

DHBW Mannheim
Studiengang Informatik

Lineare Algebra

Stand: August 2020

Autor:

Reinhold Hübl, Fakultät für Technik, DHBW Mannheim

e-mail: reinhold.huebl@dhbw-mannheim.de

©Fakultät für Technik DHBW Mannheim

Inhaltsverzeichnis

1	Logik	3
1.1	Elementare Aussagenlogik	3
1.2	Prädikatenlogik	16
1.3	Das Prinzip der vollständigen Induktion	23
2	Algebraische Strukturen	29
2.1	Mengen und Relationen	29
2.2	Gruppen	40
2.3	Ringe	51
2.4	Körper	55
2.5	Ganze Zahlen und ihre Restklassen	62
3	Allgemeine Vektorräume	81
3.1	Vektorräume und Untervektorräume	81
3.2	Der n -dimensionale reelle Raum	86
3.3	Allgemeine reelle Vektorräume	105
3.4	Vektorräume über beliebigen Körpern	118
4	Lineare Gleichungssysteme	125
4.1	Matrizen und Gleichungssysteme	125
4.1.1	Lineare Gleichungssysteme	125
4.1.2	Matrizen	142
4.2	Matrizenoperationen	154
4.3	Matrizen und Abbildungen	162

5	Determinanten	175
5.1	Definition der Determinanten	175
5.1.1	Zweidimensionale Determinanten	175
5.1.2	Dreidimensionale Determinanten	188
5.1.3	Die n -dimensionale Determinante	200
5.2	Determinanten und invertierbare Matrizen	219
6	Eigenwerte und Normalformen	239
6.1	Eigenwerte und Eigenvektoren	240
6.2	Diagonalisierbarkeit	248

Kapitel 1

Logik

1.1 Elementare Aussagenlogik

Mit der Logik, also der *Lehre vom Argumentieren* oder der *Lehre vom Schlussfolgern* beschäftigten sich bereits die Philosophen des antiken Griechenlands. Ziel der Logik ist es, Sicherheit in die Regeln des Schließens und des Argumentierens zu bringen. Oft führen (scheinbar) zwingende Schlüsse zu offensichtlich absurden Ergebnissen (Paradoxien). Die Logik hat sich das Ziel gesetzt, die Regeln des Argumentierens so streng zu fassen, dass solche Paradoxien und Widersprüche möglichst ausgeschlossen sind. Wesentliche Komponenten der Aussagenlogik gehen schon zurück auf Aristoteles, der erstmals aussagenlogische Grundsätze diskutierte. Die klassische Aussagenlogik beschäftigt sich dabei mit der inneren Struktur von Sätzen und der Frage aus welchen Teilsätzen ein Satz aufgebaut ist. Die Grundidee dabei ist, dass jede Aussage genau einen Wahrheitswert (in der Regel *wahr* oder *falsch*) hat, und dass sich der Wahrheitswert einer zusammengesetzten Aussage eindeutig aus den Wahrheitswerten ihrer Komponenten ergibt.

Wir wollen hier zunächst einen intuitiven und umgangssprachlichen Zugang vorstellen.

Definition 1.1.1. Eine **Aussage** A ist ein Satz der (in einem gegebenen Kontext) eindeutig entweder wahr w oder falsch f ist.

Beispiel 1.1.1. Beispiele für Aussagen sind etwa:

- A_1 : Berlin liegt nördlich von Hamburg.
- A_2 : 3 ist eine Quadratzahl (in \mathbb{Z})
- A_3 : 3 ist eine Quadratzahl (in \mathbb{R})

Solche Elementaraussagen können nun durch Verknüpfungen (sogenannten **Junktoren**) miteinander zu neuen, komplexeren Aussagen verbunden werden.

Die Verneinung (Negation):

Die **Negation** einer Aussage A ist eine Aussage, die mit $\neg A$ bezeichnet wird. Negation ist also ein Junktor, der sich nur auf eine Aussage bezieht. $\neg A$ ist genau dann wahr, wenn A falsch ist und genau dann falsch, wenn A wahr ist.

Die Beziehung zwischen A und $\neg A$ lässt sich also durch folgende Wahrheitstafel beschreiben:

A	$\neg A$
w	f
f	w

Beispiel 1.1.2. Die Verneinungen der Aussagen aus Beispiel 1.1.1 sind:

- $\neg A_1$: Berlin liegt nicht nördlich von Hamburg.
- $\neg A_2$: 3 ist keine Quadratzahl (in \mathbb{Z})
- $\neg A_3$: 3 ist keine Quadratzahl (in \mathbb{R})

Die Und-Verknüpfung (Konjunktion):

Die **Konjunktion** von zwei Aussagen A und B ist eine Aussage, die mit $A \wedge B$ bezeichnet wird. Konjunktion ist also ein Junktor, der sich auf zwei Aussagen bezieht. $A \wedge B$ ist genau dann wahr, wenn sowohl A als auch B wahr sind. Ist eine der Aussagen A oder B falsch, so auch deren Konjunktion.

Die Oder–Verknüpfung (Disjunktion):

Die **Disjunktion** von zwei Aussagen A und B ist eine Aussage, die mit $A \vee B$ bezeichnet wird. Disjunktion ist also ebenfalls ein Junktorkonzept, das sich auf zwei Aussagen bezieht. $A \vee B$ ist genau dann wahr, wenn entweder A oder B wahr ist (oder beide), und $A \vee B$ ist genau dann falsch, wenn sowohl A als auch B falsch sind.

Die Entweder–Oder–Verknüpfung (Kontravalenz):

Die **Kontravalenz** von zwei Aussagen A und B ist eine Aussage, die mit $A \oplus B$ bezeichnet wird. Kontravalenz ist also ein weiterer Junktorkonzept, das sich auf zwei Aussagen bezieht. $A \oplus B$ ist genau dann wahr, wenn entweder A oder B wahr ist, aber nicht beide, und $A \vee B$ ist genau dann falsch, wenn sowohl A als auch B falsch sind oder sowohl A als auch B richtig sind.

Umgangssprachlich wird die Kontravalenz oft mit der Disjunktion verwechselt und ein *oder* gerne in dem Sinne verwendet, dass genau eine der beiden Alternativen korrekt ist. In der Logik muss der Unterschied zwischen dem einfachen *oder*, bei dem beide Aussagen wahr sein dürfen und dem *entweder - oder*, bei dem nur eine Aussage wahr sein darf, damit eine korrekte Aussage herauskommt, immer beachtet werden.

In der Informatik wird die Kontravalenz oft auch als XOR–Verknüpfung bezeichnet.

Konjunktion, Disjunktion und Kontravalenz werden durch folgende Wahrheitstafel beschreiben:

A	B	$A \wedge B$	$A \vee B$	$A \oplus B$
w	w	w	w	f
w	f	f	w	w
f	w	f	w	w
f	f	f	f	f

Beispiel 1.1.3. Betrachte die beiden Aussagen

A : 5 ist das Quadrat von 2.

B : $8 = 4 \cdot 2$.

Dann ist A eine falsche Aussage, B ist wahr. Also gilt:

$$\begin{aligned}
A \wedge B = f & \quad (\text{da } A \text{ falsch ist}) \\
A \vee B = w & \quad (\text{da } B \text{ wahr ist}) \\
A \oplus B = w & \quad (\text{da } B \text{ wahr und } A \text{ falsch ist})
\end{aligned}$$

Die Wenn–Dann–Verknüpfung (Implikation):

Die **Implikation** von A nach B , geschrieben $A \implies B$, ist eine neue Aussage, die eine hinreichende Bedingung für die Wahrheit der Aussage B ausdrückt: Immer dann, wenn A wahr ist, ist auch B wahr. Man sagt auch, A ist eine **hinreichende** Bedingung für B oder B ist eine **notwendige** Bedingung für A .

Die Genau–dann–wenn–Verknüpfung (Äquivalenz):

Die **Äquivalenz** der Aussagen A und B , geschrieben $A \iff B$, ist eine neue Aussage, die eine hinreichende und notwendige Bedingung für die Wahrheit der Aussage B ausdrückt: Genau dann, wenn A wahr ist, ist auch B wahr, und genau dann, wenn A falsch ist, ist auch B falsch. Man sagt auch, A ist eine **hinreichende und notwendige** Bedingung für B . Implikation und Äquivalenz werden durch folgende Wahrheitstafel beschrieben:

A	B	$A \implies B$	$A \iff B$
w	w	w	w
w	f	f	f
f	w	w	f
f	f	w	w

Die Äquivalenz von A und B lässt sich durch Implikationen wie folgt beschreiben: Genau dann sind A und B äquivalent, wenn $A \implies B$ und $B \implies A$.

Beispiel 1.1.4. Betrachte die beiden Aussagen

A : Die Zahl a erfüllt $a = 5$.

B : Die Zahl a erfüllt $a^2 = 25$.

Dann ist A hinreichend für B und B notwendig für A , dh. die Aussage $A \implies B$ ist wahr.

Aber A ist nicht notwendig für B (auch für $a = -5$ ist B wahr), dh. die Aussage $A \iff B$ ist falsch.

Beispiel 1.1.5. Betrachte die beiden Aussagen:

A: Die Zahl a erfüllt $a = 5$ oder $a = -5$.

B: Die Zahl a erfüllt $a^2 = 25$.

Dann ist A hinreichend und notwendig für B , dh. es gilt $A \iff B$.

Beispiel 1.1.6. Anton sagt: *Bertram lügt*. Bertram sagt: *Anton lügt*. Offensichtlich kann nur einer von beiden recht haben (wenn wir annehmen, dass sich der Satz des einen jeweils auf die Aussage des anderen bezieht). Das können wir jetzt auch mit den Regeln der Logik begründen. Dazu betrachten wir die beiden Aussagen

A: Antons Aussage "Bertram lügt" ist wahr.

B: Bertrams Aussage "Anton lügt" ist wahr.

Dann können wir den Text als $A \iff \neg B$ und $B \iff \neg A$ formulieren.

Betrachten wir die Wahrheitswertetafel, so erhalten wir aus unseren Regeln für Negation und Äquivalenz

A	B	$A \iff \neg B$	$B \iff \neg A$
w	w	f	f
w	f	w	w
f	w	w	w
f	f	f	f

Damit sind die beiden Aussagen des Textes nur dann wahr, wenn entweder A wahr und B falsch ist oder umgekehrt, wenn also entweder Anton die Wahrheit sagt und Bertram lügt oder umgekehrt.

Aufgabe 1.1.1. Anton sagt *Bertram lügt*, Bertram sagt *Claus lügt* und Claus sagt *Anton und Bertram lügen*. Wer von den Dreien sagt die Wahrheit und wer lügt?

Dieser naive Zugang zur Aussagenlogik lässt natürlich noch viele Fragen offen, speziell was die Regeln des Schließens und Schlussfolgerns betrifft. Daher wollen wir hier auch noch einen formalen Zugang skizzieren.

Definition 1.1.2. Die Sprache der Aussagenlogik besteht aus

- einer (abzählbaren) Menge von Aussagevariablen oder atomaren Formeln $V = \{\mathcal{A}_0, \mathcal{A}_1, \mathcal{A}_2, \dots\}$

- den Junktoren oder Verknüpfungssymbolen $\neg, \wedge, \vee, \rightarrow$ und \leftrightarrow .
- den Gliederungssymbolen (und).

Aus diesen Bausteinen können nach folgenden Regeln aussagenlogische Sätze oder Formeln gebildet werden:

1. Alle Aussagevariablen $\mathcal{A}_i \in V$ sind Formeln.
2. Ist α eine Formel, so auch $\neg\alpha$ (**Negation**).
3. Sind α und β Formeln so auch $\alpha \wedge \beta$ (**Konjunktion**), $\alpha \vee \beta$ (**Disjunktion**), $\alpha \rightarrow \beta$ (**materiale Implikation**) und $\alpha \leftrightarrow \beta$ (**Bikonditional**).
4. Die Gliederungssymbole kennzeichnen die Bildungsschritte der Formeln.
5. Alle Formeln können mit Hilfe der Regeln (1) - (4) in endlich vielen Schritten aus den Aussagevariablen gebildet werden.

Die Menge aller aussagenlogischen Sätze bezeichnen wir mit \mathcal{A} .

Die Gliederungssymbole werden dabei verwendet um bei iterierten Formeln die ersten Schritte zusammenzufassen. Die Konjunktion von α und $\beta \vee \gamma$ etwa wird mit $\alpha \wedge (\beta \vee \gamma)$ bezeichnet, um sie von der Disjunktion von $\alpha \wedge \beta$ und γ zu unterscheiden, die mit $(\alpha \wedge \beta) \vee \gamma$ bezeichnet wird. Um zu vermeiden, dass die Formeln zu unübersichtlich werden, gibt es Präzedenz- oder Vorfahrtsregeln, analog zur "Punkt vor Strich" Regel in der Arithmetik, die es erlauben, Klammern wegzulassen. Diese besagen, dass Negation stärker bindet als Konjunktion, und diese wiederum stärker bindet als Disjunktion. Damit können wir also statt $((\alpha \wedge \beta) \vee (\neg\gamma))$ einfach $\alpha \wedge \beta \vee \neg\gamma$ schreiben. Zu beachten ist aber, dass das eine andere Formel ist als $\alpha \wedge (\beta \vee \neg\gamma)$.

Um mit diesen Formeln arbeiten zu können, ist es notwendig, ihnen Wahrheitswerte zuzuweisen. Dazu sei $\mathbb{B} = \{\text{wahr, falsch}\}$ oder kurz $\mathbb{B} = \{w, f\}$ die Menge der Wahrheitswerte.

Definition 1.1.3. Eine Abbildung $f : V \longrightarrow \mathbb{B}$, die jeder Aussagenvariable einen Wert zuordnet, heißt **Belegung der Aussagevariablen**

Aus jeder Belegung der Aussagenvariablen lässt sich auch eine eindeutige Belegung aller Formeln, also eine Abbildung $f : \mathcal{A} \rightarrow \mathbb{B}$ herleiten. Dazu verwenden wir die Wahrheitstabeln, die wir bereits im ersten Teil dieses Abschnitts betrachtet haben.

Für die Negation bedeutet das

$f(\alpha)$	$f(\neg\alpha)$
w	f
f	w

und für die verbleibenden Junktoren

$f(\alpha)$	$f(\beta)$	$f(\alpha \wedge \beta)$	$f(\alpha \vee \beta)$	$f(\alpha \rightarrow \beta)$	$f(\alpha \leftrightarrow \beta)$
w	w	w	w	w	w
w	f	f	w	f	f
f	w	f	w	w	f
f	f	f	f	w	w

Hieraus lässt sich in endlich vielen Schritten eine eindeutige Belegung für jede Formel herleiten.

Definition 1.1.4. Wir betrachten eine Formel $\alpha \in \mathcal{A}$.

- α heißt **erfüllbar**, wenn es eine Belegung f mit $f(\alpha) = w$ gibt.
- α heißt **Tautologie**, wenn für jede Belegung f gilt: $f(\alpha) = w$.
- α heißt **Kontradiktion** oder **widersprüchlich**, wenn für jede Belegung f gilt $f(\alpha) = f$.

Beispiel 1.1.7. Die Aussage $\alpha \rightarrow \alpha$ ist eine Tautologie, die Aussage $\alpha \rightarrow \neg\alpha$ dagegen ist eine Kontradiktion. Das folgt sofort aus obigen Verknüpfungstabellen.

Beispiel 1.1.8. Die Aussage $\alpha \vee \neg\alpha$ ist eine Tautologie. Auch davon überzeugen wir uns sofort.

Beispiel 1.1.9. Die bekannte Bauernregel *Wenn der Hahn kräht auf dem Mist, ändert sich das Wetter, oder es bleibt wie es ist* ist eine Tautologie. Bezeichnet etwa α die Aussage *der Hahn kräht auf dem Mist* und β die Aussage *Das Wetter ändert sich*, so kann die Bauernregel durch die Formel

$$\alpha \rightarrow (\beta \vee \neg\beta)$$

abgebildet werden. Diese Formel ist für alle Belegungen richtig.

Ändern wir die Wetterregel ab zu *Wenn der Hahn kräht auf dem Mist, so ändert sich das Wetter*, so ist das keine Tautologie mehr. Als Formel ist das nämlich $\alpha \Rightarrow \beta$, und die ist nicht richtig, wenn α wahr ist (also der Hahn auf dem Mist kräht), aber β falsch ist, sich das Wetter also nicht ändert. Diese Regel ist aber noch erfüllbar. Kräht nämlich der Hahn auf dem Mist und ändert sich das Wetter tatsächlich, so ist unsere Formel wahr. Genauso ist sie wahr, wenn der Hahn schweigt (unabhängig davon was das Wetter macht).

Wandeln wir die Regel zu *Kräht der Hahn auf dem Mist oder tut er es nicht, so ändert sich das Wetter und es ändert sich auch nicht* ab, so erhalten wir eine Kontradiktion. Als Formel ist das nämlich $\alpha \vee \neg\alpha \rightarrow \beta \wedge \neg\beta$. Da die Aussage $\alpha \vee \neg\alpha$ immer wahr ist, unabhängig davon, was der Gockel macht, und da $\beta \wedge \neg\beta$ immer falsch ist, unabhängig davon, wie sich das Wetter entwickelt, ist die gesamte Formel immer falsch.

Aufgabe 1.1.2. Zeigen Sie, dass die Regel *Wenn der Hahn kräht auf dem Mist, so ändert sich das Wetter und es ändert sich auch nicht* erfüllbar ist aber keine Tautologie.

Bemerkung 1.1.1. Die Verneinung einer Tautologie ist eine Kontradiktion, die Verneinung einer Kontradiktion ist eine Tautologie.

Definition 1.1.5. Wir betrachten $\alpha, \beta \in \mathcal{A}$.

Ist $\alpha \rightarrow \beta$ eine Tautologie, so schreiben wir hierfür $\alpha \Rightarrow \beta$ (α impliziert β , β folgt aus α).

Ist $\alpha \leftrightarrow \beta$ eine Tautologie, so schreiben wir hierfür $\alpha \Leftrightarrow \beta$ (α und β sind äquivalent).

Einige einfache Rechenregeln helfen bei der Arbeit mit aussagenlogischen Formeln.

Satz 1.1.1. *Wir betrachten logische Ausdrücke $\alpha, \beta, \gamma \in \mathcal{A}$.*

1. *Es gelten die **Kommutativgesetze***

- $\alpha \wedge \beta \iff \beta \wedge \alpha.$
- $\alpha \vee \beta \iff \beta \vee \alpha.$

2. *Es gelten die **Assoziativgesetze***

- $(\alpha \wedge \beta) \wedge \gamma \iff \alpha \wedge (\beta \wedge \gamma).$
- $(\alpha \vee \beta) \vee \gamma \iff \alpha \vee (\beta \vee \gamma).$

3. *Es gelten die **Distributivgesetze***

- $(\alpha \wedge \beta) \vee \gamma \iff (\alpha \vee \gamma) \wedge (\beta \vee \gamma).$
- $(\alpha \vee \beta) \wedge \gamma \iff (\alpha \wedge \gamma) \vee (\beta \wedge \gamma).$

4. *Es gelten die **Absorptionsgesetze***

- $\alpha \vee (\alpha \wedge \beta) \iff \alpha.$
- $\alpha \wedge (\alpha \vee \beta) \iff \alpha.$

5. *Es gelten die **Gesetze von de Morgan***

- $\neg(\alpha \wedge \beta) \iff \neg\alpha \vee \neg\beta.$
- $\neg(\alpha \vee \beta) \iff \neg\alpha \wedge \neg\beta.$

6. *Ist γ eine Tautologie, so gilt*

- $\alpha \vee \gamma \iff \gamma.$
- $\alpha \wedge \gamma \iff \alpha.$

7. *Ist γ eine Kontradiktion, so gilt*

- $\alpha \vee \gamma \iff \alpha$.
- $\alpha \wedge \gamma \iff \gamma$.

8. Es gelten die **Gesetze über Negationen**

- $\neg(\neg\alpha) \iff \alpha$.
- $\alpha \vee \neg\alpha$ ist eine Tautologie
- $\alpha \wedge \neg\alpha$ ist eine Kontradiktion.

Beweis: Alle Aussagen ergeben sich leicht durch direktes Überprüfen anhand der Wahrheitswertetabellen. Wir führen die Argumente daher nur in einigen Fällen aus.

Für das erste Distributivgesetz etwa ist zu zeigen, dass, unabhängig davon ob α , β oder γ wahr sind oder nicht, die zusammengesetzten Aussagen $(\alpha \wedge \beta) \vee \gamma$ und $(\alpha \vee \gamma) \wedge (\beta \vee \gamma)$ immer den gleichen Wahrheitswert haben. Dazu spielen wir alle Kombinationen durch:

α	β	γ	$\alpha \wedge \beta$	$\alpha \vee \gamma$	$\beta \vee \gamma$	$(\alpha \wedge \beta) \vee \gamma$	$(\alpha \vee \gamma) \wedge (\beta \vee \gamma)$
w	w	w	w	w	w	w	w
w	w	f	w	w	w	w	w
w	f	w	f	w	w	w	w
w	f	f	f	w	f	f	f
f	w	w	f	w	w	w	w
f	w	f	f	f	w	f	f
f	f	w	f	w	w	w	w
f	f	f	f	f	f	f	f

Wir sehen also, dass in den letzten beiden Spalten immer derselbe Wahrheitswert steht, und damit sind die letzten beiden Spalten in der Tat äquivalent. Als weiteres Beispiel zeigen wir noch ein Absorptionsgesetz, auch hier durch Auswertung der zugehörigen Wahrheitswertetafeln, wobei wir hier die zu überprüfende Formel jeweils in der letzten Spalte direkt angeben.

α	β	$\alpha \wedge \beta$	$\alpha \vee (\alpha \wedge \beta)$	$\alpha \vee (\alpha \wedge \beta) \Leftrightarrow \alpha$
w	w	w	w	w
w	f	f	w	w
f	w	f	f	w
f	f	f	f	w

Damit ist also nachgewiesen, dass $\alpha \vee (\alpha \wedge \beta) \Leftrightarrow \alpha$ immer wahr ist, unabhängig davon, ob das für α oder β gilt.

Aufgabe 1.1.3. Zeigen Sie, dass für Formeln α, κ gilt: Ist κ eine Kontradiktion, so gilt

- $\alpha \vee \kappa \Leftrightarrow \alpha$.
- $\alpha \wedge \kappa \Leftrightarrow \kappa$.

Aufgabe 1.1.4. Zeigen Sie, dass für Formeln α, β, γ die Assoziativgesetze gelten:

- $(\alpha \wedge \beta) \wedge \gamma \Leftrightarrow \alpha \wedge (\beta \wedge \gamma)$.
- $(\alpha \vee \beta) \vee \gamma \Leftrightarrow \alpha \vee (\beta \vee \gamma)$.

Aufgabe 1.1.5. Zeigen Sie die Absorptionsgesetze der Logik

- $\alpha \vee (\alpha \wedge \beta) \Leftrightarrow \alpha$.
- $\alpha \wedge (\alpha \vee \beta) \Leftrightarrow \alpha$.

Aufgabe 1.1.6. Zeigen Sie die Gesetze von de Morgan

- $\neg(\alpha \wedge \beta) \Leftrightarrow \neg\alpha \vee \neg\beta$.
- $\neg(\alpha \vee \beta) \Leftrightarrow \neg\alpha \wedge \neg\beta$.

und die Gesetze über Negationen

- $\neg(\neg\alpha) \Leftrightarrow \alpha$.
- $\alpha \vee \neg\alpha$ ist eine Tautologie

- $\alpha \wedge \neg\alpha$ ist eine Kontradiktion.

Besonders wichtig in der Mathematik sind Implikationen und damit die Frage, wann eine Implikation allgemein gültig oder eine Tautologie ist. Sehr viele mathematische Sätze und Formeln sind ja von der Bauart, dass sie eine Voraussetzung formulieren und dann eine Aussage treffen, die unter dieser Voraussetzung gültig ist. So kennen Sie aus der Schule sicherlich den Satz: *Die Summe der Innenwinkel in einem Dreieck Δ beträgt 180° .* Hier ist also die Voraussetzung, dass Δ ein Dreieck ist und daraus wird gefolgert, dass die Summe der Innenwinkel in Δ immer 180° ist. Betrachten wir also die Aussagen

α : Δ ist ein Dreieck

β : Die Summe der Innenwinkel in Δ ist 180° .

so bedeutet unser Satz gerade, dass $\alpha \implies \beta$ (im Rahmen der klassischen euklidischen Geometrie) also dass $\alpha \longrightarrow \beta$ allgemeingültig oder eine Tautologie ist. Das Arbeiten mit Implikationen ist daher ein wichtiger Bestandteil der Mathematik, und viele schwierige Folgerungen werden dadurch gezeigt, dass einfachere Implikationen in geschickter Weise zusammengesetzt und kombiniert werden. Oft helfen dabei die folgenden Regeln:

Regel 1.1.2. *Der **Modus Barbara**: Für aussagenlogische Formeln α und β und γ gilt:*

$$(\alpha \rightarrow \beta) \wedge (\beta \rightarrow \gamma) \implies (\alpha \rightarrow \gamma)$$

*Wenn also die Gültigkeit von α die Gültigkeit von β impliziert, und wenn aus β dann γ folgt, so folgt aus der Gültigkeit von α schon die Gültigkeit von γ . Diese Regel nennt man auch den **Kettenschluss**.*

Beispiel 1.1.10. Wir können aus den Aussagen *Wenn es regnet, ist die Straße nass* und *Wenn die Straße nass ist, besteht Rutschgefahr* folgern, dass gilt: *Wenn es regnet, besteht Rutschgefahr.*

Dazu betrachten wir

α : Es regnet

β : Die Straße ist nass

γ : Es besteht Rutschgefahr

Dann sind gegeben $\alpha \implies \beta$ und $\beta \implies \gamma$, also können wir $\alpha \implies \gamma$ ableiten.

Regel 1.1.3. *Der **Modus Ponens**: Für aussagenlogische Formeln α und β gilt:*

$$\alpha \wedge (\alpha \rightarrow \beta) \implies \beta$$

*Wenn also α gilt, und wenn aus α schon β folgt, so gilt auch β . Diese Regel heißt auch **Abtrennungsregel**.*

Beispiel 1.1.11. Wir können aus den Aussagen *Wenn es regnet, ist die Straße nass* und *Es regnet* die Aussage *Die Straße ist nass* ableiten.

Beispiel 1.1.12. Wir können aus den Aussagen *Wenn $a = 5$, dann ist $a^2 = 25$* und $a = 5$ die Aussage $a^2 = 25$ ableiten.

Regel 1.1.4. *Der **Modus Tollens**: Für aussagenlogische Formeln α und β gilt:*

$$\neg\alpha \wedge (\beta \longrightarrow \alpha) \implies \neg\beta$$

*Wenn also β nicht gilt, und wenn aus der Formel α die Formel β folgen würde, dann kann auch α nicht gelten. Diese Folgerung heißt auch **Aufhebungsregel**.*

Beispiel 1.1.13. Aus den Aussagen *Wenn es regnet, ist die Straße nass* und *Die Straße ist nicht nass* können wir die Aussage *Es regnet nicht* ableiten.

Beispiel 1.1.14. Aus den Aussagen $a^2 \neq 25$ und *Wenn $a = 5$ dann ist $a^2 = 25$* können wir ableiten, dass $a \neq 5$ gilt.

Regel 1.1.5. *Die **Kontraposition**: Für aussagenlogische Formeln α und β gilt:*

$$(\alpha \longrightarrow \beta) \implies (\neg\beta \longrightarrow \neg\alpha)$$

*Wenn also aus der Gültigkeit von α die von β folgt, und wenn wir wissen, dass β nicht wahr ist, dann kann auch α nicht wahr sein. Diese Folgerung nennt man auch **Umkehrschluß**.*

Beispiel 1.1.15. Aus der Aussage *Wenn es regnet, ist die Straße nass* kann die Aussage *Wenn die Straße nicht nass ist, dann regnet es nicht* abgeleitet werden. Falsch wäre es aber, die Aussage *Wenn es nicht regnet, dann ist die Straße nicht nass* zu folgern.

Beispiel 1.1.16. Aus den Aussagen *Alle Griechen sind Menschen* und *Alle Menschen sind sterblich* können wir mit Hilfe des Kettenschlusses und des Umkehrschlusses die Aussage *Wer nicht sterblich ist, ist kein Grieche* ableiten.

1.2 Prädikatenlogik

Wir haben jetzt die Grundzüge der Aussagen kennengelernt. Nun wollen wir die Sprache der Aussagenlogik um die Einführung von Unbekannten sowie Quantoren erweitern und uns noch kurz der Prädikatenlogik zuwenden.

Mit der elementaren Aussagenlogik lassen sich nur Aussagen der Form

$$\alpha : \quad 7 > 0$$

betrachten. In der Prädikatenlogik ist es auch möglich, Prädikate oder Aussageformen der Art

$$\alpha(x) : \quad x > 0$$

zu formulieren und zu behandeln.

Informell gesprochen ist ein Prädikat eine Folge von Wörtern mit definierten Platzhaltern (oder Variablen), die zu einer (falschen oder wahren) Aussage wird, wenn für jeden Platzhalter ein Wert (aus einem vorgegebenen Konstantenraum) eingesetzt wird.

Beispiel 1.2.1. Der Ausdruck *x ist ein Mensch* ist ein Prädikat, das durch Einsetzen des konkreten Namens eines Menschen zu einer wahren Aussage wird.

Beispiel 1.2.2. $\alpha(x) : x > 0$ ist in diesem Sinn ein Prädikat mit Platzhalter x . Setzt man etwa den Wert 5 für diesen Platzhalter ein, so erhält man eine wahre Aussage, für -3 dagegen eine falsche.

Beispiel 1.2.3. $\beta(x, y) : y = x^2 + 2$ ist ein Prädikat mit zwei Platzhaltern (zweistelliges Prädikat). Durch Einsetzen eines Platzhalters, etwa $x = 2$ wird daraus ein einstelliges Prädikat $y = 6$ (wie in Beispiel 1.2.2). Durch weiteres Einsetzen des Platzhalters $y = 6$ erhält man hieraus eine wahre Aussage, alle anderen Belegungen für y führen zu falschen Aussagen.

Die Menge aller Formeln mit einem Platzhalter x bezeichnen wir mit $\mathcal{A}(x)$, entsprechend die mit zwei Platzhaltern x und y mit $\mathcal{A}(x, y)$ usw.

Bei Prädikaten ist es nicht nur interessant, zu untersuchen, für welche Belegungen der Platzhalter eine wahre Aussage besteht, es ist auch von Bedeutung, festzustellen, ob es überhaupt Belegungen gibt, die zu wahren Aussagen führen, oder ob alle Belegungen wahre Aussagen ergeben, d.h. bei einem einstelligen Prädikat $\alpha(x)$ können die folgenden Fälle auftreten:

- $\alpha(x)$ ist für jeden Wert von x wahr. Das ist etwa der Fall für

$$\alpha : \mathbb{R} \longrightarrow \mathcal{A}(x), \alpha(x) : x^2 \geq 0.$$

- $\alpha(x)$ ist für keinen Wert von x wahr. Das ist etwa der Fall für

$$\alpha : \mathbb{R} \longrightarrow \mathcal{A}(x), \alpha(x) : x^2 \leq -1.$$

- $\alpha(x)$ ist für mindestens einen Wert von x wahr. Das ist der Fall für

$$\alpha : \mathbb{R} \longrightarrow \mathcal{A}(x), \alpha(x) : x^2 \leq 4.$$

- $\alpha(x)$ ist für mindestens einen Wert von x falsch. Das ist ebenfalls der Fall für das Beispiel

$$\alpha : \mathbb{R} \longrightarrow \mathcal{A}(x), \alpha(x) : x^2 \leq 4.$$

Um diese Untersuchungen anstellen zu können, verwendet man in der Prädikatenlogik die **Quantoren** \forall (**Allquantor**) und \exists (**Existenzquantor**). Damit schreiben sich die obigen Aussagen wie folgt:

- $\forall x \alpha(x)$ (für alle Werte von x ist $\alpha(x)$ wahr).
- $\forall x \neg \alpha(x)$ (für alle Werte von x ist $\alpha(x)$ falsch).

- $\exists x \alpha(x)$ (es gibt einen Wert von x , für den $\alpha(x)$ wahr ist).
- $\exists x \neg \alpha(x)$ (es gibt einen Wert von x , für den $\alpha(x)$ falsch ist).

Ein Quantor bindet bei einem Prädikat immer eine Variable, und macht also aus einem einstufigen Prädikat eine Aussage. Aus einem Prädikat $\alpha(x)$ wird also eine Aussage $\forall x \alpha(x)$, die genau dann wahr ist, wenn für jede Wahl von x aus dem zugrundeliegenden Individuenraum $\alpha(x)$ eine wahre Aussage wird. Entsprechend macht ein Quantor aus einem $n+1$ -stelligen Prädikat ein n -stelliges Prädikat. So wird etwa aus dem zweistelligen Prädikat $\beta(x, y) : y = x^2 + 2$ durch die Quantifizierung $\forall x \beta(x, y)$ ein einstelliges Prädikat und durch $\forall x \exists y \beta(x, y)$ eine (wahre) Aussage (mit \mathbb{R} als Individuenraum). Hierfür schreiben wir auch ausführlicher

$$\forall x \in \mathbb{R} \exists y \in \mathbb{R} : y = x^2 + 2$$

Zu beachten ist dabei, dass die Reihenfolge eine entscheidende Rolle spielt und dass $\forall x \exists y B(x, y)$ eine andere Aussage ist als $\exists y \forall x B(x, y)$. Letztere ist nämlich falsch, denn es gibt kein y , das für alle x die Gleichung $y = x^2 + 2$ löst.

Umgekehrt setzt aber die Anwendung eines Quantors auf ein Prädikat auch voraus, dass dieses Prädikat eine freie Variable hat, auf die sich der Quantor beziehen kann.

Für die Prädikatenlogik ist es also zunächst einmal wesentlich, einen **Individuenraum** festzulegen, also eine Gesamtheit von möglichen Belegungen für die Variablen. In unseren Beispielen haben wir dafür immer die reellen Zahlen \mathbb{R} genommen.

Der **Prädikatenlogik** liegen zwei **Alphabete** zugrunde, ein Alphabet der logischen Zeichen, bestehend aus Variablen, Konnektoren, Trennzeichen, Quantoren und logischen Atomen, und ein Alphabet der Theoriezeichen, bestehend aus einer Menge von Konstanten (also Objekten oder Individuum aus einem vorgegebenen Individuenbereich, der betrachtet wird), einer Menge von Funktionszeichen und einer Menge von Relationszeichen.

Unter Funktionszeichen f können wir uns dabei ganz konkret mathematische Funktionszeichen, etwa $f(x, y) = x \cdot y$, und unter Relationszeichen r entsprechend mathematische Relationen wie $r(x, y) : x > y$ vorstellen.

Variablen sind Platzhalter, die für noch nicht spezifizierte Objekte des Individuenbereichs stehen und zu einem gegebenen Zeitpunkt genau einen Wert aus dem Individuenbereich annehmen können.

Ein Term ist nun alles, was sich aus Konstanten, Variablen und Funktionszeichen sukzessive (in endlich vielen Schritten) bilden lässt. Mit Konstanten \mathbb{R} etwa können wir den Term $t_1 = f(x, 2) = 2 \cdot x$ bilden, mit diesem Term wiederum den Term $f(x, t_1) = x \cdot t_1 = 2x^2$. Konstanten und Variablen selbst sind auch Terme.

Die Sprache \mathcal{P} der Prädikatenlogik besteht also aus den folgenden Elementen

- Termen:
 1. Jede Konstante ist ein Term.
 2. Jede Variable ist ein Term.
 3. Sind t_1, \dots, t_n Terme und ist f ein n -stelliges Funktionszeichen, so ist auch $f(t_1, \dots, t_n)$ ein Term.
- Ausdrücken oder Prädikaten
- Verknüpfungssymbole \neg , \wedge , \vee , \rightarrow und \leftrightarrow .
- Quantoren \forall und \exists .

wobei für die Formeln und Ausdrücke der Prädikatenlogik folgende Bildungsgesetze gelten:

1. Alle logischen Terme sind Ausdrücke der Prädikatenlogik.
2. Sind t_1, \dots, t_n Terme und ist r eine n -stellige Relation, so ist $r(t_1, \dots, t_n)$ ein Ausdruck der Prädikatenlogik.
3. Ist α eine Formel, so auch $\neg\alpha$ (**Negation**).
4. Sind α und β Formeln, so auch $\alpha \wedge \beta$, $\alpha \vee \beta$, $\alpha \rightarrow \beta$ und $\alpha \leftrightarrow \beta$.
5. Ist α eine Formel der Prädikatenlogik mit mindestens einer freien Variable x , so sind auch $\forall x \alpha$ und $\exists x \alpha$ Formeln der Prädikatenlogik.

6. Die Gliederungssymbole kennzeichnen die Bildungsschritte der Formeln.
7. Alle Formeln der Prädikatenlogik können mit Hilfe der Regeln (1) - (6) in endlich vielen Schritten aus den Prädikaten gebildet werden.

Durch eine Variablenbelegung wird aus einem Prädikat eine Aussage, und durch eine Wahrheitswertebelegung lässt sich dann wiederum entscheiden, ob diese wahr oder falsch ist.

Mit der Sprache der Prädikatenlogik lassen sich mathematische Fragenstellungen knapp und präzise formulieren und definieren. Die Aussage

Die Zahlenfolge $a_n = \frac{n+1}{n}$ konvergiert gegen 1 für $n \rightarrow \infty$.

formuliert sich dann etwa so

$$\forall \varepsilon > 0 \exists N \forall n \geq N : |a_n - 1| < \varepsilon$$

Ein anderes Beispiel ist die Aussage

Jede gerade Zahl, die größer als zwei ist, ist die Summe zweier Primzahlen.

Diese lässt sich wie folgt formulieren:

$$\forall n \in \mathbb{N} ((2n > 2) \rightarrow (\exists x \in \mathbb{P} \exists y \in \mathbb{P} : 2n = x + y))$$

wobei \mathbb{N} die natürlichen Zahlen und \mathbb{P} die Primzahlen bezeichnet. Diese Aussage ist die sogenannte **Goldbachvermutung**. Sie geht zurück auf den Mathematiker Christian Goldbach (1690 - 1764). Es ist bis heute nicht bekannt, ob sie wahr ist.

Beispiel 1.2.4. Sind $\alpha(x)$ und $\beta(x)$ Prädikate über reelle Zahlen, so fragt die Aussage $\forall x \in \mathbb{R} \alpha(x) \rightarrow \beta(x)$, danach, ob für alle reellen Zahlen x , für die $\alpha(x)$ wahr ist, auch $\beta(x)$ eine wahre Aussage wird. Ist etwa $\alpha(x) : x \geq 5$ und $\beta(x) : x^2 \geq 25$ so ist $\forall x \in \mathbb{R} \alpha(x) \rightarrow \beta(x)$ eine wahre Aussage, denn für jede reelle Zahl a , für die $a \geq 5$ wahr ist, gilt auch $a^2 \geq 25$. Dagegen ist die Aussage $\forall x \in \mathbb{R} \beta(x) \rightarrow \alpha(x)$ falsch, denn $\beta(-6)$ ist wahr, aber $\alpha(-6)$ ist falsch. Die Aussage $\exists x \geq 0 \beta(x) \rightarrow \alpha(x)$ dagegen ist eine wahre Aussage.

Bei Implikationen mit parametrisierten Aussagen interessiert uns also vor allem der Allquantor, wir wollen wissen, ob die Gültigkeit einer Aussage $\alpha(x)$ für alle x (aus einem bestimmten Bereich) die Gültigkeit von $\beta(x)$ für diese x impliziert. Ist dabei klar, auf welche x wir uns beziehen, so lassen wir hier den Allquantor sogar weg und schreiben kurz $\alpha(x) \Rightarrow \beta(x)$, also etwa

$$x \geq 5 \Rightarrow x^2 \geq 25$$

und meinen damit, dass alle reellen Zahlen x , die die Ungleichung $x \geq 5$ erfüllen, auch die Ungleichung $x^2 \geq 25$ erfüllen, dass also die Aussage $\forall x \in \mathbb{R} (x \geq 5) \Rightarrow (x^2 \geq 25)$ eine Tautologie ist.

Bemerkung 1.2.1. Implikationen, die von einer Variable x abhängen, sind erst dann allgemein bewiesen, wenn für jede zulässige Belegung der Variable x gezeigt ist, dass die resultierende Implikation richtig ist.

Betrachten wir nun folgenden mathematischen Satz: **Das Quadrat einer geraden Zahl ist eine gerade Zahl.** Dazu setzen wir

$\alpha(x) :$ x ist eine gerade natürliche Zahl

$\beta(x) :$ x^2 ist eine gerade Zahl

Dann können wir unsere Aussage wie folgt formulieren

$$\forall x \in \mathbb{Z} (\alpha(x) \longrightarrow \beta(x))$$

und zum Nachweis dieser Aussage muss nun wirklich für jede einzelne Zahl $x \in \mathbb{Z}$ gezeigt werden, dass $\alpha(x) \implies \beta(x)$ eine wahre Aussage ist. Allerdings ist dabei für ungerade x nichts zu zeigen, denn für ungerade x ist ja $\alpha(x)$ eine falsche Aussage und damit $\alpha(x) \Rightarrow \beta(x)$ daher immer richtig.

Für gerade x können wir in diesem Fall wie folgt vorgehen: Wir schreiben $x = 2 \cdot n$ mit einer ganzen Zahl n . Das geht, denn x ist eine gerade Zahl. Dann gilt

$$x^2 = (2 \cdot n)^2 = 4 \cdot n^2 = 2 \cdot (2 \cdot n^2)$$

und damit ist x^2 das zweifache einer ganzen Zahl, also wieder eine gerade Zahl.

In diesem Fall haben wir aus der Gültigkeit von $\alpha(x)$ für ein gegebenes x unmittelbar die Gültigkeit von $\beta(x)$ für dieses x abgeleitet. Diese Art des Beweises nennt man **direkter Beweis**. Bei einem direkten Beweis gehen wir

also von einer uns bereits bekannten Tatsache aus (z.B. dass sich jede gerade ganze Zahl x als $x = 2 \cdot n$ mit einer ganzen Zahl n) und arbeiten mit dieser solange, bis wir daraus die Folgerung ableiten können.

Ein weiteres wichtiges mathematisches Beweisprinzip liefert uns der Umkehrschluss, nämlich den sogenannten **indirekten Beweis** oder **Widerspruchsbeweis**. Dieses Prinzip können wir anwenden, wenn wir eine Implikation $\alpha \implies \beta$ zeigen wollen. Nehmen wir nämlich an, dass β falsch ist, und können wir daraus ableiten, dass dann auch α falsch sein muss, können wir also $\neg\beta \implies \neg\alpha$ zeigen, so ist nach dem Umkehrschluss die Implikation $\alpha \implies \beta$ eine Tautologie. Eine Variante davon ist, dass wir annehmen, dass die Folgerung nicht richtig ist und daraus einen Widerspruch ableiten.

Dazu betrachten wir als Beispiel die Aussage:

Ist n eine natürliche Zahl, deren Quadratwurzel \sqrt{n} eine ganze Zahl ist, und ist n ungerade, so ist auch \sqrt{n} ungerade.

Dazu setzen wir

$\alpha(x) :$ x ist eine ungerade natürliche Zahl und \sqrt{x} ist eine ganze Zahl

$\beta(x) :$ \sqrt{x} ist eine ungerade natürliche Zahl

Dann können wir unsere Aussage wie folgt formulieren

$$\forall x \in \mathbb{Z} (\alpha(x) \implies \beta(x))$$

Zum Beweis nehmen wir an, wir hätten eine ungerade natürliche Zahl n gefunden, deren Quadratwurzel \sqrt{n} wieder eine natürliche Zahl ist (so dass also $\alpha(n)$ wahr ist), für die aber \sqrt{n} gerade ist (also $\beta(n)$ falsch ist). Nun haben wir ja im ersten Teil dieser Bemerkung schon gesehen, dass dann auch

$$n = \sqrt{n}^2$$

eine gerade Zahl ist, im Widerspruch zu unserer Annahme, dass $\alpha(n)$ wahr ist. Damit ist die Aussage bewiesen.

Um eine von einer Variable x abhängige Implikation zu widerlegen und als nicht allgemeingültig zu erkennen, genügt es eine einzige Belegung von x zu finden, für die die Implikation falsch wird. Wir betrachten dazu die mathematische Aussage: *Jede ungerade Zahl ist eine Primzahl*. Zunächst setzen wir

$\alpha(x) :$ x ist eine ungerade Zahl

$\beta(x) :$ x ist eine Primzahl

Dann können wir unsere Aussage wie folgt formulieren

$$\forall x \in \mathbb{Z} (\alpha(x) \longrightarrow \beta(x))$$

Diese Aussage ist sicherlich richtig für $x = 3$ oder $x = 5$, aber für $x = 9$ gilt zwar, dass x ungerade ist, sodass also $\alpha(9)$ wahr ist, aber $x = 3 \cdot 3$, und damit ist 9 keine Primzahl (also $\beta(9)$ falsch). Damit haben wir eine Belegung x gefunden, für die $\alpha(x) \longrightarrow \beta(x)$ eine falsche Aussage ist, und damit ist unsere Behauptung widerlegt. Eine Belegung von x für die $\alpha(x) \longrightarrow \beta(x)$ falsch wird, nennen wir **Gegenbeispiel**. Eine Implikation kann also durch Angabe eines einzigen Gegenbeispiels widerlegt werden.

1.3 Das Prinzip der vollständigen Induktion

Von besonderem Interesse sind auch Aussagen mit den natürlichen Zahlen als Variablen, also Aussagen $A(0), A(1), A(2), \dots, A(n), \dots$, etwa

$$A(n) : \sum_{i=1}^n i = \frac{n(n+1)}{2}$$

wobei \sum das Summenzeichen ist, also

$$\sum_{i=1}^n a_i = a_1 + a_2 + \dots + a_n$$

Ein wichtiges Hilfsmittel zur Überprüfung der Wahrheit der Aussagen $A(n)$ ist

Satz 1.3.1 (Das Prinzip der vollständigen Induktion). *Ist n_0 eine natürliche Zahl, für die gilt*

1. $A(n_0)$ ist wahr.
2. $A(n) \implies A(n+1)$ für alle $n \geq n_0$.

so ist $A(n)$ wahr für alle $n \geq n_0$

Im Regelfall werden wir mit $n_0 = 0$ oder $n_0 = 1$ arbeiten, also

Folgerung 1.3.2. *Gilt*

- $A(0)$ ist wahr.
- $A(n) \implies A(n+1)$ für alle $n \in \mathbb{N}$.

so ist $A(n)$ wahr für alle $n \in \mathbb{N}$

Gilt

- $A(1)$ ist wahr.
- $A(n) \implies A(n+1)$ für alle $n \geq 1$.

so ist $A(n)$ wahr für alle $n \geq 1$

Beweis: Das Prinzip der vollständigen Induktion ist nichts anders als ein sukzessive Anwendung des Kettenschlusses:

Ist $n > n_0$, so folgt aus Voraussetzung (2)

$$\begin{array}{rcl} & (A(n_0) & \longrightarrow A(n_0 + 1)) \\ \wedge & ((A(n_0 + 1) & \longrightarrow A(n_0 + 2)) \\ & \vdots \\ \wedge & (A(n-1) & \longrightarrow A(n)) \end{array}$$

und daher durch wiederholtes Anwenden des Modus Barbara $A(n_0) \longrightarrow A(n)$
Nach Voraussetzung (1) ist aber $A(n_0)$ wahr, und damit auch $A(n)$.

Beispiel 1.3.1. Wir wollen mit vollständiger Induktion die Aussagen

$$A(n) : \sum_{i=1}^n i = \frac{n \cdot (n+1)}{2}$$

für alle $n \geq 1$ beweisen:

Für $n = 1$ gilt:

$$\sum_{i=1}^1 i = 1 = \frac{(1+1) \cdot 1}{2},$$

also ist $A(1)$ wahr.

Wir haben noch zu zeigen: Ist $A(n)$ wahr, so auch $A(n+1)$. Wir nehmen also an, dass gilt

$$\sum_{i=1}^n i = \frac{n(n+1)}{2} \tag{1.3.1}$$

und erhalten daraus

$$\begin{aligned}
 \sum_{i=1}^{n+1} i &= \sum_{i=1}^n i + (n+1) && \text{Abspaltung des letzten Summanden} \\
 &= \frac{((n+1)n}{2} + n+1 && \text{Induktionsvoraussetzung 1.3.1} \\
 &= \frac{n^2+n}{2} + \frac{2n+2}{2} \\
 &= \frac{n^2+3n+2}{2} \\
 &= \frac{((n+1)+1)(n+1)}{2}
 \end{aligned}$$

und das ist gerade die behauptete Formel für $A(n+1)$, so dass wir also gezeigt haben

$$A(n) \implies A(n+1)$$

und damit folgt aus Satz 1.3.1 die Behauptung.

Beispiel 1.3.2. Wir wollen mit vollständiger Induktion die Aussagen

$$A(n) : \sum_{i=1}^n i^2 = \frac{n(n+1)(2n+1)}{6}$$

für alle $n \geq 1$ beweisen:

Induktionsanfang: Für $n = 1$ gilt:

$$\sum_{i=1}^1 i^2 = 1 = \frac{1(1+1)(2 \cdot 1 + 1)}{6},$$

also ist $A(1)$ wahr.

Induktionsschluss: Wir haben noch für $n \geq 1$ zu zeigen: Ist $A(n)$ wahr, so auch $A(n+1)$. Wir nehmen also an, dass gilt

$$\sum_{i=1}^n i^2 = \frac{n(n+1)(2n+1)}{6} \tag{1.3.2}$$

und erhalten daraus

$$\begin{aligned}
 \sum_{i=1}^{n+1} i^2 &= \sum_{i=1}^n i^2 + (n+1)^2 && \text{Abspaltung des letzten Summanden} \\
 &= \frac{n(n+1)(2n+1)}{6} + (n+1)^2 && \text{Induktionsvoraussetzung 1.3.2} \\
 &= \frac{n(n+1)(2n+1)}{6} + \frac{6(n+1)^2}{6} \\
 &= \frac{n(n+1)(2n+1) + 6(n+1)^2}{6} \\
 &= \frac{(n+1)((2n^2+n) + 6(n+1))}{6} \\
 &= \frac{(n+1)(2n^2+7n+6)}{6} \\
 &= \frac{(n+1)(n+2)(2n+3)}{6} \\
 &= \frac{(n+1)((n+1)+1)(2(n+1)+1)}{6}
 \end{aligned}$$

und das ist gerade die behauptete Formel für $A(n+1)$, so dass wir also gezeigt haben

$$A(n) \implies A(n+1)$$

und damit folgt aus Satz 1.3.1 die Behauptung.

Aufgabe 1.3.1. Welche der Aussagen $A \vee B$, $A \wedge B$, $\neg A \vee B$, $A \wedge \neg B$ sind wahr:

- a) $A : 16$ ist durch 2 teilbar, $B : 16$ ist eine Quadratzahl.
- b) $A : 5 < 3$, $B : 7$ ist eine gerade Zahl.
- c) $A : 4 \cdot 2 = 8$, $B : 4 + 2 = 7$.

Aufgabe 1.3.2. Gegeben seien die Aussagen

- A : "Wenn es regnet, ist die Straße nass."
- B : "Wenn die Straße nass ist besteht Rutschgefahr."

Kann man daraus die Aussage

- C : "Wenn keine Rutschgefahr besteht, dann regnet es nicht." ableiten?

Aufgabe 1.3.3. Zeigen Sie, dass für ein Prädikat $A(x)$ folgende Äquivalenzen gelten:

$$\begin{aligned}\neg(\forall x A(x)) &\iff \exists x \neg A(x) \\ \neg(\exists x A(x)) &\iff \forall x \neg A(x)\end{aligned}$$

Aufgabe 1.3.4. Welche der Beziehungen $A(x) \rightarrow B(x)$, $B(x) \rightarrow A(x)$, $A(x) \leftrightarrow B(x)$ sind für alle $x \in \mathbb{R}$ wahr:

- a) $A(x) : x = 5$, $B(x) : x^6 = 625$.
- b) $A(x) : 2x + 3 = 5$, $B(x) : x + 7 = 8$.
- c) $A(x) : x^2 \geq 10$, $B(x) : x \geq 5$.

Aufgabe 1.3.5. Zeigen Sie, dass für alle $n \geq 1$ gilt:

$$\sum_{i=1}^n (2i - 1) = n^2$$

Aufgabe 1.3.6. Zeigen Sie, dass für alle $n \geq 1$ gilt:

$$\sum_{i=1}^n i^3 = \frac{n^2(n+1)^2}{4}$$

Aufgabe 1.3.7. Zeigen Sie, dass für jede reelle Zahl $q \neq 1$ und jedes $n \geq 0$ gilt:

$$\sum_{i=0}^n q^i = \frac{1 - q^{n+1}}{1 - q}$$

Aufgabe 1.3.8. Zeigen Sie, dass für alle Zahlen $n \geq 3$ gilt:

$$2n + 1 < 2^n$$

Aufgabe 1.3.9. Zeigen Sie, dass $n^3 + 2n$ für alle $n \geq 1$ durch 3 teilbar ist.

Aufgabe 1.3.10. An einem Tanzkurs nehmen n Männer und n Frauen teil. Wählen wir aus den Männern eine Gruppe von a Männern aus (für irgendein $a \in \{1, \dots, n\}$), so kennen diese a Männer mindestens a der Frauen schon aus ihrer Schulzeit. Ist es möglich, die Tanzpaare so einzuteilen, dass jeder Mann mit einer Bekannten aus seiner Schulzeit zusammenkommt?

Kapitel 2

Algebraische Strukturen

2.1 Mengen und Relationen

Mengen spielen eine zentrale Rolle in der Mathematik (und nicht nur dort). Wir folgen hier dem Zugang Georg Cantors (1845 - 1918) und verzichten auf eine axiomatische Einführung.

Definition 2.1.1. Eine **Menge** ist eine Zusammenfassung M von bestimmten wohlunterschiedenen Objekten m , genannt die Elemente von M , unseres Anschauungsraums oder unseres Denkens zu einem Ganzen.

Ist m ein Element von M , so schreiben wir $m \in M$, andernfalls schreiben wir $m \notin M$. Für jedes Objekt m unserer Anschauung und jede Menge M gilt also genau entweder $m \in M$ oder $m \notin M$, nicht aber beides.

Beispiel 2.1.1. • Die Menge, die keine Elemente enthält, bezeichnet man als **leere Menge**. Hierfür schreiben wir \emptyset oder $\{ \}$.

- Eine Menge M , die nur endlich viele Elemente hat, lässt sich durch vollständige Auflistung aller ihrer Elemente zwischen Mengenklammern beschreiben, z.B. $M = \{1, 3, 5, 7, 9\}$ ist die Menge, die als Elemente die Zahlen 1, 3, 5, 7, 9 enthält, und $M = \{\text{rot, blau, grün}\}$ ist die Menge, die als Elemente die Farben *rot*, *blau* und *grün* enthält.
- Auch unendliche Mengen lassen sich oft in aufzählender Form darstellen.

So wird die Menge \mathbb{N} der natürlichen Zahlen oft in der Form

$$\mathbb{N} = \{0, 1, 2, 3, 4, \dots\}$$

dargestellt, und die Menge \mathbb{Z} der ganzen Zahlen als

$$\mathbb{Z} = \{\dots - 4, -3, -2, -1, 0, 1, 2, 3, 4, \dots\}$$

- Manche Menge lassen sich nicht oder nur schwer aufzählend darstellen. Häufig ist hierfür jedoch eine beschreibende Darstellung möglich. Die Menge $\mathbb{P} = \{x \in \mathbb{N} \mid x \text{ ist eine Primzahl}\}$ etwa enthält als Elemente alle natürlichen Zahlen x , die Primzahlen sind, d.h. alle natürlichen Zahlen $x \neq 1$, die nur durch 1 und sich selbst teilbar sind.
- Eine wichtige Rolle in diesem Skript spielt die Menge \mathbb{Q} der rationalen Zahlen,

$$\mathbb{Q} = \left\{ \frac{m}{n} \mid m, n \in \mathbb{Z} \text{ und } n \neq 0 \right\}.$$

Definition 2.1.2. Eine Menge N heißt **Teilmenge** einer Menge M , geschrieben $N \subseteq M$, wenn jedes Element von N auch Element von M ist, d.h. wenn die Implikation $x \in N \implies x \in M$ gilt. In diesem Fall heißt M auch **Obermenge** von N .

Definition 2.1.3. Zwei Mengen N und M heißen gleich, geschrieben $M = N$, wenn ein Objekt x genau dann Element von N ist, wenn es Element von M ist, d.h. wenn die Äquivalenz $x \in N \iff x \in M$ gilt.

Bemerkung 2.1.1. Eine Menge N heißt **echte Teilmenge** von M , geschrieben $N \subset M$, wenn N Teilmenge von M mit $N \neq M$ ist.

Die Gleichheit von Mengen lässt sich durch die Teilmengenbeziehung beschreiben.

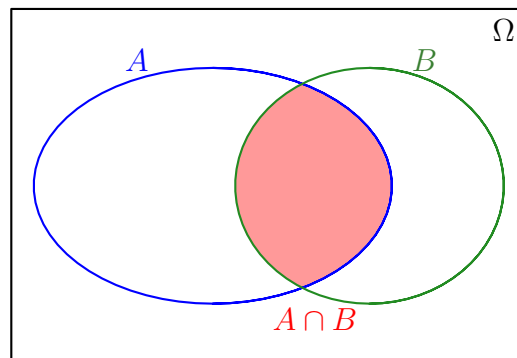
Satz 2.1.1. Für zwei Mengen M und N gilt genau dann $M = N$, wenn $M \subseteq N$ und $N \subseteq M$.

Für Teilmengen A und B einer gegebenen Menge M lassen sich verschiedene Beziehungen definieren:

Definition 2.1.4. Die **Schnittmenge** von A und B , geschrieben $A \cap B$ besteht aus den Elementen von M , die sowohl in A als auch in B sind:

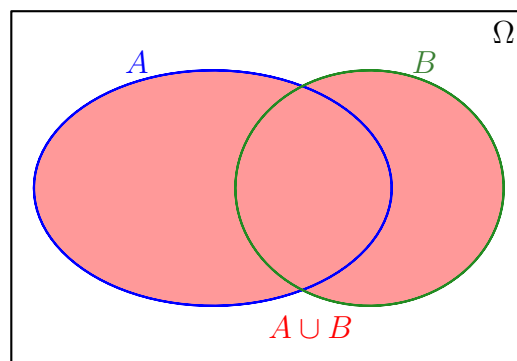
$$A \cap B = \{x \in M \mid x \in A \text{ und } x \in B\}$$

Falls $A \cap B = \emptyset$, so nennen wir A und B **disjunkt**.



Definition 2.1.5. Die **Vereinigungsmenge** von A und B , geschrieben $A \cup B$ besteht aus den Elementen von M , die entweder in A oder in B sind:

$$A \cup B = \{x \in M \mid x \in A \text{ oder } x \in B\}$$



Für die Bildung von Vereinigungs- und Durchschnittsmengen gelten folgende wichtige Regeln

Satz 2.1.2. Für Teilmengen $A, B, C \subseteq M$ gilt

<i>Kommutativgesetz</i>	$A \cup B = B \cup A$ $A \cap B = B \cap A$
<i>Assoziativgesetz</i>	$(A \cup B) \cup C = A \cup (B \cup C)$ $(A \cap B) \cap C = A \cap (B \cap C)$
<i>Distributivgesetz</i>	$(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$ $(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$
<i>Verschmelzungsgesetz</i>	$A \cap (A \cup B) = A$ $A \cup (A \cap B) = A$

Beweis: Kommutativ- und Assoziativgesetze sind klar.

Wir zeigen zunächst das erste Distributivgesetz unter Ausnutzung des ersten Distributivgesetzes der Logik aus Satz 1.1.1:

$$\begin{aligned}
 x \in (A \cup B) \cap C &\iff (x \in (A \cup B)) \wedge (x \in C) \\
 &\iff ((x \in A) \vee (x \in B)) \wedge (x \in C) \\
 &\iff ((x \in A) \wedge (x \in C)) \vee ((x \in B) \wedge (x \in C)) \\
 &\iff (x \in A \cap C) \vee (x \in B \cap C) \\
 &\iff x \in (A \cap C) \cup (B \cap C)
 \end{aligned}$$

Wir haben also gezeigt:

$$x \in (A \cup B) \cap C \iff x \in (A \cap C) \cup (B \cap C)$$

und damit

$$(A \cup B) \cap C = (A \cap C) \cup (B \cap C),$$

wie gewünscht.

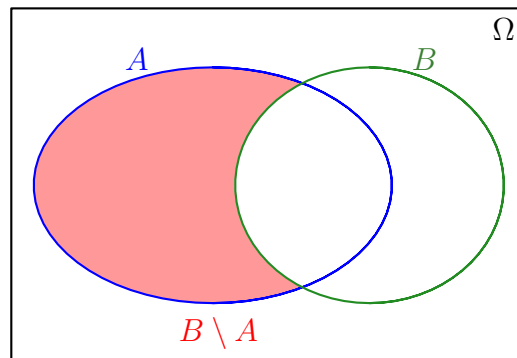
Das zweite Distributivgesetz zeigt man analog mit dem entsprechenden Distributivgesetz der Logik, siehe auch Aufgabe 2.1.2.

Den Beweis der Verschmelzungsgesetze überlassen wir dem Leser als Übungsaufgabe, siehe Aufgabe 2.1.3.

Definition 2.1.6. Die **Differenzmenge** von A und B , geschrieben $A \setminus B$ besteht aus den Elementen von M , die in A aber nicht in B sind:

$$A \setminus B = \{x \in M \mid x \in A \text{ und } x \notin B\}$$

Falls $B \subseteq A$ nennen wir $A \setminus B$ auch das **Komplement** von B in A und schreiben hierfür \overline{B}^A . Falls $A = M$, schreiben wir hierfür auch kurz \overline{B} und nennen es das Komplement von B .



Auch für die Komplementbildung gelten einige einfache Regeln:

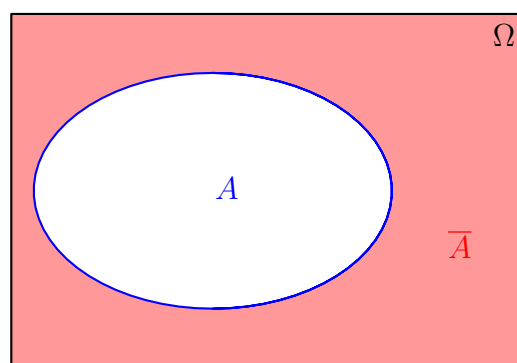
Regel 2.1.3. Für eine Teilmenge $A \subseteq M$ gilt:

1. $\overline{\overline{A}} = A$.
2. $\overline{A} \cup A = M$.
3. $\overline{A} \cap A = \emptyset$.
4. $\overline{\emptyset} = M$
5. $\overline{M} = \emptyset$.

Auch zwei beliebige Mengen können miteinander verknüpft werden:

Definition 2.1.7. Das **kartesische Produkt** zweier Mengen M und N ist die Menge $M \times N$, deren Elemente die geordneten Paare (m, n) sind, wobei $m \in M$ und $n \in N$, also

$$M \times N = \{(m, n) | m \in M \text{ und } n \in N\}$$



Beispiel 2.1.2. Es seien $M = \{1, 2, 3, 4\}$ und $N = \{3, 5\}$. Dann gilt:

$$M \times N = \{(1, 3), (1, 5), (2, 3), (2, 5), (3, 3), (3, 5), (4, 3), (4, 5)\}$$

Beachten Sie, dass es dabei keine Rolle spielt, dass manche Elemente (hier etwa die 3) in beiden Mengen vorkommen.

Beispiel 2.1.3. Ist $N = \emptyset$, so gilt $M \times N = \emptyset$ (unabhängig von M).

Definition 2.1.8. Eine **Relation** R zwischen zwei Mengen M und N ist eine Beziehung zwischen Elementen von M und N , dargestellt durch geordnete Paare (m, n) mit $m \in M$ und $n \in N$. Wir schreiben hierfür $m \sim_R n$ oder mRn und sagen *m steht in Relation mit n (bezüglich R)*.

Ist $M = N$, so heißt R auch Relation auf M . In diesem Fall nennen wir R auch **homogen**.

Bemerkung 2.1.2. Eine Relation R zwischen M und N ist ein Teilmenge $R \subseteq M \times N$ des kartesischen Produktes.

Bemerkung 2.1.3. Eine Funktion $f : D \rightarrow \mathbb{R}$ (mit $D \subseteq \mathbb{R}$) ist eine Relation R zwischen D und \mathbb{R} mit folgender besonderen Eigenschaft:

Für jedes $x \in D$ gibt es genau ein $y \in \mathbb{R}$ mit $(x, y) \in R$.

Entsprechend ist auch jede Abbildung $f : M \rightarrow N$ eine Relation R zwischen M und N mit der Eigenschaft:

Für jedes $x \in M$ gibt es genau ein $y \in N$ mit $(x, y) \in R$.

Umgekehrt definiert jede Relation R zwischen M und N mit dieser Eigenschaft auch eine Abbildung $f : M \rightarrow N$, und zwar wie folgt:

Ist $x \in M$, so gibt es genau ein Tupel $(x, y) \in M \times N$ mit $(x, y) \in R$. Setzen wir

$$f(x) = y$$

so wird dadurch eine (eindeutig bestimmte) Abbildung $f : M \rightarrow N$ definiert.

Beispiel 2.1.4. Die Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$ mit $f(x) = x^2 + x^3$ entspricht der Relation

$$R = \{(x, x^2 + x^3) \mid x \in \mathbb{R}\} \subseteq \mathbb{R} \times \mathbb{R}$$

Beispiel 2.1.5. Die Relation

$$R = \{(x, x^3) \mid x \in \mathbb{R}\} \subseteq \mathbb{R} \times \mathbb{R}$$

beschreibt eine Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$, die explizit durch $f(x) = x^3$ beschrieben wird. Diese Relation (bzw. diese Funktion) kann auch dargestellt werden als

$$R = \{(\sqrt[3]{x}, x) \mid x \in \mathbb{R}\} \subseteq \mathbb{R} \times \mathbb{R}$$

Beispiel 2.1.6. Die Relation

$$R = \{(x^2, x^3) \mid x \in \mathbb{R}\} \subseteq \mathbb{R} \times \mathbb{R}$$

beschreibt keine Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$, denn zu -1 gibt es kein y mit $(-1, y) \in R$. Allerdings definiert die Einschränkung

$$R = \{(x^2, x^3) \mid x \in \mathbb{R}_+\} \subseteq \mathbb{R}_+ \times \mathbb{R}_+$$

eine Funktion $f : \mathbb{R}_+ \rightarrow \mathbb{R}_+$, die explizit durch $f(x) = \sqrt{x^3}$ beschrieben wird.

Beispiel 2.1.7. Ist M die Menge aller Menschen, so definiert die Beziehung R : *ist Kind von* eine Relation auf M .

Beispiel 2.1.8. Ist $M = \mathbb{Z}$ die Menge der ganzen Zahlen, so definiert die Beziehung R : *ist größer oder gleich* eine Relation auf M , die mit \geq bezeichnet wird. Ein Zahlenpaar (a, b) ist also genau dann in $R \subseteq \mathbb{Z} \times \mathbb{Z}$, wenn $a \geq b$.

Beispiel 2.1.9. Ist M die beliebige Menge, so definiert die Beziehung R : *ist gleich* eine Relation auf M , die mit $=$ bezeichnet wird. Ein Paar $(a, b) \in M \times M$ ist also genau dann in R , wenn $a = b$.

Beispiel 2.1.10. Ist wieder $M = \mathbb{Z}$ die Menge der ganzen Zahlen und ist $n \in \mathbb{Z}$ eine vorgegebene Zahl, so definiert die Beziehung R : *unterscheiden sich um ein Vielfaches von n* eine Relation auf \mathbb{Z} . Ein Zahlenpaar (a, b) ist also genau dann in R wenn $a - b$ durch n teilbar ist.

Wir interessieren uns vor allem für Relation auf einer Menge M .

Definition 2.1.9. Betrachte eine Relation R auf einer Menge M .

- R heißt **reflexiv**, wenn für alle m aus M gilt: $m \sim_R m$.
- R heißt **transitiv**, wenn gilt:

$$m_1 \sim_R m_2 \text{ und } m_2 \sim_R m_3 \implies m_1 \sim_R m_3$$

- R heißt **symmetrisch**, wenn gilt:

$$m_1 \sim_R m_2 \implies m_2 \sim_R m_1$$

- R heißt **Äquivalenzrelation**, wenn R reflexiv, transitiv und symmetrisch ist.
- R heißt **antisymmetrisch**, wenn gilt:

$$m_1 \sim_R m_2 \text{ und } m_2 \sim_R m_1 \implies m_1 = m_2$$

- R heißt **asymmetrisch**, wenn gilt:

$$m_1 \sim_R m_2 \implies \neg(m_2 \sim_R m_1)$$

Beispiel 2.1.11. Die Relation R : *ist Kind von* aus Beispiel 2.1.7 ist weder transitiv noch reflexiv noch symmetrisch. Betrachten wir stattdessen die Relation R' : *ist Nachkomme von*, so ist R' transitiv, aber nicht reflexiv (wenn wir eine Person nicht als ihren eigenen Nachkommen auffassen wollen) und nicht symmetrisch. Betrachten wir R'' : *ist verwandt mit*, so ist R'' reflexiv, transitiv und symmetrisch, also eine Äquivalenzrelation.

Die Relation \geq aus Beispiel 2.1.8 ist reflexiv, transitiv und antisymmetrisch. Sie ist nicht symmetrisch.

Die Relation $=$ aus Beispiel 2.1.9 ist reflexiv, transitiv, symmetrisch und antisymmetrisch.

Die Relation R : *unterscheiden sich um ein Vielfaches von n* aus Beispiel 2.1.10 ist eine Äquivalenzrelation. Sie ist nicht antisymmetrisch

Jede Äquivalenzrelation \sim auf M liefert uns eine natürliche Zerlegung von M in disjunkte Teilmengen:

Definition 2.1.10. Wir betrachten eine Äquivalenzrelation \sim_R auf M .

Zwei Elemente $m, n \in M$ heißen **äquivalent** (bezüglich \sim_R), wenn $m \sim_R n$.

Eine Teilmenge $A \subseteq M$ heißt **Äquivalenzklasse**, wenn gilt

- Sind $m, n \in A$, so ist $m \sim_R n$.
- Ist $m \in A$ und $n \in M$ mit $n \sim_R m$, so ist $n \in A$.

Ist $m \in M$, so heißt $[m]_R := \{n \in M : n \sim_R m\}$ die **Äquivalenzklasse von m** .

