Lineare Algebra für Physiker – WS 2007/8

Axel Schüler

31. Januar 2008

Inhaltsverzeichnis

Kapitel 1

Lineare Gleichungssysteme

Literaturhinweise im Netz
Was kommt auf euch zu: nicht nur Algorithmen –> neue Begriffe, Definitionen, Sätze, Beweise
Verfahren bilden nur einen kleinen Teil der lin. Algebra.

1.1 Mengen und Abbildungen

1.1.1 Logische Operationen und Mengen

Wir werden die folgenden Symbole benutzen: Sind p und q Aussagen, so schreiben wir

$p \wedge q$	„p und q“,
$p \vee q$	„p oder q“,
$p \Rightarrow q$	p impliziert q bzw. „Wenn p, dann q. “
$p \iff q$	Äquivalenz bzw. „p genau dann, wenn q. “
$\forall x \colon \varphi(x)$	Für alle x gilt $\varphi(x)$.
$\exists a \colon \psi(a)$	Es gibt ein a, sodass $\psi(a)$ gilt.
$:=$	definierendes Gleichheitszeichen – die linke Seite wird durch die rechte definiert.

Bezeichnung von Mengen: $M = \{x, y, z, \dots\}$. Die Reihenfolge der Aufzählung spielt keine Rolle, $\{1,2,1,2\} = \{2,1\}$. Wir schreiben auch $M = \{x \in N \mid \varphi(x)\}$ für die Menge der Elemente x von N, welche eine Eigenschaft $\varphi(x)$ erfüllen. sind $A-1, A_2, \dots, A_n$ Mengen, so bezeichnet

$$A_1 \times A_2 \times \cdots \times A_n = \{(a_1, a_2, \dots, a_n) \mid \ \forall\, i = 1, \dots, n \colon a_i \in A_i\}$$

die Menge aller n-*Tupel* von Elementen aus $A_1, \dots, A_n$ (kartesisches Produkt). Im Spezialfall $A_i = \mathbb{R}$ für alle $i = 1, \dots, n$ schreiben wir

$$\mathbb{R}^n := \mathbb{R} \times \cdots \times \mathbb{R} \quad \text{gesprochen: „Err Enn“.}$$

Sei nun $\{A_i \mid i \in I\}$ eine beliebige Familie von Teilmengen einer Menge X, dann heißen

$$\bigcap_{i \in I} A_i := \{x \in X \mid \ \forall\, k \in I \colon x \in A_k\}, \quad \bigcup_{i \in I} A_i := \{x \in X \mid \ \exists\, n \in I \colon x \in A_n\}$$

Durchschnitt und *Vereinigung* der Familie von Mengen. Es gelten die *de Morganschen Regeln,*

$$\left(\bigcap_{i\in I} A_i\right)^{\mathsf{c}} = \bigcup_{i\in I} A_i^{\mathsf{c}}, \quad \left(\bigcup_{i\in I} A_i\right)^{\mathsf{c}} = \bigcap_{i\in I} A_i^{\mathsf{c}}.$$

1.1.2 Vollständige Induktion

Es sei $A(n)$ eine Aussage über natürliche Zahlen; $n \in \mathbb{N}$.

Satz 1.1 (Prinzip der vollständigen Induktion) *Es sei $n_0 \in \mathbb{N}$. Um eine Aussage $A(n)$ für alle $n \geq n_0$ zu beweisen, genügt es zu zeigen, dass*

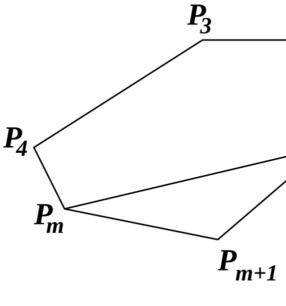
(IA) Induktionsanfang: $A(n_0)$ gilt.
(IS) Induktionsschritt: $\forall\, k \in \mathbb{N}, k \geq n_0 : A(k) \Longrightarrow A(k+1)$.

Beispiel 1.1 (a) $\forall\, n \in \mathbb{N}: \sum_{i=1}^{n} i = \frac{1}{2} n(n+1)$.
Beweis. Durch vollständige Induktion über n.
Induktionsanfang: $n = 1$. $\sum_{i=1}^{1} i = 1 = \frac{1}{2} 1 \cdot 2$ ist erfüllt.
Induktionsvoraussetzung: Die Behauptung gelte für $n = m$: $\sum_{i=1}^{m} i = \frac{1}{2} m(m+1)$.
Induktionsbehauptung: $\sum_{i=1}^{m+1} i = \frac{1}{2}(m+1)(m+2)$.
Induktionsbeweis: Es gilt

$$\sum_{i=1}^{m+1} i = \sum_{i=1}^{m} i + (m+1) \underset{\text{IV}}{=} \frac{1}{2} m(m+1) + (m+1) = (m+1)\left(\frac{m}{2} + 1\right) = \frac{1}{2}(m+1)(m+2).$$

∎

(b) Die Innenwinkelsumme in einem n-Eck beträgt $(n-2)180°$.



Beweis. durch vollständige Induktion über n. Für $n = 3$ ist die Behauptung klar, denn das Dreieck hat die Innenwinkelsumme von 180°. Die Behauptung gelte für jedes m-Eck. Wir zeigen, dass jedes $(m+1)$-Eck die Innenwinkelsumme $((m+1)-2)180°$ besitzt. Dazu zeichnen wir die Diagonale $\overline{P_1 P_n}$ ein, die das $(m+1)-Eck$ in ein m-Eck und ein Dreieck zerllegt.

Die Innenwinkelsumme w im $(m+1)$-Eck ist offenbar die Innenwinkelsumme im m-Eck $P_1 P_2 \cdots P_m$ plus die Innenwinkelsumme in Dreieck $P_m P_{m+1} P_1$. Folglich gilt für w nach Induktionsvoraussetzung

$$w = (m-2)180° + 180° = (m-1)180° = ((m+1)-2)180°.$$

Die Induktionsbehauptung ist gezeigt, also gilt die Behauptung für alle $n \in \mathbb{N}$.

1.1.3 Summen- und Produktzeichen

Wir wollen die Bedeutung von Summen- und Produktzeichen, $\sum$ und $\prod$, wiederholen. Es seine $m \le n$ ganze Zahlen und a_k, $k = m, \dots, n$ beliebige reelle Zahlen. Dann setzen wir

$$\sum_{k=m}^{n} a_k := a_m + a_{m+1} + \cdots + a_n, \qquad \prod_{k=m}^{n} a_k := a_m a_{m+1} \cdots a_n.$$

Im Falle $m = n$ bestehen Summe und Produkt nur aus einem einzigen Summanden bzw. aus einem einzigen Faktor Im Falle $n < m$ setzen wir

$$\sum_{k=m}^{n} a_k := 0, \text{(leere Summe)} \qquad \prod_{k=m}^{n} a_k := 1 \quad \text{(leeres Produkt)}.$$

Es gelten die folgenden Rechenregeln Falls $m \le n \le p$ und $d \in \mathbb{Z}$,

$$\sum_{k=m}^{n} a_k + \sum_{k=n+1}^{p} a_k = \sum_{k=m}^{p} a_k \quad \text{(Verkettung, Assoziativgesetz)}$$

$$\sum_{k=m}^{n} a_k = \sum_{k=m+d}^{n+d} a_{k-d} \quad \sum_{k=m}^{n} a_k = \sum_{k=m-d}^{n-d} a_{k+d} \quad \text{(Indexverschiebung)},$$

$$\sum_{k=1}^{n} a_k + \sum_{k=1}^{n} b_k = \sum_{k=1}^{n} (a_k + b_k) \quad \text{(Kommutativgesetz)}.$$

Für $a \in \mathbb{R}$ und $n \ge m$ haben wir $\displaystyle\sum_{k=m}^{n} a = (n - m + 1)a$.

Beispiel 1.2 (a) Teleskopsummen.

$$\sum_{k=5}^{100} (a_k - a_{k-2}) = \sum_{k=5}^{100} a_k - \sum_{k=3}^{98} a_k = -a_3 - a_4 + a_{99} + a_{100}.$$

(b) $\prod_{k=1}^{n} k =: n!$ (lies „n Fakultät"). In Beispiel 1.1 haben wir bewiesen, dass $\sum_{k=1}^{n} k = n(n+1)/2$.

(c) Definition von $\binom{n}{k}$ für $n \ge k$ natürliche Zahlen als $\binom{n}{k} = \frac{n!}{k!(n-k)!}$.
Eigenschaften: $\binom{n}{0} = \binom{n}{n} = 1$, $\binom{n}{1} = \binom{n}{n-1} = n$, $\binom{n}{k} = \binom{n}{n-k}$ (nach Definition), $\binom{n}{k} = \frac{n(n-1)\cdots(n-k+1)}{k!}$. Wir definieren $\binom{n}{m} = 0$, falls $m > n$.
Eigenschaft des Pascalschen Dreiecks: $\binom{n+1}{k+1} = \binom{n}{k} + \binom{n}{k+1}$.
(d) Doppelsummen. Es seien a_{ij}, $i = 1, \dots, m$, $j = 1, \dots, n$ gegeben. Dann gilt:

$$\sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} = \sum_{j=1}^{m} \sum_{i=1}^{n} a_{ij}.$$

Das heißt, Doppelsummen kann man vertauschen.
Beweis. Durch vollständige Induktion über m.
Induktionsanfang: $m = 1$. Dann gilt $\sum_{i=1}^{1} \sum_{j=1}^{n} a_{ij} = \sum_{j=1}^{n} a_{1j} = \sum_{j=1}^{n} \sum_{i=1}^{1} a_{ij}$.

Die Behauptung gelte nun für ein festes m (und alle $n \in \mathbb{N}$). Wir zeigen, dass die Behauptung dann auch für $m+1$ gilt. Nach Definition ist nämlich

$$\sum_{i=1}^{m+1}\sum_{j=1}^{n} a_{ij} = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij} + \sum_{j=1}^{n} a_{m+1,j} \underset{\mathrm{IV}}{=} \sum_{j=1}^{n}\sum_{i=1}^{m} a_{ij} + \sum_{j=1}^{n} a_{m+1,j} = \sum_{j=1}^{n}\left(\sum_{i=1}^{m} a_{ij} + a_{m+1,j}\right) = \sum_{j=1}^{n}\sum_{i=1}^{m+1} a_{ij}.$$

■

1.1.4 Abbildungen

Wir beginnen mit einem der wichtigsten Begriffe der Mathematik, der erst im 18. und 19. Jahrhundert fixiert wurde. Maßgeblich beteiligt waren daran Leonard Euler (1727) und Dirichlet (1837). Von Euler stammt das noch heute verwendete Symbol $f(x)$.
Motivation: Permutationen, Kurven, Flächen, Simplexe, Matrizen sind in erster Linie *Abbildungen*. Erst durch diese Sichtweise kann man Permutationen zu Gruppen machen, Tangentialvektoren an Kurven aufschreiben, die Krümmung einer Fläche berechnen, vom Rang der Matrix reden.

Definition 1.1 Es seien X und Y Mengen. Eine *Abbildung* oder *Funktion* f von X nach Y ist eine Vorschrift, die jedem $x \in X$ genau ein $f(x) \in Y$ zuordnet. Dafür schreiben wir $f: X \to Y$ oder auch $X \ni x \mapsto f(x) \in Y$.

Wichtig: Zu jeder Abbildung gehören immer zwei Mengen X und Y sowie die genannte Vorschrift.

Beispiel 1.3 (a) Es sei M eine Menge. Dann heißt die Abbildung $\mathrm{id}_M: M \to M$ mit $\mathrm{id}_M(x) = x$ für alle $x \in M$ die *identische Abbildung* auf M.
(b) Es seien X, Y Mengen und $y_0 \in Y$. Dann nennt man $f: X \to Y$ mit $f(x) = y_0$ eine *konstante* Abbildung.

Definition 1.2 Es sei $f: X \to Y$ eine Abbildung und $A \subset X$, $B \subset Y$.
Dann heißen die Mengen

$$f(A) := \{f(x) \mid x \in A\} \qquad \text{und} \qquad f^{-1}(B) := \{x \in X \mid f(x) \in B\}$$

Bild von A bzw. *Urbild von B*.

Bemerkung 1.1 Es ist klar, dass $f(A) \subset Y$ und $f^{-1}(B) \subset X$ gilt. Man beachte, dass f^{-1} in keiner Weise eine „Umkehrabbildung“ meint; das Symbol f^{-1} ohne (B) hat keinen Sinn. Beispiel: $f: \mathbb{R} \to \mathbb{R}$, $f(x) = \mathrm{e}^x$. Es gilt $f(\{0,1\}) = \{1, \mathrm{e}\}$, $f((-\infty, 0]) = (0,1]$, $f^{-1}([-2,-1]) = \emptyset$, $f^{-1}(\{0,1,2\}) = \{0, \ln(2)\}$.

Definition 1.3 Es sei $f: X \to Y$ eine Abbildung.
(a) f heißt *injektiv*, wenn keine zwei Elemente von X auf dasselbe Element von Y abgebildet werden.
(b) f heißt *surjektiv*, wenn jedes Element $y \in Y$ ein Urbild in X besitzt.
(c) f heißt *bijektiv*, wenn f injektiv und surjektiv ist.

Statt f surjektiv, sagt man auch, dass f eine Abbildung *auf* Y ist. Als Formel: f ist surjektiv gdw.

$$\boxed{f(X) = Y.}$$

Als Formel: f ist injektiv gdw.

$$\boxed{\forall x_1, x_2 \in X\colon f(x_1) = f(x_2) \quad \Longrightarrow \quad x_1 = x_2.}$$

Genau dann, wenn $f: X \to Y$ bijektiv ist, besitzt f eine *Umkehrabbildung* $f^{-1}: Y \to X$ mit $f^{-1}(y) := x$, falls $f(x) = y$. Zu jedem $y \in Y$ gibt es ein solches x, da f surjektiv ist. Dieses x ist eindeutig bestimmt, da f injektiv ist.

Beispiel 1.4 (a) id_M ist stets bijektiv. Die konstante Abbildung aus Beispiel 1.3 ist surjektiv gdw. $Y = \{y_0\}$ und injektiv gdw. $|X| = 1$.
(b) Sei $f: [-\pi/2, \pi/2] \to [-1, 1]$ gegeben durch $f(x) = \sin x$. Dann ist f bijektiv, denn alle Werte y mit $-1 \le y \le 1$ werden im Intervall $-\pi/2 \le x \le \pi/2$ genau einmal angenommen.
(c) Die Betragsfunktion $|\cdot|: \mathbb{R} \to \mathbb{R}$, $|x| = \begin{cases} x, & x \ge 0 \\ -x, & x < 0 \end{cases}$ ist weder injektiv noch surjektiv, denn $|-1| = |1|$ und $|x| = -1$ hat keine Lösung.

Aufgabe 1. Es sei $f: M \to N$ eine Abbildung zwischen zwei endlichen Mengen M und N mit derselben Anzahl von Elementen $|M| = |N| = n$.
Zeigen Sie, f ist injektiv genau dann, wenn f surjektiv.

Definition 1.4 Es seien $f: X \to Y$ und $g: Y \to Z$ Abbildungen. Dann ist die *Komposition* oder *zusammengesetzte Abbildung* $g \circ f: X \to Z$ definiert für alle $x \in X$ gemäß $g \circ f(x) = g(f(x))$.

$X \xrightarrow{f} Y \xrightarrow{g} Z$

Beispiel 1.5 (a) Es seinen $f, g: \mathbb{R} \to \mathbb{R}$ gegeben durch $f(x) = \sin x$ und $g(z) = \mathrm{e}^z$.
Dann ist die Komposition $g \circ f(x) = \mathrm{e}^{\sin x}$ und $f \circ g(z) = \sin(\mathrm{e}^z)$. Die Komposition ist i.a. nicht kommutativ.
(b) Ist $f: X \to Y$ bijektiv mit der Umkehrabbildung $f^{-1}: Y \to X$, so gilt $f^{-1} \circ f = \mathrm{id}_X$ und $f \circ f^{-1} = \mathrm{id}_Y$.
(c) Sind f und g beide injektiv, surjektiv bzw. bijektiv, dann ist auch $g \circ f$ injektiv, surjektiv bzw. bijektiv. **Beweis in der Übung**

$$\begin{tikzcd} X \arrow[r, "f"] \arrow[dr, "h"'] & Y \arrow[d, "g"] \\ & Z \end{tikzcd}$$

Hat man mit mehreren Abbildungen zwischen verschiedenen Mengen zu tun, so ist es oft übersichtlicher, sie in einem Diagramm anzuordnen: so kann man etwa die Abbildungen $f: X \to Y$, $g: Y \to Z$ und $h: X \to Z$ wie im nebenstehenden Diagramm anordnen.

Wenn in einem Diagramm alle zusammengesetzten Abbildungen zwischen beliebigen zwei Mengen des Diagramms übereinstimmen, so sagt man, dass das Diagramm *kommutativ* ist. So ist etwa das obige Diagramm kommutativ, genau dann, wenn h die Komposition von f und g ist, $h = g \circ f$.

Lemma 1.2 (Assoziativität der Komposition) *Es seien $f: X \to Y$, $g: Y \to Z$ und $h: Z \to U$ Abbildungen. Dann gilt $h \circ (g \circ f) = (h \circ g) \circ f$.*

$X \xrightarrow{f} Y \xrightarrow{g} Z \xrightarrow{h} U$.

Beweis. Zwei Abbildungen sind genau dann gleich, wenn ihre Definitions- und Wertebereiche übereinstimmen und wenn die Abbildungsvorschriften gleich sind. Das erste ist sicherlich erfüllt, da beide Abbildungen von X nach U abbilden. Ferner gilt für alle $x \in X$:

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

Damit gilt die Gleichheit. ∎

Da es egal ist, wie wir die Klammern setzten, schreiben wir für diese Abbildung $g \circ h \circ f$.

1.2 Matrizen

In diesem Abschnitt sind $m, n \in \mathbb{N}$ und mit Zahlen meinen wir die Elemente eines fixierten Körpers $\mathbb{K}$, etwa $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$ oder $\mathbb{K} = \mathbb{Q}$.
Eine *$m \times n$-Matrix* ist ein rechteckiges Zahlenschema

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

mit m *Zeilen* und n *Spalten*. Das Zahlenpaar (m, n) heißt *Typ* der Matrix. Die Zahlen a_{ij} heißen *Einträge* der Matrix. Der Eintrag a_{kl} steht in der $k$ten Zeile und $l$ten Spalte der Matrix. Wir schreiben auch

$$A = (a_{ij})_{\substack{i=1,\dots,m \\ j=1,\dots,n}}$$

Eine Matrix aus nur einer Zeile heißt auch *Zeilenvektor*, eine Matrix aus nur einer Spalte heißt *Spaltenvektor*. Zwei Matrizen sind genau dann gleich, wenn sie den gleichen Typ haben und in allen Einträgen übereinstimmen.
Die Menge der $m \times n$-Matrizen mit Einträgen aus $\mathbb{K}$ bezeichnen wir mit $\mathbb{K}^{m \times n}$.

1.2.1 Operationen mit Matrizen

(a) Addition und Vervielfachung von Matrizen

Es seien $A = (a_{ij})$ und $B = (b_{ij})$ zwei Matrizen vom selben Typ (m, n). Dann definieren wir die *Summe* von A und B als

$$A + B = \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \cdots & a_{n1} + b_{n1} \\ a_{21} + b_{21} & a_{22} + b_{22} & \cdots & a_{2n} + b_{2n} \\ \vdots & & & \vdots \\ a_{m1} + b_{m1} & a_{m2} + b_{m2} & \cdots & a_{mn} + b_{mn} \end{pmatrix}$$

Kurzschreibweise: $A + B = (a_{ij} + b_{ij})_{\substack{i=1...m \\ j=1...m}}$. Zwei Matrizen werden addiert, indem man alle Eintragungen an denselben Positionen addiert. Eine *Nullmatrix*, 0, ist eine Matrix deren sämtliche Einträge gleich Null sind. Ändert man bei allen Einträgen einer Matrix A das Vorzeichen, so erhält man die Matrix $-A = (-a_{ij})$.
Ist $\lambda \in \mathbb{K}$, so definieren wir das *Vielfache* von A als $\lambda A := (\lambda a_{ij})$, das heißt, alle Einträge von A werden mit λ multipliziert. Insbesondere gilt $(-1)A = -A$.

(b) Eigenschaften von Addition und Vervielfachung

Die folgenden Eigenschaften sind genau die eines *Vektorraumes*. Die Matrizen fungieren somit als *das* Standardbeispiel für Vektorräume.
Für beliebige Matrizen A, B , C vom Typ (m, n) und alle $\lambda, \mu \in \mathbb{K}$ gilt:

1. $A + (B + C) = (A + B) + C$ (Assoziativität)
2. $A + B = B + A$ (Kommutativität)
3. $A + 0 = 0 + A = A$ (0 ist neutral)
4. $A + (-A) = 0$ (Existenz des Inversen)
5. $(\lambda(\mu A)) = (\lambda\mu)A$
6. $(\lambda + \mu)A = \lambda A + \mu A$
7. $\lambda(A + B) = \lambda A + \lambda B$
8. $1 \cdot A = A$.

Beweis. Wir zeigen modellhaft die Eigenschaft 7. Zum Nachweis müssen wir zeigen, dass die Typen der Matrizen rechts und links gleich sind, nämlich gleich (m, n) und dass die Matrixeinträge an entsprechenden Stellen (i, j), $i = 1, \ldots, m$, $n = 1, \ldots, n$, übereinstimmen:

$$(\lambda(A + B))_{ij} = \lambda(A + B)_{ij} = \lambda(a_{ij} + b_{ij}).$$

Die rechte Seite liefert dagegen:

$$(\lambda A + \lambda B)_{ij} = (\lambda A)_{ij} + (\lambda B)_{ij} = \lambda a_{ij} + \lambda b_{ij}.$$

Damit ist die Gleichheit gezeigt. ∎

1.2.2 Das Produkt von Matrizen

Es seien $A=(a_{ij})$ und $B=(b_{kl})$ Matrizen vom Typ (m,n) bzw. (p,q). Voraussetzung für die Berechnung des Produktes ist $n=p$, das heißt, die Spaltenzahl n von A muss mit der Zeilenzahl p von B übereinstimmen. Man sagt, A und B sind *verkettet*. Dann existiert das *Produkt* $C=AB$ der Matrizen und hat den Typ (m,q), wobei $C=(c_{il})_{\substack{i=1\dots m\\ l=1\dots p}}$ gegeben ist durch

$$c_{il}=\sum_{j=1}^{n} a_{ij}\, b_{jl}.$$

Bemerkung 1.2 (a) In der obigen Situation muss das Produkt BA nicht unbedingt definiert sein, nur wenn zusätzlich $q=m$ gilt, ist das der Fall.
(b) Für zwei quadratische Matrizen A, B vom Typ (n,n) sind AB und BA immer definiert.
(c) Eine quadratische Matrix, bei der auf der Hauptdiagonalen alles Einsen stehen und sonst alles Nullen

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

heißt *Einheitsmatrix* der Ordnung n. Für ihre Matrixelemente gilt

$$(I_n)_{ij}=\begin{cases}0, & i\neq j,\\ 1, & i=j\end{cases} =: \delta_{ij}, \quad i,j=1,\dots,n.$$

Man bezeichnet δ_{ij} als *Kronecker-Symbol*.

Beispiel 1.6 (a)

$$\begin{pmatrix}2 & 3\\ 1 & -1\end{pmatrix}\begin{pmatrix}1 & -1 & 4\\ 5 & 2 & 3\end{pmatrix}=\begin{pmatrix}17 & 4 & 17\\ -4 & -3 & 1\end{pmatrix}.$$
$$\text{Typ}: \quad (2,2)\times(2,3)=(2,3)$$

$$\begin{pmatrix}1 & 2\end{pmatrix}\begin{pmatrix}3\\ 4\end{pmatrix}=11, \qquad \begin{pmatrix}3\\ 4\end{pmatrix}\begin{pmatrix}1 & 2\end{pmatrix}=\begin{pmatrix}3 & 6\\ 4 & 8\end{pmatrix},$$
$$\text{Typ}: \quad (1,2)\times(2,1)=(1,1) \quad (2,1)\times(1,2)=(2,2).$$

(b)

$$\begin{pmatrix}a_{11} & a_{12} & a_{13}\\ a_{21} & a_{22} & a_{23}\end{pmatrix}\begin{pmatrix}x_1\\ x_2\\ x_3\end{pmatrix}=\begin{pmatrix}a_{11}x_1+a_{12}x_2+a_{13}x_3\\ a_{21}x_1+a_{22}x_2+a_{23}x_3\end{pmatrix}$$

Eigenschaften der Matrixmultiplikation

1. Das Matrixprodukt ist *nicht* kommutativ. Im Allgemeinen sind AB und BA nicht gleichzeitig definiert. Selbst wenn, sind die Produkte i.a. nicht gleich:

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

2. Das Produkt von Null verschiedener Matrizen kann Null sein:
$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = 0.$$

3. Die Matrixmultiplikation ist assoziativ, das heißt, sind $A=(a_{ij})$, $B=(b_{kl})$ und $C=(c_{gh})$ Matrizen vom Typ (m,n), (n,r) bzw. (r,s), dann existieren die Produkte AB, BC, $A(BC)$ und $(AB)C$ und es gilt $A(BC)=(AB)C$.
Beweis. Die Existenz der Produkte folgt aus der Verkettung von A und B und von B und C. Wir zeigen, dass die Matrizen $(AB)C$ und $A(BC)$ an allen Einträgen überein stimmen. Für $D=AB=(d_{il})$ gilt $d_{il}=\sum_{j=1}^{n} a_{ij}b_{jl}$ und für $E=BC=(e_{jh})$ gilt $e_{jh}=\sum_{l=1}^{r} b_{jl}c_{lh}$. Daher ist
$$((AB)C)_{ih}=(DC)_{ih}=\sum_{l=1}^{r} d_{il}c_{lh}=\sum_{l=1}^{r}\sum_{j=1}^{n}(a_{ij}b_{jl})c_{lh}.$$
Andererseits ist
$$(A(BC))_{ih}=(AE)_{ih}=\sum_{j=1}^{n} a_{ij}e_{jh}=\sum_{j=1}^{n}\sum_{l=1}^{r} a_{ij}(b_{jl}c_{lh}).$$
Wegen der Assoziativität der Zahlenmultiplikation (im Körper $\mathbb{K}$) und der Vertauschbarkeit der Summationen über l und j, siehe Beispiel 1.1 (d), gilt $(AB)C=A(BC)$. ■

4. Die Matrixmultiplikation ist distributiv, das heißt, sind B und C jeweils vom Typ (n,p) und ist A vom Typ (m,n), so gilt $A(B+C)=AB+AC$.
Sind B und C jeweils vom Typ (m,n) und A vom Typ (n,p), so ist $(B+C)A=BA+CA$. (Beweis: Übungsaufgabe 2.1).

5. Für $A\in\mathbb{K}^{m\times n}$, $A=(a_{ij})$ definieren wir die zu A *transponierte* Matrix $A^\top\in\mathbb{K}^{n\times m}$ durch Vertauschen von Zeilen und Spalten: Für alle $i=1,\dots,m$, $j=1,\dots,n$ setzt man
$$\left(A^\top\right)_{ji}=a_{ij}.$$
So ist z. B. $(x_1\ \ x_1\ \ x_3)^\top=\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ und $(I_n)^\top=I_n$ und
$$\begin{pmatrix} 1 & 2 & -1 \\ 0 & 5 & b \end{pmatrix}^\top=\begin{pmatrix} 1 & 0 \\ 2 & 5 \\ -1 & b \end{pmatrix}.$$

6. Ist A vom Typ (m,n), so gilt für das Produkt mit den entsprechenden Einheitsmatrizen $AI_n=A=I_mA$. In der Tat gilt für alle i,j, dass
$$(AI_n)_{ij}=\sum_{k=1}^{n} a_{ik}\delta_{kj}=a_{i1}\delta_{1j}+\dots+a_{ij}\delta_{jj}+\dots+a_{in}\delta_{nj}=a_{ij}=(A)_{ij}.$$

7. *Eigenschaften des Transponierens.* Für beliebige Matrizen A, B und $\lambda, \mu \in \mathbb{K}$ gilt (1) $(A^\top)^\top = A$, (2) $(\lambda A + \mu B)^\top = \lambda A^\top + \mu B^\top$, (3) $(AB)^\top = B^\top A^\top$, vgl. ÜA 2.1.

1.2.3 Lineare Gleichungssysteme

Das Studium linearer Gleichungssysteme ist eines der wichtigsten Themen der linearen Algebra.
Eine Gleichung der Form $a_1x_1 + \dots + a_nx_n = b$ mit konstanten Koeffizienten $a_1, \dots, a_n, b \in \mathbb{K}$ heißt *lineare Gleichung* mit n Variablen. Alle Variablen kommen in der ersten Potenz vor. Die Gleichungen $xy = 1$, $x + y^2 = 5$ und $\sin x_1 - \cos x_2 = 1$ sind nichtlinear. Hingegen ist für eine Konstante $k \in \mathbb{R}$ die Gleichung $\sqrt{2}x + k^2y + \sin(k)z = \frac{1}{k}$ linear, da die Variablen x, y, z alle in der ersten Potenz auftreten. Im allgemeinsten Fall sind nun mehrere solche linearen Gleichungen gegeben, etwa mit m Gleichungen und n Variablen:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n &= b_2 \\ &\dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n &= b_m. \end{aligned} \tag{1.1}$$

Die Matrix $A = (a_{ij})_{\substack{i=1\dots m\\ j=1\dots n}}$ heißt *Koeffizientenmatrix* des Gleichungssystems. Hängt man *den Spaltenvektor der rechten Seite* $b := (b_1, \dots, b_n)^\top$ an diese Matrix an, so erhält man die *erweiterte Koeffizientenmatrix*

$$(A, b) := \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{pmatrix} \tag{1.2}$$

Diese hat den Typ $(m, n+1)$. Mitunter trennen wir zur besseren Übersicht bei linearen Gleichungssystemen die letzte Spalte mit einem senkrechten Strich ab. Der Vektor

$$x := \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

heißt *Spaltenvektor der Unbestimmten.* Mit Hilfe der Matrixmultiplikation kann man nun das GS schreiben als eine einzige Gleichung $A \cdot x = b$.
In der Tat sind die Matrizen A vom Typ (m, n) und x vom Typ $(n, 1)$ verkettet, sodass das Produkt existiert und den Typ $(m, 1)$ hat, An der Stelle $(i, 1)$, $i = 1, \dots, m$, des Produktes steht auf der linken Seite:

$$(A \cdot x)_{i1} = \sum_{j=1}^{n} a_{ij}x_j,$$

wogegen auf der rechten Seite b_i steht. Die Gleichung für die $i$te Zeile lautet also

$$\sum_{j=1}^{n} a_{ij} x_j = b_i.$$

Wenn man aber die obige Summe ausschreibt — als Summe von Summanden — so hat man

$$a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n = b_i.$$

1.3 Lineare Gleichungssysteme

1.3.1 Die Cramersche Regel I

Wir betrachten das lineare Gleichungssystem

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= b_1 \quad |\cdot a_{21} \\ a_{21}x_1 + a_{22}x_2 &= b_2. \quad |\cdot a_{11} \end{aligned} \tag{1.3}$$

Subtrahiert man nach Multiplikation die erste Gleichung von der zweiten, so hat man

$$x_2(-a_{12}a_{21} + a_{11}a_{22}) = b_2 a_{11} - b_1 a_{21}.$$

Wenn nun $D := a_{11}a_{22} - a_{12}a_{21} \neq 0$, kann man dividieren und erhält

$$x_2 = \frac{1}{D}(b_2 a_{11} - b_1 a_{21}), \quad x_1 = \frac{1}{D}(b_1 a_{22} - b_2 a_{12}). \tag{1.4}$$

Umgekehrt erfüllen die angegebenen Zahlen x_1 und x_2 das Gleichungssystem (1.3). Zusammengefasst, wenn $D = a_{11}a_{22} - a_{12}a_{21} \neq 0$, dann hat (1.3) eine eindeutig bestimmte Lösung.; diese ist durch (1.4) gegeben. Es sei $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ eine 2×2-Matrix. Wir definieren

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} := \begin{vmatrix} a & b \\ c & d \end{vmatrix} := ad - bc$$

und nennen diese Zahl die *Determinante* der Matrix A. Dann schreibt sich die Lösung (1.4) als

$$x_1 = \frac{\begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}} \quad x_2 = \frac{\begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}.$$

Dieses Lösungsverfahren nennt man *Cramersche Regel*. Zum Beispiel erhält man die Lösung von

$$\begin{aligned} 3x_1 + 2x_2 &= 5, \\ 5x_1 + 4x_2 &= 7 \end{aligned}$$

gemäß der Cramerschen Regel über

$$x_1 = \frac{\begin{vmatrix} 5 & 2 \\ 7 & 4 \end{vmatrix}}{\begin{vmatrix} 3 & 2 \\ 5 & 4 \end{vmatrix}} = \frac{20-14}{12-10} = 3, \quad x_2 = \frac{\begin{vmatrix} 3 & 5 \\ 5 & 7 \end{vmatrix}}{\begin{vmatrix} 3 & 2 \\ 5 & 4 \end{vmatrix}} = \frac{21-25}{2} = -2.$$

Man beachte, dass bei den beiden Gleichungssystemen

$$\begin{aligned} 3x_1 + 2x_2 &= 5, \qquad & 3x_1 + 2x_2 &= 5 \\ 6x_1 + 4x_2 &= 7, \qquad & 6x_1 + 4x_2 &= 10 \end{aligned}$$

die Koeffizientendeterminante $\begin{vmatrix} 3 & 2 \\ 6 & 4 \end{vmatrix} = 0$ gleich Null ist; die Cramersche Regel ist somit nicht anwendbar. Das erste System ist widersprüchlich, es besitzt keine Lösung; das zweite System besitzt unendlich viele Lösungen.

Geometrische Beschreibung. Im ersten Fall haben die beiden Geraden $3x_1 + 2x_2 = 5$, $5x_1 + 4x_2 = 7$ genau einen Schnittpunkt $(3, -2)$; im zweiten Fall sind die Geraden $3x_1 + 2x_2 = 5$, $6x_1 + 4x_2 = 7$ parallel und haben daher keinen Punkt gemeinsam, die Lösungsmenge ist leer. Im dritten Fall schließlich fallen die beiden Geraden zusammen — es gibt unendlich viele Lösungen.

Wir werden sehen, dass dieses Lösungsverhalten — genau eine Lösung, keine Lösung, unendlich viele Lösungen — allen linearen Gleichungssystemen zugrunde liegt.

1.3.2 Der Gauß-Algorithmus

In diesem Abschnitt stellen wir **das** Verfahren zur praktischen Lösung von Gleichungssystemen vor. Es hat aber noch viel mehr Anwendungen in der linearen Algebra: Berechnung des Ranges einer Matrix, Berechnung der Determinante, Berechnung der Inversen einer Matrix, Bestimmung der Dimension eines linearen Unterraumes, Berechnung von Eigenvektoren, Berechnung der Fläche von Dreiecken, Überprüfen der Kollinearität von Punkten im Raum usw.

Eine sehr ausführliche Beschreibung des Algorithmus findet man in [Ant98, Kapitel 1].

(a) Elementare Zeilenoperationen Gegeben sei ein lineares Gleichungssystem mit m Gleichungen und n Variablen

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m. \end{aligned} \tag{1.5}$$

Verändert man das Gleichungssystem dadurch, dass man

(a) zwei Gleichungen vertauscht,

(b) eine Gleichung mit einer von Null verschiedenen Zahl λ multipliziert,

(c) ein Vielfaches einer Zeile zu einer anderen Zeile addiert ($z'_j := z_j + \mu z_i$, $z'_i = z_i$),

so ändert sich die Lösungsmenge **nicht:** offenbar sind alle Lösungen des alten Systems auch wieder Lösungen des neuen Systems. Da die drei Operationen rückgängig gemacht werden können (Rücktausch, Multiplikation mit $1/\lambda$, Subtraktion des entsprechenden Vielfaches von einer Zeile, $z_j = z'_j - \mu z_i$), sind alle Lösungen des neuen Systems auch Lösungen des alten Systems. Diese Beobachtung liegt dem Gauß-Algorithmus zugrunde. Diese drei Umformungen nennen wir *elementare Zeilenumformungen.* Wir bezeichnen die Menge der Lösungen $x = (x_1, \ldots, x_n)$ des Gleichungssystems $Ax = b$ mit $\text{Lös}(A, b)$.
Zusammengefasst: Verändert man die erweiterte Koeffizientenmatrix (A, b) durch eine der drei elementaren Zeilenumformungen (a), (b) oder (c), zu einer Matrix (A', b'), so bleibt die Lösungsmenge unverändert, $\boxed{\text{Lös}(A, b) = \text{Lös}(A', b').}$

(b) Die reduzierte Zeilenstufenform einer erweiterten Koeffizientenmatrix. Wir streben die folgende Form, die *reduzierte Zeilenstufenform* der Matrix an.

Beispiel 1.7

$$\begin{pmatrix} 1 & 6 & 0 & 0 & 4 & | & -2 \\ 0 & 0 & 1 & 0 & 3 & | & 1 \\ 0 & 0 & 0 & 1 & 5 & | & 2 \\ 0 & 0 & 0 & 0 & 0 & | & 0 \end{pmatrix}$$

Das dieser Matrix zugeordnete Gleichungssystem lautet

$$\begin{array}{llllll} x_1 & +6x_2 & & & +4x_5 & = -2 \\ & & x_3 & & +3x_5 & = 1 \\ & & & x_4 & +5x_5 & = 2. \end{array}$$

Insbesondere können Nullzeilen gestrichen werden. Die führenden Variablen sind x_1, x_3 und x_4 während x_2 und x_5 frei wählbar sind. Auflösen nach x_1, x_3 und x_4 liefert:

$$\begin{aligned} x_1 &= -2 - 6x_2 - 4x_5 \\ x_3 &= 1 - 3x_5 \\ x_4 &= 2 - 5x_5. \end{aligned}$$

Diese Lösung lässt sich auch schreiben als

$$\text{Lös}(A, b) = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} -2 \\ 0 \\ 1 \\ 2 \\ 0 \end{pmatrix} + s \begin{pmatrix} -6 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + t \begin{pmatrix} -4 \\ 0 \\ -3 \\ -5 \\ 1 \end{pmatrix}, \quad s, t \in \mathbb{R} \right\}.$$

Es handelt sich um eine *Ebene* im $\mathbb{R}^5$ durch den Punkt $(-2, 0, 1, 2, 0)^\top$ mit den beiden Spannvektoren $v_1 = (-6, 1, 0, 0, 0)^\top$ und $v_2 = (-4, 0, -3, -5, 1)^\top$.
Es gibt unendlich viele Lösungen mit den zwei freien Parametern $x_2 = s$ und $x_5 = t$.

Die reduzierte Zeilenstufenform ist durch folgende Eigenschaften charakterisiert:

1. Wenn eine Zeile nicht nur aus Nullen besteht, so ist die erste von Null verschiedene Zahl eine Eins. Sie wird als *führende Eins* bezeichnet.
2. Alle Zeilen, die nur Nullen enthalten stehen am Ende der Matrix.
3. In zwei aufeinanderfolgenden Zeilen, die nicht nur Nullen enthalten, steht die führende Eins der unteren Zeile rechts von der führenden Eins der oberen Zeile.
4. Eine Spalte, die eine führende Eins enthält, hat keine weiteren von Null verschiedenen Einträge.

Eine Matrix, die nur die Bedingungen 1. bis 3. erfüllt, liegt in *Zeilenstufenform* vor. Auch hier ist schon die Lösung des GS ablesbar. Die Spalten der führenden Einsen entsprechen abhängigen Variablen, die Spalten ohne führende Einsen entsprechen unabhängigen, frei wählbaren Variablen.

(c) Gauß-Jordan-Algorithmus für $m \times n$ lineare Gleichungssysteme $Ax = b$ mit $A \in \mathbb{K}^{m\times n}$ und $b \in \mathbb{K}^m$. Ziel ist es, durch elementare Zeilenumformungen aus der erweiterten Koeffizientenmatrix (A, b) eine Matrix in reduzierter Zeilenstufenform herzustellen.

1.Schritt. Man bestimme die am weitesten links stehende Spalte, die von Null verschiedene Elemente enthält.
2.Schritt. Man vertausche nötigenfalls die Zeilen, sodass in der obersten Zeile der gefundenen Spalte eine von Null verschiedene Zahl a steht. Die oberste Zeile wird dann durch a dividiert, sodass wir dort eine führende Eins erzeugt haben.
3. Schritt. Man addiere passende Vielfache der ersten Zeile zu den darunter stehenden Zeilen, sodass unter der führenden Eins Nullen erzeugt werden.
4. Schritt. Man wende die ersten drei Schritte auf die Untermatrix an, die durch Streichen der ersten Zeile entsteht und wiederhole das Verfahren bis die Matrix die Zeilenstufenform hat. Hier endet der Gauß-Algorithmus. Was folgt läuft unter dem Namen *Gauß-Jordan-Algorithmus.*
5. Schritt. Mit der letzten nichtverschwindenden Zeile beginnend addiere man geeignete Vielfache zu den darüberliegenden Zeilen, um über der führenden Eins Nullen zu erzeugen.

Die reduzierte Zeilenstufenform ist erreicht; die Lösung lässt sich nun einfach ablesen.

Beispiel 1.8 Gegeben sei das 3×5 lineare GS

$$\begin{array}{cccccl} & & -2x_3 & & +7x_5 & =12 \\ 2x_1 & +4x_2 & -10x_3 & +6x_4 & +12x_5 & =28 \\ 2x_1 & +4x_2 & -5x_3 & +6x_4 & -5x_5 & =-1 \end{array}$$

Wir demonstrieren den Gauß-Jordan-Algorithmus. Zunächst ordnen wir dem Gleichungssystem die erweiterte Koeffizientenmatrix zu.

0	0	-2	0	7	12	Vertauschen der ersten beiden Zeilen
2	4	-10	6	12	28	
2	4	-5	6	-5	-1	

2	4	-10	6	12	28	$\vert\cdot(-1)$
0	0	-2	0	7	12	
2	4	-5	6	-5	- 1	

2	4	-10	6	12	28	$\vert\cdot(-\frac{1}{2})$
0	0	-2	0	7	12	$\vert\cdot(-\frac{1}{2})$
0	0	5	0	-17	-29	

1	2	-5	3	6	14	
0	0	1	0	$-\frac{7}{2}$	-6	$\vert(-5)$
0	0	5	0	-17	-29	

1	2	-5	3	6	14	
0	0	1	0	$-\frac{7}{2}$	-6	
0	0	0	0	$\frac{1}{2}$	1	$\vert\cdot 2$

1	2	-5	3	6	14	
0	0	1	0	$-\frac{7}{2}$	-6	
0	0	0	0	1	2	$\vert\cdot\frac{7}{2}$

Damit ist die Zeilenstufenform erreicht. Um die reduzierte Zeilenstufenform zu erhalten fahren wir mit dem Gauß-Jordan-Algorithmus fort:

1	2	-5	3	6	14	
0	0	1	0	$-\frac{7}{2}$	-6	
0	0	0	0	1	2	$\vert\cdot\frac{7}{2}$

1	2	-5	3	6	14	
0	0	1	0	0	1	
0	0	0	0	1	2	$\vert\cdot(-6)$

1	2	-5	3	0	2	
0	0	1	0	0	1	$\vert\cdot 5$
0	0	0	0	1	2	

1	2	0	3	0	7	
0	0	1	0	0	1	
0	0	0	0	1	2	

Damit ist die reduzierte Zeilenstufenform erreicht. Die r Variablen, in deren Spalten führende Einsen stehen, sind *abhängige* Variable. Die $n-r$ Variablen, in deren Spalten keine führenden Einsen sind, sind *unabhängige*, das heißt, frei wählbare Variablen. Sie werden beim Ablesen der Lösung auf die rechte Seite gebracht.

Wir lesen die Lösung ab: $x_5 = 2$, $x_3 = 1$, $x_1 = -2x_2 - 3x_4 + 7$ mit frei wählbaren Parametern x_2 und x_4. Führt man für x_2 und x_4 die neuen Parameter s und t ein, so kann man die Lösung aufschreiben als

$$\begin{aligned} x_1 &= 7-2s-3t \\ x_2 &= s \\ x_3 &= 1 \\ x_4 &= t \\ x_5 &= 2. \end{aligned} \qquad x = \begin{pmatrix} 7\\0\\1\\0\\2 \end{pmatrix} + s \begin{pmatrix} -2\\1\\0\\0\\0 \end{pmatrix} + t \begin{pmatrix} -3\\0\\0\\1\\0 \end{pmatrix}$$

Die Lösungsmenge ist eine Ebene im $\mathbb{R}^5$ durch den Punkt $(7,0,1,0,2)^\top$ mit den Richtungs- oder Spannvektoren $v_1 = (-2,1,0,0,0)^\top$ und $v_2 = (-3,0,0,1,0)^\top$.

(d) Das Lösungsverhalten.
Fall 1. Besitzt die reduzierte Zeilenstufenform Zeilen, bei denen nur der Eintrag der letzten Spalte (rechte Seite) ungleich Null ist und alle anderen Einträge verschwinden, dann ist die Lösungsmenge leer, $\text{Lös}(A,b) = \emptyset$. man sagt, das Gleichungssystem ist *inkonsistent*.
Fall 2. Fall 1 trete nicht ein. Das Gleichungssystem ist *konsistent*.
Fall 2.1 Es gibt genau $r = n$ führende Einsen. In jeder Spalte der Koeffizientenmatrix gibt es eine führende Eins. Dann hat die reduzierte Zeilenstufenform die folgende Gestalt:

$$\begin{pmatrix} 1 & 0 & \cdots & 0 & | & c_1 \\ 0 & 1 & \cdots & 0 & | & c_2 \\ \vdots & & & & | & \vdots \\ 0 & 0 & \cdots & 1 & | & c_n \\ 0 & 0 & \cdots & 0 & | & 0 \\ \vdots & & & 0 & | & \vdots \\ 0 & 0 & \cdots & 0 & | & 0 \end{pmatrix}$$

und das Gleichungssystem hat die eindeutige Lösung $x = (c_1, \dots, c_n)$.
Fall 2.2 Es gibt $r < n$ führende Einsen. Dann hat die reduzierte Zeilenstufenform der Matrix $n - r$ freie Parameter; insbesondere gibt es unendlich viele Lösungen.
Wir werden später in der Lage sein, die Beschreibung zu präzisieren.

Bemerkung 1.3 (a) Für die Anzahl r der führenden Einsen nach Beendigung des Gauß-Jordan-Algorithmus gilt:

$$r \le n \quad \text{und} \quad r \le m,$$

denn jede führende Eins ist einer abhängigen Variablen zugeordnet (Spalte) und es kann nicht mehr abhängige Variablen als Variablen überhaupt geben. Außerdem steht jede führende Eins in einer gewissen Zeile und es gibt nur m Zeilen (Anzahl der Gleichungen).
(b) Insbesondere gilt für die Anzahl $n - r$ der freien Variablen, $n - r \ge n - m$. Im Falle $n > m$ heißt das, $n - r \ge n - m \ge 1$ — es gibt mindestens eine freie Variable. Ein *konsistentes* lineares Gleichungssystem mit mehr Variablen als Unbekannten besitzt stets unendlich viele Lösungen. Mit anderen Worten, für ein lineares (m,n)-GS mit $m > n$ kann nur der Fall 1 eintreten (keine Lösung, Inkonsistenz) oder der Fall 2.2 (unendlich viele Lösungen).
(c) Die Festlegung (bzw. Zuordnung) von freien und abhängigen Variablen, wie sie der Gauß-Jordan-Algorithmus liefert, ist nicht zwingend, das heißt, man kann die Rollen von abhängigen und unabhängigen Variablen auch vertauschen. So lässt sich $x + 2y = 2$ auflösen zu

$x = 2 - 2y$ oder zu $y = 1 - x/2$.
Wir werden später sehen, dass allerdings die *Anzahl* der freien und abhängigen Variablen fest liegt.

