

Transposition führt zur Matrix

$$C := \begin{pmatrix} (Ae_1)^T \\ \vdots \\ (Ae_m)^T \end{pmatrix}$$

mit $Cx = 0$ und $\text{Rang } C = m$. Also $x \in \text{Kern } C$ und $\dim \text{Kern } C = n - m$ und $\dim U^\perp = n - m$.

10 Euklidische Vektorräume

- (ii) $\dim U^{\perp\perp} = n - \dim U \perp = n - (n - \dim U) = \dim U$. Nun folgt aus $U \subseteq U^{\perp\perp}$ die Gleichheit.

□

Definition.

- (i) $K_{\text{sym.}}^{n \times n} := \{A \in K^{n \times n} \mid A = A^T\}$
(ii) $\mathbb{R}_{\text{pos.def.}}^{n \times n} := \{A \in \mathbb{R}_{\text{sym.}}^{n \times n} \mid \forall x \in \mathbb{R}^n \setminus \{0\} : x^T A x > 0\}$

Bemerkung.

- (i) Für alle Bilinearformen B gilt:

$$B \text{ symmetrisch} \Leftrightarrow \text{Gram } B \in K_{\text{sym.}}^{n \times n}.$$

- (ii) Für alle Bilinearformen auf \mathbb{R} -Vektorräumen gilt:

$$B \text{ Skalarprodukt} \Leftrightarrow \text{Gram } B \in \mathbb{R}_{\text{pos.def.}}^{n \times n}.$$

Definition.

- (i) $(V, \langle \cdot, \cdot \rangle)$ heißt „**euklidischer Vektorraum**“, falls V endlichdimensionaler \mathbb{R} -Vektorraum und $\langle \cdot, \cdot \rangle$ Skalarprodukt auf V ist.
(ii) Sei $(V, \langle \cdot, \cdot \rangle)$ euklidischer Vektorraum und $U \subseteq V$ Untervektorraum. Dann heißt U^\perp „**das orthogonale Komplement von U** “.
(iii) Sei $(V, \langle \cdot, \cdot \rangle)$ euklidischer Vektorraum. Dann heißt $\|\cdot\| : v \mapsto \mathbb{R}, v \mapsto \|v\| := \sqrt{\langle v, v \rangle}$ die (von $\langle \cdot, \cdot \rangle$ induzierte) „**Norm**“ auf V .

Bemerkung. Die Norm eines Vektors v heißt auch „Länge von v “.

Satz (Ungleichung von CAUCHY²-BUNYAKOWSKI³-SCHWARZ⁴). Sei $(V, \langle \cdot, \cdot \rangle)$ ein euklidischer Vektorraum. Dann gilt für alle $v, w \in V$:

$$|\langle v, w \rangle| \leq \|v\| \cdot \|w\|,$$

d. h.

$$\langle v, w \rangle^2 \leq \langle v, v \rangle \langle w, w \rangle.$$

Beweis.

I. Falls $v = 0$ oder $w = 0$ ✓.

²1789–1875

³1804–1889

⁴1843–1921

II. Falls $v \neq 0$ und $w \neq 0$:

Studiere $f : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto f(x) := \langle xv - w, xv - w \rangle$. Offenbar $f(x) = x^2 \langle v, v \rangle - 2x \langle v, w \rangle + \langle w, w \rangle$. Minimum suchen durch Differenzieren:

$$0 = f'(x_0) = 2x_0 \langle v, v \rangle - 2 \langle v, w \rangle \Rightarrow x_0 = \frac{\langle v, w \rangle}{\langle v, v \rangle}$$

Also:

$$0 \leq f(x_0) = f\left(\frac{\langle v, w \rangle}{\langle v, v \rangle}\right) = \frac{\langle v, w \rangle^2}{\langle v, v \rangle} - 2 \frac{\langle v, w \rangle^2}{\langle v, v \rangle} + \langle w, w \rangle$$

Daraus: $\langle v, w \rangle^2 \leq \langle v, v \rangle \langle w, w \rangle$.

□

Zugabe. Gleichheit $\Leftrightarrow x_0 v = w \Leftrightarrow v, w$ linear abhängig

Satz.

(i) $\forall k \in \mathbb{R} \forall v \in V : \|kv\| = |k| \|v\|$ („positive Homogenität“)

(ii) $\forall v, w \in V : \|v + w\| \leq \|v\| + \|w\|$ („Dreiecksungleichung“)

(iii) $\forall v \in V : \|v\| = 0 \Leftrightarrow v = 0$

Beweis.

(i) aus Definition

(ii)

$$\begin{aligned} \|v + w\|^2 &= \langle v + w, v + w \rangle \\ &= \langle v, v \rangle + 2 \langle v, w \rangle + \langle w, w \rangle \\ &\stackrel{\text{CS}}{\leq} \|v\|^2 + 2 \|v\| \|w\| + \|w\|^2 \\ &= (\|v\| + \|w\|)^2 \end{aligned}$$

(iii) ✓

□

Alternativer Beweis der Cauchy-Schwarz-Ungleichung.

2. 6. 2004

$$\begin{aligned} \left(\sum_{i=1}^n a_i^2\right) \left(\sum_{i=1}^n b_i^2\right) - \left(\sum_{i=1}^n a_i b_i\right)^2 &= \sum_i \sum_j a_i^2 b_j^2 - \left(\sum_i a_i b_i\right) \left(\sum_j a_j b_j\right) \\ &= \frac{1}{2} \left(\sum_i \sum_j a_i^2 b_j^2 + \sum_i \sum_j a_j^2 b_i^2 - 2 \sum_i \sum_j a_i b_j a_j b_i \right) \\ &= \frac{1}{2} \sum_i \sum_j (a_i b_j - a_j b_i)^2 \geq 0 \end{aligned}$$

Gleichheit $\Leftrightarrow \forall i, j : a_i b_j = a_j b_i$, d. h. $\frac{a_i}{b_i} = \frac{a_j}{b_j} =: c$ und $a_i = c b_i$ für alle i .

□

10.3.1 Exkurs über Normen

Definition. Auf einem endlichdimensionalen \mathbb{R} -Vektorraum heißt $\|\cdot\| : V \rightarrow \mathbb{R}$ eine Norm, falls gilt:

- (i) $\forall k \in \mathbb{R} \forall v \in V : \|kv\| = |k|\|v\|$
- (ii) $\forall v, w \in V : \|v + w\| \leq \|v\| + \|w\|$
- (iii) $\forall v \in V : \|v\| = 0 \Leftrightarrow v = 0$

Beispiel. Die „ L_p -Normen“⁵ auf \mathbb{R}^n :

Sei $p \in [1, \infty)$ fest. Dann ist

$$\|x\|_p := \left(\sum_{i=1}^n |x_i|^p \right)^{1/p}$$

eine Norm. Speziell:

$p = 2$: $\|x\|_2 = (\sum x_i^2)^{1/2} = \sqrt{\langle x, x \rangle}$ ist die euklidische Norm.

$p = 1$: $\|x\|_1 = \sum |x_i|$ ist die „Betragnorm“.

„ $p = \infty$ “: $\|x\|_\infty := \max_{i=1}^n |x_i|$ ist die „Maximumnorm“ und paßt als letztes zur Familie der L_p -Normen wegen:

I. $\forall x \in \mathbb{R}^n \setminus \{0\} : |x_{i_0}| := \max_{i=1}^n |x_i| > 0$. Damit:

$$\begin{aligned} \lim_{p \rightarrow \infty} \|x\|_p &= \lim_{p \rightarrow \infty} \left(\sum_{i=1}^n |x_i|^p \frac{|x_{i_0}|^p}{|x_{i_0}|^p} \right)^{1/p} \\ &= |x_{i_0}| \lim_{p \rightarrow \infty} \exp \left(\underbrace{\frac{1}{p} \log \sum_{i=1}^n \underbrace{\frac{|x_i|^p}{|x_{i_0}|^p}}_{\in [0,1]}}_{\in [1,n]} \right) = \max_{i=1}^n |x_i| \end{aligned}$$

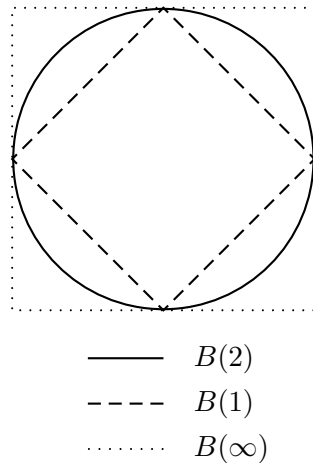
II. $x = 0$: ✓

Veranschaulichung der „**Einheitskreise**“ (siehe Abbildung 10.1 auf der nächsten Seite):

$$V = \mathbb{R}^2, B(p) := \{x \in \mathbb{R}^2 \mid \|x\|_p \leq 1\}$$

Speziell:

- $B(2) = \{x \mid \sqrt{x_1^2 + x_2^2} \leq 1\}$
- $B(1) = \{x \mid |x_1| + |x_2| \leq 1\}$
- $B(\infty) = \{x \mid \max\{|x_1|, |x_2|\} \leq 1\}$

Abbildung 10.1: Einheitskreise bzgl. verschiedener L_p -Normen

$B(1) \subseteq B(2) \subseteq B(\infty)$ und allgemein $1 \leq p \leq \tilde{p} \leq \infty \Rightarrow B(p) \subseteq B(\tilde{p})$ und $\|x\|_p \leq \|x\|_{\tilde{p}}$.

Satz. Sei V ein \mathbb{R} -Vektorraum und $\|\cdot\|$ eine Norm auf V . Dann gibt es ein Skalarprodukt $\langle \cdot, \cdot \rangle$, das die vorgegebene Norm $\|\cdot\|$ induziert genau dann, wenn die „Parallelogrammgleichung“ gilt:

$$\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2.$$

Beweis.

„ \Rightarrow “ ✓
 „ \Leftarrow “

$$\langle x, y \rangle := \frac{1}{4} (\|x + y\|^2 - \|x - y\|^2) \quad \checkmark$$

□

Im Beispiel: $\|\cdot\|_p$ wird genau für $p = 2$ von einem Skalarprodukt induziert.

Fazit: Es gibt mehr Normen als Skalarprodukte.

Norm \equiv „nur“ Längenmessung, Skalarprodukt \equiv „auch“ Winkelmessung, denn:

$$\forall x, y \neq 0 : -1 \leq \frac{\langle x, y \rangle}{\|x\| \|y\|} \leq 1$$

Definition. Die Zahl $\varphi \in [0, \pi)$ mit $\cos \varphi = \frac{\langle x, y \rangle}{\|x\| \|y\|}$ heißt „**Winkel zwischen x und y** “.

Umgekehrt: $\langle x, y \rangle = \|x\| \|y\| \cos \varphi$.

⁵nach Henri LEBESQUE (1875–1941)

10.3.2 Orthonormalbasen in euklidischen Vektorräumen

Sei $(V, \langle \cdot, \cdot \rangle)$ ein euklidischer Vektorraum.

Definition.

- (i) $\{v_1, \dots, v_n\} \subseteq V$ heißt „**orthogonal**“, falls gilt $\forall i \neq j : \langle v_i, v_j \rangle = 0$.
- (ii) $\{v_1, \dots, v_n\} \subseteq V$ heißt „**normiert**“, falls gilt $\forall i : \|v_i\| = 1$.
- (iii) $\{v_1, \dots, v_n\} \subseteq V$ heißt „**orthonormiert**“, falls $\{v_1, \dots, v_n\}$ orthogonal und normiert ist.
- (iv) $\{v_1, \dots, v_n\} \subseteq V$ heißt **Orthonormalbasis** („ONB“), falls $\{v_1, \dots, v_n\}$ Basis und orthonormiert ist.

Bemerkung.

- (i) In \mathbb{R}^n ist $\{e_1, \dots, e_n\}$ eine Orthonormalbasis.
- (ii) Ist $\{v_1, \dots, v_n\}$ eine Orthonormalbasis (bezgl. des gegebenen Skalarprodukts $\langle \cdot, \cdot \rangle$), dann ist die Gram-Matrix von $\langle \cdot, \cdot \rangle$ gegeben durch

$$\text{Gram}(\langle \cdot, \cdot \rangle) = (\langle v_i, v_j \rangle)_{i,j=1,\dots,n} = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix} = E_n.$$

- (iii) Ist $\{v_1, \dots, v_n\}$ Basis, dann ist $\forall i : v_i \neq 0$ und $\left\{ \frac{1}{\|v_1\|}v_1, \dots, \frac{1}{\|v_n\|}v_n \right\}$ normierte Basis.
- (iv) Ist $\{v_1, \dots, v_n\}$ orthogonal und $\forall i : v_i \neq 0$, dann ist $\{v_1, \dots, v_n\}$ linear unabhängig, denn:

$$\sum_{i=1}^n k_i v_i = 0 \Rightarrow \forall j : 0 = \left\langle \sum_{i=1}^n k_i v_i, v_j \right\rangle = \sum_{i=1}^n k_i \langle v_i, v_j \rangle = k_j \|v_j\|^2 \Rightarrow \forall j : k_j = 0$$

3. 6. 2004 (v) $\{b_1, \dots, b_n\}$ ist genau dann eine Orthonormalbasis, wenn $\forall i, j = 1, \dots, n : \langle b_i, b_j \rangle = \delta_{ij}$ gilt.