Bemerkung 2.1.4. Ist \sim_R eine Äquivalenzrelation auf M und sind $m, n \in M$, so gilt entweder $[m]_R = [n]_R$ oder $[m]_R$ und $[n]_R$ sind disjunkt. Das folgt sofort aus den Definitionen.

Eine Äquivalenzrelation \sim_R auf M zerlegt also M in disjunkte Teilmengen, die Äquivalenzklassen. Wir schreiben M / \sim_R für die Menge der Äquivalenzklassen, also

$$M / \sim_R = \{[m]_R \mid m \in M\}$$

Definition 2.1.11. Ist R eine Äquivalenzrelation auf M und A eine Äquivalenzklasse (bezüglich R), so heißt ein beliebiges Element $a \in A$ ein **Repräsentant** der Äquivalenzklasse A .

Ein **Repräsentantensystem** der Äquivalenzrelation R ist eine Teilmenge $N \subseteq M$ die genau einen Repräsentanten jeder Äquivalenzklasse enthält.

Beispiel 2.1.12. Für die Äquivalenzrelation \sim_R : *unterscheiden sich um ein Vielfaches von n* aus Beispiel 2.1.10 gilt:

- Falls $n = 0$:

$$\mathbb{Z} / \sim_R = \mathbb{Z}$$

- Falls $n > 0$:

$$\mathbb{Z} / \sim_R = \{[0]_R, [1]_R, \dots, [n-1]_R\}$$

- Falls $n < 0$:

$$\mathbb{Z} / \sim_R = \{[0]_R, [1]_R, \dots, [-n-1]_R\}$$

Für $n \neq 0$ sind die angegebenen Äquivalenzklassen paarweise disjunkt. Wir schreiben in diesem Fall auch \mathbb{Z}_n , $\mathbb{Z}/(n)$ oder $\mathbb{Z}/n\mathbb{Z}$ für \mathbb{Z}/\sim_R .

Die Menge $\{0, 1, \dots, n-1\}$ bildet also ein Repräsentantensystem der Relation R . Es gibt aber noch viele weitere Repräsentantensysteme, etwa $\{1, 2, \dots, n\}$ oder $\{n, n+1, \dots, 2n-1\}$ oder $\{0, n+1, 2n+2, 3n+3, \dots, n^2-1\}$.

Neben Äquivalenzrelation eine besondere Rolle spielen Vergleichsrelationen, wie etwa die Relationen \geq oder $>$ auf den reellen Zahlen.

Definition 2.1.12. Es sei R eine Relation auf eine Menge M .

R heißt **Ordnungsrelation** oder **Ordnung** auf M , wenn sie reflexiv, transitiv und antisymmetrisch ist.

R heißt **strikte Ordnungsrelation** oder **strikte Ordnung** auf M , wenn sie asymmetrisch und transitiv ist.

Beispiel 2.1.13. Die Relation \geq aus Beispiel 2.1.8 ist eine Ordnungsrelation aber keine strikte Ordnungsrelation.

Beispiel 2.1.14. Die Relation $>$ auf den ganzen Zahlen ist eine strikte Ordnungsrelation.

Beispiel 2.1.15. Wir betrachten eine Menge $M = \{m_1, m_2, \dots, m_t\}$ von Menschen. Auf M definieren wir die Relation R durch $m_1 R m_2$ genau dann, wenn m_1 älter ist als m_2 . Dann ist R eine strikte Ordnung auf M . Die Relation R' definiert durch $m_1 R' m_2$ genau dann, wenn m_1 mindestens so alt ist wie m_2 ist eine Ordnung auf M .

Aufgabe 2.1.1. Bestimmen Sie jeweils die Mengen $A \cup B$, $A \cap B$, $A \setminus B$ und $(A \cup B) \setminus (A \cap B)$ möglichst explizit:

a) $A = \{x \in \mathbb{Z} : -5 \leq x \leq 10\}$, $B = \{x \in [-2, 8] : 2x \in \mathbb{N}\}$.

b) $A = \{x \in \mathbb{R} : x^2 \leq 1\}$, $B = \{x \in \mathbb{R} : (x-1)^2 \leq 1\}$.

Aufgabe 2.1.2. Zeigen Sie das zweite Distributivgesetz der Mengenlehre:

$$(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$$

Aufgabe 2.1.3. Zeigen Sie die Verschmelzungsgesetze der Mengenlehre:
Für Teilmengen $A, B \subseteq M$ gilt:

$$\begin{aligned} A \cap (A \cup B) &= A \\ A \cup (A \cap B) &= A \end{aligned}$$

Aufgabe 2.1.4. Zeigen Sie die **de Morganschen Regeln** der Mengenlehre:
Für Teilmengen $A, B \subseteq M$ gilt:

$$\begin{aligned} \overline{A \cap B} &= \overline{A} \cup \overline{B} \\ \overline{A \cup B} &= \overline{A} \cap \overline{B} \end{aligned}$$

Aufgabe 2.1.5. Zeigen Sie, dass für zwei Teilmengen $A, B \subseteq M$ gilt:

$$A \setminus B = \overline{\overline{A} \cup B}$$

Aufgabe 2.1.6. Auf der Menge $\mathbb{Z} \times \mathbb{Z}$ definieren wir ein Relation R durch $(a, b)R(c, d)$ genau dann, wenn $a + b = c + d$ ist. Zeigen sie, dass R eine Äquivalenzrelation auf $\mathbb{Z} \times \mathbb{Z}$ ist und bestimmen Sie ein Repräsentantensystem von R .

Aufgabe 2.1.7. Auf der Menge $M := \{(a, b) \mid a \in \mathbb{Z} \wedge b \in \mathbb{Z} \setminus \{0\}\}$ definieren wir ein Relation R durch $(a, b)R(c, d)$ genau dann, wenn $a \cdot d = c \cdot b$ ist. Zeigen sie, dass R eine Äquivalenzrelation auf M ist.

Bemerkung 2.1.5. Die Menge M / \sim_R der Äquivalenzklassen der Relation R aus Aufgabe 2.1.7 ist die Menge der rationalen Zahlen.

Aufgabe 2.1.8. Heute ist ein Donnerstag. Welcher Wochentag ist in 1000 Tagen? Welcher in 3572? Welcher Wochentag war vor 8000 Tagen?

Aufgabe 2.1.9. Auf der Menge $\mathbb{Z} \times \mathbb{Z}$ definieren wir ein Relation R durch $(a, b)R(c, d)$ genau dann, wenn $a < c$ oder $a = c$ und $b \leq d$ ist. Zeigen sie, dass R eine Ordnungsrelation auf $\mathbb{Z} \times \mathbb{Z}$ ist.

Aufgabe 2.1.10. Es sei M eine Menge und $\mathfrak{P}(M)$ ihre Potenzmenge. Zeigen Sie, dass die Relation R auf $\mathfrak{P}(M)$, gegeben durch $N_1 R N_2$ genau dann, wenn $N_1 \subseteq N_2$ eine Ordnung auf $\mathfrak{P}(M)$ definiert, und dass die Relation R' auf $\mathfrak{P}(M)$, gegeben durch $N_1 R' N_2$ genau dann, wenn $N_1 \subset N_2$ eine strikte Ordnung auf $\mathfrak{P}(M)$ definiert.

2.2 Gruppen

Wir haben in den vorangegangenen Abschnitten verschiedene Mengen mit unterschiedlichen Strukturen kennengelernt. So haben wir etwa gesehen, dass wir auf den reellen Zahlen \mathbb{R} eine Addition und eine Multiplikation haben, für die einige Gesetze erfüllt sind, nämlich die Gesetze $A1 - A4$, $M1 - M4$ und D . Ähnliches gilt für die rationalen Zahlen \mathbb{Q} . Wenn wir dagegen die ganzen Zahlen \mathbb{Z} betrachten, so sehen wir, dass hier nicht mehr alle Gesetze gelten, denn nicht zu jeder ganzen Zahl $m \in \mathbb{Z}$ gibt es eine ganze Zahl $n \in \mathbb{Z}$ mit $m \cdot n = 1$. Also gilt $M4$ in \mathbb{Z} nicht. Wenn wir uns noch weiter einschränken und nur noch die natürlichen Zahlen \mathbb{N} betrachten, dann sehen wir, dass hier auch $A4$ nicht mehr gilt. Wir werden noch viele weitere Beispiele kennenlernen, in denen einige dieser Gesetze gelten, aber nicht alle. Daher wollen wir diese Gesetze und Gruppen dieser Gesetze in diesem Abschnitt isoliert voneinander betrachten.

Definition 2.2.1. Es sei M eine Menge. Eine Abbildung

$$\circ : M \times M \longrightarrow M$$

also eine Abbildung mit Definitionsbereich $M \times M$ und Bildbereich M heißt **(innere) Verknüpfung** von M .

Wir schreiben in diesem Fall $m \circ n$ für $\circ(m, n)$.

Beispiel 2.2.1. Es sei $M = \mathbb{Z}$ und

$$\circ = ' + ' : \mathbb{Z} \times \mathbb{Z} \longrightarrow \mathbb{Z}, \quad (a, b) \longmapsto a + b$$

die Addition ganzer Zahlen. Dann ist $' + '$ eine innere Verknüpfung auf \mathbb{Z} .

Beispiel 2.2.2. Es sei $M = \{f : \mathbb{R} \longrightarrow \mathbb{R}\}$ die Menge aller Abbildungen $f : \mathbb{R} \longrightarrow \mathbb{R}$ und es sei

$$\circ : M \times M \longrightarrow M, \quad (f, g) \longmapsto g \circ f$$

die Komposition von zwei Abbildungen. Dann ist \circ eine innere Verknüpfung von M .

Beispiel 2.2.3. Es sei $M = \mathbb{R}$ und es sei

$$\circ : M \times M \longrightarrow M, \quad (a, b) \longmapsto \max\{a, b\}$$

die Zuordnung, die jedem Paar reeller Zahlen die größere von beiden zuordnet. Dann ist \circ eine innere Verknüpfung von M . (Analoges gilt für das Minimum).

Beispiel 2.2.4. Es sei M eine beliebige Menge und es sei $\mathfrak{P}(M)$ ihre Potenzmenge. Dann wird durch

$$\circ : \mathfrak{P}(M) \times \mathfrak{P}(M) \longrightarrow \mathfrak{P}(M), \quad (A, B) \longmapsto A \cup B$$

also das Bilden der Vereinigungsmenge, eine innere Verknüpfung auf $\mathfrak{P}(M)$ definiert. (Analoges gilt für die Durchschnittsbildung).

Beispiel 2.2.5. Auf der Menge $M = \mathbb{R}$ definieren wir

$$\circ : M \times M \longrightarrow M$$

durch

$$\circ(a, b) = \begin{cases} 1 & \text{falls } a \geq b \\ 0 & \text{falls } a < b \end{cases}$$

Dann ist \circ eine innere Operation auf M .

Beispiel 2.2.6. Es sei $M = \mathbb{Z}$ und

$$\circ = '0' : \mathbb{Z} \times \mathbb{Z} \longrightarrow \mathbb{Z}, \quad (a, b) \longmapsto 0$$

die die Nullabbildung. Dann ist $'0'$ eine innere Verknüpfung auf \mathbb{Z} .

Für eine Menge M mit einer Verknüpfung \circ schreiben wir kurz (M, \circ) .

Definition 2.2.2. Eine nichtleere Menge (M, \circ) mit einer Verknüpfung \circ heißt **Halbgruppe**, wenn \circ das Assoziativgesetz erfüllt, also wenn gilt:

$$(n \circ m) \circ l = n \circ (m \circ l) \quad \text{für alle } l, m, n \in M$$

Beispiel 2.2.7. $(\mathbb{N}, +)$ ist eine Halbgruppe.

Beispiel 2.2.8. $(\mathbb{N} \setminus \{0\}, +)$ ist eine Halbgruppe.

Beispiel 2.2.9. $(\mathbb{N} \setminus \{0\}, \cdot)$ ist eine Halbgruppe.

Beispiel 2.2.10. (\mathbb{Z}, \cdot) ist eine Halbgruppe

Beispiel 2.2.11. Ist $M = \{f : \mathbb{R} \rightarrow \mathbb{R}\}$ die Menge aller Abbildungen aus Beispiel 2.2.2 mit Verknüpfung \circ , so ist (M, \circ) eine Halbgruppe.

Beispiel 2.2.12. $(\mathbb{Z}, '0')$ aus Beispiel 2.2.6 ist eine Halbgruppe

Beispiel 2.2.13. (\mathbb{R}, \circ) aus Beispiel 2.2.3 ist eine Halbgruppe.

Beispiel 2.2.14. $(\mathfrak{P}(M), \circ)$ aus Beispiel 2.2.4 ist eine Halbgruppe.

Beispiel 2.2.15. (\mathbb{R}, \circ) aus Beispiel 2.2.5 ist keine Halbgruppe. Hierfür gilt nämlich

$$((-1) \circ (-2)) \circ (-3) = 1 \neq 0 = (-1) \circ ((-2) \circ (-3))$$

Definition 2.2.3. Ein Element e einer Halbgruppe (M, \circ) heißt **neutrales Element** der Halbgruppe, wenn gilt

$$m \circ e = m, \quad e \circ m = m \quad \text{für alle } m \in M$$

Eine Halbgruppe (M, \circ) mit neutralem Element e heißt **Monoid**. Wir schreiben hierfür auch (M, \circ, e)

Bemerkung 2.2.1. Das neutrale Element eines Monoids (M, \circ) ist eindeutig. Sind nämlich e und e' zwei Elemente aus M mit der Eigenschaft des neutralen Elements, so folgt aus ebendieser Eigenschaft

$$e' = e \circ e' = e$$

Beispiel 2.2.16. $(\mathbb{N}, +, 0)$ ist ein Monoid mit neutralem Element 0.

Beispiel 2.2.17. $(\mathbb{N} \setminus \{0\}, +)$ ist kein Monoid.

Beispiel 2.2.18. $(\mathbb{N} \setminus \{0\}, \cdot)$ ist ein Monoid mit neutralem Element 1.

Beispiel 2.2.19. $(\mathbb{Z}, +, 0)$ ist ein Monoid mit neutralem Element 0. (\mathbb{Z}, \cdot) ist ein Monoid mit neutralem Element 1.

Beispiel 2.2.20. Ist $M = \{f : \mathbb{R} \longrightarrow \mathbb{R}\}$ die Menge aller Abbildungen aus Beispiel 2.2.2 mit Verknüpfung \circ , so ist (M, \circ) ein Monoid mit neutralem Element

$$e : \mathbb{R} \longrightarrow \mathbb{R}, \quad x \longmapsto x.$$

Beispiel 2.2.21. $(\mathbb{Z}, '0')$ aus Beispiel 2.2.6 ist kein Monoid.

Definition 2.2.4. Ist (M, \circ) ein Monoid mit neutralem Element e und ist $m \in M$, so heißt ein Element $n \in M$ **inverses Element** zu m wenn gilt

$$m \circ n = e, \quad n \circ m = e$$

In diesem Fall schreiben wir m^{-1} für n . Ein Monoid (G, \circ) heißt **Gruppe**, wenn es zu jedem Element $m \in G$ ein inverses Element in G gibt.

Beispiel 2.2.22. $(\mathbb{N}, +)$ ist keine Gruppe. So gibt es etwa keine natürliche Zahl n mit $1 + n = 0$.

Beispiel 2.2.23. $(\mathbb{Z}, +)$ ist eine Gruppe. Für jede Zahl z gilt nämlich

$$z + (-z) = 0$$

und mit z ist auch $-z$ in \mathbb{Z} . (\mathbb{Z}, \cdot) ist keine Gruppe. So gibt es etwa zu 0 keine ganze Zahl z mit $0 \cdot z = 1$.

Beispiel 2.2.24. $(\mathbb{R}, +), (\mathbb{Q}, +)$ sind Gruppen, aber $(\mathbb{R}, \cdot), (\mathbb{Q}, \cdot)$ sind keine Gruppen, da es bezüglich \cdot kein inverses Element zu 0 gibt.

$(\mathbb{R} \setminus \{0\}, \cdot), (\mathbb{Q} \setminus \{0\}, \cdot)$ sind Gruppen, da es zu jeder von Null verschiedenen reellen oder rationalen Zahl ein multiplikatives Inverses in \mathbb{R} oder \mathbb{Q} gibt, aber $(\mathbb{Z} \setminus \{0\}, \cdot)$ ist keine Gruppe, denn 2 hat kein multiplikatives Inverses in \mathbb{Z} .

Beispiel 2.2.25. Ist $M = \{f : \mathbb{R} \longrightarrow \mathbb{R}\}$ die Menge aller Abbildungen aus Beispiel 2.2.2 mit Verknüpfung \circ , so ist (M, \circ) keine Gruppe. Ist nämlich $f \in M$ die Nullabbildung, $f(x) = 0$, so gibt es kein inverses Element zu f . Ist aber $M' = \{f : \mathbb{R} \longrightarrow \mathbb{R}\}$ die Menge aller bijektiven Abbildungen aus M mit Verknüpfung \circ , so ist (M', \circ) eine Gruppe. Ist nämlich $f \in M'$, so existiert eine inverse Abbildung $f^{-1} \in M'$ mit $f \circ f^{-1} = e$ und $f^{-1} \circ f = e$, wie wir im Abschnitt über Funktionen gesehen haben.

Ganz allgemein gilt:

Ist A eine beliebige Menge, so ist

$$\text{Aut}(A) = \{f : A \longrightarrow A \mid f \text{ ist bijektiv}\}$$

zusammen mit der Komposition von Abbildungen eine Gruppe.

Beispiel 2.2.26. Ist $M_n = \{1, 2, \dots, n\}$ die Menge der Zahlen $1, 2, \dots, n$, und ist S_n die Menge der bijektiven Abbildungen auf M_n , so heißt S_n **Permutationsgruppe** der Zahlen $1, \dots, n$ und ihre Elemente heißen Permutationen von $1, \dots, n$.

Ein Element $\sigma \in S_n$ lässt sich am besten tabellarisch darstellen

1	2	...	n
$\sigma(1)$	$\sigma(2)$...	$\sigma(n)$

Hierfür schreiben wir auch

$$\sigma = \begin{pmatrix} 1 & 2 & \dots & n \\ \sigma(1) & \sigma(2) & \dots & \sigma(n) \end{pmatrix}$$

In der zweiten Zeile dieser Darstellung tauchen wieder alle Zahlen $1, \dots, n$ auf, allerdings in geänderter, also permutierter Reihenfolge. Darum nennen wir die Elemente $\sigma \in S_n$ auch **Permutationen** der Zahlen $1, \dots, n$.

So ist etwa

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 2 & 3 & 1 \end{pmatrix}$$

die Permutation der Zahlen $1, 2, 3, 4$, die die Zahlen 1 und 4 vertauscht (und alle anderen festlässt. Hierfür schreiben wir auch kurz $\langle 1\ 4 \rangle$. Ganz allgemein schreiben wir für eine Permutation $\sigma \in S_n$, die nur die Zahlen i und j ($i \neq j$) vertauscht kurz $\langle i\ j \rangle$ und nennen sie die **Transposition** der Zahlen i und j . Wegen $\langle i\ j \rangle = \langle j\ i \rangle$ können wir dabei stets annehmen, dass $i < j$.

Jede Permutation lässt sich (in nicht-eindeutiger Weise) aus Transpositionen zusammensetzen, d.h. ist $\sigma \in S_n$ so gibt es Zahlenpaare $(i_1, j_1), (i_2, j_2), \dots, (i_t, j_t)$ mit $1 \leq i_\tau < j_\tau \leq n$, so dass

$$\sigma = \langle i_1\ j_1 \rangle \circ \langle i_2\ j_2 \rangle \circ \dots \circ \langle i_t\ j_t \rangle$$

Mit $\text{id} = \text{id}_n$ bezeichnen wir das Element von S_n , das die Reihenfolge der Zahlen nicht verändert, also

$$\text{id} = \begin{pmatrix} 1 & 2 & \dots & n \\ 1 & 2 & \dots & n \end{pmatrix}$$

Eine Permutation entspricht einer Anordnung der Zahlen $1, \dots, n$, also einem geordnete n -Tupel (a_1, \dots, a_n) mit $a_i \in \{1, \dots, n\}$ und $a_i \neq a_j$ für $i \neq j$. Eine einfache kombinatorische Übung zeigt, dass

$$|S_n| = n!$$

Permutationen werden uns noch öfter begegnen, etwa bei den Determinanten. Daher wollen wir sie noch etwas genauer untersuchen.

Ist $\tau = \langle i, j \rangle$ die Transposition der Zahlen i, j , so gilt hierfür

$$\tau \circ \tau = \text{id}$$

und damit gilt allgemein für eine beliebige Permutation σ mit Darstellung

$$\sigma = \langle i_1 j_1 \rangle \circ \langle i_2 j_2 \rangle \circ \dots \circ \langle i_t j_t \rangle$$

die Beziehung

$$\sigma \circ \langle i_t j_t \rangle \circ \dots \circ \langle i_1 j_1 \rangle = \text{id}$$

so dass wir also auch für σ^{-1} eine Darstellung mit Transpositionen gefunden haben:

$$\sigma^{-1} = \langle i_t j_t \rangle \circ \dots \circ \langle i_1 j_1 \rangle$$

Eine wichtige Größe bei der Betrachtung von Permutationen ist die Anzahl der Fehlstände. Dabei heißt ein Paar $i, j \in \{1, \dots, n\}$ eine Fehlstand von σ , wenn $i < j$ aber $\sigma(i) > \sigma(j)$. Wir definieren die **Signatur** $\text{sign}(\sigma)$ von σ durch

$$\text{sign}(\sigma) = \begin{cases} +1 & \text{falls die Anzahl der Fehlstände gerade ist} \\ -1 & \text{falls die Anzahl der Fehlstände ungerade ist} \end{cases}$$

Wir nennen eine Permutation σ **gerade**, wenn $\text{sign}(\sigma) = +1$ und **ungerade**, wenn $\text{sign}(\sigma) = -1$.

Die Signatur hat auch folgende Beschreibung

$$\text{sign}(\sigma) = \prod_{i < j} \frac{\sigma(j) - \sigma(i)}{j - i}$$

wie man sehr leicht nachrechnet (siehe auch Aufgabe 2.4.7). Dabei steht \prod für das Produktzeichen, also

$$\prod_{i=1}^n a_i = a_1 \cdot a_2 \cdots a_n$$

Wir notieren die folgenden Eigenschaften der Signatur:

- Ist $\tau = \langle i j \rangle$ eine Transposition, so gilt $\text{sign}(\tau) = -1$
- Sind σ, τ zwei Permutationen, so gilt $\text{sign}(\sigma \circ \tau) = \text{sign}(\sigma) \cdot \text{sign}(\tau)$ (vergleiche Aufgabe 2.4.8).
- $\text{sign}(\text{id}) = 1$ und $\text{sign}(\sigma^{-1}) = \text{sign}(\sigma)$.
- Ist $\sigma = \langle i_1 j_1 \rangle \circ \langle i_2 j_2 \rangle \circ \cdots \circ \langle i_t j_t \rangle$, so ist $\text{sign}(\sigma) = (-1)^t$.
- Für $n \geq 2$ gibt es genauso viele gerade Permutationen wie ungerade Permutationen.

Beispiel 2.2.27. Sei $\mathcal{PM} = \{1, -1\}$ mit der Komposition $a \circ b = a \cdot b$. Dann ist \mathcal{PM} eine Gruppe.

Definition 2.2.5. Eine Gruppe (G, \circ) heißt **kommutativ** oder **abelsch** wenn für je zwei Elemente $g, h \in G$ gilt:

$$g \circ h = h \circ g$$

Beispiel 2.2.28. $(\mathbb{Z}, +)$, $(\mathbb{R}, +)$ und $(\mathbb{Q}, +)$ sind kommutative Gruppen. $(\mathbb{R} \setminus \{0\}, \cdot)$ und $(\mathbb{Q} \setminus \{0\}, \cdot)$ sind kommutative Gruppen

Beispiel 2.2.29. Ist G die Gruppe der bijektiven Abbildungen von \mathbb{R} in sich, so ist G nicht kommutativ, wie wir schon im Beispiel ?? in Abschnitt ?? gesehen haben.

Beispiel 2.2.30. Für $n \geq 3$ ist die Gruppe S_n der Permutationen aus Beispiel 2.2.26 nicht kommutativ. So ist etwa $\langle 1 2 \rangle \circ \langle 1 3 \rangle \neq \langle 1 3 \rangle \circ \langle 1 2 \rangle$.

Definition 2.2.6. Eine Gruppe (G, \circ) heißt **endlich**, wenn $|G| < \infty$, also wenn G nur endlich viele Elemente hat.

In diesem Fall heißt $\text{ord}(G) = |G|$ die **Ordnung** der Gruppe G .

Bemerkung 2.2.2. Eine endliche Gruppe G kann durch ihre Verknüpfungstafel beschrieben werden. Dazu schreiben wir $G = \{g_1, g_2, \dots, g_n\}$ und erstellen eine Tabelle der Form

\circ	g_1	g_2	\dots	g_n
g_1	$g_{1,1}$	$g_{1,2}$	\dots	$g_{1,n}$
g_2	$g_{2,1}$	$g_{2,2}$	\dots	$g_{2,n}$
\vdots	\vdots	\vdots	\ddots	\vdots
g_n	$g_{n,1}$	$g_{n,2}$	\dots	$g_{n,n}$

wobei $g_{i,j}$ kurz für das Element $g_i \circ g_j$ steht. Genau dann ist die Gruppe kommutativ, wenn die Verknüpfungstafel symmetrisch (bezüglich der Diagonale) ist.

Beispiel 2.2.31. Die Gruppe \mathcal{PM} aus Beispiel 2.2.27 hat genau zwei Elemente, ist also endlich. Ihre Verknüpfungstafel ist

\circ	-1	1
-1	1	-1
1	-1	1

Beispiel 2.2.32. Wir betrachten die Menge $G = \{0, 1\}$ mit der durch die folgende Tafel gegebenen Verknüpfung

\circ	0	1
0	0	1
1	1	0

Hierdurch wird eine kommutative Gruppe definiert (die Gruppe $\mathbb{Z}/2\mathbb{Z}$).

Beispiel 2.2.33. Wir betrachten die Menge $G = \{0, 1, 2\}$ mit der durch die folgende Tafel gegebenen Verknüpfung

\circ	0	1	3
0	0	1	2
1	1	2	0
2	2	0	1

Hierdurch wird eine kommutative Gruppe definiert (die Gruppe $\mathbb{Z}/3\mathbb{Z}$).

Beispiel 2.2.34. Für $n \geq 1$ betrachten wir $\mathbb{Z}/n\mathbb{Z}$, die Menge der Äquivalenzklassen der Relation "unterscheiden sich um ein Vielfaches von n " aus Beispiel 2.1.12. Hierauf definieren wir eine Verknüpfung "+" durch

$$[a] + [b] = [a + b]$$

Mit dieser Verknüpfung wird $\mathbb{Z}/n\mathbb{Z}$ eine kommutative Gruppe (siehe auch Aufgabe 2.4.1).

Beispiel 2.2.35. Die Permutationsgruppe S_n aus Beispiel 2.2.26 ist endlich.

Für ein Element $g \in (G, \circ)$ mit neutralem Element e setzen wir $g^0 = e$, $g^2 = g \circ g$, und allgemein für $n \geq 1$:

$$\begin{aligned} g^n &= \underbrace{g \circ g \circ \cdots \circ g}_{n\text{-mal}} \\ g^{-n} &= \underbrace{g^{-1} \circ g^{-1} \circ \cdots \circ g^{-1}}_{n\text{-mal}} \end{aligned}$$

(so dass also $g^n \circ g^m = g^{n+m}$ für alle $n, m \in \mathbb{Z}$).

Hilfssatz 2.2.1. Ist G eine endliche Gruppe und $g \in G$, so gibt es ein $n \geq 1$ mit $g^n = e$.

Beweis: Wir betrachten die Menge $B(g) := \{g = g^1, g^2, \dots, g^n, \dots\} \subseteq G$. Da G selbst endlich ist, ist es sicherlich auch $B(g)$, und daher gibt es $l, k \geq 1$, $l < k$, mit $g^l = g^k$. Dann gilt

$$e = g^l \circ g^{-l} = g^k \circ g^{-l} = g^{k-l}$$

wir können also $n = k - l$ setzen.

Definition 2.2.7. Ist G eine endliche Gruppe und $g \in G$, so heißt

$$\text{ord}(g) := \min\{n \geq 1 : g^n = e\}$$

die **Ordnung** von g .

Bemerkung 2.2.3. Wir betrachten eine endliche Gruppe G .

1. Genau dann gilt $\text{ord}(g) = 1$, wenn $g = e$.

2. Für alle $1 \leq i < j \leq \text{ord}(g)$ gilt: $g^i \neq g^j$.

Um den Ordnungsbegriff besser studieren zu können, benötigen wir eine Eigenschaft, die in den ganzen Zahlen bzw. im Monoid $(\mathbb{N} \setminus \{0\}, \cdot)$ definiert ist:

Definition 2.2.8. Sind $m, n \in \mathbb{N} \setminus \{0\}$, so heißt m **Teiler** von n , wenn es ein $a \in \mathbb{N}$ gibt mit $a \cdot m = n$, und m heißt **echter Teiler** von n , wenn es ein $a \in \mathbb{N} \setminus \{1, n\}$ gibt mit $a \cdot m = n$.

Ist m ein Teiler von n so schreiben wir hierfür $m|n$. Entsprechend schreiben wir $m \nmid n$, falls m kein Teiler von n ist.

Beispiel 2.2.36. Ist $n = 6$, so sind $m = 1, 2, 3, 6$ die Teiler von n und $m = 2, 3$ die echten Teiler von n .

Hilfssatz 2.2.2. Ist G eine endliche Gruppe und $g \in G$, so gilt

$$\text{ord}(g) \mid \text{ord}(G)$$

also die Ordnung des Elements g teilt die Ordnung der Gruppe G . Speziell gilt also

$$g^{\text{ord}(G)} = e$$

Beweisidee.:

Es sei $n = \text{ord}(g)$. Wie wir schon im Beweis von Lemma 2.2.1 gesehen haben, ist die Menge $B(g) = \{g = g^1, g^2, \dots\}$ endlich, und zwar gilt genauer

$$B(g) = \{g, g^2, \dots, g^{n-1}, g^n = e\}$$

und $B(g)$ hat genau $n = \text{ord}(g)$ verschiedene Elemente. (Wären nämlich zwei Elemente von in dieser Auflistung gleich, so könnten wir den Beweis von Lemma 2.2.1 imitieren und fänden ein $l < n$ mit $g^l = e$, im Widerspruch zur Definition der Ordnung des Elements).

Wir überdecken nun G mit den **Bahnen** von g , dh. wir definieren für $h \in G$ beliebig

$$B_h(g) = \{h \circ g, h \circ g^2, \dots, h \circ g^{n-1}, h \circ e = h\}$$

Hierfür rechnet man leicht nach

- $B_h(g)$ hat genau n Elemente.
- Für $h, h' \in G$ gilt entweder $B_h(g) = B_{h'}(g)$ oder $B_h(g) \cap B_{h'}(g) = \emptyset$.

Da G endlich ist, gibt es auch nur endlich viele verschiedene Bahnen, d.h. wir finden $h_1 = e, h_2, \dots, h_t$ mit $B_{h_i}(g) \cap B_{h_j}(g) = \emptyset$ für $i \neq j$ und $G = B_{h_1}(g) \cup \dots \cup B_{h_t}(g)$.

Damit gilt aber

$$\begin{aligned} \text{ord}(G) &= |B_{h_1}(g) \cup \dots \cup B_{h_t}(g)| \\ &= |B_{h_1}(g)| + \dots + |B_{h_t}(g)| \\ &= t \cdot n \end{aligned}$$

und es folgt $n | \text{ord}(G)$ und

$$g^{\text{ord}(G)} = g^{nt} = (g^n)^t = e^t = e,$$

also die Behauptung.

Bezeichnung.:

Ist G eine endliche Gruppe und $g \in G$ so schreiben wir

$$\langle g \rangle := B(g) = \{g = g^1, g^2, g^3, \dots\}$$

für die Teilmenge von G , bestehend aus allen Potenzen von g .

Bemerkung 2.2.4. Die Menge $\langle g \rangle$ ist selbst wieder eine (endliche) Gruppe der Ordnung $\text{ord}(g)$.

Bemerkung 2.2.5. Ist g ein Element einer endlichen Gruppe G und ist $n \geq 1$, so gilt

$$\text{ord}(g^n) | \text{ord}(g)$$

Beweis: Das Element g^n ist Element der endlichen Gruppe $\langle g \rangle$, und damit folgt die Behauptung aus Lemma 2.2.2.

Definition 2.2.9. Eine endliche Gruppe G heißt **zyklisch**, wenn es ein $g \in G$ gibt mit

$$\langle g \rangle = G$$

In diesem Fall heißt g ein **Erzeuger** von G .

Beispiel 2.2.37. Die Gruppe $\mathbb{Z}/(n)\mathbb{Z}$ ist eine zyklische Gruppe mit Erzeuger $[1]$.

Bemerkung 2.2.6. Ist G eine endliche Gruppe mit $\text{ord}(G) = p$, wobei p eine Primzahl ist, so ist G zyklisch. Dazu wählen wir ein beliebige $g \neq e$ in G . Da es nur ein neutrales Element e in G gibt, existiert ein solches g . Gemäß Lemma 2.2.2 gilt dann $\text{ord}(g) | p$, also, da p eine Primzahl ist, $\text{ord}(g) = 1$ oder $\text{ord}(g) = p$. Da aber $g \neq e$ muss gelten $\text{ord}(g) \neq 1$, und damit $\text{ord}(g) = p$, also $\langle g \rangle = G$.

2.3 Ringe

Besonders interessant für uns sind Strukturen wie \mathbb{Z} oder \mathbb{R} , die zwei Verknüpfungen tragen.

Definition 2.3.1. Ein **Ring** $(R, +, \cdot)$ ist eine nichtleere Menge R mit zwei Verknüpfungen $+$ und \cdot , die zwei Elemente $0, 1$ enthält, wobei $0 \neq 1$ ist und wobei gilt

- $(R, +, 0)$ ist eine kommutative Gruppe.
- $(R, \cdot, 1)$ ist ein Monoid.
- Es gelten die Distributivgesetze

$$\begin{aligned} (r + s) \cdot t &= r \cdot t + s \cdot t && \text{für alle } r, s, t \in R \\ r \cdot (s + t) &= r \cdot s + r \cdot t && \text{für alle } r, s, t \in R \end{aligned}$$

Ist $(R, \cdot, 1)$ ein kommutatives Monoid, so nennen wir $(R, +, \cdot)$ einen kommutativen Ring.

Bemerkung 2.3.1. Wenn die Operationen klar sind, werden wir kurz R für $(R, +, \cdot)$ schreiben.

Bemerkung 2.3.2. Wir setzen nicht voraus, dass $(R, \cdot, 1)$ kommutativ ist, da wir mit den Matrizen auch nicht-kommutative Ringe kennenlernen werden.

Beispiel 2.3.1. $(\mathbb{Z}, +, \cdot)$, $(\mathbb{R}, +, \cdot)$ und $(\mathbb{Q}, +, \cdot)$ sind kommutative Ringe.

Beispiel 2.3.2. Es sei $n > 1$ eine ganze Zahl $\mathbb{Z}/(n)$ die Menge der Äquivalenzklassen der Relation "unterscheiden sich um ein Vielfaches von n " aus Beispiel 2.1.12 in Abschnitt 2.1, also

$$\mathbb{Z}/(n) = \{[0], [1], \dots, [n-1]\}.$$

Wir definieren auf $\mathbb{Z}/(n)$ Verknüpfungen $+$ und \cdot durch

$$\begin{aligned} [r] + [s] &:= [r + s] && \text{für alle } [r], [s] \in \mathbb{Z}/(n) \\ [r] \cdot [s] &:= [r \cdot s] && \text{für alle } [r], [s] \in \mathbb{Z}/(n) \end{aligned}$$

Die Operation "+" haben wir ja schon in Beispiel 2.2.34 betrachtet.

Zunächst müssen wir uns überzeugen, dass diese Operationen sinnvoll und wohldefiniert sind. Wir haben also zu zeigen:

Sind $r' \sim_R r$ und $s' \sim_R s$ so gilt

$$[r' + s'] = [r + s], \quad [r' \cdot s'] = [r \cdot s]$$

Das ist aber leicht einzusehen. Ist etwa $r' = r + an$ und $s' = s + bn$, so gilt

$$r' \cdot s' = r \cdot s + (as + rb + abn) \cdot n$$

also unterscheidet sich $r' \cdot s'$ nur um ein Vielfaches von n von $r \cdot s$. Der Nachweis für die Summation ist noch einfacher.

Mit diesen Operationen gilt:

$(\mathbb{Z}/(n), +, \cdot)$ ist ein kommutativer Ring mit Nullelement $[0]$ und Einselement $[1]$. Die Ringeigenschaften ergeben sich dabei leicht aus den entsprechenden Eigenschaften von $(\mathbb{Z}, +, \cdot)$, siehe etwa Aufgabe 2.4.2.

Bemerkung 2.3.3. Die Ringoperationen in $\mathbb{Z}/(n)$ können auch direkt mit Hilfe des Repräsentantensystems $0, 1, \dots, n-1$ erklärt werden: Sind $r, s \in$

$\{0, 1, \dots, n-1\}$, so können wir Division durch n mit Rest anwenden und erhalten

$$\begin{aligned} r + s &= a \cdot n + d \\ r \cdot s &= \alpha \cdot n + \delta \end{aligned}$$

für geeignete $a, \alpha \in \mathbb{Z}$ und $d, \delta \in \{0, 1, \dots, n-1\}$, und es gilt

$$\begin{aligned} [r] + [s] &= [d] \\ [r] \cdot [s] &= [\delta] \end{aligned}$$

Das ist unmittelbar klar nach Definition, da ja $[a \cdot n] = [\alpha \cdot n] = [0]$.

Beispiel 2.3.3. Es sei jetzt speziell $n = 4$. Dann gilt in $\mathbb{Z}/(4)$:

$$[2] \cdot [2] = [4] = [0]$$

Wir haben also zwei von Null verschiedene Elemente in $\mathbb{Z}/(4)$ deren Produkt $[0]$ ergibt. Diese Situation tritt in \mathbb{Z} oder \mathbb{R} nicht auf.

Beispiel 2.3.4. Ist R ein kommutativer Ring, so betrachten wir die Menge

$$R[X] := \{r_n X^n + r_{n-1} X^{n-1} + \dots + r_1 X + r_0 \mid n \in \mathbb{N}, r_0, \dots, r_n \in R\}$$

aller formalen Linearkombinationen von $1, X, X^2, \dots$ über R . Sind

$$f(X) = r_n X^n + r_{n-1} X^{n-1} + \dots + r_1 X + r_0, \quad g(X) = s_m X^m + s_{m-1} X^{m-1} + \dots + s_1 X + s_0$$

So definieren wir Summe und Produkt von $f(X)$ und $g(X)$ wie folgt:

Falls $n \geq m$, so setzen wir

$$f(X) + g(X) = t_n X^n + t_{n-1} X^{n-1} + \dots + t_1 X + t_0$$

wobei

$$t_i = \begin{cases} r_i & \text{für } i > m \\ r_i + s_i & \text{für } i \leq m \end{cases}$$

und falls $n < m$, so setzen wir

$$f(X) + g(X) = t_m X^m + t_{m-1} X^{m-1} + \dots + t_1 X + t_0$$

wobei

$$t_i = \begin{cases} s_i & \text{für } i > n \\ r_i + s_i & \text{für } i \leq n \end{cases}$$

Mit $l = n \cdot m$ definieren wir

$$f(X) \cdot g(X) = t_l X^l + t_{l-1} X^{l-1} + \cdots + t_1 X + t_0$$

wobei

$$t_i = \sum_{k=0}^i r_k \cdot s_{i-k}$$

(mit $r_k = 0$ für $k > n$ und $s_k = 0$ für $k > m$).

Dann ist $R[X]$, zusammen mit dieser Addition und Multiplikation ein kommutativer Ring. Das Einselement ist das Element $1(X) = 1$ und das Nullelement ist das Element $0(X) = 0$. Wir nennen $R[X]$ den **Polynomring** (in der Variable X) über R und seine Elemente $f(X)$ **Polynome** mit Koeffizienten aus R .

Ist $f(X) = r_n X^n + r_{n-1} X^{n-1} + \cdots + r_1 X + r_0$ ein Polynom mit $r_n \neq 0$, so nennen wir n den **Grad** des Polynoms $f(X)$ und wir nennen $f(X)$ ein Polynom vom Grad n . Wir schreiben auch $\deg(f(X))$ für den Grad des Polynoms $f(X)$.

Jedes vom Nullpolynom verschiedene Polynom hat also einen eindeutigen Grad. Das Nullpolynom bildet eine Ausnahme, denn ihm kann kein Grad in vernünftiger Weise zugeordnet werden. Daher lässt man für das Nullpolynom jeden Grad zu, und damit ist es das einzige Polynom, das mehr als einen Grad hat.

Aufgabe 2.3.1. Zeigen Sie, dass $R[X]$ ein Ring ist.

Aufgabe 2.3.2. Die Polynome $f(X) \in R[X]$ vom Grad 0 entsprechen genau den Elementen von R .

Definition 2.3.2. Es sei $(R, +, \cdot)$ ein kommutativer Ring. Ein Element $r \in R \setminus \{0\}$ heißt **Nullteiler** von R wenn es ein $s \in R \setminus \{0\}$ gibt mit $r \cdot s = 0$. Ein kommutativer Ring, in dem es keine Nullteiler gibt, heißt **nullteilerfrei** oder **Integritätsbereich**.

Beispiel 2.3.5. Die Ringe \mathbb{R} , \mathbb{Q} und \mathbb{Z} sind nullteilerfrei.

Beispiel 2.3.6. Die Ringe $\mathbb{Z}/(2)$ und $\mathbb{Z}/(3)$ sind nullteilerfrei. Das sieht man sofort indem man die wenigen Elemente dieser Ringe einzeln durchgeht.

Beispiel 2.3.7. Die Ringe $\mathbb{Z}/(4)$, $\mathbb{Z}/(6)$ und $\mathbb{Z}/(15)$ sind nicht nullteilerfrei.

Aufgabe 2.3.3. Wir betrachten einen nullteilerfreien Ring R . Zeigen Sie

- a) Der Polynomring $R[X]$ ist nullteilerfrei.
- b) Sind $f(X), g(X) \in R[X]$ Polynome, die verschieden vom Nullpolynom sind, so gilt $\deg(f(X) \cdot g(X)) = \deg(f(X)) + \deg(g(X))$.

Nullteilerfrei Ringe haben eine für uns interessante Kürzungseigenschaft:

Satz 2.3.1. *Ist R ein nullteilerfreier Ring und $a \in R \setminus \{0\}$, so gilt für alle $x, y \in R$.*

$$ax = ay \implies x = y$$

Beweis: Wir haben die Beziehungskette

$$ax = ay \implies ax - ay = 0 \implies a \cdot (x - y) = 0 \longrightarrow x - y = 0$$

wobei wir für die letzte Implikation die Nullteilerfreiheit von R ausgenutzt haben (und die Tatsache, dass $a \neq 0$). Aber $x - y = 0$ ist äquivalent zu $x = y$ und es folgt die Behauptung.

Beispiel 2.3.8. In $\mathbb{Z}/(8)$ gilt

$$[2] \cdot [7] = [2] \cdot [3]$$

aber es gilt

$$[7] \neq [3]$$

Die Nullteilerfreiheit ist also wichtig für die Kürzungsregel.

2.4 Körper

Die Ringe \mathbb{Z} und $\mathbb{Z}/(n)$ spielen eine wichtige Rolle in der Informatik, und die Frage nach ihrer Nullteilerfreiheit bzw. ihren Nullteilern ist dabei von entscheidender Bedeutung.

Wir haben also jetzt unter den kommutativen Ringen eine Klasse mit einer besonderen Eigenschaft, nämlich der Nullteilerfreiheit, isoliert und gesehen,

dass nicht alle Ringe diese Eigenschaft haben. Einige Beispiele nullteilerfreier Ringe sind \mathbb{F}_p für Primzahlen p , \mathbb{Z} , \mathbb{Q} und \mathbb{R} . Trotzdem gibt es noch einen entscheidenden Unterschied zwischen \mathbb{Z} und \mathbb{Q} oder \mathbb{R} , nämlich die Tatsache, dass wir in \mathbb{Q} oder \mathbb{R} multiplikative Inverse bilden können. Das führt uns zur letzten Definition in diesem Abschnitt

Definition 2.4.1. Ein **Körper** $(K, +, \cdot)$ ist eine nichtleere Menge K mit zwei Verknüpfungen $+$ und \cdot , die zwei Elemente $0, 1$ enthält, wobei $0 \neq 1$ ist und wobei gilt

- $(K, +, 0)$ ist eine kommutative Gruppe.
- $(K \setminus \{0\}, \cdot, 1)$ ist eine kommutative Gruppe.
- Es gilt das Distributivgesetz

$$(r + s) \cdot t = r \cdot t + s \cdot t \quad \text{für alle } r, s, t \in R$$

Bemerkung 2.4.1. Da wir voraussetzen, dass $(R \setminus \{0\}, \cdot, 1)$ kommutativ ist, brauchen wir nur ein Distributivgesetz zu fordern. Das zweite Distributivgesetz ($t \cdot (r + s) = tr + ts$) folgt sofort aus der Kommutativität.

Bemerkung 2.4.2. Jeder Körper ist ein nullteilerfreier kommutativer Ring. Die Umkehrung gilt nicht.

Bemerkung 2.4.3. Schreiben wir alle Bedingungen aus, so sehen wir, dass $(K, +, \cdot)$ genau dann ein Körper ist, wenn die Gesetze $A1 - A4$, $M1 - M4$ und D , die wir bei der Einführung der reellen Zahlen notiert haben, für K gelten.

Beispiel 2.4.1. \mathbb{Q} und \mathbb{R} sind Körper.

Beispiel 2.4.2. \mathbb{Z} ist kein Körper.

Bemerkung 2.4.4. Ist $(R, +, \cdot)$ ein kommutativer Ring, so heißt ein $r \in R$ **Einheit**, wenn es ein $s \in R$ gibt mit $r \cdot s = 1$. Die Einheiten von R bilden eine Gruppe (bezüglich \cdot), die **Einheitengruppe** $E(R)$ von R .

Die Unterscheidung zwischen Körpern und nullteilerfreien Ringen wird durch den folgenden Satz gegeben, der trivialerweise aus der Definition folgt:

Satz 2.4.1. *Genau dann ist ein nullteilerfreier Ring R ein Körper wenn jedes Element $x \in R \setminus \{0\}$ ein multiplikatives Inverses hat, wenn also $E(R) = R \setminus \{0\}$.*

Ist K ein Körper, so ist der Polynomring $K[X]$ kein Körper, da das Element $X \in K[X]$ kein Inverses besitzt. Trotzdem ist der Polynomring von großem Interesse für das Studium von Körpern. Eine wichtige Eigenschaft von Polynomen ist nämlich, dass wir Elemente aus K einsetzen können: Ist

$$f(X) = r_n X^n + r_{n-1} X^{n-1} + \dots + r_1 X + r_0 \in K[X]$$

ein Polynom, und ist $a \in K$, so ist

$$f(a) := r_n a^n + r_{n-1} a^{n-1} + \dots + r_1 a + r_0$$

eine wohldefiniertes Element von K , so dass also $f(X)$ eine Abbildung

$$f : K \longrightarrow K, \quad a \mapsto f(a)$$

definiert.

Definition 2.4.2. Ein $a \in K$ heißt Nullstelle eines Polynoms $f(X) \in K[X]$, wenn $f(a) = 0$

Eine wichtige Eigenschaft von Polynomringen über Körpern ist

Satz 2.4.2. *Ist K ein Körper und $f(X)$ ein vom Nullpolynom verschiedenes Polynom vom Grad n , so hat $f(X)$ höchstens n Nullstellen.*

Beweis: Wir beweisen die Aussage durch Induktion nach n .

Die Aussage ist klar für ein Polynom vom Grad 0. Da das Polynom $f(X)$ nach Voraussetzung nicht das Nullpolynom ist, ist $f(X) = r$ mit einem Element $r \in K \setminus \{0\}$. Hierfür gilt offensichtlich

$$f(a) = r \neq 0$$

für jedes $a \in K$, so dass also $f(X)$ hier 0 Nullstellen hat.

Der Induktionsschritt beruht darauf, dass wir Polynomdivision über beliebigen Körpern durchführen können. Sind also $f(X)$, $g(X)$ zwei Polynome mit Koeffizienten aus K , und hat $g(X)$ den Grad n , so können wir immer schreiben

$$f(X) = q(X) \cdot g(X) + r(X)$$

mit einem Polynom $r(X)$ mit $\deg(r(X)) < n$ (also $f(X) = g(X) = q(X)$ Rest $r(X)$). Das kann man genauso einsehen wie die Polynomdivision über \mathbb{Q} oder \mathbb{R} .

Ist also jetzt $f(X)$ ein vom Nullpolynom verschiedenes Polynom vom Grad n und a eine Nullstelle von $f(X)$, so können wir schreiben

$$f(X) = q(X) \cdot (X - a) + r(X)$$

wobei $r(X) = r$ ein Polynom vom Grad 0, also eine Element aus K ist. Setzen wir a auf beiden Seiten ein, so erhalten wir $f(a) = 0$ (da ja a eine Nullstelle von $f(X)$ und $q(a) \cdot (a - a) + r = r$, und damit

$$0 = r$$

Das bedeutet $f(X) = g(X) \cdot (X - a)$. Notwendigerweise ist $g(X)$ nicht das Nullpolynom, und aus Aufgabe 2.3.3 folgt, dass $\deg(q(X)) = n - 1$. Daher wissen wir gemäß Induktionsvoraussetzung bereits, dass $q(X)$ höchstens $n - 1$ Nullstellen hat, und hieraus folgt leicht, dass $f(X)$ höchstens n Nullstellen hat.

Die Körper \mathbb{R} und \mathbb{Q} sind aus der Schule bekannt, im Abschnitt 2.5 werden wir auch noch einige Körper mit nur endliche vielen Elementen kennenlernen. Ein weiterer wichtiger Körper ist der Körper \mathbb{C} der komplexen Zahlen. Wie die reellen Zahlen als Erweiterung der rationalen Zahlen konstruiert werden, um Wurzeln aus (positiven) Zahlen zu ziehen und besondere Zahlen wie e oder π definieren zu können, werden die komplexen Zahlen als Erweiterung der reellen Zahlen betrachtet um einen Mangel der reellen Zahlen zu beheben, nämlich die Tatsache, dass in \mathbb{R} nicht aus allen Zahlen eine Quadratwurzel gezogen werden kann. Die Konstruktion dieses Körpers ist sehr explizit:

Wir definieren \mathbb{C} als die Menge alle Zahlenpaare (x, y)

in \mathbb{R}^2 und mit der folgenden Addition und Multiplikation.

- $(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2)$.
- $(x_1, y_1) \cdot (x_2, y_2) = (x_1 \cdot x_2 - y_1 \cdot y_2, x_1 \cdot y_2 + x_2 \cdot y_1)$.

Zur Vereinfachung der Schreibweise führen wir die folgenden Notationen ein:

$$0 = (0, 0) \text{ (das Nullelement).}$$

$$1 = (1, 0) \text{ (das Einselement).}$$

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

$$(x, y) = x + i \cdot y.$$

Für eine komplexe Zahl $z = x + i \cdot y$ heißt dann x der **Realteil** und y der **Imaginärteil** von z .

Addition und Multiplikation der komplexen Zahlen schreiben sich damit wie folgt

- $(x_1 + i \cdot y_1) + (x_2 + i \cdot y_2) = (x_1 + x_2 + i \cdot (y_1 + y_2))$.
- $(x_1 + i \cdot y_1) \cdot (x_2 + i \cdot y_2) = x_1 \cdot x_2 - y_1 \cdot y_2 + i \cdot (x_1 \cdot y_2 + x_2 \cdot y_1)$.

Satz 2.4.3 (Der Körper der komplexen Zahlen). *Die komplexen Zahlen \mathbb{C} zusammen mit dieser Addition und Multiplikation bilden einen Körper.*

Das inverse Element $\frac{1}{z}$ zu einer komplexen Zahl $z = x + i \cdot y \neq 0$ ist gegeben durch

$$\frac{1}{z} = \frac{x - i \cdot y}{x^2 + y^2}$$

das Nullelement bezüglich der Addition ist $0 = 0 + i \cdot 0$, das Einselement bezüglich der Multiplikation ist $1 = 1 + i \cdot 0$.

Beweis: Die Körperaxiome können direkt nachgerechnet werden.

Bemerkung 2.4.5. Sofort aus der Definition der Multiplikation erhalten wir

$$i^2 = (0 + i \cdot 1) \cdot (0 + i \cdot 1) = 0 \cdot 0 - 1 \cdot 1 + i \cdot (0 \cdot 1 + 1 \cdot 0) = -1$$

also ist i eine Quadratwurzel aus -1 .

Allgemeiner gilt für jede positive Zahl $r \in \mathbb{R}$ mit (positiver) Quadratwurzel \sqrt{r} :

$$(i \cdot \sqrt{r})^2 = i^2 \cdot \sqrt{r}^2 = (-1) \cdot r = -r$$

dh. $i \cdot \sqrt{r}$ ist eine Quadratwurzel von $-r$, und damit haben in \mathbb{C} auch alle negativen Zahlen eine Quadratwurzel

Es gilt sogar allgemeiner

Satz 2.4.4 (Fundamentalsatz der Algebra). *Der Körper \mathbb{C} ist algebraisch abgeschlossen, dh. jedes Polynom*

$$f(x) = a_0 + a_1 \cdot x + \cdots + a_n \cdot x^n \in \mathbb{C}[X]$$

vom Grad ≥ 1 hat eine Wurzel, d.h. es gibt ein $z_0 \in \mathbb{C}$ mit $f(z_0) = 0$.

Aufgabe 2.4.1. Zeigen Sie, dass $(\mathbb{Z}/(n), +)$ mit den in Beispiel 2.3.2 definierten Operationen tatsächlich eine kommutative Gruppe ist.

Aufgabe 2.4.2. Zeigen Sie, dass $(\mathbb{Z}/(n), +, \cdot)$ mit den in Beispiel 2.3.2 definierten Operationen tatsächlich ein Ring ist.

Aufgabe 2.4.3. In einem Spiel für zwei Spieler werden zu Beginn n Spielsteine auf einem Spielplan aufgestellt. Jeder Spieler kann in seinem Zug 1 - 4 Steine vom Spielplan entfernen. Gewonnen hat der Spieler, der den letzten Stein vom Brett nimmt. Zeigen Sie: Ist n durch 5 teilbar, so kann der nachziehende Spieler den Sieg erzwingen. In allen anderen Situationen gibt es für den anziehenden Spieler einen Gewinnweg.

Aufgabe 2.4.4. Wir betrachten die Menge $G = \{a, b, c, d\}$ mit der Verknüpfung

\circ	a	b	c	d
a	a	b	c	d
b	b	a	d	c
c	c	d	a	b
d	d	c	b	a

Zeigen Sie, dass G eine kommutative Gruppe der Ordnung 4 ist. Was ist das neutrale Element von G ? Was ist der wesentliche Unterschied zwischen dieser Gruppe und der Gruppe $\mathbb{Z}/4\mathbb{Z}$?

Aufgabe 2.4.5. Zeigen Sie: Ist G eine endliche zyklische Gruppe, so ist G kommutativ.

Aufgabe 2.4.6. Schreiben Sie die Permutation

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 5 & 1 & 4 \end{pmatrix}$$

als Komposition von Vertauschungen $\langle i j \rangle$.

Aufgabe 2.4.7. Es sei G die Permutationsgruppe S_n . Definiert man $\text{sign}(\sigma)$ als

$$\text{sign}(\sigma) = \begin{cases} +1 & \text{falls die Anzahl der Fehlstände gerade ist} \\ -1 & \text{falls die Anzahl der Fehlstände ungerade ist} \end{cases}$$

so gilt:

$$\text{sign}(\sigma) = \prod_{i < j} \frac{\sigma(j) - \sigma(i)}{j - i}$$

Aufgabe 2.4.8. Es sei G die Permutationsgruppe S_n . Dann gilt:

$$\text{sign}(\sigma \circ \tau) = \text{sign}(\sigma) \circ \text{sign}(\tau)$$

Aufgabe 2.4.9. Bestimmen Sie alle Nullteiler des Rings $\mathbb{Z}/(24)$.

Aufgabe 2.4.10. Zeigen Sie, dass $\mathbb{Z}/5\mathbb{Z}$ ein Körper ist.

2.5 Ganze Zahlen und ihre Restklassen

Die ganzen Zahlen und ihre Restklassenringe sind von fundamentaler Bedeutung für die Mathematik, aber auch weit darüberhinaus, etwa in der Informatik. Daher wollen wir diesen Ringen einen eigenen Abschnitt widmen. Ein fundamentaler Begriff in diesem Zusammenhang ist der Begriff der Teilbarkeit, und wesentlich dabei wiederum ist der Begriff der Primzahl:

Definition 2.5.1. Eine Zahl $p \in \mathbb{N} \setminus \{0, 1\}$ heißt **Primzahl**, wenn p keine echten Teiler hat.

Beispiel 2.5.1. 2, 3, 5, 7 und 11 sind Primzahlen, 9 ist keine.

Wir notieren hier zwei wichtige Eigenschaften von Primzahlen und ganzen Zahlen, deren Nachweis sich in jedem Buch zur Algebra oder zur elementaren Zahlentheorie findet (etwa in *Gerhard Frey: Elementare Zahlentheorie*; Vieweg Verlag 1984, und dort speziell die Aussagen 2.2, 3.1, 3.2 und 3.4)).

Bemerkung 2.5.1. Eine Zahl $p \in \mathbb{N} \setminus \{0, 1\}$ ist genau dann eine Primzahl, wenn für alle ganzen Zahlen $a, b \in \mathbb{Z}$ gilt

$$p \mid a \cdot b \iff p \mid a \text{ oder } p \mid b$$

also p teilt genau dann ein Produkt von zwei Zahlen, wenn es schon eine der beiden Zahlen teilt.

Bemerkung 2.5.2. Jede ganze Zahl $z \in \mathbb{Z} \setminus \{0\}$ schreibt sich in eindeutiger Weise als

$$z = \varepsilon \cdot p_1^{n_1} \cdot p_2^{n_2} \cdots p_t^{n_t}$$

mit $\varepsilon \in \{-1, 1\}$, Primzahlen $p_1 < p_2 < \dots < p_t$ und positiven natürlichen Zahlen n_1, \dots, n_t . (Dabei lassen wir den Spezialfall $t = 0$ für $z = \pm 1$ zu).

Definition 2.5.2. Sind $m, n \in \mathbb{N} \setminus \{0\}$, so heißt eine Zahl $g \in \mathbb{N}$ der **größte gemeinsame Teiler** von m und n , wenn gilt:

- g ist ein Teiler von m und ein Teiler von n .
- Ist h ein weiterer Teiler von m und n , so ist h auch ein Teiler von g .

Wir schreiben

$$\text{ggT}(m, n) := g$$

für den größten gemeinsame Teiler von m und n .

Zwei Zahlen $m, n \in \mathbb{N} \setminus \{0\}$ heißen **teilerfremd**, wenn $\text{ggT}(m, n) = 1$, wenn sie also keinen echten gemeinsamen Teiler besitzen.

Bemerkung 2.5.3. Ist die Primfaktorzerlegung von m und n bekannt, so lässt sich daraus der größte gemeinsame Teiler von m und n sofort ablesen. Sind nämlich p_1, \dots, p_t die Primzahlen, die entweder m oder n teilen, und schreiben wir

$$m = p_1^{a_1} \cdots p_t^{a_t}, \quad n = p_1^{b_1} \cdots p_t^{b_t}$$

mit $b_i, a_i \geq 0$, so gilt

$$\text{ggT}(m, n) = p_1^{\min\{a_1, b_1\}} \cdots p_t^{\min\{a_t, b_t\}}$$

dh. im größten gemeinsamen Teiler von m und n kommen alle Primzahlen vor, die sowohl m als auch n teilen, und zwar so oft, wie sie mindestens beide teilen.

Eine unmittelbare Konsequenz der Definition ist

Satz 2.5.1. *Ist p eine Primzahl und $m \in \mathbb{N} \setminus \{0\}$ kein Vielfaches von p , so sind m und p schon teilerfremd.*

Bemerkung 2.5.4. Wir hätten die Begriffe *Teiler* und *Primzahlen* auch für ganze Zahlen definieren können. In diesem Fall wäre etwa m ein echter Teiler von n wenn es ein $a \in \mathbb{Z} \setminus \{\pm 1, \pm n\}$ gibt mit $a \cdot m = n$. Das ändert wenig an den grundsätzlichen Eigenschaften, es sind lediglich Vorzeichen zu berücksichtigen. So ist etwa mit p auch $-p$ eine Primzahl, und mit m ist auch $-m$ ein Teiler von n . Wir erhalten aber weder grundsätzlich neue Teiler noch einen grundsätzlich anderen Begriff von Primzahlen. So ist etwa p genau dann eine Primzahl in \mathbb{Z} wenn entweder p oder $-p$ eine Primzahl im Sinne unserer Definition ist. Für uns sollen aber etwa 2 und -2 - lax gesprochen - die gleiche Primzahl darstellen, während 2 und 3 für uns grundsätzlich verschiedene Primzahlen sind.

Um daher lästige Vorzeichenbetrachtungen zu vermeiden, verzichten wir auf diesen augenscheinlich allgemeineren Begriff.

Der größte gemeinsame Teiler zweier natürlicher Zahlen $m, n \in \mathbb{N} \setminus \{0\}$ lässt sich mit Hilfe des **euklidischen Algorithmus** schnell berechnen.

Der euklidische Algorithmus:

- **Vorbereitungsschritt:** Wir ordnen m und n so an, dass $m \geq n$. Gegebenenfalls vertauschen wir dazu die Rollen von m und n , denn $\text{ggT}(m, n) = \text{ggT}(n, m)$. Wir setzen $i = 0$ und $r_0 = m$, $r_1 = n$.
- **Verarbeitungsschritt:** Wir dividieren r_i durch r_{i+1} mit Rest:

$$r_i = q \cdot r_{i+1} + b$$

mit einer natürlichen Zahl q und einem Rest $b \in \{0, 1, \dots, r_{i+1} - 1\}$.

- Falls $b = 0$ (dh. die Division geht ohne Rest auf) \rightarrow **STOPP**.
- Falls $b \neq 0$ setze $r_{i+2} = b$ und $i = i + 1$. Wiederhole den Verarbeitungsschritt.
- **Ergebnisschritt:** Nach endlich vielen Verarbeitungsschritten (höchstens m vielen) geht die Division erstmals ohne Rest auf, d.h.

$$r_i = q \cdot r_{i+1} + 0$$

mit $r_{i+1} \neq 0$, Das STOPP-Kriterium wird also immer erreicht. Dann ist r_{i+1} der größte gemeinsame Teiler von m und n , $r_{i+1} = \text{ggT}(m, n)$.

Der Nachweis, dass dieser Algorithmus tatsächlich zum größten gemeinsamen Teiler führt, kann etwa mit Aufgabe 2.5.2 geführt werden.

Beispiel 2.5.2. Wir betrachten die Zahlen $m = 222$ und $n = 156$. Hier gilt bereits $m \geq n$, und wir setzen $r_0 = 222$ und $r_1 = 156$.

$$i = 0: \quad 222 = 1 \cdot 156 + 66. \text{ Wir setzen } r_2 = 66.$$

$$i = 1: \quad 156 = 2 \cdot 66 + 24. \text{ Wir setzen } r_3 = 24.$$

$$i = 2: \quad 66 = 2 \cdot 24 + 18. \text{ Wir setzen } r_4 = 18.$$

$i = 3$: $24 = 1 \cdot 18 + 6$. Wir setzen $r_5 = 6$.

$i = 4$: $18 = 3 \cdot 6 + 0$. \longrightarrow **STOPP**.

Ergebnis: $\text{ggT}(222, 156) = 6$.

Beispiel 2.5.3. Wir betrachten die Zahlen $m = 19$ und $n = 234$. Hier müssen wir zuerst die Rollen von m und n vertauschen, setzen also $r_0 = 234$ und $r_1 = 19$.

$i = 0$: $234 = 12 \cdot 19 + 6$. Wir setzen $r_2 = 6$.

$i = 1$: $19 = 3 \cdot 6 + 1$. Wir setzen $r_3 = 1$.

$i = 2$: $6 = 6 \cdot 1 + 0$. \longrightarrow **STOPP**.

Ergebnis: $\text{ggT}(234, 19) = 1$. Die Zahlen 234 und 19 sind also teilerfremd.

Bemerkung 2.5.5. Der euklidische Algorithmus liefert ein effizientes Verfahren zur Berechnung des größten gemeinsamen Teilers, das für große Zahlen sehr viel schneller ist als der Weg über die Primfaktorzerlegung und die gemeinsamen Primfaktoren.

Zunächst scheint es ja ein Problem zu sein, dass wir nicht wissen, wie viele Schritte der Algorithmus braucht, dh. wie viele r_i berechnet werden müssen (und damit auch, wie viel Speicher bereitgestellt werden soll). Allerdings sieht man sofort, dass in jedem Verarbeitungsschritt aus dem vorherigen Teil der Verarbeitung nur die beiden Zahlen, die mit Rest durcheinander dividiert werden müssen, benötigt werden, daher brauchen wir immer nur diese beiden Zahlen (und den neuen Rest). Damit kann die Bestimmung des größten gemeinsamen Teilers effizient umgesetzt werden:

```
function [g] = la_euklid01(m,n)
% Funktionsbaustein zur Berechnung des
% groessten gemeinsamen Teilers von m
% und n mit dem euklidischen Algorithmus

% Annahme: m >= n > 0
```

```

%%% Initialisierung

r0 = m;
r1 = n;
% berechne den Rest der ganzzahligen Division
r = mod(r0,r1);
% solange dieser Rest nicht verschwindet, fuehre weitere
% Division mit Rest durch
while r ~= 0
    r0 = r1;
    r1 = r;
    r = mod(r0,r1);
end
% der ggT ist der letzte nichtverschwindende Rest
g = r1;
end

```

Eine interessante und wichtige Folgerung aus dem euklidischen Algorithmus ist

Satz 2.5.2. *Sind $m, n \in \mathbb{N} \setminus \{0\}$ mit $\text{ggT}(m, n) = g$, so gibt es ganze Zahlen a, b mit*

$$a \cdot m + b \cdot n = g$$

Der Beweis des Satzes besteht darin, dass wir aus dem vorletzten Schritt des euklidischen Algorithmus rückwärtsrechnen, also die Beziehung $r_{i-1} = q \cdot r_i + r_{i+1}$, in der r_{i+1} der größte gemeinsame Teiler von m und n ist, nach r_{i+1} auflösen, $r_{i+1} = r_{i-1} - q \cdot r_i$ und dann sukzessive die verschiedenen Schritte des Algorithmus zurückgehen. Für die Details siehe Aufgabe 2.5.4

Beispiel 2.5.4. In Beispiel 2.5.2 haben wir gesehen, dass $\text{ggT}(234, 156) = 6$. Um 6 mit 234 und 156 zu beschreiben, gehen wir vor wie folgt:
Aus dem Schritt $i = 3$ erhalten wir

$$6 = 24 - 1 \cdot 18$$

Aus $i = 2$ folgt $18 = 66 - 2 \cdot 24$. Setzen wir das ein, so ergibt sich

$$6 = 24 - 1 \cdot 18 = 24 - 1 \cdot (66 - 2 \cdot 24) = 3 \cdot 24 - 1 \cdot 66$$

Aus $i = 1$ erhalten wir $24 = 156 - 2 \cdot 66$. Setzen wir das ein, so ergibt sich

$$6 = 3 \cdot 24 - 1 \cdot 66 = 3 \cdot (156 - 2 \cdot 66) - 1 \cdot 66 = 3 \cdot 156 - 7 \cdot 66$$

Aus $i = 0$ folgt schließlich $66 = 222 - 1 \cdot 156$. Setzen wir das ein, so ergibt sich

$$6 = 3 \cdot 156 - 7 \cdot 66 = 3 \cdot 156 - 7 \cdot (222 - 156) = 10 \cdot 156 - 7 \cdot 222$$

Damit haben wir eine Darstellung

$$6 = 10 \cdot 156 - 7 \cdot 222$$

wie gewünscht gefunden.

Beispiel 2.5.5. In Beispiel 2.5.3 haben wir gesehen, dass $\text{ggT}(234, 19) = 1$. Wir wollen nun 1 mit 19 und 234 darstellen.

1. Aus Schritt $i = 1$ erhalte: $1 = 19 - 3 \cdot 6$.
2. Aus Schritt $i = 0$ erhalte:

$$1 = 19 - 3 \cdot 6 = 19 - 3 \cdot (234 - 12 \cdot 19) = 37 \cdot 19 - 3 \cdot 234$$

Bemerkung 2.5.6. Die Bestimmung der beiden Zahlen a und b , für die gilt

$$a \cdot m + b \cdot n = \text{ggT}(m, n)$$

ist sehr wichtig für Anwendungen in der Kryptographie oder der Codierungstheorie. Einer effizienten Umsetzung des oben beschriebenen Vorgehens scheint die Rückwärtsrechnung, die dabei durchgeführt wurde, entgegenzustehen, denn diese kann offensichtlich nicht effizient umgesetzt werden. Bei genauerer Betrachtung sehen wir, dass dieser Rückwärtsrechenschritt allerdings vermieden werden kann und alle notwendige Information beim Vorwärtsrechnen im Rahmen des eigentlichen euklidischen Algorithmus mit übergeben werden kann. Das soll nun am Beispiel 2.5.2 gezeigt werden.

Beispiel 2.5.6. Wir greifen noch einmal Beispiel 2.5.2 mit $m = 222$ und $n = 156$ auf.

