

Mathematik I für MB, WI/MB und andere

Profs. Drs. Ulrich Reif, Wilhelm Stannat und Michael Joswig

Winter 2010/11

Das vorliegende Skript ist eine Zusammenfassung der Vorlesungen “Mathematik I für MB, WI/MB und andere,” die im WS 2005/06, WS 2006/07, WS 2007/08 und WS 2010/11 an der TU Darmstadt gehalten wurden.

Inhaltlich teilweise angelehnt an das folgende Buch: K. Graf Finck von Finckenstein, J. Lehn, H. Schellhaas, H. Wegmann: Arbeitsbuch Mathematik für Ingenieure Band I: Analysis und Lineare Algebra, Teubner-Verlag.

Korrekturen bitte per Email an joswig@mathematik.tu-darmstadt.de

Inhaltsverzeichnis

1	Zahlen	1
1.1	Natürliche Zahlen	2
1.2	Ganze Zahlen	2
1.3	Rationale Zahlen	2
1.4	Reelle Zahlen	3
1.5	Komplexe Zahlen	5
1.6	Allgemeines	7
2	Vektorrechnung	11
2.1	Vektoren in \mathbb{R}^n	12
2.2	Geraden im \mathbb{R}^n	16
2.3	Ebenen in \mathbb{R}^3	20
2.4	Lineare Unterräume, Basis und Dimension	23
3	Lineare Gleichungssysteme	25
3.1	Einige geometrisch motivierte Beispiele	25
3.2	Lineare Gleichungssysteme (LGS)	26
3.3	Einschub: Matrizen	27
3.4	Elementare Umformungen	30
3.5	Gestaffelte Form eines LGS	33
3.6	Der Gaußsche Algorithmus	34
3.7	Homogene LGS	35
3.8	Inhomogene LGS	37
3.9	Determinanten	38
3.10	Matrix-Gleichungssysteme	41
3.11	Matrizen mit anderen Koeffizienten	44
4	Lineare Abbildungen	45
4.1	Allgemeines	45
4.2	Spezielle lineare Abbildungen in \mathbb{R}^n	47
4.3	Eigenwerte und -vektoren	51
5	Komplexe Zahlen	57
6	Folgen und Reihen	63
6.1	Folgen	63
6.2	Reihen	67

7	Abbildungen	71
7.1	Grundlegendes	71
7.2	Reelle Funktionen einer Variablen	73
7.3	Elementare Funktionen	75
7.4	Grenzwerte von Funktionen	77
8	Differentialrechnung	81
9	Integralrechnung	89
9.1	Das bestimmte Integral	89
9.2	Hauptsatz der Differential- und Integralrechnung	92
9.3	Uneigentliche Integrale	96
10	Potenzreihen und Taylorreihen	99

Notation für Mengen von Zahlen

$\mathbb{N} = \{1, 2, 3, \dots\}$	natürliche Zahlen
$\mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\}$	ganze Zahlen
$\mathbb{Q} = \left\{\frac{a}{b} : a \in \mathbb{Z} \text{ und } b \in \mathbb{N}\right\}$	rationale Zahlen
$\mathbb{R} =$ Menge aller Dezimalbrüche	reelle Zahlen
$[a, b] = \{x \in \mathbb{R} : a \leq x \leq b\}$	abgeschlossenes Intervall
$]a, b[= \{x \in \mathbb{R} : a < x < b\}$	offenes Intervall
$[a, b[= \{x \in \mathbb{R} : a \leq x < b\}$	halboffenes Intervall
$]a, b] = \{x \in \mathbb{R} : a < x \leq b\}$	halboffenes Intervall
$\mathbb{C} = \{a + ib : a, b \in \mathbb{R}\}$	Menge der komplexen Zahlen

Kapitel 1

Zahlen

Am Anfang aller Mathematik stehen die natürlichen Zahlen

$$1, 2, 3, 4, \dots$$

Für die Menge der natürlichen Zahlen schreiben wir

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

Dabei versteht man unter einer **Menge** M allgemein eine Zusammenfassung von unterscheidbaren Objekten zu einer Gesamtheit. Objekte dieser Gesamtheit heißen **Elemente**.

Beschreibung von Mengen

Wir beschreiben Mengen durch Mengenklammern $\{ \}$, zwischen denen die Elemente angegeben werden:

1. durch Aufzählen

$$M = \{ 1, 2, 3, 4 \}, \quad M = \{ 1, 2, 3, \dots, 100 \}$$

$$M = \{ \text{He, Ne, Ar, Kr, Xe, Ra} \} \quad (= \text{alle Edelgase, VIII. HG})$$

$$M = \{ \} = \emptyset = \text{leere Menge}$$

2. durch charakterisierende Eigenschaft

$$M = \{ x : x \text{ hat Eigenschaft E} \}$$

Etwa:

$$M = \{ x : x \text{ ist natürliche Zahl zwischen 2 und 5} \}$$

Mengen sind gleich, wenn sie dieselben Elemente enthalten. Darstellung und Anordnung spielen dabei keine Rolle:

$$\{ x : x \text{ ist Primzahl kleiner 10} \} = \{ 2, 3, 5, 7 \} = \{ 3, 7, 2, 5 \}$$

Wir schreiben:

$$x \in M \quad \text{falls } x \text{ Element der Menge } M$$

$$x \notin M \quad \text{falls } x \text{ kein Element der Menge } M$$

Mengen können endlich viele oder unendlich viele Elemente enthalten. Die Anzahl der Elemente einer Menge M heißt auch **Kardinalität** von M .

1.1 Natürliche Zahlen

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

Natürliche Zahlen kann man addieren und multiplizieren, die Summe und das Produkt zweier natürlichen Zahlen ist wieder eine natürliche Zahl, d.h.

\mathbb{N} ist abgeschlossen unter **Addition** und **Multiplikation**

1.2 Ganze Zahlen

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

\mathbb{Z} ist ebenfalls abgeschlossen unter Addition und Multiplikation, aber auch bezüglich **Subtraktion**.

Folglich besitzt die Gleichung

$$a + x = b \quad a, b \in \mathbb{Z}$$

eine Lösung $x \in \mathbb{Z}$, nämlich $x = b - a$.

Dies muss für die Lösung der Gleichung $a \cdot x = b$ im allgemeinen **nicht** gelten.

1.3 Rationale Zahlen

$$\mathbb{Q} = \left\{ x : x = \frac{b}{a}, b \in \mathbb{Z}, a \in \mathbb{N} \right\}$$

\mathbb{Q} ist abgeschlossen unter Addition, Multiplikation und Subtraktion, aber auch bezüglich **Division**.

Folglich besitzt die Gleichung

$$a \cdot x = b \quad a, b \in \mathbb{Q}, a \neq 0$$

eine Lösung $x \in \mathbb{Q}$, nämlich $x = \frac{b}{a}$.

Darstellung rationaler Zahlen

Die Darstellung rationaler Zahlen als Quotient ganzer Zahlen ist **nicht eindeutig**:

$$\frac{10}{40} = \frac{1}{4}, \quad \frac{372}{468} = \frac{31 \cdot 12}{39 \cdot 12} = \frac{31}{39}.$$

Die Darstellung $x = \frac{b}{a}$ wird erst dann eindeutig, wenn man fordert, dass a und b teilerfremd sind, $b \in \mathbb{Z}$ und $a \in \mathbb{N}$.

Das **Divisionsverfahren** liefert eine Darstellung als **Dezimalbruch**:

$$\begin{aligned} \frac{7}{4} &= 1.75 && \text{(abbrechender Dezimalbruch)} \\ \frac{1}{3} &= 0.333 \dots = 0.\bar{3}, && \frac{2}{70} = 0.\overline{0285714} \quad \text{(periodischer Dezimalbruch)} \end{aligned}$$

Beim Divisionsverfahren, angewandt auf $x = \frac{b}{a}$, tritt als **Divisionsrest** eine Zahl zwischen 0 und $a - 1$ auf. Bei 0 bricht der Dezimalbruch ab. Wiederholt sich einer der Divisionsreste, so liegt ein periodischer Dezimalbruch vor.

Da es höchstens $a - 1$ verschiedene Divisionsreste gibt, hat die zugehörige Periode höchstens Länge $a - 1$.

Jede rationale Zahl lässt sich also durch einen endlichen oder periodischen Dezimalbruch darstellen.

Umwandlung periodischer Dezimalbrüche in Brüche:

Ist $x = 0.\overline{d_1 d_2 \dots d_k}$ mit Ziffern $d_i \in \{0, 1, \dots, 9\}$, so gilt

$$x = \frac{d_1 d_2 \dots d_k}{10^k - 1}.$$

Beispiele

$$0.\bar{3} = \frac{3}{9} = \frac{1}{3}$$

$$0.\bar{18} = \frac{18}{99} = \frac{2}{11}$$

$$0.1\bar{6} = \frac{1}{10} \cdot 1.\bar{6} = \frac{1}{10} \cdot (1 + 0.\bar{6}) = \frac{1}{10} \cdot \left(1 + \frac{6}{9}\right) = \frac{1}{10} \cdot \frac{15}{9} = \frac{1}{6}.$$

1.4 Reelle Zahlen

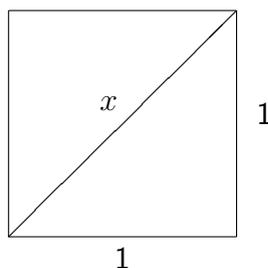
Der Zahlbereich der rationalen Zahlen ist nicht hinreichend mächtig. Zum Beispiel ist der Dezimalbruch

$$0.101001000100001\dots$$

weder endlich noch periodisch. Es kann sich also hierbei um keine rationale Zahl handeln.

Weiteres Beispiel

Für die Länge x der Diagonalen im Quadrat mit Seitenlänge 1 gilt nach dem Satz von Pythagoras $x^2 = 1^2 + 1^2 = 2$.



Wir werden in 1.6 sehen, dass $\sqrt{2}$ keine rationale Zahl ist. Daher besitzt die Gleichung

$$x^2 = 2$$

in \mathbb{Q} keine Lösung.

Da $\sqrt{2}$ keine rationale Zahl ist, bezeichnet man sie als **Irrationalzahl**. $\sqrt{2}$ lässt sich auf folgende Weise durch rationale Zahlen **beliebig gut** annähern:

Weil $1^2 = 1 < 2$ und $2^2 = 4 > 2$ gilt $\sqrt{2} \in [1, 2] =: I_0$. Unterteile nun I_0 in zehn gleich große Teilintervalle und suche dann $\sqrt{2}$:

es gilt $1.4^2 = 1.96 < 2$ und $1.5^2 = 2.25 > 2 \implies \sqrt{2} \in [1.4, 1.5] =: I_1$

Wiederhole das Verfahren

$$\begin{aligned}\sqrt{2} &\in I_2 := [1.41, 1.42] \\ \sqrt{2} &\in I_3 := [1.414, 1.415] \quad \dots\end{aligned}$$

und erhalte damit eine absteigende Folge $I_0 \supset I_1 \supset I_2 \supset \dots$ mit

$$\text{Länge des Intervalls } I_n = 10^{-n} = \frac{1}{10^n} \xrightarrow{n \rightarrow \infty} 0$$

d.h. eine **Intervallschachtelung**. Wir erhalten $\sqrt{2}$ als das Objekt, das durch diese Intervallschachtelung approximiert wird.

Die **Menge der reellen Zahlen** \mathbb{R} ist nun definiert als Gesamtheit aller Objekte, die sich durch derartige Intervallschachtelungen approximieren lassen, und die Zahlengerade liefert die richtige Anschauung für \mathbb{R} :

- in jeder (noch so kleinen) Umgebung einer reellen Zahl gibt es unendlich viele rationale Zahlen.
- jede reelle Zahl lässt sich durch eine (unendliche) Dezimalbruchentwicklung darstellen:

$$\begin{aligned}\mathbb{R} &= \left\{ k + r : k \in \mathbb{Z}, r = 0.d_1d_2d_3\dots = \frac{d_1}{10} + \frac{d_2}{100} + \frac{d_3}{1000} + \dots \right. \\ &\quad \left. \text{mit } d_1, d_2, d_3, \dots \in \{0, 1, \dots, 9\} \right\}\end{aligned}$$

- die vier Grundrechenarten auf \mathbb{Q} („+“, „·“, „-“, „/“) lassen sich auf \mathbb{R} fortsetzen.

Ordnung reeller Zahlen

Ein Blick auf die Zahlengerade zeigt:

(A1) Für zwei reelle Zahlen a, b gilt **genau eine der drei Beziehungen**

$$a < b, \quad a = b, \quad b < a$$

„ $<$ “ gibt eine Ordnung der reellen Zahlen mit folgenden weiteren Eigenschaften

(A2) Aus $a < b$ und $b < c$ folgt $a < c$

(A3) Aus $a < b$ folgt $a + c < b + c$ für alle c .

(A4) Aus $a < b$ und $0 < c$ folgt $a \cdot c < b \cdot c$.

Vereinfachende Schreibweisen

$$a > b \text{ falls } b < a$$

$$a \leq b \text{ falls } a < b \text{ oder } a = b$$

$$a \geq b \text{ falls } b \leq a$$

a heißt **positiv**, falls $a > 0$

a heißt **negativ**, falls $a < 0$

Der **Betrag** einer reellen Zahl a ist definiert durch

$$|a| := \begin{cases} a & \text{falls } a \geq 0 \\ -a & \text{falls } a < 0 \end{cases}$$

Es ist also stets $|a| \geq 0$.

Rechenregeln

$$(i) |a \cdot b| = |a| \cdot |b|$$

$$(ii) |a + b| \leq |a| + |b| \quad (\text{Dreiecksungleichung})$$

1.5 Komplexe Zahlen

In \mathbb{R} besitzt die Gleichung

$$x^2 = -1$$

keine Lösung, denn Quadrate reeller Zahlen sind stets ≥ 0 .

Wir erweitern daher den Zahlbereich \mathbb{R} um ein Element

$$i \quad (\text{„imaginäre Einheit“})$$

mit der Eigenschaft

$$i^2 = -1$$

und rechnen mit i wie mit „normalen“ Zahlen. Wegen $i^2 = -1$ schreibt man auch $i = \sqrt{-1}$.

Die Menge

$$\mathbb{C} = \{z = a + ib : a, b \in \mathbb{R}\}$$

heißt die Menge der **komplexen Zahlen**.

Für $z = a + ib$ heißt

$a =: \operatorname{Re}(z)$ **Realteil** von z

$b =: \operatorname{Im}(z)$ **Imaginärteil** von z

Rechnen mit komplexen Zahlen

Addition $(a + ib) + (c + id) = (a + c) + i(b + d)$

Subtraktion $(a + ib) - (c + id) = (a - c) + i(b - d)$

Multiplikation

$$(a + ib) \cdot (c + id) = ac + aid + ib \cdot c + \underbrace{ib \cdot id}_{=i^2 bd = -bd} = (ac - bd) + i(ad + bc)$$

Division

Für eine komplexe Zahl $z = a + ib \neq 0$ (d.h. a und b nicht beide 0) ist

$$\frac{1}{z} = \frac{a}{a^2 + b^2} - i \frac{b}{a^2 + b^2}$$

der **Kehrwert** (bzw. Reziprokwert bzw. die Inverse), denn

$$\begin{aligned} (a + ib) \cdot \left(\frac{a}{a^2 + b^2} - i \frac{b}{a^2 + b^2} \right) \\ = \left(a \cdot \frac{a}{a^2 + b^2} - b \cdot \left(-\frac{b}{a^2 + b^2} \right) \right) + i \left(a \cdot \left(-\frac{b}{a^2 + b^2} \right) + b \cdot \frac{a}{a^2 + b^2} \right) \\ = 1 + i0 = 1. \end{aligned}$$

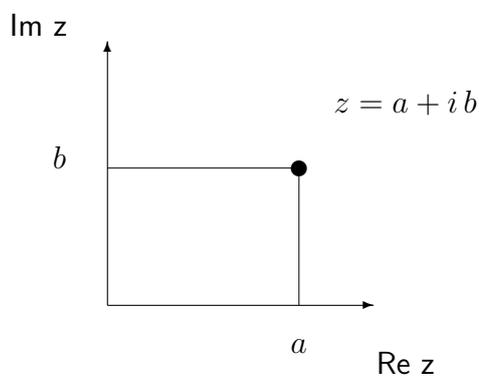
Durch

$$\frac{z_1}{z_2} = z_1 \cdot \frac{1}{z_2}$$

ist dann die Division zweier komplexer Zahlen z_1, z_2 mit $z_2 \neq 0$ erklärt.

Die Gaußsche Zahlenebene

Komplexe Zahlen lassen sich als Punkte in der Ebene veranschaulichen:



Der **Betrag** einer komplexen Zahl $z = a + i b$ ist definiert als die reelle Zahl

$$|z| = \sqrt{a^2 + b^2}$$

Offenbar gilt nach Pythagoras: $|z|$ ist der Abstand des Punktes z zum Nullpunkt.

Rechenregeln

$$(i) |z_1 \cdot z_2| = |z_1| \cdot |z_2|$$

$$(ii) |z_1 + z_2| \leq |z_1| + |z_2| \quad (\text{Dreiecksungleichung})$$

$$(iii) |z| = 0 \text{ genau dann wenn } z = 0$$

Bemerkungen

- Die Addition und die Multiplikation komplexer Zahlen ist sowohl **assoziativ** als auch **kommutativ**. Das heißt, es gilt:

$$z_1 + (z_2 + z_3) = (z_1 + z_2) + z_3 \quad \text{und} \quad z_1 \cdot (z_2 \cdot z_3) = (z_1 \cdot z_2) \cdot z_3$$

sowie

$$z_1 + z_2 = z_2 + z_1 \quad \text{und} \quad z_1 \cdot z_2 = z_2 \cdot z_1.$$

Wegen $\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$ gilt dies für alle Zahlen mit denen wir hier rechnen.

- Die schrittweise Erweiterung der Zahlbereiche $\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$ ist motiviert dadurch, dass man immer kompliziertere Gleichungen lösen will bzw. kann. Bei den komplexen Zahlen ist ein gewisser Schlusspunkt erreicht: Man sagt, dass \mathbb{C} ein **algebraisch abgeschlossener Körper** ist.
- Die reellen Zahlen sind **(an-)geordnet**, die komplexen Zahlen jedoch nicht!

1.6 Allgemeines

Zum Abschluss diese Kapitels noch einige allgemeine Ergänzungen.

Vereinfachende Schreibweisen

Sind a_m, a_{m+1}, \dots, a_n reelle (oder komplexe) Zahlen, so schreiben wir

$$\sum_{i=m}^n a_i := a_m + a_{m+1} + \dots + a_n$$

für die Summe und

$$\prod_{i=m}^n a_i := a_m \cdot a_{m+1} \cdot \dots \cdot a_n$$

für das Produkt dieser Zahlen.

Im Falle $m > n$ setzen wir

$$\sum_{i=m}^n a_i := 0 \quad \text{und} \quad \prod_{i=m}^n a_i := 1.$$

Beweismethoden

Mathematische Aussagen (Theoreme, Sätze, Hilfssätze, Lemmata und Folgerungen) bedürfen eines Beweises im Unterschied zu Axiomen („Annahmen“), die sich nicht beweisen lassen. Alle Aussagen gelten stets nur relativ zum gewählten Axiomensystem.

Wir wollen auf drei Beweisschemata im Folgenden näher eingehen:

1. Direkter Beweis

Ausgehend von den Voraussetzungen führen schrittweise Folgerungen direkt zur Aussage.

Beispiele

(i) **Aussage:** Für $q \in \mathbb{R}$ und $n = 0, 1, 2, \dots$ gilt

$$\sum_{i=0}^n q^i = \begin{cases} \frac{1-q^{n+1}}{1-q} & \text{für } q \neq 1 \\ n+1 & \text{für } q = 1 \end{cases}$$

Beweis: Es sei $s_n = \sum_{i=0}^n q^i$. Für $q = 1$ ist $s_n = n + 1$. Für $q \neq 1$ gilt

$$\begin{aligned} (1-q)s_n &= s_n - qs_n \\ &= 1 + q + q^2 + \dots + q^n - (q + q^2 + q^3 + \dots + q^{n+1}) \\ &= 1 - q^{n+1}. \end{aligned}$$

Division durch $1 - q$ ergibt

$$s_n = \frac{1 - q^{n+1}}{1 - q}.$$

(ii) **Aussage:** Für $n \geq 3$ ist $2n^2 \geq (n+1)^2$.

Beweis: Für $n \geq 3$ ist

$$n^2 = n \cdot n \geq 3 \cdot n = 2n + n \geq 2n + 1,$$

also

$$2n^2 = n^2 + n^2 \geq n^2 + 2n + 1 = (n+1)^2.$$

2. Indirekter Beweis

Zum Beweis der Aussage nimmt man an, dass das logische Gegenteil richtig sei (Gegenannahme). Ausgehend von dieser Gegenannahme führt man dann durch schrittweise Folgerungen einen Widerspruch herbei. Das logische Gegenteil zur Aussage ist daher falsch und die Aussage somit richtig.

Beispiel

Aussage: Es gibt keine rationale Zahl x mit $x^2 = 2$.

Beweis: Gegenannahme: Es gibt ein $x \in \mathbb{Q}$ mit $x^2 = 2$.

Dann gibt es $a \in \mathbb{Z}, b \in \mathbb{N}$ mit $x = \frac{a}{b}$. Also gilt weiter $x^2 = \frac{a^2}{b^2} = 2$ oder $a^2 = 2b^2$. Durch Kürzen können wir annehmen, dass a und b teilerfremd sind.

Wegen $a^2 = 2b^2$ muss a^2 gerade sein und damit auch a , also $a = 2k$ für $k \in \mathbb{Z}$. Also $a^2 = 4k^2 = 2b^2$ oder $2k^2 = b^2$. Damit ist aber b^2 gerade, also auch b gerade.

Dies ist ein Widerspruch dazu, dass a und b teilerfremd sein sollen. Die Gegenannahme ist also falsch, die Aussage somit bewiesen.

Bemerkung

Das indirekte Beweisprinzip beruht auf dem Axiom "tertium non datur" (lat.: "ein Drittes gibt es nicht") der **aristotelischen Logik**.

3. Beweis durch vollständige Induktion

Es sei $n_0 \in \mathbb{Z}$ und für alle $n \geq n_0$ eine Aussage $A(n)$ gegeben. Um die Richtigkeit aller $A(n)$ zu beweisen, kann man wie folgt vorgehen:

Schritt 1: Man zeigt, dass $A(n_0)$ richtig ist (Induktionsanfang)

Schritt 2: Man nimmt an, $A(n)$ sei richtig für eine Zahl $n \geq n_0$ und zeigt, dass daraus die Richtigkeit von $A(n+1)$ folgt (Induktionsschritt)

Dann ist die Aussage $A(n)$ für alle $n \geq n_0$ richtig.

Beispiele

(i) **Aussage:** $A(n) : \sum_{i=1}^n i = \frac{n(n+1)}{2}$ für alle $n \geq 1$.

Beweis: (durch vollständige Induktion)

Induktionsanfang: $n = 1$

$\sum_{i=1}^1 i = 1$ und $\frac{1(1+1)}{2} = 1$, also gilt $A(1)$.

Induktionsschritt: $n \rightarrow n+1$

Angenommen es gilt

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

Dann folgt

$$\begin{aligned}\sum_{i=1}^{n+1} i &= \sum_{i=1}^n i + n + 1 = \frac{n(n+1)}{2} + n + 1 \\ &= \frac{1}{2}(n^2 + n + 2n + 2) = \frac{(n+1)(n+2)}{2}.\end{aligned}$$

Also ist $A(n+1)$ richtig.

Nach dem Beweisprinzip der vollständigen Induktion ist damit die Aussage $A(n)$ richtig für alle $n \geq 1$.

(ii) Die zu beweisenden Aussagen müssen nicht unbedingt Gleichungsform haben:

Bernoullische Ungleichung

Es sei $x \in \mathbb{R}, x \geq -1$. Dann gilt

$$(1+x)^n \geq 1+nx \text{ für alle } n \in \mathbb{N}.$$

Beweis:

Induktionsanfang: $n = 1$ gilt offensichtlich.

Induktionsschritt: $n \rightarrow n+1$

Angenommen es gilt

$$A(n) : \quad (1+x)^n \geq 1+nx.$$

Dann folgt

$$\begin{aligned}(1+x)^{n+1} &= \underbrace{(1+x)^n}_{\geq (1+nx)} \underbrace{(1+x)}_{\geq 0} \geq (1+nx)(1+x) \\ &= 1+nx+x+\underbrace{nx^2}_{\geq 0} \geq 1+(n+1)x.\end{aligned}$$

Also ist auch $A(n+1)$ richtig.

Nach dem Beweisprinzip der vollständigen Induktion ist damit die Bernoullische Ungleichung bewiesen.

Kapitel 2

Vektorrechnung

Es sei $n \in \mathbb{N}$. Die Menge aller Punkte der Form

$$P = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = [x_1, x_2, \dots, x_n]^T$$

mit $x_i \in \mathbb{R}$ wird mit \mathbb{R}^n bezeichnet. Die x_i heißen **Koordinaten** des Punktes P .

Spezialfälle

Für $n = 1$ erhalten wir die Zahlengerade $\mathbb{R}^1 = \mathbb{R}$, für $n = 2$ die Ebene

$$\mathbb{R}^2 = \left\{ \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} : x_1, x_2 \in \mathbb{R} \right\}$$

und für $n = 3$ den Raum

$$\mathbb{R}^3 = \left\{ \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} : x_1, x_2, x_3 \in \mathbb{R} \right\}.$$

Rechnen mit Punkten

Es seien

$$P = [x_1, \dots, x_n]^T, \quad Q = [y_1, \dots, y_n]^T \in \mathbb{R}^n$$

Addition, Subtraktion

$$P + Q = \begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{bmatrix}, \quad P - Q = \begin{bmatrix} x_1 - y_1 \\ x_2 - y_2 \\ \vdots \\ x_n - y_n \end{bmatrix}$$

Skalarmultiplikation

$$\alpha P = \begin{bmatrix} \alpha x_1 \\ \alpha x_2 \\ \vdots \\ \alpha x_n \end{bmatrix}, \quad \alpha \in \mathbb{R}$$

Insbesondere ist $1P = P$, $(-1)P = -P$ und $0P = 0 = [0, \dots, 0]^T$ (= **Nullpunkt**).

Distributivgesetze:

$$(\alpha + \beta)P = \alpha P + \beta P, \quad \alpha(P + Q) = \alpha P + \alpha Q, \quad \alpha, \beta \in \mathbb{R}$$

Bemerkung. Die Addition und Subtraktion von Vektoren in \mathbb{R}^2 entspricht genau der Addition und Subtraktion komplexer Zahlen.

2.1 Vektoren in \mathbb{R}^n

Für zwei Punkte $P, Q \in \mathbb{R}^n$ heißt die Differenz

$$\vec{x} = P - Q$$

der **Vektor** mit **Anfangspunkt** Q und **Endpunkt** P .

Verschiedene Paare von Anfangs- und Endpunkten können denselben Vektor definieren, etwa

$$\vec{x} = \begin{bmatrix} 5 \\ 5 \end{bmatrix} - \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 7 \\ 4 \end{bmatrix} - \begin{bmatrix} 4 \\ 2 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}.$$

Die Lage eines Vektors im Raum ist also nicht eindeutig bestimmt, sondern kann durch Parallelverschiebung beliebig verändert werden.

Wählt man den Nullpunkt 0 als Anfangspunkt, ist der zugehörige Endpunkt P durch den Vektor \vec{x} eindeutig bestimmt. Man nennt dann \vec{x} den **Ortsvektor von P**. Wir können also durch

$$\vec{x} = P - 0$$

die Menge aller Vektoren im \mathbb{R}^n mit der Menge aller Punkte im \mathbb{R}^n identifizieren. Für $P = 0$ erhält man als zugehörigen Ortsvektor speziell den **Nullvektor** $\vec{0} = [0, \dots, 0]^T$.

Durch diese Identifikation übertragen sich Addition und Skalarmultiplikation von Punkten im \mathbb{R}^n auf Vektoren.

Norm eines Vektors

Die **euklidische Norm** des Vektors $\vec{x} = [x_1, \dots, x_n]^T$ ist definiert durch

$$\|\vec{x}\| := \sqrt{x_1^2 + \dots + x_n^2} = \sqrt{\sum_{i=1}^n x_i^2}.$$

$\|\vec{x}\|$ ist also der Abstand zwischen Anfangs- und Endpunkt des Vektors \vec{x} , also seine Länge.

Eigenschaften der Norm

(i) (Positiv-Definitheit) $\|\vec{x}\| \geq 0$ und
 $\|\vec{x}\| = 0$ genau dann wenn $\vec{x} = \vec{0}$.

(ii) (Homogenität)
 $\|\alpha\vec{x}\| = |\alpha| \|\vec{x}\|$ für $\alpha \in \mathbb{R}$.

(iii) (Dreiecksungleichung)
 $\|\vec{x} + \vec{y}\| \leq \|\vec{x}\| + \|\vec{y}\|$

Übung: Aus der Dreiecksungleichung folgt: $|\|\vec{x}\| - \|\vec{y}\|| \leq \|\vec{x} - \vec{y}\|$.

Normierung Ist $\vec{x} \neq \vec{0}$, so erhält man durch

$$\vec{x}_0 := \frac{1}{\|\vec{x}\|} \vec{x}$$

einen Vektor der Länge 1 der in dieselbe Richtung wie \vec{x} zeigt.

Vektoren der Länge 1 heißen **Einheitsvektoren**. Spezielle Einheitsvektoren im \mathbb{R}^n sind

$$e_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ \vdots \\ 0 \end{bmatrix}, e_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \dots, e_n = \begin{bmatrix} 0 \\ \vdots \\ \vdots \\ 0 \\ 1 \end{bmatrix}.$$

Weitere Operationen auf Vektoren

1) **Skalarprodukt** Für $\vec{x} = [x_1, \dots, x_n]^T$, $\vec{y} = [y_1, \dots, y_n]^T \in \mathbb{R}^n$ ist das **Skalarprodukt** (auch **inneres Produkt** genannt) definiert durch

$$\langle \vec{x}, \vec{y} \rangle := x_1 y_1 + \dots + x_n y_n = \sum_{i=1}^n x_i y_i.$$

Eigenschaften des Skalarproduktes

(i) (Symmetrie) $\langle \vec{x}, \vec{y} \rangle = \langle \vec{y}, \vec{x} \rangle$

(ii) (Linearität)

$$\begin{aligned} \alpha \langle \vec{x}, \vec{y} \rangle &= \langle \alpha \vec{x}, \vec{y} \rangle = \langle \vec{x}, \alpha \vec{y} \rangle \quad \text{für } \alpha \in \mathbb{R} \\ \langle \vec{x} + \vec{y}, \vec{z} \rangle &= \langle \vec{x}, \vec{z} \rangle + \langle \vec{y}, \vec{z} \rangle \\ \langle \vec{x}, \vec{y} + \vec{z} \rangle &= \langle \vec{x}, \vec{y} \rangle + \langle \vec{x}, \vec{z} \rangle \end{aligned}$$

(iii) $\langle \vec{x}, \vec{x} \rangle = \|\vec{x}\|^2$, $\langle \vec{x}, \vec{y} \rangle = \frac{1}{4} (\|\vec{x} + \vec{y}\|^2 - \|\vec{x} - \vec{y}\|^2)$

Geometrische Interpretation des Skalarproduktes

$$\langle \vec{x}, \vec{y} \rangle = \|\vec{x}\| \cdot \|\vec{y}\| \cdot \cos \varphi$$

wobei φ den Winkel zwischen \vec{x} und \vec{y} bezeichnet.

Insbesondere gilt für Vektoren $\vec{x} \neq \vec{0}$ und $\vec{y} \neq \vec{0}$:

$$\langle \vec{x}, \vec{y} \rangle = 0 \quad \Leftrightarrow \quad \cos \varphi = 0 \quad \Leftrightarrow \quad \varphi = \pm \frac{\pi}{2}$$

d.h., genau dann wenn \vec{x} und \vec{y} senkrecht aufeinander stehen. Hierfür schreibt man auch $\vec{x} \perp \vec{y}$. Noch eine andere Sprechweise für dasselbe ist: \vec{x} ist orthogonal zu \vec{y} .

