# Lineare Algebra I

Wintersemester 2014/15

W. Ebeling

©Wolfgang Ebeling
Institut für Algebraische Geometrie
Leibniz Universität Hannover
Postfach 6009
30060 Hannover
E-mail: ebeling@math.uni-hannover.de

# Literatur

[1] H. Anton: Lineare Algebra. 3., durchgesehener Nachdruck, Spektrum Akademischer Verlag, Heidelberg Berlin 2004, ISBN 3-8274-0324-3.

[2] A. Beutelspacher: Lineare Algebra. 7., aktualisierte Aufl., Vieweg Verlag, Wiesbaden, 2010, ISBN 978-3-528-66508-1.

[3] G. Fischer: Lineare Algebra. 18., aktualisierte Aufl., Springer Spektrum, Wiesbaden, 2014, ISBN 978-3-658-03944-8.

[4] D. Wille: Repetitorium der Linearen Algebra – Teil 1. Binomi Verlag, Springe, 1997, ISBN 3-923923-40-6.

# 1 Lineare Gleichungssysteme und der $\mathbb{R}^n$

Ein Hauptgegenstand dieser Vorlesung werden *lineare Gleichungssysteme* sein. So bezeichnet man ein Gleichungssystem der folgenden Art:

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

Dabei sind die $a_{ij}$ und die $b_i$ reelle Zahlen und gesucht sind die Unbekannten $x_1, \ldots, x_n$.

Wir geben ein Beispiel an, das zeigt, wie man auf ein solches Gleichungssystem stößt.

**Beispiel 1.1** In einer Fabrik werden $n$ verschiedene Produkte hergestellt. Die Herstellung einer Mengeneinheit des $i$-ten Produkts kostet $a_{1i}$ Euro. Eine Mengeneinheit des $i$-ten Produkts wird zum Preis von $a_{2i}$ Euro verkauft. Vom $i$-ten Produkt werden $x_i$ Mengeneinheiten hergestellt. Die Frage, wieviel Mengeneinheiten von jedem Produkt bei einem Einsatz von $b_1$ Euro hergestellt werden müssen, um einen Verkaufserlös von $b_2$ Euro zu erzielen, führt auf das Gleichungssystem

$$\begin{array}{rcl} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n & = & b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n & = & b_2 \end{array}$$

Im allgemeinen hat man mehrere Bedingungen, die auch durch Ungleichungen gegeben werden können, und es geht darum, einen "Gewinn" zu optimieren. Mit solchen Aufgaben beschäftigt man sich in der *linearen Optimierung*.

Wir beginnen nun mit der Darstellung der Theorie solcher linearer Gleichungssysteme.

Wir gehen aus von der Menge

$$\mathbb{R} := \text{Menge der reellen Zahlen.}$$

Die reellen Zahlen kennen Sie aus der Schule, sie werden in Analysis I eingeführt werden.

Wir stellen uns die reellen Zahlen als *Zahlengerade* vor. Jedem Punkt der Geraden entspricht eine reelle Zahl. Jedem Punkt der Ebene entspricht ein Paar $(x, y)$, jedem Punkt des Raumes ein Tripel $(x, y, z)$ von reellen Zahlen. Allgemein nennt man

$$(x_1, x_2, \ldots, x_n), \text{ wobei } x_1, \ldots, x_n \text{ reelle Zahlen sind,}$$

ein $n$-*Tupel* ($n$ ist hierbei eine beliebige natürliche Zahl). Wir setzen

$$\mathbb{R}^n := \{\vec{x} = (x_1, \ldots, x_n) \,|\, x_1, \ldots, x_n \in \mathbb{R}\}.$$

Man beachte, dass bei einem $n$-Tupel die Reihenfolge wichtig ist, d.h. zwei Tupel $(x_1, \ldots, x_n)$ und $(y_1, \ldots, y_n)$ sind genau dann gleich, wenn $x_1 = y_1, \ldots, x_n = y_n$. Man nennt den $\mathbb{R}^n$ auch den *reellen Standardraum der Dimension* $n$ und ein Element dieses Raumes auch einen *Vektor*. Die Zahlen $x_1, \ldots, x_n$ heissen die *Komponenten* von $\vec{x}$.

Der $\mathbb{R}^1$ ist die Zahlengerade, $\mathbb{R}^2$ entspricht der Ebene, $\mathbb{R}^3$ dem Raum. Für größere $n$ hat man keine geometrische Vorstellung mehr.

Mit den reellen Zahlen kann man rechnen, man kann sie nach den üblichen Regeln addieren und multiplizieren. Auch mit $n$-Tupeln kann man rechnen. Für $(x_1, \ldots, x_n) \in \mathbb{R}^n$ und $(y_1, \ldots, y_n) \in \mathbb{R}^n$ definieren wir eine *Addition*

$$(x_1, \ldots, x_n) + (y_1, \ldots, y_n) := (x_1 + y_1, \ldots, x_n + y_n)$$

und für $(x_1, \ldots, x_n) \in \mathbb{R}^n$ eine *Multiplikation mit einer Zahl* $\lambda \in \mathbb{R}$

$$\lambda \cdot (x_1, \ldots, x_n) := (\lambda \cdot x_1, \ldots, \lambda \cdot x_n).$$

Wir setzen außerdem:

$$\begin{aligned} \vec{0} &:= (0, \ldots, 0) \\ -\vec{x} &:= (-x_1, \ldots, -x_n). \end{aligned}$$

Statt $\vec{x} + (-\vec{y})$ schreibt man kürzer $\vec{x} - \vec{y}$.

Man kann diese Operationen geometrisch deuten: Dazu sehen wir ein $n$-Tupel $\vec{x} = (x_1, \ldots, x_n)$ als Vektor an, d.h. als einen Pfeil mit Fußpunkt in $\vec{0} := (0, \ldots, 0)$ und Spitze in $\vec{x}$. Zwei Vektoren $\vec{x}$ und $\vec{y}$ spannen dann ein Parallelogramm auf (siehe Abbildung 1 für $n = 2$) und dem Vektor $\vec{x} + \vec{y}$ entspricht dann der Pfeil, der auf der Diagonale mit dem Fußpunkt $\vec{0}$ liegt und als Spitze den anderen Eckpunkt hat. Der Multiplikation mit der Zahl $\lambda$ entspricht die Streckung des Vektors $\vec{x}$ um den Faktor $\lambda$.

Man sieht sofort die folgenden Regeln ein:

$$\begin{aligned} \vec{x} + \vec{0} &= (x_1, \ldots, x_n) + (0, \ldots, 0) = (x_1, \ldots, x_n), \\ 0 \cdot \vec{x} &= 0 \cdot (x_1, \ldots, x_n) = (0, \ldots, 0) = \vec{0}, \\ \vec{x} + (-\vec{x}) &= \vec{0}. \end{aligned}$$

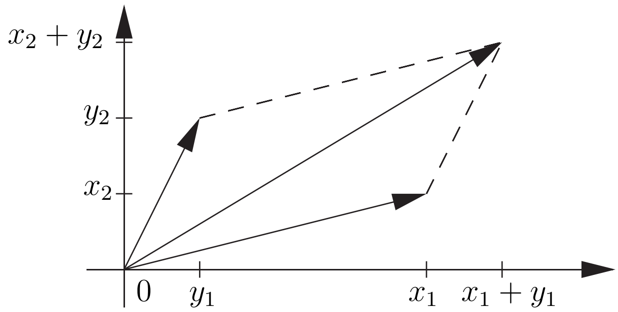



Abbildung 1: Addition zweier Vektoren ($n = 2$)

# 2 Geraden und Ebenen

Wir betrachten nun einfachste Beispiele von linearen Gleichungen und Gleichungssystemen. Zunächst betrachten wir eine lineare Gleichung der Form

$$a_1x_1 + a_2x_2 = b.$$

Dabei sind $a_1, a_2, b \in \mathbb{R}$ vorgegeben und $x_1$, $x_2$ sind die *Unbekannten*. Wir betrachten nun die Menge der *Lösungen* dieser Gleichung

$$L := \{(x_1, x_2) \in \mathbb{R}^2 \mid a_1x_1 + a_2x_2 = b\}.$$

Wir machen nun eine Fallunterscheidung.

**1. Fall:** $a_1 = a_2 = 0$. Dann ist $L = \emptyset$ für $b \neq 0$ und $L = \mathbb{R}^2$ für $b = 0$. Diesen Fall betrachten wir als ausgeartet.

**2. Fall:** $a_2 = 0$ und $a_1 \neq 0$. Dann kann man die Gleichung umformen zu

$$x_1 = \frac{b}{a_1}.$$

Die Lösungsmenge ist also eine zur $x_2$-Achse parallele Gerade.

**3. Fall:** $a_1 = 0$ und $a_2 \neq 0$. Dann kann man die Gleichung umformen zu

$$x_2 = \frac{b}{a_2}.$$

Die Lösungsmenge ist also eine zur $x_1$-Achse parallele Gerade.

**4. Fall:** $a_1 \neq 0$ und $a_2 \neq 0$. Dann kann man die Gleichung umformen zu

$$x_2 = \frac{b}{a_2} - \frac{a_1x_1}{a_2}.$$

Diese Gleichung beschreibt eine Gerade mit der Steigung $-\frac{a_1}{a_2}$ und dem Achsenabschnitt $\frac{b}{a_2}$.

Dies führt uns zu der Definition:

**Definition** Eine Teilmenge $L \subset \mathbb{R}^2$ heißt *Gerade*, genau dann, wenn es $a_1, a_2, b \in \mathbb{R}$ mit $(a_1, a_2) \neq (0,0)$ gibt, so dass

$$L = \{(x_1, x_2) \in \mathbb{R}^2 \,|\, a_1x_1 + a_2x_2 = b\}.$$

Nun kann man eine Gerade auch in *Parameterform* angeben. Für $\vec{v}, \vec{w} \in \mathbb{R}^2$ definieren wir

$$\vec{v} + \mathbb{R}\vec{w} := \{\vec{x} = \vec{v} + \lambda\vec{w} \,|\, \lambda \in \mathbb{R}\}.$$

Der Vektor $\vec{v}$ heißt *Ortsvektor* und $\vec{w}$ *Richtungsvektor* der Geraden $L = \vec{v} + \mathbb{R}\vec{w}$.

Jede Gerade lässt sich nun sowohl in Gleichungsform als auch in Parameterform angeben. Dies besagt unser erster Satz.

**Satz 2.1** *Eine Teilmenge $L \subset \mathbb{R}^2$ ist genau dann eine Gerade, wenn es $\vec{v}, \vec{w} \in \mathbb{R}^2$ mit $\vec{w} \neq \vec{0}$ gibt, so dass*

$$L = \vec{v} + \mathbb{R}\vec{w}.$$

*Beweis.* 1) Es sei

$$L := \{(x_1, x_2) \in \mathbb{R}^2 \,|\, a_1x_1 + a_2x_2 = b\}$$

mit $(a_1, a_2) \neq (0,0)$. Wir betrachten den Fall $a_2 \neq 0$, der Fall $a_1 \neq 0$ geht analog. Wir setzen

$$\vec{v} := \left(0, \frac{b}{a_2}\right), \quad \vec{w} := \left(1, -\frac{a_1}{a_2}\right), \quad L' = \vec{v} + \mathbb{R}\vec{w}.$$

Wir müssen zeigen, dass $L = L'$ gilt. Dazu müssen wir $L \subset L'$ und $L' \subset L$ zeigen.

a) $L \subset L'$: Es sei $(x_1, x_2) \in L$. Dann gilt $a_1x_1 + a_2x_2 = b$. Wir setzen $\lambda := x_1$. Dann gilt

$$\begin{aligned}(x_1, x_2) &= \left(\lambda, \frac{b}{a_2} - \frac{a_1\lambda}{a_2}\right) \\ &= \left(0, \frac{b}{a_2}\right) + \lambda\left(1, -\frac{a_1}{a_2}\right) \\ &= \vec{v} + \lambda\vec{w}.\end{aligned}$$

Also ist $(x_1, x_2) \in L'$.

b) $L' \subset L$: Es sei $(x_1, x_2) \in L'$. Dann gibt es ein $\lambda \in \mathbb{R}$ mit

$$(x_1, x_2) = \vec{v} + \lambda\vec{w} = \left(\lambda, \frac{b}{a_2} - \frac{a_1\lambda}{a_2}\right).$$

Setzt man $x_1 = \lambda$ und $x_2 = \frac{b}{a_2} - \frac{a_1\lambda}{a_2}$ in die Gleichung von $L$ ein, so erhält man

$$a_1x_1 + a_2x_2 = a_1\lambda + b - a_1\lambda = b.$$

Also ist $(x_1, x_2) \in L$.

2) Es sei nun $L = \vec{v} + \mathbb{R}\vec{w}$ gegeben. Wir müssen eine Gleichung finden. Ist $\vec{v} = (v_1, v_2)$ und $\vec{w} = (w_1, w_2)$ mit $w_1 \neq 0$, so überlegt man sich leicht, dass folgende Gleichung gilt

$$\begin{aligned} & \frac{x_2 - v_2}{x_1 - v_1} = \frac{w_2}{w_1} \\ \Leftrightarrow \quad & w_2x_1 - w_1x_2 = w_2v_1 - w_1v_2. \end{aligned}$$

Wir definieren daher

$$a_1 := w_2, \quad a_2 := -w_1, \quad b := w_2v_1 - w_1v_2,$$

und

$$L' = \{(x_1, x_2) \in \mathbb{R}^2 \mid a_1x_1 + a_2x_2 = b\}.$$

Wir müssen wieder $L = L'$ zeigen.

a) $L \subset L'$: Es sei $\vec{x} = \vec{v} + \lambda\vec{w} \in L$. Dann gilt

$$\vec{x} = (x_1, x_2) = (v_1 + \lambda w_1, v_2 + \lambda w_2).$$

Diesen Vektor müssen wir in die Gleichung $a_1x_1 + a_2x_2 = b$ einsetzen. Das ergibt

$$a_1x_1 + a_2x_2 = w_2(v_1 + \lambda w_1) - w_1(v_2 + \lambda w_2) = w_2v_1 - w_1v_2 = b.$$

Also ist $\vec{x} \in L'$.

b) $L' \subset L$: Es sei $\vec{x} = (x_1, x_2) \in L'$. Wir müssen ein $\lambda \in \mathbb{R}$ finden, so dass

$$(x_1, x_2) = \vec{v} + \lambda\vec{w} = (v_1 + \lambda w_1, v_2 + \lambda w_2).$$

Wegen $w_1 \neq 0$ kann man aus der ersten Komponente dieses Vektors ausrechnen, dass

$$\lambda = \frac{x_1 - v_1}{w_1}$$

sein muss. Also definieren wir $\lambda$ durch diese Gleichung. Wegen $\vec{x} = (x_1, x_2) \in L'$ gilt

$$w_2 x_1 - w_1 x_2 = w_2 v_1 - w_1 v_2.$$

Daraus folgt

$$x_2 = v_2 + \frac{w_2 x_1 - w_2 v_1}{w_1} = v_2 + \lambda w_2.$$

Also ist $\vec{x} \in L$. $\square$

Zwei Geraden in der Ebene sind gleich, parallel oder schneiden sich in genau einem Punkt. Wie kann man nun entscheiden, welcher Fall vorliegt, und wie kann man den möglichen Schnittpunkt finden? Wir betrachten dazu einige Beispiele.

**Beispiel 2.1** Die Geraden seien gegeben durch

$$\begin{aligned} x_1 + x_2 &= 2, \\ x_1 - x_2 &= -2. \end{aligned}$$

Addiert man diese Gleichungen zueinander, so erhält man $x_1 = 0$ und somit $x_2 = 2$. Also ist der Schnittpunkt der Punkt $(0, 2)$.

**Beispiel 2.2** Die Geraden seien gegeben durch

$$\begin{aligned} x_1 + x_2 &= 2, \\ x_1 + x_2 &= -2. \end{aligned}$$

Zieht man die zweite Gleichung von der ersten ab, so ergibt sich die Gleichung

$$0 = 4,$$

die nie erfüllt ist. Somit haben die Gleichungen keine gemeinsamen Lösungen, also die Geraden keine Schnittpunkte. Also sind die Geraden parallel.

**Beispiel 2.3** Nun nehmen wir zu den beiden Geraden aus Beispiel 2.1 noch eine dritte hinzu:

$$\begin{aligned} x_1 + x_2 &= 2, \quad (1) \\ x_1 - x_2 &= -2, \quad (2) \\ 2x_1 + 2x_2 &= 1. \quad (3) \end{aligned}$$

Dieses Gleichungssystem können wir wie folgt umformen:

$$\begin{aligned} x_1 + x_2 &= 2, \quad (1) \\ x_1 &= 0, \quad (2) + (1) \\ 0 &= -3. \quad (3) - 2 \times (1) \end{aligned}$$

Damit ergibt sich ein Widerspruch. Das Gleichungssystem besitzt keine Lösung. Die drei Geraden besitzen also keinen gemeinsamen Schnittpunkt.

Nun betrachten wir den $\mathbb{R}^3$, den dreidimensionalen Anschauungsraum. Dann definiert eine lineare Gleichung der Form

$$a_1x_1 + a_2x_2 + a_3x_3 = b$$

mit $(a_1, a_2, a_3) \neq (0, 0, 0)$ keine Gerade mehr, sondern eine Ebene. denn ist zum Beispiel $a_3 \neq 0$, so können wir die Gleichung nach $x_3$ auflösen

$$x_3 = \frac{1}{a_3}(b - a_1x_1 - a_2x_2),$$

und wir sehen, dass dies eine Ebene definiert.

**Definition** Eine Teilmenge $E \subset \mathbb{R}^3$ heißt *Ebene*, genau dann, wenn es $a_1, a_2, a_3, b \in \mathbb{R}$ mit $(a_1, a_2, a_3) \neq (0, 0, 0)$ gibt, so dass

$$E = \{(x_1, x_2, x_3) \in \mathbb{R}^3 \,|\, a_1x_1 + a_2x_2 + a_3x_3 = b\}.$$

Auch für eine Ebene gibt es eine Parameterform. Für Vektoren $\vec{u}, \vec{v}, \vec{w} \in \mathbb{R}^3$ definieren wir

$$\begin{aligned} \vec{v} + \mathbb{R}\vec{w} &:= \{\vec{x} = \vec{v} + \lambda\vec{w} \,|\, \lambda \in \mathbb{R}\}, \\ \vec{u} + \mathbb{R}\vec{v} + \mathbb{R}\vec{w} &:= \{\vec{x} = \vec{u} + \lambda_1\vec{v} + \lambda_2\vec{w} \,|\, \lambda_1, \lambda_2 \in \mathbb{R}\}. \end{aligned}$$

**Satz 2.2** *Zu einer Ebene $E \subset \mathbb{R}^3$ gibt es $\vec{u}, \vec{v}, \vec{w} \in \mathbb{R}^3$ mit $\vec{v}, \vec{w} \neq \vec{0}$, so dass*

$$E = \vec{u} + \mathbb{R}\vec{v} + \mathbb{R}\vec{w}.$$

*Beweis.* Es sei

$$E = \{(x_1, x_2, x_3) \in \mathbb{R}^3 \,|\, a_1x_1 + a_2x_2 + a_3x_3 = b\}$$

und o.B.d.A. (ohne Beschränkung der Allgemeinheit) $a_3 \neq 0$. Setzt man $x_1 = \lambda_1$ und $x_2 = \lambda_2$ in die Ebenengleichung ein, so erhält man

$$x_3 = \frac{1}{a_3}(b - a_1\lambda_1 - a_2\lambda_2).$$

Analog zum Beweis von Satz 2.1 setzen wir

$$\begin{aligned} \vec{u} &:= (0, 0, \frac{b}{a_3}), \\ \vec{v} &:= (1, 0, -\frac{a_1}{a_3}), \\ \vec{w} &:= (0, 1, -\frac{a_2}{a_3}). \end{aligned}$$

Wie im Beweis von Satz 2.1 kann man dann zeigen

$$E = \vec{u} + \mathbb{R}\vec{v} + \mathbb{R}\vec{w}.$$

$\square$

Wir betrachten nun den Schnitt von Ebenen.

**Beispiel 2.4** Wir betrachten die Gleichungen

$$\begin{array}{rcll} x_1 + x_2 + x_3 & = & -2 & (1) \\ x_1 - x_2 + 2x_3 & = & -4. & (2) \end{array}$$

Eine Umformung ergibt

$$\begin{array}{rcll} x_1 + x_2 + x_3 & = & -2 & (1) \\ -2x_2 + x_3 & = & -2. & (2) - (1) \end{array}$$

Dies bedeutet: Man kann $x_3$ frei wählen, und dann sind die anderen beiden Variablen $x_1$ und $x_2$ festgelegt. Wählen wir einen reellen Parameter $\lambda$ und setzen $x_3 = \lambda$, so folgt

$$\begin{array}{rcl} x_2 & = & \frac{1}{2}\lambda + 1, \\ x_1 & = & -\frac{3}{2}\lambda - 3. \end{array}$$

Wir setzen

$$\vec{v} = (-3, 1, 0)\,, \quad \vec{w} = \left(-\frac{3}{2}, \frac{1}{2}, 1\right).$$

Der Schnitt der beiden Ebenen ist dann die Gerade $L = \vec{v} + \mathbb{R}\vec{w}$.

**Beispiel 2.5** Wir betrachten die Gleichungen

$$\begin{array}{rcll} x_1 + x_2 + x_3 & = & -2 & (1) \\ 3x_1 + 3x_2 + 3x_3 & = & -3. & (2) \end{array}$$

Dann führt eine Umformung zu

$$\begin{array}{rcll} x_1 + x_2 + x_3 & = & -2 & (1) \\ 0 & = & 3. & (2) - 3 \times (1) \end{array}$$

Diese beiden Gleichungen sind nie gleichzeitig erfüllt. Also besteht der Schnitt der beiden Ebenen aus der leeren Menge, die beiden Ebenen sind parallel.

**Beispiel 2.6** Wir nehmen zu den beiden Gleichungen von Beispiel 2.4 noch eine Gleichung dazu:

$$\begin{array}{rcll} x_1 + x_2 + x_3 & = & -2 & (1) \\ x_1 - x_2 + 2x_3 & = & -4 & (2) \\ x_1 + 3x_2 + 2x_3 & = & 4. & (3) \end{array}$$

Diese Gleichungen formen wir um zu

$$\begin{array}{rcll} x_1 + x_2 + x_3 & = & -2 & (1) \\ -2x_2 + x_3 & = & -2 & (2*) = (2) - (1) \\ 2x_2 + x_3 & = & 6. & (3*) = (3) - (1) \end{array}$$

Wir formen noch einmal um:

$$\begin{array}{rcll} x_1 + x_2 + x_3 & = & -2 & (1) \\ -2x_2 + x_3 & = & -2 & (2*) \\ 2x_3 & = & 4. & (3*) + (2*) \end{array}$$

Die letzte Gleichung können wir nun nach $x_3$ auflösen und dann von unten nach oben einsetzen. Wir erhalten

$$\begin{array}{rcl} x_3 & = & 2 \\ x_2 & = & 2 \\ x_1 & = & -6. \end{array}$$

Der einzige Schnittpunkt der drei Ebenen ist also der Punkt $(-6, 2, 2)$. Bei drei Ebenen muss es nicht immer einen Schnittpunkt geben. Nehmen wir zum Beispiel zu den beiden Gleichungen von Beispiel 2.5 eine dritte hinzu, so gibt es keine Lösung, da zwei der Ebenen parallel sind.

# 3 Das Skalarprodukt im $\mathbb{R}^n$

Wir wollen nun noch andere Darstellungen von Ebenen im $\mathbb{R}^3$ kennenlernen. Deswegen unterbrechen wir kurz die Diskussion von linearen Gleichungssystemen.

Wir wollen nun auch Längen und Winkel definieren. Deswegen führen wir das Skalarprodukt im $\mathbb{R}^n$ ein.

**Definition** Für Vektoren $\vec{x} = (x_1, \ldots, x_n)$ und $\vec{y} = (y_1, \ldots, y_n)$ des $\mathbb{R}^n$ ist das *Skalarprodukt* $\vec{x} \cdot \vec{y}$ definiert als

$$\vec{x} \cdot \vec{y} := x_1 y_1 + x_2 y_2 + \cdots + x_n y_n.$$

Man beachte, dass $\vec{x} \cdot \vec{y}$ eine reelle Zahl ist. Bei der Multiplikation eines Vektors $\vec{x}$ mit einem Skalar $\lambda$ erhält man dagegen einen Vektor $\lambda\vec{x} \in \mathbb{R}^n$.

**Satz 3.1 (Eigenschaften des Skalarprodukts)** (i) *Für alle $\vec{x} \in \mathbb{R}^n$ gilt*

$$\vec{x} \cdot \vec{x} \geq 0$$

*und es gilt $\vec{x} \cdot \vec{x} = 0$ genau dann, wenn $\vec{x} = \vec{0}$. (Das Skalarprodukt ist* positiv definit.*)*

(ii) *Für alle $\vec{x}, \vec{y} \in \mathbb{R}^n$ gilt*

$$\vec{x} \cdot \vec{y} = \vec{y} \cdot \vec{x}.$$

*(Das Skalarprodukt ist* symmetrisch.*)*

(iii) *Für alle $\vec{x}, \vec{y}, \vec{z} \in \mathbb{R}^n$ und $\lambda \in \mathbb{R}$ gilt:*

$$\begin{aligned}(\vec{x} + \vec{z}) \cdot \vec{y} &= \vec{x} \cdot \vec{y} + \vec{z} \cdot \vec{y},\\ (\lambda\vec{x}) \cdot \vec{y} &= \lambda(\vec{x} \cdot \vec{y}),\\ \vec{x} \cdot (\vec{y} + \vec{z}) &= \vec{x} \cdot \vec{y} + \vec{x} \cdot \vec{z},\\ \vec{x} \cdot (\lambda\vec{y}) &= \lambda(\vec{x} \cdot \vec{y}).\end{aligned}$$

*(Das Skalarprodukt ist* bilinear.*)*

*Beweis.* (i) Für $\vec{x} = (x_1, \ldots, x_n)$ gilt

$$\vec{x} \cdot \vec{x} = x_1^2 + \cdots + x_n^2 \geq 0.$$

Daran sieht man auch, dass $\vec{x} \cdot \vec{x}$ genau dann gleich 0 ist, wenn $x_1 = \ldots = x_n = 0$ , also $\vec{x} = \vec{0}$ gilt.

Die Formeln von (ii) und (iii) rechnet man einfach nach. □

**Definition** Die *Länge* oder *Norm* eines Vektors $\vec{x} = (x_1, \ldots, x_n)$ ist definiert durch

$$|\vec{x}| := \sqrt{\vec{x} \cdot \vec{x}} = \sqrt{x_1^2 + \cdots + x_n^2}.$$

Nach Satz 3.1 (i) folgt

$$|\vec{x}| = 0 \Leftrightarrow \vec{x} = \vec{0}.$$

Es gelten die beiden folgenden Sätze.

**Satz 3.2 (Satz von Pythagoras)** *Für $\vec{x}, \vec{y} \in \mathbb{R}^n$ gilt*

$$|\vec{x} + \vec{y}|^2 = |\vec{x}|^2 + |\vec{y}|^2 + 2(\vec{x} \cdot \vec{y}).$$

**Satz 3.3 (Parallelogrammgleichung)** *Für $\vec{x}, \vec{y} \in \mathbb{R}^n$ gilt*

$$|\vec{x}+\vec{y}|^2 + |\vec{x}-\vec{y}|^2 = 2|\vec{x}|^2 + 2|\vec{y}|^2.$$

Die Bewiese dieser beiden Sätze überlassen wir als Übungsaufgabe.

**Satz 3.4 (Cauchy-Schwarzsche Ungleichung)** *Für $\vec{x}, \vec{y} \in \mathbb{R}^n$ gilt*

$$|\vec{x}\cdot\vec{y}| \leq |\vec{x}||\vec{y}|.$$

*Für $\vec{y} \neq \vec{0}$ gilt $|\vec{x}\cdot\vec{y}| = |\vec{x}||\vec{y}|$ genau dann, wenn es ein $\lambda \in \mathbb{R}$ gibt, so dass $\vec{x} = \lambda\vec{y}$.*

*Beweis.* Man kann diese Ungleichung durch Einsetzen von $\vec{x} = (x_1, \ldots, x_n)$ und $\vec{y} = (y_1, \ldots, y_n)$ nachrechnen, das führt aber zu einer sehr unübersichtlichen Rechnung. Deswegen benutzen wir einen kleinen Trick.

Für $\vec{y} = \vec{0}$ sind beide Seiten der Ungleichung gleich 0, die Ungleichung ist daher erfüllt. Es genügt daher, den Fall $\vec{y} \neq \vec{0}$ zu behandeln.

Für $\lambda, \mu \in \mathbb{R}$ gilt

$$\begin{aligned} 0 &\leq (\lambda\vec{x}+\mu\vec{y})\cdot(\lambda\vec{x}+\mu\vec{y}) \\ &= \lambda^2(\vec{x}\cdot\vec{x}) + 2\lambda\mu(\vec{x}\cdot\vec{y}) + \mu^2(\vec{y}\cdot\vec{y}). \end{aligned}$$

Setzen wir nun $\lambda := \vec{y}\cdot\vec{y}$ und $\mu := -(\vec{x}\cdot\vec{y})$, so folgt

$$\begin{aligned} 0 &\leq (\vec{y}\cdot\vec{y})((\vec{x}\cdot\vec{x})(\vec{y}\cdot\vec{y}) - 2(\vec{x}\cdot\vec{y})^2 + (\vec{x}\cdot\vec{y})^2) \\ &= (\vec{y}\cdot\vec{y})((\vec{x}\cdot\vec{x})(\vec{y}\cdot\vec{y}) - (\vec{x}\cdot\vec{y})^2). \end{aligned}$$

Wegen $\vec{y}\cdot\vec{y} \geq 0$ folgt daraus

$$(\vec{x}\cdot\vec{y})^2 \leq (\vec{x}\cdot\vec{x})(\vec{y}\cdot\vec{y}).$$

Nun ziehen wir auf beiden Seiten die Quadratwurzel. Dann bleibt das Ungleichheitszeichen erhalten und wir erhalten die behauptete Ungleichung.

Für den Beweis des Zusatzes bemerken wir (für $\lambda := \vec{y}\cdot\vec{y}$ und $\mu := -(\vec{x}\cdot\vec{y})$):

$$\begin{aligned} & |\vec{x}\cdot\vec{y}| = |\vec{x}||\vec{y}| \\ \Leftrightarrow\ & (\lambda\vec{x}+\mu\vec{y})\cdot(\lambda\vec{x}+\mu\vec{y}) = 0 \\ \Leftrightarrow\ & \lambda\vec{x}+\mu\vec{y} = \vec{0} \\ \Leftrightarrow\ & \vec{x} = -\frac{\mu}{\lambda}\vec{y}. \end{aligned}$$

□

Die Norm hat die folgenden Eigenschaften.

**Satz 3.5 (Eigenschaften der Norm)** (i) *Für $\vec{x} \in \mathbb{R}^n$ gilt*

$$|\vec{x}| \geq 0 \text{ und } |\vec{x}| = 0 \Leftrightarrow \vec{x} = \vec{0}.$$

(ii) *Für $\vec{x} \in \mathbb{R}^n$ und $\lambda \in \mathbb{R}$ gilt*

$$|\lambda \vec{x}| = |\lambda||\vec{x}|.$$

(iii) *(*Dreiecksungleichung*) Für $\vec{x}, \vec{y} \in \mathbb{R}^n$ gilt*

$$|\vec{x} + \vec{y}| \leq |\vec{x}| + |\vec{y}|.$$

*Beweis.* (i) folgt aus Satz 3.1 (i).
(ii) folgt aus Satz 3.1 (iii).
Zu (iii): Es gilt

$$\begin{aligned} |\vec{x} + \vec{y}|^2 &= (\vec{x} + \vec{y}) \cdot (\vec{x} + \vec{y}) \\ &= (\vec{x} \cdot \vec{x}) + 2(\vec{x} \cdot \vec{y}) + (\vec{y} \cdot \vec{y}) \\ &\leq |\vec{x}|^2 + 2|\vec{x}||\vec{y}| + |\vec{y}|^2 \quad \text{nach Satz 3.4} \\ &= (|\vec{x}| + |\vec{y}|)^2. \end{aligned}$$

Zieht man auf beiden Seiten die Wurzel, so erhält man

$$|\vec{x} + \vec{y}| \leq |\vec{x}| + |\vec{y}|.$$

□

Mit Hilfe der Norm kann man zwischen zwei Punkten $\vec{x}, \vec{y} \in \mathbb{R}^n$ einen Abstand erklären:

**Definition** Der *Abstand* zwischen zwei Punkten $\vec{x} = (x_1, \ldots, x_n)$ und $\vec{y} = (y_1, \ldots, y_n)$ ist definiert durch

$$d(\vec{x}, \vec{y}) := |\vec{y} - \vec{x}| = \sqrt{(y_1 - x_1)^2 + \cdots + (y_n - x_n)^2}.$$

Aus den Eigenschaften der Norm folgt:

**Satz 3.6 (Eigenschaften des Abstands)** (i) *Für alle $\vec{x}, \vec{y} \in \mathbb{R}^n$ gilt*

$$d(\vec{x}, \vec{y}) \geq 0 \text{ und } d(\vec{x}, \vec{y}) = 0 \Leftrightarrow \vec{x} = \vec{y}.$$

(ii) *Für alle $\vec{x}, \vec{y} \in \mathbb{R}^n$ gilt*

$$d(\vec{x}, \vec{y}) = d(\vec{y}, \vec{x}).$$

(iii) *(*Dreiecksungleichung*) Für alle $\vec{u}, \vec{v}, \vec{w} \in \mathbb{R}^n$ gilt*

$$d(\vec{u}, \vec{w}) \leq d(\vec{u}, \vec{v}) + d(\vec{v}, \vec{w}).$$

*Beweis.* (i) und (ii) folgen direkt aus Satz 3.5 (i).

(iii) folgt aus Satz 3.5 (iii) mit $\vec{x} := \vec{v} - \vec{u}$ und $\vec{y} := \vec{w} - \vec{v}$, also $\vec{x} + \vec{y} = \vec{w} - \vec{u}$. □

Nun wollen wir auch Winkel zwischen zwei Vektoren $\vec{x}, \vec{y} \in \mathbb{R}^n$ mit $\vec{x} \neq \vec{0}$ und $\vec{y} \neq \vec{0}$ erklären. Aus der Cauchy-Schwarzschen Ungleichung folgt dann nämlich

$$-1 \leq \frac{\vec{x} \cdot \vec{y}}{|\vec{x}||\vec{y}|} \leq 1.$$

Es gibt also ein $\theta \in [0, \pi]$, so dass

$$\cos\theta = \frac{\vec{x} \cdot \vec{y}}{|\vec{x}||\vec{y}|}.$$

**Definition** Der *Winkel zwischen $\vec{x}$ und $\vec{y}$*, in Zeichen $\angle(\vec{x}, \vec{y})$, ist definiert als diejenige Zahl $\theta \in [0, \pi]$, so dass

$$\cos\theta = \frac{\vec{x} \cdot \vec{y}}{|\vec{x}||\vec{y}|}.$$

**Satz 3.7 (Eigenschaften des Winkels)** (i) *Für $\vec{x}, \vec{y} \in \mathbb{R}^n$ mit $\vec{x} \neq \vec{0}$ und $\vec{y} \neq \vec{0}$ gilt*

$$\vec{x} \cdot \vec{y} = |\vec{x}||\vec{y}| \cos\angle(\vec{x}, \vec{y}).$$

(ii) *Für $\vec{x}, \vec{y} \in \mathbb{R}^n$ mit $\vec{x} \neq \vec{0}$ und $\vec{y} \neq \vec{0}$ gilt*

$$\angle(\vec{x}, \vec{y}) = \angle(\vec{y}, \vec{x}).$$

(iii) *Für $\vec{x}, \vec{y} \in \mathbb{R}^n$ mit $\vec{x} \neq \vec{0}$ und $\vec{y} \neq \vec{0}$ und $\lambda, \mu \in \mathbb{R}$ mit $\lambda\mu > 0$ gilt*

$$\angle(\lambda\vec{x}, \mu\vec{y}) = \angle(\vec{x}, \vec{y}).$$

(iv) *Für $\vec{x}, \vec{y} \in \mathbb{R}^n$ mit $\vec{x} \neq \vec{0}$ und $\vec{y} \neq \vec{0}$ gilt $\angle(\vec{x}, \vec{y}) = 0$ oder $\angle(\vec{x}, \vec{y}) = \pi$ genau dann, wenn es ein $\lambda \in \mathbb{R}$ mit $\vec{y} = \lambda\vec{x}$ gibt.*

*Beweis.* Diese Eigenschaften folgen unmittelbar aus der Definition und der Cauchy-Schwarzschen Ungleichung. □

Wir wollen nun zeigen, dass die Definition des Winkels mit der anschaulichen Definition übereinstimmt. Dazu beschränken wir uns auf den Fall $n = 2$. Es seien also $\vec{x}, \vec{y} \in \mathbb{R}^2$ von Null verschiedene Vektoren und

$$\vec{x}' := \frac{1}{|\vec{x}|}\vec{x}, \quad \vec{y}' := \frac{1}{|\vec{y}|}\vec{y}.$$

Da $|\vec{x}'| = |\vec{y}'| = 1$, gibt es $\alpha, \beta \in [0, 2\pi]$, so dass

$$\begin{aligned} \vec{x}' &= (\cos\alpha, \sin\alpha), \\ \vec{y}' &= (\cos\beta, \sin\beta). \end{aligned}$$

Aus Satz 3.7 (iii) und der Definition des Winkels folgt

$$\cos\angle(\vec{x}, \vec{y}) = \cos\angle(\vec{x}', \vec{y}') = \vec{x}' \cdot \vec{y}'.$$

Es gilt

$$\vec{x}' \cdot \vec{y}' = \cos\alpha\cos\beta + \sin\alpha\sin\beta = \cos(\alpha - \beta) = \cos(\beta - \alpha),$$

also

$$\angle(\vec{x}, \vec{y}) = \beta - \alpha.$$

**Definition** Zwei Vektoren $\vec{x}, \vec{y} \in \mathbb{R}^n$ heißen *orthogonal*, in Zeichen $\vec{x} \perp \vec{y}$, genau dann, wenn $\vec{x} \cdot \vec{y} = 0$ gilt.

Der Nullvektor ist orthogonal zu jedem Vektor. Ist $\vec{x} \neq \vec{0}$ und $\vec{y} \neq \vec{0}$, so gilt

$$\vec{x} \perp \vec{y} \Leftrightarrow \angle(\vec{x}, \vec{y}) = \frac{\pi}{2}.$$

# 4 Das Vektorprodukt

Wir wollen nun das Vektorprodukt einführen. Dies ist aber nur für Vektoren aus dem $\mathbb{R}^3$ definiert.

