Christian Pötzsche

Lineare Algebra für Informatik (Kapitel 1–3)

SS 2011
TU München

5. Juli 2011

Christian Pötzsche
Zentrum Mathematik
Technische Universität München
Boltzmannstraße 3
D-85748 Garching

poetzsch@ma.tum.de
http://www-m12.ma.tum.de/web/poetzsch

Vorwort

Um den Untertitel des bekannten Lehrbuches [Beu10] zu zitieren, ist die Lineare Algebra die „Wissenschaft der Vektoren, Abbildungen und Matrizen." In der Tat ist sie noch mehr, nämlich der erste Kontakt der Studierenden mit abstrakten mathematischen Konzepten. Im Gegensatz zur klassischen Analysis ist die lineare Algebra als eigenständige mathematische Disziplin und wesentlicher Bestandteil des Grundstudiums noch relativ jung.

Vor dem zweiten Weltkrieg waren die heute in der Linearen Algebra verwendeten Rechentechniken (lineare Gleichungssysteme, Matrizen, Determinanten, etc.) wohlbekannt, wenn auch noch nicht zu einer konsistenten Theorie zusammengefasst. Dies geschah erst unter dem Einfluss der modernen Algebra, die sich von einer Lehre über Gleichungen hin zu abstrakteren mathematischen Strukturen entwickelte.

Das Ziel dieser Vorlesung ist eine Fokussierung auf abstrakte Konzepte und damit eine weitgehende Loslösung von geometrischer Visualisierung. Es geht uns eher darum, Strukturen in der Mathematik zu erkennen, als diese etwa graphisch im drei-dimensionalen Anschauungsraum darzustellen. Dies ist nicht zuletzt durch die vielfältigen Anwendungen der linearen Algebra motiviert. Als Beispiele seien Codierungstheorie, Verschlüsselung, Analyse von (elektrischen) Netzwerken, Filter-Theorie oder die Bildverarbeitung genannt. Leider ist es uns im einzelnen nicht möglich auf diese Anwendungen detailliert einzugehen, obgleich wir an entsprechenden Stellen Hinweise geben werden.

An Vorkenntnissen gehen wir davon aus, dass der Leser mit den Rechenregeln in den rationalen oder reellen Zahlen vertraut ist. Kenntnis und Verständnis des naiven Mengenbegriffs wird sich als hilfreich erweisen, wie auch eine Vertrautheit mit den Grundkonzepten der Differenzial- und Integralrechnung.

München, 5. Juli 2011 Christian Pötzsche

Inhaltsverzeichnis

Kapitel 1
Grundlegendes

Wir gehen davon aus, dass der Leser mit dem naiven Konzept einer Menge, sowie der Inklusion und Verknüpfungen (Schnitt, Vereinigung, Differenz oder Produkt) von Mengen vertraut ist; es sei auf [Beu10, S. 1ff] oder [Fis10, S. 32ff] verwiesen.

1.1 Relationen

Zentral für die gesamte Mathematik — und vielleicht auch das tägliche Leben — ist das Konzept einer Relation. Wir beginnen mit einer abstrakten Definition und der Betrachtung einiger Beispiele:

Definition 1.1.1 (Relation)**.** Gegeben seien zwei Mengen X und Y. Eine Teilmenge R des kartesischen Produkts $X \times Y := \{(x, y) : x \in X,\ y \in Y\}$ heißt *Relation* zwischen X und Y; im Fall $X = Y$ spricht man von einer *Relation auf* X. Ist $R \subseteq X \times Y$ eine Relation, so bezeichnet

$$R^{-1} := \{(y, x) \in Y \times X : (x, y) \in R\}$$

die *Umkehrrelation* von R.

Beispiel 1.1.2. Die Menge $R_0 := \{(x, y) \in X \times Y : y \text{ ist Hauptstadt von } x\}$ ist eine Relation zwischen der Menge X aller Länder und Y aller Städte. Ihre graphische Darstellung findet man in Abb. 1.1.

Beispiel 1.1.3. Mit den Mengen $X = \mathbb{R}$, $Y = [0, \infty)$ ist $R_1 := \{(x, |x|) \in X \times Y : x \in X\}$ eine Relation mit der Umkehrrelation $R^{-1} = \{(|x|, x) : x \in X\}$ (siehe Abb. 1.2).

Beispiel 1.1.4. Mit den Mengen $X = Y = \mathbb{R}$ definiert $R_2 := \{(x, y) \in X \times Y : x \le y\}$ eine Relation auf $\mathbb{R}$ mit der Umkehrrelation $R_2^{-1} = \{(y, x) : x \le y\}$ (siehe Abb. 1.3).

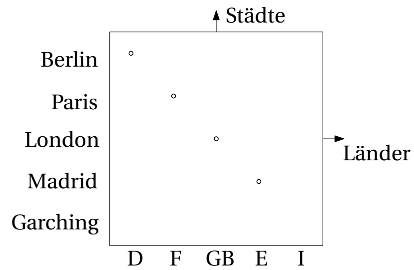

Abb. 1.1 Teilmenge der „Hauptstadt von“-Relation R_0 aus Bsp. 1.1.2

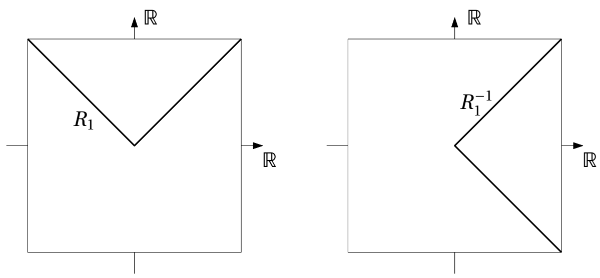

Abb. 1.2 Relation R_1 aus Bsp. 1.1.3 (links) und zugehörige Umkehrrelation R_1^{-1} (rechts)

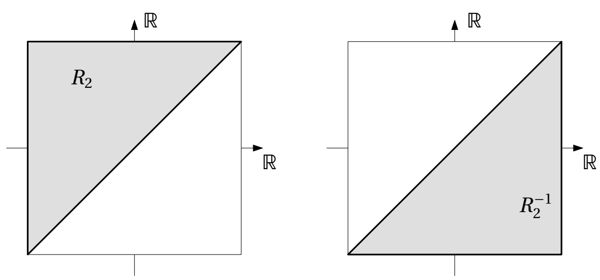

Abb. 1.3 Relation R_2 aus Bsp. 1.1.4 (links) und zugehörige Umkehrrelation R_2^{-1} (rechts)

Beispiel 1.1.5. Die Menge

$$R_3 := \{(x, y) \in C \times C : x \text{ und } y \text{ haben den gleichen Hersteller}\}$$

ist eine Relation auf der Menge aller Computer $C := \{\text{Computer}\}$; sie stimmt mit ihrer Umkehrrelation überein $R_3^{-1} = R_3$.

Bemerkung 1.1.6 (gerichtete Graphen). Relationen R auf endlichen Mengen X können alternativ wie folgt dargestellt werden: Man repräsentiert die Elemente aus X als Punkte in der Ebene und verbindet $x, y \in X$ genau dann durch einen gerichteten Pfeil, wenn $(x, y) \in R$. In diesem Kontext nennt man die Elemente von X *Punkte* (oder *Knoten*) und die Elemente von R *gerichtete Kanten*. Das Paar (X, R) heißt *gerichteter Graph* oder *Digraph* (siehe Abb. 1.4).

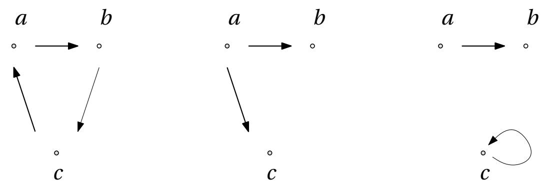

Abb. 1.4 Gerichteter Graph (X,R) mit der Menge $X = \{a,b,c\}$ versehen mit den Relationen R gegeben durch $\{(a,b),(b,c),(c,a)\}$ (links), $\{(a,b),(a,c)\}$ (mitte) und $\{(a,b),(c,c)\}$ (rechts)

Der in Def. 1.1.1 eingeführte Relationsbegriff ist noch zu allgemein, um damit substantielle Mathematik betreiben zu können. Erst das Vorhandensein bestimmter Eigenschaften einer Relation macht diese zu einem interessanten Objekt. Eine Relation R auf X heißt in diesem Sinne

$$\begin{aligned} \textit{reflexiv} \quad &:\Leftrightarrow \quad (x,x) \in R \quad \text{für alle } x \in X \\ \textit{transitiv} \quad &:\Leftrightarrow \quad \big\langle (x,y) \in R \text{ und } (y,z) \in R \Rightarrow (x,z) \in R \big\rangle \quad \text{für alle } x,y,z \in X \\ \textit{symmetrisch} \quad &:\Leftrightarrow \quad \big\langle (x,y) \in R \Rightarrow (y,x) \in R \big\rangle \quad \text{für alle } x,y \in X. \end{aligned}$$

Beispiel 1.1.7. Die Relation R_2 aus Bsp. 1.1.4 ist reflexiv, transitiv, aber nicht symmetrisch. Die Relation R_3 aus Bsp. 1.1.5 ist reflexiv, transitiv und symmetrisch.

Wir führen nun einen Relationstyp ein, bei dem alle drei obigen Eigenschaften erfüllt sind. Mit diesem Begriff lässt sich präzise beschreiben, wann an sich verschiedene (mathematische) Objekte in einem vorgegebenen Zusammenhang als gleichwertig angesehen werden können. Dies ermöglicht ursprünglich zu große Mengen auf relevante Teilmengen zu reduzieren.

Definition 1.1.8 (Äquivalenzrelation)**.** Eine Relation A auf einer Menge X heißt *Äquivalenzrelation*, falls sie reflexiv, transitiv und symmetrisch ist. Für ein Paar $(x,y) \in A$ schreiben wir $x \sim y$ und nennen x und y *äquivalent*.

Beispiel 1.1.9. (1) Ist X eine beliebige Menge, so ist $\{(x,y) \in X \times X : x = y\}$ eine Äquivalenzrelation, die sog. *Identitätsrelation*. Ebenso ist auch das gesamte Produkt $X \times X$ eine Äquivalenzrelation, welche als *Allrelation* bezeichnet wird.

(2) Die Relation R_3 aus Bsp. 1.1.5 ist eine Äquivalenzrelation. Mit ihr kann man Computer nach ihrem Hersteller klassifizieren.

Für jedes $x \in X$ ist die Menge

$$[x] := \{y \in X : y \sim x\}$$

die von x erzeugte *Äquivalenzklasse* und ein $y \in [x]$ heißt *Repräsentant* von $[x]$.

Beispiel 1.1.10. (1) Für die Identitätsrelation aus Bsp. 1.1.9(1) sind die Äquivalenzklassen ein-elementig $[x] = \{x\}$ für alle $x \in X$. Die Allrelation besitzt dagegen genau eine Äquivalenzklasse $[x] = X$ für alle $x \in X$.

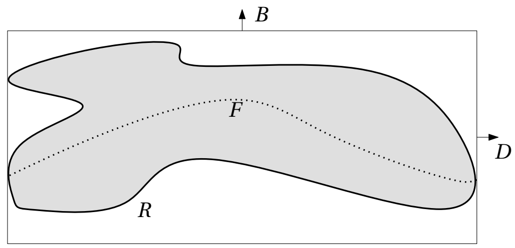

Abb. 1.5 Unterschied zwischen einer Relation $R \subseteq D \times B$ (graue Fläche) und einer Abbildung $F \subseteq D \times B$ (gepunktete Linie) von D nach B

(2) Bei der Äquivalenzrelation R_3 aus Bsp. 1.1.5 sind die Äquivalenzklassen die Menge aller Computer-Hersteller.

Aufgaben 1.1.11. (1) Man untersuche, ob die folgenden Relationen reflexiv, transitiv oder symmetrisch sind.

$$\begin{aligned}
&X_1 = \mathbb{R}, && R_1 = \big\{(x,y) \in X_1 \times X_1 : x < y\big\}, \\
& && R_2 = \big\{(x,y) \in X_1 \times X_1 : x \geq y\big\}, \\
&X_3 = \{\text{Menschen}\}, && R_3 = \big\{(x,y) \in X_3 \times X_3 : x \text{ ist verwandt mit } y\big\}, \\
&\begin{cases} X_4 = \{\text{Frauen}\}, \\ Y_4 = \{\text{Männer}\} \end{cases} && R_4 = \big\{(x,y) \in X_4 \times Y_4 : x \text{ ist Tochter von } y\big\}.
\end{aligned}$$

Im Falle von Äquivalenzrelationen bestimme man die Äquivalenzklassen.

(2) Gegeben sei die Menge $X = \{(m,n) \in \mathbb{Z} \times \mathbb{Z} : n \neq 0\}$ auf der wir die Relation

$$A := \big\{\big((m_1,n_1),(m_2,n_2)\big) \in X \times X : m_1 n_2 = m_2 n_1\big\}$$

definieren. Man untersuche, ob A eine Äquivalenzrelation ist und beschreibe ggf. die Äquivalenzklassen.

1.2 Abbildungen

Eine wesentliche Klasse von Relationen $F \subseteq D \times B$ sind Abbildungen oder Funktionen. Sie sind dadurch charakterisiert, dass es zu jedem $x \in D$ genau ein $y \in B$ mit $(x,y) \in F$ gibt (siehe Abb. 1.5). Präzise formuliert bedeutet dies:

Definition 1.2.1 (Abbildung, Funktion)**.** Eine Relation F zwischen nichtleeren Mengen D und B heißt *Abbildung* oder *Funktion* von D nach B, falls für alle $x \in D$ gilt:

(i) Es existiert ein $y \in B$ mit $(x,y) \in F$

(ii) Mit $y_1, y_2 \in B$ folgt aus $(x, y_1) \in F$ und $(x, y_2) \in F$, dass $y_1 = y_2$.

Die Menge D nennt man *Definitionsbereich* und B *Bildbereich* von F. Im Fall $D = B$ spricht man von einer *Abbildung von D in sich* oder einer *Selbstabbildung* auf D.

Bemerkung 1.2.2. Veranschaulicht man Funktionen auf (endlichen) Mengen D als gerichtete Graphen wie in Bem. 1.1.6 beschrieben, so geht von jedem Knoten $x \in D$ genau eine Kante ab. Daher ist von den in Abb. 1.4 dargestellten Relationen nur die linke eine Funktion.

Anstelle der obigen relationalen Notation $F \subseteq D \times B$, $(x, y) \in F$ schreibt man üblicherweise $f : D \to B$,

$$x \mapsto f(x), \qquad \text{oder} \qquad y := f(x)$$

und stellt dies wie in Abb. 1.6 dar. Mit einer weiteren nichtleeren Menge C und

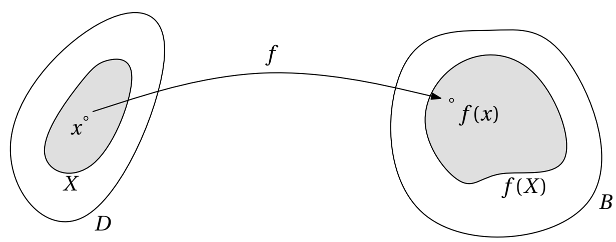

Abb. 1.6 Visualisierung einer Abbildung als Zuordnung $x \mapsto f(x)$, sowie der Bildmenge $f(X)$

einer Abbildung $g : B \to C$ ist die *Verknüpfung* oder *Komposition* von g und f definiert als

$$g \circ f : D \to C, \qquad g \circ f(x) := g(f(x)) \quad \text{für alle } x \in D; \tag{1.2a}$$

im Fall von Abbildungen f, g auf D gilt i.A. nicht $f \circ g = g \circ f$.

Statt einzelner Punkte $x \in D$ lassen sich auch ganze Mengen $X \subseteq D$ gemäß

$$f(X) := \{y \in B : \text{es gibt ein } x \in X \text{ mit } y = f(y)\}$$

abbilden; dann heißt $f(X) \subseteq B$ das *Bild* von X unter f (siehe auch Abb. 1.6). Das *Urbild* einer Menge $Y \subseteq B$ ist dagegen definiert durch (siehe Abb. 1.7)

$$f^{-1}(Y) := \{x \in D : f(x) \in Y\}.$$

Eine Abbildung $f : D \to B$ bezeichnet man als

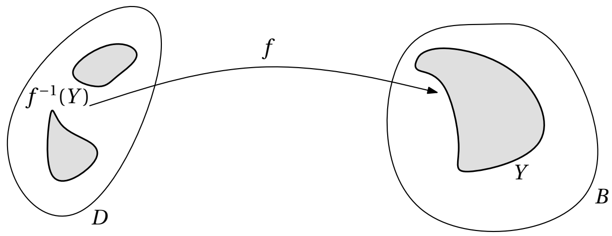

Abb. 1.7 Veranschaulichung des Urbilds $f^{-1}(Y)$ einer Menge $Y \subseteq B$

- *injektiv*, falls $f^{-1}(\{y\})$ für alle $y \in B$ höchstens ein Element enthält
- *surjektiv*, falls $f^{-1}(\{y\})$ für alle $y \in B$ mindestens ein Element enthält
- *bijektiv*, falls $f^{-1}(\{y\})$ für alle $y \in B$ genau ein Element enthält;

eine Abbildung ist genau dann bijektiv, wenn sie injektiv und surjektiv ist.

Beispiel 1.2.3 (identische Abbildung). Die *identische Abbildung* auf einer Menge $D \neq \emptyset$ ist definiert durch $\mathrm{id}_D : D \to D$, $\mathrm{id}_D(x) := x$; sie ist bijektiv.

Beispiel 1.2.4. Die Relation R_0 aus Bsp. 1.1.2 zwischen $X = \{\text{Land}\}$ und $Y = \{\text{Stadt}\}$ ist eine Funktion $r_0 : X \to Y$ mit $r_0(\text{Land}) :=$ Hauptstadt — zumindest unter der Voraussetzung, dass jedes Land genau eine Hauptstadt besitzt. Ihr Bild $r_0(X) \subseteq Y$ ist die Menge aller Hauptstädte und die Urbilder lauten

$$r_0^{-1}(\{s\}) = \begin{cases} \emptyset, & s \text{ ist keine Hauptstadt,} \\ \{l\}, & s \text{ ist Hauptstadt des Landes } l. \end{cases}$$

Folglich ist r_0 injektiv,[1] aber nicht surjektiv. Die obige Abbildung zwischen X und der Menge aller Hauptstädte als Bildbereich ist dagegen bijektiv.

Beispiel 1.2.5. Die Relation R_1 zwischen $\mathbb{R}$ und $[0,\infty)$ aus Bsp. 1.1.3 ist eine Abbildung und lässt sich auch schreiben als $r_1 : \mathbb{R} \to [0,\infty)$, $r_1(x) := |x|$. Für sie gilt etwa $r_1(\mathbb{R}) = [0,\infty)$ und $r_1^{-1}(\{y\}) = \{-y, y\}$ für alle $y \in [0,\infty)$, weshalb die Funktion $r_1 : \mathbb{R} \to [0,\infty)$ surjektiv, aber nicht injektiv ist. Betrachtet man r_1 dagegen als Abbildung von $\mathbb{R}$ in den Bildbereich $\mathbb{R}$, so ist r_1 nicht mehr surjektiv, denn es gilt $r_1^{-1}(\{y\}) = \emptyset$ für alle $y < 0$.

Beispiel 1.2.6 (ASCII-Code). Der ASCII-Code zur Codierung alpha-numerischer Zeichen ist gegeben durch eine bijektive Abbildung $f : \{0, 1, \ldots, 127\} \to \{\text{Zeichen}\}$.

Einfache Beispiele (etwa Bsp. 1.2.5) zeigen, dass die Umkehrrelation F^{-1} einer Abbildung $F \subseteq D \times B$ bzw. $f : D \to B$ nicht mehr unbedingt eine Abbildung ist.

Definition 1.2.7 (Umkehrabbildung)**.** Eine Abbildung $f : D \to B$ heißt *umkehrbar*, falls ihre Umkehrrelation F^{-1} wieder eine Abbildung ist. Für letztere schreibt man $f^{-1} : B \to D$ und nennt sie *Umkehrabbildung* von f.

[1] insofern unterschiedliche Länder unterschiedliche Hauptstädte besitzen

Bemerkung 1.2.8. Mit einer umkehrbaren Abbildung $f: D \to B$ ist auch ihre Umkehrabbildung $f^{-1}: B \to D$ umkehrbar und es gilt

$$f^{-1} \circ f = \mathrm{id}_D, \qquad f \circ f^{-1} = \mathrm{id}_B\,.$$

Korollar 1.2.9. *Eine Abbildung $f: D \to B$ ist genau dann umkehrbar, wenn f bijektiv ist. Für injektives f existiert die Umkehrfunktion nur auf $f(D)$.*

Beweis. Wegen $F^{-1} = \{(y,x) \in B \times D : (x,y) \in F\} = \{(y,x) \in B \times D : y = f(x)\}$ ist die Umkehrrelation F^{-1} genau dann eine Funktion, wenn es zu jedem $y \in B$ genau ein $x \in D$ gibt mit $y = f(x)$, d.h. wenn $f^{-1}(\{y\}) = \{x\}$ gilt. □

Aufgaben 1.2.10. (1) Für die folgenden Abbildungen $f_i: D_i \to B_i$, $i = 1,2,3,4$, bestimme man das Bild $f(D_i)$, untersuche sie auf Injektivität und Surjektivität:

$$\begin{array}{lll} D_1 = \mathbb{R}, & B_1 = \mathbb{R}, & f_1(x) := x^3 \\ D_2 = \mathbb{R}, & B_2 = \mathbb{R}, & f_2(x) := \sin(\pi x) \\ D_3 = \mathbb{R}, & B_3 = [-1,1], & f_3(x) := \sin(\pi x) \\ D_4 = \left[-\frac{1}{2}, \frac{1}{2}\right], & B_4 = [-1,1], & f_4(x) := \sin(\pi x). \end{array}$$

Geben Sie ggf. die Umkehrabbildung an.

(2) Es seien D, B nichtleer, $A_1, A_2 \subseteq D$, $B_1, B_2 \subseteq X$ und $f: D \to B$ eine Abbildung. Man zeige die Beziehungen

$$\begin{array}{ll} f(A_1 \cup A_2) = f(A_1) \cup f(A_2), & f^{-1}(B_1 \cup B_2) = f^{-1}(B_1) \cup f^{-1}(B_2), \\ f(A_1 \cap A_2) \subseteq f(A_1) \cap f(A_2), & f^{-1}(B_1 \cap B_2) = f^{-1}(B_1) \cap f^{-1}(B_2) \end{array}$$

und finde ein Beispiel, dass i.A. nicht $f(A_1 \cap A_2) = f(A_1) \cap f(A_2)$ gilt.

1.3 Matrizen

Wir führen kurz die Menge $\mathbb{C}$ der *komplexen Zahlen* ein: Darunter versteht man alle Paare $z = (x, y)$ reeller Zahlen $x, y \in \mathbb{R}$ versehen mit der *Addition*

$$z_1 + z_2 := (x_1 + x_2, y_1 + y_2)$$

und der *Multiplikation*

$$z_1 \cdot z_2 = z_1 z_2 := (x_1 x_2 - y_1 y_2, x_1 y_2 + x_2 y_1),$$

wobei $z_1 = (x_1, y_1), z_2 = (x_2, y_2) \in \mathbb{C}$. Hieraus ergibt sich leicht die Differenz bzw. der Quotient von z_1 und z_2 als

$$z_1 - z_2 := (x_1 - x_2, y_1 - y_2), \qquad \frac{z_1}{z_2} := \left(\frac{x_1 x_2 + y_1 y_2}{x_2^2 + y_2^2}, \frac{x_2 y_1 - x_1 y_2}{x_2^2 + y_2^2} \right),$$

wobei die letztere Beziehung nur im Fall $x_2^2 + y_2^2 \neq 0$ gilt. Ferner wird die erste Komponente des Paares $z = (x, y)$ als *Realteil* $\operatorname{Re} z = x$ und die zweite Komponente als *Imaginärteil* $\operatorname{Im} z = y$ von z bezeichnet. Aus rechentechnischer Sicht ist mit der Konvention $i^2 = -1$ die Schreibweise $z = (x, y) = x + i\,y$ hilfreich und üblich; zur Anschauung dient die bekannte komplexe Zahlenebene in Abb. 1.8.