Cauchy-Schwarz Ungleichung

$$\langle \vec{x}, \vec{y} \rangle \leq \|\vec{x}\| \cdot \|\vec{y}\|$$

Beweis:

$$|\langle \vec{x}, \vec{y} \rangle| = \|\vec{x}\| \cdot \|\vec{y}\| \cdot \underbrace{|\cos \varphi|}_{\leq 1} \leq \|\vec{x}\| \cdot \|\vec{y}\|.$$

2) Vektorprodukt im \mathbb{R}^3

Für Vektoren $\vec{x} = [x_1, x_2, x_3]^T$, $\vec{y} = [y_1, y_2, y_3]^T \in \mathbb{R}^3$ ist das **Vektorprodukt** (oder auch das **äußere Produkt**) definiert durch

$$\vec{x} \times \vec{y} = \begin{bmatrix} x_2 y_3 - x_3 y_2 \\ x_3 y_1 - x_1 y_3 \\ x_1 y_2 - x_2 y_1 \end{bmatrix}.$$

Eigenschaften des Vektorprodukts

(i) (Antisymmetrie)

$$\vec{x} \times \vec{y} = -\vec{y} \times \vec{x}$$

Daraus folgt insbesondere $\vec{x} \times \vec{x} = \vec{0}$.

(ii) (Linearität)

$$\begin{aligned} \alpha(\vec{x} \times \vec{y}) &= (\alpha\vec{x}) \times \vec{y} = \vec{x} \times (\alpha\vec{y}) \quad \text{für } \alpha \in \mathbb{R} \\ (\vec{x} + \vec{y}) \times \vec{z} &= \vec{x} \times \vec{z} + \vec{y} \times \vec{z} \\ \vec{x} \times (\vec{y} + \vec{z}) &= \vec{x} \times \vec{y} + \vec{x} \times \vec{z} \end{aligned}$$

(iii) (Orthogonalität)

$$\begin{aligned} \langle \vec{x} \times \vec{y}, \vec{x} \rangle &= 0 \\ \langle \vec{x} \times \vec{y}, \vec{y} \rangle &= 0 \end{aligned}$$

(iv) $\|\vec{x} \times \vec{y}\| = \|\vec{x}\| \cdot \|\vec{y}\| \cdot |\sin \varphi|$, wobei φ der Winkel zwischen \vec{x} und \vec{y}

Insbesondere ist also $\|\vec{x} \times \vec{y}\|$ die Fläche des von den beiden Vektoren \vec{x} und \vec{y} aufgespannten Parallelogramms.

Die drei Vektoren $\vec{x}, \vec{y}, \vec{x} \times \vec{y}$ bilden in dieser Reihenfolge ein **Rechtssystem**, d.h. es gilt die **rechte-Hand-Regel**: Zeigt der Daumen der rechten Hand in Richtung \vec{x} und der Zeigefinger in Richtung \vec{y} , so zeigt der Mittelfinger in Richtung $\vec{x} \times \vec{y}$.

3) Spatprodukt im \mathbb{R}^3

Für $\vec{x}, \vec{y}, \vec{z} \in \mathbb{R}^3$ heißt

$$\langle \vec{x}, \vec{y} \times \vec{z} \rangle$$

das **Spatprodukt** von \vec{x}, \vec{y} und \vec{z} .

In Koordinatenschreibweise gilt:

$$\langle \vec{x}, \vec{y} \times \vec{z} \rangle = x_1(y_2z_3 - y_3z_2) + x_2(y_3z_1 - y_1z_3) + x_3(y_1z_2 - z_1y_2)$$

$|\langle \vec{x}, \vec{y} \times \vec{z} \rangle|$ ist das Volumen des von \vec{x}, \vec{y} und \vec{z} aufgespannten Spats (oder auch Parallelepipeds).

Linearkombination

Es seien $\vec{p}_1, \dots, \vec{p}_m$ Vektoren in \mathbb{R}^n und $\lambda_1, \dots, \lambda_m$ reelle Zahlen. Dann nennt man den Vektor

$$\vec{x} = \lambda_1 \vec{p}_1 + \dots + \lambda_m \vec{p}_m = \sum_{i=1}^m \lambda_i \vec{p}_i$$

eine **Linearkombination** der Vektoren $\vec{p}_1, \dots, \vec{p}_m$ mit Koeffizienten $\lambda_1, \dots, \lambda_m$.

Offensichtlich gilt

$$0 \cdot \vec{p}_1 + \dots + 0 \cdot \vec{p}_m = \vec{0}.$$

Diese spezielle Linearkombination des Nullvektors $\vec{0}$ heißt **triviale Linearkombination** der Vektoren $\vec{p}_1, \dots, \vec{p}_m$.

Die Vektoren $\vec{p}_1, \dots, \vec{p}_m$ heißen **linear unabhängig**, wenn gilt

$$\sum_{i=1}^m \lambda_i \vec{p}_i = \vec{0} \quad \Leftrightarrow \quad \lambda_1 = \dots = \lambda_m = 0.$$

Es gilt also: Die Vektoren $\vec{p}_1, \dots, \vec{p}_m$ sind genau dann linear unabhängig, wenn sich der Nullvektor nur als triviale Linearkombination der $\vec{p}_1, \dots, \vec{p}_m$ darstellen lässt.

Sind die Vektoren $\vec{p}_1, \dots, \vec{p}_m$ nicht linear unabhängig, so heißen sie linear abhängig. In diesem Fall gibt es also $\lambda_1, \dots, \lambda_m$ nicht alle 0 mit $\lambda_1 \vec{p}_1 + \dots + \lambda_m \vec{p}_m = \vec{0}$. Ist etwa $\lambda_1 \neq 0$, so folgt durch Umformung

$$\vec{p}_1 = -\frac{\lambda_2}{\lambda_1} \vec{p}_2 - \frac{\lambda_3}{\lambda_1} \vec{p}_3 - \dots - \frac{\lambda_m}{\lambda_1} \vec{p}_m.$$

Der Vektor \vec{p}_1 lässt sich also in diesem Falle aus den übrigen Vektoren $\vec{p}_2, \dots, \vec{p}_m$ linear kombinieren.

Insbesondere gilt:

- Zwei Vektoren \vec{p}_1, \vec{p}_2 sind genau dann linear abhängig, wenn sie auf einer Geraden liegen.
- Drei Vektoren $\vec{p}_1, \vec{p}_2, \vec{p}_3$ sind genau dann linear abhängig, wenn sie in einer Ebene liegen.

Beispiel

Die Einheitsvektoren $\vec{e}_1, \dots, \vec{e}_n$ sind linear unabhängig, denn für jeden Vektor $\vec{x} = [x_1, \dots, x_n]^T$ gilt

$$\vec{x} = \sum_{i=1}^n x_i \vec{e}_i$$

und damit $\sum_{i=1}^n x_i \vec{e}_i = 0$ genau dann wenn $x_1 = \dots = x_n = 0$.

2.2 Geraden im \mathbb{R}^n

Im Folgenden machen wir zumeist Gebrauch von der oben diskutierten Identifikation von Punkten und Vektoren.

Es sei $\vec{p}, \vec{r} \in \mathbb{R}^n$, $\vec{r} \neq \vec{0}$. Die Menge aller Punkte der Form

$$g : \vec{x} = \vec{p} + \lambda \vec{r}, \quad \lambda \in \mathbb{R} \tag{2.2.1}$$

beschreibt eine **Gerade** in \mathbb{R}^n . Die Darstellung (2.2.1) heißt **parametrisierte Darstellung**:

- \vec{p} heißt **Aufpunkt**,
- \vec{r} **Richtungsvektor**,
- λ **Parameter** der Geraden g .

Beispiel 2.2.1 Gegeben sei die Geradengleichung

$$g : y = 2x + 1 \quad \text{im } \mathbb{R}^2.$$

Wir wollen eine parametrisierte Darstellung von g bestimmen.

Dazu bestimmen wir zwei Punkte auf der Geraden, etwa

$$P = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad \text{und} \quad Q = \begin{bmatrix} 1 \\ 3 \end{bmatrix}.$$

Der Ortsvektor \vec{p} zu P liefert dann einen Aufpunkt, $\vec{r} = Q - P$ den Richtungsvektor. Also ist

$$g : \begin{bmatrix} 0 \\ 1 \end{bmatrix} + \lambda \begin{bmatrix} 1 \\ 2 \end{bmatrix} \quad \lambda \in \mathbb{R}$$

eine parametrisierte Darstellung von g .

Für $\lambda = -\frac{1}{2}$ erhält man z.B. den Punkt $\begin{bmatrix} -\frac{1}{2} \\ 0 \end{bmatrix}$, für $\lambda = 3$ den Punkt $\begin{bmatrix} 3 \\ 7 \end{bmatrix}$.

Abstand Punkt-Gerade

Der Abstand vom Punkt \vec{q} zur Geraden $g : \vec{x} = \vec{p} + \lambda\vec{r}$ ist definiert als

$$d(\vec{q}, g) := \min_{\lambda \in \mathbb{R}} \|\vec{x} - \vec{q}\|.$$

Dies ist der kleinste Abstand, den ein Punkt auf der Geraden von \vec{q} haben kann.

Ist $\vec{x}^* = \vec{p} + \lambda^*\vec{r}$ ein Punkt, für den dieses Minimum angenommen wird, so gilt dass der Verbindungsvektor $\vec{x}^* - \vec{q}$ zum Punkt \vec{q} senkrecht zum Richtungsvektor \vec{r} der Geraden ist, also

$$0 = \langle \vec{x}^* - \vec{q}, \vec{r} \rangle = \langle \vec{p} + \lambda^*\vec{r} - \vec{q}, \vec{r} \rangle.$$

Auflösen nach λ^* ergibt

$$\lambda^* = \frac{\langle \vec{q} - \vec{p}, \vec{r} \rangle}{\langle \vec{r}, \vec{r} \rangle}$$

also

$$\vec{x}^* = \vec{p} + \frac{\langle \vec{q} - \vec{p}, \vec{r} \rangle}{\langle \vec{r}, \vec{r} \rangle} \cdot \vec{r},$$

und damit

$$d(\vec{q}, g) = \|\vec{x}^* - \vec{q}\| = \left(\|\vec{p} - \vec{q}\|^2 - \frac{\langle \vec{p} - \vec{q}, \vec{r} \rangle^2}{\langle \vec{r}, \vec{r} \rangle} \right)^{\frac{1}{2}}.$$

Beispiel

Für $g : \vec{x} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} + \lambda \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ und $\vec{q} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ ist

$$\lambda^* = \frac{\left\langle \begin{bmatrix} 2 \\ 1 \end{bmatrix} - \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \end{bmatrix} \right\rangle}{\left\langle \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \end{bmatrix} \right\rangle} = \frac{2}{5},$$

und

$$\vec{x}^* = \begin{bmatrix} 0 \\ 1 \end{bmatrix} + \frac{2}{5} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \frac{1}{5} \begin{bmatrix} 2 \\ 9 \end{bmatrix},$$

sowie

$$d(\vec{q}, g) = \|\vec{x}^* - \vec{q}\| = \left\| \frac{1}{5} \begin{bmatrix} 2 \\ 9 \end{bmatrix} - \frac{1}{5} \begin{bmatrix} 10 \\ 5 \end{bmatrix} \right\| = \frac{4}{\sqrt{5}}.$$

Abstand Gerade-Gerade

Für zwei Geraden

$$g_1 : \vec{x}_1 = \vec{p}_1 + \lambda_1 \vec{r}_1, \quad \lambda_1 \in \mathbb{R}$$

$$g_2 : \vec{x}_2 = \vec{p}_2 + \lambda_2 \vec{r}_2, \quad \lambda_2 \in \mathbb{R}$$

ist der Abstand definiert durch

$$d(g_1, g_2) := \min_{\lambda_1, \lambda_2 \in \mathbb{R}} \|\vec{x}_1 - \vec{x}_2\|.$$

Dies ist der kleinste Abstand, den zwei Punkte auf den Geraden voneinander haben können.

Sind

$$\begin{aligned}\vec{x}_1^* &= \vec{p}_1 + \lambda_1^* \vec{r}_1 \\ \vec{x}_2^* &= \vec{p}_2 + \lambda_2^* \vec{r}_2\end{aligned}$$

zwei Punkte auf den Geraden mit minimalem Abstand, so steht der Verbindungsvektor

$$\vec{x}_1^* - \vec{x}_2^*$$

senkrecht auf den Richtungsvektoren \vec{r}_1 und \vec{r}_2 :

$$0 = \langle \vec{x}_1^* - \vec{x}_2^*, \vec{r}_1 \rangle = \langle \vec{x}_1^* - \vec{x}_2^*, \vec{r}_2 \rangle.$$

Dies liefert die beiden linearen Gleichungen

$$\begin{aligned}0 &= \langle \vec{p}_1 - \vec{p}_2, \vec{r}_1 \rangle + \lambda_1^* \langle \vec{r}_1, \vec{r}_1 \rangle - \lambda_2^* \langle \vec{r}_2, \vec{r}_1 \rangle \\ 0 &= \langle \vec{p}_1 - \vec{p}_2, \vec{r}_2 \rangle + \lambda_1^* \langle \vec{r}_1, \vec{r}_2 \rangle - \lambda_2^* \langle \vec{r}_2, \vec{r}_2 \rangle\end{aligned}$$

die man mit den Methoden des nächsten Kapitels über lineare Gleichungssysteme löst.

Im Falle $n = 3$ gibt es einen **weiteren Lösungsweg**:

Das Vektorprodukt $\vec{n} = \vec{r}_1 \times \vec{r}_2$ steht ebenfalls senkrecht auf den Richtungsvektoren \vec{r}_1 und \vec{r}_2 . Ist $\vec{n} \neq 0$, so müssen \vec{n} und $\vec{x}_1^* - \vec{x}_2^*$ in dieselbe oder in die entgegengesetzte Richtung zeigen, d.h. es gilt

$$\vec{x}_1^* - \vec{x}_2^* = \mu \vec{n} \quad \text{für ein } \mu \in \mathbb{R}.$$

Nimmt man das Skalarprodukt dieser Gleichung mit \vec{n} , dann erhält man

$$\underbrace{\langle \vec{x}_1^* - \vec{x}_2^*, \vec{n} \rangle}_{= \langle \vec{p}_1 - \vec{p}_2, \vec{n} \rangle} = \mu \langle \vec{n}, \vec{n} \rangle$$

denn $\langle \vec{r}_1, \vec{n} \rangle = \langle \vec{r}_2, \vec{n} \rangle = 0$. Auflösen nach μ ergibt

$$\mu = \frac{\langle \vec{p}_1 - \vec{p}_2, \vec{n} \rangle}{\langle \vec{n}, \vec{n} \rangle}$$

und damit

$$d(g_1, g_2) = \|\vec{x}_1^* - \vec{x}_2^*\| = |\mu| \|\vec{n}\| = \frac{|\langle \vec{p}_1 - \vec{p}_2, \vec{n} \rangle|}{\|\vec{n}\|}.$$

Geraden in impliziter Form

Ist $g : \vec{x} = \vec{p} + \lambda \vec{r}$ eine Gerade im \mathbb{R}^2 , so gibt es genau eine hierzu senkrechte Richtung.

Ist $\vec{r} = \begin{bmatrix} r_1 \\ r_2 \end{bmatrix}$, so ist $\vec{n} = \begin{bmatrix} -r_2 \\ r_1 \end{bmatrix}$ ein Vektor, der senkrecht zu g steht. Einen solchen Vektor nennt man **Normalenvektor zu g** .

Bildet man in der Geradengleichung $\vec{x} = \vec{p} + \lambda \vec{r}$ auf beiden Seiten das Skalarprodukt mit \vec{n} , so erhält man die **implizite Darstellung**

$$g : \langle \vec{x}, \vec{n} \rangle = \langle \vec{p}, \vec{n} \rangle. \quad (2.2.2)$$

Die Gerade g ist also die Menge aller Punkte $\vec{x} \in \mathbb{R}^2$, die diese Gleichung erfüllen.

Mit $\vec{x} = \begin{bmatrix} x \\ y \end{bmatrix}$, $\vec{n} = \begin{bmatrix} a \\ b \end{bmatrix}$ und $c = \langle \vec{p}, \vec{n} \rangle$ kann man Gleichung (2.2.2) in der Form

$$g : ax + by = c, \quad a, b \text{ nicht beide } 0$$

schreiben. Ist $b \neq 0$, so erhält man hieraus durch Auflösen nach y die vertraute Geradengleichung

$$y = -\frac{a}{b}x + \frac{c}{b}.$$

Hessesche Normalform

Eine Gerade g im \mathbb{R}^2 ist eindeutig bestimmt durch ihren Abstand d zum Ursprung und den Winkel β , den das vom Ursprung aus auf die Gerade gefällte Lot mit der positiven x -Achse bildet.

Ist $d = 0$, also g eine Gerade durch den Ursprung, so ist β nur bis auf Vielfache von π eindeutig bestimmt.

Die **Hessesche Normalform** der Geraden lautet dann

$$g : \cos \beta \cdot x + \sin \beta \cdot y = d.$$

Umrechnung: implizite Form \rightarrow Hessesche Normalform

Ist $g : ax + by = c$ eine Gerade in impliziter Form, so lautet die zugehörige Hessesche Normalform

$$\begin{aligned} \frac{a}{\sqrt{a^2 + b^2}}x + \frac{b}{\sqrt{a^2 + b^2}}y &= \frac{c}{\sqrt{a^2 + b^2}} && \text{falls } c \geq 0 \\ -\frac{a}{\sqrt{a^2 + b^2}}x - \frac{b}{\sqrt{a^2 + b^2}}y &= -\frac{c}{\sqrt{a^2 + b^2}} && \text{falls } c < 0. \end{aligned}$$

Ist g in Hessescher Normalform, lässt sich der Abstand eines Punktes $\vec{q} = \begin{bmatrix} q_1 \\ q_2 \end{bmatrix}$ zu g besonders einfach ausrechnen:

$$d(\vec{q}, g) = |\cos \beta \cdot q_1 + \sin \beta \cdot q_2 - d|.$$

Beispiel

Zu $g : \vec{x} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} + \lambda \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ lautet die implizite Form (mit $\vec{n} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$)

$$g : -2x + y = 1$$

und damit die Hessesche Normalform

$$g : -\frac{2}{\sqrt{5}}x + \frac{1}{\sqrt{5}}y = \frac{1}{\sqrt{5}}.$$

Für den Abstand zum Ursprung erhalten wir also $\frac{1}{\sqrt{5}}$. Für den Abstand zum Punkt $\vec{q} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ wie zuvor

$$d(\vec{q}, g) = \left| -\frac{4}{\sqrt{5}} + \frac{1}{\sqrt{5}} - \frac{1}{\sqrt{5}} \right| = \frac{4}{\sqrt{5}}.$$

2.3 Ebenen in \mathbb{R}^3

Es seien $\vec{p}, \vec{r}_1, \vec{r}_2$ Vektoren in \mathbb{R}^3 und $\vec{n} := \vec{r}_1 \times \vec{r}_2 \neq \vec{0}$. Die Menge aller Vektoren der Form

$$E : \vec{x} = \vec{p} + \lambda_1 \vec{r}_1 + \lambda_2 \vec{r}_2, \quad \lambda_1, \lambda_2 \in \mathbb{R} \quad (2.3.1)$$

beschreibt eine Ebene in \mathbb{R}^3 in **parametrisierter Form**:

- \vec{p} heißt **Aufpunkt**,
- \vec{r}_1, \vec{r}_2 **Richtungsvektoren**,
- λ_1, λ_2 **Parameter** der Ebene E .

\vec{n} heißt **Normalenvektor** der Ebene E . Er steht senkrecht auf der Ebene und zwar so, dass die Vektoren $\vec{r}_1, \vec{r}_2, \vec{n}$ ein Rechtssystem bilden.

Abstand Punkt-Ebene

Ist \vec{q} ein Punkt, so ist der Abstand zur Ebene $E : \vec{p} + \lambda_1 \vec{r}_1 + \lambda_2 \vec{r}_2$ definiert als

$$\begin{aligned} d(\vec{q}, E) &:= \min_{\lambda_1, \lambda_2 \in \mathbb{R}} \|\vec{x} - \vec{q}\| \\ &= \text{kleinster Abstand zwischen } \vec{q} \text{ und einem Punkt der Ebene.} \end{aligned}$$

Ist $\vec{x}^* = \vec{p} + \lambda_1^* \vec{r}_1 + \lambda_2^* \vec{r}_2$ ein Punkt mit minimalem Abstand, so steht der Verbindungsvektor $\vec{x}^* - \vec{q}$ senkrecht auf den beiden Richtungsvektoren \vec{r}_1 und \vec{r}_2 . Daher gibt es ein $\mu \in \mathbb{R}$ mit

$$\vec{x}^* - \vec{q} = \mu \vec{n}.$$

Multipliziert man diese Gleichung skalar mit \vec{n} , dann erhält man

$$\begin{aligned} \underbrace{\langle \vec{x}^* - \vec{q}, \vec{n} \rangle}_{=} &= \mu \langle \vec{n}, \vec{n} \rangle \\ &= \langle \vec{p} - \vec{q}, \vec{n} \rangle \end{aligned}$$

denn $\langle \vec{r}_1, \vec{n} \rangle = \langle \vec{r}_2, \vec{n} \rangle = 0$. Man erhält

$$\mu = \frac{\langle \vec{p} - \vec{q}, \vec{n} \rangle}{\langle \vec{n}, \vec{n} \rangle}$$

und damit für den Abstand

$$d(\vec{q}, E) = |\mu| \|\vec{n}\| = \frac{|\langle \vec{p} - \vec{q}, \vec{n} \rangle|}{\|\vec{n}\|}.$$

Beachte die Analogie zur Berechnung des Abstandes Gerade-Gerade.

Beispiel

Es sei

$$E : \vec{x} = \underbrace{\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}}_{=\vec{p}} + \lambda_1 \underbrace{\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}}_{=\vec{r}_1} + \lambda_2 \underbrace{\begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}}_{=\vec{r}_2}, \quad \lambda_1, \lambda_2 \in \mathbb{R}.$$

Normalenvektor ist

$$\vec{n} = \vec{r}_1 \times \vec{r}_2 = \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix}, \quad \|\vec{n}\| = \sqrt{6}.$$

Für den Abstand der Ebene zum Ursprung ergibt sich

$$d(\vec{0}, E) = \frac{|\langle \vec{p}, \vec{n} \rangle|}{\|\vec{n}\|} = \frac{3}{\sqrt{6}}.$$

Hierbei ist

$$\mu = \frac{3}{6} = \frac{1}{2} \quad \text{und} \quad \vec{x}^* = \frac{1}{2} \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} + \frac{3}{2} \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} - \frac{3}{2} \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}.$$

Ebenen in impliziter Form

Es sei $E : \vec{x} = \vec{p} + \lambda_1 \vec{r}_1 + \lambda_2 \vec{r}_2$ eine Ebene und \vec{n} der Normalenvektor. Dann erhält man durch Multiplikation von $\vec{x} = \vec{p} + \lambda_1 \vec{r}_1 + \lambda_2 \vec{r}_2$ mit \vec{n} die **implizite Darstellung**

$$E : \langle \vec{x}, \vec{n} \rangle = \langle \vec{p}, \vec{n} \rangle. \quad (2.3.2)$$

Die Ebene E ist also die Menge aller Punkte $\vec{x} \in \mathbb{R}^3$, die diese Gleichung erfüllen.

Mit $\vec{n} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$ und $d = \langle \vec{p}, \vec{n} \rangle$ kann man Gleichung (2.3.2) in der Form

$$E : ax + by + cz = d, \quad a, b, c \text{ nicht alle } 0 \quad (2.3.3)$$

schreiben.

Hessesche Normalform

Die spezielle implizite Form

$$E : ax + by + cz = d$$

mit $d \geq 0$ und $a^2 + b^2 + c^2 = 1$ heißt **Hessesche Normalform** der Ebene. In diesem Falle ist d gleich dem Abstand der Ebene zum Ursprung.

Ist $E : ax + by + cz = d$ eine Ebene in impliziter Form, so gelangt man zur Hesseschen Normalform, indem man den Normalenvektor normiert und die Gleichung gegebenenfalls noch mit -1 multipliziert:

$$\begin{aligned} \frac{a}{\sqrt{a^2 + b^2 + c^2}}x + \frac{b}{\sqrt{a^2 + b^2 + c^2}}y + \frac{c}{\sqrt{a^2 + b^2 + c^2}}z &= \frac{d}{\sqrt{a^2 + b^2 + c^2}} && \text{falls } d \geq 0 \\ -\frac{a}{\sqrt{a^2 + b^2 + c^2}}x - \frac{b}{\sqrt{a^2 + b^2 + c^2}}y - \frac{c}{\sqrt{a^2 + b^2 + c^2}}z &= -\frac{d}{\sqrt{a^2 + b^2 + c^2}} && \text{falls } d < 0. \end{aligned}$$

Der Abstand eines Punktes $\vec{q} = \begin{bmatrix} q_1 \\ q_2 \\ q_3 \end{bmatrix}$ zu E lässt sich wiederum sehr einfach ausrechnen:

$$d(\vec{q}, E) = |a \cdot q_1 + b \cdot q_2 + c \cdot q_3 - d|.$$

Beispiel

Gegeben sei die Ebene $E : \vec{x} = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} + \lambda_1 \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} + \lambda_2 \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}$. Durch Multiplikation mit dem Normalenvektor $\vec{n} = \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix}$ erhält man die implizite Form

$$E : -x + y + 2z = 3.$$

Die Hessesche Normalform hierzu lautet also

$$-\frac{1}{\sqrt{6}}x + \frac{1}{\sqrt{6}}y + \frac{2}{\sqrt{6}}z = \frac{3}{\sqrt{6}}.$$

Der Abstand zum Ursprung beträgt also $\frac{3}{\sqrt{6}}$.

Schnitt Ebene-Gerade

Zur Berechnung des Schnittpunkts \vec{x}^* einer Ebene E mit einer Geraden g in \mathbb{R}^3 verwendet man am besten

- für die Ebene die implizite Form $E : \langle \vec{x}, \vec{n} \rangle = d$
- für die Gerade die parametrisierte Form $g : \vec{x} = \vec{p} + \lambda \vec{r}$.

Der Schnittpunkt \vec{x}^* erfüllt beide Gleichungen, also

$$\vec{x}^* = \vec{p} + \lambda^* \vec{r} \quad \text{und} \quad \langle \vec{x}^*, \vec{n} \rangle = d,$$

also

$$\langle \vec{p}, \vec{n} \rangle + \lambda^* \langle \vec{r}, \vec{n} \rangle = d. \quad (2.3.4)$$

1. Fall: $\langle \vec{r}, \vec{n} \rangle \neq 0$. In diesem Falle gibt es genau einen Schnittpunkt, den man wie folgt berechnen kann:

$$\lambda^* = \frac{d - \langle \vec{p}, \vec{n} \rangle}{\langle \vec{r}, \vec{n} \rangle}$$

und der Schnittpunkt ist durch

$$\vec{x}^* = \vec{p} + \frac{d - \langle \vec{p}, \vec{n} \rangle}{\langle \vec{r}, \vec{n} \rangle} \vec{r}$$

gegeben.

2. Fall: $\langle \vec{r}, \vec{n} \rangle = 0$

1. Unterfall: $\langle \vec{p}, \vec{n} \rangle = d$.

Dann ist Gleichung (2.3.4) für alle $\lambda^* \in \mathbb{R}$ erfüllt. Jeder Punkt der Geraden liegt also zugleich in der Ebenen, also verläuft die Gerade ganz in E .

2. Unterfall: $\langle \vec{p}, \vec{n} \rangle \neq d$.

In diesem Falle ist Gleichung (2.3.4) für kein $\lambda^* \in \mathbb{R}$ erfüllt. Damit liegt kein Punkt der Geraden in E . Es gibt also keinen Schnittpunkt. Dies bedeutet, dass die Gerade parallel zur Ebene ist und nicht in dieser liegt.

2.4 Lineare Unterräume, Basis und Dimension

Geraden und Ebenen, die $\vec{0}$ enthalten sind Beispiele für lineare Teilräume:

Definition: Eine nichtleere Teilmenge $U \subset \mathbb{R}^n$ heißt **(linearer) Teilraum** (oder linearer Unterraum) falls gilt

$$\vec{x}, \vec{y} \in U; \alpha, \beta \in \mathbb{R} \implies \alpha \vec{x} + \beta \vec{y} \in U.$$

Mit anderen Worten: Teilräume U sind nichtleere Teilmengen, die abgeschlossen sind unter Vektoraddition und Skalarmultiplikation.

Für Vektoren $\vec{p}_1, \vec{p}_2, \dots, \vec{p}_k \in \mathbb{R}^n$ heißt die Menge

$$\text{Lin}(\vec{p}_1, \vec{p}_2, \dots, \vec{p}_k) := \{\lambda_1 \vec{p}_1 + \dots + \lambda_k \vec{p}_k : \lambda_1, \dots, \lambda_k \in \mathbb{R}\}$$

aller Linearkombinationen von $\vec{p}_1, \dots, \vec{p}_k$ die **lineare Hülle** der Vektoren $\vec{p}_1, \dots, \vec{p}_k$.

Man prüft leicht nach, dass $\text{Lin}(\vec{p}_1, \dots, \vec{p}_k)$ ein linearer Teilraum ist.

Beispiel

(i) Für $\vec{p} \neq \vec{0}$ ist $\text{Lin}(\vec{p}) = \{\lambda \vec{p} : \lambda \in \mathbb{R}\}$ die Gerade durch $\vec{0}$ mit Richtungsvektor \vec{p} .

(ii) Für \vec{p}_1, \vec{p}_2 linear unabhängig ist

$$\text{Lin}(\vec{p}_1, \vec{p}_2) = \{\lambda_1 \vec{p}_1 + \lambda_2 \vec{p}_2 : \lambda_1, \lambda_2 \in \mathbb{R}\}$$

die Ebene durch $\vec{0}$ mit Richtungsvektoren \vec{p}_1, \vec{p}_2 .

Definition: Es sei $U \subset \mathbb{R}^n$ ein linearer Teilraum. Eine Menge $\{\vec{p}_1, \dots, \vec{p}_k\} \subset U$ heißt **Erzeugendensystem** von U , falls $\text{Lin}(\vec{p}_1, \dots, \vec{p}_k) = U$.

Ein Erzeugendensystem $\{\vec{p}_1, \dots, \vec{p}_k\}$ heißt **Basis** von U , falls $\vec{p}_1, \dots, \vec{p}_k$ linear unabhängig sind. Der Basisaustauschsatz der linearen Algebra impliziert: Die Anzahl der Vektoren zweier Basen von U ist gleich, d.h. eine „Invariante“ des Unterraums U .

Die **Dimension** $\dim U$ des Unterraums U ist definiert als die Anzahl der Vektoren einer (beliebigen) Basis von U .

Beispiel

(i) $\dim(\mathbb{R}^n) = n$ (Basis: $\{\vec{e}_1, \dots, \vec{e}_n\}$)

(ii) $\vec{p}_1, \dots, \vec{p}_k$ linear unabhängig $\implies \dim(\text{Lin}(\vec{p}_1, \dots, \vec{p}_k)) = k$

Insbesondere ist die Dimension einer Geraden = 1, Ebenen = 2, usw.

Kapitel 3

Lineare Gleichungssysteme

Das Lösen linearer Gleichungssysteme ist eines der zentralen Themen dieser Vorlesung. Zum einen gibt es viele direkte Anwendungen. Andererseits bauen hierauf viele andere Verfahren auf.

3.1 Einige geometrisch motivierte Beispiele

Beispiel 3.1.1 Berechne den Schnittpunkt der drei Ebenen

$$E_1 : x + y + z = 1$$

$$E_2 : 4x + 4y + 3z = 5$$

$$E_3 : 2x + y + z = 2$$

Subtrahiert man das dreifache der ersten Zeile von der zweiten Zeile, so erhält man

$$x + y = 2.$$

Subtrahiert man die erste Zeile von der dritten Zeile, so erhält man

$$x = 1.$$

Einsetzen in die vorherige Gleichung liefert $y = 1$. Einsetzen der Werte für x und y in die erste Gleichung führt auf $z = -1$. Der Schnittpunkt der drei Ebenen ist also

$$\vec{x} = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}.$$

Beispiel 3.1.2 Berechnet werden soll in Beispiel 3.1.1 die Schnittmenge der beiden Ebenen E_1 und E_2 .

Wie zuvor erhält man zunächst die Bedingung

$$x + y = 2.$$

Da nun aber keine weitere Bedingung vorliegt, kann man einer der beiden Variablen einen **beliebigen Wert** zuordnen, etwa

$$y = t, \quad t \in \mathbb{R}.$$

Dann ergibt sich $x + t = 2$, also

$$x = 2 - t$$

und Einsetzen in die Gleichung für E_1 führt auf

$$(2 - t) + t + z = 1$$

und damit $z = -1$. Die Menge aller Schnittpunkte \vec{x} der beiden Ebenen E_1 und E_2 ist also gegeben durch

$$\vec{x} = \begin{bmatrix} 2 - t \\ t \\ -1 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ -1 \end{bmatrix} + t \cdot \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \quad t \in \mathbb{R}.$$

Die Lösungsmenge ist also eine Gerade.