**Definition** Für Vektoren $\vec{x} = (x_1, x_2, x_3)$ und $\vec{y} = (y_1, y_2, y_3)$ aus $\mathbb{R}^3$ ist das *Vektorprodukt* $\vec{x} \times \vec{y}$ definiert durch

$$\vec{x} \times \vec{y} := (x_2y_3 - x_3y_2, x_3y_1 - x_1y_3, x_1y_2 - x_2y_1).$$

Man beachte, dass $\vec{x} \times \vec{y}$ wieder ein Vektor des $\mathbb{R}^3$ ist. Um sich diese Definition leichter merken zu können, geben wir noch eine Merkregel an. Wir schreiben die Vektoren als Spaltenvektoren:

$$\vec{x} \times \vec{y} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \times \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} x_2y_3 - x_3y_2 \\ x_3y_1 - x_1y_3 \\ x_1y_2 - x_2y_1 \end{pmatrix}.$$

Wir erhalten die erste Komponente (Zeile) des Vektors $\vec{x} \times \vec{y}$, indem wir die erste Zeile abdecken und die Determinante der verbleibenden $2 \times 2$-Matrix

$$\begin{vmatrix} x_2 & y_2 \\ x_3 & y_3 \end{vmatrix} = x_2y_3 - x_3y_2$$

berechnen. Entsprechend ist die zweite Komponente gleich minus der Determinante der Matrix, die man erhält, wenn man die zweite Zeile streicht. Schließlich erhält man die dritte Komponente, indem man die dritte Zeile streicht und die Determinante der verbleibenden Matrix ausrechnet.

Wir notieren nun einige Eigenschaften des Vektorprodukts.

**Satz 4.1 (Eigenschaften des Vektorprodukts)** (i) *Für* $\vec{x}, \vec{y}, \vec{z} \in \mathbb{R}^3$ *und* $\lambda \in \mathbb{R}$ *gilt*

$$\begin{aligned} (\vec{x}+\vec{y}) \times \vec{z} &= \vec{x} \times \vec{z} + \vec{y} \times \vec{z}, \\ \vec{x} \times (\vec{y}+\vec{z}) &= \vec{x} \times \vec{y} + \vec{x} \times \vec{z}, \\ (\lambda \vec{x}) \times \vec{y} &= \lambda(\vec{x} \times \vec{y}) = \vec{x} \times (\lambda \vec{y}). \end{aligned}$$

(ii) *Für* $\vec{x}, \vec{y} \in \mathbb{R}^3$ *gilt*

$$(\vec{x} \times \vec{y}) \cdot \vec{x} = (\vec{x} \times \vec{y}) \cdot \vec{y} = 0.$$

(iii) *Für* $\vec{x}, \vec{y} \in \mathbb{R}^3$ *gilt*

$$|\vec{x} \times \vec{y}|^2 = |\vec{x}|^2 |\vec{y}|^2 - (\vec{x} \cdot \vec{y})^2.$$

(iv) *Für* $\vec{x}, \vec{y} \in \mathbb{R}^3$ *gilt*

$$\vec{x} \times \vec{y} = -\vec{y} \times \vec{x} \text{ und } \vec{x} \times \vec{x} = \vec{0}.$$

*Beweis.* Alle Eigenschaften kann man einfach nachrechnen. Wir führen den Beweis von (iii) vor. Es sei $\vec{x} = (x_1, x_2, x_3)$ und $\vec{y} = (y_1, y_2, y_3)$. Dann gilt

$$\begin{aligned} |\vec{x} \times \vec{y}|^2 &= (x_2 y_3 - x_3 y_2)^2 + (x_3 y_1 - x_1 y_3)^2 + (x_1 y_2 - x_2 y_1)^2 \\ &= (x_1 y_2)^2 + (x_2 y_1)^2 + (x_3 y_1)^2 + (x_1 y_3)^2 + (x_2 y_3)^2 + (x_3 y_2)^2 \\ &\quad -2x_1 x_2 y_1 y_2 - 2x_1 x_3 y_1 y_3 - 2x_2 x_3 y_2 y_3 \\ &= (x_1^2 + x_2^2 + x_3^2)(y_1^2 + y_2^2 + y_3^2) - (x_1 y_1 + x_2 y_2 + x_3 y_3)^2 \\ &= |\vec{x}|^2 |\vec{y}|^2 - (\vec{x} \cdot \vec{y})^2. \end{aligned}$$

□

**Korollar 4.1** *Für vom Nullvektor verschiedene* $\vec{x}, \vec{y} \in \mathbb{R}^3$ *gilt*

$$|\vec{x} \times \vec{y}| = |\vec{x}||\vec{y}| \sin \angle(\vec{x}, \vec{y}).$$

*Beweis.* Es sei $\theta = \angle(\vec{x}, \vec{y})$. Nach Satz 3.7 (i) gilt

$$\vec{x} \cdot \vec{y} = |\vec{x}||\vec{y}| \cos \theta.$$

Nach Satz 4.1 (iii) folgt also

$$\begin{aligned} |\vec{x} \times \vec{y}|^2 &= |\vec{x}|^2|\vec{y}|^2 - (\vec{x} \cdot \vec{y})^2 \\ &= |\vec{x}|^2|\vec{y}|^2(1 - \cos^2\theta) \\ &= |\vec{x}|^2|\vec{y}|^2 \sin^2\theta. \end{aligned}$$

Wegen $\theta \in [0, \pi]$ ist $\sin\theta \geq 0$. Also können wir auf beiden Seiten die Quadratwurzel ziehen und die Behauptung folgt. □

Wir erhalten damit die übliche geometrische Beschreibung des Vektorprodukts. Nach Satz 4.1 (i),(ii) steht der Vektor $\vec{x} \times \vec{y}$ senkrecht auf der von den Vektoren $\vec{x}$ und $\vec{y}$ aufgespannten Ebene. Seine Länge ist nach der Formel von Korollar 4.1 gleich dem Flächeninhalt des von $\vec{x}$ und $\vec{y}$ aufgespannten Parallelogramms. Nun gibt es, wenn $\vec{x} \times \vec{y} \neq \vec{0}$, genau zwei Vektoren mit diesen Eigenschaften. Der Vektor $\vec{x} \times \vec{y}$ berechnet sich nach der *Rechte-Hand-Regel*: Zeigt der Daumen der rechten Hand in die Richtung von $\vec{x}$ und der Zeigefinger in die Richtung von $\vec{y}$, so zeigt der Mittelfinger in die Richtung von $\vec{x} \times \vec{y}$.

Wir kommen nun zurück zu Geraden und Ebenen. Später wird der Begriff der linearen Unabhängigkeit eine große Rolle spielen. Diesen Begriff führen wir nun in einem Spezialfall ein.

**Definition** Zwei Vektoren $\vec{x}$ und $\vec{y}$ des $\mathbb{R}^n$ heißen *linear unabhängig* genau dann, wenn für alle $\lambda, \mu \in \mathbb{R}$ mit der Eigenschaft, dass

$$\lambda\vec{x} + \mu\vec{y} = \vec{0}$$

gilt, notwendigerweise $\lambda = \mu = 0$ folgt. Die Vektoren $\vec{x}$ und $\vec{y}$ heißen *linear abhängig* genau dann, wenn sie nicht linear unabhängig sind, d.h. wenn es $\lambda, \mu \in \mathbb{R}$ mit $\lambda \neq 0$ oder $\mu \neq 0$ gibt, so dass

$$\lambda\vec{x} + \mu\vec{y} = \vec{0}.$$

Der folgende Satz macht diese Definition etwas verständlicher.

**Satz 4.2** *Für $\vec{x}, \vec{y} \in \mathbb{R}^n$ sind folgende Bedingungen gleichwertig:*

(i) *$\vec{x}, \vec{y}$ sind linear abhängig.*

(ii) *$\vec{x} = \vec{0}$ oder es gibt ein $\rho \in \mathbb{R}$, so dass $\vec{y} = \rho\vec{x}$.*

(iii) *$\vec{y} = \vec{0}$ oder es gibt ein $\rho \in \mathbb{R}$, so dass $\vec{x} = \rho\vec{y}$.*

*Beweis.*
(i) $\Rightarrow$ (ii): Sind $\vec{x}, \vec{y}$ linear abhängig, so gibt es $\lambda, \mu \in \mathbb{R}$ mit $\lambda \neq 0$ oder $\mu \neq 0$, so dass $\lambda\vec{x} + \mu\vec{y} = \vec{0}$. Ist $\mu = 0$, so muss $\lambda \neq 0$ sein. Aus $\lambda\vec{x} = \vec{0}$ folgt dann aber $\vec{x} = \vec{0}$. Ist $\mu \neq 0$, dann gilt

$$\vec{y} = -\frac{\lambda}{\mu}\vec{x},$$

also $\vec{y} = \rho\vec{x}$ mit $\rho := -\lambda/\mu$.
(ii) $\Rightarrow$ (i): Ist $\vec{x} = \vec{0}$, so gilt $1\vec{x} + 0\vec{y} = \vec{0}$. Ist $\vec{y} = \rho\vec{x}$, so gilt $-\rho\vec{x} + \vec{y} = \vec{0}$. In beiden Fällen sind also $\vec{x}, \vec{y}$ linear abhängig.
Der Beweis von (i) $\Leftrightarrow$ (iii) geht analog. □

**Satz 4.3** *Zwei Vektoren $\vec{x}, \vec{y} \in \mathbb{R}^3$ sind genau dann linear abhängig, wenn $\vec{x} \times \vec{y} = \vec{0}$ gilt.*

*Beweis.* $\Rightarrow$: Die Vektoren $\vec{x}, \vec{y} \in \mathbb{R}^3$ seien linear abhängig. Aus Satz 4.2 folgt, dass dann $\vec{x} = \vec{0}$ gilt oder es ein $\rho \in \mathbb{R}$ gibt, so dass $\vec{y} = \rho\vec{x}$. Ist $\vec{x} = \vec{0}$, so gilt

$$\vec{x} \times \vec{y} = \vec{0} \times \vec{y} = \vec{0}.$$

Ist $\vec{y} = \rho\vec{x}$, so gilt nach Satz 4.1

$$\vec{x} \times \vec{y} = (\rho\vec{y}) \times \vec{y} = \rho(\vec{y} \times \vec{y}) = \vec{0}.$$

$\Leftarrow$: Es sei $\vec{x} \times \vec{y} = \vec{0}$. Ist $\vec{x} = \vec{0}$ oder $\vec{y} = \vec{0}$, so sind $\vec{x}, \vec{y}$ nach Satz 4.2 linear abhängig. Es sei also $\vec{x} \neq \vec{0}$ und $\vec{y} \neq \vec{0}$. Nach Korollar 4.1 gilt

$$0 = |\vec{x} \times \vec{y}| = |\vec{x}||\vec{y}| \sin\angle(\vec{x}, \vec{y}).$$

Ist $\theta := \angle(\vec{x}, \vec{y})$, so folgt $\sin\theta = 0$, also $\theta = 0$ oder $\theta = \pi$. Das bedeutet aber, dass es ein $\rho \in \mathbb{R}$ gibt, so dass $\vec{y} = \rho\vec{x}$. Nach Satz 4.2 sind $\vec{x}, \vec{y}$ daher linear abhängig. □

Nun sind wir in der Lage, eine Charakterisierung der Ebenen im $\mathbb{R}^3$ zu geben.

**Satz 4.4** *Eine Teilmenge $E \subset \mathbb{R}^3$ ist genau dann eine Ebene, wenn es Vektoren $\vec{u}, \vec{v}, \vec{w} \in \mathbb{R}^3$ gibt, so dass $\vec{v}$ und $\vec{w}$ linear unabhängig sind und*

$$E = \vec{u} + \mathbb{R}\vec{v} + \mathbb{R}\vec{w}.$$

*Beweis.* $\Rightarrow$: Es sei

$$E = \{(x_1, x_2, x_3) \in \mathbb{R}^3 \mid a_1x_1 + a_2x_2 + a_3x_3 = b\}$$

und o.B.d.A. $a_3 \neq 0$. Es sei

$$\begin{aligned} \vec{u} &:= (0, 0, \frac{b}{a_3}), \\ \vec{v} &:= (1, 0, -\frac{a_1}{a_3}), \\ \vec{w} &:= (0, 1, -\frac{a_2}{a_3}), \end{aligned}$$

wie im Beweis von Satz 2.2. Dann sind $\vec{v}$ und $\vec{w}$ linear unabhängig. Denn aus

$$\lambda\vec{v} + \mu\vec{w} = (\lambda, \mu, -\lambda\frac{a_1}{a_3} - \mu\frac{a_2}{a_3}) = (0, 0, 0)$$

folgt $\lambda = \mu = 0$.

$\Leftarrow$: Es seien $\vec{v}$ und $\vec{w}$ linear unabhängig und $E = \vec{u} + \mathbb{R}\vec{v} + \mathbb{R}\vec{w}$. Wir betrachten den Vektor

$$\vec{a} = (a_1, a_2, a_3) := \vec{v} \times \vec{w}.$$

Nach Satz 4.3 gilt $(a_1, a_2, a_3) \neq (0, 0, 0)$. Es sei $\vec{x} = (x_1, x_2, x_3) \in E$, $\vec{x} = \vec{u} + \lambda\vec{v} + \mu\vec{w}$. Dann gilt

$$\vec{a} \cdot \vec{x} = a_1x_1 + a_2x_2 + a_3x_3 = \vec{a} \cdot \vec{u}.$$

□

**Definition** Es sei $E = \vec{u} + \mathbb{R}\vec{v} + \mathbb{R}\vec{w} \subset \mathbb{R}^3$ eine Ebene. Ein Vektor $\vec{n}$ mit $|\vec{n}| = 1$ und $\vec{n} \perp (\vec{x} - \vec{u})$ für alle $\vec{x} \in E$ heißt ein *Einheitsnormalenvektor von* $E$.

Es sei $E = \vec{u} + \mathbb{R}\vec{v} + \mathbb{R}\vec{w}$ mit linear unabhängigen Vektoren $\vec{v}$ und $\vec{w}$. Der Vektor

$$\vec{n} = \frac{\vec{v} \times \vec{w}}{|\vec{v} \times \vec{w}|}$$

ist ein Einheitsnormalenvektor von $E$. Wir haben im Beweis von Satz 4.4 die folgende Beschreibung von $E$ hergeleitet:

$$E = \{\vec{x} \in \mathbb{R}^3 \mid (\vec{x} - \vec{u}) \cdot \vec{n} = 0\}.$$

Diese Darstellung nennt man auch eine *Normalengleichung* der Ebene.

Wir betrachten nun den Abstand eines Punktes von einer Geraden oder einer Ebene.

**Definition** Ist $A \subset \mathbb{R}^3$ eine beliebige Teilmenge und $\vec{u} \in \mathbb{R}^3$, so ist für jedes $\vec{x} \in A$ der Abstand $d(\vec{u}, \vec{x}) = |\vec{u} - \vec{x}|$ erklärt. Der *Abstand zwischen $\vec{u}$ und $A$* ist dann definiert als

$$d(\vec{u}, A) := \inf\{d(\vec{u}, \vec{x}) \,|\, \vec{x} \in A\}.$$

**Lemma 4.1** *Ist $L = \vec{v} + \mathbb{R}\vec{w}$ mit $\vec{w} \neq \vec{0}$ und $\vec{u} \in \mathbb{R}^3$, so gibt es ein eindeutig bestimmtes $\vec{x}' \in L$, so dass $\vec{u} - \vec{x}'$ orthogonal zu $\vec{w}$ ist. Es gilt*

$$\vec{x}' = \vec{v} + \frac{(\vec{u} - \vec{v}) \cdot \vec{w}}{\vec{w} \cdot \vec{w}} \vec{w}.$$

**Definition** Ist $L = \vec{v} + \mathbb{R}\vec{w}$ mit $\vec{w} \neq \vec{0}$ und $\vec{u} \in \mathbb{R}^3$, so heißt das $\vec{x}'$ aus Lemma 4.1 die *orthogonale Projektion von $\vec{u} - \vec{v}$ auf $L$.*

*Beweis.* Ist $\vec{x}' = \vec{v} + \lambda\vec{w}$, so gilt

$$\begin{aligned} & (\vec{u} - \vec{x}') \cdot \vec{w} = 0 \\ \Leftrightarrow \quad & (\vec{u} - \vec{v}) \cdot \vec{w} - \lambda \vec{w} \cdot \vec{w} = 0 \\ \Leftrightarrow \quad & \lambda = \frac{(\vec{u} - \vec{v}) \cdot \vec{w}}{\vec{w} \cdot \vec{w}}. \end{aligned}$$

Also ist $\vec{x}'$ eindeutig bestimmt und es gilt

$$\vec{x}' = \vec{v} + \frac{(\vec{u} - \vec{v}) \cdot \vec{w}}{\vec{w} \cdot \vec{w}} \vec{w}.$$

□

**Lemma 4.2** *Es sei $L = \vec{v} + \mathbb{R}\vec{w}$ und $\vec{u} \in \mathbb{R}^3$. Für die orthogonale Projektion $\vec{x}'$ von $\vec{u} - \vec{v}$ auf $L$ gilt $d(\vec{u}, L) = d(\vec{u}, \vec{x}')$.*

*Beweis.* Ist $\vec{x} \in L$ beliebig, so gilt $\vec{x} - \vec{x}' = \mu\vec{w}$ für ein $\mu \in \mathbb{R}$, also $(\vec{x} - \vec{x}') \perp (\vec{u} - \vec{x}')$. Damit folgt aus dem Satz von Pythagoras (Satz 3.2)

$$|\vec{x} - \vec{u}|^2 = |\vec{x} - \vec{x}'|^2 + |\vec{x}' - \vec{u}|^2,$$

also

$$|\vec{x} - \vec{u}| \geq |\vec{x}' - \vec{u}|.$$

Also gilt $d(\vec{u}, L) = d(\vec{u}, \vec{x}')$. □

**Satz 4.5** *Ist $L = \vec{v} + \mathbb{R}\vec{w}$ und $\vec{u} \in \mathbb{R}^3$, so gilt*

$$d(\vec{u}, L) = \frac{|(\vec{u} - \vec{v}) \times \vec{w}|}{|\vec{w}|}.$$

*Beweis.* Nach Lemma 4.2 gilt

$$d(\vec{u}, L) = |\vec{x}' - \vec{u}| = |\vec{u} - \vec{x}'|$$

für die orthogonale Projektion von $\vec{x} := \vec{u} - \vec{v}$ auf $L$. Nach Lemma 4.1 gilt

$$\begin{aligned}
|\vec{u} - \vec{x}'|^2 &= |\vec{x} - \frac{\vec{x} \cdot \vec{w}}{\vec{w} \cdot \vec{w}}\vec{w}|^2 \\
&= \frac{1}{|\vec{w}|^4} ||\vec{w}|^2\vec{x} - (\vec{x} \cdot \vec{w})\vec{w}|^2 \\
&= \frac{1}{|\vec{w}|^4} (|\vec{x}|^2|\vec{w}|^4 - 2(\vec{x} \cdot \vec{w})^2|\vec{w}|^2 + (\vec{x} \cdot \vec{w})^2|\vec{w}|^2) \\
&= \frac{1}{|\vec{w}|^2} (|\vec{x}|^2|\vec{w}|^2 - (\vec{x} \cdot \vec{w})^2)
\end{aligned}$$

Nach Satz 4.1 (iii) folgt daraus

$$|\vec{u} - \vec{x}'|^2 = \frac{|\vec{x} \times \vec{w}|^2}{|\vec{w}|^2}.$$

Daraus folgt die Behauptung. (Man kann die Formel auch so einsehen: $|\vec{x} \times \vec{w}|$ misst den Flächeninhalt des von $\vec{x}$ und $\vec{w}$ aufgespannten Parallelogramms. Nach einer bekannten Formel ist dieser Flächeninhalt gleich Grundseite $|\vec{w}|$ mal Höhe $|\vec{u} - \vec{x}'|$.) □

**Satz 4.6** *Es sei $E = \vec{u} + \mathbb{R}\vec{v} + \mathbb{R}\vec{w}$ mit linear unabhängigen Vektoren $\vec{v}$ und $\vec{w}$,$\vec{n}$ ein Einheitsnormalenvektor von $E$ und $\vec{y} \in \mathbb{R}^3$. Dann gilt*

$$d(\vec{y}, E) = |(\vec{y} - \vec{u}) \cdot \vec{n}|.$$

*Beweis.* Wie im Beweis von Lemma 4.2 zeigt man, dass der senkrechte Abstand der kürzeste ist. Für die orthogonale Projektion $\vec{u}'$ des Vektors $\vec{y} - \vec{u}$ auf die Gerade $\mathbb{R}\vec{n}$ gilt

$$\vec{u}' = \frac{(\vec{y} - \vec{u}) \cdot \vec{n}}{|\vec{n}|^2}\vec{n}.$$

Wegen $|\vec{n}| = 1$ folgt daraus

$$|\vec{u}'| = |(\vec{y} - \vec{u}) \cdot \vec{n}|.$$

□

Es sei $d := d(\vec{0}, E)$. Ist $0 \leq \angle(\vec{u}, \vec{n}) \leq \pi/2$, so gilt bereits

$$d = \vec{u} \cdot \vec{n}.$$

Damit erhalten wir die *Hessesche Normalform* der Ebenengleichung

$$E = \{\vec{x} \in \mathbb{R}^3 \,|\, \vec{x} \cdot \vec{n} = d\}.$$

# 5 Der Gaußsche Algorithmus

Nach diesem Exkurs in die analytische Geometrie kehren wir wieder zu linearen Gleichungssystemen zurück.

Ein solches lineares Gleichungssystem hatte die Form

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

Dabei sind die $a_{ij}$ und die $b_i$ reelle Zahlen und gesucht ist die Menge der $(x_1, \ldots, x_n) \in \mathbb{R}^n$, die alle Gleichungen erfüllen.

Dieses System wollen wir nun zunächst übersichtlicher aufschreiben. Die Koeffizienten $a_{ij}$ schreibt man in einem rechteckigen Schema

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

Ein solches Schema nennt man eine $m \times n$-Matrix. Wir schreiben auch zur Abkürzung $A = (a_{ij})$. Die Matrix $A$ heißt die *Koeffizientenmatrix* des linearen Gleichungssystems.

Die Unbekannten $x_1, \ldots, x_n$ schreiben wir als Spaltenvektor

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

Die *Multiplikation einer Matrix $A$ mit dem Vektor $\vec{x}$* erklären wir durch

$$\begin{aligned} A\vec{x} &= \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \\ &:= \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{pmatrix}. \end{aligned}$$

Anschaulich gesprochen bedeutet dies "Zeile mal Spalte". Schreiben wir nun auch noch

$$\vec{b} := \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix},$$

so wird unser Gleichungssystem zu

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

oder kurz

$$A\vec{x} = \vec{b}.$$

Wir sehen, dass es vorteilhaft ist, Vektoren als Spaltenvektoren aufzufassen. Das werden wir in Zukunft tun. Wir werden von nun an Vektoren als Spaltenvektoren schreiben.

Wir wollen uns nun mit der Lösung eines solchen Gleichungssystems befassen. Die *Lösungsmenge* ist gleich

$$\mathcal{L}(A, b) := \{\vec{x} \in \mathbb{R}^n \mid A\vec{x} = \vec{b}\}.$$

Man kann diese Lösungsmenge mit Hilfe des Gaußschen Algorithmus ermitteln. Dieses Verfahren wollen wir nun darstellen. Anstelle der Koeffizientenmatrix $A$ betrachtet man die *erweiterte Koeffizientenmatrix*

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

Wir sehen uns noch einmal Beispiel 2.6 an. Dort hatten wir die erweiterte Koeffizientenmatrix wie folgt umgeformt:

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

Ähnlich hatten wir die erweiterte Koeffizientenmatrix auch in den anderen Beispielen umgeformt. Diese Matrix ist nun in Zeilenstufenform.

**Definition** Eine $m \times n$-Matrix $A = (a_{ij})$ heißt *in Zeilenstufenform* genau dann, wenn sie von der folgenden Form ist:

$$A = \begin{pmatrix} & \bullet & & & & & & \\ & & & \bullet & & & & \\ & & & & & \bullet & & \\ & & & & & & \ddots & \\ & & 0 & & & & & \bullet \end{pmatrix}.$$

Dabei müssen die Einträge an den mit $\bullet$ markierten Stellen von Null verschieden sein. Unterhalb der "Stufen" dürfen nur Nullen stehen.

Eine Matrix $A$ ist also genau dann in Zeilenstufenform, wenn gilt

1. Es gibt eine Zahl $r$ mit $0 \leq r \leq m$, so dass es für jedes $i$ mit $1 \leq i \leq r$ ein $j$ mit $1 \leq j \leq n$ gibt, so dass $a_{ij} \neq 0$, und so dass für alle $i, j$ mit $r < i \leq m$ und $1 \leq j \leq n$ gilt: $a_{ij} = 0$.

2. Für jedes $i$ mit $1 \leq i \leq r$ sei

$$j_i := \min\{j \mid a_{ij} \neq 0\}.$$

   Dann muss gelten

$$j_1 < j_2 < \cdots < j_r.$$

Die Einträge

$$a_{1j_1}, a_{2j_2}, \ldots, a_{rj_r}$$

stehen an den mit $\bullet$ markierten Stellen und heißen *Pivots* (auf deutsch *Angelpunkte*) von $A$.

**Beispiel 5.1** Für

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

ist $m = 4$ , $n = 5$, $r = 3$, $j_1 = 1$, $j_2 = 2$, $j_3 = 3$.

Wir betrachten zwei Spezialfälle

1. Der Fall $r = 0$: In diesem Fall ist $A$ die Nullmatrix.

2. Der Fall

$$j_1 = 1,\ j_2 = 2, \ldots j_r = r.$$

Dann hat $A$ die Form

$$A = \begin{pmatrix} a_{11} & * & \cdots & \cdots & * & \cdots & * \\ 0 & a_{22} & * & \cdots & * & \cdots & * \\ \vdots & \ddots & \ddots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{rr} & * & \cdots & * \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \end{pmatrix}.$$

Dabei steht $*$ für einen beliebigen Eintrag.

Durch eine Umordnung der Spalten von $A$ kann man stets erreichen, dass einer dieser beiden Fälle vorliegt. Einer Umordnung der Spalten entspricht eine Umnummerierung der Unbekannten des zugehörigen Gleichungssystems.

Nun betrachten wir ein lineares Gleichungssystem, bei dem die zugehörige Koeffizientenmatrix $A$ bereits in Zeilenstufenform ist. Zur Vereinfachung nehmen wir an, dass $A$ in der Form von 2. ist. Dann hat die erweiterte Koeffizientenmatrix die Gestalt

$$(A, \vec{b}) = \left(\begin{array}{ccccccc|c} a_{11} & * & \cdots & \cdots & * & \cdots & * & b_1 \\ 0 & a_{22} & * & \cdots & * & \cdots & * & b_2 \\ \vdots & \ddots & \ddots & \ddots & \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & a_{rr} & * & \cdots & * & b_r \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & b_{r+1} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & b_m \end{array}\right)$$

mit $a_{ii} \neq 0$, $i = 1, \ldots, r$.

**Satz 5.1** *Unter den obigen Voraussetzungen gilt:*

(i) *Gibt es ein $r + 1 \leq i \leq m$ mit $b_i \neq 0$, so ist $\mathcal{L}(A, \vec{b}) = \emptyset$.*

(ii) *Gilt $m = n = r$, so gibt es genau eine Lösung. Ist obendrein*

$$b_1 = \ldots = b_n = 0,$$

*so hat das Gleichungssystem $A\vec{x} = \vec{0}$ nur die* triviale Lösung $\vec{x} = \vec{0}$.

(iii) *Gilt $r < n$ und $b_{r+1} = \ldots = b_m = 0$, so gibt es unendlich viele Lösungen.*

*Beweis von Satz 5.1 (i).* Die $i$-te Gleichung lautet dann

$$0x_1 + \cdots 0x_n = b_i \neq 0.$$

Das ist ein Widerspruch. Das Gleichungssystem ist also nicht zu erfüllen. □

Um Satz 5.1 (ii), (iii) zu beweisen, geben wir ein Lösungsverfahren an. Es sei

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

Dann können wir für die Variablen $x_{r+1}, \ldots, x_n$ beliebige Werte einsetzen. Diese Variablen werden daher als *freie Variablen* bezeichnet. Die Variablen $x_1, \ldots, x_r$ sind dann durch die Werte der freien Variablen festgelegt. Man bezeichnet sie daher als *gebundene Variablen*. Man berechnet sie durch *Rückwärtssubstitution*.

Genauer lässt sich das Lösungsverfahren wie folgt beschreiben. Wir setzen $k := n - r$, wählen $\lambda_1, \ldots, \lambda_k \in \mathbb{R}$ als *Parameter* und setzen

$$x_{r+1} = \lambda_1,\ x_{r+2} = \lambda_2,\ \ldots, x_n = \lambda_k.$$

Dies setzen wir in die $r$-te Gleichung ein, die damit lautet:

$$a_{rr}x_r + a_{r,r+1}\lambda_1 + \cdots + a_{rn}\lambda_k = b_r.$$

Da $a_{rr} \neq 0$ ist, können wir diese Gleichung nach $x_r$ auflösen und erhalten

$$x_r = \frac{1}{a_{rr}}(b_r - a_{r,r+1}\lambda_1 - \cdots - a_{rn}\lambda_k).$$

Setzen wir dies in die $(r-1)$-te Gleichung ein, so können wir entsprechend $x_{r-1}$ durch $\lambda_1, \ldots, \lambda_k$ ausdrücken. Fährt man so fort, so kann man schließlich auch $x_1$ durch $\lambda_1, \ldots, \lambda_k$ ausdrücken.

**Beispiel 5.2** Das Gleichungssystem

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

hat die erweiterte Koeffizientenmatrix

$$(A, \vec{b}) = \left(\begin{array}{ccccc|c} 1 & 3 & 4 & 2 & -2 & 2 \\ 0 & 2 & -1 & 2 & 0 & 1 \\ 0 & 0 & 2 & -6 & 0 & -2 \end{array}\right).$$

Setzen wir $x_4 = \lambda_1$ und $x_5 = \lambda_2$, so erhalten wir

$$\begin{aligned} x_3 &= -1 + 3\lambda_1 \\ x_2 &= \frac{1}{2}(1 + x_3 - 2\lambda_1) = \frac{1}{2}\lambda_1 \\ x_1 &= 2 - 3x_2 - 4x_3 - 2\lambda_1 + 2\lambda_2 = 6 - \frac{31}{2}\lambda_1 + 2\lambda_2. \end{aligned}$$

*Beweis von Satz 5.1(ii) und (iii).*

(ii) Gilt $m = n = r$, so sieht die Matrix $A$ wie folgt aus:

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

Wegen $k = n - r = 0$ gibt es dann keine freien Variablen, also gibt es nur eine einzige Lösung. Ist obendrein

$$b_1 = \ldots = b_n = 0,$$

so folgt durch Rückwärtssubstitution

$$x_n = \ldots = x_1 = 0.$$

Also hat das Gleichungssystem $A\vec{x} = \vec{0}$ nur die triviale Lösung $\vec{x} = \vec{0}$.

(iii) Gilt $r < n$ und $b_{r+1} = \ldots = b_m = 0$, so gibt es mindestens eine freie Variable und die Lösungen hängen von mindestens einem Parameter ab. □

Wir versuchen nun die erweiterte Koeffizientenmatrix eines beliebigen linearen Gleichungssystems auf Zeilenstufenform zu bringen. Dazu verwenden wir die folgenden *elementaren Zeilenumformungen.*

(Z1) Vertauschung von zwei Zeilen

(Z2) Addition des $\lambda$-fachen einer Zeile ($\lambda \neq 0$) zu einer anderen Zeile.

Der Gaußsche Algorithmus basiert darauf, dass diese Umformungen nichts an der Lösungsmenge eines Gleichungssystems ändern.

**Satz 5.2** *Es sei $(A, \vec{b})$ die erweiterte Koeffizientenmatrix eines linearen Gleichungssystems und $(\widetilde{A}, \widetilde{\vec{b}})$ die aus $(A, \vec{b})$ durch endlich viele Zeilenumformungen entstandene Matrix. Dann haben die linearen Gleichungssysteme*

$$A\vec{x} = \vec{b} \text{ und } \widetilde{A}\vec{x} = \widetilde{\vec{b}}$$

*die gleichen Lösungsmengen.*

*Beweis.* Es reicht zu zeigen, dass eine einzige elementare Zeilenumformung nichts an der Lösungsmenge des Gleichungssystems ändert.

Einer Vertauschung von zwei Zeilen entspricht die Vertauschung zweier Gleichungen. Es ist klar, dass dadurch die Lösungsmenge nicht geändert wird.

Der Addition des $\lambda$-fachen der $i$-Zeile zur $\ell$-ten Zeile entspricht die Umformung der beiden Gleichungen

$$
\begin{aligned}
a_{i1}x_1 + \cdots + a_{in}x_n &= b_i & (1)\\
a_{\ell 1}x_1 + \cdots + a_{\ell n}x_n &= b_\ell & (2)
\end{aligned}
$$

zu

$$
\begin{aligned}
a_{i1}x_1 + \cdots + a_{in}x_n &= b_i & (3)\\
(a_{\ell 1} + \lambda a_{i1})x_1 + \cdots + (a_{\ell n} + \lambda a_{in})x_n &= b_\ell + \lambda b_i. & (4)
\end{aligned}
$$

Diese beiden Gleichungssysteme haben aber die gleichen Lösungsmengen. Denn erfüllt $x_1, \ldots, x_n$ die Gleichungen (1) und (2), so erfüllt es auch die Gleichungen (3) und (4). Umgekehrt sieht man durch Subtraktion des $\lambda$-fachen der Gleichung (3) von Gleichung (4), dass wenn $x_1, \ldots, x_n$ (3) und (4) erfüllt, dann auch die Gleichungen (1) und (2). □

Entscheidend für die Lösung von linearen Gleichungssystemen ist nun der folgende Satz.

**Satz 5.3** *Jede Matrix $A$ kann durch elementare Zeilenumformungen in eine Matrix $\widetilde{A}$ in Zeilenstufenform überführt werden.*

*Beweis.* Wir geben einen Algorithmus an, mit dem $A$ in Zeilenstufenform überführt werden kann, der leicht in ein Computerprogramm umgewandelt werden kann.

Es sei $A$ eine $m \times n$-Matrix. Ist $A$ die Nullmatrix, so hat $A$ schon Zeilenstufenform mit $r = 0$ und wir sind fertig.

Andernfalls gibt es mindestens einen von Null verschiedenen Eintrag. Also gibt es mindestens eine Spalte, in der nicht nur Nullen stehen. Wir nehmen die mit dem kleinsten Index. Dieser Index sei $j_1$. In dieser Spalte betrachten wir nun den ersten von Null verschiedenen Eintrag. Dieser Eintrag sei $a_{i_1 j_1}$. Ist $i_1 \neq 1$, dann vertauschen wir die erste und die $i_1$-te Zeile. Dann ist $\widetilde{a}_{1j_1}$ unser erster Pivot.

Nun machen wir durch elementare Umformungen vom Typ (Z2) alle Einträge in der $j_1$-ten Spalte unterhalb von $\widetilde{a}_{1j_1}$ zu Null. Ist $a_{i_2 j_1}$ ein solcher Eintrag, so addieren wir das $\lambda$-fache der ersten Zeile zur $i_2$-Zeile, wobei

$$\lambda := -\frac{a_{i_2 j_1}}{\widetilde{a}_{1j_1}}$$

ist. Nach diesen Umformungen erhalten wir eine Matrix $\widetilde{A}_1$ von der folgenden Gestalt

$$\widetilde{A}_1 = \left(\begin{array}{cccc|ccc} 0 & \cdots & 0 & \widetilde{a}_{1j_1} & * & \cdots & * \\ \hline 0 & \cdots & 0 & 0 & & & \\ \vdots & \ddots & \vdots & \vdots & & A_2 & \\ 0 & \cdots & 0 & 0 & & & \end{array}\right).$$

Dabei stehen wieder an den mit $*$ markierten Stellen irgendwelche Einträge. Die Matrix $A_2$ hat $m-1$ Zeilen und $n-j_1$ Spalten.

Im nächsten Schritt macht man mit $A_2$ das Gleiche wie oben. Man erhält dann $A_3$, usw. Das Verfahren bricht irgendwann ab, da die Anzahl der Zeilen und Spalten bei jedem Schritt abnimmt oder die Nullmatrix entsteht. □

**Beispiel 5.3** Wir betrachten die folgenden Umformungen

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

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

Die Matrix $\widetilde{A}$ ist in Zeilenstufenform mit $r = 3$.

Der Gaußsche Algorithmus kann wie folgt zusammengefasst werden. Er besteht aus drei Schritten.

(1) **Vorwärtselimination**
Durch elementare Zeilenumformungen wird die Matrix $A$ auf Zeilenstufenform gebracht. Der Vektor $\vec{b}$ wird dabei mit transformiert. Aus der erweiterten Koeffizientenmatrix $(A, \vec{b})$ wird dabei $(\widetilde{A}, \widetilde{\vec{b}})$. Sind die Pivots die Einträge $\widetilde{a}_{11}, \dots, \widetilde{a}_{rr}$, so ist dieser Schritt fertig, ansonsten vertauschen wir die Spalten von $\widetilde{A}$ so, dass dies gilt.

(2) **Lösbarkeitsentscheidung**
Ist $\widetilde{b}_{r+1} = \ldots = \widetilde{b}_m = 0$? Wenn nicht, ist das Gleichungssystem nicht lösbar.

(3) **Rückwärtssubstitution**
Falls $\widetilde{b}_{r+1} = \ldots = \widetilde{b}_m = 0$, setze $k := n - r$ und $x_{r+1} = \lambda_1, \dots, x_n = \lambda_k$ und löse von unten nach oben nach $x_r, x_{r-1}, \dots, x_1$ auf.

# 6 Matrizen

Wir behandeln nun allgemein Matrizen. Wir legen zunächst einige Bezeichnungen fest.

**Definition** Eine $m \times n$-*Matrix* ist ein rechteckiges Zahlenschema

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

Die Zahlen $a_{ij}$ bezeichnet man als die *Einträge* oder die *Elemente* der Matrix. Eine $m \times 1$-Matrix bezeichnet man auch als *Spaltenvektor* und eine $1 \times n$-Matrix als einen *Zeilenvektor*. Eine $n \times n$-Matrix nennt man eine *quadratische Matrix*. In diesem Fall heißen die Elemente $a_{11}, \ldots, a_{nn}$ die *Diagonalelemente* von $A$.

