

EINFÜHRUNG IN
DIE MATHEMATIK FÜR



INGENIEURE UND NATURWISSENSCHAFTLER

Michael Bildhauer

Teil II

Matrizen

UND

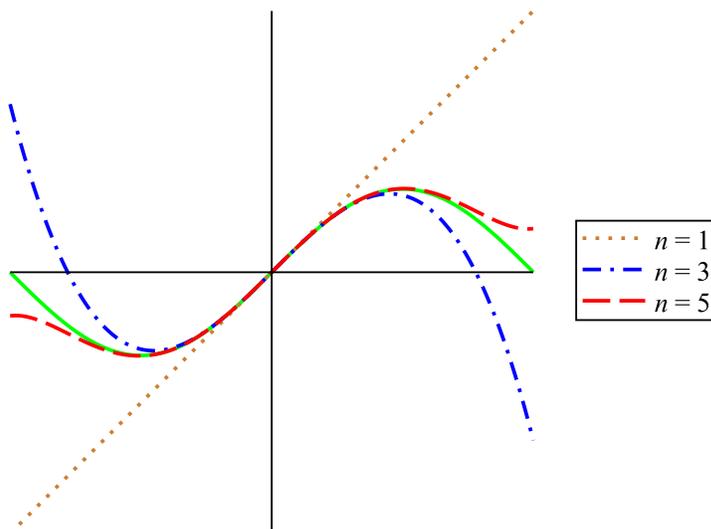
Lineare Gleichungssysteme

Lineare Abbildungen

Differential-

UND

Integralrechnung I



Inhaltsverzeichnis

1	Matrizen und lineare Gleichungssysteme	7
1.1	Matrizenkalkül (Vektorraum $M(n,m)$; Matrixmultiplikation; Transposition; Spalten- und Zeilenvektoren)	7
1.2	Zur Lösbarkeit linearer Gleichungssysteme (homogenes und inhomogenes System; Superposition von Lösungen; Struktur der Lösungsmenge; Kern einer Matrix; Rang einer Matrix; erweiterte Matrix; Dimensionsformel)	17
1.3	Überbestimmte lineare Gleichungssysteme – lineare Ausgleichsprobleme (Residuum; Methode der kleinsten Quadrate; Ausgleich nach Tschebyscheff; Normalgleichung)	24
1.4	Übungsaufgaben zu Kapitel 1	31
2	Quadratische Matrizen – Inverse und Determinante	39
2.1	Invertierbare Matrizen (reguläre und singuläre Matrizen; Dreiecksmatrizen; LR-Zerlegung; Cholesky-Zerlegung)	39
2.2	Die Determinante (orientiertes Volumen; Regel von Sarrus, Laplacescher Entwicklungssatz; alternierende n -Linearform; Multiplikationssatz; Cramersche Regel; Orientierung; orthogonale Matrizen)	47
2.3	Übungsaufgaben zu Kapitel 2	61
3	Lineare Abbildungen	67
3.1	Definition und erste Eigenschaften (Kern und Bild einer linearen Abbildung; Rang; Rangsatz; Injektivität, Surjektivität und Bijektivität einer linearen Abbildung)	68
3.2	Matrixdarstellung (Koordinaten; darstellende Matrix; Basiswechsel; Koordinatentransformation; Transformation der darstellenden Matrix; Tensor; kovariant; kontravariant)	72
3.3	Übungsaufgaben zu Kapitel 3	85

4	Stetige Funktionen	91
4.1	Beispiele, Definition und erste Eigenschaften (Lipschitz-stetige Funktion; Grenzwert und Stetigkeit einer Funktion; gleichmäßige Stetigkeit)	91
4.2	Zwei Sätze über stetige Funktionen (Stetigkeit und Kompaktheit; Zwischenwertsatz; Stetigkeit der Umkehrfunktion)	101
4.3	Übungsaufgaben zu Kapitel 4	105
5	Differentialrechnung in einer Veränderlichen	111
5.1	Grundlagen (Differenzenquotient; Ableitung; Differenzierbarkeit und Stetigkeit; Produkt- Quotientenregel; Kettenregel; Ableitung der Umkehrfunktion; höhere Ableitungen)	111
5.2	Lokale und globale Extrema, Mittelwertsatz (notwendige und hinreichende Bedingung für lokale Extrema; Suche nach globalen Extrema; Satz von Rolle; Mittelwertsatz; Ableitung und Monotonie; konvexe und konkave Funktionen; Regeln von l'Hospital)	125
5.3	Numerische Differentiation (Extrapolation zum Limes $h \rightarrow 0$; zentraler Differenzenquotient)	141
5.4	Übungsaufgaben zu Kapitel 5	149
6	Integralrechnung in einer Veränderlichen	161
6.1	Das bestimmte Riemannsche Integral (Zerlegung; Feinheit; Untersumme; Obersumme; Unterintegral; Oberintegral; Riemann integrierbare Funktionen; Integrabilitätskriterium; Rechenregeln für integrierbare Funktionen; Integration auf Teilintervallen; orientiertes Riemannsches Integral; Flächeninhalt)	161
6.2	Der Hauptsatz der Differential- und Integralrechnung (Stammfunktion; unbestimmtes Integral)	170
6.3	Integrationstechniken (einfache Integrationstechniken; partielle Integration; Substitutionsregel; Partialbruchzerlegung)	175
6.3.1	Einfache Integrationstechniken	175
6.3.2	Partielle Integration	176
6.3.3	Substitutionsregel	177
6.3.4	Partialbruchzerlegung	180
6.4	Uneigentliche Integrale (lokal integrierbare Funktionen; Konvergenzkriterien)	185

6.5	Numerische Integration (Newton-Cotes Formeln; Newton-Cotes Summenformeln; Integrationsverfahren von Romberg)	193
6.6	Übungsaufgaben zu Kapitel 6	203
7	Der Satz von Taylor	211
7.1	Taylor-Formel und Taylor-Reihe (Taylor-Polynom; Restglied; Integraldarstellung des Restgliedes; Lagrangesche Restgliedformel; die Klasse C^∞ ; reell analytische Funktionen)	211
7.2	Übungsaufgaben zu Kapitel 7	223
8	Fourier-Reihen	227
8.1	Einführung (Spektrum; harmonische Analyse; Periode einer Funktion; trigonometrische Reihen; trigonometrische Polynome; Fourier-Koeffizienten)	227
8.2	Der Satz (Sägezahnfunktion; Gibbs-Phänomen; stückweise glatte Funktion; mittlere quadratische Abweichung; Parsevalsche Gleichung)	235
8.3	Übungsaufgaben zu Kapitel 8	243
A	Tensoren: Weitere Bemerkungen zur Definition	245
A.1	Kovariante und kontravariante Tensoren (kovariant; kontravariant; lineares Funktional; Dualraum; Multilinearform; Tensorprodukt; Tensor zweiter Stufe)	245
B	Konditionierung einer numerischen Aufgabe – Stabilität eines Algorithmus	253
B.1	Konditionierung einer numerischen Aufgabe (differentielle Fehleranalyse; Landausche Symbole; relative Konditionszahlen; Problemfehler)	253
B.2	Stabilität eines Algorithmus (gutartiger Algorithmus)	259
B.3	Übungsaufgaben zum Anhang	267
	Literaturverzeichnis	273
	Index	275

Kapitel 1

Matrizen und lineare Gleichungssysteme

1.1 Matrizenkalkül (Vektorraum $M(n,m)$; Matrixmultiplikation; Transposition; Spalten- und Zeilenvektoren)

Matrizen sind im Prinzip schon bei der schematischen Lösung linearer Gleichungssysteme mithilfe des Gaußschen Eliminationsverfahrens aufgetaucht. Sie werden auch eine zentrale Rolle bei der Diskussion **linearer Abbildungen** und insbesondere als Repräsentant von Ableitungen im Fall von Funktionen mehrerer Veränderlicher spielen.

In diesem Abschnitt werden nach der Definition einer Matrix zunächst die wichtigsten Operationen mit Matrizen vorgestellt. Dabei handelt es sich **o.E. stets um reelle Matrizen**. Alle Betrachtungen übertragen sich unmittelbar auf Matrizen mit beispielsweise komplexen Einträgen.

Definition 1.1. $n \times m$ MATRIZEN

Es seien $n, m \in \mathbb{N}$. Ein $n \times m$ Koeffizientenschema (das ist nicht anderes als eine **Tabelle aus n Zeilen und m Spalten**) der Form

$$(a_{ij})_{i=1, \dots, n}^{j=1, \dots, m} := A := \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{pmatrix},$$

$a_{ij} \in \mathbb{R}$ für $1 \leq i \leq n$, $1 \leq j \leq m$, heißt eine $n \times m$ Matrix.

Notation:

- i) Können bzgl. der Zeilenzahl und der Spaltenzahl keine Missverständnisse auftreten, so schreibt man oft abkürzend (a_{ij}) für die Matrix.*
- ii) Die Menge der $n \times m$ -Matrizen wird mit $M(n, m)$ bezeichnet; weitere Bezeichnungen: $\mathbb{R}^{(n,m)}$, $\mathbb{R}^{n \times m}$.*

Merke. Der Eintrag a_{ij} steht in der i^{ten} Zeile und in der j^{ten} Spalte der Matrix (a_{ij}) .

Bereits bekannte Beispiele.

- i) Die Elemente des \mathbb{R}^n (Spaltenvektoren) können als $n \times 1$ Matrizen aufgefasst werden:*

$$\underline{\mathbf{x}} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in M(n, 1) .$$

- ii) Das Koeffizientenschema eines linearen Gleichungssystems (vgl. Gaußsches Eliminationsverfahren, Teil I, Anhang A) ohne die rechte Seite ist eine $n \times m$ Matrix.*

Operationen mit Matrizen.

Eine Matrix $(a_{ij}) \in M(n, m)$ kann mit einem Skalar $\lambda \in \mathbb{R}$ multipliziert werden:

$$\begin{aligned} \lambda(a_{ij}) &= \lambda \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{pmatrix} = (\lambda a_{ij}) \\ &= \begin{pmatrix} \lambda a_{11} & \lambda a_{12} & \dots & \lambda a_{1m} \\ \lambda a_{21} & \lambda a_{22} & \dots & \lambda a_{2m} \\ \vdots & \vdots & & \vdots \\ \lambda a_{n1} & \lambda a_{n2} & \dots & \lambda a_{nm} \end{pmatrix}. \end{aligned}$$

Ebenso einfach können zwei Matrizen $(a_{ij}), (b_{ij}) \in M(n, m)$ addiert werden:

$$\begin{aligned} (a_{ij}) + (b_{ij}) &= \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1m} \\ b_{21} & b_{22} & \dots & b_{2m} \\ \vdots & \vdots & & \vdots \\ b_{n1} & b_{n2} & \dots & b_{nm} \end{pmatrix} \\ &= \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \dots & a_{1m} + b_{1m} \\ a_{21} + b_{21} & a_{22} + b_{22} & \dots & a_{2m} + b_{2m} \\ \vdots & \vdots & & \vdots \\ a_{n1} + b_{n1} & a_{n2} + b_{n2} & \dots & a_{nm} + b_{nm} \end{pmatrix}. \end{aligned}$$

Die **Nullmatrix** (0) ist

$$(0) = \begin{pmatrix} 0 & \dots & 0 \\ \vdots & & \vdots \\ 0 & \dots & 0 \end{pmatrix} \in M(n, m)$$

und man setzt

$$(a_{ij}) - (a_{ij}) := (a_{ij}) + (-a_{ij}) = (0).$$

Mit diesen Bemerkungen kann schließlich leicht verifiziert werden:

Die Menge der $n \times m$ Matrizen $M(n, m)$ ist ein Vektorraum (hier über \mathbb{R}) im Sinne von Teil I, Definition 9.1.

Multiplikation von Matrizen.

Matrizen mit **zueinander passenden "Formaten"** können multipliziert werden:

Definition 1.2. MATRIXPRODUKT

Es seien $n, m, l \in \mathbb{N}$ und $A \in M(n, m)$, $B \in M(m, l)$, d.h. die Spaltenzahl der Matrix A sei gleich der Zeilenzahl der Matrix B .

Dann ist das *Matrixprodukt* AB per definitionem die Matrix

$$C = (c_{ij})_{i=1, \dots, n}^{j=1, \dots, l} \in M(n, l),$$

deren Eintrag in der i^{ten} Zeile ($i = 1, \dots, n$) und der j^{ten} Spalte ($j = 1, \dots, l$) gegeben ist durch

$$c_{ij} := \sum_{k=1}^m a_{ik} b_{kj}.$$

Merkregel.

- i) Das Produkt der i^{ten} Zeile der Matrix A mit der j^{ten} Spalte der Matrix B (gemäß $\sum_{k=1}^m a_{ik} b_{kj}$) ergibt den Eintrag in der i^{ten} Zeile und der j^{ten} Spalte der resultierenden Matrix C .
- ii) Schematisch sieht das wie folgt aus:

$$\begin{aligned}
& \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{im} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{pmatrix} \begin{pmatrix} b_{11} & \cdots & b_{1j} & \cdots & b_{1l} \\ b_{21} & \cdots & b_{2j} & \cdots & b_{2l} \\ \vdots & & \vdots & & \vdots \\ b_{m1} & \cdots & b_{mj} & \cdots & b_{ml} \end{pmatrix} \\
&= \begin{pmatrix} c_{11} & \cdots & c_{1j} & \cdots & c_{1l} \\ \vdots & & \vdots & & \vdots \\ c_{i1} & \cdots & c_{ij} & \cdots & c_{il} \\ \vdots & & \vdots & & \vdots \\ c_{n1} & \cdots & c_{nj} & \cdots & c_{nl} \end{pmatrix}.
\end{aligned}$$

Beispiele.

i) Es seien

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 1 & 2 \end{pmatrix} \in M(2, 3),$$

$$B = \begin{pmatrix} 0 & 1 & 2 & 0 \\ 1 & 0 & 0 & 1 \\ 3 & -1 & -1 & 0 \end{pmatrix} \in M(3, 4).$$

Dann ist

$$AB = \begin{pmatrix} 5 & 0 & 1 & 2 \\ 7 & 0 & 2 & 1 \end{pmatrix} \in M(2, 4).$$

ii) Es bezeichne I_m die quadratische $m \times m$ Einheitsmatrix,

$$I_m := \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 & 1 \end{pmatrix},$$

d.h. die Eintragungen a_{ii} , $i = 1, \dots, m$, auf der Hauptdiagonalen sind 1, alle anderen Eintragungen sind 0.

Für $A \in M(n, m)$ ist $AI_m \in M(n, m)$ und es gilt

$$AI_m = A.$$

Für $A \in M(n, m)$ ist $I_n A \in M(n, m)$ und es gilt

$$I_n A = A .$$

- iii) (a) Eine Matrix $A = (a_{ij}) \in M(n, m)$ kann mit einer $m \times 1$ Matrix, d.h. **mit einem Vektor $\underline{x} \in \mathbb{R}^m$ multipliziert werden.**

Das Ergebnis ist ein **Vektor im \mathbb{R}^n** :

$$A \underline{x} := \underbrace{\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{pmatrix}}_{\in M(n,m)} \underbrace{\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix}}_{\in \mathbb{R}^m} := \underbrace{\begin{pmatrix} \sum_{k=1}^m a_{1k} x_k \\ \sum_{k=1}^m a_{2k} x_k \\ \vdots \\ \sum_{k=1}^m a_{nk} x_k \end{pmatrix}}_{\in \mathbb{R}^n} .$$

- (b) Dementsprechend kann das lineare Gleichungssystem (vgl. wieder Gaußsches Eliminationsverfahren, Teil I, Anhang A)

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 \cdots + a_{1m}x_m &= b_1 \\ a_{21}x_1 + a_{22}x_2 \cdots + a_{2m}x_m &= b_2 \\ &\vdots \quad \vdots \quad \vdots \\ a_{n1}x_1 + a_{n2}x_2 \cdots + a_{nm}x_m &= b_n \end{aligned}$$

in der Schreibweise

$$A \underline{x} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = \underline{\mathbf{b}}$$

formuliert werden.

- (c) Eine sogenannte **lineare Abbildung** $L: V \rightarrow W$ von einem Vektorraum V in einen Vektorraum W ist durch die Eigenschaften

$$L(\underline{\mathbf{u}} + \underline{\mathbf{v}}) = L(\underline{\mathbf{u}}) + L(\underline{\mathbf{v}}) \quad \text{für alle } \underline{\mathbf{u}}, \underline{\mathbf{v}} \in V ,$$

$$L(\lambda \underline{\mathbf{v}}) = \lambda L(\underline{\mathbf{v}}) \quad \text{für alle } \underline{\mathbf{v}} \in V , \lambda \in \mathbb{R}$$

gekennzeichnet.

Lineare Abbildungen f von \mathbb{R} nach \mathbb{R} sind **durch eine reelle Zahl $a \in \mathbb{R}$ charakterisiert** (siehe Übungskapitel 3.3): Für alle $x \in \mathbb{R}$ ist

$$f(x) = ax .$$

Analog ist für $A \in M(n, m)$ die Abbildung $F: \mathbb{R}^m \rightarrow \mathbb{R}^n$,

$$\underline{\mathbf{x}} \mapsto F(\underline{\mathbf{x}}) := A\underline{\mathbf{x}}$$

eine lineare Abbildung.

Rechenregeln für die Matrizenmultiplikation.

Es seien A, B, C Matrizen und $\lambda \in \mathbb{R}$. Dann gilt:

- i)* $(A + B)C = AC + BC$ (Distributivgesetz);
- ii)* $A(B + C) = AB + AC$ (Distributivgesetz);
- iii)* $A(BC) = (AB)C$ (Assoziativgesetz);
- iv)* $A(\lambda B) = (\lambda A)B = \lambda(AB)$.

Übung. Wie müssen in den einzelnen Regeln jeweils die Zeilen- und Spaltenzahlen gewählt sein, damit die Aussagen sinnvoll definiert sind?

Vorsicht. Es **übertragen sich nicht alle bekannten Regeln** der Multiplikation etwa reeller Zahlen.

Beispielsweise sei

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} \in M(2, 2, \mathbb{R}), \quad B = \begin{pmatrix} 2 & 4 \\ -1 & -2 \end{pmatrix} \in M(2, 2, \mathbb{R}).$$

Dann gilt

$$AB = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix},$$

es ist aber

$$BA = \begin{pmatrix} 10 & 20 \\ -5 & -10 \end{pmatrix}.$$

Mit anderen Worten:

- i) Aus $AB = (0)$ kann im Allgemeinen nicht gefolgert werden, dass A oder B eine Nullmatrix ist.
- ii) Im Falle **quadratischer Matrizen** ($n = m$), bei denen sowohl das Produkt AB als auch das Produkt BA definiert ist, kann **kein Kommutativgesetz** gelten.

Transposition von Matrizen.

Eine weitere wichtige Operation mit Matrizen ist die **Transposition**. Dabei werden die **Zeilen und Spalten einer Matrix vertauscht**, d.h.: Ist

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{pmatrix} \in M(n, m),$$

so heißt

$$A^T := \begin{pmatrix} a_{11} & a_{21} & \dots & a_{n1} \\ a_{12} & a_{22} & \dots & a_{n2} \\ \vdots & \vdots & & \vdots \\ a_{1m} & a_{2m} & \dots & a_{nm} \end{pmatrix} \in M(m, n),$$

die zu A transponierte Matrix.

Beispiele.

- i) Es sei $A \in M(3, 2)$,

$$A = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}.$$

Dann ist

$$A^T = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \in M(2, 3, \mathbb{R}).$$

ii) Im Spezialfall eines Spaltenvektors $\underline{\mathbf{x}} \in M(n, 1)$,

$$\underline{\mathbf{x}} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n ,$$

produziert die Transposition einen **Zeilenvektor**

$$\underline{\mathbf{x}}^T = (x_1 \ x_2 \ \dots \ x_n) = \underline{\mathbf{x}} \in M(1, n) .$$

iii) Das **Skalarprodukt** (vgl. Teil I, Kapitel 9.2) zweier Vektoren $\underline{\mathbf{x}}, \underline{\mathbf{y}} \in \mathbb{R}^n$ kann somit auch als

$$\langle \underline{\mathbf{x}}, \underline{\mathbf{y}} \rangle = \underline{\mathbf{y}}^T \underline{\mathbf{x}} = (y_1 \ y_2 \ \dots \ y_n) \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \sum_{k=1}^n x_k y_k$$

geschrieben werden.

Rechenregeln.

i) Man erkennt sofort für $A, B \in M(n, m)$ und $\lambda \in \mathbb{R}$:

(a) $(A + B)^T = A^T + B^T$;

(b) $(\lambda A)^T = \lambda A^T$;

(c) $(A^T)^T = A$.

ii) Für $A \in M(n, m)$ und $B \in M(m, l)$ ist das Matrizenprodukt $AB \in M(n, l)$ definiert.

Wegen $B^T \in M(l, m)$ und $A^T \in M(m, n)$ ist aber ebenso die Bildung $B^T A^T \in M(l, n)$ erlaubt und es gilt (siehe Übungskapitel 1.4)

$$(AB)^T = B^T A^T .$$

Betrachtet man Matrizen mit **komplexen Einträgen**, $a_{ij} \in \mathbb{C}$, $i = 1, \dots, n$, $j = 1, \dots, m$, so geht man zur **adjungierten Matrix** über:

$$A^* := \bar{A}^T,$$

wobei \bar{A} die Matrix mit den konjugiert komplexen Einträgen bezeichnet – dementsprechend heißt \bar{A} die **konjugierte Matrix**.

Ersetzt man auf der rechten Seite der Regel i), (b) die komplexe Zahl λ durch die konjugiert komplexe Zahl $\bar{\lambda}$, so ergeben sich die obigen Rechenregeln ansonsten unverändert.

Beispiel. Motiviert ist der Übergang zur Adjungierten beispielsweise durch das so genannte **Hermiteische Skalarprodukt** für komplexe Vektoren

$$\underline{\mathbf{z}}, \underline{\mathbf{w}} \in \mathbb{C}^n = \underbrace{\mathbb{C} \times \mathbb{C} \times \dots \times \mathbb{C}}_{n\text{-mal}}.$$

Mit der Definition

$$\langle \underline{\mathbf{z}}, \underline{\mathbf{w}} \rangle := \underline{\mathbf{w}}^* \underline{\mathbf{z}} = \sum_{k=1}^n z_k \bar{w}_k$$

für alle

$$\underline{\mathbf{z}} = \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{pmatrix} \in \mathbb{C}^n, \quad \underline{\mathbf{w}} = \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix} \in \mathbb{C}^n$$

ist insbesondere

$$\|\underline{\mathbf{z}}\|^2 = \underline{\mathbf{z}}^* \underline{\mathbf{z}} = \langle \underline{\mathbf{z}}, \underline{\mathbf{z}} \rangle = \sum_{k=1}^n |z_k|^2.$$

Man beachte im komplexen Fall ($\underline{\mathbf{z}}, \underline{\mathbf{w}} \in \mathbb{C}^n$, $\lambda \in \mathbb{C}$):

$$\langle \underline{\mathbf{z}}, \underline{\mathbf{w}} \rangle = \overline{\langle \underline{\mathbf{w}}, \underline{\mathbf{z}} \rangle}, \quad \langle \underline{\mathbf{z}}, \lambda \underline{\mathbf{w}} \rangle = \bar{\lambda} \langle \underline{\mathbf{z}}, \underline{\mathbf{w}} \rangle.$$

1.2 Zur Lösbarkeit linearer Gleichungssysteme (homogenes und inhomogenes System; Superposition von Lösungen; Struktur der Lösungsmenge; Kern einer Matrix; Rang einer Matrix; erweiterte Matrix; Dimensionsformel)

Das Gaußsche Eliminationsverfahren zur expliziten Lösung (falls existent) linearer Gleichungssysteme ist bereits mehrfach angesprochen.

Nun geht es um die systematische Untersuchung der Frage nach der **Existenz und Eindeutigkeit** von Lösungen sowie um die **Struktur der Lösungsmenge**.

Im Folgenden werden wieder beispielhaft lineare Gleichungssysteme mit reellen Koeffizienten betrachtet.

Für $n, m \in \mathbb{N}$, $n \leq m$ und $1 \leq i \leq n$, $1 \leq j \leq m$ seien a_{ij} und $b_i \in \mathbb{R}$ fixiert. Zu diesen Daten ist

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 \cdots + a_{1m}x_m &= b_1 \\ a_{21}x_1 + a_{22}x_2 \cdots + a_{2m}x_m &= b_2 \\ &\vdots \quad \vdots \quad \vdots \\ a_{n1}x_1 + a_{n2}x_2 \cdots + a_{nm}x_m &= b_n \end{aligned}$$

ein lineares System aus n Gleichungen in m Unbekannten x_1, \dots, x_m .

Ist $A = (a_{ij}) \in M(n, m)$ die **Koeffizientenmatrix** und $\underline{\mathbf{b}} \in \mathbb{R}^n$ der Vektor $\begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$, so lautet die äquivalente Matrixschreibweise

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

Homogenes versus inhomogenes System.

i) Beim **homogenen System** verschwindet die rechte Seite:

$$A\underline{\mathbf{x}} = \underline{\mathbf{0}}.$$

Beobachtung.

(a) Das homogene System ist **immer lösbar**, nämlich mit der **trivialen Lösung** $\underline{\mathbf{x}} = \underline{\mathbf{0}}$:

$$A\underline{\mathbf{0}} = \underline{\mathbf{0}}.$$

Es bleibt die Frage nach der Existenz bzw. nach der Menge **nicht-trivialer Lösungen**.

(b) **Aufgrund der Linearität** des Systems gilt das **Superpositionsprinzip**, d.h. die **Summe von zwei Lösungen** und das **Vielfache einer Lösung** sind **ebenfalls Lösungen**.

Denn: Gilt für $\underline{\mathbf{x}} \in \mathbb{R}^m$ und für $\underline{\mathbf{y}} \in \mathbb{R}^m$

$$A\underline{\mathbf{x}} = \underline{\mathbf{0}} \quad \text{und} \quad A\underline{\mathbf{y}} = \underline{\mathbf{0}},$$

so folgt für $\underline{\mathbf{z}} = \lambda_1\underline{\mathbf{x}} + \lambda_2\underline{\mathbf{y}}$, $\lambda_1, \lambda_2 \in \mathbb{R}$,

$$A\underline{\mathbf{z}} = A(\lambda_1\underline{\mathbf{x}} + \lambda_2\underline{\mathbf{y}}) = \lambda_1 A\underline{\mathbf{x}} + \lambda_2 A\underline{\mathbf{y}} = \lambda_1\underline{\mathbf{0}} + \lambda_2\underline{\mathbf{0}} = \underline{\mathbf{0}}$$

und $\underline{\mathbf{z}}$ ist gleichfalls Lösung des homogenen Systems.

ii) Beim **inhomogenen System** ist eine Lösung von

$$A\underline{\mathbf{x}} = \underline{\mathbf{b}} \neq \underline{\mathbf{0}}$$

gesucht.

Beobachtung.

(a) Der **Nullvektor ist keine Lösung des inhomogenen Systems**.

Hier stellt sich die Frage, **ob bzw. für welche** $\underline{\mathbf{b}} \in \mathbb{R}^n$ eine Lösung **existiert und wie die Lösungsmenge aussieht**.

- (b) Beim inhomogenen System ist **die Summe von zwei Lösungen keine Lösung** (analog das Vielfache einer Lösung):

Aus $\underline{\mathbf{x}} \in \mathbb{R}^m$ und $\underline{\mathbf{y}} \in \mathbb{R}^m$ mit

$$A\underline{\mathbf{x}} = \underline{\mathbf{b}}, \quad \text{und} \quad A\underline{\mathbf{y}} = \underline{\mathbf{b}},$$

folgt nämlich für $\underline{\mathbf{z}} = \underline{\mathbf{x}} + \underline{\mathbf{y}}$

$$A\underline{\mathbf{z}} = A(\underline{\mathbf{x}} + \underline{\mathbf{y}}) = A\underline{\mathbf{x}} + A\underline{\mathbf{y}} = \underline{\mathbf{b}} + \underline{\mathbf{b}} = 2\underline{\mathbf{b}} \neq \underline{\mathbf{b}}.$$

- (c) Es gilt jedoch:

Ist $\underline{\mathbf{x}}$ eine **Lösung des homogenen Systems** $A\underline{\mathbf{x}} = \underline{\mathbf{0}}$ und ist $\underline{\mathbf{y}}$ eine **Lösung des inhomogenen Systems** $A\underline{\mathbf{y}} = \underline{\mathbf{b}}$, so ist $\underline{\mathbf{z}} = \underline{\mathbf{x}} + \underline{\mathbf{y}}$ eine **Lösung des inhomogenen Systems**:

$$A\underline{\mathbf{z}} = A(\underline{\mathbf{x}} + \underline{\mathbf{y}}) = A\underline{\mathbf{x}} + A\underline{\mathbf{y}} = \underline{\mathbf{0}} + \underline{\mathbf{b}} = \underline{\mathbf{b}}.$$

Die obigen Beobachtungen werden zusammengefasst in

Satz 1.1. STRUKTUR DER LÖSUNGSMENGE

- i) Die Menge der Lösungen des homogenen Systems $A\underline{\mathbf{x}} = \underline{\mathbf{0}}$ ist ein **Unterraum des \mathbb{R}^m** , genannt der **Kern der Matrix A** .

Notation: **kern A** .

Entweder ist $\text{kern } A = \{\underline{\mathbf{0}}\}$ oder es existiert eine natürliche Zahl $1 \leq k \leq m$ mit $\dim(\text{kern } A) = k$.

In zweiten Fall existiert eine Basis $(\underline{\mathbf{x}}^{(1)}, \underline{\mathbf{x}}^{(2)}, \dots, \underline{\mathbf{x}}^{(k)})$ von kern A und die **Menge aller Lösungen des homogenen Systems** ist gegeben durch

$$\underline{\mathbf{x}}_h = \sum_{j=1}^k \lambda_j \underline{\mathbf{x}}^{(j)}, \quad \lambda_i \in \mathbb{R}, \quad i = 1, \dots, k.$$

Bezeichnung: $\underline{\mathbf{x}}_h$ heißt die **allgemeine Lösung des homogenen Systems**.

ii) Es sei $\underline{\mathbf{x}}_s$ (*irgend-*) eine Lösung des inhomogenen Gleichungssystems $A\underline{\mathbf{x}} = \underline{\mathbf{b}}$.

Dann heißt $\underline{\mathbf{x}}_s$ eine *spezielle Lösung* und

$$\underline{\mathbf{x}} = \underline{\mathbf{x}}_s + \underline{\mathbf{x}}_h$$

ist die *allgemeine Lösung des inhomogenen Systems*.

Bemerkungen.

i) „Allgemeine Lösung“ bedeutet in beiden Fällen:

Für beliebige λ_i ist eine Lösung gegeben, und jede Lösung ist von dieser Form.

ii) Kennt man also *alle Lösungen des homogenen Systems* und nur *eine Lösung des inhomogenen System*, so kennt man *alle Lösungen des inhomogenen Systems* (siehe auch Übungskapitel 1.4).

Der Rang einer Matrix.

Satz 1.1 beantwortet zwar die Frage nach der Struktur der Lösungsmenge eines linearen Gleichungssystems, die *Existenz* nicht-trivialer Lösungen im homogenen Fall bzw. spezieller Lösungen im inhomogenen Fall wird im Satz jedoch nicht angesprochen.

Idee. Man schreibe die Matrix $A \in M(n, m)$ formal als *Zeile aus Spaltenvektoren* (vgl. Übungskapitel 1.4):

$$A = (\underline{\mathbf{a}}^{(1)} \quad \underline{\mathbf{a}}^{(2)} \quad \dots \quad \underline{\mathbf{a}}^{(m)}) , \quad \underline{\mathbf{a}}^{(j)} \in \mathbb{R}^n , \quad j = 1, \dots, m .$$

Hier ist für jedes fixierte $j = 1, \dots, m$

$$\underline{\mathbf{a}}^{(j)} = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{nj} \end{pmatrix} \in \mathbb{R}^n .$$

Weiterhin sei $(\underline{\mathbf{e}}^{(1)}, \underline{\mathbf{e}}^{(2)}, \dots, \underline{\mathbf{e}}^{(m)})$ wie üblich die kanonische Basis des \mathbb{R}^m .

Dann gilt für alle $j = 1, 2, \dots, m$ (nachrechnen!)

$$A\underline{\mathbf{e}}^{(j)} = \underline{\mathbf{a}}^{(j)} .$$

Identifiziert man, wie oben bereits angesprochen, die Matrix A mit der linearen Abbildung

$$F(\underline{\mathbf{x}}) = A\underline{\mathbf{x}} ,$$

so bedeutet das:

Die Bilder $F(\underline{\mathbf{e}}^{(j)})$ der kanonischen Basisvektoren sind genau die Spaltenvektoren der Matrix A .

Nun kann jeder Vektor $\underline{\mathbf{x}} \in \mathbb{R}^m$ geschrieben werden als

$$\underline{\mathbf{x}} = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} = \sum_{j=1}^m x_j \underline{\mathbf{e}}^{(j)} , \quad x_j \in \mathbb{R} , \quad j = 1, \dots, m ,$$

d.h.

$$A\underline{\mathbf{x}} = \sum_{j=1}^m x_j A\underline{\mathbf{e}}^{(j)} = \sum_{j=1}^m x_j \underline{\mathbf{a}}^{(j)} .$$

Damit das System

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

lösbar sein kann, muss es also Koeffizienten $x_i, i = 1, 2, \dots, m$, geben (die Unbekannten), sodass

$$\underline{\mathbf{b}} = \sum_{j=1}^m x_j \underline{\mathbf{a}}^{(j)} .$$

Mit anderen Worten: $\underline{\mathbf{b}}$ muss eine Linearkombination der Spaltenvektoren $\underline{\mathbf{a}}^{(i)}$ der Matrix A sein.

Zusammenfassung. Das lineare Gleichungssystem $A\underline{\mathbf{x}} = \underline{\mathbf{b}}$ ist genau dann lösbar, wenn

$$\underline{\mathbf{b}} \in \text{Spann}(\underline{\mathbf{a}}^{(1)}, \dots, \underline{\mathbf{a}}^{(m)}),$$

was äquivalent ist zu

$$\text{Spann}(\underline{\mathbf{a}}^{(1)}, \dots, \underline{\mathbf{a}}^{(m)}) = \text{Spann}(\underline{\mathbf{a}}^{(1)}, \dots, \underline{\mathbf{a}}^{(m)}, \underline{\mathbf{b}}).$$

Definition 1.3. RANG EINER MATRIX

Es seien $A \in M(n, m)$ und $\underline{\mathbf{b}} \in \mathbb{R}^n$.

- i) Der *Spaltenrang* oder der *Rang* der Matrix A ist die *maximale Anzahl linear unabhängiger Spaltenvektoren* von A .

Notation: $\text{rg } A$.

- ii) Die $n \times (m + 1)$ Matrix

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} & b_1 \\ a_{21} & a_{22} & \dots & a_{2m} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} & b_n \end{pmatrix}$$

heißt *erweiterte Matrix*.

Schreibweise:

$$(A|\underline{\mathbf{b}}) := \left(\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1m} & b_1 \\ a_{21} & a_{22} & \dots & a_{2m} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} & b_n \end{array} \right).$$

Mit Definition 1.3 liest sich obige Zusammenfassung als

Satz 1.2. KRITERIUM ZUR LÖSBARKEIT

Das lineare Gleichungssystem $A\underline{\mathbf{x}} = \underline{\mathbf{b}}$ ist *genau dann lösbar, wenn gilt*

$$\operatorname{rg} A = \operatorname{rg} (A|\underline{\mathbf{b}}) .$$

Beispiel. Es sei $n = m = 2$ und

$$A = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} .$$

Dann ist $\operatorname{rg} A = 1$, ebenso ist

$$\operatorname{rg} \left(\begin{array}{cc|c} 1 & 2 & 3 \\ 1 & 2 & 3 \end{array} \right) = 1 ,$$

das System

$$\begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} \underline{\mathbf{x}} = \begin{pmatrix} 3 \\ 3 \end{pmatrix}$$

ist lösbar (wie lauten die Lösungen?).

Dahingegen ist

$$\operatorname{rg} \left(\begin{array}{cc|c} 1 & 2 & 1 \\ 1 & 2 & 2 \end{array} \right) = 2 ,$$

das System

$$\begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} \underline{\mathbf{x}} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

ist nicht lösbar (ausprobieren!).

Bemerkungen.

- i) Analog zum Spaltenrang kann der **Zeilenrang** einer Matrix A definiert werden. Wegen (siehe Übungskapitel 1.4)

$$\text{Zeilenrang } A = \text{Spaltenrang } A$$

- (a) spricht man einfach vom **Rang einer Matrix** (siehe Definition 1.3);
 (b) folgt insbesondere

$$\operatorname{rg} A \leq \min\{n, m\} .$$

- ii) Der Rang einer Matrix kann auch nach einer Umformung auf Zeilenstufenform (vgl. Gaußsches Eliminationsverfahren, siehe Übungskapitel 1.4) abgelesen werden.

Die eingangs gestellten Fragen beantwortet abschließend

Satz 1.3. DIMENSIONSFORMEL

Es sei $A \in M(n, m)$. Dann sind die folgenden Aussagen richtig:

- i) Es gilt die *Dimensionsformel*

$$\dim(\text{kern } A) + \text{rg } A = m .$$

- ii) Ist das homogene Gleichungssystem $A\underline{\mathbf{x}} = \underline{\mathbf{0}}$ *unterbestimmt*, d.h. ist $n < m$ (weniger Gleichungen als Unbekannte), so hat es *stets nicht-triviale Lösungen*.

- iii) Ist $n = m$, so ist das lineare Gleichungssystem $A\underline{\mathbf{x}} = \underline{\mathbf{b}}$ *für jede rechte Seite $\underline{\mathbf{b}}$ genau dann eindeutig lösbar, wenn $\text{rg } A = n$.*

1.3 Überbestimmte lineare Gleichungssysteme – lineare Ausgleichsprobleme (Residuum; Methode der kleinsten Quadrate; Ausgleich nach Tschebyscheff; Normalgleichung)

In diesem Paragraphen wird eine typische Anwendung vorgestellt, die auf ein *überbestimmtes* ($n > m$, mehr Gleichungen als Unbekannte) lineares Gleichungssystem führt.

Ein Beispiel zur Motivation.

Eine quadratische Gesetzmäßigkeit der Form

$$f(\underline{\mathbf{x}}) = a_1 + a_2x + a_3x^2$$

soll experimentell gefunden werden, d.h. mithilfe von k experimentellen Messungen zu gegebenen Daten x_k und mit gemessenen Werten $y_k = f(x_k)$ sollen die unbekannt Parameter $a_1, a_2, a_3 \in \mathbb{R}$ bestimmt werden.

Aufgrund von Messfehlern ist es sicherlich nicht angeraten, die gesuchten Parameter a_1, a_2 und a_3 nur mit Hilfe von drei Messungen bestimmen zu wollen. Man wird eher eine Vielzahl von Messungen heranziehen.

Betrachtet man als einfaches Beispiel vier Messungen mit den Datenpaaren (x_k, y_k) , $k = 1, \dots, 4$,

$$\begin{array}{c|cccc} x_k & 0 & 1 & 2 & 3 \\ \hline y_k & 0 & 1 & 1 & 4/3 \end{array},$$

so kann die Aufgabe, die Parameter a_i , $i = 1, 2, 3$, zu bestimmen, als lineares Gleichungssystem **in den a_i als gesuchte Größen** formuliert werden:

$$\begin{pmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ 1 & x_3 & x_3^2 \\ 1 & x_4 & x_4^2 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} =: A\mathbf{a} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} =: \mathbf{y}.$$

Man erhält in diesem Beispiel vier Gleichungen in den drei Unbekannten a_1, a_2, a_3 .

Die Aufgabe ist also in der Regel nicht exakt lösbar, was wegen evtl. Messfehler auch nicht erwartet werden kann.

Um dennoch ein quadratisches Polynom als Gesetzmäßigkeit **möglichst gut an die Daten anzupassen**, wird das **Residuum \mathbf{r}** (in gewissem Sinne die „Abweichung von der Exaktheit“) betrachtet:

$$\mathbf{r} := A\mathbf{a} - \mathbf{y}.$$

Dieses Residuum soll möglichst klein gemacht werden, wobei **„möglichst klein“ auf unterschiedliche Arten gemessen werden kann**:

i) Versucht man, die Euklidische Norm

$$\left(\sum_{i=1}^4 r_i^2 \right)^{\frac{1}{2}} = \|\underline{\mathbf{r}}\|_2$$

zu minimieren, so spricht man von der [Methode der kleinsten Quadrate](#).

ii) Die Minimierung von

$$\max_{1 \leq i \leq 4} |r_i| = \|\underline{\mathbf{r}}\|_\infty$$

heißt der [Ausgleich nach Tschebyscheff](#).

iii) Allgemein lautet die Aufgabe,

$$\|\underline{\mathbf{r}}\| = \|A\underline{\mathbf{a}} - \underline{\mathbf{y}}\|$$

bzgl. einer gegebenen Norm zu minimieren, ein [lineares Ausgleichsproblem](#).

Die allgemeine Problemstellung.

In Anlehnung an das Beispiel sei $A \in M(n, m, \mathbb{R})$, $n > m \in \mathbb{N}$, $\underline{\mathbf{y}} \in \mathbb{R}^n$ und für alle $\underline{\mathbf{a}} \in \mathbb{R}^m$ sei

$$\underline{\mathbf{r}}(\underline{\mathbf{a}}) = A\underline{\mathbf{a}} - \underline{\mathbf{y}}.$$

Weiter betrachte man [im Folgenden den Fall](#) $\|\cdot\| = \|\cdot\|_2$, d.h. die Methode der kleinsten Quadrate.

Für alle $\underline{\mathbf{b}} \in \mathbb{R}^m$ ist mit der Notation $\underline{\mathbf{c}} := \underline{\mathbf{b}} - \underline{\mathbf{a}}$

$$\begin{aligned} \underline{\mathbf{r}}(\underline{\mathbf{b}}) &= \underline{\mathbf{r}}(\underline{\mathbf{a}} + \underline{\mathbf{c}}) \\ &= A(\underline{\mathbf{a}} + \underline{\mathbf{c}}) - \underline{\mathbf{y}} \\ &= \underline{\mathbf{r}}(\underline{\mathbf{a}}) + A\underline{\mathbf{c}}. \end{aligned}$$

Es ergibt sich

$$\begin{aligned}
 \|\underline{\mathbf{r}}(\underline{\mathbf{b}})\|^2 &= \underline{\mathbf{r}}(\underline{\mathbf{b}})^T \underline{\mathbf{r}}(\underline{\mathbf{b}}) \\
 &= (\underline{\mathbf{r}}(\underline{\mathbf{a}}) + A\underline{\mathbf{c}})^T (\underline{\mathbf{r}}(\underline{\mathbf{a}}) + A\underline{\mathbf{c}}) \\
 &= \|\underline{\mathbf{r}}(\underline{\mathbf{a}})\|^2 + \underline{\mathbf{r}}(\underline{\mathbf{a}})^T A\underline{\mathbf{c}} + (A\underline{\mathbf{c}})^T \underline{\mathbf{r}}(\underline{\mathbf{a}}) + \|A\underline{\mathbf{c}}\|^2 \\
 &= \|\underline{\mathbf{r}}(\underline{\mathbf{a}})\|^2 + \langle \underline{\mathbf{r}}(\underline{\mathbf{a}}), A\underline{\mathbf{c}} \rangle + \langle A\underline{\mathbf{c}}, \underline{\mathbf{r}}(\underline{\mathbf{a}}) \rangle + \|A\underline{\mathbf{c}}\|^2 \\
 &= \|\underline{\mathbf{r}}(\underline{\mathbf{a}})\|^2 + 2\langle A\underline{\mathbf{c}}, \underline{\mathbf{r}}(\underline{\mathbf{a}}) \rangle + \|A\underline{\mathbf{c}}\|^2 \\
 &= \|\underline{\mathbf{r}}(\underline{\mathbf{a}})\|^2 + 2\underline{\mathbf{c}}^T (A^T \underline{\mathbf{r}}(\underline{\mathbf{a}})) + \|A\underline{\mathbf{c}}\|^2.
 \end{aligned} \tag{1}$$

Ist nun **einerseits** $\text{rg } A = m$, was nach der Dimensionsformel aus Satz 1.3 (siehe Übungskapitel 1.4, vgl. auch die Überlegungen zum Spaltenrang einer Matrix)

$$\|A\underline{\mathbf{c}}\|^2 > 0 \quad \text{für } \underline{\mathbf{c}} \neq \underline{\mathbf{0}}$$

impliziert, und gilt

$$A^T \underline{\mathbf{r}}(\underline{\mathbf{a}}) = \underline{\mathbf{0}},$$

so folgt für alle $\underline{\mathbf{b}} \in \mathbb{R}^m$, $\underline{\mathbf{b}} \neq \underline{\mathbf{a}}$,

$$\|\underline{\mathbf{r}}(\underline{\mathbf{b}})\|^2 > \|\underline{\mathbf{r}}(\underline{\mathbf{a}})\|^2.$$

Mit anderen Worten: Ist

$$A^T \underline{\mathbf{r}}(\underline{\mathbf{a}}) = \underline{\mathbf{0}}, \tag{2}$$

so minimiert $\underline{\mathbf{a}}$ das Residuum nach der Methode der kleinsten Quadrate.

Umgekehrt folgt aus (1) auch, dass ein Minimierer der Gleichung (2) genügen muss:

Wäre für einen Minimierer $\underline{\mathbf{a}}$ nämlich $A^T \underline{\mathbf{r}}(\underline{\mathbf{a}}) \neq \underline{\mathbf{0}}$, so existierte ein Vektor $\tilde{\underline{\mathbf{c}}} \in \mathbb{R}^m$ mit $\tilde{\underline{\mathbf{c}}}^T A^T \underline{\mathbf{r}}(\underline{\mathbf{a}}) = -1$ und demnach für beliebiges $\varepsilon > 0$

$$(\varepsilon \tilde{\underline{\mathbf{c}}})^T A^T \underline{\mathbf{r}}(\underline{\mathbf{a}}) = -\varepsilon.$$

Aus (1) folgte dann

$$\|\underline{\mathbf{r}}(\underline{\mathbf{a}} + \varepsilon \tilde{\underline{\mathbf{c}}})\|^2 = \|\underline{\mathbf{r}}(\underline{\mathbf{a}})\|^2 - 2\varepsilon + \varepsilon^2 \|A\tilde{\underline{\mathbf{c}}}\|^2$$

und damit für hinreichend kleine $\varepsilon > 0$

$$\|\underline{\mathbf{r}}(\underline{\mathbf{a}} + \varepsilon \underline{\tilde{\mathbf{c}}})\| < \|\underline{\mathbf{r}}(\underline{\mathbf{a}})\| ,$$

mit anderen Worten ein Widerspruch zur Minimalität von $\underline{\mathbf{a}}$.

Schließlich wird die Gleichung (2) noch umformuliert,

$$\begin{aligned} A^T \underline{\mathbf{r}}(\underline{\mathbf{a}}) = \underline{\mathbf{0}} &\Leftrightarrow A^T (A \underline{\mathbf{a}} - \underline{\mathbf{y}}) = \underline{\mathbf{0}} \\ &\Leftrightarrow A^T A \underline{\mathbf{a}} = A^T \underline{\mathbf{y}} , \end{aligned}$$

und man definiert

Definition 1.4. NORMALGLEICHUNG

Es sei $A \in M(n, m, \mathbb{R})$, $n > m$, $\text{rg } A = m$ und $\underline{\mathbf{y}} \in \mathbb{R}^n$.

Dann heißt das lineare Gleichungssystem

$$A^T A \underline{\mathbf{a}} = A^T \underline{\mathbf{y}}$$

die *Normalgleichung* des linearen Ausgleichsproblems.

Aus $\text{rg } A = m$ folgt $\text{rg } A^T A = m$ (vgl. Übungskapitel 1.4) und daraus wiederum die eindeutige Lösbarkeit der Normalgleichung:

Satz 1.4. METHODE DER KLEINSTEN QUADRATE

Unter den Voraussetzungen aus Definition 1.4 hat das lineare Ausgleichsproblem nach der Methode der kleinsten Quadrate eine *eindeutige Lösung*, die sich aus der Normalgleichung berechnen lässt.

Bemerkung. Bei numerischen Berechnungen ist die Normalgleichung oft mit Vorsicht zu betrachten, da sich Rundungsfehler sehr stark auf das Ergebnis auswirken können. Dies wird später nochmals aufgegriffen.

Zurück zum Beispiel.

In diesem speziellen Fall ist

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \end{pmatrix}, \quad A^T = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 1 & 4 & 9 \end{pmatrix},$$

und man berechnet

$$A^T A = \begin{pmatrix} 4 & 6 & 14 \\ 6 & 14 & 36 \\ 14 & 36 & 98 \end{pmatrix}, \quad A^T \underline{\mathbf{y}} = \begin{pmatrix} 10/3 \\ 7 \\ 17 \end{pmatrix}.$$

Die Lösung der Normalgleichung

$$\begin{pmatrix} 4 & 6 & 14 \\ 6 & 14 & 36 \\ 14 & 36 & 98 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} 10/3 \\ 7 \\ 17 \end{pmatrix}$$

lautet

$$a_1 = \frac{1}{15}, \quad a_2 = \frac{9}{10}, \quad a_3 = -\frac{1}{6},$$

als Lösung des Ausgleichsproblems findet man (vgl. Abbildung 1.1)

$$f(x) = \frac{1}{15} + \frac{9}{10}x - \frac{1}{6}x^2.$$

Passt man mit der Methode der kleinsten Quadrate eine Gerade an die Daten an, so ergibt sich im Beispiel

$$A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{pmatrix}, \quad A^T = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \end{pmatrix},$$

und man berechnet

$$A^T A = \begin{pmatrix} 4 & 6 \\ 6 & 14 \end{pmatrix}, \quad A^T \underline{\mathbf{y}} = \begin{pmatrix} 10/3 \\ 7 \end{pmatrix}.$$

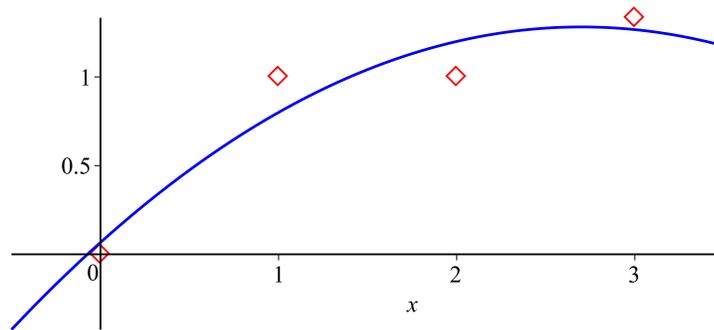


Abbildung 1.1: Das obige Ausgleichsproblem nach der Methode der kleinsten Quadrate.

Die Lösung der Normalgleichung

$$\begin{pmatrix} 4 & 6 \\ 6 & 14 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 10/3 \\ 7 \end{pmatrix}$$

lautet nun

$$a_1 = \frac{7}{30}, \quad a_2 = \frac{2}{5},$$

als Ausgleichsgerade findet man (vgl. Abbildung 1.2)

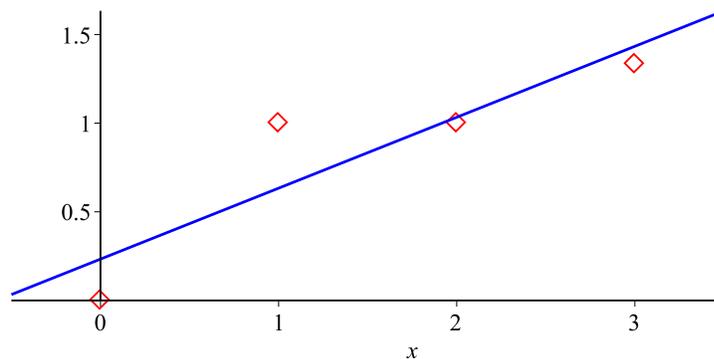


Abbildung 1.2: Die Ausgleichsgerade nach der Methode der kleinsten Quadrate.

$$f(x) = \frac{7}{30} + \frac{2}{5}x.$$

1.4 Übungsaufgaben zu Kapitel 1

Aufgabe 1.

- i) Es seien $A \in M(n_1, n_2)$, $B \in M(n_3, n_4)$ und $C \in M(n_5, n_6)$. Unter welchen Bedingungen an die $n_i \in \mathbb{N}$, $i = 1, \dots, 6$, ist das Matrizenprodukt

$$B^T C A^T$$

definiert? Welche Zeilen- und Spaltenzahl hat diese Matrix?

- ii) (a) Finden Sie eine Matrix $A \in M(?, ?)$ mit

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

- (b) Gibt es eine Matrix $A \in M(2, 3)$ mit

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

Aufgabe 2. Zeigen Sie die obigen Rechenregeln für die Matrixmultiplikation und geben Sie dabei die richtigen Zeilen- und Spaltenzahlen an.

Aufgabe 3. Zeigen Sie die obigen Rechenregeln für die Transposition von Matrizen.

Aufgabe 4.

- i) Eine Matrix $B \in M(m, l)$, $m, l \in \mathbb{N}$, kann als Tupel von Spaltenvek-

toren geschrieben werden:

$$B = \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1l} \\ b_{21} & b_{22} & \dots & b_{2l} \\ \vdots & \vdots & & \vdots \\ b_{m1} & b_{m2} & \dots & b_{ml} \end{pmatrix} = (\underline{\mathbf{b}}^{(1)} \quad \underline{\mathbf{b}}^{(2)} \quad \dots \quad \underline{\mathbf{b}}^{(l)}),$$

wobei für alle $i = 1, \dots, l$ gesetzt ist:

$$\underline{\mathbf{b}}^{(i)} = \begin{pmatrix} b_{1i} \\ b_{2i} \\ \vdots \\ b_{mi} \end{pmatrix} \in \mathbb{R}^m.$$

Zeigen Sie für $A \in M(n, m)$:

$$AB = (A\underline{\mathbf{b}}^{(1)} \quad A\underline{\mathbf{b}}^{(2)} \quad \dots \quad A\underline{\mathbf{b}}^{(l)}) \in M(n, l).$$

ii) Eine Matrix $B \in M(n, m)$ kann auch in der Form $(\tilde{\underline{\mathbf{b}}}^{(i)} \in \mathbb{R}^m, i = 1, \dots, n)$

$$B = \begin{pmatrix} (\tilde{\underline{\mathbf{b}}}^{(1)})^T \\ (\tilde{\underline{\mathbf{b}}}^{(2)})^T \\ \vdots \\ (\tilde{\underline{\mathbf{b}}}^{(n)})^T \end{pmatrix}$$

mit Zeilenvektoren

$$(\tilde{\underline{\mathbf{b}}}^{(i)})^T = (\tilde{b}_{i1} \quad \tilde{b}_{i2} \quad \dots \quad \tilde{b}_{im}),$$

$i = 1, \dots, n$, geschrieben werden.

Bestimmen Sie, ohne die Eigenschaft ‘‘Zeilenrang = Spaltenrang’’ zu zitieren, die maximale Anzahl linear unabhängiger Spaltenvektoren und die maximale Anzahl linear unabhängiger Zeilenvektoren für die Matrizen

$$A = \begin{pmatrix} 1 & 3 & 4 & 6 \\ 2 & 2 & 5 & 5 \\ 3 & 1 & 6 & 4 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 2 \\ 1 & 3 \\ 0 & 1 \\ 1 & -1 \end{pmatrix}.$$

Aufgabe 5. Es seien $n, m \in \mathbb{N}$, $n \leq m$, $A \in M(n, m)$ und $\underline{\mathbf{b}} \in \mathbb{R}^n$. Weiter sei (für ein $k \leq m$) $(\underline{\mathbf{x}}^{(1)}, \underline{\mathbf{x}}^{(2)}, \dots, \underline{\mathbf{x}}^{(k)})$ eine Basis von kern A .

Mit $\underline{\mathbf{y}}_s^{(1)}$ und $\underline{\mathbf{y}}_s^{(2)}$ seien zwei spezielle Lösungen des inhomogenen Systems $A\underline{\mathbf{x}} = \underline{\mathbf{b}}$ bezeichnet (falls existent) und es seien

$$L_1 := \left\{ \underline{\mathbf{x}} \in \mathbb{R}^m : \underline{\mathbf{x}} = \underline{\mathbf{y}}_s^{(1)} + \sum_{j=1}^k \lambda_j \underline{\mathbf{x}}^{(j)}, \lambda_i \in \mathbb{R}, i = 1, 2, \dots, k \right\},$$

$$L_2 := \left\{ \underline{\mathbf{x}} \in \mathbb{R}^m : \underline{\mathbf{x}} = \underline{\mathbf{y}}_s^{(2)} + \sum_{j=1}^k \mu_j \underline{\mathbf{x}}^{(j)}, \mu_i \in \mathbb{R}, i = 1, 2, \dots, k \right\}.$$

Zeigen Sie: $L_1 = L_2$.

Aufgabe 6.* Man betrachte die Matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} \in M(2, 3)$$

und zeige in diesem Beispiel, dass der Spaltenrang gleich dem Zeilenrang ist.

Aufgabe 7. Für $n, m \in \mathbb{N}$, betrachte man Vektoren $\underline{\mathbf{a}}^{(j)}$, $1 \leq j \leq m$ und $\underline{\mathbf{b}} \in \mathbb{R}^m$. Zeigen Sie:

$$\underline{\mathbf{b}} \in \text{Spann}(\underline{\mathbf{a}}^{(1)}, \dots, \underline{\mathbf{a}}^{(m)})$$

$$\Leftrightarrow \text{Spann}(\underline{\mathbf{a}}^{(1)}, \dots, \underline{\mathbf{a}}^{(m)}) = \text{Spann}(\underline{\mathbf{a}}^{(1)}, \dots, \underline{\mathbf{a}}^{(m)}, \underline{\mathbf{b}}).$$

Aufgabe 8. Es sei $a \in \mathbb{R}$ fixiert und

$$A = \begin{pmatrix} 1 & -1 & 0 \\ a & 0 & 1 \\ 1 & 1 & 2 \end{pmatrix} \in M(3, 3).$$

- i) Bestimmen Sie $\text{rg } A$.
- ii) Wie lautet die allgemeine Lösung des homogenen Gleichungssystems $A\underline{\mathbf{x}} = \underline{\mathbf{0}}$?
- iii) (a) Gibt es ein $\underline{\mathbf{b}} \in \mathbb{R}^3$, sodass das lineare Gleichungssystem $A\underline{\mathbf{x}} = \underline{\mathbf{b}}$ keine Lösung hat?
- (b) Gibt es ein $\underline{\mathbf{b}} \in \mathbb{R}^3$, sodass das lineare Gleichungssystem $A\underline{\mathbf{x}} = \underline{\mathbf{b}}$ eine eindeutige Lösung hat?
-

Aufgabe 9.

- i) Geben Sie eine Matrix $A \in M(6, 4)$ an mit $\dim(\text{kern } A) = 3$.
- ii) Man betrachte die Matrizen

$$A = \begin{pmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \\ 1 & 1 & 1 \end{pmatrix}.$$

Sind die linearen Gleichungssysteme $A\underline{\mathbf{x}} = \underline{\mathbf{b}}$, $B\underline{\mathbf{x}} = \underline{\mathbf{b}}$, $C\underline{\mathbf{x}} = \underline{\mathbf{b}}$ ($\underline{\mathbf{x}} \in \mathbb{R}^k$, $\underline{\mathbf{b}} \in \mathbb{R}^l$) für jede rechte Seite $\underline{\mathbf{b}}$ lösbar? Wie sind dabei k und l zu wählen, damit die Ausdrücke definiert sind?

Aufgabe 10. Es sei $a \in \mathbb{R}$ fixiert und

$$A = \begin{pmatrix} 1 & 2 \\ 3 & a \end{pmatrix} \in M(2, 2).$$

Für welche $\underline{\mathbf{b}} \in \mathbb{R}^2$ ist das lineare Gleichungssystem

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

lösbar? Ist die Lösung (falls existent) eindeutig?