1.3.3 Homogene und inhomogene Gleichungssysteme

Ein lineares Gleichungssystem, bei dem der Vektor der rechten Seite gleich dem Nullvektor ist, heißt *homogenes* Gleichungssystem; also ist $Ax = 0$ ein homogenes GS. Es besitzt stets die *triviale* Lösung $x_1 = x_2 = \cdots = x_n = 0$. Alle anderen Lösungen des Systems, falls es noch welche gibt, heißen *nichttriviale Lösungen*. Ist der Vektor der rechten Seite b ungleich Null, so heißt $Ax = b$ *inhomogenes Gleichungssystem*. Wegen des Distributivgesetzes für Matrizen können wir nun schon folgenden wichtigen Satz beweisen

Satz 1.3 *Es sei $Ax = b$ ein lineares Gleichungssystem in Matrixschreibweise und $Ax = 0$ das zugehörige homogene Gleichungssystem.*
(a) *Sind u und v Lösungen des homogenen Systems, so ist auch $\lambda u + \mu v$ für beliebige Zahlen λ, μ eine Lösung des homogenen Systems.*
(b) *Ist u eine Lösung des homogenen Systems und w eine Lösung des inhomogenen Systems, so ist $u + w$ ebenfalls eine Lösung des inhomogenen Systems.*
(c) *Sind w und w' Lösungen des inhomogenen Systems, so ist $w - w'$ eine Lösung des homogenen Systems.*

Fasst man die Aussagen von (b) und (c) zusammen, so bedeutet dies, dass sich die allgemeine Lösung w des inhomogenen Systems schreiben lässt als Summe aus einer speziellen Lösung w_0 des inhomogenen Systems und der allgemeinen Lösung u, des homogenen Systems: $w = w_0 + u$.

$$\boxed{\text{Lös}(A, b) = w_0 + \text{Lös}(A, 0),}$$

dabei ist w_0 eine spezielle Lösung des inhomogenen Systems.
Beweis. (a) Aus $Au = 0$ und $Av = 0$ folgt wegen des Distributivgesetzes $A(\lambda u + \mu v) = \lambda Au + \mu Av = \lambda 0 + \mu 0 = 0$. Also erfüllt $\lambda u + \mu v$ das homogene System $Ax = 0$.
(b) Aus $Au = 0$ und $Aw = b$ folgt durch Addition $A(u + w) = Au + Aw = 0 + b = b$.
(c) Aus $Aw = b$ und $Aw' = b$ folgt durch Differenzbildung $A(w - w') = Aw - aw' = b - b = 0$. Somit erfüllt $w - w'$ das homogene System. ■

Beispiel 1.9 Wir betrachten das inhomogene lineare GS $2x + 3y = 5$. Es hat offenbar eine spezielle Lösung $w_0 = (1, 1)^\top$. Das zugehörige homogene System $2x + 3y = 0$ hat die allgemeine Lösung:

$$\text{Lös}(A, 0) = \left\{ \begin{pmatrix} -3 \\ 2 \end{pmatrix} t \mid \quad t \in \mathbb{R} \right\}$$

Also lautet die allgemeine Lösung des inhomogenen Systems

$$\text{Lös}(A, b) = \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \begin{pmatrix} -3 \\ 2 \end{pmatrix} t \mid \quad t \in \mathbb{R} \right\}.$$

Bemerkung 1.4 Jedes *homogene* lineare GS $Ax = 0$ mit mehr Variablen als Gleichungen, $n > m$, besitzt unendlich viele Lösungen. Dies folgt aus Bemerkung 1.3 (c) und der Tatsache, dass homogene GS stets konsistent sind (es existiert stets eine Lösung, nämlich die triviale).

Kapitel 2

Vektorräume

Vektorräume, nicht Vektoren, bilden den Hauptgegenstand der linearen Algebra. Vektoren heißen die Elemente des Vektorraumes. Um zu klären, was ein Vektor ist, benötigt man also vorher den Begriff des *Vektorraumes*.

2.1 Definition, Eigenschaften, Beispiele

Die im Folgenden rot markierten Abschnitte sind in der Vorlesung nicht dran gewesen und sind als ergänzende Bemerkung eingefügt.
Es sei im Folgenden $\mathbb{K}$ der Körper $\mathbb{R}$ oder $\mathbb{C}$ oder $\mathbb{Q}$.

Definition 2.1 Eine Menge V heißt *Vektorraum über* $\mathbb{K}$, wenn in V eine Addition $+$ und eine skalare Vervielfachung $\cdot$ definiert sind, wobei die folgenden Bedingungen erfüllt sind:

1. **Axiome der Addition**. $(V,+,0)$ ist eine abelsche Gruppe, das heißt,

   (A1) $\forall x,y \in V: x+y \in V$. (Abgeschlossenheit)

   (A2) $\forall x,y \in V: x+y=y+x$ (Kommutativgesetz).

   (A3) $\forall x,y,z \in V: x+(y+z)=(x+y)+z$. (Assoziativgesetz)

   (A4) Es gibt ein neutrales Element $0 \in V$, sodass $\forall x \in V: x+0=0+x=x$. (Nullelement)

   (A5) Zu jedem $x \in V$ gibt es ein Inverses $-x \in V$ mit $x+(-x)=(-x)+x=0$.

2. **Axiome der skalaren Vervielfachung**

   (S1) $\forall x \in V, \lambda \in \mathbb{K}: \lambda x \in V$ (Abgeschlossenheit)

   (S2) $\forall x \in V, \lambda, \mu \in \mathbb{K}: \lambda(\mu x)=(\lambda\mu)x$.

   (S3) $\forall x \in V: 1x=x$

3. **Distributivgesetze**

(D1) $\forall x,y \in V, \lambda \in \mathbb{K}: \lambda(x+y)=\lambda x+\lambda y$.

(D2) $\forall x \in V, \lambda, \mu \in \mathbb{K}: (\lambda + \mu)x = \lambda x + \mu x$.

Andere Sprechweisen: Ist $\mathbb{K} = \mathbb{R}$, so heißt V *reeller Vektorraum* oder *reeller linearer Raum*; im Falle $\mathbb{K} = \mathbb{C}$ heißt V *komplexer Vektorraum* oder *komplexer linearer Raum*.

2.1.1 Einfachste Eigenschaften

Wir schreiben $v - w$ für $v + (-w)$. Wie bei den Schlussfolgerungen aus den Körperaxiomen (A1) bis (A5)(vgl. Analysis 1), gilt auch hier

1. $\forall x, y, z: x + y = x + z \Longrightarrow y = z$. (Kürzungsregel)
2. $\forall x, y \in V: x + y = x \Longrightarrow y = 0$. (Eindeutigkeit des Nullvektors)
3. $\forall x, y \in V: x + y = 0 \Longrightarrow y = -x$ (Eindeutigkeit des Inversen)
4. $\forall x \in V: -(-x) = x$.
5. $0 \cdot x = 0$.
6. $\alpha \cdot 0 = 0$.
7. $(-1) \cdot x = -x$.

2.1.2 Beispiele für Vektorräume

Beispiel 2.1 (a)$\mathbb{K}$ selbst ist ein Vektorraum über $\mathbb{K}$.
(b) $\mathbb{K}^{m\times n}$. Die Menge aller $m \times n$-Matrizen mit Einträgen aus $\mathbb{K}$ bilden einen $\mathbb{K}$-Vektorraum. Insbesondere ist der $\mathbb{R}^n = \mathbb{R}^{1\times n} = \{(x_1, \dots, x_n) \mid x_i \in \mathbb{R}\ \forall\, i = 1, \dots, n\}$ der Raum der Vektoren mit n reellen Koordinaten ein reeller Vektorraum. Ebenso sind der $\mathbb{C}^n$ und der $\mathbb{Q}^n$ definiert als Raum der Spalten- oder Zeilenvektoren.
Vereinbarung: Wir bezeichnen sowohl den Raum der Zeilenvektoren $\mathbb{R}^{1\times n}$ als auch den Raum der Spaltenvektoren $\mathbb{R}^{n\times 1}$ mit $\mathbb{R}^n$. Aus dem Kontext heraus sollte jeweils klar sein, welchen Vektorraum $\mathbb{R}^n$ wir meinen.
Wir identifizieren auch die klassischen „geometrischen“ Vektoren der Ebene und des Raumes mit den Vektoren aus $\mathbb{R}^2$ bzw. aus $\mathbb{R}^3$. Ein geometrischer Vektor der Ebene ist eine Pfeilklasse, die durch Länge, Richtung und Richtungssinn charakterisiert ist. Unter allen Pfeilen- einer Klasse gibt es einen, der im Ursprung O (eines gedachten Koordinatensystems)angreift, $\overrightarrow{OP}$. Wenn P die Koordinaten (x_1, x_2) hat, so identifizieren wir $\overrightarrow{OP}$ und $(x_1, x_2) \in \mathbb{R}^2$. Man überzeugt sich leicht, dass die Addition und Vervielfachung geometrischer Vektoren mit den Vektorraumoperationen im $\mathbb{R}^2$ und $\mathbb{R}^3$ verträglich sind. Das bedeutet, wenn $\overrightarrow{OP} + \overrightarrow{OQ} = \overrightarrow{OR}$ mit $P = (x_1, x_2)$, $Q = (y_1, y_2)$, so gilt $R = (x_1 + y_1, x_2 + y_2)$.
(c) $\mathbb{C}^n$ ist auch ein Vektorraum über $\mathbb{R}$.
(d) $\mathbf{Abb}(X, V)$. Es sei X eine beliebige Menge und V ein Vektorraum über $\mathbb{K}$. Dann ist $\mathrm{Abb}(X, V)$ die Menge aller Abbildungen von X nach V. In dieser Menge kann man wie folgt addieren und skalar vervielfachen. Für $f, g \in \mathrm{Abb}(X, V)$ und $\lambda \in \mathbb{K}$ sei :

$$(f + g)(x) := f(x) + g(x), \quad (\lambda f)(x) := \lambda f(x).$$

Auf diese Weise wird $\mathrm{Abb}(X, V)$ zu einem $\mathbb{K}$-Vektorraum.

Wir beweisen, dass etwa Axiom (D2) erfüllt ist. Seien also $\lambda, \mu \in K$ und $f \in \mathrm{Abb}(X, V)$. Wir müssen zeigen, dass die Abbildungen $(\lambda+\mu)f = \lambda f + \mu f$ übereinstimmen. Dazu müssen Definitionsbereiche und Wertebereiche der Abbildungen übereinstimmen — in unserem Falle sind dies rechts und links Abbildungen $X \to V$. Außerdem müssen die *Abbildungsvorschriften* gleich sein. In der Tat gilt für alle $x \in X$:

$$((\lambda+\mu)f)(x) = (\lambda+\mu)f(x) \underset{\text{wegen (D2) in } V}{=} \lambda f(x) + \mu f(x) = (\lambda f)(x) + (\mu f)(x) = (\lambda f + \mu f)(x).$$

Das neutrale Element in $\mathrm{Abb}(X, V)$ ist die (konstante) Nullfunktion $=: X \to V$, $0(x) = 0$; das zu f inverse Element ist $-f \colon X \to V$, wobei $(-f)(x) = -f(x)$.
Identifiziert man einen Vektor $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ mit der Funktion $x \colon \{1, \dots, n\} \to \mathbb{R}$, $x(k) = x_k$, $k = 1, \dots, n$, so sieht man, dass $\mathrm{Abb}(\{1, \dots, n\}, \mathbb{R}) = \mathbb{R}^n$.
(e) $\boldsymbol{\omega}$. Die Räume der reellen Zahlenfolgen $\omega := \mathrm{Abb}(\mathbb{N}, \mathbb{R}) = \{(x_n) \mid x_n \in \mathbb{R}\}$, der konvergenten reellen Folgen c, der Nullfolgen c_0, der finiten Folgen φ, bilden reelle Vektorräume. Es gilt $\varphi \subset c_0 \subset c \subset \omega$.
(f) $\mathbb{R}[\boldsymbol{x}]$ sei die Menge der Polynome $p(x) = a_n x^n + a_{n-1}x^{n-1} + \dots + a_1 x + a_0$ mit reellen Koeffizienten $a_0, \dots, a_n \in \mathbb{R}$. Fasst man $\mathbb{R}[x] \subset \mathrm{Abb}(\mathbb{R}, \mathbb{R})$ auf, so sind in natürlicher Weise Addition und skalare Multiplikation erklärt.
(g) $\mathbf{C}^{\boldsymbol{m}}(\mathbb{R})$. Es ist $\mathrm{C}^m(\mathbb{R})$ die Menge der m-fach stetig differenzierbaren Funktionen auf $\mathbb{R}$. Im Falle $m = 0$ schreiben wir $\mathrm{C}(\mathbb{R})$ für die Menge der stetigen Funktionen auf $\mathbb{R}$. Als Teilmenge $\mathrm{C}^m(\mathbb{R}) \subset \mathrm{Abb}(\mathbb{R}, \mathbb{R})$ sind Addition und skalare Vervielfachung wohldefiniert. So gilt etwa für die Funktion $f \colon \mathbb{R} \to \mathbb{R}$, die gegeben ist durch

$$f(x) = |x|^3 = \begin{cases} x^3, & x \ge 0, \\ -x^3, & x < 0 \end{cases} \Longrightarrow f'(x) = \begin{cases} 3x^2, & x \ge 0, \\ -3x^2, & x < 0 \end{cases} \Longrightarrow f''(x) = 6|x|.$$

Da die Betragsfunktion bei 0 nicht differenzierbar ist, gilt $f \in \mathrm{C}^2(\mathbb{R})$ aber $f \notin \mathrm{C}^3(\mathbb{R})$.
(h) **Keinen** reellen Vektorraum bilden

endliche Intervalle $[a, b]$,
die Menge der Vektoren $\{(\lambda, 1+\lambda) \in \mathbb{R}^2 \mid \lambda \in \mathbb{R}\}$ — der Nullvektor $(0,0)$ ist nicht enthalten,
Die Menge der Polynome mit *ganzzahligen* Koeffizienten.

2.2 Lineare Unterräume

Definition 2.2 Eine Teilmenge W eines Vektorraumes V heißt *linearer Unterraum* oder *linearer Teilraum* von V, wenn W zusammen mit der Addition von V und der skalaren Multiplikation aus V wieder ein Vektorraum über $\mathbb{K}$ ist.

Lemma 2.1 (Unterraumkriterium) *Eine Teilmenge W eines Vektorraumes V ist genau dann linearer Teilraum von V, wenn gilt*

(N) $W \neq \emptyset$.
(A) *Für alle* $w_1, w_2 \in W$ *ist* $w_1 + w_2 \in W$. *(Abgeschlossenheit der Addition)*
(S) *Für* $\lambda \in \mathbb{K}$ *und* $w \in W$ *gilt* $\lambda w \in W$. *(Abgeschlossenheit der skalaren Vervielfachung)*

Äquivalent zu (A) und (S) ist die Bedingung

(L) $\forall\, w_1, w_2 \in W, \lambda_1, \lambda_2 \in \mathbb{K}\colon \lambda_1 w_1 + \lambda_2 w_2 \in W.$

Beweis. Es ist lediglich zu zeigen, dass sämtliche Axiome aus der Vektorraumdefinition 2.1 sich auf W übertragen. Assoziativität und Kommutativität von Addition und skalarer Multiplikation sowie die Distributivgesetze sind unmittelbar klar, da sie für V und somit erst recht für eine Teilmenge W gelten. Eine Ausnahme machen lediglich die Existenzforderungen (Nullvektor und Inverses).
Nun ist aber W nichtleer, enthält also zumindest einen Vektor $a \in W$. Nach Eigenschaft 5 (s.o.) gilt aber $0 \cdot a = 0$ und wegen der Abgeschlossenheit der skalaren Multiplikation in W also $0 \in W$. 0 ist dann auch neutral in W, weil neutral in V. Ferner gilt nach Eigenschaft 7 auch $(-1)a = -a$. Somit liegt mit $a \in W$ auch das Inverse $-a \in W$. ∎

Bemerkung 2.1 (a) Die leere Menge ist kein Unterraum; jeder Unterraum enthält zumindest den Nullvektor. Triviale Unterräume von V sind $\{0\}$ und V.
(b) Sind W_1 und W_2 lineare Teilräume von V, so ist auch ihr Durchschnitt $W_1 \cap W_2$ ein linearer Teilraum (Beweis: Unterraumkriterium). Diese Eigenschaft lässt sich auf beliebige Durchschnitte von endlich oder unendlich vielen Unterräumen verallgemeinern.
(c) Die Vereinigung von Teilräumen ist i. a. kein Teilraum.
(d) Ist $U \subset W$ ein Unterraum und $W \subset V$ ein Unterraum, so ist auch U ein Unterraum von V.

Beispiel 2.2 (a) Für $A \in \mathbb{R}^{m \times n}$ ist die Menge der Lösungen eines homogenen Gleichungssystems

$$U := \text{Lös}(A, 0) = \{x \in \mathbb{R}^n \mid Ax = 0\}$$

ein linearer Unterraum von $\mathbb{R}^n$. Der Beweis folgt aus Satz 1.3, wo gezeigt wird, dass (L) gilt, nämlich wenn $Au = 0$ und $Av = 0$, dann ist auch $A(\lambda u + \mu v) = 0$. Ferner ist $U \neq \emptyset$, da der Nullvektor in U liegt.
(b) Es sei $\mathbb{R}_d[x]$ die Menge der reellen Polynome, die *höchstens* den Grad d haben; $p(x) = a_d x^d + \dots + a_1 x + a_0$, $a_i \in \mathbb{R}$ für alle $i = 0, \dots, d$. $\mathbb{R}_d[x]$ ist ein linearer Teilraum von $\text{Abb}(\mathbb{R}, \mathbb{R})$, denn das Nullpolynom $0(x) = 0$ liegt in $\mathbb{R}_d[x]$ und wenn $q(x) = b_d x^d + \dots + b_1 x + b_0$ ein weiteres Polynom in $\mathbb{R}_d[x]$ ist, so ist auch $(\lambda p + \mu q)(x) = \sum_{i=0}^{d} (\lambda a_i + \mu b_i) x^i$ ein Polynom in $\mathbb{R}_d[x]$.
Wir haben die folgenden Inklusionen von Unterräumen in $\text{Abb}(\mathbb{R}, \mathbb{R})$:

$$\begin{aligned} &\mathbb{R}_0[x] \subset \dots \subset \mathbb{R}_d[x] \subset \mathbb{R}[x] \subset \\ &\qquad \subset C^{\omega}(\mathbb{R}) \subset C^{\infty}(\mathbb{R}) \subset \cdots C^m(\mathbb{R}) \subset \dots \subset C^2(\mathbb{R}) \subset C^1(\mathbb{R}) \subset C(\mathbb{R}) \subset \text{Abb}(\mathbb{R}, \mathbb{R}). \end{aligned} \tag{2.1}$$

Die Polynome genau $n$ten Grades bilden keinen Teilraum, da $p(x) = x^n$ und $q(x) = -x^n + 1$ zwar beide zu W gehören, jedoch nicht ihre Summe $p(x) + q(x) = 1$.
(c) Es sei $V \subset \mathrm{Abb}(\mathbb{R}, \mathbb{R})$ ein Teilraum, $\alpha \in \mathbb{R}$ fixiert. Dann ist

$$U := \{f \in V \mid f(\alpha) = 0\}$$

ein linearer Teilraum von V. Hingegen ist $W = \{f \in V \mid f(\alpha) = 1\}$ kein linearer Teilraum.
(d) $V = \mathbb{R}^n$, i mit $1 \leq i \leq n$ sei eine fixierte natürlich Zahl. W sei die Menge aller reellen n-Tupel der Form $(x_1, \dots, x_{i-1}, 0, x_{i+1}, \dots, x_n)$. Dann ist W ein linearer Teilraum von $\mathbb{R}^n$. Dies ist ein Spezialfall von (a) mit der einzigen homogenen Gleichung $x_i = 0$; hierbei ist $A = (0 \cdots 1 \cdots 0)$ mit einer einzigen 1 an der Stelle k und $n-1$ Nullen.
(e) $V = \mathrm{C}^2(\mathbb{R})$, $a_1, a_2 \in \mathbb{R}$ fix. Es sei

$$U := \{u \in \mathrm{C}^2(\mathbb{R}) \mid u'' + a_1 u' + a_2 u = 0\}.$$

die Menge der Lösungen der gewöhnlichen Differentialgleichung zweiter Ordnung mit konstanten Koeffizienten a_1, a_2. U ist ein Unterraum in $\mathrm{C}^2(\mathbb{R})$. In der Tat ist das Unterraumkriterium erfüllt, denn aus $u_1, u_2 \in U$ folgt, dass auch $\lambda_1 u_1 + \lambda_2 u_2$ die Differentialgleichung erfüllt. Im Falle der Gleichung $u'' + u = 0$ erhalten wir $u(x) = \lambda_1 \sin x + \lambda_2 \cos x$ als allgemeine Lösung. Das heißt, $\{\sin x, \cos x\}$ bildet ein Erzeugendensystem für U.

2.2.1 Die lineare Hülle und Erzeugendensysteme

Definition 2.3 Es sei V ein Vektorraum und $v_1, \dots, v_n \in V$.
(a) Jede Summe $\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n$ mit $\lambda_k \in \mathbb{K}$, heißt *Linearkombination* der Vektoren $v_1, \dots, v_n$.
(b) Es sei $M \subset V$ eine beliebige Teilmenge des Vektorraumes V. Als *lineare Hülle* von M in V bezeichnen wir die Menge aller möglichen Linearkombinationen von Vektoren aus M:

$$\operatorname{lin} M := \{\lambda_1 v_1 + \cdots + \lambda_n v_n \mid n \in \mathbb{N}_0, \lambda_1, \dots, \lambda_n \in \mathbb{K}, v_1, \dots, v_n \in M\}$$

(c) Gilt $\operatorname{lin} M = V$, so heißt M ein *Erzeugendensystem* von V.

Man beachte, dass eine Linearkombination immer eine *endliche Summe* von Vektoren ist. Ist $n = 0$, so ist nach Vereinbarung die leere Linearkombination gleich dem Nullvektor und entsprechend gilt: $\operatorname{lin} \emptyset = \{0\}$.

Beispiel 2.3 (a) Im $\mathbb{K}^n$ bilden die n *Einheitsvektoren* $\mathrm{e}_k = (0, 0, \dots, 1, 0, \dots, 0)$ mit einer Eins an der $k$ten Stelle, $k = 1, \dots, n$ ein Erzeugendensystem, denn jeder Vektor $x \in \mathbb{R}^n$ lässt sich schreiben als Linearkombination der Einheitsvektoren, $x = x_1 \mathrm{e}_1 + \cdots + x_n \mathrm{e}_n$.
(b) Im Raum der $m \times n$-Matrizen $\mathbb{K}^{m \times n}$ bilden die *Matrixeinsen* $\{E_{rs} \mid r = 1, \dots, m,\ s = 1, \dots, n\}$ ein Erzeugendensystem. Sie sind definiert als

$$E_{rs} = \begin{pmatrix} & & 0 & & & \\ & & \vdots & & & \\ 0 & \cdots & 1 & 0 & \cdots & 0 \\ & & \vdots & & & \\ & & 0 & & & \end{pmatrix} = (\delta_{ir}\delta_{js})_{\substack{i=1\dots m \\ j=1\dots n}},$$

wobei die einzige Eins in der $r$ten Zeile und $s$ten Spalte steht. In der Tat ist diese Menge erzeugend für $\mathbb{K}^{m\times n}$, da für jede Matrix $A=(a_{ij})$ gilt: $A=\sum_{i=1}^m\sum_{j=1}^n a_{ij}E_{ij}$.
(c) Die $d+1$ Polynome $\{1,x,x^2,\dots,x^d\}$ bilden ein Erzeugendensystem von $\mathbb{R}_d[x]$ während die Polynome $\{1,x,\dots\}$ ein Erzeugendensystem in $\mathbb{R}[x]$ bilden.

Bemerkung 2.2 (a) Die lineare Hülle $\operatorname{lin} M$ ist ein linearer Teilraum von V. *Beweis.* (a) Wir zeigen, dass das Unterraumkriterium für $\operatorname{lin} M$ erfüllt ist. Im Fall der leeren Menge ist $\operatorname{lin}\varnothing=\{0\}$ ein Unterraum. Ist $M\neq\varnothing$, so ist auch $\operatorname{lin} M\supseteq M$ nichtleer. Seien nun v und w jeweils endliche Linearkombinationen von Elementen aus M. Dann ist auch $v+w$ eine Linearkombination von Elementen aus M und auch λv. Das Unterraumkriterium ist erfüllt. ∎

(b) Es gilt $\operatorname{lin} M=\bigcap_{W\supset M} W$, wobei der Durchschnitt über alle linearen Teilräume W von V genommen wird, die M enthalten. Mit anderen Worten, $\operatorname{lin} M$ ist der kleinste lineare Teilraum von V, der M enthält.

Beispiel 2.4 (Anwendung auf lineare Gleichungssysteme) Sehr häufig sind wir mit der folgenden Fragestellung konfrontiert: Gegeben seien Vektoren $v_1,\dots,v_r\in\mathbb{R}^m$

(a) Liegt ein gegebener Vektor $b\in\mathbb{R}^m$ in der linearen Hülle der $\{v_i\}$, gilt also $b\in\operatorname{lin}\{v_1,\dots,v_r\}$?
(b) Sind die Vektoren sogar ein Erzeugendensystem von $\mathbb{R}^m$?

Schreibt man die Vektoren v_i als Spaltenvektoren

$$v_i=\begin{pmatrix}a_{1i}\\a_{2i}\\\vdots\\a_{mi}\end{pmatrix},\quad b=\begin{pmatrix}b_1\\b_2\\\vdots\\b_m\end{pmatrix},\quad i=1,\dots,r.$$

dann gilt $b\in\operatorname{lin}\{v_1,\dots,v_r\}$ genau dann, wenn es reelle Zahlen $x_1,\dots,x_r\in\mathbb{R}$ gibt mit $x_1v_1+\dots+x_rv_r=b$. Mit Hilfe der Spaltenvektoren kann man das so schreiben:

$$x_1\begin{pmatrix}a_{11}\\a_{21}\\\vdots\\a_{m1}\end{pmatrix}+x_2\begin{pmatrix}a_{12}\\a_{22}\\\vdots\\a_{m2}\end{pmatrix}+\dots+x_r\begin{pmatrix}a_{1r}\\a_{2r}\\\vdots\\a_{mr}\end{pmatrix}=\begin{pmatrix}b_1\\b_2\\\vdots\\b_m\end{pmatrix},$$

was wiederum äquivalent zum linearen Gleichungssystem

$$\begin{aligned}a_{11}x_1+a_{12}x_2+\dots+a_{1r}x_r&=b_1\\a_{21}x_1+a_{22}x_2+\dots+a_{2r}x_r&=b_2\\&\dots\\a_{m1}x_1+a_{m2}x_2+\dots+a_{mr}x_r&=b_m.\end{aligned}\tag{2.2}$$

ist. Somit gilt

$$\boxed{b \in \operatorname{lin}\{v_1, \ldots, v_r\} \Leftrightarrow \text{Lös}(A, b) \neq \emptyset.}$$

Das heißt, genau dann liegt b in der linearen Hülle, wenn das zugeordnete lineare Gleichungssystem konsistent ist. Jede Lösung $x = (x_1, \ldots, x_r)$ liefert eine Darstellung von b als Linearkombination der Vektoren $v_1, \ldots, v_r$. Es gilt ferner

$$\boxed{\operatorname{lin}\{v_1, \ldots, v_r\} = \mathbb{R}^m \Leftrightarrow \ \forall\, b \in \mathbb{R}^m : \text{Lös}(A, b) \neq \emptyset.}$$

(2) Nun kann man die Konsistenz eines linearen Gleichungssystems mit Hilfe der Spaltenvektoren der Matrix A ausdrücken:

Das lineare GS (2.2) ist genau dann konsistent, wenn die rechte Seite b in der linearen Hülle der Spaltenvektoren der Koeffizientenmatrix A liegt.

2.3 Die Basis eines Vektorraumes

2.3.1 Lineare Unabhängigkeit

Es sei V ein Vektorraum über $\mathbb{K}$.

Definition 2.4 (a) Endlich viele Elemente $v_1, \ldots, v_n$ aus V heißen *linear abhängig* (über $\mathbb{K}$), wenn es $\alpha_1, \ldots, \alpha_n \in \mathbb{K}$ gibt, die nicht alle gleich Null sind und für die

$$\alpha_1 v_1 + \cdots + \alpha_n v_n = 0$$

gilt, das heißt, wenn die Vektoren $v_1, \ldots, v_n$ den Nullvektor nicht-trivial darstellen.
(b) Endlich viele Elemente $v_1, \ldots, v_n$ aus V heißen *linear unabhängig* über $\mathbb{K}$, wenn sie nicht linear abhängig sind, das heißt, wenn für alle $\alpha_1, \ldots, \alpha_n \in \mathbb{K}$ gilt:

$$\alpha_1 v_1 + \cdots + \alpha_n v_n = 0 \quad \Longrightarrow \quad \alpha_1 = \alpha_2 = \cdots = \alpha_n = 0. \tag{2.3}$$

Mit anderen Worten, eine Linearkombination der v_i, die den Nullvektor darstellt, muss die triviale Linearkombination sein.
(c) Eine nichtleere Teilmenge $M \subset V$ heißt *linear unabhängig* über $\mathbb{K}$, wenn jede endliche Teilmenge von M linear unabhängig ist. Die leere Menge gilt als linear unabhängig.

Bemerkung 2.3 (a) In der obigen Definition ist bei (a) und (b) auch der Fall $n = 1$ eingeschlossen: $\{v\}$ ist linear unabhängig gdw. $v \neq 0$. Ist einer der Vektoren $v_1, \ldots, v_n$ gleich dem Nullvektor, so ist die Menge linear abhängig, denn sei etwa $v_1 = 0$, dann ist $v_1 + 0v_2 + \cdots + 0v_n = 0$ eine nichttriviale Linearkombination, die die Null darstellt.
(b) Ist $v_1, \ldots, v_k$ linear abhängig mit $k < n$, so ist auch $v_1, \ldots, v_n$ linear abhängig. Umgekehrt ist jede Teilmenge einer linear unabhängigen Menge wieder linear unabhängig.

Beispiel 2.5 (a) Die Vektoren $(2,4,6),(3,6,9) \in \mathbb{R}^3$ sind linear abhängig, denn $3(2,4,6)+(-2)(3,6,9)=(0,0,0)$.
(b) Die Vektoren $v_1=(1,2,3)$ und $v_2=(3,2,1)$ sind im $\mathbb{R}^3$ linear unabhängig, denn das homogene lineare Gleichungssystem

$$x_1\begin{pmatrix}1\\2\\3\end{pmatrix}+x_2\begin{pmatrix}3\\2\\1\end{pmatrix}=\begin{pmatrix}x_1+3x_3\\2x_1+2x_2\\3x_1+x_2\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix}$$

geht durch den Gauß-Algorithmus über in äquivalente homogene Systeme mit Koeffizientenmatrizen

$$\begin{pmatrix}1&3&|0\\2&2&|0\\3&1&|0\end{pmatrix}\sim\begin{pmatrix}1&3&|0\\0&-4&|0\\0&-8&|0\end{pmatrix}\sim\begin{pmatrix}1&3&|0\\0&1&|0\\0&0&|0\end{pmatrix}\sim\begin{pmatrix}1&0&|0\\0&1&|0\\0&0&|0\end{pmatrix}$$

über. Hieraus folgt, dass die einzige Lösung $x_1=x_2=0$ ist; die Vektoren sind also linear unabhängig.
(c) Die Vektoren $v_1=(1,\mathrm{i})^\top$ und $v_2=(\mathrm{i},-1)^\top$ sind $\mathbb{R}$-linear unabhängig als Elemente des reellen Vektorraumes $\mathbb{C}^2$, denn es gibt keine nichttriviale reelle Linearkombination $\alpha_1 v_1+\alpha_2 v_2=0$, $\alpha_1,\alpha_2\in\mathbb{R}$.
Sie sind aber linear abhängig als Elemente des komplexen Vektorraumes $\mathbb{C}^2$, denn $v_2=\mathrm{i}v_1$.
(d) Die Polynome $\{1,x,x^2,\dots\}$ bilden eine linear unabhängige Menge in $\mathbb{R}[x]$ (Beweis: später).

Aufgabe 2. Es sei $p\in V=\mathbb{R}_d[x]$ und $a\in\mathbb{R}$.
(a) Zeigen Sie, dass $B:=\{1,x-a,(x-a)^2,\dots,(x-a)^d\}$ eine Basis in V bilden.
(b) Bestimmen Sie die Koordinaten von p bezüglich B. *Lösung.* Die Taylorentwicklung von p an der Stelle a gibt die Koordinaten:

$$p(x)=p(a)+\frac{p'(a)}{1!}(x-a)+\cdots+\frac{p^{(d)}(a)}{d!}(x-a)^d.$$

Aufgabe 3. Zeigen Sie dass $\{1,\sqrt{2},\sqrt{3}\}$ als Elemente des $\mathbb{Q}$-Vektorraumes $\mathbb{R}$ linear unabhängig sind.
Aufgabe 4. Zeigen Sie, dass $v_1(x)=\sin x$ und $v_2(x)=\cos x$ und $v_0(x)=1$ als Elemente des reellen Vektorraumes $\mathrm{Abb}(\mathbb{R},\mathbb{R})$ linear unabhängig sind.
Aufgabe 5. Zeigen Sie, dass $\{v_1,\dots,v_n\}\subset V$ genau dann linear unabhängig ist, wenn die Abbildung $\mathbb{K}^n\to V$ gegeben durch $f((\alpha_1,\dots,\alpha_n))=\sum_{i=1}^n\alpha_i v_i$ injektiv ist.

2.3.2 Der Begriff einer Basis

Definition 2.5 Eine Teilmenge B eines Vektorraumes V heißt *Basis* von V, wenn gilt

1. B ist ein Erzeugendensystem von V.
2. B ist eine linear unabhängige Menge.

Beispiel 2.6 (a) Im $\mathbb{K}^n$ bilden die n Einheitsvektoren $\mathrm{e}_k = (0,0,\ldots,1,0,\ldots,0)$ mit einer Eins an der $k$ten Stelle, $k = 1,\ldots,n$ eine Basis. In der Tat sind sie linear unabhängig, da aus $\alpha_1 \mathrm{e}_1 + \cdots + \alpha_n \mathrm{e}_n = 0$, folgt $(\alpha_1,\alpha_2,\ldots,\alpha_n) = 0$. Somit ist α der Nullvektor, also $\alpha_1 = \alpha_2 = \cdots = 0$. Diese Basis heißt *kanonische Basis* oder *Standardbasis* des $\mathbb{K}^n$.
(b) Im Raum der (m,n)-Matrizen $\mathbb{K}^{m\times n}$ bilden die Matrixeinsen E_{rs}, $r = 1,\ldots,m$, $s = 1,\ldots,n$ eine Basis, denn die Menge ist erzeugend und linear unabhängig. Die lineare Unabhängigkeit zeigt man genauso wie in (a).
(c) Die Polynome $\{1,x,\ldots,x^d\}$ bilden eine Basis in $\mathbb{R}_d[x]$.
(d) Es seien

$$\begin{aligned} v_1 &= (x_{11},x_{12},\ldots,x_{1n}),\\ v_2 &= (0,x_{22},\ldots,x_{2n}),\\ \ldots & \quad \ldots\\ v_n &= (0,0,\ldots,0,x_{nn}) \end{aligned}$$

Vektoren des $\mathbb{R}^n$. Dann ist $\{v_1,\ldots,v_n\}$ genau dann Basis von $\mathbb{R}^n$, wenn $x_{11}x_{22}\cdots x_{nn} \neq 0$.
Beweis. Wir zeigen, dass das System $\{v_1,\ldots,v_n\}$ ein Erzeugendensystem im $\mathbb{R}^n$ ist, wenn die Bedingung $x_{11}x_{22}\cdots x_{nn} \neq 0$ erfüllt ist. Dazu müssen wir nach Beispiel 2.4 zeigen, dass das lineare Gleichungssystem mit der Koeffizientenmatrix, gebildet aus den Spaltenvektoren $v_n, v_{n-1},\ldots,v_1$ für beliebige rechte Seiten lösbar ist. Das lineare GS liegt aber bereits in Dreiecksform vor. Zum Erreichen der Zeilenstufenform muss nur noch jede Zeile i durch x_{ii} geteilt werden, was nach Voraussetzung möglich ist. Somit ist das $n\times n$ lineare GS äquivalent zu einer Zeilenstufenform mit genau n führenden Einsen. Das System ist eindeutig lösbar; die Vektoren erzeugen den $\mathbb{R}^n$.
Lineare Unabhängigkeit. Hierzu hat man das homogene lineare Gleichungssystem zu lösen. Die obigen Argumentationen mit den n führenden Einsen zeigen, dass das homogene System ebenfalls eindeutig lösbar ist; es gibt nur die triviale Lösung und somit sind die Vektoren linear unabhängig.
Ist umgekehrt das Produkt der x_{ii} gleich Null, so ist mindestens ein Faktor Null, etwa $x_{kk} = 0$, $1 \le k \le n$. Dann gibt es höchstens $n-1$ führende Einsen. Somit hat das homogene System unendlich viele Lösungen; die Vektoren sind also nicht linear unabhängig. ■

2.3.3 Die Unabhängigkeit der Dimension von der Basis

In diesem Abschnitt wollen wir zeigen, dass zwei Vektorraumbasen in V stets dieselbe Anzahl von Elementen haben.

Lemma 2.2 (Charakterisierung der Basis) *Es sei V ein Vektorraum, $V \neq \{0\}$, $n \in \mathbb{N}$ und $B = \{b_1,\ldots,b_n\}$ eine Teilmenge von V. Dann sind äquivalent:*
(a) *B ist eine Basis von V.*
(b) *Zu jedem $v \in V$ existieren eindeutig bestimmte Zahlen $\alpha_1,\ldots,\alpha_n \in \mathbb{K}$ mit $v = \alpha_1 b_1 + \cdots + \alpha_n b_n$.*

(c) *B ist ein minimales Erzeugendensystem von V, d. h., für eine echte Teilmenge $C \subsetneq B$ ist* $\operatorname{lin} C \subsetneq V$.
(d) *B ist eine maximale linear unabhängige Menge, d. h. jede Menge $D \supsetneq B$ ist linear abhängig.*

Beweis. Der Beweis folgt dem Schema $(a) \Longrightarrow (b) \Longrightarrow (c) \Longrightarrow (d) \Longrightarrow (a)$.
$(a) \Longrightarrow (b)$. Die Existenz der α_i ist klar, da B ein Erzeugendensystem für V ist. Wir zeigen die Eindeutigkeit. Seien $\alpha_i' \in \mathbb{K}$ andere Zahlen mit $v = \sum_{i=1}^n \alpha_i' b_i$, so hat man durch Gleichsetzen

$$v = \sum_{i=1}^n \alpha_i' b_i = \sum_{i=1}^n \alpha_i b_i \Longrightarrow \sum_{i=1}^n (\alpha_i - \alpha_i') b_i = 0,$$

das heißt, dass der Nullvektor als Linearkombination der b_i dargestellt ist. Wegen der linearen Unabhängigkeit von B folgt $\alpha_i - \alpha_i' = 0$ für alle i — die Darstellung ist eindeutig. Man nennt sie die *Koordinatendarstellung* von v bezüglich der Basis B.
$(b) \Longrightarrow (c)$. Wegen (b) gilt $\operatorname{lin} B = V$. Angenommen, es gilt auch $\operatorname{lin} C = V$. Wegen $B \supsetneq C$ existiert $v \in B \setminus C$, also $v \in B$, $v \notin C$; Wegen $\operatorname{lin} C = V$ existieren eindeutig α_i, $i = 1, \dots, m$ mit $v = \sum_i \alpha_i c_i$ mit Vektoren $c_1, \dots, c_m \in C$. Andererseits gilt aber in B, $v = 1 \cdot v$. Somit ist die Eindeutigkeit der Darstellung von v als Linearkombination von Elementen aus B verletzt. Die Annahme ist falsch; es gilt $\operatorname{lin} C \neq V$.
$(c) \Longrightarrow (d)$. Wir müssen zeigen: B minimal erzeugend impliziert B maximal unabhängig. Das stimmt sicher, wenn B nur aus einem Element besteht, etwa $B = \{b\}$. Wegen $V \neq \{0\}$ ist $b \neq 0$ und damit ist B linear unabhängig. Andernfalls hat B mindestens 2 Elemente. Wäre B linear abhängig, so könnte ein Element von B als Linearkombination der anderen dargestellt werden, etwa $b = \sum_i \alpha_i b_i$. Dann ist aber auch $B \setminus \{b\}$ nach wie vor erzeugend; im Widerspruch zur Annahme in (c). Damit ist B linear unabhängig.
Wir zeigen die Maximalität. Da B erzeugend ist, lässt sich jeder weitere Vektor $d \in D \setminus B$ als Linearkombination von Elementen aus B darstellen. Damit ist aber $B \cup \{d\}$ linear abhängig.
$(d) \Longrightarrow (a)$. Sei B maximal linear unabhängig und $v \in V$. Wir müssen zeigen, dass B erzeugend ist, also $v \in \operatorname{lin} B$. Da B maximal linear unabhängig ist, ist $B \cup \{v\}$ linear abhängig, das heißt, es gibt eine nichttriviale Linearkombination, die den Nullvektor darstellt, etwa

$$\alpha_1 b_1 + \cdots + \alpha_n b_n + \alpha v = 0. \tag{2.4}$$

Wäre hier $\alpha = 0$, so $\alpha v = 0$ und damit $\alpha_1 b_1 + \cdots + \alpha_n b_n = 0$, was aber $\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$ impliziert, da B linear unabhängig ist. Damit wäre (2.4) eine triviale Linearkombination, im Widerspruch zur Annahme. Folglich gilt $\alpha \neq 0$ und somit $v = -(\alpha_1 b_1 + \cdots + \alpha_n b_n)/\alpha \in \operatorname{lin} B$. Da $v \in V$ beliebig war, heißt das $\operatorname{lin} B = V$; B ist ein Erzeugendensystem und damit Basis.

■

Lemma 2.3 *Es sei $B = \{b_1, \dots, b_n\}$ eine Basis von V und es sei $c = \sum_{i=1}^n \alpha_i b_i$ die Koordinatendarstellung eines Vektors $c \in V$, wobei $\alpha_1 \neq 0$ gelte.*
Dann ist $B' = \{c, b_2, \dots, b_n\}$ ebenfalls eine Basis von V.

Beweis. Da $\alpha_1 \neq 0$, kann man nach b_1 umstellen und hat $b_1 = (c - \alpha_2 b_2 - \alpha_3 b_3 - \cdots - \alpha_n b_n)/\alpha_1$. Also gilt $\operatorname{lin}\{c, b_2, \ldots, b_n\} = \operatorname{lin}\{b_1, b_2, \ldots, b_n\} = V$. Die Menge B' ist erzeugend.
Weiter ist B' aber auch linear unabhängig, denn

$$\beta c + \beta_2 b_2 + \cdots + \beta_n b_n = 0 \Longrightarrow \beta \alpha_1 b_1 + (\beta \alpha_2 + \beta_2) b_2 + \cdots + (\beta \alpha_n + \beta_n) b_n = 0$$

impliziert $\beta \alpha_1 = 0$, also $\beta = 0$ und weiter $\beta_2 = \cdots = \beta_n = 0$. Somit ist B' linear unabhängig. ■

Der nächste Satz stellt eine Verallgemeinerung des obigen Lemmas dar.

Satz 2.4 (Austauschsatz von Steinitz) *Es sei $\{b_1, \ldots, b_n\}$ eine Basis des Vektorraumes V und $\{c_1, \ldots, c_k\} \subset V$ sei linear unabhängig.*
Dann gilt $k \leq n$ und bei geeigneter Nummerierung der Vektoren $b_1, \ldots, b_n$ ist auch $\{c_1, \ldots, c_k, b_{k+1}, \ldots, b_n\}$ eine Basis von V. Man erhält also wieder eine Basis, wenn man $c_1, \ldots, c_k$ durch $n - k$ geeignete Vektoren unter den $\{b_1, \ldots, b_n\}$ ergänzt.

Beweis. Der Beweis erfolgt durch vollständige Induktion über k. Als Induktionsanfang kann man den Fall $k = 0$ zulassen, in dem überhaupt keine Vektoren ausgetauscht werden und der Satz trivial ist.
Im allgemeinen Fall von k linear unabhängigen Vektoren $c_1, \ldots, c_k$ sind auch die $k - 1$ Vektoren $c_1, \ldots, c_{k-1}$ linear unabhängig und nach Induktionsvoraussetzung folgt $k - 1 \leq n$, und, bei geeigneter Nummerierung ist $B^* := \{c_1, \ldots, c_{k-1}, b_k, b_{k+1}, \ldots, b_n\}$ eine Basis von V. Wäre $k - 1 = n$, so wäre bereits $\{c_1, \ldots, c_{k-1}\}$ eine Basis von V (alle Vektoren b_i werden durch die c_i ausgetauscht). Dann ist aber c_k eine Linearkombination von $c_1, \ldots, c_{k-1}$. Das widerspricht der linearen Unabhängigkeit der Vektoren $c_1, \ldots, c_k$. Daher gilt sogar $k - 1 < n$, also $k \leq n$.
Da B^* eine Basis von V ist gibt es Zahlen α_i, $i = 1, \ldots, n$ mit

$$c_k = \alpha_1 c_1 + \cdots + \alpha_{k-1} c_{k-1} + \alpha_k b_k + \cdots + \alpha_n b_n.$$

Würden hierbei $\alpha_k = \alpha_{k+1} = \cdots = \alpha_n = 0$ gelten, so wäre $c_k \in \operatorname{lin}\{c_1, \ldots, c_{k-1}\}$ im Widerspruch zur linearen Unabhängigkeit der Vektoren $\{c_1, \ldots, c_k\}$. Folglich ist mindestens einer der Koeffizienten $\alpha_k, \alpha_{k+1}, \ldots, \alpha_n$ von Null verschieden. Bei geeigneter Nummerierung der b_i können wir $\alpha_k \neq 0$ annehmen. Nach Lemma 2.3 ist dann $B' := \{c_1, \ldots, c_k, b_{k+1}, \ldots, b_n\}$ eine Basis von V. ■

Folgerung 2.5 *Sind $\{b_1, \ldots, b_n\}$ und $\{c_1, \ldots, c_k\}$ Basen des Vektorraumes V, so gilt $k = n$.*

Beweis. Nach dem Steinitzschen Austauschsatz gilt einerseits $k \leq n$, andererseits, wegen der Gleichberechtigung der Basen $\{c_i\}$ und $\{b_i\}$ auch $n \leq k$. Wenn also ein Vektorraum überhaupt eine endliche Basis besitzt, dann bestehen alle seine Basen aus gleichvielen Vektoren; besitzt aber ein Vektorraum eine unendliche Basis, so sind auch alle anderen Basen unendlich. ■

Definition 2.6 Wenn ein Vektorraum V eine endliche Basis besitzt, wird die allen Basen gemeinsame Anzahl von Basisvektoren die *Dimension* von V genannt, symbolisch $\dim V$. Besitzt V keine endliche Basis, so heißt V *unendlich-dimensional* ($\dim V = \infty$). Die Dimension des Nullraumes wird gleich 0 gesetzt.

Wir schreiben $\dim_{\mathbb{K}} V$ für die Dimension des $\mathbb{K}$-Vektorraumes V, wenn wir den Skalarenkörper $\mathbb{K}$ hervorheben wollen. Mitunter ist V ein Vektorraum über verschiedenen Körpern.

Beispiel 2.7 (a) Es gilt $\dim \mathbb{K}^n = n$ und $\dim \mathbb{K}^{m\times n} = mn$. Dies folgt aus Beispiel 2.3 und Beispiel2.6.
(b) Es ist $\dim \mathbb{R}[x] = \infty$, denn $\{1, x, x^2, x^3, \dots\}$ ist eine unendliche linear unabhängige Menge in $\mathbb{R}[x]$. Es gilt $\dim \mathbb{R}_d[x] = d+1$.
(c) $\dim_{\mathbb{R}} \mathbb{C}^n = 2n$, denn die Vektoren $\{\mathrm{e}_1, \dots, \mathrm{e}_n, \mathrm{i}\mathrm{e}_1, \dots, \mathrm{i}\mathrm{e}_n\}$ bilden eine Basis des reellen Vektorraumes $\mathbb{C}^n$. **Beweis in der Übung**

Lemma 2.6 *Es sei* $\dim V = n < \infty$.
(a) *Eine linear unabhängige Menge C von V kann zu einer Basis von V ergänzt werden.*
(b) *Eine linear unabhängige Teilmenge C von V ist genau dann Basis von V, wenn $|C| = n$.*
(c) *Für jeden Unterraum $U \subset V$ gilt* $\dim U \le \dim V$. *Aus $U \subset V$ und* $\dim U = \dim V$ *folgt* $U = V$.