Wir definieren nun Rechenoperationen mit Matrizen. Es seien $A = (a_{ij})$ und $B = (b_{ij})$ zwei $m \times n$-Matrizen. Dann ist ihre *Summe* $A + B$ wie folgt definiert:

$$\begin{aligned} A + B &= \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} + \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{m1} & \cdots & b_{mn} \end{pmatrix} \\ &:= \begin{pmatrix} a_{11} + b_{11} & \cdots & a_{1n} + b_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} + b_{m1} & \cdots & a_{mn} + b_{mn} \end{pmatrix}. \end{aligned}$$

Ist $A$ eine $m \times n$-Matrix und $\lambda \in \mathbb{R}$, so ist das *Produkt* $\lambda A$ definiert durch

$$\lambda A = \lambda \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} := \begin{pmatrix} \lambda a_{11} & \cdots & \lambda a_{1n} \\ \vdots & \ddots & \vdots \\ \lambda a_{m1} & \cdots & \lambda a_{mn} \end{pmatrix}.$$

Das *Produkt* $AB$ einer $m \times r$-Matrix $A = (a_{ij})$ mit einer $r \times n$-Matrix $B = (b_{ij})$ ist definiert durch

$$\begin{aligned} AB &= \begin{pmatrix} a_{11} & \cdots & a_{1r} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mr} \end{pmatrix} \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{r1} & \cdots & b_{rn} \end{pmatrix} \\ &:= \begin{pmatrix} a_{11}b_{11} + \cdots a_{1r}b_{r1} & \cdots & a_{11}b_{1n} + \cdots + a_{1r}b_{rn} \\ \vdots & \ddots & \vdots \\ a_{m1}b_{11} + \cdots + a_{mr}b_{r1} & \cdots & a_{m1}b_{1n} + \cdots + a_{mr}b_{rn} \end{pmatrix}. \end{aligned}$$

Das bedeutet,

$$AB = C = (c_{ij})$$

und $c_{ij}$ erhält man, indem man paarweise die Einträge der $i$-ten Zeile von $A$ und der $j$-ten Spalte von $B$ multipliziert und die entstehenden Produkte addiert, also

$$c_{ij} = a_{i1}b_{1j} + \cdots + a_{ir}b_{rj}.$$

Sind $a_1, a_2, \ldots a_n$ beliebige Zahlen, so schreibt man abkürzend für die Summe dieser Zahlen

$$\sum_{k=1}^{n} a_k := a_1 + a_2 + \cdots + a_n.$$

Also können wir abkürzend schreiben

$$c_{ij} = \sum_{k=1}^{r} a_{ik}b_{kj}.$$

Die neue Matrix $C$ ist eine $m \times n$-Matrix. Man merke sich: $AB$ ist nur erklärt, wenn die Spaltenzahl von $A$ gleich der Zeilenzahl von $B$ ist. Es gilt die folgende Merkregel:

$$m \times r \text{ mal } r \times n \text{ ergibt } m \times n.$$

Die Multiplikation einer Matrix mit einem Vektor ist der Spezialfall $n = 1$.

**Beispiel 6.1**

$$\begin{pmatrix} 1 & 5 & 0 & -1 \\ 0 & 3 & 2 & 0 \\ -4 & 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} 2 & -1 \\ 0 & 3 \\ -4 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 14 \\ -8 & 11 \\ -13 & 5 \end{pmatrix}.$$

Eine besondere Rolle spielen die $n \times n$-Nullmatrix

$$0 = 0_n := \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}.$$

und die $n \times n$-Einheitsmatrix

$$E = E_n := \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}.$$

**Satz 6.1 (Rechenregeln)** *Es seien alle Matrizen so gewählt, dass die Operationen definiert sind. Dann gelten die folgenden Rechenregeln*

(a) $A + B = B + A$ *(Kommutativgesetz der Addition)*

(b) $A + (B + C) = (A + B) + C$ *(Assoziativgesetz der Addition)*

(c) $A + 0 = 0 + A = A$ *(neutrales Element der Addition)*

(d) $(AB)C = A(BC)$ *(Assoziativgesetz der Multiplikation)*

(e) $AE = EA = A$ *(neutrales Element der Multiplikation)*

(f) $A(B + C) = AB + AC$ *(linkes Distributivgesetz)*

(g) $(A + B)C = AC + BC$ *(rechtes Distributivgesetz)*

*Beweis.* Alle Aussagen betreffen die Gleichheit von Matrizen und werden bewiesen, indem man beide Seiten ausrechnet und zeigt, dass die einander entsprechenden Einträge übereinstimmen. Wir führen dies nur für die Aussage (d) vor.

Es sei $A = (a_{ij})$ eine $m \times r$-Matrix, $B = (b_{ij})$ eine $r \times s$-Matrix und $C = (c_{ij})$ eine $s \times n$-Matrix. Es gilt

$$AB = (\alpha_{il}) \text{ mit } \alpha_{il} = \sum_{k=1}^{r} a_{ik} b_{kl},$$

also

$$(AB)C = (d_{ij})$$

mit

$$d_{ij} = \sum_{l=1}^{s} \alpha_{il} c_{lj} = \sum_{l=1}^{s} \left( \sum_{k=1}^{r} a_{ik} b_{kl} \right) c_{lj} = \sum_{k=1}^{r} \sum_{l=1}^{s} a_{ik} b_{kl} c_{lj}.$$

Auf der anderen Seite gilt

$$BC = (\beta_{kj}) \text{ mit } \beta_{kj} = \sum_{l=1}^{s} b_{kl} c_{lj},$$

also

$$A(BC) = (d'_{ij})$$

mit

$$d'_{ij} = \sum_{k=1}^{r} a_{ik} \beta_{kj} = \sum_{k=1}^{r} a_{ik} \left( \sum_{l=1}^{s} b_{kl} c_{lj} \right) = \sum_{k=1}^{r} \sum_{l=1}^{s} a_{ik} b_{kl} c_{lj}.$$

$\square$

Die Matrizenmultiplikation ist nicht kommutativ, wie das folgende Beispiel zeigt.

**Beispiel 6.2** Es sei

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

Dann gilt

$$AB = \begin{pmatrix} -1 & -1 \\ 1 & 1 \end{pmatrix}, \quad BA = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix},$$

also $AB \neq BA$.

**Definition** Es sei $A$ eine quadratische Matrix. Die Matrix $A$ heißt *invertierbar* genau dann, wenn es eine Matrix $B$ mit $AB = BA = E$ gibt. In diesem Fall heißt $B$ eine *inverse Matrix* zu $A$.

**Beispiel 6.3** Es sei

$$A := \begin{pmatrix} -1 & -2 \\ 3 & 5 \end{pmatrix}, \quad B := \begin{pmatrix} 5 & 2 \\ -3 & -1 \end{pmatrix}.$$

Dann ist $A$ invertierbar und $B$ ist eine inverse Matrix zu $A$, da

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

und

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

**Beispiel 6.4** Die Matrix

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

ist nicht invertierbar. Denn ist

$$B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$$

eine beliebige $2 \times 2$-Matrix, so gilt

$$AB = \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} b_{11} + 2b_{12} & b_{12} + 2b_{22} \\ 0 & 0 \end{pmatrix} \neq E.$$

Es ist zunächst ja nicht ausgeschlossen, dass es zu einer Matrix zwei inverse Matrizen geben kann. Der folgende Satz zeigt, dass das aber nicht möglich ist.

**Satz 6.2** *Sind $B$ und $C$ inverse Matrizen zu $A$, so ist $B = C$.*

*Beweis.* Nach Voraussetzung gilt

$$AB = BA = E \text{ und } AC = CA = E.$$

Also gilt einerseits

$$(BA)C = EC = C$$

und andererseits

$$(BA)C = B(AC) = BE = B.$$

Also folgt $B = C$. □

Also ist die inverse Matrix zu einer Matrix $A$ eindeutig bestimmt und wir bezeichnen sie mit $A^{-1}$. Es gilt also

$$AA^{-1} = A^{-1}A = E.$$

**Beispiel 6.5** Die Matrix

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

ist für $ad - bc \neq 0$ invertierbar und in diesem Fall gilt

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

**Satz 6.3** (i) *Für zwei invertierbare $n \times n$-Matrizen $A$ und $B$ ist auch das Produkt $AB$ invertierbar und es gilt $(AB)^{-1} = B^{-1}A^{-1}$.*

(ii) *Mit $A$ ist auch $A^{-1}$ invertierbar und es gilt $(A^{-1})^{-1} = A$.*

*Beweis.*

(i) Es gilt

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AEA^{-1} = AA^{-1} = E.$$

Entsprechend zeigt man $(B^{-1}A^{-1})(AB) = E$. Daraus folgt, dass $AB$ invertierbar ist und $(AB)^{-1} = B^{-1}A^{-1}$ gilt.

(ii) Wegen $A^{-1}A = AA^{-1} = E$ folgt, dass $A^{-1}$ invertierbar mit $(A^{-1})^{-1} = A$ ist. □

**Definition** Jeder $m \times n$-Matrix $A$ zugeordnet ist die *transponierte Matrix* $A^T$, die sich aus $A$ durch Vertauschen von Zeilen und Spalten ergibt, d.h. $A^T$ ist die $n \times m$-Matrix, deren $i$-te Spalte für $i = 1, 2, \dots, m$ die $i$-te Zeile von $A$ ist:

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

Die transponierte Matrix $A^T$ entsteht durch eine Spiegelung der Matrix $A$ an der Hauptdiagonale.

**Beispiel 6.6**

$$A = \begin{pmatrix} 1 \\ 2 \\ 3 \\ -1 \end{pmatrix}, \quad A^T = \begin{pmatrix} 1 & 2 & 3 & -1 \end{pmatrix},$$

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}, \quad A^T = \begin{pmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{pmatrix}.$$

**Satz 6.4 (Rechenregeln für die transponierte Matrix)** *Es seien alle Matrizen so gewählt, dass die Operationen definiert sind. Dann gelten die folgenden Rechenregeln*

(a) $(A + B)^T = A^T + B^T$.

(b) $(\lambda A)^T = \lambda A^T$ *für jedes* $\lambda \in \mathbb{R}$.

(c) $(A^T)^T = A$.

(d) $(AB)^T = B^T A^T$.

(e) *Mit* $A$ *ist auch* $A^T$ *invertierbar und es gilt* $(A^T)^{-1} = (A^{-1})^T$.

*Beweis.* Die Aussagen (a)-(c) sind leicht nachzurechnen.

(d) Es sei $A = (a_{ij})$ eine $m \times r$-Matrix und $B = (b_{ij})$ eine $r \times n$-Matrix. Dann gilt

$$AB = (c_{ij}) \text{ mit } c_{ij} = \sum_{k=1}^{r} a_{ik} b_{kj},$$

also

$$(AB)^T = (c_{ji}) = (\sum_{k=1}^{r} a_{jk}b_{ki}).$$

Auf der anderen Seite gilt

$$B^T A^T = (d_{ij}) \text{ mit } d_{ij} = \sum_{k=1}^{r} b_{ki}a_{jk},$$

also $(AB)^T = B^T A^T$.

(e) Es gilt nach (d)

$$\begin{aligned} A^T(A^{-1})^T &= (A^{-1}A)^T = E^T = E, \\ (A^{-1})^T A^T &= (AA^{-1})^T = E^T = E. \end{aligned}$$

Daraus folgt die Behauptung. □

In § 5 haben wir die folgenden Zeilenoperationen betrachtet:

(Z1) Vertauschung der $i$-ten mit der $k$-ten Zeile

(Z2) Addition des $\lambda$-fachen der $k$-ten Zeile zur $i$-ten Zeile

Wir betrachten noch zusätzlich die Operation

(Z3) Multiplikation der $i$-ten Zeile mit einer reellen Zahl $\lambda \neq 0$

Diesen Operationen ordnen wir wie folgt Matrizen zu:

**Definition** Die folgenden $m \times m$-Matrizen nennen wir *Elementarmatrizen*:

$$F_{ik} := \begin{array}{cl} \begin{pmatrix} 1 & & & & & & & & & & \\ & \ddots & & & & & & & & & \\ & & 1 & & & & & & & & \\ & & & 0 & & & & 1 & & & \\ & & & & 1 & & & & & & \\ & & & & & \ddots & & & & & \\ & & & & & & 1 & & & & \\ & & & 1 & & & & 0 & & & \\ & & & & & & & & 1 & & \\ & & & & & & & & & \ddots & \\ & & & & & & & & & & 1 \end{pmatrix} & \begin{array}{l} \\ \\ \\ \leftarrow i\text{-te Zeile} \\ \\ \\ \\ \leftarrow k\text{-te Zeile} \\ \\ \\ \\ \end{array} \end{array}$$

$$F_{ik}^{\lambda} := \left(\begin{array}{ccccccccccc}
1 & & & & & & & & & & \\
& \ddots & & & & & & & & & \\
& & 1 & & & & & & & & \\
& & & 1 & & & & \lambda & & & \\
& & & & 1 & & & & & & \\
& & & & & \ddots & & & & & \\
& & & & & & 1 & & & & \\
& & & & & & & 1 & & & \\
& & & & & & & & 1 & & \\
& & & & & & & & & \ddots & \\
& & & & & & & & & & 1
\end{array}\right)
\begin{array}{l}
\\ \\ \\ \leftarrow i\text{-te Zeile} \\ \\ \\ \\ \leftarrow k\text{-te Zeile} \\ \\ \\ \\
\end{array}$$

$$F_{i}^{\lambda} := \left(\begin{array}{ccccccc}
1 & & & & & & \\
& \ddots & & & & & \\
& & 1 & & & & \\
& & & \lambda & & & \\
& & & & 1 & & \\
& & & & & \ddots & \\
& & & & & & 1
\end{array}\right)
\begin{array}{l}
\\ \\ \\ \leftarrow i\text{-te Zeile}\ , \quad \lambda \neq 0. \\ \\ \\ \\
\end{array}$$

Dabei sind alle Einträge außer den angegebenen oder durch Punkte angedeuteten Einträgen gleich Null.

Es sei nun $A$ eine beliebige $m \times n$-Matrix. Dann entsprechen die folgenden Produkte gerade den Matrizen, die durch die angegebenen Zeilenoperationen aus $A$ hervorgegangen sind:

$$\begin{array}{rcl}
F_{ik}A & \leftrightarrow & \text{(Z1)} \\
F_{ik}^{\lambda}A & \leftrightarrow & \text{(Z2)} \\
F_{i}^{\lambda}A & \leftrightarrow & \text{(Z3)}
\end{array}$$

**Bemerkung 6.1** Man beachte, dass wir die Elementarmatrizen von links mit der Matrix $A$ multiplizieren. Der Multiplikation der Matrix $A$ mit Elementarmatrizen von rechts entsprechen Spaltenumformungen.

**Satz 6.5** *Die Elementarmatrizen sind invertierbar und ihre Inversen sind wieder Elementarmatrizen. Genauer gilt*

$$F_{ik}^{-1} = F_{ik}, \quad (F_{ik}^{\lambda})^{-1} = F_{ik}^{-\lambda}, \quad (F_{i}^{\lambda})^{-1} = F_{i}^{\frac{1}{\lambda}}.$$

*Beweis.* Dies rechnet man leicht durch Ausmultiplizieren nach. □

**Satz 6.6** *Für eine $n \times n$-Matrix $A$ sind folgende Aussagen äquivalent:*

(a) *$A$ ist invertierbar.*

(b) *Das Gleichungssystem $A\vec{x} = \vec{0}$ hat nur die triviale Lösung.*

(c) *$A$ lässt sich als Produkt von Elementarmatrizen schreiben.*

*Beweis.*
(a) $\Rightarrow$ (b): Es sei $A$ invertierbar. Dann können wir die Gleichung $A\vec{x} = \vec{0}$ von links mit $A^{-1}$ multiplizieren:

$$\begin{aligned} & A^{-1}A\vec{x} = A^{-1}\vec{0} \\ \Rightarrow\quad & E\vec{x} = \vec{0} \\ \Rightarrow\quad & \vec{x} = \vec{0}. \end{aligned}$$

Also hat $A\vec{x} = \vec{0}$ hat nur die triviale Lösung.
(b) $\Rightarrow$ (c): Das Gleichungssystem $A\vec{x} = \vec{0}$ habe nur die triviale Lösung. Mit Hilfe des Gaußschen Algorithmus können wir $A$ nach Satz 5.3 durch elementare Zeilenumformungen auf Zeilenstufenform bringen. Aus Satz 5.1 (ii) folgt, dass dabei eine Matrix $B$ der Form

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

mit von Null verschiedenen Diagonalelementen $b_{11}, \ldots, b_{nn}$ entsteht.

Durch weitere Zeilenumformungen können wir nun $B$ zur Einheitsmatrix $E_n$ machen. Zunächst kann man mit Hilfe der letzten Zeile die Einträge in der letzten Spalte oberhalb von $b_{nn}$ zu Null machen, dann mit Hilfe der vorletzten Zeile die Einträge der vorletzten Spalte oberhalb von $b_{n-1,n-1}$, etc. Damit erhält man eine Matrix, bei der nur noch die Diagonalelemente von Null verschieden sind. Durch Zeilenoperationen vom Typ (Z3) kann man schließlich die Diagonalelemente zu 1 machen. Den Zeilenoperationen entspricht aber die Multiplikation von $A$ mit Elementarmatrizen. Also gibt es Elementarmatrizen $F_1, \ldots, F_r$ mit

$$E_n = F_r \cdots F_1 A.$$

Also folgt

$$A = F_1^{-1} \cdots F_r^{-1} E_n.$$

(c) $\Rightarrow$ (a): Da Elementarmatrizen nach Satz 6.5 invertierbar sind, ist auch $A$ als Produkt invertierbarer Matrizen invertierbar. $\square$

Der Beweis von Satz 6.6 gestattet nun, ein einfaches Verfahren zur Bestimmung der inversen Matrix einer invertierbaren Matrix $A$ anzugeben. Es sei $A$ eine invertierbare $n \times n$-Matrix. Wie im Beweis von Satz 6.6 seien $F_1, \ldots, F_r$ Elementarmatrizen mit

$$E_n = F_r \cdots F_1 A.$$

Indem wir diese Gleichung von rechts mit $A^{-1}$ multiplizieren, erhalten wir daraus

$$A^{-1} = F_r \cdots F_1 E_n.$$

Wir erhalten also $A^{-1}$, indem wir die Einheitsmatrix von links der Reihe nach mit den Elementarmatrizen $F_1, \ldots, F_r$ multiplizieren. Da jede dieser Multiplikationen einer Zeilenoperation entspricht, folgt: *Die Folge von Zeilenoperationen, die $A$ in $E_n$ überführt, überführt gleichzeitig $E_n$ in $A^{-1}$.*

Damit erhalten wir das folgende *Verfahren zur Matrizeninversion*: Wir schreiben die Einheitsmatrix rechts neben $A$

$$(A|E).$$

Nun führen wir geeignete Zeilenoperationen aus, bis auf der linken Seite $E$ steht. Die Zeilenoperationen werden gleichzeitig an $E$ durchgeführt. Dann ergibt sich auf der rechten Seite $A^{-1}$, so dass wir die Matrix

$$(E|A^{-1})$$

erhalten.

**Beispiel 6.7**

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 2 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 2 & 0 & 0 & 1 \end{array}\right)$$

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & -3 & -6 & -2 & 1 & 0 \\ 0 & -2 & -1 & -1 & 0 & 1 \end{array}\right)$$

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & -3 & -6 & -2 & 1 & 0 \\ 0 & 0 & 3 & \frac{1}{3} & -\frac{2}{3} & 1 \end{array}\right)$$

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 0 & \frac{2}{3} & \frac{2}{3} & -1 \\ 0 & -3 & 0 & -\frac{4}{3} & -\frac{1}{3} & 2 \\ 0 & 0 & 3 & \frac{1}{3} & -\frac{2}{3} & 1 \end{array}\right)$$

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -\frac{2}{9} & \frac{4}{9} & \frac{1}{3} \\ 0 & -3 & 0 & -\frac{4}{3} & -\frac{1}{3} & 2 \\ 0 & 0 & 3 & \frac{1}{3} & -\frac{2}{3} & 1 \end{array}\right)$$

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -\frac{2}{9} & \frac{4}{9} & \frac{1}{3} \\ 0 & 1 & 0 & \frac{4}{9} & \frac{1}{9} & -\frac{2}{3} \\ 0 & 0 & 1 & \frac{1}{9} & -\frac{2}{9} & \frac{1}{3} \end{array}\right)$$

# 7 Mengen und Abbildungen

Wir wollen nun einige Grundlagen über Mengen und Abbildungen zusammenstellen.

**Definition (G. Cantor)** Eine *Menge* ist die Zusammenfassung wohlunterschiedener Objekte unseres Denkens zu einem Ganzen.

**Beispiel 7.1** $\mathbb{R} :=$ Menge der reellen Zahlen.

**Notation** Die folgenden Notationen haben wir schon in der Vorlesung benutzt:

- $a \in A$: $a$ ist *Element* von $A$.
- $A \subset B$: $A$ ist *Teilmenge* von $B$ ($a \in A \Rightarrow a \in B$).

Es seien $A$ und $B$ Mengen und $I$ eine beliebige Indexmenge. Dann sind definiert

- *Durchschnitt* $A \cap B := \{x \,|\, x \in A \text{ und } x \in B\}$.
  Allgemeiner:
  $$\bigcap_{i \in I} A_i := \{x \,|\, x \in A_i \text{ für alle } i \in I\}.$$
- *Vereinigung* $A \cup B := \{x \,|\, x \in A \text{ oder } x \in B\}$.
  Allgemeiner:
  $$\bigcup_{i \in I} A_i := \{x \,|\, x \in A_i \text{ für mindestens ein } i \in I\}.$$
- *Differenz* $A \setminus B := \{x \in A \,|\, x \notin B\}$.

- *Kartesisches Produkt*

$$A \times B := \{(a,b) \,|\, a \in A, b \in B\} \text{ (Menge von Paaren).}$$

Beispiel: $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$. Allgemeiner:

$$A_1 \times \cdots \times A_n := \{(a_1, \ldots, a_n) \,|\, a_i \in A_i\} \text{ (Menge der } n\text{-Tupel).}$$

Beispiel: $\mathbb{R}^n = \underbrace{\mathbb{R} \times \cdots \times \mathbb{R}}_{n\text{-mal}}$.

Es seien nun $X, Y$ Mengen.

**Definition** Eine *Abbildung* $f : X \to Y$ ist eine Vorschrift, die jedem Element $x \in X$ genau ein Element $f(x) \in Y$ zuordnet.
Die Menge $X$ heißt *Definitionsbereich* , die Menge $Y$ heißt *Zielbereich* oder *Wertebereich* von $f$.

**Notation** Wir schreiben

$$f : \begin{array}{ccc} X & \longrightarrow & Y \\ x & \longmapsto & f(x) \end{array}.$$

**Beispiel 7.2** (a) Es sei $X$ die Menge aller Studierenden in diesem Hörsaal, $\mathbb{N}$ die Menge der natürlichen Zahlen und $f : X \to \mathbb{N}$ die Abbildung, die jedem Studierenden seine Matrikelnummer zuordnet.

(b) Es sei $X$ die Menge aller Studierenden in diesem Hörsaal, $Y$ die Menge aller Studiengänge an dieser Universität. Die Zuordnung Studierender $\mapsto$ Studiengang ist keine Abbildung, da ein Studierender in zwei Studiengängen eingeschrieben sein kann.

(c) Es sei $f : \mathbb{R} \to \mathbb{R}$ definiert durch $f(x) = x^2$.

(d) Es sei $A$ eine $m \times n$-Matrix und

$$f : \begin{array}{ccc} \mathbb{R}^n & \longrightarrow & \mathbb{R}^m \\ \vec{x} & \longmapsto & A\vec{x} \end{array}.$$

**Definition** Für jede Menge $X$ ist die *Identität*

$$\mathrm{id}_X : \begin{array}{ccc} X & \longrightarrow & X \\ x & \longmapsto & x \end{array}$$

definiert.

**Definition** Es sei $f : X \to Y$ eine Abbildung.

(a) Das *Bild* von $f$ ist die Menge

$$f(X) := \{f(x) \mid x \in X\}.$$

Ist $M \subset X$ eine Teilmenge, so ist entsprechend

$$f(M) := \{f(x) \mid x \in M\}.$$

(b) Ist $N \subset Y$ eine Teilmenge, so ist das *Urbild* von $N$ unter $f$ definiert als

$$f^{-1}(N) := \{x \in X \mid f(x) \in N\}.$$

**Beispiel 7.3** Wir betrachten die Abbildung $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x^2$. Es sei

$$\mathbb{R}_{\geq 0} := \{x \in \mathbb{R} \mid x \geq 0\}.$$

Mit dieser Bezeichnung gilt $f(\mathbb{R}) = \mathbb{R}_{\geq 0}$, $f([0,2]) = [0,4]$, $f^{-1}([0,4]) = [-2,2]$.

**Definition** Es sei $f : X \to Y$ eine Abbildung und $M \subset X$ eine Teilmenge. Dann ist die *Einschränkung von $f$ auf $M$* definiert durch

$$f|_M : \begin{array}{ccc} M & \longrightarrow & Y \\ x & \longmapsto & f(x) \end{array}.$$

**Definition** Es seien $f : X \to Y$, $g : Y \to Z$ Abbildungen. Die *Hintereinanderschaltung* (*Komposition*) von $f$ und $g$ ist die Abbildung

$$g \circ f : \begin{array}{ccc} X & \longrightarrow & Z \\ x & \longmapsto & g(f(x)) \end{array}.$$

Ausführlich

$$\begin{array}{ccccc} X & \stackrel{f}{\to} & Y & \stackrel{g}{\to} & Z \\ x & \mapsto & f(x) & \mapsto & g(f(x)) \end{array}$$

**Beispiel 7.4**

$$\begin{aligned} f : \mathbb{R} \to \mathbb{R}_{\geq 0}, &\qquad f(x) = x^2 \\ g : \mathbb{R}_{\geq 0} \to \mathbb{R}, &\qquad g(x) = \sqrt{x} \\ (g \circ f) : \mathbb{R} \to \mathbb{R}, &\qquad (g \circ f)(x) = g(f(x)) = \sqrt{x^2} = |x|. \end{aligned}$$

**Lemma 7.1** *Sind $f : X \to Y$, $g : Y \to Z$, $h : Z \to W$ Abbildungen, so gilt*

$$h \circ (g \circ f) = (h \circ g) \circ f \text{ (Assoziativität von Abbildungen).}$$

*Beweis.* Beides sind Abbildungen von $X$ nach $W$. Für alle $x \in X$ gilt:

$$\begin{aligned} ((h \circ g) \circ f)(x) &= (h \circ g)(f(x)) = h(g(f(x))) \\ (h \circ (g \circ f))(x) &= h((g \circ f)(x)) = h(g(f(x))) \end{aligned}$$

□

**Definition** Eine Abbildung $f : X \to Y$ heißt

- *injektiv*, falls für alle $x, x' \in X$ gilt: Ist $x \neq x'$, so ist auch $f(x) \neq f(x')$.
- *surjektiv*, falls es zu jedem $y \in Y$ ein $x \in X$ gibt mit $f(x) = y$, d.h. falls $f(X) = Y$ gilt.
- *bijektiv*, falls $f$ injektiv und surjektiv ist.

**Lemma 7.2** *Die folgenden Aussagen sind äquivalent:*

(i) $f$ *ist bijektiv.*

(ii) *Zu jedem* $y \in Y$ *gibt es genau ein* $x \in X$ *mit* $f(x) = y$.

(iii) *Es gibt eine Abbildung* $g : Y \to X$ *mit* $g \circ f = \mathrm{id}_X$ *und* $f \circ g = \mathrm{id}_Y$.

**Definition** Die Abbildung $g$ von Lemma 7.2 heißt *Umkehrabbildung* von $f$ und wird mit $f^{-1}$ bezeichnet.

**Warnung** Das Urbild $f^{-1}(N)$ einer Menge $N \subset Y$ unter einer Abbildung $f : X \to Y$ ist nicht mit der Umkehrabbildung $f^{-1}$, die nur in bestimmten Fällen existiert, zu verwechseln!

*Beweis von Lemma 7.2.* (i) $\Rightarrow$ (ii): Es sei $y \in Y$. Da $f$ bijektiv ist, ist $f$ injektiv und surjektiv. Da $f$ surjektiv ist, gibt es ein $x \in X$ mit $f(x) = y$. Da $f$ injektiv ist, ist dieses $x$ eindeutig bestimmt.

(ii) $\Rightarrow$ (iii): Es sei $y \in Y$ und $x \in X$ das eindeutig bestimmte Element von $X$ mit $f(x) = y$. Wir setzen $g(y) := x$. Dann gilt nach Konstruktion von $g$:

$$(g \circ f)(x) = g(f(x)) = x, \quad (f \circ g)(y) = f(g(y)) = f(x) = y.$$

(iii) $\Rightarrow$ (i): Wir zeigen zunächst, dass $f$ injektiv ist:

$$f(x) = f(x') \Rightarrow g(f(x)) = g(f(x')) \Rightarrow x = x'.$$

Die Abbildung $f$ ist auch surjektiv: Es sei $y \in Y$. Dann gilt

$$f(g(y)) = y.$$

Also ist $g(y) \in X$ ein Element von $X$, das durch $f$ auf $y$ abgebildet wird. □

**Beispiel 7.5** Die Abbildung $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x^2$ ist weder injektiv noch surjektiv. Die Abbildung $f|_{\mathbb{R}_{\geq 0}} : \mathbb{R}_{\geq 0} \to \mathbb{R}$ ist injektiv, aber nicht surjektiv. Dagegen ist die Abbildung

$$\begin{array}{rccc} \overline{f}: & \mathbb{R}_{\geq 0} & \longrightarrow & \mathbb{R}_{\geq 0} \\ & x & \longmapsto & x^2 \end{array}$$

bijektiv mit Umkehrabbildung

$$\begin{array}{rccc} \overline{f}^{-1}: & \mathbb{R}_{\geq 0} & \longrightarrow & \mathbb{R}_{\geq 0} \\ & x & \longmapsto & \sqrt{x} \end{array} .$$

# 8 Algebraische Grundstrukturen

Unser Ziel ist es nun, allgemeine Vektorräume einzuführen. Als Vorbereitung betrachten wir algebraische Grundstrukturen.

**Definition** Eine *Verknüpfung* auf einer Menge $G$ ist eine Abbildung

$$\begin{array}{rccc} *: & G \times G & \longrightarrow & G \\ & (a,b) & \longmapsto & a * b \end{array} .$$

**Notation** $a * b$, $a \cdot b$, $ab$, $a + b$.

**Beispiel 8.1** Für $G = \mathbb{R}$ und $a, b \in G$ sind durch

$$a * b := a + b \text{ oder } a * b := a \cdot b$$

Verknüpfungen auf $\mathbb{R}$ erklärt.

**Beispiel 8.2** Es sei $\mathrm{Mat}(m, n)$ die Menge aller $m \times n$-Matrizen. Für $A, B \in \mathrm{Mat}(m, n)$ definiert

$$A * B := A + B$$

eine Verknüpfung auf $\mathrm{Mat}(m, n)$. Durch

$$A * B := AB \text{ für } A, B \in \mathrm{Mat}(n, n)$$

wird eine Verknüpfung auf $\mathrm{Mat}(n, n)$ erklärt.

**Beispiel 8.3** Es sei $M$ eine beliebige Menge und

$$G := \mathrm{Abb}(M, M) := \text{Menge aller Abbildungen } f : M \to M.$$

Als Verknüpfung wählen wir die Komposition von Abbildungen

$$\begin{array}{rccc} \circ: & G \times G & \longrightarrow & G \\ & (f,g) & \longmapsto & f \circ g \end{array} .$$

**Definition** Eine *Halbgruppe* ist eine Menge $G$ zusammen mit einer Verknüpfung $*$, die das folgende Axiom erfüllt:

(A) $(a * b) * c = a * (b * c)$ für alle $a, b, c \in G$ (*Assoziativgesetz*).

Alle angeführten Beispiele sind Halbgruppen.

**Definition** Eine Halbgruppe $(G, *)$ heißt eine *Gruppe*, wenn gilt:

(N) Es gibt ein $e \in G$ mit $a * e = a$ für alle $a \in G$ (*Neutrales Element*).

(I) Zu jedem $a \in G$ gibt es ein $a' \in G$ mit $a * a' = e$ (*Inverses Element*).

Die Gruppe heißt *abelsch* (oder *kommutativ*), falls zusätzlich folgendes Axiom erfüllt ist:

(K) $a * b = b * a$ für alle $a, b \in G$ (*Kommutativgesetz*).

**Beispiel 8.4** $\mathbb{R}$ mit der Verknüpfung $+$ bildet eine abelsche Gruppe. Die Menge $\mathbb{R}^* := \mathbb{R} \setminus \{0\}$ mit der Verknüpfung $\cdot$ bildet ebenfalls eine abelsche Gruppe. Warum muss man die Null herausnehmen?

**Beispiel 8.5** Die Menge $\mathrm{Mat}(m, n)$ mit der Verknüpfung $A * B := A + B$ für $A, B \in \mathrm{Mat}(m, n)$ ist eine abelsche Gruppe. Es sei $\mathrm{GL}(n)$ die Menge aller invertierbaren $n \times n$-Matrizen. Dann ist $\mathrm{GL}(n)$ mit der Verknüpfung $A * B := AB$ für $A, B \in \mathrm{GL}(n)$ eine Gruppe.

**Beispiel 8.6** Die Halbgruppe $(\mathrm{Abb}(M, M), \circ)$ ist keine Gruppe. Zwar gilt stets

$$f \circ \mathrm{id}_M = f,$$

aber im Allgemeinen hat eine Abbildung $f : M \to M$ keine Umkehrabbildung, also existiert im Allgemeinen zu $f$ kein inverses Element.

Es sei

$$\mathrm{Bij}(M, M) := \{f \in \mathrm{Abb}(M, M) \mid f \text{ bijektiv}\} \subset \mathrm{Abb}(M, M).$$

Dies ist eine Gruppe, denn nach Lemma 7.2 gibt es zu jedem $f$ ein $g$ mit $f \circ g = \mathrm{id}_M$.

**Definition** Es sei $M = \{1, \ldots, n\}$. Die Menge aller bijektiven Abbildungen von $M$ nach $M$ bezeichnet man mit $S_n$ und nennt sie die *symmetrische Gruppe in $n$ Variablen*. Die Elemente von $S_n$ heißen *Permutationen* von $n$ Elementen.

**Bemerkung 8.1** Für $n \geq 3$ ist die Gruppe $S_n$ nicht abelsch: Definiere $f : M \to M$ und $g : M \to M$ durch

$$\begin{aligned} f(1) &:= 1, f(2) := 3, f(3) := 2, f(i) = i \text{ für } 4 \leq i \leq n, \\ g(1) &:= 3, g(2) := 2, g(3) := 1, g(i) = i \text{ für } 4 \leq i \leq n, \end{aligned}$$

Dann ist

$$\begin{aligned} (f \circ g)(1) &= f(g(1)) = f(3) = 2, \\ (g \circ f)(1) &= g(f(1)) = g(1) = 3. \end{aligned}$$

Also ist $f \circ g \neq g \circ f$.

Wir leiten nun einige einfache Folgerungen aus den Gruppenaxiomen ab.

**Satz 8.1** *Ist $G$ eine Gruppe, so gilt:*

(a) *Das neutrale Element $e \in G$ ist eindeutig bestimmt und hat auch die Eigenschaft $e * a = a$ für alle $a \in G$.*

(b) *Das inverse Element $a'$ zu einem Element $a \in G$ ist eindeutig bestimmt und hat auch die Eigenschaft $a' * a = e$. Wir bezeichnen es mit $a^{-1}$.*

(c) $(a^{-1})^{-1} = a$ *für alle $a \in G$.*

(d) $(a * b)^{-1} = b^{-1} * a^{-1}$ *für alle $a, b \in G$.*

(e) *Es gelten die folgenden Kürzungsregeln:*

$$a * x = a * \widetilde{x} \Rightarrow x = \widetilde{x} \text{ und } y * a = \widetilde{y} * a \Rightarrow y = \widetilde{y}.$$

*Beweis.*

(a) Es sei $e \in G$ ein neutrales Element und $a \in G$. Zu $a$ gibt es ein inverses Element $a' \in G$. Zu $a'$ gibt es wiederum ein inverses Element $a'' \in G$ mit $a' * a'' = e$. Damit gilt

$$\begin{aligned} a' * a &= (a' * a) * e = (a' * a) * (a' * a'') = a' * (a * (a' * a'')) \\ &= a' * ((a * a') * a'') = a' * (e * a'') = (a' * e) * a'' = a' * a'' = e. \end{aligned}$$

Daraus folgt

$$e * a = (a * a') * a = a * (a' * a) = a * e = a.$$

Ist $\widetilde{e}$ ebenfalls ein neutrales Element, so gilt $e * \widetilde{e} = e$ und $e * \widetilde{e} = \widetilde{e}$, also $e = \widetilde{e}$.

(b) Dass das inverse Element $a'$ zu $a$ die Eigenschaft $a' * a = e$ hat, wurde bereits in (a) gezeigt. Ist $\widetilde{a}'$ ebenfalls ein inverses Element zu $a \in G$, so folgt daher

$$\widetilde{a}' = e * \widetilde{a}' = (a' * a) * \widetilde{a}' = a' * (a * \widetilde{a}') = a' * e = a'.$$

(c) Aus $a^{-1} * a = e$ folgt, dass $a$ inverses Element zu $a^{-1}$ ist, d.h. $(a^{-1})^{-1} = a$.

(d) Es gilt

$$\begin{aligned} (a * b) * (b^{-1} * a^{-1}) &= a * (b * (b^{-1} * a^{-1})) = a * ((b * b^{-1}) * a^{-1}) \\ &= a * (e * a^{-1}) = a * a^{-1} = e. \end{aligned}$$

(e) Es gilt

$$\begin{aligned} a * x = a * \widetilde{x} &\Rightarrow a^{-1} * (a * x) = a^{-1} * (a * \widetilde{x}) \\ &\Rightarrow (a^{-1} * a) * x = (a^{-1} * a) * \widetilde{x} \\ &\Rightarrow e * x = e * \widetilde{x} \Rightarrow x = \widetilde{x}. \end{aligned}$$

Die andere Kürzungsregel zeigt man analog. □

**Definition** Eine Teilmenge $U$ einer Gruppe $(G, *)$ heißt *Untergruppe* von $G$ genau dann, wenn die folgenden Bedingungen erfüllt sind:

(UG0) $U \neq \emptyset$.

(UG1) Für alle $a, b \in U$ gilt $a * b \in U$.

(UG2) Für alle $a \in U$ gilt $a^{-1} \in U$.

**Satz 8.2** *Eine Untergruppe $U$ einer Gruppe $G$ ist mit der induzierten Verknüpfung eine Gruppe.*

*Beweis.* Wir müssen zeigen, dass die Gruppenaxiome in $U$ erfüllt sind. Nach (UG1) induziert die Verknüpfung auf $G$ eine Verknüpfung auf $U$. Das Assoziativgesetz gilt in $U$, da es in $G$ gilt. Da $U$ nicht leer ist, enthält $U$ mindestens ein Element $u \in U$. Nach (UG2) ist auch $u^{-1} \in U$. Damit ist auch $e = u * u^{-1} \in U$. Nach (UG2) liegt zu jedem Element $a \in U$ das inverse Element $a^{-1}$ in $U$. Also erfüllt $U$ die Gruppenaxiome. □

**Bemerkung 8.2** Die Bedingungen (UG1) und (UG2) sind zu der folgenden einzigen Bedingung äquivalent

(UG) Für alle $a, b \in U$ gilt $a * b^{-1} \in U$.