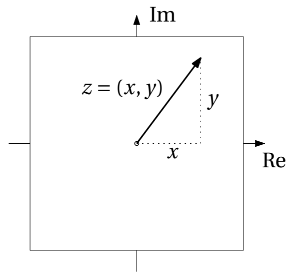

Abb. 1.8 Die Darstellung einer komplexen Zahl

$$z = (x, y) = x + i\,y$$

in der komplexen Zahlenebene $\mathbb{C} = \mathbb{R} \times \mathbb{R}$

Im Folgenden stehe $\mathbb{K}$ für eine der drei Mengen $\mathbb{Q}$ (rationale Zahlen), $\mathbb{R}$ (reelle Zahlen) oder $\mathbb{C}$; wir bezeichnen die Elemente von $\mathbb{K}$ als *Zahlen*.

Die primären „Rechenobjekte" in der linearen Algebra sind sog. Matrizen.

Definition 1.3.1 (Matrix)**.** Eine $m \times n$-*Matrix* über $\mathbb{K}$ ist ein rechteckiges Schema von Zahlen $a_{ij} \in \mathbb{K}$ der Form

$$A = (a_{ij})_{1 \le i \le m, 1 \le j \le n} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}.$$

Der erste Index $i \in \{1, \dots, m\}$ nummeriert die m Zeilen, der zweite Index $j \in \{1, \dots, n\}$ die n Spalten der Matrix A; das *Element* $a_{ij} \in \mathbb{K}$ steht daher in der $i$ten Zeile und der $j$ten Spalte. Für die Menge aller solchen Matrizen schreiben wir $\mathbb{K}^{m \times n}$. Für eine *quadratische Matrix* A gilt $m = n$ und die Zahlen a_{ii} heißen *Diagonalelemente* von A.

Beispiel 1.3.2 (n-Tupel und m-Spalte). Ein n-*Tupel* $x = (x_1, \dots, x_n)$ von Zahlen x_i aus $\mathbb{K}$ wird als $1 \times n$-Matrix interpretiert. Analog versteht man eine m-*Spalte*

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} \tag{1.3a}$$

als $m \times 1$-Matrix und wir identifizieren $\mathbb{K}^m = \mathbb{K}^{m\times 1}$.

Beispiel 1.3.3 (Kronecker-Symbol, Einheits- und Nullmatrix). Wir definieren das *Kronecker-Symbol* durch

$$\delta_{ij} := \begin{cases} 1, & i = j, \\ 0, & i \neq j \end{cases}$$

und $I_n := (\delta_{ij})_{1\le i,j\le n} \in \mathbb{K}^{n\times n}$ bezeichnet die *Einheits-* oder *Identitätsmatrix*. Bei der *Nullmatrix* $0 = (0)_{1\le i\le m, 1\le j\le n} \in \mathbb{K}^{m\times n}$ sind alle Elemente gleich $0 \in \mathbb{K}$.

Beispiel 1.3.4 (Diagonal- und Dreiecksmatrix). Man nennt eine quadratische Matrix $A = (a_{ij})_{1\le i,j\le n}$ *diagonal*, falls $a_{ij} = 0$ für alle $i \neq j$. Wir schreiben dann auch

$$A = \begin{pmatrix} a_{11} & 0 & \dots & \dots & 0 \\ 0 & a_{22} & 0 & \dots & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & 0 & \dots & a_{nn} \end{pmatrix} = \operatorname{diag}(a_{11}, a_{22}, \dots, a_{nn}).$$

Eine *obere Dreiecksmatrix* ist quadratisch und erfüllt $a_{ij} = 0$ für $i > j$, wogegen bei einer *unteren Dreiecksmatrix* $a_{ij} = 0$ für $i < j$ gilt. Sie sind also von der Form

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & \dots & a_{1n} \\ 0 & a_{22} & a_{23} & \dots & a_{2n} \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & 0 & \dots & a_{nn} \end{pmatrix} \quad \text{bzw.} \quad A = \begin{pmatrix} a_{11} & 0 & \dots & \dots & 0 \\ a_{21} & a_{22} & 0 & \dots & 0 \\ \vdots & & \ddots & & \vdots \\ a_{n1} & a_{n2} & 0 & \dots & a_{nn} \end{pmatrix}.$$

Für Matrizen lässt sich eine Reihe mathematischer Operationen erklären. So bezeichnet man die Abbildung $\cdot : \mathbb{K} \times \mathbb{K}^{m\times n} \to \mathbb{K}^{m\times n}$,

$$\alpha A := \alpha \cdot A := (\alpha a_{ij})_{1\le i\le m, 1\le j\le n} \tag{1.3b}$$

als *skalare Multiplikation* einer Zahl $\alpha \in \mathbb{K}$ mit einer Matrix $A \in \mathbb{K}^{m\times n}$ und wir schreiben $-A := (-1)\cdot A$. Weiter ist $+ : \mathbb{K}^{m\times n} \times \mathbb{K}^{m\times n} \to \mathbb{K}^{m\times n}$,

$$A + B := (a_{ij} + b_{ij})_{1\le i\le m, 1\le j\le n} \tag{1.3c}$$

die *Addition* zweier Matrizen $A, B \in \mathbb{K}^{n\times m}$ und $A - B := A + (-B)$ bezeichnet ihre *Subtraktion*. Man beachte, dass nur Matrizen mit jeweils gleicher Zeilen- und Spalten-Zahl addiert werden können.

Genau für $m \times n$-Matrizen A und $n \times p$-Matrizen B lässt sich auch eine *Multiplikation* $\cdot : \mathbb{K}^{m\times n} \times \mathbb{K}^{n\times p} \to \mathbb{K}^{m\times p}$,

$$AB := \left(\sum_{k=1}^{n} a_{ik} b_{kj} \right)_{1\le i\le m, 1\le j\le p}$$

von Matrizen definieren; das Produkt $C := AB$ ist eine $m \times p$-Matrix. Um Matrizen zu multiplizieren muss also die Spaltenzahl des ersten mit der Zeilenzahl des zweiten Faktors übereinstimmen.

Bemerkung 1.3.5. (1) Um Produkte von Matrizen $A \in \mathbb{K}^{m\times n}$ und $B \in \mathbb{K}^{n\times p}$ zu berechnen, eignet sich das Rechenschema

$$\begin{array}{c|c} & B \\ \hline A & C \end{array} \qquad \text{mit} \qquad C := \left(\sum_{k=1}^{n} a_{ik} b_{kj}\right)_{1\le i\le m,\, 1\le j\le p}.$$

(2) Ein wichtiger Spezialfall ist das Produkt $Ax \in \mathbb{K}^m$ einer $m \times n$-Matrix A mit einer n-Spalte x, welches sich ergibt zu

$$Ax = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \sum_{k=1}^{n} \begin{pmatrix} a_{1k} x_k \\ a_{2k} x_k \\ \vdots \\ a_{mk} x_k \end{pmatrix}.$$

Beispiel 1.3.6. Das Produkt C der Matrizen $A = \left(\begin{smallmatrix} 0 & 1 \\ 2 & 3 \end{smallmatrix}\right)$ und $B = \left(\begin{smallmatrix} 4 & 5 \\ 6 & 7 \end{smallmatrix}\right)$ ergibt sich aus

$$\begin{array}{cc|cc} & & 4 & 5 \\ & & 6 & 7 \\ \hline 0 & 1 & 0\cdot 4 + 1\cdot 6 & 0\cdot 5 + 1\cdot 7 \\ 2 & 3 & 2\cdot 4 + 3\cdot 6 & 2\cdot 5 + 3\cdot 7 \end{array} \qquad \text{zu} \qquad C = \begin{pmatrix} 6 & 7 \\ 26 & 31 \end{pmatrix},$$

wogegen man in umgedrehter Reihenfolge

$$\begin{pmatrix} 4 & 5 \\ 6 & 7 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 2 & 3 \end{pmatrix} = \begin{pmatrix} 10 & 19 \\ 14 & 27 \end{pmatrix}$$

erhält; die Multiplikation von Matrizen nicht kommutativ, d.h. i.A. gilt $AB \neq BA$.

Beispiel 1.3.7. (1) Für $A \in \mathbb{K}^{m\times n}$ gilt $I_m A = A = A I_n$.

(2) Für die Matrizen $A = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$ und $B = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right)$ gilt $AB = 0$ womit das Produkt von Matrizen nicht *nullteilerfrei* ist, d.h. es kann $AB = 0$ gelten, ohne dass einer der Faktoren gleich 0 ist.

(3) Das Produkt von $\left(\begin{smallmatrix} 0 & 1 & 2 \\ 3 & 4 & 5 \end{smallmatrix}\right)$ und $\left(\begin{smallmatrix} 6 & 7 \\ 8 & 9 \end{smallmatrix}\right)$ ist nicht definiert, $\left(\begin{smallmatrix} 6 & 7 \\ 8 & 9 \end{smallmatrix}\right)\left(\begin{smallmatrix} 0 & 1 & 2 \\ 3 & 4 & 5 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 21 & 34 & 47 \\ 27 & 44 & 61 \end{smallmatrix}\right)$ dagegen schon.

Beispiel 1.3.8 (RGB-Farbmodell). Im RGB-Farbmodell werden Farben durch ein Tripel $(r, g, b) \in \mathbb{R}^3$ beschrieben, wobei r für den Rot-, g für den Grün- und b für den Blauanteil steht. In diesem Sinne beschreibt $(1,0,0)$ die Farbe rot, $(0,0,1)$ die Farbe blau und $(1,1,0)$ „gelb“. In der im amerikanischen oder japanischen Fernsehen üblichen YIQ-Kodierung arbeitet man dagegen mit Tripeln (y, i, q), von denen die erste Komponente y alle im Schwarz-Weiß-Fernsehen nötigen Informationen enthält. Für die Umrechnung vom RGB-Farbraum in den YIQ-Farbraum gilt die Formel (auf eine Nachkommastelle gerundet)

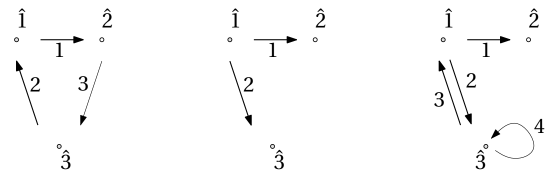

Abb. 1.9 Gerichtete Graphen ohne (links, mitte) und mit Schleifen (rechts)

$$\begin{pmatrix} y \\ i \\ q \end{pmatrix} = \begin{pmatrix} 0.3 & 0.6 & 0.1 \\ 0.6 & -0.3 & -0.3 \\ 0.2 & -0.5 & 0.3 \end{pmatrix} \begin{pmatrix} r \\ g \\ b \end{pmatrix}.$$

Beispiel 1.3.9 (Inzidenzmatrix). Gerichtete Graphen (siehe die obige Bem. 1.1.6) ohne Schleifen (d.h. kein Knoten wird mit sich selbst durch eine Kante verbunden, siehe Abb. 1.9 (rechts)) mit den Knoten $\{\hat{1}, \ldots, \hat{m}\}$ und den Kanten $\{1, \ldots, n\}$ lassen sich durch eine sog. *Inzidenzmatrix* $A \in \mathbb{R}^{m\times n}$ beschreiben, deren Elemente definiert sind durch

$$a_{ij} := \begin{cases} 1, & \text{von Knoten } \hat{\imath} \text{ geht die Kante } j \text{ aus,} \\ -1, & \text{in Knoten } \hat{\imath} \text{ mündet die Kante } j, \\ 0, & \text{Knoten } \hat{\imath} \text{ und Kante } j \text{ berühren sich nicht.} \end{cases}$$

Beispielsweise haben die in Abb. 1.9 links und mittig dargestellten gerichteten Graphen folgende respektiven Inzidenzmatrizen

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

Abschließend formulieren wir noch wichtigste Rechenregeln für Matrizen:

Satz 1.3.10 (Rechenregeln für Matrizen)**.** *Für Zahlen* $\alpha \in \mathbb{K}$ *und Matrizen* $A \in \mathbb{K}^{m\times n}$, $B \in \mathbb{K}^{n\times p}$ *gelten das* Distributiv-Gesetz

$$A(B + C) = AB + AC \quad \textit{für alle } C \in \mathbb{K}^{n\times p}$$

und die Assoziativ-Gesetze

$$(\alpha A)B = A(\alpha B), \qquad A(BC) = (AB)C \quad \textit{für alle } C \in \mathbb{K}^{p\times q}.$$

Beweis. Der Beweis sei dem interessierten Leser überlassen. □

Aufgaben 1.3.11. (1) Geben Sie eine Formel zur Umrechnung vom YIQ- in den RGB-Farbraum an.

(2) Die *Spur* einer Matrix $A = (a_{ij})_{1\le i,j\le n}$ ist definiert als Summe ihrer Diagonalelemente

$$\operatorname{tr} A := \sum_{i=1}^{n} a_{ii}.$$

Man zeige $\operatorname{tr} AB = \operatorname{tr} BA$ für alle $A, B \in \mathbb{K}^{n\times n}$.

(3) Beweisen Sie Satz 1.3.10.

1.4 Lineare Gleichungen

Als erste inner-mathematische Anwendung von Matrizen dient eine kompakte Notation für Systeme linearer Gleichungen.

Definition 1.4.1 (lineare Gleichung)**.** Es sei $A \in \mathbb{K}^{m\times n}$ und $b \in \mathbb{K}^m$. Dann bezeichnet man

$$Ax = b \tag{L_b}$$

als *lineares Gleichungssystem* mit m Gleichungen für die n Unbekannten $x_1, \dots, x_n \in \mathbb{K}$, oder kurz als *lineare Gleichung* in $\mathbb{K}^n$. Hierbei heißt A *Koeffizientenmatrix* und b *Inhomogenität* von (L_b). Im Fall $b \neq 0$ nennt man (L_b) *inhomogen* und erhält andernfalls die *homogene Gleichung*

$$Ax = 0. \tag{L_0}$$

Eine Lösung von (L_b) ist ein Element $x \in \mathbb{K}^n$ mit $Ax = b$ und

$$L_b := \left\{x \in \mathbb{K}^n : Ax = b\right\}$$

steht für die *Lösungsmenge* von (L_b).

Streng genommen ist (L_b) ein lineares *algebraisches* Gleichungssystem bzw. eine lineare *algebraische* Gleichung — im Gegensatz zu einer linearen Integral- oder Differenzial-Gleichung. Da wir jedoch nur mit Gleichungen der Form (L_b) zu tun haben, kann auf den Zusatz „algebraisch“ verzichtet werden.

Bemerkung 1.4.2. (1) Ausgeschrieben lautet die lineare Gleichung (L_b)

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2, \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m \end{cases} \tag{1.4a}$$

oder in noch unübersichtlicherer Form $\sum_{j=1}^{n} a_{ij}x_j = b_i$ für alle $1 \le i \le m$.

(2) Eine homogene Gleichung (L_0) hat stets die *triviale Lösung* $0 \in \mathbb{K}^n$.

Inhomogene lineare Gleichungen (L_b) müssen dagegen nicht unbedingt lösbar sein, wie triviale Beispiele (etwa $0x = 1$) demonstrieren. Bevor wir Lösungen explizit berechnen wollen, seien zwei charakteristische Eigenschaften linearer Gleichungen formuliert. Ihr Beweis folgt durch einfaches Einsetzen.

Satz 1.4.3 (Superpositionsprinzip)**.** *Es seien $x, y \in \mathbb{K}^n$ Lösungen der homogenen Gleichung (L_0). Dann ist auch $\alpha x + \beta y$ eine Lösung von (L_0), d.h.*

$$\alpha x + \beta y \in L_0 \quad \text{für alle } \alpha, \beta \in \mathbb{K}.$$

Beweis. Siehe Aufgabe 1.4.10(1). □

Für gegebene $x \in \mathbb{K}^n$ und $A \subseteq \mathbb{K}^n$ verwenden wir die Notation

$$x + A := \left\{ y \in \mathbb{K}^n : \text{es gibt ein } a \in A \text{ mit } y = x + a \right\}.$$

Satz 1.4.4. *Ist $\hat{x} \in \mathbb{K}^n$ eine Lösung der inhomogenen Gleichung (L_b), so gilt*

$$L_b = \hat{x} + L_0.$$

Beweis. Siehe Aufgabe 1.4.10(1). □

Nach diesen theoretischen Vorbereitungen widmen wir uns der expliziten Lösung linearer Gleichungen (L_b). Besonders einfach ist dies im Fall einer Diagonalmatrix $A \in \mathbb{K}^{n \times n}$. Gilt nämlich $a_{ii} \neq 0$, $1 \leq i \leq n$, so besitzt (L_b) die eindeutige Lösung $x \in \mathbb{K}^n$ mit Komponenten

$$x_i = \frac{b_i}{a_{ii}} \quad \text{für alle } 1 \leq i \leq n;$$

ist dagegen $a_{ii} = 0$ für einen Index $1 \leq i \leq n$, so besitzt (L_b) für $b_i = 0$ unendlich viele und für $b_i \neq 0$ keine Lösung.

Eine weitere einfach zu lösende Klasse linearer Gleichungen ist diejenige mit Koeffizientenmatrix in *Zeilen-Stufen-Form*; sie verallgemeinern obere Dreiecksmatrizen und wir definieren präzise: Eine Matrix $A \in \mathbb{K}^{m \times n}$ ist in *Zeilen-Stufen-Form*, falls in jeder Zeile gilt:

(i) Beginnt sie mit k Nullen, so stehen unter diesen Nullen lediglich Nullen
(ii) Unter dem ersten Element $\neq 0$ stehen ausnahmslos Nullen.

Man spricht von *strenger Zeilen-Stufen-Form*, insofern zusätzlich gilt:

(iii) Über dem ersten Element $\neq 0$ stehen lediglich Nullen.

Beispiel 1.4.5. (1) Obere Dreiecksmatrizen sind in Zeilen-Stufen-Form, Diagonalmatrizen in strenger Zeilen-Stufen-Form.

(2) Bezeichnet $*$ ein Element $\neq 0$, so gilt:

- $\left(\begin{smallmatrix} * & * & * \\ * & * & * \\ 0 & 0 & * \end{smallmatrix}\right)$, $\left(\begin{smallmatrix} 0 & * & * \\ 0 & * & * \\ 0 & 0 & * \end{smallmatrix}\right)$, $\left(\begin{smallmatrix} * & 0 & 0 \\ * & * & 0 \\ * & * & * \end{smallmatrix}\right)$ sind nicht in Zeilen-Stufen-Form
- $\left(\begin{smallmatrix} * & * & * & * \\ 0 & * & * & * \\ 0 & 0 & 0 & * \end{smallmatrix}\right)$ ist zwar in Zeilen-Stufen-Form, aber nicht in strenger
- $\left(\begin{smallmatrix} * & * & 0 & 0 \\ 0 & 0 & * & 0 \\ 0 & 0 & 0 & 0 \end{smallmatrix}\right)$ ist in strenger Zeilen-Stufen-Form

Zur Illustration der Lösung einer linearen Gleichung in Zeilen-Stufen-Form betrachten wir statt eines formalen Schemas das

Beispiel 1.4.6 (Rückwärts-Substitution). Die inhomogene lineare Gleichung

$$\begin{cases} x_1 + 2x_2 + 3x_3 + 4x_4 = 1, \\ x_2 + 2x_3 + 3x_4 = 1, \\ x_3 + 2x_4 = 1 \end{cases} \tag{1.4b}$$

besitzt die 3×4-Koeffizientenmatrix bzw. Inhomogenität

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

Wir lösen das Problem (1.4b) durch *Rückwärts-Substitution*: Aus der dritten Gleichung $x_3 + 2x_4 = 1$ sehen wir, dass $x_4 = t$ frei mit einem $t \in \mathbb{K}$ gewählt werden kann, was $x_3 = 1 - 2x_4 = 1 - 2t$ liefert. Nun können die bekannten Komponenten x_3, x_4 in die zweite Gleichung eingesetzt werden, d.h. $x_2 = 1 - 2x_3 - 3x_4 = t - 1$ und analog folgt schließlich aus der ersten Gleichung $x_1 = 1 - 2x_2 - 3x_3 - 4x_4 = 0$. Dies liefert die Lösungsmenge

$$L_b = \left\{ \begin{pmatrix} 0 \\ t-1 \\ 1-2t \\ t \end{pmatrix} : t \in \mathbb{K} \right\} =: \begin{pmatrix} 0 \\ -1 \\ 1 \\ 0 \end{pmatrix} + \mathbb{K} \begin{pmatrix} 0 \\ 1 \\ -2 \\ 1 \end{pmatrix}.$$

Es ist nun leicht einzusehen, dass sich die Lösungsmenge von (L_b) nicht ändert, wenn folgende Operationen auf (1.4a) angewandt werden:

- Vertauschen von Gleichungen
- Multiplikation einer Gleichung mit $\alpha \in \mathbb{K} \setminus \{0\}$
- Addition des α-fachen der $k$ten Gleichung zur $j$ten;

man spricht auch von *elementaren Zeilenoperationen*.

Diese Beobachtung ermöglicht uns, eine gegebene lineare Gleichung (L_b) sukzessive in eine simplere (etwa in Dreiecks-, Diagonal- oder Zeilen-Stufen-) Form zu transformieren, an welcher die Lösungsmenge direkt oder durch Rückwärts-Substitution abgelesen werden kann.

Das entsprechende Verfahren heißt *Gauß-Algorithmus* oder *Gauß'sches Eliminationsverfahren.* Zu seiner Beschreibung gehen wir davon aus, dass die erste Spalte der Matrix A von 0 verschieden ist; andernfalls sind die Variablen $x_1, \ldots, x_n$ entsprechend umzunummerieren. Ohne Sonderfälle zu berücksichtigen, kann das Gauß-Verfahren durch folgende Schritte beschrieben werden:

(1) Ordne die Gleichungen in (1.4a) so an, dass $a_{11} \neq 0$ ist. In der gängigen Kurznotation erhalten wir das formale Schema

$$\begin{array}{cccc|c} a_{11} & a_{12} & \ldots & a_{1n} & b_1 \\ a_{21} & a_{22} & \ldots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} & b_m \end{array}$$

(2) Subtrahiere von der $i$ten Gleichung, $2 \leq i \leq m$, in (1.4a) das $\frac{a_{i1}}{a_{11}}$-fache der ersten Gleichung, was auf die Gleichung

$$\begin{array}{cccc|c} a_{11} & a_{12} & \ldots & a_{1n} & b_1 \\ 0 & a_{22}^{(1)} & \ldots & a_{2n}^{(1)} & b_2^{(1)} \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & a_{m2}^{(1)} & \ldots & a_{mn}^{(1)} & b_m^{(1)} \end{array} \qquad \begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1, \\ A^{(1)}x^{(1)} = b^{(1)} \end{cases} \tag{1.4c}$$

mit $A^{(1)} \in \mathbb{K}^{(m-1)\times(n-1)}$ und einer Inhomogenität $b^{(1)} \in \mathbb{K}^{m-1}$ führt.