Beweis. (a) Es sei B eine Basis von V. Dann gilt $|C| \le |B| = n$. Nach dem Steinitzschen Austauschsatz kann die linear unabhängige Menge C durch geeignete Vektoren der Basis B zu einer Basis von V ergänzt werden.
(b) Ist C noch keine Basis, so kann nach (a) C durch Vektoren zu einer Basis ergänzt werden, die dann n Elemente hat; also ist $|C| < n$. Wenn C bereits aus n Vektoren besteht, so ist werden keine Vektoren zum Ergänzen benötigt; somit ist C bereits eine Basis.
(c) Es sei $\{b_1, \dots, b_n\}$ eine Basis von V und $\{c_1, \dots, c_k\}$ eine Basis von U. Nach dem Steinitzschen Austauschsatz gilt dann $k \le n$. Ist aber $k = n$, so folgt nach Teil (b), dass $\{c_1, \dots, c_k\}$ bereits eine Basis von V ist, also $U = V$. ∎

2.3.4 Direkte Summe und Summe von Unterräumen

Definition 2.7 Es sei V_i, $i = 1, \dots, n$ eine endliche Familie von Vektorräumen über $\mathbb{K}$. Dann wird die Menge

$$W := V_1 \times V_2 \times \cdots \times V_n = \{(v_1, \dots, v_n) \mid v_i \in V_i \ \forall\, i = 1, \dots, n\}$$

zu einem Vektorraum über $\mathbb{K}$, wenn man die Summe und die skalare Multiplikation koordinatenweise erklärt:

$$(v_1, \dots, v_n) + (w_1, \dots, w_n) := (v_1 + w_1, \dots, v_n + w_n), \quad \lambda(v_1, \dots, v_n) := (\lambda v_1, \lambda v_2, \dots, \lambda v_n),$$

wobei für alle $i = 1, \dots, n$ gilt $v_i, w_i \in V_i$.

In diesem Falle heißt W die *(äußere) direkte Summe* der Vektorräume V_i und wir bezeichnen sie mit $V_1 \oplus V_2 \oplus \cdots \oplus V_n$.

Bemerkung 2.4 Für eine Familie $\{V_i \mid i \in I\}$ von $\mathbb{K}$-Vektorräumen sei $V = \bigcup V_i$ deren Vereinigung. Man definiert

(a) das *direkte Produkt* der Vektorräume V_i als

$$P = \bigotimes_{i\in I} V_i := \{f : I \to V \mid \ \forall\, i \in I : f(i) \in V_i\}$$

(b) die *direkte Summe* der Vektorräume V_i als

$$S = \bigoplus_{i\in I} V_i = \{f \in \bigotimes_{i\in I} V_i \mid f(i) = 0 \text{ für fast alle } i\}.$$

In diesem Sinne ist $\bigotimes_{n\in\mathbb{N}} \mathbb{R} = \omega$ der Raum aller reellen Folgen und $\bigoplus_{n\in\mathbb{N}} \mathbb{R} = \varphi$ der Raum der finiten Folgen. Außerdem ist $\bigotimes_{r\in\mathbb{R}} \mathbb{R} = \mathrm{Abb}(\mathbb{R}, \mathbb{R})$ der Raum der auf ganz $\mathbb{R}$ definierten reellwertigen Funktionen.

Definition 2.8 Es sei V ein Vektorraum.
(a) Es seien U_1 und U_2 lineare Teilräume von V. Dann heißt

$$U_1 + U_2 = \{u_1 + u_2 \mid u_1 \in U_1, u_2 \in U_2\}$$

die *Summe* der Teilräume U_1 und U_2.
(b) Es sei $\{U_i \mid i \in I\}$ eine Familie von Teilräumen von V. Dann heißt

$$\operatorname{lin} \bigcup_{i\in I} U_i$$

Summe der Teilräume $\{U_i \mid i \in I\}$. Wir bezeichnen sie mit $\sum_{i\in I} U_i$.
(c) Die Summe von Unterräumen $W := \sum_{i\in I} U_i$ heißt *direkt* bzw. *innere direkte Summe*, wenn sich jedes $w \in W$, $w \neq 0$, auf genau eine Art als endliche Linearkombination von Elementen aus U_i schreiben lässt. Genauer, für jedes $w \in W$ gibt es eindeutig bestimmte, endlich viele Indizes $i_1, \ldots, i_r \in I$ und eindeutig bestimmte Elemente $u_k \in U_{i_k}$, $k = 1, \ldots, r$, mit $w = u_1 + \cdots + u_r$. In diesem Falle schreiben wir auch $W = \bigoplus_{i\in I} U_i$.

Bemerkung 2.5 (a) $\mathbb{K}^n = \mathbb{K} \oplus \mathbb{K} \oplus \cdots \oplus \mathbb{K}$.
(b) Eine Summe $W = U_1 + U_2$ von zwei Teilräumen ist direkt gdw. $U_1 \cap U_2 = \{0\}$.
Achtung: Eine Summe $U_1 + U_2 + U_3$ ist genau dann direkt, wenn

$$U_1 \cap U_2 = \{0\} \quad \text{und} \quad (U_1 + U_2) \cap U_3 = \{0\}.$$

Die Bedingung $U_1 \cap U_2 \cap U_3 = \{0\}$ ist nicht hinreichend für die Direktheit der Summe, wie folgendes Beispiel (c) zeigt.
(c) Wir bezeichnen mit $\mathbb{R}v = \{\lambda v \mid \lambda \in \mathbb{R}\}$ die lineare Hülle eines einzelnen Vektors v. Die Summe von Teilräumen $\mathbb{R}v_1 + \cdots + \mathbb{R}v_n$ ist direkt genau dann, wenn $\{v_1, \ldots, v_n\}$ linear unabhängig ist. Dies folgt direkt aus Lemma 2.2 (b)

(d) Ist $V = V_1 \oplus V_2 \oplus \cdots \oplus V_n$ die äußere direkte Summe von Vektorräumen V_i, so kann man die Teilmengen

$$\tilde{V}_k := \{(0, \ldots, v_i, 0, \ldots, 0) \in V \mid v_i \in V_i\}$$

definieren. Dann ist $\tilde{V}_i$ ein Unterraum von V und V ist die innere direkte Summe der $\tilde{V}_i$, $i = 1, \ldots, n$.

2.3.5 Dimensionsformel für Summen von Untervektorräumen

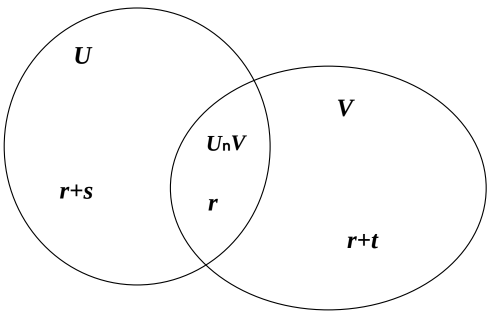
Nach dem vorigen Abschnitt ist klar, dass für eine Familie $\{V_1, \ldots, v_n\}$ von endlichdimensionalen Unterräumen von V gilt

$$\dim(V_1 + V_2 + \cdots + V_n) \leq \dim V_1 + \dim V_2 + \cdots + \dim V_n,$$

denn sind B_k Basen von V_k, $k = 1, \ldots, n$, so ist $B_1 \cup B_2 \cup \cdots \cup B_n$ sicherlich ein Erzeugendensystem für die Unterraumsumme. Genau dann, wenn die Summe direkt ist, gilt Gleichheit. Wir untersuchen den Fall $n = 2$.

Satz 2.7 (Dimensionsformel) *Es seien U und V zwei endlichdimensionale Unterräume eines Vektorraumes W. Dann gilt*

$$\dim U + \dim V = \dim(U + V) + \dim(U \cap V). \tag{2.5}$$



Beweis. Es sei $B_D = \{d_1, \ldots, d_r\}$ eine Basis des Durchschnitts $U \cap V$, wobei auch $r = 0$ zugelassen ist, also $B_D = \emptyset$ und $U \cap V = \{0\}$. Da B_D linear unabhängig sowohl in U als auch in V ist, kann diese Menge nach Satz 2.4 sowohl zu einer Basis $B_1 = \{d_1, \ldots, d_r, a_1, \ldots, a_s\}$ von U als auch zu einer Basis $B_2 = \{d_1, \ldots, d_r, b_1, \ldots, b_t\}$ von V ergänzt werden.

Wir wollen nun zeigen, dass $B = \{d_1, \ldots, d_r, a_1, \ldots, a_s, b_1, \ldots, b_t\}$ eine Basis von $U + V$ ist. Jeder Vektor $w \in U + V$ lässt sich als $w = u + v$, $u \in U$, $v \in V$ schreiben. Da sich u als Linearkombination von B_1 und v als Linearkombination von B_2 schreiben lässt, ist $w = u + v$ eine Linearkombination von B. Also gilt $U + V = \operatorname{lin} B$. Zum Nachweis der linearen Unabhängigkeit von B nehmen wir an, dass

$$\alpha_1 d_1 + \cdots + \alpha_r d_r + \beta_1 a_1 + \cdots + \beta_s a_s + \gamma_1 b_1 + \cdots + \gamma_t b_t = 0, \tag{2.6}$$

also

$$\alpha_1 d_1 + \cdots + \alpha_r d_r + \beta_1 a_1 + \cdots + \beta_s a_s = -(\gamma_1 b_1 + \cdots + \gamma_t b_t) =: x.$$

Da die linke Seite der obigen Gleichung in U liegt, die rechte aber in V, liegt der Vektor x in $U \cap V$ und ist demnach eine Linearkombination der Vektoren aus B_D. Also gibt es Zahlen $\delta_1, \ldots, \delta_r$ mit

$$x = \delta_1 d_1 + \cdots + \delta_r d_r.$$

Wegen der linearen Unabhängigkeit von B_1 folgt hieraus aber $\alpha_1 - \delta_1 = \cdots = \alpha_r - \delta_r = 0$ und $\beta_1 = \cdots = \beta_s = 0$. Setzt man $\beta_i = 0$ in (2.6) ein, so hat man

$$\alpha_1 d_1 + \cdots + \alpha_r d_r + \gamma_1 b_1 + \cdots + \gamma_t b_t = 0.$$

Da B_2 eine linear unabhängige Menge ist, folgt $\alpha_1 = \cdots = \gamma_t = 0$. Damit ist B eine Basis von $U + V$. Somit gilt

$$\dim U + \dim V = (r+s) + (r+t) = r + (r+s+t) = \dim(U \cap V) + \dim(U+V).$$

■

Bemerkung 2.6 (a) Jeder Vektorraum besitzt eine Basis. Dies beweist man mit Hilfe des *Auswahlaxioms* der Mengenlehre:

> Es sei **P** eine disjunkte Zerlegung der nichtleeren Menge X in nichtleere Teilmengen $P \subset X$. Dann existiert eine Auswahlfunktion $f \colon \mathbf{P} \to X$ mit $f(P) \in P$ für alle $P \in \mathbf{P}$.

(b) Auch im unendlichdimensionalen Fall gilt der Austauschsatz. Zu je zwei Basen B und C eines Vektorraumes V gibt es bijektive Abbildung $\varphi \colon B \to C$. Lemma 2.6 hingegen bleibt nicht gültig für unendlich-dimensionale Räume.

2.3.6 Konstruktion von Basen zu Erzeugendensystemen

Eine typische Fragestellung ist: Finden Sie zu einem gegebenen Erzeugendensystem E eines Unterraumes U eine Basis von U. Nach Lemma 2.2 (c) kann man durch schrittweises Weglassen einzelner Vektoren aus E eine Basis erzeugen. Nach jedem Schritt müsste man kontrollieren, ob noch ein Erzeugendensystem übrig geblieben ist. Dieses Verfahren ist nicht praktikabel. Vielmehr wollen wir hier ein auf dem Gauß-Algorithmus beruhendes Verfahren vorstellen, das die gegebenen Vektoren linear kombiniert zu einer möglichst einfachen Basis.

Es seien $a_1, \dots, a_m \in \mathbb{R}^n$ gegebene Vektoren, die den Unterraum $U = \operatorname{lin}\{a_1, \dots, a_m\} \subset \mathbb{R}^n$ aufspannen. Wir betrachten die Matrix $A = (a_{ij}) \in \mathbb{R}^{m \times n}$, deren m Zeilen die Vektoren $a_1, \dots, a_m$ sind, also $a_i = (a_{i1}, a_{i2}, \dots, a_{in})$, $i = 1, \dots, m$. Die drei elementaren Zeilenoperationen des Gauß-Algorithmus, verändern den Raum U, der durch die Zeilen der Matrix aufgespannt wird nicht. Ist die Zeilenstufenform erreicht, so kann man durch Weglassen der Nullzeilen sofort eine Basis von U angeben.

Beispiel 2.8 Man ermittle eine Basis des von den Vektoren $a_1 = (0,0,0,2,-1)$, $a_2 = (0,1,-2,1,0)$, $a_3 = (0,-1,2,1,-1)$ und $a_4 = (0,0,0,1,2)$ aufgespannten Teilraumes von $\mathbb{R}^5$.

0	0	0	2	-1	Vertauschen der ersten beiden Zeilen
0	1	-2	1	0	
0	-1	2	1	-1	
0	0	0	1	2	
0	1	-2	1	0	$\mid\cdot(+1)$ zur 3.Z.
0	0	0	2	-1	
0	-1	2	1	-1	
0	0	0	1	2	
0	1	-2	1	0	$\mid\cdot(-1)$ zur 3.Z.
0	0	0	2	-1	Streichen
0	0	0	2	-1	$\mid:2$
0	0	0	1	2	Vertauschen mit 2.Z.

0	1	-2	1	0	
0	0	0	1	2	
0	0	0	1	-1/2	
0	0	0	0	0	
0	1	-2	0	0	
0	0	0	1	0	
0	0	0	0	1	
0	0	0	0	0	

Somit ist $b_1 = (0,1,-2,0,0)$, $b_2 = (0,0,0,1,0)$, $b_3 = (0,0,0,0,1)$ eine Basis von W.

Kapitel 3

Lineare Abbildungen

3.1 Definition, Beispiele, Eigenschaften

Einleitung: [Koe97, §6, S. 35]: Seit Erwachen der Analysis im 17. Jahrhundert haben die Mathematiker immer wieder nach Funktionen gesucht, die sich durch besonders schöne Eigenschaften auszeichnen. Eine interessante Klasse solcher Eigenschaften fasst man unter dem Stichwort „Funktionalgleichungen“ zusammen. Für Funktionen $f\colon \mathbb{R} \to \mathbb{R}$ sind die folgenden Funktionalgleichungen denkbar:

$$\begin{aligned} f(x+1) &= f(x), \\ f(xy) &= f(x)f(y), \\ f(x+y) &= f(x)f(y), \\ f(x+y) &= f(x)+f(y). \end{aligned} \tag{3.1}$$

Die erste Funktionalgleichung wird von allen *periodischen* Funktionen mit der Periode 1 erfüllt; die zweite Gleichung von Potenzfunktionen $f(x) = x^{\alpha}$; in der dritten Klasse liegen die Exponentialfunktionen $f(x) = a^x$ und schließlich wird die vierte Klasse von den linearen Funktionen $f(x) = cx$ erfüllt. Doch, sind dies alle Lösungen? Bereits Augustin Louis Cauchy (1789–1857) zeigte, dass alle *stetigen* Lösungen der Gleichung (3.1) die Form $f(x) = cx$ mit einer Konstanten $c \in \mathbb{R}$ haben. Wir werden sehen, dass dies nicht alle Lösungen sind.

Definition 3.1 Es seien V und W Vektorräume über $\mathbb{K}$. Eine $\mathbb{K}$-*lineare Abbildung* ist eine Abbildung $T\colon V \to W$, so dass für alle $v_1, v_2 \in V$ und $\lambda \in \mathbb{K}$ gilt

$$T(v_1+v_2) = T(v_1)+T(v_2) \quad \text{(Additivität)} \quad \text{und} \quad T(\lambda v_1) = \lambda T(v_1) \quad \text{(Homogenität)}. \tag{3.2}$$

Die beiden Bedingungen lassen sich zu einer Bedingung zusammenfassen, die für alle $v_1, v_2 \in V$ und alle $\lambda_1, \lambda_2 \in \mathbb{K}$ gelten muss: $T(\lambda_1 v_1 + \lambda_2 v_2) = \lambda_1 T(v_1) + \lambda_2 T(v_2)$. Wir bezeichnen die Menge der $\mathbb{K}$-linearen Abbildungen von V nach W mit $\mathrm{L}_{\mathbb{K}}(V,W)$ oder $\mathrm{L}(V,W)$ oder $\mathrm{Hom}_{\mathbb{K}}(V,W)$.

Wenn der Grundkörper $\mathbb{K}$ eindeutig fixiert ist, sprechen wir einfach von *linearen Abbildungen*. Andere Bezeichnungen für eine lineare Abbildung sind *linearer Operator* oder *lineare Transformation* oder *Homomorphismus* von Vektorräumen.

Eine lineare Abbildung $T \in \mathrm{L}(V, W)$ heißt

> *Monomorphismus*, wenn sie injektiv ist, *Epimorphismus*, wenn sie surjektiv ist, *Isomorphismus*, falls sie bijektiv ist, *Endomorphismus*, wenn $V = W$ und *Automorphismus*, wenn sie bijektiv ist und $V = W$.

Die Menge der Endomorphismen bezeichnen wir mit $\mathrm{L}(V)$, die Menge der Automorphismen $T\colon V \to V$ mit $\mathrm{GL}(V)$.

Bemerkung 3.1 (a) $\mathrm{L}(V, W) \subset \mathrm{Abb}(V, W)$ ist ein Untervektorraum von $\mathrm{Abb}(V, W)$. In der Tat ist $\mathrm{L}(V, W)$ nichtleer, da die Nullabbildung linear ist $(0(\lambda_1 v_1 + \lambda_2 v_2) = 0 = \lambda_1 0(v_1) + \lambda_2 0(v_2))$ und stets zu $\mathrm{L}(V, W)$ gehört. Seien nun $S, T \in \mathrm{L}(V, W)$; wir müssen zeigen, dass dann auch $S + T$ und λS linear sind. In der Tat gilt für alle $v_1, v_2 \in V$, $\lambda_1, \lambda_2 \in \mathbb{K}$

$$\begin{aligned}(S+T)(\lambda_1 v_1 + \lambda_2 v_2) &= S(\lambda_1 v_1 + \lambda_2 v_2) + T(\lambda_1 v_1 + \lambda_2 v_2)\\ &= \lambda_1 S(v_1) + \lambda_2 S(v_2) + (\lambda_1 T(v_1) + \lambda_2 T(v_2)) = \lambda_1 (S+T)(v_1) + \lambda_2 (S+T)(v_2).\end{aligned}$$

Somit ist $S + T$ linear. Außerdem gilt wegen der Linearität von S

$$(\lambda S)(\lambda_1 v_1 + \lambda_2 v_2) = \lambda(S(\lambda_1 v_1 + \lambda_2 v_2)) = \lambda\lambda_1 S(v_1) + \lambda\lambda_2 S(v_2) = \lambda_1(\lambda S)(v_1) + \lambda_2(\lambda S)(v_2).$$

Also ist λS linear. Damit sind Linearkombinationen von linearen Abbildungen linear; $\mathrm{L}(V, W)$ ist ein linearer Teilraum.
(b) Sind $U \xrightarrow{f} V \xrightarrow{g} W$ lineare Abbildungen, so auch $g \circ f\colon U \to W$. In der Tat ist

$$g(f(\lambda v_1 + v_2)) = g(\lambda f(v_1) + f(v_2)) = \lambda g(f(v_1)) + g(f(v_2)) = \lambda g \circ f(v_1) + g \circ f(v_2).$$

(c) Die Identität $\mathrm{id}_V\colon V \to V$ ist stets linear. In der Tat ist $\mathrm{id}_V(\lambda_1 v_1 + \lambda_2 v_2) = \lambda_1 v_1 + \lambda_2 v_2 = \lambda_1 \mathrm{id}_V(v_1) + \lambda_2 \mathrm{id}_V(v_2)$.

Beispiel 3.1 (a1) **Die lineare Abbildung einer Matrix.** Es sei $A \in \mathbb{K}^{m\times n}$, dann definieren wir $T_A \in \mathrm{L}(\mathbb{K}^n, \mathbb{K}^m)$ für $x = (x_1, \dots, x_n)^\top$ über $T_A(x) := A \cdot x$. Das Ergebnis ist ein Spaltenvektor in $\mathbb{K}^m$. In der Tat ist T_A linear, denn wegen der Distributivität der Matrixmultiplikation ist $T_A(\lambda x + \mu y) = A(\lambda x + \mu y) = \lambda A x + \mu A y = \lambda T_A(x) + \mu T_A(y)$ für alle $x, y \in \mathbb{K}^n$, $\lambda, \mu \in \mathbb{K}$.
So ist für die Matrix $A = \begin{pmatrix} 2 & 3 & 1\\ -7 & 4 & 0\end{pmatrix} \in \mathbb{R}^{2\times 3}$ die zugehörige lineare Abbildung $T_A \in \mathrm{L}(\mathbb{R}^3, \mathbb{R}^2)$ gegeben durch

$$T_A(x) = Ax = \begin{pmatrix} 2 & 3 & 1\\ -7 & 4 & 0\end{pmatrix}\begin{pmatrix} x_1\\ x_2\\ x_3\end{pmatrix} = \begin{pmatrix} 2x_1 + 3x_2 + x_3\\ -7x_1 + 4x_2\end{pmatrix}.$$

Die Bilder der Einheitspaltenvektoren e_j, $j = 1, \dots, n$ unter T_A sind genau die Spalten der Matrix A:

$$(A\mathrm{e}_j)_i = \sum_{k=1}^{n} a_{ik}\delta_{kj} = a_{ij}, \qquad A\mathrm{e}_j = A \cdot \begin{pmatrix} 0\\ \vdots\\ 1\\ \vdots\\ 0\end{pmatrix} = \begin{pmatrix} a_{1j}\\ a_{2j}\\ \vdots\\ a_{nj}\end{pmatrix} = \sum_{i=1}^{m} a_{ij}\mathrm{e}_i.$$

(a2) **Die Komposition.** Es seien $A=(a_{ij})\in\mathbb{K}^{m\times n}$ und $B=(b_{jk})\in\mathbb{K}^{n\times p}$ verkettete Matrizen. Dann ist $AB\in\mathbb{K}^{m\times p}$ und es gilt

$$T_A\circ T_B = T_{AB}$$

als lineare Abbildungen von $\mathbb{K}^p\to\mathbb{K}^m$.
Beweis. Es sei $x=(x_1,\dots,x_p)^\top\in\mathbb{K}^p$. Dann gilt wegen der Assoziativität der Matrixmultiplikation:

$$(T_A\circ T_B)(x)=T_A(T_B(x))=A\cdot(B\cdot x)=(A\cdot B)\cdot x=T_{AB}(x).$$

∎

(b) **Differentiation.** Es sei $V=\mathrm{C}^1(\mathbb{R})$ der Vektorraum der stetig differenzierbaren Funktionen und $W=\mathrm{C}(\mathbb{R})$ die stetigen Funktionen auf $\mathbb{R}$. Dann ist $D\colon V\to W$, $f\mapsto f'$, linear, denn $(f+g)'=f'+g'$ und $(\lambda f)'=\lambda f'$.
(c) **Integration.** Es sei $V=\mathrm{C}(\mathbb{R})$, $W=\mathbb{R}$ und $a<b$, $a,b\in\mathbb{R}$. Dann ist die Abbildung $V\to W$, $f\mapsto\int_a^b f(x)\,\mathrm{d}x$, linear, denn $\int_a^b(f(x)+g(x))\,\mathrm{d}x=\int_a^b f(x)\,\mathrm{d}x+\int_a^b g(x)\,\mathrm{d}x$ und $\int_a^b\lambda f(x)\,\mathrm{d}x=\lambda\int_a^b f(x)\,\mathrm{d}x$.
(d) **Evaluation.** Es sei V ein reeller Vektorraum, X eine beliebige Menge, $\alpha\in X$ und $W=\mathrm{Abb}(X,V)$. Dann ist die Abbildung $\mathrm{ev}_\alpha\colon W\to V$, $\mathrm{ev}_\alpha(f)=f(\alpha)$, linear.
(e) **Grenzwerte.** Es sei c der Vektorraum der konvergenten reellen Zahlenfolgen. Dann ist $\lim\colon c\to\mathbb{R}$, $(x_n)\mapsto\lim_{n\to\infty}x_n$, eine lineare Abbildung, denn für konvergente Folgen (x_n) und (y_n) gilt der Grenzwertsatz $\lim_{n\to\infty}(\lambda_1x_n+\lambda_2y_n)=\lambda_1\lim_{n\to\infty}x_n+\lambda_2\lim_{n\to\infty}y_n$.

Aufgabe 6. Für welche $m,n\in\mathbb{R}$ ist die Abbildung $f\colon\mathbb{R}\to\mathbb{R}$, $f(x)=mx+n$, $x\in\mathbb{R}$, linear?

Lemma 3.1 *Es sei $T\colon V\to W$ linear. Dann gilt:*
(a) $T(0)=0$, $T(-v)=-T(v)$, $T(v-w)=T(v)-T(w)$ *für alle* $v,w\in V$.
(b) $T(\lambda_1v_1+\cdots+\lambda_nv_n)=\lambda_1T(v_1)+\cdots+\lambda_nT(v_n)$ *für alle* $v_i\in V$, $\lambda_i\in\mathbb{K}$.
(c) *Ist für eine Familie* $\{v_i\mid i\in I\}$ *die Menge* $\{T(v_i)\mid i\in I\}$ *linear unabhängig in* W, *so ist* $\{v_i\mid i\in I\}$ *linear unabhängig in* V.
(d) *Für lineare Teilräume* $V'\subset V$ *und* $W'\subset W$ *sind Bild* $T(V')\subset W$ *und Urbild* $T^{-1}(W')\subset V$ *lineare Teilräume.*
(e) *Es gilt* $\dim T(V)\le\dim V$.
(f) *Ist T ein Isomorphismus, so ist auch T^{-1} linear (und ein Isomorphismus).*

Beweis. (a) Wegen $(-1)v=-v$ und der Linearität von T gilt, $T(v-w)=T(v+(-1)w)=T(v)+(-1)T(w)=T(v)-T(w)$. Setzt man hier $v=w=0$ ein, so hat man $T(0)=T(0-0)=T(0)-T(0)=0$; setzt man $v=0$ ein, so hat man $T(-w)=T(0-w)=T(0)-T(w)=0-T(w)=-T(w)$.
(b) folgt durch vollständige Induktion über n.
(c) Sei $\sum_{k=1}^n\alpha_kv_k=0$. Nach $T(0)=0$ und (b) folgt durch Anwenden von T,

$$T(0)=0=T\left(\sum_{k=1}^n\alpha_kv_k\right)=\sum_{k=1}^n\alpha_kT(v_k).$$

Wegen der linearen Unabhängigkeit von $\{T(v_k) \mid k = 1,\dots,n\}$ folgt $\alpha_1 = \cdots = \alpha_n = 0$; also waren die $\{v_i\}$ bereits linear unabhängig.
(d) Wegen $0 \in V'$, $0 \in W'$ und $T(0) = 0$ sind die Mengen $T(V')$ und $T^{-1}(W')$ nichtleer. Wir zeigen, dass das Unterraumkriterium Lemma 2.1 für $T(v')$ erfüllt ist. Seien dazu $w_1, w_2 \in T(V')$, etwa $w_1 = T(v_1)$ und $w_2 = T(v_2)$ mit $v_1, v_2 \in V'$. Dann ist

$$\lambda_1 w_1 + \lambda_2 w_2 = \lambda_1 T(v_1) + \lambda_2 T(v_2) = T(\lambda_1 v_1 + \lambda_2 v_2) = T(v_3),$$

dabei ist $v_3 = \lambda_1 v_1 + \lambda_2 v_2 \in V'$ wegen der Abgeschlossenheit der linearen Operationen in V'. Also $\lambda_1 w_1 + \lambda_2 w_2 \in T(V')$. Der Beweis des Unterraumkriteriums für $T^{-1}(W')$ verläuft analog.
(e) Sind $T(v_1),\dots,T(v_n)$ linear unabhängig in W, so sind nach (c) $v_1,\dots,v_n$ linear unabhängig in V; also $n \le \dim V$. Dies gilt insbesondere für eine Basis $\{T(v_1),\dots,T(v_n)\}$ von $T(V)$.
(f) Es seien $w, w' \in W$. Da T bijektiv ist, existieren eindeutig bestimmte Vektoren $v, v' \in V$ mit $w = T(v)$, $w' = T(v')$, so dass $v = T^{-1}(w)$, $v' = T^{-1}(w')$. Dann gilt $T(\lambda v + \mu v') = \lambda w + \mu w'$. Wendet man darauf T^{-1} an, so hat man $\lambda T^{-1}(w) + \mu T^{-1}(w') = T^{-1}(\lambda w + \mu w')$; also ist T^{-1} linear. ∎

Lemma 3.2 (Existenz- und Eindeutigkeitssatz) *Es seien V und W Vektorräume und $B \subset V$ eine Basis von V. Ferner sei eine Abbildung $T_0\colon B \to W$ gegeben.*
Dann lässt sich T_0 eindeutig zu einer linearen Abbildung $T\colon V \to W$ fortsetzen. Das heißt, es gibt genau eine lineare Abbildung $T\colon V \to W$ mit $T(b) = T_0(b)$ für alle $b \in B$.

Beweis. (a) Es sei $v \in V$. Da B erzeugend ist, existieren $n \in \mathbb{N}$, $\alpha_k \in \mathbb{K}$ und $b_k \in B$ für $k = 1,\dots,n$, sodass $v = \alpha_1 b_1 + \cdots + \alpha_n b_n$. Wir definieren

$$T(v) := \sum_{k=1}^{n} \alpha_k T_0(b_k).$$

Wegen der Eindeutigkeit der Linearkombination für v, ist $T(v)$ wohldefiniert; Man zeigt leicht die Linearität von T. ∎

Aufgabe 7. Zeigen Sie, dass der obige Satz nicht gilt, wenn (a) die Menge B nur erzeugend ist; (b) wenn B nur linear unabhängig ist. **Beweis in der Übung**

Beispiel 3.2 Es sei V ein $\mathbb{K}$-Vektorraum und $B = \{b_1,\dots,b_n\}$ eine Basis von V. Dann gibt es nach Lemma 3.2 genau eine lineare Abbildung

$$\Phi_B\colon V \to \mathbb{K}^n, \quad \Phi_B(b_k) = \mathrm{e}_k, \quad k = 1,\dots,n.$$

Offenbar ordnet diese Abbildung jedem Vektor $v = \alpha_1 b_1 + \alpha_2 b_2 + \cdots + \alpha_n b_n$ seine Koordinaten $(\alpha_1,\dots,\alpha_n)$ bezüglich der gegebenen Basis B zu. Diese Abbildung heißt *Koordinatenabbildung* von V bezüglich B. Die Koordinatenabbildung Φ_B ist bijektiv. Ihre Umkehrabbildung $\Phi_B^{-1}\colon \mathbb{K}^n \to V$ ist eindeutig bestimmt durch $\Phi_B^{-1}(\mathrm{e}_k) = b_k$, $k = 1,\dots,n$.

3.1.1 Die Matrix einer linearen Abbildung

Mit Hilfe von Koordinatenabbildungen lässt sich zu jeder linearen Abbildung T unter Auszeichnung von Basen im Ausgangsraum und im Bildraum eine Matrix zuordnen.

Definition 3.2 Gegeben sei eine lineare Abbildung $S\colon V \to W$ und Basen $B = \{b_1, \dots, b_n\}$ und $C = \{c_1, \dots, c_m\}$ von V bzw. von W. Dann gibt es genau eine Matrix $A \in \mathbb{K}^{m\times n}$, so dass das folgende Diagramm kommutativ ist

$$\begin{array}{ccc} V & \xrightarrow{S} & W \\ \Phi_B \Big\downarrow & & \Big\downarrow \Phi_C. \\ \mathbb{K}^n & \xrightarrow{T_A} & \mathbb{K}^m \end{array}$$

A heißt *Matrix von S bezüglich B und C*.

Hierbei ist T_A die der Matrix A zugeordnete lineare Abbildung $T_A\colon \mathbb{K}^n \to \mathbb{K}^m$, (vgl Beispiel 3.1). Laut Diagramm gilt $T_A = \Phi_C \circ S \circ \Phi_B^{-1}$. Dabei gilt mit $A = (a_{ij})_{\substack{i=1\dots m\\ j=1\dots n}}$

$$\boxed{S(b_j) = \sum_{i=1}^m a_{ij} c_i, \quad j = 1, \dots, n.}$$

In der Tat ist ja umgekehrt $S = \Phi_C^{-1} \circ T_A \circ \Phi_B$ und somit

$$S(b_j) = \Phi_C^{-1}(T_A(\Phi_B(b_j))) = \Phi_C^{-1}(A\cdot \mathrm{e}_j) = \Phi_C^{-1}\left(\sum_{i=1}^m a_{ij}\mathrm{e}_i\right) = \sum_{i=1}^m a_{ij}\Phi_C^{-1}(\mathrm{e}_i) = \sum_{i=1}^m a_{ij} c_i.$$

Wir bezeichnen die Matrix von S bezüglich B und C mit $A = M_{B,C}(S)$.

Beispiel 3.3 (a) Es sei $c \in \mathbb{K}$ fixiert und $S\colon V \to V$ gegeben durch $S(v) = cv$. Ferner sei $B = \{b_1, \dots, b_n\}$ eine Basis von V. Dann gilt $S = c\,\mathrm{id}_V$ ist linear und die Matrix von S bezüglich der Basis B lautet:

$$M_{B,B}(S) = \begin{pmatrix} c & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & c \end{pmatrix} = cI_n.$$

Dies ist das c-fache der Einheitsmatrix I_n. Es gilt nämlich $S(b_i) = cb_i$, $i = 1, \dots, n$.
(b) Es sei $P\colon \mathbb{R}^3 \to \mathbb{R}^2$ gegeben durch $P(x_1, x_2, x_3) = (2x_1 - x_3, x_1 + 2x_2 + 3x_3)$. Dann gilt

$$A := M_{B_3,B_2}(P) = \begin{pmatrix} 2 & 0 & -1 \\ 1 & 2 & 3 \end{pmatrix}$$

und $P = T_A$.
(c) Es sei $W = \mathrm{Abb}(\mathbb{R}, \mathbb{R})$ und $V = \mathbb{R}_2[x]$; $a \in \mathbb{R}$ sei fixiert. Die *Verschiebung* V_a (auch Translation oder Shift genannt), gegeben durch $V_a\colon W \to W$, $(V_a f)(x) = f(x-a)$, $f \in W$, ist eine bijektive lineare Abbildung. Wir betrachten die Einschränkung von V_a auf V und berechnen die

Matrix von V_a bezüglich der Basen $B = \{1, x, x^2\}$ im Ausgangsraum und $C = \{1, 1+x, 1+x+x^2\}$ im Bildraum. Es ist

$$\begin{aligned} V_a(1) &= 1, \\ V_a(x) &= x - a = (x+1) + (-a-1)1, \\ V_a(x^2) &= (x-a)^2 = x^2 - 2ax + a^2 = (x^2 + x + 1) + (-2a-1)(x+1) + (a^2 + 2a) \end{aligned}$$

Hieraus kann man die Matrixeinträge von $A = (a_{ij}) \in \mathbb{R}^{3\times 3}$ spaltenweise sofort ablesen:

$$M_{B,C}(V_a) = A = \begin{pmatrix} 1 & -1-a & a^2+2a \\ 0 & 1 & -2a-1 \\ 0 & 0 & 1 \end{pmatrix}.$$

Man beachte, dass $M_{B,C}(T)$ nicht nur von den Basen B und C abhängt, sondern auch von der Reihenfolge der Vektoren in der Basis. Wir betrachten also hier immer *geordnete* Basen.

3.2 Kern und Bild

3.2.1 Definition und Eigenschaften

Definition 3.3 Es sei $T\colon V \to W$ eine lineare Abbildung. Dann nennt man

$\operatorname{Im} T := T(V)$ das *Bild* von T,
$\operatorname{Ker} T := T^{-1}(\{0\})$ den *Kern* von T.

Als *Rang* von T bezeichnet man die Zahl $\operatorname{rg} T := \dim \operatorname{im} T = \dim T(V)$; als *Defekt* von T bezeichnet man die Zahl $\operatorname{def} T := \dim \operatorname{Ker} T$.

In der Tat sind nach Lemma 3.1 (d) $\operatorname{Im} T$ und $\operatorname{Ker} T$ lineare Teilräume von W bzw. von V, da V und $\{0\}$ Teilräume sind. Daher hat es Sinn, von deren Dimensionen zu reden.

Lemma 3.3 *Es sei $T \in \mathrm{L}(V, W)$ eine lineare Abbildung. Dann sind äquivalent:*

(a) T ist surjektiv.
(b) Ist B eine Basis von V, so ist $T(B)$ ein Erzeugendensystem von W.

Ist darüber hinaus W endlichdimensional, so ist mit (a) und (b) gleichwertig

(c) $\operatorname{rg} T = \dim W$.

Beweis. (a) $\Leftrightarrow$ (b). Wegen der Linearität von T ist

$$\operatorname{lin} T(B) = T(\operatorname{lin} B) = T(V).$$

Ist T surjektiv, so ist $T(V) = W$ und $T(B)$ ist ein Erzeugendensystem für W. Ist umgekehrt $\operatorname{lin} T(B) = W$, so ist T surjektiv. Ist nun $\dim W$ endlich, so folgt aus der Surjektivität von T, $\dim T(V) = \dim W$, also $\operatorname{rg} T = \dim W$. Ist umgekehrt $\dim W = \dim T(V)$, so folgt aus $T(V) \subset W$ und Lemma 2.6 (c) $T(V) = W$, also ist T surjektiv. ∎

Lemma 3.4 *Es sei $T \in \mathrm{L}(V,W)$ eine lineare Abbildung. Dann sind äquivalent:*

(a) T *ist injektiv.*
(b) $\operatorname{Ker} T = \{0\}$.
(c) $\operatorname{def} T = 0$.
(d) Ist B eine Basis von V, so ist $T(B)$ linear unabhängig.

Ist V endlichdimensional, so ist außerdem gleichwertig

(e) $\operatorname{rg} T = \dim V$.

Beweis. Die Gleichwertigkeit von (b) und (c) ist klar, da der einzige nulldimensionale Raum der nur aus dem Nullvektor bestehende Raum $\{0\}$ ist.
(a) $\Longrightarrow$ (b): Wegen der Injektivität von T folgt aus $T(x) = T(0) = 0$ sofort $x = 0$; also ist $\operatorname{Ker} T = \{0\}$.
(b) $\Longrightarrow$ (d): Es sei

$$\alpha_1 T(b_1) + \cdots + \alpha_n T(b_n) = 0$$

eine Linearkombination von $T(B)$ in W. Dann gilt $T(\alpha_1 b_1 + \cdots + \alpha_n b_n) = 0$ und wegen $\operatorname{Ker} T = \{0\}$ folgt weiter $\alpha_1 b_1 + \cdots + \alpha_n b_n = 0$. Da aber B linear unabhängig in V war, ist $\alpha_1 = \cdots = \alpha_n = 0$. Somit ist $T(B)$ linear unabhängig.
(d) $\Longrightarrow$ (a): Es sei nun B eine Basis von V und $v, v' \in V$ mögen die Darstellungen $v = \sum_{i=1}^n \alpha_i b_i$ und $v' = \sum_{i=1}^n \beta_i b_i$ mit $b_i \in B$ besitzen. Wenn nun $T(v) = T(v')$ gilt, so folgt

$$T\left(\sum_i \alpha_i b_i\right) = T\left(\sum_i \beta_i b_i\right) \Longrightarrow \sum_i \alpha_i T(b_i) = \sum_i \beta_i T(b_i).$$

Nun ist aber nach der Annahme in (d) auch $T(B)$ eine Basis in W. Nach Lemma 2.2 (b) ist aber die Koordinatendarstellung bezüglich einer Basis eindeutig; folglich gilt $\alpha_i = \beta_i$ für alle $i = 1, \dots, n$. Damit gilt $v = v'$ und T ist injektiv.
(d) $\Longrightarrow$ (e): Sei nun V endlichdimensional mit einer Basis $B = \{b_1, \dots, b_n\}$. Nach (d) ist dann auch $\{T(b_1), \dots, T(b_n)\}$ linear unabhängig in W, also $\dim T(V) = \operatorname{rg} T \geq n$. Umgekehrt gilt nach Lemma 3.1 (e) $\operatorname{rg} T \leq n$. Folglich ist $\operatorname{rg} T = n$.
(e) $\Longrightarrow$ (b): Sei $\operatorname{rg} T = n = \dim V$ und $w_1 = T(b_1), \dots, w_n = T(b_n)$ sei eine Basis von $\operatorname{Im}(T)$. Nach Lemma 3.1 (c) ist dann $B := \{b_1, \dots, b_n\}$ linear unabhängig in V. Wegen $n = \dim V$ ist B sogar Basis in V. Sei nun $T(v) = 0$ für ein $v \in V$, etwa für $v = \alpha_1 b_1 + \cdots + \alpha_n b_n$. Wendet man T darauf an und benutzt die Linearität von T und $T(0) = 0$, so folgt

$$\alpha_1 T(b_1) + \cdots + \alpha_n T(b_n) = 0.$$

Da aber $\{T(b_1), \dots, T(b_n)\}$ Basis in $T(V)$ ist, folgt $\alpha_1 = \cdots = \alpha_n = 0$. Somit ist $v = 0$ und $\operatorname{Ker} T = \{0\}$. ■

Beispiel 3.4 (a) $T\colon V \to V$ sei gegeben durch $T = c\,\mathrm{id}_V$. Im Fall $c = 0$ gilt $T = 0$ und damit $\operatorname{Ker} T = V$, $\operatorname{Im} T = \{0\}$, $\operatorname{rg} T = 0$, $\operatorname{def} T = n$. Im Falle $c \neq 0$ gilt:

$x \in \operatorname{Ker} T$ genau dann, wenn $T(x) = 0$, also $cx = 0$. Wegen $c \neq 0$ folgt $x = 0$; also $\operatorname{Ker} T = \{0\}$. Andererseits ist $\operatorname{Im} T = V$, denn jeder Vektor $w \in V$ hat als Urbild $\frac{1}{c}\, w \in V$. Somit gilt $\operatorname{Im} T = V$, $\operatorname{def} T = 0$, $\operatorname{rg} T = \dim V$.
(b) $T\colon \mathbb{R}^2 \to \mathbb{R}$ sei gegeben durch $T(x_1, x_2) = 2x_1 - x_2$. Dann gilt $\operatorname{Ker} T = \text{Lös}((2\ \ -1), 0) = \mathbb{R}\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ und $\operatorname{im} T = \mathbb{R}$, also $\operatorname{def} T = \operatorname{rg} T = 1$.

Wir fassen die Aussagen von Lemma 3.3 und Lemma 3.4 zusammen und erhalten eine Charakterisierung der bijektiven linearen Abbildungen (Isomorphismen).

Lemma 3.5 *Es sei $T \in \mathrm{L}(V, W)$ eine lineare Abbildung. Dann sind die folgenden Aussagen äquivalent:*

(a) T ist ein Isomorphismus.
(b) Wenn B eine Basis von V ist, so ist $T(B)$ eine Basis von W.

Besitzen darüber hinaus V und W endliche Dimension, so ist außerdem äquivalent dazu

(c) $\dim V = \operatorname{rg} T = \dim W$.

Definition 3.4 Zwei Vektorräume V und W über $\mathbb{K}$ heißen *isomorph*, symbolisch $V \cong W$, wenn es einen Isomorphismus $T\colon V \to W$ gibt.

Folgerung 3.6 *Isomorphe Vektorräume besitzen dieselbe Dimension. Gilt umgekehrt* $\dim V = \dim W < \infty$, *so sind V und W isomorph. Insbesondere sind alle Vektorräume V mit* $\dim V = n$ *isomorph zu $\mathbb{K}^n$.*

Beweis. Der erste Teil folgt direkt aus dem obigen Lemma. Seien nun V und W endlichdimensionale Vektorräume mit derselben Dimension n und seien etwa $B = \{b_1, \dots, b_n\}$ und $C = \{c_1, \dots, c_n\}$ Basen in V bzw. in W. Dann lässt sich die Abbildung $T_0\colon B \to W$, $T(b_k) = c_k$ eindeutig zu einem Isomorphismus $T\colon V \to W$ fortsetzen. Mit Hilfe der Koordinatenabbildungen lässt sich T schreiben als $T = \Phi_C^{-1} \circ \Phi_B$. ■

Beispiel 3.5 Es sei B_n die Standardbasis von $\mathbb{K}^n$, $A \in \mathbb{K}^{m\times n}$. Dann gilt $A = M_{B_n, B_m}(T_A)$. Das heißt, dass die Matrix zur linearen Abbildung T_A bezüglich der Standardbasis genau die Matrix A selbst ist. Dies bedeutet, dass es eine bijektive Beziehung zwischen den Matrizen $\mathbb{K}^{m\times n}$ und den linearen Abbildungen $\mathrm{L}(\mathbb{K}^n, \mathbb{K}^m)$ gibt. In der einen Richtung ist sie gegeben durch die Zuordnung $A \mapsto T_A$ und in der anderen Richtung durch $S \mapsto M_{B_n, B_m}(S)$. Diese beiden Zuordnungen sind sogar zueinander inverse *lineare* Abbildungen zwischen Matrizen und linearen Abbildungen. Da bei linearen Isomorphismen Basen in Basen über gehen sind die Dimensionen beider Räume gleich

$$\dim \mathrm{L}(\mathbb{K}^n, \mathbb{K}^m) = \dim \mathbb{K}^{m\times n} = m\,n.$$

Unendlichdimensionale Räume sind nicht notwendig isomorph, etwa $\mathrm{Abb}(\mathbb{R}, \mathbb{R}) \not\cong \mathbb{R}[x]$. Aber $T\colon \mathbb{R}[x] \to \varphi$, $T(a_n x^n + \cdots + a_1 x + a_0) = (a_0, a_1, \dots, a_n, 0, 0, \dots)$ ist ein Isomorphismus.

3.2.2 Die Dimensionsformel für lineare Abbildungen

Ist $T \in \mathrm{L}(V, W)$, so kann man die Zahlen $\operatorname{def} T$ und $\operatorname{rg} T$ betrachten. Erstaunlicherweise hängt $\operatorname{def} T + \operatorname{rg} T$ gar nicht von T ab.

Satz 3.7 (Dimensionsformel für lineare Abbildungen) *Es sei V ein endlichdimensionaler Vektorraum. Für jede lineare Abbildung $T\colon V \to W$ gilt dann:*

$$\boxed{\operatorname{rg} T + \operatorname{def} T = \dim V.}$$

Beweis. Als Teilraum von V besitzt $\operatorname{Ker} T$ auch eine endliche Dimension; sei etwa $K = \{b_1, \ldots, b_d\}$ eine Basis von $\operatorname{Ker} T$. Nach Lemma 2.6 kann man K zu einer Basis $\{b_1, \ldots, b_d, b_{d+1}, \ldots, b_n\}$ von V ergänzen. Wir werden zeigen, dass $E := \{T(b_{d+1}), \ldots, T(b_n)\}$ eine Basis von $\operatorname{Im} T$ ist.
E ist erzeugend für $\operatorname{im} T$. Es sei $T(v) \in \operatorname{im} T$ mit $v = \sum_{i=1}^n \alpha_i b_i$. Wegen der Linearität von T und wegen $T(b_1) = T(b_2) = \cdots = T(b_d) = 0$ folgt

$$T(v) = T\left(\sum_{i=1}^n \alpha_i b_i\right) = \sum_{i=1}^n \alpha_i T(b_i) = \sum_{i=d+1}^n \alpha_i T(b_i) \in \operatorname{lin} E.$$

E ist linear unabhängig. Sei

$$\sum_{i=d+1}^n \alpha_i T(b_i) = 0 \Longrightarrow T\left(\sum_{i=d+1}^n \alpha_i b_i\right) = 0 \Longrightarrow \sum_{i=d+1}^n \alpha_i b_i \in \operatorname{Ker} T.$$

Folglich existieren $\alpha_i \in \mathbb{K}$, $i = 1, \ldots, d$ mit

$$\sum_{i=d+1}^n \alpha_i b_i = \sum_{i=1}^d \alpha_i b_i.$$

Da aber $\{b_1, \ldots, b_n\}$ eine Basis von V ist, folgt hieraus $\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$. Also ist E linear unabhängig und damit eine Basis von $\operatorname{Im} T$. Somit gilt

$$\operatorname{rg} T = \dim \operatorname{Im} T = |\{b_{d+1}, \ldots, b_n\}| = n - d = \dim V - \operatorname{def} T.$$

∎

Beispiel 3.6 Bestimmen Sie $\operatorname{Im} T$, $\operatorname{Ker} T$, $\operatorname{rg} T$ und $\operatorname{def} T$.
(a) Es sei $T = T_A$, $c \neq 0$, mit

$$A = \begin{pmatrix} c & \ldots & c \\ c & \ldots & c \\ \vdots & & \vdots \\ c & \ldots & c \end{pmatrix} \in \mathbb{R}^{n\times n}.$$

Dann gilt $\operatorname{Im} A = \mathbb{R}\begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$, $\operatorname{Ker} T = \{x \in \mathbb{R}^n \mid Ax = 0\} = \{(x_1, \ldots, x_n) \in \mathbb{R}^n \mid x_1 + x_2 + \cdots + x_n = 0\}$.
Folglich gilt $\operatorname{rg} T = 1$ und $\operatorname{def} T = n - 1$. Die Matrix besitzt nur eine linear unabhängige Zeile oder Spalte. Für $c = 0$ wäre $\operatorname{rg} T = 0$.