Beispiel 3.1.3 Gesucht ist nun der Schnittpunkt der Ebenen E_1 , E_2 und E_4 , wobei

$$E_4 : 2x + 2y + 3z = -1.$$

Subtrahiert man die Gleichung für E_4 von der Gleichung für E_2 , so erhält man

$$2x + 2y = 6 \text{ und damit } x + y = 3.$$

Außerdem folgt aus den Gleichungen für E_1 und E_2 wie zuvor

$$x + y = 2.$$

Zieht man diese Gleichung von der vorherigen Gleichung ab, so erhält man den Widerspruch

$$0 = 1.$$

Es gibt also keinen Schnittpunkt.

3.2 Lineare Gleichungssysteme (LGS)

Ein System von Gleichungen der Form

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

wobei die Koeffizienten a_{ij} und die Werte b_i vorgegebene reelle Zahlen sind, heißt **lineares Gleichungssystem** mit m Gleichungen für den Vektor $\vec{x} = [x_1, x_2, \dots, x_n]^T$ der n **Unbekannten**.

Gesucht ist die Menge aller Vektoren \vec{x} , für die alle Gleichungen erfüllt sind. Diese Menge heißt **Lösungsmenge** des LGS.

Die Koeffizienten a_{ij} auf der linken Seite kann man zu einem rechteckigen Zahlenschema der Form

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

zusammenfassen. A heißt **Koeffizientenmatrix** des Gleichungssystems.

Ebenso lassen sich die Werte b_1, b_2, \dots, b_m auf der rechten Seite zu einem Vektor $\vec{b} = [b_1, \dots, b_m]^T$ zusammenfassen. \vec{b} heißt **Zielvektor** und man schreibt für das LGS kurz

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

Das LGS heißt

- **unterbestimmt**, falls $m < n$,
- **quadratisch**, falls $m = n$,
- **überbestimmt**, falls $m > n$.
- Beispiel 3.1.1 führt auf ein quadratisches LGS mit

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 4 & 4 & 3 \\ 2 & 1 & 1 \end{bmatrix} \quad \text{und} \quad \vec{b} = \begin{bmatrix} 1 \\ 5 \\ 2 \end{bmatrix},$$

- Beispiel 3.1.2 führt auf ein unterbestimmtes LGS mit

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 4 & 4 & 3 \end{bmatrix} \quad \text{und} \quad \vec{b} = \begin{bmatrix} 1 \\ 5 \end{bmatrix}.$$

3.3 Einschub: Matrizen

Eine $m \times n$ -Matrix ist ein Zahlenschema der Form

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

Die Einträge a_{ij} heißen **Komponenten** (oder **Elemente**) der Matrix. Das Element a_{ij} steht am Schnittpunkt der i -ten Zeile mit der j -ten Spalte.

Die Zeilen der Matrix

$$\vec{a}_1^T, \vec{a}_2^T, \dots, \vec{a}_m^T \text{ heißen } \mathbf{Zeilenvektoren}$$

und die Spalten

$$\vec{a}_{\cdot 1}, \vec{a}_{\cdot 2}, \dots, \vec{a}_{\cdot n} \text{ heißen } \mathbf{Spaltenvektoren}$$

Spezialfälle:

Eine $1 \times n$ -Matrix $[a_{11}, a_{12}, \dots, a_{1n}]$ ist ein Zeilenvektor

Eine $m \times 1$ -Matrix $\begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix}$ ist ein Spaltenvektor.

In diesem Sinne fassen wir Vektoren $\vec{x} \in \mathbb{R}^n$ in dieser Vorlesung stets als **Spaltenvektoren** auf!

Rechenoperationen**1) Matrixaddition**

Sind A, B $m \times n$ -Matrizen, so ist die Summenmatrix $A + B$ definiert durch

$$A + B := [a_{ij} + b_{ij}].$$

2) Skalarmultiplikation

Ist A eine $m \times n$ -Matrix und $\lambda \in \mathbb{R}$ ein Skalar, so ist $\lambda \cdot A$ definiert durch

$$\lambda \cdot A = [\lambda a_{ij}].$$

Insbesondere folgt: $1 \cdot A = A$, $(-1) \cdot A = -A$, $0 \cdot A = 0_{mn}$, wobei

$$0_{mn} = \begin{bmatrix} 0 & \dots & 0 \\ \vdots & & \vdots \\ 0 & \dots & 0 \end{bmatrix} \text{ die } m \times n\text{-Nullmatrix ist.}$$

Es gelten wieder die **Distributivgesetze**:

$$(\alpha + \beta)A = \alpha A + \beta A, \quad \alpha(A + B) = \alpha A + \alpha B$$

3) Matrixmultiplikation

Es sei A eine $m \times l$ -Matrix und B eine $l \times n$ -Matrix, d.h. die Anzahl der Spalten von A ist gleich der Anzahl der Zeilen von B .

Das Produkt $C = A \cdot B$ ist dann definiert durch

$$c_{ij} = \sum_{k=1}^l a_{ik} b_{kj} \quad \text{für } i = 1, \dots, m, j = 1, \dots, n, \text{ d.h.}$$

$$\begin{bmatrix} a_{11} & \dots & a_{1l} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{ml} \end{bmatrix} \cdot \begin{bmatrix} b_{11} & \dots & b_{1n} \\ \vdots & & \vdots \\ b_{l1} & \dots & b_{ln} \end{bmatrix} \stackrel{\text{„Zeile} \times \text{Spalte“}}{=} \begin{bmatrix} \sum_{k=1}^l a_{1k} b_{k1} & \dots & \sum_{k=1}^l a_{1k} b_{kn} \\ \vdots & & \vdots \\ \sum_{k=1}^l a_{mk} b_{k1} & \dots & \sum_{k=1}^l a_{mk} b_{kn} \end{bmatrix}$$

Das Element c_{ij} der Produktmatrix erhält man also als **Skalarprodukt** der **i -ten Zeile von A** mit der **j -ten Spalte von B** . Das Resultat C ist also eine $m \times n$ -Matrix.

Beispiele

$$(i) \quad A = \begin{bmatrix} 6 & 3 & -1 \\ 0 & 2 & 4 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 3 \\ 2 & -2 \\ 0 & 1 \end{bmatrix} \text{ ergibt } A \cdot B = \begin{bmatrix} 12 & 11 \\ 4 & 0 \end{bmatrix}$$

$$(ii) \quad \vec{x} = \begin{bmatrix} 1 \\ -3 \\ -4 \end{bmatrix}, \quad \vec{y} = \begin{bmatrix} -2 \\ 2 \end{bmatrix}. \text{ Dann gilt}$$

$$A\vec{x} = \begin{bmatrix} 1 \\ -22 \end{bmatrix}, \quad \vec{y}^T A = [-2, 2] \cdot \begin{bmatrix} 6 & 3 & -1 \\ 0 & 2 & 4 \end{bmatrix} = [-12, -2, 10]$$

$$\vec{x}^T B = [1, -3, -4] \cdot \begin{bmatrix} 1 & 3 \\ 2 & -2 \\ 0 & 1 \end{bmatrix} = [-5, 5].$$

$$(iii) \text{ Ist } A \text{ eine } m \times n\text{-Matrix, } \vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n, \text{ so ist}$$

$$A \cdot \vec{x} = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + \dots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \dots + a_{mn}x_n \end{bmatrix}$$

Mit anderen Worten: In einem LGS $A\vec{x} = \vec{b}$ ist die linke Seite gerade das Matrixprodukt aus Koeffizientenmatrix A mit dem Vektor \vec{x} der Unbekannten.

Rechenregeln der Matrixmultiplikation**1) Assoziativgesetz**

$$A(BC) = (AB)C$$

Da also die Reihenfolge der Berechnung des Matrixproduktes beliebig ist, lässt man die Klammern weg und schreibt nur ABC .

2) Distributivgesetz

$$A(B + C) = AB + AC \quad \text{und} \quad (A + B)C = AC + BC$$

3) Das Kommutativgesetz gilt nicht! Im Allgemeinen ist $AB \neq BA$.**Beispiel**

$$A = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$$

$$\implies AB = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad \text{aber } BA = \begin{bmatrix} -1 & 1 \\ -1 & 1 \end{bmatrix}$$

Dieses Beispiel zeigt auch zugleich:

4) Aus $AB = 0$ folgt **nicht notwendigerweise** $A = 0$ oder $B = 0$.

Der Rang einer Matrix

Der **Rang** $\text{rang } A$ einer Matrix A ist definiert als die Dimension der von den Spaltenvektoren $\vec{a}_1, \dots, \vec{a}_n$ aufgespannten linearen Hülle

$$\text{rang } A = \dim(\text{Lin}(\vec{a}_1, \dots, \vec{a}_n))$$

Beispiele

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \implies \text{Lin}(\vec{a}_1, \vec{a}_2, \vec{a}_3) = \text{Lin}(\vec{a}_1) \quad \text{also } \text{rang } A = 1$$

$$A = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 3 \end{bmatrix} \implies \text{Lin}(\vec{a}_1, \vec{a}_2, \vec{a}_3) = \text{Lin}(\vec{a}_1, \vec{a}_2) \quad \text{also } \text{rang } A = 2$$

allgemeiner: Ist A eine $m \times n$ -Matrix in gestaffelter Form, d.h.

$$A = \begin{bmatrix} \bullet & * & \dots & \dots & * & * & \dots & * \\ 0 & \bullet & * & \dots & * & * & \dots & * \\ 0 & 0 & \bullet & \ddots & \vdots & \vdots & & \vdots \\ \vdots & \vdots & \ddots & \ddots & * & * & \dots & * \\ \vdots & \vdots & & \ddots & \bullet & * & \dots & * \\ 0 & 0 & \dots & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & & & \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & \dots & 0 & 0 & \dots & 0 \end{bmatrix}$$

wobei gilt:

- alle mit \bullet markierten Einträge sind von Null verschieden
- alle mit $*$ markierten Einträge sind beliebig

so ist $\text{rang } A$ gleich der Anzahl der Zeilen mit \bullet .

Man kann zeigen, dass die Dimension der von den Zeilenvektoren von A aufgespannten linearen Hülle mit $\text{rang } A$ übereinstimmt.

3.4 Elementare Umformungen

Zurück zu den linearen Gleichungssystemen. Gegeben sei wieder ein LGS $A\vec{x} = \vec{b}$. Hierfür verwendet man zweckmäßigerweise folgendes Schema:

$$\begin{array}{c|cccc|c} & x_1 & x_2 & \cdots & x_n & \vec{b} \\ \hline \boxed{1} & a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ \boxed{2} & a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \boxed{m} & a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}$$

Unser Ziel ist nun, dieses LGS in ein einfach zu lösendes LGS zu überführen, ohne dabei die Lösungsmenge zu verändern. Die folgenden elementaren Umformungen sind dabei erlaubt, d.h. sie **verändern nicht die Lösungsmenge** des LGS:

I **Zeilenvertauschung** Zwei Zeilen dürfen vertauscht werden:

$$\boxed{i} \leftrightarrow \boxed{j}$$

II **Spaltenvertauschung** Zwei Spalten dürfen vertauscht werden:

$$x_i \leftrightarrow x_j$$

Dabei ist zu beachten, dass auch die Einträge in der Kopfzeile vertauscht werden.

III **Skalierung** Jede Gleichung darf mit einer beliebigen von Null verschiedenen Zahl multipliziert werden:

$$\boxed{i} \leftarrow \lambda \cdot \boxed{i} \quad \lambda \neq 0$$

IV **Addition** Zu jeder Gleichung darf ein **beliebiges Vielfaches** einer **anderen** Gleichung addiert werden:

$$\boxed{i} \leftarrow \boxed{i} + \lambda \cdot \boxed{j} \quad \lambda \in \mathbb{R}, i \neq j$$

Mit Hilfe dieser Elementarumformungen ist es möglich, ein beliebiges LGS in ein LGS in gestaffelter Form zu überführen, dessen Lösungsmenge man unmittelbar bestimmen kann.

Beispiel 3.4.1 Das Schema zum LGS aus Beispiel 3.1.1 hat die Form

$$\begin{array}{c|ccc|c} & x & y & z & \vec{b} \\ \hline \boxed{1} & 1 & 1 & 1 & 1 \\ \boxed{2} & 4 & 4 & 3 & 5 \\ \boxed{3} & 2 & 1 & 1 & 2 \end{array}$$

Durch Elementarumformungen vom Typ IV können die jeweils ersten Koeffizienten der zweiten und dritten Zeile zu Null gemacht werden. Dadurch wird die Variable x aus $\boxed{2}$ und $\boxed{3}$ eliminiert.

$$\begin{array}{c|ccc|c} & x & y & z & \vec{b} \\ \hline & \boxed{1} & & & \\ \boxed{2} - 4 \cdot \boxed{1} = \boxed{4} & 0 & 0 & -1 & 1 \\ \boxed{3} - 2 \cdot \boxed{1} = \boxed{5} & 0 & -1 & -1 & 0 \end{array}$$

Vertauscht man zweite und dritte Zeile, so erhält man das LGS

$$\begin{array}{c|ccc|c} & x & y & z & \vec{b} \\ \hline \boxed{1} & 1 & 1 & 1 & 1 \\ \boxed{5} & 0 & -1 & -1 & 0 \\ \boxed{4} & 0 & 0 & -1 & 1 \end{array}$$

Das Schema hat nun **gestaffelte Form** und kann nun **von unten nach oben** schrittweise aufgelöst werden:

$$\boxed{4} : -z = 1 \text{ also } z = -1$$

$$\boxed{5} : -y - z = 0 \Rightarrow y = 1$$

$$\boxed{1} : x + y + z = 1 \Rightarrow x + 1 - 1 = 1 \text{ also } x = 1$$

Die Lösung ist demnach $\vec{x} = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$.

Beispiel 3.4.2 Das Schema zum LGS aus Beispiel 3.1.2 hat die Form

$$\begin{array}{c|ccc|c} & x & y & z & \vec{b} \\ \hline \boxed{1} & 1 & 1 & 1 & 1 \\ \boxed{2} & 4 & 4 & 3 & 5 \end{array}$$

Wie vorhin erhält man

$$\begin{array}{c|ccc|c} & x & y & z & \vec{b} \\ \hline & \boxed{1} & & & 1 \\ \boxed{2} - 4 \cdot \boxed{1} = \boxed{4} & & & -1 & 1 \end{array}$$

Es ergibt sich

$$\boxed{4} : -z = 1 \text{ also } z = -1$$

$$\boxed{1} : x + y + z = 1 \Rightarrow x + y = 2$$

In der Zeile $\boxed{1}$ kann nun entweder der Wert von x oder der Wert von y frei gewählt werden. Setzt man wie in Beispiel 3.1.2 $y = t$, $t \in \mathbb{R}$, so folgt

$$x = 2 - t$$

und damit ist die Lösungsmenge eine Gerade

$$\vec{x} = \begin{bmatrix} 2 - t \\ t \\ -1 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ -1 \end{bmatrix} + t \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \quad t \in \mathbb{R}.$$

Beispiel 3.4.3 Beispiel 3.1.3 führt auf das Schema

$$\begin{array}{c|ccc|c} & x & y & z & \vec{b} \\ \hline \boxed{1} & 1 & 1 & 1 & 1 \\ \boxed{2} & 4 & 4 & 3 & 5 \\ \boxed{3} & 2 & 2 & 3 & -1 \end{array}$$

Typ IV und Typ II Umformungen überführen das LGS in die Form

$$\begin{array}{r|ccc|c}
 & x & y & z & \vec{b} \\
 \hline
 & & & & \\
 \boxed{1} & 1 & 1 & 1 & 1 \\
 \boxed{2} - 4 \cdot \boxed{1} = \boxed{4} & 0 & 0 & -1 & 1 \\
 \boxed{3} - 2 \cdot \boxed{1} = \boxed{5} & 0 & 0 & 1 & -3
 \end{array}$$

und

$$\begin{array}{r|ccc|c}
 & x & z & y & \vec{b} \\
 \hline
 & & & & \\
 \boxed{1} & 1 & 1 & 1 & 1 \\
 \boxed{4} & 0 & -1 & 0 & 1 \\
 \boxed{5} + \boxed{4} = \boxed{6} & 0 & 0 & 0 & -2
 \end{array}$$

Aus der letzten Zeile ergibt sich der Widerspruch

$$0x + 0y + 0z = -2.$$

Es existiert daher keine Lösung des LGS.

3.5 Gestaffelte Form eines LGS

Wie in den Beispielen zuvor gesehen, lässt sich die Lösung eines LGS einfach bestimmen, indem man es durch elementare Umformungen in gestaffelte Form überführt:

$$\begin{array}{cccccccc|c}
 \tilde{x}_1 & \tilde{x}_2 & \tilde{x}_3 & \cdots & \tilde{x}_r & \tilde{x}_{r+1} & \cdots & \tilde{x}_n & \vec{b} \\
 \bullet & * & \cdots & \cdots & * & * & \cdots & * & * \\
 0 & \bullet & * & \cdots & * & * & \cdots & * & * \\
 0 & 0 & \bullet & \ddots & \vdots & \vdots & & \vdots & \vdots \\
 \vdots & \vdots & \ddots & \ddots & * & * & \cdots & * & * \\
 \vdots & \vdots & & \ddots & \bullet & * & \cdots & * & * \\
 0 & 0 & \cdots & \cdots & 0 & 0 & \cdots & 0 & \times \\
 \vdots & \vdots & & & \vdots & \vdots & & \vdots & \vdots \\
 0 & 0 & \cdots & \cdots & 0 & 0 & \cdots & 0 & \times
 \end{array}$$

Dabei sind

- $\tilde{x}_1, \dots, \tilde{x}_n$ eine Umordnung der x_1, \dots, x_n , die durch Spaltenvertauschungen entsteht,
- alle mit \bullet markierten Einträge von Null verschieden,
- alle mit $*$ oder \times markierten Einträge beliebig.

Lösbarkeitsentscheidung

Das LGS besitzt

- keine Lösung, wenn nur ein einziger der mit \times markierten Einträge nicht Null ist.
- Lösungen, wenn alle mit \times markierten Einträge Null sind (oder überhaupt keine Nullzeilen existieren). In diesem Falle kann die Lösungsmenge wie folgt beschrieben werden:

Die Werte für $\tilde{x}_{r+1}, \dots, \tilde{x}_n$ können beliebig vorgegeben werden,

$$\tilde{x}_{r+1} = t_1, \dots, \tilde{x}_n = t_{n-r}, \quad t_1, \dots, t_{n-r} \in \mathbb{R}.$$

Danach können die restlichen Werte von $\tilde{x}_r, \tilde{x}_{r-1}, \dots, \tilde{x}_1$ durch Auflösen des LGS von unten nach oben bestimmt werden.

Als Spezialfall ergibt sich hieraus: Das LGS besitzt eine eindeutige Lösung, falls die prinzipielle Lösbarkeit gesichert ist (alle mit \times markierten Einträge sind Null) und keine Werte frei gewählt werden können, wenn also $r = n$ ist.

3.6 Der Gaußsche Algorithmus

Der Gaußsche Algorithmus (oder auch Gauß-Jordansches Eliminationsverfahren) ist ein Algorithmus zur Transformation eines LGS in gestaffelte Form mit Hilfe von Elementarumformungen:

- 1. Schritt:** Bestimme ein Element $a_{ij} \neq 0$ und überführe es durch Zeilenvertauschung und Spaltenvertauschung an die erste Position der ersten Zeile.
- 2. Schritt:** Eliminiere in der ersten Spalte die Einträge $a_{21}, a_{31}, \dots, a_{m1}$ durch Subtraktion des $\frac{a_{i1}}{a_{11}}$ -fachen der 1. Zeile von der i -ten Zeile:

$$\boxed{i} \leftarrow \boxed{i} - \frac{a_{i1}}{a_{11}} \cdot \boxed{1}.$$

Nach Abschluss des 2. Schrittes hat das Schema zum LGS die Form

$$\begin{array}{cccc|c} \tilde{x}_1 & \tilde{x}_2 & \cdots & \tilde{x}_n & \vec{b} \\ \bullet & * & \cdots & * & * \\ 0 & * & \cdots & * & * \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & * & \cdots & * & * \end{array}$$

Damit haben die erste Zeile und die erste Spalte die gewünschte Form. Sie werden im weiteren Verlauf des Algorithmus nicht mehr verändert. Nun wendet man das Verfahren auf das Teilschema S' an, das man aus dem ursprünglichen Schema durch Streichen der ersten Gleichung und der ersten Spalte erhält, und setzt die Umformungen solange fort, bis die gestaffelte Form erreicht ist.

Beispiel 3.6.1 Für einen reellen Parameter $\alpha \in \mathbb{R}$ sei das folgende LGS gegeben:

$$\begin{array}{c|cccc|c} & x_1 & x_2 & x_3 & x_4 & \vec{b} \\ \hline \boxed{1} & 1 & 2 & 2 & -1 & -1 \\ \boxed{2} & 0 & 0 & 1 & 0 & 5 \\ \boxed{3} & -2 & -2 & -3 & 0 & 3 \\ \boxed{4} & 2 & 4 & 4 & -2 & \alpha \end{array}$$

Elimination der Variablen x_1 in $\boxed{2}$ – $\boxed{4}$ ergibt

$$\begin{array}{c|cccc|c} & x_1 & x_2 & x_3 & x_4 & \vec{b} \\ \hline \boxed{1} & 1 & 2 & 2 & -1 & -1 \\ \boxed{2} & 0 & 0 & 1 & 0 & 5 \\ \boxed{5} & 0 & 2 & 1 & -2 & 1 \\ \boxed{6} & 0 & 0 & 0 & 0 & 2 + \alpha \end{array}$$

Vertauschung der Zeilen $\boxed{2}$ und $\boxed{5}$ ergibt

$$\begin{array}{c|cccc|c} & x_1 & x_2 & x_3 & x_4 & \vec{b} \\ \hline \boxed{1} & 1 & 2 & 2 & -1 & -1 \\ \boxed{5} & 0 & 2 & 1 & -2 & 1 \\ \boxed{2} & 0 & 0 & 1 & 0 & 5 \\ \boxed{6} & 0 & 0 & 0 & 0 & 2 + \alpha \end{array}$$

Damit ist die gestaffelte Form erreicht und man hat zwei Fälle zu unterscheiden:

1. **Fall** $\alpha \neq -2$: dann gibt es keine Lösung.
2. **Fall** $\alpha = -2$: dann gibt es einen Lösungsraum mit einem freien Parameter

$$x_4 = t_1, \quad t_1 \in \mathbb{R}$$

und Auflösen von unten nach oben ergibt die Lösung

$$\vec{x} = \begin{bmatrix} -7 \\ -2 \\ 5 \\ 0 \end{bmatrix} + t_1 \begin{bmatrix} -1 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \quad t_1 \in \mathbb{R}.$$

3.7 Homogene LGS

Ein LGS $A\vec{x} = \vec{b}$ heißt **homogen**, wenn der Zielvektor \vec{b} der Nullvektor ist, andernfalls heißt das LGS **inhomogen**.

Homogener Fall $A\vec{x} = \vec{0}$

- $\vec{0}$ ist stets eine Lösung, denn $A\vec{0} = \vec{0}$
- die Lösungsmenge ist ein **linearer Teilraum**. Er wird als **Kern von A** bezeichnet:

$$\text{kern } A := \{\vec{x} : A\vec{x} = \vec{0}\}.$$

Beweis: $\vec{x}, \vec{y} \in \text{kern } A$, $\alpha, \beta \in \mathbb{R}$ impliziert

$$A(\alpha\vec{x} + \beta\vec{y}) = \alpha \underbrace{A\vec{x}}_{=\vec{0}} + \beta \underbrace{A\vec{y}}_{=\vec{0}} = \vec{0}$$

und damit $\alpha\vec{x} + \beta\vec{y} \in \text{kern } A$.

Die Dimension $\dim \text{kern } A$ des Kerns von A ist an der zugehörigen gestaffelten Form abzulesen. Im homogenen Fall sind alle mit \times gekennzeichneten Einträge Null und damit die Werte für

$$\tilde{x}_{r+1}, \dots, \tilde{x}_n$$

frei wählbar. Man erhält eine Lösungsmenge mit $n - r$ freien Parametern, also ist

$$\dim \text{kern } A = n - r.$$

Da außerdem $\text{rang } A = r$ folgt die **Dimensionsformel**

$$\dim \text{kern } A + \text{rang } A = n.$$

Die Dimension des Kerns und der Rang der Matrix ergeben also zusammen die Spaltenzahl der Matrix.

Beispiel 3.7.1 Im Falle des Beispiels 3.6.1 erhält man für das homogene LGS $A\vec{x} = \vec{0}$ die gestaffelte Form

$$\begin{array}{c|cccc|c} & x_1 & x_2 & x_3 & x_4 & \vec{b} \\ \hline \boxed{1} & 1 & 2 & 2 & -1 & 0 \\ \boxed{5} & 0 & 2 & 1 & -2 & 0 \\ \boxed{2} & 0 & 0 & 1 & 0 & 0 \\ \boxed{6} & 0 & 0 & 0 & 0 & 0 \end{array}.$$

Es ist also $r = 3$ und damit

$$\text{rang } A = 3 \quad \text{und} \quad \dim \text{kern } A = 1.$$

Der Kern von A ist gegeben durch

$$\text{kern } A = \left\{ t_1 \begin{bmatrix} -1 \\ 1 \\ 0 \\ 1 \end{bmatrix}, t_1 \in \mathbb{R} \right\}.$$

3.8 Inhomogene LGS

Es sei

- \vec{x}_s Lösung des LGS $A\vec{x} = \vec{b}$
- $\vec{x}_h \in \text{kern } A$ eine Lösung des zugehörigen homogenen Systems $A\vec{x} = \vec{0}$

dann ist $\vec{x} = \vec{x}_s + \vec{x}_h$ ebenfalls Lösung des LGS $A\vec{x} = \vec{b}$, denn

$$A\vec{x} = A(\vec{x}_s + \vec{x}_h) = \underbrace{A\vec{x}_s}_{=\vec{b}} + \underbrace{A\vec{x}_h}_{=\vec{0}} = \vec{b}.$$

Umgekehrt: Sind \vec{x} und \vec{x}_s Lösungen von $A\vec{x} = \vec{b}$, dann ist $\vec{x}_h = \vec{x} - \vec{x}_s \in \text{kern } A$, denn

$$A(\vec{x} - \vec{x}_s) = A\vec{x} - A\vec{x}_s = \vec{b} - \vec{b} = \vec{0}.$$

Man kann somit jede Lösung des LGS $A\vec{x} = \vec{b}$ in der Form

$$\vec{x} = \vec{x}_s + \vec{x}_h, \quad \vec{x}_h \in \text{kern } A$$

darstellen.

Mit anderen Worten: Die allgemeine Lösung eines inhomogenen Systems erhält man als Summe einer speziellen Lösung dieses Systems und der allgemeinen Lösung des zugehörigen homogenen Systems. Dieser grundlegende Sachverhalt wird als **Superpositionsprinzip** bezeichnet.

Beispiel 3.8.1 Im Falle des Beispiels 3.6.1 löst der Vektor

$$\vec{x}_s = \begin{bmatrix} -5 \\ -4 \\ 5 \\ -2 \end{bmatrix}$$

das inhomogene LGS

$$\begin{bmatrix} 1 & 2 & 2 & -1 \\ 0 & 0 & 1 & 0 \\ -2 & -2 & -3 & 0 \\ 2 & 4 & 4 & -2 \end{bmatrix} \vec{x} = \begin{bmatrix} -1 \\ 5 \\ 3 \\ -2 \end{bmatrix}.$$

Zusammen mit dem in Beispiel 3.7.1 bestimmten Kern von A erhält man gemäß dem Superpositionsprinzip die Lösungsmenge des LGS

$$\vec{x} = \begin{bmatrix} -5 \\ -4 \\ 5 \\ -2 \end{bmatrix} + t_1 \begin{bmatrix} -1 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \quad t_1 \in \mathbb{R}.$$

Im Beispiel 3.6.1 hatten wir andererseits für die Lösungsmenge die Darstellung

$$\vec{x} = \begin{bmatrix} -7 \\ -2 \\ 5 \\ 0 \end{bmatrix} + t_1 \begin{bmatrix} -1 \\ 1 \\ 0 \\ 1 \end{bmatrix}, t_1 \in \mathbb{R}$$

erhalten. Die hierdurch beschriebene Menge ist aber dieselbe, wie man durch Ersetzen von t_1 durch $t_1 - 2$ sofort einsieht.

3.9 Determinanten

Im Falle eines quadratischen LGS $A\vec{x} = \vec{b}$ kann man sehr einfach die eindeutige Lösbarkeit entscheiden. Dies geschieht mit Hilfe der **Determinante** der Koeffizientenmatrix A , die wie folgt definiert ist:

- Ist $A = [a_{11}]$ eine 1×1 -Matrix, so ist

$$\det A := a_{11}.$$

- Ist A eine $n \times n$ -Matrix, so ist

$$\det A := \sum_{j=1}^n (-1)^{i+j} a_{ij} \det A_{ij}. \quad (\text{Entwicklung nach der } i\text{-ten Zeile})$$

Hierbei ist i ein beliebiger Zeilenindex und A_{ij} diejenige $(n-1) \times (n-1)$ -Streichmatrix, die man aus A durch Streichen der i -ten Zeile und der j -ten Spalte erhält.

Die Definition der Determinante $\det A$ einer $n \times n$ -Matrix A (und damit auch ihre Berechnung) wird also mit Hilfe der Zeilenentwicklung auf die Definition der Determinanten von $(n-1) \times (n-1)$ -Matrizen zurückgeführt. Durch wiederholte Anwendung der Zeilenentwicklung führt man die Definition schließlich zurück auf die Definition von Determinanten von 1×1 -Matrizen.

Das Vorzeichen $(-1)^{i+j}$ in der Entwicklungsformel folgt einem Schachbrettmuster

$$\begin{bmatrix} + & - & + & - & + & \cdots \\ - & + & - & + & - & \cdots \\ + & - & + & - & + & \cdots \\ - & + & - & + & - & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

Alternativ kann man die Determinante auch **nach der j -ten Spalte entwickeln**:

$$\det A = \sum_{i=1}^n (-1)^{i+j} a_{ij} \det A_{ij}.$$

Hierbei ist j ein beliebiger Spaltenindex.

Spezialfälle

- $n = 2$:

$$\det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = a_{11} \cdot a_{22} - a_{21} \cdot a_{12}$$

(= "Hauptdiagonale" - "Nebendiagonale")

- $n = 3$:

$$\det \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\ - a_{31}a_{22}a_{13} - a_{32}a_{23}a_{11} - a_{33}a_{21}a_{12}$$

(= "Hauptdiagonalen" - "Nebendiagonalen")
(Regel von Sarrus)

Beispiele

$$\det \begin{bmatrix} 1 & 3 \\ 4 & 5 \end{bmatrix} = 1 \cdot 5 - 4 \cdot 3 = 5 - 12 = -7.$$

$$\det \begin{bmatrix} 1 & 2 & 4 \\ 2 & 1 & 3 \\ 3 & 1 & 1 \end{bmatrix} = 1 \cdot 1 \cdot 1 + 2 \cdot 3 \cdot 3 + 4 \cdot 2 \cdot 1 - 3 \cdot 1 \cdot 4 - 1 \cdot 3 \cdot 1 - 1 \cdot 2 \cdot 2 \\ = 1 + 18 + 8 - 12 - 3 - 4 = 8.$$

Entwicklung nach der 4. Zeile liefert in folgendem Beispiel

$$\det \begin{bmatrix} 2 & 3 & 0 & 1 \\ 3 & 2 & 1 & 0 \\ 0 & 1 & 0 & 3 \\ 3 & 2 & 0 & 0 \end{bmatrix} = (-1)^{4+1} \cdot 3 \cdot \det \begin{bmatrix} 3 & 0 & 1 \\ 2 & 1 & 0 \\ 1 & 0 & 3 \end{bmatrix} + (-1)^{4+2} \cdot 2 \cdot \det \begin{bmatrix} 2 & 0 & 1 \\ 3 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix} \\ = -3 \cdot 8 + 2 \cdot 6 = -24 + 12 = -12.$$

Ist A eine **obere Dreiecksmatrix**, d.h.

$$A = \begin{bmatrix} a_{11} & * & \cdots & \cdots & * \\ 0 & a_{22} & * & \cdots & * \\ 0 & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & * \\ 0 & 0 & \cdots & 0 & a_{nn} \end{bmatrix}$$

oder **untere Dreiecksmatrix**, d.h.

$$A = \begin{bmatrix} a_{11} & 0 & 0 & \cdots & 0 \\ * & a_{22} & 0 & \cdots & 0 \\ \vdots & * & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ * & * & \cdots & * & a_{nn} \end{bmatrix},$$

so gilt

$$\det A = a_{11} \cdot a_{22} \cdot \dots \cdot a_{nn} = \text{Produkt der Diagonalelemente.}$$

Eine **effiziente** Methode zur **praktischen** Berechnung der Determinante liefert der Gaußsche Algorithmus:

Eliminiere mit Hilfe der Elementarumformung IV alle bis auf ein Element der 1. Spalte. **Hierbei ändert sich** $\det A$ **nicht!** Entwickle danach $\det A$ nach der 1. Spalte: Hierbei hat man dann nur noch eine einzige Determinante einer $(n-1) \times (n-1)$ -Matrix zu berechnen, usw.

