

Kapitel 2

Lineare Abbildungen

2.1 Beispiele und Definitionen

2.1.1. a) Besonders wichtige Funktionen $f: \mathbb{R} \rightarrow \mathbb{R}$ sind die Polynome mit

$$f(x) = a_0 + a_1x + \dots + a_nx^n,$$

denn die ganze Abbildungsvorschrift ist durch Angabe der Koeffizienten $a_0, \dots, a_n$, d.h. von $n+1$ Zahlen, festgelegt. Ist $a_n \neq 0$, so nennt man n den Grad von f . Für Grad 0 hat man die konstanten und für Grad 1 die linearen Funktionen. In der Analysis untersucht man, ob sich eine beliebige Funktion bei kleinen Veränderungen des Arguments x nahezu wie eine lineare Funktion verhält; das führt zum Begriff der Differenzierbarkeit. Daher sind die linearen Funktionen unentbehrliches Hilfsmittel der Analysis. Wir betrachten also ein

$$f: \mathbb{R} \rightarrow \mathbb{R}, \quad x \mapsto ax + b = f(x),$$

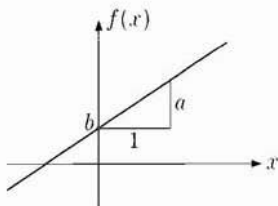


Bild 2.1

mit $a, b \in \mathbb{R}$. Ihr Graph ist eine Gerade mit Steigung a , die durch $(0, b)$ geht. Setzt man den konstanten Anteil $b = 0$, so bleibt eine Funktion $f(x) = ax$, die durch eine einzige reelle Zahl a festgelegt ist, und offensichtlich die Eigenschaften

$$f(x + x') = f(x) + f(x') \quad \text{und} \quad f(\lambda x) = \lambda f(x) \quad (*)$$

für beliebige $x, x', \lambda \in \mathbb{R}$ hat.

b) Viel interessanter und anschaulich einfach zu verstehen ist die Situation in der Ebene. Zunächst betrachten wir eine *Drehung* um den Nullpunkt mit dem Winkel ϑ . Das ist eine Abbildung

$$F: \mathbb{R}^2 \rightarrow \mathbb{R}^2,$$

die sich wie folgt beschreiben läßt. Es ist

$$F(0) = 0, \quad F(e_1) = (\cos \vartheta, \sin \vartheta), \quad F(e_2) = (-\sin \vartheta, \cos \vartheta),$$

und für ein beliebiges $x = (x_1, x_2) = x_1 e_1 + x_2 e_2$ ist

$$F(x) = x_1 F(e_1) + x_2 F(e_2) = (x_1 \cos \vartheta - x_2 \sin \vartheta, x_1 \sin \vartheta + x_2 \cos \vartheta).$$

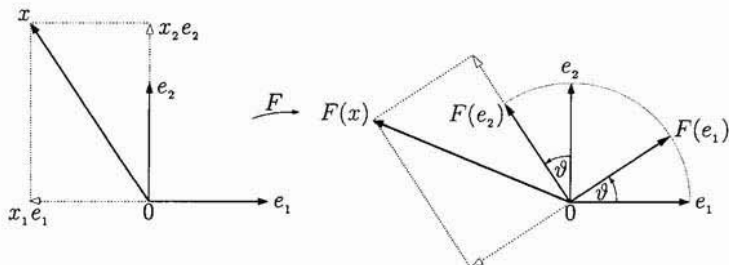


Bild 2.2

Benutzen wir die Matrix

$$A = \begin{pmatrix} \cos \vartheta & -\sin \vartheta \\ \sin \vartheta & \cos \vartheta \end{pmatrix},$$

und schreiben wir x und $F(x)$ als Spaltenvektoren, so ist mit der in 0.4.1 erklärten Multiplikation $F(x) = A \cdot x$.

An dieser Überlegung ist zu sehen, daß man für $F(e_1)$ und $F(e_2)$ beliebige Vektoren $a = (a_1, a_2)$, $b = (b_1, b_2)$ vorschreiben kann. Dann läßt sich F mit der Konstruktion aus Bild 2.3 zu einer Abbildung von $\mathbb{R}^2$ auf sich ausdehnen.

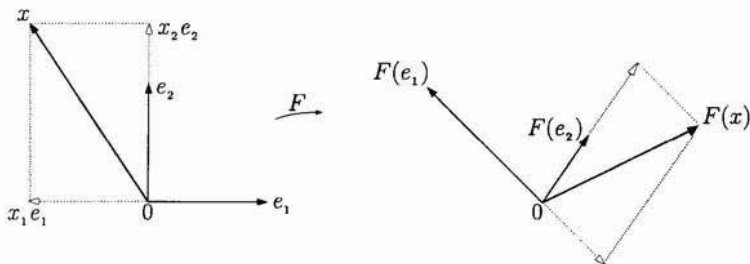


Bild 2.3

Mit Hilfe der Matrix

$$A = \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix}$$

ist F wieder beschrieben durch $F(x) = A \cdot x$.

c) Betrachtet man nun anstelle von $\mathbb{R}$ einen beliebigen Körper K , und nimmt man m und n anstelle von 2, so ist man sofort in einer recht allgemeinen Situation. Ist nämlich $A = (a_{ij})$ eine $m \times n$ -Matrix, so erhält man daraus eine Abbildung

$$F: K^n \rightarrow K^m, \quad \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto \begin{pmatrix} a_{11}x_1 + \dots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \dots + a_{mn}x_n \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}.$$

Wie man leicht nachrechnet, hat sie wieder die Eigenschaft

$$F(x + x') = F(x) + F(x'), \quad F(\lambda x) = \lambda F(x) \quad (*)$$

für $x, x' \in K^n$ und $\lambda \in K$.

Mit Hilfe der in 0.4.1 erklärten Multiplikation einer Matrix A mit einer Spalte kann man diese Abbildung einfacher beschreiben als

$$F: K^n \rightarrow K^m, \quad x \mapsto A \cdot x,$$

wenn man die Vektoren von K^n und K^m als Spalten entsprechender Höhe schreibt. Setzt man für x die Basisvektoren $e_1, \dots, e_n$ als Spalten ein und rechnet man die Produkte aus, so stellt man fest, daß

$$A \cdot e_1, \dots, A \cdot e_n$$

die Spalten von A in dieser Reihenfolge sind. Diese Beobachtung:

Die Spaltenvektoren der Matrix sind die Bilder der Basisvektoren

sollte man sich einprägen.

d) Die Transposition von Matrizen als Abbildung

$$F: M(m \times n; K) \rightarrow M(n \times m; K)$$

hat nach den Rechenregeln in 1.5.8 die zu (*) analogen Eigenschaften.

e) Der Vektorraum $V = \mathcal{C}(I; \mathbb{R})$ der auf einem Intervall $I = [a, b]$ stetigen Funktionen ist unendlichdimensional, und die Abbildung

$$S: \mathcal{C}(I; \mathbb{R}) \rightarrow \mathbb{R}, \quad f \mapsto \int_a^b f(x) dx,$$

hat nach den Rechenregeln für Integrale die zu (*) analogen Eigenschaften

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

f) Ist $V = \mathcal{D}(I; \mathbb{R})$ der Vektorraum der beliebig oft differenzierbaren Funktionen, so hat die durch Differentiation erklärte Abbildung

$$D: V \rightarrow V, \quad f \mapsto f',$$

die zu (*) analoge Eigenschaft.

2.1.2. Die obigen Beispiele motivieren die folgende

Definition. Eine Abbildung $F: V \rightarrow W$ zwischen K -Vektorräumen V und W heißt *linear* (genauer *K -linear* oder *Homomorphismus von K -Vektorräumen*), wenn

$$\mathbf{L1} \quad F(v + w) = F(v) + F(w),$$

$$\mathbf{L2} \quad F(\lambda v) = \lambda F(v)$$

für alle $v, w \in V$ und alle $\lambda \in K$. Diese beiden Bedingungen kann man zusammenfassen zu einer:

$$\mathbf{L} \quad F(\lambda v + \mu w) = \lambda F(v) + \mu F(w)$$

für alle $v, w \in V$ und $\lambda, \mu \in K$. Man überlegt sich ganz leicht, daß L1 und L2 zusammen mit L gleichwertig sind.

Es ist üblich, den Begriff Homomorphismus zu verschärfen. Man nennt eine lineare Abbildung $F: V \rightarrow W$ einen

Isomorphismus, wenn F bijektiv ist,

Endomorphismus, wenn $V = W$,

Automorphismus, wenn $V = W$ und F bijektiv ist.

Wir notieren einige einfache Folgerungen aus den Axiomen:

Bemerkung. Ist $F: V \rightarrow W$ linear, so gilt:

$$a) \quad F(0) = 0 \text{ und } F(v - w) = F(v) - F(w).$$

$$b) \quad F(\lambda_1 v_1 + \dots + \lambda_n v_n) = \lambda_1 F(v_1) + \dots + \lambda_n F(v_n).$$

c) Ist die Familie $(v_i)_{i \in I}$ in V linear abhängig, so ist $(F(v_i))_{i \in I}$ in W linear abhängig.

d) Sind $V' \subset V$ und $W' \subset W$ Untervektorräume, so sind auch $F(V') \subset W$ und $F^{-1}(W') \subset V$ Untervektorräume.

$$e) \quad \dim F(V) \leq \dim V.$$

f) Ist F ein Isomorphismus, so ist auch $F^{-1}: W \rightarrow V$ linear.

Beweis. a) $F(0) = F(0 \cdot 0) = 0 \cdot F(0) = 0$ und

$$F(v - w) = F(v + (-1)w) = F(v) + (-1)F(w) = F(v) - F(w).$$

b) folgt durch wiederholte Anwendung der Regel L.

c) Ist für $i_1, \dots, i_n \in I$ und $\lambda_1, \dots, \lambda_n \in K$

$$\lambda_1 v_{i_1} + \dots + \lambda_n v_{i_n} = 0,$$

so folgt nach b)

$$\lambda_1 F(v_{i_1}) + \dots + \lambda_n F(v_{i_n}) = 0.$$

d) Wegen $0 \in V'$ ist $0 = F(0) \in F(V')$. Sind $w, w' \in F(V')$, so gibt es $v, v' \in V'$ mit $F(v) = w$ und $F(v') = w'$. Also ist

$$w + w' = F(v) + F(v') = F(v + v') \in F(V'),$$

denn $v + v' \in V'$. Für $\lambda \in K$ gilt wegen $\lambda v \in V'$

$$\lambda w = \lambda F(v) = F(\lambda v) \in F(V').$$

Sind $v, v' \in F^{-1}(W')$, so bedeutet das $F(v), F(v') \in W'$. Also ist

$$F(v + v') = F(v) + F(v') \in W',$$

also $v + v' \in F^{-1}(W')$ und analog sieht man $\lambda v \in F^{-1}(W')$.

e) Hat man $w_1 = F(v_1), \dots, w_n = F(v_n) \in F(V)$ linear unabhängig, so sind nach c) auch $v_1, \dots, v_n$ in V linear unabhängig.

f) Seien $w, w' \in W$ und $\lambda, \mu \in K$. Ist $w = F(v)$ und $w' = F(v')$, so ist $v = F^{-1}(w)$, $v' = F^{-1}(w')$, und es folgt aus

$$F(\lambda v + \mu v') = \lambda w + \mu w'$$

durch Anwendung von F^{-1} auf beiden Seiten

$$\lambda F^{-1}(w) + \mu F^{-1}(w') = F^{-1}(\lambda w + \mu w').$$

□

2.1.3. Lineare Abbildungen kann man in verschiedener Weise miteinander verknüpfen. Zunächst betrachten wir die Hintereinanderschaltung.

Bemerkung 1. Sind U, V, W Vektorräume und $G: U \rightarrow V$, $F: V \rightarrow W$ lineare Abbildungen, so ist auch

$$F \circ G: U \rightarrow W \quad \text{linear.}$$

Man beachte dabei wie immer die Reihenfolge der Abbildungen, was man sich in einem Diagramm aufzeichnen kann:

$$\begin{array}{ccccc} & & G & \rightarrow & V & & F & \rightarrow & \\ U & \xrightarrow{\quad} & & & & & & & W \\ & & F \circ G & \rightarrow & & & & & \end{array}$$

Beweis. Für $u, u' \in U$ ist

$$\begin{aligned} (F \circ G)(u + u') &= F(G(u + u')) = F(G(u) + G(u')) \\ &= F(G(u)) + F(G(u')) \\ &= (F \circ G)(u) + (F \circ G)(u'). \end{aligned}$$

Ganz analog zeigt man $(F \circ G)(\lambda u) = \lambda(F \circ G)(u)$. □

In Aufgabe 3 zu 1.4 hatten wir gesehen, wie man für eine Menge X und einen Vektorraum W die Menge $\text{Abb}(X, W)$ zu einem Vektorraum machen kann. Ist auch $X = V$ ein K -Vektorraum, so definiert man

$$\text{Hom}_K(V, W) := \{F: V \rightarrow W: F \text{ ist } K\text{-linear}\}.$$

Falls klar ist, welcher Körper K gemeint ist, schreibt man einfacher $\text{Hom}(V, W)$.

Bemerkung 2. Für Vektorräume V und W über demselben Körper K ist

$$\text{Hom}_K(V, W) \subset \text{Abb}(V, W)$$

ein Untervektorraum.

Beweis. Für $F, G \in \text{Hom}_K(V, W)$ und $\lambda \in K$ ist zu zeigen, daß $F + G$ und λF wieder K -linear sind. Das ist aber klar, denn für alle $\sigma, \tau \in K$ und $v, w \in V$ ist

$$\begin{aligned} (F + G)(\sigma v + \tau w) &= F(\sigma v + \tau w) + G(\sigma v + \tau w) \\ &= \sigma F(v) + \tau F(w) + \sigma G(v) + \tau G(w) \\ &= \sigma(F(v) + G(v)) + \tau(F(w) + G(w)) \\ &= \sigma(F + G)(v) + \tau(F + G)(w) \end{aligned}$$

und

$$\begin{aligned} (\lambda \cdot F)(\sigma v + \tau w) &= \lambda F(\sigma v + \tau w) = \lambda(\sigma F(v) + \tau F(w)) \\ &= \sigma \lambda F(v) + \tau \lambda F(w) = \sigma(\lambda \cdot F)(v) + \tau(\lambda \cdot F)(w). \end{aligned}$$

Wir vermerken noch, daß der Nullvektor in $\text{Hom}_K(V, W)$ die Nullabbildung

$$0: V \rightarrow W \quad \text{mit} \quad 0(v) := 0 \quad \text{für alle } v \in V$$

ist, und die zu $F: V \rightarrow W$ negative Abbildung gegeben ist durch

$$-F: V \rightarrow W \quad \text{mit} \quad (-F)(v) := -F(v) \quad \text{für alle } v \in V. \quad \square$$

Die Dimension von $\text{Hom}(V, W)$ werden wir in 2.4.2 berechnen (siehe auch Aufgabe 6 zu 2.4).

2.1.4. Im Spezialfall $V = W$ setzt man

$$\text{End}(V) := \text{Hom}(V, V),$$

das sind die *Endomorphismen* von V . Diese Menge wird, wie wir in 2.1.3 gesehen haben, zu einem Vektorraum mit Addition und Multiplikation mit Skalaren. Man kann überdies eine *Multiplikation* durch die Hintereinanderschaltung erklären, in Zeichen

$$F \cdot G := F \circ G \quad \text{für} \quad F, G \in \text{End}(V).$$

Satz. Ist V ein K -Vektorraum, so ist $\text{End}(V)$ zusammen mit der oben erklärten Addition und Multiplikation ein Ring.

Der einfache *Beweis* sei dem Leser überlassen. In 2.6.4 werden wir sehen, daß dieser *Endomorphismenring* für endlichdimensionales V zu einem Matrizenring isomorph ist. Damit erhält man eine Methode, viele der interessanten Unterringe von $\text{End}(V)$ durch die Gestalt der entsprechenden Matrizen zu beschreiben.

Aufgaben zu 2.1

1. Sei X eine Menge und V der $\mathbb{R}$ -Vektorraum aller Funktionen $f: X \rightarrow \mathbb{R}$. Beweisen Sie: Ist $\varphi: X \rightarrow X$ eine beliebige Abbildung, so ist die Abbildung

$$F_\varphi: V \rightarrow V, \quad f \mapsto f \circ \varphi$$

$\mathbb{R}$ -linear.

2. Untersuchen Sie die folgenden Abbildungen auf Linearität:

- a) $\mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad (x, y) \mapsto (3x + 2y, x),$ b) $\mathbb{R} \rightarrow \mathbb{R}, \quad x \mapsto ax + b,$
 c) $\mathbb{Q}^2 \rightarrow \mathbb{R}, \quad (x, y) \mapsto x + \sqrt{2}y$ (über $\mathbb{Q}$), d) $\mathbb{C} \rightarrow \mathbb{C}, \quad z \mapsto \bar{z},$
 e) $\text{Abb}(\mathbb{R}, \mathbb{R}) \rightarrow \mathbb{R}, \quad f \mapsto f(1),$ f) $\mathbb{C} \rightarrow \mathbb{C}, \quad z \mapsto \bar{z}$ (über $\mathbb{R}$).

3. Für einen Endomorphismus $F: V \rightarrow V$ ist die Menge der *Fixpunkte* von F definiert durch $\text{Fix } F := \{v \in V: F(v) = v\}.$

- a) Zeigen Sie, dass $\text{Fix } F \subset V$ ein Untervektorraum ist.
 b) Sei der Endomorphismus F gegeben durch

$$\text{i) } F: \mathbb{R}^3 \rightarrow \mathbb{R}^3, x \mapsto \begin{pmatrix} 1 & 2 & 2 \\ 0 & 1 & 0 \\ 3 & 0 & 1 \end{pmatrix} \cdot x,$$

$$\text{ii) } F: \mathbb{R}[t] \rightarrow \mathbb{R}[t], P \mapsto P',$$

$$\text{iii) } F: \mathcal{D}(\mathbb{R}, \mathbb{R}) \rightarrow \mathcal{D}(\mathbb{R}, \mathbb{R}), f \mapsto f'.$$

Bestimmen Sie jeweils eine Basis von $\text{Fix } F$.

4. Zeigen Sie, dass die Menge $\text{Aut}(V)$ der Automorphismen eines Vektorraums V mit der Komposition von Abbildungen als Verknüpfung eine Gruppe ist.

5. Sei $F: V \rightarrow V$ ein Endomorphismus des Vektorraums V und $v \in V$, so dass für eine natürliche Zahl n gilt:

$$F^n(v) \neq 0 \quad \text{und} \quad F^{n+1}(v) = 0.$$

Beweisen Sie, dass dann $v, F(v), \dots, F^n(v)$ linear unabhängig sind.

6. Ist $F: V \rightarrow W$ ein Isomorphismus und $V = U_1 \oplus U_2$, so ist $W = F(U_1) \oplus F(U_2)$.

2.2 Bild, Fasern und Kern, Quotientenvektorräume*

Nachdem wir eine ganze Reihe formaler Eigenschaften von linearen Abbildungen behandelt haben, soll nun versucht werden, die „Geometrie“ solcher Abbildungen etwas besser zu verstehen.

2.2.1. Ist $F: V \rightarrow W$ eine lineare Abbildung, so nennen wir

$$\begin{aligned} \operatorname{Im} F &:= F(V) && \text{das Bild von } F \text{ (vgl. 1.1.3),} \\ F^{-1}(w) &:= \{v \in V: F(v) = w\} && \text{die Faser über } w \in W \text{ und} \\ \operatorname{Ker} F &= F^{-1}(0) && \text{den Kern von } F. \end{aligned}$$

Bemerkung. Ist $F: V \rightarrow W$ linear, so gilt:

- a) $\operatorname{Im} F \subset W$ und $\operatorname{Ker} F \subset V$ sind Untervektorräume.
- b) F surjektiv $\Leftrightarrow \operatorname{Im} F = W$.
- c) F injektiv $\Leftrightarrow \operatorname{Ker} F = \{0\}$.
- d) Ist F injektiv, und sind $v_1, \dots, v_n \in V$ linear unabhängig, so sind auch die Bilder $F(v_1), \dots, F(v_n)$ linear unabhängig.

Beweis. a) und b) und die „Hinrichtung“ von c) sind ganz klar. Gibt es umgekehrt zwei verschiedene $v, v' \in V$ mit $F(v) = F(v')$, so folgt $F(v - v') = 0$, also $0 \neq v - v' \in \operatorname{Ker} F$.

