Probabilistische Datenwissenschaft für die Psychologie

Dirk Ostwald

2024-01-25

Inhaltsverzeichnis

Willkommen

Vorwort

Teil I.

Mathematische Grundlagen

1. Sprache und Logik

1.1. Mathematik ist eine Sprache

Mathematik ist die Sprache der naturwissenschaftlichen Modellbildung. So entspricht zum Beispiel der Ausdruck

$$F = ma \tag{1.1}$$

im Sinne des zweiten Newtonschen Axioms einer Theorie zur Bewegung von Objekten unter der Einwirkung von Kräften (Newton (1687)). Gleichermaßen entspricht der Ausdruck

$$\max_{q(z)} \int q(z) \ln \left(\frac{p(y,z)}{q(z)} \right) dz \tag{1.2}$$

im Sinne der Variational Inference der zeitgenössischen Theorie zur Funktionsweise des Gehirns (Friston (2005), Friston et al. (2023), Ostwald et al. (2014), Blei et al. (2017)). Mathematische Symbolik dient dabei insbesondere der genauen Kommunikation wissenschaftlicher Erkenntnisse und zielt darauf ab, komplexe Sachverhalte exakt und effizient zu beschreiben. Wie beim reflektierten Umgang mit jeder Form von Sprache steht also die Frage "Was soll das heißen?" als Leitfrage im Umgang mit mathematischen Inhalten und Symbolismen immer im Vordergrund.

Als Sprachgebäude weist die Mathematik einige Besonderheiten auf. Zum einen sind ihre Inhalte oft abstrakt. Dies rührt daher, dass sich die Mathematik um eine möglichst breite Allgemeinverständlichkeit und Anwendbarkeit bemüht. Mathematische Zugänge zu den Phänomenen der Welt sind dabei an einer möglichst einfache Transferierbarkeit von Erkenntnissen in andere Kontexte interessiert. Um dies zu ermöglichen, versucht die Mathematik möglichst genau und verständlich, also im Sinne präziser Begriffsbildungen zu arbeiten. Sie geht dabei insbesondere streng hierarchisch vor, so dass an späterer Stelle eingeführte Begrifflichkeiten oft ein gutes Verständnis der ihnen zugrundeliegenden und an früherer Stelle eingeführten Begrifflichkeiten voraussetzen.

Die Genauigkeit der mathematischen Sprache impliziert dabei eine hohe Informationsdichte. Sie ist daher eher nüchtern und lässt überflüssiges weg, so dass in mathematischen Texten im besten Fall *alles* für die Kommunikation einer Idee relevant ist. Als Rezipient:in mathematischer Texte nimmt man die Informationsdichte mathematischer Texte anhand des hohen Verbrauchs an kognitiver Energie beim Lesen eines Textes wahr. Dieser hohe Energieverbrauch gebietet insbesondere Ruhe und Langsamkeit bei einem auf ein gutes Verständnis abzielenden Lesen. Als Leitsatz im Umgang mit mathematischen Texten mag dabei folgendes Zitat dienen: "Einen mathematischen Text kann man nicht lesen wie einen Roman, man muss ihn sich erarbeiten" (Unger (2000)). Nach dem Lesen eines kurzen mathematischen Textes sollte man sich immer kritisch fragen, ob man das Gelesene wirklich verstanden hat oder ob man zur Klärung des Sachverhaltes weitere Quellen heranziehen sollte. Auch ist es hilfreich, sich im Sinne des berühmten Zitats

"What I cannot create, I do not understand" von Richard Feynman eigene Aufzeichnungen anzufertigen und mathematische Sprachgebäude selbst nachzubauen.

Möchte man sich also die Welt der naturwissenschaftliche Modellbildung erschließen, so ist es hilfreich, beim Umgang mit ihrer mathematischen Ausdrucksweise und Symbolik die gleichen Strategien wie beim Erlernen einer Fremdsprache anzuwenden. Hierzu gehört neben dem Eintauchen in den entsprechenden Sprachraum, also der ständige Exposition mit mathematischen Ausdrucksweisen, sicherlich auch zunächst einmal das Auswendiglernen von Begriffen und das aktive Lesen und das Übersetzen von Texten in die Alltagssprache. Ein tiefes und sicheres Verständnis mathematischer Modellbildung ergibt sich dann insbesondere durch die Anwendung mathematischer Herangehensweisen in schriftlicher und mündlicher Form.

1.2. Grundbausteine mathematischer Kommunikation

In diesem Abschnitt stellen wir mit den Begriffen der *Definition*, des *Theorems* und des *Beweises* drei Grundbausteine mathematischer Kommunikation vor, die uns durchgängig begleiten.

Definition

Eine *Definition* ist eine Grundannahme eines mathematischen Systems, die innerhalb dieses Systems weder begründet noch deduktiv abgeleitet wird. Definitionen können nur nach ihrer Nützlichkeit innerhalb eines mathematischen Systems bewertet werden. Eine Definition lernt man am besten erst einmal auswendig und hinterfragt sie erst dann, wenn man ihren Nutzen in der Anwendung verstanden hat oder von diesem nicht überzeugt ist. Etwas Entspannung und Ruhe beim Umgang mit auf den ersten Blick komplexen Definitionen ist generell hilfreich. Um zu kennzeichnen, dass wir ein Symbol als etwas definieren, nutzen wir die Schreibweise ":=". Zum Beispiel definiert der Ausdruck "$a := 2$" das Symbol a als die Zahl Zwei. Definitionen enden in diesem Text immer mit dem Symbol $\bullet$.

Theorem

Ein *Theorem* ist eine mathematische Aussage, die mittels eines Beweises als wahr (richtig) erkannt werden kann. Dass heißt, ein Theorem wird immer aus Definitionen und/oder anderen Theoremen hergeleitet. Theoreme sind in diesem Sinne die empirischen Ergebnisse der Mathematik. Im Deutschen werden Theoreme auch oft als *Sätze* bezeichnet. In der angewandten, datenanalytischen Mathematik sind Theoreme oft für Berechnungen hilfreich. Es lohnt sich also, sie auswendig zu lernen, da sie meist die Grundlage für Datenauswertung und Dateninterpretation bilden. Oft tauchen in Theoremen Gleichungen auf. Diese ergeben sich dabei aus den Voraussetzungen des Theorems. Um Gleichungen zu kennzeichnen nutzen wir das Gleichheitszeichen "=". So besagt also zum Beispiel der Ausdruck "$a = 2$" in einem gegebenen Kontext, dass aufgrund bestimmter Voraussetzungen das Symbol oder die Variable a den Wert zwei hat. Theoreme enden in diesem Text immer mit dem Symbol $\circ$.

Beweis

Ein *Beweis* ist eine logische Argumentationskette, die auf bekannte Definitionen und Theoreme zurückgreift, um die Wahrheit (Richtigkeit) eines Theorems zu belegen. Kurze

Beweise tragen dabei oft zum Verständnis eines Theorems bei, lange Beweise eher nicht. Beweise sind also insbesondere die Antwort auf die Frage, warum eine mathematische Aussage gilt ("Warum ist das so?"). Beweise lernt man nicht auswendig. Wenn Beweise kurz sind, ist es sinnvoll, sie durchzuarbeiten, da sie meist als bekannt vorausgesetzte Inhalte wiederholen. Wenn sie lang sind, ist es sinnvoller sie zunächst zu übergehen, um sich nicht in Details zu verlieren und vom eigentlichen Weg durch das entsprechende mathematische Gebäude abzukommen. Beweise enden in diesem Text immer mit dem Symbol $\square$.

Neben den oben vorgestellten Begriffen gibt es mit *Axiomen*, *Lemmata*, *Korollaren* und *Vermutungen* noch weitere typische Grundbausteine mathematischer Texte. Wir werden diese Begriff nicht verwenden und geben deshalb für sie nur einen kurzen Überblick.

Axiome sind unbeweisbare Theoreme, in dem Sinne, als dass sie als Grundannahmen zum Aufbau mathematischer Systeme dienen. Der Übergang zwischen Definitionen und Axiomen ist dabei oft fließend. Da wir mathematisch nicht besonders tief arbeiten, bevorzugen wir in den allermeisten Fällen den Begriff der Definition.

Ein *Lemma* ist ein "Hilfstheorem", also eine mathematische Aussage, die zwar bewiesen wird, aber nicht so bedeutend ist wie ein Theorem. Da wir einerseits auf bedeutende Inhalte fokussieren und andererseits mathematische Aussagen nicht diskriminieren wollen, verzichten wir auf diesen Begriff und nutzen stattdessen den Begriff des Theorems.

Ein *Korollar* ist eine mathematische Aussage, die sich durch einen einfachen Beweis aus einem Theorem ergibt. Da die "Einfachheit" mathematischer Beweise eine relative Eigenschaft ist, verzichten wir auf diesen Begriff und nutzen stattdessen auch hier den Begriff des Theorems.

Vermutungen sind mathematische Aussagen von denen unbekannt ist, ob sie beweisbar oder widerlegbar sind. Da wir im Bereich der angewandten Mathematik arbeiten, treffen wir nicht auf Vermutungen.

1.3. Aussagenlogik

Nachdem wir nun einige Grundbausteine mathematischer Modellbildung kennengelernt haben, wollen wir uns mit der *Aussagenlogik* einem einfachem System nähern, das es erlaubt, Beziehungen zwischen mathematischen Aussagen herzustellen und zu formalisieren. Im Folgenden spielt die Aussagenlogik zum Beispiel in der Definition von Mengenoperationen, bei Optimierungsbedingungen von Funktionen und in vielen Beweisen einen tragende Rolle. In der mathematischen Anwendung ist Aussagenlogik die Grundlage der Booleschen Logik der Programmierung. In der mathematischen Psychologie ist die Aussagenlogik zum Beispiel die Grundlage der Repräsentationstheorie des Messens.

Wir beginnen mit der Definition des Begriffs der mathematischen *Aussage*.

Definition 1.1 (Aussage)**.** Eine *Aussage* ist ein Satz, dem eindeutig die Eigenschaft *wahr* oder *falsch* zugeordnet werden kann.

•

Das Adjektiv *wahr* kann auch als *richtig* verstanden werden. Wir kürzen wahr mit "w" und falsch mit "f" ab. Im Körper der reellen Zahlen ist zum Beispiel die Aussage $1+1=2$ wahr und die Aussage $1+1=3$ falsch. Man beachte, dass die Binärität des Wahrheitsgehalts von Aussagen eine Grundannahme der Aussagenlogik und damit formal wissenschaftlich und nicht empirisch zu verstehen ist. Wahrheitsgehalte beziehen sich nicht auf Definitionen, Definitionen sind immer wahr. Eine erste Möglichkeit, mit Aussagen zu arbeiten, ist, sie zu negieren. Dies führt auf folgende Definition.