Durch Vertauschen zweier Zeilen (Elementarumformung I) oder Spalten (Elementarumformung II) **ändert sich das Vorzeichen** der Determinante.

Aus diesen Überlegungen folgt insbesondere: Sollten alle Elemente einer beliebigen Zeile oder Spalte von A gleich 0 sein, so ist $\det A = 0$.

Rechenregeln Es seien A und B zwei $n \times n$ -Matrizen. Dann gilt:

1) $\det(\lambda A) = \lambda^n \cdot \det A$, $\lambda \in \mathbb{R}$.

Beachte hierbei den Exponenten von λ !

2) (Determinantenmultiplikationssatz) $\det(AB) = \det A \cdot \det B$

3) (Determinante der Transponierten) $\det(A^T) = \det A$

wobei A^T die **Transponierte** von A bezeichnet, d.h.

$$A^T = \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{nn} \end{bmatrix}$$

Die Zeilen von A werden also zu den Spalten von A^T , die Spalten von A zu den Zeilen von A^T .

Der Zusammenhang zwischen quadratischen LGS und Determinanten ist in folgendem Satz enthalten:

Satz: Ein quadratisches LGS $A\vec{x} = \vec{b}$ ist genau dann eindeutig lösbar, wenn $\det A \neq 0$.

Weiter gilt

$$\det A \neq 0 \Leftrightarrow \dim \text{kern } A = 0 \Leftrightarrow \text{rang } A = n.$$

Zusatz: Ist $\det A = 0$ so kann es keine oder unendlich viele Lösungen geben.

3.10 Matrix-Gleichungssysteme

Ein LGS der Form

$$AX = B$$

heißt **Matrix-Gleichungssystem**. Dabei sind A eine $m \times n$ -Matrix, B eine $m \times k$ -Matrix und gesucht ist eine $n \times k$ -Matrix X in den Unbekannten x_{ij} . Das zugehörige Schema ist von der Form

$$\begin{array}{c|cccc|ccc} & \vec{x}_1 & \vec{x}_2 & \cdots & \vec{x}_n & \vec{b}_1 & \cdots & \vec{b}_k \\ \hline \boxed{1} & a_{11} & a_{12} & \cdots & a_{1n} & b_{11} & \cdots & b_{1k} \\ \vdots & \vdots & \vdots & & \vdots & \vdots & & \vdots \\ \boxed{m} & a_{m1} & a_{m2} & \cdots & a_{mn} & b_{m1} & \cdots & b_{mk} \end{array}$$

Durch Elementarumformungen wird dieses Schema wieder in gestaffelte Form überführt und durch Auflösen von unten nach oben für jeden Spaltenvektor \vec{b}_j gelöst.

Die Kriterien für die Lösbarkeit sind analog zu den Lösbarkeitskriterien von LGS (siehe Abschnitt 3.4). Insbesondere ist ein Matrix-Gleichungssystem mit quadratischer Koeffizientenmatrix A genau dann eindeutig lösbar, wenn $\det A \neq 0$.

Beispiel 3.10.1 Gegeben sei das Matrix-Gleichungssystem $AX = B$ mit

$$A = \begin{bmatrix} 1 & 0 & 1 \\ 3 & 1 & 0 \\ 2 & 0 & 4 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 5 \\ 1 & 3 \\ 2 & 12 \end{bmatrix}.$$

Die Lösung X ist also eine 3×2 -Matrix. Das zugehörige Schema hat die Form

$$\begin{array}{c|ccc|cc} & \vec{x}_1 & \vec{x}_2 & \vec{x}_3 & \vec{b}_1 & \vec{b}_2 \\ \hline \boxed{1} & 1 & 0 & 1 & 1 & 5 \\ \boxed{2} & 3 & 1 & 0 & 1 & 3 \\ \boxed{3} & 2 & 0 & 4 & 2 & 12 \end{array}$$

Der Gauß-Algorithmus liefert die gestaffelte Form

$$\begin{array}{c|ccc|cc} & \vec{x}_1 & \vec{x}_2 & \vec{x}_3 & \vec{b}_1 & \vec{b}_2 \\ \hline \boxed{1} & 1 & 0 & 1 & 1 & 5 \\ \boxed{4} & 0 & 1 & -3 & -2 & -12 \\ \boxed{5} & 0 & 0 & 2 & 0 & 2 \end{array}$$

Damit ergibt sich

$$\boxed{5} : 2\vec{x}_3 = [0, 2] \Rightarrow \vec{x}_3 = [0, 1]$$

$$\boxed{4} : \vec{x}_2 - 3\vec{x}_3 = [-2, -12] \Rightarrow \vec{x}_2 = [0, 3] + [-2, -12] = [-2, -9]$$

$$\boxed{1} : \vec{x}_1 = [1, 5] - [0, 1] = [1, 4]$$

und schließlich die Lösung

$$X = \begin{bmatrix} 1 & 4 \\ -2 & -9 \\ 0 & 1 \end{bmatrix}.$$

Inverse Matrix

Eine besondere Rolle spielt das Matrix-Gleichungssystem

$$AX = E$$

wobei A eine $n \times n$ -Matrix ist und

$$E = E_n := \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} = [\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n].$$

E_n heißt **Einheitsmatrix** und man prüft leicht nach, dass für jede $n \times n$ -Matrix B gilt

$$E_n \cdot B = B = B \cdot E_n.$$

Im Falle $\det A \neq 0$ ist die Lösung X des Matrix-Gleichungssystems $AX = E$ eindeutig bestimmt. Die eindeutig bestimmte Lösung X wird auch mit A^{-1} bezeichnet und heißt **Inverse** von A . Ist $\det A = 0$, so besitzt A **keine Inverse**.

Eigenschaften der inversen Matrix A^{-1}

- 1) $AA^{-1} = E = A^{-1}A$
- 2) $(A^{-1})^{-1} = A$
- 3) $(AB)^{-1} = B^{-1}A^{-1}$ (Reihenfolge kehrt sich um!)
- 4) $(A^T)^{-1} = (A^{-1})^T$
- 5) $\det A^{-1} = \frac{1}{\det A}$, denn $1 = \det E = \det(AA^{-1}) = \det A \cdot \det(A^{-1})$.

Ist $AX = B$ ein beliebiges Gleichungssystem mit $\det A \neq 0$, so folgt durch Multiplikation des LGS von links mit A^{-1}

$$X = A^{-1}B.$$

Mit anderen Worten: Kenntnis der Inversen A^{-1} liefert durch Multiplikation mit B die Lösung des Matrix-Gleichungssystems $AX = B$.

Die Berechnung der Inversen lohnt sich immer dann, wenn wiederholt Gleichungssysteme mit derselben Matrix A und verschiedenen rechten Seiten gelöst werden müssen.

Beispiel

Für $n = 2$ gilt

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad \det A = ad - bc.$$

Ist $\det A \neq 0$, so folgt

$$A^{-1} = \frac{1}{\det A} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

Für $A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$ ergibt sich etwa $A^{-1} = \frac{1}{3} \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$.

Beispiel 3.10.2 Es sei nun A wie in Beispiel 3.10.1. Es gilt $\det A = 2$, also ist A invertierbar. Der Gauß-Algorithmus liefert für das Matrixgleichungssystem mit Schema

$$\begin{array}{c|ccc|ccc} & \vec{x}_1. & \vec{x}_2. & \vec{x}_3. & \vec{e}_1 & \vec{e}_2 & \vec{e}_3 \\ \hline \boxed{1} & 1 & 0 & 1 & 1 & 0 & 0 \\ \boxed{2} & 3 & 1 & 0 & 0 & 1 & 0 \\ \boxed{3} & 2 & 0 & 4 & 0 & 0 & 1 \end{array}$$

die gestaffelte Form

$$\begin{array}{c|ccc|ccc} & \vec{x}_1. & \vec{x}_2. & \vec{x}_3. & \vec{e}_1 & \vec{e}_2 & \vec{e}_3 \\ \hline & \boxed{1} & & & 1 & 0 & 0 \\ \boxed{2} - 3 \cdot \boxed{1} = \boxed{4} & & 1 & -3 & -3 & 1 & 0 \\ \boxed{3} - 2 \cdot \boxed{1} = \boxed{5} & & 0 & 2 & -2 & 0 & 1 \end{array}$$

Wiederum durch Auflösen und Einsetzen von unten nach oben ergeben sich hieraus der Reihe nach die Zeilen $\vec{x}_3.$, $\vec{x}_2.$ und $\vec{x}_1.$ der gesuchten Matrix

$$A^{-1} = X = \frac{1}{2} \begin{bmatrix} 4 & 0 & -1 \\ -12 & 2 & 3 \\ -2 & 0 & 1 \end{bmatrix}.$$

Nachdem man so die Inverse zu A bestimmt hat, kann man dann auch die Lösung

$$AX = \begin{bmatrix} 1 & 5 \\ 1 & 3 \\ 2 & 12 \end{bmatrix}$$

aus Beispiel 3.10.1 direkt berechnen:

$$X = A^{-1} \begin{bmatrix} 1 & 5 \\ 1 & 3 \\ 2 & 12 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 4 & 0 & -1 \\ -12 & 2 & 3 \\ -2 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 5 \\ 1 & 3 \\ 2 & 12 \end{bmatrix} = \begin{bmatrix} 1 & 4 \\ -2 & -9 \\ 0 & 1 \end{bmatrix}$$

Orthogonale Matrizen

Eine $n \times n$ -Matrix A heißt **orthogonal**, wenn

$$A^T A = E.$$

Für orthogonale Matrizen A gilt:

- 1) $A^{-1} = A^T$
- 2) $AA^T = E$
- 3) A^T ist orthogonal.
- 4) Ist B orthogonal, so ist auch AB orthogonal.
- 5) $\det A = \pm 1$, denn $1 = \det E = \det(A^T A) = \det(A^T) \cdot \det A = (\det A)^2$.

Eine Matrix ist genau dann orthogonal, wenn ihre Spaltenvektoren ein **Orthonormalsystem** bilden, d.h., wenn die Spaltenvektoren Länge 1 haben und paarweise orthogonal sind, d.h.

$$\langle \vec{a}_i, \vec{a}_j \rangle = \begin{cases} 1 & \text{für } i = j \\ 0 & \text{für } i \neq j. \end{cases}$$

Beispiel

Für beliebige Winkel φ ist die Matrix

$$D(\varphi) = \begin{bmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{bmatrix}$$

orthogonal.

3.11 Matrizen mit anderen Koeffizienten

Statt Vektoren und Matrizen mit reellen Koeffizienten kann man auch (allgemeiner) komplexe oder (spezieller) rationale Koeffizienten betrachten. Es gelten jeweils die genannten Rechenregeln, lineare Gleichungssysteme betrachten, und der Gauß-Algorithmus berechnet Lösungen.

Bei den drei Zahlbereichen \mathbb{Q} , \mathbb{R} und \mathbb{C} handelt es sich jeweils um *Körper*.

Kapitel 4

Lineare Abbildungen

4.1 Allgemeines

Definition Eine Abbildung $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ heißt **linear**, falls gilt

- 1) $f(\vec{x} + \vec{y}) = f(\vec{x}) + f(\vec{y})$ für alle $\vec{x}, \vec{y} \in \mathbb{R}^n$,
- 2) $f(\lambda\vec{x}) = \lambda f(\vec{x})$ für alle $\vec{x} \in \mathbb{R}^n, \lambda \in \mathbb{R}$.

Im Falle $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ nennen wir f auch eine lineare Abbildung **in** \mathbb{R}^n .

Eine lineare Abbildung ist also verträglich mit Vektoraddition und Skalarmultiplikation.

Beachte: Ist $W \subset \mathbb{R}^n$ ein linearer Teilraum, so ist auch das Bild $f(W) \subset \mathbb{R}^m$ wieder ein linearer Teilraum, d.h. eine lineare Abbildung überführt Geraden in Geraden (oder Punkte), Ebenen in Ebenen (oder Geraden oder Punkte), usw.

Es sei A eine $m \times n$ -Matrix. Dann ist die Abbildung

$$T_A : \mathbb{R}^n \rightarrow \mathbb{R}^m, \vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \mapsto A\vec{x} = \begin{bmatrix} \sum_{j=1}^n a_{1j}x_j \\ \vdots \\ \sum_{j=1}^n a_{mj}x_j \end{bmatrix}$$

offensichtlich linear.

Insbesondere gilt

$$T_A(\vec{e}_i) = A\vec{e}_i = \begin{bmatrix} a_{1i} \\ \vdots \\ a_{mi} \end{bmatrix} = \vec{a}_{.i}.$$

Mit anderen Worten: In den Spalten der Matrix A stehen die Bilder der Basisvektoren $\vec{e}_1, \dots, \vec{e}_n$ unter der Abbildung T_A .

Beispiel 4.1.1 ($n = m = 2$)

$$D(\varphi) = \begin{bmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{bmatrix}, \quad \varphi \in [0, 2\pi[$$

$T_{D(\varphi)}$ beschreibt dann eine Drehung des \mathbb{R}^2 um den Winkel φ .

Umgekehrt gilt: Ist f linear und

$$A = [f(\vec{e}_1), \dots, f(\vec{e}_n)]$$

so folgt für einen beliebigen Vektor

$$\vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \sum_{i=1}^n x_i \vec{e}_i$$

aufgrund der Linearität

$$f(\vec{x}) = f\left(\sum_{i=1}^n x_i \vec{e}_i\right) = \sum_{i=1}^n x_i f(\vec{e}_i) = A\vec{x}.$$

In diesem Sinne ist f die zur Matrix A gehörende lineare Abbildung T_A . A heißt **Abbildungsmatrix von f** .

Verkettung

Sind $f : \mathbb{R}^n \rightarrow \mathbb{R}^l$ und $g : \mathbb{R}^l \rightarrow \mathbb{R}^m$ linear, so ist auch die verkettete Abbildung

$$g \circ f : \mathbb{R}^n \rightarrow \mathbb{R}^m \\ \vec{x} \mapsto g(f(\vec{x}))$$

linear.

Ist A Abbildungsmatrix von f , B Abbildungsmatrix von g , so folgt

BA ist die Abbildungsmatrix von $g \circ f$

Die Matrizenmultiplikation der Abbildungsmatrizen entspricht somit der Verkettung der zugehörigen linearen Abbildungen, denn

$$g \circ f(\vec{x}) = g(\underbrace{f(\vec{x})}_{=A\vec{x}}) = B(A\vec{x}) = (BA)\vec{x}.$$

Insbesondere gilt: Die lineare Abbildung $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ ist genau dann **bijektiv** (also **umkehrbar**), wenn die zugehörige Abbildungsmatrix A **invertierbar** ist. Die **inverse Matrix** A^{-1} ist gerade die **Abbildungsmatrix der Umkehrabbildung** f^{-1} .

Beispiel 4.1.2 Drehungen $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ um den Winkel φ sind umkehrbar. Die Umkehrabbildung f^{-1} ist eine Drehung um den Winkel $-\varphi$.

Die Abbildungsmatrix der Umkehrabbildung ist also

$$D(-\varphi) = \begin{bmatrix} \cos(-\varphi) & -\sin(-\varphi) \\ \sin(-\varphi) & \cos(-\varphi) \end{bmatrix}.$$

In der Tat gilt

$$D(\varphi)D(-\varphi) = E \quad \text{also} \quad D(-\varphi) = D(\varphi)^{-1}.$$

Determinante und lineare Abbildungen

Ist $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ linear und A die zugehörige Abbildungsmatrix, so ist

$$|\det A| = \text{Volumen des von } f(\vec{e}_1), \dots, f(\vec{e}_n) \text{ aufgespannten Parallelepipeds}$$

$|\det A|$ beschreibt also die **Volumenänderung** des Einheitswürfels unter der zugehörigen linearen Abbildung f .

4.2 Spezielle lineare Abbildungen in \mathbb{R}^n

Drehungen im \mathbb{R}^n

Lineare Abbildungen $f(\vec{x}) = A\vec{x}$ im \mathbb{R}^n heißen **Drehungen**, falls A orthogonal und $\det A = 1$.

Drehungen sind normerhaltend, d.h. $\|A\vec{x}\| = \|\vec{x}\|$, denn

$$\|A\vec{x}\|^2 = (A\vec{x})^T \cdot A\vec{x} = \vec{x}^T (A^T A) \vec{x} = \vec{x}^T E \vec{x} = \vec{x}^T \cdot \vec{x} = \|\vec{x}\|^2.$$

- Im \mathbb{R}^2 sind alle Drehungen von der Form $f(\vec{x}) = D(\varphi)\vec{x}$ (siehe Beispiele 4.1.1 und 4.1.2)
- Im \mathbb{R}^3 ist

$$A = \begin{bmatrix} D(\varphi) & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \cos \varphi & -\sin \varphi & 0 \\ \sin \varphi & \cos \varphi & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

eine Drehung um die z -Achse um den Winkel φ . Für allgemeine Drehmatrizen A ist die Drehachse durch die **Fixpunktgerade**

$$\{\vec{x} : A\vec{x} = \vec{x}\}$$

gegeben. Der **Drehwinkel** φ bestimmt sich gemäß der Formel

$$2 \cos \varphi + 1 = \text{spur } A,$$

wobei $\text{spur } A := a_{11} + a_{22} + a_{33}$ die **Spur von** A bezeichnet.

Beispiel Die Matrix

$$A = \frac{1}{3} \begin{bmatrix} 2 & 2 & -1 \\ -1 & 2 & 2 \\ 2 & -1 & 2 \end{bmatrix}$$

ist orthogonal und $\det A = 1$. Also ist A eine Drehung. Die Drehgerade bestimmt sich als Lösungsmenge des LGS

$$A\vec{x} = \vec{x} \Leftrightarrow (A - E)\vec{x} = \vec{0},$$

also durch den Kern der Matrix $A - E$:

$$A - E = \frac{1}{3} \begin{bmatrix} 2 & 2 & -1 \\ -1 & 2 & 2 \\ 2 & -1 & 2 \end{bmatrix} - \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} -1 & 2 & -1 \\ -1 & -1 & 2 \\ 2 & -1 & -1 \end{bmatrix}$$

und man erkennt sofort

$$\text{kern}(A - E) = \left\{ t \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} : t \in \mathbb{R} \right\}.$$

Für den Drehwinkel φ folgt

$$2 \cos \varphi + 1 = 2, \text{ also } \cos \varphi = \frac{1}{2}, \text{ damit } \varphi = \pm \frac{\pi}{3}.$$

Das Vorzeichen des Drehwinkels hängt davon ab, aus welcher Richtung man auf die Drehachse schaut.

Projektionen im \mathbb{R}^n

Eine lineare Abbildung $f(\vec{x}) = A\vec{x}$ im \mathbb{R}^n heißt **Projektion**, falls $A^2 = A$ ist.

Jeder Bildpunkt $\vec{v} = A\vec{x}$ von f ist zugleich ein Fixpunkt, denn

$$A\vec{v} = A(A\vec{x}) = A^2\vec{x} = A\vec{x} = \vec{v}.$$

Jeder Punkt \vec{x} wird durch einmalige Anwendung der Abbildung f auf die Fixpunktmenge abgebildet und bleibt bei weiterer Anwendung der Abbildung dann unverändert.

Die Projektion heißt **orthogonal**, falls

$$\langle A\vec{x} - \vec{x}, A\vec{x} \rangle = 0 \quad \text{für alle } \vec{x} \in \mathbb{R}^n.$$

Beispiel 4.2.1 Die Matrix $A = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ beschreibt eine orthogonale Projektion auf die Gerade $g : t \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

Die Matrix $A = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$ beschreibt eine Projektion auf dieselbe Gerade, sie ist aber nicht orthogonal, denn z.B.

$$A\vec{e}_1 - \vec{e}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Also $\langle A\vec{e}_1 - \vec{e}_1, A\vec{e}_1 \rangle = 1 \neq 0$.

Projektionen auf Geraden

Ist $g : t\vec{v}$ eine Gerade im \mathbb{R}^n so beschreibt

$$A_g := \frac{\vec{v} \cdot \vec{v}^T}{\|\vec{v}\|^2} = \frac{1}{\|\vec{v}\|^2} \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} \cdot [v_1 \quad \dots \quad v_n] = \frac{1}{\|\vec{v}\|^2} \begin{bmatrix} v_1^2 & v_1 v_2 & \dots & v_1 v_n \\ v_2 v_1 & v_2^2 & \dots & v_2 v_n \\ \vdots & \vdots & \dots & \vdots \\ v_n v_1 & v_n v_2 & \dots & v_n^2 \end{bmatrix}$$

die orthogonale Projektion auf g .

Insbesondere ist $\|A_g \vec{x} - \vec{x}\| = d(\vec{x}, g)$ der Abstand von \vec{x} zur Geraden g .

Projektionen auf implizit definierte Mengen

Ist $M : \langle \vec{x}, \vec{n} \rangle = 0$ für $\vec{n} \neq 0$ eine implizit definierte Menge im \mathbb{R}^n , so ist

$$A_M = E - \frac{\vec{n} \cdot \vec{n}^T}{\|\vec{n}\|^2}$$

die orthogonale Projektion auf M .

Insbesondere ist $\|A_M \vec{x} - \vec{x}\| = d(\vec{x}, M)$ der Abstand von \vec{x} zu M (vgl. Kapitel 2 für den Fall implizit definierter Geraden und Ebenen).

Wie M implizit gegebene Mengen heißen auch **(lineare) Hyperebenen** in \mathbb{R}^n . Sie verallgemeinern Geraden durch den Ursprung in \mathbb{R}^2 und Ebenen durch den Ursprung in \mathbb{R}^3 . Lineare Hyperebenen in \mathbb{R}^n sind lineare Unterräume der Dimension $n - 1$.

Beispiel 4.2.2

$$\vec{v} = \begin{bmatrix} 3 \\ 4 \end{bmatrix} \Rightarrow A_g = \frac{1}{25} \begin{bmatrix} 9 & 12 \\ 12 & 16 \end{bmatrix}$$

beschreibt die orthogonale Projektion auf die Gerade $g : t \begin{bmatrix} 3 \\ 4 \end{bmatrix}$.

Ein Normalenvektor zu g ist $\vec{n} = \begin{bmatrix} -4 \\ 3 \end{bmatrix}$, also $g : \langle \vec{x}, \vec{n} \rangle = 0$ die implizite Darstellung. In der Tat ist auch

$$E - \frac{\vec{n} \cdot \vec{n}^T}{\|\vec{n}\|^2} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \frac{1}{25} \begin{bmatrix} 16 & -12 \\ -12 & 9 \end{bmatrix} = \frac{1}{25} \begin{bmatrix} 9 & 12 \\ 12 & 16 \end{bmatrix}.$$

Spiegelungen im \mathbb{R}^n

Eine lineare Abbildung $f(\vec{x}) = B\vec{x}$ im \mathbb{R}^n heißt **Spiegelung**, falls $B^2 = E$ ist. Zweimaliges Spiegeln führt also auf den Ausgangspunkt zurück.

Ist A eine Projektion, so ist $B := 2A - E$ eine Spiegelung (an der Fixpunktmenge von A), denn

$$B^2 = (2A - E) \cdot (2A - E) = 4A^2 - 2AE - 2EA + E^2 = E.$$

Umgekehrt: Ist B eine Spiegelung, so ist $A = \frac{1}{2}(B + E)$ eine Projektion (auf die Fixpunktmenge von B), denn

$$A^2 = \frac{1}{4} (B + E) \cdot (B + E) = \frac{1}{4} (B^2 + BE + EB + E^2) = \frac{1}{4} (2B + 2E) = A.$$

Beispiel 4.2.3 Die Spiegelung an der Geraden $g : t\vec{v}$ ist gegeben durch

$$B_g := 2A_g - E = 2 \frac{\vec{v} \cdot \vec{v}^T}{\|\vec{v}\|^2} - E.$$

Im Beispiel 4.2.2 ist die Spiegelung an der Geraden $g : t \begin{bmatrix} 3 \\ 4 \end{bmatrix}$ gegeben durch $B_g = \frac{1}{25} \begin{bmatrix} -7 & 24 \\ 24 & 7 \end{bmatrix}$.

Die Spiegelung an der implizit definierten Menge $M : \langle \vec{x}, \vec{n} \rangle = 0$ ist gegeben durch

$$B_M := 2A_M - E = E - 2 \frac{\vec{n} \cdot \vec{n}^T}{\|\vec{n}\|^2}.$$

Diese Abbildung wird auch **Householder-Transformation** genannt.

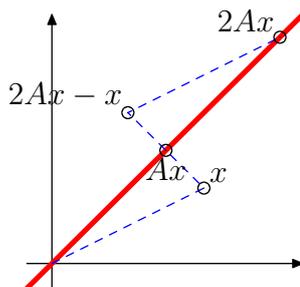


Abbildung 4.1: Projektion zur Matrix A und Spiegelung $x \mapsto 2Ax - x$.

Basiswechsel

Die Komponenten x_1, \dots, x_n des Vektors $\vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n$ heißen auch **kartesische**

Koordinaten von \vec{x} . Sie beschreiben die Koordinaten des Vektors \vec{x} bezüglich der Einheitsvektoren $\vec{e}_1, \dots, \vec{e}_n$, denn

$$\vec{x} = \sum_{i=1}^n x_i \vec{e}_i.$$

Mitunter kann es nützlich sein, eine andere Basis

$$\vec{v}_1, \dots, \vec{v}_n$$

des \mathbb{R}^n zu betrachten.

Zur **Umrechnung** der kartesischen Koordinaten von \vec{x} in die Koordinaten bezüglich der neuen Basis, bilden wir die Matrix $V = [\vec{v}_1, \dots, \vec{v}_n]$. V ist invertierbar, denn $\text{rang } V = n$, und die Komponenten y_i des Vektors $\vec{y} = V^{-1}\vec{x}$ beschreiben die Koordinaten von \vec{x} bezüglich der Basisvektoren $\vec{v}_1, \dots, \vec{v}_n$.

(Beweis: $\sum_{i=1}^n y_i \vec{v}_i = V\vec{y} = V(V^{-1}\vec{x}) = \vec{x}$.)

Die Matrix V^{-1} beschreibt also den **Basiswechsel** von der Standardbasis $\vec{e}_1, \dots, \vec{e}_n$ zu $\vec{v}_1, \dots, \vec{v}_n$ und die Matrix V entsprechend den umgekehrten Basiswechsel.

Ist $f(\vec{x}) = A\vec{x}$ eine lineare Abbildung in \mathbb{R}^n , so gilt im neuen Koordinatensystem

$$\tilde{f}(\vec{y}) = \tilde{A}\vec{y} \quad \text{mit } \tilde{A} := V^{-1}AV.$$

Die Matrizen A und \tilde{A} heißen **ähnlich** (bzw. **äquivalent**), da sie dieselbe lineare Abbildung (allerdings bezüglich verschiedener Koordinatensysteme) beschreiben.

Beispiele 4.2.4 Es sei $\vec{v}_1 = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$ und $\vec{v}_2 = \begin{bmatrix} -4 \\ 3 \end{bmatrix}$ Basis des \mathbb{R}^2 . Dann ist

$$V = \begin{bmatrix} 3 & -4 \\ 4 & 3 \end{bmatrix}, \quad V^{-1} = \frac{1}{25} \begin{bmatrix} 3 & 4 \\ -4 & 3 \end{bmatrix}.$$

Die Projektion A_g und die Spiegelung B_g für die Gerade $g : t\vec{v}_1$ haben bezüglich der V -Koordinaten die Abbildungsmatrix

$$\tilde{A}_g = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad \tilde{B}_g = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

Beispiel 4.2.5 Die Abbildung

$$f(\vec{x}) := \vec{x} \times \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} x_2 - x_3 \\ x_3 - x_1 \\ x_1 - x_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 & -1 \\ -1 & 0 & 1 \\ 1 & -1 & 0 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

ist linear. Für die Basis

$$\vec{v}_1 = \frac{1}{\sqrt{6}} \begin{bmatrix} \sqrt{3} \\ -\sqrt{3} \\ 0 \end{bmatrix}, \quad \vec{v}_2 = \frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}, \quad \vec{v}_3 = \frac{1}{\sqrt{6}} \begin{bmatrix} \sqrt{2} \\ \sqrt{2} \\ \sqrt{2} \end{bmatrix}$$

ist die Matrix $V = [\vec{v}_1, \vec{v}_2, \vec{v}_3]$ orthogonal, d.h. $V^{-1} = V^T$.

Die Abbildungsmatrix $\tilde{A} = V^{-1}AV = V^TAV$ bezüglich der neuen Basis ist gegeben durch

$$\begin{aligned} \tilde{A} &= \frac{1}{\sqrt{6}} \begin{bmatrix} \sqrt{3} & -\sqrt{3} & 0 \\ 1 & 1 & -2 \\ \sqrt{2} & \sqrt{2} & \sqrt{2} \end{bmatrix} \cdot \begin{bmatrix} 0 & 1 & -1 \\ -1 & 0 & 1 \\ 1 & -1 & 0 \end{bmatrix} \cdot \frac{1}{\sqrt{6}} \begin{bmatrix} \sqrt{3} & 1 & \sqrt{2} \\ -\sqrt{3} & 1 & \sqrt{2} \\ 0 & -2 & \sqrt{2} \end{bmatrix} \\ &= \begin{bmatrix} 0 & \sqrt{3} & 0 \\ -\sqrt{3} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \end{aligned}$$

Diese Matrix lässt sich in das Produkt $\tilde{A} = \tilde{A}_3 \cdot \tilde{A}_2 \cdot \tilde{A}_1$ zerlegen, wobei

$$\tilde{A}_1 := \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \text{Projektion auf die } (\vec{v}_1, \vec{v}_2) \text{ - Ebene,}$$

$$\tilde{A}_2 := \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \text{Drehung der } (\vec{v}_1, \vec{v}_2) \text{ - Ebene um den Winkel } -\frac{\pi}{2},$$

$$\tilde{A}_3 := \sqrt{3}E = \text{Streckung um den Faktor } \sqrt{3}.$$

4.3 Eigenwerte und -vektoren

Definition Es sei A eine $n \times n$ -Matrix. Ein Vektor $\vec{v} \in \mathbb{R}^n$, $\vec{v} \neq \vec{0}$, heißt **Eigenvektor von** A zum **Eigenwert** $\lambda \in \mathbb{R}$, wenn

$$A\vec{v} = \lambda\vec{v}.$$

Ist λ Eigenwert von A , so ist also das homogene LGS

$$(A - \lambda E)\vec{x} = \vec{0}$$

nicht eindeutig lösbar, also $\det(A - \lambda E) = 0$. Ist umgekehrt $\det(A - \lambda E) = 0$, so ist

$$\dim \text{kern}(A - \lambda E) > 0.$$

Also enthält $\text{kern}(A - \lambda E)$ einen Vektor $\vec{x} \neq 0$. Für \vec{x} gilt $A\vec{x} = \lambda\vec{x}$, \vec{x} ist also Eigenvektor von A zum Eigenwert λ .

Die Funktion

$$p(\lambda) := \det(A - \lambda E), \quad \lambda \in \mathbb{R}$$

definiert ein Polynom vom Grad n in der Variablen λ . Das Polynom p heißt **charakteristisches Polynom von A** .