(b) Sei $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$. Dann liefert der Gauß-Algorithmus

$$A \sim \begin{pmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & -6 & -12 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix}$$

Folglich ist $\operatorname{rg} A = 2$ und $\operatorname{def} A = 1$.
(c) Es sei $x = (x_1, \dots, x_n) \in \mathbb{K}^{1\times n}$ ein Zeilenvektor und $y = (y_1, \dots, y_m)^\top \in \mathbb{K}^{m\times 1}$ ein Spaltenvektor. Dann hat die Matrix

$$A = y \cdot x = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix} \cdot (x_1, \dots, x_n) = (y_i x_j)_{\substack{i=1\dots m \\ j=1\dots n}}$$

den Rang 1 oder 0. Dies folgt aus $\operatorname{rg} x, \operatorname{rg} y \le 1$ und der Übungsaufgabe 6.5.

Satz 3.8 (Äquivalenzsatz) *Es sei $T \in \mathrm{L}(V, W)$ eine lineare Abbildung zwischen endlichdimensionalen Vektorräumen gleicher Dimension,* $\dim V = \dim W < \infty$. *Dann sind äquivalent:*

(a) T ist injektiv.
(b) T ist surjektiv.
(c) T ist bijektiv.

Beweis. (a) $\to$ (b). Nach Lemma 3.4(e) gilt $\operatorname{rg} T = \dim V$, also nach Voraussetzung auch $\operatorname{rg} T = \dim W$. Nach Lemma 3.3 folgt dann T ist surjektiv.
(b) $\to$ (c). Nach den beiden Lemmas und nach Voraussetzung ist wieder $\operatorname{rg} T = \dim W = \dim V$, also ist T injektiv und damit bijektiv. Schließlich ist die Richtung (c) $\to$ (a) trivial. ∎

Der Satz gilt nicht mehr in unendlich-dimensionalen Räumen, wie das folgende Beispiel des *Verschiebungsoperators* $T\colon \omega \to \omega$ zeigt: $T(x) := T((x_1, x_2, \dots)) := (0, x_1, x_2, \dots)$. Offenbar ist T injektiv aber nicht surjektiv.

3.3 Matrizenrechnung

3.3.1 Rang und Defekt einer Matrix

Wir haben bereits in Kapitel 1.2 gesehen, wie Matrizen addiert, multipliziert und skalar vervielfacht werden. Unter dem Gesichtspunkt, dass wir Matrizen $A \in \mathbb{K}^{m\times n}$ mit ihren linearen Abbildungen $T_A \in \mathrm{L}(\mathbb{K}^n, \mathbb{K}^m)$ identifizieren, können wir die eben eingeführten Begriffe, Kern, Bild, Rang, Defekt aber auch Invertierbarkeit direkt auf Matrizen übertragen.

Definition 3.5 Für eine Matrix $A \in \mathbb{K}^{m\times n}$ definieren wir
(a) den *Kern* von A $\operatorname{Ker} A = \operatorname{Ker} T_A \subset \mathbb{K}^n$, das *Bild* von A, $\operatorname{Im} A = \operatorname{Im} T_A \subset \mathbb{K}^m$, den *Rang* $\operatorname{rg} A = \operatorname{rg} T_A$ und den *Defekt* von A über $\operatorname{def} A = \operatorname{def} T_A$.
(b) Eine Matrix heißt *quadratisch*, wenn ihre Zeilen- und Spaltenzahl überein stimmen.
(c) $\operatorname{Im} A$ heißt auch *Spaltenraum* von A und $\operatorname{Im} A^\top$ nennt man den *Zeilenraum* von A.

Bemerkung 3.2 (a) Es ist klar, dass $\operatorname{Ker} A = \{x \in \mathbb{K}^n \mid Ax = 0\} = \text{Lös}(A,0)$ genau die Lösungsmenge des homogenen linearen Gleichungssystems $Ax = 0$ ist.
(b) Dagegen ist $\operatorname{Im} A = \{b \in \mathbb{K}^m \mid \exists x \in \mathbb{K}^n : Ax = b\}$ die Menge der Vektoren b in $\mathbb{K}^m$, für die das inhomogene lineare Gleichungssystem $Ax = b$ eine Lösung besitzt.
(c) Nach unserer Definition ist der Spaltenraum $\operatorname{Im} A$ der von den Spalten der Matrix A aufgespannte Teilraum im $\mathbb{K}^m$, $\operatorname{Im} A = \operatorname{lin}\{A\mathrm{e}_1, \dots, A\mathrm{e}_n\}$. Somit gilt einerseits

$$\operatorname{rg} A = \dim \operatorname{lin}\{A\mathrm{e}_1, \dots, A\mathrm{e}_n\} = \text{Dimension des Spaltenraumes von } A.$$

Da n Vektoren höchstens einen Raum der Dimension n aufspannen, gilt $\operatorname{rg} A \le n$. Andererseits ist $\operatorname{Im} A \subset \mathbb{K}^m$ und damit auch $\operatorname{rg} A \le m$. Folglich gilt $\operatorname{rg} A \le \min\{m,n\}$.

Ebenso könnte man aber den *Zeilenraum* von A betrachten, die lineare Hülle des von den Zeilen von A aufgespannten Unterraumes von $\mathbb{K}^n$. Natürlich ist der Zeilenraum von A gleich dem Spaltenraum von $A^\top$, *Zeilenraum* von $A = \operatorname{Im} A^\top$.
Der Gauß-Jordan-Algorithmus verändert den Zeilenraum nicht, denn durch die elementaren Zeilenoperationen bleibt die lineare Hülle der Zeilen erhalten.

Satz 3.9 (Zeilenrang = Spaltenrang) *Für jede Matrix $A \in \mathbb{K}^{m\times n}$ gilt*

$$\operatorname{rg} A = \operatorname{rg} A^\top.$$

Beweis. Wir wollen eine Zeile oder Spalte von A als *linear überflüssig* bezeichnen, wenn sie die Linearkombination von anderen Zeilen bzw. Spalten ist. Verkleinert man eine Matrix durch Weglassen einer linear überflüssigen Spalte, so ändert sich der (Spalten)rang von A nicht.
(a) Wir werden zeigen, dass sich auch *der Zeilenrang dabei nicht ändert.* — Angenommen, die 1.Spalte der Matrix sei linear überflüssig, dann lässt sich die erste Komponente jeder Zeile linear kombinieren aus allen anderen Komponenten der Zeile. Dasselbe gilt für jede Linearkombination von Zeilen: auch hier lässt sich die erste Komponente linear aus den anderen Komponenten kombinieren (und zwar mit denselben Koeffizienten, wie bei allen Zeilen). Somit ist eine Linearkombination von Zeilen genau dann Null, wenn schon die entsprechende Linearkombination der Zeilen, die durch Weglassen der ersten Komponente entstehen, Null ist. Daher ändert sich der Zeilenrang nicht durch weglassen linear überflüssiger Spalten.
(b) Dasselbe Argument gilt natürlich, wenn man die Rolle von Zeilen und Spalten vertauscht: Durch Weglassen linear überflüssiger Zeilen verringert sich der Spaltenrang nicht.
(c) Nun verkleinern wir unsere Matrix A schrittweise durch Weglassen linear überflüssiger Zeilen und Spalten solange, bis es nicht weiter geht. Wir erhalten eine möglicherweise kleinere Matrix A', die denselben Zeilenrang und denselben Spaltenrang wie A hat. Also

$\mathrm{rg}\,A = \mathrm{rg}\,A'$ und $\mathrm{rg}\,A^\top = \mathrm{rg}\,A'^\top$ und A' hat nur linear unabhängige Zeilen und linear unabhängige Spalten.
(d) *A' ist eine quadratische Matrix.* Angenommen, A' hat s Spalten und z Zeilen, dann gilt wegen der linearen Unabhängigkeit der Spalten $\mathrm{rg}\,A' = s \leq z$. Dasselbe Argument gilt aber für den Zeilenraum, da die Zeilen von A' linear unabhängig sind: $\mathrm{rg}\,A'^\top = z \leq s$. Also ist $\mathrm{rg}\,A = \mathrm{rg}\,A' = s = z =: r$ und es gilt Zeilenrang von A = Spaltenrang von $A = r$= Format der quadratischen Matrix A'. ■

Bemerkung 3.3 Wir sahen bereits in Abschnitt 2.3.6, dass durch die 3 elementaren Zeilenoperationen des Gauß-Algorithmus, der Zeilenraum nicht verändert wird. Wir erhalten daher:

$$\boxed{\mathrm{rg}\,A = r = \text{Anzahl der führenden Einsen im Gauß-Algorithmus.}}$$

Der Gauß-Algorithmus (ohne Gauß-Jordan) ist daher bestens geeignet, den Rang einer Matrix zu ermitteln. Nach dem Dimensionssatz für lineare Abbildungen gilt ferner

$$\boxed{\begin{array}{l}\mathrm{def}\,A = n - r = \\ = \text{Anzahl der frei wählbaren Variablen} = \\ = \text{Dimension des Lösungsraumes des homogenen Systems } Ax = 0 \\ = \text{Anzahl der Spalten ohne eine führende Eins.}\end{array}}$$

Insbesondere ist eine Matrix

$$\begin{array}{lcll} \text{injektiv} & \Leftrightarrow & \text{Jede Spalte hat eine führende Eins.} & \Leftrightarrow r = n. \\ \text{surjektiv} & \Leftrightarrow & \text{Jede Zeile hat eine führende Eins.} & \Leftrightarrow r = m. \end{array}$$

Die Matrix $A \in \mathbb{R}^{4\times 5}$, $m = 4$, $n = 5$, liegt in reduzierter Zeilenstufenform vor:

$$\begin{pmatrix} 1 & 2 & 0 & 4 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Da genau zwei führende Einsen vorliegen, ist $\mathrm{rg}\,A = 2$. Es gibt 3 Spalten ohne führende Einsen zu den freien Variablen x_2, x_4, x_5. Daher gilt $\mathrm{def}\,A = 5 - 2 = 3$. Man liest das Bild der Matrix als lineare Hülle der Spalten ab:

$$\mathrm{Im}\,A = \mathrm{lin}\left\{\begin{pmatrix}1\\0\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\\0\end{pmatrix}\right\} \subset \mathbb{R}^4.$$

Als Lösungsmenge von $Ax = 0$ liest man ebenfalls ab:

$$\mathrm{Ker}\,A = \left\{\begin{pmatrix}x_1\\x_2\\x_3\\x_4\\x_5\end{pmatrix} \,\middle|\, x_1 = -2x_2 - 4x_4, \quad x_3 = 0\right\} = \mathbb{R}\begin{pmatrix}-2\\1\\0\\0\\0\end{pmatrix} + \mathbb{R}\begin{pmatrix}0\\0\\0\\1\\0\end{pmatrix} + \mathbb{R}\begin{pmatrix}0\\0\\0\\0\\1\end{pmatrix}.$$

3.3.2 Die inverse Matrix

Wir wissen aus Abschnitt 1.1.4, dass nur die bijektiven Abbildungen invertierbar sind. Das gilt natürlich auch für lineare Abbildungen zwischen Vektorräumen. Nach Lemma 3.5 (c) kann T_A, $A \in \mathbb{K}^{m\times n}$, nur dann ein Isomorphismus sein, wenn $m = n$ ist und somit muss A quadratisch sein. Nur für quadratische Matrizen hat der Begriff der Invertierbarkeit Sinn. die zu $(T_A)^{-1}$ korrespondierende Matrix bezeichnen wir mit B (falls sie existiert). Schreibt man die Gleichung in linearen Abbildungen $T_A\circ(T_A)^{-1} = (T_A)^{-1}\circ T_A = \mathrm{id}_{\mathbb{K}^n}$ matriziell, so lautet sie $A\cdot B = B\cdot A = I_n$, wobei B die zu $(T_A)^{-1}$ gehörige Matrix ist. Wir definieren B als die zu A inverse Matrix.

Definition 3.6 (a) Eine Matrix $A \in \mathbb{K}^{n\times n}$ heißt *invertierbar*, wenn es eine Matrix $A^{-1} \in \mathbb{K}^{n\times n}$ gibt, für die

$$A\cdot A^{-1} = A^{-1}\cdot A = I_n.$$

Dabei ist I_n die Einheitsmatrix der Ordnung n. Wir nennen A^{-1} die zu A *inverse Matrix*.
(b) Die Menge der invertierbaren $n \times n$-Matrizen bezeichnen wir mit $\mathrm{GL}(n,\mathbb{K})$.

Invertierbare Matrizen nennt man auch *regulär* während die nichtinvertierbaren Matrizen *singulär* heißen.

Bemerkung 3.4 (a) In der Tat ist A^{-1} eindeutig durch A bestimmt, denn gäbe es ein $B \in \mathbb{K}^{n\times n}$ mit derselben Eigenschaft, so wäre $B = BI_n = B(AA^{-1}) = (BA)A^{-1} = I_nA^{-1} = A^{-1}$.
(b) Aus der Gleichung für die entsprechenden linearen Abbildungen folgt (nach Beispiel 3.1 (a2))

$$T_A\circ T_{A^{-1}} = \mathrm{id}_{\mathbb{K}^n} = T_{A^{-1}}\circ T_A.$$

Das heißt, A ist invertierbar genau dann, wenn T_A invertierbar ist und in diesem Falle gilt $(T_A)^{-1} = T_{A^{-1}}$.

Lemma 3.10 *Es seien $A, B \in \mathrm{GL}(n,\mathbb{K})$.*
(a) Dann ist auch $AB \in \mathrm{GL}(n,\mathbb{K})$ und es gilt $(AB)^{-1} = B^{-1}A^{-1}$.
(b) $A^{-1}, A^\top \in \mathrm{GL}(n,\mathbb{K})$ und $(A^{-1})^{-1} = A$ und $(A^\top)^{-1} = (A^{-1})^\top$.

Beweis. (a) Wegen $AB\cdot B^{-1}A^{-1} = A(BB^{-1})A^{-1} = A\cdot I_n\cdot A^{-1} = A\cdot A^{-1} = I_n$ und analog $B^{-1}A^{-1}\cdot AB = I_n$ folgt die Behauptung. (b) Folgt genauso wie $-(-x) = x$ und $\frac{1}{\frac{1}{x}} = x$. Ferner ist $(A^{-1})^\top\cdot A^\top = (AA^{-1})^\top = I_n^\top = I_n$. Hieraus folgt die Behauptung. ∎

Lemma 3.11 *Es sei $A \in \mathbb{K}^{n\times n}$. Dann sind äquivalent*
(a) A ist invertierbar.
(b) $\mathrm{rg}\, A = n$.
(c) $\mathrm{def}\, A = 0$.
(d) Die Spalten von A bilden eine Basis von $\mathbb{K}^n$.
(e) Die Zeilen von A bilden eine Basis von $\mathbb{K}^n$.
(f) T_A ist injektiv.
(g) T_A ist surjektiv.

Beweis. (a) $\longleftrightarrow$ (f) $\longleftrightarrow$ (g): Ist A invertierbar, so ist T_A invertierbar, also bijektiv. Nach dem Äquivalenzsatz ist dies gleichwertig mit (f) und (g). Nach Lemma 3.4 sind (c) und (f) gleichwertig und nach Lemma 3.3 sind (b) und (g) äquivalent. Ferner ist (g) äquivalent zu (d). Die Äquivalenz von (d) und (e) folgt schließlich aus Satz 3.9. ∎

3.3.3 Invertieren einer Matrix

Nach Lemma 3.11 sind genau die quadratischen Matrizen invertierbar, die den vollen Rang n haben. Anders ausgedrückt, der Gauß-Algorithmus, angewandt auf A liefert n führende Einsen; jedes lineare Gleichungssystem $A\cdot x = b$, $b \in \mathbb{R}^n$ ist eindeutig lösbar (universelle Lösbarkeit für alle b). Man kann diese Lösung sofort formal angeben, da ja A^{-1} existiert. Multipliziert man die Gleichung $A\cdot x = b$ von links mit A^{-1}, so erhält man

$$\begin{aligned} A^{-1}\,| \quad & A\cdot x = b \\ & A^{-1}\cdot(A\cdot x) = (A^{-1}\cdot A)\cdot x = I_n\cdot x = x = A^{-1}\cdot b. \end{aligned}$$

Die eindeutig bestimmte Lösung von $A \cdot x = b$ ist also $x = A^{-1} \cdot b$. Genau das nutzt man zur effektiven Berechnung der Matrix A^{-1}. Ist nämlich $b = \mathrm{e}_j$, $j = 1, \ldots, n$, und ist $x^{(j)}$ die Lösung des linearen Gleichungssystems $Ax = \mathrm{e}_j$, so gilt

$$Ax^{(j)} = \mathrm{e}_j \Leftrightarrow x^{(j)} = A^{-1}\mathrm{e}_j \Leftrightarrow x^{(j)} \text{ ist die } j\text{te Spalte der Matrix } A^{-1}.$$

Die letzte Feststellung folgt aus Beispiel 3.1 (a). Löst man also alle Gleichungssysteme $Ax = \mathrm{e}_j$, $j = 1, \ldots, n$, so erhält man die Spalten von A^{-1}. Somit ist das Verfahren zur Matrixinvertierung klar:

Verfahren zur Berechnung von A^{-1}. Gegeben sei eine Matrix $A \in \mathbb{K}^{n\times n}$. Man wendet den Gauß-Jordan-Algorithmus auf die erweiterte Koeffizientenmatrix an, wobei als rechte Seiten die n Spalteneinheitsvektoren genommen werden.
Entsteht dabei auf der Koeffizientenseite (linke Seite) eine halbe Nullzeile, so ist der Rang der Matrix A kleiner als n und die Matrix ist nicht invertierbar. Andernfalls entstehen n führende Einsen und die Koeffizientenmatrix A wird in die Einheitsmatrix I_n transformiert. Auf der rechten Seite kann man dann die inverse Matrix A^{-1} ablesen.

Beispiel 3.7 Gegeben sei die reelle 4×4- Matrix

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

Wir berechnen A^{-1} durch simultanes Lösen von 4 linearen Gleichungssystemen:

1	0	1	1	1	0	0	0	$\cdot(-1)$ Add. zu Zeile 2 und 4
1	1	2	1	0	1	0	0	
0	-1	0	1	0	0	1	0	
1	0	0	2	0	0	0	1	
1	0	1	1	1	0	0	0	
0	1	1	0	-1	1	0	0	Addition zu Zeile 3
0	-1	0	1	0	0	1	0	
0	0	-1	1	-1	0	0	1	
1	0	1	1	1	0	0	0	
0	1	1	0	-1	1	0	0	
0	0	1	1	-1	1	1	0	
0	0	-1	1	-1	0	0	1	
1	0	1	1	1	0	0	0	
0	1	1	0	-1	1	0	0	
0	0	1	1	-1	1	1	0	(-1) Add zu Zeile 2 und 1
0	0	0	2	-2	1	1	1	
1	0	0	0	2	-1	-1	0	
0	1	0	-1	0	0	-1	0	
0	0	1	1	-1	1	1	0	
0	0	0	1	-1	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	Add zur 2., Subtr. von der 3.Zeile
1	0	0	0	2	-1	-1	0	
0	1	0	0	-1	$\frac{1}{2}$	$-\frac{1}{2}$	$\frac{1}{2}$	
0	0	1	0	0	$\frac{1}{2}$	$\frac{1}{2}$	$-\frac{1}{2}$	
0	0	0	1	-1	$\frac{1}{2}$	$\frac{1}{2}$	$-\frac{1}{2}$	

Somit gilt:

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

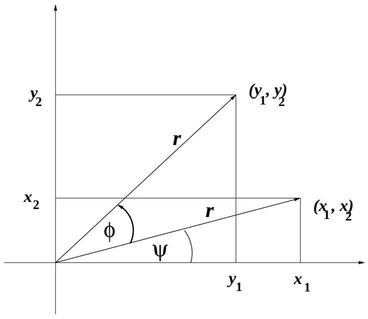
3.3.4 Spezielle Klassen von Matrizen

Es sei $\mathbb{K} = \mathbb{R}$ oder oder $\mathbb{K} = \mathbb{C}$.

Definition 3.7 Eine quadratische Matrix $A \in \mathbb{K}^{n \times n}$ heißt

symmetrisch, falls $A = A^\top$, also $a_{ij} = a_{ji}$ für alle $i, j = 1, \ldots, n$.
hermitesch, falls $A = A^*$, dabei ist $A^* = (\overline{A})^\top$, also $a_{ij} = \overline{a_{ji}}$ (komplexe Konjugation) für alle $i, j = 1, \ldots, n$.
unitär, falls $A^* A = A A^* = I_n$.
Eine unitäre Matrix mit nur reellen Einträgen heißt *orthogonal*. Für orthogonale Matrizen gilt also $A A^\top = A^\top A = I_n$.

Beispiel 3.8 (Drehung um φ) Es sei D_φ die Drehung des $\mathbb{R}^2$ um den Ursprung um den Winkel φ. Wir werden sehen, dass dies in der Tat eine lineare Abbildung ist, deren Matrix bestimmen und zeigen, dass D_φ orthogonal ist.



Die Koordinate y_1 erhält man als $y_1 = r\cos(\varphi+\psi) = r\cos\varphi\sin\psi - r\sin\varphi\cos\psi$. Nun ist aber $r\cos\psi = x_1$ und $r\sin\psi = x_2$. Also hat man $y_1 = \cos\varphi x_1 - \sin\varphi x_2$. Analog berechnet man y_2

$$\begin{aligned} y_1 &= r\cos(\varphi+\psi) = x_1\cos\varphi - x_2\sin\varphi, \\ y_2 &= r\sin(\varphi+\psi) = x_1\sin\varphi + x_2\cos\varphi. \end{aligned}$$

Man erkennt, dass die Abbildung $D_\varphi\colon \mathbb{R}^2 \to \mathbb{R}^2$, $D_\varphi\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$ linear ist mit der Matrix M_φ:

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

Zur Vereinfachung der Symbolik identifizieren wir meist die lineare Abbildung D_φ und ihre Matrix M_φ bezüglich der Standardbasis des $\mathbb{R}^2$. Man nennt M_φ die Drehmatrix um den Winkel φ. Es gilt

$$M_\varphi \cdot M_\varphi^\top = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix} \cdot \begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix} = \begin{pmatrix} \cos^2\varphi + \sin^2\varphi & 0 \\ 0 & \cos^2\varphi + \sin^2\varphi \end{pmatrix} = I_2.$$

Das heißt, M_φ ist eine orthogonale 2×2-Matrix. In der Mechanik werden Drehbewegungen im Raum , das heißt, orthogonale Transformationen $T \in \mathrm{GL}(3,\mathbb{R})$ durch die *Eulerschen Winkel* beschrieben, siehe auch `http://de.wikipedia.org/wiki/Eulersche_Winkel`. Jede Drehung im dreidimensionalen Raum $\mathbb{R}^3$ um den Ursprung ist die Hintereinanderausführung von drei ebenen Drehungen: eine um den Winkel ψ in der xy-Ebene, eine um den Winkel θ in der yz-Ebene und schließlich erneut eine Drehung in der xy-Ebene um den Winkel φ.

Später werden wir die in Definition 3.7 eingeführten Begriffe auch für lineare Abbildungen definieren.

3.3.5 Koordinatentransformationen und Basistransformationen

Problem: wie verändern sich die Koordinaten eines Vektors, wenn man zu einer anderen Basis über geht?

Es sei V ein Vektorraum mit der Basis $E = \{e_1, \ldots, e_n\}$. Dann gibt es für jeden Vektor $v \in V$ eindeutig bestimmte Zahlen, die Koordinaten von v bezüglich E, $(x_1, \ldots, x_n)$ mit $v = x_1 e_1 + \cdots + x_n e_n$. Sei nun $F = \{f_1, \ldots, f_n\}$ eine weitere Basis von V und $v = y_1 f_1 + \cdots + y_n f_n$.
Problem: Wie berechnen wir die neuen Koordinaten $(y_1, \ldots, y_n)$ aus den alten Koordinaten $(x_1, \ldots, x_n)$? Die Basistransformation sei gegeben durch eine Matrix $A \in \mathbb{K}^{n\times n}$ in der folgen-

den Weise:

$$f_j = \sum_{i=1}^{n} a_{ij} e_i, \quad j = 1, \ldots, n; \tag{3.3}$$

dabei sei $A = (a_{ij}) \in \mathbb{K}^{n\times n}$ die *Matrix der Basistransformation.* Für festes j schreiben wir also die Koordinaten a_{ij} von f_j bezüglich E in die $j$te *Spalte* der Matrix A. Da F eine Basis ist, gibt es auch eine Matrix $B = (b_{ij}) \in \mathbb{K}^{n\times n}$ mit

$$e_i = \sum_{k=1}^{n} b_{ki} f_k, \quad i = 1, \ldots, n, \tag{3.4}$$

welche die umgekehrte Basistransformation beschreibt. Setzt man (3.4) in (3.3) ein, so hat man:

$$f_j = \sum_{i=1}^{n} \sum_{k=1}^{n} a_{ij} b_{ki}\, f_k, \quad j = 1, \ldots, n.$$

Vertauscht man die Summationen und beachtet, dass F eine Basis in V ist, so hat man

$$f_j = \sum_{k=1}^{n} \left(\sum_{i=1}^{n} b_{ki} a_{ij} \right) f_k = \sum_{k=1}^{n} (BA)_{kj} f_k \longrightarrow (BA)_{kj} = \delta_{kj}, \quad k, j = 1, \ldots, n.$$

Somit gilt $BA = I_n$ — die Matrizen A und B sind zueinander invers (und insbesondere regulär); $B = A^{-1}$. Durch Einsetzen der Basistransformation (3.3) in die Gleichung $v = y_1 f_1 + \cdots + y_n f_n$ erhalten wir eine Beziehung zwischen den Koordinaten x_i und y_j:

$$v = \sum_{j=1}^{n} y_j f_j = \sum_{j=1}^{n} y_j \sum_{i=1}^{n} a_{ij} e_i = \sum_{i=1}^{n} \left(\sum_{j=1}^{n} a_{ij} y_j \right) e_i = \sum_{i=1}^{n} x_i e_i.$$

Da die Koordinaten bezüglich der Basis $\{e_i\}$ eindeutig bestimmt sind, folgt $x_i = \sum_{j=1}^{n} a_{ij} y_j$, $i = 1, \ldots, n$, oder matriziell $x = Ay$. Da A regulär ist, gilt auch $y = A^{-1}x$. Man beachte, dass in der Definition der Matrix A über (3.3) die *transponierte* Matrix einging. 'Formal' gilt $f = A^\top e$ und $e = (A^{-1})^\top f$.

Beispiel 3.9 Wir betrachten das Beispiel aus der Übungsaufgabe 5.3 genauer: Die Vektoren $b_1 = (1, 2, 1)$, $b_2 = (2, 9, 0)$ und $b_3 = (3, 3, 4)$ bilden eine Basis B des $\mathbb{R}^3$. Bestimmen Sie die neuen Koordinaten (y_1, y_2, y_3) bezüglich B, wenn die alten Koordinaten (x_1, x_2, x_3) in der Standardbasis gegeben sind.
Wir lesen die Matrix $A = (a_{ij}) \in \mathbb{R}^{3\times 3}$ der Basistransformation spaltenweise ab:

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 9 & 3 \\ 1 & 0 & 4 \end{pmatrix}$$

Mittels Gauß-Algorithmus bestimmt man die inverse Matrix

$$A^{-1} = \begin{pmatrix} -36 & 8 & 21 \\ 5 & -1 & -3 \\ 9 & -2 & -5 \end{pmatrix}$$

Somit gilt

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = y = A^{-1}x = \begin{pmatrix} -36x_1 + 8x_2 + 21x_3 \\ 5x_1 - x_2 - 3x_3 \\ 9x_1 - 2x_2 - 5x_3 \end{pmatrix}$$

Nun gilt $y_1 b_1 + y_2 b_2 + y_3 b_3 = x_1 \mathrm{e}_1 + x_2 \mathrm{e}_2 + x_3 \mathrm{e}_3$.

3.3.6 Verhalten von Matrizen bei Basistransformation

Problem: Gegeben sei eine lineare Abbildung $T \in \mathrm{L}(V, W)$. Man ermittle Basen B und C in V bzw. in W, so dass die zugehörige Matrix $M_{B,C}(T)$ eine *möglichst einfache* Gestalt hat. Wir werden sehen, dass dieses Problem eine sehr einfache Lösung besitzt (Normalformenproblem).
Wir haben gesehen, wie sich die Koordinaten ändern, wenn man zu einer neuen Basis über geht. Nun wollen wir uns überlegen, wie die neue Matrix M' einer linearen Abbildung $T \in \mathrm{L}(V; W)$ aussieht, wenn $M = M_{E,F}(T)$ die Darstellung von T bezüglich gegebener Basen E und F von V bzw. von W ist und wenn man in V und W zu neuen Basen E' bzw. F' über geht. Wir wollen also $M' = M_{E',F'}(T)$ aus M und den Transformationsmatrizen bestimmen.

A, B	Matrizen der Basistransformationen in V bzw. in W.
x, y	Koordinaten von $v \in V$ bzw. von $T(v) \in W$ bzgl. der Basen E und F.
x', y'	Koordinaten von $v \in V$ und $T(v) \in W$ bzgl. der Basen E' und F'.
M, M'	Matrizen von T bezüglich (E, F) bzw. (E', F').

Nach diesen Setzungen gilt mit Hilfe der Transformationsformeln

$$y = Mx \quad y' = M'x' \quad x' = A^{-1}x \quad y' = B^{-1}y.$$

Multipliziert man die erste Gleichung mit B^{-1} von links und setzt $x = Ax'$ ein, so hat man

$$B^{-1}y = B^{-1}Mx = B^{-1}MAx' \quad \Longrightarrow \quad y' = (B^{-1}MA)x',$$

sodass man unmittelbar die Transformationsformel erhält

$$M' = B^{-1} M A. \tag{3.5}$$

Besonders wichtig ist der Fall von Endomorphismen $T \in \mathrm{L}(V)$, wo man üblicherweise $E = F$ und $E' = F'$ wählt, sodass $A = B$ gilt. Hier lautet die Transformationsformel

$$M' = A^{-1} M A.$$

Definition 3.8 Zwei quadratische Matrizen $M, M' \in \mathbb{K}^{n \times n}$ heißen einander *ähnlich*, wenn es eine invertierbare Matrix $A \in \mathrm{GL}(n, \mathbb{K})$ gibt mit $M' = A^{-1}MA$.

Ähnliche Matrizen beschreiben ein und dieselbe lineare Abbildung $T\colon \mathbb{K}^n \to \mathbb{K}^n$ nur in verschiedenen Basen.

Beispiel 3.10 In Übungsaufgabe 5.3 war eine lineare Abbildung $T\colon \mathbb{R}^3 \to \mathbb{R}^2$ gegeben durch

$$T(b_1)=(1,0), \quad T(b_2)=(-1,1), \quad T(b_3)=(0,1),$$

wobei die Basis $B=\{b_1,b_2,b_3\}$ gegeben ist durch $b_1=(1,2,1)$, $b_2=(2,9,0)$ und $b_3=(3,3,4)$. Somit gilt

$$M=M_{B,B_2}(T)=\begin{pmatrix}1 & -1 & 0\\ 0 & 1 & 1\end{pmatrix},$$

wobei B_2 die Standardbasis vom $\mathbb{R}^2$ ist. Gesucht ist die Darstellung von T bezüglich der Standardbasen B_3 und B_2. Die Transformationsmatrix A von B_3 nach B und deren Inverse A^{-1} sind dort angegeben. Da im Zielraum $\mathbb{R}^2$ gar keine Transformation stattfindet entfällt der Faktor $B=I_2$. Somit gilt:

$$M'=M_{B_3,B_2}(T)=MA^{-1}=\begin{pmatrix}1 & -1 & 0\\ 0 & 1 & 1\end{pmatrix}\begin{pmatrix}-36 & 8 & 21\\ 5 & -1 & -3\\ 9 & -2 & -5\end{pmatrix}=\begin{pmatrix}-41 & 9 & 24\\ 14 & -3 & -8\end{pmatrix}$$

3.3.7 Das Normalformproblem für $T\in \mathrm{L}(V,W)$, $V\neq W$

Es seien V und W endlichdimensionale Vektorräume und $T\in \mathrm{L}(V,W)$. Wir kommen nun zurück zur Frage, wie man Basen in V und W wählen kann, dass die Matrix M' von T möglichst einfache Gestalt hat. Es sei $r=\operatorname{rg} T$. Wie im Beweis von Satz 3.7 wählen wir zunächst eine Basis $\{b_{r+1},b_{r+2},\dots,b_n\}$ von $\operatorname{Ker} T$. Ergänzt man diese Basis zu einer Basis $\{b_1,\dots,b_r,b_{r+1},\dots,b_n\}$ von V, so ist $\{T(b_1),\dots,T(b_r)\}$ eine Basis von $\operatorname{Im} T$. Ergänzt man wiederum diese r Vektoren beliebig zu einer Basis $C=\{T(b_1),\dots,T(b_r),w_{r+1},\dots,w_m\}$ von W, so hat $M'=M_{B,C}(T)$ die Gestalt

$$M'=\begin{pmatrix}I_r & 0\\ 0 & 0\end{pmatrix},$$

wobei I_r die Einheitsmatrix der Ordnung r ist. Dies folgt direkt aus der eingerahmten Formel in Definition 3.2.

Startet man mit einer Matrix $M\in\mathbb{K}^{m\times n}$ vom Rang $\operatorname{rg} M=r$, so gibt es nach den Überlegungen des vorigen Abschnitts reguläre Matrizen $A\in \mathrm{GL}(m,\mathbb{K})$ und $B\in\mathrm{GL}(n,\mathbb{K})$, sodass

$$M=AM'B=A\begin{pmatrix}I_r & 0\\ 0 & 0\end{pmatrix}B.$$

Wir nehmen M' als Normalform von M.

Wesentlich schwieriger ist das Normalformenproblem für $T\in\mathrm{L}(V)$, wenn wir nämlich nur *eine* Basis in V wählen können, die für den Ausgangs- und Zielraum V gleichzeitig gültig ist. Wir kommen darauf im Kapitel Eigenwerttheorie zurück.

Kapitel 4

Determinanten

4.1 Definition und einfache Eigenschaften

Determinanten bilden ein äußerst wichtiges Hilfsmittel in der linearen Algebra. Mit ihrer Hilfe kann man unmittelbar die Konsistenz von linearen $n \times n$-Gleichungssystemen erkennen, Volumina von Parallelepipeden ausrechnen, Geraden- oder Kreisgleichungen formulieren u.v.a.
Es gibt im Wesentlichen zwei Möglichkeiten, um Determinanten einzuführen, zum einen die Leibnizsche Definition, die einige zusätzliche Erläuterungen bedarf und zum zweiten eine axiomatische Charakterisierung, durch ihre wichtigsten Eigenschaften. Die letzte ist eleganter und beinhaltet gleichzeitig eine einfache Berechnungsmethode.
Wir benutzen zunächst die axiomatische Charakterisierung der Determinante. Für eine gegebene Matrix $A = (a_{ij}) \in \mathbb{K}^{n\times n}$ seien $a_i = \sum_{j=1}^{n} a_{ij} \mathrm{e}_j$, $i = 1, \dots, n$, die Zeilenvektoren der Matrix. Hier seien e_j die Zeileneinheitsvektoren im $\mathbb{K}^n$. Dann kann man die Matrix auch schreiben als

$$A = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}.$$

Definition 4.1 Eine Abbildung $\det\colon \mathbb{K}^{n\times n} \to \mathbb{K}$, $A \mapsto \det A$, heißt *Determinante*, falls folgendes gilt:

(D1) det ist linear in jeder Zeile $i = 1, \dots, n$, das heißt, wenn $a_i = \lambda b_i + \mu c_i$, $\lambda, \mu \in \mathbb{K}$, $b_i, c_i \in \mathbb{K}^n$, dann gilt

$$\det \begin{pmatrix} a_1 \\ \vdots \\ \lambda b_i + \mu c_i \\ \vdots \\ a_n \end{pmatrix} = \lambda \det \begin{pmatrix} a_1 \\ \vdots \\ b_i \\ \vdots \\ a_n \end{pmatrix} + \mu \det \begin{pmatrix} a_1 \\ \vdots \\ c_i \\ \vdots \\ a_n \end{pmatrix}.$$

Dabei seien bei den drei beteiligten Matrizen alle Matrixeinträge außerhalb der $i$ten Zeile jeweils gleich.

(D2) det ist *alternierend*, das heißt, vertauscht man zwei Zeilen a_i und a_j, so ändert sich das Vorzeichen:

$$\det\begin{pmatrix} a_1 \\ \vdots \\ a_i \\ \vdots \\ a_j \\ \vdots \\ a_n \end{pmatrix} = -\det\begin{pmatrix} a_1 \\ \vdots \\ a_j \\ \vdots \\ a_i \\ \vdots \\ a_n \end{pmatrix}.$$

(D3) det ist *normiert*, das heißt, $\det I_n = 1$.

Eine andere Schreibweise für $\det A$ ist

$$\det\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} = \begin{vmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix}.$$

Diese Definition der Determinante ist recht einfach. Es bleibt aber zu zeigen, dass eine solche Abbildung überhaupt *existiert*. Zunächst wollen wir einige Eigenschaften, u. a. die Eindeutigkeit, aus der Definition ableiten. Dazu setzen wir die Existenz voraus.

Satz 4.1 (Eigenschaften der Determinante) *Es sei* $\det\colon \mathbb{K}^{n\times n} \to \mathbb{K}$ *eine Determinantenabbildung, dann gilt für alle* $A \in \mathbb{K}^{n\times n}$ *und* $\lambda \in \mathbb{K}$*:*

(D4) $\det(\lambda A) = \lambda^n \det A$.

(D5) Besitzt A zwei identische Zeilen oder eine Nullzeile, so gilt $\det A = 0$.

(D6) Entsteht die Matrix A' aus A durch die dritte elementare Zeilenoperation, das heißt, man erhält A' aus A indem man das λfache der $i$ten Zeile von A wurde zur $j$ten Zeile von A addiert (wobei $i \neq j$), so gilt

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

(D7) Ist A eine obere Dreiecksmatrix, also $A = (a_{ij})$, $a_{ij} = 0$ für alle $i > j$,

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & a_{nn} \end{pmatrix},$$

so ist $\det A = a_{11} a_{22} \cdots a_{nn}$.

(D8) Es sei $n \geq 2$ *und* $A = \begin{pmatrix} A_1 & C \\ 0 & A_2 \end{pmatrix}$ *mit quadratischen Matrizen* A_1 *und* A_2*. Dann gilt*

$$\det A = \det A_1 \det A_2.$$

(D9) $\operatorname{rg} A < n$ *ist äquivalent zu* $\det A = 0$.

(D10) det *ist eindeutig bestimmt.*

Beweis. (D4) folgt unmittelbar aus (D1) indem man λ aus jeder Zeile als Faktor herauszieht.
(D5) Vertauscht man identische Zeilen, so ändert sich die Matrix nicht, wohl aber das Vorzeichen der Determinante wegen (D3); also gilt $\det A = -\det A$ und somit $\det A = 0$. Aus einer Nullzeile kann man durch Herausziehen des Faktors $\lambda = 0$ erhalten, dass $\det A = 0 \det A = 0$.
(D6) Sei $i < j$. Nach (D1) gilt

$$\det A' = \det \begin{pmatrix} a_1 \\ \vdots \\ a_i \\ \vdots \\ \lambda a_i + a_j \\ \vdots \\ a_n \end{pmatrix} = \lambda \det \begin{pmatrix} a_1 \\ \vdots \\ a_i \\ \vdots \\ a_i \\ \vdots \\ a_n \end{pmatrix} + \det \begin{pmatrix} a_1 \\ \vdots \\ a_i \\ \vdots \\ a_j \\ \vdots \\ a_n \end{pmatrix} = \det A.$$

Wir halten fest, dass durch die elementare Zeilenoperation (a) (Vertauschen zweier Zeilen) ein Faktor -1 hinzukommt, durch Multiplikation einer Zeile mit λ der Faktor λ hinzukommt und durch Zeilenoperation (c) sich die Determinante nicht ändert.
(D7) Ist ein Diagonalelement von A gleich Null, so kann der Gauß-Algorithmus keine n führenden Einsen mehr erzeugen (vgl. auch Beispiel 2.8 (d)), es entsteht mindestens eine Nullzeile, sodass nach (D5), $\det A = 0 = a_{11} \cdots a_{nn}$ gilt. Sind andererseits alle Diagonalelemente von A von Null verschieden, so kann man zunächst aus jeder Zeile i den Faktor a_{ii} wegen (D1) herausziehen und erhält

$$\det A = a_{11} \cdots a_{nn} \det \begin{pmatrix} 1 & a'_{12} & \cdots & a'_{1n} \\ 0 & 1 & \cdots & a'_{2n} \\ \vdots & & \vdots & \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

Die nun folgenden Jordan-Schritte benutzen nur noch die elementaren Zeilenumformungen vom Typ (c), ändern also die Determinante nicht. Als Endergebnis erhalten wir die Einheitsmatrix I_n, deren Determinante nach (D3) gleich 1 ist. Damit ist (D7) gezeigt.
(D8) Zunächst bringen wir A_1 durch die elementaren Zeilenoperationen (a) und (c) auf Dreiecksgestalt A'_1. Wir benutzten dabei *keine* Multiplikation von Zeilen. Angenommen, wir machten k Zeilenvertauschungen, so gilt $\det A'_1 = (-1)^k \det A_1$, nach (D3). Dabei ging C in eine Matrix C' über wogegen A_2 unverändert blieb. Ferner gilt nach (D1)

$$\det A = (-1)^k \det \begin{pmatrix} A'_1 & C' \\ 0 & A_2 \end{pmatrix}$$

Nun bringt man analog A_2 durch die elementaren Zeilentransformationen (a) und (c) (ohne (b)) auf Dreiecksgestalt A_2', etwa mit l Zeilenvertauschungen. Dabei ändern sich A_1' und C' nicht. Somit gilt einerseits $\det A_2 = (-1)^l \det A_2'$ und andererseits:

$$\det A = (-1)^{k+l} \det A' = (-1)^{k+l} \det \begin{pmatrix} A_1' & C' \\ 0 & A_2' \end{pmatrix}$$

Da aber A_1' und A_2' beides obere Dreiecksmatrizen sind und im linken unteren Block nur Nullen stehen, ist insgesamt A' eine obere Dreiecksmatrix. Nach (D7) gilt dann $\det A' = \det A_1' \det A_2'$, also

$$\det A = (-1)^k \det A_1' (-1)^l \det A_2' = \det A_1 \det A_2.$$

(D9) Ist $\operatorname{rg} A < n$, so liefert der Gauß-Algorithmus mindestens eine Nullzeile und damit ist nach (D4), $\det A = 0$.
Sei umgekehrt $\operatorname{rg} A = n$. Wendet man den Gauß-Jordan-Algorithmus auf A an, so wird in jedem Schritt die Determinante mit einer von Null verschiedenen Zahl multipliziert, bei (a) mit -1, bei (b) mit $\lambda \neq 0$ und bei (c) mit 1, und man endet mit $r = n$ führenden Nullen bzw. mit der Einheitsmatrix. Nach (D3) ist aber $\det I_n = 1$, also war $\det A \neq 0$.
(D10) Wir argumentieren ähnlich wie in (D9) und zeigen, dass sämtliche Werte $\det A$ festliegen. Angenommen, $\operatorname{rg} A < n$, dann ist nach (D9), $\det A = 0$. Ist hingegen $\operatorname{rg} A = n$, so liefert der Gauß-Jordan-Algorithmus nach endlich vielen elementaren Zeilenoperationen vom Typ (a), (b) und (c) die Einheitsmatrix. Dabei gilt für den Übergang von A nach A' mittels Zeilenoperation:

$$(a)\ \det A = -\det A', \quad (b)\ \det A = \lambda \det A', \quad (c)\ \det A = \det A'.$$

In jedem Fall erhalten wir $\det A$ als Produkt von Faktoren -1, $\lambda \neq 0$ und $+1$, wobei der letzte Faktor gleich $\det I_n = 1$ ist. Somit ist $\det A$ berechenbar und liegt fest. ■

Beispiel 4.1 (a)

$$\begin{vmatrix} 0 & 1 & 2 \\ 3 & 2 & 1 \\ 1 & 1 & 0 \end{vmatrix} \underset{(D2)}{=} - \begin{vmatrix} 1 & 1 & 0 \\ 3 & 2 & 1 \\ 0 & 1 & 2 \end{vmatrix} \underset{(D1)}{=} - \begin{vmatrix} 1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 1 & 2 \end{vmatrix} \underset{(D1)}{=} - \begin{vmatrix} 1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & 3 \end{vmatrix} \underset{(D7)}{=} 3.$$

(b)

$$\begin{vmatrix} 0 & b \\ c & d \end{vmatrix} = - \begin{vmatrix} c & d \\ 0 & b \end{vmatrix} = -bc.$$

Ist $a \neq 0$, so addiert man das $-\frac{c}{a}$fache der ersten Zeile zur zweiten Zeile und erhält dann mit (D7)

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = \begin{vmatrix} a & b \\ 0 & d - \frac{bc}{a} \end{vmatrix} = ad - bc.$$

Somit gilt allgemein

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

auch für $a=0$; wir leiteten die bereits aus Abschnitt 1.3 bekannte Formel ab.
(c) Eine Matrix $A \in \mathbb{K}^{n\times n}$ heißt *schiefsymmetrisch*, wenn $A^\top = -A$. Nach (b) gilt für die schiefsymmetrische Matrix $A = \begin{pmatrix} 0 & b \\ -b & 0 \end{pmatrix}$, $\det A = b^2$. Wir berechnen die Determinante einer schiefsymmetrischen 3×3-Matrix:

$$\begin{vmatrix} 0 & a & b \\ -a & 0 & c \\ -b & -c & 0 \end{vmatrix}$$

Addiert man nun zur Zeile 3 das $\frac{c}{a}$ fache der Zeile 1 und das $-\frac{b}{a}$ fache der Zeile 2, so ist die neue dritte Zeile eine Nullzeile und daher folgt aus (D5), dass $\det A = 0$.

4.2 Gruppen

Um einige Eigenschaften der Determinante beweisen zu können, benötigen wir den Begriff der Permutationsgruppe und des Charakters einer Permutation.

Definition 4.2 (a) Eine *Gruppe* ist eine Menge G zusammen mit einer Operation $G \times G \to G$, $(a,b) \mapsto ab$, genannt *Gruppenoperation* oder Gruppenmultiplikation, sodass die folgenden Bedingungen erfüllt sind:

(G1) Es gilt das Assoziativgesetz $a(bc) = (ab)c$ für alle $a,b,c \in G$.

(G2) Es gibt ein *neutrales Element* $e \in G$, sodass für alle $a \in G$ gilt $ae = ea = a$.

(G3) Zu jedem Element $a \in G$ gibt es ein *Inverses* Element a^{-1}, sodass gilt $aa^{-1} = a^{-1}a = e$.

Gilt außerdem das Kommutativgesetz $ab = ba$ für alle $a, b \in G$, so heißt G *abelsch*.
(b) Ein *Gruppenhomomorphismus* ist eine Abbildung $f\colon G \to H$ einer Gruppe G in eine Gruppe H, für die gilt

$$f(gg') = f(g)f(g') \quad \text{für alle } g, g' \in G. \tag{4.1}$$

Beispiel 4.2 (a) Jeder Körper $\mathbb{K}$ bildet bezüglich der Addition mit dem neutralen Element 0 eine abelsche Gruppe $(\mathbb{K},+,0)$. Ist $\mathbb{K}$ ein Körper, so ist $(\mathbb{K}\setminus\{0\},\cdot,1)$ eine abelsche Gruppe. Jeder Vektorraum V bildet bezüglich der Addition eine abelsche Gruppe $(V,+,0)$. Jede lineare Abbildung zwischen Vektorräumen ist ein Gruppenhomomorphismus.
(b) Es sei X eine beliebige nichtleere Menge. Dann ist

$$(X) = \{f\colon X \to X \mid f \text{ ist bijektiv}\}$$

bezüglich der Hintereinanderausführung von Abbildungen eine Gruppe, $((X),\circ,\mathrm{id}_X)$, wobei das neutrale Element die identische Abbildung id_X ist und die Umkehrabbildung f^{-1} die zu f inverse Abbildung ist.