Definition 1.2 (Negation)**.** A sei eine Aussage. Dann ist die *Negation von A* die Aussage, die falsch ist, wenn A wahr ist und die wahr ist, wenn A falsch ist. Die Negation von A wird mit $\neg A$, gesprochen als "nicht A", bezeichnet.

•

Beispielsweise ist die Negation der Aussage "Die Sonne scheint" die Aussage "Die Sonne scheint nicht". Die Negation der Aussage $1+1=2$ ist die Aussage $1+1\neq 2$ und die Negation der Aussage $x>1$ ist die Aussage $x\leq 1$. Tabellarisch stellt man die Definition der Negation einer Aussage A wie folgt dar.

A	$\neg A$
w	f
f	w

Tabellen dieser Form nennt man *Wahrheitstafeln.* Sie sind ein beliebtes Hilfsmittel in der Aussagenlogik. Möchte man zwei Aussagen logisch verbinden, so bieten sich zunächst die Begriffe der *Konjunktion* und *Disjunktion* an.

Definition 1.3 (Konjunktion)**.** A und B seien Aussagen. Dann ist die *Konjunktion von A und B* die Aussage, die dann und nur dann wahr ist, wenn A und B beide wahr sind. Die Konjunktion von A und B wird mit $A\wedge B$, gesprochen als "A und B", bezeichnet.

•

Die Definition der Konjunktion impliziert folgende Wahrheitstafel.

A	B	$A\wedge B$
w	w	w
w	f	f
f	w	f
f	f	f

Als Beispiel sei A die Aussage $2\geq 1$ und B die Aussage $2>1$. Da sowohl A und B wahr sind, ist auch die Aussage $2\geq 1\wedge 2>1$ wahr. Als weiteres Beispiel sei A die Aussage $1\geq 1$ und B die Aussage $1>1$. Hier ist nun A wahr und B falsch. Also ist die Aussage $1\geq 1\wedge 1>1$ falsch.

Definition 1.4 (Disjunktion)**.** A und B seien Aussagen. Dann ist die *Disjunktion von A und B* die Aussage, die dann und nur dann wahr ist, wenn mindestens eine der beiden Aussagen A und B wahr ist. Die Disjunktion von A und B wird mit $A \vee B$, gesprochen als "A oder B", bezeichnet.

•

Die Definition der Disjunktion impliziert folgende Wahrheitstafel

A	B	$A \vee B$
w	w	w
w	f	w
f	w	w
f	f	f

$A \vee B$ ist also insbesondere auch dann wahr, wenn A und B beide wahr sind. Damit ist das hier betrachtete "oder" genauer ein "und/oder". Man nennt die Disjunktion daher auch ein "nicht-exklusives oder". Als Beispiel sei A die Aussage $2 \geq 1$ und B die Aussage $2 > 1$. A ist wahr und B ist wahr. Also ist die Aussage $2 \geq 1 \vee 2 > 1$ wahr. Sei nun wiederrum A die Aussage $1 \geq 1$ wahr und B die Aussage $1 > 1$. Dann ist A wahr und B falsch. Also ist die Aussage $1 \geq 1 \vee 1 > 1$ wahr.

Eine Möglichkeit, Aussagen in einen mechanischen logischen Zusammenhang zu stellen, ist die *Implikation*. Diese ist wie folgt definiert.

Definition 1.5 (Implikation)**.** A und B seien Aussagen. Dann ist die *Implikation*, bezeichnet mit $A \Rightarrow B$, die Aussage, die dann und nur dann falsch ist, wenn A wahr und B falsch ist. A heißt dabei die *Voraussetzung (Prämisse)* und B der *Schluss (Konklusion)* der Implikation. $A \Rightarrow B$ spricht man als "aus A folgt B", "A impliziert B", oder "wenn A, dann B".

•

Man mag $\Rightarrow$ auch als "daraus folgt" lesen. Die Definition der Implikation impliziert folgende Wahrheitstafel.

A	B	$A \Rightarrow B$
w	w	w
w	f	f
f	w	w
f	f	w

Ein Verständnis der Definition der Implikation im Sinne obiger Wahrheitstafel ergibt sich am ehesten, indem man sie als Versuch liest, die intuitive Vorstellung einer Folgerung im Kontext der Aussagenlogik abzubilden und zu formalisieren. Betrachtet man obige Wahrheitstafel unter diesem Gesichtspunkt, so sieht man, dass wenn A wahr ist und $A \Rightarrow B$ wahr ist, B wahr ist. Konstruiert man basierend auf einer wahren Aussage also (zum Beispiel durch das Umformen von Gleichungen) eine wahre Implikation so folgt, dass auch B wahr ist. Ist dies nicht möglich (dass also gilt, wenn A wahr ist, dass

$A \Rightarrow B$ immer falsch ist), dann ist auch B falsch. So mag man Aussagen widerlegen. Schließlich sieht man, dass wenn A falsch ist und $A \Rightarrow B$ wahr ist, B wahr oder falsch sein kann. Aus einer wahren Voraussetzung folgt also nur bei wahrer Implikation eine wahre Konklusion. Insbesondere genügt die Definition der Implikation damit der Forderung “Aus Falschem folgt beliebiges (ex falso sequitur quodlibet)”. Aus falschen Aussagen kann man also mithilfe der Implikation nichts richtiges folgern.

Im Kontext der Implikation ergeben sich die Begriffe der *hinreichenden* und der *notwendigen Aussagen (Bedingungen)*. Diese sind definiert wie folgt: wenn $A \Rightarrow B$ wahr ist, sagt man “A ist *hinreichend* für B” und “B ist *notwendig* für A”. Diese Sprachregelung erklärt sich folgendermaßen. Wenn $A \Rightarrow B$ wahr ist, gilt dass, wenn A wahr ist auch B wahr ist. Die Wahrheit von A reicht also für die Wahrheit von B aus. A ist also hinreichend (ausreichend) für B. Weiterhin gilt, dass wenn $A \Rightarrow B$ wahr ist, dass wenn B falsch ist, dann auch A falsch ist. Die Wahrheit von B ist also für die Wahrheit von A notwendig.

Eine sehr häufig autretender Zusammenhang zwischen zwei Aussagen ist ihre *Äquivalenz*.

Definition 1.6 (Äquivalenz)**.** A und B seien Aussagen. Die *Äquivalenz von A und B* ist die Aussage, die dann und nur dann wahr ist,wenn A und B beide wahr sind oder wenn A und B beide falsch sind. Die Äquivalenz von A und B wird mit $A \Leftrightarrow B$ bezeichnet und gesprochen als “A genau dann wenn B” oder “A ist äquivalent zu B”.

•

Die Definition der Äquivalenz impliziert folgende Wahrheitstafel

A	B	$A \Leftrightarrow B$
w	w	w
w	f	f
f	w	f
f	f	w

Die Definition des Begriffes der *logischen Äquivalenz* erlaubt es unter anderem, die Äquivalenz zweier Aussagen mithilfe von Implikationen nachzuweisen.

Definition 1.7 (Logische Äquivalenz)**.** Zwei Aussagen heißen *logisch äquivalent*, wenn ihre Wahrheitstafeln gleich sind.

•

Als Beispiele für logische Äquivalenzen, die häufig in Beweisargumentationen genutzt werden, zeigen wir folgendes Theorem.

Theorem 1.1 (Logische Äquivalenzen)**.**

A und B seien zwei Aussagen. Dann sind folgende Aussagen logisch äquivalent

(1) $A \Leftrightarrow B$ *und* $(A \Rightarrow B) \wedge (B \Rightarrow A)$

(2) $A \Rightarrow B$ *und* $(\neg B) \Rightarrow (\neg A)$

∘

Beweis. Nach Definition des Begriffs der logischen Äquivalenz müssen wir zeigen, dass die Wahrheitstafeln der betrachteten Aussagen gleich sind. Wir zeigen erst (1), dann (2).

(1) Wir erinnern an die Wahrheitstafel von $A \Leftrightarrow B$:

A	B	$A \Leftrightarrow B$
w	w	w
w	f	f
f	w	f
f	f	w

Wir betrachten weiterhin die Wahrheitstafel von $(A \Rightarrow B) \wedge (B \Rightarrow A)$:

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

Der Vergleich der Wahrheitstafel von $A \Leftrightarrow$ mit den ersten beiden und der letzten Spalte der Wahrheitstafel von $(A \Rightarrow B) \wedge (B \Rightarrow A)$ zeigt ihre Gleichheit.

(2) Wir erinnern an die Wahrheitstafel von $A \Rightarrow B$:

A	B	$A \Rightarrow B$
w	w	w
w	f	f
f	w	w
f	f	w

Wir betrachten weiterhin die Wahrheitstafel von $(\neg B) \Rightarrow (\neg A)$:

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

Der Vergleich der Wahrheitstafel von $A \Rightarrow B$ mit den ersten beiden und der letzten Spalte der Wahrheitstafel von $(\neg B) \Rightarrow (\neg A)$ zeigt ihre Gleichheit.

□

Die erste Aussage von Theorem 1.1 besagt, dass die Aussage “A und”B sind äquivalent” logisch äquivalent zur Aussage “Aus A folgt B” und aus “B folgt A” ist. Dies ist die Grundlage für viele sogenannte *direkte Beweise* mithilfe von Äquivalenzumformungen. Die zweite Aussage von Theorem 1.1 besagt, dass die Aussage “Aus A folgt”B” logisch äquivalent zur Aussage “Aus nicht B folgt nicht A” ist. Dies ist die Grundlage für die Technik des *indirekten Beweises*.

1.4. Beweistechniken

Im letzten Abschnitt wollen wir mit den Begriffen der *direkten* und *indirekten Beweise* sowie des *Beweises durch Widerspruch* kurz drei Beweistechniken skizzieren, von denen vor allem die erste in diesem Text immer wieder zur Begründung von Theoremen herangezogen wird. Dabei haben typische Theoreme die Form $A \Rightarrow B$ für Aussagen A und B.

Es gilt dabei

- *Direkte Beweise* nutzen Äquivalenzumformungen, um $A \Rightarrow B$ zu zeigen.
- *Indirekte Beweise* nutzen die logische Äquivalenz von $A \Rightarrow B$ und $(\neg B) \Rightarrow (\neg A)$.
- *Beweise durch Widerspruch* zeigen, dass $(\neg B) \wedge A$ falsch ist.