Beispiel 4.3.1 Es sei

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} \quad \text{also} \quad A - \lambda E = \begin{bmatrix} 1 - \lambda & 2 \\ 2 & 1 - \lambda \end{bmatrix}$$

und

$$p(\lambda) = \det(A - \lambda E) = (1 - \lambda)^2 - 4 = \lambda^2 - 2\lambda - 3 = (\lambda - 3)(\lambda + 1).$$

A besitzt also die Eigenwerte 3 und -1 . Zugehörige Eigenvektoren sind

$$\vec{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \text{und} \quad \vec{v}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

Normiert man die Eigenvektoren noch zu 1, also

$$\vec{v}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \text{und} \quad \vec{v}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

so erhält man **in diesem Fall** ein Orthonormalsystem. Bezüglich dieser Basis hat die durch A beschriebene lineare Abbildung die Abbildungsmatrix $\begin{bmatrix} 3 & 0 \\ 0 & -1 \end{bmatrix}$, d.h. es ist

$$V^T \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} V = \begin{bmatrix} 3 & 0 \\ 0 & -1 \end{bmatrix}.$$

Einschub: Fundamentalsatz der Algebra

Die quadratische Gleichung

$$x^2 + px + q = 0$$

besitzt die beiden Lösungen

$$x_{\pm} = -\frac{p}{2} \pm \sqrt{\underbrace{\frac{p^2}{4} - q}_{=: \Delta = \text{Diskriminante}}}$$

und es gilt:

1. Fall: $\Delta \geq 0 \Rightarrow x_{\pm}$ reelle Lösungen

2. Fall: $\Delta < 0 \Rightarrow x_{\pm} = -\frac{p}{2} \pm i\sqrt{-\Delta}$ zwei zueinander konjugiert komplexe Lösungen

Die **komplex Konjugierte** \bar{z} einer komplexen Zahl $z = a + ib$ ist dabei gegeben durch

$$\bar{z} := a - ib$$

Man erhält \bar{z} aus z durch Spiegelung an der reellen Achse.

Im Unterschied zu \mathbb{R} sind in \mathbb{C} auch Gleichungen höheren Grades stets lösbar. Dazu betrachten wir die Gleichung

$$z^n + a_{n-1}z^{n-1} + \dots + a_1z + a_0 = 0 \quad (4.3.1)$$

mit einem Polynom p n -ten Grades mit Koeffizienten $a_0, \dots, a_{n-1} \in \mathbb{C}$. Die Nullstellen des Polynoms p sind dann gerade die Lösungen der Gleichung (4.3.1). Dann gilt der

Fundamentalsatz der Algebra

Jedes Polynom n -ten Grades lässt sich als Produkt von n **Linearfaktoren** schreiben:

$$z^n + a_{n-1}z^{n-1} + \dots + a_1z + a_0 = (z - b_1)(z - b_2) \dots (z - b_n).$$

Die komplexen Zahlen b_1, \dots, b_n sind die **Nullstellen** des Polynoms.

Fallen mehrere Nullstellen zusammen, so spricht man von Nullstellen höherer Vielfachheit. Jedes Polynom vom Grad n besitzt also genau n komplexe Nullstellen, wenn man jede Nullstelle samt ihrer Vielfachheit zählt.

Für Polynome mit **reellen** Koeffizienten a_0, \dots, a_{n-1} gilt zusätzlich:

Die nicht reellen Nullstellen treten als Paare zueinander konjugiert komplexer Nullstellen gleicher Vielfachheit auf.

Beispiel

$$z^3 + 5z^2 + z + 5 = 0$$

Rate Nullstelle: $b_1 = i$

$$i^3 + 5i^2 + i + 5 = -i - 5 + i + 5 = 0.$$

Dann ist auch $\bar{b}_1 = -i$ eine Nullstelle. Polynomdivision durch $(z - i)(z + i) = z^2 + 1$ ergibt

$$(z^3 + 5z^2 + z + 5) : (z^2 + 1) = z + 5.$$

Also folgt

$$z^3 + 5z^2 + z + 5 = (z - i)(z + i)(z + 5).$$

Zurück zu Eigenwerten und Eigenvektoren.

Es sei A wieder eine $n \times n$ -Matrix, p sein charakteristisches Polynom. Der Fundamentalsatz der Algebra liefert die Darstellung

$$p(\lambda) = \det(A - \lambda E) = (-1)^n (\lambda - \lambda_1) \cdot \dots \cdot (\lambda - \lambda_n).$$

Hierbei sind $\lambda_1, \dots, \lambda_n$ die (komplexen) Nullstellen von p . Insbesondere folgt:

- 1) Es gibt höchstens n verschiedene Eigenwerte zur Matrix A .
- 2) $\det A = \lambda_1 \cdot \dots \cdot \lambda_n$, wobei λ_i die (komplexen) Nullstellen des charakteristischen Polynoms sind.
- 3) A ist also insbesondere genau dann invertierbar, wenn alle Nullstellen $\lambda_1, \dots, \lambda_n$ von 0 verschieden sind.
- 4) Eigenwerte von A und A^T (aber im allgemeinen **nicht** die Eigenvektoren!) stimmen überein.

Weiter gilt

- 5) Eigenvektoren zu **verschiedenen** Eigenwerten sind linear unabhängig.
- 6) Die Spur von A ist gleich der Summe der Nullstellen von p

$$\text{spur } A = a_{11} + \dots + a_{nn} = \lambda_1 + \dots + \lambda_n.$$

Ist λ k -fache reelle Nullstelle von p , also λ insbesondere Eigenwert von A , so ist die Dimension des zugehörigen **Eigenraumes**

$$\{\vec{x} : A\vec{x} = \lambda\vec{x}\}$$

mindestens 1 und höchstens k .

Beispiel 4.3.2 Gegeben sie die $n \times n$ -Matrix

$$A = \begin{bmatrix} 2 & 1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \dots & \dots & 0 & 2 \end{bmatrix} \Rightarrow p(\lambda) = (2 - \lambda)^n.$$

Also ist 2 n -fache Nullstelle des charakteristischen Polynoms, aber

$$\dim\{\vec{x} : A\vec{x} = 2\vec{x}\} = 1.$$

Ein zugehöriger Eigenvektor ist \vec{e}_1 .

Wichtiger Spezialfall

Ist A reell und **symmetrisch**, d.h. $A^T = A$, so sind alle Eigenwerte des zugehörigen charakteristischen Polynoms reell. Weiterhin stimmen Dimension des zugehörigen Eigenraums und Vielfachheit des Eigenwertes überein. Weiterhin können Eigenvektoren so gewählt werden, dass sie eine Basis aus orthogonalen Einheitsvektoren bilden (siehe Beispiel 4.3.1).

Diagonalisierung

Ist A eine $n \times n$ -Matrix A mit genau n linear unabhängigen Eigenvektoren $\vec{v}_1, \dots, \vec{v}_n$ zu den Eigenwerten $\lambda_1, \dots, \lambda_n$, so folgt mit dem Basiswechsel

$$V = [\vec{v}_1, \dots, \vec{v}_n]$$

$$\tilde{A} = V^{-1}AV = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix} =: \text{diag}(\lambda_1, \dots, \lambda_n).$$

A ist also ähnlich zu einer Diagonalmatrix, bei der auf der Diagonalen die Eigenwerte der Matrix A stehen. Man sagt dann, dass A **diagonalisierbar** ist. Dies ist beispielsweise der Fall, wenn

- n paarweise verschiedene Eigenwerte existieren oder
- A reell symmetrisch ist (siehe oben).

Beispiel 4.3.3 Die Matrix

$$A = \begin{bmatrix} 2 & 1 & 4 \\ 0 & 1 & 2 \\ 0 & 0 & 4 \end{bmatrix}$$

besitzt das charakteristische Polynom $p(\lambda) = (2 - \lambda)(1 - \lambda)(4 - \lambda)$. Zugehörige Eigenvektoren sind

$$\vec{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \quad \vec{v}_3 = \begin{bmatrix} 7 \\ 2 \\ 3 \end{bmatrix}.$$

Die Matrix

$$V = [\vec{v}_1, \vec{v}_2, \vec{v}_3] = \begin{bmatrix} 1 & 1 & 7 \\ 0 & -1 & 2 \\ 0 & 0 & 3 \end{bmatrix}$$

ist invertierbar mit

$$V^{-1} = \begin{bmatrix} 1 & 1 & -3 \\ 0 & -1 & \frac{2}{3} \\ 0 & 0 & \frac{1}{3} \end{bmatrix}$$

und man prüft leicht nach, dass

$$V^{-1}AV = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{bmatrix}.$$

Beispiel 4.3.4 Viele Matrizen sind aber nicht diagonalisierbar. Beispielsweise ist die obere Dreiecksmatrix

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

nicht diagonalisierbar: Weil A eine (obere) Dreiecksmatrix ist, sieht man sofort, dass 1 der einzige Eigenwert (mit der algebraischen Vielfachheit 2) ist. Die zugehörigen Eigenvektoren sind genau die von Null verschiedenen Vektoren des Kerns der Matrix

$$A - 1 \cdot E_2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

Diese Matrix hat Rang 1. Aus der Rangformel folgt unmittelbar, dass ihr Kern ebenfalls eindimensional ist. Daher kann es keine zwei voneinander linear unabhängigen Eigenvektoren geben.

Kapitel 5

Komplexe Zahlen

In Kapitel 1 hatten wir bereits die Menge der komplexen Zahlen kennengelernt:

$$\mathbb{C} = \{z = a + ib : a, b \in \mathbb{R}\}$$

Komplexe Zahlen lassen sich als Punkte der Gaußschen Zahlenebene veranschaulichen (siehe Kapitel 1). Der Betrag $|z|$ der komplexen Zahl $z = a + ib$ ist definiert durch

$$|z| = \sqrt{a^2 + b^2}.$$

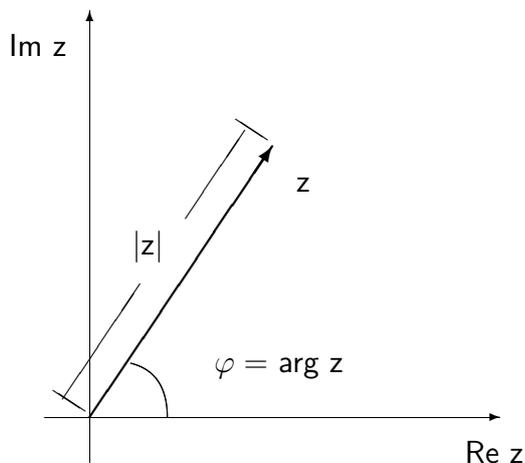
Indem wir $z = a + ib$ mit dem Ortsvektor $\begin{bmatrix} a \\ b \end{bmatrix}$ im \mathbb{R}^2 identifizieren, ist also

$$|z| = \left\| \begin{bmatrix} a \\ b \end{bmatrix} \right\|$$

d.h. die euklidische Norm des zugehörigen Ortsvektors.

Polardarstellung komplexer Zahlen

Komplexe Zahlen $z = a + ib$ sind durch ihren Betrag $|z|$ und ihren Winkel φ mit der reellen Achse eindeutig bestimmt. φ heißt **Argument** von z (Schreibweise $\arg(z)$).

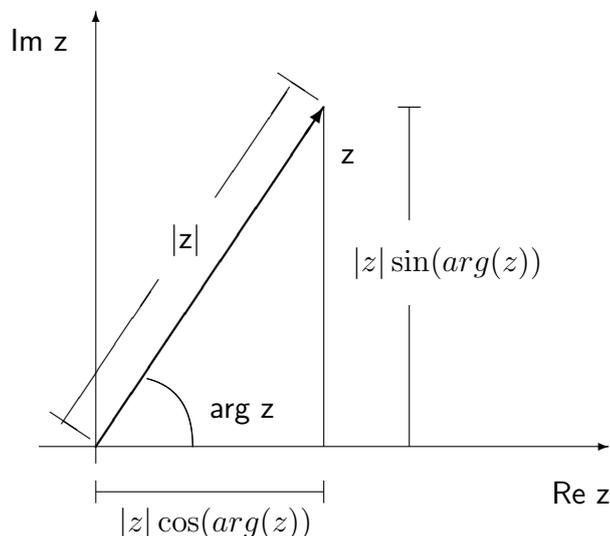


Beispiele $\arg(i) = \frac{\pi}{2}$, $\arg(-i) = \frac{3}{2}\pi$

$\arg(z)$ ist nur bis auf Vielfaches von 2π eindeutig bestimmt und für $z = 0$ undefiniert.
Es gilt:

$$\operatorname{Re}(z) = |z| \cos(\arg(z))$$

$$\operatorname{Im}(z) = |z| \sin(\arg(z))$$



Damit erhält man die Polardarstellung komplexer Zahlen

$$z = |z| (\cos(\arg(z)) + i \sin(\arg(z)))$$

Geometrische Deutung der Multiplikation

Es seien

$$z_1 = r_1(\cos \varphi_1 + i \sin \varphi_1), \quad r_1 = |z_1|, \quad \varphi_1 = \arg(z_1)$$

$$z_2 = r_2(\cos \varphi_2 + i \sin \varphi_2), \quad r_2 = |z_2|, \quad \varphi_2 = \arg(z_2)$$

komplexe Zahlen in Polardarstellung.

Dann folgt aus den Additionstheoremen für Winkelfunktionen

$$\begin{aligned} z_1 \cdot z_2 &= r_1 \cdot r_2 ((\cos(\varphi_1) \cdot \cos(\varphi_2) - \sin(\varphi_1) \cdot \sin(\varphi_2)) + i (\cos(\varphi_1) \cdot \sin(\varphi_2) + \sin(\varphi_1) \cdot \cos(\varphi_2))) \\ &= r_1 \cdot r_2 (\cos(\varphi_1 + \varphi_2) + i \sin(\varphi_1 + \varphi_2)) \end{aligned}$$

Es gilt also

$$\begin{aligned} |z_1 \cdot z_2| &= |z_1| \cdot |z_2| \\ \arg(z_1 \cdot z_2) &= \arg(z_1) + \arg(z_2) \end{aligned}$$

Mit anderen Worten: Komplexe Zahlen werden multipliziert, indem man ihre

- Beträge multipliziert
- Argumente addiert.

Geometrisch entspricht dies einer **Drehstreckung**.

Beispiel

$$z_1 = i, \text{ also } |z_1| = 1, \arg(z_1) = \frac{\pi}{2}$$

$$z_2 = -2i, \text{ also } |z_2| = 2, \arg(z_2) = \frac{3}{2}\pi$$

$$\implies |z_1 \cdot z_2| = 2 \quad \text{und} \quad \arg(z_1 \cdot z_2) = \frac{\pi}{2} + \frac{3}{2}\pi = 2\pi \quad (\hat{=} 0)$$

Also gilt $z_1 \cdot z_2 = 2(\cos(2\pi) + i \sin(2\pi)) = 2$.

Geometrische Deutung der Division

Ist $z_2 \neq 0$, so gilt

$$\frac{z_1}{z_2} = z_1 \cdot \frac{1}{z_2}$$

Für den Reziprokwert $\frac{1}{z_2}$ gilt die Formel

$$\frac{1}{z_2} = \frac{\bar{z}_2}{|z_2|^2}$$

wobei \bar{z}_2 die konjugierte komplexe Zahl von z_2 ist. Die Polardarstellung des Reziprokwertes ist

$$\frac{1}{z_2} = \frac{1}{r_2}(\cos(-\varphi_2) + i \sin(-\varphi_2))$$

und damit

$$\frac{z_1}{z_2} = \frac{r_1}{r_2}(\cos(\varphi_1 - \varphi_2) + i \sin(\varphi_1 - \varphi_2))$$

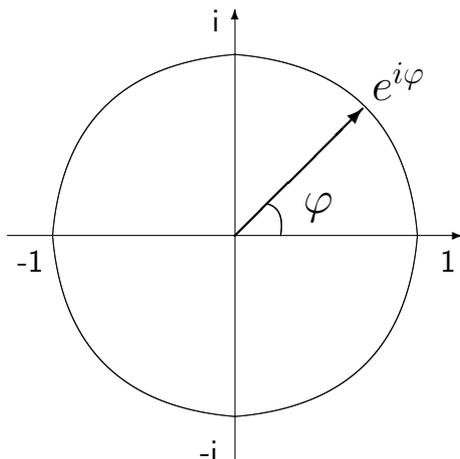
Zwei komplexe Zahlen werden also dividiert, indem man ihre

- Beträge dividiert
- Argumente subtrahiert.

Komplexe Exponentialfunktion

Für $\varphi \in \mathbb{R}$ definiert man

$$e^{i\varphi} := \cos(\varphi) + i \sin(\varphi) \quad (\text{Eulersche Formel})$$



Damit gilt insbesondere $\arg(e^{i\varphi}) = \varphi$, also lässt sich jede komplexe Zahl in der Form

$$z = |z|e^{i\arg(z)} = re^{i\varphi}, \quad r = |z|, \varphi = \arg(z)$$

schreiben. Es gilt insbesondere

$$e^{i\varphi_1} \cdot e^{i\varphi_2} = e^{i(\varphi_1+\varphi_2)}$$

und damit hat die komplexe Multiplikation die besonders einfache Gestalt

$$z_1 \cdot z_2 = (r_1 e^{i\varphi_1}) \cdot (r_2 e^{i\varphi_2}) = r_1 r_2 e^{i(\varphi_1+\varphi_2)}$$

Für beliebige komplexe Zahlen $z = a + ib$ definiert man schließlich

$$e^z = e^a \cdot e^{ib} = e^a (\cos(b) + i \sin(b))$$

Dann gilt insbesondere:

$$e^{z_1} \cdot e^{z_2} = e^{z_1+z_2}, \quad \frac{e^{z_1}}{e^{z_2}} = e^{z_1-z_2}$$

Beispiele $e^0 = e^{2\pi i} = 1$, $e^{i\frac{\pi}{2}} = i$, $e^{i\pi} = -1$, $e^{-i\frac{\pi}{2}} = -i$, $e^{2+i\pi} = -e^2$

Komplexer Logarithmus

Die Umkehrfunktion der e -Funktion heißt **natürlicher Logarithmus** (Bezeichnung: \ln).

Es gilt

$$\ln(z) = \ln(|z|) + i \arg(z)$$

wobei $\ln(z)$ den natürlichen Logarithmus der **reellen** Zahl $|z|$ meint (siehe Kapitel 7).

(Begründung: $e^{\ln(|z|)+i\arg(z)} = |z|e^{i\arg(z)} = z$)

Der (komplexe) Logarithmus ist für 0 nicht definiert und die Mehrdeutigkeit des Arguments überträgt sich auf den Logarithmus. Eindeutigkeit erhält man durch die Forderung $\arg(z) \in [0, 2\pi[$.

Beispiele $\ln(-1) = i\pi$, $\ln(-e) = 1+i\pi$, $\ln(i) = i\frac{\pi}{2}$, $\ln(3+4i) = \ln(5)+i \arctan(\frac{4}{3})$

Komplexe Potenzfunktion

Für $x \neq 0, y \in \mathbb{C}$ definiert man

$$x^y = e^{y \ln(x)} \quad (5.0.1)$$

Die komplexe Wurzelfunktion ist gegeben durch

$$\sqrt[n]{z} := z^{\frac{1}{n}} = e^{\frac{1}{n} \ln(z)}, \quad n \in \mathbb{N}$$

Beispiele

$$i^i = e^{i \ln(i)} = e^{-\frac{\pi}{2}}$$

$$\sqrt{-1} = e^{\frac{1}{2} \ln(-1)} = e^{i\frac{\pi}{2}} = i$$

$$\sqrt{-16} = i\sqrt{16} = 4i$$

Kapitel 6

Folgen und Reihen

6.1 Folgen

Eine **Folge** reeller Zahlen ist eine Abbildung

$$\begin{aligned}\mathbb{N} &\rightarrow \mathbb{R} \\ n &\mapsto a_n\end{aligned}$$

Schreibweisen: $(a_n)_{n \in \mathbb{N}}$ oder auch a_1, a_2, a_3, \dots . Das Element a_n heißt n -tes **Folgenglied**, n der zugehörige **Index**.

Allgemeiner kann man auch Folgen

$$(a_n)_{n \geq n_0} \quad \text{oder} \quad a_{n_0}, a_{n_0+1}, a_{n_0+2}, \dots$$

für beliebiges $n_0 \in \mathbb{Z}$ betrachten.

Folgen $(a_n)_{n \geq 1}$ können **explizit** oder **rekursiv** definiert sein:

- **explizite Definition**

$$a_n = c, \quad c \in \mathbb{R}, \quad \text{also } c, c, \dots \text{ (konstante Folge)}$$

$$a_n = \frac{1}{n}, \quad \text{also } 1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \dots$$

$$a_n = \frac{(-1)^n}{n}, \quad \text{also } -1, +\frac{1}{2}, -\frac{1}{3}, +\frac{1}{4}, \dots$$

$$a_n = n^2, \quad \text{also } 1, 4, 9, 16, 25, \dots$$

$$a_n = (-1)^n, \quad \text{also } -1, +1, -1, +1, \dots$$

- **rekursive Definition**

$$a_1 = 1, \quad a_{n+1} = (n+1)a_n \quad \text{für } n \geq 1, \quad \text{also } 1, 2, 6, 24, \dots$$

$$a_1 = \sqrt{2}, \quad a_{n+1} = \frac{1}{2}\left(a_n + \frac{2}{a_n}\right) \quad \text{für } n \geq 1, \quad \text{also } \sqrt{2}, \sqrt{2}, \dots$$

$$a_0 = 1, \quad a_1 = 1, \quad a_{n+2} = a_{n+1} + a_n \quad \text{für } n \geq 0, \quad \text{also } 1, 1, 2, 3, 5, 8, \dots$$

(Fibonacci Folge)

Eine Folge $(a_n)_{n \geq 1}$ heißt

- **positiv / negativ**, falls alle Folgenglieder a_n positiv/negativ sind
- **alternierend**, falls aufeinanderfolgende Folgenglieder verschiedene Vorzeichen haben (also gilt, dass $a_{n+1} \cdot a_n < 0$)
- **monoton wachsend/fallend**, falls $a_{n+1} \geq a_n$ bzw. $a_{n+1} \leq a_n$ für alle n
- **streng monoton wachsend/fallend**, falls $a_{n+1} > a_n$ bzw. $a_{n+1} < a_n$ für alle n
- **beschränkt**, falls eine Zahl $M \in \mathbb{R}$ existiert mit $|a_n| \leq M$ für alle n

Definition Es sei $(a_n)_{n \geq n_0}$ eine Folge reeller Zahlen. Die Folge heißt **konvergent** gegen $a \in \mathbb{R}$, falls gilt: Für alle $\varepsilon > 0$ existiert ein $N_\varepsilon \in \mathbb{N}$ mit

$$|a_n - a| < \varepsilon \quad \text{für alle } n \geq N_\varepsilon$$

a heißt **Grenzwert** der Folge $(a_n)_{n \geq n_0}$ und wir schreiben

$$\lim_{n \rightarrow \infty} a_n = a$$

Nicht konvergente Folgen heißen **divergent**.

Die Folge $(a_n)_{n \geq n_0}$ konvergiert also genau dann gegen a , falls für **alle** $\varepsilon > 0$ alle, bis auf endlich viele, Folgenglieder im Intervall $]a - \varepsilon, a + \varepsilon[$ liegen.

Beispiele

(i) Die Folge $a_n = \frac{1}{n}$ konvergiert gegen 0, denn für alle $\varepsilon > 0$ gibt es ein $N_\varepsilon \in \mathbb{N}$ mit

$$N_\varepsilon > \frac{1}{\varepsilon} \quad \Leftrightarrow \quad \varepsilon > \frac{1}{N_\varepsilon}$$

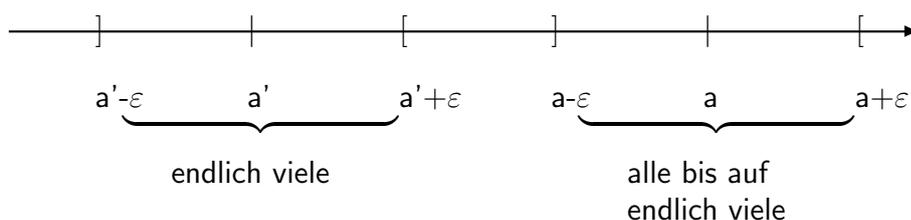
und für $n \geq N_\varepsilon$ folgt

$$\left| \frac{1}{n} - 0 \right| = \frac{1}{n} \leq \frac{1}{N_\varepsilon} < \varepsilon$$

Folgen $(a_n)_{n \geq n_0}$ mit $\lim_{n \rightarrow \infty} a_n = 0$ heißen auch **Nullfolgen**.

(ii) Die konstante Folge c, c, \dots konvergiert offensichtlich gegen c .

Der Grenzwert a einer konvergenten Folge $(a_n)_{n \geq n_0}$ ist **eindeutig bestimmt**, denn ist $a' \neq a$, so folgt für $\varepsilon = \frac{1}{3} |a' - a| (> 0)$ dass nur endlich viele Folgenglieder außerhalb von $]a - \varepsilon, a + \varepsilon[$ liegen, also nur endlich viele Folgenglieder innerhalb $]a' - \varepsilon, a' + \varepsilon[$



Konvergente Folgen $(a_n)_{n \geq n_0}$ sind **beschränkt**, denn:

Gilt $\lim_{n \rightarrow \infty} a_n = a$, so gibt es zu $\varepsilon = 1$ ein $N \in \mathbb{N}$ mit $|a_n - a| < 1$ für $n \geq N$. Insbesondere

$$|a_n| \leq |a_n - a| + |a| < 1 + |a| \quad \text{für } n \geq N$$

Also gilt

$$|a_n| \leq \max \{ |a_{n_0}|, |a_{n_0+1}|, \dots, |a_{N-1}|, 1 + |a| \} =: M$$

für alle n . Hierbei bezeichnet $\max\{\dots\}$ das Maximum der in der Menge $\{\dots\}$ enthaltenen Zahlen.

Beispiele

- Die Folgen $a_n = n$, $(-1)^n n$, n^2 sind divergent, da unbeschränkt.
- Die Folge $a_n = (-1)^n$ ist divergent, denn für jedes $a \in \mathbb{R}$ gilt entweder

$$|a_{2n} - a| \geq \frac{1}{2} \quad \text{oder} \quad |a_{2n+1} - a| \geq \frac{1}{2}$$

Rechenregeln für Grenzwerte

Es seien (a_n) und (b_n) zwei konvergente Folgen mit Grenzwerten a und b . Dann gilt

$$(i) \quad \lim_{n \rightarrow \infty} (a_n + b_n) = \lim_{n \rightarrow \infty} a_n + \lim_{n \rightarrow \infty} b_n = a + b$$

$$(ii) \quad \lim_{n \rightarrow \infty} (a_n \cdot b_n) = \lim_{n \rightarrow \infty} a_n \cdot \lim_{n \rightarrow \infty} b_n = a \cdot b$$

$$(iii) \quad \lim_{n \rightarrow \infty} c \cdot a_n = c \cdot \lim_{n \rightarrow \infty} a_n = c \cdot a \quad \text{für alle } c \in \mathbb{R}$$

(iv) Ist $b \neq 0$, so gibt es ein $n_0 \in \mathbb{N}$ mit $b_n \neq 0$ für $n \geq n_0$ und dann

$$\lim_{n \rightarrow \infty} \frac{a_n}{b_n} = \frac{\lim_{n \rightarrow \infty} a_n}{\lim_{n \rightarrow \infty} b_n} = \frac{a}{b}$$

Wichtige Grenzwerte

•

$$\lim_{n \rightarrow \infty} \frac{p_k n^k + p_{k-1} n^{k-1} + \dots + p_0}{q_l n^l + q_{l-1} n^{l-1} + \dots + q_0} = \begin{cases} 0 & \text{für } k < l \\ \frac{p_k}{q_k} & \text{für } k = l \end{cases}$$

mit $k, l \in \mathbb{N} \cup \{0\}$, $p_i, q_j \in \mathbb{R}$ und $p_k, q_l \neq 0$.

Für $k > l$ ist die Folge divergent.

- $\lim_{n \rightarrow \infty} \frac{\ln n}{n^\alpha} = 0 \quad \text{für } \alpha > 0$
- $\lim_{n \rightarrow \infty} \frac{n^\alpha}{e^{\beta n}} = \lim_{n \rightarrow \infty} n^\alpha e^{-\beta n} = 0 \quad \text{für } \alpha \in \mathbb{R}, \beta > 0$
- $\lim_{n \rightarrow \infty} n^\alpha q^n = 0 \quad \text{für } \alpha \in \mathbb{R}, |q| < 1$
- $\lim_{n \rightarrow \infty} \sqrt[n]{n} = 1$

- $\lim_{n \rightarrow \infty} \left(1 + \frac{x}{n}\right)^n = e^x$ für alle $x \in \mathbb{R}$
- $\lim_{n \rightarrow \infty} \frac{n}{\sqrt[n]{n!}} = e$

Konvergenzkriterien

- **Monotoniekriterium** Jede monoton wachsende/fallende beschränkte Zahlenfolge ist konvergent.
- **Cauchy-Kriterium** Eine Folge (a_n) ist konvergent, genau dann wenn für alle $\varepsilon > 0$ ein $N_\varepsilon \in \mathbb{N}$ existiert mit

$$|a_n - a_m| < \varepsilon \text{ für alle } m, n \geq N_\varepsilon$$

(Beachte: Mithilfe des Cauchy-Kriteriums kann man eine Folge auf Konvergenz testen, ohne den Grenzwert a kennen zu müssen.)

- **Intervallschachtelung** Es seien (a_n) monoton wachsend, (b_n) monoton fallend und $(b_n - a_n)$ eine Nullfolge. Dann konvergieren $(a_n), (b_n)$ gegen denselben Grenzwert und es gilt

$$\lim_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} b_n \in [a_k, b_k] \text{ für alle } k$$

Bestimmte Divergenz

Eine Folge $(a_n)_{n \geq n_0}$ heißt **bestimmt divergent** gegen $+\infty$ (bzw. $-\infty$) falls zu jedem $M \in \mathbb{R}$ ein $N \in \mathbb{N}$ existiert mit

$$a_n > M \quad (\text{ bzw. } a_n < M) \text{ für alle } n \geq N$$

Wir schreiben

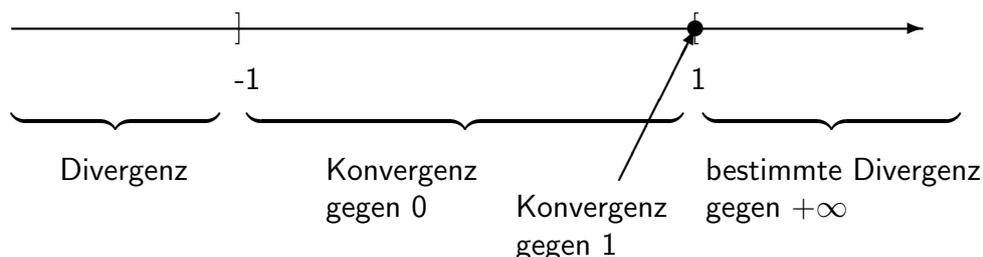
$$\lim_{n \rightarrow \infty} a_n = +\infty \quad (\text{ bzw. } \lim_{n \rightarrow \infty} a_n = -\infty)$$

Beispiele

(i) $a_n = n, n^2, n^3$, usw. sind bestimmt divergent gegen $+\infty$

(ii) Die Folge q^n , $n \in \mathbb{N}$ ist

- konvergent gegen 0 für $|q| < 1$
- konstant 1 für $q = 1$
- bestimmt divergent gegen $+\infty$ für $q > 1$
- divergent für $q \leq -1$



Vektorfolgen

Eine **Folge von Vektoren** $(\vec{a}_n)_{n \geq n_0}$ ist eine Abbildung $n \mapsto \vec{a}_n$, die jeder Zahl n einen Vektor $\vec{a}_n \in \mathbb{R}^d$ zuordnet. Die Vektorfolge heißt konvergent mit Grenzwert \vec{a} , falls die Komponentenfolgen $(a_{i,n})_{n \geq n_0}$ von \vec{a}_n gegen die Komponenten a_i von \vec{a} konvergieren.

$$\vec{a}_n = \begin{bmatrix} a_{1,n} \\ \vdots \\ a_{d,n} \end{bmatrix} \quad \vec{a} = \begin{bmatrix} a_1 \\ \vdots \\ a_d \end{bmatrix}$$

Wir schreiben dann $\lim_{n \rightarrow \infty} \vec{a}_n = \vec{a}$.

Funktionenfolgen

Es sei $I \subset \mathbb{R}$ ein Intervall. Eine **Funktionenfolge** $(f_n)_{n \geq n_0}$ ist eine Abbildung $n \mapsto f_n$ die jeder Zahl n eine Funktion $f_n : I \rightarrow \mathbb{R}$ zuordnet. Die Funktionenfolge $(f_n)_{n \geq n_0}$ heißt **punktweise konvergent auf I mit Grenzfunktion f** , wenn für alle $x \in I$ die Folge der Funktionswerte $(f_n(x))_{n \geq n_0}$ gegen $f(x)$ konvergiert.