Vorbereitungsschritt: Es gilt $m \geq n$, also setzen wir

$$r_0 = 234 \quad \text{und} \quad r_1 = 156$$

Außerdem setzen wir $a_0 = 1$, $b_0 = 0$ und $a_1 = 0$, $b_1 = 1$. Dann gilt hierfür

$$\begin{aligned} r_0 &= a_0 \cdot m + b_0 \cdot n \\ r_1 &= a_1 \cdot m + b_1 \cdot n \end{aligned}$$

Verarbeitung:

$i = 0$: $222 = 1 \cdot 156 + 66$, also $q = 1$, $r = 66$. Wir setzen $r_2 = 66$ und wir beachten, dass

$$r_2 = r_0 - q \cdot r_1 = (a_0 \cdot m + b_0 \cdot n) - q \cdot (a_1 \cdot m + b_1 \cdot n) = (a_0 - q \cdot a_1) \cdot m + (b_0 - q \cdot b_1) \cdot n$$

Setzen wir also

$$a_2 = a_0 - q \cdot a_1 = 1, \quad b_2 = b_0 - q \cdot b_1 = -1$$

so erhalten wir

$$r_2 = a_2 \cdot m + b_2 \cdot n = 1 \cdot 222 + (-1) \cdot 156$$

$i = 1$: $156 = 2 \cdot 66 + 24$, also $q = 2$, $r = 24$. Wir setzen $r_3 = 24$ und wir beachten, dass

$$r_3 = r_1 - q \cdot r_2 = (a_1 \cdot m + b_1 \cdot n) - q \cdot (a_2 \cdot m + b_2 \cdot n) = (a_1 - q \cdot a_2) \cdot m + (b_1 - q \cdot b_2) \cdot n$$

Setzen wir also

$$a_3 = a_1 - q \cdot a_2 = -2, \quad b_3 = b_1 - q \cdot b_2 = 3$$

so erhalten wir

$$r_3 = a_3 \cdot m + b_3 \cdot n = (-2) \cdot 222 + 3 \cdot 156$$

$i = 2$: $66 = 2 \cdot 24 + 18$, also $q = 2$, $r = 18$. Wir setzen $r_4 = 18$ und (wie oben)

$$a_4 = a_2 - q \cdot a_3 = 5, \quad b_4 = b_2 - q \cdot b_3 = -7$$

Damit erhalten wir

$$r_4 = a_4 \cdot m + b_4 \cdot n = 5 \cdot 222 + (-7) \cdot 156$$

$i = 3$: $24 = 1 \cdot 18 + 6$, also $q = 1$, $r = 6$. Wir setzen $r_5 = 6$ und (wie oben)

$$a_5 = a_3 - q \cdot a_4 = -7, \quad b_5 = b_3 - q \cdot b_4 = 10$$

Damit erhalten wir

$$r_5 = a_5 \cdot m + b_5 \cdot n = (-7) \cdot 222 + 10 \cdot 156$$

$i = 4$: $18 = 3 \cdot 6 + 0$. \longrightarrow **STOPP**.

Ergebnis: $\text{ggT}(222, 156) = r_5 = 6$. und

$$6 = (-7) \cdot 222 + 10 \cdot 156$$

Bemerkung 2.5.7. Wie bei der Umsetzung des euklidischen Algorithmus kann auch hier die Abhängigkeit von dem Schrittzähler i noch eliminiert werden, so dass jeder Schritt für seine Verarbeitung nur die beiden Zahlen, die mit Rest durcheinander dividiert werden, und die Darstellung dieser beiden Zahlen mithilfe von m und n enthält. Daraus ergibt sich dann insgesamt der folgende Algorithmus

```
function [g,a,b] = la_euklid02(m,n)
% Funktionsbaustein zur Berechnung des
% groessten gemeinsamen Teilers von m und n
% mit dem euklidischen Algorithmus und einer
% Darstellung g = a*m + b*n mit ganzen Zahlen a,b,

% Annahme: m >= n > 0
```

```

%%% Initialisierung
r0 = m;
r1 = n;
%%% Initialisierung der Koeffizienten der Darstellung
a0 = 1; b0 = 0;
a1 = 0; b1 = 1;

% berechne das Ergebnis der ganzzahligen Division
r = mod(r0,r1);
q = (r0-r)/r1;      % Ergebnis ist ganzzahlig
% solange dieser Rest nicht verschwindet, fuehre weitere
% Division mit Rest durch
while r ~= 0
    % aktualisiere die Koeffizienten der Darstellung
    c0 = a0 - q*a1;      % zwischenspeichern
    c1 = b0 - q*b1;
    a0 = a1;
    b0 = b1;
    a1 = c0;
    b1 = c1;
    % fuehre neue Division mit Rest durch
    r0 = r1;
    r1 = r;
    r = mod(r0,r1);
    q = (r0-r)/r1;
end
% der ggT ist der letzte nichtverschwindende Rest
g = r1;
% Darstellung ist die Darstellung von r1
a = a1;
b = b1;
end

```

Nach diesen Überlegungen können wir wieder zur Betrachtung der Ringe $\mathbb{Z}/(n)$ zurückkehren:

Satz 2.5.3. *Für Zahlen $m, n \in \mathbb{N} \setminus \{0\}$, für die m kein Vielfaches von n ist, sind die folgenden Aussagen äquivalent:*

- i) $[m]$ ist eine Nullteiler in $\mathbb{Z}/(n)$.

ii) $\text{ggT}(m, n) \neq 1$.

Beweis: Ist $[m]$ ein Nullteiler von $\mathbb{Z}/(n)$, so gibt es ein $a \in \{1, \dots, n-1\}$ mit

$$[m] \cdot [a] = [0]$$

Das bedeutet aber, $m \cdot a$ ist ein Vielfaches von n , also $m \cdot a = l \cdot n$. Nehmen wir an, dass ii) nicht gilt, dass also m und n teilerfremd sind, so finden wir nach Satz 2.5.2 Zahlen r und s mit $r \cdot m + s \cdot n = 1$. Damit gilt aber

$$r \cdot l \cdot n = a \cdot r \cdot m = a \cdot (1 - s \cdot n) = a - a \cdot s \cdot n$$

also $a = n \cdot (r \cdot l + a \cdot s)$ ist ein Vielfaches von n , im Widerspruch zur Wahl von a .

Ist umgekehrt $g = \text{ggT}(m, n)$, so können wir $n = g \cdot n_1$ und $m = g \cdot m_1$ schreiben, wobei notwendig $n_1 \in \{2, \dots, n-1\}$. Dann gilt aber

$$m \cdot n_1 = g \cdot m_1 \cdot n_1 = g \cdot n_1 \cdot m_1 = n \cdot m_1$$

ist ein Vielfaches von n , also $[m] \cdot [n_1] = [0]$. Da $[n_1] \neq [0]$ ist $[m]$ ein Nullteiler.

Folgerung 2.5.4. *Genau dann ist $\mathbb{Z}/(n)$ nullteilerfrei, wenn n eine Primzahl ist.*

Beweis: Dazu reicht es nach Satz 2.5.3 zu zeigen, dass genau dann jede Zahl $m \in \{1, \dots, n-1\}$ teilerfremd zu n ist, wenn n eine Primzahl ist.

Ist dazu zunächst n eine Primzahl, so folgt aus Satz 2.5.1, dass m und n teilerfremd sind. Ist umgekehrt jedes $m \in \{1, \dots, n-1\}$ teilerfremd zu n , so kann n insbesondere keine echten Teiler haben und ist also eine Primzahl.

Bemerkung 2.5.8. Um die Notation zu vereinfachen, schreiben wir häufig für die Restklasse eine Zahl $m \in \mathbb{Z}$ in $\mathbb{Z}/(n)$ ebenfalls m anstelle von $[m]$. Aus dem Kontext wird immer klar sein, um was es sich handelt. Ferner haben sich für $\mathbb{Z}/(n)$ auch die Notationen $\mathbb{Z}/n\mathbb{Z}$ und \mathbb{Z}_n eingebürgert. Auch diese werden häufig verwendet.

Bezeichnung::

Ist p eine Primzahl, so schreiben wir auch \mathbb{F}_p für $\mathbb{Z}/(p)$.

In Abschnitt 2.3 haben wir neben dem Begriff des Rings auch den des Körpers studiert. Neben den bekannten Beispielen \mathbb{Q} und \mathbb{R} haben wir nun einen weiteren Körper kennengelernt ohne das bis jetzt zu wissen:

Satz 2.5.5. *Für jede Primzahl p ist \mathbb{F}_p ein Körper.*

Beweis: Nach Korollar 2.5.4 ist \mathbb{F}_p nullteilerfrei, und wegen Satz 2.4.1 brauchen wir daher nur noch zu zeigen, dass jedes $x \neq 0$ in \mathbb{F}_p ein multiplikatives Inverses hat. Sei dazu $x = [m]$ mit $m \in \mathbb{Z}$. Da $[m] \neq [0]$ sind m und p teilerfremd. Deshalb finden wir Zahlen r, s mit $r \cdot m + s \cdot p = 1$, und daraus folgt

$$[r][m] = [1 - s \cdot p] = [1] - [s] \cdot [p] = [1]$$

also das Gewünschte.

Beispiel 2.5.7. Die Ermittlung des inversen Elements zu einem Element $m \in \mathbb{F}_p$ (mit Repräsentanten $m \in \mathbb{Z}$) erfolgt am besten mithilfe des erweiterten euklidischen Algorithmus:

Wir betrachten den endlichen Körper $k = \mathbb{F}_{71}$ mit 71 Elementen und wollen die folgenden Elemente

$$a = \frac{1}{24}, \quad b = \frac{17}{30}$$

von k mit den Repräsentanten $0, 1, \dots, 70$ darstellen.

Zur Bestimmung von $\frac{1}{24}$ in \mathbb{F}_{71} gehen wir vor wie folgt:

$$71 = 2 \cdot 24 + 23$$

$$24 = 1 \cdot 23 + 1$$

$$23 = 23 \cdot 1 + 0$$

Damit ergibt Rückwärtsrechnen

$$1 = 24 - 1 \cdot 23$$

$$= 24 - 1 \cdot (71 - 2 \cdot 24)$$

$$= 3 \cdot 24 - 1 \cdot 71$$

und damit gilt in \mathbb{F}_{59} :

$$1 = 3 \cdot 24$$

also

$$\frac{1}{24} = 3$$

Zur Bestimmung von $b = \frac{17}{30}$. Berechnen wir zunächst $\frac{1}{30}$. Hier ist (wieder nach dem erweiterten euklidischen Algorithmus)

$$(-26) \cdot 30 + 11 \cdot 71 = 1$$

also in \mathbb{F}_{71} :

$$\frac{1}{30} = -26 = 45$$

und damit

$$\frac{17}{30} = 17 \cdot \frac{1}{30} = 17 \cdot 45 = 55$$

Diese Körper \mathbb{F}_p haben eine für die Anwendungen sehr interessante Eigenschaft.

Satz 2.5.6. (Kleiner Satz von Fermat) Für jedes $a \in \mathbb{F}_p \setminus \{[0]\}$ gilt

$$a^{p-1} = 1 \quad \text{in } \mathbb{F}_p$$

und für jedes $a \in \mathbb{F}_p$ gilt

$$a^p = a \quad \text{in } \mathbb{F}_p$$

Beweis: Wir wissen, dass $(\mathbb{F}_p \setminus \{[0]\}, \cdot)$ eine Gruppe ist, und zwar eine Gruppe der Ordnung $p - 1$. Damit folgt aus Lemma 2.2.2 schon $a^{p-1} = 1$. Damit gilt auch $a^p = a^{p-1} \cdot a = a$. Da sicherlich $[0]^p = [0]$, folgt auch der Zusatz.

Die additive Gruppe des Körpers \mathbb{F}_p ist $(\mathbb{Z}/p\mathbb{Z}, +)$, also eine sehr einfache zyklische Gruppe. Die Einheitengruppe $E(\mathbb{F}_p)$ hat ebenfalls eine sehr einfache Struktur

Satz 2.5.7. Für jede Primzahl p ist die Einheitengruppe $E(\mathbb{F}_p)$ von \mathbb{F}_p eine zyklische Gruppe.

Beweis: Im Abschnitt 2.3 haben wir schon den Begriff der Ordnung eines Elementes g einer endlichen Gruppe G kennengelernt

$$\text{ord}(g) = \min\{n \geq 1 \mid g^n = e\}$$

und aus Lemma 2.2.2 wissen wir schon, dass

$$\text{ord}(g) \mid \text{ord}(G)$$

Wenden wir das jetzt speziell auf die Einheitengruppe $E(\mathbb{F}_p)$ von \mathbb{F}_p an, so erhalten wir, dass für jedes $x \in E(\mathbb{F}_p)$ gilt

$$\text{ord}(x) \mid p - 1$$

Wenn wir nun ein $x \in E(\mathbb{F}_p)$ finden mit $\text{ord}(x) = p - 1$, so muss dieses x notwendigerweise ein Erzeuger der gesamten Einheitengruppe sein (und diese damit zyklisch), denn in Bemerkung 2.2.3 haben wir gesehen, dass $x^l \neq x^k$ für alle $1 \leq l < k \leq \text{ord}(x)$ und daher sind die Elemente der Menge

$$\langle x \rangle = \{x, x^2, x^3, \dots, x^{p-1}\}$$

alle paarweise verschieden, wenn $\text{ord}(x) = p - 1$, müssen also schon die ganze Gruppe $E(\mathbb{F}_p)$ sein.

Um ein solches Element zu finden, setzen wir

$$M := \max\{n \in \mathbb{N} \mid \text{es gibt ein } a \in E(\mathbb{F}_p) \text{ mit } \text{ord}(a) = n\}$$

und wir wählen ein $x \in E(\mathbb{F}_p)$ mit $\text{ord}(x) = M$.

Zunächst behaupten wir: Ist $y \in E(\mathbb{F}_p)$ beliebig, und ist $\text{ord}(y) = m$, so gilt $m \mid M$. Dabei gilt sicherlich $m \leq M$.

Nehmen wir an, dass $m \nmid M$, so gilt

$$g := \text{ggT}(m, M) < m$$

also $m = m' \cdot g$ mit $m' > 1$ und m' teilerfremd zu M . Setzen wir $y' = y^g$, so hat y' die Ordnung m' , denn

$$(y')^{m'} = y^{gm'} = y^m = 1$$

und für $n \geq 1$, $n < m'$ gilt

$$(y')^{n'} = y^{gn} \neq 1$$

da $gn < m$. Setzen wir nun $a = \frac{x}{y'}$ und $n = \text{ord}(a)$, so gilt $a^n = 1$, also

$$x^n = (y')^n$$

und damit insbesondere $\text{ord}(x^n) = \text{ord}((y')^n)$. Nun gilt aber

$$\text{ord}((y')^n) \mid m', \quad \text{ord}(x^n) \mid M$$

Da m' und M teilerfremd sind, muss notwendig

$$\text{ord}(x^n) = \text{ord}((y')^n) = 1$$

gelten, also $x^n = (y')^n = 1$. Da $\text{ord}(x) = M$, muss also gelten $n \geq M$, und da $\text{ord}(y') = m'$ teilerfremd zu M ist, muss sogar gelten $n > M$, im Widerspruch zur Wahl von M .

Wir haben also jetzt gezeigt, dass $\text{ord}(y) \mid M$ für alle $y \in E(\mathbb{F}_p)$, also insbesondere, dass $y^M = 1$ für alle $y \in E(\mathbb{F}_p)$. Damit ist jedes $y \in E(\mathbb{F}_p)$ eine Nullstelle von $X^M - 1$. In Satz 2.4.2 haben wir gesehen, dass $X^M - 1$ höchstens M Nullstellen hat. Da jedes Element $y \in E(\mathbb{F}_p)$ aber Nullstelle von $X^M - 1$ ist, muss also $M \geq p - 1$ gelten. Da aber andererseits $M \mid (p - 1)$, folgt hieraus $M = p - 1$, und damit ist unser Satz gezeigt.

Aufgabe 2.5.1. Unter $n+1$ ganzen Zahlen $\{a_1, \dots, a_{n+1}\}$ gibt es mindestens zwei, deren Differenz durch n teilbar ist.

Aufgabe 2.5.2. Sind $m, n, l \in \mathbb{N} \setminus \{0\}$, so gilt

$$\text{ggT}(m, n) = \text{ggT}(m, n + km)$$

Aufgabe 2.5.3. Bestimmen Sie mit Hilfe des euklidischen Algorithmus den größten gemeinsamen Teiler von

$$a) 1008 \text{ und } 840, \quad b) 481 \text{ und } 1755, \quad c) 2940 \text{ und } 1617$$

Aufgabe 2.5.4. Benutzen Sie den euklidischen Algorithmus, um folgende Aussage zu zeigen:

Sind $m, n \in \mathbb{N} \setminus \{0\}$ mit $\text{ggT}(m, n) = g$, so gibt es ganze Zahlen a, b mit

$$a \cdot m + b \cdot n = g$$

Aufgabe 2.5.5. Berechnen Sie den größten gemeinsamen Teiler g der Zahlen 318 und 99 und stellen Sie diesen in der Form

$$g = a \cdot 318 + b \cdot 99$$

mit ganzen Zahlen a, b dar.

Aufgabe 2.5.6. Für welche natürlichen Zahlen n ist $n^2 + n + 1$ durch 3 teilbar?

Aufgabe 2.5.7. a) Zeigen Sie: Ist die natürliche Zahl $m \geq 1$ ein Teiler von a und von $a + b$, so teilt m auch b .

b) Für welche natürlichen Zahlen $n \geq 1$ teilt $n + 1$ die Zahl $n^2 + 1$?

Aufgabe 2.5.8. Bestimmen Sie (in Abhängigkeit von n)

$$a) \text{ ggT}(2^n + 1, 3), \quad b) \text{ ggT}(2^n + 1, 9)$$

Aufgabe 2.5.9. Ist $p \geq 3$ eine Primzahl und $n \geq 2$ eine beliebige natürliche Zahl, so ist $2 \cdot p$ ein Teiler von $n^p - n$.

Aufgabe 2.5.10. Ist $R = \mathbb{Z}/(n)$ mit $n = p \cdot q$ wobei $p \neq q$ zwei Primzahlen sind, so hat die Einheitengruppe von R die Ordnung $(p-1) \cdot (q-1)$. Speziell gilt also für jede Zahl $a \in \mathbb{Z}$, die teilerfremd zu n ist:

$$[a^{(p-1)(q-1)}] = [1]$$

und

$$[a^{(p-1)(q-1)+1}] = [a]$$

Für die Anwendung in der Kryptographie, speziell die sogenannten *public key Kryptosysteme* wichtiger als die Körper \mathbb{F}_p sind die Ringe $\mathbb{Z}/(n)$ wobei $n = p \cdot q$ mit zwei Primzahlen $p \neq q$. Die Mathematik, die dem sogenannten RSA-Algorithmus zugrunde liegt, wird in den folgenden Aufgaben beschrieben:

Aufgabe 2.5.11. Ist $R = \mathbb{Z}/(n)$, so ist ein $[a]$ genau dann eine Einheit in R , wenn a und n teilerfremd sind.

Aufgabe 2.5.12. Ist $R = \mathbb{Z}/(n)$ mit $n = p \cdot q$ wobei $p \neq q$ zwei Primzahlen sind, so hat die Einheitengruppe von R die Ordnung $(p-1) \cdot (q-1)$. Speziell gilt also für jede Zahl $a \in \mathbb{Z}$, die teilerfremd zu n ist:

$$[a^{(p-1)(q-1)}] = [1]$$

und

$$[a^{(p-1)(q-1)+1}] = [a]$$

Damit sind wir in der Lage, den RSA-Algorithmus zu beschreiben und zu erklären:

Es seien p und q zwei voneinander verschiedene Primzahlen, und es sei $N = p \cdot q$. Wir wählen eine zu $(p-1) \cdot (q-1)$ teilerfremde ganze Zahl $e \in \mathbb{Z}$. Mit Hilfe des euklidischen Algorithmus ermitteln wir $d, t \in \mathbb{Z}$ mit

$$d \cdot e + t \cdot (p-1) \cdot (q-1) = 1.$$

Aufgabe 2.5.13. Für jedes $a \in \mathbb{Z}$ gilt in $\mathbb{Z}/(N)$:

$$[a^{e \cdot d}] = [a]$$

Damit nennt man (N, e) den **öffentlichen Schlüssel** und (N, d) den **privaten Schlüssel** des Verfahrens. Wir veröffentlichen unseren öffentlichen Schlüssel (N, e) und halten die Zahl d geheim. Will uns eine andere Partei eine Nachricht zukommen lassen, die sich durch eine Zahl m aus $\mathbb{Z}/(N)$ bzw. eine Zahl aus $\{0, 1, \dots, N-1\}$ darstellen lässt (z.B. eine Kontonummer), so kann diese Partei wie folgt vorgehen:

- Sie benutzt unseren öffentlichen Schlüssel (N, e) und berechnet

$$[a] = [m^e] \text{ in } \mathbb{Z}/(N).$$

- Sie schickt uns die Zahl a auf einem öffentlichen Kanal, etwa per Post.

Wir gehen dann wie folgt vor

- Wir benutzen unseren privaten Schlüssel d und berechnen $[b] = [a^d]$ in $\mathbb{Z}/(N)$.
- Wir benutzen Aufgabe 2.5.13 und schließen, dass

$$[b] = [(m^e)^d] = [m^{e \cdot d}] = [m]$$

erhalten also das gewünschte m .

Die Sicherheit des RSA-Kryptosystems besteht darin, dass es im allgemeinen sehr schwierig ist, alle Primteiler einer sehr großen Zahl zu ermitteln. Sind

etwa p und q Primzahlen mit mindestens 100 Stellen, so ist gegenwärtig kein effizientes Verfahren bekannt, dass aus der Kenntnis von $N = p \cdot q$ die Zahlen p und q ermittelt. Damit kann aber auch $(p - 1) \cdot (q - 1)$ nicht effizient berechnet werden, und ohne Kenntnis dieser Zahl wiederum lässt sich d nicht ermitteln. Wichtig ist also, dass wir nur N veröffentlichen und nicht seine Teiler p und q .

Beispiel 2.5.8. Wir wählen die Primzahlen $p = 7$ und $q = 13$, also $N = 91$. Dann ist $(p - 1) \cdot (q - 1) = 72$, und wir können $e = 5$ wählen. Hierfür gilt

$$5 \cdot 29 + (-2) \cdot 72 = 1$$

so dass also $d = 29$ unser privater Schlüssel ist.

Wir veröffentlichen also das Zahlenpaar $(91, 5)$.

Ein Freund will uns eine Nachricht zukommen lassen, die durch die Zahl $m = 20$ dargestellt wird. Er berechnet

$$[a] = [20^5] = [76]$$

und schickt uns die Zahl 76 auf dem Postweg. Wir ermitteln

$$[76^{29}] = [20]$$

und haben also tatsächlich die Nachricht entschlüsselt.

Bemerkung 2.5.9. die Zahl $N = 91$, die wir in unserem Beispiel gewählt haben, bietet natürlich keinerlei Sicherheit, da jeder aus ihr die Primfaktoren $p = 7$ und $q = 13$ ablesen kann. Trotzdem ist die direkte Berechnung hier schon kompliziert und kein Taschenrechner ist in der Lage, 76^{29} korrekt anzugeben. Selbst Computer kommen beim Potenzieren schnell an die Grenzen ihrer Rechengenauigkeit. Daher behilft man sich hier mit der **Methode des iterierten Quadrierens**. In unserem Beispiel etwa schreiben wir

$$29 = 16 + 8 + 4 + 1 = 2^4 + 2^3 + 2^2 + 1$$

und erhalten damit

$$76^{29} = \left(\left((76^2)^2 \right)^2 \right)^2 \cdot \left((76^2)^2 \right)^2 \cdot (76^2)^2 \cdot 76$$

Wir berechnen nun jedes Quadrat einzeln und reduzieren nach dem Quadrieren wieder modulo 91. Dadurch bleiben die Zahlen überschaubar. In unserem Beispiel etwa erhalten wir dadurch (immer modulo 91):

$$\begin{aligned} 76^2 &= 43, \\ (76^2)^2 &= 43^2 = 29, \\ \left((76^2)^2\right)^2 &= 29^2 = 22, \\ \left(\left((76^2)^2\right)^2\right)^2 &= 22^2 = 29 \end{aligned}$$

Damit ergibt sich

$$\begin{aligned} 76^{16} &= \left(\left(\left((76^2)^2\right)^2\right)^2\right)^2 = 29 \\ 76^{16} \cdot 76^8 &= 29 \cdot \left((76^2)^2\right)^2 = 29 \cdot 22 = 1 \\ 76^{16} \cdot 76^8 \cdot 76^4 &= 1 \cdot 76^4 = 1 \cdot 29 = 29 \\ 76^{29} = 76^{16} \cdot 76^8 \cdot 76^4 \cdot 76 &= 29 \cdot 76^1 = 29 \cdot 76 = 20 \end{aligned}$$

Mit diesen Techniken bleiben die Zahlen, selbst bei sehr hohen Potenzen, immer in einem Bereich, den die Prozessorarithmetik verarbeiten kann. Genauere Analysen führen noch zu einer weiteren Verbesserung dieser Methode, die in den Übungen behandelt wird.

Wir wollen hier noch eine weitere kryptographische Anwendung der elementaren Zahlentheorie untersuchen, nämlich das sogenannte *Diffie-Hellman-Protokoll*, das auf Satz 2.5.7 beruht. Dieses Verfahren wird in erster Linie zum Austausch von Schlüsseln verwendet. Dabei werden also nicht die Nachrichten selbst ausgetauscht. Die beiden Kommunikationspartner einigen sich vielmehr vorab auf eine Liste von (klassischen) Verschlüsselungsverfahren (z.B. Verfahren die von Parametern abhängen). In einer gegebenen Situation müssen sich die beiden dann nur noch auf ein Verfahren einigen. Dadurch kann das Verfahren bei jeder Kommunikation geändert werden, was einen Angriff sehr viel schwieriger macht als bei einem immer wieder verwendeten Verfahren.

Wir geben uns dazu eine Primzahl p vor (die in der praktischen Anwendung mindestens 100 Stellen haben sollte), und wir nehmen an, dass die Liste unserer Schlüssel durch die Elemente der Einheitengruppe $E(\mathbb{F}_p)$ von \mathbb{F}_p gegeben

ist. Die Einheitengruppe $E(\mathbb{F}_p)$ ist gemäß Satz 2.5.7 zyklisch, wird also von einem Element erzeugt, und wir wählen einen Erzeuger g dieser Gruppe. Diese Daten (also p und g) sind öffentlich und bekannt. Wollen nun zwei Kommunikationspartner A und B einen Schlüssel austauschen, so wählen A und B jeweils eine beliebige Zahl $a, b \in \{2, \dots, p-2\}$. Dann berechnet A die Zahl $\alpha = g^a \bmod p$ und B die Zahl $\beta = g^b \bmod p$. Diese beiden Zahlen α und β werden nun ausgetauscht. Die Zahlen a und b halten die beiden geheim. Nun berechnet B die Zahl $S_1 = \alpha^b \bmod p$ und A die Zahl $S_2 = \beta^a \bmod p$. Da (in \mathbb{F}_p) gilt

$$S_1 = \alpha^b = (g^a)^b = g^{ab} = (g^b)^a = \beta^a = S_2$$

haben also die beiden auf diese Art und Weise denselben Schlüssel S berechnet, den sie jetzt für ihre weitere Kommunikation benutzen.

Die Sicherheit des Verfahrens beruht darauf, dass es kein effizientes Verfahren gibt, aus der Kenntnis von p , g und $\alpha = g^a$ die Zahl a zu berechnen (*Problem des diskreten Logarithmus*). Die Kenntnis von g^a und g^b reicht aber nicht aus, um g^{ab} zu berechnen. Bei hinreichend großen Primzahlen kann also der Schlüssel nicht in akzeptabler Zeit berechnet werden und die Kommunikation von A und B kann daher ungestört erfolgen.

Beispiel 2.5.9. Wir erläutern hier das Diffie–Hellman–Protokoll mit Hilfe der Zahlen $p = 13$ und $g = 2$.

1. A wählt $a = 5$ und B wählt $b = 7$.
2. A berechnet $\alpha = 2^5 \bmod 13 = 32 \bmod 13 = 6 \bmod 13$ und B berechnet $\beta = 2^7 \bmod 13 = 128 \bmod 13 = 11 \bmod 13$.
3. A und B tauschen die Zahlen 6 und 11 aus.
4. A berechnet $S = 11^5 \bmod 13 = 7 \bmod 13$ und B berechnet $S = 6^7 \bmod 13 = 7 \bmod 13$.
5. A und B benutzen den Schlüssel 7 für ihre weitere Kommunikation.

Kapitel 3

Allgemeine Vektorräume

3.1 Vektorräume und Untervektorräume

Im letzten Kapitel haben wir gesehen, dass Mengen unterschiedliche Strukturen haben können und mithilfe dieser Strukturen besser untersucht werden können. Je mehr Struktur ein Objekt dabei hat, desto einfacher ist es zu untersuchen. Am meisten kann man deshalb auch über Körper sagen. Es gibt aber auch viele Menge, die selbst keine Körper sind, die aber sehr gut mit Hilfe von Körpern untersucht werden können. Mit diesen Objekten, den **Vektorräumen** wollen wir uns in diesem Kapitel beschäftigen.

Dazu betrachten wir zunächst einen beliebigen Körper K .

Definition 3.1.1. Ein K -**Vektorraum** ist eine nichtleere Menge V zusammen mit einer Addition

$$+' : V \times V \longrightarrow V, \quad (\vec{v}, \vec{w}) \longmapsto \vec{v} + \vec{w}$$

und einer Skalarmultiplikation

$$'\cdot : K \times V \longrightarrow V, \quad (r, \vec{v}) \longmapsto r \cdot \vec{v}$$

so dass für \vec{u}, \vec{v} und \vec{w} in V und Skalare r und s in K gilt:

- $V1: (\vec{u} + \vec{v}) + \vec{w} = \vec{u} + (\vec{v} + \vec{w})$ (1. Assoziativgesetz)
 $V2: \vec{v} + \vec{w} = \vec{w} + \vec{v}$ (Kommutativgesetz)
 $V3: \text{Es existiert ein Element } \vec{0} \in V$
 mit $\vec{v} + \vec{0} = \vec{v}$ (neutrales Element der Addition)

 $V4: \text{Zu jedem } \vec{v} \in V \text{ existiert ein}$
 $-\vec{v} \in V \text{ mit } \vec{v} + (-\vec{v}) = \vec{0}$ (inverses Element der Addition)

 $V5: (r \cdot s) \cdot \vec{v} = r \cdot (s \cdot \vec{v})$ (2. Assoziativgesetz)
 $V6: r \cdot (\vec{v} + \vec{w}) = r \cdot \vec{v} + r \cdot \vec{w}$ (1. Distributivgesetz)
 $V7: (r + s) \cdot \vec{v} = r \cdot \vec{v} + s \cdot \vec{v}$ (2. Distributivgesetz)
 $V8: 1 \cdot \vec{v} = \vec{v}$ (neutrales Element der Multiplikation)

Die Elemente $\vec{v} \in V$ heißen **Vektoren**.

Beispiel 3.1.1. Das n -fache kartesische Produkt von \mathbb{R} , also die Menge

$$\mathbb{R}^n := \left\{ \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \mid v_i \in \mathbb{R} \right\}$$

zusammen mit der komponentenweisen Addition

$$\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} + \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} = \begin{pmatrix} v_1 + w_1 \\ \vdots \\ v_n + w_n \end{pmatrix}$$

($v_i, w_i \in \mathbb{R}$) und der komponentenweisen Skalarmultiplikation

$$r \cdot \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} r \cdot v_1 \\ \vdots \\ r \cdot v_n \end{pmatrix}$$

($r, v_i \in \mathbb{R}$) ist ein \mathbb{R} -Vektorraum.

Ein Vektor

$$\vec{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$$

heißt n -**dimensionaler reeller Vektor**.

Der Nullvektor

$$\vec{0} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

ist dabei das neutrale Element der Addition, und der negative Vektor $-\vec{v}$ zu einem n -dimensionalen reellen Vektor ist gegeben durch

$$-\vec{v} = \begin{pmatrix} -v_1 \\ \vdots \\ -v_n \end{pmatrix}$$

Die Axiome $V1$ bis $V8$ lassen sich in diesem Fall sehr einfach nachrechnen. Der Vektorraum \mathbb{R}^n wird auch als n -**dimensionaler reeller Vektorraum** bezeichnet.

Beispiel 3.1.2. Die Ebene kann mit \mathbb{R}^2 identifiziert werden, der Raum mit \mathbb{R}^3 . Daher sind Ebene und Raum reelle Vektorräume. Die geometrischen Vektoren (Pfeile im Raum und in der Ebene), wie Sie sie aus der Schule (oder auch aus dem Vorkurs *Lineare Algebra*) kennen, sind Vektoren im Sinne unserer neuen Definition.

Beispiel 3.1.3. Das n -fache kartesische Produkt von \mathbb{C} , also die Menge

$$\mathbb{C}^n := \left\{ \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} \mid z_i \in \mathbb{C} \right\}$$

zusammen mit der komponentenweisen Addition

$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} + \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} = \begin{pmatrix} y_1 + z_1 \\ \vdots \\ y_n + z_n \end{pmatrix}$$

$(y_i, z_i \in \mathbb{C})$ und der komponentenweisen Skalarmultiplikation

$$a \cdot \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} = \begin{pmatrix} a \cdot z_1 \\ \vdots \\ a \cdot z_n \end{pmatrix}$$

$(a, z_i \in \mathbb{C})$ ist ein \mathbb{C} -Vektorraum.

Beispiel 3.1.4. Das n -fache kartesische Produkt von \mathbb{F}_2 , also die Menge

$$\mathbb{F}_2^n := \left\{ \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \mid a_i \in \mathbb{F}_2 \right\}$$

zusammen mit komponentenweiser Addition und Skalarmultiplikation ist ein \mathbb{F}_2 -Vektorraum.

Bei \mathbb{F}_2^n handelt es sich um die n -Tupel binärer Zahlen, die sich komplett auflisten lassen. So gilt etwa

$$\mathbb{F}_2^2 := \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}$$

Insbesondere ist also \mathbb{F}_2^n eine endliche Menge mit

$$|\mathbb{F}_2^n| = 2^n$$

Der Körper \mathbb{F}_2 und die Vektorräume über ihm spielen eine fundamentale Rolle in der Informatik und der Computerprogrammierung.

Beispiel 3.1.5. Wir betrachten die Menge

$$V = \text{Abb}(\mathbb{N}, \mathbb{R}) = \{f : \mathbb{N} \longrightarrow \mathbb{R} \text{ Abbildung} \}$$

aller Abbildungen von \mathbb{N} nach \mathbb{R} . Für $f, g \in V$ definieren wir $f + g$ durch

$$(f + g)(n) = f(n) + g(n) \quad \text{für alle } n \in \mathbb{N}$$

und für $f \in V$ und $r \in \mathbb{R}$ definieren wir $r \cdot f$ durch

$$(r \cdot f)(n) = r \cdot f(n) \quad \text{für alle } n \in \mathbb{N}$$

Dann ist V ein \mathbb{R} -Vektorraum.

In dieser Situation schreibt man häufig f_n für $f(n)$ und $(f_n)_{n \in \mathbb{N}}$ für f und nennt ein Element $(f_n)_{n \in \mathbb{N}}$ eine **reelle Zahlenfolge**. Der Raum V heißt auch Vektorraum der Folgen.

Beispiel 3.1.6. Die Menge $V = \{\vec{0}\}$ mit den Operationen

$$\vec{0} + \vec{0} = \vec{0}, \quad r \cdot \vec{0} = \vec{0}$$

ist ein Vektorraum (über jedem Körper), der **Nullvektorraum** bzw. der **triviale Vektorraum**.

Auch Teilmengen von Vektorräumen können interessante Strukturen tragen:

Definition 3.1.2. Ist V ein K -Vektorraum und ist $U \subseteq V$ eine lTeilmenge, so heißt U **Untervektorraum** von V wenn gilt:

1. $U \neq \emptyset$.
2. Sind $\vec{v}, \vec{w} \in U$ so ist auch $\vec{v} + \vec{w} \in U$.
3. Ist $\vec{v} \in U$ und ist $\kappa \in K$ ein Skalar, so ist auch $\kappa \cdot \vec{v} \in U$.

Beispiel 3.1.7. Jede Gerade durch den Koordinatenursprung ist ein Untervektorraum der Ebene (als reeller Vektorraum betrachtet).

Alle Geraden und alle Ebenen durch den Koordinatenursprung sind Untervektorräume des Raums (als reeller Vektorraum betrachtet).

Beispiel 3.1.8. Ist V ein beliebiger K -Vektorraum, so sind $U_1 = \{\vec{0}\}$ und $U_2 = V$ (offensichtlich) Untervektorräume von V . Diese beiden Untervektorräume werden auch die trivialen Untervektorräume genannt.

Beispiel 3.1.9. Bezeichnen wir mit

$$V = \{f : \mathbb{N} \rightarrow \mathbb{R} \text{ Abbildung} \}$$

den Vektorraum der reellen Zahlenfolgen (aus Beispiel 3.1.5), und mit

$$U = \{(f_n)_{n \in \mathbb{N}} \mid \lim_{n \rightarrow \infty} f_n = 0\}$$

die Teilmenge der Nullfolgen, so ist $U \subseteq V$ ein Untervektorraum.

Dabei ist $U \neq \emptyset$, denn die Nullfolge $(f_n)_{n \in \mathbb{N}}$ mit $f_n = 0$ ist in U . Die weiteren Bedingungen für einen Untervektorraum werden wir in der Vorlesung *Analysis in einer Veränderlichen* im zweiten Semester nachrechnen.

Bemerkung 3.1.1. Ein Untervektorraum $U \subseteq V$ ist selbst wieder ein Vektorraum. Offensichtlich gelten nämlich die Axiome V1 bis V8 auch in U .

Aufgabe 3.1.1. Zeigen Sie, dass jede Gerade durch den Koordinatenursprung tatsächlich ein Untervektorraum der Ebene (aufgefasst als reeller Vektorraum) ist.

Aufgabe 3.1.2. Zeigen Sie, dass jede Gerade bzw. jede Ebene durch den Koordinatenursprung tatsächlich ein Untervektorraum des Raums (aufgefasst als reeller Vektorraum) ist.

Aufgabe 3.1.3. Wir betrachten einen K -Vektorraum V und zwei Untervektorräume $U_1, U_2 \subseteq V$ von V . Zeigen Sie dass dann auch $U_1 \cap U_2$ wieder ein Untervektorraum von V ist. Geben Sie ein Beispiel an, das zeigt, dass $U_1 \cup U_2$ nicht notwendigerweise ein Untervektorraum von V ist.

3.2 Der n -dimensionale reelle Raum

Im ersten Abschnitt haben wir als Beispiel (Beispiel 3.1.1) den reellen Vektorraum \mathbb{R}^n betrachtet. Dieser Vektorraum ist für viele Anwendungen auch der wichtigste und wird daher nun in diesem Abschnitt ausführlicher behandelt werden.

Die Vektorraumoperationen können in diesem Fall sehr einfach und explizit beschrieben werden:

Beispiel 3.2.1. Es ist

$$\begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix} + \begin{pmatrix} 2 \\ -3 \\ 4 \\ -5 \end{pmatrix} = \begin{pmatrix} 1+2 \\ 2+(-3) \\ 3+4 \\ 4+(-5) \end{pmatrix} = \begin{pmatrix} 3 \\ -1 \\ 7 \\ -1 \end{pmatrix}$$

und

$$\begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix} - \begin{pmatrix} 2 \\ -3 \\ 4 \\ -5 \end{pmatrix} = \begin{pmatrix} 1-2 \\ 2-(-3) \\ 3-4 \\ 4-(-5) \end{pmatrix} = \begin{pmatrix} -1 \\ 5 \\ -1 \\ 9 \end{pmatrix}$$

Ferner gilt

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

und

$$3 \cdot \begin{pmatrix} 2 \\ 3 \\ -1 \\ 2 \end{pmatrix} = \begin{pmatrix} 3 \cdot 2 \\ 3 \cdot 3 \\ 3 \cdot (-1) \\ 3 \cdot 2 \end{pmatrix} = \begin{pmatrix} 6 \\ 9 \\ -3 \\ 6 \end{pmatrix}$$

Definition 3.2.1. Ein Vektor

$$\vec{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \in \mathbb{R}^n$$

heißt n -dimensionaler reeller Vektor, die Größe

$$\|\vec{v}\| = \sqrt{v_1^2 + \cdots + v_n^2}$$

heißt **Länge** des Vektors \vec{v} .

Wir betrachten nun Vektoren $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m$ in \mathbb{R}^n .

Definition 3.2.2. Die n -Vektoren $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m$ heißen **linear abhängig**, wenn es reelle Zahlen r_1, r_2, \dots, r_m gibt, von denen mindestens eine von Null verschieden ist, mit

$$r_1 \cdot \vec{v}_1 + r_2 \cdot \vec{v}_2 + \cdots + r_m \cdot \vec{v}_m = \vec{0}$$

Andernfalls heißen sie **linear unabhängig**.

Bemerkung 3.2.1. Die Bedingung für lineare Unabhängigkeit kann auch so formuliert werden: Sind r_1, r_2, \dots, r_m reelle Zahlen mit

$$r_1 \cdot \vec{v}_1 + r_2 \cdot \vec{v}_2 + \cdots + r_m \cdot \vec{v}_m = \vec{0}$$

so muss schon gelten: $r_1 = r_2 = \cdots = r_m = 0$.

Beispiel 3.2.2. Die Vektoren $\vec{v}_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 1 \end{pmatrix}$ und $\vec{v}_2 = \begin{pmatrix} -2 \\ -4 \\ -6 \\ -2 \end{pmatrix}$ sind linear abhängig: Für $r_1 = 2$ und $r_2 = 1$ gilt

$$r_1 \cdot \vec{v}_1 + r_2 \cdot \vec{v}_2 = 2 \cdot \begin{pmatrix} 1 \\ 2 \\ 3 \\ 1 \end{pmatrix} + 1 \cdot \begin{pmatrix} -2 \\ -4 \\ -6 \\ -2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

Beispiel 3.2.3. Die Vektoren $\vec{v}_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 1 \end{pmatrix}$, $\vec{v}_2 = \begin{pmatrix} 2 \\ 3 \\ 4 \\ 2 \end{pmatrix}$ und $\vec{v}_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$ sind linear abhängig: Für $r_1 = 1$ und $r_2 = -1$ und $r_3 = 1$ gilt:

$$r_1 \cdot \vec{v}_1 + r_2 \cdot \vec{v}_2 + r_3 \cdot \vec{v}_3 = \vec{0}$$

Beispiel 3.2.4. Die Vektoren $\vec{v}_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}$ und $\vec{v}_2 = \begin{pmatrix} 3 \\ 2 \\ 1 \\ 0 \end{pmatrix}$ sind linear unabhängig: Sind nämlich $r_1, r_2 \in \mathbb{R}$ Skalare mit

$$r_1 \cdot \vec{v}_1 + r_2 \cdot \vec{v}_2 = \vec{0}$$

so bedeutet dies

$$\begin{pmatrix} r_1 \cdot 1 \\ r_1 \cdot 2 \\ r_1 \cdot 3 \\ r_1 \cdot 4 \end{pmatrix} + \begin{pmatrix} r_2 \cdot 3 \\ r_2 \cdot 2 \\ r_2 \cdot 1 \\ r_2 \cdot 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

also

$$\begin{aligned} 1 \cdot r_1 + 3 \cdot r_2 &= 0 \\ 2 \cdot r_1 + 2 \cdot r_2 &= 0 \\ 3 \cdot r_1 + 1 \cdot r_2 &= 0 \\ 4 \cdot r_1 + 0 \cdot r_2 &= 0 \end{aligned}$$

Aus der letzten Gleichung folgt sofort, dass $r_1 = 0$ gelten muss. Setzt man das in die vorletzte Gleichung ein, so ergibt sich unmittelbar $r_2 = 0$, und damit sind \vec{v}_1 und \vec{v}_2 linear unabhängig.

Beispiel 3.2.5. Die n Vektoren $\vec{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$, $\vec{e}_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}$, \dots , $\vec{e}_n = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$ im \mathbb{R}^n sind linear unabhängig. Sind nämlich r_1, r_2, \dots, r_n Skalare mit

$$r_1 \cdot \vec{e}_1 + r_2 \cdot \vec{e}_2 + \dots + r_n \cdot \vec{e}_n = \vec{0}$$

so bedeutet dies

$$\begin{pmatrix} r_1 \cdot 1 + r_2 \cdot 0 + \dots + r_n \cdot 0 \\ r_1 \cdot 0 + r_2 \cdot 1 + \dots + r_n \cdot 0 \\ \vdots \\ r_1 \cdot 0 + r_2 \cdot 0 + \dots + r_n \cdot 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

also

$$\begin{pmatrix} r_1 \\ r_2 \\ \vdots \\ r_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

und damit $r_1 = r_2 = \dots = r_n = 0$.

Bemerkung 3.2.2. Ist $\vec{v}_k = \vec{0}$ für ein k , so sind die Vektoren $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m$ schon linear abhängig. Hierzu können wir etwa $r_k = 1$ und $r_i = 0$ für $i \neq k$ setzen und erhalten eine nicht-triviale Linearkombination

$$r_1 \cdot \vec{v}_1 + \dots + r_n \cdot \vec{v}_n = \vec{0}$$

Bemerkung 3.2.3. Ein einzelner Vektor $\vec{v} \in \mathbb{R}^n$ ist genau dann linear abhängig, wenn $\vec{v} = \vec{0}$. Entsprechend ist er genau dann linear unabhängig, wenn $\vec{v} \neq \vec{0}$.

Bemerkung 3.2.4. Sind die Vektoren $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m$ linear unabhängig, so auch jede Teilmenge davon, d.h. für $1 \leq i_1 < i_2 < \dots < i_t \leq m$ sind auch die Vektoren $\vec{v}_{i_1}, \vec{v}_{i_2}, \dots, \vec{v}_{i_t}$ linear unabhängig.

Die entsprechende Aussage für linear Abhängigkeit ist nicht richtig.

Beispiel 3.2.6. Die Vektoren $\vec{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$, $\vec{e}_4 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$ im \mathbb{R}^4 sind linear unabhängig. Sie sind Teil des Systems linear unabhängiger Vektoren aus Beispiel 3.2.5.

Natürlich kann die lineare Unabhängigkeit in diesem Fall auch direkt nachgerechnet werden.

Beispiel 3.2.7. Die Vektoren $\vec{v}_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 1 \end{pmatrix}$ und $\vec{v}_2 = \begin{pmatrix} 2 \\ 3 \\ 4 \\ 2 \end{pmatrix}$ sind linear unabhängig, wie man leicht nachrechnet, obwohl sie eine Teilmenge der linear abhängigen Vektoren $\vec{v}_1, \vec{v}_2, \vec{v}_3$ aus Beispiel 3.2.3 bilden.

Bemerkung 3.2.5. Zwei Vektoren \vec{v} und \vec{w} sind genau dann linear abhängig, wenn einer von beiden ein Vielfaches des anderen ist. Gilt nämlich

$$r_1 \cdot \vec{v} + r_2 \cdot \vec{w} = \vec{0}$$

und ist etwa $r_1 \neq 0$, so gilt

$$\vec{v} = -\frac{r_2}{r_1} \cdot \vec{w}$$

Analog ist natürlich \vec{w} ein Vielfaches von \vec{v} falls $r_2 \neq 0$.

Bemerkung 3.2.6. Die lineare Unabhängigkeit von drei oder mehr Vektoren ist in der Regel schwer elementar und direkt festzustellen. Mit den linearen Gleichungssystemen werden wir das richtige Werkzeug zur Behandlung dieser Frage noch kennenlernen.

Eine wichtige Rolle beim Studium linearer Gleichungssysteme werden die Untervektorräume der \mathbb{R}^n bilden.

Beispiel 3.2.8. Ist $V = \left\{ \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \right\} \subseteq \mathbb{R}^4$, so ist V ein Untervektorraum von \mathbb{R}^4 .

Beispiel 3.2.9. Ist $V = \left\{ \begin{pmatrix} 0 \\ r \\ 0 \\ 0 \end{pmatrix} \mid r \in \mathbb{R} \right\} \subseteq \mathbb{R}^4$, so ist V ein Untervektorraum von \mathbb{R}^4 .

Beispiel 3.2.10. Ist $V = \left\{ \begin{pmatrix} 0 \\ r \\ 0 \\ 0 \end{pmatrix} \mid r \in \mathbb{R} \text{ und } r \geq 0 \right\} \subseteq \mathbb{R}^4$, so ist V kein Untervektorraum. Es ist nämlich $\begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \in V$ aber nicht $(-1) \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$.

Beispiel 3.2.11. Ist $V = \left\{ \begin{pmatrix} 0 \\ r \\ 1 \\ 0 \end{pmatrix} \mid r \in \mathbb{R} \right\} \subseteq \mathbb{R}^4$, so ist V kein Untervektorraum. Es ist nämlich $\begin{pmatrix} 0 \\ 2 \\ 1 \\ 0 \end{pmatrix} \in V$ aber nicht $2 \cdot \begin{pmatrix} 0 \\ 2 \\ 1 \\ 0 \end{pmatrix}$.

Beispiel 3.2.12. Ist $\vec{v} \in \mathbb{R}^n$ ein beliebiger Vektor, so ist

$$V := \{r \cdot \vec{v} \mid r \in \mathbb{R}\}$$

ein Untervektorraum (auch, falls \vec{v} der Nullvektor ist).

Beispiel 3.2.13. Ist $V = \left\{ \begin{pmatrix} r \\ s \\ r+s \end{pmatrix} \mid r, s \in \mathbb{R} \right\} \subseteq \mathbb{R}^3$, so ist V ein Untervektorraum.

Gesucht ist häufig eine effiziente und knappe Beschreibung von Untervektorräumen mit möglichst wenig Daten. Dazu geben wir uns einen Untervektorraum $V \subseteq \mathbb{R}^n$ vor und betrachten Vektoren $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m$ in V .

Definition 3.2.3. Die Vektoren $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m$ heißen **Erzeugendensystem** von V , wenn gilt:

1. $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m \in V$.
2. Zu jedem $\vec{w} \in V$ gibt es Skalare r_1, r_2, \dots, r_m mit

$$\vec{w} = r_1 \cdot \vec{v}_1 + r_2 \cdot \vec{v}_2 + \dots + r_m \cdot \vec{v}_m$$

Wir sagen in diesem Fall auch, $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m$ erzeugen V .

Beispiel 3.2.14. Die Vektoren $\vec{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$, $\vec{e}_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \dots, \vec{e}_n = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$ erzeugen $V = \mathbb{R}^n$ (als Untervektorraum von \mathbb{R}^n). Es gilt nämlich für einen beliebigen Vektor $\vec{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$ in \mathbb{R}^n :

$$\vec{v} = v_1 \cdot \vec{e}_1 + v_2 \cdot \vec{e}_2 + \dots + v_n \cdot \vec{e}_n$$

Beispiel 3.2.15. Ist $V = \left\{ \begin{pmatrix} 0 \\ r \\ 0 \\ 0 \end{pmatrix} \mid r \in \mathbb{R} \right\} \subseteq \mathbb{R}^4$ der Untervektorraum aus Beispiel 3.2.9 und ist $\vec{v} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$, so erzeugt \vec{v} den Untervektorraum V . Ebenso wird V erzeugt von $\vec{w} = \begin{pmatrix} 0 \\ -3 \\ 0 \\ 0 \end{pmatrix}$, und auch $\{\vec{v}, \vec{w}\}$ bildet ein Erzeugendensystem von V . Erzeugendensysteme sind also nicht eindeutig und können unterschiedlich viele Elemente enthalten.

Beispiel 3.2.16. Ist $V = \{\vec{0}\}$ der Nullvektorraum, so bildet der Nullvektor $\vec{0}$ ein Erzeugendensystem von V . Es ist aber auch üblich, die leere Menge als Erzeugendensystem von V zu betrachten.

Beispiel 3.2.17. Ist E die Ebene durch die Punkte $O = (0, 0, 0)$, $P = (1, 2, 3)$ und $Q = (-3, 2, -1)$, so ist E ein Untervektorraum von \mathbb{R}^3 (siehe Beispiel 3.1.7), der erzeugt wird von den Vektoren

$$\vec{v}_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \quad \vec{v}_2 = \begin{pmatrix} -3 \\ 2 \\ -1 \end{pmatrix}$$

also von den beiden Vektoren, die (zusammen mit $\vec{0}$ als Stützvektor) E auch als Ebene erzeugen. Ebenso bildet auch jedes andere System von Vektoren, das mit $\vec{0}$ als Stützvektor die Ebene E erzeugt, ein Untervektorraum erzeugendensystem von E , etwa

$$\vec{w}_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \quad \vec{w}_2 = \begin{pmatrix} 0 \\ 4 \\ 4 \end{pmatrix}, \quad \vec{w}_3 = \begin{pmatrix} 2 \\ 0 \\ 2 \end{pmatrix}$$

Wir betrachten nun Vektoren $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m \in \mathbb{R}^n$ gegeben, und setzen

$$U = \{ \vec{v} \in \mathbb{R}^n \mid \exists r_1, \dots, r_m \in \mathbb{R} \text{ mit } \vec{v} = r_1 \cdot \vec{v}_1 + \dots + r_m \cdot \vec{v}_m \}$$

Satz 3.2.1. *U ist ein Untervektorraum von \mathbb{R}^n .*

Beweis: Die Unterraumeigenschaften sind sehr leicht nachzurechnen. Wir zeigen, dass U abgeschlossen unter Addition ist. Seien dazu \vec{v} und \vec{w} in U . Schreibe

$$\begin{aligned}\vec{v} &= r_1 \cdot \vec{v}_1 + \cdots + r_m \cdot \vec{v}_m \\ \vec{w} &= s_1 \cdot \vec{v}_1 + \cdots + s_m \cdot \vec{v}_m\end{aligned}$$

Dann gilt

$$\vec{v} + \vec{w} = (r_1 + s_1) \cdot \vec{v}_1 + \cdots + (r_m + s_m) \cdot \vec{v}_m$$

und da $r_i + s_i \in \mathbb{R}$ folgt

$$\vec{v} + \vec{w} \in U$$

Die Abgeschlossenheit unter Skalarmultiplikation zeigt man ähnlich.

Bezeichnung::

U heißt das **Erzeugnis** von $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m$ oder der von $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m$ **aufgespannte Unterraum** von \mathbb{R}^n . Wir schreiben hierfür $\langle \{\vec{v}_i\}_{i=1, \dots, m} \rangle$ oder auch $\text{Span}(\{\vec{v}_i\}_{i=1, \dots, m})$.

Beispiel 3.2.18. Wir betrachten die Vektoren

$$\vec{v}_1 = \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix}, \quad \vec{v}_2 = \begin{pmatrix} 0 \\ 4 \\ 0 \end{pmatrix}$$

in \mathbb{R}^3 . Dann gilt hierfür

$$\langle \{\vec{v}_1, \vec{v}_2\} \rangle = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ 0 \end{pmatrix} \mid x_1, x_2 \in \mathbb{R} \right\}$$

Klar ist dabei, dass

$$\langle \{\vec{v}_1, \vec{v}_2\} \rangle \subseteq \left\{ \begin{pmatrix} x_1 \\ x_2 \\ 0 \end{pmatrix} \mid x_1, x_2 \in \mathbb{R} \right\}$$

denn mit \vec{v}_1 und \vec{v}_2 wird auch jede Linearkombination von \vec{v}_1 und \vec{v}_2 in der letzten Komponente eine "0" stehen haben. Ist umgekehrt ein Vektor $\vec{v} = \begin{pmatrix} x_1 \\ x_2 \\ 0 \end{pmatrix}$ gegeben, so können wir schreiben

$$\vec{v} = (-x_1) \cdot \vec{v}_1 + \frac{x_2 - 2x_1}{4} \cdot \vec{v}_2$$

und erhalten $\vec{v} \in \langle \{\vec{v}_1, \vec{v}_2\} \rangle$, also

$$\langle \{\vec{v}_1, \vec{v}_2\} \rangle \subseteq \left\{ \begin{pmatrix} x_1 \\ x_2 \\ 0 \end{pmatrix} \mid x_1, x_2 \in \mathbb{R} \right\}$$

und damit Gleichheit der beiden Mengen.

Definition 3.2.4. Die Vektoren $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m \in V$ heißen **Basis** von V , wenn sie ein Erzeugendensystem von V bilden, und wenn sie linear unabhängig sind.

Beispiel 3.2.19. Die Vektoren $\vec{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$, $\vec{e}_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \dots, \vec{e}_n = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$ bilden eine Basis von \mathbb{R}^n . In Beispiel 3.2.5 haben wir schon gesehen, dass sie linear unabhängig sind, und in Beispiel 3.2.14 haben wir festgestellt, dass sie den \mathbb{R}^n erzeugen.

Die Vektoren $\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n$ heißen **Standardbasis** des \mathbb{R}^n .

Beispiel 3.2.20. Ist E die Ebene aus Beispiel 3.2.17, aufgefasst als Untervektorraum von \mathbb{R}^n . Dann bilden die Vektoren

$$\vec{v}_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \quad \vec{v}_2 = \begin{pmatrix} -3 \\ 2 \\ -1 \end{pmatrix}$$

eine Basis von E . Wie wir in Beispiel 3.1.7 gesehen haben, erzeugen sie E , und wie man leicht sieht, sind sie linear unabhängig, da keiner der Vektoren ein Vielfaches des anderen ist. Dagegen bilden die Vektoren

$$\vec{w}_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \quad \vec{w}_2 = \begin{pmatrix} 0 \\ 4 \\ 4 \end{pmatrix}, \quad \vec{w}_3 = \begin{pmatrix} 2 \\ 0 \\ 2 \end{pmatrix}$$

keine Basis von E . Sie erzeugen zwar, aber es gilt

$$2 \cdot \vec{w}_1 + (-1) \cdot \vec{w}_2 + (-1) \cdot \vec{w}_3 = \vec{0},$$

sie sind also nicht linear unabhängig.

Im nächsten Abschnitt werden wir den folgenden hilfreichen Satz beweisen:

Satz 3.2.2. *Für einen Untervektorraum V von \mathbb{R}^n gilt:*

1. *Ist $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m\}$ ein Erzeugendensystem von V , so enthält $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m\}$ eine Basis von V , d.h. es gibt $1 \leq i_1 < i_2 < \dots < i_t \leq m$ so dass $\vec{v}_{i_1}, \vec{v}_{i_2}, \dots, \vec{v}_{i_t}$ eine Basis von V ist.*
2. *V hat eine Basis.*
3. *Je zwei Basen von V sind gleich lang, d.h. sind $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m\}$ und $\{\vec{w}_1, \vec{w}_2, \dots, \vec{w}_t\}$ Basen von V , so gilt $m = t$.*

Dieser Satz rechtfertigt die folgende Festsetzung

Definition 3.2.5. Die Länge einer Basis eines Untervektorraums V heißt die **Dimension** von V und wird mit $\dim(V)$ bezeichnet.

Beispiel 3.2.21. Der \mathbb{R}^n hat die Dimension n .

Beispiel 3.2.22. Die Ebene E aus Beispiel 3.1.7 hat die Dimension 2, da die Vektoren \vec{v}_1, \vec{v}_2 aus Beispiel 3.2.20 eine Basis von E bilden. Es ist also $\dim(V) = 2$.

Die Vektoren \vec{w}_1, \vec{w}_2 und \vec{w}_3 aus Beispiel 3.2.20 bilden ein Erzeugendensystem von V aber keine Basis. \vec{w}_2 und \vec{w}_3 bilden eine in diesem Erzeugendensystem enthaltene Basis von E .

Beispiel 3.2.23. Sei $\vec{v} \in \mathbb{R}^n$ ein beliebiger Vektor und sei

$$V := \{r \cdot \vec{v} \mid r \in \mathbb{R}\}$$

der Untervektorraum aus 3.2.12. Ist $\vec{v} \neq \vec{0}$ so ist \vec{v} eine Basis von V und $\dim(V) = 1$.

Beispiel 3.2.24. Wie wir in Beispiel 3.2.16 festgesetzt haben, bildet die leere Menge ein Erzeugendensystem des Nullvektorraums $V = \{\vec{0}\}$. Die leere Menge ist auch ein System linear unabhängiger Vektoren (da keine Bedingungen zu verifizieren sind), und daher ist die leere Menge eine Basis des Nullvektorraums. Damit gilt $\dim(V) = 0$.

Eine Besonderheit der Vektorräume \mathbb{R}^n ist die Möglichkeit, dort auch geometrische Überlegungen durchzuführen. Speziell in der Ebene und im Raum sind Begriffe wie Abstand und Winkel bekannt. Teilweise übertragen sich diese Konzepte auch auf höhere Dimensionen. Ein wichtiges Hilfsmittel dabei ist das Skalarprodukt:

Definition 3.2.6. Für zwei n -Vektoren $\vec{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$ und $\vec{w} = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}$ heißt

$$\langle \vec{v}, \vec{w} \rangle := v_1 w_1 + \cdots + v_n w_n$$

das **Skalarprodukt** von \vec{v} und \vec{w} .

Beispiel 3.2.25. $\langle \begin{pmatrix} 1 \\ 2 \\ 3 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 3 \\ 4 \\ 2 \end{pmatrix} \rangle = 1 \cdot 2 + 2 \cdot 3 + 3 \cdot 4 + 1 \cdot 2 = 22$.

Bemerkung 3.2.7. Für einen Vektor $\vec{v} \in \mathbb{R}^n$ gilt

$$|\vec{v}| = \sqrt{\langle \vec{v}, \vec{v} \rangle}$$

Zunächst ist nämlich

$$\langle \vec{v}, \vec{v} \rangle = v_1 \cdot v_1 + \cdots + v_n \cdot v_n = v_1^2 + \cdots + v_n^2 \geq 0$$

(sodass wir die Wurzel daraus ziehen können) und damit

$$\sqrt{\langle \vec{v}, \vec{v} \rangle} = \sqrt{v_1^2 + \cdots + v_n^2} = |\vec{v}|$$

Einige elementaren Eigenschaften des Skalarprodukts folgen sofort aus der Definition

Regel 3.2.3. Für Vektoren \vec{v} , \vec{w} , \vec{w}_1 und \vec{w}_2 und einen Skalar $r \in \mathbb{R}$ gilt:

$$1. \langle \vec{v}, \vec{w} \rangle = \langle \vec{w}, \vec{v} \rangle \quad (\text{Kommutativgesetz}).$$

2. $\langle \vec{v}, \vec{w}_1 + \vec{w}_2 \rangle = \langle \vec{v}, \vec{w}_1 \rangle + \langle \vec{v}, \vec{w}_2 \rangle$ (Distributivgesetz).
3. $\langle r \cdot \vec{v}, \vec{w} \rangle = r \cdot \langle \vec{v}, \vec{w} \rangle$ (Skalarmultiplikation).
4. $\langle \vec{v}, \vec{v} \rangle = |\vec{v}|^2$.

Wie im zwei- oder dreidimensionalen Fall gilt auch hier der Satz von Cauchy-Schwarz

Satz 3.2.4 (Satz von Cauchy-Schwarz). *Für zwei Vektoren \vec{v} und \vec{w} gilt:*

1. $|\langle \vec{v}, \vec{w} \rangle| \leq |\vec{v}| \cdot |\vec{w}|$.
2. Genau dann sind \vec{v} und \vec{w} parallel, wenn $\langle \vec{v}, \vec{w} \rangle = |\vec{v}| \cdot |\vec{w}|$.
3. Genau dann sind \vec{v} und \vec{w} anti-parallel, wenn $\langle \vec{v}, \vec{w} \rangle = -|\vec{v}| \cdot |\vec{w}|$.
4. Genau dann sind \vec{v} und \vec{w} kollinear, wenn $|\langle \vec{v}, \vec{w} \rangle| = |\vec{v}| \cdot |\vec{w}|$.

Beweis: Der direkte Beweis dieser Aussage ist - sogar schon im einfachen zweidimensionalen Fall - recht komplex und unübersichtlich. Daher benutzen wir einen kleinen Trick und setzen $r = \langle \vec{w}, \vec{w} \rangle$ und $s = -\langle \vec{v}, \vec{w} \rangle$. Wir können annehmen, dass $\vec{w} \neq 0$, also $r > 0$. Dann gilt

$$\begin{aligned}
 0 &\leq |r\vec{v} + s\vec{w}|^2 \\
 &= \langle r\vec{v} + s\vec{w}, r\vec{v} + s\vec{w} \rangle \\
 &= r^2\langle \vec{v}, \vec{v} \rangle + 2rs\langle \vec{v}, \vec{w} \rangle + s^2\langle \vec{w}, \vec{w} \rangle \\
 &= r^2\langle \vec{v}, \vec{v} \rangle - 2r\langle \vec{v}, \vec{w} \rangle^2 + s^2\langle \vec{w}, \vec{w} \rangle \\
 &= r^2\langle \vec{v}, \vec{v} \rangle - 2r\langle \vec{v}, \vec{w} \rangle^2 + \langle \vec{v}, \vec{w} \rangle^2 \cdot \langle \vec{w}, \vec{w} \rangle \\
 &= r\langle \vec{w}, \vec{w} \rangle \cdot \langle \vec{v}, \vec{v} \rangle - 2r\langle \vec{v}, \vec{w} \rangle^2 + r\langle \vec{v}, \vec{w} \rangle^2 \\
 &= r \cdot (\langle \vec{w}, \vec{w} \rangle \cdot \langle \vec{v}, \vec{v} \rangle - \langle \vec{v}, \vec{w} \rangle^2)
 \end{aligned}$$

Da $r > 0$, folgt

$$\langle \vec{v}, \vec{w} \rangle^2 \leq \langle \vec{w}, \vec{w} \rangle \cdot \langle \vec{v}, \vec{v} \rangle = |\vec{w}|^2 \cdot |\vec{v}|^2$$

und daher (durch Ziehen der Quadratwurzel)

$$|\langle \vec{v}, \vec{w} \rangle| \leq |\vec{v}| \cdot |\vec{w}|$$

also Aussage 1., und genau dann ist

$$|\langle \vec{v}, \vec{w} \rangle| = |\vec{v}| \cdot |\vec{w}|$$

wenn

$$|r \vec{v} + s \vec{w}|^2 = 0$$

also

$$r \vec{v} + s \vec{w} = 0$$

oder

$$\vec{v} = -\frac{s}{r} \cdot \vec{w}$$

d.h. \vec{v} und \vec{w} sind kollinear.

Falls $-\frac{s}{r} > 0$, so sind \vec{v} und \vec{w} parallel, und man rechnet unmittelbar nach, dass in diesem Fall

$$\langle \vec{v}, \vec{w} \rangle = |\vec{v}| \cdot |\vec{w}|$$

während im Fall $-\frac{s}{r} < 0$ die Vektoren \vec{v} und \vec{w} anti-parallel sind und

$$\langle \vec{v}, \vec{w} \rangle = -|\vec{v}| \cdot |\vec{w}|$$

Aus den Regeln ergeben sich sofort einige interessante Konsequenzen für die Länge von Vektoren:

Satz 3.2.5. Für Vektoren \vec{v}, \vec{w} und einen Skalar $r \in \mathbb{R}$ gilt:

1. $|r \cdot \vec{v}| = |r| \cdot |\vec{v}|$.
2. $|\vec{v} + \vec{w}| \leq |\vec{v}| + |\vec{w}|$.

Beweis: In der Tat gilt

$$\begin{aligned} |r \cdot \vec{v}|^2 &= \langle r \cdot \vec{v}, r \cdot \vec{v} \rangle \\ &= r^2 \cdot \langle \vec{v}, \vec{v} \rangle \\ &= r^2 \cdot |\vec{v}|^2 \end{aligned}$$

und durch Ziehen der Quadratwurzel folgt die erste Behauptung.

Zum Nachweis der zweiten Aussage gehen wir ähnlich vor:

$$\begin{aligned}
 |\vec{v} + \vec{w}|^2 &= \langle \vec{v} + \vec{w}, \vec{v} + \vec{w} \rangle \\
 &= \langle \vec{v}, \vec{v} \rangle + \langle \vec{v}, \vec{w} \rangle + \langle \vec{w}, \vec{v} \rangle + \langle \vec{w}, \vec{w} \rangle \\
 &= |\vec{v}|^2 + 2 \cdot \langle \vec{v}, \vec{w} \rangle + |\vec{w}|^2 \\
 &\leq |\vec{v}|^2 + 2 \cdot |\vec{v}| \cdot |\vec{w}| + |\vec{w}|^2 \\
 &= (|\vec{v}| + |\vec{w}|)^2
 \end{aligned}$$

wobei wir für die Ungleichung den Satz von Cauchy–Schwarz benutzt haben. Wiederum erhalten wir die Behauptung durch Ziehen der Quadratwurzel.

Bemerkung 3.2.8. Aus der ersten Bedingung des Satzes von Cauchy–Schwarz folgt wieder, dass es ein $\varphi \in [0, \pi]$ gibt mit

$$|\langle \vec{v}, \vec{w} \rangle| = \cos(\varphi) |\vec{v}| \cdot |\vec{w}|$$

Dieses φ heißt der **Winkel** zwischen \vec{v} und \vec{w} . In der Ebene handelt es sich dabei in der Tat um den Winkel zwischen den beiden Vektoren \vec{v} und \vec{w} .

Definition 3.2.7. Zwei n -dimensionale Vektoren \vec{v} und \vec{w} heißen **orthogonal**, wenn gilt

$$\langle \vec{v}, \vec{w} \rangle = 0$$

Wir sagen in diesem Fall auch, dass \vec{v} senkrecht auf \vec{w} steht und schreiben $\vec{v} \perp \vec{w}$.

Beispiel 3.2.26. Die Vektoren $\vec{v} = \begin{pmatrix} 1 \\ 2 \\ -3 \\ 1 \end{pmatrix}$ und $\vec{w} = \begin{pmatrix} 2 \\ 4 \\ 4 \\ 2 \end{pmatrix}$ sind orthogonal.

Beispiel 3.2.27. Die Vektoren $\vec{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$, $\vec{e}_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}$, \dots , $\vec{e}_n = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ 0 \end{pmatrix}$ sind paarweise orthogonal, d.h. für $i \neq j$ gilt

$$\langle \vec{e}_i, \vec{e}_j \rangle = 0$$

Definition 3.2.8. Ist $V \subseteq \mathbb{R}^n$ ein Untervektorraum, so heißt die Menge

$$V^\perp = \{ \vec{u} \in \mathbb{R}^n \mid \langle \vec{u}, \vec{v} \rangle = 0 \text{ für alle } \vec{v} \in V \}$$

das **orthogonale Komplement** von V .

Bemerkung 3.2.9. Ist $V \subseteq \mathbb{R}^n$ ein Untervektorraum der Dimension l , so ist $V^\perp \subseteq \mathbb{R}^n$ ein Untervektorraum der Dimension $n - l$.

Beweis: Wir weisen die Untervektorraumaxiome für V^\perp nach:

1. Sicherlich ist $\vec{0} \in V^\perp$, und damit ist V^\perp nicht leer.

2. Sind $\vec{u}_1, \vec{u}_2 \in V^\perp$, und ist $\vec{v} \in V$ beliebig, so gilt

$$\langle \vec{u}_1 + \vec{u}_2, \vec{v} \rangle = \langle \vec{u}_1, \vec{v} \rangle + \langle \vec{u}_2, \vec{v} \rangle = 0$$

und damit ist $\vec{u}_1 + \vec{u}_2 \in V^\perp$.