Aufgabe 11.

- i) Bestimmen Sie den Rang der folgenden Matrizen mithilfe des Gaußschen Eliminationsverfahrens. ($\lambda \in \mathbb{R}$ fixiert)

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 \\ 3 & 4 & 5 & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & 1 & 3 & 2 \\ 1 & 1 & 4 & 2 \\ 1 & 1 & 5 & 2 \\ 1 & 0 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 & 3 & 0 \\ 3 & 2 & 1 & 0 \\ 1 & 0 & \lambda & \lambda \end{pmatrix}.$$

- ii) Was ist jeweils die Dimension des Kerns?

Aufgabe 12.* Es sei $n \geq m$, $A \in M(n, m)$ und $\text{rg } A = m$. Zeigen Sie:

- i) $\|A\underline{\mathbf{x}}\| > 0 \Leftrightarrow \underline{\mathbf{x}} \neq \underline{\mathbf{0}}$.
- ii) $\text{rg}(A^T A) = m$.

Aufgabe 13. Betrachten Sie die Daten

$$\begin{array}{c|cccc} x_i & -1 & 0 & 1 & 2 \\ \hline y_i & 0 & 1 & 2 & 2 \end{array}.$$

- i) Es sei $f(x) = a_1 + a_2x$. Bestimmen Sie a_1, a_2 nach der Methode der kleinsten Quadrate zu diesen Daten.
- ii) Es sei $f(x) = a_1 + a_2x + a_3x^2$. Bestimmen Sie a_1, a_2, a_3 nach der Methode der kleinsten Quadrate zu diesen Daten.

Aufgabe 14. Vollziehen Sie die Rechnungen dieses Kapitels mithilfe Ihres Computeralgebrasystems nach.

Lösungshinweise zu den Übungsaufgaben.

Aufgabe 6. Der Spaltenrang muss in diesem Beispiel kleiner oder gleich zwei sein (im \mathbb{R}^2 gibt es höchstens zwei linear unabhängige Vektoren).

Ist der Spaltenrang gleich 0, so sind alle Eintragungen Null und die Aussage ist trivial.

Nun sei der Spaltenrang gleich 1. Dann kann die Matrix geschrieben werden als (o.E. seien der zweite und der dritte Spaltenvektor linear abhängig vom ersten)

$$A = \begin{pmatrix} a_{11} & \alpha a_{11} & \beta a_{11} \\ a_{21} & \alpha a_{21} & \beta a_{21} \end{pmatrix}, \quad \alpha, \beta \in \mathbb{R}.$$

Damit sind die Zeilenvektoren (die Transponierten)

$$a_{11} \begin{pmatrix} 1 \\ \alpha \\ \beta \end{pmatrix}, \quad a_{21} \begin{pmatrix} 1 \\ \alpha \\ \beta \end{pmatrix},$$

linear abhängig und der Zeilenrang ist ebenfalls gleich 1.

Ist der Spaltenrang gleich 2, so folgt die Aussage analog.

Aufgabe 12. *i)* und *ii)*. Nach der Dimensionsformel aus Satz 1.3 ist kern $A = \{\underline{\mathbf{0}}\}$, d.h. für alle $\underline{\mathbf{x}} \in \mathbb{R}^m$, $\underline{\mathbf{x}} \neq \underline{\mathbf{0}}$, ist $A\underline{\mathbf{x}} \neq \underline{\mathbf{0}}$, d.h. wiederum

$$\|A\underline{\mathbf{x}}\|^2 = \underline{\mathbf{x}}^T A^T A \underline{\mathbf{x}} > 0.$$

Wäre kern $A^T A \neq \{\underline{\mathbf{0}}\}$, so gäbe es ein $\underline{\mathbf{x}} \in \mathbb{R}^m$, $\underline{\mathbf{x}} \neq \underline{\mathbf{0}}$, mit $A^T A \underline{\mathbf{x}} = \underline{\mathbf{0}}$ und folglich $\|A\underline{\mathbf{x}}\|^2 = 0$, was zu einem Widerspruch führt.

Die Dimensionsformel zeigt schließlich $\text{rg } A^T A = m$. □

Kapitel 2

Quadratische Matrizen – Inverse und Determinante

In diesem Abschnitt sei $A \in M(n, n)$ stets eine quadratische $n \times n$ Matrix. Für nicht-quadratische Matrizen ergeben die folgenden Betrachtungen keinen Sinn.

2.1 Invertierbare Matrizen (reguläre und singuläre Matrizen; Dreiecksmatrizen; LR-Zerlegung; Cholesky-Zerlegung)

Zunächst wird der Frage nachgegangen, ob und wann zu einer gegebenen Matrix $A \in M(n, n)$ eine **inverse Matrix** A^{-1} existiert, wobei A^{-1} per definitionem bedeutet:

$$A^{-1}A = AA^{-1} = I_n .$$

Man startet mit einer

Beobachtung. Existiert eine solche Matrix, so ist das lineare Gleichungssystem $A\underline{\mathbf{x}} = \underline{\mathbf{b}}$ für jede rechte Seite eindeutig lösbar.

Es gilt nämlich in diesem Fall

$$A\underline{\mathbf{x}} = \underline{\mathbf{b}} \quad \Leftrightarrow \quad \underline{\mathbf{x}} = A^{-1}\underline{\mathbf{b}} .$$

Insbesondere folgt nach Satz 1.3 aus der Existenz einer inversen Matrix $\text{rg } A = n$.

Die Umkehrung ist auch richtig:

Satz 2.1. EXISTENZ EINER INVERSEN

Eine Matrix $A \in M(n, n)$ besitzt *genau dann* eine inverse Matrix (*diese ist, falls existent, immer eindeutig bestimmt*), wenn A den maximal möglichen *Rang* $\text{rg } A = n$ hat.

Dann heißt die Matrix *regulär* oder *invertierbar*.

Ist die Matrix nicht regulär, so heißt sie *singulär*.

Beispiel. Es sei

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M(2, 2).$$

Man verifiziert sofort im Fall $ad - bc \neq 0$

$$A A^{-1} = I_2 \quad \text{mit} \quad A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Die Bedingung $ad - bc \neq 0$ ist wiederum äquivalent dazu, dass die Vektoren

$$\begin{pmatrix} a \\ c \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} b \\ d \end{pmatrix}$$

linear unabhängig sind (vgl. Übungskapitel 2.3), was genau der Aussage des Satzes entspricht.

Eigenschaften. Es seien $A, B \in M(n, n)$ reguläre Matrizen. Dann gilt

- i) $(A^{-1})^{-1} = A$;
- ii) $(A^T)^{-1} = (A^{-1})^T$;
- iii) $(AB)^{-1} = B^{-1}A^{-1}$.

Explizite Berechnung.

Die Inverse einer regulären Matrix kann mithilfe des Gaußschen Eliminationsverfahrens explizit berechnet werden, wie hier anhand eines Beispiels illustriert werden soll.

Man betrachte dazu die reguläre Matrix

$$\begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \\ 1 & 0 & 1 \end{pmatrix}$$

und ergänze diese um die Einheitsmatrix:

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 3 & 2 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \end{array} \right).$$

Durch elementare Zeilenumformungen soll nun auf die Gestalt

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & c_{11} & c_{12} & c_{13} \\ 0 & 1 & 0 & c_{21} & c_{22} & c_{23} \\ 0 & 0 & 1 & c_{31} & c_{32} & c_{33} \end{array} \right) \quad (1)$$

transformiert werden.

Ist C die Matrix (c_{ij}) aus (1), so ist $C = A^{-1}$.

Im Beispiel zieht man etwa von der zweiten Zeile das Dreifache der ersten ab und dividiert das Ergebnis durch (-4) . Von der dritten Zeile zieht man die erste ab und dividiert durch (-2) :

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & 2 & 3/4 & -1/4 & 0 \\ 0 & 1 & 1 & 1/2 & 0 & -1/2 \end{array} \right).$$

Jetzt kann beispielsweise das Zweifache der zweiten Zeile von der ersten abgezogen werden und von der dritten Zeile wird die zweite abgezogen. Mit

(−1) multipliziert ergibt sich

$$\left(\begin{array}{ccc|ccc} 1 & 0 & -1 & -1/2 & 1/2 & 0 \\ 0 & 1 & 2 & 3/4 & -1/4 & 0 \\ 0 & 0 & 1 & 1/4 & -1/4 & 1/2 \end{array} \right).$$

Schließlich wird die dritte Zeile zur ersten addiert und das Zweifache der dritten Zeile von der zweiten abgezogen. Man erhält die gesuchte Form

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -1/4 & 1/4 & 1/2 \\ 0 & 1 & 0 & 1/4 & 1/4 & -1 \\ 0 & 0 & 1 & 1/4 & -1/4 & 1/2 \end{array} \right).$$

Man verifiziert sofort mit einer Probe

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

LR-Zerlegung invertierbarer Matrizen.

Für reguläre Matrizen sind so genannte **Dreieckszerlegungen** wegen ihrer numerischen Effizienz von besonderem Interesse, wenn beispielsweise ein Gleichungssystem für verschiedene rechte Seiten zu lösen ist.

Unter einer **LR-Zerlegung** einer regulären Matrix A versteht man eine Zerlegung der Form

$$A = LR = \begin{pmatrix} 1 & 0 & \dots & 0 & 0 \\ l_{21} & 1 & \dots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ l_{n1} & l_{n2} & \dots & l_{n(n-1)} & 1 \end{pmatrix} \begin{pmatrix} r_{11} & r_{12} & \dots & r_{1n} \\ 0 & r_{22} & \dots & r_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & r_{nn} \end{pmatrix}.$$

Wie eine solche Zerlegung mithilfe des Gaußschen Eliminationsverfahrens gewonnen werden kann und wie damit lineare Gleichungssysteme gelöst

werden können, soll hier ebenfalls exemplarisch anhand eines Beispiels illustriert werden.

Es sei dazu

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix} .$$

Wie üblich wird A mit Hilfe von elementaren Zeilenumformungen transformiert – **zunächst ohne Zeilenvertauschungen** (vgl. Bemerkung i) unten).

Man merkt sich hier jedoch die entsprechenden Vorfaktoren und es wird wie folgt umgeformt.

- i) Von der zweiten und der dritten Zeile werden jeweils Vielfache der ersten Zeile abgezogen, um in den ersten Einträgen eine Null zu produzieren.
- ii) Statt der führenden Nullen fügt man die Faktoren, mit denen die erste Zeile multipliziert wurde, in das Schema ein.
Dabei wird das Vorzeichen beachtet: Addiert man ein Vielfaches der ersten Zeile, so wird dies mit einem negativen Faktor berücksichtigt.
- iii) In der dritten Zeile produziert man zwei führende Nullen, indem ein entsprechendes Vielfaches der zweiten Zeile nach i) abgezogen wird, wobei die Einträge aus ii) nicht berücksichtigt werden.
- iv) Analog zu ii) wird der fehlende Faktor eingetragen.

Konkret bedeutet das im Beispiel:

Von der zweiten Zeile wird dreimal die erste Zeile abgezogen und die **3** wird statt der 0 ins Schema eingetragen.

Von der dritten Zeile wird die erste (einmal) abgezogen, die **1** ersetzt nun die 0:

$$\begin{pmatrix} 1 & 2 & 3 \\ \mathbf{3} & -4 & -8 \\ \mathbf{1} & -1 & -1 \end{pmatrix} .$$

Schließlich ist von der dritten Zeile ein viertel der zweiten Zeile abzuziehen, wobei die **rot markierten Einträge unverändert bleiben**.

Man merkt sich das viertel wiederum und gelangt zu

$$\begin{pmatrix} 1 & 2 & 3 \\ 3 & -4 & -8 \\ 1 & 1/4 & 1 \end{pmatrix}.$$

Die schwarzen Einträge des Ergebnisses bilden eine **rechte obere Dreiecksmatrix R** .

Die roten Einträge füllt man mit Einsen auf der Hauptdiagonalen auf und es entsteht eine **linke untere Dreiecksmatrix L** .

Es verbleibt zu verifizieren:

$$LR := \begin{pmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ 1 & 1/4 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 0 & -4 & -8 \\ 0 & 0 & 1 \end{pmatrix} = A.$$

Mithilfe der LR -Zerlegung kann ein Gleichungssystem sehr einfach durch **Vorwärtseinsetzen** und durch anschließendes **Rückwärtseinsetzen** gelöst werden:

Zunächst wird das Gleichungssystem

$$L\underline{\mathbf{y}} = \underline{\mathbf{b}}$$

gelöst (Vorwärtseinsetzen).

Ist etwa $\underline{\mathbf{b}} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$, so ergibt das im obigen Beispiel

$$\begin{aligned} y_1 &= 1, \\ 3y_1 + y_2 &= 1, & \text{also } y_2 &= -2, \\ y_1 + \frac{1}{4}y_2 + y_3 &= 1, & \text{also } y_3 &= \frac{1}{2}. \end{aligned}$$

Anschließend ist das System

$$R\underline{\mathbf{x}} = \underline{\mathbf{y}}$$

zu lösen (Rückwärtseinsetzen). Insgesamt hat man

$$A\underline{\mathbf{x}} = LR\underline{\mathbf{x}} = L(R\underline{\mathbf{x}}) = L\underline{\mathbf{y}} = \underline{\mathbf{b}}.$$

Im Beispiel berechnet man

$$\begin{aligned} x_1 + 2x_2 + 3x_3 &= 1 \\ -4x_2 - 8x_3 &= -2 \\ x_3 &= \frac{1}{2}, \end{aligned}$$

und in der Tat ist

$$\underline{\mathbf{x}} = \frac{1}{2} \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$$

die gesuchte Lösung des Systems $A\underline{\mathbf{x}} = \underline{\mathbf{b}}$ (Probe!).

Bemerkungen.

- i)* Nicht für jede reguläre Matrix existiert eine solche Dreieckszerlegung. Erst nach evtl. Zeilenvertauschungen ist dies immer möglich.

Anstelle einer LR -Zerlegung von

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

wird beispielsweise die LR -Zerlegung von

$$\tilde{A} = \begin{pmatrix} 3 & 1 & 1 \\ 2 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

wie oben konstruiert.

ii) Ist

$$D = \begin{pmatrix} r_{11} & 0 & \dots & 0 & 0 \\ 0 & r_{22} & \dots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \dots & 0 & r_{nn} \end{pmatrix},$$

$$\tilde{R} = \begin{pmatrix} 1 & \tilde{r}_{12} & \dots & \tilde{r}_{1(n-1)} & \tilde{r}_{1n} \\ 0 & 1 & \dots & \tilde{r}_{2(n-1)} & \tilde{r}_{2n} \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 1 \end{pmatrix},$$

wobei $\tilde{r}_{ij} = r_{ij}/r_{ii}$, $r_{ii} \neq 0$, $i, j = 1, \dots, n$, $i < j$, so ist

$$D\tilde{R} = R.$$

Wird zusätzlich vorausgesetzt, dass A **symmetrisch** ist, d.h. $A = A^T$, so folgt

$$A = A^T = (LDR)^T = \tilde{R}^T DL^T.$$

Nun ist aber \tilde{R}^T eine normierte untere Dreiecksmatrix, DL^T ist eine obere Dreiecksmatrix und aus der Eindeutigkeit der LR -Zerlegung folgt

$$\tilde{R}^T = L \quad \text{und} \quad R = DL^T.$$

Zusammenfassend ist gezeigt, dass die LR -Zerlegung für symmetrische Matrizen von der besonderen Gestalt

$$A = LDL^T,$$

(D wie oben Diagonalmatrix) ist.

Die Zerlegung heißt in diesem Fall **Cholesky-Zerlegung**.

2.2 Die Determinante (orientiertes Volumen; Regel von Sarrus, Laplacescher Entwicklungssatz; alternierende n -Linearform; Multiplikationssatz; Cramersche Regel; Orientierung; orthogonale Matrizen)

Die formale Definition der **Determinante** einer quadratischen Matrix sieht auf den ersten Blick aus wie eine „rein technische“ Rechenvorschrift ohne anschauliche Interpretation.

Tatsächlich täuscht dieser Eindruck gewaltig:

Die Determinante ist als orientierte Volumenfunktion eine fundamentale geometrische Größe, die in den unterschiedlichsten Bereichen von entscheidender Bedeutung ist.

Die Determinante als geometrisches Objekt.

Zur Motivation sei hier der einfachste Fall $n = 2$, d.h. der Fall einer 2×2 Matrix

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

betrachtet.

Ist A regulär, so sind – wie bereits gesehen – die Spaltenvektoren

$$\begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix}, \quad \begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix}$$

linear unabhängig, was genau dann der Fall ist, wenn der **Flächeninhalt des von ihnen aufgespannten Parallelogramms nicht verschwindet**.

Dieser Flächeninhalt soll nun mithilfe einer orientierten Volumenfunktion quantifiziert werden, wobei zunächst zu klären ist, welche „natürlichen“ Eigenschaften ein **orientiertes Volumen** (im Fall $n = 2$ ein **orientierter Flächeninhalt**) haben sollte.

Dazu wird der orientierte Flächeninhalt als Abbildung

$$\tau : \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}$$

interpretiert:

Zwei Vektoren im \mathbb{R}^2 wird eine reelle Zahl, der orientierte Flächeninhalt des von den beiden Vektoren aufgespannten Parallelogramms, zugeordnet.

Das Vorzeichen dieser Zahl korrespondiert mit der **Orientierung** der beiden Vektoren.

Die vier fundamentalen geometrischen Forderungen an die Funktion τ lauten der “natürlichen” intuitiven Vorstellung folgend:

- i) Die Vertauschung der Reihenfolge von zwei Vektoren bewirkt lediglich einen Vorzeichenwechsel, d.h. man fordert (vgl. Abbildung 2.1)

$$\tau(\underline{\mathbf{a}}^{(1)}, \underline{\mathbf{a}}^{(2)}) = -\tau(\underline{\mathbf{a}}^{(2)}, \underline{\mathbf{a}}^{(1)}) .$$

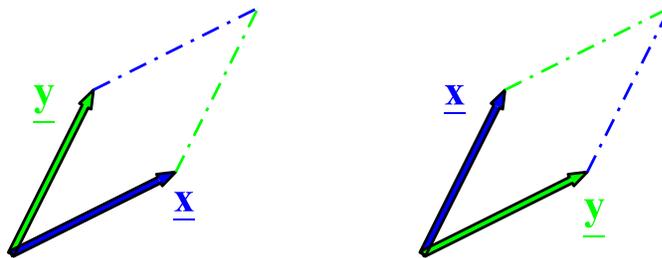


Abbildung 2.1: $\tau(\underline{\mathbf{x}}, \underline{\mathbf{y}}) > 0$ entspricht der auf der linken Seite angedeuteten Orientierung. Auf der rechten Seite ist $\tau(\underline{\mathbf{x}}, \underline{\mathbf{y}}) = -\tau(\underline{\mathbf{y}}, \underline{\mathbf{x}}) < 0$.

- ii) Zweitens ist zu fordern, dass der **Flächeninhalt proportional zu einer Verlängerung bzw. Verkürzung einer Seite** ist.

Mit anderen Worten soll für alle $\lambda \in \mathbb{R}$

$$\tau(\lambda \underline{\mathbf{a}}^{(1)}, \underline{\mathbf{a}}^{(2)}) = \lambda \tau(\underline{\mathbf{a}}^{(1)}, \underline{\mathbf{a}}^{(2)})$$

erfüllt sein. Es folgt sofort

$$\tau(\underline{\mathbf{a}}^{(1)}, \lambda \underline{\mathbf{a}}^{(2)}) = -\tau(\lambda \underline{\mathbf{a}}^{(2)}, \underline{\mathbf{a}}^{(1)}) = -\lambda \tau(\underline{\mathbf{a}}^{(2)}, \underline{\mathbf{a}}^{(1)}) = \lambda \tau(\underline{\mathbf{a}}^{(1)}, \underline{\mathbf{a}}^{(2)})$$

für alle $\lambda \in \mathbb{R}$.

iii) Um zu gegebenen Vektoren $\underline{\mathbf{a}}^{(1)}$, $\tilde{\underline{\mathbf{a}}}^{(1)}$, $\underline{\mathbf{a}}^{(2)}$ eine geometrische Rechenregel für

$$\tau(\underline{\mathbf{a}}^{(1)} + \tilde{\underline{\mathbf{a}}}^{(1)}, \underline{\mathbf{a}}^{(2)})$$

abzuleiten, betrachte man Abbildung 2.2.

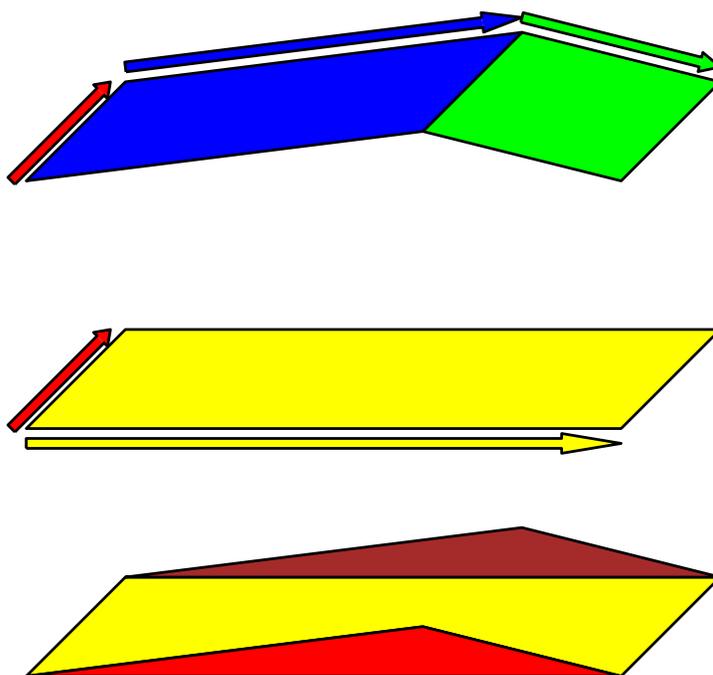


Abbildung 2.2: $\tau(\underline{\mathbf{a}}^{(1)} + \tilde{\underline{\mathbf{a}}}^{(1)}, \underline{\mathbf{a}}^{(2)}) = \tau(\underline{\mathbf{a}}^{(1)}, \underline{\mathbf{a}}^{(2)}) + \tau(\tilde{\underline{\mathbf{a}}}^{(1)}, \underline{\mathbf{a}}^{(2)})$.

Der Vektor $\underline{\mathbf{a}}^{(1)}$ ist dort blau symbolisiert, der Vektor $\tilde{\underline{\mathbf{a}}}^{(1)}$ grün und $\underline{\mathbf{a}}^{(2)}$ rot.

Im oberen Teil der Abbildung entspricht der Flächeninhalt des blau dargestellten Parallelogramms $\tau(\underline{\mathbf{a}}^{(1)}, \underline{\mathbf{a}}^{(2)})$, $\tau(\tilde{\underline{\mathbf{a}}}^{(1)}, \underline{\mathbf{a}}^{(2)})$ ist grün angedeutet.

Der Vektor $\underline{\mathbf{a}}^{(1)} + \tilde{\underline{\mathbf{a}}}^{(1)}$ ist ebenso wie $\tau(\underline{\mathbf{a}}^{(1)} + \tilde{\underline{\mathbf{a}}}^{(1)}, \underline{\mathbf{a}}^{(2)})$ im mittleren Teil der Abbildung gelb dargestellt.

Im unteren Teil der Abbildung ist anhand des braunen und des roten Dreiecks verdeutlicht, dass die blau und grün markierten Flächeninhalte oben genau dem gelben Flächeninhalt in der Mitte entsprechen.

Demnach muss als Drittes gefordert werden:

$$\tau(\underline{\mathbf{a}}^{(1)} + \tilde{\underline{\mathbf{a}}}^{(1)}, \underline{\mathbf{a}}^{(2)}) = \tau(\underline{\mathbf{a}}^{(1)}, \underline{\mathbf{a}}^{(2)}) + \tau(\tilde{\underline{\mathbf{a}}}^{(1)}, \underline{\mathbf{a}}^{(2)}) ,$$

was unmittelbar auch

$$\tau(\underline{\mathbf{a}}^{(1)}, \underline{\mathbf{a}}^{(2)} + \tilde{\underline{\mathbf{a}}}^{(2)}) = \tau(\underline{\mathbf{a}}^{(1)}, \underline{\mathbf{a}}^{(2)}) + \tau(\underline{\mathbf{a}}^{(1)}, \tilde{\underline{\mathbf{a}}}^{(2)}) .$$

impliziert.

iv) Ist schließlich $(\underline{\mathbf{e}}^{(1)}, \underline{\mathbf{e}}^{(2)})$ die kanonische Basis des \mathbb{R}^2 , so **normiert** man als letzte Forderung den Flächeninhalt des Quadrates mit der Seitenlänge 1:

$$\tau(\underline{\mathbf{e}}^{(1)}, \underline{\mathbf{e}}^{(2)}) = 1 .$$

Für beliebige Vektoren $\begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix}, \begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix} \in \mathbb{R}^2$ folgt aus *i)* bis *iv)*:

$$\begin{aligned} \tau\left(\begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix}, \begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix}\right) &= \tau(a_{11}\underline{\mathbf{e}}^{(1)} + a_{21}\underline{\mathbf{e}}^{(2)}, a_{12}\underline{\mathbf{e}}^{(1)} + a_{22}\underline{\mathbf{e}}^{(2)}) \\ &= a_{11}\tau(\underline{\mathbf{e}}^{(1)}, a_{12}\underline{\mathbf{e}}^{(1)} + a_{22}\underline{\mathbf{e}}^{(2)}) \\ &\quad + a_{21}\tau(\underline{\mathbf{e}}^{(2)}, a_{12}\underline{\mathbf{e}}^{(1)} + a_{22}\underline{\mathbf{e}}^{(2)}) \\ &= a_{11}a_{12}\tau(\underline{\mathbf{e}}^{(1)}, \underline{\mathbf{e}}^{(1)}) + a_{11}a_{22}\tau(\underline{\mathbf{e}}^{(1)}, \underline{\mathbf{e}}^{(2)}) \\ &\quad + a_{21}a_{12}\tau(\underline{\mathbf{e}}^{(2)}, \underline{\mathbf{e}}^{(1)}) + a_{21}a_{22}\tau(\underline{\mathbf{e}}^{(2)}, \underline{\mathbf{e}}^{(2)}) \\ &= a_{11}a_{22} - a_{12}a_{21} . \end{aligned}$$

In der letzten Gleichheit wurde ausgenutzt, dass für einen beliebigen Vektor $\underline{\mathbf{a}}$ aus der ersten Forderung $\tau(\underline{\mathbf{a}}, \underline{\mathbf{a}}) = -\tau(\underline{\mathbf{a}}, \underline{\mathbf{a}})$ und damit $\tau(\underline{\mathbf{a}}, \underline{\mathbf{a}}) = 0$ folgt.

Insbesondere stellt man anhand der obigen Rechnung fest:

Beobachtung. Die Funktion τ ist allein durch die Forderungen *i)* bis *iv)* eindeutig festgelegt.

Bemerkung. Die Rechnung zeigt auch:

- i)* Das positive Vorzeichen des Terms $a_{11}a_{22}$ rührt daher, dass die blauen auf die roten Indizes mit der Permutation $(1, 2)$ abgebildet werden, es werden also keine Vertauschungen vorgenommen, $\underline{\mathbf{e}}^{(1)}$ und $\underline{\mathbf{e}}^{(2)}$ sind in der Rechnung in der richtigen Reihenfolge aufgetreten.
- ii)* Das negative Vorzeichen des Terms $a_{12}a_{21}$ kommt daher, dass die Permutation $(2, 1)$ durch eine Vertauschung (**Transposition**) aus der $(1, 2)$ entsteht, $\underline{\mathbf{e}}^{(1)}$ und $\underline{\mathbf{e}}^{(2)}$ sind sozusagen in der falschen Reihenfolge (in negativer Orientierung) aufgetreten.

Definition der Determinante.

Zur Vereinfachung der Darstellung wird die Determinante nun abhängig von n über die entsprechenden Rechenvorschriften definiert. Die formale Definition wird in den anschließenden Bemerkungen zur Definition angedeutet.

Definition 2.1. DETERMINANTE

Es sei $A = (a_{ij}) \in M(n, n)$. Dann ist die Determinante

$$\det A =: \begin{vmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix}$$

wie folgt definiert:

i) Ist $n = 2$, so ist

$$\det A := a_{11}a_{22} - a_{12}a_{21} .$$

ii) Ist $n = 3$, so berechnet sich die Determinante aus der *Regel von Sarrus* (vgl. Tabelle 2.1)

$$\det A := a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\ - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31} .$$

iii) Zur *induktiven Definition* im allgemeinen Fall $n \geq 3$ werden zunächst die *Streichungsmatrizen* A_{ij} von A definiert:

Für $i, j = 1, \dots, n$ geht die $(n-1) \times (n-1)$ Matrix A_{ij} aus A durch die *Streichung der i^{ten} Zeile und der j^{ten} Spalte* hervor,

$$A = \begin{pmatrix} a_{11} & \dots & a_{1(j-1)} & a_{1j} & a_{1(j+1)} & \dots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{(i-1)1} & \dots & a_{(i-1)(j-1)} & a_{(i-1)j} & a_{(i-1)(j+1)} & \dots & a_{(i-1)n} \\ a_{i1} & \dots & a_{i(j-1)} & a_{ij} & a_{i(j+1)} & \dots & a_{in} \\ a_{(i+1)1} & \dots & a_{(i+1)(j-1)} & a_{(i+1)j} & a_{(i+1)(j+1)} & \dots & a_{(i+1)n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \dots & a_{n(j-1)} & a_{nj} & a_{n(j+1)} & \dots & a_{nn} \end{pmatrix} ,$$

$$A_{ij} := \begin{pmatrix} a_{11} & \dots & a_{1(j-1)} & a_{1(j+1)} & \dots & a_{1n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{(i-1)1} & \dots & a_{(i-1)(j-1)} & a_{(i-1)(j+1)} & \dots & a_{(i-1)n} \\ a_{(i+1)1} & \dots & a_{(i+1)(j-1)} & a_{(i+1)(j+1)} & \dots & a_{(i+1)n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{n1} & \dots & a_{n(j-1)} & a_{n(j+1)} & \dots & a_{nn} \end{pmatrix} .$$

Dann gilt der folgende *Laplacesche Entwicklungssatz*.

(a) *Entwicklung nach der i^{ten} Zeile* :

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

(b) *Entwicklung nach der j^{ten} Spalte:*

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

$$\begin{array}{ccc|ccc}
 a_{11} & a_{12} & a_{13} & a_{11} & a_{12} & \\
 & \searrow & \searrow \swarrow & \searrow \swarrow & & \swarrow \\
 a_{21} & a_{22} & a_{23} & a_{21} & a_{22} & a_{23} \\
 & \swarrow & \swarrow \searrow & \swarrow \searrow & & \searrow \\
 a_{31} & a_{32} & a_{33} & a_{31} & a_{32} & \\
 \end{array}$$

Tabelle 2.1: Zur Regel von Sarrus: Positive Vorzeichen hat man für die Produkte in Richtung der blauen Pfeile, negative Vorzeichen in Richtung der roten Pfeile.

Beispiele.

i) Es sei

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix} \in M(3,3) .$$

Dann ist

$$\begin{aligned}
 \det A &= 1 \cdot 1 \cdot 1 + 2 \cdot 1 \cdot 0 + 3 \cdot 1 \cdot 1 - 3 \cdot 1 \cdot 0 - 1 \cdot 1 \cdot 1 - 2 \cdot 1 \cdot 1 \\
 &= 1 .
 \end{aligned}$$

ii) Es sei

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 2 & 2 \\ 0 & 1 & 1 \end{pmatrix} \in M(3,3) .$$

Eine **Entwicklung nach der ersten Zeile** liefert

$$\begin{aligned}
 \det A &= 1 \begin{vmatrix} 2 & 2 \\ 1 & 1 \end{vmatrix} - 2 \begin{vmatrix} 2 & 2 \\ 0 & 1 \end{vmatrix} + 3 \begin{vmatrix} 2 & 2 \\ 0 & 1 \end{vmatrix} \\
 &= 0 - 4 + 6 = 2 .
 \end{aligned}$$

Eine **Entwicklung nach der zweiten Spalte** liefert

$$\begin{aligned} \det A &= -2 \begin{vmatrix} 2 & 2 \\ 0 & 1 \end{vmatrix} + 2 \begin{vmatrix} 1 & 3 \\ 0 & 1 \end{vmatrix} - 1 \begin{vmatrix} 1 & 3 \\ 2 & 2 \end{vmatrix} \\ &= -4 + 2 + 4 = 2. \end{aligned}$$

iii) Es sei

$$A = \begin{pmatrix} 0 & 1 & 2 & 3 \\ 0 & 2 & 2 & 2 \\ 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix} \in M(4, 4).$$

Bei der Entwicklung von $\det A$ wird man **die Nullen in der ersten Spalte** ausnutzen, d.h. nach der ersten Spalte entwickeln.

Dies ergibt (vgl. Beispiel *ii*)

$$\det A = 1 \begin{vmatrix} 1 & 2 & 3 \\ 2 & 2 & 2 \\ 0 & 1 & 1 \end{vmatrix} = 2.$$

Bemerkungen zur Definition.

i) In der Regel von Sarrus (*ii*), Definition 2.1) sind analog zum Fall $n = 2$ die Permutationen $(1, 2, 3)$, $(2, 3, 1)$, $(3, 1, 2)$ blau angedeutet, weil sie durch eine gerade Anzahl an paarweise Vertauschungen aus $(1, 2, 3)$ produziert werden können – das zugehörige Vorzeichen in der Regel von Sarrus ist positiv.

Das Vorzeichen wird nämlich durch das so genannte **Signum** $\text{sgn}(\sigma)$ der Permutation σ gegeben, welches per definitionem bei einer geraden Anzahl an benötigten Vertauschungen $+1$, bei einer ungeraden -1 ist.

Dementsprechend sind die in der Regel von Sarrus rot gekennzeichneten Permutationen $(1, 3, 2)$, $(2, 1, 3)$, $(3, 2, 1)$ mit einem negativen Vorzeichen versehen.

- ii)* Wie bereits angedeutet, sind die Regel von Sarrus und der Laplace'sche Entwicklungssatz Folgerungen aus der formalen Definition der Determinante, deren Grundidee in der vorangegangenen Bemerkung zusammengefasst ist.

Die übliche Definition der Determinante lautet:

Es sei $A = (a_{ij}) \in M(n, n)$. Dann heißt

$$\det A := \begin{vmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix} := \sum_{\sigma} \operatorname{sgn}(\sigma) a_{1\sigma_1} \dots a_{n\sigma_n}$$

die Determinante der Matrix A . Dabei ist die Summe so zu verstehen, dass über alle unterschiedlichen Permutationen von $(1, 2, \dots, n)$ zu summieren ist.

- iii)* Im Übungskapitel 2.3 wird aufgezeigt, wie die Determinante auch mithilfe des Gaußschen Eliminationsverfahrens berechnet werden kann.
- iv)* Anders formuliert können die definierenden Eigenschaften der Determinante als orientiertes Volumen wie folgt ausgedrückt werden:

- (a) $\det I_n = 1$ (vgl. *iv*);
- (b) die Determinante ist eine **n -Linearform**, d.h. in Abhängigkeit von einem Spaltenvektor (analog Zeilenvektor) bei fixierten $(n - 1)$ Spaltenvektoren (Zeilenvektoren) ist die Determinante eine lineare Abbildung $\mathbb{R}^n \rightarrow \mathbb{R}$ (vgl. *ii* und *iii*);
- (c) die Determinante ist **alternierend**, d.h. bei der Vertauschung zweier Spalten (Zeilen) ändert sich das Vorzeichen (vgl. *i*).

- v)* Der Begriff n -Linearform bedeutet **nicht, dass die Determinante eine lineare Abbildung ist:**

Im Allgemeinen gilt für $A, B \in M(n, n)$

$$\det(A + B) \neq \det A + \det B .$$

Im Gegensatz dazu bedeutet Linearität bzgl. einer Spalte beispielsweise $(\alpha, \beta \in \mathbb{R}, \underline{\mathbf{x}}, \underline{\mathbf{y}} \in \mathbb{R}^3)$

$$\begin{aligned} \det \begin{pmatrix} \alpha x_1 + \beta y_1 & 1 & 3 \\ \alpha x_2 + \beta y_2 & 2 & 3 \\ \alpha x_3 + \beta y_3 & 1 & 1 \end{pmatrix} \\ = \alpha \det \begin{pmatrix} x_1 & 1 & 3 \\ x_2 & 2 & 3 \\ x_3 & 1 & 1 \end{pmatrix} + \beta \det \begin{pmatrix} y_1 & 1 & 3 \\ y_2 & 2 & 3 \\ y_3 & 1 & 1 \end{pmatrix}. \end{aligned}$$

vi) Genau wie es von einer Volumenform in \mathbb{R}^n zu erwarten ist, gilt für $A \in M(n, n)$ und $\lambda \in \mathbb{R}$

$$\det(\lambda A) = \lambda^n \det A.$$

Weitere Eigenschaften der Determinante.

Zu den wichtigsten Eigenschaften der Determinante gehört

Satz 2.2. DETERMINANTE, RANG UND REGULARITÄT

Für eine quadratische Matrix $A \in M(n, n)$ sind die folgenden Aussagen äquivalent:

- i) Es ist $\det A \neq 0$.
- ii) Die Zeilenvektoren von A sind linear unabhängig.
- iii) Die Spaltenvektoren von A sind linear unabhängig.
- iv) Die Matrix A ist regulär.

Man kann auch leicht zeigen:

i) Für alle $A \in M(n, n)$ ist

$$\det A^T = \det A .$$

ii) Es gilt der **Determinantenmultiplikationssatz**: Sind $A, B \in M(n, n)$, so ist

$$\det (AB) = \det A \det B .$$

iii) Ist $A \in M(n, n)$ regulär, so ist

$$\det A^{-1} = \frac{1}{\det A} .$$

Schließlich sei noch eine geschlossene Formel zur Lösung linearer Gleichungssysteme angegeben. Wegen des hohen Rechenaufwandes ist diese für konkrete Rechnungen jedoch meist nicht zu empfehlen.

Satz 2.3. CRAMERSCHE REGEL

Es sei $A = (a_{ij}) = (\underline{\mathbf{a}}^{(1)} \ \underline{\mathbf{a}}^{(2)} \ \dots \ \underline{\mathbf{a}}^{(n)}) \in M(n, n)$ regulär und $\underline{\mathbf{b}} \in \mathbb{R}^n$.

Für alle $j = 1, \dots, n$ ist die j^{te} Komponente der eindeutigen Lösung des linearen Gleichungssystems $A\underline{\mathbf{x}} = \underline{\mathbf{b}}$ gegeben durch

$$x_j = \frac{\det (\underline{\mathbf{a}}^{(1)} \ \dots \ \underline{\mathbf{a}}^{(j-1)} \ \underline{\mathbf{b}} \ \underline{\mathbf{a}}^{(j+1)} \ \underline{\mathbf{a}}^{(n)})}{\det A} .$$

Determinante und Orientierung.

In Teil I, Kapitel 9.2.2, wurde das **Vektorprodukt** diskutiert und die Richtung des Vektors $\underline{\mathbf{x}} \times \underline{\mathbf{y}}$ ($\underline{\mathbf{x}}, \underline{\mathbf{y}} \in \mathbb{R}^3$) über die „**Rechte Hand Regel**“ festgelegt.

Mit der Determinante steht nun ein Hilfsmittel zur Verfügung, um dies mathematisch zu präzisieren und eine Definition zu geben, die in beliebiger Dimension $n \in \mathbb{N}$ greift.

Das Vorzeichen der Orientierung korrespondiert mit dem Vorzeichenwechsel der Determinante bei der Vertauschung von zwei Spaltenvektoren (Zeilenvektoren).

Definition 2.2. ORIENTIERUNG

Eine Basis $(\underline{\mathbf{v}}^{(1)}, \dots, \underline{\mathbf{v}}^{(n)})$ des \mathbb{R}^n heißt *positiv orientiert*, falls

$$\det(\underline{\mathbf{v}}^{(1)} \dots \underline{\mathbf{v}}^{(n)}) > 0 .$$

Andernfalls heißt die Basis *negativ orientiert*.

Beispiele. Im \mathbb{R}^3 mit der kanonischen Basis $(\underline{\mathbf{e}}^{(1)}, \underline{\mathbf{e}}^{(2)}, \underline{\mathbf{e}}^{(3)})$ und mit zwei linear unabhängigen Vektoren $\underline{\mathbf{x}}, \underline{\mathbf{y}} \in \mathbb{R}^3$ ist

- i) $(\underline{\mathbf{e}}^{(1)}, \underline{\mathbf{e}}^{(2)}, \underline{\mathbf{e}}^{(3)})$ positiv orientiert;
- ii) $(\underline{\mathbf{e}}^{(2)}, \underline{\mathbf{e}}^{(1)}, \underline{\mathbf{e}}^{(3)})$ negativ orientiert;
- iii) $(\underline{\mathbf{x}}, \underline{\mathbf{y}}, \underline{\mathbf{x}} \times \underline{\mathbf{y}})$ positiv orientiert.

Determinante und orthogonale Matrizen.

Eine spezielle Klasse quadratischer Matrizen sind so genannte **orthogonale Matrizen** $A \in M(n, n)$, die per definitionem die Identitäten

$$AA^T = A^T A = I_n .$$

erfüllen.

Für fixiertes $\varphi \in [0, 2\pi)$ betrachte man als typisches Beispiel die orthogonale Matrix

$$A = \begin{pmatrix} \cos(\varphi) & -\sin(\varphi) \\ \sin(\varphi) & \cos(\varphi) \end{pmatrix} \in M(2, 2) .$$

Die Abbildung $F: \mathbb{R}^2 \rightarrow \mathbb{R}^2$

$$\underline{\mathbf{x}} \mapsto A\underline{\mathbf{x}}$$

beschreibt in diesem Fall eine **Rotation** mit Winkel φ des Vektors $\underline{\mathbf{x}}$ um den Ursprung.

Nun schreibe man eine orthogonale Matrix $A \in M(n, n)$ in der mittlerweile vertrauten Weise als **Zeile von Spaltenvektoren**,

$$A = (\underline{\mathbf{a}}^{(1)} \ \underline{\mathbf{a}}^{(2)} \ \dots \ \underline{\mathbf{a}}^{(n)}) , \quad \underline{\mathbf{a}}^{(i)} \in \mathbb{R}^n , \quad i = 1, \dots, n ,$$

womit $A^T \in M(n, n)$ als **Spalte von Zeilenvektoren** geschrieben werden kann:

$$A^T = \begin{pmatrix} (\underline{\mathbf{a}}^{(1)})^T \\ (\underline{\mathbf{a}}^{(2)})^T \\ \vdots \\ (\underline{\mathbf{a}}^{(n)})^T \end{pmatrix} .$$

Aus der Orthogonalität folgt

$$(\delta_{ij}) = I_n = A^T A = \begin{pmatrix} (\underline{\mathbf{a}}^{(1)})^T \\ (\underline{\mathbf{a}}^{(2)})^T \\ \vdots \\ (\underline{\mathbf{a}}^{(n)})^T \end{pmatrix} (\underline{\mathbf{a}}^{(1)} \ \underline{\mathbf{a}}^{(2)} \ \dots \ \underline{\mathbf{a}}^{(n)}) = (\langle \underline{\mathbf{a}}^{(i)}, \underline{\mathbf{a}}^{(j)} \rangle) .$$

Mit anderen Worten: **Die Spaltenvektoren einer orthogonalen Matrix bilden eine Orthonormalbasis des \mathbb{R}^n .**

Aus der Definition folgt für orthogonale Matrizen

$$A^T = A^{-1} ,$$

und wegen

$$\det A^T = \det A$$

ergibt sich zusammen mit dem Determinantenmultiplikationssatz

$$\det A = \det A^T = \det A^{-1} = \frac{1}{\det A} ,$$

d.h.: Für jede orthogonale Matrix $A \in M(n, n)$ gilt

$$\det A = \pm 1 .$$

Aus $\det A = \pm 1$ folgt hingegen nicht, dass A orthogonal ist (vgl. Übungskapitel 2.3).

Bemerkung. Die Begriffsbildungen bei Matrizen mit komplexen Einträgen lauten:

Eine quadratische Matrix A mit komplexen Einträgen heißt

- i)* Hermitesch, falls $A = A^*$;
- ii)* unitär, falls $AA^* = A^*A = I_n$.

2.3 Übungsaufgaben zu Kapitel 2

Aufgabe 1.* Gegeben seien $a, b, c, d \in \mathbb{R}$. Wann sind die Vektoren

$$\begin{pmatrix} a \\ c \end{pmatrix}, \begin{pmatrix} b \\ d \end{pmatrix} \in \mathbb{R}^2$$

linear unabhängig?

Aufgabe 2. Zeigen Sie die obigen Eigenschaften *i)* bis *iii)* invertierbarer Matrizen.

Aufgabe 3. Existieren die Inversen der folgenden Matrizen? Falls ja, so berechnen Sie diese.

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 2 \\ 1 & 1 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} \cos(\varphi) & -\sin(\varphi) \\ \sin(\varphi) & \cos(\varphi) \end{pmatrix}, \quad \varphi \in \mathbb{R} \text{ fixiert.}$$

Aufgabe 4. Es sei

$$A = \begin{pmatrix} 1 & 1 & 2 \\ 1 & -1 & 1 \\ 2 & -1 & 0 \end{pmatrix} ?$$

- i)* Existiert die inverse Matrix? Falls ja, so berechnen Sie diese und machen Sie eine Probe.
- ii)* Berechnen Sie die *LR*-Zerlegung von A und lösen Sie mithilfe dieser Zerlegung das lineare Gleichungssystem $A\underline{\mathbf{x}} = \begin{pmatrix} 1 \\ 3 \\ 1 \end{pmatrix}$. Machen Sie eine Probe.
-

Aufgabe 5.

i) Zeigen Sie den Determinatenmultiplikationssatz im Spezialfall $A, B \in M(2, 2)$.

ii) Es seien

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \quad C = \begin{pmatrix} 0 & 2 \\ 1 & 2 \end{pmatrix} \in M(2, 2).$$

Berechnen Sie (falls existent) $\det(AC^{-1})$ sowie $\det(ABC)$.

iii) * Zeigen Sie für $A, B \in M(n, n)$:

$$A \text{ und } B \text{ sind regulär.} \quad \Leftrightarrow \quad AB \text{ ist regulär.}$$

Aufgabe 6. Zeigen Sie mithilfe vollständiger Induktion nach der Dimension n

$$\det \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ & \ddots & \vdots \\ 0 & & a_{nn} \end{pmatrix} = \prod_{i=1}^n a_{ii}.$$

Aufgabe 7.* Berechnen Sie

$$\det \begin{pmatrix} 1 & 2 & 1 \\ 2 & 1 & 2 \\ 1 & 0 & 2 \end{pmatrix}$$

i) mithilfe der Regel von Sarrus;

ii) mithilfe des Laplaceschen Entwicklungssatzes;

iii) durch elementare Umformungen.

Aufgabe 8. Es sei $a \in \mathbb{R}$ und

$$A = \begin{pmatrix} 1 & a & 1 \\ 0 & 2 & a \\ a & a & 1 \end{pmatrix} \in M(3, 3).$$

Berechnen Sie die Determinante von A . Ist A invertierbar?

Aufgabe 9.

i) Berechnen Sie (für $r > 0$, $\varphi \in \mathbb{R}$)

$$\det \begin{pmatrix} \cos(\varphi) & -r \sin(\varphi) & 0 \\ \sin(\varphi) & r \cos(\varphi) & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

ii) Berechnen Sie (für $r > 0$, $\varphi, \theta \in \mathbb{R}$)

$$\det \begin{pmatrix} \cos(\varphi) \cos(\theta) & -r \sin(\varphi) \cos(\theta) & -r \cos(\varphi) \sin(\theta) \\ \sin(\varphi) \cos(\theta) & r \cos(\varphi) \cos(\theta) & -r \sin(\varphi) \sin(\theta) \\ \sin(\theta) & 0 & r \cos(\theta) \end{pmatrix}.$$

Aufgabe 10.

i) Ergänzen Sie die Vektoren

$$\underline{\mathbf{x}} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad \underline{\mathbf{y}} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$$

zu einer negativ orientierten Basis des \mathbb{R}^3 .

ii) Finden Sie eine Matrix $A \in M(n, n)$ mit $\det A = 1$, die nicht orthogonal ist.

Aufgabe 11. Überzeugen Sie sich davon, dass Ihr Computeralgebrasystem auch inverse Matrizen und Determinanten berechnen kann.

Lösungshinweise zu den Übungsaufgaben.

Aufgabe 1. Um die lineare Unabhängigkeit zu zeigen, kann wegen

$$\begin{pmatrix} a \\ c \end{pmatrix} \neq \underline{\mathbf{0}}$$

o.E. $a \neq 0$ angenommen werden ($a = 0$ und $c \neq 0$ analog). In diesem Fall gilt

$$\begin{pmatrix} a \\ c \end{pmatrix}, \begin{pmatrix} b \\ d \end{pmatrix} \text{ linear abhängig} \Leftrightarrow \begin{pmatrix} a \\ c \end{pmatrix} = \lambda \begin{pmatrix} b \\ d \end{pmatrix}$$

für ein $\lambda \in \mathbb{R}$ und $b \neq 0$. Dann ist aber

$$\lambda = \frac{a}{b} \text{ und } c = \frac{a}{b}d, \quad \text{d.h. } ad - bc = 0.$$

Lineare Unabhängigkeit bedeutet also $ad - bc \neq 0$ (vgl. Determinante).

Aufgabe 5. *iii)* Nach dem Determinantenmultiplikationssatz und Satz 2.2 gilt

$$AB \text{ regulär} \Leftrightarrow \det A \cdot \det B \neq 0,$$

d.h. A und B sind regulär. □

Aufgabe 7.

i) Nach der Regel von Sarrus ist

$$\det \begin{pmatrix} 1 & 2 & 1 \\ 2 & 1 & 2 \\ 1 & 0 & 2 \end{pmatrix} = 2 + 4 + 0 - 1 - 0 - 8 = -3.$$

ii) Eine Entwicklung beispielsweise nach der ersten Spalte liefert

$$\det \begin{pmatrix} 1 & 2 & 1 \\ 2 & 1 & 2 \\ 1 & 0 & 2 \end{pmatrix} = 1 \cdot 2 - 2 \cdot 4 + 1 \cdot 3 = -3.$$

iii) Durch Addition von Vielfachen einer Zeile zu einer anderen transformiert man die Matrix auf die Form

$$\begin{pmatrix} 1 & 2 & 1 \\ 0 & -3 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

wobei die Determinante wegen

$$\det(\underline{\mathbf{a}}(\underline{\mathbf{b}} + \lambda\underline{\mathbf{a}})\underline{\mathbf{c}}) = \det(\underline{\mathbf{a}}\underline{\mathbf{b}}\underline{\mathbf{c}}) + \lambda \underbrace{\det(\underline{\mathbf{a}}\underline{\mathbf{a}}\underline{\mathbf{c}})}_{=0}$$

unverändert bleibt.

Eine Entwicklung nach der ersten Zeile ergibt wieder den Wert 3.

Statt einer Entwicklung nach der ersten Zeile kann auch Aufgabe 6 zitiert werden.

Vorsicht. Ersetzt man eine Zeile durch ein Vielfaches (im Gegensatz zur Addition eines Vielfachen zu einer weiteren Zeile), so ändert sich die Determinante um diesen Faktor.

Entsprechend ist auch bei der Vertauschung von Zeilen der Vorzeichenwechsel nach den Eigenschaften der Determinante zu beachten.

Kapitel 3

Lineare Abbildungen

Lineare Abbildungen sind eine natürliche Klasse von Abbildungen zwischen zwei Vektorräumen, denn sie vertragen sich per definitionem mit der Struktur linearer Räume.

Viele wichtige Operationen können als lineare Abbildungen interpretiert werden. Beispielsweise beschreibt die lineare Abbildung

$$\mathbb{R}^2 \ni \underline{\mathbf{x}} \mapsto \begin{pmatrix} \cos(\varphi) & -\sin(\varphi) \\ \sin(\varphi) & \cos(\varphi) \end{pmatrix} \underline{\mathbf{x}}$$

eine Rotation im \mathbb{R}^2 .

Aber auch beim Studium nicht-linearer Probleme spielen lineare Abbildungen eine gewichtige Rolle: Nicht-lineare Probleme sind in der Regel nur schwer oder gar nicht explizit lösbar. Deshalb approximiert man diese häufig zu einer gegebenen Fehlertoleranz mit (affin) linearen Funktionen (für „kleine“ x setzt man $e^x \approx 1 + x$, wie es die Potenzreihe nahelegt, etc.), sodass eine mehr oder weniger explizite Rechnung möglich ist.

Nach der Definition sowie der Klärung der grundlegenden Begriffe und Eigenschaften steht in diesem Kapitel der Zusammenhang zwischen linearen Abbildungen und Matrizen im Vordergrund.

3.1 Definition und erste Eigenschaften (Kern und Bild einer linearen Abbildung; Rang; Rangsatz; Injektivität, Surjektivität und Bijektivität einer linearen Abbildung)

Im einfachsten Fall ist f eine lineare Abbildung $f: \mathbb{R} \rightarrow \mathbb{R}$. Diese ist von der Form (vgl. Übungskapitel 3.3)

$$x \mapsto ax, \quad a \in \mathbb{R} \text{ fixiert.}$$

Im allgemeinen Fall ist die oben erwähnte Verträglichkeit mit der Struktur von Vektorräumen in Definition 3.1 festgehalten, wobei stets daran zu denken ist, dass sich die Additionen „+“ auf der linken und der rechten Seite unterscheiden, obwohl sie mit dem gleichen Symbol bezeichnet werden.

Auf der einen Seite handelt es sich um die Addition im Vektorraum V (z.B. im \mathbb{R}^m), auf der anderen Seite um die Addition im Vektorraum W (z.B. im \mathbb{R}^n oder auch in einem Vektorraum von Funktionen).

Definition 3.1. LINEARE ABBILDUNG

Es seien V und W zwei Vektorräume über einem Körper, hier der Einfachheit halber stets über \mathbb{R} .

Eine Abbildung

$$L : V \rightarrow W$$

*heißt lineare Abbildung oder **Homomorphismus**, falls für alle $\underline{\mathbf{v}}, \underline{\mathbf{w}} \in V$ und für alle $\lambda \in \mathbb{R}$ gilt:*

$$\begin{aligned} L(\underline{\mathbf{v}} + \underline{\mathbf{w}}) &= L(\underline{\mathbf{v}}) + L(\underline{\mathbf{w}}), \\ L(\lambda \underline{\mathbf{v}}) &= \lambda L(\underline{\mathbf{v}}). \end{aligned}$$

Ebenso wie in einem linearen Raum (in einem Vektorraum) stets die $\underline{\mathbf{0}}$ liegen muss, impliziert die Definition einer linearen Abbildung

$$L(\underline{\mathbf{0}}) = \underline{\mathbf{0}}.$$

Die Funktion $f: \mathbb{R} \rightarrow \mathbb{R}$,

$$x \mapsto ax + b, \quad a, b \in \mathbb{R}, b \neq 0,$$

ist demnach nicht linear, man nennt sie **affin linear**.

Zur Vereinfachung der Darstellung werden **im Folgenden lineare Abbildungen vom \mathbb{R}^m in den \mathbb{R}^n betrachtet**, obwohl in gleicher Weise auch ganz andere Beispiele diskutiert werden können.

Ist $A \in M(n, m)$, so ist die Abbildung $L: \mathbb{R}^m \rightarrow \mathbb{R}^n$,

$$\underline{\mathbf{x}} \mapsto A\underline{\mathbf{x}}$$

linear, wie man leicht nachrechnet bzw. wie bereits gezeigt wurde.

Daher liegt es nahe, einige Begriffsbildungen für Matrizen auf lineare Abbildungen zu übertragen:

Es sei $L: \mathbb{R}^m \rightarrow \mathbb{R}^n$ eine lineare Abbildung.

i) Dann ist der **Kern** der Abbildung L per definitionem

$$\text{kern } L := L^{-1}(\{\underline{\mathbf{0}}\}) := \{\underline{\mathbf{x}} \in \mathbb{R}^m : L(\underline{\mathbf{x}}) = \underline{\mathbf{0}}\}.$$

ii) Das **Bild** der Abbildung ist

$$\text{bild } L := \{\underline{\mathbf{y}} \in \mathbb{R}^n : \underline{\mathbf{y}} = L(\underline{\mathbf{x}}) \text{ für ein } \underline{\mathbf{x}} \in \mathbb{R}^m\}.$$

iii) Als **Rang** der Abbildung L definiert man

$$\text{rg } L := \dim(\text{bild } L).$$

Die Motivation für die Bezeichnung „Rang einer linearen Abbildung“ im Vergleich zur Definition 1.3 des Rangs einer Matrix wird in Kürze deutlich.

Betrachtet man beispielsweise die lineare Abbildung $L: \mathbb{R}^2 \rightarrow \mathbb{R}$, $L(\underline{\mathbf{x}}) := x_1 - x_2$, so ist $\underline{\mathbf{x}} \in \text{kern } L$ genau dann, wenn gilt $x_1 = x_2$, d.h.

$$\text{kern } L = \left\{ \lambda \begin{pmatrix} 1 \\ 1 \end{pmatrix} : \lambda \in \mathbb{R} \right\}.$$

In dem Beispiel ist bild $L = \mathbb{R}$.

Man erkennt hier, dass der Kern und das Bild Unterräume sind (eine Vektorraumstruktur haben), wodurch der Begriff „Dimension“ überhaupt erst definiert ist.

Tatsächlich ist als Übungsaufgabe zu diesem Kapitel zu zeigen, dass der Kern und das Bild einer linearen Abbildung $L: \mathbb{R}^m \rightarrow \mathbb{R}^n$ Unterräume des \mathbb{R}^m bzw. des \mathbb{R}^n sind.

Der Rangsatz.

Um den so genannten **Rangsatz** als Gegenstück zu Satz 1.3 zu verstehen, beantworte man zunächst die folgenden beiden Fragen (vgl. Übungskapitel 3.3):

Es sei L eine lineare Abbildung vom \mathbb{R}^m in den \mathbb{R}^n .

- i) Ist es möglich, dass zwei linear abhängige Vektoren $\underline{\mathbf{x}}, \underline{\mathbf{y}} \in \mathbb{R}^m$ auf zwei linear unabhängige Vektoren $\underline{\mathbf{u}}, \underline{\mathbf{v}} \in \mathbb{R}^n$ abgebildet werden?
- ii) Ist es möglich, dass zwei linear unabhängige Vektoren $\underline{\mathbf{x}}, \underline{\mathbf{y}} \in \mathbb{R}^m$ auf zwei linear abhängige Vektoren $\underline{\mathbf{u}}, \underline{\mathbf{v}} \in \mathbb{R}^n$ abgebildet werden?

Falls ja, kann man daraus eine Aussage über kern L ableiten?

Satz 3.1. RANGSATZ

Es sei $L: \mathbb{R}^m \rightarrow \mathbb{R}^n$ eine lineare Abbildung. Dann ist

$$\dim(\text{kern } L) + \text{rg } L = m .$$

Der Rangsatz ist in Abbildung 3.1 anhand des einfachen Beispiels

$$L : \mathbb{R}^3 \rightarrow \mathbb{R}^3, \quad L(\underline{\mathbf{x}}) = \begin{pmatrix} x_1 \\ x_2 \\ 0 \end{pmatrix}$$

illustriert: Der Kern ist ein eindimensionaler Unterraum (die rot dargestellte x_3 -Achse), das Bild ist die blau skizzierte x_1, x_2 -Ebene, also ein zweidimensionaler Unterraum.