Wir schreiben dann

$$\lim_{n \rightarrow \infty} f_n = f \quad \text{oder} \quad \lim_{n \rightarrow \infty} f_n(x) = f(x), x \in I$$

Beispiele

(i) $f_n(x) = (1 + \frac{x}{n})^n$ ist punktweise konvergent auf \mathbb{R} mit Grenzfunktion e^x

(ii) $f_n(x) = x^n$ ist

- punktweise konvergent auf $] -1, 1[$ gegen 0
- nicht konvergent auf $[-1, +1]$, da $(f_n(-1))_{n \geq 1}$ divergent.

6.2 Reihen

Es sei $(a_n)_{n \geq n_0}$ eine Folge reeller Zahlen. Die Folge

$$s_n := \sum_{k=n_0}^n a_k, \quad n \geq n_0$$

der Partialsummen heißt (**unendliche**) **Reihe** und wird mit $\sum_{k=n_0}^{\infty} a_k$ bezeichnet.

Definition Die Reihe $\sum_{k=n_0}^{\infty} a_k$ heißt **konvergent** (bzw. **divergent** oder **bestimmt divergent**), falls die Folge der Partialsummen $(s_n)_{n \geq n_0}$ konvergiert (bzw. divergiert oder bestimmt divergiert).

Im Falle $\lim_{n \rightarrow \infty} s_n = \lim_{n \rightarrow \infty} \sum_{k=n_0}^n a_k = s$ mit $s \in \mathbb{R} \cup \{-\infty, \infty\}$ nennt man s den **Wert** oder die **Summe** der unendlichen Reihe und man schreibt

$$\sum_{k=n_0}^{\infty} a_k = s$$

Beispiele

- (i) geometrische Reihe: $\sum_{k=0}^{\infty} q^k$ für $q \in \mathbb{R}$

Für die Partialsummen gilt (siehe Kapitel 1):

$$s_n = \begin{cases} \frac{1-q^{n+1}}{1-q} & \text{für } q \neq 1 \\ n+1 & \text{für } q = 1 \end{cases}$$

Folglich

$$\sum_{k=0}^{\infty} q^k = \begin{cases} \frac{1}{1-q} & \text{für } |q| < 1 \\ \infty & \text{für } q \geq 1 \\ \text{divergent} & \text{für } q \leq -1 \end{cases}$$

- (ii) harmonische Reihe: $\sum_{k=1}^{\infty} \frac{1}{k}$ ist bestimmt divergent gegen $+\infty$, denn für die Partialsummen s_n gilt die Abschätzung:

$$\underbrace{\ln(1+n)}_{\rightarrow \infty} \leq s_n \leq 1 + \ln(n) \quad \text{für } n \geq 1$$

Konvergenzkriterien

- **Cauchy-Konvergenzkriterium** Die Reihe $\sum_{k=n_0}^{\infty} a_k$ ist genau dann konvergent, wenn es zu jedem $\varepsilon > 0$ ein $N_\varepsilon \in \mathbb{N}$ gibt mit

$$|s_n - s_m| = \left| \sum_{k=m+1}^n a_k \right| < \varepsilon \quad \text{für alle } m, n \geq N_\varepsilon$$

Bemerkung Insbesondere gilt

$$\sum_{k=n_0}^{\infty} a_k \text{ konvergent} \quad \implies \quad (a_n)_{n \geq n_0} \text{ ist eine Nullfolge}$$

Umgekehrt ist die Bedingung, dass $(a_n)_{n \geq n_0}$ eine Nullfolge ist, jedoch **nicht hinreichend** für die Konvergenz der Reihe $\sum_{k=n_0}^{\infty} a_k$, wie man am Beispiel der harmonischen Reihe sieht. Jedoch gilt das

- **Leibniz-Kriterium für alternierende Reihen** Ist $(a_n)_{n \geq n_0}$ eine **monoton fallende Nullfolge**, dann ist die alternierende Reihe

$$\sum_{k=n_0}^{\infty} (-1)^k a_k$$

konvergent.

Für Reihen mit nichtnegativen Reihengliedern ist Konvergenz einfacher zu untersuchen:

Definition Die Reihe $\sum_{k=n_0}^{\infty} a_k$ heißt **absolut konvergent**, falls die Reihe $\sum_{k=n_0}^{\infty} |a_k|$ der Absolutbeträge der Reihenglieder konvergiert.

Bemerkung

- (i) $\sum_{k=n_0}^{\infty} a_k$ absolut konvergent $\implies \sum_{k=n_0}^{\infty} a_k$ konvergent und $|\sum_{k=n_0}^{\infty} a_k| \leq \sum_{k=n_0}^{\infty} |a_k|$
- (ii) $\sum_{k=n_0}^{\infty} a_k$ absolut konvergent \Leftrightarrow Folge der Partialsummen $s_n = |a_{n_0}| + |a_{n_0+1}| + \dots + |a_n|$ ist beschränkt (Anwendung des Monotoniekriteriums auf die Folge der Partialsummen s_n !)

Beispiel Die Reihe $\sum_{k=1}^{\infty} \frac{1}{k(k+1)}$ ist absolut konvergent, denn für die n -te Partialsumme s_n gilt:

$$s_n = \sum_{k=1}^n \underbrace{\frac{1}{k(k+1)}}_{\frac{1}{k} - \frac{1}{k+1}} = \sum_{k=1}^n \left(\frac{1}{k} - \frac{1}{k+1} \right) = 1 - \frac{1}{n+1}$$

Es ist $\lim_{n \rightarrow \infty} s_n = 1$, also hat die Reihe den Wert 1.

Kriterien für absolute Konvergenz**Vergleichskriterien**

Gilt für die Reihenglieder der Reihen $\sum_{k=n_0}^{\infty} a_k$ und $\sum_{k=n_0}^{\infty} b_k$ ab einem Index n_1 die Ungleichung

$$|a_n| \leq b_n$$

dann gilt

- (i) **Majorantenkriterium** $\sum_{k=n_0}^{\infty} b_k$ konvergent $\implies \sum_{k=n_0}^{\infty} a_k$ absolut konvergent
- (ii) **Minorantenkriterium** $\sum_{k=n_0}^{\infty} |a_k| = +\infty \implies \sum_{k=n_0}^{\infty} b_k = +\infty$

Beispiele

(i) Die Reihe $\sum_{k=1}^{\infty} \frac{1}{k^2}$ ist (absolut) konvergent, denn

$$\frac{1}{k^2} \leq 2 \frac{1}{k(k+1)} \quad \text{für alle } k \geq 1$$

und die Reihe $\sum_{k=1}^{\infty} 2 \frac{1}{k(k+1)} = 2 \sum_{k=1}^{\infty} \frac{1}{k(k+1)}$ ist (absolut) konvergent.

Damit sind dann auch die Reihen

$$\sum_{k=1}^{\infty} \frac{1}{k^3}, \sum_{k=1}^{\infty} \frac{1}{k^4}, \dots \quad \text{und} \quad \sum_{k=1}^{\infty} \frac{k}{k^3+1}, \sum_{k=1}^{\infty} \frac{2k^2+4}{k^5+1}, \dots$$

absolut konvergent.

(ii) Die Reihe $\sum_{k=1}^{\infty} \frac{1}{\sqrt{k}}$ ist divergent, denn

$$\frac{1}{\sqrt{k}} \geq \frac{1}{k} \quad \text{für alle } k \geq 1$$

und die harmonische Reihe $\sum_{k=1}^{\infty} \frac{1}{k}$ ist divergent.

Durch **Vergleich mit der geometrischen Reihe** erhält man aus den Vergleichskriterien zwei wichtige **Spezialfälle**:

a) **Quotientenkriterium** Es gebe ein n_1 mit $a_n \neq 0$ für $n \geq n_1$. Dann gilt

$$(i) \lim_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| < 1 \implies \sum_{k=n_0}^{\infty} a_k \text{ absolut konvergent}$$

$$(ii) \lim_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| > 1 \implies \sum_{k=n_0}^{\infty} a_k \text{ divergent}$$

Den angesprochenen Vergleich mit der geometrischen Reihe erhält man wie folgt:

Im Falle von (i) gibt es ein $q < 1$ und ein $N \in \mathbb{N}$ mit

$$\left| \frac{a_{n+1}}{a_n} \right| \leq q \text{ für alle } n \geq N.$$

Daher gilt

$$|a_n| = \underbrace{\left| \frac{a_n}{a_{n-1}} \right|}_{\leq q} \cdot \underbrace{\left| \frac{a_{n-1}}{a_{n-2}} \right|}_{\leq q} \cdots \underbrace{\left| \frac{a_{N+1}}{a_N} \right|}_{\leq q} \cdot |a_N| \leq q^{n-N} \cdot |a_N|$$

und damit

$$\sum_{k=n_0}^{\infty} |a_k| \leq \sum_{k=n_0}^{N-1} |a_k| + \underbrace{\sum_{k=N}^{\infty} q^{k-N} \cdot |a_N|}_{= \frac{1}{1-q}} < \infty$$

Im Falle von (ii) gibt es ein $q > 1$ und ein $N \in \mathbb{N}$ mit

$$\left| \frac{a_{n+1}}{a_n} \right| \geq q \text{ für alle } n \geq N.$$

Daher gilt $|a_n| \geq q^{n-N} |a_N|$ für $n \geq N$, und damit ist (a_n) keine Nullfolge. Also ist die Reihe $\sum_{k=n_0}^{\infty} a_k$ divergent.

Beispiele Folgende Reihen konvergieren für alle $x \in \mathbb{R}$ absolut:

$$(1) \sum_{k=0}^{\infty} \frac{x^k}{k!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots, \text{ denn } \left| \frac{a_{k+1}}{a_k} \right| = \frac{|x|}{k+1} < 1 \text{ für } k \geq |x|$$

$$(2) \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \dots$$

$$(3) \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{(2k)!} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \dots$$

b) **Wurzelkriterium**

$$(i) \lim_{n \rightarrow \infty} \sqrt[n]{|a_n|} < 1 \implies \sum_{k=n_0}^{\infty} a_k \text{ absolut konvergent}$$

$$(ii) \lim_{n \rightarrow \infty} \sqrt[n]{|a_n|} > 1 \implies \sum_{k=n_0}^{\infty} a_k \text{ divergent}$$

Beispiel

$\sum_{k=1}^{\infty} k \frac{1}{2^k}$ ist absolut konvergent, denn

$$\lim_{n \rightarrow \infty} \sqrt[n]{n \frac{1}{2^n}} = \underbrace{\lim_{n \rightarrow \infty} \sqrt[n]{n}}_{=1} \cdot \lim_{n \rightarrow \infty} \underbrace{\sqrt[n]{\frac{1}{2^n}}}_{=\frac{1}{2}} = \frac{1}{2}$$

Kapitel 7

Abbildungen

7.1 Grundlegendes

Definition Es seien D, Y zwei nichtleere Mengen. Eine Vorschrift f , die **jedem** $x \in D$ **genau ein** Element $f(x) \in Y$ zuordnet, heißt **Abbildung** (oder **Funktion**) von D nach Y .

Schreibweise:

$$\begin{aligned} f : D &\rightarrow Y \\ x &\mapsto f(x) \end{aligned}$$

Dabei ist

D = Definitionsbereich (von f)

Y = Wertebereich (von f)

x = unabhängige Variable (**Argument**)

$f(x)$ = abhängige Variable (**Funktionswert**)

$x \mapsto f(x)$ heißt **Abbildungsvorschrift**

$f(D) := \{f(x) : x \in D\} \subset Y$ heißt **Bild** von f

Ist $f(x) = y$, so heißt x **Urbild** von y .

Ist $D, Y \subset \mathbb{R}$ (bzw. $D, Y \subset \mathbb{C}$), so spricht man von reellen (bzw. komplexen) Funktionen.

Beispiele

(i) $D, Y = \mathbb{R}$, $x \mapsto x^2$ also $f(x) = x^2$, Bild = $[0, \infty[$

(ii) $D, Y = [0, \infty[$, $x \mapsto \sqrt{x}$

(iii) Die Abbildung

$$D \rightarrow Y, \quad z \mapsto (|z|, \arg(z))$$

mit $D = \mathbb{C} \setminus \{0\}$, $Y =]0, \infty[\times [0, 2\pi[:= \{(r, \varphi) \mid r \in]0, \infty[, \varphi \in [0, 2\pi[\}$ ordnet der komplexen Zahl $z \neq 0$ ihre Polarkoordinaten zu.

Verkettung von Abbildungen

Definition Für Abbildungen $f : D \rightarrow Y, g : Y \rightarrow X$ heißt die Abbildung

$$g \circ f : D \rightarrow X, \quad x \mapsto g(f(x))$$

die **Verkettung** von g und f .

Beispiel $D, Y, X = \mathbb{R}, f(x) = \sin x, g(x) = x^2$

$$\implies (g \circ f)(x) = g(f(x)) = g(\sin x) = (\sin x)^2, \text{ d.h.}$$

$$g \circ f : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto (\sin x)^2$$

Aber: $f \circ g(x) = f(g(x)) = f(x^2) = \sin(x^2)$, d.h. im Allgemeinen also $g \circ f \neq f \circ g$.

Umkehrabbildungen

Definition Es sei $f : D \rightarrow Y$ eine Abbildung. f heißt

- **injektiv**, falls $f(x_1) \neq f(x_2)$ für alle $x_1, x_2 \in D, x_1 \neq x_2$
- **surjektiv**, wenn $f(D) = Y$, also wenn jedes $y \in Y$ mindestens ein Urbild besitzt
- **bijektiv**, wenn f injektiv und surjektiv ist.

Beispiele $f : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto x^2$, ist weder injektiv noch surjektiv. Durch Einschränkung auf das Bild $[0, \infty[$ wird

$$f : \mathbb{R} \rightarrow [0, \infty[, x \mapsto x^2 \quad \text{surjektiv.}$$

f ist nach wie vor nicht injektiv, denn $f(\sqrt{y}) = f(-\sqrt{y}) = y$. Jedes $y > 0$ besitzt also zwei Urbilder. Durch Einschränkung des Definitionsbereiches auf $[0, \infty[$ wird

$$f : [0, \infty[\rightarrow [0, \infty[, x \mapsto x^2$$

bijektiv mit Umkehrabbildung

$$f^{-1} : [0, \infty[\rightarrow [0, \infty[, y \mapsto \sqrt{y}.$$

Aber auch $f :]-\infty, 0] \rightarrow [0, \infty[, x \mapsto x^2$ bijektiv mit Umkehrabbildung $f^{-1}(y) = -\sqrt{y}$.

Ist f bijektiv, so gibt es zu jedem $y \in Y$ **genau ein** Urbild (Schreibweise: $f^{-1}(y)$). Die Zuordnung

$$y \mapsto f^{-1}(y)$$

definiert eine Abbildung

$$f^{-1} : Y \rightarrow D$$

f^{-1} heißt **Umkehrabbildung** von f .

Es gilt:

$$\begin{aligned} f(f^{-1}(y)) &= y && \text{für alle } y \in f(D) = Y \\ f^{-1}(f(x)) &= x && \text{für alle } x \in D \end{aligned}$$

Man erhält $f^{-1}(y)$ durch **Auflösen** der Gleichung

$$y = f(x)$$

nach x .

Beispiel Für $\alpha > 0$ ist $f(x) = e^{-\alpha \cdot x^2}$, $[0, \infty[\rightarrow]0, 1]$ bijektiv. Die Gleichung

$$y = e^{-\alpha x^2}$$

besitzt die positive Lösung $x = \sqrt{-\frac{1}{\alpha} \ln y}$, denn:

$$y = e^{-\alpha x^2} \Leftrightarrow \ln y = -\alpha x^2 \Leftrightarrow x^2 = -\frac{1}{\alpha} \ln y$$

Beachte Ist $f : D \rightarrow Y$ injektiv, so ist $f : D \rightarrow f(D)$ bijektiv und damit **umkehrbar**. In *diesem* Sinne ist jede injektive Abbildung umkehrbar und wir werden den Begriff der Umkehrbarkeit in diesem erweiterten Sinne ebenfalls verwenden.

7.2 Reelle Funktionen einer Variablen

In diesem Abschnitt gilt stets $D, Y \subset \mathbb{R}$. $f : D \rightarrow Y$ ist also eine reelle Funktion einer Variablen.

Typischerweise ist $D = \mathbb{R}$ oder D Intervall der Form

$$\underbrace{[a, b]}_{\text{abgeschlossen}} \quad \underbrace{]a, b[}_{\text{offen}} \quad \underbrace{]a, b], [a, b[}_{\text{halboffen}}$$

oder

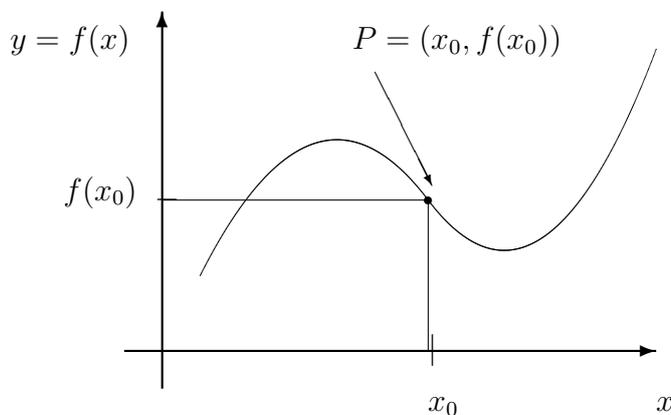
$$[a, \infty[:= \{x \in \mathbb{R} : a \leq x\}, \quad]-\infty, a] := \{x \in \mathbb{R} : x \leq a\}$$

Analoges gilt für den Wertebereich Y .

Graphische Darstellung

Der **Funktionsgraph** von f ist die Menge

$$\{(x, f(x)) : x \in D\} \subset D \times Y \quad (\subset \mathbb{R}^2)$$



Nullstellen

$x_0 \in D$ mit $f(x_0) = 0$ heißt **Nullstelle** der Funktion f . $(x_0, f(x_0))$ ist in diesem Falle Schnittpunkt des Funktionsgraphen mit der x -Achse.

Symmetrie

f heißt **gerade**, wenn ihr Funktionsgraph **spiegelsymmetrisch** zur y -Achse ist:

$$f(-x) = f(x)$$

Beispiel $f(x) = x^n$, n gerade; $f(x) = \cos x$

f heißt **ungerade**, wenn ihr Funktionsgraph **punktsymmetrisch** zum Nullpunkt ist:

$$f(-x) = -f(x)$$

Beispiel $f(x) = x^n$, n ungerade; $f(x) = \sin x$

Monotonie

f heißt **monoton wachsend** (bzw. **monoton fallend**), falls

$$f(x_1) \leq f(x_2) \quad (\text{ bzw. } f(x_1) \geq f(x_2)) \text{ für alle } x_1, x_2 \in D, x_1 < x_2$$

f heißt **streng monoton wachsend** (bzw. **streng monoton fallend**), falls

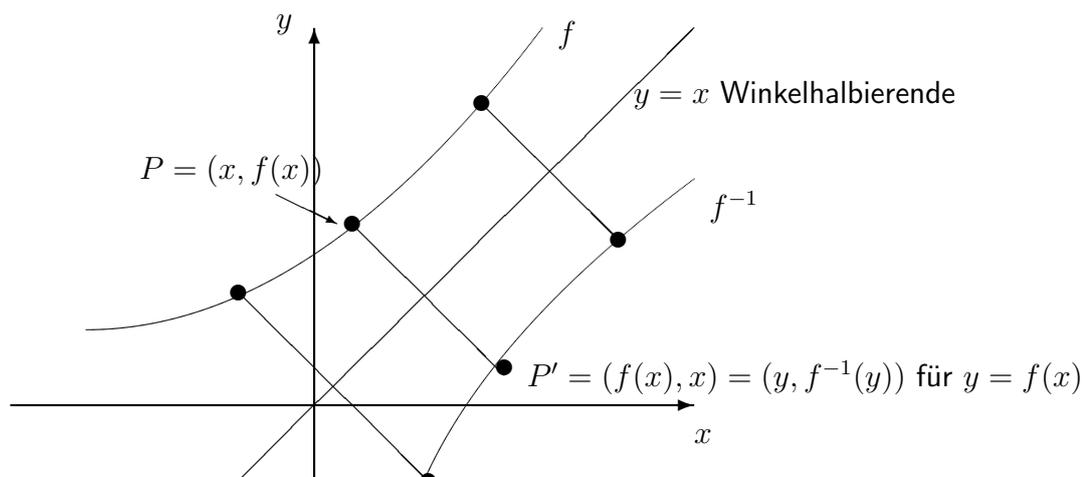
$$f(x_1) < f(x_2) \quad (\text{ bzw. } f(x_1) > f(x_2)) \text{ für alle } x_1, x_2 \in D, x_1 < x_2$$

Regeln

- streng monotone Funktionen sind injektiv, also umkehrbar
Beispiel $f(x) = x^n$ auf $[0, \infty[$ mit Umkehrfunktion $f^{-1}(y) = \sqrt[n]{y}$
- Verkettungen monoton wachsender Funktionen sind wieder monoton wachsend
- Verkettungen monoton fallender Funktionen sind wieder monoton fallend
- Verkettungen gerader Funktionen sind wieder gerade
- Verkettungen ungerader Funktionen sind wieder ungerade

Funktionsgraph der Umkehrfunktion

Den Funktionsgraph der Umkehrfunktion f^{-1} erhält man durch **Spiegelung** des Funktionsgraphen von f an der **Winkelhalbierenden**



7.3 Elementare Funktionen

A) Polynome

Eine (auf ganz \mathbb{R} definierte) Funktion

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0$$

mit $a_i \in \mathbb{R}$ und $a_n \neq 0$ heißt Polynom n -ten Grades. Aus dem Fundamentalsatz der Algebra folgt, dass sich f schreiben lässt als Produkt von

- Linearfaktoren $(x - x_i)$ (wobei x_i die reellen Nullstellen)
- quadratischen Faktoren $(x^2 + p_j x + q_j)$, wobei die Nullstellen von $x^2 + p_j x + q_j$ ein konjugiert komplexes Nullstellenpaar von f sind.

$$\begin{aligned} f(x) &= a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0 \\ &= a_n (x - x_1) \cdot \dots \cdot (x - x_k) \cdot (x^2 + p_1 x + q_1) \cdot \dots \cdot (x^2 + p_l x + q_l) \end{aligned}$$

mit $k + 2l = n$, $x_1, \dots, x_k \in \mathbb{R}$, $p_1, q_1, \dots, p_l, q_l \in \mathbb{R}$

B) Potenz- und Wurzelfunktionen

$$f(x) = x^\alpha \text{ für } \alpha \in \mathbb{R} \text{ mit } D(f) =]0, \infty[$$

Für ganzzahlige Exponenten α kann man den Definitionsbereich erweitern:

- für $\alpha \in \mathbb{N} \cup \{0\}$ können wir $D(f) = \mathbb{R}$ wählen,
- für $\alpha \in -\mathbb{N}$ können wir $D(f) = \mathbb{R} \setminus \{0\}$ wählen.

Potenzgesetze Für $x, y > 0, \alpha, \beta \in \mathbb{R}$ gilt:

$$(x \cdot y)^\alpha = x^\alpha \cdot y^\alpha \qquad \left(\frac{x}{y}\right)^\alpha = \frac{x^\alpha}{y^\alpha}$$

$$x^{\alpha+\beta} = x^\alpha \cdot x^\beta \qquad x^{\alpha-\beta} = \frac{x^\alpha}{x^\beta}$$

$$x^{\alpha \cdot \beta} = (x^\alpha)^\beta = (x^\beta)^\alpha$$

C) Exponentialfunktion und Logarithmus

Hält man in der Potenzfunktion a^x die Basis a konstant, so erhält man eine Funktion vom Exponenten x :

$$f(x) = a^x, \quad x \in \mathbb{R}, a > 0$$

f heißt **Exponentialfunktion zur Basis a** .

Spezialfall Natürliche Basis $e = \lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n \approx 2,71828$ (Eulersche Zahl)

$$f(x) = e^x = \exp(x) \quad \text{e-Funktion}$$

Rechenregeln für Exponentialfunktionen

Analog zu den Potenzgesetzen gelten die folgenden Eigenschaften:

$$a^{x+y} = a^x \cdot a^y$$

$$a^x \cdot b^x = (ab)^x$$

$$(a^x)^y = a^{x \cdot y}$$

Für $a > 0, a \neq 1$, ist a^x injektiv, also umkehrbar mit Umkehrfunktion

$$\log_a :]0, \infty[\rightarrow \mathbb{R}, x \mapsto \log_a(x)$$

Logarithmusfunktion zur Basis a

Insbesondere gilt also

$$y = \log_a(x) \quad \Leftrightarrow \quad a^y = x$$

Im Spezialfall $a = e$ spricht man von der natürlichen Logarithmusfunktion und schreibt $\ln x$:

$$\ln x := \log_e(x)$$

Rechenregeln für Logarithmusfunktionen

Wiederum aus den Potenzgesetzen ergibt sich:

$$\log_a(x \cdot y) = \log_a(x) + \log_a(y)$$

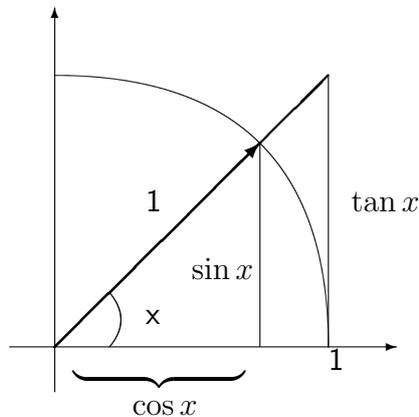
$$\log_a(x^b) = b \log_a(x)$$

also

$$\log_a\left(\frac{1}{x}\right) = -\log_a(x) \quad \text{und} \quad \log_a\left(\frac{x}{y}\right) = \log_a(x) - \log_a(y)$$

Umrechnung

$$a^x = e^{\ln a \cdot x} \qquad \log_a(x) = \frac{\ln x}{\ln a}$$

D) Trigonometrische Funktionen**Im Einheitskreis**

\sin ist injektiv auf $[-\frac{\pi}{2}, \frac{\pi}{2}]$, da streng monoton wachsend. Die Umkehrfunktion

$$\arcsin : [-1, 1] \rightarrow [-\frac{\pi}{2}, \frac{\pi}{2}]$$

heißt **Arcussinus**.

\cos ist injektiv auf $[0, \pi]$, da streng monoton fallend. Die Umkehrfunktion

$$\arccos : [-1, 1] \rightarrow [0, \pi]$$

heißt **Arcuscosinus**.

\tan ist injektiv auf $]-\frac{\pi}{2}, \frac{\pi}{2}[$, da streng monoton wachsend. Die Umkehrfunktion

$$\arctan : \mathbb{R} \rightarrow]-\frac{\pi}{2}, \frac{\pi}{2}[$$

heißt **Arcustangens**.

7.4 Grenzwerte von Funktionen

Definition Die Funktion f sei in einer Umgebung von x_0 definiert. Gibt es dann ein $c \in \mathbb{R} \cup \{\pm\infty\}$ mit folgender Eigenschaft: für **jede** Folge $(x_n)_{n \geq 1}$ mit $x_n \in D, n \geq 1$, und $\lim_{n \rightarrow \infty} x_n = x_0$ gilt

$$\lim_{n \rightarrow \infty} f(x_n) = c.$$

Dann sagen wir f **besitzt den Grenzwert c an der Stelle x_0** und wir schreiben in diesem Falle

$$\lim_{x \rightarrow x_0} f(x) = c.$$

Beispiele

- $\lim_{x \rightarrow 0} x \sin\left(\frac{1}{x}\right) = 0$

Denn für **jede** Folge $(x_n)_{n \geq 1}$ mit $\lim_{n \rightarrow \infty} x_n = 0$ gilt

$$\lim_{n \rightarrow \infty} \left| x_n \sin\left(\frac{1}{x_n}\right) \right| \leq \lim_{n \rightarrow \infty} |x_n| = 0$$

- **Heaviside-Funktion:** $H : \mathbb{R} \rightarrow \mathbb{R}$

$$H(x) = \begin{cases} 0 & \text{für } x < 0 \\ 1 & \text{für } x \geq 0 \end{cases}$$

H besitzt an der Stelle 0 keinen Grenzwert, denn

$$\lim_{n \rightarrow \infty} H\left(-\frac{1}{n}\right) = 0 \quad \text{aber} \quad \lim_{n \rightarrow \infty} H\left(\frac{1}{n}\right) = 1$$

Im Sinne folgender Definition besitzt H an der Stelle 0 den linksseitigen Grenzwert 0 und den rechtsseitigen Grenzwert +1.

Definition f besitzt in x_0 den **linksseitigen** (bzw. **rechtsseitigen**) Grenzwert c , falls

$$\lim_{n \rightarrow \infty} f(x_n) = c$$

für jede Folge $(x_n)_{n \geq 1}$ mit $x_n \in D$, $n \geq 1$, $\lim_{n \rightarrow \infty} x_n = x_0$ **und** $x_n < x_0$ (bzw. $x_n > x_0$) für alle n . In diesem Falle schreiben wir

$$\lim_{x \rightarrow x_0^-} f(x) = c, \quad (\text{bzw. } \lim_{x \rightarrow x_0^+} f(x) = c)$$

Grenzwerte für $x \rightarrow \pm\infty$

Wir schreiben $\lim_{x \rightarrow \infty} f(x) = c$ (bzw. $\lim_{x \rightarrow -\infty} f(x) = c$) falls

$$\lim_{n \rightarrow \infty} f(x_n) = c$$

für jede Folge $(x_n)_{n \geq 1}$ mit $x_n \in D$, $n \geq 1$, und $\lim_{n \rightarrow \infty} x_n = +\infty$ (bzw. $\lim_{n \rightarrow \infty} x_n = -\infty$).

Beispiel Für $\alpha > 0$ gilt $\lim_{x \rightarrow \infty} \frac{1}{x^\alpha} = 0$, denn für jede Folge $(x_n)_{n \geq 1}$ mit $\lim_{n \rightarrow \infty} x_n = +\infty$ gilt

$$\lim_{n \rightarrow \infty} x_n^\alpha = \infty, \quad \text{also} \quad \lim_{n \rightarrow \infty} \frac{1}{x_n^\alpha} = 0$$

Rechenregeln für Grenzwerte

Es gelten dieselben Rechenregeln wie für Grenzwerte von Folgen:

(i) $\lim_{x \rightarrow x_0} f(x) + g(x) = \lim_{x \rightarrow x_0} f(x) + \lim_{x \rightarrow x_0} g(x)$

$$(ii) \lim_{x \rightarrow x_0} f(x) \cdot g(x) = \lim_{x \rightarrow x_0} f(x) \cdot \lim_{x \rightarrow x_0} g(x)$$

$$(iii) \lim_{x \rightarrow x_0} c \cdot f(x) = c \lim_{x \rightarrow x_0} f(x)$$

$$(iv) \lim_{x \rightarrow x_0} \frac{f(x)}{g(x)} = \frac{\lim_{x \rightarrow x_0} f(x)}{\lim_{x \rightarrow x_0} g(x)}, \text{ falls } \lim_{x \rightarrow x_0} g(x) \neq 0$$

Beispiel $\lim_{x \rightarrow \infty} \frac{4x^3 + 2x - 10}{8x^3 - 1} = \frac{1}{2}$, denn

$$\frac{4x^3 + 2x - 10}{8x^3 - 1} = \frac{4 + \frac{2}{x^2} - \frac{10}{x^3}}{8 - \frac{1}{x^3}}$$

und

$$\lim_{x \rightarrow \infty} \frac{4 + \frac{2}{x^2} - \frac{10}{x^3}}{8 - \frac{1}{x^3}} = \frac{\lim_{x \rightarrow \infty} 4 + \frac{2}{x^2} - \frac{10}{x^3}}{\lim_{x \rightarrow \infty} 8 - \frac{1}{x^3}} = \frac{4}{8} = \frac{1}{2}$$

Stetigkeit von Funktionen

Definition. Die Funktion $f : D \rightarrow \mathbb{R}$ heißt **stetig im Punkt** $x_0 \in D$, falls

$$\lim_{x \rightarrow x_0} f(x) = f(x_0)$$

f heißt **stetig auf D**, falls f stetig in jedem Punkt aus D .