Satz (Orthonormalisierungsverfahren von GRAM und SCHMIDT). Sei $(V, \langle \cdot, \cdot \rangle)$ ein euklidischer Vektorraum, $n := \dim V$ und $\{v_1, \dots, v_n\}$ Basis von V . Dann gilt:

- (i) Setzt man

$$b_1 := v_1$$

$$b_i := v_i - \sum_{j=1}^{i-1} \frac{\langle b_j, v_i \rangle}{\langle b_j, b_j \rangle} b_j \quad \text{für } i = 2, \dots, n,$$

so ist $\{b_1, \dots, b_n\}$ eine Orthogonalbasis von V und es gilt

$$\text{span}\{b_1, \dots, b_k\} = \text{span}\{v_1, \dots, v_k\}$$

für alle $k = 1, \dots, n$.

(ii) Setzt man

$$c_1 := \frac{1}{\|v_1\|} v_1$$

$$c_i := \frac{1}{\|\tilde{c}_i\|} \left(v_i - \underbrace{\sum_{j=1}^{i-1} \langle c_j, v_i \rangle c_j}_{=: \tilde{c}_i} \right) \quad \text{für } i = 2, \dots, n,$$

so ist $\{c_1, \dots, c_n\}$ eine Orthonormalbasis von V und es gilt

$$\text{span}\{c_1, \dots, c_k\} = \text{span}\{v_1, \dots, v_k\}$$

für alle $k = 1, \dots, n$.

Beweis.

(i) Zeige $\langle b_i, b_j \rangle \stackrel{!}{=} 0$ für alle $i, j = 1, \dots, n, i \neq j$ durch Induktion:
Induktionsanfang. $\langle b_1, b_2 \rangle = 0$, denn:

$$\begin{aligned} \langle b_1, b_2 \rangle &= \left\langle v_1, v_2 - \frac{\langle b_1, v_2 \rangle}{\langle b_1, b_1 \rangle} b_1 \right\rangle \\ &= \left\langle v_1, v_2 - \frac{\langle v_1, v_2 \rangle}{\langle v_1, v_1 \rangle} v_1 \right\rangle \\ &= \langle v_1, v_2 \rangle - \frac{\langle v_1, v_2 \rangle}{\langle v_1, v_1 \rangle} \langle v_1, v_1 \rangle = 0. \end{aligned}$$

Induktionsschritt. Sei $k \in \{2, \dots, n-1\}$ und sei $\langle b_i, b_j \rangle = 0$ für alle $i, j = 1, \dots, k$ mit $i \neq j$ (Ind.ann.):

Dann gilt für alle $i = 1, \dots, k$:

$$\begin{aligned} \langle b_i, b_{k+1} \rangle &= \left\langle b_i, v_{k+1} - \sum_{j=1}^k \frac{\langle b_j, v_{k+1} \rangle}{\langle b_j, b_j \rangle} b_j \right\rangle \\ &= \langle b_i, v_{k+1} \rangle - \sum_{j=1}^k \frac{\langle b_j, v_{k+1} \rangle}{\langle b_j, b_j \rangle} \underbrace{\langle b_i, b_j \rangle}_{\substack{=0, \\ \text{falls } i \neq j \\ \text{(Ind.ann.)}}} \\ &= \langle b_i, v_{k+1} \rangle - \frac{\langle b_i, v_{k+1} \rangle}{\langle b_i, b_i \rangle} \langle b_i, b_i \rangle = 0 \end{aligned}$$

Zeige durch Induktion:

$$\forall k = 1, \dots, n : \text{span}\{b_1, \dots, b_k\} \stackrel{!}{=} \text{span}\{v_1, \dots, v_k\}.$$

Induktionsanfang. $k = 1. b_1 = v_1. \checkmark$

10 Euklidische Vektorräume

Induktionsschritt. Sei $k \geq 1$ und sei $\text{span}\{b_1, \dots, b_k\} = \text{span}\{v_1, \dots, v_k\}$ (Ind.ann).

Dann:

$$b_{k+1} = v_{k+1} - \underbrace{\sum_{j=1}^k \frac{\langle b_j, v_{k+1} \rangle}{\langle b_j, b_j \rangle} b_j}_{\substack{\in \text{span}\{b_1, \dots, b_k\} \\ = \text{span}\{v_1, \dots, v_k\} \\ \text{(Ind.ann.)}}} \in \text{span}\{v_1, \dots, v_{k+1}\}$$

Zusammen mit Induktionsannahme: $\text{span}\{b_1, \dots, b_{k+1}\} \subseteq \text{span}\{v_1, \dots, v_{k+1}\}$.

Umgekehrt:

$$v_{k+1} = b_{k+1} + \underbrace{\sum_{j=1}^k \frac{\langle b_j, v_{k+1} \rangle}{\langle b_j, b_j \rangle} b_j}_{\in \text{span}\{b_1, \dots, b_k\}} \in \text{span}\{b_1, \dots, b_{k+1}\}$$

Mit der Induktionsannahme folgt $\text{span}\{v_1, \dots, v_{k+1}\} \subseteq \text{span}\{b_1, \dots, b_{k+1}\}$. Insgesamt folgt Gleichheit. Basiseigenschaft folgt aus dem Fall $k = n$.

- (ii) Analog. Unterschied zu (i) ist, daß die Vektoren c_1, \dots, c_n gleich *normiert* werden (d. h. $c_i = 1/\|\tilde{c}_i\| \cdot \tilde{c}_i$); wegen $\langle c_j, c_j \rangle = 1$ fällt in den Summanden jeweils der Nenner weg.

□

Bemerkung.

- a) Im Beweiss von (i) werden nur Symmetrie und die Nichtausgeartetheit von $\langle \cdot, \cdot \rangle$ verwendet. Die Aussage von (i) bleibt auch wahr, wenn K ein beliebiger Körper und V endlichdimensionaler K -Vektorraum ist.
- b) Geometrische Anschauung zur Konstruktion der b_i :

$$b_2 = v_2 - \frac{\langle b_1, v_2 \rangle}{\langle b_1, b_1 \rangle} b_1 = v_2 - \underbrace{\frac{\|v_2\|}{\|b_1\|} \frac{\langle b_1, v_2 \rangle}{\|v_2\| \|b_1\|}}_{=\cos \varphi(b_1, v_2)} b_1$$

„Anteil von v_2 längs b_1 “

Allgemein: b_i geht aus v_i hervor, indem aus v_i die Anteile längs b_1, \dots, b_{i-1} subtrahiert werden.

- c) (i) und (ii) liefern praktische Verfahren zur Konstruktion von Orthogonal-/Orthonormalbasen von V . Das Ergebnis hängt i. a. von der Reihenfolge der v_1, \dots, v_n ab.
- d) Ist $\{v_1, \dots, v_n\}$ bereits orthogonal (bzw. orthonormal), so gilt $b_i = v_i$ (bzw. $b_i = c_i = v_i$) für alle $i = 1, \dots, n$.

Jeder euklidische Vektorraum besitzt folglich eine Orthonormalbasis.

Korollar 1. Sei $(V, \langle \cdot, \cdot \rangle)$ ein euklidischer Vektorraum und $\{v_1, \dots, v_k\} \subseteq V$ orthonormal. Dann kann $\{v_1, \dots, v_k\}$ zu einer Orthonormalbasis ergänzt werden.

Beweis. Da $\{v_1, \dots, v_k\}$ orthonormal ist, ist $\{v_1, \dots, v_k\}$ linear unabhängig. Mit $n := \dim V$ gibt es dann Vektoren $v_{k+1}, \dots, v_n \in V$ so, daß $\{v_1, \dots, v_n\}$ Basis von V ist. Wende dann GRAM/SCHMIDT auf $\{v_1, \dots, v_n\}$ in dieser Reihenfolge an. Dies liefert eine Orthonormalbasis $\{c_1, \dots, c_n\}$ von V . Weil $\{v_1, \dots, v_k\}$ bereits orthonormal ist, gilt $c_i = v_i$ für $i = 1, \dots, k$. \square

10.3.3 Orthogonale Abbildungen (Isometrien)

Definition. Sei $(V, \langle \cdot, \cdot \rangle)$ ein euklidischer Vektorraum und $f \in \text{End}_{\mathbb{R}}(V)$. Dann heißt f „**orthogonale Abbildung**“ oder auch „**Isometrie**“, falls gilt:

$$\forall v, w \in V : \langle f(v), f(w) \rangle = \langle v, w \rangle.$$

In anderen Worten: $\langle \cdot, \cdot \rangle$ bleibt invariant unter f .

Die Menge aller orthogonalen Abbildungen auf V nennen wir $O(V)$.

Bemerkung.

- (i) Für $\text{id}_V \in \text{End}_{\mathbb{R}}(V)$ gilt trivialerweise $\langle \text{id}_V(v), \text{id}_V(w) \rangle = \langle v, w \rangle$ für alle $v, w \in V$.
 $\text{id}_V \in O(V)$, also $O(V) \neq \emptyset$.
- (ii) Nichttriviales Beispiel: Für $\varphi \in [0, 2\pi)$ definiere die Matrix

$$D_\varphi =: \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix} \in \mathbb{R}^{2 \times 2}$$

und den Endomorphismus $f_\varphi : \mathbb{R}^2 \rightarrow \mathbb{R}^2, v \mapsto D_\varphi v$. Dann gilt:

$$\begin{aligned} f_\varphi(e_1) &= D_\varphi \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \cos \varphi \\ \sin \varphi \end{pmatrix} \\ f_\varphi(e_2) &= D_\varphi \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -\sin \varphi \\ \cos \varphi \end{pmatrix}. \end{aligned}$$

Dann gilt

$$\langle f_\varphi(e_i), f_\varphi(e_i) \rangle = (\sin \varphi)^2 + (\cos \varphi)^2 = 1 = \langle e_i, e_i \rangle$$

für $i = 1, 2$ und

$$\langle f_\varphi(e_1), f_\varphi(e_2) \rangle = (\cos \varphi)(\sin \varphi) - (\sin \varphi)(\cos \varphi) = 0 = \langle e_1, e_2 \rangle.$$

Wegen der Bilinearität des Skalarprodukts folgt $\forall v, w \in \mathbb{R}^2 : \langle f_\varphi(v), f_\varphi(w) \rangle = \langle v, w \rangle$. Also $f_\varphi \in O(\mathbb{R}^2)$. Geometrisch: Drehung um den Ursprung mit Winkel φ .

Satz. Sei $(V, \langle \cdot, \cdot \rangle)$ ein euklidischer Vektorraum. Dann gilt:

- (i) $O(V)$ ist eine Untergruppe von $\text{GL}_{\mathbb{R}}(V)$.

10 Euklidische Vektorräume

(ii) $\forall f \in O(V) : \text{Spec } f \subseteq \{\pm 1\}$.

Beweis.

(i) Zunächst gilt $O(V) \subseteq GL_{\mathbb{R}}(V)$, denn: Sei $\{b_1, \dots, b_n\}$ eine ONB von V , $n := \dim V$ und $f \in O(V)$. Dann gilt für alle $i, j = 1, \dots, n$:

$$\langle f(b_i), f(b_j) \rangle = \langle b_i, b_j \rangle = \delta_{ij},$$

d. h. $\{f(b_1), \dots, f(b_n)\}$ ist ebenfalls ONB von V . Insbesondere gilt $\text{Bild } f = V$, d. h. $f \in GL_{\mathbb{R}}(V)$.

Wir wissen $\text{id}_V \in O(V)$. Seien außerdem $f, g \in O(V)$. Dann gilt für alle $v, w \in V$:

$$\langle (f \circ g)(v), (f \circ g)(w) \rangle = \langle g(v), g(w) \rangle = \langle v, w \rangle,$$

d. h. $f \circ g \in O(V)$. Schließlich für $v, w \in V$ und $f \in O(V)$:

$$\langle f^{-1}(v), f^{-1}(w) \rangle = \langle f(f^{-1}(v)), f(f^{-1}(w)) \rangle = \langle v, w \rangle,$$

d. h. $f^{-1} \in O(V)$.

Dies zeigt, daß $(O(V), \circ)$ eine Untergruppe von $GL_{\mathbb{R}}(V)$ ist.

(ii) Sei $f \in O(V)$, $x \in \text{Spec } f$ und $v \in E_x(f) \setminus \{0\}$. Dann:

$$\langle v, v \rangle = \langle f(v), f(v) \rangle = \langle xv, xv \rangle = x^2 \langle v, v \rangle$$

Mit $\langle v, v \rangle \neq 0$ folgt $x^2 = 1$, also $x \in \{\pm 1\}$. □

9. 6. 2004 **Satz (Darstellungsmatrizen orthogonaler Abbildungen).** Sei $(V, \langle \cdot, \cdot \rangle)$ ein euklidischer Vektorraum, $\{b_1, \dots, b_n\}$ eine Orthonormalbasis von V und $f \in \text{End}(V)$ mit $A := M(f)$. Dann gilt:

$$f \in O(V) \Leftrightarrow A^{-1} = A^T \Leftrightarrow A^T A = E_n.$$

Beweis. „ \Rightarrow “

f orthogonal; $f(b_j) = \sum_i a_{ij} b_i$. Auch $\{f(b_1), \dots, f(b_n)\}$ ist eine Orthonormalbasis, denn:

$$\begin{aligned} \delta_{ij} &= \langle f(b_i), f(b_j) \rangle = \left\langle \sum_k a_{ki} b_k, \sum_l a_{lj} b_l \right\rangle = \sum_k \sum_l a_{ki} a_{lj} \underbrace{\langle b_k, b_l \rangle}_{\delta_{kl}} = \sum_k a_{ki} a_{kj} \\ &= (A^T A)_{ij}, \end{aligned}$$

d. h. $A^T A = E_n$.

