

Skript zur Vorlesung

Lineare Algebra 2 für
Lehramtsstudierende (GS/HS/RS)

(Sommersemester 2010)

Dieses Geheft enthält in kompakter Form die wesentlichen Inhalte, wie sie in der Vorlesung „Lineare Algebra für Lehramtsstudierende (GS/HS/RS)“ vorgestellt werden.

Es ist zum Gebrauch neben der Vorlesung gedacht und erhebt nicht den Anspruch, „in sich selbst verständlich“ oder vollständig zu sein.

S. Hilger

Dieses Skript liegt in einer jeweils aktualisierten Form im Internet vor:

<http://www.ku-eichstaett.de/Fakultaeten/MGF/Didaktiken/dphys/Lehre.de>

Inhaltsverzeichnis

5	Permutationen	4
5.1	Einführung	4
5.2	Zyklen und Transpositionen	6
5.3	Zerlegung von Permutationen	7
5.4	Das Signum einer Permutation	8
6	Determinanten	11
6.1	Definition und Beispiele	11
6.2	Eigenschaften	13
6.3	Der Multiplikationssatz	16
6.4	Die Entwicklung einer Determinante	17
6.5	Einschub: Transposition einer Matrix	19
6.6	Invertierbare Matrizen und Determinanten	20
6.7	Beispiel	23
7	Euklidische Vektorräume	25
7.1	Grundlegende Definitionen	25
7.2	Norm und Orthogonalität	28
7.3	Das Vektorprodukt im euklidischen \mathbb{R}^3	33
8	Diagonalisierbarkeit	34
8.1	Vektoren beim Basiswechsel	35
8.2	Homomorphismen beim Basiswechsel, Äquivalenz von Matrizen	37
8.3	Beispiel: Orthogonale Polynome	39
8.4	Endomorphismen beim Basiswechsel, Ähnlichkeit von Matrizen	44
8.5	Exkurs: Polynome	47
8.6	Das charakteristische Polynom	49
8.7	Eigenwerte und Eigenvektoren	51
8.8	Diagonalisierbarkeit	54
8.8.1	Algorithmus zur Diagonalisierung	56
9	Abbildungen in euklidischen Vektorräumen	64
9.1	Das Skalarprodukt beim Basiswechsel	64
9.2	Die adjungierte Abbildung	64
9.3	Orthogonale Abbildungen	66
9.3.1	Orthogonale Abbildungen im euklidischen \mathbb{R}^2	68
9.4	Diagonalisierung in euklidischen Vektorräumen	72
10	Elementare affine Geometrie	77
10.1	Affine Räume	77
10.2	Affine Unterräume	79
10.3	Euklidische affine Räume	81
10.3.1	Der Abstand zwischen einem Punkt und einem affinen Unterraum	81
10.3.2	Hyperebenen und Normalenvektoren	83
10.3.3	Isometrien	84

Literatur

5 Permutationen

5.1 Einführung

Definitionen:

1. Es sei X eine Menge. Eine bijektive Abbildung $f : X \rightarrow X$ heißt auch *Permutation von X* .
2. Die Menge aller Permutationen einer Menge X bezeichnen wir mit

$$\mathcal{S}(X) := \{f : X \rightarrow X \mid f \text{ ist bijektiv}\}.$$

Beobachtung: Die Menge $\mathcal{S}(X)$ bildet mit der Hintereinanderausführung als Verknüpfung

$$\circ \begin{cases} \mathcal{S}(X) \times \mathcal{S}(X) & \rightarrow \mathcal{S}(X) \\ (f, g) & \mapsto f \circ g \end{cases}$$

eine Gruppe.

Begründung: Die Hintereinanderausführung von Funktionen ist immer assoziativ: Für drei Abbildungen $f, g, h : X \rightarrow X$ gilt:

$$[(h \circ g) \circ f](x) = (h \circ g)(f(x)) = (h(g(f(x)))) = h((g \circ f)(x)) = [h \circ (g \circ f)](x).$$

Das neutrale Element ist durch die identische Abbildung id_X gegeben. Die zu einer Abbildung $f \in \mathcal{S}(X)$ inverse Abbildung ist gegeben durch die Umkehrabbildung f^{-1} .

Beispiel: Wir betrachten auf der Menge \mathbb{R} die beiden Abbildungen $f, g \in \mathcal{S}(\mathbb{R})$

$$f(x) = 2x, \quad g(x) = x + 1.$$

Dann gilt:

$$(f \circ g)(x) = f(x + 1) = 2(x + 1), \quad (g \circ f)(x) = g(2x) = 2x + 1.$$

Es ist also $f \circ g \neq g \circ f$. Die Gruppe $\mathcal{S}(\mathbb{R})$ ist nicht kommutativ.

Definition: Ist X eine endliche Menge der Mächtigkeit n , so heißt $\mathcal{S}(X)$ die *Symmetrische Gruppe auf n Elementen (oder n Buchstaben)*.

Im allgemeinen hat man es mit der Menge $X = \{1, 2, \dots, n\}$ zu tun.

Man kürzt $\mathcal{S}_n = \mathcal{S}(\{1, 2, \dots, n\})$ ab. Die Elemente werden im allgemeinen durch den Buchstaben π gekennzeichnet. Die Hintereinanderausführung zweier Permutationen schreiben wir ab jetzt als Multiplikation:

$$\pi_1 \cdot \pi_2 = \pi_1 \circ \pi_2.$$

Wie sonst auch üblich, kann das Multiplikationszeichen \cdot auch weggelassen werden.

Beispiele:

- \mathcal{S}_1 besteht nur aus der identischen Abbildung $\text{id} : \{1\} \rightarrow \{1\}$. Als Gruppe ist dies wieder die triviale Gruppe. $\mathcal{S}_1 = \{\text{id}\}$.

- \mathcal{S}_2 enthält zwei Abbildungen, nämlich die identische Abbildung $\iota = \text{id}_X$ und die Vertauschung der beiden Elemente 1 und 2. Bezeichnen wir letztere mit τ , so gilt für die Verknüpfungen:

$$\begin{array}{c|c|c} \cdot & \iota & \tau \\ \hline \iota & \iota & \tau \\ \hline \tau & \tau & \iota \end{array}$$

Diese Struktur ist wieder die gleiche wie die von $\mathbb{Z}_2 = \{g, u\}$. Man sagt, die Gruppen (\mathbb{Z}_2, \oplus) und (\mathcal{S}_2, \circ) sind *isomorph (gleichstrukturiert)*. Auch die Gruppe $(\{-1, +1\}, \cdot)$ ist isomorph zu diesen beiden Gruppen.

- Wir betrachten jetzt die Gruppe \mathcal{S}_3 der bijektiven Abbildungen in $\{1, 2, 3\}$. Wir bezeichnen für $a, b, c \in \{1, 2, 3\}$ die Abbildung

$$\pi : \begin{cases} 1 \mapsto a \\ 2 \mapsto b \\ 3 \mapsto c \end{cases} \quad \text{mit} \quad \begin{pmatrix} 1 & 2 & 3 \\ a & b & c \end{pmatrix}.$$

Dann gilt:

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

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

Die Gruppe \mathcal{S}_3 ist nicht kommutativ.

Allgemein ist $\mathcal{S}(X)$ nicht kommutativ, falls $|X| \geq 3$.

- Allgemeiner bezeichnen wir ein Element $f \in \mathcal{S}_n$

$$\pi : \begin{cases} 1 \mapsto x_1 \\ 2 \mapsto x_2 \\ \vdots \\ n \mapsto x_n \end{cases} \quad \text{mit} \quad \begin{pmatrix} 1 & 2 & \cdots & n \\ x_1 & x_2 & \cdots & x_n \end{pmatrix}.$$

- Es gilt $|\mathcal{S}_n| = n!$ (Induktions-Beweis in der Übung)

5.2 Zyklen und Transpositionen

Definitionen:

1. Eine Permutation $\pi \in \mathcal{S}_n$ heißt *Zyklus* (der Länge k , wenn die Menge $\{1, 2, \dots, n\}$ in zwei disjunkte Teilmengen zerfällt

$$\{1, 2, \dots, n\} = \{m_1, m_2, \dots, m_k\} \dot{\cup} \{\ell_1, \ell_2, \dots, \ell_{n-k}\},$$

so dass

$$\begin{aligned} \pi(m_i) &= m_{i+1} \quad \text{für } i = 1, \dots, k-1, & \pi(m_k) &= m_1 \\ \pi(\ell_i) &= \ell_i \quad \text{für } i = 1, \dots, n-k. \end{aligned}$$

2. Die Zahlen m_1, m_2, \dots, m_k heißen die *Elemente* des Zyklus. Zwei Zyklen heißen *disjunkt*, wenn sie keine zwei gleichen Elemente haben.
3. Ein Zyklus der Länge 2 heißt *Transposition*.
4. In diesem Zusammenhang wird die identische Permutation als Zyklus der Länge 1 aufgefasst.

Wenn die Zahl n aus dem Kontext bekannt ist, werden hier andere Schreibweise verwendet:

$$\begin{aligned} (13) &:= \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 1 & 4 \end{pmatrix} \quad (\text{Transposition}) \\ (145) = (451) = (514) &:= \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 2 & 3 & 5 & 1 \end{pmatrix} \\ (25374) &:= \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 1 & 5 & 7 & 2 & 3 & 6 & 4 & 8 \end{pmatrix} \\ (1) &:= \text{id.} \end{aligned}$$

Beobachtung: Sind zwei Zyklen disjunkt, so ist ihr Produkt von der Reihenfolge unabhängig:

$$(p_1 p_2, \dots, p_k) \cdot (q_1 q_2, \dots, q_j) = (q_1 q_2, \dots, q_j) \cdot (p_1 p_2, \dots, p_k).$$

Bei der Multiplikation von nicht disjunkten Zyklen muss man „höllisch“ aufpassen. Dies wollen wir aber nur für Transpositionen ausprobieren:

$$(23)(35) = (235).$$

(Teste dabei nacheinander die Wirkung dieser Abbildung auf die einzelnen Zahlen aus $\{1, 2, \dots, n\}$ aus. Beachte, dass Permutationen Abbildungen sind und daher von rechts nach links ausgeführt werden.)

$$(35)(23) = (253).$$

5.3 Zerlegung von Permutationen

Satz 30 (P1 Zerlegung von Permutationen)

- (i) Jede Permutation π , kann als Produkt von paarweise disjunkten Zyklen geschrieben werden. Die Faktoren sind eindeutig.
- (ii) Jeder Zyklus der Länge k kann als Produkt von $k - 1$ Transpositionen geschrieben werden.
- (iii) Folgerung: Jede Permutation π , kann als Produkt von Transpositionen geschrieben werden.

Beispiele:

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 5 & 8 & 2 & 7 & 1 & 6 & 4 & 3 \end{pmatrix} = (1\ 5) \cdot (2\ 8\ 3) \cdot (4\ 7).$$

Beweis Zu (i):

1. Für die gegebene Permutation π bestimme man zunächst die Menge M der Zahlen, die nicht auf sich selbst abgebildet werden.
2. Wähle eine Zahl $j \in M$ und „starte“ den zugehörigen Zyklus

$$j \mapsto \pi(j) \mapsto \pi^2(j) \mapsto \dots \mapsto \pi^k(j) = j,$$

wir haben den Zyklus

$$(j \ \pi(j) \ \pi^2(j) \ \dots \ \pi^{k-1}(j))$$

der Länge k „herausgefiltert“. Der Zyklus endet mit der Zahl j , ohne dass zwischenzeitlich eine der beteiligten Zahlen zweimal aufgetreten wäre. (Warum?)

3. Wähle aus M solange weitere Zahlen aus und konstruiere die zugehörigen Zyklen, bis alle Zahlen aus M erfasst sind.

Zu (ii):

$$(j_1\ j_2 \ \dots \ j_k) = (j_1\ j_2) \cdot (j_2\ j_3) \cdot \dots \cdot (j_{k-2}\ j_{k-1}) \cdot (j_{k-1}\ j_k).$$



5.4 Das Signum einer Permutation

Definition:

Für eine Permutation $\pi \in \mathcal{S}_n$ definieren wir das *Signum* (oder die *Signatur*) durch

$$\sigma(\pi) := \prod_{1 \leq i < j \leq n} \frac{\pi(j) - \pi(i)}{j - i}. \quad (*)$$

Für jede zwei-elementige Teilmenge $\{i, j\} \subseteq \{1, \dots, n\}$ tritt in dem Produkt ein Faktor $\frac{\pi(j) - \pi(i)}{j - i}$. Die Reihenfolge von i und j ist dabei unwesentlich, da eine Vertauschung eine Erweiterung mit -1 in dem Bruch bedeutet.

Beispiele:

- Für $\pi = (1\ 2\ 3)$ ergibt sich

$$\sigma(\pi) = \frac{\pi(2) - \pi(1)}{2 - 1} \cdot \frac{\pi(3) - \pi(1)}{3 - 1} \cdot \frac{\pi(3) - \pi(2)}{3 - 2} = \frac{3 - 2}{2 - 1} \cdot \frac{1 - 2}{3 - 1} \cdot \frac{1 - 3}{3 - 2} = 1.$$

- Für $\pi = (1\ 4) \in \mathcal{S}_4$ gilt

$$\begin{aligned} \sigma(\pi) &= \\ & \frac{\pi(2) - \pi(1)}{2 - 1} \cdot \frac{\pi(3) - \pi(1)}{3 - 1} \cdot \frac{\pi(4) - \pi(1)}{4 - 1} \cdot \frac{\pi(3) - \pi(2)}{3 - 2} \cdot \frac{\pi(4) - \pi(2)}{4 - 2} \cdot \frac{\pi(4) - \pi(3)}{4 - 3} = \\ & \frac{2 - 4}{2 - 1} \cdot \frac{3 - 4}{3 - 1} \cdot \frac{1 - 4}{4 - 1} \cdot \frac{3 - 2}{3 - 2} \cdot \frac{1 - 2}{4 - 2} \cdot \frac{1 - 3}{4 - 3} = \\ & \frac{-2}{1} \cdot \frac{-1}{2} \cdot \frac{-3}{3} \cdot \frac{1}{1} \cdot \frac{-1}{2} \cdot \frac{-2}{1} = -1 \end{aligned}$$

Satz 31 (P2 Eigenschaften des Signums)

(i) Es gilt für eine beliebige Permutation π :

$$\sigma(\pi) \in \{-1, +1\}.$$

(ii) Für zwei Permutationen $\pi_1, \pi_2 \in \mathcal{S}_n$ gilt

$$\sigma(\pi_1 \cdot \pi_2) = \sigma(\pi_1) \cdot \sigma(\pi_2).$$

(iii) Für die identische Permutation gilt $\sigma((1)) = +1$.

(iv) Für eine Permutation π und ihre Umkehr-Permutation π^{-1} gilt

$$\sigma(\pi) = \sigma(\pi^{-1}).$$

(v) Für eine Transposition $(k\ \ell)$ gilt

$$\sigma((k\ \ell)) = -1.$$

Beweis (i) Betrachte die definierende Formel (*). Da π bijektiv ist, gibt es zu jedem Paar (i, j) mit $i < j$ zwei Zahlen $k, \ell \in \{1, \dots, n\}$, so dass $\pi(k) = i$ und $\pi(\ell) = j$. Also gibt es zu jeder Nenner-Kombination $j - i$ die Zählerkombination $\pi(k) - \pi(\ell) = i - j$ oder $\pi(\ell) - \pi(k) = j - i$. Diese Zuordnung ist umkehrbar eindeutig. Also kürzen sich alle Beträge ungleich 1 heraus. Es bleibt „nur das Vorzeichen“ übrig.

(ii) Es gilt:

$$\begin{aligned} \sigma(\pi_1 \cdot \pi_2) &\stackrel{\text{def}}{=} \prod_{1 \leq i < j \leq n} \frac{\pi_1(\pi_2(j)) - \pi_1(\pi_2(i))}{j - i} \\ &= \underbrace{\prod_{1 \leq i < j \leq n} \frac{\pi_1(\pi_2(j)) - \pi_1(\pi_2(i))}{\pi_2(j) - \pi_2(i)}}_{\sigma(\pi_1)} \cdot \underbrace{\prod_{1 \leq i < j \leq n} \frac{\pi_2(j) - \pi_2(i)}{j - i}}_{\sigma(\pi_2)}. \end{aligned}$$

Dass auch die erste geschweifte Unterklammer stimmt, liegt daran, dass man statt über alle Paare $\{i, j\}$ zu multiplizieren, auch über alle Paare $\{\pi_2(i), \pi_2(j)\}$ multiplizieren kann.

(iii)

$$\sigma((1)) = \sigma((1) \cdot (1)) = \sigma((1)) \cdot \sigma((1)) \implies \sigma((1)) = 1.$$

(iv) Es gilt

$$\sigma(\pi) \cdot \sigma(\pi^{-1}) = \sigma(\pi \cdot \pi^{-1}) = \sigma((1)) = +1.$$

(v) Dies erfordert mehrere Schritte:

(1) Lässt eine Permutation $\pi \in \mathcal{S}_n$ die Zahl n fest, $\pi(n) = n$, so gilt

$$\begin{aligned} \sigma(\pi) &= \prod_{1 \leq i < j \leq n} \frac{\pi(j) - \pi(i)}{j - i} = \prod_{1 \leq i < j < n} \frac{\pi(j) - \pi(i)}{j - i} \cdot \prod_{1 \leq i < j = n} \frac{\pi(j) - \pi(i)}{j - i} \\ &= \prod_{1 \leq i < j < n} \frac{n - \pi(i)}{n - i} \cdot \underbrace{\prod_{1 \leq i \leq n-1} \frac{n - \pi(i)}{n - i}}_{=1} \end{aligned}$$

Das zweite Produkt ist gleich 1, da in den Zählern die gleichen Faktoren auftreten wie in den Nennern. Dies zeigt also, dass man bei der Berechnung des Signums nur die Zahlen $1, \dots, n - 1$ berücksichtigen muss.

(2) Durch einen einfachen Induktionsschluss über $n, n - 1, \dots, 3$ sieht man, dass für $\pi = (1\ 2)$ gilt

$$\sigma(\pi) = \prod_{1 \leq i < j \leq 2} \frac{\pi(j) - \pi(i)}{j - i} = \frac{1 - 2}{2 - 1} = -1.$$

(3) Für eine beliebige Transposition $(k\ \ell)$ sei jetzt $\tilde{\pi}$ eine Permutation mit $\tilde{\pi}(1) = k$ und $\tilde{\pi}(2) = \ell$. Dann gilt

$$\tilde{\pi} \cdot (1\ 2) \cdot \tilde{\pi}^{-1} = (k\ \ell).$$

(4) Mit Eigenschaft (iv) des Satzes folgt:

$$\sigma((k\ \ell)) = \sigma(\tilde{\pi} \cdot (1\ 2) \cdot \tilde{\pi}^{-1}) = \sigma(\tilde{\pi}) \cdot \sigma((1\ 2)) \cdot \sigma(\tilde{\pi}^{-1}) = \sigma((1\ 2)) = -1.$$



Folgerung 32 (und Definition) Eine Permutation π heißt gerade, wenn eine der folgenden äquivalenten Bedingungen erfüllt ist:

- (i) Es ist $\sigma(\pi) = +1$.
- (ii) π läßt sich als Produkt einer geraden Anzahl von Transpositionen darstellen.
- (iii) π läßt sich nicht als Produkt einer ungeraden Anzahl von Transpositionen darstellen.

Beweis (i) \implies (ii). Nach Satz P1 kann π als Produkt von Transpositionen geschrieben werden:

$$\pi = \vartheta_1 \cdot \vartheta_2 \cdot \dots \cdot \vartheta_m.$$

Dann folgt mit den Eigenschaften (iii) und (v):

$$+1 = \sigma(\pi) = \sigma(\vartheta_1) \cdot \sigma(\vartheta_2) \cdot \dots \cdot \sigma(\vartheta_m) = (-1)^m,$$

also ist m gerade. Den anderen Implikationen liegen ähnliche Überlegungen zugrunde. \blacklozenge

Folgerung 33 Ein Zykel, dessen Länge $\begin{cases} \text{gerade} \\ \text{ungerade} \end{cases}$ ist, ist $\begin{cases} \text{ungerade} \\ \text{gerade} \end{cases}$.

Folgerung 34 Die Menge der geraden Permutationen

$$\mathcal{A}_n := \{\pi \in \mathcal{S}_n \mid \sigma(\pi) = +1\}$$

bildet eine Untergruppe von \mathcal{S}_n , d.h. das Produkt zweier Permutationen aus \mathcal{A}_n liegt wieder in \mathcal{A}_n .

6 Determinanten

6.1 Definition und Beispiele

Eine lineare Abbildung $\mathbb{K}^n \rightarrow \mathbb{K}^n$ kann mit einer quadratischen $n \times n$ -Matrix A identifiziert werden.

Wir betrachten eine quadratische Matrix

$$A = (a_{jk}) = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \in \mathbb{K}^{n \times n}.$$

1. Wähle aus den n^2 Einträgen a_{jk} der Matrix eine n -Kombination aus, darunter versteht man n Einträge so, dass jede Zeile und jede Spalte genau einmal vorkommt.

Man kann auch sagen, dass zu jedem Zeilenindex $j \in \{1, \dots, n\}$ ein Spaltenindex k — bijektiv — ausgewählt wird. Das bedeutet aber gerade dass die Auswahl durch eine Permutation $\pi \in \mathcal{S}_n$ über $k = \pi(j)$ gegeben ist:

$$a_{1,\pi(1)}, a_{2,\pi(2)}, \dots, a_{n,\pi(n)}.$$

Mit gleicher Berechtigung kann man sagen, dass zu jedem Spaltenindex k ein Zeilenindex j durch die Permutation π^{-1} ausgewählt wird:

$$a_{\pi^{-1}(1),1}, a_{\pi^{-1}(2),2}, \dots, a_{\pi^{-1}(n),n}.$$

2. Die Elemente der n -Kombination werden multipliziert und mit dem Vorzeichen $\sigma(\pi)$ versehen:

$$\sigma(\pi) \cdot a_{1,\pi(1)} \cdot a_{2,\pi(2)} \cdot \dots \cdot a_{n,\pi(n)} = \sigma(\pi^{-1}) \cdot a_{\pi^{-1}(1),1} \cdot a_{\pi^{-1}(2),2} \cdot \dots \cdot a_{\pi^{-1}(n),n}.$$

(Vgl. Satz P2 (iv)).

3. Dann werden alle nach diesem Verfahren herstellbare Produkte aufaddiert. Das heißt, man muss über alle möglichen Permutationen $\pi \in \mathcal{S}_n$ addieren: Es ergibt sich die Gesamtformel

$$\begin{aligned} \det A &:= \sum_{\pi \in \mathcal{S}_n} \sigma(\pi) \cdot \prod_{j=1}^n a_{j,\pi(j)} && (Z) \\ &= \sum_{\pi \in \mathcal{S}_n} \sigma(\pi) \cdot \prod_{j=1}^n a_{\pi^{-1}(j),j} \\ &\stackrel{(*)}{=} \sum_{\pi \in \mathcal{S}_n} \sigma(\pi) \cdot \prod_{j=1}^n a_{\pi(j),j} && (S) \end{aligned}$$

Da es einerlei ist, ob man über alle Permutationen oder über alle inversen Permutationen aus \mathcal{S}_n addiert, ergibt sich die Umformung (*).

Definition: Es sei A eine quadratische Matrix. Die durch die *Leibniz-Formel* (Z) bzw. (S) definierte Zahl heißt *Determinante* der Matrix A . Die zugehörige Abbildung

$$\det : \mathbb{K}^{n \times n} \rightarrow \mathbb{K}$$

heißt *Determinantenfunktion*.

Beispiele:

- Im Fall $n = 1$ ist die Matrix $A = (a)$ einfach eine Zahl. Es gilt

$$\det A = a.$$

- Es sei $n = 2$ und $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$. Dann gilt

$$\begin{aligned} \det A &= \sum_{\pi \in \mathcal{S}_2} \sigma(\pi) \cdot \prod_{j=1}^n a_{j,\pi(j)} = \underbrace{(+1) \cdot a_{11}a_{22}}_{\pi=(1)} + \underbrace{(-1) \cdot a_{12}a_{21}}_{\pi=(12)} \\ &= a_{11} \cdot a_{22} - a_{12} \cdot a_{21}. \end{aligned}$$

- $n = 3$ und $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$. Die Gruppe \mathcal{S}_3 enthält sechs Permutationen, nämlich

$$\mathcal{S}_3 = \{(1), (123), (132), (12)(13), (23)\}.$$

Damit ergibt sich

$$\begin{aligned} \det(A) &= +a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{32}a_{21} \\ &\quad - a_{12}a_{21}a_{33} - a_{13}a_{31}a_{22} - a_{23}a_{32}a_{11}. \end{aligned}$$

Diese Formel lässt sich graphisch durch die *Regel von Sarrus* repräsentieren:

$$\begin{pmatrix} \overset{+}{a_{11}} & \overset{+}{a_{12}} & \overset{+}{a_{13}} & \overset{-}{a_{11}} & \overset{-}{a_{12}} \\ \overset{-}{a_{21}} & \overset{+}{a_{22}} & \overset{-}{a_{23}} & \overset{+}{a_{21}} & \overset{-}{a_{22}} \\ \overset{+}{a_{31}} & \overset{-}{a_{32}} & \overset{+}{a_{33}} & \overset{-}{a_{31}} & \overset{+}{a_{32}} \end{pmatrix}$$

6.2 Eigenschaften

Satz 35 (D1 Eigenschaften) Die Determinantenfunktion hat folgende Eigenschaften:

(i) \det ist linear in jeder **einzelnen** Zeile. Bezeichnen wir eine j -te Zeile mit $\mathbf{a}_j = (a_{j1}, \dots, a_{jn})$, so gilt für jedes $\ell \in \{1, \dots, n\}$:

$$\det \begin{pmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_{\ell-1} \\ \mathbf{a}_\ell + \mathbf{a}'_\ell \\ \mathbf{a}_{\ell+1} \\ \vdots \\ \mathbf{a}_n \end{pmatrix} = \det \begin{pmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_{\ell-1} \\ \mathbf{a}_\ell \\ \mathbf{a}_{\ell+1} \\ \vdots \\ \mathbf{a}_n \end{pmatrix} + \det \begin{pmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_{\ell-1} \\ \mathbf{a}'_\ell \\ \mathbf{a}_{\ell+1} \\ \vdots \\ \mathbf{a}_n \end{pmatrix}$$

$$\det \begin{pmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_{\ell-1} \\ \alpha \cdot \mathbf{a}_\ell \\ \mathbf{a}_{\ell+1} \\ \vdots \\ \mathbf{a}_n \end{pmatrix} = \alpha \cdot \det \begin{pmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_{\ell-1} \\ \mathbf{a}_\ell \\ \mathbf{a}_{\ell+1} \\ \vdots \\ \mathbf{a}_n \end{pmatrix}$$

(ii) Die Determinantenfunktion \det ist alternierend in den Zeilen, d.h. sie wechselt bei der Vertauschung zweier Zeilen das Vorzeichen:

$$\det \begin{pmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_m \\ \vdots \\ \mathbf{a}_\ell \\ \vdots \\ \mathbf{a}_n \end{pmatrix} = -\det \begin{pmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_\ell \\ \vdots \\ \mathbf{a}_m \\ \vdots \\ \mathbf{a}_n \end{pmatrix}$$

(iii) \det ist normiert, das heißt die Determinante der Einheitsmatrix I ist gleich 1

$$\det I = 1.$$

(iv) Ohne Beweis: Umgekehrt legen diese drei Eigenschaften die Determinantenfunktion eindeutig fest.