3. Ist $\vec{u} \in V^\perp$ und $r \in \mathbb{R}$, und ist $\vec{v} \in V$ beliebig, so gilt

$$\langle r \cdot \vec{u}, \vec{v} \rangle = r \cdot \langle \vec{u}, \vec{v} \rangle = 0$$

und damit ist $r \cdot \vec{u} \in V^\perp$.

Also ist V^\perp ein Untervektorraum von \mathbb{R}^n . Die Dimensionsformel ergibt sich aus dem Gram–Schmidt Orthonormalisierungsverfahren, dass wir in Satz 3.2.6 behandeln werden (vergleiche auch Aufgabe 3.2.11).

Mit dem Begriff der Basis haben wir ein gutes Mittel kennengelernt, um Untervektorräume $V \subseteq \mathbb{R}^n$ zu beschreiben. Doch auch Basen sind nicht alle gleich gut geeignet.

Beispiel 3.2.28. Eine Basis der \mathbb{R}^3 ist die Standardbasis

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

eine weitere Basis des \mathbb{R}^3 ist

$$\vec{v}_1 = \begin{pmatrix} 5 \\ 4 \\ 4 \end{pmatrix}, \quad \vec{v}_2 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad \vec{v}_3 = \frac{1}{10} \cdot \begin{pmatrix} 4 \\ 4 \\ 5 \end{pmatrix}$$

wie man sich leicht überzeugt. Für viele Anwendungen ist die erste Basis besser geeignet als die zweite. So sind etwa in der ersten Basis alle Vektoren normiert, haben also die Länge 1. In der zweiten Basis sind die Vektoren nicht nur nicht normiert, sie sind auch noch alle unterschiedlich lang. Ein weiterer entscheidender Unterschied fällt sofort ins Auge, wenn wir uns die Basen graphisch vorstellen. Für die Standardbasis des \mathbb{R}^3 erhalten wir dann folgendes Bild

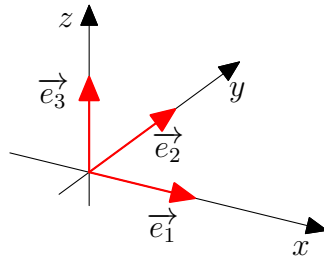


Abbildung 3.1: Standardbasis

und für die zweite Basis

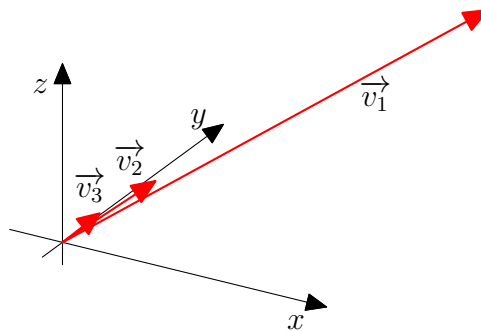


Abbildung 3.2: andere Basis

Wir sehen also, dass die Vektoren der ersten Basis jeweils senkrecht aufeinanderstehen, während die Vektoren der zweiten Basis relativ eng beieinander liegen und alle in einen Quadranten des Raums zeigen. Solche Basen sind vor

allen für numerische Berechnungen und Näherungen sehr schlecht geeignet, da sie dazu tendieren, Messfehler zu verstärken.

Dieses Beispiel motiviert die folgende Definition:

Definition 3.2.9. Eine Basis $\vec{v}_1, \dots, \vec{v}_m$ eines Untervektorraums $U \subseteq \mathbb{R}^n$ heißt **Orthonormalbasis** von U oder kurz **ONB** von U , wenn gilt:

- $|\vec{v}_i| = 1$ für $i = 1, \dots, m$, d.h. alle Vektoren der Basis sind normiert.
- $\langle \vec{v}_i, \vec{v}_j \rangle = 0$ für $i \neq j$, d.h. die Vektoren stehen paarweise senkrecht aufeinander.

Beispiel 3.2.29. Die Standardbasis $\vec{e}_1, \vec{e}_2, \vec{e}_3$ des \mathbb{R}^3 ist eine ONB von \mathbb{R}^3 , die Basis $\vec{v}_1, \vec{v}_2, \vec{v}_3$ aus Beispiel 3.2.28 ist es nicht.

Beispiel 3.2.30. Die Standardbasis $\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n$ des \mathbb{R}^n ist eine ONB.

Ganz offensichtlich sind Orthonormalbasen anderen Basen vorzuziehen. Damit stellt sich die Frage, wann solche existieren, und wie wir sie finden können. Antwort darauf liefert uns das Orthonormalisierungsverfahren von Gram-Schmidt und der folgende

Satz 3.2.6. *Jeder Untervektorraum U des \mathbb{R}^n besitzt eine Orthonormalbasis.*

Beweis: Der Beweis liefert und gleichzeitig einen Algorithmus, mit dem wir aus einer beliebigen Basis von U eine Orthonormalbasis von U machen können.

Wir betrachten dazu eine beliebige Basis $\vec{u}_1, \vec{u}_2, \dots, \vec{u}_m$ von U und wandeln diese Basis in zwei Schritten in eine Orthonormalbasis um. Zunächst bilden wir die Vektoren:

$$\begin{aligned}
 \vec{v}_1 &= \vec{u}_1 \\
 \vec{v}_2 &= \vec{u}_2 - \frac{\langle \vec{u}_2, \vec{v}_1 \rangle}{\langle \vec{v}_1, \vec{v}_1 \rangle} \cdot \vec{v}_1 \\
 \vec{v}_3 &= \vec{u}_3 - \frac{\langle \vec{u}_3, \vec{v}_1 \rangle}{\langle \vec{v}_1, \vec{v}_1 \rangle} \cdot \vec{v}_1 - \frac{\langle \vec{u}_3, \vec{v}_2 \rangle}{\langle \vec{v}_2, \vec{v}_2 \rangle} \cdot \vec{v}_2 \\
 &\vdots \\
 \vec{v}_m &= \vec{u}_m - \frac{\langle \vec{u}_m, \vec{v}_1 \rangle}{\langle \vec{v}_1, \vec{v}_1 \rangle} \cdot \vec{v}_1 - \frac{\langle \vec{u}_m, \vec{v}_2 \rangle}{\langle \vec{v}_2, \vec{v}_2 \rangle} \cdot \vec{v}_2 - \dots - \frac{\langle \vec{u}_m, \vec{v}_{m-1} \rangle}{\langle \vec{v}_{m-1}, \vec{v}_{m-1} \rangle} \cdot \vec{v}_{m-1}
 \end{aligned}$$

dh. wir übernehmen \vec{u}_1 unverändert als \vec{v}_1 , wir ziehen von \vec{u}_2 seinen Anteil "in Richtung \vec{v}_1 " ab usw. Dann gilt:

Die Vektoren $\vec{v}_1, \dots, \vec{v}_m$ bilden eine Basis von U , bestehend aus Vektoren, die paarweise senkrecht aufeinander stehen.

Der Nachweis, dass $\langle \vec{v}_i, \vec{v}_j \rangle = 0$ für $i \neq j$ ist eine einfache aber sehr langwierige Rechnung, die wir dem Leser überlassen. Um zu zeigen, dass $\vec{v}_1, \dots, \vec{v}_m$ den Unterraum U erzeugen, reicht es zu zeigen, dass sich die ursprünglichen Basisvektoren $\vec{u}_1, \vec{u}_2, \dots, \vec{u}_m$ damit darstellen lassen. Das ist aber klar, da sich ja jede der definierenden Gleichungen für ein \vec{v}_i direkt nach \vec{u}_i auflösen lässt. Die lineare Unabhängigkeit der Vektoren folgt etwa aus ihrer paarweisen Orthogonalität (siehe auch Aufgabe 3.2.10). Es ist aber auch nicht schwer, sie direkt nachzurechnen.

Wir haben also nun eine Basis von U gefunden, die aus Vektoren besteht, die paarweise senkrecht aufeinander stehen. Um eine Orthonormalbasis zu finden, reicht es also, diese Basis zu normieren:

$$\begin{aligned} \vec{w}_1 &= \frac{\vec{v}_1}{|\vec{v}_1|} \\ \vec{w}_2 &= \frac{\vec{v}_2}{|\vec{v}_2|} \\ &\vdots \\ \vec{w}_m &= \frac{\vec{v}_m}{|\vec{v}_m|} \end{aligned}$$

Es ist klar, dass die Vektoren $\vec{w}_1, \dots, \vec{w}_m$ normiert sind. Ebenfalls klar ist, dass sich durch die Normierung nichts daran ändert, dass sie paarweise orthogonal sind und dass sich dadurch auch nichts an ihrer Eigenschaft, Basis von U zu sein, ändert.

Beispiel 3.2.31. Wir betrachten den Untervektorraum $U \subseteq \mathbb{R}^3$, der von den Vektoren

$$\vec{u}_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \quad \vec{u}_2 = \begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix}$$

erzeugt wird. Zunächst überzeugen wir uns, dass diese beiden Vektoren auch eine Basis von U bilden. Das ist klar, denn keiner ist ein Vielfaches des anderen.

Orthogonalisierung:

$$\begin{aligned}\vec{v}_1 &= \vec{u}_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \\ \vec{v}_2 &= \vec{u}_2 - \frac{\langle \vec{u}_2, \vec{v}_1 \rangle}{\langle \vec{v}_1, \vec{v}_1 \rangle} \cdot \vec{v}_1 = \begin{pmatrix} \frac{8}{14} \\ \frac{2}{14} \\ -\frac{4}{14} \end{pmatrix}\end{aligned}$$

Normalisierung:

$$\begin{aligned}\vec{w}_1 &= \frac{\vec{v}_1}{|\vec{v}_1|} = \begin{pmatrix} \frac{1}{\sqrt{14}} \\ \frac{2}{\sqrt{14}} \\ \frac{3}{\sqrt{14}} \end{pmatrix} \\ \vec{w}_2 &= \frac{\vec{v}_2}{|\vec{v}_2|} = \begin{pmatrix} \frac{8}{\sqrt{84}} \\ \frac{2}{\sqrt{84}} \\ -\frac{4}{\sqrt{84}} \end{pmatrix}\end{aligned}$$

und wir haben die gewünschte Orthonormalbasis gefunden.

Aufgabe 3.2.1. Zeigen Sie, dass die Menge $V = \left\{ \vec{v} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{pmatrix} \mid 2v_1 + 3v_2 - 4v_3 + v_4 = 0 \right\}$ ein Untervektorraum von \mathbb{R}^4 ist.

Aufgabe 3.2.2. Zeigen Sie, dass die Menge $V = \left\{ \vec{v} = \begin{pmatrix} t \\ t^2 \\ t^3 \\ t^4 \end{pmatrix} \mid t \in \mathbb{R} \right\}$ kein Untervektorraum von \mathbb{R}^4 ist.

Aufgabe 3.2.3. Überprüfen Sie, ob die Menge $V = \left\{ \vec{v} = \begin{pmatrix} r \\ s \\ r+s \\ r-s \end{pmatrix} \mid r, s \in \mathbb{R} \right\}$ ein Untervektorraum von \mathbb{R}^4 ist.

Aufgabe 3.2.4. Zeigen Sie, dass die Vektoren $\vec{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 2 \\ 0 \end{pmatrix}$, $\vec{v}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 2 \end{pmatrix}$ und $\vec{v}_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$ linear unabhängig sind.

Aufgabe 3.2.5. Zeigen Sie, dass die Vektoren $\vec{v}_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}$, $\vec{v}_2 = \begin{pmatrix} 4 \\ 2 \\ 0 \\ 0 \end{pmatrix}$ und $\vec{v}_3 = \begin{pmatrix} 3 \\ 3 \\ 3 \\ 4 \end{pmatrix}$ linear abhängig sind.

Aufgabe 3.2.6. Zeigen Sie, dass die Vektoren $\vec{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$, $\vec{v}_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$ und $\vec{v}_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ eine Basis von \mathbb{R}^3 bilden.

Aufgabe 3.2.7. Gegeben seien die beiden Vektoren

$$\vec{v}_1 = \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix}, \quad \vec{v}_2 = \begin{pmatrix} -3 \\ 0 \\ 1 \end{pmatrix}$$

Zeigen Sie

$$\langle \{\vec{v}_1, \vec{v}_2\} \rangle = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^n \mid x_1 + 2x_2 + 3x_3 = 0 \right\}$$

Bilden \vec{v}_1 und \vec{v}_2 auch eine Basis dieses Untervektorraums?

Aufgabe 3.2.8. Eine Basis von \mathbb{R}^3 ist

$$\vec{v}_1 = \begin{pmatrix} 3 \\ 4 \\ 5 \end{pmatrix}, \quad \vec{v}_2 = \begin{pmatrix} 3 \\ 5 \\ 6 \end{pmatrix}, \quad \vec{v}_3 = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}$$

(vergleiche Beispiel 3.2.28). Orthonormalisieren Sie diese Basis.

Aufgabe 3.2.9. Zeigen Sie, dass

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

eine Basis von \mathbb{R}^3 ist und orthonormalisieren Sie diese Basis.

Aufgabe 3.2.10. Zeigen Sie: Stehen zwei vom Nullvektor verschiedene Vektoren \vec{v} und \vec{w} senkrecht aufeinander, so sind sie linear unabhängig. Verallgemeinern Sie diese Aussage auf n Vektoren, die paarweise orthogonal sind.

Aufgabe 3.2.11. Weisenn Sie die Dimensionsformel aus Bemerkung 3.2.9 nach.

3.3 Allgemeine reelle Vektorräume

Im vergangenen Abschnitt haben wir speziell den reellen Vektorraum \mathbb{R}^n untersucht, da diese für die Anwendungen die wichtigsten reellen Vektorräume bilden. Es gibt jedoch auch noch viele andere Mengen, die sich mithilfe der Vektorraumtheorie studieren und strukturieren lassen. Deshalb wollen wir uns in diesem Abschnitt mit allgemeinen reellen Vektorräumen beschäftigen.

Beispiel 3.3.1. Wir betrachten die Menge V aller Abbildungen

$$f : [0, 1] \longrightarrow \mathbb{R}$$

Für zwei Abbildungen $f, g \in V$ definieren wir die Abbildung $f + g$ durch

$$(f + g)(x) = f(x) + g(x) \quad \text{für alle } x \in [0, 1]$$

und für eine Abbildung $f \in V$ und ein $r \in \mathbb{R}$ erklären wir $r \cdot f$ durch

$$(r \cdot f)(x) = r \cdot f(x) \quad \text{für alle } x \in [0, 1]$$

Dann ist $(V, +, \cdot)$ ein reeller Vektorraum, wie man leicht nachrechnet.

Für diesen Vektorraum schreibt man auch $\text{Abb}([0, 1], \mathbb{R})$ oder $[0, 1]^{\mathbb{R}}$ und nennt ihn den Vektorraum der (reellwertigen) Abbildungen auf dem Intervall $[0, 1]$.

Beispiel 3.3.2. Eine Abbildung

$$f : [a, b] \longrightarrow \mathbb{R}$$

auf einem Intervall $[a, b]$ heißt Polynomabbildung, wenn sich f in der Form

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$$

für ein geeignetes n und mit geeigneten $a_0, a_1, \dots, a_n \in \mathbb{R}$ schreiben lässt, vergleiche Beispiel ???. Beispiele hierfür sind etwa $f(x) = x^2 + 2$ oder $f(x) = 6x^3 - 3x + 4$.

Wir betrachten die Menge U der Polynomabbildungen

$$f : [0, 1] \longrightarrow \mathbb{R}$$

Wie in Beispiel 3.3.1 definieren wir für zwei polynome Abbildungen $f, g \in U$ die Abbildung $f + g$ durch

$$(f + g)(x) = f(x) + g(x) \quad \text{für alle } x \in [0, 1]$$

und für eine polynome Abbildung $f \in U$ und ein $r \in \mathbb{R}$ erklären wir $r \cdot f$ durch

$$(r \cdot f)(x) = r \cdot f(x) \quad \text{für alle } x \in [0, 1]$$

Dann ist zunächst zu beachten, dass sowohl $f + g$ als auch $r \cdot f$ wieder polynomiale Abbildungen sind. So ist etwa $(x^2 + 2) + (6x^3 - 3x + 4) = 6x^3 + x^2 - 2x + 6$ oder $3(6x^3 - 3x + 4) = 18x^3 - 9x + 12$.

Hierfür rechnen wir sofort nach, dass $U \subseteq \text{Abb}([0, 1], \mathbb{R})$ ein Untervektorraum ist, und damit insbesondere auch selbst ein reeller Vektorraum.

Für diesen Vektorraum schreiben wir $\text{Pol}([0, 1], \mathbb{R})$ und nennen ihn den Vektorraum der (reellwertigen) polynomalen Abbildungen auf dem Intervall $[0, 1]$. (Genauso gut hätten wir natürlich jedes andere Intervall $[a, b]$ nehmen können.)

Wir betrachten nun einen Vektorraum V und Vektoren $\vec{w}, \vec{v}_1, \dots, \vec{v}_n \in V$.

Definition 3.3.1. \vec{w} heißt Linearkombination von $\vec{v}_1, \dots, \vec{v}_n$, wenn es $r_1, \dots, r_n \in \mathbb{R}$ gibt mit

$$\vec{w} = r_1 \cdot \vec{v}_1 + \dots + r_n \cdot \vec{v}_n$$

Beispiel 3.3.3. Sei $V = \mathbb{R}^n$ und seien $\vec{w} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$, $\vec{v}_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ und $\vec{v}_2 = \begin{pmatrix} 8 \\ 6 \\ 4 \\ 2 \end{pmatrix}$. Dann ist \vec{w} eine Linearkombination von \vec{v}_1 und \vec{v}_2 , denn es gilt:

$$\vec{w} = \frac{1}{5} \cdot \vec{v}_1 + \frac{1}{10} \cdot \vec{v}_2$$

Beispiel 3.3.4. Sei $V = \mathbb{R}^n$ und seien $\vec{w} = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}$, $\vec{v}_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ und $\vec{v}_2 = \begin{pmatrix} 8 \\ 6 \\ 4 \\ 2 \end{pmatrix}$. Dann ist \vec{w} keine Linearkombination von \vec{v}_1 und \vec{v}_2 . Das ist so unmittelbar nicht sofort einzusehen. Wir werden diese Beispiel später im Zusammenhang mit linearen Gleichungssystemen noch einmal untersuchen und überlassen es hier dem Leser, zu zeigen, dass es keine reellen Zahlen r, s gibt mit

$$\vec{w} = r \cdot \vec{v}_1 + s \cdot \vec{v}_2$$

Beispiel 3.3.5. Sei $V = \text{Pol}([0, 1], \mathbb{R})$ der Vektorraum aus Beispiel 3.3.2 und seien $f(x) = 2x^2 - 4$, $f_0(x) = 1$, $f_1(x) = x$, $f_2(x) = x^2$ und $f_3(x) = x^3$. Dann ist f Linearkombination von f_0, f_1, f_2, f_3 , denn es gilt

$$f(x) = -4 \cdot f_0(x) + 2 \cdot f_2(x) = -4 \cdot f_0(x) + 0 \cdot f_1(x) + 2 \cdot f_2(x) + 0 \cdot f_3(x)$$

Das Konzept der Linearkombination lässt sich auch auf unendliche Familien von Vektoren ausdehnen. Dazu sei Λ eine Indexmenge, etwa $\Lambda = \{1, \dots, n\}$ oder $\Lambda = \mathbb{N}$ und es sei $\{\vec{v}_\lambda\}_{\lambda \in \Lambda}$ eine Familie von Elementen von V (d.h. für jedes $\lambda \in \Lambda$ ist ein Element $\vec{v}_\lambda \in V$ gegeben). Für $\Lambda = \{1, \dots, n\}$ ist das einfach eine Menge $\{\vec{v}_1, \dots, \vec{v}_n\}$ von n Elementen aus V . Für $\Lambda = \mathbb{N} = \{0, 1, 2, \dots\}$ ist eine Menge $\{\vec{v}_0, \vec{v}_1, \vec{v}_2, \dots\}$ von unendlich vielen Elementen aus V .

Definition 3.3.2. Ein Vektor \vec{w} heißt **Linearkombination** von $\{\vec{v}_\lambda\}_{\lambda \in \Lambda}$, wenn es endlich viele $\lambda_1, \dots, \lambda_n \in \Lambda$ und $r_1, \dots, r_n \in \mathbb{R}$ gibt mit

$$\vec{w} = r_1 \cdot \vec{v}_{\lambda_1} + \dots + r_n \cdot \vec{v}_{\lambda_n}$$

Bemerkung 3.3.1. Ist Λ endlich, $\lambda = \{1, \dots, n\}$, so stimmt diese Definition mit der aus Definition 3.3.1 überein.

Beispiel 3.3.6. Sei $V = \text{Pol}([0, 1], \mathbb{R})$ und sei $\Lambda = \mathbb{N} = \{0, 1, 2, \dots\}$. Definiere f_n durch

$$f_n(x) = x^n \quad \text{für alle } n \in \Lambda = \mathbb{N}$$

(also $f_0(x) = 1$, $f_1(x) = x$, $f_2(x) = x^2, \dots$). Ferner sei $f \in V$ die Polynomfunktion mit $f(x) = 3x^5 - 2x^2 + 7$. Dann ist f eine Linearkombination der $(f_\lambda)_{\lambda \in \Lambda}$.

Dazu betrachte die endliche Teilmenge von Λ mit den Elementen $\lambda_1 = 0$, $\lambda_2 = 2$ und $\lambda_3 = 5$. Dann gilt hierfür

$$f(x) = 7 \cdot f_{\lambda_1}(x) + (-2) \cdot f_{\lambda_2}(x) + 3 \cdot f_{\lambda_3}(x)$$

Allgemein ist jede beliebige Polynomfunktion $f : [0, 1] \rightarrow \mathbb{R}$ eine Linearkombination der $\{f_\lambda\}_{\lambda \in \Lambda}$. Schreibt sich dazu $f(x)$ etwa als

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0,$$

so betrachte die endliche Teilmenge $\{0, 1, \dots, n\} \subseteq \Lambda$ und erhalte mit

$$f(x) = a_n \cdot f_n(x) + a_{n-1} \cdot f_{n-1}(x) + \dots + a_1 \cdot f_1(x) + a_0 \cdot f_0(x)$$

eine Darstellung wie gewünscht.

Die Exponentialfunktion e^x , eingeschränkt auf das Intervall $[0, 1]$ hat eine Beschreibung durch die sogenannte Exponentialreihe

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \frac{x^4}{24} + \dots$$

hat also auch eine Summendarstellung mit den f_λ . Trotzdem ist das keine Darstellung der Exponentialfunktion als Linearkombination der f_λ , da unendlich viele der f_λ benötigt werden.

Man benötigt etwas Analysis um zu zeigen, dass es auch keine andere Darstellung der Exponentialfunktion als Linearkombination der f_λ gibt, die Exponentialfunktion ist also keine Polynomfunktion.

Definition 3.3.3. Eine Familie $\{\vec{v}_\lambda\}_{\lambda \in \Lambda}$ von Vektoren in V heißt ein **Erzeugendensystem** von V , wenn sich jedes Element $\vec{v} \in V$ als Linearkombination von $\{\vec{v}_\lambda\}_{\lambda \in \Lambda}$ schreiben lässt.

Beispiel 3.3.7. sei $V = \left\{ \begin{pmatrix} x \\ y \\ 0 \end{pmatrix} \mid x, y \in \mathbb{R} \right\}$. Dann ist V ein Vektorraum (als Untervektorraum von \mathbb{R}^3) und

$$\vec{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \vec{v}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

ist ein Erzeugendensystem von V . Ebenso ist

$$\vec{w}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad \vec{w}_2 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \quad \vec{w}_3 = \begin{pmatrix} 3 \\ 2 \\ 0 \end{pmatrix}$$

ein Erzeugendensystem von V .

Beispiel 3.3.8. Ist $U = \text{Pol}([0, 1], \mathbb{R})$ der Vektorraum der polynomialen Funktionen aus Beispiel 3.3.2 und ist $(f_n)_{n \in \mathbb{N}}$ die Familie polynomialer Funktionen aus Beispiel 3.3.6, so ist $(f_n)_{n \in \mathbb{N}}$ ein Erzeugendensystem von U , wie wir in Beispiel 3.3.6 nachgerechnet haben.

Beispiel 3.3.9. Ist $V = \text{Abb}([0, 1], \mathbb{R})$ der Vektorraum aus Beispiel 3.3.1 und ist $(f_n)_{n \in \mathbb{N}}$ die Familie (polynomialer) Abbildungen aus Beispiel 3.3.6, so ist $(f_n)_{n \in \mathbb{N}}$ kein Erzeugendensystem von V , da die Exponentialfunktion in V liegt, aber keine Linearkombination der $(f_n)_{n \in \mathbb{N}}$ ist.

Definition 3.3.4. Eine Familie $\{\vec{v}_\lambda\}_{\lambda \in \Lambda}$ von Vektoren in V heißt **linear unabhängig**, wenn gilt:

Ist $\Lambda_0 = \{\lambda_1, \dots, \lambda_n\} \subseteq \Lambda$ eine beliebige endliche Teilmenge von Λ und sind r_1, \dots, r_n reelle Zahlen mit

$$r_1 \cdot \vec{v}_{\lambda_1} + r_2 \cdot \vec{v}_{\lambda_2} + \dots + r_n \cdot \vec{v}_{\lambda_n} = \vec{0}$$

so muss schon gelten:

$$r_1 = r_2 = \dots = r_n = 0$$

Andernfalls heißt die Familie $\{\vec{v}_\lambda\}_{\lambda \in \Lambda}$ linear abhängig.

Bemerkung 3.3.2. Ist Λ insgesamt schon eine endlich Menge, so reicht es, die Bedingung aus der Definition nur für $\Lambda_0 = \Lambda$ zu prüfen. Wir erhalten dann wieder die Definition 3.2.2, die wir im letzten Abschnitt für Vektoren im \mathbb{R}^n gemacht haben.

Bemerkung 3.3.3. Ist $\{\vec{v}_\lambda\}_{\lambda \in \Lambda}$ eine Familie linear unabhängiger Vektoren und ist $\Gamma \subseteq \Lambda$ eine Teilmenge von Λ , so ist auch $\{\vec{v}_\lambda\}_{\lambda \in \Gamma}$ linear unabhängig.

Beispiel 3.3.10. Die Vektoren $\vec{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ und $\vec{v}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$ sind linear unabhängig, wie man sofort sieht.

Beispiel 3.3.11. Die Vektoren $\vec{v}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ und $\vec{v}_2 = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$ und $\vec{v}_3 = \begin{pmatrix} 4 \\ 3 \\ 4 \end{pmatrix}$ sind linear abhängig, denn

$$7 \cdot \vec{v}_1 + \vec{v}_2 + (-2) \cdot \vec{v}_3 = \vec{0}$$

Beispiel 3.3.12. Alle Beispiele linear unabhängiger Vektoren aus dem letzten Abschnitt sind auch linear unabhängig in der neuen Definition; alle Beispiele linear abhängiger Vektoren aus dem letzten Abschnitt sind auch linear abhängig im Sinne der neuen Definition.

Beispiel 3.3.13. Ist $U = \text{Pol}([0, 1], \mathbb{R})$ der Vektorraum der polynomialen Funktionen aus Beispiel 3.3.2 und ist $(f_n)_{n \in \mathbb{N}}$ die Familie polynomialer Funktionen aus Beispiel 3.3.6, so ist $(f_n)_{n \in \mathbb{N}}$ eine Familie linear unabhängiger Vektoren. Dies nachzurechnen erfordert ein wenig Analysis: Sei $\{i_1, \dots, i_n\} \subseteq \mathbb{N}$ eine endliche Teilmenge, wobei wir annehmen können, dass $i_1 < i_2 < \dots < i_n$, und seien $r_1, \dots, r_n \in \mathbb{R}$ mit

$$0 = r_1 \cdot f_{i_1}(x) + \dots + r_n \cdot f_{i_n}(x)$$

Wir differenzieren nun beide Seiten wiederholt und beachten dabei, dass $f_{i_l}^{(i_n)}(x) = 0$ für $l < n$, da $i_l < i_n$, und dass $f_{i_n}^{(i_n)}(x) = i_n! (= i_n(i_n - 1) \dots 1)$. Damit erhalten wir also durch i_n -faches Differenzieren

$$0 = r_n i_n!$$

woraus folgt, dass $r_n = 0$ gelten muss. Damit haben wir schon eine Linearkombination

$$0 = r_1 \cdot f_{i_1}(x) + \dots + r_{n-1} \cdot f_{i_{n-1}}(x)$$

mit der wir genauso vorgehen und durch i_{n-1} -faches Ableiten zeigen, dass auch $r_{n-1} = 0$. Nach n Schritten haben wir gezeigt, dass $r_1 = r_2 = \dots = r_n = 0$.

Beispiel 3.3.14. Ist wieder $U = \text{Pol}([0, 1], \mathbb{R})$ der Vektorraum der polynomialen Funktionen aus Beispiel 3.3.2 und ist $(g_n)_{n \in \mathbb{N}}$ die Familie polynomialer Funktionen mit

$$\begin{aligned} g_0(x) &= 2x^2 + 3x^4 \\ g_n(x) &= x^{2n} \quad \text{für } n \geq 1 \end{aligned}$$

so sind die $(g_n)_{n \in \mathbb{N}}$ linear abhängig, denn $g_0(x) - 2g_1(x) - 3g_2(x) = 0$. Definieren wir stattdessen

$$g_0(x) = 1 + 2x^2 + 3x^4$$

so ist die Familie $(g_n)_{n \in \mathbb{N}}$ linear unabhängig. Das nachzurechnen überlassen bleibt Ihnen als Übung überlassen.

Definition 3.3.5. Eine Familie $\{\vec{v}_\lambda\}_{\lambda \in \Lambda}$ von Vektoren in V heißt **Basis** von V , wenn die Vektoren $\{\vec{v}_\lambda\}_{\lambda \in \Lambda}$ sowohl linear unabhängig als auch ein Erzeugendensystem von V ist.

Beispiel 3.3.15. Die Vektoren $\vec{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$, $\vec{e}_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \dots, \vec{e}_n = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$ bilden eine Basis von \mathbb{R}^n wie wir schon im Beispiel 3.2.19 aus dem letzten Abschnitt gesehen haben.

Beispiel 3.3.16. Die Vektoren $\vec{v}_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 3 \end{pmatrix}$, $\vec{v}_2 = \begin{pmatrix} 0 \\ 1 \\ 2 \\ 3 \end{pmatrix}$, $\vec{v}_3 = \begin{pmatrix} 0 \\ 1 \\ 2 \\ 2 \end{pmatrix}$ und $\vec{v}_4 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$ bilden eine Basis von \mathbb{R}^4 , wie man leicht nachrechnet.

Beispiel 3.3.17. Die Vektoren $\vec{w}_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}$, $\vec{w}_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}$, $\vec{w}_3 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}$, $\vec{w}_4 = \begin{pmatrix} 2 \\ 0 \\ 1 \\ 0 \end{pmatrix}$ und $\vec{w}_5 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 2 \end{pmatrix}$ bilden keine eine Basis von \mathbb{R}^4 . Diese Vektoren erzeugen zwar den \mathbb{R}^4 , aber sie sind nicht linear unabhängig, da etwa

$$\vec{w}_1 - 5 \cdot \vec{w}_2 + 2 \cdot \vec{w}_3 - 2 \cdot \vec{w}_4 = \vec{0}$$

Beispiel 3.3.18. Die Vektoren $\vec{u}_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}$, $\vec{u}_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}$ und $\vec{u}_3 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}$ sind keine Basis von \mathbb{R}^4 . Sie sind zwar linear unabhängig, erzeugen aber \mathbb{R}^4 nicht. So ist etwa $\vec{v} = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 5 \end{pmatrix}$ keine Linearkombination von $\vec{u}_1, \vec{u}_2, \vec{u}_3$.

Beispiel 3.3.19. Sei $U = \text{Pol}([0, 1], \mathbb{R})$ der Vektorraum der polynomialen Funktionen aus Beispiel 3.3.2 und sei $(f_n)_{n \in \mathbb{N}}$ die Familie polynomialer Funktionen aus Beispiel 3.3.6. Dann ist $(f_n)_{n \in \mathbb{N}}$ eine Basis von U , wie aus den Beispielen 3.3.8 und 3.3.13 folgt.

Ein wichtiges allgemeines Ergebnis über Basen liefert der folgende Satz (das im Fall unendlicher Indexfamilien auf dem sogenannten Auswahlaxiom beruht):

Satz 3.3.1. Für eine Familie $\{\vec{v}_\lambda\}_{\lambda \in \Lambda}$ von Vektoren in V sind äquivalent:

1. $\{\vec{v}_\lambda\}_{\lambda \in \Lambda}$ ist eine Basis von V .
2. $\{\vec{v}_\lambda\}_{\lambda \in \Lambda}$ ist ein unverkürzbares Erzeugendensystem von V , d.h. $\{\vec{v}_\lambda\}_{\lambda \in \Lambda}$ erzeugt V und für jede echte Teilmenge $\Lambda_0 \subset \Lambda$ gilt: $\{\vec{v}_\lambda\}_{\lambda \in \Lambda_0}$ erzeugt V nicht mehr.

3. $\{\vec{v}_\lambda\}_{\lambda \in \Lambda}$ ist eine maximale linear unabhängige Familie von Vektoren on V , d.h. nimmt man noch einen Vektor \vec{w} hinzu, so ist die resultierende Familie niemals mehr linear unabhängig.
4. $\{\vec{v}_\lambda\}_{\lambda \in \Lambda}$ ist ein Erzeugendensystem von V mit dem sich jedes Element von V auf genau eine Weise darstellen lässt.

Beweis: Wir zeigen $1. \implies 2. \implies 4. \implies 3. \implies 1.$ Damit sind alle Äquivalenzen gezeigt.

$1. \implies 2.:$ Angenommen wir könnten ein \vec{v}_{λ_0} weglassen und erhielten immer noch ein Erzeugendensystem $\{\vec{v}_\lambda\}_{\lambda \in \Lambda \setminus \{\lambda_0\}}$ von V . Dann könnten wir insbesondere \vec{v}_{λ_0} als Linearkombination dieses Erzeugendensystems schreiben:

$$\vec{v}_{\lambda_0} = r_1 \vec{v}_{\lambda_1} + r_2 \vec{v}_{\lambda_2} + \dots + r_n \vec{v}_{\lambda_n}$$

mit $\lambda_1, \dots, \lambda_n \in \Lambda \setminus \{\lambda_0\}$. Dann ist aber

$$\vec{v}_{\lambda_0} - r_1 \vec{v}_{\lambda_1} - r_2 \vec{v}_{\lambda_2} - \dots - r_n \vec{v}_{\lambda_n} = \vec{0}$$

im Widerspruch zur linearen Unabhängigkeit von $\{\vec{v}_\lambda\}_{\lambda \in \Lambda}$.

$2. \implies 4.:$ Angenommen, ein \vec{v} hat zwei Darstellungen,

$$\begin{aligned} \vec{v} &= r_1 \vec{v}_{\lambda_1} + \dots + r_n \vec{v}_{\lambda_n} \\ \vec{v} &= s_1 \vec{v}_{\mu_1} + \dots + s_n \vec{v}_{\mu_n} \end{aligned}$$

Dann erhalten wir durch Differenzbilden

$$\vec{0} = (r_1 \vec{v}_{\lambda_1} + \dots + r_n \vec{v}_{\lambda_n}) - (s_1 \vec{v}_{\mu_1} + \dots + s_n \vec{v}_{\mu_n})$$

also (nach eventueller Umbenennung und Umsortierung) eine Linearkombination

$$\vec{0} = t_1 \vec{v}_{\nu_1} + \dots + t_l \vec{v}_{\nu_l}$$

in der nach Voraussetzung mindestens einer der Koeffizienten t_i von Null verschieden ist. Wir können annehmen, dass das t_1 ist. Dann gilt aber

$$\vec{v}_{\nu_1} = -\frac{1}{t_1} (t_2 \vec{v}_{\nu_2} + \dots + t_l \vec{v}_{\nu_l})$$

und damit könnte \vec{v}_{ν_1} aus dem Erzeugendensystem entfernt werden und die verbleibenden vektoren wären immer noch erzeugend. Ein Widerspruch zur Annahme.

4. \implies 3.: Wenn sich jedes Element in V eindeutig mit $\{\vec{v}_\lambda\}_{\lambda \in \Lambda}$ darstellen lässt, dann gilt das insbesondere für den Nullvektor.

$$\vec{0} = r_1 \vec{v}_{\lambda_1} + \dots + r_n \vec{v}_{\lambda_n}$$

Da sich der Nullvektor aber in trivialer Weise

$$\vec{0} = 0 \cdot \vec{v}_{\lambda_1} + \dots + 0 \cdot \vec{v}_{\lambda_n}$$

darstellen lässt, folgt aus der Eindeutigkeit schon, dass $r_1 = \dots = r_n = 0$. Damit ist die Familie $\{\vec{v}_\lambda\}_{\lambda \in \Lambda}$ linear unabhängig. Da sie aber schon erzeugt, kann man kein weiteres Element aus V mehr dazu nehmen, ohne die lineare Unabhängigkeit zu zerstören. Ist \vec{w} ein beliebiger weiterer Vektor, so schreibt sich \vec{w} als

$$\vec{w} = r_1 \vec{v}_{\lambda_1} + \dots + r_n \vec{v}_{\lambda_n}$$

was zu einer nicht-trivialen Linearkombination

$$\vec{w} = \vec{w} - r_1 \vec{v}_{\lambda_1} - \dots - r_n \vec{v}_{\lambda_n}$$

so dass die Familie $\{\vec{v}_\lambda\}_{\lambda \in \Lambda} \cup \{\vec{w}\}$ nicht mehr linear unabhängig ist.

3. \implies 1.: Es bleibt noch zu zeigen, dass $\{\vec{v}_\lambda\}_{\lambda \in \Lambda}$ den Vektorraum V erzeugt. Nehmen wir dazu an, dass es ein \vec{v} gibt, dass keine Linearkombination von $\{\vec{v}_\lambda\}_{\lambda \in \Lambda}$ ist, so wäre auch die Familie $\{\vec{v}_\lambda\}_{\lambda \in \Lambda} \cup \{\vec{v}\}$ linear unabhängig: Klar ist, dass es keine Beziehung der Form

$$\vec{0} = r_1 \cdot \vec{v}_{\lambda_1} + \dots + r_n \cdot \vec{v}_{\lambda_n}$$

mit $\lambda_1, \dots, \lambda_n \in \Lambda$ geben kann, da ja $\{\vec{v}_\lambda\}_{\lambda \in \Lambda}$ eine Familie linear unabhängiger Vektoren ist. Es kann also höchstens eine Beziehung der Form

$$\vec{0} = r_1 \cdot \vec{v}_{\lambda_1} + \dots + r_n \cdot \vec{v}_{\lambda_n} + s \cdot \vec{v}$$

mit $s \neq 0$ geben. Aber dann lässt sich diese Beziehung auflösen zu

$$\vec{v} = -\frac{r_1}{s} \cdot \vec{v}_{\lambda_1} - \dots - \frac{r_n}{s} \cdot \vec{v}_{\lambda_n}$$

und \vec{v} wäre eine Linearkombination der $\{\vec{v}_\lambda\}_{\lambda \in \Lambda}$, im Widerspruch zur Wahl von \vec{v} .

Das wiederum ist ein Widerspruch dazu, dass $\{\vec{v}_\lambda\}_{\lambda \in \Lambda}$ ein nicht-verlängerbares System linear unabhängiger Vektoren ist.

Folgerung 3.3.2. Ist $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m\}$ ein Erzeugendensystem eines Vektorraums V , so enthält es eine Basis, d.h. es gibt $i_1, i_2, \dots, i_n \in \{1, \dots, m\}$, so dass $\{\vec{v}_{i_1}, \vec{v}_{i_2}, \dots, \vec{v}_{i_n}\}$ eine Basis von V ist.

Beweis: Das folgt sofort aus Satz 3.3.1, indem man aus dem Erzeugendensystem so lange Vektoren weglässt, bis es nicht mehr verkürzt werden kann.

Folgerung 3.3.3. Ist $\{\vec{v}_\lambda\}_{\lambda \in \Lambda}$ eine Familie linear unabhängiger Vektoren in V , so kann $\{\vec{v}_\lambda\}_{\lambda \in \Lambda}$ zu einer Basis von V ergänzt werden.

Beweis: Das folgt sofort aus Satz 3.3.1, indem man das System linear unabhängiger Vektoren so lange verlängert, bis es nicht maximal ist.

Hieraus erhält man nun unmittelbar

Folgerung 3.3.4. Jeder Vektorraum hat eine Basis.

Zur Erinnerung: Für eine Menge M bezeichnet $|M|$ die Mächtigkeit dieser Menge. Das kann eine endliche Zahl sein oder auch ∞ .

Definition 3.3.6. Ist $\{\vec{v}_\lambda\}_{\lambda \in \Lambda}$ eine Basis von V , so heißt $|\Lambda|$ die Länge der Basis.

Beispiel 3.3.20. Die Standardbasis $\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$ des \mathbb{R}^n hat die Länge n .

Beispiel 3.3.21. Die Basis $\{f_n\}_{n \in \mathbb{N}}$ des Vektorraums $U = \text{Pol}([0, 1], \mathbb{R})$ der polynomialen Funktionen aus Beispiel 3.3.19 hat die Länge ∞ .

Satz 3.3.5. Je zwei Basen $\{\vec{v}_\lambda\}_{\lambda \in \Lambda}$ und $\{\vec{w}_\gamma\}_{\gamma \in \Gamma}$ sind gleich lang, d.h. sie sind entweder beide unendlich oder beide endlich mit der selben Anzahl von Elementen.

Definition 3.3.7. Die Länge einer Basis von V heißt die **Dimension** des Vektorraums V und wird mit $\dim(V)$ bezeichnet.

Bemerkung 3.3.4. $\dim(V)$ ist eine endliche Zahl oder ∞ .

Der entscheidende Punkt im Beweis von Satz 3.3.5 ist der *Basisaustauschsatz von Steinitz*, den wir nur im endlichdimensionalen Fall formulieren:

Satz 3.3.6 (Basisaustauschsatz von Steinitz). *Ist $\{\vec{v}_1, \dots, \vec{v}_n\}$ eine Basis von V und sind $\{\vec{w}_1, \dots, \vec{w}_m\}$ linear unabhängige Vektoren, so gibt es $i_1, \dots, i_m \in \{1, \dots, n\}$ derart, dass die \vec{v}_{i_t} durch die \vec{w}_t ersetzt werden können und wieder eine Basis entsteht. In anderen Worten: Die Vektoren $\vec{u}_1, \dots, \vec{u}_m$ mit*

$$\vec{u}_j = \begin{cases} \vec{w}_t & \text{falls } j = i_t \text{ für ein } t \\ \vec{v}_j & \text{sonst} \end{cases}$$

bilden wieder eine Basis von V .

Speziell gilt also $m \leq n$.

Begründung von 3.3.6: Zunächst zeigen wir: Ist $\vec{w} \in V$ ein beliebiger, vom Nullvektor verschiedener Vektor, und gilt

$$\vec{w} = r_1 \vec{v}_1 + \dots + r_n \vec{v}_n \quad (3.3.1)$$

und gilt $r_i \neq 0$ für ein i , so ist $\{\vec{v}_1, \dots, \vec{v}_{i-1}, \vec{w}, \vec{v}_{i+1}, \dots, \vec{v}_n\}$ wieder eine Basis von V .

Dazu nehmen wir an, dass $i = 1$. Das können wir erreichen, indem wir die \vec{v}_i anders durchnummerieren. Dann gilt also $\vec{w} = r_1 \vec{v}_1 + \dots + r_n \vec{v}_n$ mit $r_1 \neq 0$ und daher

$$\vec{v}_1 = \frac{1}{r_1} \cdot (\vec{w} - r_2 \vec{v}_2 - \dots - r_n \vec{v}_n)$$

und es folgt leicht, dass $\{\vec{w}, \vec{v}_2, \dots, \vec{v}_n\}$ den Vektorraum V erzeugen. Es bleibt zu zeigen, dass sie auch linear unabhängig sind. Dazu sei eine Gleichung

$$\vec{0} = s_1 \vec{w} + s_2 \vec{v}_2 + \dots + s_n \vec{v}_n$$

gegeben. Für \vec{w} setzen wir nun (3.3.1) ein und erhalten

$$\vec{0} = s_1 r_1 \vec{v}_1 + (s_2 + s_1 r_2) \vec{v}_2 + \dots + (s_n + s_1 r_n) \vec{v}_n$$

Aus der linearen Unabhängigkeit von $\vec{v}_1, \dots, \vec{v}_n$ folgt

$$r_1 s_1 = 0, \quad s_i + s_1 r_i = 0 \quad \text{für } i \geq 2$$

Da $r_1 \neq 0$ erhalten wir hieraus zunächst $s_1 = 0$, und damit durch Einsetzen in die anderen Gleichungen auch sofort $s_i = 0$ für $i = 2, \dots, n$.

Wir benutzen vollständige Induktion über die Anzahl m der linear unabhängigen Vektoren.

Für $m = 0$ ist nichts zu zeigen.

Sei also $m \geq 1$ und die Behauptung für $m - 1$ linear unabhängige Vektoren schon bewiesen. Durch geschicktes Umm Nummerieren und Anwendung der Induktionsvoraussetzung können wir erreichen, dass $\{\vec{w}_1, \dots, \vec{w}_{m-1}, \vec{v}_m, \dots, \vec{v}_n\}$ eine Basis von V ist. Mit dieser Basis können wir auch \vec{w}_m darstellen und schreiben

$$\vec{w}_m = r_1 \vec{w}_1 + \dots + r_{m-1} \vec{w}_{m-1} + r_m \vec{v}_m + \dots + r_n \vec{v}_n$$

Dabei muss mindestens einer der Koeffizienten r_m, \dots, r_n von Null verschieden sein, den andernfalls bekämen wir eine nichttriviale Relation zwischen $\vec{w}_1, \dots, \vec{w}_m$,

$$\vec{w}_m - r_1 \vec{w}_1 - \dots - r_{m-1} \vec{w}_{m-1} = \vec{0}$$

im Widerspruch zu deren linearer Unabhängigkeit. Nach dem ersten Teil des Beweises können wir das entsprechende \vec{v}_i durch \vec{w}_m ersetzen und erhalten wieder eine Basis. Damit ist der Basisaustauschsatz von Steinitz bewiesen.

Begründung von 3.3.5: Sei zunächst $\{\vec{v}_1, \dots, \vec{v}_n\}$ eine endliche Basis von V und sei $\{\vec{w}_\lambda\}_{\lambda \in \Lambda}$ eine beliebige weitere Basis von V . Dann muss nach dem Satz von Steinitz schon gelten, dass $|\lambda| \leq n$ ist. (Falls Λ endlich ist, folgt das sofort aus 3.3.6. Wäre Λ unendlich, so könnten wir darin eine $n + 1$ -elementige Teilmenge Λ_0 auswählen, und nach 3.3.3 wäre dann auch $\{\vec{w}_\lambda\}_{\lambda \in \Lambda_0}$ linear unabhängig, im Widerspruch zu 3.3.6). Damit können wir also $\{\vec{w}_\lambda\}_{\lambda \in \Lambda}$ schreiben als $\{\vec{w}_1, \dots, \vec{w}_m\}$ mit $m \leq n$. Wenn wir nun die Rollen von $\{\vec{w}_1, \dots, \vec{w}_m\}$ und $\{\vec{v}_1, \dots, \vec{v}_n\}$ vertauschen und wieder den Satz von Steinitz anwenden, so erhalten wir auch $n \leq m$, also insgesamt $n = m$. Damit haben wir also gezeigt: Hat V eine endliche Basis, so sind alle anderen Basen von V ebenfalls endlich und haben die gleiche Länge wie die vorgegebene Basis. Hieraus folgt aber auch: Hat V eine Basis unendlicher Länge, so sind auch alle anderen Basen unendlich lang.

Folgerung 3.3.7. *Ist $U \subseteq V$ ein Untervektorraum, so gilt:*

$$\dim(U) \leq \dim(V)$$

Beweis: Eine Basis von U ist eine Familie linear unabhängiger Vektoren in V , lässt sich also nach Korollar 3.3.3 zu einer Basis von V ergänzen. Daraus folgt die Behauptung.

Aufgabe 3.3.1. Sei $V = \text{Abb}(I, \mathbb{R})$ die Menge der Abbildungen

$$f : I \longrightarrow \mathbb{R}$$

wobei I ein beliebiges Intervall ist (geschlossen, offen oder halboffen). Zeigen Sie, dass V ein Vektorraum wird, wenn man die Vektorraumoperationen wie im Beispiel 3.3.1 für $\text{Abb}([0, 1], \mathbb{R})$ definiert.

Aufgabe 3.3.2. Sei $V = \text{Abb}(\mathbb{R}, \mathbb{R})$ der Vektorraum der Abbildungen

$$f : \mathbb{R} \longrightarrow \mathbb{R}$$

(vergleiche Aufgabe 3.3.1). Zeigen Sie, dass die Elemente $f, g \in V$ mit

$$f(x) = \cos(x), \quad g(x) = \sin(x) \quad \text{für alle } x \in \mathbb{R}$$

linear unabhängig sind.

Aufgabe 3.3.3. Sei $V = \text{Abb}((0, \infty), \mathbb{R})$ der Vektorraum der Abbildungen

$$f : (0, \infty) \longrightarrow \mathbb{R}$$

(vergleiche Aufgabe 3.3.1). Zeigen Sie, dass die Elemente $f, g \in V$ mit

$$f(x) = \exp(x), \quad g(x) = \ln(x) \quad \text{für alle } x \in (0, \infty)$$

linear unabhängig sind.

Aufgabe 3.3.4. Sei $V = \text{Abb}((1, \infty), \mathbb{R})$ der Vektorraum der Abbildungen

$$f : (1, \infty) \longrightarrow \mathbb{R}$$

(vergleiche Aufgabe 3.3.1). Zeigen Sie, dass die Elemente $f, g, h \in V$ mit

$$f(x) = \frac{1}{x+1}, \quad g(x) = \frac{1}{x-1}, \quad h(x) = \frac{x}{x^2-1} \quad \text{für alle } x \in (1, \infty)$$

linear abhängig sind.

Aufgabe 3.3.5. Benutzen Sie den Satz von Steinitz 3.3.6 um folgende Aussage zu beweisen: Ist E_1 parallel zu E_2 , so ist auch E_2 parallel zu E_1 .

3.4 Vektorräume über beliebigen Körpern

Im vorangegangenen Abschnitt haben wir - weder in den Aussagen noch in den Beweisen - die spezielle Struktur von \mathbb{R} , also der Anordnung oder dem Dedekindschen Schnittgesetz, gemacht. Alle Aussagen und Beweise benutzen lediglich die Regeln für Addition, subtraktion, Multiplikation oder Division, also die allgemeinen Axiome, die für jeden Körper gelten. Genauso gut hätten wir anstelle von \mathbb{R} also auch \mathbb{C} , \mathbb{Q} oder gar einen der endlichen Körper \mathbb{F}_p , die wir in Abschnitt 2.3 kennengelernt haben, nehmen können. Das wollen wir in diesem Abschnitt auch tun. Dazu fixieren wir einen beliebigen Körper K .

Beispiel 3.4.1. Wir haben im Beispiel 3.3.2 in Abschnitt 3.3 schon allgemeine polynomiale Abbildungen

$$f : [0, 1] \longrightarrow \mathbb{R}$$

betrachtet. In der Praxis treten dabei meistens Polynome mit rationalen Koeffizienten auf, also Abbildungen $f(x)$, die sich als

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$$

mit $a_i \in \mathbb{Q}$ schreiben lassen. Die Menge dieser Polynome mit rationalen Koeffizienten wollen wir mit $\text{Pol}_{\mathbb{Q}}([0, 1], \mathbb{R})$ bezeichnen. Dann ist $\text{Pol}_{\mathbb{Q}}([0, 1], \mathbb{R})$ mit der Addition wie wir sie in Beispiel 3.3.2 definiert haben und der Skalarmultiplikation

$$(q \cdot f)(x) = q \cdot f(x) \quad \text{für alle } x \in [0, 1]$$

für $q \in \mathbb{Q}$ ein \mathbb{Q} -Vektorraum.

Beachten Sie dabei, dass $\text{Pol}_{\mathbb{Q}}([0, 1], \mathbb{R})$ kein reeller Vektorraum ist, denn die Abbildung f mit $f(x) = x$ ist sicherlich ein Element von $\text{Pol}_{\mathbb{Q}}([0, 1], \mathbb{R})$, aber $\sqrt{2} \cdot f$ ist es nicht.

Beispiel 3.4.2. Die Menge \mathbb{R} der reellen Zahlen, zusammen mit der üblichen Addition reeller Zahlen und der (üblichen) Multiplikation mit rationalen Zahlen ist ein \mathbb{Q} -Vektorraum.

Nun betrachten wir in einem beliebigen K -Vektorraum noch eine Familie $\{\vec{v}_\lambda\}_{\lambda \in \Lambda}$ von Elementen von V .

Definition 3.4.1.

- Ein Vektor \vec{w} heißt **Linearkombination** von $\{\vec{v}_\lambda\}_{\lambda \in \Lambda}$, wenn es *endlich viele* $\lambda_1, \dots, \lambda_n \in \Lambda$ und $\kappa_1, \dots, \kappa_n \in K$ gibt mit

$$\vec{w} = \kappa_1 \cdot \vec{v}_{\lambda_1} + \dots + \kappa_n \cdot \vec{v}_{\lambda_n}$$

- Eine Familie $\{\vec{v}_\lambda\}_{\lambda \in \Lambda}$ von Vektoren in V heißt ein **Erzeugendensystem** von V , wenn sich jedes Element $\vec{v} \in V$ als Linearkombination von $\{\vec{v}_\lambda\}_{\lambda \in \Lambda}$ schreiben lässt.
- Eine Familie $\{\vec{v}_\lambda\}_{\lambda \in \Lambda}$ von Vektoren in V heißt **linear unabhängig**, wenn gilt:

Ist $\Lambda_0 = \{\lambda_1, \dots, \lambda_n\} \subseteq \Lambda$ eine beliebige endliche Teilmenge von Λ und sind $\kappa_1, \kappa_2, \dots, \kappa_n \in K$ Skalare mit

$$\kappa_1 \cdot \vec{v}_{\lambda_1} + \kappa_2 \cdot \vec{v}_{\lambda_2} + \dots + \kappa_n \cdot \vec{v}_{\lambda_n} = \vec{0}$$

so muss schon gelten:

$$\kappa_1 = \kappa_2 = \dots = \kappa_n = 0$$

Andernfalls heißt die Familie $\{\vec{v}_\lambda\}_{\lambda \in \Lambda}$ linear abhängig.

- Eine Familie $\{\vec{v}_\lambda\}_{\lambda \in \Lambda}$ von Vektoren in V heißt **Basis** von V , wenn die Vektoren $\{\vec{v}_\lambda\}_{\lambda \in \Lambda}$ sowohl linear unabhängig als auch ein Erzeugendensystem von V ist.

Beispiel 3.4.3. Über jeden Körper K hat das n -dimensionale kartesische Produkt K^n die Basis

$$\vec{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \vec{e}_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \dots, \vec{e}_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

Beispiel 3.4.4. Der \mathbb{Q} -Vektorraum $\text{Pol}_{\mathbb{Q}}([0, 1], \mathbb{R})$ hat die Basis $1, x, x^2, x^3, \dots$

Aufgabe 3.4.1. Der Begriff der linearen Unabhängigkeit hängt ab von den Skalaren, die wir betrachten. Zeigen Sie dazu:

Betrachten wir \mathbb{R} als \mathbb{Q} -Vektorraum, so sind die Elemente $1, \sqrt{2} \in \mathbb{R}$ linear unabhängig. Betrachten wir $\mathbb{R} = \mathbb{R}^1$ als \mathbb{R} -Vektorraum, so sind sie linear abhängig.

Alle Aussagen, die wir über allgemeine reelle Vektorräume getroffen haben, gelten vollkommen analog auch für Vektorräume über beliebigen Körpern. Speziell gilt das für Satz 3.3.1, der sich - inklusive Beweis - wortwörtlich auf unsere Situation überträgt. Wir wollen hier seine für uns wichtigsten Folgerungen auflisten:

Satz 3.4.1. *Für einen beliebigen K -Vektorraum V gilt:*

1. *Ist $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m\}$ ein Erzeugendensystem eines Vektorraums V , so enthält es eine Basis, d.h. es gibt $i_1, i_2, \dots, i_n \in \{1, \dots, m\}$, so dass $\{\vec{v}_{i_1}, \vec{v}_{i_2}, \dots, \vec{v}_{i_n}\}$ eine Basis von V ist.*
2. *Ist $\{\vec{v}_\lambda\}_{\lambda \in \Lambda}$ eine Familie linear unabhängiger Vektoren in V , so kann $\{\vec{v}_\lambda\}_{\lambda \in \Lambda}$ zu einer Basis von V ergänzt werden.*
3. *Jeder Vektorraum hat eine Basis.*

Wie auch für reelle Vektorräume zeigt man im allgemeinen Fall

Satz 3.4.2. *Je zwei K -Basen $\{\vec{v}_\lambda\}_{\lambda \in \Lambda}$ und $\{\vec{w}_\gamma\}_{\gamma \in \Gamma}$ eines K -Vektorraums V sind gleich lang, d.h. sie sind entweder beide unendlich oder beide endlich mit der selben Anzahl von Elementen.*

Definition 3.4.2. Die Länge einer K -Basis von V heißt die **Dimension** des K -Vektorraums V und wird mit $\dim_K(V)$ bezeichnet.

Bemerkung 3.4.1. $\dim_K(V)$ ist eine endliche Zahl oder ∞ .

Beispiel 3.4.5. $\dim_K(K^n) = n$.

Allgemeine reelle Vektorräume unterscheiden sich von den speziellen reellen Vektorräumen \mathbb{R}^n in ihrer geometrischen Interpretation. Auf \mathbb{R}^n konnten wir ein Skalarprodukt definieren und damit Begriffe wie *Abstand*, *Winkel* oder *Orthogonalität* einführen. Wir können uns jetzt natürlich ebenso fragen, ob eine geometrische Interpretation auch für die Körper \mathbb{K}^n möglich ist.

Zunächst scheint zumindest die formale Definition eines Skalarproduktes auf K^n möglich zu sein, indem wir einfach die Definition aus dem \mathbb{R}^n übertragen

und setzen

$$\left\langle \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}, \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \right\rangle' = a_1 b_1 + \cdots + a_n b_n$$

wir sehen aber sofort einen fundamentalen Unterschied zur reellen Situation:

Beispiel 3.4.6. Im \mathbb{F}_2 -Vektorraum \mathbb{F}_2^2 gilt:

$$\left\langle \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\rangle' = 1 + 1 = 0$$

Im \mathbb{C} -Vektorraum \mathbb{C}^2 gilt:

$$\left\langle \begin{pmatrix} 1 \\ i \end{pmatrix}, \begin{pmatrix} 1 \\ i \end{pmatrix} \right\rangle' = 1 + i^2 = 1 - 1 = 0$$

Die von uns betrachtete Paarung $\langle \cdot, \cdot \rangle'$ hat also nicht mehr die Eigenschaft $\langle \vec{v}, \vec{v} \rangle' \neq 0$ (sogar $\langle \vec{v}, \vec{v} \rangle' > 0$) für $\vec{v} \neq \vec{0}$. Und sogar in \mathbb{Q}^n (wo diese Nichtverschwindungsaussage noch gilt) hat diese Paarung einen entscheidenden Nachteil:

Beispiel 3.4.7. Wir betrachten den \mathbb{Q} -Vektorraum \mathbb{Q}^2 und den Vektor $\vec{v} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. Dann gibt es keine rationale Zahl q mit

$$\langle q \cdot \vec{v}, q \cdot \vec{v} \rangle' = 1$$

Im \mathbb{R}^n konnten wir diese Normierung von \vec{v} immer erreichen, indem wir \vec{v} mit dem (reellen) Skalar $r = \frac{1}{|\vec{v}|}$ multiplizierten. Diese Eigenschaft war sehr wichtig für uns, etwa beim Gram-Schmidt-Orthonormalisierungsverfahren. Die Definition des Skalarprodukts überträgt sich also nicht unmittelbar auf beliebige K -Vektorräume K^n . Aber lässt sich die Definition dann vielleicht so abändern, dass wir eine dem Skalarprodukt ähnliche Paarung erhalten? Zumindest im Fall der komplexen Vektorräume \mathbb{C}^n ist das möglich:

Definition 3.4.3. Für zwei Vektoren $\vec{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}, \vec{w} = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} \in \mathbb{C}^n$ setzen wir

$$\langle \vec{v}, \vec{w} \rangle = v_1 \cdot \overline{w_1} + \cdots + v_n \cdot \overline{w_n}$$

(wobei \bar{z} die zu einer komplexen Zahl $z = x + i \cdot y$ konjugierte Zahl $x - i \cdot y$ bezeichnet) und nennen es das **Skalarprodukt** von \vec{v} und \vec{w} .

Eine ganz wesentliche Eigenschaft dieses Skalarprodukts ist

Satz 3.4.3. Für $\vec{v} \neq \vec{0}$ ist $\langle \vec{v}, \vec{v} \rangle$ eine reelle Zahl mit

$$\langle \vec{v}, \vec{v} \rangle > 0$$

Beweis: Wir haben

$$\langle \vec{v}, \vec{v} \rangle = v_1 \cdot \overline{v_1} + \cdots + v_n \cdot \overline{v_n} = |v_1|^2 + \cdots + |v_n|^2$$

wobei $|z| = \sqrt{x^2 + y^2}$ den Betrag einer komplexen Zahl $z = x + i \cdot y$ bezeichnet. Hieraus folgt die Aussage.

Definition 3.4.4. Für ein $\vec{v} \in \mathbb{C}^n$ heißt

$$|\vec{v}| := \sqrt{\langle \vec{v}, \vec{v} \rangle}$$

die **Länge** oder die **Norm** von \vec{v} .

Für das komplexe Skalarprodukt gelten die folgenden Regeln:

Regel 3.4.4. Für Vektoren $\vec{v}, \vec{w}, \vec{w}_1$ und \vec{w}_2 und einen Skalar $z \in \mathbb{C}$ gilt:

1. $\langle \vec{v}, \vec{w} \rangle = \overline{\langle \vec{w}, \vec{v} \rangle}$.
2. $\langle \vec{v}, \vec{w}_1 + \vec{w}_2 \rangle = \langle \vec{v}, \vec{w}_1 \rangle + \langle \vec{v}, \vec{w}_2 \rangle$ (Distributivgesetz).
3. $\langle z \cdot \vec{v}, \vec{w} \rangle = z \cdot \langle \vec{v}, \vec{w} \rangle$ (Skalarmultiplikation im ersten Faktor).
4. $\langle \vec{v}, z \cdot \vec{w} \rangle = \overline{z} \cdot \langle \vec{v}, \vec{w} \rangle$ (Skalarmultiplikation im zweiten Faktor).

Auch die Dreiecksungleichung und der Satz von Cauchy–Schwarz übertragen sich auf diese Situation:

Satz 3.4.5. Für Vektoren \vec{v}, \vec{w} und einen Skalar $z \in \mathbb{C}$ gilt:

1. $|z \cdot \vec{v}| = |z| \cdot |\vec{v}|$.
2. $|\vec{v} + \vec{w}| \leq |\vec{v}| + |\vec{w}|$.
3. $|\langle \vec{v}, \vec{w} \rangle| \leq |\vec{v}| \cdot |\vec{w}|$.

4. Genau dann sind \vec{v} und \vec{w} kollinear, wenn $|\langle \vec{v}, \vec{w} \rangle| = |\vec{v}| \cdot |\vec{w}|$.

Definition 3.4.5. Zwei Vektoren \vec{v} und \vec{w} in \mathbb{C}^n heißen **orthogonal**, wenn gilt

$$\langle \vec{v}, \vec{w} \rangle = 0$$

Wir sagen in diesem Fall auch, dass \vec{v} senkrecht auf \vec{w} steht und schreiben $\vec{v} \perp \vec{w}$.

Definition 3.4.6. Eine Basis $\vec{v}_1, \dots, \vec{v}_m$ eines Untervektorraums $U \subseteq \mathbb{C}^n$ heißt **Orthonormalbasis** von U oder kurz **ONB** von U , wenn gilt:

- $|\vec{v}_i| = 1$ für $i = 1, \dots, m$, d.h. alle Vektoren der Basis sind normiert.
- $\langle \vec{v}_i, \vec{v}_j \rangle = 0$ für $i \neq j$, d.h. die Vektoren stehen paarweise senkrecht aufeinander.

Aufgabe 3.4.2. Wieviele Basen hat der \mathbb{F}_3 -Vektorraum \mathbb{F}_3^2 ?

Aufgabe 3.4.3. Zeigen Sie $\dim_{\mathbb{Q}}(\mathbb{R}) = \infty$.

Aufgabe 3.4.4. Beweisen Sie die Regeln 3.4.4.

Aufgabe 3.4.5. Übertagen Sie das Gram–Schmidt–Orthonormalisierungsverfahren auf den komplexen Vektorraum \mathbb{C}^n und benutzen Sie es, um aus den Vektoren

$$\vec{v}_1 = \begin{pmatrix} 1 \\ i \\ i \end{pmatrix}, \quad \vec{v}_2 = \begin{pmatrix} 0 \\ i \\ -i \end{pmatrix}, \quad \vec{v}_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

eine Orthonormalbasis des \mathbb{C}^n zu machen.

Kapitel 4

Lineare Gleichungssysteme

4.1 Matrizen und Gleichungssysteme

Ein entscheidender Grund für die Beschäftigung mit Vektoren und Vektorräume sind lineare Gleichungssysteme und damit zusammenhängende Fragen. In vielen praktischen Anwendungen hängen unterschiedliche Größen etwa eines physikalischen Systems in linearer Weise voneinander ab. In diesem Abschnitt wollen wir das Studium solcher Zusammenhänge formalisieren und vorantreiben.

Der Einfachheit halber beschäftigen wir uns hier nur mit reellen Gleichungssysteme und reellen Matrizen, wir notieren jedoch, dass Matrizen und Gleichungssystemen, genauso wie die Vektorräume, über beliebigen Körpern betrachtet werden können. Dadurch ändern sich jedoch die Methoden, Techniken und Beweise in diesem Abschnitt nicht, und daher überlassen wir die nötigen Anpassungen dem interessierten Leser.

4.1.1 Lineare Gleichungssysteme

Bereits in den letzten Abschnitten haben wir gesehen, dass oft Beziehungen zwischen Vektoren zu betrachten sind. Bei der Untersuchung auf lineare Unabhängigkeit der Vektoren $\vec{v}_1, \dots, \vec{v}_m$ etwa stehen wir vor dem Problem, alle Werte r_1, \dots, r_m zu finden, für die

$$r_1 \vec{v}_1 + \dots + r_m \vec{v}_m = \vec{0}$$

gilt. Handelt es sich dabei um Vektoren aus dem \mathbb{R}^n und schreiben wir sie in Komponentenschreibweise, so erhalten wir daraus eine Gleichung für jede Komponente, und damit insgesamt n Gleichungen

$$\begin{array}{ccccccc} r_1 v_{1,1} & + & \dots & + & r_m v_{m,1} & = & 0 \\ r_1 v_{1,2} & + & \dots & + & r_m v_{m,2} & = & 0 \\ \vdots & & \ddots & & \vdots & & \vdots \\ r_1 v_{1,n} & + & \dots & + & r_m v_{m,n} & = & 0 \end{array}$$

Die Frage nach der linearen Unabhängigkeit der Vektoren $\vec{v}_1, \dots, \vec{v}_m$ ist also äquivalent zur Frage ob dieses Gleichungssystem ausser der trivialen Lösung $r_1 = \dots = r_m = 0$ noch andere Lösungen hat.

Ähnliche Problemstellungen treten in vielen anderen Bereichen auf:

Beispiel 4.1.1. Alpaka (Neusilber) ist eine Legierung aus Kupfer, Nickel und Zink. Im Lager liegen vier Legierungen I - IV mit den durch folgende Tabelle angegebenen Zusammensetzungen

	<i>I</i>	<i>II</i>	<i>III</i>	<i>IV</i>
Kupfer	40%	50%	60%	70%
Nickel	26%	22%	25%	18%
Zink	34%	28%	15%	12%

Wie lässt sich aus diesen Legierungen Alpaka mit einem Gehalt von 55% Kupfer, 23% Nickel und 22% Zink herstellen? Schreiben wir hier zunächst die Komponenten der vorhandenen Legierungen jeweils als Vektoren, also

$$\vec{v}_1 = \begin{pmatrix} 40 \\ 26 \\ 34 \end{pmatrix}, \vec{v}_2 = \begin{pmatrix} 50 \\ 22 \\ 28 \end{pmatrix}, \vec{v}_3 = \begin{pmatrix} 60 \\ 25 \\ 15 \end{pmatrix}, \vec{v}_4 = \begin{pmatrix} 70 \\ 18 \\ 12 \end{pmatrix}$$

und ebenso für die Ziellegierung

$$\vec{b} = \begin{pmatrix} 55 \\ 23 \\ 22 \end{pmatrix}$$

so suchen wir Zahlen a, b, c und d , die die Gleichung

$$a\vec{v}_1 + b\vec{v}_2 + c\vec{v}_3 + d\vec{v}_4 = 100 \vec{b}$$

erfüllen. Auch das führt zu einer Gleichung für jede einzelne Komponente, und damit zu einem Gleichungssystem

$$\begin{aligned}40a + 50b + 60c + 70d &= 5500 \\26a + 22b + 25c + 18d &= 2300 \\34a + 28b + 15c + 12d &= 2200\end{aligned}$$

Man sieht dieser Gleichung sicher nicht unmittelbar an, welche Lösungen sie hat, oder ob sie überhaupt Lösungen hat.

Bei der Betrachtung linearer Gleichungssysteme ist es naheliegend, aus einer Gleichung Information über eine der Unbekannten zu ziehen und damit die anderen Gleichungen zu vereinfachen.

Beispiel 4.1.2. Wir betrachten das lineare Gleichungssystem

$$\begin{aligned}1x + 4y &= 8 \\3x + 2y &= 4\end{aligned}$$

Durch Betrachtung der ersten Gleichung erhalten wir

$$x = 8 - 4y$$

Setzen wir das in die zweite Gleichung ein, so erhalten wir

$$3(8 - 4y) + 2y = 20$$

also

$$24 - 10y = 4$$

Das ist eine lineare Gleichung, mit einer Unbekannten, die wir leicht lösen können. Durch Zusammenfassen der Terme wird sie zu

$$10y = 20$$

hat also die Lösung $y = 2$. Setzen wir das nun wiederum in das ursprüngliche Gleichungssystem ein, so erhalten wir

$$\begin{aligned}1x + 8 &= 8 \\3x + 4 &= 4\end{aligned}$$

also $x = 0$, und wir haben gezeigt, dass $x = 0$, $y = 4$ die eindeutige Lösung unseres Gleichungssystems ist.

