Universität Stuttgart

Vorlesung Mathematik für Informatiker
und Softwaretechniker I, WS 2003/2004

Prof. Dr. Anna-Margarete Sändig

Berichte aus dem Institut für Angewandte Analysis und Numerische Simulation

Vorlesungsskript 2004/005

Universität Stuttgart

Vorlesung Mathematik für Informatiker
und Softwaretechniker I, WS 2003/2004

Prof. Dr. Anna-Margarete Sändig

Berichte aus dem Institut für Angewandte Analysis und Numerische Simulation

Vorlesungsskript 2004/005

Institut f¨ur Angewandte Analysis und Numerische Simulation (IANS)
Fakultät Mathematik und Physik
Fachbereich Mathematik
Pfaffenwaldring 57
D-70 569 Stuttgart

E-Mail: `ians-preprints@mathematik.uni-stuttgart.de`
WWW: `http://preprints.ians.uni-stuttgart.de`

ISSN **1611-4176**

Inhaltsverzeichnis

1 Grundlagen **9**

1.1 Aussagenlogik 9
1.2 Mengen, Relationen, Abbildungen 18
1.3 Zahlenmengen 32
1.4 Gruppen, Ringe, Körper 51

2 Lineare Algebra **57**

2.1 Vektorräume 57
2.2 Lineare Abbildungen von Vektorräumen 73
2.3 Matrizen 78
2.4 Lineare Gleichungssysteme 91
2.5 Determinante einer Matrix 105
2.6 Eigenwerte und Diagonalisierbarkeit 115
2.7 Skalarprodukt 127

3 Analysis I (Grundbegriffe) **135**

3.1 Konvergenz in metrischen Räumen 135
3.2 Zahlenfolgen und Zahlenreihen 142
3.3 Stetige Abbildungen in metrischen Räumen 155
3.4 Folgen und Reihen von Funktionen 160
3.5 Spezielle Funktionen 168

Einleitung

Die vierstündige Mathematikvorlesung, die durch zweistündige Übungen ergänzt wird, soll mathematische Grundkenntnisse aus Linearer Algebra und Analysis vermitteln, die für das Studium der Informatik bzw. Softwaretechnik wichtig sind.
Im vorliegenden Skript ist der Stoff des 1. Semesters der Vorlesung und Übungen, die im Wintersemester 2003/04 an der Universität Stuttgart gehalten wurden, zu finden. Die Vorlesung wurde in 3 Blöcke von jeweils 5 Wochen unterteilt, deren Inhalte in den Kapiteln 1, 2 und 3 dargelegt sind.

- **Kapitel 1** ist den **Grundlagen** gewidmet. Es beginnt mit der Aussagenlogik und behandelt danach die Begriffe Menge, Relation und Abbildung. Es werden weiterhin Zahlenmengen eingeführt sowie algebraische Grundbegriffe wie Gruppe, Ring und Körper definiert und an Beispielen erläutert.

- **Kapitel 2** befasst sich mit **Linearer Algebra**. Es werden allgemeine Vektorräume eingeführt und lineare Abbildungen von Vektorräumen untersucht. Daran anschließend werden endlichdimensionale Vektorräume, insbesondere der $\mathbb{R}^n$ und $\mathbb{C}^n$, genauer behandelt. Matrizen und Determinanten werden eingeführt und die Lösungstheorie für lineare Gleichungssysteme ausführlich diskutiert. Ein Abschnitt ist den Eigenwerten, Eigenvektoren und der Diagonalisierbarkeit einer Matrix gewidmet. Das Kapitel über lineare Algebra endet mit der Einführung des Skalarproduktes, othonormierter Systeme sowie orthogonaler Matrizen.

- **Kapitel 3** trägt die Überschrift **Analysis I (Grundbegriffe)**. Am Anfang steht der fundamentale Begriff der Konvergenz. Er wird in metrischen Räumen definiert und an Beispielen erläutert. Es schließt sich ein Abschnitt über Zahlenfolgen und Zahlenreihen an. Der Begriff der Stetigkeit einer Abbildung wird ebenfalls in metrischen Räumen definiert und diskutiert. Die Verbindung zur klassischen Analysis wird betont und in einem gesonderten Abschnitt werden Folgen und Reihen von reellen (komplexen) Funktionen einer rellen (komplexen) Variablen betrachtet. Durch die Einführung von Potenzreihen wird die Definition spezieller Funktionen (Exponential, trigonometrischer und Hyperbelfunktionen sowie ihrer Umkehrfunktionen) vorbereitet und in einem letzten Abschnitt werden diese diskutiert.

Herr Dr. Norbert Röhrl war für die begleitenden zweistündigen Gruppenübungen verantwortlich. Er hat wesentlichen Anteil an der Zusammenstellung der Übungsaufgaben. Ich möchte mich bei ihm für die konstruktive Zusammenarbeit und vor allem für das Korrekturlesen des Skriptes bedanken.

Das vorliegende Skript wurde von Frau stud.math. Iris Pflieger in LaTeXgesetzt. Ich bedanke mich bei ihr für das sorgfältige Schreiben und die sachkundigen Bemerkungen.

Anna-Margarete Sändig — Stuttgart, im Februar 2004

Kapitel 1

Grundlagen

In diesem Kapitel werden Grundbegriffe wie Aussage, Menge, Relation, Abbildung, Zahlenmengen, Gruppe, Ring und Körper eingeführt. Zusammenhänge werden erläutert, einige Eigenschaften werden bewiesen und Beispiele illustrieren die Begriffe. Als Literatur wird empfohlen [3, Kapitel 1-4,7] und [5, Kapitel 1-5].

1.1 Aussagenlogik

Ich beginne mit einem Zitat aus [3].
„Ohne Aussagenlogik keine Schaltkreise und ohne Schaltkreise keine Computer.“
Was verstehen wir unter einer Aussage? Sie sollte in einer Sprache formuliert werden können und eindeutig als wahr oder falsch identifizierbar sein. Die Bewertung „wahr “ oder „falsch “ entspricht bei Schaltkreisen der Zuordnung 1 oder 0.

Definition 1.1. Eine Aussage ist ein Satz in einer Sprache, der entweder wahr (w) oder falsch (f) ist.

Bemerkung 1.2. Diese Definition besagt, dass eine zweiwertige Logik (Aristotelische Logik, tertium non datur) benutzt wird.

Beispiele:

(i) „Was studierst Du?“ Dieser Satz ist keine Aussage

(ii) „11 ist eine Primzahl.“ Dieser Satz ist eine wahre Aussage.

(iii) „Meine Mutter ist jünger als ich.“ Dieser Satz ist eine falsche Aussage.

Wir können neue Aussagen erzeugen, indem wir Abbildungen und Verknüpfungen von Aussagen einführen. Die Werte der Abbildungen bzw. Verknüpfungen können in sogenannten Wahrheitstafeln beschrieben werden.

1. Fall:
Wir gehen von einer Aussage A aus. Die Abbildung r ordne der Aussage A die Aussage $r(A)$ zu. In der entsprechenden Wahrheitstafel werden die Wahrheitswerte w und f eingetragen:

A	$r_1(A)$	$r_2(A)$	$r_3(A)$	$r_4(A)$
w	w	f	w	f
f	f	w	w	f

Es gibt 4 verschiedene Möglichkeiten die Wahrheitswerte der Bilder $r_i(A)$ zu beschreiben. Wir sehen uns die Abbildung r_2 genauer an.

Definition 1.3 (Negation). Als Negation einer Aussage A bezeichnen wir die Aussage $\neg A$ (sprich: „nicht A“), die den Wahrheitswert f annimmt, wenn A wahr ist und den Wahrheitswert w annimmt, wenn A falsch ist.

Beispiele

- Die Aussage A laute: „Die Kreide ist weiß.“ Die Negation ist: „Die Kreide ist nicht weiß.“
- Die Aussage A laute: „$1 = 2$.“ Die Negation ist: „$1 \neq 2$.“

2. Fall:
Wir betrachten zwei Aussagen A und B und wollen diese zu einer Aussage C verknüpfen. Wir schreiben

$$R(A, B) = C.$$

Alle möglichen Verknüpfungen $R_i(A, B), i = 1, \ldots, 16$, können mit Hilfe von Wahrheitstafeln beschrieben werden:

A	B	$R_1(A,B)$	$R_2(A,B)$	...	$R_{16}(A,B)$
w	w	w	w	...	f
w	f	w	w	...	f
f	w	w	w	...	f
f	f	w	f	...	f

Wir sehen uns auch hier einige Verknüpfungen etwas genauer an, die Konjunktion, die Disjunktion, die Implikation und die Äquivalenz:

A	B	Konjunktion $\wedge$	Disjunktion $\vee$	Implikation $A \Rightarrow B$	Äquivalenz $A \Leftrightarrow B$
w	w	w	w	w	w
w	f	f	w	f	f
f	w	f	w	w	f
f	f	f	f	w	w

Tabelle 1.1: Konjunktion, Disjunktion, Implikation und Äquivalenz

Zunächst wird die Konjunktion durch eine Definition eingeführt.

Definition 1.4. Die Verknüpfung der Aussagen A und B durch das logische „und“ nennen wir Konjunktion:

$$R(A, B) = A \wedge B.$$

Sie ist nur dann wahr, wenn beide Aussagen wahr sind.

Umgangssprachlich entspricht die Konjunktion dem Ausdruck „sowohl als auch“.
Die Disjunktion entspricht dem „oder“, mindestens eine Aussage muss wahr sein, damit die Verknüpfung wahr ist. Die entsprechende Definition lautet:

Definition 1.5. Die Verknüpfung der Aussagen A und B durch das logische „oder“ nennen wir Disjunktion:

$$R(A, B) = A \vee B.$$

Die Disjunktion ist wahr, wenn mindestens eine der beiden Aussagen wahr ist, sonst ist sie falsch.

Die Implikation (logische Folgerung) ist nur dann falsch, wenn aus einer wahren eine falsche Aussage gefolgert wird.

Definition 1.6. Die Verknüpfung der Aussagen A und B durch eine logische Folgerung, nennen wir Implikation:

$$R(A, B) = (A \Rightarrow B).$$

Sie ist nur dann falsch, wenn B falsch ist.

In der Mathematik ist man bestrebt, aus wahren Aussagen neue wahre Aussagen abzuleiten. Dies wird in Formulierungen „wenn ... , dann“ ausgedrückt.

Beispiel

Wenn f im Punkt x_0 differenzierbar ist, dann ist f im Punkt x_0 stetig.

$$A = \{f \text{ ist im Punkt } x_0 \text{ differenzierbar}\}$$
$$B = \{f \text{ ist im Punkt } x_0 \text{ stetig}\}.$$

Die Aussage $A \Rightarrow B$ besagt:

$$A \text{ ist hinreichende Bedingung für } B,$$
$$B \text{ ist notwendige Bedingung für } A.$$

Beispiel:

Nicht jede stetige Funktion ist differenzierbar. A kann falsch sein, trotzdem ist B wahr.

Ein weiteres Beispiel aus dem täglichen Leben:

$$\text{Wenn } \underbrace{\text{es regnet}}_{A}\text{, dann } \underbrace{\text{ist die Straße nass}}_{B}.$$

A ist hinreichend für B, aber nicht notwendig, d.h. $B \not\Rightarrow A$. Die Straße kann auch anders nass geworden sein. Ist die Straße nicht nass, dann regnet es auch nicht.

Die Äquivalenz ist nur dann wahr, wenn beide Aussagen gleiche Wahrheitswerte besitzen.

Definition 1.7. Die Verknüpfung der Aussagen A und B durch eine logische Folgerung in beiden Richtungen wird Äquivalenz genannt

$$R(A, B) = (A \Leftrightarrow B) = C.$$

C ist wahr, falls A und B gleiche Wahrheitswerte besitzen.

In der Mathematik wird die Äquivalenz durch die Formulierungen „... genau dann, wenn ...“ ausgedrückt.

Beispiel

A ist eine gerade Zahl, genau dann, wenn A durch 2 teilbar ist.

Die Verknüpfungen können durch logische Ausdrücke aneinander gereiht werden.

Beispiel

A	B	$A \Rightarrow B$	$B \Rightarrow A$	$(A \Rightarrow B) \wedge (B \Rightarrow A)$
w	w	w	w	w
w	f	f	w	f
f	w	w	f	f
f	f	w	w	w

Daraus folgt unmittelbar unter Beachtung von Tabelle 1.1 nachfolgender Satz.

Satz 1.8. *Es gilt:*

$$(A \Rightarrow B) \wedge (B \Rightarrow A) \Leftrightarrow (A \Leftrightarrow B).$$

Satz 1.9. *Es gilt:*

$$A \Rightarrow B \Leftrightarrow (\neg B \Rightarrow \neg A) \Leftrightarrow \neg(A \wedge (\neg B)).$$

Beweis. Wir betrachten die Wahrheitstafel:

A	B	$A \Rightarrow B$	$\neg A$	$\neg B$	$\neg B \Rightarrow \neg A$	$\neg(A \wedge (\neg B))$
w	w	w	f	f	w	w
w	f	f	f	w	f	f
f	w	w	w	f	w	w
f	f	w	w	w	w	w

□

Der Satz 1.9 ist grundlegend für mathematische Beweistechniken, die wir jetzt einführen.

1. Direkter Beweis

A sei wahr. Wir zeigen (evtl. schrittweise), $A \Rightarrow B$ ist wahr.

2. Indirekter Beweis

Wir nehmen an $\neg B$ sei wahr. Wir zeigen $\neg B \Rightarrow \neg A$ ist wahr.

3. Widerspruchsbeweis

Wir zeigen, dass $A \wedge (\neg B)$ falsch ist, d.h. diese Aussage ist ein Widerspruch. Daher muss $\neg(A \wedge (\neg B))$ wahr sein.

Beispiel

Man zeige: Aus $\{|x-1| < 1\} = A$ folgt $\{x < 2\} = B$ für eine feste reelle Zahl x.

1. Direkter Beweis.
Fall a) Es sei $x \geq 1$. Dann ist $|x-1| = x-1 < 1$ und $x < 2$.
Fall b) Es sei $x < 1$. Dann ist $x < 1 < 2$.

2. Indirekter Beweis.
Es gelte $\neg B = \{x \geq 2\}$. Dann ist $|x-1| = x-1 \geq 1$, d.h. $\neg B \Rightarrow \neg A$.

3. Widerspruchsbeweis.
Die Aussage $A \wedge (\neg B) = \{|x-1| < 1\} \wedge \{x \geq 2\}$ ist falsch, da in diesem Fall $|x-1| = x-1 \geq 1$ ist.

Wir haben gesehen, dass $(A \Rightarrow B) \Leftrightarrow \neg(A \wedge (\neg B))$ ist, d.h. dass die Implikation durch die Verknüpfungen $\neg, \wedge$ ersetzt werden kann. Dies gilt auch für die Äquivalenz

$$(A \Leftrightarrow B) \Leftrightarrow (A \Rightarrow B) \wedge (B \Rightarrow A).$$

Die Frage tritt auf: Können wir alle 16 Möglichkeiten der Verknüpfungen von 2 Aussagen auf logische Ausdrücke in denen nur $\neg, \wedge, \vee$ auftreten, zurückführen?

Satz 1.10. *Jede Verknüpfung von 2 Aussagen kann durch logische Ausdrücke, in denen nur $\neg, \wedge, \vee$ auftreten, erzeugt werden.*

Beweis. Es gibt 16 verschiedene Möglichkeiten, wie die Wahrheitswerte einer Verknüpfung angeordnet sind. 8 davon werden durch Negation erzeugt. Von den verbleibenden 8 sind 4 durch Disjunktion, Konjunktion, Implikation und Äquivalenz erklärt. Die restlichen 4 sind:

A	B	$(A \vee \neg B) \vee B$	$\neg B \vee A$	$A \wedge (A \vee \neg B \vee B)$	$B \wedge (A \vee \neg B \vee B)$
w	w	w	w	w	w
w	f	w	w	w	f
f	w	w	f	f	w
f	f	w	w	f	f

$\square$

Folgende Rechenregeln können leicht verifiziert werden:

(1) De Morgan'sche Regeln (Augustus de Morgan(1806-1871))

$$\neg(A \wedge B) = \neg A \vee \neg B, \tag{1.1}$$

$$\neg(A \vee B) = \neg A \wedge \neg B. \tag{1.2}$$

(2) Distributivgesetze

$$A \vee (B \wedge C) = (A \vee B) \wedge (A \vee C), \tag{1.3}$$

$$A \wedge (B \vee C) = (A \wedge B) \vee (A \wedge C). \tag{1.4}$$

(3) Transitivität der Implikation

$$(A \Rightarrow B) \wedge (B \Rightarrow C) \Rightarrow (A \Rightarrow C). \tag{1.5}$$

Quantoren

In der Mathematik benutzt man auch die Symbole
$\forall$: *für alle* (Elemente),
$\exists$: *es existiert* ein (Element),
$\exists\,!$: *es existiert genau* ein (Element).

Beispiele

- $(\forall\, x \in \mathbb{N} :\ x^2 \in \mathbb{N})$ bedeutet: für alle x aus der Menge der natürlichen Zahlen $\mathbb{N}$ gilt, dass x^2 aus $\mathbb{N}$ ist.
- $(\exists x \in \mathbb{N} :\ x \leq 4)$ bedeutet: es existiert ein x aus $\mathbb{N}$ mit der Eigenschaft, dass $x \leq 4$ ist.
- $(\exists\,!x \in \mathbb{N} :\ x - 1 = 5)$ bedeutet: es gibt genau ein $x \in \mathbb{N}$, so dass $x - 1 = 5$ ist.

Negation von Aussagen, die Quantoren enthalten

Sei $A = A(x)$ eine Aussage, die sich auf Elemente x aus einer Menge M bezieht.
Die Negation der Aussage

$$\forall\, x \,\text{aus}\, M:\; A(x)\,\text{ist wahr} \tag{1.6}$$

lautet:

$$\exists\, x \,\text{aus}\, M:\; A(x)\,\text{ist falsch.} \tag{1.7}$$

Die Negation der Aussage

$$\exists\, x \,\text{aus}\, M:\; A(x)\,\text{ist wahr} \tag{1.8}$$

lautet:

$$\forall\, x \,\text{aus}\, M:\; A(x)\,\text{ist falsch.} \tag{1.9}$$

Logische Ausdrücke und Schaltkreise

In der Hardware-Entwicklung wird vorgegeben, welche Werte durch einen Schaltkreis realisiert werden sollen, d.h. die Wahrheitstafel ist damit festgelegt. Dann wird der Schaltkreis entworfen, der zu diesem Ergebnis führt. Dabei soll eine möglichst „optimale“ Schaltung gefunden werden.
Als Beispiel sehen wir uns einen Multiplexer an.

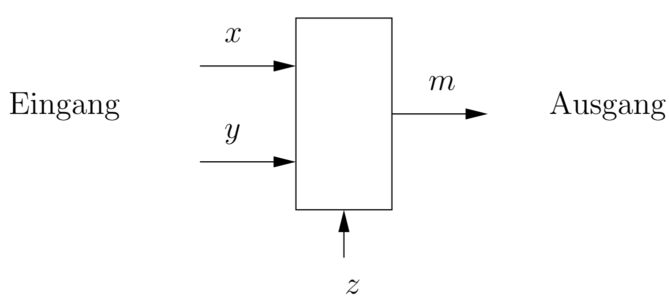


Abbildung 1.1: Multiplexer

z wählt aus, welche Daten weitergeleitet werden, z heißt Steuergröße. Die Steuerregel laute: Falls z wahr ist, dann soll x übertragen werden, falls z falsch ist, dann soll y übertragen werden. Die Wahrheitstafel für den Multiplexer lautet also:

z	x	y	m	
w	w	w	w	$z \wedge x \wedge y = A_1$
w	w	f	w	$z \wedge x \wedge (\neg y) = A_2$
w	f	w	f	
w	f	f	f	
f	w	w	w	$\neg z \wedge x \wedge y = A_3$
f	w	f	f	
f	f	w	w	$\neg z \wedge \neg x \wedge y = A_4$
f	f	f	f	

Wir konstruieren den entsprechenden logischen Ausdruck. Dazu sehen wir uns die Zeilen an, in denen m den Wert w annimmt und verbinden die logischen Ausdrücke durch Konjunktionen. Der Ausgang m ist durch den Wert w belegt, falls

$$A = A_1 \vee A_2 \vee A_3 \vee A_4$$

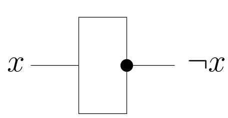
wahr ist. Durch Nachprüfen kann man sich überzeugen, dass alle A_i, $i = 1, 2, 3, 4$, in den übrigen Fällen den Wert f annehmen und damit A den Wert f besitzt. A stellt also den logischen Ausdruck dar, der den Ausgang m beschreibt.
Mit Hilfe der Rechenregel (1.4) kann A vereinfacht werden. Dazu setzen wir $z \wedge x = p$, $\neg z \wedge y = q$ und erhalten:

$$\begin{aligned} A &= (p \wedge y) \vee (p \wedge \neg y) \vee (q \wedge x) \vee (q \wedge \neg x) \\ &\stackrel{(1.4)}{=} p \wedge (y \vee \neg y) \vee (q \wedge (x \vee \neg x) \\ &= p \vee q = (z \wedge x) \vee (\neg z \wedge y). \end{aligned} \tag{1.10}$$

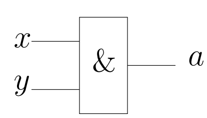
A ist durch eine entsprechende Schaltung realisierbar. Wir führen dazu den Negator, die Parallel- und Reihenschaltung als Bausteine ein.

- Negator: x wird in $\neg x$ umgewandelt.
- Reihenschaltung: Die Eingänge x und y gehen in $a = x \wedge y$ über.
- Parallelschaltung: Die Eingänge x und y gehen in $a = x \vee y$ über.

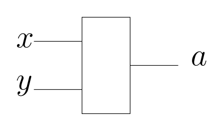
Diese Bausteine werden schematisch wie folgt dargestellt.



Negator

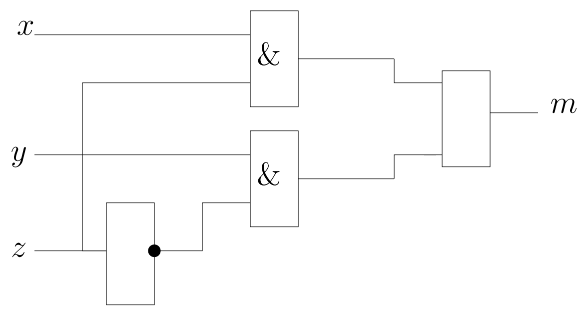


Reihenschaltung



Parallelschaltung

Der folgende Schaltkreis entspricht dem logischen Ausdruck A.



Bemerkung 1.11. Es können sehr viel mehr Eingänge auftreten, der Multiplexer wird dann komplizierter. Im Allgemeinen ist das Konstruieren einer optimalen Schaltung ein schwieriges Problem.

Übungsaufgaben

1) Beweisen Sie mit Hilfe von Wahrheitstafeln, dass die folgenden Aussagen wahr sind:

1. De Morgan'sche Regeln: $\neg(A \wedge B) = \neg A \vee \neg B$ und $\neg(A \vee B) = \neg A \wedge \neg B$,
2. Distributivgesetze: $A \vee (B \wedge C) = (A \vee B) \wedge (A \vee C)$ und $A \wedge (B \vee C) = (A \wedge B) \vee (A \wedge C)$,
3. Transitivität der Implikation: $\big[(A \Rightarrow B) \wedge (B \Rightarrow C)\big] \Rightarrow (A \Rightarrow C)$.

2) Die Aussage A sei gegeben durch $x^2 < 4$, wobei $x \in \mathbb{R}$ ist. Geben Sie jeweils eine Bedingung für x an, die für A

1. notwendig, aber nicht hinreichend,
2. hinreichend, aber nicht notwendig,
3. hinreichend und notwendig ist.

3) Einer der vier Herren Krause, Lehmann, Meier und Schulze ist von Beruf Arzt, ein anderer Ingenieur, ein dritter Lehrer und der vierte Notar. Welchen Beruf übt jeder dieser vier aus, wenn die drei folgenden Aussagen falsch sind?

1. Herr Meier ist nicht Lehrer und auch nicht Ingenieur.
2. Herr Meier ist nicht Notar und Herr Schulze nicht Ingenieur.
3. Herr Lehmann ist Notar.

4) Eine Gruppe von Rittern und Lügnern fuhr auf einen Camping-Trip. Nachdem sie ihre Zelte am Ende eines langen Tages aufgeschlagen hatten, fing Thomas (der mit Abstand beste Koch von allen) an, den Eintopf am Lagerfeuer anzurühren, während alle anderen sich im Kreis um das Feuer niederließen und ihm zuschauten. Ihm fiel auf, dass jeder im Kreis die beiden Personen neben sich zu kennen schien. Thomas aber kannte niemanden außer seinem guten Freund Richard.
An dieser Stelle sollte man bemerken, dass Ritter wegen ihres Ehrenkodex immer die Wahrheit sagen und Lügner immer verlässlich lügen. Ritter und Lügner unterscheiden sich nicht äußerlich, aber wenn sich zwei Leute kennen, wissen sie voneinander, ob sie Ritter oder Lügner sind.

Um die Runde etwas näher kennenzulernen, fragte er eine Person im Kreis:
"Du und die beiden, die neben dir sitzen: ist eine ungerade Zahl von Lügnern in dieser kleinen Menge?"
Die Person antwortete. Thomas fragte eine andere beliebig gewählte Person und erhielt die gleiche Antwort. Wen er auch fragte, er erhielt immer wieder diese Antwort. Schließlich, nachdem er bereits alle anderen gefragt hatte, wandte er sich mit derselben Frage an Richard. Überraschenderweise war dessen Antwort anders als all die anderen.
Thomas hielt einen Moment inne und fragte dann Richard: "Sitzt du zwischen zwei Rittern?", worauf ihm Richard abermals so antwortete.
Nickend erklärte Thomas: "So, dann sind die Lügner hier in der Überzahl", und kümmerte sich wieder um das Essen.
Falls n die Zahl der Leute auf dem Camping-Ausflug ist, wie viele davon sind Ritter bzw. Lügner, und was sind Thomas und Richard? Gibt es Einschränkungen an möglichen Zahlen n?
(Abgedruckt mit freundlicher Genehmigung des Autors Nick Reed, Southampton, England.)

5) Seien A, B, C nichtleere Teilmengen von $\mathbb{R}$. Formulieren Sie die folgenden Aussagen durch Quantoren und verneinen Sie sie:

1. Für alle $x \in A$ gibt es mindestens ein $y \in B$, so dass $y = x + 10$.
2. Für alle $z \in C$ gibt es $x \in A$ und $y \in B$, so dass $z \leq x + y \leq z + 5$.

6) Das aus der Schule bekannte ε-δ Kriterium für Stetigkeit einer Funktion f an der Stelle x_0 lautet

$$\forall \varepsilon > 0 \; \exists \delta \; \forall x : \big(|x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \varepsilon \big).$$

Negieren Sie diese Aussage!

7) Vereinfachen Sie den folgenden Ausdruck

$$(\neg x \wedge y \wedge z) \vee (\neg x \wedge y \wedge \neg z) \vee (x \wedge \neg y \wedge z).$$

1.2 Mengen, Relationen, Abbildungen

Wir beschäftigen uns zunächst mit Mengen.

Mengen

Bei der Einführung der Quantoren hatten wir bereits von Mengen gesprochen. Wir geben jetzt eine Definition an, die vom Begründer der Mengenlehre, Georg Cantor (1845 - 1918), stammt.

Definition 1.12. Unter einer Menge M verstehen wir jede Zusammenfassung von bestimmten wohlunterschiedenen Objekten der Anschauung oder des Denkens, welche die Elemente von M genannt werden, zu einem Ganzen.

Kurz: Eine Menge M ist eine Sammlung von Objekten (Elementen).

Gehört ein Element x zur Menge M, dann schreiben wir $x \in M$, gehört ein Element x nicht zur Menge M, dann wird dies durch $x \notin M$ ausgedrückt.
Mengen können wie folgt beschrieben werden:

- explizit, durch Aufzählen der Elemente, die in geschweiften Klammern gesetzt werden, z.B.
$$M = \{1, 2, 3\} = \{3, 2, 1\}, \quad M = \{1, 2, \{1, 2\}\},$$

- durch Eigenschaften, z.B.
$$M = \{n \in \mathbb{N} : n \text{ ist gerade}\}, \quad M = \{n \in \mathbb{N} : n < 10\}.$$

Beispiele für Zahlenmengen

$$\begin{aligned} &\mathbb{N} = \{1, 2, 3, \dots\} && \text{Menge der natürlichen Zahlen,} \\ &\mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\} && \text{Menge der ganzen Zahlen,} && (1.11) \\ &\mathbb{Q} = \{\frac{a}{b} : a \in \mathbb{Z} \wedge b \in \mathbb{N}\} && \text{Menge der rationalen Zahlen.} && (1.12) \end{aligned}$$

$\emptyset$ ist die leere Menge. Diese Menge enthält kein Element.

Vergleich von Mengen

Gleichheit: Zwei Mengen M_1 und M_2 sind gleich, wenn sie die gleichen Elemente enthalten, d.h.
$$x \in M_1 \Leftrightarrow x \in M_2.$$

Teilmengen: M_1 ist Teilmenge von M_2, geschrieben als $M_1 \subseteq M_2$, falls für alle $x \in M_1$
$$x \in M_1 \Rightarrow x \in M_2$$
gilt. Falls $M_1 \neq M_2$ dann ist M_1 eine echte Teilmenge von M_2, geschrieben als $M_1 \subset M_2$. M_2 wird auch Obermenge von M_1 genannt.

Beispiele

- Es ist $\mathbb{N} \subset \mathbb{Z}$.

- Sei M eine beliebige Menge, dann ist $\emptyset \subseteq M$.

Verknüpfungen von Mengen

Seien M_1 und M_2 zwei Mengen. Die Menge

$$\begin{aligned} M_1 \cup M_2 &= \{x \,:\, x \in M_1 \vee x \in M_2\} && \text{heißt Vereinigung von } M_1 \text{ und } M_2, \\ M_1 \cap M_2 &= \{x \,:\, x \in M_1 \wedge x \in M_2\} && \text{heißt Durchschnitt von } M_1 \text{ und } M_2, \\ M_1 \setminus M_2 &= \{x \,:\, x \in M_1 \wedge x \notin M_2\} && \text{heißt Differenz von } M_1 \text{ und } M_2. \end{aligned}$$

Falls $M_2 \subseteq M_1$ ist, dann heißt $M_1 \setminus M_2$ auch Komplement von M_2 in M_1 und wir schreiben $M_1 \setminus M_2 = M_2^c$. Der obere Index bezieht sich somit auf die Grundmenge M_1.
M_1 und M_2 heißen disjunkt, falls $M_1 \cap M_2 = \emptyset$ gilt.

Potenzmenge

Für jede Menge M können wir eine neue Menge bilden, die als Elemente alle Teilmengen von M besitzt.

Definition 1.13. Die Potenzmenge $\mathcal{P}(M)$ einer Menge ist die Menge aller Teilmengen von M.

Beispiel

Es sei $M = \{1, 2, 3\}$. Dann ist

$$\mathcal{P}(M) = \{\emptyset, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\}.$$

Wir sehen, die Menge M, die aus 3 Elementen besteht, besitzt eine Potenzmenge, die aus 8 Elementen besteht. Es gilt allgemein.

Satz 1.14. *Die Potenzmenge einer Menge M von n Elementen besitzt 2^n Elemente.*

Beweis. Diese Aussage kann durch vollständige Induktion bewiesen werden. □

Rechengesetze

Satz 1.15. *Für eine Menge M und $S, T, V \in \mathcal{P}(M)$ gelten folgende Rechengesetze.*
Kommutativität:

$$\begin{aligned} S \cup T &= T \cup S, \\ S \cap T &= T \cap S. \end{aligned}$$

Assoziativität:

$$\begin{aligned} (S \cup T) \cup V &= S \cup (T \cup V), \\ (S \cap T) \cap V &= S \cap (T \cap V). \end{aligned}$$

Distributivität:

$$S \cup (T \cap V) = (S \cup T) \cap (S \cup V),$$
$$S \cap (T \cup V) = (S \cap T) \cup (S \cap V).$$

De Morgan'sche Regeln:

$$(S \cup T)^c = S^c \cap T^c,$$
$$(S \cap T)^c = S^c \cup T^c.$$

Weitere Rechengesetze:

$$\begin{aligned}
S \cup \emptyset &= S, & S \cap \emptyset &= \emptyset,\\
S \cup M &= M, & S \cap M &= S,\\
S \cup S &= S, & S \cap S &= S,\\
S \cup S^c &= M, & S \cap S^c &= \emptyset,\\
M^c &= \emptyset, & \emptyset^c &= M,\\
(S \setminus T) \setminus V &= S \setminus (T \cup V), & S \setminus (T \setminus V) &= (S \setminus T) \cup (S \cap V),\\
S \cup (T \setminus V) &= S \cup T \setminus (V \setminus S), & S \cap (T \setminus V) &= S \cap T \setminus (S \cap V).
\end{aligned}$$

Kartesisches Produkt

Die Verknüpfung zweier Mengen zu einer Produktmenge spielt eine wichtige Rolle bei der Einführung von Relationen.

Definition 1.16. Das kartesische Produkt zweier Mengen M_1 und M_2 ist die Menge aller geordneten Paare (m_1, m_2) mit $m_1 \in M_1$ und $m_2 \in M_2$:

$$M_1 \times M_2 = \{(m_1, m_2) \,:\, m_1 \in M_1, m_2 \in M_2\}.$$

Entsprechend ist

$$M_1 \times \cdots \times M_r = \{(m_1, m_2, \ldots, m_r) \,:\, m_j \in M_j, 1 \le j \le r\},$$

wobei $(m_1, m_2, \ldots, m_r)$ ein geordnetes r-Tupel ist.

Beispiel

Wir betrachten das kartesische Produkt zweier Intervalle,

$$M_1 = [1,2] = \{x \in \mathbb{R} \,:\, 1 \le x \le 2\}, \quad M_2 = [3,5] = \{y \in \mathbb{R} \,:\, 3 \le y \le 5\}.$$

Es ist $M_1 \times M_2 = \{(x,y) \,:\, 1 \le x \le 2 \wedge 3 \le y \le 5\}$ ein Rechteck (siehe Abbildung 1.2).

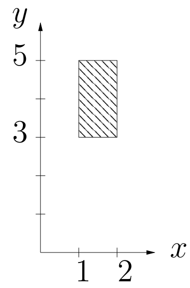


Abbildung 1.2: $M_1 \times M_2$

Relationen

Die Elemente einer gegebenen Menge können eine oder mehrere Eigenschaften besitzen. Sehen wir z.B. die Menge der Studenten in diesem Hörsaal an. Wir können die Merkmale größer oder kleiner gleich als 1.80 m, weiblich oder männlich, Einkommen kleiner gleich oder größer als 300 Euro, Brille oder nicht Brille jedem Studenten zuordnen.
Wir stellen damit eine Relation (Beziehung) zwischen der Menge der Studenten M und einer Menge N von Merkmalen auf, was der Anlage einer Datenbank entspricht.

Bemerkung 1.17. Tabellen (Datenbanken), in denen Informationen aufgelistet sind, entsprechen Relationen. Die Relationen sind Teilmengen von $M_1 \times \cdots \times M_r$, wobei r z.B. die Anzahl der Spalten der Tabellen sind und M_i die Elemente enthalten, die in den Spalten auftreten. Um Informationen zu extrahieren oder neue zu gewinnen, können mengentheoretische Verknüpfungen herangezogen werden, z.B. für Teilmengen R_1 und R_2 von $M_1 \times \cdots \times M_r$ ist erklärt:

1. $R_1 \cup R_2$ (R_1 UNION R_2) erzeugt alle Tupel, die in R_1 oder R_2 vorkommen,
2. $R_1 \setminus R_2$ (R_1 MINUS R_2) enthält alle Tupel, die in R_1 aber nicht in R_2 vorkommen,
3. $R_1 \times R_2$ (R_1 TIMES R_2) ist das kartesische Produkt.

Beispiel

Es sei $M = \{1, 2, 3, 4\}$ eine Menge von 4 Studenten, $N = \{a, b, c\}$ eine Menge von 3 Merkmalen. Eine Relation zwischen M und N kann sein

$x \in M$	$y \in N$	$(x, y) \in M \times N$
1	a	$(1, a)$
1	b	$(1, b)$
2	a	$(2, a)$
3	b	$(3, b)$
3	c	$(3, c)$
4	c	$(4, c)$

d.h. Student 1 besitzt die Merkmale a und b, Student 2 das Merkmal a, Student 3 die Merkmale b und c und Student 4 das Merkmal c. Dies kann auch folgendermaßen dargestellt werden.

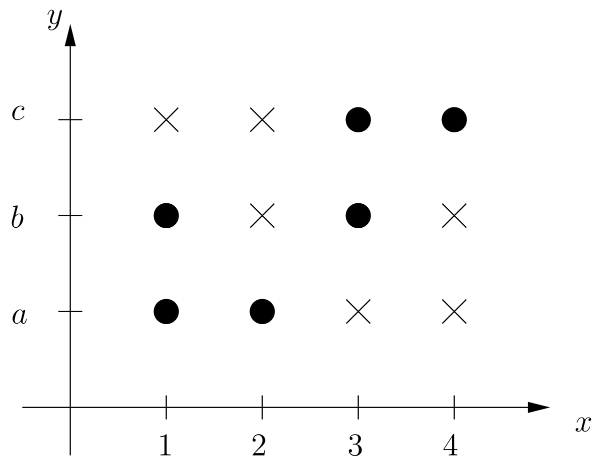


Abbildung 1.3: Merkmalsverteilung

Die dicken runden Punkte kennzeichnen die obige Relation. Sie bilden eine Teilmenge von $M \times N$.

Definition 1.18 (Relation). Eine Teilmenge R des kartesischen Produkts $M \times N$ der Mengen M und N heißt binäre Relation. Eine n-näre Relation zwischen den Mengen $M_1, M_2, \dots, M_n$ ist eine Teilmenge des kartesischen Produkts $M_1 \times M_2 \times \dots \times M_n$

$$R \subseteq M_1 \times M_2 \times \dots \times M_n.$$

In unserem obigen Beispiel ist R eine binäre Relation

$$R = \{(1, a), (1, b), (2, a), (3, b), (3, c), (4, c)\}.$$

Aus gegebenen Relationen können neue durch Durchschnittsbildung, Vereinigungen, Differenzen, u.s.w. erzeugt werden. Eine weitere Möglichkeit ist die Komposition.

Definition 1.19 (Komposition). Seien $R \subseteq M_1 \times M_2$, $S \subseteq M_2 \times M_3$ binäre Relationen. Dann ist die Komposition $S \circ R$ eine Relation auf $M_1 \times M_3$, die durch

$$S \circ R = \{(m_1, m_3) \in M_1 \times M_3 \,:\, \exists\, m_2 \in M_2 \text{ sodass } (m_1, m_2) \in R, (m_2, m_3) \in S \text{ ist}\}$$

gegeben ist.

Bemerkung 1.20. $S \circ R$ bedeutet: zuerst R dann S ausführen, d.h. von rechts nach links die Relationen abarbeiten.

Beispiel

Wir betrachten eine Familie im eigentlichen Sinne und die Mengen

$$\begin{aligned} M_1 &= \{m_{11}, m_{12}, m_{13}\} = \{\text{männliche Personen, deren Alter} \geq 30 \text{ ist}\}, \\ M_2 &= \{m_{21}, m_{22}, m_{23}, m_{24}\} = \{\text{männliche Personen, deren Alter} < 30 \text{ ist}\}, \\ M_3 &= \{m_{31}, m_{32}, m_{33}, m_{34}\} = \{\text{weibliche Personen, deren Alter} < 30 \text{ ist}\}. \end{aligned}$$

Wir sehen uns die Relation „Vater von“ an:

$$R = \{(m_{11}, m_{21}), (m_{11}, m_{23}), (m_{12}, m_{24})\} \subset M_1 \times M_2,$$

d.h. m_{11} hat die Söhne m_{21} und m_{23}, m_{12} hat den Sohn m_{24}. Weiterhin sehen wir uns die Relation „verheiratet mit “ an

$$S = \{(m_{21}, m_{32}), (m_{24}, m_{33})\} \subset M_2 \times M_3,$$

d.h. der Sohn m_{21} ist mit m_{32} verheiratet, der Sohn m_{24} ist mit m_{33} verheiratet.
Dann ist

$$S \circ R = \{(m_{11}, m_{32}), (m_{12}, m_{33})\} \subset M_1 \times M_3$$

die Relation „Schwiegervater von “ . Die verbindenden Elemente wären die Söhne m_{21} und m_{24}.

Definition 1.21 (Inverse Relation). Ist $R \subseteq M \times N$ eine binäre Relation, dann nennen wir

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

die zu R inverse Relation.

Beispiel

Es sei $R = \{(x, y) \in \mathbb{R} \times \mathbb{R} \,:\, x \geq y\}$. Dann ist $R^{-1} = \{(y, x) \in \mathbb{R} \times \mathbb{R} \,:\, y \geq x\}$. Es ist

$$R^{-1} \circ R = R \circ R^{-1} = I = \{(x, x),\, x \in \mathbb{R}\}.$$

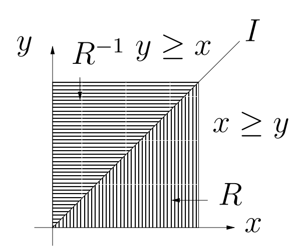


Abbildung 1.4: Die Mengen R und R^{-1}.

Äquivalenz- und Ordnungsrelationen

Wir betrachten spezielle binäre Relationen $R \subset M \times M$.

Definition 1.22. Die Relation $R \subset M \times M$ heißt

- reflexiv, falls für alle $x \in M$ $(x, x) \in R$ ist,
- symmetrisch, falls aus $(x, y) \in R$ folgt $(y, x) \in R$,
- antisymmetrisch, falls aus $(x, y) \in R \wedge (y, x) \in R$ folgt $x = y$,
- transitiv, falls für $(x, y) \in R$ und $(y, z) \in R$ gilt $(x, z) \in R$.

Definition 1.23 (Äquivalenzrelation). Eine reflexive, symmetrische und transitive Relation $R \subset M \times M$ heißt Äquivalenzrelation. Für $(x, y) \in R$ schreibt man $x \sim y$.

Beispiel

Es sei $M = \mathbb{Z}$ die Menge der ganzen Zahlen weiterhin sei $R = \{(x, y) \in \mathbb{Z} \times \mathbb{Z} \,:\, x \equiv y \mod 5\}$. Hierbei bedeutet, dass zwei ganze Zahlen x und y kongruent modulo einer natürlichen Zahl m sind, falls m ein Teiler der Differenz $x - y$ ist (also $x - y = qm$ ist, wobei $q \in \mathbb{Z}$ beliebig ist).
Wir überprüfen die obigen Eigenschaften.

- Reflexivität:
  Es ist $x \equiv x \mod 5$, denn $x - x = 0 \cdot 5$.
- Symmetrie:
  Es sei $x \equiv y \mod 5$, d.h. $x - y = q \cdot 5$. Daraus folgt $y - x = -q \cdot 5$ und $y \equiv x \mod 5$.
- Transitivität:
  Es sei $x \equiv y \mod 5$ und $y \equiv z \mod 5$. Dann ist $x - y = q_1 \cdot 5$ und $y - z = q_2 \cdot 5$. Daraus folgt
  $$x - z = x - y + (y - z) = (q_1 + q_2) \cdot 5 = q \cdot 5.$$

Durch eine Äquivalenzrelation kann man die Menge M in zueinander disjunkte Teilmengen, die Äquivalenzklassen genannt werden, zerlegen. Dabei gehören die Elemente x und y, für die $(x, y) \in R$, bzw. für die $x \sim y$ ist, zu einer Klasse.

Definition 1.24 (Äquivalenzklasse). Für $x \in M$ nennen wir die Menge

$$R[x] = \{y \in M \,:\, x \sim y\}$$

die Äquivalenzklasse von x bezüglich der Äquivalenzrelation R.

Beispiel

Im obigen Beispiel ist

$$
\begin{aligned}
R[0] &= \{y \in \mathbb{Z} : y = x - q5, x = 0, q \in \mathbb{Z}\} \\
&= \{\dots, -10, -5, 0, 5, 10, \dots\} \\
R[1] &= \{\dots, -9, -4, 1, 6, 11, \dots\} \\
R[2] &= \{\dots, -8, -3, 2, 7, 12, \dots\} \\
R[3] &= \{\dots, -7, -2, 3, 8, 13, \dots\} \\
R[4] &= \{\dots, -6, -1, 4, 9, 14, \dots\}.
\end{aligned}
$$

Damit ist $\mathbb{Z} = R[0] \cup R[1] \cup \dots \cup R[4]$ und $R[i] \cap R[j] = \emptyset$ für $i \neq j$, $i, j \in \{0, 1, 2, 3, 4\}$. Die Zahlen i mit $i \in \{0, 1, 2, 3, 4\}$ heißen Repräsentanten von $R[i]$.
Die Menge $\mathbb{Z}$ ist damit in 5 disjunkte Teilmengen zerlegt worden, allgemeiner spricht man von einer Partition $\{R[i]\}_{i=0,1,\dots,4}$ von $\mathbb{Z}$.

Definition 1.25 (Partition einer Menge M). Eine Menge von Teilmengen $\{K_i \subset M\}_{i \in I}$ mit $K_i \cap K_j = \emptyset$ für alle $i, j \in I, i \neq j$, nennen wir Partition von M falls

$$\bigcup_{i \in I} K_i = M$$

ist.

Bemerkung 1.26. Die Indexmenge I kann aus endlich vielen oder unendlich vielen Elementen bestehen. So kann z.B. die Menge der natürlichen Zahlen $\mathbb{N}$ wie folgt dargestellt werden

$$\mathbb{N} = \bigcup_{i=1}^{\infty} \{i\}.$$

Satz 1.27. *Jede Äquivalenzrelation $R \subset M \times M$ erzeugt eine Partition von M, jede Partition bestimmt eine Äquivalenzrelation*

Es wird empfohlen, den Beweis als Übung auszuführen (siehe Aufgabe 9).

Beispiel

Sei M die Menge aller Wörter. Wir führen in M folgende Äquivalenzrelation ein: zwei Wörter sind äquivalent, falls sie den gleichen Anfangsbuchstaben besitzen. Dadurch wird eine Partition erzeugt.

Ordnungsrelation

Die Äquivalenzrelation kann auch als „Gleichheitsrelation“ angesehen werden. Ordnungsrelationen sollten Vergleiche wie „$\leq$ “ bzw. „$\geq$ “ ermöglichen.

Definition 1.28 (Ordnungsrelation). Eine reflexive, antisymmetrische und transitive Relation $R \subset M \times M$ heißt Ordnungsrelation. Wir schreiben in diesem Fall

$$(x, y) \in R \text{ als } x \leq y.$$

Beispiele

1° Sei M die Menge der natürlichen Zahlen $\mathbb{N}$ und $R = \{(n_1, n_2) : n_1 \leq n_2\}$. Dann gilt

- $(n_1, n_1) \in R$,
- $(n_1, n_2) \in R \wedge (n_2, n_1) \in R$, d.h. $n_1 \leq n_2 \wedge n_2 \leq n_1$, dann ist $n_1 = n_2$,
- $(n_1, n_2) \in R \wedge (n_2, n_3) \in R$, d.h. $n_1 \leq n_2, n_2 \leq n_3$, dann ist $n_1 \leq n_3$ und $(n_1, n_3) \in R$.

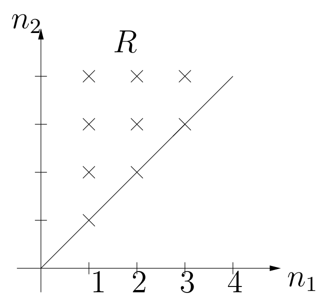


Abbildung 1.5: Die Relation R

2° Sei $M = \mathbb{N} \times \mathbb{N} = \mathbb{N}^2$. Für zwei Elemente $\boldsymbol{x} = (x_1, x_2)$, $\boldsymbol{y} = (y_1, y_2) \in \mathbb{N}^2$ sei folgende Relation definiert:

$$R = \{(\boldsymbol{x}, \boldsymbol{y}) : x_1 \leq y_1\}.$$

Diese Relation ist **keine** Ordnungsrelation, da sie nicht antisymmetrisch ist:

$$(\boldsymbol{x}, \boldsymbol{y}) \in R \ \wedge \ (\boldsymbol{y}, \boldsymbol{x}) \in R$$

bedeutet $x_1 = y_1$, aber nicht $\boldsymbol{x} = \boldsymbol{y}$.
Wir führen in $\mathbb{N}^2$ eine lexikographische Ordnungsrelation ein:

$$R_{\text{lex}} = \{(\boldsymbol{x}, \boldsymbol{y}) : x_1 \leq y_1 \ \wedge \text{ falls } x_1 = y_1, \text{ dann } x_2 \leq y_2\}.$$

Es gilt tatsächlich:

- $(\boldsymbol{x}, \boldsymbol{x}) \in R_{\text{lex}}$.
- Seien $(\boldsymbol{x}, \boldsymbol{y}) \in R_{\text{lex}} \ \wedge \ (\boldsymbol{y}, \boldsymbol{x}) \in R_{\text{lex}}$. Dann ist $y_1 = x_1$ und $x_2 \leq y_2 \ \wedge \ y_2 \leq x_2$, woraus $\boldsymbol{x} = \boldsymbol{y}$ folgt.
- Seien $(\boldsymbol{x}, \boldsymbol{y}) \in R_{\text{lex}} \ \wedge \ (\boldsymbol{y}, \boldsymbol{z}) \in R_{\text{lex}}$. Es ist $x_1 \leq y_1 \leq z_1$. Falls $x_1 = y_1 = z_1$, dann ist $x_2 \leq y_2 \leq z_2$, also $x_2 \leq z_2$ und somit $(\boldsymbol{x}, \boldsymbol{z}) \in R_{\text{lex}}$.

Somit ist diese lexikographische Ordungsrelation im mathematischen Sinn wohl definiert.

Abbildungen

Seinen M und N Mengen. Wir führen den Begriff einer Abbildung von Elementen $x \in M$ auf Elemente $y \in N$ ein.

Definition 1.29 (Abbildungen). Die Zuordnungsvorschrift f, die Elemente $x \in M$ eindeutig auf Elemente $y \in N$ abbildet, heißt Abbildung. Es wird kurz geschrieben

$$\begin{aligned} f\, : M \to N \quad & (f \text{ bildet } M \text{ in } N \text{ ab}),\\ x \to f(x) = y \quad & (\text{jedem } x \in M \text{ wird eindeutig ein Element } f(x) = y \in N \text{ zugeordnet}). \end{aligned}$$

M heißt Definitionsbereich von f, N heißt Wertebereich von f,
$f(M) = \{y \in N \,:\, \exists x \in M \,:\, f(x) = y\}$ ist das **Bild** von M;
$f^{-1}(Y) = \{x \in M \,:\, f(x) \in Y \subset N\}$ ist das **Urbild** einer Menge $Y \subset f(M)$.

Bemerkung 1.30. Sind $M = N = \mathbb{R}$ die Mengen der reellen Zahlen, dann bezeichnet man f auch als Funktion. Der Graph einer Abbildung ist eine spezielle Relation.

Definition 1.31. Der Graph $G = G(f) \subset M \times N$ einer Abbildung $f \,:\, M \to N$ ist durch

$$G = G(f) = \{(x, y) \in M \times N, y = f(x)\}$$

gegeben.

Beispiel

Es sei $M = N = \mathbb{R}$ die Menge der reellen Zahlen und

$$\begin{aligned} f\, &: \mathbb{R} \to \mathbb{R},\\ f\, &: x \to x^2 \end{aligned}$$

die Funktion $f(x) = x^2$. Der Graph $G = G(f) = \{(x, y) \in \mathbb{R} \times \mathbb{R} = \mathbb{R}^2 \,:\, y = x^2\}$ ist in Abbildung 1.6 dargestellt.

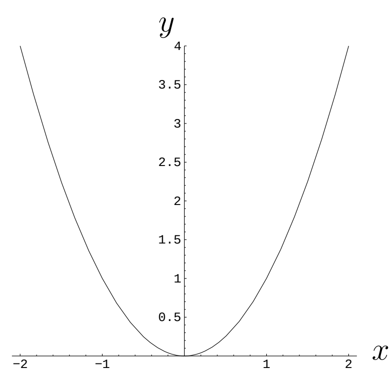


Abbildung 1.6: $f(x) = x^2$

Definition 1.32. Die Abbildung $f : M \to N$ heißt **surjektiv**, falls $f(M) = N$ ist, kurz: f ist eine Abbildung von M auf N.
Die Abbildung $f : M \to N$ heißt **injektiv**, falls gilt: $f(x_1) = f(x_2) \Leftrightarrow x_1 = x_2$, d.h. f ist eine eineindeutige Abbildung von M auf $f(M)$.
Die Abbildung $f : M \to N$ heißt **bijektiv**, falls sie sowohl injektiv als auch surjektiv ist.

Bemerkung 1.33. Eine bijektive Abbildung f zwischen M und N wird kurz geschrieben als $f : M \leftrightarrow N$.

Beispiel

Die obige Abbildung $f : \mathbb{R} \to \mathbb{R}, f(x) = x^2$, ist weder surjektiv noch injektiv.

Mächtigkeit von Mengen

Besitzen zwei Mengen endlich viele Elemente, so können wir durch die Anzahl der Elemente ihre Mächtigkeit definieren, d.h. wir können sofort sehen, ob sie gleich groß (gleich mächtig) sind oder nicht. Bei Mengen, die unendlich viele Elemente enthalten, ist das schwieriger. Betrachten wir z.B. $\mathbb{N}$ und die Teilmenge $\mathbb{N}_{\text{ger}} = \{n \in \mathbb{N} : n \text{ gerade}\}$. Beide enthalten unendlich viele Elemente; sind diese Mengen $\mathbb{N}$ und $\mathbb{N}_{\text{ger}}$ gleich mächtig?

Definition 1.34. Zwei Mengen A und B heißen gleich groß oder gleich mächtig, falls es eine bijektive Abbildung $f : A \leftrightarrow B$ gibt.
Die Menge A heißt endlich, falls es ein $n \in \mathbb{N}$ und eine Bijektion $f : A \to \{1, 2, \ldots, n\}$ gilt, oder $A = \emptyset$ ist.
A heißt unendlich, falls A nicht endlich ist.
A heißt abzählbar unendlich, falls A gleich mächtig zu $\mathbb{N}$ ist.
A heißt überabzählbar, falls A unendlich und nicht abzählbar unendlich ist.

Beispiel

1° Die Menge der geraden natürlichen Zahlen ist gleichmächtig zu $\mathbb{N}$:

$$f : \mathbb{N}_{\text{ger}} \leftrightarrow \mathbb{N}, \quad f : m = 2n \to \frac{m}{2} = n.$$

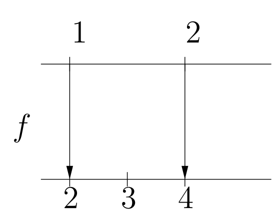


Abbildung 1.7: Zuordnungsvorschrift $f(m)$

2° $\mathbb{Z}$ ist gleichmächtig zu $\mathbb{N}$. Eine Bijektion wäre

$$f : \mathbb{Z} \leftrightarrow \mathbb{N}, \quad f(m) = \begin{cases} 2m & \text{falls } m > 0, \\ -2m+1 & \text{falls } m \leq 0. \end{cases}$$

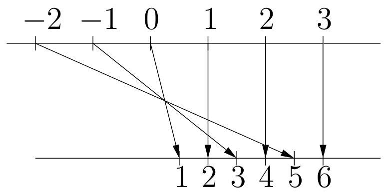


Abbildung 1.8: Zuordnungsvorschrift $f(m)$

3° $\mathbb{N} \times \mathbb{N}$ und $\mathbb{N}$ sind gleichmächtig. Wir schreiben die Paare aus $\mathbb{N} \times \mathbb{N}$ folgendermaßen auf:

$$\begin{array}{ccccccccc}
(1,1) & & (1,2) & & (1,3) & & (1,4) & \dots \\
& \swarrow & & \swarrow & & \swarrow & & \\
(2,1) & & (2,2) & & (2,3) & & (2,4) & \dots \\
& \swarrow & & \swarrow & & & & \\
(3,1) & & (3,2) & & (3,3) & & (3,4) & \dots \\
& \swarrow & & & & & & \\
(4,1) & & (4,2) & & (4,3) & & (4,4) & \dots \\
\vdots & & \vdots & & \vdots & & \vdots &
\end{array}$$

Dies ergibt ein zweidimensionales Nummerierungsschema.

n_2 / n_1	1	2	3	4	5	6	...
1	1	2	4	7	11	16	
2	3	5	8	12	17		
3	6	9	13	18			
4	10	14	19				
5	15	20					
6	21						
⋮							

Die Nummerierung ist eineindeutig. Dieses Vorgehen wird Cantorsches Diagonalverfahren genannt. Durch sie wird eine bijektive Abbildung $f : \mathbb{N} \times \mathbb{N} \leftrightarrow \mathbb{N}$ erzeugt.

