Lineare Algebra A

Wintersemester 2001/2002

W. Ebeling

Institut für Mathematik
Universität Hannover
Postfach 6009
30060 Hannover
E-mail: ebeling@math.uni-hannover.de

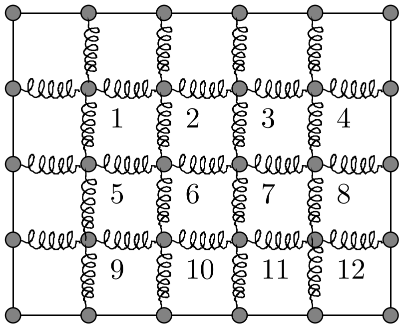

Abbildung 1: Federrost ($n = 12$)

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

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

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

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

Wir geben drei Beispiele an, die zeigen, wie man auf solche Gleichungssysteme 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

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

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*.

Beispiel 1.2 Es soll das Verhalten einer Brücke, eines Flächentragwerks oder einer schwingfähigen Membran unter Einwirkung von äusseren Kräften studiert werden. Dazu bedient man sich eines Modells. Man ersetzt die Membran durch ein Federrost ähnlich einem alten Bettgestell. Das Federrost bestehe aus einem Gitter regelmäßig im Schachbrettmuster verspannter Federn (siehe Abbildung 1 für $n = 12$). Die Anzahl der Gitterpunkte sei n. Die Kraft, die in einem Gitterpunkt wirkt, ist proportional zur Auslenkung. Der Einfachheit halber lassen wir die Proportionalitätsfaktoren weg. Mit x_i bezeichnen wir die Auslenkung im i-ten Gitterpunkt in Bezug auf eine Bezugsebene. Von aussen wirke eine Kraft b_i auf den i-ten Gitterpunkt. Dann erhalten wir die Gleichung

$$\begin{aligned} b_i &= (x_{\text{rechts}} - x_i) + (x_{\text{links}} - x_i) + (x_{\text{oben}} - x_i) + (x_{\text{unten}} - x_i) \\ &= x_{\text{rechts}} + x_{\text{links}} + x_{\text{oben}} + x_{\text{unten}} - 4x_i \end{aligned}$$

Solche Gleichungssysteme spielen bei *Mehrgitterverfahren* eine Rolle. Je feiner man das Gitter wählt, desto genauer stimmt das Modell mit der Wirklichkeit überein.

Beispiel 1.3 Wir betrachten das Problem, einen optimalen Fahrplan zu finden: Es sei ein Streckennetz mitsamt Fahrzeiten für jede Einzelstrecke, zu garantierenden Anschlüssen und Mindestumsteigezeiten gegeben. Wie setzt man die Abfahrtszeiten der Züge so, dass sie sich zum Beispiel im Stundentakt wiederholen, alle Anschlüsse erreicht werden und Verspätungen keine allzu großen Auswirkungen haben? Abstrakt gesprochen geht es um die Organisation von Prozessen (hier eine Zugfahrt von A), die erst stattfinden können, nachdem gewisse andere Prozesse (Anschlussfahrten nach A) abgeschlossen sind. Gleiches gilt für die Organisation von Prozeduren in einem Parallelrechner, wo auch die einzelnen Prozessoren Zwischenergebnisse von anderen Prozessoren benötigen. Man spricht von dynamischen Systemen mit diskreten Ereignissen (*discrete event dynamical systems*, DEDS). Es haben nun Mathematiker herausgefunden, dass man das Problem des optimalen Fahrplans auch auf die Lösung eines geeigneten linearen Gleichungssystems zurückführen kann. Allerdings darf dabei nicht wie gewöhnlich gerechnet werden, sondern man muss eine geignete neue Addition und Multiplikation einführen, man rechnet in einer sogenannten Max-Plus-Algebra.

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 Calculus A 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

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

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

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

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

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 $(x_1, \dots, x_n) \in \mathbb{R}^n$ und $(y_1, \dots, y_n) \in \mathbb{R}^n$ definieren wir eine *Addition*

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

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

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

Wir setzen ausserdem:

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

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

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

Man sieht sofort die folgenden Regeln ein:

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

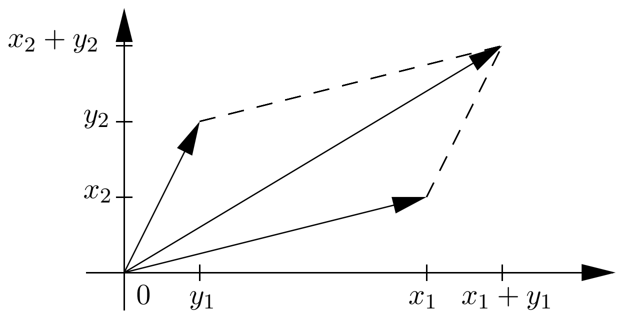

Abbildung 2: Addition zweier Vektoren ($n = 2$)

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

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

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

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

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

$$\mathbf{v} + \mathbb{R}\mathbf{w} := \{\mathbf{x} = \mathbf{v} + \lambda\mathbf{w} \mid \lambda \in \mathbb{R}\}.$$

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

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 $\mathbf{v}, \mathbf{w} \in \mathbb{R}^2$ mit $\mathbf{w} \neq \mathbf{0}$ gibt, so dass*

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

Beweis. 1) Es sei

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

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

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

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

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

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

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

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 $L = \mathbf{v} + \mathbb{R}\mathbf{w}$ gegeben. Wir müssen eine Gleichung finden. Ist $\mathbf{v} = (v_1, v_2)$ und $\mathbf{w} = (w_1, w_2)$ 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

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

Wir müssen wieder $L = L'$ zeigen.

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

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

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 $\mathbf{x} \in L'$.

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

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

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 λ durch diese Gleichung. Wegen $\mathbf{x} = (x_1, x_2) \in L'$ gilt

$$w_2x_1 - w_1x_2 = w_2v_1 - w_1v_2.$$

Daraus folgt

$$x_2 = v_2 + \frac{w_2x_1 - w_2v_1}{w_1} = v_2 + \lambda w_2.$$

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

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

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

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

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

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

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

Beweis. 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. (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

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

Wie im Beweis von Satz 2.1 kann man dann zeigen

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

□

Wir betrachten nun den Schnitt von Ebenen.

Beispiel 2.4 Wir betrachten die Gleichungen

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

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 λ 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

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

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

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:

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

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

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

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 $\mathbf{x} = (x_1, \dots, x_n)$ und $\mathbf{y} = (y_1, \dots, y_n)$ des $\mathbb{R}^n$ ist das *Skalarprodukt* $\mathbf{x} \cdot \mathbf{y}$ definiert als

$$\mathbf{x} \cdot \mathbf{y} := x_1y_1 + x_2y_2 + \cdots + x_ny_n.$$

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

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

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

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

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

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

(Das Skalarprodukt ist symmetrisch.*)*

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

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

(Das Skalarprodukt ist bilinear.*)*

Beweis. (i) Für $\mathbf{x} = (x_1, \dots, x_n)$ gilt

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

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

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

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

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

Nach Satz 3.1 (i) folgt

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

Es gelten die beiden folgenden Sätze.

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

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

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

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

Die Bewiese dieser beiden Sätze überlassen wir als Übungsaufgabe.

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

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

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

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

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

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

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

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

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

Wegen $\lambda \geq 0$ folgt daraus

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

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 $\lambda := \mathbf{y} \cdot \mathbf{y}$ und $\mu := -(\mathbf{x} \cdot \mathbf{y})$) :

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

□

Die Norm hat die folgenden Eigenschaften.

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

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

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

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

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

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

Beweis. (i) folgt aus Satz 3.1 (i).

(ii) folgt aus Satz 3.1 (iii).

Zu (iii): Es gilt

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

Zieht man auf beiden Seiten die Wurzel, so erhält man

$$|\mathbf{x}+\mathbf{y}| \le |\mathbf{x}| + |\mathbf{y}|.$$

□

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

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

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

Aus den Eigenschaften der Norm folgt:

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

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

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

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

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

$$d(\mathbf{u}, \mathbf{w}) \le d(\mathbf{u}, \mathbf{v}) + d(\mathbf{v}, \mathbf{w}).$$

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

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

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

$$-1 \le \frac{\mathbf{x}\cdot\mathbf{y}}{|\mathbf{x}||\mathbf{y}|} \le 1.$$

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

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

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

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

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

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

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

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

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

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

(iv) *Für* $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ *mit* $\mathbf{x} \neq \mathbf{0}$ *und* $\mathbf{y} \neq \mathbf{0}$ *gilt* $\angle(\mathbf{x}, \mathbf{y}) = 0$ *oder* $\angle(\mathbf{x}, \mathbf{y}) = \pi$ *genau dann, wenn es ein* $\lambda \in \mathbb{R}$ *mit* $\mathbf{y} = \lambda\mathbf{x}$ *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 $\mathbf{x}, \mathbf{y} \in \mathbb{R}^2$ von Null verschiedene Vektoren und

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

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

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

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

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

Es gilt

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

also

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

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

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

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

4 Das Vektorprodukt

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

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

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

Man beachte, dass $\mathbf{x} \times \mathbf{y}$ 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:

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

Wir erhalten die erste Komponente (Zeile) des Vektors $\mathbf{x} \times \mathbf{y}$, indem wir die erste Zeile abdecken und die Determinante der verbleibenden 2×2-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* $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^3$ *und* $\lambda \in \mathbb{R}$ *gilt*

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

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

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

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

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

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

$$\mathbf{x} \times \mathbf{y} = -\mathbf{y} \times \mathbf{x} \textit{ und } \mathbf{x} \times \mathbf{x} = \mathbf{0}.$$