**Beispiel 8.7** Die Gruppe $(\mathbb{Z}, +)$ ist eine Untergruppe von $(\mathbb{R}, +)$.

**Definition** Es seien $(G, *)$ und $(H, \cdot)$ zwei Gruppen. Eine Abbildung $f : G \to H$ heißt *(Gruppen)homomorphismus* genau dann, wenn für alle $a, b \in G$ gilt

$$f(a * b) = f(a) \cdot f(b).$$

**Beispiel 8.8** Es sei $(G, *) = (H, \cdot) = (\mathbb{Z}, +)$.

(a) Die Abbildung

$$\begin{array}{rccc} f: & \mathbb{Z} & \longrightarrow & \mathbb{Z} \\ & n & \longmapsto & 3n \end{array}$$

ist ein Homomorphismus, denn

$$f(n+m) = 3(n+m) = 3n + 3m = f(n) + f(m).$$

(b) Die Abbildung

$$\begin{array}{rccc} f: & \mathbb{Z} & \longrightarrow & \mathbb{Z} \\ & n & \longmapsto & n+1 \end{array}$$

ist kein Homomorphismus, denn

$$f(n+m) = n + m + 1, \text{ aber } f(n) + f(m) = n + 1 + m + 1 = n + m + 2.$$

**Notation** Wir lassen im Folgenden das Verknüpfungszeichen $*$ weg, d.h. $a * b$ wird einfach als $ab$ geschrieben.

**Definition** Ein Homomorphismus $f : G \to H$ heißt

- *Monomorphismus*, falls $f$ injektiv ist,
- *Epimorphismus*, falls $f$ surjektiv ist,
- *Isomorphismus*, falls es einen Homomorphismus $g : H \to G$ gibt mit $f \circ g = \mathrm{id}_H$ und $g \circ f = \mathrm{id}_G$.

**Lemma 8.1** *Für einen Gruppenhomomorphismus $f : G \to H$ sind äquivalent:*

(i) *$f$ ist bijektiv.*

(ii) *$f$ ist ein Isomorphismus.*

*Beweis.* (ii) $\Rightarrow$ (i) folgt aus Lemma 7.2.

(i) $\Rightarrow$ (ii): Nach Lemma 7.2 gibt es eine Abbildung $g : H \to G$ mit $g \circ f = \mathrm{id}_G$, $f \circ g = \mathrm{id}_H$. Es ist ist noch zu zeigen, dass $g$ ein Gruppenhomomorphismus ist, d.h. dass gilt:

$$g(ab) = g(a)g(b).$$

Nun gilt aber

$$\begin{aligned} f(g(a)g(b)) &= f(g(a))f(g(b)) = ab \text{ (da } f \text{ Gruppenhomomorphismus)}, \\ f(g(ab)) &= ab. \end{aligned}$$

Da $f$ injektiv ist, folgt $g(a)g(b) = g(ab)$. □

**Satz 8.3** *Es sei $f : G \to H$ ein Gruppenhomomorphismus. Dann gilt:*

(i) $f(e_G) = e_H$, *wobei $e_G$ bzw. $e_H$ das neutrale Element von $G$ bzw. $H$ ist.*

(ii) $f(a^{-1}) = f(a)^{-1}$ *für alle $a \in G$.*

*Beweis.* (i) Da $f$ ein Gruppenhomomorphismus ist, gilt

$$f(e_G) = f(e_G e_G) = f(e_G)f(e_G).$$

Daraus folgt

$$e_H = f(e_G)^{-1}f(e_G) = f(e_G)^{-1}f(e_G)f(e_G) = e_H f(e_G) = f(e_G).$$

(ii) Nach (i) gilt

$$f(a)f(a^{-1}) = f(aa^{-1}) = f(e_G) = e_H.$$

Da das inverse Element zu $f(a)$ eindeutig bestimmt ist, folgt $f(a^{-1}) = f(a)^{-1}$. □

**Definition** Es sei $f : G \to H$ ein Gruppenhomomorphismus, $e_G, e_H$ seien die neutralen Elemente von $G$ bzw. $H$.

- Der *Kern* von $f$ ist definiert als

$$\operatorname{Ker} f := \{g \in G \mid f(g) = e_H\}.$$

- Das *Bild* von $f$ ist

$$\operatorname{Im} f := f(G) = \{f(g) \in H \mid g \in G\}.$$

**Beispiel 8.9** Es sei

$$H := C[a,b] := \{f : [a,b] \to \mathbb{R} \mid f \text{ ist stetig}\}.$$

Auf $H$ definieren wir eine Verknüpfung $+$ durch

$$(f+g)(x) := f(x) + g(x).$$

Dann ist $(C[a,b], +)$ eine Gruppe. Ferner sei

$$G := C^1[a,b] := \{f : [a,b] \to \mathbb{R} \mid f \text{ ist stetig differenzierbar}\}.$$

Dann ist $G$ eine Untergruppe von $H$. Die Abbildung

$$\begin{array}{rccl} \dfrac{d}{dx}: & G & \longrightarrow & H \\ & f & \longmapsto & f' = \dfrac{df}{dx} \end{array}$$

ist ein Homomorphismus, denn es gilt

$$(f+g)' = f' + g'.$$

Es gilt

$$\operatorname{Ker} \frac{d}{dx} = \{f \mid f \equiv c\},$$

d.h. der Kern dieser Abbildung besteht gerade aus der Menge der konstanten Funktionen.

**Satz 8.4** *Es sei $f : G \to H$ ein Gruppenhomomorphismus. Dann gilt:*

(i) $\operatorname{Ker} f$ *ist eine Untergruppe von $G$.*

(ii) $\operatorname{Im} f$ *ist eine Untergruppe von $H$.*

(iii) *$f$ ist genau dann injektiv, wenn* $\operatorname{Ker} f = \{e_G\}$.

*Beweis.* (i): Nach Satz 8.3(i) ist $e_G \in \operatorname{Ker} f$. Also ist $\operatorname{Ker} f \neq \emptyset$. Nach Bemerkung 8.2 reicht es zu zeigen, dass mit $a, b \in \operatorname{Ker} f$ auch $ab^{-1} \in \operatorname{Ker} f$. Dies gilt, da

$$f(ab^{-1}) = f(a)f(b^{-1}) \overset{\text{Satz 8.3(ii)}}{=} f(a)f(b)^{-1} = e_H e_H^{-1} = e_H.$$

(ii): Nach Satz 8.3(i) ist $e_H \in \operatorname{Im} f$. Also ist $\operatorname{Im} f \neq \emptyset$. Es sei $u, v \in \operatorname{Im} f$. Wieder reicht es zu zeigen, dass $uv^{-1} \in \operatorname{Im} f$. Da $u, v \in \operatorname{Im} f$, gibt es $a, b \in G$ mit $f(a) = u$, $f(b) = v$. Dann gilt

$$uv^{-1} = f(a)f(b)^{-1} \overset{\text{Satz 8.3(ii)}}{=} f(a)f(b^{-1}) = f(ab^{-1}) \in \operatorname{Im} f.$$

(iii): "$\Rightarrow$": Nach Satz 8.3(i) ist $e_G \in \operatorname{Ker} f$. Da $f$ injektiv ist, folgt $\operatorname{Ker} f = \{e_G\}$.

"$\Leftarrow$": Es seien $a, b \in G$ mit $f(a) = f(b)$. Wir müssen $a = b$ zeigen. Es gilt

$$f(ab^{-1}) = f(a)f(b^{-1}) \overset{\text{Satz 8.3(ii)}}{=} f(a)f(b)^{-1} = f(a)f(a)^{-1} = e_H.$$

Also folgt $ab^{-1} \in \operatorname{Ker} f$. Da $\operatorname{Ker} f = \{e_G\}$, folgt $ab^{-1} = e_G$, also $a = b$, was zu zeigen war. □

**Definition** Es sei $R$ eine Menge mit zwei Verknüpfungen $+, \cdot$. Das Tripel $(R, +, \cdot)$ heißt ein *Ring*, genau dann, wenn die folgenden Axiome erfüllt sind:

(R1) $(R, +)$ ist eine abelsche Gruppe.

(R2) Für alle $r, s, t \in R$ gilt:

$$(r \cdot s) \cdot t = r \cdot (s \cdot t).$$

(R3) Für alle $r, s, t \in R$ gilt:

$$\begin{aligned} (r + s) \cdot t &= r \cdot t + s \cdot t, \\ r \cdot (s + t) &= r \cdot s + r \cdot t. \end{aligned}$$

**Beispiel 8.10** $(\mathbb{Z}, +, \cdot)$, $(\mathbb{Q}, +, \cdot)$. $(\mathbb{R}, +, \cdot)$ sind Ringe. Ebenso ist $\operatorname{Mat}(n, n)$ mit der Matrizenaddition und Matrizenmultiplikation ein Ring.

**Definition** (a) Ein Ring $(R, +, \cdot)$ heißt *kommutativ*, falls

$$r \cdot s = s \cdot r \qquad (\text{für alle } r, s, \in R).$$

(b) Ein *Ring mit Einselement* ist ein Ring $(R, +, \cdot)$, in dem es ein Element $1 \neq 0$ (0 ist das neutrale Element von $(R, +)$) gibt mit

$$1 \cdot r = r \cdot 1 = r \qquad (\text{für alle } r \in R).$$

**Beispiel 8.11** $(\mathbb{Z}, +, \cdot)$, $(\mathbb{Q}, +, \cdot)$. $(\mathbb{R}, +, \cdot)$ sind kommutative Ringe mit Einselement. Der Ring $\operatorname{Mat}(n, n)$ ist nicht kommutativ, besitzt aber ein Einselement: die Einheitsmatrix $E$.

**Notation** Statt $r \cdot s$ schreibt man meist $rs$.

**Definition** Es seien $R, R'$ Ringe. Eine Abbildung $f : R \to R'$ heißt ein *Ringhomomorphismus*, falls für alle $r, s \in R$ gilt:

$$\begin{aligned} &\text{(i)} \qquad f(r + s) = f(r) + f(s) \\ &\text{(ii)} \qquad f(rs) = f(r)f(s) \end{aligned}$$

**Definition** Ein *Körper* ist ein kommutativer Ring $(K, +, \cdot)$ mit Einselement, für den zusätzlich gilt: Für jedes $a \in K$, $a \neq 0$ (0 ist das neutrale Element von $(K, +)$), gibt es ein $a' \in K$ mit

$$a \cdot a' = 1 \qquad \text{(1 ist das Einselement von } K\text{)}.$$

**Bemerkung 8.3** Ein Körper ist ein Tripel $(K, +, \cdot)$, wobei $+, \cdot$ Verknüpfungen auf $K$ sind, für das gilt:

(K1) $(K, +)$ ist eine abelsche Gruppe.

(K2) Ist 0 das neutrale Element von $(K, +)$, so setzt man

$$K^* := K \setminus \{0\}.$$

Für $a, b \in K^*$ gilt $a \cdot b \in K^*$ und $(K^*, \cdot)$ ist eine abelsche Gruppe.

(K3) Es gelten die *Distributivgesetze*: Für alle $a, b, c \in K$

$$\begin{aligned} a \cdot (b + c) &= a \cdot b + a \cdot c \\ (a + b) \cdot c &= a \cdot c + b \cdot c \end{aligned}$$

**Beispiel 8.12** $(\mathbb{Q}, +, \cdot)$ und $(\mathbb{R}, +, \cdot)$ sind Körper, $(\mathbb{Z}, +, \cdot)$ ist kein Körper.

**Beispiel 8.13** Der Körper $(\mathbb{F}_2, +, \cdot)$ ist wie folgt erklärt: Es ist $\mathbb{F}_2 = \{0, 1\}$ und Addition und Multiplikation sind durch die folgenden Tafeln erklärt:

| $+$ | 0 | 1 |
|---|---|---|
| 0 | 0 | 1 |
| 1 | 1 | 0 |

| $\cdot$ | 0 | 1 |
|---|---|---|
| 0 | 0 | 0 |
| 1 | 0 | 1 |

Für konkrete Anwendungen dieses Körpers siehe
http://www.uni-hannover.de/imperia/md/content/alumni/10_hulek.pdf

**Notation**

$$\begin{aligned} 0 &:= \text{neutrales Element von } (K, +) \\ 1 &:= \text{neutrales Element von } (K^*, \cdot) \quad (\text{Es gilt } 1 \neq 0) \\ -a &:= \text{inverses Element zu } a \text{ in } (K, +) \\ \frac{1}{a} &:= a^{-1} := \text{inverses Element zu } a \neq 0 \text{ in } (K^*, \cdot) \\ a - b &:= a + (-b) \\ \frac{a}{b} &:= ab^{-1} \quad (b \neq 0) \end{aligned}$$

**Satz 8.5** *Für einen Körper $K$ gilt:*

(i) $0 \cdot a = a \cdot 0 = 0$ *für alle* $a \in K$.

(ii) *Gilt für $a, b \in K$ die Gleichung $a \cdot b = 0$, so folgt $a = 0$ oder $b = 0$.*

(iii) *Ist $a \in K$, $a \neq 0$, so gibt es genau eine Lösung von $a \cdot x = b$, nämlich $x = \frac{b}{a}$.*

(iv) $a \cdot (-b) = (-a) \cdot b = -(a \cdot b)$, $(-a) \cdot (-b) = a \cdot b$ *für alle* $a, b \in K$.

*Beweis.* (i): $0 \cdot a = (0+0) \cdot a = 0 \cdot a + 0 \cdot a$. Daraus folgt $0 \cdot a = 0$. Die Aussage $a \cdot 0 = 0$ zeigt man analog.

(ii): Aus $a \cdot b = 0$ und $a \neq 0$ folgt $b = a^{-1} \cdot a \cdot b = 0$.

(iii): Für $a \neq 0$ gilt $a^{-1} \cdot a \cdot x = a^{-1} \cdot b$, also $x = a^{-1} \cdot b$.

(iv): $a \cdot b + a \cdot (-b) = a \cdot (b + (-b)) = a \cdot 0 \overset{\text{(i)}}{=} 0$. Daraus folgt $a \cdot (-b) = -(a \cdot b)$. Analog zeigt man $(-a) \cdot b = -(a \cdot b)$. Schließlich gilt

$$(-a) \cdot (-b) = -(a \cdot (-b)) = -(-(a \cdot b)) \overset{\text{Satz 8.1(iii)}}{=} a \cdot b.$$

□

**Definition** Es seien $K, K'$ Körper. Eine Abbildung $f : K \to K'$ heißt ein *Körperhomomorphismus*, falls für alle $a, b \in K$ gilt:

$$\begin{aligned} f(a+b) &= f(a) + f(b) \\ f(a \cdot b) &= f(a) \cdot f(b) \end{aligned}$$

Ein Körperhomomorphismus $f$ heißt *Körpermonomorphismus* (*-epimorphismus*), falls $f$ injektiv (surjektiv) ist.

Ein Körperhomomorphismus $f$ heißt ein *Körperisomorphismus*, falls es einen Körperhomomorphismus $g : K' \to K$ gibt mit $f \circ g = \mathrm{id}_{K'}$ und $g \circ f = \mathrm{id}_K$.

**Warnung** Wir verwenden hier missbräuchlich dasselbe Zeichen für die Operationen in $K$ und $K'$!

**Lemma 8.2** *Ein Körperhomomorphismus $f : K \to K'$ ist genau dann ein Körperisomorphismus, wenn $f$ bijektiv ist.*

*Beweis.* Der Beweis geht genau wie der von Lemma 8.1. □

**Satz 8.6** *Die Menge $\mathbb{R} \times \mathbb{R}$ aller (geordneten) Paare reeller Zahlen bildet zusammen mit der Addition und Multiplikation*

$$\begin{aligned}(x, y) + (x', y') &:= (x + x', y + y') \\ (x, y) \cdot (x', y') &:= (xx' - yy', xy' + yx')\end{aligned}$$

*einen Körper.*

**Definition** Dieser Körper heißt der Körper $\mathbb{C}$ der *komplexen Zahlen.*

*Beweis von Satz 8.6.* Wir weisen die Eigenschaften (K1), (K2), (K3) von Bemerkung 8.3 nach.

(K1): $(\mathbb{C}, +)$ ist eine abelsche Gruppe mit neutralem Element $e := (0, 0)$ und inversem Element $-(x, y) = (-x, -y)$ zu einem Element $(x, y) \in \mathbb{C}$.

(K2): Es ist zunächst zu zeigen: Ist $(x, y) \neq (0, 0)$ und $(x', y') \neq (0, 0)$, so folgt $(x, y) \cdot (x', y') \neq (0, 0)$.
Angenommen

$$(x, y) \cdot (x', y') = (xx' - yy', xy' + yx') = (0, 0).$$

Wir nehmen zunächst an: $x \neq 0$. Dann gilt:

$$xx' - yy' = 0 \Rightarrow x' = \frac{yy'}{x}.$$

Einsetzen ergibt:

$$xy' + yx' = 0 \Rightarrow xy' + \frac{y^2 y'}{x} = 0 \Rightarrow \underbrace{(x^2 + y^2)}_{\neq 0} y' = 0 \Rightarrow y' = 0.$$

Dann folgt auch $x' = 0$, also $(x', y') = (0, 0)$. Der Fall $y \neq 0$ geht analog.
Neutrales Element: Das neutrale Element ist $(1, 0)$, denn es gilt

$$(1, 0) \cdot (x, y) = (1x - 0y, 1y - 0x) = (x, y).$$

Inverses Element: Das zu $(x, y) \neq (0, 0)$ inverse Element ist

$$(x, y)^{-1} = \left( \frac{x}{x^2 + y^2}, \frac{-y}{x^2 + y^2} \right).$$

Es gilt nämlich

$$\begin{aligned}\left( \frac{x}{x^2 + y^2}, \frac{-y}{x^2 + y^2} \right) \cdot (x, y) &= \left( \frac{x^2}{x^2 + y^2} + \frac{y^2}{x^2 + y^2}, \frac{xy}{x^2 + y^2} - \frac{xy}{x^2 + y^2} \right) \\ &= (1, 0).\end{aligned}$$

Das Assoziativgesetz und das Kommutativgesetz für die Multiplikation rechnet man leicht nach.

(K3) Die Distributivgesetze rechnet man auch leicht nach. □

Wir betrachten die Abbildung

$$\begin{array}{rccc} f: & \mathbb{R} & \longrightarrow & \mathbb{C} \\ & x & \longmapsto & (x,0) \end{array}.$$

Dies ist ein Körpermonomorphismus: Die Injektivität ist klar. Es gilt außerdem:

$$\begin{aligned} f(x+y) &= (x+y,0) = (x,0)+(y,0) = f(x)+f(y) \\ f(xy) &= (xy,0) = (x,0)\cdot(y,0) = f(x)\cdot f(y) \end{aligned}$$

Man sagt, mittels $f$ wird der Körper der reellen Zahlen in den Körper der komplexen Zahlen *eingebettet*.

**Definition** Das Element

$$i := (0,1)$$

heißt *imaginäre Einheit*.

Es gilt

$$i^2 = (0,1)\cdot(0,1) = (0-1,0) = (-1,0) = -(1,0) = f(-1).$$

Identifiziert man $\mathbb{R}$ mit $f(\mathbb{R})$, so kann man dafür kurz schreiben:

$$i^2 = -1.$$

**Definition** Wir schreiben nun

$$x+iy := (x,0)+(0,1)\cdot(y,0) = (x,0)+(0,y) = (x,y).$$

Wir nennen $x$ den *Realteil* und $y$ den *Imaginärteil* der komplexen Zahl $z = x+iy$.

**Notation** $x = \operatorname{Re} z$, $y = \operatorname{Im} z$.

Geometrisch kann man $\mathbb{C}$ mit der Ebene $\mathbb{R}^2$ identifizieren. Diese Ebene nennt man in diesem Fall auch die *Gaußsche Zahlenebene*. Die $x$-Achse nennt man die *reelle Achse*, die $y$-Achse die *imaginäre Achse*.

Addition und Multiplikation drücken sich in der neuen Schreibweise wie folgt aus:

$$\begin{aligned} (x+iy)+(x'+iy') &= (x+x')+i(y+y') \\ (x+iy)\cdot(x'+iy') &= (xx'-yy')+i(xy'+yx'). \end{aligned}$$

**Definition** Ist $z = x + iy$ eine komplexe Zahl, so heißt

$$\overline{z} := x - iy$$

die zu $z$ *konjugiert komplexe* Zahl.

**Bemerkung 8.4** Es gilt $\overline{z+w} = \overline{z} + \overline{w}$, $\overline{zw} = \overline{z}\,\overline{w}$, $\overline{\overline{z}} = z$, $z = \overline{z} \Leftrightarrow z \in \mathbb{R}$.

**Definition** Ist $z = x + iy$, so heißt

$$|z| := \sqrt{z\overline{z}} = \sqrt{(x+iy)(x-iy)} = \sqrt{x^2+y^2}$$

der *Betrag* von $z$.

Der Betrag von $z = x + iy$ ist also die Länge des Vektors $(x, y)$.

In vielen Fällen ist es nützlich, komplexe Zahlen durch Polarkoordinaten darzustellen. Eine komplexe Zahl $z \neq 0$ wird durch ihren Betrag $r = |z|$ und den Winkel $\varphi$ zur reellen Achse eindeutig bestimmt ($\varphi \in [0, 2\pi)$). Setzt man

$$e^{i\varphi} := \cos\varphi + i \sin\varphi,$$

so ist

$$z = re^{i\varphi}.$$

Mit Hilfe von Polarkoordinaten kann man nun gut die Multiplikation beschreiben: Es sei

$$z = re^{i\varphi}, \quad z' = r'e^{i\varphi'}.$$

Dann folgt mit Hilfe der Additionstheoreme für sin und cos

$$\begin{aligned}
zz' &= (re^{i\varphi})(r'e^{i\varphi'}) = rr'(\cos\varphi + i\sin\varphi)(\cos\varphi' + i\sin\varphi') \\
&= rr'((\cos\varphi\cos\varphi' - \sin\varphi\sin\varphi') + i(\sin\varphi\cos\varphi' + \cos\varphi\sin\varphi')) \\
&= rr'(\cos(\varphi+\varphi') + i\sin(\varphi+\varphi')) \\
&= rr'e^{i(\varphi+\varphi')}.
\end{aligned}$$

Das bedeutet: Bei der Multiplikation zweier komplexer Zahlen multiplizieren sich die Beträge und die Winkel addieren sich.

# 9 Vektorräume

Wir wollen nun allgemeine Vektorräume einführen. Es sei $K$ ein Körper.

**Definition** Eine Menge $V$ zusammen mit einer inneren Verknüpfung (*Addition* genannt)

$$\begin{array}{rccl} + : & V \times V & \longrightarrow & V \\ & (x, y) & \longmapsto & x + y \end{array}$$

und einer äußeren Verknüpfung (*Multiplikation mit Skalaren* oder *skalare Multiplikation* genannt)

$$\begin{array}{rccl} \cdot : & K \times V & \longrightarrow & V \\ & (\lambda, x) & \longmapsto & \lambda \cdot x =: \lambda x \end{array}$$

heißt $K$-*Vektorraum* (oder *Vektorraum über* $K$) genau dann, wenn die folgenden Axiome erfüllt sind:

(V1) $(V, +)$ ist eine abelsche Gruppe.

(V2) Für alle $\lambda, \mu \in K$ und $x, y \in V$ gilt:

- (a) $\lambda \cdot (x + y) = \lambda \cdot x + \lambda \cdot y$.
- (b) $(\lambda + \mu) \cdot x = \lambda \cdot x + \mu \cdot x$.
- (c) $\lambda \cdot (\mu \cdot x) = (\lambda\mu) \cdot x$.
- (d) $1 \cdot x = x$ $(1 = 1_K)$.

Die Elemente $\lambda, \mu \in K$ heißen *Skalare*. Die Elemente von $V$ werden als *Vektoren* bezeichnet. Wir bezeichnen mit $0 = 0_K$ bzw. $1 = 1_K$ das additive bzw. multiplikative neutrale Element in $K$. Das neutrale Element von $V$ bezüglich der Addition wird *Nullvektor* genannt und mit $0 = 0_V$ bezeichnet. Das inverse Element zu einem Vektor $x$ bezüglich der Addition wird das *Negative* von $x$ genannt und mit $-x$ bezeichnet.

**Beispiel 9.1** Für einen Körper $K$ setzen wir

$$K^n := \underbrace{K \times \cdots \times K}_{n\text{-mal}}.$$

Im Fall $K = \mathbb{R}$ erhalten wir den früher betrachteten $\mathbb{R}^n$. Wir schreiben die Elemente von $K^n$ wieder als Spaltenvektoren.Wie in §1 definieren wir:
Addition:

$$x + y = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} := \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix}.$$

Skalare Multiplikation:

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

Dann ist $K^n$ mit dieser Addition und skalaren Multiplikation ein Vektorraum. Die Axiome prüft man leicht nach. Als Beispiel beweisen wir (V2) (a). Es sei $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, $y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$. Dann gilt

$$\begin{aligned} \lambda \cdot (x+y) &= \lambda \left( \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \right) = \lambda \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix} \\ &= \begin{pmatrix} \lambda(x_1 + y_1) \\ \vdots \\ \lambda(x_n + y_n) \end{pmatrix} = \begin{pmatrix} \lambda x_1 + \lambda y_1 \\ \vdots \\ \lambda x_n + \lambda y_n \end{pmatrix} \\ &= \lambda \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \lambda \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \lambda x + \lambda y. \end{aligned}$$

**Beispiel 9.2** Die Menge $\mathrm{Mat}(m,n)$ zusammen mit der in §6 definierten Addition und skalaren Multiplikation ist ein Vektorraum. Wieder prüft man die Axiome leicht nach. Der Nullvektor ist die $m \times n$-Nullmatrix, das Negative zu $A \in \mathrm{Mat}(m,n)$ ist die Matrix $-A$.

**Beispiel 9.3** (a) Für einen Körper $K$ und eine Menge $M$ definieren wir:

$$\mathrm{Abb}(M,K) := \text{Menge aller Abbildungen } f : M \to K.$$

Auf $\mathrm{Abb}(M,K)$ führen wir eine Addition und eine skalare Multiplikation wie folgt ein: Für zwei Elemente $f, g \in \mathrm{Abb}(M,K)$ und ein $\lambda \in K$ definieren wir die Summe $f+g$ und das skalare Vielfache $\lambda f$ durch

$$\begin{aligned} (f+g)(x) &= f(x) + g(x), \\ (\lambda f)(x) &= \lambda f(x). \end{aligned}$$

für alle $x \in M$. Mit diesen Verknüpfungen ist $\mathrm{Abb}(M,K)$ ein Vektorraum:
(V1): $(\mathrm{Abb}(M,K), +)$ ist eine abelsche Gruppe mit 0 definiert durch $0(x) :=$

0 für alle $x \in M$, inverses Element $-f$ definiert durch $(-f)(x) := -(f(x))$ für alle $x \in M$.
Die Axiome (V2) prüft man leicht nach. Als Beispiel zeigen wir
(V2) (a): $\lambda(x+y) = \lambda x + \lambda y$.
Es gilt

$$\begin{aligned} (\lambda(f+g))(x) &= \lambda((f+g)(x)) = \lambda(f(x)+g(x)) \\ &= \lambda f(x) + \lambda g(x) = (\lambda f)(x) + (\lambda g)(x) \\ &= (\lambda f + \lambda g)(x). \end{aligned}$$

(b) Spezialfall: $K = \mathbb{R}$, $M := [a,b] \subset \mathbb{R}$.

$$\mathrm{Abb}(M,K) = \mathrm{Abb}([a,b],\mathbb{R}) := \{f \mid f : [a,b] \to \mathbb{R}\}.$$

Dies kann man weiter spezialisieren:

$$\begin{aligned} C[a,b] &:= \{f \in \mathrm{Abb}(M,K) \mid f : [a,b] \to \mathbb{R} \text{ stetig}\} \\ C^1[a,b] &:= \{f \in \mathrm{Abb}(M,K) \mid f : [a,b] \to \mathbb{R} \text{ stetig differenzierbar}\} \\ C^\infty[a,b] &:= \{f \in \mathrm{Abb}(M,K) \mid f : [a,b] \to \mathbb{R} \text{ beliebig oft differenzierbar}\} \end{aligned}$$

(c) Es sei $K = \mathbb{R}$,

$$V = \mathbb{R}[x] := \{a_0 + a_1 x + \cdots + a_n x^n \mid a_i \in \mathbb{R}\}$$

der *Raum der Polynome* in der Variablen $x$. Für $\lambda \in \mathbb{R}$ und

$$\begin{aligned} p(x) &= a_0 + a_1 x + \cdots + a_n x^n \\ q(x) &= b_0 + b_1 x + \cdots + b_m x^m \end{aligned}$$

definieren wir (falls $n \leq m$)

$$\begin{aligned} p(x) + q(x) &:= (a_0 + b_0) + (a_1 + b_1)x + \cdots + (a_m + b_m)x^m, \\ \lambda p(x) &:= (\lambda a_0) + (\lambda a_1)x + \cdots + (\lambda a_n)x^n, \end{aligned}$$

wobei wir $a_i = 0$ für $n < i \leq m$ gesetzt haben. Dann ist $V$ ein $\mathbb{R}$-Vektorraum.

**Beispiel 9.4** Die Menge $V$ enthalte nur ein Element, das wir mit 0 bezeichnen. Definieren wir

$$\begin{aligned} 0 + 0 &= 0 \\ \lambda \cdot 0 &= 0 \end{aligned}$$

für alle $\lambda \in K$, so erfüllt $V$ die Vektorraumaxiome. Der $K$-Vektorraum $V$ heißt der *Nullvektorraum*.

**Satz 9.1** *Es sei $V$ ein Vektorraum über $K$, $x \in V$ und $\lambda \in K$. Dann gilt:*

(a) $0 \cdot x = 0$.

(b) $\lambda \cdot 0 = 0$.

(c) $(-1) \cdot x = -x$.

(d) *Aus* $\lambda \cdot x = 0$ *folgt* $\lambda = 0$ *oder* $x = 0$.

*Beweis.* (a) Aus

$$0 \cdot x = (0+0) \cdot x \overset{\text{(V2b)}}{=} 0 \cdot x + 0 \cdot x$$

folgt $0 \cdot x = 0$.

(b) Aus

$$\lambda \cdot 0 = \lambda \cdot (0+0) \overset{\text{(V2a)}}{=} \lambda \cdot 0 + \lambda \cdot 0$$

folgt $\lambda \cdot 0 = 0$.

(c) Aus

$$x + (-1) \cdot x \overset{\text{(V2d)}}{=} 1 \cdot x + (-1) \cdot x \overset{\text{(V2b)}}{=} (1-1) \cdot x = 0 \cdot x \overset{\text{(a)}}{=} 0$$

folgt $(-1) \cdot x = -x$.

(d) Ist $\lambda \cdot x = 0$ und $\lambda \neq 0$, so folgt

$$x \overset{\text{(V2d)}}{=} 1 \cdot x = (\frac{1}{\lambda}\lambda) \cdot x \overset{\text{(V2c)}}{=} \frac{1}{\lambda} \cdot (\lambda \cdot x) = \frac{1}{\lambda} \cdot 0 \overset{\text{(b)}}{=} 0.$$

□

Wir haben bisher zwischen der Multiplikation in $K$, die wir ohne Zeichen geschrieben haben, und der skalaren Multiplikation, die wir mit $\cdot$ bezeichnet haben, unterschieden. Der Satz zeigt, dass dies eigentlich nicht nötig ist, und deswegen werden wir in Zukunft das Zeichen $\cdot$ weglassen.

In § 5 hatten wir die Lösungsmengen $\mathcal{L}(A, \vec{b})$ von linearen Gleichungssystemen $A\vec{x} = \vec{b}$ betrachtet. Wir interessieren uns nun für den Fall $\vec{b} = \vec{0}$. Wir werden zeigen, dass in diesem Fall die Lösungsmenge $\mathcal{L}(A, 0) \subset \mathbb{R}^n$ von $\mathbb{R}^n$ die Vektorraumstruktur erbt. Dazu führen wir den Begriff des Unterraums ein.

**Definition** Es sei $V$ ein $K$-Vektorraum. Eine Teilmenge $U \subset V$ heißt *Unterraum von $V$* genau dann, wenn folgende Bedingungen erfüllt sind:

(U0) $U \neq \emptyset$.

(U1) $x, y \in U \Rightarrow x + y \in U$
(d.h. *U ist abgeschlossen unter der Addition*).

(U2) $\lambda \in K, x \in U \Rightarrow \lambda x \in U$
(d.h. *U ist abgeschlossen unter der Multiplikation mit Skalaren*).

Aus dieser Definition folgt, dass wir auf einem Unterraum $U \subset V$ eine Addition und eine Multiplikation mit Skalaren haben, die von der Addition und Multiplikation mit Skalaren in $V$ *induziert* werden.

**Satz 9.2** *Ein Unterraum $U \subset V$ eines Vektorraums $V$ ist zusammen mit der induzierten Addition und skalaren Multiplikation wieder ein Vektorraum.*

*Beweis.* Nach (U0) gibt es ein $x \in U$. Nach Satz 9.1(a) gilt $0 \cdot x = 0$. Nach (U2) liegt daher auch der Nullvektor 0 in $U$. Zu jedem $x \in U$ ist nach (U2) auch $-x = (-1) \cdot x \in U$. Die übrigen Vektorraumaxiome gelten in $U$, da sie in $V$ gelten. □

**Beispiel 9.5** Es sei $A$ eine $m \times n$-Matrix. Ein lineares Gleichungssystem $A\vec{x} = \vec{0}$ bezeichnet man als ein *homogenes* lineares Gleichungssystem. Die Lösungsmenge $\mathcal{L}(A, \vec{0})$ eines homogenen linearen Gleichungssystems ist ein Unterraum des $\mathbb{R}^n$. Wir prüfen die Axiome (U0)–(U2) nach:

(U0) $\mathcal{L}(A, \vec{0})$ ist nicht leer, da $\mathcal{L}(A, \vec{0})$ mindestens den Nullvektor $\vec{0}$ enthält.

(U1) Sind $\vec{x}$ und $\vec{x}'$ Vektoren aus $\mathcal{L}(A, \vec{0})$, so gilt

$$A\vec{x} = \vec{0} \text{ und } A\vec{x}' = \vec{0}.$$

Daraus folgt

$$A(\vec{x} + \vec{x}') = A\vec{x} + A\vec{x}' = \vec{0} + \vec{0} = \vec{0}.$$

Also ist $\vec{x} + \vec{x}' \in \mathcal{L}(A, \vec{0})$.

(U2) Ist $\vec{x}$ ein Vektor aus $\mathcal{L}(A, \vec{0})$ und $\lambda \in \mathbb{R}$, so gilt

$$A(\lambda\vec{x}) = \lambda A\vec{x} = \lambda\vec{0} = \vec{0}.$$

Also ist auch $\lambda\vec{x} \in \mathcal{L}(A, \vec{0})$.

Insbesondere sind Geraden durch den Ursprung im $\mathbb{R}^2$ Unterräume des $\mathbb{R}^2$ und Geraden und Ebenen im $\mathbb{R}^3$, die den Ursprung enthalten, Unterräume des $\mathbb{R}^3$.

**Beispiel 9.6** Es sei $U = \{(x, y) \in \mathbb{R}^2 \,|\, x \geq 0, y \geq 0\}$. Dann ist $U$ kein Unterraum des $\mathbb{R}^2$, da zum Beispiel $x = (1, 1)$ in $U$ liegt, aber nicht $(-1)x = -x = (-1, -1)$.

# 10 Lineare Hülle und lineare Unabhängigkeit

Wir betrachten nun wichtige Begriffsbildungen in Vektorräumen. Es sei $V$ ein Vektorraum über $K$.

**Definition** Ein Vektor $w \in V$ heißt *Linearkombination* der Vektoren $v_1$, …, $v_r$ genau dann, wenn es $\lambda_1, \ldots, \lambda_r \in K$ gibt, so dass gilt:

$$w = \lambda_1 v_1 + \cdots + \lambda_r v_r.$$

Die Linearkombination heißt *trivial*, falls $\lambda_1 = \cdots = \lambda_r = 0$ ist.

**Beispiel 10.1** Jeder Vektor $\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3$ ist eine Linearkombination der Standardbasisvektoren

$$\vec{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \vec{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \vec{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},$$

denn es gilt

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = x_1\vec{e}_1 + x_2\vec{e}_2 + x_3\vec{e}_3.$$

**Beispiel 10.2** Es sei $\vec{u} = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}$ und $\vec{v} = \begin{pmatrix} 3 \\ 1 \\ -1 \end{pmatrix}$. Dann ist $\vec{w} = \begin{pmatrix} 4 \\ 1 \\ 1 \end{pmatrix}$ keine Linearkombination von $\vec{u}$ und $\vec{v}$.

*Beweis.* Wenn $\vec{w}$ Linearkombination von $\vec{u}$ und $\vec{v}$ wäre, so müsste es $\lambda_1, \lambda_2 \in \mathbb{R}$ geben, so dass

$$\lambda_1 \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} + \lambda_2 \begin{pmatrix} 3 \\ 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 4 \\ 1 \\ 1 \end{pmatrix}.$$

Daraus ergibt sich das Gleichungssystem

$$\begin{aligned} \lambda_1 + 3\lambda_2 &= 4 \\ -\lambda_1 + \lambda_2 &= 1 \\ 2\lambda_1 - \lambda_2 &= 1 \end{aligned}$$

Dieses Gleichungssystem besitzt aber keine Lösung. □

**Definition** Es sei $S \subset V$ eine nicht leere Teilmenge. Die Menge

$$\operatorname{Span} S := \{v \in V \mid v = \lambda_1 v_1 + \ldots + \lambda_r v_r, v_i \in S, \lambda_i \in K\},$$

also die Menge aller Linearkombinationen von Elementen aus $S$, heißt *Spann* oder *lineare Hülle* von $S$. Wir setzen $\operatorname{Span} \emptyset = \{0\}$.

**Satz 10.1** *Es sei $S \subset V$. Dann gilt:*

(a) *Die Menge* $\operatorname{Span} S$ *ist ein Unterraum von $V$.*

(b) $\operatorname{Span} S$ *ist der kleinste Unterraum von $V$, der $S$ enthält, d.h. jeder Unterraum von $V$, der $S$ enthält, enthält auch* $\operatorname{Span} S$.

*Beweis.* (a) ist klar.

(b) Es sei $W$ ein Unterraum von $V$, der $S$ enthält. Sind $v_1, \ldots, v_r \in S \subset W$, so ist auch jede Linearkombination $\lambda_1 v_1 + \ldots + \lambda_r v_r$ dieser Vektoren in $W$. Also gilt $\operatorname{Span} S \subset W$. □

**Notation** Ist $S = \{v_1, \ldots, v_r\}$, so schreibt man auch

$$\operatorname{Span} S = \operatorname{Span}(v_1, \ldots, v_r).$$

**Beispiel 10.3** Sind $\vec{v}_1, \vec{v}_2$ zwei linear unabhängige Vektoren im $\mathbb{R}^3$, so ist $\operatorname{Span}(\vec{v}_1, \vec{v}_2)$ eine Ebene durch den Nullpunkt.

**Definition** Eine Teilmenge $S \subset V$ heißt *Erzeugendensystem* von $V$, falls $V = \operatorname{Span} S$ ist.

Ein Vektorraum $V$ heißt *endlich erzeugt*, falls es eine endliche Menge $S \subset V$ mit $V = \operatorname{Span} S$ gibt.