4° Die Menge der rationalen Zahlen ist abzählbar unendlich. Um dies zu zeigen, betrachten wir zunächst, ähnlich wie bei den ganzen Zahlen, eine Bijektion f_1:

$$f_1 : \mathbb{Q} \leftrightarrow \mathbb{Q}^+ = \{\frac{a}{b}, a \in \mathbb{N}, b \in \mathbb{N}, a \,\text{und}\, b \,\text{teilerfremd}\}.$$

Danach bilden wir $\mathbb{Q}^+$ injektiv in $\mathbb{N} \times \mathbb{N}$ durch f_2 ab:

$$f_2(\frac{a}{b}) = (a, b).$$

Da a und b teilerfremd sind, ist die Bildmenge $f_2(\mathbb{Q}^+)$ eine echte Teilmenge von $\mathbb{N}\times\mathbb{N}$. Es ist leicht zu sehen, dass $f_2(\mathbb{Q}^+)$ abzählbar ist, indem wir bei der Nummerierung durch das Cantorsche Diagonalverfahren, die Paare überspringen, bei denen Zähler und Nenner nicht teilerfremd sind.
Durch diese Nummerierung haben wir eine bijektive Abbildung von $f_3 : f_2(\mathbb{Q}^+) \leftrightarrow \mathbb{N}$ beschrieben.
Fügen wir die Abbildungen f_1, f_2, f_3 nacheinander aus, erhalten wir, dass

$$f = f_3 \circ f_2 \circ f_1 \, : \mathbb{Q} \leftrightarrow \mathbb{N}$$

die gesuchte bijektive Abbildung ist.

Übungsaufgaben

8) Beweisen Sie folgende Aussagen über Mengen S, T und V

1. $(S \cap T)^c = S^c \cup T^c$ und $(S \cap T)^c = S^c \cup T^c$,
2. $S \cup (T \cap V) = (S \cup T) \cap (S \cup V)$ und $S \cap (T \cup V) = (S \cap T) \cup (S \cap V)$,
3. $(S \subseteq T) \wedge (T \subseteq V) \Rightarrow S \subseteq V$.

9) Beweisen Sie folgenden Satz:
Jede Äquivalenzrelation $R \subseteq M \times M$ erzeugt eine Partition von M. Umgekehrt bestimmt jede Partition eine Äquivalenzrelation.

10) Lexikographische Ordnung. Sei $\Sigma = \{\sqcup, A, B, \ldots, Z\}$ ein Alphabet mit der üblichen alphabetischen Ordnung und $\sqcup \leq x$ für alle $x \in \Sigma$. Das Symbol Σ^* bezeichnet die endlichen Zeichenfolgen (Strings) aus diesem Alphabet. Eine Zeichenfolge a der Länge $|a|$ ist also eine geordnete Menge $a = a_1 a_2 \ldots a_{|a|}$ von Elementen $a_i \in \Sigma$, wobei a_1 das erste, a_2 das zweite, usw. Zeichen darstellt.
Betrachten wir nun Zeichenfolgen a und b der Länge $|a| = |b| = l$.
Wir definieren die Menge

$$M_{ab} = \{i | 1 \leq i \leq l, a_i \neq b_i\}$$

der Stellen, an denen sich die Zeichenfolgen a und b unterscheiden und

$$m_{ab} = \min M_{ab}$$

sei die erste solche Stelle, oder ∞ falls $M_{ab} = \emptyset$ ist.

Wir definieren die Lexikographische Ordnung durch die Relation

$$a \leq b \Leftrightarrow (M_{ab} = \emptyset) \vee (a_{m_{ab}} \leq b_{m_{ab}}).$$

Beweisen Sie, dass die Lexikographische Ordnung eine Ordnungsrelation ist. Geben Sie die dazu inverse Relation an.

11) Sei M eine Menge, dann nennen wir die Relation

$$I_M = \{(x,x) | x \in M\} \subseteq M \times M$$

identische Relation oder Identität in M. Seien N und O weitere Mengen. Zeigen Sie,

1. dass nicht jede Relation $R \subseteq M \times N$ die Gleichungen

$$\begin{aligned} R \circ R^{-1} =& I_N \,, \\ R^{-1} \circ R =& I_M \end{aligned}$$

   erfüllt,

2. dass gilt
$$\left(R^{-1}\right)^{-1} = R,$$

3. und für eine weitere Relation $R_2 \subseteq N \times O$
$$(R_2 \circ R)^{-1} = R^{-1} \circ R_2^{-1}$$

   gilt.

12) Seien X,Y Mengen, $U, V \subseteq X$ und $\tilde{U}, \tilde{V} \subseteq Y$. Zeigen Sie, dass jede Abbildung $f : X \longrightarrow Y$ folgende Eigenschaften erfüllt:

1. $f(U \cap V) \subseteq f(U) \cap f(V)$,
2. $f(U \cup V) = f(U) \cup f(V)$,
3. $f^{-1}(\tilde{U} \cap \tilde{V}) = f^{-1}(\tilde{U}) \cap f^{-1}(\tilde{V})$,
4. $f^{-1}(\tilde{U} \cup \tilde{V}) = f^{-1}(\tilde{U}) \cup f^{-1}(\tilde{V})$.

Geben Sie ein Beispiel für eine Funktion f mit $f(U \cap V) \neq f(U) \cap f(V)$ an.

1.3 Zahlenmengen

Wir haben bereits in vorangegangenen Abschnitten von den Mengen der natürlichen, ganzen, rationalen und reellen Zahlen gesprochen.
Insbesondere die rationalen und reellen Zahlen können im Allgemeinen nur näherungsweise mit Hilfe von Stellenwertsystemen (endliche Darstellung!) auf dem Computer realisiert werden. Über diese Realisierung werden wir später sprechen. Wir sehen uns zunächst an, wie die Zahlenmengen $\mathbb{N}, \mathbb{Z}, \mathbb{Q}$ und $\mathbb{R}$ eingeführt werden können.

Natürliche Zahlen, vollständige Induktion

In der Mathematik werden aus gegebenen Aussagen mit Hilfe logischer Schlussfolgerungen neue Aussagen gewonnen, die Sätze. Dieser Prozess muss irgendwann beginnen. Am Anfang müssen Tatsachen stehen, die als wahr angenommen werden, ohne dass diese bewiesen werden. Diese Grundtatsachen nennt man Axiome.
Guiseppe Peano (1858-1932) hat 1889 das folgende Axiomensystem aufgestellt, wodurch natürliche Zahlen charakterisiert werden.

Definition 1.35 (Peano-Axiome). Die natürlichen Zahlen bilden eine Menge $\mathbb{N}$, auf der eine Abbildung f, die jedem n einen Nachfolger zuordnet;

$$f : \mathbb{N} \to \mathbb{N}$$

erklärt ist, die folgende Eigenschaften besitzt:

($\mathbb{N}$1) $\exists ! 1 \in \mathbb{N} : \ 1 \notin f(\mathbb{N})$, d.h. es existiert genau eine Zahl, die nicht Nachfolger einer anderen Zahl ist.

($\mathbb{N}$2) Die Abbildung f ist injektiv, d.h. aus

$$f(n_1) = f(n_2) \quad \text{folgt } n_1 = n_2.$$

($\mathbb{N}$3) Ist $M \subseteq \mathbb{N}$ und gilt

a) $1 \in M$,

b) aus $n \in M$ folgt $f(n) \in M$,

dann ist $M = \mathbb{N}$.

Bemerkungen 1.36.

- Die Zahl 1 ist die kleinste natürliche Zahl. Das ist Konvention. Man könnte auch mit der 0 beginnen. Um die 0 einzubeziehen führt man $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$ ein.
- Das Axiom ($\mathbb{N}$3) wird auch Induktionsaxiom genannt. Es bildet die Grundlage für ein Beweisverfahren, das vollständige Induktion genannt wird.

Satz 1.37 (Vollständige Induktion). *Für $n \in \mathbb{N}$ sei $A = A(n)$ eine Aussage. Es gelte*

- *$A(1)$ ist wahr (Induktionsanfang),*
- *Für alle $n \in \mathbb{N}$ gilt: Ist $A(n)$ wahr, so folgt auch, dass $A(n+1)$ wahr ist (Induktionsschritt).*

Dann ist auch $A(n)$ wahr für alle $n \in \mathbb{N}$ (Induktionsschluss).

Beweis. Sei $M = \{n \in A(n) : A(n) \text{ ist wahr}\}$. Die Eigenschaften a) und b) des Axioms ($\mathbb{N}$3) sind erfüllt. Daher ist $M = \mathbb{N}$. □

Beispiel

Man beweise: Es sei $q \neq 1$. Dann ist für alle $n \in \mathbb{N}$

$$q^0 + q^1 + \cdots + q^n = \frac{1-q^{n+1}}{1-q}. \tag{1.13}$$

Beweis. Wir betrachten die Aussage

$$A(n) : \left(q^0 + q^1 + \cdots + q^n = \frac{1-q^{n+1}}{1-q} \right).$$

- Die Aussage $A(1) : q^0 + q^1 = 1 + q = \frac{1-q^2}{1-q}$ ist wahr.
- Es sei $A(n)$ wahr. Dann gilt

$$\begin{aligned} \frac{1+q^{n+1}}{1-q} + q^{n+1} &= \frac{1-q^{n+1}+(1-q)q^{n+1}}{1-q} \\ &= \frac{1+(-1+1-q)q^{n+1}}{1-q} \\ &= \frac{1-q^{n+2}}{1-q}, \end{aligned}$$

  was bedeutet, dass $A(n+1)$ wahr ist.

□

Teilbarkeit und Primzahlen

Der Begriff „Teiler“ einer ganzen Zahl ist aus der Schule bekannt.

Definition 1.38. Sei $n \in \mathbb{Z}, m \in \mathbb{N}$. Die Zahl m heißt Teiler von n, in Zeichen $m|n$, wenn

$$\exists\, k \in \mathbb{Z} \,:\, km = n.$$

In diesem Fall heißt n teilbar durch m.

Es gilt:

- Die Zahl 0 ist durch alle m teilbar.
- Ist $m|n_1, m|n_2$ dann folgt $m|(n_1 + n_2)$.

Definition 1.39. Sei $a \in \mathbb{Z}$. Die Menge aller Teiler ist

$$D(a) = \{d \in \mathbb{N} : d|a\}.$$

Seien $a, b \in \mathbb{Z}$. Die Menge aller gemeinsamen Teiler von a und b ist

$$D(a, b) = D(a) \cap D(b).$$

Die Zahl

$$\mathrm{ggT}(a, b) := \max\{d \in D(a) \cap D(b)\}$$

heißt größter gemeinsamer Teiler von a und b.

Beispiel

Sei $a = 30$ und $b = 24$. Dann ist

$$D(30) = \{1, 2, 3, 5, 6, 10, 15, 30\},$$
$$D(24) = \{1, 2, 3, 4, 6, 8, 12, 24\}.$$

Weiterhin ist $D(30, 24) = D(30) \cap D(24) = \{1, 2, 3, 6\}$ und somit ist der größte gemeinsame Teiler von 30 und 24:

$$\mathrm{ggT}(30, 24) = 6.$$

Berechnung des größten gemeinsamen Teilers

Wir gehen vom „Teilen“ mit Rest aus. Sind a und $b \in \mathbb{N}, a > b$. Dann gibt es natürliche Zahlen q, r mit $0 \leq r < b$ und

$$a = qb + r.$$

Beispiel

Wie bei vorherigem Beispiel sei $a = 30$ und $b = 24$. Es gilt

$$30 = 1 \cdot 24 + 6.$$

Hilfssatz 1.40. *Seien $a \in \mathbb{N}$ und $b \in \mathbb{N}$ und $a = qb + r$. Dann ist*

$$D(a, b) = D(b, r), \tag{1.14}$$
$$\mathrm{ggT}(a, b) = \mathrm{ggT}(b, r). \tag{1.15}$$

Beweis. Ist $d \in D(a, b)$, d.h. $d|a \wedge d|b$, dann gilt $d|r$ mit $r = a - qb$. Damit ist $d \in D(b, r)$. Umgekehrt, sei $d \in D(b, r)$ d.h. $d|b \wedge d|r$. Dann ist $d|a$ und $d \in D(a, b)$. Daher gilt (1.14) und (1.15). $\square$

Euklidischer Algorithmus zur Berechnung von $\mathrm{ggT}(a, b)$

Starte mit (a, b).

1. Teste, ob $b|a$.
2. Wenn ja: $D(a, b) = D(b)$ und $\mathrm{ggT}(a, b) = b$; Stopp.
3. Wenn nein: Teile mit Rest $a = qb + r, 0 < r < b$. Zurück zum Anfang mit dem Paar (b, r).

Beispiele

Berechne $\mathrm{ggT}(30, 24)$.

1. Gilt $24|30$? Nein. Daraus folgt $30 = 1 \cdot 24 + 6$.
2. Gilt $6|24$? Ja. Daraus folgt $\mathrm{ggT}(30, 24) = 6$.

Berechne $\mathrm{ggT}(210, 25)$.

1. Gilt $25|210$? Nein. Daraus folgt $210 = 8 \cdot 25 + 10$.
2. Gilt $10|25$? Nein. Daraus folgt $25 = 2 \cdot 10 + 5$.
3. Gilt $5|10$? Ja. Daraus folgt $\mathrm{ggT}(210, 25) = 5$.

Dieses Vorgehen kann allgemein folgendermaßen beschrieben werden.
Seien $a \in \mathbb{N}$ und $b \in \mathbb{N}, a > b$. Dann existieren eindeutig bestimmte Zahlen $r_0, r_1, \ldots, r_{k+2} \in \mathbb{N}$ mit

$$\begin{aligned} r_0 = a &= q_1 r_1 + r_2, && 0 < r_2 < r_1 = b, \\ r_1 = b &= q_2 r_2 + r_3, && 0 < r_3 < r_2 \leq b - 1, \\ &\vdots \\ r_k &= q_{k+1} r_{k+1} + r_{k+2}, && 0 < r_{k+2} < r_{k+1}, \\ r_{k+1} &= q_{k+2} r_{k+2} + 0. \end{aligned}$$

Es ist $\mathrm{ggT}(a, b) = r_{k+2}$.
Das Flussbild dieses Algorithmus ist in Abbildung 1.9 dargestellt.

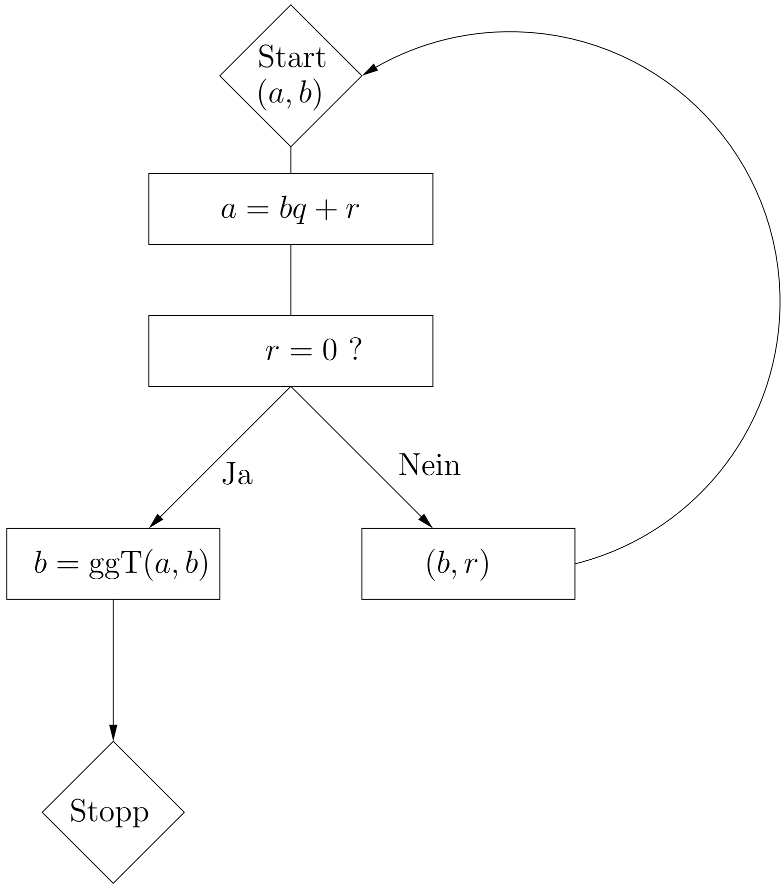


Abbildung 1.9: Berechnung von $\mathrm{ggT}(a, b)$, $a > b$.

Primzahlen

Definition 1.41. Eine natürliche Zahl $p \geq 2$ heißt Primzahl, falls

$$D(p) = \{1, p\}$$

ist.

Lemma 1.42. *Jede natürliche Zahl $n \geq 2$ ist ein Produkt von Primzahlen.*

Beweis. Vollständige Induktion.
Wir betrachten die Aussage $A(n) : n$ ist Produkt von Primzahlen.

Induktionsanfang: $A(2)$ ist wahr.

Induktionsannahme: $A(n)$ sei wahr für $n \geq 3$.

Induktionsschluss: Wir zeigen $A(n + 1)$ ist wahr.
Wir unterscheiden die Fälle

a) $n+1$ sei prim, d.h. $n+1 = p$.

b) $n+1$ ist nicht prim, d.h. $n+1 = n_1 \cdot n_2$ mit $n_1 < n+1, n_2 < n+1$. Nach Induktionsannahme ist $n+1 = n_1 \cdot n_2 = p_1 \dots p_{k_1} q_1 \dots q_{k_2}$, wobei $p_i, q_j, i = 1, \dots, k_1, j = 1, \dots, k_2$ Primzahlen sind.

$\square$

Der folgende Satz wird als Übungsaufgabe formuliert (Aufgabe 16).

Satz 1.43 (Fundamentalsatz der Arithmetik). *Jede natürliche Zahl $n \geq 2$ lässt sich bis auf die Reihenfolge der Faktoren auf genau eine Weise als Produkt von Primzahlen schreiben.*

Satz 1.44 (Satz von Euklid). *Es gibt unendlich viele Primzahlen.*

Beweis. Wir nehmen an, es gäbe r verschiedene Primzahlen $p_1, p_2, \dots, p_r$. Wir konstruieren eine weitere Primzahl. Dazu zerlegen wir die Zahl $p_1 p_2 \dots p_r + 1$ in Primfaktoren

$$p_1 p_2 \dots p_r + 1 = q_1 q_2 \dots q_s.$$

Wir zeigen durch einen Widerspruchsbeweis, dass $p_i \neq q_j$ ist für $1 \leq i \leq r$, $1 \leq j \leq s$. Wäre für ein $i = i_0$ und ein $j = j_0$ $p_{i0} = q_{j0}$, dann wäre $p_{i0} | q_1 \dots q_s$ und damit $p_{i0} | p_1 \dots p_r + 1$. Es würde folgen, dass $p_{i0} | 1$, was nicht sein kann. Daher sind $q_1, \dots, q_s$ weitere Primzahlen, die wir der Liste $p_1, p_2, \dots, p_r$ hinzufügen können. $\square$

Rationale und reelle Zahlen

Wir erinnnern zunächst an die Menge der rationalen Zahlen $\mathbb{Q}$ und führen die Menge der reellen Zahlen $\mathbb{R}$ ein, indem wir diese Mengen beschreiben, aber nicht alle technischen Details ausführen. In der Mathematik führt man die Menge der reellen Zahlen mit Hilfe Dedekindscher Schnitte oder mit Intervallschachtelung ein, die in $\mathbb{Q}$ definiert werden. Eine weitere Möglichkeit wäre, die Menge der reellen Zahlen durch ein Axiomensystem zu beschreiben. Wir gehen hier einen anderen Weg und führen, etwas weniger mathematisch rigoros $\mathbb{Q}$ und $\mathbb{R}$ so ein, wie sie historisch entstanden sind.
Mit Hilfe der Peano-Axiome können in $\mathbb{N}$ eine Addition, Multiplikation und eine Ordnungsrelation eingeführt werden. Die Umkehrung dieser Rechenoperationen, die Subtraktion und Division ist nicht immer möglich. Um diese ausführen zu können, wird die Menge der ganzen Zahlen $\mathbb{Z}$ (1.11) und die Menge der rationalen Zahlen $\mathbb{Q}$ (1.12) eingeführt:

$$\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q}.$$

Die Menge $\mathbb{Q}$ hat folgende Eigenschaften:

1. Die Menge $\mathbb{Q}$ ist total geordnet, d.h. für je zwei verschiedene rationale Zahlen a und b lässt sich angeben, welche die kleinere ist.

2. Die Menge $\mathbb{Q}$ ist in sich überall dicht, d.h. zwischen je zwei verschiedenen rationalen Zahlen a und b liegt mindestens eine weitere rationale Zahl c. Daher liegen zwischen zwei verschiedenen rationalen Zahlen unendlich viele weitere rationalen Zahlen.

3. Die Rechenoperationen Addition, Subtraktion, Multiplikation und Division (außer durch 0) sind in $\mathbb{Q}$ ausführbar.

4. Jede rationale Zahl lässt sich in Form eines endlichen oder unendlichen periodischen Dezimalbruches darstellen.

Die Menge der rationalen Zahlen reicht jedoch nicht aus, um die gesamte Zahlengerade zu füllen. So ist das Potenzieren von rationalen Zahlen $q = \frac{a}{b}, q^n = \left(\frac{a}{b}\right)^n = q \cdot \cdots \cdot q, n \in \mathbb{N}$ in $\mathbb{Q}$ definiert, die inverse Operation, das Wurzelziehen, liefert jedoch nicht immer eine rationale Zahl.

Beispiel

Die Ausdrücke $\sqrt{2}$ und $\sqrt[3]{10}$ sind keine rationalen Zahlen.
Wir führen einen Widerspruchsbeweis um zu zeigen, dass $\sqrt{2}$ keine rationale Zahl ist. Dazu betrachten wir die Aussagen

$$A = \{x = \sqrt{2}\}, \quad B = \{x \text{ ist nicht rational}\}.$$

Wir behaupten, dass die Aussage $\{A \Rightarrow B\}$ wahr ist, d.h. die Aussage $A \wedge \neg B$ ist falsch.
Die Aussage $\neg B$ lautet: x ist rational, genauer: $x = \frac{a}{b}$, wobei a und b teilerfremd sind. Die Aussage

$$A \wedge \neg B = \left(x = \sqrt{2}\right) \wedge \left(x = \frac{a}{b} \text{ (teilerfremd)}\right)$$

liefert

$$\sqrt{2}\sqrt{2} = \frac{a^2}{b^2}, \quad 2b^2 = a^2.$$

Damit ist a^2 eine ganze Zahl und a muss gerade sein. Wir setzen $a = 2k$ und erhalten $2b^2 = 4k^2$ und somit ist $b^2 = 2k^2$ gerade. Es folgt, dass b gerade ist und 2 Teiler von a und b ist. Dies ist ein Widerspruch und $\neg(A \wedge \neg B)$ ist wahr, was äquivalent dazu ist, dass $A \Rightarrow B$ wahr ist.

In der Übungsaufgabe 20 soll gezeigt werden, dass $\sqrt[3]{10}$ keine rationale Zahl ist.

Die Einführung irrationaler Zahlen gestattet es, jedem Punkt der Zahlengeraden eine reelle Zahl zuzuordnen. Dabei ist eine Zahl irrational, wenn sie sich durch einen nichtperiodischen unendlichen Dezimalbruch darstellen lässt.
Wir können jetzt die Menge der reellen Zahlen $\mathbb{R}$ einführen:

$$\mathbb{R} = \mathbb{Q} \cup \{\text{irrationale Zahlen}\}.$$

Die Menge $\mathbb{R}$ ist nicht abzählbar.

Satz 1.45. *Das Intervall* $I = \{x \in \mathbb{R} : 0 < x < 1\}$ *ist überabzählbar und damit ist* $\mathbb{R}$ *überabzählbar.*

Beweis. Wir nehmen an, dass $f : I \to \mathbb{N}$ bijektiv sei, d.h. wir können die Elemente aus I durchnummerieren $x_1, x_2, x_3, \ldots, x_n, \ldots$.
Sei in der Dezimaldarstellung

$$x_n = 0, a_{-1}^{(n)} a_{-2}^{(n)} \ldots a_{-n}^{(n)} \ldots \ . \tag{1.16}$$

Wir setzen

$$b_n = \begin{cases} 0 & \text{falls } a_{-n}^{(n)} \neq 0, \\ 1 & \text{falls } a_{-n}^{(n)} = 0 \end{cases}$$

und betrachten das Element $y = 0, b_1 b_2 b_3 \ldots$. Es ist $y \neq x_1$, da $b_1 \neq a_{-1}^{(1)}$ ist; $y \neq x_2$, da $b_2 \neq a_{-1}^{(2)}$ ist; usw. Das Element $y = 0, b_1 b_2 b_3 \ldots$ gehört also nicht zu der Menge der Elemente x_n, die durch (1.16) beschrieben werden. Dieser Widerspruch besagt, dass I nicht abzählbar ist. □

Haupteigenschaften der reellen Zahlen

1. Die Menge der reellen Zahlen ist geordnet.

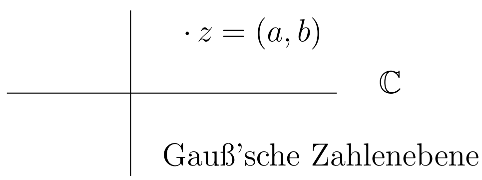
2. Die Menge der rationalen Zahlen $\mathbb{Q}$ ist dicht in $\mathbb{R}$, d.h. sei $r \in \mathbb{R}$, dann gilt: Für alle $\varepsilon > 0$ existiert ein $q = q(\varepsilon) \in \mathbb{Q} : |q - r| < \varepsilon$.

3. Die Menge der reellen Zahlen ist stetig (ein Kontinuum), d.h. jeder Punkt der Zahlengeraden entspricht einer reellen Zahl.

4. Die arithmetischen Operationen (Addition, Subtraktion, Multiplikation, Division – außer durch 0 –) sind mit reellen Zahlen stets ausführbar und ergeben wieder eine reelle Zahl. Das Potenzieren ist in $\mathbb{R}$ stets ausführbar, wobei das Wurzelziehen im Bereich $\mathbb{R}^+ = \{x \in \mathbb{R} : x \geq 0\}$ definiert ist.

Um Wurzeln aus negativen Zahlen zu definieren, muss die Menge der komplexen Zahlen eingeführt werden.

Übersicht, welche Rechenoperationen in den einzelnen Zahlenmengen ausgeführt werden können



$\mathbb{N}$	Addition, Multiplikation
$\mathbb{Z}$	zusätzlich: Subtraktion
$\mathbb{Q}$	zusätzlich: Division
$\mathbb{R}$	zusätzlich: Wurzelziehen $\sqrt[n]{b}, b \geq 0,$
$\mathbb{C}$	zusätzlich: Wurzelziehen $\sqrt[n]{-b}$

Die Menge der komplexen Zahlen $\mathbb{C}$

Die Einführung der komplexen Zahlen ist notwendig, um Nullstellen von Polynomen bestimmen, bzw. Wurzeln aus negativen reellen Zahlen ziehen zu können.

Definition 1.46. Sei $\mathbb{R}$ die Menge der reellen Zahlen. Die Menge der komplexen Zahlen $\mathbb{C}$ entspricht dem kartesischen Produkt $\mathbb{R} \times \mathbb{R} = \mathbb{R}^2$, in der folgende Addition und Multiplikation eingeführt sind:

$$\begin{aligned} z_1 + z_2 &= (a_1 + a_2, b_1 + b_2), \\ z_1 \cdot z_2 &= (a_1, b_1) \cdot (a_2, b_2) = (a_1 a_2 - b_1 b_2, a_1 b_2 + a_2 b_1), \end{aligned}$$

wobei $z_1 = (a_1, b_1) \in \mathbb{R}^2$, $z_2 = (a_2, b_2) \in \mathbb{R}^2$ sind.

Bemerkung 1.47. Die Menge der reellen Zahlen kann als Teilmenge von $\mathbb{C}$ aufgefasst werden. Sei $\mathbb{R}^* = \{z \in \mathbb{C} \,:\, z(a, 0)\}$. Dann ist die Abbildung $f\,:\, \mathbb{R} \to \mathbb{R}^*$ bijektiv, d.h.

$$a \in \mathbb{R} \Leftrightarrow z = (a, 0) := a. \tag{1.17}$$

Wir führen die imaginäre Einheit ein:

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

Es gilt

$$i^2 = (0, 1)(0, 1) = (0 - 1, 0) = (-1, 0) \stackrel{(1.17)}{=} -1$$

und es wird kurz geschrieben $i = \sqrt{-1}$.

Satz 1.48. *Für jede komplexe Zahl $z \in \mathbb{C}$ ist*

$$z = (a, b) = a + ib.$$

Beweis. Wir haben

$$a + ib = (a, 0) + (b, 0)(0, 1) = (a, 0) + (0, b) = (a, b).$$

□

Definition 1.49. Sind a und b reelle Zahlen und ist $z = a + ib$, dann heißt $\bar{z} := a - ib$ die konjugiert komplexe Zahl zu z. Die Zahlen $a = \mathrm{Re}(z)$, $b = \mathrm{Im}(z)$ werden Realteil bzw. Imaginärteil von z genannt.

Komplexe Zahlen können in der Gauß'schen Zahlenebene dargestellt werden.

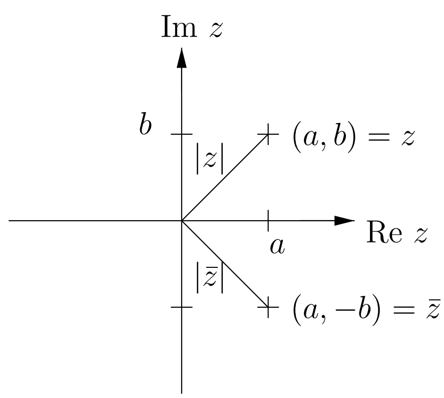


Abbildung 1.10: Gauß'sche Zahlenebene

Satz 1.50. *Sind z und w komplexe Zahlen, dann gilt*

a) $\overline{z + w} = \bar{z} + \bar{w}$,

b) $\overline{z \cdot w} = \bar{z} \cdot \bar{w}$, *c)* $z + \bar{z} = 2\mathrm{Re}(z),\ z - \bar{z} = 2i\mathrm{Im}(z),$

d) $z \cdot \bar{z}$ *ist eine reelle nichtnegative Zahl.*

Der Beweis wird als Übung empfohlen (siehe auch Abbildung 1.10).

Definition 1.51. Der Betrag von $z = a + ib \in \mathbb{C}$ ist definiert als:

$$|z| := (z \cdot \bar{z})^{\frac{1}{2}} = \sqrt{a^2 + b^2}.$$

Es ist $|z| = |\bar{z}|$.

Trigonometrische Darstellung der komplexen Zahlen

Die komplexen Zahlen $z = a + ib$ mit $|z| = 1$ liegen auf dem Einheitskreis.

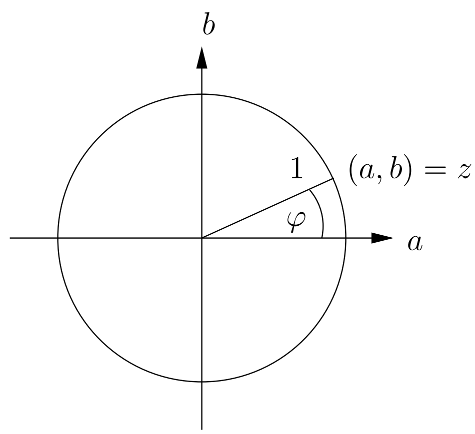


Abbildung 1.11: Einheitskreis

Daher gilt $a = \cos\varphi$ und $b = \sin\varphi$. In diesem Fall ist

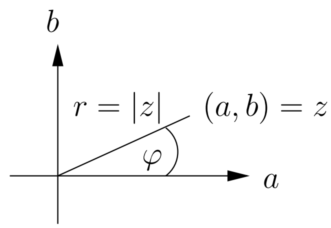
$$z = \cos\varphi + i\,\sin\varphi.$$

Sei nun $z \neq (0,0)$ ein Element aus $\mathbb{C}$. Es ist $|z| \neq 0$. Die komplexe Zahl $w = \frac{z}{|z|}$ besitzt den Betrag $|w| = 1$ und daher ist

$$\begin{aligned} w &= \cos\varphi + i\,\sin\varphi \quad \text{und} \\ z &= |z|\left(\cos\varphi + i\,\sin\varphi\right). \end{aligned} \tag{1.18}$$

Schreibt man $|z| = r$ erhält man die Darstellung

$$z = r\left(\cos\varphi + i\,\sin\varphi\right).$$



Dabei ist $\cos\varphi = \frac{a}{\sqrt{a^2+b^2}}, \sin\varphi = \frac{b}{\sqrt{a^2+b^2}}$.

Beispiele

Es ist

$$z = i = \cos\left(\frac{\pi}{2}\right) + i\sin\left(\frac{\pi}{2}\right),$$
$$z = 1 + i = \sqrt{2}\left(\cos\left(\frac{\pi}{4}\right) + i\sin\left(\frac{\pi}{4}\right)\right).$$

Zahlendarstellung mit g-adischen Brüchen

Wir betrachten eine natürliche Zahl $n \in \mathbb{N}$. Sie lässt sich als Summe von Potenzen der Zahl 10 darstellen

$$n = a_0 \cdot 10^0 + a_1 \cdot 10^1 + a_2 \cdot 10^2 + \cdots + a_N \cdot 10^N, \quad 0 \leq a_i \leq 9.$$

Beispiel

Die Zahl 2003 lässt sich darstellen als

$$2003 = 3 \cdot 10^0 + 0 \cdot 10^1 + 0 \cdot 10^2 + 2 \cdot 10^3.$$

Die Zifferndarstellung besteht darin, die Koeffizienten in umgekehrter Reihenfolge anzuordnen:

$$n = a_N a_{N-1} \dots a_0. \tag{1.19}$$

Etwas schwieriger ist es, eine reelle Zahl als Zehnerpotenzen darzustellen. Wir beschreiben diese Dezimaldarstellung als Algorithmus:
Sei $x \in \mathbb{R}$. Wir finden ein $y \in \mathbb{Z}$ mit

$$y \leq x < y + 1.$$

Falls $y = x$ ist, dann ist der Algorithmus beendet und wir schreiben $x = \pm a_N a_{N-1} \dots a_0$ wie in (1.19). Falls nicht, wird das Intervall $[y, y+1)$ in 10 gleiche Teile zerlegt. Wir betrachten das Teilintervall $I_{a_{-1}}, a_{-1} \in \{0, 1, \dots, 9\}$, das x enthält, d.h.

$$r_1 = y + a_{-1} \cdot 10^{-1} \leq x < y + (a_{-1} + 1)10^{-1} = r_2.$$

Falls $x = y + a_{-1} \cdot 10^{-1}$, dann ist der Algorithmus beendet; falls nicht, wird das Intervall $[r_1, r_2)$ in 10 gleiche Teile zerlegt und das Teilintervall $I_{a_{-2}}, a_{-2} \in \{0, 1, \dots, 9\}$ betrachtet, das x enthält. Dieser Prozess wird so oft wiederholt, bis es entweder ein Intervall gibt, dessen linker Randpunkt gleich x ist (endlicher Dezimalbruch) oder er hat kein endliches Ende (unendlicher Dezimalbruch). Wir schreiben

$$x = y, a_{-1}a_{-2}a_{-3}\cdots = \pm a_N a_{N-1} \dots a_0, a_{-1}a_{-2}a_{-3} \dots \tag{1.20}$$

Es gilt folgende Vereinbarung:
Ist $a_{-i} = 9$ für alle $i : i \geq i_0$ und $a_{-i_0+1} \neq 9$, dann schreiben wir

$$\begin{aligned} x &= y, a_{-1} \dots a_{-i_0+1}999 \dots \\ &= y, a_{-1} \dots (a_{-i_0+1} + 1). \end{aligned} \tag{1.21}$$

Beispiele

Es ist

$$\begin{aligned} 0,9999 \dots &= 1, \\ 3,264999 \dots &= 3.625. \end{aligned}$$

Mit dieser Konvention sind zwei reelle Zahlen gleich, wenn ihre Ziffern gleich sind.

g-adische Darstellung von reellen Zahlen

Neben der Dezimaldarstellung (1.20) von reellen Zahlen, die auf der Darstellung durch Zehner-Potenzen beruht, können auch andere Ziffernbasen verwandt werden.
Sei $g \in \mathbb{N}, g \geq 2$.

Satz 1.52. *Jede natürliche Zahl $n \in \mathbb{N}$ lässt sich eindeutig in der Form*

$$n = n_g = \sum_{i=0}^{N} a_i g^i = a_0 g^0 + a_1 g^1 + \dots a_N g^N, \quad a_N \neq 0, \quad a_i \in \{0, 1, \dots, g-1\},$$

darstellen. Die Ziffern $a_i, i = 0, \dots, N$, können mit Hilfe eines Divisionsalgorithmus berechnet werden. Die Zifferndarstellung lautet

$$n = n_g = a_N a_{N-1} \dots a_0.$$

Beweis. Wir zeigen zunächst die Eindeutigkeit.
Es gelte

$$n = a_0 g^0 + a_1 g^1 + \dots + a_N g^N = b_0 g^0 + b_1 g^1 + \dots + b_M g^M, \quad M \geq N, \tag{1.22}$$

$$a_i, b_j \in \{0, 1, \dots, g-1\},\ i = 0, \dots, N,\ j = 1, \dots, M. \tag{1.23}$$

Dann ist

$$\begin{aligned}a_0 - b_0 &= b_1 g^1 + \cdots + b_M g^M - \left(a_1 g^1 + \cdots + a_N g^N\right)\\ &= g\left[b_1 - a_1 + \cdots + (b_N - a_N)\, g^{N-1} + b_{N+1} g^N + \cdots + b_M g^{M-1}\right]\\ &= gk,\end{aligned}$$

mit $k = b_1 - a_1 + \cdots + (b_N - a_N)\, g^{N-1} + b_{N+1} g^N + \cdots + b_M g^{M-1} \in \mathbb{Z}$.
Wegen (1.23) ist

$$-g < a_0 - b_0 = gk < g$$

und $k = 0$. Es folgt $a_0 = b_0$.
In (1.22) ziehen wir $a_0 = b_0$ auf beiden Seiten ab und dividieren durch g. Es folgt analog, dass $a_1 = b_1$ ist und schließlich dass $a_i = b_i$ für $i = 1, \ldots, N$ und $b_{N+1} = \cdots = b_M = 0$ sind. Durch die Bedingung $a_N \neq 0$ ist N eindeutig bestimmt.

2. Schritt: Existenzbeweis durch Konstruktion

0° Es existiert ein $N \in \mathbb{N}$, so dass

$$g^N \leq n < g^{N+1}$$

ist. Dieses N ist zu bestimmen. Wir führen dazu folgenden Divisionsalgorithmus aus.

1° Teile n durch g^N (Teilen mit Rest):

$$n = a_N g^N + r_N, \quad \text{mit } 0 \leq r_N < g^N, 0 < a_N \leq g - 1.$$

2° Teile r_N durch g^{N-1}:

$$r_N = a_{N-1} g^{N-1} + r_{N-1}, \quad \text{mit } 0 \leq r_{N-1} < g^N, 0 \leq a_{N-1} \leq g - 1.$$

3° Teile r_{N-1} durch g^{N-2}:

$$\begin{aligned}r_{N-1} &= a_{N-2} g^{N-2} + r_{N-2}, \quad \text{mit } 0 \leq r_{N-2} < g^N, 0 \leq a_{N-2} \leq g - 1,\\ &\vdots\end{aligned}$$

Allgemeiner:

Teile r_k durch g^{k-1}:

$$\begin{aligned}r_k &= a_{k-1} g^{k-1} + r_{k-1},\\ &\vdots\\ r_2 &= a_1 g^1 + r_1,\\ r_1 &= a_0, \quad \text{mit } 0 \leq a_0 \leq g - 1.\end{aligned}$$

□

Bemerkung 1.53. Im Schritt 1° ist $a_N > 0$ wegen $a_N > g^N$. Für $i \in \{1, 2, \dots, N-1\}$ ist auch $a_i = 0$ zugelassen.

Beispiele

Die Zahl $n = 2003$ kann dargestellt werden als

$$\begin{aligned} g = 2: \quad 2003_2 &= 1 \cdot 2^0 + 1 \cdot 2^1 + 1 \cdot 2^4 + 1 \cdot 2^6 + 1 \cdot 2^7 + 1 \cdot 2^8 + 1 \cdot 2^9 + 1 \cdot 2^{10} \\ &= 11111010011, \\ g = 4: \quad 2003_4 &= 3 \cdot 4^0 + 1 \cdot 4^2 + 3 \cdot 4^3 + 3 \cdot 4^4 +^{\cdot} 4^5 \\ &= 133103. \end{aligned}$$

Die Darstellung der reellen Zahlen im g-adischen Ziffernsystem folgt dem Algorithmus, der zuvor für $g = 10$ erklärt wurde.
Wir erhalten das Ergebnis: Jede reelle Zahl x lässt sich eindeutig im g-adischen Ziffernsystem darstellen:

$$\begin{aligned} x = x_g &= \pm a_N a_{N-1} \dots a_0, a_{-1} a_{-2} \dots, \quad a_N \neq 0, \\ x &= a_N g^N + a_{N-1} g^{N-1} + \dots + a_0 g^0 + a_{-1} g^{-10} + a_{-2} g^{-2} + \dots \\ &= \sum_{j=-N}^{\infty} a_{-j} g^{-j}. \end{aligned} \tag{1.24}$$

Dabei vereinbaren wir, wie zuvor, dass

$$a_{-j} = g - 1 \text{ für } j \geq j_0$$

verboten ist (vergleiche mit (1.21)) und setzen

$$\begin{aligned} &\pm a_N a_{N-1} \dots a_0, a_{-1} a_{-2} \dots a_{-j_0+1} (g-1)(g-1)(g-1) \dots \\ &= \pm a_N a_{N-1} \dots a_0, a_{-1} a_{-2} \dots (a_{-j_0+1} + 1). \end{aligned}$$

Bemerkung 1.54. Die Babylonier benutzten 1600 v. Chr. ein Zahlensystem mit $g = 60$, die Mayas hatten 665 n. Chr. ein Zahlensystem mit $g = 20$.

Zahlendarstellung im Computer

Die Darstellung der ganzen Zahlen im Computer erfolgt zumeist für $g = 2$, d.h. in einem binären Stellenwertsystem. Es können nur Zahlen einer endlichen Länge L benutzt werden. Daher können nur maximal 2^L verschiedene Zahlen als Bitmuster dargestellt werden. Außerdem muss das Vorzeichen berücksichtigt werden [3, S. 49].

Bitmuster:

$$
\begin{aligned}
\overbrace{000\ldots000}^{L} &\cong 0,\\
000\ldots001 &\cong 1,\\
000\ldots010 &\cong 2,\\
&\vdots\\
111\ldots111 &\cong \sum_{i=0}^{L-1} 2^i = 1+2+2^2+\cdots+2^{L-1} = \frac{1-2^L}{1-2} = 2^L-1.
\end{aligned}
$$

Die Bitmuster können addiert und subtrahiert werden.
Rationale Zahlen können durch Multiplikation der Bitmuster mit dem Faktor 2^{-k} erfasst werden (Festpunkt-Zahlensystem) oder der Skalierungsfaktor wird variabel gehalten (Gleitpunkt-Zahlensystem). Numerische Berechnungen, die eine Vielzahl von Gleitpunktoperationen benötigen, können beträchtliche Fehler aufweisen. Ursache können Rundungsfehler sein, wie sie bei Subtraktion entstehen. Das Assoziativgesetz ist verletzt im Gleitkommasystem (festes L).

Beispiel (Dezimalsystem , siehe [3, S. 53])

$$
\begin{aligned}
x &= 0.23371258 \cdot 10^{-4},\\
y &= 0.33678429 \cdot 10^{2},\\
z &= -0.33677811 \cdot 10^{2}.
\end{aligned}
$$

Das exakte Ergebnis lautet $x + y + z = 0.641371258 \cdot 10^{-3}$.
Durch Rundung erhalten wir

$$
\begin{aligned}
x + (y + z) &= 0.23371258 \cdot 10^{-4}\\
&\quad + 0.00000618 \cdot 10^{2}\\
&\approx 0.02337126 \cdot 10^{-3}\\
&\quad + 0.618 \qquad\quad \cdot 10^{-3}\\
&= \underline{0.64137126 \cdot 10^{-3}},\\
(x + y) + z &\approx 0.00000023 \cdot 10^{2}\\
&\quad + 0.33678429 \cdot 10^{2}\\
&\quad + z\\
&= 0.33678452 \cdot 10^{2}\\
&\quad - 0.33677811 \cdot 10^{2}\\
&= 0.00000641 \cdot 10^{2} = \underline{0.641 \cdot 10^{-3}}.
\end{aligned}
$$

Die Kunst der Numerik besteht darin, diese Fehler klein zu halten.

Übungsaufgaben

13) Beweisen Sie durch vollständige Induktion für alle $n \in \mathbb{N}$

$$\sum_{k=1}^{n} k^2 = \frac{1}{3} n(n+1)(n+\frac{1}{2}) .$$

14) Pascalsches Dreieck und Binomischer Satz.
Für ganze Zahlen $n \geq k \geq 0$ definieren wir den **Binomialkoeffizienten**

$$\binom{n}{k} := \frac{n!}{(n-k)!k!} = \frac{n(n-1)(n-2)\ldots(n-k+1)}{k(k-1)(k-2)\ldots 2\cdot 1} ,$$

(lies: n über k) wobei $n! = 1 \cdot 2 \cdots (n-1)n$ und $0! = 1$ ist.

1. Man betrachte das Pascalsche Dreieck

$$\begin{array}{ccccccccccc} &&&&& 1 &&&&& \\ &&&& 1 && 1 &&&& \\ &&& 1 && 2 && 1 &&& \\ && 1 && 3 && 3 && 1 && \\ & 1 && 4 && 6 && 4 && 1 & \\ 1 && 5 && 10 && 10 && 5 && 1 \end{array}$$

   Das Dreieck hat am Rand Einsen und die inneren Zahlen sind die Summe der beiden darüber liegenden Zahlen. Man zeige durch vollständige Induktion nach n, dass die k-te Zahl der n-ten Zeile durch den Binomialkoeffizienten $\binom{n}{k}$ gegeben ist, wobei die Nummerierung von n und k bei 0 startet.

2. Man zeige für $a, b \in \mathbb{R}$ den **Binomischen Satz**

$$(a+b)^n = \sum_{i=0}^{n} \binom{n}{i} a^{n-i} b^i .$$

15) Größter gemeinsame Teiler.

1. Bestimmen Sie $\mathrm{ggT}(a,b)$ von den Paaren $a = 8942, b = 2312$ und $a = 7543, b = 2344$ durch Ausführen des Euklidischen Algorithmus auf dem Papier.

2. Bestimmen Sie $\mathrm{ggT}(a,b)$ von den Paaren $a = 1234567, b = 2345678$ und $a = 123456789, b = 896745321$ wahlweise auf Papier oder mit Hilfe eines (selbst geschriebenen) Programms. Das Programm ist ggf. mit anzugeben.

3. Finden Sie eine Folge a_n, $n \geq 1$, von ganzen Zahlen, so dass der Euklidische Algorithmus genau n Schritte zur Berechnung des $\mathrm{ggT}(a_{n+1}, a_n)$ braucht. Hinweis: Man starte mit einer beliebigen Zahl a_1 und finde mit Hilfe des Algorithmus eine passende Zahl a_2.

16) Seien $a, b, c, d, x, y \in \mathbb{N}$ und p eine Primzahl. Beweisen Sie folgende Aussagen:

1. Zeigen Sie $\mathrm{ggT}(ab, cb) = b\,\mathrm{ggT}(a,c)$. Als unmittelbare Folgerung halten wir

$$d|x \wedge d|y \Rightarrow d|\mathrm{ggT}(x,y)$$

   fest. Tipp: Ein eleganter Beweis hierfür benutzt den Euklidischen Algorithmus.

2. Bestimmen Sie $\mathrm{ggT}(ab, pb)$, wenn p kein Teiler von a ist.

3. Zeigen Sie
$$p|ab \Rightarrow p|a \vee p|b\,.$$
(Teilt eine Primzahl ein Produkt, dann auch einen der Faktoren.)

4. **(Fundamentalsatz der Arithmetik)** Jede natürliche Zahl $n \geq 2$ lässt sich bis auf die Reihenfolge der Faktoren auf genau eine Weise als Produkt von Primzahlen schreiben.

17) Zeigen Sie, dass die Dezimaldarstellung jeder rationalen Zahl periodisch ist, und dass jede Zahl mit einer periodischen Dezimaldarstellung rational ist.

18) Zahlensysteme. Die in der Informatik gebräuchlichsten Zahlensysteme sind: Binärsystem (g=2), Oktalsystem (g=8) und Hexadezimalsystem (g=16). Um in letzterem die Zahlen 10, 11, 12, 13, 14 und 15 darzustellen, verwendet man die ersten Buchstaben des Alphabets $\mathrm{A}, \mathrm{B}, \mathrm{C}, \mathrm{D}, \mathrm{E}$ und F.

1. Wandeln Sie die Dezimalzahl 42_{10} in das Binär-, Oktal- und Hexadezimalsystem um.

2. Wandeln Sie die Hexadezimalzahl $1\mathrm{FCE}_{16}$ in das Binär-, Oktal- und Dezimalsystem um.

3. Wandeln Sie die Oktalzahl 7353_8 in das Binär-, Dezimal- und Hexadezimalsystem um.

19) Babylonisches Wurzelziehen.
Wir untersuchen einen Algorithmus zum Berechnen der Quadratwurzel einer Zahl $x \geq 1$, der schon in den Gesetzestafeln des Hammurabi im Jahre 1950 v.Chr stand.
Wir definieren rekursiv eine Folge reeller Zahlen w_i mit
$$w_0 = x \quad w_{n+1} = \frac{1}{2}\left(w_n + \frac{x}{w_n}\right).$$

1. Zeigen Sie, dass für alle $n \in \mathbb{N}$ die Zahlen w_n positiv sind und
$$w_n^2 \geq w_{n+1}^2 \geq x$$
gilt, also jedes w_{n+1} die Wurzel mindestens so gut annähert wie w_n.

2. Berechnen Sie $\sqrt{3}$ auf 4 Stellen hinter dem Komma durch Verwendung des obigen Algorithmus, d.h. es ist ein w_n mit $w_n - \sqrt{3} < 0.5 \cdot 10^{-4}$ zu berechnen. Zur Abschätzung der Genauigkeit kann man folgende Ungleichung verwenden:
$$w_n - \sqrt{3} = \frac{w_n^2 - 3}{w_n + \sqrt{3}} \leq \frac{w_n^2 - 3}{w_n + 1}\,,$$
da $1 \leq \sqrt{3}$.

20) Zeigen Sie, dass folgende Zahlen irrational sind

1. $\sqrt[3]{10}$,

2. $\sqrt[5]{30}$.

21) Berechnen Sie das Produkt wz der komplexen Zahlen

$$w = |w|(\cos\eta + \mathrm{i}\,\sin\eta)\,,\quad z = |z|(\cos\varphi + \mathrm{i}\,\sin\varphi)\,.$$

Fertigen Sie eine Skizze an und deuten Sie die Verknüpfung geometrisch. Berechnen Sie w/z für $z \neq (0,0)$.

22) g-adische Darstellung.
Berechnen Sie die g-adische Darstellung der Zahlen

1. $\left(\frac{1}{3}\right)_{10}$ für $g = 2$, $g = 3$ und $g = 10$.
2. $\left(\frac{1}{5}\right)_{10}$ für $g = 2$, $g = 8$ und $g = 9$.

23) Berechnen Sie die folgenden Ausdrücke im Binärsystem auf 5 Stellen (in jedem Schritt jede Zahl nach der 5. Stelle abbrechen, nicht runden) genau.

1. $\frac{1}{3} + \frac{1}{5}$,
2. $\frac{8}{15} - (\frac{1}{3} + \frac{1}{5})$.

Ermitteln Sie in beiden Fällen den relativen Fehler $\frac{\tilde{x}}{x}$, wobei x das exakte und $\tilde{x}$ das berechnete Ergebnis ist.

1.4 Gruppen, Ringe, Körper

Die Menge der komplexen Zahlen wurde eingeführt, indem wir im $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ eine Addition und eine Multiplikation von zwei Elementen definiert hatten. Die Addition und Multiplikation sind Beispiele für Verknüpfungen, die folgendermaßen definiert werden.

Definition 1.55. Sei M eine Menge. Eine Abbildung

$$f : M \times M \to M, \quad (m_1, m_2) \to m = m_1 * m_2$$

heißt Verknüpfung. Eine Verknüpfung heißt assoziativ, wenn gilt

$$(m_1 * m_2) * m_3 = m_1 * (m_2 * m_3) \quad \forall\, m_1, m_2, m_3 \in M.$$

Sie heißt kommutativ, wenn

$$m_1 * m_2 = m_2 * m_1 \quad \forall\, m_1, m_2 \in M$$

gilt. Ein Element $e \in M$ heißt neutral, falls gilt

$$e * m = m * e = m \quad \forall\, m \in M.$$

Beispiele

1° Sei $M = \mathcal{P}(X)$ die Potenzmenge einer Menge X. Die Abbildungen „Vereinigung" und „Durchschnitt"

$$\begin{aligned} &f : M \times M \to M, \quad && g : M \times M \to M, \\ &f : U \cup V = W, && g : U \cap V = W, \quad U, V \subset M \end{aligned}$$

sind kommutative und assoziative Verknüpfungen.
Das neutrale Element für f ist die leere Menge $\emptyset$; das neutrale Element für g ist die Menge X.

2° Wir betrachten die Menge der rationalen Zahlen $\mathbb{Q}$ und die Abbildungen Addition und Multiplikation:

$$\begin{aligned} &f : \mathbb{Q} \times \mathbb{Q} \to \mathbb{Q}, \quad && g : \mathbb{Q} \times \mathbb{Q} \to \mathbb{Q}, \\ &f(a,b) = a + b, && g(a,b) = a \cdot b. \end{aligned}$$

Das neutrale Element für die Addition ist die Null, für die Multiplikation die Eins.

Definition 1.56 (Halbgruppe, Gruppe). Eine Menge G mit einer assoziativen Verknüpfung „$*$" heißt Halbgruppe, geschrieben als $(G, *)$.
Existiert in einer Halbgruppe ein neutrales Element e und gibt es zu jedem $g \in G$ ein inverses Element g^{-1}, so dass

$$g * g^{-1} = g^{-1} * g = e$$

ist, dann ist $(G, *)$ eine Gruppe.
Ist die Verknüpfung kommutativ, dann heißt $(G, *)$ abelsche Gruppe. Hat eine Gruppe $(G, *)$ endlich viele Elemente, so sprechen wir von einer endlichen Gruppe.

Beispiele

1° Die natürlichen Zahlen bilden eine Halbgruppe mit der Multiplikation oder der Addition.

2° $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$ und $\mathbb{C}$ sind Gruppen bezüglich der Addition; $\mathbb{Q}^+ = \{x \in \mathbb{Q} : x > 0\}$ ist eine Gruppe bezüglich der Multiplikation.

3° Die endliche Menge $G = \{-1, +1\}$ ist ebenfalls eine Gruppe bezüglich der Multiplikation in $\mathbb{Z}$.

4° **Permutationsgruppen**
Etwas ausführlicher betrachten wir Permutationsgruppen. Es sei M eine Menge und $S(M)$ eine Menge von Abbildungen von M in M:

$$S(M) = \{f : M \leftrightarrow M, f \text{ bijektiv}\}.$$

Die Verknüpfung in $S(M)$ sei die Komposition

$$f * g = f \circ g.$$

Das neutrale Element ist die identische Abbildung e

$$\begin{aligned} e &: M \to M, \\ e &: m \to m, \quad \forall\, m \in M. \end{aligned}$$

Da f bijektiv ist, existiert die inverse Abbildung $f \circ f^{-1} = f^{-1} \circ f = e$. Daher ist $(S(M), \circ)$ eine Gruppe.
Als Spezialfall betrachten wir: $S_n = S(\{1, 2, 3, \ldots, n\})$. S_n heißt auch symmetrische Gruppe auf n Elementen. Diese Gruppe hat $n!$ Elemente, die Permutationen genannt werden. Dieser Beweis ist als Übung zu erbringen (siehe Übungsaufgabe 25).