Anstatt die erste Gleichung nach x aufzulösen und das Resultat in die zweite Gleichung einzusetzen, hätten wir auch hingehen können und ein geeignetes Vielfaches der ersten Gleichung (nämlich das Dreifache) von der zweiten Gleichung abzuziehen. Das ergibt

$$(3x + 4y) - 3 \cdot (1x + 4y) = 4 - 3 \cdot 8$$

also

$$-10y = -20$$

und liefert uns wieder $y = 2$, $x = 0$. Diese beiden Ansätze sind äquivalent.

Beispiel 4.1.3. Wir wollen nun das Gleichungssystem aus Beispiel 4.1.1 mit der zweiten Methode aus Beispiel 4.1.2 behandeln. Zunächst subtrahieren wir das $\frac{26}{40}$ -fache der ersten Zeile von der zweiten, und das $\frac{34}{40}$ -fache der ersten Zeile von der dritten und erhalten

$$\begin{array}{rclclcl} 40a & + & 50b & + & 60c & + & 70d & = & 5500 \\ & & -\frac{21}{2}b & - & 14c & - & \frac{55}{2}d & = & -1275 \\ & & -\frac{29}{2}b & - & 36c & - & \frac{95}{2}d & = & -2475 \end{array}$$

Dadurch haben wir die zweite und die dritte Zeile dahingehend vereinfacht, dass wir in diesen Gleichungen nur noch drei Variablen zu betrachten haben. Wir gehen weiter, indem wir den selben Ansatz auf das Gleichungssystem, das nur noch aus der zweiten und dritten Gleichung besteht, anwenden. Dazu ziehen wir das $\frac{29}{21}$ -fache der zweiten Gleichung von der dritten Gleichung ab und erhalten

$$\begin{array}{rclclcl} 40a & + & 50b & + & 60c & + & 70d & = & 5500 \\ & & -\frac{21}{2}b & - & 14c & - & \frac{55}{2}d & = & -1275 \\ & & & & -\frac{50}{3}c & - & \frac{250}{21}d & = & -\frac{5000}{7} \end{array}$$

Wir haben also die letzte Gleichung auf zwei Unbekannte reduziert und das ganze Gleichungssystem auf eine Dreiecksform gebracht. Da wir keine weiteren Gleichungen mehr zur Verfügung haben, lässt sich auch keine weitere Elimination mehr durchführen. Wir können nun so vorgehen, dass wir für d

einen beliebigen Wert x einsetzen. Wir erhalten dann aus der letzten Gleichung

$$c = \frac{300}{7} - \frac{5}{7}x$$

Diesen Wert können wir (zusammen mit x für d) für c in die zweite Gleichung einsetzen und erhalten

$$b = \frac{2}{21} \cdot \left(1275 - 14 \left(\frac{300}{7} - \frac{5}{7}x \right) - \frac{55}{2}x \right) = \frac{450}{7} - \frac{5}{3}x$$

Setzen wir nun alles in die erste Gleichung ein, so ergibt sich

$$a = \frac{1}{40} \left(5500 - 60 \left(\frac{450}{7} - \frac{5}{3}x \right) - 70 \left(\frac{300}{7} - \frac{5}{7}x \right) - 70x \right) = -\frac{9500}{7} + 80x$$

Und damit haben wir erhalten: Alle Lösungen unseres Gleichungssystems sind von der Form

$$a = -\frac{9500}{7} + 80x, \quad b = \frac{450}{7} - \frac{5}{3}x, \quad c = \frac{300}{7} - \frac{5}{7}x, \quad d = x$$

mit einem beliebigen x . Es gibt also viele Lösungen. Um unser ursprüngliches Problem zu behandeln ist es natürlich erforderlich, x so zu wählen, dass $a \geq 0, b \geq 0, c \geq 0$ und $d \geq 0$. Aber auch das ist möglich, etwa durch die Wahl von $x = 30$.

Die in den Beispielen 4.1.2 und 4.1.3 skizzierten Ansätze lassen sich zu einem allgemeinen Algorithmus zur Lösung linearer Gleichungssysteme ausbauen. Allerdings ist dazu noch ein zusätzlicher Fall zu betrachten.

Beispiel 4.1.4. Gegeben sei das Gleichungssystem

$$\begin{aligned} 2z &= 8 \\ 2x + 2y + 4z &= 4 \end{aligned}$$

Dieses System ist zwar einfach zu lösen, fügt sich aber insofern nicht in unser Schema aus Beispiel 4.1.3 oder 4.1.2 ein, als es nicht möglich ist, durch Addition eines geeigneten Vielfachen der ersten Gleichung die erste Unbekannte aus der zweiten Gleichung zu eliminieren. Um eine allgemeingültigen und computertauglichen Algorithmus zu erhalten, ist das aber erforderlich. Deshalb schieben wir einen Zwischenschritt ein und vertauschen zunächst mal die beiden Zeilen

$$\begin{aligned} 2x + 2y + 4z &= 4 \\ 2z &= 8 \end{aligned}$$

Nun haben wir schon fast was wir wünschen. Für eine perfekte Dreiecksform ist allerdings auch noch dafür Sorge zu tragen, dass in der zweiten Gleichung die zweite Unbekannte (und nicht nur die dritte) vorkommt. Das können wir am einfachsten dadurch erreichen, dass wir die Rollen von y und z vertauschen, als schreiben

$$\begin{array}{rclcl} 2x & + & 4z & + & 2y & = & 4 \\ & & 2z & & & = & 8 \end{array}$$

Schließlich nehmen wir noch eine Umbenennung vor und schreiben y für z und z für y und erhalten

$$\begin{array}{rclcl} 2x & + & 4y & + & 2z & = & 4 \\ & & 2y & & & = & 8 \end{array}$$

also ein Gleichungssystem, das genau in unser Schema passt. Um uns das Leben noch einfacher zu machen, normieren wir die Gleichungen dahingehend, dass wir jede Gleichung durch den Koeffizienten der ersten auftretenden Variable teilen und bekommen das Gleichungssystem

$$\begin{array}{rclcl} x & + & 2y & + & z & = & 2 \\ & & y & & & = & 4 \end{array}$$

aus dem wir sofort ablesen, dass seine allgemeine Lösung von der Form

$$x = -2 - \lambda, y = 4, z = \lambda$$

mit einem freien Parameter λ ist. Zu beachten ist dabei, dass die Vertauschungen rückgängig gemacht werden müssen, wenn wir unser ursprüngliches Gleichungssystem betrachten. In diesem Fall bedeutet das, dass das Gleichungssystem

$$\begin{array}{rclcl} & & 2z & = & 8 \\ 2x & + & 2y & + & 4z & = & 4 \end{array}$$

die allgemeine Lösung

$$x = -2 - \lambda, y = \lambda, z = 4$$

mit einem freien Parameter λ hat.

Bemerkung 4.1.1. Wir haben also vier Operationen kennengelernt, die wir an Gleichungssystemen durchführen können

- (i) Subtrahiere das Vielfache einer Gleichung von einer anderen Gleichung
- (ii) Vertausche zwei Gleichungen
- (iii) Vertausche zwei Variablen
- (iv) Multipliziere eine Gleichung mit einer Zahl $r \neq 0$.

Diese Operationen führen zu einem äquivalenten Gleichungssystem, also zu einem Gleichungssystem das die selben Lösungen hat wie das ursprüngliche (wobei im Fall der Variablenvertauschung diese rückgängig zu machen ist). Diese Operationen reichen auch aus, um jedes Gleichungssystem auf eine Form zu bringen in der

- a) sofort entschieden werden kann, ob das Gleichungssystem lösbar ist.
- b) im Fall der Lösbarkeit alle Lösungen sofort angegeben werden können.

Dazu betrachten wir jetzt ein allgemeines System linearer Gleichungen, bestehend aus m Gleichungen mit n Unbekannten. Dafür benutzen wir folgende Notation:

$$\begin{array}{ccccccc} a_{1,1}x_1 & + & a_{1,2}x_2 & + & \cdots & + a_{1,n}x_n & = & b_1 \\ a_{2,1}x_1 & + & a_{2,2}x_2 & + & \cdots & + a_{2,n}x_n & = & b_2 \\ \vdots & & & & \ddots & & & \vdots \\ a_{m,1}x_1 & + & a_{m,2}x_2 & + & \cdots & + a_{m,n}x_n & = & b_m \end{array} \quad (4.1.1)$$

Dabei stehen x_1, \dots, x_n für die zu bestimmenden Unbekannten, $a_{1,1}, \dots, a_{m,n}$ für die Koeffizienten des Gleichungssystems und b_1, \dots, b_m für die Ergebnisse. In Beispiel 4.1.2 etwa war $x_1 = x, x_2 = y, a_{1,1} = 1, a_{1,2} = 4, a_{2,1} = 3$ und $a_{2,2} = 2$ sowie $b_1 = 8$ und $b_2 = 4$.

Definition 4.1.1. Eine Gleichungssystem (4.1.1) heißt **homogen**, wenn $b_i = 0$ für alle i .

Bemerkung 4.1.2. Jedes homogene Gleichungssystem mit n Unbekannten hat mindestens eine Lösung, nämlich

$$x_1 = 0, x_2 = 0, \dots, x_n = 0$$

Diese Lösung heißt auch **triviale Lösung** des homogenen Gleichungssystems.

Definition 4.1.2. Wir sagen, ein Gleichungssystem (4.1.1) liegt in (**Gaußscher**) **Normalform** vor, wenn es ein $t \in \{1, \dots, m\}$ gibt mit

- $a_{i,j} = 0$ für alle $i > t$ und alle j .
- $a_{i,j} = 0$ für $i \in \{2, \dots, t\}$ und $j < i$.
- $a_{i,i} = 1$ für $i = 1, \dots, t$.

In diesem Fall heißt t der **Rang** des linearen Gleichungssystems.

Bemerkung 4.1.3. Ein lineares Gleichungssystem in Normalform hat also die Gestalt

$$\begin{array}{cccccccc}
 x_1 & + & a_{1,2}x_2 & + & a_{1,3}x_3 & + & \cdots & + & a_{1,n}x_n & = & b_1 \\
 0 & + & x_2 & + & a_{2,3}x_3 & + & \cdots & + & a_{2,n}x_n & = & b_2 \\
 \vdots & & & & & & \ddots & & & & \vdots \\
 0 & + & 0 & + & 0 & + & \cdots & + & x_t & + & \cdots & + & a_{t,n}x_n & = & b_t \\
 0 & + & 0 & + & 0 & + & \cdots & + & 0 & = & b_{t+1} \\
 \vdots & & & & & & \ddots & & & & \vdots \\
 0 & + & 0 & + & 0 & + & \cdots & + & 0 & = & b_m
 \end{array} \quad (4.1.2)$$

In dieser allgemeinen Form sieht die Normalform sehr unübersichtlich aus. Die Situation wird aber einfacher, wenn wir einige Beispielen betrachten.

Beispiel 4.1.5. Das Gleichungssystem aus Beispiel 4.1.2 hat die Normalform

$$\begin{array}{rcl}
 x_1 & + & 4x_2 = 8 \\
 x_2 & & = 2
 \end{array}$$

Beispiel 4.1.6. Das Gleichungssystem aus Beispiel 4.1.1 hat Normalform

$$\begin{array}{rcl}
 x_1 & + & \frac{5}{4}x_2 + \frac{3}{2}x_3 + \frac{7}{4}x_4 = \frac{550}{4} \\
 & & x_2 + \frac{4}{3}x_3 + \frac{55}{21}x_4 = \frac{2550}{21} \\
 & & x_3 + \frac{5}{7}x_4 = \frac{300}{7}
 \end{array}$$

Gleichungssysteme in Normalform lassen sich sehr einfach weiterverarbeiten:

Satz 4.1.1. *Liegt ein Gleichungssystem (4.1.1) in Normalform (4.1.2) vor, so gilt*

- *Genau dann ist das Gleichungssystem lösbar, wenn $b_i = 0$ für $i > t$.*
- *Genau dann hat das Gleichungssystem eine eindeutige Lösung, wenn $b_i = 0$ für $i > t$ und $t = n$.*
- *Genau dann hat das Gleichungssystem unendlich viele Lösungen, wenn $b_i = 0$ für $i > t$ und $t < n$.*

Beweis: Es ist klar, dass wir keine Lösung haben können, wenn $b_i \neq 0$ für ein $i > t$, etwa $b_{t+1} \neq 0$. Dann haben wir nämlich eine Gleichung

$$0 = b_{t+1}$$

in unserem Gleichungssystem, die durch keine Wahl der x_i erfüllt werden kann.

Ist $b_i = 0$ für alle $i > t$, so können wir für unser Gleichungssystem wie in den Beispielen einfach Lösungen angeben indem wir mit der letzten Gleichung starten. Falls auch noch $t = n$, so ergibt sich aus der letzten Gleichung $x_n = b_n$ für x_n der eindeutige Lösungswert $x_n = b_n$, und durch sukzessives Einsetzen in die vorangehenden Gleichungen ergeben sich eindeutige Belegungen für x_1, \dots, x_n .

Falls dagegen $t < n$, so hat die letzte Gleichung die Gestalt

$$x_t + a_{t,t+1}x_{t+1} + \dots + a_{t,n}x_n = b_t,$$

mit $t \neq n$. Wir nennen in diesem Fall die Variablen x_{t+1}, \dots, x_n die **freien Variablen** des Gleichungssystems und können für x_{t+1}, \dots, x_n beliebige Werte einsetzen. Daraus erhalten wir eine eindeutige Lösung für x_t . Durch sukzessives Einsetzen in die vorangehenden Gleichungen ergeben sich eindeutige Belegungen für x_1, \dots, x_{t-1} . Deshalb werden x_1, \dots, x_t in diesem Fall auch als **gebundene Variablen** bezeichnet. Damit haben wir in dieser Situation $n - t + 1$ Freiheitsgrade, also unendlich viele Lösungen. Dieser

Beweis hat uns also auch gleichzeitig einen Algorithmus geliefert, um Gleichungssysteme in Normalform zu lösen (falls sie lösbar sind).

Beachten Sie dabei, dass der Fall $t > n$ nach Definition der Normalform nicht auftreten kann!

Um allgemeine Gleichungssysteme zu lösen, reicht es daher, einen Weg zu finden, sie in äquivalente Gleichungssysteme in Normalform zu überführen.

Satz 4.1.2. *Jedes Gleichungssystem kann durch die Operationen (i) - (iv) aus Bemerkung 4.1.1 in ein Gleichungssystem in Normalform überführt werden.*

Beweis: Wir geben einen Algorithmus an, der ein beliebiges Gleichungssystem in die gewünschte Form bringt.

Gegeben sei ein allgemeines lineares Gleichungssystem (4.1.1).

Startschritt: Setze $\rho = 0$.

Verarbeitungsschritt: Setze $\rho = \rho + 1$.

Gibt es ein $i = \{\rho, \dots, m\}, j \in \{\rho, \dots, n\}$ mit $a_{i,j} \neq 0$?

- Nein: STOPP, das Gleichungssystem ist in Normalform, und sein Rang ist $\rho - 1$.
- Ja: Führe folgende Schritte durch
 1. Falls $i > \rho$, wende Operation (ii) auf Gleichungen ρ und i an, dh. vertausche Zeile ρ und Zeile i des Gleichungssystems; falls $j > \rho$ wende Operation (iii) auf die Variablen x_ρ und x_j an, d.h. vertausche Spalte ρ und Spalte j des Gleichungssystems. Nach diesem Verarbeitungsschritt erhalten wir ein äquivalentes Gleichungssystem mit $a_{\rho,\rho} \neq 0$.
 2. Wende Operation (iv) mit $r = \frac{1}{a_{\rho,\rho}}$ auf Zeile ρ an, dh. multipliziere diese Gleichung mit $\frac{1}{a_{\rho,\rho}}$. Nach diesem Verarbeitungsschritt erhalten wir ein äquivalentes Gleichungssystem mit $a_{\rho,\rho} = 1$.
 3. Für $i = \rho + 1, \dots, m$ wende Operation (i) auf Gleichung ρ und Gleichung i mit Faktor $a_{i,\rho}$ an, dh. subtrahiere das $a_{i,\rho}$ -fache der ρ -ten Zeile von der i -ten Zeile. Nach diesem Verarbeitungsschritt

erhalten wir ein äquivalentes Gleichungssystem mit $a_{i,\rho} = 0$ für alle $i \in \{\rho + 1, \dots, m\}$.

Der Algorithmus endet also nach höchstens $\min\{n, m\}$ Verarbeitungsschritten.

Folgerung 4.1.3. *Ist ein Gleichungssystem homogen, so auch seine Normalform.*

Um den Algorithmus wirklich zu verstehen ist es am besten, ein Beispiel durchzuführen:

Beispiel 4.1.7. Wir betrachten das lineare Gleichungssystem

$$\begin{array}{rrrrcl} & 2x_2 & + & 4x_3 & = & 4 \\ x_1 & + & x_2 & + & x_3 & = & 6 \\ 2x_1 & + & x_2 & & & = & 10 \end{array}$$

Wir starten mit $\rho = 0$.

1. Schritt: $\rho = 1$: Es gibt $i \in \{1, 2, 3\}$ und $j \in \{1, 2, 3\}$ mit $a_{i,j} \neq 0$, etwa $i = 2, j = 1$.

1. Spaltenvertauschungen sind nicht durchzuführen, da $j = 1 = \rho$. Vertausche die ersten beiden Zeilen des Gleichungssystem und erhalte

$$\begin{array}{rrrrcl} x_1 & + & x_2 & + & x_3 & = & 6 \\ & & 2x_2 & + & 4x_3 & = & 4 \\ 2x_1 & + & x_2 & & & = & 10 \end{array}$$

2. Teile Gleichung 1 durch $a_{1,1} = 1$; das ändert das Gleichungssystem nicht.
3. Subtrahiere das 0-fache der ersten Gleichung von der zweiten Gleichung und das 2-fache der ersten Gleichung von der dritten Gleichung und erhalte

$$\begin{array}{rrrrcl} x_1 & + & x_2 & + & x_3 & = & 6 \\ & & 2x_2 & + & 4x_3 & = & 4 \\ & & -x_2 & + & -2x_3 & = & -2 \end{array}$$

2. Schritt: $\rho = 2$: Es gibt $i \in \{2, 3\}$ und $j \in \{2, 3\}$ mit $a_{i,j} \neq 0$, etwa $i = 2, j = 2$.

1. Vertauschungsoperationen sind nicht durchzuführen, da $i = \rho, j = \rho$.
2. Teile Gleichung 2 durch 2 und erhalte

$$\begin{array}{rrrr} x_1 & + & x_2 & + & x_3 & = & 6 \\ & & x_2 & + & 2x_3 & = & 2 \\ & & -x_2 & + & -2x_3 & = & -2 \end{array}$$

3. Subtrahiere das (-1) -fache der zweiten Zeile von der dritten Zeile und erhalte

$$\begin{array}{rrrr} x_1 & + & x_2 & + & x_3 & = & 6 \\ & & x_2 & + & 2x_3 & = & 2 \\ & & 0 & & 0 & = & 0 \end{array}$$

3. Schritt: $\rho = 3$: Es gibt kein $i \in \{3\}$ und $j \in \{3\}$ mit $a_{i,j} \neq 0$.

Der Algorithmus ist also beendet, und das Gleichungssystem ist in der Tat in Normalform. Sein Rang ist 2 und x_3 ist die freie Variable des Gleichungssystems. Seine allgemeine Lösung ist $x_3 = \lambda, x_2 = 2 - 2\lambda$ und $x_1 = 4 + \lambda$ mit einem freien Parameter λ .

Bemerkung 4.1.4. Die Normalform, die man am Ende des Algorithmus erhält, ist abhängig von Auswahl des **Pivotelements**, d.h. von der Auswahl von i, j mit $a_{i,j} \neq 0$, die in jedem Schritt getroffen wird.

Wir hätten in unserem Beispiel im ersten Schritt auch $i = 1, j = 2$ nehmen können und hätten dann als Normalform

$$\begin{array}{rrrr} x_1 & & + & 2x_3 & = & 2 \\ & x_2 & - & x_3 & = & 4 \\ & & 0 & & = & 0 \end{array}$$

erhalten. Das allgemeine Ergebnis für dieses Gleichungssystem ist $x_3 = \lambda, x_2 = 4 + \lambda$ und $x_1 = 2 - 2\lambda$ mit einem freien Parameter λ . Bei Betrachtung des Originalgleichungssystems ist wiederum zu beachten, dass wir die erste und die zweite Variable im Laufe des Algorithmus vertauscht haben. Machen wir

diese Vertauschung rückgängig, so erhalten wir als Lösung des Originalsystems $x_3 = \lambda$, $x_2 = 2 - 2\lambda$ und $x_1 = 4 + \lambda$ mit einem freien Parameter λ , genau wie oben.

Um einen Algorithmus mit reproduzierbarem Ergebnis zu bekommen, ist es also noch erforderlich, eine Strategie zum Auffinden eines Pivotelements anzugeben.

Bemerkung 4.1.5. Der Eliminationsalgorithmus ist anfällig für Rundungsfehler. Diese Anfälligkeit kann durch eine geschickte Wahl des Pivotelements in jedem Verarbeitungsschritt eliminiert werden.

Bei manueller (und exakter) Lösung des Gleichungssystems ist es hilfreich, wenn das Pivotelement den Wert 1 hat, für numerische Zwecke ist es aber günstiger, immer den betragsmäßig größten Wert zu wählen.

Das kann im Verarbeitungsschritt etwa wie folgt implementiert werden (mit $r = \rho$):

```

r = r+1;
a = abs(A(r,r)); i = r; j = r;
for o = r:m
    for p = r:n
        if abs(A(o,p)) > a
            a = abs(A(o,p));
            i = o;
            j = p;
        end
    end
end
end

```

- Falls $a = 0$: STOPP, das Gleichungssystem ist in Normalform, und sein Rang ist $\rho - 1$.
- Falls $a \neq 0$:
 1. Vertausche die Zeilen i und ρ und die Spalten j und ρ .
 2. Subtrahiere das $\frac{a_{i,\rho}}{a_{\rho,\rho}}$ -fache der ρ -ten Zeile von der i -ten Zeile.

In dieser Normalform sind die Diagonalelemente nicht notwendig normiert, sie erweist sich aber in der Praxis als deutlich weniger anfällig für Rundungsfehler wie das klassische Gaußverfahren.

Die hier angegebene Suche eines Pivotelements ist sehr aufwendig. Daher beschränken sich die meisten Verfahren in der Regel darauf, nur in der aktuellen Spalte ein passendes Element zu finden:

```

r = r+1;
a = abs(A(r,r)); i = r;
for o = r:m
    if abs(A(o,r)) > a
        a = abs(A(o,r));
        i = o;
    end
end

```

Dann werden die Zeilen i und ρ vertauscht. Nur wenn $a = 0$, kommt es noch zu Spaltenvertauschungen.

Satz 4.1.4. *Hat ein Gleichungssystem mit n Unbekannten den Rang n , so kann es nur mit den Operationen (i), (ii) und (iv) aus Bemerkung 4.1.1 auf Normalform gebracht werden. Es ist in diesem Fall also nicht nötig, Variablen zu vertauschen.*

Beweis: Das Vertauschen von zwei Variablen x_i und x_j oder von zwei Gleichungen wird benötigt, wenn im Schritt ρ der Eintrag $a_{\rho,\rho} = 0$ auftritt. Gibt es in diesem Fall ein $i \in \{\rho, \dots, m\}$ mit $a_{i,\rho} \neq 0$, so können wir die Zeilen i und ρ vertauschen und den Algorithmus ohne Operation (iii) fortsetzen. Gibt es kein solches i , so taucht in den Gleichungen ρ, \dots, m die Variable x_i nicht mehr auf. Da wir aber mit keiner der Operationen (i) - (iv) eine neue Variable generieren können, wird nach jedem Schritt eine der verbleibenden Variablen in den verbleibenden Gleichungen fehlen. Damit können wir aber in keine Normalform vom Rang n erreichen, denn in ihr müßte auch die fehlende Variable als führende Variable in einer Zeile auftauchen.

Bemerkung 4.1.6. Für die Behandlung von Gleichungssystemen ist es nicht erforderlich, dass die Koeffizienten reelle Zahlen sind. Der Eliminationsalgorithmus lässt sich genauso über jedem anderen Körper durchführen. Für uns relevant sind dabei die komplexen Zahlen. Betrachten wir etwa die Beziehung

$$\begin{aligned} x_1 + i \cdot x_2 &= 1 + i \\ (1 + i) \cdot x_1 - 2 \cdot x_2 &= 1 + 3 \cdot i \end{aligned}$$

so Können wir hier die zum reellen Fall analogen Operationen durchführen: Durch Subtraktion des $(i+1)$ -fachen der ersten Zeile von der zweiten erhalten wir

$$\begin{aligned} x_1 + i \cdot x_2 &= 1 + i \\ - (1 + i) \cdot x_2 &= 1 + i \end{aligned}$$

Division der zweiten Zeile durch $-(1 + i)$ ergibt

$$\begin{aligned} x_1 + i \cdot x_2 &= 1 + i \\ x_2 &= -1 \end{aligned}$$

und durch Rückwärtsrechnen erhalten wir

$$x_2 = -1, \quad x_1 = 1 + 2 \cdot i$$

Die Aussagen über die Lösbarkeit eines Gleichungssystems und die Anzahl der Lösungen gelten auch für Gleichungssysteme mit komplexen Koeffizienten.

Bemerkung 4.1.7. Das Gauß-Eliminationsverfahren ist (bei richtiger Pivotwahl) relativ stabil und von allen allgemeinen Verfahren das schnellste und effizienteste. Trotzdem ist es für viele Anwendungsprobleme zu langsam. Bei technischen Problemen, etwa Anwendungen linearer Näherungsverfahren im Maschinen- und Fahrzeugbau, kommt es schnell zu Gleichungssystemen mit 10^6 Gleichungen und Unbekannten. Im solche Gleichungssysteme mit dem Gaußalgorithmus zu lösen braucht selbst eine Rechner mit einer Rechenleistung von hundert GFlops mehrere Jahre. Obwohl die Gleichungssysteme, die dabei benutzt werden, aufgrund der speziellen Struktur in jeder Gleichung nur sehr wenige von 0 verschiedene Einträge stehen (oft weniger als hundert), führt das Eliminationsverfahren dann häufig zu relativ stark

besetzten Dreiecksmatrizen und damit auch zu einem Speicherbedarf von mehreren tausend Gigabytes.

Aus diesem Grund werden in solchen Situationen häufig Näherungsverfahren verwendet, die speziell auf diese Situation zugeschnitten und darauf optimiert sind. Eines dieser Verfahren ist das **Jacobi-Verfahren**. Erste Voraussetzung hierfür ist, dass $n = m$, dass es also genau so viele Gleichungen wie Unbekannte gibt. Es beruht auf der folgenden Beobachtung:

Ist x_1, \dots, x_n eine Lösung eines allgemeinen Gleichungssystems (4.1.1), so gilt für jedes i :

$$\sum_{j=1}^n a_{i,j} x_j = b_i$$

also

$$a_{i,i} \cdot x_i = b_i - \sum_{j \neq i} a_{i,j} \cdot x_j$$

und damit, falls $a_{i,i} \neq 0$,

$$x_i = \frac{1}{a_{i,i}} \cdot \left(b_i - \sum_{j \neq i} a_{i,j} \cdot x_j \right) \quad (4.1.3)$$

Die Überlegung des Mathematikers Carl Gustav Jacob Jacobi war nun, mit diesem Ansatz (4.1.3) aus einer ersten Schätzung der Lösung sukzessive Werte x_i zu bekommen, die immer näher an einer Lösung sind. Als Startwert wird dabei häufig $x_i^{(0)} = b_i$ gewählt. Für $k > 0$ setzen wir dann iterativ

$$x_i^{(k)} = \frac{1}{a_{i,i}} \cdot \left(b_i - \sum_{j \neq i} a_{i,j} \cdot x_j^{(k-1)} \right) \quad (4.1.4)$$

Dieses Verfahren führt nicht immer zu einer Lösung, funktioniert jedoch in einer Vielzahl von Fällen, die im Rahmen der sogenannten *finite-Elemente-Methode* auftreten.

Regel 4.1.5. *Erfüllt das Gleichungssystem (4.1.1) die Bedingung $n = m$ und*

$$\sum_{j \neq i} |a_{i,j}| < |a_{i,i}| \quad \text{für alle } i = 1, \dots, n$$

*(ein solches Gleichungssystem heißt **diagonaldominant**), so konvergiert das Jacobi-Verfahren (unabhängig von den gewählten Startwerten $x_1^{(0)}, \dots, x_n^{(0)}$),*

dh. die in (4.1.4) bestimmten $x_1^{(k)}, \dots, x_n^{(k)}$ nähern sich mit zunehmendem k einer Lösung des Gleichungssystem (4.1.1) an.

Bemerkung 4.1.8.

1. Die Bedingung der Diagonaldominanz stellt nicht nur sicher, dass wir immer durch $a_{i,i}$ teilen können, sie stellt auch sicher, dass das Gleichungssystem eine eindeutige Lösung hat.
2. In vielen Anwendungen konzentrieren sich die Koeffizienten des Gleichungssystems auf ein sehr enges Band um die Diagonale, dh. es gibt ein k , das sehr viel kleiner als n ist, mit

$$a_{i,j} = 0 \quad \text{falls } |j - i| > k$$

In diesem Fall müssen von dem Gleichungssystem nur die Koeffizienten $a_{i,j}$ mit $|j - i| \leq k$ gespeichert werden, was auch das Speicherplatzproblem adressiert.

3. Das Jacobi-Verfahren kann sehr gut parallelisiert werden, da in jedem Verarbeitungsschritt die $x_i^{(k)}$ unabhängig voneinander berechnet werden (sie hängen nur ab den dem Werten $x_i^{(k-1)}$ des Vorgängerschrittes).
4. Die Voraussetzungen von Regel 4.1.5 stellen sicher, dass für

$$\delta := \max \left\{ \frac{\sum_{j \neq i} |a_{i,j}|}{|a_{i,i}|} \mid i = 1, \dots, n \right\}$$

gilt: $\delta < 1$. Wie schnell das Verfahren einen guten Näherungswert für die Lösung liefert, hängt stark von diesem δ ab. Je näher δ an 1 liegt, desto langsamer verbessert sich die Lösung, je näher δ an der 0 ist, desto schneller nähern sich die Werte der tatsächlichen Lösung an.

Aufgabe 4.1.1. Bestimmen Sie alle Lösungen des linearen Gleichungssystems

$$\begin{array}{ccccccccc} x_1 & + & 2x_2 & + & 3x_3 & + & x_4 & = & 0 \\ 2x_1 & + & x_2 & + & 3x_3 & - & x_4 & = & 0 \\ 3x_1 & & & + & x_3 & + & x_4 & = & 0 \end{array}$$

Aufgabe 4.1.2. Bestimmen Sie alle Lösungen des linearen Gleichungssystems

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

Aufgabe 4.1.3. Überprüfen Sie, ob das lineare Gleichungssystem

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

Lösungen hat und bestimmen Sie diese gegebenenfalls.

Aufgabe 4.1.4. Überprüfen Sie, ob das lineare Gleichungssystem

$$\begin{aligned} x_1 + 2x_2 - x_3 &= 2 \\ 2x_1 + 3x_3 &= 0 \\ x_1 + 2x_2 + x_3 &= 1 \end{aligned}$$

Lösungen hat und bestimmen Sie diese gegebenenfalls.

4.1.2 Matrizen

Die Vorgehensweise beim Gauß-Algorithmus ist in der Notation ein wenig kompliziert und umständlich, da wir immer die Variablen mitführen, obwohl diese für den Algorithmus keine Rolle spielen. Mit der Matrizenschreibweise werden wir nun einen Formalismus kennenlernen, der es erlaubt, lineare Gleichungssysteme und damit zusammenhängende Probleme effizienter zu behandeln.

Definition 4.1.1. Eine **Matrix** A vom Typ (m, n) oder eine Matrix vom Typ $m \times n$ ist ein Schema der Form

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,j} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,j} & \dots & a_{2,n} \\ \vdots & & & \ddots & & \vdots \\ a_{i,1} & a_{i,2} & \dots & a_{i,j} & \dots & a_{i,n} \\ \vdots & & & \ddots & & \vdots \\ a_{m,1} & a_{m,2} & \dots & a_{m,j} & \dots & a_{m,n} \end{pmatrix}$$

bestehend aus m Zeilen und n Spalten von reelle Zahlen $a_{i,j}$.

$a_{i,j}$ heißt dabei das **Matrixelement** an der Stelle (i, j) , i heißt **Zeilenindex** von $a_{i,j}$ und j heißt **Spaltenindex** von $a_{i,j}$.

Wir schreiben kurz $A = (a_{i,j})_{(m,n)}$ oder - falls m und n klar sind - $A = (a_{i,j})$.

Zwei $m \times n$ Matrizen $A = (a_{i,j})$ und $B = (b_{i,j})$ heißen gleich, wenn

$$a_{i,j} = b_{i,j} \quad \text{für alle } i = 1, \dots, m; j = 1, \dots, n$$

Beispiel 4.1.1.

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

ist eine 2×4 -Matrix mit Einträgen

$$\begin{array}{cccc} a_{1,1} = 1 & a_{1,2} = 2 & a_{1,3} = 3 & a_{1,4} = 4 \\ a_{2,1} = 7 & a_{2,2} = 6 & a_{2,3} = 5 & a_{2,4} = 4 \end{array}$$

Beispiel 4.1.2. Die $m \times n$ -Matrix $A = (a_{i,j})$ mit $a_{i,j} = 0$ für alle i, j heißt **Nullmatrix**. Hierfür schreiben wir auch $0_{(m,n)}$ oder kurz 0 .

Beispiel 4.1.3. Eine $m \times 1$ -Matrix A ist ein Schema der Gestalt

$$A = \begin{pmatrix} a_{1,1} \\ \vdots \\ a_{m,1} \end{pmatrix}$$

entspricht also einem Spaltenvektor der Länge m .

Ein $1 \times n$ -Matrix B ist ein Schema der Gestalt

$$B = \begin{pmatrix} b_{1,1} & \dots & b_{1,n} \end{pmatrix}$$

entspricht also einem Zeilenvektor der Länge n .

Damit sind Vektoren Spezialfälle von Matrizen.

Definition 4.1.2. Ist $A = (a_{i,j})$ eine $m \times n$ -Matrix, so heißt

$$A_{i,\bullet} = \begin{pmatrix} a_{i,1} & a_{i,2} & \dots & a_{i,n} \end{pmatrix}$$

die i -te Zeile oder der i -te Zeilenvektor von A und

$$A_{\bullet,j} = \begin{pmatrix} a_{1,j} \\ a_{2,j} \\ \vdots \\ a_{m,j} \end{pmatrix}$$

die j -te Spalte oder der j -te Spaltenvektor von A .

Bezeichnung.:

Eine $n \times n$ -Matrix A heißt **quadratische Matrix** (der Größe n).

Beispiel 4.1.4. Die Matrix

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

ist eine quadratische Matrix.

Beispiel 4.1.5. Die quadratische $n \times n$ -Matrix $A = (a_{i,j})$ mit

$$a_{i,j} = \begin{cases} 1 & \text{falls } i = j \\ 0 & \text{falls } i \neq j \end{cases}$$

heißt $n \times n$ -Einheitsmatrix. Hierfür schreiben wir auch \mathbf{E}_n oder $\mathbf{1}_n$.

Die 2×2 -Einheitsmatrix hat also folgende Gestalt:

$$\mathbf{E}_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

In quadratischen Matrizen gibt es neben den Zeilen und Spalten noch weitere ausgezeichnete Gruppen von Matrixelementen, die Diagonale und die Gegendiagonale:

$$\begin{pmatrix} a_{1,1} & \dots & a_{1,n} \\ & \diagdown & \\ a_{n,1} & \dots & a_{n,n} \end{pmatrix} \quad \text{bzw.} \quad \begin{pmatrix} a_{1,1} & \dots & a_{1,n} \\ & \diagup & \\ a_{n,1} & \dots & a_{n,n} \end{pmatrix}$$

Die Diagonale umfasst also die Element $a_{1,1}, a_{2,2}, \dots, a_{n,n}$ und die Gegendiagonale die Elemente $a_{1,n}, a_{2,n-1}, \dots, a_{n,1}$.

Definition 4.1.3. Eine quadratische Matrix $A = (a_{i,j})$ heißt **Diagonalmatrix**, wenn alle Matrixelemente außerhalb der Diagonale verschwinden, d.h. $a_{i,j} = 0$ für $i \neq j$.

Eine quadratische Matrix $A = (a_{i,j})$ heißt **(obere) Dreiecksmatrix**, wenn nur auf und oberhalb der Diagonale von Null verschiedene Elemente stehen, d.h. $a_{i,j} = 0$ für $i > j$.

Entsprechend heißt eine Matrix $A = (a_{i,j})$ **untere Dreiecksmatrix**, wenn nur auf und unterhalb der Diagonale von Null verschiedene Elemente stehen, d.h. $a_{i,j} = 0$ für $i < j$.

Beispiel 4.1.6. Die Matrizen

$$E_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 3 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} -2 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 7 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}$$

sind Diagonalmatrizen, die Matrizen

$$E_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 2 & 1 & 3 \\ 0 & 4 & -3 \\ 0 & 0 & 9 \end{pmatrix}, \quad \begin{pmatrix} -2 & 0 & -1 & -4 \\ 0 & 3 & 3 & 0 \\ 0 & 0 & 7 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix}$$

sind (obere) Dreiecksmatrizen.

Beispiel 4.1.7. Jedes allgemeine lineare Gleichungssystem

$$\begin{array}{ccccccc} a_{1,1}x_1 & + & a_{1,2}x_2 & + & \cdots & + a_{1,n}x_n & = & b_1 \\ a_{2,1}x_1 & + & a_{2,2}x_2 & + & \cdots & + a_{2,n}x_n & = & b_2 \\ \vdots & & & & \ddots & & & \vdots \\ a_{m,1}x_1 & + & a_{m,2}x_2 & + & \cdots & + a_{m,n}x_n & = & b_m \end{array}$$

mit n Unbekannten und m -Gleichungen führt zu einer $m \times n$ -Matrix $A = (a_{i,j})$, der sogenannten **Koeffizientenmatrix** des Gleichungssystems, und einem Vektor (oder einer $m \times 1$ -Matrix) $\vec{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$. Darüberhinaus können wir auch die Unbekannten zu einem Vektor $\vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ der Länge n zusammenfassen.

Im Fall des linearen Gleichungssystems

$$\begin{array}{cccccc} 40x_1 & + & 50x_2 & + & 60x_3 & + & 70x_4 & = & 5500 \\ 26x_1 & + & 22x_2 & + & 25x_3 & + & 18x_4 & = & 2300 \\ 34x_1 & + & 28x_2 & + & 15x_3 & + & 12x_4 & = & 2200 \end{array}$$

aus Beispiel 4.1.1 ergibt das

$$A = \begin{pmatrix} 40 & 50 & 60 & 70 \\ 26 & 22 & 25 & 18 \\ 34 & 28 & 15 & 12 \end{pmatrix}, \quad \vec{b} = \begin{pmatrix} 5500 \\ 2300 \\ 2200 \end{pmatrix}$$

Umgekehrt definieren auch jede $m \times n$ -Matrix $A = (a_{i,j})$ und jeder Vektor $\vec{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$ der Länge m ein lineares Gleichungssystem in n -Unbekannten und mit m Gleichungen via

$$\begin{array}{ccccccccc} a_{1,1}x_1 & + & a_{1,2}x_2 & + & \cdots & + & a_{1,n}x_n & = & b_1 \\ a_{2,1}x_1 & + & a_{2,2}x_2 & + & \cdots & + & a_{2,n}x_n & = & b_2 \\ \vdots & & & & & & & & \vdots \\ a_{m,1}x_1 & + & a_{m,2}x_2 & + & \cdots & + & a_{m,n}x_n & = & b_m \end{array}$$

Definition 4.1.4. Die **transponierte Matrix** zu einer $m \times n$ -Matrix $A = (a_{i,j})$ ist die $n \times m$ -Matrix $A^T = (a_{i,j}^T)$ mit

$$a_{i,j}^T = a_{j,i} \quad \text{für alle } i = 1, \dots, n, j = 1, \dots, m$$

Bemerkung 4.1.1. A^T entsteht also aus A indem wir Zeilen und Spalten von A miteinander vertauschen.

Beispiel 4.1.8. Es sei

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

Dann ist

$$A^T = \begin{pmatrix} 1 & 6 \\ 2 & 5 \\ 3 & 4 \\ 4 & 3 \end{pmatrix}$$

Beispiel 4.1.9. Es sei

$$A = \begin{pmatrix} 2 & 3 \\ 4 & 5 \end{pmatrix}$$

Dann ist

$$A = \begin{pmatrix} 2 & 4 \\ 3 & 5 \end{pmatrix}$$

die Transponierte einer quadratischen Matrix ist also wieder eine quadratische Matrix der gleichen Größe.

Definition 4.1.5. Eine quadratische $n \times n$ -Matrix A heißt **symmetrisch**, wenn $A = A^T$.

Bemerkung 4.1.2. Eine quadratische Matrix ist also symmetrisch, wenn gilt

$$a_{i,j} = a_{j,i} \quad \text{für alle } i, j = 1, \dots, n$$

Beispiel 4.1.10. Die Matrix

$$A = \begin{pmatrix} 2 & 3 \\ 3 & 2 \end{pmatrix}$$

ist symmetrisch.

Bezeichnung:

Die Menge aller reellen $m \times n$ -Matrizen bezeichnen wir $\text{Matr}(m, n)$ oder $\text{Matr}(m, n; \mathbb{R})$.

Mit Hilfe der Matrizenschreibweise lässt sich nun der Gaußsche Eliminationsalgorithmus zum Lösen linearer Gleichungssystem einfacher hinschreiben:

Wie wir gesehen haben, liefert jedes lineare Gleichungssystem

$$\begin{array}{ccccccc} a_{1,1}x_1 & + & a_{1,2}x_2 & + & \cdots & + a_{1,n}x_n & = & b_1 \\ a_{2,1}x_1 & + & a_{2,2}x_2 & + & \cdots & + a_{2,n}x_n & = & b_2 \\ \vdots & & & & \ddots & & & \vdots \\ a_{m,1}x_1 & + & a_{m,2}x_2 & + & \cdots & + a_{m,n}x_n & = & b_m \end{array}$$

eine $m \times n$ -Matrix $A = (a_{i,j})$ und einen Spaltenvektor $\vec{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$ der Länge m . Daraus bilden wir die **augmentierte Matrix**

$$(A | \vec{b}) = \left(\begin{array}{ccc|c} a_{1,1} & \dots & a_{1,n} & b_1 \\ a_{2,1} & \dots & a_{2,n} & b_2 \\ \vdots & \ddots & & \vdots \\ a_{m,1} & \dots & a_{m,n} & b_m \end{array} \right)$$

Die augmentierte Matrix $(A | \vec{b})$ ist also ein $m \times (n+1)$ -Matrix, die aus A dadurch entsteht, dass wir ein Spalte, bestehend aus den Elementen von \vec{b} anfügen. In Analogie zu linearen Gleichungssystemen definieren wir:

Definition 4.1.6. Die augmentierte Matrix $(A | \vec{b})$ liegt in **(Gaußscher) Normalform** vor, wenn es ein $t \in \{1, \dots, m\}$ gibt mit

- $a_{i,j} = 0$ für alle $i > t$ und alle j .
- $a_{i,j} = 0$ für $i \in \{2, \dots, t\}$ und $j < i$.
- $a_{i,i} = 1$ für $i = 1, \dots, t$.

Eine augmentierte Matrix in Normalform hat also die Gestalt

$$\left(\begin{array}{cccccc|c} 1 & a_{1,2} & \dots & a_{1,t} & \dots & a_{1,n} & b_1 \\ 0 & 1 & \dots & a_{2,t} & \dots & a_{2,n} & b_2 \\ \vdots & & \ddots & & & & \vdots \\ 0 & 0 & \dots & 1 & \dots & a_{t,n} & b_t \\ 0 & 0 & \dots & 0 & \dots & 0 & b_{t+1} \\ \vdots & & \ddots & & & & \vdots \\ 0 & 0 & \dots & 0 & \dots & 0 & b_m \end{array} \right)$$

und zu einer augmentierten Matrix in Normalform gehört ein lineares Gleichungssystem

chungssystem in Normalform, nämlich

$$\begin{array}{cccccccc}
 x_1 & + & a_{1,2}x_2 & + & a_{1,3}x_3 & + & \cdots & + & a_{1,n}x_n & = & b_1 \\
 0 & + & x_2 & + & a_{2,3}x_3 & + & \cdots & + & a_{2,n}x_n & = & b_2 \\
 \vdots & & & & & & \ddots & & & & \vdots \\
 0 & + & 0 & + & 0 & \cdots & + x_t & + \cdots & + a_{t,n}x_n & = & b_t \\
 0 & + & 0 & + & 0 & + & \cdots & + & 0 & = & b_{t+1} \\
 \vdots & & & & & & \ddots & & & & \vdots \\
 0 & + & 0 & + & 0 & + & \cdots & + & 0 & = & b_m
 \end{array}$$

Satz 4.1.6. Wir betrachten die folgenden vier Operationen auf einer augmentierten Matrix $(A | \vec{b})$

- (i) Subtrahiere das Vielfache einer Zeile von $(A | \vec{b})$ von einer anderen Zeile
- (ii) Vertausche zwei Zeilen von $(A | \vec{b})$
- (iii) Vertausche zwei Spalten von $(A | \vec{b})$, von denen keine die letzte sein darf.
- (iv) Multipliziere eine Zeile von $(A | \vec{b})$ mit einer Zahl $r \neq 0$.

Jede augmentierte Matrix $(A | \vec{b})$ lässt sich durch wiederholtes Anwenden der Operationen (i) - (iv) auf Normalform bringen. Das zu dieser Normalform gehörige lineare Gleichungssystem ist eine Normalform des ursprünglichen linearen Gleichungssystems.

Beweis: Der Beweis ist identisch mit dem Beweis Satz 4.1.2.

Beispiel 4.1.11. Wir betrachten das lineare Gleichungssystem

$$\begin{array}{rcl}
 x_1 & + & 2x_2 = 3 \\
 2x_1 & + & 3x_2 = 4
 \end{array}$$

Hierzu gehört die augmentierte Matrix

$$(A | \vec{b}) = \left(\begin{array}{cc|c} 1 & 2 & 3 \\ 2 & 3 & 4 \end{array} \right)$$

Subtrahieren wir das Doppelte der ersten Zeile von der zweiten, so erhalten wir

$$(A'|\vec{b}') = \left(\begin{array}{cc|c} 1 & 2 & 3 \\ 0 & -1 & -2 \end{array} \right)$$

Multiplizieren wir die zweite Zeile mit -1, so ergibt sich

$$(A''|\vec{b}'') = \left(\begin{array}{cc|c} 1 & 2 & 3 \\ 0 & 1 & 2 \end{array} \right)$$

also eine augmentierte Matrix in Normalform. Das zugehörige lineare Gleichungssystem ist

$$\begin{array}{rcrcrcrcrcrcl} x_1 & + & 2x_2 & = & 3 \\ & & x_2 & = & 2 \end{array}$$

also ein Gleichungssystem in Normalform. Seine eindeutige Lösung ist

$$x_1 = -1, \quad x_2 = 2$$

und das ist auch die Lösung des ursprünglichen Gleichungssystems

Bemerkung 4.1.3. Die Normalform einer augmentierten Matrix $(A|\vec{b})$ ist nicht eindeutig, genauso wie die Normalform eines Gleichungssystems nicht eindeutig war, und nicht alle Normalformen sind gleich einfach zu behandeln. Wir können aber immer eine Normalform $(A'|\vec{b}')$ erreichen, die die Gestalt

$$(A|\vec{b}) = \left(\begin{array}{cccccc|c} 1 & 0 & \dots & 0 & a'_{1,t+1} & \dots & a'_{1,n} & b'_1 \\ 0 & 1 & \dots & 0 & a'_{2,t+1} & \dots & a'_{2,n} & b'_2 \\ \vdots & & & & \ddots & & & \\ 0 & 0 & \dots & 1 & a'_{t,t+1} & \dots & a'_{t,n} & b'_t \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 & b'_{t+1} \\ \vdots & & & & \ddots & & & \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 & b'_m \end{array} \right)$$

hat. Dazu müssen wir lediglich, nach Erreichen einer Normalform, ein geeignetes Vielfaches der zweiten Zeile von der ersten Zeile abziehen, dann geeignete Vielfache der dritten Zeile von der ersten und der zweiten, usw. Diese Art der Normalform hat den Vorteil, dass sich die Lösungen des zugehörigen Gleichungssystems

$$A' \cdot \vec{x} = \vec{b}'$$

(falls sie existieren), und damit auch (bis auf Vertauschungen) die Lösungen des Ausgangssystems sofort aus der augmentierten Matrix ablesen lassen:

ist $x_{t+1} = \lambda_{t+1}, \dots, x_n = \lambda_n$ die Belegung der freien Variablen, so berechnet sich die komplette Lösung als

$$\begin{aligned} x_1 &= b'_1 - a'_{1,t+1}\lambda_{t+1} - \dots - a'_{1,n}\lambda_n \\ x_2 &= b'_2 - a'_{2,t+1}\lambda_{t+1} - \dots - a'_{2,n}\lambda_n \\ &\vdots \\ x_t &= b'_t - a'_{t,t+1}\lambda_{t+1} - \dots - a'_{t,n}\lambda_n \end{aligned}$$

Ist speziell auch noch $t = m = n$ so erhält man direkt die (eindeutige) Lösung als

$$x_1 = b'_1, x_2 = b'_2, \dots, x_m = b'_m$$

Beispiel 4.1.12. Wir betrachten nochmal das Gleichungssystem

$$\begin{array}{rcl} x_1 & + & 2x_2 = 3 \\ 2x_1 & + & 3x_2 = 4 \end{array}$$

aus Beispiel 4.1.11. Wir haben schon die Normalform

$$(A'' | \vec{b}'') = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \end{pmatrix}$$

erhalten. Indem wir jetzt noch das Doppelte der zweiten Zeile von der ersten Zeile subtrahieren, erhalten wir

$$(A''' | \vec{b}''') = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \end{pmatrix}$$

und lesen sofort die eindeutige Lösung $x_1 = -1$ und $x_2 = 2$ ab.

Bemerkung 4.1.4. Die Methode mit den augmentierten Matrizen eignet sich auch ausgezeichnet, um mehrere Gleichungssysteme (mit denselben Koeffizienten) simultan zu lösen. Betrachten wir etwa die Gleichungen

$$A \cdot \vec{x} = \vec{b}_1, \quad A \cdot \vec{x} = \vec{b}_2$$

so können wir daraus die (erweiterte) augmentierte Matrix $(A | \vec{b}_1 \vec{b}_2)$ bilden und diese nach den Methoden von Satz 4.1.6 auf Normalform bringen

(wobei hier lediglich zu beachten ist, dass bei den Vertauschungsoperationen die letzten beiden Spalten ausgeschlossen sind). Ist dann $\left(A'|\vec{b}_1\vec{b}_2\right)$ eine Normalform dieser erweiterten augmentierten Matrix, so sind $\left(A'|\vec{b}_1\right)$ und $\left(A'|\vec{b}_2\right)$ Normalformen der jeweiligen einfachen augmentierten Matrizen.

Bemerkung 4.1.5. Wie bei linearen Gleichungssystemen können auch bei Matrizen komplexe Koeffizienten auftreten. Betrachten wir etwa wieder das Gleichungssystem aus Bemerkung 4.1.6 im letzten Abschnitt, so erhalten wir die Koeffizientenmatrix

$$A = \begin{pmatrix} 1 & i \\ 1+i & -2 \end{pmatrix}$$

und die augmentierte Matrix

$$\left(A|\vec{b}\right) = \begin{pmatrix} 1 & i & 1+i \\ 1+i & -2 & 1+3 \cdot i \end{pmatrix}$$

Alle Operationen, die wir mit reellen Matrizen durchgeführt haben, lassen sich auch auf komplexe Matrizen ausdehnen, und komplexe Gleichungssysteme lassen sich mit Hilfe der augmentierten Matrix lösen.

Aufgabe 4.1.5. Bestimmen Sie die Transponierte von

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

Aufgabe 4.1.6. Bestimmen Sie die Transponierte von

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

Ist A symmetrisch?

Aufgabe 4.1.7. Wir betrachten das lineare Gleichungssystem

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

Bestimmen Sie die Koeffizientenmatrix und die augmentierte Matrix dieses Gleichungssystems, überprüfen Sie, ob es Lösungen hat, und bestimmen Sie diese gegebenenfalls.

Aufgabe 4.1.8. Wir betrachten das lineare Gleichungssystems

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

Bestimmen Sie die Koeffizientenmatrix und die augmentierte Matrix dieses Gleichungssystems, überprüfen Sie, ob es Lösungen hat, und bestimmen Sie diese gegebenenfalls.

Aufgabe 4.1.9. Wir betrachten das lineare Gleichungssystems

$$\begin{array}{cccccccl} i \cdot x_1 & + & (1+i) \cdot x_2 & + & (1-i) \cdot x_3 & = & 2 \\ 2 \cdot x_1 & & & + & 3 \cdot i \cdot x_3 & = & 1 + 2 \cdot i \\ & & i \cdot x_2 & + & x_3 & = & 3 \cdot i \end{array}$$

Bestimmen Sie die Koeffizientenmatrix und die augmentierte Matrix dieses Gleichungssystems, überprüfen Sie, ob es Lösungen hat, und bestimmen Sie diese gegebenenfalls.

Aufgabe 4.1.10. Bestimmen Sie alle Zahlen $a \in \mathbb{R}$, für die das Gleichungssystem

$$\begin{array}{cccccccl} x_1 & + & x_2 & - & x_3 & + & x_4 & = & 2 \\ 2x_1 & + & 2x_2 & - & x_3 & - & x_4 & = & 1 \\ -x_1 & + & x_2 & - & x_3 & + & x_4 & = & 2 \\ x_1 & + & 3x_2 & + & a \cdot x_3 & + & a \cdot x_4 & = & 1 \end{array}$$

eine Lösung hat.

4.2 Matrizenoperationen

Wichtig wird es für uns sei, mit Matrizen zu rechnen, ähnlich wie wir das schon mit Vektoren können. Bei der Definition der Operationen gehen wir auch so vor wie bei den Vektoren im \mathbb{R}^n .

Definition 4.2.1. Es seien $A = (a_{i,j})$, $B = (b_{i,j}) \in \text{Matr}(m, n)$ und $\lambda \in \mathbb{R}$. Wir definieren die Matrizenaddition durch

$$A + B = (a_{i,j} + b_{i,j}) = \begin{pmatrix} a_{1,1} + b_{1,1} & a_{1,2} + b_{1,2} & \dots & a_{1,n} + b_{1,n} \\ a_{2,1} + b_{2,1} & a_{2,2} + b_{2,2} & \dots & a_{2,n} + b_{2,n} \\ \vdots & & \ddots & \vdots \\ a_{m,1} + b_{m,1} & a_{m,2} + b_{m,2} & \dots & a_{m,n} + b_{m,n} \end{pmatrix}$$

$A + B$ ist also wieder eine $n \times m$ -Matrix, die in der i -ten Zeile und der j -ten Spalte den Eintrag $a_{i,j} + b_{i,j}$ hat. Analog definieren wir Matrixsubtraktion durch

$$A - B = (a_{i,j} - b_{i,j}) = \begin{pmatrix} a_{1,1} - b_{1,1} & a_{1,2} - b_{1,2} & \dots & a_{1,n} - b_{1,n} \\ a_{2,1} - b_{2,1} & a_{2,2} - b_{2,2} & \dots & a_{2,n} - b_{2,n} \\ \vdots & & \ddots & \vdots \\ a_{m,1} - b_{m,1} & a_{m,2} - b_{m,2} & \dots & a_{m,n} - b_{m,n} \end{pmatrix}$$

Wir definieren die Multiplikation einer Matrix mit einem Skalar durch

$$\lambda \cdot A = (\lambda \cdot a_{i,j}) = \begin{pmatrix} \lambda \cdot a_{1,1} & \lambda \cdot a_{1,2} & \dots & \lambda \cdot a_{1,n} \\ \lambda \cdot a_{2,1} & \lambda \cdot a_{2,2} & \dots & \lambda \cdot a_{2,n} \\ \vdots & & \ddots & \vdots \\ \lambda \cdot a_{m,1} & \lambda \cdot a_{m,2} & \dots & \lambda \cdot a_{m,n} \end{pmatrix}$$

$\lambda \cdot A$ ist also wieder eine $n \times m$ -Matrix, die in der i -ten Zeile und der j -ten Spalte den Eintrag $\lambda \cdot a_{i,j}$ hat.

Beispiel 4.2.1. Für

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

gilt

$$A + B = \begin{pmatrix} 1+2 & 2+1 & 3+3 & 4+2 \\ 5+1 & 6+4 & 5+1 & 4+5 \end{pmatrix} = \begin{pmatrix} 3 & 3 & 6 & 6 \\ 6 & 10 & 6 & 9 \end{pmatrix}$$

und

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

Beispiel 4.2.2. Für

$$\lambda = 3, \quad A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 5 & 4 \end{pmatrix}$$

gilt

$$\lambda \cdot A = \begin{pmatrix} 3 \cdot 1 & 3 \cdot 2 & 3 \cdot 3 & 3 \cdot 4 \\ 3 \cdot 5 & 3 \cdot 6 & 3 \cdot 5 & 3 \cdot 4 \end{pmatrix} = \begin{pmatrix} 3 & 6 & 9 & 12 \\ 15 & 18 & 15 & 12 \end{pmatrix}$$

Definition 4.2.2. Die negative Matrix $-A$ zu einer Matrix $A = (a_{i,j})$ ist definiert als

$$-A = (-a_{i,j})$$

also $-A = (-1) \cdot A$.

Für das Rechnen mit Matrizen gelten die folgenden Regeln:

Regel 4.2.1. Für $m \times n$ -Matrizen A, B, C und Skalare r, s gilt:

- $(A + B) + C = A + (B + C)$ (1. Assoziativgesetz)
- $A + B = B + A$ (Kommutativgesetz)
- $A + 0_{(m,n)} = A$ (neutrales Element Addition)
- $A + (-A) = 0_{(m,n)}$ (inverses Element Addition)
- $(r \cdot s) \cdot A = r \cdot (s \cdot A)$ (2. Assoziativgesetz)
- $r \cdot (A + B) = r \cdot A + r \cdot B$ (1. Distributivgesetz)
- $(r + s) \cdot A = r \cdot A + s \cdot A$ (2. Distributivgesetz)
- $1 \cdot A = A$ (neutrales Element Skalarmultiplikation)

Folgerung 4.2.2. Die Menge $\text{Matr}(m, n)$ zusammen mit der Matrizenaddition und der Skalarmultiplikation von Matrizen ist ein Vektorraum.

Mit $E_{l,k}$ bezeichnen wir die $m \times n$ -Matrix mit den Matrixelementen

$$a_{i,j} = \begin{cases} 1 & \text{falls } i = l, j = k \\ 0 & \text{sonst} \end{cases}$$

Beispiel 4.2.3. Für $m = 2, n = 2$ gilt

$$E_{1,1} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, E_{1,2} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, E_{2,1} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, E_{2,2} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

Satz 4.2.3. Die Matrizen $\{E_{l,k}\}_{l=1,\dots,m, k=1,\dots,n}$ bilden eine Basis von $\text{Matr}(m,n)$. Insbesondere gilt also

$$\dim(\text{Matr}(m,n)) = m \cdot n$$

Beweis: D as lässt sich leicht nachrechnen.

Wir haben nun auf den Matrizen Vektorraumoperationen eingeführt. Darüberhinaus gibt es noch eine weitere, sehr interessante Operation auf Matrizen:

Definition 4.2.3. Für eine $m \times n$ -Matrix $A = (a_{i,j})$ und eine $n \times l$ -Matrix $B = (b_{i,j})$ definieren wir das **Matrizenprodukt** $A \cdot B$ von A und B als diejenige $m \times l$ -Matrix $C = (c_{i,j})$ mit

$$c_{i,j} = a_{i,1} \cdot b_{1,j} + a_{i,2} \cdot b_{2,j} + \cdots + a_{i,n} \cdot b_{n,j} = \sum_{k=1}^n a_{i,k} \cdot b_{k,j}$$

Bemerkung 4.2.1. Der entscheidende Punkt bei der Matrizenmultiplikation von A mit B ist die Tatsache, dass die Anzahl der Spalten von A mit der Anzahl der Zeilen von B übereinstimmt.

Beispiel 4.2.4. Für die Matrizen

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

gilt

$$\begin{aligned} A \cdot B &= \begin{pmatrix} 1 \cdot 2 + 2 \cdot 3 + 3 \cdot (-1) & 1 \cdot 1 + 2 \cdot 1 + 3 \cdot (-2) & 1 \cdot 4 + 2 \cdot 2 + 3 \cdot 1 \\ 4 \cdot 2 + 2 \cdot 3 + 0 \cdot (-1) & 4 \cdot 1 + 2 \cdot 1 + 0 \cdot (-2) & 4 \cdot 4 + 2 \cdot 2 + 0 \cdot 1 \end{pmatrix} \\ &= \begin{pmatrix} 5 & -3 & 11 \\ 14 & 6 & 20 \end{pmatrix} \end{aligned}$$

Beispiel 4.2.5. Für die Matrizen

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

gilt

$$\begin{aligned} A \cdot B &= \begin{pmatrix} 2 \cdot 1 + 1 \cdot 3 + 3 \cdot (-1) & 2 \cdot 4 + 1 \cdot 2 + 3 \cdot (-2) \\ 1 \cdot 1 + 2 \cdot 3 + 0 \cdot (-1) & 1 \cdot 4 + 2 \cdot 2 + 0 \cdot (-2) \end{pmatrix} \\ &= \begin{pmatrix} 2 & 4 \\ 7 & 8 \end{pmatrix} \end{aligned}$$

Beachten Sie, dass wir in diesem Beispiel auch $B \cdot A$ bilden können. Hierfür erhalten wir

$$\begin{aligned} B \cdot A &= \begin{pmatrix} 1 \cdot 2 + 4 \cdot 1 & 1 \cdot 1 + 4 \cdot 2 & 1 \cdot 3 + 4 \cdot 0 \\ 3 \cdot 2 + 2 \cdot 1 & 3 \cdot 1 + 2 \cdot 2 & 3 \cdot 3 + 2 \cdot 0 \\ (-1) \cdot 2 + (-2) \cdot 1 & (-1) \cdot 1 + (-2) \cdot 2 & (-1) \cdot 3 + (-2) \cdot 0 \end{pmatrix} \\ &= \begin{pmatrix} 6 & 5 & 3 \\ 8 & 7 & 9 \\ -4 & -3 & -3 \end{pmatrix} \end{aligned}$$

Dieses Beispiel zeigt, dass im allgemeinen $A \cdot B \neq B \cdot A$, selbst wenn beide Multiplikationen definiert sind. Es stimmen weder die Größen der resultierenden Ergebnissen noch deren Einträge überein.

Bemerkung 4.2.2. Fassen wir einen Vektor \vec{v} als eine $n \times 1$ -Matrix auf, so können wir eine $m \times n$ -Matrix A mit \vec{v} multiplizieren. Das Ergebnis $A \cdot \vec{v}$ ist eine $m \times 1$ -Matrix, also ein Vektor der Länge m .

Ist $\vec{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \dots, \vec{e}_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$ die Standardbasis des \mathbb{R}^n , so gilt:

$$A \cdot \vec{e}_j = A_{\bullet, j} \quad \text{für alle } j \in \{1, \dots, n\}$$

ist der j -te Spaltenvektor von A . Das folgt sofort aus der Definition der Matrizenmultiplikation.

Beispiel 4.2.6.

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 6 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \cdot 0 + 2 \cdot 0 + 3 \cdot 1 + 4 \cdot 0 \\ 2 \cdot 0 + 3 \cdot 0 + 4 \cdot 1 + 5 \cdot 0 \\ 3 \cdot 0 + 4 \cdot 0 + 5 \cdot 1 + 6 \cdot 0 \end{pmatrix} = \begin{pmatrix} 3 \\ 4 \\ 5 \end{pmatrix}$$

ist die dritte Spalte von A .

Für die Matrizenmultiplikation gelten eine Reihe einfacher Regeln:

Regel 4.2.4. 1. Für eine $m \times n$ -Matrix A , eine $n \times l$ -Matrix B und eine $l \times k$ -Matrix C gilt

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

2. Für eine $m \times n$ -Matrix A , eine $n \times l$ -Matrix B und einen Skalar $\lambda \in \mathbb{R}$ gilt:

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

3. Für eine $m \times n$ -Matrix A , eine $n \times l$ -Matrix B und eine $n \times l$ -Matrix C gilt

$$A \cdot (B + C) = A \cdot B + A \cdot C$$

4. Für eine $m \times n$ -Matrix A , eine $m \times n$ -Matrix B und eine $n \times l$ -Matrix C gilt

$$(A + B) \cdot C = A \cdot C + B \cdot C$$

5. Für eine $m \times n$ -Matrix A , eine $n \times l$ -Matrix B gilt

$$(A \cdot B)^T = B^T \cdot A^T$$

6. Für eine $m \times n$ -Matrix A gilt

$$A \cdot E_n = A = E_m \cdot A$$

wobei E_n die $n \times n$ -Einheitsmatrix und E_m die $m \times m$ -Einheitsmatrix aus Beispiel 4.1.5 bezeichnet.

Folgerung 4.2.5. Ist $A = (a_{i,j})$ eine $m \times n$ -Matrix mit den Spaltenvektoren $\vec{a}_1 = A_{\bullet,1}, \dots, \vec{a}_n = A_{\bullet,n}$, und ist $\vec{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$ ein n -Vektor, so gilt

$$A \cdot \vec{v} = v_1 \cdot \vec{a}_1 + \dots + v_n \cdot \vec{a}_n$$

Beweis: Wir können schreiben

$$\vec{v} = v_1 \cdot \vec{e}_1 + \dots + v_n \cdot \vec{e}_n$$

wobei $\vec{e}_1, \dots, \vec{e}_n$ die Standardbasis des \mathbb{R}^n bezeichnet. Dann gilt nach Regeln 4.2.4 (2) und (3):

$$\begin{aligned} A \cdot \vec{v} &= A \cdot (v_1 \cdot \vec{e}_1 + \dots + v_n \cdot \vec{e}_n) \\ &= v_1 \cdot A \cdot \vec{e}_1 + \dots + v_n \cdot A \cdot \vec{e}_n \\ &= v_1 \cdot \vec{a}_1 + \dots + v_n \cdot \vec{a}_n \end{aligned}$$

Eine interessante Beziehung erhalten wir zwischen Matrizenprodukt und Skalarprodukt:

Regel 4.2.6. Ist A eine $n \times n$ -Matrix und sind \vec{v} und \vec{w} zwei n -Vektoren, so gilt

$$\langle A \cdot \vec{v}, \vec{w} \rangle = \langle \vec{v}, A^T \cdot \vec{w} \rangle$$

Ist also speziell A eine symmetrische Matrix, ist also $A^T = A$, so gilt

$$\langle A \cdot \vec{v}, \vec{w} \rangle = \langle \vec{v}, A \cdot \vec{w} \rangle$$

Beweis: Ist $A = (a_{i,j})$, und schreiben wir $\vec{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$ und $\vec{w} = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}$ so gilt

$$A \cdot \vec{v} = \begin{pmatrix} \sum_{i=1}^n a_{1,i} \cdot v_i \\ \vdots \\ \sum_{i=1}^n a_{n,i} \cdot v_i \end{pmatrix}$$

also

$$\langle A \cdot \vec{v}, \vec{w} \rangle = \sum_{j=1}^n \sum_{i=1}^n a_{j,i} \cdot v_i \cdot w_j$$

Andererseits ist $A^T = (a_{j,i})$, und daher

$$A \cdot \vec{w} = \begin{pmatrix} \sum_{j=1}^n a_{j,1} \cdot w_j \\ \vdots \\ \sum_{j=1}^n a_{j,n} \cdot w_j \end{pmatrix}$$

also

$$\langle \vec{v}, A^T \cdot \vec{w} \rangle = \sum_{j=1}^n \sum_{i=1}^n a_{j,i} \cdot w_j \cdot v_i = \langle A \cdot \vec{v}, \vec{w} \rangle$$

wie gewünscht.

Bemerkung 4.2.3. Genau dann ist eine $n \times n$ -Matrix symmetrisch, wenn für alle n -Vektoren \vec{v} und \vec{w} gilt

$$\langle A \cdot \vec{v}, \vec{w} \rangle = \langle \vec{v}, A \cdot \vec{w} \rangle$$

Wir wissen nämlich schon aus Regel 4.2.6, dass

$$\langle A \cdot \vec{v}, \vec{w} \rangle = \langle \vec{v}, A^T \cdot \vec{w} \rangle$$

Ist also A symmetrisch, so gilt $A = A^T$ und diese Beziehung wird

$$\langle A \cdot \vec{v}, \vec{w} \rangle = \langle \vec{v}, A \cdot \vec{w} \rangle$$

wie gewünscht.

Gilt nun umgekehrt diese Beziehung für alle Vektoren, so gilt sie speziell auch für die Einheitsvektoren, also

$$\langle A \cdot \vec{e}_i, \vec{e}_j \rangle = \langle \vec{e}_i, A \cdot \vec{e}_j \rangle \quad (4.2.1)$$

für alle $i, j \in \{1, \dots, n\}$. Nun gilt aber

$$\begin{aligned} \langle A \cdot \vec{e}_i, \vec{e}_j \rangle &= a_{j,i} \\ \langle \vec{e}_i, A \cdot \vec{e}_j \rangle &= a_{i,j} \end{aligned}$$

so dass aus Gleichung 4.2.1 folgt, dass A symmetrisch ist.

Bemerkung 4.2.4. Addition und Multiplikation sind auch für komplexe Matrizen definiert.

Aufgabe 4.2.1. Gegeben seien die Matrizen

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

Bestimmen Sie A^T , B^T , $A \cdot B$ und $B^T \cdot A$.

Aufgabe 4.2.2. Gegeben seien die Matrizen

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

Bestimmen Sie A^T , B^T , $A + B$, $A \cdot B$ und $B \cdot A$.