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

$$\begin{aligned}
|\mathbf{x} \times \mathbf{y}|^2 &= (x_2y_3 - x_3y_2)^2 + (x_3y_1 - x_1y_3)^2 + (x_1y_2 - x_2y_1)^2 \\
&= (x_1y_2)^2 + (x_2y_1)^2 + (x_3y_1)^2 + (x_1y_3)^2 + (x_2y_3)^2 + (x_3y_2)^2 \\
&\quad -2x_1x_2y_1y_2 - 2x_1x_3y_1y_3 - 2x_2x_3y_2y_3 \\
&= (x_1^2 + x_2^2 + x_3^2)(y_1^2 + y_2^2 + y_3^2) - (x_1y_1 + x_2y_2 + x_3y_3)^2 \\
&= |\mathbf{x}|^2|\mathbf{y}|^2 - (\mathbf{x} \cdot \mathbf{y})^2.
\end{aligned}$$

□

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

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

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

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

Nach Satz 4.1 (iii) folgt also

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

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 $\mathbf{x} \times \mathbf{y}$ senkrecht auf der von den Vektoren $\mathbf{x}$ und $\mathbf{y}$ aufgespannten Ebene. Seine Länge ist nach der Formel von Korollar 4.1 gleich dem Flächeninhalt des von $\mathbf{x}$ und $\mathbf{y}$ aufgespannten Parallelogramms. Nun gibt es, wenn $\mathbf{x} \times \mathbf{y} \neq \mathbf{0}$, genau zwei Vektoren mit diesen Eigenschaften. Der Vektor $\mathbf{x} \times \mathbf{y}$ berechnet sich nach der *Rechte-Hand-Regel*: Zeigt der Daumen der rechten Hand in die Richtung von $\mathbf{x}$ und der Zeigefinger in die Richtung von $\mathbf{y}$, so zeigt der Mittelfinger in die Richtung von $\mathbf{x} \times \mathbf{y}$.

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

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

gilt, notwendigerweise $\lambda = \mu = 0$ folgt. Die Vektoren $\mathbf{x}$ und $\mathbf{y}$ 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

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

Der folgende Satz macht diese Definition etwas verständlicher.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Ist $\theta := \angle(\mathbf{x}, \mathbf{y})$, so folgt $\sin\theta = 0$, also $\theta = 0$ oder $\theta = \pi$. Das bedeutet aber, dass es ein $\rho \in \mathbb{R}$ gibt, so dass $\mathbf{y} = \rho\mathbf{x}$. Nach Satz 4.2 sind $\mathbf{x}, \mathbf{y}$ 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 $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^3$ gibt, so dass $\mathbf{v}$ und $\mathbf{w}$ linear unabhängig sind und*

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

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

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

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

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

folgt $\lambda = \mu = 0$.

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

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

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

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

□

Definition Es sei $E \subset \mathbb{R}^3$ eine Ebene. Ein Vektor $\mathbf{n}$ mit $|\mathbf{n}| = 1$ und $\mathbf{n} \perp \mathbf{x}$ für alle $\mathbf{x} \in E$ heißt ein *Einheitsnormalenvektor von E*.

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

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

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

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

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

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

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

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

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

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

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

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

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

□

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

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

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

also

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

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

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

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

Beweis. Nach Lemma 4.2 gilt

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

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

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

Nach Satz 4.1 (iii) folgt daraus

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

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

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

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

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

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

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

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

□

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

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

Damit erhalten wir die *Hessesche Normalform* der Ebenengleichung

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

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

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

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

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

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

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

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

$$\mathbf{b} := \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix},$$

so wird unser Gleichungssystem zu

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

oder kurz

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

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

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

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*

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

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 $r \leq i \leq m$ 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

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

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 $\mathcal{L}(A, \mathbf{b}) = \emptyset$.*

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

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

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

(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 $r-1$-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

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

hat die erweiterte Koeffizientenmatrix

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

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 $A\mathbf{x} = \mathbf{0}$ nur die triviale Lösung $\mathbf{x} = \mathbf{0}$.
(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 λ-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 $(A, \mathbf{b})$ die erweiterte Koeffizientenmatrix eines linearen Gleichungssystems und $(\tilde{A}, \tilde{\mathbf{b}})$ die aus $(A, \mathbf{b})$ durch endlich viele Zeilenumformungen entstandene Matrix. Dann haben die linearen Gleichungssysteme*

$$A\mathbf{x} = \mathbf{b} \text{ und } \tilde{A}\mathbf{x} = \tilde{\mathbf{b}}$$

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 λ-fachen der i-Zeile zur j-ten Zeile entspricht die Umformung der beiden Gleichungen

$$\begin{aligned} a_{i1}x_1 + \cdots + a_{in}x_n &= b_i & (1) \\ a_{j1}x_1 + \cdots + a_{jn}x_n &= b_j & (2) \end{aligned}$$

zu

$$\begin{aligned} a_{i1}x_1 + \cdots + a_{in}x_n &= b_i & (3) \\ (a_{j1} + \lambda a_{i1})x_1 + \cdots + (a_{jn} + \lambda a_{in})x_n &= b_j + \lambda b_i. & (4) \end{aligned}$$

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 λ-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 $\tilde{A}$ 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 $\tilde{a}_{1 j_1}$ unser erster Pivot.

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

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

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

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

Dabei stehen wieder an den mit $*$ markierten Stellen irgendweldche 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. $\square$

Beispiel 5.3 Wir betrachten die folgenden Umformungen

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

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

Die Matrix $\tilde{A}$ 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 $\mathbf{b}$ wird dabei mit transformiert. Aus der erweiterten Koeffizientenmatrix $(A, \mathbf{b})$ wird dabei $(\tilde{A}, \tilde{\mathbf{b}})$. Sind die Pivots die Einträge $\tilde{a}_{11}, \dots, \tilde{a}_{rr}$, so ist dieser Schritt fertig, ansonsten vertauschen wir die Spalten von $\tilde{A}$ so, dass dies gilt.

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

(3) **Rückwärtssubstitution**
Falls $\tilde{b}_{r+1} = \dots = \tilde{b}_m = 0$, 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 $a_{11}, \dots, a_{nn}$ 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* λA 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 $r = 1$.

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

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

□

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 2×2-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 $i = 1, 2, \ldots, m$ 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 λ-fachen der k-ten Zeile zur i-ten Zeile

Wir betrachten noch zusätzlich die Operation

(Z3) Multiplikation der i-ten Zeile mit λ

Diesen Operationen ordnen wir wie folgt Matrizen zu:

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

$$F_{ik} := \left(\begin{array}{ccccccccccc}
1 &&&&&&&&&& \\
& \ddots &&&&&&&&& \\
&& 1 &&&&&&&& \\
&&& 0 &&&& 1 &&& \\
&&&& 1 &&&&&& \\
&&&&& \ddots &&&&& \\
&&&&&& 1 &&&& \\
&&& 1 &&&& 0 &&& \\
&&&&&&&& 1 && \\
&&&&&&&&& \ddots & \\
&&&&&&&&&& 1
\end{array}\right)
\begin{array}{l}
\\ \\ \\ \leftarrow i\text{-te Zeile} \\ \\ \quad\vdots \\ \\ \leftarrow k\text{-te Zeile} \\ \\ \\ \\
\end{array}$$

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

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

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 $A\mathbf{x} = \mathbf{0}$ 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 $A\mathbf{x} = \mathbf{0}$ von links mit A^{-1} multiplizieren:

$$\begin{array}{rl} & A^{-1}A\mathbf{x} = A^{-1}\mathbf{0} \\ \Rightarrow & E\mathbf{x} = \mathbf{0} \\ \Rightarrow & \mathbf{x} = \mathbf{0}. \end{array}$$

Also hat $A\mathbf{x} = \mathbf{0}$ hat nur die triviale Lösung.

(b) $\Rightarrow$ (c): Das Gleichungssystem $A\mathbf{x} = \mathbf{0}$ 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 $b_{11}, \dots, b_{nn}$ 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 $F_1, \dots, F_r$ 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. □

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 $F_1, \dots, F_r$ 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 $F_1, \dots, F_r$ 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 Gruppen

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

Definition Unter einer *Verknüpfung* auf einer Menge G versteht man eine Vorschrift $*$, die zwei gegebenen Elementen $a, b \in G$ ein neues Element $a * b$ zuordnet.

Beispiel 7.1 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 7.2 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.

Definition Eine Menge G zusammen mit einer Verknüpfung $*$ heißt *Gruppe* genau dann, wenn folgende Axiome erfüllt sind:

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

(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 7.3 $\mathbb{R}$ mit der Verknüpfung $+$ bildet eine abelsche Gruppe. Die Menge $\mathbb{R}^* := \mathbb{R} \setminus \{0\}$ mit der Verknüfung $\cdot$ bildet ebenfalls eine abelsche Gruppe. Warum muss man die Null herausnehmen?

Beispiel 7.4 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.

Wir leiten nun einige einfache Folgerungen aus den Gruppenaxiomen ab.

Satz 7.1 *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:*

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

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

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

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

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

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

(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

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

Die andere Kürzungsregel zeigt man analog. □

8 Vektorräume

Wir wollen nun allgemeine Vektorräume einführen. Es sei $K := \mathbb{R}$ oder $K := \mathbb{C}$.

Definition Eine Menge V zusammen mit einer inneren Verknüpfung (*Addition* genannt) $+$, die je zwei Elementen $\mathbf{v}, \mathbf{w} \in V$ ihre *Summe* $\mathbf{v} + \mathbf{w}$ zuordnet, und einer äußeren Verknüpfung (*Multiplikation mit Skalaren* oder *skalare Multiplikation* genannt) $\cdot$, die einem Element $\lambda \in K$ und einem Element $\mathbf{v} \in V$ das *skalare Vielfache* $\lambda \cdot \mathbf{v}$ zuordnet, heißt *K-Vektorraum* (oder *Vektorraum über K*) genau dann, wenn die folgenden Axiome erfüllt sind:

(V1) V bildet zusammen mit der Addition eine abelsche Gruppe.