Zu d) nehmen wir an, daß

$$\mu_1 F(v_1) + \dots + \mu_n F(v_n) = 0.$$

Die linke Seite hat wegen der Linearität das Urbild $\mu_1 v_1 + \dots + \mu_n v_n$, wegen der Injektivität folgt

$$\mu_1 v_1 + \dots + \mu_n v_n = 0,$$

also $\mu_1 = \dots = \mu_n = 0$. □

Eine besonders wichtige Zahl für eine lineare Abbildung ist die Dimension ihres Bildes. Man nennt sie den *Rang*, in Zeichen

$$\operatorname{rang} F := \dim \operatorname{Im} F.$$

Insbesondere beschreibt eine Matrix $A \in M(m \times n; K)$ eine lineare Abbildung

$$A: K^n \rightarrow K^m, \quad x \mapsto y = Ax,$$

wobei die Vektoren $x \in K^n$ und $y \in K^m$ als Spalten geschrieben sind, und als *Rang von A*, in Zeichen

$$\operatorname{rang} A,$$

bezeichnet man den Rang dieser linearen Abbildung. Ist $(e_1, \dots, e_n)$ die kanonische Basis des K^n , so sind

$$Ae_1, \dots, Ae_n$$

die Spalten von A , also ist

$$\operatorname{Im} A = A(K^n) = \operatorname{span}(Ae_1, \dots, Ae_n),$$

das ist der Spaltenraum von A . Also ist der gerade erklärte Rang von A gleich dem in 1.5.8 eingeführten Spaltenrang. Will man ihn berechnen, so genügt es, A auf Spaltenstufenform, d.h. A auf Zeilenstufenform zu bringen (vgl. 1.5.7).

2.2.2. Die Begriffe *Bild* und *Faser* hat man analog für eine beliebige Abbildung $F: X \rightarrow Y$ zwischen Mengen, und X wird durch die Fasern in disjunkte Teilmengen zerlegt:

$$X = \bigcup_{y \in \operatorname{Im} F} F^{-1}(y).$$

Wir wollen untersuchen, wie diese *Faserung* im Fall einer linearen Abbildung aussieht. Dazu zunächst ein einfaches, aber typisches

Beispiel. Wir betrachten die Abbildung

$$F: \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \begin{pmatrix} -2x_1 + 2x_2 \\ -x_1 + x_2 \end{pmatrix}.$$

Es ist $\operatorname{Im} F = \mathbb{R} \cdot (2, 1)$, $\operatorname{Ker} F = \mathbb{R} \cdot (1, 1)$, und für $(2b, b) \in \operatorname{Im} F$ ist die Faser die Gerade mit der Gleichung $x_2 = x_1 + b$, also

$$\begin{aligned} F^{-1}(2b, b) &= (0, b) + \mathbb{R} \cdot (1, 1) \\ &= \{(\lambda, b + \lambda) : \lambda \in \mathbb{R}\}. \end{aligned}$$

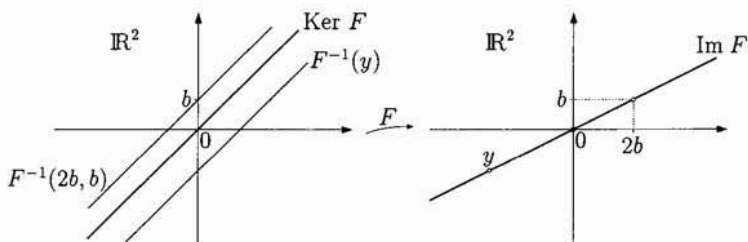


Bild 2.4

Die Fasern sind also parallele Geraden, der Kern ist die einzige Faser durch den Nullpunkt. Allgemein gilt die

Bemerkung. Ist $F: V \rightarrow W$ linear, $w \in \text{Im } F$ und $u \in F^{-1}(w)$ beliebig, so ist

$$F^{-1}(w) = u + \text{Ker } F = \{u + v : v \in \text{Ker } F\}.$$

Beweis. Ist $v' \in F^{-1}(w)$, so folgt

$$\begin{aligned} F(v') &= F(u) \Rightarrow F(v' - u) = 0 \Rightarrow v := v' - u \in \text{Ker } F \\ &\Rightarrow v' = u + v \in u + \text{Ker } F \end{aligned}$$

Ist umgekehrt $v' = u + v \in u + \text{Ker } F$, so ist $F(v') = F(u) = w$, also $v' \in F^{-1}(w)$. $\square$

2.2.3. Teilmengen, die durch „Parallelverschiebung“ eines Untervektorraumes entstehen, erhalten einen eigenen Namen.

Definition. Eine Teilmenge X eines K -Vektorraumes V heißt ein *affiner Unterraum*, falls es ein $v \in V$ und einen Untervektorraum $W \subset V$ gibt, so daß

$$X = v + W := \{u \in V : \text{es gibt ein } w \in W \text{ mit } u = v + w\}$$

(Bild 2.5). Es ist vorteilhaft, auch die leere Menge einen affinen Unterraum zu nennen.

Beispiele für affine Unterräume des $\mathbb{R}^n$ sind Punkte, Geraden und Ebenen (vgl. Kap. 0).

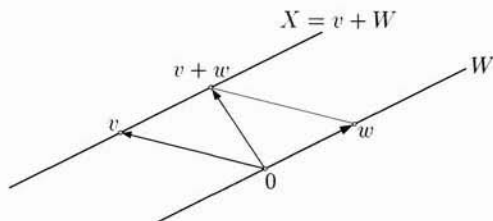


Bild 2.5

Bemerkung. Sei $X = v + W \subset V$ ein affiner Unterraum. Dann gilt:

- Für ein beliebiges $v' \in X$ ist $X = v' + W$.
- Ist $v' \in V$ und $W' \subset V$ ein Untervektorraum mit $v + W = v' + W'$, so folgt $W = W'$ und $v' - v \in W$.

Kurz ausgedrückt: Zu einem affinen Unterraum $v + W$ ist der Untervektorraum W eindeutig bestimmt, und der Aufhängepunkt v kann beliebig in X gewählt werden.

Beweis. a) Wir schreiben $v' = v + w'$.

$$\begin{aligned} X \subset v' + W: \quad u \in X &\Rightarrow u = v + w \quad \text{mit } w \in W \\ &\Rightarrow u = v' + (w - w') \\ &\Rightarrow u \in v' + W \end{aligned}$$

$$v' + W \subset X: \quad u = v' + w \in v' + W \Rightarrow u = v + (w + w') \in v + W.$$

b) Definiert man

$$X - X := \{u - u' : u, u' \in X\}$$

als die Menge der Differenzen (man beachte den Unterschied zu der in 1.1.2 definierten Differenzmenge $X \setminus X = \emptyset$), so sieht man ganz leicht, daß

$$X - X = W \quad \text{und} \quad X - X = W'$$

sein muß. Also ist $W = W'$.

Wegen $v + W = v' + W$ gibt es ein $w \in W$ mit $v' = v + w$. Also ist $v' - v = w \in W$. $\square$

Da für einen affinen Unterraum $X = v + W \subset V$ der Untervektorraum W eindeutig bestimmt ist, können wir durch

$$\dim X := \dim W$$

die *Dimension* von X erklären.

2.2.4. Wir zeigen nun, wie man Basen wählen kann, die maßgeschneidert sind für eine lineare Abbildung (vgl. 2.4.3).

Satz. Sei $F: V \rightarrow W$ linear und $\dim V < \infty$. Sind Basen

$$(v_1, \dots, v_k) \quad \text{von } \operatorname{Ker} F, \quad (w_1, \dots, w_r) \quad \text{von } \operatorname{Im} F,$$

sowie beliebige Vektoren $u_1 \in F^{-1}(w_1), \dots, u_r \in F^{-1}(w_r)$ gegeben, so ist

$$\mathcal{A} := (u_1, \dots, u_r, v_1, \dots, v_k)$$

eine Basis von V . Insbesondere gilt die Dimensionsformel

$$\dim V = \dim \operatorname{Im} F + \dim \operatorname{Ker} F.$$

Beweis. Für $v \in V$ sei

$$F(v) = \mu_1 w_1 + \dots + \mu_r w_r \quad \text{und} \quad v' := \mu_1 u_1 + \dots + \mu_r u_r.$$

Wegen $F(v) = F(v')$ folgt $v - v' \in \operatorname{Ker} F$, also

$$v - v' = \lambda_1 v_1 + \dots + \lambda_k v_k$$

und

$$v = \mu_1 u_1 + \dots + \mu_r u_r + \lambda_1 v_1 + \dots + \lambda_k v_k.$$

Also wird V durch $\mathcal{A}$ erzeugt. Ist

$$\mu_1 u_1 + \dots + \mu_r u_r + \lambda_1 v_1 + \dots + \lambda_k v_k = 0, \quad (*)$$

so folgt durch Anwendung von F

$$\mu_1 w_1 + \dots + \mu_r w_r = 0, \quad \text{also} \quad \mu_1 = \dots = \mu_r = 0,$$

da $w_1, \dots, w_r$ linear unabhängig sind. In $(*)$ eingesetzt ergibt sich

$$\lambda_1 v_1 + \dots + \lambda_k v_k = 0, \quad \text{also} \quad \lambda_1, \dots, \lambda_k = 0$$

da $v_1, \dots, v_k$ linear unabhängig sind. $\square$

Als unmittelbare Folgerung aus der Dimensionsformel notieren wir:

Korollar 1. Ist V endlichdimensional und $F: V \rightarrow W$ linear, so gilt für alle nichtleeren Fasern

$$\dim F^{-1}(w) = \dim V - \dim \operatorname{Im} F. \quad \square$$

Korollar 2. Zwischen zwei endlichdimensionalen Vektorräumen V und W gibt es genau dann einen Isomorphismus, wenn $\dim V = \dim W$. $\square$

Korollar 3. Sei $\dim V = \dim W < \infty$ und $F: V \rightarrow W$ linear. Dann sind folgende Bedingungen gleichwertig:

- i) F injektiv
- ii) F surjektiv
- iii) F bijektiv $\square$

2.2.5. Durch weiteres Spielen mit den Basen aus Satz 2.2.4 erhält man folgenden

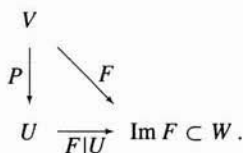
Faktorisierungssatz. Sei $F: V \rightarrow W$ linear und

$$\mathcal{A} = (u_1, \dots, u_r, v_1, \dots, v_k) \quad \text{eine Basis von } V$$

mit $\operatorname{Ker} F = \operatorname{span}(v_1, \dots, v_k)$. Definieren wir $U = \operatorname{span}(u_1, \dots, u_r)$, so gilt

- 1) $V = U \oplus \operatorname{Ker} F$.
- 2) Die Einschränkung $F|_U: U \rightarrow \operatorname{Im} F$ ist ein Isomorphismus.
- 3) Bezeichnet $P: V = U \oplus \operatorname{Ker} F \rightarrow U$, $v = u + v' \mapsto u$, die Projektion auf den ersten Summanden, so ist $F = (F|_U) \circ P$.

In Form eines Diagrammes hat man



Insbesondere hat jede nichtleere Faser $F^{-1}(w)$ mit U genau einen Schnittpunkt, und es ist

$$P(v) = F^{-1}(F(v)) \cap U.$$

Man kann also $F: V \rightarrow W$ zerlegen (oder *faktorisieren*) in drei Anteile: eine *Parallelprojektion*, einen *Isomorphismus* und die *Inklusion* des Bildes. Der zur Konstruktion erforderliche direkte Summand U ist allerdings nicht eindeutig bestimmt, er hängt ab von der Wahl der Basisvektoren $u_1, \dots, u_r$. Wenn in V eine Winkelmessung möglich ist (vgl. Kapitel 5), kann man U eindeutig machen durch die Vorschrift, auf $\operatorname{Ker} F$ senkrecht zu stehen. Die Umkehrung $\operatorname{Im} F \rightarrow U$ von $F|U$ nennt man einen *Schnitt*, da sie aus jeder Faser genau einen Punkt ausschneidet. Als gute Illustration kann Beispiel 2.2.2 mit $k = r = 1$ dienen.

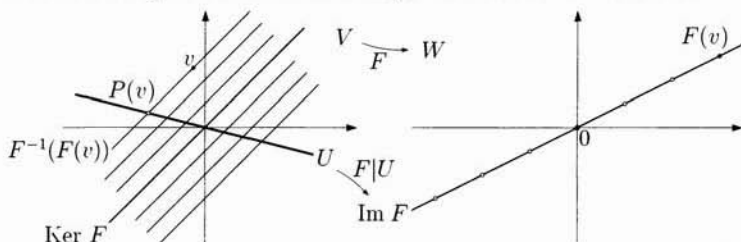


Bild 2.6

Beweis. 1) folgt aus der Charakterisierung direkter Summen in 1.6.3.

Wegen $\operatorname{Ker} F|U = (\operatorname{Ker} F) \cap U = \{0\}$ ist $F|U$ auch injektiv, also Isomorphismus mit Bild $\operatorname{Im} F$. 3) folgt aus der Konstruktion von P . Ist schließlich

$$v \in V \text{ und } v = u + v' \text{ mit } u \in U \text{ und } v' \in \operatorname{Ker} F,$$

so ist $u = P(v)$, also $F(v) = F(u) = F(P(v)) =: w$. Ist überdies $w_1, \dots, w_r$ eine Basis von $\operatorname{Im} F$ mit $F(u_i) = w_i$,

$$w = \mu_1 w_1 + \dots + \mu_r w_r \text{ und } v \in F^{-1}(w) \cap U,$$

so folgt $v = \mu_1 u_1 + \dots + \mu_r u_r$. □

Zur Vorbereitung auf den gleich folgenden Abschnitt über lineare Gleichungssysteme ist es nützlich, die gerade beschriebene allgemeine Situation für eine

durch eine Matrix $A \in M(m \times n; K)$ in Zeilenstufenform gegebene Abbildung

$$A: K^n \rightarrow K^m$$

zu betrachten. Sind (in der Notation von 0.4.3) $j_1, \dots, j_r$ die Indizes der Pivotspalten, und sind $e_{j_1}, \dots, e_{j_r}$ die zu diesen Indizes gehörigen Basisvektoren des K^n , so sind die Bilder

$$A(e_{j_1}), \dots, A(e_{j_r}) \in K^m$$

(das sind gerade die Pivotspalten) eine Basis von

$$\text{Im}(A) = \text{span}(e'_1, \dots, e'_r).$$

Dabei ist mit $(e'_1, \dots, e'_r)$ die kanonische Basis des K^r bezeichnet. Also ist

$$U := \text{span}(e_{j_1}, \dots, e_{j_r})$$

in diesem Fall ein direkter Summand zum Kern von A im Sinn von 1.6.3. Der Leser möge das zur Übung präzise begründen.

Für die erste Lektüre wird empfohlen, den Rest dieses Abschnittes zu überblättern und bei 2.3 wieder einzusteigen.

2.2.6. Ist $F: V \rightarrow W$ eine lineare Abbildung, so sind die Fasern von F nach 2.2.2 die zum Untervektorraum $\text{Ker } F \subset V$ parallelen affinen Räume. Wir wollen nun umgekehrt zu jedem vorgegebenen Untervektorraum $U \subset V$ eine lineare Abbildung mit Kern U konstruieren. Dazu benötigt man einen Vektorraum W als Bild; wir zeigen, daß es dafür einen kanonischen Kandidaten gibt, den „Quotientenvektorraum“ $W = V/U$. Da die Konstruktion ziemlich abstrakt ist, wollen wir zunächst etwas inhaltlichen Hintergrund bereitstellen.

Beispiel 1. Sei $V = \mathbb{R}^2$ und $U \subset V$ eine Gerade durch den Ursprung. Man nennt zwei Punkte $v, v' \in \mathbb{R}^2$ *äquivalent*, wenn die Differenz in U liegt, in Zeichen

$$v \sim v' \Leftrightarrow v' - v \in U.$$

Geometrisch bedeutet das, daß v und v' *gleich weit entfernt* von U sind, wobei die Entfernung von Punkten links von U negativ und rechts von U positiv gerechnet sein soll.

Es ist ganz einfach zu sehen, daß dadurch in V eine Äquivalenzrelation im Sinn von 1.1.8 erklärt wird. Die Äquivalenzklassen sind die zu U parallelen Geraden, das sind die affinen Räume aus 2.2.3.

Der Leser mache sich auch die Analogie zu den Restklassen modulo m aus 1.2.7 klar: Dort wurde die *Gleichheit* abgeschwächt zur *Kongruenz*, hier wird *gleich* ersetzt durch *gleich weit entfernt*.

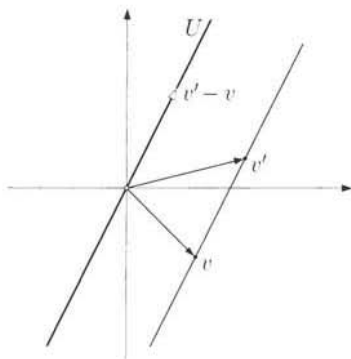


Bild 2.7

Beispiel 2. a) Wir betrachten den unendlich-dimensionalen Vektorraum

$$\mathcal{C}(\mathbb{R}) = \{f: \mathbb{R} \rightarrow \mathbb{R}: f \text{ stetig}\}.$$

Eine beliebige Teilmenge $X \subset \mathbb{R}$ sei vorgegeben, ihr Komplement $A := \mathbb{R} \setminus X$ soll die Rolle einer *Ausnahmemenge* spielen, d.h. die Werte von f auf A werden als unwesentlich angesehen. Damit können wir den Untervektorraum

$$\mathcal{I}(X) := \{f \in \mathcal{C}(\mathbb{R}): f(x) = 0 \text{ für alle } x \in X\} \subset \mathcal{C}(\mathbb{R})$$

der „unwesentlichen“ Funktionen betrachten und für $f, g \in \mathcal{C}(\mathbb{R})$

$$f \sim g : \Leftrightarrow g - f \in \mathcal{I}(X)$$

erklären. In Worten bedeutet das, f und g sind im wesentlichen (d.h. außerhalb A) gleich. Auch diese Äquivalenz ist eine kontrollierte (von A abhängige) Abschwächung der Gleichheit.

b) Eine Variante davon ist die folgende: Man benutzt auf $\mathbb{R}$ ein Integral (etwa das Riemann- oder besser das Lebesgue-Integral), d.h. ein Integral, mit dem möglichst viele Funktionen integrierbar sind. Sei

$$\mathcal{L}(\mathbb{R}) := \{f: \mathbb{R} \rightarrow \mathbb{R}: f \text{ integrierbar}\}$$

und

$$\mathcal{N} := \{f \in \mathcal{L}(\mathbb{R}): \int_{\mathbb{R}} |f(t)| dt = 0\} \subset \mathcal{L}(\mathbb{R}).$$

Nach den Rechenregeln für ein Integral folgt, daß $\mathcal{N} \subset \mathcal{L}(\mathbb{R})$ ein Untervektorraum ist. Man beachte, daß $\mathcal{N}$ unendliche Dimension hat, denn etwa die Funktionen f_i mit $f_i(t) = 0$ für $t \neq i$ und $f_i(i) = 1$ sind für $i \in \mathbb{N}$ in $\mathcal{N}$ linear unabhängig.

Für $f, g \in \mathcal{L}(\mathbb{R})$ bedeutet $f \sim_{\mathcal{N}} g$ dann

$$\int_{\mathbb{R}} |f(t) - g(t)| dt = 0.$$

Dafür sagt man auch, f und g sind „fast überall“ gleich, denn die Menge

$$\{t \in \mathbb{R} : f(t) \neq g(t)\}$$

muß sehr klein sein.

2.2.7. Sei nun ganz allgemein V ein K -Vektorraum und $U \subset V$ ein Untervektorraum. Für $v, v' \in V$ erklären wir die Äquivalenz modulo U

$$v \sim_U v' :\Leftrightarrow v' - v \in U.$$

Aus den Eigenschaften eines Untervektorraumes folgt ganz einfach, daß die Bedingungen für eine Äquivalenzrelation aus 1.1.8 erfüllt sind.

Die Äquivalenzklasse eines $v \in V$ ist gleich dem affinen Unterraum, also

$$\{v' \in V : v' \sim_U v\} = v + U,$$

denn

$$v' \sim_U v \Leftrightarrow v' - v \in U \Leftrightarrow \text{es gibt ein } u \in U \text{ mit } v' = v + u.$$

Die Menge der Äquivalenzklassen wird mit V/U bezeichnet, die kanonische Abbildung sei

$$\varrho: V \rightarrow V/U = \{v + U : v \in V\}, \quad v \mapsto \varrho(v) = v + U.$$

Dabei wird jedem Punkt der ihn enthaltende affine Raum zugeordnet, oder anders ausgedrückt wird jeder Vektor ersetzt durch die Menge all der zu ihm gleichwertigen Vektoren. Im Extremfall $U = 0$ ist die Äquivalenz die Gleichheit und ϱ wird bijektiv. Für $U = V$ ist alles äquivalent, und V/U besteht nur aus einem Element.

Nun kommt der entscheidende Schritt, nämlich die Beobachtung, daß man mit den affinen Räumen rechnen kann wie mit Vektoren.

Satz. Sei V ein K -Vektorraum und $U \subset V$ ein Untervektorraum. Dann kann man die Menge V/U auf genau eine Weise so zu einem K -Vektorraum machen, daß die kanonische Abbildung

$$\varrho: V \rightarrow V/U, \quad v \mapsto v + U,$$

linear wird. Weiter gilt:

1) ϱ ist surjektiv.

2) $\text{Ker } \varrho = U$.

3) $\dim V/U = \dim V - \dim U$, falls $\dim V < \infty$.

4) Der Quotientenvektorraum V/U hat die folgende universelle Eigenschaft: Ist $F: V \rightarrow W$ eine lineare Abbildung mit $U \subset \text{Ker } F$, so gibt es genau eine lineare Abbildung $\bar{F}: V/U \rightarrow W$ mit $F = \bar{F} \circ \varrho$. Das kann man in Form eines kommutativen Diagramms schreiben:

$$\begin{array}{ccc} V & \xrightarrow{F} & W \\ \varrho \downarrow & \nearrow \bar{F} & \\ V/U & & \end{array}$$

Weiter ist $\text{Ker } \bar{F} = (\text{Ker } F)/U$.