„ \Leftarrow “

Aus $A^T A = E_n$ folgt $\delta_{ij} = \sum_k a_{ki} a_{kj} = \dots = \langle f(b_i), f(b_j) \rangle$. Wegen der Bilinearität dann auch

$$\begin{aligned} \forall v, w \in V : \langle f(v), f(w) \rangle &= \sum_i \sum_j \langle f(x_i b_i), f(y_j b_j) \rangle = \sum_i \sum_j x_i y_j \delta_{ij} \\ &= \sum_i \sum_j x_i y_j \langle b_i, b_j \rangle = \left\langle \sum_i x_i b_i, \sum_j y_j b_j \right\rangle = \langle v, w \rangle. \end{aligned}$$

□

Definition. Die Menge $O_{\mathbb{R}}(n) := \{A \in GL_{\mathbb{R}}(n) \mid A^T A = E_n\}$ heißt „**orthogonale Gruppe**“.

Bemerkung.

a) Ist B eine Basis von V , so ist die Abbildung $M : GL(V) \rightarrow GL(n)$, die also jedem Vektorraum-Automorphismus die Darstellungsmatrix bzgl. der Basis B zuordnet, ein Gruppenisomorphismus. Ist B Orthonormalbasis von V , so gilt nach obigem Satz $O_{\mathbb{R}}(n) = M(O(V))$. Laut Übung 19 ist also $O(n)$ eine Untergruppe von $GL(n)$.

b) Man kann natürlich die Gruppenstruktur von $O(n)$ auch direkt nachrechnen.

c) $\forall A \in O(n) : \det A \in \{\pm 1\}$, denn $(\det A)^2 = \det(A^T) \det A = \det(A^T A) = \det E_n = 1$.

d) Überblick zu den vier (klassischen) Gruppen reeller Matrizen:

- $GL_{\mathbb{R}}(n) = \{A \in \mathbb{R}^{n \times n} \mid \det A \neq 0\}$ offene, unbeschränkte Teilmenge von $\mathbb{R}^{n \times n}$.
- $SL_{\mathbb{R}}(n) = \{A \in \mathbb{R}^{n \times n} \mid \det A = 1\}$ abgeschlossene, unbeschränkte Teilmenge von $\mathbb{R}^{n \times n}$.
- $O_{\mathbb{R}}(n) = \{A \in \mathbb{R}^{n \times n} \mid A^T A = E_n\}$ abgeschlossene, beschränkte (also kompakte) Teilmenge von $\mathbb{R}^{n \times n}$.
- $SO_{\mathbb{R}}(n) := SL_{\mathbb{R}}(n) \cap O_{\mathbb{R}}(n) = \{A \in O_{\mathbb{R}}(n) \mid \det A = 1\}$ kompakte, zusammenhängende Teilmenge von $\mathbb{R}^{n \times n}$.

Satz (Existenz und Eindeutigkeit von adjungierten Endomorphismen). Sei $(V, \langle \cdot, \cdot \rangle)$ ein euklidischer Vektorraum. Dann gilt:

$$\forall f \in \text{End}(V) \exists! f^T \in \text{End}(V) \forall v, w \in V : \langle f(v), w \rangle = \langle v, f^T(w) \rangle.$$

Beweis. Die Abbildung $\tilde{B}(v, w) := \langle f(v), w \rangle$ ist bilinear. Mit Satz über Bilinearformen, Endomorphismen und quadratische Matrizen auf Seite 40 folgt die Existenz eines Endomorphismus f^T mit $\tilde{B}(v, w) = \langle v, f^T(w) \rangle$. Dort vergessen, die Eindeutigkeit zu beweisen. Deshalb nachhaken:

Für jeden anderen Endomorphismus g mit $\tilde{B}(v, w) = \langle v, g(w) \rangle$ gilt bei festem $w \in V$:

$$\begin{aligned} \forall v \in V : \langle v, g(w) - f^T(w) \rangle &= \tilde{B}(v, w) - \tilde{B}(v, w) = 0 && \Rightarrow g(w) - f^T(w) \perp V \\ &&& \Rightarrow g(w) = f^T(w). \end{aligned}$$

□

Zusatz. Sei $\{b_1, \dots, b_n\}$ Orthonormalbasis von V . Dann gilt $\text{Gram}(\langle \cdot, \cdot \rangle) = E_n$. Wenn f die Darstellungsmatrix $A = M(f)$ hat, dann hat f^T die Darstellungsmatrix $M(f^T) = A^T$.

Beweis. Mit $v = \sum x_i b_i$ und $w = \sum y_j b_j$ folgt:

$$\begin{aligned} \langle f(v), w \rangle &= (Ax)^T E_n y = x^T A^T E_n y = x^T E_n A^T y \\ &= \langle v, \tilde{f}(w) \rangle \quad \text{mit } \tilde{f} \text{ definiert durch } A^T. \end{aligned}$$

Wegen Eindeutigkeit folgt $\tilde{f} = f^T$ und $M(f^T) = A^T$.

□

10 Euklidische Vektorräume

Definition. Sei $f \in \text{End}(V)$.

- (i) f^T aus dem Satz heißt der „**adjungierte Endomorphismus**“ von f (bzgl. $\langle \cdot, \cdot \rangle$).
- (ii) f heißt „**selbstadjungiert**“, falls $f^T = f$.

Bemerkung. Bezüglich einer beliebigen ONB gilt:

- (i) $M(f^T) = (M(f))^T$
- (ii) f selbstadjungiert $\Leftrightarrow M(f)$ symmetrisch
- (iii) f orthogonal $\Leftrightarrow \langle v, w \rangle = \langle f(v), f(w) \rangle = \langle v, f^T(f(w)) \rangle \Leftrightarrow f \circ f^T = \text{id}_V \Leftrightarrow f^T \circ f = \text{id}_V \Leftrightarrow f^{-1} = f^T$.

Satz (Hauptsatz über Normaldarstellungsmatrizen orthogonaler Abbildungen).

Sei $(V, \langle \cdot, \cdot \rangle)$ ein euklidischer Vektorraum und $f \in O(V)$. Dann existiert eine ONB $\{b_1, \dots, b_n\}$, bzgl. der die Darstellungsmatrix von f Blockdiagonalgestalt hat:

$$M(f) = \text{Blockdiag}(E_{n_1}, -E_{n_2}, D_{\varphi_1}, \dots, D_{\varphi_t})$$

mit 2×2 -Blöcken $D_\varphi = \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix}$.

Beweis. Siehe Beutelspacher S. 268 f. □

Nachtrag. Transposition funktioniert bei allen Matrizen in $K^{n \times m}$. Warum hier speziell *reelle* und quadratische Matrizen $\mathbb{R}^{n \times n}$? Auf zutreffende Allgemeinheit verzichtet zugunsten eines stromlinienförmigen Stoffaufbaus.

Volle Allgemeinheit wäre das folgende:

- a) $B : V \times W \rightarrow K$ „bilinear“ macht Sinn für **zwei** K -Vektorräume V und W . Bezüglich zweier Basen dann $\text{Gram } B = (B(v_i, w_j))_{\substack{i=1, \dots, n \\ j=1, \dots, m}} \in K^{n \times m}$.

Es ist zu unterscheiden:

B „nicht entartet **in der ersten Variable**“ : $\Leftrightarrow \bigcap_{w \in W} \text{Kern } B(\cdot, w) = \{0\}$

Analog für zweite Variable.

Dazu ist äquivalent: $\text{Rang Gram } B = n$, bzw. $\text{Rang Gram } B = m$.

- b) Bei $\dim V = \dim W$ braucht die Entartung nicht nach erster und zweiter Variable unterschieden werden.
- c) $V = W$ gibt Sinn für „symmetrische Bilinearformen“ $B(v, w) = B(w, v)$.
- d) $K = \mathbb{R}$ gibt Sinn für „positiv definite Bilinearformen“.
- e) Wichtig $K = \mathbb{C}$: Dann $\langle z, w \rangle := \sum_{j=1}^n z_j \overline{w_j}$ für $z, w \in \mathbb{C}^n$ ist leider nicht bilinear, denn

$$\forall \alpha \in \mathbb{C} : \langle z, \alpha w \rangle = \overline{\alpha} \langle z, w \rangle \quad \text{„sesqui-linear“.}$$

16. 6. 2004

Satz (Darstellungssatz für symmetrische Matrizen, Hauptachsentransformation, Spektralsatz).

Sei $(V, \langle \cdot, \cdot \rangle)$ euklidischer Vektorraum. Dann gilt:

$$\forall f \in \text{End}(V) : f \text{ selbstadjungiert} \Leftrightarrow f \text{ diagonalisierbar (siehe Kapitel 8.11).}$$

Konkret in Matrixsprache:

$$\forall A \in \mathbb{R}_{\text{sym.}}^{n \times n} \exists S \in O(n) : S^T A S \text{ diagonal.}$$

Bemerkung.

- a) Hauptproblem ist, Existenz von Eigenwerten und Eigenvektoren zu sichern. Wenn so etwas existiert, dann folgt notwendig, daß Eigenvektoren zu verschiedenen Eigenwerten orthogonal sind: Sei $f(v) = \lambda v$ und $f(w) = \mu w$, dann $0 = \langle f(v), w \rangle - \langle v, f(w) \rangle = (\lambda - \mu) \langle v, w \rangle$.
 $\lambda \neq \mu \Rightarrow \langle v, w \rangle = 0$.
- b) S orthogonal $\Rightarrow S^T A S = S^{-1} A S$, d. h. S ist Matrix eines Basiswechsels (siehe Kapitel 8.1 auf Seite 8). Hier interessieren besonders Orthonormalbasen, dafür sind die zugehörigen Konvertierungsmatrizen S orthogonal. In diesen Fällen $\tilde{M}(f) = S^T M(f) S$.
- c) Damit äquivalente Formulierung des Satzes:

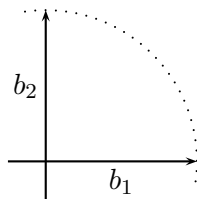
$$\forall f \in \text{End}(V) \text{ selbstadjungiert} \exists \{\tilde{b}_1, \dots, \tilde{b}_n\} \text{ ONB} \exists \lambda_1, \dots, \lambda_n :$$

$$f(\tilde{b}_j) = \lambda_j \tilde{b}_j \text{ für } j = 1, \dots, n$$

$$\text{mit } \tilde{M}(f) = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

Diese Basisvektoren heißen auch Hauptachsen⁶ von f .

- d) Anschaulich: Sei $f(\tilde{b}_1) = 3\tilde{b}_1$, $f(\tilde{b}_2) = -2\tilde{b}_2$. Siehe Abbildungen 10.2 bis 10.3 auf Seiten 53–54.



Wirkung von f unklar.

Abbildung 10.2: Hauptachsentransformation I

⁶engl. *principle axes* oder *principle components*

Gemäß Satz gibt es gedrehtes Orthonormalsystem:

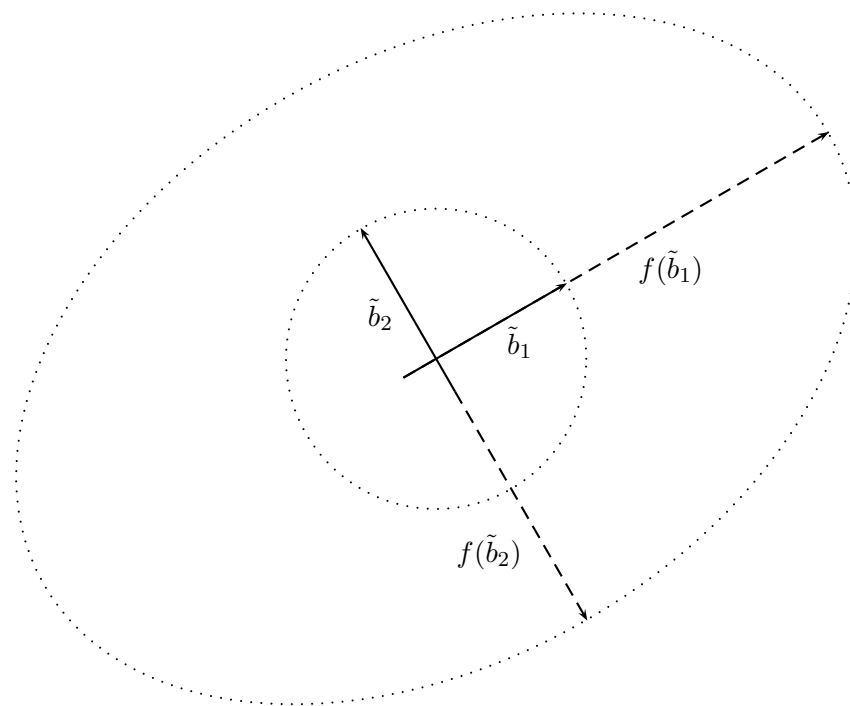


Abbildung 10.3: Hauptachsentransformation II

Beweis. I. Wenn f einen Eigenwert λ_1 mit Eigenvektor v_1 hat, dann setze $\tilde{b}_1 := \frac{1}{\|v_1\|}v_1$ und mache weiter mit $V_1 := \{v_1\}^\perp = \{\tilde{b}_1\}^\perp$.

Es gilt: $\dim V_1 = \dim V - 1$. V_1 mit $\langle \cdot, \cdot \rangle|_{V_1}$ ist wieder euklidischer Vektorraum. $f_1 := f|_{V_1}$ ist selbstadjungierter Endomorphismus auf V_1 . Und so weiter. . .