Aufgabe 4.2.3. Gegeben seien die Matrizen

$$A = \begin{pmatrix} 1+i & 1-2 \cdot i \\ 4 & 3 \cdot i \end{pmatrix} \quad B = \begin{pmatrix} 3 \cdot i & 2+i & 1 \\ 5-i & i-4 & -2-i \end{pmatrix}$$

Bestimmen Sie $A \cdot B$ und $B^T \cdot A$.

Aufgabe 4.2.4. Gegeben seien die Matrizen

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & -1 & 1 \\ -2 & 3 & -4 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 1 \\ 2 & 1 \\ 1 & 2 \end{pmatrix}$$

Berechnen Sie (falls der Ausdruck existiert):

a) $A \cdot B$

b) $B \cdot A$

c) $B \cdot C$

d) $B^T \cdot A$

e) $B^T + C$

4.3 Matrizen und Abbildungen

Wir haben schon gesehen, dass sich lineare Gleichungssysteme mit Matrizen und augmentierten Matrizen beschreiben lassen. In diesem Abschnitt wollen wir diese Beziehungen weiter studieren und vertiefen.

Bemerkung 4.3.1. Fassen wir einen Vektor der Länge n als eine $n \times 1$ -Matrix auf, so können wir insbesondere eine $m \times n$ -Matrix A mit einem Vektor \vec{v} der Länge n multiplizieren. Das Ergebnis $A \cdot \vec{v}$ ist eine $m \times 1$ -Matrix, also ein Vektor der Länge m . Damit definiert A eine Abbildung

$$f = f_A : \mathbb{R}^n \longrightarrow \mathbb{R}^m, \quad \vec{v} \longmapsto A \cdot \vec{v}$$

Aus Regel 4.2.4 folgt sofort, dass diese Abbildung f_A folgende interessante Eigenschaften hat:

(i) Für $\vec{v}, \vec{w} \in \mathbb{R}^n$ gilt:

$$f_A(\vec{v} + \vec{w}) = f_A(\vec{v}) + f_A(\vec{w})$$

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

$$f_A(\lambda \cdot \vec{v}) = \lambda \cdot f_A(\vec{v})$$

Abbildungen, die die beiden Eigenschaften aus Bemerkung 4.3.1 erfüllen, sind von besonderem Interesse in der Mathematik

Definition 4.3.1. Eine Abbildung $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ heißt **lineare Abbildung** von \mathbb{R}^n nach \mathbb{R}^m , wenn sie die folgenden beiden Eigenschaften hat

(i) Für $\vec{v}, \vec{w} \in \mathbb{R}^n$ gilt:

$$f_A(\vec{v} + \vec{w}) = f_A(\vec{v}) + f_A(\vec{w})$$

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

$$f_A(\lambda \cdot \vec{v}) = \lambda \cdot f_A(\vec{v})$$

Bemerkung 4.3.2. Wir haben gesehen, dass jede $m \times n$ -Matrix A eine lineare Abbildung f_A von \mathbb{R}^n nach \mathbb{R}^m definiert. Umgekehrt gilt aber auch, dass jede lineare Abbildung f von \mathbb{R}^n nach \mathbb{R}^m auf diese Art und Weise von einer $m \times n$ -Matrix A kommt, d.h. $f = f_A$ für eine $m \times n$ -Matrix A .

Dazu sei $\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n$ die Standardbasis des \mathbb{R}^n aus Beispiel 3.2.19 in Abschnitt 3.2. Dann ist $f(\vec{e}_i)$ ein Vektor in \mathbb{R}^m für alle $i \in \{1, \dots, n\}$, wir können also schreiben

$$f(\vec{e}_i) = \begin{pmatrix} a_{1,i} \\ \vdots \\ a_{m,i} \end{pmatrix}$$

und bilden daraus die Matrix

$$A = (a_{i,j}) = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} \end{pmatrix}$$

in der also die $f(\vec{e}_i)$ gerade die i -ten Spalten bilden. Dann erhalten wir für einen beliebigen Vektor $\vec{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \in \mathbb{R}^n$ aufgrund der Linearität von f :

$$\begin{aligned} f(\vec{v}) &= f(v_1 \cdot \vec{e}_1 + \dots + v_n \cdot \vec{e}_n) \\ &= v_1 \cdot f(\vec{e}_1) + \dots + v_n \cdot f(\vec{e}_n) \\ &= v_1 \cdot \begin{pmatrix} a_{1,1} \\ \vdots \\ a_{m,1} \end{pmatrix} + \dots + v_n \cdot \begin{pmatrix} a_{1,n} \\ \vdots \\ a_{m,n} \end{pmatrix} \\ &= \begin{pmatrix} v_1 \cdot a_{1,1} + \dots + v_n \cdot a_{1,n} \\ \vdots \\ v_1 \cdot a_{m,1} + \dots + v_n \cdot a_{m,n} \end{pmatrix} \\ &= A \cdot \vec{v} \end{aligned}$$

und damit ist gezeigt, dass $f = f_A$.

A heißt **darstellende Matrix** von f (bezüglich der Standardbasis).

Zu jeder linearen Abbildung gehört auf diese Art und Weise also eine eindeutige Matrix, und umgekehrt bestimmt auch jede Matrix genau eine lineare Abbildung. Matrizen und lineare Abbildungen entsprechen sich also eindeutig.

Wie wir schon gesehen haben, gibt jedes allgemeine lineare Gleichungssystem

$$\begin{array}{ccccccc} a_{1,1}x_1 & + & a_{1,2}x_2 & + & \cdots & + a_{1,n}x_n & = & b_1 \\ a_{2,1}x_1 & + & a_{2,2}x_2 & + & \cdots & + a_{2,n}x_n & = & b_2 \\ \vdots & & & & \ddots & & & \vdots \\ a_{m,1}x_1 & + & a_{m,2}x_2 & + & \cdots & + a_{m,n}x_n & = & b_m \end{array} \quad (4.3.1)$$

Anlaß zu einer $m \times n$ -Koeffizientenmatrix $A = (a_{i,j})$, einem Spaltenvektor $\vec{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$ der Länge m und einem Vektor von Unbekannten $\vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ der Länge n . wir rechnen leicht nach, dass sich linke Seite des Gleichungssystems 4.3.1 nichts anderes ist als $A \cdot \vec{x}$. Damit schreibt sich das Gleichungssystem 4.3.1 kompakt als

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

Im Sinne von Bemerkung 4.3.1 ist also das Finden der Lösungsmenge von Gleichungssystem 4.3.1 äquivalent zum Finden der Urbildmenge $f_A^{-1}(\vec{b})$.

Definition 4.3.2. Für eine $m \times n$ -Matrix A nennen wir

$$\text{Ker}(A) := \{ \vec{v} \in \mathbb{R}^n \mid A \cdot \vec{v} = \vec{0} \}$$

den **Kern** von A und

$$\text{Im}(A) := \{ \vec{b} \in \mathbb{R}^m \mid \exists \vec{v} \in \mathbb{R}^n \text{ mit } A \cdot \vec{v} = \vec{b} \}$$

das **Bild** oder den **Spaltenraum** von A .

Satz 4.3.1. Für jede $m \times n$ -Matrix A gilt

- $\text{Ker}(A)$ ist ein Untervektorraum von \mathbb{R}^n .
- $\text{Im}(A)$ ist ein Untervektorraum von \mathbb{R}^m .

Beweis: S sicherlich gilt $\vec{0} \in \text{Ker}(A)$ und $\vec{0} \in \text{Im}(A)$, also sind beide Mengen nicht leer.

Seien jetzt $\vec{v}, \vec{w} \in \text{Ker}(A)$ und sei $\lambda \in \mathbb{R}$. Dann gilt nach den allgemeinen Eigenschaften 4.2.4 der Matrizenmultiplikation oder der Linearität von A :

$$\begin{aligned} A \cdot (\vec{v} + \vec{w}) &= A \cdot \vec{v} + A \cdot \vec{w} = \vec{0} + \vec{0} = \vec{0} \\ A \cdot (\lambda \vec{v}) &= \lambda \cdot A \cdot \vec{v} = \lambda \cdot \vec{0} = \vec{0} \end{aligned}$$

also sind auch $\vec{v} + \vec{w} \in \text{Ker}(A)$ und $\lambda \cdot \vec{v} \in \text{Ker}(A)$ und damit ist $\text{Ker}(A)$ ein Untervektorraum.

Sind $\vec{b}, \vec{c} \in \text{Im}(A)$ und ist $\lambda \in \mathbb{R}$, so schreibe zunächst

$$\vec{b} = A \cdot \vec{v}, \quad \vec{c} = A \cdot \vec{w}$$

für geeignete $\vec{v}, \vec{w} \in \mathbb{R}^n$. Damit gilt, wieder nach den allgemeinen Regeln 4.2.4 für die Matrixmultiplikation oder der Linearität von A :

$$\begin{aligned} \vec{b} + \vec{c} &= A \cdot \vec{v} + A \cdot \vec{w} = A \cdot (\vec{v} + \vec{w}) \\ \lambda \cdot \vec{b} &= \lambda \cdot (A \cdot \vec{v}) = A \cdot (\lambda \cdot \vec{v}) \end{aligned}$$

also sind auch $\vec{b} + \vec{c} \in \text{Im}(A)$ und $\lambda \cdot \vec{b} \in \text{Im}(A)$ und damit ist $\text{Im}(A)$ ein Untervektorraum.

Definition 4.3.3. $\text{rg}(A) := \dim(\text{Im}(A))$ heißt der **Rang** der Matrix A .
 $\text{nul}(A) = \dim(\text{Ker}(A))$ heißt die **Nullität** von A .

Bemerkung 4.3.3. Ist A eine $m \times n$ -Matrix, sind $\vec{a}_1 = A_{1,\bullet}^T, \dots, \vec{a}_m = A_{m,\bullet}^T$ die Zeilenvektoren von A (als Spalten geschrieben) und ist

$$\mathcal{Z} = \langle \{\vec{a}_1, \dots, \vec{a}_m\} \rangle$$

der hiervon erzeugte Untervektorraum von \mathbb{R}^n (der **Zeilenraum** von A), so gilt

$$\text{Ker}(A) = \mathcal{Z}^\perp$$

wobei \mathcal{Z}^\perp das orthogonale Komplement von \mathcal{Z} bezeichnet (vergleiche hierzu Definition 3.2.8 aus Abschnitt 3.2).

In der Tat sehen wir, dass

$$A \cdot \vec{v} = \begin{pmatrix} \langle \vec{v}, \vec{a}_m \rangle \\ \vdots \\ \langle \vec{v}, \vec{a}_1 \rangle \end{pmatrix}$$

und damit gilt

$$\begin{aligned} A \cdot \vec{v} = \vec{0} &\iff \langle \vec{v}, \vec{a}_i \rangle = 0 \quad \text{für alle } i \in \{1, \dots, m\} \\ &\iff \vec{v} \in \mathcal{Z}^\perp \end{aligned}$$

Satz 4.3.2. Genau dann ist x_1, \dots, x_n eine Lösung des homogenen linearen Gleichungssystems

$$A \cdot \vec{x} = \vec{0}$$

wenn $\vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \text{Ker}(A)$.

Genau dann hat das allgemeine lineare Gleichungssystem

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

eine Lösung, wenn $\vec{b} \in \text{Im}(A)$.

Ist \vec{v} eine Lösung von

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

so ist jede weitere Lösung dieses Gleichungssystems von der Form $\vec{v} + \vec{w}$ für ein $\vec{w} \in \text{Ker}(A)$ und umgekehrt ist auch für jedes $\vec{w} \in \text{Ker}(A)$ der Vektor $\vec{v} + \vec{w}$ eine Lösung dieses Gleichungssystems.

Beweis: Die ersten beiden Aussagen sind nur eine Umformulierung von Bemerkung 4.3.1.

Ist \vec{v} eine Lösung des gegebenen Gleichungssystems, und ist $\vec{w} \in \text{Ker}(A)$, so gilt

$$A \cdot (\vec{v} + \vec{w}) = A \cdot \vec{v} + A \cdot \vec{w} = \vec{b} + \vec{0} = \vec{b}$$

und damit ist $\vec{v} + \vec{w}$ eine Lösung des Gleichungssystems. Ist \vec{v}' eine weitere Lösung des Gleichungssystems, so gilt für den Vektor $\vec{w} = \vec{v}' - \vec{v}$:

$$A \cdot \vec{w} = A \cdot \vec{v}' - A \cdot \vec{v} = \vec{b} - \vec{b} = \vec{0}$$

und damit ist $\vec{w} \in \text{Ker}(A)$ mit

$$\vec{v}' = \vec{v} + \vec{w}$$

wie gewünscht.

Bemerkung 4.3.4. Ist A die Matrix, die zu einem homogenen Gleichungssystem in Normalform gehört, so kann eine Basis von $\text{Ker}(A)$ leicht ermittelt

werden. Hat das homogene Gleichungssystem die Gestalt

$$\begin{array}{cccccccc}
 x_1 & + & a_{1,2}x_2 & + & a_{1,3}x_3 & + & \cdots & + & a_{1,n}x_n & = & b_1 \\
 0 & + & x_2 & + & a_{2,3}x_3 & + & \cdots & + & a_{2,n}x_n & = & b_2 \\
 \vdots & & & & & & \ddots & & & & \vdots \\
 0 & + & 0 & + & 0 & + & \cdots & + & x_t & + & \cdots & + & a_{t,n}x_n & = & b_t \\
 0 & + & 0 & + & 0 & + & \cdots & + & 0 & = & b_{t+1} \\
 \vdots & & & & & & \ddots & & & & \vdots \\
 0 & + & 0 & + & 0 & + & \cdots & + & 0 & = & b_m
 \end{array}$$

so haben wir gesehen, dass wir genau $n-t$ Freiheitsgrade $x_{t+1} = \lambda_{t+1}, x_{t+2} = \lambda_{t+2}, \dots, x_n = \lambda_n$ haben und zu jeder dieser Parametervorgaben genau eine Lösung (also passende x_1, \dots, x_t) finden. Speziell gilt das für die Kombinationen

$$\begin{array}{ll}
 1: & x_{t+1} = 1, \quad x_{t+2} = 0, \quad x_{t+3} = 0, \quad \dots, \quad x_n = 0 \\
 2: & x_{t+1} = 0, \quad x_{t+2} = 1, \quad x_{t+3} = 0, \quad \dots, \quad x_n = 0 \\
 \vdots & \ddots \\
 n-t: & x_{t+1} = 0, \quad x_{t+2} = 0, \quad x_{t+3} = 0, \quad \dots, \quad x_n = 1
 \end{array}$$

Bezeichnen wir die zugehörigen Lösungen mit $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_{n-t}$, so rechnen wir leicht nach, dass die Lösung \vec{v} zu den Parametern $x_{t+1} = \lambda_{t+1}, x_{t+2} = \lambda_{t+2}, \dots, x_n = \lambda_n$ gegeben ist durch

$$\vec{v} = \lambda_{t+1}\vec{v}_1 + \lambda_{t+2}\vec{v}_2 + \cdots + \lambda_n\vec{v}_{n-t}$$

Aufgrund der speziellen Wahl der Kombinationen 1 bis $n-t$ rechnet man sofort nach, dass $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_{n-t}$ linear unabhängig sind. Damit ist gezeigt, dass sie ein Basis von $\text{Ker}(A)$ bilden.

Im Rest dieses Abschnitts wollen wir noch die Vielfältigkeit der Lösungen genauer analysieren.

Satz 4.3.3. *Es sei $A = (a_{i,j})$ eine $m \times n$ -Matrix, und es seien $\vec{a}_1 = A_{\bullet,1}, \dots, \vec{a}_n = A_{\bullet,n}$ die Spaltenvektoren von A . Dann gilt*

$$\text{Im}(A) = \langle \vec{a}_1, \dots, \vec{a}_n \rangle$$

Beweis: Nach Definition der Matrizenmultiplikation und Bemerkung 4.2.2 ist

$$A \cdot \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = v_1 \cdot \vec{a}_1 + \cdots + v_n \cdot \vec{a}_n$$

und damit gilt

$$\text{Im}(A) \subseteq \langle \vec{a}_1, \dots, \vec{a}_n \rangle$$

Umgekehrt gilt aber auch

$$\lambda_1 \cdot \vec{a}_1 + \cdots + \lambda_n \cdot \vec{a}_n = A \cdot \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}$$

und hieraus folgt die andere Inklusion

$$\text{Im}(A) \supseteq \langle \vec{a}_1, \dots, \vec{a}_n \rangle$$

also Gleichheit der beiden Mengen.

Wir haben also jetzt gesehen, dass wir aus A unmittelbar ein Erzeugendensystem von $\text{Im}(A)$ ablesen können. Wie aber können wir daraus eine Basis von $\text{Im}(A)$ ableiten? Wie finden wir eine Basis von $\text{Ker}(A)$? Und wie können wir den Rang von A schnell bestimmen.

Zunächst halten wir hierzu fest, dass gilt

$$\text{rg}(A) = \text{rg}(A \mid \vec{0})$$

wobei $(A \mid \vec{0})$ die augmentierte Matrix zum homogenen Gleichungssystem

$$A \cdot \vec{x} = \vec{0}$$

ist, denn durch Hinzunahme des Nullvektors ändert sich ein Untervektorraum nicht. Augmentierte Matrizen haben wir in Satz 4.1.6 durch die vier Operationen

- (i) Subtrahiere das Vielfache einer Zeile von $(A \mid \vec{b})$ von einer anderen Zeile

- (ii) Vertausche zwei Zeilen von $(A | \vec{b})$
- (iii) Vertausche zwei Spalten von $(A | \vec{b})$, von denen keine die letzte sein darf.
- (iv) Multipliziere eine Zeile von $(A | \vec{b})$ mit einer Zahl $r \neq 0$.

auf Normalform gebracht. Der entscheidende Punkt für unsere Überlegungen ist

Hilfssatz 4.3.4. *Führen wir an einer augmentierten Matrix $(A | \vec{0})$ eine der Operationen (i) - (iv) durch, so ändert sich dadurch der Rang von $(A | \vec{0})$ nicht.*

Beweis: Klar ist, dass die Operation (iii) den Rang nicht ändert, denn das Erzeugnis von Vektoren hängt nicht von deren Reihenfolge ab.

Für den Nachweis in den anderen Fällen führen wir eine der Operation (i), (ii) oder (iv) an der Matrix $(A | \vec{0})$ durch und nennen die dadurch erhaltene augmentierte Matrix $(A' | \vec{0})$. Mit \vec{a}_j bezeichnen wir den j -ten Spaltenvektor von $(A | \vec{0})$ und mit \vec{a}'_j den entsprechenden Spaltenvektor von $(A' | \vec{0})$. Es reicht dann zu zeigen: Sind $1 \leq j_1 < \dots < j_t \leq n+1$, so sind die Vektoren $\vec{a}'_{j_1}, \dots, \vec{a}'_{j_t}$ genau dann linear unabhängig, wenn es die Vektoren $\vec{a}_{j_1}, \dots, \vec{a}_{j_t}$ sind. Wir können dabei auch immer annehmen, dass $j_t \leq n$, denn $\vec{a}'_{n+1} = \vec{a}_{n+1} = \vec{0}$.

Wir betrachten Operation (i). Zur Vereinfachung der Schreibweise nehmen wir an, dass wir dabei das r -fache der zweiten Zeile zur ersten Zeile addieren. Dann ist die lineare Unabhängigkeit von $\vec{a}'_{j_1}, \dots, \vec{a}'_{j_t}$ äquivalent dazu, dass das Gleichungssystem

$$\begin{array}{cccccc}
 a_{1,j_1}x_1 & + & a_{1,j_2}x_2 & + & \dots & + & a_{1,j_t}x_t & = & 0 \\
 a_{2,j_1}x_1 & + & a_{2,j_2}x_2 & + & \dots & + & a_{2,j_t}x_t & = & 0 \\
 \vdots & & & & \ddots & & & & \vdots \\
 a_{m,j_1}x_1 & + & a_{m,j_2}x_2 & + & \dots & + & a_{m,j_t}x_t & = & 0
 \end{array} \tag{4.3.2}$$

nur die triviale Lösung hat, und die lineare Unabhängigkeit von $\vec{a}_{j_1}', \dots, \vec{a}_{j_t}'$ ist äquivalent dazu, dass das Gleichungssystem

$$\begin{array}{ccccccc}
 ((a_{1,j_1} + ra_{2,j_1})x_1 & + & (a_{1,j_2} + ra_{2,j_2})x_2 & + & \dots & + & (a_{1,j_t} + ra_{2,j_t})x_t & = & 0 \\
 a_{2,j_1}x_1 & & + & a_{2,j_2}x_2 & & + & \dots & + & a_{2,j_t}x_t & = & 0 \\
 \vdots & & & & & & \ddots & & & & \vdots \\
 a_{m,j_1}x_1 & & + & x_{m,j_2}x_2 & & + & \dots & + & a_{m,j_t}x_t & = & 0
 \end{array} \tag{4.3.3}$$

nur die triviale Lösung hat.

Aber die Gleichungssysteme (4.3.2) und (4.3.3) sind äquivalent, haben also die gleichen Lösungen.

Die anderen beiden Fälle behandelt man ähnlich.

Bemerkung 4.3.5. Anstelle der Zeilenoperationen

- (i) Subtrahiere das Vielfache einer Zeile von $(A | \vec{b})$ von einer anderen Zeile
- (iv) Multipliziere eine Zeile von $(A | \vec{b})$ mit einer Zahl $r \neq 0$.

können wir an der augmentierten Matrix genauso die Spaltenoperationen

- (i) Subtrahiere das Vielfache einer Spalte von $(A | \vec{b})$ die nicht die letzte ist von einer anderen Spalte, die nicht die letzte ist.
- (iv) Multipliziere eine Spalte von $(A | \vec{b})$, die nicht die letzte ist, mit einer Zahl $r \neq 0$.

Auch diese Operationen verändern die Lösbarkeit des Gleichungssystems $A \cdot \vec{x} = \vec{b}$ (wenn auch das Rückwärtsrechnen komplizierter wird).

Bemerkung 4.3.6. Die Operationen (i) - (iv) lassen die letzte Spalte der augmentierten Matrix $(A | \vec{0})$ unberührt; sie bleibt immer die Nullspalte. Daher verzichten wir in diesem Fall darauf, diese Spalte mitzuführen und führen die Operationen an der Matrix A durch. Wir sprechen dann auch von der Normalform von A .

Benutzen wir für die Findung der Normalform sowohl Zeilen– als auch Spaltenoperationen, so können wir A in eine rangäquivalente Matrix der Gestalt

$$A' = \begin{pmatrix} 1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & 0 & \dots & 0 \\ \vdots & & \ddots & & & & \vdots \\ 0 & 0 & \dots & 1 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & & \ddots & & & & \vdots \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \end{pmatrix}$$

In diesem Fall ist der Rang von A die Anzahl der 1–Einträge in der Matrix A' .

Folgerung 4.3.5. *Der Rang einer Matrix A ist der Rang des zugehörigen Gleichungssystems $A \cdot \vec{x} = \vec{0}$, kann also mit Hilfe des Gaußschen Eliminationsalgorithmus ermittelt werden.*

Folgerung 4.3.6. *Genau dann hat das Gleichungssystem*

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

eine Lösung, wenn

$$\text{rg}(A) = \text{rg}\left(A \mid \vec{b}\right)$$

Aus Satz 4.1.4 in Verbindung mit Satz 4.1.6 erhalten wir damit

Folgerung 4.3.7. *Hat die Matrix A den Rang n , so kann jede augmentierte Matrix $\left(A \mid \vec{b}\right)$ durch wiederholtes Anwenden der Operationen (i) (ii) und (iv) auf Normalform gebracht werden.*

Folgerung 4.3.8. *Ist A eine $m \times n$ –Matrix, so gilt*

$$\text{rg}(A) = \text{rg}(A^T)$$

Beweis: Wie wir in 4.3.4 und 4.3.5 gesehen haben, ändert sich der Rang von A nicht, wenn wir die erlaubten Zeilen– oder Spaltenoperationen an A durchführen. Aber Spaltenoperationen an A entsprechen Zeilenoperationen an A^T und umgekehrt. Ist daher A' eine rangäquivalente Form von A wie wir sie in 4.3.6 erreicht haben, so ist $(A')^T$ eine rangäquivalente Form von A^T , und daraus folgt die Behauptung.

Folgerung 4.3.9. *Ist A eine $m \times n$ -Matrix, so gilt*

$$\operatorname{rg}(A) + \operatorname{nul}(A) = n$$

Beweis: Wir können dazu annehmen, dass A in Normalform vorliegt. Nach Lemma 4.3.4 ändert sich dadurch der Rang der Matrix nicht, und auch die Lösungsmenge von $A \cdot \vec{x} = \vec{0}$ bleibt von den Operationen (i) bis (iv) bis auf eventuelles Vertauschen von Komponenten unberührt. Hat aber die Normalform von A die Gestalt

$$\begin{array}{cccccc} 1 & a_{1,2} & \dots & a_{1,t} & \dots & a_{1,n} \\ 0 & 1 & \dots & a_{2,t} & \dots & a_{2,n} \\ \vdots & & \ddots & & & \vdots \\ 0 & 0 & \dots & 1 & \dots & a_{t,n} \\ 0 & 0 & \dots & 0 & \dots & 0 \\ \vdots & & \ddots & & & \vdots \\ 0 & 0 & \dots & 0 & \dots & 0 \end{array}$$

so hat A den Rang t , und in Bemerkung 4.3.4 haben wir bereits festgestellt, dass $\dim(\operatorname{Ker}(A)) = n - t$.

Folgerung 4.3.10. *Hat ein lineares Gleichungssystem*

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

mit einer $m \times n$ -Matrix A eine eindeutige Lösung für jede Wahl von $\vec{b} \in \mathbb{R}^m$, so muss gelten $n = m$.

Beweis: Damit es überhaupt eine Lösung von

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

gibt, muss auf jeden Fall jeder Vektor $\vec{b} \in \mathbb{R}^m$ im Bild $\operatorname{Im}(A)$ von A sein, und damit muss $\operatorname{rg}(A) = m$ gelten. Andererseits wollen wir das jedes Gleichungssystem nur eine Lösung hat, speziell also auch

$$A \cdot \vec{x} = \vec{0}$$

Das bedeutet aber, dass $\operatorname{nul}(A) = 0$, und aus 4.3.9 folgt jetzt sofort, dass $n = m$.

Wir können die Ergebnisse unserer Überlegungen wie folgt zusammenfassen:

Regel 4.3.11. *Es sei A eine $m \times n$ -Matrix und es sei*

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

das zugehörige allgemeine Gleichungssystem. Dann gilt:

1. *Ist $\vec{b} \in \text{Im}(A)$, so hat das Gleichungssystem (mindestens) eine Lösung, ist $\vec{b} \notin \text{Im}(A)$, so hat das Gleichungssystem keine Lösung.*
2. *Ist $\text{nul}(A) = 0$, so hat das Gleichungssystem höchstens eine Lösung.*
3. *Ist $\text{nul}(A) > 0$, so hat das Gleichungssystem entweder keine oder unendlich viele Lösungen.*

Bemerkung 4.3.7. Betrachten wir eine komplexe $m \times n$ -Matrix A , so definiert diese keine Abbildung $f_A: \mathbb{R}^n \rightarrow \mathbb{R}^n$, sondern eine Abbildung $f_A: \mathbb{C}^n \rightarrow \mathbb{C}^n$. Wie wir im Abschnitt 3.4 gesehen haben, können wir \mathbb{C}^n als komplexen Vektorraum der Dimension n auffassen. Mit dieser Betrachtung gelten die Aussagen dieses Abschnitts analog auch für komplexe Matrizen.

Aufgabe 4.3.1. Gegeben sei die Matrix

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

Bestimmen Sie $\text{rg}(A)$ und $\text{nul}(A)$.

Aufgabe 4.3.2. Gegeben sei die Matrix

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

Bestimmen Sie $\text{rg}(A)$ und $\text{nul}(A)$.

Aufgabe 4.3.3. Gegeben sei die Matrix

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

Bestimmen Sie alle $\vec{b} \in \mathbb{R}^3$, für die das Gleichungssystem

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

Lösungen hat.

Aufgabe 4.3.4. Gegeben sei die Matrix

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

Bestimmen Sie eine Basis von $\ker(A)$ und eine Basis von $\text{Im}(A)$.

Aufgabe 4.3.5. Zeigen Sie: Ist A eine $n \times n$ -Matrix mit $\text{nul}(A) = 0$, so hat das Gleichungssystem

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

für jedes $\vec{b} \in \mathbb{R}^n$ eine eindeutige Lösung.

Aufgabe 4.3.6. Wir betrachten eine 3×3 -Matrix A mit $\text{rg}(A) = 2$, und wir nehmen an, dass die ersten beiden Zeilen \vec{a}_1 und \vec{a}_2 von A linear unabhängig sind. Zeigen Sie, dass der Vektor

$$\vec{v} = \vec{a}_1 \times \vec{a}_2$$

(das Vektorprodukt der Vektoren \vec{a}_1 und \vec{a}_2) eine Basis von $\ker(A)$ ist.

Kapitel 5

Determinanten

5.1 Definition der Determinanten

Von besonderem Interesse für uns und für viele Anwendungen sind quadratische Matrizen bzw. Gleichungssysteme mit genauso vielen Unbekannten wie Gleichungen. Wie wir bereits in 4.3.9 gesehen haben, muss eine Matrix A quadratisch sein, damit alle zugehörigen Gleichungssysteme $A \cdot \vec{x} = \vec{b}$ eine eindeutige Lösung besitzen. In diesem Abschnitt werden wir jeder quadratischen Matrix in subtiler Weise einen Skalar, ihre Determinante, zuordnen, der uns erlaubt, die Frage nach der eindeutigen Lösbarkeit allgemein zu entscheiden, und die uns auch hilft, diese Lösung gegebenenfalls zu finden.

Determinanten waren wohl schon G. Leibniz (1646 - 1716) bekannt. Wiederentdeckt und weiterentwickelt wurde sie vom Mathematiker G. Cramer (1704 - 1752). Beiträge zur Determinantentheorie stammen auch von dem englischen Mathematiker C. L. Dodgson (1832 - 1898), der allerdings besser als Lewis Carroll und für seine Kinderliteratur (etwa *Alice im Wunderland*) bekannt ist.

5.1.1 Zweidimensionale Determinanten

Der einfachste Fall eines linearen Gleichungssystems ist natürlich der einer einzigen Gleichung in einer Unbekannten,

$$a \cdot x = b \tag{5.1.1}$$

In dieser Situation ist die Lage offensichtlich:

Regel 5.1.1. *Genau dann hat (5.1.1) für jede Wahl von b eine eindeutige Lösung, wenn $a \neq 0$; in diesem Fall ist $x = \frac{b}{a}$.*

Unser Ziel ist es, diese Aussage in möglichst großen Umfang auf $n \times n$ -Matrizen zu verallgemeinern.

Wir wollen uns in diesem Abschnitt mit dem Spezialfall $n = 2$ beschäftigen, den nach (5.1.1) einfachsten Fall. Ein solches Gleichungssystem hat die allgemeine Gestalt

$$\begin{aligned} a_{1,1}x_1 + a_{1,2}x_2 &= b_1 \\ a_{2,1}x_1 + a_{2,2}x_2 &= b_2 \end{aligned} \quad (5.1.2)$$

Wenn wir versuchen, dieses Gleichungssystem nach den Eliminationsmethoden von Gauß zu lösen, kommen wir schnell auf Fallunterscheidungen (der Fall $a_{1,1} = 0$ wird anders behandelt als der Fall $a_{1,1} \neq 0$). Wir wollen auf diese Fallunterscheidung verzichten und multiplizieren die erste Gleichung mit $a_{2,2}$ und die zweite mit $a_{1,2}$. Damit erhalten wir

$$\begin{aligned} a_{2,2}a_{1,1}x_1 + a_{2,2}a_{1,2}x_2 &= a_{2,2}b_1 \\ a_{1,2}a_{2,1}x_1 + a_{1,2}a_{2,2}x_2 &= a_{1,2}b_2 \end{aligned} \quad (5.1.3)$$

wobei wir aber beachten müssen, dass das erhaltene Gleichungssystem nicht mehr äquivalent zum Ausgangssystem ist, wenn $a_{2,2} = 0$ oder $a_{1,2} = 0$. Addieren wir in (5.1.3) die erste Gleichung zum Negativen der zweiten, so erhalten wir

$$\begin{aligned} a_{2,2}a_{1,1}x_1 + a_{2,2}a_{1,2}x_2 &= a_{2,2}b_1 \\ (a_{1,1}a_{2,2} - a_{1,2}a_{2,1}) \cdot x_1 &= a_{2,2}b_1 - a_{1,2}b_2 \end{aligned} \quad (5.1.4)$$

Falls jetzt $(a_{1,2}a_{2,1} - a_{1,1}a_{2,2}) \neq 0$, so hat die zweite Zeile, wie wir schon im Spezialfall $n = 1$ gesehen haben, die eindeutige Lösung

$$x_1 = \frac{a_{2,2}b_1 - a_{1,2}b_2}{a_{1,1}a_{2,2} - a_{1,2}a_{2,1}}$$

und hieraus ergibt sich aus der ersten Gleichung

$$x_2 = \frac{a_{1,1}b_2 - a_{2,1}b_1}{a_{1,1}a_{2,2} - a_{1,2}a_{2,1}}$$

als eindeutiger Wert für x_2 . Damit haben wir gesehen, dass (5.1.4) für jede Wahl von b_1 und b_2 genau eine Lösung hat, wenn $a_{1,1}a_{2,2} - a_{1,2}a_{2,1} \neq 0$. Setzen wir diese Lösung in das ursprüngliche Gleichungssystem (5.1.2) ein, so rechnen wir leicht nach, dass auch dieses gelöst wird. Damit haben wir also eine Lösung von (5.1.2) gefunden, falls $a_{1,1}a_{2,2} - a_{1,2}a_{2,1} \neq 0$. Wenn aber $a_{2,2} = 0$ oder $a_{1,2} = 0$, so ist das transformierte Gleichungssystem nicht äquivalent zum ursprünglichen. Allerdings kann sich durch unsere Transformationen die Anzahl der Lösungen höchstens vergrößern, jede Lösung von (5.1.2) ist auch eine Lösung von (5.1.4). Damit haben wir schon gesehen: Falls $a_{1,2}a_{2,1} - a_{1,1}a_{2,2} \neq 0$ so hat (5.1.2) für jede Wahl von b_1 und b_2 eine eindeutig bestimmte Lösung, die gegeben wird durch

$$\begin{aligned} x_1 &= \frac{a_{2,2}b_1 - a_{1,2}b_2}{a_{1,1}a_{2,2} - a_{1,2}a_{2,1}} \\ x_2 &= \frac{a_{1,1}b_2 - a_{2,1}b_1}{a_{1,1}a_{2,2} - a_{1,2}a_{2,1}} \end{aligned}$$

Es bleibt noch der Fall $a_{1,1}a_{2,2} - a_{1,2}a_{2,1} = 0$ zu betrachten. In diesem Fall hat das Gleichungssystem (5.1.4) keine Lösung, falls $a_{1,2}b_2 - a_{2,2}b_1 \neq 0$, oder falls $a_{2,2}a_{1,1} = 0$ und $a_{2,2}a_{1,2} = 0$ und $a_{2,2}b_1 \neq 0$. Andernfalls hat (5.1.4) unendlich viele Lösungen.

Aber was ist aber in diesem Fall mit der ursprünglichen Gleichung (5.1.2)? Wir betrachten hierzu alle in Frage kommenden Fälle. Klar ist der triviale Fall $a_{1,1} = a_{1,2} = a_{2,1} = a_{2,2} = 0$: In diesem Fall gibt es für $b_1 = b_2 = 0$ unendlich viele Lösungen und in allen anderen Fällen keine. Wir können uns daher auf die Situation beschränken, dass mindestens ein Koeffizient $a_{i,j} \neq 0$. Ist aber etwa $a_{1,1} \neq 0$, so muss wegen $a_{1,1}a_{2,2} - a_{1,2}a_{2,1} = 0$ schon $a_{2,2} = \frac{a_{1,2}a_{2,1}}{a_{1,1}}$ gelten, und daher ist $x_1 = a_{1,2}$ und $x_2 = -a_{1,1}$ eine nichttriviale Lösung von

$$\begin{aligned} a_{1,1}x_1 + a_{1,2}x_2 &= 0 \\ a_{2,1}x_1 + a_{2,2}x_2 &= 0 \end{aligned} \tag{5.1.5}$$

wie wir sofort nachrechnen. Entsprechend finden wir im Fall $a_{2,2} \neq 0$ die nichttriviale Lösung $x_1 = a_{2,2}$ und $x_2 = -a_{2,1}$, ist $a_{2,1} \neq 0$, so löst ebenfalls $x_1 = a_{2,2}$ und $x_2 = -a_{2,1}$ das Gleichungssystem (5.1.5) und ist nicht-trivial, und ist $a_{1,2} \neq 0$, so liefert wiederum $x_1 = a_{1,2}$ und $x_2 = -a_{1,1}$ eine nicht-triviale Lösung von (5.1.5). Da es in diesem Fall auch noch die triviale Lösung

gibt, hat also das homogene System (5.1.5) mehrere Lösungen und entsprechend hat das System (5.1.2) wegen Regel 4.3.11 entweder gar keine oder unendlich viele Lösungen.

Definition 5.1.1. Die **Determinante** $\det(A)$ einer 2×2 -Matrix $A = \begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix}$ ist der Ausdruck

$$\det(A) = a_{1,1} \cdot a_{2,2} - a_{2,1} \cdot a_{1,2}$$

Wir haben in unseren Vorüberlegungen schon das folgende Resultat abgeleitet:

Satz 5.1.2. Ist A eine 2×2 -Matrix, so hat das Gleichungssystem

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

genau dann eine eindeutige Lösung für jedes $\vec{b} \in \mathbb{R}^2$, wenn $\det(A) \neq 0$.

Darüberhinaus haben wir sogar eine Formel für die eindeutig bestimmte Lösung gefunden, die sogenannte **Cramersche Regel** (für $n = 2$).

Satz 5.1.3. Ist A eine 2×2 -Matrix mit $\det(A) \neq 0$, so ist die eindeutige Lösung von

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

für einen beliebigen Vektor $\vec{b} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$ gegeben durch

$$\begin{aligned} x_1 &= \frac{a_{2,2} \cdot b_1 - a_{1,2} \cdot b_2}{\det(A)} \\ x_2 &= \frac{a_{1,1} \cdot b_2 - a_{2,1} \cdot b_1}{\det(A)} \end{aligned}$$

Beispiel 5.1.1. Wir betrachten die Matrix

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

Hierfür gilt

$$\det(A) = 1 \cdot 4 - 2 \cdot 3 = 4 - 6 = -2$$

Insbesondere ist also $\det(A) \neq 0$ und damit hat jedes Gleichungssystem

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

eine eindeutige Lösung. Betrachten wir etwa

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

so gilt

$$\begin{aligned} x_1 &= \frac{4 \cdot 2 - 2 \cdot 4}{-2} = 0 \\ x_2 &= \frac{1 \cdot 4 - 3 \cdot 2}{-2} = 1 \end{aligned}$$

Natürlich führt auch der Gaußalgorithmus zu diesem Ergebnis.

Beispiel 5.1.2. Wir betrachten die Matrix

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

Hierfür gilt

$$\det(A) = 1 \cdot 1 - 1 \cdot (-1) = 1 - (-1) = 2$$

Insbesondere ist also $\det(A) \neq 0$ und damit hat jedes Gleichungssystem

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

eine eindeutige Lösung. Betrachten wir etwa

$$\begin{aligned} x_1 - x_2 &= 4 \\ x_1 + x_2 &= 2 \end{aligned}$$

so gilt

$$\begin{aligned} x_1 &= \frac{1 \cdot 4 - (-1) \cdot 2}{2} = 3 \\ x_2 &= \frac{1 \cdot 2 - 1 \cdot 4}{2} = -1 \end{aligned}$$

Beispiel 5.1.3. Wir betrachten die Matrix

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

Hierfür gilt

$$\det(A) = 1 \cdot 4 - 2 \cdot 2 = 4 - 4 = 0$$

Insbesondere ist also $\det(A) = 0$ und damit gibt es kein Gleichungssystem

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

mit einer eindeutigen Lösung. Betrachten wir das zugehörige homogene System

$$\begin{array}{rcrcrcrcrcl} x_1 & + & 2x_2 & = & 0 \\ 2x_1 & + & 4x_2 & = & 0 \end{array}$$

so hat dieses die Normalform

$$\begin{array}{rcrcrcrcrcl} x_1 & + & 2x_2 & = & 0 \\ & & 0 & = & 0 \end{array}$$

und wir sehen, dass $x_1 = -2$ und $x_2 = 1$ eine nichttriviale Lösung ist. Aus der Normalform lesen wir auch ab, dass $\text{rg}(A) = 1$, also wegen Satz 4.3.9 auch $\text{nul}(A) = 1$. Also ist $\text{Ker}(A)$ ein Vektorraum der Dimension 1, und $\begin{pmatrix} -2 \\ 1 \end{pmatrix}$ ist eine Basis.

In Satz 5.1.3 haben wir eine Formel für die Lösungen eines Gleichungssystems kennengelernt, falls die Determinante der Koeffizientenmatrix von Null verschieden ist. In dieser Formel haben auch die Zähler eine Gestalt, die an eine Determinante erinnert. Dazu betrachten wir wieder ein Gleichungssystem

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

mit Koeffizientenmatrix $A = \begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix}$ und $\vec{b} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$ und wir bezeichnen mit $A_1(\vec{b})$ bzw. $A_2(\vec{b})$ die 2×2 -Matrix, die wir aus A erhalten indem wir die erste bzw. zweite Spalte von A durch den Vektor \vec{b} ersetzen, also

$$A_1(\vec{b}) = \begin{pmatrix} b_1 & a_{1,2} \\ b_2 & a_{2,2} \end{pmatrix}, \quad A_2(\vec{b}) = \begin{pmatrix} a_{1,1} & b_1 \\ a_{2,1} & b_2 \end{pmatrix}$$

Dann folgt sofort aus Satz 5.1.3, dass auch die folgende Form der Cramerschen Regel gilt:

Satz 5.1.4. *Ist A eine 2×2 -Matrix mit $\det(A) \neq 0$, so ist die eindeutige Lösung von*

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

für einen beliebigen Vektor \vec{b} gegeben durch

$$\begin{aligned} x_1 &= \frac{\det(A_1(\vec{b}))}{\det(A)} \\ x_2 &= \frac{\det(A_2(\vec{b}))}{\det(A)} \end{aligned}$$

Die Determinante einer 2×2 -Matrix ist sehr einfach zu berechnen. Im Hinblick auf spätere Verallgemeinerungen wollen wir trotzdem schon in diesem Fall Regeln für die Berechnung von Determinanten aufstellen, die die Berechnungen weiter vereinfachen.

Regel 5.1.5. *Für die 2×2 -Einheitsmatrix $E_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ gilt*

$$\det(E_2) = 1$$

Regel 5.1.6. *Für eine 2×2 -Matrix A gilt $\det(A) = \det(A^T)$.*

Beweis: Ist $A = \begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix}$, so ist $A^T = \begin{pmatrix} a_{1,1} & a_{2,1} \\ a_{1,2} & a_{2,2} \end{pmatrix}$. Jetzt folgt die Aussage aus der Definition.

Beispiel 5.1.4.

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

Regel 5.1.7. *Für eine 2×2 -Matrix A gilt $\det(A) = 0$, wenn*

1. *Eine Zeile von A die Nullzeile ist.*
2. *Eine Spalte von A die Nullspalte ist.*
3. *Eine Zeile von A ein Vielfaches der anderen Zeile ist.*

4. Eine Spalte von A ein Vielfaches der anderen Spalte ist.

Beispiel 5.1.5. Ist $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ so gilt

$$\det(A) = 0$$

Beispiel 5.1.6. Ist $A = \begin{pmatrix} 1 & 2 \\ 4 & 8 \end{pmatrix}$ so gilt

$$\det(A) = 0$$

Beweis: Die ersten beiden Formeln sind klar. Für die dritte betrachten wir den Fall $a_{2,1} = \lambda a_{1,1}$, $a_{2,2} = \lambda a_{1,2}$ und erhalten

$$\det(A) = a_{1,1} \cdot \lambda \cdot a_{1,2} - a_{1,2} \cdot \lambda \cdot a_{1,1} = 0$$

Genauso behandeln wir den Fall, dass die erste Zeile ein Vielfaches der zweiten ist.

Die vierte Regel kann analog bewiesen werden. Sie folgt aber wegen Regel 5.1.6 auch sofort aus der dritten!

Regel 5.1.8. Ist A eine obere oder untere Dreiecksmatrix (d.h. gilt $a_{1,2} = 0$ oder $a_{2,1} = 0$), so gilt

$$\det(A) = a_{1,1} \cdot a_{2,2}$$

Beweis: Allgemein gilt

$$\det(A) = a_{1,1} \cdot a_{2,2} - a_{1,2} \cdot a_{2,1}$$

Ist A eine Dreiecksmatrix, so ist $a_{1,2} = 0$ oder $a_{2,1} = 0$, und unsere Formel folgt.

Beispiel 5.1.7.

$$\det \begin{pmatrix} 2 & 2 \\ 0 & 3 \end{pmatrix} = 2 \cdot 3 = 6$$

Regel 5.1.9. *Entsteht die 2×2 -Matrix A' aus A durch Vertauschen der beiden Zeilen, so gilt*

$$\det(A') = -\det(A)$$

Das Gleiche gilt, wenn A' durch Vertauschung der Spalten aus A entsteht.

Beweis: Ist $A = \begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix}$ so ist $A' = \begin{pmatrix} a_{2,1} & a_{2,2} \\ a_{1,1} & a_{1,2} \end{pmatrix}$ und damit

$$\det(A') = a_{2,1}a_{1,2} - a_{1,1}a_{2,2} = -\det(A)$$

Beispiel 5.1.8. Ist $A = \begin{pmatrix} 1 & 2 \\ 3 & 0 \end{pmatrix}$, so ist $A' = \begin{pmatrix} 3 & 0 \\ 1 & 2 \end{pmatrix}$ und

$$\det(A') = 3 \cdot 2 - 0 \cdot 1 = 6 = -(1 \cdot 0 - 2 \cdot 3) = \det(A)$$

Regel 5.1.10. *Entsteht die 2×2 -Matrix A' aus A durch Multiplikation einer Zeile von A mit einer Zahl r , so gilt*

$$\det(A') = r \cdot \det(A)$$

Das Gleiche gilt, wenn A' aus A durch Multiplikation einer Spalte von A mit einer Zahl r hervorgeht.

Beweis: Ist $A = \begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix}$ und entsteht A' durch Multiplikation der ersten Zeile von A mit r aus A , so ist $A' = \begin{pmatrix} r \cdot a_{1,1} & r \cdot a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix}$ und damit

$$\det(A') = r \cdot a_{1,1} \cdot a_{2,2} - r \cdot a_{1,2} \cdot a_{2,1} = r \cdot (a_{1,1}a_{2,2} - a_{1,2}a_{2,1}) = r \cdot \det(A)$$

Die anderen Fälle behandelt man genauso.

Beispiel 5.1.9. Ist $A = \begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix}$, und ist $A' = \begin{pmatrix} 4 & 8 \\ 1 & 3 \end{pmatrix}$, so ist

$$\det(A') = 4 \cdot 3 - 8 \cdot 1 = 4 = 4 \cdot (1 \cdot 3 - 2 \cdot 1) = 4 \cdot \det(A)$$

Die erste nichttriviale Regel ist die folgende

Regel 5.1.11. *Entsteht die 2×2 -Matrix A' aus A dadurch, dass wir ein Vielfaches einer Zeile von A zur anderen Zeile von A addieren, so gilt*

$$\det(A') = \det(A)$$

Das Gleiche gilt, wenn A' aus A dadurch entsteht, dass wir ein Vielfaches einer Spalte von A zur anderen Spalte von A addieren.

Beweis: Sei $A = \begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix}$. Wir betrachten den Fall, dass das r -fache der ersten Zeile von A zur zweiten addiert wird, so dass

$$A' = \begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} + r \cdot a_{1,1} & a_{2,2} + r \cdot a_{1,2} \end{pmatrix}.$$

Dann gilt

$$\begin{aligned} \det(A') &= a_{1,1} \cdot (a_{2,2} + r \cdot a_{1,2}) - a_{1,2} \cdot (a_{2,1} + r \cdot a_{1,1}) \\ &= a_{1,1} \cdot a_{2,2} - a_{1,2}a_{2,1} + (a_{1,1} \cdot r \cdot a_{1,2} - a_{1,2} \cdot r \cdot a_{1,1}) \\ &= a_{1,1} \cdot a_{2,2} - a_{1,2}a_{2,1} \\ &= \det(A) \end{aligned}$$

Beispiel 5.1.10. Ist $A = \begin{pmatrix} -1 & 1 \\ 2 & 3 \end{pmatrix}$, und ist $A' = \begin{pmatrix} -1 & 1 \\ 0 & 1 \end{pmatrix}$, so entsteht A' aus A durch Addition des doppelten der ersten Zeile zur zweiten, und es gilt

$$\det(A') = (-1) \cdot 1 - 0 \cdot 1 = 1 = (-1) \cdot 3 - 1 \cdot 2 = \det(A)$$

Bemerkung 5.1.1. Die Regeln erlauben es immer, die Berechnung von Determinanten auf Dreiecksmatrizen zu reduzieren. In diesem Fall ist die Determinantenberechnung besonders einfach.

Regel 5.1.12. *Sind A, B zwei 2×2 -Matrizen, so gilt*

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

Beweis: Falls $A = \begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix}$ und $B = \begin{pmatrix} b_{1,1} & b_{1,2} \\ b_{2,1} & b_{2,2} \end{pmatrix}$, so ist

$$A \cdot B = \begin{pmatrix} a_{1,1}b_{1,1} + a_{1,2}b_{2,1} & a_{1,1}b_{1,2} + a_{1,2}b_{2,2} \\ a_{2,1}b_{1,1} + a_{2,2}b_{2,1} & a_{2,1}b_{1,2} + a_{2,2}b_{2,2} \end{pmatrix}$$

und damit

$$\begin{aligned} \det(A \cdot B) &= (a_{1,1}b_{1,1} + a_{1,2}b_{2,1}) \cdot (a_{2,1}b_{1,2} + a_{2,2}b_{2,2}) \\ &= -(a_{1,1}b_{1,2} + a_{1,2}b_{2,2}) \cdot (a_{2,1}b_{1,1} + a_{2,2}b_{2,1}) \\ &= a_{1,1}b_{1,1}a_{2,1}b_{1,2} + a_{1,1}b_{1,1}a_{2,2}b_{2,2} \\ &\quad + a_{1,2}b_{2,1}a_{2,1}b_{1,2} + a_{1,2}b_{2,1}a_{2,2}b_{2,2} \\ &\quad - a_{1,1}b_{1,2}a_{2,1}b_{1,1} - a_{1,1}b_{1,2}a_{2,2}b_{2,1} \\ &\quad - a_{1,2}b_{2,2}a_{2,1}b_{1,1} - a_{1,2}b_{2,2}a_{2,2}b_{2,1} \\ &= a_{1,1}b_{1,1}a_{2,2}b_{2,2} + a_{1,2}b_{2,1}a_{2,1}b_{1,2} \\ &\quad - a_{1,1}b_{1,2}a_{2,2}b_{2,1} - a_{1,2}b_{2,2}a_{2,1}b_{1,1} \\ &= (a_{1,1}a_{2,2} - a_{1,2}a_{2,1}) \cdot (b_{1,1}b_{2,2} - b_{1,2}b_{2,1}) \\ &= \det(A) \cdot \det(B) \end{aligned}$$

Beispiel 5.1.11. Für $A = \begin{pmatrix} 1 & 3 \\ 0 & 2 \end{pmatrix}$ und $B = \begin{pmatrix} 3 & 0 \\ 4 & 2 \end{pmatrix}$ gilt

$$A \cdot B = \begin{pmatrix} 15 & 6 \\ 8 & 4 \end{pmatrix}$$

und

$$\det(A \cdot B) = 15 \cdot 4 - 6 \cdot 8 = 12 = 2 \cdot 6 = \det(A) \cdot \det(B)$$

Die Determinante kann auch geometrisch interpretiert werden:

Satz 5.1.13. Sind

$$\vec{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}, \quad \vec{w} = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}$$

zwei nicht kollineare Vektoren, ist P das Parallelogramm, dessen Seiten durch die beiden Vektoren \vec{v} und \vec{w} gegeben ist, und ist

$$A = \begin{pmatrix} v_1 & w_1 \\ v_2 & w_2 \end{pmatrix}$$

die Matrix mit \vec{v} und \vec{w} als Spalten, so ist $|\det(A)|$ die Fläche des Parallelogramms P .

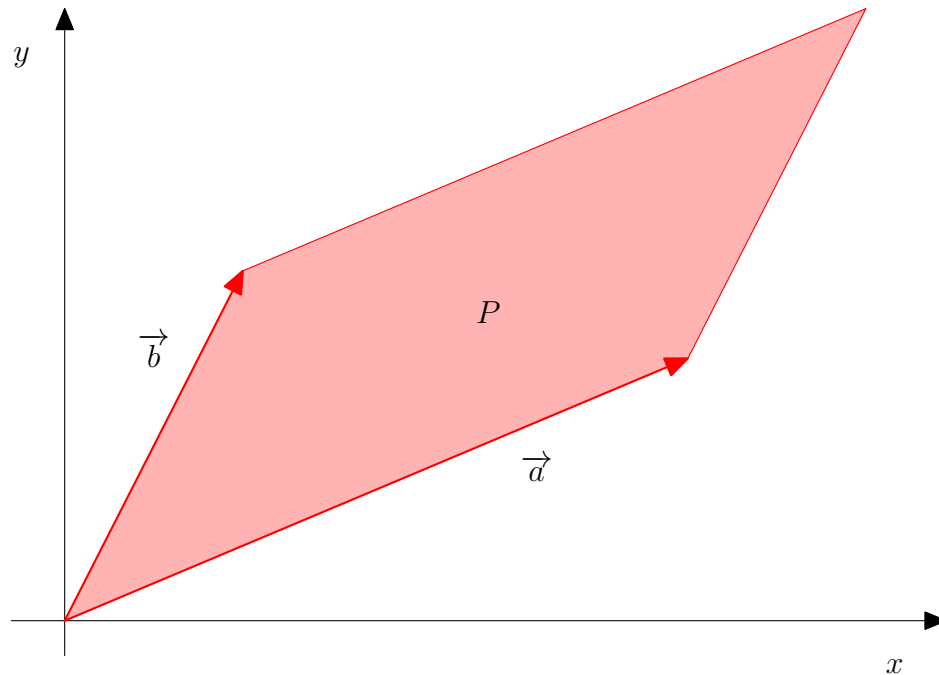


Abbildung 5.1: Determinante und Parallelogramm

Beweis: Wir schreiben die Vektoren in ihrer Polarkoordinatendarstellung

$$\vec{v} = r \cdot \begin{pmatrix} \cos(\varphi) \\ \sin(\varphi) \end{pmatrix}, \quad \vec{w} = s \cdot \begin{pmatrix} \cos(\psi) \\ \sin(\psi) \end{pmatrix}$$

und erhalten damit nach den Additionstheoremen für Sinus und Kosinus (vergleiche Vorkurs *Trigonometrie*)

$$\begin{aligned} \det(A) &= r \cdot \cos(\varphi) \cdot s \cdot \sin(\psi) - s \cdot \cos(\psi) \cdot r \cdot \sin(\varphi) \\ &= r \cdot s \cdot \sin(\psi - \varphi) \end{aligned}$$

Dabei ist entweder $\psi - \varphi$ oder $-(\psi - \varphi)$ der Winkel zwischen \vec{v} und \vec{w} . Damit ist $s \cdot |\sin(\psi - \varphi)|$ die Höhe des Parallelogramms P dessen Grundlinie durch den Vektor \vec{v} und dessen Seitenlinie durch den Vektor \vec{w} gegeben ist, wie sofort aus der Definition des Sinus folgt, und dementsprechend ist $r \cdot s \cdot |\sin(\psi - \varphi)| = |\det(A)|$ seine Fläche.

Folgerung 5.1.14. *In der Situation von Satz 5.1.13 gilt:*

$$\det(A)^2 + \langle \vec{v}, \vec{w} \rangle^2 = |\vec{v}|^2 \cdot |\vec{w}|^2$$

Beweis: Das kann man direkt nachrechnen. Wegen Satz 5.1.13 gilt aber auch

$$|\det(A)| = |\vec{v}| \cdot |\vec{w}| \cdot \sin(\alpha)$$

wobei $\alpha \in [0, \pi]$ den Winkel zwischen \vec{v} und \vec{w} bezeichnet. Damit folgt

$$\begin{aligned} \det(A)^2 + \langle \vec{v}, \vec{w} \rangle^2 &= |\vec{v}|^2 \cdot |\vec{w}|^2 \cdot \sin^2(\alpha) + |\vec{v}|^2 \cdot |\vec{w}|^2 \cdot \cos^2(\alpha) \\ &= |\vec{v}|^2 \cdot |\vec{w}|^2 (\cos^2(\alpha) + \sin^2(\alpha)) \\ &= |\vec{v}|^2 \cdot |\vec{w}|^2 \end{aligned}$$

wie gewünscht.

Aufgabe 5.1.1. Berechnen Sie $\det(A)$ für $A = \begin{pmatrix} 2 & 2 \\ -2 & 2 \end{pmatrix}$.

Aufgabe 5.1.2. Berechnen Sie $\det(A)$ für $A = \begin{pmatrix} 2 & 2 \\ 3 & 3 \end{pmatrix}$.

Aufgabe 5.1.3. Berechnen Sie $\det(A)$ für $A = \begin{pmatrix} 1 & 3 \\ 2 & 2 \end{pmatrix}$. Benutzen Sie die Cramersche Regel, um die Lösung von

$$A \cdot \vec{x} = \begin{pmatrix} 4 \\ 6 \end{pmatrix}$$

zu bestimmen.

Aufgabe 5.1.4. Berechnen Sie $\det(A)$ für $A = \begin{pmatrix} 2 & 3 \\ 1 & -1 \end{pmatrix}$. Hat das Gleichungssystem

$$A \cdot \vec{x} = \begin{pmatrix} 0 \\ 5 \end{pmatrix}$$

eine eindeutige Lösung?

Aufgabe 5.1.5. Zeigen Sie: Ist $A = \begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix}$ eine 2×2 -Matrix mit $a_{1,1} = 0$ oder $a_{2,2} = 0$, so gilt

$$\det(A) = -a_{1,2}a_{2,1}$$

Aufgabe 5.1.6. Zeigen Sie: Genau dann sind die beiden Zeilen von A kollinear, wenn $\det(A) = 0$.

5.1.2 Dreidimensionale Determinanten

Der Fall von Gleichungssystemen mit zwei Unbekannten und zwei Gleichungen oder - äquivalent - der Fall von 2×2 -Matrizen war also noch recht einfach, und wir wollen uns jetzt dem dreidimensionalen Fall zuwenden.

Ausgangspunkt unserer Überlegungen ist wieder ein Gleichungssystem, diesmal in drei Unbekannten und mit drei Gleichungen:

$$\begin{aligned} a_{1,1}x_1 + a_{1,2}x_2 + a_{1,3}x_3 &= b_1 \\ a_{2,1}x_1 + a_{2,2}x_2 + a_{2,3}x_3 &= b_2 \\ a_{3,1}x_1 + a_{3,2}x_2 + a_{3,3}x_3 &= b_3 \end{aligned} \quad (5.1.6)$$

dessen Koeffizientenmatrix wir wie üblich mit A bezeichnen.

Unser Ziel ist es wieder, dem Gleichungssystem (5.1.6) bzw. seiner Koeffizientenmatrix A eine Zahl $\det(A)$ zuzuordnen, die genau dann von Null verschieden ist, wenn das Gleichungssystem für jede Wahl von b_1, b_2 und b_3 eine eindeutige Lösung hat. Darüberhinaus sollen auch die Regeln, die wir für zweidimensionale Determinanten abgeleitet haben, in analoger Weise im dreidimensionalen Fall gelten.

Für die Berechnung der Determinante einer 2×2 -Matrix haben die Diagonale und die Gegendiagonale eine besondere Rolle gespielt. Betrachten wir im dreidimensionalen die speziellen Gleichungssysteme

$$\begin{aligned} a_{1,1}x_1 &= b_1 \\ a_{2,2}x_2 &= b_2 \\ a_{3,3}x_3 &= b_3 \end{aligned} \quad (5.1.7)$$

mit $a_{1,1} \cdot a_{2,2} \cdot a_{3,3} \neq 0$ und

$$\begin{aligned} a_{1,3}x_3 &= b_1 \\ a_{2,2}x_2 &= b_2 \\ a_{3,1}x_1 &= b_3 \end{aligned} \quad (5.1.8)$$

mit $a_{1,3} \cdot a_{2,2} \cdot a_{3,1} \neq 0$, so haben auch diese stets eine eindeutige Lösung. Damit ist zu erwarten, dass auch bei der Bildung der dreidimensionalen Determinante die Produkte der Diagonalelemente, $a_{1,1} \cdot a_{2,2} \cdot a_{3,3}$ und der Nebendiagonalelemente $a_{1,3} \cdot a_{2,2} \cdot a_{3,1}$ eine besondere Rolle spielen. Eine Bildung

der Form

$$"det"(A) = a_{1,1} \cdot a_{2,2} \cdot a_{3,3} - a_{1,3} \cdot a_{2,2} \cdot a_{3,1}$$

die der zweidimensionalen Determinante recht nahe kommt, greift aber leider zu kurz. Sie liefert zwar für die Matrix

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

ein Ergebnis wie gewünscht, nämlich den Wert 1, aber für Matrizen

$$A' = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}, \quad A'' = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \quad \text{oder} \quad A''' = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

erhalten wir immer den Wert 0, obwohl auch hier die zugehörigen Gleichungssysteme eindeutig lösbar sind, und obwohl diese Matrizen aus A durch Vertauschung von Zeilen oder Spalten hervorgegangen sind und wir daher aufgrund der Regel 5.1.9 im Dreidimensionalen den Wert 1 oder -1 erwarten. Diese Beispiele zeigen bereits, dass die Definition der dreidimensionalen Determinante nicht nur die Diagonale und die Gegendiagonale berücksichtigen muss, sondern auch die Nebendiagonalen und die Nebengegendiagonalen der Matrix. Es bleibt zu klären, wie diese anzuordnen sind, und ob noch weitere Einflussfaktoren zu berücksichtigen sind. Das ist nicht unmittelbar klar. Darum wollen wir uns noch einen anderen Zugang zuwenden.

Sind \vec{v} und \vec{w} zwei nicht kollineare ebene Vektoren, und ist A die 2×2 -Matrix, die diese Vektoren als Spalten hat, so haben wir in Satz 5.1.13 gesehen, dass $|\det(A)|$ der Flächeninhalt des von \vec{v} und \vec{w} aufgespannten Parallelograms ist. Daher ist es naheliegend, im Dreidimensionalen das folgende zu verlangen: Sind \vec{u} , \vec{v} und \vec{w} drei nicht komplanare räumliche Vektoren, und ist A die 3×3 -Matrix, die diese Vektoren als Spalten hat, so ist $|\det(A)|$ der Inhalt des von \vec{u} , \vec{v} und \vec{w} aufgespannten Parallelotops. Erfreulicherweise haben Sie im Vorkurs *Lineare Algebra* schon eine Größe kennengelernt, die diese Bedingung erfüllt, nämlich das Spatprodukt $[\vec{u}, \vec{v}, \vec{w}]$ der Vektoren \vec{u} , \vec{v} und \vec{w} . Wollen wir also diese geometrische Interpretation auch im Dreidimensionalen, so ergibt sich daraus:

Sind die drei Spaltenvektoren \vec{u} , \vec{v} und \vec{w} der Matrix A linear unabhängig, so muss gelten

$$\det(A) = \pm[\vec{u}, \vec{v}, \vec{w}]$$

Wenn wir auch noch die Normierungseigenschaft der Regel 5.1.5 im Dreidimensionalen, also $\det(E_3) = 1$, so erhalten wir sogar

$$\det(A) = [\vec{u}, \vec{v}, \vec{w}]$$

und das ist tatsächlich der richtige Ansatz, der zu folgender Formel führt:

Definition 5.1.1. Die **Determinante** $\det(A)$ einer 3×3 -Matrix

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{pmatrix}$$

ist der Ausdruck

$$\begin{aligned} \det(A) &= a_{1,1}a_{2,2}a_{3,3} + a_{1,2}a_{2,3}a_{3,1} + a_{1,3}a_{2,1}a_{3,2} \\ &\quad - a_{1,3}a_{2,2}a_{3,1} - a_{1,2}a_{2,1}a_{3,3} - a_{1,1}a_{2,3}a_{3,2} \end{aligned}$$

Bemerkung 5.1.1. Wie wir schon in unseren Vorüberlegungen festgestellt haben, spielen alle Diagonalen und alle Nebendiagonalen bei der Bildung der Determinante eine Rolle.

Beispiel 5.1.1.

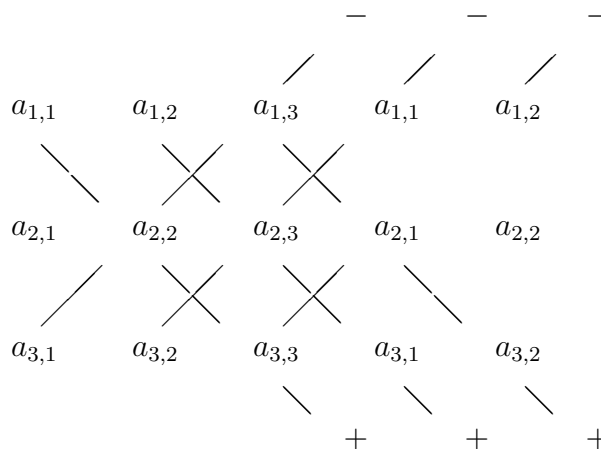
$$\begin{aligned} \det \begin{pmatrix} 1 & 2 & 0 \\ 2 & 3 & 4 \\ 1 & 0 & 3 \end{pmatrix} &= 1 \cdot 3 \cdot 3 + 2 \cdot 4 \cdot 1 + 0 \cdot 2 \cdot 0 - 0 \cdot 3 \cdot 1 - 2 \cdot 2 \cdot 3 - 1 \cdot 4 \cdot 0 \\ &= 5 \end{aligned}$$

Beispiel 5.1.2.

$$\begin{aligned} \det \begin{pmatrix} 1 & 2 & 3 \\ 0 & 3 & 4 \\ 0 & 0 & 3 \end{pmatrix} &= 1 \cdot 3 \cdot 3 + 2 \cdot 4 \cdot 0 + 3 \cdot 0 \cdot 0 - 3 \cdot 3 \cdot 0 - 2 \cdot 0 \cdot 3 - 1 \cdot 4 \cdot 0 \\ &= 9 \end{aligned}$$

Die Determinante lässt sich ganz allgemein nach einem geometrischen Schema, der **Regel von Sarrus** ermitteln

Regel 5.1.15. Zu einer 3×3 -matrix $A = (a_{i,j})$ betrachte das Schema

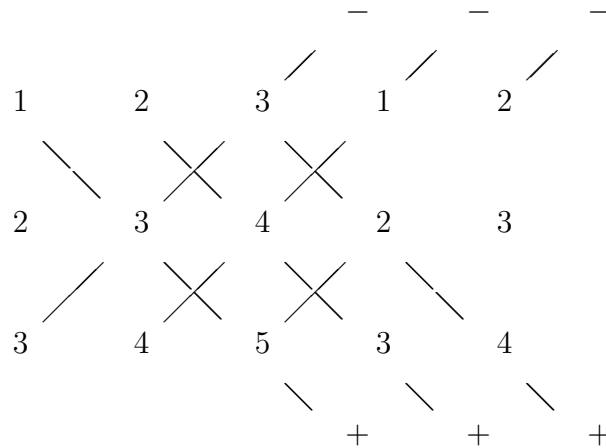


das man erhält indem man die Spalten von A hinschreibt und die ersten beiden Spalten von A noch einmal anfügt. Die Determinante von A berechnet sich dadurch, dass man jeweils die Elemente entlang der drei Diagonalen dieses 3×5 -Schemas miteinander multipliziert und aufaddiert und jeweils die Elemente entlang der drei Gegendiagonalen miteinander multipliziert und davon abzieht.

Beispiel 5.1.3. Wir betrachten die Matrix

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \end{pmatrix}$$

Dazu gehört das Sarrus-Schema



und damit die Determinante

$$1 \cdot 3 \cdot 5 + 2 \cdot 4 \cdot 3 + 3 \cdot 2 \cdot 4 - 3 \cdot 3 \cdot 3 - 1 \cdot 4 \cdot 4 - 2 \cdot 2 \cdot 5 = 0$$

Die Berechnung der Determinante im Dreidimensionalen ist also wesentlich komplizierter als im Zweidimensionalen. Daher ist es in diesem Fall besonders wichtig, Formeln und Regeln zu finden, die die Berechnungen vereinfachen. Eine Beziehung, die folgende **Entwicklungsformel**, ergibt sich unmittelbar aus der Definition des Spatprodukts, die ja der Determinante zugrunde liegt:

Satz 5.1.16 (Entwicklungssatz von Laplace (Spezialfall)). *Ist $A = (a_{i,j})$ eine 3×3 -Matrix, so gilt*

$$\begin{aligned} \det(A) &= a_{1,1} \cdot (a_{2,2}a_{3,3} - a_{2,3}a_{3,2}) - a_{2,1} \cdot (a_{1,2}a_{3,3} - a_{3,2}a_{1,3}) \\ &\quad + a_{3,1} \cdot (a_{1,2}a_{2,3} - a_{2,2}a_{1,3}) \\ &= a_{1,1} \cdot \det \begin{pmatrix} a_{2,2} & a_{2,3} \\ a_{3,2} & a_{3,3} \end{pmatrix} - a_{2,1} \cdot \det \begin{pmatrix} a_{1,2} & a_{1,3} \\ a_{3,2} & a_{3,3} \end{pmatrix} \\ &\quad + a_{3,1} \cdot \det \begin{pmatrix} a_{1,2} & a_{1,3} \\ a_{2,2} & a_{2,3} \end{pmatrix} \end{aligned}$$

Die Berechnung der dreidimensionalen Determinante lässt sich also auf die Berechnung von zweidimensionalen Determinanten zurückführen. Diese Re-

gel gilt sogar noch etwas allgemeiner. Dazu sei wieder

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{pmatrix}$$

eine beliebige 3×3 -Matrix.

Definition 5.1.2. Für $i, j \in \{1, 2, 3\}$ bezeichnen wir mit $A_{i,j}$ die i, j -Streichungsmatrix, also die 2×2 -Matrix, die wir aus A durch Streichen der i -ten Zeile und der j -ten Spalte erhalten.