Um diese Techniken an einem Beispiel zu erläutern, erinnern wir kurz an folgende *Äquivalenzumformungen von Gleichungen*:

- Addition oder Subtraktion einer Zahl auf beiden Seiten der Gleichung, zum Beispiel

$$2x + 4 = 10 \Leftrightarrow 2x = 6, \tag{1.3}$$

- Multiplikation mit einer oder Division durch eine von Null verschiedene Zahl auf beiden Seiten der Gleichung, zum Beispiel

$$2x = 6 \Leftrightarrow x = 3, \tag{1.4}$$

- Anwendung einer injektiven Funktion auf beiden Seiten der Gleichung, zum Beispiel

$$\exp(x) = 2 \Leftrightarrow x = \ln(2), \tag{1.5}$$

sowie an folgende elementaren *Äquivalenzumformungen von Ungleichungen*:

- Addition oder Subtraktion einer Zahl auf beiden Seiten der Ungleichung, zum Beispiel

$$-2x + 4 \geq 10 \Leftrightarrow -2x \geq 6, \tag{1.6}$$

- Multiplikation mit einer Zahl oder Division durch eine von Null verschiedene Zahl auf beiden Seiten der Ungleichung, wobei die Multiplikation oder Division mit einer negativen Zahl die Umkehrung der Ungleichung impliziert, zum Beispiel

$$-2x \geq 6 \Leftrightarrow x \leq -3, \tag{1.7}$$

- Anwendung monotoner Funktionen auf beiden Seiten der Ungleichung

$$\exp(x) \geq 2 \Leftrightarrow x \geq \ln(2). \tag{1.8}$$

Damit ausgestattet wollen wir nun folgendes Theorem mithilfe eines direkten Beweises, eines indirekten Beweises und eines Beweises durch Widerspruch beweisen (vgl. Arens et al. (2018)).

Theorem 1.2 (Quadrate positiver Zahlen)**.** *Es seien a und b zwei positive Zahlen. Dann gilt $a^2 < b^2 \Rightarrow a < b$.*

$\circ$

Beweis. Wir geben zunächst einen *direkten Beweis*. Dazu sei $a^2 < b^2$ die Aussage A und $a < b$ die Aussage B. Dann gilt

$$a^2 < b^2 \Leftrightarrow 0 < b^2 - a^2 \Leftrightarrow 0 < (b+a)(b-a) \Leftrightarrow 0 < (b-a) \Leftrightarrow a < b. \tag{1.9}$$

Wir geben nun einen *indirekten Beweis*. Es sei $a^2 \geq b^2$ die Aussage $\neg A$. Weiterhin sei $a \geq b$ die Aussage $\neg B$. Dann gilt

$$a \geq b \Leftrightarrow a^2 \geq ab \wedge ab \geq b^2 \Leftrightarrow a^2 \geq b^2. \tag{1.10}$$

Schließlich geben wir einen *Beweis durch Widerspruch*. Wir zeigen, dazu, dass die Annahme $(\neg B) \wedge A$ auf eine falsche Aussage führt. Es gilt

$$a \geq b \wedge a^2 < b^2 \Leftrightarrow a^2 \geq ab \wedge a^2 < b^2 \Leftrightarrow ab \leq a^2 < b^2. \tag{1.11}$$

Weiterhin gilt

$$a \geq b \wedge a^2 < b^2 \Leftrightarrow ab \geq b^2 \wedge a^2 < b^2 \Leftrightarrow a^2 < b^2 \leq ab. \tag{1.12}$$

Insgesamt gilt dann also die falsche Aussage

$$ab \leq a^2 < b^2 \leq ab \Leftrightarrow ab < ab. \tag{1.13}$$

□

1.5. Selbstkontrollfragen

1. Erläutern Sie die Besonderheiten der mathematischen Sprache.
2. Was sind wesentliche Tätigkeiten zum Erlernen einer Sprache?
3. Erläutern Sie den Begriff der Definition.
4. Erläutern Sie den Begriff des Theorems.
5. Erläutern Sie den Begriff des Beweises.
6. Geben Sie die Definition einer mathematischen Aussage wieder.
7. Geben Sie die Definition der Negation einer mathematischen Aussage wieder.
8. Geben Sie die Definition der Konjunktion zweier mathematischer Aussagen wieder.
9. Geben Sie die Definition der Disjunktion zweier mathematischer Aussagen wieder.
10. Geben Sie die Definition der Implikation wieder.
11. Geben Sie die Definition der Äquivalenz wieder.
12. Geben Sie die Definition der logischen Äquivalenz wieder.
13. Erläutern Sie die Begriffe des direkten Beweises, des indirekten Beweises und des Beweises durch Widerspruch.

2. Mengen

2.1. Grundlegende Definitionen

Mengen fassen mathematische Objekte wie beispielsweise Zahlen zusammen und bilden die Grundlage der modernen Mathematik. Wir beginnen mit folgender Definition.

Definition 2.1 (Mengen)**.** Nach Cantor (1895) ist eine *Menge* definiert als "eine Zusammenfassung M von bestimmten wohlunterschiedenen Objekten m unsere Anschauung oder unseres Denken (welche die Elemente der Menge genannt werden) zu einem Ganzen". Wir schreiben

$$m \in M \text{ bzw. } m \notin M \tag{2.1}$$

um auszudrücken, dass m ein Element bzw. kein Element von M ist.

•

Zur Definition von Mengen gibt es mindestens folgende Möglichkeiten:

- Auflisten der Elemente in geschweiften Klammern, z.B. $M := \{1, 2, 3\}$.
- Angabe der Eigenschaften der Elemente, z.B. $M := \{x \in \mathbb{N} | x < 4\}$.
- Gleichsetzen mit einer anderen eindeutig definieren Menge, z.B. $M := \mathbb{N}_3$.

Die Schreibweise $\{x \in \mathbb{N} | x < 4\}$ wird gelesen als "$x \in \mathbb{N}$, für die gilt, dass $x < 4$ ist", wobei die Bedeutung von $\mathbb{N}$ im Folgenden noch zu erläutern sein wird. Es ist wichtig zu erkennen, dass Mengen *ungeordnete* mathematische Objekte sind, dass heißt die Reihenfolge der Auflistung der Elemente einer Menge spielt keine Rolle. Zum Beispiel bezeichnen $\{1, 2, 3\}$, $\{1, 3, 2\}$ und $\{2, 3, 1\}$ dieselbe Menge, nämlich die Menge der ersten drei natürlichen Zahlen.

Grundlegende Beziehungen zwischen mehreren Mengen werden in der nächsten Definition festgelegt.

Definition 2.2 (Teilmengen und Mengengleichheit)**.** A und B seien zwei Mengen.

- Eine Menge A heißt *Teilmenge* einer Menge B, wenn für jedes Element $a \in A$ gilt, dass auch $a \in B$. Ist A eine Teilmenge von B, so schreibt man

$$A \subseteq B \tag{2.2}$$

  und nennt A *Untermenge* von B und B *Obermenge* von A.
- Eine Menge A heißt *echte Teilmenge* einer Menge B, wenn für jedes Element $a \in A$ gilt, dass auch $a \in B$, es aber zumindest ein Element $b \in B$ gibt, für das gilt $b \notin A$. Ist A eine echte Teilmenge von B, so schreibt man

$$A \subset B. \tag{2.3}$$

- Zwei Mengen A und B heißen *gleich*, wenn für jedes Element $a \in A$ gilt, dass auch $a \in B$, und wenn für jedes Element $b \in B$ gilt, dass auch $b \in A$. Sind die Mengen A und B gleich, so schreibt man
$$A = B. \tag{2.4}$$

•

Betrachten wir zum Beispiel die Mengen $A := \{1\}$, $B := \{1, 2\}$, und $C := \{1, 2\}$. Dann gilt mit obigen Definitionen, dass $A \subset B$, weil $1 \in A$ und $1 \in B$, aber $2 \in B$ und $2 \notin A$. Weiterhin gilt, dass $B \subseteq C$, weil $1 \in B$ und $1 \in C$ sowie $2 \in B$ und $2 \in C$ und es kein Element von C gibt, welches nicht in B ist. Ebenso gilt $C \subseteq B$, weil $1 \in C$ und $1 \in B$ sowie $2 \in C$ und $2 \in B$ und es kein Element von B gibt, welches nicht in C ist. Schließlich gilt sogar $B = C$, weil für jedes Element $b \in B$ gilt, dass auch $b \in C$, und gleichzeitig für jedes Element $c \in C$ gilt, dass auch $c \in B$.

Eine wichtige Eigenschaft einer Menge ist die Anzahl der in ihr enthaltenen Elemente. Diese wird als *Kardinalität* der Menge bezeichnet.

Definition 2.3 (Kardinalität)**.** Die Anzahl der Elemente einer Menge M heißt *Kardinalität* und wird mit $|M|$ bezeichnet.

•

Eine besondere Menge ist die Menge ohne Elemente.

Definition 2.4. Eine Menge mit Kardinalität Null heißt *leere Menge* und wird mit $\emptyset$ bezeichnet.

•

Als Beispiele seien $A := \{1, 2, 3\}$, $B = \{a, b, c, d\}$ und $C := \{\,\}$. Dann gelten $|A| = 3$, $B = 4$ und $|C| = 0$.

Zu jeder Menge kann man die Menge aller Teilmengen dieser Menge betrachten. Dies führt auf den wichtigen Begriff der *Potenzmenge*.

Definition 2.5 (Potenzmenge)**.** Die Menge aller Teilmengen einer Menge M heißt *Potenzmenge von M* und wird mit $\mathcal{P}(M)$ bezeichnet.

•

Man beachte, dass die leere Untermenge von M und M selbst auch immer Elemente von $\mathcal{P}(M)$ sind. Wir betrachten vier Beispiele zum Begriff der Potenzmenge.

- $M_0 := \emptyset$ sei die leere Menge. Dann gilt
$$\mathcal{P}(M_0) = \emptyset. \tag{2.5}$$
- M_1 sei die einelementige Menge $M_1 := \{a\}$. Dann gilt
$$\mathcal{P}(M_1) = \{\emptyset, \{a\}\}. \tag{2.6}$$

- Es sei $M_2 := \{a, b\}$. Dann hat M_2 sowohl ein- als auch zweielementige Teilmengen und es gilt
$$\mathcal{P}(M_2) = \{\emptyset, \{a\}\,\{b\}, \{a, b\}\}. \tag{2.7}$$

- Schließlich sei $M_3 := \{a, b, c\}$. Dann hat M ein-, zwei-, als auch dreielementige Teilmengen und es gilt
$$\mathcal{P}(M_3) = \{\emptyset, \{a\}, \{b\}, \{c\}, \{a, b\}, \{a, c\}, \{b, c\}, \{a, b, c\}\}. \tag{2.8}$$

Hinsichtlich der Kardinalitäten einer Menge und ihrer Potenzmenge kann man beweisen, dass aus $|M| = n$ mit $n > 0$ folgt, dass die Kardinalität der Potenzmenge $|\mathcal{P}(M)| = 2^n$ ist. In den obigen Beispielen haben wir die Fälle $|M_1| = 1$ und somit $|\mathcal{P}(M_1)| = 2^1 = 2$, $|M_2| = 2$ und somit $|\mathcal{P}(M_3)| = 2^2 = 4$ und schließlich $|M_3| = 3$ und somit $|\mathcal{P}(M_3)| = 2^3 = 8$, wovon man sich durch Nachzählen schnell überzeugt.