II. Also Knackpunkt ist die Aussage: Jeder selbstadjungierte Endomorphismus besitzt (mindestens) einen Eigenwert, d. h. $\text{Spec } f \neq \emptyset$.

Wir skizzieren drei Beweise:

II A. Komplexer Beweis:

Also Einbettung des \mathbb{R} -Vektorraums V in einen (größeren) \mathbb{C} -Vektorraum V^* . Ebenso Fortsetzung des (reellen) inneren Produkts $\langle \cdot, \cdot \rangle$ zu einer HERMITE-Form auf V^* . Existenz eines komplexen Eigenwerts $\lambda^* \in \mathbb{C}$ von f^* auf V^* geschenkt. Dann nachweisen, daß λ^* tatsächlich reell ist und daß $\lambda \equiv \lambda^*$ ein Eigenwert von f auf V ist. Siehe Beutelspacher S. 276/8.

II B. Analytischer Beweis:

Maximiere $x^T Ax$ unter $\|x\| = 1$ für $x \in \mathbb{R}^n$.

$\|x\| = 1 \Leftrightarrow \|x\|^2 = 1 \Leftrightarrow \langle x, x \rangle = 1$ und bedeutet, daß x auf der Einheitskugel in \mathbb{R}^n liegt. Existenz des Maximums folgt aus Stetigkeit und Kompaktheit der Einheitskugel.

Lösung mit Lagrangeschem Multiplikator λ :

$$\text{grad}(x^T Ax) = Ax + A^T x \stackrel{A^T = A}{=} 2Ax$$

$$\text{grad}(x^T x) = 2x$$

Charakterisierung mittels Lagrange-Multiplikator:

x_0 (lokales) Extremum

$$\Leftrightarrow \exists \lambda_0 \in \mathbb{R} : \lambda_0 \text{grad}(x^T Ax) = \text{grad}(x^T x)$$

$$\Leftrightarrow \lambda_0 \in \text{Spec } A \text{ und } x_0 \in E_{\lambda_0}(A) \setminus \{0\}$$

II C. Numerischer Beweis:

Siehe Übungsblatt 9.

□

Satz (Trägheitssatz von Sylvester⁷). $\forall A \in \mathbb{R}_{sym}^{n \times n} \exists ! p, q \in \mathbb{N} \exists S \in \text{GL}(n) :$

17. 6. 2004

$$S^T AS = \begin{pmatrix} E_p & & 0 \\ & -E_q & \\ 0 & & 0 \end{pmatrix}$$

⁷1814–1897

10 Euklidische Vektorräume

Beweis.

I. Sei $Q \in O(n)$ so, daß $Q^T A Q = \text{Diag}(\lambda_1, \dots, \lambda_n)$. Sei $p := \#\{i | \lambda_i > 0\}$ und $p := \#\{j | \lambda_j < 0\}$. Setze Diagonalmatrix

$$D := \begin{pmatrix} \frac{1}{\sqrt{|\lambda_1|}} & & 0 \\ & \ddots & \\ 0 & & \frac{1}{\sqrt{|\lambda_n|}} \end{pmatrix},$$

(außer für $\lambda_i = 0$: dafür setze Diagonaleintrag auf 1).

Für $S_i := QD$ gilt $S^T A S = DQ^T A Q D =$ wie gewünscht.

II. Eindeutigkeit siehe Jänich S. 239.

□

Unterscheide **Ähnlichkeitstransformationen** $A \mapsto S^{-1} A S$, gedeutet als Basiskonvertierung bei Endomorphismen, und **Kongruenztransformationen** $A \mapsto S^T A S$, gedeutet als Basiswechsel bei Bilinearformen!

11 Projektionen und verallgemeinerte Matrixinverse

23. 6. 2004

Voraussetzung: V endlichdimensionaler \mathbb{R} -Vektorraum.

11.1 Projektionen

Sei V direkte Summe zweier Untervektorräume U_1 und U_0 , d. h. $V = \text{span}(U_1 \cup U_0)$ und $U_1 \cap U_0 = \{0\}$.

Satz. $\forall v \in V \exists! u_1 \in U_1, u_0 \in U_0 : v = u_1 + u_0$

Beweis. Existenz aus span-Bedingung und Eindeutigkeit aus \cap -Bedingung. □

Die Abbildung $P : V \rightarrow U_1$ mit $P(v) = u_1$ heißt „**Projektion auf U_1 entlang U_0** “.

Satz.

(i) $P \in \text{End}(V)$, d. h. P lineare Abbildung von V in sich selbst

(ii) $\text{Bild}(P) = U_1$, $\text{Kern}(P) = U_0$

(iii) $P \circ P = P$ „Idempotenz“

Beweis.

(i) Skalarmultiplikation: $v = u_1 + u_0 \Rightarrow \lambda v = \underbrace{\lambda u_1}_{\in U_1} + \underbrace{\lambda u_0}_{\in U_0} \Rightarrow P(\lambda v) = \lambda u_1 = \lambda P(v)$.

Vektoraddition: $v = u_1 + u_0$ und $\tilde{v} = \tilde{u}_1 + \tilde{u}_0 \Rightarrow v + \tilde{v} = \underbrace{(u_1 + \tilde{u}_1)}_{\in U_1} + \underbrace{(u_0 + \tilde{u}_0)}_{\in U_0}$

$\Rightarrow P(v + \tilde{v}) = u_1 + \tilde{u}_1 = P(v) + P(\tilde{v})$

(ii) $\text{Bild}(P) = \{P(v) | v \in V\} \subseteq U_1$. Umgekehrt $\forall u_1 \in U_1 : u_1 = P(u_1) \in \text{Bild}(P)$.

$\text{Kern}(P) = \{v \in V | P(v) = 0\} \supseteq U_0$. Umgekehrt $\forall v \in V : P(v) = 0 \Rightarrow v = \underbrace{u_1}_{=0} + u_0 \in U_0$.

(iii) $\forall v \in V : P(v) = u_1 \in U_1 \Rightarrow P(P(v)) = u_1 = P(v) \Rightarrow P \circ P = P$. □

Definition. (i) Ein Endomorphismus heißt „**Projektion**“, falls er idempotent ist.

11 Projektionen und verallgemeinerte Matrixinverse

(ii) Eine quadratische Matrix heißt „**Projektion**“, falls sie idempotent ist.

$$\mathbb{R}_{\text{proj}}^{n \times n} := \{A \in \mathbb{R}^{n \times n} \mid A^2 = A\}.$$

A ist Projektionsmatrix auf Bild A entlang $\ker A$.

Satz (Darstellungssatz für Projektionen). *Projektionen haben nur Eigenwerte 1 und 0, sind diagonalisierbar und erfüllen $\text{Rang} = \text{Spur}$.*

Beweis. Offenbar $\forall \lambda \in \mathbb{R} \forall x \in V \setminus \{0\} : \lambda x = P(x) = P^2 x = P(Px) = P(\lambda x) = \lambda P(x) = \lambda \lambda x = \lambda^2 x \Rightarrow \lambda^2 - \lambda = 0 \Rightarrow \lambda = 1$ oder $\lambda = 0$.

Offenbar $E_1(P) = U_1$ und $E_0(P) = U_0$. Aus $V = U_1 \oplus U_0 = E_1(P) \oplus E_0(P)$ folgt Diagonalisierbarkeit.

$$\begin{aligned} \text{Spur } P &= \text{Spur}(\text{Darstellungsmatrix von } P) = \text{Spur} \begin{pmatrix} 1 & & & & 0 \\ & \ddots & & & \\ & & 1 & & \\ & & & 0 & \\ 0 & & & & \ddots \\ & & & & & 0 \end{pmatrix} \\ &= \dim E_1(P) = \dim U_1 = \dim \text{Bild}(P) = \text{Rang } P. \end{aligned}$$

□

11.2 Orthogonale Projektionen

Sei $(V, \langle \cdot, \cdot \rangle)$ euklidischer Vektorraum.

Satz. Sei $A \in \mathbb{R}^{n \times k}$ mit den Spalten $A = (s_1 \mid \dots \mid s_k)$. Dann gilt:

(i) $\text{Kern } A = \text{Kern } A^T A$

$A^T A$ heißt *Gram-Matrix* zur Matrix A und AA^T heißt auch *Gram-Matrix* zur Matrix A .

(ii) $\text{Kern } A = (\text{Bild } A^T)^\perp$

(iii) $\text{Bild } A = \text{Bild } AA^T$

Beweis.

(i) $\text{Kern } A \subseteq \text{Kern } A^T A$. Umgekehrt: $x \in \text{Kern } A^T A \Rightarrow A^T A x = 0 \Rightarrow \underbrace{x^T A^T A x}_{=\|Ax\|^2} = 0 \Rightarrow$

$$Ax = 0 \Rightarrow x \in \text{Kern } A.$$

(ii) $x \in \text{Kern } A \Leftrightarrow Ax = 0 \Leftrightarrow \forall y \in \mathbb{R}^n : 0 = y^T Ax = (A^T y)^T x \Leftrightarrow x \perp \text{Bild } A^T \Leftrightarrow x \in (\text{Bild } A^T)^\perp$

(iii) $\text{Bild } A \stackrel{(ii)}{=} (\text{Kern } A^T)^\perp \stackrel{(i)}{=} (\text{Kern } AA^T)^\perp \stackrel{(ii)}{=} \text{Bild}(AA^T)^T = \text{Bild}(AA^T)$

□

Definition. $P : V \rightarrow U$ heißt „orthogonale Projektion auf U “, falls P Projektion auf U entlang U^\perp .

Satz. P orthogonale Projektion $\Leftrightarrow P = P^2$ und $P = P^T$.

Beweis.

„ \Rightarrow “

$\forall v, \tilde{v} \in V:$

$$\langle v, P(\tilde{v}) \rangle = \langle u_1 + u_0, \underbrace{P(\tilde{u}_1 + \tilde{u}_0)}_{\tilde{u}_1} \rangle = \langle u_1, \tilde{u}_1 \rangle + \underbrace{\langle u_0, \tilde{u}_1 \rangle}_{=0}$$

und ebenso $\langle P(v), \tilde{v} \rangle = \langle u_1, \tilde{u}_1 \rangle = \langle v, P(\tilde{v}) \rangle$. Also P selbstadjungiert.

„ \Leftarrow “

P Projektion auf Bild P entlang Kern $P = (\text{Bild } P^T)^\perp = (\text{Bild } P)^\perp$. □

Satz (Satz von Pythagoras \equiv orthogonale Projektion als Abstandsminimierung). Sei $U \subseteq V$ Untervektorraum und P orthogonale Projektion auf U . Dann gilt $\forall v \in V$:

$$\min_{u \in U} \|v - u\| = \|v - Pv\|$$

Beweis.

$$\begin{aligned} \|v - u\|^2 &= \langle v - Pv + Pv - u, v - Pv + Pv - u \rangle \\ &= \langle v - Pv, v - Pv \rangle + 2 \underbrace{\langle \underbrace{v + Pv}_{\in U_0 = U^\perp}, \underbrace{Pv - u}_{\in U} \rangle}_{=0} + \langle Pv - u, Pv - u \rangle \\ &= \|v - Pv\|^2 + \|Pv - u\|^2 \geq \|v - Pv\|^2 \end{aligned}$$

und „ $=$ “ $\Leftrightarrow \|Pv - u\| = 0 \Leftrightarrow u = Pv$. □

Satz (Iterierte orthogonale Projektionen). Für alle orthogonalen Projektionen $A, B \in \mathbb{R}^{n \times n}$ gilt:

$$\begin{aligned} AB \text{ orthogonale Projektion} &\Leftrightarrow AB = BA \\ &\Leftrightarrow AB \text{ orthogonale Projektion auf } (\text{Bild } A) \cap (\text{Bild } B) \end{aligned}$$

Beweis.

„ \Rightarrow “ AB orthogonale Projektion $\Rightarrow AB$ symmetrisch $\Rightarrow (AB) = (AB)^T = B^T A^T = BA$

„ \Rightarrow “ $AB = BA \Rightarrow AB$ symmetrisch und Bild $AB = (\text{Bild } A) \cap (\text{Bild } B)$, denn $x \in \text{Bild}(AB) = \text{Bild}(BA) \Rightarrow x \in (\text{Bild } A) \cap (\text{Bild } B)$

Umgekehrt: $x \in (\text{Bild } A) \cap (\text{Bild } B) \Rightarrow x = ax$ und $x = Bx \Rightarrow x = A(Bx) = ABx \in \text{Bild } AB$. □

11.3 Generalisiertinverse Matrizen

Definition. Für $A \in \mathbb{R}^{n \times k}$ definiere

$$A^- := \{G \in \mathbb{R}^{k \times n} \mid AGA = A\}.$$

Jede Matrix $G \in A^-$ heißt „eine generalisiertinverse Matrix von A “.

Satz. $\forall G \in \mathbb{R}^{k \times n}$:

$$G \in A^- \Leftrightarrow AG \text{ Projektion auf Bild } A$$

Beweis. „ \Rightarrow “

$$AG \cdot AG = AG \quad \checkmark$$

$$\text{Bild } A = \text{Bild } AGA \subseteq \text{Bild } AG \subseteq \text{Bild } A$$

„ \Leftarrow “

Sei $AG = (AG)^2$ und $\text{Bild } AG = \text{Bild } A$. Dann sei $x \in \mathbb{R}^k$ beliebig. Für $Ax \in \text{Bild } A$ existiert $y \in \mathbb{R}^n$ mit $Ax = AGy$. Daher $AGAx = AG \cdot AGy = Ax$ und $AGA = A$ \square

Satz (Bildinklusionslemma). Seien $A \in \mathbb{R}^{n \times k}$, $B \in \mathbb{R}^{n \times l}$. Dann gilt:

$$\text{Bild } A \subseteq \text{Bild } B \Leftrightarrow \forall G \in B^- : BGA = A$$

$$\Leftrightarrow \exists G \in B^- : BGA = A$$

Beweis.