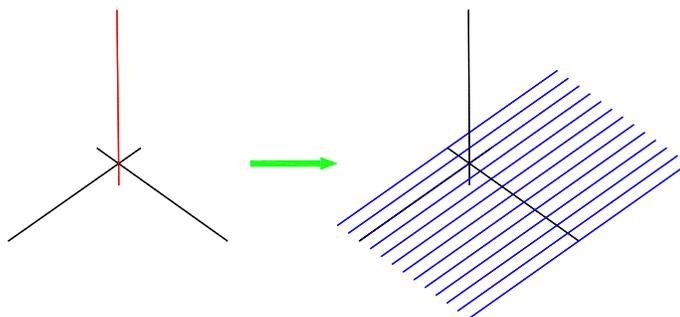


Abbildung 3.1: Ein einfaches Beispiel um Rangsatz.

Als Korollar folgt aus dem Rangsatz unmittelbar die Äquivalenz von [Injektivität](#), [Surjektivität](#) und damit [Bijektivität](#) linearer Abbildungen.

Korollar 3.1.

INJEKTIVITÄT VERSUS SURJEKTIVITÄT

Es sei $n = m$ und $L: \mathbb{R}^n \rightarrow \mathbb{R}^n$ sei eine lineare Abbildung.

Dann ist L injektiv (d.h. $\dim(\text{kern } L) = 0$) genau dann, wenn L surjektiv ist (d.h. $\text{rg } L = n$).

Beweis. Übungskapitel 3.3. □

3.2 Matrixdarstellung (Koordinaten; darstellende Matrix; Basiswechsel; Koordinatentransformation; Transformation der darstellenden Matrix; Tensor; kovariant; kontravariant)

Wie gesagt ist eine lineare Abbildung $f: \mathbb{R} \rightarrow \mathbb{R}$ von der Form

$$f(x) = ax, \quad a \in \mathbb{R} \text{ fixiert.}$$

Insbesondere ist die Abbildung allein durch den Parameter a charakterisiert.

Kann diese Beobachtung verallgemeinert werden?

Im Folgenden sei $L: \mathbb{R}^m \rightarrow \mathbb{R}^n$ eine lineare Abbildung.

Betrachtet seien weiter eine Basis

$$\mathcal{V} = (\underline{\mathbf{v}}^{(1)}, \dots, \underline{\mathbf{v}}^{(m)})$$

des \mathbb{R}^m und eine Basis

$$\mathcal{W} = (\underline{\mathbf{w}}^{(1)}, \dots, \underline{\mathbf{w}}^{(n)})$$

des \mathbb{R}^n .

Bzgl. der Basis \mathcal{V} des \mathbb{R}^m ist jedes $\underline{\mathbf{x}} \in \mathbb{R}^m$ darstellbar als

$$\underline{\mathbf{x}} = \sum_{j=1}^m \alpha_j \underline{\mathbf{v}}^{(j)}, \quad \alpha_j \in \mathbb{R},$$

und für die lineare Abbildung L folgt nach Definition 3.1

$$L(\underline{\mathbf{x}}) = L\left(\sum_{j=1}^m \alpha_j \underline{\mathbf{v}}^{(j)}\right) = \sum_{j=1}^m \alpha_j L(\underline{\mathbf{v}}^{(j)}).$$

Beobachtung. $L(\underline{x})$ ist demnach für alle $\underline{x} \in \mathbb{R}^m$ festgelegt, wenn nur die Bilder $L(\underline{v}^{(j)})$, $j = 1, \dots, m$, irgendeiner Basis des \mathbb{R}^m bekannt sind.

Für jedes fixierte $1 \leq j \leq m$ kann das Bild $L(\underline{v}^{(j)}) \in \mathbb{R}^n$ aber wiederum bzgl. der Basis \mathcal{W} des \mathbb{R}^n in der Form

$$L(\underline{v}^{(j)}) = \sum_{i=1}^n a_{ij} \underline{w}^{(i)}, \quad a_{ij} \in \mathbb{R},$$

dargestellt werden, wobei die a_{ij} von den gewählten Basen \mathcal{V} und \mathcal{W} abhängen.

Zwei Beispiele.

i) Es sei $\mathcal{E} = (\underline{e}^{(1)}, \underline{e}^{(2)})$ die kanonische Basis des \mathbb{R}^2 , $\mathcal{F} = (\underline{f}^{(1)}, \underline{f}^{(2)}, \underline{f}^{(3)})$ sei die kanonische Basis des \mathbb{R}^3 .

Weiterhin sei $L: \mathbb{R}^2 \rightarrow \mathbb{R}^3$ eine lineare Abbildung und es gelte (wie bereits gesagt, ist durch diese Wahl die lineare Abbildung eindeutig bestimmt)

$$L(\underline{e}^{(1)}) = \underline{f}^{(1)} + \underline{f}^{(3)} =: \underline{y}, \quad L(\underline{e}^{(2)}) = \underline{f}^{(1)} + 2\underline{f}^{(2)} =: \underline{z}.$$

Mit obiger Notation ist

$$a_{11} = 1, \quad a_{21} = 0, \quad a_{31} = 1; \quad a_{12} = 1, \quad a_{22} = 2, \quad a_{32} = 0.$$

Bezüglich der kanonischen Basis \mathcal{E} des \mathbb{R}^2 wird ein beliebiges $\underline{x} \in \mathbb{R}^2$ nun in **Koordinaten** als

$$\underline{x} = x_1 \underline{e}^{(1)} + x_2 \underline{e}^{(2)} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}_{\mathcal{E}},$$

geschrieben.

Die Vektoren $\underline{\mathbf{y}}, \underline{\mathbf{z}} \in \mathbb{R}^3$ werden in Koordinaten bzgl. der kanonischen Basis \mathcal{F} des \mathbb{R}^3 geschrieben:

$$\underline{\mathbf{y}} = \begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}_{\mathcal{F}},$$

$$\underline{\mathbf{z}} = \begin{pmatrix} a_{12} \\ a_{22} \\ a_{32} \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}_{\mathcal{F}}.$$

Bildet man aus diesen Spaltenvektoren eine Matrix A , so folgt für alle $\underline{\mathbf{x}} \in \mathbb{R}^2$

$$L(\underline{\mathbf{x}}) = x_1 L(\underline{\mathbf{e}}^{(1)}) + x_2 L(\underline{\mathbf{e}}^{(2)}) = \begin{pmatrix} 1 & 1 \\ 0 & 2 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = A \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

Beobachtung. Die lineare Abbildung L kann mit einer Matrix A identifiziert werden.

Die Bilder der Basis des \mathbb{R}^2 werden als Spaltenvektoren im \mathbb{R}^3 geschrieben und aus diesen Spaltenvektoren setzt sich die Matrix A zusammen.

ii) Man betrachte die gleiche lineare Abbildung wie im ersten Beispiel.

Als Basis des \mathbb{R}^2 diene aber jetzt $\mathcal{G} := (\underline{\mathbf{g}}^{(1)}, \underline{\mathbf{g}}^{(2)})$,

$$\underline{\mathbf{g}}^{(1)} = \underline{\mathbf{e}}^{(1)} + \underline{\mathbf{e}}^{(2)}, \quad \underline{\mathbf{g}}^{(2)} = \underline{\mathbf{e}}^{(1)} - \underline{\mathbf{e}}^{(2)},$$

als Basis des \mathbb{R}^3 werde $\mathcal{H} = (\underline{\mathbf{h}}^{(1)}, \underline{\mathbf{h}}^{(2)}, \underline{\mathbf{h}}^{(3)})$,

$$\underline{\mathbf{h}}^{(1)} = \underline{\mathbf{f}}^{(1)} + \underline{\mathbf{f}}^{(3)}, \quad \underline{\mathbf{h}}^{(2)} = \underline{\mathbf{f}}^{(2)} + \underline{\mathbf{f}}^{(3)}, \quad \underline{\mathbf{h}}^{(3)} = \underline{\mathbf{f}}^{(3)},$$

gewählt.

Man berechnet

$$\begin{aligned} L(\underline{\mathbf{g}}^{(1)}) &= L(\underline{\mathbf{e}}^{(1)} + \underline{\mathbf{e}}^{(2)}) = L(\underline{\mathbf{e}}^{(1)}) + L(\underline{\mathbf{e}}^{(2)}) \\ &= 2\underline{\mathbf{f}}^{(1)} + 2\underline{\mathbf{f}}^{(2)} + \underline{\mathbf{f}}^{(3)} =: \underline{\mathbf{v}} \end{aligned}$$

und analog

$$L(\underline{\mathbf{g}}^{(2)}) = \underline{\mathbf{f}}^{(3)} - 2\underline{\mathbf{f}}^{(2)} =: \underline{\mathbf{w}}.$$

Nun sollen $\underline{\mathbf{v}}$ und $\underline{\mathbf{w}}$ bzgl. der neuen Basis \mathcal{H} dargestellt werden. Dazu beachtet man

$$\begin{aligned} \underline{\mathbf{f}}^{(1)} &= \underline{\mathbf{h}}^{(1)} - \underline{\mathbf{h}}^{(3)}, \\ \underline{\mathbf{f}}^{(2)} &= \underline{\mathbf{h}}^{(2)} - \underline{\mathbf{h}}^{(3)}, \\ \underline{\mathbf{f}}^{(3)} &= \underline{\mathbf{h}}^{(3)}, \end{aligned}$$

was eingesetzt

$$\begin{aligned} \underline{\mathbf{v}} &= 2\underline{\mathbf{h}}^{(1)} + 2\underline{\mathbf{h}}^{(2)} - 3\underline{\mathbf{h}}^{(3)}, \\ \underline{\mathbf{w}} &= 3\underline{\mathbf{h}}^{(3)} - 2\underline{\mathbf{h}}^{(2)} \end{aligned}$$

ergibt.

Schließlich wird für beliebiges $\underline{\mathbf{x}} \in \mathbb{R}^2$ bzw. für beliebiges $\underline{\mathbf{u}} \in \mathbb{R}^3$ die **Koordinatendarstellung bzgl. der Basis \mathcal{G}** bzw. bzgl. der Basis \mathcal{H} eingeführt, d.h.

$$\begin{aligned} \underline{\mathbf{x}} &= \alpha_1 \underline{\mathbf{g}}^{(1)} + \alpha_2 \underline{\mathbf{g}}^{(2)} =: \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix}_{\mathcal{G}}, \\ \underline{\mathbf{u}} &= \beta_1 \underline{\mathbf{h}}^{(1)} + \beta_2 \underline{\mathbf{h}}^{(2)} + \beta_3 \underline{\mathbf{h}}^{(3)} =: \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix}_{\mathcal{H}}. \end{aligned}$$

Insbesondere gilt

$$\underline{\mathbf{g}}^{(1)} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}_{\mathcal{G}}, \quad \underline{\mathbf{g}}^{(2)} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}_{\mathcal{G}}$$

sowie

$$\underline{\mathbf{v}} = \begin{pmatrix} 2 \\ 2 \\ -3 \end{pmatrix}_{\mathcal{H}}, \quad \underline{\mathbf{w}} = \begin{pmatrix} 0 \\ -2 \\ 3 \end{pmatrix}_{\mathcal{H}}.$$

Im dieser Notation kann die lineare Abbildung bei fixierten Basen \mathcal{G} und \mathcal{H} als

$$\begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix}_{\mathcal{G}} \mapsto \begin{pmatrix} 2 & 0 \\ 2 & -2 \\ -3 & 3 \end{pmatrix}_{\mathcal{H}}^{\mathcal{G}} \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix}_{\mathcal{G}}$$

geschrieben werden.

Mit anderen Worten: Die Matrix

$$\begin{pmatrix} 2 & 0 \\ 2 & -2 \\ -3 & 3 \end{pmatrix}_{\mathcal{H}}^{\mathcal{G}}$$

repräsentiert die lineare Abbildung L bzgl. der Basen \mathcal{G} und \mathcal{H} .

In der Tat berechnet man

$$\begin{pmatrix} 2 & 0 \\ 2 & -2 \\ -3 & 3 \end{pmatrix}_{\mathcal{H}}^{\mathcal{G}} \begin{pmatrix} 1 \\ 0 \end{pmatrix}_{\mathcal{G}} = \begin{pmatrix} 2 \\ 2 \\ -3 \end{pmatrix}_{\mathcal{H}},$$

$$\begin{pmatrix} 2 & 0 \\ 2 & -2 \\ -3 & 3 \end{pmatrix}_{\mathcal{H}}^{\mathcal{G}} \begin{pmatrix} 0 \\ 1 \end{pmatrix}_{\mathcal{G}} = \begin{pmatrix} 0 \\ -2 \\ 3 \end{pmatrix}_{\mathcal{H}}$$

als alternative Schreibweise für $L(\underline{\mathbf{g}}^{(1)}) = \underline{\mathbf{v}}$ und $L(\underline{\mathbf{g}}^{(2)}) = \underline{\mathbf{w}}$, wodurch die lineare Abbildung eindeutig bestimmt ist (s.o.).

Das allgemeine Schema.

Das allgemeine Schema zur **Darstellung einer linearen Abbildung als Matrix** bzgl. gegebener Basen lautet:

- i) Es sei \mathcal{G} eine Basis des \mathbb{R}^m , \mathcal{H} eine Basis \mathbb{R}^n und $L: \mathbb{R}^m \rightarrow \mathbb{R}^n$ sei eine lineare Abbildung.
- ii) Man berechne die Bilder (unter L) der Basisvektoren von \mathcal{G} .
- iii) Von diesen Bildern wird die Koordinatendarstellung bzgl. \mathcal{H} berechnet.
- iv) Die Matrixdarstellung $A_{\mathcal{H}}^{\mathcal{G}} \in M(n, m)$ der linearen Abbildung L bzgl. der Basen \mathcal{G} und \mathcal{H} erhält man, indem man aus diesen Spaltenvektoren eine Matrix bildet.

Die darstellende Matrix enthält mit anderen Worten die **Koordinaten der Bilder der Basisvektoren als Spaltenvektoren**.

- v) Ein beliebiges $\underline{\mathbf{x}} \in \mathbb{R}^m$ wird **mit seiner Koordinatendarstellung**

$$\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_m \end{pmatrix}_{\mathcal{G}} \text{ bzgl. der Basis } \mathcal{G} \text{ identifiziert.}$$

- vi) Dann liefert $A_{\mathcal{H}}^{\mathcal{G}} \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_m \end{pmatrix}_{\mathcal{G}}$ die **Koordinatendarstellung von $L(\underline{\mathbf{x}})$ bzgl. der Basis \mathcal{H}** .

Basiswechsel und Koordinatentransformation.

Beim Übergang von Beispiel i) zu Beispiel ii) wurde dasselbe Problem bzgl. anderer Basen betrachtet, d.h. mithilfe eines **Basiswechsels** transformiert.

Ein solcher Basiswechsel kann konkrete Rechnungen erheblich vereinfachen, falls eine neue Basis die physikalischen oder geometrischen

Eigenschaften eines Problems widerspiegelt.

Mit dem Basiswechsel ändern sich die Koordinaten eines Vektors und man spricht von einer **Koordinatentransformation**.

Der Zusammenhang zwischen Basiswechsel und Koordinatentransformation wird nun systematisiert.

Es seien dazu $\mathcal{A} = (\underline{\mathbf{a}}^{(1)}, \dots, \underline{\mathbf{a}}^{(m)})$ und $\mathcal{B} = (\underline{\mathbf{b}}^{(1)}, \dots, \underline{\mathbf{b}}^{(m)})$ zwei Basen des \mathbb{R}^m .

Die „neuen“ Basisvektoren der Basis \mathcal{B} seien in Abhängigkeit der „alten“ Basisvektoren aus \mathcal{A} gegeben, d.h. mit Koeffizienten $\gamma_{ij} \in \mathbb{R}$, $1 \leq i, j \leq m$, gelte

$$\begin{aligned} \underline{\mathbf{b}}^{(1)} &= \gamma_{11}\underline{\mathbf{a}}^{(1)} + \dots + \gamma_{1m}\underline{\mathbf{a}}^{(m)}, \\ &\vdots \\ \underline{\mathbf{b}}^{(m)} &= \gamma_{m1}\underline{\mathbf{a}}^{(1)} + \dots + \gamma_{mm}\underline{\mathbf{a}}^{(m)}. \end{aligned}$$

Aus diesen Koeffizienten entsteht die so genannte **Transformationsmatrix** als die **transponierte Matrix** der Koeffizienten, d.h. man definiert

$$S := (\gamma_{ij})^T = \begin{pmatrix} \gamma_{11} & \dots & \gamma_{m1} \\ \vdots & & \vdots \\ \gamma_{1m} & \dots & \gamma_{mm} \end{pmatrix} \in M(m, m).$$

Hat nun $\underline{\mathbf{x}} \in \mathbb{R}^m$ bzgl. \mathcal{A} bzw. \mathcal{B} die Koordinatendarstellungen

$$\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_m \end{pmatrix}_{\mathcal{A}}, \quad \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_m \end{pmatrix}_{\mathcal{B}},$$

so berechnet man

$$\begin{aligned}\underline{\mathbf{x}} &= \alpha_1 \underline{\mathbf{a}}^{(1)} + \cdots + \alpha_m \underline{\mathbf{a}}^{(m)} = \beta_1 \underline{\mathbf{b}}^{(1)} + \cdots + \beta_m \underline{\mathbf{b}}^{(m)} \\ &= \sum_{j=1}^m \beta_j \left(\sum_{i=1}^m \gamma_{ji} \underline{\mathbf{a}}^{(i)} \right) \\ &= \sum_{i=1}^m \left(\sum_{j=1}^m \gamma_{ji} \beta_j \right) \underline{\mathbf{a}}^{(i)}.\end{aligned}$$

Mit anderen Worten gilt die Beziehung

$$\alpha_i = \sum_{j=1}^m \gamma_{ji} \beta_j.$$

zwischen den Koordinaten bzgl. \mathcal{A} und \mathcal{B} .

Die zugehörige Matrixschreibweise lautet

$$\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_m \end{pmatrix}_{\mathcal{A}} = S \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_m \end{pmatrix}_{\mathcal{B}}, \quad \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_m \end{pmatrix}_{\mathcal{B}} = S^{-1} \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_m \end{pmatrix}_{\mathcal{A}},$$

wobei die Invertierbarkeit von S unmittelbar daraus folgt, dass sowohl \mathcal{A} als auch \mathcal{B} Basen des \mathbb{R}^m sind.

Beispiel. Sind die „neuen“ Basisvektoren lediglich Vielfache der „alten“ Basisvektoren im Sinne von $\underline{\mathbf{b}}^{(j)} = \lambda \underline{\mathbf{a}}^{(j)}$, $j = 1, \dots, m$, $\lambda \in \mathbb{R}$, $\lambda \neq 0$, so ist

$$S = \lambda I_m, \quad S^{-1} = \frac{1}{\lambda} I_m,$$

und man erkennt, dass die Koordinaten mit dem Kehrwert von λ skaliert werden.

In gewissem Sinne verhalten sich also Basis- und Koordinatentransformationen konträr.

Transformation der darstellenden Matrix beim Basiswechsel.

Vollzieht man einen Basiswechsel im Urbild \mathbb{R}^m und/oder im Bild \mathbb{R}^n einer linearen Abbildung, so transformiert sich mit den Koordinaten auch die darstellende Matrix der linearen Abbildung.

Die Situation ist schematisch in Abbildung 3.2 dargestellt.

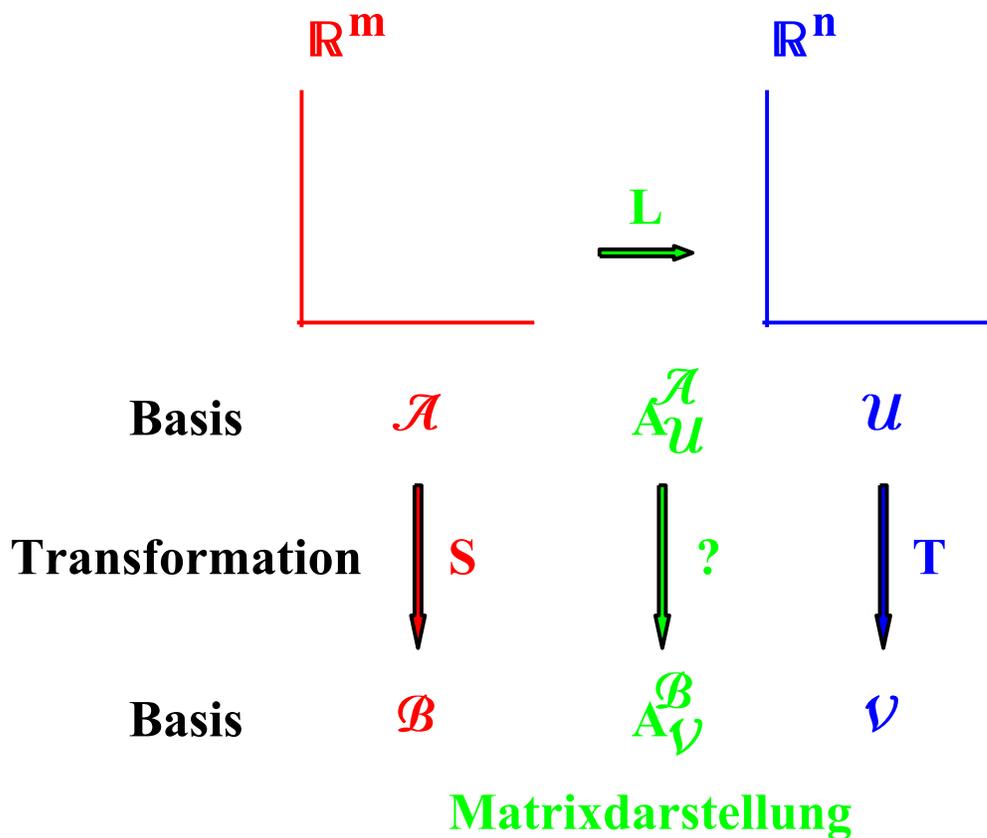


Abbildung 3.2: Zur Transformation der darstellenden Matrix.

Wie in der Abbildung angedeutet, sei $L: \mathbb{R}^m \rightarrow \mathbb{R}^n$ eine lineare Abbildung, $\mathcal{A} = (\underline{\mathbf{a}}^{(1)}, \dots, \underline{\mathbf{a}}^{(m)})$ sei eine Basis des \mathbb{R}^m , $\mathcal{U} = (\underline{\mathbf{u}}^{(1)}, \dots, \underline{\mathbf{u}}^{(n)})$ sei eine Basis des \mathbb{R}^n und $A_{\mathcal{U}}^{\mathcal{A}} \in M(n, m)$ sei die darstellende Matrix der linearen Abbildung L bzgl. dieser Basen.

Man betrachte weiterhin „neue“ Basen $\mathcal{B} = (\underline{\mathbf{b}}^{(1)}, \dots, \underline{\mathbf{b}}^{(m)})$ des \mathbb{R}^m bzw. $\mathcal{V} = (\underline{\mathbf{v}}^{(1)}, \dots, \underline{\mathbf{v}}^{(n)})$ des \mathbb{R}^n .

Gesucht ist die Matrixdarstellung $A_{\mathcal{V}}^{\mathcal{B}}$ von L bzgl. dieser Basen.

Der Basiswechsel von \mathcal{A} nach \mathcal{B} werde durch die Transformationsmatrix $S \in M(m, m)$ beschrieben, der Basiswechsel von \mathcal{U} nach \mathcal{V} analog durch die Transformationsmatrix $T \in M(n, n)$.

Ist $L(\underline{\mathbf{x}}) = \underline{\mathbf{y}}$ und

$$\begin{aligned}\underline{\mathbf{x}} &= \alpha_1 \underline{\mathbf{a}}^{(1)} + \dots + \alpha_m \underline{\mathbf{a}}^{(m)} \\ &= \beta_1 \underline{\mathbf{b}}^{(1)} + \dots + \beta_m \underline{\mathbf{b}}^{(m)}, \\ \underline{\mathbf{y}} &= \mu_1 \underline{\mathbf{u}}^{(1)} + \dots + \mu_n \underline{\mathbf{u}}^{(n)} \\ &= \nu_1 \underline{\mathbf{v}}^{(1)} + \dots + \nu_n \underline{\mathbf{v}}^{(n)},\end{aligned}$$

so gilt mit der darstellenden Matrix $A_{\mathcal{U}}^{\mathcal{A}} \in M(n, m)$

$$\begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix}_{\mathcal{U}} = A_{\mathcal{U}}^{\mathcal{A}} \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_m \end{pmatrix}_{\mathcal{A}} = A_{\mathcal{U}}^{\mathcal{A}} S \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_m \end{pmatrix}_{\mathcal{B}},$$

wobei die zweite Gleichheit aus der bereits gezeigten Koordinatentransformation folgt.

Eine Transformation der linken Seite ergibt schließlich

$$\begin{aligned}\begin{pmatrix} \nu_1 \\ \vdots \\ \nu_n \end{pmatrix}_{\mathcal{V}} &= T^{-1} \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix}_{\mathcal{U}} \\ &= T^{-1} A_{\mathcal{U}}^{\mathcal{A}} S \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_m \end{pmatrix}_{\mathcal{B}} = A_{\mathcal{V}}^{\mathcal{B}} \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_m \end{pmatrix}_{\mathcal{B}}.\end{aligned}$$

Zusammenfassend ist folgender Satz gezeigt:

Satz 3.2. BASISWECHSEL

- i) Es seien \mathcal{A} , \mathcal{B} Basen des \mathbb{R}^m und die Transformationsmatrix $S \in M(m, m)$ sei wie oben gegeben.

Ist

$$\underline{\mathbf{x}} = \alpha_1 \underline{\mathbf{a}}^{(1)} + \cdots + \alpha_m \underline{\mathbf{a}}^{(m)} = \beta_1 \underline{\mathbf{b}}^{(1)} + \cdots + \beta_m \underline{\mathbf{b}}^{(m)},$$

so gilt

$$\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_m \end{pmatrix} = S \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_m \end{pmatrix}, \quad \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_m \end{pmatrix} = S^{-1} \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_m \end{pmatrix}.$$

- ii) Sind weiterhin \mathcal{U} , \mathcal{V} Basen des \mathbb{R}^n , ist die Transformation dieser Basen durch die Matrix $T \in M(n, n)$ gegeben und ist $L: \mathbb{R}^m \rightarrow \mathbb{R}^n$ eine lineare Abbildung mit der Matrixdarstellung $A_{\mathcal{U}}^{\mathcal{A}} \in M(n, m)$ bzgl. der Basen \mathcal{A} , \mathcal{U} , so lautet die Matrixdarstellung von L bzgl. der Basen \mathcal{B} und \mathcal{V}

$$A_{\mathcal{V}}^{\mathcal{B}} = T^{-1} A_{\mathcal{U}}^{\mathcal{A}} S.$$

Bemerkung. Da die Spaltenvektoren aus den Koordinaten selbst als Vektoren aus dem \mathbb{R}^m bzw. \mathbb{R}^n aufgefasst werden, werden im obigen Satz die Indizes zur Basisbenennung weggelassen.

Zu den einführenden Beispielen.

Die Situation in den einführenden Beispielen i), ii) zu Beginn des Kapitels sieht wie folgt aus:

Der Basiswechsel von i) nach ii) wird beschrieben durch

$$S = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \quad T = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix},$$

die Inverse von T berechnet sich zu

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

Es folgt wie behauptet

$$\begin{aligned} A_{\mathcal{H}}^{\mathcal{G}} &= T^{-1} \begin{pmatrix} 1 & 1 \\ 0 & 2 \\ 1 & 0 \end{pmatrix} S \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 2 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \\ &= \begin{pmatrix} 2 & 0 \\ 2 & -2 \\ -3 & 3 \end{pmatrix}. \end{aligned}$$

Eine kurze Bemerkung zu kovarianten Tensoren erster Stufe.

Tensoren sind **Multilinearformen**, wobei es sich im einfachsten Fall eines **kovarianten Tensors erster Stufe** um eine lineare Abbildung vom \mathbb{R}^m in die reellen Zahlen handelt.

Eine solche lineare Abbildung wird bzgl. einer gegebenen Basis des \mathbb{R}^m durch einen Zeilenvektor repräsentiert. Dieser Zeilenvektor ändert sich wie oben gesehen bei einem Basiswechsel.

Eine physikalisch relevante Größe sollte jedoch unabhängig von der speziellen Wahl einer Basis sein, d.h.:

Der Tensor als lineare Abbildung ist ein physikalisch sinnvolles Objekt, wohingegen es sich bei einem Zeilenvektor nur um eine spezielle Darstellung bzgl. einer fixierten Basis handelt.

Man betrachte nun eine lineare Abbildung $L: \mathbb{R}^m \rightarrow \mathbb{R}$, eine Basis \mathcal{A} des \mathbb{R}^m sowie die Darstellung

$$(\xi_1 \ \dots \ \xi_m)_{\mathcal{A}}$$

der linearen Abbildung bzgl. \mathcal{A} als Zeilenvektor.

Ist \mathcal{B} eine weitere Basis des \mathbb{R}^m , so ist nach Satz 3.2 die Darstellung bzgl. dieser Basis

$$(\psi_1 \ \dots \ \psi_m)_{\mathcal{B}} = (\xi_1 \ \dots \ \xi_m) \begin{pmatrix} \gamma_{11} & \dots & \gamma_{m1} \\ \vdots & & \vdots \\ \gamma_{1m} & \dots & \gamma_{mm} \end{pmatrix},$$

wobei der Basiswechsel wie oben durch die Koeffizienten γ_{ij} , $1 \leq i, j \leq m$, gegeben sei.

Man erkennt für alle $i = 1, \dots, n$:

$$\psi_i = \sum_{j=1}^m \gamma_{ij} \xi_j.$$

Mit anderen Worten: Zeilenvektoren (die Koordinaten der linearen Abbildung) transformieren sich bei einem Basiswechsel genau wie die Basis selbst, weshalb man von einem kovarianten Tensor spricht.

Spaltenvektoren verhalten sich hingegen konträr, wie bei Koordinatentransformationen bereits festgestellt wurde.

Sie stellen so genannte **kontravariante Tensoren** erster Stufe bzgl. einer gegebenen Basis dar.

3.3 Übungsaufgaben zu Kapitel 3

Aufgabe 1.

- i) Geben Sie eine nicht-lineare Abbildung $f: \mathbb{R}^2 \rightarrow \mathbb{R}^4$ an.
- ii) Finden Sie eine lineare Abbildung $L: \mathbb{R} \rightarrow \mathbb{R}$ mit $L(3) = -2$. Zeigen Sie, dass es keine weitere lineare Abbildung $\tilde{L}: \mathbb{R} \rightarrow \mathbb{R}$ mit $\tilde{L}(3) = -2$ gibt.
- iii) Finden Sie eine lineare Abbildung $L: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ mit

$$L\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right) = \begin{pmatrix} 2 \\ 3 \end{pmatrix}, \quad L\left(\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right) = \begin{pmatrix} -2 \\ -3 \end{pmatrix}.$$

Ist diese eindeutig bestimmt? Bestimmen Sie den Kern und das Bild von L .

Aufgabe 2.* Zeigen Sie: Kern und Bild einer linearen Abbildung $L: \mathbb{R}^m \rightarrow \mathbb{R}^n$ sind Unterräume des \mathbb{R}^m bzw. des \mathbb{R}^n .

Aufgabe 3. Es sei L eine lineare Abbildung vom \mathbb{R}^m in den \mathbb{R}^n .

- i) Ist es möglich, dass zwei linear abhängige Vektoren $\underline{\mathbf{x}}, \underline{\mathbf{y}} \in \mathbb{R}^m$ auf zwei linear unabhängige Vektoren $\underline{\mathbf{u}}, \underline{\mathbf{v}} \in \mathbb{R}^n$ abgebildet werden?
- ii) Ist es möglich, dass zwei linear unabhängige Vektoren $\underline{\mathbf{x}}, \underline{\mathbf{y}} \in \mathbb{R}^m$ auf zwei linear abhängige Vektoren $\underline{\mathbf{u}}, \underline{\mathbf{v}} \in \mathbb{R}^n$ abgebildet werden?
Falls ja, kann man daraus eine Aussage über kern L ableiten?
-

Aufgabe 4.* Zeigen Sie Korollar 3.1.

Aufgabe 5. Es sei $a = 0$, $a = 1$ oder $a = 2$ fixiert und $L: \mathbb{R}^4 \rightarrow \mathbb{R}^5$ sei gegeben durch

$$L(\underline{\mathbf{x}}) = \begin{pmatrix} 0 \\ x_2 - x_3 \\ (x_1 + x_4)^a \\ x_1 + x_2 + x_3 + x_4 \\ x_1 + x_4 \end{pmatrix}.$$

i) Ist L eine lineare Abbildung?

ii) Nun sei $a = 1$: Durch welche Matrix wird L bzgl. der kanonischen Basen dargestellt? Bestimmen Sie kern L , $\text{rg } L$ und bild L .

Aufgabe 6.* Es bezeichne $\mathcal{E} = (\underline{\mathbf{e}}^{(1)}, \underline{\mathbf{e}}^{(2)})$ die kanonische Basis des \mathbb{R}^2 . Gegeben seien weiter die Basen $\mathcal{V} = (\underline{\mathbf{v}}^{(1)}, \underline{\mathbf{v}}^{(2)})$ und $\mathcal{W} = (\underline{\mathbf{w}}^{(1)}, \underline{\mathbf{w}}^{(2)})$ des \mathbb{R}^2 , wobei gelte

$$\begin{aligned} \underline{\mathbf{v}}^{(1)} &= \underline{\mathbf{e}}^{(1)} + \underline{\mathbf{e}}^{(2)}, & \underline{\mathbf{v}}^{(2)} &= \underline{\mathbf{e}}^{(1)} - \underline{\mathbf{e}}^{(2)}; \\ \underline{\mathbf{w}}^{(1)} &= \underline{\mathbf{e}}^{(2)}, & \underline{\mathbf{w}}^{(2)} &= 2\underline{\mathbf{e}}^{(1)} + \underline{\mathbf{e}}^{(2)}. \end{aligned}$$

Es sei $L: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ die lineare Abbildung mit

$$A_{\mathcal{W}}^{\mathcal{V}} = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$$

Bestimmen Sie $A_{\mathcal{V}}^{\mathcal{V}}$ und $A_{\mathcal{V}}^{\mathcal{W}}$.

Aufgabe 7.* Es sei

$$A = \begin{pmatrix} 1 & 2 & 2 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix} \in M(3, 3).$$

Weiterhin bezeichne $\mathcal{E} = (\underline{\mathbf{e}}^{(1)}, \underline{\mathbf{e}}^{(2)}, \underline{\mathbf{e}}^{(3)})$ die kanonische Basis des \mathbb{R}^3 und gegeben sei zudem die Basis $\mathcal{V} = (\underline{\mathbf{v}}^{(1)}, \underline{\mathbf{v}}^{(2)}, \underline{\mathbf{v}}^{(3)})$ mit

$$\underline{\mathbf{v}}^{(1)} = \underline{\mathbf{e}}^{(1)}, \quad \underline{\mathbf{v}}^{(2)} = \underline{\mathbf{e}}^{(1)} + \underline{\mathbf{e}}^{(2)}, \quad \underline{\mathbf{v}}^{(3)} = \underline{\mathbf{e}}^{(1)} + \underline{\mathbf{e}}^{(2)} + \underline{\mathbf{e}}^{(3)}.$$

Die Matrixdarstellung $A_{\mathcal{E}}^{\mathcal{E}}$ einer linearen Abbildung $L: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ sei durch obige Matrix A gegeben. Bestimmen Sie $A_{\mathcal{V}}^{\mathcal{V}}$. Berechnen Sie die Koordinatendarstellungen von $L(\underline{\mathbf{v}}^{(3)})$ bzgl. der Basen \mathcal{E} und \mathcal{V} .

Aufgabe 8.* Es bezeichne $\mathcal{E} = (\underline{\mathbf{e}}^{(1)}, \underline{\mathbf{e}}^{(2)})$ die kanonische Basis des \mathbb{R}^2 . Es sei weiter $L: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ die lineare Abbildung mit $L(\underline{\mathbf{e}}^{(1)}) = \underline{\mathbf{e}}^{(1)}$ und $L(\underline{\mathbf{e}}^{(2)}) = \underline{\mathbf{e}}^{(2)} - \underline{\mathbf{e}}^{(1)}$. Bezüglich welcher Basis $\mathcal{V} = (\underline{\mathbf{v}}^{(1)}, \underline{\mathbf{v}}^{(2)})$ hat L die Matrixdarstellung

$$A_{\mathcal{V}}^{\mathcal{E}} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} ?$$

Aufgabe 9. Man betrachte die lineare Abbildung $L: \mathbb{R}^3 \rightarrow \mathbb{R}^2$, die bestimmt ist durch

$$L(\underline{\mathbf{e}}^{(1)}) = \underline{\mathbf{f}}^{(1)} + \underline{\mathbf{f}}^{(2)}, \quad L(\underline{\mathbf{e}}^{(2)}) = 2\underline{\mathbf{f}}^{(1)} + \underline{\mathbf{f}}^{(2)}, \quad L(\underline{\mathbf{e}}^{(3)}) = -\underline{\mathbf{f}}^{(1)} + \underline{\mathbf{f}}^{(2)}.$$

Dabei bezeichne $(\underline{\mathbf{e}}^{(1)}, \underline{\mathbf{e}}^{(2)}, \underline{\mathbf{e}}^{(3)})$ die kanonische Basis des \mathbb{R}^3 , $(\underline{\mathbf{f}}^{(1)}, \underline{\mathbf{f}}^{(2)})$ bezeichne die kanonische Basis des \mathbb{R}^2 .

- i) Bestimmen Sie die Matrixdarstellung der linearen Abbildung bzgl. der kanonischen Basen.
- ii) Bestimmen Sie die Matrixdarstellung der linearen Abbildung bzgl. der Basen $(\underline{\mathbf{v}}^{(1)}, \underline{\mathbf{v}}^{(2)}, \underline{\mathbf{v}}^{(3)})$ des \mathbb{R}^3 und $(\underline{\mathbf{w}}^{(1)}, \underline{\mathbf{w}}^{(2)})$ des \mathbb{R}^2 , wobei gelte

$$\begin{aligned} \underline{\mathbf{v}}^{(1)} &= \underline{\mathbf{e}}^{(1)} + \underline{\mathbf{e}}^{(2)} + \underline{\mathbf{e}}^{(3)}, & \underline{\mathbf{v}}^{(2)} &= \underline{\mathbf{e}}^{(1)} - \underline{\mathbf{e}}^{(3)}, & \underline{\mathbf{v}}^{(3)} &= 5\underline{\mathbf{e}}^{(2)} + \underline{\mathbf{e}}^{(3)}; \\ \underline{\mathbf{w}}^{(1)} &= -\underline{\mathbf{f}}^{(1)} - \underline{\mathbf{f}}^{(2)}, & \underline{\mathbf{w}}^{(2)} &= \underline{\mathbf{f}}^{(1)} + 4\underline{\mathbf{f}}^{(2)}. \end{aligned}$$
- iii) Finden Sie eine Basis $(\underline{\mathbf{g}}^{(1)}, \underline{\mathbf{g}}^{(2)})$ des \mathbb{R}^2 , sodass die Matrixdarstellung von L bzgl. der kanonischen Basis des \mathbb{R}^3 und bzgl. $(\underline{\mathbf{g}}^{(1)}, \underline{\mathbf{g}}^{(2)})$ gegeben ist durch

$$\begin{pmatrix} 2 & 3 & 0 \\ 0 & 1 & -2 \end{pmatrix}.$$

Lösungshinweise zu den Übungsaufgaben.

Aufgabe 2. Man vergleiche die Aufgabe mit der Diskussion linearer Gleichungssysteme (z.B. mit dem Superpositionsprinzip im homogenen Fall).

Aufgabe 4. Ist L injektiv, so ist für alle $\underline{\mathbf{x}} \in \mathbb{R}^n$, $\underline{\mathbf{x}} \neq \underline{\mathbf{0}}$,

$$L(\underline{\mathbf{x}}) \neq L(\underline{\mathbf{0}}) = \underline{\mathbf{0}},$$

d.h. kern $L = \{\underline{\mathbf{0}}\}$. Aus dem Rangsatz folgt, dass L auch surjektiv ist.

Ist L surjektiv, so ist nach dem Rangsatz kern $L = \{\underline{\mathbf{0}}\}$. Aus $L(\underline{\mathbf{x}}) = L(\underline{\mathbf{y}})$ folgt aber $L(\underline{\mathbf{x}} - \underline{\mathbf{y}}) = \underline{\mathbf{0}}$ und somit $\underline{\mathbf{x}} = \underline{\mathbf{y}}$. \square

Aufgabe 6. Man berechnet

$$\underline{\mathbf{w}}^{(1)} = \frac{\underline{\mathbf{v}}^{(1)} - \underline{\mathbf{v}}^{(2)}}{2}, \quad \underline{\mathbf{w}}^{(2)} = \frac{3}{2}\underline{\mathbf{v}}^{(1)} + \frac{1}{2}\underline{\mathbf{v}}^{(2)},$$

woraus sich als Transformationsmatrix S von \mathcal{V} nach \mathcal{W} bzw. als deren Inverse S^{-1}

$$S = \begin{pmatrix} \frac{1}{2} & \frac{3}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix}, \quad S^{-1} = \begin{pmatrix} \frac{1}{2} & -\frac{3}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

ergibt. Ist T die Transformationsmatrix von \mathcal{W} nach \mathcal{V} , so ist $T = S^{-1}$ (Probe!).

Es folgt

$$A_{\mathcal{V}}^{\mathcal{V}} = SA_{\mathcal{W}}^{\mathcal{V}}I_2 = \frac{1}{2} \begin{pmatrix} 7 & 5 \\ 1 & -1 \end{pmatrix}$$

und

$$A_{\mathcal{V}}^{\mathcal{W}} = SA_{\mathcal{W}}^{\mathcal{W}}S = \frac{1}{2} \begin{pmatrix} 1 & 13 \\ 1 & 1 \end{pmatrix}.$$

Aufgabe 7. Für die Transformationsmatrizen gilt mit der Notation dieses Kapitels

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

Als Inverse berechnet man

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

Damit ergibt sich

$$A_{\mathcal{V}}^{\mathcal{Y}} = S^{-1} A_{\mathcal{E}}^{\mathcal{X}} S = \begin{pmatrix} 1 & 2 & 3 \\ -1 & 0 & 0 \\ 1 & 1 & 2 \end{pmatrix} .$$

Es folgt

$$L(\underline{\mathbf{v}}^{(3)}) = \begin{pmatrix} 3 \\ 0 \\ 2 \end{pmatrix}_{\mathcal{V}}$$

und wegen

$$3\underline{\mathbf{v}}^{(1)} + 2\underline{\mathbf{v}}^{(3)} = 5\underline{\mathbf{e}}^{(1)} + 2\underline{\mathbf{e}}^{(2)} + 2\underline{\mathbf{e}}^{(3)}$$

ist

$$L(\underline{\mathbf{v}}^{(3)}) = \begin{pmatrix} 5 \\ 2 \\ 2 \end{pmatrix}_{\mathcal{E}}$$

Aufgabe 8. Aus

$$A_{\mathcal{V}}^{\mathcal{E}} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$$

folgt

$$\begin{aligned} \underline{\mathbf{e}}^{(1)} &= L(\underline{\mathbf{e}}^{(1)}) = \underline{\mathbf{v}}^{(1)} + \underline{\mathbf{v}}^{(2)} , \\ \underline{\mathbf{e}}^{(2)} - \underline{\mathbf{e}}^{(1)} &= L(\underline{\mathbf{e}}^{(2)}) = \underline{\mathbf{v}}^{(1)} + 2\underline{\mathbf{v}}^{(2)} , \end{aligned}$$

also

$$\begin{aligned} \underline{\mathbf{v}}^{(1)} &= 3\underline{\mathbf{e}}^{(1)} - \underline{\mathbf{e}}^{(2)} , \\ \underline{\mathbf{v}}^{(2)} &= \underline{\mathbf{e}}^{(2)} - 2\underline{\mathbf{e}}^{(1)} . \end{aligned}$$

Kapitel 4

Stetige Funktionen

An die Diskussion linearer Abbildungen schließt sich das Studium der nächsten besonderen Klasse von Funktionen an, der Klasse der **stetigen Funktionen**.

Während lineare Abbildungen mit Methoden der **linearen Algebra** untersucht wurden, bilden nun Grenzwerte von Folgen und Reihen die Grundlage für den tieferen Einstieg in die **Infinitesimalrechnung**, d.h. in die Differential- und Integralrechnung.

Dabei wird in diesem Kapitel der Fall **reellwertiger Funktionen einer Veränderlichen** betrachtet.

4.1 Beispiele, Definition und erste Eigenschaften

(Lipschitz-stetige Funktion; Grenzwert und Stetigkeit einer Funktion; gleichmäßige Stetigkeit)

Der intuitive Begriff der **Stetigkeit** einer Funktion ist der folgende:

Der Graph einer stetigen Funktion kann „ohne Absetzen des Stiftes“ gezeichnet werden (vgl. Abbildungen 4.1, 4.2, 4.3).

Die Grenzen dieses intuitiven Begriffes werden aber deutlich, wenn etwa die Funktion $f: \mathbb{R} \rightarrow \mathbb{R}$,

$$f(x) = \begin{cases} x \cos(1/x), & \text{falls } x \neq 0, \\ 0, & \text{falls } x = 0, \end{cases}$$

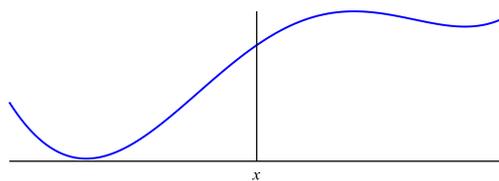


Abbildung 4.1: Der Graph einer stetigen Funktion.

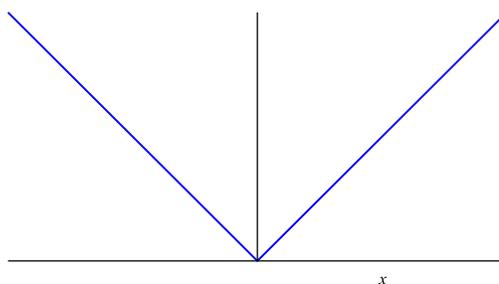


Abbildung 4.2: Der Graph der Betragsfunktion.

betrachtet wird (vgl. Abbildung 4.4).

Hier kann dem Graphen nicht unmittelbar angesehen werden, ob die Funktion stetig ist oder nicht.

Lange Zeit ist man jedoch davon ausgegangen, dass die Natur zumindest für „vernünftige“ Problemstellungen „glatte“ (d.h. stetige und je nach Daten bessere) Lösungen liefert.¹

In der Tat gibt es aber „natürliche Probleme“, die unstetige Lösungen produzieren, d.h. die Lösungen haben so genannte **Singularitäten**. Das prominenteste Beispiel ist von der Struktur $x/|x|$.

¹David Hilbert formulierte z.B. 1900 auf dem „Internationalen Mathematikerkongreß“ in Paris die berühmten 23 Probleme, die die Mathematik des zwanzigsten Jahrhunderts angehen sollte. Das 19. Problem lautet: „Sind die Lösungen regulärer Variationsprobleme stets notwendig analytisch?“

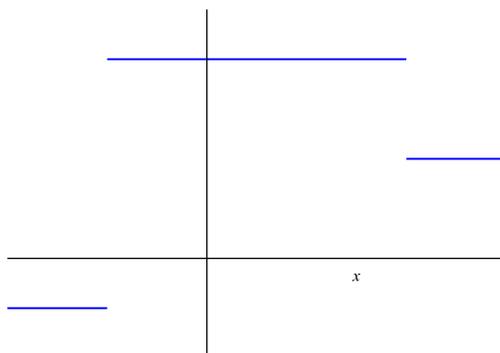
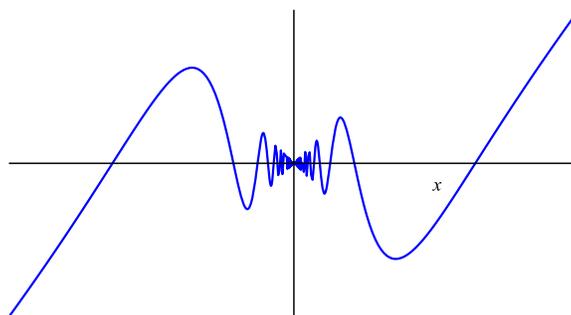


Abbildung 4.3: Der Graph einer unstetigen Funktion mit Sprungstellen.

Abbildung 4.4: Die Funktion $f(x) = x \cos(1/x)$.

Eine besondere Klasse stetiger Funktionen.

Unter den stetigen Funktionen gibt es eine besondere Klasse von Funktionen, deren Graph anschaulich gesprochen „Knickstellen“ haben kann.

Es handelt sich um so genannte **Lipschitz-stetige** Funktionen, bei denen der Abstand der Funktionswerte nach oben durch den Abstand der Argumente abgeschätzt werden kann.

D.h.: Ist $I \subset \mathbb{R}$ ein (evtl. verallgemeinertes) Intervall, so heißt eine Funktion $f: I \rightarrow \mathbb{R}$ Lipschitz-stetig, falls es eine Konstante $K \in \mathbb{R}$ gibt, sodass

$$|f(y) - f(x)| \leq K|y - x| \quad \text{für alle } x, y \in I .$$

Mit anderen Worten: Ist $\varepsilon > 0$ beliebig (klein) vorgegeben, so haben die Funktionswerte $f(y)$ und $f(x)$ einen Abstand kleiner als ε , falls nur die Argumente x und y einen Abstand kleiner als $\delta = \varepsilon/K$ haben.

Der bekannteste und wichtigste Vertreter der Klasse der Lipschitz-stetigen Funktionen ist der Betrag:

Die **Betragsfunktion** ist Lipschitz-stetig, wobei $K = 1$ gewählt werden kann (vgl. Abbildung 4.5 und Übungskapitel 4.3).

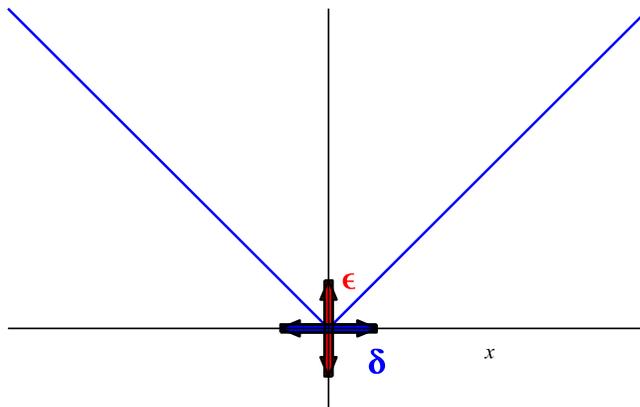


Abbildung 4.5: Die Betragsfunktion ist Lipschitz-stetig mit $K = 1$.

Grenzwert und Stetigkeit einer Funktion.

Aber nicht jede stetige Funktion ist Lipschitz-stetig (vgl. Übungskapitel 4.3).

Die allgemeine Definition der Stetigkeit erfordert zunächst den Begriff des **Grenzwertes** einer Funktion.

Im Folgenden sei $I \subset \mathbb{R}$ **stets ein (evtl. verallgemeinertes) Intervall** – dies dient lediglich der Vereinfachung der Notation und ist nicht grundlegend für die Definition.

Weiter sei in diesem Kapitel auch ohne besondere Erwähnung $f: I \rightarrow \mathbb{R}$.

Beobachtung. Ist $\{x_n\}$ eine Folge von Elementen aus I , so ist $\{f(x_n)\}$ eine reelle Zahlenfolge, die auf Konvergenz untersucht werden kann.

Definition 4.1. GRENZWERT UND STETIGKEIT

i) Es seien $x_0 \in I$ und $\alpha \in \mathbb{R}$ fixiert. Dann hat f an der Stelle x_0 den Grenzwert α , falls **für jede Folge** $\{x_n\}$ aus I mit $x_n \neq x_0$ für alle $n \in \mathbb{N}$ und mit $x_n \rightarrow x_0$ für $n \rightarrow \infty$ gilt

$$\lim_{n \rightarrow \infty} f(x_n) = \alpha .$$

Notation:

$$\lim_{x \rightarrow x_0} f(x) = \alpha .$$

ii) Die Funktion f heißt stetig in einem Punkt $x_0 \in I$, falls

$$\lim_{x \rightarrow x_0} f(x) = f(x_0) .$$

iii) Die Funktion heißt stetig auf I , falls sie in jedem Punkt $x_0 \in I$ stetig ist.

Bemerkungen.

- i) „Wie immer“ gilt: Die Existenz eines Grenzwertes impliziert per definitionem, dass dieser **eindeutig bestimmt** ist.
- ii) Ist der Definitionsbereich im Gegensatz zur Generalvoraussetzung dieses Kapitels kein Intervall und f in einem Punkt nicht definiert, so kann trotzdem der Grenzwert existieren. Gleiches gilt für die Randpunkte.
- iii) Stetigkeit im Komplexen ist völlig analog definiert.

Reicht es in der Definition denn nicht, nur eine Folge $\{x_n\}$ zu betrachten?

Zur Beantwortung dieser Frage sei (vgl. Abbildung 4.6) $f: \mathbb{R} \rightarrow \mathbb{R}$,

$$f(x) = \begin{cases} 0, & \text{falls } x < 0, \\ 1, & \text{falls } x \geq 0. \end{cases}$$

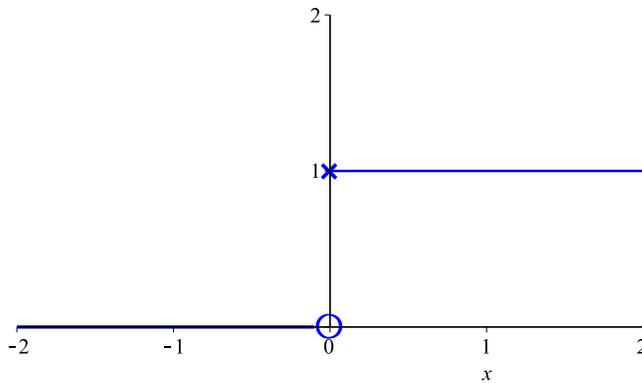


Abbildung 4.6: Eine Sprungfunktion ist nicht stetig.

Man betrachte nun zunächst etwa die gegen Null konvergente Folge $\{1/n\}$.

Es ist

$$\lim_{n \rightarrow \infty} f\left(\frac{1}{n}\right) = \lim_{n \rightarrow \infty} 1 = 1,$$

die Folge $\{f(1/n)\}$ konvergiert also gegen 1.

Als nächstes betrachte man beispielsweise die ebenfalls gegen Null konvergente Folge $\{-1/n\}$ mit

$$\lim_{n \rightarrow \infty} f\left(-\frac{1}{n}\right) = \lim_{n \rightarrow \infty} 0 = 0 .$$

Auch die Folge $\{f(-1/n)\}$ konvergiert, allerdings mit Grenzwert 0.

Obwohl beide betrachteten Folgen konvergieren, hat f im Nullpunkt keinen Grenzwert und ist dort insbesondere nicht stetig, was auch der intuitive Stetigkeitsbegriff vermuten lässt.

Anhand dieses Beispiels wird deutlich, dass es **nicht ausreicht, nur eine spezielle Folge zu betrachten**, um eine Funktion in einem Punkt auf Stetigkeit zu untersuchen.

Unstetigkeit trotz der Existenz eines Grenzwertes.

Die Existenz des Grenzwertes einer Funktion in einem Punkt $x_0 \in I$ impliziert nicht die Stetigkeit.

Dies belegt das einfache Beispiel (vgl. Abbildung 4.7)

$$f(x) = \begin{cases} 1, & \text{falls } x \neq 1, \\ 0, & \text{falls } x = 1. \end{cases}$$

In dem Beispiel gilt

$$\lim_{x \rightarrow 1} f(x) = 1 \neq 0 = f(1) .$$

Beispiele stetiger Funktionen.

i) Konstante Funktionen und die Identität $f(x) = x$ sind stetig.

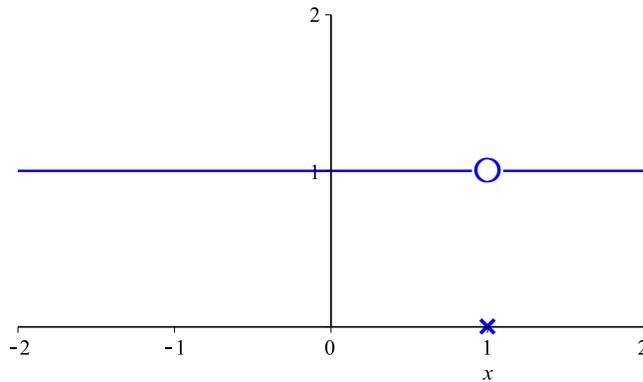


Abbildung 4.7: Obwohl der Grenzwert in 1 existiert, ist die Funktion dort nicht stetig.

ii) Es sei $f(x) = x^2$ und $x_0 \in \mathbb{R}$ sei fixiert. Weiter sei $\{x_n\}$ eine beliebige Folge mit $x_n \rightarrow x_0$ für $n \rightarrow \infty$.

Die Stetigkeit von f im Punkt x_0 ist gezeigt, falls

$$f(x_n) \xrightarrow{n \rightarrow \infty} f(x_0) .$$

Tatsächlich ist

$$\begin{aligned} |f(x_n) - f(x_0)| &= |x_n^2 - x_0^2| = |(x_n - x_0)(x_n + x_0)| \\ &\leq |x_n + x_0| |x_n - x_0| . \end{aligned}$$

Wegen (Dreiecksungleichung)

$$|x_n + x_0| \leq |x_n| + |x_0|$$

und da die Folge $\{x_n\}$ als konvergente Folge beschränkt ist (vgl. Teil I, Kapitel 7.1), existiert eine Konstante $K \in \mathbb{R}$ mit

$$|x_n + x_0| \leq K .$$

Es folgt die Behauptung

$$|f(x_n) - f(x_0)| \leq K|x_n - x_0| \xrightarrow{n \rightarrow \infty} 0 .$$

□

iii) Lipschitz-stetige Funktionen (und damit die Betragsfunktion) sind stetig (vgl. Übungskapitel 4.3).

Rechenregeln für stetige Funktionen.

Mithilfe des folgenden Satzes (Notation $(f + g)(x) := f(x) + g(x)$ etc.) verifiziert man unmittelbar die Stetigkeit einer Vielzahl von Funktionen (z.B. aller Polynome).

Satz 4.1. RECHENREGELN FÜR STETIGE FUNKTIONEN

- i) Es seien f, g zwei auf I stetige reellwertige Funktionen und $c \in \mathbb{R}$ sei fixiert.
- (a) Dann sind $cf, f + g$ und $f \cdot g$ stetige Funktionen.
- (b) In allen Punkten, in denen g nicht den Wert 0 annimmt, ist $\frac{f}{g}$ stetig.
- ii) Die Verkettung $g \circ f$ stetiger Funktionen f, g ist stetig, sofern sie definiert ist.

Beweis. Siehe Übungskapitel 4.3. □

Bemerkung. An dieser Stelle sei ein wesentlicher Zusammenhang zwischen gleichmäßiger Konvergenz (vgl. Teil I, Definition 8.2) und Stetigkeit erwähnt (Beweisskizze: Siehe Übungskapitel 4.3.)

- i) Eine **gleichmäßig konvergente** Folge $\{f_n\}$ stetiger Funktionen $f_n: I \rightarrow \mathbb{R}$ konvergiert gegen eine stetige Grenzfunktion $f: I \rightarrow \mathbb{R}$.

Diese Eigenschaft verdeutliche man sich anhand der Beispiele aus Teil I, Kapitel 8.

ii) Analog zu i) können gleichmäßig konvergente Funktionenreihen mit stetigen Funktionen als Glieder der Reihe betrachtet werden.

Insbesondere sind **exp, sin, cos, sinh, cosh** stetige Funktionen.