2.2. Verknüpfungen von Mengen

Zwei Mengen können auf unterschiedliche Weise miteinander verknüpft werden. Das Ergebnis einer solchen Verknüpfung ist eine weitere Menge. Wir bezeichnen die Verknüpfung zweier Mengen als *Mengenoperation* und geben folgende Definitionen.

Definition 2.6 (Mengenoperationen)**.** M und N seien zwei Mengen.

- Die *Vereinigung von M und N* ist definiert als die Menge
$$M \cup N := \{x|x \in M \vee x \in N\}, \tag{2.9}$$
wobei $\vee$ wie immer im inklusiven Sinne als und/oder zu verstehen ist.

- Der *Durchschnitt von M und N* ist definiert als die Menge
$$M \cap N := \{x|x \in M \wedge x \in N\}. \tag{2.10}$$
Wenn für M und N gilt, dass $M \cap N = \emptyset$, dann heißen M und N *disjunkt*.

- Die *Differenz von M und N* ist definiert als die Menge
$$M \setminus N := \{x|x \in M \wedge x \notin N\}. \tag{2.11}$$
Die Differenz M und N heißt, insbesondere bei $M \subseteq N$, auch das *Komplement von N bezüglich M* und wird mit N^c bezeichnet.

- Die *symmetrische Differenz von M und N* ist definiert als die Menge
$$M \Delta N := \{x|(x \in M \vee x \in N) \wedge x \notin M \cap N\}, \tag{2.12}$$
Die symmetrische Differenz kann also als *exklusives oder* verstanden werden.

•

Als Beispiel betrachten wir die Mengen $M := \{1, 2, 3\}$und $N := \{2, 3, 4, 5\}$. Dann gelten

- $M \cup N = \{1,2,3,4,5\}$, weil $1 \in M$, $2 \in M$, $3 \in M$, $4 \in N$ und $5 \in N$.
- $M \cap N = \{2,3\}$, weil nur für 2 und 3 gilt, dass $2 \in M, 3 \in M$ und auch $2 \in N, 3 \in N$. Für 1 gilt lediglich, dass $1 \in M$ und für 4 und 5 gelten lediglich, dass $4 \in N$ und $5 \in N$.
- $M \setminus N = \{1\}$, weil $1 \in M$, aber $1 \notin N$ und $2 \in M$, aber auch $2 \in N$.
- $N \setminus M = \{4,5\}$, weil $2 \in N$ und $3 \in N$, aber auch $2 \in M$ und $3 \in M$. Dies zeigt insbesondere, dass die Differenz von M Und N *nicht* symmetrisch ist, also dass *nicht* zwangsläufig gilt, dass $M \setminus N$ gleich $N \setminus M$ ist.
- $M \Delta N = \{1,4,5\}$, weil $1 \in M$, aber $1 \notin \{2,3\}$, $2 \in M$, aber $2 \in \{2,3\}$, $3 \in M$, aber $3 \in \{2,3\}$, $4 \in N$, aber $4 \notin \{2,3\}$ und $5 \in N$, aber $5 \notin \{2,3\}$.

Schließlich wollen wir noch den Begriff der Partition einer Menge einführen.

Definition 2.7 (Partition)**.** M sei eine Menge und $P := \{N_i\}$ sei eine Menge von Mengen N_i mit $i = 1, ..., n$, so dass gilt

$$M = \cup_{i=1}^{n} N_i \wedge N_i \cap N_j = \emptyset \text{ für } i = 1, ..., n, j = 1, ..., n, i \neq j. \tag{2.13}$$

Dann heißt P eine *Partition von M*.

•

Intuitiv entspricht die Partition einer Menge also dem Aufteilen der Menge in disjunkte Teilmengen. Partitionen sind generell nicht eindeutig, d.h. es gibt meist verschiedene Möglichkeiten eine gegebene Menge zu partitionieren. Betrachten wir zum Beispiel die Menge $M := \{1,2,3,4,5,6\}$. Dann sind $P_1 := \{\{1\},\{2,3,4,5,6\}\}$, $P_2 := \{\{1,2,3\},\{4,5,6\}\}$ und $P_3 := \{\{1,2\},\{3,4\},\{5,6\}\}$ drei mögliche Partitionen von M.

2.3. Spezielle Mengen

In der Naturwissenschaft versucht man, Phänomene der Welt mit Zahlen zu beschreiben. Je nach Phänomen bieten sich dazu *diskrete* oder *kontinuierliche* Zahlenmengen an. Die Mathematik stellt dazu unter anderem die in folgender Definition gegebenen Zahlenmengen bereit.

Definition 2.8 (Zahlenmengen)**.** Es bezeichnen

- $\mathbb{N} := \{1,2,3,...\}$ die *natürlichen Zahlen*,
- $\mathbb{N}_n := \{1,2,3,...,n\}$ die *natürlichen Zahlen der Ordnung n*,
- $\mathbb{N}^0 := \mathbb{N} \cup \{0\}$ die *natürlichen Zahlen* und Null,
- $\mathbb{Z} := \{..., -3, -2, -1, 0, 1, 2, 3...\}$ die *ganzen Zahlen*,
- $\mathbb{Q} := \{\frac{p}{q} | p, q \in \mathbb{Z}, q \neq 0\}$ die *rationalen Zahlen*,
- $\mathbb{R}$ die *reellen Zahlen*, und
- $\mathbb{C} := \{a + ib | a, b \in \mathbb{R}, i := \sqrt{-1}\}$ die *komplexen Zahlen*.

•

Die natürlichen und ganzen Zahlen eignen sich insbesondere zum Quantifizieren diskreter Phänomene. Die rationalen und insbesondere die reellen Zahlen eignen sich zum Quantifizieren kontinuierlicher Phänomene. $\mathbb{R}$ umfasst dabei die rationalen Zahlen und die sogenannten *irrationalen Zahlen* $\mathbb{R} \setminus \mathbb{Q}$. Rationale Zahlen sind Zahlen, die sich, wie oben definiert, durch Brüche ganzer Zahlen ausdrücken lassen. Dies sind alle ganzen Zahlen sowie die negativen und positiven Dezimalzahlen wie z.B. $-\frac{9}{10} = -0.9$, $\frac{1}{3} = 1.3\bar{3}$, und $\frac{196}{100} = 1.96$. Irrationale Zahlen sind Zahlen, die sich nicht als rationale Zahlen ausdrücken lassen. Beispiele für irrationale Zahlen sind die *Eulersche Zahl* $e \approx 2.71$, die *Kreiszahl* $\pi \approx 3.14$ und die Quadratwurzel von 2, $\sqrt{2} \approx 1.41$.

Die reellen Zahlen enthalten als Teilmengen die natürlichen, ganzen, und die rationalen Zahlen. Es gibt also sehr viele reelle Zahlen. Tatsächlich kann man beweisen (Cantor (1892)), dass es mehr reelle Zahlen als natürliche Zahlen gibt, obwohl es sowohl unendlich viele reelle Zahlen als auch unendlich viele natürliche Zahlen gibt. Diese Eigenschaft der reellen Zahlen bezeichnet man als die *Überabzählbarkeit* der reellen Zahlen. Insbesondere gilt

$$\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}. \tag{2.14}$$

Zwischen zwei reellen Zahlen gibt es unendlich viele weitere reelle Zahlen. Positiv-Unendlich (∞) und Negativ-Unendlich ($-\infty$) sind keine Zahlen, mit denen in der Standardmathematik gerechnet werden kann. Sie gehören auch nicht zu den in obiger Definition gegebenen Zahlenmengen, es gilt also sowohl $\infty \notin \mathbb{R}$ als auch $-\infty \notin \mathbb{R}$.

Komplexe Zahlen eignen sich zur Beschreibung zweidimensionaler kontinuierlicher Phänomene. Dabei werden die Werte der ersten Dimension im reellen Teil a und die Werte der zweiten Dimension im komplexen Teil b einer komplexen Zahl repräsentiert. Komplexe Zahlen kommen insbesondere bei der Modellierung physikalischer Phänomene und im Bereich der Fourieranalyse zum Einsatz. Wir vertiefen die Theorie komplexer Zahlen an dieser Stelle nicht.

Wichtige Teilmengen der reellen Zahlen sind die sogenannten *Intervalle*. Wir geben folgende Definitionen.

Definition 2.9. Zusammenhängende Teilmengen der reellen Zahlen heißen *Intervalle*. Für $a, b \in \mathbb{R}$ unterscheidet man

- das *abgeschlossene Intervall*

$$[a, b] := \{x \in \mathbb{R} | a \le x \le b\}, \tag{2.15}$$

- das *offene Interval*

$$]a, b[:= \{x \in \mathbb{R} | a < x < b\}, \tag{2.16}$$

- und die *halboffenen Intervalle*

$$]a, b] := \{x \in \mathbb{R} | a < x \le b\} \text{ und } [a, b[:= \{x \in \mathbb{R} | a \le x < b\}. \tag{2.17}$$

•

Wie oben erwähnt sind Positiv-Unendlich (∞) und Negativ-Unendlich ($-\infty$) keine Elemente von $\mathbb{R}$. Es gilt also immer $]-\infty, b]$ oder $]-\infty, b[$ bzw. $]a, \infty[$ oder $[a, \infty[$, sowie $\mathbb{R} =]-\infty, \infty[$.

Oft möchte man mehrere Eigenschaften eines Phänomens gleichzeitig quantitativ beschreiben. Zu diesem Zweck können die oben definierten eindimensionalen Zahlenmenge durch Bildung *Kartesischer Produkte* auf mehrdimensionale Zahlenmengen erweitert werden. Die Elemente Kartesischer Produkte nennt man *geordnete Tupel* oder auch *Vektoren*.

Definition 2.10 (Kartesische Produkte)**.** M und N seien zwei Mengen. Dann ist das *Kartesische Produkt der Mengen M und N* die Menge aller geordneten Tupel (m, n) mit $m \in M$ und $n \in N$, formal

$$M \times N := \{(m, n) | m \in M, n \in N\}. \tag{2.18}$$

Das Kartesische Produkt einer Menge M mit sich selbst wird bezeichnet mit

$$M^2 := M \times M. \tag{2.19}$$

Seien weiterhin $M_1, M_2, ..., M_n$ Mengen. Dann ist das *Kartesische Produkt der Mengen $M_1, ..., M_n$* die Menge aller geordneten n-Tupel $(m_1, ..., m_n)$ mit $m_i \in M_i$ für $i = 1, ..., n$, formal

$$\prod_{i=1}^{n} M_i := M_1 \times \cdots \times M_n := \{(m_1, ..., m_n) | m_i \in M_i \text{ für } i = 1, ..., n\}. \tag{2.20}$$

Das n-fache Kartesische Produkt einer Menge M mit sich selbst wird bezeichnet mit

$$M^n := \prod_{i=1}^{n} M := \{(m_1, , ..., m_n) | m_i \in M\}. \tag{2.21}$$

•

Im Gegensatz zu Mengen sind die in Definition 2.10 eingeführten Tupel *geordnet*. Das heißt, für Mengen gilt zum Beispiel $\{1, 2\} = \{2, 1\}$, aber für Tupel gilt $(1, 2) \neq (2, 1)$.