Beispiel

Wir betrachten

$$S_3 = S(\{1, 2, 3\}).$$

Die möglichen Permutationen können folgendermaßen beschrieben werden:

$$\begin{pmatrix} 1 & 2 & 3 \\ f(1) & f(2) & f(3) \end{pmatrix}.$$

Es ergeben sich $3! = 6$ Permutationen:

$$f_1 = e \hat{=} \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \quad f_2 \hat{=} \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \quad f_3 \hat{=} \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix},$$

$$f_4 \hat{=} \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \quad f_5 \hat{=} \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}, \quad f_6 \hat{=} \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}.$$

5° **Gruppe der Restklassen**
Wir hatten in der Menge der ganzen Zahlen eine Äquivalenzrelation eingeführt (kongruent modulo m, $m \in \mathbb{N}$)

$$a \sim b \;\Leftrightarrow\; a \equiv b \text{ modulo } m \;\Leftrightarrow\; a - b = qm, \quad q \in \mathbb{Z}.$$

Dadurch erhalten wir eine Partition von $\mathbb{Z}$, deren m Elemente die Äquivalenzklassen sind.
Die Äquivalenzklasse von $a \in \mathbb{Z}$ wird auch Restklasse von a modulo m genannt

$$\begin{aligned} R[a] &= \{b \in \mathbb{Z} \,:\, b \sim a\}, \\ &= a + m\mathbb{Z}. \end{aligned}$$

Das Element a heißt Vertreter bzw. Repräsentant der Restklasse. Die Menge aller Restklassen $\{R[0], R[1], \dots R[m-1]\}$ wird mit

$$\mathbb{Z}/m\mathbb{Z}$$

bezeichnet. Auf $\mathbb{Z}/m\mathbb{Z}$ kann eine Addition $\oplus$ und eine Multiplikation $\odot$ definiert werden:

$$\begin{aligned} R[a] \oplus R[b] &= R[a+b], \\ R[a] \odot R[b] &= R[ab]. \end{aligned}$$

Beispiel

Es sei $m = 2$, dann gibt es zwei Restklassen:

$$\begin{aligned} R[0] &= \{b \in \mathbb{Z} : 0 - b = 2q,\ q \in \mathbb{Z}\} \\ &= \{\dots, -4, -2, 0, 2, 4, 6, \dots\} \quad \text{ganze gerade Zahlen,} \\ R[1] &= \{b \in \mathbb{Z} : 1 - b = 2q,\ q \in \mathbb{Z}\} \\ &= \{\dots, -3, -1, 0, 1, 3, 5, \dots\} \quad \text{ganze ungerade Zahlen.} \end{aligned}$$

Es ist $e = R[0]$ das neutrale Element bezüglich der Addition und $e = R[1]$ das neutrale Element bezüglich der Multiplikation. Insbesondere haben wir:

$$\begin{array}{ll} R[0] \oplus R[1] = R[1], & \text{d.h. gerade Zahl + ungerade Zahl = ungerade Zahl,} \\ R[0] \oplus R[0] = R[0], & \text{d.h. gerade Zahl + gerade Zahl = gerade Zahl,} \\ R[1] \oplus R[1] = R[2] = R[0], & \text{d.h. ungerade Zahl + ungerade Zahl = gerade Zahl,} \\ R[0] \odot R[0] = R[0], & \text{d.h. gerade Zahl} \cdot \text{gerade Zahl = gerade Zahl,} \\ R[0] \odot R[1] = R[0], & \text{d.h. gerade Zahl} \cdot \text{ungerade Zahl = gerade Zahl,} \\ R[1] \odot R[1] = R[1], & \text{d.h. ungerade Zahl} \cdot \text{ungerade Zahl = ungerade Zahl.} \end{array}$$

Die Menge $(\mathbb{Z}/m\mathbb{Z}, \oplus)$ ist eine Gruppe; die Menge $(\mathbb{Z}/m\mathbb{Z}, \odot)$ ist eine Gruppe falls m eine Primzahl ist.

Definition 1.57 (Ring). Eine Menge mit zwei Verknüpfungen „+ “ und „· “ und den Eigenschaften:

1. $(M, +)$ ist abelsche Gruppe mit dem neutralen Element e_+,
2. die Verknüpfung „· “ ist assoziativ, es gibt ein neutrales Element $e_\cdot$,
3. $(a + b) \cdot c = a \cdot c + b \cdot c$, $c \cdot (a + b) = c \cdot a + c \cdot b$ für alle $a, b, c \in M$,
4. $e_+ \neq e_\cdot$.

heißt Ring.

Beispiele für Ringe sind: $\mathbb{Z}$ und $\mathbb{Z}/p\mathbb{Z}$, wobei p eine Primzahl ist.

Definition 1.58 (Körper). Sei M ein Ring. Ist M kommutativ bezüglich der Multiplikation und $M\backslash\{e_+\}$ eine Gruppe unter der Verknüpfung „$\cdot$“, dann heißt M Körper.

Beispiele

$\mathbb{Q}, \mathbb{R}$ und $\mathbb{C}$ sind Körper.

Übungsaufgaben

24) Sei M eine Menge und $\mathcal{R}$ die Menge der Relationen auf M. Welche Eigenschaften einer Gruppe hat $(\mathcal{R}, \circ)$ und welche nicht?

25) Zeigen Sie, dass die Mächtigkeit der Gruppe $S_n = S(\{1, 2, \ldots, n\})$ der Permutationen von n Elementen gleich $n!$ ist.

26) Eine Permutation $\tau \in S_n$ heißt **Transposition**, falls τ zwei Elemente vertauscht und alle anderen fest lässt, d.h. es gibt $l, k \in \{1, 2, \ldots, n\}$ mit $l \neq k$ und

$$\tau(k) = l\,, \quad \tau(l) = k\,, \quad \tau(i) = i \;\forall\, i \in \{1, 2, \ldots, n\} \setminus \{l, k\}\,.$$

Zeigen Sie, dass

1. sich jede beliebige Permutation $\sigma \in S_n$ als Komposition von Transpositionen $\sigma = \tau_r \circ \tau_{r-1} \circ \cdots \circ \tau_1$ schreiben lässt.
2. sich jede Transposition τ als Komposition von Transpositionen zweier benachbarter Elemente i und $i+1$ schreiben lässt.

27) Sei $\sigma \in S_n$. Man betrachte die Paare (i, j) mit $1 \leq i < j \leq n$. Ein Paar (i, j) heißt **Fehlstand**, wenn $\sigma(i) > \sigma(j)$ ist. Wir definieren das **Signum** (also Vorzeichen) der Permutation durch

$$\operatorname{sign} \sigma = \begin{cases} +1 & \sigma \text{ hat eine gerade Anzahl von Fehlständen,} \\ -1 & \sigma \text{ hat eine ungerade Anzahl von Fehlständen.} \end{cases}$$

Zeigen Sie:

1. Für jede Transposition $\tau \in S_n$ gilt $\operatorname{sign} \tau = -1$.
2. Für zwei beliebige Permutationen σ_1 und σ_2 gilt

   $$\operatorname{sign} \sigma_2 \circ \sigma_1 = (\operatorname{sign} \sigma_2)(\operatorname{sign} \sigma_1)\,.$$

   Insbesondere ist $\operatorname{sign}\left(\sigma_1^{-1}\right) = \operatorname{sign} \sigma_1$. Falls sich σ_1 als Produkt von r Transpositionen $\sigma_1 = \tau_r \circ \tau_{r-1} \circ \cdots \circ \tau_1$ schreiben lässt, ist $\operatorname{sign} \sigma_1 = (-1)^r$.

Kapitel 2

Lineare Algebra

In der linearen Algebra werden lineare Räume (Vektorräume) beschrieben und lineare Abbildungen in diesen Räumen untersucht. Vektoren werden verwendet, um die Lage und Bewegung von Punkten in der Ebene bzw. in drei oder n-dimensionalen Räumen zu beschreiben. Sie treten in vielen Anwendungen, z.B. in der Computer-Graphik, Computer-Tomographie, in der Strömungs- und Festkörpermechanik (Kräfte, Geschwindigkeiten, Felder) in der Elektromechanik, Meteorologie, usw. auf.
In diesem Kapitel werden Vektorräume eingeführt, lineare Abbildungen zwischen verschiedenen Vektorräumen beschrieben und in Form linearer Gleichungssysteme untersucht. Dazu werden grundlegende Begriffe wie lineare Unabhängigkeit, Basis, Dimension, Matrix und Determinante eingeführt. Weiterhin werden Eigenwerte, Eigenvektoren und Diagonalisierbarkeit von Matrizen behandelt, sowie Vektorräume mit Skalarprodukt untersucht. Als Literatur wird empfohlen [3, Kapitel 5], [4], [5, Kapitel 6-10], [8, 2. Teil].

2.1 Vektorräume

Wir erinnern zunächst an das kartesische Produkt $\mathbb{R}^2 = \mathbb{R}\times\mathbb{R}$, bzw. $\mathbb{R}^n = \underbrace{\mathbb{R}\times\mathbb{R}\times\cdots\times\mathbb{R}}_{n-\text{mal}}$ als Mengen geordneter Paare (x_1, x_2) bzw. geordneter n-Tupel $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$, $x_i \in \mathbb{R}$, $i = 1, 2, \ldots, n$. In $\mathbb{R}^n$ kann eine Addition definiert werden:

$$\boldsymbol{x} + \boldsymbol{y} = (x_1, x_2, \ldots, x_n) + (y_1, y_2, \ldots, y_n) = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n),$$

die $(\mathbb{R}^n, +)$ zu einer kommutativen Gruppe macht. Das neutrale Element ist $\mathbf{0} = (0, 0, \ldots, 0)$, das inverse Element zu $\boldsymbol{x}$ ist $-\boldsymbol{x} = (-x_1, -x_2, \ldots, -x_n)$.
Weiterhin kann die Multiplikation auf $\mathbb{R}^n$ mit „skalaren“ Elementen $\alpha \in \mathbb{R}$ eingeführt werden

$$\alpha\boldsymbol{x} = \alpha(x_1, x_2, \ldots, x_n) := (\alpha x_1, \ldots, \alpha x_n).$$

Es gilt für $\alpha, \beta \in \mathbb{R}$:

$$\begin{aligned} \alpha(\boldsymbol{x}+\boldsymbol{y}) &= \alpha\boldsymbol{x}+\alpha\boldsymbol{y}, \\ (\alpha\beta)\boldsymbol{x} &= \alpha(\beta\boldsymbol{x}), \\ (\alpha+\beta)\boldsymbol{x} &= \alpha\boldsymbol{x}+\beta\boldsymbol{x}. \end{aligned}$$

Diese Verknüpfungen (Addition in $\mathbb{R}^n$ und Multiplikation auf $\mathbb{R}^n$ mit skalaren Elementen aus einem Zahlenkörper) motivieren die Definition eines Vektorraums, der auch linearer Raum genannt wird.

Definition 2.1. Sei K ein Körper, dessen neutrales Element bezüglich der Multiplikation mit 1 bezeichnet wird. V sei eine Menge mit einer Verknüpfung „+“ $: V \times V \to V$, so dass $(V, +)$ eine kommutative Gruppe ist.
Die Verbindung von K und X sei durch eine Multiplikation

$$K \times V \to V, \quad (\alpha, x) \to \alpha x$$

beschrieben, wobei für alle $\alpha, \beta \in K$ und für alle x, y die folgenden Eigenschaften gelten:

$$(\alpha\beta)x = \alpha(\beta x), \tag{2.1}$$

$$1x = x, \tag{2.2}$$

$$(\alpha+\beta)x = \alpha x + \beta x, \tag{2.3}$$

$$\alpha(x+y) = \alpha x + \alpha y. \tag{2.4}$$

Dann heißt V Vektorraum bzw. linearer Raum über dem Körper K, kurz K-Vektorraum geschrieben. Die Elemente von V heißen Vektoren. Ist $K = \mathbb{R}$, bzw. $K = \mathbb{C}$, dann spricht man von reellen oder komplexen Vektorräumen.

Beispiele

1° $\mathbb{R}^n$ mit $K = \mathbb{R}$, $\mathbb{Q}^n$ mit $K = \mathbb{Q}$ und $\mathbb{C}^n$ mit $K = \mathbb{C}$ sind Vektorräume, ebenso $\mathbb{C}^n$ mit $K = \mathbb{R}$.

2° Sei $M = \{f : \mathbb{R} \to \mathbb{R}\}$ die Menge aller Abbildungen (Funktionen) von $\mathbb{R}$ in $\mathbb{R}$. Durch die Verknüpfungen

$$\begin{aligned} (f+g)(x) &:= f(x)+g(x) \quad && \forall\, f, g \in M, \forall\, x \in \mathbb{R}, \\ (\alpha f)(x) &:= \alpha f(x) && \forall\, f \in M, \forall\, \alpha, x \in \mathbb{R} \end{aligned}$$

wird die Menge M zu einem reellen Vektorraum.

Definition 2.2. Sei V ein K-Vektorraum. Eine nichtleere Teilmenge $V_0 \subseteq V$ heißt Teilvektorraum oder Untervektorraum falls gilt:

1. Sind x, y Elemente aus V_0, dann ist auch $x+y \in V_0$, d.h. V_0 ist bezüglich der Addition in V abgeschlossen.

2. Ist $x \in V_0$, dann ist auch $\alpha x \in V_0$ für alle $\alpha \in K$, d.h. V_0 ist unter der Multiplikation mit Skalaren abgeschlossen, insbesondere gilt $0 \in V_0$.

Beispiel

$V = \mathbb{R}^2$ ist ein $\mathbb{R}$-Vektorraum, $M = \mathbb{R} \times \{0\} \subset \mathbb{R} \times \mathbb{R}$ ist Teilvektorraum, d.h.

$$\mathbb{R} = \{(a, 0) \,:\, a \in \mathbb{R}\} \subset \{(a, b) \,:\, a \in \mathbb{R}, b \in \mathbb{R}\}.$$

Lineare Abhängigkeit, Basen

Im Vektorraum V waren die Addition und die Multiplikation mit „Skalaren“ aus dem Körper K definiert. Diese Verknüpfungen können nicht nur auf ein Paar von Elementen angewandt werden, sondern auf endlich viele Elemente in Form einer Linearkombination übertragen werden.

Definition 2.3. Sei V ein Vektorraum über dem Körper K. $M \subset V$ sei eine Teilmenge. Eine Element $x \in V$ heißt Linearkombination von Elementen aus M, falls es eine endliche Menge $M_x \subset M$ gibt, so dass x in der Gestalt

$$x = \sum_{m \in M_x} \alpha_m m$$

mit $\alpha_m \in K$, darstellbar ist.
$\text{Span}(M) = \{x \in V \,:\, x\,\text{ist Linearkombination von Elementen aus}\,M\}$ heißt der von M aufgespannte Teilvektorraum von V. Span(M) wird auch lineare Hülle von M genannt.

Wir überzeugen uns davon, dass Span(M) ein Vektorraum über K ist. Dazu zeigen wir zunächst folgendes Lemma.

Lemma 2.4. *Sei V ein Vektorraum über dem Körper K, $\alpha \in K$ und $x \in V$. Dann ist $\alpha x = 0$ genau dann, wenn $\alpha = 0$ oder $x = 0$ ist.*

Beweis.

a) Wir nehmen an, dass $\alpha = 0$ oder $x = 0$ ist. Aus $\alpha = 0$ folgt für ein $\beta \in K$ $\beta x = (\beta + 0)x = \beta x + 0x$, also ist $0x = 0$ das neutrale Element. Aus $x = 0$ folgt, $\alpha y = \alpha(0 + y) = \alpha \cdot 0 + \alpha \cdot y$. Damit ist $\alpha 0 = 0$.

b) Es sei $\alpha x = 0$ und $\alpha \neq 0$. Dann ist $x = 1 \cdot x = (\alpha^{-1}\alpha)x = \alpha^{-1}(\alpha x) = 0$. Es sei $\alpha x = 0$ und $x \neq 0$. Wäre $\alpha \neq 0$, dann existiert α^{-1} mit $\alpha^{-1}(\alpha x) = \alpha^{-1}0 = 0 = x$, was zu einem Widerspruch führt.

□

Lemma 2.5. *Span(M) ist ein K-Teilvektorraum von V.*

Beweis. Die Addition, die in V erklärt ist, gilt auch für Elemente $(x, y) \in \mathrm{Span}(M) \times \mathrm{Span}(M)$:

$$\begin{aligned} x + y &= \sum_{m \in M_x} \alpha_m m + \sum_{m \in M_y} \alpha_m m \\ &= \sum_{m \in M_x \cup M_y} \alpha_m m. \end{aligned}$$

Die Addition ist kommutativ.
Das neutrale Element

$$0 = 0m, \quad m \in M$$

befindet sich in $\mathrm{Span}(M)$ nach Lemma 2.4.
Ist $x \in \mathrm{Span}(M)$, d.h. $x = \sum_{m \in M_x} \alpha_m m$, dann ist das inverse Element bezüglich der Addition

$$x^{-1} = \sum_{m \in M_x} (-\alpha_m) m$$

ebenfalls in $\mathrm{Span}(M)$.
Die Eigenschaften (2.1)-(2.4) gelten. □

Beispiele

1° Wir betrachten den Vektorraum $\mathbb{R}^3 = \mathbb{R} \times \mathbb{R} \times \mathbb{R}$. Die Menge M besteht aus 2 Dreiertupeln

$$M = \{(1, 0, 0), (0, 1, 0)\}.$$

Dann ist

$$\mathrm{Span}(M) = \{\boldsymbol{x} = (x_1, x_2, x_3) \,:\, \boldsymbol{x} = \alpha(1, 0, 0) + \beta(0, 1, 0) = (\alpha, \beta, 0),\ \alpha, \beta \in \mathbb{R}\},$$

d.h. $\mathrm{Span}(M) \subset \mathbb{R}^3$ beschreibt eine Ebene, die durch die Vektoren (3er-Tupel) $(1, 0, 0)$, $(0, 1, 0)$ aufgespannt wird, siehe Abbildung 2.1.

2° Wir betrachten wieder den Vektorraum $\mathbb{R}^3$ und die Menge

$$M = \{(1, 0, 0), (-1, 0, 0)\}.$$

Dann ist

$$\mathrm{Span}(M) = \{\boldsymbol{x} = (x_1, x_2, x_3) \in \mathbb{R}^3 \,:\, \boldsymbol{x} = (\alpha, 0, 0),\ \alpha \in \mathbb{R}\},$$

d.h. $\mathrm{Span}(M)$ stimmt mit der x_1-Achse überein, siehe Abbildung 2.2.

Im ersten Beispiel waren die Elemente von M linear unabhängig, im 2. Beispiel linear abhängig. Die lineare Unabhängigkeit von Elementen einer Menge M wird wie folgt definiert.

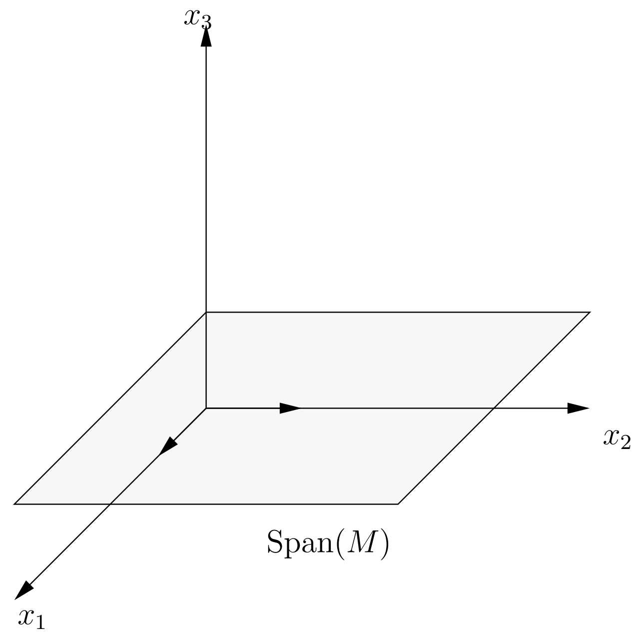


Abbildung 2.1: $x_1 - x_2$-Ebene

Definition 2.6. Sei V ein K-Vektorraum. Eine Menge $M \subset V$ besteht aus linear unabhängigen Elementen, falls für jede endliche nichtleere Teilmenge $\tilde{M} \subset M$ gilt:
Aus

$$\sum_{m \in \tilde{M}} \alpha_m m = 0, \quad \alpha_m \in K \tag{2.5}$$

folgt $\alpha_m = 0$ für alle $m \in \tilde{M}$.
Ist M eine endliche Menge, dann braucht in (2.5) nur $\tilde{M} = M$ betrachtet werden.

Wir sehen uns die vorherigen Beispiele an.

1° Es sei

$$M = \{(1,0,0),(0,1,0)\}.$$

Wir betrachten die Linearkombination (2.5)

$$\sum_{m \in M} \alpha_m m = (\alpha_1, \alpha_2, 0) = \mathbf{0} = (0,0,0).$$

Es folgt $\alpha_1 = \alpha_2 = 0$.

2° Es sei

$$M = \{(1,0,0),(-1,0,0)\}.$$

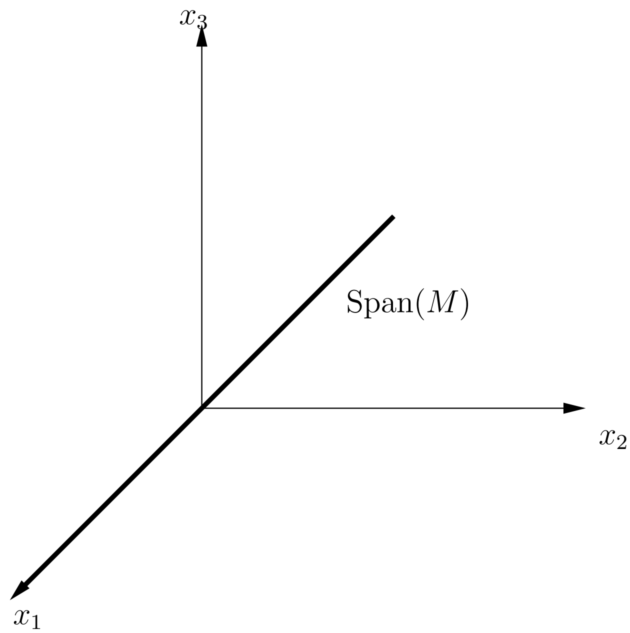


Abbildung 2.2: Span(M)

Die Linearkombination (2.5) lautet

$$\alpha(1,0,0) + \beta(-1,0,0) = (\alpha - \beta, 0, 0) = (0,0,0),$$

woraus $\alpha = \beta$ folgt.

Definition 2.7. Eine Menge $B \subset V$ heißt Basis des Vektorraums V, falls B aus linear unabhängigen Elementen besteht und

$$\text{Span}(B) = V$$

ist.

Beispiel

1° Im $\mathbb{R}^n$ bildet die Menge der Vektoren

$$B = \{(1,0,\ldots,0), (0,1,0,\ldots,0), \ldots, (0,0,\ldots,1)\}$$

eine Basis, aus Einheitsvektoren, die sehr häufig betrachtet und im Folgenden Standardbasis genannt wird.

2° Im $\mathbb{R}^2$ bilden die Vektoren

$$B = \{(a,b), (-b,a)\} \quad (a,b) \neq (0,0)$$

eine Basis.
Wir zeigen, dass B aus linear unabhängigen Vektoren besteht. Sei

$$\alpha(a,b)+\beta(-b,a)=(\alpha a-\beta b,\alpha b+\beta a)=(0,0).$$

Dann muss

$$\alpha a-\beta b=0, \tag{2.6}$$
$$\alpha b+\beta a=0 \tag{2.7}$$

sein. Es folgt aus (2.6), dass $\alpha=\frac{\beta b}{a}$ ist. Durch Einsetzen von α in (2.7) erhalten wir $\beta(b^2+a^2)=0$, woraus $\beta=0$ und schließlich $\alpha=\frac{\beta b}{a}=0$ folgt.
Weiterhin enthält Span(B) die Einheitsvektoren $(1,0)$ und $(0,1)$, durch die der ganze Raum $\mathbb{R}^2$ aufgespannt wird.

Satz 2.8. *Sei $B\subset V$ eine Basis des K-Vektorraums V. Jedes $x\in V$ kann eindeutig als*

$$x=\sum_{e\in B}\alpha_e e$$

mit $\alpha_e\in K, e\in B, \alpha_e\neq 0$ nur für endlich viele $e\in B$, dargestellt werden.

Beweis. Es sei

$$x=\sum_{e\in B}\alpha_e e=\sum_{e\in B}\beta_e e.$$

Dann ist

$$0=\sum_{e\in B}(\alpha_e-\beta_e)e.$$

Wegen der linearen Unabhängigkeit der Elemente $e\in B$ folgt

$$\alpha_e=\beta_e.$$

□

Die Elemente $\alpha_e\in K$ heißen Koordinaten von x bezüglich der Basis B.

Beispiel

Wir betrachten den $\mathbb{R}^n$ und schreiben die Elemente (n-Tupel) des kartesischen Produktes in Spaltenform

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

Sei B die Standardbasis im $\mathbb{R}^n$

$$B = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} \right\}. \tag{2.8}$$

Dann ist $\boldsymbol{x}$ folgenderweise darstellbar

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \dots + x_n \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix},$$

d.h. x_i ist die i-te Koordinate bezüglich der Standardbasis.

Lemma 2.9 (Austausch eines Basiselements). *Sei $B \subset V$ eine Basis des K-Vektorraums V. Wir betrachten ein Element $e_0 \in B$ und ein Element $f_0 \neq 0$ aus V mit der Darstellung*

$$f_0 = \sum_{e \in B} \alpha_e e, \quad \alpha_{e_0} \neq 0. \tag{2.9}$$

Dann ist auch $\hat{B} = (B \setminus \{e_0\}) \cup \{f_0\}$ eine Basis.

Beweis.

a) Wir zeigen zunächst, dass $\hat{B}$ aus linear unabhängigen Elementen besteht. Dazu betrachten wir die Gleichung

$$\beta_{f_0} f_0 + \sum_{\substack{e \in B \\ e \neq e_0}} \beta_e e = 0. \tag{2.10}$$

Mit der Darstellung (2.9) erhalten wir

$$\beta_{f_0}(\alpha_{e_0} e_0 + \sum_{\substack{e \in B \\ e \neq e_0}} \alpha_e e) + \sum_{\substack{e \in B \\ e \neq e_0}} \beta_e e = \beta_{f_0} \alpha_{e_0} e_0 + \sum_{\substack{e \in B \\ e \neq e_0}} (\beta_{f_0} \alpha_e + \beta_e) e = 0.$$

Da die Elemente $e \in B$ linear unabhängig sind, folgt

$$\begin{aligned} \beta_{f_0} \alpha_{e_0} &= 0, \\ \beta_{f_0} \alpha_e + \beta_e &= 0 \quad \text{für } e \neq e_0. \end{aligned}$$

Da $\alpha_{e_0} \neq 0$ ist, muss $\beta_{f_0} = 0$ sein und daher $\beta_e = 0$ für $e \neq e_0$. Damit verschwinden alle Koeffizienten in (2.10).

b) Wir zeigen, dass jedes $x \in V$ Linearkombination von Elementen aus $\hat{B}$ ist. Aus der Darstellung (2.9) von f_0 folgt

$$\begin{aligned} \alpha_{e_0} e_0 &= f_0 - \sum_{\substack{e \in B \\ e \neq e_0}} \alpha_e e, \\ e_0 &= \frac{f_0}{\alpha_{e_0}} - \sum_{\substack{e \in B \\ e \neq e_0}} \frac{\alpha_e}{\alpha_{e_0}} e. \end{aligned} \tag{2.11}$$

Sei x beliebig aus V. Die Darstellung von x bezüglich der Basis B lautet

$$x = \gamma_{e_0} e_0 + \sum_{\substack{e \in B \\ e \neq e_0}} \gamma_e e.$$

Unter der Beachtung von (2.11) erhalten wir daraus eine Darstellung bezüglich der Basis $\hat{B}$:

$$x = \frac{\gamma_{e_0}}{\alpha_{e_0}} f_0 + \sum_{\substack{e \in B \\ e \neq e_0}} \left(\gamma_e - \frac{\gamma_{e_0}}{\alpha_{e_0}} \alpha_e \right) e.$$

$\square$

Beispiel

In der Standardbasis (2.8) können wir z.B. das Element

$$\boldsymbol{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

durch das Element

$$\boldsymbol{f}_1 = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$$

ersetzen. Damit wird

$$\hat{B} = \left\{ \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} \right\}$$

ebenfalls eine Basis.

Das Lemma 2.9 kann verallgemeinert werden.

Lemma 2.10 (Austausch von k Basiselementen). *Sei V ein K-Vektorraum mit der Basis B. Seien $f_1, \ldots, f_k$ linear unabhängige Elemente aus V, $f_i \neq 0, i = 1, \ldots, k$. Dann gibt es $e_1, \ldots, e_k$ Elemente aus B, so dass auch*

$$\hat{B} = (B \setminus \{e_1, \ldots, e_k\}) \cup \{f_1, \ldots, f_k\}$$

eine Basis ist.

Beweis. Das Element f_1 besitzt eine Darstellung bezüglich B:

$$f_1 = \sum_{e \in B} \alpha_e e \neq 0.$$

Es gibt ein Element $e_1 \in B$, so dass $\alpha_{e_1} \neq 0$ ist. Daher ist nach Lemma 2.9 $\hat{B} = (B \setminus \{e_1\}) \cup f_1$ eine Basis. Nun betrachten wir $f_2 \in V$ und ein Element $e_2 \in \hat{B} \cap B$, so dass

$$f_2 = \sum_{\hat{e} \in \hat{B}} \beta_{\hat{e}} \hat{e} \quad \text{und } \beta_{e_2} \neq 0$$

ist. Ein solches Element e_2 muss existieren, da die $f_i, i = 1, \ldots, k$, linear unabhängig sind. Diesen Prozess setzen wir fort und erhalten die Behauptung. □

Aus Lemma 2.10 folgt der Satz:

Satz 2.11. *Sei V ein K-Vektorraum mit einer Basis von n Elementen. Jede weitere Basis besitzt ebenfalls n Elemente.*

Beweis. Seien $B = \{e_1, \ldots, e_n\}$ und $\hat{B} = \{\hat{e}_1, \ldots, \hat{e}_m\}$ zwei Basen. Wir nehmen an, dass $m > n$ ist. Nach Lemma 2.10 ist jedoch $\{\hat{e}_1, \ldots, \hat{e}_n\}$ bereits eine Basis und die Menge $\hat{B}$ wäre linear abhängig. Dies ist ein Widerspruch und m muss gleich n sein. □

Existenz von Basen

In unseren Beispielen hatten wir bisher gegebene Basen betrachtet. Erinnern wir uns an den Vektorraum

$$V = \{f \ : \ f : \mathbb{R} \to \mathbb{R}\},$$

so sieht man nicht unmittelbar, dass in diesem Raum eine Basis existiert. Die Frage nach der Existenz von Basen ist also durchaus sinnvoll.
Es gilt: Jeder Vektorraum besitzt eine Basis. Um dies zu beweisen benötigt man Axiome aus der Mengenlehre. Wir verzichten hier auf den Beweis und verweisen auf [2, S. 261]. Jedoch können wir in sogenannten „endlich erzeugten“ Vektorräumen leicht eine Basis konstruieren.

Definition 2.12. Ein Vektorraum heißt endlich erzeugt, wenn es endlich viele Elemente $v_1, \ldots, v_n \in v$ gibt, so dass

$$\text{Span}(\{v_1, \ldots, v_n\}) = V$$

ist.

Satz 2.13. *Jeder endlich erzeugte Vektorraum V besitzt eine Basis.*

Beweis. Wir wählen aus der Menge $\{v_1, \ldots, v_n\}$ eine größte Teilmenge aus, die linear unabhängig ist. Dies sei $B = \{v_1, \ldots, v_r\}, r \leq n$. B ist Basis von V. □

Da wir festgestellt haben, dass jeder Vektorraum eine Basis besitzt, können wir jetzt definieren, was wir unter der Dimension eines Vektorraums verstehen. Dazu bemerken wir, dass zwei Basen in V stets die gleiche Mächtigkeit besitzen.

Definition 2.14. Besitzt $V \neq \{0\}$ eine Basis aus n Elementen, dann hat V die Dimension n, d.h. $\dim V = n$.
$v = \{0\}$ besitzt die Dimension 0.
Besitzt V eine unendliche Basis, dann ist V unendlich-dimensional. Wir schreiben $\dim V = \infty$.

Wir können also kurz formulieren:
Ist $V \neq \{0\}$, dann ist die Dimension von V gleich der Anzahl der Basiselemente.

Vektorrechnung im $\mathbb{R}^n$

In diesem Abschnitt gehen wir auf die geometrische Veranschaulichung der Rechenoperatoren, Koordinatendarstellung und der Beschreibung von Geraden und Ebenen im $\mathbb{R}^n$ ein, indem wir die Darstellung durch Pfeile benutzen.
Es sei $V = \mathbb{R}^n$ der n-dimensionale Vektorraum mit der bereits zu Beginn dieses Abschnittes eingeführten Addition und der Multiplikation mit Skalaren aus $K = \mathbb{R}$.
Wir betrachten die Standardbasis B. Die Koordinatendarstellung eines Elements (Vektors) aus V lautet also

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \sum_{i=1}^{n} x_i \boldsymbol{e}_i, \quad \boldsymbol{e}_i = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \\ 0 \end{pmatrix} \quad \leftarrow i\text{-te Stelle.} \tag{2.12}$$

Die geometrische Darstellung des Elements x erfolgt im $\mathbb{R}^2$ und $\mathbb{R}^3$ auch durch den sogenannten Ortsvektor $\vec{x}$, einem Pfeil vom 0-Element (Nullpunkt) zum $\boldsymbol{x}$-Element (Punkt $\boldsymbol{x}$).

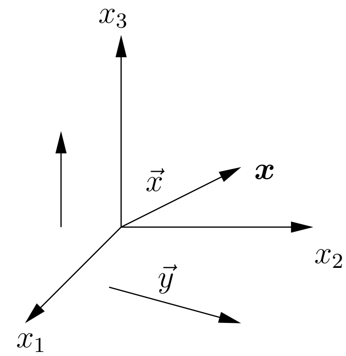


Abbildung 2.3: Ortsvektor $\vec{x}$ und freier Vektor $\vec{y}$

Die Ortsvektoren können im $\mathbb{R}^n$ durch Drehungen und Verschiebungen bewegt werden. Der Anfangspunkt eines solchen entstandenen „freien “ Vektors braucht also nicht im Nullpunkt zu liegen. Um eine bijektive Abbildung zwischen der Menge der „freien Vektoren“ und der Koordinatendarstellung (2.12) zu erhalten, wird eine Äquivalenzrelation in der Menge der freien „Pfeilvektoren “ eingeführt:
Zwei freie Vektoren $\vec{x}$ und $\vec{y}$ sind äquivalent, falls Länge und Richtung der entsprechenden Pfeile übereinstimmen (siehe Abbildung 2.4).

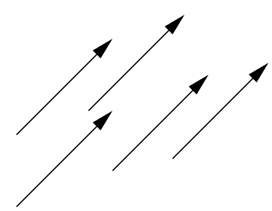


Abbildung 2.4: Äquivalente freie Pfeilvektoren

Wir können stets einen Vertreter der Äquivalenzklasse finden, dessen Ausgangspunkt im Nullpunkt liegt und der durch den Index 0 gekennzeichnet wird. Damit wird die Menge der Äquivalenzklassen der freien Pfeilvektoren auf den $\mathbb{R}^2$ oder den $\mathbb{R}^3$ bijektiv abgebildet. Dies sei durch die Abbildung

$$f : \{\text{Äquivalenzklassen freier Vektoren}\} \to \mathbb{R}^n, \quad n = 2, 3$$

realisiert. Wir beschreiben Summe und Differenz zweier freier Pfeilvektoren.
Die Summe zweier freier Pfeilvektoren $\vec{x}$ und $\vec{y}$ (o.B.d.A. Endpunkt von $\vec{x}$=Ausgangspunkt von $\vec{y}$) ist der Pfeilvektor, der vom Ausgangspunkt von $\vec{x}$ zum Endpunkt von $\vec{y}$ führt.
Wir überzeugen uns, dass für die bijektive Abbildung $f : \{\text{Äquivalenzklassen freier Vektoren}\} \to \mathbb{R}^2$ gilt

$$f(\vec{x} + \vec{y}) = f(\vec{x}) + f(\vec{y}) = \boldsymbol{x} + \boldsymbol{y}.$$

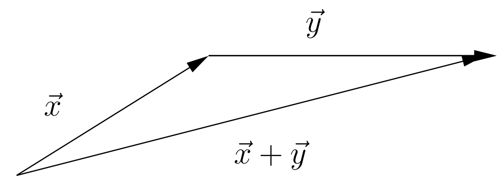


Abbildung 2.5: Addition zweier freier Pfeilvektoren

Dazu betrachten wir die zu $\vec{x}$ und $\vec{y}$ äquivalenten Ortsvektoren $\vec{x}_0, \vec{y}_0$ deren Ausgangspunkt der Nullpunkt ist. Dann ist $\vec{x}_0 + \vec{y}_0$ der Ortsvektor zu $\boldsymbol{x}+\boldsymbol{y}$ (siehe Abbildung 2.6).

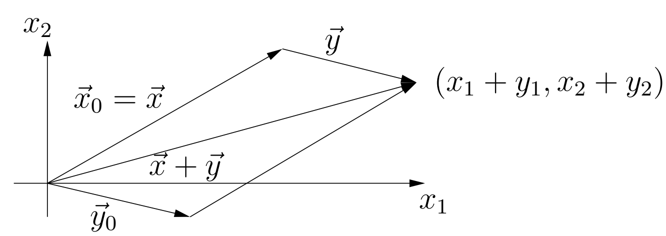


Abbildung 2.6: Ortsvektor zu $\boldsymbol{x}+\boldsymbol{y}$

Die Differenz zweier freier Pfeilvektoren ist durch

$$\vec{x}-\vec{y} = \vec{x} + (-\vec{y})$$

erklärt (siehe Abbildung 2.7).

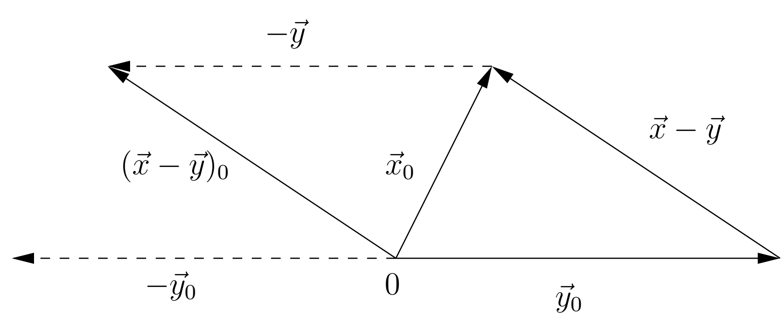


Abbildung 2.7: Differenz zweier freier Vektoren

Die Multiplikation eine Pfeilvektors $\vec{x}$ mit $\alpha \in \mathbb{R}$ ergibt einen Pfeilvektor, dessen Länge das $|\alpha|$-fache der Länge von $\vec{x}$ beträgt. Seine Richtung stimmt mit der Richtung von $\vec{x}$ überein, falls $\alpha > 0$ ist; für $\alpha < 0$ ist sie entgegengesetzt.

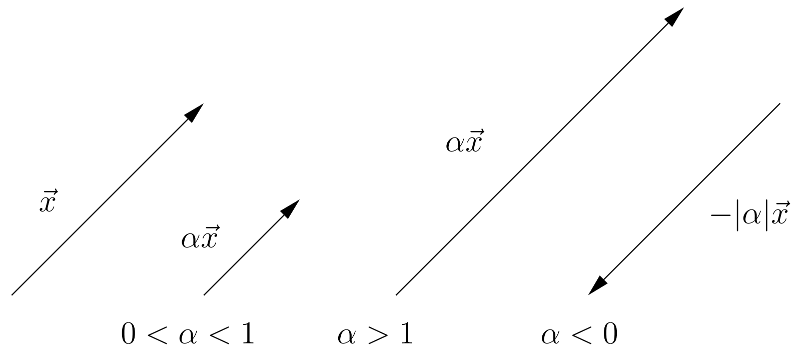


Abbildung 2.8: Multiplikation eines Pfeilvektors mit einem Skalar

Geraden und Ebenen im $\mathbb{R}^n$

Wir beschreiben zunächst Geraden im $\mathbb{R}^n$.

Definition 2.15. Die Teilmenge $G \subset \mathbb{R}^n$ ist eine Gerade, wenn es Elemente $\boldsymbol{x}, \boldsymbol{y}$ aus V gibt, so dass

$$G = \{\boldsymbol{z} \in \mathbb{R}^n \,:\, \exists\, \alpha \in \mathbb{R}, \text{ so dass } \boldsymbol{z} = \boldsymbol{x} + \alpha\boldsymbol{y}\}.$$

Diese Darstellung heißt Parameterdarstellung ($\alpha \in \mathbb{R}$ ist der Parameter). Ist $\boldsymbol{x} = \boldsymbol{0}$, so ist G eine Gerade durch den Nullpunkt.

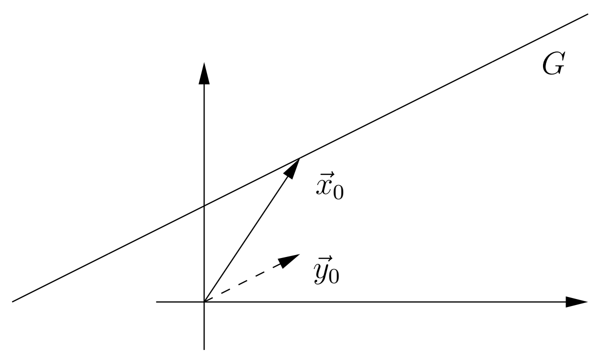


Abbildung 2.9: Parameterdarstellung der Geraden G

Eine Gerade, die durch die Punkte $\boldsymbol{x}$ und $\boldsymbol{y}$ verläuft hat die Parameterdarstellung

$$G = \{\boldsymbol{z} \in \mathbb{R}^n \,:\, \exists\, \alpha \in \mathbb{R}, \text{ so dass } \boldsymbol{z} = \boldsymbol{x} + \alpha(\boldsymbol{y} - \boldsymbol{x})\}.$$

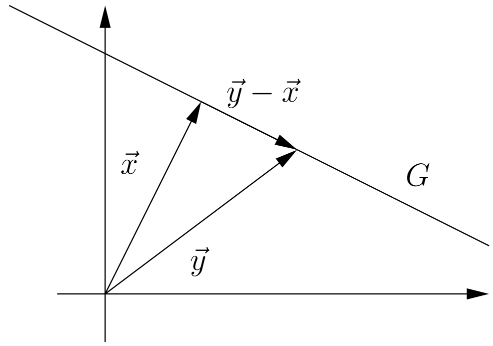


Abbildung 2.10: Gerade durch $\boldsymbol{x}$ und $\boldsymbol{y}$

Diese Definition gilt auch in allgemeinen Vektorräumen V, d.h. eine Gerade ist durch zwei Vektoren $\boldsymbol{x}$ und $\boldsymbol{y}$ eindeutig bestimmt. Bei der Definition von Ebenen werden wir auf drei Vektoren zurückgreifen.

Definition 2.16. Eine Teilmenge $E \subset \mathbb{R}^n$ nennen wir Ebene, wenn es drei Vektoren $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in \mathbb{R}^n$ gibt, mit

$$E = \{\boldsymbol{x} \in \mathbb{R}^n \,:\, \exists\, \alpha, \beta \in \mathbb{R}, \text{ so dass } \boldsymbol{x} = \boldsymbol{u} + \alpha \boldsymbol{v} + \beta \boldsymbol{w}\}.$$

Dies ist ebenfalls eine Parameterdarstellung.

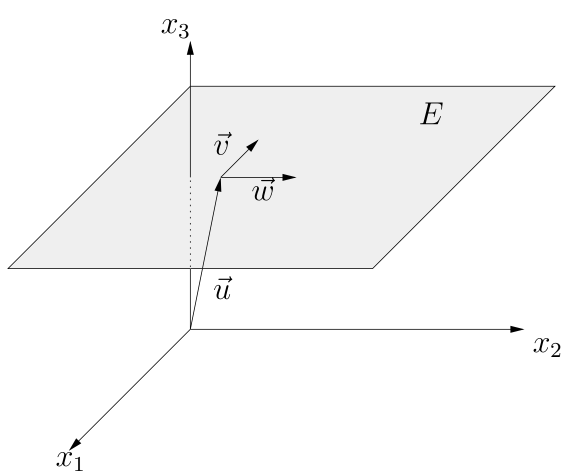


Abbildung 2.11: Ebene durch $\boldsymbol{u}$, $\boldsymbol{v}$, $\boldsymbol{w}$

Eine Ebene, die durch die Punkte $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}$ verläuft, hat die Darstellung

$$E = \{\boldsymbol{x} \in \mathbb{R}^n \,:\, \exists\, \alpha, \beta \in \mathbb{R}, \text{ so dass } \boldsymbol{x} = \boldsymbol{u} + \alpha(\boldsymbol{v} - \boldsymbol{u}) + \beta(\boldsymbol{w} - \boldsymbol{u})\}.$$

Es ist

$$\boldsymbol{x} = \boldsymbol{u} + \alpha(\boldsymbol{v} - \boldsymbol{u}) + \beta(\boldsymbol{w} - \boldsymbol{u}) = \begin{cases} \boldsymbol{u} & \text{für } \alpha = \beta = 0, \\ \boldsymbol{v} & \text{für } \alpha = 1, \beta = 0, \\ \boldsymbol{w} & \text{für } \alpha = 0, \beta = 1. \end{cases}$$

Übungsaufgaben

28) Lineare Splines. Um aus einer (endlichen) Wertetabelle $\{(x_0, y_0), (x_1, y_1), \ldots, (x_n, y_n)\}$ eine Funktion zu konstruieren, verwendet man vielfach so genannte Splines. Im einfachsten Fall der linearen Splines, verbindet man die Punkte im Graphen der Wertetabelle durch Geraden. Zur weiteren Vereinfachung betrachten wir äquidistante Stützstellen im Einheitsintervall, also

$$x_i = \frac{i}{n}\,.$$

Die Splinefunktion $f_{\boldsymbol{y}} : [0,1] \to \mathbb{R}$ ist dann für $\boldsymbol{y} = (y_0, y_1, \ldots, y_n)$ gegeben durch

$$f_{\boldsymbol{y}}(x) = y_i(i + 1 - nx) + y_{i+1}(nx - i)\,, \qquad \text{falls } \frac{i}{n} \le x \le \frac{i+1}{n}\,.$$

Zeigen Sie, dass

1. $f_{\boldsymbol{y}}(x_i) = y_i$ für alle $0 \le i \le n$ und $f_{\boldsymbol{y}}$ auf dem Intervall $[x_i, x_{i+1}]$ eine Gerade ist.

2. $V_{\frac{1}{n}} = \{f_{\boldsymbol{y}} | \boldsymbol{y} \in \mathbb{R}^{n+1}\}$ ein $\mathbb{R}$-Vektorraum ist, wobei Addition und Multiplikation mit Skalaren in folgender Weise definiert sind

$$\begin{aligned} \big(f_{\boldsymbol{y}} + f_{\boldsymbol{z}}\big)(x) &:= f_{\boldsymbol{y}}(x) + f_{\boldsymbol{z}}(x), \\ \big(\alpha f_{\boldsymbol{y}}\big)(x) &:= \alpha f_{\boldsymbol{y}}(x). \end{aligned}$$

Geben Sie eine Basis des Vektorraums an und skizzieren Sie diese Basiselemente.

29) Sei $V = \mathbb{R}^3$. Gegeben seien die vom Parameter $t \in \mathbb{R}$ abhängigen Vektoren

$$\boldsymbol{a} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \qquad \boldsymbol{b} = \begin{pmatrix} 1 \\ 3 \\ 1 \end{pmatrix}, \qquad \boldsymbol{c} = \begin{pmatrix} 1 \\ 2 \\ t^2 \end{pmatrix}.$$

Untersuchen Sie, für welche Werte von t diese Vektoren linear abhängig bzw. unabhängig sind.

30) Sei $V = \mathbb{R}^4$. Gegeben seien die vier Vektoren

$$\boldsymbol{a_1} = \begin{pmatrix} 2 \\ 0 \\ 1 \\ 3 \end{pmatrix}, \qquad \boldsymbol{a_2} = \begin{pmatrix} 0 \\ 1 \\ 2 \\ 3 \end{pmatrix}, \qquad \boldsymbol{a_3} = \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix}, \qquad \boldsymbol{a_4} = \begin{pmatrix} 0 \\ -2 \\ 1 \\ 0 \end{pmatrix}.$$

1. Zeigen Sie, dass $\{\boldsymbol{a_1}, \boldsymbol{a_2}, \boldsymbol{a_3}, \boldsymbol{a_4}\}$ eine Basis von V ist.

2. Stellen Sie den Vektor $\boldsymbol{b} = \begin{pmatrix} 0 \\ 5 \\ 2 \\ 6 \end{pmatrix}$ in dieser Basis dar.

31) Gegeben sei für alle $t \in \mathbb{R}$ die Geradenschar

$$g_t := \left\{ \begin{pmatrix} 1 \\ 2 \\ -5 \end{pmatrix} + \lambda \begin{pmatrix} 0 \\ 1 \\ t \end{pmatrix} \; : \; \lambda \in \mathbb{R} \right\}.$$

1. Zeigen Sie, dass alle Geraden in einer Ebene E liegen und geben Sie eine Parameterdarstellung dieser Ebene an.

2. Die Hessesche Normalform einer Ebene ist gegeben durch die Gleichung

$$a_1x_1 + a_2x_2 + a_3x_3 = b\,.$$

   Berechnen Sie die Hessesche Normalform der Ebene aus der vorigen Aufgabe.

3. Gibt es Geraden dieser Schar, die parallel zu einer der Koordinatenachsen des $\mathbb{R}^3$ verlaufen?

4. Gibt es zu jeder der Geraden eine senkrechte Gerade, die ebenfalls zu dieser Geradenschar gehört?

2.2 Lineare Abbildungen von Vektorräumen

Sind Elemente, Geraden bzw. Ebenen in einem Vektorraum gegeben, so können wir sie bewegen bzw. verkleinern oder vergrößern, indem wir eine Verschiebung (Translation), eine Drehung (Rotation), eine Streckung oder Stauchung vornehmen. Drehungen, Stauchungen oder Streckungen können durch lineare Abbildungen beschrieben werden.
Wir beginnen mit der Definition einer linearen Abbildung.

Definition 2.17. Seien V und W K-Vektorräume. Eine Abbildung f

$$f : V \to W$$

heißt linear, falls gilt

- $f(x + y) = f(x) + f(y)$ für alle $x, y \in V$,
- $f(\alpha x) = \alpha f(x)$ für alle $\alpha \in K$ und $x \in V$.

Die Menge aller linearen Abbildungen von V in W wird als $L(V, W)$ bezeichnet.

Beispiele

1° Es sei $V = W = \mathbb{R}^2$. Die Abbildung $f : \mathbb{R}^2 \to \mathbb{R}^2$,

$$f\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix}$$

ist linear. Geometrisch beschreibt f eine Drehung „nach links" um 90° (siehe Abbildung 2.12).

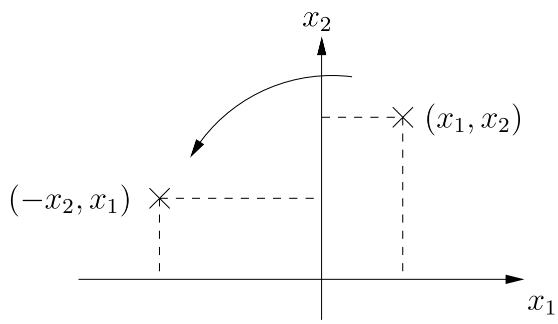


Abbildung 2.12: Drehung um 90°

2° Es sei $V = \mathbb{R}^3, W = \mathbb{R}^2$. Die Abbildung $f : \mathbb{R}^3 \to \mathbb{R}^2$

$$f\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

ist eine Projektion (siehe Abbildung 2.13), f ist linear.

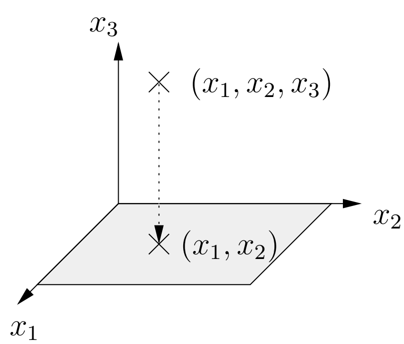


Abbildung 2.13: Projektion auf die x_1x_2-Ebene

3° Es sei $V = \mathbb{R}^2, W = \mathbb{R}^3$. Wir betrachten zwei feste Elemente $\boldsymbol{u}$ und $\boldsymbol{v} \in W$. Wir betrachten die Abbildung

$$f : V \to W,$$
$$f \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = x_1\boldsymbol{u} + x_2\boldsymbol{w}.$$

Die Abbildung f ist linear.

4° Es sei $V = W = \mathbb{R}, f : V \to W, f(x) = ax + b$, wobei a und b feste reelle Zahlen sind und $a \neq 0, b \neq 0$. Die Abbildung f ist nicht linear:

$$f(x + y) = a(x + y) + b \neq f(x) + f(y) = a(x + y) + 2b.$$

5° Es sei $V = W = \mathbb{R}^2, f : V \to W, f \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2^2 \end{pmatrix}$. Die Abbildung f ist nicht linear

$$f(\alpha\boldsymbol{x}) = f \begin{pmatrix} \alpha x_1 \\ \alpha^2 x_2^2 \end{pmatrix} \neq \alpha f \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \alpha x_1 \\ \alpha x_2^2 \end{pmatrix}$$

für $\alpha \neq 0, \alpha \neq 1$.

Wir charakterisieren jetzt Vektorräume, die als „gleich“ angesehen werden können.

Definition 2.18. Seien V und W K-Vektorräume. Eine bijektive lineare Abbildung $f : V \to W$ heißt Isomorphismus. Zwei K-Vektorräume V und W heißen isomorph (geschrieben als $V \cong W$), falls es einen Isomorphismus $f : V \to W$ gibt.

Beispiel

Sei V ein n-dimensionaler K-Vektorraum mit der Basis $B = \{e_1, \ldots, e_n\}$. Jedes $x \in V$ ist eindeutig darstellbar als

$$x = \sum_{i=1}^{n} \alpha_i e_i = \sum_{i=1}^{n} \alpha_i(x) e_i.$$

Die Abbildung

$$f : V \to K^n, f(x) = \begin{pmatrix} \alpha_1(x) \\ \vdots \\ \alpha_n(x) \end{pmatrix} \tag{2.13}$$

ist ein Isomorphismus. Dieser Isomorphismus heißt **Koordinatendarstellung zur Basis $\boldsymbol{B}$**. Lineare Abbildungen f sind durch ihre Werte auf einer Basis vollständig bestimmt. Dies wird durch nachfolgendes Lemma genauer ausgedrückt.

Lemma 2.19. *Seien V und W K-Vektorräume und $B \subset V$ eine Basis. Jedem $e \in B$ sei durch f ein Element $\hat{e} \in W$ zugeordnet. Dann gibt es genau eine lineare Abbildung*

$$f : V \to W$$

mit

$$f(e) = \hat{e}.$$

Beweis. Jedes $x \in V$ hat die Darstellung

$$x = \sum_{e \in B} \alpha_e e, \quad \alpha_e = \alpha_e(x) \in K,$$

wobei nur endlich viele $\alpha_e(x) \neq 0$ sind. Für eine beliebige lineare Abbildung $f : V \to W$ gilt

$$f(x) = f\left(\sum_{e \in B} \alpha_e(x) e\right) = \sum_{e \in B} \alpha_e(x) f(e). \tag{2.14}$$

Sei $f(e) = \hat{e}$. Dann ist $f(x) := \sum_{e \in B} \alpha_e(x)\hat{e}$ linear. Für jede weitere lineare Abbildung $g : V \to W$ mit $g(e) = \hat{e}$ gilt (2.14)

$$g(x) = \sum_{e \in B} \alpha_e(x)\hat{e}$$

und damit muss $f = g$ sein. □

Satz 2.20. *Wir betrachten die K-Vektorräume V und W und eine lineare Abbildung $f : V \to W$. Dann sind die Mengen*

$$\begin{aligned} \mathrm{im}(f) &= f(V) = \{f(x) \,:\, x \in V\} \subset W, && \textit{(Bildmenge, image)} \\ \ker(f) &= \{x \in V \,:\, f(x) = 0\} \subset V && \textit{(Kern, kernel)} \end{aligned}$$

Teilvektorräume.