(3) Transformiere $A^{(1)}x^{(1)} = b^{(1)}$ entsprechend und fahre sukzessive fort, bis idealerweise eine Dreiecks- oder Zeilen-Stufen-Form entstanden ist.

(4) Löse das resultierende System durch Rückwärts-Einsetzen.

Beispiel 1.4.7. Als Kurzschreibweise für die homogene Gleichung

$$\begin{cases} x_1 + 2x_2 + 3x_3 = 0, \\ 4x_1 + 5x_2 + 6x_3 = 0, \\ 7x_1 + 8x_2 + 9x_3 = 0 \end{cases} \tag{1.4d}$$

verwenden wir das formale Schema

$$\begin{array}{ccc|cl} 1 & 2 & 3 & 0 & \\ 4 & 5 & 6 & 0 & |II-4I \\ 7 & 8 & 9 & 0 & |III-7I \end{array} \quad \Leftrightarrow \quad \begin{array}{ccc|cl} 1 & 2 & 3 & 0 & \\ 0 & -3 & -6 & 0 & |\cdot(-\frac{1}{3}) \\ 0 & -6 & -12 & 0 & |III-2II \end{array} \quad \Leftrightarrow \quad \begin{array}{ccc|c} 1 & 2 & 3 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{array}$$

womit (1.4d) äquivalent zur Gleichung

$$\begin{cases} x_1 + 2x_2 + 3x_3 = 0, \\ x_2 + 2x_3 = 0 \end{cases}$$

ist. Letztere wird durch Rückwärts-Substitution gelöst. Wir wählen $x_3 = t$ mit beliebigen $t \in \mathbb{K}$ und erhalten aus $x_2 = -2x_3 = -2t$, $x_1 = -2x_2 - 3x_3 = 4t - 3t = t$ die

Lösungsmenge von (1.4d) zu

$$L_0 = \left\{ t \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} : t \in \mathbb{K} \right\} =: \mathbb{K} \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}.$$

Abschließend präsentieren wir noch zwei theoretische Ergebnisse, deren Beweis auf dem Gauß-Verfahren basiert.

Satz 1.4.8. *Hat die homogene Gleichung* (L_0) *weniger Gleichungen als Unbekannte, d.h.* $m < n$*, so besitzt sie unendlich viele verschiedene Lösungen.*

Beweis. (I) Wir zeigen zunächst

$$(L_0) \text{ hat eine nichttriviale Lösung.} \qquad (*)$$

Dazu sei $m \in \mathbb{N}$ fest und wir führen eine vollständige Induktion über n aus. Im Induktionsanfang sei $m = 1$ und $n = 2$. Die Gleichung $a_{11}x_1 + a_{22}x_2 = 0$ besitzt für beliebige $a_{11}, a_{22} \in \mathbb{K}$ eine nichttriviale Lösung. Im Induktionsschritt $n-1 \to n$ sei die Behauptung $(*)$ für $n-1$ Unbekannte und $m < n-1$ Gleichungen zutreffend. Sind in der homogenen Gleichung (L_0) mit $m < n$ alle Elemente $a_{ij} = 0$, so ist jedes n-Tupel x eine Lösung. Wir dürfen also $A \neq 0$ annehmen und nach einer geeigneten Umnummerierung sei $a_{11} \neq 0$. Nun addieren wir zur $i$ten Gleichung, $2 \leq i$, das $-\frac{a_{i1}}{a_{ii}}$-fache der ersten Gleichung und erhalten

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = 0, \\ A^* \begin{pmatrix} x_2 \\ \vdots \\ x_n \end{pmatrix} = 0 \end{cases} \tag{1.4e}$$

mit einer $(m-1) \times (n-1)$-Matrix A^*. Das Teilsystem $A^*x^* = 0$ besitzt nach Induktionsvoraussetzung eine nichttriviale Lösung x^* mit Komponenten $x_2, \ldots, x_n$. Durch Einsetzen in die erste Gleichung von (1.4e) folgt ein Wert x_1 und damit eine nichttriviale Lösung x von (L_0).

(II) Da (L_0) nach Schritt (I) eine nichttriviale Lösung x hat, ist nach dem Superpositionsprinzip aus Satz 1.4.3 jedes αx, $\alpha \in \mathbb{K}$, eine Lösung. □

Satz 1.4.9. *Hat die inhomogene Gleichung* (L_b) *genauso viele Gleichungen wie Unbekannte, d.h.* $m = n$*, so gilt:*

(a) Ist $L_0 = \{0\}$*, so besitzt* (L_b) *für alle* $b \in \mathbb{K}^m$ *genau eine Lösung.*

(b) Besitzt (L_0) eine nichttriviale Lösung, so existieren entweder keine oder unendlich viele verschiedene Lösungen von (L_b).

Beweis. (a) Wir gehen mittels vollständiger Induktion vor. Für $n = 1$ ist die Behauptung sicherlich zutreffend. Im Induktionsschritt sei die Aussage (a) richtig für $n-1$. Da (L_0) nur die triviale Lösung besitzt, gilt sicherlich $A \neq 0$ und wir nummerieren die Variablen wieder derart um, dass $a_{11} \neq 0$. Dann wird zur $i$ten Gleichung, $2 \leq i$, in (1.4a) das $-\frac{a_{i1}}{a_{ii}}$-fache der ersten Gleichung addiert, was auf

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1, \\ A^* \begin{pmatrix} x_2 \\ \vdots \\ x_n \end{pmatrix} = b^* \end{cases} \tag{1.4f}$$

mit $A^* \in \mathbb{K}^{(n-1)\times(n-1)}$ und einer Inhomogenität $b^* \in \mathbb{K}^{n-1}$ führt. Die homogene Gleichung $A^*x^* = 0$ besitzt nur die triviale Lösung, denn sonst hätte (L_0) eine nichttriviale Lösung. Das Teilsystem $A^*x^* = b^*$ besitzt nach Induktionsvoraussetzung genau eine Lösung x^* mit Komponenten $x_2, \ldots, x_n$ und durch Einsetzen in die erste Gleichung von (1.4f) folgt ein eindeutiger Wert für x_1. Damit sind alle Komponenten der Lösung $x \in \mathbb{K}^n$ von (L_b) eindeutig bestimmt.

(b) Es sei $\hat{x} \in \mathbb{K}^n$ eine Lösung von (L_b) und $x \in \mathbb{K}^n$ eine nichttriviale Lösung von (L_0). Dann garantieren die Sätze 1.4.3 und 1.4.4, dass $\hat{x} + \alpha x$ die inhomogene Gleichung (L_b) für jedes Skalar $\alpha \in \mathbb{K}$ löst. In diesem Fall gibt es unendlich viele Lösungen. Als einzige Möglichkeit verbleibt, dass (L_b) nicht lösbar ist. □

Aufgaben 1.4.10. (1) Beweise die Sätze 1.4.3 und 1.4.4.

(2) Bestimmen Sie die Lösungsmenge der linearen Gleichung

$$\begin{cases} \alpha x_1 + \beta^2 x_2 = \alpha - \beta \\ \alpha^2 x_1 + \beta x_2 = \beta - \alpha \end{cases}$$

in Abhängigkeit der Parameter $\alpha, \beta \in \mathbb{R}$.

(3) Löse die folgenden linearen Gleichungen

$$\begin{cases} i x_1 + 2x_2 + 3x_3 = 0, \\ 4x_1 + 5x_2 + 6x_3 = 0, \end{cases} \qquad \begin{cases} x_1 + 2x_2 + 3x_3 = 0, \\ 4x_1 + 5x_2 + 6x_3 = 0, \\ 7x_1 + 8x_2 + 9x_3 = 0, \\ 10x_1 + 11x_2 + 12x_3 = 0. \end{cases}$$

mit dem Gauß-Algorithmus bzw. einer adäquaten Modifikation.

Kapitel 2
Lineare Räume

2.1 Algebraische Strukturen

Bezeichnet $M \neq \emptyset$ eine Menge und $F(M)$ die Menge aller Selbstabbildungen von M, so kann die Komposition $\circ$ aus (1.2a) als Abbildung $\circ : F(M) \times F(M) \to F(M)$ interpretiert werden — man spricht von einer *Verknüpfung*.

Derartige Paare aus einer Menge und einer Verknüpfung treten vielfach in der Mathematik auf. Ein besonders wichtiges Beispiel liefert

Definition 2.1.1 (Gruppe)**.** Eine *Gruppe* $(\mathbb{G}, *)$ ist eine nichtleere Menge $\mathbb{G}$ zusammen mit einer Abbildung $* : \mathbb{G} \times \mathbb{G} \to \mathbb{G}$ mit den Eigenschaften:

(G_1) Die Verknüpfung $*$ ist *assoziativ*, d.h. $a * (b * c) = (a * b) * c$ für $a, b, c \in \mathbb{G}$
(G_2) es existiert ein *neutrales Element* $e \in \mathbb{G}$ mit $a * e = a = e * a$ für alle $a \in \mathbb{G}$
(G_3) zu jedem $a \in \mathbb{G}$ gibt es ein *Inverses* $a^{-1} \in \mathbb{G}$ mit $a * a^{-1} = a^{-1} * a = e$

Bei einer *kommutativen* oder *Abel'schen Gruppe* wird zusätzlich

(G_4) $a * b = b * a$ für alle $a, b \in \mathbb{G}$

gefordert. Für eine *Halbgruppe* müssen lediglich (G_1) und (G_2) gelten.

Falls aus dem Kontext klar hervorgeht, wie die Operation $* : \mathbb{G} \times \mathbb{G} \to \mathbb{G}$ definiert ist, spricht man lediglich von einer Gruppe, ohne $*$ explizit anzugeben. Bei einer *additiven Gruppe* ist die Verknüpfung $*$ durch eine Addition gegeben und entsprechend ist eine *multiplikative Gruppe* erklärt.

Bemerkung 2.1.2. (1) Das neutrale Element $e \in \mathbb{G}$ einer Halbgruppe ist eindeutig bestimmt. In der Tat, bezeichnen $e_1, e_2 \in \mathbb{G}$ zwei neutrale Elemente in $\mathbb{G}$, so folgt vermöge (G_2), dass $e_2 = e_1 * e_2$ und $e_1 * e_2 = e_1$ ist, womit $e_1 = e_2$ resultiert.

(2) Zu gegebenem $a \in \mathbb{G}$ ist auch das inverse Element a^{-1} eindeutig. Für inverse Elemente $a_1^{-1}, a_2^{-1} \in \mathbb{G}$ zu a folgt nämlich

$$a_1^{-1} \overset{(G_2)}{=} a_1^{-1} * e \overset{(G_3)}{=} a_1^{-1} * (a * a_2^{-1}) \overset{(G_1)}{=} (a_1^{-1} * a) * a_2^{-1} \overset{(G_3)}{=} e * a_2^{-1} \overset{(G_2)}{=} a_2^{-1}.$$

(3) Entsprechend zeigt man $e = e^{-1}$ und $a = (a^{-1})^{-1}$.

Bemerkung 2.1.3 (Potenzen). Die *Potenzen* $a^n \in \mathbb{G}$ eines Elements $a \in \mathbb{G}$ einer multiplikativen Halbgruppe $\mathbb{G}$ sind rekursiv definiert durch

$$a^0 := e, \qquad a^{n+1} := a * (a^n) \quad \text{für alle } n \in \mathbb{N}_0;$$

in einer Gruppe setzen wir $a^n := (a^{-n})^{-1}$ für $n < 0$.

Beispiel 2.1.4. (1) Das Paar $(\mathbb{Z}, +)$ ist eine kommutative additive Gruppe mit neutralem Element 0 und dem zu $a \in \mathbb{Z}$ inversen Element $-a$. Dagegen ist $(\mathbb{Z}, \cdot)$ keine Gruppe, denn das multiplikative Inverse lässt sich innerhalb der ganzen Zahlen $\mathbb{Z}$ nicht definieren. Die natürlichen Zahlen $(\mathbb{N}, +)$ sind mangels eines neutralen Elements keine Gruppe, $(\mathbb{N}_0, +)$ aber eine additive Halbgruppe.

(2) Es sei $\mathbb{K} \in \{\mathbb{Q}, \mathbb{R}, \mathbb{C}\}$. Dann ist $(\mathbb{K}, +)$ eine kommutative additive Gruppe mit neutralem Element 0 und dem Inversen $-a$ zu $a \in \mathbb{K}$. Auch $(\mathbb{K} \setminus \{0\}, \cdot)$ ist eine kommutative Gruppe mit neutralem Element 1 und dem zu $a \in \mathbb{K} \setminus \{0\}$ Inversen $\frac{1}{a}$.

Beispiel 2.1.5. Mit $\mathbb{K} \in \{\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}\}$ bilden die Matrizen $(\mathbb{K}^{m \times n}, +)$ eine kommutative additive Gruppe mit neutralem Element 0 und dem Inversen $-A$ zu $A \in \mathbb{K}^{m \times n}$. Die quadratischen reellen, rationalen oder komplexen Matrizen $(\mathbb{K}^{n \times n} \setminus \{0\}, \cdot)$ bilden keine Gruppe, da etwa $\mathrm{diag}(1, 0) \neq 0$ kein multiplikatives Inverses besitzt.

Beispiel 2.1.6 (modulo). Es sei $p \geq 2$ ein ganze Zahl und $\mathbb{Z}_p := \{0, \ldots, p-1\}$. Für beliebige $a, b \in \mathbb{Z}$ gibt es vermöge der Division mit Rest eindeutige $m \in \mathbb{Z}$ und $k \in \mathbb{Z}_p$ derart, dass $a + b = mp + k$; wir schreiben auch

$$k = a + b \mod p \qquad \text{oder} \qquad k =: a +_p b.$$

Dann ist $(\mathbb{Z}_p, +_p)$ eine kommutative Gruppe mit neutralem Element $0 \in \mathbb{Z}_p$.

Beispiel 2.1.7 (symmetrische Gruppe). Es sei M eine nichtleere Menge und $S(M)$ bezeichne die Menge aller bijektiven Abbildungen $f : M \to M$. Dann ist die *symmetrische Gruppe* $(S(M), \circ)$ eine i.A. nicht-kommutative Gruppe mit der Identität id_M als neutralem Element und der Umkehrfunktion $f^{-1} : M \to M$ als Inversem zu f. Insbesondere für $M = \{1, \ldots, n\}$ schreiben wir $S_n := S(\{1, \ldots, n\})$. Die Menge aller nicht notwendig bijektiven Selbstabbildungen $F(M)$ bildet dagegen eine Halbgruppe bzgl. der Komposition.

Kürzungsregeln gelten auch in allgemeinen Gruppen:

Korollar 2.1.8 (Rechenregeln in Gruppen)**.** *Für alle $a, b, c \in \mathbb{G}$ gilt die Beziehung $(a * b)^{-1} = b^{-1} * a^{-1}$, wie auch die Implikationen*

$$a * b = a * c \quad \Rightarrow \quad b = c, \qquad a * b = e \quad \Rightarrow \quad a = b^{-1}.$$

Beweis. Es seien $a, b, c \in \mathbb{G}$. Wir zeigen zunächst, dass $b^{-1} * a^{-1}$ das inverse Element zu $a * b$ ist. Dazu gilt

$$\begin{aligned}(b^{-1} * a^{-1}) * (a * b) &\overset{(G_1)}{=} b^{-1} * (a^{-1} * (a * b)) \overset{(G_1)}{=} b^{-1} * ((a^{-1} * a) * b) \\ &\overset{(G_3)}{=} b^{-1} * (e * b) \overset{(G_2)}{=} b^{-1} * b \overset{(G_3)}{=} e\end{aligned}$$

und entsprechend $(a * b) * (b^{-1} * a^{-1}) = e$. Die erste Implikation ergibt sich aus der Voraussetzung durch

$$\begin{aligned}b &\overset{(G_2)}{=} e * b \overset{(G_3)}{=} (a^{-1} * a) * b \overset{(G_1)}{=} a^{-1} * (a * b) \\ &= a^{-1} * (a * c) \overset{(G_1)}{=} (a^{-1} * a) * c \overset{(G_3)}{=} e * c \overset{(G_2)}{=} c\end{aligned}$$

und für $c = b^{-1}$ folgt hieraus die letzte Implikation. □

Im Laufe der Vorlesung wird uns noch eine Vielzahl weiterer Gruppen begegnen. Wir fahren daher mit einer weiteren algebraischen Struktur fort, in der eine Addition und eine Multiplikation erklärt ist, welche die bekannten Rechenregeln aus den reellen Zahlen $\mathbb{R}$ abstrahiert:

Definition 2.1.9 (Körper)**.** Ein *Körper* $(\mathbb{K}, +, \cdot)$ ist eine Menge $\mathbb{K}$ mit mindestens zwei Elementen versehen mit den beiden *arithmetischen Operationen* $+ : \mathbb{K} \times \mathbb{K} \to \mathbb{K}$ (*Addition*) und $\cdot : \mathbb{K} \times \mathbb{K} \to \mathbb{K}$ (*Multiplikation*), so dass:

(K_1) $(\mathbb{K}, +)$ ist eine kommutative Gruppe mit neutralem Element 0 und dem zu $\alpha \in \mathbb{K}$ Inversem $-\alpha$, d.h. für alle $\alpha, \beta, \gamma \in \mathbb{K}$ gilt
- (K_1^1) $\alpha + (\beta + \gamma) = (\alpha + \beta) + \gamma$
- (K_1^2) $\alpha + 0 = 0 + \alpha = \alpha$
- (K_1^3) $\alpha + (-\alpha) = (-\alpha) + \alpha = 0$
- (K_1^4) $\alpha + \beta = \beta + \alpha$

(K_2) $(\mathbb{K} \setminus \{0\}, \cdot)$ ist eine kommutative Gruppe mit neutralem Element 1 und dem zu $\alpha \in \mathbb{K} \setminus \{0\}$ Inversem $\frac{1}{\alpha}$, d.h. für alle $\alpha, \beta, \gamma \in \mathbb{K} \setminus \{0\}$ gilt
- (K_2^1) $\alpha \cdot (\beta \cdot \gamma) = (\alpha \cdot \beta) \cdot \gamma$
- (K_2^2) $\alpha \cdot 1 = 1 \cdot \alpha = \alpha$
- (K_2^3) $\alpha \cdot \frac{1}{\alpha} = \frac{1}{\alpha} \cdot \alpha = 1$
- (K_2^4) $\alpha \cdot \beta = \beta \cdot \alpha$

(K_3) es gelten die *Distributiv-Gesetze*

$$\alpha \cdot (\beta + \gamma) = \alpha \cdot \beta + \alpha \cdot \gamma, \quad (\alpha + \beta) \cdot \gamma = \alpha \cdot \gamma + \beta \cdot \gamma \quad \text{für alle } \alpha, \beta, \gamma \in \mathbb{K}.$$

Falls die Definition der arithmetischen Operationen aus dem Kontext klar ist, schreibt man auch nur $\mathbb{K}$ für einen Körper. Üblicherweise wird $\alpha\beta := \alpha \cdot \beta$ abgekürzt und man verwendet *Subtraktion* $\alpha - \beta := \alpha + (-\beta)$ und *Division* $\frac{\alpha}{\beta} := \alpha \cdot \frac{1}{\beta}$

für $\beta \neq 0$. Ferner gilt die Konvention „Punkt vor Strich“ und man verwendet auch in allgemeinen Körpern die Summen- und Produktsymbole in bekannter Weise.

Beispiel 2.1.10. Die rationalen Zahlen $\mathbb{Q}$, die reellen Zahlen $\mathbb{R}$ und die komplexen Zahlen $\mathbb{C}$ sind Körper bzgl. der üblichen Addition und Multiplikation.

Beispiel 2.1.11 (Restklassenkörper modulo p). Mit einer fest gegebenen Primzahl $p \in \mathbb{N}$ definieren wir die Menge $\mathbb{Z}_p := \{0, \dots, p-1\}$. Dann gibt es für beliebige $\alpha, \beta \in \{0, \dots, p-1\}$ eindeutige Zahlen $m, n \in \mathbb{Z}$ und $k, l \in \mathbb{Z}_p$ derart, dass

$$\alpha + \beta = mp + k, \qquad \alpha\beta = np + l$$

(Division mit Rest). Mit der Addition (vgl. Bsp. 2.1.6) bzw. der Multiplikation

$$\alpha +_p \beta := k, \qquad \alpha \cdot_p \beta := l \tag{2.1a}$$

wird $(\mathbb{Z}_p, +_p, \cdot_p)$ dann ein Körper, der sog. *Restklassenkörper modulo p*. Für das konkrete Beispiel $\mathbb{Z}_2$ sind die arithmetischen Operationen durch die Tabellen

$$\begin{array}{c|cc} +_2 & 0 & 1 \\ \hline 0 & 0 & 1 \\ 1 & 1 & 0 \end{array} \qquad \begin{array}{c|cc} \cdot_2 & 0 & 1 \\ \hline 0 & 0 & 0 \\ 1 & 0 & 1 \end{array}$$

gegeben und im Fall $\mathbb{Z}_3$ gilt

$$\begin{array}{c|ccc} +_3 & 0 & 1 & 2 \\ \hline 0 & 0 & 1 & 2 \\ 1 & 1 & 2 & 0 \\ 2 & 2 & 0 & 1 \end{array} \qquad \begin{array}{c|ccc} \cdot_3 & 0 & 1 & 2 \\ \hline 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 2 \\ 2 & 0 & 2 & 1 \end{array}$$

Man erkennt, dass die Addition in $\mathbb{Z}_2$ der logischen XOR-Verknüpfung entspricht.

Körper sind insbesondere nullteilerfrei und erfüllen die Rechenregeln:

Korollar 2.1.12. *Ist $(\mathbb{K}, +, \cdot)$ ein Körper, so gilt für alle $\alpha, \beta, \gamma \in \mathbb{K}$, dass*

$$0 \cdot \alpha = \alpha \cdot 0 = 0, \qquad \beta \cdot (-\alpha) = -(\beta \cdot \alpha) = (-\beta) \cdot \alpha, \tag{2.1b}$$

$$(-1) \cdot \alpha = -\alpha, \qquad (-\alpha) \cdot (-\beta) = \alpha \cdot \beta \tag{2.1c}$$

und ferner die Implikation $\alpha \cdot \beta = 0 \Rightarrow \alpha = 0$ oder $\beta = 0$.

Bemerkung 2.1.13. Es gilt $1 \neq 0$, da die Annahme $1 = 0$ folgenden Widerspruch impliziert: Da $\mathbb{K}$ mindestens zwei Elemente enthält, gibt es ein $\alpha \in \mathbb{K}$, $\alpha \neq 0$ mit

$$\alpha \overset{(K_2^2)}{=} \alpha \cdot 1 = \alpha \cdot 0 \overset{(2.1b)}{=} 0.$$

Daher ist der Restklassenkörper modulo 2 aus Bsp. 2.1.11 der kleinste Körper.

Beweis. Wir wählen beliebige $\alpha, \beta, \gamma \in \mathbb{K}$. Es gilt

$$0 \cdot \alpha \overset{(K_1^2)}{=} (0+0) \cdot \alpha \overset{(K_3)}{=} 0 \cdot \alpha + 0 \cdot \alpha$$

und mittels Kor. 2.1.8 (für die Verknüpfung $+$, $a = b = 0 \cdot \alpha$ und $c = 0$) folgt sofort $0 \cdot \alpha = 0$; die Kommutativität (K_2^4) liefert auch $\alpha \cdot 0 = 0$.

Aus der oben gezeigten Behauptung resultiert

$$(-\beta) \cdot \alpha + \beta \cdot \alpha \overset{(K_3)}{=} (-\beta + \beta) \cdot \alpha \overset{(K_1^3)}{=} 0 \cdot \alpha = 0$$

und wir wenden hierauf Kor. 2.1.8 mit der Verknüpfung $+$ und $a = (-\beta) \cdot \alpha$, $b = \beta \cdot \alpha$ an. Dies liefert $-(\beta \cdot \alpha) = (-\beta) \cdot \alpha$ und die Identität $\beta \cdot (-\alpha) = -(\beta \cdot \alpha)$ folgt analog.