Man nennt V/U den Quotientenvektorraum von V nach U . Diese Bezeichnung entspricht der Vorstellung, daß man U aus V „herausdividiert“, weil U in V/U zur Null wird.

Beweis. Zur vorübergehenden Unterscheidung werden die neu zu definierenden Verknüpfungen in V/U mit $\dot{+}$ und $\cdot$, die alten in V mit $+$ und ohne Symbol bezeichnet. Soll ϱ linear werden, so muß

$$(v + U) \dot{+} (w + U) = \varrho(v) \dot{+} \varrho(w) = \varrho(v + w) = (v + w) + U,$$

$$\lambda \cdot (v + U) = \lambda \cdot \varrho(v) = \varrho(\lambda v) = \lambda v + U$$

gelten. Also gibt es nur eine Möglichkeit, die gesuchten Verknüpfungen in V/U zu erklären:

$$(v + U) \dot{+} (w + U) := (v + w) + U, \quad \lambda \cdot (v + U) := \lambda v + U.$$

Es ist jedoch keineswegs klar, daß diese „Definition“ sinnvoll ist. Man muß noch zeigen, daß sie von der Wahl der Repräsentanten v und w unabhängig ist; dann sagt man, durch $\dot{+}$ und $\cdot$ seien Verknüpfungen in V/U wohldefiniert.

Seien also weitere Repräsentanten v', w' gegeben, d.h.

$$v + U = v' + U \quad \text{und} \quad w + U = w' + U,$$

Dann ist $v' - v \in U$ und $w' - w \in U$, also $(v' + w') - (v + w) \in U$, und somit (siehe Bild 2.8)

$$(v + w) + U = (v' + w') + U.$$

Analog zeigt man, daß

$$\lambda v + U = \lambda v' + U,$$

d.h. auch die Multiplikation mit Skalaren ist in V/U wohldefiniert.

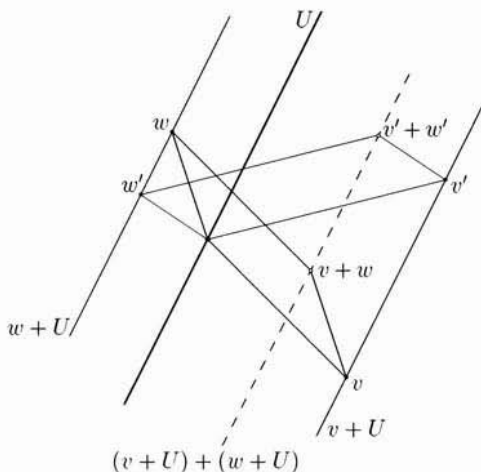


Bild 2.8

Der Nachweis der Vektorraumaxiome in V/U mit Hilfe der entsprechenden Rechenregeln in V bereitet keinerlei Probleme, das sei dem Leser zur Übung empfohlen. Nullvektor in V/U ist U , denn

$$(v + U) \dot{+} U = (v + U) \dot{+} (0 + U) = (v + 0) + U = v + U,$$

und der zu $v + U$ negative Vektor ist $-v + U$. Diese Rechnungen zeigen, daß die Unterscheidung von $+$ und $\dot{+}$ überflüssig ist.

Die zusätzlichen Aussagen sind ganz einfach. 1) folgt aus der Definition von ϱ . Ist $v + U = U$, so ist $v \in U$, also folgt 2). 3) folgt aus der Dimensionsformel in 2.2.4.

Zu 4) bemerkt man zunächst, daß wegen der Forderung $F = \overline{F} \circ \varrho$ für alle $v \in V$

$$F(v) = \overline{F}(\varrho(v)) = \overline{F}(v + U)$$

sein muß. Dadurch ist $\overline{F}$ auch wohldefiniert: denn ist $v + U = v' + U$, so folgt

$$v' - v \in U \subset \text{Ker } F, \quad \text{also} \quad F(v) = F(v').$$

Die Linearität von $\overline{F}$ ist klar. Die Gleichung $\text{Ker } \overline{F} = \text{Ker } F/U$ folgt aus

$$v + U \in \text{Ker } \overline{F} \Leftrightarrow v \in \text{Ker } F \Leftrightarrow v + U \in \text{Ker } F/U,$$

wobei zu bedenken ist, daß $\text{Ker } F/U \subset V/U$ ein Untervektorraum ist. $\square$

Eine Schwierigkeit beim Verständnis der Quotientenstruktur besteht wohl darin, daß Mengen von Vektoren (in diesem Fall affine Räume) zu neuen Vektoren

werden. Aber Vektor zu sein hat keine individuelle Bedeutung; ein Vektor muß sich nur innerhalb einer Gesamtheit von Vektoren (d.h. in einem Vektorraum) nach den dort geltenden Spielregeln (den Axiomen) verhalten. In diesem Sinne ist z.B. auch eine Funktion ein Vektor, d.h. ein Element oder „Punkt“ eines Vektorraumes (vgl. 1.4.1, Beispiel e).

2.2.8. Manchmal mag es beruhigend sein, wenn man einen abstrakten Quotientenvektorraum durch etwas Konkretes ersetzen kann. Dazu betrachten wir noch einmal die Beispiele aus 2.2.6.

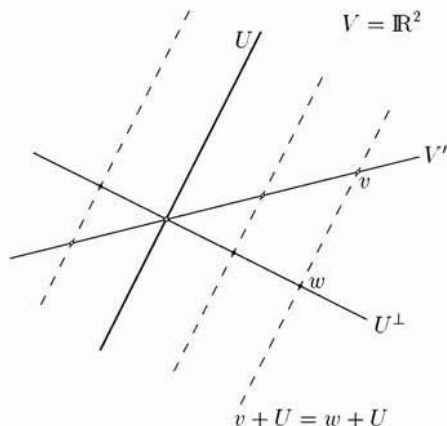


Bild 2.9

Beispiel 1. Für eine Gerade $U \subset V = \mathbb{R}^2$ ist der Quotient V/U eindimensional. Jeder affine Raum $v + U \in V/U$ kann durch einen Repräsentanten $v \in V$ gegeben werden, und man kann die Repräsentanten alle auf einen Streich in folgender Weise erhalten: Ist $V' \subset V$ eine von U verschiedene Gerade durch 0, so schneidet V' jeden affinen Raum $v + U$ in genau einem Punkt (Bild 2.9). Bezeichnet man mit

$$q': V' \rightarrow V/U, \quad v \mapsto v + U,$$

die Beschränkung der kanonischen Abbildung q , so wird q' zu einem Isomorphismus. Man kann also in gewisser Weise den abstrakten Quotientenvektorraum V/U durch einen konkreten Untervektorraum V' ersetzen. V' ist direkter Summand im Sinne von 1.6.3, d.h. es ist

$$V = U \oplus V',$$

und die Umkehrung von q' ist ein Schnitt im Sinn von 2.2.5. Aber V' hat den Nachteil, nicht eindeutig zu sein. Ein besonders ausgezeichneter direkter Summand ist die zu U senkrechte Gerade $U^\perp$ (vgl. dazu 5.4.8).

Daß die elementargeometrische Vorstellung hier nicht immer hilfreich ist, sieht man an

Beispiel 2. a) Die Elemente aus $\mathcal{C}(\mathbb{R})/\mathcal{I}(X)$ sind Klassen auf $\mathbb{R}$ stetiger Funktionen, die auf X gleich sind. Eine solche Klasse kann man stetige Funktion auf X nennen, damit hat man Stetigkeit auch auf nicht-offenen Teilmengen $X \subset \mathbb{R}$ erklärt.

Das geht zum Glück auch etwas weniger abstrakt. Sei

$$\mathcal{F}(X) = \{\varphi: X \rightarrow \mathbb{R}\}$$

der Vektorraum aller auf X definierten Funktionen und

$$\sigma: \mathcal{C}(\mathbb{R}) \rightarrow \mathcal{F}(X), \quad f \mapsto f|X,$$

der Einschränkungshomomorphismus. Wir definieren

$$\mathcal{C}(X) := \text{Im } \sigma = \{\varphi \in \mathcal{F}(X): \text{es gibt ein } f \in \mathcal{C}(\mathbb{R}) \text{ mit } \varphi = f|X\} \subset \mathcal{F}(X)$$

als den Vektorraum der auf X stetigen, d.h. auf $\mathbb{R}$ stetig fortsetzbaren Funktionen. Offenbar ist $\text{Ker } \sigma = \mathcal{I}(X)$, also hat man nach der universellen Eigenschaft des Quotientenvektorraumes ein Diagramm

$$\begin{array}{ccc} \mathcal{C}(\mathbb{R}) & \xrightarrow{\sigma} & \mathcal{F}(X) \\ q \downarrow & \nearrow \bar{\sigma} & \\ \mathcal{C}(\mathbb{R})/\mathcal{I}(X) & & \end{array}$$

wobei $\bar{\sigma}$ wegen $\text{Ker } \bar{\sigma} = \text{Ker } \sigma / \mathcal{I}(X) = 0$ injektiv ist. Der abstrakte Quotientenvektorraum $\mathcal{C}(\mathbb{R})/\mathcal{I}(X)$ kann also als Untervektorraum des konkreteren Vektorraumes $\mathcal{F}(X)$ aufgefaßt werden.

b) Der Quotientenvektorraum

$$L(\mathbb{R}) := \mathcal{C}(\mathbb{R})/\mathcal{N}$$

besteht aus den Klassen fast überall gleicher Funktionen. Im Gegensatz zu a) kann man ihn nicht als Untervektorraum von $\mathcal{F}(\mathbb{R})$ realisieren (warum?). In Aufgabe 6 zu 5.1 wird er mit einer Norm versehen. Das ergibt einen brauchbaren Begriff der Konvergenz; der Preis dafür ist, daß man Funktionen durch Äquivalenzklassen ersetzen muß.

2.2.9. Nach diesen Beispielen wieder zurück zur allgemeinen Theorie. Wir zeigen, daß man den Quotientenvektorraum weitgehend durch einen direkten Sum-

manden ersetzen kann. Dessen Existenz war im endlich-dimensionalen Fall in 1.6.3 gezeigt worden. Im allgemeinen Fall ist das zwar auch noch richtig, aber für die Praxis nutzlos.

Satz. Sei $V = V_1 \oplus V_2$ und $\varrho: V \rightarrow V/V_2$ die kanonische Abbildung. Dann ist

$$\varrho' := \varrho|_{V_1}: V_1 \rightarrow V/V_2$$

ein Isomorphismus.

Beweis. Jedes $v \in V$ hat eine eindeutige Darstellung $v = v_1 + v_2$ mit $v_1 \in V_1$ und $v_2 \in V_2$. Weiter ist

$$\varrho(v) = \varrho(v_1 + v_2) = v_1 + v_2 + V_2 = v_1 + V_2 = \varrho'(v_1).$$

Daraus folgt sofort, daß ϱ' bijektiv ist. □

Aufgaben zu 2.2

1. Sei $F: \mathbb{R}^n \rightarrow \mathbb{R}^m$ gegeben durch die folgenden Matrizen:

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}, \quad \begin{pmatrix} 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 \end{pmatrix}.$$

Bestimmen Sie jeweils Basen von $\text{Ker } F$ und $\text{Im } F$.

2. Sei $I \subset \mathbb{R}$ ein Intervall und

$$d: \mathcal{D}(I; \mathbb{R}) \rightarrow \mathcal{D}(I; \mathbb{R}), \quad f \mapsto f'.$$

Zeigen Sie, dass d eine $\mathbb{R}$ -lineare Abbildung ist, und geben Sie eine Basis von $\text{Ker } d$ an. Wie sieht $\text{Ker } d$ aus im Fall, dass I disjunkte Vereinigung von Intervallen ist?

3. Sei V ein endlichdimensionaler Vektorraum und $F: V \rightarrow V$ ein Endomorphismus. Es sei definiert: $W_0 := V$ und $W_{i+1} := F(W_i)$ für $i \in \mathbb{N}$. Dann gilt: Es gibt ein $m \in \mathbb{N}$ mit $W_{m+i} = W_m$ für alle $i \in \mathbb{N}$.

4. Sei $F: V \rightarrow V$ linear mit $F^2 = F$. Zeigen Sie, dass es Untervektorräume U, W von V gibt mit $V = U \oplus W$ und $F(W) = 0$, $F(u) = u$ für alle $u \in U$.

5. Sei $F: \mathbb{R}^3 \rightarrow \mathbb{R}^2$ gegeben durch die Matrix

$$\begin{pmatrix} 2 & 1 & 3 \\ -4 & -2 & -6 \end{pmatrix}.$$

a) Bestimmen Sie Basen $\mathcal{A} = (u, v_1, v_2)$ des $\mathbb{R}^3$ und $\mathcal{B} = (w, w')$ des $\mathbb{R}^2$, so dass

$$\text{Ker } F = \text{span}(v_1, v_2), \quad \text{Im } F = \text{span}(w) \quad \text{und} \quad F(u) = w.$$

- b) Geben Sie für $x \in \text{Im } F$ eine Parametrisierung der Faser $F^{-1}(x)$ an und zeigen Sie, dass jede nichtleere Faser $F^{-1}(x)$ genau einen Schnittpunkt mit $U = \text{span}(u)$ hat (vgl. 2.2.5).
6. Beweisen Sie das Lemma aus 1.5.8 noch einmal, aber benutzen Sie nun, dass die Projektion $\pi: W \rightarrow K^{m-1}$ linear und injektiv ist.
7. Sei $F: V \rightarrow W$ linear und $U \subset W$ ein Untervektorraum. Zeigen Sie, dass dann
- $$\dim F^{-1}(U) = \dim(U \cap \text{Im } F) + \dim \text{Ker } F.$$
8. Geben Sie einen neuen Beweis von Teil a) der Bemerkung aus 2.2.3 unter Benutzung der Äquivalenzrelation $\sim_W$ in V .
9. Zeigen Sie mit Hilfe der universellen Eigenschaft des Quotientenvektorraumes, dass für Vektorräume V, W sowie einen Untervektorraum $U \subset V$ die lineare Abbildung
- $$\{F \in \text{Hom}(V, W): F|_U = 0\} \rightarrow \text{Hom}(V/U, W) \quad \text{mit} \quad F \mapsto \bar{F}$$
- (vgl. Satz 2.2.7) ein Isomorphismus von Vektorräumen ist.

2.3 Lineare Gleichungssysteme

Mit Hilfe der bisher entwickelten Techniken von Vektorräumen und linearen Abbildungen können wir nun lineare Gleichungssysteme noch einmal von einem etwas abstrakteren Standpunkt aus behandeln und dabei auch die in Kapitel 0 versprochenen Begründungen nachliefern.

2.3.1. In Kapitel 0 hatten wir nur den reellen Vektorraum $\mathbb{R}^n$ betrachtet. Man kann ihn für einen beliebigen Körper K durch K^n ersetzen, da bei der Behandlung von linearen Gleichungssystemen nur die Körpereigenschaften der reellen Zahlen verwendet wurden.

Wir betrachten jetzt allgemein eine Matrix $A = (a_{ij}) \in M(m \times n; K)$ und eine Spalte $b = {}^t(b_1, \dots, b_m) \in M(m \times 1; K)$. Daraus ergibt sich das Gleichungssystem

$$A \cdot x = b, \quad \text{d.h.} \quad \sum_{j=1}^n a_{ij} x_j = b_i \quad \text{für } i = 1, \dots, m. \quad (**)$$

Man nennt

$$A \cdot x = 0, \quad \text{d.h.} \quad \sum_{j=1}^n a_{ij} x_j = 0 \quad \text{für } i = 1, \dots, m. \quad (*)$$

das zu (**) gehörige *homogene System*; ist $b \neq 0$, so nennt man das System (**) *inhomogen*. Die Mengen

$$\text{Lös}(A, b) := \{x \in K^n : A \cdot x = b\} \subset K^n$$

nennt man *Lösungsräume*. Entscheidend für die Theorie ist die Beziehung zu der durch A beschriebenen linearen Abbildung

$$F: K^n \rightarrow K^m, \quad x \mapsto A \cdot x,$$

denn es ist $\text{Lös}(A, b) = F^{-1}(b)$, also insbesondere $\text{Lös}(A, 0) = \text{Ker } F$.

Nach 2.2 ist die „Größe“ der Lösungsräume festgelegt durch die Zahl

$$r := \text{rang } F = \text{rang } A = \text{Spaltenrang } A.$$

Genauer folgt aus 2.2.3 und 2.2.4 das

Korollar. Gegeben sei das lineare Gleichungssystem $A \cdot x = b$ mit m Gleichungen und n Unbekannten, es sei $r = \text{rang } A$. Dann gilt für die Lösungsräume:

- 1) $\text{Lös}(A, 0) \subset K^n$ ist ein Untervektorraum der Dimension $n - r$.
- 2) $\text{Lös}(A, b) \subset K^n$ ist entweder leer oder ein affiner Raum der Dimension $n - r$.
Ist $v \in \text{Lös}(A, b)$ beliebig, so ist

$$\text{Lös}(A, b) = v + \text{Lös}(A, 0).$$

□

Anders ausgedrückt sagt man dafür: Die *allgemeine Lösung* eines inhomogenen linearen Gleichungssystems erhält man durch die Addition einer *speziellen Lösung* des inhomogenen Gleichungssystems und der *allgemeinen Lösung* des zugehörigen homogenen Gleichungssystems.

2.3.2. Der Lösungsraum eines homogenen linearen Gleichungssystems enthält als Untervektorraum stets die *triviale* Lösung 0, für ein inhomogenes System gibt es keine triviale Lösung. In 0.4.4 hatten wir durch Umformung der erweiterten Koeffizientenmatrix am Ergebnis ablesen können, ob es Lösungen gibt. Auch das kann man etwas abstrakter ansehen.

Wir vergleichen dazu die Matrizen A vom Rang r und (A, b) . Die erweiterte Matrix hat eine Spalte mehr, also ist

$$r \leq \text{rang}(A, b) \leq r + 1.$$

Satz. Der Lösungsraum des linearen Gleichungssystems $A \cdot x = b$ ist genau dann nicht leer, wenn

$$\text{rang } A = \text{rang}(A, b).$$

Diese Bedingung wurde in den Jahren 1875/76 von G. FONTENÉ, E. ROUCHÉ und F.G. FROBENIUS gefunden (vgl. [Fr]).

Beweis. A beschreibt eine lineare Abbildung

$$A: K^n \rightarrow K^m, \quad x \mapsto A \cdot x,$$

und (A, b) beschreibt eine lineare Abbildung

$$A': K^{n+1} \rightarrow K^m, \quad x' \mapsto A' \cdot x'.$$

Sind $(e_1, \dots, e_n)$ und $(e'_1, \dots, e'_{n+1})$ die kanonischen Basen, so gilt

$$A(e_1) = A'(e'_1), \dots, A(e_n) = A'(e'_n) \text{ und } A'(e'_{n+1}) = b.$$

Bei der Abbildung A' kommt also b nach Konstruktion im Bild vor, während das bei A gerade zu entscheiden ist. Wegen $\text{Im } A \subset \text{Im } A'$ ist stets

$$\text{rang } A \leq \text{rang } A'.$$

Also ist $\text{rang } A = \text{rang } A'$ gleichbedeutend mit $\text{Im } A = \text{Im } A'$, d.h. nach Definition von A' mit $b \in \text{Im } A$. $\square$

Hat man (A, b) auf Zeilenstufenform gebracht, so ist der Zeilenrang gleich der Anzahl der von Null verschiedenen Zeilen. Verwenden wir die Gleichheit von Zeilenrang und Spaltenrang (vgl. 1.5.8 und 2.6.6), so folgt die

Bemerkung. Sei

$$(A, b) = \left(\begin{array}{ccc|c} \circledast & & & \\ & \circledast & & \\ & & \ddots & \\ & & & \circledast & \\ 0 & & & & b_{r+1} \\ & & & & \vdots \\ & & & & b_m \end{array} \right)$$

Dann ist $\text{rang } A = \text{rang } (A, b)$ genau dann, wenn

$$b_{r+1} = \dots = b_m = 0.$$

□

2.3.3. Wie im reellen Fall zeigt man, daß sich jede Matrix durch elementare Zeilenumformungen auf Zeilenstufenform $\tilde{A}$ bringen läßt (vgl. 0.4.7), und daß sich der Lösungsraum nicht ändert, wenn man dabei die Spalte b zu $\tilde{b}$ mit umformt, d.h.

$$\text{Lös}(\tilde{A}, \tilde{b}) = \text{Lös}(A, b)$$

(vgl. 0.4.6).

2.3.4. Schließlich betrachten wir noch einmal den Fall, daß die erweiterte Koeffizientenmatrix in Zeilenstufenform und der Lösungsraum $\text{Lös}(A, b)$ nicht leer ist. Nach eventueller Umordnung der Spalten von A können wir annehmen, daß

$$(A, b) = \left(\begin{array}{ccc|c} a_{11} & & & b_1 \\ & a_{22} & & \vdots \\ & & \ddots & \vdots \\ & & & a_{rr} & b_r \\ 0 & & & & b_{r+1} \\ & & & & \vdots \\ & & & & b_m \end{array} \right)$$

mit $a_{ii} \neq 0$ für $i = 1, \dots, r$ und $b_{r+1} = \dots = b_m = 0$.