(v) Die gleichen Regeln gelten analog bzgl. der Spalten.

Beweis Eine Multiplikation über alle Indices $j \in \{1, \dots, n\}$ außer $j = \ell$ schreiben wir mit dem Zeichen $\prod_{j \neq \ell}$.

$$\begin{aligned} \det \begin{pmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_{\ell-1} \\ \mathbf{a}_\ell + \mathbf{a}'_\ell \\ \mathbf{a}_{\ell+1} \\ \vdots \\ \mathbf{a}_n \end{pmatrix} &= \sum_{\pi \in \mathcal{S}_n} \sigma(\pi) \cdot (a_{\ell, \pi(\ell)} + a'_{\ell, \pi(\ell)}) \cdot \prod_{j \neq \ell} a_{j, \pi(j)} \\ &= \sum_{\pi \in \mathcal{S}_n} \sigma(\pi) \cdot a_{\ell, \pi(\ell)} \cdot \prod_{j \neq \ell} a_{j, \pi(j)} + \sum_{\pi \in \mathcal{S}_n} \sigma(\pi) \cdot a'_{\ell, \pi(\ell)} \cdot \prod_{j \neq \ell} a_{j, \pi(j)} \\ &= \det \begin{pmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_{\ell-1} \\ \mathbf{a}_\ell \\ \mathbf{a}_{\ell+1} \\ \vdots \\ \mathbf{a}_n \end{pmatrix} + \det \begin{pmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_{\ell-1} \\ \mathbf{a}'_\ell \\ \mathbf{a}_{\ell+1} \\ \vdots \\ \mathbf{a}_n \end{pmatrix} \end{aligned}$$

$$\begin{aligned} \det \begin{pmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_{\ell-1} \\ \alpha \cdot \mathbf{a}_\ell \\ \mathbf{a}_{\ell+1} \\ \vdots \\ \mathbf{a}_n \end{pmatrix} &= \sum_{\pi \in \mathcal{S}_n} \sigma(\pi) \cdot \alpha a_{\ell, \pi(\ell)} \cdot \prod_{j \neq \ell} a_{j, \pi(j)} \\ &= \alpha \cdot \sum_{\pi \in \mathcal{S}_n} \sigma(\pi) \cdot a_{\ell, \pi(\ell)} \cdot \prod_{j \neq \ell} a_{j, \pi(j)} = \alpha \cdot \det \begin{pmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_{\ell-1} \\ \mathbf{a}_\ell \\ \mathbf{a}_{\ell+1} \\ \vdots \\ \mathbf{a}_n \end{pmatrix} \end{aligned}$$

(ii) Es sei $\vartheta = (k \ell)$ und A' die durch Vertauschung von k - und ℓ -ter Zeile aus A hervorgehende Matrix, d.h.

$$a'_{jk} = a_{\vartheta(j), k}.$$

Dann gilt

$$\det A' = \sum_{\pi \in \mathcal{S}_n} \sigma(\pi) \cdot \prod_{j=1}^n a'_{j, \pi(j)} = \sum_{\pi \in \mathcal{S}_n} \sigma(\pi) \cdot \prod_{j=1}^n a_{\vartheta(j), \pi(j)}$$

$$= \sum_{\pi \in \mathcal{S}_n} (-\sigma(\pi \circ \vartheta)) \cdot \prod_{j=1}^n a_{j, \pi \circ \vartheta(j)} = - \sum_{\pi \in \mathcal{S}_n} \sigma(\pi) \cdot \prod_{j=1}^n a_{j, \pi(j)} = -\det(A).$$

(iii) Es gibt nur eine n -Kombination der Einträge von I , bei der das zugehörige Produkt ungleich Null ist. Dies ist das Produkt, das zur identischen Permutation gehört:

$$a_{11} \cdot a_{22} \cdot \dots \cdot a_{nn} = 1,$$

alle anderen Produkte enthalten mindestens einen Eintrag a_{jk} mit $j \neq k$, also $a_{jk} = 0$. Das bedeutet aber:

$$\det A = \sum_{\pi \in \mathcal{S}_n} \sigma(\pi) \cdot \prod_{j=1}^n a_{j, \pi(j)} = \sum_{\pi=(1)} \sigma(\pi) \cdot \prod_{j=1}^n a_{j, \pi(j)} = 1.$$

(v) Die Beweise erfolgen genau analog zu den Beweisen von (i) und (ii). Man verwendet einfach die Leibniz-Formel (S) anstelle der Formel (Z). \blacklozenge

Folgerung 36 (D2 Weitere Eigenschaften) *Die Determinantenfunktion hat die folgenden weiteren Eigenschaften:*

(i) Für jedes $\alpha \in \mathbb{K}$ gilt:

$$\det(\alpha \cdot A) = \alpha^n \cdot \det A.$$

(ii) Ist eine Zeile oder Spalte von A gleich Null, so gilt:

$$\det A = 0.$$

(iii) Hat A zwei gleiche Zeilen oder zwei gleich Spalten, so gilt:

$$\det A = 0.$$

(iv) Entsteht die Matrix A' aus der Matrix A dadurch, dass eine Zeile (Spalte) von A zu einer anderen Zeile (Spalte) von A addiert wird, so gilt:

$$\det A' = \det A.$$

(Vgl. LAL01, Abschnitt 4.4, S. 63)

Beweis In der Übung. \blacklozenge

6.3 Der Multiplikationssatz

Satz 37 (D3: Multiplikation von Determinanten) *Es seien A, B zwei $n \times n$ -Matrizen. Dann gilt*

$$\det(A \cdot B) = \det A \cdot \det B.$$

Beispiel: Das Produkt zweier 2×2 -Matrizen:

$$\begin{pmatrix} 2 & -3 \\ 1 & 6 \end{pmatrix} \cdot \begin{pmatrix} 0 & 3 \\ 1 & -2 \end{pmatrix} = \begin{pmatrix} -3 & 12 \\ 6 & -9 \end{pmatrix}$$

Entsprechend gilt für die Determinanten:

$$15 \cdot (-3) = -45.$$

Beweis Der Beweis dieser Beziehung für beliebige n ist mathematisch-elementar, leider aber schreibtechnisch sehr aufwendig. Unter Benutzung der Spalten-Linearität und der Alternierungseigenschaft rechnen wir nach:

$$\begin{aligned} \det(AB) &= \det(A b_1, A b_2, \dots, A b_n) \\ &= \sum_{j_1=1}^n b_{j_1,1} \cdot \det(A e_{j_1}, A b_2, \dots, A b_n) \\ &= \sum_{j_1=1}^n b_{j_1,1} \cdot \sum_{j_2=1}^n b_{j_2,2} \cdot \det(A e_{j_1}, A e_{j_2}, \dots, A b_n) \\ &= (\text{Entwicklung nach den anderen Argumenten}) \\ &= \sum_{j_1=1}^n b_{j_1,1} \cdot \sum_{j_2=1}^n b_{j_2,2} \cdot \dots \cdot \sum_{j_n=1}^n b_{j_n,n} \cdot \det(A e_{j_1}, A e_{j_2}, \dots, A e_{j_n}) \\ &= \sum_{1 \leq j_1, j_2, \dots, j_n \leq n} b_{j_1,1} \cdot b_{j_2,2} \cdot \dots \cdot b_{j_n,n} \cdot \det(A e_{j_1}, A e_{j_2}, \dots, A e_{j_n}) \\ &\stackrel{(*)}{=} \sum_{\pi \in \mathcal{S}_n} b_{\pi(1),1} \cdot b_{\pi(2),2} \cdot \dots \cdot b_{\pi(n),n} \cdot (-1)^{\text{sgn}(\pi)} \det A \\ &= \det B \cdot \det A. \end{aligned}$$

In der vorletzten Zeile bezeichnet $\pi \in \mathcal{S}_n$ die Permutation der Elemente von $\{1, \dots, n\}$, für die $\pi(i) = j_i$ ist. In der dritt-letzten Zeile sind die Summanden, für die ein j_k mit einem j_ℓ , ($k, \ell \in \{1, \dots, n\}$) übereinstimmt, gleich Null, da die Matrix unter der Determinante dann zwei gleiche Spalten enthält. Daraus ergibt sich die Umformung (*). \blacklozenge

6.4 Die Entwicklung einer Determinante

Es sei A eine $n \times n$ -Matrix.

Für $m, \ell \in \{1, \dots, n\}$ sei die Matrix $A^{(m\ell)}$ die $(n-1) \times (n-1)$ -Matrix, die aus A durch Streichung der m -ten Zeile und ℓ -ten Spalte entsteht.

Satz 38 (D4: Entwicklung einer Determinante)

(i) Für eine Zeilennummer m gilt:

$$\det A = \sum_{k=1}^n a_{mk} \cdot (-1)^{m+k} \det A^{(mk)}.$$

(ii) Für eine Spaltennummer ℓ gilt:

$$\det A = \sum_{j=1}^n a_{j\ell} \cdot (-1)^{j+\ell} \det A^{(j\ell)}.$$

Beispiel: Wir entwickeln die Determinante einer 4×4 -Matrix. Dabei benutzen wir eine vereinfachende Schreibweise für Determinanten:

$$\begin{aligned} \begin{vmatrix} 0 & 2 & 3 & 0 \\ 2 & -1 & 0 & -2 \\ 4 & 0 & 0 & -3 \\ 1 & 1 & 2 & 0 \end{vmatrix} &= -2 \cdot \begin{vmatrix} 2 & 0 & -2 \\ 4 & 0 & -3 \\ 1 & 2 & 0 \end{vmatrix} + 3 \cdot \begin{vmatrix} 2 & -1 & -2 \\ 4 & 0 & -3 \\ 1 & 1 & 0 \end{vmatrix} \\ &= (-2) \cdot (-2) \begin{vmatrix} 2 & -2 \\ 4 & -3 \end{vmatrix} + 3 \cdot [1 \cdot 3 - 1] \begin{vmatrix} 2 & -2 \\ 4 & -3 \end{vmatrix} \\ &= 9 + \begin{vmatrix} 2 & -2 \\ 4 & -3 \end{vmatrix} = 9 + 2 = 11. \end{aligned}$$

Beweis Wir zeigen die Aussage (i) für $m = n$. Die allgemeine Aussage folgt dann mit Satz D1 (ii).

$$\begin{aligned} \det A &= \sum_{\pi \in \mathcal{S}_n} \sigma(\pi) \cdot \prod_{j=1}^n a_{j,\pi(j)} = \sum_{k=1}^n \sum_{\pi \in \mathcal{S}_n, \pi(n)=k} \sigma(\pi) \cdot \prod_{j=1}^n a_{j,\pi(j)} \\ &= \sum_{k=1}^n \sum_{\pi \in \mathcal{S}_n, \pi(n)=k} \sigma(\pi) \cdot a_{nk} \cdot \prod_{j=1}^{n-1} a_{j,\pi(j)} \\ &= \sum_{k=1}^n a_{nk} \cdot \sum_{\pi \in \mathcal{S}_n, \pi(n)=k} \sigma(\pi) \cdot \prod_{j=1}^{n-1} a_{j,\pi(j)} \end{aligned}$$

Es bleibt zu zeigen, dass

$$\sum_{\pi \in \mathcal{S}_n, \pi(n)=k} \sigma(\pi) \cdot \prod_{j=1}^{n-1} a_{j,\pi(j)} = (-1)^{n+k} \cdot \det A^{(nk)}.$$

Dazu sei ϑ die Permutation, die die Zahl k an die letzte Stelle n „verschiebt“:

$$\vartheta = (n \ n-1) \dots (k+2 \ k+1)(k+1 \ k), \quad \sigma(\vartheta) = (-1)^{n-k} = (-1)^{n+k}.$$

Dann sind die Einträge von $A^{(nk)}$ gegeben durch

$$a_{j\ell}^{(nk)} = a_{j,\vartheta^{-1}(\ell)}, \quad 1 \leq j, \ell \leq n-1.$$

Es gilt dann:

$$\begin{aligned} \sum_{\pi \in \mathcal{S}_n, \pi(n)=k} \sigma(\pi) \cdot \prod_{j=1}^{n-1} a_{j,\pi(j)} &= \quad (\text{Transformation } \pi = \vartheta^{-1}\tilde{\pi}) \\ \sum_{\tilde{\pi} \in \mathcal{S}_n, \vartheta^{-1}\tilde{\pi}(n)=k} \sigma(\vartheta^{-1}\tilde{\pi}) \cdot \prod_{j=1}^{n-1} a_{j,\vartheta^{-1}\tilde{\pi}(j)} &= \\ (-1)^{n+k} \cdot \sum_{\tilde{\pi} \in \mathcal{S}_n, \tilde{\pi}(n)=n} \sigma(\tilde{\pi}) \cdot \prod_{j=1}^{n-1} a_{j,\vartheta^{-1}\tilde{\pi}(j)} &= \\ (-1)^{n+k} \cdot \sum_{\tilde{\pi} \in \mathcal{S}_n, \tilde{\pi}(n)=n} \sigma(\tilde{\pi}) \cdot \prod_{j=1}^{n-1} a_{j,\tilde{\pi}(j)}^{(nk)} &= \\ (-1)^{n+k} \cdot \sum_{\tilde{\pi} \in \mathcal{S}_{n-1}} \sigma(\tilde{\pi}) \cdot \prod_{j=1}^{n-1} a_{j,\tilde{\pi}(j)}^{(nk)} &= \\ (-1)^{n+k} \cdot \det A^{(nk)}. & \end{aligned}$$



6.5 Einschub: Transposition einer Matrix

Definition: Es sei A eine $m \times n$ -Matrix. Vertauscht man die Rollen von Zeilen und Spalten, so entsteht die $n \times m$ -Matrix $A^T = {}^tA$ mit den Einträgen

$$(A^T)_{jk} = a_{kj}, \quad j = 1, \dots, n, \quad k = 1, \dots, m.$$

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & & a_{2n} \\ \vdots & & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}^T = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & & a_{m2} \\ \vdots & & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{pmatrix}$$

Sie heißt die zu A *transponierte* Matrix.

Eine quadratische Matrix A heißt *symmetrisch*, wenn $A^T = A$, sie heißt *schief-symmetrisch*, wenn $A^T = -A$.

Satz 39 (Eigenschaften der Transposition)

Für das Produkt zweier Matrizen $A \in \mathbb{K}^{m \times n}$ und $B \in \mathbb{K}^{n \times \ell}$ gilt:

$$(A \cdot B)^T = B^T \cdot A^T.$$

Für eine quadratische Matrix A gilt:

$$\det A^T = \det A.$$

Beweis (i) Für die einzelnen Einträge von $A \cdot B$ gilt:

$$((A \cdot B)^T)_{jk} = (A \cdot B)_{kj} = \sum_{i=1}^n (A)_{ki} (B)_{ij} = \sum_{i=1}^n (B^T)_{ji} (A^T)_{ik} = (B^T \cdot A^T)_{jk}.$$

(ii)

$$\det A^T \stackrel{(Z)}{=} \sum_{\pi \in \mathcal{S}_n} \sigma(\pi) \cdot \prod_{j=1}^n (A^T)_{j\pi(j)} = \sum_{\pi \in \mathcal{S}_n} \sigma(\pi) \cdot \prod_{j=1}^n (A)_{\pi(j)j} \stackrel{(S)}{=} \det A.$$



6.6 Invertierbare Matrizen und Determinanten

Satz 40 Zu einer gegebenen quadratischen Matrix A sei die Matrix A^\diamond definiert durch die Einträge

$$a_{jk}^\diamond = (-1)^{j+k} \det A^{(jk)}.$$

Ausgeschrieben ergibt sich die Matrix mit einer Vorzeichen-Schachbrettverteilung:

$$A^\diamond = \begin{pmatrix} + \det A^{(11)} & - \det A^{(12)} & \dots & \dots & (-1)^{1+n} \det A^{(1n)} \\ - \det A^{(12)} & + \det A^{(22)} & \dots & \dots & (-1)^{2+n} \det A^{(2n)} \\ \vdots & \vdots & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \vdots \\ (-1)^{n+1} \det A^{(n1)} & (-1)^{n+2} \det A^{(n2)} & \dots & \dots & + \det A^{(nn)} \end{pmatrix}$$

Dann gilt für das Produkt der Matrizen A und $A^{\diamond T}$

$$A \cdot A^{\diamond T} = A^{\diamond T} \cdot A = \det A \cdot I = \begin{pmatrix} \det A & 0 & \dots & \dots & \dots & 0 \\ 0 & \det A & \ddots & & & \vdots \\ \vdots & \ddots & \det A & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & 0 \\ 0 & \dots & \dots & \dots & 0 & \det A \end{pmatrix}$$

Beweis Wir rechnen einfach die Einträge des Matrixprodukts aus:

$$(A \cdot A^{\diamond T})_{jk} = \sum_{i=1}^n (A)_{ji} (A^{\diamond T})_{ik} = \sum_{i=1}^n a_{ji} \cdot (-1)^{k+i} \det A^{(ki)}$$

(Entwicklung der folgenden Determinante nach der k -ten Zeile)

$$= \det \begin{pmatrix} a_{11} & \dots & \dots & a_{1n} \\ \vdots & & & \vdots \\ a_{k-1,1} & \dots & \dots & a_{k-1,n} \\ a_{j1} & \dots & \dots & a_{jn} \\ a_{k+1,1} & \dots & \dots & a_{k+1,n} \\ \vdots & & & \vdots \\ a_{n1} & \dots & \dots & a_{nn} \end{pmatrix} \longleftarrow k\text{-te Zeile}$$

$$= \begin{cases} \det A, & \text{falls } j = k, \\ 0, & \text{falls } j \neq k. \end{cases}$$

Die Matrix in der vorletzten Zeile entsteht aus der Matrix A dadurch, dass die k -te Zeile von A durch die j -te Zeile von A ersetzt ist. Für $j = k$ ist diese Matrix gleich der Matrix A . Für $i \neq k$ sind in dieser Matrix die Zeilen j und k identisch. Daraus folgt die letzte Gleichung.

Die andere Eigenschaft $A^{\diamond T} \cdot A = \det A \cdot I$ folgt mit einer dualen Betrachtung, d.h. einer Spaltenentwicklung. ◆

Definition Eine quadratische Matrix A heißt *invertierbar* oder *regulär*, wenn es eine Matrix B gibt, so dass $A \cdot B = I$ (Einheitsmatrix). Eine Matrix, die nicht regulär ist, heißt auch *singulär*.

Satz 41 (D5 Matrixgruppe) Die Menge der invertierbaren $n \times n$ -Matrizen über \mathbb{K} bildet eine Gruppe. Sie wird mit $\mathbf{GL}(\mathbb{K}, n)$ (General linear group) bezeichnet. Für $n \geq 2$ ist die Gruppe nicht kommutativ.

Beweis Er ist Ihnen zur Übung überlassen. ◆

Aus der Gruppentheorie ist bekannt, dass inverse Elemente einer Gruppe eindeutig bestimmt ist. Die zu einer gegebenen invertierbaren Matrix A inverse Matrix wird daher mit A^{-1} bezeichnet.

Folgerung 42 (D6 Matrixinversion) Hat eine quadratische Matrix A eine nicht-verschwindende Determinante

$$\det A \neq 0,$$

so ist sie invertierbar. Es gilt

$$A^{-1} = \frac{1}{\det A} \cdot A^{\circ T}.$$

Folgerung 43 (D7 Cramer'sche Regel) Ist die Matrix $A \in \mathbb{K}^{n \times n}$ in dem LGS

$$Ax = b$$

invertierbar, so ist die Lösung x gegeben durch

$$x_j = \frac{1}{\det A} \cdot \det \begin{pmatrix} a_{11} & \cdots & a_{1,j-1} & b_1 & a_{1,j+1} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,j-1} & b_n & a_{n,j+1} & \cdots & a_{nn} \end{pmatrix}, \quad j = 1, \dots, n.$$

Beweis Es ist

$$\frac{1}{\det A} \cdot \det \begin{pmatrix} a_{11} & \cdots & a_{1,j-1} & b_1 & a_{1,j+1} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,j-1} & b_n & a_{n,j+1} & \cdots & a_{nn} \end{pmatrix}$$

(Entwicklung nach der j -ten Spalte)

$$\begin{aligned} &= \frac{1}{\det A} \cdot \sum_{i=1}^n b_i \cdot (-1)^{i+j} \det A^{(ij)} \\ &= \frac{1}{\det A} \cdot \sum_{i=1}^n (A^{\circ T})_{ji} b_i = \frac{1}{\det A} \cdot (A^{\circ T} b)_j = (A^{-1} b)_j = x_j. \end{aligned}$$

◆

Satz 44 (D8 Charakterisierung regulärer Matrizen) Die folgenden Aussagen über eine $n \times n$ -Matrix A sind äquivalent:

- (i) Der Rang von A ist gleich n (= Zeilenrang = Spaltenrang = Vollrang).
- (ii) Das Gleichungssystem $A \cdot x = b$ hat für jedes $b \in \mathbb{K}^n$ genau eine Lösung $x \in \mathbb{K}^n$.
- (iii) A ist invertierbar (= regulär = nicht-singulär).
- (iv) Es ist $\det A \neq 0$.

Beweis Die Implikation (i) \implies (ii) ist durch Satz LA 7 (Vorlesung LAL01) gegeben.

Zu (ii) \implies (iii): Existenz und Eindeutigkeit einer Lösung ermöglichen die Definition einer Abbildung

$$B \begin{cases} \mathbb{K}^n & \rightarrow & \mathbb{K}^n \\ b & \mapsto & x. \end{cases}$$

Es gilt dann $A \cdot Bb = A \cdot x = b$ und $B \cdot Ax = Bb = x$, also ist die Abbildung B die Umkehrabbildung zur Abbildung A . Die Abbildung ist linear:

$$B(\alpha_1 b_1 + \alpha_2 b_2) = \alpha_1 x_1 + \alpha_2 x_2,$$

da der Linearkombination zweier rechten Seiten eines LGS die entsprechende Linearkombination der Lösungen gehört. Also kann B durch eine Matrix repräsentiert werden, die dann die Inverse zu A ist.

(iii) \implies (iv): Wenn A invertierbar ist, so gilt

$$\det A \cdot \det(A^{-1}) = \det(A \cdot A^{-1}) = \det I = 1,$$

also kann nicht $\det A = 0$ sein.

(iv) \implies (i): Wir nehmen an, dass A nicht Vollrang hat. Dann ist eine — sagen wir die k -te — Spalte eine Linearkombination der anderen Spalten. Zieht man diese Linearkombination von der k -ten Spalte ab, so entsteht eine Matrix A' mit lauter Nullen in der k -ten Spalte. Dann gilt aber mit Satz D2 (iv) und (ii)

$$\det A = \det A' = 0$$

im Widerspruch zu (iv). ◆

6.7 Beispiel

Als Beispiel betrachten wir die Matrix

$$A = \begin{pmatrix} 5 & 2 & -1 \\ 0 & 1 & -4 \\ -3 & 4 & 2 \end{pmatrix}$$

Ihre Determinante ist gemäß Regel von Sarrus

$$\det A = 10 + 24 + 0 - 3 + 80 - 0 = 111.$$

Die Matrizen A^\diamond , $A^{\diamond T}$ und A^{-1} sind dann

$$A^\diamond = \begin{pmatrix} 18 & 12 & 3 \\ -8 & 7 & -26 \\ -7 & 20 & 5 \end{pmatrix}, \quad A^{\diamond T} = \begin{pmatrix} 18 & -8 & -7 \\ 12 & 7 & 20 \\ 3 & -26 & 5 \end{pmatrix}, \quad A^{-1} = \begin{pmatrix} \frac{18}{111} & -\frac{8}{111} & -\frac{7}{111} \\ \frac{12}{111} & \frac{7}{111} & \frac{20}{111} \\ \frac{3}{111} & -\frac{26}{111} & \frac{5}{111} \end{pmatrix},$$

Es empfiehlt sich, dies drei Schritte trotz erhöhten Schreibaufwands tatsächlich durchzuführen. Eine Abkürzung ist mit deutlich erhöhter Fehlergefahr verbunden.