„ \Rightarrow “

Seien $A = (s_1 \mid \dots \mid s_k)$ die Spalten von A . Dann gilt $\forall i \leq k : s_i \in \text{Bild } B$, d. h. $Bx = s_i$ ist lösbar. Wähle $c_i \in L(B, s_i)$, d. h. $Bc_i = s_i$, und baue zusammen zur Matrix $C := (c_1 \mid \dots \mid c_k) \in \mathbb{R}^{l \times k}$. Per Konstruktion gilt $A = BC$. Daraus $\forall G \in B^- : BGA = BGBC = BC = A$.

„ \Rightarrow “ \checkmark

„ \Leftarrow “

$$\text{Bild } A = \text{Bild } BGA \subseteq \text{Bild } A \quad \square$$

Frage: Ist überhaupt $A^- \neq \emptyset$

Schnelle Antwort: Sei U_0 ein komplementärer Untervektorraum zu $\text{Bild } A$ und sei P die Projektionsmatrix auf $\text{Bild } A$ entlang U_0 . Zerlege $P = (s_1 \mid \dots \mid s_n)$. Dann gilt $\forall j \leq n : s_j \in \text{Bild } P = \text{Bild } A$, d. h. $\exists g_j : Ag_j = s_j$. Zusammen erhält man $G = (g_1 \mid \dots \mid g_n) \in \mathbb{R}^{k \times n}$ mit $AG = P$. Daraus $AGA = PA = A$.

11.4 Pseudoinverse Matrizen¹

Penrose-Kriterien:

Gegeben $A \in \mathbb{R}^{n \times k}$ und $B \in \mathbb{R}^{k \times n}$.

(PK1) AB symmetrisch

¹Siehe E. H. MOORE, Bull. Amer. Math. Soc. 26 (1920), 394–395 und R. PENROSE, Proc. Cambridge Philos. Soc. 51 (1955), 406–413

(PK2) BA symmetrisch

(PK3) $ABA = A$

(PK4) $BAB = B$

Bemerkung.

(PK1) und (PK3) $\Rightarrow AB$ orthogonale Projektion

(PK2) und (PK4) $\Rightarrow BA$ orthogonale Projektion

Satz. $\forall A \in \mathbb{R}^{n \times k} \exists! B \in \mathbb{R}^{k \times n} : A$ und B erfüllen die vier Penrose-Kriterien.

Definition. Dieses B heißt „pseudoinverse Matrix von A “ und wird bezeichnet mit A^+

*Eindeutigkeitsbeweis*². Seien PK1–4 erfüllt von $B \in \mathbb{R}^{k \times n}$ und $A^+ \in \mathbb{R}^{k \times n}$.

$$\begin{aligned}
 B &= BAB && \text{(PK4)} \\
 &= (ABB^T)^T && \text{(PK1)} \\
 &= (AA^+ABB^T)^T && \text{(PK3)} \\
 &= (AA^+(BAB)^T)^T && \text{(PK1)} \\
 &= (AA^+B^T)^T && \text{(PK4)} \\
 &= BAA^+ && \text{(PK1)} \\
 &= (BA)^T A^+ && \text{(PK2)} \\
 &= (AA^+A)^T B^T A^+ && \text{(PK3)} \\
 &= A^+ AA^T B^T A^+ && \text{(PK2)} \\
 &= A^+ ABAA^+ && \text{(PK2)} \\
 &= A^+ AA^+ && \text{(PK3)} \\
 &= A^+ && \text{(PK4)}
 \end{aligned}$$

□

Nichteindeutigkeitsbeispiele.

PK 1, 2, 3: Sei $A = 0 \in \mathbb{R}$. $\forall A^+ = \alpha \in \mathbb{R}$:

PK1 ✓, PK2 ✓, PK3: $AA^+A = 0 = A$.

PK 1, 2, 4: Sei $A = 1 \in \mathbb{R}$. $\forall A^+ = \alpha \in \{0, 1\}$:

PK1 ✓, PK2 ✓, PK4: $A^+AA^+ = \alpha^2 = A^+$.

PK 1, 3, 4: Sei $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \in \mathbb{R}^{2 \times 2}$. $\forall A^+ = \begin{pmatrix} 1 & 0 \\ \alpha & 0 \end{pmatrix}, \alpha \in \mathbb{R}$:

PK1: $AA^+ = A$ symmetrisch ✓, PK3: $AA^+A = A^2 = A$ ✓, PK4: $A^+AA^+ = A^+$

²nach A. ALBERT, Regression and the Moore-Penrose inverse, N. Y. 1972, S. 72

11 Projektionen und verallgemeinerte Matrixinverse

PK 2, 3, 4: $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, A^+ = \begin{pmatrix} 1 & \alpha \\ 0 & 0 \end{pmatrix}, \alpha \in \mathbb{R} \dots$

Satz (konstruktiver Existenzsatz).

(i) $\forall \alpha \in \mathbb{R} : \alpha^+ = \begin{cases} \alpha^{-1} & \text{für } \alpha \neq 0 \\ 0 & \text{für } \alpha = 0 \end{cases} \quad \checkmark$

(ii) $\forall n \in \mathbb{N} : \forall D = \text{Diag}(\alpha_1, \dots, \alpha_n) \in \mathbb{R}^{n \times n} : D^+ = \text{Diag}(\alpha_1^+, \dots, \alpha_n^+) \quad \checkmark$

(iii) $\forall n \in \mathbb{N} \forall S \in \mathbb{R}_{\text{sym}}^{n \times n} : Q^T S Q \text{ diagonal} \Rightarrow S^+ = Q(Q^T S Q)^+ Q^T \text{ (symmetrisch) mit } Q \in O(n).$

(iv) $\forall n, k \in \mathbb{N} \forall A \in \mathbb{R}^{n \times k} : A^+ = (A^T A)^+ A^T = A^T (A A^T)^+$

Beweis.

(iii) Sei $Q^T S Q = D$ diagonal. Dann $S = Q D Q^T$.

$$S S^+ = Q D Q^T Q (Q^T S Q)^+ Q^T = Q \underbrace{D D^+}_{\text{diagonal}} Q^T \text{ symmetrisch}$$

$$S^+ S = Q D^+ D Q^T \text{ symmetrisch}$$

$$S S^+ S = Q \underbrace{D D^+ Q^T Q D}_{=D} Q^T = S$$

$$S^+ S S^+ = Q D^+ Q^T = S^+$$

(iv) Setze $S = A^T A$. Dann für $A^+ = S^+ A^T$:

$$A A^+ = A S^+ A^T \text{ symmetrisch}$$

$$A^+ A = S^+ A^T A = S^+ S \text{ symmetrisch}$$

$$A A^+ A = (A S^+ S)^{TT} = (S S^+ A^T)^T = (A^T)^T = A,$$

denn $S S^+$ ist orthogonale Projektion auf $\text{Bild } S = \text{Bild } A^T A = \text{Bild } A^T$, daraus folgt $S S^+ A^T = A^T$.

$$A^+ A A^+ = S^+ S S^+ A^T = S^+ A^T = A^+$$

□

30. 6. 2004 **Satz (Spezialfälle).**

(i) $\forall v \in \mathbb{R}^n : v^+ = \begin{cases} \frac{1}{\|v\|^2} v^T & \text{für } v \neq 0 \\ 0 & \text{für } v = 0 \end{cases}$

(ii) $\forall P \in \mathbb{R}^{n \times n} : P = P^2 = P^T \Rightarrow P^+ = P$

(iii) $\forall G \in \text{GL}(n) : G^+ = G^{-1}$

(iv) $\forall A \in \mathbb{R}^{n \times k} : A^+ = \begin{cases} A^T(AA^T)^{-1} & \text{für Rang } A = n \\ (A^T A)^{-1}A^T & \text{für Rang } A = k \end{cases}$

Beweis. Penrose-Kriterien nachprüfen... □

Satz (Rechenregeln).

(i) $\forall A \in \mathbb{R}^{n \times k} \forall \alpha \in \mathbb{R} : (\alpha A)^+ = \alpha^+ A^+$

(ii) $\forall A \in \mathbb{R}^{n \times k} :$
 $A^{++} = A$
 $A^{T+} = A^{+T}$
 $(A^T A)^+ = A^+ A^{T+}$

(iii) $\forall S \in \mathbb{R}_{sym.}^{n \times n} : (S^2)^+ = (S^+)^2$
 $SS^+ = S^+S$

(iv) $\forall P \in \mathbb{R}_{proj., sym.}^{n \times n} \forall A \in \mathbb{R}^{n \times k} :$
 $(PA)^+ = (PA)^+ P$
 $(A^T P)^+ = P(A^T P)^+$

In Worten: Orthogonale Projektionen gehen beim Pseudoinvertieren auf der anderen Seite als zusätzlicher Faktor heraus.

Beweis. Verifizieren von PK1–4, außer bei (iii),2:

$$SS^+ \stackrel{\text{PK1}}{=} (SS^+)^T = S^{+T} S^T = S^{T+} S^T = S^+ S$$

□

Satz (Bild, Kern und Projektionen).

(i) AA^+ ist orthogonale Projektion auf Bild A .
 $E_n - AA^+$ ist orthogonale Projektion auf $(\text{Bild } A^T)^\perp = \text{Kern } A$.

(ii) A^+A ist orthogonale Projektion auf Bild A^T .
 $E_k - A^+A$ ist orthogonale Projektion auf $(\text{Bild } A)^\perp = \text{Kern } A^T$.

(iii) $\text{Bild } A^+ = \text{Bild } A^T$
 $\text{Kern } A^+ = \text{Kern } A^T$

Beweis.

(i) AA^+ Projektion auf Bild A (denn A^+ ist insbesondere eine g-Inverse) und AA^+ symmetrisch (PK1). $\Rightarrow AA^+$ ist orthogonale Projektion auf Bild A .
 Für $E_n - AA^+$ siehe Übung 41.

(ii) A^+A ist symmetrisch (PK2) und idempotent (PK4) $\Rightarrow A^+A$ ist orthogonale Projektion auf Bild $A^+ = \text{Bild } A^T (AA^T)^+ \subseteq \text{Bild } A^T = \text{Bild}(AA^+A)^T = \text{Bild } A^+ AA^T \subseteq \text{Bild } A^+.$

□

Satz. Sei $P = P^2 = P^R \in \mathbb{R}^{n \times n}$. Dann gilt $\forall i = 1, \dots, n \forall j \neq i$:

$$p_{ij}^2 \leq p_{ii}(1 - p_{ii}).$$

Insbesondere gilt $p_{ii} \in [0, 1]$.

Erinnerung: Spur $P = \underbrace{\sum p_{ii}}_{\text{ganzzahlig}} = \text{Rang } P$.

Beweis.

$$p_{ii} = e_i^T P e_i = e_i^T P^2 e_i = e_i^T P^T P e_i = \|P e_i\|^2 = \sum_{j=1}^n p_{ij}^2 = p_{ii}^2 + \sum_{j \neq i} p_{ij}^2 \geq p_{ii}^2$$

$$\Rightarrow p_{ii} \geq 0 \text{ und } p_{ii} \leq 1 \text{ und } p_{ij}^2 \leq \sum_{j \neq i} p_{ij}^2 = p_{ii} - p_{ii}^2$$

□

11.5 Vektorhalbordnungen

Definition.

(i) Sei M Menge und \geq eine zweistellige Relation auf M , d. h. \geq Teilmenge von $M \times M$. Dann heißt \geq eine „**Halbordnung**“³, falls gilt:

- (1) $\forall x \in M : x \geq x$ (Reflexivität)
- (2) $\forall x, y \in M : x \geq y \text{ und } y \geq x \Rightarrow x = y$ (Antisymmetrie)
- (3) $\forall x, y, z \in M : x \geq y \text{ und } y \geq z \Rightarrow x \geq z$ (Transitivität)

(ii) Sei V ein euklidischer Vektorraum und \geq eine zweistellige Relation auf V . Dann heißt \geq eine „**Vektorhalbordnung**“, falls \geq Halbordnung ist und gilt:

- (4) $\forall x \in V : x \geq 0 \text{ und } -x \geq 0 \Rightarrow x = 0$
- (5) $\forall x, y \in V \forall \alpha > 0 : x \geq y \Rightarrow \alpha x \geq \alpha y$
- (6) $\forall x, y, z \in V : x \geq y \Rightarrow x + z \geq y + z$

(iii) Sei V ein euklidischer Vektorraum und $K \subseteq V$. Dann heißt K ein „**Ordnungskegel**“⁴, falls gilt

- (7) $0 \in K$
- (8) $\forall x \in V : x \in K \text{ und } -x \in K \Rightarrow x = 0$ (Geradenfreiheit)
- (9) $\forall x \in V \forall \alpha > 0 : x \in K \Rightarrow \alpha x \in K$ (Kegeleigenschaft)

³engl. *partial ordering*

⁴engl. *ordering cone*

$$(10) \quad \forall x, y \in V \quad \forall \alpha \in (0, 1) : x, y \in K \Rightarrow (1 - \alpha)x + \alpha y \in K \quad (\text{Konvexität})$$

Bemerkung.

a) $y \leq x \Leftrightarrow x \geq y$

b) Eine Halbordnung \geq , in der je zwei Punkte vergleichbar sind, d. h.

$$\forall x, y \in M : x \geq y \text{ oder } y \geq x,$$

heißt „**Totalordnung**“. Die Zahlengerade \mathbb{R} ist totalgeordnet.

c) K ist Kegel und K ist konvex $\Leftrightarrow \forall x, y \in K \quad \forall \alpha, \beta > 0 : \alpha x + \beta y \in K$

d) Wenn eine Menge $K \subseteq V$ mit je zwei Punkten $x, y \in K$

1) die *Strecke* zwischen x und y enthält, d. h. $\forall \alpha \in (0, 1) : (1 - \alpha)x + \alpha y \in K$, dann ist K konvex;

2) die *Gerade* durch x und y enthält, d. h. $\forall \alpha \in \mathbb{R} : (1 - \alpha)x + \alpha y \in K$, dann ist K „**lineare Menge**“, d. h. ein Untervektorraum.