Nach Satz 4.1 sind folglich auch Funktionen wie $\sin^2(x)$, $\sin(x) \cosh(x)$ oder $e^{\sin(x)} = (\exp \circ \sin)(x)$ etc. stetig.

Äquivalente Grenzwertdefinition und gleichmäßige Stetigkeit.

In Abbildung 4.5 ist bereits ein anderer Zugang zur Definition des Grenzwertes einer Funktion angedeutet:

Satz 4.2. ÄQUIVALENTE GRENZWERTDEFINITION

Es seien $x_0 \in I$ und $\alpha \in \mathbb{R}$ fixiert.

Dann hat f genau dann an der Stelle x_0 den Grenzwert α , falls es zu jedem $\varepsilon > 0$ ein $\delta = \delta(\varepsilon) > 0$ gibt, sodass

$$x \in I, \quad x \neq x_0, \quad |x - x_0| < \delta \quad \Rightarrow \quad |f(x) - \alpha| < \varepsilon.$$

Beweis. Siehe Übungskapitel 4.3. □

Mithilfe der Formulierung aus Satz 4.2 kann eine Funktion auf „gleichmäßige Konvergenzgeschwindigkeit“ in allen Punkten untersucht werden:

Definition 4.2. GLEICHMÄSSIGE STETIGKEIT

Eine Abbildung $f: I \rightarrow \mathbb{R}$ heißt **gleichmäßig stetig**, falls es zu jedem $\varepsilon > 0$ ein $\delta > 0$ gibt, sodass für alle $x, y \in I$ mit $|x - y| < \delta$ gilt:

$$|f(x) - f(y)| < \varepsilon.$$

Demnach sind insbesondere Lipschitz-stetige Funktionen gleichmäßig stetig (siehe Übungskapitel 4.3).

4.2 Zwei Sätze über stetige Funktionen (Stetigkeit und Kompaktheit; Zwischenwertsatz; Stetigkeit der Umkehrfunktion)

In Teil I, Übungskapitel 5, ist bereits angedeutet, dass die Existenz von Extremwerten einer Funktion im Allgemeinen vom Definitionsbereich abhängt. Deshalb ist der folgende Satz von fundamentaler Bedeutung.

Satz 4.3. STETIGKEIT UND KOMPAKTHEIT

Es sei $I = [a, b]$ ein *kompaktes* Intervall (d.h. abgeschlossen und beschränkt) und $f: I \rightarrow \mathbb{R}$ sei stetig auf I .

Dann nimmt die Funktion f in I ihr *absolutes Maximum* und ihr *absolutes Minimum* an.

Das bedeutet: Es existieren Punkte x_1 (genannt *Minimalstelle* oder *Minimierer*), $x_2 \in I$ (genannt *Maximalstelle* oder *Maximierer*) mit

$$f(x_1) \leq f(x) \leq f(x_2) \quad \text{für alle } x \in I.$$

Beweis. Siehe Übungskapitel 4.3. □

Schließlich wird der anschauliche Stetigkeitsbegriff im so genannten *Zwischenwertsatz* präzisiert (vgl. Abbildungen 4.8, 4.9, 4.10), der mithilfe eines geeigneten Intervallschachtelungsprinzips gezeigt werden kann.

Satz 4.4. ZWISCHENWERTSATZ

Die Funktion $f: \mathbb{R} \supset [a, b] \rightarrow \mathbb{R}$ sei stetig und o.E. sei $f(a) < f(b)$.

Dann *nimmt f jeden Wert zwischen $f(a)$ und $f(b)$ an*, was bedeutet:

Aus $f(a) < \xi < f(b)$ folgt die Existenz von (mindestens) einem Punkt $x_0 \in (a, b)$ mit $f(x_0) = \xi$.

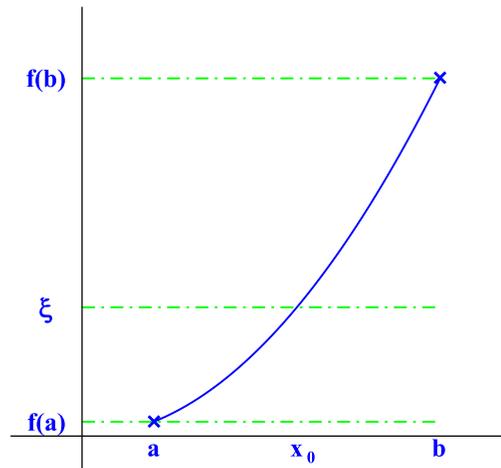


Abbildung 4.8: Veranschaulichung des Zwischenwertsatz.

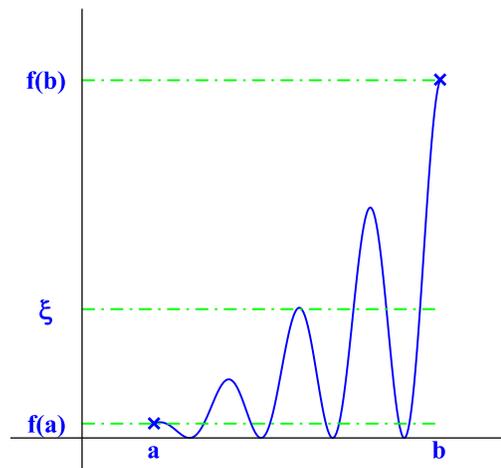


Abbildung 4.9: Der Punkt x_0 im Zwischenwertsatz muss nicht eindeutig bestimmt sein.

Der Zwischenwertsatz impliziert ein Korollar, welches beispielsweise die Stetigkeit der Logarithmusfunktion (und vieler bereits angesprochener Umkehrfunktionen) belegt.

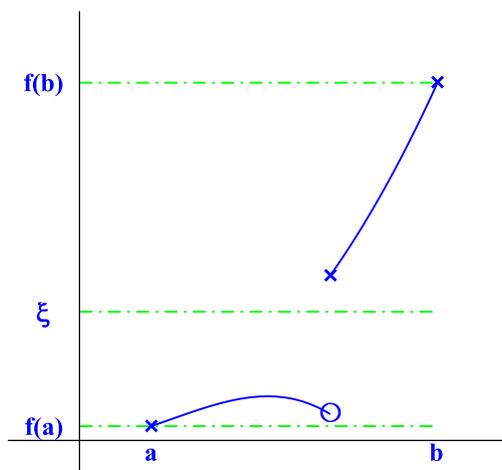


Abbildung 4.10: Ohne Stetigkeit kein Zwischenwertsatz.

Korollar 4.1. STETIGKEIT UND UMKEHRFUNKTION

Es sei $f: I \rightarrow \mathbb{R}$ stetig und streng monoton.

Dann besitzt f eine stetige, streng monotone Umkehrfunktion

$$f^{-1} : f(I) \rightarrow I .$$

4.3 Übungsaufgaben zu Kapitel 4

Aufgabe 1. Zeigen Sie:

- i) Die Betragsfunktion ist Lipschitz-stetig.
- ii) Lipschitz-stetige Funktionen (und damit insbesondere die Betragsfunktion) sind stetig.
- iii) Lipschitz-stetige Funktionen sind gleichmäßig stetig.
- iv) Die Funktion $f: \mathbb{R} \rightarrow \mathbb{R}$, $f(x) = \sqrt{|x|}$, ist im Nullpunkt stetig aber nicht Lipschitz-stetig.

Aufgabe 2.* Es seien $f, g: \mathbb{R} \rightarrow \mathbb{R}$ definiert als

$$f(x) = \begin{cases} \cos\left(\frac{1}{x}\right) & \text{für } x \neq 0 \\ 0 & \text{für } x = 0 \end{cases}, \quad g(x) = \begin{cases} x \cos\left(\frac{1}{x}\right) & \text{für } x \neq 0 \\ 0 & \text{für } x = 0 \end{cases}.$$

Sind die Funktionen im Punkt $x = 0$ stetig?

Aufgabe 3. Ist die Funktion $f: \mathbb{R} \rightarrow \mathbb{R}$ stetig, wobei

$$f(x) = \begin{cases} x \sin(\ln(|x|)) & \text{für } x \neq 0, \\ 0 & \text{für } x = 0. \end{cases}$$

Aufgabe 4. Es sei f eine Funktion $I \rightarrow \mathbb{R}$, wobei $I \subset \mathbb{R}$ ein (verallgemeinertes) Intervall bezeichne.

Dann lassen sich auch die **einseitigen Grenzwerte** definieren:

$$\lim_{x \rightarrow x_0^-} f(x) = y_0 \quad :\Leftrightarrow \quad \text{für alle } \{x_n\}, \quad x_n \in I, \quad x_n < x_0 \text{ gilt:}$$

$$\lim_{n \rightarrow \infty} x_n = x_0 \quad \Rightarrow \quad \lim_{n \rightarrow \infty} f(x_n) = y_0 ;$$

$$\lim_{x \rightarrow x_0^+} f(x) = y_0 \quad :\Leftrightarrow \quad \text{für alle } \{x_n\}, \quad x_n \in I, \quad x_n > x_0 \text{ gilt:}$$

$$\lim_{n \rightarrow \infty} x_n = x_0 \quad \Rightarrow \quad \lim_{n \rightarrow \infty} f(x_n) = y_0 .$$

Weierhin definiere man für eine reelle Zahlenfolge $\{x_n\}$ (analog der Fall “ $-\infty$ ”):

$$\lim_{n \rightarrow \infty} x_n = \infty$$

$:\Leftrightarrow$ Zu jedem $M > 0$ existiert ein $N \in \mathbb{N}$ mit $M < x_n$ für alle $n > N$,

$$\lim_{x \rightarrow \infty} f(x) = y_0$$

$:\Leftrightarrow$ für alle $\{x_n\}, \quad x_n \in I, \quad \text{gilt: } \lim_{n \rightarrow \infty} x_n = \infty \Rightarrow \lim_{n \rightarrow \infty} f(x_n) = y_0 .$

Existieren die folgenden Grenzwerte:

$$\lim_{x \rightarrow x_0^+} \frac{x^2 + 1}{x^3}, \quad \lim_{x \rightarrow x_0^-} \frac{x^2 + 1}{x^3}, \quad \lim_{x \rightarrow \infty} \frac{x^2 + 1}{x^3},$$

$$\lim_{x \rightarrow x_0^+} \frac{x}{|x|}, \quad \lim_{x \rightarrow x_0^-} \frac{x}{|x|},$$

$$\lim_{x \rightarrow x_0^+} \frac{x \sin(x)}{|x|}, \quad \lim_{x \rightarrow x_0^-} \frac{x \sin(x)}{|x|}, \quad \lim_{x \rightarrow \infty} \frac{x \sin(x)}{|x|} ?$$

Aufgabe 5. Es seien $f, g: \mathbb{R} \rightarrow \mathbb{R}$ stetige Funktionen. Zeigen Sie, dass dann auch die Funktionen $f + g$, $f \cdot g$ und $f \circ g$ stetig sind.

Aufgabe 6. Zeigen Sie Satz 4.2.

Aufgabe 7.* Zeigen Sie Satz 4.3.

Aufgabe 8.

- i) * Zeigen Sie: Eine gleichmäßig konvergente Folge stetiger Funktionen hat eine stetige Grenzfunktion.
 - ii) Konvergiert die Funktionenfolge $\{f_n\}$, $f_n: [-1, 1] \rightarrow \mathbb{R}$, $f_n(x) = (1 - x^2)^{n/2}$, gegen eine stetige Grenzfunktion?
-

Aufgabe 9. Zeigen Sie, dass die Funktion $f: \mathbb{R}^+ \rightarrow \mathbb{R}$, $f(x) = 1/x$ nicht gleichmäßig stetig ist.

Aufgabe 10. Es sei $f: [0, 1] \rightarrow [0, 1]$ eine stetige Funktion.

- i) Zeigen Sie, dass ein $\xi \in [0, 1]$ existiert mit $f(\xi) = \xi$ – ein solches ξ nennt man Fixpunkt von f .

Hinweis. Verwenden Sie den Zwischenwertsatz für die Funktion $g(x) = f(x) - x$.

- ii) Zeigen Sie weiter, dass die Gleichung

$$\exp(x) = 3x$$

(mindestens) eine Lösung $x \in [0, 1]$ hat.

Lösungshinweise zu den Übungsaufgaben.

Aufgabe 2. Die Funktion f ist nicht stetig. Dazu finde man eine gegen Null konvergente Folge $\{x_k\}$ mit $f(x_k) = 1$ für alle $k \in \mathbb{N}$.

Die Funktion g ist stetig. Dazu beachte man für $x \neq 0$

$$|g(x) - 0| = |x \cos(1/x)| \leq |x| .$$

Aufgabe 7. Es sei

$$s := \sup\{f(x) : a \leq x \leq b\} .$$

Dieses s existiert, wenn der „Wert $+\infty$ “ zugelassen wird.

Nach der Definition von s existiert eine Folge $\{x_k\}$, $a \leq x_k \leq b$ für alle $k \in \mathbb{N}$, mit

$$f(x_k) \xrightarrow{k \rightarrow \infty} s . \quad (*)$$

Der Satz von Bolzano-Weierstraß (Teil I, Kapitel 7.1) liefert weiter eine konvergente Teilfolge $\{x_{k_j}\}$ und ein $\hat{x} \in [a, b]$ mit

$$x_{k_j} \xrightarrow{j \rightarrow \infty} \hat{x} , \quad (**)$$

wobei die Konvergenz (*) beim Übergang zu der Teilfolge erhalten bleibt.

Andererseits ist f nach Voraussetzung stetig und aus (**) ergibt sich

$$f(x_{k_j}) \xrightarrow{j \rightarrow \infty} f(\hat{x}) .$$

Es ist gezeigt

$$s = f(\hat{x}) = \max_{a \leq x \leq b} f(x) ,$$

in der Tat ist s also endlich und f nimmt sein Maximum in \hat{x} an. Die Annahme des Minimums folgt ebenso. \square

Aufgabe 8. *i)* Es seien $x_0 \in I$ und $\varepsilon > 0$ fixiert. Nach der Dreiecksungleichung gilt für alle $x \in I$ und für alle $n \in \mathbb{N}$

$$|f(x) - f(x_0)| \leq |f(x) - f_n(x)| + |f_n(x) - f_n(x_0)| + |f_n(x_0) - f(x_0)|.$$

Gleichmäßige Konvergenz der Funktionenfolge bedeutet

$$\lim_{n \rightarrow \infty} \|f_n - f\|_\infty = \lim_{n \rightarrow \infty} \sup_{x \in I} |f_n(x) - f(x)| = 0.$$

Ist also $\tilde{\varepsilon} = \varepsilon/3$ gegeben, so existiert ein $N = N(\tilde{\varepsilon}) \in \mathbb{N}$ mit

$$\sup_{x \in I} |f_n(x) - f(x)| < \tilde{\varepsilon} = \frac{\varepsilon}{3} \quad \text{für alle } n \geq N.$$

Ein solches n sei nun **fixiert**. Dann gilt für alle $x \in I$

$$|f(x) - f_n(x)| < \frac{\varepsilon}{3}, \quad \text{und} \quad |f_n(x_0) - f(x_0)| < \frac{\varepsilon}{3}.$$

Zudem ist die Funktion f_n nach Voraussetzung stetig in x_0 , d.h. zu $\tilde{\varepsilon}$ existiert ein $\delta > 0$, sodass für alle $x \in I$ mit $|x - x_0| < \delta$

$$|f_n(x) - f_n(x_0)| < \tilde{\varepsilon} = \frac{\varepsilon}{3}.$$

Insgesamt folgt für alle $x \in I$ mit $|x - x_0| < \delta$:

$$|f(x) - f(x_0)| < \varepsilon$$

und die Stetigkeit der Grenzfunktion ist bewiesen. □

Kapitel 5

Differentialrechnung in einer Veränderlichen

5.1 Grundlagen (Differenzenquotient; Ableitung; Differenzierbarkeit und Stetigkeit; Produkt- Quotientenregel; Kettenregel; Ableitung der Umkehrfunktion; höhere Ableitungen)

Die Differential- und Integralrechnung geht gleichzeitig auf [Newton](#) und [Leibniz](#) zurück und ist die Grundlage aller modernen Naturwissenschaften.

Der intuitive Begriff der [Differenzierbarkeit](#) einer Funktion ist der folgende:

Der Graph einer differenzierbaren Funktion hat keinen „Knick“.

Dementsprechend sollte die Betragsfunktion (vgl. Abbildung 4.2) zwar stetig aber nicht differenzierbar im Nullpunkt sein, was auch tatsächlich der Fall ist.

Für die Tauglichkeit des intuitiven Begriffs der Differenzierbarkeit gelten aber ähnliche Einschränkungen wie im letzten Kapitel zur Stetigkeit.

Geometrische Interpretation der Ableitung.

Bei der geometrischen Interpretation der Ableitung wird, wie in Abbildung

5.1 angedeutet, in einem festen Punkt x_0 die **Tangente** an den Graphen der Funktion als **affin lineare Approximation** der Funktion betrachtet.

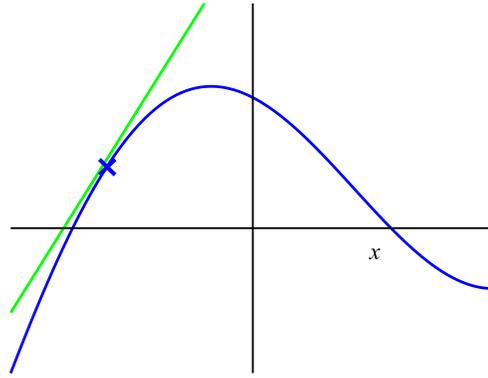


Abbildung 5.1: Zur geometrischen Interpretation der Ableitung.

Mit anderen Worten: Der Graph der Funktion wird in der Nähe des Punktes $(x_0, f(x_0))$ „möglichst gut“ mit einer Geraden angenähert.

Kinematische Interpretation der Ableitung.

Bei der kinematischen Interpretation der Ableitung betrachtet man etwa die Bewegung eines Massenpunktes im dreidimensionalen Raum in Abhängigkeit von der Zeit.

Diese Bewegung wird durch eine **Kurve** im \mathbb{R}^3 beschrieben. Das ist eine Abbildung $\gamma: \mathbb{R} \supset I \rightarrow \mathbb{R}^3$, wobei I ein Zeitintervall bezeichnet.

Die Ableitung entspricht in diesem Fall dem **Geschwindigkeitsvektor** des Massenpunktes zu einem gegebenen Zeitpunkt t_0 (vgl. Abbildung 5.2).

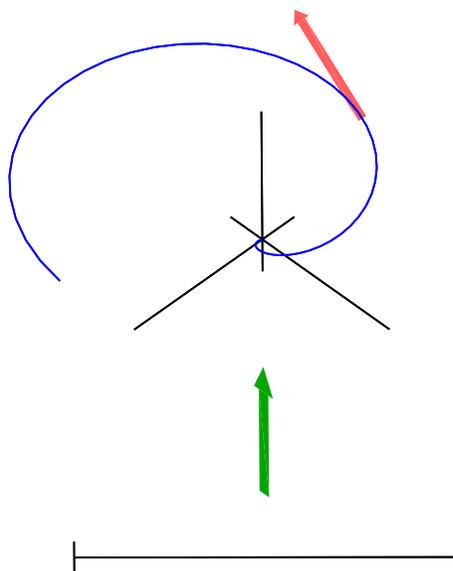


Abbildung 5.2: Zur kinematischen Interpretation der Ableitung.

Differenzenquotient und Ableitung.

Im Folgenden sei $I \subset \mathbb{R}$ stets ein offenes (evtl. verallgemeinertes) Intervall und $f: I \rightarrow \mathbb{R}$ eine Funktion von I nach \mathbb{R} .

Definition 5.1. ABLEITUNG

Es sei f wie oben und $x_0 \in I$ fixiert.

i) Ist $x \in I$, $x \neq x_0$, so heißt der Quotient

$$\frac{f(x) - f(x_0)}{x - x_0}$$

Differenzenquotient von f bzgl. der Punkte x und x_0 .

ii) Die Funktion f heißt *differenzierbar* im Punkt $x_0 \in I$, falls der Grenzwert

$$\lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{h \rightarrow 0} \frac{f(x_0 + h) - f(x_0)}{h}$$

existiert.

Der Grenzwert heißt die *Ableitung* oder der *Differentialquotient* von f bei x_0 .

Notation:

$$f'(x_0) \quad \text{oder} \quad \frac{df}{dx}(x_0) \quad \text{oder} \quad \frac{df}{dx}|_{x=x_0} .$$

iii) Die Funktion heißt differenzierbar auf I , wenn sie in jedem Punkt $x_0 \in I$ differenzierbar ist.

Dann kann eine weitere Funktion (ebenfalls als Ableitung von f bezeichnet) betrachtet werden, die jedem Punkt $x \in I$ die Ableitung von f an der Stelle x zuordnet:

$$f' : I \rightarrow \mathbb{R}, \quad I \ni x \mapsto f'(x) .$$

Bemerkungen.

- i) Die obige Definition **kann nicht auf den Fall mehrerer Veränderlicher übertragen werden**, da man „nicht durch einen Vektor teilen“ kann.
- ii) Wie im Übungskapitel 4.3 definiert, können auch hier die **einseitigen Grenzwerte**

$$f'(x_0^+) := \lim_{x \rightarrow x_0^+} \frac{f(x) - f(x_0)}{x - x_0} ,$$

$$f'(x_0^-) := \lim_{x \rightarrow x_0^-} \frac{f(x) - f(x_0)}{x - x_0} ,$$

genannt die **rechtsseitige Ableitung** und die **linksseitige Ableitung**, betrachtet werden.

Diese sind insbesondere am Rande des Definitionsbereiches von Bedeutung, wo die Ableitung selbst nicht definiert ist.

Zu den obigen Interpretationen.

- i) Geometrisch entspricht der Differenzenquotient von f bzgl. x_1 und x_0 der Steigung der **Sekante** s durch die Punkte $(x_0, f(x_0))$ und $(x_1, f(x_1))$ (vgl. Abbildung 5.3),

$$s(x) = f(x_0) + \frac{f(x_1) - f(x_0)}{x_1 - x_0}(x - x_0).$$

In der Tat ist $s(x_0) = f(x_0)$ und $s(x_1) = f(x_1)$ und es handelt sich um eine Gerade durch die gegebenen Punkte mit der Steigung $(f(x_1) - f(x_0))/(x_1 - x_0)$.

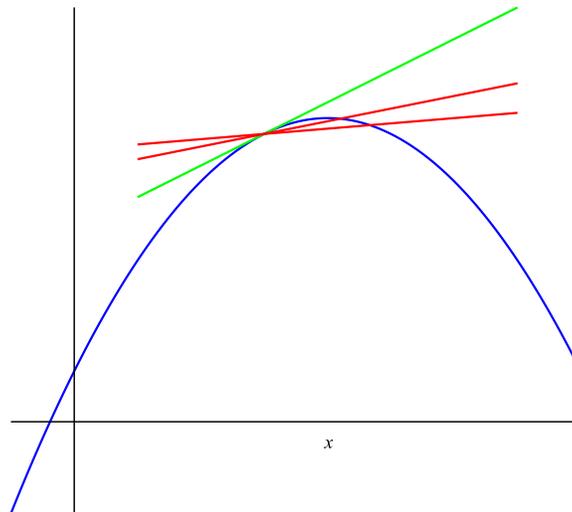


Abbildung 5.3: Die Sekantensteigungen gehen in die Steigung der Tangente über.

Im Grenzwert $x_1 \rightarrow x_0$ erhält man die Steigung der Tangente an $(x_0, f(x_0))$.

- ii) Kinematisch entspricht der Differenzenquotient der **Durchschnittsgeschwindigkeit** im Zeitraum zwischen x_0 und x_1 .

Der Grenzwert $x \rightarrow x_0$ gibt die **Momentangeschwindigkeit** zum Zeitpunkt x_0 an.

Beispiele differenzierbarer Funktionen.

- i) Im Fall einer konstanten Funktion verschwindet der Differenzenquotient identisch.

Demnach ist die Funktion überall differenzierbar mit der Ableitung $f'(x) = 0$ für alle $x \in I$.¹

- ii) Es sei $f(x) = x$ für alle $x \in \mathbb{R}$.

Der Differenzenquotient für ein fixiertes x_0 lautet

$$\frac{x - x_0}{x - x_0} \equiv 1 .$$

Die Funktion ist überall differenzierbar mit der Ableitung 1.

- iii) Es sei $f(x) = x^2$ für alle $x \in \mathbb{R}$ und $x_0 \in \mathbb{R}$ sei fixiert.

Dann gilt

$$\frac{f(x) - f(x_0)}{x - x_0} = \frac{x^2 - x_0^2}{x - x_0} = \frac{(x - x_0)(x + x_0)}{x - x_0} = x + x_0 .$$

Um zum Grenzwert $x \rightarrow x_0$ und damit zur Ableitung überzugehen, ist in diesen Differenzenquotienten für x eine beliebige Folge $\{x_n\}$ mit Grenzwert x_0 einzusetzen (vgl. Grenzwertdefinition einer Funktion, Definition 4.1).

Da die Folge $\{x_n\}$ nach Voraussetzung aber gegen x_0 konvergiert, konvergiert die rechte Seite der obigen Gleichheit, d.h. auch der Differenzenquotient, gegen $2x_0$.

Zusammenfassend ist gezeigt: Die Funktion $f(x) = x^2$ ist in jedem $x_0 \in \mathbb{R}$ differenzierbar mit der Ableitung $f'(x_0) = 2x_0$.

¹Man beachte, dass die Frage nach einer möglichen Umkehrung der Aussage (Impliziert eine überall verschwindende Ableitung die Konstanz einer Funktion?) an dieser Stelle noch nicht beantwortet werden kann.

iv) Es sei $f(x) = \sqrt{x}$ für $x > 0$ und $x_0 > 0$ sei fixiert.

In der zweiten Schreibweise des Differenzenquotienten ($x = x_0 + h$) gilt für $|h|$ hinreichend klein

$$\begin{aligned} \frac{f(x_0 + h) - f(x_0)}{h} &= \frac{1}{h}(\sqrt{x_0 + h} - \sqrt{x_0}) \\ &= \frac{1}{h} \frac{(\sqrt{x_0 + h} - \sqrt{x_0})(\sqrt{x_0 + h} + \sqrt{x_0})}{\sqrt{x_0 + h} + \sqrt{x_0}} \\ &= \frac{1}{\sqrt{x_0 + h} + \sqrt{x_0}} \\ &\xrightarrow{h \rightarrow 0} \frac{1}{2\sqrt{x_0}}. \end{aligned}$$

Die Funktion $f(x) = \sqrt{x}$ ist demnach in allen Punkten $x_0 > 0$ differenzierbar mit Ableitung $f'(x_0) = 1/(2\sqrt{x_0})$.

Im Punkt $x_0 = 0$ ist die Wurzelfunktion aber nicht differenzierbar:

Die Tangentensteigungen werden für beliebig kleine x_0 beliebig groß, d.h. auch der einseitige Grenzwert des Differenzenquotienten existiert im Nullpunkt nicht.

Stetigkeit und Differenzierbarkeit.

Als typisches Beispiel sei $f(x) = |x|$ im Punkt $x_0 = 0$ betrachtet.

Für die Nullfolge $\{x_n\} = \{1/n\}$ gilt

$$\frac{f(x_n) - f(x_0)}{x_n - x_0} = \frac{(1/n) - 0}{(1/n) - 0} = 1.$$

Für die Nullfolge $\{\tilde{x}_n\} = \{-1/n\}$ gilt hingegen

$$\frac{f(\tilde{x}_n) - f(x_0)}{\tilde{x}_n - x_0} = \frac{(1/n) - 0}{(-1/n) - 0} = -1.$$

Der Grenzwert des Differenzenquotienten kann also im Nullpunkt nicht existieren und wie erwartet ist die **Betragsfunktion im Nullpunkt nicht differenzierbar**.

Das Beispiel der stetigen Betragsfunktion belegt: **Aus der Stetigkeit einer Funktion folgt nicht die Differenzierbarkeit**.

Es gilt aber die Umkehrung im Sinne von

Satz 5.1. DIFFERENZIERBARKEIT IMPLIZIERT STETIGKEIT

Ist die Funktion $f: I \rightarrow \mathbb{R}$ im Punkt $x_0 \in I$ differenzierbar, so ist f im Punkt x_0 stetig.

Beweis Siehe Übungskapitel 5.4. □

Rechenregeln für differenzierbare Funktionen.

Wie bei der Betrachtung stetiger Funktionen sind im nächsten Schritt Summe und Produkt differenzierbarer Funktionen zu untersuchen. Dabei werden explizite Vorschriften zur Berechnung der Ableitung mitgeliefert.

Satz 5.2. RECHENREGELN FÜR DIFF' BARE FUNKTIONEN

Es seien f, g zwei auf I differenzierbare reellwertige Funktionen und $c \in \mathbb{R}$ sei fixiert.

i) Dann sind auch $cf, f + g$ und fg auf I differenzierbar mit

(a) $(cf)' = cf'$;

(b) $(f + g)' = f' + g'$;

*(c) $(fg)' = f'g + fg'$ (*Produktregel*).*

ii) Ist $g(x_0) \neq 0$ für $x_0 \in I$, so ist auch f/g in x_0 differenzierbar mit (Quotientenregel)

$$\left(\frac{f}{g}\right)'(x_0) = \frac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{g^2(x_0)}.$$

Beweisidee. Exemplarisch wird hier nur gezeigt, wie die zwei Summanden bei der Ableitung eines Produktes entstehen:

Es seien $x, x_0 \in I$. Dann gilt

$$\begin{aligned} & \frac{f(x)g(x) - f(x_0)g(x_0)}{x - x_0} \\ &= \frac{f(x)g(x) - f(x_0)g(x) + f(x_0)g(x) - f(x_0)g(x_0)}{x - x_0} \\ &= \frac{f(x) - f(x_0)}{x - x_0}g(x) + f(x_0)\frac{g(x) - g(x_0)}{x - x_0} \\ &\xrightarrow{x \rightarrow x_0} f'(x_0)g(x_0) + f(x_0)g'(x_0). \end{aligned}$$

Das ist aber genau die Behauptung. □

Bemerkung. Es gilt sogar:

i) Für alle $n \in \mathbb{N}$ sei $f_n: I \rightarrow \mathbb{R}$ auf I differenzierbar.

Sind die Funktionenfolgen $\{f_n\}$ und $\{f'_n\}$ auf I **gleichmäßig konvergent** und bezeichnet f den Grenzwert der Folge $\{f_n\}$, so ist auch f differenzierbar mit

$$f'_n \rightrightarrows f' \quad \text{für } n \rightarrow \infty.$$

ii) Dies gilt analog für Funktionenreihen (d.h. bei gleichmäßiger Konvergenz der Reihe und der Reihe der Ableitungen):

$$\left[\sum_{k=0}^{\infty} f_k(x) \right]' = \sum_{k=0}^{\infty} f'_k(x).$$

Mit anderen Worten: Unter diesen Voraussetzungen „darf die Ableitung in die Reihe hereingezogen werden“.

Weitere Beispiele differenzierbarer Funktionen.

Satz 5.2 und die anschließenden Bemerkungen liefern beispielsweise:

i) Für alle $k \in \mathbb{Z}$ ($x \neq 0$ für $k < 0$) gilt (vgl. Übungskapitel 5.4)

$$\frac{d}{dx} x^k = kx^{k-1}.$$

ii) Im Falle der Exponentialfunktion berechnet man für alle $x \in \mathbb{R}$:

$$\begin{aligned} \frac{d}{dx} \exp(x) &= \left[\sum_{k=0}^{\infty} \frac{1}{k!} x^k \right]' = \sum_{k=0}^{\infty} \frac{1}{k!} [x^k]' = \sum_{k=1}^{\infty} \frac{1}{k!} kx^{k-1} \\ &= \sum_{k=1}^{\infty} \frac{1}{(k-1)!} x^{k-1} = \sum_{j=0}^{\infty} \frac{1}{j!} x^j = \exp(x). \end{aligned}$$

iii) In Tabelle 5.1 sind die Ableitungen einiger bekannter Funktionen festgehalten.

iv) Es sei $f(x) = \tan(x)$, $x \in (-\pi/2, \pi/2)$. Dann ist nach der Quotientenregel

$$f'(x) = \frac{\cos(x) \cos(x) - \sin(x)(-\sin(x))}{\cos^2(x)} = \frac{1}{\cos^2(x)}.$$

Ableitung der Verkettung von Funktionen.

Auch bei der Verkettung differenzierbarer Funktionen kann man auf Differenzierbarkeit schließen und die Ableitung berechnen, ohne explizit

$f(x)$	$f'(x)$	Def. Bereich	Bem.
x^k	kx^{k-1}	$\mathbb{R}, x \neq 0$ für $k < 0$	$k \in \mathbb{Z}$
\sqrt{x}	$\frac{1}{2} \frac{1}{\sqrt{x}}$	\mathbb{R}^+	vgl. x^α
x^α	$\alpha x^{\alpha-1}$	\mathbb{R}^+	$\alpha \in \mathbb{R}$
e^x	e^x	\mathbb{R}	
$\sin(x)$	$\cos(x)$	\mathbb{R}	
$\cos(x)$	$-\sin(x)$	\mathbb{R}	
$\sinh(x)$	$\cosh(x)$	\mathbb{R}	
$\cosh(x)$	$\sinh(x)$	\mathbb{R}	

Tabelle 5.1: Die Ableitungen einiger bekannter Funktionen.

die Definition nachzuvollziehen.²

Satz 5.3. KETTENREGEL

Die Verkettung $g \circ f$ zweier reellwertiger differenzierbarer Funktionen f, g sei wohl definiert (d.h. das Bild von f liege im Definitionsbereich von g).

Dann ist auch $g \circ f$ differenzierbar und es gilt

$$(g \circ f)'(x) = g'(f(x))f'(x) .$$

Sprechweise. Man formuliert die Kettenregel oft als „äußere Ableitung mal innere Ableitung“.

²Im so genannten [Leibnizkalkül](#) schreibt man **formal** für die Ableitung der Funktion $z(y(x))$:

$$\frac{dz}{dx} = \frac{dz}{dy} \frac{dy}{dx} .$$

Diese Notation wird jedoch in der Regel im Folgenden nicht verwendet.

Beispiele.

- i) Für alle $x \in \mathbb{R}$ sei $h(x) = \sqrt{1+x^2}$, d.h. $h = g \circ f$ mit $g(y) = \sqrt{y}$ und $f(x) = 1+x^2$.

Aus der Kettenregel folgt

$$\begin{aligned} h'(x) &= g'(f(x))f'(x) \\ &= \frac{1}{2\sqrt{1+x^2}} \cdot 2x \\ &= \frac{x}{\sqrt{1+x^2}}. \end{aligned}$$

- ii) Für fixiertes $a > 0$ und für alle $x \in \mathbb{R}$ ist die Funktion $f(x) = a^x$ definiert als (vgl. Teil I, Kapitel 8.2)

$$f(x) = \exp(x \ln(a)).$$

Mithilfe der Kettenregel folgt

$$f'(x) = \exp(x \ln(a)) \ln(a) = a^x \ln(a).$$

Analog wird die Funktion x^x abgeleitet (siehe Übungskapitel 5.4).

Ableitung der Umkehrfunktion.

Der folgende Satz zur Ableitung der Umkehrfunktion folgt unmittelbar aus der Kettenregel durch Differentiation der Identität $(f^{-1} \circ f)(x) = x$:

$$(f^{-1})'(f(x))f'(x) = 1.$$

Satz 5.4. ABLEITUNG DER UMKEHRFUNKTION

Die Funktion $f: I \rightarrow f(I)$ ($y = f(x)$) sei bijektiv und differenzierbar mit $f'(x) \neq 0$ für alle $x \in I$.

Dann ist für alle $y \in f(I)$ die Ableitung der Umkehrfunktion f^{-1} gegeben durch

$$(f^{-1})'(y) = \frac{1}{f'(x)} = \frac{1}{f'(f^{-1}(y))}.$$

Bemerkung. Die Funktion $f(x) = x^3$ ist zwar bijektiv, die Ableitung verschwindet aber im Ursprung. Deshalb ist die Umkehrfunktion in Null nicht differenzierbar.

Beispiele

i) Für $x > 0$ sei die Funktion $f(x) = x^2$ betrachtet. Die Umkehrfunktion f^{-1} ist die Wurzelfunktion und Satz 5.4 liefert wegen $f'(x) = 2x$

$$\frac{d}{dy} \sqrt{y} = (f^{-1})'(y) = \frac{1}{2f^{-1}(y)} = \frac{1}{2\sqrt{y}}.$$

Das die bekannte Ableitungsregel für die Wurzelfunktion, wobei üblicherweise die unabhängige Variable wieder mit x bezeichnet wird.

ii) Auf \mathbb{R}^+ betrachte man die Funktion $\ln(y) = f^{-1}(y)$ als Umkehrfunktion der Exponentialfunktion $e^x = f(x)$.

Nach Satz 5.4 gilt ($y = e^x$, $x = \ln(y)$)

$$\ln'(y) = \frac{1}{e^x} = \frac{1}{e^{\ln(y)}} = \frac{1}{y}.$$

iii) Die Ableitungen der trigonometrischen Umkehrfunktionen sind ohne Beweis in Tabelle 5.2 festgehalten.

$f(x)$	$f'(x)$	Def. Bereich
$\arcsin(x)$	$\frac{1}{\sqrt{1-x^2}}$	$-1 < x < 1$
$\arccos(x)$	$-\frac{1}{\sqrt{1-x^2}}$	$-1 < x < 1$
$\arctan(x)$	$\frac{1}{1+x^2}$	\mathbb{R}

Tabelle 5.2: Die Ableitungen der trigonometrischen Umkehrfunktionen.

Höhere Ableitungen.

In Definition 5.1, *iii*), ist bereits festgehalten, dass die Ableitung einer auf I differenzierbaren Funktion f **selbst wieder eine Funktion** $f': I \rightarrow \mathbb{R}$ ist.

Somit kann auch die Funktion f' auf Differenzierbarkeit untersucht werden.

Sagt die erste Ableitung etwas über die Steigung des Graphen einer Funktion aus, so gibt die zweite Ableitung Informationen über die **Krümmung**, wie im nächsten Abschnitt noch deutlich wird.

Definition 5.2. HÖHERE ABLEITUNGEN

- i*) Es sei $f: I \rightarrow \mathbb{R}$ differenzierbar auf I mit der Ableitung $f': I \rightarrow \mathbb{R}$.
Ist auch f' differenzierbar auf I , so heißt die Ableitung von f' die **zweite Ableitung** von f .

Notation:

$$f''(x) \quad \text{oder} \quad f^{(2)}(x) \quad \text{oder} \quad \frac{d^2}{dx^2}f(x).$$

- ii*) Analog wird die dritte, vierte etc. Ableitung von f definiert.
- iii*) Eine Funktion $f: I \rightarrow \mathbb{R}$ heißt von der Klasse $C^k(I)$, falls f k -mal differenzierbar ist und die Ableitung $f^{(k)}$ eine stetige Funktion ist.
-

Bemerkungen.

- i) $C^0(I)$ bezeichnet die Klasse der auf I stetigen Funktionen.
- ii) Nicht alle auf I differenzierbaren Funktionen sind von der Klasse C^1 . Dies belegen Beispiele wie $x^2 \sin(1/x)$, die zwar überall differenzierbar sind, deren Ableitungen aber keine stetigen Funktionen sind (vgl. Übungskapitel 5.4).

Beispiel. Es sei wieder $f(x) = \sqrt{1+x^2}$ für alle $x \in \mathbb{R}$. Dann gilt

$$\begin{aligned} f''(x) &= \frac{d}{dx} \frac{x}{\sqrt{1+x^2}} = \frac{d}{dx} (1+x^2)^{-\frac{1}{2}} x \\ &= \left(-\frac{1}{2}\right) (1+x^2)^{-\frac{3}{2}} 2xx + (1+x^2)^{-\frac{1}{2}} \\ &= \frac{1}{(1+x^2)^{\frac{3}{2}}}. \end{aligned}$$

5.2 Lokale und globale Extrema, Mittelwertsatz (notwendige und hinreichende Bedingung für lokale Extrema; Suche nach globalen Extrema; Satz von Rolle; Mittelwertsatz; Ableitung und Monotonie; konvexe und konkave Funktionen; Regeln von l'Hospital)

Im Folgenden sei $I = [a, b]$ bzw. $I = (a, b)$. Die Fälle $I = (-\infty, \infty)$ etc. werden analog behandelt.

Bei der Beschreibung von Naturvorgängen gehört es zu den wesentlichen Aufgaben, die Maxima oder Minima bestimmter Größen zu finden (z.B. den Flächeninhalt, die Energie bei der Deformation eines Festkörpers oder die sogenannte Wirkung nach dem [Prinzip der kleinsten Wirkung](#) ...).

Hier liefert die Differentialrechnung die wesentlichen Hilfsmittel.

Notwendige Bedingung für lokale Extrema.

In Satz 4.3 wird untersucht, ob eine Funktion in einem Punkt ihren insgesamt größten bzw. kleinsten Wert annimmt.

Man spricht vom **absoluten Maximum** bzw. **Minimum** oder auch vom **globalen Maximum** bzw. **Minimum** – Abbildung 5.4 zeigt die Skizze einer „zweidimensionalen Landschaftskontur“, bei der das absolute Maximum bzw. Minimum jeweils am Rand realisiert wird. (Maximum und Minimum müssen aber natürlich nicht am Rand liegen.)

Ebenso relevant sind aber die Hügel und Täler mit zumindest lokal maximaler bzw. lokal minimaler „Energie“ im Innern von I :

Die in x_0 bzw. x_1 dargestellten Kugeln werden dort liegen bleiben, wobei ein noch so kleiner Anstoß reicht, die Kugel im Punkt x_0 ins Rollen zu bringen.

Bei kleinen Anstößen wird die Kugel in x_1 wieder in ihre Position zurückrollen. Stößt man sie aber mit viel Schwung an, so kann sie jedoch beispielsweise den „Hügel“ in x_0 überwinden.

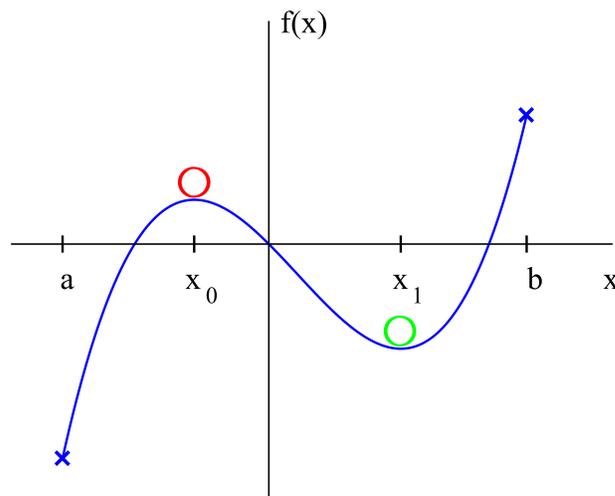


Abbildung 5.4: Lokale und globale Maxima (Minima) sind zu unterscheiden.

Definition 5.3. LOKALE EXTREMA

Eine Funktion $f: I \rightarrow \mathbb{R}$ hat an der Stelle $x_0 \in I$ ein *lokales Maximum* (*lokales Minimum*) $f(x_0)$, falls ein $r > 0$ existiert mit

$$f(x) \leq f(x_0) \quad (\text{lok. Max.}) \quad \text{bzw.} \quad f(x_0) \leq f(x) \quad (\text{lok. Min.})$$

für alle $x \in I \cap (x_0 - r, x_0 + r)$.

Gilt für $x \neq x_0$ jeweils die strikte Ungleichung, so spricht man von einem *strengen lokalen Maximum* (*Minimum*).

Sprechweise: Lokale Maxima und lokale Minima heißen *lokale Extrema*.

Punkte x , in denen lokale Maxima bzw. lokale Minima angenommen werden, heißen *lokale Maximierer* bzw. *lokale Minimierer* (*lokale Maximalstellen* bzw. *lokale Minimalstelle*).

Bemerkung. Ein Vergleich mit Satz 4.3 zeigt unmittelbar, dass globale Extrema automatisch auch lokale Extrema sind.

Anhand von Abbildung 5.5 wird deutlich, dass bei der Suche nach lokalen Extrema im Inneren des Definitionsbereiches zunächst Punkte mit **horizontaler Tangente an den Graphen** zu finden sind.

Aber nicht alle Punkte mit horizontaler Tangente sind lokale Extrema, wie Abbildung 5.6 zeigt (typisches Beispiel: $f(x) = x^3$ im Punkt $x = 0$). Trotz horizontaler Tangente liegt im Punkt x_0 kein Extremum vor, es handelt sich um einen sogenannten **Sattelpunkt**.

Mit anderen Worten: Die Bedingung „horizontale Tangente“ in einem inneren Punkt ist eine **notwendige** aber keine **hinreichende**.

Schließlich ist bei der Diskussion von Abbildung 5.5 noch zu beachten, dass die Steigung des Graphen selbst im Falle eines globalen Extremums am Rand von I (hier ist die einseitige Ableitung zu betrachten) nicht verschwinden muss.

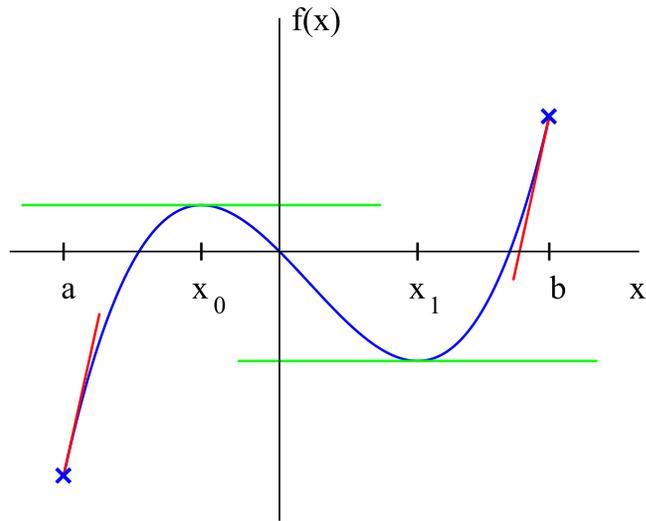


Abbildung 5.5: Die Tangente in einem lokalen Extremum und die Steigung am Rande.

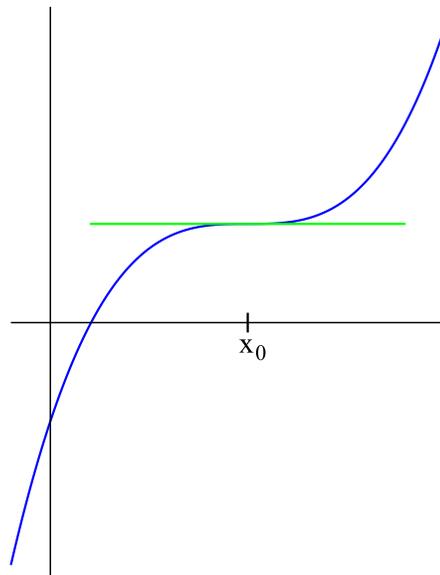


Abbildung 5.6: Kein lokales Extremum trotz horizontaler Tangente.

Satz 5.5. NOTWENDIGE BEDINGUNG

Die Funktion $f: I \rightarrow \mathbb{R}$ habe in einem Punkt $x_0 \in (a, b)$ ein lokales Extremum und sei dort differenzierbar.

Dann gilt

$$f'(x_0) = 0 .$$

Sprechweise: Punkte mit horizontaler Tangente an den Graphen, d.h. mit $f'(x_0) = 0$ heißen stationäre oder kritische Punkte.

Beweis. Es sei $x_0 \in (a, b)$ ein lokaler Minimierer von f (Maximierer analog).

Es existiert ein $\delta > 0$ mit $(x_0 - \delta, x_0 + \delta) \subset (a, b)$, wobei δ hinreichend klein mit $\delta < r$ (vgl. Definition 5.3) gewählt wird.

Aus der Eigenschaft „lokaler Minimierer“ folgt

$$f(x_0 + h) - f(x_0) \geq 0$$

für alle h mit $0 < |h| < \delta$, d.h.

$$\frac{f(x_0 + h) - f(x_0)}{h} \geq 0 \quad \text{für alle } 0 < h < \delta ,$$

$$\frac{f(x_0 + h) - f(x_0)}{h} \leq 0 \quad \text{für alle } -\delta < h < 0 .$$

Dabei hat sich in der zweiten Zeile das Vorzeichen vertauscht, da die Abschätzung $f(x_0 + h) - f(x_0) \geq 0$ durch die negative Zahl h geteilt wurde.

Aus der ersten Zeile und der Differenzierbarkeit von f folgt weiter $f'(x_0) \geq 0$, aus der zweiten $f'(x_0) \leq 0$, und wie behauptet ist $f'(x_0) = 0$. \square

Die Suche nach globalen Extrema.

Zur Herleitung einer hinreichenden Bedingung für lokale Extrema sind noch einige Vorbereitungen notwendig. An dieser Stelle kann aber schon festgehalten werden, wie globale Extrema ermittelt werden können:

- i) Im ersten Schritt wird untersucht, ob solche überhaupt existieren. Hier ist Satz 4.3 das wichtigste Hilfsmittel.
- ii) Anschließend wird dort, wo die Funktion differenzierbar ist, nach Punkten mit horizontaler Tangente gesucht. Die Funktionswerte an diesen Stellen sind die ersten Kandidaten für globale Extrema.
- iii) Im nächsten Schritt wird das Randverhalten untersucht: Ist beispielsweise das Definitionsintervall kompakt, $I = [a, b]$, so berechnet man die Funktionswerte in a und b . Auch diese sind Kandidaten.
- iv) In den Punkten, in denen die Funktion nicht differenzierbar ist, sind die Funktionswerte als weitere Kandidaten zu berechnen (z.B. wird das Minimum der Funktion $f(x) = |x|$ im Punkt $x = 0$ angenommen, in dem f nicht differenzierbar ist).
- v) Vergleicht man nun alle Kandidaten miteinander, so findet man die globalen Maxima bzw. Minima (falls existent).

Beispiel. Betrachtet sei die Funktion $f: [-2, 2] \rightarrow \mathbb{R}$, $f(x) = (1 - x^2)^2$.

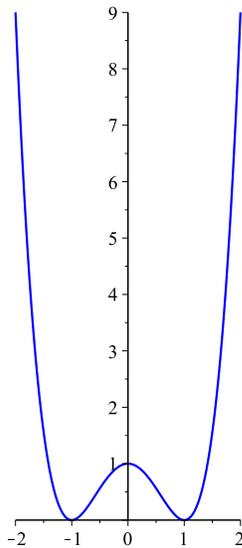


Abbildung 5.7: Die Funktion $f(x) = (1 - x^2)^2$ in $[-2, 2]$.

Das Definitionsintervall ist kompakt, die Funktion ist stetig und Satz 4.3 zeigt, dass zumindest ein globales Maximum und zumindest ein globales

Minimum existiert.

Es gilt

$$f'(x) = -4x(1 - x^2) ,$$

woraus sich die kritischen Punkte $x_1 = -1$, $x_2 = 0$ und $x_3 = 1$ ergeben.

Wegen $f \geq 0$ und $f(x_1) = f(x_3) = 0$ kann in diesem speziellen Beispiel direkt festgestellt werden, dass f in x_1 und x_3 ein globales Minimum (und damit automatisch auch ein lokales Minimum) hat.

Da für alle $x \in [-1, 1]$ die Ungleichung $f(x) \leq 1 = f(x_2)$ gilt, liegt in x_2 ein lokales Maximum vor – dies wird später auch mithilfe einer hinreichenden Bedingung verifiziert.

Zurück zur Suche nach globalen Maxima: Die Funktion ist in allen inneren Punkten differenzierbar und als Kandidaten sind lediglich die Funktionswerte in den kritischen Punkten und die Funktionswerte in den Randpunkten zu ermitteln:

$f(-2) = 9 = f(2)$ zeigt, dass das globale Maximum in den Randpunkten 2 und -2 und nicht im Punkt $x_2 = 0$ realisiert wird.

Der Satz von Rolle.

Der **Satz von Rolle** ist der nächste Schritt bei der Analyse einer Funktion $f: [a, b] \rightarrow \mathbb{R}$.

Wie in Abbildung 5.8 illustriert, besagt der Satz u.a., dass eine differenzierbare Funktion zwischen zwei Nullstellen einen kritischen Punkt haben muss, genauer gesagt gilt:

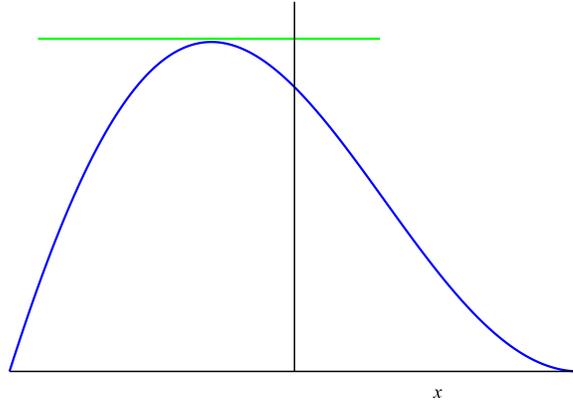


Abbildung 5.8: Zum Satz von Rolle.

Satz 5.6. SATZ VON ROLLE

Es sei $f: [a, b] \rightarrow \mathbb{R}$ stetig und differenzierbar auf (a, b) .

Ist $f(a) = f(b)$, so existiert ein $x_0 \in (a, b)$ mit $f'(x_0) = 0$.

Beweis. Nach Satz 4.3 nimmt die stetige Funktion f auf dem Kompaktum $[a, b]$ sowohl ein globales Maximum M als auch ein globales Minimum m an.

Im Fall $m = M$ ist die Funktion konstant und es gilt $f' \equiv 0$.

Nun zum Fall $m < M$: In diesem Fall kann wegen $f(a) = f(b)$ höchstens einer dieser Werte in einem Randpunkt angenommen werden.

In (a, b) wird also entweder das absolute Maximum oder das absolute Minimum (oder beides) angenommen.

Nach Satz 5.5 verschwindet dort die Ableitung von f . □

Der Mittelwertsatz.

In Verallgemeinerung des Satzes von Rolle besagt der **Mittelwertsatz**, dass für eine gegebene differenzierbare Funktion f im Intervall (a, b) (mindestens) ein Punkt x_0 existiert, in dem die **Tangente an den Graphen von f die gleiche Steigung hat wie die Sekante durch die Punkte $(a, f(a))$ und $(b, f(b))$** (vgl. Abbildung 5.9).

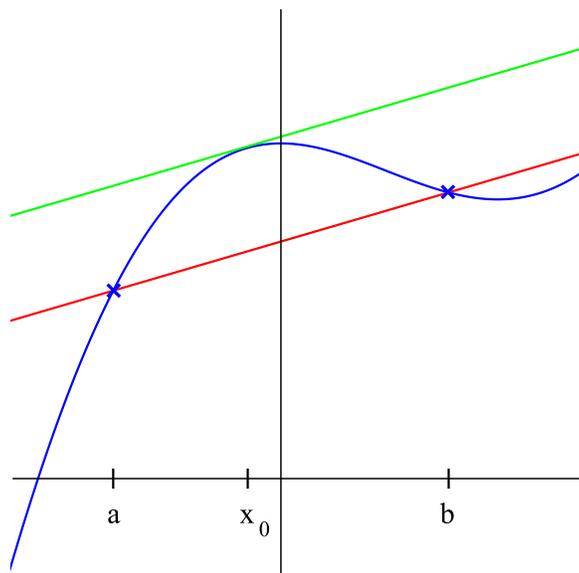


Abbildung 5.9: Zum Mittelwertsatz.

Satz 5.7. MITTELWERTSATZ

Die Funktion $f: [a, b] \rightarrow \mathbb{R}$ sei stetig und auf (a, b) differenzierbar.

Dann existiert ein $x_0 \in (a, b)$ mit

$$f'(x_0) = \frac{f(b) - f(a)}{b - a}.$$

Erste Konsequenzen aus dem Mittelwertsatz sind im folgenden Korollar festgehalten, welches als Übungsaufgabe gezeigt werden kann.

Korollar 5.1.**FOLGERUNGEN AUS DEM MITTELWERTSATZ**

Es sei f wie in Satz 5.7.

- i) Ist $f'(x) = 0$ für alle $x \in (a, b)$, so ist f konstant.
- ii) Es ist $f'(x) \geq 0$ für alle $x \in (a, b)$ *genau dann, wenn* f auf $[a, b]$ *monoton wachsend* ist (analog „ \leq “ und *monoton fallend*).
- iii) Ist $f'(x) > 0$ für alle $x \in (a, b)$, so ist f *streng monoton wachsend* auf $[a, b]$ (analog: „ $<$ “ und *streng monoton fallend*).

Hier ist die Umkehrung i.A. falsch, wie die Funktion $f(x) = x^3$ belegt.

Nun können Funktionen auch nach ihrem Krümmungsverhalten klassifiziert werden.

Konvexe und konkave Funktionen.

Eine typische (streng) **konvexe Funktion** ist die Funktion $f: \mathbb{R} \rightarrow \mathbb{R}$, $f(x) = x^2$ – der Graph ist eine „Linkskurve“.

Eine typische (streng) **konkave Funktion** ist der Logarithmus $\ln: \mathbb{R}^+ \rightarrow \mathbb{R}$ – der Graph ist eine „Rechtskurve“.

Die geometrische Vorstellung ist dabei, dass der Graph konvexer (konkaver) Funktionen stets oberhalb (unterhalb) aller seiner **Tangenten** liegt. Dies ist den Abbildungen 5.10 und 5.11 verdeutlicht.

Umgekehrt liegt die **Sekante** durch die Punkte $(a, f(a))$ und $(b, f(b))$ bei einer konvexen (konkaven) Funktion im Intervall (a, b) oberhalb (unterhalb) des Graphen (vgl. Abbildung 5.12 und Abbildung 5.13).

Mit anderen Worten ist der Funktionswert einer konvexen (konkaven) Funktion kleiner oder gleich (größer oder gleich) dem Sekantenwert an dieser Stelle im Sinne von:

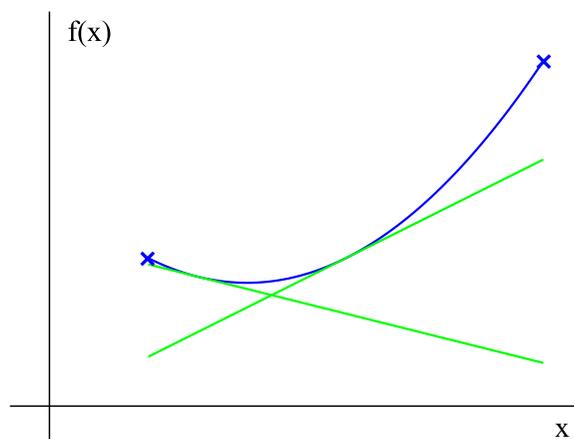


Abbildung 5.10: Der Graph einer konvexen Funktion liegt oberhalb seiner Tangenten.

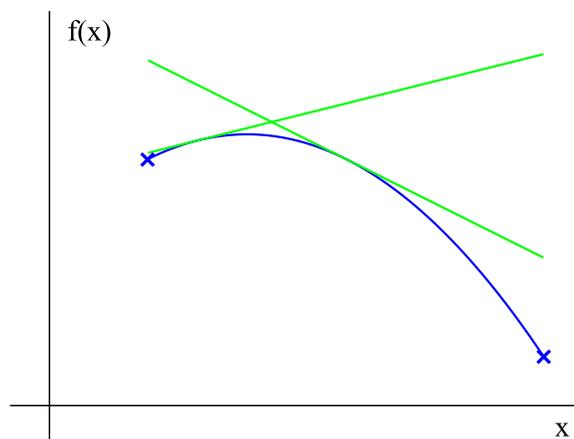


Abbildung 5.11: Der Graph einer konkaven Funktion liegt unterhalb seiner Tangenten.