Wie oben beschrieben eignen sich insbesondere die reellen Zahlen zur Beschreibung kontinuierlicher Phänomene. Zur simultanen Beschreibung mehrere Aspekte eines kontinuierlichen Phänomens bietet sich entsprechend die *Menge der reellen Tupel n-ter Ordnung* an.

Definition 2.11 (Menge der reellen Tupel n-ter Ordnung)**.** Das n-fache Kartesische Produkt der reellen Zahlen mit sich selbst wird bezeichnet mit

$$\mathbb{R}^n := \prod_{i=1}^{n} \mathbb{R} := \{x := (x_1, , ..., x_n) | x_i \in \mathbb{R}\} \tag{2.22}$$

und wird "$\mathbb{R}$ hoch n" gesprochen. Wir schreiben die Elemente von $\mathbb{R}^n$ als Spalten

$$x := \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \tag{2.23}$$

und nennen sie *n-dimensionale Vektoren*. Zu Abgrenzung nennen wir die Elemente von $\mathbb{R}^1 = \mathbb{R}$ auch *Skalare*.

•

Ein Beispiel für $x \in \mathbb{R}^4$ ist

$$x = \begin{pmatrix} 0.16 \\ 1.76 \\ 0.23 \\ 7.11 \end{pmatrix}. \tag{2.24}$$

2.4. Selbstkontrollfragen

1. Geben Sie die Definition einer Menge nach Cantor (1895) wieder.
2. Nennen Sie drei Möglichkeiten zur Definition einer Menge.
3. Erläutern Sie die Ausdrücke $m \in M, m \notin N, M \subseteq N, M \subset N$ für zwei Mengen M und N.
4. Geben Sie die Definition der Kardinalität einer Menge wieder.
5. Geben Sie die Definition der Potenzmenge einer Menge wieder.
6. Es sei $M := \{1, 2\}$. Bestimmen Sie $\mathcal{P}(M)$.
7. Es seien $M := \{1, 2\}, N := \{1, 4, 5\}$. Bestimmen Sie $M \cup N, M \cap N, M \setminus N, M \Delta N$.
8. Erläutern Sie die Symbole $\mathbb{N}$, $\mathbb{N}_n$, und $\mathbb{N}^0$.
9. Erläutern Sie die Unterschiede zwischen $\mathbb{N}$ und $\mathbb{Z}$ und zwischen $\mathbb{R}$ und $\mathbb{Q}$.
10. Geben Sie die Definition abgeschlossener, offener, und halboffener Intervalle wieder.
11. Es seien M und N Mengen. Erläutern Sie die Notation $M \times N$.
12. Geben Sie die Definition von $\mathbb{R}^n$ wieder.

3. Summen, Produkte, Potenzen

3.1. Summen

Diese Einheit führt einige Schreibweisen für die Grundrechenarten ein.

Definition 3.1 (Summenzeichen)**.** Es bezeichnet

$$\sum_{i=1}^{n} x_i = x_1 + x_2 + \dots + x_n. \tag{3.1}$$

Dabei stehen

- Σ für das griechische *S*igma, mnemonisch für*S*umme,
- das Subskript $i = 1$ für den *Laufindex* und den *Startindex*,
- das Superskript n für den *Endindex* und
- $x_1, x_2, ..., x_n$ für die *Summanden.*

•

Für die sinnvolle Benutzung des Summenzeichens ist es essentiell, dass mit mithilfe des Subskripts und des Superskripts Anfang und Ende der Summation festgelegt werden. Die genaue Bezeichnung des Laufindexes ist dagegen für den Wert der Summe irrelevant, es gilt

$$\sum_{i=1}^{n} x_i = \sum_{j=1}^{n} x_j. \tag{3.2}$$

Manchmal wird der Laufindex auch als Element einer *Indexmenge* angegeben. Ist z.B. die Indexmenge $I := \{1, 5, 7\}$ definiert, so ist

$$\sum_{i \in I} x_i := x_1 + x_5 + x_7. \tag{3.3}$$

Im Folgenden wollen wir kurz einige Beispiele für die Benutzung des Summenzeichens betrachten.

- *Summation vordefinierter Summanden.* Es seien $x_1 := 2$, $x_2 := 10$, $x_3 := -4$. Dann gilt
$$\sum_{i=1}^{3} x_i = x_1 + x_2 + x_3 = 2 + 10 - 4 = 8. \tag{3.4}$$

- *Summation natürlicher Zahlen.* Es gilt
$$\sum_{i=1}^{5} i = 1 + 2 + 3 + 4 + 5 = 15. \tag{3.5}$$

- *Summation gerader natürlicher Zahlen.* Es gilt

$$\sum_{i=1}^{5} 2i = 2 \cdot 1 + 2 \cdot 2 + 2 \cdot 3 + 2 \cdot 4 + 2 \cdot 5 = 2 + 4 + 6 + 8 + 10 = 30. \tag{3.6}$$

- *Summation ungerader natürlicher Zahlen.* Es gilt

$$\sum_{i=1}^{5} (2i-1) = 2\cdot1-1+2\cdot2-1+2\cdot3-1+2\cdot4-1+2\cdot5-1 = 1+3+5+7+9 = 25. \tag{3.7}$$

Der Umgang mit dem Summenzeichen wird oft durch die Anwendung folgender Rechenregeln vereinfacht.

Theorem 3.1 (Rechenregeln für Summen)**.**

(1) Summen gleicher Summanden

$$\sum_{i=1}^{n} x = nx \tag{3.8}$$

(2) Assoziativität bei Summen gleicher Länge

$$\sum_{i=1}^{n} x_i + \sum_{i=1}^{n} y_i = \sum_{i=1}^{n} (x_i + y_i) \tag{3.9}$$

(3) Distributivität bei Multiplikation mit einer Konstante

$$\sum_{i=1}^{n} a x_i = a \sum_{i=1}^{n} x_i \tag{3.10}$$

(4) Aufspalten von Summen mit $1 < m < n$

$$\sum_{i=1}^{n} x_i = \sum_{i=1}^{m} x_i + \sum_{i=m+1}^{n} x_i \tag{3.11}$$

(5) Umindizierung

$$\sum_{i=0}^{n} x_i = \sum_{j=m}^{n+m} x_{j-m} \tag{3.12}$$

∘

Beweis. Man überzeugt sich von diesen Rechenregeln durch Ausschreiben der Summen und Anwenden der Rechenregeln von Addition und Multiplikation. Wir zeigen hier exemplarisch die Assoziativität bei Summen gleicher Länge und die Distributivität bei Multiplikation mit einer Konstante. Hinsichtlich ersterer haben wir

$$\begin{aligned}
\sum_{i=1}^{n} x_i + \sum_{i=1}^{n} y_i &= x_1 + x_2 + \cdots + x_n + y_1 + y_2 + \cdots + y_n \\
&= x_1 + y_1 + x_2 + y_2 + \cdots + x_n + y_n \\
&= \sum_{i=1}^{n} (x_i + y_i).
\end{aligned} \tag{3.13}$$

Hinsichtlich letzterer gilt

$$\begin{aligned}\sum_{i=1}^{n} ax_i &= ax_1 + ax_2 + \cdots + ax_n \\ &= a(x_1 + x_2 + \cdots + x_n) \\ &= a\sum_{i=1}^{n} x_i.\end{aligned} \tag{3.14}$$

□

Als Beispiel für die Anwendung einer Rechenregel betrachten wir die Auswertung eines *Mittelwertes* (manchmal auch *Durchschnitt* genannt). Dazu seien $x_1, x_2, ..., x_n$ reelle Zahlen. Der Mittelwert dieser Zahlen entspricht der Summe von $x_1, x_2, ..., x_n$ geteilt durch die Anzahl der Zahlen n. Dabei ist es nach obiger Rechenregel (3) irrelevant, ob zunächst die Zahlen aufaddiert werden und dann die resultierende Summe durch n geteilt wird, oder die Zahlen jeweils einzeln durch n geteilt werden und die entsprechenden Ergebenisse dann aufaddiert werden. Genauer gilt durch Anwendung von Rechenregel (3) mit $a = 1/n$, dass

$$\frac{1}{n}\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} \frac{x_i}{n}. \tag{3.15}$$

So ist zum Beispiel der Mittelwert von $x_1 := 1, x_2 := 4, x_3 := 2\ x_4 := 1$ gegeben durch

$$\frac{1}{4}\sum_{i=1}^{4} x_i = \frac{1}{4}(1+4+2+1) = \frac{8}{4} = 2 = \frac{8}{4} = \frac{1}{4} + \frac{4}{4} + \frac{2}{4} + \frac{1}{4} = \sum_{i=1}^{4} \frac{x_i}{4}. \tag{3.16}$$

3.2. Produkte

Eine analoge Schreibweise zum Summenzeichen bietet das Produktzeichen für Produkte.

Definition 3.2 (Produktzeichen)**.** Es bezeichnet

$$\prod_{i=1}^{n} x_i = x_1 \cdot x_2 \cdot \cdots \cdot x_n. \tag{3.17}$$

Dabei stehen

- $\prod$ für das griechische *P*i, mnemonisch für *P*rodukt,
- das Subskript $i = 1$ für den *Laufindex* und den *Startindex*,
- das Superskript n für den *Endindex*,
- $x_1, x_2, ..., x_n$ für die *Produktterme*

•

Analog zum Summenzeichen gilt, dass das Produktzeichen nur mit Subskript und Superskripten zu Lauf- und Endindex Sinn ergibt. Die genaue Bezeichnung des Laufindizes ist wiederum irrelevant, es gilt

$$\prod_{i=1}^{n} x_i = \prod_{j=1}^{n} x_j. \tag{3.18}$$

Auch hier wird in seltenen Fällen der Laufindex als Element einer Indexmenge angegeben. Ist z.B. die Indexmenge $J := \mathbb{N}_2^0$ definiert, so ist

$$\prod_{j \in J} x_j := x_0 \cdot x_1 \cdot x_2. \tag{3.19}$$

3.3. Potenzen

Produkte von Zahlen mit sich selbst können mithilfe der Potenzschreibweise abgekürzt werden.

Definition 3.3 (Potenz)**.** Für $a \in \mathbb{R}$ und $n \in \mathbb{N}^0$ ist die *n-te Potenz von a* definiert durch

$$a^0 := 1 \text{ und } a^{n+1} := a^n \cdot a. \tag{3.20}$$

Weiterhin ist für $a \in \mathbb{R} \setminus 0$ und $n \in \mathbb{N}^0$ die *negative n-te Potenz von a* definiert durch

$$a^{-n} := (a^n)^{-1} := \frac{1}{a^n}. \tag{3.21}$$

a wird dabei *Basis* und n wird *Exponent* genannt.