Satz. Sei V ein euklidischer Vektorraum.

(i) Sei \geq Vektorhalbordnung. Dann ist $K := \{v \in V | v \geq 0\}$ ein Ordnungskegel.

(ii) Sei $K \subseteq V$ Ordnungskegel. Dann wird durch $x \geq y \Leftrightarrow x - y \in K$ eine Vektorhalbordnung definiert.

Beweis.

1. 7. 2004

(i) Aus (1) folgt (7), aus (4) folgt (8) und aus (5) folgt (9). Zu (10):

$$x \geq y \text{ und } \tilde{x} \geq \tilde{y} \underset{(6)}{\Rightarrow} \left. \begin{array}{l} \tilde{x} + y \geq \tilde{y} + y \\ x + \tilde{x} \geq y + \tilde{x} \end{array} \right\} \Rightarrow x + \tilde{x} \geq y + \tilde{x} \geq y + \tilde{y} \underset{(3)}{\Rightarrow} x + \tilde{x} \geq y + \tilde{y}$$

Also mit (9):

$$x \geq 0 \text{ und } y \geq 0 \Rightarrow (1 - \alpha)x \geq 0 \text{ und } \alpha y \geq 0 \Rightarrow (1 - \alpha)x + \alpha y \geq 0$$

(ii) (1) $x \geq x \Leftrightarrow x - x \in K \quad \checkmark$ wegen (7)

(2) aus (8): $\underbrace{x - y}_z \in K, \underbrace{y - x}_{-z} \in K \Rightarrow z = 0 \quad \checkmark$

(3) aus (9), (10): $x - y \in K, y - z \in K \Rightarrow x - z = 2 \cdot \left(\frac{1}{2}(x - y) + \frac{1}{2}(y - z)\right) \in K$

(4) aus (8) \checkmark

(5) aus (0) \checkmark

(6) aus Definition: $x - y \in K \Rightarrow (x + z) - (y + z) \in K \quad \checkmark$

□

Beispiel.

- a) \mathbb{R} mit Standardordnung \geq hat Ordnungskegel $[0, \infty)$ und ist Vektorhalbordnung, sogar Vektortotalordnung.
- b) \mathbb{R}^n mit Ordnungskegel $[0, \infty)^n$ induziert Vektorhalbordnung

$$x \geq y \Leftrightarrow x - y \in [0, \infty)^n \Leftrightarrow \forall j \leq n : x_j - y_j \geq 0.$$

Diese „komponentenweise Vektorhalbordnung“ ist nicht total, denn für $n \geq 2$ sind z. B. $x = (1, 0, \dots, 0, -1)^T$ und $y = 0$ unvergleichbar. Diese Vektorhalbordnung ist „abgeschlossen“ im Sinne von $\forall n : x_n \geq 0$ und $(x_n)_{n \geq 1}$ konvergent $\Rightarrow \lim x_n \geq 0$. Das selbe ist, daß $[0, \infty)^n$ als Menge abgeschlossen ist.

- c) \mathbb{R}^n mit Ordnungskegel $(0, \infty)^n \cup \{0\}$ hat Vektorhalbordnung. Diese ist nicht abgeschlossen.
- d) Im \mathbb{R}^n induziert der konvexe Kegel $(0, \infty)^n$ keine Vektorhalbordnung, sondern die Relation $x \geq y \Leftrightarrow x - y \in (0, \infty)^n$ ist

$$\forall x \in V : x \not\geq x \quad \text{irreflexiv,}$$

$$\forall x, y \in V : x > y \Rightarrow y \not\geq x \quad \text{antisymmetrisch}$$

und transitiv.

- 7.7.2004 e) \mathbb{R}^n mit Ordnungskegel $K := \{x \in \mathbb{R}^n \mid x_1 \geq x_2 \geq \dots \geq x_n \geq 0\}$ („Vektoren mit absteigend geordneten Komponenten“) induziert die Vektorhalbordnung

$$x \geq y \Leftrightarrow x - y \in K \Leftrightarrow x_1 - y_1 \geq x_2 - y_2 \geq \dots \geq x_n - y_n \geq 0.$$

- f) \mathbb{R}^n mit „Eistütenkegel“ $\{x \in \mathbb{R}^n \mid \sqrt{x_1^2 + \dots + x_{n-1}^2} \leq x_n\}$. Für Bilder dieser Kegel siehe F. PUKELSHEIM, *Optimal Design of Experiments*, S. 38.

Fazit: Es gibt eine Unmenge von Ordnungskegel und somit auch von Vektorhalbordnungen.

- g) Im Matrizenraum $\mathbb{R}^{n \times k}$ ist (gelegentlich) die komponentenweise Halbordnung von Interesse, mit Ordnungskegel $[0, \infty)^{n \times k}$. Das heißt $A \geq 0 \Leftrightarrow \forall i, j : a_{ij} \geq 0$.

- h) Im Matrizenraum $\mathbb{R}_{\text{sym.}}^{n \times n}$ wird durch

$$\mathbb{R}_{\text{n.neg.def.}}^{n \times n} := \{A \in \mathbb{R}_{\text{sym.}}^{n \times n} \mid \forall x \in \mathbb{R}^n : x^T A x \geq 0\} =: \text{NND}(n)$$

ein Ordnungskegel definiert.

Beweis.

$$(7) 0 \in \text{NND}(n) \quad \checkmark$$

(9) $\forall A \in \text{NND}(n) \forall \alpha > 0 : (\forall x \in \mathbb{R}^n : x^T(\alpha A)x = \underbrace{\alpha}_{>0} \underbrace{x^T Ax}_{\geq 0} \geq 0) \Rightarrow \alpha A \in \text{NND}(n).$

(10) $\forall A, B \in \text{NND}(n) \forall \alpha \in (0, 1) : (1 - \alpha)A + \alpha B \stackrel{?}{\in} \text{NND}(n)$
 $x^T((1 - \alpha)A + \alpha B)x = \underbrace{(1 - \alpha)}_{>0} \underbrace{x^T Ax}_{\geq 0} + \underbrace{\alpha}_{>0} \underbrace{x^T Bx}_{\geq 0} \geq 0$

(8) Sei $A \in \text{NND}(n)$ so, daß $-A \in \text{NND}(n)$. Wähle zu Eigenwert $\lambda \in \text{Spec } A$ einen Eigenvektor x . Dann

$$0 \leq x^T Ax = \lambda x^T x \Rightarrow \lambda \geq 0$$

und

$$0 \leq x^T(-A)x = -\lambda x^T x \Rightarrow \lambda \leq 0.$$

$$\Rightarrow \lambda = 0$$

□

Satz (Charakterisierung nichtnegativdefiniter Matrizen). Sei $A \in \mathbb{R}_{\text{sym.}}^{n \times n}$. Dann sind folgende Aussagen äquivalent:

(i) $A \in \text{NND}(n)$

(ii) $\forall x \in \mathbb{R}^n : x^T Ax \geq 0$

(iii) $\exists W \in \mathbb{R}_{\text{sym.}}^{n \times n} : A = W^2$

(iv) $\exists V \in \mathbb{R}^{\text{Rang } A \times n} : A = V^T V$

(v) $\text{Spec } A \subseteq [0, \infty)$, in Worten: Alle Eigenwerte des Endomorphismus A sind nichtnegativ.

Beweis. Definition und Übungen 37 und 38. □

Definition.

$$\text{PD}(n) := \mathbb{R}_{\text{pos.def.}}^{n \times n} := \{A \in \mathbb{R}_{\text{sym.}}^{n \times n} \mid \forall x \neq 0 : x^T Ax > 0\}$$

heißt der „Kegel der positiv definiten Matrizen“.

Bemerkung. $\text{PD}(n)$ ist kein Ordnungskegel. Er ist das Innere des Ordnungskegels $\text{NND}(n)$.

Satz (Charakterisierung positivdefiniter Matrizen). Sei $A \in \mathbb{R}_{\text{sym.}}^{n \times n}$. Dann sind folgende Aussagen äquivalent:

(i) $A \in \text{PD}(n)$

(ii) $\forall x \in \mathbb{R}^n \setminus \{0\} : x^T Ax > 0$

(iii) $\exists W \in \mathbb{R}_{\text{sym.}}^{n \times n} : A = W^2$ **und** $\text{Rang } W = n$

(iv) $\exists V \in \text{GL}(n) : A = V^T V$

11 Projektionen und verallgemeinerte Matrixinverse

(v) $\text{Spec } A \subseteq (0, \infty)$

(vi) $A \in \text{NND}(n)$ **und** $\text{Rang } A = n$

(vii) (Abschnittsdeterminantenkriterium.) $\forall m = 1, \dots, n : \det(A_{(m)}) > 0$, wobei

$$A_{(m)} := \begin{pmatrix} a_{11} & \dots & a_{1m} \\ \dots & \dots & \dots \\ a_{m1} & \dots & a_{mm} \end{pmatrix}.$$

Beweis. (vii) mit Satz von Sylvester. □

Bemerkung. Übergang von $\text{PD}(n)$ zu $\text{NND}(n)$ gelegentlich durch „Regularisierung“:

$$\underbrace{A}_{\in \text{NND}(n)} = \lim_{\epsilon \rightarrow 0} \underbrace{(A + \epsilon E_n)}_{\in \text{PD}(n)}$$

z. B. $AA^+ = A \cdot \lim_{\epsilon \rightarrow 0} (A + \epsilon E_n)^{-1}$:

Diagonalisierung: $A = Q^t \text{Diag}(\lambda_1, \dots, \lambda_n) Q$.

$$A + \epsilon E_n = Q^T \text{Diag}(\lambda_1, \dots, \lambda_n) Q + \epsilon Q^T Q = Q^T \text{Diag}(\epsilon + \lambda_1, \dots, \epsilon + \lambda_n) Q$$

$$(A + \epsilon E_n)^{-1} = (Q^T \text{Diag}(\epsilon + \lambda_1, \dots, \epsilon + \lambda_n) Q)^{-1} = Q^T \text{Diag}\left(\frac{1}{\epsilon + \lambda_1}, \dots, \frac{1}{\epsilon + \lambda_n}\right) Q$$

$$\begin{aligned} A(A + \epsilon E_n)^{-1} &= Q^T \text{Diag}(\lambda_i) Q \cdot Q^T \text{Diag}\left(\frac{1}{\epsilon + \lambda_i}\right) Q \\ &= Q^T \text{Diag}\left(\frac{\lambda_i}{\epsilon + \lambda_i}\right) Q \xrightarrow{\epsilon \rightarrow 0} Q^T \text{Diag}(\lambda_i \lambda_i^+) Q = AA^+ \end{aligned}$$

Ausblick:

a) Entweder weiter mit „Vektroptimierungsproblemen“ bzwgl. einer vorgegebenen Vektorhalbordnung:

Gegeben: (V, \geq) euklidischer Vektorraum mit Vektorhalbordnung \geq . Sei $M \subseteq V$.

Problem I: Finde einen „Maximalvektor $x \in M$ “ so, daß „kein anderer Vektor $y \in M$ besser als x ist“, d. h. $\forall y \in M : y \geq x \Rightarrow y = x$ („paretooptimal“).

Problem II: Finde einen „Maximumsvektor $x \in M$ “ so, daß x „besser ist als alle Konkurrenten $y \in M$ “, d. h. $\forall y \in M : x \geq y$.

Bemerkung. Bei Totalordnungen ist Problem I gleich Problem II. Bei Halbordnungen ist dies i. a. nicht so.

Beispiel. \mathbb{R} mit komponentenweiser Halbordnung mit Ordnungskegel $K := [0, \infty)^2$.

$$M := \text{Einheitskreis} = \{x \in \mathbb{R}^2 \mid x_1^2 + x_2^2 \leq 1\}$$

$$\forall z \in M : \{x \in \mathbb{R}^2 \mid x \geq z\} = \{x \in \mathbb{R}^2 \mid x - z \geq 0\} = \{x \in \mathbb{R}^2 \mid x \in z + K\}$$

Die Menge der Maximalvektoren von M bzgl. \geq ist der Rand des Viertelkreises im ersten Quadranten. Die Menge der Maximumsvektoren ist leer.

b) Oder weiter mit multilinearen Abbildungen.

12 Tensorprodukte oder Multilineare Algebra

Sei K ein Körper.

Definition. Seien V_0, V_1, \dots, V_n K -Vektorräume. Dann heißt eine Abbildung $f : V_1 \times \dots \times V_n \rightarrow V_0$ „ n -linear“, falls gilt:

$$\forall j = 1, \dots, n \forall v_1 \in V_1, \dots, v_{j-1} \in V_{j-1}, v_{j+1} \in V_{j+1}, \dots, v_n \in V_n :$$

$$\begin{aligned} f_j : V_j &\rightarrow V_0 \\ v_j &\mapsto f(v_1, \dots, v_{j-1}, v_j, v_{j+1}, \dots, v_n) \text{ ist linear.} \end{aligned}$$

Bemerkung.