Die in 0.4.4 konstruierte Parametrisierung müssen wir nun genauer ansehen.

Die dort berechneten Koeffizienten tragen wir in Matrizen ein:

$$D' := (d_{ij}) \in M(r \times r; K), \quad C' := (c_{ij}) \in M(r \times k; K).$$

Diese werden vergrößert zu

$$C := \begin{pmatrix} C' \\ E_k \end{pmatrix} \in M(n \times k; K) \quad \text{und} \quad D := \begin{pmatrix} D' \\ 0 \end{pmatrix} \in M(n \times r; K),$$

das ergibt lineare Abbildungen

$$\varphi: K^r \rightarrow K^n, \quad b \mapsto D \cdot b, \quad \text{und} \quad \Phi_0: K^k \rightarrow K^n, \quad \lambda \mapsto C \cdot \lambda,$$

wobei $b = (b_1, \dots, b_r)$, und nach der in 0.4.4 ausgeführten Rechnung ist

$$\Phi_b: K^k \rightarrow K^n \quad \text{gegeben durch} \quad \Phi_b(\lambda) = \varphi(b) + \Phi_0(\lambda).$$

Nach Konstruktion von Φ_b gilt für alle $b \in K^r$

$$\Phi_b(K^k) \subset \text{Lös}(A, b) = \varphi(b) + \text{Lös}(A, 0).$$

Für $b = 0$ sieht man an der Matrix C , daß Φ_0 injektiv ist. Weiter ist nach 2.3.1 $\dim \text{Lös}(A, 0) = k$, also folgt

$$\Phi_0(K^k) = \text{Lös}(A, 0),$$

und durch Verschiebung um $\varphi(b)$ ergibt sich daraus

$$\Phi_b(K^k) = \text{Lös}(A, b).$$

Man beachte, daß hierbei wieder die Gleichheit von Zeilenrang und Spaltenrang benutzt wird: Die Dimension von $\text{Lös}(A, b)$ ist nach der Dimensionsformel aus 2.2.4 durch den Spaltenrang bestimmt, an der Zeilenstufenform liest man den Zeilenrang ab. Insgesamt ist folgendes bewiesen:

Satz. Sei (A, b) in Zeilenstufenform mit $r = \text{Zeilenrang } A$ und $b \in K^r$. Dann hat die in 0.4.4 konstruierte von b abhängige Parametrisierung

$$\Phi_b: K^{n-r} \rightarrow \text{Lös}(A, b) \subset K^n$$

folgende Eigenschaften:

- 1) $\Phi_0: K^{n-r} \rightarrow \text{Lös}(A, 0) \subset K^n$ ist ein Vektorraumisomorphismus.
- 2) $\Phi_b: K^{n-r} \rightarrow \text{Lös}(A, b) \subset K^n$ ist für jedes $b \in K^r$ bijektiv.
- 3) Es gibt einen Homomorphismus $\varphi: K^r \rightarrow K^n$, so daß für alle $b \in K^r$

$$\Phi_b = \varphi(b) + \Phi_0 \quad \text{und} \quad \text{Lös}(A, b) = \varphi(b) + \text{Lös}(A, 0). \quad \square$$

Damit ist die Frage nach Lösbarkeit von linearen Gleichungssystemen beantwortet und auch die Abhängigkeit der Lösungsmenge von der „rechten Seite“ b explizit beschrieben. Die oben angegebene Abbildung $\varphi: K^r \rightarrow K^n$ ist ein Schnitt im Sinne von 2.2.5.

Ist das System nicht in Zeilenstufenform, so kann man zeigen, daß der Übergang vom ursprünglichen b zum umgeformten $\tilde{b}$ durch einen Isomorphismus

$$S: K^m \rightarrow K^m, \quad b \mapsto \tilde{b},$$

beschrieben wird. In 2.7.7 zeigen wir, wie man S berechnen kann.

2.3.5. Zur Beschreibung der Lösungen eines linearen Gleichungssystems benutzt man oft eine weitere Sprechweise. Eine Basis $(w_1, \dots, w_k)$ von $\text{Lös}(A, 0)$ heißt *Fundamentalsystem von Lösungen des homogenen Systems*. Ein beliebiges $v \in \text{Lös}(A, b)$ heißt *spezielle Lösung des inhomogenen Systems*. Dann hat man eine Darstellung

$$\text{Lös}(A, b) = v + K w_1 + \dots + K w_k,$$

und die Linearkombinationen der Lösungen sind eindeutig. Man erhält die Vektoren $w_1, \dots, w_k$ als Spalten der in 2.3.4 konstruierten Matrix C . Dabei ist jedoch Vorsicht geboten, weil die Zeilen von C anders angeordnet sind, wenn die Pivotspalten nicht die Voraussetzung $j_i = i$ erfüllen.

Es lohnt sich nicht, den allgemeinen Fall aufzuschreiben, wir geben lieber ein typisches

Beispiel. Wir betrachten das Gleichungssystem in Zeilenstufenform und mit beliebiger rechter Seite:

$$\left(\begin{array}{cccccc|c} 0 & 1 & 0 & 2 & -1 & -4 & 0 & b_1 \\ 0 & 0 & 1 & -1 & -1 & 2 & 1 & b_2 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & b_3 \\ 0 & 0 & 0 & 0 & 0 & 1 & 2 & b_4 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right)$$

mit $n = 7$, $m = 5$, $r = 4$, $j_1 = 2$, $j_2 = 3$, $j_3 = 5$, $j_4 = 6$ und $k = 3$. Dann ist $x_1 = \lambda_1$, $x_4 = \lambda_2$, $x_7 = \lambda_3$, und daraus ergibt sich

$$\begin{aligned} x_6 &= b_4 - 2x_7 = b_4 - 2\lambda_3, \\ x_5 &= b_3 - x_6 = b_3 - b_4 + 2\lambda_3, \\ x_3 &= b_2 + x_4 + x_5 - 2x_6 - x_7 = b_2 + b_3 - 3b_4 + \lambda_2 + 5\lambda_3, \\ x_2 &= b_1 - 2x_4 + x_5 + 4x_6 = b_1 + b_3 + 3b_4 - 2\lambda_2 - 6\lambda_3. \end{aligned}$$

Trägt man die erhaltenen Koeffizienten in die entsprechenden Matrizen ein, so

wird

$$D = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 3 \\ 0 & 1 & 1 & -3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -2 & -6 \\ 0 & 1 & 5 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & -2 \\ 0 & 0 & 1 \end{pmatrix}.$$

Die Spalten von C sind das Fundamentalsystem, $D \cdot b$ ist für jedes $b \in K^4$ eine spezielle Lösung.

2.3.6. Zwei wichtige Spezialfälle linearer Gleichungssysteme haben eigene Namen. Besteht der Lösungsraum aus genau einem Element, so nennt man das System *eindeutig lösbar*. Aus dem bisher Bewiesenen folgt sofort die

Bemerkung. Für $A \in M(m \times n; K)$ und $b \in K^m$ sind folgende Bedingungen gleichwertig:

- i) Das lineare Gleichungssystem $A \cdot x = b$ ist *eindeutig lösbar*.
- ii) $\text{rang } A = \text{rang } (A, b) = n$. □

In diesem Fall besitzt das zugehörige homogene Gleichungssystem $A \cdot x = 0$ nur die *triviale Lösung* 0.

Ist $m = n$, so kann man ii) ersetzen durch $\text{rang } A = n$. Dies bedeutet, daß die lineare Abbildung $A: K^n \rightarrow K^n$ surjektiv, also nach Korollar 3 aus 2.2.4 sogar bijektiv ist. Bezeichnet A^{-1} die inverse Abbildung, so ist die eindeutige Lösung x gegeben durch

$$x = A^{-1}(b).$$

Nach 2.5.5 ist A^{-1} beschrieben durch die inverse Matrix.

Ist die Matrix $A \in M(m \times n; K)$ vom Rang m , so ist die lineare Abbildung $A: K^n \rightarrow K^m$ surjektiv, also ist der Lösungsraum von $A \cdot x = b$ für jedes $b \in K^m$ nicht leer. Ein solches Gleichungssystem nennt man *universell lösbar*.

Ist der Rang von A kleiner als m , so ist das System nur für spezielle b lösbar. Ein Rechenverfahren, dies bei festem A für ein gegebenes b zu entscheiden, leiten wir in 2.7.7 ab.

Aufgaben zu 2.3

1. Ein Nahrungsmittel enthält Schadstoffe $S_1, \dots, S_5$, die bei der Produktion und Lagerung als Bestandteile von Pflanzenschutzmitteln auftreten. Auf den einzelnen Stationen werden die folgenden Pflanzenschutzmittel benutzt:

Station	Mittel
1. Landwirt	A
2. Rohproduktlagerung	B
3. Veredelungsbetrieb	C
4. Grossist und Transport	D
5. Einzelhändler	E

Die folgende Tabelle gibt die prozentuale Zusammensetzung der Mittel A, ..., E wieder:

	S_1	S_2	S_3	S_4	S_5
A	0.2	0.5	0	0.3	0
B	0.1	0.6	0.3	0	0
C	0.1	0.2	0.2	0.3	0.2
D	0	0	0.1	0.4	0.5
E	0	0.1	0.3	0.3	0.3

Für das fertige Produkt ergibt die Nahrungsmittelanalyse die folgenden Werte (in Gewichtseinheiten):

S_1	S_2	S_3	S_4	S_5
0.75	2.25	0.65	1.60	0.75

Ermitteln Sie, wieviel (in Gewichtseinheiten) die einzelnen Stationen zur Schadstoffbelastung beitragen.

2. Es seien Metall-Legierungen M_1, M_2 und M_3 gegeben, die alle Kupfer, Silber und Gold enthalten, und zwar in folgenden Prozentsätzen:

	Kupfer	Silber	Gold
M_1	20	60	20
M_2	70	10	20
M_3	50	50	0

Kann man diese Legierungen so mischen, dass eine Legierung entsteht, die 40% Kupfer, 50% Silber und 10% Gold enthält?

3. Zeigen Sie: Ist die Matrix $A \in M(m \times n; K)$ in Zeilenstufenform und r der Rang von A , so ist $(e_1, \dots, e_r)$ eine Basis von $\text{Im } A \subset K^m$.

4. Bestimmen Sie für das folgende Gleichungssystem in Zeilenstufenform mit beliebiger rechter Seite Matrizen C und D wie in 2.3.4, so dass die Spalten von C ein Fundamentalsystem bilden und $D \cdot b$ für jedes $b \in \mathbb{R}^5$ eine spezielle Lösung ist.

$$\left(\begin{array}{cccccc|c} 0 & 1 & -1 & 2 & 0 & 3 & b_1 \\ 0 & 0 & 2 & -1 & 2 & 0 & b_2 \\ 0 & 0 & 0 & -1 & 4 & 0 & b_3 \\ 0 & 0 & 0 & 0 & 0 & -7 & b_4 \\ 0 & 0 & 0 & 0 & 0 & 0 & b_5 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right)$$

5. Gegeben seien die Matrizen

$$A = \begin{pmatrix} 3 & 5 & 7 \\ 4 & 6 & 8 \\ 1 & 3 & 4 \end{pmatrix}, \quad B = \begin{pmatrix} 3 & 2 & 6 & 3 \\ 2 & 1 & 3 & 2 \\ 2 & 3 & 1 & 4 \end{pmatrix}.$$

a) Untersuchen Sie die folgenden Gleichungssysteme darauf, ob sie eindeutig lösbar sind:

$$Ax = \begin{pmatrix} 2 \\ 4 \\ 9 \end{pmatrix}, \quad Bx = \begin{pmatrix} 4 \\ 1 \\ 7 \end{pmatrix}.$$

b) Untersuchen Sie die Gleichungssysteme $Ax = b$ und $Bx = b$ für beliebige $b \in \mathbb{R}^3$ darauf, ob sie universell lösbar sind.

6. Sei der Untervektorraum $W \subset \mathbb{R}^n$ gegeben durch m lineare Gleichungen $\varphi_1, \dots, \varphi_m$, d. h.

$$W = \{x \in \mathbb{R}^n : \varphi_1(x) = \dots = \varphi_m(x) = 0\}.$$

Zeigen Sie, dass dann W bereits durch eine einzige (nicht notwendig lineare) Gleichung beschrieben werden kann. Genauer gilt: Es existiert ein Polynom $f \in \mathbb{R}[t_1, \dots, t_n]$ mit

$$W = \{(x_1, \dots, x_n) \in \mathbb{R}^n : f(x_1, \dots, x_n) = 0\}.$$

Zeigen Sie, dass diese Aussage auch gilt, falls man $\mathbb{R}$ durch einen endlichen Körper K ersetzt.

7. Finden Sie neue (kürzere) Beweise für Satz 0.2.4 und Aufgabe 2a) zu 0.3.

8. Zeigen Sie, dass eine Teilmenge L des $\mathbb{R}^3$ eine Gerade ist (d. h. es existieren $v, w \in \mathbb{R}^3$, $w \neq 0$, mit $L = v + \mathbb{R}w$) genau dann, wenn es eine Matrix $A \in M(2 \times 3; \mathbb{R})$ mit $\text{rang } A = 2$ und ein $b \in \mathbb{R}^2$ gibt, so dass $L = \{x \in \mathbb{R}^3 : Ax = b\}$. Was bedeutet das geometrisch?

2.4 Lineare Abbildungen und Matrizen

Eine Abbildung $F: X \rightarrow Y$ von Mengen ist nach Definition eine Vorschrift, die jedem Argument $x \in X$ einen Wert $F(x) \in Y$ zuordnet. Dabei ist im allgemeinen nichts darüber ausgesagt, wie die Vorschrift auszusehen hat, also kann man die Bilder verschiedener Argumente völlig unabhängig voneinander wählen.

Ganz anders ist die Situation für eine lineare Abbildung $F: V \rightarrow W$ zwischen Vektorräumen. Kennt man einen Wert $F(v)$, so ist F auf der ganzen Geraden Kv festgelegt. Will man für einen weiteren Vektor $v' \in V$ den Wert beliebig vorschreiben, so darf also v' nicht auf der Geraden Kv liegen. Durch $F(v)$ und $F(v')$ ist dann F auf der ganzen Ebene $Kv + Kv'$ festgelegt, usw.

2.4.1. Die Frage, durch wieviele Vorgaben eine lineare Abbildung festgelegt ist, hat eine einfache Antwort:

Satz. Gegeben seien endlichdimensionale Vektorräume V und W , sowie Vektoren $v_1, \dots, v_r \in V$ und $w_1, \dots, w_r \in W$. Dann gilt

- 1) Sind $v_1, \dots, v_r$ linear unabhängig, so gibt es mindestens eine lineare Abbildung

$$F: V \rightarrow W \quad \text{mit} \quad F(v_i) = w_i \quad \text{für} \quad i = 1, \dots, r.$$

- 2) Ist $(v_1, \dots, v_r)$ eine Basis, so gibt es genau eine lineare Abbildung

$$F: V \rightarrow W \quad \text{mit} \quad F(v_i) = w_i \quad \text{für} \quad i = 1, \dots, r.$$

Dieses F hat folgende Eigenschaften:

- a) $\text{Im } F = \text{span}(w_1, \dots, w_r)$.
 b) F injektiv $\Leftrightarrow w_1, \dots, w_r$ linear unabhängig.

Beweis. Wir beginnen mit Teil 2). Jedes $v \in V$ hat eine eindeutige Darstellung

$$v = \lambda_1 v_1 + \dots + \lambda_r v_r,$$

wegen $F(v_i) = w_i$ und der Linearität von F muß also

$$F(v) = \lambda_1 w_1 + \dots + \lambda_r w_r \tag{*}$$

sein. Also gibt es höchstens ein solches F , nämlich das durch (*) erklärte. Man darf nun allerdings nicht versäumen zu zeigen, daß die durch (*) erklärte Abbildung wirklich linear ist. Das folgt aus den Rechnungen

$$\begin{aligned} F(v + v') &= F(\lambda_1 v_1 + \dots + \lambda_r v_r + \lambda'_1 v_1 + \dots + \lambda'_r v_r) \\ &= F((\lambda_1 + \lambda'_1)v_1 + \dots + (\lambda_r + \lambda'_r)v_r) \\ &= (\lambda_1 + \lambda'_1)w_1 + \dots + (\lambda_r + \lambda'_r)w_r \\ &= \lambda_1 w_1 + \dots + \lambda_r w_r + \lambda'_1 w_1 + \dots + \lambda'_r w_r \\ &= F(v) + F(v') \end{aligned}$$

und

$$F(\lambda v) = F(\lambda \lambda_1 v_1 + \dots + \lambda \lambda_r v_r) = \lambda \lambda_1 w_1 + \dots + \lambda \lambda_r w_r = \lambda F(v).$$

Die Inklusion $\text{Im } F \subset \text{span}(w_1, \dots, w_r)$ ist klar. Ist umgekehrt

$$w = \mu_1 w_1 + \dots + \mu_r w_r, \quad \text{so folgt} \quad w = F(\mu_1 v_1 + \dots + \mu_r v_r).$$

Zu b) nehmen wir an, $w_1, \dots, w_r$ sei linear abhängig. Dann gibt es $(\mu_1, \dots, \mu_r) \neq (0, \dots, 0)$ mit

$$\mu_1 w_1 + \dots + \mu_r w_r = 0,$$

und es folgt $F(\mu_1 v_1 + \dots + \mu_r v_r) = 0$; also ist F nicht injektiv. Umgekehrt sei $F(v) = 0$. Wir schreiben

$$v = \lambda_1 v_1 + \dots + \lambda_r v_r, \quad \text{dann ist} \quad \lambda_1 w_1 + \dots + \lambda_r w_r = 0.$$

Wegen der linearen Unabhängigkeit von $w_1, \dots, w_r$ folgt $\lambda_1 = \dots = \lambda_r = 0$, also $v = 0$. Damit ist 2) bewiesen.

Ist $v_1, \dots, v_r$ nun linear unabhängig, so können wir diese Familie zu einer Basis

$$(v_1, \dots, v_r, v_{r+1}, \dots, v_n)$$

ergänzen und durch Vorgabe beliebiger weiterer Werte $w_{r+1}, \dots, w_n$ entsprechend 2) ein F mit

$$F(v_i) = w_i \quad \text{für } i = 1, \dots, n$$

finden. An dieser Konstruktion kann man erkennen, wie weit F von der Eindeutigkeit entfernt ist: ein Maß dafür ist die Zahl $n - r$. $\square$

2.4.2. Der Satz aus 2.4.1 über die *Erzeugung von linearen Abbildungen* hat zahlreiche Folgerungen.

Korollar 1. Ist V ein Vektorraum mit einer Basis $\mathcal{B} = (v_1, \dots, v_n)$, so gibt es dazu genau einen Isomorphismus

$$\Phi_{\mathcal{B}}: K^n \rightarrow V \quad \text{mit} \quad \Phi_{\mathcal{B}}(e_j) = v_j \quad \text{für } j = 1, \dots, n,$$

wobei $(e_1, \dots, e_n)$ die kanonische Basis von K^n bezeichnet. $\square$

$\Phi_{\mathcal{B}}$ heißt *Koordinatensystem*, damit werden wir uns in 2.6.1 weiter beschäftigen.

Korollar 2. Zu jeder linearen Abbildung $F: K^n \rightarrow K^m$ gibt es genau eine Matrix $A \in M(m \times n; K)$, so daß

$$F(x) = A \cdot x$$

für alle Spaltenvektoren $x \in K^n$.

Man braucht also in diesem Fall zwischen linearen Abbildungen und Matrizen nicht mehr zu unterscheiden.

Beweis. Man schreibe $F(e_1), \dots, F(e_n)$ als Spaltenvektoren nebeneinander, das ergibt A . $\square$

Einen solchen Zusammenhang zwischen linearen Abbildungen und Matrizen gibt es nicht nur in den Standardräumen:

Satz. Gegeben seien K -Vektorräume

V mit Basis $\mathcal{A} = (v_1, \dots, v_n)$ und W mit Basis $\mathcal{B} = (w_1, \dots, w_m)$.

Dann gibt es zu jeder linearen Abbildung $F: V \rightarrow W$ genau eine Matrix $A = (a_{ij}) \in M(m \times n; K)$, so daß

$$F(v_j) = \sum_{i=1}^m a_{ij} w_i \quad \text{für } j = 1, \dots, n, \quad (*)$$

und die so erhaltene Abbildung

$$M_{\mathcal{B}}^{\mathcal{A}}: \text{Hom}(V, W) \rightarrow M(m \times n; K), \quad F \mapsto A = M_{\mathcal{B}}^{\mathcal{A}}(F),$$

ist ein Isomorphismus von K -Vektorräumen. Insbesondere gilt

$$M_{\mathcal{B}}^{\mathcal{A}}(F+G) = M_{\mathcal{B}}^{\mathcal{A}}(F) + M_{\mathcal{B}}^{\mathcal{A}}(G) \quad \text{und} \quad M_{\mathcal{B}}^{\mathcal{A}}(\lambda F) = \lambda M_{\mathcal{B}}^{\mathcal{A}}(F).$$

Kurz gesagt: Nach Wahl fester Basen kann man lineare Abbildungen durch Matrizen (d.h. relativ abstrakte durch konkrete Objekte) ersetzen. Man bezeichnet $M_{\mathcal{B}}^{\mathcal{A}}(F)$ als die Matrix, die F bezüglich der Basen $\mathcal{A}$ und $\mathcal{B}$ darstellt.