(V2) Für alle $\lambda, \mu \in K$ und $\mathbf{v}, \mathbf{w} \in V$ gilt:

- (a) $\lambda \cdot (\mathbf{v} + \mathbf{w}) = \lambda \cdot \mathbf{v} + \lambda \cdot \mathbf{w}$.
- (b) $(\lambda + \mu) \cdot \mathbf{v} = \lambda \cdot \mathbf{v} + \mu \cdot \mathbf{v}$.

(c) $\lambda \cdot (\mu \cdot \mathbf{v}) = (\lambda\mu) \cdot \mathbf{v}$.

(d) $1 \cdot \mathbf{v} = \mathbf{v}$.

Die Elemente von V werden als *Vektoren* bezeichnet. Das neutrale Element von V bezüglich der Addition wird *Nullvektor* genannt und mit $\mathbf{0}$ bezeichnet. Das inverse Element zu einem Vektor $\mathbf{v}$ bezüglich der Addition wird das *Negative* von $\mathbf{v}$ genannt und mit $-\mathbf{v}$ bezeichnet. Einen $\mathbb{R}$-Vektorraum bezeichnen wir auch einfach als *Vektorraum*.

Beispiel 8.1 $V = \mathbb{R}^n$ mit der in § 1 erklärten Addition und skalaren Multiplikation ist ein Vektorraum. Die Axiome prüft man leicht nach. Als Beispiel beweisen wir (V2) (a). Es sei $\mathbf{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$, $\mathbf{w} = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}$. Dann gilt

$$\begin{aligned} \lambda \cdot (\mathbf{v} + \mathbf{w}) &= \lambda \left(\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} + \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} \right) = \lambda \begin{pmatrix} v_1 + w_1 \\ \vdots \\ \lambda v_n + w_n \end{pmatrix} \\ &= \begin{pmatrix} \lambda(v_1 + w_1) \\ \vdots \\ \lambda(v_n + w_n) \end{pmatrix} = \begin{pmatrix} \lambda v_1 + \lambda w_1 \\ \vdots \\ \lambda v_n + \lambda w_n \end{pmatrix} \\ &= \lambda \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} + \lambda \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} = \lambda \cdot \mathbf{v} + \lambda \cdot \mathbf{w}. \end{aligned}$$

Beispiel 8.2 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 8.3 Es sei V die Menge aller auf $\mathbb{R}$ definierten reellwertigen Funktionen $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto f(x)$. Für zwei Elemente $f, g \in V$ und ein $\lambda \in \mathbb{R}$ definieren wir die Summe $f + g$ und das skalare Vielfache $\lambda \cdot f$ durch

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

Mit diesen Verknüpfungen ist V ein Vektorraum.

Beispiel 8.4 Die Menge V enhalte nur ein Element, das wir mit $\mathbf{0}$ bezeichnen. Definieren wir

$$\begin{aligned} \mathbf{0} + \mathbf{0} &= \mathbf{0} \\ \lambda \cdot \mathbf{0} &= \mathbf{0} \end{aligned}$$

für alle $\lambda \in K$, so erfüllt V die Vektorraumaxiome. Der K-Vektorraum V heißt der *Nullvektorraum*.

Satz 8.1 *Es sei V ein Vektorraum über K, $\mathbf{v} \in V$ und $\lambda \in K$. Dann gilt:*

(a) $0 \cdot \mathbf{v} = \mathbf{0}$.

(b) $\lambda \cdot \mathbf{0} = \mathbf{0}$.

(c) $(-1) \cdot \mathbf{v} = -\mathbf{v}$.

(d) *Aus $\lambda \cdot \mathbf{v} = \mathbf{0}$ folgt $\lambda = 0$ oder $\mathbf{v} = \mathbf{0}$.*

Beweis. (a) Aus

$$0 \cdot \mathbf{v} = (0+0) \cdot \mathbf{v} \overset{\text{(V2b)}}{=} 0 \cdot \mathbf{v} + 0 \cdot \mathbf{v}$$

folgt $0 \cdot \mathbf{v} = \mathbf{0}$.

(b) Aus

$$\lambda \cdot \mathbf{0} = \lambda \cdot (\mathbf{0} + \mathbf{0}) \overset{\text{(V2a)}}{=} \lambda \cdot \mathbf{0} + \lambda \cdot \mathbf{0}$$

folgt $\lambda \cdot \mathbf{0} = \mathbf{0}$.

(c) Aus

$$\mathbf{v} + (-1) \cdot \mathbf{v} \overset{\text{(V2d)}}{=} 1 \cdot \mathbf{v} + (-1) \cdot \mathbf{v} \overset{\text{(V2b)}}{=} (1-1) \cdot \mathbf{v} = 0 \cdot \mathbf{v} \overset{\text{(a)}}{=} \mathbf{0}$$

folgt $(-1) \cdot \mathbf{v} = -\mathbf{v}$.

(d) Ist $\lambda \cdot \mathbf{v} = \mathbf{0}$ und $\lambda \neq 0$, so folgt

$$\mathbf{v} \overset{\text{(V2d)}}{=} 1 \cdot \mathbf{v} = (\frac{1}{\lambda}\lambda) \cdot \mathbf{v} \overset{\text{(V2c)}}{=} \frac{1}{\lambda} \cdot (\lambda \cdot \mathbf{v}) = \frac{1}{\lambda} \cdot \mathbf{0} \overset{\text{(b)}}{=} \mathbf{0}.$$

□

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 $\mathcal{L}(A, \mathbf{b})$ von linearen Gleichungssystemen $A\mathbf{x} = \mathbf{b}$ betrachtet. Wir interessieren uns nun für den Fall $\mathbf{b} = \mathbf{0}$. Wir werden zeigen, dass in diesem Fall die Lösungsmenge $\mathcal{L}(A, \mathbf{0}) \subset \mathbb{R}^n$ 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:

(U1) $U \neq \emptyset$.

(U2) $\mathbf{v}, \mathbf{w} \in U \Rightarrow \mathbf{v} + \mathbf{w} \in U$
(d.h. *U ist abgeschlossen unter der Addition*).

(U3) $\lambda \in K, \mathbf{v} \in U \Rightarrow \lambda\mathbf{v} \in U$
(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 8.2 *Ein Unterraum $U \subset V$ eines Vektorraums V ist zusammen mit der induzierten Addition und skalaren Multiplikation wieder ein Vektorraum.*

Beweis. Nach (U1) gibt es ein $\mathbf{v} \in U$. Nach Satz 8.1(a) gilt $0 \cdot \mathbf{v} = \mathbf{0}$. Nach (U3) liegt daher auch der Nullvektor $\mathbf{0}$ in U. Zu jedem $\mathbf{v} \in U$ ist nach (U3) auch $-\mathbf{v} = (-1) \cdot \mathbf{v} \in U$. Die übrigen Vektorraumaxiome gelten in U, da sie in V gelten. □

Beispiel 8.5 Es sei A eine $m \times n$-Matrix. Ein lineares Gleichungssystem $A\mathbf{x} = \mathbf{0}$ bezeichnet man als ein *homogenes* lineares Gleichungssystem. Die Lösungsmenge $\mathcal{L}(A, \mathbf{0})$ eines homogenen linearen Gleichungssystems ist ein Unterraum des $\mathbb{R}^n$. Wir prüfen die Axiome (U1)–(U3) nach:

(U1) $\mathcal{L}(A, \mathbf{0})$ ist nicht leer, da $\mathcal{L}(A, \mathbf{0})$ mindestens den Nullvektor $\mathbf{0}$ enthält.

(U2) Sind $\mathbf{x}$ und $\tilde{\mathbf{x}}$ Vektoren aus $\mathcal{L}(A, \mathbf{0})$, so gilt

$$A\mathbf{x} = \mathbf{0} \text{ und } A\tilde{\mathbf{x}} = \mathbf{0}.$$

Daraus folgt

$$A(\mathbf{x} + \tilde{\mathbf{x}}) = A\mathbf{x} + A\tilde{\mathbf{x}} = \mathbf{0} + \mathbf{0} = \mathbf{0}.$$

Also ist $\mathbf{x} + \tilde{\mathbf{x}} \in \mathcal{L}(A, \mathbf{0})$.

(U3) Ist $\mathbf{x}$ ein Vektor aus $\mathcal{L}(A, \mathbf{0})$ und $\lambda \in \mathbb{R}$, so gilt

$$A(\lambda\mathbf{x}) = \lambda A\mathbf{x} = \lambda\mathbf{0} = \mathbf{0}.$$

Also ist auch $\lambda\mathbf{x} \in \mathcal{L}(A, \mathbf{0})$.

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 8.6 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 $\mathbf{v} = (1, 1)$ in U liegt, aber nicht $(-1)\mathbf{v} = -\mathbf{v} = (-1, -1)$.

Beispiel 8.7 Es sei n eine natürliche Zahl und P_n die Menge aller *Polynome* (*ganzrationalen Funktionen*)

$$p(x) = a_0 + a_1 x + \cdots + a_n x^n$$

mit reellen Koeffizienten $a_0, a_1, \dots, a_n$, deren Grad höchstens n ist. Dann ist P_n ein Unterraum des Vektorraums V aller auf $\mathbb{R}$ definierten reellwertigen Funktionen. Denn sind

$$\begin{aligned} p(x) &= a_0 + a_1 x + \cdots + a_n x^n, \\ q(x) &= b_0 + b_1 x + \cdots + b_n x^n \end{aligned}$$

Polynome und $\lambda \in \mathbb{R}$, so sind auch

$$(p+q)(x) = p(x) + q(x) = (a_0 + b_0) + (a_1 + b_1)x + \cdots + (a_n + b_n)x^n$$

und

$$(\lambda p)(x) = \lambda p(x) = (\lambda a_0) + (\lambda a_1)x + \cdots + (\lambda a_n)x^n$$

wieder Polynome vom Grad kleiner oder gleich n.