•

Die Art der Definition von a^{n+1} mit Rückbezug auf die Potenz a^n in obiger Definition nennt man *rekursiv*. Die Definition $a^0 := 1$ nennt man dabei den *Rekursionsanfang*; er macht die rekursive Definition von a^{n+1} erst möglich. Die Definition $a^{n+1} := a^n \cdot a$ nennt man auch *Rekursionsschritt*. Folgende Rechenregeln vereinfachen das Rechnen mit Potenzen.

Theorem 3.2 (Rechenregeln für Potenzen)**.** *Für $a, b \in \mathbb{R}$ und $n, m \in \mathbb{Z}$ mit $a \neq 0$ bei negativen Exponenten gelten folgende Rechenregeln:*

$$a^n a^m = a^{n+m} \tag{3.22}$$

$$(a^n)^m = a^{nm} \tag{3.23}$$

$$(ab)^n = a^n b^n \tag{3.24}$$

∘

Wir verzichten auf einen Beweis. Beispielsweise gelten also

$$2^2 \cdot 2^3 = (2 \cdot 2) \cdot (2 \cdot 2 \cdot 2) = 2^5 = 2^{2+3}, \tag{3.25}$$

$$(3^2)^3 = (3 \cdot 3)^3 = (3 \cdot 3) \cdot (3 \cdot 3) \cdot (3 \cdot 3) = 3^6 = 3^{2 \cdot 3}, \tag{3.26}$$

und

$$(2 \cdot 4)^2 = (2 \cdot 4) \cdot (2 \cdot 4) = (2 \cdot 2) \cdot (4 \cdot 4) = 2^2 \cdot 4^2. \tag{3.27}$$

In enger Beziehung zur Potenz steht die Definition der $n$ten Wurzel:

Definition 3.4 (n-te Wurzel)**.** Für $a \in \mathbb{R}$ und $n \in \mathbb{N}$ ist die *n-te Wurzel von a* definiert als die Zahl r, so dass

$$r^n = a. \tag{3.28}$$

•

Beim Rechnen mit Wurzeln ist die Potenzschreibweise von Wurzeln oft hilfreich, da sie die direkte Anwendung der Rechenregeln für Potenzen ermöglicht.

Theorem 3.3 (Potenzschreibweise der n-ten Wurzel)**.** *Es sei $a \in \mathbb{R}$, $n \in \mathbb{N}$, und r die n-te Wurzel von a. Dann gilt*

$$r = a^{\frac{1}{n}} \tag{3.29}$$

∘

Beweis. Es gilt

$$\left(a^{\frac{1}{n}}\right)^n = a^{\frac{1}{n}} \cdot a^{\frac{1}{n}} \cdot \dots \cdot a^{\frac{1}{n}} = a^{\sum_{i=1}^n \frac{1}{n}} = a^1 = a. \tag{3.30}$$

Also gilt mit der Definition der n-ten Wurzel, dass $r = a^{\frac{1}{n}}$.

□

Das Rechnen mit Quadratwurzeln wird durch die Potenzschreibweise $\sqrt{x} = x^{\frac{1}{2}}$ sehr erleichtert. Zum Beispiel gilt

$$\frac{2\pi}{\sqrt{2\pi}} = \frac{2\pi}{(2\pi)^{\frac{1}{2}}} = (2\pi)^1 \cdot (2\pi)^{-\frac{1}{2}} = (2\pi)^{1-\frac{1}{2}} = (2\pi)^{\frac{1}{2}} = \sqrt{2\pi}. \tag{3.31}$$

3.4. Selbstkontrollfragen

1. Geben Sie die Definition des Summenzeichens wieder.
2. Berechnen Sie die Summen $\sum_{i=1}^3 2$, $\sum_{i=1}^3 i^2$, und $\sum_{i=1}^3 \frac{2}{3}i$.
3. Schreiben Sie die Summe $1+3+5+7+9+11$ mithilfe des Summenzeichens.
4. Schreiben Sie die Summe $0+2+4+6+8+10$ mithilfe des Summenzeichens.
5. Geben Sie die Definition des Produktzeichens wieder.
6. Geben Sie die Definition der n-ten Potenz von $a \in \mathbb{R}$ wieder.
7. Berechnen Sie $2^2 \cdot 2^3$ und 2^5 und geben Sie die zugehörige Potenzregel wieder.
8. Berechnen Sie 6^2 und $2^2 \cdot 3^2$ und geben Sie die zugehörige Potenzregel wieder.
9. Begründen Sie, warum die n-te Wurzel von a als $a^{\frac{1}{n}}$ geschrieben werden kann.
10. Berechnen Sie $(\sqrt{2})^{\frac{2}{3}}, 9^{\frac{1}{2}}$, und $4^{-\frac{1}{2}}$.

4. Funktionen

Funktionen bilden zusammen mit den Mengen die Grundpfeiler mathematischer Modellierung. In dieser Einheit definieren wir den Begriff der Funktion, führen erste Eigenschaften von Funktionen ein und geben eine Übersicht über einige elementare Funktionen. Funktionen werden äquivalent auch als Abbildungen bezeichnet.

4.1. Definition und Eigenschaften

Definition 4.1 (Funktion)**.** Eine *Funktion* oder *Abbildung* f ist eine Zuordnungsvorschrift, die jedem Element einer Menge D genau ein Element einer Zielmenge Z zuordnet. D wird dabei *Definitionsmenge von* f und Z wird *Zielmenge von* f genannt. Wir schreiben

$$f : D \to Z, x \mapsto f(x), \tag{4.1}$$

wobei $f : D \to Z$ gelesen wird als "die Funktion f bildet alle Elemente der Menge D eindeutig auf Elemente in Z ab" und $x \mapsto f(x)$ gelesen wird als "x, welches ein Element von D ist, wird durch die Funktion f auf $f(x)$ abgebildet, wobei $f(x)$ ein Element von Z ist". Der Pfeil $\to$ steht für die Abbildung zwischen den Mengen D und Z, der Pfeil $\mapsto$ steht für die Abbildung zwischen einem Element von D und einem Element von Z.

•

Es ist zentral, zwischen der *Funktion* f als Zuordnungsvorschrift und einem *Wert der Funktion* $f(x)$ als Element von Z zu unterscheiden. x ist das *Argument* der Funktion (der *Input* der Funktion), $f(x)$ der Wert, den die Funktion f für das Argument x annimmt (der *Output* der Funktion). Üblicherweise folgt in der Definition einer Funktion $f(x)$ die Definition der *funktionalen Form von* f, also einer Regel, wie aus x der Wert $f(x)$ zu bilden ist. Zum Beispiel wird in folgender Definition einer Funktion

$$f : \mathbb{R} \to \mathbb{R}_{\geq 0}, x \mapsto f(x) := x^2 \tag{4.2}$$

die Definition der Potenz genutzt.

Funktionen sind immer *eindeutig*, in dem Sinne dass sie jedem $x \in D$ bei jeder Anwendung der Funktion immer dasselbe $f(x) \in Z$ zuordnen. Funktionen setzen dabei Elemente von Mengen miteinander in Beziehung. Die Mengen dieser Elemente erhalten spezielle Bezeichnungen.

Definition 4.2 (Bildmenge und Urbildmenge)**.** Es sei $f : D \to Z, x \mapsto f(x)$ eine Funktion und es seien $D' \subseteq D$ und $Z' \subseteq Z$. Die Menge

$$f(D') := \{z \in Z | \text{Es gibt ein } x \in D' \text{ mit } z = f(x)\} \tag{4.3}$$

heißt die *Bildmenge von D'* und $f(D) \subseteq Z$ heißt der *Wertebereich* von f. Weiterhin heißt die Menge

$$f^{-1}(Z') := \{x \in D | f(x) \in Z'\} \tag{4.4}$$

die *Urbildmenge von Z'*. $x \in D$ mit $z = f(x) \in Z$ heißt auch *Urbild von z*.

•

Man beachte, dass der Wertebereich $f(D)$ von f und die Zielmenge Z von f sind nicht notwendigerweise identisch sein müssen. Grundlegende Eigenschaften von Funktionen werden in folgender Definition festgelegt.

Definition 4.3 (Injektivität, Surjektivität, Bijektivität)**.** $f : D \to Z, x \mapsto f(x)$ sei eine Funktion. f heißt *injektiv*, wenn es zu jedem Bild $z \in f(D)$ genau ein Urbild $x \in D$ gibt. Äquivalent gilt, dass f injektiv ist, wenn aus $x_1, x_2 \in D$ mit $x_1 \neq x_2$ folgt, dass $f(x_1) \neq f(x_2)$ ist. f heißt *surjektiv*, wenn $f(D) = Z$ gilt, wenn also jedes Element der Zielmenge Z ein Urbild in der Definitionsmenge D hat. Schließlich heißt f *bijektiv*, wenn f injektiv und surjektiv ist. Bijektive Funktionen werden auch *eineindeutige Funktionen* (engl. *one-to-one mappings*) genannt.

•

Abbildung 4.1 verdeutlicht diese Definitionen anhand dreier (Gegen)beispiele.

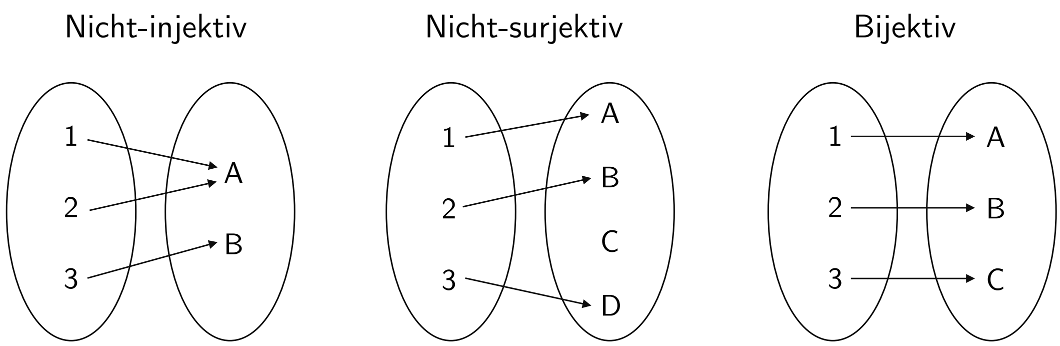

Abbildung 4.1. Injektivität, Surjektivität, Bijektivität.

Abbildung 4.1 A visualisiert die *nicht-injektive* Funktion

$$f : \{1,2,3\} \to \{A, B\}, x \mapsto f(x) := \begin{cases} f(1) & := A \\ f(2) & := A \\ f(3) & := B \end{cases}. \tag{4.5}$$

Die Funktion ist nicht-injektiv, weil es zum Element A in der Bildmenge von f mehr als ein Urbild in der Definitionsmenge von f gibt, nämlich 1 und 2.