Beweis. Wir müssen zeigen, dass $\mathrm{im}(f)$ und $\ker(f)$ bezüglich der Addition und der Multiplikation mit Skalaren abgeschlossen sind. Wir sehen uns zunächst $\mathrm{im}(f)$ an. Seien $w_1, w_2 \in \mathrm{im}(f)$. Es gibt Elemente x_1 und $x_2 \in V$, so dass $w_1 = f(x_1)$ und $w_2 = f(x_2)$ ist. Da f linear ist, gilt

$$w_1 + w_2 = f(x_1) + f(x_2) = f(x_1 + x_2) \in \mathrm{im}(f).$$

Weiterhin ist für ein $w = f(x) \in W$:

$$\alpha w = \alpha f(x) = f(\alpha x) \in \mathrm{im}(f).$$

Für $\ker(f)$ werden analoge Überlegungen ausgeführt. □

Die Dimensionen der Teilvektorräume $\mathrm{im}(f)$ und $\ker(f)$ hängen wie folgt zusammen:

Satz 2.21 (Dimensionssatz). *Sei V ein n-dimensionaler und W ein beliebiger K-Vektorraum; $f : V \to W$ eine lineare Abbildung. Dann ist*

$$n = \dim(V) = \dim(\ker(f)) + \dim(\mathrm{im}(f)).$$

Beweis. Nach Lemma 2.10 ist stets eine solche Basis

$$B = \{e_1, e_2, \ldots, e_r, e_{r+1}, \ldots, e_n\}$$

von V wählbar, dass

$$B_{\ker(f)} = \{e_1, e_2, \ldots, e_r\}, \quad r \leq n,$$

Basis von $\ker(f)$ ist. Wegen der Linearität von f ist

$$\mathrm{im}(f) = \mathrm{Span}\{f(e_{r+1}), \ldots, f(e_n)\}.$$

Wir zeigen, dass die Elemente $f(e_{r+1}), \ldots, f(e_n)$ linear unabhängig sind und damit eine Basis von $\mathrm{im}(f)$ bilden. Dazu betrachten wir eine verschwindende Linearkombination

$$0 = \sum_{j=1}^{n-r} \alpha_{r+j} f(e_{r+j}) = f\left(\sum_{j=1}^{n-r} \alpha_{r+j} e_{r+j}\right).$$

Damit ist $\sum_{j=1}^{n-r} \alpha_{r+j} e_{r+j}$ aus $\ker(f)$ und kann als Linearkombination der Basiselemente $e_1, \ldots, e_r$ dargestellt werden,

$$\sum_{j=1}^{n-r} \alpha_{r+j} e_{r+j} = \sum_{i=1}^{r} \beta_i e_i.$$

Es folgt, dass

$$\sum_{i=1}^{r} \beta_i e_i - \sum_{j=1}^{n-r} \alpha_{r+j} e_{r+j} = 0$$

ist. Da $B = \{e_1, \ldots, e_n\}$ aus linear unabhängigen Elementen besteht, müssen die Koeffizienten $\alpha_{r+j} = 0$ für $j = 1, \ldots, n-r$ sein. □

Bezeichnung: Sei $f \in L(V, W)$. V, W K-Vektorräume. Wir bezeichnen als **Rang von** f

$$\dim(\mathrm{im}(f)) = \dim(f(V)) = \mathrm{Rang}(f).$$

Folgerung 2.22. Ein Isomorphismus zwischen zwei endlichdimensionalen Vektorräumen bildet Basen aufeinander ab. Umgekehrt, falls f Basen zweier n-dimensionaler Räume aufeinander abbildet, dann ist f ein Isomorphismus ($\ker(f) = \{0\}$).

Übungsaufgaben

32) Zeigen Sie, dass folgende Abbildungen linear sind. Bestimmen Sie deren Kern, Bild und die zugehörige Matrix bzgl. der Standardbasis.

1. $f : \mathbb{R}^3 \to \mathbb{R}^2\,, (x, y, z) \mapsto (x + y, y + z)$,
2. $g : \mathbb{R}^2 \to \mathbb{R}^3\,, (x, y) \mapsto (x, x + y, y)$.

33) Seien U, V, W Vektorräume über einem Körper K; das neutrale Element der Addition sei jeweils $0_V, 0_W$ und $g : U \to V$, $f : V \to W$ seien lineare Abbildungen. Zeigen Sie

1. $f(0_V) = 0_W$.
2. $f \circ g : U \to W$ ist linear.
3. Ist f bijektiv, so ist auch $f^{-1} : W \to V$ linear.

2.3 Matrizen

Wir hatten lineare Abbildungen eines K-Vektorraums V in den K-Vektorraum W im vorherigen Abschnitt betrachtet. Wir betrachten jetzt endlichdimensionale K-Vektorräume V und W; $\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_s\}$ sei Basis von V, $\{\boldsymbol{w}_1, \dots, \boldsymbol{w}_r\}$ sei Basis von W. Sei f eine lineare Abbildung von V in W. Es gibt eindeutig bestimmt Zahlen $a_{ij} \in K$, die die Koordinaten der Basisdarstellung von $f(\boldsymbol{v}_j)$ sind:

$$f(\boldsymbol{v}_j) = \sum_{i=1}^{r} a_{ij}\boldsymbol{w}_i, \quad j = 1, \dots, s. \tag{2.15}$$

Nach Lemma 2.19 ist die lineare Abbildung f durch die Elemente $a_{ij} \in K, 1 \le i \le r, 1 \le j \le s$ eindeutig bestimmt. Wir ordnen die Menge $\{a_{ij}\}_{\substack{1\le i\le r\\ 1\le j\le s}}$ in Form einer Matrix an.

Definition 2.23. Sei K ein Körper, $r, s \in \mathbb{N}$ und $a_{ij} \in K, 1 \le i \le r, 1 \le j \le s$. Ein rechteckiges Schema

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1s} \\ a_{21} & a_{22} & \dots & a_{2s} \\ \vdots & \vdots & \vdots & \vdots \\ a_{r1} & a_{r2} & \dots & a_{rs} \end{pmatrix}$$

heißt $(r \times s)$ Matrix über K. A besitzt r Zeilen und s Spalten. Der erste Index ist der Zeilenindex, der zweite der Spaltenindex. Die Vektoren $\begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{rj} \end{pmatrix}$ heißen Spaltenvektoren, $j = 1, \dots, s$. Die Vektoren $(a_{i1}, \dots, a_{is}), i = 1, \dots, r$, heißen Zeilenvektoren.

Aus den obigen Überlegungen folgt:

Lemma 2.24. *Sei V ein s-dimensionaler K-Vektorraum mit der Basis $\{v_1, \dots, v_s\}$ und W ein r-dimensionaler K-Vektorraum mit der Basis $\{w_1, \dots, w_r\}$. Jeder linearen Abbildung $f : V \to W$ wird durch (2.15) eine eindeutig bestimmte Matrix $A = (a_{ij})_{\substack{1 \le i \le r \\ 1 \le j \le s}}$ zugeordnet. Umgekehrt wird durch eine gegebene $r \times s$ Matrix A durch (2.15) eine lineare Abbildung $f : V \to W$ definiert. Dabei sind die Spaltenvektoren von (a_{ij}) die Koordinatenvektoren der Bilder der Basisvektoren von V.*

Beispiele

1° Wir betrachten den $\mathbb{R}^2$ mit der Standardbasis und die lineare Abbildung $f : \mathbb{R}^2 \to \mathbb{R}^2$ mit

$$f\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_2 \\ -x_1 \end{pmatrix}.$$

Die Gleichungen (2.15) lauten

$$f\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \end{pmatrix} = 0\begin{pmatrix} 1 \\ 0 \end{pmatrix} - 1\begin{pmatrix} 0 \\ 1 \end{pmatrix},$$
$$f\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 1\begin{pmatrix} 1 \\ 0 \end{pmatrix} + 0\begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Damit ist $a_{11} = 0, a_{21} = -1, a_{12} = 1, a_{22} = 0$ und

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

2° Es sei $V = \mathbb{R}^3, W = \mathbb{R}^2$ mit den entsprechenden Standardbasen. Wir betrachten die Projektion

$$f\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

Wir sehen uns die Gleichungen (2.15) an:

$$f\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 1\begin{pmatrix} 1 \\ 0 \end{pmatrix} + 0\begin{pmatrix} 0 \\ 1 \end{pmatrix},$$
$$f\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} = 0\begin{pmatrix} 1 \\ 0 \end{pmatrix} + 1\begin{pmatrix} 0 \\ 1 \end{pmatrix},$$
$$f\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} = 0\begin{pmatrix} 1 \\ 0 \end{pmatrix} + 0\begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Die entsprechende (2×3) Matrix A lautet

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

3° Es sei $V = \mathbb{R}^3$ mit der Standardbasis und $W = \mathbb{R}^2$, versehen mit der Basis $\left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$.
Sei wie in Beispiel 2°

$$f \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

Dann ist

$$f \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} - 1 \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

$$f \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} = 0 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + 1 \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

$$f \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} = 0 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + 0 \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

und somit

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

Dieses Beispiel zeigt, dass A und $\hat{A}$ verschieden sind und somit von der Wahl der Basen abhängen.

Das Bild eines allgemeinen Vektors

Bisher hatten wir uns angesehen, wohin die Basisvektoren durch eine lineare Abbildung abgebildet werden. Wir sehen uns jetzt das Bild eines beliebigen Elementes $\boldsymbol{v} \in V$ an. Es ist

$$f(\boldsymbol{v}) = f\left(\sum_{j=1}^{s} \alpha_j \boldsymbol{v}_j\right) = \sum_{j=1}^{s} \alpha_j f(\boldsymbol{v}_j) \overset{(2.15)}{=} \sum_{j=1}^{s} \alpha_j \sum_{i=1}^{r} a_{ij} \boldsymbol{w}_i = \sum_{i=1}^{r} \left(\sum_{j=1}^{s} a_{ij}\alpha_j\right) \boldsymbol{w}_i.$$

Damit werden die Koordinaten $\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_s \end{pmatrix}$ bezüglich der Basis $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_s\}$ auf die Koordinaten

$$\begin{pmatrix} \sum_{j=1}^{s} a_{1j}\alpha_j \\ \vdots \\ \sum_{j=1}^{s} a_{rj}\alpha_j \end{pmatrix} = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_r \end{pmatrix} \tag{2.16}$$

bezüglich der Basis $\{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_r\}$ abgebildet.
In (2.16) treten die Einträge a_{ij} der $(r \times s)$ Matrix A auf. Wir schreiben

$$\begin{aligned} A\,\boldsymbol{\alpha} &= \begin{pmatrix} a_{11} & \ldots & a_{1s} \\ a_{21} & \ldots & a_{2s} \\ \vdots & \vdots & \vdots \\ a_{r1} & \ldots & a_{rs} \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_s \end{pmatrix} := \begin{pmatrix} \sum_{j=1}^s a_{1j}\alpha_j \\ \vdots \\ \sum_{j=1}^s a_{rj}\alpha_j \end{pmatrix} \\ &= \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_r \end{pmatrix} = \boldsymbol{\beta}. \end{aligned} \tag{2.17}$$

Wir haben damit eine Matrix-Vektor Multiplikation eingeführt, indem wir rechnen:
1. Koordinate des Bildvektors = 1. Zeile der Matrix „mal “ Vektor,
2. Koordinate des Bildvektors = 2. Zeile der Matrix „mal “ Vektor,
$\vdots$
r. Koordinate des Bildvektors = r. Zeile der Matrix „mal “ Vektor.
Hierbei bedeutet „mal “ das Produkt der Zeilenvektoren der Matrix mit dem Spaltenvektor $\boldsymbol{\alpha}$.

Bemerkung 2.25. Sind $V = K^s$ und $W = K^r$ mit der Standardbasis versehen, dann ist

$$f(\boldsymbol{x}) = f\left(\sum_{j=1}^s x_j \boldsymbol{e}_j\right) = \sum_{i=1}^r \left(\sum_{j=1}^s a_{ij}x_j\right) \boldsymbol{e}_i = A\boldsymbol{x}. \tag{2.18}$$

In diesem Fall identifizieren wir f mit A und schreiben $f = f_A$.

Beispiele

1° Wir sehen uns das erste vorherige Beispiel an:

$$f\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_2 \\ -x_1 \end{pmatrix}.$$

In der Standardbasis ist

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix}$$

und

$$\begin{pmatrix} x_2 \\ -x_1 \end{pmatrix} = x_2 \begin{pmatrix} 1 \\ 0 \end{pmatrix} - x_1 \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}.$$

Mit Hilfe der Matrix $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ erhalten wir

$$f\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = A\boldsymbol{x} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_2 \\ -x_1 \end{pmatrix} = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}.$$

2° Wir betrachten die lineare Abbildung $f : \mathbb{R}^2 \to \mathbb{R}^3$, wobei in $\mathbb{R}^2$ und $\mathbb{R}^3$ die Standardbasen betrachtet werden,

$$f\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 + 2x_2 \\ 3x_1 + 4x_2 \\ 5x_1 + 6x_2 \end{pmatrix}.$$

Sie kann durch eine (3×2) Matrix A dargestellt werden

$$f(\boldsymbol{x}) = A\boldsymbol{x} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

Produkt von Matrizen

Man kann in natürlicher Weise ein Produkt von Matrizen erklären. Sei K ein Körper und $f : K^s \to K^r, g : K^r \to K^t$ lineare Abbildungen. Die Komposition

$$g \circ f : K^s \to K^t$$

ist linear. Zu f gehöre die $(r \times s)$ Matrix $A = (a_{ij})$, zu g die $(t \times r)$ Matrix $B = (b_{kl})$. Dann soll das Produkt von BA die zu $g \circ f$ gehörige $(t \times s)$ Matrix sein. Nach (2.17) erhalten wir (man beachte, dass $(x_1, \ldots, x_s) = (\alpha_1, \ldots, \alpha_s)$ ist)

$$f(\boldsymbol{x}) = f(x_1, \ldots, x_s) = A\boldsymbol{x} = \begin{pmatrix} \sum_{j=1}^s a_{1j} x_j \\ \vdots \\ \sum_{j=1}^s a_{sj} x_j \end{pmatrix},$$

$$g(\boldsymbol{y}) = g(y_1, \ldots, y_r) = B\boldsymbol{y} = \begin{pmatrix} \sum_{l=1}^r b_{1l} y_l \\ \vdots \\ \sum_{l=1}^r b_{tl} y_l \end{pmatrix}.$$

Für die Komposition gilt:

$$\begin{aligned} (g \circ f)(\boldsymbol{x}) &= g(\sum_{j=1}^s a_{1j} x_j, \ldots, \sum_{j=1}^s a_{rj} x_j) = \begin{pmatrix} \sum_{l=1}^r (b_{1l} \sum_{j=1}^s a_{lj} x_j) \\ \vdots \\ \sum_{l=1}^r (b_{tl} \sum_{j=1}^s a_{lj} x_j) \end{pmatrix} \\ &= \begin{pmatrix} \sum_{j=1}^s (\sum_{l=1}^r b_{1l} a_{lj}) x_j \\ \vdots \\ \sum_{j=1}^s (\sum_{l=1}^r b_{tl} a_{lj}) x_j \end{pmatrix} = BA\boldsymbol{x} = C\boldsymbol{x}, \end{aligned}$$

wobei C die Einträge

$$c_{kj} = \sum_{l=1}^r b_{kl} a_{lj}, \quad 1 \le k \le t, 1 \le j \le s$$

besitzt.
Die Multiplikation zweier Matrizen ist also folgendermaßen definiert:
Sei A eine $(r \times s)$ Matrix, B eine $(t \times r)$ Matrix. Dann ist das Produkt $BA = C$ eine $(t \times s)$ Matrix, deren Einträge

$$c_{kj} = \sum_{l=1}^{r} b_{kl} a_{lj} \tag{2.19}$$

(Produkt der k-ten Zeile von B mit der j-ten Spalte von A) sind:

$$\begin{array}{ccc} B & A & = C, \\ (t \times r) & (r \times s) & (t \times s) \end{array}$$

$$\begin{pmatrix} b_{11} & \dots & b_{1r} \\ \vdots & \vdots & \vdots \\ b_{k1} & \dots & b_{kr} \\ \vdots & \vdots & \vdots \\ b_{t1} & \dots & b_{tr} \end{pmatrix} \begin{pmatrix} a_{11} & \dots & a_{1j} & \dots & a_{1s} \\ a_{21} & \dots & a_{2j} & \dots & a_{2s} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{r1} & \dots & a_{rj} & \dots & a_{rs} \end{pmatrix} = \begin{pmatrix} c_{11} & \dots & c_{1s} \\ \vdots & c_{kj} & \vdots \\ c_{t1} & \dots & c_{ts} \end{pmatrix}.$$

Produkt zweier Vektoren

Als Spezialfall erhalten wir für das Produkt eines Zeilenvektors = Matrix $(1 \times r)$, mit einem Spaltenvektor, Matrix $(r \times 1)$, eine (1×1) Matrix, also einen Skalar:

$$\begin{pmatrix} b_1 & b_2 & \dots & b_r \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_r \end{pmatrix} = \sum_{i=1}^{r} b_i a_i = c.$$

Dieses Produkt wird auch als Skalarprodukt zweier Vektoren im $\mathbb{R}^r$ bezeichnet.

Produkt einer Matrix mit einem Vektor

Durch die Formel (2.17) wurde bereits die Matrix-Vektor Multiplikation eingeführt. Sie kann auch als Spezialfall der Matrizenmultiplikation (2.19) aufgefasst werden.
Sei B eine $(t \times r)$ Matrix und A eine $(r \times 1)$ Matrix (Spaltenvektor). Dann ist

$$\begin{array}{ccc} B & A & = C \\ (t \times r) & (r \times 1) & (t \times 1) \end{array}$$

ein Spaltenvektor ($(t \times 1)$ Matrix). Ausführlich geschrieben, heißt das

$$\begin{pmatrix} b_{11} & \dots & b_{1r} \\ \vdots & \vdots & \vdots \\ b_{k1} & \dots & b_{kr} \\ \vdots & \vdots & \vdots \\ b_{t1} & \dots & b_{tr} \end{pmatrix} \begin{pmatrix} a_1 \\ \vdots \\ \vdots \\ \vdots \\ a_r \end{pmatrix} = \begin{pmatrix} c_1 \\ \vdots \\ c_k \\ \vdots \\ c_t \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^r b_{1i} a_i \\ \vdots \\ \sum_{i=1}^r b_{ki} a_i \\ \vdots \\ \sum_{i=1}^r b_{ti} a_i \end{pmatrix}.$$

Die lineare Gruppe der Matrizen, inverse Matrizen

Wir haben gesehen, dass lineare Abbildungen $f : K^s \to K^r$ (wähle z.B. $K = \mathbb{R}$) durch $(r \times s)$ Matrizen darstellbar sind. Weiterhin haben wir eine Verknüpfung, die Multiplikation von $(t \times r)$ Matrizen mit $(r \times s)$ Matrizen, eingeführt.
Wir klären nun, ob die Menge der linearen Abbildungen von $K^s \to K^r$, bzw. die Menge der Matrizen $M(K, r \times s)$ ein Vektorraum oder sogar eine multiplikative Gruppe ist.

Lemma 2.26. *Die Menge der linearen Abbildungen von K^s in K^r, bzw. die Menge der Matrizen $M(K, r \times s)$ ist ein Vektorraum mit der Addition*

$$\begin{aligned} (f+g)(\boldsymbol{x}) &= f(\boldsymbol{x}) + g(\boldsymbol{x}) \quad \forall\, \boldsymbol{x} \in K^s, \\ A + B &:= (a_{ij} + b_{ij})_{\substack{1 \le i \le r \\ 1 \le j \le s}} \end{aligned}$$

und der Multiplikation mit Elementen aus K:

$$\begin{aligned} (\alpha f)(\boldsymbol{x}) :=& \alpha f(\boldsymbol{x}) \quad \forall\, \boldsymbol{x} \in K^s, \\ \alpha A :=& (\alpha a_{ij})_{\substack{1 \le i \le r \\ 1 \le j \le s}}. \end{aligned}$$

Das neutrale Element bezüglich der Addition ist die Nullabbildung $f(\boldsymbol{x}) = 0$ für alle $\boldsymbol{x} \in K^s$, bzw. die Nullmatrix $A = (a_{ij} = 0)_{\substack{1 \le i \le r \\ 1 \le j \le s}}$.

Wir sehen uns jetzt etwas genauer die Multiplikation in $M(K, r \times s)$ an:
Zunächst stellen wir fest. Die Komposition von Abbildungen

$$f : K^s \to K^r, g : K^r \to K^t, h : K^t \to K^l$$

ist assoziativ

$$h \circ (g \circ f) = (h \circ g) \circ f.$$

Daraus folgt, dass das Produkt von Matrizen $A \in M(K, r \times s)$, $B \in M(K, r \times s)$, $C \in M(K, r \times s)$ assoziativ ist,

$$C(BA) = (CB)A.$$

Um das Produkt in $M(K, r \times s)$ definieren zu können, nehmen wir an, dass $r = s$ ist. Dann sind AB und BA für $A, B \in M(K, r \times r)$ definiert. Jedoch ist die Multiplikation nicht immer kommutativ, d.h. es gibt Beispiele mit

$$AB \neq BA.$$

Beispiel

Es sei $r = 2$, $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$. Dann ist

$$
\begin{aligned}
AB &= \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \\
BA &= \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.
\end{aligned}
$$

Ein neutrales Element bezüglich der Multiplikation existiert in $M(K, r \times r)$. Es ist die Einheitsmatrix

$$
E_r = \begin{pmatrix} 1 & 0 & \dots & 0 \\ \vdots & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \dots & 0 & 1 \end{pmatrix}.
$$

Definition 2.27 (invertierbare Matrix). Man nennt eine Matrix $A \in M(K, r \times r)$ invertierbar, falls es eine Matrix $A^{-1} \in M(K, r \times r)$ gibt, so dass

$$
A^{-1}A = AA^{-1} = E_r
$$

ist. A^{-1} heißt inverse Matrix zu A.

Nicht alle Matrizen sind invertierbar.

Beispiel

Sei $r = 2$, $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ und $A^{-1} = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$. Dann müsste gelten

$$
\begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & b_{11} \\ 0 & b_{21} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}
$$

und

$$
\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} b_{21} & b_{22} \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},
$$

was nicht sein kann.

Lemma 2.28. *A ist invertierbar genau dann, wenn die lineare Abbildung $f : K^r \to K^r$*

$$
f_A(\boldsymbol{x}) = A\boldsymbol{x}
$$

bijektiv ist, d.h. $\ker f_A = \ker A = \{0\}$ ist.

Beweis.

a) Wir nehmen an, dass A^{-1} existiert. Falls $A\boldsymbol{x} = 0$ ist, folgt

$$A^{-1}A\boldsymbol{x} = \boldsymbol{x} = A^{-1}0 = 0$$

und $\ker A = \{0\}$.

b) Wir nehmen an, dass $\ker A = \{0\}$ ist. Nach dem Dimensionssatz 2.21 ist

$$r = \dim(\ker f_A) + \dim(\operatorname{im} f_A)$$

und daher $r = \dim f_A$.
Damit ist die Gleichung $A\boldsymbol{x} = \boldsymbol{b}$ für alle $\boldsymbol{b}$ aus K^r eindeutig lösbar, d.h. die inverse lineare Abbildung f_A^{-1} existiert, d.h.

$$f_A^{-1} \circ f_A(\boldsymbol{x}) = \boldsymbol{x} = f_A \circ f_A^{-1}(\boldsymbol{x}).$$

Die f_A^{-1} zugeordnete Matrix ist A^{-1}.

□

Lemma 2.29. *Die invertierbaren Matrizen aus $M(K, r \times r)$ bilden eine Gruppe bezüglich der Multiplikation mit dem neutralen Element E_r. Diese Gruppe heißt lineare Gruppe $GL_r(K)$.*

Der Beweis dieses Lemmas besteht darin, die Gruppenaxiome zu überprüfen. Es muss u.a. gezeigt werden: Falls $A, B \in GL_r(K)$, dann ist auch $AB \in GL_r(K)$ und

$$(AB)^{-1} = B^{-1}A^{-1}.$$

Bemerkung 2.30. Die Berechnung von inversen Matrizen ist keine einfache Aufgabe. Wir werden hierzu einen Algorithmus kennen lernen.

Wir fassen folgende Rechengesetze für Matrizen, die von ihrer Größe her zusammenpassen, zusammen:

$$\begin{aligned} A + B &= B + A, \\ (A + B)C &= AC + BC, \\ (AB)C &= A(BC), \\ AB &\neq BA \ \text{ (im Allgemeinen).} \end{aligned}$$

Basiswechsel, Koordinatentransformation

Am Beginn dieses Abschnittes hatten wir lineare Abbildungen eines K-Vektorraumes V mit der Basis $\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_r\}$ in den K-Vektorraum W mit der Basis $\{\boldsymbol{w}_1, \boldsymbol{w}_2, \dots, \boldsymbol{w}_s\}$ betrachtet. Wir betrachten nun einen Basiswechsel in V, d.h. in $V = W$ seien die Basen

$B_1 = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_r\}$ und $B_2 = \{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_r\}$ gegeben und wir möchten die lineare Abbildung beschreiben, die den Übergang zwischen den Basen angibt.
Für ein beliebiges $\boldsymbol{v} \in V$ gilt

$$\boldsymbol{v} = \sum_{i=1}^{r} \alpha_i \boldsymbol{v}_i = \sum_{i=1}^{r} \beta_i \boldsymbol{w}_i. \tag{2.20}$$

Sei H_{B_1} die $(r \times r)$ Matrix, die aus den Spaltenvektoren $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_r$ besteht und H_{B_2} die $(r \times r)$ Matrix, die die Spaltenvektoren $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_r$ besitzt. Dann können wir (2.20) schreiben als

$$\boldsymbol{v} = H_{B_1}\boldsymbol{\alpha} = H_{B_2}\boldsymbol{\beta}, \tag{2.21}$$

wobei $\boldsymbol{\alpha} = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_r \end{pmatrix}$, $\boldsymbol{\beta} = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_r \end{pmatrix}$ sind.
(2.21) liefert uns die Koordinatentransformation

$$\boldsymbol{\alpha} = H_{B_1}^{-1} H_{B_2}\boldsymbol{\beta}, \quad \boldsymbol{\beta} = H_{B_2}^{-1} H_{B_1}\boldsymbol{\alpha}. \tag{2.22}$$

Ist insbesondere $B_1 = \{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_r\}$ die Standardbasis, dann ist $H_{B_1} = E = E_r$ und

$$\boldsymbol{\alpha} = H_{B_2}\boldsymbol{\beta}, \tag{2.23}$$

$$\boldsymbol{\beta} = H_{B_2}^{-1}\boldsymbol{\alpha}. \tag{2.24}$$

Beispiel

Sei $V = W = \mathbb{R}^2$. Wir wollen die Koordinatentransformation von der Standardbasis $B_1 = \{\boldsymbol{e}_1, \boldsymbol{e}_2\}$ in die Basis $B_2 = \left\{ \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \end{pmatrix} \right\}$ beschreiben. Es ist $H_{B_2} = \begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix}$ und $H_{B_2}^{-1} \overset{(2.42)}{=} \frac{1}{5} \begin{pmatrix} 3 & -1 \\ -1 & 2 \end{pmatrix}$, wovon man sich durch Nachrechnen überzeugen kann.
Somit wird ein $\boldsymbol{v} \in V$, zunächst in der Standardbasis gegeben, in der Basis B_2 folgendermaßen dargestellt

$$\begin{aligned} \boldsymbol{v} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} &= x_1\boldsymbol{e}_1 + x_2\boldsymbol{e}_2 \overset{(2.21)}{=} \beta_1 \begin{pmatrix} 2 \\ 1 \end{pmatrix} + \beta_2 \begin{pmatrix} 1 \\ 3 \end{pmatrix} \\ &= \frac{1}{5}(3x_1 - x_2) \begin{pmatrix} 2 \\ 1 \end{pmatrix} + \frac{1}{5}(-x_1 + 2x_2) \begin{pmatrix} 1 \\ 3 \end{pmatrix}. \end{aligned}$$

Hier haben wir die Gleichung (2.24) benutzt:

$$\begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} = \frac{1}{5} \begin{pmatrix} 3 & -1 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \frac{1}{5}(3x_1 - x_2) \\ \frac{1}{5}(-x_1 + 2x_2) \end{pmatrix}.$$

Der Rang einer Matrix

Der Rang einer linearen Abbildung $f : V \to W$ war bereits eingeführt worden:

$$\operatorname{Rang} f = \dim(\operatorname{im} f) = \dim(f(V)).$$

Zu jeder Matrix $A = (a_{ij}) \in M(K, r \times s)$ gibt es eine lineare Abbildung f_A, so dass gilt

$$f_A : K^s \to K^r, \quad \boldsymbol{x} \to f_A(\boldsymbol{x}) = A\boldsymbol{x} = \begin{pmatrix} \sum_{j=1}^s a_{1j}x_j \\ \vdots \\ \sum_{j=1}^s a_{rj}x_j \end{pmatrix}.$$

Damit können wir definieren

$$\operatorname{Rang} A := \operatorname{Rang}(f_A).$$

Lemma 2.31. *Der Rang einer Matrix ist gleich der maximalen Anzahl linear unabhängiger Spaltenvektoren von A. Man spricht kurz vom Spaltenrang.*

Beweis. Es ist

$$\begin{aligned} A\boldsymbol{x} &= \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1s}x_s \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2s}x_s \\ \vdots \\ a_{r1}x_1 + a_{r2}x_2 + \cdots + a_{rs}x_s \end{pmatrix} \\ &= x_1 \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{r1} \end{pmatrix} + x_2 \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{r2} \end{pmatrix} + \cdots + x_s \begin{pmatrix} a_{1s} \\ a_{2s} \\ \vdots \\ a_{rs} \end{pmatrix} \end{aligned}$$

und daher

$$\begin{aligned} \operatorname{im} f_A &= \operatorname{Span}(\text{Spaltenvektoren}), \\ \dim(\operatorname{im} f_A) &= \text{maximale Anzahl linear unabhängiger Spaltenvektoren von } A \\ &= \text{Spaltenrang}. \end{aligned}$$

□

Beispiele

1° Für $A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$ ist $\operatorname{Rang} A = 3$.

2° Für $A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 2 & 0 & 2 \end{pmatrix}$ ist $\operatorname{Rang} A = 2$.

Es gilt jedoch überraschend:

Lemma 2.32. *Der Rang einer Matrix A ist gleich der maximalen Anzahl linear unabhängiger Zeilenvektoren, kurz als Zeilenrang bezeichnet.*

Beweis. A hat r Zeilenvektoren $\{z_1, \ldots, z_r\}$ der Länge s. Die maximale Anzahl der linear unabhängigen Zeilenvektoren sei $r_0 \leq r$. Wir bezeichnen sie mit

$$\boldsymbol{b}_j = (b_{j1}, \ldots, b_{js}), \quad 1 \leq j \leq r_0.$$

Die Zeilenvektoren $\boldsymbol{z}_i$, $i = 1, \ldots, r$, lassen sich als Linearkombination der $\boldsymbol{b}_j$ darstellen

$$\begin{aligned} \boldsymbol{z}_i &= k_{i1}\boldsymbol{b}_1 + k_{i2}\boldsymbol{b}_2 + \cdots + k_{ir_0}\boldsymbol{b}_{r_0} \\ &= (k_{i1}b_{11} + k_{i2}b_{21} + \cdots + k_{ir_0}b_{r_01}, \ldots, k_{i1}b_{1s} + \cdots + k_{ir_0}b_{r_0s}). \end{aligned} \tag{2.25}$$

Wir betrachten in jedem Zeilenvektor $\boldsymbol{z}_i$, $i = 1, \ldots, r$, den l-ten Eintrag, $1 \leq l \leq s$. Dann lautet (2.25):

$$\begin{aligned} a_{1l} &= k_{11}b_{1l} + \cdots + k_{1r_0}b_{r_0l} \\ &\vdots \\ a_{il} &= k_{i1}b_{1l} + \cdots + k_{ir_0}b_{r_0l} \\ &\vdots \\ a_{rl} &= k_{r1}b_{1l} + \cdots + k_{rr_0}b_{r_0l}. \end{aligned}$$

Damit ist jeder Spaltenvektor in der Form

$$\begin{pmatrix} a_{1l} \\ \vdots \\ a_{il} \\ \vdots \\ a_{rl} \end{pmatrix} = b_{1l} \begin{pmatrix} k_{11} \\ \vdots \\ k_{i1} \\ \vdots \\ k_{r1} \end{pmatrix} + \cdots + b_{r_0l} \begin{pmatrix} k_{1r_0} \\ \vdots \\ k_{ir_0} \\ \vdots \\ k_{rr_0} \end{pmatrix}$$

darstellbar. Es folgt, dass

$$\text{Rang}\, A = \text{ Spaltenrang} \leq r_0 = \text{ Zeilenrang}$$

ist. Das gleiche Verfahren können wir anwenden, indem wir die Zeilenvektoren durch Spaltenvektoren ersetzen. Es gilt dann

$$r_0 = \text{ Zeilenrang} \leq \text{ Spaltenrang} = \text{Rang}\, A.$$

Damit erhalten wir, dass Spaltenrang = Zeilenrang ist und damit

$$\begin{aligned} \text{Rang}\, A = &\text{ maximale Anzahl der linear unabhängigen Spaltenvektoren} \\ = &\text{ maximale Anzahl der linear unabhängigen Zeilenvektoren.} \end{aligned} \tag{2.26}$$

□

Übungsaufgaben

34) Es sei die folgende Menge $\mathcal{M}$ von Matrizen gegeben:

$$\mathcal{M} := \left\{ \begin{pmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{pmatrix} \middle| \, a, b, c \in \mathbb{R} \right\}.$$

1. Zeigen Sie: $\quad A, B \in \mathcal{M} \Rightarrow A \cdot B \in \mathcal{M}$.
2. Wann gilt für $A, B \in \mathcal{M}$: $\quad A \cdot B = B \cdot A$?
3. Für welche Matrizen $A \in \mathcal{M}$ gilt: $\quad A \cdot B = B \cdot A \quad \forall B \in \mathcal{M}$?
4. Wie lautet die inverse Matrix $A^{-1} \in \mathcal{M}$ zu einer beliebigen Matrix $A \in \mathcal{M}$?

35)

1. Seien V ein endlichdimensionaler Vektorraum über $\mathbb{R}$ oder $\mathbb{C}$ und $T : V \to V$ eine lineare Abbildung. Zeigen Sie:

   T ist genau dann injektiv, wenn T surjektiv ist.

2. Seien U und W zwei 4-dimensionale Unterräume eines Vektorraums V mit der Dimension 6. Welche Werte kann $\dim(U \cap W)$ annehmen? Man gebe zu jeder Möglichkeit ein Beispiel an.

36) Seien $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$ und $\boldsymbol{e}_4$ die Vektoren der Standardbasis des $\mathbb{R}^4$.

1. Beweisen Sie, dass auch

   $$B := \left\{ \boldsymbol{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \boldsymbol{v}_2 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \boldsymbol{v}_3 = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \boldsymbol{v}_4 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 2 \end{pmatrix} \right\}$$

   eine Basis des $\mathbb{R}^4$ ist.

2. Bestimmen Sie die Matrizen der Abbildungen $f, g : \mathbb{R}^4 \to \mathbb{R}^4$ mit $f(\boldsymbol{e}_i) = \boldsymbol{v}_i$ und $g(\boldsymbol{v}_i) = \boldsymbol{e}_i$.
3. Bestimmen Sie die Matrix der Abbildung $h : \mathbb{R}^4 \to \mathbb{R}^4$ mit $h\big((x_1, x_2, x_3, x_4)\big) = (x_2 + x_3, x_2, 0, 0)$ bezüglich der Standardbasis und die Matrix von h bezüglich B. Berechnen Sie den Rang der beiden Matrizen.

2.4 Lineare Gleichungssysteme

Die Theorie der linearen Abbildungen, die wir in den Abschnitten 2.2 und 2.3 kennen gelernt haben, findet eine schöne Anwendung bei der Lösung linearer Gleichungssysteme.

Definition 2.33. Unter einem linearen Gleichungssystem mit s Unbekannten und r Gleichungen verstehen wir das System

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1s}x_s &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2s}x_s &= b_2 \\ &\vdots \\ a_{r1}x_1 + a_{r2}x_2 + \cdots + a_{rs}x_s &= b_r, \end{aligned} \tag{2.27}$$

wobei die Koeffizientenmatrix $A = (a_{ij})_{\substack{1\le i\le r\\ 1\le j\le s}} \in M(K, r\times s)$ (K ist ein Körper, z.B. $K = \mathbb{R}$) und die rechte Seite

$$\boldsymbol{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_r \end{pmatrix} \in K^r$$

gegeben sind. Gesucht ist die Lösungsmenge L aller Vektoren $\boldsymbol{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_s \end{pmatrix} \in K^s$, die allen Gleichungen (2.27) genügen.

Das Gleichungssystem (2.27) kann kürzer als

$$A\boldsymbol{x} = \boldsymbol{b} \tag{2.28}$$

geschrieben werden. Ist $\boldsymbol{b} = \boldsymbol{0}$, dann spricht man von einem homogenen Gleichungssystem,

$$A\boldsymbol{x} = \boldsymbol{0},$$

für $\boldsymbol{b} \neq \boldsymbol{0}$ von einem inhomogenen Gleichungssystem.
Die Existenz und Eindeutigkeit der Lösungen des Gleichungssystems (2.27) wird durch folgenden Lösbarkeitssatz beschrieben.

Satz 2.34 (Existenz- und Eindeutigkeit für $r \times s$ Systeme).

1. *Das Gleichungssystem (2.27) ist genau dann lösbar, wenn* $\boldsymbol{b}$ *im Bild der linearen Abbildung*
$$f_A : K^s \to K^r, \quad x \to f_A(\boldsymbol{x}) = A\boldsymbol{x}$$
liegt, d.h.
$$\operatorname{Rang} A = \operatorname{Rang}(A|\boldsymbol{b}),$$

wobei $A|\boldsymbol{b}$ die erweiterte $(r \times (s+1))$ Koeffizienten-Matrix ist:

$$A|\boldsymbol{b} = \left(\begin{array}{ccc|c} a_{11} & \dots & a_{1s} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{r1} & \dots & a_{rs} & b_s \end{array}\right).$$

2. *Das Gleichungssystem (2.27) ist für alle $\boldsymbol{b} \in K^r$ lösbar, falls gilt*

$$\operatorname{Rang} A = r.$$

3. *Ist das Gleichungssystem (2.27) lösbar, dann ist eine Lösung $\boldsymbol{x}$ eindeutig bestimmt, falls gilt*

$$\operatorname{Rang} A = s.$$

Beweis.

zu 1. Das Bild der linearen Abbildung f_A ist

$$\operatorname{im}(f_A) = \{\boldsymbol{b} \in K^r \,:\, \exists\, \boldsymbol{x} \in K^s \text{ mit } A\boldsymbol{x} = \boldsymbol{b}\}.$$

Nach Lemma 2.31 ist

$$\operatorname{im}(f_A) = \operatorname{Span}\{\text{linear unabhängige Spaltenvektoren von } A\}.$$

Daher muss eine gegeben rechte Seite $\boldsymbol{b}$ von (2.27) durch eine Linearkombination der linear unabhängigen Spaltenvektoren von A darstellbar sein, woraus

$$\operatorname{Rang} A = \operatorname{Rang}(A|\boldsymbol{b})$$

folgt.

zu 2. In diesem Fall ist $\operatorname{im}(f_A) = K^r$.

zu 3. Nach dem Dimensionssatz 2.21 ist

$$s = \dim(\ker f_A) + \operatorname{Rang} A.$$

Ist $\operatorname{Rang} A = s$, dann ist $\dim(\ker f_A) = 0$ und $\ker f_A = \{0\}$. Daraus folgt die Eindeutigkeit.

□

Folgerung 2.35. Falls $r > s$, dann ist (2.27) nicht für alle $\boldsymbol{b} \in K^r$ lösbar, falls $r < s$, dann besitzt (2.27) keine eindeutig bestimmte Lösung.

Im Fall $r = s$ kann man die eindeutige Lösbarkeit einfacher charakterisieren.

Satz 2.36 (Eindeutige Lösbarkeit von $(r \times r)$ **Systemen).** *Sei A eine $(r \times r)$ Matrix. Dann sind folgende Aussagen äquivalent.*

(i) A ist invertierbar, d.h. A^{-1} existiert.

(ii) Für alle $\boldsymbol{b} \in K^r$ ist $A\boldsymbol{x} = \boldsymbol{b}$ eindeutig lösbar.

(iii) Das homogene Gleichungssystem $A\boldsymbol{x} = \mathbf{0}$ besitzt nur die triviale Lösung $\boldsymbol{x} = \mathbf{0}$.

Beweis. Wir zeigen $(i) \Rightarrow (ii) \Rightarrow (iii) \Rightarrow (i)$.

a) Es gelte (i). Wir betrachten das System $A\boldsymbol{x} = \boldsymbol{b}$. Dann ist

$$A^{-1}A\boldsymbol{x} = \boldsymbol{x} = A^{-1}\boldsymbol{b}$$

und (ii) ist gezeigt.

b) Es gelte (ii). Dann besitzt auch das homogene System $A\boldsymbol{x} = \mathbf{0}$ nur eine Lösung $\boldsymbol{x} = \mathbf{0}$.

c) Es gelte (iii). Nach dem Dimensionssatz 2.21 ist

$$r = \dim(\ker f_A) + \operatorname{Rang} A = \operatorname{Rang} A.$$

Aus Satz 2.34, Aussagen 2 und 3 folgt, dass (2.27) für jede rechte Seite eindeutig lösbar ist. Damit muss

$$\boldsymbol{x} = A^{-1}\boldsymbol{b}$$

sein und (i) gilt.

$\square$

Das Gaußsche Eliminationsverfahren

Nach den theoretischen Sätzen 2.34 und 2.36 wenden wir uns jetzt einem praktischen Lösungsweg zu, der von Carl Friedrich Gauß (1777 - 1855) vorgeschlagen wurde. Dabei wird die erweiterte Koeffizientenmatrix $(A|\boldsymbol{b})$ in eine Matrix besonders einfacher Gestalt, in eine sogenannte Zeilenstufenmatrix, umgewandelt. Aus der Zeilenstufenmatrix lässt sich die Lösung leicht gewinnen, der Rang kann abgelesen werden und es ist damit zu sehen, ob das Gleichungssystem überhaupt eine Lösung besitzt.
Die Umwandlung der erweiterten Koeffizientenmatrix in eine Zeilenstufenform erfolgt durch elementare Zeilenoperationen. Dabei verstehen wir unter einer Zeilenstufenmatrix folgendes:

Definition 2.37. Sei $A = (a_{ij}) \in M(K, r \times s)$. Sei I_i die Anzahl der aufeinander folgenden Nullen der i-ten Zeile:

$$I_i = \begin{cases} \max\{I : a_{i1} = a_{i2} = \cdots = a_{iI} = 0\} & (1 \leq i \leq r), \\ 0 & \text{falls } a_{i1} \neq 0. \end{cases}$$

Ist $I_1 < I_2 < \cdots < I_r$ oder $I_1 < I_2 < \cdots < I_t < I_{t+1} = \cdots = I_r = s$ für ein $t < r$ (d.h. ab der $(t+1)$-ten Zeile treten nur Nullzeilen auf), dann heißt A Zeilenstufenmatrix.

Beispiele:

1° $A = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix}$ ist Zeilenstufenmatrix. Es ist $I_1 = 0, I_2 = 3, I_3 = 4$.

2° $A = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 2 & 3 \end{pmatrix}$ ist keine Zeilenstufenmatrix, denn $I_1 = 0, I_2 = 2, I_3 = 2$.

Hat die Koeffizientenmatrix eines linearen Gleichungssystems Zeilenstufenform, dann kann die Lösungsmenge explizit angegeben werden:

1. In den Zeilen, wo nur Nullen stehen, müssen die Einträge von $\boldsymbol{b}$ verschwinden, sonst existiert keine Lösung. Diese können dann gestrichen werden.

2. In den verbleibenden Zeilen beginnen wir dann mit der t-ten Zeile, $t \leq r$, mit $I_1 < I_2 < \cdots < I_t \neq s, I_{t+1} = s$. Die Zeilengleichung lautet

$$\sum_{I_t < j \leq s} a_{tj} x_j = b_t. \tag{2.29}$$

   Hier ist $a_{tj} \neq 0$ für $j = I_t + 1$ und wir können alle Lösungen $\begin{pmatrix} x_{I_t+1} \\ \vdots \\ x_s \end{pmatrix}$ angeben.

3. Wir setzen eine dieser Lösungen in die $(t-1)$-te Zeilengleichung ein

$$\sum_{I_{t-1} < j \leq s} a_{(t-1)j} x_j = b_{t-1}. \tag{2.30}$$

   Wegen $I_{t-1} < I_t$ ist (2.30) eine Gleichung zur Bestimmung der noch freien Unbekannten $x_{I_{t-1}+1}, \ldots, x_{I_t}$.

4. Die Vektoren $\begin{pmatrix} x_{I_{t-1}+1} \\ \vdots \\ x_{I_t} \\ x_{I_t+1} \\ \vdots \\ x_s \end{pmatrix}$ werden in die $(t-2)$-te Zeilengleichung eingesetzt. Daraus wird der Vektor
$$\begin{pmatrix} x_{I_{t-2}+1} \\ \vdots \\ x_{I_{t-1}} \\ x_{I_{t-1}+1} \\ \vdots \\ x_{I_t} \\ x_{I_t+1} \\ \vdots \\ x_s \end{pmatrix}$$
gewonnen.

5. Dieser Prozess wird fortgesetzt, bis die erste Zeile erreicht ist. Daraus erhalten wir eine Lösung $\boldsymbol{x}$.

Beispiel:

Wir betrachten das Gleichungssystem

$$A\boldsymbol{x} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 = 0 \end{pmatrix}. \tag{2.31}$$

Erster Schritt (Wegstreichen der Nullzeilen).
Durch Wegstreichen der Nullzeile erhalten wir

$$\hat{A}\boldsymbol{x} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}.$$

Zweiter Schritt.
Wir betrachten die zweite Zeile ($t = 2$):

$$2x_4 = b_2 \quad \Rightarrow \quad x_4 = \frac{b_2}{2}.$$

Dritter Schritt.
Einsetzen von x_4 in die erste Zeilengleichung

$$x_1 + x_2 + x_3 + \frac{b_2}{2} = b_1 \tag{2.32}$$

ergibt

$$x_1 + x_2 + x_3 = b_1 - \frac{b_2}{2}.$$

Damit ist

$$x_1 = b_1 - \frac{b_2}{2} - x_2 - x_3$$

und

$$\begin{pmatrix} b_1 - \frac{b_2}{2} - 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} b_1 - \frac{b_2}{2} - 1 \\ 1 \\ 0 \end{pmatrix}$$

sind Lösungen von (2.32).
Es folgt, dass

$$\begin{pmatrix} b_1 - \frac{b_2}{2} - 1 \\ 0 \\ 1 \\ \frac{b_2}{2} \end{pmatrix}, \begin{pmatrix} b_1 - \frac{b_2}{2} - 1 \\ 1 \\ 0 \\ \frac{b_2}{2} \end{pmatrix} \tag{2.33}$$

Lösungen von $A\boldsymbol{x} = \boldsymbol{b}$ sind.

Die Frage ob wir alle Lösungen von $A\boldsymbol{x} = \boldsymbol{b}$ erfasst haben, diskutieren wir später. Wir erklären jetzt, was wir unter elementaren Zeilenoperationen verstehen, die eine Matrix A in eine Zeilenstufenform überführen.

Definition 2.38. Sei $A|\boldsymbol{b}$ die erweiterte Matrix zum Gleichungssystem

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

Elementare Zeilenumformungen von $A|\boldsymbol{b}$ sind:

- Vertauschungen von Zeilen,
- Multiplikation von Zeilen mit $\alpha \in K\backslash\{0\}$,
- Addition einer α- fachen Zeile zu einer andere Zeile.

Wir sehen sofort: die Lösungsmenge L des linearen Gleichungssystems ändert sich nicht durch Anwendung elementarer Zeilenumformungen. So ist z.B. $\boldsymbol{x}$ Lösung von $A\boldsymbol{x} = \boldsymbol{b}$, genau dann falls $\boldsymbol{x}$ Lösung des folgende Gleichungsystems ist:

$$\sum_{j=1}^{s} a_{ij}x_j = b_j, \qquad 1 \leq i \leq r, i \neq l,$$

$$\sum_{j=1}^{s} (a_{lj} + \alpha a_{kl})x_j = b_l + \alpha b_k.$$

Lemma 2.39 (Gauß Elimination). *Jede Matrix $A \in M(K, r\times s)$ lässt sich durch endlich viele elementare Zeilenoperationen auf Zeilenstufenform bringen.*

Beweis.
1.Schritt: Erzeugen von $a_{11} \neq 0$.
Wir dürfen annehmen, dass die erste Spalte von A mindestens ein Element $\neq 0$ enthält (sonst gehen wir zur ersten Spalte, die von Null verschieden ist, über). Durch Vertauschen der Zeilen erreichen wir, dass $a_{11} \neq 0$ ist.
2. Schritt: Erzeugen von Nullen (außer a_{11}) in der ersten Spalte.
Wir bestimmen Elemente $\alpha_i \in K$, so dass gilt

$$\alpha_i a_{11} + a_{i1} = 0, \quad 2 \leq i \leq r. \tag{2.34}$$

Durch addieren des α_i-fachen der ersten Zeile zur i-ten Zeile erhalten wir

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1s} \\ 0 & \alpha_2 a_{12} + a_{22} & \dots & \alpha_2 a_{1s} + a_{2s} \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \alpha_r a_{12} + a_{r2} & \dots & \alpha_r a_{1s} + a_{rs} \end{pmatrix}. \tag{2.35}$$

3. Schritt: Erzeugen von Nullen in der zweiten Spalten von (2.35).
Wir unterscheiden 2 Fälle:
Sind alle Einträge unter a_{12} in der zweiten Spalte von (2.35) Null, so gehen wir zur 3. Spalte über.
Ist zumindest ein Eintrag unter a_{12} in der zweiten Spalte von (2.35) nicht Null, so setzen wir ihn durch Vertauschen der Zeilen unter a_{12}. Dann gehen wir wie oben vor. Wir erhalten

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1s} \\ 0 & \alpha_2 a_{12} + a_{22} & \dots & \dots & \alpha_2 a_{1s} + a_{2s} \\ 0 & 0 & * & \dots & * \\ 0 & 0 & * & \dots & * \\ \vdots & \vdots & & & \vdots \\ 0 & 0 & \dots & \dots & * \end{pmatrix}.$$

Dieses Verfahren wird so oft angewandt, bis die Zeilenstufenform erreicht ist. □

Beispiel

Wir betrachten die „erweiterte“ Matrix

$$\left(\begin{array}{ccc|c} 1 & 0 & 0 & 2 \\ -1 & 3 & 2 & 0 \\ 2 & 1 & -2 & 1 \end{array}\right). \tag{2.36}$$

a_{11} ist von Null verschieden. Wir bestimmen α_2 und α_3:

$$\begin{aligned} \alpha_2 a_{11} + a_{21} = \quad &\Rightarrow \ \alpha_2 = -\frac{a_{21}}{a_{11}} = 1, \\ \alpha_3 a_{11} + a_{31} = 0 &\Rightarrow \ \alpha_3 = -\frac{a_{31}}{a_{11}} = -2. \end{aligned}$$

Damit erhalten wir nach Addition des α_2-fachen der ersten Zeile zur zweiten Zeile und des α_3-fachen der ersten Zeile zur dritten Zeile

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

Wir bestimmen jetzt eine Zahl μ, so dass gilt

$$\mu 3 + 1 = 0, \qquad \mu = -\frac{1}{3},$$

und addieren das μ-fache der zweiten Zeile zur dritten Zeile:

$$\left(\begin{array}{ccc|c} 1 & 0 & 0 & 2 \\ 0 & 3 & 2 & 2 \\ 0 & 0 & -\frac{8}{3} & -\frac{11}{3} \end{array}\right). \tag{2.37}$$

Die Lösungsmenge von (2.36)

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

stimmt mit der Lösungsmenge von (2.37)

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 2 \\ 0 & 0 & -\frac{8}{3} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \\ -\frac{11}{3} \end{pmatrix} \tag{2.38}$$

überein. Wir erhalten sofort aus (2.38)

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2 \\ -\frac{1}{4} \\ \frac{11}{8} \end{pmatrix}.$$

Bemerkung 2.40.

(*i*) Das Gaußsche Eliminationsverfahren, das eine Matrix in eine Zeilenstufenmatrix überführt, liefert gleichzeitig den Rang einer Matrix.
Der Rang einer Zeilenstufenmatrix ist gleich der Anzahl der Zeilen, die einen von Null verschiedenen Eintrag enthalten.

(*ii*) Außerdem gilt: Der Rang einer Matrix ändert sich nicht bei elementaren Zeilenoperationen. Diese Aussage folgt aus Lemma 2.32.

Die vollständige Lösung linearer Gleichungssysteme

Wir diskutieren jetzt, wie man die vollständige Lösungsmenge eines linearen Gleichungssystem

$$A\boldsymbol{x} = \boldsymbol{b}, \quad A \in M(K, r \times s),\ \boldsymbol{b} \in K^r$$

beschreiben kann.

Satz 2.41. *Sei $\boldsymbol{x}_0 \in K^s$ eine spezielle Lösung von $A\boldsymbol{x} = \boldsymbol{b}$. Jede Lösung $\boldsymbol{x}$ von $A\boldsymbol{x} = \boldsymbol{b}$ hat die Gestalt*

$$\boldsymbol{x} = \boldsymbol{y} + \boldsymbol{x}_0,$$

wobei $\boldsymbol{y} \in \ker(A) = \{\boldsymbol{y} : A\boldsymbol{y} = \boldsymbol{0}\}$ ist.
Kurz ausgedrückt: $\{\boldsymbol{x} \in K^s : A\boldsymbol{x} = \boldsymbol{b}\} = L = \{\boldsymbol{x}_0 + \boldsymbol{y} : A\boldsymbol{y} = \boldsymbol{0}\}$.

Beweis.

a) Seien $\boldsymbol{x}$ und $\boldsymbol{x}' \in L$. Dann ist $A(\boldsymbol{x} - \boldsymbol{x}') = \boldsymbol{0}$, d.h. $\boldsymbol{x}$ ist darstellbar als

$$\begin{aligned} \boldsymbol{x} &= \boldsymbol{x} - \boldsymbol{x}' + \boldsymbol{x}' = \boldsymbol{y} + \boldsymbol{x}_0, \\ \boldsymbol{y} &= \boldsymbol{x} - \boldsymbol{x}' \in \ker A, \\ \boldsymbol{x}' &= \boldsymbol{x}_0 \ \text{ ist Lösung von } A\boldsymbol{x}_0 = \boldsymbol{b}. \end{aligned}$$

b) Sei $\boldsymbol{x} = \boldsymbol{y} + \boldsymbol{x}_0$. Dann ist

$$A\boldsymbol{x} = A\boldsymbol{y} + A\boldsymbol{x}_0 = \boldsymbol{0} + \boldsymbol{b} = \boldsymbol{b}.$$

□

Wir beschreiben nun, wie man vorgehen kann, um die vollständige Lösung eines Gleichungssystems $A\boldsymbol{x} = \boldsymbol{b}$ zu erhalten.

1. Bringe $A|\boldsymbol{b}$ in eine Zeilenstufenform.

2. Prüfe ob $\mathrm{Rang}(A) = \mathrm{Rang}(A|\boldsymbol{b})$ ist.
   Falls nein, dann existiert keine Lösung. Falls ja, konstruiere aus der Zeilenstufenmatrix eine spezielle Lösung $\boldsymbol{x}_0$.

3. Bestimme $\ker(A)$. Es ist nach der Dimensionsformel 2.21 $\dim(\ker(A)) = s - \mathrm{Rang}\, A$.
   Aus der Zeilenstufenform von A kann die Basis von $\ker(A)$ bestimmt werden.

Beispiel

Wir sehen uns noch einmal das Beispiel (2.31) an:

$$\begin{aligned} x_1 + x_2 + x_3 + x_4 &= b_1 \\ 2x_4 &= b_2. \end{aligned}$$

1.Schritt: Zeilenstufenform von $A|\boldsymbol{b}$.