**Beispiel 10.4** Wir betrachten in dem Vektorraum $K^n$ die Vektoren

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, e_2 := \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \ldots, e_n = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}.$$

Dann gilt $K^n = \operatorname{Span}(e_1, \ldots, e_n)$, da für $x = (x_1, \ldots, x_n)^T \in K^n$ gilt

$$x = x_1 e_1 + \ldots + x_n e_n.$$

Der Vektorraum $K^n$ ist endlich erzeugt.

**Beispiel 10.5** Die Menge $S = \{1, x, x^2, \ldots\}$ ist ein Erzeugendensystem des Vektorraums $\mathbb{R}[x]$. Denn jedes Polynom $p(x)$ ist eine Linearkombination

$$p(x) = a_0 + a_1 x + \cdots + a_n x^n$$

von solchen Monomen. Der Vektorraum $\mathbb{R}[x]$ ist nicht endlich erzeugt, da es Polynome von beliebig hohem Grad gibt.

Ein weiterer wichtiger Begriff ist der Begriff der linearen Unabhängigkeit. Für zwei Vektoren des $\mathbb{R}^n$ haben wir diesen Begriff schon in § 4 eingeführt.

**Definition** Die Vektoren $v_1, \ldots, v_r \in V$ heißen *linear unabhängig* genau dann, wenn gilt: Sind $\lambda_1, \ldots, \lambda_r \in K$ und ist

$$\lambda_1 v_1 + \cdots + \lambda_r v_r = 0,$$

so folgt

$$\lambda_1 = \cdots = \lambda_r = 0.$$

Anders ausgedrückt, bedeutet dies, dass sich der Nullvektor nur als triviale Linearkombination

$$0 = 0v_1 + 0v_2 + \cdots + 0v_r$$

der Vektoren $v_1, v_2, \ldots, v_r$ darstellen läßt.

Die Vektoren $v_1, \ldots, v_r \in V$ heißen *linear abhängig* genau dann, wenn sie nicht linear unabhängig ist, d.h. wenn es Skalare $\lambda_1, \ldots, \lambda_r$ gibt, die nicht alle gleich Null sind, so dass

$$\lambda_1 v_1 + \cdots + \lambda_r v_r = 0.$$

**Definition** Eine Teilmenge $S \subset V$ heißt *linear unabhängig* genau dann, wenn je endlich viele Vektoren $v_1, \ldots, v_r \in S$ linear unabhängig sind.

**Beispiel 10.6** Im $K^n$ sind die Vektoren $e_1, \ldots, e_n$ linear unabhängig. Denn aus der Gleichung

$$\lambda_1 e_1 + \cdots + \lambda_n e_n = 0$$

ergibt sich $(\lambda_1, \ldots, \lambda_n)^T = 0$, also $\lambda_1 = \cdots = \lambda_n = 0$.

**Beispiel 10.7** Die Vektoren $\vec{v}_1, \vec{v}_2, \vec{v}_3 \in \mathbb{R}^3$ mit

$$\vec{v}_1 := \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \vec{v}_2 := \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \vec{v}_3 := \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$$

sind linear abhängig, denn es gilt $\vec{v}_1 + \vec{v}_2 - \vec{v}_3 = \vec{0}$.

**Beispiel 10.8** Die Menge $S := \{1, x, x^2, \dots\} \subset \mathbb{R}[x]$ ist linear unabhängig in $\mathbb{R}[x]$. Denn aus

$$a_0 + a_1x + a_2x^2 + \cdots + a_nx^n = 0 \text{ für alle } x \in \mathbb{R}$$

folgt

$$a_0 = a_1 = a_2 = \cdots = a_n = 0,$$

da ein Polynom $n$-ten Grades höchstens $n$ Nullstellen hat.

**Satz 10.2** *Ist $r \geq 2$, so sind die Vektoren $v_1, \dots, v_r \in V$ genau dann linear abhängig, wenn mindestens einer davon als Linearkombination der anderen dargestellt werden kann.*

*Beweis.*
"$\Rightarrow$": Sind $v_1, \dots, v_r$ linear abhängig, so gibt es $\lambda_1, \lambda_2, \dots, \lambda_r \in K$ mit $\lambda_1 v_1 + \cdots + \lambda_r v_r = 0$ und ein $k \in \{1, \dots, r\}$ mit $\lambda_k \neq 0$. Dann gilt

$$v_k = -\frac{\lambda_1}{\lambda_k} v_1 - \cdots - \frac{\lambda_{k-1}}{\lambda_k} v_{k-1} - \frac{\lambda_{k+1}}{\lambda_k} v_{k+1} - \cdots - \frac{\lambda_r}{\lambda_k} v_r.$$

"$\Leftarrow$": Ist $v_k = \mu_1 v_1 + \cdots + \mu_{k-1} v_{k-1} + \mu_{k+1} v_{k+1} + \cdots + \mu_r v_r$, so gilt

$$\mu_1 v_1 + \cdots + \mu_{k-1} v_{k-1} + (-1) v_k + \mu_{k+1} v_{k+1} + \cdots + \mu_r v_r = 0.$$

□

**Bemerkung 10.1** Aus Satz 10.2 folgt, dass Satz 4.2 auch in einem allgemeinen $K$-Vektorraum gilt. Wir erhalten als Folgerung aus Satz 10.2:

(a) *Eine Menge, die den Nullvektor enthält, ist linear abhängig.*

(b) *Zwei Vektoren sind genau dann linear abhängig, wenn einer der Vektoren ein skalares Vielfaches des anderen ist.*

**Satz 10.3** *Die Vektoren $v_1, \dots, v_r \in V$ sind genau dann linear unabhängig, wenn sich jeder Vektor aus $\operatorname{Span}(v_1, \dots, v_n)$ in eindeutiger Weise als Linearkombination der Vektoren $v_1, \dots, v_r$ schreiben lässt.*

*Beweis.*
"$\Leftarrow$": Angenommen, $v_1, \dots, v_r$ sind linear abhängig. Dann lässt sich der Nullvektor $0 \in \operatorname{Span}(v_1, \dots, v_r)$ auf mindestens zwei verschiedene Weisen als Linearkombination der Vektoren $v_1, \dots, v_r$ schreiben.

"$\Rightarrow$": Die Vektoren $v_1, \ldots, v_r$ seien linear unabhängig. Angenommen, wir hätten für den Vektor $v \in \mathrm{Span}(v_1, \ldots, v_n)$ die Darstellungen

$$\begin{aligned} v &= \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_r v_r, \\ v &= \mu_1 v_1 + \mu_2 v_2 + \cdots + \mu_r v_r. \end{aligned}$$

Durch Subtraktion dieser Gleichungen folgt

$$0 = (\lambda_1 - \mu_1) v_1 + (\lambda_2 - \mu_2) v_2 + \cdots + (\lambda_r + \mu_r) v_r.$$

Da $S$ linear unabhängig ist, ergibt sich

$$\lambda_1 - \mu_1 = \lambda_2 - \mu_2 = \cdots = \lambda_r - \mu_r = 0,$$

also

$$\lambda_1 = \mu_1,\ \lambda_2 = \mu_2,\ \ldots, \lambda_r = \mu_r.$$

Also ist die Darstellung von $v$ eindeutig. $\square$

# 11 Basis und Dimension

Wir wollen nun die grundlegenden Begriffe Basis und Dimension einführen. Wie stets sei $V$ ein $K$-Vektorraum.

**Definition** Eine Teilmenge $B \subset V$ heißt (ungeordnete) *Basis* von $V$ genau dann, wenn sie folgende Bedingungen erfüllt:

(B1) $B$ ist linear unabhängig.

(B2) $B$ ist ein Erzeugendensystem von $V$.

**Satz 11.1** *Ist $B$ eine Basis des $K$-Vektorraums $V$, so lässt sich jeder Vektor $v \in V$ auf genau eine Weise als Linearkombination von Vektoren aus $B$ darstellen.*

*Beweis.* Da $B$ ein Erzeugendensystem von $V$ ist, kann jeder Vektor $v \in V$ auf mindestens eine Art als Linearkombination von Vektoren aus $B$ dargestellt werden. Angenommen, der Vektor $v \in V$ lässt sich auf zwei verschiedenen Arten als Linearkombination von Vektoren aus $B$ schreiben. Indem wir fehlende Vektoren mit dem Koeffizienten 0 in die Summe mit aufnehmen, können wir annehmen, dass in beiden Linearkombinationen die gleichen Vektoren $v_1, \ldots, v_r \in B$ auftreten. Da $B$ linear unabhängig ist, sind $v_1, \ldots, v_r$ linear unabhängig. Damit haben wir einen Widerspruch zu Satz 10.3. $\square$

**Beispiel 11.1** (1) Der Vektorraum $V = K^n$ hat die Basis $\{e_1, e_2, \ldots, e_n\}$. Man nennt diese Basis die *Standardbasis* des $K^n$.

(2) Der Vektorraum $\mathbb{R}[x]$ hat die (unendliche) Basis $\{1, x, x^2, \ldots\}$.

**Satz 11.2** *Für eine Teilmenge $B = \{v_1, \ldots, v_n\}$ eines Vektorraums $V \neq \{0\}$ sind die folgenden Bedingungen äquivalent:*

(i) *$B$ ist eine Basis von $V$.*

(ii) *$B$ ist ein "unverkürzbares" Erzeugendensystem, d.h. $B$ ist ein Erzeugendensystem und für jedes $r \in \{1, \ldots, n\}$ gilt, dass $B \setminus \{v_r\}$ kein Erzeugendensystem mehr ist.*

(iii) *$B$ ist "unverlängerbar" linear unabhängig, d.h. $B$ ist linear unabhängig und für jedes $v \in V$ gilt, dass $B \cup \{v\}$ linear abhängig ist.*

*Beweis.* "(i) $\Rightarrow$ (ii)": Es sei $B$ eine Basis. Angenommen, $B \setminus \{v_r\}$ sei ein Erzeugendensystem von $V$, $1 \leq r \leq n$. Dann ist $v_r$ eine Linearkombination der restlichen Vektoren $v_1, \ldots, v_{r-1}, v_{r+1}, \ldots, v_n$. Also ist $B$ linear abhängig, im Widerspruch dazu, dass $B$ eine Basis ist.

"(ii) $\Rightarrow$ (iii)": Es sei $B$ ein unverkürzbares Erzeugendensystem. Wir zeigen, dass sich jeder Vektor $v \in V$ auf eindeutig bestimmte Weise als Linearkombination der Vektoren $v_1, \ldots, v_n$ schreiben lässt. Aus Satz 10.3 folgt dann, dass $B$ linear unabhängig ist. Angenommen

$$v = \lambda_1 v_1 + \cdots + \lambda_n v_n = \mu_1 v_1 + \cdots + \mu_n v_n,$$

und o.B.d.A. $\lambda_1 \neq \mu_1$. Dann folgt

$$v_1 = \frac{\lambda_2 - \mu_2}{\lambda_1 - \mu_1} v_2 + \cdots + \frac{\lambda_n - \mu_n}{\lambda_1 - \mu_1} v_n,$$

also ist $B \setminus \{v_1\}$ ein Erzeugendensystem von $V$, Widerspruch!

"(iii) $\Rightarrow$ (i)": Es sei $B$ unverlängerbar linear unabhängig. Für jedes $v \in V$ gibt es dann $\lambda_1, \ldots, \lambda_n, \lambda \in K$ mit

$$\lambda_1 v_1 + \cdots + \lambda_n v_n + \lambda v = 0.$$

Da $B$ linear unabhängig ist, muss $\lambda \neq 0$ sein, also

$$v = -\frac{\lambda_1}{\lambda} v_1 - \cdots - \frac{\lambda_n}{\lambda} v_n.$$

Also ist $B$ auch ein Erzeugendensystem von $V$. $\square$

**Korollar 11.1 (Basisauswahlsatz)** *Aus jedem endlichen Erzeugendensystem eines Vektorraums kann man eine Basis auswählen. Insbesondere hat jeder endlich erzeugte Vektorraum eine Basis.*

*Beweis.* Aus einem endlichen Erzeugendensystem kann man in endlich vielen Schritten so viele einzelne Vektoren weglassen, bis es unverkürzbar geworden ist. Nach Satz 11.2 ist das so verkürzte Erzeugendensystem eine Basis von $V$. □

Allgemeiner gilt

**Theorem 11.1** *Jeder Vektorraum hat eine Basis.*

Diesen Satz werden wir nicht beweisen, da der Beweis ungleich schwieriger ist und er Hilfsmittel aus der Mengenlehre, etwa das Zornsche Lemma, benutzt.

**Lemma 11.1 (Austauschlemma)** *Es sei $V$ ein endlich erzeugter $K$-Vektorraum mit der Basis $B = \{v_1, \ldots, v_n\}$ und $w = \lambda_1 v_1 + \ldots + \lambda_n v_n$. Ist $k \in \{1, \ldots, n\}$ mit $\lambda_k \neq 0$, so ist*

$$B' := \{v_1, \ldots, v_{k-1}, w, v_{k+1}, \ldots, v_n\}$$

*wieder eine Basis von $V$. Man kann also $v_k$ gegen $w$ austauschen.*

*Beweis.* O.B.d.A. sei $k = 1$. Es ist zu zeigen: $B' = \{w, v_2, \ldots, v_n\}$ ist eine Basis von $V$.

(B1) $B'$ ist linear unabhängig: Es seien $\mu, \mu_2, \ldots, \mu_n \in K$ mit

$$\mu w + \mu_2 v_2 + \cdots + \mu_n v_n = 0.$$

Einsetzen von $w = \lambda_1 v_1 + \ldots + \lambda_n v_n$ ergibt

$$\mu\lambda_1 v_1 + (\mu\lambda_2 + \mu_2)v_2 + \cdots + (\mu\lambda_n + \mu_n)v_n = 0.$$

Da $B$ linear unabhängig ist , ergibt sich $\mu\lambda_1 = \mu\lambda_2 + \mu_2 = \ldots = \mu\lambda_n + \mu_n = 0$. Wegen $\lambda_1 \neq 0$ folgt $\mu = 0$ und damit $\mu_2 = \ldots = \mu_n = 0$.

(B2) $B'$ ist ein Erzeugendensystem von $V$: Es sei $v \in V$. Dann gilt

$$v = \mu_1 v_1 + \cdots + \mu_n v_n \text{ mit } \mu_1, \ldots, \mu_n \in K.$$

Wegen $\lambda_1 \neq 0$ ist

$$v_1 = \frac{1}{\lambda_1} w - \frac{\lambda_2}{\lambda_1} v_2 - \cdots - \frac{\lambda_n}{\lambda_1} v_n,$$

also
$$v = \frac{\mu_1}{\lambda_1} w + \left(\mu_2 - \frac{\mu_1 \lambda_2}{\lambda_1}\right) v_2 + \cdots + \left(\mu_n - \frac{\mu_1 \lambda_n}{\lambda_1}\right) v_n .$$
□

Durch Iteration erhält man

**Satz 11.3 (Steinitzscher Austauschsatz)** *Es sei $B = \{v_1, \ldots, v_n\}$ eine Basis des Vektorraums $V$ und die Vektoren $w_1, \ldots, w_k \in V$ seien linear unabhängig. Dann gilt $k \leq n$, und bei geeigneter Nummerierung der Vektoren $v_1, \ldots, v_n$ ist auch*
$$B' = \{w_1, \ldots, w_k, v_{k+1}, \ldots, v_n\}$$
*eine Basis von $V$. Man erhält also wieder eine Basis, wenn man $k$ geeignete unter den Vektoren $v_1, \ldots, v_n$ gegen die Vektoren $w_1, \ldots, w_k$ austauscht.*

*Beweis.* Der Beweis erfolgt durch Induktion über $k$. Als Induktionsbeginn kann man den Fall $k = 0$ zulassen, in dem überhaupt keine Vektoren auszutauschen sind und die Behauptung trivial ist.

Induktionsschritt $k - 1 \to k$: Die Vektoren $w_1, \ldots, w_k$ seien linear unabhängig. Dann sind auch die Vektoren $w_1, \ldots, w_{k-1}$ linear unabhängig. Nach Induktionsannahme folgt daher $k - 1 \leq n$ und bei geeigneter Nummerierung ist $B^* = \{w_1, \ldots, w_{k-1}, v_k, \ldots, v_n\}$ eine Basis von $V$. Würde $k - 1 = n$ gelten, so wäre bereits $\{w_1, \ldots, w_{k-1}\}$ eine Basis von $V$ und $w_k$ müsste sich als Linearkombination von $w_1, \ldots, w_{k-1}$ darstellen lassen. Dann wären aber $w_1, \ldots, w_k$ nicht linear unabhängig. Daher gilt sogar $k - 1 < n$. Da $B^*$ eine Basis von $V$ ist, gilt weiter
$$w_k = \lambda_1 w_1 + \cdots + \lambda_{k-1} w_{k-1} + \lambda_k v_k + \ldots + \lambda_n v_n$$
für $\lambda_1, \ldots, \lambda_n \in K$. Wäre $\lambda_k = \ldots = \lambda_n = 0$, so hätten wir einen Widerspruch gegen die lineare Unabhängigkeit von $w_1, \ldots, w_k$. Bei geeigneter Nummerierung kann daher $\lambda_k \neq 0$ angenommen werden. Nach dem Austauschlemma ist dann $B' = \{w_1, \ldots, w_k, v_{k+1}, \ldots, v_n\}$ eine Basis von $V$. □

**Korollar 11.2** *Hat ein $K$-Vektorraum $V$ eine endliche Basis, so ist jede Basis von $V$ endlich.*

*Beweis.* Es sei $\{v_1, \ldots, v_n\}$ eine endliche Basis und $B'$ eine unendliche Basis. Dann gäbe es Elemente $w_1, \ldots, w_{n+1} \in B'$, die linear unabhängig wären. Das widerspricht aber dem Austauschsatz. □

**Korollar 11.3** *Je zwei endliche Basen eines $K$-Vektorraums haben die gleiche Anzahl von Elementen.*

*Beweis.* Sind $\{v_1, \ldots, v_n\}$ und $\{w_1, \ldots, w_k\}$ zwei Basen, so kann man den Austauschsatz zweimal anwenden und erhält $k \leq n$ und $n \leq k$, also $k = n$. □

**Definition** Für einen $K$-Vektorraum $V$ definieren wir

$$\dim_K V := \begin{cases} n, & \text{falls } V \text{ eine Basis mit } n \text{ Elementen besitzt,} \\ \infty, & \text{falls } V \text{ keine endliche Basis besitzt.} \end{cases}$$

Die Zahl $\dim V = \dim_K V$ heißt die *Dimension* von $V$ über $K$.

**Bemerkung 11.1** $\dim V = 0 \Leftrightarrow V = \{0\}$.

**Beispiel 11.2** (1) $\dim K^n = n$. (2) $\dim \mathbb{R}[x] = \infty$.

**Korollar 11.4** *Ist $U \subset V$ ein Unterraum eines endlich dimensionalen $K$-Vektorraums $V$, so ist auch $U$ endlich dimensional und es gilt* $\dim U \leq \dim V$. *Es gilt* $\dim U = \dim V$ *genau dann, wenn $U = V$ gilt.*

*Beweis.* Es sei $\dim V = n$. Wäre $U$ nicht endlich dimensional, so hätte $U$ eine unendliche Basis $B$. Dann gäbe es aber $n+1$ linear unabhängige Vektoren in $V$, ein Widerspruch zum Austauschsatz. Also hat $U$ eine endliche Basis und nach dem Austauschsatz hat sie höchstens $n$ Elemente.

Es sei $\dim U = n = \dim V$ und $B = \{u_1, \ldots, u_n\}$ eine Basis von $U$. Ist $U \neq V$, so gibt es ein $v \in V \setminus U$ und $u_1, \ldots, u_n, v$ sind linear unabhängig im Widerspruch zum Austauschsatz. □

**Satz 11.4 (Basisergänzungssatz)** *Es sei $V$ ein endlich erzeugter $K$-Vektorraum und $w_1, \ldots, w_r \in V$ linear unabhängige Vektoren. Dann kann man diese Vektoren zu einer Basis ergänzen, d.h. man kann $w_{r+1}, \ldots, w_n \in V$ finden, so dass*

$$B = \{w_1, \ldots, w_r, w_{r+1}, \ldots, w_n\}$$

*eine Basis von $V$ ist.*

*Beweis.* Es sei $\{v_1, \ldots, v_m\}$ ein Erzeugendensystem von $V$. Nach dem Basisauswahlsatz kann man daraus eine Basis auswählen. Bei geeigneter Nummerierung können wir annehmen, dass $\{v_1, \ldots, v_n\}$ mit $n \leq m$ eine Basis ist. Aus dem Austauschsatz folgt dann, dass bei geeigneter Nummerierung $\{w_1, \ldots, w_r, v_{r+1}, \ldots, v_n\}$ eine Basis von $V$ ist. □

# 12 Rang einer Matrix

Wir betrachten nun wieder Matrizen. Es sei $K$ ein Körper. Wir betrachten Matrizen über $K$. Eine solche Matrix ist von der Form

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}, \quad a_{ij} \in K.$$

Die Menge aller $m \times n$-Matrizen über dem Körper $K$ bezeichnen wir mit $\mathrm{Mat}(m, n; K)$.

Die Vektoren

$$z_i := (a_{i1}, a_{i2}, \ldots, a_{in}), \quad i = 1, \ldots, m,$$

heißen die *Zeilenvektoren von* $A$ und die Vektoren

$$s_j := \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix}, \quad j = 1, \ldots, n,$$

heißen die *Spaltenvektoren von* $A$.

**Definition** Es sei $A$ eine $m \times n$-Matrix über $K$. Der von den Zeilenvektoren aufgespannte Unterraum von $K^n$, also $\mathrm{Span}(z_1, \ldots, z_m)$, heißt der *Zeilenraum von* $A$. Der von den Spaltenvektoren aufgespannte Unterraum von $K^m$, also $\mathrm{Span}(s_1, \ldots, s_n)$, heißt der *Spaltenraum von* $A$.

**Definition** Es sei $A$ eine $m \times n$-Matrix über $K$. Die Dimension des Zeilenraums von $A$ bezeichnen wir als den *Zeilenrang* von $A$. Die Dimension des Spaltenraums von $A$ bezeichnen wir als den *Spaltenrang* von $A$.

Unser Ziel ist es, zu zeigen, dass der Zeilenrang einer Matrix gleich dem Spaltenrang ist. Dazu untersuchen wir, wie sich Zeilen- und Spaltenraum unter elementaren Zeilenoperationen der Matrix verändern. Wir haben bisher solche Zeilenoperationen nur für reelle Matrizen betrachtet. Ebenso haben wir bisher nur lineare Gleichungssysteme über $\mathbb{R}$ betrachtet. Alles, was wir für diesen Spezialfall gesagt haben, bleibt aber auch über einem beliebigen Körper gültig, da wir nur die Körpereigenschaften von $\mathbb{R}$ benutzt haben.

**Satz 12.1** *Elementare Zeilenoperationen ändern den Zeilenraum einer Matrix nicht.*

*Beweis.* Es seien
$$z_1, \dots, z_m$$
die Zeilenvektoren von $A$.

Es ist klar, dass die Vertauschung von zwei Zeilen (Z1) den Zeilenraum nicht ändert.

Als Zeilenoperation vom Typ (Z2) betrachten wir die Addition des $\lambda$-fachen der $k$-ten Zeile zur $i$-ten Zeile. Wenden wir diese Operation auf die Matrix $A$ an, so erhalten wir eine Matrix $B$ mit den Zeilenvektoren
$$z_1, \dots, z_{i-1}, z_i + \lambda z_k, z_{i+1}, \dots, z_m.$$
Da der Zeilenraum von $A$ mit $z_1, \dots, z_m$ auch alle Linearkombinationen dieser Vektoren enthält, enthält er auch den Zeilenraum von $B$.

Da man $A$ durch Anwenden der inversen Zeilenoperation aus $B$ enthält, folgt, dass der Zeilenraum von $A$ im Zeilenraum von $B$ enthalten ist.

Entsprechend argumentiert man auch im Fall einer Zeilenoperation (Z3).
□

Der Spaltenraum einer Matrix kann sich allerdings bei elementaren Zeilenoperationen ändern.

**Beispiel 12.1** Der Spaltenraum der Matrix
$$A = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \in \mathrm{Mat}(2, 2; \mathbb{R})$$
ist die Gerade $y = 0$ im $\mathbb{R}^2$, der Spaltenraum von
$$B = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \in \mathrm{Mat}(2, 2; \mathbb{R})$$
ist die Gerade $y = x$.

Es gilt aber

**Satz 12.2** *Elementare Zeilenoperationen ändern den Spaltenrang einer Matrix nicht.*

*Beweis.* Es sei $\{s_{i_1}, \dots, s_{i_k}\}$ eine linear unabhängige Menge von Spaltenvektoren von $A$. Das bedeutet, dass das lineare Gleichungssystem
$$\lambda_1 s_{i_1} + \cdots + \lambda_k s_{i_k} = 0$$

nur die triviale Lösung $\lambda_1 = \ldots = \lambda_k = 0$ hat. Die Matrix $\widetilde{A}$ sei aus $A$ durch endlich viele elementare Zeilenoperationen entstanden. Die entsprechenden Spaltenvektoren von $\widetilde{A}$ bezeichnen wir mit $\widetilde{s}_{i_1}, \ldots, \widetilde{s}_{i_k}$. Nach Satz 5.2 verändern Zeilenoperationen vom Typ (Z1) und (Z2) die Lösungsmenge eines linearen Gleichungssystems nicht. Das gleiche gilt für eine Zeilenoperation vom Typ (Z3), da dies ja nur die Multiplikation einer Gleichung mit einer Zahl $\lambda$ bedeutet. Also hat auch das Gleichungssystem

$$\lambda_1 \widetilde{s}_{i_1} + \cdots + \lambda_k \widetilde{s}_{i_k} = x$$

nur die triviale Lösung. Damit ist auch die Menge $\{\widetilde{s}_{i_1}, \ldots, \widetilde{s}_{i_k}\}$ linear unabhängig. □

**Satz 12.3** *Der Zeilenrang einer Matrix $A$ ist gleich dem Spaltenrang.*

*Beweis.* Nach Satz 5.3 können wir die Matrix $A$ durch elementare Zeilenumformungen in eine Matrix $\widetilde{A}$ in Zeilenstufenform überführen. Nach Satz 12.1 und Satz 12.2 ändern elementare Zeilenumformungen weder den Zeilenrang noch den Spaltenrang. Wir nehmen an, dass die Pivots die Einträge

$$\widetilde{a}_{1j_1}, \widetilde{a}_{2j_2}, \ldots, \widetilde{a}_{rj_r}$$

der Matrix $\widetilde{A}$ sind. Dann sieht man leicht, dass die ersten $r$ Zeilenvektoren eine Basis des Zeilenraums von $\widetilde{A}$ bilden. Also ist der Zeilenrang von $\widetilde{A}$ und damit auch von $A$ gleich $r$. Die Spaltenvektoren $\widetilde{s}_{j_1}, \widetilde{s}_{j_2}, \ldots, \widetilde{s}_{j_r}$ bilden eine Basis des Spaltenraums. Also ist auch der Spaltenrang von $\widetilde{A}$ und damit auch von $A$ gleich $r$. □

Satz 12.3 führt zu der folgenden Definition.

**Definition** Der Zeilen- oder Spaltenrang einer Matrix $A$ wird auch *Rang* von A, in Zeichen $\operatorname{Rang} A$, genannt.

**Beispiel 12.2** Der Rang der Matrix

$$A = \begin{pmatrix} 1 & 3 & 4 & 2 & -2 \\ 0 & 2 & -1 & 2 & 0 \\ -1 & 1 & -4 & -4 & 2 \\ 1 & 1 & 3 & 6 & -2 \end{pmatrix} \in \operatorname{Mat}(4, 5; \mathbb{R})$$

ist gleich 3. Denn nach Beispiel 5.3 können wir $A$ durch elementare Zeilenumformungen in die folgende Matrix in Zeilenstufenform überführen

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

Die Anzahl der Pivots ist 3, also hat diese Matrix den Rang 3.

**Satz 12.4** *Für jede Matrix A gilt:*

$$\operatorname{Rang} A = \operatorname{Rang} A^T.$$

*Beweis.*

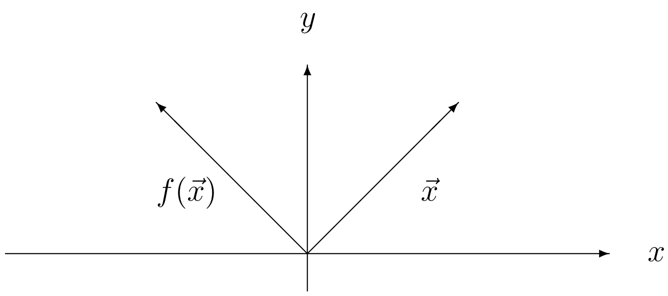
$$\operatorname{Rang} A = \operatorname{Zeilenrang} A = \operatorname{Spaltenrang} A^T = \operatorname{Rang} A^T.$$

□

# 13 Lineare Abbildungen

Wir wollen nun lineare Abbildungen einführen. Dazu betrachten wir zunächst einige Beispiele.

**Beispiel 13.1** Wir betrachten Abbildungen der Ebene $\mathbb{R}^2$. Es sei $f : \mathbb{R}^2 \to \mathbb{R}^2$ die Abbildung, die jedem Vektor $\vec{x} \in \mathbb{R}^2$ sein Spiegelbild bezüglich der $y$-Achse zuordnet.



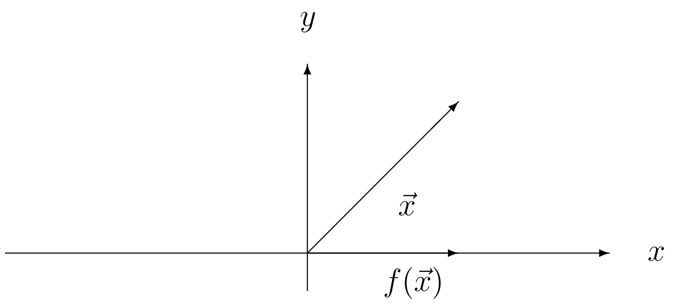

Das Bild $f(\vec{x})$ ist wieder ein Vektor. Wir bezeichnen die Komponenten dieses Vektors mit $f_1(\vec{x})$ und $f_2(\vec{x})$. Es gilt

$$\begin{pmatrix} f_1(\vec{x}) \\ f_2(\vec{x}) \end{pmatrix} = \begin{pmatrix} f_1(x,y) \\ f_2(x,y) \end{pmatrix} = \begin{pmatrix} -x \\ y \end{pmatrix}.$$

Dies können wir auch so ausdrücken:

$$\begin{pmatrix} f_1(x,y) \\ f_2(x,y) \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

**Beispiel 13.2** Wir betrachten die Abbildung $f : \mathbb{R}^2 \to \mathbb{R}^2$, die jedem Vektor $\vec{x}$ seine Orthogonalprojektion auf die $x$-Achse zuordnet.



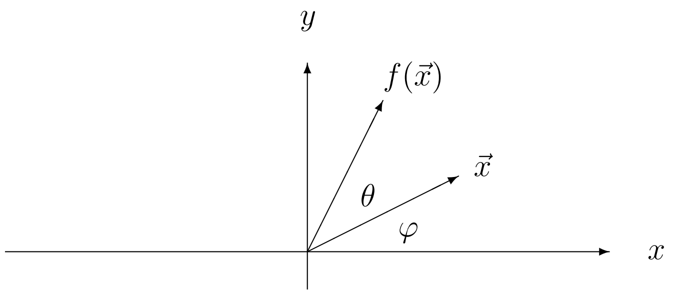

Es gilt

$$\begin{pmatrix} f_1(x,y) \\ f_2(x,y) \end{pmatrix} = \begin{pmatrix} x \\ 0 \end{pmatrix}.$$

Mit Hilfe einer Matrix können wir dies so schreiben:

$$\begin{pmatrix} f_1(x,y) \\ f_2(x,y) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

**Beispiel 13.3** Wir betrachten nun die Abbildung $f : \mathbb{R}^2 \to \mathbb{R}^2$, die einen Vektor $\vec{x}$ um den Winkel $\theta$ dreht. Um eine Beschreibung für $f_1(x,y)$ und $f_2(x,y)$ abzuleiten, betrachten wir eine Drehung um einen positiven Winkel $\theta$. Es sei $\varphi$ der Winkel zwischen dem Vektor $\vec{x}$ und der positiven $x$-Achse und $r$ die Länge von $\vec{x}$.




Dann gilt

$$x = r\cos\varphi, \quad y = r\sin\varphi$$

und

$$f_1(x,y) = r\cos(\varphi+\theta), \quad f_2(x,y) = r\sin(\varphi+\theta).$$

Durch Anwendung der Additionstheoreme von sin und cos ergibt sich hieraus

$$\begin{aligned} f_1(x,y) &= r\cos\theta\cos\varphi - r\sin\theta\sin\varphi \\ f_2(x,y) &= r\sin\theta\cos\varphi + r\cos\theta\sin\varphi, \end{aligned}$$

und schließlich

$$\begin{aligned} f_1(x,y) &= (\cos\theta)x - (\sin\theta)y \\ f_2(x,y) &= (\sin\theta)x + (\cos\theta)y. \end{aligned}$$

In Matrizenschreibweise lautet dies

$$\begin{pmatrix} f_1(x,y) \\ f_2(x,y) \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

**Beispiel 13.4** Nun betrachten wir auch Abbildungen des Raumes. Es sei $f : \mathbb{R}^3 \to \mathbb{R}^3$ eine Drehung um die $x$-Achse um den Winkel $\theta$. Wie im vorigen Beispiel leitet man her, dass diese Abbildung durch die folgende Vorschrift gegeben wird:

$$\begin{pmatrix} f_1(x,y,z) \\ f_2(x,y,z) \\ f_3(x,y,z) \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

**Beispiel 13.5** Die Beispiele von Abbildungen der Ebene und des Raumes sind Spezialfälle der folgenden Konstruktion. Einer $m \times n$-Matrix

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

über einem Körper $K$ kann man wie folgt eine Abbildung $f : K^n \to K^m$ zuordnen: Wir definieren

$$f(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}) = \begin{pmatrix} a_{11}x_1 + \cdots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n \end{pmatrix},$$

oder anders ausgedrückt

$$f(x) = Ax.$$

Nach den Rechenregeln für Matrizen gilt

$$A(x+y) = Ax + Ay, \quad A(\lambda x) = \lambda Ax.$$

Also hat diese Abbildung die Eigenschaften

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

Eine Abbildung mit den Eigenschaften des letzten Beispiels heißt *linear*. Diesen Begriff wollen wir nun ganz allgemein definieren.

**Definition** Es seien $V$ und $W$ $K$-Vektorräume. Eine Abbildung $f : V \to W$ heißt *linear* genau dann, wenn gilt

(L1) $f(v + v') = f(v) + f(v')$ für alle $v, v' \in V$,

(L2) $f(\lambda v) = \lambda f(v)$ für alle $v \in V$ und $\lambda \in K$.

**Bemerkung 13.1** Die beiden Bedingungen (L1) und (L2) kann man auch zu der Bedingung

(L) $f(\lambda v + \mu v') = \lambda f(v) + \mu f(v')$ für alle $v, v' \in V$ und $\lambda, \mu \in K$.

zusammenfassen. Man überlegt sich leicht, dass die beiden Bedingungen (L1) und (L2) zusammen äquivalent zu der Bedingung (L) sind.

**Beispiel 13.6** Die Abbildungen von Beispiel 13.5 und damit aus allen vorherigen Beispielen sind linear. Die Funktion $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x^2$, ist nicht linear, denn es gilt

$$f(\lambda x) = (\lambda x)^2 = \lambda^2 x^2 \text{ für alle } \lambda \in \mathbb{R}$$

und etwa für $\lambda = 2$ ist $\lambda^2 \neq \lambda$.

**Beispiel 13.7** Es sei $[a, b] \subset \mathbb{R}$ ein Intervall und $C^1[a, b]$ und $C[a, b]$ die in Beispiel 9.3 eingeführten $\mathbb{R}$-Vektorräume. Dann ist die Abbildung

$$\begin{array}{rccc} \varphi : & C^1[a,b] & \longrightarrow & C[a,b] \\ & f & \longmapsto & f' \end{array}$$

linear.

**Definition** Man nennt eine lineare Abbildung $f : V \to W$ einen

- *Monomorphismus*, falls $f$ injektiv ist,
- *Epimorphismus*, falls $f$ surjektiv ist,
- *Isomorphismus*, falls es eine lineare Abbildung $g : W \to V$ gibt mit $f \circ g = \mathrm{id}_W$ und $g \circ f = \mathrm{id}_V$.
- *Endomorphismus*, falls $V = W$ ist,
- *Automorphismus*, falls $V = W$ und $f$ ein Isomorphismus ist.

Wieder gilt:

**Lemma 13.1** *Eine lineare Abbildung $f : V \to W$ ist genau dann ein Isomorphismus, wenn $f$ bijektiv ist.*

*Beweis.* Der Beweis geht analog zum Beweis von Lemma 8.1. □

Wir geben nun einige Eigenschaften von linearen Abbildungen an.

**Satz 13.1** *Es sei $f : V \to W$ eine lineare Abbildung. Dann gilt für alle Vektoren $v, v' \in V$:*

(a) $f(0) = 0$.

(b) $f(-v) = -f(v)$.

(c) $f(v - v') = f(v) - f(v')$.

*Beweis.* (a) $f(0) = f(0v) \overset{(L2)}{=} 0f(v) = 0$.

(b) $f(-v) = f((-1)v) \overset{(L2)}{=} (-1)f(v) = -f(v)$.

(c) $f(v - v') = f(v + (-1)v') \overset{(L)}{=} f(v) + (-1)f(v') = f(v) - f(v')$. □

**Satz 13.2** *Es sei $V$ ein endlich dimensionaler $K$-Vektorraum. Dann ist eine lineare Abbildung $f : V \to W$ bereits durch die Bilder der Vektoren einer Basis vollständig festgelegt.*

*Beweis.* Es sei $S = \{v_1, \ldots, v_n\}$ eine Basis von $V$. Dann lässt sich jeder Vektor $v \in V$ als Linearkombination bezüglich dieser Basis schreiben:

$$v = \lambda_1 v_1 + \cdots + \lambda_n v_n.$$

Mit (L) erhalten wir

$$\begin{aligned} f(v) &= f(\lambda_1 v_1 + \cdots + \lambda_n v_n) \\ &= \lambda_1 f(v_1) + \cdots + \lambda_n f(v_n). \end{aligned}$$

□

**Beispiel 13.8** Es sei $A$ eine $m \times n$-Matrix über $K$ und $f : K^n \to K^m$ die lineare Abbildung $f(x) = Ax$. Es sei $\{e_1, \ldots, e_n\}$ die Standardbasis von $K^n$. Dann ist $f$ durch die Bilder $f(e_i)$ der Standardbasisvektoren festgelegt. Aber

$$Ae_1, \ldots, Ae_n$$

sind gerade die Spalten der Matrix $A$ in dieser Reihenfolge. Man merke sich also: *Die Spalten der Matrix $A$ sind die Bilder der Standardbasisvektoren.*

**Satz 13.3** *Sind $U$, $V$, $W$ $K$-Vektorräume und $f : U \to V$, $g : V \to W$ lineare Abbildungen, so ist auch die Abbildung*

$$g \circ f : U \to W$$

*linear.*

*Beweis.* Für $u, u' \in U$ gilt

$$\begin{aligned}
(g \circ f)(u + u') &= g(f(u + u')) \\
&= g(f(u) + f(u')) \quad \text{(da } f \text{ linear)} \\
&= g(f(u)) + g(f(u')) \quad \text{(da } g \text{ linear)} \\
&= (g \circ f)(u) + (g \circ f)(u').
\end{aligned}$$

Analog zeigt man $(g \circ f)(\lambda u) = \lambda(g \circ f)(u)$ für $\lambda \in K$. □

# 14 Kern und Bild

Wir versuchen nun, die Geometrie linearer Abbildungen besser zu verstehen. Dazu führen wir die Begriffe Kern und Bild ein.