Beispiele

- Stetig sind die durch die Abbildungsvorschriften (mit den jeweils zuvor diskutierten Definitions- und Wertebereichen) gegebenen Funktionen $x \mapsto c$, $x \mapsto x$, $x \mapsto x^2$, ..., allgemein $x \mapsto x^\alpha$, allgemeiner beliebige Polynomabbildungen, sowie die trigonometrischen Funktionen \sin , \cos , \tan , ..., die Exponentialfunktionen \exp , $x \mapsto a^x$ und deren Umkehrfunktionen (jeweils auf ihrem Definitionsbereich).
- Die Heaviside Funktion ist nicht stetig in $x = 0$.

Aus den Grenzwertsätzen für Funktionen folgt:

Sind $f, g : D \rightarrow \mathbb{R}$ stetig, so sind auch die folgenden Funktionen stetig:

$$(i) f + g$$

$$(ii) f \cdot g$$

$$(iii) c \cdot f \text{ für alle } c \in \mathbb{R}$$

$$(iv) \frac{f}{g} \text{ auf } D_0 := \{x : g(x) \neq 0\}$$

$$(v) \text{ Ist } h : E \rightarrow \mathbb{R} \text{ stetig und } f(D) \subset E, \text{ so ist auch } h \circ f \text{ stetig}$$

Eigenschaften stetiger Funktionen

Im Folgenden sei $f : [a, b] \rightarrow \mathbb{R}$ stetig. Dann gilt:

(i) f ist beschränkt.

(ii) **Existenz des Maximums/Minimums.** Es existieren $x_{min}, x_{max} \in [a, b]$ mit

$$f(x_{min}) \leq f(x) \leq f(x_{max}) \quad \forall x \in [a, b].$$

Achtung: für die Existenz von x_{min} und x_{max} ist es wichtig, dass f auf einem **abgeschlossenen** Intervall definiert ist.

(iii) **Zwischenwertsatz.** Die Funktion f nimmt alle Werte zwischen Minimum und Maximum an, d.h. zu $c \in [f(x_{min}), f(x_{max})]$ gibt es ein $x_c \in [a, b]$ mit

$$f(x_c) = c.$$

Insbesondere

- Ist $f(a) \cdot f(b) \leq 0$, so besitzt f eine Nullstelle in $[a, b]$.
- Ist f ein Polynom ungeraden Grades, so besitzt f mindestens eine reelle Nullstelle.

(iv) Der Funktionsgraph einer stetigen Funktion weist keine Sprünge auf.

Kapitel 8

Differentialrechnung

Bekanntlich ist die Steigung einer Geraden $f(x) = mx + b$ definiert durch

$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} = \frac{mx_2 + b - (mx_1 + b)}{x_2 - x_1} = m.$$

Für allgemeine f definiert man die Steigung in x_0 als Grenzwert von Sekantensteigungen für $x \rightarrow x_0$.

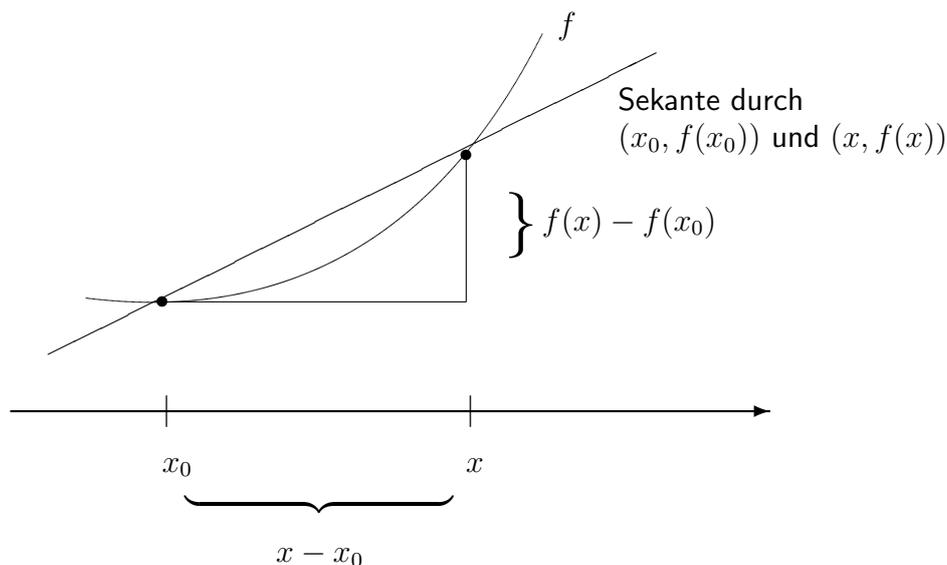


Abbildung 8.1: Geometrische Interpretation der Differenzenquotienten.

Die Steigung $\frac{f(x) - f(x_0)}{x - x_0}$ der Sekante bezeichnet man auch als Differenzenquotient. Für $x \rightarrow x_0$ geht die Sekante in die Tangente an f im Punkt $(x_0, f(x_0))$ über und damit die Sekantensteigung in die Steigung der Tangente an f in x_0 .

Definition Eine Funktion $f : D \rightarrow \mathbb{R}$ heißt **differenzierbar in** $x_0 \in D$, falls der Grenzwert

$$\frac{df}{dx}(x_0) := f'(x_0) := \lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

existiert. $f'(x_0)$ heißt **Ableitung** der Funktion f in x_0 und f heißt **differenzierbar** auf D , falls f differenzierbar in jedem Punkt $x_0 \in D$ ist.

Bemerkung Mit $h = x - x_0$ erhält man: f ist differenzierbar in x_0 , genau dann wenn

$$\lim_{h \rightarrow 0} \frac{f(x_0 + h) - f(x_0)}{h} = f'(x_0)$$

Beispiele

(i) Lineare Funktionen: $f(x) = mx + b \Rightarrow f'(x) = m$

Denn es gilt

$$f'(x) = \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h} = \lim_{h \rightarrow 0} \frac{m(x+h) + b - (mx+b)}{h} = m$$

(ii) $f(x) = x^2 \Rightarrow f'(x) = 2x$

Denn es gilt

$$f'(x) = \lim_{h \rightarrow 0} \frac{\overbrace{(x+h)^2 - x^2}^{\frac{1}{h}(x^2 + 2hx + h^2 - x^2)}}{h} = \lim_{h \rightarrow 0} 2x + h = 2x$$

(iii) Die Betragsfunktion $f(x) = |x|$ ist nicht differenzierbar in $x = 0$, denn

$$\frac{f(x) - f(0)}{x - 0} = \begin{cases} 1 & \text{für } x > 0 \\ -1 & \text{für } x < 0 \end{cases}$$

damit existiert kein Grenzwert $\lim_{x \rightarrow 0} \frac{f(x) - f(0)}{x - 0}$, also ist f nicht differenzierbar.

Bemerkung Ist f in x_0 differenzierbar, so ist f auch stetig in x_0 , denn

$$\lim_{x \rightarrow x_0} f(x) = \lim_{x \rightarrow x_0} \left(\underbrace{\frac{f(x) - f(x_0)}{x - x_0}}_{\rightarrow f'(x_0)} \cdot \underbrace{(x - x_0)}_{\rightarrow 0} \right) + f(x_0) = f(x_0)$$

Rechenregeln der Differentialrechnung

Sind $f, g : D \rightarrow \mathbb{R}$ in $x_0 \in D$ differenzierbar, so sind auch die folgenden Funktionen differenzierbar in x_0 :

(i) **Linearität** $\alpha f + \beta g$ für $\alpha, \beta \in \mathbb{R}$ mit

$$(\alpha f + \beta g)'(x_0) = \alpha f'(x_0) + \beta g'(x_0)$$

(ii) **Produktregel** $f \cdot g$ mit

$$(f \cdot g)'(x_0) = f'(x_0) \cdot g(x_0) + f(x_0) \cdot g'(x_0)$$

Tabelle 8.1: Ableitung elementarer Funktionen, Teil I

	$f(x)$	$f'(x)$	
Potenz- und	x^n	nx^{n-1}	für $n = 0, 1, 2, \dots$
Wurzelfunktionen	x^α	$\alpha x^{\alpha-1}$	für $\alpha \in \mathbb{R}, x > 0$
	\sqrt{x}	$\frac{1}{2\sqrt{x}}$	für $x > 0$
Exponential-	e^x	e^x	für $x \in \mathbb{R}$
und Logarithmusfunktion	$\ln x$	$\frac{1}{x}$	für $x > 0$
Trigonometrische	$\sin x$	$\cos x$	für $x \in \mathbb{R}$
Funktionen	$\cos x$	$-\sin x$	für $x \in \mathbb{R}$
	$\tan x$	$1 + (\tan x)^2$	für $x \neq (k + \frac{1}{2})\pi, k \in \mathbb{Z}$
Arkusfunktionen	$\arcsin x$	$\frac{1}{\sqrt{1-x^2}}$	für $x \in]-1, 1[$
	$\arccos x$	$-\frac{1}{\sqrt{1-x^2}}$	für $x \in]-1, 1[$
	$\arctan x$	$\frac{1}{1+x^2}$	für $x \in \mathbb{R}$

(iii) **Quotientenregel** Ist $g(x) \neq 0$ für $x \in D$, so ist auch $\frac{f}{g}$ differenzierbar in x_0 mit

$$\left(\frac{f}{g}\right)'(x_0) = \frac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{g^2(x_0)}$$

(iv) **Kettenregel** Ist $h : E \rightarrow \mathbb{R}$ differenzierbar und $g(D) \subset E$, so ist auch $f = h \circ g : D \rightarrow \mathbb{R}$ differenzierbar in x_0 und es gilt

$$f'(x_0) = (h \circ g)'(x_0) = \underbrace{h'(g(x_0))}_{\text{äußere}} \cdot \underbrace{g'(x_0)}_{\text{innere Ableitung}}$$

Ausschließlich als Merkhilfe geeignet ist die Aussage: " $\frac{df}{dx} = \frac{df}{dg} \cdot \frac{dg}{dx}$."

(v) **Ableitung der Umkehrfunktion** Ist $f : D \rightarrow \mathbb{R}$ streng monoton wachsend (bzw. fallend), so ist die Umkehrfunktion f^{-1} differenzierbar in $x_0 \in f(D)$ mit

$$(f^{-1})'(x_0) = \frac{1}{f'(f^{-1}(x_0))}$$

Das ist eine Anwendung der Kettenregel, denn aus $x = f(f^{-1}(x))$ für $x \in f(D)$ folgt durch differenzieren $1 = f'(f^{-1}(x)) \cdot (f^{-1})'(x)$.

Tabelle 8.2: Ableitung elementarer Funktionen, Teil II

	$f(x)$	$f'(x)$	
Hyperbelfunktionen	$\sinh x$	$\cosh x$	für $x \in \mathbb{R}$
	$\cosh x$	$\sinh x$	für $x \in \mathbb{R}$
	$\tanh x$	$1 - (\tanh x)^2$	für $x \in \mathbb{R}$
Areafunktionen	$\operatorname{arsinh} x$	$\frac{1}{\sqrt{x^2+1}}$	für $x \in \mathbb{R}$
	$\operatorname{arcosh} x$	$\frac{1}{\sqrt{x^2-1}}$	für $x > 1$
	$\operatorname{artanh} x$	$\frac{1}{1-x^2}$	für $x \in]-1, 1[$

Beispiele

zu (i) $(4 \ln x + 2 \cos x)' = 4 \frac{1}{x} - 2 \sin x$ für $x > 0$

zu (ii) $(x^2 \cdot e^x)' = 2xe^x + x^2e^x$

zu (iii) $(\tan x)' = \left(\frac{\sin x}{\cos x}\right)' = \frac{(\sin x)' \cdot \cos x - \sin x \cdot (\cos x)'}{(\cos x)^2} = \frac{(\cos x)^2 + (\sin x)^2}{(\cos x)^2} = 1 + (\tan x)^2$

zu (iv) a) $f(x) = e^{ax} = \exp(ax) = h(g(x))$, für $h(x) = e^x, g(x) = ax$. Also folgt

$$f'(x) = \underbrace{h'(g(x))}_{=e^{ax}} \cdot \underbrace{g'(x)}_{=a} = ae^{ax}$$

b) $f(x) = x^\alpha = \exp(\alpha \ln x) = h(g(x))$, für $h(x) = e^x, g(x) = \alpha \ln x$. Also folgt

$$f'(x) = \underbrace{h'(g(x))}_{=e^{\alpha \ln x}} \cdot \underbrace{g'(x)}_{=\alpha \cdot \frac{1}{x}} = \alpha \frac{1}{x} \cdot \exp(\alpha \ln x) = \alpha \cdot \frac{1}{x} x^\alpha = \alpha x^{\alpha-1}$$

zu (v) a) $(\ln x)' = \frac{1}{\underbrace{\exp'(\ln x)}_{=\exp(\ln x)=x}} = \frac{1}{x}$

b) $(\arcsin x)' = \frac{1}{\sin'(\arcsin x)} = \frac{1}{\cos(\arcsin x)} = \frac{1}{\sqrt{1-(\sin(\arcsin x))^2}} = \frac{1}{\sqrt{1-x^2}}$

Hyperbelfunktionen Die Hyperbelfunktionen sind folgendermaßen definiert:

$$\sinh x = \frac{1}{2}(e^x - e^{-x}) \quad \text{und} \quad \cosh x = \frac{1}{2}(e^x + e^{-x}),$$

sowie $\tanh x = \sinh x / \cosh x$. Die Umkehrfunktionen heißen Areafunktionen.

Mittelwertsatz Es sei $f : D \rightarrow \mathbb{R}$ differenzierbar und $[a, b] \subset D$. Dann gibt es eine Zwischenstelle $x \in]a, b[$ mit

$$\frac{f(b) - f(a)}{b - a} = f'(x)$$

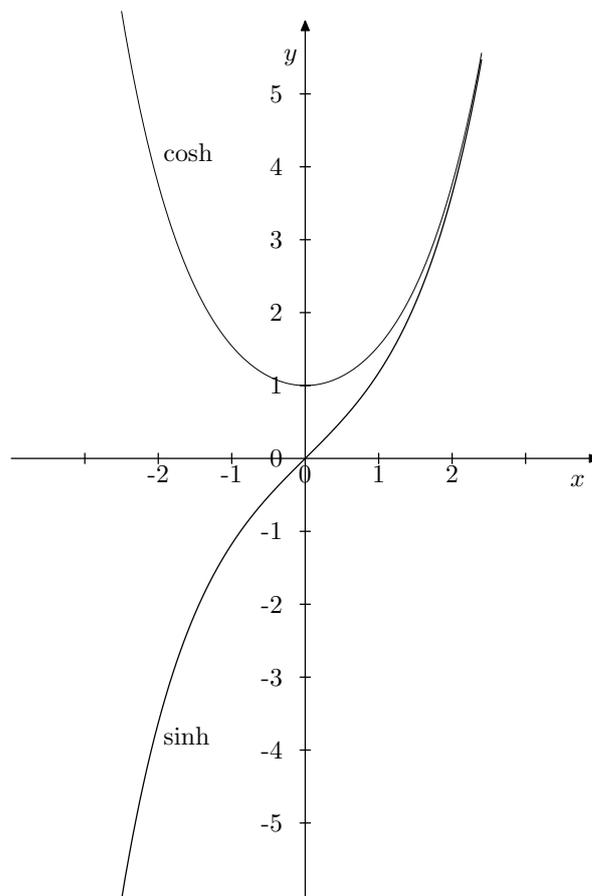


Abbildung 8.2: Funktionsgraphen von sinh und cosh.

Folgerungen aus dem Mittelwertsatz

Im Folgenden: $I \subset \mathbb{R}$ ein Intervall und $f : I \rightarrow \mathbb{R}$ differenzierbar.

- $f'(x) > 0$ ($f'(x) < 0$) auf $I \Rightarrow f$ streng monoton wachsend (fallend) auf I , denn es gilt

$$a, b \in I, a < b \Rightarrow \frac{f(b) - f(a)}{b - a} = f'(x) > 0 \Rightarrow f(b) > f(a)$$

- $f'(x) \geq 0$ ($f'(x) \leq 0$) auf $I \Rightarrow f$ monoton wachsend (fallend) auf I
- $f'(x) = 0$ auf $I \Rightarrow f$ konstant auf I

Regel von l'Hospital

Die Regel von l'Hospital kann dabei helfen, **Grenzwerte vom Typ $\frac{0}{0}$ (bzw. $\frac{\infty}{\infty}$)** zu bestimmen. Beispiele für solche Grenzwerte sind etwa

$$\lim_{x \rightarrow 0} \frac{\sin x}{x}, \quad \lim_{x \rightarrow \infty} \frac{x}{e^x}$$

d.h. Grenzwerte vom Typ

$$\lim_{x \rightarrow x_0} \frac{f(x)}{g(x)}$$

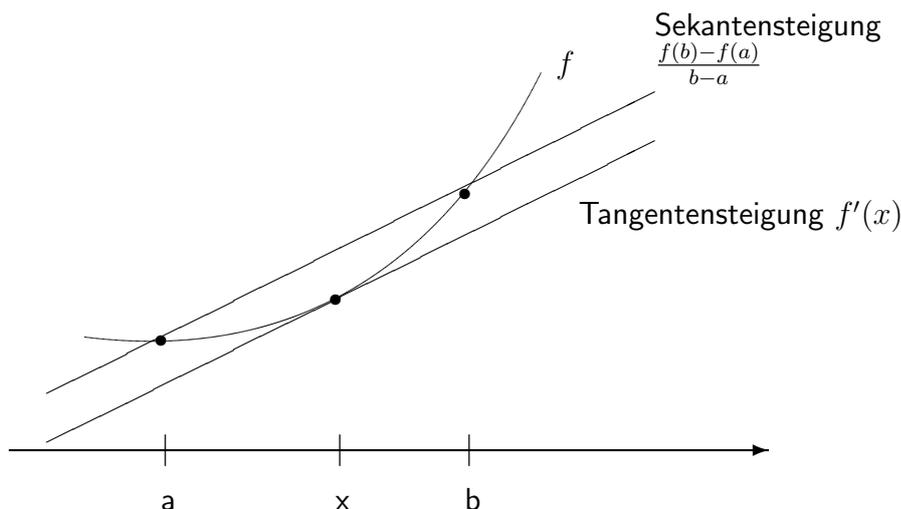


Abbildung 8.3: Geometrische Interpretation der Ableitung.

für Funktionen mit

$$\lim_{x \rightarrow x_0} f(x) = \lim_{x \rightarrow x_0} g(x) = 0 \quad (\text{bzw. } = \pm\infty)$$

Regel von l'Hospital Es sei $I \subset \mathbb{R}$ ein Intervall und $x_0 \in I$ oder Randpunkt von I (auch $x_0 = +\infty$ oder $x_0 = -\infty$ bei unbeschränkten Intervallen zugelassen) sowie

$$f, g : I \setminus x_0 \rightarrow \mathbb{R} \quad \text{differenzierbar}$$

Ist dann

$$\lim_{x \rightarrow x_0} f(x) = \lim_{x \rightarrow x_0} g(x) = 0 \quad (\text{bzw. } = \pm\infty)$$

und existiert $\lim_{x \rightarrow x_0} \frac{f'(x)}{g'(x)}$, so existiert auch $\lim_{x \rightarrow x_0} \frac{f(x)}{g(x)}$ und es gilt

$$\lim_{x \rightarrow x_0} \frac{f(x)}{g(x)} = \lim_{x \rightarrow x_0} \frac{f'(x)}{g'(x)}$$

Beispiele

$$\begin{aligned} \frac{0}{0}: \lim_{x \rightarrow 0} \frac{\sin x}{x} &= \lim_{x \rightarrow 0} \frac{\cos x}{1} = \cos(0) = 1 \\ \lim_{x \rightarrow 0} \frac{1 - \cos x}{x^2} &= \lim_{x \rightarrow 0} \frac{\sin x}{2x} = \lim_{x \rightarrow 0} \frac{\cos x}{2} = \frac{1}{2} \end{aligned}$$

Also: eine einmalige Anwendung der Regel von l'Hospital muss nicht zum Ziel führen.

$$\begin{aligned} \frac{\infty}{\infty}: \lim_{x \rightarrow \infty} \frac{x}{e^x} &= \lim_{x \rightarrow \infty} \frac{1}{e^x} = 0 \\ \lim_{x \rightarrow 0} \underbrace{x \ln x}_{= \frac{\ln x}{\frac{1}{x}}} &= \lim_{x \rightarrow 0} -\frac{1}{x^2} = \lim_{x \rightarrow 0} -x = 0 \end{aligned}$$

$$\lim_{x \rightarrow \infty} \frac{\ln x}{x} = \lim_{x \rightarrow \infty} \frac{\frac{1}{x}}{1} = 0$$

Höhere Ableitungen

Ist $f : D \rightarrow \mathbb{R}$ differenzierbar und die Ableitung $f' : D \rightarrow \mathbb{R}$ ebenfalls, so heißt f zweimal differenzierbar. Wir schreiben f'' oder $\frac{d^2 f}{dx^2}$ für die **zweite Ableitung** von f . Entsprechend definieren wir die höheren Ableitungen von f . Wir schreiben $f^{(n)}$ oder $\frac{d^n f}{dx^n}$ für die **n -te Ableitung** von f .

Beispiele

$$(i) \quad f(x) = x^2 \Rightarrow f^{(1)}(x) = 2x, f^{(2)} = 2, f^{(n)}(x) = 0, n \geq 3$$

$$(ii) \quad f(x) = e^x \Rightarrow f^{(1)}(x) = e^x = f(x) \Rightarrow f^{(n)}(x) = e^x \text{ für } n = 1, 2, 3, \dots$$

$$(iii) \quad f(x) = \sin x \Rightarrow f^{(1)}(x) = \cos x, f^{(2)}(x) = -\sin x = -f(x), \text{ usw.}$$

Die zweite Ableitung f'' bestimmt die **Krümmung** des Funktionsgraphen. In Punkten mit $f'' > 0$ ($f'' < 0$) ist der Funktionsgraph nach links (rechts) gekrümmt. Punkte x_0 mit $f''(x_0) = 0$ heißen entsprechend **Wendepunkte**.

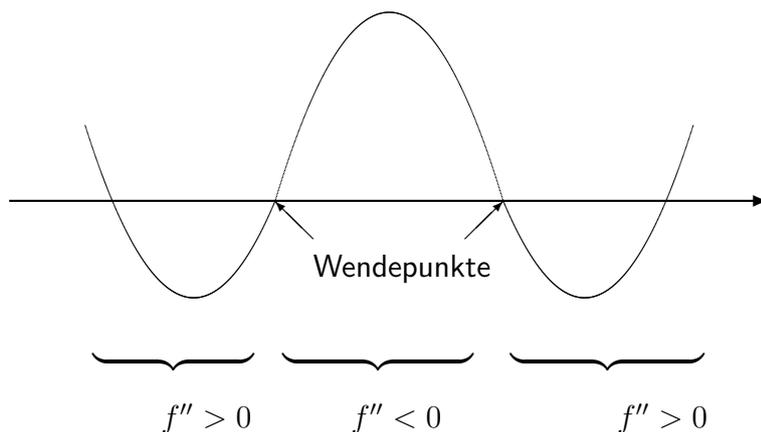


Abbildung 8.4: Krümmung einer Kurve.

Anwendung: Bestimmung von Extremalstellen

Definition Es sei $f : D \rightarrow \mathbb{R}$ eine Funktion. f hat in $x_0 \in D$ ein **lokales Maximum (Minimum)**, falls ein $\varepsilon > 0$ existiert mit

$$f(x_0) \geq f(x) \quad (\text{ bzw. } f(x_0) \leq f(x))$$

für alle $x \in D$ mit $|x - x_0| < \varepsilon$.

Gilt sogar $f(x_0) \geq f(x)$ (bzw. $f(x_0) \leq f(x)$) für alle $x \in D$, so heißt x_0 **absolutes Maximum (Minimum)**

Notwendiges Kriterium Ist $f : D \rightarrow \mathbb{R}$ differenzierbar und x_0 lokales Extremum (=Maximum oder Minimum) im Inneren von D , so ist $f'(x_0) = 0$.

Umgekehrt gilt:

Satz Ist f zweimal differenzierbar und $x_0 \in D$ mit $f'(x_0) = 0$, dann gilt

$$f''(x_0) > 0 \Rightarrow x_0 \text{ ist lokales Minimum}$$

$$f''(x_0) < 0 \Rightarrow x_0 \text{ ist lokales Maximum}$$

Verschwindet auch die 2. Ableitung in x_0 und gibt es ein $n \in \mathbb{N}$, $n > 2$ mit

$$f^{(2)}(x_0) = \dots = f^{(n-1)}(x_0) = 0, \text{ aber } f^{(n)}(x_0) \neq 0$$

so folgt

$$n \text{ gerade und } \begin{cases} f^{(n)}(x_0) > 0 \Rightarrow x_0 \text{ ist lokales Minimum} \\ f^{(n)}(x_0) < 0 \Rightarrow x_0 \text{ ist lokales Maximum} \end{cases}$$

n ungerade $\Rightarrow x_0$ ist kein lokales Extremum sondern Sattelpunkt (d.h. Wendepunkt mit waagerechter Tangente).

Beispiel $f(x) = x^n$, $n \geq 2$, besitzt in $x_0 = 0$ ein (lokales) Minimum für n gerade und einen Sattelpunkt für n ungerade.

Kapitel 9

Integralrechnung

9.1 Das bestimmte Integral

Es sei $f : [a, b] \rightarrow \mathbb{R}$ stetig und nicht negativ.

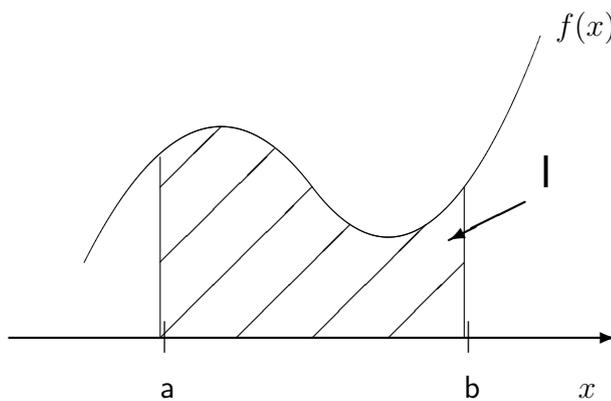


Abbildung 9.1: Flächeninhalt unter einer Kurve.

Wie berechnet man den Flächeninhalt I des schraffierten Bereichs?

Wähle eine **Zerlegung** Z von $[a, b]$ in n Teilintervalle:

$$[x_0, x_1], [x_1, x_2], \dots, [x_{n-1}, x_n]$$

wobei $a = x_0 < x_1 < \dots < x_{n-1} < x_n = b$

Für jedes Teilintervall $[x_{i-1}, x_i]$ sei

$$m_i := \text{Minimum von } f \text{ auf } [x_{i-1}, x_i]$$

$$M_i := \text{Maximum von } f \text{ auf } [x_{i-1}, x_i]$$

Hierzu bilden wir die **Untersumme** von f bzgl. Z

$$s_f(Z) := \sum_{i=1}^n m_i (x_i - x_{i-1})$$

und die **Obersumme** von f bzgl. Z

$$S_f(Z) := \sum_{i=1}^n M_i(x_i - x_{i-1}).$$

Offensichtlich gilt $s_f(Z) \leq I \leq S_f(Z)$. Verfeinert man Z durch weiteres Unterteilen der Intervalle $[x_{i-1}, x_i]$, so wächst offenbar die Untersumme und fällt die Obersumme.

Es sei

$$\begin{aligned} \delta(Z) &:= \text{größte Länge eines Teilintervalls von } Z \\ &= \underbrace{\max}_{\text{Maximum}} \{x_i - x_{i-1} : 1 \leq i \leq n\} \end{aligned}$$

die **Feinheit der Zerlegung** Z .

Lässt man $\delta(Z)$ gegen 0 konvergieren, so sind die Folgen der Unter- und Obersummen also jeweils monoton und (durch einander) beschränkt, also konvergent. Bemerkenswert hieran ist, dass für dieses Argument die Stetigkeit von f nicht notwendig ist.

In unserem Fall *ist* die Funktion f aber stetig, und dann gilt zusätzlich, dass diese beiden Grenzwerte übereinstimmen. d.h. es gilt

$$\lim_{\delta(Z) \rightarrow 0} s_f(Z) = I = \lim_{\delta(Z) \rightarrow 0} S_f(Z).$$

Wichtig Der gemeinsame Grenzwert I ist **unabhängig** von der Wahl der Folge der Zerlegungen Z .

Beispiel $f(x) = x$ auf $[0, 1]$. Zu $n \in \mathbb{N}$ wähle Zerlegung in Teilintervalle

$$\left[0, \frac{1}{n}\right], \left[\frac{1}{n}, \frac{2}{n}\right], \dots, \left[\frac{n-1}{n}, 1\right]$$

d.h. $x_i = \frac{i}{n}$. Da f monoton wachsend, folgt

$$m_i = f(x_{i-1}) = \frac{i-1}{n} \quad \text{und} \quad M_i = f(x_i) = \frac{i}{n}.$$

Also

$$s_f(Z) = \sum_{i=1}^n \underbrace{m_i}_{=\frac{i-1}{n}} \underbrace{(x_i - x_{i-1})}_{=\frac{1}{n}} = \frac{1}{n^2} \underbrace{\sum_{i=0}^{n-1} i}_{=\frac{1}{2}(n-1)n} = \frac{1}{2} \frac{n-1}{n} \xrightarrow{n \rightarrow \infty} \frac{1}{2}$$

Dasselbe gilt für die Folge der Obersummen:

$$S_f(Z) = \sum_{i=1}^n \underbrace{M_i}_{=\frac{i}{n}} \underbrace{(x_i - x_{i-1})}_{=\frac{1}{n}} = \frac{1}{n^2} \underbrace{\sum_{i=1}^n i}_{=\frac{1}{2}n(n+1)} = \frac{1}{2} \frac{n+1}{n} \xrightarrow{n \rightarrow \infty} \frac{1}{2}$$

Definition Es sei $f : [a, b] \rightarrow \mathbb{R}$ eine beschränkte Funktion. Gilt dann

$$I := \lim_{\delta(Z) \rightarrow 0} s_f(Z) = \lim_{\delta(Z) \rightarrow 0} S_f(Z)$$

so heißt f **(Riemann-)integrierbar (auf $[a, b]$)** und der gemeinsame Grenzwert I der Unter- bzw. Obersummen heißt das **(Riemann-) Integral von f (auf $[a, b]$)**.

Schreibweise

$$\int_a^b f(x) dx := I$$

Bezeichnungen

f = Integrand

x = Integrationsvariable

a/b = untere/obere Integrationsgrenze

Integrale über negative Funktionen sind negativ ("negativer Flächeninhalt"). Nimmt die Funktion sowohl negative als auch positive Werte an, so ergibt die Summe der positiven und negativen Flächeninhalte das Integral.

Vereinbarung Vertauscht man die Integrationsgrenzen, kehrt sich das Vorzeichen des Integrals um:

$$\int_a^b f(x) dx := - \int_b^a f(x) dx$$

Rechenregeln für bestimmte Integrale

(i) **Linearität**

$$\int_a^b (\alpha f(x) + \beta g(x)) dx = \alpha \int_a^b f(x) dx + \beta \int_a^b g(x) dx \quad \text{für } \alpha, \beta \in \mathbb{R}$$

(ii)

$$\int_a^b f(x) dx = \int_a^c f(x) dx + \int_c^b f(x) dx \quad \text{für } a \leq c \leq b$$

(iii) **Monotonie**

$$f(x) \leq g(x) \text{ für alle } x \in [a, b] \quad \Rightarrow \quad \int_a^b f(x) dx \leq \int_a^b g(x) dx$$

Satz Es sei $f : [a, b] \rightarrow \mathbb{R}$ stückweise stetig (d.h., es gibt eine Unterteilung $a = x_0 < x_1 < \dots < x_{n-1} < x_n = b$ von $[a, b]$, so dass f auf $]x_{i-1}, x_i[$ stetig und stetig fortsetzbar auf $[x_{i-1}, x_i]$ für alle i). Dann ist f (Riemann-)integrierbar.

Zum Beweis: Die Überlegungen zu Beginn dieses Abschnitts beweisen die Aussage für f stetig. Durch Zerlegung des Definitionsbereichs in Intervalle, in denen eine stückweise stetige Funktion stetig ist, ergibt sich die hier formulierte allgemeinere Behauptung.

9.2 Hauptsatz der Differential- und Integralrechnung

Die Berechnung bestimmter Integrale als Grenzwert von Untersummen bzw. Obersummen ist äußerst mühsam. Weitaus häufiger führt man die Berechnung bestimmter Integrale mit Hilfe des Hauptsatzes der Differential- und Integralrechnung auf die Berechnung von Stammfunktionen zurück:

Definition Eine differenzierbare Funktion $F : [a, b] \rightarrow \mathbb{R}$ heißt **Stammfunktion** von f , falls $F'(x) = f(x)$ für alle $x \in [a, b]$.