Beweis. Da $\mathcal{B}$ Basis ist, sind die Linearkombinationen aus (*) und somit auch die Spalten der Matrix A eindeutig bestimmt. Gehört zur Abbildung G die Matrix $B = (b_{ij})$, so ist

$$(F+G)(v_j) = F(v_j) + G(v_j) = \sum_{i=1}^m a_{ij} w_i + \sum_{i=1}^m b_{ij} w_i = \sum_{i=1}^m (a_{ij} + b_{ij}) w_i,$$

und für $\lambda \in K$ ist

$$(\lambda F)(v_j) = \lambda \cdot F(v_j) = \lambda \sum_{i=1}^m a_{ij} w_i = \sum_{i=1}^m (\lambda a_{ij}) w_i.$$

Daher ist die Abbildung $M_{\mathcal{B}}^{\mathcal{A}}$ linear. Da $\mathcal{A}$ eine Basis ist, gibt es nach 2.4.1 genau ein F , das darauf die durch Bedingung (*) festgelegten Werte annimmt. Also ist $M_{\mathcal{B}}^{\mathcal{A}}$ bijektiv. $\square$

Zusatz. Im Spezialfall $V = K^n$ und $W = K^m$ mit den kanonischen Basen $\mathcal{K}$ und $\mathcal{K}'$ ist der kanonische Isomorphismus

$$M_{\mathcal{K}'}^{\mathcal{K}}: \text{Hom}(K^n, K^m) \rightarrow M(m \times n; K), \quad F \mapsto A,$$

die in Korollar 2 beschriebene Beziehung. □

Mit Hilfe einer Basis kann man Vektoren eindeutig als Linearkombinationen darstellen. Der obige Satz zeigt, wie man Abbildungen als Vektoren betrachten und mit Hilfe zweier Basen analog verfahren kann. Dazu sei

$$F_i^j: V \rightarrow W \quad \text{erklärt durch} \quad F_i^j(v_k) := \begin{cases} w_i & \text{für } k = j, \\ 0 & \text{sonst.} \end{cases}$$

Dann ist $M_{\mathcal{B}}^{\mathcal{A}}(F_i^j) = E_j^i$ (mit der Bezeichnung aus 1.5.1), und die $m \cdot n$ Abbildungen F_i^j bilden eine Basis von $\text{Hom}(V, W)$. Die zur Linearkombination eines beliebigen F nötigen Skalare stehen an passender Stelle in $M_{\mathcal{B}}^{\mathcal{A}}(F)$.

Die naheliegende Frage, wie sich die Matrix A ändert, wenn man in V und W neue Basen einführt, wird in 2.6.5 beantwortet.

2.4.3. Als Folgerung aus 2.2.4 erhält man, daß bei Benutzung der dort konstruierten Basen auch die darstellende Matrix besonders einfach wird.

Korollar. Sei $F: V \rightarrow W$ linear, $n = \dim V$, $m = \dim W$ und $r = \dim \text{Im } F$. Dann gibt es Basen $\mathcal{A}$ von V und $\mathcal{B}$ von W , so daß

$$M_{\mathcal{B}}^{\mathcal{A}}(F) = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}.$$

Dabei bezeichnet E_r die in 1.5.7 eingeführte r -reihige Einheitsmatrix.

Beweis. Es genügt, die in 2.2.4 gewählte Basis von $\text{Im } F$ zu einer Basis

$$\mathcal{B} = (w_1, \dots, w_r, w_{r+1}, \dots, w_m)$$

von W zu ergänzen. □

2.4.4. Ist der Bildraum W gleich dem Urbildraum V (d.h. hat man einen Endomorphismus), so setzt man am besten auch $\mathcal{A} = \mathcal{B}$ und zur Vereinfachung der Notation $M_{\mathcal{B}} := M_{\mathcal{B}}^{\mathcal{B}}$, sowie $\text{End}(V) = \text{Hom}(V, V)$. Der Vektorraumisomorphismus

$$M_{\mathcal{B}}: \text{End}(V) \rightarrow M(n \times n; K)$$

ist dann charakterisiert durch die Gleichungen

$$F(v_j) = \sum_{i=1}^n a_{ij} v_i \quad \text{für } j = 1, \dots, n,$$

wenn $\mathcal{B} = (v_1, \dots, v_n)$ und $A = (a_{ij}) = M_{\mathcal{B}}(F)$.

Die n -reihige Einheitsmatrix $E_n = (\delta_{ij})$ beschreibt dabei die identische Abbildung, in Zeichen

$$M_{\mathcal{B}}(\text{id}_V) = E_n.$$

Die zu 2.4.3 analoge Frage, für einen Endomorphismus *eine* besonders einfache Basis zu finden, ist weit schwieriger zu beantworten. Sie ist Gegenstand von Kapitel 4.

Aufgaben zu 2.4

1. Gibt es eine lineare Abbildung $F: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ mit

$$F(2, 0) = (0, 1), \quad F(1, 1) = (5, 2), \quad F(1, 2) = (2, 3) ?$$

2. Sei $\mathcal{B} = (\sin, \cos, \sin \cdot \cos, \sin^2, \cos^2)$ und $V = \text{span } \mathcal{B} \subset \text{Abb}(\mathbb{R}, \mathbb{R})$. Betrachten Sie den Endomorphismus $F: V \rightarrow V$, $f \mapsto f'$, wobei f' die erste Ableitung von f bezeichnet.

a) Zeigen Sie, dass $\mathcal{B}$ eine Basis von V ist.

b) Bestimmen Sie die Matrix $M_{\mathcal{B}}(F)$.

c) Bestimmen Sie Basen von $\text{Ker } F$ und $\text{Im } F$.

3. Für $n \in \mathbb{N}$ sei $V_n = \text{span}(1, \dots, t^n) \subset \mathbb{R}[t]$ mit der Basis $\mathcal{B}_n = (1, \dots, t^n)$ und

$$\mathcal{D}_n: V_n \rightarrow V_{n-1}, \quad f \mapsto f'$$

der Ableitungshomomorphismus.

a) Bestimmen Sie die Matrix $M_{\mathcal{B}_{n-1}}^{\mathcal{B}_n}(\mathcal{D}_n)$.

b) Zeigen Sie, dass es eine lineare Abbildung $\mathcal{I}_n: V_{n-1} \rightarrow V_n$ gibt mit $\mathcal{D}_n \circ \mathcal{I}_n = \text{id}$, und bestimmen Sie $M_{\mathcal{B}_n}^{\mathcal{B}_{n-1}}(\mathcal{I}_n)$.

4. Sei $V = \{f \in \mathbb{R}[t]: \deg f \leq 3\}$ mit der Basis $\mathcal{B} = (1, t, t^2, t^3)$. Wir betrachten die linearen Abbildungen

$$F: V \rightarrow \mathbb{R}, \quad f \mapsto \int_{-1}^1 f(t) dt \quad \text{und} \quad G: V \rightarrow \mathbb{R}^3, \quad f \mapsto (f(-1), f(0), f(1)).$$

a) Es seien $\mathcal{K}$ und $\mathcal{K}'$ die kanonischen Basen von $\mathbb{R}$ und $\mathbb{R}^3$. Bestimmen Sie die Matrizen

$$M_{\mathcal{K}}^{\mathcal{B}}(F) \text{ und } M_{\mathcal{K}'}^{\mathcal{B}}(G).$$

b) Zeigen Sie: $\text{Ker } G \subset \text{Ker } F$.

c) Es gibt eine lineare Abbildung $H: \mathbb{R}^3 \rightarrow \mathbb{R}$ mit $H \circ G = F$.

5. Seien V und W endlichdimensionale Vektorräume mit $V = V_1 \oplus V_2$, $W = W_1 \oplus W_2$ sowie $F: V \rightarrow W$ linear mit $F(V_i) \subset W_i$ für $i = 1, 2$. Zeigen Sie, dass es Basen $\mathcal{A}$ von V und $\mathcal{B}$ von W gibt mit

$$M_{\mathcal{B}}^{\mathcal{A}}(F) = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix},$$

wobei $A \in M(\dim W_1 \times \dim V_1; K)$, $B \in M(\dim W_2 \times \dim V_2; K)$.

6. Zeigen Sie ohne Verwendung von Matrizen, dass die in 2.4.2 definierten Abbildungen $F_i^j: V \rightarrow W$ eine Basis von $\text{Hom}(V, W)$ bilden.

7. Sei

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

und $F: \mathbb{R}^4 \rightarrow \mathbb{R}^3$ die durch $F(x) = Ax$ definierte lineare Abbildung. Bestimmen Sie Basen $\mathcal{A}$ von $\mathbb{R}^4$ und $\mathcal{B}$ von $\mathbb{R}^3$ mit

$$M_{\mathcal{B}}^{\mathcal{A}}(F) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

8. Sei V ein endlichdimensionaler Vektorraum und $F: V \rightarrow V$ linear mit $F^2 = F$. Zeigen Sie, dass es eine Basis $\mathcal{B}$ von V gibt mit

$$M_{\mathcal{B}}(F) = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}.$$

Hinweis: Aufgabe 5 und Aufgabe 4 zu 2.2.

9. Zeigen Sie: Ist $F: V \rightarrow V$ ein Endomorphismus des endlichdimensionalen Vektorraums V mit $\dim \text{Fix } F = r$ (vgl. Aufgabe 3 zu 2.1), so existiert eine Basis $\mathcal{B}$ von V mit

$$M_{\mathcal{B}}(F) = \begin{pmatrix} E_r & * \\ 0 & * \end{pmatrix}.$$

2.5 Multiplikation von Matrizen

2.5.1. In 2.1.3 hatten wir gesehen, daß für lineare Abbildungen $G: U \rightarrow V$ und $F: V \rightarrow W$ die Komposition $F \circ G: U \rightarrow W$ wieder linear ist. Im Spezialfall $U = K^r$, $V = K^n$, $W = K^m$ sind F und G entsprechend 2.4.2 durch Matrizen A und B gegeben, und wir wollen ausrechnen, welche Matrix C die lineare Abbildung $F \circ G$ beschreibt. Das folgende Diagramm soll die Übersicht erleichtern:

$$\begin{array}{ccccc}
 K^r & \xrightarrow{B} & K^n & \xrightarrow{A} & K^m, \\
 x = \begin{pmatrix} x_1 \\ \vdots \\ x_r \end{pmatrix} & \mapsto & y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} & \mapsto & z = \begin{pmatrix} z_1 \\ \vdots \\ z_m \end{pmatrix}.
 \end{array}$$

Ist $B = (b_{jk}) \in M(n \times r; K)$ und $y = B(x)$, so folgt

$$y_j = b_{j1}x_1 + \dots + b_{jr}x_r \quad \text{für } j = 1, \dots, n, \quad (\text{b})$$

$A = (a_{ij}) \in M(m \times n; K)$ und $z = A(y)$ bedeutet

$$z_i = a_{i1}y_1 + \dots + a_{in}y_n \quad \text{für } i = 1, \dots, m, \quad (\text{a})$$

und schließlich $C = (c_{ik}) \in M(m \times r; K)$, $z = A(B(x))$ bedeutet

$$z_i = c_{i1}x_1 + \dots + c_{ir}x_r \quad \text{für } i = 1, \dots, m. \quad (\text{c})$$

Setzt man (b) in (a) ein, so erhält man

$$\begin{aligned}
 z_i &= a_{i1}(b_{11}x_1 + \dots + b_{1r}x_r) + \dots + a_{in}(b_{n1}x_1 + \dots + b_{nr}x_r) \\
 &= (a_{i1}b_{11} + \dots + a_{in}b_{n1})x_1 + \dots + (a_{i1}b_{1r} + \dots + a_{in}b_{nr})x_r.
 \end{aligned} \quad (\text{c}')$$

Vergleich von (c) und (c') ergibt

$$c_{i1} = a_{i1}b_{11} + \dots + a_{in}b_{n1}, \dots, c_{ir} = a_{i1}b_{1r} + \dots + a_{in}b_{nr}.$$

Unter Verwendung von Summenzeichen schreibt sich das so:

$$\begin{aligned}
 z_i &= \sum_{j=1}^n a_{ij}y_j = \sum_{j=1}^n a_{ij} \left(\sum_{k=1}^r b_{jk}x_k \right) = \sum_{j=1}^n \left(\sum_{k=1}^r a_{ij}b_{jk}x_k \right) \\
 &= \sum_{k=1}^r \left(\sum_{j=1}^n a_{ij}b_{jk} \right) x_k = \sum_{k=1}^r c_{ik}x_k,
 \end{aligned}$$

also ist

$$c_{ik} = \sum_{j=1}^n a_{ij}b_{jk} \quad \text{für } i = 1, \dots, m \text{ und } k = 1, \dots, r.$$

Diese kleine Rechnung, die nur auf der Umordnung einer Summe beruht, hat wichtige Konsequenzen. Man kann damit eine Multiplikation von Matrizen passender Größe erklären.

2.5.2. Zur Definition der *Multiplikation von Matrizen* nehmen wir

$$A = (a_{ij}) \in M(m \times n; K) \quad \text{und} \quad B = (b_{jk}) \in M(n \times r; K),$$

d.h. die Spaltenzahl von A muß mit der Zeilenzahl von B übereinstimmen. Dann ist das Produkt

$$A \cdot B = (c_{ik}) \in M(m \times r; K) \quad \text{erklärt durch} \quad c_{ik} := \sum_{j=1}^n a_{ij} b_{jk}.$$

Die Matrix $A \cdot B$ hat also so viele Zeilen wie A und so viele Spalten wie B , die gemeinsame Zahl n verschwindet bei der Multiplikation. Diese Größenverhältnisse kann man durch folgendes Schema zum Ausdruck bringen:

$$\underbrace{\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{i1} & \cdots & a_{in} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}}_n \cdot \underbrace{\begin{pmatrix} b_{11} & \cdots & b_{1k} & \cdots & b_{1r} \\ \vdots & & \vdots & & \vdots \\ b_{n1} & \cdots & b_{nk} & \cdots & b_{nr} \\ \vdots & & c_{ik} & & \vdots \\ c_{m1} & \cdots & c_{mr} \end{pmatrix}}_r = \begin{pmatrix} c_{11} & \cdots & c_{1r} \\ \vdots & & \vdots \\ c_{m1} & \cdots & c_{mr} \end{pmatrix}_m.$$

Hieran sieht man auch gut, wie

$$c_{ik} = a_{i1}b_{1k} + \cdots + a_{in}b_{nk}$$

aus der i -ten Zeile von A und der k -ten Spalte von B entsteht. Ansonsten ist es aber recht unpraktisch, Matrizen bei der Multiplikation so anzuschreiben. Ein Beispiel der üblichen Schreibweise ist

$$\begin{pmatrix} 1 & 2 & -1 & 1 \\ 0 & 1 & 2 & -2 \\ 2 & -1 & 0 & -3 \end{pmatrix} \cdot \begin{pmatrix} 1 & -2 \\ 0 & -1 \\ 1 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & -5 \\ 0 & 1 \\ -1 & -3 \end{pmatrix}.$$

Ist speziell $m = r = 1$ und n beliebig, so ist

$$A \cdot B = (a_1 \dots a_n) \cdot \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = a_1 b_1 + \dots + a_n b_n \in K = M(1 \times 1; K).$$

Man beachte, daß dagegen

$$\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \cdot (b_1 \dots b_n) = \begin{pmatrix} a_1 b_1 & \dots & a_1 b_n \\ \vdots & & \vdots \\ a_n b_1 & \dots & a_n b_n \end{pmatrix} \in M(n \times n; K)$$

gilt.

Ist speziell $m = n = r$, so kann man für $A, B \in M(m \times m; K)$ sowohl $A \cdot B$ als auch $B \cdot A$ bilden, im allgemeinen ist aber $A \cdot B \neq B \cdot A$. Zum Beispiel ist

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},$$

aber

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

Daran sieht man außerdem, daß das Produkt von zwei Matrizen die Nullmatrix ergeben kann, obwohl beide Matrizen von der Nullmatrix verschieden waren.

Beispiel. Ist $f = (f_1, \dots, f_m): \mathbb{R}^n \rightarrow \mathbb{R}^m$ eine differenzierbare Abbildung (das heißt $f_1, \dots, f_m: \mathbb{R}^n \rightarrow \mathbb{R}$ sind differenzierbare Funktionen) mit $f(0) = 0$ (diese Annahme dient nur zur Vereinfachung der Bezeichnungen), und sind $y_1, \dots, y_n$ die Koordinaten im $\mathbb{R}^n$, so sei

$$A := \begin{pmatrix} \frac{\partial f_1}{\partial y_1}(0) & \dots & \frac{\partial f_1}{\partial y_n}(0) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial y_1}(0) & \dots & \frac{\partial f_m}{\partial y_n}(0) \end{pmatrix}$$

die sogenannte **JACOBI-Matrix** von f im Punkte Null.

Ist $g = (g_1, \dots, g_r): \mathbb{R}^r \rightarrow \mathbb{R}^n$ eine weitere differenzierbare Abbildung mit $g(0) = 0$, und sind $x_1, \dots, x_r$ Koordinaten im $\mathbb{R}^r$, so bezeichnen wir mit

$$B := \begin{pmatrix} \frac{\partial g_1}{\partial x_1}(0) & \dots & \frac{\partial g_1}{\partial x_r}(0) \\ \vdots & & \vdots \\ \frac{\partial g_n}{\partial x_1}(0) & \dots & \frac{\partial g_n}{\partial x_r}(0) \end{pmatrix}$$

die Jacobimatrix von g im Punkte Null. Ist $h := f \circ g: \mathbb{R}^r \rightarrow \mathbb{R}^m$ und $h = (h_1, \dots, h_m)$, so gilt für die Jacobimatrix von h im Punkt Null

$$\begin{pmatrix} \frac{\partial h_1}{\partial x_1}(0) & \dots & \frac{\partial h_1}{\partial x_r}(0) \\ \vdots & & \vdots \\ \frac{\partial h_m}{\partial x_1}(0) & \dots & \frac{\partial h_m}{\partial x_r}(0) \end{pmatrix} = A \cdot B.$$

Das folgt sofort aus den Rechenregeln für die partiellen Ableitungen. Diese Verknüpfungseigenschaft der Systeme partieller Ableitungen war im 19. Jahrhundert einer der Ausgangspunkte für die Entwicklung des Matrizenkalküls gewesen.

2.5.3. Zwei Spezialfälle der Matrizenmultiplikation sollen noch besonders erwähnt werden.

a) Drehungen des $\mathbb{R}^2$ um die Winkel α und β werden nach 2.1.1 beschrieben durch

$$A = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} \quad \text{und} \quad B = \begin{pmatrix} \cos \beta & -\sin \beta \\ \sin \beta & \cos \beta \end{pmatrix}.$$

Die Hintereinanderausführung ist eine Drehung um den Winkel $\alpha + \beta$, und sie wird beschrieben durch $B \cdot A$, das bedeutet

$$\begin{pmatrix} \cos(\alpha + \beta) & -\sin(\alpha + \beta) \\ \sin(\alpha + \beta) & \cos(\alpha + \beta) \end{pmatrix} = \begin{pmatrix} \cos \alpha \cos \beta - \sin \alpha \sin \beta & -(\sin \alpha \cos \beta + \cos \alpha \sin \beta) \\ \cos \alpha \sin \beta + \sin \alpha \cos \beta & \cos \alpha \cos \beta - \sin \alpha \sin \beta \end{pmatrix}.$$

Daß die Einträge auf beiden Seiten gleich sind, ist die Aussage der sogenannten *Additionstheoreme* für Sinus und Cosinus. Dem Leser sei empfohlen, in diesem Fall die Gleichheit $A \cdot B = B \cdot A$ nachzuweisen, denn das ist bei Matrizen ein bemerkenswertes Ereignis.

b) Die Multiplikation einer $m \times n$ -Matrix A mit einer $n \times 1$ -Matrix, d.h. einer Spalte x haben wir schon lange benutzt zur Beschreibung der linearen Abbildung

$$A: K^n \rightarrow K^m, \quad x \mapsto A \cdot x.$$

Mit Hilfe der Matrizenmultiplikation geschrieben bedeutet die in 2.5.1 durchgeführte Rechnung dann

$$A \cdot (B \cdot x) = (A \cdot B) \cdot x$$

für jede $(r \times 1)$ -Matrix x . Das ist ein Spezialfall des Assoziativgesetzes für die Matrizenmultiplikation, das in der folgenden Sammlung von Regeln enthalten ist.