Abbildung 4.1 B visualisiert die nicht-surjektive Funktion

$$g : \{1,2,3\} \to \{A, B, C, D\}, x \mapsto g(x) := \begin{cases} g(1) & := A \\ g(2) & := B \\ g(3) & := D \end{cases}. \tag{4.6}$$

Die Funktion ist nicht surjektiv, weil das Element D in der Zielmenge von f kein Urbild in der Definitionsmenge von f hat. Abbildung 4.1 C schließlich visualisiert die bijektive Funktion

$$h : \{1, 2, 3\} \to \{A, B, C\}, x \mapsto g(x) := \begin{cases} h(1) & := A \\ h(2) & := B \\ h(3) & := C \end{cases}. \tag{4.7}$$

Zu *jedem* Element in der Zielmenge von h gibt es *genau ein* Urbild, die Funktion ist also injektiv und surjektiv und damit bijektiv.

Als weiteres Beispiel betrachten wir die Funktion

$$f : \mathbb{R} \to \mathbb{R}, x \mapsto f(x) := x^2 \tag{4.8}$$

Diese Funktion ist nicht injektiv, weil z.B. für $x_1 = 2 \neq -2 = x_2$ gilt, dass $f(x_1) = 2^2 = 4 = (-2)^2 = f(x_2)$. Weiterhin ist f auch nicht surjektiv, weil z.B. $-1 \in \mathbb{R}$ kein Urbild unter f hat. Schränkt man die Definitionsmenge von f allerdings auf die nicht-negativen reellen Zahlen ein, definiert man also die Funktion

$$\tilde{f} : [0, \infty[\to [0, \infty[, x \mapsto \tilde{f}(x) := x^2, \tag{4.9}$$

so ist $\tilde{f}$ im Gegensatz zu f injektiv und surjektiv, also bijektiv.

4.2. Funktionentypen

Durch Verkettung lassen sich aus Funktionen weitere Funktionen bilden.

Definition 4.4 (Verkettung von Funktionen)**.** Es seien $f : D \to Z$ und $g : Z \to S$ zwei Funktionen, wobei die Wertemenge von f mit der Definitionsmenge von g übereinstimmen sollen. Dann ist durch

$$g \circ f : D \to S, x \mapsto (g \circ f)(x) := g(f(x)) \tag{4.10}$$

eine Funktion definiert, die die *Verkettung von f und g* genannt wird.

•

Die Schreibweise für verkettete Funktionen ist etwas gewöhnungsbedürftig. Wichtig ist es zu erkennen, dass $g \circ f$ die verkette Funktion und $(g \circ f)(x)$ ein Element in der Zielmenge der verketten Funktion bezeichnen. Intuitiv wird bei der Auswertung von $(g \circ f)(x)$ zunächst die Funktion f auf x angewendet und dann die Funktion g das Element auf $f(x)$ von R angewendet. Dies ist in der funktionalen Form $g(f(x))$ festgehalten. Der Einfachheit halber benennt man die Verkettung zweier Funktionen auch oft mit einem einzelnen Buchstaben und schreibt beispielsweise, $h := g \circ f$ mit $h(x) = g(f(x))$.

Leicht zur Verwirrung kann es führen, wenn Elemente in der Zielmenge von f mit y bezeichnet werden, also die Schreibweise $y = f(x)$ und $h(x) = g(y)$ genutzt wird. Allerdings ist diese Schreibweise manchmal zur notationellen Vereinfachung nötig.

Als Beispiel für die Verkettung zweier Funktionen betrachten wir

$$f : \mathbb{R} \to \mathbb{R}, x \mapsto f(x) := -x^2 \tag{4.11}$$

und

$$g : \mathbb{R} \to \mathbb{R}, x \mapsto g(x) := \exp(x). \tag{4.12}$$

Die Verkettung von f und g ergibt sich in diesem Fall zu

$$g \circ f : \mathbb{R} \to \mathbb{R}, x \mapsto (g \circ f)(x) := g(f(x)) = \exp\left(-x^2\right). \tag{4.13}$$

Eine erste Anwendung der Verkettung von Funktionen findet sich in folgender Definition.

Definition 4.5 (Inverse Funktion)**.** Es sei $f : D \to Z, x \mapsto f(x)$ eine bijektive Funktion. Dann heißt die Funktion f^{-1} mit

$$f^{-1} \circ f : D \to D, x \mapsto (f^{-1} \circ f)(x) := f^{-1}(f(x)) = x \tag{4.14}$$

inverse Funktion, *Umkehrfunktion* oder einfach *Inverse von f*.

•

Inverse Funktionen sind immer bijektiv. Dies folgt, weil f bijektiv ist und damit jedem $x \in D$ genau ein $f(x) = z \in Z$ zugeordnet wird. Damit wird aber auch jedem $z \in Z$ genau ein $x \in D$, nämlich $f^{-1}(f(x)) = x$ zugeordnet.

Intuitiv macht die inverse Funktion von f den Effekt von f auf ein Element x rückgängig. Betrachtet man den Graphen einer Funktion in einem Kartesischen Koordinatensystem, so führt die Anwendung von einem Wert auf der x-Achse zu einem Wert auf der y-Achse. Die Anwendung der inversen Funktion führt dementsprechend von einem Wert auf der y-Achse zu einem Wert auf der x-Achse. Betrachten wir zum Beispiel die Funktion

$$f : \mathbb{R} \to \mathbb{R}, x \mapsto f(x) := 2x =: y. \tag{4.15}$$

Dann ist die inverse Funktion von f gegeben durch

$$f^{-1} : \mathbb{R} \to \mathbb{R}, y \mapsto f^{-1}(y) := \frac{1}{2}y, \tag{4.16}$$

weil für jedes $x \in \mathbb{R}$ gilt, dass

$$(f^{-1} \circ f)(x) := f^{-1}(f(x)) = f^{-1}(2x) = \frac{1}{2} \cdot 2x = x. \tag{4.17}$$

Eine wichtige Klasse von Funktionen sind *lineare Abbildungen*.

Definition 4.6 (Lineare Abbildung)**.** Eine Abbildung $f : D \to Z, x \mapsto f(x)$ heißt *lineare Abbildung*, wenn für $x, y \in D$ und einen Skalar c gelten, dass

$$f(x + y) = f(x) + f(y)f(cx) = cf(x) \tag{Additivität}$$

und

$$f(cx) = cf(x) \tag{Homogenität}$$

Eine Abbildung, für die obige Eigenschaften nicht gelten, heißt *nicht-lineare Abbildung*.

•

Lineare Abbildungen sind oft als “gerade Linien” bekannt. Die allgemeine Definition linearer Abbildungen ist mit dieser Intuition nicht komplett kongruent. Insbesondere sind lineare Abbildungen nur solche Funktionen, die den Nullpunkt auf den Nullpunkt abbilden. Wir zeigen dazu folgendes Theorem.

Theorem 4.1 (Lineare Abbildung der Null)**.** *$f : D \to Z$ sei eine lineare Abbildung. Dann gilt*

$$f(0) = 0. \tag{4.18}$$

∘

Beweis. Wir halten zunächst fest, dass mit der Additivität von f gilt, dass

$$f(0) = f(0 + 0) = f(0) + f(0). \tag{4.19}$$

Addition von $-f(0)$ auf beiden Seiten obiger Gleichung ergibt dann

$$\begin{aligned} f(0) - f(0) &= f(0) + f(0) - f(0) \\ 0 &= f(0) \end{aligned} \tag{4.20}$$

und damit ist alles gezeigt.

□

Wir wollen den Begriff der linearen Abbildung noch an zwei Beispielen verdeutlichen.

- Für $a \in \mathbb{R}$ ist die Abbildung

  $$f : \mathbb{R} \to \mathbb{R}, x \mapsto f(x) := ax \tag{4.21}$$

  eine lineare Abbildung, weil gilt, dass

  $$f(x+y) = a(x+y) = ax+ay = f(x)+f(y) \text{ und } f(cx) = acx = cax = cf(x). \tag{4.22}$$

- Für $a, b \in \mathbb{R}$ ist dagegen die Abbildung

  $$f : \mathbb{R} \to \mathbb{R}, x \mapsto f(x) := ax + b \tag{4.23}$$

  nicht-linear, weil z.B. für $a := b := 1$ gilt, dass

  $$f(x + y) = 1(x + y) + 1 = x + y + 1 \neq x + 1 + y + 1 = f(x) + f(y). \tag{4.24}$$

Eine Abbildung der Form $f(x) := ax + b$ heißt *linear-affine Abbildung* oder *linear-affine Funktion*. Etwas unsauber werden Funktionen der Form $f(x) := ax + b$ auch manchmal als *lineare Funktionen* bezeichnet.

Neben den bisher diskutierten Funktionentypen gibt es noch viele weitere Klassen von Funktionen. In folgender Definition klassifizieren wir Funktionen anhand der Dimensionalität ihrer Definitions- und Zielmengen. Diese Art der Funktionsklassifikation ist oft hilfreich, um sich einen ersten Überblick über ein mathematisches Modell zu verschaffen.

Definition 4.7 (Funktionenarten)**.** Wir unterscheiden

- *univariate reellwertige Funktionen* der Form
$$f : \mathbb{R} \to \mathbb{R}, x \mapsto f(x), \tag{4.25}$$
- *multivariate reellwertige Funktionen* der Form
$$f : \mathbb{R}^n \to \mathbb{R}, x \mapsto f(x) = f(x_1, ..., x_n), \tag{4.26}$$
- und *multivariate vektorwertige Funktionen* der Form
$$f : \mathbb{R}^n \to \mathbb{R}^m, x \mapsto f(x) = \begin{pmatrix} f_1(x_1, ..., x_n) \\ \vdots \\ f_m(x_1, ..., x_n) \end{pmatrix}, \tag{4.27}$$
wobei $f_i, i = 1, ..., m$ die *Komponenten(funktionen)* von f genannt werden.

•

In der Physik werden multivariate reellwertige Funktionen *Skalarfelder* und multivariate vektorwertige Funktionen *Vektorfelder* genannt. In manchen Anwendungen treten zum Beispiel auch *matrixvariate matrixwertige Funktionen* auf.

4.3. Elementare Funktionen

Als *elementare Funktionen* bezeichnen wir eine kleine Schar von univariaten reellwertigen Funktionen, die häufig als Bausteine komplexerer Funktionen auftreten. Dies sind die *Polynomfunktionen*, die *Exponentialfunktion*, die *Logarithmusfunktion* und die *Gammafunktion*. Im Folgenden geben wir wesentliche Eigenschaften dieser Funktionen und ihre Graphen an. Für Beweise der Eigenschaften der hier vorgestellten F unktionen verweisen wir auf die weiterführende Literatur.

Definition 4.8 (Polynomfunktionen)**.** Eine Funktion der Form
$$f : \mathbb{R} \to \mathbb{R}, x \mapsto f(x) := \sum_{i=0}^{k} a_i x^i = a_0 + a_1 x^1 + a_2 x^2 + \cdots + a_k x^k \tag{4.28}$$
heißt *Polynomfunktion* k-ten Grades mit *Koeffizienten* $a_0, a_1, ..., a_k \in \mathbb{R}$.