- i)* Eine Funktion $f: I \rightarrow \mathbb{R}$ heißt konvex, falls für alle $x, y \in I$ und für alle $\lambda \in (0, 1)$ gilt

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y) .$$

- ii)* Eine Funktion $f: I \rightarrow \mathbb{R}$ heißt konkav, falls für alle $x, y \in I$ und für

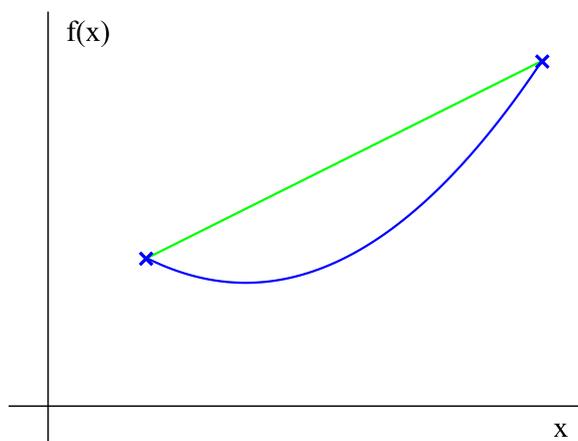


Abbildung 5.12: Zwischen zwei Punkten liegt die Sekante (durch diese Punkte) einer konvexen Funktion oberhalb des Graphen.

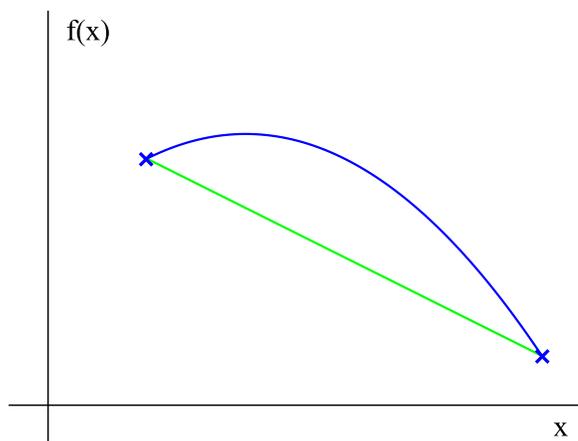


Abbildung 5.13: Zwischen zwei Punkten liegt die Sekante (durch diese Punkte) einer konkaven Funktion unterhalb des Graphen.

alle $\lambda \in (0, 1)$ gilt

$$f(\lambda x + (1 - \lambda)y) \geq \lambda f(x) + (1 - \lambda)f(y) .$$

iii) Bei **streng konvexen** bzw. **streng konkaven** Funktionen gelten die

strikten Ungleichungen.

Ist f konvex (konkav) und zweimal differenzierbar, so erkennt man anhand von Abbildung 5.10 (Abbildung 5.11), dass die **erste Ableitung f' monoton wachsend (fallend) ist**.

Für die Ableitung f'' der monoton wachsenden (fallenden) Funktion f' gilt folglich $f'' \geq 0$ ($f'' \leq 0$).

Satz 5.8. ZWEITE ABLEITUNG UND KONVEXITÄT/KONKAVITÄT

Es sei $f: (a, b) \rightarrow \mathbb{R}$ zweimal differenzierbar.

- i) Ist $f''(x) > 0$ für alle $x \in (a, b)$, so ist f **streng konvex**.
- ii) Ist $f''(x) < 0$ für alle $x \in (a, b)$, so ist f **streng konkav**.

Hinreichende Bedingung für lokale Extrema.

Mit diesen Vorbereitungen kann nun eine hinreichende Bedingung für die Existenz lokaler Extrema formuliert werden. Wieder ist die geometrische Vorstellung in den Abbildungen 5.10 und 5.11 wiedergegeben:

Strenge Konvexität (Konkavität) in der Nähe eines kritischen Punktes impliziert die Existenz eines lokalen Minimums (Maximums).

Satz 5.9. HINREICHENDE BEDINGUNG

Es sei $f: I = (a, b) \rightarrow \mathbb{R}$ von der Klasse $C^2(I)$.

Im Punkt $x_0 \in (a, b)$ gelte weiter $f'(x_0) = 0$.

- i) Ist $f''(x_0) > 0$, so hat f in x_0 ein **strenges lokales Minimum**.

ii) Ist $f''(x_0) < 0$, so hat f in x_0 ein *strenges lokales Maximum*.

Bemerkung. Der Satz macht *keine Aussage* über den Fall $f''(x_0) = 0$. In diesem Fall ist das Verhalten der Funktion bzw. der Ableitung genauer zu analysieren (vgl. Übungskapitel 5.4).

Im oben diskutierten Beispiel $f(x) = (1 - x^2)^2$, $x \in [-2, 2]$ ist

$$f''(x) = 12x^2 - 4.$$

Für die kritischen Punkte $x_1 = -1$ und $x_3 = 1$ gilt demnach $f''(x_1) > 0$, $f''(x_3) > 0$ und es handelt sich (wie bereits gesehen) um strenge lokale Minima.

Im Punkt $x_2 = 0$ ist $f''(x_2) < 0$, es handelt sich um ein strenges lokales Maximum.

Wendepunkte.

Neben Extrema sind *Wendepunkte* charakteristisch für eine Funktion.

In einem Wendepunkt geht der Graph der Funktion von *einer „Linkskurve in eine Rechtskurve“ über* (oder umgekehrt), d.h. von einem *konvexen in einen konkaven Bereich* (oder umgekehrt).

Wendepunkte können als (strenge) lokale Extrema der ersten Ableitung interpretiert werden – die Situation ist in Abbildung 5.14 angedeutet.

Es sei $f: (a, b) \rightarrow \mathbb{R}$ von der Klasse $C^3((a, b))$. Analog zur Diskussion lokaler Extrema gilt:

- i) In einem Wendepunkt $x_0 \in (a, b)$ gilt immer $f''(x_0) = 0$.
- ii) Ist $f''(x_0) = 0$ und $f'''(x_0) \neq 0$, so ist x_0 ein Wendepunkt.

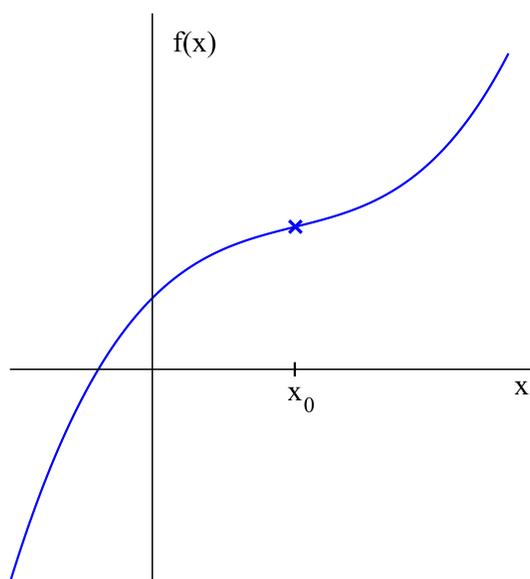


Abbildung 5.14: Ein Wendepunkt.

Die Regeln von l'Hospital.

Zum Abschluss dieses Paragraphen wird nun eine wichtige Folgerung aus dem (verallgemeinerten) Mittelwertsatz vorgestellt, mit deren Hilfe Grenzwerte von Quotienten berechnet werden können, die sich nicht aus den bisher bekannten Regeln erschließen.

Beispielsweise liefert Satz 5.10 den Grenzwert vom Typ „ $\frac{0}{0}$ “

$$\lim_{x \rightarrow 0} \frac{\sin(x)}{x}.$$

Satz 5.10. REGELN VON L'HOSPITAL

Es seien $f, g: (a, b) \rightarrow \mathbb{R}$ differenzierbar und $x_0 \in (a, b)$.

Ist $f(x_0) = g(x_0) = 0$, ist $g'(x_0) \neq 0$ für $x \neq x_0$ und existiert der Grenzwert

$$\lim_{x \rightarrow x_0} \frac{f'(x)}{g'(x)},$$

so gilt

$$\lim_{x \rightarrow x_0} \frac{f(x)}{g(x)} = \lim_{x \rightarrow x_0} \frac{f'(x)}{g'(x)} .$$

Beispiele.

i) Es ist

$$\lim_{x \rightarrow 0} \frac{\sin(x)}{x} = \lim_{x \rightarrow 0} \frac{\cos(x)}{1} = 1 .$$

ii) Die Regel kann auch **mehrfach hintereinander angewandt werden**:

$$\lim_{x \rightarrow 0} \frac{1 - \cos(x)}{x^2} = \lim_{x \rightarrow 0} \frac{\sin(x)}{2x} = \lim_{x \rightarrow 0} \frac{\cos(x)}{2} = \frac{1}{2} .$$

iii) Eine analoge Aussage gilt für **Grenzwerte $x \rightarrow \infty$** , falls f und g im Unendlichen gegen Null konvergieren, z.B.

$$\begin{aligned} \lim_{x \rightarrow \infty} [\sqrt{x(x+1)} - x] &= \lim_{x \rightarrow \infty} \frac{\sqrt{1 + \frac{1}{x}} - 1}{\frac{1}{x}} \\ &= \lim_{x \rightarrow \infty} \frac{\frac{1}{2\sqrt{1 + \frac{1}{x}}} \left[-\frac{1}{x^2} \right]}{\left[-\frac{1}{x^2} \right]} \\ &= \lim_{x \rightarrow \infty} \frac{1}{2\sqrt{1 + \frac{1}{x}}} = \frac{1}{2} . \end{aligned}$$

iv) Ebenso gelten die Regeln für **Grenzwerte der Form $x \rightarrow x_0$, $f(x)$, $g(x) \rightarrow \pm\infty$** . Wie überall sind dabei auch einseitige Grenzwerte zugelassen:

$$\lim_{x \rightarrow 0^+} x \ln(x) = \lim_{x \rightarrow 0^+} \frac{\ln(x)}{\frac{1}{x}} = \lim_{x \rightarrow 0^+} \frac{\frac{1}{x}}{-\frac{1}{x^2}} = \lim_{x \rightarrow 0^+} (-x) = 0 .$$

5.3 Numerische Differentiation (Extrapolation zum Limes $h \rightarrow 0$; zentraler Differenzenquotient)

Grenzwerte vom Typ „ $\frac{0}{0}$ “, die nicht analytisch mit den Regeln von l’Hospital berechnet werden können, approximiert man mit numerischen Mitteln.

Man beachte, dass die Ableitung als Grenzwert des Differenzenquotienten von diesem Typ ist.

Die Grundidee ist eine Polynominterpolation nach Teil I, Kapitel 6, zusammen mit einer [Extrapolation zum Limes \$h \rightarrow 0\$](#) .

Die allgemeine Aufgabenstellung lautet:

Gegeben sei eine Funktion $g: (0, b] \rightarrow \mathbb{R}$, $b > 0$, wobei für $0 < h \leq b$ die Funktion g explizit ausgewertet werden könne. Zudem sei g als stetige Funktion in den Nullpunkt fortsetzbar.

Gesucht ist eine numerische Näherung für

$$\lim_{h \rightarrow 0^+} g(h) =: g(0) .$$

Beispiele.

i) Man approximiere numerisch

$$\lim_{h \rightarrow 0^+} \frac{h^{3/2}}{\sin(h)} .$$

ii) Man approximiere numerisch

$$\lim_{h \rightarrow 0^+} \frac{\cosh(h) - 1}{\sinh(h)} .$$

iii) Für eine differenzierbare Funktion $f: I \rightarrow \mathbb{R}$ setzt man in einem fixierten inneren Punkt $x_0 \in I$ und für $|h|$ hinreichend klein

$$g(h) := \frac{f(x_0 + h) - f(x_0)}{h} .$$

Gesucht ist in diesem Beispiel eine Näherung für den Grenzwert des Differenzenquotienten, d.h. für die Ableitung von f in x_0 .

Die allgemeine Lösungsidee.

- i) Zunächst werden die Werte $g(h_i)$, $i = 0, \dots, n$, für geeignete Stützstellen $h_0 > h_1 > \dots > h_n > 0$ und für geeignetes $n \in \mathbb{N}$ (nicht zu groß, um starke Oszillationen zu vermeiden, vgl. Teil I, Kapitel 6) berechnet.
 - ii) Man löse im nächsten Schritt die Interpolationsaufgabe von Lagrange (vgl. Teil I, Kapitel 6) zu diesen Daten, d.h. betrachtet sei das eindeutig bestimmte Interpolationpolynom p_n vom Grad n durch die Punkte $(h_0, g(h_0)), \dots, (h_n, g(h_n))$.
 - iii) Ist die Güte der Approximation von g mit p_n in der Nähe des Nullpunkts hinreichend gut, so ist $p_n(0)$ eine gute Näherung für $g(0)$.
-

Zu den ersten beiden Beispielen.

- i) Gesucht ist eine Näherung für

$$\lim_{h \rightarrow 0^+} \frac{h^{3/2}}{\sin(h)}.$$

Bekanntlich ist dieser Grenzwert 0.

In den Stützstellen $h_0 = 1/8$, $h_1 = 1/16$, $h_2 = 1/32$, $h_3 = 1/64$ liefert die Berechnung der Funktionswerte

h_i	1/8	1/16	1/32	1/64
$h_i^{3/2} / \sin(h_i)$	0.35447578	0.25016283	0.17680547	0.12500509

Anstatt den Wert 0.12500509 als Näherung zu wählen, legt man durch die berechneten Werte das Interpolationpolynom vom Grad 3,

wie es in der Abbildung 5.15 rot angedeutet ist.

Der exakte Verlauf ist grün gekennzeichnet und belegt eine recht gute Übereinstimmung.

Für $p_3(0)$ ergibt sich der Wert 0.05938205, der zwar nicht um Größenordnungen, aber dennoch deutlich näher bei der 0 liegt als der Wert 0.12500509.

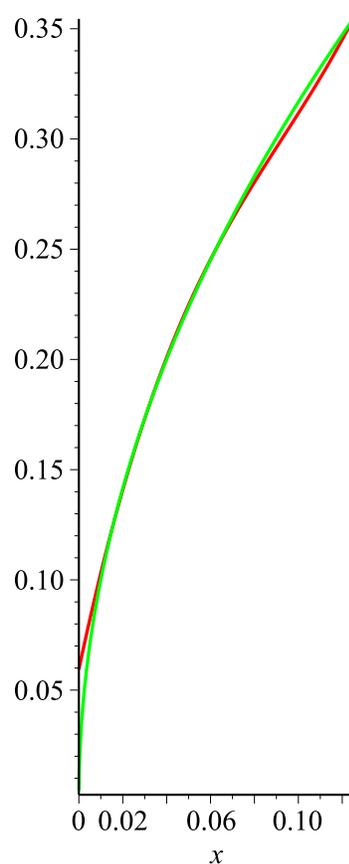


Abbildung 5.15: Extrapolation von $h^{3/2}/\sin(h)$ im Limes $h \rightarrow 0$.

ii) Mit den gleichen Stützstellen wie oben soll

$$\lim_{h \rightarrow 0^+} \frac{\cosh(h) - 1}{\sinh(h)} = 0$$

näherungsweise berechnet werden.

Die Daten lauten in diesem Beispiel

h_i	1/8	1/16	1/32	1/64
$\frac{\cosh(h_i)-1}{\sinh(h_i)}$	0.06241875	0.03123983	0.01562373	0.00781235

Die Extrapolation zum Limes $h \rightarrow 0$ liefert $p_3(0) = 3.638 \cdot 10^{-8}$, die Approximation ist also tatsächlich um Größenordnungen besser als der Wert 0.00781235.

In der für die h_i natürlichen Skalierung aus Abbildung 5.16 kann der Unterschied zwischen dem Approximationspolynom und der exakten Lösung nicht mehr aufgelöst werden.

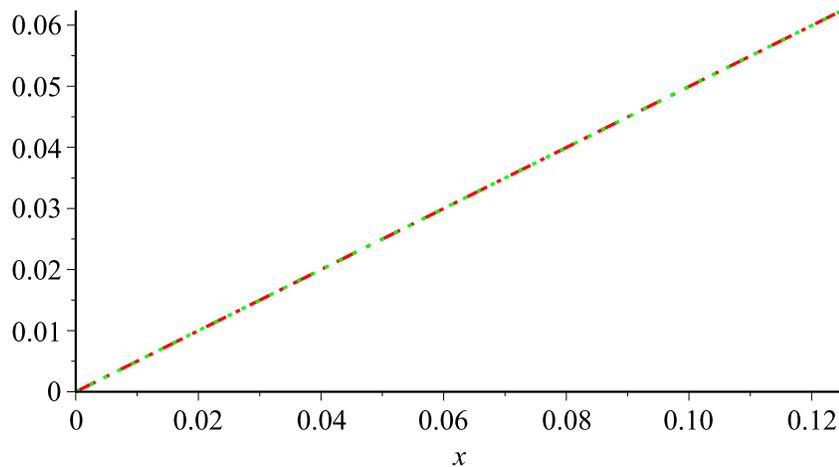


Abbildung 5.16: Extrapolation von $(\cosh(h) - 1) / \sinh(h)$ im Limes $h \rightarrow 0$.

Erst in der extremen Vergrößerung aus Abbildung 5.17 erkennt man die Abweichungen in der Nähe des Nullpunkts (wegen zu vieler

Rechenschritte ist die exakte Lösung nicht bis in den Ursprung dargestellt).

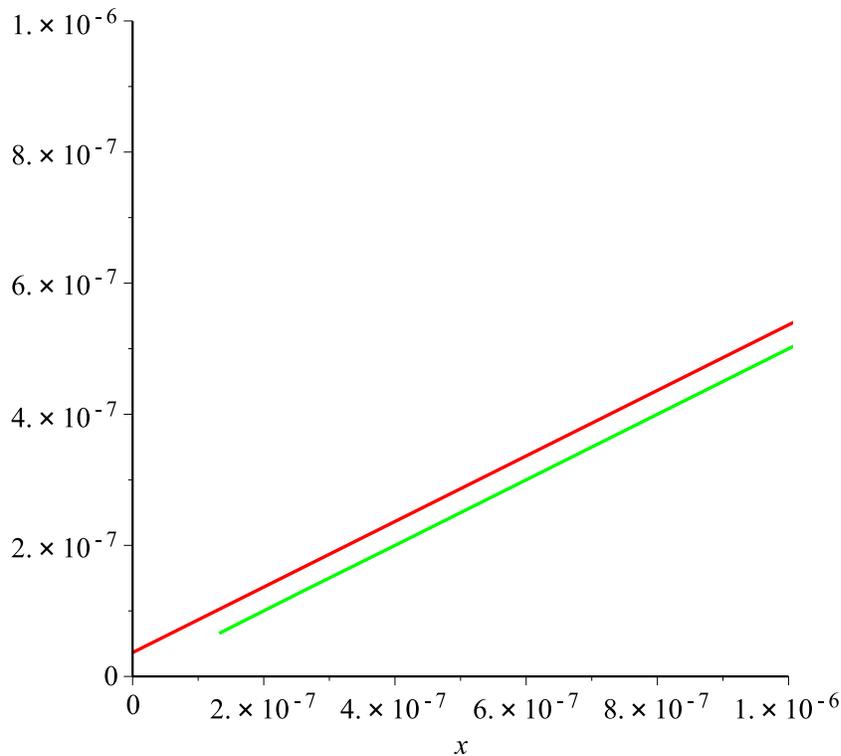


Abbildung 5.17: Extrapolation von $(\cosh(h) - 1) / \sinh(h)$ im Limes $h \rightarrow 0$.

Natürlich approximiert das Interpolationspolynom die Funktion nur in der Nähe des Ursprungs so gut.

Zur Illustration dieser Tatsache sind die Verläufe von Interpolationspolynom und exakter Lösung in Abbildung 5.18 in einem größeren Bereich dargestellt.

Die Abweichungen für große h sind für die gegebene Fragestellung jedoch irrelevant.

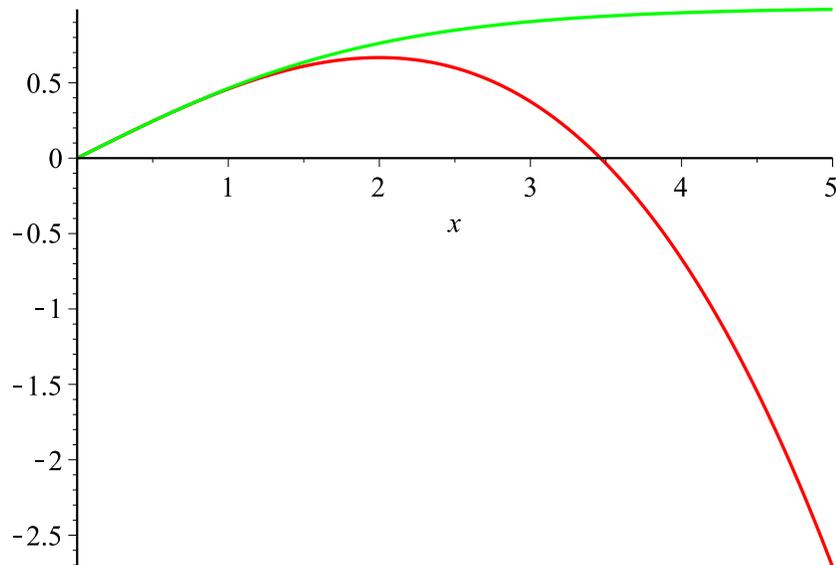


Abbildung 5.18: Extrapolation von $(\cosh(h) - 1) / \sinh(h)$ im Limes $h \rightarrow 0$.

Beispiel: Numerische Differentiation.

Zur numerischen Differentiation einer C^1 -Funktion $f: I = (a, b) \rightarrow \mathbb{R}$ in einem Punkt $x_0 \in I$ kann wie oben angemerkt der Differenzenquotient

$$g(h) = \frac{f(x_0 + h) - f(x_0)}{h}$$

herangezogen werden.

Die Extrapolation zum Limes $h \rightarrow 0$ liefert wie beschrieben einen Näherungswert für die Ableitung von f im Punkt x_0 . Aber:

Wichtige Beobachtung. Man betrachte den [zentralen Differenzenquotienten](#)

$$\delta(h) := \frac{f(x_0 + h) - f(x_0 - h)}{2h}.$$

Der zentrale Differenzenquotient $\delta(h)$ konvergiert im Limes $h \rightarrow 0$ ebenfalls gegen $f'(x_0)$ (Beweis: Übungskapitel 5.4).

Der wesentliche Vorteil im Hinblick auf diesen Paragraphen ist, dass der zentrale Differenzenquotient ein **gerade Funktion in h** ist, d.h.

$$\frac{f(x_0 + (-h)) - f(x_0 - (-h))}{-2h} = \frac{f(x_0 + h) - f(x_0 - h)}{2h}.$$

Dies legt die Idee nahe, mit geraden Polynomen, d.h. mit **Polynomen in Abhängigkeit von h^2 zu approximieren**.

Tatsächlich ist die Güte der Approximation in der Regel deutlich besser, wenn mit dem zentralen Differenzenquotienten in h^2 anstelle des Differenzenquotienten in h approximiert wird.

Beispiel. Es sei $f(x) = \sin(x)$. Mittels des zentralen Differenzenquotienten soll eine Näherung für $f'(0) = 1$ berechnet werden.

Der zentrale Differenzenquotient von f in 0 lautet

$$\delta(h) = \frac{\sin(h) - \sin(-h)}{2h} = \frac{\sin(h)}{h} = g(h^2).$$

Obwohl in den Differenzenquotienten das Argument h einzusetzen ist, handelt es sich tatsächlich um eine Funktion von h^2 .

In der Tat ist $\sin(h) = h - \frac{1}{3}h^3 + \dots$ und

$$\frac{\sin(h)}{h} = 1 - \frac{1}{3}h^2 + \dots$$

Man setzt

$$g(h) := 1 - \frac{1}{3}h + \dots, \quad \text{d.h. } g(h^2) = 1 - \frac{1}{3}h^2 + \dots = \frac{\sin(h)}{h}.$$

Es sei jetzt $h_0 = 1/8$, $h_1 = 1/16$, $h_2 = 1/32$.

Zu berechnen ist als Näherungswert für $(\sin(x))' = \cos(x)$ an der Stelle 0 das Interpolationspolynom der Ordnung 2 von g zu den Stützstellen h_i^2 , $i = 0, 1, 2$.

Dieses ist an der Stelle 0 auszuwerten.

Der [Algorithmus von Neville](#) (vgl. wieder Teil I, Kapitel 6) liefert die Werte

h_i^2	$g(h_i^2) = \sin(h_i)/h_i$	$k = 1$	$k = 2$
$(1/8)^2$	$p_{0,0}(0) = 0.997397867$		
$(1/16)^2$	$p_{1,0}(0) = 0.999349085$	$p_{1,1}(0) = 0.999999491$	
$(1/32)^2$	$p_{2,0}(0) = 0.999837248$	$p_{2,1}(0) = 0.999999969$	$p_{2,2}(0) = 1.000000001$

Man erkennt deutlich die Güte der berechneten Approximation.

Ein weiteres Beispiel (insbesondere zum Vergleich zwischen der Extrapolation mit dem Differenzenquotienten und der Extrapolation mit dem zentralen Differenzenquotienten) findet sich im Übungskapitel 5.4.

5.4 Übungsaufgaben zu Kapitel 5

Aufgabe 1. Es sei $\alpha > 0$ fixiert und $f: \mathbb{R} \rightarrow \mathbb{R}$, $f(x) = |x|^\alpha$ für alle $x \in \mathbb{R}$. Ist f differenzierbar im Nullpunkt?

Aufgabe 2.* Ist die Funktion $f: \mathbb{R} \rightarrow \mathbb{R}$,

$$f(x) = \begin{cases} x^2 \sin(1/x) & \text{für } x \neq 0, \\ 0 & \text{für } x = 0, \end{cases}$$

differenzierbar im Punkt $x_0 = 0$? Ist die Ableitung eine stetige Funktion?

Aufgabe 3.* Zeigen Sie Satz 5.1.

Aufgabe 4.* Zeigen Sie: Für alle $k \in \mathbb{Z}$ ($x \neq 0$ für $k < 0$) gilt

$$\frac{d}{dx} x^k = kx^{k-1}.$$

Aufgabe 5.

i) Bestimmen Sie mit Hilfe der Reihendarstellungen die Ableitungen von $\sin(x)$, $\cos(x)$, $\sinh(x)$, $\cosh(x)$.

Benutzen Sie dann die Quotientenregel, um die Ableitung von $\cot(x)$ zu berechnen (wo sie existiert).

ii) Berechnen Sie (dort wo sie existieren) die Ableitungen der Umkehrfunktionen $\arcsin(x)$, $\arccos(x)$, $\arctan(x)$ der trigonometrischen Funktionen.

Aufgabe 6. Berechnen Sie die Ableitungen der folgenden Funktionen (wo sie definiert sind):

i) $f(x) = 2^{\ln(x)}$;

ii) $f(x) = a^{(x^2)}$, $a > 0$ fixiert;

iii) $f(x) = \exp(\exp(\exp(x)))$;

iv) $f(x) = x^x$;

v) $f(x) = \frac{\sqrt{x^2 + 1}}{\sqrt{2x + 2}}$;

vi) $f(x) = \arcsin(1/(1 + x^2))$.

Aufgabe 7. Es sei $I \subset \mathbb{R}$ ein offenes Intervall und es sei $[a, b] \subset I$. Die Funktion $f: I \rightarrow \mathbb{R}$ sei differenzierbar auf ganz I und die Funktion f' sei eine auf dem Intervall $[a, b]$ stetige Funktion. Zeigen Sie, dass f eine auf $[a, b]$ Lipschitz-stetige Funktion ist.

Hinweis. Man benutze Satz 4.3 der Vorlesung.

Aufgabe 8.

i) Die Funktionen f und $g: \mathbb{R} \rightarrow \mathbb{R}$ seien dreimal differenzierbar. Berechnen Sie $(f \cdot g)^{(3)}$.

ii) Finden Sie eine Funktion $f: \mathbb{R} \rightarrow \mathbb{R}$, die auf \mathbb{R} genau zweimal differenzierbar ist.

iii) Es seien $a < b \in \mathbb{R}$ fixiert und $f: (a, b) \rightarrow \mathbb{R}$ die durch $f(x) = \ln(x)/x$ definierte Funktion. Ist f konvex bzw. konkav?

- iv)* Berechnen Sie die zweite Ableitung der Funktion $f: f(x) = \ln(\ln(x))$.
Für welche $x \in \mathbb{R}$ ist die Funktion definiert?
-

Aufgabe 9.*

- i)* Zeigen Sie Korollar 5.1.
- ii)* Es sei $U = (0, 1) \cup (2, 3)$ und $f: U \rightarrow \mathbb{R}$ sei differenzierbar mit $f'(x) = 0$ für alle $x \in U$. Ist f eine konstante Funktion?
-

Aufgabe 10.* Betrachten Sie die Funktionen $f, g: \mathbb{R} \rightarrow \mathbb{R}$,

$$f(x) = x^2 \sin(x), \quad g(x) = x^2 \cos(x).$$

Ist der Punkt $x_0 = 0$

- i)* ein kritischer Punkt;
ii) eine lokale Minimalstelle bzw. Maximalstelle;
iii) ein Sattelpunkt

von f bzw. g ?

Aufgabe 11. Es seien $f: I \rightarrow \mathbb{R}$ und $I \subset \mathbb{R}$ jeweils gegeben durch

- i)* $f(x) = x^2 - 2x + 1, I = [-2, 2]$;
ii) $f(x) = x^2 - 2|x| + 1, I = [-2, 2]$;
iii) $f(x) = x^2 - 2|x| + 1, I = \mathbb{R}$.

Bestimmen Sie alle lokalen und globalen Maxima und Minima von f .

Aufgabe 12.* Es sei $I = [-1, 1] \subset \mathbb{R}$ und $f: I \rightarrow \mathbb{R}$ gegeben durch

$$f(x) = x^2|x + x^2|.$$

Ist f stetig auf I ? Ist f differenzierbar auf $(-1, 1)$? Bestimmen Sie alle lokalen und globalen Maxima und Minima von f auf I .

Aufgabe 13.* Es sei $I = [-2, 2] \subset \mathbb{R}$ und die stetige Funktion $f: I \rightarrow \mathbb{R}$ sei gegeben durch

$$f(x) = \begin{cases} xe^{x-1} & \text{für } x \leq 0, \\ xe^{1-x} & \text{für } x > 0. \end{cases}$$

Ist f differenzierbar auf $(-2, 2)$? Bestimmen Sie alle lokalen und globalen Maxima und Minima von f auf I .

Aufgabe 14.

i) Berechnen Sie (falls existent) die Grenzwerte

$$\lim_{x \rightarrow 0} x^x, \quad \lim_{x \rightarrow \infty} x^{\frac{1}{x}}.$$

ii) Berechnen Sie (falls existent)

$$\lim_{x \rightarrow 0} \left[\frac{1}{\sin(x)} - \frac{1}{x} \right].$$

iii) Berechnen Sie (falls existent)

$$\lim_{x \rightarrow 0} \frac{1 - \frac{x^2}{2} - \cos(x)}{x \sin(x)}.$$

Aufgabe 15. Es sei $f: I \rightarrow \mathbb{R}$ im Punkt $x_0 \in I$ differenzierbar. Zeigen Sie: Der zentrale Differenzenquotient in x_0 konvergiert gegen $f'(x_0)$.

Aufgabe 16. Es sei $x > -1$, $f(x) = \sqrt{1+x}$, $n = 2$, $h_0 = 1/8$, $h_1 = 1/16$, $h_2 = 1/32$.

Berechnen Sie einen Näherungswert („Extrapolation zum Limes $h \rightarrow 0$ “, 8 Nachkommastellen) für $f'(0)$ mittels

i) des Differenzenquotienten;

ii) des zentralen Differenzenquotienten (als Polynom in h_i^2).

Aufgabe 17. Studieren Sie die Möglichkeiten Ihres Computeralgebrasystems in der Differentialrechnung.

Lösungshinweise zu den Übungsaufgaben.

Aufgabe 2. Für $x_0 \neq 0$ ist die Funktion differenzierbar mit

$$\begin{aligned} f'(x_0) &= 2x_0 \sin(1/x_0) + x_0^2 \cos(1/x_0)(-1/x_0)^2 \\ &= 2x_0 \sin(1/x_0) - \cos(1/x_0) . \end{aligned}$$

Für $x_0 = 0$

$$\begin{aligned} \left| \frac{f(x) - f(x_0)}{x - 0} \right| &= |x| |\sin(1/x)| \\ &\leq |x| \rightarrow 0 , \end{aligned}$$

falls $x \rightarrow 0$, d.h. f ist im Nullpunkt differenzierbar mit $f'(0) = 0$.

Aufgabe 3. Es sei $\{x_n\}$ eine Folge aus I , $x_n \neq x_0$, mit $x_n \rightarrow x_0$ für $n \rightarrow \infty$.
Zu zeigen ist:

$$f(x_n) \xrightarrow{n \rightarrow \infty} f(x_0) .$$

Nach Voraussetzung ist f in x_0 differenzierbar, was

$$\frac{f(x_n) - f(x_0)}{x_n - x_0} \xrightarrow{n \rightarrow \infty} f'(x_0) .$$

bedeutet.

Aus den Rechenregeln für Folgen ergibt sich

$$f(x_n) - f(x_0) = \frac{f(x_n) - f(x_0)}{x_n - x_0} (x_n - x_0) \xrightarrow{n \rightarrow \infty} f'(x_0) \cdot 0 = 0 ,$$

also genau die Behauptung. □

Aufgabe 4. Der Beweis wird mit vollständiger Induktion geführt.

Induktionsanfang: Für $k = 0$ und $k = 1$ ist die Aussage bereits verifiziert.

Induktionsschluss: Es sei nun angenommen, dass die Aussage für ein $k \in \mathbb{N}$ richtig sei. Dann folgt aus der Produktregel nach der Induktionsannahme

$$\frac{d}{dx}x^{k+1} = \frac{d}{dx}(x^k x) = kx^{k-1}x + x^k 1 = (k+1)x^k,$$

die Aussage gilt somit auch für $k+1$ und die Behauptung ist für alle $k \in \mathbb{N}_0$ bewiesen.

Für negative $k \in \mathbb{Z}$ folgt die Behauptung aus der Quotientenregel. \square

Aufgabe 9.

- i)* *i)* Ist $f'(x) = 0$ für alle $x \in (a, b)$, so folgt aus dem Mittelwertsatz für alle $x_1, x_2 \in (a, b)$ (o.E. $x_1 < x_2$) und für ein $\xi \in (x_1, x_2)$

$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} = f'(\xi) = 0,$$

also die erste Behauptung.

- ii)* Ist $f'(x) \geq 0$ für alle $x \in (a, b)$, so ergibt sich genau wie in *i)*: $f(x_2) \geq f(x_1)$ für $x_2 > x_1$, d.h. die Monotonie.

Ist umgekehrt f monoton wachsend und in $x_0 \in (a, b)$ differenzierbar, so gilt ($x > x_0$)

$$f'(x_0) = \lim_{x \rightarrow x_0^+} \frac{f(x) - f(x_0)}{x - x_0} \geq 0.$$

- iii)* Die Aussage *iii)* folgt analog zum ersten Teil von *ii)* und das Korollar ist bewiesen. \square

- ii)* Finden Sie ein einfaches Gegenbeispiel.

Aufgabe 10. Für alle $x \in \mathbb{R}$ ist

$$f'(x) = 2x \sin(x) + x^2 \cos(x), \quad g'(x) = 2x \cos(x) - x^2 \sin(x)$$

und wegen $f'(0) = g'(0) = 0$ handelt es sich in beiden Fällen um einen kritischen Punkt.

Die zweiten Ableitungen berechnen sich zu

$$\begin{aligned} f''(x) &= 2 \sin(x) + 4x \cos(x) - x^2 \sin(x), \\ g''(x) &= 2 \cos(x) - 4x \sin(x) - x^2 \cos(x). \end{aligned}$$

Aus $g''(0) = 2 > 0$ folgt mit der hinreichenden Bedingung, dass der Punkt $x_0 = 0$ eine lokale Minimalstelle von g ist.

Wegen $f''(0) = 0$ liefert Satz 5.9 aber keine Aussage für die Funktion f .

Allerdings ist für $x < 0$, $|x|$ hinreichend klein, $f(x) < 0$ und für $x > 0$, $|x|$ hinreichend klein, $f(x) > 0$, d.h. im Nullpunkt kann kein lokales Extremum vorliegen – der Punkt $x_0 = 0$ ist ein Sattelpunkt der Funktion f .

Aufgabe 12. Als Summe, Produkt und Verkettung stetiger Funktionen ist die Funktion f stetig und nimmt auf dem kompakten Intervall $[-1, 1]$ nach Satz 4.3 ihr globales Maximum bzw. Minimum an.

Analog folgt, dass die Funktion auf $(-1, 1)$ für $x \neq 0$ differenzierbar ist.

Nach der Definition des Betrages kann f geschrieben werden als

$$\begin{aligned} f(x) &= \begin{cases} x^2(x + x^2) & \text{für } 0 \leq x + x^2 = x(1 + x), \\ x^2(-(x + x^2)) & \text{für } 0 > x + x^2 = x(1 + x), \end{cases} \\ &= \begin{cases} x^3(1 + x) & \text{für } 0 \leq x \leq 1, \\ -x^3(1 + x) & \text{für } -1 \leq x < 0. \end{cases} \end{aligned}$$

Wegen

$$\lim_{x \rightarrow 0} \frac{\pm x^3(1+x) - 0}{x - 0} = \lim_{x \rightarrow 0} \pm x^2(1+x) = 0$$

ist die Funktion in $x_1 = 0$ differenzierbar mit der Ableitung 0.

Da $f(x) \geq 0$ für alle $x \in I$ und da $f(0) = 0$ gilt, handelt es sich im Punkt $x_1 = 0$ um ein globales (und damit auch lokales) Minimum.

Für $x \neq 0$ ist

$$f'(x) = \pm(3x^2(1+x) + x^3) = \pm x^2(4x+3) = 0$$

im Punkt $x_2 = -3/4$.

Für $-1 \leq x < 0$ ist aber $f''(x) = -x(12x+6)$ und $f''(-3/4) < 0$ impliziert, dass im Punkt $x_1 = -3/4$ ein lokales Maximum vorliegt.

Bei der Suche nach evtl. weiteren globalen Minima bzw. globalen Maxima ist die Funktion noch in den Randpunkten $x_3 = -1$ und $x_4 = 1$ auszuwerten und mit den beiden bisherigen Kandidaten zu vergleichen.

Aus

$$f(x_1) = f(x_3) = 0 < f(x_2) = \frac{27}{256} < f(x_4) = 2$$

folgt, dass in x_3 ebenfalls ein globales Minimum angenommen wird, in x_4 das globale Maximum.

Aufgabe 13. Die Funktion ist als Summe, Produkt und Verkettung stetiger Funktionen stetig und nimmt auf dem kompakten Intervall $[-2, 2]$ nach Satz 4.3 ihr globales Maximum bzw. Minimum an.

Analog folgt, dass die Funktion auf $(-1, 1)$ für $x \neq 0$ differenzierbar ist.

Anhand von

$$\frac{f(x) - f(0)}{x - 0} = \begin{cases} e^{x-1} & \text{für } x \leq 0, \\ e^{1-x} & \text{für } x > 0 \end{cases}$$

erkennt man, dass die Funktion im Punkt $x_0 = 0$ nicht differenzierbar ist.

Für $x \neq 0$ ist

$$f'(x) = \begin{cases} e^{x-1}(1+x) & \text{für } x < 0, \\ e^{1-x}(1-x) & \text{für } x > 0, \end{cases}$$

d.h. die Ableitung verschwindet in den Punkten $x_1 = -1$ und $x_2 = 1$.

Schließlich berechnet man

$$f''(x) = \begin{cases} e^{x-1}(2+x) & \text{für } x < 0, \\ e^{1-x}(x-2) & \text{für } x > 0. \end{cases}$$

Wegen $f''(x_1) > 0$ liegt dort ein lokales Minimum vor, analog handelt es sich wegen $f''(x_2) < 0$ an der Stelle x_2 um ein lokales Maximum.

Der Vergleich mit den Funktionswerten in den Randpunkten $x_3 = -2$ und $x_4 = 2$ ergibt

$$f(x_1) = -e^{-2} < f(x_3) = -2e^{-3} < f(x_0) = 0 < f(x_4) = 2e^{-1} < f(x_2) = 1.$$

Damit ist das globale Minimum im Punkt $x_1 = -1$ gefunden, das globale Maximum im Punkt $x_2 = 1$.

Es bleibt noch anzumerken, dass im Punkt $x_0 = 0$ kein lokales Minimum vorliegt, da die Funktion im Nullpunkt ihr Vorzeichen wechselt.

Kapitel 6

Integralrechnung in einer Veränderlichen

6.1 Das bestimmte Riemannsches Integral (Zerlegung; Feinheit; Untersumme; Obersumme; Unterintegral; Oberintegral; Riemann integrierbare Funktionen; Integritätskriterium; Rechenregeln für integrierbare Funktionen; Integration auf Teilintervallen; orientiertes Riemannsches Integral; Flächeninhalt)

Im Folgenden ist **stets** $I = [a, b]$ und $f: I \rightarrow \mathbb{R}$ eine **beschränkte Funktion**.

Idee zur Einführung des bestimmten Integrals.

Das **bestimmte Integral** $\int_a^b f(x) dx$ soll den Flächeninhalt der vom Graphen von f und der x -Achse eingeschlossenen Punktmenge beschreiben, so wie es in Abbildung 6.1 angedeutet ist.

Das Problem bei dieser suggestiven Vorstellung ist aber:

Wie ist der Flächeninhalt überhaupt definiert, wenn es sich nicht um ein elementargeometrisches Objekt handelt?

In der Tat sieht die Vorgehensweise umgekehrt aus:

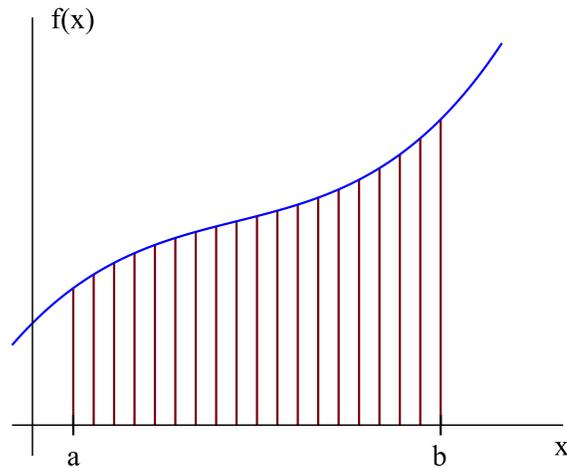


Abbildung 6.1: Zur suggestiven Vorstellung des bestimmten Integrals.

- i)* Man definiert zunächst das bestimmte Integral.
- ii)* Dabei soll „einfachen“ geometrischen Objekten (Rechtecken, Dreiecken etc.) der Flächeninhalt zugeordnet werden, der mit der Elementargeometrie konsistent ist.
- iii)* Um dies zu erreichen, approximiert man bei der Definition des Integrals die von einem Graphen eingeschlossene Menge mit Rechtecken, denen elementargeometrisch ein Flächeninhalt zugeordnet ist.
- iv)* Es ist zu untersuchen, ob zum Grenzwert übergegangen werden kann.
- v)* Mithilfe des bestimmten Integrals können schließlich Flächeninhalte definiert werden, die elementargeometrisch nicht bekannt sind.

Die Definition des bestimmten Integrals

Die oben angesprochene Approximation mit Rechtecken erfolgt via

Definition 6.1. UNTER- UND OBERSUMME

i) Eine *Zerlegung* \mathcal{Z} des Intervalls $I = [a, b]$ in Teilintervalle I_j , $j = 1, \dots, n$, der Länge $|I_j|$ ist eine *Menge von Punkten*

$$\mathcal{Z} = \{a = x_0 < x_1 < x_2 < \dots < x_n = b\}.$$

Hierbei ist $I_j := [x_{j-1}, x_j]$, $\Delta x_j := x_j - x_{j-1} = |I_j|$, $j = 1, \dots, n$, und

$$\Delta(\mathcal{Z}) := \max\{\Delta x_1, \dots, \Delta x_n\}$$

heißt die *Feinheit der Zerlegung* \mathcal{Z} .

ii) Ist für $j = 1, \dots, n$

$$\underline{m}_j := \inf\{f(x) : x \in I_j\},$$

$$\overline{m}_j := \sup\{f(x) : x \in I_j\},$$

so heißt

$$\underline{S}_{\mathcal{Z}}(f) := \sum_{j=1}^n \underline{m}_j \Delta x_j$$

die *Untersumme*,

$$\overline{S}_{\mathcal{Z}}(f) := \sum_{j=1}^n \overline{m}_j \Delta x_j$$

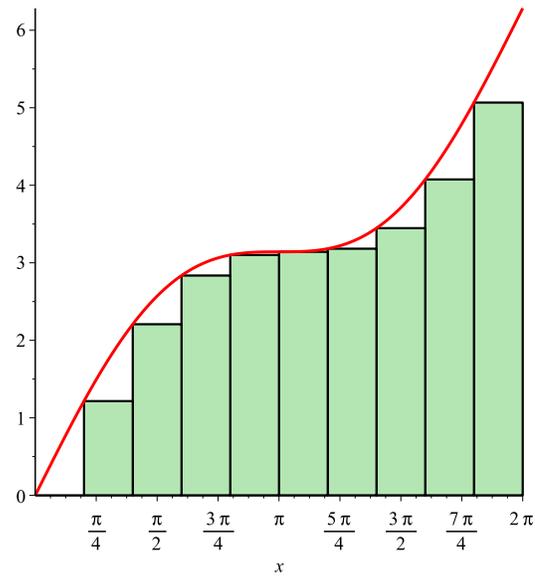
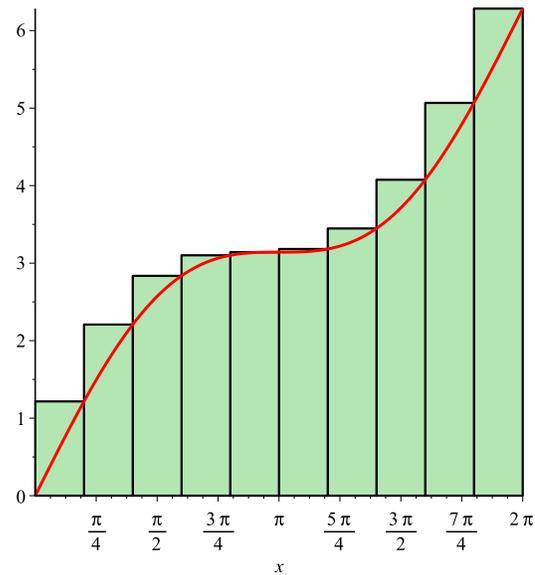
die *Obersumme* von f zur Zerlegung \mathcal{Z} .

In den Abbildungen 6.2 und 6.3 sind eine Unter- und eine Obersumme dargestellt.

Man erkennt für jede Zerlegung als unmittelbare Folgerung aus der Definition

$$\underline{S}_{\mathcal{Z}}(f) \leq \overline{S}_{\mathcal{Z}}(f).$$

Bei einer sukzessiven *Verfeinerung* (vgl. Übungskapitel 6.6) lassen die Abbildungen 6.2 und 6.3 im Falle von „gutartigen Funktionen“ erwarten,

Abbildung 6.2: Eine Untersumme von $f(x) = \sin(x) + x$.Abbildung 6.3: Eine Obersumme von $f(x) = \sin(x) + x$.

dass die Folge der Obersummen von oben und die Folge der Untersummen von unten gegen einen gemeinsamen Grenzwert konvergieren (vgl. wieder Übungskapitel 6.6).

Dies motiviert

Definition 6.2. BESTIMMTES INTEGRAL

Die Funktion $f: I = [a, b] \rightarrow \mathbb{R}$ sei beschränkt.

Dann sind das *Unterintegral* $\underline{\mathcal{I}}$ und das *Oberintegral* $\overline{\mathcal{I}}$ definiert als

$$\underline{\mathcal{I}}(f) := \sup\{\underline{S}_{\mathcal{Z}}(f) : \mathcal{Z} \text{ ist Zerlegung von } I\},$$

$$\overline{\mathcal{I}}(f) := \inf\{\overline{S}_{\mathcal{Z}}(f) : \mathcal{Z} \text{ ist Zerlegung von } I\}.$$

Die Funktion heißt (*Riemann-*) *integrierbar* auf I , falls gilt

$$\underline{\mathcal{I}}(f) = \overline{\mathcal{I}}(f) =: \mathcal{I}(f) =: \int_I f(x) \, dx =: \int_a^b f(x) \, dx.$$

Notation: Die Klasse der auf I integrierbaren Funktionen wird mit $\mathcal{R}(I)$ bezeichnet und $\int_a^b f(x) \, dx$ heißt das *bestimmte Integral* von f zwischen den Grenzen a und b .

Die Funktion f nennt man auch den *Integranden*, x heißt die *Integrationsvariable*.

Der direkte Umgang mit der Definition kann recht schwierig sein.

Im Folgenden erkennt man allerdings, dass die Klasse der auf einem kompakten Intervall $[a, b]$ integrierbaren Funktionen noch deutlich größer ist als die Klasse der auf $[a, b]$ stetigen Funktionen.

Dennoch gibt es nicht integrierbare Funktionen, wie in Übungskapitel 6.6 zu zeigen ist.

Erste Beispiele.

- i) Für eine Konstante $c \in \mathbb{R}$ sei $f(x) = c$ auf $[a, b]$. Es folgt für jede Zerlegung \mathcal{Z} von I

$$\underline{S}_{\mathcal{Z}}(f) = \sum_{j=1}^n c \Delta x_j = c(b-a) = \overline{S}_{\mathcal{Z}}(f).$$

Somit ist f auf $[a, b]$ integrierbar mit

$$\int_a^b c \, dx = c(b-a).$$

- ii) Es sei $I = [0, 1]$ und $f(x) = x$ auf I . Zu $n \in \mathbb{N}$ betrachte man die äquidistante Zerlegung

$$\mathcal{Z}_n = \left\{ 0, \frac{1}{n}, \frac{2}{n}, \dots, 1 \right\}.$$

Für die Ober- bzw. Untersummen gilt

$$\begin{aligned} \underline{S}_{\mathcal{Z}_n}(f) &= \sum_{j=1}^n \frac{j-1}{n} \left[\frac{j}{n} - \frac{j-1}{n} \right] = \frac{1}{n^2} \sum_{j=1}^n (j-1) \\ &= \frac{1}{n^2} \frac{n(n-1)}{2} = \frac{1}{2} - \frac{1}{2n} \xrightarrow{n \rightarrow \infty} \frac{1}{2}, \\ \overline{S}_{\mathcal{Z}_n}(f) &= \frac{1}{2} + \frac{1}{2n} \xrightarrow{n \rightarrow \infty} \frac{1}{2}. \end{aligned}$$

Da hier nur eine spezielle Zerlegungsfolge betrachtet wird, ist die Integrierbarkeit von f nach wie vor unklar.

Man benötigt noch:

Satz 6.1. INTEGRABILITÄTSKRITERIUM

Für eine beschränkte Funktion $f: I \rightarrow \mathbb{R}$ gilt

$$f \in \mathcal{R}(I) \Leftrightarrow \begin{cases} \text{Zu jedem } \varepsilon > 0 \text{ gibt es eine Zerlegung } \mathcal{Z} \text{ von } I \text{ mit} \\ \overline{S}_{\mathcal{Z}}(f) - \underline{S}_{\mathcal{Z}}(f) < \varepsilon. \end{cases}$$

Konsequenz. Die Funktion $f(x) = x$ ist auf $I = [0, 1]$ integrierbar mit

$$\int_0^1 x \, dx = \frac{1}{2}.$$

Rechenregeln für integrierbare Funktionen.

Die folgenden Rechenregeln erleichtern die Suche nach weiteren integrierbaren Funktionen erheblich. Insbesondere folgt beispielsweise die Integrierbarkeit von Polynomen auf $[a, b]$.

Man beachte allerdings, dass im Gegensatz zu den Rechenregeln für differenzierbare Funktionen (Satz 5.2) keine konstruktive Vorschrift zur Berechnung des Integrals gegeben wird.

Satz 6.2. RECHENREGELN ZUR INTEGRIERBARKEIT

Es seien $f, g \in \mathcal{R}(I)$ und $\alpha, \beta \in \mathbb{R}$.

Dann gilt:

i) $\alpha f + \beta g \in \mathcal{R}(I)$ mit

$$\mathcal{I}(\alpha f + \beta g) = \alpha \mathcal{I}(f) + \beta \mathcal{I}(g).$$

ii) $fg \in \mathcal{R}(I)$.

iii) $|f| \in \mathcal{R}(I)$, wobei $|f|(x) := |f(x)|$.

iv) Ist $|g| \geq c$ für eine Konstante $c > 0$, so ist auch $f/g \in \mathcal{R}(I)$.

v) Ist für alle $x \in I$ die Ungleichung $f(x) \leq g(x)$ richtig, so folgt

$$\mathcal{I}(f) \leq \mathcal{I}(g) .$$

vi) $|\mathcal{I}(f)| \leq \mathcal{I}(|f|)$.

Beweis. Der Beweis folgt unmittelbar aus der Definition des Riemannsches Integrals. □

Weitere Klassen integrierbarer Funktionen.

Satz 6.3.

WEITERE KLASSEN INTEGRIERBARER FUNKTIONEN

Für eine beschränkte Funktion $f: [a, b] \rightarrow \mathbb{R}$ gilt:

i) Ist f *monoton*, so folgt $f \in \mathcal{R}(I)$.

ii) Ist f *stetig*, so folgt $f \in \mathcal{R}(I)$.

In der Tat sind sogar *stückweise stetige Funktionen* integrierbar.

Dabei ist der Begriff „stückweise stetig“ in Abbildung 6.4 illustriert. Man beachte, dass die einseitigen Grenzwerte in den Sprungstellen und am Rande des Intervalls existieren müssen.

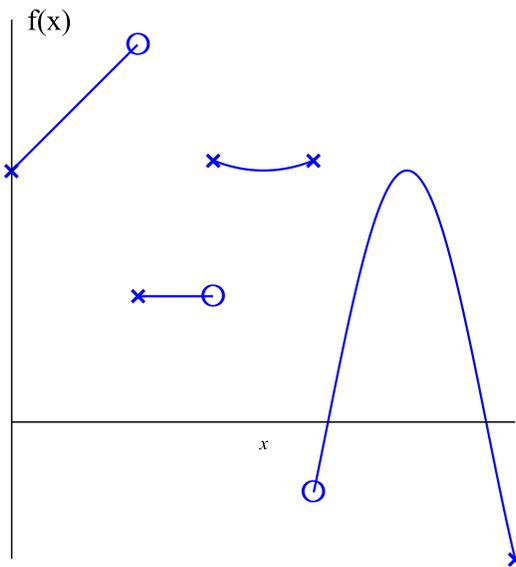


Abbildung 6.4: Eine stückweise stetige Funktion.

Abschließende Bemerkungen zur Definition.

- i) Eine Funktion $f \in \mathcal{R}(I)$ kann auch über Teilintervalle integriert werden im Sinne von:
- (a) f ist auch auf jedem Teilintervall $I' \subset I$ integrierbar.
 - (b) Ist I in endlich viele Teilintervalle I_1, I_2, \dots, I_n zerlegt, die höchstens Randpunkte gemeinsam haben, so gilt

$$\int_I f(x) \, dx = \sum_{j=1}^n \int_{I_j} f(x) \, dx .$$

- ii) Das Integral $\int_a^b f(x) \, dx$ ist bisher nur auf Intervallen $[a, b]$, d.h. für $a < b$ definiert.

Das **orientierte Integral** ist in Erweiterung des obigen Integralbegriffs für $a = b$ als $\int_a^b f(x) \, dx := 0$ definiert, im Fall $b < a$ setzt man

$$\int_a^b f(x) \, dx := - \int_b^a f(x) \, dx .$$

- iii) Ist f stetig auf $I = [a, b]$ und ist $f(x) > 0$ für alle $x \in I$, so ist

$$A := \int_a^b f(x) \, dx$$

positiv und **definiert** den **Flächeninhalt** der Menge

$$F := \{\mathbf{x} \in \mathbb{R}^2 : a \leq x_1 \leq b, 0 \leq x_2 \leq f(x_1)\}$$

(vgl. Abbildung 6.1).

6.2 Der Hauptsatz der Differential- und Integralrechnung (Stammfunktion; unbestimmtes Integral)

Der **Hauptsatz der Differential- und Integralrechnung** ist das Bindeglied zwischen diesen beiden Disziplinen und gleichzeitig ein starkes Werkzeug, um Integrale analytisch zu berechnen.

Die Idee des Satzes sei hier anhand eines einfachen Beispiels erläutert:

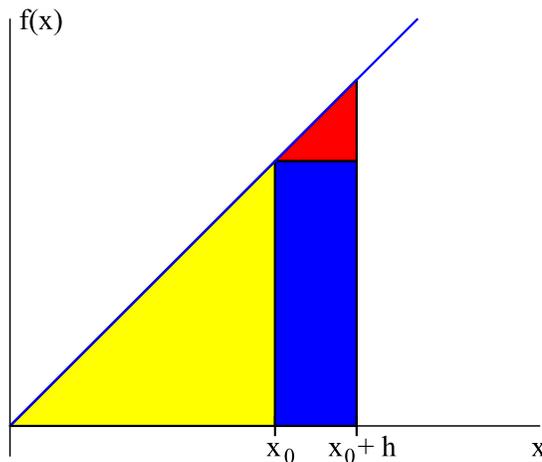


Abbildung 6.5: Zum Hauptsatz der Differential- und Integralrechnung.

Man betrachte für $x \geq 0$ die Funktion $f(x) = \alpha x$, $\alpha > 0$, man betrachte den Flächeninhalt ($x_0 > 0$)

$$A(x_0) = \int_0^{x_0} f(x) \, dx$$

und ebenso für $h > 0$

$$A(x_0 + h) = \int_0^{x_0+h} f(x) \, dx .$$

Nun wird die relative Änderung “ $\Delta A/\Delta x$ ” betrachtet, d.h. gesucht ist der Differenzenquotient von A an der Stelle x_0 ,

$$\frac{1}{h} \left[\int_0^{x_0+h} f(x) \, dx - \int_0^{x_0} f(x) \, dx \right].$$

Die Differenz $A(x_0+h) - A(x_0)$ setzt sich zusammen aus dem Flächeninhalt des in Abbildung 6.5 blau gekennzeichneten Rechtecks R und dem des roten Dreiecks D .

Der Flächeninhalt von R ist $\alpha x_0 h$, der von D ist $\alpha h^2/2$.

Es ergibt sich

$$\frac{1}{h} \left[\int_0^{x_0+h} f(x) \, dx - \int_0^{x_0} f(x) \, dx \right] = \frac{1}{h} \left[\alpha x_0 h + \frac{\alpha h^2}{2} \right]$$

$$\xrightarrow{h \rightarrow 0} \alpha x_0 = f(x_0).$$

Die Ableitung des Flächeninhalts $A(x_0)$ nach x_0 entspricht dem Funktionswert $f(x_0)$.

Präzise Formulierung.

Definition 6.3. STAMMFUNKTION

Betrachtet seien zwei Funktionen $F, f: [a, b] \rightarrow \mathbb{R}$.

Ist F stetig auf $[a, b]$, differenzierbar auf (a, b) und gilt $F'(x) = f(x)$ für alle $x \in (a, b)$, so heißt F eine *Stammfunktion* von f .

Stammfunktionen sind nicht eindeutig bestimmt, da mit F für eine beliebige Konstante $C \in \mathbb{R}$ auch die Funktion $F_C = F + C$ eine Stammfunktionen von f ist.

Dies ist aber die einzige Form der Mehrdeutigkeit: Nach Korollar 5.1 ist die Differenz zweier Stammfunktionen immer konstant.¹

Satz 6.4. HAUPTSATZ

Es sei $f: I = [a, b] \rightarrow \mathbb{R}$ eine stetige Funktion.