2.5.4. Rechenregeln für Matrizen. Sind Matrizen $A, A' \in M(m \times n; K)$ und $B, B' \in M(n \times r; K)$, $C \in M(r \times s; K)$ und $\lambda \in K$ gegeben, so gilt:

$$1) A \cdot (B + B') = A \cdot B + A \cdot B' \text{ und } (A + A') \cdot B = A \cdot B + A' \cdot B. \\ \text{(Distributivgesetze)}$$

$$2) A \cdot (\lambda B) = (\lambda A) \cdot B = \lambda(A \cdot B).$$

$$3) (A \cdot B) \cdot C = A \cdot (B \cdot C). \quad \text{(Assoziativgesetz)}$$

$$4) {}^t(A \cdot B) = {}^t B \cdot {}^t A.$$

$$5) E_m \cdot A = A \cdot E_n = A. \quad \text{(Neutralität der Einheitsmatrix)}$$

Beweis. 1), 2) und 5) sind ganz einfach und erfordern höchstens etwas Schreibarbeit. Für den Beweis von 4) muß man sorgfältig mit den Buchstaben umgehen: Ist $A = (a_{ij})$ und $B = (b_{jk})$, so ist

$$A \cdot B = (c_{ik}) \quad \text{mit } c_{ik} = \sum_{j=1}^n a_{ij} b_{jk}, \quad \text{also } {}^t(A \cdot B) = (c'_{ki}) \quad \text{mit } c'_{ki} = c_{ik}.$$

Weiter ist

$${}^t B = (b'_{kj}) \quad \text{mit } b'_{kj} = b_{jk} \quad \text{und} \quad {}^t A = (a'_{ji}) \quad \text{mit } a'_{ji} = a_{ij},$$

also

$${}^t B \cdot {}^t A = (d_{ki}) \quad \text{mit } d_{ki} = \sum_{j=1}^n b'_{kj} a'_{ji} = \sum_{j=1}^n b_{jk} a_{ij} = \sum_{j=1}^n a_{ij} b_{jk}.$$

Also ist $c'_{ki} = d_{ki}$.

Bleibt der Beweis des Assoziativgesetzes 3). Dazu betrachten wir die beteiligten Matrizen als lineare Abbildungen:

$$K^s \xrightarrow{C} K^r \xrightarrow{B} K^n \xrightarrow{A} K^m.$$

Nach 1.1.5 gilt das Assoziativgesetz für die Hintereinanderschaltung von Abbildungen, also ist

$$(A \circ B) \circ C = A \circ (B \circ C).$$

In 2.5.1 haben wir gezeigt, daß die Hintereinanderschaltung der Abbildungen durch das Produkt der Matrizen ausgedrückt wird. Das liefert die Behauptung.

Wer bei diesem Kniff Unwohlsein empfindet, möge zur Linderung noch einmal Summen umordnen: Sei $A = (a_{ij})$, $B = (b_{jk})$ und $C = (c_{kl})$. Dann ist

$$A \cdot B = (\alpha_{ik}) \quad \text{mit } \alpha_{ik} = \sum_{j=1}^n a_{ij} b_{jk},$$

also

$$(A \cdot B) \cdot C = (d_{il}) \quad \text{mit } d_{il} = \sum_{k=1}^r \left(\sum_{j=1}^n a_{ij} b_{jk} \right) c_{kl}.$$

Weiter ist

$$B \cdot C = (\beta_{jl}) \quad \text{mit } \beta_{jl} = \sum_{k=1}^r b_{jk} c_{kl},$$

also

$$A \cdot (B \cdot C) = (d'_{il}) \quad \text{mit } d'_{il} = \sum_{j=1}^n a_{ij} \cdot \left(\sum_{k=1}^r b_{jk} c_{kl} \right).$$

Die beiden Summen für d_{il} und d'_{il} enthalten genau die gleichen Summanden, also ist $d_{il} = d'_{il}$. $\square$

Anstelle von $A \cdot B$ schreibt man für das Produkt von Matrizen meist nur AB . Nach den Regeln 2) und 3) kann man auch Klammern weglassen und einfach

$$\lambda AB \quad \text{bzw.} \quad ABC$$

schreiben.

Vorsicht! In der Rechenregel 4) steht auf der rechten Seite der Gleichung *nicht* ${}^t A \cdot {}^t B$. In dieser Reihenfolge könnte man die Matrizen im allgemeinen nicht einmal miteinander multiplizieren. Aber selbst wenn $A, B \in M(n \times n; K)$ gilt, ist im allgemeinen

$${}^t(A \cdot B) \neq {}^t A \cdot {}^t B.$$

Man kontrolliere das an Beispielen nach. Auch von der Richtigkeit des Assoziativgesetzes sollte man sich anhand von einigen Beispielen überzeugen, denn es ist gar nicht selbstverständlich (und eine gute Kontrollmöglichkeit für die Rechnung).

Im Spezialfall quadratischer Matrizen folgt aus diesen Regeln das

Korollar. Die Menge $M(n \times n; K)$ mit der Addition aus 1.4.1 und der Multiplikation aus 2.5.2 ist ein Ring. $\square$

2.5.5. Es ist eine naheliegende Frage, wie der Rang der Produktmatrix von den Rängen der Faktoren abhängt. Man hat folgende Abschätzungen:

Lemma. Ist $A \in M(m \times n; K)$ und $B \in M(n \times r; K)$, so gilt

$$\text{rang } A + \text{rang } B - n \leq \text{rang } (A \cdot B) \leq \min \{ \text{rang } A, \text{rang } B \} .$$

Beweis. Wir betrachten die Matrizen als lineare Abbildungen, das ergibt ein Diagramm

$$\begin{array}{ccc} K^r & \xrightarrow{A \cdot B} & K^m \\ & \searrow B & \nearrow A \\ & K^n & \end{array}$$

Weiter definieren wir $F' := A|_{\text{Im } B}$. Dann ist

$$\text{Im } F' = \text{Im } (A \cdot B) \quad \text{und} \quad \text{Ker } F' = \text{Ker } A \cap \text{Im } B .$$

Die Dimensionsformel aus 2.2.4 angewandt auf F' ergibt

$$\text{rang } (A \cdot B) = \text{rang } F' = \dim \text{Im } B - \dim \text{Ker } F' = \text{rang } B - \dim \text{Ker } F' . \quad (*)$$

Daraus folgt $\text{rang } (A \cdot B) \leq \text{rang } B$. Da $\text{Im } (A \cdot B) \subset \text{Im } A$, folgt die Abschätzung nach oben. Wegen $\text{Ker } F' \subset \text{Ker } A$ folgt aus (*) weiter

$$\text{rang } (A \cdot B) \geq \text{rang } B - \dim \text{Ker } A = \text{rang } B + \text{rang } A - n ,$$

wobei die letzte Gleichung nach der Dimensionsformel für die Abbildung A gilt. Das ergibt die Abschätzung nach unten.

Sie ist scharf, wenn $\text{Ker } F' = \text{Ker } A$, also wenn $\text{Ker } A \subset \text{Im } B$. $\square$

2.5.6. Wir betrachten den kanonischen Isomorphismus aus 2.4.2 im Fall $m = n$, also

$$\text{Hom}(K^n, K^n) \rightarrow M(n \times n; K)$$

und fragen, welche quadratischen Matrizen die Isomorphismen

$$F: K^n \rightarrow K^n$$

beschreiben. Ein Isomorphismus hat eine Umkehrung F^{-1} mit

$$F \circ F^{-1} = F^{-1} \circ F = \text{id}_{K^n} .$$

Übersetzt in die entsprechenden Matrizen ergibt das die

Definition. Eine Matrix $A \in M(n \times n; K)$ heißt *invertierbar*, wenn es ein $A' \in M(n \times n; K)$ gibt mit

$$A \cdot A' = A' \cdot A = E_n .$$

Bemerkung 1. Die Menge

$$\mathrm{GL}(n; K) = \{A \in \mathrm{M}(n \times n; K) : A \text{ invertierbar}\}$$

mit der Multiplikation von Matrizen als Verknüpfung ist eine Gruppe mit neutralem Element E_n .

Sie heißt die *allgemeine lineare Gruppe* (general linear).

Beweis. Zunächst ist zu zeigen, daß die Matrizenmultiplikation tatsächlich eine Multiplikation in $\mathrm{GL}(n; K)$ induziert, d.h. daß für $A, B \in \mathrm{GL}(n; K)$ auch $A \cdot B \in \mathrm{GL}(n; K)$ gilt. Seien A' und B' so gewählt, daß

$$AA' = A'A = E_n = BB' = B'B$$

gilt. Dann ist

$$(B'A')(AB) = E_n = (AB)(B'A')$$

nach dem Assoziativgesetz für die Matrizenmultiplikation, also ist AB invertierbar. Es bleiben die Gruppenaxiome G1 und G2 nachzuweisen (vgl. 1.2.2).

Das Assoziativgesetz gilt in $\mathrm{GL}(n; K)$, denn die Multiplikation ist sogar in $\mathrm{M}(n \times n; K)$ assoziativ.

Die n -reihige Einheitsmatrix hat die Eigenschaft eines neutralen Elementes, und zu $A \in \mathrm{GL}(n; K)$ gibt es nach Definition ein Inverses A' . $\square$

Wie wir in Abschnitt 1.2.3 gesehen haben, ist das Inverse A' eindeutig bestimmt, und wie üblich schreibt man dafür A^{-1} . Es gilt dann

$$(A^{-1})^{-1} = A \quad \text{und} \quad (AB)^{-1} = B^{-1}A^{-1}.$$

Bemerkung 2. Für eine Matrix $A \in \mathrm{M}(n \times n; K)$ sind folgende Bedingungen gleichwertig:

- i) A ist invertierbar.
- ii) $'A$ ist invertierbar.
- iii) Spaltenrang $A = n$.
- iv) Zeilenrang $A = n$.

Außerdem ist $('A)^{-1} = '(A^{-1})$.

Beweis. i) $\Rightarrow$ ii) folgt aus $'(A^{-1})'A = '(AA^{-1}) = 'E_n = E_n$, und ii) $\Rightarrow$ i) ergibt sich daraus durch Transposition. i) $\Leftrightarrow$ iii) ist eine Folgerung aus 2.2.4, und ii) $\Leftrightarrow$ iv) ergibt sich wieder durch Transposition. $\square$

Aufgaben zu 2.5

1. Gegeben seien die Matrizen

$$A := \begin{pmatrix} 1 & -1 & 2 \\ 0 & 3 & 5 \\ 1 & 8 & -7 \end{pmatrix}, \quad B := \begin{pmatrix} -1 & 0 & 1 & 0 \\ 0 & 1 & 0 & -1 \\ 1 & 0 & -1 & 0 \end{pmatrix}, \quad C := \begin{pmatrix} 1 \\ 0 \\ 8 \\ -7 \end{pmatrix},$$

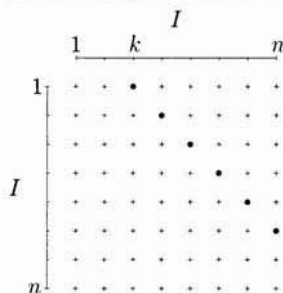
$$D := \begin{pmatrix} -1 & 2 & 0 & 8 \end{pmatrix}, \quad E := \begin{pmatrix} 1 & 4 \\ 0 & 5 \\ 6 & 8 \end{pmatrix}.$$

Berechnen Sie alle möglichen Produkte.

2. In dieser Aufgabe betrachten wir Eigenschaften „dünn besetzter“ Matrizen, in denen viele Einträge null sind.

a) Sei $n \in \mathbb{N} \setminus \{0\}$ und $I = \{1, \dots, n\}$. Wir betrachten die Menge $I \times I \subset \mathbb{N} \times \mathbb{N}$.

Finden Sie für $k \in \mathbb{N}$ Gleichungen für die „Gerade“ L in $I \times I$ durch $(1, k)$ und $(2, k+1)$ sowie für die Gerade L' durch $(k, 1)$ und $(k+1, 2)$. Finden Sie weiter Ungleichungen für den Halbraum H in $I \times I$, der oberhalb von L liegt und den Halbraum H' , der unterhalb von L' liegt.



b) Formulieren und beweisen Sie folgende Aussagen:

$$\begin{pmatrix} \diagup & 0 \\ & * \\ 0 & \diagdown \end{pmatrix} \cdot \begin{pmatrix} \diagup & 0 \\ & * \\ 0 & \diagdown \end{pmatrix} = \begin{pmatrix} \diagup & 0 \\ & * \\ 0 & \diagdown \end{pmatrix}$$

$$\begin{pmatrix} & \begin{array}{|c|} \hline * \\ \hline \end{array} \\ 0 & \end{pmatrix} \cdot \begin{pmatrix} \begin{array}{|c|} \hline * \\ \hline \end{array} & \\ 0 & \end{pmatrix} = \begin{pmatrix} & \begin{array}{|c|} \hline * \\ \hline \end{array} \\ 0 & \end{pmatrix}$$

- c) Eine Matrix $A = (a_{ij}) \in M(n \times n; K)$ heißt echte obere Dreiecksmatrix, falls $a_{ij} = 0$ für $i \geq j$. Zeigen Sie, dass eine echte obere Dreiecksmatrix A nilpotent ist, d.h. es existiert ein $m \in \mathbb{N}$ mit $A^m = 0$.

3. Sind die folgenden Teilmengen Unterringe?

- a) $\{(a_{ij}) \in M(n \times n; K) : a_{ij} = 0 \text{ für } i \geq j\} \subset M(n \times n; K)$
 b) $\{(a_{ij}) \in M(n \times n; K) : a_{ij} = 0 \text{ für } i \geq j+k \text{ oder } j \geq i+k\} \subset M(n \times n; K)$, wobei $k \in \mathbb{N}$
 c) $\left\{ \begin{pmatrix} a & b \\ 0 & c \end{pmatrix} \in M(2 \times 2; \mathbb{R}) : a \in \mathbb{Q}, b, c \in \mathbb{R} \right\} \subset M(2 \times 2; \mathbb{R})$
 d) $\left\{ \begin{pmatrix} 0 & a \\ 0 & b \end{pmatrix} \in M(2 \times 2; K) : a, b \in K \right\} \subset M(2 \times 2; K)$
 e) $\{(a_{ij}) \in M(n \times n; K) : a_{ij} = 0 \text{ für } i \neq j \text{ oder } i \geq k\} \subset M(n \times n; K)$, wobei $k \in \mathbb{N}$.

4. Sei K ein Körper und $n \in \mathbb{N} \setminus \{0\}$.

- a) Für $\lambda \in K$ gilt: $(\lambda E_n)B = B(\lambda E_n)$ für alle $B \in M(n \times n; K)$.
 b) Zeigen Sie: Ist $A \in M(n \times n; K)$ mit $AB = BA$ für alle $B \in M(n \times n; K)$, so existiert ein $\lambda \in K$ mit $A = \lambda E_n$.

5. Sei $C = \left\{ \begin{pmatrix} a & -b \\ b & a \end{pmatrix} : a, b \in \mathbb{R} \right\} \subset M(2 \times 2; \mathbb{R})$.

- a) Zeigen Sie, dass C ein Körper ist.
 b) In C ist die Gleichung $X^2 + 1 = 0$ lösbar.
 c) C ist als Körper isomorph zu $\mathbb{C}$.

6. Zeigen Sie, dass für eine Matrix $B \in M(n \times k; \mathbb{R})$ die Abbildung

$$\Phi: M(m \times n; \mathbb{R}) \rightarrow M(m \times k; \mathbb{R}), \quad A \mapsto A \cdot B,$$

stetig ist.

7. Zeigen Sie, dass die Abschätzung

$$\text{rang} A + \text{rang} B - n \leq \text{rang}(AB) \leq \min\{\text{rang} A, \text{rang} B\}$$

aus 2.5.5 für den Rang der Produktmatrix in beide Richtungen scharf ist, d. h. finden Sie Beispiele für

$$\text{rang} A + \text{rang} B - n = \text{rang}(AB) \quad \text{und} \quad \text{rang}(AB) = \min\{\text{rang} A, \text{rang} B\}.$$

8. Wir wollen eine Methode angeben, um die Inverse einer Matrix auszurechnen:

Sei dazu $A \in M(n \times n; K)$ invertierbar, d. h. $\text{rang } A = n$. Zeigen Sie: Ist

$$x^i = \begin{pmatrix} x_{1i} \\ \vdots \\ x_{ni} \end{pmatrix}$$

die Lösung des Gleichungssystems $Ax = e_i$, so ist

$$A^{-1} = \begin{pmatrix} x_{11} & \dots & x_{1n} \\ \vdots & & \vdots \\ x_{n1} & \dots & x_{nn} \end{pmatrix}.$$

Berechnen Sie auf diese Weise die inverse Matrix von

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

9. Für eine differenzierbare Abbildung

$$f: \mathbb{R}^n \rightarrow \mathbb{R}^m, \quad x \mapsto (f_1(x), \dots, f_m(x)),$$

ist die Jacobi-Matrix von f im Punkt x definiert durch

$$\text{Jac}_x f := \left(\frac{\partial f_i}{\partial x_j}(x) \right).$$

Ist $m = 1$ und f zweimal stetig partiell differenzierbar, so versteht man unter der Hesse-Matrix von f im Punkt x die Matrix

$$\text{Hess}_x f := \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(x) \right).$$

a) Berechnen Sie die Jacobi-Matrix einer linearen Abbildung $F: \mathbb{R}^n \rightarrow \mathbb{R}^m$, $x \mapsto Ax$, wobei $A \in M(m \times n; \mathbb{R})$.

b) Sei

$$P: \mathbb{R}^n \rightarrow \mathbb{R}, \quad (x_1, \dots, x_n) \mapsto \sum_{i \leq j} a_{ij} x_i x_j + \sum_{i=1}^n b_i x_i,$$

wobei $a_{ij}, b_i \in \mathbb{R}$. Berechnen Sie die Jacobi-Matrix und die Hesse-Matrix von P .

2.6 Koordinatentransformationen

Eine immer wiederkehrende Methode in der linearen Algebra ist es, durch Anwendung passender Begriffe langweilige Rechnungen mit Schlachten gegen Indizes zu vermeiden. Das ist etwas gefährlich, weil dadurch ein Trainingsrückstand im Rechnen entstehen kann. Daher vertrauen wir in diesem Abschnitt besonders darauf, daß der Leser nebenbei zur Übung genügend viele Beispiele rechnet.

2.6.1. Sei wieder V ein Vektorraum mit einer Basis $\mathcal{B} = (v_1, \dots, v_n)$. Entsprechend 2.4.2 gehört dazu genau ein Isomorphismus

$$\Phi_{\mathcal{B}}: K^n \rightarrow V \quad \text{mit} \quad \Phi_{\mathcal{B}}(e_j) = v_j \quad \text{für } j = 1, \dots, n,$$

wobei $(e_1, \dots, e_n)$ wie immer die kanonische Basis des K^n bezeichnet. Nach Definition ist

$$\Phi_{\mathcal{B}}(x_1, \dots, x_n) = x_1 v_1 + \dots + x_n v_n.$$

Man nennt $\Phi_{\mathcal{B}}$ das durch $\mathcal{B}$ bestimmte *Koordinatensystem* in V und

$$x = (x_1, \dots, x_n) = \Phi_{\mathcal{B}}^{-1}(v) \in K^n$$

die *Koordinaten* von $v = x_1 v_1 + \dots + x_n v_n$.

2.6.2. Für Anwendungen ist es wichtig, verschiedene Koordinaten ineinander umzurechnen. Dazu nimmt man an, es seien in V zwei Basen $\mathcal{A} = (v_1, \dots, v_n)$ und $\mathcal{B} = (w_1, \dots, w_n)$ gegeben. Dann hat man ein Diagramm von Isomorphismen

$$\begin{array}{ccc} K^n & & \\ & \searrow \Phi_{\mathcal{A}} & \\ T_{\mathcal{B}}^{\mathcal{A}} := \Phi_{\mathcal{B}}^{-1} \circ \Phi_{\mathcal{A}} & \downarrow & V \\ & \nearrow \Phi_{\mathcal{B}} & \\ K^n & & \end{array}$$

Man nennt die als lineare Abbildung angesehene Matrix $T_{\mathcal{B}}^{\mathcal{A}} \in \text{GL}(n; K)$ die *Transformationsmatrix* des Basiswechsels. Sie hat nach Definition die folgende Eigenschaft: ist

$$v = x_1 v_1 + \dots + x_n v_n = y_1 w_1 + \dots + y_n w_n \in V, \quad \text{so ist}$$

$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = T_{\mathcal{B}}^{\mathcal{A}} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Kennt man die Matrix $T_{\mathcal{B}}^{\mathcal{A}}$, so kann man also die „neuen“ Koordinaten y aus den „alten“ x berechnen.

Das wichtigste Beispiel ist $V = K^n$. Sind A und B die Matrizen mit den Vektoren aus $\mathcal{A}$ und $\mathcal{B}$ als Spalten, so wird obiges Diagramm zu

$$\begin{array}{ccc} K^n & \xrightarrow{A} & K^n \\ T \downarrow & & \nearrow B \\ K^n & & \end{array} \quad \text{also} \quad T = B^{-1} \cdot A.$$

Ist insbesondere $\mathcal{A}$ die kanonische Basis, so folgt $T = B^{-1}$.