Die Beziehung $(-1) \cdot \alpha = -\alpha$ resultiert aus dem eben Gezeigten mit $\beta = 1$ und

$$(-1) \cdot \alpha = 1 \cdot (-\alpha) \overset{(K_2^2)}{=} -\alpha$$

Die verbleibende Gleichung in (2.1c) ergibt sich mittels Bem. 2.1.2(3) aus

$$(-\alpha) \cdot (-\beta) \overset{(2.1b)}{=} -(\alpha \cdot (-\beta)) \overset{(2.1b)}{=} -(-(\alpha \cdot \beta)) = \alpha \cdot \beta.$$

Schließlich sei noch $\alpha \cdot \beta = 0$. Wir nehmen $\alpha \neq 0$ und $\beta \neq 0$ an, was aber mit

$$1 \overset{(K_2^3)}{=} \frac{1}{\beta} \cdot \frac{1}{\alpha} \cdot \alpha \cdot \beta \overset{(2.1b)}{=} 0$$

einen Widerspruch zur Folge hat (vgl. Bem. 2.1.13). □

Aufgaben 2.1.14. (1) Beweisen Sie Bem. 2.1.2(3).

(2) Es sei $\Delta_u \subseteq \mathbb{K}^{n \times n}$ die Menge aller oberen Dreiecksmatrizen, deren Diagonalelemente sämtlich ungleich 0 sind. Ist Δ_u eine Gruppe bezüglich der Matrix-Multiplikation?

(3) Ist auch $\mathbb{Z}_4 := \{0, 1, 2, 3\}$ mit den beiden in (2.1a) definierten arithmetischen Operationen $+_4, \cdot_4$ ein Körper?

(4) Man zeige, dass $\mathbb{K} := \left\{a + \sqrt{3}b : a, b \in \mathbb{Q}\right\}$ ein Körper bzgl. der von $\mathbb{R}$ vererbten arithmetischen Operationen ist.

(5) Welche der in den Abschnitten 1.3 und 1.4 gemachten Aussagen gelten, wenn $\mathbb{K}$ durch einen beliebigen Körper ersetzt wird?

(6) Über den ganzen Zahlen $\mathbb{Z}$ und dem Körper $\mathbb{Z}_3$ mit den in (2.1a) definierten arithmetischen Operationen löse man die lineare Gleichung

$$\begin{cases} x_1 + 2x_2 = 0, \\ x_2 + 2x_3 = 1, \\ 2x_1 + x_3 = 2. \end{cases}$$

2.2 Vektorräume

Die in der linearen Algebra primär studierten Objekte sind Vektoren, also Elemente von linearen Räumen und Abbildungen zwischen solchen.

Definition 2.2.1 (linearer Raum, Vektorraum)**.** Es sei $\mathbb{K}$ ein Körper. Ein *Vektorraum* oder *linearer Raum* $(X,+,\cdot)$ (über $\mathbb{K}$) ist eine nichtleere Menge X mit den arithmetischen Operationen

(a) *Addition* $+ : X \times X \to X$ derart, dass $(X,+)$ eine kommutative Gruppe ist; deren neutrales Element wird als *Null* oder *Nullvektor* 0 bezeichnet

(b) *Skalare Multiplikation* $\cdot : \mathbb{K} \times X \to X$ derart, dass für alle $\alpha, \beta \in \mathbb{K}$ und $x, y \in X$ gilt

- (V_1) $\alpha \cdot (x+y) = \alpha \cdot x + \alpha \cdot y$ (*Distributiv-Gesetz*)
- (V_2) $(\alpha+\beta) \cdot x = \alpha \cdot x + \beta \cdot x$ (*Distributiv-Gesetz*)
- (V_3) $(\alpha\beta) \cdot x = \alpha \cdot (\beta \cdot x)$ (*Assoziativ-Gesetz*)
- (V_4) $1 \cdot x = x$

Die Elemente von $\mathbb{K}$ heißen *Skalare* und die von X *Vektoren.*

Oftmals sind die arithmetischen Operationen $+, \cdot$ aus dem Kontext klar und man schreibt nur X für einen Vektorraum. Der Einfachheit halber wird das Produkt $\alpha \cdot x$ häufig als αx geschrieben, man definiert die *Subtraktion* $x - y := x + (-y)$ und folgt der Konvention „Punkt vor Strich“ auch in linearen Räumen.

Beispiel 2.2.2. Es sei $(\mathbb{K},+,\cdot)$ ein Körper.

(0) Der triviale Raum $\{0\}$ besteht nur aus der Null 0.

(1) Weiter ist $\mathbb{K}$ ein linearer Raum über sich selbst.

(2) Die Menge aller $m \times n$-Matrizen $\mathbb{K}^{m \times n}$ ist ein linearer Raum über $\mathbb{K}$ bezüglich der in (1.3b) und (1.3c) definierten arithmetischen Operationen. Ein n-Tupel $(x_1,\dots,x_n) \in \mathbb{K}^{1\times n}$ bezeichnet man auch als *Zeilenvektor* und eine m-Spalte (1.3a) als *Spaltenvektor*.

In der Codierungstheorie und Kryptographie spielen die folgenden Vektorräume eine große Rolle.

Beispiel 2.2.3. Es sei p eine Primzahl und $n \in \mathbb{N}$. Dann sind die n-Spalten $\mathbb{Z}_p^n$ in $\mathbb{Z}_p$ mit der komponentenweisen Addition $+_p$ und der entsprechenden skalaren Multiplikation $\cdot_p$ aus Bsp. 2.1.11 ein linearer Raum. Insbesondere für $\mathbb{Z}_2^2$ erhält man die Additions- bzw. Multiplikationstabellen

$+_2$	$\binom{0}{0}$	$\binom{1}{0}$	$\binom{0}{1}$	$\binom{1}{1}$
$\binom{0}{0}$	$\binom{0}{0}$	$\binom{1}{0}$	$\binom{0}{1}$	$\binom{1}{1}$
$\binom{1}{0}$	$\binom{1}{0}$	$\binom{0}{0}$	$\binom{1}{1}$	$\binom{0}{1}$
$\binom{0}{1}$	$\binom{0}{1}$	$\binom{1}{1}$	$\binom{0}{0}$	$\binom{1}{0}$
$\binom{1}{1}$	$\binom{1}{1}$	$\binom{0}{1}$	$\binom{0}{0}$	$\binom{0}{0}$

$\cdot_2$	$\binom{0}{0}$	$\binom{1}{0}$	$\binom{0}{1}$	$\binom{1}{1}$
0	$\binom{0}{0}$	$\binom{0}{0}$	$\binom{0}{0}$	$\binom{0}{0}$
1	$\binom{0}{0}$	$\binom{1}{0}$	$\binom{0}{1}$	$\binom{1}{1}$

Beispiel 2.2.4 (Lösungsmenge L_0). Mit dem Superpositionsprinzip in Satz 1.4.3 ist die Lösungsmenge L_0 einer homogenen Gleichung (L_0) ein Vektorraum über $\mathbb{K}$. Für die Lösungsmenge L_b inhomogener Systeme trifft diese Aussage nicht zu.

Beispiel 2.2.5 (Funktionen- und Folgenräume). Es sei Ω eine nichtleere Menge und X ein linearer Raum über dem Körper $\mathbb{K}$. Dann ist $F(\Omega, X) := \{u : \Omega \to X\}$ ein Vektorraum über $\mathbb{K}$ mit punktweise definierten arithmetischen Operationen

$$(u+v)(t) := u(t) + v(t), \qquad (\alpha \cdot u)(t) := \alpha u(t) \quad \text{für alle } t \in \Omega;$$

die Menge $F(\Omega, X)$ wird als *Funktionenraum* bezeichnet. Falls der Definitionsbereich Ω durch die natürlichen Zahlen $\mathbb{N}$ oder die ganzen Zahlen $\mathbb{Z}$ gegeben ist, spricht man von einem *Folgenraum*.

Als nächstes formulieren wir einfache Eigenschaften von Vektorräumen:

Korollar 2.2.6. *Ist $(X, +, \cdot)$ ein linearer Raum über $\mathbb{K}$, so gilt für alle Skalare $\alpha, \beta \in \mathbb{K}$ und Vektoren $x, y \in X$:*

(a) $0 \cdot x = \alpha \cdot 0 = 0$
(b) *Falls $\alpha \cdot x = 0$, so folgt $\alpha = 0 \in \mathbb{K}$ oder $x = 0 \in X$*
(c) $(-\alpha) \cdot x = \alpha \cdot (-x) = -(\alpha \cdot x)$
(d) $\alpha \cdot (x - y) = \alpha \cdot x - \alpha \cdot y$ *und* $(\alpha - \beta) \cdot x = \alpha \cdot x - \beta \cdot x$

Beweis. Es seien $\alpha \in \mathbb{K}$ und $x \in X$.
(a) Es gilt $0 \cdot x = (0+0) \cdot x = 0 \cdot x + 0 \cdot x$ wegen (V_2). Nach Def. 2.2.1(a) existiert zum Vektor $z := 0 \cdot x$ ein Vektor $-z$ mit $0 \cdot x + (-z) = 0$ und wir erhalten

$$0 = 0 \cdot x + (-z) = (0 \cdot x + 0 \cdot x) + (-z) = 0 \cdot x + (0 \cdot x + (-z)) = 0 \cdot x + 0 = 0 \cdot x$$

und die Beziehung $\alpha \cdot 0 = 0$ folgt analog.
(b) Es gelte $\alpha \cdot x = 0$ mit $\alpha \neq 0$ und wir müssen $x = 0$ zeigen. Wegen $\alpha \neq 0$ existiert $\frac{1}{\alpha}$. Nach der bewiesenen Aussage (a) folgt $\frac{1}{\alpha}(\alpha \cdot x) = \frac{1}{\alpha} \cdot 0 = 0$ und andererseits

$$\frac{1}{\alpha}(\alpha \cdot x) = \left(\frac{1}{\alpha}\alpha\right) \cdot x = 1 \cdot x = x.$$

Den Beweis der Aussagen (c) und (d) überlassen wir dem Leser. □

Wir betrachten nun Teilmengen linearer Räume, die selbst wieder Vektorräume bzgl. der vererbten arithmetischen Operationen sind.

Definition 2.2.7 (Unterraum)**.** Eine nichtleere Teilmenge $Y \subseteq X$ eines linearen Raumes $(X, +, \cdot)$ über $\mathbb{K}$ heißt *Unterraum* von X, falls gilt

$$\alpha_1 y_1 + \alpha_2 y_2 \in Y \quad \text{für alle } \alpha_1, \alpha_2 \in \mathbb{K},\ y_1, y_2 \in Y.$$

Bemerkung 2.2.8. Jeder lineare Raum X hat die trivialen Unterräume $\{0\}$ und X. Ferner enthält jeder Unterraum von X das Element $0 \in X$.

Beispiel 2.2.9 (stetige und stetig-differenzierbare Funktionen). Es sei $I \subseteq \mathbb{R}$ ein Intervall. Die Menge aller stetigen Funktionen $C(I, \mathbb{R}^n)$ auf I mit Bildern in $\mathbb{R}^n$ ist ein Unterraum von $F(I, \mathbb{R}^n)$. Ebenso sind die stetig-differenzierbaren Funktionen $C^1(I, \mathbb{R}^n)$ ein Unterraum von $C(I, \mathbb{R}^n)$ und $F(I, \mathbb{R}^n)$.

Beispiel 2.2.10 (Polynome). Mit gegebenem Körper $\mathbb{K}$ definieren wir den *Raum aller Polynome* (über $\mathbb{K}$) durch

$$P(\mathbb{K}) := \left\{ p \in F(\mathbb{K}, \mathbb{K}) \,|\, \exists n \in \mathbb{N}_0 : \exists \alpha_0, \dots, \alpha_n \in \mathbb{K} : p(t) \equiv \sum_{k=0}^{n} \alpha_k t^k \quad \text{auf } \mathbb{K} \right\};$$

seine Elemente heißen *Polynome* und die α_k deren *Koeffizienten*. Dann ist $P(\mathbb{K})$ ein Unterraum von $F(\mathbb{K}, \mathbb{K})$. Der *Grad* $\deg p$ eines Polynoms $p \in P(\mathbb{K})$ ist der maximale Index $k \in \mathbb{N}_0$ für den $\alpha_k \neq 0$ gilt. Für jedes $m \in \mathbb{N}_0$ sind die Mengen

$$P_m(\mathbb{K}) := \{ p \in P(\mathbb{K}) : \deg p \le m \}$$

Unterräume von $P(\mathbb{K})$, wogegen $\{ p \in P(\mathbb{K}) : \deg p = m \}$ für $m \neq 0$ kein Unterraum ist. Ferner ist jedes $P_n(\mathbb{K})$ Unterraum von $P_m(\mathbb{K})$ für $0 \le n \le m$.

Es ist einfach zu sehen, dass die Vereinigung von Unterräumen i.A. kein linearer Raum ist. Anders verhält es sich mit Schnitten und Summen:

Satz 2.2.11 (Schnitte und Summen von Unterräumen)**.** *Ist I eine nichtleere Indexmenge und $\{Y_i\}_{i \in I}$ eine Familie von Unterräumen von X, so gilt:*

(a) Der Durchschnitt $\bigcap_{i \in I} Y_i$ ist ein Unterraum von X.
(b) Für endliches I ist die Summe

$$\sum_{i \in I} Y_i := \left\{ \sum_{i \in I} y_i \in X : y_i \in Y_i \text{ with } i \in I \right\}$$

der kleinste Unterraum von X, welcher jedes Y_i enthält.

Für $I = \{1, \dots, n\}$ schreibt man auch $Y_1 + \dots + Y_n = \sum_{i \in \mathbb{I}} Y_i$.

Beweis. (a) Es seien $\alpha, \beta \in \mathbb{K}$ und $x, y \in \bigcap_{i \in I} Y_i$. Dann gilt $x, y \in Y_i$ für alle $i \in I$ und da jedes Y_i ein Unterraum ist, folgt $\alpha x + \beta y \in Y_i$ für jeden Index $i \in I$. Dies impliziert $\alpha x + \beta y \in \bigcap_{i \in I} Y_i$.

(b) Wir zeigen zunächst, dass die Summe $Y := \sum_{i \in I} Y_i$ ein Unterraum von X ist. Dazu sei $x = \sum_{i \in I} x_i$ und $y = \sum_{i \in I} y_i$ mit $x_i, y_i \in Y_i$ und wir erhalten

$$\alpha x + \beta y = \alpha \sum_{i\in I} x_i + \beta \sum_{i\in I} y_i = \sum_{i\in I} \underbrace{(\alpha x_i + \beta y_i)}_{\in Y_i} \in Y \quad \text{für alle } \alpha, \beta \in \mathbb{K}.$$

Also ist Y ein Unterraum, der sicherlich jedes Y_i, $i \in I$, enthält. Es bleibt zu verifizieren, dass Y der kleinste solche Unterraum ist. Dazu sei $Z \subseteq X$ ein weiterer jedes Y_i umfassender Unterraum. Für $x_i \in Y_i$ ist dann auch $x_i \in Z$ für alle $i \in I$, da X_i in Z enthalten ist. Aus der Unterraum-Eigenschaft von Z resultiert die Inklusion $\sum_{i\in I} x_i \in Z$ und folglich ist $Y \subseteq Z$. Somit ist Satz 2.2.11 bewiesen. □

2.3 Lineare Abhängigkeit

Gegeben sei eine nichtleere Menge $\mathscr{S}$ von Vektoren aus einem linearen Raum X über dem Körper $\mathbb{K}$. Existieren zu gegebenem $x \in X$ dann endlich viele *Koeffizienten* $\alpha_i \in \mathbb{K}$ und Vektoren $x_i \in \mathscr{S}$, $1 \le i \le n$, mit

$$x = \sum_{i=1}^{n} \alpha_i x_i,$$

so bezeichnet man x als *Linearkombination* der Vektoren aus $\mathscr{S}$.

Definition 2.3.1 (Spann)**.** Es sei $\mathscr{S} \subseteq X$ gegeben. Der *Spann* oder die *lineare Hülle* $\operatorname{span}\mathscr{S}$ von $\mathscr{S}$ ist die Menge aller Linearkombinationen der Vektoren aus $\mathscr{S}$. Ferner setzt man $\operatorname{span}\emptyset = \{0\}$.

Beispiel 2.3.2. Für endliches $\mathscr{S} = \{x_1, \dots, x_n\}$ ist

$$\operatorname{span}\mathscr{S} = \left\{ \sum_{i=1}^{n} \alpha_i x_i \in X : \alpha_i \in \mathbb{K} \right\}$$

Insbesondere im Fall $\mathbb{K} = \mathbb{R}$ und $e_1 = \binom{1}{0}$, $e_2 = \binom{0}{1}$ gilt $\operatorname{span}\{e_1, e_2\} = \mathbb{R}^2$. Ebenso erhält man $\operatorname{span}\{x_1, x_2\} = \mathbb{R}^2$ mit den Vektoren $x_1 = \binom{1}{1}$, $e_2 = \binom{1}{-1}$, wogegen für $y_1 = \binom{1}{1}$, $y_2 = \binom{2}{2}$ lediglich $\operatorname{span}\{y_1, y_2\} = \mathbb{R}\binom{1}{1} \subset \mathbb{R}^2$ gilt.

Beispiel 2.3.3 (Monome). Die Polynome $m_n(t) := t^n$, $n \in \mathbb{N}_0$, heißen *Monome*. Dann lassen sich die Polynome über $\mathbb{K}$ darstellen als lineare Hülle der Monome, d.h. $\operatorname{span}\{m_n\}_{n\in\mathbb{N}_0} = P(\mathbb{K})$ und insbesondere $\operatorname{span}\{m_0, \dots, m_n\} = P_n(\mathbb{K})$. Ebenso erlauben im Fall $\mathbb{K} = \mathbb{R}$ die geraden resp. ungeraden Polynome die Darstellung

$$\begin{aligned} \operatorname{span}\{m_{2n}\}_{n\in\mathbb{N}_0} &= \{p \in P(\mathbb{R}) : p(t) \equiv p(-t) \quad \text{auf } \mathbb{R}\}, \\ \operatorname{span}\{m_{2n+1}\}_{n\in\mathbb{N}_0} &= \{p \in P(\mathbb{R}) : p(t) \equiv -p(-t) \quad \text{auf } \mathbb{R}\}. \end{aligned}$$

Proposition 2.3.4. *Es sei $\mathscr{S} \subseteq X$ nichtleer. Dann ist die lineare Hülle* $\operatorname{span}\mathscr{S}$ *der kleinste $\mathscr{S}$ umfassende Unterraum von X.*

Beweis. Mit $x, y \in \mathscr{S}$ ist auch jede Linearkombination $\alpha x + \beta y$, $\alpha, \beta \in \mathbb{K}$, in $\operatorname{span}\mathscr{S}$ enthalten. Also ist $\operatorname{span}\mathscr{S}$ ein Unterraum von X. Weiter enthält $\operatorname{span}\mathscr{S}$ die Vektoren aus $\mathscr{S}$ und somit $\mathscr{S} \subseteq \operatorname{span}\mathscr{S}$. Es sei nun $Y \subseteq X$ ein Unterraum mit $x \in Y$ für sämtliche $x \in \mathscr{S}$. Dann liegen auch sämtliche Linearkombinationen von Vektoren aus $\mathscr{S}$ in Y. Also ist $\operatorname{span}\mathscr{S}$ in Y enthalten und $\operatorname{span}\mathscr{S}$ ist der kleinste Unterraum, der alle Vektoren aus $\mathscr{S}$ enthält. □

Korollar 2.3.5. *Ist x eine Linearkombination von Vektoren aus $\mathscr{S} \subseteq X$, so gilt* $\operatorname{span}\mathscr{S} = \operatorname{span}(\mathscr{S} \cup \{x\})$.

Beweis. Wir zeigen die Behauptung durch zwei Inklusionen:

($\subseteq$) Es ist klar, dass $\operatorname{span}\mathscr{S} \subseteq \operatorname{span}(\mathscr{S} \cup \{x\})$ gilt.

($\supseteq$) Als Linearkombination von Vektoren aus $\mathscr{S}$ liegt x auch in $\operatorname{span}\mathscr{S}$. Demnach ist $\operatorname{span}\mathscr{S}$ derjenige Unterraum, welcher $\mathscr{S}$ und $\{x\}$ umfasst. Damit folgt mit Prop. 2.3.4 sofort $\operatorname{span}(\mathscr{S} \cup \{x\}) \subseteq \operatorname{span}\mathscr{S}$. □

Im Folgenden wird der in der linearen Algebra grundlegende Begriff der linearen Unabhängigkeit eingeführt. Als Vorbetrachtung betrachten wir Vektoren $x, y \in X$ mit $y = \xi x$ für ein $\xi \in \mathbb{K} \setminus \{0\}$. Dann ist x eine Linearkombination von y, nämlich $x = \xi^{-1} y$ und es gilt $\xi x + (-1) y = 0$ — man nennt x, y linear abhängig.

Definition 2.3.6 (linear unabhängig)**.** Eine endliche Menge $\{x_1, \ldots, x_n\}$ von Vektoren aus X heißt *linear unabhängig*, falls die Implikation

$$\sum_{k=1}^{n} \xi_k x_k = 0 \quad \Rightarrow \quad \xi_k = 0 \quad \text{für alle } 1 \leq k \leq n$$

gilt. Eine beliebige Menge $\mathscr{S} \subseteq X$ wird *linear unabhängig* genannt, insofern jede endliche Teilmenge von $\mathscr{S}$ linear unabhängig ist; die leere Menge wird als linear unabhängig angesehen. Schließlich bezeichnet man eine Teilmenge von X als *linear abhängig*, falls sie nicht linear unabhängig ist.

Man nennt auch die Vektoren $x_1, x_2, \ldots$ linear unabhängig, falls die aus ihnen geformte Menge $\{x_1, x_2, \ldots\}$ linear unabhängig ist.

Bemerkung 2.3.7. (1) Lineare Abhängigkeit einer endlichen Menge $\{x_1, \ldots, x_n\}$ bedeutet, dass eine nichttriviale Darstellung der Null aus den Vektoren $x_1, \ldots, x_n$ existiert. Man kann also

$$\sum_{k=1}^{n} \xi_k x_k = 0 \tag{2.3a}$$

schreiben, ohne dass alle Koeffizienten $\xi_1, \ldots, \xi_n \in \mathbb{K}$ verschwinden.

(2) Jede Obermenge einer linear abhängigen Menge ist linear abhängig. Jede Teilmenge einer linear unabhängigen Menge ist linear unabhängig.

Beispiel 2.3.8. Die Menge $\{0\}$ ist linear abhängig, wogegen $\{x\}$, $x \neq 0$, linear unabhängig ist.

Proposition 2.3.9. *Es sei $\mathscr{S} \subseteq X$ nichtleer und $x, x_1, \ldots, x_n \in X$.*

(a) Ist $\mathscr{S} = \{x_1, \ldots, x_n\}$ linear abhängig, so lässt sich mindestens ein Vektor aus $\mathscr{S}$ als Linearkombination der übrigen darstellen.
(b) Für jede Linearkombination x aus $\mathscr{S}$ ist $\mathscr{S} \cup \{x\}$ linear abhängig.