**Definition** Für eine lineare Abbildung $f : V \to W$ definieren wir

$$\begin{aligned}
\operatorname{Im} f &:= f(V) := \{f(v) \mid v \in V\}, \quad \textit{Bild} \text{ von } f, \\
\operatorname{Ker} f &:= \{v \in V \mid f(v) = 0\}, \quad \textit{Kern} \text{ von } f.
\end{aligned}$$

**Satz 14.1** *Für eine lineare Abbildung $f : V \to W$ gilt:*

(a) $\operatorname{Im} f$ *ist ein Unterraum von $W$.*

(b) $\operatorname{Ker} f$ *ist ein Unterraum von $V$.*

*Beweis.* (a) Wegen $f(0) = 0$ enthält $\operatorname{Im} f$ mindestens den Nullvektor. Es seien $w$, $w'$ Vektoren aus $\operatorname{Im} f$. Dann gibt es $v$, $v' \in V$ mit $f(v) = w$, $f(v') = w'$. Dann gilt

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

Also ist auch $w + w' \in \operatorname{Im} f$.

Analog zeigt man, dass mit $w \in \operatorname{Im} f$ und $\lambda \in K$ auch $\lambda w \in \operatorname{Im} f$ ist.

(b) Wegen $f(0) = 0$ enthält $\operatorname{Ker} f$ mindestens den Nullvektor. Es seien $v$, $v' \in \operatorname{Ker} f$. Dann gilt

$$f(v + v') = f(v) + f(v') = 0 + 0 = 0.$$

Also ist auch $v + v' \in \operatorname{Ker} f$.

Analog zeigt man, dass mit $v \in \operatorname{Ker} f$ und $\lambda \in K$ auch $\lambda v \in \operatorname{Ker} f$ ist. □

**Satz 14.2 (Dimensionsformel)** *Es sei $f : V \to W$ eine lineare Abbildung von einem endlich dimensionalen $K$-Vektorraum $V$ in einen $K$-Vektorraum $W$. Dann gilt*

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

*Beweis.* Es sei $\{v_1, \ldots, v_k\}$ eine Basis von $\operatorname{Ker} f$ und $\{w_1, \ldots, w_r\}$ eine Basis von $\operatorname{Im} f$. Es seien $u_1, \ldots, u_r$ beliebige Vektoren aus $V$ mit $f(u_1) = w_1, \ldots, f(u_r) = w_r$. Wir zeigen

**Behauptung** $S = \{u_1, \ldots, u_r, v_1, \ldots, v_k\}$ ist eine Basis von $V$.

(a) Es sei $v \in V$. Dann gilt

$$f(v) = \mu_1 w_1 + \cdots + \mu_r w_r.$$

Es sei

$$v' = \mu_1 u_1 + \cdots + \mu_r u_r.$$

Dann gilt

$$f(v') = \mu_1 f(u_1) + \cdots + \mu_r f(u_r) = f(v).$$

Also ist $v - v' \in \operatorname{Ker} f$. Das bedeutet, dass es $\lambda_1, \ldots, \lambda_k \in K$ gibt mit

$$v - v' = \lambda_1 v_1 + \cdots + \lambda_k v_k.$$

Dann folgt aber

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

Also ist $S$ ein Erzeugendensystem von $V$.

(b) Um zu zeigen, dass $S$ linear unabhängig ist, betrachten wir die Gleichung

$$\mu_1 u_1 + \cdots + \mu_r u_r + \lambda_1 v_1 + \cdots + \lambda_k v_k = 0. \tag{5}$$

Es gilt

$$0 = f(\mu_1 u_1 + \cdots + \mu_r u_r + \lambda_1 v_1 + \cdots + \lambda_k v_k) = \mu_1 w_1 + \cdots + \mu_r w_r.$$

Da $w_1, \ldots, w_r$ linear unabhängig sind, folgt $\mu_1 = \cdots = \mu_r = 0$. In (5) eingesetzt ergibt sich

$$\lambda_1 v_1 + \cdots + \lambda_k v_k = 0.$$

Da $v_1, \ldots, v_k$ linear unabhängig sind, folgt daraus $\lambda_1 = \cdots = \lambda_k = 0$. Daraus folgt die Behauptung.

Aus der Behauptung folgt $r + k = \dim V$ und daraus die Behauptung des Satzes. $\square$

**Definition** Es sei $f : V \to W$ eine lineare Abbildung. Der *Rang* von $F$, in Zeichen $\operatorname{Rang} f$, ist die Dimension von $\operatorname{Im} f$.

**Beispiel 14.1** Es sei $A$ eine $m \times n$-Matrix über $K$. Für den Rang der linearen Abbildung $f : K^n \to K^m$ mit $f(x) = Ax$ gilt $\operatorname{Rang} f = \operatorname{Rang} A$.

**Satz 14.3** *Eine lineare Abbildung $f : V \to W$ ist genau dann ein Monomorphismus, wenn der Kern von $f$ nur den Nullvektor enthält.*

*Beweis.*

"$\Rightarrow$": Es sei $v \in V$, $v \neq 0$. Da $f(0) = 0$ und $f$ injektiv ist, gilt $f(v) \neq 0$. Also enthält der Kern von $f$ nur den Nullvektor.

"$\Leftarrow$": Es seien $v, v' \in V$ mit $v \neq v'$. Wir müssen zeigen, dass $f(v) \neq f(v')$ gilt. Angenommen, $f(v) = f(v')$. Dann gilt

$$0 = f(v) - f(v') = f(v - v').$$

Also liegt $v - v'$ im Kern von $f$. Da der Kern von $f$ nur den Nullvektor enthält, folgt $v - v' = 0$, ein Widerspruch. □

**Satz 14.4** *Es seien $V$, $W$ endlich dimensionale $K$-Vektorräume. Eine lineare Abbildung $f : V \to W$ ist genau dann ein Epimorphismus, wenn* $\operatorname{Rang} f = \dim W$ *gilt.*

*Beweis.* Die lineare Abbildung $f$ ist genau dann surjektiv, wenn $\operatorname{Im} f = W$ gilt. Da $\operatorname{Im} f$ ein Unterraum von $W$ ist, ist dies nach Korollar 11.4 genau dann der Fall, wenn $\operatorname{Rang} f = \dim \operatorname{Im} f = \dim W$ gilt. □

**Satz 14.5** *Für eine lineare Abbildung $f : V \to W$ zwischen endlich dimensionalen $K$-Vektorräumen* gleicher Dimension *sind die folgenden Bedingungen äquivalent:*

(a) *$f$ ist ein Monomorphismus.*

(b) *$f$ ist ein Epimorphismus.*

(c) *$f$ ist ein Isomorphismus.*

*Beweis.* Dies folgt sofort aus Satz 14.3, Satz 14.4 und der Dimensionsformel (Satz 14.2). □

# 15 Lineare Gleichungssysteme

Wir wenden nun unsere bisher entwickelte Theorie auf lineare Gleichungssysteme an und leiten die wichtigsten Sätze über Lösbarkeit und Lösungen von linearen Gleichungssystemen her.

Es sei $A$ eine $m \times n$-Matrix über $K$ und $b \in K^m$. Wir betrachten nun das Gleichungssystem $Ax = b$ für $x \in K^n$. Wie wir schon in §12 bemerkt haben, bleibt alles, was wir für den Spezialfall $K = \mathbb{R}$ gesagt haben, auch über einem beliebigen Körper gültig, da wir bei der Behandlung von linearen Gleichungssystemen nur die Körpereigenschaften von $\mathbb{R}$ benutzt haben.

**Satz 15.1 (Lösbarkeitsentscheidung)** *Ein lineares Gleichungssystem $Ax = b$ ist genau dann lösbar, wenn der Rang von $A$ gleich dem Rang der erweiterten Koeffizientenmatrix $(A, b)$ ist.*

*Beweis.* Sind $s_1, \ldots, s_n$ die Spaltenvektoren von $A$, so können wir das Gleichungssystem $Ax = b$ auch so schreiben:

$$x_1 s_1 + \cdots + x_n s_n = b.$$

Das bedeutet aber, dass das Gleichungssystem $Ax = b$ genau dann lösbar ist, wenn $b$ eine Linearkombination der Spaltenvektoren von $A$ ist, also im Spaltenraum von $A$ liegt. Also ist das Gleichungssystem $Ax = b$ genau dann lösbar, wenn die Spaltenräume von $A$ und $(A, b)$ übereinstimmen. Das ist aber genau dann der Fall, wenn die Ränge der entsprechenden Matrizen übereinstimmen. □

Entscheidend für die weitere Theorie ist der Zusammenhang zu der durch $A$ bestimmten linearen Abbildung

$$f: \begin{array}{ccc} K^n & \longrightarrow & K^m \\ x & \longmapsto & Ax \end{array}.$$

Es gilt

$$\mathcal{L}(A, b) = f^{-1}(b), \quad \text{insbesondere } \mathcal{L}(A, 0) = \operatorname{Ker} f.$$

Es sei

$$r = \operatorname{Rang} f = \operatorname{Rang} A = \text{Zeilenrang } A.$$

Aus Satz 14.1 und Satz 14.2 folgt:

**Satz 15.2 (Lösung des homogenen Systems)** *Es sei $A$ eine $m \times n$-Matrix über $K$ mit* $\operatorname{Rang} A = r$. *Dann gilt für den Lösungsraum des homogenen linearen Gleichungssystem $Ax = 0$: $\mathcal{L}(A, 0)$ ist ein Unterraum von $K^n$ der Dimension $n - r$.*

Wir leiten nun eine Beziehung zwischen der Lösungsmenge eines inhomogenen Gleichungssystems $Ax = b$ und dem Lösungsraum des zugehörigen homogenen Gleichungssystems $Ax = 0$ her.

**Satz 15.3 (Lösung des inhomogenen Systems)** *Es sei $x_0$ eine Lösung des inhomogenen Gleichungssystems $Ax = b$ und $\{v_1, \ldots, v_r\}$ eine Basis von $\mathcal{L}(A, 0)$. Ein Vektor $x$ ist genau dann eine Lösung von $Ax = b$, wenn er die Darstellung*

$$x = x_0 + \lambda_1 v_1 + \cdots + \lambda_r v_r$$

*mit Skalaren $\lambda_1, \ldots, \lambda_r$ besitzt.*

*Beweis.*

"$\Rightarrow$": Es sei $x$ eine Lösung von $Ax = b$. Dann gilt

$$Ax = b.$$

Da $x_0$ ebenfalls eine Lösung von $Ax = b$ ist, gilt

$$Ax_0 = b.$$

Subtraktion der beiden Gleichungen liefert

$$\begin{aligned} & Ax - Ax_0 = x \\ \Leftrightarrow \quad & A(x - x_0) = 0. \end{aligned}$$

Also ist $x - x_0$ eine Lösung des zugehörigen homogenen Gleichungssystems $Ax = 0$. Da $\{v_1, \ldots, v_r\}$ eine Basis des Lösungsraums von $Ax = 0$ ist, existieren Skalare $\lambda_1, \ldots, \lambda_r$ mit

$$x - x_0 = \lambda_1 v_1 + \cdots + \lambda_r v_r.$$

Daraus folgt

$$x = x_0 + \lambda_1 v_1 + \cdots + \lambda_r v_r.$$

"$\Leftarrow$": Es sei

$$x = x_0 + \lambda_1 v_1 + \cdots + \lambda_r v_r.$$

Dann gilt

$$\begin{aligned} Ax &= A(x_0 + \lambda_1 v_1 + \cdots + \lambda_r v_r) \\ &= Ax_0 + \lambda_1 A(v_1) + \cdots + \lambda_r A(v_r) \\ &= b. \end{aligned}$$

Also ist $x$ eine Lösung von $Ax = 0$. $\square$

**Bemerkung 15.1** Der Vektor $x_0$ in Satz 15.3 heißt eine *spezielle* Lösung des linearen Gleichungssystems $Ax = b$. Der Ausdruck

$$\lambda_1 v_1 + \cdots + \lambda_r v_r$$

wird als *allgemeine* Lösung des homogenen Systems $Ax = 0$ bezeichnet. Satz 15.3 besagt, dass man *die allgemeine Lösung des inhomogenen Gleichungssystems $Ax = b$ erhält, indem man zu einer speziellen Lösung von $Ax = b$ die allgemeine Lösung des zugehörigen homogenen Gleichungssystems $Ax = 0$ addiert.*

Geometrisch bedeutet dies, dass man die Lösungsmenge des inhomogenen Gleichungssystems $Ax = b$ erhält, indem man den Lösungsraum des zugehörigen homogenen Systems $Ax = 0$ um den Vektor $x_0$ verschiebt. Der Lösungsraum von $Ax = 0$ ist nach Beispiel 9.5 ein Unterraum von $K^n$. Ist $b \neq 0$, so ist auch $x_0 \neq 0$ und die Lösungsmenge $\mathcal{L}(A, b)$ enthält nicht die Null. Die Lösungsmenge eines inhomogenen linearen Gleichungssystems ist also im Allgemeinen kein Unterraum des $K^n$. Im $\mathbb{R}^3$ können Lösungsmengen von $Ax = b$ zum Beispiel Geraden oder Ebenen, die nicht durch den Nullpunkt gehen, sein.

**Beispiel 15.1** Nach Beispiel 5.2 hat das inhomogene lineare Gleichungssystem

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

eine spezielle Lösung

$$x_0 = \begin{pmatrix} 6 \\ 0 \\ -1 \\ 0 \\ 0 \end{pmatrix}.$$

Wir bestimmen nun eine Basis und die Dimension des Lösungsraums des zugehörigen homogenen Gleichungssystems:

$$\begin{aligned} x_1 + 3x_2 + 4x_3 + 2x_4 - 2x_5 &= 0 \\ 2x_2 - x_3 + 2x_4 &= 0 \\ 2x_3 - 6x_4 &= 0. \end{aligned}$$

Wie in Beispiel 5.2 erhalten wir als allgemeine Lösung:

$$x_1 = -\frac{31}{2}\lambda_1 + 2\lambda_2,\ x_2 = \frac{1}{2}\lambda_1,\ x_3 = 3\lambda_1,\ x_4 = \lambda_1,\ x_5 = \lambda_2.$$

Damit erhalten wir als allgemeinen Lösungsvektor

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} -\frac{31}{2}\lambda_1 + 2\lambda_2 \\ \frac{1}{2}\lambda_1 \\ 3\lambda_1 \\ \lambda_1 \\ \lambda_2 \end{pmatrix} = \lambda_1 \begin{pmatrix} -\frac{31}{2} \\ \frac{1}{2} \\ 3 \\ 1 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 2 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

Also spannen die Vektoren

$$v_1 := \begin{pmatrix} -\frac{31}{2} \\ \frac{1}{2} \\ 3 \\ 1 \\ 0 \end{pmatrix}, \; v_2 := \begin{pmatrix} 2 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

den Lösungsraum auf. Sie sind linear unabhängig. Damit ist $S := \{v_1, v_2\}$ eine Basis des Lösungsraums. Der Lösungsraum hat also die Dimension 2. Also hat die allgemeine Lösung des inhomogenen Systems die Gestalt

$$x = \begin{pmatrix} 6 \\ 0 \\ -1 \\ 0 \\ 0 \end{pmatrix} + \lambda_1 \begin{pmatrix} -\frac{31}{2} \\ \frac{1}{2} \\ 3 \\ 1 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 2 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

Wir betrachten nun noch einige Folgerungen aus diesen Sätzen.

**Satz 15.4** *Es sei $A$ eine $m \times n$-Matrix über $K$. Dann gilt:*

(a) *$Ax = b$ ist genau dann für alle $b \in K^m$ lösbar, wenn* $\operatorname{Rang} A = m$.

(b) *$Ax = 0$ besitzt genau dann nur die triviale Lösung, wenn* $\operatorname{Rang} A = n$.

*Beweis.* Wir betrachten wieder die zu $A$ gehörige lineare Abbildung

$$\begin{array}{rccc} f: & K^n & \longrightarrow & K^m \\ & x & \longmapsto & Ax \end{array}.$$

Dann folgt (a) aus Satz 14.4 und (b) aus Satz 14.3. □

**Satz 15.5** *Es sei $A$ eine $n \times n$-Matrix über $K$. Dann ist $A$ genau dann invertierbar, wenn* $\operatorname{Rang} A = n$ *gilt.*

*Beweis.* Nach Satz 6.6 ist $A$ genau dann invertierbar, wenn das Gleichungssystem $Ax = 0$ nur die triviale Lösung besitzt. Nach Satz 15.4 (b) ist dies dann und nur dann der Fall, wenn $\operatorname{Rang} A = n$ gilt. □

# 16 Lineare Abbildungen und Matrizen

Wir wollen nun den Zusammenhang zwischen linearen Abbildungen und Matrizen studieren. Es sei $V$ ein $K$-Vektorraum der Dimension $n$.

**Vereinbarung** Im Folgenden verstehen wir unter einer Basis von $V$ immer eine *geordnete Basis*, d.h. ein $n$-Tupel

$$B = (v_1, \ldots, v_n).$$

Ist $B = (v_1, \ldots, v_n)$ eine Basis von $V$, dann besitzt jeder Vektor $v \in V$ nach Satz 11.1 eine eindeutige Darstellung

$$v = x_1 v_1 + \cdots x_n v_n.$$

**Definition** Der Vektor

$$q_B(v) = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in K^n$$

heißt der *Koordinatenvektor von $v$ bezüglich der Basis $B$.*

Wir erhalten eine Abbildung

$$\begin{array}{rccc} q_B: & V & \longrightarrow & K^n \\ & v & \longmapsto & \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \end{array}.$$

Es gilt $q_B(v_i) = e_i$ und die Abbildung $q_B$ ist ein Isomorphismus von Vektorräumen.

Es seien nun $V$ und $W$ endlich dimensionale $K$-Vektorräume und $f : V \to W$ eine lineare Abbildung. Wir wollen dieser Abbildung eine Matrix zuordnen. Dies geschieht wie folgt:

(1) Wir wählen eine Basis $B = (v_1, \ldots, v_n)$ von $V$ und eine Basis $B' = (w_1, \ldots w_m)$ von $W$.

(2) Wir stellen die Bilder $f(v_j)$ der Vektoren der Basis $B$ von $V$ in der Basis $B'$ von $W$ dar:

$$f(v_j) = \sum_{i=1}^{m} a_{ij} w_i, \quad a_{ij} \in K.$$

Also ist

$$q_{B'}(f(v_j)) = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix}$$

der Koordinatenvektor von $f(v_j)$ bezüglich der Basis $B'$ von $W$.

(3) Wir bilden die Matrix $A$, die die Koordinatenvektoren $q_{B'}(f(v_1)), \ldots,$ $q_{B'}(f(v_n))$ als Spaltenvektoren hat:

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

**Definition** Die Matrix $A \in \mathrm{Mat}(m, n; K)$ heißt *Darstellungsmatrix von $f$ bezüglich der Basen $B$ und $B'$*. Wir bezeichnen sie mit $M_{B'}^{B}(f)$.

Wir bezeichnen im Folgenden für eine Matrix $A \in \mathrm{Mat}(m, n; K)$ die Abbildung $K^n \to K^m$, $x \mapsto Ax$, ebenfalls mit $A$. Dann können wir uns die Abbildungsmatrix in dem folgenden Diagramm veranschaulichen:

$$\begin{tikzcd} V \arrow[r, "f"] \arrow[d, "q_B"'] & W \arrow[d, "q_{B'}"] \\ K^n \arrow[r, "M_{B'}^{B}(f)"] & K^m \end{tikzcd}$$

Es gilt

$$q_{B'} \circ f = M_{B'}^{B}(f) \circ q_B.$$

Man sagt dazu, das Diagramm ist *kommutativ*. Das bedeutet, dass wir zwei verschiedene Wege haben, um entlang der Pfeilrichtungen von $V$ nach $K^m$ zu gelangen, und bei der Komposition der entsprechenden Abbildungen kommen die gleichen Abbildungen heraus.

**Beispiel 16.1** Wir betrachten die lineare Abbildung $f : \mathbb{R}^3 \to \mathbb{R}^2$, die definiert ist durch

$$\begin{aligned} f_1(x, y, z) &= 2x - 3y + z, \\ f_2(x, y, z) &= -x + 2y - z. \end{aligned}$$

Mit $e_1$, $e_2$, $e_3$ bezeichnen wir die Standardbasisvektoren von $\mathbb{R}^3$. Dann gilt

$$f\ (e_1) = \begin{pmatrix} 2 \\ -1 \end{pmatrix},\ f(e_2) = \begin{pmatrix} -3 \\ 2 \end{pmatrix},\ f(e_3) = \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Also hat $f$ bezüglich der Standardbasen von $\mathbb{R}^3$ und $\mathbb{R}^2$ die Darstellungsmatrix

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

**Bemerkung 16.1** Die lineare Abbildung $f$ ist durch ihre Darstellungsmatrix $A$ bezüglich der Basen $B$ und $B'$ eindeutig bestimmt, denn nach Satz 13.2 ist eine lineare Abbildung durch die Bilder der Vektoren einer Basis vollständig festgelegt. Daraus folgt, dass es nach Wahl von Basen $B$ von $V$ und $B'$ von $W$ eine eineindeutige Beziehung zwischen linearen Abbildungen und Matrizen gibt: Eine lineare Abbildung bestimmt nach der obigen Konstruktion eine Matrix. Eine Matrix $A$ bestimmt umgekehrt eine lineare Abbildung $f : V \to W$ wie folgt: Es sei $v \in V$ gegeben. Ist $x = q_B(v)$ der Koordinatenvektor des Vektors $v$ bezüglich der Basis $B$ von $V$, so erhält man den Koordinatenvektor $y = q_{B'}(f(v))$ des Vektors $f(v)$ bezüglich der Basis $B'$ von $W$ wie folgt:

$$y = Ax.$$

**Satz 16.1** *Es seien $U$, $V$, $W$ endlich dimensionale $K$-Vektorräume, $f : U \to V$, $g : V \to W$ lineare Abbildungen, $B$ Basis von $U$, $B'$ Basis von $V$, $B''$ Basis von $W$. Dann gilt*

$$M^B_{B''}(g \circ f) = M^{B'}_{B''}(g) M^B_{B'}(f).$$

*Beweis.* Es sei

$$B = (u_1, \ldots, u_p),\ B' = (v_1, \ldots, v_n),\ B'' = (w_1, \ldots, w_m),$$

$$f(u_j) = \sum_{k=1}^{n} a_{kj} v_k,\ g(v_k) = \sum_{i=1}^{m} b_{ik} w_i,\ (g \circ f)(u_j) = \sum_{i=1}^{m} c_{ij} w_i.$$

Dann gilt

$$\begin{aligned} (g \circ f)(u_j) &= g(f(u_j)) \\ &= g\left( \sum_{k=1}^{n} a_{kj} v_k \right) \\ &= \sum_{k=1}^{n} a_{kj} g(v_k) \\ &= \sum_{k=1}^{n} a_{kj} \left( \sum_{i=1}^{m} b_{ik} w_i \right) \\ &= \sum_{i=1}^{m} \left( \sum_{k=1}^{n} b_{ik} a_{kj} \right) w_i. \end{aligned}$$

Daraus folgt

$$c_{ij} = \sum_{k=1}^{n} b_{ik} a_{kj},$$

was zu zeigen war. $\square$

**Satz 16.2** *Es sei $V$ ein endlich dimensionaler $K$-Vektorraum, $f : V \to V$ ein Automorphismus und $B$ eine Basis von $V$. Dann existiert die Umkehrabbildung $f^{-1} : V \to V$ und es gilt*

$$M_B^B(f^{-1}) = M_B^B(f)^{-1}.$$

*Beweis.* Es gilt $f \circ f^{-1} = \mathrm{id}$, wobei $\mathrm{id} : V \to V$ die identische Abbildung mit $\mathrm{id}(v) = v$ ist. Nach Satz 16.1 gilt dann

$$M_B^B(f) M_B^B(f^{-1}) = M_B^B(\mathrm{id}) = E.$$

Daraus folgt die Behauptung. $\square$

Man beachte, dass die Darstellungsmatrix einer linearen Abbildung von den Basen $B$ und $B'$ abhängt. Wir diskutieren nun, wie sich die Darstellungsmatrix einer linearen Abbildung ändert, wenn wir zu anderen Basen übergeben.

Dazu nehmen wir an, dass in dem endlich dimensionalen $K$-Vektorraum $V$ zwei Basen $B = (v_1, \ldots, v_n)$ und $B' = (v'_1, \ldots, v'_n)$ gegeben sind. Dann können wir jeden Vektor $v_j$ als Linearkombination der Basisvektoren $v'_1, \ldots, v'_n$ darstellen:

$$v_j = \sum_{i=1}^{n} t_{ij} v'_i.$$

**Definition** Die Matrix $T_{B'}^{B} := (t_{ij}) \in \mathrm{Mat}(n, n; K)$, deren Spalten die Koordinatenvektoren $q_{B'}(v_j)$ der Basisvektoren $v_j$ bezüglich der Basis $B'$ sind, heißt die *Transformationsmatrix von $B$ nach $B'$.*

Auch die Transformationsmatrix können wir uns in einem kommutativen Diagramm veranschaulichen:

$$\begin{tikzcd} & V \arrow[dl, "q_B"'] \arrow[dr, "q_{B'}"] & \\ K^n \arrow[rr, "T_{B'}^{B}"] & & K^n \end{tikzcd}$$

**Bemerkung 16.2** Es sei $\mathrm{id} : V \to V$ die identische Abbildung, die durch $\mathrm{id}(v) = v$ definiert ist. Dann gilt $T_{B'}^{B} = M_{B'}^{B}(\mathrm{id})$. Denn das obige kommutative Diagramm entspricht dem folgenden Diagramm:

$$\begin{array}{ccc} V & \xrightarrow{\mathrm{id}} & V \\ {\scriptstyle q_B}\downarrow & & \downarrow{\scriptstyle q_{B'}} \\ K^n & \xrightarrow{M_{B'}^{B}(\mathrm{id})} & K^n \end{array}$$

**Beispiel 16.2** Es sei $B = (e_1, e_2, e_3)$ die Standardbasis des $\mathbb{R}^3$ und $B' = (v_1, v_2, v_3)$ die Basis mit

$$v_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 2 \\ 3 \\ 3 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 3 \\ 4 \\ 5 \end{pmatrix}.$$

Dann gilt

$$\begin{aligned} e_1 &= 3v_1 - v_2, \\ e_2 &= -v_1 + 2v_2 - v_3, \\ e_3 &= -v_1 - v_2 + v_3. \end{aligned}$$

Also lautet die Transformationsmatrix von $B$ nach $B'$

$$T_{B'}^{B} = \begin{pmatrix} 3 & -1 & -1 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix}.$$

Die Transformationsmatrix von $B'$ nach $B$ lautet

$$T_{B}^{B'} = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 4 \\ 1 & 3 & 5 \end{pmatrix}.$$

Man rechnet leicht aus, dass

$$T_{B}^{B'} = (T_{B'}^{B})^{-1}.$$

Ist $x$ der Koordinatenvektor eines Vektors $v \in V$ bezüglich der Basis $B$, so erhält man den Koordinatenvektor $y$ von $v$ bezüglich $B'$ durch

$$y = T_{B'}^{B} x.$$

Das bedeutet, dass die Transformationsmatrix $T_{B'}^{B}$ angibt, wie sich die Koordinaten zur Basis $B$ in die Koordinaten zur Basis $B'$ transformieren. Durch Auflösen der obigen Gleichung nach $x$ erhalten wir

$$x = (T_{B'}^{B})^{-1} y.$$

Das bedeutet, dass die Matrix $(T^B_{B'})^{-1}$ die Transformation von den Koordinaten zu $B'$ in die Koordinaten zu $B$ angibt. Damit erhalten wir

**Satz 16.3** *Es seien $B$ und $B'$ Basen des endlich dimensionalen $K$-Vektorraums $V$. Dann ist die Transformationsmatrix $T^B_{B'}$ invertierbar und es gilt*

$$T^{B'}_B = (T^B_{B'})^{-1}.$$

Wir können nun eine Antwort auf die Frage geben, wie sich die Darstellungsmatrix einer linearen Abbildung bei Einführung neuer Basen ändert.

**Satz 16.4 (Transformationsformel)** *Es sei $f : V \to W$ eine lineare Abbildung, $B$ und $B'$ Basen von $V$ und $C$ und $C'$ Basen von $W$. Dann gilt*

$$M^{B'}_{C'}(f) = T^C_{C'} M^B_C(f)(T^B_{B'})^{-1}.$$

*Setzen wir $A := M^B_C(f)$, $A' := M^{B'}_{C'}(f)$, $T := T^B_{B'}$ und $S := T^C_{C'}$, so können wir diese Gleichung auch so ausdrücken*

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

*Beweis.* Dies folgt leicht aus Satz 16.1. □

Für den Spezialfall $W = V$ ergibt sich

**Korollar 16.1** *Es sei $f : V \to V$ eine lineare Abbildung und es seien $B$, $B'$ Basen von $V$. Wir setzen $A := M^B_B(f)$, $A' := M^{B'}_{B'}(f)$, $T := T^B_{B'}$. Dann gilt*

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

**Beispiel 16.3** Es sei $f : \mathbb{R}^3 \to \mathbb{R}^2$ die lineare Abbildung von Beispiel 16.1. Es sei $B$ die Standardbasis von $\mathbb{R}^3$ und $C$ die Standardbasis von $\mathbb{R}^2$. Es sei $B' = (v_1, v_2, v_3)$ die Basis von $\mathbb{R}^3$ mit

$$v_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 2 \\ 3 \\ 3 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 3 \\ 4 \\ 5 \end{pmatrix}$$

(vgl. Beispiel 16.2). Schließlich sei $C' = (w_1, w_2)$ die Basis von $\mathbb{R}^2$ mit

$$w_1 = \begin{pmatrix} 2 \\ 3 \end{pmatrix}, \quad w_2 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

Nach Beispiel 16.2 gilt

$$(T_{\mathcal{B}'}^{\mathcal{B}})^{-1} = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 4 \\ 1 & 3 & 5 \end{pmatrix}.$$

Für $T_{\mathcal{C}'}^{\mathcal{C}}$ berechnen wir

$$T_{\mathcal{C}'}^{\mathcal{C}} = (T_{\mathcal{C}}^{\mathcal{C}'})^{-1} = \begin{pmatrix} 2 & 1 \\ 3 & 2 \end{pmatrix}^{-1} = \begin{pmatrix} 2 & -1 \\ -3 & 2 \end{pmatrix}.$$

Dann gilt für die Darstellungsmatrix von $f$ bezüglich der Basen $\mathcal{B}'$ und $\mathcal{C}'$ nach Satz 16.4

$$\begin{aligned} M_{\mathcal{C}'}^{\mathcal{B}'}(f) &= \begin{pmatrix} 2 & -1 \\ -3 & 2 \end{pmatrix} \begin{pmatrix} 2 & -3 & 1 \\ -1 & 2 & -1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 4 \\ 1 & 3 & 5 \end{pmatrix} \\ &= \begin{pmatrix} 2 & -1 \\ -3 & 2 \end{pmatrix} \begin{pmatrix} 0 & -2 & -1 \\ 0 & 1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 & -5 & -2 \\ 0 & 8 & 3 \end{pmatrix}. \end{aligned}$$

**Notation**

$$\mathrm{GL}(n;K) := \{A \in \mathrm{Mat}(n,n;K) \mid A \text{ ist invertierbar}\}.$$

**Definition** Zwei Matrizen $A, A' \in \mathrm{Mat}(m,n;K)$ heißen *äquivalent*, genau dann, wenn es eine Matrix $S \in \mathrm{GL}(m;K)$ und eine Matrix $T \in \mathrm{GL}(n;K)$ gibt, so dass

$$A' = SAT^{-1}$$

gilt.

Zwei Matrizen $A, A' \in \mathrm{Mat}(n,n;K)$ heißen *ähnlich*, genau dann, wenn es eine Matrix $T \in \mathrm{GL}(n;K)$ gibt, so dass

$$A' = TAT^{-1}$$

gilt.

**Bemerkung 16.3** Damit können wir festhalten: *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 die gleiche lineare Selbstabbildung beschreiben.*

**Satz 16.5** *Es sei $f : V \to W$ eine lineare Abbildung zwischen endlich dimensionalen $K$-Vektorräumen vom Rang $r$. Dann gibt es eine Basis $B$ von $V$ und eine Basis $C$ von $W$, so dass*

$$M_C^B(f) = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}.$$

*Beweis.* Es sei $B$ die Basis $B := (u_1, \ldots, u_r, v_1, \ldots, v_k)$ aus dem Beweis von Satz 14.2. Wir ergänzen die Basis $(w_1, \ldots, w_r)$ von $\operatorname{Im} f$ aus diesem Beweis zu einer Basis $C$ von $W$. Wegen $f(u_i) = w_i$, $i = 1, \ldots, r$, hat $f$ dann bezüglich der Basen $B$ und $C$ die angebene Darstellungsmatrix. □

**Korollar 16.2** *Eine Matrix $A \in \mathrm{Mat}(m, n; K)$ vom Rang $r$ ist äquivalent zu der Matrix*

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

*Beweis.* Es sei $f : K^n \to K^m$ die lineare Abbildung mit $f(x) = Ax$. Dann hat $f$ den Rang $r$. Nach Satz 16.5 gibt es eine Basis $B$ von $K^n$ und eine Basis $C$ von $K^m$, so dass $A' := M_{\mathcal{C}}^B(f)$ die angebene Form hat. Es sei $T$ die Transformationsmatrix der Standardbasis von $K^n$ nach $B$ und $S$ die Transformationsmatrix der Standardbasis von $K^m$ nach $C$. Dann gilt nach Satz 16.4

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

Also sind die Matrizen $A$ und $B$ äquivalent. □

# 17 Determinanten

Wir wollen nun den Begriff der Determinante einer Matrix einführen. Dazu erinnern wir zunächst an den Begriff der symmetrischen Gruppe (vgl. §8)

**Definition** Es sei $M = \{1, \ldots, n\}$. Die Menge aller bijektiven Abbildungen von $M$ nach $M$ bezeichnet man mit $S_n$ und nennt sie die *symmetrische Gruppe in n Variablen.* Die Elemente von $S_n$ heißen *Permutationen* von $n$ Elementen.

**Satz 17.1** *Die Gruppe $S_n$ enthält*

$$n! := n \cdot (n-1) \cdot \ldots \cdot 2 \cdot 1 \quad (n\text{-Fakultät})$$

*Elemente.*

*Beweis.* Wir beweisen die Behauptung durch Induktion nach $n$.

Induktionsanfang $n = 1$: Klar.

Induktionsschritt $n \to n+1$: Für $i = 1, \dots, n+1$ betrachte die Menge

$$S_{n+1}^i := \{\sigma \in S_{n+1} \mid \sigma(1) = i\}.$$

Jeder Permutation in dieser Menge entspricht aber eine bijektive Abbildung der Menge $\{2, \dots, n+1\}$ nach der Menge $\{1, \dots, i-1, i+1, \dots, n+1\}$. Dies sind aber beides $n$-elementige Mengen. Also können wir jedem Element von $S_{n+1}^i$ eindeutig ein Element von $S_n$ zuordnen. Nach Induktionsannahme hat $S_n$ genau $n!$ Elemente. Nun ist aber $S_{n+1}$ die disjunkte Vereinigung

$$S_{n+1} = S_{n+1}^1 \cup \ldots \cup S_{n+1}^{n+1}.$$

Also hat $S_{n+1}$ genau $(n+1)n! = (n+1)!$ Elemente. □

**Notation** Eine Permutation $\sigma \in S_n$ schreibt man als

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(n) \end{pmatrix}$$

**Beispiel 17.1** Wir betrachten die folgenden Elemente von $S_3$:

$$\sigma = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, \quad \tau = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}.$$

Dann gilt

$$\sigma \circ \tau = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}, \quad \tau \circ \sigma = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} \neq \sigma \circ \tau.$$

**Definition** Eine Permutation $\tau \in \mathcal{S}_n$ heißt *Transposition*, falls $\tau$ zwei Elemente aus $\{1, \dots, n\}$ vertauscht und alle anderen festlässt, d.h. wenn es $k, \ell \in \{1, \dots, n\}$ mit $k \neq \ell$ gibt, so dass gilt:

$$\begin{aligned} \tau(k) &= \ell, \\ \tau(\ell) &= k, \\ \tau(i) &= i \text{ für } i \in \{1, \dots, n\}, i \neq k, \ell. \end{aligned}$$

Offensichtlich gilt $\tau^{-1} = \tau$ für jede Transposition $\tau \in \mathcal{S}_n$.

**Lemma 17.1** *Ist $n \geq 2$, so lässt sich jede Permutation $\sigma \in S_n$ als Hintereinanderschaltung von Transpositionen darstellen, d.h. es gibt Transpositionen $\tau_1, \dots, \tau_m \in S_n$ mit*

$$\sigma = \tau_1 \circ \cdots \circ \tau_m.$$

*Beweis.* Ist $\sigma = \mathrm{id}$, so gilt $\sigma = \tau \circ \tau$ für irgendeine Transposition.