Dann gelten die beiden folgenden Aussagen:

i) Die Funktion

$$F(x) := \int_a^x f(t) \, dt$$

ist eine Stammfunktion von $f(x)$.

ii) Ist umgekehrt F eine Stammfunktion von f , so ist

$$\int_a^b f(t) \, dt = F(b) - F(a) =: \left[F(x) \right]_a^b =: F(x) \Big|_a^b .$$

Bezeichnung: Die Gesamtheit aller Stammfunktionen einer stetigen Funktion f heißt das *unbestimmte Integral* der Funktion f .

Notation:

$$\{F : F \text{ ist Stammfunktion von } f\} =: \int f(x) \, dx =: F(x) + C .$$

Sprechweise: Das Aufsuchen einer Stammfunktion wird als *Integration* von f bezeichnet. (Nach dem Hauptsatz “macht die Differentiation die Integration wieder rückgängig”.)

Beweis. Ad i). Zu $h \neq 0$, $x, x + h \in (a, b)$ betrachtet man den Differenzenquotienten

$$\Delta_h F(x) := \frac{1}{h} [F(x + h) - F(x)] , \quad F(x) = \int_a^x f(t) \, dt .$$

¹Man beachte, dass hier als Definitionsbereich immer ein Intervall gewählt ist.

Zu zeigen ist für das fixierte x :

$$\Delta_h F(x) \xrightarrow{h \rightarrow 0} f(x) .$$

In der Tat gilt

$$\Delta_h F(x) = \frac{1}{h} \int_x^{x+h} f(t) dt ,$$

$$f(x) = \frac{1}{h} \int_x^{x+h} f(x) dt .$$

In der zweiten Gleichung ist zu beachten, dass $f(x)$ für fixiertes x eine Konstante ist und bzgl. der Integrationsvariablen t integriert wird.

Es folgt

$$\begin{aligned} |\Delta_h F(x) - f(x)| &\leq \frac{1}{|h|} \left| \int_x^{x+h} (f(t) - f(x)) dt \right| \\ &\leq \frac{1}{|h|} |h| \sup \{ |f(t) - f(x)| : \\ &\quad t \in [a, b], |x - t| \leq |h| \} \\ &\xrightarrow{h \rightarrow 0} 0 , \end{aligned}$$

wobei im letzten Schritt die Stetigkeit von f auf dem kompakten Intervall $[a, b]$ ausgenutzt wurde. Damit ist der erste Teil des Satzes gezeigt.

Ad *ii*). Da die Differenz zweier Stammfunktionen konstant ist, kann nach *i*) jede Stammfunktion F von f in der Form

$$F(x) = \int_a^x f(t) dt + C , \quad C \in \mathbb{R} .$$

geschrieben werden.

Insbesondere ist

$$F(b) = \int_a^b f(t) dt + C ,$$

$$F(a) = \int_a^a f(t) dt + C = C ,$$

d.h.

$$F(b) - F(a) = \int_a^b f(t) dt + C - C = \int_a^b f(t) dt ,$$

und der Hauptsatz ist bewiesen. □

$f(x)$	$\int f(x) dx$	gültig, falls
x^k	$\frac{1}{k+1}x^{k+1} + C$	$k \in \mathbb{Z}, k \neq -1, x \in \mathbb{R}$
$\frac{1}{x}$	$\ln(x) + C$	$x \neq 0$
x^α	$\frac{1}{\alpha+1}x^{\alpha+1} + C$	$\alpha \neq -1, x > 0$
e^x	$e^x + C$	$x \in \mathbb{R}$
a^x	$\frac{a^x}{\ln(a)} + C$	$x \in \mathbb{R}, 0 < a, a \neq 1$
$\sin(x)$	$-\cos(x) + C$	$x \in \mathbb{R}$
$\cos(x)$	$\sin(x) + C$	$x \in \mathbb{R}$
$\tan(x)$	$-\ln(\cos(x)) + C$	$x \neq (2k+1)\pi/2, k \in \mathbb{Z}$
$\cot(x)$	$\ln(\sin(x)) + C$	$x \neq k\pi, k \in \mathbb{Z}$
$\frac{1}{\cos^2(x)}$	$\tan(x) + C$	$x \neq (2k+1)\pi/2, k \in \mathbb{Z}$
$\frac{1}{\sin^2(x)}$	$-\cot(x) + C$	$x \neq k\pi, k \in \mathbb{Z}$
$\frac{1}{\sqrt{1-x^2}}$	$\arcsin(x) + C$	$-1 < x < 1$
$\frac{1}{\sqrt{1-x^2}}$	$\arccos(x) + C$	$-1 < x < 1$
$\frac{1}{1+x^2}$	$\arctan(x) + C$	$x \in \mathbb{R}$

Tabelle 6.1: Einige unbestimmte Integrale.

Beispiele.

i) Als ein einfaches Beispiel sei hier die Funktion $f(x) = x^n, n \in \mathbb{N}$,

betrachtet. Diese hat als eine Stammfunktion

$$F(x) = \frac{x^{n+1}}{n+1},$$

woraus

$$\int_a^b x^n dx = \frac{1}{n+1} [b^{n+1} - a^{n+1}]$$

folgt.

ii) Ohne Beweis sind in Tabelle 6.3 einige unbestimmte Integrale aufgelistet.

6.3 Integrationstechniken (einfache Integrationstechniken; partielle Integration; Substitutionsregel; Partialbruchzerlegung)

Im Gegensatz zur Differentiation ist eine geeignete Vorgehensweise bei der Integration oft nicht ersichtlich – manche Integrale können analytisch überhaupt nicht berechnet werden.

Dennoch gibt es einige Werkzeuge zur Integration bzw. zur Umformung von Integralen.

Die wesentlichen Techniken werden in diesem Paragraphen vorgestellt.

6.3.1 Einfache Integrationstechniken

Es seien f, g stetig auf $[a, b]$. Dann gilt (auf $[a, b]$)

$$i) \int cf(x) dx = c \int f(x) dx, \quad c \in \mathbb{R},$$

$$ii) \int (f(x) + g(x)) dx = \int f(x) dx + \int g(x) dx,$$

$$iii) \int \frac{f'(x)}{f(x)} dx = \ln(|f(x)|) + C, \text{ falls } f \text{ keine Nullstellen hat,}$$

wobei die Funktion f in der letzten Zeile zusätzlich von der Klasse C^1 sei.

Beispiele.

i) Mit $a_j \in \mathbb{R}$, $j = 1, \dots, n$, ist

$$\int \left[\sum_{j=0}^n a_j x^j \right] dx = \sum_{j=0}^n a_j \int x^j dx = \sum_{j=0}^n a_j \frac{x^{j+1}}{j+1} + C .$$

ii) Es sei $f(x) = \tan(x)$, $x \in [a, b] \subset (-\pi/2, \pi/2)$. In diesem Intervall hat der Kosinus keine Nullstelle und wegen $\cos(x) > 0$ für alle $x \in (-\pi/2, \pi/2)$ gilt

$$\begin{aligned} \int \tan(x) dx &= \int \frac{\sin(x)}{\cos(x)} dx = - \int \frac{(-\sin(x))}{\cos(x)} dx \\ &= -\ln(\cos(x)) + C . \end{aligned}$$

6.3.2 Partielle Integration

Hierbei handelt es sich um die Folgerung aus der Produktregel.

Satz 6.5. PARTIELLE INTEGRATION

Es seien $f, g: [a, b] \rightarrow \mathbb{R}$ von der Klasse C^1 .

Dann gilt für das unbestimmte Integral

$$\int f(x)g'(x) dx = f(x)g(x) - \int f'(x)g(x) dx$$

bzw. für das bestimmte Integral

$$\int_a^b f(x)g'(x) dx = \left[f(x)g(x) \right]_a^b - \int_a^b f'(x)g(x) dx .$$

Beispiele.

i) Mit $f(x) = x$ und $g'(x) = e^x$ folgt

$$\int x e^x dx = x e^x - \int e^x dx = x e^x - e^x + C .$$

ii) Mit $f(x) = \ln(x)$ und $g'(x) = 1$ ergibt sich auf $[a, b] \subset (0, \infty)$

$$\begin{aligned} \int \ln(x) \, dx &= \int \ln(x) 1 \, dx = \ln(x)x - \int \frac{1}{x} x \, dx \\ &= x \ln(x) - x + C . \end{aligned}$$

iii) Ist $f(x) = g'(x) = \cos(x)$, so sieht man

$$\begin{aligned} \int \cos^2(x) \, dx &= \int \cos(x) \cos(x) \, dx \\ &= \cos(x) \sin(x) + \int \sin^2(x) \, dx \\ &= \sin(x) \cos(x) + \int 1 \, dx - \int \cos^2(x) \, dx , \end{aligned}$$

also

$$\int \cos^2(x) \, dx = \frac{1}{2} [\sin(x) \cos(x) + x] + C .$$

6.3.3 Substitutionsregel

Die Regel zur Substitution ist eine Folgerung aus der Kettenregel.

Satz 6.6. SUBSTITUTIONSREGEL

Es sei $f: [a, b] \rightarrow \mathbb{R}$ stetig und $g: [c, d] \rightarrow [a, b]$ sei von der Klasse C^1 .

Dann gilt für alle $t_0, t_1 \in [c, d]$

$$\int_{g(t_0)}^{g(t_1)} f(x) \, dx = \int_{t_0}^{t_1} f(g(t))g'(t) \, dt .$$

Für die Substitutionsregel gibt es zwei verschiedene Arten der Anwendung.

Erste Anwendungsvariante.

Gesucht sei $\int_{t_0}^{t_1} h(t) dt$, wobei der Integrand h die spezielle Struktur

$$h(t) = f(g(t))g'(t)$$

habe.

Mit anderen Worten erkennt man (bis auf evtl. konstante Faktoren) das **Produkt aus einer Verkettung von Funktionen und der inneren Ableitung**

.

Dann gilt

$$\int_{t_0}^{t_1} h(t) dt = \int_{g(t_0)}^{g(t_1)} f(x) dx .$$

Kann hier die rechte Seite explizit berechnet werden, so ist auch die linke Seite bekannt.

Beispiel. Gesucht sei das bestimmte Integral

$$\int_0^2 te^{1+t^2} dt .$$

Die Funktion $h(t) := te^{1+t^2}$ lässt sich in der Form

$$h(t) = \frac{1}{2}e^{1+t^2} \frac{d}{dt}(1+t^2)$$

schreiben, also mit $f(x) := e^x$ und $g(t) := 1+t^2$:

$$2h(t) = f(g(t))g'(t)$$

Es folgt

$$\begin{aligned} \int_0^2 te^{1+t^2} dt &= \frac{1}{2} \int_0^2 f(g(t))g'(t) dt \\ &= \frac{1}{2} \int_1^5 e^x dx = \frac{e^5 - e}{2} . \end{aligned}$$

Zweite Anwendungsvariante.

In dieser Variante versucht man, $\int_{x_0}^{x_1} f(x) dx$ durch eine geeignete Transformation der Integrationsvariablen zu berechnen, d.h. man sucht eine Transformation

$$x = g(t), \quad t_0 \leq t \leq t_1,$$

mit der Eigenschaft

$$x_0 = g(t_0) \quad \text{und} \quad x_1 = g(t_1).$$

Hierbei sei $g: [t_0, t_1] \rightarrow [x_0, x_1]$ (o.E. gelte $t_0 < t_1$ und $x_0 < x_1$) bijektiv (etwa $g'(t) > 0$ für alle $t \in [t_0, t_1]$) und von der Klasse C^1 .

Bezeichnet dann $\psi := g^{-1}$ die Inverse von g , so gilt mit $t_0 = \psi(x_0)$ und $t_1 = \psi(x_1)$

$$\int_{x_0}^{x_1} f(x) dx = \int_{\psi(x_0)}^{\psi(x_1)} f(g(t))g'(t) dt.$$

Bemerkung. Das **formale Vorgehen** sieht wie folgt aus:

Man substituiert $x = g(t)$ in $\int_{x_0}^{x_1} f(x) dx$ und schreibt in der **Leibnizschen** Weise

$$g'(t) = \frac{dx}{dt}, \quad \text{also} \quad dx = g'(t) dt.$$

Die untere Integrationsgrenze $x_0 = g(t_0)$ wird zu $t_0 = \psi(x_0)$, die obere transformiert sich analog.

Beispiel. Gesucht sei zu fixiertem $r > 0$

$$\int_0^r \sqrt{r^2 - x^2} dx.$$

Wegen $r^2 - r^2 \sin^2(x) = r^2 \cos^2(x)$ bietet sich die Substitution

$$x = g(t) := r \sin(t), \quad 0 \leq t \leq \frac{\pi}{2},$$

an, wobei $g'(t) = r \cos(t)$. Man beachtet ($\psi = g^{-1}$)

$$0 = g(0), \quad r = g(\pi/2), \quad \text{d.h.} \quad \psi(0) = 0, \quad \psi(r) = \frac{\pi}{2}.$$

Mit der Substitutionsregel berechnet man

$$\begin{aligned} \int_0^r \sqrt{r^2 - x^2} \, dx &= \int_0^{\pi/2} \sqrt{r^2 - r^2 \sin^2(t)} \, r \cos(t) \, dt \\ &= r^2 \int_0^{\pi/2} \cos^2(t) \, dt = \frac{\pi r^2}{4}. \end{aligned}$$

6.3.4 Partialbruchzerlegung

Mithilfe einer sogenannten [Partialbruchzerlegung](#) können gebrochenrationale Funktionen (vgl. Teil I, Kapitel 5.1)

$$f(x) = \frac{p(x)}{q(x)}, \quad p, q \text{ Polynome vom Grad } m \text{ bzw. } n,$$

integriert werden.

Die präzise Darstellung aller möglichen Fälle erfordert eine etwas technische Notation.

Zum prinzipiellen Verständnis genügt es aber bereits, sich von den folgenden Beispielen leiten zu lassen.

Falls nötig: Polynomdivision.

Wie in Teil I, Kapitel 5.1, kann ohne Beschränkung der Allgemeinheit

$$\text{grad } p < \text{grad } q,$$

angenommen werden.

Ansonsten wird eine [Polynomdivision](#) vorgeschaltet.

Beispiel. Ist

$$f(x) = \frac{x^3 + 3x^2 - 4x + 2}{x^3 - x^2 - x + 1},$$

so liefert die Rechnung

$$f(x) = \underbrace{1}_{r(x)} + \frac{4x^2 - 3x + 1}{\underbrace{x^3 - x^2 - x + 1}_{p(x)/q(x)}},$$

es ist also

$$f(x) = r(x) + \frac{p(x)}{q(x)},$$

wobei das Polynom $r(x)$ elementar zu integrieren ist und wobei $\text{grad } p < \text{grad } q$.

Nullstellen von q mit Vielfachheit 1.

Beispiel. Es sei

$$\frac{p(x)}{q(x)} = \frac{x}{2x^2 - 6x + 4}.$$

Zunächst sucht man die Nullstellen x_1, x_2 des Nennerpolynoms q : $x_1 = 1$, $x_2 = 2$, in denen die Funktion nicht definiert ist.

Beide Nullstellen sind verschieden und haben die **Vielfachheit 1**, d.h. $q(x)$ kann als

$$q(x) = 2(x - 1)^1(x - 2)^1 = 2(x - 1)(x - 2)$$

in **Linearfaktoren** zerlegt werden.

In dieser Situation macht man den **Ansatz**

$$\frac{p(x)}{q(x)} = \frac{A}{x - 1} + \frac{B}{x + 2},$$

wobei A und B reelle Konstanten bezeichnen.

Der Ansatz liefert

$$\begin{aligned} \frac{p(x)}{q(x)} &= \frac{x}{2(x-1)(x-2)} \\ &= \frac{A}{x-1} + \frac{B}{x+2} \\ &= \frac{A(x-2) + B(x-1)}{(x-1)(x-2)} \\ &= \frac{x(A+B) + (-2A-B)}{(x-1)(x-2)}. \end{aligned}$$

Ein [Koeffizientenvergleich](#) (man vergleicht die Koeffizienten zu allen auftretenden Potenzen von x) zeigt

$$\begin{aligned} A + B &= \frac{1}{2}, \\ -2A - B &= 0, \end{aligned}$$

mit der Lösung (Probe!)

$$A = -\frac{1}{2}, \quad B = 1.$$

Insgesamt ist

$$\begin{aligned} \int \frac{x}{2x^2 - 6x + 4} dx &= -\frac{1}{2} \int \frac{1}{x-1} dx + \int \frac{1}{x-2} dx \\ &= -\frac{1}{2} \ln|x-1| + \ln(|x-2|) + C. \end{aligned}$$

Dabei ist im unbestimmten Integral als Gesamtheit der Stammfunktionen wie immer die Integrationskonstante $C \in \mathbb{R}$ zu berücksichtigen und die Stammfunktionen sind in den Polstellen von f nicht definiert.

Nullstellen von q mit Vielfachheit 2 und größer.

Beispiel. Es sei

$$\frac{p(x)}{q(x)} = \frac{x^2 + 1}{x^3 - 2x^2 + x} .$$

Zunächst werden wieder die Nullstellen x_1, x_2, x_3 des Nennerpolynoms q gesucht: $x_1 = 0, x_2 = 1, x_3 = 1$.

Die Nullstelle $x_1 = 0$ hat die Vielfachheit **1**. Die Nullstellen $x_2 = x_3 = 1$ sind gleich, d.h. 1 ist Nullstelle der Vielfachheit **2** und $q(x)$ wird gemäß

$$q(x) = x^1(x - 1)^2 = x(x - 1)^2$$

in Linearfaktoren zerlegt.

Nun macht man den Ansatz (A, B, C reelle Konstanten)

$$\frac{p(x)}{q(x)} = \frac{A}{x} + \frac{B}{x - 1} + \frac{C}{(x - 1)^2} .$$

Dieser liefert

$$\begin{aligned} \frac{p(x)}{q(x)} &= \frac{x^2 + 1}{x(x - 1)^2} \\ &= \frac{A}{x} + \frac{B}{x - 1} + \frac{C}{(x - 1)^2} \\ &= \frac{A(x - 1)^2 + Bx(x - 1) + Cx}{x(x - 1)^2} \\ &= \frac{x^2(A + B) + x(-2A - B + C) + A}{x(x - 1)^2} . \end{aligned}$$

Aus einem Koeffizientenvergleich folgt

$$\begin{aligned} A + B &= 1 , \\ -2A - B + C &= 0 , \\ A &= 1 \end{aligned}$$

mit der Lösung (Probe!)

$$A = 1, \quad B = 0, \quad C = 2.$$

Insgesamt ist

$$\int \frac{x^2 + 1}{x^3 - 2x^2 + x} dx = \ln |x| - \frac{2}{x - 1} + C.$$

Irreduzible quadratische Polynome.

Beispiel. Es sei

$$\frac{p(x)}{q(x)} = \frac{1}{x^3 - x^2 + x - 1}.$$

Bei der Suche nach Nullstellen des Nennerpolynoms findet man nur die eine Nullstelle $x_1 = 1$ der Vielfachheit 1.

Das Polynom **zerfällt nicht in Linearfaktoren** und kann lediglich in der Form

$$q(x) = (x - 1)(x^2 + 1)$$

zerlegt werden, wobei $x^2 + 1$ **irreduzibel** ist, d.h. keine reelle Nullstellen hat.

Der Ansatz lautet in diesem Fall

$$\frac{p(x)}{q(x)} = \frac{A}{x - 1} + \frac{B + Cx}{x^2 + 1}$$

und führt auf

$$\begin{aligned} \frac{p(x)}{q(x)} &= \frac{1}{(x-1)(x^2+1)} \\ &= \frac{A}{x-1} + \frac{B+Cx}{x^2+1} \\ &= \frac{A(x^2+1) + B(x-1) + Cx(x-1)}{(x-1)(x^2+1)} \\ &= \frac{x^2(A+C) + x(B-C) + (A-B)}{(x-1)(x^2+1)}. \end{aligned}$$

Hier ergibt ein Koeffizientenvergleich

$$\begin{aligned} A + C &= 0, \\ B - C &= 0, \\ A - B &= 1 \end{aligned}$$

mit der Lösung (Probe!)

$$A = \frac{1}{2}, \quad B = C = -\frac{1}{2}.$$

Man berechnet

$$\int \frac{1}{x^3 - x^2 + x - 1} dx = \frac{1}{2} \ln|x-1| - \frac{1}{2} \arctan(x) - \frac{1}{4} \ln(|x^2+1|) + C.$$

6.4 Uneigentliche Integrale (lokal integrierbare Funktionen; Konvergenzkriterien)

In der Definition 6.2 des bestimmten Integrals sind zwei verschiedene Arten der **Beschränktheit** vorausgesetzt:

- i) Das Intervall, über welches integriert wird, ist beschränkt.
- ii) Der Integrand f ist eine beschränkte Funktion.

Nun wird der Frage nachgegangen, ob bzw. unter welchen Voraussetzungen man auf diese Einschränkungen verzichten kann.

Das typische Beispiel dieses Kapitels ist das Integral

$$\int_I \frac{dx}{x^\alpha}, \quad \alpha \in \mathbb{R}^+ \text{ fixiert.}$$

Hier bezeichnet I ein evtl. verallgemeinertes Intervall.

Ist etwa $I = [1, \infty)$, so ist über ein Intervall „unendlicher Länge“ zu integrieren und es ist unklar, ob das Integral existiert.

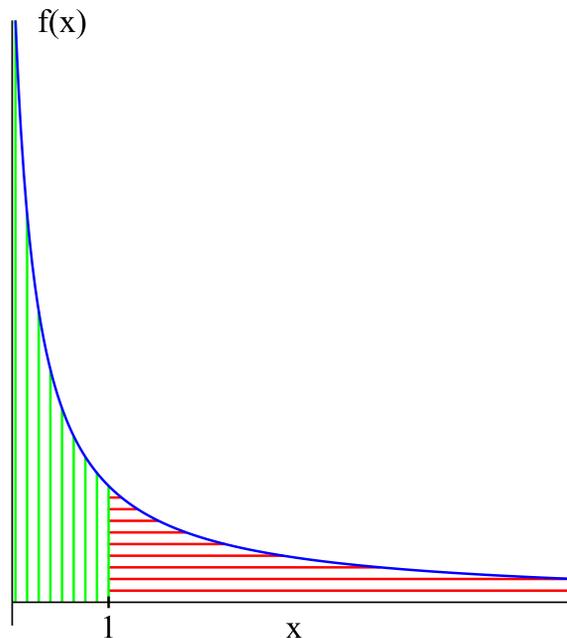


Abbildung 6.6: Unbeschränktes Intervall versus unbeschränkte Funktion.

Ist beispielsweise $I = (0, 1]$, so ist zwar das Integrationsintervall beschränkt, die Funktion wird jedoch in der Nähe des Nullpunkts beliebig groß. Damit ist ebenso unklar, ob das Integral existiert.

Geometrisch lautet die Frage, ob der in Abbildung 6.6 grün bzw. rot schraffierte Flächeninhalt endlich ist oder nicht.

Unbeschränktes Integrationsintervall.

Es sei $I = [a, \infty)$ und $f \in \mathcal{R}([a, b])$ für alle b mit $a < b < \infty$.

Man nennt die Funktion f in diesem Fall **lokal integrierbar**.

Mit anderen Worten: Für alle $b \in I$ existiere das bestimmte Integral

$$\int_a^b f(x) \, dx .$$

Definition 6.4. UNEIGENTLICHES INTEGRAL

Es sei f wie oben definiert lokal integrierbar.

i) Falls der Grenzwert

$$\lim_{b \rightarrow \infty} \int_a^b f(x) \, dx$$

existiert, so heißt dieser das **uneigentliche Integral** von f über $[a, \infty)$.

Sprechweise: Das uneigentliche Integral existiert oder **konvergiert**.

Notation:

$$\int_a^\infty f(x) \, dx := \lim_{b \rightarrow \infty} \int_a^b f(x) \, dx .$$

Andernfalls heißt $\int_a^\infty f(x) \, dx$ **divergent**.

ii) Das uneigentliche Integral $\int_a^\infty f(x) \, dx$ heißt **absolut konvergent**, falls $\int_a^\infty |f(x)| \, dx$ konvergiert.

Absolute Konvergenz impliziert Konvergenz des uneigentlichen Integrals.

Bemerkung. Uneigentliche Integrale der Form

$$\int_{-\infty}^b f(x) \, dx$$

sind analog zu diskutieren.

Cauchyscher Hauptwert.

Bei einem uneigentliches Integral der Form

$$\int_{-\infty}^{\infty} f(x) \, dx$$

ist jedoch Vorsicht geboten:

Das Integral konvergiert per definitionem, falls für ein beliebiges $a \in \mathbb{R}$ sowohl $\int_a^{\infty} f(x) \, dx$ als auch $\int_{-\infty}^a f(x) \, dx$ konvergiert:

$$\int_{-\infty}^{\infty} f(x) \, dx := \int_a^{\infty} f(x) \, dx + \int_{-\infty}^a f(x) \, dx .$$

Existiert eines der Integrale $\int_a^{\infty} f(x) \, dx$, $\int_{-\infty}^a f(x) \, dx$ nicht, so kann trotzdem der sogenannte **Cauchysche Hauptwert** existieren (vgl. Übungskapitel 6.6).

Das Standardbeispiel.

Es ist für alle $1 < b < \infty$

$$\begin{aligned} \int_1^b \frac{dx}{x^\alpha} &= \begin{cases} \frac{1}{1-\alpha} \frac{1}{x^{\alpha-1}} \Big|_1^b & \text{falls } \alpha \neq 1, \\ \ln(|x|) \Big|_1^b & \text{falls } \alpha = 1, \end{cases} \\ &= \begin{cases} \frac{b^{1-\alpha} - 1}{1-\alpha} & \text{falls } \alpha \neq 1, \\ \ln(b) & \text{falls } \alpha = 1. \end{cases} \end{aligned}$$

Demnach ist

$$\int_1^{\infty} \frac{dx}{x^\alpha} \quad \text{divergent für } \alpha \leq 1 .$$

Es ist aber

$$\int_1^{\infty} \frac{dx}{x^\alpha} = \frac{1}{\alpha - 1} \quad \text{für } \alpha > 1 .$$

Konvergenzkriterien.

In der Regel kann jedoch nicht wie im obigen Beispiel explizit gerechnet werden und man benötigt **Konvergenzkriterien** um zu entscheiden, ob ein uneigentliches Integral existiert oder nicht.

Satz 6.7. KONVERGENZKRITERIEN

Es sei $f: I = [a, \infty) \rightarrow \mathbb{R}$ lokal integrierbar.

- i) Das uneigentliche Integral $\int_a^\infty f(x) dx$ existiert **genau dann**, wenn zu jedem $\varepsilon > 0$ ein $\xi \geq a$ existiert mit

$$\left| \int_b^{b'} f(x) dx \right| < \varepsilon \quad \text{für alle } b, b' > \xi .$$

- ii) **Majorantenkriterium:** Ist für alle $x \in I$ die Ungleichung $|f(x)| \leq g(x)$ richtig und ist $\int_a^\infty g(x) dx$ konvergent, so ist $\int_a^\infty f(x) dx$ absolut konvergent.
- iii) **Minorantenkriterium:** Gilt umgekehrt $0 \leq g(x) \leq f(x)$ für alle $x \in I$ und divergiert das uneigentliche Integral $\int_a^\infty g(x) dx$, so divergiert auch $\int_a^\infty f(x) dx$.

Beispiele.

i) Man betrachte

$$\int_2^{\infty} \frac{\sqrt{1+x^2}}{x^3} dx .$$

Für $x \in [2, \infty)$ ist $1 < x^2$, d.h.

$$|\sqrt{1+x^2}| = \sqrt{1+x^2} \leq \sqrt{2x^2} \leq \sqrt{2}x .$$

Wegen der Konvergenz von (vgl. obiges Standardbeispiel)

$$\int_1^{\infty} \frac{dx}{x^2}$$

ist auch

$$\int_2^{\infty} \frac{\sqrt{2}}{x^2} dx$$

konvergent und das Majorantenkriterium belegt die Konvergenz von

$$\int_2^{\infty} \frac{\sqrt{1+x^2}}{x^3} dx .$$

ii) Betrachtet sei das uneigentliche Integral

$$\int_0^{\infty} \frac{\sin(x)}{x} dx ,$$

wobei es im Nullpunkt keine Schwierigkeiten gibt, da $\sin(x)/x$ bekanntlich stetig in 0 fortgesetzt werden kann.

Dieses uneigentliche Integral ist **konvergent, aber nicht absolut konvergent**:

Zum Konvergenzbeweis beachtet man, dass für beliebige $0 < b < b'$ gilt (partielle Integration)

$$\int_b^{b'} \frac{\sin(x)}{x} dx = -\frac{\cos(x)}{x} \Big|_b^{b'} - \int_b^{b'} \frac{\cos(x)}{x^2} dx .$$

Somit kann abgeschätzt werden

$$\begin{aligned} \left| \int_b^{b'} \frac{\sin(x)}{x} dx \right| &\leq \frac{1}{b} + \frac{1}{b'} + \int_b^{b'} \frac{1}{x^2} dx \\ &< \frac{2}{b} \xrightarrow{b \rightarrow \infty} 0. \end{aligned}$$

Nach dem ersten Konvergenzkriterium aus Satz 6.7 konvergiert das Integral.²

Es gilt jedoch für alle $k \in \mathbb{N}$:

$$\begin{aligned} \int_0^{k\pi} \left| \frac{\sin(x)}{x} \right| dx &= \sum_{j=1}^k \int_{(j-1)\pi}^{j\pi} \left| \frac{\sin(x)}{x} \right| dx \\ &\geq \sum_{j=1}^k \frac{1}{j\pi} \int_{(j-1)\pi}^{\pi} |\sin(x)| dx \\ &= \frac{2}{\pi} \sum_{j=1}^k \frac{1}{j} \\ &\rightarrow \infty \quad \text{falls } k \rightarrow \infty, \end{aligned}$$

das Integral konvergiert wie behauptet nicht absolut.

Unbeschränkte Funktionen.

Es sei nun $f: I = [a, b) \rightarrow \mathbb{R}$ bzw. $I = (a, b] \rightarrow \mathbb{R}$ auf allen kompakten Teilintervallen von $[a, b)$ (bzw. von $(a, b]$) integrierbar, aber nicht notwendigerweise für “ $x \rightarrow b^-$ ” (bzw. für “ $x \rightarrow a^+$ ”) beschränkt.

²Eulers Formel (1781) besagt $\int_0^{\infty} \frac{\sin(x)}{x} dx = \frac{\pi}{2}$.

Definition 6.5. UNEIGENTLICHES INTEGRAL

Es sei f wie oben im Fall unbeschränkter Funktionen.

i) Falls der Grenzwert

$$\lim_{\xi \rightarrow b^-} \int_a^\xi f(x) \, dx \quad \text{bzw.} \quad \lim_{\xi \rightarrow a^+} \int_\xi^b f(x) \, dx$$

existiert, so heißt dieser das *uneigentliche Integral* von f über I .

Sprechweise: Das uneigentliche Integral existiert oder *konvergiert*.

Notation:

$$\int_a^b f(x) \, dx .$$

Andernfalls heißt $\int_a^b f(x) \, dx$ *divergent*.

ii) Das uneigentliche Integral $\int_a^b f(x) \, dx$ heißt *absolut konvergent*, falls $\int_a^b |f(x)| \, dx$ konvergiert.

Absolute Konvergenz impliziert Konvergenz des uneigentlichen Integrals.

Im Fall unbeschränkter Funktionen gelten analoge Bemerkungen und Kriterien wie Fall eines unbeschränkten Integrationsintervalls (insbesondere das *Majorantenkriterium*).

Beispiele.

i) Für $0 < \xi < 1$ ist

$$\int_\xi^1 \frac{dx}{\sqrt{x}} = \left[2\sqrt{x} \right]_\xi^1 = 2 - 2\sqrt{\xi} \xrightarrow{\xi \rightarrow 0^+} 2 ,$$

dennach ist

$$\int_0^1 \frac{dx}{\sqrt{x}} = 2 ,$$

insbesondere ist das Integral konvergent.

ii) Analog folgt die **Konvergenz** von

$$\int_0^1 \frac{dx}{x^\alpha} = \frac{1}{1-\alpha} \quad \text{im Fall} \quad 0 < \alpha < 1$$

und die **Divergenz** dieses unbestimmten Integrals im Fall $\alpha \geq 1$ (vgl. Übungskapitel 6.6).

6.5 Numerische Integration (Newton-Cotes Formeln; Newton-Cotes Summenformeln; Integrationsverfahren von Romberg)

Obwohl etwa jede stetige Funktion nach dem Hauptsatz der Differential- und Integralrechnung eine Stammfunktion besitzt, ist diese häufig nicht analytisch darstellbar.

Selbst elementar aussehende Integrale wie

$$\int e^{x^2} dx$$

können nur näherungsweise bestimmt werden.

Die Idee.

Im ersten Schritt zur **numerischen Integration**, d.h. zur numerischen Approximation von

$$I(f) = \int_a^b f(x) dx$$

wird der Integrand f durch ein Interpolationspolynom p_n ersetzt und

$$\int_a^b p_n(x) dx$$

als Näherung zu berechnen.

Im einfachsten Fall approximiert man $\int_a^b f(x) dx$ durch die “Fläche unter der Verbindungsstrecke zwischen $(a, f(a))$ und $(b, f(b))$ ” (vgl. Abbildung

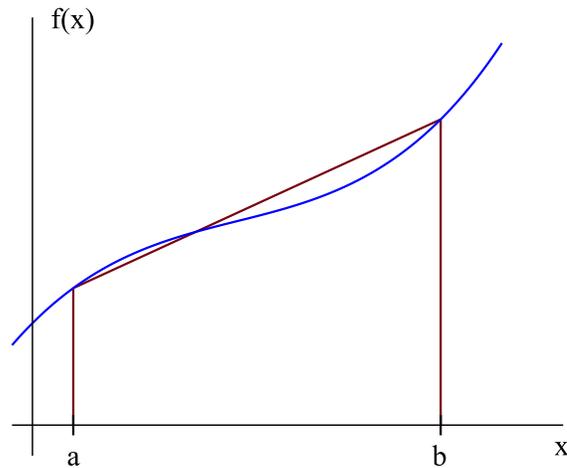


Abbildung 6.7: Zur numerischen Integration.

6.7).

Dies entspricht der unten diskutierten [Trapezregel](#).

Die Konstruktion.

Es seien

$$a \leq x_0 < x_1 < \cdots < x_n \leq b$$

gegeben und p_n bezeichne das (eindeutig bestimmte) Interpolationspolynom vom Grad $\leq n$ mit

$$f(x_i) = p_n(x_i), \quad i = 0, \dots, n.$$

An dieser Stelle sei an die Lagrangesche Darstellung des Interpolationspolynoms (vgl. Teil I, Kapitel 6) erinnert:

$$p_n(x) = \sum_{i=0}^n f(x_i) L_i^{(n)}(x),$$

$$L_i^{(n)}(x) = \prod_{k=0, k \neq i}^n \frac{x - x_k}{x_i - x_k}.$$

Entsprechend obiger Idee setzt man

$$I_n(f) := \int_a^b p_n(x) \, dx = \sum_{i=0}^n f(x_i) \left[\int_a^b L_i^{(n)}(x) \, dx \right]$$

$$= \sum_{i=0}^n f(x_i) \alpha_i^{(n)}.$$

Dabei heißen die

$$\alpha_i^{(n)} := \int_a^b L_i^{(n)}(x) \, dx$$

Quadraturgewichte.

Sie sind **unabhängig von f** und nur **abhängig von den gewählten Stützstellen**.

Wird speziell $f \equiv 1$ gewählt, so folgt

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

Die Newton-Cotes Quadraturformeln.

Im Fall **äquidistanter Stützstellen** spricht man von den sogenannten **Newton-Cotes Quadraturformeln**:

i) Es sei für $n \in \mathbb{N}$ und für $i = 0, 1, \dots, n$

$$x_i = a + ih, \quad h = \frac{b-a}{n}.$$

Hier ist $x_0 = a$, $x_n = b$, d.h. die Intervallgrenzen sind Teilpunkte und man spricht von den [abgeschlossenen Newton-Cotes Formeln](#).

ii) Es sei für $n \in \mathbb{N}_0$ und für $i = 0, 1, \dots, n$

$$x_i = a + (i+1)h, \quad h = \frac{b-a}{n+2}.$$

In diesem Fall ist $x_0 > a$ und $x_n < b$, d.h. die Intervallgrenzen sind keine Teilpunkte und man spricht von den [offenen Newton-Cotes Formeln](#).

Zur Berechnung der Quadraturgewichte.

Die konkrete Berechnung der Quadraturgewichte $\alpha_i^{(n)}$ sei hier am Beispiel der abgeschlossenen Newton-Cotes Formeln vorgestellt.

Dabei soll gleichzeitig zu Größen $\sigma_i^{(n)}$, $i = 0, \dots, n$, übergegangen werden, die nicht mehr vom Intervall $[a, b]$ sondern lediglich noch von i und n abhängen.

Mit der Transformation $t: [a, b] \mapsto [0, n]$,

$$t(x) = \frac{n}{b-a}(x-a) = \frac{1}{h}(x-a),$$

folgt

$$\begin{aligned} L_i^{(n)}(x) &= \prod_{k=0, k \neq i}^n \frac{x - x_k}{x_i - x_k} = \frac{a + t(x)h - (a + kh)}{a + ih - (a + kh)} \\ &= \prod_{k=0, k \neq i}^n \frac{t(x) - k}{i - k}. \end{aligned}$$

Für die Quadraturgewichte ergibt sich für $i = 0, \dots, n$ mittels der Substitution $\tau = t(x)$

$$\alpha_i^{(n)} = \int_a^b L_i^{(n)}(x) dx = h \int_0^n \prod_{k=0, k \neq i}^n \frac{\tau - k}{i - k} d\tau,$$

mit anderen Worten

$$\alpha_i^{(n)} = h\sigma_i^{(n)}, \quad \sigma_i^{(n)} := \int_0^n \prod_{k=0, k \neq i}^n \frac{t - k}{i - k} dt.$$

Es gilt

$$\sum_{i=0}^n \sigma_i^{(n)} = \frac{1}{h} \sum_{i=0}^n \alpha_i^{(n)} = \frac{1}{h}(b - a) = n$$

und die $\sigma_i^{(n)}$ hängen zudem wie angestrebt nur von i und n ab.

Sie werden **einmal berechnet und dann tabelliert**:

Abgeschlossene Newton-Cotes Formeln.

n	h	$\frac{1}{n}\sigma_i^{(n)}, i = 0, 1, \dots, n$	Name
1	$b - a$	$\frac{1}{2} \quad \frac{1}{2}$	Trapez-Regel
2	$\frac{b-a}{2}$	$\frac{1}{6} \quad \frac{4}{6} \quad \frac{1}{6}$	Simpson-Regel
3	$\frac{b-a}{3}$	$\frac{1}{8} \quad \frac{3}{8} \quad \frac{3}{8} \quad \frac{1}{8}$	$\frac{3}{8}$ -Regel
4	$\frac{b-a}{4}$	$\frac{7}{90} \quad \frac{32}{90} \quad \frac{12}{90} \quad \frac{32}{90} \quad \frac{7}{90}$	Milner-Regel
5	$\frac{b-a}{5}$	$\frac{19}{288} \quad \frac{75}{288} \quad \frac{50}{288} \quad \frac{50}{288} \quad \frac{75}{288} \quad \frac{19}{288}$	
6	$\frac{b-a}{6}$	$\frac{41}{840} \quad \frac{216}{840} \quad \frac{27}{840} \quad \frac{272}{840} \quad \frac{27}{840} \quad \frac{216}{840} \quad \frac{41}{840}$	Weddle-Regel

$$\begin{aligned}
n = 1 \quad \text{Trapez-Regel} \quad I_1(f) &= \frac{h}{2}(f(a) + f(b)) \\
n = 2 \quad \text{Simpson-Regel} \quad I_2(f) &= \frac{h}{3}\left(f(a) + 4f\left(\frac{a+b}{2}\right) + f(b)\right) \\
n = 3 \quad \frac{3}{8}\text{-Regel} \quad I_3(f) &= \frac{3}{8}h(f_0 + 3f_1 + 3f_2 + f_3) \\
n = 4 \quad \text{Milner-Regel} \quad I_4(f) &= \frac{2}{45}h(7f_0 + 32f_1 + 12f_2 + 32f_3 + 7f_4) \\
n = 5 \quad I_5(f) &= \frac{5}{288}h(19f_0 + 75f_1 + 50f_2 + 50f_3 \\
&\quad + 75f_4 + 19f_5) \\
n = 6 \quad \text{Weddle-Regel} \quad I_6(f) &= \frac{1}{140}h(41f_0 + 216f_1 + 27f_2 + 272f_3 \\
&\quad + 27f_4 + 216f_5 + 41f_6).
\end{aligned}$$

Hierbei ist $f_i := f(a + ih)$, $i = 0, \dots, n$.

Beispiel. Zu berechnen sei

$$\int_0^1 \frac{x}{1+x^2} dx.$$

Der exakte Wert ist (auf 9 Stellen)

$$\int_0^1 \frac{x}{1+x^2} dx = \frac{1}{2} \ln(1+x^2) \Big|_0^1 = 0.346573590.$$

Die Trapez-Regel liefert in diesem Beispiel

$$I_1(f) = \frac{1}{2} [f(0) + f(1)] = 0.25.$$

Die Simpson-Regel ergibt

$$I_2(f) = \frac{1}{3} \left[f(0) + 4f\left(\frac{1}{2}\right) + f(1) \right] = \frac{21}{60} = 0.35.$$

Summenregeln.

Nun wurde bereits bei der Polynominterpolation diskutiert, dass es im Allgemeinen nicht günstig ist, mit einer großen Anzahl von Stützstellen zu arbeiten.

Es besteht also nicht die Hoffnung, ein Integral durch Newton-Cotes Formeln mit wachsender Anzahl n von Stützstellen immer besser zu approximieren.

Idee. Stattdessen zerlegt man das Intervall $[a, b]$ in N Teilintervalle $[x_{i-1}, x_i]$,

$$x_i = a + i \frac{b-a}{N} =: a + ih_N, \quad i = 0, 1, \dots, N.$$

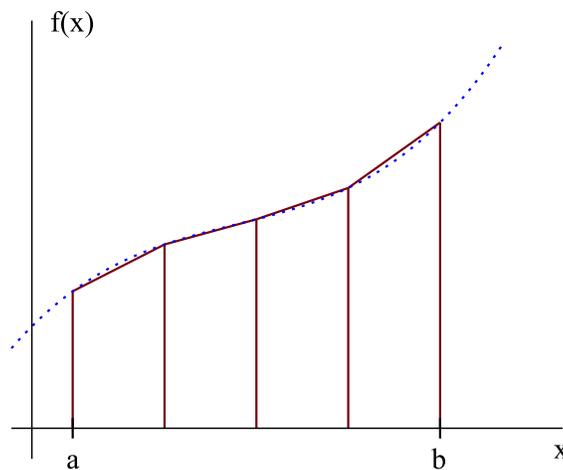


Abbildung 6.8: Zu den Newton-Cotes Summenformeln.

Auf diesen Teilintervallen wendet man dann eine Newton-Cotes Formel für kleines n an (vgl. Abbildung 6.8):

$$I_n^\Sigma(f) := \sum_{i=1}^N I_{n|[x_{i-1}, x_i]}(f)$$

ist eine geeignete Näherung, wobei für $i = 1, \dots, n$ die Größe $I_{n|[x_{i-1}, x_i]}$ das Integral über $[x_{i-1}, x_i]$ des Interpolationspolynoms p_n zu $(n+1)$

äquidistanten Stützstellen in $[x_{i-1}, x_i]$ bezeichne.

Es ergeben sich die **Newton-Cotes Summenformeln**.

Beispielsweise sieht die Situation im Fall der **Trapez-Summenregel** wie folgt aus:

Es ist $N \in \mathbb{N}$, $h_N = (b - a)/N$ und $x_i = a + ih_N$ für $i = 0, 1, \dots, N$.

Für die Trapez-Regel auf dem Intervall $[x_{i-1}, x_i]$ ist anzusetzen

$$h = x_i - x_{i-1} = a + ih_N - (a + (i - 1)h_N) = h_N = \frac{b - a}{N}.$$

Es folgt

$$\begin{aligned} T(f, h = h_N) &:= I_1^\Sigma(f) = \sum_{i=1}^N \frac{h}{2} [f(x_{i-1}) + f(x_i)] \\ &= h \left[\frac{1}{2}f(x_0) + f(x_1) + f(x_2) + \dots + f(x_{N-1}) + \frac{1}{2}f(x_N) \right] \\ &= h \left[\frac{1}{2}f(a) + f(a + h) + f(a + 2h) + \dots \right. \\ &\quad \left. + f(b - h) + \frac{1}{2}f(b) \right]. \end{aligned}$$

Beispiel. Eine Näherung von $\int_0^1 \frac{x}{1+x^2} dx$ über die Trapez-Summenregel mit $N = 4$ ist

$$\begin{aligned} T(f, h_4) &= \frac{1}{4} \left[\frac{1}{2}f(0) + f\left(\frac{1}{4}\right) + f\left(\frac{1}{2}\right) + f\left(\frac{3}{4}\right) + \frac{1}{2}f(1) \right] \\ &= \frac{1}{4} \left[0 + \frac{1}{4} \frac{16}{17} + \frac{1}{2} \frac{4}{5} + \frac{3}{4} \frac{16}{25} + \frac{1}{4} \right] \\ &= 0.341323529. \end{aligned}$$

Bemerkung. Die Idee des [Integrationsverfahrens von Romberg](#) ist es schließlich, $T(f, h_{N_i})$ zu einer absteigenden Folge von Stützweiten h_{N_i} , $i = 1, \dots, m$, zu berechnen und anschließend zu diesen Daten eine Extrapolation zum Limes $h \rightarrow 0$ (vgl. Kapitel 5.3) durchzuführen.

Dabei kann wieder mit “Polynomen in h^2 ” gearbeitet werden.

6.6 Übungsaufgaben zu Kapitel 6

Aufgabe 1. Eine Zerlegung \mathcal{Z}^* von I heißt **Verfeinerung** der Zerlegung \mathcal{Z} von I ($\mathcal{Z} \subseteq \mathcal{Z}^*$), falls alle Teilpunkte von \mathcal{Z} auch Teilpunkte von \mathcal{Z}^* sind.

Zeigen Sie:

i) Ist $\mathcal{Z} \subseteq \mathcal{Z}^*$, so folgt

$$\underline{S}_{\mathcal{Z}}(f) \leq \underline{S}_{\mathcal{Z}^*}(f) \leq \overline{S}_{\mathcal{Z}^*}(f) \leq \overline{S}_{\mathcal{Z}}(f) .$$

ii) Für zwei beliebige Zerlegungen $\mathcal{Z}_1, \mathcal{Z}_2$ von I gilt

$$\underline{S}_{\mathcal{Z}_1}(f) \leq \overline{S}_{\mathcal{Z}_2}(f) .$$

iii) Für jede Zerlegung \mathcal{Z} von I ist

$$|I| \inf_I f \leq \underline{S}_{\mathcal{Z}}(f) \leq \overline{S}_{\mathcal{Z}}(f) \leq |I| \sup_I f .$$

iv) Für eine beliebige Zerlegung \mathcal{Z} von I gilt

$$\underline{S}_{\mathcal{Z}}(f) \leq \underline{I}(f) \leq \overline{I}(f) \leq \overline{S}_{\mathcal{Z}}(f) .$$

Aufgabe 2. Es sei $I = [0, 1]$. Gibt es eine beschränkte Funktion $f: I \rightarrow \mathbb{R}$, die nicht von der Klasse $\mathcal{R}(I)$ ist?

Aufgabe 3.

i) Zeigen Sie mithilfe des Zwischenwertsatzes 4.4 den so genannten **Mittelwertsatz** der Integralrechnung:

Es sei f stetig auf $I = [a, b]$ und es sei $p \in \mathcal{R}(I)$ mit $p(x) \geq 0$ für alle $x \in I$. Dann gibt es ein $\xi \in (a, b)$ mit

$$\int_a^b f(x)p(x) \, dx = f(\xi) \int_a^b p(x) \, dx .$$

ii) Wie lautet der Spezialfall $p \equiv 1$ und wie ist dieser geometrisch zu interpretieren?

Aufgabe 4. Es sei $g: [0, 1] \rightarrow \mathbb{R}$ stetig und $f: [0, 1] \rightarrow \mathbb{R}$ für alle $x \in (0, 1)$ gegeben durch

$$f(x) = \int_0^{x^2} g(t) dt .$$

Berechnen Sie $f'(x)$.

Aufgabe 5.* Berechnen Sie

$$\int_0^1 (1 + t^2)^n t dt .$$

Aufgabe 6. Berechnen Sie die bestimmten bzw. unbestimmten Integrale

$$\int_0^{2\pi} \sin^2(x) dx , \quad \int \frac{x^2 + 1}{x^3 - 2x^2 + x} dx ,$$

$$\int \frac{x \sin(x^2)}{\cos(x^2) + 1} dx , \quad \int x^2 e^x dx ,$$

$$\int \frac{e^x}{e^{2x} + 1} dx , \quad \int \sqrt{1 + x^2} dx ,$$

$$\int \frac{x}{2x^2 - 6x + 4} dx , \quad \int_1^{e^\pi} \frac{\cos(\ln(x)) \ln(x)}{x} dx ,$$

$$\int \frac{1}{x^3 - x^2 + x - 1} dx , \quad \int e^{1 + \ln(x) + x^2} dx .$$

Aufgabe 7. Es sei f eine lokal integrierbare Funktion auf \mathbb{R} . Der Cauchysche Hauptwert

$$\int_{-\infty}^{\infty} f(x) \, dx$$

ist (sofern existent)

$$\lim_{R \rightarrow \infty} \int_{-R}^R f(x) \, dx .$$

- i) Berechnen Sie den Cauchyschen Hauptwert $\int_{-\infty}^{\infty} x^3 \, dx$.
 ii) Zeigen Sie, dass $\int_{-\infty}^{\infty} x^3 \, dx$ divergiert.
-

Aufgabe 8.

- i) Es sei $\alpha \in \mathbb{R}$ fixiert. Konvergiert das uneigentliche Integral

$$\int_0^1 \frac{dx}{x^\alpha} ?$$

- ii) Konvergiert das uneigentliche Integral

$$\int_0^{\infty} \frac{x\sqrt{1+x^2}}{x^3} \, dx ?$$

- iii) Konvergieren die uneigentlichen Integrale

$$\int_0^2 \frac{e^{\sin(x)}}{x(2+x)} \, dx , \quad \int_2^{\infty} \frac{3 - \sin(x)}{x(x+1)} \, dx ?$$

Aufgabe 9.* Konvergiert das uneigentliche Integral

$$\int_0^{\infty} e^{-x^2} \, dx ?$$

Aufgabe 10.*

i) Berechnen Sie die unbestimmten Integrale

$$\int \frac{x \sinh(x^2)}{1 + \cosh(x^2)} dx, \quad \int \frac{4x - 4}{x^2 - 4} dx, \quad \int x \ln(x) dx .$$

ii) Berechnen Sie das bestimmte Integral

$$\int_0^{\sqrt{\pi}} \cos(\pi + x^2)(\pi x + x^3) dx .$$

iii) Konvergiert das uneigentliche Integral

$$\int_0^2 \frac{e^{\sin(x)}}{x(2+x)} dx ?$$

Aufgabe 11.

i) Bestimmen Sie eine Approximation (8 Nachkommastellen) von

$$\int_0^1 \frac{x}{1+x^2} dx$$

i) mit der Trapez-Regel; *ii)* mit der Simpson-Regel; *iii)* mit der 3/8-Regel; *iv)* mit der summierten Trapez-Regel zu $N = 4$.

Wie groß ist jeweils der (absolute) Fehler?

ii) Berechnen Sie eine Approximation (8 Nachkommastellen) von

$$\int_0^1 \sqrt{1-x^2} dx$$

mit der Simpson-Regel und mit der summierten Trapez-Regel zu $N = 4$.

Wie groß ist jeweils der (absolute) Fehler?

Lösungshinweise zu den Übungsaufgaben.

Aufgabe 5. Gesucht ist das bestimmte

$$\int_0^1 (1+t^2)^n t \, dt ,$$

zu dessen Berechnung sich die Substitutionsregel anbietet:

Die Funktion $h(t) := (1+t^2)^n t$ lässt sich in der Form

$$h(t) = \frac{1}{2}(1+t^2)^n \frac{d}{dt}(1+t^2)$$

schreiben, also

$$2h(t) = f(g(t))g'(t)$$

mit $f(x) := x^n$ und $g(t) := 1+t^2$.

Es folgt

$$\begin{aligned} \int_0^1 (1+t^2)^n t \, dt &= \frac{1}{2} \int_0^1 f(g(t))g'(t) \, dt \\ &= \frac{1}{2} \int_1^2 x^n \, dx \\ &= \frac{x^{n+1}}{2(n+1)} \Big|_1^2 = \frac{2^{n+1} - 1}{2(n+1)} . \end{aligned}$$

Aufgabe 9. Zur Analyse des Integrals

$$\int_0^\infty e^{-x^2} \, dx$$

beachtet man zunächst, dass $x^2 \geq 2x - 1$ aus $(x-1)^2 \geq 0$ folgt, d.h. für alle $x > 0$ gilt

$$e^{-x^2} \leq e^{1-2x} =: g(x) .$$

Weiter ist für alle $0 < b$

$$\int_0^b g(x) \, dx = \int_0^b e^{1-2x} \, dx = -\frac{1}{2} e^{1-2x} \Big|_0^b = \frac{e}{2} (1 - e^{-2b}) ,$$

die Funktion g ist somit eine konvergente Majorante von e^{-x^2} und das Majorantenkriterium impliziert die (absolute) Konvergenz von $\int_0^\infty e^{-x^2} dx$.

Aufgabe 10.

i) Es ist

$$\begin{aligned} \int \frac{x \sinh(x^2)}{1 + \cosh(x^2)} dx &= \int \frac{\frac{1}{2} \frac{d}{dx} (1 + \cosh(x^2))}{1 + \cosh(x^2)} dx \\ &= \frac{1}{2} \ln (1 + \cosh(x^2)) . \end{aligned}$$

Beim zweiten Integral beachtet man $x^2 - 4 = (x + 2)(x - 2)$ und macht den Ansatz einer Partialbruchzerlegung:

$$\begin{aligned} \frac{4x - 4}{x^2 - 4} &= \frac{A}{x + 2} + \frac{B}{x - 2} \\ &= \frac{A(-2) + B(x + 2)}{x^2 - 4} \\ &= \frac{x(A + B) - 2A + 2B}{x^2 - 4} . \end{aligned}$$

Man erhält unmittelbar die Lösung $A = 3$, $B = 1$.

Es folgt

$$\begin{aligned} \int \frac{4x - 4}{x^2 - 4} dx &= 3 \int \frac{1}{x + 2} dx + \int \frac{1}{x - 2} dx \\ &= 3 \ln (|x + 2|) + \ln (|x - 2|) . \end{aligned}$$

Eine partielle Integration liefert schließlich für das dritte Integral

$$\begin{aligned} \int x \ln(x) dx &= \frac{1}{2} x^2 \ln(x) - \frac{1}{2} \int x^2 \frac{1}{x} dx \\ &= \frac{1}{2} x^2 \ln(x) - \frac{1}{4} x^2 = \frac{1}{2} x^2 \left(\ln(x) - \frac{1}{2} \right) . \end{aligned}$$

ii) Es ist $\pi x + x^3 = x(\pi + x^2)$ und die Substitution $y = \pi + x^2$ liefert zusammen mit einer partiellen Integration

$$\begin{aligned} \int_0^{\sqrt{\pi}} \cos(\pi + x^2)(\pi x + x^3) dx &= \frac{1}{2} \int_{\pi}^{2\pi} \cos(y)y dy \\ &= \frac{1}{2} \left[\sin(y)y \Big|_{\pi}^{2\pi} - \int_{\pi}^{2\pi} \sin(y) dy \right] \\ &= -\frac{1}{2} \left[-\cos(y) \right]_{\pi}^{2\pi} = 1 . \end{aligned}$$

iii) Für $x < 2$ gilt $2 + x \leq 4$ und wegen $0 \leq x$ hat man sofort

$$\frac{1}{x(2+x)} \geq \frac{1}{4x} .$$

Weiter impliziert $-1 \leq \sin(x)$ aufgrund der Monotonie der Exponentialfunktion

$$e^{\sin(x)} \geq \frac{1}{e} .$$

Insgesamt ergibt sich für alle $0 < \varepsilon < 2$

$$\int_0^2 \frac{e^{\sin(x)}}{x(2+x)} dx \geq \frac{1}{4e} \int_{\varepsilon}^2 \frac{1}{x} dx ,$$

was die Divergenz des Integrals zeigt.

Kapitel 7

Der Satz von Taylor

7.1 Taylor-Formel und Taylor-Reihe (Taylor-Polynom; Restglied; Integraldarstellung des Restgliedes; Lagrangesche Restgliedformel; die Klasse C^∞ ; reell analytische Funktionen)

In Kapitel 5 wird eine Funktion f in der Nähe eines fixierten Punktes x_0 mit einer (affin) linearen Funktion approximiert, deren Steigung die erste Ableitung von f im Punkt x_0 ist.

Für eine “bessere” Approximation, die beispielsweise auch das Krümmungsverhalten von f berücksichtigt, liegt es nahe, die Funktion mit quadratischen Polynomen bzw. mit Polynomen höherer Ordnung zu approximieren (nicht zu verwechseln mit der Polynominterpolation aus Teil I, Kapitel 6).

Als Preis ist der höhere Rechenaufwand zu zahlen (sei es analytisch oder numerisch).

Entscheidend ist bei diesen Betrachtungen die Güte der Approximation, d.h. die Abweichung der Approximation von der gegebenen Funktion f . Dabei hängt der Fehler im Allgemeinen davon ab, wie weit man sich vom gegebenen Punkt x_0 entfernt.

In Abbildung 7.1 ist zur Veranschaulichung der Sinus mit einem Polynom ersten, dritten und fünften Grades approximiert.

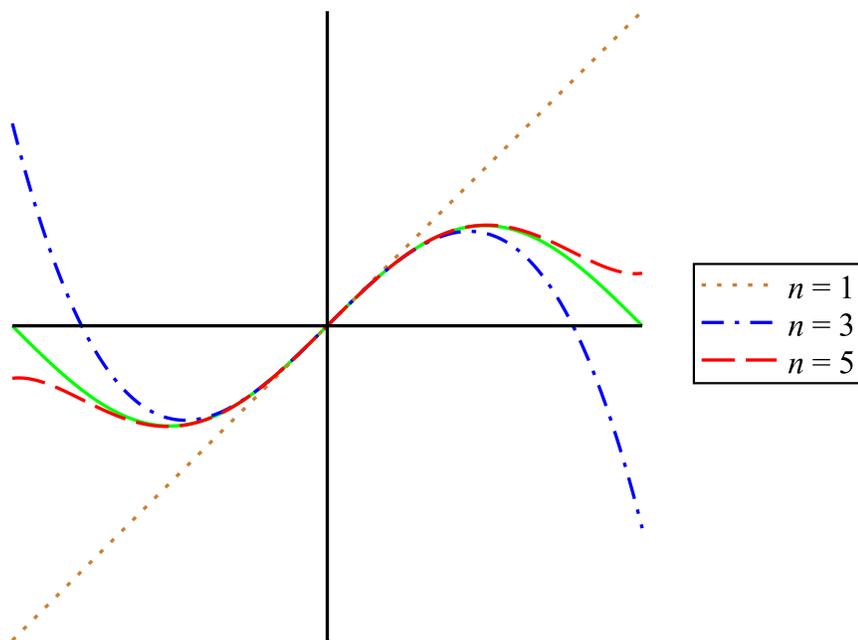


Abbildung 7.1: Taylor-Polynome des Sinus.