9 Lineare Hülle und lineare Unabhängigkeit

Wir betrachten nun wichtige Begriffsbildungen in Vektorräumen. Es sei V ein Vektorraum über K.

Definition Ein Vektor $\mathbf{w} \in V$ heißt *Linearkombination* der Vektoren $\mathbf{v}_1$, $\mathbf{v}_2$, …, $\mathbf{v}_r$ genau dann, wenn es $\lambda_1, \lambda_2, \dots, \lambda_r \in K$ gibt, so dass gilt:

$$\mathbf{w} = \lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2 + \cdots + \lambda_r \mathbf{v}_r.$$

Beispiel 9.1 Jeder Vektor $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3$ ist eine Linearkombination der Standardbasisvektoren

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},$$

denn es gilt

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = x_1 \mathbf{e}_1 + x_2 \mathbf{e}_2 + x_3 \mathbf{e}_3.$$

Beispiel 9.2 Es sei $\mathbf{u} = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}$ und $\mathbf{v} = \begin{pmatrix} 3 \\ 1 \\ -1 \end{pmatrix}$. Dann ist $\mathbf{w} = \begin{pmatrix} 4 \\ 1 \\ 1 \end{pmatrix}$ keine Linearkombination von $\mathbf{u}$ und $\mathbf{v}$.

Beweis. Wenn $\mathbf{w}$ Linearkombination von $\mathbf{u}$ und $\mathbf{v}$ 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. □

Wir untersuchen nun die Menge aller Linearkombinationen von vorgegebenen Vektoren $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$.

Satz 9.1 *Es sei $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$ eine Menge von Vektoren eines Vektorraums V über K. Dann gilt:*

(a) *Die Menge U aller Linearkombinationen von $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$ bildet einen Unterraum von V.*

(b) *U ist der kleinste Unterraum von V, der $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$ enthält, d.h. jeder Unterraum von V, der $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$ enthält, enthält auch U.*

Beweis. (a) Wir prüfen die Eigenschaften (U1)–(U3) für U nach:
(U1): $U \neq \emptyset$, da $\mathbf{0} = 0\mathbf{v}_1 + 0\mathbf{v}_2 + \cdots + 0\mathbf{v_r} \in U$.
(U2), (U3): Es seien $\mathbf{v}, \mathbf{w} \in U$ und $\lambda \in K$. Dann gilt

$$\begin{aligned} \mathbf{v} &= \lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2 + \cdots + \lambda_r\mathbf{v}_r, \\ \mathbf{w} &= \mu_1\mathbf{v}_1 + \mu_2\mathbf{v}_2 + \cdots + \mu_r\mathbf{v}_r \end{aligned}$$

für $\lambda_1, \lambda_2, \ldots, \lambda_r, \mu_1, \mu_2, \ldots, \mu_r \in K$. Dann gilt

$$\begin{aligned} \mathbf{v} + \mathbf{w} &= (\lambda_1 + \mu_1)\mathbf{v}_1 + (\lambda_2 + \mu_2)\mathbf{v}_2 + \cdots + (\lambda_r + \mu_r)\mathbf{v}_r \in U, \\ \lambda\mathbf{v} &= (\lambda\lambda_1)\mathbf{v}_1 + (\lambda\lambda_2)\mathbf{v}_2 + \cdots + (\lambda\lambda_r)\mathbf{v}_r \in U. \end{aligned}$$

(b) Jeder Vektor $\mathbf{v}_i$, $i = 1, \ldots, r$, ist wegen

$$\mathbf{v}_i = 0\mathbf{v}_1 + \cdots + 0\mathbf{v}_{i-1} + 1\mathbf{v}_i + 0\mathbf{v}_{i+1} + \cdots + 0\mathbf{v}_r$$

eine Linearkombination von $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$. Also enthält U die Vektoren $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$. Es sei U' ein Unterraum von V, der $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$ enthält. Aufgrund der Eigenschaften (U2) und (U3) von U' liegen dann auch alle Linearkombinationen von $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$ in U'. Damit folgt $U \subset U'$. □

Definition Es sei V ein K-Vektorraum, $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$ eine Teilmenge von V und U der Unterraum von V, der aus den Linearkombinationen von $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$ besteht. Dann sagt man, U wird von $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$ *aufgespannt* oder *erzeugt* und man nennt U die *lineare Hülle* von $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$, in Zeichen

$$U = \operatorname{span}(S) \text{ oder } U = \operatorname{span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}.$$

Die Menge S heißt *Erzeugendensystem* von U.

Beispiel 9.3 Zwei linear unabhängige Vektoren im $\mathbb{R}^3$ spannen eine Ebene durch den Nullpunkt auf.

Beispiel 9.4 Die Monome $1, x, x^2, \ldots, x^n$ spannen den Vektorraum P_n auf. Denn jedes Polynom $p(x)$ aus P_n ist eine Linearkombination

$$p(x) = a_0 + a_1 x + \cdots + a_n x^n$$

dieser Monome. Es gilt also

$$P_n = \operatorname{span}\{1, x, x^2, \ldots, x^n\}.$$

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 Eine Teilmenge $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$ eines K-Vektorraums V heißt *linear unabhängig* genau dann, wenn gilt: Sind $\lambda_1, \lambda_2, \ldots, \lambda_r \in K$ und ist

$$\lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2 + \cdots + \lambda_r \mathbf{v}_r = \mathbf{0},$$

so folgt

$$\lambda_1 = \lambda_2 = \cdots = \lambda_r = 0.$$

Anders ausgedrückt, bedeutet dies, dass sich der Nullvektor nur als triviale Linearkombination

$$\mathbf{0} = 0\mathbf{v}_1 + 0\mathbf{v}_2 + \cdots + 0\mathbf{v}_r$$

der Vektoren $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$ darstellen läßt.

Die Teilmenge S heißt *linear abhängig* genau dann, wenn S nicht linear unabhängig ist, d.h. wenn es Skalare $\lambda_1, \lambda_2, \ldots, \lambda_r$ gibt, die nicht alle gleich Null sind, so dass

$$\lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2 + \cdots + \lambda_r \mathbf{v}_r = \mathbf{0}.$$

Beispiel 9.5 Die Menge $S := \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ mit

$$\mathbf{e}_1 := \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{e}_2 := \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{e}_3 := \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

ist linear unabhängig. Denn aus der Gleichung

$$\lambda_1 \mathbf{e}_1 + \lambda_2 \mathbf{e}_2 + \lambda_3 \mathbf{e}_3 = \mathbf{0}$$

ergibt sich

$$\lambda_1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + \lambda_3 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

also

$$\begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

also $\lambda_1 = \lambda_2 = \lambda_3 = 0$.

Beispiel 9.6 Die Menge $S := \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ mit

$$\mathbf{v}_1 := \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{v}_2 := \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{v}_3 := \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$$

ist linear abhängig, denn es gilt $\mathbf{v}_1 + \mathbf{v}_2 - \mathbf{v}_3 = \mathbf{0}$.

Beispiel 9.7 Die Menge $S := \{1, x, x^2, \ldots, x^n\}$ ist linear unabhängig in P_n. Denn aus

$$a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n = 0 \text{ für alle } x \in \mathbb{R}$$

folgt

$$a_0 = a_1 = a_2 = \cdots = a_n = 0,$$

da ein Polynom n-ten Grades höchstens n Nullstellen hat.

Satz 9.2 *Es sei V ein Vektorraum über K und $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$ eine Teilmenge von V mit $r \geq 2$. Die Teilmenge S ist genau dann linear abhängig, wenn mindestens einer der Vektoren von S als Linearkombination der anderen dargestellt werden kann.*

Beweis.

"$\Rightarrow$": Die Teilmenge S sei linear abhängig. Dann existieren Skalare $\lambda_1, \lambda_2, \ldots, \lambda_r$, die nicht alle gleich Null sind, so dass

$$\lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2 + \cdots + \lambda_r \mathbf{v}_r = \mathbf{0}.$$

Ist etwa $\lambda_1 \neq 0$, so folgt daraus

$$\mathbf{v}_1 = -\frac{\lambda_2}{\lambda_1}\mathbf{v}_2 - \cdots - \frac{\lambda_r}{\lambda_1}\mathbf{v}_r.$$

Also ist $\mathbf{v}_1$ Linearkombination von $\mathbf{v}_2, \ldots, \mathbf{v}_r$. Entsprechend erhält man für $\lambda_i \neq 0$ den Vektor $\mathbf{v}_i$ als Linearkombination der übrigen.

"$\Leftarrow$": Angenommen, $\mathbf{v}_1$ lässt sich als Linearkombination der anderen darstellen:

$$\mathbf{v}_1 = \lambda_2 \mathbf{v}_2 + \cdots + \lambda_r \mathbf{v}_r.$$

Daraus folgt

$$\mathbf{v}_1 - \lambda_2 \mathbf{v}_2 - \cdots - \lambda_r \mathbf{v}_r = \mathbf{0}.$$

Also lässt sich der Nullvektor als nichttriviale Linearkombination von Vektoren aus S schreiben. Daraus folgt, dass S linear abhängig ist. Lässt sich nicht $\mathbf{v}_1$, sondern der Vektor $\mathbf{v}_i$ als Linearkombination der anderen Vektoren aus S darstellen, so folgt die Behauptung analog. □

Bemerkung 9.1 Aus Satz 9.2 folgt, dass Satz 4.2 auch in einem allgemeinen K-Vektorraum gilt. Wir erhalten als Folgerung aus Satz 9.2:

(a) *Eine Menge, die den Nullvektor enthält, ist linear abhängig.*

(b) *Eine Menge mit zwei Elementen ist genau dann linear abhängig, wenn einer der Vektoren ein skalares Vielfaches des anderen ist.*

Satz 9.3 *Es sei V ein K-Vektorraum. Eine Teilmenge $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$ von V ist genau dann linear unabhängig, wenn sich jeder Vektor aus* $\operatorname{span}(S)$ *in eindeutiger Weise als Linearkombination der Vektoren $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$ schreiben lässt.*

Beweis.