Beweis. (1) Abgeschlossenheit. Es seien $g, f \in (X)$. Nach Übungsaufgabe 1.4 ist die Komposition $g \circ f$ wieder bijektiv; also $g \circ f \in (X)$.
(2) Die Komposition von Abbildungen ist assoziativ nach Lemma 1.2.
(3) Die Identität ist bijektiv, $\mathrm{id}_X \in (X)$ und es gilt $\mathrm{id}_X \circ f = f \circ \mathrm{id}_X = f$ für alle $f \in (X)$.
(4) Wenn $f \in (X)$, so existiert die inverse Abbildung f^{-1} und ist ebenfalls bijektiv; ferner gilt $f \circ f^{-1} = f^{-1} \circ f = \mathrm{id}_X$. Damit sind alle Gruppenaxiome erfüllt, $((X), \circ, \mathrm{id}_X)$ ist eine Gruppe. ∎

(c) Ist $\mathbb{K}$ ein Körper, so bildet $\mathrm{GL}(n, \mathbb{K})$, die Menge der invertierbaren Matrizen, eine Gruppe bezüglich der Matrixmultiplikation. Das neutrale Element ist die Einheitsmatrix I_n und A^{-1} ist das zu A inverse Element. Ist V ein Vektorraum, so ist die Menge der Automorphismen, $\mathrm{GL}(V)$ eine Gruppe bezüglich der Hintereinanderausführung. Es ist $\mathrm{GL}(V) \subset (V)$ eine Untergruppe.

4.2.1 Die symmetrische Gruppe

Permutationen

In diesem Abschnitt wollen wir das Beispiel 4.2 (b) mit einer endlichen Menge $X = \{1, 2, \ldots, n\}$ genauer studieren. In diesem Falle schreiben wir $(X) =: {}_n$ und nennen diese Gruppe die *symmetrische Gruppe* von n Elementen. Die Elemente der Menge sind die bijektiven Abbildungen $f : X \to X$. Diese Abbildungen heißen auch *Permutationen* von X. Wir wollen Permutationen als zweireihige Schemas aufschreiben, wobei in der ersten Zeile die Elemente von X (möglicherweise ungeordnet) stehen und darunter in der zweiten Zeile ihre Bilder:

$$f = \begin{pmatrix} 1 & 2 & 3 & \cdots & n \\ f(1) & f(2) & f(3) & \cdots & f(n) \end{pmatrix}.$$

Beispiel 4.3 Die 6 Elemente von S_3 sind

$$\mathrm{id} = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \qquad s_1 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, \qquad s_2 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix},$$
$$s_3 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}, \qquad z_1 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \qquad z_2 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix},$$

In diesem Falle ist $s_1 \circ s_1 = s_2 \circ s_2 = s_3 \circ s_3 = \mathrm{id}_X$ und ferner

$$s_1 \circ s_2 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} = z_1,$$

während $s_2 \circ s_1 = z_2$ gilt. Man beachte, dass wie immer bei Abbildungen zuerst die rechtsstehende und dann die links stehende Permutation ausgeführt wird.

Bemerkung 4.1 (a) Die Gruppe $_n$ ist für $n \geq 3$ nicht abelsch. Die Gruppen $_1$ und $_2$ bestehen jeweils nur aus einem bzw. aus zwei Elementen und sind beide abelsch.
(b) Durch vollständige Induktion beweist man, dass $|{}_n| = n!$.

Wir werden Permutationen oft mit kleinen griechischen Buchstaben bezeichnen. So ist zum Beispiel die zu $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 4 & 1 & 5 & 6 & 3 \end{pmatrix}$ inverse Permutation gleich

$$\sigma^{-1} = \begin{pmatrix} 2 & 4 & 1 & 5 & 6 & 3 \\ 1 & 2 & 3 & 4 & 5 & 6 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 1 & 6 & 2 & 4 & 5 \end{pmatrix}.$$

Der Charakter einer Permutation

Definition 4.3 Es sei $\pi \in {}_n$ eine Permutation.
(a) Als *Inversion* von π bezeichnen wir ein Paar (i,j), $i,j \in X$, $i<j$ mit $\pi(i) > \pi(j)$.
Wir bezeichnen mit $\operatorname{inv}(\pi)$ die Anzahl der Inversionen von π und nennen $\operatorname{sign}(\pi) = (-1)^{\operatorname{inv}(\pi)}$ den *Charakter* der Permutation π. Ist $\operatorname{sign}(\pi) = 1$, so heißt π *gerade*, ist hingegen $\operatorname{sign}(\pi) = -1$, so nennt man π *ungerade*.

So haben die Permutationen s_1 und s_2 aus Beispiel 4.3 jeweils eine Inversion, nämlich $(1,2)$ bzw. $(2,3)$, wogegen s_3 drei Inversionen $(1,2)$, $(1,3)$ und $(2,3)$ besitzt:

$$\operatorname{inv}(\mathrm{id}) = 0, \quad \operatorname{inv}(s_1) = \operatorname{inv}(s_2) = 1, \quad \operatorname{inv}(s_3) = 3, \quad \operatorname{inv}(z_1) = \operatorname{inv}(z_2) = 2.$$

Somit gilt für die Charaktere der Permutationen

$$\operatorname{sign}(\mathrm{id}) = \operatorname{sign}(z_1) = \operatorname{sign}(z_2) = 1, \quad \operatorname{sign}(s_1) = \operatorname{sign}(s_2) = \operatorname{sign}(s_3) = -1.$$

Lemma 4.2 *Es seien* $\pi, \sigma \in {}_n$.
(a) Dann gilt

$$\operatorname{sign} \pi = \prod_{1 \le i < j \le n} \frac{\pi(j) - \pi(i)}{j - i}.$$

(b) $\operatorname{sign}(\pi \circ \sigma) = \operatorname{sign} \pi \operatorname{sign} \sigma$. *Insbesondere gilt* $\operatorname{sign} \pi = \operatorname{sign} \pi^{-1}$.

Beweis. (a) Zunächst überlegt man sich, dass im Zähler und im Nenner des Bruchs jeweils sämtliche Differenzen zweier voneinander verschiedener Zahlen j und i vorkommen, eventuell aber mit unterschiedlichem Vorzeichen. Damit ist dass Produkt gleich 1 oder -1. Genau dann, wenn $\pi(j) < \pi(i)$, wenn also (i,j) eine Inversion ist, kommt ein negatives Vorzeichen hinzu. Die Gesamtzahl der negativen Vorzeichen ist also $\operatorname{inv}(\pi)$, sodass das Produkt gleich $\operatorname{sign} \pi$ ist.
(b) Nach (a) gilt

$$\operatorname{sign}(\pi\sigma) = \prod_{i<j} \frac{\pi(\sigma(j)) - \pi(\sigma(i))}{j-i} = \prod_{i<j} \frac{\pi(\sigma(j)) - \pi(\sigma(i))}{\sigma(j) - \sigma(i)} \prod_{j<i} \frac{\sigma(j) - \sigma(i)}{j-i}.$$

Der zweite Faktor ist nach (a) gleich $\operatorname{sign} \sigma$. Also genügt es zu zeigen, dass das erste Produkt gleich $\operatorname{sign} \pi$ ist. Nun gilt aber

$$\prod_{i<j} \frac{\pi(\sigma(j)) - \pi(\sigma(i))}{\sigma(j) - \sigma(i)} = \prod_{\substack{i<j \\ \sigma(i)<\sigma(j)}} \frac{\pi(\sigma(j)) - \pi(\sigma(i))}{\sigma(j) - \sigma(i)} \prod_{\substack{i<j \\ \sigma(i)>\sigma(j)}} \frac{\pi(\sigma(j)) - \pi(\sigma(i))}{\sigma(j) - \sigma(i)}$$

Vertauscht man im zweiten Faktor die Rolle von i und j, so ändert sich am Bruch $\frac{\pi(\sigma(j))-\pi(\sigma(i))}{\sigma(j)-\sigma(i)}$ und man kann fortsetzen mit

$$= \prod_{\substack{i<j\\ \sigma(i)<\sigma(j)}} \frac{\pi(\sigma(j))-\pi(\sigma(i))}{\sigma(j)-\sigma(i)} \prod_{\substack{i>j\\ \sigma(i)<\sigma(j)}} \frac{\pi(\sigma(j))-\pi(\sigma(i))}{\sigma(j)-\sigma(i)}.$$

Nun kann man beide Produkte wieder zu einem Produkt zusammenfassen:

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

Da σ bijektiv ist, durchlaufen $r=\sigma(i)$ und $s=\sigma(j)$ alle Zahlen von 1 bis n, also

$$= \prod_{r<s} \frac{\pi(s)-\pi(r)}{s-r} = \operatorname{sign}\pi.$$

Damit ist die Behauptung gezeigt. Wegen $\operatorname{sign}\mathrm{id} = 1$ folgt $\operatorname{sign}\sigma \operatorname{sign}\sigma^{-1} = \operatorname{sign}\mathrm{id} = 1$ und damit folgt die zweite Behauptung $\operatorname{sign}\sigma = \operatorname{sign}\sigma^{-1}$. ■

Bemerkung 4.2 (a) Ist $H=\{-1,1\}$ die Gruppe aus 2 Elementen bezüglich der gewöhnlichen Multiplikation, so ist $\operatorname{sign}\colon {}_n \to H$ nach (b) ein Gruppenhomomorphismus.
(b) Multipliziert man nacheinander die Menge der geraden Permutationen $\mathbf{A}_n$ mit der Permutation $\tau = \begin{pmatrix} 1 & 2 & 3 & \cdots & n \\ 2 & 1 & 3 & \cdots & n \end{pmatrix}$, so entstehen nach obigem Lemma nur ungerade Permutationen, denn $\operatorname{sign}\tau = -1$ (τ hat genau eine Inversion). Umgekehrt werden dadurch gerade Permutationen in ungerade Permutationen überführt.
Die Abbildung $f\colon SS_n \to {}_n$, $f(\pi) = \tau\pi$ ist bijektiv mit $f^{-1} = f$, denn $\tau^2 0 \mathrm{id}_X$ und vertauscht die Menge der geraden mit der Menge der uungeraden Permutationen. Es gibt also genauso viele gerade wie ungerade Permutationen. Also gilt $|\mathbf{A}_n| = |{}_n \setminus \mathbf{A}_n| = \frac{1}{2}|{}_n| = \frac{1}{2}n!$.
(c) Um den Charakter einer Permutation schnell zu bestimmen, ist die Zyklenschreinbweise günstig:

$$\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 4 & 1 & 3 & 5 & 6 & 2 & 10 & 9 & 8 & 7 \end{pmatrix} = (1\,4\,5\,6\,2)(3)(7\,10)(8\,9).$$

Zyklen ungerader Länge sind gerade, Zyklen gerader Länge sind ungerade. Also gilt hier $\operatorname{sign}(\pi) = 1 \cdots (-1)(-1) = 1$; π ist gerade.

4.3 Rechnen mit Determinanten

4.3.1 Leibnizdefinition der Determinante

Zunächst wollen wir (D1) simultan auf mehrere Zeilen anwenden.

Lemma 4.3 *Es sei* $\det\colon \mathbb{K}^{n\times n} \to \mathbb{K}$ *eine Determinantenfunktion. Dann gilt für alle* $A = (a_{ij}) \in \mathbb{K}^{n\times n}$

$$\det A = \sum_{j_1, j_2, \cdots, j_n = 1}^{n} a_{1j_1} a_{2j_2} \cdots a_{nj_n} \det \begin{pmatrix} \mathrm{e}_{j_1} \\ \mathrm{e}_{j_2} \\ \vdots \\ \mathrm{e}_{j_n} \end{pmatrix}. \tag{4.2}$$

Dabei haben wir n Summen, die jeweils von $1, \ldots, n$ *laufen, also* n^n *Summanden.*

Beweis. Wir fixieren die Zeilen 2 bis n und schreiben die erste Zeile mit Hilfe der Basiszeilenvektoren $\mathrm{e}_{j_1} \in \mathbb{K}^{1\times n}$ als $a_1 = \sum_{j_1=1}^{n} a_{1j_1} \mathrm{e}_{j_1}$. Wegen (D1) gilt dann

$$\det A = \det \begin{pmatrix} \sum_{j_1=1}^{n} a_{1j_1} \mathrm{e}_{j_1} \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = \sum_{j_1=1}^{n} a_{1j_1} \det \begin{pmatrix} \mathrm{e}_{j_1} \\ a_2 \\ \vdots \\ a_n \end{pmatrix}.$$

Macht man dasselbe mit der 2.Zeile $a_2 = \sum_{j_2=1}^{n} a_{2j_2} \mathrm{e}_{j_2}$, so erhält man

$$\det A = \sum_{j_1=1}^{n} \sum_{j_2=1}^{n} a_{1j_1} a_{2j_2} \det \begin{pmatrix} \mathrm{e}_{j_1} \\ \mathrm{e}_{j_2} \\ a_3 \\ \vdots \\ a_n \end{pmatrix}.$$

Setzt man dies fort bis zur $n$ten Zeile, so erhält man die Behauptung. ∎

Zum Beispiel haben wir

$$\begin{aligned} \begin{vmatrix} a & b \\ c & d \end{vmatrix} &= a \begin{vmatrix} 1 & 0 \\ c & d \end{vmatrix} + b \begin{vmatrix} 0 & 1 \\ c & d \end{vmatrix} = a \left(c \begin{vmatrix} 1 & 0 \\ 1 & 0 \end{vmatrix} + d \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} \right) + b \left(c \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix} + d \begin{vmatrix} 0 & 1 \\ 0 & 1 \end{vmatrix} \right) \\ &= ac \begin{vmatrix} 1 & 0 \\ 1 & 0 \end{vmatrix} + ad \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} + bc \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix} + bd \begin{vmatrix} 0 & 1 \\ 0 & 1 \end{vmatrix}. \end{aligned}$$

Wir haben im vorigen Abschnitt gesehen, wie man die Berechnung jeder Determinante zurückführen kann auf die Berechnung der n^n Determinanten

$$\det \begin{pmatrix} \mathrm{e}_{j_1} \\ \mathrm{e}_{j_2} \\ \vdots \\ \mathrm{e}_{j_n} \end{pmatrix}, \quad j_1, \ldots, j_n = 1, \ldots, n, \tag{4.3}$$

die nur aus Nullen und n Einsen bestehen. Nach Determinateneigenschaft (D4) sind aber alle diejenigen Determinanten gleich Null, die mindestens zwei gleiche Zeilen haben, wo also der Rang der entsprechenden Matrix kleiner als n ist. Die Determinanten (4.3) sind daher von Null verschieden, wenn sie vollen Rang n haben. Das ist genau dann der Fall, wenn in jeder Spalte der Matrix (4.3) genau eine Eins steht bzw., wenn alle Indizes $j_1, j_2, \dots, j_n$ paarweise voneinander verschieden sind. Das n-Tupel $\sigma = (j_1, \dots, j_n)$ stellt also eine Permutation von $\{1, \dots, n\}$ dar. Somit vereinfacht sich die Formel (4.2) zu

$$\det A = \sum_{\sigma\in{}_n} a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)} \det \begin{pmatrix} e_{\sigma(1)} \\ e_{\sigma(2)} \\ \vdots \\ e_{\sigma(n)} \end{pmatrix}. \tag{4.4}$$

Wir kehren nun zum Gauß-Algorithmus zurück, um die Determinante der *Permutationsmatrix* $P_\sigma = \begin{pmatrix} e_{\sigma(1)} \\ e_{\sigma(2)} \\ \vdots \\ e_{\sigma(n)} \end{pmatrix}$ zu berechnen. Dabei gehen wir wiefolgt vor: Wir vertauschen zwei benachbarte Zeilen, wenn in der oberen Zeile i die Eins rechts von der Eins in der darunter liegenden Zeile $i+1$ steht, wenn also $\sigma(i) > \sigma(i+1)$. In diesem Falle ist $(i, i+1)$ eine Inversion der Permutation σ. Durch den Tausch geht die Permutation σ in eine neue Permutation σ' über, die *genau eine* Inversion weniger hat als σ, $\operatorname{inv}\sigma' = \operatorname{inv}\sigma - 1$. In der Tat, ist (k, i) eine Inversion von σ, so ist $(k, i+1)$ eine Inversion von σ'; war $(k, i+1)$ eine Inversion von σ, so ist nach dem Tausch (k, i) eine Inversion von σ'. Waren schließlich (i, j) bzw. $(i+1, j)$ Inversionen von σ, so sind $(i+1, j)$ bzw. (i, j) Inversionen von σ'. Weitere Inversionen kann σ' nicht haben. Nach genau $\operatorname{inv}\sigma$ Nachbarzeilenvertauschungen, haben wir also die Einheitsmatrix aus P_σ hergestellt. Somit gilt nach (D2) und (D3):

$$\det \begin{pmatrix} e_{\sigma(1)} \\ e_{\sigma(2)} \\ \vdots \\ e_{\sigma(n)} \end{pmatrix} = (-1)^{\operatorname{inv}\sigma} \det I_n = \operatorname{sign}\sigma.$$

Satz 4.4 (Leibnizdefinition) *Für jedes* $n \in \mathbb{N}$ *gibt es genau eine Determinantenfunktion* $\det\colon \mathbb{K}^{n\times n} \to \mathbb{K}$. *Sie ist gegeben durch*

$$\det A = \sum_{\sigma\in{}_n} \operatorname{sign}\sigma\ a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)}. \tag{4.5}$$

Beweis. Wir haben oben gesehen: Wenn es überhaupt eine Determinantenfunktion gibt, so muss sie (4.5) erfüllen. Wir müssen also nur noch zeigen, dass die durch (4.5) gegebene Funktion tatsächlich die Bedingungen (D1), (D2) und (D3) einer Determinante erfüllt.
In der Tat ist (D1) erfüllt, denn ersetzt man bei fixiertem Zeilenindex i die Matrixelemente $a_{i\sigma(i)}$ durch $\lambda b_{i\sigma(i)} + \mu c_{i\sigma(i)}$, so ist

$$\sum_{\sigma\in{}_n} \operatorname{sign}\sigma\, a_{1\sigma(1)} \cdots (\lambda b_{i\sigma(i)} + \mu c_{i\sigma(i)}) \cdots a_{n\sigma(n)} = \lambda \sum_{\sigma\in{}_n} \operatorname{sign}\sigma\, a_{1\sigma(1)} \cdots b_{i\sigma(i)} \cdots a_{n\sigma(n)} +$$
$$\mu \sum_{\sigma\in{}_n} \operatorname{sign}\sigma\, a_{1\sigma(1)} \cdots c_{i\sigma(i)} \cdots a_{n\sigma(n)} = \lambda \det A' + \mu \det A''.$$

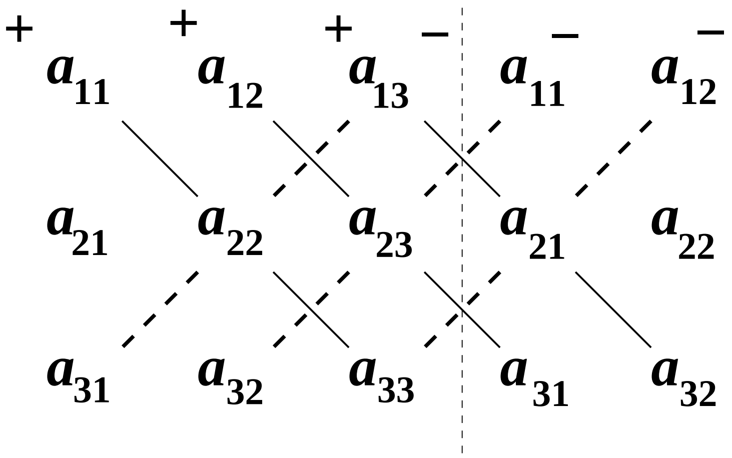
Die Vertauschung zweier Zeilen i und j, $i \neq j$, kann man durch die Multiplikation von σ mit der Permutation $\tau = \begin{pmatrix} 1 & \cdots & i & \cdots & j & \cdots & n \\ 1 & \cdots & j & \cdots & i & \cdots & n \end{pmatrix}$ korrigieren. Man beachte, dass $\operatorname{sign}\tau = -1$ eine ungerade Permutation ist. Es sei A' die Matrix mit den vertauschten Zeilen a_i und a_j, wobei $i < j$. Dann ist

$$\begin{aligned} \det A' &= \sum_{\sigma \in {}_n} \operatorname{sign}\,\sigma\, a_{1\sigma(1)} \dots a_{j\sigma(i)} \dots a_{i\sigma(j)} \dots a_{n\sigma(n)} \\ &= \sum_{\sigma \in {}_n} \operatorname{sign}\sigma\ a_{1,(\sigma\tau)(1)} \dots a_{j,(\sigma\tau)(j)} \dots a_{i,(\sigma\tau)(i)} \dots a_{n,(\sigma\tau)(n)} \\ &= \sum_{\sigma\tau \in {}_n} (-1)\operatorname{sign}(\sigma\tau)\,\sigma\, a_{1,\sigma\tau(1)} \dots\dots a_{n,\sigma\tau(n)} = -\det A. \end{aligned}$$

Schließlich ist klar, dass die Leibnizdefinition die Normierungsbedingung erfüllt, denn in (4.5) bleibt ein einziger Summand übrig, nämlich der zur identischen Permutation: $\det I_n \underset{\sigma=\mathrm{id}}{=} \operatorname{sign}(\mathrm{id})\, a_{11} \cdots a_{nn} = 1$. ∎

Beispiel 4.4 Mit Hilfe der Leibnizdefinition berechnen wir $\det A$, $A = (a_{ij}) \in \mathbb{K}^{3\times 3}$. Die symmetrische Gruppe $_3$ besitzt drei gerade Permutationen id, z_1 und z_2 und drei ungerade Permutationen s_1, s_2 und s_3, also gilt

$$\det A = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32}.$$



Dasselbe Ergebnis erhält man mit der Sarrusschen Regel: Achtung, sie gilt nur für 3×3-Matrizen. Man schreibt zunächst die beiden ersten Spalten als vierte und fünfte Spalte hinten an und berechnet die sechs Produkte auf den entstehenden Haupt- und Nebendiagonalen.

Lemma 4.5 *Für eine quadratische Matrix A gilt* $\det A = \det A^\top$.

Beweis. Es sei $A = (a_{ij}) \in \mathbb{K}^{n\times n}$. Dann ist $A^\top = (a_{ji})$ und

$$\begin{aligned} \det A^\top &= \sum_{\sigma \in {}_n} \operatorname{sign}\sigma\ a_{\sigma(1),1} \dots a_{\sigma(n),n} \\ &= \sum_{\sigma \in {}_n} \operatorname{sign}\sigma\ a_{1,\sigma^{-1}(1)} \dots a_{n,\sigma^{-1}(n)} \\ &= \sum_{\sigma^{-1} \in {}_n} \operatorname{sign}\sigma^{-1}\ a_{1,\sigma^{-1}(1)} \dots a_{n,\sigma^{-1}(n)} = \det A. \end{aligned}$$

Hierbei benutzen wir Lemma 4.2 (b) und den Fakt, dass mit σ auch σ^{-1} die gesamte $_n$ durchläuft.

■

Folgerung 4.6 *Für die Determinantenfunktion* $\det\colon \mathbb{K}^{n\times n} \to \mathbb{K}^n$ *gilt*

(D1') $\det$ *ist linear in jeder Spalte* j, $j = 1,\ldots,n$, *das heißt, ist* $\alpha_j = \lambda\beta_j + \mu\gamma_j$, $\beta_j, \gamma_j \in \mathbb{K}^{n\times 1}$, *so ist*

$$\det(\alpha_1,\ldots,\lambda\beta_j + \mu\gamma_j,\ldots\alpha_n) = \lambda\det(\alpha_1,\ldots,\beta_j,\ldots,\alpha_n) + \mu\det(\alpha_1,\ldots,\gamma_j,\ldots,\alpha_n).$$

(D2') Vertauscht man in A *zwei Spalten* α_i *und* α_j*, so ändert sich bei der Determinante das Vorzeichen.*

(D5') Besitzt A *zwei identische Spalten oder eine Nullspalte, so ist* $\det A = 0$.

(D6') Entsteht die Matrix A' *aus* A *indem man zur* j *ten Spalte von* A *das* λ*fache der* i *ten Spalte addiert und alle anderen Spalten unverändert lässt, so gilt* $\det A' = \det A$.

(D7') Für eine untere Dreiecksmatrix

$$A = \begin{pmatrix} a_{11} & 0 & 0 & \cdots & 0 \\ a_{21} & a_{22} & 0 & \cdots & 0 \\ \vdots & & & & \vdots \\ a_{n1} & a_{n2} & \cdots & & a_{nn} \end{pmatrix}$$

gilt $\det A = a_{11}a_{22}\cdots a_{nn}$.

Beweis. Die Eigenschaften folgen alle aus den entsprechenden „Zeileneigenschaften", da beim Transponieren Zeilen in Spalten über gehen und $\det A = \det A^\top$. ■

Bemerkung 4.3 Analog zum Gauß-Jordan-Algorithmus kann man daher die Determinante auch durch die drei elementaren *Spaltenoperationen* berechnen: (a) Multiplikation einer Spalte mit einer von Null verschiedenen Zahl λ , (b) Vertauschung zweier Spalten und (c) Addition des μ-fachen einer Spalte zu einer anderen Spalte. Dabei muss man bei (a) und (b) den Faktor λ bzw. (-1) beachten.
Natürlich kann man auch beliebig elementare Zeilen- und Spaltenoperationen miteinander koppeln.

4.3.2 Entwicklungssätze und inverse Matrix

Wir geben hier eine dritte praktische Möglichkeit an, Determinanten zu berechnen.
Wir beginnen mit einer einfachen Überlegung, die aus (D8) und (D8') folgt. Dazu sei $A \in \mathbb{K}^{n\times n}$. Dann gilt

$$\det\begin{pmatrix} 1 & * \\ 0 & A \end{pmatrix} = \det\begin{pmatrix} 1 & 0 \\ 0 & A \end{pmatrix} = \det\begin{pmatrix} 1 & 0 \\ * & A \end{pmatrix} = \det A. \tag{4.6}$$

Dabei steht $*$ für einen beliebigen Zeilen- bzw. Spaltenvektor.
Beispiel. Wir vereinfachen die Determinante der Matrix $X \in \mathbb{R}^{4\times 4}$. Addiert man Vielfache der 2. Spalte zu den anderen Spalten, so erhält man:

$$\det X = \begin{vmatrix} a_1 & 0 & c_{11} & c_{12} \\ a_2 & 0 & c_{21} & c_{22} \\ 2 & 5 & 3 & 4 \\ b & 0 & d_1 & d_2 \end{vmatrix} = \begin{vmatrix} a_1 & 0 & c_{11} & c_{12} \\ a_2 & 0 & c_{21} & c_{22} \\ \mathbf{0} & 5 & \mathbf{0} & \mathbf{0} \\ b & 0 & d_1 & d_2 \end{vmatrix} = 5\begin{vmatrix} a_1 & 0 & c_{11} & c_{12} \\ a_2 & 0 & c_{21} & c_{22} \\ 0 & \mathbf{1} & 0 & 0 \\ b & 0 & d_1 & d_2 \end{vmatrix}$$

Vertauscht man jetzt die erste und zweite Spalte und dann die 2. und 3. Zeile und 1. und 2. Zeile, so hat man insgesamt 3 Vertauschungen vorgenommen, so dass die Determinante ihr Vorzeichen wechselt:

$$\det X = -5\begin{vmatrix} 0 & a_1 & c_{11} & c_{12} \\ 0 & a_2 & c_{21} & c_{22} \\ 1 & 0 & 0 & 0 \\ 0 & b & d_1 & d_2 \end{vmatrix} = -5\begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & a_1 & c_{11} & c_{12} \\ 0 & a_2 & c_{21} & c_{22} \\ 0 & b & d_1 & d_2 \end{vmatrix}$$

Wendet man nun (D7) an, so hat man

$$\det X = 5\begin{vmatrix} a_1 & c_{11} & c_{12} \\ a_2 & c_{21} & c_{22} \\ b & d_1 & d_2 \end{vmatrix}.$$

Die Anzahl der Zeilenvertauschungen ist $i-1$, die Anzahl der Spaltenvertauschungen ist $j-1$, wobei $(i,j)=(3,2)$ die Position des Eintrages $x_{32}=5$ ist. Daher erhält man $(-1)^{i+j-2}=(-1)^{i+j}$ als Vorzeichen.

Satz 4.7 (Zeilenentwicklungssatz) *Es sei $A=(a_{ij})\in\mathbb{K}^{n\times n}$ und $A_{ij}\in\mathbb{K}^{(n-1)\times(n-1)}$ bezeichne die Matrix, die aus A dadurch entsteht, dass man die i te Zeile und die j te Spalte streicht. Dann gilt für alle $i=1,\dots,n$:*

$$\det A = \sum_{j=1}^{n}(-1)^{i+j} a_{ij}\det A_{ij}. \tag{4.7}$$

Man nennt (4.7) die *Entwicklung der Determinante nach der i ten Zeile.*
Beweis. Für alle i,j bezeichne A'_{ij} die $n\times n$- Matrix, die aus A entsteht, indem man alle Einträge der i ten Zeile und alle Einträge der j ten Spalte Null setzt bis auf das Element $a'_{ij}=1$,

das Eins gesetzt wird. Wegen der Linearität von $\det A$ in der $i$ten Zeile folgt wie im Beweis von (4.2)

$$\det A = \sum_{j=1}^{n} a_{ij} \det A'_{ij},$$

wobei wir die Elemente der $j$ten Spalte zu Null gemacht haben (bis auf $a_{ij} = 1$) mit Hilfe der dritten elementaren Spaltenoperation. Nun tauschen wir die $i$te Zeile von A'_{ij} schrittweise nach oben in die erste Zeile und die $j$te Spalte schrittweise nach vorn in die erste Spalte, wobei wir die Reihenfolge der jeweils anderen Zeilen und Spalten beibehalten. Dazu sind genau $i-1$ Nachbarzeilenvertauschungen und $j-1$ Nachbarspaltenvertauschungen nötig, sodass

$$\det A'_{ij} = (-1)^{i-1+j-1} \det \begin{pmatrix} 1 & 0 \\ 0 & A_{ij} \end{pmatrix} = (-1)^{i+j} \det A_{ij}.$$

Setzt man dies oben ein, so folgt die Behauptung. ■

Folgerung 4.8 (Spaltenentwicklungssatz) *Mit den Bezeichnungen des Zeilenentwicklungssatzes gilt für alle $j = 1, \dots, n$:*

$$\det A = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det A_{ij}. \tag{4.8}$$

Man nennt (4.8) die *Entwicklung der Determinante nach der $j$ten Spalte*. Der Beweis folgt sofort aus dem Zeilenentwicklungssatz, zusammen mit $\det A = \det A^\top$.
Man beachte bei den Entwicklungssätzen die schachbrettartige Anordnung der Vorzeichen:

$$\begin{pmatrix} + & - & + \\ - & + & - \\ + & - & + \end{pmatrix}$$

Beispiel 4.5 (a) Die Entwicklung von $A = (a_{ij}) \in \mathbb{K}^{3\times 3}$ nach der 2. Zeile liefert

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

(b) Entwicklung nach der 2. Spalte und dann nach der 3. Zeile liefert:

$$\begin{vmatrix} 1 & 5 & 7 & 1 \\ 0 & 0 & 2 & 3 \\ 8 & 0 & -1 & 2 \\ 9 & 0 & 0 & 0 \end{vmatrix} = (-5) \begin{vmatrix} 0 & 2 & 3 \\ 8 & -1 & 2 \\ 9 & 0 & 0 \end{vmatrix} = (-5)9 \begin{vmatrix} 2 & 3 \\ -1 & 2 \end{vmatrix} = -45 \cdot 7 = -315.$$

Satz 4.9 (Inverse Matrix) *Es sei $A = (a_{ij}) \in \mathrm{GL}(n, \mathbb{K})$ eine invertierbare Matrix, und es sei $d_{ji} = (-1)^{i+j} \det A_{ij}$, $i, j = 1, \dots, n$.*
Dann gilt

$$A^{-1} = \frac{1}{\det A} (d_{ij})_{i,j=1\dots n}. \tag{4.9}$$

Beweis. Wir müssen zeigen, dass $(A\,B)_{ij} = \delta_{ij}$ für alle $i, j = 1, \dots, n$, dabei ist B die Matrix auf der rechten Seite von (4.9). Für $i = j$ ist in der Tat nach dem Zeilenentwicklungssatz

$$(A\,B)_{ii} = \frac{1}{\det A} \sum_{k=1}^{n} a_{ik} (-1)^{k+i} \det A_{ik} = \frac{\det A}{\det A} = 1.$$

Für $i \neq j$ haben wir

$$(A\,B)_{ij} = \frac{1}{\det A} \sum_{k=1}^{n} (-1)^{k+j} a_{ik} \det A_{jk}. \tag{4.10}$$

Ersetzt man in der Matrix A die $j$te Zeile durch die $i$te Zeile, so erhält man eine Matrix A' mit $\det A' = 0$ wegen (D5). Entwickelt man nun A' nach der $j$ten Zeile, so erhält man die Unterdeterminanten $\det A_{jk}$ mit den Koeffizienten $(-1)^{k+j} a_{ik}$, also genau die Formel der rechten Seite von (4.10). Folglich gilt

$$(A\,B)_{ij} = \frac{1}{\det A} \det A' = 0.$$

■

Beispiel 4.6 (a) Für eine reguläre Matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{GL}(2, \mathbb{K})$ gilt

$$A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

(b) Für eine reguläre Matrix $A = (a_{ij}) \in \mathrm{GL}(3, \mathbb{K})$ gilt

$$(A^{-1})_{13} = \frac{1}{\det A} (-1)^{1+3} \det A_{31} = \frac{1}{\det A} \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix}.$$

4.3.3 Multiplikationssatz und die Determinante und Spur von $T \in \mathrm{L}(V)$

Satz 4.10 (Multiplikationssatz) *Es seien $A, B \in \mathbb{K}^{n\times n}$. Dann gilt* $\det(AB) = \det A \det B$.

Beweis. Wir benutzen die Leibnizdefinition für $C = AB$, $c_{ik} = \sum_{j=1}^{n} a_{ij} b_{jk}$:

$$\begin{aligned} \det C &= \sum_{\sigma \in_n} \operatorname{sign} \sigma \sum_{j_1=1}^{n} a_{1j_1} b_{j_1\sigma(1)} \sum_{j_2=1}^{n} a_{2j_2} b_{j_2\sigma(2)} \cdots \sum_{j_n=1}^{n} a_{nj_n} b_{j_n\sigma(n)} \\ &= \sum_{j_1, j_2, \dots, j_n = 1}^{n} a_{1j_1} \cdots a_{nj_n} \sum_{\sigma \in_n} \operatorname{sign} \sigma\, b_{j(1)\sigma(1)} b_{j(2)\sigma(2)} \cdots b_{j(n)\sigma(n)} \end{aligned}$$

Ist nun $j = (j_1, \dots, j_n)$ keine Permutation von $\{1, \dots, n\}$, so ist der entsprechende Summand $\sum_{\sigma \in_n} \operatorname{sign} \sigma\, b_{j(1)\sigma(1)} b_{j(2)\sigma(2)} \cdots b_{j(n)\sigma(n)}$ gleich Null, da die Matrix zwei identische Zeilen hat.

Wir können uns also auf Permutationen $j = (j(1),\dots,j(n)) \in {}_n$ beschränken. Durch Umordnen der Faktoren $b_{j(1)\sigma(1)}b_{j(2)\sigma(2)}\cdots b_{j(n)\sigma(n)}$ erhalten wir $b_{1,j^{-1}\sigma(1)}b_{2,j^{-1}\sigma(2)}\cdots b_{n,j^{-1}\sigma(n)}$. Wegen $\operatorname{sign} j^{-1} = \operatorname{sign} j$ gilt:

$$\begin{aligned}\det C &= \sum_{j\in_n} a_{1j_1}\cdots a_{nj_n} \operatorname{sign} j \sum_{\sigma\in_n} \operatorname{sign} j^{-1} \operatorname{sign}\sigma\, b_{1,j^{-1}\sigma(1)}b_{2,j^{-1}\sigma(2)}\cdots b_{n,j^{-1}\sigma(n)}\\ &= \sum_{j\in_n} \operatorname{sign}(j) a_{1j_1}\cdots a_{nj_n} \det B = \det A \det B.\end{aligned}$$

■

Folgerung 4.11 *(a) Es sei $A \in \mathrm{GL}(n,\mathbb{K})$ eine invertierbare Matrix. Dann gilt* $\det A^{-1} = \frac{1}{\det A}$.
(b) Für alle Matrizen $M \in \mathbb{K}^{n\times n}$ mit $M' = A^{-1}MA$ gilt:

$$\det M' = \det(A^{-1}MA) = \det M.$$

Matrizen M und M', für die es eine Matrix $A \in \mathrm{GL}(n,\mathbb{K})$ gibt, so dass $M' = A^{-1}MA$, heißen einander ähnlich. *M. a. W., ähnliche Matrizen haben dieselbe Determinante.*

Beweis. (a) Nach dem Multiplikationssatz und wegen $\det I_n = 1$ gilt $1 = \det I_n = \det(AA^{-1}) = \det A \det A^{-1}$. Hieraus folgt (a).
(b) Nach dem Multiplikationssatz und (a) gilt $\det AMA^{-1} = \det A \det M (\det A)^{-1} = \det M$.
■

Bemerkung 4.4 Nach Abschnitt 3.3.6 sind die Matrizen $M = M_{E,E}(T)$ und $M' = M_{E',E'}(T)$ einander ähnlich, denn $M' = A^{-1}MA$, wobei A die Matrix der Basistransformation von E nach E' ist.

Definition 4.4 Es sei $T \in \mathrm{L}(V)$ eine lineare Abbildung von V in sich. Ferner sei B eine Basis in V, $\dim V = n < \infty$.
(a) Dann bezeichnet man als *Determinante von T* die Zahl

$$\mathbf{det}\, \boldsymbol{T} = \mathbf{det}\, \boldsymbol{M_{B,B}(T)}.$$

(b) Für eine quadratische Matrix $A = (a_{ij}) \in \mathbb{K}^{n\times n}$ definieren wir die *Spur von A* als

$$\mathbf{tr}\, \boldsymbol{A} = \sum_{\boldsymbol{k=1}}^{\boldsymbol{n}} \boldsymbol{a_{kk}}$$

; die Spur von A ist also die Summe der Hauptdiagonalelemente von A.
(c) Wir definieren die *Spur von T* über die Spur einer Darstellungsmatrix $M_{B,B}(T)$:

$$\mathbf{tr}\, \boldsymbol{T} = \mathbf{tr}\, \boldsymbol{M_{B,B}(T)}.$$

Bemerkung 4.5 (a) Die Definition der Determinante hängt wegen der Transformationsformel für Endomorphismen, $M' = AMA^{-1}$ und wegen Folgerung 4.11 (b) nicht von der speziellen Wahl der Basis B ab.
(b) Für alle $A, B \in \mathbb{K}^{n\times n}$ gilt:

$$\operatorname{tr} AB = \operatorname{tr} BA.$$

In der Tat ist

$$\operatorname{tr} AB = \sum_{k=1}^{n} (AB)_{kk} = \sum_{k=1}^{n}\sum_{j=1}^{n} a_{kj} b_{jk} = \sum_{j=1}^{n}\sum_{k=1}^{n} b_{jk} a_{kj} = \sum_{j=1}^{n} (BA)_{jj} = \operatorname{tr} BA.$$

Hieraus folgt, dass einander ähnliche Matrizen M und M' die gleiche Spur haben. Es ist nämlich

$$\operatorname{tr} M' = \operatorname{tr}(A^{-1}{\cdot}MA) = \operatorname{tr}(MA{\cdot}A^{-1}) = \operatorname{tr}(MI_n) = \operatorname{tr} M.$$

Daher hängt $\operatorname{tr} T$ nicht von der speziellen Wahl der Basis ab.

Beispiel 4.7 (Fortsetzung von Beispiel 3.1 und 3.3) (a) Sei $\dim V = n$. Dann gilt $\det(\mathrm{id}_V) = 1$ und $\operatorname{tr}(\mathrm{id}_V) = n$.
(b) **Verschiebungsoperators** $V_a \in \mathrm{L}(\mathbb{R}_2[x])$. $\det V_a = 1$ und $\operatorname{tr} V_a = 3$.
(c) **Differentiation** $D\colon V \to V$ auf den trigonometrischen Polynomen höchstens zweiten Grades (siehe Übungsaufgabe 6.3): $\det D = \operatorname{tr} D = 0$.
(d) **Symmetriesator**. Es sei $S \in \mathrm{L}(\mathbb{R}^{2\times 2})$ die lineare Abbildung $S(A) = \frac{1}{2}(A + A^\top)$, $A \in \mathbb{R}^{2\times 2}$ die jeder Matrix ihren symmetrischen Teil zuordnet. Wir bestimmen Determinanten, Spur, Rang und Defekt von S. Dazu zeichnen wir die Basis von Matrixeinsen in $\mathbb{R}^{2\times 2}$ aus, $B = \{E_{11}, E_{12}, E_{21}, E_{22}\}$. Wegen $S(E_{11}) = E_{11}$, $S(E_{22}) = E_{22}$ und $S(E_{12}) = S(E_{21}) = \frac{1}{2}(E_{12} + E_{21})$, gilt

$$M = M_{B,B}(S) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Folglich gilt $\det S = 0$, $\operatorname{tr} S = 3 = \operatorname{rg} S$, $\operatorname{def} S = 1$.

4.4 Einfache Anwendungen der Determinante

4.4.1 Cramersche Regel

Wie im Abschnitt über Entwicklungssätze bezeichnen wir die Spalten einer Matrix $A = (a_{ij}) \in \mathbb{K}^{n\times n}$ mit $\alpha_j = \sum_{i=1}^{n} a_{ij} \mathrm{e}_i$, wobei hier $\mathrm{e}_i \in \mathbb{K}^{n\times 1}$ den $i$ten Spalteneinheitsvektor bezeichnet.

Satz 4.12 (Cramersche Regel – Gabriel Cramer, 1704 –1752) *Es sei $A \in \mathrm{GL}(n, \mathbb{K})$ eine reguläre Matrix und $b \in \mathbb{K}^{n\times 1}$ ein Spaltenvektor. Dann erhält man die Lösung $x \in \mathbb{K}^{n\times 1}$ des linearen Gleichungssystems $Ax = b$ folgendermaßen: Man ersetzt in der Matrix A die i te Spalte*

α_i durch die rechte Seite b und berechnet die entsprechende Determinante. Für $i = 1, \dots, n$ ist dann

$$x_i = \frac{1}{\det A} \det(\alpha_1, \dots, \alpha_{i-1}, b, \alpha_{i+1}, \dots, \alpha_n).$$

Also ist

$$x_i = \frac{1}{\det A} \begin{vmatrix} a_{11} & \dots & b_1 & \dots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots & \\ a_{n1} & \dots & b_n & \dots & a_{nn} \end{vmatrix}.$$

Beweis. Für die Lösung x hat man die folgende Identität in Spaltenvektoren: $x_1\alpha_1 + x_2\alpha_2 + \dots + x_n\alpha_n = b$. Da die Determinante linear in den Spalten ist, (D1)', gilt:

$$\begin{aligned} \det(b, \alpha_2, \dots, \alpha_n) &= \det(x_1\alpha_1 + x_2\alpha_2 + \dots + x_n\alpha_n, \alpha_2, \dots, \alpha_n) \\ &= x_1 \det(\alpha_1, \alpha_2, \dots, \alpha_n) + x_2 \det(\alpha_2, \alpha_2, \dots, \alpha_n) + \dots + x_n \det(\alpha_n, \dots, \alpha_n) \\ &= x_1 \det A. \end{aligned}$$

Daraus folgt die Behauptung für x_1; analog beweist man sie für $x_2, \dots, x_n$. ∎

4.4.2 Ebene Geometrie

Mit Hilfe von Determinanten lassen sich Geraden, Kreise, Kegelschnitte, Flächeninhalte besonders elegant darstellen.

Lemma 4.13 *Sind $A = (a_1, a_2)$ und $B = (b_1, b_2)$ zwei verschiedene Punkte im $\mathbb{R}^2$, dann liegt $(x_1, x_2) \in \mathbb{R}^2$ genau dann auf der Geraden durch A und B, wenn*

$$\begin{vmatrix} x_1 & x_2 & 1 \\ a_1 & a_2 & 1 \\ b_1 & b_2 & 1 \end{vmatrix} = 0. \tag{4.11}$$

Beweis. In (4.11) subtrahieren wir die zweite Zeile von der ersten und der dritten und entwickeln dann die Determinante nach der 3. Spalte:

$$|A| = \begin{vmatrix} x_1 - a_1 & x_2 - a_2 & 0 \\ a_1 & a_2 & 1 \\ b_1 - a_1 & b_2 - a_2 & 0 \end{vmatrix} = - \begin{vmatrix} x_1 - a_1 & x_2 - a_2 \\ b_1 - a_1 & b_2 - a_2 \end{vmatrix} = 0.$$

Dies ist äquivalent zu $(x_1 - a_1)(b_2 - a_2) - (x_2 - a_2)(b_1 - a_1) = 0$ und falls $b_2 \neq a_2$ weiter zu $\frac{x_1 - a_1}{x_2 - a_2} = \frac{a_2 - a_1}{b_2 - b_1}$, der bekannten 2-Punkte-Gleichung der Geraden. ∎

Weitere Formeln mit Determinanten

(a) Gegeben seien drei Punkte $A=(a_1,a_2)$, $B=(b_1,b_2)$ und $C=(c_1,c_2)$ der Ebene $\mathbb{R}^2$, die nicht alle zusammen fallen. Dann gibt es genau einen Kreis h oder eine Gerade h durch A, B und C. Ein Punkt $X=(x_1,x_2)$ liegt auf h genau dann, wenn

$$\begin{vmatrix} x_1^2+x_2^2 & x_1 & x_2 & 1 \\ a_1^2+a_2^2 & a_1 & a_2 & 1 \\ b_1^2+b_2 & b_1 & b_2 & 1 \\ c_1^2+c_2^2 & c_1 & c_2 & 1 \end{vmatrix} = 0.$$

(b) Ein *Kegelschnitt* ist die Menge der Punkte $(x,y)\in\mathbb{R}^2$, für die gilt

$$Ax^2+By^2+Cxy+Dx+Ey+F=0,$$

wobei A,B,C,D,E,F fixierte reelle Zahlen sind. Diese Menge kann eine Ellipse, Parabel, Hyperbel, ein Geradenpaar, eine Doppelgerade, ein Punkt oder die leere Menge sein. Bis auf die Normierung (division durch A) hat man also fünf freie Parameter. Durch 5 Punkte verschiedene Punkte a, b, c, d, e der Ebene verläuft daher genau ein Kegelschnitt $\mathbf{E}$. Der Punkt $x=(x_1,x_2)$ liegt genau dann auf $\mathbf{E}$, wenn gilt

$$\begin{aligned} x_1^2A+x_2^2B+x_1x_2C+x_1D+x_2E+1F&=0\\ a_1^2A+a_2^2B+a_1a_2C+a_1D+a_2E+1F&=0\\ b_1^2A+b_2^2B+b_1b_2C+b_1D+b_2E+1F&=0\\ c_1^2A+c_2^2B+c_1c_2C+c_1D+c_2E+1F&=0\\ d_1^2A+d_2^2B+d_1d_2C+d_1D+d_2E+1F&=0\\ e_1^2A+e_2^2B+e_1e_2C+e_1D+e_2E+1F&=0. \end{aligned}$$

Dies ist ein homogenes 6×6 lineares Gleichungssystem. Es besitzt genau dann nichtttriviale Lösungen (A,B,C,D,E,F), wenn die Koeffizientendeterminante gleich Null ist:

$$\begin{vmatrix} x_1^2 & x_2^2 & x_1x_2 & x_1 & x_2 & 1 \\ a_1^2 & a_2^2 & a_1a_2 & a_1 & a_2 & 1 \\ b_1^2 & b_2^2 & b_1b_2 & b_1 & b_2 & 1 \\ c_1^2 & c_2^2 & c_1c_2 & c_1 & c_2 & 1 \\ d_1^2 & d_2^2 & d_1d_2 & d_1 & d_2 & 1 \\ e_1^2 & e_2^2 & e_1e_2 & e_1 & e_2 & 1 \end{vmatrix} = 0.$$

(c) Die Fläche eines Dreiecks mit den Seitenlängen a, b, c beträgt

$$F=\frac{1}{4}\left|\det\begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & a^2 & b^2 \\ 1 & a^2 & 0 & c^2 \\ 1 & b^2 & c^2 & 0 \end{pmatrix}\right|^{\frac{1}{2}}.$$

(d) Liegen 4 Punkte in der Ebene und sind a_{ij}, $1 \le i,j \le 4$, die Abstände zwischen je zweien von ihnen, so gilt

$$\begin{vmatrix} 0 & a_{12}^2 & a_{13}^2 & a_{14}^2 & 1 \\ a_{21}^2 & 0 & a_{23}^2 & a_{24}^2 & 1 \\ a_{31}^2 & a_{32}^2 & 0 & a_{34}^2 & 1 \\ a_{41}^2 & a_{42}^2 & a_{43}^2 & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 \end{vmatrix} = 0.$$

4.4.3 Lineare Unabhängigkeit von Polynomen

Lemma 4.14 *Die Menge der Monome $B = \{1, x, x^2, \dots\}$ bildet eine Basis im Vektorraum $\mathbb{R}[x]$.*

Beweis. Nach Beispiel 2.3 (c) ist B erzeugend für $\mathbb{R}[x]$. Wir müssen die lineare Unabhängigkeit noch zeigen. dazu müssen wir die lineare Unabhängigkeit jeder endlichen Teilmenge $B_n = \{1, x, x^2, \dots, x^{n-1}\}$ nachweisen. Angenommen

$$\alpha_0 + \alpha_1 x + \alpha_2 x^2 + \cdots + \alpha_{n-1} x^{n-1} = 0$$

in $\mathbb{R}[x]$ für eine gewisse Linearkombination der Monome. Dann muss diese Identität insbesondere für n verschiedene paarweise verschiedene x-Werte $x_1, x_2, \dots, x_n$ gelten:

$$\alpha_0 + \alpha_1 x_i + \cdots + \alpha_{n-1} x_i^{n-1} = 0, \quad i = 1, \dots, n.$$

Dies ist nun ein lineares Gleichungssystem in den Variablen α_i, $i = 0, \dots, n-1$ mit der quadratischen $n \times n$-Kooeffizientenmatrix

$$A = \begin{pmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-1} \\ \cdots & \cdots & \cdots & \cdots & \\ 1 & x_n & x_n^2 & \cdots & x_n^{n-1} \end{pmatrix}$$

Die Determinante ist eine Vandermondsche Determinante, diese ist nach Übungsaufgabe 9.2 gleich

$$\det A = \prod_{1 \le i < j \le n} (x_j - x_i).$$

Da die x_i als paarweise verschieden gewählt wurden, ist $x_j - x_i \ne 0$ für alle Paare (i, j), $i \ne j$, also ist $\det A \ne 0$; das homogen lineare Gleichungssystem $A\alpha = 0$ besitzt nur die triviale Lösung $\alpha_0 = \alpha_1 = \cdots = \alpha_{n-1} = 0$. Damit ist B_n linear unabhängig und B ist eine Basis. ∎

Kapitel 5

Euklidische und unitäre Räume

5.1 Euklidische Räume

In diesem Abschnitt sei $\mathbb{K} = \mathbb{R}$ und wir betrachten nur reelle Vektorräume.
Mit Hilfe des Skalarproduktes verallgemeinern wir die Begriffe „Länge“ eines Vektors, „Winkel“ zwischen zwei Vektoren, wobei wir von den bekannten Begriffen im $\mathbb{R}^2$ und $\mathbb{R}^3$ starten.