Beweis. (a) Weil $\{x_1, \ldots, x_n\}$ linear abhängig ist, besitzt 0 die Darstellung (2.3a) in welcher nicht alle Koeffizienten $\xi_1, \ldots, \xi_n$ verschwinden. Also existiert ein Index $1 \leq k^* \leq n$ mit $\xi_{k^*} \neq 0$ und damit

$$x_{k^*} = -\xi_{k^*}^{-1} \sum_{\substack{k=1,\\ k\neq k^*}}^{n} \xi_k x_k = \sum_{\substack{k=1,\\ k\neq k^*}}^{n} \left(-\xi_{k^*}^{-1} \xi_k\right) x_k$$

(b) Mit $x = \sum_{k=1}^{n} \xi_k x_k$ ist $\sum_{k=1}^{n} \xi_k x_k - x$ eine nichttriviale Darstellung der 0. □

In den linearen Räumen $X = \mathbb{K}^m$ besteht ein enger Zusammenhang zwischen der linearen Unabhängigkeit von Mengen $\mathscr{S} = \{a_1, \ldots, a_n\} \subseteq \mathbb{K}^m$ und linearen homogenen Gleichungen. Mit der $m \times n$-Matrix $A := (a_1, \ldots, a_n)$ ist die Beziehung

$$\sum_{k=1}^{n} \xi_k a_k = 0$$

(vgl. (2.3a)) nämlich äquivalent zur linearen homogenen Gleichung

$$Ax = 0 \in \mathbb{K}^d, \qquad x := \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}. \tag{2.3b}$$

Demzufolge ist die Menge $\mathscr{S}$ genau dann linear unabhängig, wenn $Ax = 0$ nur die triviale Lösung besitzt. Aus Satz 1.4.8 (in Verbindung mit Aufgabe 2.1.14(5)) erhalten wir daher, dass mehr als m Vektoren im $\mathbb{K}^m$ stets linear abhängig sind.

Beispiel 2.3.10. (1) Für die *kanonischen Einheitsvektoren* in $\mathbb{K}^m$,

$$e_1 := \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \qquad e_2 := \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \qquad \dots \qquad e_m := \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

gilt mit obiger Terminologie $A = I_m$. Also besitzt $Ax = 0$ nur die triviale Lösung und $\{e_1, \dots, e_m\}$ ist linear unabhängig.

(2) Es sei $\lambda \in \mathbb{R}$. Um die lineare Unabhängigkeit der Vektoren

$$x_1 := \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \qquad x_2 := \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix}, \qquad x_3 := \begin{pmatrix} 7 \\ 8 \\ \lambda \end{pmatrix}$$

in $\mathbb{R}^3$ zu untersuchen, betrachten wir die Gleichung (2.3b) mit

$$A := \begin{pmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & \lambda \end{pmatrix}$$

und lösen sie mit dem in Bsp. 1.4.7 beschriebenen Schema:

$$\begin{array}{ccc|cl} 1 & 4 & 7 & 0 & \\ 2 & 5 & 8 & 0 & |2I - II \\ 3 & 6 & \lambda & 0 & |3I - III \end{array} \quad \Leftrightarrow \quad \begin{array}{ccc|cl} 1 & 4 & 7 & 0 & \\ 0 & 3 & 6 & 0 & \\ 0 & 6 & 21-\lambda & 0 & |III - 2II \end{array} \quad \Leftrightarrow \quad \begin{array}{ccc|c} 1 & 4 & 7 & 0 \\ 0 & 3 & 6 & 0 \\ 0 & 0 & 9-\lambda & 0 \end{array}$$

Hieraus erkennt man sofort, dass die Gleichung $Ax = 0$ für $\lambda \neq 9$ nur die triviale Lösung besitzt und $\{x_1, x_2, x_3\}$ linear unabhängig ist. Für $\lambda = 9$ ist $\{x_1, x_2, x_3\}$ dagegen linear abhängig; es gilt etwa $x_1 - 2x_2 + x_3 = 0$.

Satz 2.3.11. *Eine Menge $\mathscr{S} \subseteq X$ ist genau dann linear unabhängig, wenn jedes $x \in \operatorname{span} \mathscr{S}$ auf nur eine Art (bis auf Glieder mit Null-Koeffizienten) als Linearkombination von Vektoren aus $\mathscr{S}$ dargestellt werden kann.*

Beweis. ($\Rightarrow$) Die Menge $\mathscr{S}$ sei linear unabhängig und ein beliebig gegebenes $x \in X$ sei auf zwei Arten als Linearkombination von Vektoren aus $\mathscr{S}$ dargestellt. Indem wir geeignete Null-Koeffizienten einführen, sei also

$$x = \sum_{k=1}^{n} \xi_k x_k = \sum_{k=1}^{n} \eta_k x_k$$

mit $\xi_k, \eta_k \in \mathbb{K}$ und $x_k \in \mathscr{S}$. Dies impliziert sofort $0 = x - x = \sum_{k=1}^{n} (\xi_k - \eta_k) x_k$ und weil $\{x_1, \dots, x_n\}$ linear unabhängig ist, resultiert $\xi_k = \eta_k$ für alle $1 \leq k \leq n$. Die oben gegebenen Linearkombinationen von x mittels Vektoren aus $\mathscr{S}$ unterscheiden sich also nur durch Glieder mit Koeffizient 0.

($\Leftarrow$) Umgekehrt besitze $x \in \operatorname{span}\mathscr{S}$ genau eine Darstellung als Linearkombination von Vektoren $x_k \in \mathscr{S}$, nämlich

$$x = \sum_{k=1}^{n} \xi_k x_k \quad \text{mit } \xi_k \in \mathbb{K}.$$

Andererseits gilt für den Nullvektor $0x_1 + \ldots + 0x_n = 0$. Daraus folgt wegen der vorausgesetzten eindeutigen Darstellung $\xi_k = 0$ für $1 \le k \le n$. □

Aufgaben 2.3.12. (1) Es sei $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$. Mit den durch $m_k(t) := t^k$ definierten Monomen $m_k : \mathbb{K} \to \mathbb{K}$ zeige man, dass die Familie $\{m_k\}_{k \in \mathbb{N}_0}$ linear unabhängig in $P(\mathbb{K})$ ist.

(2) Eine Matrix $A \in \mathbb{K}^{m \times n}$ besitze die Zeilen $a^1, \ldots, a^m \in \mathbb{K}^{1 \times n}$ und $B \in \mathbb{K}^{m \times n}$ die Zeilen $b^1, \ldots, b^m \in \mathbb{K}^{1 \times n}$. Man zeige $\operatorname{span}\{a^1, \ldots, a^m\} = \operatorname{span}\{b^1, \ldots, b^n\}$, insofern B aus A durch elementare Zeilenoperationen hervorgegangen ist.

2.4 Basis und Dimension

Wieder sei X ein linearer Raum über $\mathbb{K}$.

Definition 2.4.1 (Basis)**.** Eine Menge $\mathscr{X} \subseteq X$ heißt *Basis* von X, falls $\mathscr{X}$ linear unabhängig und $X = \operatorname{span}\mathscr{X}$ ist.

Eine Menge $\mathscr{X}$ mit $X = \operatorname{span}\mathscr{X}$ wird auch *Erzeugendensystem* von X genannt. Man sagt X ist *endlich erzeugt*, falls X ein endliches Erzeugendensystem besitzt.

Beispiel 2.4.2. Die Basis des Nullraumes $\{0\}$ ist die leere Menge.

Beispiel 2.4.3 (Standardbasis). Die mittels der kanonischen Einheitsvektoren aus Bsp. 2.3.10(1) gebildete Menge $\mathscr{E}_m := \{e_1, \ldots, e_m\}$ ist eine Basis des Raumes $\mathbb{K}^m$, die sog. *Standardbasis*. Damit ist $\mathbb{K}^m$ endlich erzeugt.

Beispiel 2.4.4 (Polynome). Für die beiden Körper $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$ sind die Monome $\mathscr{M}_n := \{m_0, \ldots, m_n\}$ aus Bsp. 2.3.3 eine Basis der Polynome $P_n(\mathbb{K})$ vom maximalen Grad n und $\{m_n\}_{n \in \mathbb{N}_0}$ bildet eine Basis aller Polynome $P(\mathbb{K})$. Also ist $P_n(\mathbb{K})$ endlich erzeugt, $P(\mathbb{K})$ dagegen nicht. Im Fall des Körpers $\mathbb{K} = \mathbb{Z}_2$ ist dagegen $t^2 - t = 0$ für alle $t \in \mathbb{Z}_2$ und in $P_2(\mathbb{Z}_2)$ ist damit $0 = m_2 - m_1$, d.h. $0 \in P_2(\mathbb{K}_2)$ besitzt eine nichttriviale Darstellung.

Lemma 2.4.5. *Es sei $\mathscr{S} \subseteq X$ linear unabhängig. Gilt dann $x \notin \operatorname{span}\mathscr{S}$, so ist auch $\mathscr{S} \cup \{x\}$ linear unabhängig.*

Beweis. Es ist nachzuweisen, dass jede endliche Teilmenge von $\mathscr{S} \cup \{x\}$ linear unabhängig ist. Dazu sei $\{x_1, \ldots, x_n, x\}$ eine solche Menge und ferner

$$\sum_{k=1}^{n} \xi_k x_k + \eta x = 0$$

eine Darstellung der Null. Wäre $\eta \neq 0$, so ließe sich x als Linearkombination der Vektoren x_k, $1 \leq k \leq n$, darstellen; dies widerspricht der vorausgesetzten Inklusion $x \notin \operatorname{span} \mathscr{S}$. Also gilt $\eta = 0$. Da aber $\{x_1, \dots x_n\}$ linear unabhängig ist, folgt auch $\xi_k = 0$ für alle Indizes $1 \leq k \leq n$, was zu beweisen war. □

In trivialer Weise ist X ein Erzeugendensystem von sich selbst. Unser Interesse besteht dagegen in möglichst „kleinen“ solchen Mengen.

Satz 2.4.6. *Mit nichtleerem $\mathscr{X} \subseteq X$ sind folgende Aussagen äquivalent:*

(a) $\mathscr{X}$ ist eine Basis von X
(b) Jeder Vektor $x \in X$ lässt sich eindeutig als Linearkombination von Vektoren aus $\mathscr{X}$ darstellen
(c) $\mathscr{X}$ ist maximal linear unabhängig, *d.h. $\mathscr{X}$ ist linear unabhängig und für jedes $x \in X \setminus \mathscr{X}$ ist $\mathscr{X} \cup \{x\}$ linear abhängig*
(d) $\mathscr{X}$ ist ein minimales Erzeugendensystem, *d.h. keine echte Teilmenge von $\mathscr{X}$ ist ein Erzeugendensystem.*

Bemerkung 2.4.7 (Koordinaten). Besitzt $x \in X$ bzgl. der Basis $\mathscr{X} = \{x_1, \dots, x_n\}$ die nach Satz 2.4.6(b) eindeutige Darstellung

$$x = \sum_{k=1}^{n} \xi_k x_k \qquad \text{mit Koeffizienten} \qquad \xi_k \in \mathbb{K},$$

so bezeichnet man das n-Tupel $(\xi_1, \dots, \xi_n)$ als seine *Koordinaten*.

Beweis. Die Äquivalenz der Aussagen (a) und (b) folgt sofort aus Satz 2.3.11.

$(c) \Rightarrow (a)$ Für eine maximal linear unabhängige Familie $\mathscr{X} \subseteq X$ zeigen wir $\operatorname{span} \mathscr{X} = X$. Angenommen, es gibt ein $x \in X \setminus \operatorname{span} \mathscr{X}$, so wäre nach Lemma 2.4.5 auch $\operatorname{span} \mathscr{X} \cup \{x\}$ linear unabhängig im Widerspruch zur Maximalität.

$(c) \Rightarrow (a)$ Als Basis von X ist $\mathscr{X}$ per definitionem linear unabhängig. Da jeder Vektor $x \in X$ eine Linearkombination von Vektoren aus $\mathscr{X}$ ist, muss $\mathscr{X} \cup \{x\}$ nach Prop. 2.3.9(b) linear abhängig sein. Also ist $\mathscr{X}$ maximal.

$(a) \Rightarrow (d)$ Es sei $\mathscr{X}$ eine Basis und folglich auch ein Erzeugendensystem. Wir zeigen seine Minimalität indirekt: Dazu nehmen wir an, es gäbe eine echte Teilmenge $\mathscr{X}' \subset \mathscr{X}$, die ebenfalls ein Erzeugendensystem von X ist. Der Vektor $x \in \mathscr{X} \setminus \mathscr{X}'$ habe nun die Darstellung $x = \sum_{k=1}^{n} \xi_k x_k$ mit Koeffizienten $\xi_k \in \mathbb{K}$ und Vektoren $x_k \in \mathscr{X}'$. Folglich ist

$$-x + \sum_{k=1}^{n} \xi_k x_k = 0$$

eine nichttriviale Darstellung der 0 und $\mathscr{X}$ muss linear abhängig sein. Dies widerspricht der Basis-Eigenschaft von $\mathscr{S}$.

$(d) \Rightarrow (a)$ Sei $\mathscr{X}$ ein minimales Erzeugendensystem und wir zeigen, dass $\mathscr{X}$ linear unabhängig ist. Andernfalls gäbe es nach Prop. 2.3.9(a) ein $x \in \mathscr{X}$, das sich als Linearkombination der Vektoren aus $\mathscr{X}$ darstellen ließe. Dann wäre aber auch $\mathscr{X} \setminus \{x\}$ ein Erzeugendensystem. Dies widerspricht der Minimalität von $\mathscr{X}$. □

Im Folgenden beschränken wir uns auf endlich erzeugbare lineare Räume X. Der interessierte Leser sei jedoch darauf hingewiesen, dass der nachfolgende Satz 2.4.8 und die Prop. 2.4.9 auch für beliebige Vektorräume gelten. Allerdings basiert deren Beweis dann auf einem tiefliegenden Mengen- bzw. Ordnungstheoretischen Resultat, dem Lemma von Zorn (vgl. etwa [Bos08, S. 42]), auf welches wir hier nicht eingehen wollen.

Satz 2.4.8. *Jedes endliche Erzeugendensystem eines Vektorraumes enthält eine Basis. Insbesondere hat jeder endlich erzeugbare lineare Raum eine Basis.*

Beweis. Es sei $\mathscr{X}$ ein endliches Erzeugendensystem eines linearen Raumes X. Ist $\mathscr{X}$ keine Basis, so kann $\mathscr{X}$ nicht minimal sein und es existiert eine echte Teilmenge $\mathscr{X}^1 \subset \mathscr{X}$, die ebenfalls Erzeugendensystem ist. Ist $\mathscr{X}^1$ wiederum keine Basis, so existiert erneut eine echte Teilmenge $\mathscr{X}^2 \subset \mathscr{X}^1$, die X erzeugt. Durch Iteration erhält man eine echt absteigende Folge von Teilmengen

$$\ldots \subset \mathscr{X}^2 \subset \mathscr{X}^1 \subset \mathscr{X}$$

Diese Folge bricht nach endlich vielen Schritten ab, da $\mathscr{X}$ endlich ist, d.h. es gibt ein minimales $\mathscr{X}^k$. Dieses $\mathscr{X}^k$ ist nach Satz 2.4.6 eine Basis von X. □

Proposition 2.4.9. *Ist X endlich erzeugbar und $\mathscr{S} \subseteq X$ linear unabhängig, so existiert eine Basis von X, welche $\mathscr{S}$ als Teilmenge enthält.*

Beweis. Für jedes Erzeugendensystem $\bar{\mathscr{S}} = \{x_1, \ldots, x_n\}$ ist auch $\mathscr{S} \cup \bar{\mathscr{S}}$ ein solches. Wir konstruieren nun eine aufsteigende Folge von Mengen $\mathscr{S}^k$ derart, dass

$$\mathscr{S} = \mathscr{S}^0 \subset \mathscr{S}^1 \subset \ldots \subset \mathscr{S} \cup \bar{\mathscr{S}}.$$

Dazu definieren wir

$$\mathscr{S}^1 := \begin{cases} \mathscr{S}^0, & x_1 \in \operatorname{span} \mathscr{S}^0, \\ \mathscr{S}^0 \cup \{x_1\}, & \text{sonst}, \end{cases} \qquad \mathscr{S}^2 := \begin{cases} \mathscr{S}^1, & x_1 \in \operatorname{span} \mathscr{S}^1, \\ \mathscr{S}^1 \cup \{x_2\}, & \text{sonst} \end{cases}$$

$$\ldots \qquad \mathscr{S}^n := \begin{cases} \mathscr{S}^{n-1}, & x_n \in \operatorname{span} \mathscr{S}^{n-1}, \\ \mathscr{S}^{n-1} \cup \{x_n\}, & \text{sonst} \end{cases}$$

und nach Lemma 2.4.5 sind die Mengen $\mathscr{S}^1,\dots,\mathscr{S}^n$ linear unabhängig. Wegen $x_1,\dots,x_n \in \operatorname{span}\mathscr{S}^n$ ist $\mathscr{S}^n$ ein Erzeugendensystem und folglich eine Basis. □

Das nächste Hilfsresultat wird als *Austauschsatz von Steinitz* bezeichnet.

Lemma 2.4.10. *Ist* $\{x_1,\dots,x_p\}$ *linear unabhängig und* $\{y_1,\dots,y_n\}$ *ein Erzeugendensystem von* X*, so gilt* $p \le n$ *und nach einer geeigneten Umnummerierung der* y_k *ist* $\{x_1,\dots,x_p,y_{p+1},\dots,y_n\}$ *ein Erzeugendensystem von* X.

Beweis. Wir beweisen durch Induktion die Existenz einer Nummerierung der $y_1,\dots,y_n$ derart, dass $\mathscr{S}' := \{x_1,\dots,x_k,y_{k+1},\dots,y_n\}$ ein Erzeugendensystem von X ist und $k \le p$. Für $k=0$ ist nichts zu beweisen. Für $k>0$ gilt nach Induktionsannahme (und Umnummerierung) der y_i, dass $\mathscr{S}'' := \{x_1,\dots,x_{k-1},y_k,\dots,y_n\}$ ein Erzeugendensystem ist. Der Vektor x_k ist nun darstellbar als Linearkombination

$$x_k = \sum_{i=1}^{k-1} \xi_i x_i + \sum_{j=k}^{n} \eta_j y_j$$

mit $x_i, y_j \in \mathscr{S}''$. Wegen $x_k \ne 0$ gilt $\eta_j \ne 0$ für mindestens ein $k \le j \le n$, denn andernfalls wären $x_1,\dots,x_{k-1},x_k$ linear abhängig. Insbesondere ist $k \le n$. Vermöge einer Umnummerierung der y_j können wir $\eta_k \ne 0$ annehmen und es folgt

$$y_k = \sum_{i=1}^{k-1} \left(-\tfrac{\xi_i}{\eta_k}\right) x_i + \frac{1}{\eta_k} x_k + \sum_{j=k+1}^{n} \left(-\tfrac{b_j}{b_k}\right) y_j .$$

Dies garantiert $y_k \in \operatorname{span}\{x_1,\dots,x_k,y_{k+1},\dots,y_n\}$ und Kor. 2.3.5 impliziert

$$\operatorname{span}\mathscr{S}'' = \operatorname{span}\left(\mathscr{S}'' \cup \{x_k\}\right) = \operatorname{span}\left(\mathscr{S}'' \cup \{y_k\}\right) = \operatorname{span}\mathscr{S}'.$$

Mit $\mathscr{S}''$ ist also auch $\mathscr{S}'$ ein Erzeugendensystem von X, was zu beweisen war. □

Satz 2.4.11 (Dimension)**.** *Falls* X *eine Basis von* n *Elementen besitzt, so enthält jede Basis von* X *genau* n *Elemente. Wir bezeichnen* n *als* Dimension *von* X *und schreiben* $\dim X = n$.

Bemerkung 2.4.12. Ein linearer Raum X heißt *unendlich dimensional* (symbolisch $\dim X = \infty$), falls er kein endliches Erzeugendensystem besitzt; andernfalls heißt er *endlich dimensional*.

Beweis. Es seien $\{x_1,\dots,x_n\}$ und $\{y_1,\dots,y_m\}$ Basen eines linearen Raumes X. Mit Lemma 2.4.10 folgt dann $m \le n$ und $n \le m$, also $m = n$. □

Beispiel 2.4.13. Der Nullraum erfüllt $\dim\{0\} = 0$. Für die weiteren bisher betrachteten linearen Räume gilt beispielsweise

$$\dim \mathbb{K}^n = n, \qquad \dim \mathbb{K}^{m\times n} = mn, \qquad \dim P_n(\mathbb{R}) = n+1$$

und $\dim P(\mathbb{R}) = \dim C^1(\mathbb{R},\mathbb{R}) = \dim C(\mathbb{R},\mathbb{R}) = \dim F(\mathbb{R},\mathbb{R}) = \infty$.

Beispiel 2.4.14. Die komplexen Zahlen $\mathbb{C}$ sind ein 2-dimensionaler Vektorraum über dem Körper $\mathbb{R}$ und ein 1-dimensionaler Raum über $\mathbb{C}$.

Korollar 2.4.15. *In linearen Räumen X mit $n := \dim X < \infty$ gilt:*

(a) Weniger als n Vektoren aus X sind kein Erzeugendensystem
(b) Mehr als n Vektoren aus X sind linear abhängig
(c) Jedes Erzeugendensystem mit n Elementen ist eine Basis von X
(d) Jede linear unabhängige Menge mit n Elementen ist eine Basis von X.

Beweis. (a) Jedes Erzeugendensystem enthält laut Satz 2.4.8 eine Basis. Für jedes aus weniger als n Vektoren bestehendem Erzeugendensystem gäbe es dann auch eine Basis mit weniger als n Elementen. Dies widerspricht Satz 2.4.11.

(b) Laut Prop. 2.4.9 ist jede linear unabhängige Menge Teil einer Basis. Somit hätte man mit einer linear unabhängigen Familie von mehr als n Vektoren ebenso eine Basis mit mehr als n Elementen — im Widerspruch zu Satz 2.4.11.

(c) Ein Erzeugendensystem enthält wegen Satz 2.4.8 eine Basis und ist aufgrund von Satz 2.4.11 bereits eine solche.

(d) Mit Prop. 2.4.9 ist eine linear unabhängige Familie Teilmenge einer Basis und wegen Satz 2.4.11 eine Basis. ◻

Korollar 2.4.16. *Für jeden Unterraum Y eines endlich dimensionalen linearen Raumes X ist $\dim Y \le \dim X$; Gleichheit gilt genau für $X = Y$.*

Beweis. Es sei $\mathscr{Y}$ eine Basis von Y und $n = \dim X$. Insbesondere ist $\mathscr{Y}$ linear unabhängig und damit Teilmenge einer Basis von X (vgl. Prop. 2.4.9). Somit kann die Basis $\mathscr{Y}$ maximal n Vektoren enthalten. Enthält $\mathscr{Y}$ gerade n Vektoren, dann ist $\mathscr{Y}$ aufgrund von Kor. 2.4.15(d) auch eine Basis von X. ◻

Aufgaben 2.4.17. (1) Es sei $X = \mathbb{R}^4$. Untersuchen Sie die Familie auf lineare Unabhängigkeit:

$$\mathscr{S} := \left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ -4 \\ 11 \\ 2 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \\ -3 \\ 0 \end{pmatrix} \right\}$$

Geben Sie ferner eine Teilfamilie an, welche eine Basis von $\operatorname{span}\mathscr{S}$ ist.

(2) Ergänzen Sie die Familie

$$\mathscr{S} := \left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \\ -3 \\ 0 \end{pmatrix} \right\}$$

zu einer Basis von $\mathbb{R}^4$.

(3) Man untersuche, ob die Menge

$$\Pi_n(\mathbb{R}) := \{p \in P_n(\mathbb{R}) : p(0) = 0 = p(1)\}$$

ein Unterraum von $P_n(\mathbb{R})$ ist und bestimme ggf. seine Dimension. Ergänzen Sie $\Pi_n(\mathbb{R})$ ferner zu einer Basis von $P_n(\mathbb{R})$.