Die inverse Matrix A kann man auch mit Hilfe des Gauß-Algorithmus auffinden. Dazu schreibt man rechts neben die gegebene Matrix A die Einheitsmatrix gleicher Dimension und fasst diese als eine Zusammenstellung von drei rechten Seiten eines inhomogenen LGS auf. Der Gauß-Algorithmus besteht nun darin, Zeilenumformungen für die dreifach-erweiterte Matrix $(A|I)$ so auszuführen, dass am Ende die Matrix $(I|A^{-1})$ steht.

$$\begin{array}{ccc|ccc|c}
 5 & 2 & -1 & 1 & 0 & 0 & \\
 0 & 1 & -4 & 0 & 1 & 0 & \\
 -3 & 4 & 2 & 0 & 0 & 1 & +\frac{3}{5} \text{ I} \\
 \hline
 5 & 2 & -1 & 1 & 0 & 0 & \\
 0 & 1 & -4 & 0 & 1 & 0 & \\
 0 & \frac{26}{5} & \frac{7}{5} & \frac{3}{5} & 0 & 1 & 5 \cdot \\
 \hline
 5 & 2 & -1 & 1 & 0 & 0 & \\
 0 & 1 & -4 & 0 & 1 & 0 & \\
 0 & 26 & 7 & 3 & 0 & 5 & -26 \text{ II} \\
 \hline
 5 & 2 & -1 & 1 & 0 & 0 & \\
 0 & 1 & -4 & 0 & 1 & 0 & \\
 0 & 0 & 111 & 3 & -26 & 5 & \frac{1}{111} \cdot \\
 \hline
 5 & 2 & -1 & 1 & 0 & 0 & + \text{ III} \\
 0 & 1 & -4 & 0 & 1 & 0 & +4 \text{ II} \\
 0 & 0 & 1 & \frac{3}{111} & -\frac{26}{111} & \frac{5}{111} & \\
 \hline
 5 & 2 & 0 & \frac{114}{111} & -\frac{26}{111} & \frac{5}{111} & -2 \text{ II} \\
 0 & 1 & 0 & \frac{12}{111} & \frac{7}{111} & \frac{20}{111} & \\
 0 & 0 & 1 & \frac{3}{111} & -\frac{26}{111} & \frac{5}{111} & \\
 \hline
 5 & 0 & 0 & \frac{90}{111} & -\frac{40}{111} & -\frac{35}{111} & \frac{1}{5} \cdot \\
 0 & 1 & 0 & \frac{12}{111} & \frac{7}{111} & \frac{20}{111} & \\
 0 & 0 & 1 & \frac{3}{111} & -\frac{26}{111} & \frac{5}{111} & \\
 \hline
 1 & 0 & 0 & \frac{18}{111} & -\frac{8}{111} & -\frac{7}{111} & \\
 0 & 1 & 0 & \frac{12}{111} & \frac{7}{111} & \frac{20}{111} & \\
 0 & 0 & 1 & \frac{3}{111} & -\frac{26}{111} & \frac{5}{111} &
 \end{array}$$

7 Euklidische Vektorräume

7.1 Grundlegende Definitionen

Definition Es sei V ein reeller Vektorraum. Wir betrachten eine Abbildung

$$\langle \cdot, \cdot \rangle : \begin{cases} V \times V & \rightarrow \mathbb{R} \\ (v, w) & \mapsto \langle v, w \rangle \end{cases}$$

mit zwei Argumenten in V und Werten in \mathbb{R} . Es gibt auch andere Bezeichnungen anstelle von $\langle \cdot, \cdot \rangle$:

$$\langle v, w \rangle = \langle v|w \rangle = (v|w) = v \cdot w.$$

(Die letzte Version ist in der Schule üblich, aber sehr ungünstig, weil sie die Gültigkeit eines Assoziativgesetzes suggeriert: $(v \cdot w) \cdot u \stackrel{\text{falsch}}{=} v \cdot (w \cdot u)$. Das ist aber Quatsch!)

- (i) Die Abbildung $\langle \cdot, \cdot \rangle$ heißt *bilinear* oder eine *Bilinearform*, wenn sie in jedem einzelnen Argument linear ist, d.h. wenn für alle $v, v_1, v_2 \in V$ und $\alpha_1, \alpha_2 \in \mathbb{R}$ gilt

$$\begin{aligned} \langle \alpha_1 v_1 + \alpha_2 v_2, v \rangle &= \alpha_1 \langle v_1, v \rangle + \alpha_2 \langle v_2, v \rangle \\ \langle v, \alpha_1 v_1 + \alpha_2 v_2 \rangle &= \alpha_1 \langle v, v_1 \rangle + \alpha_2 \langle v, v_2 \rangle. \end{aligned}$$

Anders ausgedrückt: Für festes $v \in V$ sind die Abbildungen

$$\langle \cdot, v \rangle : V \rightarrow \mathbb{R} \quad \text{und} \quad \langle v, \cdot \rangle : V \rightarrow \mathbb{R}$$

linear.

- (ii) Die Abbildung $\langle \cdot, \cdot \rangle$ heißt *symmetrisch*, wenn für alle $v, w \in V$ gilt:

$$\langle v, w \rangle = \langle w, v \rangle.$$

- (iii) $\langle \cdot, \cdot \rangle$ heißt *positiv definit*, wenn für alle $v \in V \setminus \{\vec{0}\}$ gilt:

$$\langle v, v \rangle > 0.$$

- (iv) Die Abbildung $\langle \cdot, \cdot \rangle$ heißt ein *Skalarprodukt* oder ein *Inneres Produkt*, wenn sie bilinear, symmetrisch und positiv definit ist.

- (v) Ein endlich-dimensionaler reeller Vektorraum V , auf dem ein Skalarprodukt $\langle \cdot, \cdot \rangle$ definiert ist, heißt ein *euklidischer* oder *Euklid'scher Vektorraum*. Man schreibt auch $(V, \langle \cdot, \cdot \rangle)$.

Beispiele: Wir betrachten den Vektorraum $V = \mathbb{R}^n$ mit der Standardbasis $\{e_1, e_2, \dots, e_n\}$. Ein Vektor $v \in V$ kann eindeutig als $v = v_1 e_1 + v_2 e_2 + \dots + v_n e_n$ geschrieben werden. Wir definieren das *kanonische* oder *Standard-Skalarprodukt* auf V durch

$$\langle v, w \rangle := v_1 w_1 + v_2 w_2 + \dots + v_n w_n = v^T \cdot w.$$

Probiere das für zwei Vektoren aus \mathbb{R}^4 aus.

Satz 45 (und Beispielklasse EU1)

Es sei $\mathbf{B} = (b_{jk})$ eine $n \times n$ -Matrix. Durch

$$\langle v, w \rangle_{\mathbf{B}} := v^T \cdot \mathbf{B} \cdot w = \sum_{j,k=1}^n v_j b_{jk} w_k$$

ist eine Bilinearform auf V definiert. Ist umgekehrt eine Bilinearform $\langle \cdot, \cdot \rangle$ auf V gegeben, so wird durch

$$b_{jk} = \langle e_j, e_k \rangle$$

eine Matrix \mathbf{B} definiert.

Die so gegebene Korrespondenz zwischen Bilinearformen und Matrizen ist umkehrbar eindeutig.

Die in diesem Zusammenhang gegebene Matrix \mathbf{B} heißt auch *Gram'sche Matrix* oder *Strukturmatrix*.

Beispiele

1. Zum Standard-Skalarprodukt gehört im obigen Sinne die Einheitsmatrix $\mathbf{B} = I = (\delta_{jk})$. (Erinnern Sie sich an die Definition des Kronecker-Symbols in Abschnitt 4.3 LAL01.)

$$\langle v, w \rangle_I = \sum_{j,k=1}^n v_j \delta_{jk} w_k = \sum_{j=1}^n v_j w_j.$$

2. (Staatsexamen F00 T2, A3a) Auf dem \mathbb{R}^3 sei die Bilinearform

$$(x, y) := x_1 y_1 + 3x_2 y_2 + 4x_3 y_3 + x_1 y_2 + x_2 y_1 + x_1 y_3 + x_3 y_1 + x_2 y_3 + x_3 y_2$$

gegeben. Zeigen Sie, dass diese Bilinearform ein inneres Produkt auf dem \mathbb{R}^3 definiert.

Bei einer Vertauschung von x und y gilt:

$$(y, x) := y_1 x_1 + 3y_2 x_2 + 4y_3 x_3 + y_1 x_2 + y_2 x_1 + y_1 x_3 + y_3 x_1 + y_2 x_3 + y_3 x_2$$

Dieser Ausdruck ist offenbar mit dem obigen identisch.

Die Linearität im ersten Argument ist einfach nachzurechnen:

$$\begin{aligned} (\alpha x, y) &= \\ \alpha x_1 y_1 + 3\alpha x_2 y_2 + 4\alpha x_3 y_3 + \alpha x_1 y_2 + \alpha x_2 y_1 + \alpha x_1 y_3 + \alpha x_3 y_1 + \alpha x_2 y_3 + \alpha x_3 y_2 &= \\ \alpha(x, y) \end{aligned}$$

$$\begin{aligned} (x + z, y) &= \\ (x_1 + z_1)y_1 + 3(x_2 + z_2)y_2 + 4(x_3 + z_3)y_3 + (x_1 + z_1)y_2 + (x_2 + z_2)y_1 + \\ &\quad (x_1 + z_1)y_3 + (x_3 + z_3)y_1 + (x_2 + z_2)y_3 + (x_3 + z_3)y_2 = \\ x_1 y_1 + z_1 y_1 + 3x_2 y_2 + 3z_2 y_2 + 4x_3 y_3 + 4z_3 y_3 + x_1 y_2 + z_1 y_2 + x_2 y_1 + z_2 y_1 + \\ &\quad x_1 y_3 + z_1 y_3 + x_3 y_1 + z_3 y_1 + x_2 y_3 + z_2 y_3 + x_3 y_2 + z_3 y_2 = \\ (x, y) + (z, y). \end{aligned}$$

Bezüglich der Linearität im zweiten Argument genügt es, auf die bereits bewiesene Symmetrie zu verweisen.

Zum Schluss muss die Positiv-Definitheit gezeigt werden:

$$\begin{aligned}(x, x) &= x_1x_1 + 3x_2x_2 + 4x_3x_3 + x_1x_2 + x_2x_1 + x_1x_3 + x_3x_1 + x_2x_3 + x_3x_2 \\ &= (x_1 + x_2 + x_3)^2 + 2x_2^2 + 3x_3^2.\end{aligned}$$

Dieser Ausdruck ist niemals negativ. Ist er gleich Null, so kann man in der folgenden Reihenfolge schließen:

$$x_2 = 0 \quad \text{und} \quad x_3 = 0 \quad \implies \quad x_1 = 0.$$

Satz 46

(i) Die Bilinearform ist symmetrisch genau dann, wenn die zugehörige Gram'sche Matrix \mathbf{B} symmetrisch ist: $\mathbf{B}^T = \mathbf{B}$.

(ii) (Ohne Beweis) Die Bilinearform ist positiv definit genau dann, wenn die n Bedingungen

$$\begin{vmatrix} b_{11} & \cdots & b_{1m} \\ \vdots & & \vdots \\ b_{m1} & \cdots & b_{mm} \end{vmatrix} > 0 \quad m = 1, \dots, n$$

erfüllt sind.

Beweis Ist \mathbf{B} symmetrisch, so gilt:

$$\langle v, w \rangle_{\mathbf{B}} = v^T \cdot \mathbf{B} \cdot w = v^T \cdot \mathbf{B}^T \cdot w = (w^T \cdot \mathbf{B} \cdot v)^T = w^T \cdot \mathbf{B} \cdot v = \langle w, v \rangle_{\mathbf{B}}.$$

Ist umgekehrt die Bilinearform $\langle \cdot, \cdot \rangle_{\mathbf{B}}$ symmetrisch, so gilt:

$$b_{jk} = \langle e_j, e_k \rangle = \langle e_k, e_j \rangle = b_{kj}.$$



7.2 Norm und Orthogonalität

Definition Es sei $(V, \langle \cdot, \cdot \rangle)$ ein euklidischer Vektorraum.

(i) Die Abbildung

$$\|\cdot\| \begin{cases} V & \rightarrow \mathbb{R}_0^+ \\ v & \mapsto \|v\| := \sqrt{\langle v, v \rangle} \end{cases}$$

heißt (*euklidische*) *Norm* oder (*euklidische*) *Länge* auf $(V, \langle \cdot, \cdot \rangle)$.

(ii) Ein Vektor $v \in V$ heißt *normiert* oder *Einheitsvektor*, wenn er „Einheitslänge“ hat:
 $\|v\| = 1$.

(iii) Zwei einzelne Vektoren $u, w \in V$ heißen *orthogonal*, wenn $\langle u, w \rangle = 0$.

(iv) Zwei Teilmengen U, W heißen *orthogonal*, wenn gilt:

$$\langle u, w \rangle = 0 \quad \text{für alle } u \in U, w \in W.$$

(v) Ist $W \subseteq V$ eine Teilmenge von V , so heißt die Menge

$$W^\perp := \{v \in V \mid \langle v, w \rangle = 0 \text{ für alle } w \in W\}$$

das orthogonale *Komplement* von W .

(vi) Eine Teilmenge $W \subseteq V$ heißt *orthonormiert* oder *orthonormal*, wenn die Vektoren in W normiert und paarweise orthogonal sind:

$$\begin{aligned} \|w\| &= 1 && \text{für alle } w \in W, \\ \langle w_1, w_2 \rangle &= 0 && \text{für alle } w_1, w_2 \in W \text{ mit } w_1 \neq w_2. \end{aligned}$$

(vii) Eine Basis W von V heißt *Orthonormalbasis* (ONB), wenn sie orthonormiert ist.

Satz 47 (EU 2: Eigenschaften der euklidischen Norm) Für $v, w \in V$ und $\alpha \in \mathbb{R}$ sind die folgenden Eigenschaften erfüllt.

$$(i) \quad \|v\| = 0 \implies v = \vec{0}.$$

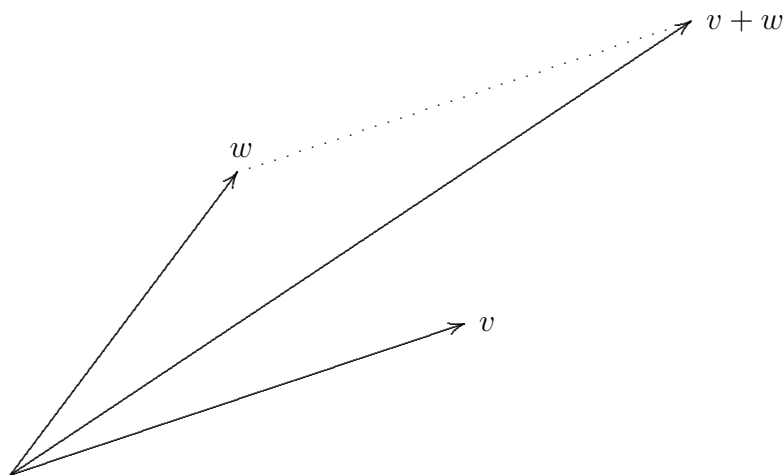
$$(ii) \quad \|\alpha \cdot v\| = |\alpha| \cdot \|v\|,$$

$$(iii) \quad |\langle v, w \rangle| \leq \|v\| \cdot \|w\| \quad (\text{Ungleichung von Cauchy-Schwarz (CSU)})$$

Dabei tritt Gleichheit genau dann ein, wenn v und w linear abhängig sind.

$$(iv) \quad \|v + w\| \leq \|v\| + \|w\| \quad (\text{Dreiecksungleichung})$$

$$|\|v\| - \|w\|| \leq \|v - w\|$$



Bemerkungen

1. Eine Abbildung $\|\cdot\|$ auf einem beliebigen (evtl. nicht mit einem Skalarprodukt versehenen) Vektorraum heißt (allgemeine) *Norm*, wenn sie die Eigenschaften (i),(ii) und (iv) hat.
2. Gemäß CSU existiert für $v, w \in V \setminus \{\vec{0}\}$ eine Zahl $c \in \mathbb{R}$ mit $|c| \leq 1$, so dass

$$\langle v, w \rangle = \|v\| \cdot \|w\| \cdot c.$$

In der Analysis mit komplexen Zahlen kann man den *Winkel* $\gamma := \sphericalangle(v, w)$ zwischen zwei Vektoren definieren. Es stellt sich dann heraus, dass die Zahl c gleich dem Cosinus des Zwischenwinkels ist:

$$\langle v, w \rangle = \|v\| \cdot \|w\| \cdot \cos \sphericalangle(v, w).$$

Beweis

(i) folgt direkt aus der Positiv-Definitheit des Skalarprodukts.

(ii) Es ist

$$\|\alpha \cdot v\| = \sqrt{\langle \alpha \cdot v, \alpha \cdot v \rangle} = \sqrt{\alpha \cdot \langle v, \alpha \cdot v \rangle} = \sqrt{\alpha^2 \cdot \langle v, v \rangle} = |\alpha| \cdot \sqrt{\langle v, v \rangle} = |\alpha| \cdot \|v\|.$$

(iii) Wir gehen in mehreren Schritten vor.

(1) Für $w = \vec{0}$ ist die Aussage des Satzes trivial.

(2) Für beliebiges $\alpha \in \mathbb{R}$ gilt:

$$\begin{aligned} 0 &\leq \langle v - \alpha w, v - \alpha w \rangle = \langle v, v \rangle - \alpha \langle v, w \rangle - \alpha \langle w, v \rangle + \alpha^2 \langle w, w \rangle \\ &= \langle v, v \rangle - 2\alpha \langle v, w \rangle + \alpha^2 \langle w, w \rangle. \end{aligned}$$

(3) Für $w \neq \vec{0}$ setzen wir in (2) $\alpha := \frac{\langle v, w \rangle}{\langle w, w \rangle}$. Dann lautet die obige Ungleichung:

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

Die Multiplikation dieser Ungleichung mit $\langle w, w \rangle > 0$ liefert

$$0 \leq \langle v, v \rangle \cdot \langle w, w \rangle - \langle v, w \rangle^2$$

und dann

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

Zieht man daraus die Wurzel, so hat man die CSU.

(4) Gleichheit gilt in der letzten Ungleichung genau dann, wenn Gleichheit in (2) gilt, wenn also

$$\vec{0} = v - \alpha w = v - \frac{\langle v, w \rangle}{\langle w, w \rangle} w.$$

Das bedeutet aber gerade lineare Abhängigkeit der Vektoren v und w .

(iv) Ein typischer Trick beim Umgang mit Ungleichungen bei Bilinearformen ist, zum Quadrat überzugehen:

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

Die andere Ungleichung wird in der Übung bewiesen. ◆

Satz 48 (EU 3: Orthonormale Mengen und Orthonormalbasen)

- (i) Eine orthonormale Menge $W = \{w_1, w_2, \dots, w_m\}$ ist linear unabhängig.
(ii) (Orthonormalisierungsverfahren von Erhard Schmidt) Ist $W = \{w_1, w_2, \dots, w_m\}$ linear unabhängig, so kann man eine orthonormale Menge $U = \{u_1, u_2, \dots, u_m\}$ finden, so dass

$$\text{span } U = \text{span } W.$$

- (iii) Jede orthonormale Teilmenge $W = \{w_1, w_2, \dots, w_m\}$ eines euklidischen Vektorraums kann durch eine Menge $W' = \{w_{n-m+1}, \dots, w_n\}$ von $n - m$ Vektoren zu einer ONB von V ergänzt werden. Es gilt dann (vgl. Definition in Abschnitt 3.5 LAL01):

$$(\text{span } W)^\perp = \text{span } W' \quad \text{und} \quad \text{span } W + \text{span } W' = V.$$

- (iv) Jeder euklidische Vektorraum hat eine ONB.

Beweis (i) Wir stellen den Nullvektor als Linearkombination von $\{w_1, w_2, \dots, w_m\}$ dar.

$$\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_m v_m = \vec{0} \quad (*)$$

Es sei jetzt $j \in \{1, \dots, n\}$ beliebig. Wir bilden auf beiden Seiten der Gleichung (*) das Skalarprodukt mit v_j .

$$\alpha_1 \langle v_1, v_j \rangle + \alpha_2 \langle v_2, v_j \rangle + \dots + \alpha_m \langle v_m, v_j \rangle = \langle \vec{0}, v_j \rangle = 0.$$

Da die Vektoren paarweise orthogonal sind, gilt $\langle v_k, v_j \rangle = 0$ für alle $k \neq j$. Daraus folgt

$$\alpha_j \langle v_j, v_j \rangle = 0 \quad \implies \quad \alpha_j = 0.$$

j war beliebig, es ist also $\alpha_j = 0$ für alle j . Also ist W linear unabhängig.

- (ii) Der Fall $m = 1$ ist einfach: Man setzt

$$u_1 = \frac{w_1}{\|w_1\|}.$$

Wir nehmen jetzt als Induktionsvoraussetzung an, dass bereits eine orthonormale Menge $\{u_1, \dots, u_k\}$, $k < m$, mit

$$\text{span } \{u_1, \dots, u_k\} = \text{span } \{w_1, \dots, w_k\}$$

konstruiert ist. Wir setzen jetzt

$$u'_{k+1} := w_{k+1} - \langle u_1, w_{k+1} \rangle u_1 - \dots - \langle u_k, w_{k+1} \rangle u_k.$$

Damit ist für $\ell = 1, \dots, k$

$$\begin{aligned} \langle u'_{k+1}, u_\ell \rangle &= \langle w_{k+1} - \langle u_1, w_{k+1} \rangle u_1 - \dots - \langle u_k, w_{k+1} \rangle u_k, u_\ell \rangle \\ &= \langle w_{k+1}, u_\ell \rangle - \langle u_1, w_{k+1} \rangle \cdot \langle u_1, u_\ell \rangle - \dots - \langle u_k, w_{k+1} \rangle \cdot \langle u_k, u_\ell \rangle \\ &= \langle w_{k+1}, u_\ell \rangle - \langle u_k, w_{k+1} \rangle = 0. \end{aligned}$$

Es ist also u_{k+1} orthogonal zu $\{u_1, \dots, u_k\}$. Setzen wir jetzt noch

$$u_{k+1} := \frac{u'_{k+1}}{\|u'_{k+1}\|},$$

so ist $\{u_1, \dots, u_{k+1}\}$ orthonormal. Nach (i) ist $\{u_1, \dots, u_{k+1}\}$ linear unabhängig und deshalb:

$$\text{span}\{u_1, \dots, u_{k+1}\} = \text{span}\{w_1, \dots, w_{k+1}\}.$$

(iii) Man ergänze zunächst die Menge W zu einer Basis von V (Basisergänzungssatz) und wende dann das obige Verfahren auf diese Basis von V an.

(iv) ist in (iii) enthalten. ◆

Bei der konkreten Durchführung des Verfahrens genügt es natürlich, schrittweise die Vektoren u'_{k+1} und dann u_{k+1} zu definieren. Die Orthonormalität von $\{u_1, \dots, u_{k+1}\}$ ist durch den abstrakten Beweis sichergestellt.

7.3 Das Vektorprodukt im euklidischen \mathbb{R}^3

Wir betrachten den euklidischen Vektorraum $V = \mathbb{R}^3$ mit der kanonischen Basis $\{e_1, e_2, e_3\}$, die bezüglich des kanonischen Skalarprodukts eine ONB ist.

Definition Wir definieren auf V eine zweistellige Verknüpfung, das *Vektorprodukt*

$$\times : \begin{cases} V \times V & \rightarrow V \\ (v, w) & \mapsto v \times w \end{cases}$$

durch

$$v \times w := \begin{pmatrix} v_2 w_3 - v_3 w_2 \\ v_3 w_1 - v_1 w_3 \\ v_1 w_2 - v_2 w_1 \end{pmatrix}$$

Anders als beim Skalarprodukt ist das Ergebnis wieder ein Vektor. Beachte, dass das Vektorprodukt nicht assoziativ ist:

$$\begin{aligned} (e_1 \times e_2) \times e_2 &= e_3 \times e_2 = -e_1 \\ e_1 \times (e_2 \times e_2) &= e_1 \times \vec{0} = \vec{0} \end{aligned}$$

Satz 49 (EU 4: Eigenschaften des Vektorprodukts)

(i) Das Vektorprodukt ist bilinear, d.h. für alle $v, \tilde{v}, w, \tilde{w} \in V$ und $\alpha, \beta \in \mathbb{R}$ gilt

$$\begin{aligned} (\alpha v + \beta \tilde{v}) \times w &= \alpha v \times w + \beta \tilde{v} \times w \\ v \times (\alpha w + \beta \tilde{w}) &= \alpha v \times w + \beta v \times \tilde{w}. \end{aligned}$$

(ii) Das Vektorprodukt ist schiefsymmetrisch, d.h. für alle $v, w \in V$ gilt

$$w \times v = -v \times w.$$

(iii) Für $u, v, w \in V$ gilt

$$\langle u \times v, w \rangle = \det(u, v, w).$$

(iv) Für $v, w \in V$ gilt:

$$\|v \times w\|^2 = \|v\|^2 \|w\|^2 - \langle v, w \rangle^2.$$

(v) v und w sind linear abhängig genau dann, wenn $v \times w = \vec{0}$.

Beweis (i) – (iv) sind einfach nachzurechnen. (v) zeigt man mit Hilfe von (iv) und der Cauchy–Schwarz–Ungleichung. \blacklozenge

Bemerkung Aus der Gleichung (v) und mit Hilfe der Bemerkung 2 im Anschluss an Satz EU 2 folgt, dass

$$\begin{aligned} \|v \times w\|^2 &= \|v\|^2 \|w\|^2 - \langle v, w \rangle^2 = \|v\|^2 \|w\|^2 \cdot [1 - \cos^2 \angle(v, w)] \\ &= \|v\|^2 \|w\|^2 \cdot \sin^2 \angle(v, w), \end{aligned}$$

und damit weiter

$$\|v \times w\| = \|v\| \|w\| \cdot |\sin \angle(v, w)|.$$

Aufgrund der Betragsstriche beim Sinus braucht man sich hier nicht um Probleme der Orientierung bei der Definition des Winkels zu kümmern.