"$\Leftarrow$": Angenommen, S ist linear abhängig. Dann lässt sich der Nullvektor $\mathbf{0} \in \operatorname{span}(S)$ auf mindestens zwei verschiedene Weisen als Linearkombination der Vektoren $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$ schreiben.

"$\Rightarrow$": Es sei S linear unabhängig. Angenommen, wir hätten für den Vektor $\mathbf{v}$ die Darstellungen

$$\begin{aligned} \mathbf{v} &= \lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2 + \cdots + \lambda_r \mathbf{v}_r, \\ \mathbf{v} &= \mu_1 \mathbf{v}_1 + \mu_2 \mathbf{v}_2 + \cdots + \mu_r \mathbf{v}_r. \end{aligned}$$

Durch Subtraktion dieser Gleichungen folgt

$$\mathbf{0} \;=\; (\lambda_1 - \mu_1)\mathbf{v}_1 + (\lambda_2 - \mu_2)\mathbf{v}_2 + \cdots + (\lambda_r + \mu_r)\mathbf{v}_r.$$

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 $\mathbf{v}$ eindeutig. □

10 Basis und Dimension

Wir wollen nun die grundlegenden Begriffe Basis und Dimension einführen.

Definition Eine Teilmenge $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ eines K-Vektorraums V heißt *Basis* von V genau dann, wenn sie folgende Bedingungen erfüllt:

(a) S ist linear unabhängig.

(b) S ist ein Erzeugendensystem von V.

Aus Satz 9.3 folgt

Satz 10.1 *Ist $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ eine Basis des K-Vektorraums V, so besitzt jeder Vektor $\mathbf{v} \in V$ eine eindeutig bestimmte Darstellung*

$$\mathbf{v} = \lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2 + \cdots + \lambda_n\mathbf{v}_n.$$

Definition Es sei V ein $\mathbb{R}$-Vektorraum. Hat der Vektor $\mathbf{v}$ bezüglich $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ die Darstellung

$$\mathbf{v} = \lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2 + \cdots + \lambda_n\mathbf{v}_n,$$

so heißen die Skalare $\lambda_1, \lambda_2, \ldots, \lambda_n$ die *Koordinaten* von $\mathbf{v}$ bezüglich S. Der Vektor

$$\begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} \in \mathbb{R}^n$$

heißt *Koordinatenvektor* von $\mathbf{v}$ bezüglich S.

Beispiel 10.1 Wir haben gesehen, dass $S = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ mit

$$\mathbf{e}_1 := \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{e}_2 := \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{e}_3 := \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

linear unabhängig ist und den $\mathbb{R}^3$ erzeugt. Also ist S eine Basis des $\mathbb{R}^3$, die sogenannte *Standardbasis* des $\mathbb{R}^3$. Ein Vektor $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ hat wegen

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3$$

bezüglich S die Koordinaten x_1, x_2, x_3.

Beispiel 10.2 Wir betrachten die Teilmenge $S = \{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ des $\mathbb{R}^n$ mit

$$\mathbf{e}_1 := \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \mathbf{e}_2 := \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \dots, \mathbf{e}_n := \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

Wie im Fall $n = 3$ kann man zeigen, dass S linear unabhängig ist und den $\mathbb{R}^n$ erzeugt. Also ist S eine Basis des $\mathbb{R}^n$, die sogenannte *Standardbasis* des $\mathbb{R}^n$. Ein Vektor

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

hat die Koordinaten $x_1, \dots, x_n$.

Beispiel 10.3 Die Menge $S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ mit

$$\mathbf{v}_1 := \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{v}_2 := \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{v}_3 := \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}$$

ist eine Basis des $\mathbb{R}^3$.

Beweis. (a) Wir zeigen zunächst, dass S ein Erzeugendensystem des $\mathbb{R}^3$ ist. Dazu müssen wir zeigen, dass wir für einen beliebigen Vektor $\mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$ Skalare λ_1, λ_2, λ_3 finden können, so dass

$$\mathbf{b} = \lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2 + \lambda_3\mathbf{v}_3$$

gilt. Wir müssen also λ_1, λ_2, λ_3 finden, so dass gilt

$$\lambda_1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} + \lambda_3 \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}.$$

Dies führt zu dem linearen Gleichungssystem

$$\begin{aligned} \lambda_1 + 2\lambda_2 + 3\lambda_3 &= b_1 \\ \lambda_2 + 2\lambda_3 &= b_2 \\ \lambda_3 &= b_3. \end{aligned}$$

Dieses Gleichungssystem ist aber für alle $\mathbf{b}$ lösbar.

(b) Nun zeigen wir, dass S linear unabhängig ist. Dazu müssen wir zeigen, dass aus

$$\lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2 + \lambda_3 \mathbf{v}_3 = \mathbf{0}$$

$\lambda_1 = \lambda_2 = \lambda_3 = 0$ folgt. Wir werden auf das obige lineare Gleichungssystem mit $\mathbf{b} = \mathbf{0}$ geführt. Dies hat aber nur die triviale Lösung $\lambda_1 = \lambda_2 = \lambda_3 = 0$. $\square$

Die Koordinaten des Vektors $\mathbf{b} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ bezüglich S sind $\lambda_1 = 0$, $\lambda_2 = -4$, $\lambda_3 = 3$.

Beispiel 10.4 Die Menge $S := \{1, x, x^2, \dots, x^n\}$ ist eine Basis des Raumes P_n. Denn nach Beispiel 9.7 ist S linear unabhängig und nach Beispiel 9.4 ein Erzeugendensystem von P_n. Die Koordinaten eines Vektors

$$\mathbf{p} = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n$$

bezüglich S sind $(a_0, a_1, a_2, \dots, a_n)$.

Definition Ein Vektorraum V über K heißt *endlich dimensional*, wenn V eine endliche Basis besitzt, sonst heißt er *unendlich dimensional*.

Bemerkung 10.1 Der Nullvektorraum besitzt die leere Menge als Basis. Er ist deswegen auch endlich dimensional.

Beispiel 10.5 Die Vektorräume $\mathbb{R}^n$ und P_n sind endlich dimensional. Der Vektorraum der reellen Funktionen ist unendlich dimensional.

Satz 10.2 *Es sei $S = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ eine Basis des endlich dimensionalen K-Vektorraums V. Dann gilt:*

(i) *Jede Teilmenge von V, die mehr als n Elemente enthält, ist linear abhängig.*

(ii) *Jedes Erzeugendensystem von V enthält mindestens n Elemente.*

Beweis.

(i) Es sei $M = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m\}$ eine m-elementige Teilmenge von V mit $m > n$. Da nach Voraussetzung S eine Basis von V ist, lassen sich die Vektoren $\mathbf{w}_i$ aus M als Linearkombinationen von $\mathbf{v}_1, \ldots, \mathbf{v}_n$ darstellen:

$$\begin{array}{rcl} \mathbf{w}_1 & = & a_{11}\mathbf{v}_1 + \cdots + a_{n1}\mathbf{v}_n \\ \vdots & \vdots & \vdots \\ \mathbf{w}_m & = & a_{1m}\mathbf{v}_1 + \cdots + a_{nm}\mathbf{v}_n. \end{array}$$

Es ist zu zeigen, dass M linear abhängig ist. Dazu müssen wir zeigen, dass wir den Nullvektor als nichttriviale Linearkombination der Vektoren $\mathbf{w}_1, \ldots, \mathbf{w}_m$ darstellen können, d.h. dass es Skalare $\lambda_1, \ldots, \lambda_m$ gibt, die nicht alle gleich Null sind, so dass

$$\lambda_1\mathbf{w}_1 + \ldots + \lambda_m\mathbf{w}_m = \mathbf{0}$$

gilt. Setzen wir die obige Darstellung der $\mathbf{w}_i$ ein, so erhalten wir die Gleichung

$$(a_{11}\ \lambda_1 + \cdots + a_{1m}\ \lambda_m)\mathbf{v}_1 + \cdots + (a_{n1}\lambda_1 + \cdots + a_{nm}\lambda_m)\mathbf{v}_n = \mathbf{0}.$$

Da S linear unabhängig ist, muss gelten

$$\begin{array}{rcl} a_{11}\lambda_1 + \cdots + a_{1m}\lambda_m & = & 0 \\ \vdots & \vdots & \vdots \\ a_{n1}\lambda_1 + \cdots + a_{nm}\lambda_m & = & 0. \end{array}$$

Dies ist aber ein lineares Gleichungssystem für $\lambda_1, \ldots, \lambda_m$, das wegen $m > n$ mehr Unbekannte als Gleichungen hat. Ein solches lineares Gleichungssystem besitzt aber eine nichttriviale Lösung.

(ii) Es sei $M = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m\}$ eine m-elementige Teilmenge von V mit $m < n$. Wir zeigen, dass M kein Erzeugendensystem von V ist. Dazu nehmen wir an, dass M ein Erzeugendensystem ist, und führen diese Aussage zum Widerspruch, indem wir zeigen, dass dann S linear abhängig ist.