(4) Mit $m, n \in \mathbb{N}$ bestimme man eine Basis und die Dimension der linearen Räume

$$\{A \in \mathbb{K}^{n\times n} : A \text{ ist obere Dreiecks-Matrix}\}, \qquad \mathbb{K}^{m\times n}, \qquad F(\{1,\dots,n\},\mathbb{R}).$$

2.5 Komplemente und direkte Summen

Wieder sei X ein linearer Raum über $\mathbb{K}$.

Definition 2.5.1 (direkte Summe)**.** Es seien $Y_1, Y_2 \subseteq X$ Unterräume von X. Dann heißt Y_2 *Komplement* von Y_1 in X, falls gilt

$$Y_1 + Y_2 = X, \qquad\qquad Y_1 \cap Y_2 = \{0\};$$

man schreibt $X = Y_1 \oplus Y_2$ und nennt X die *direkte Summe* von Y_1 und Y_2.

Ein Unterraum Y_1 besitzt i.A. mehrere Komplemente Y_2 in X. Konkrete Beispiele im Anschauungsraum $\mathbb{R}^3$ finden sich in Abb. 2.1

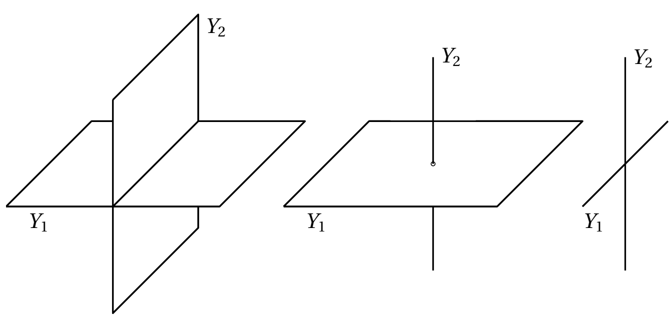

Abb. 2.1 Unterräume Y_1 und Y_2 des $\mathbb{R}^3$: Direkte Summe (mitte), keine direkte Summe wegen $Y_1 \cap Y_2 \neq \{0\}$ (rechts) und $Y_1 + Y_2 \neq \mathbb{R}^3$ (links)

Beispiel 2.5.2. Im Raum $X = \mathbb{R}^3$ ist die Gerade $Y_2 := \left\{\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in X : x_1 = x_2 = x_3\right\}$ ein Komplement zur Ebene $Y_1 := \{x \in X : x_1 - x_2 + x_3 = 0\}$. In der Tat, liegt $x \in Y_1 \cap Y_2$, so erfüllen die Elemente des Schnittes

$$\begin{cases} x_1 - x_2 + x_3 = 0 \\ x_1 - x_2 = 0 \\ x_2 - x_3 = 0 \end{cases}$$

und folglich $x = 0$; dies bedeutet $Y_1 \cap Y_2 = \{0\}$. Andererseits liegen die Vektoren $y_1 := \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$, $y_2 := \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$ in Y_1 und $y_3 := \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ in Y_2. Da $\{y_1, y_2, y_3\}$ eine Basis des Anschauungsraumes $X = \mathbb{R}^3$ bilden, gilt auch $Y_1 + Y_2 = X$.

Beispiel 2.5.3. Wir betrachten den Unterraum $Y_1 := \{p \in P(\mathbb{K}) : p(0) = 0\}$ der Polynome $X = P(\mathbb{K})$. Dann gilt $P(\mathbb{K}) = Y_1 \oplus P_0(\mathbb{K})$, d.h. der lineare Raum aller konstanten Polynome ist ein Komplement zu Y_1.

Satz 2.5.4. *Es seien Y_1, Y_2 Unterräume von X. Es ist $Y_1 \oplus Y_2 = X$ genau dann, wenn es zu jedem $x \in X$ eindeutige $y_1 \in Y_1$ und $y_2 \in Y_2$ mit $x = y_1 + y_2$ gibt.*

Beweis. ($\Rightarrow$) Es sei Y_2 ein Komplement zu Y_1 in X. Wegen $Y_1 + Y_2 = X$ lässt sich dann jedes $x \in X$ darstellen als $x = y_1 + y_2$ mit $y_1 \in Y_1$ und $y_2 \in Y_2$. Um die Eindeutigkeit von y_1, y_2 zu verifizieren, seien $\hat{y}_1 \in Y_1$, $\hat{y}_2 \in Y_2$ zwei weitere Vektoren mit $x = \hat{y}_1 + \hat{y}_2$. Dies impliziert $y_1 - \hat{y}_1 = \hat{y}_2 - y_2$ mit $y_i - \hat{y}_i \in Y_i$ für $i = 1,2$ und folglich $y_i - \hat{y}_i \in Y_1 \cap Y_2 = \{0\}$ für $i = 1,2$. Dies liefert sofort $y_i = \hat{y}_i$, $i = 1,2$.

($\Leftarrow$) Umgekehrt seien $Y_1, Y_2 \subseteq X$ Unterräume derart, dass sich jedes $x \in X$ eindeutig als Summe $x = y_1 + y_2$ mit $y_i \in Y_i$, $i = 1,2$, darstellen lässt. Dann gilt sicherlich $Y_1 + Y_2 = X$. Ist nun $x \in Y_1 \cap Y_2$, so gilt $x = x + 0 = 0 + x$ und da die Darstellung eindeutig sein muss, resultiert $x = 0$. Also ist auch $Y_1 \cap Y_2 = \{0\}$. □

Satz 2.5.5. *Jeder Unterraum eines endlich dimensionalen linearen Raumes besitzt ein Komplement.*

Beweis. Es sei Unterraum $Y_1 \subseteq X$ ein endlich dimensionaler Unterraum von X. Nach Kor. 2.4.16 besitzt Y_1 eine Basis $\{x_1, \dots, x_p\}$, die wir laut Prop. 2.4.9 zu einer Basis $\mathscr{X} = \{x_1, \dots, x_p, x_{p+1}, \dots, x_n\}$ von X ergänzen können. Mit den linearen Raum $Y_2 := \operatorname{span}\{x_{p+1}, \dots, x_n\}$ gilt dann sicherlich $X = Y_1 + Y_2$. Es bleibt der triviale Schnitt $Y_1 \cap Y_2 = \{0\}$ zu zeigen. Für $x \in Y_1 \cap Y_2$ gelten die Darstellungen

$$x = \sum_{i=1}^{p} \xi_i y_i = \sum_{i=p+1}^{n} \eta_i y_i$$

und folglich $\sum_{i=1}^{p} \xi_i y_i - \sum_{i=p+1}^{n} \eta_i y_i = 0$. Da die Vektoren in $\mathscr{X}$ linear unabhängig sind, muss somit $\xi_i = 0$, $1 \le i \le p$ und $\eta_i = 0$, $p < i \le n$ gelten; also $x = 0$. □

2.6 Anwendung: Matrizen und lineare Gleichungen

Es sei $\mathbb{K}$ ein Körper. In diesem Abschnitt illustrieren wir die bisherigen Konzepte anhand von Matrizen $A \in \mathbb{K}^{m\times n}$ mit den n Spalten $a_1, \dots, a_n \in \mathbb{K}^m$ und den m Zeilen $a^1, \dots, a^m \in \mathbb{K}^{1\times n}$. In diesem Kontext bezeichnet man den Unterraum $\operatorname{span}\{a_k\}_{1\le k\le n} \subseteq \mathbb{K}^m$ als *Spaltenraum* und $\operatorname{span}\left\{a^k\right\}_{1\le k\le m} \subseteq \mathbb{K}^{1\times n}$ als *Zeilenraum* von A.

Definition 2.6.1 (Rang einer Matrix)**.** Der *Rang* $\operatorname{rk} A$ einer Matrix $A \in \mathbb{K}^{m\times n}$ ist die Dimension ihres Zeilenraumes.

Bemerkung 2.6.2. Offensichtlich gilt $0 \le \operatorname{rk} A \le m$.

Proposition 2.6.3. *Der Lösungsraum $L_0 \subseteq \mathbb{K}^n$ einer homogenen linearen Gleichung (L_0) mit n Unbekannten erfüllt*

$$\dim L_0 = n - \operatorname{rk} A.$$

Beweis. Wir können o.B.d.A. annehmen, dass die Koeffizientenmatrix $A \in \mathbb{K}^{m\times n}$ in strenger Zeilen-Stufen-Form ist (s. Aufgabe 2.3.12(2)). Es sei r die Anzahl der Zeilen, welche mindestens einen Element $\neq 0$ haben — dies ist der Rang von A. Für $1 \le i \le r$ sei $j_i \in \{1, \dots, n+1\}$ derjenige Spaltenindex, in welcher das erste Element $\neq 0$ der $i$ten Zeile steht. Weiter seien $k_1, \dots, k_{n-r}$ diejenigen Elemente von $\{1, \dots, n\}$, welche nicht in $\left\{j_1, \dots, j_r\right\}$ sind. Die Lösungsmenge von (L_0) lautet

$$L_0 = \left\{ x = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} \in \mathbb{K}^n : \xi_{k_1}, \dots, \xi_{k_{n-r}} \in \mathbb{K} \text{ und } \xi_{j_i} = -a_{i,j_i}^{-1} \sum_{j=1}^{n-r} a_{i,k_j} \xi_{k_j} \text{ für } 1 \le i \le r \right\}$$

und $x_1, \dots, x_{n-r} \in \mathbb{K}^n$ bezeichne die Vektoren in L_0 mit $\xi_{k_j} = 1$ und $\xi_{k_i} = 0$ für Indizes $i \neq j$. Man überlegt sich nun, dass $\{x_1, \dots, x_{n-r}\}$ eine Basis von L_0 ist und es folgt insbesondere $\dim L_0 = n - r$. □

Kapitel 3
Lineare Abbildungen

Dieses Kapitel beschäftigt sich mit Abbildungen zwischen Vektorräumen X und Y, welche die lineare Struktur von X erhalten. Dabei seien X, Y lineare Räume über demselben Körper $\mathbb{K}$.

3.1 Grundlagen

Definition 3.1.1 (lineare Abbildung)**.** Eine Abbildung $T : X \to Y$ heißt *linear*, falls gilt

$$T(\alpha_1 x_2 + \alpha_2 x_2) = \alpha_1 T x_1 + \alpha_2 T x_2 \quad \text{für alle } \alpha_1, \alpha_2 \in \mathbb{K},\ x_1, x_2 \in X. \tag{3.1a}$$

Für die Menge aller solchen linearen Abbildungen schreiben wir $L(X, Y)$.

In obiger Definition haben wir schon die für lineare Abbildungen übliche Notation $Tx := T(x)$ verwendet.

Bemerkung 3.1.2. (1) Jede lineare Abbildung T erfüllt $T0 = 0 \in Y$.

(2) Die Menge $L(X, Y)$ ist ein Unterraum aller Abbildungen $F(X, Y)$ zwischen X und Y; wir kürzen ferner $L(X) := L(X, X)$ ab. Ist Z ein weiterer linearer Raum und $T \in L(X, Y)$, $S \in L(Y, Z)$, so ist auch die Komposition $S \circ T : X \to Z$ linear. $(L(X), \circ)$ bildet eine Halbgruppe mit dem neutralen Element id_X.

Beispiel 3.1.3. Die *Nullabbildung* $0 : X \to Y$, $0x := 0 \in Y$ ist linear, wie auch die *identische Abbildung* $\mathrm{id}_X : X \to X$ aus Bsp. 1.2.3.

Beispiel 3.1.4 (affine Abbildung). Eine Abbildung $S : X \to Y$ heißt *affin*, falls es ein $T \in L(X, Y)$ und ein $y \in Y$ derart gibt, dass $Sx = Tx + y$ für alle $x \in X$. Folglich ist eine affine Abbildung genau dann linear, wenn $y = 0$ gilt.

Beispiel 3.1.5 (die Abbildung T_A). Die wichtigsten linearen Abbildungen dieser Vorlesung sind von der Form $T_A : \mathbb{K}^n \to \mathbb{K}^m$, $T_A x := Ax$ mit einer Matrix $A \in \mathbb{K}^{m\times n}$ über einem Körper $\mathbb{K}$. Auch die Abbildung $\mathbb{K}^{m\times n} \to L(\mathbb{K}^n, \mathbb{K}^m)$, $A \mapsto T_A$ ist linear.

Beispiel 3.1.6. (1) Es sei $\Omega \neq \emptyset$ eine Menge und $x \in \Omega$. Dann ist die *Auswertung* $\mathrm{ev}_x : F(\Omega, X) \to X$, $\mathrm{ev}_x\, u := u(x)$ linear.

(2) Es sei I ein Intervall. Dann ist die *Differenziation* $D : C^1(I, \mathbb{R}^n) \to C(I, \mathbb{R}^n)$, $Du := u'$ eine lineare Abbildung.

(3) Mit einem Intervall I, dem fixen Punkt $t_0 \in I$ und reellen Zahlen $a < b$ definieren auch nachfolgende Integrale lineare Abbildungen:

$$T_1 : C([a,b], \mathbb{R}^n) \to \mathbb{R}^n, \qquad (T_1 u)(t) := \int_a^b u(s)\, ds,$$

$$T_2 : C(I, \mathbb{R}^n) \to C^1(I, \mathbb{R}^n), \qquad (T_2 u)(t) := \int_{t_0}^t u(s)\, ds.$$

Beispiel 3.1.7 (Vorwärts-Shift). Es sei X ein linearer Raum und $\mathbb{I} \in \{\mathbb{N}_0, \mathbb{Z}\}$ gegeben. Bezeichnet dann $\ell(\mathbb{I})$ den linearen Raum aller Folgen $F(\mathbb{I}, X)$ in X, so ist der durch $(S\phi)_k := \phi_{k+1}$ definierte *Vorwärts-Shift* eine Abbildung $S \in L(\ell(\mathbb{I}))$.

Definition 3.1.8 (Kern, Bild, Rang)**.** Ist $T \in L(X, Y)$ eine lineare Abbildung, so bezeichnet man $N(T) := \{x \in X : Tx = 0\}$ als den *Kern*, $R(T) := TX$ als das *Bild* und $\mathrm{rk}\, T := \dim R(T)$ als den *Rang* von T.

Man beachte, dass der Begriff „Rang“ sowohl für Matrizen (siehe Def. 2.6.1), als hiermit auch für lineare Abbildungen eingeführt wurde. Ein Zusammenhang zwischen beiden Begriffen wird in Satz 3.3.8 hergestellt. Ferner wurde ein Dimensionsbegriff nur für lineare Räume erklärt. Dass $R(T)$ tatsächlich diese Struktur hat, und der Rang einer linearen Abbildung T wohldefiniert ist, ergibt sich aus

Proposition 3.1.9. *Für jede lineare Abbildung $T \in L(X, Y)$ ist der Kern $N(T)$ ein Unterraum von X und das Bild $R(T)$ ein Unterraum von Y.*

Beweis. Übungsaufgabe. □

Satz 3.1.10. *Ist $T \in L(X, Y)$ eine lineare Abbildung, so gilt:*

(a) T ist genau dann injektiv, wenn $N(T) = \{0\}$
(b) T ist genau dann surjektiv, wenn $R(T) = Y$.

Beweis. (a) Die Abbildung T ist genau dann nicht injektiv, wenn es $y \in Y$ und Vektoren $x_1, x_2 \in X$ derart gibt, dass $x_1 \neq x_2$ und $y = Tx_1 = Tx_2$. Dies ist wiederum äquivalent zu $T(x_1 - x_2) = 0$ mit $x_1 - x_2 \neq 0$ bzw. $N(T) \neq \{0\}$.

(b) Die Behauptung folgt direkt aus der Definition der Surjektivität. □

Beispiel 3.1.11. (1) Die Auswertung $\mathrm{ev}_x : F(\Omega, X) \to X$ aus obigem Bsp. 3.1.6(1) hat den Kern $N(\mathrm{ev}_x) := \{u \in F(\Omega, X) : u(x) = 0\}$ und das Bild $R(\mathrm{ev}_x) = X$; das Urbild zu einem beliebigen $y \in X$ ist etwa die konstante Funktion $u(x) \equiv y$ auf Ω.

(2) Für die Nullabbildung $0 \in T(X, Y)$ gilt $N(0) = X$ und $R(0) = \{0\}$; sie ist weder surjektiv,[1] noch injektiv.[2] Die identische Abbildung ist aufgrund von $N(\mathrm{id}_X) = \{0\}$ und $R(\mathrm{id}_X) = X$ dagegen bijektiv.

(3) Mit einer Matrix $A \in \mathbb{K}^{m \times n}$ ist die Abbildung $T_A \in L(\mathbb{K}^n, \mathbb{K}^m)$ aus Bsp. 3.1.5 ist genau dann

- injektiv, wenn die linear homogene Gleichung (L_0) nur die triviale Lösung hat,
- surjektiv, wenn für jedes $b \in \mathbb{K}^m$ mindestens eine Lösung der linear inhomogenen Gleichung (L_b) existiert.

(4) Bei der Differenziation $D : C^1([a,b], \mathbb{R}^n) \to C([a,b], \mathbb{R}^n)$ aus Bsp. 3.1.6(2) besteht der Kern $N(D)$ aus allen konstanten Funktionen. Für das Bild $R(D)$ erhalten wir dagegen ganz $C([a,b], \mathbb{R}^n)$, denn ein beliebiges $v \in C([a,b], \mathbb{R}^n)$ besitzt nach dem Hauptsatz der Differenzial- und Integralrechnung das Urbild

$$u(t) := v(a) + \int_a^t v(s)\,ds \quad \text{für alle } t \in [a,b].$$

Für lineare Abbildungen auf endlich-dimensionalen Räumen gilt

Satz 3.1.12 (Dimensionssatz)**.** *Für jedes $T \in L(X, Y)$ mit $\dim X < \infty$ gilt*

$$\dim N(T) + \dim R(T) = \dim X.$$

Beweis. Es sei $\{x_1, \ldots, x_n\}$ eine Basis von $N(T)$ und $\{y_1, \ldots, y_m\}$ eine Basis des Bildes $R(T)$. Wir wählen Vektoren $\hat{x}_1, \ldots, \hat{x}_m \in X$ derart, dass $T\hat{x}_i = y_i$ für alle $1 \le i \le m$ gilt und weisen nach, dass $\mathscr{X} := \{x_1, \ldots, x_n, \hat{x}_1, \ldots, \hat{x}_m\}$ eine Basis von X ist:

(I) $\mathscr{X}$ ist linear unabhängig: Dazu sei

$$\sum_{i=1}^n \alpha_i x_i + \sum_{i=1}^m \beta_i \hat{x}_i = 0 \tag{3.1b}$$

mit Koeffizienten $\alpha_i, \beta_i \in \mathbb{K}$ und durch Anwendung von T resultiert

---

[1] außer Y ist der Nullraum

[2] es sei denn X ist der Nullraum

$$0 = T0 \overset{(3.1b)}{=} \sum_{i=1}^{n} \alpha_i T x_i + \sum_{i=1}^{m} \beta_i T \hat{x}_i = \sum_{i=1}^{m} \beta_i y_i.$$

Weil $\{y_1, \dots, y_m\}$ linear unabhängig ist, folgt $\beta_i = 0$, $1 \le i \le m$ und aus (3.1b) resultiert somit $\sum_{i=1}^{n} \alpha_i x_i = 0$. Die lineare Unabhängigkeit von $\{x_1, \dots, x_n\}$ garantiert daher $\alpha_i = 0$, $1 \le i \le n$.

(II) $\mathscr{X}$ ist ein Erzeugendensystem: Für beliebiges $x \in X$ betrachten wir die Darstellung $Tx = \sum_{i=1}^{m} \beta_i y_i$ mit Koeffizienten $\beta_i \in \mathbb{K}$. Für $\tilde{x} := x - \sum_{i=1}^{m} \beta_i \hat{x}_i$ folgt dann

$$T\tilde{x} = Tx - \sum_{i=1}^{m} \beta_i T \hat{x}_i = Tx - \sum_{i=1}^{m} \beta_i y_i = 0$$

und somit $\tilde{x} \in N(T)$. Also existieren $\alpha_i \in \mathbb{K}$ mit $\tilde{x} = \sum_{i=1}^{n} \alpha_i x_i$ und insgesamt ist

$$x = \tilde{x} + \sum_{i=1}^{m} \beta_i \hat{x}_i = \sum_{i=1}^{n} \alpha_i x_i + \sum_{i=1}^{m} \beta_i \hat{x}_i \in \operatorname{span} \mathscr{X}.$$

Damit ist $\mathscr{X}$ eine Basis von X, also $\dim X = n + m = \dim N(T) + \dim R(T)$. □

Korollar 3.1.13 (Berechnung des Bildes)**.** *Es sei $T \in L(X,Y)$. Ist $\{x_1, \dots, x_n\}$ eine Basis von $N(T)$ und $\{x_1, \dots, x_n, x_{n+1}, \dots, x_d\}$ eine Basis von X mit $n < d$, so besitzt das Bild $R(T)$ die Basis $\{Tx_{n+1}, \dots, Tx_d\}$.*

Beweis. Mit $d = \dim X$ und $n = \dim N(T)$ resultiert nach Satz 3.1.12 die Beziehung $\dim R(T) = d - n$. Es gilt daher $d - n$ linear unabhängige Vektoren in $R(T)$ zu finden. Wir wählen dazu $y_j := Tx_j \in R(T)$ für $n < j \le d$ und müssen nur noch zeigen, dass $\mathscr{Y} := \{y_{n+1}, \dots, y_d\}$ linear unabhängig ist. Wir gehen indirekt vor und nehmen an, dass $\mathscr{Y}$ linear abhängig ist. Dann existiert ein x_{j^*}, $n < j^* \le d$, mit

$$Tx_{j^*} = y_{j^*} = \sum_{\substack{j=n+1,\\ j \neq j^*}}^{d} \eta_j y_j = \sum_{\substack{j=n+1,\\ j \neq j^*}}^{d} T(\eta_j x_j)$$

und Koeffizienten $\eta_{n+1}, \dots, \eta_d \in \mathbb{K}$, was wiederum äquivalent ist zu

$$0 = \sum_{j=n+1}^{d} T(\eta_j x_j) \overset{(3.1a)}{=} T\left(\sum_{j=n+1}^{d} \eta_j x_j\right)$$

mit $\eta_{j^*} = -1$. Nach Definition des Kerns $N(T)$ gilt also $\sum_{j=n+1}^{d} \eta_j x_j \in N(T)$ und weil $\{x_1, \dots, x_n\}$ eine Basis von $N(T)$ ist, gibt es Koeffizienten $\eta_1, \dots, \eta_n \in \mathbb{K}$ mit

$$\sum_{j=1}^{n} \eta_j x_j = \sum_{j=n+1}^{d} \eta_j x_j \quad \Leftrightarrow \quad \sum_{j=1}^{n} \eta_j x_j - \sum_{j=n+1}^{d} \eta_j x_j = 0.$$

Wegen $\eta_{j^*} \neq 0$ muss dann $\{x_1, \dots, x_d\}$ linear abhängig sein, im Widerspruch zur vorausgesetzten Basis-Eigenschaft. □

Zum Abschluss beweisen wir noch ein im weiteren Text zentrales Resultat. Es besagt, dass sich lineare Abbildungen (zwischen endlich dimensionalen Räumen) eindeutig durch die Bilder der Basisvektoren festlegen lassen:

Satz 3.1.14 (Prinzip der linearen Fortsetzung)**.** *Mit einer Basis* $\{x_1, \dots, x_n\}$ *von* X *und beliebigen* $\hat{y}_1, \dots, \hat{y}_n \in Y$ *gilt:*

(a) Sind $T, S \in L(X,Y)$ *Abbildungen mit* $Tx_i = Sx_i$ *für* $1 \le i \le n$*, so folgt* $T = S$
(b) Es existiert genau ein $T \in L(X,Y)$ *mit* $Tx_i = \hat{y}_i$ *für alle* $1 \le i \le n$*.*

Bemerkung 3.1.15 (lineare Fortsetzung). Für gegebenes $x = \sum_{k=1}^n \xi_k x_k$, $n \in \mathbb{N}$, $\xi_k \in \mathbb{K}$, gilt

$$Tx = T \sum_{k=1}^n \xi_k x_k \overset{(3.1a)}{=} \sum_{k=1}^n \xi_k T x_k$$

und Kenntnis der Koeffizienten $\xi_1, \dots, \xi_n \in \mathbb{K}$, sowie der Werte $Tx_1, \dots, Tx_n \in Y$ ermöglicht uns, den Wert Tx zu berechnen. Man sagt auch, Tx lässt sich durch *lineare Fortsetzung* bestimmen.