Definition 5.1 Es sei V ein reeller Vektorraum. Ein *Skalarprodukt* in V ist eine Abbildung $\langle \cdot , \cdot \rangle : V \times V \to \mathbb{R}$, welche die folgenden Eigenschaften besitzt:

(i) $\langle \lambda v + \mu w \, , \, u \rangle = \lambda \langle v \, , \, u \rangle + \mu \langle w \, , \, u \rangle$ (Linearität im 1. Argument)

(ii) $\langle v \, , \, w \rangle = \langle w \, , \, v \rangle$ (Symmetrie).

(iii) $\langle v \, , \, v \rangle \geq 0$ für alle v und $\langle v \, , \, v \rangle = 0$ genau dann, wenn $v = 0$. (Positivität und Definitheit).

Ein reeller Vektorraum V, zusammen mit einem Skalarprodukt heißt *Euklidischer Raum*.

Aus den Eigenschaften (i) und (ii) folgt sofort, dass das Skalarprodukt linear in *beiden* Argumenten ist, also dass auch gilt:

$$\text{(i)}' \quad \langle u \, , \, \lambda v + \mu w \rangle = \lambda \langle u \, , \, v \rangle + \mu \langle u \, , \, w \rangle .$$

Diese Eigenschaft lässt sich auf beliebig viele Summanden auf beiden Seiten verallgemeinern, so dass für $v = \sum_{i=1}^{n} x_i b_i$, $w = \sum_{j=1}^{m} y_j c_j$ gilt

$$\langle v \, , \, w \rangle = \sum_{i=1}^{n} \sum_{j=1}^{m} x_i y_j \left\langle b_i \, , \, c_j \right\rangle . \tag{5.1}$$

Bemerkung 5.1 (a) Es sei $B = \{b_1, \dots, b_n\}$ eine Basis von V. Setzt man in (5.1) $m = n$ und $c_j = b_j$, $j = 1, \dots, n$, so erkennt man, dass das Skalarprodukt $\langle v \, , \, w \rangle$ für alle $v, w \in V$ *eindeutig bestimmt* ist, wenn man die n^2 Skalarprodukte $\left\langle b_i \, , \, b_j \right\rangle$, $i, j = 1, \dots, n$ kennt

$$\langle v \, , \, w \rangle = \sum_{i,j=1}^{n} x_i \, y_j \, \left\langle b_i \, , \, b_j \right\rangle \tag{5.2}$$

Wir nennen

$$A = (\langle b_i\,,\,b_j\rangle)_{i,j=1\ldots n} \in \mathbb{R}^{n\times n}$$

die *Matrix des Skalarproduktes* bezüglich der Basis B. Wegen der Symmetrie des Skalarproduktes gilt $A^\top = A$.

(b) **Matrixschreibweise.** Setzt man $x = (x_1,\ldots,x_n)^\top$, $y = (y_1,\ldots,y_n)^\top$ und die Matrix A in (5.2) ein, so hat man

$$\langle v\,,\,w\rangle = \sum_{i,j=1}^{n} x_i\,a_{ij}\,y_j = \sum_{i=1}^{n} x_i(Ay)_i = x^\top Ay.$$

(c) Ist umgekehrt $A \in \mathbb{R}^{n\times n}$ eine beliebige symmetrische Matrix und $v = \sum_{i=1}^n x_i b_i$, $w = \sum_{j=1}^n y_j b_j$, so definiert

$$\langle v\,,\,w\rangle = x^\top Ay,$$

eine Abbildung, die symmetrisch und bilinear ist. Eigenschaft (iii), die positive Definitheit, muss extra geprüft werden.

Beispiel 5.1 (a) **Standardskalarprodukt im $\mathbb{R}^n$.** Die Matrix des Standardskalarproduktes bzgl. der Standardbasis ist die Einheitsmatrix. Für Vektoren $x = (x_1,\ldots,x_n)$ und $y = (y_1,\ldots,y_n)$ des $\mathbb{R}^n$ setzen wir:

$$\langle x\,,\,y\rangle = \sum_{i=1}^{n} x_i\,y_i.$$

Im Falle des euklidischen Raumes mit Standard-SKP schreiben wir auch $\langle x\,,\,y\rangle = x{\cdot}y$. Es gilt $x{\cdot}y = x^\top y$, $A = I_n$. Wegen $\langle x\,,\,x\rangle = \sum_{i=1}^n x_i^2 \geq 0$ ist das Standard-SKP positiv. Wenn $\langle x\,,\,x\rangle = 0$, so $\sum_{i=1}^n x_i^2 = 0$, so $x_i^2 = 0$ für alle i; also $x = 0$; das SKP ist positiv-definit.

(b) $V = \mathbb{R}^2$. Für $x = (x_1,x_2) \in V$ und $y = (y_1,y_2) \in \mathbb{R}^2$ definieren wir

$$\langle x\,,\,y\rangle := x_1y_1 + 5x_1y_2 + 5x_2y_1 + 27x_2y_2.$$

Die Eigenschaften (i) und (ii) sind sofort klar; Dieses Skalarprodukt ist auch nichtnegativ, denn für alle $x = (x_1,x_2) \in \mathbb{R}^2$ ist $\langle x\,,\,x\rangle = x_1^2 + 10x_1x_2 + 27x_2^2 = (x_1+5x_2)^2 + 2x_2^2 \geq 0$. Auch die positive Definitheit ist jetzt klar, denn

$$\langle x\,,\,x\rangle = 0 \Longrightarrow (x_1+5x_2)^2 + 2x_2^2 = 0 \Longrightarrow x_1 + 5x_2 = x_2 = 0 \Longrightarrow x_1 = x_2 = 0.$$

Die Matrix des SKP bzgl. der Standardbasis ist $A = \begin{pmatrix} 1 & 5 \\ 5 & 27 \end{pmatrix}$.

(c1) $\mathrm{C}[a,b]$ mit $\langle f\,,\,g\rangle = \int_a^b f(x)g(x)\,\mathrm{d}x$. (i) Wir zeigen die Linearität im 1. Argument:

$$\langle \lambda_1 f_1 + \lambda_2 f_2\,,\,g\rangle = \int_a^b (\lambda_1 f_1(t) + \lambda_2 f_2(t))g(t)\,\mathrm{d}t = \lambda_1\int_a^b f_1(t)g(t)\,\mathrm{d}t + \lambda_2\int_a^b f_2(t)g(t)\,\mathrm{d}t.$$

In der letzten Gleichung benutzten wir die Linearität des Integrals. Hieraus folgt die Linearität. Die Symmetrie (ii) ist klar, da $f(t)g(t) = g(t)f(t)$ für alle t.

(iii) Die Klammer ist nichtnegativ, denn das Integral einer nichtnegativen Funktion $f(t)^2 \geq 0$ ist wieder nichtnegativ:

$$\langle f\,,\,f\rangle = \int_a^b f(t)f(t)\,\mathrm{d}t = \int_a^b f(t)^2\,\mathrm{d}t \geq 0.$$

Für die Definitheit benutzt man die Stetigkeit von f entscheidend.
(c2) Es sei $V = \mathbb{R}_2[x]$. Dann definiert $\langle p\,,\,q\rangle = \int_{-1}^{1} p(x)q(x)\,\mathrm{d}x$ ein SKP auf V. Allgemein gilt nämlich: Ist V ein Raum mit SKP und $U \subset V$ ein linearer Teilraum, so definiert die Einschränkung des SKP von V auch auf U ein Skalarprodukt. In unserem Falle kann man V auffassen als 3-dimensionalen Teilraum von $\mathrm{C}[-1,1]$, denn in der Tat ist ein Polynom $p \in \mathbb{R}[x]$ bereits eindeutig festgelegt durch seine Werte im Intervall $[-1,1]$.
Ist $B := \{b_1, b_2, b_3\} := \{1, x, x^2\}$ eine Basis in V, so ist die Matrix des SKP bezüglich dieser Basis gleich

$$G = (\langle b_i\,,\,b_j\rangle)_{i,j=1,2,3} = \begin{pmatrix} 2 & 0 & \frac{2}{3} \\ 0 & \frac{2}{3} & 0 \\ \frac{2}{3} & 0 & \frac{2}{5} \end{pmatrix}.$$

(d) $\ell_2 = \{(x_n) \mid x_n \in \mathbb{R},\quad \sum_{n=1}^{\infty} x_n^2 < \infty\}$ mit SKP $\langle x\,,\,y\rangle = \sum_{n=1}^{\infty} x_n y_n$ heißt „klein ell 2“ oder *Hilbertscher Folgenraum*.

5.1.1 Unitäre Räume

Völlig analog kann man in komplexen Vektorräumen ein Skalarprodukt definieren, wobei die Symmetrie leicht abgeändert wird.

Definition 5.2 Es sei V ein komplexer Vektorraum. Ein *Skalarprodukt* in V ist eine Funktion $\langle\cdot\,,\,\cdot\rangle : V \times V \to \mathbb{C}$, welche die folgenden Eigenschaften besitzt:

(i) $\langle \lambda v + \mu w\,,\,u\rangle = \lambda\langle v\,,\,u\rangle + \mu\langle w\,,\,u\rangle$ (linear im ersten Argument)

(ii) $\langle v\,,\,w\rangle = \overline{\langle w\,,\,v\rangle}$ (hermitesch).

(iii) $\langle v\,,\,v\rangle \geq 0$ für alle v und $\langle v\,,\,v\rangle = 0$ genau dann, wenn $v = 0$ (positiv definit).

Ein komplexer Vektorraum V, zusammen mit einem Skalarprodukt heißt *unitärer Raum*.

Bemerkung 5.2 Aus den Eigenschaften (i) und (ii) folgt sofort, dass das Skalarprodukt *antilinear* im zweiten Argumenten ist, das heißt, es gilt:

$$\langle u\,,\,\lambda v + \mu w\rangle = \overline{\langle \lambda v + \mu w\,,\,u\rangle} = \overline{\lambda\langle v\,,\,u\rangle + \mu\langle w\,,\,u\rangle} = \overline{\lambda}\langle u\,,\,v\rangle + \overline{\mu}\langle u\,,\,w\rangle.$$

Wie im euklidischen Fall ist auch hier das SKP schon eindeutig bestimmt durch seine Werte auf einer Basis, also durch die Matrix $(\langle b_i\,,\,b_j\rangle)$ des SKP.
Die Matrix G eines komplexen Skalarproduktes ist hermitesch, das heißt, es gilt $G^* = G$, vgl. auch Definition 3.7. In der Tat ist $g_{ij} = \langle e_i\,,\,e_j\rangle = \overline{\langle e_j\,,\,e_i\rangle} = \overline{g_{ji}}$.

Beispiel 5.2 (a) Standardskalarprodukt im $\mathbb{C}^n$. Die Matrix des Standardskalarproduktes bzgl. der Standardbasis ist die Einheitsmatrix. Für Vektoren $x = (x_1, \ldots, x_n)$ und $y = (y_1, \ldots, y_n)$ des $\mathbb{C}^n$ gilt

$$\langle x \,, y \rangle = \sum_{i=1}^{n} x_i \overline{y_i}.$$

(b) Es sei V der Vektorraum der komplexwertigen stetigen Funktionen auf dem Intervall $[a, b]$, $V = \mathrm{C}([a, b], \mathbb{C})$. Dann definiert

$$\langle f \,, g \rangle = \int_a^b f(x)\overline{g(x)}\,\mathrm{d}x$$

ein Skalarprodukt auf V.

5.2 Die Cauchy-Schwarzsche Ungleichung und die Norm eines Vektors

Der Begriff der *Norm* eines Vektors verallgemeinert die *Länge* eines Vektors auf beliebige euklidische und unitäre Räume.

Satz 5.1 (Cauchy-Schwarzsche Ungleichung) *Es sei V ein Vektorraum mit Skalarprodukt. Dann gilt für alle $x, y \in V$*

$$\left| \langle x \,, y \rangle \right|^2 \le \langle x \,, x \rangle \, \langle y \,, y \rangle \,. \tag{5.3}$$

Gleichheit tritt genau dann ein, wenn $\{x, y\}$ linear abhängig ist.

Beweis. Wir führen den Beweis nur für den Fall eines unitären Raumes durch. Für $y = 0$ ist die Aussage richtig, denn auf beiden Seiten steht 0. In der Tat ist $\langle x \,, 0\rangle = \langle x \,, 0 \cdot 0\rangle = \overline{0}\,\langle x \,, 0\rangle = 0$ für alle $x \in V$.
Sei nun $y \neq 0$. Für beliebiges $t \in \mathbb{C}$ gilt:

$$0 \le \langle x - t y \,, x - t y \rangle = \langle x \,, x \rangle - t\,\overline{\langle x \,, y \rangle} - \overline{t}\,\langle x \,, y \rangle + |\,t\,|^2 \langle y \,, y \rangle \,.$$

Setzten wir hier $t := \frac{\langle x \,, y \rangle}{\langle y \,, y \rangle}$, so erhalten wir

$$0 \le \langle x \,, x \rangle - \frac{1}{\langle y \,, y \rangle} \left| \langle x \,, y \rangle \right|^2 - \frac{1}{\langle y \,, y \rangle} \left| \langle x \,, y \rangle \right|^2 + \frac{\left| \langle x \,, y \rangle \right|^2}{\langle y \,, y \rangle}.$$

Multipliziert man diese Ungleichung mit $\langle y \,, y \rangle$, so erhält man die behauptete Ungleichung. Notwendig für die Gleichheit ist, dass in allen Abschätzungen, oangefangen mit der ersten, Gleichheit herrscht, also $0 = \langle x - t y \,, x - t y \rangle$. Wegen der Definitheit folgt hieraus $x - t y = 0$. Das heißt aber, $x = t y$ für ein $t \in \mathbb{C}$. Umgekehrt, ist $x = t y$ für ein $t \in \mathbb{C}$, so folt $\left| \langle x \,, y \rangle \right|^2 = \langle x \,, x \rangle \langle y \,, y \rangle$. Es gilt also Gleichheit genau dann, wenn x und y linear abhängig sind. ∎

Die Cauchy-Schwarz-Ungleichung für das Standardskalarprodukt im $\mathbb{R}^n$ ist schon aus der Analysis-Vorlesung bekannt: Für Vektoren $x = (x_1, \ldots, x_n)$ und $y_1 = (y_1, \ldots, y_n)$ des $\mathbb{R}^n$ gilt

$$\left(\sum_{i=1}^{n} x_i y_i\right)^2 \leq \sum_{i=1}^{n} x_i^2 \sum_{i=1}^{n} y_i^2.$$

Im Falle von $\mathrm{C}([a,b],\mathbb{C})$, den stetigen, komplexwertigen Funktionen auf $[a,b]$ lautet die Cauchy-Schwarz-Ungleichung:

$$\left|\int_a^b f(t)\overline{g(t)}\,\mathrm{d}t\right|^2 \leq \int_a^b \left|f(t)\right|^2 \mathrm{d}t \int_a^b \left|g(t)\right|^2 \mathrm{d}t.$$

Definition 5.3 Es sei V ein reeller oder komplexer Vektorraum. Eine Funktion $\|\cdot\| : V \to \mathbb{R}$ heißt *Norm*, wenn die folgenden drei Bedingungen erfüllt sind:

(a) $\|x\| \geq 0$ und $\|x\| = 0$ genau dann, wenn $x = 0$ (Positivität und Definitheit).

(b) $\|\lambda x\| = |\lambda| \, \|x\|$, $\lambda \in \mathbb{K}$, $x \in V$ (Homogenität).

(c) $\left\|x + y\right\| \leq \|x\| + \left\|y\right\|$ für alle $x, y \in V$ (Dreiecksungleichung)

Ein linearer Raum mit Norm heißt *normierter Raum*.

Satz 5.2 *Es sei V ein Raum mit SKP $\langle\cdot\,,\cdot\rangle$. Dann wird durch $\|x\| := \sqrt{\langle x\,, x\rangle}$ auf V eine Norm definiert.*

Benutzt man die Norm, so hat die Cauchy-Schwarz-Ungleichung (5.3) die Form

$$\left|\langle x\,, y\rangle\right| \leq \|x\| \left\|y\right\|. \tag{5.4}$$

Beweis. (a) Die angegebene Norm ist wohldefiniert und nichtnegativ, da $r = \langle x\,, x\rangle \geq 0$ und $\sqrt{r} \geq 0$. Wenn $\|x\| = \sqrt{\langle x\,, x\rangle} = 0$, so folgt $\langle x\,, x\rangle = 0$, und hieraus wegen der Definitheit des Skalarproduktes $x = 0$.
(b) Für $\lambda \in \mathbb{K}$ und $x \in V$ gilt

$$\|\lambda x\| = \sqrt{\langle \lambda x\,, \lambda x\rangle} = \sqrt{\lambda\overline{\lambda}\,\langle x\,, x\rangle} = \sqrt{|\lambda|^2}\,\sqrt{\langle x\,, x\rangle} = |\lambda|\,\|x\|.$$

(c) Wir benutzen, dass für alle $z \in \mathbb{C}$ gilt: $z + \overline{z} = 2\,\mathrm{Re}\,z$ und $\mathrm{Re}(z) \leq |z|$. Für alle $x, y \in V$ gilt

$$\begin{aligned}
\left\|x + y\right\|^2 &= \langle x + y\,, x + y\rangle = \langle x\,, x\rangle + \langle x\,, y\rangle + \langle y\,, x\rangle + \langle y\,, y\rangle \\
&= \|x\|^2 + 2\,\mathrm{Re}\,\langle x\,, y\rangle + \left\|y\right\|^2 \leq \|x\|^2 + 2\left|\langle x\,, y\rangle\right| + \left\|y\right\|^2 \\
&\leq \|x\|^2 + 2\,\|x\|\,\left\|y\right\| + \left\|y\right\|^2 = (\|x\| + \left\|y\right\|)^2.
\end{aligned}$$

Dabei haben wir in der letzten Ungleichung die CSU benutzt. Gleichheit tritt genau dann ein, wenn $x = \lambda y$ oder $y = \lambda x$ mit einem $\lambda \geq 0$. ∎

Bemerkung 5.3 (a) Nicht jeder normierte Raum besitzt ein SKP. So ist zum Beispiel der $\mathbb{R}^2$ mit $\|(x_1,x_2)\| := |x_1|+|x_2|$ ein normierter Raum, zu dem es kein Skalarprodukt $\langle\cdot\,,\cdot\rangle$ gibt, mit $\|x\| = \sqrt{\langle x\,,x\rangle}$.
(b) Für einen normierten Raum mit SKP gilt nämlich notwendig die *Parallelogrammidentität*:

$$\left\|x-y\right\|^2+\left\|x+y\right\|^2 = 2(\|x\|^2+\left\|y\right\|^2) \qquad \forall\, x,y\in V. \tag{5.5}$$

Diese Bedingung ist auch hinreichend. Ist sie erfüllt, so definiert in einem euklidischen Raum

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

das zugehörige Skalarprodukt. Für unitäre Räume gibt es eine ähnliche Formel.

5.3 Orthonormalbasen

Wir definieren den Begriff „Orthogonalität“ (oder des „senkrecht stehens“) und zeigen, dass jeder endlichdimensionale Raum mit SKP eine Basis aus paarweise orthogonalen Vektoren besitzt, die die Länge 1 haben. Das Schmidtsche Orthonormalisierungsverfahren ist eine Methode, wie man aus einem Erzeugendensystem eine Orthonormalbasis machen kann.

5.3.1 Winkel und Orthogonalität

Definition 5.4 (a) Es sei V ein euklidischer Raum mit Skalarprodukt $\langle\cdot\,,\cdot\rangle$ und a, c von Null verschiedene Vektoren in V. Dann heißt $\varphi\in[0,\pi]$ der *Öffnungswinkel* zwischen a und c, wenn gilt

$$\cos\varphi = \frac{\langle a\,,c\rangle}{\|a\|\,\|c\|}.$$

(b) Es sei V ein Raum mit Skalarprodukt. Dann heißen $a,c\in V$ zueinander *orthogonal*, wenn $\langle a\,,c\rangle = 0$. Man schreibt in diesem Falle auch $a\perp c$.

Man beachte, dass (a) sinnvoll ist, denn nach der Cauchy-Schwarz-Ungleichung gilt für alle $a,c\neq 0$

$$-\|a\|\,\|c\|\le\langle a\,,c\rangle\le\|a\|\,\|c\| \Longrightarrow -1\le\frac{\langle a\,,c\rangle}{\|a\|\,\|c\|}\le 1.$$

Da die Kosinusfunktion auf $[0,\pi]$ bijektiv ist und alle Werte zwischen -1 und 1 annimmt, gibt es einen eindeutig bestimmten Winkel φ zwischen a und c. Der Begriff der Orthogonalität, das heißt, der Öffnungswinkel ist $\pi/2$ und $\cos\varphi = 0$, hat auch in unitären Räumen Sinn und auch für Vektoren, die gleich Null sind.

Beispiel 5.3 (a) Die Vektoren $a = (2,0)$ und $c = (1,-\sqrt{3})$ im euklidischen Raum $\mathbb{R}^2$ (mit dem Standardskalarprodukt) schließen einen Winkel von $\frac{\pi}{3}$ ein, denn es ist $\|a\| = \|c\| = 2$ und

$$\cos\frac{\pi}{3} = \frac{1}{2} = \frac{1}{\|a\|\,\|c\|}(2,0)\cdot(1,-\sqrt{3}) = \frac{1}{2\cdot 2}(2+0\cdot(-\sqrt{3})) = \frac{1}{2}.$$

(b) Wir betrachten den unitären Raum der komplexwertigen, stetigen Funktionen auf $[0,2\pi]$, $V = \mathrm{C}([0,2\pi],\mathbb{C},)$ mit dem SKP $\langle f\,,\,g\rangle = \int_0^{2\pi} f(t)\overline{g(t)}\,\mathrm{d}t$. Die Vektoren $x(t) = \sin t$ und $y(t) = \cos t$ sind orthogonal in V, denn

$$\langle x\,,\,y\rangle = \int_0^{2\pi} \sin t\,\overline{\cos t}\,\mathrm{d}t = \frac{1}{2}\int_0^{2\pi} \sin(2t)\,\mathrm{d}t = -\frac{1}{4}\cos(2t)|_0^{2\pi} = 0.$$

Das heißt, $x \perp y$.
(c) Man beachte, dass aus $x \perp y$ folgt $y \perp x$ und dass aus $x \perp x$ folgt $x = 0$ (positive Definitheit des SKP). Der Nullvektor steht auf allen Vektoren senkrecht $0 \perp x$ für alle $x \in V$.

5.3.2 Orthogonalsysteme

Definition 5.5 Es sei V ein Raum mit Skalarprodukt und $C = \{c_i \mid i \in I\} \subset V$ eine beliebige Teilmenge. Hierbei sei stets $I \subset \mathbb{N}$ eine nichtleere Menge.
(a) C heißt *Orthogonalsystem*, abgekürzt OS, wenn für alle $i, j \in I$ mit $i \neq j$ gilt $\langle c_i\,,\,c_j\rangle = 0$.
(b) C heißt *Orthonormalsystem*, abgekürzt ONS, wenn gilt

$$\langle c_i\,,\,c_j\rangle = \delta_{ij}, \qquad \forall\, i, j \in I.$$

(c) C heißt *Orthonormalbasis*, ONB, wenn C ein Orthonormalsystem ist und gleichzeitig eine Basis in V.

Bemerkung 5.4 Es sei C eine orthogonale Familie im Skalarproduktraum V, die nicht den Nullvektor enthält. Dann kann man C in eine orthonormale Familie C' überführen, indem man die einzelnen Vektoren *normiert*: Man setzt

$$c_i' := \frac{c_i}{\|c_i\|}, \quad i \in I.$$

Dann ist $C' = \{c_i' \mid i \in I\}$ ein ONS. In der Tat ist jeder Vektor $x/\|x\|$, $x \neq 0$, normiert, denn

$$\left\|\frac{x}{\|x\|}\right\| = \left\|\frac{1}{\|x\|}\,x\right\| = \left|\frac{1}{\|x\|}\right|\,\|x\| = \frac{\|x\|}{\|x\|} = 1.$$

Beispiel 5.4 (a) Im euklidischen (unitären) Raum $\mathbb{R}^n$ ($\mathbb{C}^n$) mit dem Standardskalarprodukt ist die Standardbasis $\{\mathrm{e}_1, \dots, \mathrm{e}_n\}$ eine ONB.
(b) Im $\mathbb{R}^3$ mit Standard-SKP sind $b_1 = (1,0,1)$, $b_2 = (-1,0,1)$ und $b_3 = (0,1,0)$ orthogonal. Die zugehörige orthonormale Familie B' ist

$$B' := \left\{\frac{1}{\sqrt{2}}(1,0,1),\ \frac{1}{\sqrt{2}}(-1,0,1),\ (0,1,0)\right\}.$$

Da B' linear unabhängig ist, ist B' neben $\{\mathrm{e}_1, \mathrm{e}_2, \mathrm{e}_3\}$ eine weitere ONB im $\mathbb{R}^3$.
(c) Es sei $V = \mathrm{C}[0,2\pi]$ mit dem SKP aus Beispiel 5.1 (c1). Die Familie

$$\{1, \cos x, \sin x, \cos 2x, \sin 2x, \dots\}$$

ist orthogonal in V, es gilt nämlich für alle nichtnegativen ganzen Zahlen $m, n \in \mathbb{N}_0$:

$$\int_0^{2\pi} \cos(mt)\cos(nt)\,\mathrm{d}t = 0, \quad \text{falls} \quad m \neq n,$$
$$\int_0^{2\pi} \sin(mt)\sin(nt)\,\mathrm{d}t = 0, \quad \text{falls} \quad m \neq n,$$
$$\int_0^{2\pi} \cos(mt)\sin(nt)\,\mathrm{d}t = 0.$$

Dieses Resultat spielt in der Theorie der Fourierreihen eine große Rolle.

Lemma 5.3 (Pythagoras) *Es sei* $\{c_1, \dots, c_n\}$ *ein OS in* V. *Dann gilt*

$$\|c_1 + \cdots + c_n\|^2 = \|c_1\|^2 + \cdots + \|c_n\|^2. \tag{5.6}$$

Beweis. durch vollständige Induktion über n. Für $n = 1$ ist die Behauptung offenbar. Die Behauptung möge nun richtig sein für jedes OS mit $n-1$ Vektoren, wir zeigen sie für das OS $C = \{c_1, \dots, c_n\}$. Dazu bemerken wir, dass $C' = \{c_1 + c_2, c_3, \dots, c_n\}$ ein OS aus $n-1$ Vektoren ist, also gilt nach Induktionsvoraussetzung

$$\begin{aligned}\|(c_1+c_2)+c_3+\cdots+c_n\|^2 &= \|c_1+c_2\|^2 + \|c_3\|^2 + \cdots + \|c_n\|^2 = \langle c_1+c_2\,,\, c_1+c_2\rangle + \sum_{i=3}^{n} \|c_i\|^2 \\ &= \langle c_1\,,\,c_1\rangle + \langle c_1\,,\,c_2\rangle + \langle c_2\,,\,c_1\rangle + \langle c_2\,,\,c_1\rangle + \sum_{i=3}^{n}\|c_i\|^2 = \sum_{i=1}^{n}\|c_i\|^2.\end{aligned}$$

■

Lemma 5.4 *Jede orthogonale Menge, die den Nullvektor nicht enthält, ist linear unabhängig.*

Beweis. Es sei $C = \{c_i \mid i \in I\}$ ein OS, das den Nullvektor nicht enthält. Angenommen, es gibt eine Linearkombination $\sum_{j=1}^{n} \alpha_j c_j = 0$, $\alpha_j \in \mathbb{K}$. Bilden wir mit dieser Identität das SKP mit einem festen c_k, $k = 1, \dots, n$, so erhalten wir

$$0 = \langle c_k\,,\,0\rangle = \left\langle c_k\,,\, \sum_{j=1}^{n} \alpha_j c_j \right\rangle = \sum_{j=1}^{n} \alpha_j \langle c_j\,,\,c_k\rangle = \alpha_k \langle c_k\,,\,c_k\rangle.$$

Wegen $c_k \neq 0$ ist $\langle c_k\,,\,c_k\rangle \neq 0$ und daher $\alpha_k = 0$. Da dies für alle $k = 1, \dots, n$ gilt, ist C linear unabhängig. ■

Für SKP-Räume ist es sehr einfach, die Koordinatendarstellung eines Vektors bezüglich einer ONB aufzuschreiben.

Satz 5.5 *Es sei V ein endlichdimensionaler Vektorraum mit Skalarprodukt, und es sei $B = \{b_1, \dots, b_n\}$ eine Orthonormalbasis von V.*
(a) Dann gilt für jedes $x \in V$

$$x = \sum_{i=1}^{n} \langle x\,,\, b_i \rangle\, b_i. \tag{5.7}$$

Für alle $x, y \in V$ gilt

$$\text{(b)}\quad \langle x\,,\, y \rangle = \sum_{i=1}^{n} \langle x\,,\, b_i \rangle\, \langle b_i\,,\, y \rangle\,. \tag{5.8}$$

$$\text{(c)}\quad \|x\|^2 = \sum_{i=1}^{n} |\langle x\,,\, b_i \rangle|^2\,. \tag{5.9}$$

Beweis. Es sei $x = \sum_{i=1}^{n} \alpha_i b_i$ die Koordinatendarstellung von x bezüglich der Basis B. Wir bilden das SKP mit dem Vektor b_k und erhalten

$$\langle x\,,\, b_k \rangle = \left\langle \sum_{i=1}^{n} \alpha_i b_i\,,\, b_k \right\rangle = \sum_{i=1}^{n} \alpha_i \langle b_i\,,\, b_k \rangle = \sum_{i=1}^{n} \alpha_i \delta_{ik} = \alpha_k.$$

Hieraus folgt die Behauptung (a).
(b) Folgt aus (a) durch Bildung des SKP mit y. (c) folgt aus (b) indem man $y = x$ setzt und beachtet, dass $\langle x\,,\, b_i \rangle \langle b_i\,,\, x \rangle = |\langle x\,,\, b_i \rangle|^2$. ■

Bemerkung 5.5 Der obige Satz spielt in der Theorie der Fourierreihen eine wichtige Rolle: die rechte Seite von (5.7) heißt im unendliche-dimensionalen Fall *Fourierreihe* von x und $\langle x\,,\, b_k \rangle$, $k \in \mathbb{N}$, sind die *Fourierkoeffizienten.*

Beispiel 5.5 Bestimmen Sie die Koordinaten von $v = (3, 4, -8, 12) \in \mathbb{R}^4$ bezüglich der Orthonormalbasis $B = \{\frac{1}{2}(1,1,1,1), \frac{1}{2}(-1,1,-1,1), \frac{1}{2}(1,-1,-1,1), \frac{1}{2}(1,1,-1,-1)\}$ des Standardskalarproduktes. Nach (5.7) ist

$$\begin{aligned} v &= \langle v\,,\, b_1 \rangle\, b_1 + \langle v\,,\, b_2 \rangle\, b_2 + \langle v\,,\, b_3 \rangle\, b_3 + \langle v\,,\, b_4 \rangle\, b_4 \\ &= \frac{1}{2}((3+4-8+12)b_1 + (-3+4+8+12)b_2 + (3-4+8+12)b_3 + (3+4+8-12)b_4) \\ &= \frac{1}{2}(11, 21, 19, 3). \end{aligned}$$

Die Vektoren $\{b_1, b_2, b_3, -b_4\}$ (Beispiel aus der Vorlesung) bilden ebenso eine ONB des $\mathbb{R}^4$.

5.3.3 Orthogonalisierung

Wir werden sehen, dass jeder endlichdimensionale SKP-Raum eine Orthonormalbasis besitzt. Der Beweis besteht in der Angabe eines expliziten Verfahrens, das auf Jørgen Pedersen Gram (dänischer Mathematiker, 1850 — 1916) und Erhardt Schmidt (Universität Berlin, 1876 –1959) zurück geht.

Satz 5.6 (Schmidtsches Orthogonalisierungsverfahren) *Es sei V ein linearer Raum mit Skalarprodukt und $C=\{c_1,\dots,c_n\}$ eine linear unabhängige Menge in V.*
Dann gibt es eine orthonormale Menge $B=\{e_1,e_2,\dots,e_n\}$ in V mit

$$\operatorname{lin}\{c_1,\dots,c_k\}=\operatorname{lin}\{e_1,\dots,e_k\},\quad k=1,\dots,n. \tag{5.10}$$

Beweis. Die Vektoren $e_1,\dots,e_n$ werden rekursiv definiert. (a) Wegen der linearen Unabhängigkeit von C ist $c_1\neq 0$ und wir setzten

$$e_1:=\frac{c_1}{\|c_1\|}.$$

Dann ist e_1 ein normierten Vektor und $\mathbb{K}c_1=\mathbb{K}e_1$; die Bedingung (5.10) ist erfüllt.
(b) Angenommen, die Vektoren $e_1,\dots,e_j$ seien bereits so definiert, dass (5.10) für $k=1,\dots,j$ gilt. Wir konstruieren nun

$$f_{j+1}:=c_{j+1}-\sum_{k=1}^{j}\langle c_{j+1}\,,\,e_k\rangle\; e_k.$$

Die Mengen $D=\{e_1,\dots,e_j,c_{j+1}\}$ und $D'=\{e_1,\dots,e_j,f_{j+1}\}$ haben dieselbe lineare Hülle, da die Koeffizienten oben vor f_{j+1} und c_{j+1} beide gleich 1 sind:

$$\operatorname{lin}D'=\operatorname{lin}D=\operatorname{lin}\{c_1,c_2,\dots,c_{j+1}\}.$$

Insbesondere ist $f_{j+1}\neq 0$, da sonst $c_{j+1}\in\operatorname{lin}\{c_1,\dots,c_j\}$. Wir zeigen, dass $f_{j+1}\perp e_i$ für alle $i=1,\dots,j$. Nach Induktionsvoraussetzung gilt $\langle e_i\,,\,e_k\rangle=\delta_{ik}$ für alle $i,k=1,\dots,j$. Daher ist für $i=1,\dots,j$:

$$\begin{aligned}\langle f_{j+1}\,,\,e_i\rangle&=\left\langle c_{j+1}-\sum_{k=1}^{j}\langle c_{j+1}\,,\,e_k\rangle\; e_k\,,\,e_i\right\rangle=\langle c_{j+1}\,,\,e_i\rangle-\sum_{k=1}^{j}\langle c_{j+1}\,,\,e_k\rangle\langle e_k\,,\,e_i\rangle\\&=\langle c_{j+1}\,,\,e_i\rangle-\sum_{k=1}^{j}\langle c_{j+1}\,,\,e_k\rangle\,\delta_{ki}=\langle c_{j+1}\,,\,e_i\rangle-\langle c_{j+1}\,,\,e_i\rangle=0.\end{aligned}$$

Setzt man daher

$$e_{j+1}=\frac{f_{j+1}}{\left\|f_{j+1}\right\|},$$

so bilden die Vektoren $\{e_1,\dots,e_j,e_{j+1}\}$ ein ONS, das den selben Unterraum aufspannt, wie $\{c_1,\dots,c_j,c_{j+1}\}$.
(c) Eindeutigkeit. Angenommen, die Vektoren $e_1,\dots,e_j$ seien bereits so definiert, dass (5.10) für $k=1,\dots,j$ gilt. Angenommen, es gibt einen weiteren Vektor e'_{j+1} mit derselben Eigenschaft. Dann gilt

$$e'_{j+1}=\alpha c_{j+1}-\sum_{k=1}^{j}\alpha'_k e_k=\alpha\left(c_{j+1}-\sum_{k=1}^{j}\alpha_k e_k\right)$$

mit einem $\alpha\neq 0$ (sonst wäre e'_{j+1} in der linearen Hülle der $c_1,\dots,c_j$ und könnte nicht $\operatorname{lin}\{c_1,\dots,c_{j+1}\}$ aufspannen). Für alle $k=1,\dots,j$ folgt dann aus der Orthogonalität

$$\left\langle e'_{j+1}\,,\,e_k\right\rangle=\alpha\left(\langle c_{j+1}\,,\,e_k\rangle-\alpha_k\right)\Longrightarrow\alpha_k=\langle c_{j+1}\,,\,e_k\rangle.$$

Also gilt $e'_{j+1} = \alpha f_{j+1}$. Wegen der Normiertheit von e_{j+1} und e'_{j+1} folgt hieraus $e'_{j+1} = \beta e_{j+1}$, wobei $|\beta| = 1$ ist — im komplexen Fall ist $\beta = \mathrm{e}^{\mathrm{i}\varphi}$ ein Phasenfaktor, im reellen Fall ist $\beta = \pm 1$. ∎

Wendet man das Schmidtsche Verfahren auf eine beliebige Basis eines SKP-Raumes an, so erhält man eine orthonormierte Basis.

Folgerung 5.7 *Jeder endlichdimensionale Raum mit Skalarprodukt besitzt eine Orthonormalbasis.*

Beispiel 5.6 (a) Wenden Sie das Schmidtsche Verfahren an auf die Vektoren

$$c_1 = (4, 2, -2, -1), \quad c_2 = (2, 2, -4, -5), \quad c_3 = (0, 8, -2, -5)$$

im $\mathbb{R}^4$ mit dem Standardskalarprodukt. Man erhält

$$\begin{aligned}
e_1 &= \frac{c_1}{\|c_1\|} = \frac{1}{5}(4, 2, -2, -1)\\
f_2 &= c_2 - \langle c_2, e_1\rangle\, e_1 = (2, 2, -4, -5) - \frac{1}{5}\cdot 25 \frac{1}{5}(4, 2, -2, -1) = (-2, 0, -2, -4)\\
e_2 &= \frac{f_2}{\|f_2\|} = \frac{1}{\sqrt{24}}(-2, 0, -2, -4),\\
f_3 &= c_3 - \langle c_3, e_1\rangle\, e_1 - \langle c_3, e_2\rangle\, e_2 = (0, 8, -2, -5) - \frac{25}{5}\cdot\frac{1}{5}(4, 2, -2, -1) - \frac{24}{\sqrt{24}}\frac{1}{\sqrt{24}}(-2, 0, -2, -4)\\
&= (-2, 6, 2, 0),\\
e_3 &= \frac{f_3}{\|f_3\|} = \frac{1}{\sqrt{44}}(-2, 6, 2, 0).
\end{aligned}$$

(b) **(Fortsetzung von Beispiel 5.1 (c2))**. Die Basis $B = \{1, x, x^2\}$ im Raum $\mathbb{R}_2[x]$ soll orthonormalisiert werden. Das SKP ist gegeben durch $\langle p, q\rangle = \int_{-1}^{1} p(x)q(x)\,\mathrm{d}x$. Man beachte, dass $\|1\| = \sqrt{2}$, $\|x\| = \sqrt{2/3}$.

$$\begin{aligned}
e_1(x) &= \frac{b_1}{\|b_1\|} = \frac{1}{\sqrt{2}},\\
f_2(x) &= b_2(x) - \left(\int_{-1}^{1} x\frac{1}{2}\,\mathrm{d}x\right) e_1(x) = b_2(x) = x,\\
e_2(x) &= \frac{f_2(x)}{\|f_2\|} = \sqrt{\frac{3}{2}}x,\\
f_3(x) &= b_3(x) - \left(\int_{-1}^{1} x^2\frac{1}{\sqrt{2}}\,\mathrm{d}x\right)\frac{1}{\sqrt{2}} - \left(\int_{-1}^{1} x^3\,\mathrm{d}x\right)\frac{3}{2}x = x^2 - \frac{1}{3},\\
e_3(x) &= \sqrt{\frac{5}{2}}\left(\frac{3}{2}x^2 - \frac{1}{2}\right).
\end{aligned}$$

Dies sind die ersten drei Legendre-Polynome, die später bei den Differentialgleichungen eine Rolle spielen werden, siehe auch `http://mathworld.wolfram.com/LegendrePolynomial.html`

5.4 Längenerhaltende Abbildungen

5.4.1 Orthogonale und unitäre Abbildungen

Besonders wichtig sind solche linearen Abbildungen, die Längen und Winkel erhalten.

Definition 5.6 Es sei V ein euklidischer bzw. unitärer Raum und $T \in \mathrm{L}(V)$ ein Endomorphismus. Im folgenden sei V stets endlich-dimensional.
Dann heißt T *orthogonal* bzw. *unitär* wenn für alle $v, w \in V$ gilt:

$$\langle T(v), T(w)\rangle = \langle v, w\rangle . \tag{5.11}$$

Lemma 5.8 *Es sei V ein Raum mit SKP und $T \in \mathrm{L}(V)$ sei orthogonal bzw. unitär. dann gilt*

(a) $\|T(x)\| = \|x\|, \quad x \in V.$
(b) $x \perp y$ *impliziert* $T(x) \perp T(y)$.
(c) T *ist bijektiv und* T^{-1} *ist ebenfalls orthogonal bzw. unitär.*
(d) Die Menge $\mathrm{O}(V)$ *der orthogonalen Endomorphismen eines euklidischen Vektorraumes bzw. die Menge* $\mathrm{U}(V)$ *der unitären Endomorphismen eines unitären Raumes bilden eine Gruppe bezüglich der Hintereinanderausführung (Komposition).*

Beweis. (a), (b) folgen sofort aus der Definition; nämlich

$$\|T(x)\|^2 = \langle T(x), T(x)\rangle = \langle x, x\rangle = \|x\|^2$$

und aus $\langle x, y\rangle = 0$ folgt $\langle T(x), T(y)\rangle = \langle x, y\rangle = 0$.
(c) Aus (a) folgt die Injektivität von T, denn sei $T(x) = 0$, so $\|T(x)\| = 0$, also $\|x\| = 0$ und damit $x = 0$. Somit ist $\mathrm{Ker}\, T = \{0\}$, nach Lemma 3.4 folgt hieraus die Injektivität von T. Da V endlichdimensional ist, folgt nach Satz 3.8, dass T auch surjektiv ist, also bijektiv. Setzt man in (5.11) $v = T^{-1}(x)$ und $w = T^{-1}(y)$, also $x = T(v)$, $y = T(w)$, so hat man für alle $x, y \in V$

$$\langle x, y\rangle = \langle v, w\rangle = \langle T^{-1}(x), T^{-1}(y)\rangle .$$

Somit ist auch T^{-1} orthogonal bzw. unitär.
(d) Wir beschränken uns auf den reellen Fall. Analog zum Unterraumkriterium Lemma 2.1 gibt es das *Untergruppenkriterium*:

Eine nichtleere Teilmenge $H \subset G$ einer Gruppe G ist Untergruppe, falls gilt
(1) $\forall\, h \in H: h^{-1} \in H$.
(2) $\forall\, g, h \in H: gh \in H$.

In (c) haben wir bereits gesehen, dass $\mathrm{O}(V) \subset \mathrm{GL}(V)$ und dass (1) erfüllt ist, denn mit $T \in \mathrm{O}(V)$ ist auch $T^{-1} \in \mathrm{O}(V)$. Wir zeigen, dass (2) erfüllt ist, dass also mit $T, S \in \mathrm{O}(V)$ auch $S \circ T \in \mathrm{O}(V)$. In der Tat ist für alle $x, y \in V$

$$\langle T \circ S(x), T \circ S(y)\rangle = \langle T(S(x)), T(S(y))\rangle = \langle S(x), S(y)\rangle = \langle x, y\rangle .$$

Somit ist auch $T \circ S$ orthogonal. Wir nennen $\mathrm{O}(V)$ bzw. $\mathrm{U}(V)$ die *orthogonale Gruppe* bzw. *unitäre Gruppe* über V. ■

Beispiel 5.7 $\mathrm{O}(\mathbb{R}^1) = \{-1, 1\}$, $\mathrm{U}(\mathbb{C}^1) = \{z \mid |z| = 1\}$. In der Tat, wenn T orthogonal oder unitär ist, so gilt

$$\|T(x)\| = |T x| \stackrel{!}{=} \|x\| = |x|, \qquad \forall x.$$

Insbesondere gilt dies für $x = 1$, also $|T| = |1| = 1$. Im reellen Fall bleiben nur $T = \pm 1$, im komplexen Fall bleiben alle Elemente des Einheitskreises übrig.

5.4.2 Orthogonale und unitäre Matrizen

In Definition 3.7 hatten wir bereits die Mengen $\mathrm{O}(n)$ bzw. $\mathrm{U}(n)$ der orthogonalen bzw. unitären Matrizen definiert:

$$\mathrm{O}(n) = \{A \in \mathbb{R}^{n\times n} \mid A^\top A = AA^\top = I_n\}, \quad \mathrm{U}(n) = \{A \in \mathbb{C}^{n\times n} \mid A^* A = AA^* = I_n\},$$

wobei $A^* = \overline{A^\top}$ die adjungierte Matrix zu A ist.
Die Menge der orthogonalen bzw. unitären Matrizen mit Determinante gleich 1 bezeichnen wir mit $\mathrm{SO}(n)$ bzw. $\mathrm{SU}(n)$.

Lemma 5.9 *Für eine Matrix $A \in \mathbb{K}^{n\times n}$ sind die folgenden Bedingungen gleichwertig:*
(a) *A ist orthogonal bzw. unitär.*
(b) *Die Spalten von A bilden eine Orthonormalbasis von $\mathbb{K}^n$ (bezüglich des Standardskalarproduktes).*
(c) *Die Zeilen von A bilden eine Orthonormalbasis in $\mathbb{K}^n$.*
(d) *T_A ist orthogonal bzw. unitär.*

Beweis. (a) $\longleftrightarrow$ (b). Wir führen den Beweis nur im komplexen Fall. Sei $A = (a_{ij})_{i,j=1\ldots n}$ unitär, das heißt

$$\delta_{kl} = (A^* A)_{kl} = \sum_{i=1}^{n} (A^*)_{ki} a_{il} = \sum_{i=1}^{n} \overline{a_{ik}} a_{il} = \langle a_l, a_k \rangle,$$

wobei a_l die $l$te Spalte der Matrix A bezeichne. Also ist $\{a_1, \ldots, a_n\}$ eine ONB in $\mathbb{C}^n$. Die Umkehrung folgt analog.
(a) $\longleftrightarrow$ (c) folgt aus $AA^* = I_n$.
(a) $\longleftrightarrow (d)$. Wir betrachten nur den reellen Fall. Es sei $A = (a_{ij})$ die Matrix von T bezüglich $B = \{b_1, \ldots, b_n\}$. Wegen $T(b_j) = \sum_{i=1}^n a_{ij} b_i$ ist T genau dann orthogonal, wenn für alle $j, k = 1, \ldots, n$ gilt

$$\begin{aligned}\delta_{jk} = \langle b_j, b_k \rangle = \langle T(b_j), T(b_k) \rangle &= \left\langle \sum_{i=1}^{n} a_{ij} b_i, \sum_{l=1}^{n} a_{lk} b_l \right\rangle = \sum_{i,l=1}^{n} a_{ij} a_{lk} \delta_{il} \\ &= \sum_{i=1}^{n} a_{ij} a_{ik} = (A^\top A)_{jk}.\end{aligned}$$

dies ist aber äquivalent zu $A^\top A = I_n$ bzw. zur Orthogonalität von A. ■

5.5 Orthogonalzerlegung und Projektion

In diesem Abschnitt sei V stets ein Raum mit Skalarprodukt mit $\dim V = n$.