Beispiel 5.1.4. Für

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$$

gilt etwa

$$\begin{aligned} A_{1,1} &= \begin{pmatrix} 5 & 6 \\ 8 & 9 \end{pmatrix} & A_{2,1} &= \begin{pmatrix} 2 & 3 \\ 8 & 9 \end{pmatrix} & A_{2,2} &= \begin{pmatrix} 1 & 3 \\ 7 & 9 \end{pmatrix} \\ A_{2,3} &= \begin{pmatrix} 1 & 2 \\ 7 & 8 \end{pmatrix} & A_{3,1} &= \begin{pmatrix} 2 & 3 \\ 5 & 6 \end{pmatrix} & A_{3,3} &= \begin{pmatrix} 1 & 2 \\ 4 & 5 \end{pmatrix} \end{aligned}$$

Beispiel 5.1.5. Für eine allgemeine Matrix

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{pmatrix}$$

gilt etwa

$$A_{3,1} = \begin{pmatrix} a_{1,2} & a_{1,3} \\ a_{2,2} & a_{2,3} \end{pmatrix} \quad A_{2,2} = \begin{pmatrix} a_{1,1} & a_{1,3} \\ a_{3,1} & a_{3,3} \end{pmatrix} \quad A_{2,3} = \begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{3,1} & a_{3,2} \end{pmatrix}$$

Diese Teilmatrizen können nun wie folgt in der Berechnung der Determinanten verwendet werden:

Satz 5.1.17 (Entwicklungssatz von Laplace). *Für eine allgemeine 3×3 -Matrix*

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{pmatrix}$$

kann $\det(A)$ nach den folgenden Regeln berechnet werden:

1. Entwicklung nach der i -ten Zeile:

Für jedes $i \in \{1, 2, 3\}$ gilt

$$\det(A) = (-1)^{i+1}a_{i,1}\det(A_{i,1}) + (-1)^{i+2}a_{i,2}\det(A_{i,2}) + (-1)^{i+3}a_{i,3}\det(A_{i,3})$$

2. Entwicklung nach der j -ten Spalte

Für jedes $j \in \{1, 2, 3\}$ gilt

$$\det(A) = (-1)^{1+j}a_{1,j}\det(A_{1,j}) + (-1)^{2+j}a_{2,j}\det(A_{2,j}) + (-1)^{3+j}a_{3,j}\det(A_{3,j})$$

Bemerkung 5.1.2. Die Formel aus Satz 5.1.16 ist der Spezialfall der Entwicklung nach der ersten Spalte.

Zu beachten ist, dass die Unterdeterminanten nicht einfach nur aufaddiert werden, vielmehr sind dabei Vorzeichen zu berücksichtigen.

Beweis: Der direkte Nachweis aller Formeln ist mühsam und rechenintensiv aber wenig herausfordernd und bleibt dem Leser überlassen.

Beispiel 5.1.6. Wir betrachten wieder die Matrix

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$$

Hierfür gilt

$$A_{2,1} = \begin{pmatrix} 2 & 3 \\ 8 & 9 \end{pmatrix} \quad A_{2,2} = \begin{pmatrix} 1 & 3 \\ 7 & 9 \end{pmatrix} \quad A_{2,3} = \begin{pmatrix} 1 & 2 \\ 7 & 8 \end{pmatrix}$$

wie wir schon in Beispiel 5.1.4 gesehen haben. Entwickeln wir die Determinante also nach der zweiten Spalte, so erhalten wir

$$\begin{aligned}\det(A) &= 2 \cdot (-1)^{2+1} \det \left(\begin{pmatrix} 2 & 3 \\ 8 & 9 \end{pmatrix} \right) + 5 \cdot (-1)^{2+2} \det \left(\begin{pmatrix} 1 & 3 \\ 7 & 9 \end{pmatrix} \right) \\ &\quad + 8 \cdot (-1)^{2+3} \det \left(\begin{pmatrix} 1 & 2 \\ 7 & 8 \end{pmatrix} \right) \\ &= -2 \cdot (-6) + 5 \cdot (-12) - 8 \cdot (-6) \\ &= 0\end{aligned}$$

Trotz der Entwicklungsregel bleibt das Berechnen von Determinanten mühsam und vor allem fehleranfällig. Daher werden weitere Regeln benötigt, die Vereinfachungen erlauben. Es gilt nämlich wie im zweidimensionalen Fall:

Regel 5.1.18. *Es sei A eine 3×3 -Matrix.*

1. Für die 3×3 -Einheitsmatrix $E_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ gilt

$$\det(E_3) = 1$$

2. $\det(A^T) = \det(A)$.

3. Es gilt schon $\det(A) = 0$, wenn

- Eine Zeile von A die Nullzeile ist.
- Eine Spalte von A die Nullspalte ist.
- Eine Zeile von A ein Vielfaches einer anderen Zeile ist.
- Eine Spalte von A ein Vielfaches einer anderen Spalte ist.

4. Ist A eine obere oder untere Dreiecksmatrix (d.h. gilt $a_{1,2} = a_{1,3} = a_{2,3} = 0$ oder $a_{2,1} = a_{3,1} = a_{3,2} = 0$), so gilt

$$\det(A) = a_{1,1} \cdot a_{2,2} \cdot a_{3,3}$$

5. Entsteht A' aus A durch Vertauschen von zwei Zeilen, so gilt

$$\det(A') = -\det(A)$$

Das Gleiche gilt, wenn A' durch Vertauschung von zwei Spalten aus A entsteht.

6. Entsteht A' aus A durch Multiplikation einer Zeile von A mit einer Zahl r , so gilt

$$\det(A') = r \cdot \det(A)$$

Das Gleiche gilt, wenn A' aus A durch Multiplikation einer Spalte von A mit einer Zahl r hervorgeht.

7. Entsteht A' aus A dadurch, dass wir ein Vielfaches einer Zeile von A zu einer anderen Zeile von A addieren, so gilt

$$\det(A') = \det(A)$$

Das Gleiche gilt, wenn A' aus A dadurch entsteht, dass wir ein Vielfaches einer Spalte von A zu einer anderen Spalte von A addieren.

8. Ist B eine weitere 3×3 -Matrix, so gilt

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

Beweis: Der Nachweis dieser Eigenschaften ist ähnlich, wenn auch aufwendiger als im zweidimensionalen Fall.

Bemerkung 5.1.3. Die Regeln erlauben es immer, die Berechnung der Determinante einer Matrix auf die Berechnung der Determinante einer Dreiecksmatrix zurückzuführen. Dieser Ansatz ist in der Praxis in der Tat der einfachste und praktikabelste Weg zur Berechnung von Determinanten.

Beispiel 5.1.7. Wir betrachten die Matrix

$$A = \begin{pmatrix} 2 & 3 & 4 \\ 4 & 6 & 9 \\ 2 & 4 & 5 \end{pmatrix}$$

- i) Wir subtrahieren das Doppelte der ersten Zeile von der zweiten Zeile und die erste Zeile von der dritten Zeile und erhalten

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

- ii) Wir vertauschen die Zeilen 2 und 3 der Matrix und beachten dabei, dass sich dadurch das Vorzeichen der Determinante ändert. Wir erhalten

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

- iii) A'' ist eine Dreiecksmatrix, hat also Determinante

$$\det(A'') = 2 \cdot (-2) \cdot 1 = -4$$

- iv) Die Operation in Schritt ii) hat zu einem Vorzeichenwechsel der Determinante geführt. Wir machen diesen rückgängig und erhalten

$$\det(A) = -\det(A'') = 4$$

Unser ursprüngliches Ziel war es, eine Größe zu finden, die uns erlaubt zu entscheiden, ob ein lineares Gleichungssystem

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

für alle Wahlen von \vec{b} eine eindeutige Lösung hat oder nicht. Wir wollen nun überprüfen, ob wir dieses Ziel mit dieser Definition der Determinante erreicht haben. Wir betrachten dazu die augmentierte Matrix $(A \mid \vec{b})$ und führen diese in Normalform $(A' \mid \vec{b}')$ über. Hierzu benötigen wir die Operationen

- Addition des Vielfachen einer Zeile von $(A \mid \vec{b})$ zu einer anderen Zeile.
- Multiplikation einer Zeile von $(A \mid \vec{b})$ mit einer Zahl $r \neq 0$.
- Vertauschung von zwei Zeilen von $(A \mid \vec{b})$.
- Vertauschung von zwei Spalten von $(A \mid \vec{b})$, von denen aber keine die letzte sein darf.

Bei all diesen Operationen wissen wir jetzt, wie sich die Determinante ändert, und wir erhalten aus Regel 5.1.18:

Bemerkung 5.1.4. Genau dann ist $\det(A) \neq 0$ wenn $\det(A') \neq 0$.

Wir können uns daher bei unseren Überlegungen auf Gleichungssysteme $A \cdot \vec{x} = \vec{b}$ in Normalform beschränken. In diesem Fall ist aber A eine Dreiecksmatrix und hat eine der folgenden Gestalten

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

In den ersten drei Fällen gilt $\det(A) = 0$, nur im vierten Fall haben wir

$$\det(A) = 1 \neq 0$$

Dieser Fall ist genau die Situation, in der $\text{rg}(A) = 3$, und in der das Gleichungssystem

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

eine eindeutige Lösung hat. Damit kann also die eindeutige Lösbarkeit tatsächlich mit Hilfe der Determinante entschieden werden und es gilt

Satz 5.1.19. Ist A eine 3×3 -Matrix, so hat das Gleichungssystem

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

genau dann eine eindeutige Lösung für jeden Vektor \vec{b} , wenn

$$\det(A) \neq 0$$

Aufgabe 5.1.7. Berechnen Sie die Determinante der folgenden Matrizen

$$A = \begin{pmatrix} 1 & 2 & 5 \\ 3 & 1 & 4 \\ 1 & 4 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 0 & 5 \\ 2 & 3 & 1 \\ 4 & 0 & 2 \end{pmatrix}, \quad C = \begin{pmatrix} 6 & 1 & 3 \\ 3 & 0 & 1 \\ 2 & 3 & -1 \end{pmatrix}$$

Aufgabe 5.1.8. Bestimmen Sie die Streichungsmatrizen $A_{1,1}$, $A_{1,2}$, $A_{1,3}$, $A_{2,1}$, $A_{2,2}$ und $A_{2,3}$ von

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

berechnen Sie die Determinanten dieser Streichungsmatrizen und die Determinante von A .

Aufgabe 5.1.9. Es sei

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

Entscheiden Sie, ob das Gleichungssystem

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

für jede Wahl von \vec{b} eine eindeutige Lösung besitzt.

Aufgabe 5.1.10. Es sei $A = (a_{i,j})$ eine 3×3 -Matrix mit

$$a_{1,1} = a_{1,2} = a_{2,1} = 0$$

Zeigen Sie

$$\det(A) = -a_{1,3} \cdot a_{2,2} \cdot a_{3,1}$$

Aufgabe 5.1.11. Zeigen Sie, dass genau dann $\det(A) = 0$ ist, wenn die Spaltenvektoren von A linear abhängig sind.

5.1.3 Die n -dimensionale Determinante

Analog zum Begriff der Vektoren, den wir uns auch zunächst im zwei- und dreidimensionalen Fall studierten ehe wir zur n -dimensionalen Situation übergingen, wollen wir jetzt auch die Determinantentheorie in der allgemeinen Situation betrachten. Der Übergang vom Zwei- zum Dreidimensionalen legt aber schon nahe, dass hier die Situation etwas komplizierter sein wird.

Es gibt viele Methoden, Determinanten zu motivieren und einzuführen. Wir wollen hier einen rekursiven Zugang verfolgen, mit dem wir auch schon die Berechnung dreidimensionaler Determinanten auf die Berechnung zweidimensionaler Determinanten zurückgeführt haben.

Sei dazu $A = (a_{i,j})$ eine beliebige $n \times n$ -Matrix. Für $l, k \in \{1, \dots, n\}$ bezeichnen wir wieder mit $A_{l,k}$ die l, k -Streichungsmatrix von A , also die $(n-1) \times (n-1)$ -Matrix, die aus A durch Streichung der l -ten Zeile und der k -ten Spalte hervorgeht. Wir definieren rekursiv

Definition 5.1.1. Die **Determinante** $\det(A)$ von A ist definiert wie folgt:

- Ist $n = 1$, also $A = (a_{1,1})$, so ist $\det(A) = a_{1,1}$.
- Ist $n = 2$ oder $n = 3$ so ist $\det(A)$ definiert wie in den Abschnitten 5.1.1 und 5.1.2.
- Ist $n > 3$ und die Determinante für $(n-1) \times (n-1)$ -Matrizen schon erklärt, so setzen wir

$$\begin{aligned} \det(A) &= a_{1,1} \cdot \det(A_{1,1}) - a_{1,2} \cdot \det(A_{1,2}) + \dots \\ &\quad + (-1)^{n+1} \cdot a_{1,n} \cdot \det(A_{1,n}) \\ &= \sum_{l=1}^n (-1)^{1+l} a_{1,l} \cdot \det(A_{1,l}) \end{aligned}$$

Bemerkung 5.1.1. In diese Rekursionsformel können wir weiter einsetzen. So ist etwa

$$\begin{aligned} \det(A_{1,1}) &= a_{2,2} \cdot \det((A_{1,1})_{1,1}) - a_{2,3} \cdot \det((A_{1,1})_{1,2}) + \dots \\ &\quad + (-1)^n \cdot a_{2,n} \cdot \det((A_{1,1})_{1,n-1}) \end{aligned}$$

Diesen Prozess können wir fortsetzen bis wir den Fall von Teilmatrizen der Größe 1×1 erreicht haben. Führen wir das tatsächlich durch, so erhalten wir

$$\det(A) = \sum_{\sigma \in S_n} \text{sign}(\sigma) a_{1,\sigma(1)} \cdot a_{2,\sigma(2)} \cdots a_{n,\sigma(n)}$$

mit einer Summe über alle Permutationen $\sigma \in S_n$ (vergleiche etwa Beispiel 2.2.26 aus Abschnitt 2.3), also einer Summe mit $n!$ vielen Summanden. Da $n!$ sehr schnell wächst, wenn n groß wird, kann man schon erkennen, dass der Rechenaufwand zur Ermittlung der Determinante sehr groß werden kann.

Beispiel 5.1.1. Wir betrachten die Matrix

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

Die Teilmatrizen, die wir benötigen, sind

$$A_{1,1} = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ -1 & 1 & 0 \end{pmatrix}, \quad A_{1,2} = \begin{pmatrix} 0 & 2 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \end{pmatrix}, \quad A_{1,3} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & -1 & 0 \end{pmatrix},$$

$$A_{1,4} = \begin{pmatrix} 0 & 1 & 2 \\ 1 & 0 & 1 \\ 1 & -1 & 1 \end{pmatrix}$$

mit

$$\det(A_{1,1}) = 0, \quad \det(A_{1,2}) = 0, \quad \det(A_{1,3}) = 0, \quad \det(A_{1,4}) = -2.$$

Also gilt

$$\det(A) = (-1)^{1+4} \cdot (-1) \cdot (-2) = -2$$

Aufgrund des mit n sehr schnell wachsenden Rechenaufwands ist es notwendig, Formeln und Regeln zu finden, die das Berechnen von Determinanten vereinfachen.

Hilfssatz 5.1.20. *Geht A' aus A durch das Vertauschen von zwei Spalten hervor, so gilt*

$$\det(A') = -\det(A)$$

Beweis: Wir betrachten dazu die Darstellung

$$\det(A) = \sum_{\sigma \in S_n} \text{sign}(\sigma) a_{1,\sigma(1)} a_{2,\sigma(2)} \cdots a_{n,\sigma(n)}$$

Falls A' aus A durch Vertauschen der Spalten l, k mit $l < k$ entsteht, so gilt also

$$\det(A') = \sum_{\sigma \in S_n} \text{sign}(\sigma) \cdot a_{1,\sigma(1)} \cdots a_{l,\sigma(k)} \cdots a_{k,\sigma(l)} \cdots a_{n,\sigma(n)}$$

Bezeichnet nun $\tau = \langle l k \rangle$ die Transposition von l und k , so schreibt sich dies als

$$\begin{aligned} \det(A') &= \sum_{\sigma \in S_n} \text{sign}(\sigma) \cdot a_{1,\sigma(1)} \cdots a_{l,\sigma(\tau(l))} \cdots a_{k,\sigma(\tau(k))} \cdots a_{n,\sigma(n)} \\ &= \sum_{\sigma \in S_n} \text{sign}(\sigma) \cdot a_{1,\sigma(\tau(1))} \cdots a_{l,\sigma(\tau(l))} \cdots a_{k,\sigma(\tau(k))} \cdots a_{n,\sigma(\tau(n))} \end{aligned}$$

(da ja τ die Zahlen $\neq l, k$ unberührt lässt). Nun haben wir aber, dass gilt

$$\text{sign}(\sigma) = -\text{sign}(\sigma \circ \tau)$$

und mit σ durchläuft auch $\sigma \circ \tau$ alle Permutationen. Daher schreibt sich diese Beziehung auch als

$$\begin{aligned} \det(A') &= \sum_{\sigma \in S_n} -\text{sign}(\sigma \circ \tau) \cdot a_{1,\sigma(\tau(1))} \cdots a_{l,\sigma(\tau(l))} \cdots a_{k,\sigma(\tau(k))} \cdots a_{n,\sigma(\tau(n))} \\ &= \sum_{\sigma \in S_n} -\text{sign}(\sigma) \cdot a_{1,\sigma(1)} \cdots a_{l,\sigma(l)} \cdots a_{k,\sigma(k)} \cdots a_{n,\sigma(n)} \\ &= -\det(A) \end{aligned}$$

wie gewünscht.

Hilfssatz 5.1.21. *Ist A^T die zu A transponierte Matrix, so gilt*

$$\det(A) = \det(A^T)$$

Beweis: Dazu benutzen wir wieder die Darstellung

$$\det(A) = \sum_{\sigma \in S_n} \text{sign}(\sigma) a_{1,\sigma(1)} a_{2,\sigma(2)} \cdots a_{n,\sigma(n)}$$

und erhalten daraus

$$\begin{aligned} \det(A) &= \sum_{\sigma \in S_n} \text{sign}(\sigma) \cdot a_{1,\sigma(1)} \cdot a_{2,\sigma(2)} \cdots a_{n,\sigma(n)} \\ &= \sum_{\sigma \in S_n} \text{sign}(\sigma) \cdot a_{(\sigma^{-1} \circ \sigma)(1), \sigma((\sigma^{-1} \circ \sigma)(1))} \cdot a_{(\sigma^{-1} \circ \sigma)(2), \sigma((\sigma^{-1} \circ \sigma)(2))} \cdots \\ &\quad \cdots a_{(\sigma^{-1} \circ \sigma)(n), \sigma((\sigma^{-1} \circ \sigma)(n))} \\ &= \sum_{\sigma \in S_n} \text{sign}(\sigma) \cdot a_{\sigma^{-1}(1), \sigma(\sigma^{-1}(1))} \cdot a_{\sigma^{-1}(2), \sigma(\sigma^{-1}(2))} \cdots a_{\sigma^{-1}(n), \sigma(\sigma^{-1}(n))} \\ &= \sum_{\sigma \in S_n} \text{sign}(\sigma) \cdot a_{\sigma(1), 1} \cdot a_{\sigma(2), 2} \cdots a_{\sigma(n), n} \\ &= \det(A^T) \end{aligned}$$

Dabei haben wir beim Übergang von Zeile 2 zu Zeile 3 die Terme im Produkt umsortiert, und beim Übergang von Zeile 3 zu Zeile 4 ausgenutzt, dass mit σ auch σ^{-1} alle Permutationen durchläuft, und dass $\text{sign}(\sigma) = \text{sign}(\sigma^{-1})$, wie wir schon in Beispiel 2.2.26 aus Abschnitt 2.3 gesehen haben.

Damit ist auch dieses Lemma gezeigt.

Folgerung 5.1.22. *Entsteht A' aus A durch Vertauschen von zwei Zeilen, so gilt*

$$\det(A') = -\det(A)$$

Beweis: Das folgt jetzt sofort aus den beiden Hilfssätzen 5.1.20 und 5.1.21, da das Vertauschen von Zeilen von A dem Vertauschen von Spalten von A^T entspricht.

Damit sind wir jetzt in der Lage einige Regeln für Determinanten herzuleiten und zusammenzustellen.

Regel 5.1.23. *Es sei A eine $n \times n$ -Matrix.*

1. *Für die $n \times n$ -Einheitsmatrix E_n gilt*

$$\det(E_n) = 1$$

2. $\det(A^T) = \det(A)$.

3. Es gilt schon $\det(A) = 0$, wenn

- Eine Zeile von A die Nullzeile ist.
- Eine Spalte von A die Nullspalte ist.
- Eine Zeile von A ein Vielfaches der anderen Zeile ist.
- Eine Spalte von A ein Vielfaches der anderen Spalte ist.

4. Ist A eine obere oder untere Dreiecksmatrix, so gilt

$$\det(A) = a_{1,1} \cdot a_{2,2} \cdots a_{n,n}$$

5. Entsteht A' aus A durch Vertauschen von zwei Zeilen, so gilt

$$\det(A') = -\det(A)$$

Das Gleiche gilt, wenn A' durch Vertauschung von zwei Spalten aus A entsteht.

6. Entsteht A' aus A durch Multiplikation einer Zeile von A mit einer Zahl r , so gilt

$$\det(A') = r \cdot \det(A)$$

Das Gleiche gilt, wenn A' aus A durch Multiplikation einer Spalte von A mit einer Zahl r hervorgeht.

7. Entsteht A' aus A dadurch, dass wir ein Vielfaches einer Zeile von A zu einer anderen Zeile von A addieren, so gilt

$$\det(A') = \det(A)$$

Das Gleiche gilt, wenn A' aus A dadurch entsteht, dass wir ein Vielfaches einer Spalte von A zu einer anderen Spalte von A addieren.

8. Schreiben wir A als $(\vec{a}_1 \dots \vec{a}_n)$ mit Spaltenvektoren $\vec{a}_1, \dots, \vec{a}_n$, so gilt

$$\begin{aligned} \det(\vec{a}_1 \dots \vec{a}_i + \vec{a}_i' \dots \vec{a}_n) &= \det(\vec{a}_1 \dots \vec{a}_i \dots \vec{a}_n) \\ &\quad + \det(\vec{a}_1 \dots \vec{a}_i' \dots \vec{a}_n) \end{aligned}$$

Die analoge Aussage gilt für die Zeilen von A .

Beweis: Einige der Regeln wurden ja schon in den vorangegangenen Lemmata bewiesen.

Wir zeigen hier beispielhaft (6) und überlassen die Nachweise der anderen Regeln, die nach einem ähnlichen Muster erfolgen, dem Leser. A' entsteht aus A durch Addition eines Vielfachen der l -ten Zeile zur k -ten Zeile, wobei $l \neq k$.

Wir führen den Beweis durch vollständige Induktion.

Die Fälle $n = 2$ und $n = 3$ sind uns schon aus den Regeln 5.1.11 bzw. 5.1.18 bekannt. Wir können annehmen, dass $n \geq 4$, und dass wir die Aussage für Determinanten kleinerer Dimension schon kennen.

Wir können auch annehmen, dass $l > 1$ und $k > 1$. Anderfalls vertauschen wir in A und A' geeignete Zeilen und ändern dadurch die Determinanten jeweils um ein Vorzeichen. Dann gilt aber

$$\det(A') = \sum_{l=1}^n (-1)^{1+l} \cdot a_{1,l} \cdot \det(A'_{1,l}) \quad (5.1.9)$$

wobei $A'_{i,l}$ aus $A_{i,l}$ jeweils durch Addition eines Vielfachen der $(k-1)$ -ten Zeile zur $(l-1)$ -ten Zeile entsteht. Damit gilt also nach Induktionsvoraussetzung

$$\det(A'_{i,l}) = \det(A_{i,l}) \quad \text{für } i = 1, \dots, n$$

und daher folgt aus 5.1.9

$$\det(A') = \det(A)$$

Beispiel 5.1.2. Wir betrachten die Matrix

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

Ohne jegliche explizite Rechnung erhalten wir, dass $\det(A) = 0$, denn die dritte Zeile ist das doppelte der ersten Zeile, und wir können Punkt (3) der Regel 5.1.23 anwenden.

Eine wiederholte Anwendung der Punkte (3) und (8) dieser Regel liefert sofort folgendes hilfreiche Resultat

Folgerung 5.1.24. *Ist eine Zeile der Matrix A eine Linearkombination der anderen Zeilen, so gilt*

$$\det(A) = 0$$

Beispiel 5.1.3. Wir betrachten die Matrix

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

Ohne jegliche Rechnung erhalten wir, dass $\det(A) = 0$, denn die dritte Zeile ist die Summer der ersten beiden Zeilen.

Hieraus erhalten wir das n -dimensionale Analog von Satz 5.1.17:

Satz 5.1.25 (Entwicklungssatz von Laplace). *Für eine allgemeine $n \times n$ -Matrix*

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{pmatrix}$$

kann $\det(A)$ nach den folgenden Regeln berechnet werden:

1. *Entwicklung nach der i -ten Zeile:*

Für jedes $i \in \{1, \dots, n\}$ gilt

$$\begin{aligned} \det(A) &= (-1)^{i+1} \cdot a_{i,1} \cdot \det(A_{i,1}) + (-1)^{i+2} \cdot a_{i,2} \cdot \det(A_{i,2}) \\ &\quad + \dots + (-1)^{i+n} \cdot a_{i,n} \cdot \det(A_{i,n}) \end{aligned}$$

2. *Entwicklung nach der j -ten Spalte*

Für jedes $j \in \{1, \dots, n\}$ gilt

$$\begin{aligned} \det(A) &= (-1)^{1+j} \cdot a_{1,j} \cdot \det(A_{1,j}) + (-1)^{2+j} \cdot a_{2,j} \cdot \det(A_{2,j}) \\ &\quad + \dots + (-1)^{n+j} \cdot a_{n,j} \cdot \det(A_{n,j}) \end{aligned}$$

Beweis: Durch $(i-1)$ -faches Vertauschen können wir die i -te Zeile zur ersten machen, ohne dadurch die Reihenfolge der anderen zu verändern. Dadurch ändert sich das Vorzeichen der Determinante um den Faktor $(-1)^{i-1} = (-1)^{i+1}$. Wenden wir jetzt die Definition der Determinante an, so erhalten wir die Formel für die Entwicklung nach der i -ten Zeile.

Die Formel für die Entwicklung nach der j -ten Spalte ergibt sich wegen Lemma 5.1.21 aus der Zeilenentwicklung, angewendet auf A^T .

Ferner können wir aus den Regeln folgendes nützliche Resultat ableiten:

Hilfssatz 5.1.26. *Entsteht A' aus A durch eine der elementaren Matrizenoperationen aus Satz 4.1.6, so gilt*

$$\det(A') \neq 0 \iff \det(A) \neq 0$$

Ist A'' eine Normalform von A , so gilt

$$\det(A'') \neq 0 \iff \det(A) \neq 0$$

Beweis: Nach den Regeln 5.1.23 ändert sich bei Durchführung einer elementaren Matrizenoperation aus Satz 4.1.6 allenfalls um ein Vorzeichen oder einen von Null verschiedenen Faktor. Hieraus folgt die Aussage für A' . Da A'' aus A durch wiederholte Anwendung dieser Operationen entsteht, folgt auch die Aussage für A'' .

Definition 5.1.2. Eine $n \times n$ -Matrix heißt **regulär**, wenn $\det(A) \neq 0$

Satz 5.1.27. Für eine $n \times n$ -Matrix A sind äquivalent:

1. A ist regulär.
2. $\operatorname{rg}(A) = n$.

Beweis: Wegen Lemma 5.1.26 ist A genau dann regulär, wenn das für eine Normalform von A gilt, und auch der Rang einer Matrix ändert sich nicht, wenn wir zur Normalform übergehen. Wir können deshalb annehmen, dass A in Normalform vorliegt. Dann ist aber A eine obere Dreiecksmatrix der Form

$$\begin{pmatrix} 1 & a_{1,2} & \dots & a_{1,t} & a_{1,t+1} & \dots & a_{1,n} \\ 0 & 1 & \dots & a_{2,t} & a_{2,t+1} & \dots & a_{2,n} \\ \vdots & & \ddots & & & & \vdots \\ 0 & 0 & \dots & 1 & a_{t,t+1} & \dots & a_{t,n} \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & & \ddots & & & & \vdots \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \end{pmatrix} \quad \text{für ein } t < n$$

oder

$$\begin{pmatrix} 1 & a_{1,2} & \dots & a_{1,n} \\ 0 & 1 & \dots & a_{2,n} \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}$$

Im ersten Fall gilt $\operatorname{rg}(A) < n$ und $\det(A) = 0$ (etwa wegen (3) aus Regel 5.1.23) und im zweiten Fall gilt $\operatorname{rg}(A) = n$ und $\det(A) = 1$ nach (4) aus Regel 5.1.23.

Bemerkung 5.1.2. Die Definition der Determinante lässt sich sehr einfach als Algorithmus umsetzen:

```

function [det] = determinante ( A )
% Berchnung der Determinante mit Rekursionsformel
% Pruefung
[n,m] = size(A);  %% Anzahl der Zeilen und Spalten

if n==1
    %% Rekursion endet hier ...
    det=A(1,1);
elseif n==2
    %% ... oder fuer n = 2
    det=A(1,1)*A(2,2)-A(2,1)*A(1,2);
elseif n > 2
    det=0;
    for i = 1:n
        %% A kopieren, da A selbst noch gebraucht wird
        Q = A;
        Q(:,i) = [];           %% Streichung i-te Spalte
        Q(1,:) = [];           %% Streichung erste Zeile
        det = det + (-1)^(i+1) * A(1,i) * determinante(Q);  %% Entwicklung
    end
end
end

```

Für große n ist das Berechnen der Determinante sowohl über die Laplace-Entwicklung als auch über die Darstellung mit der Permutationsgruppe sehr arbeitsintensiv. In diesem Fall wird in der Praxis die Determinante dadurch berechnet, dass man die Matrix in Normalform überführt und sich die Operationen, die dabei durchgeführt werden und den Wert der Determinante ändern, merkt. Dieses Verfahren wächst in seiner Komplexität viel langsamer als etwa die Laplace-Entwicklung.

Beispiel 5.1.4. Wir betrachten die Matrix

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

Im ersten Schritt subtrahieren wir das Doppelte der ersten Zeile von der zweiten und das Dreifache der ersten Zeile von der Dritten, wodurch sich die

Determinante nicht ändert, und erhalten

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

Im zweiten Schritt multiplizieren wir die zweite Zeile mit $-\frac{1}{2}$ und merken uns, dass dadurch die Determinante mit $-\frac{1}{2}$ multipliziert wird. Wir erhalten

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

Wir addieren jetzt das Doppelte der zweiten Zeile zur Dritten, wodurch sich die Determinante nicht ändert, und erhalten

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

Diese Matrix A''' ist eine obere Dreiecksmatrix, hat also Determinante $1 \cdot 1 \cdot (-1) = -1$. Um die Determinante von A zu bekommen, müssen wir berücksichtigen, dass in Schritt 2 die Determinante mit dem Faktor $-\frac{1}{2}$ multipliziert wurde und dies rückgängig machen. Wir erhalten also

$$\det(A) = 2$$

Beispiel 5.1.5. In Beispiel 5.1.4 hätten wir die Determinante auch noch leicht nach dem Entwicklungssatz von Laplace ausrechnen können. Schon etwas komplizierter ist das im Fall

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

Hier empfiehlt es sich bereits, die Regeln aus 5.1.23 anzuwenden.

Zunächst vertauschen wir die erste und die dritte Spalte und erhalten

$$B_1 = \begin{pmatrix} 1 & 0 & -2 & 5 \\ 4 & -4 & -3 & -2 \\ 5 & 1 & 5 & 4 \\ 3 & 1 & 2 & 4 \end{pmatrix}$$

Hierdurch ändert sich das Vorzeichen, $\det(B_1) = -\det(B)$. Wir subtrahieren

viermal die erste Zeile von der zweiten, fünfmal die erste von der dritten und dreimal die erste von der vierten und erhalten

$$B_2 = \begin{pmatrix} 1 & 0 & -2 & 5 \\ 0 & -4 & 5 & -22 \\ 0 & 1 & 15 & -21 \\ 0 & 1 & 8 & -11 \end{pmatrix}$$

Hierdurch ändert sich die Determinante nicht, $\det(B_2) = \det(B_1) = -\det(B)$.

Nun können wir $\det(B_2)$ schon recht einfach durch Entwicklung nach der ersten Spalte berechnen. Wir können aber auch weiter vereinfachen. Dazu vertauschen wir zunächst die zweite und die vierte Spalte:

$$B_3 = \begin{pmatrix} 1 & 0 & -2 & 5 \\ 0 & 1 & 8 & -11 \\ 0 & 1 & 15 & -21 \\ 0 & -4 & 5 & -22 \end{pmatrix}$$

Hierdurch ändert sich das Vorzeichen der Determinante, $\det(B_3) = -\det(B_2) =$

$\det(B)$. Jetzt subtrahieren wir die zweite Zeile von der dritten und addieren sie viermal zur vierten:

$$B_4 = \begin{pmatrix} 1 & 0 & -2 & 5 \\ 0 & 1 & 8 & -11 \\ 0 & 0 & 7 & -10 \\ 0 & 0 & 37 & -66 \end{pmatrix}$$

Hierdurch ändert sich die Determinante nicht, $\det(B_4) = \det(B_3) = \det(B)$.

Um uns das Leben noch einfacher zu machen, subtrahieren wir fünfmal die dritte Spalte von der Vierten und erhalten

$$B_5 = \begin{pmatrix} 1 & 0 & -2 & 5 \\ 0 & 1 & 8 & -11 \\ 0 & 0 & 7 & -10 \\ 0 & 0 & 2 & -16 \end{pmatrix}$$

Hierdurch ändert sich die Determinante nicht, $\det(B_5) = \det(B_4) = \det(B)$.

Nun berechnen wir $\det(B_5)$ indem wir zweimal nach der ersten Spalte entwickeln:

$$\begin{aligned} \det(B_5) &= 1 \cdot \det \begin{pmatrix} 8 & -11 \\ 7 & -10 \\ 2 & -16 \end{pmatrix} - 0 + 0 - 0 \\ &= 1 \cdot \det \begin{pmatrix} 7 & -10 \\ 2 & -16 \end{pmatrix} - 0 + 0 \\ &= 7 \cdot (-16) - (-10) \cdot 2 \\ &= -92 \end{aligned}$$

Also gilt $\det(B) = -92$.

Beispiel 5.1.6. Wir wollen die Determinante der Matrix

$$A = \begin{pmatrix} 1 & 2 & 3 & 2 & 1 \\ 2 & 1 & 4 & 2 & 3 \\ 7 & 6 & 5 & 4 & 3 \\ 1 & -1 & 1 & -1 & 1 \\ 2 & 3 & 1 & 0 & -1 \end{pmatrix}$$

berechnen. In diesem Fall ist die Entwicklung nach einer Zeile oder einer Spalte nicht zu empfehlen. Hierzu müssten zunächst fünf Determinanten von 4×4 -Matrizen berechnet werden, und jede dieser Determinanten erfordert wiederum die Ermittlung von vier Determinanten von 3×3 -Matrizen. Neben dem Rechenaufwand ist hier auch die Gefahr, sich zu verrechnen, zu berücksichtigen. Die numerische einfachere Methode ist es, die Matrix auf obere Dreiecksform zu bringen.

Dazusubtrahieren wir geeignete Vielfache der ersten Zeile von den Zeilen 2 - 5 und erhalten

$$A_1 = \begin{pmatrix} 1 & 2 & 3 & 2 & 1 \\ 0 & -3 & -2 & -2 & 1 \\ 0 & -8 & -16 & -10 & -4 \\ 0 & -3 & -2 & -3 & 0 \\ 0 & -1 & -5 & -4 & -3 \end{pmatrix}$$

Dadurch ändert sich die Determinante nicht, also $\det(A_1) = \det(A)$.

Nun vertauschen wir die zweite und die fünfte Spalte und erhalten

$$A_2 = \begin{pmatrix} 1 & 1 & 3 & 2 & 2 \\ 0 & 1 & -2 & -2 & -3 \\ 0 & -4 & -16 & -10 & -8 \\ 0 & 0 & -2 & -3 & -3 \\ 0 & -3 & -5 & -4 & -1 \end{pmatrix}$$

Dadurch ändert die Determinante das Vorzeichen, $\det(A_2) = -\det(A_1) = -\det(A)$.

Wir addieren viermal die zweite Zeile zur dritten und dreimal die zweite Zeile zur vierten, und erhalten

$$A_3 = \begin{pmatrix} 1 & 1 & 3 & 2 & 2 \\ 0 & 1 & -2 & -2 & -3 \\ 0 & 0 & -24 & -18 & -20 \\ 0 & 0 & -2 & -3 & -3 \\ 0 & 0 & -11 & -10 & -10 \end{pmatrix}$$

Dadurch ändert sich die Determinante nicht, also $\det(A_3) = \det(A_2) = -\det(A)$.

Nun bietet sich zunächst keine Zeile oder Spalte direkt zur Weiterarbeit an. Wir könnten die dritte Zeile durch 20 dividieren, uns diesen Faktor für die Determinantenberechnung merken und dann mit der modifizierten dritten Zeile weiterarbeiten. Das wäre die Methode, die ein Computerprogramm benutzen

würde, bei manuellen Berechnungen ist das jedoch aufgrund der auftretenden Brüche fehleranfällig. Daher schieben wir einen Zwischenschritt ein und subtrahieren sechsmal die vierte Zeile von der fünften:

$$A_4 = \begin{pmatrix} 1 & 1 & 3 & 2 & 2 \\ 0 & 1 & -2 & -2 & -3 \\ 0 & 0 & -24 & -18 & -20 \\ 0 & 0 & -2 & -3 & -3 \\ 0 & 0 & 1 & 8 & 8 \end{pmatrix}$$

Dadurch ändert sich die Determinante nicht, also $\det(A_4) = \det(A_3) = -\det(A)$.

Nun vertauschen wir die Zeilen 3 und 5:

$$A_5 = \begin{pmatrix} 1 & 1 & 3 & 2 & 2 \\ 0 & 1 & -2 & -2 & -3 \\ 0 & 0 & 1 & 8 & 8 \\ 0 & 0 & -2 & -3 & -3 \\ 0 & 0 & -24 & -18 & -20 \end{pmatrix}$$

Dadurch ändert die Determinante das Vorzeichen, $\det(A_5) = -\det(A_4) = \det(A)$.

Wir addieren das doppelte der dritten Zeile zur vierten und das vierundzwanzigfache der dritten zur fünften und erhalten

$$A_6 = \begin{pmatrix} 1 & 1 & 3 & 2 & 2 \\ 0 & 1 & -2 & -2 & -3 \\ 0 & 0 & 1 & 8 & 8 \\ 0 & 0 & 0 & 13 & 13 \\ 0 & 0 & 0 & 174 & 172 \end{pmatrix}$$

Dadurch ändert sich die Determinante nicht, also $\det(A_6) = \det(A_5) = \det(A)$.

Nun können wir die Determinante bereits explizit berechnen, indem wir sukzessive nach der ersten Spalte entwickeln:

$$\det(A_6) = 1 \cdot 1 \cdot 1 \cdot (13 \cdot 172 - 13 \cdot 174) = -26$$

Alternativ können wir aber auch noch die fünfte Spalte von der vierten subtrahieren und erhalten

$$A_7 = \begin{pmatrix} 1 & 1 & 3 & 0 & 2 \\ 0 & 1 & -2 & 1 & -3 \\ 0 & 0 & 1 & 0 & 8 \\ 0 & 0 & 0 & 0 & 13 \\ 0 & 0 & 0 & 2 & 172 \end{pmatrix}$$

Dadurch ändert sich die Determinante nicht, also $\det(A_7) = \det(A_6) = \det(A)$.

Vertauschen wir jetzt die Zeilen 4 und 5, so ergibt sich

$$A_8 = \begin{pmatrix} 1 & 1 & 3 & 0 & 2 \\ 0 & 1 & -2 & 1 & -3 \\ 0 & 0 & 1 & 0 & 8 \\ 0 & 0 & 0 & 2 & 172 \\ 0 & 0 & 0 & 0 & 13 \end{pmatrix}$$

und ein Vorzeichenwechsel in der Determinante, $\det(A_8) = -\det(A_7) = -\det(A)$. Jetzt haben wir obere Dreiecksform erreicht und lesen also ab, dass

$$\det(A_8) = 1 \cdot 1 \cdot 1 \cdot 2 \cdot 13 = 26$$

also auch hier $\det(A) = -26$.

Bei der Definition der Determinante haben die (i, j) -Steichungsmatrizen $A_{i,j}$ und ihre Determinanten eine wichtige Rolle gespielt. Wir setzen

$$\widetilde{a_{i,j}} = (-1)^{i+j} \cdot \det(A_{j,i})$$

(beachten Sie dabei die Vertauschung der Indizes).

Definition 5.1.3. Die Matrix $\widetilde{A} := (\widetilde{a_{i,j}})$ heißt die zu A **komplementäre** Matrix

Beispiel 5.1.7. Sei

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

Dann ist

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

Beispiel 5.1.8. Sei

$$A = \begin{pmatrix} 1 & 0 & -1 \\ 8 & -4 & -1 \\ -2 & 1 & 0 \end{pmatrix}$$

Dann ist

$$\tilde{A} = \begin{pmatrix} 1 & -1 & -4 \\ 2 & -2 & -7 \\ 0 & -1 & -4 \end{pmatrix}$$

Satz 5.1.28. *Es gilt*

$$\tilde{A} \cdot A = A \cdot \tilde{A} = \det(A) \cdot E_n$$

ist eine Diagonalmatrix mit $\det(A)$ als Eintrag an jeder Stelle der Diagonale.

Beweis: Wir berechnen dazu explizit den Matrixeintrag $(\tilde{A} \cdot A)_{k,l}$ an jeder Stelle k, l .

Ist $l = k$, so gilt:

$$\begin{aligned} (\tilde{A} \cdot A)_{l,l} &= \sum_{i=1}^n \widetilde{a_{l,i}} \cdot a_{i,l} \\ &= \sum_{i=1}^n (-1)^{i+l} \det(A_{i,l}) \cdot a_{i,l} \\ &= \sum_{i=1}^n (-1)^{i+l} a_{i,l} \cdot \det(A_{i,l}) \\ &= \det(A) \end{aligned}$$

nach dem Laplace-Entwicklungssatz (Entwicklung nach der k -ten Spalte).

Ist $l \neq k$ so betrachten wir die Matrix B , die aus A dadurch entsteht, dass wir die k -te Spalte von A durch die l -te Spalte von A ersetzen. Damit hat

B zwei identische Spalten (an der Stelle k und der Stelle l), und daher gilt nach Regel 5.1.23

$$\det(B) = 0$$

Andererseits gilt $B_{i,k} = A_{i,k}$, da B und A in allen Spalten ausser der k -ten übereinstimmen. Entwickeln wir daher B nach der k -ten Spalte, so erhalten wir

$$\begin{aligned} \det(B) &= \sum_{i=1}^n (-1)^{i+k} b_{i,k} \cdot \det(B_{i,k}) \\ &= \sum_{i=1}^n (-1)^{i+k} a_{i,k} \cdot \det(A_{i,k}) \\ &= \sum_{i=1}^n (-1)^{i+k} \widetilde{a_{k,i}} \cdot a_{i,k} \\ &= \left(\widetilde{A} \cdot A \right)_{k,k} \end{aligned}$$

Daraus folgt $\widetilde{A} \cdot A = \det(A) \cdot E_n$, und die Gleichheit $A \cdot \widetilde{A} = \det(A) \cdot E_n$ sieht man genauso.

Beispiel 5.1.9. Für die Matrix A aus Beispiel 5.1.7 gilt

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

und $\det(A) = 1$.

Beispiel 5.1.10. Für die Matrix A aus Beispiel 5.1.8 gilt

$$A \cdot \widetilde{A} = \begin{pmatrix} 1 & 0 & -1 \\ 8 & -4 & 1 \\ -2 & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & -1 \\ 8 & -4 & -1 \\ -2 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & -1 & -4 \\ 2 & -2 & -7 \\ 0 & -1 & -4 \end{pmatrix} = 1 \cdot E_3$$

und $\det(A) = 1$.

Ein wichtiges aber schwieriges Resultat ist

Satz 5.1.29. (Produktsatz)

Sind A und B zwei $n \times n$ -Matrizen, so gilt

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

Wir werden den Beweis dieses Satzes im Abschnitt 5.2 skizzieren.

Aufgabe 5.1.12. Berechnen Sie die Determinanten der Matrizen

$$A = \begin{pmatrix} 1 & 2 & 3 & 5 \\ 2 & 5 & 7 & 9 \\ -1 & 2 & -3 & 4 \\ 0 & 1 & 0 & 2 \end{pmatrix}, \quad B = \begin{pmatrix} 7 & 1 & 5 & 2 \\ 4 & -4 & 2 & -2 \\ 1 & 0 & -2 & 5 \\ 1 & 1 & 2 & 2 \end{pmatrix}, \quad C = \begin{pmatrix} 6 & -6 & 0 & 2 \\ 4 & 1 & 1 & 4 \\ -1 & -2 & 3 & 4 \\ 2 & 1 & 2 & 1 \end{pmatrix}$$

Aufgabe 5.1.13. Berechnen Sie die Determinante der Matrix

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

Aufgabe 5.1.14. Ist A eine $n \times n$ Matrix, die sich in der Form

$$A = \begin{pmatrix} B & C \\ 0 & D \end{pmatrix}$$

mit einer $n_1 \times n_1$ -Matrix B , einer $n_1 \times (n - n_1)$ -Matrix C und einer $(n - n_1) \times (n - n_1)$ -Matrix D (sowie einer $(n - n_1) \times n_1$ -Nullmatrix 0) schreiben lässt, so gilt

$$\det(A) = \det(B) \cdot \det(D)$$

Aufgabe 5.1.15. Berechnen Sie die Komplementärmatrizen zu den Matrizen

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

Aufgabe 5.1.16. Für reelle Zahlen r_1, \dots, r_n ($n \geq 2$) definieren wir die $n \times n$ -Matrix

$$V(r_1, \dots, r_n) = \begin{pmatrix} 1 & 1 & \dots & 1 \\ r_1 & r_2 & \dots & r_n \\ r_1^2 & r_2^2 & \dots & r_n^2 \\ \vdots & & \ddots & \vdots \\ r_1^{n-1} & r_2^{n-1} & \dots & r_n^{n-1} \end{pmatrix}$$

Zeigen Sie: $\det(V(r_1, \dots, r_n)) = \prod_{1 \leq i < j \leq n} (r_j - r_i)$.

5.2 Determinanten und invertierbare Matrizen

Wir haben in den vorangegangenen Abschnitten gesehen, dass $n \times n$ -Matrizen A mit $\det(A) \neq 0$ in der Menge der Matrizen eine ausgezeichnete Rolle spielen und diese Matrizen als reguläre Matrizen bezeichnet. In diesem Abschnitt wollen wir uns einer anderen ausgezeichneten Eigenschaft einiger Matrizen zuwenden.

Definition 5.2.1. Eine $n \times n$ -Matrizen A heißt **invertierbar**, wenn es eine $n \times n$ -Matrizen B gibt mit

$$A \cdot B = E_n, \quad B \cdot A = E_n$$

Bezeichnung:

Ist A invertierbar, so nennen wir die Matrix B aus der Definition die *zu A inverse Matrix* und bezeichnen sie mit A^{-1} .

Bemerkung 5.2.1. Zunächst stellt sich die Frage, ob vielleicht jede Matrix invertierbar ist. Betrachten wir dazu zunächst $A = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$, also die Nullmatrix, so gilt für jede Matrix B :

$$A \cdot B = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

und damit kann es sicherlich keine Matrix mit $A \cdot B = E_2$ geben.

Bei den reellen Zahlen ist jede von Null verschiedene reelle Zahl invertierbar. Ist also vielleicht jede von der Nullmatrix verschiedene Matrix invertierbar? Auch das ist nicht der Fall, wie das Beispiel $A = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$ zeigt. Ist nämlich B eine beliebige 2×2 -Matrix, so gilt immer

$$A \cdot B = \begin{pmatrix} * & * \\ 0 & 0 \end{pmatrix}$$

die zweite Zeile ist also immer die Nullzeile, und damit kann es auch hier keine Matrix B mit $A \cdot B = E_2$ geben.

Bemerkung 5.2.2. Auf dem Vektorraum $\text{Matr}(n, n)$ haben wir neben der Matrizenaddition auch noch die Matrizenmultiplikation. Es ist leicht einzusehen, dass $(\text{Matr}(n, n), +, \cdot)$ ein (nicht-kommutativer) Ring mit Einselement

E_n ist. Wir suchen nun die Elemente in diesem Ring, die ein multiplikatives Inverses haben.

Bemerkung 5.2.3. Es reicht, zu einer $n \times n$ -Matrix A ein Linksinverses und ein Rechtsinverses zu finden, also eine Matrix B mit $B \cdot A = E_n$ und eine Matrix C mit $A \cdot C = E_n$. Dann gilt schon

$$B = B \cdot E_n = B \cdot A \cdot C = E_n \cdot C = C$$

Genauso zeigt man, dass eine inverse Matrix, falls sie existiert, eindeutig ist: Sind B und C zwei inverse Matrizen zu A , so gilt

$$B = B \cdot E_n = B \cdot A \cdot C = E_n \cdot C = C$$

Inverse Matrizen sind ein hilfreiches Instrument bei der Lösung linearer Gleichungssysteme

Bemerkung 5.2.4. Ist A eine invertierbare Matrix mit Inverser A^{-1} so hat das Gleichungssystem

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

die eindeutige Lösung

$$\vec{x} = A^{-1} \cdot \vec{b}$$

Beweis: Ist \vec{x} eine Lösung des Gleichungssystems, so muss gelten

$$\vec{x} = A^{-1} \cdot A \cdot \vec{x} = A^{-1} \cdot \vec{b}$$

Setzen wir umgekehrt $\vec{x} = A^{-1} \cdot \vec{b}$, so erhalten wir

$$A \cdot \vec{x} = A \cdot A^{-1} \cdot \vec{b} = \vec{b}$$

Wir konzentrieren uns zunächst auf das Finden einer Rechtsinversen, also auf das Finden einer Matrix B mit $A \cdot B = E_n$. Dazu sei $\vec{e}_1, \dots, \vec{e}_n$ die Standardbasis des \mathbb{R}^n .

Satz 5.2.1. *Genau dann existiert eine rechtsinverse Matrix B von A , wenn die Gleichungssysteme*

$$A \cdot \vec{x} = \vec{e}_i$$

für jedes $i = 1, \dots, n$ lösbar ist. Ist in diesem Fall \vec{v}_i eine Lösung von $A \cdot \vec{x} = \vec{e}_i$ und ist B die Matrix mit den Spaltenvektoren $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$, so gilt

$$A \cdot B = E_n$$

Beweis: Ist B eine Rechtsinverse zu A und ist $B_{\bullet,i}$ der i -te Spaltenvektor von B , so gilt nach Definition der Matrizenmultiplikation gerade $A \cdot B_{\bullet,i} = \vec{e}_i$, und damit sind alle fraglichen Gleichungssysteme lösbar.

Sind umgekehrt $\vec{v}_1, \dots, \vec{v}_n$ Lösungen von $A \cdot \vec{x} = \vec{e}_i$ und ist B die Matrix mit den $\vec{v}_1, \dots, \vec{v}_n$ als Spaltenvektoren, so ist - wiederum nach Definition der Matrizenmultiplikation - $A \cdot B$ die Matrix mit den Spaltenvektoren $A \cdot \vec{v}_1, \dots, A \cdot \vec{v}_n$, also

$$A \cdot B = (\vec{e}_1 \ \vec{e}_2 \ \dots \ \vec{e}_n) = E_n$$

Folgerung 5.2.2. *Genau dann existiert zu A eine rechtsinverse Matrix, wenn $\text{rg}(A) = n$.*

Beweis: Nach Satz 4.3.2 aus Abschnitt 4.3 sind die Gleichungssysteme $A \vec{x} = \vec{e}_i$ genau dann für alle i lösbar, wenn $\vec{e}_i \in \text{Im}(A)$ ist für alle i , und das ist äquivalent zu

$$\text{Im}(A) = \mathbb{R}^n$$

also zu $\text{rg}(A) = n$. Aus dem Satz folgt jetzt die Behauptung.

Wir haben also die Frage nach einer rechtsinversen Matrix zu A gelöst und können uns jetzt der Suche nach einer linksinversen Matrix zuwenden. Hierfür benötigen wir einen kleinen Hilfssatz

Hilfssatz 5.2.3. *Genau dann hat A eine linksinverse Matrix, wenn A^T eine rechtsinverse Matrix hat.*

Ist in diesem Fall C eine rechtsinverse Matrix von A^T , so ist C^T eine linksinverse Matrix von A .

Beweis: Ist C eine rechtsinverse Matrix zu A^T , so gilt nach den Regeln zu Transponieren von Matrizen

$$C^T \cdot A = \left((C^T \cdot A)^T \right)^T = \left(A^T \cdot (C^T)^T \right)^T = (A^T \cdot C)^T = E_n^T = E_n$$

Ist umgekehrt D eine linksinverse Matrix zu A so rechnet man genauso nach, dass D^T eine rechtsinverse Matrix zu A^T ist.

Satz 5.2.4. *Für eine $n \times n$ -Matrix A sind äquivalent:*

1. A ist regulär.
2. A ist invertierbar.
3. $\text{rg}(A) = n$.

Beweis: Wir haben in Satz 5.1.27 schon gesehen, dass die Regularität von A äquivalent zu $\text{rg}(A) = n$ ist und brauchen also nur noch zeigen, dass A genau dann invertierbar ist, wenn $\text{rg}(A) = n$.

Nach Satz 5.2.1 existiert zu A genau dann eine rechtsinverse Matrix, wenn $\text{rg}(A) = n$. Aus Lemma 5.2.3 folgt, dass A genau dann eine linksinverse Matrix hat, wenn $\text{rg}(A^T) = n$. Da nach Satz 4.3.8 aus Abschnitt 4.3 gilt

$$\text{rg}(A) = \text{rg}(A^T)$$

folgt hieraus mit Bemerkung 5.2.3 die Behauptung.

Zur Bestimmung einer inversen Matrix reicht es also, eine rechtsinverse Matrix von A zu ermitteln. Dazu liefert Satz 5.2.1 ein probates Mittel. Sowohl um zu entscheiden, ob es eine zu A inverse Matrix gibt als auch um diese zu finden müssen wir nur die Gleichungssysteme $A \cdot \vec{x} = \vec{e}_i$ für alle $i = 1, \dots, n$ lösen. Nach Bemerkung 4.1.4 aus Abschnitt 4.1.2 können wir diese Gleichungssysteme alle in einem Durchgang durch Betrachtung der erweiterten augmentierten Matrix $(A \mid \vec{e}_1 \mid \vec{e}_2 \mid \dots \mid \vec{e}_n)$ und Überführen dieser Matrix in Normalform lösen. Ferner können wir diese erweiterte Matrix gemäß Satz 4.1.4 nur durch Zeilenoperationen auf die Normalform

$$\left(\begin{array}{cccc|cccc} 1 & 0 & \dots & 0 & \vec{b}_1 & \vec{b}_2 & \dots & \vec{b}_n \\ 0 & 1 & \dots & 0 & & & & \\ \vdots & & \ddots & \vdots & & & & \\ 0 & 0 & \dots & 1 & & & & \end{array} \right)$$

bringen, und daraus folgt

Folgerung 5.2.5. $A^{-1} = (\vec{b}_1 \ \dots \ \vec{b}_n)$ ist die Matrix mit den \vec{b}_i als Spaltenvektoren.

Beispiel 5.2.1. Wir betrachten die Matrix

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

und gehen hier vor wie folgt:

$$\begin{array}{ll} (A \quad \vec{e}_1 \ \vec{e}_2) &= \left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 2 & 3 & 0 & 1 \end{array} \right) \\ II - 2 \cdot I &\left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & -1 & -2 & 1 \end{array} \right) \\ (-1) \cdot II &\left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & 1 & 2 & -1 \end{array} \right) \\ I - 2 \cdot II &\left(\begin{array}{cc|cc} 1 & 0 & -3 & 2 \\ 0 & 1 & 2 & -1 \end{array} \right) \end{array}$$

wobei wir in der linken Spalte die durchgeführten Zeilenoperationen notieren ("II - 2 · I" bedeutet also: Subtrahiere das Doppelte der ersten Zeile von der zweiten). Wir erhalten, dass A invertierbar ist mit

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

und wir überzeugen uns davon, dass in der Tat

$$A \cdot A^{-1} = E_2 = A^{-1} \cdot A$$

Beispiel 5.2.2. Wir betrachten die Matrix

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

und gehen vor wie folgt:

$$\begin{aligned}
 (A \mid \vec{e}_1 \vec{e}_2 \vec{e}_3) &= \left(\begin{array}{ccc|ccc} 0 & 1 & 2 & 1 & 0 & 0 \\ 1 & 2 & 2 & 0 & 1 & 0 \\ 1 & 3 & 3 & 0 & 0 & 1 \end{array} \right) \\
 I \leftrightarrow II &\left(\begin{array}{ccc|ccc} 1 & 2 & 2 & 0 & 1 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 \\ 1 & 3 & 3 & 0 & 0 & 1 \end{array} \right) \\
 III - I &\left(\begin{array}{ccc|ccc} 1 & 2 & 2 & 0 & 1 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & -1 & 1 \end{array} \right) \\
 III - II &\left(\begin{array}{ccc|ccc} 1 & 2 & 2 & 0 & 1 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 \\ 0 & 0 & -1 & -1 & -1 & 1 \end{array} \right) \\
 (-1) \cdot III &\left(\begin{array}{ccc|ccc} 1 & 2 & 2 & 0 & 1 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & -1 \end{array} \right) \\
 \begin{array}{l} II - 2 \cdot III \\ I - 2 \cdot III \end{array} &\left(\begin{array}{ccc|ccc} 1 & 2 & 0 & -2 & -1 & 2 \\ 0 & 1 & 0 & -1 & -2 & 2 \\ 0 & 0 & 1 & 1 & 1 & -1 \end{array} \right) \\
 I - 2 \cdot II &\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & 3 & -2 \\ 0 & 1 & 0 & -1 & -2 & 2 \\ 0 & 0 & 1 & 1 & 1 & -1 \end{array} \right)
 \end{aligned}$$

Wir erhalten auch hier, dass A invertierbar ist mit

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

Beispiel 5.2.3. Nun betrachten wir die Matrix

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

und gehen vor wie folgt:

$$\begin{aligned}
 (A \mid \vec{e}_1 \vec{e}_2 \vec{e}_3) &= \left(\begin{array}{ccc|ccc} 0 & 1 & 2 & 1 & 0 & 0 \\ 1 & 2 & 2 & 0 & 1 & 0 \\ 1 & 3 & 4 & 0 & 0 & 1 \end{array} \right) \\
 I \leftrightarrow II &\quad \left(\begin{array}{ccc|ccc} 1 & 2 & 2 & 0 & 1 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 \\ 1 & 3 & 4 & 0 & 0 & 1 \end{array} \right) \\
 III - II &\quad \left(\begin{array}{ccc|ccc} 1 & 2 & 2 & 0 & 1 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 & -1 & 1 \end{array} \right)
 \end{aligned}$$

An dieser Stelle bricht der Algorithmus ab. Die letzte Zeile ist zur Nullzeile geworden, und damit hat die Matrix A nicht den vollen Rang 3, ist also auch nicht invertierbar.

Die elementaren Zeilenoperationen, die wir durchführen, um A zu invertieren, lassen sich auch mit Hilfe von Matrizen beschreiben.

Wir betrachten dazu drei Typen von Elementarmatrizen, $M_i(\lambda)$ für ein $\lambda \in \mathbb{R} \setminus \{0\}$, $P_{i,j}$ für $i \neq j$ und $S_{i,j}(\nu)$ für $i \neq j$ und $\nu \in \mathbb{R}$, die wie folgt definiert sind:

$M_i(\lambda) = (m_{l,k})$ mit

$$m_{l,k} = \begin{cases} 1 & \text{falls } l = k \text{ und } l \neq i \\ \lambda & \text{falls } l = k = i \\ 0 & \text{sonst} \end{cases}$$

$P_{i,j} = (p_{l,k})$ mit

$$p_{l,k} = \begin{cases} 1 & \text{falls } l = k \text{ und } l \neq i, l \neq j \\ 1 & \text{falls } l = i, k = j \\ 1 & \text{falls } l = j, k = i \\ 0 & \text{sonst} \end{cases}$$

$S_{i,j}(\nu) = (s_{l,k})$ mit

$$s_{l,k} = \begin{cases} 1 & \text{falls } l = k \\ \nu & \text{falls } l = i, k = j \\ 0 & \text{sonst} \end{cases}$$

Beispiel 5.2.4. Es sei $n = 4$.

$$M_2(3) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad P_{2,3} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad S_{1,4}(7) = \begin{pmatrix} 1 & 0 & 0 & 7 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Bemerkung 5.2.5. Diese Elementarmatrizen haben die folgende Eigenschaft:

- $M_i(\lambda) \cdot A$ ist die Matrix, die aus A durch Multiplikation der i -ten Zeile mit λ hervorgeht.
- $P_{i,j} \cdot A$ ist die Matrix, die aus A durch Vertauschen der Zeilen i und j hervorgeht.
- $S_{i,j}(\nu) \cdot A$ ist die Matrix, die aus A durch Addition des ν -fachen der j -ten Zeile zur i -ten Zeile hervorgeht.

Damit können also die Zeilenoperationen, die eine invertierbare Matrix A in die Einheitsmatrix E_n überführen, als Multiplikationen mit diesen Elementarmatrizen beschrieben werden. Es gibt also Elementarmatrizen $\mathcal{E}_1, \dots, \mathcal{E}_l$ mit

$$\mathcal{E}_1 \cdots \mathcal{E}_l \cdot A = E_n$$

und damit gilt

$$A^{-1} = \mathcal{E}_1 \cdots \mathcal{E}_l \quad \text{und} \quad A = \mathcal{E}_l^{-1} \cdots \mathcal{E}_1^{-1}$$

Da die Inverse einer Elementarmatrix wieder eine Elementarmatrix ist,

$$M_i(\lambda)^{-1} = M_i\left(\frac{1}{\lambda}\right), \quad P_{i,j}^{-1} = P_{i,j} \quad \text{und} \quad S_{i,j}(\nu)^{-1} = S_{i,j}(-\nu)$$

haben wir gezeigt: Jede invertierbare Matrix A ist Produkt von Elementarmatrizen,

$$A = \mathcal{E}_1 \cdots \mathcal{E}_l$$

Invertierbare Matrizen A sind dadurch charakterisiert, dass $\det(A) \neq 0$.

Determinanten können aber auch dazu benutzt werden, um inverse Matrizen zu berechnen:

Satz 5.2.6. *Ist A invertierbar, so gilt*

$$A^{-1} = \frac{1}{\det(A)} \cdot \tilde{A}$$

mit der zu A komplementären Matrix \tilde{A} aus Abschnitt 5.1.3.

Beweis: Das haben wir im Prinzip schon mit Satz 5.1.28 gezeigt.

Beispiel 5.2.5. Für die Matrix $A = \begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix}$ haben wir schon in Beispiel 5.1.7 gesehen, dass $\det(A) = 1$ und

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

Damit ist $A^{-1} = \tilde{A}$.

Beispiel 5.2.6. Ist $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ eine beliebige 2×2 -Matrix so ist

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

Ist also $ad - bc \neq 0$, so ist A invertierbar mit

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

Beispiel 5.2.7. Ist $A = \begin{pmatrix} 1 & 0 & -1 \\ 8 & -4 & -1 \\ -2 & 1 & 0 \end{pmatrix}$ so haben wir in Beispiel 5.1.7

gesehen, dass $\det(A) = 1$ und

$$\tilde{A} = \begin{pmatrix} 1 & -1 & -4 \\ 2 & -2 & -7 \\ 0 & -1 & -4 \end{pmatrix}$$

Daher ist $A^{-1} = \tilde{A}$.

Bemerkung 5.2.6. Für große n ist es unpraktikabel, die inverse Matrix über die Komplementärmatrix zu berechnen, da der Rechenaufwand zu hoch ist.

Im Allgemeinen ist es nicht einfach, zu entscheiden, ob eine $n \times n$ -Matrix invertierbar ist (speziell für großes n) und gegebenenfalls ihr Inverses zu berechnen. Einige Regeln helfen dabei manchmal.

Regel 5.2.7. • Sind A und B invertierbare $n \times n$ -Matrizen, so ist auch $A \cdot B$ invertierbar und

$$(A \cdot B)^{-1} = B^{-1} \cdot A^{-1}$$

• Ist A eine Diagonalmatrix,

$$A = \begin{pmatrix} d_1 & 0 & \dots & 0 \\ 0 & d_2 & \dots & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & \dots & d_n \end{pmatrix},$$

so ist A genau dann invertierbar, wenn $d_i \neq 0$ für alle i , und in diesem Fall gilt

$$A^{-1} = \begin{pmatrix} \frac{1}{d_1} & 0 & \dots & 0 \\ 0 & \frac{1}{d_2} & \dots & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & \dots & \dots & \frac{1}{d_n} \end{pmatrix}$$

• Ist A eine $n \times n$ -Matrix mit $A^l = 0$ für ein $l \in \mathbb{N}$, so ist $E_n - A$ invertierbar und $(E_n - A)^{-1} = E_n + A + A^2 + \dots + A^{l-1}$.

Beweis: Die Aussagen (1) und (2) lassen sich unmittelbar nachrechnen. Zum Nachweis von (3) rechnen wir

$$\begin{aligned} (E_n - A) &\cdot (E_n + A + A^2 + \dots + A^{l-1}) \\ &= E_n + A + A^2 + \dots + A^{l-1} - A - A^2 - A^3 - \dots - A^l \\ &= E_n - A^l \\ &= E_n \end{aligned}$$

Also ist $E_n - A$ invertierbar.

Bemerkung 5.2.7. Eine Matrix A mit $A^l = 0$ für ein $l \in \mathbb{N}$ heißt **nilpotent**.

Definition 5.2.2. Eine Matrix $n \times n$ -Matrix A heißt **orthogonal**, wenn die Spalten von A eine Orthonormalbasis von \mathbb{R}^n bilden.

Beispiel 5.2.8. Die Matrix

$$A = \begin{pmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix}$$

ist orthogonal.

Satz 5.2.8. Ist A eine orthogonale Matrix, so ist A invertierbar und

$$A^{-1} = A^T$$

ist die Transponierte von A .

Beweis: Wir bezeichnen mit $\vec{u}_1, \dots, \vec{u}_n$ die Spaltenvektoren von A . Dann gilt nach Definition der Matrizenmultiplikation

$$\begin{aligned} A^T \cdot A &= \begin{pmatrix} \langle \vec{u}_1, \vec{u}_1 \rangle & \langle \vec{u}_2, \vec{u}_1 \rangle & \dots & \langle \vec{u}_n, \vec{u}_1 \rangle \\ \langle \vec{u}_1, \vec{u}_2 \rangle & \langle \vec{u}_2, \vec{u}_2 \rangle & \dots & \langle \vec{u}_n, \vec{u}_2 \rangle \\ \vdots & & \ddots & \vdots \\ \langle \vec{u}_1, \vec{u}_n \rangle & \langle \vec{u}_2, \vec{u}_n \rangle & \dots & \langle \vec{u}_n, \vec{u}_n \rangle \end{pmatrix} \\ &= E_n \end{aligned}$$

nach Definition einer Orthonormalbasis.

Beispiel 5.2.9. Die Matrix

$$A = \begin{pmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix}$$

ist invertierbar mit inverser Matrix

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

Folgerung 5.2.9. Ist A eine orthogonale $n \times n$ -Matrix, so gilt

$$\langle A \cdot \vec{v}, A \cdot \vec{w} \rangle = \langle \vec{v}, \vec{w} \rangle \quad \text{für alle } \vec{v}, \vec{w} \in \mathbb{R}^n$$

Beweis: Wir benutzen Regel 4.2.6 und Satz 5.2.8 und erhalten

$$\begin{aligned} \langle A \cdot \vec{v}, A \cdot \vec{w} \rangle &= \langle \vec{v}, A^T \cdot A \cdot \vec{w} \rangle \\ &= \langle \vec{v}, A^{-1} \cdot A \cdot \vec{w} \rangle \\ &= \langle \vec{v}, \vec{w} \rangle \end{aligned}$$

Bemerkung 5.2.8. Orthogonale Matrizen haben folgende interessante Eigenschaft: Ist A eine orthogonale Matrix und \vec{v} ein beliebiger Vektor, so gilt

$$|A \cdot \vec{v}| = |\vec{v}|$$

Das folgt jetzt sofort aus Korollar 5.2.9, denn damit gilt

$$\begin{aligned} |A \cdot \vec{v}|^2 &= \langle A \cdot \vec{v}, A \cdot \vec{v} \rangle \\ &= \langle \vec{v}, \vec{v} \rangle \\ &= |\vec{v}|^2 \end{aligned}$$

Durch Wurzelziehen erhalten wir die Behauptung.

Eine Matrix mit $|A \cdot \vec{v}| = |\vec{v}|$ für jeden Vektor $\vec{v} \in \mathbb{R}^n$ nennen wir auch eine **Isometrie**. Jede orthogonale Matrix ist also eine Isometrie.

Beispiel 5.2.10. Ist A eine orthogonale 2×2 -Matrix, so gibt es ein $\alpha \in [0, 2\pi[$ mit

$$A = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix} \quad \text{oder} \quad A = \begin{pmatrix} \cos(\alpha) & \sin(\alpha) \\ \sin(\alpha) & -\cos(\alpha) \end{pmatrix}$$

Dazu schreiben wir zunächst ganz allgemein

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

mit reellen Zahlen a, b, c und d . Da A orthogonal ist, muss gelten $A^T \cdot A = E_2$, also

$$\begin{pmatrix} a & c \\ b & d \end{pmatrix} \cdot \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Multiplizieren wir das aus, so ergeben sich folgende drei Beziehungen

$$\begin{aligned} a^2 + c^2 &= 1 \\ ab + cd &= 0 \\ b^2 + d^2 &= 1 \end{aligned}$$

Die erste Beziehung besagt, dass $\vec{v} = \begin{pmatrix} a \\ c \end{pmatrix}$ ein Vektor der Länge a ist. Schreiben wir als \vec{v} in Polarkoordinaten, so bedeutet das, dass es genau ein

$\alpha \in [0, 2\pi)$ gibt mit

$$\vec{v} = \begin{pmatrix} \cos(\alpha) \\ \sin(\alpha) \end{pmatrix}$$

also $a = \cos(\alpha)$ und $c = \sin(\alpha)$. Genauso finden wir aus der dritten Beziehung ein β mit $b = \sin(\beta)$ und $d = \cos(\beta)$.