Ganz analog behandelt man den Fall, daß für allgemeines V die Vektoren aus $\mathcal{B}$ durch Linearkombinationen aus $\mathcal{A}$ gegeben sind. Sei

$$w_j = s_{1j}v_1 + \dots + s_{nj}v_n$$

und $S = (s_{ij})$ die Matrix mit diesen Koeffizienten als Spalten. Es gilt

$$\Phi_B = \Phi_A \circ S,$$

denn die Werte der beiden Abbildungen auf der kanonischen Basis des K^n sind gleich:

$$(\Phi_A \circ S)(e_j) = \Phi_A(s_{1j}, \dots, s_{nj}) = s_{1j}v_1 + \dots + s_{nj}v_n = w_j = \Phi_B(e_j).$$

Als Ergebnis erhalten wir

$$S = \Phi_A^{-1} \circ \Phi_B = (T_B^A)^{-1}, \quad \text{also} \quad T_B^A = S^{-1}.$$

Das kann man in Form eines Diagramms schreiben:

$$\begin{array}{ccc} K^n & \xrightarrow{\Phi_A} & V \\ S \uparrow & & \nearrow \Phi_B \\ K^n & & \end{array}.$$

Damit ist das Problem wieder auf die Bestimmung einer inversen Matrix zurückgeführt. Ein allgemeines Verfahren dafür wird in 2.7.4 angegeben. Wenigstens ein ganz einfaches Beispiel rechnen wir direkt aus.

Beispiel. Im $\mathbb{R}^2$ betrachten wir neben der kanonischen Basis $\mathcal{K} = (e_1, e_2)$ die Basis

$$\mathcal{B} = (w_1, w_2) \quad \text{mit} \quad w_1 = {}^t(2, 1), \quad w_2 = {}^t(1, 3).$$

Die Einträge von B^{-1} betrachten wir als Unbestimmte, das ergibt die Bedingung

$$\begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} x_1 & x_3 \\ x_2 & x_4 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Sie ist gleichwertig mit den linearen Gleichungssystemen

$$\begin{array}{rcl} 2x_1 + x_2 & = & 1 \\ x_1 + 3x_2 & = & 0 \end{array} \quad \text{und} \quad \begin{array}{rcl} 2x_3 + x_4 & = & 0 \\ x_3 + 3x_4 & = & 1. \end{array}$$

Die eindeutigen Lösungen sind

$$\begin{array}{lcl} x_1 = \frac{3}{5}, & x_3 = -\frac{1}{5}, & \text{also ist } B^{-1} = \frac{1}{5} \begin{pmatrix} 3 & -1 \\ -1 & 2 \end{pmatrix}. \\ x_2 = -\frac{1}{5}, & x_4 = \frac{2}{5}, & \end{array}$$

Für $v = -e_1 - e_2$ ist $x = {}^t(-1, -1)$, also $y = B^{-1} \cdot x = {}^t(-\frac{2}{5}, -\frac{1}{5})$, d.h. $v = -\frac{2}{5}w_1 - \frac{1}{5}w_2$.

2.6.3. Der Zusammenhang zwischen Koordinatensystemen und darstellenden Matrizen ergibt sich aus der

Bemerkung. Sei $F: V \rightarrow W$ eine lineare Abbildung, und seien $\mathcal{A}$ und $\mathcal{B}$ Basen von V und W . Dann hat man ein Diagramm linearer Abbildungen

$$\begin{array}{ccc} K^n & \xrightarrow{\Phi_{\mathcal{A}}} & V \\ \downarrow M_{\mathcal{B}}^{\mathcal{A}}(F) & & \downarrow F \\ K^m & \xrightarrow{\Phi_{\mathcal{B}}} & W, \end{array}$$

und es gilt

$$\Phi_{\mathcal{B}} \circ M_{\mathcal{B}}^{\mathcal{A}}(F) = F \circ \Phi_{\mathcal{A}}, \quad \text{also} \quad M_{\mathcal{B}}^{\mathcal{A}}(F) = \Phi_{\mathcal{B}}^{-1} \circ F \circ \Phi_{\mathcal{A}}.$$

Die darstellenden Matrizen sind offenbar eine Verallgemeinerung der Transformationsmatrizen, denn für $V = W$ und $F = \text{id}_V$ gilt

$$M_{\mathcal{B}}^{\mathcal{A}}(\text{id}_V) = T_{\mathcal{B}}^{\mathcal{A}}.$$

Beweis. Es genügt zu zeigen, daß die beiden Abbildungen auf der kanonischen Basis $(e_1, \dots, e_n)$ übereinstimmen. Ist $M_{\mathcal{B}}^{\mathcal{A}}(F) = A = (a_{ij})$, so ist

$$\Phi_{\mathcal{B}}(M_{\mathcal{B}}^{\mathcal{A}}(F)(e_j)) = \Phi_{\mathcal{B}}(a_{1j}, \dots, a_{mj}) = \sum_{i=1}^m a_{ij} w_i,$$

$$F(\Phi_{\mathcal{A}}(e_j)) = F(v_j) = \sum_{i=1}^m a_{ij} w_i.$$

Die zweite Gleichung folgt aus der ersten durch Multiplikation von links mit $\Phi_{\mathcal{B}}^{-1}$. $\square$

2.6.4. Im Diagramm aus 2.6.3 hat man zwei verschiedene Wege, mit Hilfe der Abbildungen in Pfeilrichtung von K^n nach W zu gelangen, und die Aussage ist, daß auf den verschiedenen Wegen die gleiche Abbildung herauskommt. Ein Diagramm mit dieser Eigenschaft heißt *kommutativ*. Wie nützlich dieser Begriff ist, werden wir gleich sehen:

Satz. Gegeben seien Vektorräume U , V und W mit Basen $\mathcal{A}$, $\mathcal{B}$ und $\mathcal{C}$, sowie lineare Abbildungen $G: U \rightarrow V$ und $F: V \rightarrow W$. Dann gilt:

$$M_{\mathcal{C}}^{\mathcal{A}}(F \circ G) = M_{\mathcal{C}}^{\mathcal{B}}(F) \cdot M_{\mathcal{B}}^{\mathcal{A}}(G).$$

Kurz ausgedrückt: Der Hintereinanderschaltung von linearen Abbildungen entspricht das Produkt der darstellenden Matrizen.

Beweis. Für die Standardräume mit den kanonischen Basen wurde das schon in 2.5.1 ausgerechnet. Der allgemeine Fall folgt daraus durch Betrachtung des Diagramms

$$\begin{array}{ccccc}
 K^n & & \xrightarrow{\Phi_{\mathcal{A}}} & & U \\
 & \searrow B & & \nearrow G & \\
 A \cdot B & K^m & \xrightarrow{\Phi_{\mathcal{B}}} & V & \\
 & \nearrow A & & \searrow F & \\
 K^n & & \xrightarrow{\Phi_{\mathcal{C}}} & & W,
 \end{array}$$

$F \circ G$ (rechts neben dem Pfeil von V nach W)

wobei $A = M_{\mathcal{C}}^{\mathcal{B}}(F)$ und $B = M_{\mathcal{B}}^{\mathcal{A}}(G)$. Alle Teildiadgramme sind kommutativ (man mache sich klar warum), daher ist das ganze Diagramm kommutativ, und insbesondere folgt die Behauptung.

Wer lieber etwas rechnet, betrachte einen Vektor $u \in U$ und seine Koordinaten $x = \Phi_{\mathcal{A}}^{-1}(u)$. Wegen 2.5.4 ist

$$\Phi_{\mathcal{B}}^{-1}(G(u)) = B \cdot x, \quad \Phi_{\mathcal{C}}^{-1}(F(G(u))) = A \cdot (B \cdot x) = (A \cdot B) \cdot \Phi_{\mathcal{A}}^{-1}(u),$$

also $\Phi_{\mathcal{C}}^{-1} \circ (F \circ G) \circ \Phi_{\mathcal{A}} = A \cdot B$. Auch daraus folgt die Behauptung. $\square$

Für den Spezialfall von Endomorphismen (vgl. 2.4.4) ergibt sich das

Korollar. In V seien eine Basis B sowie Endomorphismen F, G gegeben. Dann ist

$$M_B(F \circ G) = M_B(F) \cdot M_B(G).$$

□

Insbesondere folgt daraus, daß

$$M_B: \text{End}(V) \rightarrow M(n \times n; K)$$

ein Ringisomorphismus ist (vgl. 1.3.2, 2.1.4, 2.4.4 und 2.5.4).

2.6.5. Nun kommen wir zum wichtigsten Ergebnis dieses Abschnittes, nämlich der Antwort auf die Frage, wie sich die darstellende Matrix bei Einführung neuer Basen ändert.

Transformationsformel. Ist $F: V \rightarrow W$ eine lineare Abbildung, sind $\mathcal{A}, \mathcal{A}'$ Basen von V und $\mathcal{B}, \mathcal{B}'$ Basen von W , so ist das Diagramm

$$\begin{array}{ccc}
 K^n & \xrightarrow{M_{\mathcal{B}}^{\mathcal{A}}(F)} & K^m \\
 \downarrow T_{\mathcal{A}'}^{\mathcal{A}} & \searrow \Phi_{\mathcal{A}} & \swarrow \Phi_{\mathcal{B}} \\
 & V \xrightarrow{F} W & \\
 \uparrow \Phi_{\mathcal{A}'} & & \searrow \Phi_{\mathcal{B}'} \\
 K^n & \xrightarrow{M_{\mathcal{B}'}^{\mathcal{A}'}(F)} & K^m \\
 & \downarrow T_{\mathcal{B}'}^{\mathcal{B}} &
 \end{array}$$

kommutativ. Insbesondere gilt für die beteiligten Matrizen

$$M_{\mathcal{B}'}^{\mathcal{A}'}(F) = T_{\mathcal{B}'}^{\mathcal{B}} \cdot M_{\mathcal{B}}^{\mathcal{A}}(F) \cdot (T_{\mathcal{A}'}^{\mathcal{A}})^{-1}.$$

Anders ausgedrückt: Sind

$$A = M_{\mathcal{B}}^{\mathcal{A}}(F) \quad \text{und} \quad B = M_{\mathcal{B}'}^{\mathcal{A}'}(F)$$

die beiden Matrizen, die F bezüglich verschiedener Paare von Basen darstellen, und sind

$$T = T_{\mathcal{A}'}^{\mathcal{A}}, \quad S = T_{\mathcal{B}'}^{\mathcal{B}}$$

die Transformationsmatrizen zwischen den verschiedenen Basen, so gilt

$$B = S \cdot A \cdot T^{-1}.$$

Zum *Beweis* genügt es zu bemerken, daß nach 2.6.2 und 2.6.3 die dreieckigen und viereckigen Teile des Diagramms kommutativ sind. Also ist das Gesamtdiagramm kommutativ. $\square$

Wer diesen Beweis als Hokusfokus ansieht, möge die Formel $B = SAT^{-1}$ direkt durch Multiplikation der drei Matrizen nachrechnen (Viel Spaß mit den Indizes!). Dabei wird sich zeigen, daß nur Rechnungen wiederholt und ineinander eingesetzt werden, die vorher schon einmal ausgeführt worden waren. Der Trick besteht also darin, sich dieses zu ersparen.

Für den Spezialfall eines Endomorphismus ergibt sich mit der Notation aus 2.4.4 das

Korollar. Sind in V zwei Basen $\mathcal{A}$ und $\mathcal{B}$ sowie ein Endomorphismus F gegeben, so ist

$$M_{\mathcal{B}}(F) = T_{\mathcal{B}}^{\mathcal{A}} \cdot M_{\mathcal{A}}(F) \cdot T_{\mathcal{A}}^{\mathcal{B}},$$

oder anders ausgedrückt

$$B = SAS^{-1},$$

wenn $A = M_{\mathcal{A}}(F)$, $B = M_{\mathcal{B}}(F)$ und $S = T_{\mathcal{B}}^{\mathcal{A}}$. $\square$

2.6.6. Nun endlich können wir noch einmal (vgl. auch 1.5.8) die Gleichheit von Zeilenrang und Spaltenrang beweisen, ohne uns die Finger mit Indizes zu beschmutzen.

Satz. Für jede Matrix $A \in M(m \times n; K)$ gilt

$$\text{Zeilenrang } A = \text{Spaltenrang } A.$$

Diese Zahl ist nach 2.2.1 gleich $\text{rang } A$.

Beweis. Wir betrachten $A: K^n \rightarrow K^m$ als lineare Abbildung und wählen in K^n und K^m Basen $\mathcal{A}$ und $\mathcal{B}$ entsprechend 2.4.3, d.h. mit

$$M_{\mathcal{B}}^{\mathcal{A}}(A) = B = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}.$$

Für B ist offensichtlich

$$\text{Zeilenrang } B = r = \text{Spaltenrang } B.$$

Um zu zeigen, daß sich diese Gleichheit auf A überträgt, wählen wir entsprechend 2.6.5 invertierbare Matrizen S und T mit

$$B = SAT.$$

Es genügt also der Beweis von folgendem

Hilfssatz. Für $A \in M(m \times n; K)$, $S \in GL(m; K)$ und $T \in GL(n; K)$ gilt

- 1) Spaltenrang $SAT =$ Spaltenrang A ,
- 2) Zeilenrang $SAT =$ Zeilenrang A .

Beweis des Hilfssatzes. Zu den gegebenen Matrizen gehört ein kommutatives Diagramm linearer Abbildungen

$$\begin{array}{ccc} K^n & \xrightarrow{A} & K^m \\ T \uparrow & & \downarrow S \\ K^n & \xrightarrow{SAT} & K^m. \end{array}$$

Da S und T Isomorphismen sind, haben die linearen Abbildungen A und SAT gleichen Rang, d.h. es gilt 1). Daraus folgt 2) durch Transposition, denn

$$\text{Zeilenrang } A = \text{Spaltenrang } {}^t A \quad \text{und} \quad {}^t(SAT) = {}^t T \cdot {}^t A \cdot {}^t S.$$

Man beachte, daß bei Multiplikation von A von rechts sogar der Spaltenraum, bei Multiplikation von links nur seine Dimension erhalten bleibt. $\square$

Wie man die Matrizen S und T aus A berechnen kann, werden wir in 2.7.6 sehen.

2.6.7. Die Transformationsformel aus 2.6.5 ergibt in die Sprache der Matrizen übersetzt die folgende

Definition. Zwei Matrizen $A, B \in M(m \times n; K)$ heißen *äquivalent*, wenn es $S \in GL(m; K)$ und $T \in GL(n; K)$ gibt mit

$$B = SAT^{-1}.$$

Im Spezialfall $m = n$ nennen wir $A, B \in M(m \times m; K)$ *ähnlich*, wenn es ein $S \in GL(m; K)$ gibt mit

$$B = SAS^{-1}.$$

Aus 2.6.5 folgt sofort die

Bemerkung. Zwei Matrizen sind genau dann äquivalent, wenn sie bezüglich verschiedener Paare von Basen die gleiche lineare Abbildung beschreiben.

Zwei quadratische Matrizen sind genau dann ähnlich, wenn sie bezüglich verschiedener Basen den gleichen Endomorphismus beschreiben. $\square$

Daß dieser Begriff der Äquivalenz nichts Neues liefert, zeigt das

Lemma. Zwei Matrizen sind genau dann äquivalent, wenn sie den gleichen Rang haben. Insbesondere ist jede Matrix vom Rang r äquivalent zu

$$\begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}.$$

Diese speziellen Matrizen repräsentieren die Äquivalenzklassen und heißen *Normalformen*.

Zum *Beweis* genügt es, die Argumente aus 2.6.6 zu wiederholen. Daß äquivalente Matrizen gleichen Rang haben, folgt aus dem dort bewiesenen Hilfssatz.

Ist A vom Rang r , so sieht man durch entsprechende Wahl von Basen in K^m und K^n , daß A äquivalent zu der obigen *Normalform* ist. Also ist A äquivalent zu jeder anderen Matrix B vom Rang r (benutze Aufgabe 4). $\square$

Viel schwieriger ist die Frage nach *Normalformen für Klassen ähnlicher Matrizen*. Das ist Thema von Kapitel 4.

Aufgaben zu 2.6

1. Gegeben sei ein endlichdimensionaler Vektorraum V mit Basen $\mathcal{A}$, $\mathcal{B}$ und $\mathcal{C}$. Beweisen Sie die „Kürzungsregel“

$$T_C^A = T_C^B \cdot T_B^A.$$

2. Im $\mathbb{R}^3$ seien die Basen

$$\mathcal{A} = ((1, -1, 2), (2, 3, 7), (2, 3, 6)) \text{ und } \mathcal{B} = ((1, 2, 2), (-1, 3, 3), (-2, 7, 6))$$

gegeben.

a) Berechnen Sie die Transformationsmatrix T_B^A .

b) Bestimmen Sie die Koordinaten des Vektors

$$v = 2 \cdot (1, -1, 2) + 9 \cdot (2, 3, 7) - 8 \cdot (2, 3, 6)$$

bezüglich der Basis $\mathcal{B}$.

3. V sei ein $\mathbb{R}$ -Vektorraum mit Basis $\mathcal{A} = (v_1, \dots, v_4)$, W sei ein $\mathbb{R}$ -Vektorraum mit Basis $\mathcal{B} = (w_1, \dots, w_5)$. $F: V \rightarrow W$ sei die lineare Abbildung, die gegeben ist durch

$$M_B^A(F) = \begin{pmatrix} 3 & 1 & -2 & 2 \\ -2 & -2 & 7 & -3 \\ 4 & 0 & 3 & 1 \\ 1 & 3 & 12 & 4 \\ 0 & 4 & -17 & 5 \end{pmatrix}.$$

Schließlich seien $\mathcal{A}' = (v'_1, \dots, v'_4)$ mit $v'_1 = v_1 + v_2$, $v'_2 = v_2 + v_3$, $v'_3 = v_3 + v_4$, $v'_4 = v_4$ und $\mathcal{B}' = (w'_1, \dots, w'_5)$ mit $w'_1 = w_1$, $w'_2 = w_1 + w_2$, $w'_3 = -w_1 + w_3$, $w'_4 = w_1 + w_4$, $w'_5 = w_1 + w_5$.

- Zeigen Sie, dass $\mathcal{A}'$ eine Basis von V und $\mathcal{B}'$ eine Basis von W ist.
- Berechnen Sie $M_{\mathcal{B}}^{\mathcal{A}'}(F)$, $M_{\mathcal{B}'}^{\mathcal{A}}(F)$ und $M_{\mathcal{B}'}^{\mathcal{A}'}(F)$.
- Bestimmen Sie $F^{-1}(\text{span}(w_1, w_2, w_3))$.

4. Zeigen Sie, dass durch

$$A \sim B \Leftrightarrow A \text{ und } B \text{ sind äquivalent}$$

(vgl. 2.6.7) tatsächlich eine Äquivalenzrelation auf der Menge $M(m \times n; K)$ gegeben ist und durch

$$A \sim B \Leftrightarrow A \text{ und } B \text{ sind ähnlich}$$

(vgl. 2.6.7) eine Äquivalenzrelation auf $M(m \times m; K)$ erklärt ist.

5. Zeigen Sie, dass für $A \in M(n \times n; \mathbb{R})$ gilt: $\text{rang } A = \text{rang } (A \cdot {}^t A)$. Gilt dies auch, falls $A \in M(n \times n; \mathbb{C})$?

2.7 Elementarmatrizen und Matrizenumformungen

Die Bedeutung von Matrizenumformungen sieht man schon am Eliminationsverfahren von GAUSS. Es gestattet die Lösung durch wiederholte Anwendung elementarer Schritte. Auf diese Weise kann man in der linearen Algebra viele Probleme lösen. In diesem Abschnitt geben wir weitere Beispiele dafür: die Inversion einer quadratischen Matrix und die explizite Bestimmung der Transformationsmatrizen, die eine gegebene Matrix auf Normalform bringen. Zur theoretischen Begründung ist es dabei sehr hilfreich, Matrizenumformungen zu interpretieren als Multiplikation mit speziellen invertierbaren Matrizen, die ihrerseits die gesamte allgemeine lineare Gruppe erzeugen. Diese sogenannten „Elementarmatrizen“ kann man als die Heizeilmännchen der linearen Algebra bezeichnen: sie erledigen geräuschlos und zuverlässig die kleinen Schmutzarbeiten.

2.7.1. Ist m eine beliebige natürliche Zahl, $1 \leq i, j \leq m$ mit $i \neq j$ und $\lambda \in K^*$, so nennt man die quadratischen Matrizen

$$\begin{array}{c}
 \begin{array}{cc}
 i\text{-te} & j\text{-te} \\
 \text{Spalte} & \text{Spalte} \\
 \downarrow & \downarrow
 \end{array} \\
 S_i(\lambda) := \left(\begin{array}{cccc|ccc}
 1 & & & & & & & \\
 & \ddots & & & & & & \\
 & & 1 & & & & & \\
 - & - & \lambda & - & - & 0 & - & - \\
 & & & 1 & & & & \\
 & & & & \ddots & & & \\
 & & & & & 1 & & \\
 - & - & 0 & - & - & 1 & - & - \\
 & & & & & & 1 & \\
 & & & & & & & \ddots \\
 & & & & & & & 1
 \end{array} \right) \begin{array}{l} \\ \\ \\ \leftarrow i\text{-te Zeile} \\ \\ \\ \leftarrow j\text{-te Zeile} \\ \\ \end{array}
 \end{array}$$

$$\begin{array}{c}
 Q_i^j := \left(\begin{array}{cccc|ccc}
 1 & & & & & & & \\
 & \ddots & & & & & & \\
 & & 1 & & & & & \\
 - & - & 1 & - & - & 1 & - & - \\
 & & & 1 & & & & \\
 & & & & \ddots & & & \\
 & & & & & 1 & & \\
 - & - & 0 & - & - & 1 & - & - \\
 & & & & & & 1 & \\
 & & & & & & & \ddots \\
 & & & & & & & 1
 \end{array} \right) \begin{array}{l} \\ \\ \\ \leftarrow i\text{-te Zeile} \\ \\ \\ \leftarrow j\text{-te Zeile} \\ \\ \end{array}
 \end{array}$$

$$Q_i^j(\lambda) := \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ - & - & - & 1 & - & - & - & \lambda & - & - & - \\ & & & 1 & & & & & & & \\ & & & & \ddots & & & & & & \\ & & & & & 1 & & & & & \\ - & - & - & 0 & - & - & - & 1 & - & - & - \\ & & & & & & & 1 & & & \\ & & & & & & & & \ddots & & \\ & & & & & & & & & 1 & \end{pmatrix} \begin{array}{l} \leftarrow i\text{-te Zeile} \\ \\ \leftarrow j\text{-te Zeile} \end{array}$$

$$P_i^j := \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ - & - & - & 0 & - & - & - & 1 & - & - & - \\ & & & 1 & & & & & & & \\ & & & & \ddots & & & & & & \\ & & & & & 1 & & & & & \\ - & - & - & 1 & - & - & - & 0 & - & - & - \\ & & & & & & & 1 & & & \\ & & & & & & & & \ddots & & \\ & & & & & & & & & 1 & \end{pmatrix} \begin{array}{l} \leftarrow i\text{-te Zeile} \\ \\ \leftarrow j\text{-te Zeile} \end{array}$$

aus $M(m \times m; K)$ *Elementarmatrizen*. Außer den eingetragenen oder durch Punkte angedeuteten Komponenten sind dabei alle Komponenten gleich Null.