Beweis. (a) Es gelte $Tx_i = Sx_i$ für $1 \le i \le n$ und wir wählen $x \in X$ beliebig mit der Darstellung $x = \sum_{i=1}^n \xi_i x_i$ für Koeffizienten $\xi_1, \dots, \xi_n \in \mathbb{K}$. Dies impliziert

$$Tx = T\sum_{i=1}^n \xi_i x_i = \sum_{i=1}^n \xi_i T x_i = \sum_{i=1}^n \xi_i S x_i = S \sum_{i=1}^n \xi_i x_i = Sx$$

und weil $x \in X$ beliebig war, folgt $T = S$.

(b) Wir definieren T wie folgt: Ein beliebiger Vektor $x \in X$ habe die Darstellung $x = \sum_{i=1}^n \xi_i x_i$ mit $\xi_1, \dots, \xi_n \in \mathbb{K}$ und wir setzen $Tx := \sum_{i=1}^n \xi_i \hat{y}_i$. Die eindeutige Darstellungseigenschaft von $\mathscr{X}$ aus Satz 2.4.6(b) liefert, dass T wohldefiniert ist; Linearität ergibt sich durch einfaches Nachrechnen und unsere Konstruktion liefert schließlich $Tx_i = \hat{y}_i$ für $1 \le i \le n$. □

Aufgaben 3.1.16. (1) Für den Shift-Operator S aus Bsp. 3.1.7 bestimme man Kern, Bild und untersuche ihn auf Injektivität und Surjektivität.
(3) Man beweise Prop. 3.1.9.
(2) Bestimmen Sie Basen für den Kern und das Bild von $T_A \in L(\mathbb{R}^3, \mathbb{R}^4)$ mit

$$A := \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \\ 10 & 11 & 12 \end{pmatrix}.$$

3.2 Isomorphismen

Eine besonders wichtige Klasse linearer Abbildungen sind diejeningen, welche neben der algebraischen Struktur ihres Definitionsbereichs auch seine Dimension bewahren.

Definition 3.2.1 (Isomorphismus)**.** Eine bijektive Abbildung $T \in L(X,Y)$ heißt *Isomorphismus* und wir definieren

$$GL(X,Y) := \left\{T \in L(X,Y) : T \text{ ist bijektiv}\right\}.$$

Lineare Räume X und Y werden als *isomorph* bezeichnet, falls es einen Isomorphismus $T \in L(X,Y)$ gibt; man schreibt $X \cong Y$.

Bemerkung 3.2.2. (1) Für einen weiteren Vektorraum Z über $\mathbb{K}$ und $T \in GL(X,Y)$, $S \in L(Y,Z)$, ist auch die Verknüpfung $S \circ T : X \to Y$ ein Isomorphismus. Wir kürzen oft $GL(X) := GL(X,X)$ ab und erhalten, dass $(GL(X),\circ)$ eine Gruppe mit neutralem Element id_X definiert — man spricht von der *allgemeinen linearen Gruppe* (engl. *general linear*). Allerdings ist $GL(X,Y)$ kein Unterraum von $L(X,Y)$.

(2) Durch $A := \left\{(X,Y) : X \text{ und } Y \text{ sind isomorph}\right\}$ ist eine Äquivalenzrelation auf der Menge aller linearen Räume definiert.

Beispiel 3.2.3 (Transponierte). Die Abbildung $\cdot^T : \mathbb{K}^{m\times n} \to \mathbb{K}^{n\times m}$,

$$A = (a_{ij})_{1\le i\le n, 1\le j\le m} \mapsto A^T := (a_{ji})_{1\le i\le n, 1\le j\le m},$$

welche bei einer gegebenen Matrix A deren Zeilen und Spalten vertauscht, ist ein mit seiner Umkehrabbildung übereinstimmender Isomorphismus: $(A^T)^T = A$ für alle $A \in \mathbb{K}^{m\times n}$. Man nennt $A^T \in \mathbb{K}^{n\times m}$ die *Transponierte* der Matrix A. Damit ist der Raum der n-Spalten isomorph zum Raum der n-Zeilen, d.h. $\mathbb{K}^n \cong \mathbb{K}^{1\times n}$.

Beispiel 3.2.4 (Polynome). Es sei $\mathbb{K} \in \left\{\mathbb{Q},\mathbb{R},\mathbb{C}\right\}$.

(1) Die Polynome $P_n(\mathbb{K})$ von maximalem Grad $n \in \mathbb{N}_0$ sind isomorph zu $\mathbb{K}^{n+1}$ vermöge des Isomorphismus $T : P_n(\mathbb{K}) \to \mathbb{K}^{n+1}$,

$$p \mapsto (\alpha_0,\dots,\alpha_n) \qquad \text{mit} \qquad p(t) = \sum_{k=0}^{n} \alpha_k t^k.$$

(2) Wir bezeichnen mit

$$\ell_{00} := \left\{(\alpha_k)_{k\in\mathbb{N}_0} : \text{es gibt ein } n \in \mathbb{N}_0 \text{ mit } \alpha_k = 0 \quad \text{für alle } k > n\right\}$$

die Menge aller Folgen in $\mathbb{K}$, welche schließlich (d.h. nach endlich vielen Indizes) Null werden. Dann sind ℓ_{00} und der Raum aller Polynome $P(\mathbb{K})$ isomorph. Als Isomorphismus $T : P(\mathbb{K}) \to \ell_{00}$ wähle man dazu

$$p \mapsto (\alpha_0, \dots, \alpha_n, 0, \dots) \qquad \text{mit} \qquad p(t) = \sum_{k=0}^{n} \alpha_k t^k.$$

Lemma 3.2.5. *Für $T \in L(X,Y)$ ist mit $\mathscr{S} \subseteq X$ auch $T\mathscr{S} \subseteq Y$ linear abhängig.*

Beweis. Wir beschränken uns auf endliche Mengen $\mathscr{S}$. Es sei $\mathscr{S} = \{x_1, \dots, x_n\}$ linear abhängig und folglich existiert eine nichttriviale Darstellung $\sum_{i=1}^{n} \xi_i x_i = 0$ von $0 \in X$. Dann ist aber auch

$$\sum_{i=1}^{n} \xi_i T x_i \overset{(3.1a)}{=} T \sum_{i=1}^{n} \xi_i x_i = T0 = 0$$

eine nichttriviale Darstellung von $0 \in Y$ und $T\mathscr{S}$ ist linear abhängig. □

Satz 3.2.6. *Es sei $T \in GL(X,Y)$. Eine Menge $\mathscr{S} \subseteq X$ ist genau dann linear abhängig, wenn $T\mathscr{S} \subseteq Y$ linear abhängig ist.*

Bemerkung 3.2.7. Als logische Kontraposition erhalten wir, dass Isomorphismen linear unabhängige Mengen (oder Basen) auf ebensolche abbilden.

Beweis. (⇒) Die Behauptung folgt sofort aus Lemma 3.2.5.

(⇐) Nun sei $T\mathscr{S} \subseteq Y$ linear abhängig und wegen $T \in GL(X,Y)$ ist auch die Menge $\mathscr{S} = T^{-1}T\mathscr{S}$ nach Lemma 3.2.5 linear abhängig. □

Proposition 3.2.8. *Jeder n-dimensionale lineare Raum ist isomorph zu $\mathbb{K}^n$, $n \in \mathbb{N}_0$.*

Beweis. Es sei X ein linearer Raum mit $n = \dim X$ und der Basis $\mathscr{X} = \{x_1, \dots, x_n\}$. Wir definieren die lineare Abbildung

$$T : \mathbb{K}^n \to X, \qquad \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} \mapsto \sum_{i=1}^{n} \xi_i x_i.$$

Aufgrund der linearen Unabhängigkeit von $\mathscr{X}$ ist $N(T) = \{0\}$, nach Satz 3.1.10(a) ist T dann injektiv. Da $\mathscr{X}$ ein Erzeugendensystem ist, muss T auch surjektiv sein. Wir haben also einen Isomorphismus zwischen $\mathbb{K}^n$ und X konstruiert. □

Satz 3.2.9. *Endlich dimensionale lineare Räume X, Y sind genau dann isomorph, wenn* $\dim X = \dim Y$.

Beweis. ($\Rightarrow$) Es sei $n := \dim X = \dim Y$. Nach Prop. 3.2.8 existieren Isomorphismen $\Phi : X \to \mathbb{K}^n$ und $\Psi : Y \to \mathbb{K}^n$. Vermöge der Komposition $\Psi^{-1} \circ \Phi \in GL(X, Y)$ sind dann auch X und Y isomorph.

($\Leftarrow$) Es sei $T \in GL(X, Y)$ und $\mathscr{X} = \{x_1, \dots, x_n\}$ eine Basis von X. Wir setzen nun $y_i := Tx_i$ für $1 \le i \le n$ und Satz 3.2.6 garantiert, dass $\mathscr{Y} = \left\{y_1, \dots, y_n\right\}$ eine Basis von Y ist. Da $\mathscr{X}$ und $\mathscr{Y}$ beide n Elemente enthalten, folgt $\dim X = \dim Y$. □

Satz 3.2.10. *Für jede Abbildung $T \in L(X, Y)$ zwischen linearen Räumen X, Y mit $\dim X = \dim Y < \infty$ sind folgende Aussagen äquivalent:*

(a) T ist ein Isomorphismus
(b) T ist injektiv
(c) T ist surjektiv

Beweis. Es ist nachzuweisen, dass unter den angegebenen Voraussetzungen Surjektivität und Injektivität von T äquivalent sind. Es sei also $T \in L(X, Y)$ injektiv. Wegen Satz 3.1.10(a) ist dies äquivalent zu $N(T) = \{0\}$ bzw. $\dim N(T) = 0$. Mit dem Dimensionssatz 3.1.12 ist dann

$$\dim R(T) = \dim X - \dim N(T) = \dim Y - \dim N(T) = \dim Y$$

und T ist genau dann injektiv, wenn $\dim R(T) = \dim Y$, also $R(T) = Y$ gilt. Aufgrund von Satz 3.1.10(b) bedeutet dies Surjektivität von T. □

3.3 Lineare Abbildungen und Matrizen

In diesem Abschnitt seien X und Y endlich dimensionale lineare Räume. Ferner bezeichne durchgehend $\mathscr{X} := \{x_1, \dots, x_n\}$ eine Basis von X und $\mathscr{Y} := \left\{y_1, \dots, y_m\right\}$ eine Basis von Y.

3.3.1 Darstellende Matrizen

Wir folgern aus Satz 3.1.14, dass eine lineare Abbildung T durch die Bilder Tx_i der Basisvektoren von X bestimmt ist. Sind etwa

$$Tx_i = \sum_{j=1}^{m} a_{ij} y_j \quad \text{für alle } 1 \le i \le n \tag{3.3a}$$

und $x = \sum_{k=1}^{n} \xi_k x_k \in X$ mit $Tx = \sum_{j=1}^{m} \eta_j y_j$, so resultiert die Darstellung

$$Tx = T\left(\sum_{k=1}^{n} \xi_k x_k\right) \stackrel{(3.1a)}{=} \sum_{k=1}^{n} \xi_k T x_k \stackrel{(3.3a)}{=} \sum_{k=1}^{n} \xi_k \sum_{j=1}^{m} a_{kj} y_j = \sum_{j=1}^{m}\left(\sum_{k=1}^{n} a_{kj}\xi_k\right) y_j. \tag{3.3b}$$

Für die Koordinaten $\eta_1, \dots, \eta_m \in \mathbb{K}$ von Tx bzgl. der Basis $\mathscr{Y}$ erhalten wir

$$\eta_i = \sum_{k=1}^{n} a_{ik}\xi_k \quad \text{für alle } 1 \le i \le m \tag{3.3c}$$

und haben damit den folgenden Satz bewiesen:

Satz 3.3.1 (darstellende Matrix)**.** *Jedes $T \in L(X,Y)$ wird eindeutig durch eine Matrix $T_{\mathscr{X}}^{\mathscr{Y}} \in \mathbb{K}^{m\times n}$ beschrieben, in deren $k$ter Spalte gerade die Koordinaten von Tx_k bzgl. der Basis $\mathscr{Y}$ stehen. Man nennt $T_{\mathscr{X}}^{\mathscr{Y}}$ die T* darstellende Matrix *in den Basen $\mathscr{X}$ und $\mathscr{Y}$; im Fall $X = Y$ schreiben wir $T_{\mathscr{X}} := T_{\mathscr{X}}^{\mathscr{X}}$.*

Bemerkung 3.3.2. Wir versehen $X = \mathbb{K}^n$ und $Y = \mathbb{K}^m$ mit den Standardbasen $\mathscr{E}_n$ bzw. $\mathscr{E}_m$ aus Bsp. 2.4.3. Für jede Abbildung $T \in L(\mathbb{K}^n, \mathbb{K}^m)$ mit der darstellenden Matrix $T_{\mathscr{E}_n}^{\mathscr{E}_m}$ gilt dann

$$T = T_{T_{\mathscr{E}_n}^{\mathscr{E}_m}}.$$

In der Tat, mit $T_{\mathscr{E}_n}^{\mathscr{E}_m} =: A = (a_{ij})$ rechnen wir für $x = \sum_{k=1}^{n} \xi_k x_k$ gemäß (3.3b) die Beziehung $Tx = Ax$ nach. Insbesondere sind alle linearen Abbildungen zwischen $\mathbb{K}^n$ und $\mathbb{K}^m$ von der Form T_A mit $A \in \mathbb{K}^{m\times n}$, und A ist die darstellende Matrix von T_A in den Standardbasen. Insbesondere gilt $\mathrm{id}_{\mathscr{E}_n} = I_n$.

Beispiel 3.3.3 (Polynome). Es sei $n \in \mathbb{N}$ und $X = P_n(\mathbb{R})$ ausgestattet mit der monomialen Basis $\mathscr{M}_n = \{m_0, \dots, m_n\}$ aus Bsp. 2.4.4. Als lineare Abbildung betrachten die Ableitung $D : P_n(\mathbb{R}) \to P_n(\mathbb{R})$ mit den Bildern

$$Dm_0 = 0, \qquad Dm_k = km_{k-1} \quad \text{für alle } k \in \mathbb{N}$$

und erhalten aus Satz 3.3.1 die darstellende Matrix

$$D_{\mathscr{M}_n} = \begin{pmatrix} 0 & 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 2 & 0 & \dots & 0 \\ 0 & 0 & 0 & 3 & \dots & 0 \\ \vdots & & & & \ddots & 0 \\ 0 & 0 & 0 & 0 & \dots & n \\ 0 & 0 & 0 & 0 & \dots & 0 \end{pmatrix} \in \mathbb{R}^{(n+1)\times(n+1)}.$$

Als nächstes zeigen wir, dass es zu jeder Matrix $A \in \mathbb{K}^{m\times n}$ eine lineare Abbildung gibt, deren darstellende Matrix gerade A ist.

Proposition 3.3.4. *Zu jedem $A \in \mathbb{K}^{m\times n}$ gibt es ein $T \in L(X,Y)$ mit $A = T^{\mathscr{Y}}_{\mathscr{X}}$.*

Beweis. Es sei $A \in \mathbb{K}^{m\times n}$ gegeben. Zu jedem beliebigen $x \in X$ finden wir Koeffizienten $\xi_1,\dots,\xi_n \in \mathbb{K}$ mit $x = \sum_{k=1}^n \xi_k x_k$. Damit definieren wir die gesuchte lineare Abbildung $T : X \to Y$ durch $Tx := \sum_{i=1}^m \left(\sum_{j=1}^n a_{ij}\xi_j\right) y_i$. □

Korollar 3.3.5. *Für endlich dimensionale lineare Räume X, Y sind $L(X,Y)$ und $\mathbb{K}^{\dim X\times \dim Y}$ isomorph. Insbesondere gilt* $\dim L(X,Y) = \dim X \dim Y$.

Beweis. Es sei $m := \dim Y$ und $n := \dim X$. Wir betrachten die lineare Abbildung $\Phi : L(X,Y) \to \mathbb{K}^{m\times n}$ definiert durch $\Phi(T) := T^{\mathscr{Y}}_{\mathscr{X}}$ (vgl. Satz 3.3.1), deren Inverse gerade in Prop. 3.3.4 konstruiert wurde. Also ist Φ ein Isomorphismus und Satz 3.2.9 liefert die Aussage über die Dimension. □

Bekanntlich ist die Verknüpfung linearer Abbildungen wieder eine lineare Abbildung (vgl. Bem. 3.1.2(2)). Daher stellt sich die Frage nach den darstellenden Matrizen der Komposition.

Satz 3.3.6. *Es sei Z ein linearer Raum mit Basis $\mathscr{Z}$. Für lineare Abbildungen $T \in L(X,Y)$ und $S \in L(Y,Z)$ gilt dann*

$$(S\circ T)^{\mathscr{Z}}_{\mathscr{X}} = S^{\mathscr{Z}}_{\mathscr{Y}} T^{\mathscr{Y}}_{\mathscr{X}}.$$

Bemerkung 3.3.7. Für $A \in \mathbb{K}^{l\times m}$ und $B \in \mathbb{K}^{m\times n}$ gilt mittels Bem. 3.3.2, dass

$$T_A \circ T_B = T_{AB}. \tag{3.3d}$$

Beweis. Es sei $\mathscr{Z} := \{z_1,\dots,z_l\}$ eine Basis von Z und

$$A := T^{\mathscr{Y}}_{\mathscr{X}} \in \mathbb{K}^{m\times n}, \qquad B := S^{\mathscr{Z}}_{\mathscr{Y}} \in \mathbb{K}^{l\times m}.$$

Für jedes $1 \le j \le n$ gilt dann

$$(S\circ T)^{\mathscr{Z}}_{\mathscr{X}} x_j = S\sum_{k=1}^m b_{kj}y_k = \sum_{k=1}^m b_{kj}Sy_k = \sum_{k=1}^m b_{kj}\sum_{i=1}^l a_{ik}z_i = \sum_{i=1}^l\sum_{k=1}^m a_{ik}b_{kj}z_i.$$

Der Koeffizient von z_i in obiger Summe ist genau das (i,j)-Element des Produkts AB und es ergibt sich die Behauptung. □

3.3.2 Die Abbildung T_A

In diesem Abschnitt sei $A \in \mathbb{K}^{m\times n}$ und wir interessieren uns für die in Bsp. 3.1.5 eingeführte lineare Abbildung $T_A \in L(\mathbb{K}^n, \mathbb{K}^m)$,

$$T_A x := Ax.$$

Satz 3.3.8. *Der Ränge von $A \in \mathbb{K}^{m\times n}$ und $T_A \in L(\mathbb{K}^n, \mathbb{K}^m)$ stimmen überein*

$$\operatorname{rk} T_A = \operatorname{rk} A.$$

Bemerkung 3.3.9. Mit den Spalten $a_1, \dots, a_n \in \mathbb{K}^m$ gilt $R(T_A) = \operatorname{span}\{a_1, \dots, a_n\}$, weshalb insbesondere $R(T_A)$ und der Spaltenraum von A gleiche Dimension haben. Andererseits war $\operatorname{rk} A$ nach Def. 2.6.1 die Dimension des Zeilenraumes der Matrix A. Daher wird Satz 3.3.8 auch häufig formuliert als

„Spaltenrang = Zeilenrang“.

Beweis. Zunächst merken wir an, dass $N(T_A)$ mit dem Lösungsraum $L_0 \subseteq \mathbb{K}^n$ einer homogenen linearen Gleichung (L_0) übereinstimmt. Nach Prop. 2.6.3 gilt also $\dim N(T_A) = n - \operatorname{rk} A$. Andererseits liefert der Dimensionssatz 3.1.12, dass

$$n = \dim N(T_A) + \dim R(T_A) = \dim N(T_A) + \operatorname{rk} T_A$$

und folglich $\operatorname{rk} T_A = \operatorname{rk} A$. □

Nun sei $A \in \mathbb{K}^{n\times n}$ eine quadratische Matrix mit $\operatorname{rk} A = n$. Dies ist mit Satz 3.2.10 äquivalent dazu, dass $T_A \in L(\mathbb{K}^n)$, $T_A x := Ax$ ein Isomorphismus ist. Wir interessieren uns für die simultane Lösbarkeit der Gleichungen

$$Ax = e_i \quad \text{für alle } 1 \le i \le n,$$

welche wir mittels der augmentierten Matrix $(A, e_1, \dots, e_n) = (A, I_n) \in \mathbb{K}^{n\times 2n}$ notieren. Vermöge des Gauß-Algorithmus lässt sich (A, I_n) durch elementare Zeilenoperationen auf die Form (I_n, B) mit einem $B \in \mathbb{K}^{n\times n}$ bringen; nach Konstruktion erfüllt letztere Matrix die Beziehung $AB = I_n$.

Definition 3.3.10 (inverse Matrix)**.** Eine Matrix $A \in \mathbb{K}^{n\times n}$ heißt *invertierbar*, falls es ein $B \in \mathbb{K}^{n\times n}$ mit der Eigenschaft $AB = I_n$ gibt. Man nennt B die *Inverse* von A und schreibt $A^{-1} := B$.

Beispiel 3.3.11. Wir wollen die Matrix $A := \begin{pmatrix} 1&2&3\\2&3&4\\3&4&6\end{pmatrix}$ invertieren. Gemäß obigem Schema schreiben wir

$$\begin{array}{ccc|ccc} 1&2&3&1&0&0\\2&3&4&0&1&0\\3&4&6&0&0&1\end{array} \Leftrightarrow \begin{array}{ccc|ccc} 1&2&3&1&0&0\\0&1&2&2&-1&0\\0&2&3&3&0&-1\end{array} \Leftrightarrow \begin{array}{ccc|ccc} 1&2&3&1&0&0\\0&1&2&2&-1&0\\0&0&1&1&-2&1\end{array} \Leftrightarrow \begin{array}{ccc|ccc} 1&2&0&-2&6&-3\\0&1&0&0&3&-2\\0&0&1&1&-2&1\end{array}$$

$$\Leftrightarrow \begin{array}{ccc|ccc} 1&0&0&-2&0&1\\0&1&0&0&3&-2\\0&0&1&1&-2&1\end{array}$$

und erhalten die inverse Matrix $A^{-1} = \begin{pmatrix} -2&0&1\\0&3&-2\\1&-2&1\end{pmatrix}$. Eine testweise Multiplikation ergibt $AA^{-1} = A^{-1}A = I_3$.

Korollar 3.3.12. *Die inverse Matrix $A^{-1} \in \mathbb{K}^{n\times n}$ ist eindeutig bestimmt mit*

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

Ferner ist A genau dann invertierbar, wenn $T_A \in GL(\mathbb{K}^n)$.