Setzen wir diese Resultate in die zweite Beziehung ein und nutzen auch noch die Additionstheoreme für Sinus und Kosinus aus, so erhalten wir

$$\begin{aligned} 0 &= \cos(\alpha) \sin(\beta) + \sin(\alpha) \cos(\beta) \\ &= \sin(\alpha + \beta) \end{aligned}$$

Da $\alpha, \beta \in [0, 2\pi[$ kann das nur dann der Fall sein, wenn

$$\alpha + \beta \in \{0, \pi, 2\pi, 3\pi\}$$

Falls nun $\alpha + \beta = 0$ oder $\alpha + \beta = 2\pi$, so gilt

$$\sin(\beta) = -\sin(\alpha), \quad \cos(\beta) = \cos(\alpha)$$

und wir erhalten also

$$A = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix}$$

Ist dagegen $\alpha + \beta = \pi$ oder $\alpha + \beta = 3\pi$, so gilt

$$\sin(\beta) = \sin(\alpha), \quad \cos(\beta) = -\cos(\alpha)$$

und wir erhalten also

$$A = \begin{pmatrix} \cos(\alpha) & \sin(\alpha) \\ \sin(\alpha) & -\cos(\alpha) \end{pmatrix}$$

Die orthogonale Matrix

$$A = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix}$$

beschreibt eine **Drehung** um den Winkel α , und wir schreiben hierfür auch D_α . Ist etwa $\alpha = \frac{\pi}{3}$ (also $\alpha = 60^\circ$ im Winkelmaß), so ist

$$D_{60^\circ} = \begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix}$$

Das ergibt folgendes Bild

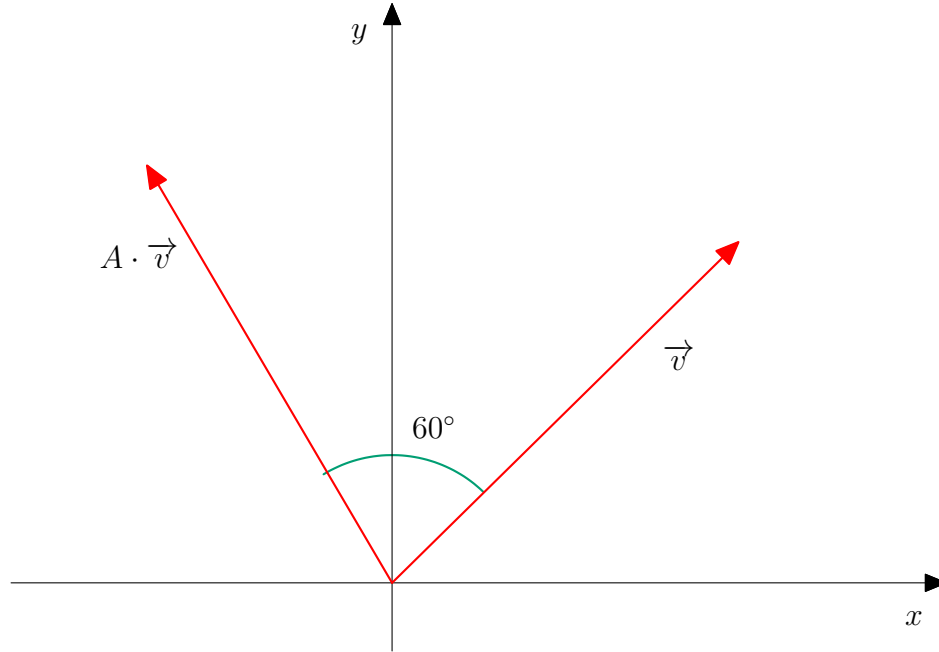


Abbildung 5.2: Drehung um 60°

Die orthogonale Matrix

$$A = \begin{pmatrix} \cos(\alpha) & \sin(\alpha) \\ \sin(\alpha) & -\cos(\alpha) \end{pmatrix}$$

dagegen beschreibt eine **Spiegelung** an der Geraden mit Winkel $\frac{\alpha}{2}$ zur x -Achse, und wir schreiben hierfür $S_{\frac{\alpha}{2}}$. Ist wieder $\alpha = \frac{\pi}{3}$ (also $\alpha = 60^\circ$ im Winkelmaß), so ist

$$S_{30^\circ} = \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}$$

Das ergibt dann Bild in Abbildung 5.3 (mit Spiegelungsachse S)

Bemerkung 5.2.9. Bei der geometrischen Betrachtung komplexer Zahlen spielen Drehungen (zusammen mit Streckungen) eine wichtige Rolle. Ist eine komplexe Zahl z in Polarkoordinatendarstellung durch die Länge r und den Winkel α gegeben, so ist die Multiplikation mit z eine Drehstreckung mit

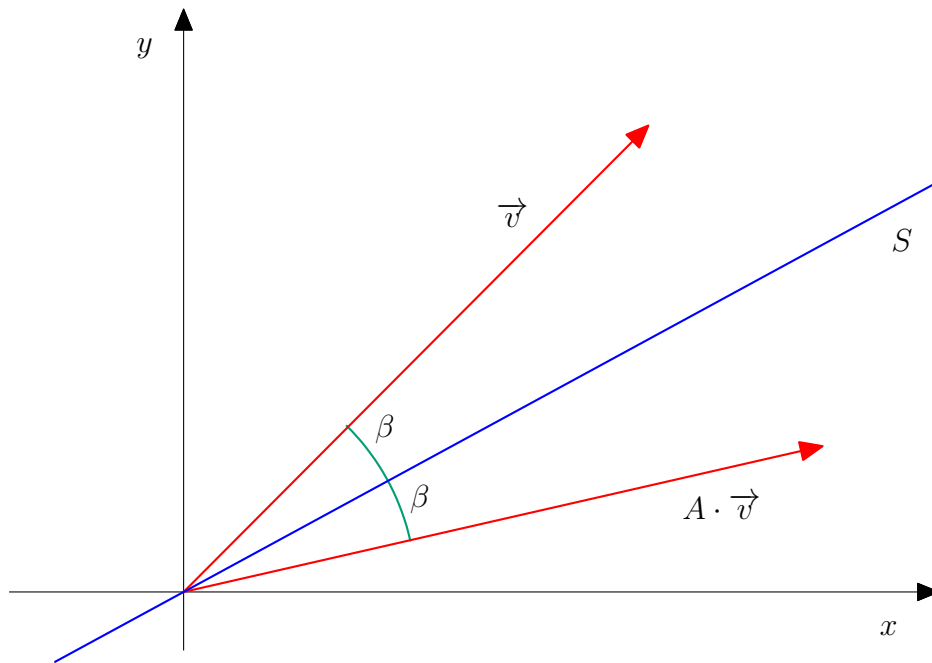


Abbildung 5.3: Spiegelung an einer Achse

Winkel α und Streckfaktor r , schreibt sich also in Matrizenschreibweise als Multiplikation mit der Matrix

$$\mu_z = r \cdot D_\alpha = r \cdot \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix}$$

Bemerkung 5.2.10. Für jeden Winkel α gilt:

$$\det(D_\alpha) = 1, \quad \det(S_\alpha) = -1$$

Im letzten Abschnitt haben wir für diesen Paragraphen eine Beweisskizze des Produktsatzes für Determinanten versprochen:

Satz 5.2.10 (Produktsatz). Sind A und B zwei $n \times n$ -Matrizen, so gilt

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

Beweisskizze :

1. Fall: Wir betrachten zunächst den Fall, dass $\text{rg}(A) < n$ oder $\text{rg}(B) < n$. Dann ist auch $\text{rg}(A \cdot B) < n$.

Ist etwa $\text{rg}(B) < n$, so ist nach dem Rangsatz $\text{nul}(B) > 0$, also gibt es ein $\vec{x} \in \mathbb{R}^n$ mit $\vec{x} \neq \vec{0}$ und $B \cdot \vec{x} = \vec{0}$. Dann gilt aber auch

$$A \cdot B \cdot \vec{x} = A \cdot \vec{0} = \vec{0},$$

also $\text{nul}(A \cdot B) > 0$ und damit $\text{rg}(A \cdot B) < n$. Ist $\text{rg}(B) = n$ und $\text{rg}(A) < n$ so gibt es zunächst ein \vec{y} mit $\vec{y} \neq \vec{0}$ und $A \cdot \vec{y} = \vec{0}$. Da $\text{rg}(B) = n$, gibt es ein \vec{x} mit $B \cdot \vec{x} = \vec{y}$ und hierfür gilt

$$A \cdot B \cdot \vec{x} = A \cdot \vec{y} = \vec{0},$$

also auch in dieser Situation $\text{nul}(A \cdot B) > 0$ und damit $\text{rg}(A \cdot B) < n$.

Damit gilt in diesem Fall also wegen Satz 5.1.27:

$$\det(A \cdot B) = 0 = \det(A) \cdot \det(B)$$

2. Fall: Es gilt $\text{rg}(A) = n$ und $\text{rg}(B) = n$. Damit brauchen wir also nur noch invertierbare Matrizen betrachten. Nach Bemerkung 5.2.5 schreibt sich dann A als Produkt von Elementarmatrizen, $A = \mathcal{E}_1 \cdots \mathcal{E}_l$, und es reicht zu zeigen

$$\det(\mathcal{E} \cdot B) = \det(\mathcal{E}) \cdot \det(B)$$

für jede (invertierbare) Matrix B und jede Elementarmatrix \mathcal{E} . Das rechnet man leicht mit den Regeln 5.1.23 nach.

Folgerung 5.2.11. *Sind A und B zwei $n \times n$ -Matrizen, und ist $A \cdot B$ invertierbar, so sind auch A und B invertierbar. Entsprechend haben in diesem Fall auch A und B den vollen Rang n .*

Beweis: Die Invertierbarkeit von $A \cdot B$ ist äquivalent zu $\det(A \cdot B) \neq 0$. Nach dem Produktsatz folgt aber daraus $\det(A) \neq 0$ und $\det(B) \neq 0$.

Zum Abschluß dieses Abschnitts und dieses Kapitels wollen wir nochmal zum Ausgangspunkt dieses Kapitels zurückkehren, nämlich zu linearen Gleichungssystemen

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

mit einer $n \times n$ -Koeffizientenmatrix A . Wir schreiben $A = (\vec{a}_1 \ \vec{a}_2 \ \dots \ \vec{a}_n)$ mit Spaltenvektoren \vec{a}_i und setzen

$$A_i(\vec{b}) = \left(\vec{a}_1 \ \dots \ \vec{a}_{i-1} \ \vec{b} \ \vec{a}_{i+1} \ \dots \ \vec{a}_n \right),$$

wir ersetzen also die i -te Spalte von A durch \vec{b} .

Satz 5.2.12 (Cramersche Regel). *Ist A invertierbar, so ist die eindeutige Lösung \vec{x} von*

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

gegeben durch

$$x_i = \frac{\det(A_i(\vec{b}))}{\det(A)}.$$

Beweis: Da A invertierbar ist, ist die Lösung des Gleichungssystems gegeben durch

$$\vec{x} = A^{-1} \vec{b}$$

Drücken wir A^{-1} gemäß Satz 5.2.6 mit der Komplementärmatrix aus, so bedeutet das

$$\begin{aligned} x_i &= \left(A^{-1} \cdot \vec{b} \right)_i \\ &= \sum_{l=1}^n \frac{\widetilde{a_{i,l}}}{\det(A)} \cdot b_l \\ &= \frac{1}{\det(A)} \cdot \sum_{l=1}^n (-1)^{i+l} \cdot \det(A_{l,i}) \cdot b_l \end{aligned}$$

Berechnen wir andererseits $\det(A_i(\vec{b}))$ durch Entwicklung nach der i -ten Spalte, so erhalten wir

$$\det(A_i(\vec{b})) = \sum_{l=1}^n (-1)^{i+l} \cdot b_l \cdot \det(A_{l,i})$$

da $A_i(\vec{b})_{l,i} = A_{l,i}$, und es folgt die Behauptung.

Beispiel 5.2.11. Wir betrachten das lineare Gleichungssystem

$$\begin{array}{rcrcrcrcrcrcl} x_1 & + & x_2 & & & & & & = & 2 \\ 3x_1 & + & 2x_2 & + & x_3 & = & 0 \\ & & x_2 & + & x_3 & = & 2 \end{array}$$

haben also

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 3 & 2 & 1 \\ 0 & 1 & 1 \end{pmatrix}, \quad \vec{b} = \begin{pmatrix} 2 \\ 0 \\ 2 \end{pmatrix}$$

Da $\det(A) = -2$ kann die Cramersche Regel angewendet werden und wir erhalten

$$\begin{aligned} x_1 &= \frac{-1}{2} \cdot \det \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 2 & 1 & 1 \end{pmatrix} = -2 & x_2 &= \frac{-1}{2} \cdot \det \begin{pmatrix} 1 & 2 & 0 \\ 3 & 0 & 1 \\ 2 & 2 & 1 \end{pmatrix} = 4 \\ x_3 &= \frac{-1}{2} \cdot \det \begin{pmatrix} 1 & 1 & 2 \\ 3 & 2 & 0 \\ 0 & 1 & 2 \end{pmatrix} = -2 \end{aligned}$$

Bemerkung 5.2.11. Die Cramersche Regel ist keine praktikable Methode, um die Lösung eines Gleichungssystems zu bestimmen (ausser vielleicht im Fall $n = 2$). Der Rechenaufwand, der damit verbunden ist, ist viel höher als der, der sich aus dem Gaußschen Eliminationsverfahren ergibt.

Bemerkung 5.2.12. Determinanten können mit den gleichen Berechnungsregeln auch für komplexe Matrizen definiert werden. Betrachten wir etwa die Matrix

$$A = \begin{pmatrix} 1 & i \\ i & i - 1 \end{pmatrix}$$

so gilt hierfür

$$\det(A) = 1 \cdot (i - 1) - i \cdot i = i - 1 - (-1) = i$$

Die Aussagen über Determinanten übertragen sich entsprechend. Insbesondere ist eine komplexe Matrix invertierbar, wenn ihre Determinante von Null

verschieden ist. Das gilt also für unsere Matrix A , und ihre Inverse berechnet sich wie im reellen Fall mit der erweiterten Matrix:

$$(A | E_2) = \left(\begin{array}{cc|cc} 1 & i & 1 & 0 \\ i & i-1 & 0 & 1 \end{array} \right)$$

Subtrahieren wir das i -fache der ersten Zeile von der zweiten, so erhalten wir

$$\left(\begin{array}{cc|cc} 1 & i & 1 & 0 \\ 0 & i & -i & 1 \end{array} \right)$$

Division der zweiten Zeile durch i ergibt

$$\left(\begin{array}{cc|cc} 1 & i & 1 & 0 \\ 0 & 1 & -1 & -i \end{array} \right)$$

Subtrahieren wir das i -fache der zweiten Zeile von der ersten, so ergibt sich

$$(E_2 | A^{-1}) = \left(\begin{array}{cc|cc} 1 & 0 & 1+i & -1 \\ 0 & 1 & -1 & -i \end{array} \right)$$

und damit

$$A^{-1} = \begin{pmatrix} 1+i & -1 \\ -1 & -i \end{pmatrix}$$

Bemerkung 5.2.13. Die Theorie der Determinanten und invertierbaren Matrizen kann genauso auf jeden anderen Körper, etwa \mathbb{Q} oder \mathbb{F}_p für eine Primzahl $p > 0$ übertragen werden. Alle Regeln und Berechnungsformeln gelten entsprechend.

Aufgabe 5.2.1. Bestimmen Sie die Inverse der Matrix

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

Aufgabe 5.2.2. Bestimmen Sie die Inverse der Matrix

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

Aufgabe 5.2.3. Überprüfen Sie, ob die Matrix

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

invertierbar ist und bestimmen Sie gegebenenfalls die inverse Matrix.

Aufgabe 5.2.4. Zeigen Sie, dass eine $n \times n$ -Matrix A genau dann invertierbar ist, wenn ihre Spaltenvektoren eine Basis des \mathbb{R}^n bilden.

Aufgabe 5.2.5. Zeigen Sie: Ist A eine invertierbare Matrix, so gilt

$$\det(A^{-1}) = \frac{1}{\det(A)}$$

Aufgabe 5.2.6. Zeigen Sie: Ist A eine orthogonale Matrix, so gilt

$$\det(A) = \pm 1$$

Kapitel 6

Eigenwerte und Normalformen

Matrizen können sehr unterschiedliche Gestalt annehmen, und ihre Gestalt ist von entscheidender Bedeutung dafür, wie leicht oder schwer es ist, mit diesen Matrizen zu arbeiten und zu rechnen. So ist es in der Regel etwa umso einfacher mit einer Matrix zu arbeiten, je weniger von Null verschiedene Einträge sie hat. Besonders leicht zu behandeln sind dabei Matrizen, die nur entlang der Diagonale von Null verschiedene Einträge haben. So ist etwa das Gleichungssystem

$$\begin{array}{rcl} 2x_1 & & = 1 \\ & 3x_2 & = 2 \\ & & -x_3 = 4 \end{array}$$

einfacher zu lösen als das Gleichungssystem

$$\begin{array}{rclcl} 2x_1 & + & 4x_2 & + & 3x_3 & = & 1 \\ x_1 & - & 3x_2 & - & 2x_3 & = & 2 \\ 2x_1 & - & 2x_3 & + & x_3 & = & 4 \end{array}$$

und auch bei der Determinantenberechnung ist

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

einfacher zu handhaben als

$$B = \begin{pmatrix} 2 & 4 & 3 \\ 1 & -3 & -2 \\ 2 & -2 & 1 \end{pmatrix}$$

Nun kommen natürlich in vielen Fragestellungen Matrizen mit vielen Einträgen vor. In diesem Fall können wir uns jedoch fragen, ob es möglich ist, diese Matrizen durch eine geschickte Transformation zu vereinfachen und auf Diagonalgestalt (oder eine Form, die zumindest nahe an der Diagonalgestalt ist) zu bringen. Dabei stellt sich natürlich als erstes die Frage, was wir unter "geschickter Transformation" verstehen wollen. Der Schlüssel hierzu liegt in der Interpretation von Matrizen als lineare Abbildung. Wir haben diese Matrix mit Hilfe der Standardbasis des \mathbb{R}^n gebildet. In diesem Abschnitt wollen wir versuchen, durch geschickte Wahl einer Basis eine einfachere Form der beschreibenden Matrix einer linearen Abbildung zu bekommen.

6.1 Eigenwerte und Eigenvektoren

Besonders einfache Matrizen sind die Diagonalmatrizen D . Diese zeichnen sich dadurch aus, dass für sie gilt:

$$D \cdot \vec{e}_i = d_i \cdot \vec{e}_i \text{ für alle } i$$

wobei $\vec{e}_1, \dots, \vec{e}_n$ die Standardbasis des \mathbb{R}^n bezeichnet und d_i das i -te Diagonalelement von D . Bei einer stark besetzten Matrix ist das nicht mehr der Fall, wir können uns aber auch in dieser Situation fragen, ob es Vektoren gibt, die eine analoge Eigenschaft haben. Dazu betrachten wir eine $n \times n$ -Matrix A und definieren

Definition 6.1.1. Ein $\lambda \in \mathbb{R}$ heißt **Eigenwert** von A , wenn es einen Vektor $\vec{v} \in \mathbb{R}^n$, $\vec{v} \neq \vec{0}$ gibt mit

$$A \cdot \vec{v} = \lambda \cdot \vec{v}$$

In diesem Fall heißt der Vektor \vec{v} **Eigenvektor** von A zum Eigenwert λ .

Beispiel 6.1.1. Ist

$$D = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$$

so sind 2 und 3 Eigenwerte von D . $\vec{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ ist ein Eigenvektor von D zum Eigenwert 2 und $\vec{e}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ ist ein Eigenvektor von D zum Eigenwert 3.

Beispiel 6.1.2. Es sei

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

Dann ist 0 ein Eigenwert von A und $\vec{v} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ ist ein Eigenvektor von A zum Eigenwert 0. Ferner ist 3 ein Eigenwert von A und $\vec{w} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ ist ein Eigenvektor von A zum Eigenwert 2.

Bemerkung 6.1.1. Der Nullvektor $\vec{0}$ ist als Eigenvektor nicht zugelassen (schon deswegen nicht, weil ja für jede reelle Zahl r gelten würde: $A \cdot \vec{0} = r \cdot \vec{0}$). Der Skalar 0 ist aber sehr wohl als Eigenwert möglich, wie wir in Beispiel 6.1.2 gesehen haben.

An Beispiel 6.1.2 können wir auch eine andere wichtige Tatsache ablesen: Der Wert 0 kommt genau dann als Eigenwert von A vor, wenn das homogene Gleichungssystem

$$A \cdot \vec{x} = \vec{0}$$

eine nichttriviale Lösung hat. Ist \vec{x} eine solche nichttriviale Lösung, so ist \vec{x} ein Eigenvektor von A zum Eigenwert 0.

Ist jetzt allgemeiner λ ein Eigenwert von A und \vec{v} ein zugehöriger Eigenvektor, so gilt hierfür

$$A \cdot \vec{v} = \lambda \cdot \vec{v}$$

Das lässt sich umschreiben zu

$$A \cdot \vec{v} = \lambda \cdot E_n \cdot \vec{v}$$

wobei E_n die $n \times n$ -Einheitsmatrix ist, und damit zu

$$\lambda \cdot E_n \cdot \vec{v} - A \cdot \vec{v} = \vec{0}$$

oder äquivalent zu

$$(\lambda \cdot E_n - A) \cdot \vec{v} = \vec{0}$$

Damit ist also $\lambda \in \mathbb{R}$ genau dann ein Eigenwert von A , wenn das homogene Gleichungssystem

$$(\lambda \cdot E_n - A) \cdot \vec{v} = \vec{0}$$

eine nichttriviale Lösung hat, und jede nicht-triviale Lösung \vec{v} dieses Gleichungssystems ist ein Eigenvektor von A zum Eigenwert λ . Damit haben wir gezeigt

Satz 6.1.1. *Genau dann ist ein Skalar $\lambda \in \mathbb{R}$ ein Eigenwert von A , wenn $\det(\lambda \cdot E_n - A) = 0$.*

In diesem Fall ist jeder nicht-triviale Vektor $\vec{v} \in \text{Ker}(\lambda \cdot E_n - A)$ ein Eigenvektor zum Eigenwert λ .

Dieser Satz motiviert die allgemeine Betrachtung von $P_A(\lambda) = \det(\lambda \cdot E_n - A)$ als Ausdruck in λ .

Satz 6.1.2. *Ist A eine $n \times n$ -Matrix, so ist $P_A(\lambda) = \det(\lambda \cdot E_n - A)$ ein normiertes Polynom vom Grad n in λ .*

Beweis: Wir benutzen dazu die vollständige Entwicklung von Determinanten, also die Formel

$$\det(A) = \sum_{\sigma \in \S_n} \text{sign}(\sigma) \cdot a_{1,\sigma(1)} \cdots a_{n,\sigma(n)}$$

wenden wir das auf die Matrix $\lambda \cdot E_n - A$ an, so sehen wir

$$\det(\lambda \cdot E_n - A) = (\lambda - a_{1,1}) \cdot (\lambda - a_{2,2}) \cdots (\lambda - a_{n,n}) + \text{weitere Terme}$$

wobei in jedem der weiteren Terme λ höchstens $(n-1)$ -mal auftaucht. Da

$$(\lambda - a_{1,1}) \cdot (\lambda - a_{2,2}) \cdots (\lambda - a_{n,n}) = \lambda^n + \text{weitere Terme}$$

wobei auch hier in jedem der weiteren Terme λ höchstens $(n-1)$ -mal auftaucht, folgt daraus die Behauptung.

Definition 6.1.2. Der Ausdruck $P_A(\lambda) = \det(\lambda \cdot E_n - A)$ heißt das **charakteristische Polynom** der Matrix A .

Satz 6.1.3. *Ist A eine $n \times n$ -Matrix, so hat A höchstens n verschiedene Eigenwerte.*

Beweis: Alle Eigenwerte von A sind Nullstellen des charakteristischen Polynoms $P_A(\lambda)$. Da ein Polynom vom Grad n höchstens n verschiedene Nullstellen hat, folgt die Behauptung.

Beispiel 6.1.3. Wir wollen nochmals die Matrix

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

betrachten. Hierfür gilt

$$\lambda \cdot E_2 - A = \begin{pmatrix} \lambda - 1 & -1 \\ -2 & \lambda - 2 \end{pmatrix}$$

und damit

$$P_A(\lambda) = (\lambda - 1)(\lambda - 2) - (-1) \cdot (-2) = \lambda^2 - 3\lambda$$

Wir sehen also, dass $P_A(\lambda)$ die beiden Nullstellen $\lambda_1 = 0$ und $\lambda_2 = 3$ hat, und dabei handelt es sich um die beiden (einzigen) Eigenwerte von A .

Beispiel 6.1.4. Wir betrachten die Matrix

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

Dann gilt hierfür

$$\begin{aligned} P_A(\lambda) &= \det \begin{pmatrix} \lambda & -1 & 1 \\ -3 & \lambda - 2 & 3 \\ -2 & -2 & \lambda + 3 \end{pmatrix} \\ &= \lambda \cdot \det \begin{pmatrix} \lambda - 2 & 3 \\ -2 & \lambda + 3 \end{pmatrix} - (-3) \cdot \det \begin{pmatrix} -1 & 1 \\ -2 & \lambda + 3 \end{pmatrix} \\ &\quad + (-2) \cdot \det \begin{pmatrix} -1 & 1 \\ \lambda - 2 & 3 \end{pmatrix} \\ &= \lambda^3 + \lambda^2 - \lambda - 1 \\ &= (\lambda + 1)^2(\lambda - 1) \end{aligned}$$

Damit hat A die Eigenwerte $\lambda_1 = 1$ und $\lambda_2 = -1$. Die Matrix A hat also nur zwei Eigenwerte, einen weniger als wir nach Satz 6.1.3 maximal bekommen könnten.

Die Eigenvektoren zum Eigenwert -1 sind die nicht-trivialen Lösungen des Gleichungssystems

$$\begin{array}{rrcr} -x_1 & - & x_2 & + & x_3 & = & 0 \\ -3x_1 & - & 3x_2 & + & 3x_3 & = & 0 \\ -2x_1 & - & 2x_2 & + & 2x_3 & = & 0 \end{array}$$

Eine Basis des zugehörigen Lösungsraums sind die beiden Vektoren

$$\vec{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad \vec{v}_2 = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$$

und damit sind die Vektoren

$$\vec{v} = r \cdot \vec{v}_1 + s \cdot \vec{v}_2$$

mit $r \neq 0$ oder $s \neq 0$ die Eigenvektoren zum Eigenwert -1 .

Die Eigenvektoren zum Eigenwert 1 sind die nicht-trivialen Lösungen des Gleichungssystems

$$\begin{array}{rrcr} x_1 & - & x_2 & + & x_3 & = & 0 \\ -3x_1 & - & x_2 & + & 3x_3 & = & 0 \\ -2x_1 & - & 2x_2 & + & 4x_3 & = & 0 \end{array}$$

Eine Basis des zugehörigen Lösungsraums ist der Vektor

$$\vec{v}_3 = \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix}$$

und damit sind die Vektoren

$$\vec{v} = r \cdot \vec{v}_3$$

mit $r \neq 0$ die Eigenvektoren zum Eigenwert 1 .

Bemerkung 6.1.2. Die Theorie der Eigenwerte und Eigenvektoren lässt sich natürlich auch auf komplexe Matrizen ausdehnen. So hat etwa die Matrix

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

das charakteristische Polynom

$$P_A(\lambda) = \det \begin{pmatrix} \lambda - 1 & -i \\ -i & \lambda - 1 \end{pmatrix} = \lambda^2 - 2\lambda + 2$$

und damit hat A die Eigenwerte $\lambda_1 = 1 + i$ und $\lambda_2 = 1 - i$. Die Eigenvektoren zu $\lambda_1 = 1 + i$ sind wie im reellen Fall die nicht-trivialen (komplexen) Lösungen des Gleichungssystems

$$(\lambda_1 \cdot E_2 - A) \vec{x} = \vec{0}$$

also die Lösungen von

$$\begin{aligned} i \cdot x_1 - i \cdot x_2 &= 0 \\ -i \cdot x_1 + i \cdot x_2 &= 0 \end{aligned}$$

Eine Basis des Lösungsraums dieses Gleichungssystems ist

$$\vec{v}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

und damit sind die Vektoren

$$\vec{v} = z \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

mit $z \in \mathbb{C} \setminus \{0\}$ die Eigenvektoren zum Eigenwert $1 + i$. Entsprechend erhalten wir für $\lambda_2 = 1 - i$ die Eigenvektoren

$$\vec{v} = z \cdot \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

mit $z \in \mathbb{C} \setminus \{0\}$.

Der Übergang zu komplexen Zahlen ist allerdings nicht nur notwendig, wenn wir mit einer komplexen Matrix starten. Betrachten wir etwa die Matrix

$$B = \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix}$$

so hat diese charakteristisches Polynom

$$P_B(\lambda) = \lambda^2 - 2\lambda + 5$$

Diese Polynom hat keine reellen Nullstellen, aber die beiden komplexen Lösungen $\lambda_1 = 1 + 2 \cdot i$ und $\lambda_2 = 1 - 2 \cdot i$. Betrachten wir die Matrix B also als komplexe Matrix (*Koeffizientenerweiterung*), so hat B Eigenwerte und die Eigenvektoren zu $\lambda_1 = 1 + 2 \cdot i$ erhalten wir aus den nicht-trivialen Lösungen des Gleichungssystems

$$\begin{aligned} 2 \cdot i \cdot x_1 - 2 \cdot x_2 &= 0 \\ 2 \cdot x_1 + 2 \cdot i \cdot x_2 &= 0 \end{aligned}$$

Damit sind die Vektoren

$$\vec{v} = z \cdot \begin{pmatrix} -i \\ 1 \end{pmatrix}$$

mit $z \in \mathbb{C} \setminus \{0\}$ die Eigenvektoren zum Eigenwert $1 + 2 \cdot i$. Entsprechend erhalten wir für $\lambda_2 = 1 - 2 \cdot i$ die Eigenvektoren

$$\vec{v} = z \cdot \begin{pmatrix} i \\ 1 \end{pmatrix}$$

mit $z \in \mathbb{C} \setminus \{0\}$. Die Matrix ist also in diesem Fall reell, die Eigenwerte und Eigenvektoren sind aber komplex.

Aufgabe 6.1.1. Bestimmen Sie Eigenwerte und Eigenvektoren der Matrix

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

Aufgabe 6.1.2. Bestimmen Sie Eigenwerte und Eigenvektoren der Matrix

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

Aufgabe 6.1.3. Bestimmen Sie Eigenwerte und Eigenvektoren der Matrix

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

Aufgabe 6.1.4. Bestimmen Sie Eigenwerte und Eigenvektoren der Matrix

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

Aufgabe 6.1.5. Bestimmen Sie das charakteristische Polynom der Matrix

$$A = \begin{pmatrix} \cos(\alpha) & \sin(\alpha) \\ -\sin(\alpha) & \cos(\alpha) \end{pmatrix}$$

Für welche Werte von α hat diese Matrix (reelle) Eigenwerte?

Aufgabe 6.1.6. Bestimmen Sie das charakteristische Polynom der Matrix

$$A = \begin{pmatrix} \cos(\alpha) & \sin(\alpha) \\ \sin(\alpha) & -\cos(\alpha) \end{pmatrix}$$

Für welche Werte von α hat diese Matrix (reelle) Eigenwerte?

Aufgabe 6.1.7. Bestimmen Sie Eigenwerte und Eigenvektoren der Matrix

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

Aufgabe 6.1.8. Bestimmen Sie Eigenwerte und Eigenvektoren der Matrix

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

Aufgabe 6.1.9. Wir betrachten die Matrizen

$$A_{a,b} = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

mit reellen Zahlen a, b . Bestimmen Sie die (komplexen) Eigenwerte und Eigenvektoren von $A_{a,b}$ (in Abhängigkeit von a, b).

Aufgabe 6.1.10. Wir betrachten eine $n \times n$ -Matrix A , (paarweise verschiedene) Eigenwerte $\lambda_1, \dots, \lambda_m$ von A , und zu jedem λ_i einen Eigenvektor \vec{v}_i . Zeigen Sie, dass die Vektoren $\vec{v}_1, \dots, \vec{v}_m$ linear unabhängig sind.

6.2 Diagonalisierbarkeit

Wie wir jetzt gesehen haben, sind Eigenvektoren zu einer Matrix A Vektoren, bezüglich derer sich die Multiplikation von A mit dem Vektor besonders einfach darstellen lässt, nämlich als

$$A \cdot \vec{v} = \lambda \cdot \vec{v}$$

Wie kann uns das aber jetzt bei der Vereinfachung der Matrizen helfen?

Beispiel 6.2.1. Wir betrachten die Matrix

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

Wie wir sofort sehen, ist 0 ein Eigenwert von A und $\vec{v} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ ist ein Eigenvektor von A zum Eigenwert 0. Ferner ist 2 ein Eigenwert von A und $\vec{w} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ ist ein zugehöriger Eigenvektor von A zum Eigenwert 2. Wir haben also

$$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix} = 0 \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} = 2 \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

oder in Matrixzenschreibweise

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

Nun ist die Matrix

$$S = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$$

invertierbar mit Inverser

$$S^{-1} = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

Multiplizieren wir also beide Seiten dieser Matrixgleichung mit S^{-1} so erhalten wir daraus

$$\begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}$$

also ist $S^{-1} \cdot A \cdot S$ eine Diagonalmatrix.

Beispiel 6.2.2. Wir betrachten wieder die Matrix

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

aus Beispiel 6.1.2. Wie wir dort gesehen haben, ist 0 ein Eigenwert von A und $\vec{v} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ ist ein Eigenvektor von A zum Eigenwert 0. Ferner ist 3 ein Eigenwert von A und $\vec{w} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ ist ein zugehöriger Eigenvektor von A .

Auch hier ist die Matrix

$$S = \begin{pmatrix} 1 & -1 \\ 2 & 1 \end{pmatrix}$$

invertierbar mit Inverser

$$S^{-1} = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} \\ -\frac{2}{3} & \frac{1}{3} \end{pmatrix}$$

und wir erhalten mit einer analogen Rechnung

$$S^{-1} \cdot A \cdot S = \begin{pmatrix} 3 & 0 \\ 0 & 0 \end{pmatrix}$$

Wir haben also in diesen Beispielen, ausgehend von den Eigenwerten von A und von zugehörigen Eigenvektoren eine Transformation gefunden, die aus A eine Diagonalmatrix macht. Wir wollen uns nun damit beschäftigen, wie weit sich dieses Beispiel verallgemeinern lässt.

Dazu sei zunächst A eine $n \times n$ -Matrix, die die maximal mögliche Anzahl von n verschiedenen (reelle) Eigenwerte hat, für die das charakteristische Polynom also in n paarweise verschiedene Linearfaktoren zerfällt. Diese Eigenwerte bezeichnen wir mit $\lambda_1, \dots, \lambda_n$, und zu jedem λ_i wählen wir einen Eigenvektor \vec{v}_i . Dann können wir die Beziehungen

$$A \cdot \vec{v}_i = \lambda_i \cdot \vec{v}_i$$

wieder als Matrizengleichung schreiben. Dazu betrachten wir die Matrix

$$S = (\vec{v}_1 \ \vec{v}_2 \ \dots \ \vec{v}_n)$$

die die Vektoren \vec{v}_i als i -te Spaltenvektoren hat. Dann erhalten wir genauso wie im Beispiel

$$A \cdot S = S \cdot D$$

wobei

$$D = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}$$

eine Diagonalmatrix ist. Nun sind nach Aufgabe 6.1.10 die Spalten von S linear unabhängig, und damit ist S invertierbar. Also erhalten wir aus dieser Gleichung durch Multiplikation mit S^{-1} wieder eine Beziehung

$$S^{-1} \cdot A \cdot S = D$$

also eine Transformation von A auf Diagonalgestalt. Wir haben also gezeigt

Satz 6.2.1. *Ist A eine $n \times n$ -Matrix mit n verschiedenen (reellen) Eigenwerten $\lambda_1, \dots, \lambda_n$, so gibt es eine invertierbare Matrix S mit*

$$S^{-1} \cdot A \cdot S = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}$$

Definition 6.2.1. Eine quadratische Matrix A heißt **diagonalisierbar**, wenn es eine invertierbare Matrix S gibt, so dass $S^{-1} \cdot A \cdot S$ eine Diagonalmatrix ist.

Wir haben gesehen, dass eine $n \times n$ -Matrix mit n verschiedenen Eigenwerten diagonalisierbar ist. Wie sieht es nun aus mit Matrizen, die weniger als n Eigenwerte haben? Betrachten wir dazu etwa die Matrix

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

aus Beispiel 6.1.4. Wie wir in diesem Beispiel gesehen haben, hat diese Matrix nur zwei Eigenwerte, -1 und 1 . Dabei haben wir zum Eigenwert -1 zwei linear unabhängige Eigenvektoren

$$\vec{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad \vec{v}_2 = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$$

und zum Eigenwert 1 den Eigenvektor

$$\vec{v}_3 = \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix}$$

Wir überzeugen uns, dass die Vektoren $\vec{v}_1, \vec{v}_2, \vec{v}_3$ linear unabhängig sind und bilden mit diesen die Matrix

$$S = (\vec{v}_1 \ \vec{v}_2 \ \vec{v}_3) = \begin{pmatrix} 1 & -1 & 1 \\ 0 & 1 & 3 \\ 1 & 0 & 2 \end{pmatrix}$$

Hierfür rechnen wir unmittelbar nach, dass

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

und, da S invertierbar ist, also auch hier

$$S^{-1} \cdot A \cdot S = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Für die Diagonalisierbarkeit von A ist es also nicht notwendig, dass A genau n verschiedene Eigenwerte hat, es reicht, dass A genügend viele linear unabhängige Eigenvektoren hat. Genauer gilt

Satz 6.2.2. *Genau dann ist eine $n \times n$ -Matrix A diagonalisierbar, wenn es n linear unabhängige Eigenvektoren von A gibt.*

Beweis: Sind $\vec{v}_1, \dots, \vec{v}_n$ linear unabhängige Eigenvektoren von A , so setzen wir

$$S = (\vec{v}_1 \ \dots \ \vec{v}_n)$$

bilden also die Matrix mit diesen Eigenvektoren als Spalten und rechnen dann wie in obigen Vorüberlegungen oder im Beweis von Satz 6.2.1 nach, dass $S^{-1} \cdot A \cdot S$ eine Diagonalmatrix ist.

Gibt es umgekehrt eine invertierbare Matrix S so dass

$$S^{-1} \cdot A \cdot S = D$$

eine Diagonalmatrix ist, so erhalten wir daraus

$$A \cdot S = S \cdot D$$

Bezeichnen wir die Spalten von S mit $\vec{v}_1, \dots, \vec{v}_n$, so sind diese linear unabhängig, und es gilt

$$A \cdot \vec{v}_i = d_i \cdot \vec{v}_i$$

wobei d_i das i -te Diagonalelement von D ist.

Damit ist der Satz beweisen.

Es ist nun naheliegend, dass wir uns fragen, ob alle Matrizen diagonalisierbar sind. Das ist nicht der Fall, wie die folgenden Beispiele zeigen.

Beispiel 6.2.3. Wir betrachten die Matrix

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

Diese Matrix hat das charakteristische Polynom

$$P_A(\lambda) = \lambda^2 + 1$$

und damit keinen reellen Eigenwert, also auch keine reellen Eigenvektoren. Nach dem Satz 6.2.2 ist sie also auch nicht (über den reellen Zahlen) diagonalisierbar.

Allerdings ist das kein "echtes" Gegenbeispiel, denn das charakteristische Polynom hat zwei komplexe Nullstellen, nämlich $\lambda_1 = i$ und $\lambda_2 = -i$, und dazu gibt es in \mathbb{C}^2 auch Eigenvektoren: Für $\vec{v}_1 = \begin{pmatrix} 1 \\ -i \end{pmatrix}$ rechnen wir nach, dass

$$A \cdot \vec{v}_1 = i \cdot \vec{v}_1$$

also ist \vec{v}_1 ein komplexer Eigenvektor zu $\lambda_1 = i$ und entsprechend ist $\vec{v}_2 = \begin{pmatrix} i \\ 1 \end{pmatrix}$ ein komplexer Eigenvektor zu $\lambda_2 = -i$. Die zugehörige komplexe Matrix

$$S := \begin{pmatrix} -i & i \\ 1 & 1 \end{pmatrix}$$

ist invertierbar mit inverser Matrix

$$S^{-1} = \begin{pmatrix} \frac{i}{2} & \frac{1}{2} \\ -\frac{i}{2} & \frac{1}{2} \end{pmatrix}$$

Damit gilt dann auch

$$S^{-1} \cdot A \cdot S = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$$

Die Matrix A ist also nicht über den reellen Zahlen diagonalisierbar, aber über den komplexen Zahlen.

Beispiel 6.2.4. Wir betrachten die Matrix

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

Diese Matrix hat das charakteristische Polynom

$$P_A(\lambda) = \lambda^2 - 4\lambda + 4$$

und damit nur eine Nullstelle $\lambda = 2$. Das zugehörige Gleichungssystem $(2E_2 - A) \cdot \vec{x} = \vec{0}$ schreibt sich explizit als

$$\begin{aligned} 0 \cdot x_1 + 1 \cdot x_2 &= 0 \\ 0 \cdot x_1 + 0 \cdot x_2 &= 0 \end{aligned}$$

und eine Basis des Lösungsraums ist gegeben durch

$$\vec{v} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

Damit sind die Eigenvektoren von A zum Eigenwert $\lambda = 2$ die nicht-trivialen Vielfachen von \vec{v} . Insbesondere gibt es keine zwei linear unabhängigen Eigenvektoren. Da es keinen weiteren Eigenwert gibt (auch nicht über den komplexen Zahlen), ist also A nicht diagonalisierbar.

In der Tat ist A bereits in der optimalen Form, die wir hierfür durch Transformationen erreichen können.

Die Theorie der Eigenwerte und Eigenvektoren hat viele Anwendungen, speziell bei den linearen Differentialgleichungssystemen mit konstanten Koeffizienten. Interessant ist sie auch für Probleme, bei denen hohe Potenzen A^k einer Matrix A berechnet werden müssen, etwa bei Markow-Ketten.

Beispiel 6.2.5. Ein Teilchen kann in zwei Energiezuständen A oder B auftreten. Innerhalb einer Zeiteinheit t geht ein Teilchen im Zustand A mit einer Wahrscheinlichkeit von 40% von Zustand A in Zustand B über (und verbleibt mit einer Wahrscheinlichkeit von 60% im Zustand A), und ein Teilchen im Zustand B geht einer Wahrscheinlichkeit von 30% von Zustand B in Zustand A über (und verbleibt mit einer Wahrscheinlichkeit von 70% im Zustand B). Wie ist - bei gegebener Anfangsverteilung und bei einer sehr großen Menge von Teilchen - die Zustandsverteilung nach k Perioden?

Wir schreiben den Anfangszustand der Teilchen als Vektor

$$\vec{v} = \vec{v}(0) = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix},$$

wobei v_1 die Anzahl der Teilchen im Zustand A und v_2 die Anzahl der Teilchen im Zustand B bezeichnet. Nach einer Periode verbleiben $0.6 \cdot v_1$ der v_1 Teilchen im Zustand A in diesem Zustand, wohingegen $0.4 \cdot v_1$ Teilchen in den Zustand B gewechselt haben. Von den v_2 Teilchen im Zustand B verbleiben $0.7 \cdot v_2$ Teilchen in diesem Zustand, $0.3 \cdot v_2$ Teilchen haben in den Zustand A gewechselt. Damit ist die Verteilung nach einer Periode

$$\vec{v}(1) = \begin{pmatrix} 0.6 \cdot v_1 + 0.3 \cdot v_2 \\ 0.4 \cdot v_1 + 0.7 \cdot v_2 \end{pmatrix} = \begin{pmatrix} 0.6 & 0.3 \\ 0.4 & 0.7 \end{pmatrix} \cdot \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 0.6 & 0.3 \\ 0.4 & 0.7 \end{pmatrix} \cdot \vec{v}(0)$$

Setzen wir

$$A = \begin{pmatrix} 0.6 & 0.3 \\ 0.4 & 0.7 \end{pmatrix}$$

so ist also der Übergang von Periode 0 zu Periode 1 gegeben durch

$$\vec{v}(1) = A \cdot \vec{v}(0)$$

und analog gilt für den Zustand nach k Perioden

$$\vec{v}(k) = A \cdot \vec{v}(k-1) = A^k \cdot \vec{v}(0)$$

wie wir uns induktiv überzeugen. Diese Matrizen A^k sind relativ kompliziert zu berechnen. Hätten wir statt A eine Diagonalmatrix,

$$D = \begin{pmatrix} d_1 & 0 \\ 0 & d_2 \end{pmatrix}$$

so wäre die Situation deutlich einfacher, denn

$$D^k = \begin{pmatrix} d_1^k & 0 \\ 0 & d_2^k \end{pmatrix}$$

wie man sehr leicht mit Induktion nachrechnet. Fast so einfach ist die Situation, wenn A diagonalisierbar ist. Dann ist nämlich $S^{-1} \cdot A \cdot S = D$ eine Diagonalmatrix, also

$$A = S \cdot D \cdot S^{-1}$$

mit einer Diagonalmatrix D . Dann gilt aber

$$\begin{aligned} A^2 &= (S \cdot D \cdot S^{-1}) \cdot (S \cdot D \cdot S^{-1}) \\ &= S \cdot D \cdot (S^{-1} \cdot S) \cdot D \cdot S^{-1} \\ &= S \cdot D \cdot E_2 \cdot D \cdot S^{-1} \\ &= S \cdot D^2 \cdot S^{-1} \end{aligned}$$

und allgemein

$$A^k = S \cdot D^k \cdot S^{-1}$$

Wir wollen daher unsere Matrix A auf Diagonalisierbarkeit untersuchen:

$$\begin{aligned} P_A(\lambda) &= \det \begin{pmatrix} \lambda - 0.6 & -0.3 \\ -0.4 & \lambda - 0.7 \end{pmatrix} \\ &= \lambda^2 - 1.3 \cdot \lambda + 0.3 \\ &= (\lambda - 1) \cdot (\lambda - 0.3) \end{aligned}$$

Also hat A zwei Eigenwerte, $\lambda_1 = 1$ und $\lambda_2 = 0.3$. Zugehörige Eigenvektoren sind etwa $\vec{v}_1 = \begin{pmatrix} 3 \\ 4 \end{pmatrix}$ zu λ_1 und $\vec{v}_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$ zu λ_2 . Die zugehörige Matrix

$$S = \begin{pmatrix} 3 & -1 \\ 4 & 1 \end{pmatrix}$$

hat Inverse

$$S^{-1} = \frac{1}{7} \cdot \begin{pmatrix} 1 & 1 \\ -4 & 3 \end{pmatrix}$$

und mit

$$D = \begin{pmatrix} 1 & 0 \\ 0 & 0.3 \end{pmatrix}$$

gilt also $A = S \cdot D \cdot S^{-1}$ und damit

$$\vec{v}(k) = S \cdot \begin{pmatrix} 1 & 0 \\ 0 & 0.3^k \end{pmatrix} \cdot S^{-1} \cdot \vec{v}(0)$$

Interessanterweise gilt auch

$$A = \begin{pmatrix} 3 \\ 4 \end{pmatrix} = \begin{pmatrix} 3 \\ 4 \end{pmatrix}$$

dh. werden eine große Anzahl von Teilchen im Verhältnis 3 : 4 auf die beiden Energiezustände aufgeteilt, so bleibt das System im Gleichgewicht und die Verteilung der Teilchen auf diese beiden Zustände ist stabil. Im allgemeinen ist es so, dass die Verhältnisse der Teilchen sich im Laufe der Zeit einem Verhältnis von 3 : 4 annähern wird.

Es ist eine mühsame Aufgabe, die Diagonalisierbarkeit einer Matrix A von Fall zu Fall durch die Berechnung von Eigenwerten und Eigenvektoren zu überprüfen. In der Regel ist das leider die einzige Möglichkeit, erfreulicherweise gibt es aber zumindest eine Klasse von Matrizen, der wir die Diagonalisierbarkeit unmittelbar ansehen:

Satz 6.2.3. *Ist A eine symmetrische $n \times n$ -Matrix mit reellen Koeffizienten, so ist A diagonalisierbar.*

Beweis: Wir zeigen die Behauptung durch Induktion nach n . Für $n = 1$ ist $A = (a)$ und wir haben nichts zu zeigen. Wir nehmen also an, dass $n > 1$ und dass wir die Behauptung für $n - 1$ schon gezeigt haben.

Wir betrachten das charakteristische Polynom $P_A(\lambda) = \det(\lambda \cdot E_n - A)$ von A . Als erstes behaupten wir, dass $P_A(\lambda)$ in Linearfaktoren zerfällt, dass also gilt

$$P_A(\lambda) = (\lambda - b_1) \cdot (\lambda - b_2) \cdots (\lambda - b_n)$$

mit reellen Zahlen b_1, \dots, b_n (die nicht notwendigerweise paarweise verschieden sind). Dabei ist klar, dass wir eine solche Zerlegung über \mathbb{C} haben, dass

es also komplexe Zahlen b_i mit dieser Eigenschaft gibt, und wir haben noch zu zeigen, dass $b_i \in \mathbb{R}$. Dabei ist zu beachten, dass alles, was wir in diesem Abschnitt über den reellen Zahlen gemacht haben, genauso über den komplexen Zahlen gilt. Ist also b_i eine komplexe Nullstelle von $P_A(\lambda)$, so ist b_i ein komplexer Eigenwert von A , dh. es gibt einen komplexen Vektor $\vec{v} \in \mathbb{C}^n$, verschieden vom Nullvektor, mit

$$A \cdot \vec{v} = b_i \cdot \vec{v}$$

Hierfür gilt nun

$$\begin{aligned} b_i \cdot |\vec{v}|^2 &= b_i \cdot \langle \vec{v}, \vec{v} \rangle \\ &= \langle b_i \cdot \vec{v}, \vec{v} \rangle \\ &= \langle A \cdot \vec{v}, \vec{v} \rangle \\ &= \langle \vec{v}, A \cdot \vec{v} \rangle \\ &= \langle \vec{v}, b_i \cdot \vec{v} \rangle \\ &= \overline{b_i} \cdot \langle \vec{v}, \vec{v} \rangle \\ &= \overline{b_i} \cdot |\vec{v}|^2 \end{aligned}$$

wobei wir Regel 3.4.4 aus Abschnitt 3.4 und Bemerkung 4.2.3 aus Abschnitt 4.2 (also die Eigenschaft, dass $\langle A \cdot \vec{v}, \vec{w} \rangle = \langle \vec{v}, A \cdot \vec{w} \rangle$ für symmetrische Matrizen gilt), benutzt haben.

Insgesamt erhalten wir daraus

$$b_i \cdot |\vec{v}|^2 = \overline{b_i} \cdot |\vec{v}|^2$$

also, da $|\vec{v}|^2 \neq 0$, auch

$$b_i = \overline{b_i}$$

Eine komplexe Zahl und ihre komplex-konjugierte Zahl stimmen aber nur dann überein, wenn die Zahl schon reell ist (also ihr Imaginärteil verschwindet). Damit haben wir gezeigt, dass $P_A(\lambda)$ nur reelle Nullstellen hat.

Nun geben wir uns eine Nullstelle b_1 von $P_A(\lambda)$, also einen Eigenwert von A vor und bestimmen einen Eigenvektor \vec{u}_1 dazu, wobei wir \vec{u}_1 auf die Länge 1 normieren, also annehmen können, dass $|\vec{u}_1| = 1$. Wir ergänzen (mit dem Gram-Schmidt-Orthonormalisierungsverfahren) \vec{u}_1 zu einer Orthonormalbasis von \mathbb{R}^n , wir finden also $\vec{u}_2, \dots, \vec{u}_n$ so, dass $\vec{u}_1, \dots, \vec{u}_n$ eine Orthonormalbasis von \mathbb{R}^n ist.

Der entscheidende Punkt für unsere Überlegungen ist nun

$$A \cdot \vec{u}_i \in \text{Span}(\vec{u}_2, \dots, \vec{u}_n) \quad \text{für alle } i \in \{2, \dots, n\}$$

Jeder Vektor $A \cdot \vec{u}_i$ ($i = 2, \dots, n$) ist also schon in dem von $\vec{u}_2, \dots, \vec{u}_n$ aufgespannten Untervektorraum enthalten. Dazu geben wir uns ein $i \in \{2, \dots, n\}$ vor und schreiben

$$A \cdot \vec{u}_i = r_1 \vec{u}_1 + r_2 \vec{u}_2 + \dots + r_n \vec{u}_n$$

und haben zu zeigen, dass $r_1 = 0$. Dazu rechnen wir

$$\begin{aligned} r_1 &= r_1 \cdot \langle \vec{u}_1, \vec{u}_1 \rangle \\ &= \sum_{i=1}^n r_i \cdot \langle \vec{u}_i, \vec{u}_1 \rangle \\ &= \langle r_1 \vec{u}_1 + r_2 \vec{u}_2 + \dots + r_n \vec{u}_n, \vec{u}_1 \rangle \\ &= \langle A \cdot \vec{u}_i, \vec{u}_1 \rangle \\ &= \langle \vec{u}_i, A \cdot \vec{u}_1 \rangle \\ &= \langle \vec{u}_i, b_1 \cdot \vec{u}_1 \rangle \\ &= b_1 \cdot \langle \vec{u}_i, \vec{u}_1 \rangle \\ &= 0 \end{aligned}$$

Dabei haben wir, neben den üblichen Eigenschaften des Skalarprodukts, in der zweiten und in der letzten Zeile ausgenutzt, dass

$$\langle \vec{u}_i, \vec{u}_1 \rangle = 0 = \langle \vec{u}_1, \vec{u}_i \rangle$$

für $i \neq 1$ und in der ersten Zeile, dass

$$\langle \vec{u}_1, \vec{u}_1 \rangle = 1$$

Wir betrachten nun die Matrix $U = (\vec{u}_1 \dots \vec{u}_n)$, die die Vektoren $\vec{u}_1, \dots, \vec{u}_n$ als Spaltenvektoren hat. Da die Spalten von U eine Orthonormalbasis von \mathbb{R}^n bilden, ist U eine orthogonale Matrix, so dass also U gemäß Satz 5.2.8 invertierbar ist mit

$$U^{-1} = U^T$$

Wir betrachten nun die Matrix

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

Dann gilt hierfür

$$\begin{aligned}
 B \cdot \vec{e}_1 &= U^{-1} \cdot A \cdot U \cdot \vec{e}_1 \\
 &= U^{-1} \cdot A \cdot \vec{u}_1 \\
 &= U^{-1} \cdot b_1 \cdot \vec{u}_1 \\
 &= b_1 \cdot U^{-1} \cdot \vec{u}_1 \\
 &= b_1 \cdot \vec{e}_1
 \end{aligned}$$

und für $i \geq 2$ gilt nach oben Gezeigtem, wenn wir $A \cdot \vec{u}_i = r_{2,i} \vec{u}_2 + \dots + r_{n,i} \vec{u}_n$ setzen

$$\begin{aligned}
 B \cdot \vec{e}_i &= U^{-1} \cdot A \cdot U \cdot \vec{e}_i \\
 &= U^{-1} \cdot A \cdot \vec{u}_i \\
 &= U^{-1} \cdot (r_{2,i} \vec{u}_2 + \dots + r_{n,i} \vec{u}_n) \\
 &= r_{2,i} \cdot U^{-1} \cdot \vec{u}_2 + \dots + r_{n,i} \cdot U^{-1} \cdot \vec{u}_n \\
 &= r_{2,i} \vec{e}_2 + \dots + r_{n,i} \vec{e}_n
 \end{aligned}$$

und damit hat $U^{-1} \cdot A \cdot U$ die Gestalt

$$U^{-1} \cdot A \cdot U = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & C & \\ 0 & & & \end{pmatrix}$$

mit einer $(n-1) \times (n-1)$ -Matrix C . Genauer gilt $C = (r_{i,j})$ mit den oben bestimmten $r_{i,j}$. Nun gilt für beliebige Vektoren \vec{v} und \vec{w} aufgrund der Bemerkung 4.2.3 und der Regel 4.2.6 aus Abschnitt 4.2 und der Orthogonalität von U :

$$\begin{aligned}
 \langle U^{-1} \cdot A \cdot U \cdot \vec{v}, \vec{w} \rangle &= \langle U^T \cdot A \cdot U \cdot \vec{v}, \vec{w} \rangle \\
 &= \langle A \cdot U \cdot \vec{v}, (U^T)^T \cdot \vec{w} \rangle \\
 &= \langle A \cdot U \cdot \vec{v}, U \cdot \vec{w} \rangle \\
 &= \langle U \cdot \vec{v}, A \cdot U \cdot \vec{w} \rangle \\
 &= \langle \vec{v}, U^T \cdot A \cdot U \cdot \vec{w} \rangle \\
 &= \langle \vec{v}, U^{-1} \cdot A \cdot U \cdot \vec{w} \rangle
 \end{aligned}$$

so dass also die Matrix $U^{-1} \cdot A \cdot U$ nach Bemerkung 4.2.3 aus Abschnitt 4.2 symmetrisch ist. Damit muss aber auch die Matrix C symmetrisch sein, und wir können unsere Induktionsvoraussetzung anwenden und erhalten, dass C diagonalisierbar ist. Wir finden also eine invertierbare $(n-1) \times (n-1)$ -Matrix

V , so dass $V^{-1} \cdot C \cdot V$ eine Diagonalmatrix D ist. Setzen wir

$$W := \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & V & \\ 0 & & & \end{pmatrix}$$

so gilt nach den Regeln für die Matrizenrechnung

$$W^{-1} \cdot U^{-1} \cdot A \cdot U \cdot W = \begin{pmatrix} b_1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & D & \\ 0 & & & \end{pmatrix}$$

Setzen wir also $S = U \cdot W$, so hat $S^{-1} \cdot A \cdot S$ Diagonalgestalt, wie gewünscht.

Bemerkung 6.2.1. Ist A eine symmetrische Matrix, so können wir die Transformationsmatrix S als orthogonale Matrix wählen. Das haben wir implizit mitgezeigt.

Beispiel 6.2.6. Wir betrachten die Matrix

$$A = \begin{pmatrix} 1 & 5 \\ 5 & 1 \end{pmatrix}$$

Dann ist A symmetrisch und damit diagonalisierbar. Hier sehen wir aber auch so sofort, dass

$$P_A(\lambda) = \lambda^2 - 2\lambda - 24 = (\lambda - 6) \cdot (\lambda + 4)$$

dass wir also zwei verschiedene Eigenwerte haben. Ein Eigenvektor zu $\lambda_1 = 6$ ist $\vec{u}_1 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$, und ein Eigenvektor zu $\lambda_2 = -4$ ist $\vec{u}_2 = \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$. Die Vektoren \vec{u}_1, \vec{u}_2 bilden in der Tat eine Orthonormalbasis von \mathbb{R}^2 und mit

$$S := \begin{pmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$$

gilt dann

$$S^T \cdot A \cdot S = \begin{pmatrix} -4 & 0 \\ 0 & 6 \end{pmatrix}$$

Beispiel 6.2.7. Wir betrachten die Matrix

$$A = \begin{pmatrix} 5 & -2 & 2 \\ -2 & 2 & 4 \\ 2 & 4 & 2 \end{pmatrix}$$

Auch hier ist A symmetrisch, und daher sehen wir ohne Rechnung, dass A diagonalisierbar ist. Hier gilt

$$P_A(\lambda) = \lambda^3 - 9 \cdot \lambda^2 + 108 = (\lambda - 6) \cdot (\lambda - 6) \cdot (\lambda + 3)$$

Diese Matrix hat also nur zwei Eigenwerte, und an der Anzahl der Eigenwerte können wir die Diagonalisierbarkeit nicht ablesen. Wenn wir jedoch die Eigenvektoren berechnen, erhalten wir für $\lambda_1 = 6$ die beiden linear unabhängigen Eigenvektoren

$$\vec{u}_1 = \begin{pmatrix} \frac{2}{3} \\ \frac{1}{3} \\ \frac{2}{3} \end{pmatrix}, \quad \vec{u}_2 = \begin{pmatrix} -\frac{2}{3} \\ \frac{2}{3} \\ \frac{1}{3} \end{pmatrix}$$

und für $\lambda_2 = -3$ den Eigenvektor

$$\vec{u}_3 = \begin{pmatrix} \frac{1}{3} \\ \frac{2}{3} \\ -\frac{2}{3} \end{pmatrix}$$

die Vektoren $\vec{u}_1, \vec{u}_2, \vec{u}_3$ bilden eine Orthonormalbasis von \mathbb{R}^3 und für die Matrix

$$S = \begin{pmatrix} \frac{2}{3} & -\frac{2}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{1}{3} & -\frac{2}{3} \end{pmatrix}$$

gilt

$$S^T \cdot A \cdot S = \begin{pmatrix} 6 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & -3 \end{pmatrix}$$

Bemerkung 6.2.2. Um die Diagonalisierbarkeit von Matrizen besser zu verstehen, ist es hilfreich, sich an die Beziehungen zwischen Matrizen und linearen Abbildungen zu erinnern (vergleiche etwa Bemerkung 4.3.2 in Abschnitt 4.3). Zu einer Matrix A gehört eine lineare Abbildung f , und A ist die darstellende Matrix von f bezüglich der Standardbasis. Die Standardbasis ist aber relativ willkürlich dadurch entstanden, dass wir einen Koordinatenursprung gewählt und dann Koordinatenachsen festgelegt haben. Genauso gut hätten wir aber auch andere Achsen wählen können. Wählen wir nämlich im \mathbb{R}^n andere Koordinatenachsen, so können wir diese (ausgehend vom ursprünglichen Koordinatensystem) durch n Vektoren $\vec{v}_1, \dots, \vec{v}_n$ beschreiben. Schreiben wir diese Vektoren als Spalten einer Matrix

$$S = (\vec{v}_1 \ \dots \ \vec{v}_n)$$

so ist diese Matrix invertierbar (das ist sogar äquivalent dazu, dass $\vec{v}_1, \dots, \vec{v}_n$ ein neues Koordinatensystem von \mathbb{R}^n definieren), und in diesem neuen Koordinatensystem wird die lineare Abbildung f durch die Matrix

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

beschrieben. Sind also $\vec{v}_1, \dots, \vec{v}_n$ speziell n linear unabhängige Eigenvektoren von A , so bedeutet das gerade, dass $\vec{v}_1, \dots, \vec{v}_n$ ein neues Koordinatensystem von \mathbb{R}^n definieren, und in diesem Koordinatensystem wird die lineare Abbildung durch die Diagonalmatrix

$$D = S^{-1} \cdot A \cdot S$$

beschrieben. Bezüglich der durch $\vec{v}_1, \dots, \vec{v}_n$ gegebenen Koordinaten gilt also

$$f(\vec{v}_i) = \lambda_i \cdot \vec{v}_i \quad \text{für alle } i \in \{1, \dots, n\}$$

Aufgabe 6.2.1. Untersuchen Sie die Matrix

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

auf Diagonalisierbarkeit. Bestimmen Sie gegebenenfalls eine Transformationsmatrix S , so dass $S^{-1} \cdot A \cdot S$ Diagonalgestalt hat.

Aufgabe 6.2.2. Untersuchen Sie die Matrix

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

auf Diagonalisierbarkeit. Bestimmen Sie gegebenenfalls eine Transformationsmatrix S , so dass $S^{-1} \cdot A \cdot S$ Diagonalgestalt hat.

Aufgabe 6.2.3. Untersuchen Sie die Matrix

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

auf Diagonalisierbarkeit. Bestimmen Sie gegebenenfalls eine Transformationsmatrix S , so dass $S^{-1} \cdot A \cdot S$ Diagonalgestalt hat.

Aufgabe 6.2.4. Untersuchen Sie die Matrix

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

auf Diagonalisierbarkeit. Bestimmen Sie gegebenenfalls eine Transformationsmatrix S , so dass $S^{-1} \cdot A \cdot S$ Diagonalgestalt hat. (Arbeiten Sie über den komplexen Zahlen, falls erforderlich).

Aufgabe 6.2.5. Untersuchen Sie die Matrix

$$A = \frac{1}{15} \begin{pmatrix} 10 & 5 & 10 \\ 5 & -14 & 2 \\ 10 & 2 & -11 \end{pmatrix}$$

auf Diagonalisierbarkeit. Bestimmen Sie gegebenenfalls eine Transformationsmatrix S , so dass $S^{-1} \cdot A \cdot S$ Diagonalgestalt hat.

Aufgabe 6.2.6. Untersuchen Sie die Matrix

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

auf Diagonalisierbarkeit. Bestimmen Sie gegebenenfalls eine Transformationsmatrix S , so dass $S^{-1} \cdot A \cdot S$ Diagonalgestalt hat.

Aufgabe 6.2.7. Untersuchen Sie die Matrix

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

auf Diagonalisierbarkeit. Bestimmen Sie gegebenenfalls eine Transformationsmatrix S , so dass $S^{-1} \cdot A \cdot S$ Diagonalgestalt hat.

Aufgabe 6.2.8. Wir betrachten die symmetrische Matrix

$$A = \begin{pmatrix} 2 & 3 \\ 3 & 2 \end{pmatrix}$$

Bestimmen Sie eine orthogonale Matrix S , so dass $S^T \cdot A \cdot S$ Diagonalgestalt hat.

Aufgabe 6.2.9. Wir betrachten die symmetrische Matrix

$$A = \begin{pmatrix} 10 & 5 & 10 \\ 5 & -14 & 2 \\ 10 & 2 & -11 \end{pmatrix}$$

Bestimmen Sie eine orthogonale Matrix S , so dass $S^T \cdot A \cdot S$ Diagonalgestalt hat.

Aufgabe 6.2.10. Bestimmen Sie (in Abhängigkeit von a) eine orthogonale Matrix S , die die symmetrische Matrix

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

diagonalisiert.

Aufgabe 6.2.11. Bestimmen Sie (in Abhängigkeit von a) eine orthogonale Matrix S , die die symmetrische Matrix

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

diagonalisiert.

Aufgabe 6.2.12. Wir betrachten die Matrizen

$$A_{a,b} = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

mit reellen Zahlen a, b (vergleiche 6.1.9). Bestimmen Sie (in Abhängigkeit von a, b) eine Transformationsmatrix $S = S_{a,b}$, so dass $S^{-1} \cdot A \cdot S$ Diagonalgestalt hat.

Index

- Äquivalenz, 6
- Äquivalenzklasse, 37
 - Repräsentant, 37
 - Repräsentantensystem, 37
- Allquantor, 17
- Aussagenlogik, 3
- Basis, 110, 119
- charakteristisches Polynom, 242
- Cramersche Regel
 - allgemein, 235
 - zweidimensional, 178
- Determinante
 - n -dimensional, 200
 - dreidimensional, 190
 - Laplace Entwicklung, 193, 206
 - Produktsatz, 217
 - Sarrus, 191
 - zweidimensional, 178
- diagonalisierbare Matrix, 250
- Differenzmenge, 32
- Disjunktion, 5
- Drehung, 231
- Eigenvektor, 240
- Eigenwert, 240
- Element, 29
- euklidischer Algorithmus, 64
- Existenzquantor, 17
- Gruppe, 43
 - endliche Gruppe, 47
 - kommutative, 46
 - Ordnung einer Gruppe, 47
 - Ordnung eines Elements, 48
 - zyklisch, 51
- Halbgruppe, 41
- Imaginärteil, 59
- Implikation, 6
- inverses Element, 43
- Isometria, 230
- Junktoren, 4
- Körper, 56
- kartesisches Produkt, 33
- Komplement, 32
- komplexe Zahlen, 58
- Konjunktion, 4
- Kontravalenz, 5
- linear Abbildung, 162
 - Matrix, 163
- linear unabhängig, 87
- lineare Unabhängigkeit, 109, 119
- lineares Gleichungssystem, 131

- Gauß–Normalform, 132
- homogen, 131
- Rang, 132
- Linearkombination, 107, 119
- Matrix, 142
 - Bild, 164
 - diagonalisierbar, 250
 - Diagonalmatrix, 145
 - Dreiecksmatrix, 145
 - Einheitsmatrix, 144
 - Gauß–Normalform, 148
 - inverse Matrix, 219
 - invertierbar, 219
 - Kern, 164
 - Koeffizientenmatrix, 145
 - komplementär, 215
 - Nullität, 165
 - orthogonal, 229
 - Produkt, 156
 - quadratisch, 144
 - Rang, 165
 - regulär, 208
 - Streichungsmatrix, 193
 - symmetrisch, 147
 - transponiert, 146
 - Zeile, 144
- Menge, 29
 - disjunkt, 31
- Monoid, 42
- Negation, 4
- neutrales Element, 42
- Obermenge, 30
- orthogonales Komplement, 98
- Permutationsgruppe, 44
 - Signatur, 45
 - Transposition, 44
- Pivotelement, 136
- Polynom, 54
 - Grad, 54
 - Polynomring, 54
- Prädikatenlogik, 18
 - Alphabet, 18
- Primzahl, 62
- Quantor, 17
- Realteil, 59
- Relation, 34
 - Äquivalenzrelation, 36
 - antisymmetrisch, 36
 - asymmetrisch, 36
 - homogen, 34
 - Ordnung, 38
 - reflexiv, 36
 - strikte Ordnung, 38
 - symmetrisch, 36
 - transitiv, 36
- Ring, 51
 - Nullteiler, 54
- RSA–Kryptosystem, 77
- Satz
 - Austauschsatz von Steinitz, 114
 - kleiner Fermatscher Satz, 73
 - Laplace–Entwicklung, 193, 206
 - von Cauchy–Schwarz, 96, 122
- Schnittmenge, 31

- Skalarprodukt, 95, 121
- Spiegelung, 232
- Standardbasis, 93
- Teiler, 49
- teilerfremd, 63
- Teilmenge, 30
 - echte, 30
- Untervektorraum, 84
 - Basis, 93
 - Dimension, 94
 - Erzeugendensystem, 90
 - Orthonormalbasis, 101, 123
- Vektor, 82
 - Länge, 87
- Vektoren
 - linear unabhängig, 109, 119
 - Lineare Unabhängigkeit, 87
 - orthogonal, 98, 123
- Vektorraum, 81
 - aufgespannter Unterraum, 92
 - Basis, 110, 119
 - Dimension, 114, 120
 - Erzeugendensystem, 108, 119
- Vektorraum \mathbb{R}^n , 83
- Vereinigungsmenge, 31
- Verknüpfung, 40
- vollständige Induktion, 23