Sind E_i^j die in 1.5.1 definierten Matrizen und ist E die m -reihige Einheitsmatrix, so ist

$$Q_i^j = E + E_i^j, \quad Q_i^j(\lambda) = E + \lambda E_i^j \quad \text{und} \quad S_i(\lambda) = E + (\lambda - 1)E_i^i.$$

Weiter ist selbstverständlich

$$Q_i^j = Q_i^j(1) \quad \text{und} \quad P_i^j = P_i^j.$$

Grundlegend ist der Zusammenhang mit den elementaren Umformungen von Matrizen. Ist $A \in M(m \times n; K)$ und $\lambda \in K^*$ gegeben, so hatten wir in 1.5.7 umgeformte Matrizen betrachtet, die wie folgt aus A entstanden waren:

A_I durch Multiplikation der i -ten Zeile mit λ ,

A_{II} durch Addition der j -ten Zeile zur i -ten Zeile,

A_{III} durch Addition der λ -fachen j -ten Zeile zur i -ten Zeile,

A_{IV} durch Vertauschen der i -ten und der j -ten Zeile.

Wie man sofort sieht, gilt

$$\begin{aligned} A_I &= S_i(\lambda) \cdot A, & A_{II} &= Q_i^j \cdot A, \\ A_{III} &= Q_i^j(\lambda) \cdot A, & A_{IV} &= P_i^j \cdot A. \end{aligned}$$

Man sollte sich davon auch anhand von Beispielen überzeugen, um mehr Zutrauen zu den eigenartigen Elementarmatrizen zu gewinnen.

Ganz Entsprechendes gilt, wenn man anstatt Zeilen immer Spalten umformt. Wir wollen es sicherheitshalber notieren. Ist $A \in M(m \times n; K)$ und $\lambda \in K^*$, so betrachten wir wieder die wie folgt aus A entstandenen Matrizen:

A^I durch Multiplikation der i -ten Spalte mit λ ,

A^{II} durch Addition der i -ten Spalte zur j -ten Spalte,

A^{III} durch Addition der λ -fachen i -ten Spalte zur j -ten Spalte,

A^{IV} durch Vertauschen der j -ten und der i -ten Spalte.

Verwenden wir die entsprechenden n -reihigen Elementarmatrizen, so gilt

$$\begin{aligned} A^I &= A \cdot S_i(\lambda), & A^{II} &= A \cdot Q_i^j, \\ A^{III} &= A \cdot Q_i^j(\lambda), & A^{IV} &= A \cdot P_i^j. \end{aligned}$$

Kurz ausgedrückt: Multiplikation von *links* mit Elementarmatrizen bewirkt *Zeilen*umformungen, und Multiplikation von *rechts* mit Elementarmatrizen bewirkt *Spalten*umformungen. Man beachte dabei die Vertauschung von i und j beim Übergang von links nach rechts.

Bemerkung. Die Elementarmatrizen $Q_i^j(\lambda)$ und P_i^j sind Produkte von Elementarmatrizen vom Typ $S_i(\lambda)$ und Q_i^j . Genauer gilt

$$Q_i^j(\lambda) = S_j\left(\frac{1}{\lambda}\right) \cdot Q_i^j \cdot S_j(\lambda), \quad P_i^j = Q_i^j \cdot Q_i^j(-1) \cdot Q_j^i \cdot S_j(-1).$$

Dies entspricht dem in 1.5.7 bemerkten Sachverhalt, daß sich Matrizenumformungen vom Typ III und IV aus Umformungen vom Typ I und II kombinieren lassen.

Der *Beweis* der Bemerkung ist ganz einfach, wenn man die Multiplikation als Zeilen- oder Spaltenumformung interpretiert. Wir wollen dies dem Leser überlassen. □

2.7.2. Lemma. Die Elementarmatrizen sind invertierbar und ihre Inversen sind wieder Elementarmatrizen. Genauer gilt:

$$\begin{aligned} (S_i(\lambda))^{-1} &= S_i\left(\frac{1}{\lambda}\right), & (Q_i^j)^{-1} &= Q_i^j(-1), \\ (Q_i^j(\lambda))^{-1} &= Q_i^j(-\lambda), & (P_i^j)^{-1} &= P_i^j. \end{aligned}$$

Zum Beweis genügt es, die rechten Seiten der Gleichungen mit den linken zu multiplizieren und festzustellen, daß die Einheitsmatrix herauskommt. $\square$

2.7.3. Satz. Jede invertierbare Matrix $A \in M(n \times n; K)$ ist (endliches) Produkt von Elementarmatrizen.

Man sagt dafür auch, daß die Gruppe $GL(n; K)$ von den Elementarmatrizen erzeugt wird.

Beweis. Nach 2.5.6 ist der Zeilenrang von A gleich n . Wie in 0.4.7 ausgeführt ist, kann man A durch elementare Zeilenumformungen zu einer Matrix der Form

$$B = \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ & \ddots & \vdots \\ 0 & & b_{nn} \end{pmatrix}$$

mit von Null verschiedenen Diagonalelementen $b_{11}, \dots, b_{nn}$ machen. Nach 2.7.1 gibt es Elementarmatrizen $B_1, \dots, B_r$, so daß

$$B = B_r \cdot \dots \cdot B_1 \cdot A.$$

Man kann nun B durch weitere Zeilenumformungen zur Einheitsmatrix E_n machen. Dazu beseitigt man zunächst $b_{1n}, \dots, b_{n-1,n}$ mit Hilfe der letzten Zeile, dann $b_{1,n-1}, \dots, b_{n-2,n-1}$ mit Hilfe der vorletzten Zeile, usw. Schließlich normiert man die Komponenten in der Diagonalen auf 1. Es gibt also nach 2.7.1 weitere Elementarmatrizen $B_{r+1}, \dots, B_s$, so daß

$$E_n = B_s \cdot \dots \cdot B_{r+1} B = B_s \cdot \dots \cdot B_1 \cdot A.$$

Daraus folgt

$$A^{-1} = B_s \cdot \dots \cdot B_1, \quad \text{also} \quad A = B_1^{-1} \cdot \dots \cdot B_s^{-1},$$

und die Behauptung folgt aus 2.7.2. $\square$

2.7.4. Der Beweis von Satz 2.7.3 gestattet nun, ein einfaches Rechenverfahren für die Bestimmung der inversen Matrix anzugeben. Es hat die angenehme zusätzliche Eigenschaft, daß man von der gegebenen quadratischen Matrix im voraus gar nicht zu wissen braucht, ob sie invertierbar ist. Das stellt sich im Laufe der Rechnung heraus.

Sei also $A \in M(n \times n; K)$ gegeben. Man schreibt die Matrizen A und E_n nebeneinander. Alle Umformungen, die im folgenden an A vorgenommen werden, führt man parallel an E_n durch.

Zunächst bringt man A durch Zeilenumformungen auf Zeilenstufenform. Dabei stellt sich heraus, ob

$$\text{Zeilenumrang } A = n,$$

d.h. ob A invertierbar ist (vgl. 2.5.6). Ist der Zeilenumrang von A kleiner als n , so kann man aufhören; die Umformungen mit E_n waren dann umsonst. Ist der Zeilenumrang von A gleich n , so führt man weitere Zeilenumformungen durch, bis aus A die Matrix E_n geworden ist. Schematisch sieht das so aus (die Umformungen sind als Multiplikation mit Elementarmatrizen beschrieben):

A	E_n
$B_1 \cdot A$	$B_1 \cdot E_n$
$\vdots$	$\vdots$
$B_s \cdot \dots \cdot B_1 \cdot A$	$B_s \cdot \dots \cdot B_1 \cdot E_n$

Ist nun links aus A die Einheitsmatrix E_n entstanden, so hat sich rechts aus E_n die inverse Matrix A^{-1} aufgebaut, denn aus

$$B_s \cdot \dots \cdot B_1 \cdot A = E_n$$

folgt

$$B_s \cdot \dots \cdot B_1 \cdot E_n = B_s \cdot \dots \cdot B_1 = A^{-1}.$$

Diese erste schöne Anwendung wird den Leser hoffentlich schon vom Wert der Elementarmatrizen überzeugen.

Vorsicht! Anstelle von Zeilenumformungen kann man auch *ausschließlich* Spaltenumformungen anwenden. Aber bei *abwechselnder* Anwendung von Zeilen- und Spaltenumformungen funktioniert das Verfahren im allgemeinen nicht.

2.7.5. Beispiele. a)

$$\begin{array}{l}
 A = \begin{array}{c} \begin{array}{|ccc|ccc|} \hline 0 & 1 & -4 & 1 & 0 & 0 \\ 1 & 2 & -1 & 0 & 1 & 0 \\ 1 & 1 & 2 & 0 & 0 & 1 \\ \hline \end{array} \\ P_2^1 \begin{array}{|ccc|ccc|} \hline 1 & 2 & -1 & 0 & 1 & 0 \\ 0 & 1 & -4 & 1 & 0 & 0 \\ 1 & 1 & 2 & 0 & 0 & 1 \\ \hline \end{array} \\ Q_3^1(-1) \begin{array}{|ccc|ccc|} \hline 1 & 2 & -1 & 0 & 1 & 0 \\ 0 & 1 & -4 & 1 & 0 & 0 \\ 0 & -1 & 3 & 0 & -1 & 1 \\ \hline \end{array} \\ Q_3^2 \begin{array}{|ccc|ccc|} \hline 1 & 2 & -1 & 0 & 1 & 0 \\ 0 & 1 & -4 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 & -1 & 1 \\ \hline \end{array} \\ S_3(-1) \begin{array}{|ccc|ccc|} \hline 1 & 2 & -1 & 0 & 1 & 0 \\ 0 & 1 & -4 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 & 1 & -1 \\ \hline \end{array} \\ Q_1^2(-2) \begin{array}{|ccc|ccc|} \hline 1 & 0 & 7 & -2 & 1 & 0 \\ 0 & 1 & -4 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 & 1 & -1 \\ \hline \end{array} \\ Q_1^3(-7) \begin{array}{|ccc|ccc|} \hline 1 & 0 & 0 & 5 & -6 & 7 \\ 0 & 1 & -4 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 & 1 & -1 \\ \hline \end{array} \\ Q_2^3(4) \begin{array}{|ccc|ccc|} \hline 1 & 0 & 0 & 5 & -6 & 7 \\ 0 & 1 & 0 & -3 & 4 & -4 \\ 0 & 0 & 1 & -1 & 1 & -1 \\ \hline \end{array} \end{array} = A^{-1}.
 \end{array}$$

Man berechne zur Kontrolle $A \cdot A^{-1}$!

b)

$$A = \begin{array}{c} \begin{array}{|cc|cc|} \hline 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 \\ \hline \end{array} \\ Q_3^1(-1) \\ \begin{array}{|cc|cc|} \hline 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & -1 & 0 & 1 \\ \hline \end{array} \\ Q_3^2 \\ \begin{array}{|cc|cc|} \hline 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 & 1 & 1 \\ \hline \end{array} \end{array}$$

A ist nicht invertierbar, denn Zeilenrang $A = 2$.

2.7.6. Ist $A \in M(m \times n; K)$ und

$$A: K^n \rightarrow K^m, \quad x \mapsto Ax,$$

die zugehörige lineare Abbildung, so gibt es nach 2.6.5 Transformationsmatrizen $S \in GL(m; K)$ und $T \in GL(n; K)$ mit

$$SAT^{-1} = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix},$$

wobei $r = \text{rang } A$. Wir leiten nun ein *Rechenverfahren für die Bestimmung von S und T* ab. Dazu betrachten wir folgendes Schema:

E_m	A	
$B_1 \cdot E_m$	$B_1 \cdot A$	
$\vdots$	$\vdots$	
$B_k \cdot \dots \cdot B_1 \cdot E_m$	$B_k \cdot \dots \cdot B_1 \cdot A$	E_n
	$B_k \cdot \dots \cdot B_1 \cdot A \cdot C_1$	$E_n \cdot C_1$
	$\vdots$	$\vdots$
	$B_k \cdot \dots \cdot B_1 \cdot A \cdot C_1 \cdot \dots \cdot C_l$	$E_n \cdot C_1 \cdot \dots \cdot C_l$

Zunächst wird A durch Zeilenumformungen auf Zeilenstufenform gebracht. Die Zeilenumformungen entsprechen der Multiplikation von links mit m -reihigen Elementarmatrizen $B_1, \dots, B_k$. Diese Umformungen führt man parallel an

E_m durch. Da die Matrix

$$B_k \cdot \dots \cdot B_1 \cdot A$$

Zeilenstufenform hat, kann man sie durch Spaltenumformungen auf die Form

$$\begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}$$

mit $r = \text{rang } A$ bringen. Dies entspricht Multiplikation von rechts mit n -reihigen Elementarmatrizen $C_1, \dots, C_l$. Diese Spaltenumformungen führt man parallel an E_n durch. Wegen

$$B_k \cdot \dots \cdot B_1 \cdot A \cdot C_1 \cdot \dots \cdot C_l = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}$$

sind durch

$$S = B_k \cdot \dots \cdot B_1 = B_k \cdot \dots \cdot B_1 E_m \quad \text{und} \quad T^{-1} = C_1 \cdot \dots \cdot C_l = E_n C_1 \cdot \dots \cdot C_l$$

Transformationsmatrizen der gewünschten Art gefunden.

Beispiel. Sei $K = \mathbb{R}$ und $A = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 2 & 1 \end{pmatrix}$.

$$S = \begin{array}{c|c|c} \begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array} & \begin{array}{ccc} 1 & 2 & 0 \\ 2 & 2 & 1 \end{array} & = A \\ \hline \begin{array}{cc} 1 & 0 \\ -2 & 1 \end{array} & \begin{array}{ccc} 1 & 2 & 0 \\ 0 & -2 & 1 \end{array} & \begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array} \\ \hline & \begin{array}{ccc} 1 & 0 & 2 \\ 0 & 1 & -2 \end{array} & \begin{array}{ccc} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{array} \\ \hline & \begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & -2 \end{array} & \begin{array}{ccc} 1 & 0 & -2 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{array} \\ \hline & \begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \end{array} & \begin{array}{ccc} 1 & 0 & -2 \\ 0 & 0 & 1 \\ 0 & 1 & 2 \end{array} \\ \hline SAT^{-1} = & & = T^{-1}. \end{array}$$

Ist

$$D = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix},$$

so erhält man auch Basen $\mathcal{A}$ und $\mathcal{B}$ von K^n und K^m , bezüglich derer A durch D beschrieben wird. Dazu betrachten wir das Diagramm

$$\begin{array}{ccc} K^n & \xrightarrow{D} & K^m \\ \uparrow T & & \uparrow S \\ K^n & \xrightarrow{A} & K^m, \end{array}$$

das wegen $D = SAT^{-1}$ kommutativ ist. $\mathcal{A}$ und $\mathcal{B}$ sind die Bilder der kanonischen Basen $\mathcal{K}$ und $\mathcal{K}'$ von K^n und K^m unter den Isomorphismen T^{-1} und S^{-1} . Also erhält man $\mathcal{A}$ und $\mathcal{B}$ als Spaltenvektoren von T^{-1} und S^{-1} . Dazu muß man S noch invertieren.

In unserem Beispiel ist

$$S^{-1} = \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix},$$

also sind

$$((1, 0, 0), (0, 0, 1), (-2, 1, 2)) \quad \text{und} \quad ((1, 2), (0, 1))$$

Basen der gesuchten Art. Zur Kontrolle prüft man nach:

$$A \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \quad A \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad \text{und} \quad A \cdot \begin{pmatrix} -2 \\ 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

2.7.7. Natürlich kann man auch das Gaußsche Eliminationsverfahren mit Hilfe von Elementarmatrizen beschreiben. Sei das System

$$A \cdot x = b$$

gegeben. Die elementaren Zeilenumformungen von A und (A, b) werden bewirkt durch Multiplikation von links mit Elementarmatrizen aus $GL(m; K)$. Ihr Produkt ergibt eine Matrix

$$S \in GL(m; K) \quad \text{mit} \quad (\tilde{A}, \tilde{b}) = S \cdot (A, b) = (SA, Sb),$$

wobei $(\tilde{A}, \tilde{b})$ die auf Zeilenstufenform gebrachte erweiterte Koeffizientenmatrix ist. Man beachte, daß S allein durch A bestimmt ist. Die Berechnung von S kann

man wieder schematisch durchführen, indem man die Zeilenumformungen von A parallel an E_m durchführt.

E_m	A
$B_1 \cdot E_m$	$B_1 \cdot A$
$\vdots$	$\vdots$
$B_s \cdot \dots \cdot B_1 \cdot E_m$	$B_s \cdot \dots \cdot B_1 \cdot A$

Ist nun aus A die Matrix $\tilde{A} = B_s \cdot \dots \cdot B_1 \cdot A$ in Zeilenstufenform entstanden, so hat sich links die Matrix

$$S = B_s \cdot \dots \cdot B_1$$

aufgebaut. Damit kann man sofort entscheiden, ob für ein $b \in K^m$ Lösungen von $Ax = b$ existieren. Man berechnet

$$S \cdot b = \tilde{b} = \begin{pmatrix} \tilde{b}_1 \\ \vdots \\ \tilde{b}_r \\ \tilde{b}_{r+1} \\ \vdots \\ \tilde{b}_m \end{pmatrix}$$

und sieht nach, ob $\tilde{b}_{r+1} = \dots = \tilde{b}_m = 0$, wobei $r = \text{rang } A = \text{rang } \tilde{A}$ (vgl. Aufgabe 5).

Aufgaben zu 2.7

1. Stellen Sie die folgende Matrix A als Produkt von Elementarmatrizen dar:

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

2. Sind die folgenden Matrizen invertierbar? Wenn ja, dann geben die inverse Matrix an.

$$\begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \in M(4 \times 4; \mathbb{R}), \quad \begin{pmatrix} 6 & 3 & 4 & 5 \\ 1 & 2 & 2 & 1 \\ 2 & 4 & 3 & 2 \\ 3 & 3 & 4 & 2 \end{pmatrix} \in M(4 \times 4; \mathbb{R}),$$

$$\begin{pmatrix} 1 & 2 & 0 \\ 1 & 1 & 1 \\ 2 & 0 & 1 \end{pmatrix} \in M(3 \times 3; \mathbb{R}), \quad \begin{pmatrix} 1 & 2 & 0 \\ 1 & 1 & 1 \\ 2 & 0 & 1 \end{pmatrix} \in M(3 \times 3; \mathbb{Z}/3\mathbb{Z}).$$

3. Zeigen Sie:

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M(2 \times 2; K) \text{ ist invertierbar} \Leftrightarrow ad - bc \neq 0.$$

Berechnen Sie in diesem Fall die Inverse von A .

4. Modifizieren Sie das Rechenverfahren aus 2.7.6 so, dass man statt S die inverse Matrix S^{-1} erhält (benutzen Sie dabei die Inversen der Elementarmatrizen aus 2.7.2).

5. Finden Sie für die Gleichungssysteme $Ax = b$ aus 0.3.5 sowie aus Aufgabe 2 in 0.4 jeweils eine Matrix S , so dass $\tilde{A} = SA$ in Zeilenstufenform ist, und berechnen Sie $\tilde{b} = Sb$.

6. Beweisen Sie:

a) Für $A \in M(n \times n; K)$ und $m \in \mathbb{N}$ gilt:

$$E_n - A^m = (E_n - A) \left(\sum_{i=0}^{m-1} A^i \right) = \left(\sum_{i=0}^{m-1} A^i \right) (E_n - A).$$

(Dabei sei $A^0 := E_n$.)

b) Ist $A \in M(n \times n; K)$ eine Matrix, für die ein $m \in \mathbb{N}$ existiert mit $A^m = 0$, so ist $E_n - A$ invertierbar. Wie sieht die inverse Matrix aus?