Andernfalls gibt es ein $i_1 \in \{1, \ldots, n\}$ mit

$$\sigma(i) = i \text{ für } i = 1, \ldots, i_1 - 1 \text{ und } \sigma(i_1) > i_1.$$

Es sei $\tau_1$ die Transposition, die $i_1$ mit $\sigma(i_1)$ vertauscht, und $\sigma_1 := \tau_1 \circ \sigma$. Dann gilt

$$\sigma_1(i) = i \text{ für } i = 1, \ldots, i_1.$$

Nun ist entweder $\sigma_1 = \mathrm{id}$ oder es gibt ein $i_2$ mit $i_2 > i_1$ und

$$\sigma_1(i) = i \text{ für } i = 1, \ldots, i_2 - 1 \text{ und } \sigma_1(i_2) > i_2.$$

Analog erhält man $\tau_2$ und $\sigma_2$, usw., bis wir schließlich ein $m \leq n$ und Transpositionen $\tau_1, \ldots, \tau_m$ mit

$$\sigma_m = \tau_m \circ \cdots \circ \tau_1 \circ \sigma = \mathrm{id}$$

erhalten. Daraus folgt

$$\sigma = (\tau_m \circ \cdots \circ \tau_1)^{-1} = \tau_1^{-1} \circ \cdots \circ \tau_m^{-1} = \tau_1 \circ \cdots \circ \tau_m.$$

□

**Lemma 17.2** *Es sei $n \geq 2$ und $\tau_{12}$ die Transposition, die 1 und 2 vertauscht, und $\tau$ eine beliebige Transposition. Dann gibt es eine Permutation $\sigma \in S_n$ mit*

$$\tau = \sigma \circ \tau_{12} \circ \sigma^{-1}.$$

*Beweis.* Es seien $k$ und $\ell$ die von $\tau$ vertauschten Elemente. Es sei $\sigma \in S_n$ eine Permutation mit

$$\sigma(1) = k \text{ und } \sigma(2) = \ell.$$

Es sei $\tau' := \sigma \circ \tau_{12} \circ \sigma^{-1}$. Es ist zu zeigen: $\tau = \tau'$. Es gilt:

$$\begin{aligned}
\tau'(k) &= \sigma(\tau_{12}(\sigma^{-1}(k))) = \sigma(\tau_{12}(1)) = \sigma(2) = \ell, \\
\tau'(\ell) &= \sigma(\tau_{12}(\sigma^{-1}(\ell))) = \sigma(\tau_{12}(2)) = \sigma(1) = k, \\
\tau'(i) &= \sigma(\tau_{12}(\sigma^{-1}(i))) = \sigma(\sigma^{-1}(i)) = i \text{ für } i \neq k, \ell.
\end{aligned}$$

Daraus folgt $\tau' = \tau$. □

**Definition** Es sei $\sigma \in S_n$. Eine *Inversion* von $\sigma$ ist ein Zahlenpaar $(i, j)$ mit $1 \leq i < j \leq n$ aber $\sigma(i) > \sigma(j)$, d.h. ein Zahlenpaar, bei der die Reihenfolge bei der Permutation vertauscht wird.

**Beispiel 17.2** Die Anzahl der Inversionen von

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

ist $2 + 1 + 1$.

**Definition** Das *Signum* einer Permutation $\sigma \in S_n$ ist definiert durch

$$\operatorname{sign} \sigma := \begin{cases} +1, & \text{falls } \sigma \text{ eine gerade Anzahl von Inversionen hat,} \\ -1, & \text{falls } \sigma \text{ eine ungerade Anzahl von Inversionen hat.} \end{cases}$$

**Beispiel 17.3**

$$\operatorname{sign} \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 4 & 1 \end{pmatrix} = 1.$$

**Lemma 17.3** *Für jedes* $\sigma \in S_n$ *gilt*

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

*Beweis.* Im Zähler und Nenner des Produkts auf der rechten Seite kommen die gleichen Differenzen vor, im Falle einer Inversion aber mit umgekehrtem Vorzeichen. □

**Satz 17.2** *Für alle* $\sigma, \tau \in S_n$ *gilt*

$$\operatorname{sign}(\tau \circ \sigma) = \operatorname{sign} \tau \cdot \operatorname{sign} \sigma.$$

**Bemerkung 17.1** Dieser Satz besagt gerade, dass die Abbildung

$$\operatorname{sign} : (S_n, \circ) \to (\{-1, +1\}, \cdot)$$

ein Gruppenhomomorphismus in die Gruppe mit zwei Elementen ist.

**Korollar 17.1** *Es gilt* $\operatorname{sign}(\sigma^{-1}) = \operatorname{sign} \sigma$.

*Beweis von Satz 17.2.* Es ist

$$\begin{aligned} \operatorname{sign}(\tau \circ \sigma) &= \prod_{i<j} \frac{\tau(\sigma(j)) - \tau(\sigma(i))}{j - i} \\ &= \prod_{i<j} \frac{\tau(\sigma(j)) - \tau(\sigma(i))}{\sigma(j) - \sigma(i)} \cdot \prod_{i<j} \frac{\sigma(j) - \sigma(i)}{j - i} \end{aligned}$$

Da das zweite Produkt gleich $\operatorname{sign}\sigma$ ist, genügt es zu zeigen, dass das erste Produkt gleich $\operatorname{sign}\tau$ ist.

$$\begin{aligned}
&\prod_{i<j} \frac{\tau(\sigma(j)) - \tau(\sigma(i))}{\sigma(j) - \sigma(i)} \\
&\quad = \prod_{\substack{i<j \\ \sigma(i)<\sigma(j)}} \frac{\tau(\sigma(j)) - \tau(\sigma(i))}{\sigma(j) - \sigma(i)} \cdot \prod_{\substack{i<j \\ \sigma(i)>\sigma(j)}} \frac{\tau(\sigma(j)) - \tau(\sigma(i))}{\sigma(j) - \sigma(i)} \\
&\quad = \prod_{\substack{i<j \\ \sigma(i)<\sigma(j)}} \frac{\tau(\sigma(j)) - \tau(\sigma(i))}{\sigma(j) - \sigma(i)} \cdot \prod_{\substack{i>j \\ \sigma(i)<\sigma(j)}} \frac{\tau(\sigma(j)) - \tau(\sigma(i))}{\sigma(j) - \sigma(i)} \\
&\quad = \prod_{\sigma(i)<\sigma(j)} \frac{\tau(\sigma(j)) - \tau(\sigma(i))}{\sigma(j) - \sigma(i)}
\end{aligned}$$

Da $\sigma$ bijektiv ist, kann man das Produkt statt über $i < j$ auch über $\sigma(i) < \sigma(j)$ bilden. Also ist das letzte Produkt gleich $\operatorname{sign}\tau$. □

**Korollar 17.2** *Es sei $n \geq 2$.*

(i) *Für jede Transposition $\tau \in S_n$ gilt* $\operatorname{sign}\tau = -1$.

(ii) *Für $\sigma = \tau_1 \circ \cdots \circ \tau_k$ mit Transpositionen $\tau_1, \ldots, \tau_k$ gilt*

$$\operatorname{sign}\sigma = (-1)^k.$$

*Beweis.* (i) Ist $\tau_{12}$ die Transposition, die 1 und 2 vertauscht, so ist $\operatorname{sign}\tau_{12} = -1$, denn $\tau_{12}$ hat genau eine Inversion. Nach Lemma 17.2 gibt es ein $\sigma \in S_n$ mit $\tau = \sigma \circ \tau_{12} \circ \sigma^{-1}$. Also folgt aus Satz 17.2

$$\operatorname{sign}\tau = \operatorname{sign}\sigma \cdot \operatorname{sign}\tau_{12} \cdot \operatorname{sign}\sigma^{-1} = \operatorname{sign}\tau_{12} = -1.$$

(ii) folgt aus (i) und Satz 17.2. □

**Definition** Die Menge

$$A_n := \{\sigma \in S_n \,|\, \operatorname{sign}\sigma = +1\} \subset S_n$$

heißt die *alternierende Gruppe.*

**Lemma 17.4** *Die alternierende Gruppe $A_n$ ist eine Untergruppe von $S_n$.*

*Beweis.* Es sei $\sigma, \tau \in A_n$. Dann gilt

$$\operatorname{sign}(\sigma \circ \tau^{-1}) = \operatorname{sign}\sigma \operatorname{sign}(\tau^{-1}) = \operatorname{sign}\sigma \operatorname{sign}\tau = 1 \cdot 1 = 1.$$

Also ist auch $\sigma \circ \tau^{-1} \in A_n$. □

**Lemma 17.5** *Es sei* $\tau \in S_n$ *mit* $\operatorname{sign}\tau = -1$. *Dann ist*

$$S_n = A_n \cup A_n\tau \text{ und } A_n \cap A_n\tau = \emptyset.$$

*Beweis.* Es sei $\sigma \in S_n$ gegeben. Ist $\operatorname{sign}\sigma = 1$, so ist $\sigma \in A_n$. Es sei $\operatorname{sign}\sigma = -1$. Dann ist $\operatorname{sign}(\sigma \circ \tau^{-1}) = 1$. Also ist $\sigma = (\sigma \circ \tau^{-1}) \circ \tau \in A_n\tau$. Die Vereinigung ist disjunkt, da für jedes $\sigma \in A_n\tau$ gilt: $\operatorname{sign}\sigma = -1$. □

Im Folgenden sei nun $A$ eine $n \times n$-Matrix über einem Körper $K$.

**Definition** Ein *elementares Produkt* von Einträgen aus $A$ ist ein Produkt von $n$ Einträgen aus $A$, wobei aus jeder Zeile und jeder Spalte genau ein Eintrag stammt. Ein elementares Produkt von Einträgen aus $A$ ist also ein Ausdruck

$$a_{1\sigma(1)} \cdots a_{n\sigma(n)},$$

wobei $\sigma$ eine Permutation der Menge $\{1, 2, \ldots, n\}$ ist.

**Beispiel 17.4** Die elementaren Produkte der Einträge der Matrix

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

sind

$$a_{11}a_{22} \text{ und } a_{12}a_{21}.$$

Die elementaren Produkte der Einträge der Matrix

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

sind

$$\begin{array}{lll} a_{11}a_{22}a_{33} & a_{12}a_{23}a_{31} & a_{13}a_{21}a_{32} \\ a_{13}a_{22}a_{31} & a_{11}a_{23}a_{32} & a_{12}a_{21}a_{33}. \end{array}$$

**Definition** Die *Determinante* der Matrix $A$, in Zeichen $\det A$, ist definiert durch die Formel

$$\det A = \sum_{\sigma \in S_n} \operatorname{sign}\sigma \, a_{1\sigma(1)} \cdots a_{n\sigma(n)}.$$

**Warnung** Man beachte, dass die Determinante nur für eine *quadratische* Matrix definiert ist!

**Beispiel 17.5** Mit Hilfe von Beispiel 17.4 erhalten wir

$$\begin{aligned}\det\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} &= a_{11}a_{22} - a_{12}a_{21}, \\ \det\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} &= \begin{array}{l} a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\ -a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}. \end{array}\end{aligned}$$

Die Berechnung der Determinante einer $3 \times 3$-Matrix merkt man sich mit Hilfe der folgenden Regel:

**Regel von Sarrus:** *Man schreibe die ersten beiden Spalten der Matrix noch einmal hinter die Matrix und multipliziere entlang der angedeuteten Pfeile, wobei die elementaren Produkte längs der nach oben gerichteten Pfeile mit dem Vorzeichen − zu versehen sind:*

$$\left(\begin{array}{ccc} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{array}\right)\begin{array}{cc} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{array}$$

**Beispiel 17.6** Die Determinante der Matrix

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

lautet

$$\det A = 2 + 0 + (-6) - (-4) - 1 - 0 = -1.$$

**Warnung** Die Regel von Sarrus funktioniert nur für $3 \times 3$-Matrizen, *nicht* für größere Matrizen!

Die Berechnung von Determinanten mit Hilfe der Definition erfordert im Allgemeinen einen großen Rechenaufwand. Um eine Determinante einer $4 \times 4$-Matrix zu berechnen, braucht man schon $4! = 24$ Terme, bei einer $10 \times 10$-Matrix $10! = 3\,628\,800$. Selbst für schnelle Computer wächst die Rechenzeit mit der Größe der Matrix sehr schnell ins Unermessliche.

**Notation** Statt $\det A$ schreibt man auch $|A|$, also z.B.

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

Wir wollen uns nun mit Eigenschaften von Determinanten beschäftigen. Es sei $A \in \mathrm{Mat}(n, n; K)$. Die Zeilenvektoren von $A$ bezeichnen wir im Folgenden mit $a_1, \dots, a_n$, also

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

**Satz 17.3** *Die Funktion*

$$\begin{array}{rccl} \det: & \mathrm{Mat}(n,n;K) & \longrightarrow & K \\ & A & \longmapsto & \det A \end{array}$$

*hat die folgenden Eigenschaften:*

(D1) det *ist* linear in jeder Zeile, *d.h. für jedes* $i = 1, \dots, n$ *gilt*

$$\text{(a)} \qquad \det \begin{pmatrix} a_1 \\ \vdots \\ a_i' + a_i'' \\ \vdots \\ a_n \end{pmatrix} = \det \begin{pmatrix} a_1 \\ \vdots \\ a_i' \\ \vdots \\ a_n \end{pmatrix} + \det \begin{pmatrix} a_1 \\ \vdots \\ a_i'' \\ \vdots \\ a_n \end{pmatrix}$$

$$\text{(b)} \qquad \det \begin{pmatrix} a_1 \\ \vdots \\ \lambda a_i' \\ \vdots \\ a_n \end{pmatrix} = \lambda \det \begin{pmatrix} a_1 \\ \vdots \\ a_i' \\ \vdots \\ a_n \end{pmatrix} \qquad \textit{für } \lambda \in K.$$

(D2) det *ist* alternierend, *d.h. hat* $A$ *zwei gleiche Zeilen, so ist* $\det A = 0$.

(D3) det *ist* normiert, *d.h.* $\det E = 1$.

*Beweis.* (D1)(a): Es gilt

$$\begin{aligned}
\det\begin{pmatrix} a_1 \\ \vdots \\ a_i' + a_i'' \\ \vdots \\ a_n \end{pmatrix} &= \sum_{\sigma\in S_n} \operatorname{sign}\sigma\, a_{1\sigma(1)}\cdots(a'_{i\sigma(i)} + a''_{i\sigma(i)})\cdots a_{n\sigma(n)} \\
&= \sum_{\sigma\in S_n} \operatorname{sign}\sigma\, a_{1\sigma(1)}\cdots a'_{i\sigma(i)}\cdots a_{n\sigma(n)} \\
&\quad + \sum_{\sigma\in S_n} \operatorname{sign}\sigma\, a_{1\sigma(1)}\cdots a''_{i\sigma(i)}\cdots a_{n\sigma(n)} \\
&= \det\begin{pmatrix} a_1 \\ \vdots \\ a_i' \\ \vdots \\ a_n \end{pmatrix} + \det\begin{pmatrix} a_1 \\ \vdots \\ a_i'' \\ \vdots \\ a_n \end{pmatrix}.
\end{aligned}$$

(D1)(b) zeigt man analog.

(D2) Angenommen $a_k = a_\ell$ für $k < \ell$. Es sei $\tau$ die Transposition, die $k$ und $l$ vertauscht. Dann gilt nach Lemma 17.5

$$S_n = A_n \cup A_n\tau$$

und diese Vereinigung ist disjunkt. Wenn $\sigma$ die Menge $A_n$ durchläuft, durchläuft $\sigma\circ\tau$ die Menge $A_n\tau$ und es gilt

$$\operatorname{sign}\sigma = +1 \text{ und } \operatorname{sign}(\sigma\circ\tau) = -1.$$

Also gilt

$$\det A = \sum_{\sigma\in A_n} a_{1\sigma(1)}\cdots a_{n\sigma(n)} - \sum_{\sigma\in A_n} a_{1\sigma(\tau(1))}\cdots a_{n\sigma(\tau(n))}.$$

Nun gilt aber wegen $a_k = a_\ell$

$$\begin{aligned}
&a_{1\sigma(\tau(1))}\cdots a_{k\sigma(\tau(k))}\cdots a_{\ell\sigma(\tau(\ell))}\cdots a_{n\sigma(\tau(n))} \\
&= a_{1\sigma(1)}\cdots a_{k\sigma(\ell)}\cdots a_{\ell\sigma(k)}\cdots a_{n\sigma(n)} \\
&= a_{1\sigma(1)}\cdots a_{k\sigma(k)}\cdots a_{\ell\sigma(\ell)}\cdots a_{n\sigma(n)} = a_{1\sigma(1)}\cdots a_{n\sigma(n)}.
\end{aligned}$$

Daraus folgt

$$\det A = 0.$$

(D3) Es sei

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

das *Kronecker-Symbol*. Dann gilt $E = (\delta_{ij})$ und damit

$$\det E = \sum_{\sigma \in S_n} \operatorname{sign} \sigma \, \delta_{1\sigma(1)} \cdots \delta_{n\sigma(n)} = \operatorname{sign}(\mathrm{id}) \, \delta_{11} \cdots \delta_{nn} = 1.$$

□

**Bemerkung 17.2** Man kann beweisen, dass es genau eine Funktion $\det : \mathrm{Mat}(n,n;K) \to K$ mit den Eigenschaften (D1)–(D3) gibt. Man kann die Determinante also auch durch die Eigenschaften (D1)–(D3) charakterisieren.

Aus den Eigenschaften (D1)–(D3) folgen weitere Eigenschaften der Determinante:

**Satz 17.4** *Es sei* $A \in \mathrm{Mat}(n,n;K)$.

(D4) *Ist $B$ die Matrix, die aus $A$ durch Addition eines Vielfachen einer Zeile zu einer anderen entsteht, so gilt* $\det B = \det A$.

(D5) *Entsteht $B$ aus $A$ durch eine Zeilenvertauschung, so ist*

$$\det B = -\det A.$$

(D6) *Enthält $A$ eine Nullzeile, so ist* $\det A = 0$.

(D7) *Enthält $A$ zwei linear abhängige Zeilen, so ist* $\det A = 0$.

(D8) *Es sei $\lambda \in K$. Dann gilt*

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

(D9) *Ist $A$ eine obere Dreiecksmatrix, also*

$$A = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix},$$

*dann ist* $\det A = \lambda_1 \cdots \lambda_n$.

(D10) *Es sei $A$ von der Gestalt*

$$A = \left(\begin{array}{c|c} A_1 & C \\ \hline 0 & A_2 \end{array}\right) \quad (A_1, A_2 \textit{ quadratisch}).$$

*Dann gilt*

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

*Beweis.* (D4) Die Matrix $B$ entstehe aus $A$ durch Addition des $\lambda$-fachen der $k$-ten Zeile zur $l$-ten Zeile. Nach Satz 17.3 gilt dann

$$\det B = \det \begin{pmatrix} \vdots \\ a_k \\ \vdots \\ a_l + \lambda a_k \\ \vdots \end{pmatrix} = \det \begin{pmatrix} \vdots \\ a_k \\ \vdots \\ a_l \\ \vdots \end{pmatrix} + \lambda \det \begin{pmatrix} \vdots \\ a_k \\ \vdots \\ a_k \\ \vdots \end{pmatrix} = \det A$$

(D5) Wir nehmen an, dass $B$ aus $A$ durch Vertauschen der Zeilen $i < j$ entsteht. Dann folgt aus Satz 17.3 (D1) und (D2):

$$\begin{aligned} \det A + \det B &= \det \begin{pmatrix} \vdots \\ a_i \\ \vdots \\ a_j \\ \vdots \end{pmatrix} + \det \begin{pmatrix} \vdots \\ a_j \\ \vdots \\ a_i \\ \vdots \end{pmatrix} \\ &= \det \begin{pmatrix} \vdots \\ a_i \\ \vdots \\ a_i \\ \vdots \end{pmatrix} + \det \begin{pmatrix} \vdots \\ a_i \\ \vdots \\ a_j \\ \vdots \end{pmatrix} + \det \begin{pmatrix} \vdots \\ a_j \\ \vdots \\ a_i \\ \vdots \end{pmatrix} + \det \begin{pmatrix} \vdots \\ a_j \\ \vdots \\ a_j \\ \vdots \end{pmatrix} \\ &= \det \begin{pmatrix} \vdots \\ a_i + a_j \\ \vdots \\ a_i + a_j \\ \vdots \end{pmatrix} = 0 \end{aligned}$$

(D6) Die $l$-te Zeile der Matrix $A$ sei eine Nullzeile. Addieren wir eine beliebige andere Zeile, etwa die $k$-te Zeile, $k \neq l$, zu dieser Zeile, so enthält die neue Matrix $B$ eine Zeile doppelt. Nach Satz 17.3 (D2) gilt $\det B = 0$.

(D7) Enthält die Matrix A zwei linear abhängige Zeilen, so kann man durch Addition eines geeigneten Vielfachen der einen Zeile zur anderen eine Nullzeile erzeugen.

(D8) Jede der $n$ Zeilen der Matrix $\lambda A$ enthält den gemeinsamen Faktor $\lambda$. Nach Satz 17.3 (D1)(b) kann er aus jeder Zeile vor die Determinante gezogen werden. Insgesamt ergibt sich also der Faktor $\lambda^n$.

(D9) Sind alle $\lambda_i \neq 0$, so können wir $A$ durch elementare Zeilenumformungen vom Typ (Z2) auf Diagonalgestalt bringen. Nach (D4) wird dabei die Determinante nicht geändert. Also gilt

$$\det A = \det \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} = \lambda_1 \cdots \lambda_n \det E = \lambda_1 \cdots \lambda_n.$$

Gibt es ein $i$ mit $\lambda_i = 0$, so sei $i$ maximal gewählt, d.h. $\lambda_k \neq 0$ für $i < k \leq n$. Mit Hilfe von elementaren Zeilenoperationen vom Typ (Z2) können wir dann die $i$-te Zeile zu einer Nullzeile machen. Aus (D5) folgt dann $\det A = 0$.

(D10) Durch elementare Zeilenumformungen vom Typ (Z1) und (Z2) machen wir $A_1$ zu einer oberen Dreiecksmatrix $B_1$. Dabei bleibt $A_2$ unverändert, aus $C$ wird $C'$. Ist $k_1$ die Anzahl der ausgeführten Zeilenvertauschungen, so folgt aus (D5):

$$\det B_1 = (-1)^{k_1} \det A_1.$$

Nun machen wir entsprechend $A_2$ zu einer oberen Dreiecksmatrix $B_2$. Dabei bleiben $B_1$ und $C'$ unverändert. Ist $k_2$ die Anzahl der ausgeführten Zeilenvertauschungen, so gilt

$$\det B_2 = (-1)^{k_2} \det A_2.$$

Die Matrix

$$B := \left(\begin{array}{c|c} B_1 & C' \\ \hline 0 & B_2 \end{array}\right)$$

ist dann eine obere Dreiecksmatrix. Nach (D9) folgt

$$\det B = \det B_1 \cdot \det B_2.$$

Anderseits gilt

$$\det A = (-1)^{k_1+k_2} \det B.$$

Daraus folgt die Behauptung. □

**Satz 17.5** *Es sei $A \in \mathrm{Mat}(n, n; K)$. Dann gilt*

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

*Beweis.* Es gilt

$$\begin{aligned}\det A^T &= \sum_{\sigma\in S_n} \operatorname{sign}\sigma\, a_{\sigma(1)1}\cdots a_{\sigma(n)n}\\ &= \sum_{\sigma\in S_n} \operatorname{sign}\sigma\, a_{1\sigma^{-1}(1)}\cdots a_{n\sigma^{-1}(n)},\end{aligned}$$

Da es zu jeder Permutation $\sigma \in S_n$ nach Beispiel 8.6 genau eine inverse Permutation $\sigma^{-1} \in S_n$ gibt, können wir die Summe genauso gut über $\sigma^{-1}$ laufen lassen und $\sigma^{-1}$ durchläuft $S_n$. Nach Korollar 17.1 gilt $\operatorname{sign}(\sigma^{-1}) = \operatorname{sign}\sigma$. Setzen wir also $\tau = \sigma^{-1}$, so folgt

$$\det A^T = \sum_{\tau\in S_n} \operatorname{sign}\tau\, a_{1\tau(1)}\cdots a_{n\tau(n)} = \det A.$$

□

Aus diesem Satz folgt, dass wir in allen Sätzen über Determinanten, die sich auf die Zeilen der Matrix beziehen, das Wort "Zeile" durch "Spalte" ersetzen können, wir also einen entsprechenden Satz für die Spalten der Matrix erhalten.

Wir berechnen nun die Determinanten der in §6 betrachteten Elementarmatrizen.

**Satz 17.6** (1) $\det F_{ik} = -1$.

(2) $\det F_{ik}^{\lambda} = 1$.

(3) $\det F_i^{\lambda} = \lambda$.

*Beweis.* Nach Satz 17.3 (D3) gilt $\det E = 1$. Man erhält die Elementarmatrizen durch elementare Zeilenumformungen aus der Einheitsmatrix $E$. Damit folgt die Behauptung aus Satz 17.3 und Satz 17.4. □

**Satz 17.7 (Determinantenmultiplikationssatz)** *Es seien $A$ und $B$ zwei $n \times n$-Matrizen über $K$. Dann gilt*

$$\det AB = \det A \det B.$$

*Beweis.* (a) Wir zeigen den Satz zunächst für den Spezialfall, dass $A$ eine Elementarmatrix ist. Es sei etwa $A = F_{ik}$. Dann entsteht $AB = F_{ik}B$ aus $B$ durch Vertauschung der $i$-ten mit der $k$-ten Zeile. Nach Satz 17.4 gilt $\det AB = -\det B$. Nach Satz 17.6 gilt $\det F_{ik} = -1$, also

$$\det F_{ik}B = \det F_{ik} \det B.$$

Analog zeigt man die Behauptung für die anderen Elementarmatrizen.

(b) Wir zeigen den Satz für den Spezialfall, dass $A$ eine Nullzeile enthält. In diesem Fall enthält auch $AB$ eine Nullzeile. Nach Satz 17.4 (D6) gilt dann $\det A = \det AB = 0$, also auch

$$\det AB = \det A \det B.$$

(c) Wir zeigen den Satz nun für den Fall, dass $A$ invertierbar ist. Nach Satz 6.6 lässt sich $A$ dann als Produkt von Elementarmatrizen schreiben. Damit folgt die Behauptung durch wiederholte Anwendung von (a).

(d) Schließlich betrachten wir den Fall, dass $A$ nicht invertierbar ist. Nach Satz 15.5 hat $A$ dann nicht maximalen Rang. Das bedeutet, dass die Zeilen von $A$ linear abhängig sind. Durch geeignete Zeilenumformungen kann man dann eine Nullzeile erzeugen. Den Zeilenumformungen entspricht aber die Multiplikation der Matrix $A$ mit Elementarmatrizen. Also folgt die Behauptung aus Teil (a) und (b). □

**Korollar 17.3** *Eine Matrix* $A \in \mathrm{Mat}(n, n; K)$ *ist genau dann invertierbar, wenn* $\det A \neq 0$ *gilt. In diesem Fall gilt*

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

*Beweis.* Ist $A$ invertierbar, so gilt $AA^{-1} = E$. Nach Satz 17.7 folgt

$$\det A \det A^{-1} = \det E = 1.$$

Also folgt $\det A \neq 0$ und wir können die Gleichung durch $\det A$ dividieren.

Ist $A$ nicht invertierbar, so sind die Zeilen von $A$ nach Satz 15.5 linear abhängig. Durch geeignete Zeilenumformungen kann man dann eine Nullzeile erzeugen. Aus Satz 17.4 folgt dann $\det A = 0$. □

Wir behandeln nun ein wichtiges Verfahren zur Berechnung von Determinanten. Dazu sehen wir uns noch einmal die Determinante einer $3 \times 3$-Matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

an. In Beispiel 17.5 hatten wir gesehen

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

Dies können wir auch wie folgt schreiben:

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

Die Ausdrücke in den Klammern sind aber Determinanten von Untermatrizen.

**Definition** Es sei $A \in \mathrm{Mat}(n,n;K)$. Die Matrix $A_{ij}$ entstehe aus $A$ durch Streichen der $i$-ten Zeile und $j$-ten Spalte von $A$, d.h.

$$A_{ij} = \begin{pmatrix} a_{11} & \cdots & a_{1,j-1} & a_{1j} & a_{1,j+1} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j} & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ a_{i1} & \cdots & a_{i,j-1} & a_{ij} & a_{i,j+1} & \cdots & a_{in} \\ a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,j-1} & a_{nj} & a_{n,j+1} & \cdots & a_{nn} \end{pmatrix}$$

Mit dieser Definition können wir also schreiben

$$\det A = a_{11} \det A_{11} - a_{21} \det A_{21} + a_{31} \det A_{31}.$$

Allgemein gilt der folgende Satz

**Satz 17.8 (Laplacescher Entwicklungssatz)** *Es sei $A \in \mathrm{Mat}(n,n;K)$, $1 \le k \le n$. Dann gilt*

$$\det A = \sum_{i=1}^{n} (-1)^{i+k} a_{ik} \det A_{ik} \quad \textit{(Entwicklung nach der $k$-ten Spalte)}$$

*und*

$$\det A = \sum_{j=1}^{n} (-1)^{k+j} a_{kj} \det A_{kj} \quad \textit{(Entwicklung nach der $k$-ten Zeile).}$$

**Bemerkung 17.3** Die Vorzeichen in Satz 17.8 sind gemäß eines Schachbrettmusters verteilt:

$$\begin{pmatrix} + & - & + & - & \cdots \\ - & + & - & + & \cdots \\ + & - & + & - & \cdots \\ - & + & - & + & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

**Beispiel 17.7** Wir entwickeln die Determinante der Matrix

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

nach der ersten Spalte:

$$\det A = -\begin{vmatrix} -1 & 1 & 2 \\ 1 & 1 & 0 \\ -1 & 0 & 1 \end{vmatrix} - \begin{vmatrix} -1 & 1 & 2 \\ 0 & 2 & -1 \\ -1 & 0 & 1 \end{vmatrix} = 0 - 3 = -3.$$

Zum Beweis dieses Satzes führen wir noch eine andere aus $A$ abgeleitete Matrix ein. Es sei $\widetilde{A}_{ij}$ die Matrix, die man aus $A$ enthält, indem man $a_{ij}$ durch 1 und alle anderen Einträge in der $i$-ten Zeile und $j$-ten Spalte durch 0 ersetzt:

$$\widetilde{A}_{ij} = \begin{pmatrix} a_{11} & \cdots & a_{1,j-1} & 0 & a_{1,j+1} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & 0 & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ a_{i+1,1} & \cdots & a_{i+1,j-1} & 0 & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,j-1} & 0 & a_{n,j+1} & \cdots & a_{nn} \end{pmatrix}.$$

**Lemma 17.6** *Es gilt*

$$\det \widetilde{A}_{ij} = (-1)^{i+j} \det A_{ij}.$$

*Beweis.* Durch $i-1$ Vertauschungen benachbarter Zeilen und $j-1$ Vertauschungen benachbarter Spalten kann man $\widetilde{A}_{ij}$ auf die Gestalt

$$\begin{pmatrix} 1 & 0 \\ 0 & A_{ij} \end{pmatrix}$$

bringen. Aus Satz 17.4 (D10) folgt dann

$$\det \widetilde{A}_{ij} = (-1)^{(i-1)+(j-1)} \det A_{ij} = (-1)^{i+j} \det A_{ij}.$$

□

*Beweis von Satz 17.8.* Wir zeigen die Formel für die Entwicklung nach der $k$-ten Spalte. Der Beweis für die Entwicklung nach der $k$-ten Zeile geht analog.

Es seien $s_1, \dots, s_n$ die Spaltenvektoren von $A$. Dann gilt

$$\det \widetilde{A}_{ij} = \det(s_1, \cdots, s_{j-1}, e_i, s_{j+1}, \cdots, s_n),$$

da man durch Addition von geeigneten Vielfachen der $j$-ten Spalte $e_i$ alle Einträge in der $i$-ten Zeile außer der 1 in der $j$-ten Spalte zu Null machen kann. Nun gilt

$$\begin{aligned}
\det A &= \det(s_1, \cdots, s_{k-1}, \sum_{i=1}^{n} a_{ik} e_i, s_{k+1}, \cdots, s_n) \\
&= \sum_{i=1}^{n} a_{ik} \det(s_1, \cdots, s_{k-1}, e_i, s_{k+1}, \cdots, s_n) \quad \text{(nach (D1))} \\
&= \sum_{i=1}^{n} a_{ik} \det \widetilde{A}_{ik} = \sum_{i=1}^{n} (-1)^{i+k} a_{ik} \det A_{ik} \quad \text{(nach Lemma 17.6).}
\end{aligned}$$

□

Wir betrachten nun Anwendungen von Determinanten.

**Satz 17.9 (Cramersche Regel)** *Es sei $A \in \mathrm{Mat}(n, n; K)$ eine invertierbare Matrix und $b \in K^n$. Mit $A_i$ bezeichnen wir die Matrix, die dadurch entsteht, dass wir die $i$-te Spalte von $A$ durch den Spaltenvektor $b$ ersetzen. Dann werden die Komponenten $x_i$ der Lösung $x$ des linearen Gleichungssystems $Ax = b$ durch die Formel*

$$x_i = \frac{\det A_i}{\det A}, \quad 1 \leq i \leq n,$$

*gegeben.*

*Beweis.* Wir bezeichnen die Spaltenvektoren der Matrix $A$ mit $s_1, \dots, s_n$. Dann gilt

$$\begin{aligned}
\det A_1 &= \det(b, s_2, \dots, s_n) \\
&= \det(x_1 s_1 + \cdots + x_n s_n, s_2, \dots, s_n) \\
&= x_1 \det(s_1, s_2, \dots, s_n) + \cdots + x_n \det(s_n, s_2, \dots, s_n) \\
&= x_1 \det A.
\end{aligned}$$

Der Beweis für die anderen Komponenten geht analog. □

Die Cramersche Regel ist für praktische Berechnungen nicht sehr geeignet, wohl aber für theoretische Überlegungen.

**Beispiel 17.8** Wir betrachten das folgende Gleichungssystem

$$\begin{aligned} 2x_1 - x_2 &= 4 \\ -3x_1 + 2x_2 &= -3. \end{aligned}$$

Dann ergibt die Cramersche Regel:

$$x_1 = \frac{\begin{vmatrix} 4 & -1 \\ -3 & 2 \end{vmatrix}}{\begin{vmatrix} 2 & -1 \\ -3 & 2 \end{vmatrix}} = 5, \quad x_2 = \frac{\begin{vmatrix} 2 & 4 \\ -3 & -3 \end{vmatrix}}{\begin{vmatrix} 2 & -1 \\ -3 & 2 \end{vmatrix}} = 6.$$

Wir wenden nun die Cramersche Regel auf die Bestimmung der inversen Matrix einer invertierbaren Matrix an. Es sei $A$ eine invertierbare $n \times n$-Matrix über $K$. Die inverse Matrix bestimmt sich durch das lineare Gleichungssystem

$$AX = E$$

oder

$$Ax_j = e_j, \quad j = 1, \ldots, n,$$

wobei wir mit $x_1, \ldots, x_n$ die Spalten der Matrix $X$ bezeichnen. Nach der Cramerschen Regel gilt

$$x_{ij} = \frac{1}{\det A} \det A_i,$$

wobei $A_i$ die Matrix ist, die aus $A$ entsteht, indem man die $i$-te Spalte durch den Vektor $e_j$ ersetzt. Nach dem Laplaceschen Entwicklungssatz gilt

$$x_{ij} = \frac{1}{\det A}(-1)^{i+j} \det A_{ji}.$$

**Definition** Die Matrix $A^* := (a^*_{ij})$ mit $a^*_{ij} = (-1)^{i+j} \det A_{ji}$ heißt die *adjungierte Matrix* von $A$.

Wir haben damit bewiesen:

**Satz 17.10** *Für eine invertierbare Matrix $A$ ist*

$$A^{-1} = \frac{1}{\det A} A^*.$$

**Bemerkung 17.4** Man beachte die Vertauschung der Indizes $i$ und $j$ bei der Definition der adjungierten Matrix.

**Beispiel 17.9** Die adjungierte Matrix zu der Matrix

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

lautet

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

Wegen $\det A = 1$ ist $A^{-1} = A^*$.

# 18 Eigenwerte und Eigenvektoren

Wir haben in §16 gesehen, dass jede Matrix äquivalent zu einer Matrix in besonders einfacher Form ist, nämlich zu der Matrix

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

Wir betrachten nun eine quadratische Matrix $A$ und suchen eine Matrix $A'$ in möglichst einfacher Form, die ähnlich zu der Matrix $A$ ist, d.h. wir suchen eine invertierbare Matrix $T$, so dass die Matrix

$$A' = TAT^{-1}$$

in möglichst einfacher Form ist. Dieses Problem ist schwieriger zu behandeln, da wir nur *eine* Matrix $T$ zur Verfügung haben. Die Behandlung dieses Problems führt zu der Betrachtung von Eigenwerten und Eigenvektoren.

Ist $f : \mathbb{R}^n \to \mathbb{R}^n$ eine lineare Abbildung und $x \in \mathbb{R}^n$, so besteht im Allgemeinen keine Beziehung zwischen den Vektoren $x$ und $f(x)$. Häufig existieren jedoch von Null verschiedene Vektoren $x$, so dass $x$ und $f(x)$ skalare Vielfache voneinander sind. Solche Vektoren nennt man Eigenvektoren. Sie spielen in Anwendungen eine große Rolle, so zum Beispiel bei der Untersuchung von Schwingungen und elektrischen Systemen.

Im Folgenden sei $V$ ein $K$-Vektorraum.

**Definition** Es sei $f : V \to V$ eine lineare Abbildung. Ein Vektor $v \in V$ heißt *Eigenvektor* von $f$, wenn $v \neq 0$ und es ein $\lambda \in K$ gibt mit

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

Der Skalar $\lambda$ heißt dann *Eigenwert* von $f$.

Wegen des engen Zusammenhangs zwischen linearen Abbildungen und Matrizen können wir auch zu quadratischen Matrizen Eigenvektoren und Eigenwerte definieren.

**Definition** Es sei $A \in \mathrm{Mat}(n, n; K)$. Ein Vektor $x \in K^n$ heißt *Eigenvektor* von $A$, wenn $x \neq 0$ und es ein $\lambda \in K$ gibt mit

$$Ax = \lambda x.$$

Der Skalar $\lambda$ heißt dann *Eigenwert* von $A$.

**Bemerkung 18.1** Man beachte, dass der Nullvektor ausdrücklich ausgeschlossen ist. Denn für den Nullvektor gilt $f(0) = 0 = \lambda 0$ bzw. $A0 = 0 = \lambda 0$ für jeden Skalar $\lambda$.

**Beispiel 18.1** Es sei $V = \mathbb{R}^2$ und $f : \mathbb{R}^2 \to \mathbb{R}^2$ die Spiegelung an der Geraden $x - y = 0$. Die Abbildung $f$ hat bezüglich der Standardbasis die Darstellungsmatrix

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

Es ist anschaulich klar, dass $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ ein Eigenvektor zum Eigenwert 1 ist und $\begin{pmatrix} -1 \\ 1 \end{pmatrix}$ ein Eigenvektor zum Eigenwert $-1$ ist.

Wir setzen von nun an voraus, dass $V$ ein endlich dimensionaler $K$-Vektorraum ist. Es sei $B$ eine Basis des $K$-Vektorraums $V$. Dann ist ein Vektor $v \in V$ genau dann Eigenvektor der linearen Abbildung $f : V \to V$ zum Eigenwert $\lambda$, wenn der Koordinatenvektor $q_B(v)$ von $v$ bezüglich der Basis $B$ ein Eigenvektor der Darstellungsmatrix $M_B^B(f)$ von $f$ bezüglich der Basis $B$ zu demselben Eigenwert ist. Ein Vektor $v \in K^n$ ist genau dann ein Eigenvektor von $A \in \mathrm{Mat}(n, n; K)$ zum Eigenwert $\lambda$, wenn $v$ Eigenvektor der linearen Abbildung $f : K^n \to K^n$, $f(x) = Ax$, zu demselben Eigenwert ist. Deswegen werden wir im Folgenden nur Eigenvektoren und Eigenwerte einer Matrix betrachten.