Motivation für die Taylorsche Formel: Polynome.

Man betrachte im einfachsten Fall ein Polynom vom Grad n ,

$$p(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n .$$

Für $k = 0, 1, \dots, n$ ist die k -te Ableitung an der Stelle Null

$$p^{(k)}(0) = k!a_k ,$$

mit anderen Worten: Es gilt

$$p(x) = \sum_{k=0}^n \frac{1}{k!} p^{(k)}(0) x^k .$$

Zu beliebigem $x_0 \in \mathbb{R}$ kann das Polynom ebenso in der Form

$$p(x) = \sum_{k=0}^n b_k (x - x_0)^k ,$$

geschrieben werden und analog gilt

$$p(x) = \sum_{k=0}^n \frac{1}{k!} p^{(k)}(x_0) (x - x_0)^k .$$

Falls eine Funktion in Polynomdarstellung „entwickelt“ werden kann, so müssen die Koeffizienten ebenfalls von dieser Form sein.

Definition 7.1. TAYLORSCHES FORMEL

Es sei $f: I = (a, b) \rightarrow \mathbb{R}$ von der Klasse $C^n(I)$ und es sei $x_0 \in (a, b)$.

Dann heißt

$$T_n(x; x_0) := \sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k$$

das *Taylor-Polynom* n -ten Grades zum *Entwicklungspunkt* x_0 .

Die *Taylorsche Formel* lautet

$$f(x) = T_n(x; x_0) + R_n(x - x_0) ,$$

wobei das *Restglied* $R_n(x - x_0)$ im Folgenden zu quantifizieren ist.

Bemerkungen.

- i) Das Taylor-Polynom ist nach Definition 7.1 so aufgebaut, dass die Ableitungen des Taylor-Polynoms an der Stelle x_0 bis zur Ordnung n mit denen von f übereinstimmen.
- ii) Die Taylorsche Formel in Definition 7.1 ist lediglich eine Definition des Restglieds. Ohne eine Abschätzung der Größe des Restglieds enthält sie keinerlei Information.

Beispiel. Es sei $\alpha \in \mathbb{R}$, $\alpha \neq 0$, fixiert und für $x > -1$ sei $f(x) = (1 + x)^\alpha$.

Dann gilt für alle $k \in \mathbb{N}$

$$\begin{aligned} f'(x) &= \alpha(1+x)^{\alpha-1}, \\ f''(x) &= \alpha(\alpha-1)(1+x)^{\alpha-2}, \\ &\vdots \\ f^{(k)}(x) &= \alpha(\alpha-1)\dots(\alpha-k+1)(1+x)^{\alpha-k}. \end{aligned}$$

Setzt man in Verallgemeinerung der Binomialkoeffizienten (Teil I, Definition 2.2) für alle $k \in \mathbb{N}$

$$\binom{\alpha}{k} := \frac{\alpha(\alpha-1)\dots(\alpha-k+1)}{k!}, \quad \binom{\alpha}{0} = 1,$$

so folgt

$$(1+x)^\alpha = \sum_{k=0}^n \binom{\alpha}{k} x^k + R_n(x).$$

Man vergleiche dies mit dem binomischen Lehrsatz (Teil I, Satz 2.2).

Zur Güte der Approximation.

Eine Antwort auf die Frage nach der Größe des Fehlers gibt

Satz 7.1. ERSTE DARSTELLUNG DES RESTGLIEDS

Es sei $f: I = (a, b) \rightarrow \mathbb{R}$ von der Klasse $C^{n+1}(I)$ und es seien $x_0, x = x_0 + h \in (a, b)$.

Dann gilt

$$R_n(h) = \int_0^1 \frac{1}{n!} (1-t)^n f^{(n+1)}(x_0 + th) h^{n+1} dt.$$

Beweis. Der Beweis ergibt sich unmittelbar aus dem folgenden Lemma (vgl. Übungskapitel 7.2). \square

Lemma 7.1.

HERLEITUNG DER RESTGLIEDDARSTELLUNG

Für eine stetige Funktion $\varphi: [0, 1] \rightarrow \mathbb{R}$ von der Klasse $C^{n+1}((0, 1))$ gilt

$$\varphi(1) = \sum_{k=0}^n \frac{1}{k!} \varphi^{(k)}(0) + \int_0^1 \frac{1}{n!} (1-t)^n \varphi^{(n+1)}(t) dt .$$

Beweis. Der Beweis folgt induktiv mithilfe partieller Integration (vgl. Übungskapitel 7.2) \square

Weitere Darstellungen des Restglieds.

Es gibt eine Reihe von verschiedenen Darstellungen des Restgliedes:

Über eine einfache Substitution gelangt man zur [Integraldarstellung des Restgliedes](#) in ihrer üblichen Form.

Die wohl bekannteste Darstellung ist die [Lagrangesche Restgliedformel](#). Sie folgt mit Hilfe des Mittelwertsatzes der Integralrechnung (siehe Übungskapitel 6.6).

Korollar 7.1.

WEITERE RESTGLIEDDARSTELLUNGEN

Es seien f , x und x_0 wie oben gegeben, $h = x - x_0$.

i) Integraldarstellung des Restgliedes: Es ist

$$R_n(x - x_0) = \int_{x_0}^x \frac{1}{n!} (x - t)^n f^{(n+1)}(t) dt .$$

ii) *Lagrangesche Restgliedformel*: Es ist für ein $\theta \in (0, 1)$

$$R_n(x - x_0) = \frac{1}{(n+1)!} f^{(n+1)}(x_0 + \theta h) h^{n+1} .$$

Wie in der Einleitung dieses Kapitels bereits angedeutet, erkennt man anhand der Restglieddarstellungen, dass eine “gute Approximation” der Funktion durch ein Taylor-Polynom in der Regel nur für “kleine $|h|$ ”, d.h. in der Nähe des Entwicklungspunktes x_0 , erwarten werden kann.

Umgekehrt ausgedrückt: Ist bei festem n der Punkt x sehr weit von x_0 entfernt, so kann das Restglied (der Fehler) sehr groß werden.

Beispiele.

i) Es sei $f(x) = e^x$. Dann ist

$$T_n(x; 0) = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!} = \sum_{k=0}^n \frac{x^k}{k!} ,$$

genau wie es die Definition als Potenzreihe erwarten lässt.

Für das Restglied gilt (für ein $\theta \in (0, 1)$)

$$\begin{aligned} |R_n(x)| &= \left| \frac{e^{\theta x}}{(n+1)!} \right| |x^{n+1}| \leq \frac{e^{\theta|x|}}{(n+1)!} |x|^{n+1} \\ &\leq \frac{e^{|x|}}{(n+1)!} |x|^{n+1} . \end{aligned}$$

Insbesondere ist für jedes fixierte $x \in \mathbb{R}$ bewiesen:

$$R_n(x) \xrightarrow{n \rightarrow \infty} 0 .$$

ii) Ähnliches gilt für $\sin(x)$, $\cos(x)$... (vgl. Übungskapitel 7.2). Die Taylor-Polynome des Sinus zum Entwicklungspunkt $x_0 = 0$ sind in Abbildung 7.1 veranschaulicht.

Entwicklung als Potenzreihe.

Das Beispiel Exponentialfunktion läßt vermuten, dass eine Funktion unter gewissen Voraussetzungen im Grenzwert $n \rightarrow \infty$ durch die Taylor-Polynome reproduziert wird, d.h. (zumindest lokal) als Potenzreihe geschrieben werden kann.

Es stellen sich die wesentlichen Fragen:

- i)* Konvergiert die Folge $\{T_n(x; x_0)\}$ im Grenzwert $n \rightarrow \infty$ gegen eine Funktion $T(x)$?
- ii)* Falls ja, ist der Grenzwert $T(x)$ gleich $f(x)$?

Ist $f^{(n+1)}$ eine stetige Funktion auf dem kompakten Intervall $[a, b]$, so ist $f^{(n+1)}$ nach Satz 4.3 auf $[a, b]$ insbesondere beschränkt.

Demnach existiert eine Konstante $K = K(n)$ mit

$$|f^{(n+1)}(x)| \leq K(n) \quad \text{für alle } x \in [a, b]$$

und es gilt

$$|R_n(x - x_0)| \leq \frac{K(n)}{(n + 1)!} |x - x_0|^{n+1} .$$

Da die Konstante $K(n)$ aber von der Ordnung n der Ableitung abhängt, bedeutet das noch nicht zwingend, dass das Restglied für große n bei fixiertem x klein wird.

In Satz 7.3 ist eine explizite Bedingung an die Konstante formuliert, unter der das Restglied tatsächlich gegen Null konvergiert.

Zuvor soll jedoch anhand eines prominenten Beispiels illustriert werden, dass die Folge der Taylor-Polynome (die Raylor-Reihe) konvergieren kann, der Grenzwert aber nicht mit der Funktion f übereinstimmt.

In diesem Fall konvergiert das Restglied zwar, es konvergiert jedoch nicht gegen Null.

Cauchys Beispiel.

Betrachtet sei die in Abbildung 7.2 skizzierte Funktion $f: \mathbb{R} \rightarrow \mathbb{R}$,

$$f(x) = \begin{cases} e^{-1/x^2} & \text{für } x \neq 0, \\ 0 & \text{für } x = 0. \end{cases}$$

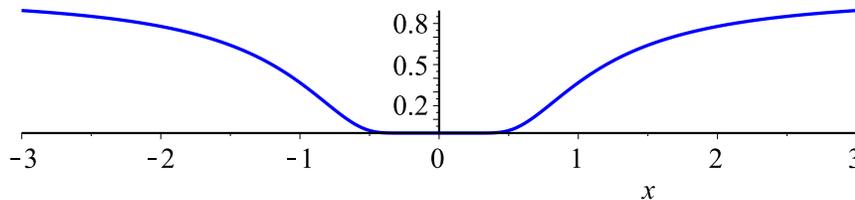


Abbildung 7.2: Cauchys Beispiel.

Es ist

$$\begin{aligned} f'(x) &= 2x^{-3}e^{-1/x^2}, \\ f''(x) &= (4x^{-6} - 6x^{-4})e^{-1/x^2}, \\ &\vdots \quad \quad \quad \vdots \end{aligned}$$

Induktiv sieht man, dass im Nullpunkt alle rechtsseitigen und linksseitigen Ableitungen existieren und gleich Null sind, die Funktion ist beliebig oft differenzierbar (von der Klasse C^∞) und für alle $k \in \mathbb{N}_0$ gilt (vgl. Übungskapitel 7.2)

$$f^{(k)}(0) = 0.$$

Dementsprechend ist für jedes $n \in \mathbb{N}$

$$T_n(x; 0) = \sum_{k=1}^n \frac{f^{(k)}(0)}{k!} x^k = 0,$$

alle Taylor-Polynome um 0 verschwinden identisch, die Funktion f hingegen nicht und das Restglied kann für große n nicht klein werden.

In der Tat gilt in Cauchys Beispiel nach der Taylorschen Formel für das Restglied um den Nullpunkt: $R_n(x) = f(x)$.

Selbst wenn die Folge $\{T_n(x; x_0)\}$ im Grenzwert $n \rightarrow \infty$ konvergiert, kann nicht auf $f(x) = \lim_{n \rightarrow \infty} T_n(x; x_0)$ geschlossen werden.

Reell analytische Funktionen.

Reell analytische Funktionen können per definitionem (lokal) als Potenzreihen dargestellt werden.

Satz 7.2. REELL ANALYTISCHE FUNKTIONEN

Es sei $f: I = (a, b) \rightarrow \mathbb{R}$ beliebig oft differenzierbar und es sei $x_0 \in (a, b)$.

Falls für jedes $x \in (a, b)$

$$\lim_{n \rightarrow \infty} R_n(x - x_0) = 0$$

gilt, so ist für alle $x \in (a, b)$

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k .$$

Sprechweisen:

- i) In diesem Fall heißt die Funktion f reell analytisch oder von der Klasse C^ω .
- ii) Die Reihe $\sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k$ heißt Taylor-Reihe von f um den Entwicklungspunkt x_0 .

Bemerkungen.

- i) Die Potenzreihendarstellung einer Funktion ist, sofern sie existiert, eindeutig bestimmt. An dieser Stelle sei an die Konvergenzkriterien für Potenzreihen aus Teil I, Kapitel 8.2, sowie an die Definitionen von Exponentialfunktion, Sinus, Kosinus ... (Teil I, Kapitel 8.3) erinnert.
- ii) Nach Satz 7.2 konvergiert die Taylor-Reihe von f gegen f , falls das Restglied gegen Null konvergiert.

Beispiel. Es sei $-1/2 < x < 1/2$, $x_0 = 0$ und $f(x) = \ln(1+x)$. Es ist

$$\begin{aligned} f'(x) &= \frac{1}{1+x}, \\ f''(x) &= -\frac{1}{(1+x)^2}, \\ &\vdots \\ f^{(k)}(x) &= (-1)^{k-1}(k-1)! \frac{1}{(1+x)^k} \quad \text{für alle } k \in \mathbb{N}. \end{aligned}$$

Die Identität

$$\begin{aligned} \ln(1+x) &= x - \frac{1}{2}x^2 + \frac{1}{3}x^3 - \dots + (-1)^{n-1} \frac{1}{n}x^n + \dots \\ &= \sum_{k=1}^{\infty} (-1)^{k-1} \frac{1}{k} x^k \end{aligned}$$

folgt aus

Satz 7.3. KRITERIUM FÜR REELL ANALYTISCHE FUNKTIONEN

Es sei $f: I = (a, b) \rightarrow \mathbb{R}$ von der Klasse $C^\infty(I)$ und mit reellen Konstanten $M, r > 0$ gelte für alle $x \in (a, b)$ und für alle $k \in \mathbb{N}_0$

$$|f^{(k)}(x)| \leq k! M r^{-k}.$$

Ist $\delta \in (0, r)$ und ist $x \in (a, b)$ derart, dass $|x - x_0| \leq \delta$ gilt, so folgt

$$f(x) = \sum_{k=0}^{\infty} \frac{1}{k!} f^{(k)}(x_0) (x - x_0)^k.$$

Beweis des Satzes. Es ist

$$R_n(h) = \frac{1}{(n+1)!} f^{(n+1)}(x_0 + \theta h) h^{n+1}, \quad h = x - x_0,$$

also für $|x - x_0| < \delta$ nach Voraussetzung

$$|R_n(h)| \leq M \left(\frac{\delta}{r}\right)^{n+1} \xrightarrow{n \rightarrow \infty} 0$$

und aus Satz 7.2 folgt die Behauptung. \square

Zurück zum Beispiel. Im obigen Beispiel $f(x) = \ln(1+x)$ bleiben die Voraussetzungen von Satz 7.3 zu verifizieren:

Für $|x| < 1/2$ und für alle $k \in \mathbb{N}$ ist

$$|f^{(k)}(x)| \leq k!(1+x)^{-k} \leq k! \left(\frac{1}{2}\right)^{-k}.$$

Der Satz kann also angewandt werden und die Konvergenz der Taylor-Reihe um $x_0 = 0$ gegen f für $|x| < 1/2$ ist bewiesen.

Bemerkung. In der Tat kann gezeigt werden, dass die Taylor-Reihe von $\ln(1+x)$ um $x_0 = 0$ für $-1 < x \leq 1$ gegen $\ln(1+x)$ konvergiert.

Konstruktion mithilfe der geometrischen Reihe.

Die explizite Berechnung einer Taylor-Reihe kann in bestimmten Fällen auf eine geometrische Reihe zurückgeführt werden. Dieser “Trick” wird bei der Reihenentwicklung in der Funktionentheorie erneut aufgegriffen.

Es sei beispielsweise die Funktion $f(x) = 1/(1+x)$ um den Punkt $x_0 = 1$ zu entwickeln.

Es ist für $|x - 1| < 2$

$$\begin{aligned}\frac{1}{1+x} &= \frac{1}{2+(x-1)} = \frac{1}{2} \frac{1}{1+\frac{x-1}{2}} \\ &= \frac{1}{2} \frac{1}{1-\left[-\frac{x-1}{2}\right]} \\ &= \frac{1}{2} \sum_{k=0}^{\infty} \frac{(-1)^k}{2^k} (x-1)^k .\end{aligned}$$

7.2 Übungsaufgaben zu Kapitel 7

Aufgabe 1.*

- i) Zeigen Sie Lemma 7.1.
 - ii) Zeigen Sie damit Satz 7.1.
-

Aufgabe 2. Zeigen Sie in Cauchys Beispiel mithilfe vollständiger Induktion: Für alle $k \in \mathbb{N}$ gilt und für $x \neq 0$ gilt

$$f^{(k)}(x) = p_k(1/x)e^{-1/x^2},$$

wobei $p_k(t)$ ein Polynom der Ordnung 3 bezeichnet. Folgern Sie daraus $f \in C^\infty(\mathbb{R})$.

Aufgabe 3.*

- i) Es sei $f: \mathbb{R} \rightarrow \mathbb{R}$ gegeben durch

$$x \mapsto f(x) = (x - 2)^2.$$

Berechnen Sie für ein fixiertes $x_0 \in \mathbb{R}$ das Taylor-Polynom $T_2(x; x_0)$ zweiten Grades zum Entwicklungspunkt x_0 von f und zeigen Sie $T_2(x; x_0) = f(x)$ (für alle $x \in \mathbb{R}$).

- ii) Bestimmen Sie für $n \in \mathbb{N}$ das Taylor-Polynom n^{ten} Grades zum Entwicklungspunkt $x_0 = 1$ von $f(x) = xe^{x-1}$.
-

Aufgabe 4.

- i) Bestimmen Sie für $n \in \mathbb{N}$ das Taylorpolynom n^{ten} Grades zum Entwicklungspunkt $x_0 = 1$ sowie das Restglied in Lagrangescher Darstellung von

$$f(x) = x^3 - 3x^2 + 3x - 1.$$

- ii) Bestimmen Sie für $n \in \mathbb{N}$ das Taylorpolynom n^{ten} Grades zum Entwicklungspunkt $x_0 = 0$ von
- (a) $f(x) = x^2 + e^x$,
- (b) $f(x) = e^{x^2}$.
- iii) Bestimmen Sie das Taylor-Polynom dritter Ordnung zum Entwicklungspunkt 0 der Funktion $f: \mathbb{R} \rightarrow \mathbb{R}$, $f(x) = \sin(x^2)$.
-

Aufgabe 5.

- i) Bestimmen Sie für $n \in \mathbb{N}$ das Taylor-Polynom n^{ten} Grades zum Entwicklungspunkt $x_0 = 0$ von $f(x) = xe^{(x^3)}$.
- ii) Bestimmen Sie das Taylor-Polynom dritter Ordnung zum Entwicklungspunkt 0 der Funktion $f: \mathbb{R} \rightarrow \mathbb{R}$, $f(x) = x \cos(\pi + x)$.
-

Aufgabe 6. Die Funktion $f: \mathbb{R} \rightarrow \mathbb{R}$ sei für alle $x \in \mathbb{R}$ gegeben durch

$$f(x) = x^2 \sin(x) .$$

Berechnen Sie das Taylorpolynom dritten Grades $T_3(x; x_0)$ von f um den Entwicklungspunkt $x_0 = 0$.

Aufgabe 7.

- i) Berechnen Sie für die Funktion $f(x) = \cos(x)$ das Restglied $R_2(x - 0)$ in Lagrangescher Darstellung.
- ii) Berechnen Sie

$$\lim_{x \rightarrow 0} \frac{\cos(x) - 1}{x^2}$$

mit Hilfe dieser Restgliedabschätzung und der Taylorschen Formel.

Lösungshinweise zu den Übungsaufgaben.

Aufgabe 1.

i) Es wird sukzessive partiell integriert mit dem Ergebnis

$$\begin{aligned}
 & \int_0^1 \frac{1}{n!} (1-t)^n \varphi^{(n+1)}(t) \, dt \\
 &= \varphi^{(n)}(t) \frac{1}{n!} (1-t)^n \Big|_0^1 + \int_0^1 \frac{1}{(n-1)!} (1-t)^{n-1} \varphi^{(n)}(t) \, dt \\
 &= -\frac{1}{n!} \varphi^{(n)}(0) + \int_0^1 \frac{1}{(n-1)!} (1-t)^{n-1} \varphi^{(n)}(t) \, dt \\
 &= \vdots \\
 &= -\left[\frac{1}{n!} \varphi^{(n)}(0) + \dots + \frac{1}{1!} \varphi'(0) \right] + \int_0^1 \varphi'(t) \, dt .
 \end{aligned}$$

Aus $\int_0^1 \varphi'(t) \, dt = \varphi(1) - \varphi(0)$ folgt das Lemma. □

ii) Für $0 \leq t \leq 1$ setzt man $\varphi(t) := f(x_0 + th)$ und es folgt

$$\varphi'(t) = hf'(x_0 + th), \quad \varphi''(t) = h^2 f''(x_0 + th), \quad \dots,$$

also für $k = 0, 1, \dots, n+1$

$$\varphi^{(k)}(t) = h^k f^{(k)}(x_0 + th).$$

Lemma 7.1 liefert

$$\begin{aligned}
 f(x_0 + h) &= \varphi(1) \\
 &= \sum_{k=0}^n \frac{1}{k!} f^{(k)}(x_0) h^k \\
 &\quad + \int_0^1 \frac{1}{n!} (1-t)^n f^{(n+1)}(x_0 + th) h^{n+1} \, dt .
 \end{aligned}$$

Dies ist genau die Behauptung des Satzes. □

Aufgabe 3.

i) Es ist

$$T_2(x; x_0) = (x_0 - 2)^2 + 2(x_0 - 2)(x - x_0) + (x - x_0)^2$$

und ausmultiplizieren ergibt die Behauptung.

ii) Man schreibe

$$\begin{aligned} xe^{x-1} &= (x-1)e^{x-1} + e^{x-1} \\ &= (x-1) \sum_{n=0}^{\infty} \frac{1}{n!} (x-1)^n + \sum_{n=0}^{\infty} \frac{1}{n!} (x-1)^n \\ &= 1 + \sum_{n=1}^{\infty} \frac{n+1}{n!} (x-1)^n \\ &= \sum_{n=0}^{\infty} \frac{n+1}{n!} (x-1)^n \end{aligned}$$

oder man zeige induktiv

$$f^{(k)}(x) = ke^{x-1} + xe^{x-1}.$$

Kapitel 8

Fourier-Reihen

8.1 Einführung (Spektrum; harmonische Analyse; Periode einer Funktion; trigonometrische Reihen; trigonometrische Polynome; Fourier-Koeffizienten)

Reell analytische Funktionen haben eine Reihendarstellung als Taylor-Reihe, die über die Ableitungen definiert ist.

Im Falle [periodischer Funktionen](#) existiert unter ansonsten sehr viel schwächeren Voraussetzungen eine weitere Reihendarstellung, nämlich die als so genannte [Fourier-Reihe](#).

Motiviert ist diese Reihendarstellung durch das Beispiel einer schwingenden Saite. Ebenso typische Beispiele periodischer Funktionen sind andere mechanische Schwingungen oder Wechselstromkreise etc.

Grund- und Oberschwingungen.

In diesem Kapitel wird der Frage nachgegangen: Lässt sich eine periodische Funktion als Überlagerung oder [Superposition](#) von [Grund- und Oberschwingungen](#) darstellen?

Man spricht von einer [harmonische Analyse](#).

Mit anderen Worten soll das [diskrete Spektrum periodischer Funktionen](#)

studiert werden. In der Technik ist diese Frage z.B. bei der Analyse von Messkurven oder Signalen von großer Relevanz.

Als exemplarisches Beispiel wird in Kapitel 8.2 die so genannte **Sägezahnfunktion** diskutiert.

Zur Verdeutlichung der Betrachtungen dieses Kapitels ist in den Abbildungen 8.1 bis 8.7 die harmonische Analyse der Sägezahnfunktion veranschaulicht.

Erinnerung: Periodische Funktionen.

Eine Funktion $f: \mathbb{R} \rightarrow \mathbb{R}$ heißt periodisch mit der Periode $T \neq 0$, falls für alle $x \in \mathbb{R}$ gilt: $f(x + T) = f(x)$.

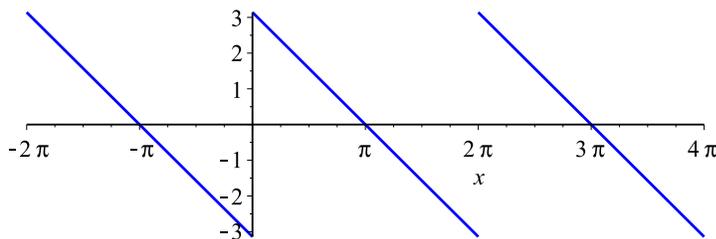


Abbildung 8.1: Die Sägezahnfunktion.

Beispiele.

- i) Als einfachstes Beispiel dienen konstante Funktionen, die bekanntesten Beispiele sind $\sin(x)$ und $\cos(x)$.
- ii) Für beliebige Amplituden $A_1, A_2 \in \mathbb{R}$ ist die Funktion

$$A_1 \sin(\omega_1 x) + A_2 \sin(\omega_2 x)$$

periodisch, falls das Verhältnis der Kreisfrequenzen (und damit der Perioden) rational ist.

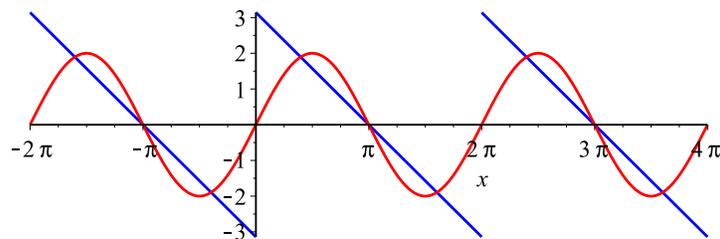


Abbildung 8.2: Die Sägezahnfunktion: Grundschwingung.

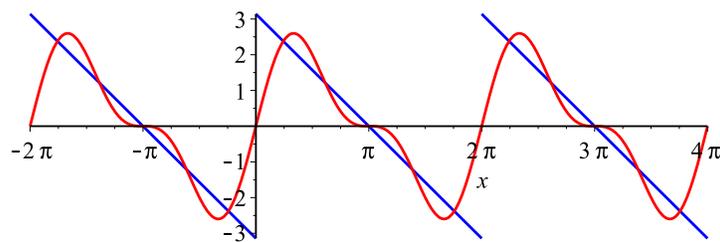


Abbildung 8.3: Die Sägezahnfunktion: Grundschwingung und eine Oberschwingung.

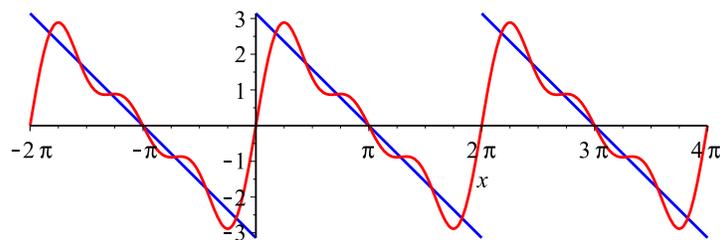


Abbildung 8.4: Die Sägezahnfunktion: Grundschwingung und zwei Oberschwingungen.

iii) Insbesondere hat die Funktion $\sin(x) + \sin(\omega x)$ die Periode 2π , falls $\omega \in \mathbb{Z}$ (vgl. Übungskapitel 8.3).

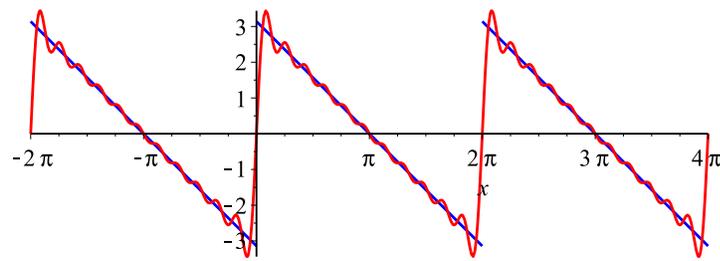


Abbildung 8.5: Die Sägezahnfunktion: Grundschwingung und zehn Oberschwingungen.

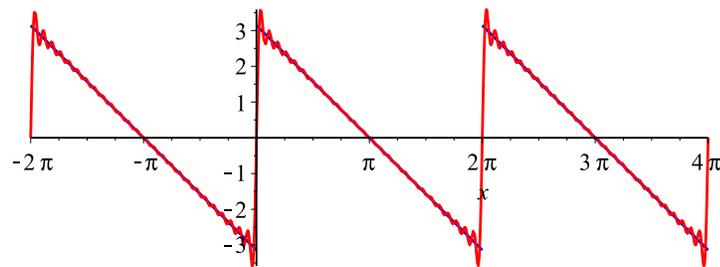


Abbildung 8.6: Die Sägezahnfunktion: Grundschwingung und 25 Oberschwingungen.

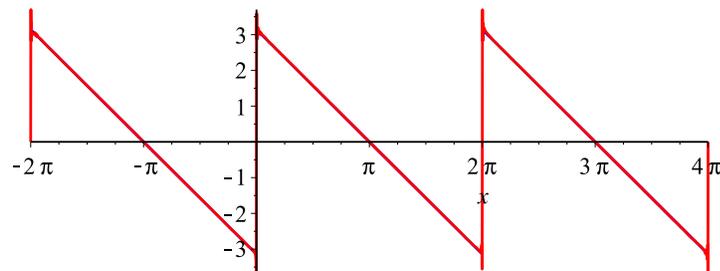


Abbildung 8.7: Die Sägezahnfunktion: Grundschwingung und 500 Oberschwingungen.

Im Folgenden werden zunächst **o.E. 2π -periodische Funktionen** betrachtet.

Die Form einer Fourier-Reihe.

Fourier-Reihen sind (konvergente) **trigonometrische Reihen** der Form

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos(nx) + b_n \sin(nx)) , \quad a_n, b_n \in \mathbb{R} ,$$

die Partialsummen

$$\frac{a_0}{2} + \sum_{n=1}^N (a_n \cos(nx) + b_n \sin(nx))$$

heißen **trigonometrische Polynome**.

Wie oben erwähnt, sollen 2π -periodische Funktionen analysiert werden.

Damit z.B. die Summe von zwei Sinus-Termen – etwa $A_1 \sin(x) + A_2 \sin(\omega x)$ mit Konstanten A_1, A_2 und ω – die zur gegebenen Funktion passende Periode 2π hat, muss nach obiger Bemerkung für jede Oberschwingung $\omega \in \mathbb{N}$ gelten, d.h. trigonometrische Reihen mit obigen **Frequenzen** sind die passende Grundlage der harmonischen Analyse.

Fragen. Präzisiert lauten die entscheidenden Frage dieses Kapitels:

- i) Lässt sich eine 2π -periodische Funktion in eine trigonometrische Reihe entwickeln?
- ii) Wenn ja, in welchem Sinne konvergiert die Reihe?

Eigenschaften trigonometrischer Funktionen.

Zur Vorbereitung benötigt man das folgende Lemma, welches als Übung nachzurechnen ist (vgl. Übungskapitel 8.3).

Lemma 8.1.**EIGENSCHAFTEN TRIGONOMETRISCHER FKT.**

i) Für $m, n \in \mathbb{N} \cup \{0\} =: \mathbb{N}_0$ gelten die *Orthogonalitätsbeziehungen*

$$\int_0^{2\pi} \cos(mx) \cos(nx) \, dx = 0 \quad \text{für } m \neq n ;$$

$$\int_0^{2\pi} \sin(mx) \sin(nx) \, dx = 0 \quad \text{für } m \neq n ;$$

$$\int_0^{2\pi} \cos(mx) \sin(nx) \, dx = 0 .$$

ii) Für $m, n \in \mathbb{N}_0$ gelten die *Normierungsbeziehungen*

$$\int_0^{2\pi} \cos^2(nx) \, dx = \begin{cases} \pi & \text{für } n \in \mathbb{N} , \\ 2\pi & \text{für } n = 0 ; \end{cases}$$

$$\int_0^{2\pi} \sin^2(nx) \, dx = \begin{cases} \pi & \text{für } n \in \mathbb{N} , \\ 0 & \text{für } n = 0 . \end{cases}$$

Zunächst wird nun **einfach angenommen**, dass eine formal gebildete Fourier-Reihe gleichmäßig gegen eine Funktion f konvergiert.

Die Fourier-Koeffizienten.

Unter dieser Konvergenzannahme folgt

Satz 8.1.**KOEFFIZIENTEN DER FORMALEN REIHE**

Konvergiert die Fourier-Reihe (die Folge der Partialsummen)

$$\frac{a_0}{2} + \sum_{n=1}^N (a_n \cos(nx) + b_n \sin(nx))$$

im Intervall $[0, 2\pi]$ **gleichmäßig** gegen eine Funktion $f(x)$, so ist f stetig und es gilt

$$a_n = \frac{1}{\pi} \int_0^{2\pi} f(x) \cos(nx) \, dx, \quad n = 0, 1, 2, \dots,$$

$$b_n = \frac{1}{\pi} \int_0^{2\pi} f(x) \sin(nx) \, dx, \quad n = 1, 2, \dots$$

Der Satz besagt also, dass aus der **angenommenen** gleichmäßigen Konvergenz der Reihe gegen eine Funktion f automatisch Darstellungsformeln für die Koeffizienten a_n und b_n folgen.

Beweisidee von Satz 8.1. Zunächst ist insbesondere mithilfe der gleichmäßigen Konvergenz zu verifizieren, dass im Beweisverlauf Integration und Summation vertauscht werden dürfen.

Ist das der Fall, so folgt aus der Darstellung

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos(nx) + b_n \sin(nx))$$

unmittelbar

$$\begin{aligned} & \int_0^{2\pi} f(x) \cos(kx) \, dx \\ &= \frac{a_0}{2} \int_0^{2\pi} \cos(kx) \, dx + \sum_{n=1}^{\infty} \int_0^{2\pi} (a_n \cos(nx) + b_n \sin(nx)) \cos(kx) \, dx \\ &= a_k \pi \end{aligned}$$

und der Satz ist bewiesen. □

Wie bereits betont, beantwortet der Satz **nicht** die Frage nach der Darstellbarkeit einer gegebenen Funktion f als Fourier-Reihe, da es keinerlei Rechtfertigung für die Voraussetzung „gleichmäßige Konvergenz“ in Satz 8.1 gibt.

Dennoch motiviert der Satz die nachfolgende Definition.

Definition 8.1. FOURIER-KOEFFIZIENTEN

Die reellwertige Funktion f sei auf $[0, 2\pi]$ definiert und integrierbar.

i) Die Zahlen

$$a_n = \frac{1}{\pi} \int_0^{2\pi} f(x) \cos(nx) \, dx, \quad n = 0, 1, 2, \dots,$$

$$b_n = \frac{1}{\pi} \int_0^{2\pi} f(x) \sin(nx) \, dx, \quad n = 1, 2, \dots$$

heißen *Fourier-Koeffizienten* von f .

ii) Die mit diesen Koeffizienten *formal* gebildete Reihe

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos(nx) + b_n \sin(nx))$$

heißt *Fourier-Reihe* von f .

Beispiel. Wie in Abbildung 8.8 angedeutet, sei $f: \mathbb{R} \rightarrow \mathbb{R}$ periodisch mit

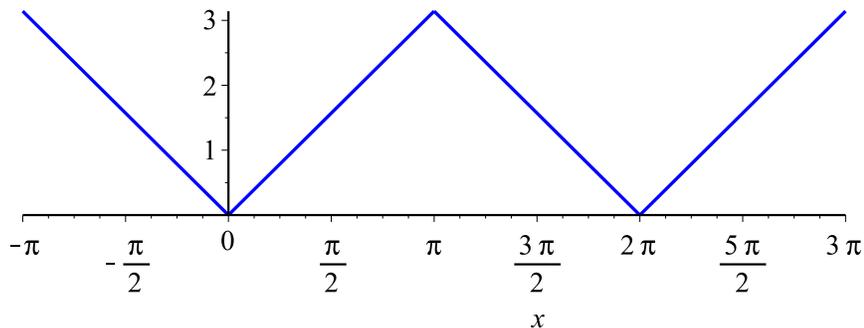


Abbildung 8.8: Die Zackenfunktion f .

der Periode 2π gegeben durch

$$f(x) = \begin{cases} x, & 0 \leq x \leq \pi, \\ 2\pi - x, & \pi \leq x \leq 2\pi. \end{cases}$$

Die Fourier-Koeffizienten lauten

$$a_0 = \pi, \quad a_n = \begin{cases} 0 & \text{für } n = 2m, \\ -\frac{4}{\pi n^2} & \text{für } n = 2m + 1, \end{cases} \quad n \in \mathbb{N},$$

$$b_n = 0, \quad n \in \mathbb{N}.$$

Die Fourier-Reihe ist gleichmäßig konvergent (siehe Übungskapitel 8.3).

Beobachtung. Schließlich sei noch angemerkt, dass sich die Berechnung der Koeffizienten für gerade bzw. ungerade Funktionen mithilfe der folgenden leicht zu verifizierenden Beobachtung vereinfachen läßt.

Ist f **gerade**, also $f(-x) = f(x)$ für alle $x \in \mathbb{R}$, so folgt

$$b_n = 0 \quad \text{für alle } n \in \mathbb{N},$$

$$a_n = \frac{2}{\pi} \int_0^\pi f(x) \cos(nx) \, dx \quad \text{für alle } n \in \mathbb{N}_0.$$

Ist f **ungerade**, also $f(-x) = -f(x)$ für alle $x \in \mathbb{R}$, so folgt

$$a_n = 0 \quad \text{für alle } n \in \mathbb{N}_0,$$

$$b_n = \frac{2}{\pi} \int_0^\pi f(x) \sin(nx) \, dx \quad \text{für alle } n \in \mathbb{N}.$$

8.2 Der Satz (Sägezahnfunktion; Gibbs-Phänomen; stückweise glatte Funktion; mittlere quadratische Abweichung; Parsevalsche Gleichung)

Das Musterbeispiel dieses Kapitels.

Es sei $f: \mathbb{R} \rightarrow \mathbb{R}$ periodisch mit Periode 2π und

$$f(x) = \pi - x \quad \text{für } 0 \leq x < 2\pi \quad (\text{Sägezahnfunktion}).$$

In diesem Fall ist f ungerade, also $a_n = 0$ für alle $n \in \mathbb{N}_0$ (wie f in den

Sprungstellen definiert ist, spielt für dieses Argument keine Rolle).

$$\begin{aligned}
 b_n &= \frac{1}{\pi} \int_0^{2\pi} f(x) \sin(nx) \, dx = \frac{1}{\pi} \int_0^{2\pi} (\pi - x) \sin(nx) \, dx \\
 &= -\frac{1}{\pi} \int_0^{2\pi} x \sin(nx) \, dx \\
 &= -\frac{1}{\pi} \left[x \left(-\frac{1}{n} \cos(nx) \right) \Big|_0^{2\pi} + \frac{1}{n} \int_0^{2\pi} \cos(nx) \, dx \right] = \frac{2}{n}.
 \end{aligned}$$

Die (formale) Fourier-Reihe lautet:

$$2 \left(\sin(x) + \frac{1}{2} \sin(2x) + \frac{1}{3} \sin(3x) + \dots \right) = 2 \sum_{n=1}^{\infty} \frac{\sin(nx)}{n}.$$

Zur Konvergenz der Reihe.

Betrachtet seien im Musterbeispiel die Partialsummen

$$S_N(x) := 2 \sum_{n=1}^N \frac{\sin(nx)}{n}$$

sowie die Differenz

$$R_N(x) := S_N(x) - f(x).$$

Man kann zeigen: Für $x \in (0, 2\pi)$ gilt

$$|R_N(x)| \leq \frac{4}{(2N+1) \sin(x/2)},$$

d.h.: Gleichmäßige Konvergenz gegen $f(x)$ gilt nur auf Teilintervallen der Form $[\varepsilon, 2\pi - \varepsilon]$ (weg von der Sprungstelle).

Für $x = 0, x = 2\pi, \dots$ ist $S_N = 0$, es liegt **keine Konvergenz der Reihe gegen die Funktion in den Sprungstellen** vor.

Man beobachtet aber, dass es sich hier genau um den **Mittelwert zwischen dem linksseitigen und dem rechtsseitigen Grenzwert** von f an der entsprechenden Stelle handelt.

Man kann weiter zeigen: Für $x_N = \pi/(N + 1/2)$ und für alle N hinreichend groß gilt

$$R_N(x_N) \stackrel{\approx}{\geq} 0.1789 \cdot \pi .$$

Mit anderen Worten: Auch wenn N sehr groß wird, **überschwingen** die Partialsummen den Wert π um ca. 18 % (**Gibbs-Phänomen** – vgl. insbesondere die Abbildungen 8.5 bis 8.7).

Eine geeignete Funktionenklasse und das Hauptresultat.

Bevor nun das Hauptergebnis dieses Kapitels formuliert werden kann, ist eine geeignete Funktionenklasse einzuführen.

Anhand von Abbildung 6.4 ist oben bereits erläutert, was unter einer **stückweise stetigen Funktion** zu verstehen ist (höchstens endlich viele **Sprungstellen**, einseitige Grenzwerte existieren in den Sprungstellen und am Rande des Intervalls).

Eine Funktion $f: [a, b] \rightarrow \mathbb{R}$ heißt **stückweise glatt**, falls sowohl f als auch f' stückweise stetig sind, wobei f' in den Sprungstellen von f nicht definiert sein muss.

Eine stückweise glatte Funktion hat also höchstens endlich viele Sprungstellen und endlich viele “Ecken”.

Die Bedeutung stückweise glatter Funktionen in den Anwendungen wird deutlich, wenn man etwa an Einschaltvorgänge denkt.

Satz 8.2. KONVERGENZ VON FOURIER-REIHEN

Die 2π -periodische Funktion f sei in $[0, 2\pi]$ stückweise glatt.

i) Dann konvergiert die Fourier-Reihe von f für alle $x \in \mathbb{R}$.

ii) Ist f *stetig an der Stelle x* , so gilt

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos(nx) + b_n \sin(nx)) .$$

iii) Ist x *eine Sprungstelle* von f , so gilt

$$\frac{f(x^+) + f(x^-)}{2} = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos(nx) + b_n \sin(nx)) .$$

iv) In jedem abgeschlossenen Intervall, in dem f stetig ist, ist die Reihe gleichmäßig konvergent (vgl. Gibbs Phänomen).

Beispiel. Im Beispiel am Ende von Kapitel 8.1 (Zackenfunktion) ist f stückweise glatt und stetig, also

$$f(x) = \left\{ \begin{array}{ll} x, & 0 \leq x < \pi \\ 2\pi - x, & \pi \leq x < 2\pi \end{array} \right\} = \frac{\pi}{2} - \frac{4}{\pi} \sum_{m=0}^{\infty} \frac{\cos((2m+1)x)}{(2m+1)^2} .$$

Insbesondere folgt mit der Wahl $x = 0$

$$\sum_{m=0}^{\infty} \frac{1}{(2m+1)^2} = \frac{\pi^2}{8} .$$

Eine Reihe von weiteren Beispielen wird im Übungskapitel 8.3 diskutiert.

Funktionen mit der Periode $T > 0$.

Die Diskussion von periodischen Funktionen f mit Periode $T > 0$, $T \neq 2\pi$, erfolgt analog (mittels der Transformation $u = 2\pi x/T$).

Die Fourier-Reihe lautet dann

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} \left(a_n \cos\left(\frac{2\pi}{T}nx\right) + b_n \sin\left(\frac{2\pi}{T}nx\right) \right)$$

mit den Fourier-Koeffizienten

$$a_n = \frac{2}{T} \int_0^T f(x) \cos\left(\frac{2\pi}{T}nx\right) dx, \quad n = 0, 1, 2, \dots,$$

$$b_n = \frac{2}{T} \int_0^T f(x) \sin\left(\frac{2\pi}{T}nx\right) dx, \quad n = 1, 2, \dots$$

Verschiebung des Integrationsintervalls.

Statt des Intervalls $[0, T]$ können analog Intervalle der Form $[a, a + T]$, $a \in \mathbb{R}$, betrachtet werden, d.h.

$$a_n = \frac{2}{T} \int_a^{a+T} f(x) \cos\left(\frac{2\pi}{T}nx\right) dx, \quad n = 0, 1, 2, \dots,$$

$$b_n = \frac{2}{T} \int_a^{a+T} f(x) \sin\left(\frac{2\pi}{T}nx\right) dx, \quad n = 1, 2, \dots$$

Vorsicht. Für die Sägezahnfunktion berechnet man beispielsweise

$$b_1 = \frac{1}{\pi} \int_0^{2\pi} (\pi - x) \sin(x) dx = 2.$$

Es gilt jedoch

$$\frac{1}{\pi} \int_{\pi}^{3\pi} (\pi - x) \sin(x) dx = -2 \neq b_1.$$

Handelt es sich hier um einen Widerspruch? (vgl. Übungskapitel 8.3)

Approximation im quadratischen Mittel.

Approximiert man eine 2π -periodische Funktion mit einem trigonometrischen Polynom

$$P_n(x) = \frac{a_0}{2} + \sum_{k=1}^n (\alpha_k \cos(kx) + \beta_k \sin(kx))$$

mit Koeffizienten derart, dass

$$Q_n := \int_0^{2\pi} [f(x) - P_n(x)]^2 dx$$

(mittlere quadratische Abweichung) minimal wird (Approximation im quadratischen Mittel), so erhält man genau die Fourier-Koeffizienten.

Die Parsevalsche Gleichung.

Gilt $\lim_{n \rightarrow \infty} Q_n = 0$ (konvergiert also die Fourier-Reihe im quadratischen Mittel gegen f), so gilt die Parsevalsche Gleichung

$$\frac{a_0^2}{2} + \sum_{k=1}^{\infty} (a_k^2 + b_k^2) = \frac{1}{\pi} \int_0^{2\pi} f^2(x) dx .$$

Man beachte: Konvergenz im quadratischen Mittel impliziert nicht punktweise Konvergenz.

Komplexe Notation.

Man setzt im Fall einer 2π -periodischen Funktion

$$c_n := \frac{1}{2\pi} \int_0^{2\pi} f(x) e^{-inx} dx .$$

Dabei beobachtet man zunächst

$$c_n = \frac{1}{2\pi} \int_0^{2\pi} f(x) \cos(nx) dx - \frac{i}{2\pi} \int_0^{2\pi} f(x) \sin(nx) dx ,$$

d.h. es gilt

$$c_n + c_{-n} = a_n \quad \text{und} \quad c_n - c_{-n} = -ib_n .$$

Man beobachtet weiter:

$$\begin{aligned} \sum_{n=-\infty}^{\infty} c_n e^{inx} &= \sum_{n=0}^{\infty} c_n (\cos(nx) + i \sin(nx)) \\ &\quad + \sum_{n=1}^{\infty} c_{-n} (\cos(nx) - i \sin(nx)) \\ &= c_0 \cos(x) + \sum_{n=1}^{\infty} (c_n + c_{-n}) \cos(nx) \\ &\quad + \sum_{n=1}^{\infty} i(c_n - c_{-n}) \sin(nx) \\ &= \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos(nx) + b_n \sin(nx)) , \end{aligned}$$

die Fourier-Reihe ist in komplexer Schreibweise dargestellt.

Hat f die Periode $T > 0$, so setzt man analog (in diesem Beispiel symmetrisch zum Ursprung)

$$c_n = \frac{1}{T} \int_{-T/2}^{T/2} f(x) e^{-in\frac{2\pi}{T}x} dx$$

und die Fourier-Reihe wird geschrieben als

$$\sum_{n=-\infty}^{\infty} c_n e^{in\frac{2\pi}{T}x}.$$

8.3 Übungsaufgaben zu Kapitel 8

Aufgabe 1. Begründen Sie:

i) Für beliebige Amplituden $A_1, A_2 \in \mathbb{R}$ ist die Funktion

$$A_1 \sin(\omega_1 x) + A_2 \sin(\omega_2 x)$$

periodisch, falls das Verhältnis ω_1/ω_2 der Kreisfrequenzen rational ist.

ii) Für beliebige Amplituden $A_1, A_2 \in \mathbb{R}$ ist die Funktion

$$A_1 \sin(x) + A_2 \sin(\omega x)$$

periodisch mit Periode 2π , falls $\omega \in \mathbb{Z}$, d.h. bei einer Fourier-Reihe kann man o.E. $\omega \in \mathbb{N}$ annehmen.

Aufgabe 2. Zeigen Sie Lemma 8.1.

Aufgabe 3. Bestimmen Sie die Fourier-Koeffizienten der Zackenfunktion.

Aufgabe 4. Es sei $f: \mathbb{R} \rightarrow \mathbb{R}$ die 2π -periodische Funktion mit $f(x) = x + \sin(x) - \pi$ für $0 \leq x < 2\pi$.

i) Skizzieren Sie f . Bestimmen Sie die explizite Abbildungsvorschrift $f(x) = \dots$ für $-2\pi \leq x < 0$. Ist f eine gerade bzw. eine ungerade Funktion?

ii) Bestimmen Sie die Fourier-Reihe von f .

iii) Gegen welche Funktion konvergiert die Fourier-Reihe von f ?

Aufgabe 5. Es sei $f: \mathbb{R} \rightarrow \mathbb{R}$ die 2π -periodische Funktion mit $f(x) = \frac{x}{2} - \pi$ für $0 \leq x < 2\pi$. Bestimmen Sie die Fourier-Reihe von f .

Aufgabe 6. Es sei $f: \mathbb{R} \rightarrow \mathbb{R}$,

$$f(x) = \frac{\pi^2 x}{12} \left(1 - \frac{x^2}{\pi^2}\right) \quad \text{für } -\pi \leq x < \pi$$

und $f(x) = f(x + 2\pi)$, d.h. f sei 2π -periodisch.

- i) Skizzieren Sie f . Deuten Sie dabei insbesondere die Lage der Extremwerte und Nullstellen von f an.
 - ii) Bestimmen Sie die Fourier-Reihe von f .
-

Aufgabe 7. Es sei $f: \mathbb{R} \rightarrow \mathbb{R}$ die Funktion mit der Periode 8, die gegeben ist durch

$$f(x) = \begin{cases} 2 - x & \text{für } 0 \leq x < 4, \\ x - 6 & \text{für } 4 \leq x < 8. \end{cases}$$

Skizzieren Sie f . Geben Sie die Fourier-Reihe von f an. Gegen welche Funktion konvergiert die Fourier-Reihe, wo konvergiert Sie gleichmäßig?

Aufgabe 8. Es sei $h: \mathbb{R} \rightarrow \mathbb{R}$ die 2π -periodische Funktion mit

$$h(x) = \begin{cases} 1 & \text{für } -\pi \leq x \leq -\pi/2, \\ 1/2 & \text{für } -\pi/2 < x < 0, \\ 0 & \text{für } x = 0, \\ -1 & \text{für } 0 < x < \pi. \end{cases}$$

Gegen welche Funktion konvergiert die Fourier-Reihen von h ?

Anhang A

Tensoren: Weitere Bemerkungen zur Definition

A.1 Kovariante und kontravariante Tensoren (kovariant; kontravariant; lineares Funktional; Dualraum; Multilinearform; Tensorprodukt; Tensor zweiter Stufe)

Im letzten Abschnitt von Kapitel 3.2 wurde bereits folgendes Szenario diskutiert:

Man betrachte eine lineare Abbildung $L: \mathbb{R}^m \rightarrow \mathbb{R}$, d.h. die darstellenden Matrizen bzgl. einer Basis \mathcal{A} bzw. einer Basis \mathcal{B} des \mathbb{R}^m sind Zeilenvektoren.

Die darstellende Matrix bzgl. der Basis \mathcal{A} sei

$$\left(\xi_1 \quad \dots \quad \xi_m \right)$$

und Satz 3.2 liefert die darstellende Matrix bzgl. der Basis \mathcal{B}

$$\left(\psi_1 \quad \dots \quad \psi_m \right) = \left(\xi_1 \quad \dots \quad \xi_m \right) \begin{pmatrix} \gamma_{11} & \dots & \gamma_{m1} \\ \vdots & & \vdots \\ \gamma_{1m} & \dots & \gamma_{mm} \end{pmatrix}$$

mit den Koeffizienten γ_{ij} , $1 \leq i, j \leq m$, des Matrixwechsels.

Aufgrund von

$$\psi_i = \sum_{j=1}^m \gamma_{ij} \xi_j,$$

handelt es sich um einen **kovarianten Tensor** erster Stufe.

Die linearen Abbildungen vom \mathbb{R}^m in die reelle Zahlen liefern das erste Beispiel eines **linearen Funktionals**.

Weitere lineare Funktionale?

In Verallgemeinerung des obigen Beispiels definiert man:

Definition A.1. LINEARES FUNKTIONAL

Es sei V ein Vektorraum über einem Körper \mathbb{K} .

Dann ist die Menge aller linearen Abbildungen

$$\varphi : V \rightarrow \mathbb{K}$$

selbst wieder ein Vektorraum über \mathbb{K} und heißt der zu V duale Vektorraum oder Dualraum von V .

Bezeichnung: V^* .

Jedes $\varphi \in V^*$ heißt *Linearform* oder *lineares Funktional* auf V .

Beispiel.

- i) Ist $V = \mathbb{R}^m$ und sind $a_1, \dots, a_m \in \mathbb{R}$ fixiert, so entsteht wie in der Einleitung dieses Anhangs beschrieben eine Linearform $\varphi \in (\mathbb{R}^m)^*$ aus der Vorschrift

$$\varphi : \mathbb{R}^m \rightarrow \mathbb{R}, \quad \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} \mapsto a_1 x_1 + \dots + a_m x_m.$$

Bezüglich der kanonischen Basis kann φ mit dem Zeilenvektor $(a_1 \dots a_m)$ identifiziert werden und es gilt

$$\varphi(\underline{\mathbf{x}}) = (a_1 \ a_2 \ \dots \ a_m) \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix} \in \mathbb{R} .$$

Wie ebenfalls bereits erwähnt, können die Elemente von $(\mathbb{R}^m)^*$ bzgl. einer beliebigen gegebenen Basis \mathcal{A} des \mathbb{R}^m als Zeilenvektoren der Länge m interpretiert werden und folglich hat der Dualraum $(\mathbb{R}^m)^*$ des \mathbb{R}^m die **gleiche Dimension** m wie der \mathbb{R}^m .

ii) Um sich eine Linearform auf $(\mathbb{R}^m)^*$ vorstellen zu können, d.h. eine Abbildung $\psi \in ((\mathbb{R}^m)^*)^*$,

$$\psi : (\mathbb{R}^m)^* \rightarrow \mathbb{R} \quad ((\mathbb{R}^m \rightarrow \mathbb{R}) \rightarrow \mathbb{R}) ,$$

identifiziert man einen Vektor $\underline{\mathbf{x}} \in \mathbb{R}^m$ mit der linearen Abbildung $\psi_{\underline{\mathbf{x}}}$ definiert durch

$$\psi_{\underline{\mathbf{x}}}(\varphi) := \varphi(\underline{\mathbf{x}}) \quad \text{für alle } \varphi \in (\mathbb{R}^m)^* .$$

Bemerkung. Von besonderem Interesse sind auch lineare Funktionale auf Funktionenräumen, was aber an dieser Stelle nicht vertieft werden soll.

Duale Basis.

Das obige Beispiel soll nun noch ein wenig systematisiert werden:

Mit der kanonischen Basis des \mathbb{R}^m existiert auch eine kanonische Basis des $(\mathbb{R}^m)^*$, die zugehörige **duale Basis**.

Sie ist gegeben durch

$$\begin{aligned}(\underline{\mathbf{e}}^{(1)})^T &= (1 \ 0 \ 0 \ \dots \ 0 \ 0) , \\(\underline{\mathbf{e}}^{(2)})^T &= (0 \ 1 \ 0 \ \dots \ 0 \ 0) , \\&\vdots \quad \quad \quad \vdots \\(\underline{\mathbf{e}}^{(m)})^T &= (0 \ 0 \ 0 \ \dots \ 0 \ 1) ,\end{aligned}$$

und jede Linearform auf \mathbb{R}^m (d.h. jeder Zeilenvektor – bei fixierter Basis) kann klarerweise als Linearkombination der $(\underline{\mathbf{e}}^{(i)})^T$, $1 \leq i \leq m$, geschrieben werden.

Man beachte, dass für alle $1 \leq i, j \leq m$ gilt

$$(\underline{\mathbf{e}}^{(j)})^T (\underline{\mathbf{e}}^{(i)}) = \delta_{ij} .$$

Die hier geschilderte Prozedur zur Konstruktion der dualen kanonischen Basis kann natürlich analog mit anderen Basen des \mathbb{R}^m durchgeführt werden.

Weiter sei analog zu *ii*) des obigen Beispiels festgehalten:

Ist V ein endlichdimensionaler Vektorraum, so kann der Bidualraum $(V^*)^*$ mit V selbst identifiziert werden.

Mit dieser Identifikation verhalten sich lineare Funktionale auf dem $(\mathbb{R}^m)^*$ wie der \mathbb{R}^m selbst.

Es sei an dieser Stelle an die Notation und die Koordinatentransformation aus Satz 3.2,

$$\beta_i = \sum_{j=1}^m \rho_{ij} \alpha_j , \quad (\rho_{ij}) = ((\gamma_{ij})^T)^{-1} , \quad i = 1, \dots, m ,$$

erinnert, man spricht von einem kontravarianten Tensor erster Stufe.

Das “konträre” Verhalten ist bereits in Kapitel 3.2 angedeutet.

Warum wird ein Tensor als (multi-) lineares Funktional definiert?

Wird beispielsweise ein Spannungstensor einfach als 3×3 -Matrix eingeführt, so ist er bzgl. einer gegebenen Basis des \mathbb{R}^3 definiert.

Eine andere Basis erfordert dabei eine andere Matrixdarstellung.

Eine physikalische Größe sollte aber **unabhängig von der zugrunde liegenden Basis sein** und deshalb wird von einer Matrixdarstellung zu Multilinearformen übergegangen.

Erinnerung. Der Begriff **Bilinearform** ($n = 2$) bzw. allgemein **n -Linearform** (**Multilinearform**) wurde im Paragraphen 2.2 erklärt.

Es handelt sich um eine Abbildung M von n Vektorräumen V_1, V_2, \dots, V_n , in den zugrunde liegenden Körper \mathbb{K} ,

$$M : V_1 \times \dots \times V_n \rightarrow \mathbb{K} ,$$

die **linear in jedem Argument ist**.

Mit diesen Vorbereitungen kann nun definiert werden, was allgemein unter einem **Tensor** zu verstehen ist.

Definition A.2. TENSOR

Gegeben sei ein endlichdimensionaler Vektorraum V über einem Körper \mathbb{K} mit dem Dualraum V^ von V .*

*Die Elemente der Menge $\mathcal{T}_q^p(V)$ der **Multilinearformen** M ,*

$$M : \underbrace{V \times V \times \dots \times V}_{q \text{ mal}} \times \underbrace{V^* \times V^* \times \dots \times V^*}_{p \text{ mal}} \rightarrow \mathbb{K} ,$$

*heißen **q -fach kovariante** und **p -fach kontravariante** Tensoren über V .*

Dabei setzt man $\mathcal{T}_0^0(V) = \mathbb{K}$.

Bemerkung. Die Begriffe kovariant und kontravariant sind wie in den Beispielen motiviert.

Konstruktion von Tensoren höherer Stufe.

Nach den Beispielen von Tensoren erster Stufe und der allgemeinen Definition eines Tensors beliebiger Stufe, zeigt nun das **Tensorprodukt** wie **Tensoren höherer Stufe** in natürlicher Weise produziert werden.

Dazu bezeichnen U und V zwei Vektorräume über einem Körper \mathbb{K} .

Jedes feste $u \in U$ kann als Linearform auf U^* interpretiert werden (wie bereits im obigen Beispiel *ii*)).

Die Vorschrift lautet:

$$u(u^*) := u^*(u) \quad \text{für alle } u^* \in U^* .$$

Man beachte, dass die rechte Seite wohl definiert ist. Analoges gilt natürlich auch für jedes feste $v \in V$.