8 Diagonalisierbarkeit

Gegeben sei die Matrix

$$A = \begin{pmatrix} \frac{16}{3} & \frac{10}{3} \\ -5 & -3 \end{pmatrix}$$

Fragen:

- Kann man auf einfache Weise A^2 , A^5 oder Polynome ausrechnen?
- Kann man aus A die (zweite, dritte, n -te) Wurzel ziehen? Wenn ja, wie lautet sie?
- Kann man die Matrix A in Potenzreihen (Exponentialfunktion, Sinusfunktion, ...) einsetzen und den Wert berechnen?

Diese Fragen kann man mit Hilfe der Diagonalisierung von Matrizen beantworten.

8.1 Vektoren beim Basiswechsel

Es sei V ein n -dimensionaler Vektorraum über dem Körper \mathbb{K} mit Basis $B = \{w_1, \dots, w_n\}$. Es kann eine bijektive lineare Abbildung

$$\eta_B : V \rightarrow \mathbb{K}^n$$

auf die folgende Weise definiert werden: Man entwickelt $v \in V$ bzgl. der Basis

$$v = \alpha_1 w_1 + \dots + \alpha_n w_n$$

und hat dann die bijektive Zuordnung

$$v \begin{array}{c} \xrightarrow{\eta_B} \\ \xleftarrow{\eta_B^{-1}} \end{array} v_B = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}.$$

Wir nennen $v_B = \eta_B(v)$ die B -Entwicklung oder B -Darstellung (von v bzgl. der Basis B).

Quintessenz: Alles, was man in V oder mit V machen kann, kann man stattdessen auch in \mathbb{K}^n oder mit dem \mathbb{K}^n machen. Der Vorteil besteht darin, dass man im \mathbb{K}^n , also mit Zahlen, rechnen kann. Von Nachteil ist, dass diese Berechnungen immer von der gerade gewählten Basis abhängen.

Wir betrachten jetzt für einen fixierten Vektorraum V zwei Basen $B = \{w_1, \dots, w_n\}$ und $B' = \{w'_1, \dots, w'_n\}$ mit den zugehörigen Abbildungen η_B und $\eta_{B'}$

$$\begin{array}{ccc} \mathbb{K}^n & \xrightarrow{\eta_B^{-1}} & V \\ \Theta_{B'B} \downarrow & & \swarrow \eta_{B'} \\ \mathbb{K}^n & & \end{array} \quad \begin{array}{ccc} v_B & \xrightarrow{\eta_B^{-1}} & v \\ \Theta_{B'B} \downarrow & & \swarrow \eta_{B'} \\ v_{B'} & & \end{array}$$

In ihnen ist die bijektive lineare Abbildung $\Theta_{B'B}$ definiert durch

$$\Theta_{B'B} := \eta_{B'} \circ \eta_B^{-1} : \mathbb{K}^n \rightarrow \mathbb{K}^n$$

Sie kann als quadratische invertierbare Matrix aufgefasst werden. Sie heißt *Übergangsmatrix (transition matrix)* für den Basiswechsel $B' \leftarrow B$.

Satz 50 (ED1, über Übergangsmatrizen)

(i) Für zwei Basen B, B' sind die beiden Übergangsmatrizen invers zueinander:

$$\Theta_{B'B} \cdot \Theta_{BB'} = I.$$

(ii) Für drei Basen B, B', B'' gilt:

$$\Theta_{B''B} = \Theta_{B''B'} \cdot \Theta_{B'B}.$$

(iii) Hat ein Vektor $v \in V$ bzgl. zweier Basen B und B' die Entwicklungen v_B bzw. $v_{B'}$, so gilt

$$v_{B'} = \Theta_{B'B} v_B \quad \text{bzw.} \quad v_B = \Theta_{B'B}^{-1} v_{B'}.$$

(iv) In der j -ten Spalte von

$$\Theta_{B'B} = \begin{pmatrix} \vartheta_{11} & \cdots & \vartheta_{1j} & \cdots & \vartheta_{1n} \\ \vdots & & \vdots & & \vdots \\ \vartheta_{n1} & \cdots & \vartheta_{nj} & \cdots & \vartheta_{nn} \end{pmatrix}$$

stehen die Entwicklungskoeffizienten des B -Basisvektors w_j bzgl. der Basis B'

$$w_j = \vartheta_{1j} w'_1 + \cdots + \vartheta_{nj} w'_n.$$

(v) In der j -ten Spalte von $\Theta_{B'B}^{-1}$ stehen die Entwicklungskoeffizienten des B' -Basisvektors w'_j bzgl. der Basis B , das heißt

$$w'_j = (\Theta_{B'B}^{-1})_{1j} w_1 + \cdots + (\Theta_{B'B}^{-1})_{nj} w_n.$$

Beweis Die ersten beiden Aussagen ergeben sich aus der Definition der Übergangsmatrix. Für (iv) schauen wir uns das folgende Diagramm an:

$$\sum_{k=1}^n (\Theta_{B'B})_{kj} w'_k \xrightarrow{\eta_{B'}} \Theta_{B'B}(e_j) \xrightarrow{\Theta_{B'B}^{-1}} e_j \xrightarrow{\eta_B^{-1}} w_j.$$

Wegen $\Theta_{B'B}^{-1} = \eta_B \circ \eta_{B'}^{-1}$ ist die gesamte Abbildung die Identität, es besteht also zwischen den Ausdrücken ganz links und ganz rechts Gleichheit.

(v) ergibt sich entsprechend aus einem Diagramm. ◆

8.2 Homomorphismen beim Basiswechsel, Äquivalenz von Matrizen

Sind nun zwei Vektorräume V und U mit $\dim V = n$, $\dim U = m$, und ein Homomorphismus (lineare Abbildung) $f : V \rightarrow U$ gegeben, so kann man das obige Spiel anwenden, um eine Matrixdarstellung für f zu finden. Ist $B = \{w_1, \dots, w_n\}$ eine Basis für V und $C = \{u_1, \dots, u_m\}$ eine Basis für U , so gibt das folgende Diagramm die Abbildungssituation wieder:

$$\mathbb{K}^n \begin{array}{c} \xrightarrow{\eta_B^{-1}} \\ \xleftarrow{\eta_B} \end{array} V \xrightarrow{f} U \begin{array}{c} \xleftarrow{\eta_C} \\ \xrightarrow{\eta_C^{-1}} \end{array} \mathbb{K}^m$$

Die gesamte Abbildung

$$f_{CB} := \eta_C \circ f \circ \eta_B^{-1} : \mathbb{K}^n \rightarrow \mathbb{K}^m$$

ist linear, sie kann also durch eine $m \times n$ -Matrix $f_{CB} \in \mathbb{K}^{m \times n}$ dargestellt werden. Beachte, dass diese *Matrixdarstellung* von den Basen C und B abhängt.

In der j -ten Spalte von f_{CB} stehen die Entwicklungskoeffizienten von $f(w_j)$, $w_j \in B$, bezüglich der Basis C .

Zur Begründung muss man diese Behauptung etwas abstrakter aufschreiben und die Definition von f_{CB} einsetzen:

$$\begin{pmatrix} (f_{CB})_{1j} \\ \vdots \\ (f_{CB})_{mj} \end{pmatrix} = f_{CB} e_j \stackrel{\text{def}}{=} \eta_C \circ f \circ \underbrace{\eta_B^{-1}(e_j)}_{w_j} = \eta_C(f(w_j))$$

Es seien jetzt jeweils zwei Basen B, B' für V und C, C' für U gegeben. Wir stellen die Frage, was mit der Matrix f_{CB} bei einem Basiswechsel passiert. Die Situation ergibt sich aus dem folgenden Diagramm:

$$\begin{array}{ccccc} \mathbb{K}^n & \xrightarrow{f_{CB}} & & & \mathbb{K}^m \\ & \searrow \eta_B^{-1} & & & \swarrow \eta_C^{-1} \\ \Theta_{B'B} \downarrow & & V & \xrightarrow{f} & U & & \downarrow \Theta_{C'C} \\ & \swarrow \eta_{B'} & & & \searrow \eta_{C'} & & \\ \mathbb{K}^n & \xrightarrow{f_{C'B'}} & & & \mathbb{K}^m \end{array}$$

Man kann ablesen, dass gilt:

$$f_{C'B'} = \Theta_{C'C} \cdot f_{CB} \cdot \Theta_{BB'} = \Theta_{C'C} \cdot f_{CB} \cdot \Theta_{B'B}^{-1}.$$

Es gibt also verschiedene Matrizen für die gleiche lineare Abbildung. Gehören zwei Matrizen zu der gleichen linearen Abbildung, so wollen wir dies mit einer Namensgebung hervorheben:

Definition Zwei Matrizen $F, F' \in \mathbb{K}^{m \times n}$ heißen *äquivalent* zueinander, wenn es eine reguläre $n \times n$ -Matrix T und eine reguläre $m \times m$ -Matrix S gibt, so dass

$$F' = S \cdot F \cdot T^{-1}.$$

Es stellt sich jetzt die Frage, ob man mit Hilfe einer geeigneten Wahl der Basen für V und U erreichen kann, dass die Matrix f_{CB} eine möglichst einfache Gestalt annimmt.

Satz 51 *Zwei $m \times n$ -Matrizen F und F' sind genau dann äquivalent, wenn sie den gleichen Rang haben. Insbesondere ist jede Matrix vom Rang r äquivalent zu der Matrix*

$$\begin{pmatrix} I_{r \times r} & 0_{r \times (n-r)} \\ 0_{(m-r) \times r} & 0_{(m-r) \times (n-r)} \end{pmatrix}$$

mit vier Untermatrizen. Der Index gibt jeweils die Dimension der Einheitsmatrix I oder der Nullmatrix 0 an.

Beweis Er besteht im wesentlichen im Gauß'schen Algorithmus, er wird nicht genau ausgeführt. \blacklozenge

Definition Der *Rang* einer linearen Abbildung $f : V \rightarrow W$ ist per definitionem gleich dem Rang einer Matrix f_{BC} , die dieser linearen Abbildung durch Basen B von V bzw. C von W zugeordnet ist.

8.3 Beispiel: Orthogonale Polynome

Es sei V_n der $(n + 1)$ -dimensionale \mathbb{R} -Vektorraum der Polynome vom Grad $\leq n$. Wir beschäftigen uns im folgenden mit V_3 .

Wir wissen bereits, dass die Polynome

$$\begin{aligned} e_0(x) &= 1 \\ e_1(x) &= x \\ e_2(x) &= x^2 \\ e_3(x) &= x^3 \end{aligned}$$

Basen des Vektorraums V_3 bilden.

Die folgenden Polynome sind Beispiele für *orthogonale* Polynome, die in weiten Teilen der Analysis, numerischen Mathematik und mathematischen Physik eine wichtige Rolle spielen.

Hermite-Polynome:

$$\begin{aligned} h_0(x) &= 1 \\ h_1(x) &= 2x \\ h_2(x) &= 4x^2 - 2 \\ h_3(x) &= 8x^3 - 12x \end{aligned}$$

Legendre-Polynome:

$$\begin{aligned} j_0(x) &= 1 \\ j_1(x) &= x \\ j_2(x) &= \frac{3}{2}x^2 - \frac{1}{2} \\ j_3(x) &= \frac{5}{2}x^3 - \frac{3}{2}x \end{aligned}$$

Laguerre-Polynome:

$$\begin{aligned} \ell_0(x) &= 1 \\ \ell_1(x) &= -x + 1 \\ \ell_2(x) &= \frac{1}{2}(x^2 - 4x + 2) \\ \ell_3(x) &= \frac{1}{6}(-x^3 + 9x^2 - 18x + 6) \end{aligned}$$

Wir bezeichnen die Basen von V_3 , die aus diesen jeweils vier Polynomen gebildet werden, mit E, H, J bzw. L .

A Berechnen Sie $\eta_B(y)$ für $B = E, H, J, L$ und $y = \dots$

$$\begin{aligned} p(x) &= 4 \\ q(x) &= x + 2 \\ r(x) &= x^2 - 3x + 2 \end{aligned}$$

Lösung:

$$\begin{aligned} \eta_E(p) &= \eta_H(p) = \eta_J(p) = \eta_L(p) = \begin{pmatrix} 4 \\ 0 \\ 0 \\ 0 \end{pmatrix} \\ \eta_E(q) &= \begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad \eta_H(q) = \begin{pmatrix} 2 \\ \frac{1}{2} \\ 0 \\ 0 \end{pmatrix}, \quad \eta_J(q) = \begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad \eta_L(q) = \begin{pmatrix} 3 \\ -1 \\ 0 \\ 0 \end{pmatrix} \\ \eta_E(r) &= \begin{pmatrix} 2 \\ -3 \\ 1 \\ 0 \end{pmatrix}, \quad \eta_H(r) = \begin{pmatrix} \frac{5}{2} \\ \frac{3}{2} \\ \frac{1}{4} \\ 0 \end{pmatrix}, \quad \eta_J(r) = \begin{pmatrix} \frac{7}{3} \\ -3 \\ \frac{2}{3} \\ 0 \end{pmatrix}, \quad \eta_L(r) = \begin{pmatrix} 1 \\ -1 \\ 2 \\ 0 \end{pmatrix} \end{aligned}$$

B Berechnen Sie die Übergangsmatrizen Θ_{EH} , Θ_{EJ} , Θ_{EL} durch Entwicklung der Basisvektoren von H nach denen von E usw.

Lösung:

$$\begin{aligned} \Theta_{EH} &= \begin{pmatrix} 1 & 0 & -2 & 0 \\ 0 & 2 & 0 & -12 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 8 \end{pmatrix} \\ \Theta_{EJ} &= \begin{pmatrix} 1 & 0 & -\frac{1}{2} & 0 \\ 0 & 1 & 0 & -\frac{3}{2} \\ 0 & 0 & \frac{3}{2} & 0 \\ 0 & 0 & 0 & \frac{5}{2} \end{pmatrix} \\ \Theta_{EL} &= \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & -1 & -2 & -3 \\ 0 & 0 & \frac{1}{2} & \frac{3}{2} \\ 0 & 0 & 0 & -\frac{1}{6} \end{pmatrix} \end{aligned}$$

C Berechnen Sie die Übergangsmatrizen Θ_{HE} , Θ_{JE} , Θ_{LE} durch

- Entwicklung der Basisvektoren von E nach denen von H usw. oder
- durch Invertierung der Matrizen aus **B**.

Lösung:

$$\begin{aligned}\Theta_{HE} &= \Theta_{EH}^{-1} = \begin{pmatrix} 1 & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{3}{4} \\ 0 & 0 & \frac{1}{4} & 0 \\ 0 & 0 & 0 & \frac{1}{8} \end{pmatrix} \\ \Theta_{JE} &= \Theta_{EJ}^{-1} = \begin{pmatrix} 1 & 0 & \frac{1}{3} & 0 \\ 0 & 1 & 0 & \frac{3}{5} \\ 0 & 0 & \frac{2}{3} & 0 \\ 0 & 0 & 0 & \frac{2}{5} \end{pmatrix} \\ \Theta_{LE} &= \Theta_{EL}^{-1} = \begin{pmatrix} 1 & 1 & 2 & 6 \\ 0 & -1 & -4 & -18 \\ 0 & 0 & 2 & 18 \\ 0 & 0 & 0 & -6 \end{pmatrix}\end{aligned}$$

D Überprüfen Sie die Gleichung $\eta_{B'}(y) = \Theta_{B'B}\eta_B(y)$ für verschiedene Basen $B, B' = E, H, J, L$ und Polynome y aus **A**.

Lösungsbeispiel:

$$\begin{aligned}\Theta_{LJ} &= \Theta_{LE} \cdot \Theta_{EJ} = \begin{pmatrix} 1 & 1 & 2 & 6 \\ 0 & -1 & -4 & -18 \\ 0 & 0 & 2 & 18 \\ 0 & 0 & 0 & -6 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & -\frac{1}{2} & 0 \\ 0 & 1 & 0 & -\frac{3}{2} \\ 0 & 0 & \frac{3}{2} & 0 \\ 0 & 0 & 0 & \frac{5}{2} \end{pmatrix} \\ &= \begin{pmatrix} 1 & 1 & \frac{5}{2} & \frac{27}{2} \\ 0 & -1 & -6 & -\frac{37}{2} \\ 0 & 0 & 3 & 45 \\ 0 & 0 & 0 & -15 \end{pmatrix}\end{aligned}$$

E Die Differentiation D kann als lineare Abbildung

$$D : \begin{cases} V_3 & \rightarrow V_3 \\ \alpha_0 + \alpha_1 x + \alpha_2 x^2 + \alpha_3 x^3 & \mapsto \alpha_1 + 2\alpha_2 x + 3\alpha_3 x^2 \end{cases}$$

aufgefasst werden. Überlegen Sie, dass bezüglich der kanonischen Basis E für V_3 dann gilt

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

Berechnen Sie D_{BB} oder $D_{B'B}$ für andere Basen $B, B' = E, H, J, L$.

Lösung:

$$\begin{aligned} D_{HH} &= \Theta_{HE} D_{EE} \Theta_{EH} \\ &= \begin{pmatrix} 1 & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{3}{4} \\ 0 & 0 & \frac{1}{4} & 0 \\ 0 & 0 & 0 & \frac{1}{8} \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & -2 & 0 \\ 0 & 2 & 0 & -12 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 8 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 2 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 6 \\ 0 & 0 & 0 & 0 \end{pmatrix} \end{aligned}$$

Testen Sie diese Matrix für die Polynome r und $Dr = r' = 2x - 3$.

Lösung: Es gilt

$$\eta_H(r) = \begin{pmatrix} \frac{5}{2} \\ -\frac{3}{2} \\ \frac{1}{4} \\ 0 \end{pmatrix} \quad \text{und} \quad \eta_H(r') = \begin{pmatrix} -3 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

Tatsächlich gilt

$$\eta_H(r') = D_{HH} \cdot \eta_H(r).$$

F Die Multiplikation X mit einer Variablen kann als lineare Abbildung

$$X : \begin{cases} V_2 & \rightarrow V_3 \\ \alpha_0 + \alpha_1 x + \alpha_2 x^2 & \mapsto \alpha_0 x + \alpha_1 x^2 + \alpha_2 x^3 \end{cases}$$

aufgefasst werden. Bezüglich der kanonischen Basen E' für V_2 und E für V_3 gilt dann

$$X_{EE'} = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Berechnen Sie $X_{BB'}$ für andere Basen $B = H, J, L$.

Lösung:

$$X_{JJ'} = \Theta_{JE} X_{EE'} \Theta_{E'J'}$$

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

Testen Sie diese Matrix für die Polynome r und $Xr = x^3 - 3x^2 + 2x$.

Lösung: Es gilt

$$\eta_{J'}(r) = \begin{pmatrix} \frac{7}{3} \\ -3 \\ \frac{2}{3} \end{pmatrix} \quad \text{und} \quad \eta_J(Xr) = \begin{pmatrix} -1 \\ \frac{13}{5} \\ -2 \\ \frac{2}{5} \end{pmatrix}$$

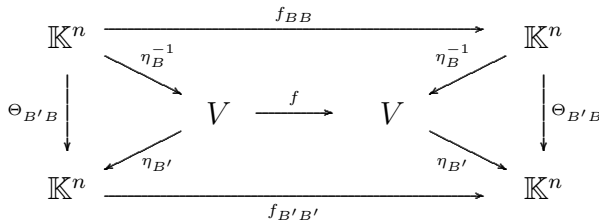
Tatsächlich gilt

$$\eta_J(Xr) = X_{JJ'} \cdot \eta_{J'}(r).$$

8.4 Endomorphismen beim Basiswechsel, Ähnlichkeit von Matrizen

Es sei V ein n -dimensionaler Vektorraum mit zwei Basen B und B' . Wir betrachten einen Endomorphismus (lineare Abbildung **innerhalb eines** Vektorraums) $f : V \rightarrow V$. Die obigen Überlegungen wollen wir jetzt unter dem sinnvollen, aber einengenden, Gesichtspunkt betrachten, dass die Basiswechsel im Definitionsraum und im Bildraum gleich sind. Der letzte Satz kann also nicht angewandt werden.

Das obige Diagramm hat jetzt die Form:



Bei vorgegebener Basis $B = \{w_1, \dots, w_n\}$ wird der Abbildung f die quadratische Matrix $f_B = f_{BB}$ dadurch (eindeutig) zugeordnet, dass

$$f(w_j) = \sum_{k=1}^n (f_B)_{jk} w_k$$

gilt.

Bei einem durch die Matrix $\Theta_{B'B}$ vermittelten Basiswechsel von B nach B' geht die Matrix f_{BB} in die Matrix F über, wobei gilt:

$$f_{B'B'} = \Theta_{B'B} \cdot f_B \cdot \Theta_{BB'} = \Theta_{B'B} \cdot f_B \cdot \Theta_{B'B}^{-1}.$$

Dieser Verwandtschaft von zwei Matrizen wird wieder ein Namen gegeben:

Definition (und Satz) Zwei Matrizen $A, A' \in \mathbb{K}^{n \times n}$ heißen *ähnlich* oder *konjugiert* zueinander, symbolisch $A \sim A'$, wenn eine der folgenden äquivalenten Bedingungen erfüllt ist:

- (1) Es gibt eine reguläre Matrix $T \in \mathbb{K}^{n \times n}$, so dass

$$A' = T \cdot A \cdot T^{-1}.$$

- (2) Es gibt eine reguläre Matrix $T \in \mathbb{K}^{n \times n}$, so dass

$$A' = T^{-1} \cdot A \cdot T.$$

- (3) Ist $f : V \rightarrow V$ ein Endomorphismus eines n -dimensionalen Vektorraums und $A = f_B$ die Matrixdarstellung von f bezüglich einer Basis B von V , so gibt es eine Basis B' von V , so dass $A' = f_{B'}$.

Satz 52 (ED 2, Eigenschaften ähnlicher Matrizen, Teil 1)

- (i) Jede Matrix A ist zu sich selbst ähnlich (Reflexivität).
 (ii) Gilt $A \sim A'$, so gilt auch $A' \sim A$ (Symmetrie).
 (iii) Es gilt für drei Matrizen A, A', A'' die Transitivität

$$\text{WENN } A \sim A' \text{ UND } A' \sim A'', \quad \text{DANN } A \sim A''.$$

- (iv) Zwei ähnliche Matrizen A und A' haben die gleiche Determinante:

$$A \sim A' \quad \implies \quad \det A = \det A'$$

- (v) Zwei ähnliche Matrizen A und A' haben den gleichen Rang:

$$A \sim A' \quad \implies \quad \text{Rang } A = \text{Rang } A'$$

- (vi) Ist eine Matrix ähnlich zu einem Vielfachen der Einheitsmatrix, so ist sie gleich diesem Vielfachen der Einheitsmatrix.

$$A \sim \alpha \cdot I \quad \implies \quad A = \alpha I.$$

(Wir schreiben das α -fache der Einheitsmatrix einfach als α , da die Wirkung auf Vektoren bzw. andere Matrizen die gleiche ist.)

Beweis (i) und (ii) sind einfach.

Zu (iii) Es seien $A \sim A'$ und $A' \sim A''$, es existieren also reguläre Matrizen S, T mit

$$A' = SAS^{-1} \quad \text{und} \quad A'' = TA'T^{-1}.$$

Dann gilt aber

$$A'' = TA'T^{-1} = TSAS^{-1}T^{-1} = (TS)A(TS)^{-1}.$$

Zu (iv):

$$\det A' = \det(TAT^{-1}) = \det T \cdot \det A \cdot \det T^{-1} = \det T \cdot \det A \cdot (\det T)^{-1} = \det A.$$

(v): Die Multiplikation mit T von links verändert nicht die Anzahl der linear unabhängigen Spaltenvektoren. Die Multiplikation mit T^{-1} von rechts verändert nicht die Anzahl der linear unabhängigen Zeilenvektoren.

(vi): Aus der Ähnlichkeit folgt so die Gleichheit

$$A = T(\alpha I)T^{-1} = \alpha TIT^{-1} = \alpha I.$$



Definition Ist $f : V \rightarrow V$ eine lineare Abbildung und f_B ihre Matrixdarstellung bezüglich einer Basis B , so heißt

$$\det f = \det f_B$$

die *Determinante der linearen Abbildung* f (Beachte, dass wir bisher nur Determinanten von Matrizen gekannt haben).

Beispiel (Siehe Anfang des Kapitels) In einem zweidimensionalen Vektorraum V sei eine lineare Abbildung $f : V \rightarrow V$ bzgl. einer Basis $B = \{w_1, w_2\}$ durch

$$f_B = \begin{pmatrix} \frac{16}{3} & \frac{10}{3} \\ -5 & -3 \end{pmatrix}$$

gegeben. Gibt es eine lineare Abbildung $g : V \rightarrow V$, für die $g^2 = g \circ g = f$ gilt?

Anders gefragt: Gibt es eine Matrix g_B mit $g_B^2 = f_B$?