Da M ein Erzeugendensystem von V ist, können wir die Vektoren $\mathbf{v}_i$ als Linearkombination von $\mathbf{w}_1, \ldots, \mathbf{w}_m$ schreiben:

$$\begin{array}{rcl} \mathbf{v}_1 & = & a_{11}\mathbf{w}_1 + \cdots + a_{m1}\mathbf{w}_m \\ \vdots & \vdots & \vdots \\ \mathbf{v}_n & = & a_{1n}\mathbf{w}_1 + \cdots + a_{mn}\mathbf{w}_m. \end{array}$$

Wir wollen nun zeigen, dass S linear abhängig ist. Dazu müssen wir zeigen, dass wir den Nullvektor als nichttriviale Linearkombination der Vektoren

$\mathbf{v}_1, \dots, \mathbf{v}_n$ darstellen können, d.h. dass es Skalare $\lambda_1, \dots, \lambda_n$ gibt, die nicht alle gleich Null sind, so dass

$$\lambda_1 \mathbf{v}_1 + \ldots + \lambda_n \mathbf{v}_n = \mathbf{0}$$

gilt. Setzen wir die obige Darstellung der $\mathbf{v}_i$ ein, so erhalten wir die Gleichung

$$(a_{11}\ \lambda_1 + \cdots + a_{1n}\ \lambda_n)\mathbf{w}_1 + \cdots + (a_{m1}\lambda_1 + \cdots + a_{mn}\lambda_n)\mathbf{w}_m = \mathbf{0}.$$

Diese Gleichung ist auf jeden Fall erfüllt, wenn alle Koeffizienten Null sind, d.h. wenn gilt

$$\begin{array}{rcl} a_{11}\lambda_1 + \cdots + a_{1n}\lambda_n & = & 0 \\ \vdots & \vdots & \vdots \\ a_{m1}\lambda_1 + \cdots + a_{mn}\lambda_n & = & 0. \end{array}$$

Dies ist aber ein lineares Gleichungssystem für $\lambda_1, \dots, \lambda_n$, das wegen $n > m$ mehr Unbekannte als Gleichungen hat. Ein solches lineares Gleichungssystem besitzt aber eine nichttriviale Lösung. $\square$

Ist $S = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ eine Basis des K-Vektorraums V, so folgt aus diesem Satz, dass alle Basen von V genau n Vektoren enthalten. Also erhalten wir als wichtige Folgerung

Satz 10.3 *Alle Basen eines endlich dimensionalen K-Vektorraums enthalten die gleiche Anzahl von Vektoren.*

Dies führt zu der folgenden Definition

Definition Die *Dimension* $\dim V$ eines endlich dimensionalen K-Vektorraums V ist die Anzahl der Vektoren einer Basis von V.

Beispiel 10.6 Für den Nullvektorraum $V = \{\mathbf{0}\}$ gilt $\dim V = 0$. Nach Beispiel 10.2 gilt $\dim \mathbb{R}^n = n$. Nach Beispiel 10.4 gilt $\dim P_n = n + 1$.

Beispiel 10.7 Wir betrachten das folgende homogene lineare Gleichungssystem (vgl. Beispiel 5.2):

$$\begin{array}{rcl} x_1 + 3x_2 + 4x_3 + 2x_4 - 2x_5 & = & 0 \\ 2x_2 - x_3 + 2x_4 & = & 0 \\ 2x_3 - 6x_4 & = & 0. \end{array}$$

Wir bestimmen nun eine Basis und die Dimension des Lösungsraums dieses Gleichungssystems. 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

$$\mathbf{v}_1 := \begin{pmatrix} -\frac{31}{2} \\ \frac{1}{2} \\ 3 \\ 1 \\ 0 \end{pmatrix}, \ \mathbf{v}_2 := \begin{pmatrix} 2 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

den Lösungsraum auf. Sie sind linear unabhängig. Damit ist $S := \{\mathbf{v}_1, \mathbf{v}_2\}$ eine Basis des Lösungsraums. Der Lösungsraum hat also die Dimension 2.

11 Rang einer Matrix

Wir betrachten nun wieder Matrizen. Es sei

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

eine $m \times n$-Matrix. Die Vektoren

$$\mathbf{z}_i := (a_{i1}, a_{i2}, \ldots, a_{in}), \quad i = 1, \ldots, m,$$

heißen die *Zeilenvektoren von* A und die Vektoren

$$\mathbf{s}_j := \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix}, \quad j = 1, \ldots, n,$$

heißen die *Spaltenvektoren von* A.

Definition Es sei A eine $m \times n$-Matrix. Der von den Zeilenvektoren aufgespannte Unterraum von $\mathbb{R}^n$ heißt der *Zeilenraum von* A. Der von den Spaltenvektoren aufgespannte Unterraum von $\mathbb{R}^n$ heißt der *Spaltenraum von* A.

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

Satz 11.1 *Elementare Zeilenoperationen ändern den Zeilenraum einer Matrix nicht.*

Beweis. Es seien

$$\mathbf{z}_1, \ldots, \mathbf{z}_m$$

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 λ-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

$$\mathbf{z}_1, \ldots, \mathbf{z}_{i-1}, \mathbf{z}_i + \lambda \mathbf{z}_k, \mathbf{z}_{i+1}, \ldots, \mathbf{z}_m.$$

Da der Zeilenraum von A mit $\mathbf{z}_1, \ldots, \mathbf{z}_m$ 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).
$\square$

Der Spaltenraum einer Matrix kann sich allerdings bei elementaren Zeilenoperationen ändern.

Beispiel 11.1 Der Spaltenraum der Matrix

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

ist die Gerade $y = 0$ im $\mathbb{R}^2$, der Spaltenraum von

$$B = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

ist die Gerade $y = x$.

Es gilt aber

Satz 11.2 *Elementare Zeilenoperationen ändern den Spaltenrang einer Matrix nicht.*

Beweis. Es sei $\{\mathbf{s}_{i_1}, \dots, \mathbf{s}_{i_k}\}$ eine linear unabhängige Menge von Spaltenvektoren von A. Das bedeutet, dass das lineare Gleichungssystem

$$\lambda_1 \mathbf{s}_{i_1} + \cdots + \lambda_k \mathbf{s}_{i_k} = \mathbf{0}$$

nur die triviale Lösung $\lambda_1 = \ldots = \lambda_k = 0$ hat. Die Matrix $\tilde{A}$ sei aus A durch endlich viele elementare Zeilenoperationen entstanden. Die Spaltenvektoren von $\tilde{A}$ bezeichnen wir mit $\tilde{\mathbf{s}}_{i_1}, \dots, \tilde{\mathbf{s}}_{i_k}$. 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 λ bedeutet. Also hat auch das Gleichungssystem

$$\lambda_1 \tilde{\mathbf{s}}_{i_1} + \cdots + \lambda_k \tilde{\mathbf{s}}_{i_k} = \mathbf{0}$$

nur die triviale Lösung. Damit ist auch die Menge $\{\tilde{\mathbf{s}}_{i_1}, \dots, \tilde{\mathbf{s}}_{i_k}\}$ linear unabhängig. □

Satz 11.3 *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 $\tilde{A}$ in Zeilenstufenform überführen. Nach Satz 11.1 und Satz 11.2 ändern elementare Zeilenumformungen weder den Zeilenrang noch den Spaltenrang. Wir nehmen an, dass die Pivots die Einträge

$$\tilde{a}_{1j_1}, \tilde{a}_{2j_2}, \dots, \tilde{a}_{rj_r}$$

der Matrix $\tilde{A}$ sind. Dann sieht man leicht, dass die ersten r Zeilenvektoren eine Basis des Zeilenraums von $\tilde{A}$ bilden. Also ist der Zeilenrang von $\tilde{A}$ und damit auch von A gleich r. Die Spaltenvektoren $\tilde{\mathbf{s}}_{j_1}, \tilde{\mathbf{s}}_{j_2}, \dots, \tilde{\mathbf{s}}_{j_r}$ bilden eine Basis des Spaltenraums. Also ist auch der Spaltenrang von $\tilde{A}$ und damit auch von A gleich r. □

Satz 11.3 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 11.2 Der Rang der Matrix

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

ist gleich 3. Denn nach Beispiel 5.3 können wir A durch elementare Zeilenumformungen in die folgende Matrix in Zeilenstufenform überführen

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

Die Anzahl der Pivots ist 3, also hat diese Matrix den Rang 3.

Satz 11.4 *Für jede Matrix A gilt:*

$$\operatorname{rang} A = \operatorname{rang} A^T.$$

Beweis.

$$\operatorname{rang} A = \text{Zeilenrang}\, A = \text{Spaltenrang}\, A^T = \operatorname{rang} A^T.$$

□

12 Einige wichtige Sätze über lineare Gleichungssysteme

Wir wenden nun unsere Resultate 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.

Satz 12.1 (Lösbarkeitsentscheidung) *Ein lineares Gleichungssystem $A\mathbf{x} = \mathbf{b}$ ist genau dann lösbar, wenn der Rang von A gleich dem Rang der erweiterten Koeffizientenmatrix $(A, \mathbf{b})$ ist.*

Beweis. Sind $\mathbf{s}_1, \ldots, \mathbf{s}_n$ die Spaltenvektoren von A, so können wir das Gleichungssystem $A\mathbf{x} = \mathbf{b}$ auch so schreiben:

$$x_1\mathbf{s}_1 + \cdots + x_n\mathbf{s}_n = \mathbf{b}.$$

Das bedeutet aber, dass das Gleichungssystem $A\mathbf{x} = \mathbf{b}$ genau dann lösbar ist, wenn $\mathbf{b}$ eine Linearkombination der Spaltenvektoren von A ist, also im Spaltenraum von A liegt. Also ist das Gleichungssystem $A\mathbf{x} = \mathbf{b}$ genau dann lösbar, wenn die Spaltenräume von A und $(A, \mathbf{b})$ übereinstimmen. Das ist aber genau dann der Fall, wenn die Ränge der entsprechenden Matrizen übereinstimmen. □

Wir leiten nun eine Beziehung zwischen der Lösungsmenge eines inhomogenen Gleichungssystems $A\mathbf{x} = \mathbf{b}$ und dem Lösungsraum des zugehörigen homogenen Gleichungssystems $A\mathbf{x} = \mathbf{0}$ her.