Definition 5.7 Es sei $M \subset V$ eine Teilmenge von V. Dann heißt

$$M^{\perp} = \{x \in V \mid \langle x\,,\,m\rangle = 0\ \forall\, m \in M\}$$

das *orthogonale Komplement* von M in V.

Insbesondere gilt $V^{\perp} = \{0\}$ und $\{0\}^{\perp} = V$.

Bemerkung 5.6 (Eigenschaften) (a) $M^{\perp}$ ist ein linearer Teilraum von V und $(\operatorname{lin} M)^{\perp} = M^{\perp}$.
(b) $M \subset N \Longrightarrow M^{\perp} \supset N^{\perp}$.
Sind U und W Teilräume von V, so gilt
(c) $(U + W)^{\perp} = U^{\perp} \cap W^{\perp}$.
(d) $U \cap U^{\perp} = \{0\}$.
(e) $(U^{\perp})^{\perp} = U$.

Satz 5.10 (Projektionssatz) *Es sei U ein linearer Teilraum von V. Dann gilt*

$$V = U \oplus U^{\perp},$$

mit anderen Worten, für jeden Vektor $v \in V$ gibt es eindeutig bestimmte Vektoren $v_1 \in U$ und $v_2 \in U^{\perp}$ mit $v = v_1 + v_2$.
v_1 heißt Projektion *des Vektors V auf U, $P_U(v) = v_1$.*

Beweis. Existenz der Zerlegung. Es sei $B = \{e_1, \dots, e_k\}$ eine ONB von U. Zu gegebenem $v \in V$ setzten wir

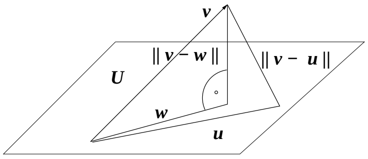
$$v_1 = \sum_{i=1}^{k} \langle v\,,\,e_i\rangle\, e_i, \quad v_2 = v - v_1. \tag{5.12}$$

dann ist $v_1 \in U$ und für alle $j = 1, \dots, k$ gilt

$$\langle v - v_1\,,\,e_j\rangle = \langle v\,,\,e_j\rangle - \left\langle \sum_{i=1}^{k} \langle v\,,\,e_i\rangle\, e_i\,,\,e_j \right\rangle = \langle v\,,\,e_j\rangle - \sum_{i=1}^{k} \langle v\,,\,e_i\rangle \langle e_i\,,\,e_j\rangle = \langle v\,,\,e_j\rangle - \langle v\,,\,e_j\rangle = 0.$$

Somit steht v_2 auf allen Basisvektoren $e_1, \dots, e_k$ von U senkrecht, also ist $v_2 \in U^{\perp}$.
Eindeutigkeit der Zerlegung. Seien $v = v_1 + v_2 = v_1' + v_2'$ Zerlegungen von v mit $v_1, v_1' \in U$, $v_2, v_2' \in U^{\perp}$. Dann gilt $v_1 - v_1' = v_2' - v_2 \in U \cap U^{\perp} = \{0\}$ (nach Bemerkung,5.6 (d)). Somit ist $v_1 = v_1'$ und $v_2 = v_2'$; die Zerlegung ist eindeutig. ∎

Die Projektion $P_U \colon V \to V$ ist wegen der Linearität des SKP in der ersten Komponente und wegen (5.12) eine **lineare** Abbildung. Es gilt außerdem $P_U + P_{U^{\perp}} = \mathrm{id}_V$, $\operatorname{rg} P_U = \dim U$ und daher $\dim U + \dim U^{\perp} = n$.



Satz 5.11 (Satz vom kleinsten Abstand) *Es sei U ein Unterraum von V und $v \in V$. Dann sind äquivalent:*

(i) $w = P_U(v)$.
(ii) $w \in U$ *und* $\min_{u \in U} \|v - u\| = \|v - w\|$.

Beweis. Wir setzen $v_1 = P_U(v)$.
(i) → (ii). Für alle $u \in U$ gilt $v - v_1 \perp u - v_1$, da $v - v_1 \in U^\perp$ und $u - v_1 \in U$. Also gilt nach Pythagoras

$$\|v - u\|^2 = \|v - v_1 + v_1 - u\|^2 = \|v - v_1\|^2 + \|v_1 - u\|^2 \geq \|v - v_1\|^2 .$$

Da dies für alle $u \in U$ gilt, ist $\|v - v_1\|^2$ eine untere Schranke für $\|v - u\|^2$, $u \in U$. Da $v_1 \in U$, gehört die untere Schranke $\|v - v_1\|^2$ zur Referenzmenge selbst dazu, ist also das Minimum:

$$\min_{u \in U} \|v - u\|^2 = \|v - v_1\|^2 .$$

Zieht man hier die Wurzel, folgt die Behauptung.
(ii) → (i). Das Minimum des Abstandes werde bei $w \in U$ angenommen. Also gilt

$$\|v - w\|^2 = \|v - v_1\|^2 + \|v_1 - w\|^2 \underset{\min}{\geq} \|v - w\|^2 + \|w - v_1\|^2$$

Subtrahiert man auf beiden Seiten $\|v - w\|^2$, so hat man $0 \geq \|w - v_1\|^2$ also $\|w - v_1\| = 0$ und damit $w = v_1$. ■

Anschaulich besagt der Satz vom kleinsten Abstand: Ist v_1 die Projektion von v auf U, dann hat der Vektor $v - v_1$ unter allen Vektoren $v - u$, $u \in U$ die kürzeste Länge. Ist umgekehrt $\|v - w\|$ minimal, so ist $v - w$ senkrecht auf U. Dieser Fakt wird bei Approximationen ausgenutzt, genauer, er bildet das Herzstück der „Kleinsten Quadratapproximation“. Angewandt wird dies bei der Fourierreihenentwicklung, Approximation stetiger Funktionen mit Polynomen und bei der Berechnung von Näherungslösungen für überbestimmte (inkonsistente) lineare GS.

Beispiel 5.8 Bestimmen Sie den Abstand des Vektors $v = (3, 4, -8, 12) \in \mathbb{R}^4$ vom Unterraum $U = \operatorname{lin}\{b_4\}^\perp = \{x \in \mathbb{R}^4 \mid x_1 + x_2 - x_3 - x_4 = 0\}$.
Im Beispiel 5.5 ist eine ONB von U angegeben, nämlich, $U = \operatorname{lin}\{b_1, b_2, b_3\}$, $U^\perp = \operatorname{lin}\{b_4\}$. Somit gilt $v_1 = P_U(v) = \frac{11}{2} b_1 + \frac{21}{2} b_2 + \frac{19}{2} b_3$ und damit $\|v - v_1\| = \left\|\frac{3}{2} b_4\right\| = \frac{3}{2}$.

5.6 Anhang: Die adjungierte Abbildung

Der Inhalt dieses Anhanges ist zur Ergänzung gedacht. Im Folgenden sei stets V ein endlichdimensionaler Raum mit SKP.

Definition 5.8 Es sei $T \in \mathrm{L}(V)$. Eine lineare Abbildung $T^* \in \mathrm{L}(V)$ heißt *adjungiert* zu T, falls für alle $v, w \in V$ gilt:

$$\langle T(v), w \rangle = \langle v, T^*(w) \rangle$$

Wenn es eine adjungierte Abbildung T^* gibt, so ist sie eindeutig bestimmt.

Satz 5.12 *Es sei $T \in \mathrm{L}(V)$ und $\{e_1, \ldots, e_n\}$ eine Orthonormalbasis in V, dann existiert der adjungierte Operator $T^* \in \mathrm{L}(V)$ und ist für $v \in V$ gegeben durch*

$$T^*(v) = \sum_{i=1}^{n} \langle v\,,\, T(e_i) \rangle\, e_i. \tag{5.13}$$

Beispiel 5.9 (a) Wir betrachten im euklidischen Raum $\mathbb{R}^3$ den Endomorphismus

$$T(x_1, x_2, x_3) = (x_1 - x_2, -x_1 + x_2 + 2x_3, x_2 + x_3) = \begin{pmatrix} 1 & -1 & 0 \\ -1 & 1 & 2 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

Ist T^* die adjungierte Abbildung zu T, so gilt für alle $x, y \in \mathbb{R}^3$

$$\begin{aligned} \langle T(x)\,,\, y \rangle &= \big\langle (x_1 - x_2, -x_1 + x_2 + 2x_3, x_2 + x_3)\,,\, (y_1, y_2, y_3) \big\rangle \\ &= x_1 y_1 - x_1 y_2 - x_2 y_1 + x_2 y_2 + x_2 y_3 + 2x_3 y_2 + x_3 y_3 \\ &= \big\langle (x_1, x_2, x_3)\,,\, (y_1 - y_2, -y_1 + y_2 + y_3, 2y_2 + y_3) \big\rangle. \end{aligned}$$

Also gilt

$$T^*(y_1, y_2, y_3) = (y_1 - y_2, -y_1 + y_2 + y_3, 2y_2 + y_3) = \begin{pmatrix} 1 & -1 & 0 \\ -1 & 1 & 1 \\ 0 & 2 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}.$$

(b) Es sei nun allgemein $A \in \mathbb{K}^{n \times n}$ und $T_A \in \mathrm{L}(\mathbb{K}^n)$ die zugehörige lineare Abbildung. Dann gilt

$$\boxed{(T_A)^* = T_{A^*} \quad (\mathbb{K} = \mathbb{C}), \qquad\qquad (T_A)^* = T_{A^\top} \quad (\mathbb{K} = \mathbb{R}).}$$

Wir betrachten dabei die Standardskalarprodukte in $\mathbb{K}^n$. Wir führen den Beweis im komplexen Fall. Für alle $x \in \mathbb{C}^n$, $y \in \mathbb{C}^m$ gilt

$$\langle A x\,,\, y \rangle = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij} x_j \right) \overline{y_i} = \sum_{j=1}^{n} x_j \sum_{i=1}^{m} \overline{\overline{a_{ij}} y_i} = \sum_{j=1}^{n} x_j \sum_{i=1}^{n} \overline{(A^*)_{ji}\, y_i} = \sum_{j=1}^{n} x_j\, \overline{(A^* y)_j} = \langle x\,,\, A^* y \rangle.$$

Im reellen Fall hat man natürlich $(T_A)^* = T_{A^\top}$.

Lemma 5.13 *Es seien $S, T \in \mathrm{L}(V)$. Dann gilt für alle $\lambda \in \mathbb{K}$*

1. $(S + T)^* = S^* + T^*$,
2. $(\lambda T)^* = \overline{\lambda}\, T^*$,
3. $(ST)^* = T^* S^*$,
4. $(S^*)^* = S$,
5. $\operatorname{tr} T^* = \overline{\operatorname{tr} T}$, $\quad \det T^* = \overline{\det T}$.
6. $\operatorname{Ker} T^* = (\operatorname{Im} T)^\perp$, *und* $\operatorname{Ker} T = (\operatorname{Im} T^*)^\perp$.
7. $\operatorname{rg} T^* = \operatorname{rg} T$.

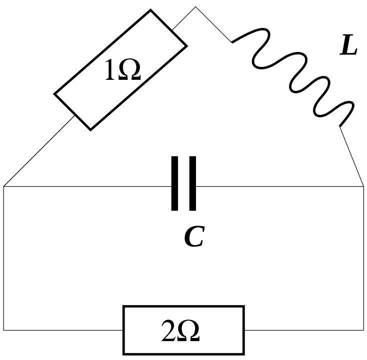
Definition 5.9 Eine lineare Abbildung T, $T \in \mathrm{L}(V)$, eines euklidischen oder unitären Raumes V heißt *selbstadjungiert*, falls $T^* = T$.

Nach Beispiel 5.9 ist klar, dass selbstadjungierte Abbildungen T durch symmetrische bzw. hermitesche Matrizen dargestellt werden: $A^\top = A$ bzw. $A^* = A$.
Selbstadjungierte lineare Endomorphismen spielen eine entscheidende Rolle in der Quantenmechanik — sie repräsentieren die beobachtbaren Größen wie Ort, Impuls, Drehimpuls, Energie.

Kapitel 6

Eigenwerte

Viele Prozesse in der Natur, wie etwa das Räuber-Beute-Modell, gekoppelte Federschwinger, elektrische Schwingkreise führen direkt auf Eigenwertprobleme.



Der gegebene Stromkreis aus zwei Widerständen $R_1 = 1\Omega$, $R_2 = 2\Omega$, einer Spule mit der Induktivität $L = 1\text{H}$ und mit einem Kondensator der Kapazität $C = \frac{1}{2}\text{F}$.
Dann genügen der Strom I durch die Spule und die Spannung U am Kondensator dem gewöhnlichen Differentialgleichungssystem

$$\frac{\mathrm{d}I}{\mathrm{d}t} = -I - U, \quad \frac{\mathrm{d}U}{\mathrm{d}t} = 2I - U.$$

Diese Modelle führen meistens auf dasselbe zentrale mathematische Problem: Zu einer gegebenen quadratischen $n \times n$-Matrix A finde man eine möglichst einfache ähnliche Matrix A', also $A' = S^{-1}AS$, $S \in \mathrm{GL}(n)$. Im optimalen Fall ist A' eine Diagonalmatrix. Wir werden aber sehen, dass dies nicht immer möglich ist.
Da in diesem Abschnitt nur Endomorphismen $T\colon V \to V$ betrachtet werden, die innerhalb *eines* Raumes wirken, schreiben wir ab hier stets $M_B(T)$ anstelle von $M_{B,B}(T)$ für die darstellende Matrix des Endomorphismus T in der Basis B von V.

6.1 Eigenwerte und Eigenvektoren

Es sei $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$.

Definition 6.1 Es sei $T \in \mathrm{L}(V)$ ein Endomorphismus.
Eine Zahl $\lambda \in \mathbb{K}$ heißt *Eigenwert* von T, wenn es einen von Null verschiedenen Vektor $x \in V$, $x \neq 0$, gibt mit $T(x) = \lambda x$. Jeder Vektor x, $x \neq 0$, mit $T(x) = \lambda x$ heißt *Eigenvektor* von T zum Eigenwert λ.

Man beachte, dass die Zahl 0 durchaus Eigenwert von T sein kann. Der Vektor 0 kann aber nie Eigenvektor sein. Das zentrale Problem ist die Existenz und die Vielfachheit von Eigenvektoren und Eigenwerten.

Beispiel 6.1 Es sei $T = 0$ die Nullabbbildung. Dann ist $\lambda = 0$ einziger Eigenwert, denn die Eigenwertgleichung $T(x) = \lambda x$ ist äquivalent zu $0 = \lambda x$ und diese ist nur für $\lambda = 0$ unter der Voraussetzung $x \neq 0$ erfüllt. Alle Vektoren (außer dem Nullvektor) sind Eigenvektoren, denn es gilt $T(x) = 0x$.
Die Abbildung $T = \mathrm{id}_V$ hat als einzigen Eigenwert $\lambda = 1$; alle Vektoren (außer 0) sind Eigenvektoren zu 1, denn $\mathrm{id}_V(x) = 1x$.

Als *Eigenwert* bzw. *Eigenvektor* einer quadratischen Matrix A bezeichnen wir Eigenwert und -vektor der linearen Abbildung T_A.

Beispiel 6.2 (a) Es sei $B = \{b_1, b_2, b_3\}$ eine Basis von V und $T \in \mathrm{L}(V)$ sei gegeben durch $T(x_1 b_1 + x_2 b_2 + x_3 b_3) = \lambda x_1 b_1 + \mu x_2 b_2 + \mu x_3 b_3$. Dann sind λ und μ Eigenwerte von T zu den Eigenvektoren b_1 bzw. b_2 und b_3, denn $T(b_1) = \lambda b_1$, $T(b_2) = \mu b_2$ und $T(b_3) = \mu b_3$ und in der Basis B hat T die Matrixdarstellung

$$M_B(T) = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \mu & 0 \\ 0 & 0 & \mu \end{pmatrix}.$$

(b) **Spiegelung.** $S\colon \mathbb{R}^3 \to \mathbb{R}^3$ bezeichne die Spiegelung des $\mathbb{R}^3$ an der $x_1 - x_2$-Ebene. Das heißt, alle Vektoren dieser Ebene bleiben unverändert (fix) und $(0,0,1)$ geht in $(0,0,-1)$ über. Somit ist $S(x,y,z) = (x,y,-z)$. Die Matrixdarstellung bezüglich der Standardbasis ist $A = \mathrm{diag}(1,1,-1)$. S hat zwei Eigenvektoren zum Eigenwert 1, nämlich e_1 und e_2 und einen Eigenvektor zu -1, nämlich e_3. S ist diagonal in der Standardbasis.
(c) $A = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$ hat nur den Eigenwert 2 und einen Eigenvektor $v = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$. Sei nämlich

$$Ax = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2x_1 + x_2 \\ 2x_2 \end{pmatrix} \stackrel{!}{=} \lambda \begin{pmatrix} x_1 \\ x_2 \end{pmatrix},$$

so hat man $2x_1 + x_2 = \lambda x_1$ und $2x_2 = \lambda x_2$. Wäre $\lambda \neq 2$, so wäre $x_2 = 0$ und damit $x_1 = 0$ kein Eigenvektor; also ist $\lambda = 2$ und $x_2 = 0$, x_1 beliebig.

6.2 Diagonalisierung

6.2.1 Eigenraum und geometrische Vielfachheit

Definition 6.2 Ein Endomorphismus $T \in \mathrm{L}(V)$ heißt *diagonalisierbar*, wenn es eine Basis $B = \{b_1, \ldots, b_n\}$ aus Eigenvektoren gibt. In diesem Fall gilt

$$M_B(T) = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} =: \mathrm{diag}(\lambda_1, \ldots, \lambda_n).$$

wobei $T(b_i) = \lambda_i b_i$, $i = 1, \ldots, n$, d. h. die λ_i sind die Eigenwerte von T zum Eigenvektor b_i.

Wir haben im obigen Beispiel gesehen: Die Spiegelung ist diagonalisierbar ebenso wie die komplexe Drehung $D_\varphi : \mathbb{C}^2 \to \mathbb{C}^2$. Die reelle Drehung $D_\varphi : \mathbb{R}^2 \to \mathbb{R}^2$, $\varphi \neq 0, \pi$, ist nicht diagonalisierbar, da es keine reellen Eigenwerte bzw. keine reellen Eigenvektoren gibt; die Matrix A aus (c) ist ebenfalls nicht diagonalisierbar, denn sie hat nur einen Eigenvektor.

Lemma 6.1 (/Definition) *Ein Vektor $v \in V$, $v \neq 0$, ist genau dann Eigenvektor zu $T \in \mathrm{L}(V)$ zum Eigenwert λ, wenn*

$$v \in \mathrm{Ker}(T - \lambda \mathrm{id}).$$

Der Raum $E_\lambda := \mathrm{Ker}(T - \lambda \mathrm{id})$ heißt Eigenraum *von T zum Eigenwert λ. Die Dimension von E_λ*

$$\dim E_\lambda = \dim \mathrm{Ker}(T - \lambda \mathrm{id}) = \mathrm{def}(T - \lambda \mathrm{id})$$

bezeichnen wir als geometrische Vielfachheit *des Eigenwertes λ.*

Beweis. Die Gleichung $T(v) = \lambda v$ lässt sich äquivalent umformen zu $T(v) - \lambda v = 0$ und mit $v = \mathrm{id}(v)$ ist dies äquivalent zu $T(v) - \lambda \mathrm{id}(v) = 0$ bzw. zu $(T - \lambda \mathrm{id})(v) = 0$ oder $v \in \mathrm{Ker}(T - \lambda \mathrm{id})$. Somit bilden alle Eigenvektoren zu λ zusammen mit dem Nullvektor einen Unterraum von V. ∎

Dass die Eigenräume E_λ und E_μ zu verschiedenen Eigenwerten $\lambda \neq \mu$ nur den Nullvektor gemeinsam haben können ist klar, denn aus $T(v) = \lambda v = \mu v$ folgt $(\lambda - \mu)v = 0$ bzw. $v = 0$. Es gilt aber noch viel mehr, die Eigenvektoren zu verschiedenen Eigenwerten sind sogar linear unabhängig.

Lemma 6.2 *Es seien $v_1, \ldots, v_r$ Eigenvektoren zu den paarweise verschiedenen Eigenwerten $\lambda_1, \ldots, \lambda_r$ von T, $\lambda_i \neq \lambda_j$.*
Dann ist $\{v_1, \ldots, v_r\}$ linear unabhängig.

Beweis. durch vollständige Induktion über r. Für $r = 1$ ist die Behauptung richtig, da $v_1 \neq 0$ linear unabhängig ist. Sei nun die Behauptung gültig für k Eigenvektoren $v_1, \ldots, v_k$. Wir zeigen die Behauptung für $r = k + 1$. Seien dazu $v_1, \ldots, v_{k+1}$ Eigenvektoren zu den paarweise verschiedenen Eigenwerten λ_1, bis λ_{k+1}. Angenommen $\alpha_1 v_1 + \cdots + \alpha_{k+1} v_{k+1} = 0$. Durch Multiplikation mit λ_{k+1} bzw. durch Anwenden von T erhält man die beiden Gleichungen:

$$\begin{aligned} \alpha_1 \lambda_{k+1} v_1 + \alpha_2 \lambda_{k+1} v_2 + \cdots + \alpha_{k+1} \lambda_{k+1} v_{k+1} &= 0, \\ \alpha_1 \lambda_1 v_1 + \alpha_2 \lambda_2 v_2 + \cdots + \alpha_{k+1} \lambda_{k+1} v_{k+1} &= 0. \end{aligned}$$

Durch Subtraktion folgt daraus die Gleichung

$$\alpha_1(\lambda_1 - \lambda_{k+1}) v_1 + \cdots + \alpha_k(\lambda_k - \lambda_{k+1}) v_k = 0,$$

die v_{k+1} gar nicht mehr enthält. Nach Induktionsvoraussetzung sind aber die $v_1, \ldots, v_k$ linear unabhängig, also

$$\alpha_1(\lambda_1 - \lambda_{k+1}) = \cdots = \alpha_k(\lambda_k - \lambda_{k+1}) = 0.$$

Da die Eigenwerte aber paarweise verschieden vorausgesetzt waren folgt hieraus $\alpha_1 = \cdots = \alpha_k = 0$, also auch $\alpha_{k+1} v_{k+1} = 0$ und somit $\alpha_{k+1} = 0$. ■

Somit ist $T\colon V \to V$, $\dim V = n$, diagonalisierbar, wenn die Abbildung n paarweise verschiedene Eigenwerte hat. Dies ist eine hinreichende aber keineswegs notwendige Bedingung, wie etwa $T = \mathrm{id}$ zeigt.

Folgerung 6.3 *Es sei $T \in \mathrm{L}(V)$, $\dim_{\mathbb{K}} V = n$ und $\lambda_1, \ldots, \lambda_r$ seien seine verschiedenen Eigenwerte mit den geometrischen Vielfachheiten $n_1, \ldots, n_r$. Ferner sei B_i, $i = 1, \ldots, r$ jeweils eine Basis des Eigenraumes E_{λ_i}.*
Dann ist auch $B = B_1 \cup B_2 \cup \cdots \cup B_r$ linear unabhängig. Insbesondere ist die Summe der geometrischen Vielfachheiten kleiner gleich n

$$n_1 + n_2 + \cdots + n_r \leq n$$

und T ist genau dann diagonalisierbar, wenn die Summe der geometrischen Vielfachheiten gleich n ist.

Beweis. Sei $B_i := \{v_1^{(i)}, \ldots, v_{n_i}^{(i)}\}$, $i = 1, \ldots, r$ die Basis des Eigenraumes E_{λ_i}. Sei

$$\sum_{i=1}^{r} \sum_{k=1}^{n_i} \alpha_k^{(i)} v_k^{(i)} = 0.$$

Dann sind nach obigem Lemma die Vektoren $\sum_{k=1}^{n_i} \alpha_k^{(i)} v_k^{(i)} \in E\lambda_i$ gleich dem Nullvektor für alle $i = 1, \ldots, r$. Weil die $v_1^{(i)}, \ldots, v_{n_i}^{(i)}$ linear unabhängig sind, verschwinden alle Koeffizienten $\alpha_k^{(i)}$. Durch Aneinanderreihung der Basen der Eigenräume entsteht also eine linear unabhängige Menge aus $n_1 + n_2 + \cdots + n_r$ Vektoren. Also gilt $n_1 + \cdots + n_r \leq n$. Im Falle der Gleichheit hat V eine Basis, die nur aus Eigenvektoren besteht und ist somit diagonalisierbar.
Ist umgekehrt T diagonalisierbar und ist m_i die Anzahl der Eigenvektoren zu λ_i in der Basis aus Eigenvektoren, so ist sicher $m_i \leq n_i$. Also gilt

$$n = m_1 + \cdots + m_r \leq n_1 + \cdots + n_r \leq n.$$

Dann muss aber in jeder Abschätzung die Gleichheit gelten, also ist die Summe der geometrischen Vielfachheiten gleich n und alle Eigenvektoren werden für die Diagonalisierung benötigt ($m_i = n_i$). ■

Wir sehen hier schon, wie man zu einer Basis aus Eigenvektoren kommt: Zunächst bestimme man alle λ_i, für die $T - \lambda_i \mathrm{id}$ nicht injektiv ist — das sind die Eigenwerte von T. Dann bestimme man zu jedem dieser λ_i eine Basis von $E_{\lambda_i} = \mathrm{Ker}(T - \lambda_i \mathrm{id})$. Wenn T überhaupt diagonalisierbar ist, so muss die Aneinanderreihung dieser Basen eine Basis von V sein.

6.2.2 Das charakteristische Polynom

Wir geben hier ein Verfahren an, die Eigenwerte eines Endomorphismus $T \in \mathrm{L}(V)$ zu bestimmen.

In Lemma 6.1 haben wir bereits gesehen, dass λ genau dann ein Eigenwert von $T \in \mathrm{L}(V)$ ist, wenn der Raum $\mathrm{Ker}(T-\lambda\mathrm{id})$ nicht nur aus dem Nullvektor besteht. Äquivalent dazu ist, dass das $T-\lambda\mathrm{id}$ nicht injektiv ist bzw. nach Lemma 3.11 und Satz 4.1 (D9) es gilt $\det(T-\lambda\mathrm{id})=0$.

Definition 6.3 Es sei $T \in \mathrm{L}(V)$. Dann heißt

$$\chi_T(\lambda) = \det(\lambda\mathrm{id} - T) \tag{6.1}$$

das *charakteristische Polynom* von T.

Man beachte, dass bei der Berechnung der Determinante auf der rechten Seite von (6.1) eine Matrixdarstellung $A \in \mathbb{K}^{n\times n}$, $A = M_B(T)$, bezüglich einer Basis von V gebraucht wird, $\det(\lambda\mathrm{id} - T) = \det(\lambda I_n - A)$. Wir zeigen, dass das charakteristische Polynom nicht von der Wahl der Basis abhängt. Sei etwa $A' = M_C(T)$ eine weitere Matrixdarstellung von T, so existiert eine invertierbare Matrix S mit $A' = SAS^{-1}$. Dann gilt nach dem Produktsatz für Determinanten:

$$\det(A'-\lambda I_n) = \det(SAS^{-1} - \lambda SS^{-1}) = \det(S(A-\lambda I_n)S^{-1}) = \det S \det(A-\lambda I_n)\det(S^{-1})$$

Wegen $\det(S^{-1}) = 1/\det S$ kürzen sich diese beiden Faktoren und es ist

$$\chi_A(\lambda) = \det(A-\lambda I_n) = \det(A'-\lambda I_n) = \chi_{A'}(\lambda).$$

Satz 6.4 *Es sei $T \in \mathrm{L}(V)$, $\dim V = n$. Dann gilt*
(a) $\mathrm{grad}\,\chi_T = n$
(b) *μ ist Eigenwert von T genau dann, wenn μ Nullstelle des charakteristischen Polynoms von T ist, also wenn $\chi_T(\mu)=0$.*
(c) *Es gilt*

$$\chi_T(\lambda) = \lambda^n - \mathrm{tr}\,T\lambda^{n-1} + a_{n-2}\lambda^{n-2} + \cdots + a_1\lambda + (-1)^n \det T,$$

das heißt, der höchste Koeffizient des charakteristischen Polynoms ist 1*, der zweithöchste ist* $-\mathrm{tr}\,T$ *und das Absolutglied ist* $(-1)^n \det T$.
(d) *Ähnliche Matrizen A und $A' = S^{-1}AS$ besitzen dasselbe charakteristische Polynom $\chi_A = \chi_{A'}$.*

Beweis. (b) folgt aus der Bemerkung am Anfang dieses Abschnitts: μ ist Eigenwert von T genau dann, wenn $\det(\mu\,\mathrm{id} - T) = 0$, das heißt, μ ist Nullstelle des charakteristischen Polynoms. (a), (c). Es sei $A = (a_{ij}) = M_B(T)$ eine T darstellende Matrix, also $\det T = \det A$, $\mathrm{tr}\,T = \mathrm{tr}\,A$ und

$$\chi_T(\lambda) = \det(\lambda\mathrm{id} - T) = \det(\lambda I_n - A) = \chi_A(\lambda) = \begin{vmatrix} \lambda - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & & \vdots \\ -a_{n1} & -a_{n2} & \cdots & \lambda - a_{nn} \end{vmatrix}.$$

Nach Leibnizdefinition der Determinante erhalten wir das charakteristische Polynom von T (bzw. von A) indem wir alle möglichen Produkte von n Faktoren dieser Matrix bilden, wobei aus jeder Zeile und jeder Spalte immer genau ein Faktor genommen wird. Spaltet man den zur identischen Permutation $\sigma = \mathrm{id} \in {}_n$ gehörigen Summanden ab, so hat man

$$\chi_A(\lambda) = (\lambda - a_{11})(\lambda - a_{22}) \cdots (\lambda - a_{nn}) + \sum_{\sigma \in {}_n \setminus \{\mathrm{id}\}} \mathrm{sign}(\sigma) b_{1\sigma(1)} \cdots b_{n\sigma(n)},$$

wobei $B = (b_{ij}) = \lambda I_n - A$. Die Summe $\sum_{\sigma \in {}_n \setminus \{\mathrm{id}\}}$ liefert ein Polynom $q(\lambda)$ vom Grad, kleiner gleich $n-2$, da für jede Permutation, die ungleich der Identität ist, mindestens zwei Außerdiagonalelemente $b_{kl} = -a_{kl}$ und $b_{rs} = -a_{rs}$, $k \neq l$, $r \neq s$ auftreten. Diese beiden Faktoren enthalten aber beide kein λ, sodass höchstens $n-2$ Faktoren λ im Produkt $b_{1\sigma(1)} \cdots b_{n\sigma(n)}$ auftreten. Durch Ausmultiplizieren von $(\lambda - a_{11})(\lambda - a_{22}) \cdots (\lambda - a_{nn})$ erhalten wir:

$$\chi_A(\lambda) = \lambda^n - a_{11}\lambda^{n-1} - a_{22}\lambda^{n-1} - \cdots - a_{nn}\lambda^{n-1} + q_1(\lambda), \quad \mathrm{grad}\, q_1 \leq n-2.$$

Also ist

$$\chi_A(\lambda) = \lambda^n - (a_{11} + a_{22} + \cdots + a_{nn})\lambda^{n-1} + q_1(\lambda) = \lambda^n - \mathrm{tr}\, A \lambda^{n-1} + q_1(\lambda).$$

Das Absolutglied a_0 eines Polynoms erhält man durch Einsetzen von $x = 0$, also $a_0 = p(0)$. Das Absolutglied a_0 des charakteristischen Polynoms ist also gleich

$$a_0 = \chi_T(0) = \det(0 - T) = \det(-T) = (-1)^n \det T.$$

(d) Für $A' = S^{-1}AS$ erhält man

$$\begin{aligned}\chi_{A'}(\lambda) &= \det(\lambda S^{-1} I_n S - S^{-1}AS) = \det\left(S^{-1}(\lambda I_n - A)S\right) \\ &= \det S^{-1} \det(\lambda I_n - A) \det S = \frac{1}{\det S} \chi_A(\lambda) \det S = \chi_A(\lambda).\end{aligned}$$

■

Bemerkung 6.1 (1) Ist α eine Nullstelle des Polynoms $p(x)$, so gibt es ein Polynom $q(x)$ mit $p(x) = (x-\alpha)q(x)$, es ist $\mathrm{grad}\, q = \mathrm{grad}\, p - 1$. Man sagt „$q$ entsteht aus p durch Abspalten des Linearfaktors $(x-\alpha)$“.

(2) Ist α eine Nullstelle von $p(x)$, so heißt α *k-fache Nullstelle*, falls es ein Polynom $q(x)$ gibt mit

$$p(x) = (x-\alpha)^k q(x), \quad q(\alpha) \neq 0.$$

(3) Der *Fundamentalsatz der Algebra* besagt, dass jedes Polynom $n$ten Grades genau n komplexe Nullstellen, gezählt in ihrer Vielfachheit, besitzt.

Es ist klar, dass die Summe aller Vielfachheiten aller Nullstellen von p gleich n ist. So hat zum Beispiel das Polynom $p(x) = (x^2+1)(x-1)^2(x+2)$ die einfachen Nullstellen i, $-$i und -2 und die doppelte Nullstelle 1.

(4) Ein Eigenwert μ von T hat die *algebraische Vielfachheit* k, wenn μ eine k-fache Nullstelle des charakteristischen Polynoms ist.

Sind m_i die algebraischen Vielfachheiten der Eigenwerte λ_i von T, so lautet das charakteristische Polynom von T

$$\chi_T(\lambda) = (\lambda-\lambda_1)^{m_1}(\lambda-\lambda_2)^{m_2}\cdots(\lambda-\lambda_r)^{m_r}.$$

Es gilt $m_1+m_2+\cdots+m_r = n = \operatorname{grad}\chi_T$.

Folgerung 6.5 *Es seien $\lambda_1,\dots,\lambda_n \in \mathbb{C}$ die Eigenwerte von T, gezählt in ihrer Vielfachheit. Dann gilt*
(a) $\operatorname{tr} T = \lambda_1+\lambda_2+\cdots+\lambda_n$.
(b) $\det T = \lambda_1\lambda_2\cdots\lambda_n$.

Beweis. Nach dem Satz gilt, dass die Eigenwerte die Nullstellen von $\chi_T(x)$ sind. Daher kann man das charakteristische Polynom auch als Produkt von Linearfaktoren aufschreiben:

$$\chi_T(x) = (x-\lambda_1)(x-\lambda_2)\cdots(x-\lambda_n).$$

(a) Setzt man $x=0$ ein, so hat man nach dem Satz

$$\chi_T(0) = (-1)^n \det T = (0-\lambda_1)(0-\lambda_2)\cdots(0-\lambda_n) = (-1)^n\lambda_1\lambda_2\cdots\lambda_n.$$

Multipliziert man das Produkt auf der rechten Seite aus, so hat man

$$\chi_T(x) = x^n - (\lambda_1+\lambda_2+\cdots+\lambda_n)x^{n-1}+\cdots+a_1x+(-1)^n\lambda_1\lambda_2\cdots\lambda_n. \tag{6.2}$$

Der Koeffizientenvergleich von (6.2) mit Satz 6.4 (c) liefert die Behauptung. ∎

Beispiel 6.3 (a) $\chi_{\mathrm{id}}(\lambda) = \det\operatorname{diag}(\lambda-1,\lambda-1,\dots,\lambda-1) = (\lambda-1)^n$; $\chi_0(\lambda) = \lambda^n$.
(b) $\chi_{D_\varphi}(\lambda) = \begin{vmatrix} \lambda-\cos\varphi & \sin\varphi \\ -\sin\varphi & \lambda-\cos\varphi \end{vmatrix} = \lambda^2 - 2\cos\varphi\,\lambda+1$. Die Nullstellen dieses Polynoms sind $\lambda_{1,2} = \cos\varphi \pm \mathrm{i}\sin\varphi$.
(c) Ist

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

so liefert die Entwicklung nach der ersten Spalte

$$\begin{aligned}\chi_A(\lambda) &= \begin{vmatrix} \lambda & 1 & -1 \\ 3 & \lambda+2 & -3 \\ 2 & 2 & \lambda-3 \end{vmatrix} = \lambda\begin{vmatrix} \lambda+2 & -3 \\ 2 & \lambda-3 \end{vmatrix} - 3\begin{vmatrix} 1 & -1 \\ 2 & \lambda-3 \end{vmatrix} + 2\begin{vmatrix} 1 & -1 \\ \lambda+2 & -3 \end{vmatrix}\\ &= \lambda(\lambda^2-\lambda-6+6) - 3(\lambda-3+2) + 2(\lambda+2-3)\\ &= \lambda^3-\lambda^2-3\lambda+3+2\lambda-2 = \lambda^3-\lambda^2-\lambda+1 = (\lambda-1)^2(\lambda+1).\end{aligned}$$

6.2.3 Diagonalisierung beliebiger Endomorphismen

Satz 6.6 *Es sei λ ein Eigenwert von T.*

(a) Dann ist seine geometrische Vielfachheit kleiner oder gleich seiner algebraischen Vielfachheit.

(b) T ist genau dann diagonalisierbar, wenn für jeden Eigenwert die geometrische und die algebraische Vielfachheit überein stimmen.

Beweis. (a) Es sei $\dim E_\lambda = k$ die geometrische Vielfachheit des Eigenwertes λ von T. Dann existieren genau k linear unabhängige Eigenvektoren zum Eigenwert λ. Ergänzt man diese zu einer Basis B von V, dann hat T in dieser Basis eine Matrixdarstellung

$$A = M_B(T) = \begin{pmatrix} \lambda & 0 & \cdots & 0 & * \\ 0 & \lambda & \cdots & 0 & * \\ \vdots & & & \lambda & * \\ 0 & \cdots & & 0 & \\ \vdots & & & & B \\ 0 & \cdots & & 0 & \end{pmatrix}.$$

Dabei steht links oben die quadratische Diagonalmatrix λI_k, darunter eine Nullmatrix und rechts unten eine quadratische Matrix $B \in \mathbb{K}^{(n-k)\times(n-k)}$. Nach Satz 4.1 (D8) gilt

$$\chi_T(x) = \det(\operatorname{diag}(x-\lambda, \cdots, x-\lambda)) \det(xI_{n-k} - B) = (x-\lambda)^k \chi_B(x),$$

Damit ist die algebraische Vielfachheit von T mindestens gleich k.

(b) Es sei n_i die geometrische und m_i die algebraische Vielfachheit von λ_i. Nach (a) ist also $n_i \le m_i$ für alle i. Angenommen, T ist diagonalisierbar, dann gilt nach Folgerung 6.3,

$$n = n_1 + \cdots + n_r \le m_1 + \cdots + m_r = n,$$

somit ist $n_i = m_i$ für alle i. Gilt umgekehrt diese Gleichheit, so ist die Summe der geometrischen Vielfachheiten gleich n und T ist diagonalisierbar. ■

Wie schon am Ende von Abschnitt 6.2.1 festgestellt sieht das Verfahren zur Diagonalisierung eines Endomorphismus $T \in \mathrm{L}(V)$, $\dim V = n$, wie folgt aus:

1. Auswahl einer Basis B in V und Bestimmung der Darstellungsmatrix $A = M_B(T)$.
2. Berechnung des charakteristischen Polynoms $\chi_T(\lambda)$.
3. Berechnung der Eigenwerte von T — die Nullstellen des charakteristischen Polynoms.
4. Berechnung des Eigenraumes $E_\lambda = \operatorname{Ker}(T - \lambda \mathrm{id})$ zu jedem Eigenwert λ. Dies erfolgt durch Bestimmung der Lösung des homogenen linearen $n \times n$ Gleichungssystems $(A - \lambda I_n)x = 0$. Stimmen für jeden Eigenwert λ die algebraische und geometrische Vielfachheit überein, so ist T diagonalisierbar, andernfalls nicht.
5. Die n Eigenvektoren bilden nun eine Basis E. Schreibt man deren Koordinatendarstellungen (zur Basis B) als Spalten einer Matrix S auf, so gilt

$$A = S \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} S^{-1} \quad \text{bzw.} \quad AS = SA',$$

mit der Diagonalmatrix $A' = M_E(T) = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$.

Beispiel 6.4 Es sei $T \in \mathrm{L}(\mathbb{R}^3)$ gegeben durch

$$T(x_1, x_2, x_3) = (-x_2 + x_3, -3x_1 - 2x_2 + 3x_3, -2x_1 - 2x_2 + 3x_3).$$

In der Standardbasis von $\mathbb{R}^3$ hat T die Matrixdarstellung $A = M_{B_3}(T) = \begin{pmatrix} 0 & -1 & 1 \\ -3 & -2 & 3 \\ -2 & -2 & 3 \end{pmatrix}$. In Beispiel 6.3 berechneten wir die Eigenwerte und erhielten $\lambda_1 = 1$ mit algebraischer Vielfachheit 2 und den einfachen Eigenwert $\lambda_2 = -1$. Wir bestimmen $E_{\lambda_1}(T) = E_1(T)$. Es ist

$$A - I_3 = \begin{pmatrix} -1 & -1 & 1 \\ -3 & -3 & 3 \\ -2 & -2 & 2 \end{pmatrix}$$

Diese Matrix hat Rang 1 und damit den Defekt 2. Geometrische und algebraische Vielfachheit stimmen hier überein. Es gibt zwei linear unabhängige Eigenvektoren, nämlich die Lösungen von $-x_1 - x_2 + x_3 = 0$, etwa $b_1 = (1, 0, 1)$ und $b_2 = (0, 1, 1)$.
Wir bestimmen den Eigenraum $E_{-1}(T)$. Es ist

$$A + I_3 = \begin{pmatrix} 1 & -1 & 1 \\ -3 & -1 & 3 \\ -2 & -2 & 4 \end{pmatrix}.$$

Addition des Dreifachen der ersten Zeile zur zweiten und des Doppelten der ersten Zeile zur der dritten Zeile liefert:

$$A + I_3 \sim \begin{pmatrix} 1 & -1 & 1 \\ 0 & -4 & 6 \\ 0 & -4 & 6 \end{pmatrix} \sim \begin{pmatrix} 1 & -1 & 1 \\ 0 & 1 & -\frac{3}{2} \\ 0 & 0 & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & -\frac{1}{2} \\ 0 & 1 & -\frac{3}{2} \\ 0 & 0 & 0 \end{pmatrix}.$$

Hiermit ist die reduzierte Zeilenstufenform erreicht. Die einparametrige Lösung mit frei wählbarem x_3 lautet $x_1 = \frac{1}{2}x_3$, $x_2 = \frac{3}{2}x_3$. Wählt man also $x_3 = 2$, so erhält man den Eigenvektor zu -1 und zwar $b_3 = (1,3,2)$. Somit lautet die Matrix der Basistransformation $S = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 3 \\ 1 & 1 & 2 \end{pmatrix}$.

In der Tat gilt dann

$$SA' = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 3 \\ 1 & 1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -3 \\ 1 & 1 & -2 \end{pmatrix} = \begin{pmatrix} 0 & -1 & 1 \\ -3 & -2 & 3 \\ -2 & -2 & 3 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 3 \\ 1 & 1 & 2 \end{pmatrix} = AS.$$

6.2.4 Die Jordansche Normalform

Gibt es weniger Eigenvektoren als die Dimension des Raumes angibt, so lässt sich der Endomorphismus nicht mehr auf Diagonalform bringen. Man kann aber erreichen, dass auf der Diagonalen die Eigenwerte von T in ihrer algebraischen Vielfachheit stehen und in der ersten Nebendiagonalen oberhalb der Diagonale Einsen und Nullen. Die Normalform setzt sich aus Jordan-Blöcken zusammen.

Definition 6.4 Es sei $T \in \mathrm{L}(V)$. Ein Teilraum $U \subset V$ heißt *invariant* (unter T), wenn $T(U) \subset U$.

Mit anderen Worten, invariante Teilräume werden *in sich selbst* abgebildet. In diesem Falle existiert die *Einschränkung* $T_1 \colon U \to U$ von T auf U, die mit T auf U übereinstimmt, $T_1(u) = T(u)$ für alle $u \in U$. Jeder Eigenraum E_λ ist invariant. Die Summe $U_1 + U_2$ und der Durchschnitt $U_1 \cap U_2$ invarianter Teilräume sind ebenfalls invariant.
Ohne Beweis erwähnen wir das folgende große Resultat.

Theorem 6.7 (Jordansche Normalform) *Es sei $T \in \mathrm{L}(V)$ ein Endomorphismus des komplexen Vektorraumes V mit $\dim V = n$. Wir setzten voraus, dass T genau k ($k \le n$) linear unabhängige Eigenvektoren*

$$e_1, f_1, \ldots, h_1$$

zu den Eigenwerten $\lambda_1, \lambda_2, \ldots, \lambda_k$ besitzt.
Dann existiert eine Basis von V aus k Gruppen von Vektoren

$$e_1, \ldots, e_p;\ f_1, \ldots, f_q;\ \ldots; h_1, \ldots, h_s,$$

derart, dass T die folgende Gestalt hat:

$$T(e_1) = \lambda_1 e_1, \quad T(e_2) = \lambda_1 e_2 + e_1, \quad \ldots \quad T(e_p) = \lambda_1 e_p + e_{p-1}, \tag{6.3}$$
$$T(f_1) = \lambda_2 f_2, \quad T(f_2) = \lambda_2 f_2 + f_1, \quad \ldots \quad T(f_q) = \lambda_2 f_q + f_{q-1}, \tag{6.4}$$
$$\ldots \quad \ldots \quad \ldots \tag{6.5}$$
$$T(h_1) = \lambda_k h_1, \quad T(h_2) = \lambda_k h_2 + h_1, \quad \ldots \quad T(h_s) = \lambda_k h_s + h_{s-1}. \tag{6.6}$$

Man sieht, dass die Basisvektoren jeder Gruppe durch T wieder in solche aus derselben Gruppe abgebildet werden. Das bedeutet, dass V in k invariante Unterräume zerfällt. Die Matrixdarstellung von T auf dem von $\{e_1, \dots, e_p\}$ aufgespannten Unterraum lautet dann (in genau dieser Basis):

$$A_1 = \begin{pmatrix} \lambda_1 & 1 & 0 & \cdots & 0 \\ 0 & \lambda_1 & 1 & \cdots & 0 \\ & \vdots & & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & \lambda_1 \end{pmatrix}$$

Diese Matrix bezeichnet man als Jordanblock. Wir zeigen, dass innerhalb einer Gruppe $V_e = \text{lin}\{e_1, \dots, e_p\}$ nur ein einziger Eigenvektor von T existiert und zwar zum Eigenwert λ_1. In der Tat, sei $v = \sum_{i=1}^p \alpha_i e_i$ ein Eigenvektor zum Eigenwert λ, so gilt: $T(\sum_{i=1}^p \alpha_i e_i) = \lambda \sum_{i=1}^p \alpha_i e_i$. Da die $\{e_i\}$ eine Basis bilden führt das mit (6.3) auf $\alpha_2 = \alpha_3 = \cdots = \alpha_p = 0$.
Die Darstellungsmatrix von T besteht also aus k Jordanblöcken:

$$A = \begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_k \end{pmatrix}.$$

Zwei Matrizen sind genau dann einander ähnlich, wenn sie (bis auf die Reihenfolge der Blöcke) dieselbe Jordansche Normalform besitzen.

6.3 Diagonalisierung von selbstadjungierten Endomorphismen

In diesem Abschnitt sei V stets ein Raum mit Skalarprodukt (unitär oder euklidisch). Wir wiederholen die Definition aus dem Anhang des vorigen Kapitels.

Definition 6.5 (a) Es sei $T \in \mathrm{L}(V)$. Die lineare Abbildung $T^* \in \mathrm{L}(V)$ heißt *adjungiert* zu T, falls für alle $v, w \in V$ gilt:

$$\langle T(v), w \rangle = \langle v, T^*(w) \rangle \tag{6.7}$$

(b) T heißt *selbstadjungiert*, wenn gilt $T^* = T$, also $\langle T(v), w \rangle = \langle v, T(w) \rangle$ für alle $v, w \in V$.