Wählen wir als andere (woher? später!) Basis $B' = \{w'_1, w'_2\}$ von V mittels

$$w'_1 = w_1 - w_2, \quad w'_2 = -2w_1 + 3w_2,$$

so stellt sich bei Berechnung der Bilder heraus:

$$f(w'_1) = f(w_1 - w_2) = \left(\frac{16}{3}w_1 - 5w_2\right) - \left(\frac{10}{3}w_1 - 3w_2\right) = 2w_1 - 2w_2 = 2w'_1$$

$$\begin{aligned} f(w'_2) &= f(-2w_1 + 3w_2) = -2\left(\frac{16}{3}w_1 - 5w_2\right) + 3\left(\frac{10}{3}w_1 - 3w_2\right) = \\ &= -\frac{2}{3}w_1 + w_2 = \frac{1}{3}(-2w_1 + 3w_2) = \frac{1}{3}w'_2 \end{aligned}$$

Bezüglich der neuen Basis wird der Abbildung f also die Matrix

$$f_{B'} = \begin{pmatrix} 2 & 0 \\ 0 & \frac{1}{3} \end{pmatrix}$$

zugeordnet. Der Basiswechsel $B \rightarrow B'$ ist auch durch die Matrizen

$$\Theta_{B'B}^{-1} = \begin{pmatrix} 1 & -2 \\ -1 & 3 \end{pmatrix}, \quad \Theta_{B'B} = \begin{pmatrix} 3 & 2 \\ 1 & 1 \end{pmatrix}$$

gegeben. Man kann testen, dass

$$f_{B'} = \Theta_{B'B} \cdot f_B \cdot \Theta_{B'B}^{-1}.$$

Jetzt kann die Frage nach der Matrixwurzel beantwortet werden: Für die Matrix

$$g_B = \Theta_{B'B}^{-1} \cdot \begin{pmatrix} \sqrt{2} & 0 \\ 0 & \frac{1}{\sqrt{3}} \end{pmatrix} \cdot \Theta_{B'B} = \begin{pmatrix} 3\sqrt{2} - \frac{2}{\sqrt{3}} & 2\sqrt{2} - \frac{2}{\sqrt{3}} \\ -3\sqrt{2} + \sqrt{3} & -2\sqrt{2} + \sqrt{3} \end{pmatrix}$$

gilt:

$$\begin{aligned} g_B \cdot g_B &= \Theta_{B'B}^{-1} \cdot \begin{pmatrix} \sqrt{2} & 0 \\ 0 & \frac{1}{\sqrt{3}} \end{pmatrix} \cdot \Theta_{B'B} \cdot \Theta_{B'B}^{-1} \cdot \begin{pmatrix} \sqrt{2} & 0 \\ 0 & \frac{1}{\sqrt{3}} \end{pmatrix} \cdot \Theta_{B'B} \\ &= \Theta_{B'B}^{-1} \cdot \begin{pmatrix} 2 & 0 \\ 0 & \frac{1}{3} \end{pmatrix} \cdot \Theta_{B'B} = f_B. \end{aligned}$$

Der springende Punkt ist, dass die neue Matrix eine Diagonalmatrix ist. Diagonalmatrizen sind viel einfacher zu lesen, zu interpretieren, mit ihnen kann man wesentlich besser rechnen. Beispielsweise kann man das der Abbildung f^2 zugeordnete Matrix-Quadrat f_B^2 oder die „Wurzel“ einer Matrix viel leichter ausrechnen.

Wie kann man Basen finden, bzgl. derer eine gegebene lineare Abbildung Diagonalgestalt annimmt?

8.5 Exkurs: Polynome

Es sei \mathbb{K} ein Körper (wir denken vor allem an $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$).

Definition

- (1) Eine Abbildung $p : \mathbb{K} \rightarrow \mathbb{K}$ heißt *Polynom (über \mathbb{K})*, wenn es Zahlen $p_n, p_{n-1}, \dots, p_1, p_0 \in \mathbb{K}$ gibt, so dass

$$p(x) = p_n x^n + p_{n-1} x^{n-1} + \dots + p_1 x + p_0.$$

- (2) Ist $p_n \neq 0$, so hat das Polynom, per Definition, den *Grad n* , symbolisch

$$\text{grad } p = n.$$

Der Grad des Null-Polynoms wird als $-\infty$ festgelegt.

- (3) Die Menge der Polynome über \mathbb{K} wird mit $\mathbb{K}[x]$ bezeichnet. (Das ist ein bisschen seltsam, da der Buchstabe x eigentlich ohne Bedeutung ist.)
- (4) Eine Zahl $\lambda \in \mathbb{K}$ heißt *Nullstelle* des Polynoms p , wenn $p(\lambda) = 0$.

Die Polynome höchstens n -ten Grades bilden einen \mathbb{K} -Vektorraum der Dimension $n + 1$. Polynome können multipliziert und nacheinander ausgeführt werden.

Satz 53 (P1 Nullstellen und Linearfaktoren) *Die beiden folgenden Aussagen über ein Polynom $p \in \mathbb{K}[x]$ n -ten Grades und eine Zahl $\lambda \in \mathbb{K}$ sind äquivalent:*

- (i) λ ist eine Nullstelle von p .
- (ii) Es gibt ein Polynom q vom Grad $n - 1$, so dass

$$p(x) = (x - \lambda) \cdot q(x).$$

Man nennt $x - \lambda$ einen *Linearfaktor* von p . Man spricht hier auch von der *Abspaltung eines Linearfaktors*.

Beweis Die Richtung (ii) \implies (i) ist trivial. Für die Umkehrung machen wir eine Vorüberlegung: Es gilt

$$\begin{aligned} (x^n - \lambda^n) &= \left[x^n + x^{n-1}\lambda + x^{n-2}\lambda^2 + \dots + x^2\lambda^{n-2} + x\lambda^{n-1} \right] \\ &\quad - \left[x^{n-1}\lambda + x^{n-2}\lambda^2 + \dots + x^2\lambda^{n-2} + x\lambda^{n-1} + \lambda^n \right] \\ &= (x - \lambda) \cdot \underbrace{(x^{n-1} + x^{n-2}\lambda + \dots + x\lambda^{n-2} + \lambda^{n-1})}_{=: Q_{n-1}(x, \lambda)} \end{aligned}$$

Jetzt können wir die Umkehrung zeigen:

$$\begin{aligned}
 p(x) &= p(x) - p(\lambda) \\
 &= \left(p_n x^n + p_{n-1} x^{n-1} + \dots + p_1 x + p_0 \right) - \left(p_n \lambda^n + p_{n-1} \lambda^{n-1} + \dots + p_1 \lambda + p_0 \right) \\
 &= p_n (x^n - \lambda^n) + p_{n-1} (x^{n-1} - \lambda^{n-1}) + \dots + p_1 (x - \lambda) \\
 &= (x - \lambda) \cdot \left(p_n Q_{n-1}(x, \lambda) + p_{n-1} Q_{n-2}(x, \lambda) + \dots + p_1 \right) \\
 &= (x - \lambda) \cdot q(x).
 \end{aligned}$$

◆

Die fortgesetzte Abspaltung von Linearfaktoren führt auf die

Folgerung 54 (P2) Zu einem Polynom p mit Nullstelle λ gibt es eine eindeutig festgelegte Zahl $k \in \mathbb{N}$ und ein Polynom q vom Grad $n - k$, so dass

$$p(x) = (x - \lambda)^k \cdot q(x), \quad \text{wobei } q(\lambda) \neq 0.$$

Die Zahl k heißt die Vielfachheit der Nullstelle λ im Polynom p .

Definition Man sagt, ein Polynom zerfällt (über \mathbb{K}) in Linearfaktoren, wenn es in der Form

$$p(x) = p_n \cdot (x - \lambda_1)^{k_1} \cdot (x - \lambda_2)^{k_2} \cdot \dots \cdot (x - \lambda_m)^{k_m}$$

mit Zahlen $\lambda_1, \dots, \lambda_m \in \mathbb{K}$ und $k_1, \dots, k_m \in \mathbb{N}$ dargestellt werden kann.

Beispiele Das Polynom $x^2 + 1$ zerfällt über \mathbb{R} nicht in Linearfaktoren, über \mathbb{C} schon:

$$x^2 + 1 = (x - i) \cdot (x + i).$$

Satz 55 (P3 Fundamentalsatz der Algebra)

Jedes Polynom über \mathbb{C} besitzt eine Nullstelle in \mathbb{C} .

Folgerung: Ein komplexes Polynom zerfällt (über \mathbb{C}) in Linearfaktoren.

Beweis Er muss mit Mitteln der Analysis geführt werden. Er ist allerdings für den Rahmen einer grundlegenden Vorlesung Analysis zu aufwändig, mit Methoden der Funktionentheorie geht es ganz schnell. Vielleicht haben Sie einmal die Möglichkeit, ihn in einem Seminar kennenzulernen. ◆

8.6 Das charakteristische Polynom

Für eine Zahl $\alpha \in \mathbb{K}$ bezeichnen wir die lineare Abbildung $\alpha \cdot \text{id}_V : V \rightarrow V, v \mapsto \alpha \cdot v$, einfach mit

$$\alpha \simeq \alpha \cdot \text{id}_V = \begin{pmatrix} \alpha & 0 & \cdots & \cdots & 0 \\ 0 & \alpha & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & \alpha \end{pmatrix}.$$

Sie wird ja bzgl. jeder Basis durch die α -fache Einheitsmatrix $\alpha \cdot I_n$ dargestellt.

Satz 56 (ED 3, und Definition) *Ist A eine quadratische Matrix, so ist der Ausdruck*

$$\chi_A(x) := \det(A - x)$$

ein Polynom n -ten Grades in x . Es heißt das charakteristische Polynom der Matrix A . Genauer gilt:

$$\chi_A(x) = (-1)^n x^n + (-1)^{n-1} (a_{11} + a_{22} + \dots + a_{nn}) x^{n-1} + \dots + \det A.$$

Die Summe der Diagonalelemente einer quadratischen Matrix A heißt die Spur der Matrix A .

$$\text{Spur } A := a_{11} + \dots + a_{nn}.$$

Beweis Wir schreiben die Determinante aus

$$\det(A - x) = \begin{vmatrix} a_{11} - x & a_{12} & \cdots & \cdots & a_{1n} \\ a_{21} & a_{22} - x & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & a_{n-1,n} \\ a_{n1} & \cdots & \cdots & a_{n,n-1} & a_{nn} - x \end{vmatrix}.$$

und überlegen, wie die Variable x über die Leibniz-Formel in dem Ausdruck $\chi_A(\lambda)$ Eingang findet. Da in der Leibniz-Formel nur Summen und Produkte der Matrix-Einträge auftreten, ist es klar, dass es sich um ein Polynom handelt. Die Potenzen x^n und x^{n-1} können nur in dem Produkt auftreten, das zu der Permutation $\pi = \text{id}$ gehört. In den Produkten, die zu den anderen Permutationen gehören, kommen mindestens zwei Nicht-Diagonalelemente der Matrix $A - x$, also höchstens $n - 2$ Diagonalelemente, vor. Das heißt aber, dass x in diesen Produkten höchstens mit dem Exponenten $n - 2$ auftritt.

Es genügt also, das zur identischen Permutation gehörende Produkt zu betrachten. Dafür gilt aber

$$\begin{aligned} & (a_{11} - x) \cdot (a_{22} - x) \cdot \dots \cdot (a_{nn} - x) = \\ & (-x)^n + (a_{11} + a_{22} + \dots + a_{nn})(-x)^{n-1} + \text{Glieder niedrigerer Ordnung} \end{aligned}$$

Das sind genau die ersten beiden Glieder im charakteristischen Polynom. Setzt man $x = 0$, so sieht man, dass das konstante Glied im charakteristischen Polynom gerade die Determinante von A ist. \blacklozenge

Definition Ist $f : V \rightarrow V$ eine lineare Abbildung und B eine Basis, so heißt

$$\chi_f(x) := \chi_{f_B}(x)$$

das *charakteristische Polynom der linearen Abbildung*. Die *Spur der linearen Abbildung* ist definitionsgemäß die Spur der Matrix f_B :

$$\text{Spur}(f) = \text{Spur } f_B.$$

Es besteht die Frage, ob diese Begriffe wohldefiniert sind und nicht von „irgendeiner Zufälligkeit“ (hier: der gerade gewählten Basis) abhängen. Die Antwort gibt der folgende

Satz 57 (ED 4: Eigenschaften ähnlicher Matrizen, Teil 2) *Sind zwei Matrizen A und A' ähnlich zueinander, so stimmen ihre charakteristischen Polynome überein:*

$$\chi_A(x) = \chi_{A'}(x).$$

Es stimmen auch ihre Spuren überein:

$$\text{Spur } A = \text{Spur } A'.$$

Beweis Es sei $A' = TAT^{-1}$, dann gilt

$$\begin{aligned} \chi_{A'}(x) &= \det(A' - x) = \det(TAT^{-1} - x) = \det(T(A - x)T^{-1}) \\ &= \det T \cdot \det(A - x) \cdot \det(T^{-1}) = \det(A - x) \\ &= \chi_A(x). \end{aligned}$$

Da die Spur in dem charakteristischen Polynom als ein Koeffizient auftritt, ist auch sie für beide Matrizen gleich. \blacklozenge

8.7 Eigenwerte und Eigenvektoren

In diesem Abschnitt sei V ein n -dimensionaler \mathbb{K} -Vektorraum, wobei $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$.

Definition Es sei f eine lineare Abbildung $V \rightarrow V$.

- (1) Man nennt $\lambda \in \mathbb{K}$ einen *Eigenwert* von f , wenn es einen Vektor $v \in V \setminus \{\vec{0}\}$ gibt, so dass

$$f(v) = \lambda \cdot v$$

erfüllt ist.

- (2) Der Vektor $v \in V \setminus \{\vec{0}\}$ heißt *Eigenvektor* von f (zum Eigenwert λ), wenn die Beziehung $f(v) = \lambda \cdot v$ gilt.
- (3) Die Menge der Vektoren

$$\text{Eig}(f, \lambda) = \{v \in V \mid f(v) = \lambda \cdot v\}$$

heißt der *Eigenraum* von f zum Eigenwert λ . Diese Menge besteht aus den Eigenvektoren zum Eigenwert λ und dem Nullvektor.

Zeige: Der Eigenraum ist ein Unterraum.

- (4) Die Dimension des Eigenraums $\text{Eig}(f, \lambda)$

$$\dim \text{Eig}(f, \lambda) = \dim \ker(f - \lambda) \stackrel{\text{LA14}}{=} n - \text{Rang}(f - \lambda)$$

(vgl. letzter Satz LA 14 in LAL01) heißt die *geometrische Vielfachheit* des Eigenwertes λ .

Satz 58 (ED 5: Charakterisierung von Eigenwerten) *Es sei $f : V \rightarrow V$ ein Endomorphismus und B eine Basis von V . Die folgenden Aussagen über eine Zahl $\lambda \in \mathbb{K}$ sind äquivalent:*

(A) *Die Zahl λ ist Eigenwert von f .*

(B) *Es gibt einen Vektor $v \in V \setminus \{\vec{0}\}$, so dass*

$$(f - \lambda)(v) = \vec{0},$$

d.h. v liegt im Kern von $f - \lambda$.

(C) *Es gibt eine Lösung $v_B \in \mathbb{K}^n \setminus \{\vec{0}\}$ des LGSs*

$$(f_B - \lambda)v_B = \vec{0}.$$

(D) *Die quadratische Matrix $f_B - \lambda$ ist nicht-invertierbar.*

(E) λ ist Nullstelle des charakteristischen Polynoms:

$$\chi_f(\lambda) = 0.$$

Definition Die Vielfachheit als Nullstelle in $\chi_f(x)$ heißt die *algebraische Vielfachheit* des Eigenwerts λ .

Beweis Der Satz fasst nur bereits bekannte Aussagen zusammen. ◆

Bemerkung Im Falle eines Vektorraums V über $\mathbb{K} = \mathbb{R}$ kann es zu einem Endomorphismus f nach unserer Definition genau genommen nur reelle Eigenwerte geben. Das charakteristische Polynom χ_f von f kann aber komplexe Nullstellen haben.

Beispiel: Die Abbildung

$$v \mapsto \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \cdot v$$

im \mathbb{R}^2 hat keine reellen Eigenwerte und damit auch keine Eigenvektoren. Das charakteristische Polynom hat zwei komplexe Nullstellen $\lambda_1 = +i$ und $\lambda_2 = -i$.

Diese Nullstellen heißen die *komplexen Eigenwerte* des Endomorphismus im reellen Vektorraum V .

Hintergrund dieser Begriffsbildung ist, dass Endomorphismen in reellen Vektorräumen „komplexifiziert“ werden können. Der komplexifizierte Endomorphismus besitzt dann die komplexen Eigenwerte im genauen Sinne der Definition.

In Bezug auf das obige Beispiel gilt:

$$\begin{aligned} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ i \end{pmatrix} &= \begin{pmatrix} -i \\ 1 \end{pmatrix} = -i \cdot \begin{pmatrix} 1 \\ i \end{pmatrix} \\ \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ -i \end{pmatrix} &= \begin{pmatrix} i \\ 1 \end{pmatrix} = i \cdot \begin{pmatrix} 1 \\ -i \end{pmatrix} \end{aligned}$$

Satz 59 (ED 6: Geometrische und algebraische Vielfachheit) *Es sei V ein Vektorraum und $f : V \rightarrow V$ eine lineare Abbildung. Für einen Eigenwert λ ist*

$$\text{geometrische Vielfachheit} \leq \text{algebraische Vielfachheit}.$$

Beweis Es sei $E = \text{Eig}(f, \lambda)$ der Eigenraum zum Eigenwert λ mit der Dimension m , die ja gleich der geometrischen Vielfachheit ist. Wir wählen eine Basis $\{w_1, \dots, w_m\}$ von E und ergänzen sie durch $\{w_{m+1}, \dots, w_n\}$ zu einer Basis B von V . Dann hat f die Darstellungsmatrix

$$f_B = \begin{pmatrix} \lambda & 0 & \cdots & 0 & * & \cdots & * \\ 0 & \ddots & \ddots & \vdots & \vdots & & \vdots \\ \vdots & \ddots & \ddots & 0 & \vdots & & \vdots \\ 0 & \cdots & 0 & \lambda & * & \cdots & * \\ 0 & \cdots & \cdots & 0 & * & \cdots & * \\ \vdots & & & \vdots & \vdots & & \vdots \\ 0 & \cdots & \cdots & 0 & * & \cdots & * \end{pmatrix}$$

Das charakteristische Polynom ist dann aber

$$\begin{aligned} \chi_f(x) &= \chi_{f_B}(x) = \det(f_B - x) = (\lambda - x)^m \cdot \underbrace{\det(\text{„Block rechts unten“} - x)}_{=:q(x)} \\ &= (-1)^m \cdot (x - \lambda)^m \cdot q(x). \end{aligned}$$

Also ist m kleiner oder gleich dem Grad der Nullstelle λ in $\chi_f(x)$. ◆

Die Gleichung in (E) bietet eine Handhabe zum Auffinden der zunächst unbekanntesten Eigenwerte λ . Ist ein Eigenwert λ bekannt, so kann man durch Lösen des LGS in (C) Eigenvektoren finden.

8.8 Diagonalisierbarkeit

Definition

- (1) Eine quadratische Matrix A heißt *Diagonalmatrix* (sie hat *Diagonalgestalt*), wenn $a_{jk} = 0$ für $j \neq k$.
- (2) Eine lineare Abbildung $f : V \rightarrow V$ heißt *diagonalisierbar*, wenn es eine Basis B gibt, für die die zugeordnete Matrix f_B Diagonalgestalt hat.
- (3) Das Verfahren zum Auffinden einer Basis B , bzgl. der die Abbildung Diagonalgestalt hat, heißt *Diagonalisierung*.

Satz 60 (ED 7: Diagonalisierbarkeit)

Die folgenden Aussagen über eine lineare Abbildung $f : V \rightarrow V$ sind äquivalent:

- (A) f kann als Diagonalmatrix mit den m verschiedenen Zahlen $\lambda_j \in \mathbb{K}$ in der Diagonale dargestellt werden. Dabei kommt die Zahl λ_j n_j -mal in der Diagonale vor.
- (B) Das charakteristische Polynom $\chi_f(x)$ von f zerfällt über \mathbb{K} in Linearfaktoren

$$\chi_f(x) = (x - \lambda_1)^{n_1} \cdot \dots \cdot (x - \lambda_m)^{n_m}$$

und für alle $j = 1, \dots, m$ ist $n_j = \dim \text{Eig}(f, \lambda_j)$ (algebraische = geometrische Vielfachheit).

- (C) f besitzt die (paarweise verschiedenen) Eigenwerte $\lambda_1, \dots, \lambda_m$ und es ist

$$\underbrace{\dim \text{Eig}(f, \lambda_1)}_{=n_1} + \dots + \underbrace{\dim \text{Eig}(f, \lambda_m)}_{=n_m} = n.$$

- (D) Der Vektorraum kann als Summe der Eigenräume $\text{Eig}(f, \lambda_j)$ mit $\dim \text{Eig}(f, \lambda_j) = n_j$ dargestellt werden:

$$V = \text{Eig}(f, \lambda_1) + \dots + \text{Eig}(f, \lambda_m).$$

Zur Erinnerung: Der Begriff „Summe von Unterräumen“ bedeutet, dass jeder Vektor $v \in V$ eindeutig als Summe

$$v = v_1 + \dots + v_m$$

mit $v_j \in \text{Eig}(f, \lambda_j)$ dargestellt werden kann.

Beweis (A) \implies (B) Für das charakteristische Polynom von f gilt

$$\chi_f(x) = \chi_{f_B}(x) = \det(f_B - x) = (x - \lambda_1)^{n_1} \cdot \dots \cdot (x - \lambda_m)^{n_m},$$

es zerfällt also in Linearfaktoren. Außerdem ist

$$\left. \begin{array}{l} f_B e_1 = \lambda_1 e_1, \\ \vdots \\ f_B e_{n_1} = \lambda_1 e_{n_1}, \end{array} \right\} \implies \dim \text{Eig}(f, \lambda_1) = n_1,$$

$$\left. \begin{array}{l} f_B e_{n_1+1} = \lambda_2 e_{n_1+1} \\ \vdots \\ f_B e_{n_1+n_2} = \lambda_2 e_{n_1+n_2}, \end{array} \right\} \implies \dim \text{Eig}(f, \lambda_2) = n_2,$$

$$\vdots$$

$$\left. \begin{array}{l} f_B e_{n_1+\dots+n_{m-1}+1} = \lambda_m e_{n_1+\dots+n_{m-1}+1} \\ \vdots \\ f_B e_n = \lambda_m e_n \end{array} \right\} \implies \dim \text{Eig}(f, \lambda_m) = n_m,$$

woraus zu sehen ist, dass die geometrischen mit den algebraischen Vielfachheiten übereinstimmen.

(B) \implies (C) Aufgrund von (ii) ist

$$\dim \text{Eig}(f, \lambda_1) + \dots + \dim \text{Eig}(f, \lambda_m) = n_1 + \dots + n_m = n.$$

(C) \implies (D) Wähle n Vektoren $w_1, \dots, w_n \in V$ so aus, dass

$$\begin{array}{lll} \{w_1, \dots, w_{n_1}\} & \text{Basis ist von} & \text{Eig}(f, \lambda_1) \\ \{w_{n_1+1}, \dots, w_{n_1+n_2}\} & \text{Basis ist von} & \text{Eig}(f, \lambda_2) \\ & \vdots & \\ \{w_{n_1+\dots+n_{m-1}+1}, \dots, w_n\} & \text{Basis ist von} & \text{Eig}(f, \lambda_m). \end{array}$$

Wir bezeichnen den zu w_j gehörigen Eigenwert mit $\tilde{\lambda}_j$. Es gilt also $w_j \in \text{Eig}(f, \tilde{\lambda}_j)$ für alle $j = 1, \dots, n$. Wir zeigen per Induktion über $k = 1, \dots, n$, dass die Menge

$$B_k = \{w_1, \dots, w_k\}$$

linear unabhängig ist.

Für $k = 1$ ist die Aussage trivial, da $w_1 \neq \vec{0}$ ist.

Wir nehmen dann als Induktionsvoraussetzung an, dass B_k linear unabhängig ist, und wollen zeigen, dass auch B_{k+1} linear unabhängig ist. Es sei also

$$\alpha_1 w_1 + \dots + \alpha_k w_k + \alpha_{k+1} w_{k+1} = \vec{0}. \quad (*)$$

Wir wenden auf diese Gleichung die Abbildung $f - \tilde{\lambda}_{k+1}$ an. Es gilt dann:

$$\alpha_1 (\tilde{\lambda}_1 - \tilde{\lambda}_{k+1}) w_1 + \dots + \alpha_k (\tilde{\lambda}_k - \tilde{\lambda}_{k+1}) w_k + \underbrace{\alpha_{k+1} (\tilde{\lambda}_{k+1} - \tilde{\lambda}_{k+1})}_{=0} w_{k+1} = \vec{0}.$$

Nach Induktionsvoraussetzung ist B_k linear unabhängig, es folgt:

$$\alpha_1(\tilde{\lambda}_1 - \tilde{\lambda}_{k+1}) = 0, \quad \dots \quad \alpha_k(\tilde{\lambda}_k - \tilde{\lambda}_{k+1}) = 0, \quad (**)$$

Es sei die Zahl $\ell \in \{1, \dots, k+1\}$ die kleinste Zahl, so dass

$$\tilde{\lambda}_\ell = \tilde{\lambda}_{k+1},$$

das heißt, der Vektor w_ℓ ist der erste in B_k , der den gleichen Eigenwert wie w_{k+1} hat. Umgekehrt heißt das aber, dass

$$\tilde{\lambda}_j \neq \tilde{\lambda}_{k+1} \quad \text{für alle } j = 1, \dots, \ell - 1.$$

Dann folgt aber mit (**), dass

$$\alpha_j = 0 \quad \text{für alle } j = 1, \dots, \ell - 1,$$

Damit reduziert sich die Gleichung (*) auf

$$\alpha_\ell w_\ell + \dots + \alpha_k w_k + \alpha_{k+1} w_{k+1} = \vec{0}.$$

Die Vektoren w_j in dieser Linearkombination gehören aber alle zum Eigenraum $\text{Eig}(f, \tilde{\lambda}_{k+1})$ und sind gemäß der Auswahl am Anfang des Beweises linear unabhängig. Es folgt also, dass

$$\alpha_\ell = \dots = \alpha_k = \alpha_{k+1} = 0.$$

(D) \implies (A) Wähle in jedem Eigenraum eine Basis (aus Eigenvektoren). Die Menge aller dieser Eigenvektoren bildet eine Basis für V . f_B hat dann Diagonalgestalt. \blacklozenge

Folgerung 61 *Besitzt eine lineare Abbildung $f : V \rightarrow V$ in einem n -dimensionalen Vektorraum n paarweise verschiedene Eigenwerte $\lambda_j \in \mathbb{K}$, so ist f diagonalisierbar.*

Beweis Setze in dem vorhergehenden Satz $m = n$ und $n_j = 1$ für $j = 1, \dots, m$. \blacklozenge

8.8.1 Algorithmus zur Diagonalisierung

Satz 62 *Vorgegeben sind:*

- Ein \mathbb{R} -Vektorraum V der Dimension n ,
- ein \mathbb{R} -Endomorphismus $f : V \rightarrow V$,
- einer der beiden Körper $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$,
- eine Basis B von V .