Satz 12.2 (Lösung des inhomogenen Systems) *Es sei* $\mathbf{x}_0$ *eine Lösung des inhomogenen Gleichungssystems* $A\mathbf{x} = \mathbf{b}$ *und* $\{\mathbf{v}_1, \dots, \mathbf{v}_r\}$ *eine Basis des Lösungsraums des zugehörigen homogenen Gleichungssystems* $A\mathbf{x} = \mathbf{0}$*. Ein Vektor* $\mathbf{x}$ *ist genau dann eine Lösung von* $A\mathbf{x} = \mathbf{b}$*, wenn er die Darstellung*

$$\mathbf{x} = \mathbf{x}_0 + \lambda_1\mathbf{v}_1 + \cdots + \lambda_r\mathbf{v}_r$$

mit Skalaren $\lambda_1, \dots, \lambda_r$ *besitzt.*

Beweis.

"$\Rightarrow$": Es sei $\mathbf{x}$ eine Lösung von $A\mathbf{x} = \mathbf{b}$. Dann gilt

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

Da $\mathbf{x}_0$ ebenfalls eine Lösung von $A\mathbf{x} = \mathbf{b}$ ist, gilt

$$A\mathbf{x}_0 = \mathbf{b}.$$

Subtraktion der beiden Gleichungen liefert

$$\begin{aligned} & A\mathbf{x} - A\mathbf{x}_0 = \mathbf{0} \\ \Leftrightarrow \quad & A(\mathbf{x} - \mathbf{x}_0) = \mathbf{0}. \end{aligned}$$

Also ist $\mathbf{x} - \mathbf{x}_0$ eine Lösung des zugehörigen homogenen Gleichungssystems $A\mathbf{x} = \mathbf{0}$. Da $\{\mathbf{v}_1, \dots, \mathbf{v}_r\}$ eine Basis des Lösungsraums von $A\mathbf{x} = \mathbf{0}$ ist, existieren Skalare $\lambda_1, \dots, \lambda_r$ mit

$$\mathbf{x} - \mathbf{x}_0 = \lambda_1\mathbf{v}_1 + \cdots + \lambda_r\mathbf{v}_r.$$

Daraus folgt

$$\mathbf{x} = \mathbf{x}_0 + \lambda_1\mathbf{v}_1 + \cdots + \lambda_r\mathbf{v}_r.$$

"$\Leftarrow$": Es sei

$$\mathbf{x} = \mathbf{x}_0 + \lambda_1\mathbf{v}_1 + \cdots + \lambda_r\mathbf{v}_r.$$

Dann gilt

$$\begin{aligned} A\mathbf{x} &= A(\mathbf{x}_0 + \lambda_1\mathbf{v}_1 + \cdots + \lambda_r\mathbf{v}_r) \\ &= A\mathbf{x}_0 + \lambda_1 A(\mathbf{v}_1) + \cdots + \lambda_r A(\mathbf{v}_r) \\ &= \mathbf{b}. \end{aligned}$$

Also ist $\mathbf{x}$ eine Lösung von $A\mathbf{x} = \mathbf{0}$. □

Bemerkung 12.1 Der Vektor $\mathbf{x}_0$ in Satz 12.2 heißt eine *spezielle* Lösung des linearen Gleichungssystems $A\mathbf{x} = \mathbf{b}$. Der Ausdruck

$$\lambda_1\mathbf{v}_1 + \cdots + \lambda_r\mathbf{v}_r$$

wird als *allgemeine* Lösung des homogenen Systems $A\mathbf{x} = \mathbf{0}$ bezeichnet. Satz 12.2 besagt, dass man *die allgemeine Lösung des inhomogenen Gleichungssystems* $A\mathbf{x} = \mathbf{b}$ *erhält, indem man zu einer speziellen Lösung von* $A\mathbf{x} = \mathbf{b}$ *die allgemeine Lösung des zugehörigen homogenen Gleichungssystems* $A\mathbf{x} = \mathbf{0}$ *addiert.*

Geometrisch bedeutet dies, dass man die Lösungsmenge des inhomogenen Gleichungssystems $A\mathbf{x} = \mathbf{b}$ erhält, indem man den Lösungsraum des zugehörigen homogenen Systems $A\mathbf{x} = \mathbf{0}$ um den Vektor $\mathbf{x}_0$ verschiebt. Der Lösungsraum von $A\mathbf{x} = \mathbf{0}$ ist nach Beispiel 8.5 ein Unterraum von $\mathbb{R}^n$. Ist $\mathbf{b} \neq \mathbf{0}$, so ist auch $\mathbf{x}_0 \neq \mathbf{0}$ und die Lösungsmenge $\mathcal{L}(A, \mathbf{b})$ enthält nicht die Null. Die Lösungsmenge eines inhomogenen linearen Gleichungssystems ist also im Allgemeinen kein Unterraum des $\mathbb{R}^n$. Im $\mathbb{R}^3$ können Lösungsmengen von $A\mathbf{x} = \mathbf{b}$ zum Beispiel Geraden oder Ebenen, die nicht durch den Nullpunkt gehen, sein.

Beispiel 12.1 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

$$\mathbf{x}_0 = \begin{pmatrix} 6 \\ 0 \\ -1 \\ 0 \\ 0 \end{pmatrix}.$$

Nach Beispiel 10.7 hat der Lösungssraum des zugehörigen homogenen Gleichungssystems

$$\begin{aligned} x_1 + 3x_2 + 4x_3 + 2x_4 - 2x_5 &= 0 \\ 2x_2 - x_3 + 2x_4 &= 0 \\ 2x_3 - 6x_4 &= 0. \end{aligned}$$

eine Basis $\{\mathbf{v}_1, \mathbf{v}_2\}$ mit

$$\mathbf{v}_1 := \begin{pmatrix} -\frac{31}{2} \\ \frac{1}{2} \\ 3 \\ 1 \\ 0 \end{pmatrix}, \ \mathbf{v}_2 := \begin{pmatrix} 2 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

Also hat die allgemeine Lösung des inhomogenen Systems die Gestalt

$$\mathbf{x} = \begin{pmatrix} 6 \\ 0 \\ -1 \\ 0 \\ 0 \end{pmatrix} + \lambda_1 \begin{pmatrix} -\frac{31}{2} \\ \frac{1}{2} \\ 3 \\ 1 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 2 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

Satz 12.3 (Lösung des homogenen Systems) *Es sei A eine $m \times n$-Matrix. Der Lösungsraum des homogenen linearen Gleichungssystems $A\mathbf{x} = \mathbf{0}$ hat die Dimension $n - \operatorname{rang} A$.*

Beweis. Das Gleichungssystem $A\mathbf{x} = \mathbf{0}$ ist ein System mit m Gleichungen und n Unbekannten. Es sei r die Anzahl der gebundenen Variablen. Dann ist $n-r$ die Anzahl der freien Variablen. Die Anzahl der freien Variablen ist aber gerade die Dimension des Lösungsraum. Andererseits entspricht die Anzahl r der gebundenen Variablen gerade der Anzahl der Pivots in der Matrix $\tilde{A}$ in Zeilenstufenform, in die man A durch elementare Zeilenumformungen überführen kann. Diese Anzahl ist aber der Rang der Matrix A. □

Beispiel 12.2 Das homogene lineare Gleichungssystem

$$\begin{aligned} x_1 + 3x_2 + 4x_3 + 2x_4 - 2x_5 &= 0 \\ 2x_2 - x_3 + 2x_4 &= 0 \\ 2x_3 - 6x_4 &= 0 \end{aligned}$$

hat die Koeffizientenmatrix

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

Diese Matrix hat den Rang 3. Also ist die Dimension des Lösungsraums gleich $5 - 3 = 2$ in Übereinstimmung mit Beispiel 10.7.

Wir betrachten nun noch einige Folgerungen aus diesen Sätzen.

Satz 12.4 *Es sei A eine $m \times n$-Matrix. Dann gilt:*

(a) *$A\mathbf{x} = \mathbf{b}$ ist genau dann für alle $\mathbf{b} \in \mathbb{R}^m$ lösbar, wenn $\operatorname{rang} A = m$.*

(b) *$A\mathbf{x} = \mathbf{0}$ besitzt genau dann nur die triviale Lösung, wenn $\operatorname{rang} A = n$.*

Beweis.
(a) "$\Rightarrow$": Es sei $A\mathbf{x} = \mathbf{b}$ für alle $\mathbf{b} \in \mathbb{R}^m$ lösbar. Nach Satz 12.1 gilt dann $\operatorname{rang} A = \operatorname{rang}(A, \mathbf{b})$ für alle $\mathbf{b} \in \mathbb{R}^m$. Daraus folgt, dass alle $\mathbf{b} \in \mathbb{R}^m$ im Spaltenraum von A liegen. Daraus folgt

$$\operatorname{rang} A = \dim \mathbb{R}^m = m.$$

"$\Leftarrow$": Es sei $\operatorname{rang} A = m$. Aus Satz 10.2 folgt dann, dass $\operatorname{rang} A = \operatorname{rang}(A, \mathbf{b})$ für alle $\mathbf{b} \in \mathbb{R}^m$ gilt. Nach Satz 12.1 folgt damit, dass $A\mathbf{x} = \mathbf{b}$ für alle $\mathbf{b} \in \mathbb{R}^m$ lösbar ist.