Ist A eine Matrixdarstellung eines selbstadjungierten Operators, dann gilt $A^* = A$. Mehr noch für alle $T \in \mathrm{L}(V)$ und alle $A \in \mathbb{K}^{n \times n}$ gilt:

$$(T_A)^* = T_{A^*}, \quad M_B(T^*) = (M_B(T))^*.$$

Lemma 6.8 *Die Eigenwerte eines selbstadjungierten Operators T sind reell.*

Beweis. Es sei $\lambda \in \mathbb{C}$ ein Eigenwert zu $T \in \mathrm{L}(V)$ und $v \in V$ ein Eigenvektor zu λ. Dann gilt

$$\lambda \langle v\,,\,v\rangle = \langle \lambda v\,,\,v\rangle = \langle T(v)\,,\,v\rangle = \langle v\,,\,T(v)\rangle = \langle v\,,\,\lambda v\rangle = \overline{\lambda}\langle v\,,\,v\rangle\,.$$

Da $v \neq 0$ ein Eigenvektor ist, gilt $\langle v\,,\,v\rangle \neq 0$; also gilt $\lambda = \overline{\lambda}$, das heißt $\lambda \in \mathbb{R}$. ■

Beispiel 6.5 Die Matrix $A = \begin{pmatrix} 0 & \mathrm{i} \\ -\mathrm{i} & 0 \end{pmatrix}$ ist hermitesch, $A^* = A$, und hat daher nur reelle Eigenwerte:

$$\chi_A(\lambda) = \begin{vmatrix} \lambda & -\mathrm{i} \\ \mathrm{i} & \lambda \end{vmatrix} = \lambda^2 - 1 = (\lambda - 1)(\lambda + 1).$$

Da alle Eigenwerte einfach sind, ist A diagonalisierbar:

$$\operatorname{Ker}(A - I_2) = \operatorname{Ker}\begin{pmatrix} -1 & \mathrm{i} \\ -\mathrm{i} & -1 \end{pmatrix} = \operatorname{lin}\begin{pmatrix} \mathrm{i} \\ 1 \end{pmatrix} =: \operatorname{lin} b_1$$

Analog ist

$$\operatorname{Ker}(A + I_2) = \operatorname{Ker}\begin{pmatrix} 1 & \mathrm{i} \\ -\mathrm{i} & 1 \end{pmatrix} = \operatorname{lin}\begin{pmatrix} \mathrm{i} \\ -1 \end{pmatrix} =: \operatorname{lin} b_2.$$

Somit ist $\{b_1, b_2\}$ eine Basis, in der A Diagonalgestalt hat, nämlich

$$A = \begin{pmatrix} 0 & \mathrm{i} \\ -\mathrm{i} & 0 \end{pmatrix} = \begin{pmatrix} \mathrm{i} & \mathrm{i} \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} S^{-1},$$

wobei $S = \begin{pmatrix} \mathrm{i} & \mathrm{i} \\ 1 & -1 \end{pmatrix}$ die spaltenweise genommenen Eigenvektoren b_1 und b_2 sind.

Für selbstadjungierte Endomorphismen gilt, dass das orthogonale Komplement eines invarianten Teilraumes wieder invariant ist.

Satz 6.9 *Es sei $T^* = T$ ein selbstadjungierter Operator und $U \subset V$ ein invarianter Teilraum.*
(a) Dann ist $U^\perp$ ebenfalls invariant.
(b) Ist $B = C \cup D$ eine Basis von V, wobei C und D Basen von U bzw. $U^\perp$ sind, so gilt

$$M_B(T) = \begin{pmatrix} A & 0 \\ 0 & A' \end{pmatrix},$$

wobei $A = M_C(T)$ und $A' = M_D(T)$.

Beweis. (a) Wegen der Invarianz von U gibt es für alle $u \in U$ ein $u_1 \in U$ mit $u_1 = T(u)$. Für ein $v \in U^\perp$ gilt dann

$$\langle u\,,\,T(v)\rangle = \langle T(u)\,,\,v\rangle = \langle u_1\,,\,v\rangle = 0.$$

Somit steht $T(v)$ senkrecht auf U, also $T(v) \in U^\perp$. Damit ist $T(U^\perp) \subset U^\perp$ ebenfalls invariant.
(b) Wir erhalten linksunten die Nullmatrix, da für keinen Vektor aus U der Vektor $T(u)$ einen Anteil in $U^\perp$ hat (Invarianz). Wir erhalten rechtsoben die Nullmatrix, da $T(v)$, $v \in U^\perp$, keinen Anteil in U hat. ■

Satz 6.10 (Hauptachsentransformation) *Es sei $T^* = T$, $T \in \mathrm{L}(V)$, ein selbstadjungierter Operator.*
(a) Dann ist T diagonalisierbar. Mehr noch, es existiert eine Orthonormalbasis B von V, so dass die Matrix $M_B(T)$ von T diagonal und reell ist.
(b) Ist $A^* = A \in \mathbb{C}^{n\times n}$ hermitesch, so existiert eine unitäre Matrix $S \in \mathrm{U}(n)$ und $\lambda_1,\ldots,\lambda_n \in \mathbb{R}$ mit

$$A = S \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} S^{-1}.$$

(c) Ist $A = A^\top \in \mathbb{R}^{n\times n}$ symmetrisch, so existiert eine orthogonale Matrix $S \in \mathrm{O}(n)$ und $\lambda_1,\ldots,\lambda_n \in \mathbb{R}$ mit

$$A = S\,\mathrm{diag}(\lambda_1,\ldots,\lambda_n)S^{-1}.$$

Beweis. (a) Der Beweis erfolgt durch vollständige Induktion über die Dimension n des Raumes V. Der Induktionsanfang, $\dim V = 1$ ist trivial, denn jeder eindimensionale Endomorphismus ist bereits diagonal. Wir setzen voraus, dass jeder selbstadjungierte Endomorphismus eines $(n-1)$-dimensionalen Raumes diagonalisierbar sei.
Sei nun T wie oben, $\dim V = n$. Nach Bemerkung 6.1 besitzt T mindestens einen Eigenwert λ_1, der nach Lemma 6.8 sogar reell ist. Sei e_1 ein zugehöriger Eigenvektor, e_1 sei normiert. Dann ist $U = \mathrm{lin}\{e_1\}$ ein invarianter Unterraum von V, da $T(e_1) = \lambda_1 e_1$. Nach Satz 6.9 ist das orthogonale Komplement $U_1 = U^\perp$ ein T-invarianter, $(n-1)$-dimensionaler Unterraum von V. Nach Induktionsvoraussetzung ist die Einschränkung $T{\restriction}U_1$ diagonalisierbar und es existiert eine orthonormale Eigenbasis $\{e_2,\ldots,e_n\}$ von U_1 zu den reellen Eigenwerten $\lambda_2,\ldots,\lambda_n$. Da $e_1 \perp U_1$ und $\|e_1\| = 1$, ist $\{e_1,\ldots,e_n\}$ eine Orthonormalbasis von Eigenvektoren von T und $M_B(T) = \mathrm{diag}(\lambda_1,\ldots,\lambda_n)$. Der Induktionsschluss ist fertig.
(b) und (c) folgen aus (a), angewandt auf den selbstadjungierten Operator $T = T_A \in \mathrm{L}(C^n)$ bzw. $T_A \in \mathrm{L}(\mathbb{R}^n)$. Die Matrix S erhält man indem man die Koordinaten der Eigenvektoren $e_1,\ldots,e_n$ spaltenweise aufschreibt. Nach Lemma 5.9 ist S unitär bzw. orthogonal. ■

Bemerkung 6.2 Führen Sie die Hauptachsentrasformation mit $T = T^*$ durch. Diese Standardaufgabe löst man, indem man den Algorithmus zur **Diagonalisierung** aus Abschnit 6.2.3 an zwei Stellen verändert bzw. ergänzt:

1. Auswahl einer **ONB** B in V und Bestimmung der Darstellungsmatrix $A = M_B(T)$.

4.5 Die Eigenvektoren jedes Eigenraumes E_λ werden *jeweils separat* mit dem Schmidtschen Verfahren orthogonalisiert.

5. Die n Eigenvektoren bilden nun eine **ONB** E. Schreibt man die Koordinaten der Vektoren aus E zur Basis B als Spalten einer Matrix S auf, so gilt

$$A = S\,\mathrm{diag}(\lambda_1,\ldots,\lambda_n)S^{-1} = SA'S^*.$$

mit der Diagonalmatrix $A' = M_E(T) = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$. Die Matrix S ist dabei unitär bzw. orthogonal, sodass $S^{-1} = S^*$.

Kapitel 7

Bilinearformen und quadratische Formen

7.1 Bilinearformen

7.1.1 Linearformen

Definition 7.1 Es sei V ein linearer Raum über dem Körper $\mathbb{K}$. Als *lineares Funktional* oder *Linearform* bezeichnet man die Elemente von $\mathrm{L}(V,\mathbb{K})$, das heißt die linearen Abbildungen von V, deren Werte skalar sind. Der Raum $\mathrm{L}(V,\mathbb{K})$ heißt auch *Dualraum* von V und wird mit V^* bezeichnet.

Beispiel 7.1 (a) $V=\mathbb{R}^n$. Es sei $B=\{b_1,\dots,b_n\}$ eine Basis von V, dann ist für jedes $i=1,\dots,n$, die Abbildung

$$V\ni x_1b_1+x_2b_2+\cdots+x_nb_n\mapsto x_i\in\mathbb{K}$$

ein lineares Funktional. Es heißt $i$tes *Koordinatenfunktional*, und man bezeichnet es mit b_i^*. Folglich gilt für alle $i,j=1,\dots,n$, dass $b_i^*(b_j)=\delta_{ij}$.
Es gilt $(\mathbb{R}^n)^*\cong\mathbb{R}^n$, denn jedes lineare Funktional F auf $\mathbb{R}^n$ ist durch einen Zeilenvektor $(a_1,\dots,a_n)\in\mathbb{R}^n$ gegeben über

$$\mathbb{R}^n\ni x=(x_1,\dots,x_n)\mapsto\sum_{i=1}^n a_ix_i.$$

In diesem Falle gilt $F=\sum_{i=1}^n a_i\mathrm{e}_i^*$. Umgekehrt gibt es für jedes lineare Funktional $F\in(\mathbb{R}^n)^*$ einen solchen Vektor $a=(a_1,\dots,a_n)$ so dass F die obige Form hat, $F=\sum_{i=1}^n F(\mathrm{e}_i)\mathrm{e}_i^*$.
(b) Es sei $V=\mathrm{C}[0,1]$, $g\in V$ fest. Setzt man für $f\in V$

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

so ist dadurch eine Linearform auf V definiert.
(c) Sei $V=\mathrm{C}(\mathbb{R})$ und $a\in\mathbb{R}$ fest. Dann definiert

$$F_a(f)=f(a),\quad f\in V$$

ein lineares Funktional auf V.
(d) Es sei $(V,\langle\cdot,\cdot\rangle)$ ein unitärer Raum und $a\in V$ fixiert. Dann definiert $F(v)=\langle v\,,a\rangle$ eine Linearform auf V.

7.1.2 Bilinearformen – Einfachste Eigenschaften

Definition 7.2 Es sei V ein Vektorraum. Eine Abbildung $B\colon V\times V\to\mathbb{K}$ heißt *Bilinearform* auf V, falls für alle $u,v,w\in V$ und $\lambda,\mu\in\mathbb{K}$ gilt:

$$B(\lambda v+\mu w,u)=\lambda B(v,u)+\mu B(w,u),\quad \text{Linearität im ersten Argument}$$
$$B(u,\lambda v+\mu w)=\lambda B(u,v)+\mu B(u,w)\quad \text{Linearität im zweiten Argument.}$$

Eine Bilinearform B heißt *symmetrisch*, wenn für alle $u,v\in V$ gilt $B(u,v)=B(v,u)$; sie heißt *schiefsymmetrisch*, wenn für alle $u,v\in V$ gilt $B(u,v)=-B(v,u)$.
B heißt *nicht-entartet* (oder nicht-ausgeartet), falls gilt

$$B(x,y)=0\quad\forall y\in V\Longrightarrow x=0.$$

Andernfalls heißt B *ausgeartet*.

Beispiel 7.2 (a) Ein euklidisches Skalarprodukt ist also eine symmetrische Bilinearform, die zusätzlich positiv definit ist.
(b) Es sei $A\in\mathbb{R}^{n\times n}$. Dann definiert $B(x,y)=\langle x\,,Ay\rangle=x^\top\cdot A\cdot y$, $x,y\in\mathbb{R}^n$, eine Bilinearform auf $\mathbb{R}^n$. Wir werden sehen, dass es für jede Bilinearform auf dem $\mathbb{R}^n$ eine Matrix $A\in\mathbb{R}^{n\times n}$ gibt, so dass $B(x,y)=\langle x\,,Ay\rangle$. Ist $A^\top=A$, so ist B symmetrisch, ist $A^\top=-A$, so ist B schiefsymmetrisch. Die Determinantenfunktion auf $\mathbb{R}^2$ ist ein Beispiel für eine schiefsymmetrische Bilinearform.
(c) Sind $f,g\in V^*$, so definiert $B(x,y)=f(x)g(y)$, $x,y\in V$ eine Bilinearform.
(d) Jede Bilinearform B ist die Summe aus einer symmetrischen Bilinearform S und einer schiefsymmetrischen Bilinearform T. Setzt man nämlich für alle $x,y\in V$,

$$S(x,y)=\frac{1}{2}\big(B(x,y)+B(y,x)\big),\qquad T(x,y)=\frac{1}{2}\big(B(x,y)-B(y,x)\big),$$

so ist $B=S+T$, wobei S symmetrisch und T schiefsymmetrisch ist. Der Raum der symmetrischen Bilinearformen (bzw. symmetrischen $n\times n$-Matrizen) hat die Dimension $\frac{1}{2}n(n+1)$; der Raum der schiefsymmetrischen Bilinearformen über V (der schiefsymmetrischen $n\times n$-Matrizen) hat die Dimension $\frac{1}{2}n(n-1)$.

7.1.3 Die Matrix einer Bilinearform

(a) Die Matrix $M_E(B)$.

In Bemerkung 5.1 haben wir zu einem Skalarprodukt die Matrix $G=(\langle e_i\,,e_j\rangle)_{\substack{i=1\dots n\\ j=1\dots n}}$ definiert, wobei $E=\{e_1,\dots,e_n\}$ eine Basis von V sei. Analog definiert man die *Matrix einer Bilinearform* B, $A=(B(e_i,e_j))_{\substack{i=1\dots n\\ j=1\dots n}}$; Bezeichnung: $A=M_E(B)$. Für $v=\sum_{i=1}^n x_ie_i$ und $w=\sum_{j=1}^n y_je_j$ ist dann

$$B(v,w)=B\left(\sum_{i=1}^n x_ie_i,\sum_{j=1}^n y_je_j\right)=\sum_{i,j=1}^n x_iy_j\,B(e_i,e_j)=\sum_{i,j}x_ia_{ij}y_j=\langle x\,,Ay\rangle=x^\top\cdot A\cdot y.$$

Weil $B(v,w)$ eine Zahl ist, gilt $B(v,w)=\langle x\,,Ay\rangle=(x^\top Ay)^\top=y^\top A^\top x$.

(b) Ausgeartetheit

Nach Definition ist eine Bilinearform B ausgeartet, wenn es einen Vektor $v \neq 0$ gibt mit $B(v,w)=0$ für alle w. In Koordinaten bezüglich einer Basis E heißt das: $\langle x\,,Ay\rangle=0$ für alle $y\in\mathbb{R}^n$ bzw. $\langle A^\top x\,,y\rangle=0$ für alle y. Also gilt dies insbesondere für $y=A^\top x$ und somit $\langle A^\top x\,,A^\top x\rangle=\left\|A^\top x\right\|^2=0$. Wegen der Definitheit des SKP (bzw. der Norm) folgt, dass $A^\top x=0$ oder $x\in\operatorname{Ker}A^\top$. Folglich ist $A^\top$ nicht injektiv; äquivalent ist $\det A=0$, A ist nicht invertierbar. Wegen $\det A^\top=\det A$ kann man daher in der Definition der Nicht-Ausgeartetheit einer Bilinearform die Rolle von x und y vertauschen.
Fazit: Für jede Basis E gilt: B ist genau dann nicht-entartet, wenn die zugehörige Matrix $A=M_E(B)$ regulär ist.

(c) Transformationsformel

Es sei V ein reeller Vektorraum mit den Basen $E=\{e_1,\dots,e_n\}$ und $F=\{f_1,\dots,f_n\}$. Die Basistransformation sei gegeben durch eine Matrix $S\in\mathbb{R}^{n\times n}$, $S=(s_{ij})$, in üblicher Weise:

$$f_j=\sum_{i=1}^n s_{ij}\,e_i,\quad j=1,\dots,n; \tag{7.1}$$

siehe auch Abschnitt 3.3.5. Die Bilinearform B werde bezüglich E beschrieben durch die Matrix $A=(a_{ij})$ und bezüglich F durch die Matrix $A'=(a'_{ij})$. Dann gilt:

$$\begin{aligned} a'_{kl}=B(f_k,f_l)&=B\left(\sum_{i=1}^n s_{ik}\,e_i,\sum_{j=1}^n s_{jl}\,e_j\right)\\ &=\sum_{i,j=1}^n s_{ik}s_{jl}\,B(e_i,e_j)=\sum_{i,j=1}^n s_{ik}s_{jl}\,a_{ij}=\sum_{i,j=1}^n (S^\top)_{ki}\,a_{ij}\,s_{jl}=(S^\top AS)_{kl}; \end{aligned}$$

also gilt die folgende *Transformationsformel für Bilinearformen*

$$\boxed{A'=S^\top AS.}$$

Man beachte, dass sich diese Formel von der Transformationsformel für Endomorphismen unterscheidet. Die Matrizen A und A' haben i. a. unterschiedliche charakteristische Polynome und damit unterschiedliche Eigenwerte.

7.2 Quadratische Formen und Bilinearformen

7.2.1 Definition und Beispiel

Definition 7.3 Es sei $B\colon V\times V\to\mathbb{R}$ eine Bilinearform auf V. Dann heißt $Q_B(x)=B(x,x)$, $x\in V$, die zugehörige *quadratische Form*. Es sei $A=M_E(B)\in\mathbb{R}^{n\times n}$ die Matrix zu B bezüglich einer Basis E, dann ist $Q_B(x)=\langle x\,,Ax\rangle=\sum_{i,j=1}^n a_{ij}\,x_i x_j$ die zugehörige quadratische Form.
Sei umgekehrt $Q\colon V\to\mathbb{R}$ eine quadratische Form, dann definiert

$$C(x,y)=\frac{1}{4}\left(Q(x+y)-Q(x-y)\right)$$

eine *symmetrische* Bilinearform C.

Wir schreiben auch Q_A für Q_B, wenn $A = M_E(B)$ die Matrix von B ist. Man beachte, dass für $\lambda \in \mathbb{R}$ gilt $Q(\lambda x) = \lambda^2 Q(x)$, $x \in V$. Außerdem ist $Q(v+w) = Q(v) + B(v,w) + B(w,v) + Q(w)$, $v, w \in V$.

Beispiel 7.3 (a) Auf $\mathbb{R}^2$ mit der Standardbasis ist durch $B(x,y) = x_1y_1 + x_1y_2 + 3x_2y_1$ eine Bilinearform gegeben. Die zugehörige Matrix von B ist dann $A = \begin{pmatrix} 1 & 1 \\ 3 & 0 \end{pmatrix}$. Die zugehöriger quadratischer Form lautet

$$Q(x_1, x_2) = x_1^2 + 4x_1x_2.$$

Die zugehörige symmetrische Bilinearform ist dann

$$\begin{aligned} C(x,y) &= \frac{1}{4}\left((x_1+y_1)^2 + 4(x_1+y_1)(x_2+y_2) - (x_1-y_1)^2 - 4(x_1-y_1)(x_2-y_2)\right) = x_1y_1 + 2x_1y_2 + 2x_2y_1 \\ &= \frac{1}{2}\left(B(x,y) + B(y,x)\right). \end{aligned}$$

Die C entsprechende Matrix ist $A' = \begin{pmatrix} 1 & 2 \\ 2 & 0 \end{pmatrix}$. Wenn man mit einer quadratischen Form startet, kann man sich also immer auf symmetrische Bilinearformen beschränken.
(b) Bestimmen Sie die zur quadratischen Form assoziierte symmetrische Bilinearform: $Q\colon \mathbb{R}^3 \to \mathbb{R}$, $Q(x) = 2x_1^2 - x_3^2 + 4x_1x_2 - 6x_1x_3$.
Dazu muss man die Matrix von Q aufschreiben (bzgl der Standardbasis), wobei die gemischten Glieder **halbiert** werden:

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

Also ist

$$B(x,y) = x^\top A y = 2x_1y_1 + 2x_1y_2 + 2x_2y_1 - x_3y_3 - 3x_1y_3 - 3x_3y_1.$$

In der Tat ist dann $B(x,x) = Q(x)$.

7.2.2 Definitheit

Für lokale Extrema von Funktionen mehrerer Veränderlicher ist von entscheidender Bedeutung, welche Werte eine gewisse quadratische Formen annehmen kann. Das Verhalten der *Hesse*-Matrix von f an der Stelle x_0, $\left(\frac{\partial^2 f}{\partial x_i \partial x_j}(x_0)\right)_{i,j}$ ist hier entscheidend.

Definition 7.4 Es sei $A \in \mathbb{R}^{n\times n}$ eine symmetrische quadratische Matrix und $Q(x) = \sum_{i,j=1}^n a_{ij}x_ix_j$, $x = (x_1, \dots, x_n) \in \mathbb{R}^n$, die zugehörige quadratische Form. Q heißt

(a) *positiv definit*, falls für alle $x \in V$ mit $x \neq 0$ gilt $Q(x) > 0$.
(b) *negativ definit*, falls für alle $x \in V$ mit $x \neq 0$ gilt, $Q(x) < 0$.
(c) *indefinit*, falls es Vektoren $x_1, x_2 \in V$ gibt mit $Q(x_1) > 0 > Q(x_2)$.
(d) *positiv semidefinit*, falls für alle $x \in V$, $Q(x) \geq 0$.

(e) *negativ semidefinit*, falls für alle $x \in V$, $Q(x) \leq 0$.

Man sagt, dass die symmetrische Matrix A bzw. die symmetrische Bilinearform B positiv definit, negativ definit usw. ist, falls Q_A positiv definit, negativ definit usw. ist.
Im komplexen Fall betrachtet man in der Regel *Sesquilinearformen* B, die im ersten Argument linear und im zweiten Argument antilinear sind. B heißt hermitesch, falls zusätzlich gilt $B(x,y) = \overline{B(y,x)}$. Diese Formen sind durch hermitesche Matrizen $A^* = A$ gegeben über $B(x,y) = \sum_{i,j=1}^n a_{ij} x_i \overline{y_j}$. Man assoziiert dann zu B die Form $Q(x) = \sum_{i,j=1}^n a_{ij} x_i \overline{x}_j$.

Lemma 7.1 *Es sei A ein quadratische symmetrische, reelle Matrix. Die quadratische Form Q_A ist positiv definit genau dann, wenn alle Eigenwerte positiv sind, negativ definit, genau dann, wenn alle Eigenwerte von A negativ sind und indefinit, falls es einen positiven und einen negativen Eigenwert gibt. Q_A ist positiv semidefinit, falls alle Eigenwerte von A nichtnegativ sind und negativ semidefinit, falls alle Eigenwerte von A kleiner gleich Null sind.*

Beweis. Nach dem Satz über die Hauptachsentransformation besitzt $\mathbb{R}^n$ eine Orthonormalbasis aus Eigenvektoren $\{e_1, \ldots, e_n\}$ zu den reellen Eigenwerten $\lambda_1, \ldots, \lambda_n$. Für $x = \sum_{i=1}^n x_i e_i$, $x \neq 0$, folgt dann

$$Q_A(x) = \langle x, Ax \rangle = \left\langle \sum_{i=1}^n x_i e_i, \sum_{j=1}^n \lambda_j x_j e_j \right\rangle = \sum_{i,j=1}^n x_i x_j \lambda_i \delta_{ij} = \sum_{i=1}^n \lambda_i x_i^2.$$

Sind nun alle Eigenwerte von A positiv, so ist offenbar auch $Q_A(x) > 0$ für alle $x \neq 0$. Ist umgekehrt $Q_A(x) > 0$ für alle von Null verschiedenen Vektoren, so ist insbesondere $Q_A(e_i) = \lambda_i > 0$. Also sind alle Eigenwerte von A positiv.
Die anderen Aussagen (A negativ definit, indefinit usw.) folgen analog. ■

7.2.3 Die Normalform einer symmetrischen Bilinearform

Satz 7.2 *Es sei V ein reeller Vektorraum mit* $\dim V = n$ *und $B\colon V \times V \to \mathbb{R}$ eine symmetrische Bilinearform auf V.*
Dann gibt es eine Basis $E = \{e_1, \ldots, e_n\}$ von V, so dass die Matrixdarstellung $A = M_E(B)$ folgende Gestalt hat

$$A = \begin{pmatrix} I_p & 0 & 0 \\ 0 & -I_q & \\ 0 & 0 & 0 \end{pmatrix},$$

wobei die angegebenen Blöcke auch die Zeilenzahl 0 *haben dürfen. Mit anderen Worten, es gibt Zahlen $p, q \geq 0$, $p + q \leq n$, so dass bezüglich E die Bilinearform B die Gestalt*

$$B((x_1, \ldots, x_n), (y_1, \ldots, y_n)) = x_1 y_1 + x_2 y_2 + \cdots + x_p y_p - x_{p+1} y_{p+1} - \cdots - x_{p+q} y_{p+q}$$

hat. Die Zahl $p + q$ heißt Rang *von B und $p - q$ heißt* Signatur *von B.*

Beweis. Wir statten V mit einem Skalarprodukt aus, wählen eine beliebige Basis E_1 und wenden die Hauptachsentransformation auf die symmetrische Matrix $A' = M_{E_1}(B)$ an. Nach dem Satz über die Hauptachsentransformation existiert eine Orthonormalbasis E_2 von V aus Eigenvektoren von A' zu den ausschließlich reellen Eigenwerten von A'. Wir ordnen nun die Eigenvektoren $\{f_1,\dots,f_n\}$ so in E_2 an, dass zunächst alle positiven Eigenwerte $\{\lambda_1,\dots,\lambda_p\}$, dann alle negativen Eigenwerte $\{\lambda_{p+1},\dots,\lambda_{p+q}\}$ und dann alle Eigenwerte 0 kommen. In der Basis E_2 hat also B die Gestalt

$$B(x,y)=\lambda_1 x_1 y_1+\lambda_2 x_2 y_2+\dots+\lambda_p x_p y_p-\left|\lambda_{p+1}\right|x_{p+1}y_{p+1}-\dots-\left|\lambda_{p+q}\right|x_{p+q}y_{p+q}.$$

Insbesondere ist also $B(f_k,f_k)=\lambda_k$ für alle k. Im letzten Schritt reskalieren wir die Basisvektoren von E_2 und erhalten die gewünschte Basis E:

$$e_k=\frac{1}{\sqrt{|\lambda_k|}}f_k,\quad k=1,\dots,p+q.$$

In der Tat ist nun $B(e_k,e_k)=B\left(\frac{f_k}{\sqrt{|\lambda_k|}},\frac{f_k}{\sqrt{|\lambda_k|}}\right)=\frac{1}{|\lambda_k|}\lambda_k=\operatorname{sign}\lambda_k$. Dann hat B in der Basis E die gewünschte Gestalt. ∎

Für symmetrische Matrizen A bedeutet das, dass eine invertierbare Matrix $S\in \mathrm{GL}(n,\mathbb{R})$ existiert mit

$$A=S^\top \operatorname{diag}(I_p,-I_q,0)S. \tag{7.2}$$

Man beachte, dass S in der Regel nicht mehr orthogonal ist und auch nicht eindeutig bestimmt. Im nächsten Abschnitt werden wir aber sehen, dass p und q nicht von der Wahl der Basis abhängen.

7.2.4 Der Sylvestersche Trägheitssatz

Satz 7.3 (Sylvesterscher Trägheitssatz) *Es sei B eine symmetrische Bilinearform auf dem reellen Vektorraum V, $\dim V=n$.*
Dann existieren Unterräume V_0, V_+ und V_- von V mit

$$V=V_0\oplus V_+\oplus V_-,$$

wobei $V_0=\{x\in V\mid B(x,y)=0\ \forall y\in V\}$, $B(x,x)>0$ für alle $x\in V_+\setminus\{0\}$ und $B(x,x)<0$ für alle $x\in V_-\setminus\{0\}$. Die Zahlen

$$p=\dim V_+\quad \textit{und}\quad q=\dim V_-$$

sind Invarianten von B und hängen nicht von der speziellen Wahl der Unterräume V_+ und V_- ab. Die Differenz $p-q$ heißt Signatur *und $p+q$ heißt* Rang *der symmetrischen Bilinearform B.*

Einzig der Nullraum V_0 ist eindeutig bestimmt. Die Räume V_+ und V_- kann man variieren indem man einzelne Basisvektoren durch Vektoren von V_0 abändert.

Beweis. (a) Existenz der Zerlegung. Die Existenz folgt sofort aus dem Satz 7.2, indem man $V_+ = \mathrm{lin}\{e_1,\dots,e_p\}$, $V_- = \mathrm{lin}\{e_{p+1},\dots,e_{p+q}\}$ und $V' := \mathrm{lin}\{e_{p+q+1},\dots,e_n\}$ setzt. Es ist klar, dass dann V_+ und V_- die Bedingungen der Aufgabe erfüllen, dass $V = V_+ \oplus V_- \oplus V'$ und dass $V' \subset V_0$. Wir müssen noch zeigen, dass $V' \supset V_0$ gilt. Dazu sei $x = \sum_{i=1}^n x_i e_i \in V_0$. Also gilt $B(x, e_i) = x_i = 0$ für alle $i = 1,\dots,p$ und $B(x, e_i) = -x_i = 0$ für alle $i = p+1,\dots,p+q$. Somit hat x nur noch von Null verschiedene Koeffizienten x_i, $i > p+q$, also $x \in V'$; also $V' = V_0$.

(b) Eindeutigkeit von p und q. Es sei $V = V_0 \oplus W_+ \oplus W_-$ eine zweite Zerlegung dieser Art. Für $v \in W_+ \cap (V_0 \oplus V_-) \setminus \{0\}$ gilt dann einerseits

$$B(v,v) > 0, \quad \text{weil } v \in W_+$$

und andererseits gilt

$$B(v,v) \le 0, \quad \text{weil } v = v_0 + v_- \in V_0 \oplus V_- \quad B(v_0 + v_-, v_0 + v_-) = 0 + 0 + 0 + B(v_-, v_-) \le 0.$$

Dies ist aber ein Widerspruch. Folglich ist $W_+ \cap (V_0 \oplus V_-) = \{0\}$. Somit gilt nach dem Dimensionssatz 2.7 wegen $W_+ + ((V_0 \oplus V_-) \subset V$

$$\begin{aligned}\dim V &\ge \dim(W_+ + (V_0 \oplus V_-)) = \dim W_+ + \dim V_0 \oplus V_- = \dim W_+ + \dim V_0 + \dim V_- \\ &= \dim W_+ + (n - p - q) + q \quad \Longrightarrow \quad p = \dim V_+ \ge \dim W_+.\end{aligned}$$

Aus Symmetriegründen (Vertauschen der Rollen von W_+, W_- und V_+, V_-) muss aber auch die umgekehrte Ungleichung gelten; somit gilt Gleichheit $\dim W_+ = \dim V_+$ und damit aus Dimensionsgründen auch $\dim V_- = \dim W_-$, was zu zeigen war. ■

Folgerung 7.4 *Es sei $A \in \mathbb{R}^{n\times n}$ eine symmetrische Matrix, $S \in \mathrm{GL}(n,\mathbb{R})$ und $A' = S^\top AS$. Dann haben A und A' die gleiche Anzahl von positiven und und die gleiche Anzahl von negativen Eigenwerten. Insbesondere ist A positiv definit (negativ definit, indefinit, usw.) genau dann, wenn A' positiv definit (negativ definit, indefinit, usw.) ist.*

Beweis. A und A' beschreiben ein und dieselbe Bilinearform nur in einer anderen Basis (weil S invertierbar ist). Daher ist $p = \dim V_+$ und $q = \dim V_-$ die Anzahl der positiven bzw. negativen Eigenwerte von A – und auch von A'. ■

7.2.5 Sylvesterkriterium für positiv definite Matrizen

Bemerkung 7.1 (a) Jede positiv (negativ) definite symmetrische Bilinearform ist nichtentartet. Das folgt sofort aus dem Trägheitssatz da alle Eigenwerte von A positiv (negativ) sind und somit A invertierbar.

(b) Die quadratische reelle, symmetrische Matrix $A = \begin{pmatrix} a & b \\ b & d \end{pmatrix}$ ist genau dann positiv definit, wenn

$$\det A = ad - b^2 > 0 \quad \text{und} \quad a > 0.$$

Zum Beweis betrachten wir die Gleichung, $x = (x_1, x_2) \in \mathbb{R}^2$,

$$a\,Q_A(x) = a(ax_1^2 + 2bx_1x_2 + dx_2^2) = (ax_1 + bx_2)^2 + (ad - b^2)x_2^2.$$

sind die obigen beiden Bedingungen erfüllt, so gilt offenbar $Q_A(x) > 0$ für alle $x \neq 0$. Ist umgekehrt $Q_A(x) > 0$, so insbesondere $Q((1,0)) = a > 0$ und somit auch $aQ((-b,a)) = (ad - b^2)a^2 > 0$, also $\det A > 0$.

(c) Ist $A = (a_{ij})$ positiv definit, so gilt
(1) $a_{ii} > 0$ für alle $i = 1, \ldots, n$.
(2) $a_{ii}a_{jj} - a_{ij}^2 > 0$ für alle $1 \leq i < j \leq n$.

Ist A negativ definit, so gilt
(1) $a_{ii} < 0$ für alle $i = 1, \ldots, n$.
(2) $a_{ii}a_{jj} - a_{ij}^2 > 0$ für alle $1 \leq i < j \leq n$.

Die erste Behauptung folgt aus $Q_A(\mathrm{e}_i) > 0$ bzw. aus $Q_A(x_1\mathrm{e}_i + x_2\mathrm{e}_j) > 0$.

(d) Ist A' eine $k \times k$ Untermatrix von A in der linken oberen Ecke von A, also

$$A = \begin{pmatrix} A' & * \\ * & * \end{pmatrix},$$

so ist A' ebenfalls positiv definit. Man nennt $\det A'$ den *k ten Hauptminor* von A.
Ist nämlich A' eine $r \times r$-Matrix, so wählt man $x = (x_1, \ldots, x_r, 0, \ldots, 0)$ und erhält $Q_A(x) = x^\top Ax = x'^\top A'x' > 0$, wobei $x' = (x_1, \ldots, x_r)$.

Wir verallgemeinern die Eigenschaft (b) für beliebige Dimension.

Satz 7.5 (Sylvestersches Definitheitskriterium) *Es sei $A = (a_{ij}) \in \mathbb{R}^{n\times n}$, $A = A^\top$, eine symmetrische reelle Matrix. Für $k = 1, \ldots, n$ bezeichen $A_k = (a_{ij})_{\substack{i=1\ldots k\\ j=1\ldots k}}$ die „abgeschnittenen" $k \times k$-Matrizen. Außerdem setzen wir $d_k = \det A_k$.*
Dann gilt:

(a) Q_A *ist genau dann positiv definit, wenn* $d_k > 0$ *für alle* $k = 1, \ldots, n$.
(b) Q_A *ist genau dann negativ definit, wenn* $(-1)^k d_k > 0$ *für alle* $k = 1, \ldots, n$

Man nennt die die Determinanten d_k die *Hauptminore* von A.
Beweis. Wir beweisen (a). (1) Die *Notwendigkeit* der Bedingung ist klar: Ist A positiv definit, so ist nach Bemerkung 7.1 (d) für alle k die Matrix A_k positiv definit. Damit sind alle ihre Eigenwerte $\lambda_1^{(k)}, \ldots, \lambda_k^{(k)}$ positiv. Ferner gilt $\det A_k = \lambda_1^{(k)} \cdots \lambda_k^{(k)} > 0$.
(2) *Hinlänglichkeit.* Beweis jdurch vollständige Induktion über n. Die Behauptung gelte schon für alle symmetrischen Matrizen der Ordnung $n - 1$ mit sämtlich positiven Hauptminoren $d_1, \ldots d_{n-1} > 0$. Nun sei auch $d_n = \det A > 0$. Insbesondere sind alle A_k invertierbar. Man schreibt nun $A = A_n$ in Kästchenform wie folgt auf:

$$A = \begin{pmatrix} A_{n-1} & v \\ v^\top & a \end{pmatrix},$$

mit der symmetrischen $(n-1)\times(n-1)$-Matrix A_{n-1}, einem Spaltenvektor $v \in \mathbb{R}^{n-1}$ und $a \in \mathbb{R}$. Man rechnet nach, dass nun

$$A = S^\top \begin{pmatrix} A_{n-1} & 0 \\ 0 & b \end{pmatrix} S, \quad S = \begin{pmatrix} I_{n-1} & A_{n-1}^{-1} v \\ 0 & 1 \end{pmatrix}, \quad b = a - v^\top A_{n-1}^{-1} v.$$

Nach Folgerung 7.4 ist A genau dann positiv definit, wenn die mit S transformierte Matrix $\begin{pmatrix} A_{n-1} & 0 \\ 0 & b \end{pmatrix}$ positiv definit ist. Dies ist wiederum äquivalent zur positiven Definitheit von A_{n-1} und $b > 0$. Nach Induktionsvoraussetzung ist aber $d_1, d_2, \dots, d_{n-1} > 0$ und damit A_{n-1} positiv definit. Außerdem folgt aus der Zerlegung

$$\det A = \det S \det A_{n-1} b \det S.$$

Nun ist aber S aber eine obere Dreiecksmatrix mit lauter Einsen auf der Diagonale; also ist $\det S = 1$. Folglich gilt $\det A = b \det A_{n-1}$. Mit $\det A > 0$ und $\det A_{n-1} > 0$ folgt $b > 0$ und die positive Definitheit von A ist gezeigt.
(3) Sei die Bedingung in (b) erfüllt. Wir betrachten die symmetrische Matrix $-A$. Alle Hauptminore von $-A$ sind dann positiv und damit ist $-A$ nach (a) positiv definit und folglich ist A negativ definit. Die Notwendigkeit folgt analog. ■

Beispiel 7.4 Die Matrix

$$A = \begin{pmatrix} 2 & -1 & -3 \\ -1 & 2 & 4 \\ -3 & 4 & 9 \end{pmatrix}$$

ist positiv definit, da ihre Hauptminore

$$\det A_1 = 2, \quad \det A_2 = \begin{vmatrix} 2 & -1 \\ -1 & 2 \end{vmatrix} = 3, \quad \det A = \begin{vmatrix} 2 & -1 & -3 \\ -1 & 2 & 4 \\ -3 & 4 & 9 \end{vmatrix} = 1$$

alle positiv sind.

Für positiv definite Matrizen können wir Folgendes zusammenfassend feststellen.

Satz 7.6 (Äquivalenzsatz für positiv definite Matrizen) *Für eine symmetrische Matrix $A \in \mathbb{R}^{n\times n}$, $A^\top = A$, sind äquivalent*

i. A ist positiv definit.
ii. Es gibt eine invertierbare quadratische Matrix $S \in \mathrm{GL}(n, \mathbb{R})$ mit $A = S^\top S$.
iii. Alle Hauptminore sind positiv.
iv. Alle Eigenwerte von A sind positiv.
v. Es gibt eine orthogonale Matrix $U \in \mathrm{O}(n)$ und eine Diagonalmatrix A' mit positiven Diagonalelementen, so dass $A = U^\top A' U = U^{-1} A' U$.
vi. A ist positiv semidefinit und $\det A \neq 0$.
vii. A ist invertierbar und A^{-1} ist positiv semidefinit.

7.3 Geometrie in euklidischen Räumen

Im Folgenden sei V ein euklidischer Raum mit Skalarprodukt.

7.3.1 Geraden und Strecken

Eine *Gerade* g in V ist durch einen Punkt $p_0 \in V$ und einen Richtungsvektor $r \in V$, $r \neq 0$, gegeben,

$$g = \{p_0 + t\,r \mid t \in \mathbb{R}\}.$$

Ist $p_0 = 0$, so ist g sogar ein eindimensionaler Teilraum. Variiert der Parameter t nur in einem endlichen Intervall $[a,b]$, so nennt man $s = \{p_0 + t\,r \mid t \in [a,b]\}$ eine *Strecke*. Die Strecke $\overline{p_0 p_1}$ von p_0 nach p_1, $p_0, p_1 \in V$ hat die Gestalt $\overline{p_0 p_1} = \{(1-\lambda)p_0 + \lambda p_1 \mid \lambda \in [0,1]\}$.

Definition 7.5 Es sei $p_0 \in V$ und $U \subset V$ ein linearer Teilraum von V, dann bezeichnet man

$$M = p_0 + U = \{p_0 + u \mid u \in U\}$$

als *affinen Teilraum* von V. Wir bezeichnen $\dim M := \dim U$ als *Dimension* des affinen Teilraumes M.

Punkte und Geraden sind also genau die null- bzw. eindimensionalen affinen Teilräume von V. Die $n-1$-dimensionalen affinen Teilräume bezeichnet man als *Hyperebenen*.

Beispiel 7.5 (a) Für jedes lineare Gleichungssystem $Ax = b$, $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$, ist $\text{Lös}(A,b)$ ein affiner Teilraum von $\mathbb{R}^n$ der Dimension $\operatorname{def} A$.
(b) Ist $v \in V$, $v \neq 0$, so ist

$$v^\perp = \{x \in V \mid \langle v\,,\, x\rangle = 0\}$$

eine Hyperebene durch den Ursprung. Jede Hyperebene im $\mathbb{R}^n$ ist gegeben durch eine einzige lineare Gleichung:

$$H = \{(x_1,\dots,x_n) \mid a_1x_1 + \cdots + a_nx_n = c\},$$

wobei $a = (a_1,\dots,a_n) \neq 0$, $c \in \mathbb{R}$. Die Hyperebenen des $\mathbb{R}^3$ sind die Ebenen, die des $\mathbb{R}^2$ sind die Geraden.

7.3.2 Kurven und Flächen zweiter Ordnung — Quadriken

Die allgemeine Gleichung 2. Grades im $\mathbb{R}^n$ hat die Gestalt (mit $a_{ij} = a_{ji}$, a_k, $c \in \mathbb{R}$)

$$\sum_{i,j=1}^n a_{ij}x_ix_j + 2\sum_{i=1}^n a_kx_k + c = 0, \quad a_{ij} = a_{ji}.$$

Fasst man die Koeffizienten zu eine symmetrischen Matrix $B = (a_{ij}) \in \mathbb{R}^{(n+1)\times(n+1)}$ zusammen, $B = \begin{pmatrix} A & a \\ a^\top & c\end{pmatrix}$, wobei a_{ij} die Koeffizienten aus der Gleichung sind, so kann man die Gleichung schreiben als

$$P_B(x) = \langle x\,,\, Ax\rangle + 2\,\langle a\,,\, x\rangle + c = x^\top A\,x + 2a^\top x + c = 0.$$

Als *Hyperfläche zweiter Ordnung* oder *Quadrik* bezeichnet man jede Menge (sofern sie nicht leer ist)

$$M_B = \{x \in \mathbb{R}^n \mid P_B(x) = 0\}, \quad A \neq 0. \tag{7.3}$$

Die Gestalt einer Quadrik ist vollständig bestimmt durch Rang und Signatur von A und B.

Satz 7.7 (Normalformen-Satz für Quadriken) *Durch Verschiebung ($x := x + d$) und Bewegung ($x := Sx$, $S^\top S = I_n$, S ist orthogonal) lässt sich jede Quadrik* (7.3) *auf eine der folgenden Normalformen transformieren:*

(I) $\lambda_1 x_1^2 + \cdots + \lambda_n x_n^2 + \alpha = 0.$
(II) $\lambda_1 x_1^2 + \cdots + \lambda_{n-1} x_{n-1}^2 + 2\beta x_n = 0.$

Dabei sind $\lambda_1, \ldots, \lambda_n$ die Eigenwerte von A, wobei im Falle (II) $\lambda_n = 0$ gilt.

Im Falle $n = 2$ erhält man die folgenden Normalformen von Quadriken (Kurven 2. Ordnung):

i. Ellipse, $a_1^2 x_1^2 + a_2^2 x_2^2 = 1$, wobei $a_1 a_2 \neq 0$.

ii. Hyperbel oder zwei sich schneidende Geraden, $a_1^2 x_1^2 - a_2^2 x_2^2 = 1$ oder $= 0$ mit $a_1 a_2 \neq 0$.

iii. Parabel, $x_1^2 + b x_2 = 0$ mit $b \neq 0$.

iv. $x_1^2 = c$ mit $c > 0$ (paralleles Geradenpaar) oder mit $c = 0$ (Doppelgerade).

7.3.3 Beispiel zur Hauptachsentransformation

Geben sei die quadratische Gleichung

$$f(x_1, x_2) = x_1^2 + 6x_1 x_2 + x_2^2 + 8x_1 + 8x_2 + 1 = 0, \quad x_1, x_2 \in \mathbb{R},$$

welche eine Kurve $C = \{(x_1, x_2) \in \mathbb{R}^2 \mid f(x_1, x_2) = 0\}$ in der Ebene $\mathbb{R}^2$ definiert. Ist C die leere Menge, ein Geradenpaar, eine Ellipse, eine Hyperbel oder eine Parabel?

Wir schreiben $f(x)$, $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, in Matrixform,

$$f(x) = x^\top A x + b^\top x + c \quad A = \begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix}, \quad b = \begin{pmatrix} 8 \\ 8 \end{pmatrix}, \quad c = 1.$$

und führen mit A die Hauptachsentransformation durch. Wegen $\det A = 1 - 9 = -8$ und $\operatorname{tr} A = 2$ ist $\chi_A(\lambda) = \lambda^2 - 2\lambda - 8 = (\lambda - 4)(\lambda + 2)$ und damit sind $\lambda_1 = 4$ und $\lambda_2 = -2$ die Eigenwerte von A. Die Eigenräume sind:

$$V_{\lambda_1} = \text{Lös}(A - 4, 0) = \{x \in \mathbb{R}^2 \mid -3x_1 + 3x_2 = 0\} = \mathbb{R}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$V_{\lambda_2} = \text{Lös}(A + 2, 0) = \{x \in \mathbb{R}^2 \mid 3x_1 + 3x_2 = 0\} = \mathbb{R}\begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Die Orthonormalbasis aus Eigenvektoren lautet also $E = \{(1,1)/\sqrt{2}, (1,-1)/\sqrt{2}\}$. Die Transformationsmatrix (vergleiche Abschnitt 3.3.6) $S = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ führt auf die neuen Koordinaten

$$y = S^{-1}x, \quad x = Sy = \begin{pmatrix} \frac{1}{\sqrt{2}}(y_1 + y_2) \\ \frac{1}{\sqrt{2}}(y_1 - y_2) \end{pmatrix}, \quad x_1 = \frac{1}{\sqrt{2}}(y_1 + y_2), \quad x_2 = \frac{1}{\sqrt{2}}(y_1 - y_2).$$

Setzt man x_1 und x_2 in die Gleichung der Kurve C ein, so erhält man

$$0 = 4y_1^2 - 2y_2^2 + 8\sqrt{2}y_1 + 1 = \left(y_1 + 4\sqrt{2}\right)^2 - 2y_2^2 - 31 = 4z_1^2 - 2z_2^2 - 31.$$

wobei $z_1 = y_1 + 4\sqrt{2}$, $z_2 = y_2$ eine weitere Koordinatentransformation (Verschiebung in Nullpunktslage) ist. Weil A jeweils einen positiven und einen negativen Eigenwert hat, handelt es sich bei C um eine Hyperbel.

Literaturverzeichnis

[Ant98] H. Anton. *Lineare Algebra. Einführung, Grundlagen, Übungen.* Spectrum Lehrbuch. Spektrum Akademischer Verlag, Heidelberg, 1998.

[Fis05] G. Fischer. *Lineare Algebra.* Vieweg Studium: Grundkurs Mathematik. Friedrich Vieweg & Sohn, Wiesbaden, 15 edition, 2005.

[Jän04] K. Jänich. *Lineare Algebra.* Springer-Lehrbuch. Springer-Verlag, Berlin, 10 edition, 2004.

[Koe97] M. Koecher. *Lineare Algebra und analytische Geometrie.* Springer-Verlag, Berlin, 4 edition, 1997.

[Kow79] H.-J. Kowalsky. *Lineare Algebra.* De Gruyter Lehrbuch. Walter de Gruyter, Berlin, New York, 1979.

PD Dr. A. Schüler
Mathematisches Institut
Universität Leipzig
04009 Leipzig
`mailto:Axel.Schueler@math.uni-leipzig.de`