Schließlich kann etwa auf $U^* \times V^*$ die Bilinearform $u \otimes v: U^* \times V^* \rightarrow \mathbb{K}$

$$(u \otimes v)(u^*, v^*) = u(u^*)v(v^*) \quad \text{für alle } u^* \in U^* , v^* \in V^* .$$

definiert werden.

Die Gesamtheit aller endlichen Summen von Termen dieser Form heißt das Tensorprodukt $U \otimes V$ und **besteht aus den Bilinearformen auf $U^* \times V^*$** .

Bemerkung. Analog wird $U_1 \otimes U_2 \otimes \cdots \otimes U_r$ definiert.

Beispiel. Es sei $(U^*)^* = U = \mathbb{R}^m$, $V = (\mathbb{R}^m)^*$ und die kanonische Basis des \mathbb{R}^m sei fixiert.

Dann kann $\underline{\mathbf{u}} \in \mathbb{R}^m$ als Spaltenvektor geschrieben werden, $\underline{\mathbf{v}}^T \in (\mathbb{R}^m)^*$ wird als Zeilenvektor interpretiert.

Damit wirkt $\underline{\mathbf{u}} \otimes \underline{\mathbf{v}}^T$ auf ein Paar $(\underline{\mathbf{x}}^T, \underline{\mathbf{y}})$ bestehend aus einem Zeilenvektor und einem Spaltenvektor.

Es ist

$$\underline{\mathbf{u}}(\underline{\mathbf{x}}^T) := (x_1 \ \dots \ x_m) \begin{pmatrix} u_1 \\ \vdots \\ u_m \end{pmatrix} \quad \text{für alle } \underline{\mathbf{x}}^T \in (\mathbb{R}^m)^* .$$

Analog ist

$$\underline{\mathbf{v}}^T(\underline{\mathbf{y}}) := (v_1 \ \dots \ v_m) \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} \quad \text{für alle } \underline{\mathbf{y}} \in \mathbb{R}^m .$$

Es ist also

$$\begin{aligned} \underline{\mathbf{u}} \otimes \underline{\mathbf{v}}^T(\underline{\mathbf{x}}^T, \underline{\mathbf{y}}) &= (x_1 \ \dots \ x_m) \begin{pmatrix} u_1 \\ \vdots \\ u_m \end{pmatrix} (v_1 \ \dots \ v_m) \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} \\ &= (x_1 \ \dots \ x_m) \left[\begin{pmatrix} u_1 \\ \vdots \\ u_m \end{pmatrix} (v_1 \ \dots \ v_m) \right] \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} \\ &= (x_1 \ \dots \ x_m) \begin{pmatrix} u_1 v_1 & \dots & u_1 v_m \\ \vdots & & \vdots \\ u_m v_1 & \dots & u_m v_m \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} . \end{aligned}$$

Auf diese Art wird der Tensor $\underline{\mathbf{u}} \otimes \underline{\mathbf{v}}^T$ zweiter Stufe für die fixierte kanonische Basis dargestellt.

Als Übungsaufgabe überlege man sich, wie der Tensor im Sinne von Definition A.2 zu bezeichnen ist.

Mit diesen Bemerkungen soll die (teils heuristische) Diskussion von Tensoren abgeschlossen werden.

Anhang B

Konditionierung einer numerischen Aufgabe – Stabilität eines Algorithmus

B.1 Konditionierung einer numerischen Aufgabe (differenzielle Fehleranalyse; Landausche Symbole; relative Konditionszahlen; Problemfehler)

Eine **numerische Aufgabe** ist die Berechnung der **Ausgabedaten**

$$\underline{y} = f(\underline{x})$$

bei einer gegebenen Funktion $f: \mathbb{R}^m \supset U \rightarrow \mathbb{R}^n$ aus den **Eingabedaten** $x \in U$.

In der Praxis involviert dabei die Berechnung von $f(\underline{x})$ in der Regel eine Vielzahl von Auswertungen.

Einfaches Beispiel. Es sei $f: \mathbb{R}^2 \rightarrow \mathbb{R}, \underline{x} \mapsto y = x_1 + x_2$.

Was versteht man unter der Konditionierung einer numerischen Aufgabe?

Eine numerische Aufgabe heißt **gut konditioniert**, falls “kleine” Störungen der Eingabedaten nur “kleine” Störungen des Ergebnisses (der Ausgabedaten) verursachen.

Beispiel. Die Eingabe in eine numerische Aufgabe sei einmal

$$\underline{\mathbf{b}} = \begin{pmatrix} 3.7041 \\ 1.8519 \end{pmatrix} .$$

Eine zweite Eingabe sei

$$\tilde{\underline{\mathbf{b}}} = \begin{pmatrix} 3.7040 \\ 1.8518 \end{pmatrix} .$$

Die numerische Aufgabe selbst laute: Man löse das Gleichungssystem

$$\begin{pmatrix} 2.4692 & 1.2345 \\ 1.2345 & 0.6172 \end{pmatrix} \mathbf{Ausgabedaten} = \mathbf{Eingabedaten} ,$$

wobei die Eingabedaten einmal durch obiges $\underline{\mathbf{b}}$, bei der zweiten Berechnung durch $\tilde{\underline{\mathbf{b}}}$ gegeben seien.

Die Ausgabedaten berechnen sich zu

$$\underline{\mathbf{y}} = \begin{pmatrix} 3 \\ -3 \end{pmatrix} \quad \text{bzw.} \quad \begin{pmatrix} 12344 \\ -6170 \end{pmatrix} .$$

Dabei ist zu betonen, dass die Ausgabedaten **exakt berechnet sind und keine numerischen Näherungen darstellen**.

Beobachtung. Störungen der Größenordnung 10^{-4} in der Eingabe führen in dem Beispiel zu Störungen der Größenordnung 10^4 in der Ausgabe.

Auch bei exakter Rechnung machen kleine Störungen in der Eingabe das Ergebnis völlig unsicher.

In dem konkreten Beispiel berechne man die Determinante der obigen Matrix und interpretiere deren Einfluss auf die numerische Aufgabe.

Um die Auswirkungen kleiner Fehler zu quantifizieren, arbeitet man bei der sogenannten **differentiellen Fehleranalyse** meist mit:

Definition B.1. LANDAUSCHE SYMBOLE

Es sei $U \subset \mathbb{R}^m$ eine Umgebung der Null, $g: U \rightarrow \mathbb{R}^n$ und $h: U \rightarrow \mathbb{R}^l$.

i) Man schreibt

$$g(\underline{\mathbf{x}}) = O(h(\underline{\mathbf{x}})) \quad \text{für } \underline{\mathbf{x}} \rightarrow \underline{\mathbf{0}},$$

falls $\varepsilon, C > 0$ existieren mit

$$\|g(\underline{\mathbf{x}})\| \leq C\|h(\underline{\mathbf{x}})\| \quad \text{für alle } \underline{\mathbf{x}} \in U \text{ mit } \|\underline{\mathbf{x}}\| < \varepsilon.$$

ii) Man schreibt

$$g(\underline{\mathbf{x}}) = o(h(\underline{\mathbf{x}})) \quad \text{für } \underline{\mathbf{x}} \rightarrow \underline{\mathbf{0}},$$

falls es ein $\varepsilon > 0$ sowie eine Funktion $\varphi: U \rightarrow [0, \infty)$ gibt mit $\varphi(\underline{\mathbf{x}}) \rightarrow \underline{\mathbf{0}}$ für $\underline{\mathbf{x}} \rightarrow \underline{\mathbf{0}}$ und

$$\|g(\underline{\mathbf{x}})\| \leq \varphi(\underline{\mathbf{x}})\|h(\underline{\mathbf{x}})\| \quad \text{für alle } \underline{\mathbf{x}} \in U \text{ mit } \|\underline{\mathbf{x}}\| < \varepsilon.$$

iii) Ist $f: U \rightarrow \mathbb{R}^n$ eine weitere Funktion, so setzt man ($o(h(\underline{\mathbf{x}}))$ analog)

$$f(\underline{\mathbf{x}}) = g(\underline{\mathbf{x}}) + O(h(\underline{\mathbf{x}})) \quad :\Leftrightarrow \quad f(\underline{\mathbf{x}}) - g(\underline{\mathbf{x}}) = O(h(\underline{\mathbf{x}})).$$

Bemerkung. Analog können auch Grenzübergänge $\underline{\mathbf{x}} \rightarrow \underline{\mathbf{x}}_0$ für beliebiges $\underline{\mathbf{x}}_0 \in U$ betrachtet werden.

Beispiele. Es ist ($m = n = l = 1$)

$$\begin{aligned} \sin(x) &= O(x), \\ \cos(x) &= O(1), \\ x^3 &= o(x^2), \\ \cos^2(x) - 1 &= o(x). \end{aligned}$$

Zur differentiellen Fehleranalyse einer numerischen Aufgabe.

Man betrachte zunächst den einfachsten Fall $m = n = 1$:

Für $f: \mathbb{R} \supset [a, b] \rightarrow \mathbb{R}$, sei die numerische Aufgabe der Berechnung von $y = f(x)$ näher analysiert.

Dabei sei $f \in C^2([a, b])$ und $x, x + \Delta x \in (a, b)$, wobei Δx eine Störung der Eingabedaten (aufgrund von Rundungsfehlern etc.) bezeichne.

Aus dem Taylorschen Satz (Lagrangesche Restglieddarstellung nach Korollar 7.1) folgt für die korrespondierende Störung Δy der Ausgabedaten

$$\begin{aligned} \Delta y &:= f(x + \Delta x) - f(x) \\ &= f'(x)\Delta x + R_1(\Delta x), \end{aligned}$$

mit

$$|R_1(\Delta x)| \leq \frac{1}{2} |\Delta x|^2 \sup_{0 < \theta < 1} f''(x + \theta \Delta x).$$

Nach Voraussetzung ist f von der Klasse C^2 und für eine Konstante c gilt

$$\sup_{0 < \theta < 1} f''(x + \theta \Delta x) \leq c,$$

es kann geschätzt werden

$$|R_1(\Delta x)| \leq \frac{c}{2} |\Delta x|^2,$$

d.h. für $|\Delta x| \rightarrow 0$ gilt

$$\Delta y = f'(x)\Delta x + O(|\Delta x|^2).$$

Ist allgemein $f: \mathbb{R}^m \subset U \rightarrow \mathbb{R}^n$, so folgt analog für alle $j = 1, \dots, n$

$$\Delta y_j = \sum_{i=1}^m \frac{\partial f_j}{\partial x_i}(\underline{\mathbf{x}}) \Delta x_i + O(|\Delta \underline{\mathbf{x}}|^2).$$

Dabei setzt man

$$f(\underline{\mathbf{x}}) = \begin{pmatrix} f_1(\underline{\mathbf{x}}) \\ \vdots \\ f_n(\underline{\mathbf{x}}) \end{pmatrix}$$

und betrachtet für $j = 1, \dots, n$ die Funktionen $f_j: U \rightarrow \mathbb{R}$ unabhängig voneinander.

Weiter betrachtet man die **partiellen Ableitungen** ($i = 1, \dots, m, j = 1, \dots, n$)

$$\frac{\partial f_j}{\partial x_i}(\underline{\mathbf{x}}),$$

deren systematisches Studium in Teil III, Kapitel 6.1 folgt.

In diesem Anhang ist die folgende Rechenregel völlig ausreichend:

Die i^{te} partielle Ableitung ($i = 1, \dots, m$) einer stetig differenzierbaren Funktion $g: \mathbb{R}^m \subset U \rightarrow \mathbb{R}$ berechnet man, indem die $x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_m$ als feste Konstanten interpretiert werden und nur nach der Variablen x_i abgeleitet wird.

Ist also beispielsweise $g: \mathbb{R}^m \rightarrow \mathbb{R}$ gegeben durch

$$g(\underline{\mathbf{x}}) = \|\underline{\mathbf{x}}\|^2 = x_1^2 + \dots + x_{i-1}^2 + x_i^2 + x_{i+1}^2 + \dots + x_m^2,$$

so ist für fixiertes i ($i = 1, \dots, m$)

$$\frac{\partial g}{\partial x_i}(\underline{\mathbf{x}}) = 2x_i.$$

Schließlich sei für $x_i \neq 0, i = 1, \dots, m$, und für $y_j \neq 0, j = 1, \dots, n$,

$$\varepsilon_{x_i} := \left| \frac{\Delta x_i}{x_i} \right| \quad (\text{relativer Fehler von } x_i),$$

$$\varepsilon_{y_j} := \left| \frac{\Delta y_j}{y_j} \right| \quad (\text{relativer Fehler von } y_j),$$

$$K_{ij}(\underline{\mathbf{x}}) := \left| \frac{\partial f_j}{\partial x_i}(\underline{\mathbf{x}}) \frac{x_i}{y_j} \right| \quad (\text{relative Konditionszahlen von } f \text{ in } \underline{\mathbf{x}}),$$

und es ist gezeigt

$$\begin{aligned}\varepsilon_{y_j} &\leq \sum_{i=1}^m \left| \frac{\partial f_j}{\partial x_i}(\underline{\mathbf{x}}) \frac{x_i}{y_j} \right| \left| \frac{\Delta x_i}{x_i} \right| + O(|\Delta \mathbf{x}|^2) \\ &= \sum_{i=1}^m K_{ij}(\underline{\mathbf{x}}) \varepsilon_{x_i} + O(|\Delta \underline{\mathbf{x}}|^2) .\end{aligned}$$

Interpretation. Die Zahl $K_{ij}(\underline{\mathbf{x}})$ gibt an, wie stark sich der relative Fehler ε_{x_i} der i^{ten} Komponente der Eingabegröße $\underline{\mathbf{x}}$ auf den relativen Fehler ε_{y_j} der j^{ten} Komponente der Ausgabegröße $\underline{\mathbf{y}}$ auswirkt.

Gibt es Indizes $1 \leq i \leq m$, $1 \leq j \leq n$ mit $K_{ij} \gg 1$, so spricht man von einem **schlecht konditionierten Problem**, andernfalls heißt das Problem **gut konditioniert**.

Der hier besprochene Fehler heißt der **Problemfehler**, der unabhängig vom konkret eingesetzten Algorithmus zur Lösung der Aufgabe ist.

Bemerkung. Die Bedingung $K_{ij} \gg 1$ bedeutet, dass K_{ij} wirklich um Größenordnungen größer als 1 ist, sind etwa alle K_{ij} ungefähr 2 (oder 3.165), so heißt das Problem nach wie vor gut konditioniert.

Beispiele.

i) Addition: Es sei $f: \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(\underline{\mathbf{x}}) = x_1 + x_2$. Dann gilt

$$K_1(\underline{\mathbf{x}}) = \left| \frac{x_1}{x_1 + x_2} \right| ,$$

$$K_2(\underline{\mathbf{x}}) = \left| \frac{x_2}{x_1 + x_2} \right| .$$

Haben also x_1 und x_2 gleiches Vorzeichen, so ist das Problem gut konditioniert.

Gilt dagegen (**Auslöschung**)

$$x_1 \approx -x_2 ,$$

so ist das Problem schlecht konditioniert: Kleine relative Eingabefehler können ähnlich wie oben zu sehr großen relativen Fehlern im Ergebnis führen.

ii) Multiplikation: Es sei $f: \mathbb{R}^2 \rightarrow \mathbb{R}$, $\underline{\mathbf{x}} \mapsto x_1 x_2$. Dann gilt

$$K_1 = 1 = K_2 ,$$

die Multiplikation ist **stets gut konditioniert**.

B.2 Stabilität eines Algorithmus (gutartiger Algorithmus)

Der im letzten Paragraphen diskutierte Problemfehler ist systematischer Natur und kann nicht abgewendet werden.

Zur konkreten Berechnung einer numerischen Aufgabe können in der Regel aber vielzählige unterschiedliche Algorithmen eingesetzt werden, die sich durch die Art und die Reihenfolge der auftretenden Gleitpunktoperationen unterscheiden.

Dabei werden sich unterschiedliche (Rundungs-) Fehler ins Ergebnis fort-pflanzen und dort widerspiegeln.

Wie hängt der Fehler vom speziell eingesetzten Algorithmus ab?

Bezeichnungen.

i) Gegeben sei die numerische Aufgabe der Berechnung von $\underline{\mathbf{y}} = f(\underline{\mathbf{x}})$.

Ein Verfahren oder **Algorithmus** zur (näherungsweise) Berechnung von $\underline{\mathbf{y}}$ aus $\underline{\mathbf{x}}$ ist eine **endliche** (manchmal auch abzählbar unendliche) Folge von “elementaren” Abbildungen $\varphi^{(k)}$, die nacheinander angewendet auf $\underline{\mathbf{x}}$ einen Näherungswert $\underline{\tilde{\mathbf{y}}}$ für $\underline{\mathbf{y}}$ liefern:

$$\underline{\mathbf{x}} = \underline{\mathbf{x}}^{(0)} \mapsto \varphi^{(1)}(\underline{\mathbf{x}}^{(0)}) = \underline{\mathbf{x}}^{(1)} \mapsto \dots \mapsto \varphi^{(k)}(\underline{\mathbf{x}}^{(k-1)}) = \underline{\mathbf{x}}^{(k)} = \underline{\tilde{\mathbf{y}}} .$$

Im einfachsten Fall ist $\varphi^{(k)} \in \{+, -, \cdot, /\}$ für alle k .

- ii) Bei einem **gutartigen Algorithmus** (oder **gut konditionierten Algorithmus**) ist der Beitrag des Rundungsfehlers (durch die Maschinenoperationen des Algorithmus) zum **Gesamtfehler höchstens von der Größenordnung des Problemfehlers**.
- iii) Die **numerische Mathematik** befasst sich mit der Suche nach gutartigen Lösungsalgorithmen.

Beispiel. Zu berechnen sei ($f: \mathbb{R}^2 \rightarrow \mathbb{R}$)

$$y = f(\underline{\mathbf{x}}) = x_1^2 - x_2^2 = (x_1 + x_2)(x_1 - x_2) .$$

Bevor konkrete Algorithmen auf ihre Gutartigkeit untersucht werden können, ist der Problemfehler abzuschätzen.

Es ist

$$\begin{aligned} \varepsilon_y = \left| \frac{\Delta y}{y} \right| &\leq \sum_{i=1}^2 K_i(\underline{\mathbf{x}}) \left| \frac{\Delta x_i}{x_i} \right| + O(|\Delta \underline{\mathbf{x}}|^2) \\ &= \sum_{i=1}^2 \left| 2x_i \frac{x_i}{x_1^2 - x_2^2} \right| \left| \frac{\Delta x_i}{x_i} \right| + O(|\Delta \underline{\mathbf{x}}|^2) \\ &\leq 2 \left| \frac{x_1^2 + x_2^2}{x_1^2 - x_2^2} \right| \text{eps} + O(\text{eps}^2) \\ &= 2 \left| \frac{(x_1/x_2)^2 + 1}{(x_1/x_2)^2 - 1} \right| \text{eps} + O(\text{eps}^2) . \end{aligned}$$

Dabei wurde (wie es im Folgenden auch getan wird) angenommen, dass relative Eingabefehler durch die Maschinengenauigkeit eps (vgl. Teil I, Kapitel 7.3) abgeschätzt werden können.

Zur konkreten Lösung der numerischen Aufgabe werden nun zwei unterschiedliche Algorithmen analysiert.

Algorithmus I.

$$\left. \begin{array}{l} x_1 \mapsto x_1 \odot x_1 \\ x_2 \mapsto x_2 \odot x_2 \end{array} \right\} \mapsto (x_1 \odot x_1) \ominus (x_2 \odot x_2) =: \tilde{y} .$$

Algorithmus II.

$$\left. \begin{array}{l} \underline{x} \mapsto x_1 \ominus x_2 \\ \underline{x} \mapsto x_1 \oplus x_2 \end{array} \right\} \mapsto (x_1 \ominus x_2) \odot (x_1 \oplus x_2) =: \tilde{y} .$$

Zum Vergleich der beiden Algorithmen sei daran erinnert (siehe wieder Teil I, Kapitel 7.3), dass die exakten Rechenoperationen durch Maschinenoperationen \oplus , \ominus , \odot , \oslash realisiert werden, die fehlerbehaftet sind.

Es gilt beispielsweise (alle anderen Operationen analog)

$$x_1 \oplus x_2 = (x_1 + x_2)(1 + \varepsilon) ,$$

wobei ε zwar von x_1 , x_2 abhängt, der Absolutbetrag ist jedoch (für x_1 , x_2 , $x_1 + x_2 \in D$, was im Folgenden stets angenommen wird) durch eps beschränkt.

Zum Algorithmus I.

Hier ergibt sich mit $|\varepsilon_1|, |\varepsilon_2|, |\varepsilon_3| < \text{eps}$ (wie oben angedeutet müssen die ε_i unterschieden werden)

$$\begin{aligned} \tilde{y} &= (x_1 \odot x_1) \ominus (x_2 \odot x_2) \\ &= [x_1^2(1 + \varepsilon_1) - x_2^2(1 + \varepsilon_2)](1 + \varepsilon_3) \\ &= \underbrace{x_1^2 - x_2^2}_{=y} + x_1^2\varepsilon_1 - x_2^2\varepsilon_2 + (x_1^2 - x_2^2)\varepsilon_3 + O(\text{eps}^2) . \end{aligned}$$

Der **relative Rundungsfehler durch die eingesetzten Maschinenoperationen**

ist abgeschätzt durch

$$\begin{aligned} \left| \frac{\Delta y}{y} \right| &= \left| \frac{y - \tilde{y}}{y} \right| \\ &\leq \frac{x_1^2 + x_2^2 + |x_1^2 - x_2^2|}{|x_1^2 - x_2^2|} \text{eps} + O(\text{eps}^2) \\ &\leq \left[1 + \left| \frac{(x_1/x_2)^2 + 1}{(x_1/x_2)^2 - 1} \right| \right] \text{eps} + O(\text{eps}^2) . \end{aligned}$$

Dabei ist $O(\text{eps}^2)$ “sehr klein” und kann vernachlässigt werden.

Hinzu kommt ein Term proportional zu (im Beispiel sogar gleich) eps.

Solche Terme sind stets gutartig, da der Fehler nie unterhalb der Maschinengenauigkeit abgeschätzt werden kann.

Der verbleibende Term

$$\left| \frac{(x_1/x_2)^2 + 1}{(x_1/x_2)^2 - 1} \right| \text{eps}$$

entspricht (bis auf einen konstanten Faktor) genau dem Problemfehler.

Insgesamt ist der durch den Algorithmus bedingte Beitrag zum Gesamtfehler von der Größenordnung des Problemfehlers, der Algorithmus ist gutartig.

Zum Algorithmus II.

Hier ist mit $|\varepsilon_i| < \text{eps}$, $i = 1, 2, 3$,

$$\begin{aligned} \tilde{y} &= [(x_1 - x_2)(1 + \varepsilon_1) (x_1 + x_2)(1 + \varepsilon_2)](1 + \varepsilon_3) \\ &= y(1 + \varepsilon_1)(1 + \varepsilon_2)(1 + \varepsilon_3) \\ &= y(1 + \varepsilon_1 + \varepsilon_2 + \varepsilon_3) + O(\text{eps}^2) , \end{aligned}$$

d.h.

$$\begin{aligned} \left| \frac{\Delta y}{y} \right| &= \left| \frac{y - \tilde{y}}{y} \right| \\ &= |\varepsilon_1 + \varepsilon_2 + \varepsilon_3| + O(\text{eps}^2) \\ &\leq 3\text{eps} + O(\text{eps}^2) . \end{aligned}$$

Dieser relative Fehler ist immer von der Größenordnung der Maschinengenauigkeit, der Algorithmus ist auf jeden Fall gutartig.

Im Gegensatz zum ersten Algorithmus tritt auch für $x_1 \approx x_2$ keine **Auslöschung** auf und der zweite Algorithmus ist zu bevorzugen.

Das obige Beispiel motiviert die erste der folgenden drei Orientierungshilfen zur konkreten Konstruktion eines gutartigen Algorithmus:

Faustregeln.

- i)* Numerisch schlecht konditionierte Operationen sollten möglichst früh im Algorithmus ausgeführt werden.
- ii)* Bei der Lösung quadratischer Gleichungen sollten nicht beide Wurzeln aus der Lösungsformel berechnet werden.
- iii)* Die Auswertung von Polynomen sollte mit Hilfe des Horner-Schemas erfolgen.

Ein Zahlenbeispiel.

Betrachtet sei wie oben die numerische Aufgabe der Berechnung von

$$y = f(\mathbf{x}) = x_1^2 - x_2^2 .$$

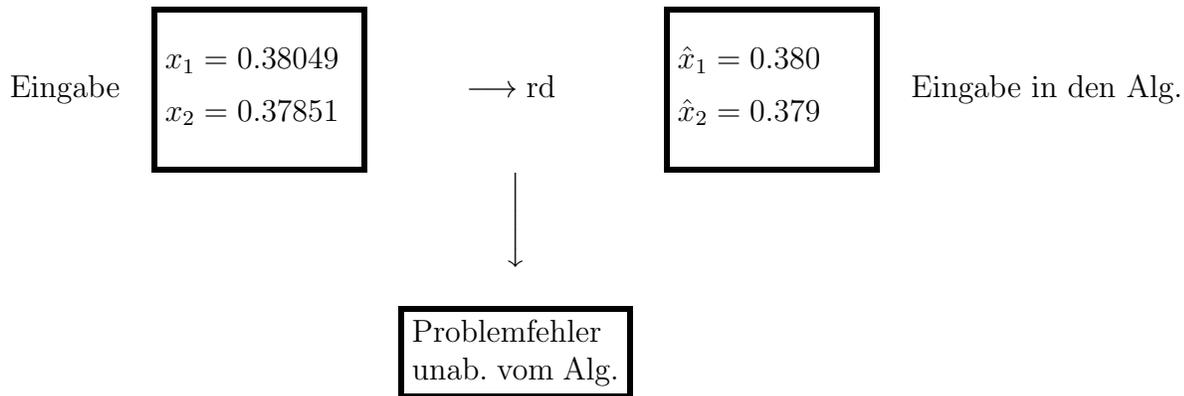


Tabelle B.1: Zum Fehler bei der Berechnung einer numerischen Aufgabe.

Die **Mantissenlänge** einer konkreten Rechenmaschine sei der Einfachheit wegen $t = 3$, was natürlich keine praxisnahe Annahme ist.

Die Eingabedaten seien

$$x_1 = 0.38049 \quad \text{und} \quad x_2 = 0.37851 ,$$

die Situation ist in Tabelle B.1 verdeutlicht.

Die **exakten** Eingabedaten x_1, x_2 werden von der Maschine gerundet, die relativen Fehler $\varepsilon_{x_1}, \varepsilon_{x_2} \approx 0.001$ bedingen einen relativen Problemfehler $\varepsilon_y \approx 0.5$.

Hier wurde ε_y explizit mittels des Vergleichs von $0.380^2 - 0.379^2$ und $0.38049^2 - 0.37851^2$ berechnet.

Man beachte, dass sich die Konditionszahlen etwa zu 190 berechnen, die theoretisch berechnete obere Schranke für ε_y ist also in guter Übereinstimmung mit der expliziten Rechnung

$$\varepsilon_y \approx 2 \cdot 190 \cdot \underbrace{5 \cdot 10^{-3}}_{\text{eps}} \approx 2 .$$

Die Werte $\hat{x}_1 = 0.380, \hat{x}_2 = 0.379$ sind die Eingabedaten für den berechnenden Algorithmus.

Algorithmus I. Im ersten Schritt wird berechnet

$$\left. \begin{array}{l} \hat{x}_1 \\ \hat{x}_2 \end{array} \right\} \mapsto \begin{cases} u = \hat{x}_1 \odot \hat{x}_1 = \text{rd}(\hat{x}_1 \hat{x}_1) = \text{rd}(0.1444) = 0.144, \\ v = \hat{x}_2 \odot \hat{x}_2 = \text{rd}(\hat{x}_2 \hat{x}_2) = \text{rd}(0.143641) = 0.144. \end{cases}$$

Dabei werden (wie bereits erwähnt) $\hat{x}_1 \hat{x}_1$ und $\hat{x}_2 \hat{x}_2$ zunächst intern mit erhöhter Stellenzahl berechnet und dann gerundet.

Es verbleibt die Operation

$$\left. \begin{array}{l} u \\ v \end{array} \right\} \mapsto u - v = 0.$$

Die fehlerbehafteten Werte u und v dienen hier als Eingabe in den schlecht konditionierten Teil des Algorithmus.

Insgesamt ist der durch den Algorithmus bedingte relative Rundungsfehler

$$\left| \frac{0 - (\hat{x}_1^2 - \hat{x}_2^2)}{\hat{x}_1^2 - \hat{x}_2^2} \right| = 1.$$

Dieser ist zwar von der gleichen Größenordnung wie der Problemfehler (der Algorithmus ist gutartig), trotzdem ist er sehr groß.

Algorithmus II. Der zweite Algorithmus beginnt mit

$$\left. \begin{array}{l} \tilde{x}_1 \\ \tilde{x}_2 \end{array} \right\} \mapsto \begin{cases} u = \hat{x}_1 \ominus \hat{x}_2 = \text{rd}(\hat{x}_1 - \hat{x}_2) = \text{rd}(0.001) = 0.001, \\ v = \hat{x}_1 \oplus \hat{x}_2 = \text{rd}(\hat{x}_1 + \hat{x}_2) = \text{rd}(0.759) = 0.759. \end{cases}$$

Der letzte Schritt ist in diesem Algorithmus

$$\left. \begin{array}{l} u \\ v \end{array} \right\} \mapsto u \odot v = \text{rd}(uv) = \text{rd}(0.759 \cdot 10^{-3}) = 0.759 \cdot 10^{-3}.$$

Im gewählten Zahlenbeispiel liefert der zweite Algorithmus sogar die exakte Lösung der vom Algorithmus zu berechnenden Aufgabe $y = \hat{x}_1^2 - \hat{x}_2^2$.

Der wesentliche Vorteil des zweiten Algorithmus ist, dass als Eingabe für den schlecht konditionierten Teil die exakte Algorithmuseingabe dient.

Bemerkung. Natürlich können auch im Verlauf von Algorithmus II (kleine) Fehler auftreten. Ist etwa

$$\hat{x}_1 = 0.106 \cdot 10^2, \quad \hat{x}_2 = 0.612,$$

so ist

$$\hat{x}_1 \oplus \hat{x}_2 = \text{rd} (0.11212 \cdot 10^2) = 0.112 \cdot 10^2.$$

B.3 Übungsaufgaben zum Anhang

Aufgabe 1. Schreiben Sie f jeweils in der Form $f(x) = O(x^m)$ und $f(x) = o(x^m)$ mit maximalem $m \in \mathbb{N}_0$, falls

i) $f(x) = 2x^3 + (\cos(x) - 1)$;

ii) $f(x) = x(x^2 + \frac{1}{1+x^2})^4 + 2x^6$;

iii) $f(x) = \ln(1 + x^2)$.

Aufgabe 2.* Es sei $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ gegeben durch die Vorschrift

$$\underline{\mathbf{x}} \mapsto x_1^2(x_1 - x_2) .$$

i) Betrachten Sie die numerische Aufgabe der Berechnung von $y = f(\underline{\mathbf{x}})$ und berechnen Sie die relativen Konditionszahlen. Ist das Problem stets gut konditioniert?

ii) Ist zur Lösung der Aufgabe der Algorithmus

$$u = x_1 \odot x_1 , \quad v = x_1 \ominus x_2 , \quad \tilde{y} = u \odot v$$

gutartig?

Aufgabe 3.* Zur Berechnung der numerischen Aufgabe $y = f(\underline{\mathbf{x}})$, $f: \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(\underline{\mathbf{x}}) = (x_1 + x_2)x_2$ betrachte man die beiden Algorithmen

$$u = x_1 \oplus x_2 , \quad \tilde{y} = u \odot x_2 ;$$

$$u = x_1 \odot x_2 , \quad v = x_2 \odot x_2 , \quad \tilde{y} = u \oplus v .$$

Welcher ist zu bevorzugen?

Aufgabe 4.* Es sei $f: \mathbb{R}^3 \rightarrow \mathbb{R}$ gegeben durch

$$\underline{\mathbf{x}} \mapsto x_1 - x_2x_3 .$$

Betrachten Sie die numerische Aufgabe der Berechnung von $y = f(\underline{\mathbf{x}})$ sowie den Algorithmus

$$u = x_2 \odot x_3, \quad \tilde{y} = x_1 \ominus u.$$

Ist der Algorithmus gutartig?

Aufgabe 5. Finden Sie eine numerische Aufgabe und einen zugehörigen Algorithmus, der nicht gutartig ist.

Lösungshinweise zu den Übungsaufgaben.**Aufgabe 2.**

i) Es ist

$$\begin{aligned} K_1(\underline{\mathbf{x}}) &= \left| \frac{\partial f}{\partial x_1}(\underline{\mathbf{x}}) \frac{x_1}{y} \right| \\ &= \left| (2x_1(x_1 - x_2) + x_1^2) \frac{x_1}{x_1^2(x_1 - x_2)} \right| \\ &= \left| 2 + \frac{x_1}{x_1 - x_2} \right| \end{aligned}$$

und

$$K_2(\underline{\mathbf{x}}) = \left| x_1^2 \frac{x_2}{x_1^2(x_1 - x_2)} \right| = \left| \frac{x_2}{x_1 - x_2} \right|.$$

Das Problem ist also nicht stets gut konditioniert, wie der Fall $x_1 \approx x_2$ zeigt.

ii) Es ist

$$\begin{aligned} u &= x_1 \odot x_1 = x_1^2(1 + \varepsilon_1), \\ v &= x_1 \ominus x_2 = (x_1 - x_2)(1 + \varepsilon_2), \\ \tilde{y} &= u \odot v = (x_1^2(1 + \varepsilon_1)(x_1 - x_2)(1 + \varepsilon_2))(1 + \varepsilon_2), \end{aligned}$$

d.h. man hat

$$\begin{aligned} \tilde{y} &= x_1^2(x_1 - x_2)(1 + \varepsilon_1)(1 + \varepsilon_2)(1 + \varepsilon_3) \\ &= y(1 + \varepsilon_1 + \varepsilon_2 + \varepsilon_3) + O(\text{eps}^2) \end{aligned}$$

und damit

$$\left| \frac{\tilde{y} - y}{y} \right| \leq 3\text{eps} + O(\text{eps}^2).$$

Der Algorithmus ist gutartig, da der relative Fehler von der Größenordnung der Maschinengenauigkeit ist.

Aufgabe 3. Beim ersten Algorithmus ist

$$((x_1 + x_2)(1 + \varepsilon_1))x_2(1 + \varepsilon_2) = y(1 + \varepsilon_1 + \varepsilon_2) + O(\text{eps}^2)$$

zu betrachten, d.h.

$$\left| \frac{\tilde{y} - y}{y} \right| \leq 2\text{eps} + O(\text{eps}^2) .$$

Beim zweiten Algorithmus erhält man

$$\begin{aligned} & (x_1 \cdot x_2(1 + \varepsilon_1) + x_2 \cdot x_2(1 + \varepsilon_2))(1 + \varepsilon_3) \\ & = x_1x_2 + x_2x_2 + x_1x_2(\varepsilon_1 + \varepsilon_3) + x_2x_2(\varepsilon_2 + \varepsilon_3) + O(\text{eps}^2) \end{aligned}$$

und als Abschätzung für den relativen Fehler

$$\begin{aligned} \left| \frac{\tilde{y} - y}{y} \right| & \leq \left| \frac{x_1x_2(\varepsilon_1 + \varepsilon_3) + x_2x_2(\varepsilon_2 + \varepsilon_3)}{(x_1 + x_2)x_2} \right| + O(\varepsilon^2) \\ & \leq \frac{|x_1| + |x_2|}{|x_1 + x_2|} 2\text{eps} + O(\text{eps}^2) . \end{aligned}$$

Dieser ist zwar von der Größenordnung des Problemfehlers, wie eine Berechnung der Konditionszahlen zeigt.

Damit ist der Algorithmus zwar gutartig, Algorithmus 1 ist aber zu bevorzugen.

Man erkennt hier auch die Faustregel wieder: Schlecht konditionierte Teile sind zuerst auszuführen.

Aufgabe 4. Zunächst berechnet man für relativen Konditionszahlen

$$K_1 = \left| \frac{x_1}{x_1 - x_2x_3} \right|, \quad K_2 = \left| \frac{x_3x_2}{x_1 - x_2x_3} \right|, \quad K_3 = \left| \frac{x_2x_3}{x_1 - x_2x_3} \right| .$$

Aus

$$u = x_2 \odot x_3 = x_2x_3(1 + \varepsilon_1)$$

folgt weiter

$$\begin{aligned}
 \tilde{y} &= x_1 \ominus u \\
 &= (x_1 - u)(1 + \varepsilon_2) \\
 &= (x_1 - x_2x_3(1 + \varepsilon_1))(1 + \varepsilon_2) \\
 &= ((x_1 - x_2x_3) - x_2x_3\varepsilon_1)(1 + \varepsilon_2) \\
 &= x_1 - x_2x_3 - x_2x_3(\varepsilon_1 + \varepsilon_2) + \varepsilon_2x_1 - x_2x_3\varepsilon_1\varepsilon_2 \\
 &= y + (x_1 - x_2x_3)\varepsilon_2 - x_2x_3\varepsilon_1 + O(\text{eps}^2) .
 \end{aligned}$$

Man erhält den Ausdruck

$$\left| \frac{\tilde{y} - y}{y} \right| = \left| \varepsilon_2 - \frac{x_2x_3\varepsilon_1}{x_1 - x_2x_3} \right| + O(\text{eps}^2)$$

und erkennt die Größenordnung des Problemfehlers wieder, der Algorithmus ist gutartig.

Literaturverzeichnis

- [AORS1] Ansorge, R., Oberle, H.J., Rothe, K., Sonar, Th.; Mathematik für Ingenieure 1 u. 2. 4. erweiterte Auflage, Wiley-VCH, Weinheim, 2010.
- [AORS2] Ansorge, R., Oberle, H.J., Rothe, K., Sonar, Th., Aufgaben und Lösungen zu Mathematik für Ingenieure 1 u. 2. Wiley-VCH, Weinheim, 2010.
- [Ba] Bärwolff, G.; Höhere Mathematik für Naturwissenschaftler und Ingenieure. 2. erweiterte Auflage, Spektrum-Elsevier, München 2005.
- [Br] Braun, R., Meise, R., Analysis mit Maple. 2.Auflage, Vieweg u. Teubner, Wiesbaden 2012.
- [BHW] Burg, K., Haf, H., Wille, F.; Höhere Mathematik für Ingenieure. I - V. Teubner/Vieweg-Teubner.
- [Di] Dirschmid, H.J.; Mathematische Grundlagen der Elektrotechnik. Vieweg, Braunschweig/Wiesbaden 1990.
- [Fi] Fischer, G.; Lineare Algebra. 17. Auflage, Vieweg u. Teubner, Wiesbaden 2010.
- [HSZ] Hackbusch, W.; Schwarz, H.R., Zeidler, E.; Teubner-Taschenbuch der Mathematik. Teubner, Wiesbaden 2003.
- [Hi1] Hildebrandt, S.; Analysis 1. Springer, Berlin/Heidelberg 2002.
- [Hi2] Hildebrandt, S.; Analysis 2. Springer, Berlin/Heidelberg 2003.
- [HMV1] Hoffmann, A., Marx, B., Vogt, W.; Mathematik für Ingenieure 1. Pearson, München 2005. eBook: ISBN: PDF-978-3-8273-7113-3

- [HVM2] Hoffmann, A., Marx, B., Vogt, W.; Mathematik für Ingenieure 2. Pearson, München 2006. eBook: ISBN: PDF-978-3-8273-7114-0
- [spektrum] Lexikon der Mathematik, Spektrum Akademischer Verlag, Heidelberg, 2001.
- [Pa] Papula, L.; Mathematik für Chemiker. Enke, 1991.
- [PaW] Pavel, W., Winkler, R.; Mathematik für Naturwissenschaftler. Pearson, 2007.
- [Re] Reinsch, E.A.; Mathematik für Chemiker. Teubner, 2004.
- [Ro] Rösch, N.; Mathematik für Chemiker. Springer, 1993.
- [We] Westermann, Th., Ingenieurmathematik kompakt mit Maple. Springer online.
- [FJ] Zachmann, H.G., Jüngel, A.; Mathematik für Chemiker. Wiley, 2007.

Index

- LR*-Zerlegung, 42
- n*-Linearform, 249
- Abbildung
 - lineare, 7, 12
 - lineare, 67
- Ableitung, 114
 - geometrische Interpretation, 111
 - höhere, 124
 - kinematische Interpretation, 112
 - linksseitige, 114
 - partielle
 - partielle Ableitung, 257
 - Produktregel, 118
 - Quotientenregel, 119
 - rechtsseitige, 114
 - Umkehrfunktion, 122
 - zweite, 124
- absolutes Maximum, 101, 126
- absolutes Minimum, 101, 126
- adjungierte Matrix, 16
- affin linear, 69
- affin lineare Approximation, 112
- Algorithmus, 259
 - gut konditioniert, 260
 - gutartig, 260
- Algorithmus von Neville, 148
- allgemeine Lösung des homogenen Systems, 19
- allgemeine Lösung des inhomogenen Systems, 20
- alternierend, 55
- Approximation im quadratischen Mittel, 240
- Assoziativgesetz, 13
- Ausgabedaten, 253
- Ausgleich nach Tschebyscheff, 26
- Ausgleichsproblem, lineares, 26
 - Ausgleich nach Tschebyscheff, 26
 - Methode der kleinsten Quadrate, 26
 - Normalgleichung, 28
 - Residuum, 25
- Auslöschung, 258, 263
- Basis
 - duale, 247
- Basiswechsel, 77
- Beschränktheit, 185
- bestimmtes Integral, 161, 165
- Betragsfunktion, 94
- Bidualraum, 248
- bijektive lineare Abbildung, 71
- Bild einer linearen Abbildung, 69
- Bilinerform, 249
- Cauchyscher Hauptwert, 188
- Cholesky-Zerlegung, 46
- Cramersche Regel, 57
- Daten
 - Ausgabedaten, 253
 - Eingabedaten, 253
- Determinante, 47
 - alternierend, 55

- Determinantenmultiplikations-
satz, 57
- Entwicklung nach der i^{ten} Zeile,
52
- Entwicklung nach der j^{ten} Spalte,
53
- Entwicklungssatz von Laplace,
52
- n-Linearform, 55
- Regel von Sarrus, 52
- Determinantenmultiplikationssatz,
57
- Differentialquotient, 114
- Differenzenquotient, 113
 - zentraler, 146
- Differenzierbarkeit, 111, 113
 - Kettenregel, 121
 - und gleichmäßige Konvergenz,
119
- Dimensionsformel, 24
- Distributivgesetz, 13
- Dreieckszerlegung, 42
- duale Basis, 247
- Dualraum, 246
- Durchschnittsgeschwindigkeit, 115
- Eingabedaten, 253
- Einheitsmatrix, 11
- einseitiger Grenzwert, 106, 114
- Entwicklung nach der i^{ten} Zeile, 52
- Entwicklung nach der j^{ten} Spalte, 53
- Entwicklungssatz von Laplace, 52
- erster Stufe, 83
- erweiterte Matrix, 22
- Extrapolation zum Limes $h \rightarrow 0$,
141
- Extremum
 - lokales, 127
- Fehler
 - Gesamtfehler, 260
 - Problemfehler, 258
 - relativer
 - Ausgabe, 257
 - Eingabe, 257
- Fehleranalyse
 - Auslöschung, 258
 - differentielle, 254
 - Konditionszahl, 257
- Feinheit, 163
- Flächeninhalt, 170
 - orientierter, 47
- Fourier
 - Koeffizienten, 234
- Fourier-Reihe, 227, 234
- Frequenz, 231
- Funktion
 - Ableitung, 114
 - affin lineare Approximation, 112
 - Differentialquotient, 114
 - Differenzenquotient, 113
 - differenzierbare, 111, 113
 - Grenzwert, 95
 - integrierbare, 165
 - konkave, 134
 - konvexe, 134
 - Lipschitz-stetige, 94
 - periodische, 227
 - reell analytische, 219
 - stückweise stetige, 168
 - stetige, 91, 95
 - Wendepunkt, 138
- Funktional
 - lineares, 246
- Gesamtfehler, 260
- Geschwindigkeitsvektor, 112

- Gibbs-Phänomen, 237
- gleichmäßig stetig, 100
- Gleichungssystem, lineares, 17
 - allgemeine Lösung homogen, 19
 - allgemeine Lösung inhomogen, 20
 - Cramersche Regel, 57
 - Dimensionsformel, 24
 - homogenes lineares Gleichungssystem, 18
 - inhomogenes lineares Gleichungssystem, 18
 - Koeffizientenmatrix, 17
 - Rückwärtseinsetzen, 44
 - spezielle Lösung, 20
 - Superpositionsprinzip, 18
 - triviale Lösung, 18
 - überbestimmt, 24
 - unterbestimmt, 24
 - Vorwärtseinsetzen, 44
- globales Maximum, 126
- globales Minimum, 126
- Grenzwert, 95
 - einseitiger, 106, 114
- Grundschwingung, 227
- gut konditioniert, 260
- gutartig, 260
- höhere Ableitung, 124
- harmonische Analyse, 227
- Hauptdiagonale, 11
- Hauptsatz der Differential- und Integralrechnung, 170, 172
- Hermiteisch, 60
- hinreichende Bedingung, 127
- homogenes lineares Gleichungssystem, 18
- Homomorphismus, 68
- Infinitesimalrechnung, 91
- inhomogenes lineares Gleichungssystem, 18
- injektive lineare Abbildung, 71
- Integrabilitätskriterium, 167
- Integral
 - bestimmtes, 161, 165
 - Flächeninhalt, 170
 - Integrand, 165
 - Integrationsvariable, 165
 - Oberintegral, 165
 - Obersumme, 163
 - orientiertes, 169
 - unbestimmtes, 172
 - uneigentliches, 187, 192
 - Unterintegral, 165
 - Untersumme, 163
- Integrand, 165
- Integration, 172
 - numerische, 193
 - Partialbruchzerlegung, 180
 - partielle, 176
 - Substitutionsregel, 177
- Integrationsvariable, 165
- Integrationsverfahren von Romberg, 201
- integrierbar, 165
 - lokal, 187
- Integralrechnung
 - Mittelwertsatz, 203
- inverse Matrix, 39
- invertierbar, 40
- irreduzibel, 184
- Kern einer linearen Abbildung, 69
- Kern einer Matrix, 19
- Kettenregel, 121
- Koeffizientenmatrix, 17

- Koeffizientenvergleich, 182
- kompakt, 101
- konditioniert
 - gut, 253, 258
 - schlecht, 258
- Konditionszahl, 257
- konjugierte Matrix, 16
- konkav, 134
- kontravariant, 84, 248
 - p -fach, 249
- Konvergenz
 - gleichmäßige, 99, 119
- konvex, 134
- Koordinaten, 73
- Koordinaten bzgl. beliebiger Basis, 75
- Koordinatentransformation, 78
 - Transformationsmatrix, 78
- kovariant, 83, 246
 - q -fach, 249
- Krümmung, 124
- kritischer Punkt, 129
- Kurve, 112
- Landausche Symbole, 255
- Leibnizkalkül, 121
- lineare Abbildung, 7, 12, 67
 - bijektive, 71
 - Bild, 69
 - injektive, 71
 - Kern, 69
 - Matrixdarstellung, 77
 - Rang, 69
 - Rangsatze, 70
 - surjektive, 71
- lineare Algebra, 91
- lineares Funktional, 246
- Linearfaktoren, 181
- Linearform, 246
- linke untere Dreiecksmatrix, 44
- linksseitige Ableitung, 114
- Lipschitz-stetig, 94
- lokal integrierbar, 187
- lokales Extremum, 127
- lokales Maximum, 127
- lokales Minimum, 127
- Majorantenkriterium, 189, 192
- Mantisse, 264
- Matrix, 7
 - LR -Zerlegung, 42
 - adjungierte, 16
 - Cholesky-Zerlegung, 46
 - Determinante, 47
 - Dreieckszerlegung, 42
 - Einheitsmatrix, 11
 - erweiterte, 22
 - Hauptdiagonale, 11
 - Hermitesche, 60
 - inverse, 39
 - invertierbare, 40
 - Kern, 19
 - konjugierte, 16
 - linke untere Dreiecksmatrix, 44
 - Nullmatrix, 9
 - orthogonale, 58
 - Produkt, 10
 - quadratische, 14
 - Rang, 22, 23, 40
 - rechte obere Dreiecksmatrix, 44
 - reguläre, 40
 - singuläre, 40
 - Spaltenrang, 22
 - Spaltenzahl, 8
 - Streichungsmatrix, 52
 - symmetrische, 46

- transponierte, 14
- unitäre, 60
- Zeilenrang, 23
- Zeilenzahl, 8
- Matrixdarstellung einer linearen Abbildung, 77
- Matrixprodukt, 10
 - Assoziativgesetz, 13
 - Distributivgesetz, 13
- Maximalstelle, 101
 - lokale, 127
- Maximierer, 101
 - lokaler, 127
- Maximum
 - absolutes, 101, 126
 - globales, 126
 - lokales, 127
- Methode der kleinsten Quadrate, 26
- Minimalstelle, 101
 - lokale, 127
- Minimierer, 101
 - lokaler, 127
- Minimum
 - absolutes, 101, 126
 - globales, 126
 - lokales, 127
- Minorantenkriterium, 189
- Mittelwertsatz, 133
 - Integralrechnung, 203
- mittlere quadratische Abweichung, 240
- Monentangeschwindigkeit, 115
- monoton fallend, 134
 - streng, 134
- monoton wachsend, 134
 - streng, 134
- Multilinearform, 83, 249
- n-Linearform, 55
- negativ orientiert, 58
- Newton-Cotes Quadraturformeln, 195
 - abgeschlossene, 196
 - offene, 196
- Newton-Cotes Summenformeln, 200
- Normalgleichung, 28
- Normierungsbeziehungen, 232
- notwendige Bedingung, 126, 127
- Nullmatrix, 9
- Nullstelle
 - Vielfachheit, 181
- numerische Aufgabe, 253
- numerische Differentiation, 146
- numerische Integration, 193
 - Newton-Cotes Quadraturformeln, 195
 - Quadraturgewichte, 195
 - Trapezregel, 194
- numerische Mathematik, 260
- Oberintegral, 165
- Oberschwingung, 227
- Obersumme, 163
- orientierter Flächeninhalt, 47
- orientiertes Integral, 169
- orientiertes Volumen, 47
- Orientierung, 48
 - negativ orientiert, 58
 - positiv orientiert, 58
- orthogonale Matrix, 58
- Orthogonalitätsbeziehungen, 232
- Parsevalsche Gleichung, 240
- Partialbruchzerlegung, 180
- partielle Ableitung, 257
- partielle Integration, 176

- Permutation
 - Signum, 54
 - Transposition, 51
- Polynom
 - irreduzibel, 184
- Polynomdivision, 180
- positiv orientiert, 58
- Prinzip der kleinsten Wirkung, 125
- Problemfelder, 258
- Produktregel, 118
- Punkt
 - kritischer, 129
 - Sattelpunkt, 127
 - stationärer, 129
- Quadraturgewichte, 195
- Quotientenregel, 119
- Rückwärtseinsetzen, 44
- Rang einer linearen Abbildung, 69
- Rang einer Matrix, 22, 23, 40
- Rangatz, 70
- Rechte Hand Regel, 57
- rechte obere Dreiecksmatrix, 44
- rechtsseitige Ableitung, 114
- reell analytisch, 219
- Regel von Sarrus, 52
- Regeln von l'Hospital, 139
- regulär, 40
- Reihe
 - Fourier, 227
 - trigonometrische, 231
- Residuum, 25
- Restglied, 213
 - Darstellung nach Lagrange, 215
 - Integraldarstellung, 215
- Riemann integrierbar, 165
- Rotation, 59
- Sägezahnfunktion, 228, 235
- Sattelpunkt, 127
- Satz
 - Rolle, 131
- Sekante, 115, 134
- Signum, 54
- singulär, 40
- Singularität, 92
- Skalar, 8
- Skalarprodukt, 15
 - Hermiteches, 16
- Spaltenrang, 22
- Spaltenvektor, 8
- Spaltenzahl, 8
- Spektrum
 - diskretes, 228
 - periodischer Funktionen, 227
- spezielle Lösung, 20
- Sprungstelle, 237
- stückweise glatt, 237
- stückweise stetig, 168, 237
- Stammfunktion, 171
- stationärer Punkt, 129
- Stetigkeit, 91, 95
 - stückweise, 168
 - stückweise stetig, 237
 - und gleichmäßige Konvergenz, 99
 - und Kompaktheit, 101
 - und Umkehrfunktion, 103
- Stetigkeit
 - gleichmäßige, 100
- Streichungsmatrix, 52
- strenges lokales Maximum, 127
- strenges lokales Minimum, 127
- Substitutionsregel, 177
- Superposition, 227
- Superpositionsprinzip, 18

- surjektive lineare Abbildung, 71
- symmetrische Matrix, 46
- Tangente, 112, 134
- Taylor
 - Formel, 213
 - Polynom, 213
 - Entwicklungspunkt, 213
 - Reihe, 219
 - Restglied, 213
- Tensor, 83, 249
 - p -fach kontravariant, 249
 - q -fach kovariant, 249
 - erster Stufe, 83, 246, 248
 - höherer Stufe, 250
 - kontravariant, 84, 248
 - kovariant, 83, 246
- Tensorprodukt, 250
- Transformationsmatrix, 78
- Transposition, 14
- Transposition einer Permutation, 51
- Trapez-Summenregel, 200
- Trapezregel, 194
- trigonometrische Reihe, 231
- trigonometrisches Polynom, 231
- triviale Lösung, 18
- überbestimmtes Gleichungssystem, 24
- Umkehrfunktion
 - Ableitung, 122
- unbestimmtes Integral, 172
- uneigentliches Integral, 187, 192
 - absolute Konvergenz, 187, 192
 - Cauchyscher Hauptwert, 188
 - Divergenz, 187, 192
 - Konvergenz, 187, 192
 - Konvergenzkriterien, 189
 - Majorantenkriterium, 189, 192
 - Minorantenkriterium, 189
- unitär, 60
- unterbestimmtes Gleichungssystem, 24
- Unterintegral, 165
- Untersumme, 163
- Vektor
 - Koordinaten, 73
 - Koordinaten bzgl. beliebiger Basis, 75
 - Koordinatentransformation, 78
 - Spaltenvektor, 8
 - Zeilenvektor, 15
- Vektorprodukt, 57
- Vektorraum
 - dualer, 246
- Verfeinerung, 163, 203
- Vielfachheit, 181
- Volumen
 - orientiertes, 47
- Vorwärtseinsetzen, 44
- Wendepunkt, 138
- Zeilenrang, 23
- Zeilenvektor, 15
- Zeilenzahl, 8
- zentraler Differenzenquotient, 146
- Zerlegung, 163
 - Feinheit, 163
 - Verfeinerung, 163, 203
- zweite Ableitung, 124
- Zwischenwertsatz, 101

PERSONENVERZEICHNIS

- Cauchy, Augustin Louis (1789-1857), 188
- Cauchy, Augustin, Louis (1789-1857), 218
- Cholesky, André-Louis (1875-1918), 46
- Cotes, Roger (1682-1716), 196
- Cramer, Gabriel (1704-1752), 57
- Euler, Leonhard (1707-1783), 191
- Fourier, J.B. Baron de (1768–1830), 227
- Gibbs, J.W. (1839-1903), 237
- Hermite, Charles (1822-1901), 16
- Hilbert, David (1862-1943), 92
- l’Hospital, Guillaume Marquis de (1661-1704), 139
- Lagrange, Joseph Louis (1736-1813), 215
- Landau, E. (1877-1938), 255
- Laplace Marquis de, Pierre Simon (1749-1827), 52
- Leibniz, Gottfried Wilhelm (1646-1716), 111, 179
- Lipschitz, Rudolf (1832-1903), 94
- Neville, Eric Harold (1889-1961), 148
- Newton, Isaac (1643-1727), 111, 196
- Riemann, Bernhard (1826-1866), 165
- Rolle, Michel (1652-1719), 131
- Sarrus, Pierre Frédérick (1798-1861), 52
- Taylor, Brook (1685-1731), 213
- Tschebyscheff, Pafnuti Lwowsch (1821-1894), 26

SYMBOLVERZEICHNIS

Fehleranalyse

$O(h(\mathbf{x}))$ Landausches Symbol,
255

$o(h(\mathbf{x}))$ Landausches Symbol,
255

Funktionen

$\frac{df}{dx}(x_0)$ Ableitung von f an der
Stelle x_0 , 114

$\frac{df}{dx}|_{x=x_0}$ Ableitung von f an der
Stelle x_0 , 114

$f'(x_0)$ Ableitung von f an der
Stelle x_0 , 114

$f^{(k)}(x)$ k^{te} Ableitung von f , 124

$f'(x_0^-)$ linksseitige Ableitung,
114

$\partial f / \partial x_i$ partielle Ableitung, 257

$f'(x_0^+)$ rechtsseitige Ableitung,
114

$\frac{d^2}{dx^2}f(x)$ zweite Ableitung von
 f , 124

$f''(x)$ zweite Ableitung von f ,
124

$f^{(2)}(x)$ zweite Ableitung von f ,
124

$C^\infty(I)$ beliebig oft differenzier-
bar, 218

$\lim_{x \rightarrow x_0} f(x)$ Grenzwert, 95

$C^k(I)$ k -mal stetig differenzier-
bar, 124

C^ω reell analytisch, 219

$T_n(x; x_0)$ Taylor Polynom, 213

$R_n(x - x_0)$ Taylor Restglied, 213

Integral

$F(x)|_a^b$ Auswertung der Stamm-
funktion, 172

$[F(x)]_a^b$ Auswertung der

Stammfunktion, 172

$\int_I f(x) dx$ bestimmtes Integral,
165

$\int_a^b f(x) dx$ bestimmtes Integral,
165

$\mathcal{I}(f)$ bestimmtes Integral, 165

$\Delta(\mathcal{Z})$ Feinheit einer Zerlegung,
163

$\bar{\mathcal{I}}$ Oberintegral, 165

$\bar{S}_{\mathcal{Z}}(f)$ Obersumme, 163

$\mathcal{R}(I)$ Riemann-integrierbare
Funktionen, 165

$\int f(x) dx$ unbestimmtes Inte-
gral, 172

$\int_a^\infty f(x) dx$ uneigentliches Inte-
gral, 187

$\int_a^b f(x) dx$ uneigentliches Inte-
gral, 192

$\int_{-\infty}^\infty f(x) dx$ uneigentliches In-
tegral, 188

$\underline{\mathcal{I}}$ Unterintegral, 165

$\underline{S}_{\mathcal{Z}}(f)$ Untersumme, 163

$\mathcal{Z} \subseteq \mathcal{Z}^*$ Verfeinerung einer Zer-
legung, 203

\mathcal{Z} Zerlegung, 163

Lineare Abbildungen

V^* Dualraum, 246

\mathcal{V} Basis des \mathbb{R}^k , 72

bild L Bild einer lin. Abb., 69

kern L Kern einer lin. Abb., 69

$\begin{pmatrix} v_1 \\ v_2 \end{pmatrix}_{\mathcal{V}}$ Koordinatendarstel-
lung bzgl. Basis, 73

rg L Rang einer lin. Abb., 69

$U \otimes V$ Tensorprodukt, 250

Matrizen

A^* adjungierte Matrix, 16

I_m Einheitsmatrix, 11
 $(A|\mathbf{b})$ erweiterte Matrix, 22
 A^{-1} Inverse, 39
kern A Kern einer Matrix, 19
 \bar{A} konjugierte Matrix, 16
 $M(n, m)$ Matrix, 8
 $\mathbb{R}^{(n,m)}$ Matrix, 8
 $\mathbb{R}^{n \times m}$ Matrix, 8
rg A Rang einer Matrix, 22
 A^T transponierte Matrix, 14