Es ist

$$A|\boldsymbol{b} = \begin{pmatrix} 1 & 1 & 1 & 1 & b_1 \\ 0 & 0 & 0 & 2 & b_2 \end{pmatrix}$$

und damit liegt bereits eine Zeilenstufenmatrix vor.

2.Schritt: Rangbestimmung und Konstruktion einer speziellen Lösung.

Es ist Rang $A =$ Rang $A|\boldsymbol{b} = 2$. Somit ist

$$\boldsymbol{x}_0 = \begin{pmatrix} b_1 - \frac{b_2}{2} \\ 0 \\ 0 \\ \frac{b_2}{2} \end{pmatrix}$$

eine spezielle Lösung der inhomogenen Gleichung.

3. Schritt: Berechnung von $\ker A$.

Wir sehen, dass $\dim(\ker(A)) = s - \text{Rang}\, A = 4 - 2 = 2$ ist. Daher müssen wir zwei linear unabhängige Lösungen von

$$\begin{aligned} x_1 + x_2 + x_3 + x_4 &= 0 \\ 2x_4 &= 0 \end{aligned}$$

bestimmen. Da $x_4 = 0$ sein muss, folgt $x_1 + x_2 + x_3 = 0$. Es sind z.B.

$$\boldsymbol{y}_1 = \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix}, \quad \boldsymbol{y}_2 = \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix}$$

zwei linear unabhängige Lösungen.
Die Menge aller Lösungen ist

$$\boldsymbol{x} = c_1 \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix} + c_2 \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix} + \begin{pmatrix} b_1 - \frac{b_2}{2} \\ 0 \\ 0 \\ \frac{b_2}{2} \end{pmatrix},$$

wobei c_1 und c_2 beliebige Konstanten aus K sind. Für $c_1 = 0, c_2 = -1$ und für $c_1 = -1, c_2 = 0$ erhalten wir die Lösungen (2.33).

Berechnung inverser Matrizen

Wir betrachten ein $r \times r$ Gleichungssystem $A\boldsymbol{x} = \boldsymbol{b}$. In Satz 2.36 wurde geklärt, wann eine inverse Matrix A^{-1} existiert ($\ker A = \{\boldsymbol{0}\}$). In diesem Fall besitzt $A\boldsymbol{x} = \boldsymbol{b}$ eine eindeutig bestimmt Lösung

$$\boldsymbol{x} = A^{-1}\boldsymbol{b}.$$

Ist A^{-1} bekannt, dann kann die Lösung $\boldsymbol{x}$ für beliebige rechte Seiten $\boldsymbol{b}$ durch Matrix-Vektor-Multiplikation berechnet werden. Um A^{-1} zu berechnen, könnte man folgendermaßen vorgehen:
Bestimme eine Matrix $B \in M(K, r \times r)$, so dass

$$BA = E_r$$

ist. Dies führt zu einem Gleichungssystem von r^2 Gleichungen zur Bestimmung der r^2 Unbekannten $(b_{ij})_{\substack{1\le i\le r\\ 1\le j\le r}}$. Wir können aber auch das Gaußsche Eliminationsverfahren anwenden, um effizienter zum Ziel zu kommen. Sei A eine $(r \times r)$-Matrix mit $\mathrm{Rang}(A) = r$.

1.Schritt

Wir betrachten die erweiterte $(r \times 2r)$-Matrix

$$(A|E_r)$$

und bringen diese durch elementare Zeilenumformungen auf eine Zeilenstufenform. Wir erhalten

$$(SA|S) = \left(\begin{array}{ccc|c} c_1 & & * & \\ & \ddots & & S \\ 0 & & c_r & \end{array}\right),$$

wobei S eine Matrix ist, die die Zeilenumformungen beschreibt. Da $\mathrm{Rang}(A) = r$ ist, müssen alle Diagonalelemente $c_i \neq 0$ sein.

2.Schritt

Wir multiplizieren die Matrix $SA|S$ mit der Diagonalmatrix

$$C = \begin{pmatrix} c_1^{-1} & & 0 \\ & \ddots & \\ 0 & & c_r^{-1} \end{pmatrix},$$

d.h. die i-te Zeile wird mit c_i^{-1} multipliziert. Wir erhalten

$$(CSA|CS) = \left(\begin{array}{ccc|c} 1 & & * & \\ & \ddots & & CS \\ 0 & & 1 & \end{array}\right). \tag{2.39}$$

3.Schritt

Wir bringen CSA durch elementare Zeilenoperationen in die Einheitsmatrix E_r, d.h. es existiert eine Matrix $\hat{S}$ mit $\hat{S}CSA = E_r$. Es wird

$$(\hat{S}CSA|\hat{S}CS) = (E_r|\hat{S}CS)$$

und $\hat{S}CS = A^{-1}$.
Zusammenfassend können wir diesen Prozess folgendermaßen beschreiben:

Transformieren wir A durch elementare Zeilenoperationen in E_r, dann wird aus E_r die inverse Matrix A^{-1}

$$\begin{array}{c}(A|E_r)\\ \downarrow \text{ Zeilenoperationen}\\ (E_r|A^{-1}).\end{array}$$

Beispiel

Sei $A = (a_{ij})_{\substack{1\leq i\leq 2\\ 1\leq j\leq 2}}$ eine (2×2)-Matrix.

1.Schritt

Wir betrachten die erweiterte (2×4)-Matrix

$$A|E_2 = \left(\begin{array}{cc|cc} a_{11} & a_{12} & 1 & 0\\ a_{21} & a_{22} & 0 & 1\end{array}\right).$$

Wir nehmen an, dass $\underline{a_{11} \neq 0}$ und bringen $A|E_2$ in eine Zeilenstufenform. Wir bestimmen nach (2.34) α_2 so, dass

$$\alpha_2 a_{11} + a_{21} = 0, \quad \alpha_2 = -\frac{a_{21}}{a_{11}}$$

ist. Durch Addition des α_2-faches der 1. Zeile zur 2. Zeile erhalten wir

$$SA|S = \left(\begin{array}{cc|cc} a_{11} & a_{12} & 1 & 0\\ 0 & \frac{-a_{21}a_{12}+a_{22}a_{11}}{a_{11}} & -\frac{a_{21}}{a_{11}} & 1\end{array}\right). \tag{2.40}$$

Wir können annehmen, dass $\underline{a_{11}a_{22} - a_{12}a_{21} = D \neq 0}$ ist, sonst wäre Rang $A \neq 2$.

2.Schritt

Die Matrix C lautet:

$$C = \begin{pmatrix} \frac{1}{a_{11}} & 0\\ 0 & \frac{a_{11}}{D}\end{pmatrix}.$$

Nach Multiplikation von (2.40) mit C ergibt sich

$$CSA|CS = \left(\begin{array}{cc|cc} 1 & \frac{a_{12}}{a_{11}} & \frac{1}{a_{11}} & 0\\ 0 & 1 & -\frac{a_{21}}{D} & \frac{a_{11}}{D}\end{array}\right). \tag{2.41}$$

3.Schritt

Wir bringen $\begin{pmatrix} 1 & \frac{a_{12}}{a_{11}}\\ 0 & 1\end{pmatrix}$ in die Einheitsmatrix, indem wir das $-\frac{a_{12}}{a_{11}}$-fache der 2. Zeile zur 1. Zeile addieren. Dann ist

$$\hat{S}CSA|\hat{S}CS = \left(\begin{array}{cc|cc} 1 & 0 & \frac{1}{a_{11}} + \frac{a_{12}a_{21}}{a_{11}D} & -\frac{a_{12}}{D}\\ 0 & 1 & -\frac{a_{21}}{D} & \frac{a_{11}}{D}\end{array}\right),$$

und

$$\begin{aligned} A^{-1} &= \frac{1}{D}\begin{pmatrix} \frac{D+a_{12}a_{21}}{a_{11}} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix} \\ &= \frac{1}{D}\begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}. \end{aligned}$$

Wir fassen zusammen: Sei $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$. Ist $D = a_{11}a_{22} - a_{12}a_{21} \neq 0$, dann existiert die inverse Matrix und es ist

$$A^{-1} = \frac{1}{D}\begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}. \tag{2.42}$$

Übungsaufgaben

37) Der **Rang** einer Matrix A (bzw. einer linearen Abbildung) ist definiert als die Dimension des Bilds der zugehörigen linearen Abbildung.

1. Man gebe die Dimension und eine Basis des Lösungsvektorraums L des folgenden Gleichungssystems an:

$$\begin{array}{ccccccccccc} x_1 & + & 2x_2 & + & 2x_3 & - & x_4 & + & 3x_5 & = & 0 \\ x_1 & + & 2x_2 & + & 3x_3 & + & x_4 & + & x_5 & = & 0 \\ 3x_1 & + & 6x_2 & + & 8x_3 & + & x_4 & + & 5x_5 & = & 0\,. \end{array}$$

   Was ist der Rang obiger Koeffizientenmatrix?

2. Man gebe ein homogenes lineares Gleichungssystem an, dessen Lösungsvektorraum L durch

$$\begin{pmatrix} 1 \\ -2 \\ 0 \\ 3 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ -1 \\ -1 \\ 4 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ 0 \\ -2 \\ 5 \end{pmatrix}$$

   aufgespannt wird. Bestimmen Sie den Rang der Koeffizientenmatrix.

38)

1. Eine $n \times n$ Matrix $A = (a_{ij})$ heißt **obere Dreiecksmatrix**, wenn $a_{ij} = 0$ für alle $i > j$ und **untere Dreiecksmatrix**, wenn $a_{ij} = 0$ für alle $i < j$. Zeigen Sie, dass das Produkt von oberen (bzw. unteren) Dreiecksmatrizen wieder eine obere (bzw. untere) Dreiecksmatrix ist.

2. Die elementaren Zeilenumformungen des Gaußschen Algorithmus

   (a) addieren des λ-fachen der Zeile i ($\lambda \neq 0$) zur Zeile j,

   (b) multiplizieren der Zeile i mit $\lambda \neq 0$,

   (c) vertauschen der Zeilen i und j

lassen sich jeweils durch Matrixmultiplikation von links mit einer so genannten Elementarmatrix $B_{(a),i,j,\lambda}$, $B_{(b),i,\lambda}$ und $B_{(c),i,j}$ darstellen. Finden Sie für jede dieser Umformungen die zugehörige Elementarmatrix $B_{(a),i,j,\lambda}$, $B_{(b),i,\lambda}$ bzw. $B_{(c),i,j}$.

3. **LU Zerlegung.** Sei A eine $n \times n$ Matrix und

$$A = LU$$

mit einer unteren $n \times n$ Dreiecksmatrix $L = (l_{ij})$ mit $l_{ii} = 1$ und einer oberen $n \times n$ Dreiecksmatrix U, dann heißt das Paar L, U die **LU Zerlegung** von A.

Ist A eine $n \times n$ Matrix mit $\mathrm{Rang}\, A = n$, die durch Gauß-Elimination ohne Zeilenvertauschungen auf obere Dreicksform gebracht werden kann, so existiert eine LU Zerlegung von A. In diesem Fall braucht man für den Gaußschen Algorithmus nur Umformungen vom Typ a) und wenn B_i die zugehörigen Elementarmatrizen bezeichnen, gilt

$$B_r B_{r-1} \dots B_2 B_1 A = U\,, \qquad \text{also } A = B_1^{-1} B_2^{-1} \dots B_{r-1}^{-1} B_r^{-1} U.$$

Begründen Sie, dass dies eine LU Zerlegung ist und leiten Sie hieraus einen Algorithmus zur Berechnung der LU Zerlegung her. Begründen Sie, warum Ihr Algorithmus das richtige Ergebnis liefert.

39) Berechnen Sie die Lösungsmengen folgender Gleichungssysteme und bestimmen Sie den Rang der Koeffizientenmatrix:

$$\text{i)}\ \begin{array}{rcrcrcr} x_1 & + & 2x_2 & + & 3x_3 & = & 1 \\ 4x_1 & + & 5x_2 & - & 6x_3 & = & 2 \\ 7x_1 & + & 8x_2 & + & 9x_3 & = & 3 \\ 5x_1 & + & 7x_2 & + & 9x_3 & = & 4 \end{array}, \qquad \text{ii)}\ \begin{array}{rcrcrcr} x_1 & + & x_2 & + & x_3 & = & 3 \\ 2x_1 & & & + & x_3 & = & 3 \\ -x_1 & + & x_2 & & & = & 1 \\ & & 2x_2 & + & x_3 & = & 4 \\ 2x_1 & + & 2x_2 & + & 2x_3 & = & 7 \end{array}$$

$$\text{iii)}\ \begin{array}{rcrcrcrcr} x_1 & - & 2x_2 & + & x_3 & - & 4x_4 & = & 4 \\ & & x_2 & - & x_3 & + & x_4 & = & -3 \\ x_1 & + & 3x_2 & & & - & 3x_4 & = & 1 \\ & - & 7x_2 & + & 3x_3 & + & x_4 & = & -3 \end{array}$$

40) Bestimmen Sie die Inversen der Matrizen, falls diese existieren:

$$\begin{pmatrix} -2 & 3 & 1 \\ 3 & 6 & 2 \\ 1 & 2 & 1 \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} 1 & 1 & 1 & 0 \\ -1 & 2 & 1 & 0 \\ 1 & 4 & 1 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}.$$

2.5 Determinante einer Matrix

Bei der Berechnung der Inversen einer (2×2)-Matrix hat sich die Zahl $D = a_{11}a_{22} - a_{12}a_{21}$ als wichtige Größe herausgestellt. Es wird sich erweisen, dass sie die Determinante der Matrix A ist.
Determinanten werden für quadratische Matrizen folgendermaßen definiert.

Definition 2.42. Eine Abbildung f von $M(K, r\times r) \to K$, $f : A \to f(A) = \det A$ heißt Determinante, falls folgende Eigenschaften gelten:

(D1) det ist linear in den Zeilenvektoren, d.h. seien $\boldsymbol{z}_1, \ldots, \boldsymbol{z}_r$ die Zeilenvektoren von A, $\boldsymbol{b}$ ein beliebiger Zeilenvektor und $\alpha, \beta \in K$. Dann ist

$$\det \begin{pmatrix} \boldsymbol{z}_1 \\ \vdots \\ \boldsymbol{z}_{k-1} \\ \alpha\boldsymbol{z}_k + \beta\boldsymbol{b} \\ \boldsymbol{z}_{k+1} \\ \vdots \\ \boldsymbol{z}_r \end{pmatrix} = \alpha \det \begin{pmatrix} \boldsymbol{z}_1 \\ \vdots \\ \boldsymbol{z}_{k-1} \\ \boldsymbol{z}_k \\ \boldsymbol{z}_{k+1} \\ \vdots \\ \boldsymbol{z}_r \end{pmatrix} + \beta \det \begin{pmatrix} \boldsymbol{z}_1 \\ \vdots \\ \boldsymbol{z}_{k-1} \\ \boldsymbol{b} \\ \boldsymbol{z}_{k+1} \\ \vdots \\ \boldsymbol{z}_r \end{pmatrix}$$

.

(D2) $\det A$ ändert bei Vertauschung zweier Zeilen von A das Vorzeichen, d.h.

$$\det \begin{pmatrix} \boldsymbol{z}_1 \\ \vdots \\ \boldsymbol{z}_j \\ \vdots \\ \boldsymbol{z}_k \\ \vdots \\ \boldsymbol{z}_r \end{pmatrix} = -\det \begin{pmatrix} \boldsymbol{z}_1 \\ \vdots \\ \boldsymbol{z}_k \\ \vdots \\ \boldsymbol{z}_j \\ \vdots \\ \boldsymbol{z}_r \end{pmatrix}.$$

(D3) $\det E_r = 1$.

Lemma 2.43. *Es gilt:*

(D4) Sind zwei Zeilenvektoren gleich, so ist $\det A = 0$.

(D5) Addition einer α-fachen Zeile zu einer anderen Zeile, ändert die Determinanten nicht.

Beweis. (D4) Es sei $\boldsymbol{z}_j = \boldsymbol{z}_k, k \neq j$. Dann ist

$$\det A = \det \begin{pmatrix} \boldsymbol{z}_1 \\ \vdots \\ \boldsymbol{z}_j \\ \vdots \\ \boldsymbol{z}_k \\ \vdots \\ \boldsymbol{z}_r \end{pmatrix} = \det \begin{pmatrix} \boldsymbol{z}_1 \\ \vdots \\ \boldsymbol{z}_k \\ \vdots \\ \boldsymbol{z}_j \\ \vdots \\ \boldsymbol{z}_r \end{pmatrix} \overset{(D2)}{=} -\det \begin{pmatrix} \boldsymbol{z}_1 \\ \vdots \\ \boldsymbol{z}_k \\ \vdots \\ \boldsymbol{z}_j \\ \vdots \\ \boldsymbol{z}_r \end{pmatrix} = -\det A,$$

woraus $\det A = 0$ folgt.

(D5) Es ist

$$\det \begin{pmatrix} \boldsymbol{z}_1 \\ \vdots \\ \boldsymbol{z}_k + \alpha \boldsymbol{z}_j \\ \vdots \\ \boldsymbol{z}_r \end{pmatrix} \overset{(D1)}{=} \det A + \alpha \det \begin{pmatrix} \boldsymbol{z}_1 \\ \vdots \\ \boldsymbol{z}_j \\ \vdots \\ \boldsymbol{z}_j \\ \vdots \\ \boldsymbol{z}_r \end{pmatrix} \overset{(D4)}{=} \det A.$$

□

Bemerkung 2.44. Die Addition des α-fachen der j-ten Zeile zur k-ten Zeile $(k \neq j)$ kann durch Multiplikation mit einer Matrix

$$S(\alpha, j, k) = \begin{array}{cc} \begin{array}{c} j - \text{ te Spalte} \\ \downarrow \end{array} & \\ \begin{pmatrix} 1 & 0 & & & 0 \\ & \ddots & & & \\ & \alpha & 1 & & \\ & & & \ddots & \\ 0 & & & & 1 \end{pmatrix} & \begin{array}{c} \\ \\ \leftarrow k - \text{ te Zeile} \\ \\ \\ \end{array} \end{array} \tag{2.43}$$

$$S_{ij} = \begin{cases} \alpha & \text{für } i = k, l = j, \\ 1 & \text{für } i = l, \\ 0 & \text{sonst} \end{cases}$$

beschrieben werden. Es ist nämlich

$$S(\alpha,j,k)A = \begin{pmatrix} 1 & 0 & & & 0 \\ & \ddots & & & \\ & \alpha & 1 & & \\ & & & \ddots & \\ 0 & & & & 1 \end{pmatrix} \begin{pmatrix} a_{11} & \dots & a_{1r} \\ \vdots & & \vdots \\ a_{j1} & \dots & a_{kr} \\ a_{k1} & \dots & a_{jr} \\ \vdots & & \vdots \\ a_{r1} & \dots & a_{rr} \end{pmatrix}$$

$$= \begin{pmatrix} \boldsymbol{z}_1 \\ \vdots \\ \boldsymbol{z}_k + \alpha \boldsymbol{z}_j \\ \vdots \\ \boldsymbol{z}_r \end{pmatrix}.$$

Lemma 2.45. *Besitzt A Zeilenstufenform, dann ist*

$$\det A = a_{11}a_{22}\cdots a_{rr}.$$

Beweis. Da A eine $(r \times r)$-Matrix ist, hat A die Gestalt

$$A = \begin{pmatrix} a_{11} & & * \\ & \ddots & \\ 0 & & a_{rr} \end{pmatrix},$$

wobei in der Diagonale auch Nullen stehen können.

1.Fall

Wir nehmen an, dass A nicht invertierbar ist, d.h. $\operatorname{Rang}(A) < r$. Daher muss mindestens die r-te Zeile aus Nullen bestehen, d.h. $a_{rr} = 0$.
Nach (D5) können wir zur letzten Zeile eine andere addieren und erhalten 2 gleiche Zeilen. (D4) liefert uns die Aussage, dass in diesem Fall $\det A = 0$ ist.

2.Fall

Es sei $\operatorname{Rang}(A) = r$, d.h. A ist invertierbar. In diesem Fall müssen alle Diagonalelemente ungleich Null sein. Durch Addition von α-fachen Zeilen zur darüberliegenden Zeile erhalten wir eine Diagonalform $\hat{A}$

$$\hat{A} = \begin{pmatrix} a_{11} & & 0 \\ & \ddots & \\ 0 & & a_{rr} \end{pmatrix}.$$

Es ist

$$\det A \overset{(D5)}{=} \det \hat{A} \overset{(D1)}{=} a_{11} \det \begin{pmatrix} 1 & & 0 \\ 0 & a_{22} & \\ & \ddots & \\ 0 & & a_{rr} \end{pmatrix} \overset{(D1)}{=} a_{11}a_{22}\cdots a_{rr} \det E_r \overset{(D3)}{=} a_{11}a_{22}\cdots a_{rr}.$$

$\square$

Beispiele

1° Wir betrachten

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}.$$

Es sei $a_{11} \neq 0$. Die Zeilenstufenform von A lautet

$$\tilde{A} = \begin{pmatrix} a_{11} & a_{12} \\ 0 & -\frac{a_{21}}{a_{11}} a_{12} + a_{22} \end{pmatrix}.$$

Es ist nach (D1) und Lemma 2.45

$$\det A = \det \tilde{A} = a_{11} \left(-\frac{a_{12}}{a_{11}} a_{21} + a_{22} \right) = a_{11}a_{22} - a_{12}a_{21} = D \text{ (Vergleiche S. 102).}$$

2° Sei $S(\alpha_{ij}, k)$ durch (2.43) gegeben. Dann ist

$$\det S(\alpha, j, k) = 1. \tag{2.44}$$

<u>1.Fall:</u> $i < l$. Dann liegt der Eintrag $S_{il} = \alpha$ in der oberen Dreiecksmatrix.

<u>2.Fall:</u> $i > l$. $S_{il} = \alpha$ befindet sich in der unteren Dreiecksmatrix und kann durch Addition des α-fachen der l-ten Zeile zur i-ten Zeile zu Null gemacht werden.

Folgerung 2.46. Folgende Aussagen sind für Matrizen $A \in M(k, r \times r)$ äquivalent:

$$\begin{aligned} \det A = 0 \Leftrightarrow \operatorname{Rang} A < r \Leftrightarrow A &\text{ ist nicht invertierbar} \\ \Leftrightarrow &\text{ die Zeilen bzw. Spalten von } A \text{ sind linear abhängig.} \end{aligned}$$

Offen ist bisher, ob die Determinante eindeutig bestimmt ist. Das folgende Lemma beantwortet diese Frage.

Lemma 2.47. *Es gibt höchstens eine Determinante zu einer Matrix $A \in M(K, r \times r)$.*

Beweis. Da durch Addition einer α-fachen Zeile zu einer anderen Zeile stets erreicht werden kann, dass die entstehende Matrix $\tilde{A}$ Zeilenstufenform hat und nach (D1)

$$\det A = \det \tilde{A}$$

ist, folgt aus Lemma 2.45 die Behauptung. $\square$

Achtung!

Es gibt Beispiele, dass $\det(A + B) \neq \det A + \det B$ ist. Sei

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}, B = \begin{pmatrix} 0 & 0 \\ -2 & 0 \end{pmatrix}.$$

Es ist

$$\det A = \det \begin{pmatrix} 1 & 2 \\ 0 & -3 \end{pmatrix} = -3,$$
$$\det B = 0.$$

Jedoch erhalten wir

$$\det(A + B) = \det \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} = 1 \neq -3 + 0.$$

Für das Produkt zweier Matrizen gilt jedoch:

Lemma 2.48. *Seien $A, B \in M(K, r \times r)$, dann gilt*
(D6) $\det(AB) = \det A \det B$.

Beweis.

1. Fall

A sei nicht invertierbar. Wir zeigen, dass dann auch AB nicht invertierbar ist:
Seien $\boldsymbol{z}_i$ die Zeilenvektoren von A und $\boldsymbol{s}_j$ die Spaltenvektoren von B, $i, j = 1, \ldots, r$. Die Zeilenvektoren von AB lauten

$$\boldsymbol{Z}_i = (\boldsymbol{z}_i \boldsymbol{s}_1, \boldsymbol{z}_i \boldsymbol{s}_2, \ldots, \boldsymbol{z}_i \boldsymbol{s}_r), \quad i = 1, \ldots, r.$$

Sind die Zeilenvektoren $\boldsymbol{z}_k$ und $\boldsymbol{z}_j$ von A linear abhängig, d.h. es existieren $\alpha, \beta \in \mathbb{R}$, $\alpha \neq 0, \beta \neq 0$, sodass

$$\alpha \boldsymbol{z}_k + \beta \boldsymbol{z}_j = (0, 0, \ldots, 0)$$

ist, dann gilt

$$\begin{aligned} \alpha \boldsymbol{Z}_k + \beta \boldsymbol{Z}_j &= (\alpha \boldsymbol{z}_k \boldsymbol{s}_1, \ldots, \alpha \boldsymbol{z}_k \boldsymbol{s}_r) \\ &\quad + (\beta \boldsymbol{z}_j \boldsymbol{s}_1, \ldots, \beta \boldsymbol{z}_j \boldsymbol{s}_r) \\ &= ((\alpha \boldsymbol{z}_k + \beta \boldsymbol{z}_j) \boldsymbol{s}_1, \ldots, (\alpha \boldsymbol{z}_k + \beta \boldsymbol{z}_j) \boldsymbol{s}_r) \\ &= (0, 0, \ldots, 0). \end{aligned}$$

Damit sind auch die Zeilenvektoren $\boldsymbol{Z}_k$ und $\boldsymbol{Z}_j$ linear abhängig. Wir erhalten dass

$$\det(AB) = 0 = \det A \det B = 0 \det B$$

ist.

2. Fall

A sei invertierbar. Wir können A nach (D1) durch Multiplikation mit einer Matrix S, die durch Multiplikation von Matrizen der Form $S(\alpha, j, k)$ (siehe Formel (2.43)) entsteht, in Diagonalform bringen, wobei

$$\det(SA) = \hat{a}_{11} \cdots \hat{a}_{rr} = \det A$$

ist. Hierbei wurde genutzt, dass $\det S = 1$ nach (2.44) und (D6) ist. Weiterhin ist

$$\begin{aligned}
\det(AB) &= \det(SAB) = \det((SA)B) \\
&= \det\left[\begin{pmatrix} \hat{a}_{11} & & 0 \\ & \ddots & \\ 0 & & \hat{a}_{rr} \end{pmatrix} \begin{pmatrix} b_{11} & \dots & b_{1r} \\ \vdots & \vdots & \vdots \\ b_{r1} & \dots & b_{rr} \end{pmatrix}\right] \\
&= \det\begin{pmatrix} \hat{a}_{11}b_{11} & \dots & \hat{a}_{11}b_{1r} \\ \vdots & \vdots & \vdots \\ \hat{a}_{rr}b_{r1} & \dots & \hat{a}_{rr}b_{rr} \end{pmatrix} \\
&\overset{(D1)}{=} \hat{a}_{11}\hat{a}_{22}\cdots\hat{a}_{rr} \det B = \det A \det B.
\end{aligned}$$

□

Berechnung von Determinanten

Wir haben bereits gesehen, dass Determinanten von (2×2)-Matrizen folgendermaßen berechnet werden können:

$$\det\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11}a_{22} - a_{12}a_{21}.$$

Die Determinanten von (3×3) Matrizen können ebenfalls durch Transformation in eine Zeilenstufenmatrix und nach Lemma 2.45 berechnet werden (Regel von Sarrus):

$$\begin{aligned}
&\det\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \\
&= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32}.
\end{aligned} \tag{2.45}$$

Es gibt ein Verfahren, wie man die Berechnung einer Determinanten einer $(r \times r)$ Matrix auf die Berechnung von Determinanten von $((r-1) \times (r-1))$ Matrizen zurückführen kann. Um dies zu erläutern, führen wir folgende Bezeichnung ein:
Es sei A eine $(r \times r)$ Matrix

$$A = \begin{pmatrix} a_{11} & \dots & a_{1r} \\ \vdots & \vdots & \vdots \\ a_{r1} & \dots & a_{rr} \end{pmatrix}.$$

A_{ij} sei die $((r-1) \times (r-1))$ Matrix, die durch Streichen der i-ten Zeile und der j-ten Spalte von A entsteht.

Weiterhin sei $\tilde{A}_{ij}$ die Matrix, in der die i-te Zeile von A durch den j-ten Einheitsvektor $(0, \ldots, 1, \ldots, 0)$ ersetzt wird,

$$\begin{array}{cc} & \begin{array}{c} j\text{-te Spalte} \\ \downarrow \end{array} \\ \tilde{A}_{ij} = & \begin{pmatrix} a_{11} & \ldots & a_{1j} & \ldots & a_{1r} \\ \vdots & & \vdots & & \vdots \\ 0 & \ldots & 1 & \ldots & 0 \\ \vdots & & \vdots & & \vdots \\ a_{r1} & \ldots & a_{rj} & \ldots & a_{rr} \end{pmatrix} \leftarrow i\text{-te Zeile}, \end{array}$$

d.h. a_{ij} wird durch 1 ersetzt, sonst stehen Nullen in der i-ten Zeile.

Lemma 2.49. *Es ist*

$$\det(\tilde{A}_{ij}) = (-1)^{i+j} \det A_{ij}.$$

Beweis.
1. Fall

Es sei $i = j = 1$. Es gilt, dass

$$\operatorname{Rang}(A_{11}) = \operatorname{Rang}(\tilde{A}_{11}) - 1$$

ist. Es ist nämlich

$$A_{11} = \begin{pmatrix} a_{22} & \ldots & a_{2r} \\ \vdots & \vdots & \vdots \\ a_{r2} & \ldots & a_{rr} \end{pmatrix}, \quad \tilde{A}_{11} = \begin{pmatrix} 1 & \mathbf{0} \\ * & \begin{pmatrix} \\ A_{11} \\ \\ \end{pmatrix} \end{pmatrix}.$$

Durch Addition von Vielfachen der ersten Zeile von $\tilde{A}_{11}$ zu den anderen Zeilen können wir erreichen, dass aus $\tilde{A}_{11}$

$$\tilde{\tilde{A}}_{11} = \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & A_{11} \end{pmatrix}$$

wird und

$$\operatorname{Rang}(\tilde{A}_{11}) = \operatorname{Rang}(\tilde{\tilde{A}}_{11}) = \operatorname{Rang}(A_{11}) + 1$$

gilt. Damit erhalten wir

$$\det A_{11} = 0 \Leftrightarrow \det \tilde{A}_{11} = 0.$$

Ist $\det A_{11} \neq 0$, so können wir A_{11} in Diagonalform bringen

$$SA_{11} = \hat{A}_{11} = \begin{pmatrix} \hat{a}_{22} & & 0 \\ & \ddots & \\ 0 & & \hat{a}_{rr} \end{pmatrix}.$$

Wir betrachten

$$\tilde{S}\tilde{A}_{11} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ * & \hat{a}_{22} & & 0 \\ * & & \ddots & \vdots \\ * & 0 & \dots & \hat{a}_{rr} \end{pmatrix},$$

wobei $\tilde{S} = \begin{pmatrix} 1 & \mathbf{0} \\ \boldsymbol{*} & S \end{pmatrix}$ ist. Wir erhalten auch in diesem Fall

$$\det A_{11} = \hat{a}_{22} \cdot \dots \cdot \hat{a}_{rr} = \det \tilde{S}\tilde{A}_{11} \stackrel{(2.44),(D6)}{=} 1 \cdot \hat{a}_{22} \cdot \dots \cdot \hat{a}_{rr} = \det \tilde{A}_{11}.$$

2. Fall

i und j seien beliebig.
Wir bringen die 1, die am Platz ij von $\tilde{A}_{ij}$ steht, durch Vertauschen benachbarter Zeilen in die Diagonale. Dies liefert nach (D2) den Faktor $(-1)^{|i-j|} = (-1)^{i+j}$. Dann gehen wir analog zum 1. Fall vor. □

Satz 2.50 (Entwicklungssatz von Laplace). *Sei A eine $(r \times r)$ Matrix. Es ist*

$$\det A = \sum_{j=1}^{r} (-1)^{i+j} a_{ij} \det A_{ij}, \tag{2.46}$$

wobei $i \in \{1, \dots, r\}$ ist. (2.46) wird als Entwicklung nach der i-ten Zeile bezeichnet.

Beweis. Wir können die Matrix A folgendermaßen schreiben:

$$A = \begin{pmatrix} \boldsymbol{z}_1 \\ \vdots \\ \boldsymbol{z}_i \\ \vdots \\ \boldsymbol{z}_r \end{pmatrix} = \begin{pmatrix} \boldsymbol{z}_1 \\ \vdots \\ a_{ij}(1, 0, \dots, 0) + \dots + a_{ir}(0, 0, \dots, 1) \\ \vdots \\ \boldsymbol{z}_r \end{pmatrix}.$$

Nach (D1) gilt

$$\det A = a_{i1} \det \tilde{A}_{i1} + \dots + a_{ir} \det \tilde{A}_{ir} \stackrel{\text{Lemma } 2.49}{=} a_{i1}(-1)^{i+1} \det A_{i1} + \dots + a_{ir}(-1)^{i+r} \det A_{ir}.$$

□

Beispiel

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

a) Druch kreuzweises Rechnen (Regel von Sarrus (2.45)) erhalten wir

$$\det A = 6 + 12 = 18.$$

b) Entwicklung nach der ersten Zeile liefert:

$$\begin{aligned} \det A &= 2\det\begin{pmatrix} 3 & 1 \\ 0 & 1 \end{pmatrix} + 0\det\begin{pmatrix} 5 & 1 \\ -1 & 1 \end{pmatrix} + 4\det\begin{pmatrix} 5 & 3 \\ -1 & 0 \end{pmatrix} \\ &= 2\cdot 3 + 4\cdot 3 = 18. \end{aligned}$$

c) Entwicklung nach der zweiten Spalte (siehe Lemma 2.53) liefert

$$\det A = 3\det\begin{pmatrix} 2 & 4 \\ -1 & 1 \end{pmatrix} = 3\cdot 6 = 18.$$

Die transponierte Matrix

Bei der Rangbestimmung haben wir gesehen, dass Rang A = Zeilenrang A = Spaltenrang A ist. Es tritt die Vermutung auf, dass wir die Determinanten von A auch durch Entwicklung nach Spalten berechnen können. Dies wird tatsächlich möglich sein.
Wir führen dazu die zu A transponierte Matrix $A^\top$ ein.

Definition 2.51. Sei A eine $(r \times s)$ Matrix, $A = (a_{ij})_{\substack{1\le i\le r \\ 1\le j\le s}}$. Die zu A transponierte Matrix ist eine $(s \times r)$ Matrix und durch

$$A^\top = (a_{ji})_{\substack{1\le j\le s \\ 1\le i\le r}}$$

gegeben.

Beispiele

1. Sei $A = \begin{pmatrix} 1 & 2 & 0 & 0 \\ 0 & 3 & 1 & 1 \end{pmatrix}$. Dann ist

$$A^\top = \begin{pmatrix} 1 & 0 \\ 2 & 3 \\ 0 & 1 \\ 0 & 1 \end{pmatrix}.$$

2. Für $A = \begin{pmatrix} a_1 \\ \vdots \\ a_r \end{pmatrix}$ gilt
$$A^\top = (a_1, \ldots, a_r).$$

Lemma 2.52. *Es ist*
$$(AB)^\top = B^\top A^\top.$$

Beweis. Es gilt $AB = (c_{ik})$, wobei $c_{ik} = \sum_{j=1}^r a_{ij} b_{jk}$ (i-te Zeile mal k-te Spalte) ist. Andererseits ist $(AB)^\top = (c_{ki})$ mit

$$\begin{aligned} c_{ki} &= \sum_{j=1}^r a_{kj} b_{ji} \quad (k\text{-te Zeile mal } i\text{-te Spalte}) \\ &= \sum_{j=1}^r b_{ij} a_{jk} \quad (i\text{-te Zeile von } B^\top \text{ mal } k\text{-te Spalte von } A^\top). \end{aligned}$$

Somit gilt $(AB)^\top = B^\top A^\top$. □

Lemma 2.53. *Sei A eine $(r \times r)$ Matrix. Dann ist*

a) $\det(A) = \det\left(A^\top\right)$,

b) $\det A = \sum_{i=1}^r (-1)^{i+j} a_{ij} \det A_{ij}$, *wobei $j \in \{1, 2, \ldots, r\}$ ist.*

Beweis.

a) **1. Fall:** Es sei $\det A = 0$. Dann folgt $\operatorname{Rang}(A) < r$ und $\operatorname{Rang}(A^\top) < r$ (nach (2.26)) und damit $\det A^\top = 0$.

2. Fall: Es sei $\det A \neq 0$.
Dann können wir durch Zeilenumformungen erreichen, dass

$$\det A = \det(SA) = \det \begin{pmatrix} \hat{a}_{11} & & 0 \\ & \ddots & \\ 0 & & \hat{a}_{rr} \end{pmatrix}. \tag{2.47}$$

Hierbei ist S das Produkt von Matrizen der Form $S(\alpha, j, k)$, die durch (2.43) gegeben sind. Da $\det S(\alpha, j, k) = \det\left(S(\alpha, j, k)^\top\right) \overset{(2.44)}{=} 1$ gilt und nach Lemma 2.48 und Lemma 2.52
$$\det S = \det S^\top = 1$$
folgt, erhalten wir

$$\det A = \det(SA) \overset{(2.47)}{=} \det(SA)^\top \overset{\text{Lemma 2.48}}{=} \det A^\top \det S^\top = \det A^\top.$$

b) Diese Eigenschaft folgt aus a).

□

Übungsaufgaben

41)

1. Gegeben seien die Matrizen

$$A = (1\ 2)\,, B = \begin{pmatrix} 1 \\ 2 \end{pmatrix}\,, C = \begin{pmatrix} 1 & 0 \\ 2 & 3 \end{pmatrix}\,, D = \begin{pmatrix} 2 & 0 & 1 \\ 4 & 1 & 3 \end{pmatrix}.$$

   Geben Sie an, welche der Matrixprodukte

$$AB, BA, CD, DC^\top, DC, D^\top C, D^\top D, DD^\top$$

   existieren und berechnen Sie diese.

2. Zeigen Sie mit vollständiger Induktion

$$\begin{pmatrix} 1 & 0 & a \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix}^n = \begin{pmatrix} 1 & 0 & na \\ 0 & 1 & nb \\ 0 & 0 & 1 \end{pmatrix}.$$

42) Gegeben sind die Matrizen

$$A = \frac{1}{\sqrt{5}} \begin{pmatrix} 2 & 1 \\ -1 & 2 \end{pmatrix} \quad \text{und} \quad B = \begin{pmatrix} 8 & -2 \\ -2 & 5 \end{pmatrix}.$$

1. Überprüfen Sie, dass $AA^\top = A^\top A$ gleich der Einheitsmatrix ist. Berechnen Sie $D = A^\top BA$ und zeigen Sie, dass D eine Diagonalform der Gestalt

$$D = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$$

   besitzt. Was erhält man für D^n $(n \in \mathbb{N})$?

2. Begründen Sie, dass $B^n = AD^nA^\top$ gilt, und berechnen Sie damit B^n für $n \in \mathbb{N}$.

2.6 Eigenwerte und Diagonalisierbarkeit

Im vorigen Abschnitt hatten wir lineare Gleichungssysteme

$$A\boldsymbol{x} = \boldsymbol{b}$$

betrachtet. Dabei konnten wir durch elementare Zeilenumformungen erreichen, dass das Gleichungssystem eine Diagonalform annimmt. Wir sehen uns jetzt die Diagonalisierbarkeit einer gegebenen Matrix A an, die durch Darstellung der Matrix in einer geeigneten

Basis charakterisiert wird. Zunächst erinnern wir uns an die Darstellung linearer Abbildungen durch Matrizen, die von der Wahl der Basis abhängt (siehe Abschnitt 2.3).
Sei f eine lineare Abbildung, die einen r-dimensionalen Vektorraum V in sich selbst abbildet:

$$f\,:\,V \to V, \quad \dim V = r.$$

Sei $\{\boldsymbol{v_1}, \ldots, \boldsymbol{v_r}\}$ eine Basis von V. Es gibt eindeutig bestimmte Zahlen $a_{ij} \in K$, so dass

$$f(\boldsymbol{v}_j) = \sum_{i=1}^{r} a_{ij}\boldsymbol{v}_i \tag{2.48}$$

ist. Die Matrix $A = (a_{ij})$ charakterisiert die Abbildung $f : V \to V$, siehe (2.17). In der Standardbasis ist, siehe (2.18),

$$f(\boldsymbol{x}) = f_A(\boldsymbol{x}) = A\boldsymbol{x} \quad \forall \boldsymbol{x} \in V = K^r. \tag{2.49}$$

Im Allgemeinen hängt A von der Wahl der Basis ab. Die Diagonalisierbarkeit ist äquivalent dazu, eine Basis zu finden, in der A Diagonalform besitzt. Es wird sich erweisen, dass eine solche Basis durch sogenannte Eigenvektoren erzeugt werden kann.

Definition 2.54. Sei $K = \mathbb{R}$ oder $\mathbb{C}$ und V ein r-dimensionaler K-Vektorraum. Sei $f : V \to V$ eine lineare Abbildung. Eine Zahl $\lambda \in K$ heißt Eigenwert (kurz EW) von f, falls es ein $\boldsymbol{v} \in V \setminus \{\boldsymbol{0}\}$ gibt mit

$$f\boldsymbol{v} = \lambda\boldsymbol{v}. \tag{2.50}$$

Jedes $\boldsymbol{v} \in V \setminus \{\boldsymbol{0}\}$, das (2.50) genügt, heißt Eigenvektor von f (kurz EV) zum Eigenwert λ. Ein Paar $\{\lambda, \boldsymbol{v}\}$, wobei λ Eigenwert ist und $\boldsymbol{v}$ der zugehörigen Eigenvektor, wird Eigenpaar genannt.

Bemerkung 2.55. Sei $(\boldsymbol{e}_1, \ldots, \boldsymbol{e}_r)$ die Standardbasis in K^r und A die durch (2.48) bzw. (2.18) zugeordnete Matrix zu $f = f_A$. Ersetzt man in der Definition 2.54 f durch A, erhält man die Begriffe Eigenwert bzw. Eigenvektor einer Matrix.

Wir führen jetzt den Begriff eines Eigenraumes und des Spektrums einer linearerAbbildung ein.

Definition 2.56.

(i) Sei f wie in der Definition 2.54 gegeben und $\lambda \in K$ ein Eigenwert von f. Der lineare Teilraum

$$E_\lambda = \{\boldsymbol{v} \in V \;:\; f(\boldsymbol{v}) = \lambda\boldsymbol{v}\} = \ker(f - \lambda\,\mathrm{id}),$$

wobei id die identische Abbildung ist, $\mathrm{id}\,\boldsymbol{v} = \boldsymbol{v}$, wird Eigenraum zum Eigenwert λ genannt. Die Dimension des Eigenraums, $\dim E_\lambda$, heißt **geometrische** Vielfachheit des Eigenwertes λ.

(ii) Die Teilmenge

$$\sigma(f) = \{\lambda \in K \,:\, \lambda \text{ ist Eigenwert von } f\} \subset K$$

heißt Spektrum von f.

Diese Definitionen können auch auf Matrizen A, die durch lineare Abbildungen f_A dargestellt werden, übertragen werden.
Für die Berechnung der Eigenwerte einer Matrix A ist folgender Satz grundlegend.

Satz 2.57. *Die Zahl λ ist genau dann Eigenwert von A, falls gilt*

$$\det(A - \lambda E) = 0.$$

Beweis.

a) Es sei λ Eigenwert von A. Dann gibt es ein $\boldsymbol{v} \neq \mathbf{0}$, so dass

$$A\boldsymbol{v} - \lambda\boldsymbol{v} = (A - \lambda E)\boldsymbol{v} = \mathbf{0}$$

ist, d.h. $\ker(A - \lambda E) \neq \{\mathbf{0}\}$, $(A - \lambda E)$ ist nicht invertierbar und aus Folgerung 2.46 erhalten wir

$$\det(A - \lambda E) = 0.$$

b) Es sei $\det(A - \lambda E) = 0$. Dann ist $\ker(A - \lambda E) \neq \{\mathbf{0}\}$ und es existiert ein $\boldsymbol{v} \neq \mathbf{0}$, so dass

$$A\boldsymbol{v} = \lambda\boldsymbol{v}$$

ist.

□

Das charakteristische Polynom

Es ist

$$\det(A - \lambda E) = \det \begin{pmatrix} a_{11} - \lambda & a_{12} & \dots & a_{1r} \\ a_{21} & a_{22} - \lambda & \dots & a_{2r} \\ \vdots & \vdots & \vdots & \vdots \\ a_{r1} & \dots & a_{rr-1} & a_{rr} - \lambda \end{pmatrix} = P(\lambda),$$

wobei $P(\lambda)$ ein Polynom in λ ist. $P(\lambda)$ wird **charakteristisches Polynom** genannt.
Damit ist λ genau dann ein Eigenwert von A, falls λ Nullstelle von $P(\lambda)$ ist. Weiterhin kann A (bzw. f_A) höchstens r Eigenwerte besitzen. Führen wir einen Basiswechsel durch, so dass die Standardbasis im K^r in eine Basis B aus r Eigenvektoren überführt wird, $B = \{\boldsymbol{v}_1, \dots, \boldsymbol{v}_r\}$, dann wird die zugehörige Matrix A^B nach (2.48) Diagonalform besitzen. Dies wird später gezeigt und motiviert folgende Definition.

Definition 2.58. A (bzw. f_A) heißt diagonalisierbar, falls K^r eine Basis besitzt, die nur aus Eigenvektoren besteht.

Bemerkung 2.59. In Lemma 2.65 wird gezeigt: Ist A diagonalisierbar, dann existiert eine Matrix H deren Spalten die Eigenvektoren sind, so dass

$$H^{-1}AH = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_r \end{pmatrix}$$

ist. $\lambda_1, \dots, \lambda_r$ sind die Eigenwerte von A.

Beispiele

(i) Wir betrachten die Matrix $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. Die Abbildung $A\boldsymbol{x} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_2 \\ x_1 \end{pmatrix}$ beschreibt die Spiegelung an der Diagonale in der $x_1 - x_2$-Ebene. Es ist

$$\det(A - \lambda E) = \det \begin{pmatrix} -\lambda & 1 \\ 1 & -\lambda \end{pmatrix} = \lambda^2 - 1 = 0,$$

falls $\lambda = \pm 1$ ist.
Wir betrachten zunächst den Eigenwert $\lambda_1 = 1$. Es ist

$$\begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -x_1 + x_2 \\ x_1 - x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

falls $x_1 = x_2$ ist. Damit ist $\boldsymbol{v}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ ein Eigenvektor. Entsprechend gilt für den Eigenwert $\lambda_2 = -1$

$$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 + x_2 \\ x_1 + x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

falls $x_1 = -x_2$ ist. Es folgt, dass $\boldsymbol{v}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ ein Eigenvektor ist. Die Eigenvektoren $\boldsymbol{v}_1$ und $\boldsymbol{v}_2$ sind linear unabhängig. Man betrachte

$$\det \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = -2 \neq 0.$$

Führen wir im $\mathbb{R}^2$ die Basis $B = \{\boldsymbol{v}_1, \boldsymbol{v}_2\}$ der Eigenvektoren ein, so erhalten wir nach (2.48)

$$A^B = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

(ii) Es sei $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. A beschreibt eine Spieglung an der Achse $x_1 = x_2$ und eine anschließende Projektion auf die x_1-Achse:

$$A\boldsymbol{x} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_2 \\ 0 \end{pmatrix}.$$

Wir erhalten

$$\det(A-\lambda E) = \det\begin{pmatrix} -\lambda & 1 \\ 0 & -\lambda \end{pmatrix} = \lambda^2 = 0,$$

falls $\lambda = 0$ ist. Wir bestimmen die zugehörigen Eigenvektoren. Es ist

$$(A - 0\lambda)\boldsymbol{x} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_2 \\ 0 \end{pmatrix}.$$

Damit ist $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ ein Eigenvektor, jedoch kann ein einzelner Eigenvektor keine Basis im $\mathbb{R}^2$ bilden.

(iii) Wir sehen uns die Drehmatrix (Drehung um den Winkel ϕ) an:

$$A = \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix}.$$

Es ist

$$\begin{aligned}\det(A-\lambda E) &= \det\begin{pmatrix} \cos\phi - \lambda & -\sin\phi \\ \sin\phi & \cos\phi - \lambda \end{pmatrix} \\ &= (\cos\phi - \lambda)^2 + (\sin\phi)^2 = \lambda^2 - 2\lambda\cos\phi + 1.\end{aligned}$$

Die Wurzeln sind $\lambda_{1,2} = \cos\phi \pm \sqrt{-(\sin\phi)^2}$. Wir erhalten

- Für $K = \mathbb{R}$:

  Ist $\sin\phi \neq 0$, dann existiert kein reeller Eigenwert. Ist $\sin\phi = 0$, dann ist $A = \pm E$.

- Für $K = \mathbb{C}$:

  Es treten die Eigenwerte $\lambda_{1,2} = \cos\phi \pm \mathrm{i}\sin\phi$ auf. Ist $\sin\phi \neq 0$, dann sind sie voneinander verschieden. Wir bestimmen in diesem Fall die komplexen Eigenvektoren.
  Für $\lambda_1 = \cos\phi + \mathrm{i}\sin\phi$ muss gelten

  $$\begin{pmatrix} -\mathrm{i}\sin\phi & -\sin\phi \\ \sin\phi & -\mathrm{i}\sin\phi \end{pmatrix}\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

  Damit erhalten wir für $z_1 = a_1 + \mathrm{i}\,b_1$, $z_2 = a_2 + \mathrm{i}\,b_2$:

  $$\begin{aligned} -\sin\phi(\mathrm{i}\,z_1 + z_2) &= 0, \\ \sin\phi(z_1 - \mathrm{i}\,z_2) &= 0, \end{aligned}$$

  woraus

  $$\begin{aligned} a_1 + \mathrm{i}\,b_1 - \mathrm{i}(a_2 + \mathrm{i}\,b_2) &= 0, \\ a_1 + b_2 &= 0, \\ b_1 - a_2 &= 0 \end{aligned}$$

folgt. Damit ist $\boldsymbol{v}_1 = \begin{pmatrix} 1+\mathrm{i} \\ 1-\mathrm{i} \end{pmatrix}$ ein Eigenvektor.

Für $\lambda_2 = \cos\phi - \mathrm{i}\sin\phi$ muss gelten

$$\begin{pmatrix} \mathrm{i}\sin\phi & -\sin\phi \\ \sin\phi & \mathrm{i}\sin\phi \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

und analog zu oben erhalten wir, dass

$$\boldsymbol{v}_2 = \begin{pmatrix} 1+\mathrm{i} \\ -1+\mathrm{i} \end{pmatrix}$$

ein entsprechender Eigenvektor ist.

Diagonalisierbarkeit

Wir müssen also die Fragen nach der Existenz einer Basis, die nur aus Eigenvektoren besteht, diskutieren. Wir sehen uns zunächst die lineare Unabhängigkeit von Eigenvektoren an, die zu verschiedenen Eigenwerten gehören.

Satz 2.60. *Sei V ein K-Vektorraum, $f : V \to V$. Seien $\lambda_1, \dots, \lambda_k \in K$ verschiedene Eigenwerte zu f, $\boldsymbol{v}_1, \dots, \boldsymbol{v}_k$ dazugehörige Eigenvektoren. Dann sind $\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_k\}$ linear unabhängig.*

Beweis. Der Beweis erfolgt durch eine vollständige Induktion nach $k \in \mathbb{N}$:

1. Es sei $k = 1$. Dann ist $\boldsymbol{v}_1 \neq 0$ und die Menge $\{\boldsymbol{v}_1\}$ besteht aus linear unabhängigen Vektoren (Induktionsanfang).

2. Die Aussage gelte für $(k-1)$ verschiedene Eigenwerte (Induktionsannahme).

3. Wir zeigen, die Aussage gilt für k verschiedene Eigenwerte (Induktionsschluss). Wären $\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_k\}$ linear abhängig, dann lässt sich einer der Vektoren, sagen wir $\boldsymbol{v}_1$, durch die anderen ausdrücken

$$\boldsymbol{v}_1 = \mu_2\boldsymbol{v}_2 + \cdots + \mu_k\boldsymbol{v}_k, \quad \mu_i \in K. \tag{2.51}$$

Wir betrachten die Eigenwertgleichung:

$$\lambda_1\boldsymbol{v}_1 = \lambda_1\mu_2\boldsymbol{v}_2 + \cdots + \lambda_1\mu_k\boldsymbol{v}_k = f(\boldsymbol{v}_1) = f(\mu_2\boldsymbol{v}_2 + \cdots + \mu_k\boldsymbol{v}_k) = \mu_2\lambda_2\boldsymbol{v}_2 + \cdots + \mu_k\lambda_k\boldsymbol{v}_k.$$

Daraus folgt

$$0 = \mu_2(\lambda_2 - \lambda_1)\boldsymbol{v}_2 + \cdots + \mu_k(\lambda_k - \lambda_1)\boldsymbol{v}_k.$$

Aus der Induktionsannahme erhalten wir, dass gilt

$$\mu_i(\lambda_i - \lambda_1) = 0 \quad \text{für } i = 2, \dots, k.$$

Da $\lambda_i \neq \lambda_j$ ist, muss $\mu_i = 0$ für $i = 2, \dots, k$ sein. Der Eigenvektor $\boldsymbol{v}_1$ ist durch (2.51) gegeben und daher muss er der Nullvektor sein, was nicht sein kann.

□

Folgerung 2.61. Sei V endlichdimensional, $\dim V = r$, $f : V \to V$ eine lineare Abbildung. Besitzt f (bzw. f_A) r verschiedene Eigenwerte, d.h. $\sigma(f) = \{\lambda_1, \ldots, \lambda_r\}$, dann ist f bzw. die zu f_A gehörige Matrix A diagonalisierbar. In diesem Fall ist $\dim(E_{\lambda_i}) = 1$, $i = 1, \ldots, r$.

Die Matrix im Beispiel (i) S. 118 ist damit diagonalisierbar.
Ist die Anzahl verschiedener Eigenwerte kleiner als r, dann kann die Diagonalisierbarkeit bei genügender Anzahl von Eigenvektoren ebenfalls garantiert werden. Es gilt offensichtlich:

Satz 2.62. *Folgende Aussagen sind äquivalent.*

1. *f_A bzw. A seien diagonalisierbar.*
2. *$\sum_{i=1}^{k} \dim(E_{\lambda_i}) = r$, wobei $\lambda_1, \ldots, \lambda_k, k \leq r$, verschiedene Eigenwerte sind.*

Diagonalisierbarkeit und charakteristisches Polynom

Wir charakterisieren nun die Diagonalisierbarkeit mit Hilfe des charakteristischen Polynoms

$$P(\lambda) = \det(A - \lambda E).$$

Jede Nullstelle von $P(\lambda)$ ist Eigenwert von A. Wir wissen, dass auch mehrfache Nullstellen auftreten können.

Definition 2.63. Ist λ_0 Nullstelle k-ter Vielfachheit von $P(\lambda)$, so heißt k algebraische Vielfachheit des Eigenwertes λ_0

Satz 2.64. [4, S.235] *Sei $f : K^r \to K^r$ eine lineare Abbildung. Folgende Aussagen sind äquivalent:*

(i) f_A (bzw. A) ist diagonalisierbar.

(ii) $P(\lambda)$ zerfällt über K in Linearfaktoren (es treten nicht notwendig r verschiedene Nullstellen = Eigenwerte auf). Für jede Nullstelle λ_0 gilt, dass die geometrische Vielfachheit ($\dim(E_{\lambda_0})$) mit der algebraischen Vielfachheit von λ_0 übereinstimmt.

Beweis. Hier wird nur eine Skizze des Beweises angegeben.

- Es wird benutzt, dass für einen Eigenwert λ gilt

$$\text{geometrische Vielfachheit von } \lambda \leq \text{ algebraische Vielfachheit von } \lambda.$$

- Außerdem besagt, der Fundamentalsatz der Algebra, dass für $K = \mathbb{C}$ gilt

$$\sum_{i=1}^{k} \text{algebraische Vielfachheit } (\lambda_i) = r.$$

□

Diagonalisierung

Wir beschreiben jetzt den Prozess, wie man Matrizen diagonalisiert.

1.Schritt:

Berechnung der Eigenwerte λ_i, $i = 1, 2, \ldots, k$, der Matrix A durch Bestimmen der Nullstellen des charakteristischen Polynoms $P(\lambda) = \det(A - \lambda E) = 0$.

2.Schritt:

Berechnung zugehöriger Eigenvektoren $\boldsymbol{v}_i$ durch Lösung der Gleichungssysteme

$$(A - \lambda_i E)\boldsymbol{v}_i = 0.$$

Beachte, dass die $\boldsymbol{v}_i$ nicht eindeutig bestimmt sind.