- a) Falls $V_0 = K$, dann multilineare „Formen“.
- b) $n = 2$ ist Spezialfall der „bilinearen Abbildungen“ (salopp auch „Produkte“ genannt).

12.1 Tensorprodukte von Vektorräumen

Definition. Ein Paar (T, φ) heißt „Tensorprodukt“ der Vektorräume V und W , falls gilt:

(TP1) T ist Vektorraum

(TP2) φ ist eine bilineare Abbildung von $V \times W$ nach T mit der folgenden „**Faktorisierungseigenschaft**“.

(TP3) Für jeden Vektorraum U und für jede bilineare Abbildung $f : V \times W \rightarrow U$ existiert genau eine lineare Abbildung $f^* : T \rightarrow U$ mit $f = f^* \circ \varphi$.

Bemerkung.

- a) Zugehörige Abbildungsmengen sind also

$$L^2(V \times W, U) := \{f : V \times W \rightarrow U \mid f \text{ bilinear}\} \quad \text{und}$$

$$L(T, U) := \{g : T \rightarrow U \mid g \text{ linear}\}.$$

- b) Im Bildchen:

$$\begin{array}{ccc} V \times W & \xrightarrow{f \text{ bilinear}} & U \\ & \searrow \varphi & \nearrow f^* \text{ linear} \\ & & T \end{array}$$

12.1.1 Reduktionssatz

Satz. Seien V, W K -Vektorräume, (T, φ) Tensorprodukt für V und W und U ebenfalls ein K -Vektorraum. Dann ist die durch die Eigenschaft (TP3) gegebene Zuordnung

$$* : L^2(V \times W, U) \rightarrow L(T, U)$$

ein Vektorraum-Isomorphismus.

Beweis. L^2 und L sind Vektorräume. ✓

- I. $*$ ist wohldefinierte Abbildung wegen (TP3).
- II. $*$ ist surjektiv: Sei $h \in L(T, U)$. Dann betrachte $f := h \circ \varphi : V \times W \rightarrow U$. Offenbar ist f bilinear und $f^* = h$ und h Bildwert unter der Abbildung $*$.
- III. $*$ ist injektiv: Aus $f^* = g^*$ folgt $f = f^* \circ \varphi = g^* \circ \varphi = g$.
- IV. $*$ ist linear: Es gilt $(\alpha f + \beta g)^* = \alpha f^* + \beta g^*$, denn $(\alpha f^* + \beta g^*) \circ \varphi = \alpha f^* \circ \varphi + \beta g^* \circ \varphi = \alpha f + \beta g$.

□

12.1.2 Satz zur Minimaleigenschaft von T

Satz. (T, φ) sei Tensorprodukt für V und W . Dann gilt $T = \text{span}(\varphi(V \times W))$.

Beweis.

- I. φ ist bilinear und (T, φ) ist Tensorprodukt. Also existiert eine lineare Abbildung $\varphi^* \in L(T, T)$. Offenbar ist $\varphi^* = \text{id}_T$.
- II. Sei $\tilde{T} := \text{span}(\varphi(V \times W))$. Es ist \tilde{T} ein Untervektorraum von T . Definiere $\tilde{\varphi} : V \times W \rightarrow \tilde{T}$ durch $\tilde{\varphi}(v, w) := \varphi(v, w)$. Dann ist $\tilde{\varphi}$ bilinear und faktorisiert gemäß (TP3):

$$\tilde{\varphi} = \tilde{\varphi}^* \circ \varphi \text{ mit } \tilde{\varphi}^* \in L(T, \tilde{T}).$$

Mit der („Inklusions-“)Abbildung $j : \tilde{T} \rightarrow T$ gegeben durch $\forall v \in \tilde{T} : j(v) := v$ gilt dann: $\varphi = j \circ \tilde{\varphi} = \underbrace{j \circ \tilde{\varphi}^*}_{\text{linear}} \circ \varphi$.

- III. Aus der Eindeutigkeit in (TP3) folgt $j \circ \tilde{\varphi}^* = \varphi^* = \text{id}_T$. Also (mit Übung 10 aus LA I) ist j surjektiv und $\tilde{T} = T$.

□

12.1.3 Tensorprodukte von Teilräumen

Satz. Sei (T, φ) ein Tensorprodukt für V und W . \mathcal{L}_V sei Untervektorraum von V , \mathcal{L}_W sei Untervektorraum von W . Setze $T_{\mathcal{L}} := \text{span}(\varphi(\mathcal{L}_V \times \mathcal{L}_W)) \subseteq T$ und $\varphi_{\mathcal{L}} : \mathcal{L}_V \times \mathcal{L}_W \rightarrow T_{\mathcal{L}}$, $(v, w) \mapsto \varphi(v, w)$. Dann ist $(T_{\mathcal{L}}, \varphi_{\mathcal{L}})$ Tensorprodukt für \mathcal{L}_V und \mathcal{L}_W .

Beweis.

(TP1) ✓

(TP2) ✓

(TP3) Sei U ein Vektorraum und $f : \mathcal{L}_V \times \mathcal{L}_W \rightarrow U$ bilinear. Man wähle Fortsetzung $F : V \times W \rightarrow U$ bilinear (d. h. $F|_{\mathcal{L}_V \times \mathcal{L}_W} = f$). Dann existiert $F^* : T \rightarrow U$ linear mit $F = F^* \circ \varphi$. Daraus $f = F|_{\mathcal{L}_V \times \mathcal{L}_W} = F^* \circ \varphi|_{\mathcal{L}_V \times \mathcal{L}_W} = F^*|_{T_{\mathcal{L}}} \circ \varphi_{\mathcal{L}}$. Diese Faktorisierung ist eindeutig, denn mit einer zweiten Faktorisierung $f = h \circ \varphi_{\mathcal{L}}$ ist $h \circ \varphi_{\mathcal{L}} = f = F^*|_{T_{\mathcal{L}}} \circ \varphi_{\mathcal{L}}$ und somit $h = F^*|_{T_{\mathcal{L}}}$ wegen Minimalität von $T_{\mathcal{L}}$.

□

12.1.4 Eindeutigkeit von Tensorprodukten

Satz. (T, φ) sei Tensorprodukt für V und W . \tilde{T} sei Vektorraum und $\tilde{\varphi} : V \times W \rightarrow \tilde{T}$ bilinear. Dann ist $(\tilde{T}, \tilde{\varphi})$ genau dann Tensorprodukt für V und W , wenn es einen Vektorraum-Isomorphismus $i : T \rightarrow \tilde{T}$ gibt mit $i \circ \varphi = \tilde{\varphi}$.

14. 7. 2004 *Beweis.*

„ \Rightarrow “

I. Da φ Tensorprodukt und $\tilde{\varphi}$ bilinear, existiert Faktorisierung $\tilde{\varphi} = i \circ \varphi$ mit $i \in L(T, \tilde{T})$.

II. Da $\tilde{\varphi}$ Tensorprodukt und φ bilinear, existiert Faktorisierung $\varphi = h \circ \tilde{\varphi}$ mit $h \in L(\tilde{T}, T)$.

III. Somit gilt:

$$\varphi(v, w) = h \circ \tilde{\varphi}(v, w) = h \circ i \circ \varphi(v, w)$$

$$\tilde{\varphi}(v, w) = i \circ \varphi(v, w) = i \circ h \circ \tilde{\varphi}(v, w)$$

Wegen Satz 12.1.2 und der Linearität von $h \circ i$ und $i \circ h$ folgt

(i) $h \circ i = \text{id}_T$ und

(ii) $i \circ h = \text{id}_{\tilde{T}}$.

Mit Übungsaufgabe 10 (LA I):

(i) i injektiv und h surjektiv und

(ii) h injektiv und i surjektiv.

Somit i Vektorraum-Isomorphismus und $h = i^{-1}$.

„ \Leftarrow “

(TP1) ✓

(TP2) $\tilde{\varphi}(v, w) = i \circ \varphi(v, w)$ ist bilinear. ✓

(TP3) Sei U Vektorraum und $f : V \times W \rightarrow U$ bilinear. Dann gilt:

$$f \stackrel{\text{Vor.}}{=} f^* \circ \varphi = \underbrace{f^* \circ i^{-1}}_{=: \tilde{f}^*} \circ \underbrace{i \circ \varphi}_{\tilde{\varphi}} = \tilde{f}^* \circ \tilde{\varphi}$$

Diese Faktorisierung ist eindeutig, denn aus $f = j \circ \tilde{\varphi}$ folgt

$$f = j \circ \tilde{\varphi} = \underbrace{j \circ i \circ \varphi}_{\text{linear}}$$

Da φ Tensorprodukt, muß gelten $j \circ i = f^*$. Daraus: $j = f^* \circ i^{-1} = \tilde{f}^*$.

□

12.1.5 Existenz von Tensorprodukten

Satz. Seien V und W K -Vektorräume mit $\dim V = n$ und $\dim W = k$. Dann existiert ein Tensorprodukt (T, φ) für $V \times W$. Es ist $\dim T = n \cdot k$.

Beweis. Es seien e_1^V, \dots, e_n^V Basisvektoren für V und e_1^W, \dots, e_k^W Basisvektoren für W .

I. (TP1) Konstruktion eines Vektorraums T :

Betrachte $S := \{(e_i^V, e_j^W) \mid i = 1, \dots, n, j = 1, \dots, k\}$ mit $\#S = n \cdot k$.

Sei $T := K^S \equiv \{t : S \rightarrow K \mid f \text{ Abbildung}\}$.

Dann ist T ein K -Vektorraum mit $\dim T = \#S = n \cdot k$ und mit Basis t_{ij} , definiert durch

$$t_{ij}(e_i^V, e_j^W) := \delta_{i\tilde{i}} \cdot \delta_{j\tilde{j}} \equiv \begin{cases} 1 & \text{für } i = \tilde{i}, j = \tilde{j} \\ 0 & \text{sonst.} \end{cases}$$

II. (TP2) Konstruktion von φ :

$V \times W \rightarrow T$ bilinear.

Dazu identifiziere $v \in V$ mit seinem Koordinatentupel $x = (x_1, \dots, x_n)^T$ gemäß $v = \sum_{i=1}^n x_i e_i^V$. Identifiziere entsprechend $w \in W$ mit $y = (y_1, \dots, y_k)^T$. Damit definiere

$$\varphi(v, w) := \sum_{i=1}^n \sum_{j=1}^k x_i y_j t_{ij} \in T.$$

Diese Abbildung ist bilinear. Insbesondere $\varphi(e_i^V, e_j^W) = t_{ij}$.

III. (TP3)

Sei U Vektorraum und $f : V \times W \rightarrow U$ bilinear. Sei $f^* \in L(T, U)$ definiert durch $f^*(t_{ij}) := f(e_i^V, e_j^W)$. Dann

$$\begin{aligned} f(v, w) &= f\left(\sum_i x_i e_i^V, \sum_j y_j e_j^W\right) = \sum_i \sum_j x_i y_j f(e_i^V, e_j^W) \\ &= \sum_i \sum_j x_i y_j f^*(t_{ij}) = \sum_i \sum_j x_i y_j f^*(\varphi(e_i^V, e_j^W)) = f^* \circ \varphi(v, w). \end{aligned}$$

Ist zudem $f = h \circ \varphi$ dann $h(t_{ij}) = f(e_i^V, e_j^W) = f^*(t_{ij})$ und somit $h = f^*$.

□

12.2 Tensorprodukte von linearen Abbildungen

Seien V_1, V_2, W_1, W_2 endlichdimensionale K -Vektorräume.

Sei $A \in L(V_1, V_2)$ eine lineare Abbildung von V_1 nach V_2

und $B \in L(W_1, W_2)$ eine lineare Abbildung von W_1 nach W_2

und (T_1, φ_1) ein Tensorprodukt für V_1 und W_1

und (T_2, φ_2) ein Tensorprodukt für V_2 und W_2 .

Betrachte $f : V_1 \times W_1 \rightarrow T_2, (v, w) \mapsto \varphi_2(Av, Bw)$. Dann offenbar f bilinear. Also Faktorisierung über φ_1 möglich, d. h. $\exists! \psi_{21}(A, B) \in L(T_1, T_2)$ mit $\varphi_2(Av, Bw) = f(v, w) = \psi_{21}(A, B) \circ \varphi_1(v, w)$.

Definition. Eine lineare Abbildung $\psi_{21}(A, B) \in L(T_1, T_2)$ heißt „**Tensorprodukt von A und B** “ (bzgl. (T_1, φ_1) und (T_2, φ_2)), falls $\forall v \in V_1 \forall w \in W_1 :$

$$\psi_{21}(A, B) \circ \varphi_1(v, w) = \varphi_2(Av, Bw).$$

Bemerkung. Die Zuordnung $(A, B) \mapsto \psi_{21}(A, B)$ definiert eine bilineare Abbildung

$$\psi_{21} : L(V_1, V_2) \times L(W_1, W_2) \rightarrow L(T_1, T_2).$$

Sind alle Räume endlichdimensional, dann ist $(L(T_1, T_2), \psi_{21})$ ein Tensorprodukt für $L(V_1, V_2)$ und $L(W_1, W_2)$. Im Fall von unendlicher Dimension ist das nicht mehr richtig.

12.2.1 Rangformel

Satz. $\text{Rang } \psi_{21}(A, B) = \text{Rang } A \cdot \text{Rang } B$

Beweis.