Wie bestimmt man nun die Eigenwerte einer $n \times n$-Matrix $A$? Dazu formen wir die Gleichung $Ax = \lambda x$ um:

$$\begin{aligned} & Ax = \lambda x \\ \Leftrightarrow\ & Ax - \lambda x = 0 \\ \Leftrightarrow\ & Ax - \lambda E x = 0 \\ \Leftrightarrow\ & (A - \lambda E)x = 0. \end{aligned}$$

Damit $\lambda$ ein Eigenwert von $A$ ist, muss die Gleichung

$$(A - \lambda E)x = 0$$

eine nichttriviale Lösung besitzen. Nach Satz 6.6 ist das genau dann der Fall, wenn die Matrix $A - \lambda E$ nicht invertierbar ist. Nach Korollar 17.3 ist dies äquivalent zu

$$\det(A - \lambda E) = 0.$$

**Definition** Die Gleichung

$$\det(A - \lambda E) = 0$$

für die Unbekannte $\lambda$ heißt die *charakteristische Gleichung* von $A$. Ihre Lösungen sind die Eigenwerte von $A$. Entwickelt man $\det(A - \lambda E)$, so ergibt sich ein Polynom $p_A(\lambda)$ in $\lambda$, das als *charakteristisches Polynom* von $A$ bezeichnet wird.

Es gilt

$$\begin{aligned} p_A(\lambda) &= \det(A - \lambda E) \\ &= \begin{vmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - \lambda \end{vmatrix} \\ &= (-1)^n \lambda^n + c_1 \lambda^{n-1} + \cdots + c_n. \end{aligned}$$

Das charakteristische Polynom einer $n \times n$-Matrix ist also ein Polynom $n$-ten Grades.

**Beispiel 18.2** Das charakteristische Polynom der Matrix

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

lautet

$$\begin{aligned} p_A(\lambda) &= \det(A - \lambda E) = \begin{vmatrix} a - \lambda & b \\ c & d - \lambda \end{vmatrix} \\ &= (a - \lambda)(d - \lambda) - bc = \lambda^2 - (a + d)\lambda + (ad - bc). \end{aligned}$$

Für die Matrix

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

erhalten wir

$$p_A(\lambda) = \lambda^2 - 1 = (\lambda - 1)(\lambda + 1).$$

Die Eigenwerte von $A$ sind also $+1$ und $-1$ in Übereinstimmung mit Beispiel 18.1.

**Satz 18.1** *Es sei*

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

*eine obere Dreiecksmatrix. Dann sind die Eigenwerte von $A$ die Hauptdiagonalelemente von $A$.*

*Beweis.* Es gilt

$$\begin{aligned} p_A(\lambda) &= \det(A - \lambda E) \\ &= \begin{vmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} - \lambda & \cdots & a_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{nn} - \lambda \end{vmatrix} \\ &= (a_{11} - \lambda)(a_{22} - \lambda) \cdots (a_{nn} - \lambda). \end{aligned}$$

Also sind die Eigenwerte

$$\lambda = a_{11}, \quad \lambda = a_{22}, \quad \ldots, \quad \lambda = a_{nn}.$$

$\square$

Das charakteristische Polynom einer Matrix mit reellen Einträgen kann auch komplexe Nullstellen haben.

**Beispiel 18.3** Die Matrix

$$A = \begin{pmatrix} \cos\frac{\pi}{2} & -\sin\frac{\pi}{2} \\ \sin\frac{\pi}{2} & \cos\frac{\pi}{2} \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

hat das charakteristische Polynom

$$\det(A - \lambda E) = \begin{vmatrix} -\lambda & -1 \\ 1 & -\lambda \end{vmatrix} = \lambda^2 + 1.$$

Dieses Polynom hat die komplexen Nullstellen $\lambda = i$ und $\lambda = -i$.

Wir befassen uns nun mit der Bestimmung der Eigenvektoren zu einem gegebenen Eigenwert.

**Definition** Zu einem Eigenwert $\lambda \in K$ nennen wir

$$\mathrm{Eig}(A, \lambda) := \{x \in K^n \mid Ax = \lambda x\}$$

den *Eigenraum* von $A$ zum Eigenwert $\lambda$.

**Bemerkung 18.2** $\mathrm{Eig}(A, \lambda)$ ist ein Unterraum des $K^n$, denn $\mathrm{Eig}(A, \lambda)$ ist der Lösungsraum des linearen Gleichungssystems $(A - \lambda E)x = 0$.

Zur Bestimmung des Eigenraums zu einem Eigenwert $\lambda$ muss man also das lineare Gleichungssystem

$$(A - \lambda E)x = 0$$

lösen.

**Beispiel 18.4** Wir bestimmen eine Basis des Eigenraums zum Eigenwert 1 der Matrix

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

Dazu müssen wir eine Basis des Lösungsraum des Gleichungssystems

$$\begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

bestimmen. Man sieht leicht, dass dieses lineare Gleichungssystem einen eindimensionalen Lösungsraum mit der Basis $\{\begin{pmatrix} 1 \\ 1 \end{pmatrix}\}$ hat.

# 19 Diagonalisierung

Wir kommen nun auf das eingangs gestellte Problem der Vereinfachung einer quadratischen Matrix zurück.

**Definition** Eine *Diagonalmatrix* $A$ ist eine quadratische Matrix, bei der alle Elemente außer eventuell den Diagonalelementen gleich Null sind, d.h. $A$ hat die Form

$$A = \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & a_{nn} \end{pmatrix}.$$

**Definition** Eine $n \times n$-Matrix $A$ über $K$ heißt *diagonalisierbar*, wenn sie ähnlich zu einer Diagonalmatrix ist, d.h. wenn eine invertierbare Matrix $T$ existiert, so dass $TAT^{-1}$ Diagonalgestalt hat.

**Satz 19.1** *Ähnliche Matrizen haben das gleiche charakteristische Polynom.*

*Beweis.* Es sei $B = TAT^{-1}$. Dann gilt

$$\begin{aligned}\det(B - \lambda E) &= \det(TAT^{-1} - \lambda TET^{-1}) = \det(T(A - \lambda E)T^{-1}) \\ &= \det T \det(A - \lambda E) \det T^{-1} = \det(A - \lambda E).\end{aligned}$$

□

**Korollar 19.1** *Ist $A$ diagonalisierbar, so ist $A$ ähnlich zu einer Diagonalmatrix, bei der auf der Diagonalen die Eigenwerte von $A$ stehen.*

*Beweis.* Dies folgt sofort aus Satz 19.1. □

**Satz 19.2** *Eine Matrix $A \in \mathrm{Mat}(n, n; K)$ ist genau dann diagonalisierbar, wenn sie $n$ linear unabhängige Eigenvektoren besitzt.*

*Beweis.* "$\Rightarrow$": Es sei $A$ diagonalisierbar. Dann gibt es ein $T \in \mathrm{GL}(n; K)$, so dass $TAT^{-1}$ Diagonalgestalt hat. Also ist $TAT^{-1} = D$ mit

$$D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix}.$$

Aus $TAT^{-1} = D$ folgt $AT^{-1} = T^{-1}D$. Es sei $(e_1, \ldots, e_n)$ die Standardbasis von $K^n$. Dann gilt

$$A(T^{-1}(e_i)) = T^{-1}(\lambda_i e_i) = \lambda_i T^{-1}(e_i) \quad \text{für } i = 1, \ldots, n.$$

Also ist $T^{-1}(e_i)$ ein Eigenvektor von $A$ zum Eigenwert $\lambda_i$, $i = 1, \ldots, n$. Da die Vektoren $e_1, \ldots, e_n$ linear unabhängig sind und $T^{-1}$ invertierbar ist, sind auch die Vektoren $T^{-1}(e_1), \ldots, T^{-1}(e_n)$ linear unabhängig.

"$\Leftarrow$": Es seien $v_1, \ldots, v_n$ die linear unabhängigen Eigenvektoren von $A$ zu den Eigenwerten $\lambda_1, \ldots, \lambda_n$. Es sei $B$ die Standardbasis von $K^n$ und $B' = (v_1, \ldots, v_n)$. Es sei $D$ die Matrix

$$D := \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix}.$$

Dann ist $D$ die Darstellungsmatrix $M_{B'}^{B'}(f)$ der linearen Abbildung $f : K^n \to K^n$ mit $f(x) = Ax$ bezüglich der Basis $B'$. Nach Satz 16.4 gilt

$$A = T_B^{B'} D (T_B^{B'})^{-1}.$$

Also ist $A$ diagonalisierbar. □

**Beispiel 19.1** Wir untersuchen die Matrix

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

auf Diagonalisierbarkeit. Das charakteristische Polynom

$$\det(A - \lambda E) = (1 - \lambda)^2$$

liefert 1 als einzigen Eigenwert, aber der Eigenraum $\mathrm{Eig}(A, 1)$ zum Eigenwert 1 wird von dem Vektor $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ erzeugt. Also besitzt die Matrix $A$ nur einen linear unabhängigen Eigenvektor. Nach Satz 19.2 ist $A$ nicht diagonalisierbar.

**Satz 19.3** *Es sei $A \in \mathrm{Mat}(n, n; K)$. Es seien $\lambda_1, \ldots, \lambda_k$ paarweise verschiedene Eigenwerte von $A$ und $v_1, \ldots, v_k$ zugehörige Eigenvektoren. Dann sind die Vektoren $v_1, \ldots, v_k$ linear unabhängig.*

*Beweis.* Wir führen Induktion über $k$ durch.

**Induktionsanfang:** Da ein Eigenvektor $v_1$ nach Definition verschieden vom Nullvektor ist, ist $v_1$ linear unabhängig. Damit ist der Satz für k=1 richtig.

**Induktionsschritt:** Es sei $k \geq 2$. Wir nehmen an, dass die Aussage bereits für $k - 1$ bewiesen ist. Wir müssen zeigen, dass $v_1, \ldots, v_k$ linear unabhängig sind. Dazu betrachten wir die Gleichung

$$\alpha_1 v_1 + \cdots + \alpha_k v_k = 0. \tag{6}$$

Wir multiplizieren diese Gleichung einerseits mit $A$ und andererseits mit $\lambda_1$:

$$\begin{aligned} \alpha_1 \lambda_1 v_1 + \alpha_2 \lambda_2 v_2 + \cdots + \alpha_k \lambda_k v_k &= 0 \\ \alpha_1 \lambda_1 v_1 + \alpha_2 \lambda_1 v_2 + \cdots + \alpha_k \lambda_1 v_k &= 0 \end{aligned}$$

Subtrahieren wir die zweite Gleichung von der ersten, so ergibt sich

$$\alpha_2(\lambda_2 - \lambda_1) v_2 + \cdots + \alpha_k(\lambda_k - \lambda_1) v_k = 0.$$

Nach Induktionsannahme sind die Vektoren $v_2, \dots, v_k$ linear unabhängig. Daraus folgt

$$\alpha_2(\lambda_2 - \lambda_1) = \dots = \alpha_k(\lambda_k - \lambda_1) = 0.$$

Da nach Voraussetzung die Eigenwerte paarweise verschieden sind, ergibt sich daraus

$$\alpha_2 = \dots = \alpha_k = 0.$$

Setzt man dies in Gleichung (6) ein, so folgt

$$\alpha_1 v_1 = 0$$

und daraus $\alpha_1 = 0$, da $v_1 \neq 0$. □

Aus Satz 19.3 erhalten wir als Folgerung:

**Satz 19.4** *Besitzt eine $n \times n$-Matrix $A$ $n$ paarweise verschiedene Eigenwerte, so ist $A$ diagonalisierbar.*

*Beweis.* Es seien $v_1, \dots, v_n$ Eigenvektoren zu den paarweise verschiedenen Eigenwerten $\lambda_1, \dots, \lambda_n$. Nach Satz 19.3 sind $v_1, \dots, v_n$ linear unabhängig. Nach Satz 19.2 ist $A$ deshalb diagonalisierbar. □

Satz 19.4 liefert nur ein hinreichendes, aber kein notwendiges Kriterium für die Diagonalisierbarkeit.

# 20 Euklidische und unitäre Vektorräume

Wir wollen das Skalarprodukt auf dem $\mathbb{R}^n$ verallgemeinern.

**Definition** Es sei $V$ ein $\mathbb{R}$-Vektorraum. Eine Abbildung

$$\begin{array}{rrcl} \langle\ ,\ \rangle : & V \times V & \longrightarrow & \mathbb{R} \\ & (v, w) & \longmapsto & \langle v, w \rangle \end{array}$$

heißt

(i) eine *Bilinearform*, falls für alle $v, v', w, w' \in V$ und $\lambda \in \mathbb{R}$ gilt

$$\begin{array}{ll} \langle v + v', w \rangle = \langle v, w \rangle + \langle v', w \rangle & \quad \langle \lambda v, w \rangle = \lambda \langle v, w \rangle \\ \langle v, w + w' \rangle = \langle v, w \rangle + \langle v, w' \rangle & \quad \langle v, \lambda w \rangle = \lambda \langle v, w \rangle, \end{array}$$

(ii) *symmetrisch*, falls für alle $v, w \in V$

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

(iii) *positiv definit*, falls

$$\langle v, v\rangle > 0 \text{ für alle } v \neq 0.$$

Eine positiv definite symmetrische Bilinearform nennt man auch ein *Skalarprodukt*.

Ein *euklidischer Vektorraum* ist ein $\mathbb{R}$-Vektorraum zusammen mit einem Skalarprodukt.

**Beispiel 20.1** Nach Satz 3.1 ist der $\mathbb{R}^n$ mit dem gewöhnlichen Skalarprodukt

$$\langle x, y\rangle := x^T y = (x_1, \ldots, x_n)\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \sum_{i=1}^{n} x_i y_i$$

ein euklidischer Vektorraum.

**Beispiel 20.2** Es sei $V = C[0,1] = \{f \mid f : [0,1] \to \mathbb{R} \text{ stetig}\}$ und

$$\langle f, g\rangle := \int_0^1 f(t)g(t)dt.$$

Nach den Rechenregeln für Integrale ist $\langle\ ,\ \rangle$ eine symmetrische Bilinearform. Wegen

$$\langle f, f\rangle = \int_0^1 f^2(t)dt > 0 \text{ für } f \not\equiv 0$$

ist diese Bilinearform positiv definit, also ein Skalarprodukt auf $V$.

Wir wollen nun auch ein Skalarprodukt auf dem $\mathbb{C}^n$ einführen. Warum definiert die Bilinearform von Beispiel 20.1 kein Skalarprodukt auf dem $\mathbb{C}^n$?

**Definition** Es sei $V$ ein $\mathbb{C}$-Vektorraum. Eine Abbildung

$$\begin{array}{rrcl} \langle\ ,\ \rangle : & V \times V & \longrightarrow & \mathbb{C} \\ & (v, w) & \longmapsto & \langle v, w\rangle \end{array}$$

heißt

(i) eine *Sesquilinearform*, falls für alle $v, v', w, w' \in V$ und $\lambda \in \mathbb{C}$ gilt

$$\begin{array}{rclcrcl} \langle v + v', w\rangle &=& \langle v, w\rangle + \langle v', w\rangle & \qquad & \langle \lambda v, w\rangle &=& \lambda\langle v, w\rangle \\ \langle v, w + w'\rangle &=& \langle v, w\rangle + \langle v, w'\rangle & \qquad & \langle v, \lambda w\rangle &=& \overline{\lambda}\langle v, w\rangle, \end{array}$$

(ii) *hermitesch*, falls für alle $v, w \in V$

$$\langle v, w\rangle = \overline{\langle w, v\rangle},$$

(iii) *positiv definit*, falls

$$\langle v, v\rangle > 0 \text{ für alle } v \neq 0.$$

Eine positiv definite hermitesche Sesquilinearform auf einem $\mathbb{C}$-Vektorraum $V$ nennt man auch ein *Skalarprodukt* auf $V$.

Ein *unitärer Vektorraum* ist ein $\mathbb{C}$-Vektorraum zusammen mit einem Skalarprodukt.

**Beispiel 20.3** Man rechnet leicht nach, dass $\mathbb{C}^n$ mit der Sesquilinearform

$$\langle z, w\rangle := z^T\overline{w} = (z_1, \ldots, z_n)\begin{pmatrix} \overline{w}_1 \\ \vdots \\ \overline{w}_n \end{pmatrix} = \sum_{i=1}^{n} z_i\overline{w}_i$$

ein unitärer Vektorraum ist.

**Beispiel 20.4** Es sei $f : [0,1] \to \mathbb{C}$ eine Funktion. Dann ist $f = u + iv$ für reelle Funktionen $u, v : [0,1] \to \mathbb{R}$. Die Funktion $f$ heißt *stetig*. falls die reellen Funktionen $u$ und $v$ stetig sind. Das Integral von $f$ über $[0,1]$ ist definiert als

$$\int_0^1 f(t)dt := \int_0^1 u(t)dt + i\int_0^1 v(t)dt \in \mathbb{C}.$$

Es sei $V = \{f : [0,1] \to \mathbb{C} \text{ stetig}\}$ und

$$\langle f, g\rangle := \int_0^1 f(t)\overline{g(t)}dt.$$

Dann definiert $\langle\ ,\ \rangle$ ein Skalarprodukt auf $V$.

**Definition** (a) Eine $n \times n$-Matrix über $\mathbb{R}$ heißt *symmetrisch*, falls

$$A = A^T.$$

(b) Eine $n \times n$-Matrix über $\mathbb{C}$ heißt *hermitesch*, falls

$$A = \overline{A}^T.$$

(Hierbei ist $\overline{A}$ wie folgt definiert: Ist $A = (a_{ij})$, so ist $\overline{A} = (\overline{a}_{ij})$)

Es sei im Folgenden $K = \mathbb{R}$ oder $K = \mathbb{C}$. Für eine symmetrische (hermitesche) $n \times n$-Matrix $A$ über $K$ setzen wir

$$\begin{array}{rccl} \langle\ ,\ \rangle_A : & K^n \times K^n & \longrightarrow & K \\ & (x, y) & \longmapsto & \langle x, y\rangle_A := x^T A\overline{y} \end{array} .$$

Man beachte, dass die komplexe Konjugation auf $\mathbb{R}$ die Identität ist. Deswegen können wir auf diese Weise den symmetrischen und den hermiteschen Fall parallel behandeln.

**Lemma 20.1** *Ist $A$ symmetrisch (hermitesch), so ist $\langle\ ,\ \rangle_A$ eine symmetrische Bilinearform (hermitesche Sesquilinearform).*

*Beweis.*

$$\langle x, y\rangle_A = x^T A\overline{y} = \overline{y}^T A^T x = \overline{y^T \overline{A}^T \overline{x}} = \overline{y^T A\overline{x}} = \overline{\langle y, x\rangle_A}.$$

□

Es sei nun $V$ ein endlich dimensionaler $K$-Vektorraum mit einer fest gewählten geordneten Basis $B = (v_1, \ldots, v_n)$. Dann können wir einer symmetrischen (hermiteschen) $n \times n$-Matrix $A$ wie folgt eine symmetrische Bilinearform (hermitesche Sesquilinearform) auf $V$ zuordnen: Es seien $v, w \in V$ und $x = q_B(v)$ der Koordinatenvektor zu $v$ und $y = q_B(w)$ der Koordinatenvektor zu $w$ bezüglich der Basis $B$. Dann setzen wir

$$\langle v, w\rangle_A := x^T A\overline{y}.$$

Es sei nun umgekehrt

$$\langle\ ,\ \rangle : V \times V \to K$$

eine symmetrische Bilinearform (hermitesche Sesquilinearform). Dann ordnen wir dieser eine $n \times n$-Matrix $A$ wie folgt zu:

$$A = (\langle v_i, v_j\rangle).$$

**Definition** Die Matrix $A$ heißt die *Darstellungsmatrix* von $\langle\ ,\ \rangle$ bezüglich der Basis $B$.

**Lemma 20.2** *Die Darstellungsmatrix $A$ einer symmetrischen Bilinearform (hermiteschen Sesquilinearform) ist symmetrisch (hermitesch).*

*Beweis.* Dies folgt aus

$$a_{ij} = \langle v_i, v_j \rangle = \overline{\langle v_j, v_i \rangle} = \overline{a}_{ji}.$$

□

Daraus ergibt sich, dass wir nach Wahl einer Basis $B$ eine bijektive Beziehung zwischen symmetrischen (hermiteschen) Matrizen und symmetrischen Bilinearformen (hermiteschen Sesquilinearformen) auf $V$ haben.

Wir diskutieren nun, wie sich die Darstellungsmatrix einer symmetrischen Bilinearform (hermiteschen Sesquilinearform) $\langle\ ,\ \rangle$ ändert, wenn wir von der Basis $B$ zu einer anderen Basis $B'$ übergehen.

**Satz 20.1 (Transformationsformel)** *Es sei $V$ ein endlich dimensionaler $K$-Vektorraum, $B$, $B'$ Basen von $V$ und $\langle\ ,\ \rangle$ eine symmetrische Bilinearform (hermitesche Sesquilinearform) auf $V$. Es sei $A$ bzw. $A'$ die Darstellungsmatrix von $\langle\ ,\ \rangle$ bezüglich $B$ bzw. $B'$. Schließlich sei $S = T^B_{B'}$ die Transformationsmatrix von $B$ nach $B'$. Dann gilt*

$$A = S^T A' \overline{S}.$$

*Beweis.* Es seien $v, w \in V$ und $x = q_B(v)$ bzw. $y = q_B(w)$ die Koordinatenvektoren von $v$ bzw. $w$ bezüglich der Basis $B$. Dann sind $Sx$ bzw. $Sy$ die Koordinatenvektoren von $v$ bzw. $w$ bezüglich der Basis $B'$. Dann gilt auf der einen Seite

$$\langle v, w \rangle = x^T A \overline{y},$$

auf der anderen Seite

$$\langle v, w \rangle = (Sx)^T A' \overline{(Sy)} = x^T (S^T A' \overline{S}) \overline{y}.$$

Speziell für $v = v_i$, $w = v_j$ erhält man

$$\begin{aligned} \langle v_i, v_j \rangle &= e_i^T A \overline{e_j} = a_{ij}, \\ \langle v_i, v_j \rangle &= e_i^T (S^T A' \overline{S}) \overline{e_j} = c_{ij}, \end{aligned}$$

wobei $A = (a_{ij})$ und $S^T A' \overline{S} = (c_{ij})$. Daraus folgt

$$A = S^T A' \overline{S}.$$

□

Man vergleiche diese Formel mit der Transformationsformel für lineare Abbildungen (Satz 16.4).

Wie in §3 definieren wir:

**Definition** Es sei $V$ ein euklidischer (unitärer) Vektorraum. Für $v \in V$ setzen wir

$$||v|| := \sqrt{\langle v, v \rangle}.$$

Wir nennen $||v||$ die *Norm* von $v$.

Wie in §3 kann man zeigen, dass die Abbildung

$$\begin{array}{rccc} ||\ ||: & V & \longrightarrow & \mathbb{R} \\ & v & \longmapsto & ||v|| \end{array}$$

die Eigenschaften von Satz 3.5 hat.

**Definition** Es sei $V$ ein euklidischer (unitärer) Vektorraum.

(a) Zwei Vektoren $v, w \in V$ heißen *orthogonal*, in Zeichen $v \perp w$, falls $\langle v.w \rangle = 0$.

(b) Eine Teilmenge $S = \{v_1, \ldots, v_n\}$ von $V$ heißt *orthogonal*, wenn die Vektoren $v_i$ paarweise orthogonal sind, d.h. wenn $v_i \perp v_j$ für alle $i \neq j$, $i, j = 1, \ldots, n$, gilt.

(c) Eine Teilmenge $S = \{v_1, \ldots, v_n\}$ von $V$ heißt *orthonormal*, wenn $S$ orthogonal ist und alle Vektoren auf die Länge 1 normiert sind, d.h. wenn

$$\langle v_i, v_j \rangle = 0 \text{ für alle } i \neq j \text{ und } \langle v_i, v_i \rangle = 1 \text{ für } i, j = 1, \ldots, n.$$

(d) Eine Basis $S = \{v_1, \ldots, v_n\}$ von $V$ heißt *Orthonormalbasis* (abgekürzt *ON-Basis*), wenn sie orthonormal ist.

Es sei $V$ ein euklidischer (unitärer) Vektorraum.

**Lemma 20.3** *Eine orthogonale Teilmenge $S = \{v_1, \ldots, v_n\}$ von $V$ mit vom Nullvektor verschiedenen Elementen ist linear unabhängig.*

*Beweis.* Es sei

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

Dann gilt

$$\langle \alpha_1 v_1 + \cdots + \alpha_n v_n, v_i \rangle = \alpha_i \langle v_i, v_i \rangle = 0.$$

Da nach Vorausetzung $v_i \neq 0$, folgt daraus $\alpha_i = 0$. Dies gilt für alle $i = 1, \ldots, n$. □

**Lemma 20.4** *Es sei $W$ ein Unterraum von $V$ und $\{v_1, \dots, v_m\}$ eine Orthonormalbasis von $W$. Es sei $u \in V$. Dann ist der Vektor*

$$v := u - \langle u, v_1 \rangle v_1 - \cdots - \langle u, v_m \rangle v_m$$

*orthogonal zu allen Vektoren von $W$.*

**Definition** Der Vektor

$$u - \langle u, v_1 \rangle v_1 - \cdots - \langle u, v_m \rangle v_m$$

heißt die *zu $W$ orthogonale Komponente von $u$.*

*Beweis.* Es gilt

$$\langle v, v_i \rangle = \langle u - \langle u, v_1 \rangle v_1 - \cdots - \langle u, v_m \rangle v_m, v_i \rangle = \langle u, v_i \rangle - \langle u, v_i \rangle = 0.$$

□

**Satz 20.2 (Orthonormalisierungssatz)** *Es sei $V$ ein endlich dimensionaler euklidischer (unitärer) Vektorraum. Es sei $W \subset V$ ein Unterraum und $\{w_1, \dots, w_m\}$ eine ON-Basis von $W$. Dann kann man diese zu einer ON-Basis $\{w_1, \dots, w_m, w_{m+1}, \dots, w_n\}$ von $V$ ergänzen.*

*Beweis.* Der Beweis wird durch das *E. Schmidt'sche Orthonormalisierungsverfahren* gegeben.

Ist $V = W$, so ist man fertig. Andernfalls gibt es ein $v \in V \setminus W$. Wir betrachten die zu $W$ orthogonale Komponente von $v$:

$$w := v - \langle v, w_1 \rangle w_1 - \cdots - \langle v, w_m \rangle w_m.$$

Da $v \notin W$, ist $w \neq 0$. Nach Lemma 20.4 ist $w$ orthogonal zu allen Vektoren aus $W$. Wir setzen nun

$$w_{m+1} := \frac{w}{\|w\|}.$$

Dann ist $\{w_1, \dots, w_{m+1}\}$ eine orthogonale Teilmenge, deren Elemente vom Nullvektor verschieden sind. Nach Lemma 20.3 ist diese Teilmenge linear unabhängig und damit eine ON-Basis von

$$W' := \operatorname{Span}\{w_1, \dots, w_{m+1}\}.$$

Ist $W' = V$, so ist man fertig. Andernfalls kommt man nach endlich vielen Schritten zum Ziel. □

**Korollar 20.1** *Jeder endlich dimensionale euklidische (unitäre) Vektorraum besitzt eine Orthonormalbasis.*

*Beweis.* Man wende das obige Verfahren mit $W = \{0\}$ an. □

# 21 Orthogonale und unitäre Endomorphismen

Im Folgenden sei wieder $K = \mathbb{R}, \mathbb{C}$ und $(V, \langle\ ,\ \rangle)$ ein euklidischer (unitärer) Vektorraum.

**Definition** Es sei $V$ ein euklidischer (unitärer) Vektorraum. Ein Endomorphismus $f : V \to V$ heißt *orthogonal* (*unitär*), falls gilt:

$$\langle f(v), f(w)\rangle = \langle v, w\rangle \quad \text{für alle } v, w \in V.$$

**Satz 21.1** *Für einen orthogonalen (unitären) Endomorphismus $f : V \to V$ gilt:*

(i) $\|f(v)\| = \|v\|$ *für alle* $v \in V$.

(ii) $v \perp w \Leftrightarrow f(v) \perp f(w)$.

(iii) $f$ *ist injektiv.*

(iv) *Ist $V$ endlich dimensional, so ist $f$ ein Automorphismus und $f^{-1}$ ist ebenfalls orthogonal (unitär).*

(v) *Ist $\lambda$ ein Eigenwert von $f$, so ist $|\lambda| = 1$.*

*Beweis.* Die Aussagen (i),(ii) sind klar. Aus (i) folgt, dass $f$ injektiv ist. Aus Satz 14.5 folgt damit (iv).

(v) Es sei $v$ Eigenvektor von $f$ zum Eigenwert $\lambda$. Dann gilt

$$\|v\| = \|f(v)\| = \|\lambda v\| = |\lambda| \|v\|.$$

Wegen $v \neq 0$ folgt daraus $|\lambda| = 1$. □

**Satz 21.2** *Ein Endomorphismus $f : V \to V$ ist genau dann orthogonal (unitär), wenn $f$ eine* Isometrie *ist, d.h. $\|f(v)\| = \|v\|$ für alle $v \in V$ gilt.*

Zum Beweis benutzen wir den folgenden Zusammenhang zwischen dem Skalarprodukt und der Norm. Man kann $\langle\ ,\ \rangle$ aus $\|\ \|$ zurückgewinnen. Dies nennt man *Polarisierung*:

$$K = \mathbb{R}: \qquad \langle v, w\rangle = \frac{1}{4}\left(\|v+w\|^2 - \|v-w\|^2\right),$$
$$K = \mathbb{C}: \qquad \langle v, w\rangle = \frac{1}{4}\left(\|v+w\|^2 - \|v-w\|^2 + i\|v+iw\|^2 - i\|(v-iw\|^2\right).$$

(Beweis durch Nachrechnen.)

*Beweis von Satz 21.2.* "$\Rightarrow$" ist Satz 21.1 (i).

"$\Leftarrow$". Es sei $K = \mathbb{R}$. (Der Fall $K = \mathbb{C}$ kann analog bewiesen werden.) Wir wenden Polarisierung an:

$$\begin{aligned} \langle f(v), f(w) \rangle &= \frac{1}{4}\left(\|f(v)+f(w)\|^2 - \|f(v)-f(w)\|^2\right) \\ &= \frac{1}{4}\left(\|f(v+w)\|^2 - \|f(v-w)\|^2\right) \\ &= \frac{1}{4}\left(\|v+w\|^2 - \|v-w\|^2\right) = \langle v, w \rangle. \end{aligned}$$

□

**Definition** (a) Eine Matrix $A \in \mathrm{GL}(n;\mathbb{R})$ heißt *orthogonal*, wenn gilt:

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

(b) Eine Matrix $A \in \mathrm{GL}(n;\mathbb{C})$ heißt *unitär*, wenn gilt:

$$A^{-1} = \overline{A}^T.$$

**Lemma 21.1** *Für eine orthogonale (unitäre) Matrix $A$ gilt $|\det A| = 1$.*

*Beweis.* Aus $A\overline{A}^T = E$ folgt:

$$\det(A\overline{A}^T) = \det A \det \overline{A}^T = (\det A)\overline{(\det A)} = |\det A|^2 = 1.$$

□

**Definition**

(a) $O(n) := \{A \in \mathrm{GL}(n;\mathbb{R}) \mid A^{-1} = A^T\}$ (*orthogonale Gruppe*)

(b) $SO(n) := \{A \in O(n) \mid \det A = 1\}$ (*spezielle orthogonale Gruppe*)

(c) $U(n) := \{A \in \mathrm{GL}(n;\mathbb{C}) \mid A^{-1} = \overline{A}^T\}$ (*unitäre Gruppe*)

(d) $SU(n) := \{A \in U(n) \mid \det A = 1\}$ (*spezielle unitäre Gruppe*)

Die angegebenen Mengen sind Untergruppen von $\mathrm{GL}(n;K)$ (Beweis Übungsaufgabe).

**Satz 21.3** *Für $A \in \mathrm{Mat}(n,n;K)$ sind die folgenden Aussagen äquivalent:*

(i) *$A$ ist orthogonal (unitär).*

(ii) *Die Spaltenvektoren von $A$ bilden eine ON-Basis von $K^n$.*

(iii) *Die Zeilenvektoren von $A$ bilden eine ON-Basis von $K^n$.*

*Beweis.* (ii) bedeutet $A^T\overline{A} = E$.
(iii) bedeutet $A\overline{A}^T = E$. □

**Satz 21.4** *Es sei $V$ ein euklidischer (unitärer) Vektorraum mit einer ON-Basis $B$ und $f$ ein Endomorphismus von $V$. Dann ist $f$ genau dann orthogonal (unitär), wenn die Darstellungsmatrix $M_B^B(f)$ bezüglich der Basis $B$ orthogonal (unitär) ist.*

*Beweis.* Es seien $v, w \in V$ und $x$ bzw. $y$ die Koordinatenvektoren von $v$ bzw. $w$ bezüglich der Basis $B$. Dann gilt

$$\langle v, w\rangle = x^T E\overline{y} = x^T\overline{y},$$

da $E$ die Darstellungsmatrix des Skalarprodukts bezüglich einer ON-Basis ist.

Setze $A := M_B^B(f)$. Dann gilt:

$$\langle f(v), f(w)\rangle = \langle v, w\rangle \Leftrightarrow (Ax)^T\overline{(Ay)} = x^T\overline{y} \Leftrightarrow A^T\overline{A} = E.$$

□

Wir betrachten nun orthogonale Abbildungen $f : \mathbb{R}^n \to \mathbb{R}^n$ für $n = 1, 2, 3$.
a) Im Fall $n = 1$ gibt es nur die Möglichkeiten

$$f(x) = \pm x.$$

b) Im Fall $n = 2$ klassifizieren wir die orthogonalen $2 \times 2$-Matrizen.

**Satz 21.5** *Ist $A$ eine orthogonale $2\times 2$-Matrix, so gibt es ein $\alpha \in [0, 2\pi)$, so dass*

$$A = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \text{ oder } A = \begin{pmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{pmatrix}.$$

*Beweis.* Es sei

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

Aus $A^T A = E$ folgt

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

Wegen $a^2 + c^2 = 1$ und $b^2 + d^2 = 1$ gibt es $\alpha, \alpha' \in [0, 2\pi)$, so dass

$$a = \cos\alpha, \quad c = \sin\alpha, \quad b = \sin\alpha', \quad d = \cos\alpha'.$$

Wegen $ab + cd = 0$ gilt

$$0 = \cos\alpha \cdot \sin\alpha' + \sin\alpha \cdot \cos\alpha' = \sin(\alpha + \alpha').$$

Also ist $\alpha+\alpha'$ entweder ein geradzahliges oder ein ungeradzahliges Vielfaches von $\pi$. Im ersten Fall ist

$$b = \sin\alpha' = -\sin\alpha \text{ und } d = \cos\alpha' = \cos\alpha$$

und im zweiten Fall ist

$$b = \sin\alpha' = \sin\alpha \text{ und } d = \cos\alpha' = -\cos\alpha.$$

□

Im ersten Fall ist die zugehörige orthogonale Abbildung eine Drehung um den Winkel $\alpha$, im zweiten Fall eine Spiegelung an der um den Winkel $\alpha/2$ gedrehten $x$-Achse. Im ersten Fall ist $A$ dann und nur dann diagonalisierbar, wenn $\alpha = 0$ oder $\alpha = \pi$ ist. Im zweiten Fall ist $A$ diagonalisierbar mit den Eigenwerten $+1$ und $-1$.

c) Es sei nun $f : \mathbb{R}^3 \to \mathbb{R}^3$ eine orthogonale Abbildung. Dann hat das charakteristische Polynom einer Darstellungsmatrix von $f$ den Grad 3. Also hat das charakteristische Polynom mindestens eine reelle Nullstelle. Also hat $f$ einen Eigenwert $\lambda_1$. Nach Satz 21.1 gilt $\lambda_1 = \pm 1$. Es sei $v_1 \in \mathbb{R}^3$ ein Eigenvektor zu dem Eigenwert $\lambda_1$ mit $\|v_1\| = 1$. Mit Hilfe des E. Schmidt'schen Orthonormalisierungsverfahrens (Satz 20.2) können wir $\{v_1\}$ zu einer Orthonormalbasis $B = (v_1, v_2, v_3)$ des $\mathbb{R}^3$ ergänzen. Es sei $W$ die von $v_2$ und $v_3$ aufgespannte Ebene. Aus Satz 21.1 folgt, dass $f(W) = W$. Es sei $A := M_B^B(f)$. Dann gilt

$$A = \begin{pmatrix} \lambda_1 & 0 \quad 0 \\ 0 & \\ 0 & \boxed{A'} \end{pmatrix}.$$

Aus Satz 21.4 folgt, dass $A'$ eine orthogonale $2 \times 2$-Matrix ist. Außerdem gilt

$$\det A = \lambda_1 \det A'.$$

Nun müssen wir Fallunterscheidungen machen.

Es sei $\det A = +1$. Ist $\lambda_1 = -1$, so muss $\det A' = -1$ gelten. Damit hat $A'$ die Eigenwerte $\lambda_2 = +1$ und $\lambda_3 = -1$ und wir können $v_2$ als Eigenvektor zu $\lambda_2$ und $v_3$ als Eigenvektor zu $\lambda_3$ wählen. Damit gilt

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

Die zugehörige orthogonale Abbildung ist eine Drehung des $\mathbb{R}^3$ um die von $v_2$ aufgespannte Gerade um den Winkel 180°.

Ist $\lambda_1 = +1$, so muss auch $\det A' = +1$ sein. Also gibt es nach Satz 21.5 ein $\alpha \in [0, 2\pi)$, so dass

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{pmatrix}.$$

Die zugehörige orthogonale Abbildung ist eine Drehung des $\mathbb{R}^3$ um die von $v_1$ aufgespannte Gerade um den Winkel $\alpha$.

Ist $\det A = -1$, so ergeben sich die Möglichkeiten

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \text{ und } A = \begin{pmatrix} -1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{pmatrix}.$$

Die zugehörige orthogonale Abbildung ist im ersten Fall eine Spiegelung an der von $v_1$ und $v_2$ aufgespannten Ebene. Im zweiten Fall handelt es sich um eine Drehspiegelung: Die Vektoren des $\mathbb{R}^3$ werden an der von $v_2$ und $v_3$ aufgespannten Ebene gespiegelt und anschließend um den Winkel $\alpha$ um die von $v_1$ aufgespannte Gerade gedreht.

Diese Überlegungen haben eine erstaunliche Konsequenz:

**Satz 21.6 (Satz vom Fußball)** *Bei jedem Fußballspiel, in dem nur ein Ball benutzt wird, gibt es zwei Punkte auf der Oberfläche des Balles, die sich zu Beginn der ersten und der zweiten Halbzeit (wenn der Ball genau auf dem Anstoßpunkt liegt) an der gleichen Stelle im umgebenden Raum befinden.*

*Beweis.* Die beiden Positionen des Balles unterscheiden sich durch eine orthogonale Abbildung $f$ des Balles mit der Determinante $+1$. Eine solche Abbildung hat stets den Eigenwert $+1$, d.h. $f$ lässt eine Gerade fest. Die Gerade schneidet die Oberfläche des Fußballs in zwei Punkten. $\square$

# Inhaltsverzeichnis