3.Schritt:

Prüfen, ob $\sum_{i=1}^{k} \dim(E_{\lambda_i}) = r$ ist.

Wir nummerieren die verschiedenen Eigenpaare in der Form $(\lambda_1, \boldsymbol{v}_1), \ldots, (\lambda_k, \boldsymbol{v}_k)$, wobei nicht alle λ_i verschieden sein müssen. Die Menge $B = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_r\}$ der Eigenvektoren ist die gewünschte Basis, in der A^B Diagonalform besitzt. Wir überzeugen uns davon:
Die Koordinatentransformation von der Standardbasis $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_r\}$ in die Basis $B = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_r\}$ wird nach (2.21) durch die Matrix $H_{B_2} = H_B = H$, deren Spalten die Eigenvektoren $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_r$ sind, realisiert.

Lemma 2.65. *Es ist*

$$A^B = H^{-1}AH = D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_r \end{pmatrix}. \tag{2.52}$$

Beweis. Die Eigenwertgleichung liefert

$$A\boldsymbol{v}_i = \lambda_i \boldsymbol{v}_i, \quad i = 1, 2, \ldots, r. \tag{2.53}$$

Die $\boldsymbol{v}_i$ lassen sich darstellen als $\boldsymbol{v}_i = H\boldsymbol{e}_i$. Wir setzen $\boldsymbol{v}_i$ in (2.53) ein und erhalten:

$$AH\boldsymbol{e}_i = \lambda_i H\boldsymbol{e}_i = H\lambda_i \boldsymbol{e}_i.$$

Daraus folgt

$$\begin{aligned} H^{-1}AH\boldsymbol{e}_i &= \lambda_i \boldsymbol{e}_i, \\ H^{-1}AH &= D. \end{aligned}$$

Mit Hilfe der Beziehung (2.48) folgt $A^B = D$. Es ist nämlich

$$f_A(\boldsymbol{v}_j) = A\boldsymbol{v}_j = \sum_{i=1}^{r} a_{ij}^B \boldsymbol{v}_j = \lambda_j \boldsymbol{v}_j, \quad j = 1, \ldots, r.$$

Da die Koordinatendarstellung bezüglich der Basis $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_r\}$ eindeutig ist, folgt

$$a_{ij}^B = \begin{cases} 0 & \text{für } i \neq j, \\ \lambda_j & \text{für } i = j. \end{cases}$$

Wir erhalten, dass $A^B = (a_{ij}^B)_{\substack{1 \le i \le r \\ 1 \le j \le r}} = D$ ist. □

Bemerkung 2.66. Die Gleichung (2.52) kann auch durch direktes Nachrechnen von

$$AH = HD$$

unter Beachtung von (2.53) verifiziert werden.

Jordansche Normalform

Wie wir im Beispiel (ii) im letzten Abschnitt gesehen haben, kann es durchaus sein, dass nicht genügend viele Eigenvektoren existieren, d.h.

$$\text{geometrische Vielfachheit von } \lambda < \text{ algebraische Vielfachheit von } \lambda.$$

In diesem Fall ist f_A nicht diagonalisierbar. Trotzdem kann man in diesem Fall noch eine Jordansche Normalform von f_A bzw. der Matrix A konstruieren, in der anstelle der Diagonalelemente *einfache* Blockmatrizen (Jordan-Matrizen) treten.

Definition 2.67. Sei $\lambda \in K$. Die $(k \times k)$ Matrix

$$J_k(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \ldots & 0 \\ 0 & \lambda & 1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \vdots & \vdots & & \ddots & 1 \\ 0 & 0 & 0 & \ldots & \lambda \end{pmatrix}$$

heißt Jordan-Matrix.

Satz 2.68. [4, S. 268] *Sei $f = f_A : K^r \to K^r$, $f_A(\boldsymbol{x}) = A\boldsymbol{x}$ eine lineare Abbildung deren charakteristisches Polynom in Faktoren zerfällt,*

$$P(\lambda) = \det(A - \lambda E) = (\lambda_1 - \lambda)^{n_1} \cdot \cdots \cdot (\lambda_k - \lambda)^{n_k},$$

wobei $\lambda_1, \ldots, \lambda_k$ verschiedene Eigenwerte von f_A bzw. A sind. Dann gibt es eine Basis B in K^r in der die zu f gehörige Matrix A^B die Form

$$\begin{pmatrix} J_{n_1 1}(\lambda_1) & \ldots & \ldots & \ldots & \ldots & \ldots & 0 \\ \vdots & \ddots & & & & & \vdots \\ \vdots & & J_{n_1 l_1}(\lambda_1) & & & & \vdots \\ \vdots & & & \ddots & & & \vdots \\ \vdots & & & & J_{n_k 1}(\lambda_k) & & \vdots \\ \vdots & & & & & \ddots & \vdots \\ 0 & \ldots & \ldots & \ldots & \ldots & \ldots & J_{n_k l_k}(\lambda_k) \end{pmatrix} \tag{2.54}$$

besitzt. Dabei ist $\sum_{j=1}^{l_i} j = n_i$, $i = 1, \ldots, k$, *und* $n_1 + n_2 + \cdots + n_k = r$.

Bemerkung 2.69. Sind alle Jordan-Matrizen eindimensional, dann hat A Diagonalform.

Beispiele

1. Im Beispiel (ii) des vorherigen Abschnittes ist $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ bereits eine Jordan-Matrix.

2. Wir betrachten die Matrix
$$A = \begin{pmatrix} 0 & 1 & 3 \\ 0 & 0 & 0 \\ 0 & 5 & 1 \end{pmatrix}.$$
Es ist
$$\det(A - \lambda E) = \det \begin{pmatrix} -\lambda & 1 & 3 \\ 0 & -\lambda & 0 \\ 0 & 5 & 1-\lambda \end{pmatrix} = (1-\lambda)\lambda^2,$$
d.h. $n_1 = 1$, $n_2 = 2$.
Die Eigenwerte sind $\lambda_1 = 1$ und $\lambda_2 = 0$. Wir betrachten zunächst den Eigenwert $\lambda_1 = 1$ und berechnen dazu einen Eigenvektor $\boldsymbol{v}_1$:
$$(A - E)\boldsymbol{v}_1 = \begin{pmatrix} -1 & 1 & 3 \\ 0 & -1 & 0 \\ 0 & 5 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} -x_1 + x_2 + 3x_3 \\ -x_2 \\ 5x_2 \end{pmatrix},$$
woraus z.B. folgt
$$\boldsymbol{v}_1 = \begin{pmatrix} 3 \\ 0 \\ 1 \end{pmatrix}.$$
Wir erhalten
$$\dim(E_{\lambda_1}) = 1.$$
Nun sehen wir uns den Eigenwert $\lambda_2 = 0$ an und berechnen Eigenvektoren aus
$$(A - 0E)\boldsymbol{v}_2 = \begin{pmatrix} 0 & 1 & 3 \\ 0 & 0 & 0 \\ 0 & 5 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} x_2 + 3x_3 \\ 0 \\ 5x_2 + x_3 \end{pmatrix}.$$
Damit ist $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \boldsymbol{v}_2$ ein Eigenvektor und $\dim(E_{\lambda_2}) = 1$.
Die Matrix A ist daher nicht diagonalisierbar und ihre Jordansche Normalform lautet
$$A^B = \hat{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \tag{2.55}$$

da die algebraische Vielfachheit von $\lambda = 0$ gleich 2 ist.

Wir beschreiben eine Basis, in der A^B gegeben ist. Zu den Eigenvektoren $\boldsymbol{v}_1 = \begin{pmatrix} 3 \\ 0 \\ 1 \end{pmatrix}$ und $\boldsymbol{v}_2 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ ist ein weiterer Basisvektor $\boldsymbol{v}_3$ zu konstruieren. Dazu sehen wir uns die Gleichungen (2.48) und (2.49) an. Es soll gelten:

$$\begin{aligned} f_A(\boldsymbol{v}_3) = \hat{A}\boldsymbol{v}_3 &= \hat{a}_{13}\boldsymbol{v}_1 + \hat{a}_{23}\boldsymbol{v}_2 + \hat{a}_{33}\boldsymbol{v}_3 \\ &= 0\boldsymbol{v}_1 + 1\boldsymbol{v}_2 + 0\boldsymbol{v}_3 \end{aligned}$$

und damit

$$\begin{pmatrix} 0 & 1 & 3 \\ 0 & 0 & 0 \\ 0 & 5 & 1 \end{pmatrix} \begin{pmatrix} v_{3,1} \\ v_{3,2} \\ v_{3,3} \end{pmatrix} = 1\boldsymbol{v}_2 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} v_{3,2} + 3v_{3,3} \\ 0 \\ 5v_{3,2} + v_{3,3} \end{pmatrix}.$$

Wir erhalten

$$\boldsymbol{v}_3 = \begin{pmatrix} 0 \\ -\frac{1}{14} \\ \frac{5}{14} \end{pmatrix}.$$

Eine Basis in der A Jordan-Normalform $\hat{A} = A^B$ besitzt, lautet

$$B = \left\{ \begin{pmatrix} 3 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \begin{pmatrix} 0 \\ -\frac{1}{14} \\ \frac{5}{14} \end{pmatrix} \right\}.$$

3. Die (6×6)-Matrix A besitze zwei 3-fache Eigenwerte $\lambda_1 = 1$ und $\lambda_2 = 2$. Es sei $\dim E_{\lambda_1} = 2$ und $\dim E_{\lambda_2} = 1$. Dann lautet eine Jordansche Normalform von A:

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

In diesem Fall ist in (2.54) $n_11 = 1, n_12 = 2$, d.h. $n_11 + n_12 = 3$, und $n_21 = n_2 = 3$.

Bemerkung 2.70. Der Algorithmus, der im 2. Beispiel zur Bestimmung des fehlenden Basisvektors $\boldsymbol{v}_3$ druchgeführt hat, kann allgemein beschrieben werden. Sei f_A eine gegebene lineare Abbildung von K^k in K^k,

$$A^B = \begin{pmatrix} \lambda & 1 & & 0 \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & 1 \\ 0 & & & \lambda \end{pmatrix} \tag{2.56}$$

ein $k \times k$ Jordanblock zum Eigenwert λ mit $\dim E_\lambda = 1$. $\boldsymbol{v} \in E_\lambda$ sei ein bekannter Eigenvektor zu λ. Die restlichen $k-1$ Basisvektoren $\boldsymbol{v}_2, \ldots, \boldsymbol{v}_k$ können Schritt für Schritt aus folgenden Gleichungssystemen bestimmt werden, indem (2.48) und (2.56) verwendet werden:

$$f_A(\boldsymbol{v}_2) = A\boldsymbol{v}_2 = \sum_{i=1}^{k} a_{ij}^B \boldsymbol{v}_i \overset{\text{2. Spalte von (2.56)}}{=} 1\boldsymbol{v}_1 + \lambda \boldsymbol{v}_2.$$

Damit ist das Gleichungssystem

$$(A - \lambda E)\boldsymbol{v}_2 = \boldsymbol{v}_1 \tag{2.57}$$

zu lösen. Nehmen wir an, dass $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_{j-1}$ bekannt seien. Dann gilt für $\boldsymbol{v}_j$

$$f_A(\boldsymbol{v}_j) = A\boldsymbol{v}_j = \sum_{i=1}^{k} a_{ij}^B \boldsymbol{v}_i \overset{j\text{−Spalte von (2.56)}}{=} \boldsymbol{v}_{j-1} + \lambda \boldsymbol{v}_j$$

und daraus folgt

$$(A - \lambda)\boldsymbol{v}_j = \boldsymbol{v}_{j-1}.$$

Übungsaufgaben

43) Gegeben sind die Matrizen

$$A := \begin{pmatrix} 1 & -3 & 3 \\ 3 & -5 & 3 \\ 6 & -6 & 4 \end{pmatrix}, \qquad B := \begin{pmatrix} -3 & 1 & -1 \\ -7 & 5 & -1 \\ -6 & 6 & -2 \end{pmatrix}, \qquad C := \begin{pmatrix} 3 & 5 & 1 \\ 1 & 5 & -3 \\ 5 & 5 & 5 \end{pmatrix}.$$

1. Berechnen Sie alle Eigenwerte und zugehörige Eigenvektoren der Matrizen A, B und C. Hinweis: Die Gleichung $\det(A - \lambda E) = 0$ kann zur Berechnung der Eigenwerte verwendet werden.

2. Können Sie für jede Matrix eine Basis aus Eigenvektoren finden? Geben Sie diese gegebenenfalls an.

44) Bestimmen Sie die Jordan-Normalform der Matrix

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

und die zugehörige Basis.

45) Diagonalisieren Sie folgende Matrizen.

$$A = \begin{pmatrix} 2 & 3 & 3 \\ 1 & 4 & 3 \\ 3 & 6 & 11 \end{pmatrix}, \qquad B = \begin{pmatrix} 3 & -2 & 1 \\ 2 & 1 & 0 \\ 2 & -1 & 2 \end{pmatrix}.$$

Bestimmen Sie jeweils eine Diagonalmatrix D und eine invertierbare Matrix S, die den Basiswechsel beschreibt, mit $D = S^{-1}AS$ bzw. $D = S^{-1}BS$.

2.7 Skalarprodukt

Wir betrachten in diesem Abschnitt abstrakte Vektorräume, die in Abschnitt 2.1 eingeführt wurden. Sie können endlich oder unendlichdimensional sein, ihre Elemente können z.B. Zahlentupel, Abbildungen oder Funktionen sein. In Vektorräumen mit Skalarprodukt können geometrische Größen, wie Betrag und Winkel eingeführt werden, wie sie in endlichdimensionalen Vektorräumen bekannt sind. Damit ist auch der Begriff der Orthogonalität zweier Elemente eines abstrakt definierten Vektorraumes V wohldefiniert.

Definition 2.71. Es sei V ein K-Vektorraum, $K = \mathbb{R}$ oder $K = \mathbb{C}$. Eine Abbildung

$$\begin{aligned} V \times V &\to K, \\ (x, y) &\to \langle x, y \rangle, \end{aligned}$$

heißt Skalarprodukt auf V, falls die folgenden Eigenschaften für alle $x, y, z \in V, \alpha, \beta \in K$ gelten:

(S1) $\langle x, x \rangle \geq 0, \quad \langle x, x \rangle = 0 \Leftrightarrow x = 0,$

(S2) $\langle x, y \rangle = \langle y, x \rangle$ für $K = \mathbb{R}$ $\qquad \langle x, y \rangle = \overline{\langle y, x \rangle}$ für $K = \mathbb{C}$,

(S3) $\langle \alpha x + \beta y, z \rangle = \alpha \langle x, z \rangle + \beta \langle y, z \rangle.$

Aus diesen Eigenschaften folgt:

- $\langle x, y \rangle = \bar{\alpha} \langle x, y \rangle,$
- $\langle x, y + z \rangle = \langle x, y \rangle + \langle x, z \rangle.$

Beispiele

(i) Die Standard-Skalarprodukte auf $\mathbb{R}^n$ oder $\mathbb{C}^n$ sind:

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle_{\mathbb{R}^n} = \langle \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \rangle = \sum_{i=1}^n x_i y_i,$$

$$\langle \boldsymbol{z}, \boldsymbol{w} \rangle_{\mathbb{C}^n} = \langle \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix}, \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} \rangle = \sum_{i=1}^n z_i \bar{w}_i.$$

(ii) Wir betrachten im $\mathbb{R}^n$ ein weiteres Skalarprodukt $\langle \boldsymbol{x}, \boldsymbol{y} \rangle_A = \langle A\boldsymbol{x}, \boldsymbol{y} \rangle_{\mathbb{R}^n}$, wobei A eine Diagonalmatrix mit positiven Einträgen $(a_{ii} > 0, i = 1, \ldots, n)$ ist. Wir überprüfen die Eigenschaften (S1), (S2) und (S3):

(S1) Es gilt $\langle \boldsymbol{x}, \boldsymbol{x} \rangle_A = \sum_{i=1}^n a_{ii} x_i^2 \geq 0$ und $\langle \boldsymbol{x}, \boldsymbol{x} \rangle_A = 0 \Leftrightarrow \boldsymbol{x} = \boldsymbol{0}$.

(S2) Es gilt $\langle \boldsymbol{x}, \boldsymbol{y} \rangle_A = \sum_{i=1}^n a_{ii} x_i y_i = \langle \boldsymbol{y}, \boldsymbol{x} \rangle_A = \sum_{i=1}^n a_{ii} y_i x_i$.

(S3) Es gilt:

$$\begin{aligned} \langle \alpha \boldsymbol{x} + \beta \boldsymbol{y}, \boldsymbol{z} \rangle_A = \langle A(\alpha \boldsymbol{x} + \beta \boldsymbol{y}), \boldsymbol{z} \rangle_{\mathbb{R}^n} &= \sum_{i=1}^n a_{ii}(\alpha x_i + \beta y_i) z_i \\ &= \alpha \sum_{i=1}^n a_{ii} x_i z_i + \beta \sum_{i=1}^n a_{ii} y_i z_i \\ &= \alpha \langle \boldsymbol{x}, \boldsymbol{z} \rangle_A + \beta \langle \boldsymbol{y}, \boldsymbol{z} \rangle_A . \end{aligned}$$

(iii) Sei $V := \{f : [a,b] \to \mathbb{C}, f \text{ ist stetig}\}$. Durch

$$\langle f, g \rangle = \int_a^b f(x) \bar{g}(x) dx$$

ist ein Skalarprodukt definiert.

Aus den Eigenschaften (S1), (S2) und (S3) kann man eine wichtige Abschätzung folgern, die **Cauchy-Schwarzsche Ungleichung**:

Satz 2.72. *In einem K-Vektorraum, $K = \mathbb{R}$ oder $K = \mathbb{C}$, mit einem Skalarprodukt gilt die Abschätzung*

$$|\langle x, y \rangle| \leq \sqrt{\langle x, x \rangle} \sqrt{\langle y, y \rangle}. \tag{2.58}$$

Beweis. Es ist im $\mathbb{C}$-Vektorraum für einen Wert $\lambda \in \mathbb{C}$

$$\langle x + \lambda y, x + \lambda y \rangle = \langle x, x \rangle + \lambda \langle y, x \rangle + \bar{\lambda} \langle x, y \rangle + |\lambda|^2 \langle y, y \rangle \geq 0. \tag{2.59}$$

Wir nehmen an, dass $y \neq 0$ sei, also $\langle y, y \rangle > 0$ ist und wählen $\lambda = -\frac{\langle x, y \rangle}{\langle y, y \rangle}$ in (2.59). Wir erhalten

$$\langle x, x \rangle - \frac{\langle x, y \rangle \langle y, x \rangle}{\langle y, y \rangle} - \frac{\langle y, x \rangle \langle x, y \rangle}{\langle y, y \rangle} + \frac{|\langle x, y \rangle|^2}{\langle y, y \rangle^2} \langle y, y \rangle \geq 0,$$

woraus

$$\langle x, x \rangle \langle y, y \rangle \geq |\langle x, y \rangle|^2$$

und schließlich (2.58) folgt. □

Norm (Betrag)

Nun sind wir in der Lage, die Länge (Betrag) eines Elements im abstrakten Vektorraum V zu definieren.

Satz 2.73. *Sei V ein K-Vektorraum, $K = \mathbb{R}$ oder $K = \mathbb{C}$, mit Skalarprodukt. Dann ist für jedes Element $x \in V$ eine Norm (Betrag)*

$$\|x\| = \sqrt{\langle x, x \rangle}$$

definiert, die folgende Eigenschaften für alle $x, y \in V, \alpha \in K$ besitzt:

(N1) $\|x\| \geq 0, \quad \|x\| = 0 \;\Leftrightarrow\; x = 0,$

(N2) $\|\alpha x\| = |\alpha|\|x\|,$

(N3) $\|x+y\| \leq \|x\| + \|y\|.$

Beweis. (N1) und (N2) sind offensichtlich und wir sehen uns nur die Dreiecksungleichung (N3) an:

$$\begin{aligned}\|x+y\|^2 &= |\langle x+y, x+y\rangle| = |\langle x,x\rangle + \langle y,x\rangle + \langle x,y\rangle + \langle y,y\rangle| \\ &\overset{(2.58)}{\leq} \|x\|^2 + 2\|y\|\|x\| + \|y\|^2 = (\|x\| + \|y\|)^2.\end{aligned}$$

□

Bemerkung 2.74. Ein K-Vektorraum, in dem eine Abbildung $\|\cdot\| : V \to \mathbb{R}$ definiert ist, die den Axiomen (N1), (N2) und (N3) genügt, heißt normierter Raum.

Winkel und Orthogonalität

Wir können auch einen Winkel $\phi \in [0, \pi]$ zwischen zwei Elementen $x \neq 0$ und $y \neq 0$ in einem abstrakten Vektorraum V mit Skalarprodukt einführen:

$$\cos\phi = \frac{\langle x, y\rangle}{\|x\|\|y\|}. \tag{2.60}$$

Da $\cos\phi = \frac{|\langle x,y\rangle|}{\|x\|\|y\|} \overset{(2.58)}{\leq} 1$, ist diese Definition sinnvoll. Aus der Beziehung (2.60) folgt unmittelbar die Erklärung, wann zwei Elemente x und y zueinander orthogonal sind.

Definition 2.75. Zwei Elemente x und y aus einem K-Vektorraum, $K = \mathbb{R}$ oder $K = \mathbb{C}$, sind orthogonal ($x \perp y$), falls gilt

$$\langle x, y\rangle = 0.$$

Folgerung 2.76. Nur das Nullelement ist orthogonal zu allen Elementen aus V.

Definition 2.77. Eine Menge von Elementen $(v_i)_{i\in I} \subseteq V$ mit $v_i \neq 0$ heißt **Orthogonalsystem**, falls $v_i \perp v_j$ für $i \neq j$ ist. Wir sprechen von einem **Orthonormalsystem**, falls

$$\langle v_i, v_j\rangle = \delta_{ij} = \begin{cases} 0 & \text{für } i \neq j, \\ 1 & \text{für } i = j \end{cases}$$

ist. Ist ein Orthonormalsystem Basis in V, so sprechen wir von einer **Orthonormalbasis**.

Bemerkung 2.78.

- Es gilt der verallgemeinerte Satz vom Pythagoras für ein Orthogonalsystem aus n Elementen:
$$\|x_1 + \cdots + x_n\|^2 = \|x_1\|^2 + \cdots + \|x_n\|^2.$$
- Ein Orthogonalsystem besteht aus linear unabhängigen Vektoren. Ein System aus linear unabhängigen Vektoren, kann orthonormiert werden (Gram-Schmidtsches Orthogonalisierungsverfahren, siehe Übungsaufgabe 47).

Lineare orthogonale Abbildungen (orthogonale Matrizen)

Wir charakterisieren lineare Abbildung (Matrizen) eines abstrakten Vektorraumes V in V, die sich aus Spiegelungen und Drehungen zusammensetzen. Es sollen also Winkel und Beträge erhalten bleiben.

Definition 2.79. Sei V ein K-Vektorraum, $K = \mathbb{R}$ oder $K = \mathbb{C}$, mit Skalarprodukt. Eine lineare Abbildung $f : V \to V$, heißt orthogonal, falls für alle v und w aus V gilt

$$\langle v, w\rangle_V = \langle f(v), f(w)\rangle. \tag{2.61}$$

Folgerung 2.80. Es ist für alle v aus V für eine orthogonale Abbildung f

$$\|v\| = \|f(v)\|.$$

Der Winkel zwischen zwei Elementen v und $w \in V$ bleibt erhalten:

$$\text{Winkel zwischen } v \text{ und } w \ = \ \text{Winkel zwischen } f(v) \text{ und } f(w).$$

Weiterhin ist eine orthogonale Abbildung injektiv.

Endlichdimensionale Vektorräume und orthogonale Matrizen

Wir betrachten jetzt endlichdimensionale Vektorräume V, $\dim V = r$, mit Skalarprodukt. In V sei eine Orthonormalbasis, $B = \{\boldsymbol{v}_1, \dots, \boldsymbol{v}_r\}$, gegeben:

$$\langle \boldsymbol{v}_i, \boldsymbol{v}_j\rangle = \delta_{ij} \quad i, j = 1, \dots, r.$$

Wir wissen, dass eine lineare Abbildung f durch eine Matrix A bezüglich dieser Basis dargestellt werden kann (vergleiche dazu (2.48), (2.49)). Wir charakterisieren die Matrizen, die zu linearen orthogonalen Abbildungen gehören.

Lemma 2.81. *Sei $V = \mathbb{R}^r$ ein Vektorraum mit Skalarprodukt und orthogonaler Basis. Matrizen A, die zu linearen orthogonalen Abbildungen $f_A : V \to V$, $f_A(\boldsymbol{x}) = A\boldsymbol{x}$, gehören, haben die Eigenschaft, dass*

$$A^\top = A^{-1}$$

ist. Solche Matrizen heißen orthogonale Matrizen.

Beweis. Für eine orthogonale Basis $\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_r\}$ gilt:

$$\langle \boldsymbol{x}, \boldsymbol{y}\rangle = \langle \sum_{i=1}^r x_i \boldsymbol{v}_i, \sum_{j=1}^r y_j \boldsymbol{v}_j\rangle = \sum_{i=1}^r x_i y_i = \boldsymbol{x}^\top \boldsymbol{y} \quad \text{für alle } \boldsymbol{x}, \boldsymbol{y} \in V.$$

Daher ist

$$\langle \boldsymbol{x}, \boldsymbol{y}\rangle \overset{(2.61)}{=} \langle A\boldsymbol{x}, A\boldsymbol{y}\rangle = (A\boldsymbol{x})^\top A\boldsymbol{y} = \boldsymbol{x}^\top A^\top A\boldsymbol{y} = \langle \boldsymbol{x}, A^\top A\boldsymbol{y}\rangle.$$

Es folgt, dass

$$\langle \boldsymbol{x}, (A^\top A - E)\boldsymbol{y} \rangle = 0 \quad \text{für alle } \boldsymbol{x} \text{ und } \boldsymbol{y} \text{ aus } V \text{ ist,}$$

und daher muss nach Folgerung 2.76

$$A^\top A = E \tag{2.62}$$

sein. Wir müssen noch zeigen, dass $AA^\top = E$ ist.
Da V endlichdimensional ist und eine orthogonale Abbildung stets injektiv ist, liegt in diesem Fall eine bijektive Abbildung vor und A^{-1} existiert. Für A^{-1} gilt

$$\langle A^{-1}\boldsymbol{z}, A^{-1}\boldsymbol{w} \rangle = \langle AA^{-1}\boldsymbol{z}, AA^{-1}\boldsymbol{w} \rangle = \langle \boldsymbol{z}, \boldsymbol{w} \rangle \quad \text{für alle } \boldsymbol{z}, \boldsymbol{w} \in V, \tag{2.63}$$

d.h. A^{-1} ist ebenfalls eine orthogonale Matrix. Jetzt können wir zeigen, dass auch $A^\top$ diese Eigenschaft besitzt. Seien $\boldsymbol{x}$ und $\boldsymbol{y}$ beliebig aus V. Es existieren $\boldsymbol{z}$ und $\boldsymbol{w}$ aus V, so dass $\boldsymbol{x} = A\boldsymbol{z}, \boldsymbol{y} = A\boldsymbol{w}$ sind. Wir erhalten

$$\langle A^\top \boldsymbol{x}, A^\top \boldsymbol{y} \rangle = \langle A^\top A\boldsymbol{z}, A^\top A\boldsymbol{w} \rangle \overset{(2.62)}{=} \langle \boldsymbol{z}, \boldsymbol{w} \rangle = \langle A^{-1}\boldsymbol{x}, A^{-1}\boldsymbol{y} \rangle \overset{(2.63)}{=} \langle \boldsymbol{x}, \boldsymbol{y} \rangle.$$

Da

$$\langle A^\top \boldsymbol{x}, A^\top \boldsymbol{y} \rangle = (A^\top \boldsymbol{x})^\top A^\top \boldsymbol{y} = \boldsymbol{x} AA^\top \boldsymbol{y} = \langle \boldsymbol{x}, AA^\top \boldsymbol{y} \rangle = \langle \boldsymbol{x}, \boldsymbol{y} \rangle$$

ist, gilt

$$AA^\top = E$$

und zusammen mit (2.62) folgt $A^\top = A^{-1}$. □

Folgerung 2.82. Spalten und Zeilen einer orthogonalen Matrix sind jeweils orthonormale Systeme.

Bemerkung 2.83. Liegt ein $\mathbb{C}$-Vektorraum vor, dann spricht man von unitären Matrizen, falls

$$A^{-1} = \bar{A}^\top$$

ist.

Beispiele

- Für jedes $\alpha \in \mathbb{R}$ sei

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

  $A(\alpha)$ beschreibt bezüglich der Standardbasis im $\mathbb{R}^2$ eine Drehung um den Winkel α. Es ist

  $$A^\top(\alpha) = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix}$$

  und

  $$A^\top(\alpha)A(\alpha) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

- Die Matrix
$$B(\alpha) = \begin{pmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{pmatrix}$$
ist ebenfalls orthogonal. Sie stellt eine Drehspiegelung dar. Es ist
$$B(\alpha)\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \cos\alpha \\ \sin\alpha \end{pmatrix}, \quad B(\alpha)\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} \sin\alpha \\ -\cos\alpha \end{pmatrix}.$$

Wir hatten bereits gesehen, siehe (2.52), falls eine $(r \times r)$-Matrix r Eigenvektoren besitzt, dann können wir A diagonalisieren, indem wir
$$H^{-1}AH = D = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \dots & \dots & \lambda_r \end{pmatrix},$$
bilden. Die Zahlen λ_i sind die Eigenwerte von A und H besitzt die Eigenvektoren als Spaltenvektoren. Betrachten wir ein orthonormales System von Eigenvektoren, dann ist $H^{-1} = H^\top$ und
$$H^\top AH = D.$$
Diese Transformation wird auch Hauptachsentransformation genannt.

Übungsaufgaben

46) Zeigen Sie, dass für jedes Orthogonalsystem $\boldsymbol{v}_1, \dots, \boldsymbol{v}_n$

1. die Vektoren $\boldsymbol{v}_1, \dots, \boldsymbol{v}_n$ linear unabhängig sind,
2. der verallgemeinerte Satz von Pythagoras
$$\|\boldsymbol{v}_1 + \dots + \boldsymbol{v}_n\|^2 = \|\boldsymbol{v}_1\|^2 + \dots + \|\boldsymbol{v}_n\|^2$$
gilt.

Überprüfen Sie beide Aussagen anhand des Orthogonalsystems $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3$ mit
$$\boldsymbol{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \boldsymbol{v}_2 = \begin{pmatrix} -1 \\ -1 \\ 1 \\ 1 \end{pmatrix}, \boldsymbol{v}_3 = \begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \end{pmatrix}.$$

47) Schmidt'sches Orthonormalisierungsverfahren. Zu jeder Menge linear unabhängiger Vektoren $\boldsymbol{v}_1, \dots, \boldsymbol{v}_n$ gibt es ein Orthonormalsystem $\boldsymbol{w}_1, \dots, \boldsymbol{w}_n$ mit
$$\mathrm{Span}\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_k\} = \mathrm{Span}\{\boldsymbol{w}_1, \dots, \boldsymbol{w}_k\}, \qquad \text{für alle } 1 \le k \le n.$$
Man beweist diese Aussage mit Induktion über n:

ist $n = 1$, so ist

$$\boldsymbol{w}_1 = \frac{\boldsymbol{v}_1}{\|\boldsymbol{v}_1\|}$$

ein Orthonormalsystem mit $\mathrm{Span}\{\boldsymbol{v}_1\} = \mathrm{Span}\{\boldsymbol{w}_1\}$.
Für $n = 2$ setzt man $\tilde{\boldsymbol{w}}_2 = \boldsymbol{v}_2 + \alpha \boldsymbol{w}_1$ und bestimmt α so, dass $\tilde{\boldsymbol{w}}_2 \perp \boldsymbol{w}_1$:

$$0 = \langle \tilde{\boldsymbol{w}}_2, \boldsymbol{w}_1 \rangle = \langle \boldsymbol{v}_2, \boldsymbol{w}_1 \rangle + \alpha \langle \boldsymbol{w}_1, \boldsymbol{w}_1 \rangle.$$

Es ergibt sich also $\alpha = -\langle \boldsymbol{v}_2, \boldsymbol{w}_1 \rangle$ und somit $\tilde{\boldsymbol{w}}_2 = \boldsymbol{v}_2 - \langle \boldsymbol{v}_2, \boldsymbol{w}_1 \rangle \boldsymbol{w}_1 \neq 0$ da $\boldsymbol{v}_2, \boldsymbol{w}_1$ linear unabhängig sind. Es ist also

$$\boldsymbol{w}_2 = \frac{\tilde{\boldsymbol{w}}_2}{\|\tilde{\boldsymbol{w}}_2\|}$$

der gesuchte Vektor.

1. Vollenden Sie den Beweis mit dem Induktionsschritt.

2. Wenden Sie das Schmidt'sche Orthonormalisierungsverfahren auf den Unterraum

$$\mathrm{Span}\{\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3\} \subset \mathbb{R}^4 \quad \text{mit } \boldsymbol{a}_1 = \begin{pmatrix} 2 \\ 1 \\ 0 \\ 2 \end{pmatrix}, \boldsymbol{a}_2 = \begin{pmatrix} 2 \\ 6 \\ 4 \\ 4 \end{pmatrix}, \boldsymbol{a}_3 = \begin{pmatrix} 1 \\ 9 \\ 5 \\ 8 \end{pmatrix}$$

   bezüglich des Standard-Skalarprodukts an.

48) Sei V ein endlichdimensionaler euklidischer $\mathbb{R}$-Vektorraum und $f_A : V \to V$ eine orthogonale Abbildung. Zeigen Sie, dass jeder Eigenwert λ der zugehörigen Matrix den Betrag 1 hat und dass deren Determinante gleich ± 1 ist.

49) Sei A eine symmetrische $n \times n$ Matrix mit reellen Einträgen ($A^\top = A$). Zeigen Sie:

1. Alle Eigenwerte von A sind reell. (Hinweis: man betrachte A als Matrix einer Abbildung des unitären Vektorraums $\mathbb{C}^n$.)

2. Eigenvektoren zu verschiedenen Eigenwerten sind zueinander orthogonal.

3. Für jeden Eigenwert λ_i mit Eigenraum E_{λ_i} gilt

$$f_A(E_{\lambda_i}) \subseteq E_{\lambda_i} \text{ und } f_A(E_{\lambda_i}^\perp) \subseteq E_{\lambda_i}^\perp,$$

   mit den Orthogonalraum

$$E_{\lambda_i}^\perp := \{\boldsymbol{x} \in \mathbb{R}^n | \langle \boldsymbol{x}, \boldsymbol{y} \rangle = 0 \ \forall \boldsymbol{y} \in E_{\lambda_i}\}.$$

4. Die Matrix A ist diagonalisierbar und hat eine Orthonormalbasis aus Eigenvektoren. (Hinweis: Die Matrix A habe k verschiedene Eigenwerte $\lambda_1, \ldots, \lambda_k$. Sei $B(E_{\lambda_i})$ die Menge der Basiselemente von E_{λ_i} und $W = E_{\lambda_1} \oplus E_{\lambda_2} \oplus \cdots \oplus E_{\lambda_k}$ der Raum mit Basis $\bigcup_1^k B(E_{\lambda_i})$. Zeigen Sie, dass $f_A(W^\perp) \subseteq W^\perp$ gilt und daraus $W^\perp = \{0\}$ und somit die Behauptung folgt.)

50) Für welche Werte λ und μ in $\mathbb{R}$ ist $\langle \boldsymbol{x}, \boldsymbol{y} \rangle_A := \langle A\boldsymbol{x}, \boldsymbol{y} \rangle_{\mathbb{R}^2}$ mit

$$A = \begin{pmatrix} 1 & \lambda \\ \mu & 1 \end{pmatrix}$$

ein Skalarprodukt auf $\mathbb{R}^2$?

Kapitel 3

Analysis I (Grundbegriffe)

In diesem Kapitel wird der Konvergenzbegriff eine zentrale Rolle spielen. Wir werden konvergente Folgen in metrischen Räumen betrachten und verschiedene Beispiele diskutieren. Es schließt sich ein Abschnitt über Zahlenfolgen und Zahlenreihen an. Der Begriff der Stetigkeit einer Abbildung wird ebenfalls in metrischen Räumen definiert und diskutiert. Die Verbindung zur klassischen Analysis wird betont und in einem gesonderten Abschnitt werden Folgen und Reihen von reellen (komplexen) Funktionen einer rellen (komplexen) Variablen betrachtet. Durch die Einführung von Potenzreihen wird die Definition spezieller Funktionen (Exponential-, trigonometrischer und Hyperbelfunktionen sowie ihrer Umkehrfunktionen) vorbereitet und in einem letzten Abschnitt werden diese diskutiert. Als Literatur wird empfohlen [3, Kapitel 8-11], [6, Kapitel 3-4], [5, Kapitel 14-16], [7, Kapitel 2-4].

3.1 Konvergenz in metrischen Räumen

Intuitiv versteht man unter Konvergenz eine schrittweise Annäherung an ein festes Element. Dazu muss man messen, wie weit man von diesem Grenzelement entfernt ist. Es muss also ein Abstandsbegriff vorhanden sein. Dies wird durch Einführung einer Metrik auf einer Menge X realisiert, durch die X zu einem metrischen Raum wird.

Definition 3.1 (Metrik, metrischer Raum). Sei X eine Menge. Eine Abbildung $\rho : X \times X \to \mathbb{R}$ heißt Metrik auf X, wenn sie folgende Eigenschaften für alle $x, y, z \in X$ besitzt:

(M1) $\rho(x, y) \geq 0, \quad \rho(x, y) = 0$, genau dann wenn $x = y$ ist.

(M2) $\rho(x, y) = \rho(y, x)$, d.h. die Metrik ist symmetrisch.

(M3) $\rho(x, z) \leq \rho(x, y) + \rho(y, z)$, d.h. eine Dreiecksungleichung ist erfüllt.

Die Menge X, versehen mit einer Metrik ρ, heißt metrischer Raum.

Beispiele

1° Im Raum $\mathbb{R}^n$ wird zumeist der Euklidische Abstand

$$\rho_2(\boldsymbol{x}, \boldsymbol{y}) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2} = \|\boldsymbol{x} - \boldsymbol{y}\| = \sqrt{\langle \boldsymbol{x} - \boldsymbol{y}, \boldsymbol{x} - \boldsymbol{y} \rangle}$$

betrachtet.
Stellt man sich jedoch vor, dass man vom Ort $\boldsymbol{x}$ zum Ort $\boldsymbol{y}$ nur über rechtwinklige Straßen gelangt, so ist folgende Definition sinnvoll:

$$\rho_1(\boldsymbol{x}, \boldsymbol{y}) = \sum_{i=1}^{n} |x_i - y_i|.$$

Auch

$$\rho_p(\boldsymbol{x}, \boldsymbol{y}) = \left(\sum_{i=1}^{n} |x_i - y_i|^p \right)^{\frac{1}{p}}, \quad p \geq 1$$

ist eine Metrik. Weiterhin spricht man von der Maximummetrik:

$$\rho_\infty(\boldsymbol{x}, \boldsymbol{y}) = \max_i |x_i - y_i|.$$

2° Im $\mathbb{R}^2$ können wir neben den oben erwähnten Metriken die Pariser Metrik einführen (französisches Eisenbahnsystem). Sei $\boldsymbol{x}_0$ ein festes Element (Paris)

$$\rho_{\text{Paris}}(\boldsymbol{x}, \boldsymbol{y}) = \begin{cases} \rho_2(\boldsymbol{x}, \boldsymbol{y}) & \text{falls } \boldsymbol{x}, \boldsymbol{y} \text{ auf einer Geraden durch } \boldsymbol{x}_0 \text{ liegen,} \\ \rho_2(\boldsymbol{x}, \boldsymbol{x}_0) + \rho_2(\boldsymbol{x}_0, \boldsymbol{y}) & \text{sonst.} \end{cases}$$

3° Auf jeder nichtleeren Menge X kann man die diskrete Metrik einführen:

$$d(x, y) = \begin{cases} 1 & \text{falls } x \neq y, \\ 0 & \text{falls } x = 0 \end{cases}$$

ist. X, versehen mit dieser Metrik, heißt diskreter Raum.

4° In der Codierungstheorie führt man in der Menge der n-stelligen Binärwörter (Strings) den sogenannten Hammingabstand ein: Sei $X = \{\boldsymbol{x} = (x_1, \ldots, x_n), x_i = 0 \text{ oder } 1\}$ versehen mit der Metrik

$$\begin{aligned} \rho_H(\boldsymbol{x}, \boldsymbol{y}) = {} & \text{Anzahl der Stellen, an denen sich} \\ & \boldsymbol{x} \text{ und } \boldsymbol{y} \text{ unterscheiden} \\ = {} & \sum_{i=1}^{n} ((x_i + y_i) \bmod 2)\,. \end{aligned}$$

Mit Hilfe der Metrik führen wir den Begriff einer Kugel-Umgebung eines festen Elements $x_0 \in X$ ein.

Definition 3.2. Sei $x_0 \in X, \varepsilon > 0$ eine reelle Zahl. Unter einer Kugelumgebung mit Radius ε und Mittelpunkt x_0 verstehen wir die Menge

$$K_\varepsilon(x_0) = \{x \in X \,:\, \rho(x, x_0) < \varepsilon\}.$$

Diese Definition bildet die Grundlage, um offene und abgeschlossene Mengen in einem metrischen Raum zu definieren.

Definition 3.3. Sei X ein metrischer Raum. Eine Teilmenge $X_0 \subset X$ ist offen, falls zu jedem $x_0 \in X_0$ eine Kugelumgebung von x_0 existiert, die in X_0 liegt.
$X_0 \subset X$ ist abgeschlossen, falls $X \setminus X_0$ offen ist.

Beispiel

Wir betrachten im $\mathbb{R}^2$ den Kreis (ohne Rand)

$$K_1(\mathbf{0}) = \{\boldsymbol{x} \in \mathbb{R}^2 \,:\, |\boldsymbol{x}| = \sqrt{x_1^2 + x_2^2} < 1\}$$

und den Kreis mit Rand

$$\overline{K_1(\mathbf{0})} = \{\boldsymbol{x} \in \mathbb{R}^2 \,:\, |\boldsymbol{x}| = \sqrt{x_1^2 + x_2^2} \leq 1\}.$$

Im ersten Fall liegt eine offene Menge vor, im zweiten Fall eine abgeschlossene Menge.

Wir sind jetzt in der Lage, einen Konvergenzbegriff einzuführen. Dazu betrachten wir eine Folgen von Elementen.

Definition 3.4. Eine Folge ist eine Abbildung von $\mathbb{N}$ in X, so dass jedem $n \in \mathbb{N}$ ein Element $x_n \in X$ zugeordnet wird. Wir schreiben

$$(x_n)_{n \in \mathbb{N}}$$

für eine Folge.

Wir sprechen von in X konvergenten Folgen, falls gilt:

Definition 3.5. Die Folgen $(x_n)_{n \in \mathbb{N}}$ aus dem metrischen Raum X konvergiert gegen das Element $x \in X$, falls es zu jedem $\varepsilon > 0$ einen Index $n_0 = n_0(\varepsilon)$ gibt, so dass

$$x_n \in K_\varepsilon(x) \quad \forall\, n \geq n_0 \tag{3.1}$$

ist. Das Element x heißt Grenzwert der Folge $(x_n)_{n \in \mathbb{N}}$.

Bemerkung 3.6. Die Aussage (3.1) kann auch folgendermaßen formuliert werden: Zu jedem $\varepsilon > 0$ gibt es einen Index $n_0 = n_0(\varepsilon)$, so dass gilt

$$\rho(x_n, x) < \varepsilon \quad \forall\, n \geq n_0(\varepsilon).$$

Man schreibt auch $x_n \xrightarrow{\rho} x$, $\lim_{n\to\infty} x_n \overset{\rho}{=} x$, oder $\rho(x_n, x) \to 0$ für $n \to \infty$.

Eine nicht konvergente Folge heißt divergent.

Folgerung 3.7. Der Grenzwert einer konvergenten Folge ist eindeutig bestimmt.

Beweis. Wir nehmen an, dass x und y verschiedene Grenzwerte der Folge (x_n) seien und setzen $\varepsilon = \frac{\rho(x,y)}{2}$. Dann gibt es ein $n_0 = n_0(\varepsilon)$, so dass $\rho(x_n, x) < \varepsilon$ für alle $n \geq n_0(\varepsilon) = n_0$ ist und ein $n_0' = n_0'(\varepsilon)$, so dass $\rho(x_n, y) < \varepsilon$ für alle $n \geq n_0'(\varepsilon) = n_0'$ ist. Es folgt, dass für $n \geq \max\{n_0, n_0'\}$ gilt

$$\varepsilon = \frac{\rho(x,y)}{2} > \rho(x_n, x) \geq \rho(x,y) - \rho(x_n, y) > 2\varepsilon - \varepsilon = \varepsilon.$$

Dies ist ein Widerspruch und es muss $x = y$ sein. □

Definition 3.8. Eine Folge (x_n) in einem metrischen Raum ist beschränkt, falls es ein $x_0 \in X$ und ein $R > 0$ gibt, so dass

$$\begin{aligned} &x_n \in K_R(x_0) \quad\ \forall\, n \in \mathbb{N}, \quad \text{bzw.} \\ &\rho(x_n, x_0) < R \quad \forall\, n \in \mathbb{N} \end{aligned}$$

ist.

Konvergente Folgen sind beschränkte Folgen.

Beispiele

1° Sei $X = \mathbb{R}$ und $\rho(x, y) = |x - y|$. Wir betrachten die Zahlenfolge $\big(x_n = \frac{1}{n}\big)_{n\in\mathbb{N}}$. Es ist

$$\lim_{n\to\infty} \frac{1}{n} = 0.$$

Wir begründen dies. Sei $\varepsilon > 0$ beliebig gegeben. Wir betrachten die natürliche Zahl $n_0(\varepsilon) = \left[\frac{1}{\varepsilon}\right] + 1 > \frac{1}{\varepsilon}$, wobei

$$[x] := \max\{k \in \mathbb{Z} \,:\, k \leq x\} \tag{3.2}$$

die sogenannte Gauß-Klammer ist, die den größten ganzzahligen Anteil von x bezeichnet.

Für $n \geq n_0(\varepsilon)$ gilt die Abschätzung

$$\rho(\frac{1}{n}, 0) = |\frac{1}{n} - 0| = \frac{1}{n} \leq \frac{1}{n_0(\varepsilon)} < \frac{1}{\frac{1}{\varepsilon}} = \varepsilon.$$

2° Sei $X = \mathbb{R}$ und $\rho(x, y) = |x - y|$. Wir betrachten die Zahlenfolge $\left(\frac{1}{n^q}\right)_{n\in\mathbb{N}}$, wobei $0 < q \in \mathbb{Q}$ ist. Wir zeigen, dass

$$\frac{1}{n^q} \xrightarrow{\rho} 0 \quad \text{für } n \to \infty$$

konvergiert.
Sei $q = \frac{a}{b} \in \mathbb{Q}$ mit $a > 0$ und $b > 0$. Zu $\varepsilon > 0$ wählen wir $n_0 = n_0(\varepsilon) = \left[\sqrt[a]{\frac{1}{\varepsilon^b}}\right] + 1$. Es ist $n_0 > \sqrt[a]{\frac{1}{\varepsilon^b}}$ und $n_0^a > \frac{1}{\varepsilon^b}$, woraus $\frac{1}{\sqrt[b]{n_0^a}} < \varepsilon$ folgt. Daher gilt für $n \geq n_0$

$$n^a \geq (n_0)^a, \quad \frac{1}{n^{\frac{a}{b}}} \leq \frac{1}{n_0^{\frac{a}{b}}} < \varepsilon.$$

3° Sei $X = \mathbb{R}$ und $\rho(x, y) = |x - y|$. Die Folge $(x_n)_{n\in\mathbb{N}}, x_n = (-1)^n$ ist divergent.

4° Sei $X = \mathbb{R}^r$ und $\rho(\boldsymbol{x}, \boldsymbol{y}) = \rho_\infty(\boldsymbol{x}, \boldsymbol{y}) = \max_{1\leq i\leq r} |x_i - y_i|$. Wir betrachten die Folge

$$\boldsymbol{x}_n = \begin{pmatrix} x_1^n \\ \vdots \\ x_r^n \end{pmatrix},$$

wobei $\boldsymbol{x} = (x_1, \ldots, x_r)^\top$ aus $K_1(\boldsymbol{0})$ ist. Wir zeigen, dass der Grenzwert der Nullvektor ist, d.h.

$$\boldsymbol{x}_n \xrightarrow{\rho} \boldsymbol{0} \quad \text{für } n \to \infty, \text{ bzw. } \lim_{n\to\infty} \max_i |x_i| = 0. \tag{3.3}$$

Dazu betrachten wir ein beliebiges $\varepsilon > 0$ und bilden

$$n_0 = \left[\frac{1}{\varepsilon(1 - \rho_\infty(\boldsymbol{x}, \boldsymbol{0}))}\right] + 1.$$

Es ist

$$n_0 > \frac{1}{\varepsilon(1 - \rho_\infty(\boldsymbol{x}, \boldsymbol{0}))}, \quad \rho_\infty(\boldsymbol{x}, \boldsymbol{0}) = \max_i |x_i| = |x_{i0}|. \tag{3.4}$$

Für $n \geq n_0$ gilt

$$|x_{i0}|^n < \frac{1}{n(1 - |x_{i0}|)} \leq \frac{1}{n_0(1 - |x_{i0}|)} \overset{(3.4}{<} \varepsilon.$$

woraus (3.3) folgt.
Um

$$|x|^n < \frac{1}{n(1 - |x|)} \quad \text{für } |x| < 1 \tag{3.5}$$

zu zeigen, greifen wir auf die Bernoulli-Ungleichung zurück:

Bernoulli-Ungleichung:

$$(1+x)^n \geq 1 + nx \quad \text{für } x \geq -1. \tag{3.6}$$

Diese Ungleichung kann durch vollständige Induktion bewiesen werden.

Folgerungen 3.9.

(i) Es gilt

$$\sqrt[n]{x} - 1 \leq \frac{x-1}{n} \quad \text{für } x \geq 0. \tag{3.7}$$

Beweis. Wir setzen $x = \sqrt[n]{y} - 1$ mit $y \geq 0$ in (3.6) ein. Wir erhalten

$$\begin{aligned}(1 + \sqrt[n]{y} - 1)^n = y \geq 1 + n(\sqrt[n]{y} - 1) &= 1 + n\sqrt[n]{y} - n,\\ -n\sqrt[n]{y} + n &\geq -y + 1,\\ \sqrt[n]{y} - 1 &\leq \frac{y-1}{n}.\end{aligned}$$

Durch Umbezeichnung $(y = x)$ folgt (3.7). □

(ii) Es gilt (3.5)

$$|x|^n < \frac{1}{n(1-|x|)} \quad \text{für } |x| < 1.$$

Beweis. Da der Betrag in (3.5) erscheint und für $x = 0$ die Ungleichung erfüllt ist, können wir uns auf den Fall $0 < x < 1$ beschränken. Es ist

$$0 < 1 - \sqrt[n]{x} = \sqrt[n]{x}\left(\frac{1}{\sqrt[n]{x}} - 1\right) \overset{(3.7)}{\leq} 1\left(\frac{\frac{1}{x} - 1}{n}\right) = \frac{1-x}{xn}.$$

Für $z = \sqrt[n]{x}$ folgt

$$\begin{aligned}0 < 1 - z &\leq \frac{1-z^n}{z^n n} < \frac{1}{nz^n},\\ z^n &\leq \frac{1}{n(1-z)}.\end{aligned}$$

Durch Umbezeichnung erhalten wir (3.5). □

Bemerkung 3.10. Im $\mathbb{R}^r$, versehen mit einer Metrik $\rho(\boldsymbol{x}, \boldsymbol{y}) = \|\boldsymbol{x} - \boldsymbol{y}\|$, konvergiert eine Folgen genau dann, wenn ihre Koordinatenfolgen (r Zahlenfolgen) konvergieren. Hierbei ist $\|x\|$ eine Norm, d.h. die Axiome N1, N2, N3 (siehe Kapitel 2.7) gelten.

Übungsaufgaben

51) Gegeben seien eine nichtleere Menge X, $n \in \mathbb{N}$ und die Abbildungen

$$\begin{aligned} d \ : \ X \times X &\to \mathbb{R} \\ (x,y) &\mapsto \begin{cases} 1 & x \neq y \\ 0 & x = y \end{cases} \end{aligned} \qquad \begin{aligned} \rho_{\text{Paris}} \ : \ \mathbb{R}^n \times \mathbb{R}^n &\to \mathbb{R} \\ (\boldsymbol{x}, \boldsymbol{y}) &\mapsto \begin{cases} \rho_2(\boldsymbol{x}, \boldsymbol{y}) & \exists \lambda : \lambda \boldsymbol{x} = \boldsymbol{y} \\ \rho_2(\boldsymbol{x}, 0) + \rho_2(\boldsymbol{y}, 0) & \text{sonst} \end{cases} \end{aligned}$$

Hierbei wird d als Diskrete Metrik und ρ_{Paris} als Pariser Metrik bezeichnet.

1. Beweisen Sie, dass sowohl die Diskrete Metrik als auch die Pariser Metrik die Axiome für Metriken erfüllen und somit im mathematischen Sinn eine Metrik sind.

2. Beweisen Sie ein notwendiges und hinreichendes Kriterium (ohne Verwendung der Metrik in dem Kriterium) für die Konvergenz einer Folge $(m_i)_{i \in \mathbb{N}}$ mit $m_i \in X$ bezüglich d.

52) Zwei weitere Metriken auf $\mathbb{R}^n$ sind gegeben durch

$$\rho_1(\boldsymbol{x}, \boldsymbol{y}) := \sum_{i=1}^{n} |x_i - y_i|\,, \ \text{und} \ \rho_\infty(\boldsymbol{x}, \boldsymbol{y}) := \max_i |x_i - y_i|.$$

1. Zeigen Sie, dass ρ_∞ eine Metrik ist.

2. Skizzieren Sie für $n = 2$ die Mengen $\rho(\mathbf{0}, \boldsymbol{x}) < 1$ für beide Metriken.

53) Zeigen Sie folgende Aussagen.

1. Für jede Metrik ρ gilt die **Vierecksungleichung**

$$|\rho(x,y) - \rho(z,w)| \leq \rho(x,z) + \rho(y,w).$$

2. Sind auf einem Raum E zwei Metriken ρ_1 und ρ_2 erklärt, so nennt man ρ_1 **stärker** als ρ_2, wenn für alle $x \in E$ und alle Folgen $(x_n)_{n \in \mathbb{N}}$ in E aus $\rho_1(x_n, x) \to 0$ stets $\rho_2(x_n, x) \to 0$ folgt. In anderen Worten, wenn Konvergenz einer Folge $(x_n)_{n \in \mathbb{N}}$ bzgl. ρ_1 die Konvergenz der Folge bzgl. ρ_2 mit dem selben Grenzwert impliziert. Ist ρ_1 stärker als ρ_2 und gleichzeitig ρ_2 stärker als ρ_1, so heißen ρ_1 und ρ_2 **äquivalent**.

   Zeigen Sie, dass die Hammingmetrik ρ_H äquivalent zur Diskreten Metrik d auf dem Raum $X = \{0,1\}^n$ für ein $n \in \mathbb{N}$ ist.

54) Sei (X, ρ) ein metrischer Raum. Zeigen Sie, dass

1. die Vereinigung einer beliebigen Anzahl offener Mengen wieder offen ist.

2. der Durchschnitt einer endlichen Anzahl offener Mengen wieder offen ist.

3. sowohl die leere Menge $\emptyset$ als auch die ganze Menge X offen und abgeschlossen sind.

Finden Sie zu 2. ein Beispiel in $\mathbb{R}$, dass der Durchschnitt unendlich vieler offener Mengen nicht mehr offen zu sein braucht.

3.2 Zahlenfolgen und Zahlenreihen

Wir wenden uns jetzt dem Spezialfall zu, dass $X = \mathbb{R}$ oder $X = \mathbb{C}$ ist und betrachten die Metrik $\rho(x, y) = |x - y|$.
Wir formulieren zunächst Sätze, wie man mit Grenzwerten von Folgen arbeitet und formulieren Konvergenzkriterien.