I. Setze $T_{\mathcal{L}} := \text{span}(\varphi_2(\text{Bild } A \times \text{Bild } B))$ und $\varphi_{\mathcal{L}} : \text{Bild } A \times \text{Bild } B \rightarrow T_{\mathcal{L}},$
 $(v, w) \mapsto \varphi_2(v, w)$.

Nach Satz 12.1.3 ist $(T_{\mathcal{L}}, \varphi_{\mathcal{L}})$ wieder ein Tensorprodukt für $\text{Bild } A$ und $\text{Bild } B$. Für die Dimension gilt gemäß Satz 12.1.5 $\dim T_{\mathcal{L}} = \dim \text{Bild } A \cdot \dim \text{Bild } B = \text{Rang } A \cdot \text{Rang } B$.

12.3 Das Kroneckerprodukt (Tensorprodukt für euklidische Vektorräume)

II. Da das Bild von φ_1 den Raum T_1 aufspannt (Satz 12.1.2), gilt:

$$\begin{aligned} \text{Bild } \psi_{21}(A, B) &= \psi_{21}(A, B)(\text{span } \varphi_1(V \times W)) \\ &=_{\psi_{21} \circ \varphi_1 = \varphi_2} \text{span } \varphi_2(\text{Bild } A \times \text{Bild } B) = T_{\mathcal{L}} \end{aligned}$$

III. $\text{Rang } \psi_{21}(A, B) = \dim \text{Bild } \psi_{21}(A, B) = \dim T_{\mathcal{L}} = \text{Rang } A \cdot \text{Rang } B.$

□

Also Tensorprodukt von linearen Abbildungen $\psi_{21}(A, B)$ definiert als Lösung von

15. 7. 2004

$$\psi_{21}(A, B) \circ \varphi_1(v, w) = \varphi_2(Av, Bw)$$

für alle $v \in V_1, w \in W_1$, bzw. wegen Bilinearität einfach für alle Basisvektoren einer Basis von V_1 und einer Basis von W_1 .

12.2.2 Tensorprodukte von Produkten linearer Abbildungen

Satz. Seien V_3, W_3 weitere Vektorräume und (T_3, φ_3) ein Tensorprodukt für V_3 und W_3 . Seien $\tilde{A} \in L(V_2, V_3)$ und $\tilde{B} \in L(W_2, W_3)$. Dann gilt:

$$\psi_{31}(\tilde{A}, \tilde{B}) = \psi_{32}(\tilde{A}, \tilde{B}) \circ \psi_{21}(A, B).$$

Beweis. Verifikation der definierenden Gleichung:

$$\psi_{32}(\tilde{A}, \tilde{B}) \circ \underbrace{\psi_{21}(A, B) \circ \varphi_1(v, w)}_{=\varphi_2(Av, Bw)} = \varphi_3(\tilde{A}Av, \tilde{B}Bw)$$

□

12.3 Das Kroneckerprodukt (Tensorprodukt für euklidische Vektorräume)

Jetzt $V = \mathbb{R}^n, W = \mathbb{R}^k, (T, \varphi)$ Tensorprodukt. Dann $\dim T = nk$. Spezielle Kandidaten: $(\mathbb{R}^{nk}, \otimes)$ oder $(\mathbb{R}^{n \times k}, (x, y) \mapsto xy^T)$.

Satz. Mit

$$\begin{aligned} \otimes : \mathbb{R}^n \times \mathbb{R}^k &\rightarrow \mathbb{R}^{nk} \\ (x, y) &\mapsto x \otimes y := \begin{bmatrix} x_1 y \\ x_2 y \\ \vdots \\ x_n y \end{bmatrix} \end{aligned}$$

ist $(\mathbb{R}^{nk}, \otimes)$ ein Tensorprodukt für \mathbb{R}^n und \mathbb{R}^k .

12 Tensorprodukte oder Multilineare Algebra

Beweis. In Satz 12.1.5 war T konstruiert mit Basisvektoren t_{ij} und bilinearer Abbildung

$$\varphi(x, y) = \sum_i \sum_j x_i y_j t_{ij}.$$

Transport der t_{ij} von T in \mathbb{R}^{nk} durch

$$s(t_{ij}) = \begin{bmatrix} (e_i^n)_1 e_j^k \\ (e_i^n)_2 e_j^k \\ \vdots \\ (e_i^n)_n e_j^k \end{bmatrix} = \text{„lauter Nullen, nur in } i\text{-tem Block auf } j\text{-ter Stelle eine 1“}.$$

Mit diesen Basisvektoren dann

$$x \otimes y = \sum_i \sum_j x_i y_j \underbrace{(e_i^n \otimes e_j^k)}_{=s(t_{ij})} = \sum_i \sum_j x_i y_j \begin{bmatrix} (e_i^n)_1 e_j^k \\ \vdots \\ (e_i^n)_n e_j^k \end{bmatrix} = \begin{bmatrix} x_1 y \\ \vdots \\ x_n y \end{bmatrix} = \begin{pmatrix} x_1 y_1 \\ \vdots \\ x_1 y_k \\ \hline x_2 y_1 \\ \vdots \\ x_2 y_k \\ \hline \vdots \\ \hline x_n y_1 \\ \vdots \\ x_n y_k \end{pmatrix}.$$

□

12.3.1 Kroneckerprodukt von zwei Matrizen

Kroneckerprodukt von linearen Abbildungen \equiv Matrizen: $A \in \mathbb{R}^{n_2 \times n_1}$, $B \in \mathbb{R}^{k_2 \times k_1}$. Dann ist $\psi_{21}(A, B)$ Lösung von $\psi_{21}(A, B)(\underbrace{x \otimes y}_{\varphi_1(x, y)}) = (Ax) \otimes (By)$.

Satz.

(i) *Darstellung als Blockmatrix:*

$$\psi_{21} =: A \otimes B = \begin{bmatrix} a_{11}B & a_{12}B & \dots & a_{1n_1}B \\ \dots & \dots & \dots & \dots \\ a_{n_2 1}B & a_{n_2 2}B & \dots & a_{n_2 n_1}B \end{bmatrix} \in \mathbb{R}^{n_2 k_2 \times n_1 k_1}$$

(ii) $\text{Rang}(A \otimes B) = \text{Rang } A \cdot \text{Rang } B$

12.3 Das Kroneckerprodukt (Tensorprodukt für euklidische Vektorräume)

(iii) $(A \otimes B)(x \otimes y) = Ax \otimes By$ und allgemeiner:

$$(\tilde{A} \otimes \tilde{B})(A \otimes B) = (\tilde{A}A) \otimes (\tilde{B}B)$$

für $\tilde{A} \in \mathbb{R}^{n_3 \times n_2}$, $\tilde{B} \in \mathbb{R}^{k_3 \times k_2}$

(iv) $(A \otimes B)^T = (A^T) \otimes (B^T)$

(v) $(A \otimes B)^+ = (A^+) \otimes (B^+)$

(vi) $P \in \mathbb{R}_{proj}^{n \times n}$ und $Q \in \mathbb{R}_{proj}^{k \times k} \Rightarrow P \otimes Q \in \mathbb{R}_{proj}^{nk \times nk}$

Beweis.

(i) Zu zeigen: $(A \otimes B)(x \otimes y) = (Ax) \otimes (By)$.

$$\begin{bmatrix} a_{11}B & \dots & a_{1n_1}B \\ \dots & \dots & \dots \\ a_{n_2 1}B & \dots & a_{n_2 n_1}B \end{bmatrix} \begin{bmatrix} x_1 y \\ \vdots \\ x_n y \end{bmatrix} = \begin{bmatrix} (\sum a_{1j} x_j) By \\ \vdots \\ (\sum a_{n_2 j} x_j) By \end{bmatrix}$$

(ii) Aus Satz 12.2.1.

(iii) Konkretisierung des Tensorprodukts von linearen Abbildungen.

(iv) Aus Blockdarstellung.

(v) PK 1–4 verifizieren. Zum Beispiel:

$$\begin{aligned} [(A \otimes B)(A^+ \otimes B^+)]^T &= [(AA^+) \otimes (BB^+)]^T = (AA^+)^T \otimes (BB^+)^T \\ &= (AA^+) \otimes (BB^+) = (A \otimes B)(A^+ \otimes B^+) \end{aligned}$$

(vi) $(P \otimes Q)^2 = (P^2) \otimes (Q^2) = P \otimes Q$. Falls orthogonal: $(P \otimes Q)^T = (P^T) \otimes (Q^T) = P \otimes Q$.

□

$$xy^T = \begin{bmatrix} x_1 y^T \\ \vdots \\ x_n y^T \end{bmatrix}$$

Definition. $\text{vec} : \mathbb{R}^{n \times k} \rightarrow \mathbb{R}^{nk}$,
 $xy^T \mapsto x \otimes y$

(ist linear).

Satz. $\forall A \in \mathbb{R}^{n \times k} : \text{vec } A = (\text{„Zeile an Zeile“})^T$

12 Tensorprodukte oder Multilineare Algebra

Beweis.

$$A = \sum_i \sum_j a_{ij} \underbrace{e_i^n (e_j^k)^T}_{\substack{= E_{ij} = n \times k\text{-Matrix} \\ \text{mit } (i, j)\text{-Eintrag } 1 \\ \text{und sonst lauter Nullen}}}$$

$$\Rightarrow \text{vec } A = \sum_i \sum_j a_{ij} \underbrace{\text{vec}(e_i^n (e_j^k)^T)}_{e_i^n \otimes e_j^k} = (\text{„Zeile an Zeile“})^T$$

□

21.7.2004 12.3.2 Der Vektorisierungsoperator

Satz.

(i) Sowohl (\mathbb{R}^n, \otimes) als auch $(\mathbb{R}^{n \times k}, (x, y) \mapsto xy^T)$ sind Tensorprodukte für \mathbb{R}^n . Ein Isomorphismus ist

$$\text{vec} : \mathbb{R}^{n \times k} \rightarrow \mathbb{R}^{nk}$$

$$A \mapsto \begin{pmatrix} a_{11} \\ \vdots \\ a_{1k} \\ \hline a_{21} \\ \vdots \\ a_{2k} \\ \vdots \end{pmatrix} = (\text{„Zeile an Zeile“})^T.$$

Dieser ist „tensorprodukttreu“, d. h. $\text{vec}(xy^T) = x \otimes y$.

(ii) Es gilt $\text{vec}(ABC) = (A \otimes C^T) \text{vec}(B)$ für zusammenpassende Matrizen A, B, C .

(iii) vec ist skalarprodukttreu für $(\mathbb{R}^{n \times k}, \text{Spur } AB^T)$ und $(\mathbb{R}^{nk}, x^T y)$.

Beweis.

(i) Aus Linearität und Tensorprodukttreue folgt die Gestalt von $\text{vec } A$, nämlich

$$\text{vec } A = \text{vec} \left(\sum_i \sum_j a_{ij} e_i^n (e_j^k)^T \right) = \sum_i \sum_j a_{ij} \underbrace{\text{vec}(e_i^n (e_j^k)^T)}_{e_i^n \otimes e_j^k} = (\text{„Zeile an Zeile“})^T.$$

12.3 Das Kroneckerprodukt (Tensorprodukt für euklidische Vektorräume)

(ii) Mit $B = \sum \sum b_{ij} e_i e_j^T$ gilt:

$$\begin{aligned} \text{vec}(ABC) &= \sum_i \sum_j b_{ij} \text{vec}(Ae_i e_j^T C) \\ &= \sum_i \sum_j b_{ij} \text{vec}(Ae_i (C^T e_j)^T) \\ &= \sum_i \sum_j b_{ij} (Ae_i) \otimes (C^T e_j) \\ &\stackrel{12.3.1(iii)}{=} \sum_i \sum_j b_{ij} (A \otimes C^T)(e_i \otimes e_j) \\ &= (A \otimes C^T)(\text{vec } B) \end{aligned}$$

(iii) $x^T y = \sum x_i y_i$ ist inneres Produkt auf \mathbb{R}^{nk} . ✓
 $\text{Spur } AB^T = \sum \sum a_{ij} b_{ij}$ ist inneres Produkt auf $\mathbb{R}^{n \times k}$. ✓
 $(\text{vec } A)^T (\text{vec } B) = \sum \sum a_{ij} b_{ij} = \text{Spur}(AB^T)$

□

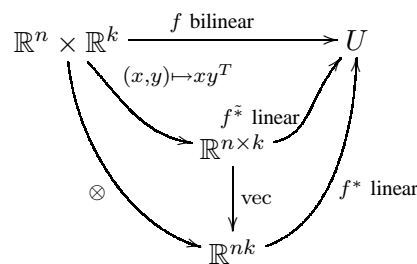
Bemerkung. Für $\|A\| := \sqrt{\text{Spur } AA^T}$ gilt $\|AB\| \leq \|A\| \cdot \|B\|$.

Beweis.

$$\begin{aligned} \|AB\|^2 &= \text{Spur } AB B^T A^T = \sum_i \sum_j \left(\sum_k a_{ik} b_{kj} \right)^2 \\ &\stackrel{(CS)}{\leq} \sum_i \sum_j \left(\sum_k a_{ik} \right)^2 \left(\sum_k b_{kj} \right)^2 = \|A\|^2 \cdot \|B\|^2 \end{aligned}$$

□

(Konkrete) Tensorprodukte für \mathbb{R}^n und \mathbb{R}^k :



mit $f(x, y) = f^*(x \otimes y)$ und $f(x, y) = f^{\tilde{}}(xy^T)$.