Beweis. Für invertierbares $A \in \mathbb{K}^{n\times n}$ gelten die Beziehung

$$T_A \circ T_{A^{-1}} \overset{(3.3d)}{=} T_{I_n} = \mathrm{id}_{\mathbb{K}^n}, \qquad T_{A^{-1}} \circ T_A \overset{(3.3d)}{=} T_{I_n} = \mathrm{id}_{\mathbb{K}^n}$$

und folglich ist $T_{A^{-1}} = T_A^{-1}$ die Umkehrabbildung von T_A. Bezüglich der Standardbasis $\mathscr{E}_n$ ist A^{-1} ihre darstellende Matrix, was die Eindeutigkeit der Inversen garantiert. Mittels $T_A^{-1} \circ T_A = \mathrm{id}_{\mathbb{K}^n}$ folgt $A^{-1}A = I_n$. □

Definition 3.3.13 (reguläre und singuläre Matrix)**.** Eine Matrix $A \in \mathbb{K}^{n\times n}$ heißt *regulär*, falls $\operatorname{rk} A = n$ gilt; andernfalls nennt man sie *singulär*.

Die Menge aller Matrizen $\{A \in \mathbb{K}^{n\times n} : A \text{ ist regulär}\}$ ist bezüglich des Matrixprodukts eine Gruppe mit dem neutralen Element I_n.

Eine Diagonalmatrix $A = \operatorname{diag}(a_1, \dots, a_n) \in \mathbb{K}^{n\times n}$ ist genau dann regulär, wenn $\prod_{k=1}^n a_k \neq 0$ gilt. Folglich bilden die regulären Diagonalmatrizen eine kommutative Gruppe mit neutralem Element I_n.

Aufgrund von folgendem Resultat nennt man reguläre Matrizen auch *invertierbar*:

Satz 3.3.14 (Charakterisierung regulärer Matrizen)**.** *Die folgenden Aussagen sind äquivalent für jede Matrix $A \in \mathbb{K}^{n\times n}$:*

(a) $T_A \in GL(\mathbb{K}^n)$
(b) *A ist regulär*
(c) *die Zeilen von A sind linear unabhängig*
(d) *die Spalten von A sind linear unabhängig*
(e) *die lineare Gleichung (L_0) besitzt lediglich die triviale Lösung*
(f) *für jedes $b \in \mathbb{K}^n$ ist (L_b) eindeutig lösbar.*

Beweis. Die Äquivalenz der Aussagen (b) und (c) ist gerade Def. 3.3.13. Aufgrund von Satz 3.3.8 und Bem. 3.3.9 sind auch (c) und (d) gleichwertig.

$(d) \Rightarrow (e)$ resultiert aus der Def. 2.3.6 der linearen Unabhängigkeit.

$(e) \Rightarrow (f)$ ergibt sich aus Satz 1.4.9(a).

$(f) \Rightarrow (a)$ Nach unserer Voraussetzung (f) ist offenbar $T_A^{-1}(\{b\})$ für jedes $b \in \mathbb{K}^n$ einpunktig. Also ist $T_A \in L(\mathbb{K}^n)$ bijektiv.

$(a) \Rightarrow (b)$ Wegen der Invertierbarkeit von T_A ist $\dim R(T_A) = n$ und folglich gilt nach Satz 3.3.8, dass $\operatorname{rk} A = \operatorname{rk} T_A = \dim R(T_A) = n$. □

Zum Abschluss dieses Kapitels beschäftigen wir uns mit linear-inhomogenen Gleichungen (L_b). Ihre *augmentierte Koeffizientenmatrix* $(A,b) \in \mathbb{K}^{m\times(n+1)}$ definieren wir durch

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

Satz 3.3.15. *Es sei $b \in \mathbb{K}^m$. Eine lineare inhomogene Gleichung (L_b) hat genau dann eine Lösung, wenn für die augmentierte Koeffizientenmatrix gilt*

$$\operatorname{rk} A = \operatorname{rk}(A,b).$$

Beweis. Mit beliebigem $b \in \mathbb{K}^m$ gilt:

$(\Rightarrow)$ Ist (L_b) lösbar, so ist b eine Linearkombination der Spalten von A, so dass die Spaltenräume der Matrizen A und (A,b) übereinstimmen. Mit Bem. 3.3.9 folgt dann die Behauptung.

$(\Leftarrow)$ Da der Spaltenraum von A im Spaltenraum von (A,b) enthalten ist, muss dann b wegen $\operatorname{rk} A = \operatorname{rk}(A,b)$ im Spaltenraum von A liegen. Also ist b eine Linearkombination der Spalten von A und es existiert eine Lösung von (L_b). □

3.3.3 Basiswechsel

Auf einem endlich dimensionalen linearen Raum X seien nun die beiden Basen $\mathscr{X} = \{x_1, \dots, x_n\}$ und $\mathscr{X}' = \{x'_1, \dots, x'_n\}$ gegeben. Wir interessieren uns für die Frage nach dem Verhalten der Darstellungsmatrix einer linearen Abbildung $T \in L(X)$, wenn man von einer Basis $\mathscr{X}$ zu einer weiteren Basis $\mathscr{X}'$ übergeht:

- Zunächst lassen sich die Elemente der Basis $\mathscr{X}'$ darstellen durch die Elemente von $\mathscr{X}$ als
$$x'_j = \sum_{i=1}^{n} s_{ij} x_i \quad \text{für alle } 1 \le j \le n. \tag{3.3e}$$
Hieraus wird die sog. *Basiswechselmatrix* $S := (s_{ij})_{1 \le i,j \le n} \in \mathbb{K}^{n \times n}$ gebildet, welche den Übergang von $\mathscr{X}$ nach $\mathscr{X}'$ beschreibt. Sie genügt der Regel
$$\text{Spalten von } S = \text{Koordinatenvektoren der „neuen" Basisvektoren.}$$
- Umgekehrt kann man die Basiselemente x_j auch mittels der x'_i ausgedrückt werden: Aus $x_j = \sum_{i=1}^{n} s'_{ij} x'_i$ erhalten wir
$$x_j = \sum_{i=1}^{n} s'_{ij} \left(\sum_{k=1}^{n} s_{ki} x_k \right) = \sum_{k=1}^{n} \left(\sum_{i=1}^{n} s_{ki} s'_{ij} \right) x_k \quad \text{für alle } 1 \le j \le n$$
und der in der letzten Klammer stehende Ausdruck ist gerade das (k, j)-te Element des Produkts $SS' \in \mathbb{K}^{n \times n}$.

Wegen der linearen Unabhängigkeit von $\mathscr{X}$ folgern wir, dass $SS' = I_n$ und somit $S' = S^{-1}$ gelten muss. Schließlich definiert auch jede reguläre Matrix $S \in \mathbb{K}^{n \times n}$ einen Basiswechsel gemäß der Beziehung (3.3e) bzw. $x'_j = Sx_j$ für alle $1 \le j \le n$.

Damit kehren wir zu unserer Ausgangsfrage zurück:

Satz 3.3.16. *Ist $S \in \mathbb{K}^{n \times n}$ die Basiswechselmatrix zwischen $\mathscr{X}$ und $\mathscr{X}'$, so gilt*
$$T_{\mathscr{X}'} = S^{-1} T_{\mathscr{X}} S \quad \text{für alle } T \in L(X).$$

Beweis. Es sei $A := T_{\mathscr{X}} \in \mathbb{K}^{n \times n}$ und wir wollen die Darstellungsmatrix $T_{\mathscr{X}'}$ bestimmen. Dazu gilt mit der oben eingeführten Notation, dass
$$\begin{aligned} Tx'_j &\overset{(3.3e)}{=} T\left(\sum_{i=1}^{n} s_{ij} x_i \right) \overset{(3.1a)}{=} \sum_{i=1}^{n} s_{ij} T x_i = \sum_{i=1}^{n} s_{ij} \left(\sum_{k=1}^{n} a_{ki} x_k \right) \\ &= \sum_{i=1}^{n} \sum_{k=1}^{n} a_{ki} s_{ij} \left(\sum_{l=1}^{n} s'_{lk} x'_l \right) = \sum_{l=1}^{n} \left(\sum_{i=1}^{n} \sum_{k=1}^{n} s'_{lk} a_{ki} s_{ij} \right) x'_l \quad \text{für alle } 1 \le j \le n. \end{aligned}$$

Der in der runden Klammer stehende Ausdruck ist das (l, j)-te Element des Produkts $S^{-1}T_{\mathscr{X}}S$, welcher mit dem entsprechenden Element der Darstellungsmatrix $T_{\mathscr{X}'}$ übereinstimmen muss. $\square$

Der obige Satz 3.3.16 gilt als Motivation für

Definition 3.3.17 (Ähnlichkeit)**.** Zwei Matrizen $A, B \in \mathbb{K}^{n\times n}$ heißen *ähnlich*, falls es eine reguläre Matrix $S \in \mathbb{K}^{n\times n}$ derart gibt, dass $B = S^{-1}AS$.

Bemerkung 3.3.18. (1) Laut Satz 3.3.16 sind die Darstellungsmatrizen bezüglich verschiedener Basen ähnlich vermöge der Basiswechselmatrix.

(2) Die Ähnlichkeit von Matrizen definiert eine Äquivalenzrelation auf dem linearen Raum aller Matrizen $\mathbb{K}^{n\times n}$. Ein Ziel der linearen Algebra besteht darin, möglichst einfache Repräsentanten zu finden.

Beispiel 3.3.19. Auf dem Raum $X = \mathbb{R}^2$ seien die beiden Basen

$$\mathscr{X} = \{e_1, e_2\}, \qquad \mathscr{X}' = \left\{\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}1\\-1\end{pmatrix}\right\}$$

und die lineare Abbildung $T \in L(\mathbb{R}^2)$, $Tx := \left(\begin{smallmatrix}x_2\\x_1\end{smallmatrix}\right)$ mit der entsprechenden Darstellungsmatrix $T_{\mathscr{X}} = \left(\begin{smallmatrix}0&1\\1&0\end{smallmatrix}\right)$ gegeben. Für die Basiswechselmatrix S erhalten wir

$$S = \begin{pmatrix}1&1\\1&-1\end{pmatrix}, \qquad S^{-1} = \tfrac{1}{2}\begin{pmatrix}1&1\\1&-1\end{pmatrix},$$

was laut Satz 3.3.16 die Darstellungsmatrix $T_{\mathscr{X}'} = S^{-1}T_{\mathscr{X}}S = \left(\begin{smallmatrix}1&0\\0&-1\end{smallmatrix}\right)$ zur Folge hat.

Lösungen zu Aufgaben

Aufgaben 1.1.11

Aufgabe 1.1.11(1)

- **Vor.:** $X_1 := \mathbb{R}$ und $R_1 := \{(x, y) \in \mathbb{R} \times \mathbb{R} : x < y\}$
  Beh.: R_1 ist weder reflexiv noch symmetrisch, dafür aber transitiv!
  Bew.: Es seien $x, y, z \in \mathbb{R}$.
  Offensichtlich gilt nicht $x < x$, d.h. $(x, x) \notin R_1$. Also ist R_1 nicht reflexiv. Für Paare $(x, y) \in R_1$ und $(y, z) \in R_1$ ist $x < y$ und $y < z$, also $x < z$ und damit $(x, z) \in R_1$. Demnach ist R_1 transitiv. Das Paar $(0, 1)$ ist wegen $0 < 1$ zwar in R_1 enthalten, aber es gilt nicht $1 < 0$ und somit $(1, 0) \notin R_1$; folglich kann die Relation R_1 nicht symmetrisch sein.
- **Vor.:** $X_1 := \mathbb{R}$ und $R_2 := \{(x, y) \in \mathbb{R} \times \mathbb{R} : x \geq y\}$
  Beh.: R_2 ist reflexiv und transitiv, aber nicht symmetrisch!
  Bew.: Es seien $x, y, z \in \mathbb{R}$.
  Da sicherlich $x \geq x$ gilt, erhalten wir $(x, x) \in R_2$ und R_2 ist reflexiv. Die Bedingungen $(x, y) \in R_2$ und $(y, z) \in R_2$ sind gleichbedeutend mit $x \geq y$ und $y \geq z$, was nunmehr $x \geq z$ impliziert. Demzufolge gilt $(x, z) \in R_2$ und R_2 ist auch transitiv. Ferner ist R_2 nicht symmetrisch, denn obwohl das Paar $(1, 0)$ in der Relation R_2 enthalten ist, gilt keineswegs $0 \geq 1$ bzw. $(0, 1) \in R_2$.
- **Vor.:** $X_3 := \{\text{Menschen}\}$ und $R_3 := \{(x, y) \in X \times X : x \text{ ist verwandt mit } y\}$
  Beh.: R_3 ist reflexiv, transitiv und symmetrisch, also eine Äquivalenzrelation!
  Bew.: Es seien x, y, z Menschen.
  Weil jeder Mensch x mit sich selbst verwandt ist, gilt $(x, x) \in R_3$ und R_3 ist reflexiv. Ist x mit y verwandt und y mit z, so sind auch x und z verwandt, d.h. $(x, z) \in R_3$ und R_3 ist transitiv. Ferner ist die Verwandtschafts-Relation offensichtlich symmetrisch.
  Die zugehörigen Äquivalenzklasse $[x]$ zu einem Menschen x ist die Menge aller mit ihm verwandten Menschen, d.h. seine Familie.

- **Vor.:** $R_4 = \{(x, y) \in X_4 \times Y_4 : x \text{ ist Tochter von } y\}$ zwischen den beiden Mengen $X_4 := \{\text{Frauen}\}$ und $Y_4 := \{\text{Männer}\}$
  Bem.: Da R_4 keine Relation auf einer Menge ist, macht die Fragestellung keinen Sinn!

Aufgabe 1.1.11(2)

Vor.: Es sei $X = \{(m, n) \in \mathbb{Z} \times \mathbb{Z} : n \neq 0\}$ und

$$A := \big\{\big((m_1, n_1), (m_2, n_2)\big) \in X \times X : m_1 n_2 = m_2 n_1\big\}.$$

Beh.: A ist eine Äquivalenzrelation!
Bew.: Es seien $(m_1, n_1), (m_2, n_2), (m_3, n_3) \in X$.
Wegen der trivialen Aussage $m_1 n_1 = m_1 n_1$ ist $\big((m_1, n_1), (m_1, n_1)\big) \in A$, d.h. die Relation A ist reflexiv. Für $\big((m_1, n_1), (m_2, n_2)\big) \in A$ und $\big((m_2, n_2), (m_3, n_3)\big) \in A$ ist

$$m_1 n_2 = m_2 n_1, \qquad m_2 n_3 = m_3 n_2,$$

was wegen $n_i \neq 0$, $1 \leq i \leq 3$, äquivalent zu

$$\frac{m_1}{n_1} = \frac{m_2}{n_2}, \qquad \frac{m_2}{n_2} = \frac{m_3}{n_3}$$

ist. Somit gilt auch $\frac{m_1}{n_1} = \frac{m_3}{n_3}$ und gleichwertig $m_1 n_3 = m_3 n_1$, was schießlich die gewünschte Inklusion $\big((m_1, n_1), (m_3, n_3)\big) \in A$ zur Folge hat; wir haben die Transitivität von A gezeigt. Nun ist A auch noch symmetrisch, denn $\big((m_1, n_1), (m_2, n_2)\big) \in A$ bedeutet

$$m_1 n_2 = m_2 n_1$$

und es resultiert $\big((m_2, n_2), (m_1, n_1)\big) \in A$.
Aus den Beziehungen

$$[(m_1, n_1)] = \{(m_2, n_2) \in X : m_1 n_2 = m_2 n_1\} = \left\{(m_2, n_2) \in X : \frac{m_1}{n_1} = \frac{m_2}{n_2}\right\}$$

erkennt man, dass die Äquivalenzklassen von A gerade die rationalen Zahlen $\mathbb{Q}$ sind.

Aufgaben 1.2.10

Aufgabe 1.2.10(1)

- **Vor.:** $D_1 = B_1 = \mathbb{R}$ und $f_1(x) := x^3$
  Beh.: f_1 ist injektiv, surjektiv und es gilt $f_1(D_1) = \mathbb{R}$!
  Bew.: Für jedes $y \in \mathbb{R}$ erhält man aus der Eindeutigkeit und Existenz der Wurzelfunktion das Urbild $f_1^{-1}(\{y\}) = \{\sqrt[3]{y}\}$. Damit ist f_1 injektiv und surjektiv. Ferner besitzt f_1 das Bild $f_1(\mathbb{R}) = \mathbb{R}$. Ihre Umkehrabbildung lautet $f_1^{-1} : B_1 \to D_1$ mit der Abbildungsvorschrift $f_1^{-1}(x) = \sqrt[3]{x}$.
- **Vor.:** $D_2 = B_2 = \mathbb{R}$ und $f_2(x) := \sin(\pi x)$
  Beh.: f_2 ist weder injektiv noch surjektiv und es gilt $f_2(D_2) = [-1, 1]$.
  Bew.: Um zu gegebenem $y \in B_2 = \mathbb{R}$ das Urbild $f_2^{-1}(\{y\})$ zu bestimmen, ist die transzendente Gleichung $\sin(\pi x) = y$ nach y zu lösen. Hierzu unterscheiden wir mehrere Fälle: Im Fall $|y| > 1$ existieren offensichtlich keine Lösungen x, denn die Werte der Sinus-Funktion liegen im Intervall $[-1, 1]$, d.h. $f_2^{-1}(\{y\}) = \emptyset$. Für $y = 1$ und jedes $k \in \mathbb{Z}$ ist $x = 2k + \frac{1}{2}$ eine Lösung und somit $f_2^{-1}(\{1\}) = \{2k + \frac{1}{2} : k \in \mathbb{Z}\}$. Analog folgt für $y = -1$ das Urbild $f_2^{-1}(\{-1\}) = \{2k + \frac{3}{2} : k \in \mathbb{Z}\}$. Schließlich erhalten wir im verbleibenden Fall $y \in (-1, 1)$ die Urbilder

  $$f_2^{-1}(\{y\}) = \left\{2k + \frac{\arcsin y}{\pi} : k \in \mathbb{Z}\right\} \cup \left\{2k + 1 - \frac{\arcsin y}{\pi} : k \in \mathbb{Z}\right\}$$

  Demzufolge ist f_2 nicht surjektiv und auch nicht injektiv. Allerdings existiert zu jedem $y \in [-1, 1]$ mindestens ein Urbild und daher $f_2(\mathbb{R}) = [-1, 1]$.
- **Vor.:** $D_3 = \mathbb{R}$, $B_3 = [-1, 1]$ und $f_3(x) := \sin(\pi x)$
  Beh.: f_3 ist nicht injektiv, aber surjektiv und es gilt $f_3(D_3) = [-1, 1]$.
  Bew.: Die Behauptung folgt aus den obigen Überlegungen.
- **Vor.:** $D_4 = [-\frac{1}{2}, \frac{1}{2}]$, $B_2 = [-1, 1]$ und $f_4(x) := \sin(\pi x)$
  Beh.: f_4 ist injektiv, surjektiv und es gilt $f_4(D_4) = [-1, 1]$.
  Bew.: Wie oben gilt $f_4^{-1}(\{y\}) = \left\{\frac{\arcsin y}{\pi}\right\}$ für alle $y \in [-1, 1] = B_4$ und damit ist f_4 surjektiv und injektiv, also bijektiv. Ferner gilt ebenso $f_4(D_4) = [-1, 1]$ und die Umkehrabbildung ergibt sich zu $f_4^{-1} : [-1, 1] \to [-\frac{1}{2}, \frac{1}{2}]$ mit $f_4^{-1}(x) = \frac{\arcsin x}{\pi}$.

Aufgabe 1.2.10(2)

Vor.: Es seien D, B nichtleer, $A_1, A_2 \subseteq D$, $B_1, B_2 \subseteq B$ und eine Abbildung $f : D \to B$ gegeben.

- **Beh.:** $f(A_1 \cup A_2) = f(A_1) \cup f(A_2)$
  Bew.: Die folgenden Aussagen sind äquivalent:

  $y \in f(A_1 \cup A_2)$
  $\Leftrightarrow$ es existiert ein $x \in A_1 \cup A_2$ mit $f(x) = y$
  $\Leftrightarrow$ es gibt ein $x \in A_1$ mit $f(x) = y$, oder es gibt ein $x \in A_2$ mit $f(x) = y$

$$\begin{aligned}&\Leftrightarrow\ y \in f(A_1) \text{ oder } y \in f(A_2)\\&\Leftrightarrow\ y \in f(A_1) \cup f(A_2)\end{aligned}$$

und folglich ergibt sich die Behauptung.

- **Beh.:** $f^{-1}(B_1 \cup B_2) = f^{-1}(B_1) \cup f^{-1}(B_2)$
  Bew.: Wir erhalten die folgenden Äquivalenzen:
  $$\begin{aligned}&\ x \in f^{-1}(B_1 \cup B_2)\\&\Leftrightarrow \text{ es existiert ein } y \in B_1 \cup B_2 \text{ mit } f(x) = y\\&\Leftrightarrow \text{ es gibt ein } y \in B_1 \text{ mit } f(x) = y, \text{ oder es gibt ein } y \in B_2 \text{ mit } f(x) = y\\&\Leftrightarrow\ x \in f^{-1}(B_1) \text{ oder } x \in f^{-1}(B_2)\\&\Leftrightarrow\ x \in f^{-1}(B_1) \cup f^{-1}(B_2),\end{aligned}$$
  was unsere Behauptung zeigt.
- **Beh.:** $f(A_1 \cap A_2) \subseteq f(A_1) \cap f(A_2)$
  Bew.: Im Fall $A_1 \cap A_2 = \emptyset$ gilt die Behauptung trivialerweise. Im Fall von nicht disjunkten Mengen A_1 und A_2 mit $y \in f(A_1 \cap A_2)$ existiert ein $x \in A_1 \cap A_2$ derart, dass $f(x) = y$. Wegen $x \in A_1$ und $x \in A_2$ gilt $y \in f(A_1)$ und $y \in f(A_2)$, also die Inklusion $y \in f(A_1) \cap f(A_2)$.
  Bsp.: Es sei $D := \{1,2\}$, $B := \{0\}$ und $f : D \to B$ definiert durch $f(1) := f(2) := 0$. Mit $A_1 := \{1\}$ und $A_2 := \{2\}$ gilt dann $f(A_1 \cap A_2) = f(\emptyset) = \emptyset$, andererseits aber die Beziehung $f(A_1) \cap f(A_2) = \{0\} \cap \{0\} = \{0\} \neq \emptyset$.
- **Beh.:** $f^{-1}(A_1 \cap A_2) = f^{-1}(A_1) \cap f^{-1}(A_2)$
  Bew.: Wiederum erhält man die Äquivalenzen
  $$\begin{aligned}&\ x \in f^{-1}(B_1 \cap B_2)\\&\Leftrightarrow\ f(x) \in B_1 \cap B_2\\&\Leftrightarrow\ f(x) \in B_1 \text{ und } f(x) \in B_2\\&\Leftrightarrow\ x \in f^{-1}(B_1) \text{ und } x \in f^{-1}(B_2)\\&\Leftrightarrow\ x \in f^{-1}(B_1) \cap f^{-1}(B_2),\end{aligned}$$
  was der Behauptung entspricht.

Literaturverzeichnis

[Beu10] A. Beutelspacher, *Lineare Algebra, Eine Einführung in die Wissenschaft der Vektoren, Abbildungen und Matrizen*, Vieweg & Teubner, 2010

[Bos08] S. Bosch, *Lineare Algebra,* 4. überarbeitete Auflage, Springer, 2008

[Fis10] G. Fischer, *Lineare Algebra, Eine Einf§hrung f§r Studienanfänger*, Vieweg & Teubner, 2010

Sachverzeichnis