Rechnen mit Grenzwerten

Satz 3.11. *Seien $(x_n)_{n\in\mathbb{N}}$ und $(y_n)_{n\in\mathbb{N}}$ zwei konvergente Folgen in $\mathbb{R}$ mit $x = \lim_{n\to\infty} x_n$, $y = \lim_{n\to\infty} y_n$ und $x_n \le y_n$. Dann ist*

$$x \le y.$$

Beweis. Es ist für ein beliebiges $\varepsilon < 0$

$$\begin{aligned} 0 &\le y_n - x_n = y_n - x_n - y + y - x + x = y - x + (y_n - y) + (x - x_n) \\ &\le y - x + 2\varepsilon \quad \text{für } n > n_0. \end{aligned}$$

Daher ist

$$0 \le \inf_{\varepsilon>0}\{y - x + 2\varepsilon\} = y - x.$$

□

Bemerkung 3.12. Die Abkürzung inf bedeutet: Man betrachte das Infimum, d.h. die größte untere Schranke der Menge $M = \{y - x + 2\varepsilon, \varepsilon > 0\}$. Analog bedeutet die Abkürzung sup: Man betrachte das Supremum, d.h. die kleinste obere Schranke einer Menge M.

Satz 3.13. *Seien $(x_n)_{n\in\mathbb{N}}, (y_n)_{n\in\mathbb{N}}$ konvergente Folgen in $\mathbb{R}$ oder $\mathbb{C}$ und α, β reelle bzw. komplexe Zahlen. Dann gilt*

$$|\lim_{n\to\infty} x_n| = \lim_{n\to\infty} |x_n|, \tag{3.8}$$

$$\lim_{n\to\infty} \alpha x_n = \alpha \lim_{n\to\infty} x_n, \tag{3.9}$$

$$\lim_{n\to\infty} (x_n + y_n) = \lim_{n\to\infty} x_n + \lim_{n\to\infty} y_n, \tag{3.10}$$

$$\lim_{n\to\infty} (x_n y_n) = \lim_{n\to\infty} x_n \lim_{n\to\infty} y_n, \tag{3.11}$$

$$\textit{falls } x_n \ge 0,\ \lim_{n\to\infty} \sqrt[p]{x_n} = \sqrt[p]{\lim_{n\to\infty} x_n} \textit{ für } p \in \mathbb{N}. \tag{3.12}$$

Beweis. Wir zeigen nur Eigenschaft (3.11). Die anderen Eigenschaften werden als Übung empfohlen.
Sei $\lim_{n\to\infty} x_n = x, \lim_{n\to\infty} y_n = y$. Es ist

$$\begin{aligned} |x_n y_n - xy| &= \frac{1}{2}|(x_n + x)(y_n - y) + (y_n + y)(x_n - x)| \\ &\le \frac{1}{2}\left(|x_n + x||y_n - y| + |y_n + y||x_n - x|\right) \\ &< \frac{1}{2}M\varepsilon. \end{aligned}$$

Hierbei haben wir benutzt, dass konvergente Folgen beschränkt sind, d.h. es gibt ein M, so dass $|x_n + x| \leq |x_n| + |x| \leq M$ und $|y_n + y| \leq |y_n| + |y| \leq M$ sind. □

Konvergenzkriterien

Wir beginnen mit einem Einschließungskriterium.

Satz 3.14. *Seien $(x_n)_{n\in\mathbb{N}}, (y_n)_{n\in\mathbb{N}}, (z_n)_{n\in\mathbb{N}}$ reelle Zahlenfolgen. Weiterhin sei*

$$x_n \leq z_n \leq y_n \quad \textit{und} \quad \lim_{n\to\infty} x_n = x = \lim_{n\to\infty} y_n.$$

Dann gilt

$$\lim_{n\to\infty} z_n = x.$$

Beweis. Es ist für ein $\varepsilon > 0$

$$|x - z_n| = \begin{cases} z_n - x \leq y_n - x \leq |y_n - x| < \varepsilon & \text{für } x \leq z_n, \\ x - z_n \leq x - x_n \leq |x - x_n| < \varepsilon & \text{für } x \geq z_n, \end{cases}$$

falls $n \geq \max\{n_x(\varepsilon), n_y(\varepsilon)\}$. □

Zur Berechnung von Grenzwerten kann das Einschließungskriterium häufig recht gut verwendet werden. Dazu braucht man Ungleichungen.

Lemma 3.15. *Es gilt*

$$(1 + x)^n > \frac{n^2}{4}x^2, \quad x > 0, \ \ n \geq 2. \tag{3.13}$$

Beweis. Für $n \geq 2$ gilt

$$(1 + x)^n = \sum_{i=0}^{n} \binom{n}{i} x^i > \binom{n}{2} x^2 = \frac{n(n-1)}{2}x^2 \geq \frac{n^2}{4}x^2.$$

□

Mit Hilfe dieser Ungleichung und der Bernoulli-Ungleichung (3.6) zeigen wir, dass

$$\lim_{n\to\infty} \sqrt[n]{n} = 1$$

ist. Wir setzen in (3.13) $x = \sqrt[n]{n} - 1 \geq 0$. Das ergibt

$$\begin{aligned} n &> \frac{n^2}{4}(\sqrt[n]{n} - 1)^2 \\ \sqrt{n} &> \frac{n}{2}(\sqrt[n]{n} - 1) \\ \frac{2}{\sqrt{n}} + 1 &> \sqrt[n]{n} > 1. \end{aligned}$$

Es ist $\lim_{n\to\infty}(1+\frac{2}{\sqrt{n}}) = 1$ und damit $\sqrt[n]{n} = 1$.
Ähnlich können wir zeigen, dass $\lim_{n\to\infty} \sqrt[n]{a} = 1$ für $a > 1$ ist. Dazu setzen wir in (3.13) $x = \sqrt[n]{a} - 1$. Es ist

$$a > \frac{n^2}{4}(\sqrt[n]{a} - 1)^2$$

also

$$\sqrt{a} > \frac{n}{2}(\sqrt[n]{a} - 1)$$

$$\frac{2\sqrt{a}}{n} + 1 > \sqrt[n]{a} > 1.$$

Daraus folgt die Behauptung für $a > 1$.

Bemerkung 3.16. Es gilt auch für $0 < a \leq 1$, dass $\lim_{n\to\infty} \sqrt[n]{a} = 1$ ist.

Weiterhin betrachten wir monoton wachsende bzw. monoton fallende Folgen in $\mathbb{R}$.

Definition 3.17. Die Folgen $(x_n)_{n\in\mathbb{N}}$, $x_n \in \mathbb{R}$, heißt

- streng monoton wachsend, falls $x_n < x_{n+1}$,
- streng monoton fallend, falls $x_{n+1} < x_n$,
- monoton wachsend, falls $x_n \leq x_{n+1}$,
- monoton fallend, falls $x_{n+1} \leq x_n$

für alle $n \in \mathbb{N}$ ist.

Satz 3.18. *Beschränkte monoton wachsende bzw. fallende Folgen sind konvergent.*

Beweis. Wir zeigen diese Aussage für monoton wachsende Folgen. Sei $x_n \leq x_{n+1} \leq c$. Dann existiert eine kleinste obere Schranke

$$c_{\min} = \sup\{x_n, n \in \mathbb{N}\}.$$

Zu jedem $\varepsilon > 0$ existiert ein $n_0 = n_0(\varepsilon) \in \mathbb{N}$, so dass

$$c_{\min} - \varepsilon < x_{n_0} \tag{3.14}$$

gilt. Würde die Verneinung von (3.14) gelten, dann existiert ein $\varepsilon_0 > 0$, so dass gilt

$$c_{\min} - \varepsilon_0 \geq x_n \quad \forall\, n \in \mathbb{N}.$$

Dann wäre $c_{\min} - \varepsilon$ kleinste obere Schranke und nicht $c_{\min}$.
Aus (3.14) folgt

$$c_{\min} - \varepsilon < x_{n_0} \leq x_n \leq c_{\min} \quad \forall\, n \geq n_0.$$

Damit ist $c_{\min} - x_n = |c_{\min} - x_n| < \varepsilon$ und

$$\lim_{n\to\infty} x_n = c_{\min}.$$

□

Beispiel

Wir sehen uns die Folge $\left(x_n = \left(1+\frac{1}{n}\right)^n\right)_{n\in\mathbb{N}}$ an. Wir zeigen zunächst, dass diese Folge monoton wachsend ist. Es ist

$$\begin{aligned}
\frac{x_{n+1}}{x_n} &= \frac{\left(\frac{n+2}{n+1}\right)^{n+1}}{\left(\frac{n+1}{n}\right)^n} = \left(\frac{(n+2)n}{(n+1)(n+1)}\right)^{n+1}\frac{n+1}{n} \\
&= \left(\frac{n^2+2n+1-1}{n^2+2n+1}\right)^{n+1}\frac{n+1}{n} \\
&= \left(1-\frac{1}{(n+1)^2}\right)^{n+1}\frac{n+1}{n} \\
&\overset{(3.6)}{\geq} \left(1-\frac{n+1}{(n+1)^2}\right)\frac{n+1}{n} \\
&= \left(1-\frac{1}{n+1}\right)\left(\frac{n+1}{n}\right) \\
&= \left(\frac{n+1-1}{n+1}\right)\frac{n+1}{n} = 1.
\end{aligned}$$

Wir überlegen nun, dass die Folge beschränkt ist. Dazu betrachten wir die Folge $(y_n)_{n\in\mathbb{N}}$ mit

$$y_n = \left(1+\frac{1}{n}\right)^{n+1} > x_n$$

und zeigen, dass sie monoton fallend ist. Es gilt

$$\begin{aligned}
\frac{y_n}{y_{n-1}} &= \frac{\left(\frac{n+1}{n}\right)^{n+1}}{\left(\frac{n}{n-1}\right)^n} = \frac{n+1}{n}\left(\frac{(n+1)(n-1)}{n^2}\right)^n \\
&= \frac{n+1}{n}\left(1-\frac{1}{n^2}\right)^n.
\end{aligned}$$

Wir wenden die Bernoullische Ungleichung (3.6) an:

$$(1+x)^n \geq 1+nx, \quad x > -1.$$

Dann ist

$$\frac{1}{1+nx} \geq (1+x)^{-n} \geq (1-x)^n,$$

da $\frac{1}{1+x} \geq 1-x$ ist. Dadurch können wir abschätzen, indem wir $x = \frac{1}{n^2}$ setzen,

$$\frac{y_n}{y_{n-1}} \leq \frac{n+1}{n}\frac{1}{1+\frac{n}{n^2}} = \frac{(n+1)n^2}{n(n^2+n)} = 1.$$

Somit haben wir erhalten

$$y_1 \geq y_2 \geq \cdots \geq y_n \geq x_n \geq x_{n-1} \geq \cdots \geq x_1,$$

d.h. $y_1 \geq x_n$ für alle $n \in \mathbb{N}$. Die Folge $\left(x_n = \left(1 + \frac{1}{n}\right)^n\right)_{n\in\mathbb{N}}$ hat als Grenzwert die Eulersche Zahl

$$\lim_{n\to\infty} \left(1 + \frac{1}{n}\right)^n = e = 2,7182818284\ldots .$$

Wir führen den Begriff einer Teilfolge ein.

Definition 3.19. Sei $(x_n)_{n\in\mathbb{N}}$ eine Folge reeller Zahlen und $\mathbb{N}' \subset \mathbb{N}$ eine abzählbar unendliche Teilmenge. $(x_{n'})_{n'\in\mathbb{N}'}$ heißt Teilfolge von $(x_n)_{n\in\mathbb{N}}$.

Wir sind jetzt in der Lage zu erklären, was wir unter einem Häufungspunkt verstehen.

Definition 3.20. Ein Element $a \in \mathbb{R}$ oder $\mathbb{C}$ heißt Häufungspunkt der reellen bzw. komplexen Zahlenfolgen $(x_n)_{n\in\mathbb{N}}$, falls es eine gegen a konvergente Teilfolge $(x_{n'})_{n'\in\mathbb{N}'}$ gibt.

Beispiel

Die reelle Zahlenfolge $((-1)^n)_{n\in\mathbb{N}}$ besitzt die Häufungspunkte ± 1.

Wir kommen nun zu einem wichtigen Konvergenzkriterium, das gestattet, aus den Abständen der Folgenglieder Rückschlüsse auf das Konvergenzverhalten zu ziehen.

Satz 3.21 (Cauchysches Konvergenzkriterium). *Eine Folge reeller Zahlen $(x_n)_{n\in\mathbb{N}}$ konvergiert genau dann, wenn sie eine Cauchyfolge ist, d.h. für alle $\varepsilon > 0$ existiert ein $n_0(\varepsilon)$, so dass gilt*

$$|x_n - x_m| < \varepsilon \quad \forall\, n, m > n_0(\varepsilon).$$

Beweisskizze.

a) $(x_n)_{n\in\mathbb{N}}$ sei eine konvergente Folge mit $\lim_{n\to\infty} x_n = x$. Dann gibt es für ein $\frac{\varepsilon}{2}$ einen Index $n_0(\frac{\varepsilon}{2})$, so das

$$\begin{aligned} |x - x_n| &< \frac{\varepsilon}{2} \quad \text{für } n \geq n_0(\frac{\varepsilon}{2}) \\ \text{und } |x - x_m| &< \frac{\varepsilon}{2} \quad \text{für } m \geq n_0(\frac{\varepsilon}{2}). \end{aligned}$$

Es folgt, dass $|x_n - x_m| \leq |x_n - x| + |x_m - x| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$ für $n, m > n_0(\frac{\varepsilon}{2}) = \hat{n}(\varepsilon)$ ist.

b) In [1, S.94] wird gezeigt: Sei $(x_n)_{n\in\mathbb{N}}$ eine Cauchyfolge, dann ist $(x_n)_{n\in\mathbb{N}}$ eine konvergente Folge. Hierbei wird auf Teilfolgen zurückgegriffen.

Bemerkung 3.22. Augustin Louis Cauchy war ein französischer Mathematiker, der von 1789 bis 1857 lebte.

Zahlenreihen

Reihen sind durch spezielle Folgen, deren Elemente Partialsummen (endliche Teilsummen) sind, erklärt.

Definition 3.23. Sei $(a_k)_{k\in\mathbb{N}}$ ein Folge in $X=\mathbb{R}, X=\mathbb{C}, X=\mathbb{R}^r$ oder $X=\mathbb{C}^r$ (oder in einem Vektorraum X mit Norm). Der Ausdruck

$$\sum_{k=0}^{\infty} a_k \tag{3.15}$$

heißt (unendliche) Reihe, a_k heißen die Glieder der Reihe. Unter einer Partialsumme versteht man die endliche Summe

$$s_n = \sum_{k=0}^{n} a_k.$$

Konvergiert die Folge der Partialsummen $(s_n)_{n\in\mathbb{N}}$ zu einem Grenzwert s in X, dann sagt man, $\sum_{k=0}^{\infty} a_k$ ist konvergent und

$$\lim_{n\to\infty} s_n = s =: \sum_{k=0}^{\infty} a_k$$

ist die Summe der Reihe (3.15). Besitzt die Folge (s_n) keinen Grenzwert, so spricht man von divergenten Reihen.

Bemerkung 3.24. Man definiert auch

$$\sum_{k=0}^{\infty} a_k := \left(\sum_{k=0}^{n} a_k\right)_{n\in\mathbb{N}},$$

d.h. die Reihe wird als Folge der Partialsummen aufgefasst. Konvergiert diese, so spricht man von konvergenten Reihen.
Die Summierung kann auch bei $k=k_0$ beginnen.

Beispiele

1° Es sei $a_k = \frac{1}{k(k-1)}, k=2,3,\ldots$. Wir wollen untersuchen, ob

$$\sum_{k=2}^{\infty} \frac{1}{k(k-1)} = \frac{1}{1\cdot 2} + \frac{1}{2\cdot 3} + \frac{1}{3\cdot 4} + \ldots$$

konvergiert. Es ist

$$\frac{1}{k(k-1)} = \frac{1}{k-1} - \frac{1}{k}$$

und

$$\begin{aligned} s_n = a_2 + \cdots + a_n &= \left(\frac{1}{1} - \frac{1}{2}\right) + \left(\frac{1}{2} - \frac{1}{3}\right) + \cdots + \left(\frac{1}{n-1} - \frac{1}{n}\right) \\ &= 1 - \frac{1}{n}. \end{aligned}$$

Da $\lim_{n\to\infty} s_n = \lim_{n\to\infty}\left(1 - \frac{1}{n}\right) = 1$ ist folgt

$$\sum_{k=2}^{\infty} \frac{1}{k(k-1)} = 1. \tag{3.16}$$

2° Wir betrachten die geometrische Reihe $\sum_{k=0}^{\infty} z^k$ für $z \in \mathbb{C}$. Wir haben bereits durch vollständige Induktion (1.13) gezeigt, dass für $z \neq 1$

$$s_n = 1 + z + z^2 + \cdots + z^n = \frac{1 - z^{n+1}}{1 - z}$$

ist. Wir erhalten

$$\sum_{k=0}^{\infty} z^k = \begin{cases} \frac{1}{1-z} & \text{für } |z| < 1, \\ \text{divergent} & \text{für } |z| \geq 1. \end{cases}$$

3° Die harmonische Reihe $\sum_{k=1}^{\infty} \frac{1}{k}$ ist divergent. Es ist

$$\frac{1}{m+1} + \cdots + \frac{1}{2m} \geq m\frac{1}{2m} = \frac{1}{2}.$$

Daher gilt

$$\sum_{k=1}^{\infty} \frac{1}{k} = 1 + \frac{1}{2} + \underbrace{\left(\frac{1}{3} + \frac{1}{4}\right)}_{m=2=2^1 \geq \frac{1}{2}} + \underbrace{\left(\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8}\right)}_{m=4=2^2 \geq \frac{1}{2}} + \underbrace{\left(\frac{1}{9} + \cdots + \frac{1}{16}\right)}_{m=8=2^3 \geq \frac{1}{2}}$$

und

$$S_{2^j} \geq \frac{j+2}{2}.$$

Es folgt:

$$\lim_{j\to\infty} (S_{2^j}) \geq \lim_{j\to\infty} \left(\frac{j}{2} + 1\right) = \infty.$$

Wir formulieren nun notwendige und hinreichende Konvergenzkriterien für Reihen. Wir erinnern, was die Implikation „$A \Rightarrow B$ “bedeutet: B ist notwendige Bedingung für A und A ist hinreichende Bedingung für B.

Satz 3.25 (Notwendiges Konvergenzkriterium). *Die Reihe $\sum_{k=0}^{\infty} a_k$ sei konvergent. Dann ist die Folge $(a_k)_{k\in\mathbb{N}_0}$ eine Nullfolge, d.h. $\lim_{k\to\infty} a_k = 0$, bzw. $|a_k| \to 0$ für $k \to \infty$.*

Beweis. Da eine konvergente Reihe vorliegt, ist die Folge der Partialsummen $(s_n)_{n\in\mathbb{N}_0}$ eine Cauchyfolge, d.h. zu jedem $\varepsilon > 0$ gibt es eine natürliche Zahl $n_0(\varepsilon)$, so dass

$$|s_m - s_n| = \left|\sum_{k=n+1}^{m} a_k\right| < \varepsilon \quad \text{für } m, n \geq n_0(\varepsilon), m \geq n. \tag{3.17}$$

Wählen wir $m = n + 1$, dann ist

$$|s_n - s_{n+1}| = |a_{n+1}| < \varepsilon \quad \text{für } n \geq n_0(\varepsilon)$$

und

$$|a_k| < \varepsilon \quad \text{für } k \geq n_0(\varepsilon) + 1,$$

woraus die Behauptung folgt. □

Bemerkung 3.26.

- Dieses Kriterium ist nicht hinreichend. Man betrachte z.B. die harmonische Reihe, die divergent ist und für die gilt

$$a_k = \frac{1}{k} \to 0 \quad \text{für } k \to \infty.$$

- Die Ungleichung (3.17) heißt Cauchysches Konvergenzkriterium für Reihen und ist **notwendig und hinreichend** dafür, dass eine Zahlenreihe konvergiert.

Lemma 3.27. *Eine Reihe mit nichtnegativen reellen Gliedern für die die Folgen der Partialsummen beschränkt ist, ist eine konvergente Reihe.*

Beweis. Wir betrachten die Reihe, die zur Folge $(a_k)_{k\in\mathbb{N}}, a_k \geq 0, a_k \in \mathbb{R}$, gehört.
Die Folge der Partialsummen $(s_n)_{n\in\mathbb{N}}$ ist monoton wachsend und beschränkt. Aus Satz 3.18 folgt die Behauptung. □

Wir betrachten jetzt alternierende Reihen und das Leibnizsche Konvergenzkriterium.

Definition 3.28. Eine Reihe $\sum_{k=0}^{\infty} a_k$ heißt alternierend, wenn die Folge $(a_k)_{k\in\mathbb{N}_0}$ wechselnde Vorzeichen besitzt.

Satz 3.29 (Hinreichendes Leibniz-Kriterium (1646-1716)). *Sei $(a_k)_{k\in\mathbb{N}_0}$ eine reelle Zahlenfolge, für die gilt $a_k \geq 0, a_k \geq a_{k+1}$ und $\lim_{k\to\infty} a_k = 0$. Dann ist die alternierende Reihe konvergent, d.h.*

$$\sum_{k=0}^{\infty} (-1)^k a_k = s$$

und es gilt

$$|s - \sum_{k=0}^{n} (-1)^k a_k| \leq a_{n+1}.$$

Beweis. Für jedes $\varepsilon > 0$ existiert ein $n_0(\varepsilon)$, so dass $|(-1)^k a_k| = a_k < \varepsilon$ für $k \geq n_0(\varepsilon)$ ist. Daher gilt für die Partialsummen s_m und s_n, $m \geq n \geq n_0(\varepsilon)$, falls n ungerade ist:

$$\begin{aligned} |s_m - s_n| &= |(-1)^{n+1}a_{n+1} + (-1)^{n+2}a_{n+2} + \cdots + (-1)^{m-1}a_{m-1} + (-1)^m a_m| \\ &= (a_{n+1} - a_{n+2}) + \cdots + \begin{cases} (a_{m-1} - a_m) & \text{für } (-1)^{m+n} = 1 \\ a_m & \text{für } (-1)^{m+n} = -1 \end{cases} \\ &= a_{n+1} - (a_{n+2} - a_{n+3}) - \cdots + \begin{cases} -(a_{m-2} - a_{m-1}) - a_m & \text{für } (-1)^{m+n} = 1 \\ -(a_{m-1} - a_m) & \text{für } (-1)^{m+n} = -1 \end{cases} \\ &\leq a_{n+1}. \end{aligned}$$

Dies kann durch eine genauere Fallunterscheidung für gerade und ungerade m und n nachgerechnet werden. □

Beispiel

Es gilt

$$\sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} = \ln 2.$$

Achtung!

Die Partialsummen und damit auch die Summationsreihenfolge in einer unendlichen Reihe sind genau vorgeschrieben. Umordnen und Klammersetzung kann gefährlich sein.

Beispiele

1° **Umordnung**

Wir betrachten

$$\sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} \pm \ldots.$$

Wir ordnen um und betrachten

$$\underbrace{1 - \frac{1}{2}}_{>0} + \underbrace{\frac{1}{3} - \frac{1}{4}}_{>0} + \underbrace{\left(\frac{1}{5} + \frac{1}{7}\right) - \frac{1}{6}}_{>0} + \underbrace{\left(\frac{1}{9} + \frac{1}{11} + \frac{1}{13} + \frac{1}{15}\right) - \frac{1}{8}}_{>0} + \ldots. \tag{3.18}$$

Da für gerades m

$$\frac{1}{m+1} + \frac{1}{m+3} + \cdots + \frac{1}{2m-1} \geq \frac{m}{2} \cdot \frac{1}{2m} = \frac{1}{4}$$

gilt, wird die umgeordnete Reihe (3.18) divergent sein.

2° **Klammerung**

Wir betrachten die Reihe

$$\sum_{k=0}^{\infty}(-1)^k = 1-1+1-1+\cdots = \begin{cases}(1-1)+(1-1)+\cdots = 0,\\ 1+(-1+1)+(-1+1)+\cdots = 1.\end{cases}$$

Die Klammerungen führen zu unterschiedlichen Ergebnissen.

Um Umordnen zu können, braucht man stärkere Eigenschaften.

Definition 3.30. Eine Reihe $\sum_{k=0}^{\infty} a_k$ heißt **absolut konvergent**, falls die Reihe $\sum_{k=0}^{\infty} |a_k|$ konvergiert.

Satz 3.31. *Jede absolut konvergente Reihe ist konvergent.*

Beweis. Für jedes $\varepsilon > 0$ existiert ein $n_0 = n_0(\varepsilon)$, so dass

$$|s_m - s_n| = |\sum_{k=n+1}^{m} a_k| \le \sum_{k=n+1}^{m} |a_k| < \varepsilon \quad \text{für}\, n, m \ge n_0(\varepsilon).$$

Aus dem Cauchy-Kriterium (3.17) folgt die Behauptung. □

Bemerkung 3.32. Die Umkehrung dieses Satzes gilt nicht. So ist die Reihe $\sum_{k=1}^{\infty} \frac{(-1)^k}{k}$ konvergent, $\sum_{k=1}^{\infty} \frac{1}{k}$ jedoch nicht.

Das folgende Kriterium ist häufig sehr hilfreich.

Satz 3.33 (Majorantenkriterium). *Seien $(a_k)_{k\in\mathbb{N}}$ und $(b_k)_{k\in\mathbb{N}}$ Folgen mit $|a_k| \le b_k$. Ist $\sum_{k=1}^{\infty} b_k$ konvergent, dann ist $\sum_{k=1}^{\infty} |a_k|$ konvergent und damit auch $\sum_{k=1}^{\infty} a_k$. $\sum_{k=1}^{\infty} b_k$ ist eine konvergente Majorante von $\sum_{k=1}^{\infty} |a_k|$ bzw. $\sum_{k=1}^{\infty} a_k$.*

Beweis. Sei $s_n = \sum_{k=1}^{n} |a_k|$. Dann ist für jedes $\varepsilon > 0$

$$|s_m - s_n| = \sum_{k=n+1}^{m} |a_k| \le \sum_{k=n+1}^{m} b_k = |\tilde{s}_m - \tilde{s}_n| < \varepsilon \quad \text{für}\, m, n \ge n_0(\varepsilon),$$

wobei $\tilde{s}_m = \sum_{k=1}^{m} b_k$, $\tilde{s}_n = \sum_{k=1}^{n} b_k$ sind. □

Es gibt auch ein **Minorantenkriterium**:

Sei $a_k \ge c_k$, $c_k \ge 0$ und $\sum_{k=1}^{\infty} c_k$ sei divergent. Dann divergiert auch $\sum_{k=1}^{\infty} a_k$.

Beispiele

1° Wir betrachten $\sum_{n=1}^{\infty} \frac{1}{n^2}$. Da $\frac{1}{n^2} \le \frac{1}{n(n-1)}$ für $n \ge 2$ und $\sum_{n=2}^{\infty} \frac{1}{n(n-1)} = 1$ ist (nach (3.16)) haben wir eine konvergente Majorante gefunden und es ist

$$\sum_{n=1}^{\infty} \frac{1}{n^2} \le 1 + \sum_{n=2}^{\infty} \frac{1}{n(n-1)} = 2.$$

2° Wir sehen uns $\sum_{n=1}^{\infty}\frac{1}{\sqrt{n}}$ an. Es ist $\frac{1}{\sqrt{n}} \geq \frac{1}{n}$ und $\sum_{n=1}^{\infty}\frac{1}{n}$ ist eine divergente Minorante. Daher ist $\sum_{n=1}^{\infty}\frac{1}{\sqrt{n}}$ divergent.

3° Wir untersuchen die Reihe $\sum_{n=1}^{\infty}\frac{1}{2^n+3^n}$ auf Konvergenz. Da $\frac{1}{2^n+3^n} \leq \frac{1}{3^n} = \left(\frac{1}{3}\right)^n$ gilt, ist $\sum_{n=1}^{\infty}\left(\frac{1}{3}\right)^n$ als geometrische Reihe eine konvergente Majorante.

Wir formulieren nun zwei weitere hinreichende Konvergenzkriterien, das Wurzelkriterium und das Quotientenkriterium.

Satz 3.34 (Wurzelkriterium). *Die Reihe $\sum_{k=1}^{\infty} a_k$ ist absolut konvergent, falls es ein q mit $0 \leq q < 1$ gibt und ein $k_0 \in \mathbb{N}$ existiert, so dass*

$$\sqrt[k]{|a_k|} \leq q \quad \text{für alle } k \text{ mit} k \geq k_0$$

gilt.

Beweis. Es ist $|a_k| \leq q^k$ und $\sum_{k=1}^{\infty} q^k$ ist konvergent (geometrische Reihe). □

Satz 3.35 (Quotientenkriterium). *Sei $a_k \neq 0$ für $k \geq k_0$. Die Reihe $\sum_{k=0}^{\infty} a_k$ ist absolut konvergent, falls es ein q mit $0 \leq q < 1$ gibt, so dass*

$$\left|\frac{a_{k+1}}{a_k}\right| \leq q \quad \text{für alle} k \geq k_0 \tag{3.19}$$

ist.

Beweis. Aus der Ungleichung (3.19) folgt

$$|a_{k+1}| \leq q|a_k| \leq q^2|a_{k-1}| \leq \cdots \leq q^{k+1-k_0}|a_{k_0}|.$$

Daher ist

$$\sum_{k=0}^{\infty} |a_k| \leq \frac{a_{k_0}}{q^{k_0}} \sum_{k=0}^{\infty} q^k,$$

und wir haben eine konvergente Majorante gefunden. □

Bemerkungen 3.36.

- Falls $\sqrt[k]{|a_k|} \geq 1$ für alle $k \geq k_0$ gilt, dann ist $\sum_{k=1}^{\infty} a_k$ divergent.
- Falls $\left|\frac{a_{k+1}}{a_k}\right| \geq 1$ für alle $k \geq k_0$ gilt, dann ist $\sum_{k=1}^{\infty} a_k$ divergent.
- Ist $\lim_{k\to\infty} \sqrt[k]{|a_k|} = q < 1$, so ist $\sum_{k=1}^{\infty} a_k$ absolut konvergent.
  Ist $\lim_{k\to\infty} \sqrt[k]{|a_k|} = Q > 1$, so ist $\sum_{k=1}^{\infty} a_k$ divergent.
- Existiert der Grenzwert $\lim_{k\to\infty}\left|\frac{a_{k+1}}{a_k}\right| = q < 1$, so ist $\sum_{k=a}^{\infty} a_k$ absolut konvergent.
  Ist $\lim_{k\to\infty}\left|\frac{a_{k+1}}{a_k}\right| = Q > 1$, so ist $\sum_{k=a}^{\infty} a_k$ divergent. Im Fall $q = Q = 1$ ist Konvergenz und Divergenz möglich.

Beispiele

1° Wir betrachten die Reihe $\sum_{k=0}^{\infty} \frac{x^k}{k!} = 1 + x + \frac{x^2}{2} + \frac{x^3}{3!} + \ldots$, wobei x eine reelle Zahl sei. Es ist nach dem Quotientenkriterium

$$\lim_{k\to\infty} \left| \frac{x^{k+1}k!}{x^k(k+1)!} \right| = \lim_{k\to\infty} \left| \frac{x}{k+1} \right| = 0$$

und die Reihe ist konvergent.

2° Wir sehen uns die divergente harmonische Reihe $\sum_{k=1}^{\infty} \frac{1}{k}$ an. Das Quotientenkriterium liefert

$$\left| \frac{a_{k+1}}{a_k} \right| = \frac{k}{k+1} < 1,$$

aber es gilt nicht $\frac{k}{k+1} \leq q < 1$ für alle $k \geq k_0$, da $\lim_{k\to\infty} \left| \frac{a_{k+1}}{a_k} \right| = 1$ ist.

Das Wurzelkriterium führt ebenfalls auf $\lim_{k\to\infty} \sqrt[k]{\frac{1}{k}} = 1$.

Satz 3.37 (Umordnungssatz). [1, S.118] *Sei $\sum_{k=1}^{\infty} a_k$ absolut konvergent und $\tau : \mathbb{N} \to \mathbb{N}$ sei eine bijektive Abbildung. Dann ist die umgeordnete Reihe $\sum_{k=1}^{\infty} a_{\tau(k)}$ ebenfalls absolut konvergent und es gilt*

$$\sum_{k=1}^{\infty} a_k = \sum_{k=1}^{\infty} a_{\tau(k)}.$$

Das Cauchy-Produkt von Reihen

Produkte von endlichen Summen werden durch multiplizieren der entsprechenden Summanden gebildet. Dabei kommt es auf die Reihenfolge der Multiplikationen nicht an. Bei der Multiplikation von unendlichen Reihen $\left(\sum_{k=0}^{\infty} a_k\right)\left(\sum_{l=0}^{\infty} b_l\right)$ müssen alle Produkte $a_k b_l$ auftreten und man muss geschickt sortieren. Die Idee von Cauchy wird durch folgendes Schema beschrieben:

$$\begin{array}{ccccccccc}
a_0b_0 & & a_0b_1 & & a_0b_2 & & a_0b_3 & \ldots \\
 & \swarrow & & \swarrow & & \swarrow & & \\
a_1b_0 & & a_1b_1 & & a_1b_2 & & a_1b_3 & \ldots \\
 & \swarrow & & \swarrow & & & & \\
a_2b_0 & & a_2b_1 & & a_2b_2 & & a_2b_3 & \ldots \\
 & \swarrow & & & & & & \\
a_3b_0 & & a_3b_1 & & a_3b_2 & & a_3b_3 & \ldots \\
\vdots & & \vdots & & \vdots & & \vdots &
\end{array}$$

Damit wird

$$\left(\sum_{k=0}^{\infty} a_k\right)\left(\sum_{l=0}^{\infty} b_l\right) = a_0b_0 + \underbrace{a_0b_1 + a_1b_0}_{n=1} + \underbrace{a_0b_2 + a_1b_1 + a_2b_0}_{n=2} + \cdots + \sum_{j=0}^{n} a_j b_{n-j} + \ldots .$$

Satz 3.38. [1, S.123] *Seien* $\sum_{k=0}^{\infty} a_k$ *und* $\sum_{l=0}^{\infty} b_l$ *absolut konvergent und* $c_n = \sum_{j=0}^{n} a_j b_{n-j}$. *Dann konvergiert das Cauchy-Produkt* $\sum_{n=0}^{\infty} c_n$ *absolut und für die Grenzwerte gilt:*

$$\left(\sum_{k=0}^{\infty} a_k\right)\left(\sum_{l=0}^{\infty} b_l\right) = \sum_{n=0}^{\infty} c_n.$$

Beispiel

Wir betrachten die absolut konvergenten Reihen $e^x := \sum_{k=0}^{\infty} \frac{x^k}{k!}$ und $e^y := \sum_{l=0}^{\infty} \frac{y^l}{l!}$ für reelle Zahlen x und y. Das Cauchy-Produkt ist $\sum_{n=0}^{\infty} c_n$ mit

$$\begin{aligned} c_n = \sum_{j=0}^{n} \frac{x^j}{j!}\frac{y^{n-j}}{(n-j)!} &= \sum_{j=0}^{\infty} \frac{n!}{n!j!(n-j)!} x^j y^{n-j} \\ &= \frac{1}{n!}\sum_{j=0}^{\infty} \left(\begin{array}{c} n \\ j \end{array}\right) x^j y^{n-j} \\ &= \frac{1}{n!}(x+y)^n. \end{aligned}$$

Damit ist

$$\sum_{n=0}^{\infty} c_n = \sum_{n=0}^{\infty} \frac{(x+y)^n}{n!} = \left(\sum_{k=0}^{\infty} \frac{x^k}{k!}\right)\left(\sum_{l=0}^{\infty} \frac{y^l}{l!}\right) = e^x e^y = e^{x+y}.$$

Übungsaufgaben

55) Zeigen Sie: Sind $(x_n)_{n\in\mathbb{N}}$ und $(y_n)_{n\in\mathbb{N}}$ konvergente Folgen in $\mathbb{R}$ oder $\mathbb{C}$, dann gilt

1. $\lim\limits_{n\to\infty}(x_n + y_n) = \lim\limits_{n\to\infty} x_n + \lim\limits_{n\to\infty} y_n$,
2. $\lim\limits_{n\to\infty} \sqrt[p]{x_n} = \sqrt[p]{\lim\limits_{n\to\infty} x_n}$ für $x_n \geq 0$ und $p \in \mathbb{N}$.

56) Zeigen Sie, dass

1. $\lim\limits_{n\to\infty} \sqrt[n]{n} = 1$.
2. für die Folge $x_0 = 1$ und $x_{n+1} = 1 + \dfrac{1}{x_n}$ gilt

   $$|x_n - g| \leq \frac{1}{g^{n+1}} \quad \text{und somit} \quad \lim_{n\to\infty} x_n = g\,,$$

   wobei g der Goldene Schnitt, also die positive Lösung der Gleichung $g = 1 + \frac{1}{g}$ ist.
3. für jedes $w > 0$ die Folge $x_0 = w$ und $x_{n+1} = \dfrac{1}{2}\left(x_n + \dfrac{w}{x_n}\right)$ gegen $\sqrt{w}$ konvergiert.

57) Untersuchen Sie die Folgen auf Konvergenz und geben Sie die Häufungspunkte und gegen die Häufungspunkte konvergierende Teilfolgen an.

1. $x_n = \mathrm{i}^n + \frac{1}{n}$,
2. $x_n = (-1)^n \frac{1}{n}$.

58) Überprüfen Sie die folgenden Reihen auf Konvergenz bzw. Divergenz. Untersuchen Sie die Reihen **e), f)** und **g)** auf absolute Konvergenz und geben Sie bei den Reihen **h)** und **i)** an, für welche $z \in \mathbb{C}$ die Reihe jeweils konvergiert.

$$\textbf{a)}\ \sum_{n=1}^{\infty} \frac{n}{(4n-3)(4n-1)}\,, \qquad \textbf{b)}\ \sum_{n=1}^{\infty} \frac{n^2}{2^n}\,, \qquad \textbf{c)}\ \sum_{n=1}^{\infty} \frac{n!}{n^n}\,,$$

$$\textbf{d)}\ \sum_{n=1}^{\infty} \left(\frac{n}{n+1}\right)^{n^2}\,, \qquad \textbf{e)}\ \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{\sqrt{n}}\,, \qquad \textbf{f)}\ \sum_{n=1}^{\infty} \frac{(-1)^{\frac{n(n-1)}{2}}}{2^n}\,,$$

$$\textbf{g)}\ \sum_{n=2}^{\infty} \frac{(-1)^n}{\sqrt{n}+(-1)^n}\,, \qquad \textbf{h)}\ \sum_{n=1}^{\infty} n^{1000} z^n\,, \qquad \textbf{i)}\ \sum_{n=1}^{\infty} \frac{z^n}{n}\,.$$

59) Für $z \in \mathbb{C}$, $|z| < 1$ gilt $\displaystyle\sum_{k=0}^{\infty} z^k = \frac{1}{1-z}$.
Bestimmen Sie das Cauchy-Produkt

$$\frac{1}{(1-z)^2} = \left(\sum_{k=0}^{\infty} z^k\right) \cdot \left(\sum_{k=0}^{\infty} z^k\right).$$

60) Wie viele Glieder der Reihenentwicklung von

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} \pm \dots\ .$$

sind zu berücksichtigen, um $\cos\left(\frac{\pi}{6}\right)$ auf 5 Dezimalstellen genau zu erhalten?

3.3 Stetige Abbildungen in metrischen Räumen

Der Begriff einer Funktion, oder allgemeiner, einer Abbildung f eines metrischen Raumes X in den metrischen Raum Y, spielt in der Analysis eine grundlegende Rolle. Durch die Metriken ρ_X in X bzw. ρ_Y in Y können wir messen, wie groß der Abstand der Urbildpunkte x_1 und x_2 in X ist und wie weit die Bildpunkte $f(x_1)$ und $f(x_2)$ in Y voneinander entfernt sind. Insbesondere möchte man die Eigenschaft erfassen: Falls eine Folge von Elemenen $(x_n)_{n\in\mathbb{N}}$ den Grenzwert x_0 besitzt, dann soll diese Eigenschaft auch auf die Bildfolge $(f(x_n))_{n\in\mathbb{N}}$ zutreffen, die den Grenzwert $f(x_0)$ besitzen sollte. Diese Eigenschaft wird als Stetigkeit der Abbildung $f : X \to Y$ im Punkt x_0 bezeichnet werden.
Der Begriff der Stetigkeit hat seinen Ursprung in der klassischen Physik und lässt sich bis ins Griechische Altertum zurückverfolgen. Jedoch erst 1817 wurde er von Bolzano, 1821

von Cauchy und schließlich 1860 von Weierstraß in die heute benutzte Form gebracht. Wir geben zwei äquivalente Definitionen für die Stetigkeit an, die erste beruht auf dem Konvergenzbegriff für Folgen, die zweite benutzt den Begriff der Kugelumgebungen.

Definition 3.39. Seien X und Y metrische Räume mit den Metriken ρ_X und ρ_Y. Die Abbildung $f : X \to Y$ heißt **stetig im Punkt** $x_0 \in X$, wenn aus $x_n \overset{\rho_X}{\to} x_0$ stets

$$f(x_n) \overset{\rho_Y}{\to} f(x_0) \tag{3.20}$$

folgt.

Kurz geschrieben lautet (3.20):

$$\lim_{n\to\infty} f(x_n) = f(\lim_{n\to\infty} x_n). \tag{3.21}$$

Definition 3.40. Seien X und Y metrische Räume mit den Metriken ρ_X und ρ_Y. Die Abbildung $f : X \to Y$ heißt **stetig im Punkt** $x_0 \in X$, wenn es zu jedem $\varepsilon > 0$ ein $\delta = \delta(\varepsilon)$ gibt, so dass gilt

$$\rho_Y(f(x), f(x_0)) < \varepsilon \quad \text{für alle } x \text{ mit } \rho_X(x, x_0) < \delta(\varepsilon) \tag{3.22}$$

ist.

Satz 3.41. *Die Definitionen 3.39 und 3.40 sind äquivalent.*

Beweis.

a) Es gelte die Definition 3.39. Wir nehmen an, dass Definition 3.40 falsch ist, d.h. es existiert ein $\varepsilon_0 > 0$, so dass für alle $\delta > 0$ ein $x = x_\delta \in K_\delta(x_0)$ existiert, so dass

$$\rho_Y(f(x_\delta), f(x_0)) \geq \varepsilon_0 \tag{3.23}$$

ist. Wir betrachten $\delta = \frac{1}{n}$. Dann besagt (3.23): zu jedem $n \in \mathbb{N}$ gibt es ein x_n mit $\rho_X(x_n, x_0) < \frac{1}{n}$, so dass $\rho_Y(f(x_n), f(x_0)) \geq \varepsilon_0$ ist. Da $x_n \overset{\rho_X}{\to} x_0$ ist, aber $\rho_Y(f(x_n), f(x_0)) \geq \varepsilon_0$ gilt, erhalten wir einen Widerspruch zur Definition 3.39.

b) Es gelte Definition 3.40. Wir betrachten eine konvergente Folge $(x_n)_{n\in\mathbb{N}}$, $x_n \in X$, mit $\lim_{n\to\infty} x_n \overset{\rho_X}{=} x_0$, d.h. für ein $\varepsilon > 0$ und das nach Definition 3.40 zugehörige $\delta(\varepsilon)$ gilt:

$$\rho_X(x_n, x_0) < \delta(\varepsilon) \quad \text{für } n = n_0(\delta(\varepsilon)) = n_0(\varepsilon)$$

und damit folgt aus (3.22) $\rho_Y(f(x_n), f(x_0)) < \varepsilon$. Somit gilt Definition 3.39.

□

Definition 3.42. Die Funktion $f : X \to Y$ ist stetig im Gebiet $D \subset X$, falls f in allen Punkten aus D stetig ist.

Beispiele

1° Es sei $X = Y = \mathbb{R}$, $\rho_X(a,b) = \rho_Y(a,b) = |a-b|$. Wir betrachten die Abbildung

$$f : \mathbb{R} \to \mathbb{R}, \quad f(x) = x^2.$$

Wir zeigen $f(x)$ ist stetig in allen Punkten $x_0 \in \mathbb{R}$. Dazu betrachten wir eine beliebige Zahlenfolge

$$(x_n)_{n\in\mathbb{N}} \quad \text{mit} \quad \lim_{n\to\infty} x_n = x_0.$$

Für die Bildfolge $(f(x_n))_{n\in\mathbb{N}}$ gilt:

$$|f(x_n) - f(x_0)| = |x_n^2 - x_0^2| = |x_n - x_0||x_n + x_0|.$$

Da eine konvergente Folge beschränkt ist, finden wir ein M, so dass gilt

$$|x_n + x_0| \leq |x_n| + |x_0| \leq M \quad \forall\, n \geq n_0.$$

Damit erhalten wir

$$0 \leq \lim_{n\to\infty} |f(x_n) - f(x_0)| \leq M \lim_{n\to\infty} |x_n - x_0| = 0.$$

2° Es sei $X = Y = \mathbb{R}$, $\rho(x,y) = |x-y|$. Die Heavisidesche Sprungfunktion

$$H(x) = \begin{cases} 0 & \text{für } x < 0, \\ 1 & \text{für } x \geq 0 \end{cases}$$

ist unstetig im Punkt $x_0 = 0$.
Betrachten wir eine Nullfolge $(x_n)_{n\in\mathbb{N}}$, $x_n \to 0$, $x_n < 0$, dann ist

$$|H(x_n) - H(x_0)| = |0 - 1| = 1.$$

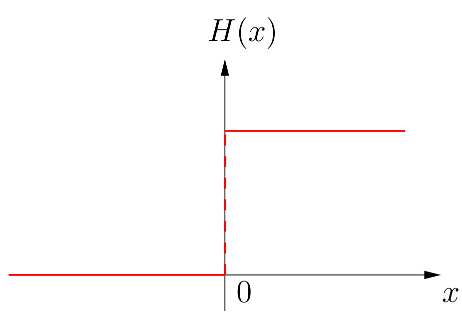


Abbildung 3.1: Die Sprungfunktion $H(x)$

3° Sei X ein $\mathbb{R}$-Vektorraum mit Skalarprodukt $\langle\cdot,\cdot\rangle$ und $Y=\mathbb{R}$, $\rho_X(x,y)=\sqrt{\langle x-y,x-y\rangle}$, $\rho_Y(a,b)=|a-b|$. Wir halten ein Element z aus X fest und betrachten die Abbildung $f:X\to\mathbb{R}$, $x\to f(x,z)=\langle x,z\rangle$. Diese Abbildung ist stetig in allen Punkten (Elementen) aus X, d.h. für in X konvergente Folgen $(x_n)_{n\in\mathbb{N}}$ gilt nach Formel (3.21):

$$\lim_{n\to\infty}\langle x,z\rangle=\langle\lim_{n\to\infty}x_n,z\rangle.$$

Um dies zu sehen, betrachten wir ein $x_0\in X$ und eine beliebige Folgen $(x_n)_{n\in\mathbb{N}}$ mit $\|x_n-x_0\|=\sqrt{\langle x_n-x_0,x_n-x_0\rangle}\to 0$. Dann gilt

$$\begin{aligned}|f(x_n)-f(x_0)|&=|\langle x_n,z\rangle-\langle x_0,z\rangle|=|\langle x_n-x_0,z\rangle|\\&\overset{(2.58)}{\le}\|x_n-x_0\|\|z\|\to 0.\end{aligned}$$

Das Skalarprodukt $g:X\times X\to\mathbb{R}$, $g(x,y)=\langle x,y\rangle$, ist auch stetig in allen Punkten $(x_0,y_0)\in X\times X$. Sei nämlich $(x_n,y_n)_{n\in\mathbb{N}}$ eine Folge von Elementen aus $X\times X$ mit

$$\|x_n-x_0\|=\sqrt{\langle x_n-x_0,x_n-x_0\rangle}\to 0,\quad \|y_n-y_0\|=\sqrt{\langle y_n-y_0,y_n-y_0\rangle}\to 0$$

für $n\to\infty$. Dann gilt

$$\begin{aligned}|g(x_n,y_n)-g(x_0,y_0)|&=|\langle x_n,y_n\rangle-\langle x_0,y_0\rangle|\\&=|\langle x_n-x_0,y_n\rangle+\langle x_0,y_n-y_0\rangle|\\&\le\|x_n-x_0\|\|y_n\|+\|x_0\|\|y_n-y_0\|\to 0\end{aligned}$$

für $n\to\infty$.

Reelle bzw. komplexe Funktionen einer reellen bzw. komplexen Variablen

Es sei $X=Y=\mathbb{R}$ oder $X=Y=\mathbb{C}$, $\rho(x,y)=|x-y|$. Man benutzt folgende Schreibweise für die Stetigkeit einer Funktion im Punkt x_0:

$$\begin{aligned}f(x_0)=\lim_{x\to x_0}f(x)\quad\Leftrightarrow\quad&\forall\,\varepsilon>0\ \exists\,\delta=\delta(\varepsilon)>0,\\&\forall\,x\text{ mit }|x-x_0|<\delta(\varepsilon)\text{ gilt }|f(x)-f(x_0)|<\varepsilon.\end{aligned}$$

Den Ausdruck $\lim_{x\to x_0}f(x)$ bezeichnet man als Funktionenlimes, bzw. Grenzwert in Punkt x_0. Der Funktionenlimes kann auch für nicht notwendig stetige Funktionen definiert werden.

Definition 3.43. Die Zahl $a=\lim_{x\to x_0}f(x)$ ist der **Funktionenlimes** (Grenzwert) der Funktion f im Punkt x_0, falls für alle $\varepsilon>0$ ein $\delta=\delta(\varepsilon)>0$ existiert, so dass für alle x mit $|x-x_0|<\delta$ gilt

$$|f(x)-a|<\varepsilon.$$

Weiterhin werden linksseitige und rechtsseitige Grenzwerte für $f : \mathbb{R} \to \mathbb{R}$ eingeführt:
f besitzt an der Stelle x_0 einen **linksseitigen Grenzwert**

$$\lim_{x \to x_0 -} f(x) = a_l,$$

falls für alle $\varepsilon > 0$ ein $\delta = \delta(\varepsilon) > 0$ existiert, so dass für alle x mit $x < x_0$ und $|x - x_0| < \delta$ gilt

$$|f(x) - a_l| < \varepsilon.$$

f besitzt an der Stelle x_0 einen **rechtsseitigen Grenzwert**

$$\lim_{x \to x_0 +} f(x) = a_r,$$

falls für alle $\varepsilon > 0$ ein $\delta = \delta(\varepsilon) > 0$ existiert, so dass für alle x mit $x > x_0$ und $|x - x_0| < \delta$ gilt

$$|f(x) - a_r| < \varepsilon.$$

Im Beispiel 2, in dem wir die Heaviside-Funktion betrachtet haben, ist

$$\lim_{x \to 0-} H(x) = a_l = 0,$$
$$\lim_{x \to 0+} H(x) = a_r = 1.$$

Wir betrachten jetzt Verknüpfungen von stetigen Funktionen.

Satz 3.44. *Seien f und g Abbildungen von X in X, $X = \mathbb{R}$ oder $X = \mathbb{C}$, die im Punkt x_0 stetig sind. Dann sind auch folgende Abbildungen in x_0 stetig:*

$$|f| : x \to |f(x)|, \tag{3.24}$$
$$cf : x \to cf(x), \quad c \in \mathbb{R}, \ \textit{oder}\, c \in \mathbb{C}, \tag{3.25}$$
$$f + g : x \to f(x) + g(x), \tag{3.26}$$
$$fg : x \to f(x)g(x), \tag{3.27}$$
$$\frac{f}{g} : x \to \frac{f(x)}{g(x)}, \quad g(x_0) \neq 0, \tag{3.28}$$
$$\sqrt[p]{f} : x \to \sqrt[p]{f(x)}, \quad f(x) \geq 0, \ p \in \mathbb{N}, \tag{3.29}$$
$$f \circ g : x \to f(g(x)). \tag{3.30}$$

Beweis. Wir führen den Beweis nur für (3.28) aus.
Wir betrachten eine beliebige Folge (x_n), die zu x_0 konvergiert. Da $g(x_0) \neq 0$ ist, finden wir einen Index n_0, so dass $g(x_n) \neq 0$ für $n \geq n_0$ ist. Wir erhalten

$$\begin{aligned}
\left| \frac{f(x_n)}{g(x_n)} - \frac{f(x_0)}{g(x_0)} \right| &= \left| \frac{f(x_n)g(x_0) - f(x_0)g(x_n)}{g(x_0)g(x_n)} \right| \\
&= \left| \frac{(f(x_n) - f(x_0))g(x_0) - f(x_0)(g(x_n) - g(x_0))}{g(x_0)g(x_n)} \right| \\
&\leq \frac{|f(x_n) - f(x_0)|}{|g(x_n)|} + \left| \frac{f(x_0)}{g(x_0)g(x_n)} \right| |g(x_n) - g(x_0)| \to 0 \quad \text{für } n \to \infty.
\end{aligned}$$

$\square$

Folgerung 3.45. Polynome, $P_k(x) = a_0 + a_1 x + \cdots + a_k x^k$, $x \in \mathbb{R}$ oder $\mathbb{C}$, und rationale Funktionen $\frac{P_k(x)}{P_m(x)}$, $P_m(x_0) \neq 0$ sind stetig in x_0.

Übungsaufgaben

61)

1. Sei (X, d) ein metrischer Raum mit der diskreten Metrik d und Y ein metrischer Raum mit einer beliebigen Metrik. Zeigen Sie, dass jede Funktion $f : X \to Y$ stetig ist.

2. Zeigen Sie, dass jede Metrik auf einer endlichen Menge X äquivalent zur diskreten Metrik ist.

62) Seinen X, Y und Z metrische Räume und $f, g : X \to Y$, $h : Y \to Z$ Abbildungen. Zeigen Sie folgende Aussagen.

1. Sind f und g stetig in $x_0 \in X$, so ist auch $fg : X \to Y$ mit $fg(x) = f(x)g(x)$ stetig in x_0.

2. Ist f stetig in x_0 und h stetig in $y_0 = f(x_0)$, so ist ist auch $h \circ f : X \to Z$ mit $h \circ f(x) = h(f(x))$ stetig in x_0.

Hinweis: Wenn man will, kann man auch $X = Y = Z = \mathbb{R}$ annehmen; der Beweis ist im Wesentlichen der Gleiche.

63) Der Graph einer Funktion $f : \mathbb{R}^2 \to \mathbb{R}$ kann geometrisch als Fläche im $\mathbb{R}^3$ interpretiert werden. Stellen Sie die folgenden Funktionen graphisch dar und überprüfen Sie diese auf Stetigkeit in $(0,0)$ (bezüglich der Euklidischen Metrik $\rho(\mathbf{x}, \mathbf{y}) = \sqrt{|x_1 - y_1|^2 + |x_2 - y_2|^2}$).

a)
$$f(x_1, x_2) = \begin{cases} \frac{x_1 + x_2}{x_1 - x_2}, & \text{für } x_1 \neq x_2 \\ 0, & \text{für } x_1 = x_2. \end{cases}$$

b)
$$f(x_1, x_2) = \begin{cases} \frac{x_1 x_2^2}{x_1^2 + x_2^2}, & \text{für } (x_1, x_2) \neq (0,0) \\ 0, & \text{für } (x_1, x_2) = (0,0). \end{cases}$$

c)
$$f(x_1, x_2) = \begin{cases} \frac{x_1 x_2^2}{x_1^2 + x_2^4}, & \text{für } (x_1, x_2) \neq (0,0) \\ 0, & \text{für } (x_1, x_2) = (0,0). \end{cases}$$

3.4 Folgen und Reihen von Funktionen

Wir betrachten Abbildungen f, $f_n : \mathbb{R} \to \mathbb{R}$, bzw. f,$f_n : \mathbb{C} \to \mathbb{C}$.

Definition 3.46 (Punktweise und gleichmäßige Konvergenz von Funktionenfolgen). Sei $D \subseteq \mathbb{R}$ oder $(\mathbb{C})$ der Definitionsbereich der Funktionen f und f_n.
Die Folge $(f_n)_{n \in \mathbb{N}}$ heißt **punktweise konvergent** in D gegen f, geschrieben als

$\lim_{n\to\infty} f_n(x) \stackrel{p.w.}{=} f(x)$, falls für alle $x \in D$ und für alle $\varepsilon > 0$ ein $n_0 = n_0(\varepsilon, x) \in \mathbb{N}$ existiert, so dass gilt

$$|f_n(x) - f(x)| < \varepsilon \quad \text{für alle } n \geq n_0. \tag{3.31}$$

Die Folge $(f_n)_{n\in\mathbb{N}}$ heißt **gleichmäßig konvergent** gegen f, falls n_0 in (3.31) unabhängig von x ist, d.h. $n_0 = n_0(\varepsilon)$ hängt nicht von der Lage der Punkte $x \in D$ ab:

$$\lim_{n\to\infty} f_n(x) \stackrel{gl.}{=} f(x).$$

Beispiele

1° Sei $f_n(x) = 1 + \frac{1}{n}x$. Wir betrachten als Definitionsbereich $D = [-1, +1] \subset \mathbb{R}$.
Die Folge (f_n) konvergiert gleichmäßig auf $[-1, +1]$ zur Funktion f, $f(x) \equiv 1$, für alle $x \in [-1, +1]$. Sei $\varepsilon > 0$ beliebig. Dann ist

$$\left|1 + \frac{1}{n}x - 1\right| = \left|\frac{1}{n}x\right| \leq \frac{1}{n} < \varepsilon \quad \text{für } n \geq n_0(\varepsilon) = \left[\frac{1}{\varepsilon}\right] + 1 > \frac{1}{\varepsilon}.$$

Hierbei bezeichnet $[\cdot]$ wieder die Gaußsche Klammer, welche in Kapitel 3.1 durch (3.2) eingeführt wurde.