(b) Das Gleichungssystem $A\mathbf{x} = \mathbf{0}$ besitzt genau dann nur die triviale Lösung, wenn $\dim \mathcal{L}(A, \mathbf{0}) = 0$. Nach Satz 12.3 gilt aber

$$\dim \mathcal{L}(A, \mathbf{0}) = 0 \Leftrightarrow n - \operatorname{rang} A = 0 \Leftrightarrow \operatorname{rang} A = n.$$

□

Satz 12.5 *Es sei A eine $n \times n$-Matrix. 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 $A\mathbf{x} = \mathbf{0}$ nur die triviale Lösung besitzt. Nach Satz 12.4 (b) ist dies dann und nur dann der Fall, wenn $\operatorname{rang} A = n$ gilt. □

13 Einige wichtige Sätze über Basen

Wir betrachten nun einige wichtige Sätze, die nicht nur theoretisch interessant sind, sondern auch in praktischen Anwendungen hilfreich sind.

Satz 13.1 *Es sei $S \neq \emptyset$ eine Teilmenge eines K-Vektorraums V.*

(a) *Ist S linear unabhängig und $\mathbf{v}$ ein Vektor aus V, der nicht in $\operatorname{span}(S)$ liegt, so ist auch $S \cup \{\mathbf{v}\}$ linear unabhängig.*

(b) *Ist $\mathbf{v}$ ein Vektor aus S, der sich als Linearkombination der anderen Elemente von S darstellen lässt, so gilt*

$$\operatorname{span}(S) = \operatorname{span}(S \setminus \{\mathbf{v}\}).$$

Beweis. Es sei $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$.

(a) Es sei S linear unabhängig und $\mathbf{v} \in V$, $\mathbf{v} \notin \operatorname{span}(S)$. Wir betrachten die Gleichung

$$\lambda_1 \mathbf{v}_1 + \cdots + \lambda_r \mathbf{v}_r + \lambda_{r+1} \mathbf{v} = \mathbf{0}.$$

Wir müssen zeigen, dass aus dieser Gleichung $\lambda_1 = \cdots = \lambda_r = \lambda_{r+1} = 0$ folgt. Angenommen, $\lambda_{r+1} \neq 0$. Dann folgt aus dieser Gleichung

$$\mathbf{v} = -\frac{1}{\lambda_{r+1}}(\lambda_1 \mathbf{v}_1 + \cdots + \lambda_r \mathbf{v}_r).$$

Das bedeutet aber, dass $\mathbf{v} \in \operatorname{span}(S)$ im Widerspruch zur Voraussetzung. Also gilt $\lambda_{r+1} = 0$. Damit reduziert sich die Gleichung zu

$$\lambda_1 \mathbf{v}_1 + \cdots + \lambda_r \mathbf{v}_r = \mathbf{0}.$$

Daraus folgt aber $\lambda_1 = \cdots = \lambda_r = 0$, da S linear unabhängig ist.

(b) Der Vektor $\mathbf{v}_r$ habe die Darstellung

$$\mathbf{v}_r = \mu_1 \mathbf{v}_1 + \cdots + \mu_{r-1}\mathbf{v}_{r-1}.$$

Wir müssen zeigen, dass jeder Vektor $\mathbf{w}$ aus $\mathrm{span}(S)$ bereits eine Linearkombination der Vektoren $\mathbf{v}_1, \ldots, \mathbf{v}_{r-1}$ ist. Es sei also

$$\mathbf{w} = \lambda_1 \mathbf{v}_1 + \cdots + \lambda_{r-1}\mathbf{v}_{r-1} + \lambda_r \mathbf{v}_r.$$

Setzen wir die obige Darstellung von $\mathbf{v}_r$ ein, so ergibt sich

$$\begin{aligned} \mathbf{w} &= \lambda_1 \mathbf{v}_1 + \cdots + \lambda_{r-1}\mathbf{v}_{r-1} + \lambda_r(\mu_1 \mathbf{v}_1 + \cdots + \mu_{r-1}\mathbf{v}_{r-1}) \\ &= (\lambda_1 + \lambda_r \mu_1)\mathbf{v}_1 + \cdots + (\lambda_{r-1} + \lambda_r \mu_{r-1})\mathbf{v}_{r-1}. \end{aligned}$$

Also ist $\mathbf{w}$ eine Linearkombination von $\mathbf{v}_1, \ldots, \mathbf{v}_{r-1}$. □

Satz 13.2 *Es sei V ein n-dimensionaler K-Vektorraum. Dann gilt:*

(a) *Eine n-elementige linear unabhängige Teilmenge S von V ist eine Basis von V.*

(b) *Ein n-elementiges Erzeugendensystem S von V ist eine Basis von V.*

Beweis.

(a) Wir müssen zeigen, dass S ein Erzeugendensystem von V ist. Angenommen, S ist kein Erzeugendensystem von V. Dann gibt es einen Vektor $\mathbf{v} \in V$, der nicht in $\mathrm{span}(S)$ liegt. Nach Satz 13.1 (a) ist dann $S \cup \{\mathbf{v}\}$ linear unabhängig. Diese Menge besitzt aber $n+1$ Elemente. Da $\dim V = n$, ist dies ein Widerspruch zu Satz 10.2 (i).

(b) Wir müssen zeigen, dass S linear unabhängig ist. Angenommen, S ist linear abhängig. Dann enthält S einen Vektor $\mathbf{v}$, der sich als Linearkombination der anderen Vektoren von S schreiben lässt. Nach Satz 13.1 (b) ist dann bereits $S \setminus \{\mathbf{v}\}$ ein Erzeugendensystem von V. Diese Menge besitzt aber $n-1$ Elemente, im Widerspruch zu Satz 10.2 (ii). □

Satz 13.2 erleichtert in vielen Fällen den Nachweis, dass eine Teilmenge S eines Vektorraums über K eine Basis ist. Denn ist $\dim V = n$ und hat S n Elemente, so genügt es nach diesem Satz, eine der beiden Eigenschaften einer Basis nachzuweisen.

Beispiel 13.1 Um zu zeigen, dass die Menge

$$\left\{ \begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 3 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} -4 \\ 0 \\ 0 \end{pmatrix} \right\}$$

eine Basis des $\mathbb{R}^3$ bildet, genügt es zu zeigen, dass diese Menge linear unabhängig ist.

Satz 13.3 *Es sei $S = \{\mathbf{v}_1, \dots, \mathbf{v}_r\}$ eine Teilmenge eines n-dimensionalen K-Vektorraums.*

(a) *Ist S linear unabhängig und $r < n$, so erhält man durch Hinzufügen geeigneter Vektoren zu S eine Basis von V.*

(b) *Ist S ein Erzeugendensystem von V und $r > n$, so kann man S durch Entfernen geeigneter Vektoren zu einer Basis von V machen.*

Beweis.
(a) Nach Satz 10.2 (ii) ist S kein Erzeugendensystem von V. Also gibt es einen Vektor $\mathbf{v} \in V$, der nicht in $\operatorname{span}(S)$ liegt. Nach Satz 13.1 (a) ist dann die Menge $\tilde{S} := S \cup \{\mathbf{v}\}$ linear unabhängig. Ist $\tilde{r} := r + 1 = n$, so ist $\tilde{S}$ nach Satz 13.2 (a) eine Basis. Andernfalls können wir den Prozess wiederholen, bis wir eine Teilmenge mit n Elementen erreicht haben.

(b) Nach Satz 10.2 (i) ist S linear abhängig. Also gibt es einen Vektor $\mathbf{v}$ in S, der sich als Linearkombination der übrigen Elemente von S darstellen lässt. Nach Satz 13.1 (b) ist dann auch die Menge $S' := S \setminus \{\mathbf{v}\}$ ein Erzeugendensystem von V. Ist $r' := r - 1 = n$, so ist S' nach Satz 13.2 (b) eine Basis. Andernfalls können wir den Prozess wiederholen, bis wir eine Teilmenge mit n Elementen erreicht haben. □

Satz 13.4 *Ist U 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 $\dim V = n$. Enthält U nur den Nullvektor, so ist U endlich dimensional. Andernfalls gibt es einen Vektor $\mathbf{u}_1 \neq \mathbf{0}$ in U. Die Menge $S := \{\mathbf{u}_1\}$ ist linear unabhängig. Ist S keine Basis von U, so gibt es einen Vektor $\mathbf{u}_2 \in U$, der nicht in $\operatorname{span}(S)$ liegt. Nach Satz 13.1 (a) ist dann auch die Menge $\tilde{S} := S \cup \{\mathbf{u}_2\}$ linear unabhängig. Ist auch $\tilde{S}$ keine Basis von U, so können wir so fortfahren. Da nach Satz 10.2 (i) eine linear unabhängige Teilmenge von V höchstens n Elemente enthält, müssen wir nach spätestens n Schritten auf eine Basis von U treffen. Also ist U endlich dimensional und es gilt $\dim U \leq \dim V$.

Es sei $\dim U = n = \dim V$ und $S = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ eine Basis von U. Dann ist S insbesondere linear unabhängig. Nach Satz 13.2 (a) ist S auch eine Basis von V. Also folgt $U = \operatorname{span}(S) = V$. □

Literatur

[1] H. Anton: Lineare Algebra. Spektrum Akademischer Verlag, Heidelberg Berlin, 1998, ISBN 3-8274-0324-3.

[2] G. Fischer: Lineare Algebra. 12., verb. Aufl., Vieweg Verlag, Wiesbaden, 2000, ISBN 3-528-87217-9.

Inhaltsverzeichnis