Der folgende Algorithmus liefert ...

- Die Entscheidung, ob es eine Basis B' von V gibt, so dass $f_{B'}$ eine Diagonalmatrix mit Einträgen aus \mathbb{K} ist.
- die Konstruktion dieser Basis B' ,

- die Berechnung der zugehörigen Transformationsmatrix $T = \Theta_{B'B}$.

1 Berechne die Matrix f_B . Sie ist durch die Gleichungen $f(w_j) = \sum_{k=1}^n (f_B)_{kj} w_k$, $j = 1, \dots, n$ festgelegt.

2 Berechne das charakteristische Polynom $\chi_f(x) = \chi_{f_B}(x)$.

3 Ermittle die m verschiedenen Nullstellen $\lambda_j \in \mathbb{K}$. Folgende Möglichkeiten stehen zur Verfügung:

- Nullstellen erraten oder erspüren,
- Lösungsformel für die Nullstellen einer quadratischen Gleichung (Mitternachtsformel),
- Satz von Vieta,
- Polynomdivision,
- Lösung für biquadratische Gleichungen $ax^4 + bx^2 + c = 0$.
- Vielleicht kennt man Eigenwerte (Nullstellen) aufgrund spezieller Eigenschaften des Endomorphismus (Siehe nächste Kapitel: Endomorphismen in euklidischen Vektorräumen).

4 Ermittle für $j = 1, \dots, m$ die algebraischen Vielfachheiten n_j der Nullstellen λ_j .

5 5a $n_1 + \dots + n_m < n$? Falls JA:
 χ_f zerfällt nicht vollständig in Linearfaktoren über \mathbb{K} , d.h.:
 f ist nicht über \mathbb{K} diagonalisierbar. \ominus
(Beachte, dass dies im fall $\mathbb{K} = \mathbb{C}$ nicht passieren kann)

5b $n_1 + \dots + n_m = n$? Falls JA:
 χ_f zerfällt vollständig in Linearfaktoren über \mathbb{K} :
Schreibe χ_f als Produkt der Linearfaktoren:

$$\chi_f(x) = (x - \lambda_1)^{n_1} \cdot \dots \cdot (x - \lambda_m)^{n_m}.$$

6 Ermittle die geometrischen Vielfachheiten g_j der Eigenwerte λ_j

$$g_j = \dim \ker(f_B - \lambda_j) = n - \text{Rang}(f_B - \lambda_j).$$

Beachte, dass $1 \leq g_j \leq n_j$ gelten muss.

7 Die Frage der Diagonalisierbarkeit kann jetzt entschieden werden:

7a Stimmen für eine Nullstelle die beiden Vielfachheiten n_j und g_j nicht überein, so ist f nicht diagonalisierbar. \ominus

7b Stimmen für alle Nullstellen die beiden Vielfachheiten n_j und g_j überein, so ist f diagonalisierbar. Es gibt eine Basis B' , bzgl. derer f die Darstellung

$$f_{B'} = \begin{pmatrix} \lambda_1 & & & \\ & \ddots & & \\ & & \ddots & \\ & & & \lambda_m \end{pmatrix}$$

hat. Dabei kommt der Eigenwert λ_j genau $n_j = g_j$ -mal vor.

Ermittle, wenn gewünscht, die neue Basis B' und die zugehörige Übergangsmatrix $T = \Theta_{B'B}$ gemäß der folgenden Schritte.

- 8] Für jede Nullstelle λ_j : Bestimme n_j linear unabhängige Vektoren $x_{j\ell} \in \mathbb{K}^n$, $\ell = 1, \dots, n_j$ durch Lösen des Gleichungssystems (vielleicht schon in Schritt 6] erledigt)

$$(f_B - \lambda_j)x = \vec{0}.$$

(Beachte, dass diese Lösungen nicht eindeutig sind).

- 9] Bilde die Matrix

$$T^{-1} = \Theta_{B'B}^{-1} := \left(x_{11}, \dots, x_{1n_1}, x_{21}, \dots, x_{2n_2}, \dots, x_{mn_1}, \dots, x_{mn_m} \right)$$

- 10] Berechne die Matrix T . (Vermeide zu viele Brüche dadurch, dass die inverse Determinante $\frac{1}{\det T^{-1}}$ als Vorfaktor stehen bleibt.)

- 11] Dann hat die Matrix

$$f_{B'} = T \cdot f_B \cdot T^{-1}$$

bezüglich der Basis $B' = \{w'_1, \dots, w'_n\}$ mit

$$w'_j = \sum_{k=1}^n (T^{-1})_{kj} w_k$$

die Diagonalgestalt wie in Schritt 7b] angegeben.

Beispiel 1

Wir betrachten das Beispiel vom Anfang des Kapitels. Es sei V ein zweidimensionaler reeller Vektorraum und der Endomorphismus $f : V \rightarrow V$ definiert durch

$$f(w_1) = \frac{16}{3}w_1 - 5w_2, \quad f(w_2) = \frac{10}{3}w_1 - 3w_2.$$

$$1. \quad f_B = \begin{pmatrix} \frac{16}{3} & \frac{10}{3} \\ -5 & -3 \end{pmatrix}$$

$$2. \quad \chi_f(x) = \begin{vmatrix} \frac{16}{3} - x & \frac{10}{3} \\ -5 & -3 - x \end{vmatrix} = \left(\frac{16}{3} - x\right)(-3 - x) - \frac{10}{3} \cdot (-5) = x^2 - \frac{7}{3}x + \frac{2}{3}$$

$$3. \quad \lambda_{1,2} = \frac{\frac{7}{3} \pm \sqrt{\frac{49}{9} - \frac{8}{3}}}{2} = \frac{\frac{7}{3} \pm \frac{5}{3}}{2}, \quad \text{also } \lambda_1 = 2, \quad \lambda_2 = \frac{1}{3}.$$

4. Die algebraischen Vielfachheiten der Eigenwerte sind $n_1 = 1, n_2 = 1$.

5b. Die Zerlegung in Linearfaktoren ist $\chi_f(x) = (x - 2)(x - \frac{1}{3})$.

6. Es muss $g_1 = 1, g_2 = 1$ gelten.

7b. Da die Vielfachheiten alle übereinstimmen, ist f diagonalisierbar.

8. Wir lösen das LGS für den ersten Eigenwert

$$(f_B - \lambda_1)x = \begin{pmatrix} \frac{16}{3} - 2 & \frac{10}{3} \\ -5 & -3 - 2 \end{pmatrix} x = \begin{pmatrix} \frac{10}{3} & \frac{10}{3} \\ -5 & -5 \end{pmatrix} x = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Wir finden ganz leicht eine Lösung

$$x_{11} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Wir lösen das LGS für den zweiten Eigenwert

$$(f_B - \lambda_2)x = \begin{pmatrix} \frac{16}{3} - \frac{1}{3} & \frac{10}{3} \\ -5 & -3 - \frac{1}{3} \end{pmatrix} x = \begin{pmatrix} 5 & \frac{10}{3} \\ -5 & -\frac{10}{3} \end{pmatrix} x = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Wir sehen ganz leicht

$$x_{21} = \begin{pmatrix} 2 \\ -3 \end{pmatrix}.$$

9. Jetzt kann die Matrix T^{-1} zusammengestellt werden:

$$T^{-1} = \begin{pmatrix} 1 & 2 \\ -1 & -3 \end{pmatrix}.$$

(Sie unterscheidet sich von der Matrix $\Theta_{B'B}^{-1}$ (Siehe Abschnitt 4.3/Beispiel) durch einen Faktor -1 in der zweiten Spalte. Beachte, dass die Übergangsmatrix T nicht eindeutig bestimmt ist.)

10. Die inverse Matrix ist

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

11. Es gilt dann

$$T \cdot f_B \cdot T^{-1} = \begin{pmatrix} 2 & 0 \\ 0 & \frac{1}{3} \end{pmatrix}.$$

Beispiel 2

Es sei f eine lineare Abbildung in einem 3-dimensionalen reellen Vektorraum. Bezüglich einer Basis habe sie die Matrixdarstellung

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

1. Dieser Schritt erübrigt sich.
2. Das charakteristische Polynom ist wegen der Dreiecksgestalt leicht zu berechnen:

$$\chi_f(x) = (3 - x)^2 \cdot (-1 - x).$$

3. Die beiden Eigenwerte sind $\lambda_1 = 3$, $\lambda_2 = -1$.
4. Die algebraischen Vielfachheiten der Eigenwerte sind $n_1 = 2$, $n_2 = 1$.
5. Die Zerlegung in Linearfaktoren liegt bereits (wegen der Dreiecksgestalt der Matrix f_B) vor.
6. Es ist

$$f_B - 3 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & -4 & c \\ 0 & 0 & 0 \end{pmatrix},$$

also

$$g_1 = n - \text{Rang}(f_B - 3) = \begin{cases} 3 - 2 = 1 \neq n_1, & \text{falls } c \neq 0, \\ 3 - 1 = 2 = n_1, & \text{falls } c = 0. \end{cases}$$

Weiter ist $g_2 = 1$.

7. f ist genau dann diagonalisierbar (über \mathbb{R} oder über \mathbb{C}), wenn $c = 0$. Wir führen für diesen Fall die Diagonalisierung durch.

8. Die Gleichung

$$(f_B - \lambda_1)x = \begin{pmatrix} 0 & 1 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & 0 \end{pmatrix} x = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

hat die beiden linear unabhängigen Lösungen

$$x_{11} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \quad \text{und} \quad x_{12} = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}.$$

Die Gleichung

$$(f_B - \lambda_2)x = \begin{pmatrix} 4 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 4 \end{pmatrix} x = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

hat die nicht-triviale Lösung

$$x_{21} = \begin{pmatrix} 1 \\ -4 \\ 0 \end{pmatrix}.$$

9. Es wird die Matrix

$$T^{-1} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & -4 \\ 1 & -1 & 0 \end{pmatrix}$$

zusammengestellt.

10. Die inverse Matrix ist

$$T = \frac{1}{-8} \cdot \begin{pmatrix} -4 & -1 & -4 \\ -4 & -1 & 4 \\ 0 & 2 & 0 \end{pmatrix}$$

11. Es gilt dann

$$T \cdot f_B \cdot T^{-1} = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

Beispiel 3

Es sei f eine lineare Abbildung in einem 3-dimensionalen Vektorraum. Bezüglich einer Basis habe sie die Matrixdarstellung

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

Ist f über \mathbb{R} oder über \mathbb{C} diagonalisierbar? Wenn ja, führe das Diagonalisierungsverfahren durch.

1. Dieser Schritt erübrigt sich.
2. Durch Entwicklung nach der zweiten Zeile erhält man das charakteristische Polynom:

$$\chi_f(x) = (3 - x) \cdot [(-2 - x)^2 + 1] = -(x - 3) \cdot [(x + 2)^2 + 1].$$

3. Der Fall $\mathbb{K} = \mathbb{R}$: Der zweite Faktor in χ_f ist positiv für alle $x \in \mathbb{R}$. Es gibt also nur eine reelle Nullstelle $\lambda_1 = 3$. f ist nicht über \mathbb{R} diagonalisierbar.
3. Der Fall $\mathbb{K} = \mathbb{C}$: Man kann das charakteristische Polynom in Linearfaktoren zerlegen:

$$\chi_f(x) = -(x - 3) \cdot [(x + 2)^2 + 1] = -(x - 3) \cdot [x - (i - 2)] \cdot [x - (-i - 2)].$$

Die Eigenwerte sind $\lambda_1 = 3$, $\lambda_2 = i - 2$, $\lambda_3 = -i - 2$.

4. Die algebraischen Vielfachheiten der Eigenwerte sind $n_1 = n_2 = n_3 = 1$.
5. Bereits erledigt.
6. Die geometrischen Vielfachheiten müssen alle gleich 1 sein.
7. f ist über \mathbb{C} diagonalisierbar.
8. Die Gleichung

$$(f_B - \lambda_1)x = \begin{pmatrix} 5 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & -5 \end{pmatrix} x = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

hat die nicht-triviale Lösung

$$x_{11} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}.$$

Die Gleichung

$$(f_B - \lambda_2)x = \begin{pmatrix} -i & 0 & 1 \\ 0 & 5 - i & 0 \\ -1 & 0 & -i \end{pmatrix} x = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

hat die nicht-triviale Lösung

$$x_{21} = \begin{pmatrix} 1 \\ 0 \\ i \end{pmatrix}.$$

Die Gleichung

$$(f_B - \lambda_3)x = \begin{pmatrix} +i & 0 & 1 \\ 0 & 5+i & 0 \\ -1 & 0 & +i \end{pmatrix} x = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

hat die nicht-triviale Lösung

$$x_{31} = \begin{pmatrix} 1 \\ 0 \\ -i \end{pmatrix}.$$

9. Wir bilden die Matrix

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

zusammengestellt.

10. Die inverse Matrix ist

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

11. Es gilt dann

$$T \cdot f_B \cdot T^{-1} = \begin{pmatrix} 3 & 0 & 0 \\ 0 & i-2 & 0 \\ 0 & 0 & -i-2 \end{pmatrix}.$$

9 Abbildungen in euklidischen Vektorräumen

9.1 Das Skalarprodukt beim Basiswechsel

Es sei $(V, \langle \cdot, \cdot \rangle)$ ein euklidischer Vektorraum und $B = \{w_1, \dots, w_n\}$ eine Orthonormalbasis. Wie bekannt, ist durch B eine Abbildung

$$\eta_B : \begin{cases} V & \rightarrow \mathbb{R}^n \\ v & \mapsto v_B = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \end{cases},$$

die Entwicklung nach den Basisvektoren, definiert.

Hinweis zur Notation: In diesem Kapitel sind für $j = 1, \dots, n$ die Zahlen $u_j, v_j \in \mathbb{K}$ Komponenten von Vektoren u bzw. v und w_j die Vektoren einer Orthonormalbasis.

Satz 63 *Ist B eine Orthonormalbasis für V , so gilt für alle $u, v \in V$*

$$\langle u, v \rangle = (u_B)^T \cdot v_B = u_1 v_1 + \dots + u_n v_n.$$

Man kann auch sagen, dass das Skalarprodukt $\langle \cdot, \cdot \rangle$ von V bei der B -Darstellung in das Standard-Skalarprodukt auf dem \mathbb{K}^n übergeht, also dort der Bilinearform $\mathbf{B} = I$ entspricht.

Beweis Ist $B = \{w_1, \dots, w_n\}$ die Orthonormalbasis, so gilt

$$\begin{aligned} \langle u, v \rangle &= \langle u_1 w_1 + \dots + u_n w_n, v_1 w_1 + \dots + v_n w_n \rangle \\ &= \langle u_1 w_1, v_1 w_1 \rangle + \dots + \langle u_n w_n, v_n w_n \rangle = u_1 v_1 + \dots + u_n v_n. \end{aligned}$$

◆

9.2 Die adjungierte Abbildung

Definition Es sei $(V, \langle \cdot, \cdot \rangle)$ ein euklidischer Vektorraum und $f : V \rightarrow V$ ein Endomorphismus. Eine weitere lineare Abbildung $g : V \rightarrow V$ heißt *adjungiert* zu f , wenn für alle $u, v \in V$ gilt:

$$\langle f(u), v \rangle = \langle u, g(v) \rangle.$$

Satz 64

(i) *Eine Abbildung $g : V \rightarrow V$ ist genau dann adjungiert zu $f : V \rightarrow V$, wenn die Matrixdarstellungen f_B und g_B bezüglich einer Orthonormalbasis B zueinander transponiert sind:*

$$g_B = (f_B)^T.$$

(ii) *Zu jeder Abbildung $f : V \rightarrow V$ gibt es genau eine adjungierte Abbildung $f^* : V \rightarrow V$.*

Beweis (i) Es gilt für alle $u, v \in V$

$$\begin{aligned}\langle f(u), v \rangle &= ((f(u))_B)^T \cdot v_B = (f_B \cdot u_B)^T \cdot v_B = u_B^T \cdot (f_B)^T \cdot v_B \\ \langle u, g(v) \rangle &= u_B^T \cdot (g(v))_B = u_B^T \cdot (g_B \cdot v_B) = u_B^T \cdot g_B \cdot v_B\end{aligned}$$

Sind also die beiden Matrixdarstellungen transponiert zueinander, so folgt, dass f und g adjungiert sind.

Gilt umgekehrt Gleichheit zwischen den beiden Ausdrücken links, so folgt, wenn man $u = e_j, v = e_k$ für $j, k \in \{1, \dots, n\}$ einsetzt, dass

$$(g_B)_{jk} = ((f_B)^T)_{jk}.$$

(ii) Eigentlich folgen Existenz und Eindeutigkeit aus der Aussage (i). Die Matrixdarstellungen der beiden Abbildungen müssen transponiert zueinander sein, die Transponierte einer Matrix existiert, sie ist eindeutig.

Wir zeigen noch einmal die Eindeutigkeit, ohne Bezug auf eine Basis zu nehmen: Wir nehmen an, es gebe zwei verschiedene zu f adjungierte Abbildungen g_1, g_2 . Dann gibt es einen Vektor $v \in V$, so dass $g_1(v) \neq g_2(v)$. Daraus lässt sich aber wie folgt

$$\begin{aligned}0 &\neq \langle g_1(v) - g_2(v), g_1(v) - g_2(v) \rangle \\ &= \langle g_1(v) - g_2(v), g_1(v) \rangle - \langle g_1(v) - g_2(v), g_2(v) \rangle \\ &= \langle f(g_1(v) - g_2(v)), v \rangle - \langle f(g_1(v) - g_2(v)), v \rangle = 0\end{aligned}$$

ein Widerspruch herleiten. ◆

Definition Es sei V ein euklidischer Vektorraum. f sei ein Endomorphismus von V und f_B die Matrixdarstellung bezüglich einer Orthonormalbasis B .

$f : V \rightarrow V$ heißt ...	wenn gilt: ...	Das ist genau dann der Fall, wenn die Matrix f_B ...
<i>selbstadjungiert,</i>	$f^* = f$	symmetrisch ist: $(f_B)^T = f_B$
<i>schief-selbstadjungiert,</i>	$f^* = -f$	schiefsymmetrisch ist: $(f_B)^T = -f_B$
<i>orthogonal,</i>	$f^* = f^{-1}$	orthogonal ist: $(f_B)^T = (f_B)^{-1}$

9.3 Orthogonale Abbildungen

Satz 65 (Charakterisierung orthogonaler Abbildungen) *Betrachte die folgenden Aussagen über eine lineare Abbildung $f : V \rightarrow V$ in einem euklidischen Vektorraum:*

(A) *f ist orthogonal.*

(B) *Das Skalarprodukt zweier beliebiger Vektoren $u, v \in V$ wird durch f nicht verändert:*

$$\langle f(u), f(v) \rangle = \langle u, v \rangle.$$

(C) *Die Länge (Norm) eines beliebigen Vektors $v \in V$ wird durch f nicht verändert:*

$$\|f(v)\| = \|v\| \quad \text{für alle } v \in V.$$

(D) *Der Cosinus des Zwischenwinkels von zwei beliebigen Vektoren wird durch f nicht verändert:*

$$\cos \sphericalangle(f(u), f(v)) = \cos \sphericalangle(u, v) \quad \text{für alle } u, v \in V.$$

(E) *Lässt man für Winkelmaße nur das Intervall $[0, \pi]$ (bzw. $[0^\circ, 180^\circ]$) zu (das heißt die Orientierung oder das Mehrfachdurchlaufen bleiben unberücksichtigt), so wird der Winkel durch f nicht verändert.*

Dann gelten die folgenden Implikationen:

$$(A) \iff (B) \iff (C) \implies (D) \iff (E)$$

Es ist bemerkenswert, dass in (B) und (C) nicht gefordert wird, dass f linear ist. Dies folgt dann aber aus der Äquivalenz (A) \iff (C).

Beweis (B) \implies (C) ist ganz einfach:

$$\|f(v)\| = \sqrt{\langle f(v), f(v) \rangle} = \sqrt{\langle v, v \rangle} = \|v\|.$$

Die Umkehrung (C) \implies (B) scheint anschaulich klar zu sein, der Beweis erfordert aber — wieder einmal — den Quadrat-Trick. Es gilt

$$\begin{aligned} \|u + v\|^2 &= \langle u + v, u + v \rangle = \langle u, u \rangle + 2\langle u, v \rangle + \langle v, v \rangle \\ &= \|u\|^2 + 2\langle u, v \rangle + \|v\|^2 \\ \|f(u) + f(v)\|^2 &= \langle f(u) + f(v), f(u) + f(v) \rangle \\ &= \langle f(u), f(u) \rangle + 2\langle f(u), f(v) \rangle + \langle f(v), f(v) \rangle \\ &= \|f(u)\|^2 + 2\langle f(u), f(v) \rangle + \|f(v)\|^2. \end{aligned}$$

Aufgrund von (C) stimmen die Normquadrate in der oberen und unteren Gleichungskette jeweils überein. Also müssen auch die Skalarprodukte in der oberen und unteren Gleichungskette übereinstimmen.

Die Implikation $(B) \implies (D)$ ist klar aufgrund der Beziehung zwischen Cosinus und Skalarprodukt

$$\cos \sphericalangle(f(u), f(v)) = \frac{\langle f(u), f(v) \rangle}{\sqrt{\langle f(u), f(u) \rangle} \sqrt{\langle f(v), f(v) \rangle}} = \frac{\langle u, v \rangle}{\sqrt{\langle u, u \rangle} \sqrt{\langle v, v \rangle}} = \cos \sphericalangle(u, v).$$

Die letzte Äquivalenz $(D) \iff (E)$ folgt daraus, dass die Kosinusfunktion auf dem Intervall $[0, \pi]$ streng monoton fallend, also bijektiv, ist. Dies wird in der Analysis genauer betrachtet. \blacklozenge

Satz 66 (Die orthogonale Gruppe $O(V)$) Die Menge der orthogonalen Abbildungen in einem euklidischen Vektorraum V bildet unter Hintereinanderausführung eine Gruppe. Sie wird mit $O(V)$ bezeichnet.

Beweis

Mit zwei orthogonalen Abbildungen f, g ist auch die Hintereinanderausführung orthogonal, da

$$\langle f(g(u)), f(g(v)) \rangle = \langle g(u), g(v) \rangle = \langle u, v \rangle.$$

Die identische Abbildung ist ohne Zweifel orthogonal, sie ist das neutrale Element der Gruppe. Es sei $f(u) = \vec{0}$. Dann gilt:

$$\langle f(u), f(u) \rangle = \langle \vec{0}, \vec{0} \rangle = 0,$$

wegen der Positiv-Definitheit des Skalarprodukts folgt $u = \vec{0}$. Das aber bedeutet, dass die Abbildung f injektiv ist. Eine injektive lineare Abbildung innerhalb eines Vektorraums V ist aber bijektiv. Es existiert also eine lineare Abbildung $g : V \rightarrow V$ mit $f \circ g = g \circ f = \text{id}_V$. Diese ist ebenfalls orthogonal, da

$$\langle g(u), g(v) \rangle = \langle f(g(u)), f(g(v)) \rangle = \langle u, v \rangle.$$

\blacklozenge

Satz 67 (Die spezielle orthogonale Gruppe $SO(V)$)

(i) Die Determinante einer orthogonalen Abbildung ist ± 1 .

(ii) Die Teilmenge von $O(V)$ der orthogonalen Abbildungen mit Determinante $+1$ bildet eine Untergruppe von $O(V)$. Sie wird mit $SO(V)$ bezeichnet.

Beweis (i) Es sei B eine Orthonormalbasis, f_B die Matrixdarstellung der orthogonalen Abbildung. Dann gilt

$$(\det f)^2 = (\det f_B)^2 = \det f_B \cdot \det f_B^T = \det f_B \cdot \det f_B^{-1} = 1.$$

(ii) Für zwei Abbildungen $f, g \in SO(V)$ gilt

$$\det(f \circ g) = \det(f_B \cdot g_B) = \det f_B \cdot \det g_B = 1 \cdot 1 = 1,$$

also ist auch $f \circ g$ in $SO(V)$. \blacklozenge

9.3.1 Orthogonale Abbildungen im euklidischen \mathbb{R}^2

Es sei

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

eine orthogonale Abbildung in dem mit dem Standard-Skalarprodukt versehenen Vektorraum $V = \mathbb{R}^2$.

Überlegen Sie mittels Anschauung: Wie schaut eine längen- und winkelerhaltende lineare Abbildung im \mathbb{R}^2 aus?

Wir überlegen gemeinsam mit Hilfe der Linearen Algebra:

(1) Die Orthogonalitätsbedingung an die lineare Abbildung lässt sich in den beiden Formen

$$I = A^T \cdot A \quad \text{und} \quad I = A \cdot A^T$$

darstellen. Ausgeschrieben lauten diese beiden Bedingungen

$$\begin{aligned} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} &= \begin{pmatrix} a & c \\ b & d \end{pmatrix} \cdot \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a^2 + c^2 & ab + cd \\ ab + cd & b^2 + d^2 \end{pmatrix} \\ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} &= \begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \begin{pmatrix} a & c \\ b & d \end{pmatrix} = \begin{pmatrix} a^2 + b^2 & ac + bd \\ ac + bd & c^2 + d^2 \end{pmatrix} \end{aligned}$$

und damit auf die sechs Gleichungen

$$a^2 + c^2 = 1 \quad b^2 + d^2 = 1 \quad a^2 + b^2 = 1 \quad c^2 + d^2 = 1 \quad ab + cd = 0 \quad ac + bd = 0.$$

(2) Daraus folgen die Gleichungen

$$a^2 = d^2 \quad b^2 = c^2$$

und damit die vier Fälle

$$a = \pm d \quad b = \pm c.$$

(3) Die beiden Fälle $+/+$ und $-/-$ widersprechen den Gleichungen mit den „gemischten“ Gliedern in (1). Es muss

$$a = d \text{ und } b = -c \quad \text{oder} \quad a = -d \text{ und } b = c.$$

gelten.