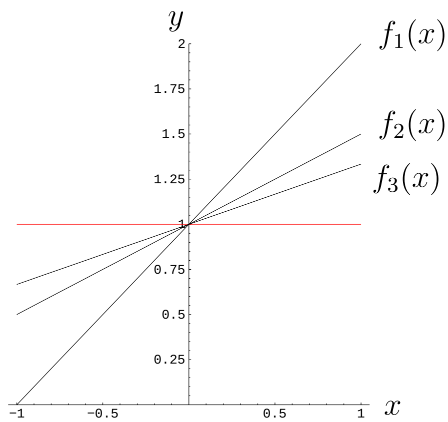


Abbildung 3.2: Die Funktionenfolge $f_n(x) = 1 + \frac{1}{n}x$

2° Sei $D = [-1, +1] \subset \mathbb{R}$, $f_n(x) = x^{2n}$.
Die Folge (f_n) konvergiert punktweise auf $[-1, +1]$, aber nicht gleichmäßig zur Funktion f:

$$f(x) = \begin{cases} 0 & \text{für } -1 < x < +1, \\ 1 & \text{für } x = \pm 1. \end{cases}$$

Sei $\varepsilon > 0$ beliebig. Dann ist für ein festes $x_0 \in (-1, +1)$ (siehe Beispiel 4 S. 139)

$$|f_n(x_0) - f(x_0)| = |x_0^{2n}| < \varepsilon \quad \text{für } n \geq n_0(\varepsilon) = \left[\frac{1}{\varepsilon(1 - |x_0|)}\right] + 1.$$

Für $x_0 = \pm 1$ ist $f_n(\pm 1) = (\pm 1)^{2n} = 1$ für alle $n \in \mathbb{N}$.

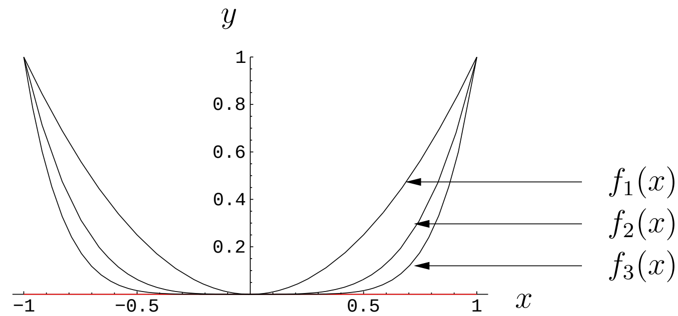


Abbildung 3.3: Die Funktionenfolge $f_n(x) = x^{2n}$

Aus der gleichmäßigen Konvergenz folgt die punktweise Konvergenz, aber nicht umgekehrt. Die gleichmäßige Konvergenz garantiert uns, dass der Grenzwert stetiger Funktionen wieder einen stetige Funktion ist. Dabei ist eine Funktion stetig im Definitionsbereich $D \subset \mathbb{R}$ ($D \subset \mathbb{C}$), falls sie in jedem Punkt aus D stetig ist.

Satz 3.47. *Konvergiert eine Folge stetiger Funktionen gleichmäßig in D, dann ist der Grenzwert dort stetig.*

Beweis. Sei $\varepsilon > 0$, $x_0 \in D$.
Wähle $n_0 = n_0(\varepsilon)$ so, dass $|f_{n_0}(x) - f(x)| < \frac{\varepsilon}{3}$ für alle $x \in D$ gilt.
Wähle $\delta = \delta(\varepsilon)$ so, dass $|f_{n_0}(x) - f_{n_0}(x_0)| < \frac{\varepsilon}{3}$ für alle $x \in D$ mit $|x - x_0| < \delta$ ist.
Dann gilt

$$\begin{aligned} |f(x) - f(x_0)| &\leq |f(x) - f_{n_0}(x)| + |f_{n_0}(x) - f_{n_0}(x_0)| + |f_{n_0}(x_0) - f(x_0)| \\ &< \frac{3}{3}\varepsilon = \varepsilon \quad \text{für } |x - x_0| < \delta. \end{aligned}$$

□

Bemerkung 3.48. Reihen von Funktionen $\sum_{n=0}^{\infty} f_n$ konvergieren punktweise bzw. gleichmäßig, wenn die Folge der Partialsummen punktweise bzw. gleichmäßig konvergieren.

Kriterien für gleichmäßige Konvergenz

Folgende Kriterien sind leicht zu verifizieren:

1. **Cauchy Kriterium**:
   Falls es zu jedem $\varepsilon > 0$ einen Index $n_0(\varepsilon)$ gibt, so dass für alle $x \in D$ gilt

   $$|f_n(x) - f_m(x)| < \varepsilon \quad \text{für alle } n, m \geq n_0(\varepsilon),$$

   dann existiert ein $f : D \to f(D)$ mit

   $$f \stackrel{\text{gl.}}{=} \lim_{n\to\infty} f_n.$$

2. **Weierstraß Kriterium** (Karl Theodor Wilhelm Weierstraß lebte 1815-1897 und lehrte in Berlin.)
   Falls für eine Reihe von Funktionen gilt:
$$|f_n(x)| \leq c_n \quad \text{für alle } x \in D \text{ und } \sum_{n=0}^{\infty} c_n \text{ ist konvergent,}$$
   dann konvergieren $\sum_{n=0}^{\infty} f_n$ und $\sum_{n=0}^{\infty} |f_n|$ gleichmäßig auf D.

Beispiel

Sei $D = \{c \in \mathbb{C} : |z| < 1\}$. Die Reihe $\sum_{n=0}^{\infty} z^n$ ist auf D nicht gleichmäßig konvergent. Betrachten wir die Folge der Partialsummen
$$s_n = \sum_{k=0}^{n} z^k$$
und wählen $\varepsilon = \frac{1}{2}$. Dann gibt es für jedes $n \in \mathbb{N}$ ein $z \in D$, so dass
$$|s_{n+1} - s_n| = |z|^{n+1} \geq \frac{1}{2}$$
ist.

Potenzreihen

Eine wichtige Rolle spielen die Potenzreihen, durch die eine große Klasse von Funktionen dargestellt werden.

Definition 3.49. Sei $z_0 \in \mathbb{C}$ und $(a_k)_{n \in \mathbb{N}}$ eine Folge komplexer Zahlen. Dann heißt
$$\sum_{k=0}^{\infty} a_k (z - z_0)^k$$
eine Potenzreihe um z_0 mit den Koeffizienten a_k. Sie ist für alle $z \in \mathbb{C}$ definiert, für die die Zahlenreihe $\sum_{k=0}^{\infty} a_k (z - z_0)^k$ konvergiert. Falls $z, z_0 \in \mathbb{R}$ und $a_k \in \mathbb{R}$ sind, dann heißt die Potenzreihe reell.

Beispiele

- $\sum_{n=0}^{\infty} \frac{z^n}{n!} =: e^z$ (konvergiert für alle $z \in \mathbb{C}$),
- $\sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} z^{2n+1} =: \sin z$ (konvergiert für alle $z \in \mathbb{C}$),
- $\sum_{n=0}^{\infty} \frac{(-1)^n}{2n!} z^{2n} =: \cos z$ (konvergiert für alle $z \in \mathbb{C}$),
- $\ln(1 + x) = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} x^n$ konvergiert für $-1 < x \leq 1$,
- $\ln x = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} (x - 1)^n$ konvergiert für $0 < x \leq 2$.

Wir überlegen, wie wir die Konvergenz von Potenzreihen überprüfen können. Dazu wird der Begriff des Konvergenzradius eingeführt.

Satz 3.50. *Es sei $\sum_{n=0}^{\infty} a_n(z-z_0)^n$ eine Potenzreihe.*

1. *Es gibt einen eindeutig bestimmten Konvergenzradius $R \geq 0$, so dass gilt*
$$\sum_{n=0}^{\infty} a_n(z-z_0)^n \begin{cases} \text{absolut konvergent} & \text{für } |z-z_0| < R, \\ \text{divergent} & \text{für } |z-z_0| > R. \end{cases} \tag{3.32}$$

2. *Falls*
$$r = \lim_{n\to\infty} \frac{|a_n|}{|a_{n+1}|} \text{ oder } r = \lim_{n\to\infty} \frac{1}{\sqrt[n]{|a_n|}} \tag{3.33}$$
existieren, dann ist $r = R$.

3. *Falls $R > 0$ und $0 < \rho < R$, dann konvergiert $\sum_{n=0}^{\infty} a_n(z-z_0)^n$ gleichmäßig für $|z-z_0| \leq \rho$.*

Beweis. Wir setzen im Beweis $z_0 = 0$. Wir betrachten die Menge
$$K := \{|z| : \sum_{n=0}^{\infty} a_n z^n \text{ konvergent}\}.$$

Wir zeigen, dass
$$R := \begin{cases} \sup K & \text{falls } K \text{ beschränkt ist,} \\ \infty & \text{falls } K \text{ unbeschränkt ist} \end{cases}$$
der oben beschriebene Konvergenzradius ist.
Wir beweisen die 1. Aussage:

1. Fall
Sei $R = 0$. Dann ist die Reihe nur für $z = 0$ konvergent.

2. Fall
Sei $R > 0$. Wir betrachten ein $z \in \mathbb{C}$ mit $|z| < R$. Wir finden ein $w \in \mathbb{C}$ mit $|z| < |w| < R$, so dass auch $\sum_{n=0}^{\infty} a_n w^n$ konvergent ist. Die Glieder $a_n w^n$ müssen ab einem Index n_0 gleichmäßig beschränkt sein, d.h. es existiert ein c, so dass
$$|a_n w^n| \leq c \quad \text{für alle } n$$
ist. Wir können mit Hilfe dieses w eine konvergente Majorante für $\sum_{n=0}^{\infty} a_n z^n$ finden. Es ist nämlich
$$|a_n z^n| = \left| a_n \frac{w^n}{w^n} z^n \right| = |a_n w^n| \left| \frac{z^n}{w^n} \right| \leq c \left| \frac{z}{w} \right|^n. \tag{3.34}$$
Da $\left|\frac{z}{w}\right| < 1$ gilt, ist $\sum_{n=0}^{\infty} c \left|\frac{z}{w}\right|^n$ eine konvergente Majorante und $\sum_{n=0}^{\infty} a_n z^n$ ist absolut konvergent.
Für $|z| > R$ ist $\sum_{n=0}^{\infty} a_n z^n$ divergent nach Definition von R.

Kommen wir nun zur 2. Aussage:
Wir wenden das Quotientenkriterium oder das Wurzelkriterium auf die Potenzreihe $\sum_{n=0}^{\infty} a_n z^n$ an. Es ist

$$\lim_{n\to\infty}\left|\frac{a_{n+1}z^{n+1}}{a_n z^n}\right| = \lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right||z| = \frac{1}{r}|z| < 1 \quad \text{für alle } z \text{ mit } |z| < r,$$

oder

$$\lim_{n\to\infty}\sqrt[n]{|a_n z^n|} = \lim_{n\to\infty}|z|\sqrt[n]{a_n} = \frac{1}{r}|z| < 1 \quad \text{für alle } z \text{ mit } |z| < r.$$

Da für $|z| > r$ Divergenz vorliegt, muss $r = R$ sein.

Wir zeigen nun die 3. Aussage.
Dazu wählen wir ein $w \in \mathbb{C}$ mit $\rho < |w| < R$. Für $|z| \leq \rho$ gilt die Abschätzung (3.34)

$$|a_n z^n| \leq c\left|\frac{z}{w}\right|^n \leq c\left|\frac{\rho}{w}\right|^n.$$

Aus dem Weierstraß-Kriterium folgt die gleichmäßige Konvergenz im abgeschlossenen Kreis: $|z| \leq \rho$. □

Bemerkung 3.51. In Satz 3.50 ist indirekt angenommen worden, dass die Koeffizienten $a_n \neq 0$ sind. Dies kann abgeschwächt werden, indem die Existenz der Grenzwerte (3.33) durch Existenz des „limes superior“

$$r = \overline{\lim_{n\to\infty}}\left|\frac{a_n}{a_{n+1}}\right| \quad \text{oder} \quad r = \overline{\lim_{n\to\infty}}\frac{1}{\sqrt[n]{|a_n|}}$$

ersetzt werden. Hierbei bedeutet (siehe Definition 3.20)

$$\overline{\lim_{n\to\infty}} x_n = \text{ größter Häufungspunkt von } (x_n)_n.$$

Wir sehen uns u.a. die Beispiele (S. 163) an.

1° Für $e^z = \sum_{n=0}^{\infty}\frac{z^n}{n!}$ haben wir

$$\lim_{n\to\infty}\left|\frac{a_n}{a_{n+1}}\right| = \lim_{n\to\infty}\frac{(n+1)!}{n!} = \lim_{n\to\infty}(n+1) = \infty,$$

woraus $R = r = \infty$ folgt.

2° Für $\sum_{n=0}^{\infty}\frac{(-1)^n}{(2n+1)!}z^{2n+1}$ gilt, dass $e^z = \sum_{n=0}^{\infty}\frac{z^n}{n!}$ eine konvergente Majorante für alle $z \in \mathbb{C}$ ist. Daher gilt auch hier, $R = \infty$.

3° Für $\ln(1+x) = \sum_{n=1}^{\infty}\frac{(-1)^{n-1}}{n}x^n$ erhalten wir

$$\lim_{n\to\infty}\left|\frac{a_n}{a_{n+1}}\right| = \lim_{n\to\infty}\frac{n+1}{n} = 1.$$

4° Für $\sum_{n=0}^{\infty} z^n = \frac{1}{1-z}$ ist

$$\lim_{n\to\infty} \left| \frac{a_n}{a_{n+1}} \right| = 1.$$

Wir sehen, dass diese Reihe z.B. für $z = 2$ divergiert, trotzdem $\frac{1}{1-z}$ für $z = 2$ erklärt ist. Das bedeutet, dass $\frac{1}{1-z}$ für $|z| > 1$ nicht durch die geometrische Reihe dargestellt wird. Folgende Frage tritt auf: Kann man im Bereich $|z| > 1$ eine andere Potenzreihendarstellung von $\frac{1}{1-z}$ finden?
Es ist

$$\begin{aligned}\frac{1}{1-z} = \frac{1}{2-(z+1)} &= \frac{1}{2}\left(\frac{1}{1-\frac{z+1}{2}}\right)\\ &= \frac{1}{2}\sum_{n=0}^{\infty}\left(\frac{z+1}{2}\right)^n \quad \text{für } \frac{|z+1|}{2} < 1, \text{ d.h.}\\ &= \sum_{n=0}^{\infty}\frac{1}{2^{n+1}}(z-(-1))^n \quad \text{für } |z-(-1)| < 2.\end{aligned}$$

Ist die Potenzreihendarstellung einer Funktion eindeutig? Diese Frage wird durch den Identitätssatz beantwortet.

Satz 3.52 (Identitätssatz). *Seien $f(z) = \sum_{n=0}^{\infty} a_n(z-z_0)^n$ und $g(z) = \sum_{n=0}^{\infty} b_n(z-z_0)^n$ Potenzreihen mit positivem Konvergenzradius. Existiert eine Folge $(z_k)_{k\in\mathbb{N}}$ mit $z_k \to z_0$, $z_k \neq z_0$ und $f(z_k) = g(z_k)$ für alle $k \in \mathbb{N}$, dann ist $a_n = b_n$ und $f = g$.*

Das heißt, eine Funktion kann auf höchstens eine Weise in eine Potenzreihe um z_0 entwickelt werden.

Beweis. Wir setzen der Einfachheit halber $z_0 = 0$ und nehmen an, dass beide Reihen in $K_R(0)$, $R > 0$ konvergent sind. Wir wenden die vollständige Induktion an.

Induktionsanfang: Wir zeigen $a_0 = b_0$.
Es ist

$$0 = f(z_k) - g(z_k) = \sum_{n=0}^{\infty}(a_n - b_n)z_k^n = a_0 - b_0 + z_k\sum_{n=1}^{\infty}(a_n - b_n)z_k^{n-1}. \tag{3.35}$$

Da $z_k \to 0$ für $k \to \infty$ und $\sum_{n=1}^{\infty}(a_n - b_n)z_k^{n-1}$ beschränkt ist, muss $a_0 = b_0$ sein.

Induktionsschritt: Sei $a_0 = b_0$, $a_1 = b_1, \ldots, a_{n_0} = b_{n_0}$. Wir zeigen, dass $a_{n_0+1} = b_{n_0+1}$ ist.
Es ist

$$0 = f(z_k) - g(z_k) = z_k^{n_0+1}\sum_{n=n_0+1}^{\infty}(a_n - b_n)z_k^{n-(n_0+1)}$$

und

$$0 = \frac{1}{z_k^{n_0+1}}\left(f(z_k) - g(z_k)\right) = (a_{n_0+1} - b_{n_0+1}) + (a_{n_0+2} - b_{n_0+2})z_k^1 + \ldots .$$

Es folgt wie in (3.35), dass $a_{n_0+1} = b_{n_0+1}$ sein muss. □

Übungsaufgaben

64) Der maximale **Definitionsbereich** einer Funktion $f : \mathbb{R} \longrightarrow \mathbb{R}$ ist die größte Teilmenge von $\mathbb{R}$, für die der gegebene Ausdruck eine endliche reelle Zahl liefert. Bestimmen Sie für jede der angegebenen Funktionen den (maximalen) Definitions– und Wertebereich, und prüfen Sie, wo f stetig bzw. rechts- oder linksseitig stetig ist. Stellen Sie die Funktionen graphisch dar:

$$\textbf{a)} f(x) = x^4 - 2 + \sqrt{2 - \frac{1}{x^4}} \qquad \textbf{b)} f(x) = \max\{z \in \mathbb{Z} \mid z \leq x\}$$

$$\textbf{c)} f(x) = \sqrt{\ln x} + \ln\sqrt{x} \qquad \textbf{d)} f(x) = \tan x \cdot \sin\frac{1}{x}.$$

65) Untersuchen Sie die folgenden Funktionenreihen auf punktweise Konvergenz und gleichmäßige Konvergenz:

$$\textbf{a)} \quad \sum_{n=0}^{\infty} x^n(1-x) \qquad \textbf{b)} \quad \sum_{n=1}^{\infty} \frac{\sin(nx)}{n^2}\,.$$

66) Bestimmen Sie den Konvergenzradius der folgenden Potenzreihen:

$$\text{a)} \quad \sum_{n=1}^{\infty} \frac{1}{(2n+5)^3}\, x^n \qquad \text{b)} \quad \sum_{n=1}^{\infty} (\ln n)^n x^n \qquad \text{c)} \quad \sum_{n=1}^{\infty} (2^n + 3^n)\, x^n$$

$$\text{d)} \quad \sum_{n=1}^{\infty} 3^n \sqrt{(3n-2)2^n}\, x^n \qquad \text{e)} \quad \sum_{n=1}^{\infty} \frac{(2+(-1)^n)^n}{n}\, x^n \qquad \text{f)} \quad \sum_{n=1}^{\infty} \frac{x^{5n+1}}{1+2^n}.$$

Für die Reihen e) und f) muss man sich (z.B. anhand des Vergleichskriteriums) überlegen, dass der Konvergenzradius allgemeiner gegeben ist durch $R^{-1} = \overline{\lim}_{n\to\infty} \sqrt[n]{|a_n|}$.

67) Bestimmen Sie von folgenden Potenzreihen den Konvergenzradius, die Menge A aller $x \in \mathbb{R}$ für die die Reihe absolut konvergiert und die Menge K aller $x \in \mathbb{R}$ für die die Reihe konvergiert

1. $\displaystyle\sum_{n=1}^{\infty} \frac{(-3)^n}{n+\sqrt{n}}\, x^n\,,$

2. $\displaystyle\sum_{n=1}^{\infty} (-1)^n \left(\frac{2}{3}\right)^n x^n\,,$

3. $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n^2 \cdot 4^n}\, x^n\,.$

3.5 Spezielle Funktionen

In diesem Abschnitt führen wir einige komplexe und relle Funktionen ein, die sehr häufig auftreten.

Exponentialfunktion und trigonometrische Funktionen

Wir hatten bereits für komplexe Zahlen z (und damit auch für reelle Zahlen) definiert

$$\begin{aligned}
e^z &:= \sum_{n=0}^{\infty} \frac{z^n}{n!} = 1 + z + \frac{z^2}{2!} + \frac{z^3}{3!} + \dots, \\
\sin z &:= \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} z^{2n+1} = z - \frac{z^3}{3!} + \frac{z^5}{5!} - \frac{z^7}{7!} \pm \dots, \\
\cos z &:= \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} z^{2n} = 1 - \frac{z^2}{2!} + \frac{z^4}{4!} - \frac{z^6}{6!} \pm \dots .
\end{aligned}$$

Es gilt

- $e^{z+w} = e^z e^w$ (bereits für das Cauchy-Produkt gezeigt),
- $\sin(-z) = -\sin z$, $\cos(z) = \cos(-z)$,
- **Eulersche Formel:**
$$\cos z + \mathrm{i} \sin z = \sum_{n=0}^{\infty} (-1)^n \frac{z^{2n}}{(2n)!} + \mathrm{i}(-1)^n \frac{z^{2n+1}}{(2n+1)!} = \sum_{k=0}^{\infty} \mathrm{i}^k \frac{z^k}{k!} = e^{\mathrm{i}\, z} \tag{3.36}$$
  Damit ist
$$e^{x+\mathrm{i}\, y} = e^x(\cos y + \mathrm{i} \sin y).$$
- **Formel von Moivre:**
$$(\cos z + \mathrm{i} \sin z)^n = e^{\mathrm{i}\, nz} = \cos(nz) + \mathrm{i} \sin(nz) \tag{3.37}$$
- $$\cos z = \frac{e^{\mathrm{i}\, z} + e^{-\mathrm{i}\, z}}{2}, \quad \sin z = \frac{e^{\mathrm{i}\, z} - e^{-\mathrm{i}\, z}}{2\,\mathrm{i}} \tag{3.38}$$
- $$\begin{aligned}
(\sin z)^2 + (\cos z)^2 &= \left(\frac{e^{\mathrm{i}\, z} - e^{-\mathrm{i}\, z}}{2\,\mathrm{i}}\right)^2 + \left(\frac{e^{\mathrm{i}\, z} + e^{-\mathrm{i}\, z}}{2}\right)^2 \\
&= \frac{e^{2\,\mathrm{i}\, z} - 2e^{\mathrm{i}\, z}e^{-\mathrm{i}\, z} + e^{-2\,\mathrm{i}\, z}}{-4} + \frac{e^{2\,\mathrm{i}\, z} + 2e^{\mathrm{i}\, z}e^{-\mathrm{i}\, z} + e^{-2\,\mathrm{i}\, z}}{4} \\
&= e^{\mathrm{i}\, z}e^{-\mathrm{i}\, z} = 1
\end{aligned}$$

- Aus den Formeln (3.38) können die Additionstheoreme hergeleitet werden:

$$\sin(z_1+z_2) = \sin z_1 \cos z_2 + \cos z_1 \sin z_2,$$
$$\cos(z_1+z_2) = \cos z_1 \cos z_2 - \sin z_1 \sin z_2.$$

- Aus (3.36) kann auch eine Exponentialdarstellung der komplexen Zahlen gewonnen werden. Es ist
$$z = x + \mathrm{i}\, y = r(\cos\phi + \mathrm{i}\sin\phi) \overset{(3.36)}{=} re^{\mathrm{i}\,\phi}.$$
Weiterhin gilt für die Potenzen von z:
$$z^n = r^n(\cos\phi + \mathrm{i}\sin\phi)^n = r^n e^{\mathrm{i}\, n\phi} = r^n(\cos(n\phi) + \mathrm{i}\sin(n\phi)).$$

Wir führen noch die trigonometrischen Funktionen **Tangens** und **Cotangens** ein:

$$\tan z := \frac{\sin z}{\cos z} \overset{(3.38)}{=} -\mathrm{i}\,\frac{e^{\mathrm{i}\,z} - e^{-\mathrm{i}\,z}}{e^{\mathrm{i}\,z} + e^{-\mathrm{i}\,z}}, \quad z \in \mathbb{C}, z \neq (\frac{1}{2} + k)\pi, \ \text{wobei } k \in \mathbb{Z} \text{ ist},$$
$$\cot : z = \frac{\cos z}{\sin z} \overset{(3.38)}{=} \mathrm{i}\,\frac{e^{\mathrm{i}\,z} + e^{-\mathrm{i}\,z}}{e^{\mathrm{i}\,z} - e^{-\mathrm{i}\,z}} \quad , \quad z \in \mathbb{C}, z \neq k\pi, \ k \in \mathbb{Z}.$$

Hyperbelfunktionen

Die Hyperbelfunktionen sind definiert als:

$$\begin{aligned}
\sinh z &= \frac{1}{2}\left(e^z - e^{-z}\right) = -\mathrm{i}\sin(\mathrm{i}\,z) && && \text{(sinus hyperbolicus)},\\
\cosh z &= \frac{1}{2}\left(e^z + e^{-z}\right) = \cos(\mathrm{i}\,z) && \text{für } z \in \mathbb{C} && \text{(cosinus hyperbolicus)},\\
\tanh z &= \frac{\sinh z}{\cosh z}, \quad \cosh z \neq 0 && && \text{(tangens hyperbolicus)},\\
\coth z &= \frac{\cosh z}{\sinh z}, \quad \sinh z \neq 0 && && \text{(cotangens hyperbolicus)},
\end{aligned}$$

wobei $z \in \mathbb{C}$ ist.Die Bezeichnung Hyperbelfunktion geht auf folgende Eigenschaft zurück:

$$(\cosh z)^2 - (\sinh z)^2 = 1.$$

Setzen wir im Reellen $a = \cosh\phi$, $b = \sinh\phi$, dann gilt die Hyperbelgleichung

$$a^2 - b^2 = 1.$$

Die Graphen der entsprechenden **reellen** Funktionen sind:

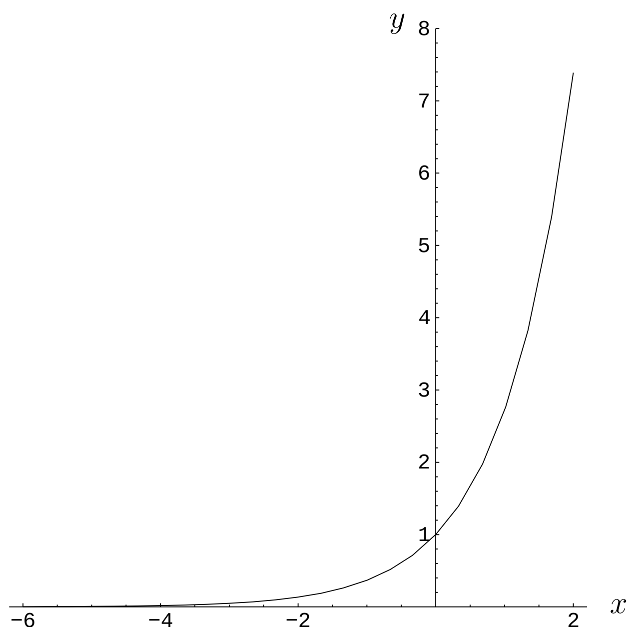


Abbildung 3.4: $y = e^x$

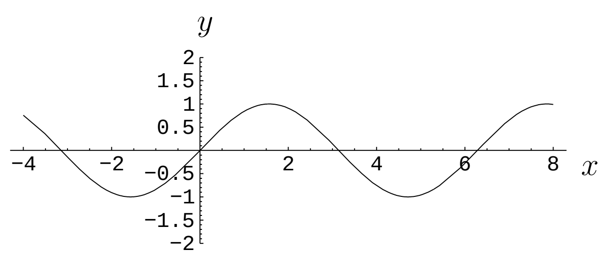


Abbildung 3.5: $y = \sin x$

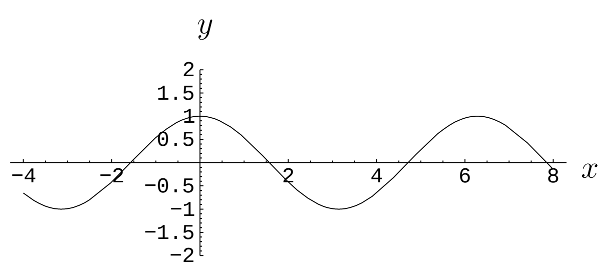


Abbildung 3.6: $y = \cos x$

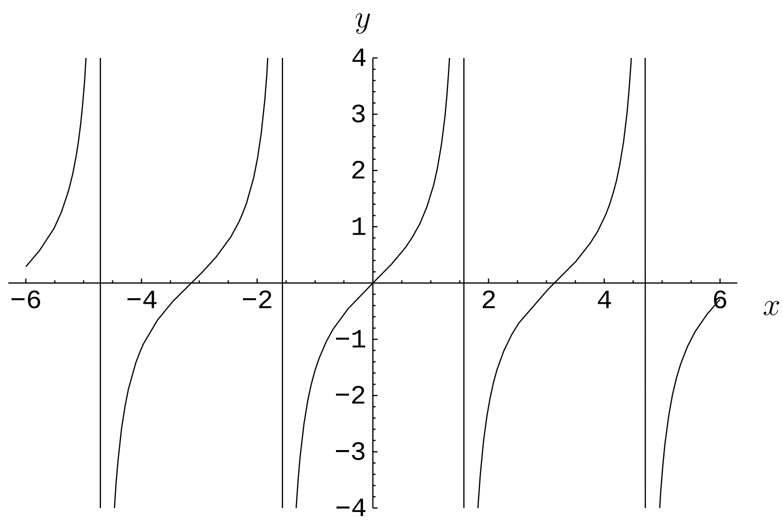


Abbildung 3.7: $y = \tan x$

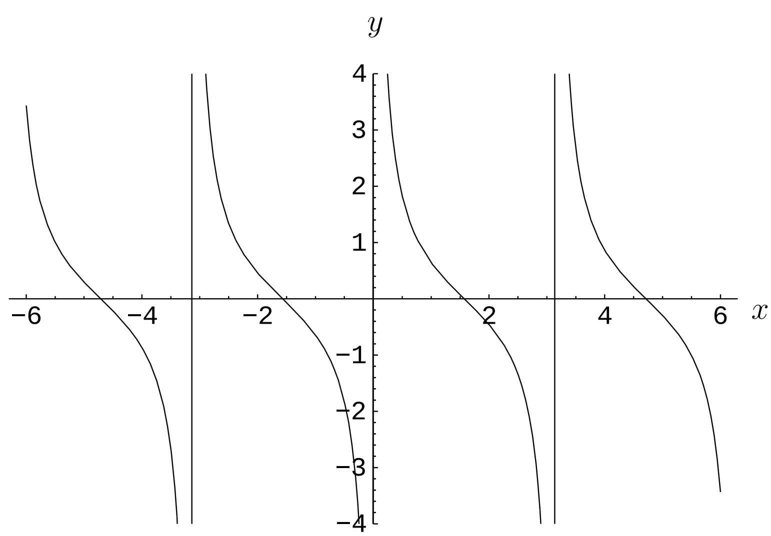


Abbildung 3.8: $y = \cot x$

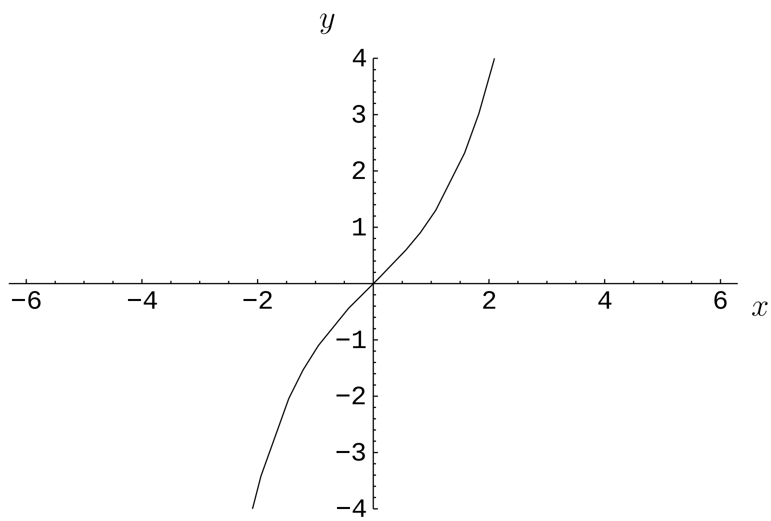


Abbildung 3.9: $y = \sinh x = \frac{1}{2}\left(e^x - e^{-x}\right)$

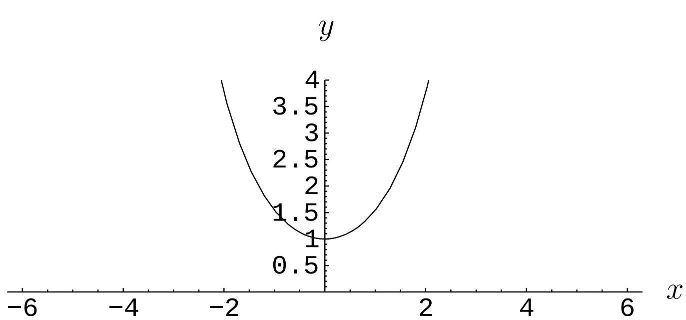


Abbildung 3.10: $y = \cosh x = \frac{1}{2}\left(e^x + e^{-x}\right)$

Monotone und inverse Funktionen

Wir hatten bereits definiert, was wir unter einer wachsenden bzw. fallenden Folge verstehen. Die Begriffe werden auf reelle Funktionen einer reellen Variablen übertragen.

Definition 3.53. Eine Funktion $f : D \to \mathbb{R}$ heißt monoton wachsend über dem Definitionsbereich D genau dann, wenn für alle $x_1, x_2 \in D$ mit $x_1 < x_2$ gilt, dass $f(x_1) \leq f(x_2)$ ist. Sie heißt streng monoton wachsend, wenn für alle $x_1, x_2 \in D$ mit $x_1 < x_2$ gilt, dass $f(x_1) < f(x_2)$ ist.
Entsprechend heißt $f : D \to \mathbb{R}$ monoton fallend, falls

$$f(x_1) \geq f(x_2) \quad \text{für } x_1 < x_2, \quad x_1, x_2 \in D$$

und streng monoton fallend, falls

$$f(x_1) > f(x_2) \quad \text{für } x_1 < x_2, \quad x_1, x_2 \in D$$

gilt. f heißt (streng) monoton, falls sie (streng) monoton wachsend bzw. fallend ist.

Beispiele

1° Die Funktion e^x ist streng monoton wachsend auf $\mathbb{R}$.

2° Die Funktion $\sin x$ ist im Intervall $D = \left[-\frac{\pi}{2}, +\frac{\pi}{2}\right]$ streng monoton wachsend und im Intervall $D = \left[\frac{\pi}{2}, +\frac{3\pi}{2}\right]$ streng monoton fallend.

Im 1. Kapitel hatten wir den Begriff der inversen Abbildung eingeführt. Wir formulieren hier den Begriff einer inversen Funktion (Umkehrfunktion):

Definition 3.54. Sei $f : D \to M$ eine injektive Abbildung, d.h. zu jedem $y \in f(D) \subseteq M$ existiert genau ein $x \in D$ mit $f(x) = y$. Dann heißt

$$\begin{aligned} f^{-1} : f(D) &\to D \\ y = f(x) &\to x \end{aligned}$$

die zu f inverse Abbildung und es gilt

$$f^{-1} \circ f(x) = f^{-1}(f(x)) = x \quad \forall\, x \in D, \tag{3.39}$$

$$f \circ f^{-1}(y) = f(f^{-1}(y)) = y \quad \forall\, y \in f(D). \tag{3.40}$$

Inverse Funktionen (Umkehrfunktionen) existieren für streng monotone Funktionen. Genauer wird diese Aussage durch folgenden Satz beschrieben:

Satz 3.55. [1, S.164] *Sei $f : [a, b] \to \mathbb{R}$ stetig und streng monoton wachsend oder fallend. Dann bildet f das Intervall $[a, b]$ bijektiv auf das Intervall $[f(a), f(b)]$ ab und die Umkehrfunktion $f^{-1} : [f(a), f(b)] \to [a, b]$ ist stetig.*

Im Beweis benutzt man den Zwischenwertsatz von Bolzano, der besagt: Sei $f : \mathbb{R} \supseteq [a, b] \to \mathbb{R}$ stetig. Dann nimmt f jeden Wert zwischen $f(a)$ und $f(b)$ wenigstens einmal an.

Bevor wir uns Beispiele von Umkehrfunktionen ansehen, die eine Rolle in den Ingenieurwissenschaften spielen, machen wir eine Bemerkung zu den Graphen.

Bemerkungen 3.56.

1. Der Graph von $y = f(x)$ stimmt mit dem Graphen von $x = f^{-1}(y)$ überein.
2. Den Graph von $y = f^{-1}(x)$ erhält man durch Spiegelung des Graphen von $y = f(x)$ an der Geraden $y = x$.

Beispiel

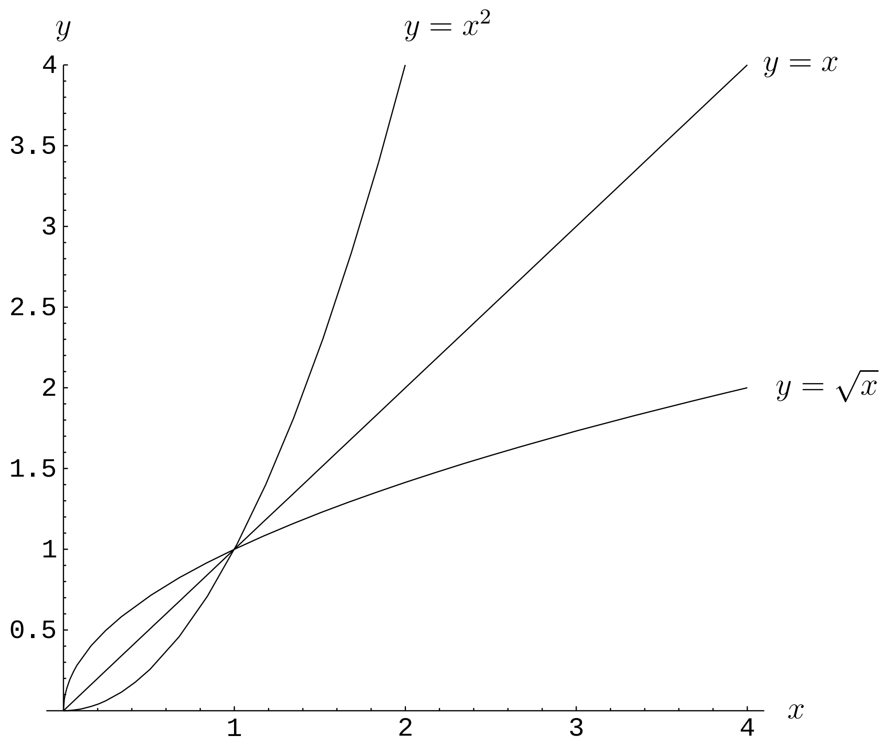


Abbildung 3.11: $y = x^2$ und die Umkehrfunktion $y = \sqrt{x}$

Umkehrfunktionen

Die streng monotone Exponentialfunktion

$$y = e^x = \exp(x), \quad -\infty < x < \infty$$

besitzt den natürlichen Logarithmus (kurz Logarithmus) als Umkehrfunktion:

$$y = \ln x, \quad 0 < x < \infty.$$

Insbesondere ist $\ln 1 = 0, \ln e = 1$ und

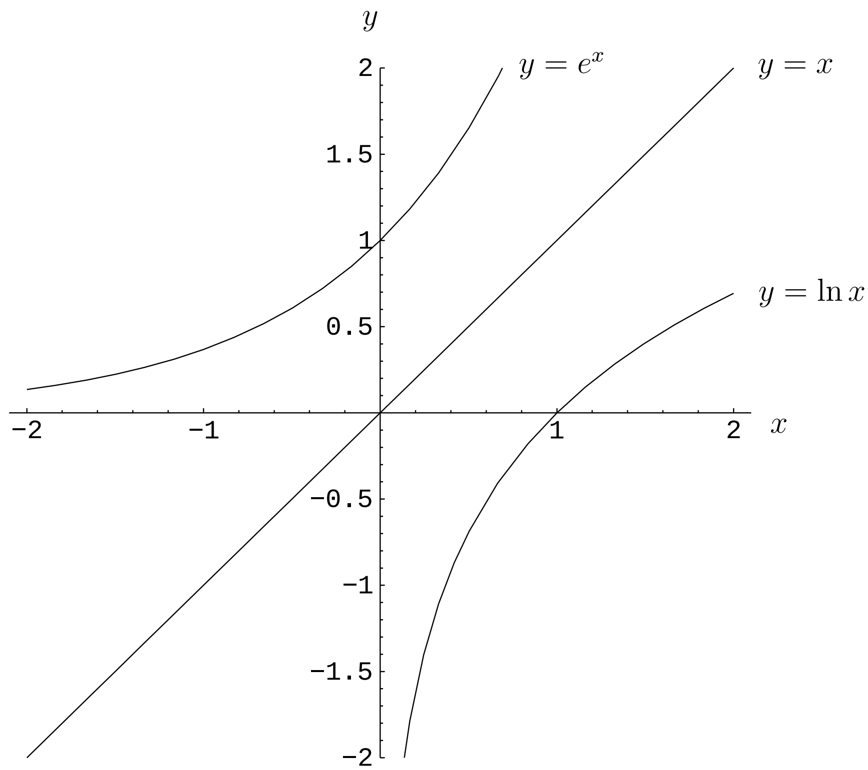


Abbildung 3.12: $y = e^x$ und die Umkehrfunktion $y = \ln(x)$

$$\lim_{x\to 0} \ln x = -\infty, \quad \lim_{x\to\infty} \ln x = \infty.$$

Da $uv = e^{\ln u} e^{\ln v} = e^{\ln u + \ln v}$ gilt, folgt

$$\begin{aligned}
\ln(uv) &= \ln u + \ln v, \\
\ln\left(\frac{u}{v}\right) &= \ln u - \ln v, \\
\ln\left(\frac{1}{x}\right) &= -\ln x, \\
\ln\left(x^n\right) &= n \ln x, \\
\ln\left(\sqrt[n]{x}\right) &= \frac{1}{n} \ln x.
\end{aligned}$$

Wir sind jetzt in der Lage, a^x für alle reelle Potenzen x zu definieren.

Definition 3.57. Sei a eine positive und x eine beliebige reelle Zahl, so versteht man unter der x-ten Potenz von a die reelle Zahl

$$a^x := e^{x \ln a}. \tag{3.41}$$

Folgerungen 3.58. Es ist

$$\begin{aligned}
a^0 &= e^{0 \ln a} = 1,\\
a^1 &= e^{1 \ln a} = a,\\
a^{\frac{m}{n}} &= e^{\frac{m}{n} \ln a} = e^{\ln\left(\sqrt[n]{a^m}\right)} = \sqrt[n]{a^m},\\
a^{u+v} &= e^{(u+v)\ln a} = e^{u \ln a} e^{v \ln a} = a^u a^v,\\
\ln\left(a^x\right) &= \ln\left(e^{x \ln a}\right) = x \ln a.
\end{aligned}$$

Trigonometrische Umkehrfunktionen (Zyklometrische Funktionen)

Da die trigonometrischen Funktionen nur auf Teilbereichen der reellen Achse monoton sind, müssen wir uns auf diese als Wahl von D beschränken. Wählt man die um $x = 0$ gelegten Monotoniebereiche aus, so spricht man von „Hauptwerten“ der zugehörigen Umkehrfunktionen. Wir geben hier eine Zusammenstellung der Hauptwerte an.

1. Die Umkehrfunktion von $y = \sin x$, $-\frac{\pi}{2} \leq x \leq \frac{\pi}{2}$ ist
$$x = \arcsin y, \quad -1 \leq y \leq 1 \quad \text{(arcussinus)}.$$

2. Die Umkehrfunktion von $y = \cos x$, $0 \leq x \leq \pi$ ist
$$x = \arccos y, \quad -1 \leq y \leq 1 \quad \text{(arcuscosinus)}.$$

3. Die Umkehrfunktion von $y = \tan x$, $-\frac{\pi}{2} < x < \frac{\pi}{2}$ ist
$$x = \arctan y, \quad -\infty < y < \infty \quad \text{arcustangens)}.$$

4. Die Umkehrfunktion von $y = \cot x$, $0 < x < \pi$ ist
$$x = \operatorname{arccot} y \quad -\infty < y < \infty \quad \text{(arcuscotangens)}.$$

5. Die Umkehrfunktion von $y = \sinh x$, $-\infty < x < \infty$ ist
$$x = \operatorname{Arsinh} y \quad -\infty < y < \infty \quad \text{(Areasinus)}.$$

6. Die Umkehrfunktion von $y = \cosh x$, $0 \leq x < \infty$ ist
$$x = \operatorname{Arcosh} y \quad 1 \leq y < \infty \quad \text{(Areacosinus)}.$$

7. Die Umkehrfunktion von $y = a^x$, $a > 0, a \neq 1, x \in \mathbb{R}$ ist
$$x = \frac{\ln y}{\ln a}.$$

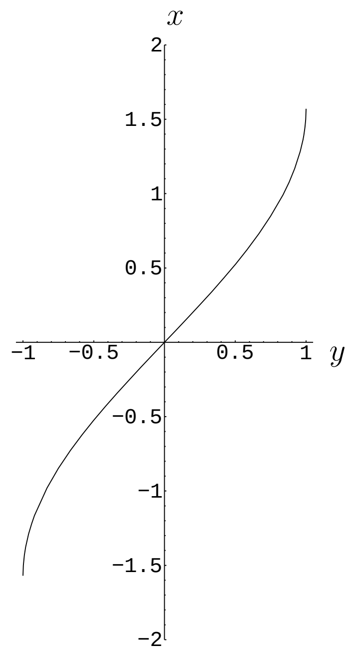


Abbildung 3.13: Die Umkehrfunktion (Hauptwert) $x = \arcsin y$ von $y = \sin x$

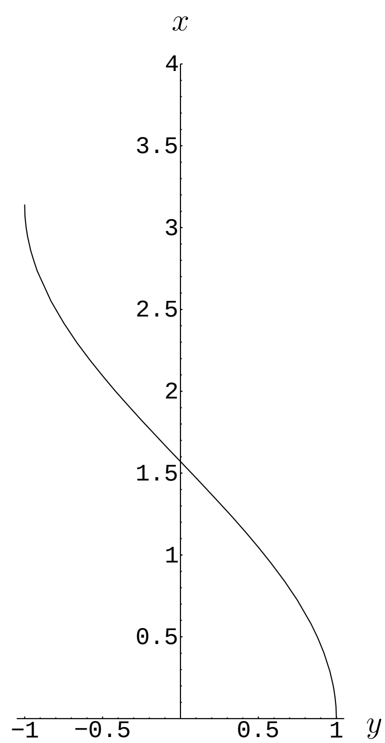


Abbildung 3.14: Die Umkehrfunktion (Hauptwert) $x = \arccos y$ von $y = \cos x$

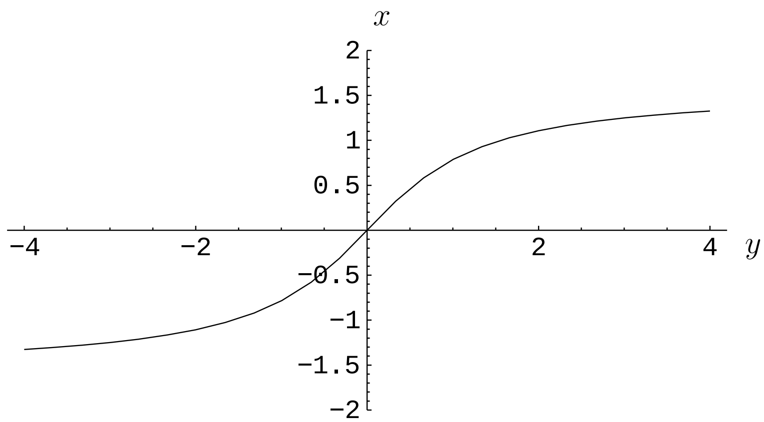


Abbildung 3.15: Die Umkehrfunktion (Hauptwert) $x = \arctan y$ von $y = \tan x$

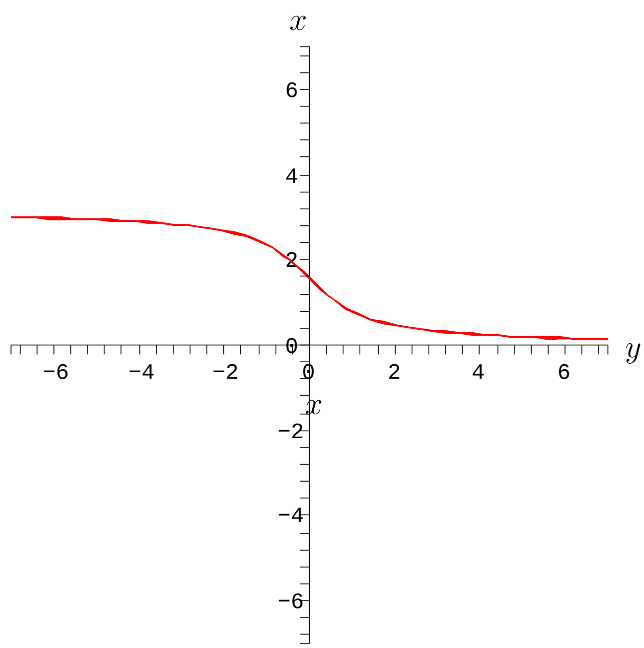


Abbildung 3.16: Die Umkehrfunktion $x = \operatorname{arccot} y$ von $y = \cot x$

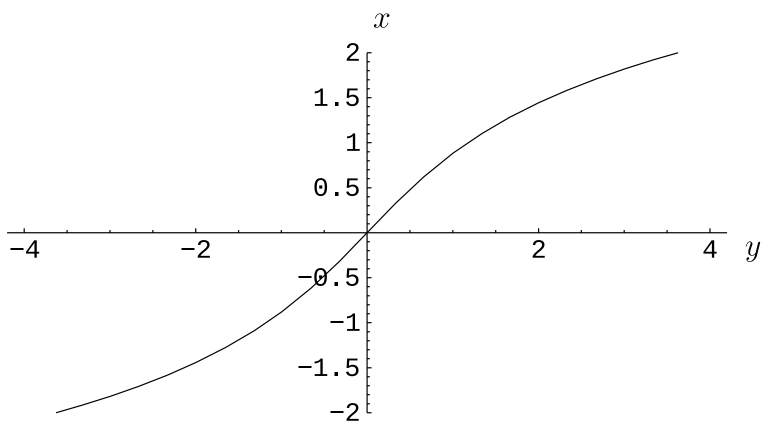


Abbildung 3.17: Die Umkehrfunktion $x = \operatorname{Arsinh} y$ von $y = \sinh x$

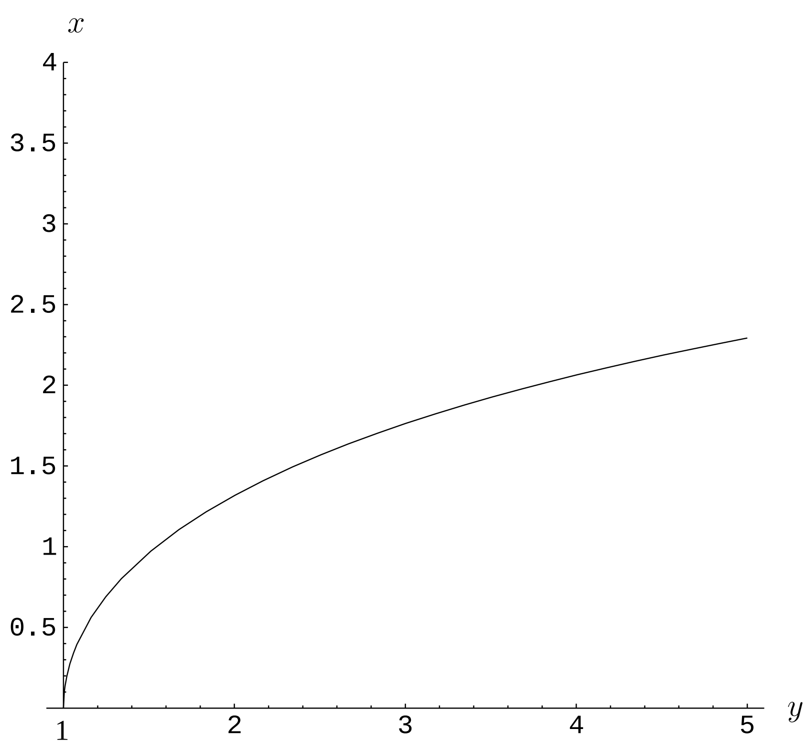


Abbildung 3.18: Die Umkehrfunktion $x = \operatorname{Arcosh} y$ von $y = \cosh x$, $0 \leq x < \infty$

Übungsaufgaben

68) Betrachten Sie die Funktionen $\cos : \mathbb{C} \to \mathbb{C}$ und $\sin : \mathbb{C} \to \mathbb{C}$ in der komplexen Ebene. Zeigen Sie, dass diese nur reelle Nullstellen haben (also $\cos(x) \neq 0 \,\forall\, x \in \mathbb{C} \setminus \mathbb{R}$ und $\sin(x) \neq 0 \,\forall\, x \in \mathbb{C} \setminus \mathbb{R}$) und bestimmen Sie die Nullstellen.

69) Leiten Sie die folgenden Additionstheoreme her:

1. $\cos^2 \dfrac{z}{2} = \dfrac{1}{2}(1 + \cos z)$,
2. $\sin^2 \dfrac{z}{2} = \dfrac{1}{2}(1 - \cos z)$,
3. $\cos a \pm \cos b = \pm 2, \begin{matrix}\cos \\ \sin\end{matrix} \left(\dfrac{a+b}{2}\right) \begin{matrix}\cos \\ \sin\end{matrix} \left(\dfrac{a-b}{2}\right)$,
4. $\sin a + \sin b = 2 \cos \left(\dfrac{a \mp b}{2}\right) \sin \left(\dfrac{a \pm b}{2}\right)$.

Literaturverzeichnis

[1] Beyer, Gottwald, Günther, and Wünsch. *Grundkurs Analysis*, vol. Teil 1. Teubner Verlag, Leipzig, 1972.

[2] Brieskorn, E. *Lineare Algebra und analytische Geometrie*. Vieweg Verlag, 1983.

[3] Brill, M. *Mathematik für Informatiker*. Hanser-Verlag, 2001.

[4] Fischer, G. *Lineare Algebra*. Vieweg Verlag, 2002.

[5] Hartmann, P. *Mathematik für Informatiker*. Vieweg, 2002.

[6] K.Burg, H.Haf, F. *Höhere Mathematik für Ingenieure*, vol. I und II. Teubner Verlag, 2001,2002.

[7] Mayberg, K., and Vachenauer, P. *Höhere Mathematik I*. Springer-Verlag, 2001.

[8] Pareigis, B. *Lineare Algebra für Informatiker*. Springer-Verlag, 2000.

Erschienene Preprints ab Nummer 2004/001

Komplette Liste: `http://preprints.ians.uni-stuttgart.de`

2004/001 *Geis, W., Mishuris, G., Sändig, A.-M.*: 3D and 2D asymptotic models for piezoelectric stack actuators with thin metal inclusions

2004/002 *Klimke, A., Wohlmuth, B., Willner, K.*: Computing expensive multivariate functions of fuzzy numbers using sparse grids

2004/003 *Klimke, A., Wohlmuth, B., Willner, K.*: Uncertainty modeling using efficient fuzzy arithmetic based on sparse grids: applications to dynamic systems

2004/004 *Flemisch, B., Mair, M., Wohlmuth, B.*: Nonconforming discretization techniques for overlapping domain decompositions

2004/005 *Sändig, A.-M.*: Vorlesung Mathematik für Informatiker und Softwaretechniker I, WS 2003/2004