Wichtig Eine Stammfunktion von f ist nur bis auf eine Konstante eindeutig bestimmt, d.h. sind F, G Stammfunktionen von f , so gibt es eine Konstante $c \in \mathbb{R}$ mit $F(x) = G(x) + c$ für alle $x \in [a, b]$.

Beispiele

f	Stammfunktion zu f
x	$\frac{x^2}{2} + c$
$\cos x$	$\sin x + c$
e^x	$e^x + c$

Satz (Hauptsatz der Differential- und Integralrechnung)

Es sei $f : [a, b] \rightarrow \mathbb{R}$ stetig. Dann gilt:

(i) Ist F Stammfunktion von f , so gilt

$$\int_a^b f(x) dx = F(x) \Big|_a^b := F(b) - F(a)$$

(ii) Umgekehrt ist $F_a(x) := \int_a^x f(t) dt$, $x \in [a, b]$, eine Stammfunktion von f , d.h. es gilt

$$\frac{d}{dx} \left(\int_a^x f(t) dt \right) = f(x)$$

Jede andere Stammfunktion von f hat die Form

$$F(x) = F_a(x) + c$$

für eine Konstante $c \in \mathbb{R}$.

Der Hauptsatz besagt also, dass Differenzieren und Integrieren zueinander inverse Probleme sind (daher spricht man manchmal auch vom *Aufleiten* statt vom Integrieren einer Funktion). Der Hauptsatz **reduziert** die **Berechnung des Integrals** einer Funktion **auf** die **Bestimmung einer Stammfunktion**.

Beispiele

1) $\int_a^b x dx = \frac{1}{2}x^2 \Big|_a^b = \frac{b^2}{2} - \frac{a^2}{2}$. Insbesondere $\int_0^1 x dx = \frac{1}{2}$.

2) $\int_a^b \cos x dx = \sin x \Big|_a^b = \sin b - \sin a$. Also etwa $\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos x dx = \sin\left(\frac{\pi}{2}\right) - \sin\left(-\frac{\pi}{2}\right) = 2$.

$$3) \int_a^b e^x dx = e^x \Big|_a^b = e^b - e^a$$

Beweisskizze von Teil (ii) des Hauptsatzes Da f stetig ist, folgt

$$\begin{aligned} \frac{1}{h}(F_a(x+h) - F_a(x)) &= \frac{1}{h} \left(\int_a^{x+h} f(t) dt - \int_a^x f(t) dt \right) \\ &= \frac{1}{h} \int_x^{x+h} f(t) dt \quad \begin{cases} \leq \max_{t \in [x, x+h]} f(t) \\ \geq \min_{t \in [x, x+h]} f(t) \end{cases} \quad \rightarrow_{h \rightarrow 0} f(x). \end{aligned}$$

Berechnung von Integralen

Jede **Stammfunktion** F einer Funktion $f : [a, b] \rightarrow \mathbb{R}$ heißt **unbestimmtes Integral von f** . Man schreibt

$$F(x) = \int f(x) dx$$

Die unbestimmten Integrale der elementaren Funktionen erhält man also durch Umkehren der Tabelle für die Ableitungen aus Kapitel 8.

Tabelle 9.1: Stammfunktionen elementarer Funktionen.

$f(x)$	$\int f(x) dx$
$x^\alpha, \alpha \in \mathbb{R}, \alpha \neq -1$	$\frac{1}{\alpha+1} x^{\alpha+1}$
$\frac{1}{x}$	$\ln x$
$e^{\alpha x}, \alpha \in \mathbb{R}, \alpha \neq 0$	$\frac{1}{\alpha} e^{\alpha x}$
$\sin x$	$-\cos x$
$\cos x$	$\sin x$
$\frac{1}{1+x^2}$	$\arctan x$
$\frac{1}{\sqrt{1-x^2}}$	$\arcsin x$

Die unbestimmten Integrale weiterer elementarer Funktionen findet man in mathematischen Formelsammlungen (Stichwort *Integraltafeln*). Die Berechnung unbestimmter Integrale anderer Funktionen führt man möglichst mit Hilfe der folgenden **Integrationsmethoden** auf bekannte Integrale zurück:

(i) **Partielle Integration** (Produktintegration)

Es seien $f, g : [a, b] \rightarrow \mathbb{R}$ stetig differenzierbar. Die Produktregel der Differentialrechnung besagt:

$$\frac{d}{dx}(fg)(x) = f'(x)g(x) + f(x)g'(x)$$

Daher ist fg Stammfunktion von $f'g + fg'$ und somit gilt

$$\int f'(x)g(x) dx = f(x)g(x) - \int f(x)g'(x) dx$$

Für das bestimmte Integral gilt also:

$$\int_a^b f'(x)g(x) dx = f(x)g(x) \Big|_a^b - \int_a^b f(x)g'(x) dx$$

Beispiel

$$\int_a^b (\cos x)^2 dx \quad \underbrace{=}_{f(x)=\sin x, g(x)=\cos x} \quad \sin x \cos x \Big|_a^b - \underbrace{\int_a^b \sin x (-\sin x) dx}_{= \int_a^b (\sin x)^2 dx = \int_a^b 1 - (\cos x)^2 dx}$$

also

$$\int_a^b (\cos x)^2 dx = \frac{1}{2} \sin x \cos x \Big|_a^b + \frac{1}{2} \underbrace{\int_a^b 1 dx}_{=x \Big|_a^b} = \frac{1}{2} (\sin x \cos x + x) \Big|_a^b$$

(ii) Substitution

Sind F und g stetig differenzierbar, so folgt aus der Kettenregel der Differentialrechnung

$$\frac{d}{dx}(F \circ g)(x) = F'(g(x)) \cdot g'(x)$$

Es folgt also mit $f(x) = F'(x)$

$$\int f(g(x))g'(x) dx = (F \circ g)(x) = \int f(y) dy \Big|_{y=g(x)}$$

Merkhilfe Mit $y = g(x)$ ist $\frac{dy}{dx} = g'(x)$ also

$$dy = g'(x) dx$$

Für das bestimmte Integral gilt also die **Substitutionsregel**

$$\int_a^b f(g(x))g'(x) dx = F(g(b)) - F(g(a)) = \int_{g(a)}^{g(b)} f(y) dy$$

Die Substitutionsregel kann in zwei Richtungen angewandt werden:

A) Berechnung von $\int_a^b f(g(x))g'(x) dx$

a) Substitution $g(x) = y$ und $g'(x) dx = dy$

- b) Berechnung der Stammfunktion $\int f(y) dy = F(y)$
 c) Rückwärtssubstitution $y = g(x)$, $F(y) = F(g(x))$, also

$$\int_a^b f(g(x))g'(x) dx = F(g(b)) - F(g(a))$$

Beispiele

- 1) $\int_a^b \sin x \cos x dx = \int_a^b g(x)g'(x) dx = \int_{g(a)}^{g(b)} y dy = \frac{1}{2}y^2 \Big|_{g(a)}^{g(b)} = \frac{1}{2}((\sin b)^2 - (\sin a)^2)$
 für $g(x) = \sin x$ und $f(y) = y$.
 2) $\int_a^b \frac{x}{(x^2+1)^2} dx = \frac{1}{2} \int_a^b \frac{1}{g(x)^2} g'(x) dx = \frac{1}{2} \int_{g(a)}^{g(b)} \frac{1}{y^2} dy = -\frac{1}{2} \frac{1}{y} \Big|_{g(a)}^{g(b)} = \frac{1}{2} \left(\frac{1}{a^2+1} - \frac{1}{b^2+1} \right)$
 für $g(x) = x^2 + 1$ und $f(y) = \frac{1}{y^2}$.

B) Berechnung von $\int_a^b f(x) dx$

- a) Substitution $x = g(y)$ und $dx = g'(y) dy$ mit einer geeigneten **umkehrbaren** Funktion g
 b) Berechnung der Stammfunktion $\int f(g(y))g'(y) dy = H(y)$
 c) Auflösen von $x = g(y)$ nach y , d.h. $y = g^{-1}(x)$, also

$$\int_a^b f(x) dx = \int_{g^{-1}(a)}^{g^{-1}(b)} f(g(y))g'(y) dy = H(g^{-1}(b)) - H(g^{-1}(a))$$

Beispiel Berechnung der Fläche des Einheitskreises

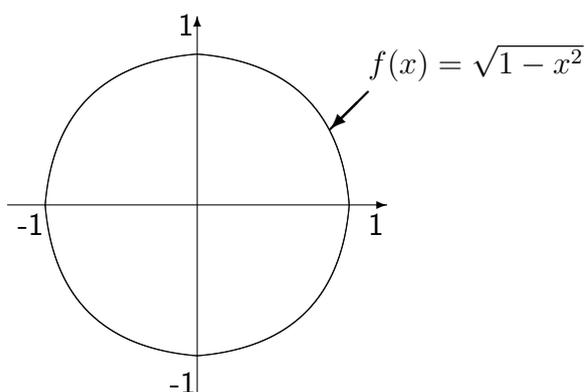


Abbildung 9.2: Der Einheitskreis als Vereinigung eines Funktionsgraphen und dessen Spiegelbilds.

Die Grafik macht deutlich, dass sich der gesuchte Flächeninhalt I schreiben lässt als

$$I = 2 \int_{-1}^{+1} \sqrt{1-x^2} dx$$

Substituiere $x = \sin y$, also $dx = \cos y \, dy$

$$\int \sqrt{1-x^2} \, dx = \int \underbrace{\sqrt{1-(\sin y)^2}}_{\cos y} \cos y \, dy = \int (\cos y)^2 \, dy = \frac{1}{2}(\sin y \cos y + y)$$

Die Rücksubstitution $y = \arcsin x$ führt auf

$$\begin{aligned} \int_{-1}^{+1} \sqrt{1-x^2} \, dx &= \frac{1}{2} (x \cos(\arcsin x) + \arcsin x) \Big|_{-1}^{+1} \\ &= \frac{1}{2} (x\sqrt{1-x^2} + \arcsin x) \Big|_{-1}^{+1} \\ &= \frac{1}{2} \underbrace{(\arcsin(1))}_{\frac{\pi}{2}} - \underbrace{(\arcsin(-1))}_{-\frac{\pi}{2}} = \frac{\pi}{2} \end{aligned}$$

Beim zweiten Gleichheitszeichen haben wir dabei verwandt, dass

$$\cos(\arcsin x) = \sqrt{1 - (\sin(\arcsin x))^2} = \sqrt{1 - x^2}.$$

Die Fläche I des Einheitskreises beträgt also π .

9.3 Uneigentliche Integrale

Bei der Integration können zwei Schwierigkeiten auftreten, die es erfordern, den Integrationsbegriff zu erweitern:

- (i) der Integrand besitzt eine Polstelle
- (ii) das Integrationsintervall ist unbeschränkt

Beispiel $f(x) = \frac{1}{\sqrt{x}}$ auf $]0, 1]$. Für $c \in]0, 1]$ gilt

$$\int_c^1 \frac{1}{\sqrt{x}} \, dx = 2\sqrt{x} \Big|_c^1 = 2 - 2\sqrt{c}$$

Also existiert $\lim_{c \rightarrow 0} \int_c^1 \frac{1}{\sqrt{x}} \, dx = 2$.

Dies motiviert:

Definition

- (i) $b \in \mathbb{R}$, $a < b$ (auch $a = -\infty$), $f :]a, b] \rightarrow \mathbb{R}$ integrierbar auf $[c, b]$ für alle $a < c \leq b$
 f heißt **uneigentlich integrierbar auf $]a, b]$** , falls

$$\lim_{c \rightarrow a} \int_c^b f(x) \, dx =: \int_a^b f(x) \, dx \quad \text{existiert.}$$

Man sagt auch, dass das uneigentliche Integral **konvergiert**.

(ii) Analog für $a \in \mathbb{R}, b > a$ (auch $b = +\infty$):

$$\lim_{c \rightarrow b} \int_a^c f(x) dx =: \int_a^b f(x) dx$$

(iii) Schließlich für $a < b$ (auch $a = -\infty, b = +\infty$): Ist

$$f :]a, b[\rightarrow \mathbb{R} \text{ integrierbar auf } [c, d]$$

für alle $a < c \leq d < b$, so heißt f **uneigentlich integrierbar auf** $]a, b[$, falls

$$\lim_{c \rightarrow a, d \rightarrow b} \int_c^d f(x) dx =: \int_a^b f(x) dx$$

existiert.

Beispiele

1) Es sei $\alpha \in [0, \infty[$

$$\int_0^1 \frac{dx}{x^\alpha} = \begin{cases} +\infty & \text{falls } \alpha \geq 1 \text{ (bestimmt divergent)} \\ \frac{1}{1-\alpha} & \text{falls } \alpha < 1 \text{ (konvergent)} \end{cases}$$

denn für $\alpha \neq 1$ gilt

$$\int_c^1 \frac{dx}{x^\alpha} = \frac{1}{1-\alpha} x^{1-\alpha} \Big|_c^1 = \frac{1}{1-\alpha} - \frac{1}{1-\alpha} c^{1-\alpha} \xrightarrow{c \rightarrow 0} \begin{cases} +\infty & \text{für } \alpha > 1 \\ \frac{1}{1-\alpha} & \text{für } \alpha < 1 \end{cases}$$

und für $\alpha = 1$

$$\int_c^1 \frac{dx}{x} = \ln x \Big|_c^1 = -\ln c \xrightarrow{c \rightarrow 0} +\infty$$

$$\int_1^\infty \frac{dx}{x^\alpha} = \begin{cases} \frac{1}{\alpha-1} & \text{falls } \alpha > 1 \text{ (konvergent)} \\ +\infty & \text{falls } \alpha \leq 1 \text{ (bestimmt divergent)} \end{cases}$$

denn für $\alpha \neq 1$ gilt

$$\int_1^c \frac{dx}{x^\alpha} = \frac{1}{1-\alpha} x^{1-\alpha} \Big|_1^c = \frac{1}{1-\alpha} c^{1-\alpha} - \frac{1}{1-\alpha} \xrightarrow{c \rightarrow \infty} \begin{cases} \frac{1}{\alpha-1} & \text{falls } \alpha > 1 \\ +\infty & \text{falls } \alpha < 1 \end{cases}$$

und für $\alpha = 1$

$$\int_1^c \frac{dx}{x} = \ln c \xrightarrow{c \rightarrow \infty} +\infty$$

2) Für $\alpha > 0$ ist $\int_0^\infty e^{-\alpha x} dx = \frac{1}{\alpha}$, denn

$$\int_0^c e^{-\alpha x} dx = -\frac{1}{\alpha} e^{-\alpha x} \Big|_0^c = -\frac{1}{\alpha} e^{-\alpha c} + \frac{1}{\alpha} \xrightarrow{c \rightarrow \infty} \frac{1}{\alpha}$$

3) $\int_0^\infty x e^{-\frac{x^2}{2}} dx = 1$, denn mit $y = \frac{x^2}{2}$ gilt

$$\int_0^\infty x e^{-\frac{x^2}{2}} dx = \int_0^\infty e^{-y} dy = 1.$$

4) $\int_e^\infty \frac{dx}{x(\ln x)^2} = 1$, denn mit $y = \ln x$ gilt

$$\int_e^\infty \frac{dx}{x(\ln x)^2} = \int_1^\infty \frac{1}{y^2} dy = 1.$$

Ist eine Stammfunktion zu f nicht bekannt, kann man die Existenz des uneigentlichen Integrals manchmal auch mit folgendem Satz rechtfertigen:

Satz (Vergleichskriterium) $a \in \mathbb{R}$, $a < b$ (auch $b = +\infty$). Die Funktionen f und g seien integrierbar auf $[a, c]$ für alle $a \leq c < b$. Dann gilt:

(i) Ist $|f(x)| \leq g(x)$ für alle $x \in [a, b[$ und konvergiert $\int_a^b g(x) dx$, so konvergieren auch die uneigentlichen Integrale

$$\int_a^b f(x) dx \quad \text{und} \quad \int_a^b |f(x)| dx.$$

(ii) Ist $0 \leq g(x) \leq f(x)$ für alle $x \in [a, b[$ und divergiert $\int_a^b g(x) dx$, so divergiert auch $\int_a^b f(x) dx$.

Beispiele $\int_1^\infty \frac{\sin x}{x^2} dx$ konvergiert, denn

$$\left| \frac{\sin x}{x^2} \right| \leq \frac{1}{x^2} \quad \text{und} \quad \int_1^\infty \frac{1}{x^2} dx \quad \text{konvergiert.}$$

$\int_1^\infty \frac{3x+2}{1+x+2x^5} dx$ konvergiert, denn

$$\left| \frac{3x+2}{1+x+2x^5} \right| \leq \frac{3}{2} \frac{1}{x^4} + \frac{1}{x^5} \quad \text{und} \quad \int_1^\infty \frac{3}{2x^4} + \frac{1}{x^5} dx \quad \text{konvergiert.}$$

Anwendung: Integralkriterium für Reihen

Ist $f : [1, \infty[\rightarrow [0, \infty[$ monoton fallend, so haben

$$\sum_{k=1}^{\infty} f(k) \quad \text{und} \quad \int_1^{\infty} f(x) dx$$

dasselbe Konvergenzverhalten, d.h. beide Terme sind entweder konvergent oder bestimmt divergent.

Beispiel Für $\alpha \in [0, \infty[$ haben

$$\sum_{k=1}^{\infty} \frac{1}{k^\alpha} \quad \text{und} \quad \int_1^{\infty} \frac{dx}{x^\alpha} \quad \text{dasselbe Konvergenzverhalten.}$$

Also gilt

$$\sum_{k=1}^{\infty} \frac{1}{k^\alpha} \text{ ist } \begin{cases} \text{konvergent für } \alpha > 1 \\ \text{bestimmt divergent für } \alpha \leq 1 \end{cases}$$

Kapitel 10

Potenzreihen und Taylorreihen

Reihen der Form

$$\sum_{k=0}^{\infty} a_k(x - x_0)^k = a_0 + a_1(x - x_0) + a_2(x - x_0)^2 + \dots \quad (10.0.1)$$

mit $x \in \mathbb{R}$ **variabel** und $x_0, a_k \in \mathbb{R}$ **konstant**, heißen **Potenzreihen**.

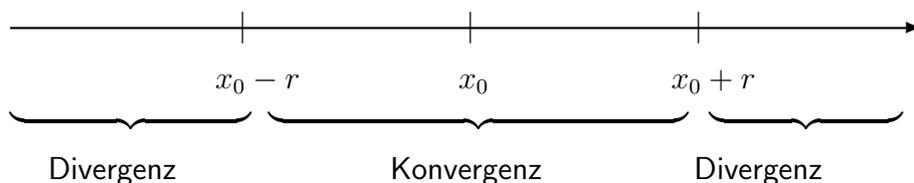
- a_0, a_1, a_2, \dots heißen **Koeffizienten**
- x_0 heißt **Entwicklungspunkt** der Potenzreihe
- Der **Konvergenzbereich** der Potenzreihe (10.0.1) ist die Menge

$$M = \left\{ x \in \mathbb{R} : \sum_{k=0}^{\infty} a_k(x - x_0)^k \text{ konvergiert} \right\}$$

Ist $M \neq \mathbb{R}$, so gibt es ein $r \geq 0$ mit

$$|x - x_0| < r \quad \Rightarrow \quad \sum_{k=0}^{\infty} a_k(x - x_0)^k \quad \text{konvergiert}$$

$$|x - x_0| > r \quad \Rightarrow \quad \sum_{k=0}^{\infty} a_k(x - x_0)^k \quad \text{divergiert}$$



Für $|x - x_0| = r$ kann die Reihe konvergieren oder divergieren. Die Zahl r heißt **Konvergenzradius** der Potenzreihe $\sum_{k=0}^{\infty} a_k(x - x_0)^k$.

Beispiele

- (i) Die geometrische Reihe $\sum_{k=0}^{\infty} x^k$ hat den Konvergenzradius $r = 1$, denn für $|x| < 1$ ist die Reihe konvergent und für $|x| \geq 1$ divergent (siehe Kapitel 6).
- (ii) Die Reihe $\sum_{k=1}^{\infty} \frac{(-1)^k}{k} x^k$ hat den Konvergenzradius 1, denn

$$\lim_{k \rightarrow \infty} \frac{\frac{|x|^{k+1}}{k+1}}{\frac{|x|^k}{k}} = \lim_{k \rightarrow \infty} |x| \cdot \frac{k}{k+1} = |x|$$

Also folgt aus dem Quotientenkriterium, dass $\sum_{k=1}^{\infty} \frac{(-1)^k}{k} x^k$ für $|x| < 1$ konvergiert und für $|x| > 1$ divergiert.

- Für $x = 1$ konvergiert die Reihe aufgrund des Leibniz-Kriteriums für alternierende Reihen.
 - Für $x = -1$ ist die Reihe divergent (harmonische Reihe).
- (iii) Die Reihe $\sum_{k=0}^{\infty} \frac{x^k}{k!}$ ist für alle $x \in \mathbb{R}$ konvergent und damit ist $r = \infty$.
- (iv) Die Reihe $\sum_{k=0}^{\infty} k! x^k$ hat Konvergenzradius 0. Sie konvergiert nur für $x = 0$.

Für den Konvergenzradius r der Potenzreihe $\sum_{k=0}^{\infty} a_k (x - x_0)^k$ gilt die Formel

$$r = \frac{1}{\lim_{k \rightarrow \infty} \sqrt[k]{|a_k|}}$$

Beispiel $\sum_{k=0}^{\infty} kx^k$ hat Konvergenzradius +1, denn

$$\lim_{k \rightarrow \infty} \sqrt[k]{k} = \lim_{k \rightarrow \infty} k^{\frac{1}{k}} = \lim_{k \rightarrow \infty} e^{\frac{1}{k} \ln k} = e^0 = 1$$

und damit folgt

$$r = \frac{1}{\lim_{k \rightarrow \infty} \sqrt[k]{k}} = 1$$

Eigenschaften von Potenzreihen

Potenzreihen konvergieren **innerhalb** des Konvergenzbereiches **absolut** und zwei Potenzreihen dürfen im gemeinsamen Konvergenzbereich **gliedweise addiert, subtrahiert** und **multipliziert** werden.

Es sei $\sum_{k=0}^{\infty} a_k (x - x_0)^k$ eine Potenzreihe mit Konvergenzradius r . Die durch

$$f :]x_0 - r, x_0 + r[\rightarrow \mathbb{R}, \quad f(x) = \sum_{k=0}^{\infty} a_k (x - x_0)^k$$

definierte Funktion heißt die durch die Potenzreihe $\sum_{k=0}^{\infty} a_k (x - x_0)^k$ **dargestellte Funktion**. Umgekehrt nennt man $\sum_{k=0}^{\infty} a_k (x - x_0)^k$ die **Potenzreihendarstellung von f (im Entwicklungspunkt x_0)**.

Die Funktion f ist **beliebig oft differenzierbar** und ihre Ableitungen erhält man durch **gliedweises Differenzieren** der Potenzreihe, also insbesondere

$$f'(x) = f^{(1)}(x) = \sum_{k=1}^{\infty} k a_k (x - x_0)^{k-1}.$$

Für die höheren Ableitungen gilt:

$$f^{(n)}(x) = \sum_{k=n}^{\infty} \underbrace{k \cdot (k-1) \cdot \dots \cdot (k-(n-1))}_{n\text{-mal}} a_k (x - x_0)^{k-n}$$

Die Reihe $\sum_{k=1}^{\infty} k a_k (x - x_0)^{k-1}$ hat denselben Konvergenzradius wie die Reihe $\sum_{k=0}^{\infty} a_k (x - x_0)^k$, denn

$$\lim_{k \rightarrow \infty} \sqrt[k]{k|a_k|} = \lim_{k \rightarrow \infty} \underbrace{\sqrt[k]{k}}_{=1} \lim_{k \rightarrow \infty} \sqrt[k]{|a_k|} = \lim_{k \rightarrow \infty} \sqrt[k]{|a_k|}.$$

Beispiel

Die durch $\sum_{k=0}^{\infty} \frac{x^k}{k!}$ dargestellte Funktion f ist für alle $x \in \mathbb{R}$ differenzierbar und es gilt

$$f'(x) = \sum_{k=1}^{\infty} k \frac{x^{k-1}}{k!} = \sum_{k=1}^{\infty} \frac{x^{k-1}}{(k-1)!} = \sum_{k=0}^{\infty} \frac{x^k}{k!} = f(x).$$

In der Tat ist f die Exponentialfunktion, d.h. es gilt

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!} \quad \text{für } x \in \mathbb{R}.$$

Weitere wichtige Potenzreihendarstellungen

$$\sin x = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!} x^{2k+1} = x - \frac{x^3}{3!} + \frac{x^5}{5!} \mp \dots \quad \text{für } x \in \mathbb{R}$$

$$\cos x = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!} x^{2k} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} \mp \dots \quad \text{für } x \in \mathbb{R}$$

$$\ln(1+x) = \sum_{k=0}^{\infty} \frac{(-1)^k}{k+1} x^{k+1} = x - \frac{x^2}{2} + \frac{x^3}{3} \mp \dots \quad \text{für } |x| < 1$$

$$\arctan x = \sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1} x^{2k+1} = x - \frac{x^3}{3} + \frac{x^5}{5} \mp \dots \quad \text{für } |x| < 1$$

Analog zur Differentiation gilt, dass Potenzreihen innerhalb ihres Konvergenzbereiches gliedweise integriert werden dürfen, d.h. für

$$f(x) = \sum_{k=0}^{\infty} a_k (x - x_0)^k$$

mit Konvergenzradius r gilt:

$$F(x) = \sum_{k=0}^{\infty} \frac{a_k}{k+1} (x-x_0)^{k+1}$$

ist Stammfunktion von f , sie hat denselben Konvergenzradius. Insbesondere folgt hieraus für bestimmte Integrale

$$\int_a^b f(x) dx = \sum_{k=0}^{\infty} a_k \int_a^b (x-x_0)^k dx = \sum_{k=0}^{\infty} \frac{a_k}{k+1} (x-x_0)^{k+1} \Big|_a^b$$

Taylorreihen

Die Potenzreihenentwicklung einer Funktion kann äußerst nützlich sein. Daher betrachtet man für beliebige Funktionen

- Approximationen durch Polynome
- Entwicklungen in Potenzreihen

Es sei dazu $f :]a, b[\rightarrow \mathbb{R}$ n -mal stetig differenzierbar und $x_0 \in]a, b[$.

Ansatz zur Approximation

Bestimme ein Polynom

$$p(x) = a_0 + a_1(x-x_0) + a_2(x-x_0)^2 + \dots + a_n(x-x_0)^n$$

mit $p^{(k)}(x_0) = f^{(k)}(x_0)$ für $k = 0, 1, \dots, n$. Das heißt, die ersten n Ableitungen von f und p sollen im Punkt x_0 übereinstimmen. Notwendigerweise gilt dann

$$\begin{aligned} p(x_0) &= a_0, \text{ also } a_0 = f(x_0) \\ p'(x_0) &= a_1, \text{ also } a_1 = f'(x_0) \\ p^{(2)}(x_0) &= 2a_2, \text{ also } a_2 = \frac{f^{(2)}(x_0)}{2} \\ &\vdots \\ p^{(n)}(x_0) &= n! a_n, \text{ also } a_n = \frac{f^{(n)}(x_0)}{n!} \end{aligned}$$

Im Allgemeinen ist natürlich $f \neq p$, aber durch die obige Approximation sollte der Fehler (bzw. das Restglied)

$$R_{n+1}(x) = f(x) - p(x)$$

für wachsendes n gegen 0 konvergieren.

Satz Es sei $f :]a, b[\rightarrow \mathbb{R}$ $(n+1)$ -mal stetig differenzierbar und $x_0 \in]a, b[$. Dann gilt für alle $x \in]a, b[$ die **Taylorformel**:

$$f(x) = f(x_0) + \frac{f'(x_0)}{1!} (x-x_0) + \frac{f''(x_0)}{2!} (x-x_0)^2 + \dots + \frac{f^{(n)}(x_0)}{n!} (x-x_0)^n + R_{n+1}(x)$$

mit

$$R_{n+1}(x) = \frac{1}{n!} \int_{x_0}^x (x-t)^n f^{(n+1)}(t) dt$$

bzw. für ein $\xi \in [x_0, x]$

$$R_{n+1}(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!} (x-x_0)^{n+1} \quad \text{(Lagrangesche Restgliedformel).}$$

Im **Spezialfall** $x_0 = 0$ spricht man auch von der **MacLaurinschen Formel**.

Das Polynom $f(x_0) + \frac{f^{(1)}(x_0)}{1!}(x-x_0) + \dots + \frac{f^{(n)}(x_0)}{n!}(x-x_0)^n$ heißt **n -tes Taylorpolynom von f im Entwicklungspunkt x_0** .

Ist $f^{(n+1)}$ beschränkt durch eine Konstante M , d.h.

$$|f^{(n+1)}(x)| \leq M \text{ für alle } x \in]a, b[$$

so folgt

$$|f(x) - p(x)| = |R_{n+1}(x)| \leq \underbrace{\frac{M}{(n+1)!} |x-x_0|^{n+1}}_{\text{i.A. sehr gute Abschätzung}}$$

Beispiele

(i) $f(x) = e^x \Rightarrow f^{(n)}(x) = e^x$ insbesondere $f^{(n)}(0) = e^0 = 1$. Für den Entwicklungspunkt $x_0 = 0$ und $n \in \mathbb{N}$ folgt

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots + \frac{x^n}{n!} + R_{n+1}(x)$$

mit $R_{n+1}(x) = e^\xi \frac{x^{n+1}}{(n+1)!}$ für ein $\xi \in [0, x]$.

Folglich gilt

$$|R_{n+1}(x)| \leq \frac{|e^\xi|}{(n+1)!} |x|^{n+1} \leq \frac{e^{|x|}}{(n+1)!} |x|^{n+1}.$$

Man erkennt: $\lim_{n \rightarrow \infty} |R_{n+1}(x)| = 0$, also konvergiert die Folge der Taylorreihen für alle x gegen e^x und das sogar außerordentlich schnell.

(ii) Für $f(x) = \sin x$ erhält man

$$\sin^{(2n)}(0) = 0, \quad \sin^{(2n+1)}(0) = (-1)^n$$

und damit

$$\begin{aligned} \sin x &= \sin(0) + \frac{\sin^{(1)}(0)}{1!}x + \frac{\sin^{(2)}(0)}{2!}x^2 + \dots \\ &= x - \frac{x^3}{3!} + \frac{x^5}{5!} \mp \dots + \frac{(-1)^n}{(2n+1)!}x^{2n+1} + R_{2n+2}(x) \end{aligned}$$

mit

$$|R_{2n+2}(x)| = \left| \frac{\sin^{(2n+2)}(\xi)}{(2n+2)!} x^{2n+2} \right| \leq \frac{|x|^{2n+2}}{(2n+2)!}$$

Ist f beliebig oft differenzierbar, so können wir für $x_0 \in]a, b[$ auch die unendliche Potenzreihe

$$T_f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k$$

bilden. T_f heißt **Taylorreihe von f mit Entwicklungspunkt x_0** (und im Spezialfall $x_0 = 0$ auch **MacLaurinsche Reihe von f**).

Vorsicht

- 1) Der Konvergenzradius von T_f kann 0 sein.
- 2) Falls die Taylorreihe konvergiert, so konvergiert sie nicht notwendigerweise gegen f .
- 3) Für $x \in]a, b[$ konvergiert $T_f(x)$ gegen $f(x)$ genau dann wenn $R_n(x)$ gegen 0 konvergiert.

Ein hinreichendes Kriterium für die Konvergenz der Taylorreihe T_f gegen f ist

$$|f^{(n)}(x)| \leq A \cdot B^n \text{ für alle } x \in]a, b[$$

für Konstanten A, B unabhängig von n .

Ist der Konvergenzradius r von $T_f > 0$ und gilt $T_f(x) = f(x)$ für $|x - x_0| < r$, so sagt man:

f lässt sich um x_0 als Taylorreihe darstellen bzw. in eine Taylorreihe entwickeln.

Für Potenzreihen $f(x) = \sum_{k=0}^{\infty} a_k (x - x_0)^k$ mit Konvergenzradius $r > 0$ gilt speziell:

Die Taylorreihe T_f von f stimmt mit f überein, d.h.

$$a_k = \frac{f^{(k)}}{k!}, \quad k = 0, 1, 2, \dots$$