(4) Die Aussage in (3) bedeutet, dass die Matrix A die Form

$$A = \underbrace{\begin{pmatrix} a & -c \\ c & a \end{pmatrix}}_{=:D} \quad \text{oder} \quad A = \begin{pmatrix} a & c \\ c & -a \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}}_{=:S} \underbrace{\begin{pmatrix} a & -c \\ c & a \end{pmatrix}}_{=:D}$$

mit $a^2 + c^2 = 1$ hat.

(6) Man kann umgekehrt nachrechnen, dass Matrizen dieser Form orthogonal sind.

Da $\det D = +1$ und $\det S = -1$ ist, können wir zusammenfassend feststellen:

$$\begin{aligned}SO(2) &:= SO(\mathbb{R}^2) = \left\{ \begin{pmatrix} a & -c \\ c & a \end{pmatrix} \mid a^2 + c^2 = 1 \right\} \\O(2) &:= O(\mathbb{R}^2) = \left\{ \begin{pmatrix} a & -c \\ c & a \end{pmatrix}, \begin{pmatrix} a & c \\ c & -a \end{pmatrix} \mid a^2 + c^2 = 1 \right\}\end{aligned}$$

Wir untersuchen weiter die hier auftretenden Matrizen D und S .

Die Matrix S bildet den ersten Einheitsvektor e_1 auf sich selbst ab, der andere Eigenvektor wird mit -1 multipliziert, also (an der durch e_1 definierten Geraden) gespiegelt. S ist eine Spiegelungsmatrix.

Wir fassen unsere Beobachtungen zusammen im

Satz 68 (Charakterisierung von Drehmatrizen) Die folgenden Aussagen über einen Endomorphismus (2×2 -Matrix) $D : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ sind äquivalent:

(A) $D \in SO(\mathbb{R}^2)$, d.h. D ist orthogonal mit Determinante $+1$.

(B) Es ist

$$D = \begin{pmatrix} a & -c \\ c & a \end{pmatrix} \quad \text{mit } a, c \in \mathbb{R} \quad \text{und} \quad a^2 + c^2 = 1.$$

(C) Es existiert ein $\varphi \in \mathbb{R}$, so dass

$$D = \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix}$$

Diese Matrix dreht einen beliebigen Vektor um den Winkel φ (in Richtung $e_2 \rightarrow e_1$), das heißt D ist eine Drehmatrix.

(D) D hat zwei Eigenwerte und Eigenvektoren (über \mathbb{C}):

$$D \begin{pmatrix} 1 \\ i \end{pmatrix} = \lambda_1 \begin{pmatrix} 1 \\ i \end{pmatrix}, \quad D \begin{pmatrix} 1 \\ -i \end{pmatrix} = \lambda_2 \begin{pmatrix} 1 \\ -i \end{pmatrix},$$

wobei

$$\lambda_1 = \overline{\lambda_2} \quad \text{und} \quad |\lambda_1| = |\lambda_2| = 1$$

gilt. (Insbesondere ist D über \mathbb{C} diagonalisierbar.)

(E) „Identifiziert“ man die Ebene \mathbb{R}^2 mit der komplexen Zahlenebene \mathbb{C} mittels der bijektiven Abbildung

$$\mathbb{R}^2 \rightarrow \mathbb{C}, \quad \begin{pmatrix} x \\ y \end{pmatrix} \mapsto x + iy,$$

so korrespondiert zu der Abbildung D im \mathbb{R}^2 die Multiplikation in \mathbb{C} mit einer komplexen Zahl d vom Betrag $|d| = 1$.

Beweis (A) \implies (B): Das wurde in den dem Satz vorhergehenden Überlegungen dargelegt.

(B) \implies (A) ist leicht nachzurechnen.

(B) \iff (C): Das ist letztlich eine Aussage der Analysis.

(B) \implies (D): Setze $\lambda_1 = a - ic$ und $\lambda_2 = a + ic$.

(B) \iff (E): Mit $d := a + ic$ gilt

$$D \cdot \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ax - cy \\ cx + ay \end{pmatrix}$$

$$\begin{aligned} d[\varphi] \cdot (x + iy) &= (a + ic) \cdot (x + iy) \\ &= (ax - cy) + i(cx + ay). \end{aligned}$$

(D) \implies (A) Wir setzen D als reelle $n \times n$ -Matrix an:

$$D = \begin{pmatrix} x & y \\ z & w \end{pmatrix}$$

Mit $\lambda_1 := a - ic$, gilt dann:

$$\begin{pmatrix} x + iy \\ z + iw \end{pmatrix} = \begin{pmatrix} x & y \\ z & w \end{pmatrix} \begin{pmatrix} 1 \\ i \end{pmatrix} = (a - ic) \begin{pmatrix} 1 \\ i \end{pmatrix} = \begin{pmatrix} a - ic \\ c + ia \end{pmatrix}.$$

Daraus folgt die Darstellung in (B). Es ist weiter $a^2 + c^2 = |\lambda_1|^2 = 1$. ◆

Reelle Diagonalisierbarkeit einer Matrix aus $SO(\mathbb{R}^2)$ liegt genau dann vor, wenn

$$\lambda_1 = \lambda_2 \in \{-1, +1\}.$$

In diesen Fällen ergibt sich:

$$D = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I \quad \text{bzw.} \quad D = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = -I.$$

Die zweite Abbildung ist die *Punktspiegelung* mit $\vec{0}$ als Zentrum.

9.4 Diagonalisierung in euklidischen Vektorräumen

Satz 69 *Es sei V ein euklidischer Vektorraum und $f : V \rightarrow V$ ein Endomorphismus. Es sei $\lambda \in \mathbb{C}$ ein beliebiger Eigenwert von f .*

- (i) *Ist f selbst-adjungiert, so ist $\lambda \in \mathbb{R}$.*
- (ii) *Ist f schief-selbst-adjungiert, so ist $\lambda \in i\mathbb{R}$.*
- (iii) *Ist f orthogonal, so ist $|\lambda| = 1$.*

Beweis Vorarbeiten: Es sei $f_B \in \mathbb{R}^{n \times n}$ die Matrixdarstellung von f bzgl. einer Orthonormalbasis. λ ist dann auch Eigenwert von f_B . Es existiert also ein Eigenvektor $v \in \mathbb{C}^n \setminus \{\vec{0}\}$, so dass

$$f_B v = \lambda \cdot v.$$

Wir formen diese Gleichung etwas um (Der Konjugiert-Komplex-Operator wird jeweils auf die Komponenten von f_B bzw. v angewandt):

$$\overline{f_B v} = \overline{\lambda \cdot v} \implies f_B \bar{v} = \bar{\lambda} \cdot \bar{v} \implies \bar{v}^T \cdot f_B^T = \bar{\lambda} \cdot \bar{v}^T.$$

Als weitere Vorbereitung bemerken wir noch für den Eigenvektor $v = (v_1, \dots, v_n)^T \neq \vec{0}$, dass

$$\bar{v}^T \cdot v = \bar{v}_1 v_1 + \dots + \bar{v}_n v_n = |v_1|^2 + \dots + |v_n|^2 > 0.$$

(i) Ist f selbst-adjungiert, so folgt weiter:

$$\begin{aligned} \bar{\lambda} \cdot (\bar{v}^T \cdot v) &= (\bar{\lambda} \cdot \bar{v}^T) \cdot v = (\bar{v}^T \cdot f_B^T) \cdot v = \\ \bar{v}^T \cdot (f_B \cdot v) &= \bar{v}^T \cdot \lambda \cdot v = \lambda \cdot (\bar{v}^T \cdot v). \end{aligned}$$

Division durch den Faktor $(\bar{v}^T \cdot v)$ liefert das Resultat:

$$\bar{\lambda} = \lambda, \quad \text{also } \lambda \in \mathbb{R}.$$

(ii) Ist f schief-selbst-adjungiert, so folgt weiter:

$$\begin{aligned} \bar{\lambda} \cdot (\bar{v}^T \cdot v) &= (\bar{\lambda} \cdot \bar{v}^T) \cdot v = (\bar{v}^T \cdot f_B^T) \cdot v = \\ -\bar{v}^T \cdot (f_B \cdot v) &= -\bar{v}^T \cdot \lambda \cdot v = -\lambda \cdot (\bar{v}^T \cdot v), \end{aligned}$$

damit

$$\bar{\lambda} = -\lambda, \quad \text{also } \lambda \in i\mathbb{R}.$$

(iii) Ist schließlich f orthogonal, so gilt:

$$\begin{aligned} \bar{\lambda} \cdot \lambda \cdot (\bar{v}^T \cdot v) &= (\bar{\lambda} \bar{v}^T) \cdot \lambda v = (\bar{v}^T f_B^T) \cdot f_B v = \\ \bar{v}^T \cdot (f_B^T \cdot f_B) \cdot v &= \bar{v}^T \cdot v. \end{aligned}$$

Daraus folgt:

$$|\lambda|^2 = \bar{\lambda} \cdot \lambda = 1.$$



Definition Es sei $f : V \rightarrow V$ ein Endomorphismus. Ein Unterraum $U \subseteq V$ heißt *invariant (unter f)* oder *f -invariant*, wenn

$$f(U) \subseteq U$$

gilt. Das heißt, die Abbildung f bildet einen beliebigen Vektor $u \in U$ wieder auf einen Vektor aus U ab.

Beispiele:

- Die Unterräume $\{\vec{0}\}$ und V sind immer invariant.
- Bei einer Achsenspiegelung im \mathbb{R}^2 ist die Drehachse invariant.
- Bei einer Drehung im \mathbb{R}^3 ist die Drehachse invariant.
- Die Eigenräume eines Endomorphismus f sind invariant.

Satz 70 (Invarianz von orthogonalen Komplementen)

Es sei V ein euklidischer Vektorraum.

Ist f selbst-adjungiert, schief-selbst-adjungiert oder orthogonal, so gilt

$$U \text{ invariant} \implies U^\perp \text{ invariant.}$$

Beweis Es sei U invariant und $u' \in U^\perp$ beliebig. Die Aussage $f(U^\perp) \subseteq U^\perp$ ist jeweils bewiesen, wenn wir

$$\langle f(u'), u \rangle = 0 \quad \text{für alle } u \in U$$

zeigen können.

- (1) Ist f selbst-adjungiert, so gilt für alle $u \in U$:

$$\langle f(u'), u \rangle = \langle u', \underbrace{f(u)}_{\in U} \rangle = 0.$$

- (2) Wenn f schief-selbst-adjungiert ist, so geht das (fast) ganz genauso:

$$\langle f(u'), u \rangle = \langle u', \underbrace{-f(u)}_{\in U} \rangle = 0.$$

- (3) Es sei jetzt f orthogonal. Da f bijektiv ist, gilt allgemein $\dim f(U) = \dim U$ (Vgl. Satz 15/LAL01). Deswegen folgt aus $f(U) \subseteq U$ sogar $f(U) = U$ (Vgl. Satz 12/LAL01), damit gilt aber auch $f^{-1}(U) = U$.

Jetzt können wir schließen:

$$\langle f(u'), u \rangle = \langle u', f^*(u) \rangle = \langle u', \underbrace{f^{-1}(u)}_{\in U} \rangle = 0.$$

**Beispiel**

Im \mathbb{R}^2 mit dem Standard-Skalarprodukt sind die beiden Unterräume

$$U = \text{span}(e_1), \quad U^\perp = \text{span}(e_2)$$

orthogonale Komplemente zueinander. Die Matrix

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

lässt $U = \text{span}(e_1)$ invariant, nicht aber $U^\perp = \text{span}(e_2)$.

Satz 71 *Es sei V ein euklidischer Vektorraum. Ist $f : V \rightarrow V$ selbstadjungiert, so existiert eine Orthonormalbasis B von V , bzgl. der die Matrixdarstellung f_B eine reelle Diagonalmatrix ist.*

Beweis Wir zeigen per Induktion über $k = 0, \dots, n = \dim V$: Es gibt eine Menge $B_k \subseteq V$ mit k paarweise orthogonalen Eigenvektoren von f .

Für $k = 0$ ist die Aussage trivial. Wähle die leere Menge.

Für den Induktionsschritt nehmen wir an, dass die Aussage für k gilt. Der Unterraum $U_k := \text{span } B_k$ wird von Eigenvektoren von f aufgespannt, er ist deshalb f -invariant. Entscheidend ist jetzt, dass nach dem vorhergehenden Satz U_k^\perp invariant ist.

$$f(U_k^\perp) \subseteq U_k^\perp.$$

Das charakteristische Polynom $\chi_{f'}$ der auf diesen Unterraum eingeschränkten Abbildung

$$f' \begin{cases} U_k^\perp & \rightarrow U_k^\perp \\ u & \mapsto f(u) \end{cases}$$

hat mindestens eine komplexe Nullstelle, also hat f einen Eigenwert $\lambda_{k+1} \in \mathbb{C}$.

Der letzte Satz zeigt aber, dass λ_{k+1} sogar reell sein muss. Es existiert dann ein reeller Eigenvektor $u_{k+1} \in \text{Eig}(f, \lambda_{k+1}) \subseteq U_k^\perp$. Man setze jetzt

$$B_{k+1} = B_k \cup \{u_{k+1}\}.$$

Die Menge $\{u_1, \dots, u_k, u_{k+1}\}$ ist eine orthogonale Menge von Eigenvektoren, da ja u_{k+1} ein Eigenvektor aus dem orthogonalen Komplement von $\text{span } U_k$ ist.

Damit ist die Aussage auch für $k = n = \dim V$ gezeigt. Normiert man jetzt noch die Eigenvektoren in U_n , so hat man eine Orthonormalbasis. ◆

Bemerkung: Die Übungsaufgabe ?? zeigt, dass die in dem zugehörigen Diagonalisierungsverfahren auftretende Transformationsmatrix T orthogonal ist.

Da $\lambda^2 \neq 1$ ist, muss $\langle w, w \rangle = 0$ sein. (Dies ist kein Widerspruch zur Positiv-Definitheit des Skalarprodukts, da w ein komplexer Vektor ist.) Daraus folgt aber auch

$$\langle \bar{w}, \bar{w} \rangle = \overline{\langle w, w \rangle} = \bar{0} = 0.$$

Jetzt sehen wir, dass Real- und Imaginärteil von w aufeinander senkrecht stehen, es gilt nämlich

$$4i \langle u, v \rangle = \langle u + iv, u + iv \rangle - \langle u - iv, u - iv \rangle = \langle w, w \rangle - \langle \bar{w}, \bar{w} \rangle = 0.$$

Außerdem gilt:

$$\begin{aligned} f(u) &= f\left(\frac{1}{2} \cdot (w + \bar{w})\right) = \frac{1}{2} \cdot (\lambda w + \bar{\lambda} \bar{w}) \\ &= \frac{1}{2} [(a - ic)(u + iv) + (a + ic)(u - iv)] = au + cv. \\ f(v) &= f\left(\frac{1}{2i} \cdot (w - \bar{w})\right) = \frac{1}{2i} \cdot [\lambda w - \bar{\lambda} \bar{w}] \\ &= \frac{1}{2i} [(a - ic)(u + iv) - (a + ic)(u - iv)] = av - cu. \end{aligned}$$

Der durch u, v aufgespannte Unterraum ist invariant, die Matrixdarstellung dieser Abbildung innerhalb dieses Unterraums ist die Drehmatrix

$$D_{\frac{k+2}{2}} = \begin{pmatrix} a & -c \\ c & a \end{pmatrix}.$$

Setzen wir $w_{k+1} := \frac{1}{\|u\|} \cdot u$ und $w_{k+2} := \frac{1}{\|v\|} \cdot v$, so ist die Aussage für $k + 2$ gezeigt.

Für $k \geq 2r$ hat das charakteristische Polynom der auf U_k^\perp eingeschränkten Abbildung f immer noch (komplexe) Nullstellen. Diese sind aber reell, nach dem Satz über die Eigenwerte orthogonaler Abbildungen also gleich 1 oder gleich -1 . Man wähle als w_{k+1} einen zugehörigen (reellen) normierten Eigenvektor aus U_k^\perp . Dann ist die Aussage für $k + 1$ gezeigt. \blacklozenge

10 Elementare affine Geometrie

10.1 Affine Räume

Was sind freie Vektoren, Verschiebungsvektoren, Ortsvektoren, Punkte, Klassen paralleler Pfeile, ...?

Definition Eine Menge \mathcal{P} heißt (reeller) *affiner Raum* der Dimension n , wenn es einen reellen Vektorraum V der Dimension n gibt und eine Abbildung

$$\tau : \begin{cases} V \times \mathcal{P} & \rightarrow \mathcal{P} \\ (v, P) & \mapsto \tau_v(P) =: P + v, \end{cases}$$

so dass die folgenden drei Eigenschaften erfüllt sind:

- (1) Für jedes feste $v \in V$ ist $\tau_v : \mathcal{P} \rightarrow \mathcal{P}$ eine bijektive Abbildung.
- (2) Zu $P, Q \in \mathcal{P}$ gibt es genau ein $v \in V$, so dass

$$\tau_v(P) = Q.$$

Dieser *Verbindungsvektor* v wird dann mit \overrightarrow{PQ} bezeichnet.

- (3) Für $v, w \in V$ gilt:

$$\tau_v \circ \tau_w = \tau_{v+w}.$$

Die Elemente von \mathcal{P} heißen die *Punkte* des affinen Raumes.

Die Abbildungen τ_v heißt *Translation* oder *Verschiebung* (in \mathcal{P}) um den Vektor v . Deshalb kann der Vektorraum V auch als der *Vektorraum der Translationen* von \mathcal{P} bezeichnet werden.

Die Zuordnung

$$\begin{cases} V & \rightarrow \{ \text{Abbildungen } \mathcal{P} \rightarrow \mathcal{P} \} \\ v & \mapsto \tau_v \end{cases}$$

heißt auch (*Translations-*)*Operation* des Vektorraums V auf der Menge \mathcal{P} .

Mit dieser Definition wird eine Unterscheidung von Punkten (den statischen Objekten) und den Verschiebungen (den dynamischen Objekten) in der Geometrie möglich.

Die Schreibweise in der Definition von τ macht bereits deutlich, dass man sich die Operation von V auf \mathcal{P} vorstellen kann als eine Art Anheften des Vektorpfeils v an einen Punkt $P \in \mathcal{P}$. Das Bild $Q = P + v$ ist dann durch die Spitze des Pfeils festgelegt.

Dass sich ein affiner Raum von dem Vektorraum seiner Translationen nur wenig unterscheidet, wird in dem folgenden Satz deutlich:

Satz 73 Wird in \mathcal{P} irgendein Punkt O ("Origin") als Ursprung ausgezeichnet, so kommt dadurch eine bijektive Abbildung

$$\begin{cases} V & \rightarrow \mathcal{P} \\ v & \mapsto \tau_v(O) = O + v \end{cases} \quad \text{mit Umkehrabbildung} \quad \begin{cases} \mathcal{P} & \rightarrow V \\ P & \mapsto \overrightarrow{OP} \end{cases}$$

zustande. Die Räume sind „als Mengen gleich“. In diesem Zusammenhang heißt der Vektor \overrightarrow{OP} auch der Ortsvektor (bezüglich O) für den Punkt P .

Jeder affine Raum erbt also in gewisser Weise Strukturen von seinem Vektorraum V . Der umgekehrte Sachverhalt kommt in dem folgenden Beispiel zum Ausdruck:

Beispiele Jeder (reelle) Vektorraum V kann als affiner Raum $\mathcal{P} = V$ (mit sich selbst als Translationenvektorraum aufgefasst werden. Die zugehörige Abbildung ist einfach durch

$$\tau \begin{cases} V \times V & \rightarrow V \\ (v, w) & \mapsto \tau_v(w) =: w + v. \end{cases}$$

gegeben. Der kanonische Vektorraum \mathbb{R}^n wird so zum kanonischen affinen Raum \mathcal{A}^n . Als Ursprung wird hier oft der Punkt $\vec{0}$ gewählt.

Der wesentliche Unterschied von affinen Räumen gegenüber Vektorräumen besteht darin, dass in ihnen kein Nullpunkt von vornherein (auf natürliche Weise) ausgezeichnet ist.

Unsere „Anschauungsräume“ der Dimension 2 (Zeichenebene) oder der Dimension 3 (der Raum, in dem wir leben) werden geeigneter als Affine Räume denn als Vektorräume modelliert.

Erste Beobachtungen und Rechenregeln werden in dem folgenden Satz zusammengefasst:

Satz 74

(i) Die Translationsoperation hat die folgenden Eigenschaften:

$$\begin{aligned} \tau_{\vec{0}} &= \text{id}_{\mathcal{P}}. \\ \tau_v^{-1} &= \tau_{-v} \end{aligned}$$

(ii) Es seien $P, Q, R \in \mathcal{P}$. Für Verbindungsvektoren gilt:

$$\begin{aligned} \overrightarrow{PP} &= \vec{0} \\ \overrightarrow{PQ} + \overrightarrow{QR} &= \overrightarrow{PR} \\ \overrightarrow{QP} &= -\overrightarrow{PQ} \end{aligned}$$

(iii) Es seien $P, Q \in \mathcal{P}$. Für Ortsvektoren gilt:

$$\overrightarrow{PQ} = \overrightarrow{OQ} - \overrightarrow{OP}.$$

Beweis (i) Es sei P ein beliebiger Punkt von \mathcal{P} . Dann gilt:

$$\tau_{\vec{0}} \circ \tau_{\vec{0}}(P) \stackrel{(3)}{=} \tau_{\vec{0}+\vec{0}}(P) = \tau_{\vec{0}}(P).$$

Da $\tau_{\vec{0}}$ gemäß Eigenschaft (1) bijektiv ist, können wir die Umkehrabbildung $\tau_{\vec{0}}^{-1}$ auf diese Gleichung anwenden. Es folgt:

$$\tau_{\vec{0}}(P) = P.$$

Also ist $\tau_{\vec{0}}$ die identische Abbildung von \mathcal{P} .

Die zweite Aussage ist jetzt ganz einfach zu zeigen: Es gilt nämlich

$$\tau_{-v} \circ \tau_v = \tau_{\vec{0}} = \text{id}_{\mathcal{P}} \quad \text{und} \quad \tau_v \circ \tau_{-v} = \tau_{\vec{0}} = \text{id}_{\mathcal{P}}.$$

und das heißt, dass $\tau_{\vec{0}}$ die Umkehrabbildung von τ_v ist.

(ii) Der Vektor \overrightarrow{PQ} ist gemäß (2) durch die Eigenschaft

$$\tau_{\overrightarrow{PQ}}(P) = Q$$

eindeutig charakterisiert.

Dann folgt die erste Behauptung mit

$$\tau_{\overrightarrow{PP}}(P) = P = \tau_{\vec{0}}(P).$$

Die zweite Behauptung sieht man so:

$$\tau_{\overrightarrow{PQ} + \overrightarrow{QR}}(P) = \tau_{\overrightarrow{QR}}(\tau_{\overrightarrow{PQ}}(P)) = \tau_{\overrightarrow{QR}}(Q) = R = \tau_{\overrightarrow{PR}}(P).$$

Die dritte Behauptung ergibt sich mit der ersten und zweiten so:

$$\overrightarrow{PQ} + \overrightarrow{QP} = \overrightarrow{PP} = \vec{0}.$$

(iii) ist klar. ◆

10.2 Affine Unterräume

Definition Es sei \mathcal{P} ein affiner Raum mit Translationsvektorraum V . Eine Teilmenge $\mathcal{Q} \subseteq \mathcal{P}$ heißt affiner Unterraum von \mathcal{P} der Dimension m , wenn es einen Unterraum $U \subseteq V$ mit $\dim U = m$ gibt, so dass

$$\overrightarrow{PQ} \in U \quad \text{für alle } P, Q \in \mathcal{Q}.$$

Dieser Unterraum U ist eindeutig bestimmt.

Ein affiner Raum heißt

- *Gerade*, wenn seine Dimension gleich 1 ist,
- *Ebene*, wenn seine Dimension gleich 2 ist,
- *Hyperebene*, wenn seine Dimension gleich $n - 1$ ist.

Eine Menge $\{P\}$, die genau einen Punkt von \mathcal{P} enthält, ist ein 0-dimensionaler affiner Unterraum.