•

Einige ausgewählte Polynomfunktionen sind in Tabelle 4.1 aufgelistet, Abbildung 4.2 zeigt die enstprechende Graphen.

Tabelle 4.1. Ausgewählte Polynomfunktionen

Name	Funktionale Form	Koeffizienten
Konstante Funktion	$f(x) = a$	$a_0 := a, a_i := 0, i > 0$
Identitätsfunktion	$f(x) = x$	$a_0 := 0, a_1 := 1, a_i := 0, i > 1$
Linear-affine Funktion	$f(x) = ax + b$	$a_0 := b, a_1 := a, a_i := 0, i > 1$
Quadratfunktion	$f(x) = x^2$	$a_0 := 0, a_1 := 0, a_2 := 1, a_i := 0, i > 2$

Ein wichtiges Funktionenpaar sind die Exponentialfunktion und die Logarithmusfunktion. Die Graphen der Exponential- und Logarithmusfunktion sind in Abbildung 4.3 abgebildet.

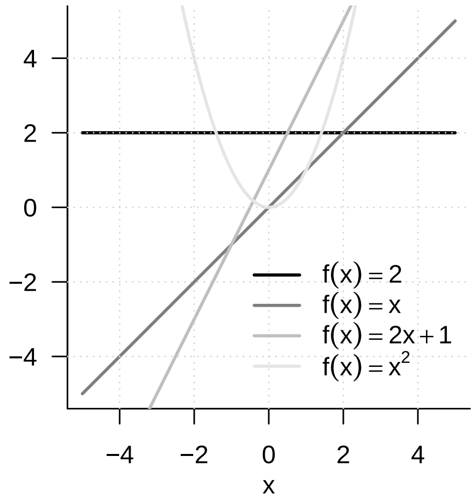

Abbildung 4.2. Ausgewählte Polynomfunktionen

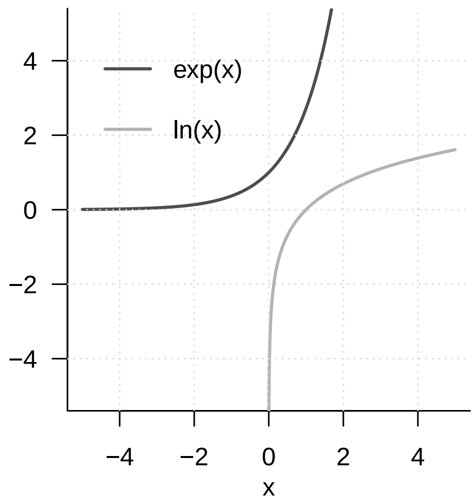

Abbildung 4.3. Exponentialfunktion und Logarithmusfunktion

Definition 4.9 (Exponentialfunktion)**.** Die *Exponentialfunktion* ist definiert als

$$\exp : \mathbb{R} \to \mathbb{R}, x \mapsto \exp(x) := e^x := \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \cdots. \tag{4.29}$$

•

Die Exponentialfunktion hat unter anderem folgende Eigenschaften.

Wertebereich der Exponentialfunktion

- $x \in]-\infty, 0[\Rightarrow \exp(x) \in]0, 1[$
- $x \in]0, \infty[\quad \Rightarrow \exp(x) \in]1, \infty[$

Insbesondere nimmt die Exponentialfunktion also nur positive Werte an.

Monotonieeigenschaft der Exponentialfunktion

- $x < y \Rightarrow \exp(x) < \exp(y)$

Spezielle Werte der Exponentialfunktion

- $\exp(0) = 1$
- $\exp(1) = e \approx 2.71$

Die Logarithmusfunktion schneidet die y-Achse also bei 0. Die Zahl e heißt *Eulersche Zahl*.

Summationseigenschaft und Subtraktionseigenschaft der Exponentialfunktion

- $\exp(x + y) = \exp(x) \exp(y)$
- $\exp(x - y) = \frac{\exp(x)}{\exp(y)}$

Mit den speziellen Werten der Exponentialfunktion gilt dann insbesondere auch

$$\exp(x) \exp(-x) = \exp(x - x) = \exp(0) = 1. \tag{4.30}$$

Definition 4.10 (Logarithmusfunktion)**.** Die *Logarithmusfunktion* ist definiert als inverse Funktion der Exponentialfunktion,

$$\ln :]0, \infty[\to \mathbb{R}, x \mapsto \ln(x) \text{ mit } \ln(\exp(x)) = x \text{ für alle } x \in \mathbb{R}. \tag{4.31}$$

•

Die Logarithmusfunktion hat unter anderem folgende Eigenschaften.

Wertebereich der Logarithmusfunktion

- $x \in]0, 1[\ \Rightarrow \ln(x) \in]-\infty, 0[$
- $x \in]1, \infty[\Rightarrow \ln(x) \in]0, \infty[$

Die Logarithmusfunktion nimmt also sowohl negative als auch positive Werte an.

Monotonie der Logarithmusfunktion

- $x < y \Rightarrow \ln(x) < \ln(y)$

Spezielle Werte der Logarithmusfunktion

- $\ln(1) = 0$ und $\ln(e) = 1$.

Die Logarithmusfunktion schneidet die x-Achse also bei 1.

Produkteigenschaft, Potenzeigenschaft und Divisionseigenschaft der Logarithmusfunktion

- $\ln(xy) = \ln(x) + \ln(y)$
- $\ln(x^c) = c\ln(x)$
- $\ln\left(\frac{1}{x}\right) = -\ln(x)$

Letztere Eigenschaft sind beim Rechnen Logarithmusfunktionen zentral. Man merkt sie sich intuitiv als "Die Logarithmusfunktion wandelt Produkte in Summen und Potenzen in Produkte um."

Ein häufiger Begleiter in der Wahrscheinlichkeitstheorie ist die *Gammafunktion*. Ein Auschnitt des Graphen der Gammafunktion ist in Abbildung 4.4 dargestellt.

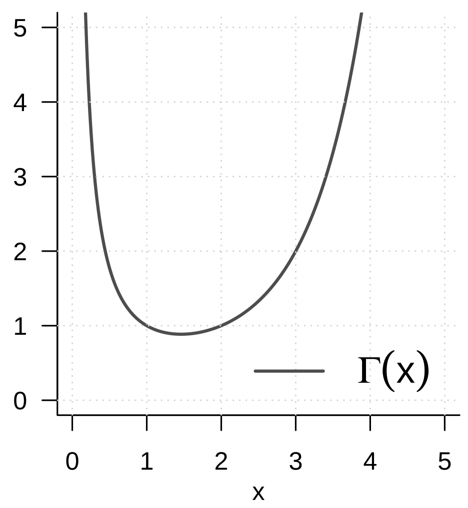

Abbildung 4.4. Gammafunktion

Definition 4.11 (Gammafunktion)**.** Die *Gammafunktion* ist definiert durch

$$\Gamma : \mathbb{R} \to \mathbb{R}, x \mapsto \Gamma(x) := \int_0^\infty \xi^{x-1} \exp(-\xi)\, d\xi \tag{4.32}$$

•

Die Gammafunktion hat folgende Eigenschaften:

Spezielle Werte der Gammafunktion

- $\Gamma(1) = 1$
- $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$
- $\Gamma(n) = (n-1)!$ für $n \in \mathbb{N}$.

Rekursionseigenschaft der Gammafunktion

- Für $x > 0$ gilt $\Gamma(x+1) = x\Gamma(x)$

4.4. Selbstkontrollfragen

1. Geben Sie die Definition einer Funktion wieder.
2. Geben Sie die Definition der Begriffe Bildmenge, Wertebereich, und Urbildmenge wieder.
3. Geben Sie die Definitionen der Begriffe Surjektivität, Injektivität, und Bijektivität wieder.
4. Erläutern Sie, warum $f : \mathbb{R} \to \mathbb{R}, x \mapsto f(x) := x^2$ weder injektiv noch surjektiv ist.
5. Erläutern Sie, warum $f : [0, \infty[\to [0, \infty[, x \mapsto f(x) := x^2$ bijektiv ist.
6. Geben Sie die Definition der Verkettung von Funktionen wieder.
7. Geben Sie die Definition des Begriffs der inversen Funktion wieder.
8. Geben Sie die inverse Funktion von x^2 auf $[0, \infty[$ an.
9. Geben Sie die Definition des Begriffs der linearen Abbildung wieder.
10. Geben Sie die Definitionen der Begriffe der univariat-reellwertigen, multivariat-reellwertigen und multivariat-vektorwertigen Funktion wieder.
11. Skizzieren Sie die Identitätsfunktion und die konstante Funktion für $a := 1$.
12. Skizzieren Sie die linear-affine Funktion $f(x) = ax + b$ für $a = 2$ und $b = 3$.
13. Skizzieren Sie die Funktionen $f(x) := (x-1)^2$ und $g(x) := (x+3)^2$.
14. Skizzieren Sie die Exponential- und Logarithmusfunktionen.
15. Geben Sie die Summations- und Subtraktionseigenschaften der Exponentialfunktion an.
16. Geben Sie die Produkt-, Potenz- und Divisionseigenschaften der Logarithmusfunktion an.

Referenzen

Arens, T., Hettlich, F., Karpfinger, C., Kockelkorn, U., Lichtenegger, K., & Stachel, H. (2018). *Mathematik*. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-56741-8

Blei, D. M., Kucukelbir, A., & McAuliffe, J. D. (2017). Variational Inference: A Review for Statisticians. *Journal of the American Statistical Association*, *112*(518), 859–877. https://doi.org/10.1080/01621459.2017.1285773

Cantor, G. (1892). Über Eine Eigenschaft Des Inbegriffes Aller Reellen Algebraischen Zahlen. *Jahresbericht der Deutschen Mathematiker-Vereinigung*, *1*.

Cantor, G. (1895). Beiträge Zur Begründung Der Transfiniten Mengenlehre. *Mathematische Annalen*, *46*(4), 481–512. https://doi.org/10.1007/BF02124929

Friston, K. (2005). A Theory of Cortical Responses. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *360*(1456), 815–836. https://doi.org/10.1098/rstb.2005.1622

Friston, K., Da Costa, L., Sakthivadivel, D. A. R., Heins, C., Pavliotis, G. A., Ramstead, M., & Parr, T. (2023). Path Integrals, Particular Kinds, and Strange Things. *Physics of Life Reviews*, *47*, 35–62. https://doi.org/10.1016/j.plrev.2023.08.016

Newton, I. (1687). *Philosophiae Naturalis Principia Mathematica*. Royal Society.

Ostwald, D., Kirilina, E., Starke, L., & Blankenburg, F. (2014). A Tutorial on Variational Bayes for Latent Linear Stochastic Time-Series Models. *Journal of Mathematical Psychology*, *60*, 1–19. https://doi.org/10.1016/j.jmp.2014.04.003

Unger, L. (2000). *Grundkurs Mathematik*.