Beobachtung: Eine Teilmenge $\mathcal{Q} \subseteq \mathcal{P}$ ist genau dann ein affiner Unterraum, wenn es einen Punkt $P \in \mathcal{P}$ und einen Unterraum $U \subseteq V$ gibt, so dass

$$\mathcal{Q} = \{\tau_u(P) \mid u \in U\} = P + U.$$

Bemerkung: Ist $B = \{w_1, \dots, w_m\}$ eine Basis des Unterraums U , so heißt der Ausdruck unter der geschweiften Klammer in

$$\mathcal{Q} = \left\{ \underbrace{P + \mu_1 w_1 + \dots + \mu_m w_m}_{\substack{\mu_1, \dots, \mu_m \in \mathbb{R}}} \right\}.$$

eine *Parameterdarstellung* des affinen Unterraums \mathcal{Q} . Für jedes $Q \in \mathcal{Q}$ gibt es ein m -Tupel $(\mu_1, \dots, \mu_m)^T \in \mathbb{R}^m$, so dass

$$Q = P + \mu_1 w_1 + \dots + \mu_m w_m.$$

Zwei affine Unterräume \mathcal{Q}_1 und \mathcal{Q}_2 heißen *parallel (zueinander)*, wenn für die zugehörigen Unterräume U_1 und U_2 gilt:

$$U_1 \subseteq U_2 \quad \text{oder} \quad U_2 \subseteq U_1 \quad (*)$$

Symbolisch schreibt man für Parallelität bzw. Nicht-Parallelität:

$$\mathcal{Q}_1 \parallel \mathcal{Q}_2, \quad \mathcal{Q}_1 \not\parallel \mathcal{Q}_2.$$

Bemerkung: Gilt $\dim \mathcal{Q}_1 = \dim \mathcal{Q}_2$, so ist die Bedingung (*) äquivalent zu

$$U_1 = U_2.$$

Zwei affine Unterräume \mathcal{Q}_1 und \mathcal{Q}_2 heißen *windschief (zueinander)*, wenn

$$\mathcal{Q}_1 \cap \mathcal{Q}_2 = \emptyset \quad \text{und} \quad \mathcal{Q}_1 \not\parallel \mathcal{Q}_2.$$

Für zwei affine Unterräume \mathcal{Q}_1 und \mathcal{Q}_2 mit $\dim \mathcal{Q}_1 = \dim \mathcal{Q}_2$ kann man dann die folgenden Fälle unterscheiden:

- $\mathcal{Q}_1 = \mathcal{Q}_2$,
- $\mathcal{Q}_1 \neq \mathcal{Q}_2$ mit $\mathcal{Q}_1 \cap \mathcal{Q}_2 \neq \emptyset$.
- $\mathcal{Q}_1 \neq \mathcal{Q}_2$ mit $\mathcal{Q}_1 \parallel \mathcal{Q}_2$,
- \mathcal{Q}_1 windschief zu \mathcal{Q}_2 .

Betrachte als Beispiele Geraden und Ebenen im zwei— oder drei-dimensionalen Anschauungsraum. Welche Fälle können dann gar nicht auftreten?

Wir können die Lösungsmengen von LGSen geometrisch beschreiben (Siehe auch Satz LA 6 aus LAL01).

Satz 75 (LGS-Lösungsmengen als affine Unterräume) *Es sei*

$$Ax = b$$

ein lineares Gleichungssystem mit einer reellen $m \times n$ -Matrix A und einem Vektor $b \in \mathbb{R}^n$.

Dann ist die Lösungsmenge $\mathcal{L}(A|b)$ genau dann nicht-leer, wenn

$$\text{Rang}(A|b) = \text{Rang } A.$$

Ist die Lösungsmenge nicht-leer, so bildet sie einen affinen Unterraum der Dimension

$$\dim \mathcal{L}(A|b) = n - \text{Rang } A.$$

Die Gleichung $Ax = b$ heißt in diesem Zusammenhang auch einfach Die Gleichung des affinen Unterraums.

10.3 Euklidische affine Räume

Definition

- (i) Ist auf dem Translationsvektorraum V eines affinen Vektorraums \mathcal{P} ein Skalarprodukt $\langle \cdot, \cdot \rangle$ gegeben, so spricht man von einem *euklidischen affinen Raum*.
- (ii) Für zwei Punkte $P, Q \in \mathcal{P}$ heißt die nicht-negative reelle Zahl

$$\overline{PQ} = d(P, Q) := \|\overrightarrow{PQ}\| = \sqrt{\langle \overrightarrow{PQ}, \overrightarrow{PQ} \rangle}$$

der *Abstand* zwischen P und Q .

Beobachtung: Für alle $P, Q, R \in \mathcal{P}$ gilt:

$$d(P, Q) = 0 \iff P = Q$$

$$d(P, Q) = d(Q, P)$$

$$d(P, R) \leq d(P, Q) + d(Q, R).$$

Ganz allgemein heißt eine Funktion $d : M \times M \rightarrow \mathbb{R}_0^+$ auf einer Menge M eine *Metrik*, wenn sie diese drei Eigenschaften erfüllt.

10.3.1 Der Abstand zwischen einem Punkt und einem affinen Unterraum

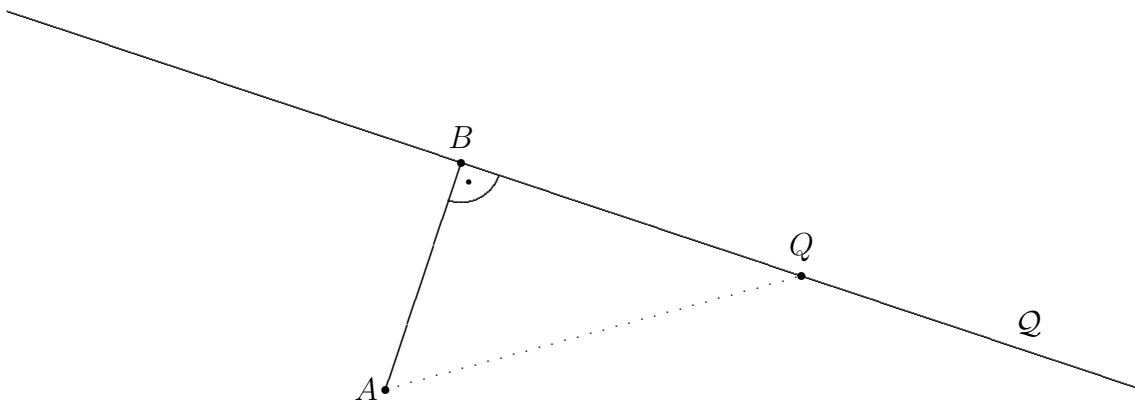
Satz 76 (Senkrechte Distanz ist minimale Distanz) *Es sei \mathcal{Q} ein m -dimensionaler affiner Unterraum von \mathcal{P} und $A \in \mathcal{P}$. Die folgenden Aussagen über einen Punkt $B \in \mathcal{Q}$ sind äquivalent.*

- (i) *Ist $\{w_1, \dots, w_m\}$ eine Orthonormalbasis des Translationsvektorraums U von \mathcal{Q} , so gilt*

$$\overrightarrow{QB} = \langle \overrightarrow{QA}, w_1 \rangle w_1 + \dots + \langle \overrightarrow{QA}, w_m \rangle w_m$$

für alle $Q \in \mathcal{Q}$.

- (ii) *Es gilt $\langle \overrightarrow{AB}, \overrightarrow{BQ} \rangle = 0$ für alle $Q \in \mathcal{Q}$, d.h. \overrightarrow{AB} steht senkrecht auf \mathcal{Q} .*
- (iii) *Es ist $\|\overrightarrow{AB}\| < \|\overrightarrow{AQ}\|$ für alle $Q \in \mathcal{Q} \setminus \{B\}$, d.h. $B \in \mathcal{Q}$ ist näher an A als alle anderen Punkte von \mathcal{Q} ,*
- (iv) *Es gilt: $\|\overrightarrow{AB}\| \leq \|\overrightarrow{AQ}\|$ für alle $Q \in \mathcal{Q}$, d.h. bei B wird der Abstand der Punkte $Q \in \mathcal{Q}$ zu A minimal.*



Beweis (i) \implies (ii): Der Vektor \overrightarrow{BQ} hat die folgende Entwicklung bzgl. der ONB

$$\overrightarrow{BQ} = \langle \overrightarrow{BQ}, w_1 \rangle w_1 + \dots + \langle \overrightarrow{BQ}, w_n \rangle w_n,$$

was man leicht durch Skalarmultiplikation dieser Gleichung mit den Vektoren w_i der Orthonormalbasis nachweisen kann. Es gilt dann weiter:

$$\begin{aligned} \langle \overrightarrow{AB}, \overrightarrow{BQ} \rangle &= \langle \overrightarrow{AQ}, \overrightarrow{BQ} \rangle - \langle \overrightarrow{AQ}, w_1 \rangle \langle w_1, \overrightarrow{BQ} \rangle - \dots - \langle \overrightarrow{AQ}, w_m \rangle \langle w_m, \overrightarrow{BQ} \rangle \\ &= \langle \overrightarrow{AQ}, \overrightarrow{BQ} - \langle w_1, \overrightarrow{BQ} \rangle w_1 - \dots - \langle w_m, \overrightarrow{BQ} \rangle w_m \rangle = \langle \overrightarrow{AQ}, \vec{0} \rangle = 0. \end{aligned}$$

(ii) \implies (i): Der Vektor \overrightarrow{BQ} ist in U enthalten, besitzt also eine Entwicklung bzgl. der Orthonormalbasis:

$$\overrightarrow{QB} = \lambda_1 w_1 + \dots + \lambda_m w_m.$$

Wir berechnen jetzt für $\overrightarrow{QB} = \overrightarrow{QA} + \overrightarrow{AB}$ die Skalarprodukte mit den Basisvektoren:

$$\lambda_i = \langle \overrightarrow{QB}, w_i \rangle = \langle \overrightarrow{QA}, w_i \rangle + \langle \overrightarrow{AB}, w_i \rangle = \langle \overrightarrow{QA}, w_i \rangle.$$

Daraus folgt $\lambda_i = -\langle \overrightarrow{AQ}, w_i \rangle$.

(ii) \implies (iii): Für $Q \in \mathcal{Q} \setminus \{B\}$ gilt:

$$\begin{aligned} \|\overrightarrow{AQ}\|^2 - \|\overrightarrow{AB}\|^2 &= \|\overrightarrow{AB} + \overrightarrow{BQ}\|^2 - \|\overrightarrow{AB}\|^2 \\ &= \langle \overrightarrow{AB} + \overrightarrow{BQ}, \overrightarrow{AB} + \overrightarrow{BQ} \rangle - \langle \overrightarrow{AB}, \overrightarrow{AB} \rangle \\ &= 2\langle \overrightarrow{AB}, \overrightarrow{BQ} \rangle + \langle \overrightarrow{BQ}, \overrightarrow{BQ} \rangle = \|\overrightarrow{BQ}\|^2 > 0. \end{aligned}$$

Die Aussage (iii) impliziert, dass $\|\overrightarrow{AB}\| \leq \|\overrightarrow{AQ}\|$ für alle $Q \in \mathcal{Q}$. Das ist aber gerade die Aussage (iv).

(iv) \implies (ii): Dies ist einfach dann, wenn man den Trick kennt: Es sei $Q \in \mathcal{P}$ beliebig. Dann ist auch der Punkt P , definiert durch

$$\overrightarrow{AP} = \overrightarrow{AB} - \frac{\langle \overrightarrow{AB}, \overrightarrow{BQ} \rangle}{\langle \overrightarrow{BQ}, \overrightarrow{BQ} \rangle} \cdot \overrightarrow{BQ}$$

in \mathcal{Q} enthalten. Daher gilt mit (iv)

$$\begin{aligned} 0 &\leq \|\overrightarrow{AP}\|^2 - \|\overrightarrow{AB}\|^2 = \|\overrightarrow{AB} - \frac{\langle \overrightarrow{AB}, \overrightarrow{BQ} \rangle}{\langle \overrightarrow{BQ}, \overrightarrow{BQ} \rangle} \cdot \overrightarrow{BQ}\|^2 - \|\overrightarrow{AB}\|^2 \\ &= -2 \frac{\langle \overrightarrow{AB}, \overrightarrow{BQ} \rangle^2}{\langle \overrightarrow{BQ}, \overrightarrow{BQ} \rangle} + \frac{\langle \overrightarrow{AB}, \overrightarrow{BQ} \rangle^2}{\langle \overrightarrow{BQ}, \overrightarrow{BQ} \rangle^2} \cdot \langle \overrightarrow{BQ}, \overrightarrow{BQ} \rangle = -\frac{\langle \overrightarrow{AB}, \overrightarrow{BQ} \rangle^2}{\langle \overrightarrow{BQ}, \overrightarrow{BQ} \rangle} \leq 0. \end{aligned}$$

Daraus folgt aber $\langle \overrightarrow{AB}, \overrightarrow{BQ} \rangle = 0$ für alle $Q \in \mathcal{Q}$. ◆

Mit Hilfe des in diesem Satz beschriebenen Punktes kann man den Abstand eines Punktes zu einem affinen Unterraum \mathcal{Q} definieren.

Definition Ist \mathcal{Q} ein affiner Unterraum von \mathcal{P} und $A \in \mathcal{P}$ ein Punkt, so heißt

$$d(A, \mathcal{Q}) := \min\{d(A, Q) \mid Q \in \mathcal{Q}\} = d(A, B) \quad (B \text{ wie im Satz beschrieben})$$

der *Abstand* zwischen dem Unterraum \mathcal{Q} und dem Punkt A .

10.3.2 Hyperebenen und Normalenvektoren

Es sei \mathcal{Q} eine Hyperebene in \mathcal{P} und U der zugehörige $(n-1)$ -dimensionale Unterraum von V . Eine Orthonormalbasis $B = \{w_1, \dots, w_{n-1}\}$ von U lässt sich durch einen Vektor \vec{n} zu einer Orthonormalbasis von V ergänzen. Er ist bis auf einen Faktor ± 1 eindeutig bestimmt. Ein solcher Vektor \vec{n} hat also die Länge 1 und steht senkrecht auf den Verbindungsvektoren in \mathcal{Q} , er heißt deshalb *Normaleneinheitsvektor* für \mathcal{Q} .

Satz 77 Für den Abstand eines Punktes $A \in \mathcal{P}$ von einer Hyperebene \mathcal{Q} gilt Ist Q ein beliebiger Punkt in der Hyperebene \mathcal{Q} , so gilt für den

$$d(A, \mathcal{Q}) = |\langle \vec{n}, \overrightarrow{AQ} \rangle|,$$

wobei Q ein beliebiger Punkt von \mathcal{Q} ist.

Beweis Ist nämlich $B \in \mathcal{Q}$ der abstandsminimierende Punkt für A (vgl. oben), so stehen sowohl \vec{n} als auch \overrightarrow{AB} (gemäß obigem Satz) auf der Hyperebene \mathcal{Q} senkrecht. Damit sind sie aber linear abhängig. Mit Satz EU 2 (iii) (über die CSU) gilt dann

$$\begin{aligned} d(A, \mathcal{Q}) &= d(A, B) = \|\overrightarrow{AB}\| = \|\vec{n}\| \cdot \|\overrightarrow{AB}\| = |\langle \vec{n}, \overrightarrow{AB} \rangle| \\ &= |\langle \vec{n}, \overrightarrow{AB} \rangle + \langle \vec{n}, \overrightarrow{BQ} \rangle| = |\langle \vec{n}, \overrightarrow{AQ} \rangle|. \end{aligned}$$

◆

Ist $P \in \mathcal{Q}$ ein fest gewählter Punkt in einer Hyperebene, so gilt für alle $Q \in \mathcal{Q}$

$$\langle \vec{n}, \overrightarrow{OQ} - \overrightarrow{OP} \rangle = 0.$$

Führt man eine Orthonormalbasis in dem Translationsvektorraum V ein, so kann diese Gleichung auch in Koordinatenform aufgeschrieben werden. Diese Gleichung heißt dann die *Hesse'sche Normalform* der Hyperebene \mathcal{Q} .

10.3.3 Isometrien

Definition Eine Abbildung $f : \mathcal{P} \rightarrow \mathcal{P}$ in einem euklidischen affinen Raum heißt eine *Isometrie* oder *Kongruenzabbildung*, wenn

$$d(f(P), f(Q)) = d(P, Q) \quad \text{für alle } P, Q \in \mathcal{P}.$$

Satz 78 (Isometrien)

(i) Translationen sind Isometrien.

(ii) Es sei $D \in \mathcal{P}$, die beiden Abbildungen $f : \mathcal{P} \rightarrow \mathcal{P}$ und $f^\sharp : V \rightarrow V$ seien durch die Beziehung

$$f^\sharp(\overrightarrow{DP}) = \overrightarrow{Df(P)} \quad \text{für alle } P \in \mathcal{P}$$

wechselseitig definiert.

Genau dann ist f eine Isometrie mit Fixpunkt D , wenn f^\sharp orthogonal ist.

(iii) Ist f eine Isometrie, so gibt es einen Vektor v und einen Punkt $D \in \mathcal{P}$, so dass

$$f = g \circ \tau_v,$$

wobei g eine Isometrie mit D als Fixpunkt ist.

Beweis (i)

$$\begin{aligned} d(\tau_v(P), \tau_v(Q)) &= \|\overrightarrow{\tau_v(P)\tau_v(Q)}\| = \|\overrightarrow{\tau_v(P)P} + \overrightarrow{PQ} + \overrightarrow{\tau_v(Q)Q}\| \\ &= \|-v + \overrightarrow{PQ} + v\| = \|\overrightarrow{PQ}\| = d(P, Q). \end{aligned}$$

(ii) Es sei f eine Isometrie mit Fixpunkt D . Dann gilt für beliebiges $v \in V$, $P := \tau_v(D)$:

$$\|f^\sharp(v)\| = \|f^\sharp(\overrightarrow{DP})\| = \|\overrightarrow{Df(P)}\| = \|\overrightarrow{f(D)f(P)}\| = \|\overrightarrow{DP}\| = \|v\|,$$

daraus folgt mit dem Satz über die Eigenschaften orthogonaler Abbildungen (Satz 36), dass f^\sharp orthogonal ist.

Ist umgekehrt eine Abbildung f^\sharp mit der im Satz angegebenen Eigenschaft gegeben, so gilt:

$$\overrightarrow{Df(D)} = f^\sharp(\overrightarrow{DD}) = f^\sharp(\vec{0}) = \vec{0} \implies f(D) = D.$$

Außerdem ist

$$\begin{aligned} d(f(P), f(Q)) &= \|\overrightarrow{f(P)f(Q)}\| = \|\overrightarrow{f(P)D} + \overrightarrow{Df(Q)}\| = \|-f^\sharp(\overrightarrow{DP}) + f^\sharp(\overrightarrow{DQ})\| \\ &= \|f^\sharp(\overrightarrow{DQ} - \overrightarrow{DP})\| = \|\overrightarrow{DQ} - \overrightarrow{DP}\| = \|\overrightarrow{PQ}\| = d(P, Q). \end{aligned}$$

(iii) Wähle als Vektor

$$v := \overrightarrow{f^{-1}(D)D}.$$

Dann gilt:

$$f \circ \tau_{-v}(D) = f \circ f^{-1}(D) = D.$$

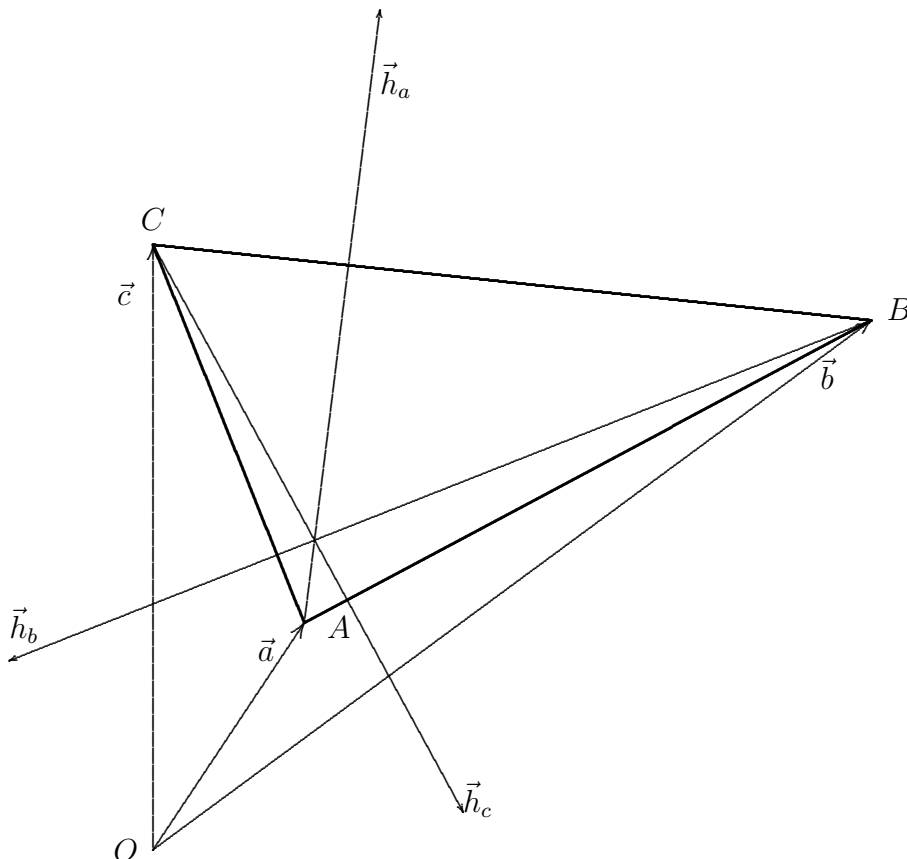
Daraus folgt, dass $g := f \circ \tau_{-v}$ eine Isometrie mit Fixpunkt D ist. ◆

Anwendung:

Satz 79 (Schnittpunkte von Transversalen in einem Dreieck) *Es sei ABC ein Dreieck in der Zeichenebene (Euklidische affine Ebene).*

- (i) *Die Höhen des Dreiecks schneiden sich in einem Punkt.*
(ii) *Die Mittelsenkrechten des Dreiecks schneiden sich in einem Punkt.*

Beweis



Wir wählen irgendeinen Punkt O als Ursprung und bezeichnen die Ortsvektoren von A, B, C mit $\vec{a}, \vec{b}, \vec{c}$. Weiter bezeichnen wir mit $\vec{\ell}_a, \vec{\ell}_b, \vec{\ell}_c$ Vektoren, die jeweils senkrecht (lotrecht) auf den Seiten des Dreiecks stehen. Ihre Länge ist ohne Belang.

Es gilt dann:

$$\langle \vec{\ell}_a, \overrightarrow{BC} \rangle = \langle \vec{\ell}_b, \overrightarrow{CA} \rangle = \langle \vec{\ell}_c, \overrightarrow{AB} \rangle = 0. \quad (*)$$

(i) Wir berechnen den Ortsvektor \vec{h}_{ab} des Schnittpunkts der Höhen auf a und b durch den Ansatz

$$\vec{h}_{ab} = \vec{a} + \mu_a \vec{\ell}_a = \vec{b} + \mu_b \vec{\ell}_b.$$

Wir bilden das Skalarprodukt mit \overrightarrow{BC} und beachten dabei (*):

$$\langle \vec{a}, \overrightarrow{BC} \rangle = \langle \vec{b}, \overrightarrow{BC} \rangle + \mu_b \langle \vec{\ell}_b, \overrightarrow{BC} \rangle$$

Daraus folgt:

$$\mu_b = \frac{\langle \vec{a} - \vec{b}, \vec{BC} \rangle}{\langle \vec{\ell}_b, \vec{BC} \rangle} = \frac{\langle \vec{BA}, \vec{BC} \rangle}{\langle \vec{\ell}_b, \vec{BC} \rangle} \quad \text{und weiter} \quad \vec{h}_{ab} = \vec{b} + \frac{\langle \vec{BA}, \vec{BC} \rangle}{\langle \vec{\ell}_b, \vec{BC} \rangle} \vec{\ell}_b.$$

Die genau gleiche Prozedur — mit den Rollen von C und A vertauscht — liefert den Ortsvektor \vec{h}_{cb} des Höhenschnittpunkt H_{cb} :

$$\vec{h}_{cb} = \vec{b} + \frac{\langle \vec{BC}, \vec{BA} \rangle}{\langle \vec{\ell}_b, \vec{BA} \rangle} \vec{\ell}_b.$$

Weiter gilt mit (*):

$$\langle \vec{\ell}_b, \vec{BC} \rangle - \langle \vec{\ell}_b, \vec{BA} \rangle = \langle \vec{\ell}_b, \vec{AC} \rangle = 0$$

und damit $\vec{h}_{ab} = \vec{h}_{cb}$.

(ii) Wir berechnen den Ortsvektor \vec{m}_{ab} des Schnittpunkts der Mittelsenkrechten auf a und b durch den folgenden Ansatz:

$$\vec{m}_{ab} = \frac{\vec{b} + \vec{c}}{2} + \mu_a \vec{\ell}_a = \frac{\vec{a} + \vec{c}}{2} + \mu_b \vec{\ell}_b.$$

Die Lote werden jetzt also von dem Mittelpunkt der Seiten aus gefällt. Wir bilden wieder das Skalarprodukt mit \vec{BC} und beachten dabei wieder (*):

$$\left\langle \frac{\vec{b} + \vec{c}}{2}, \vec{BC} \right\rangle = \left\langle \frac{\vec{a} + \vec{c}}{2}, \vec{BC} \right\rangle + \mu_b \langle \vec{\ell}_b, \vec{BC} \rangle$$

Daraus folgt:

$$\mu_b = \frac{1}{2} \cdot \frac{\langle \vec{AB}, \vec{BC} \rangle}{\langle \vec{\ell}_b, \vec{BC} \rangle} \quad \text{und weiter} \quad \vec{m}_{ab} = \frac{\vec{a} + \vec{c}}{2} + \frac{1}{2} \cdot \frac{\langle \vec{AB}, \vec{BC} \rangle}{\langle \vec{\ell}_b, \vec{BC} \rangle} \vec{\ell}_b.$$

Die genau gleiche Prozedur — mit den Rollen von C und A vertauscht — liefert den Ortsvektor \vec{m}_{cb} des Höhenschnittpunkt H_{cb} :

$$\vec{m}_{cb} = \frac{\vec{a} + \vec{c}}{2} + \frac{1}{2} \cdot \frac{\langle \vec{CB}, \vec{BA} \rangle}{\langle \vec{\ell}_b, \vec{BA} \rangle} \vec{\ell}_b.$$

Weiter gilt:

$$\langle \vec{\ell}_b, \vec{BC} \rangle - \langle \vec{\ell}_b, \vec{BA} \rangle = \langle \vec{\ell}_b, \vec{AC} \rangle = 0$$

und damit $\vec{m}_{ab} = \vec{m}_{cb}$. ◆