

# Was bisher geschah



## In diesem Kapitel ...

- ▶ Bezeichnungen und Schreibweisen
- ▶ Vektoren, Matrizen, Determinanten und lineare Gleichungssysteme
- ▶ Das Wichtigste aus der eindimensionalen Analysis

**M**ehrdimensionale Analysis, gewöhnliche Differentialgleichungen, Funktionentheorie, stochastische Prozesse – das alles klingt nach fortgeschrittener Mathematik. Und das ist es auch. Die Themen dieses Buchs bauen auf mathematischen Methoden aus der Linearen Algebra und der eindimensionalen Analysis auf. Naturgemäß können nicht alle Grundlagen ausführlich erläutert werden: das Buch hätte dann mindesten den doppelten Umfang. Ich gehe daher davon aus, dass Sie den größten Teil der benötigten Grundlagenmathematik schon kennen: Dazu gehören die Grundbegriffe der mathematischen Sprache, neben den üblichen Rechensymbolen auch die logischen Zeichen  $\exists$  für »es existiert« und  $\forall$  für »für alle«. Das Zahlensystem der reellen Zahlen  $\mathbb{R}$  mit den darin enthaltenen natürlichen Zahlen  $\mathbb{N}$ , den ganzen Zahlen  $\mathbb{Z}$  und den rationalen Zahlen  $\mathbb{Q}$  und die darauf aufbauende Vektorrechnung in den  $n$ -dimensionalen Vektorräumen  $\mathbb{R}^n$  oder die eindimensionale Analysis.

Keine Panik! Die für dieses Buch wichtigsten Begriffe und Methoden aus diesen Bereichen werde ich in diesem Kapitel kurz erläutern. Falls Sie darüber hinaus neugierig geworden sind, wie das alles im Detail aussieht, können Sie das zum Beispiel im ersten Band »Mathematik für Ingenieure 1 für Dummies« nachlesen.

## Grundlagen aus der Linearen Algebra

Das Rechnen mit Vektoren und Matrizen spielt für die mehrdimensionale Analysis eine ähnliche Rolle wie die Grundrechenarten für die eindimensionale Analysis. Anstelle der reellen Zahlen aus  $\mathbb{R}$  treten hierbei Vektoren als Variablen und Funktionswerte auf und bei der Klassifizierung von Extremstellen helfen Ihnen bestimmte Eigenschaften von Matrizen weiter. Die bei der Optimierung häufig auftretenden oft sehr schwierig zu lösenden nichtlinearen Gleichungssysteme können Sie mit Hilfe des *Newtonverfahrens* durch die Lösungen linearer Gleichungssysteme zumindest approximieren. Grund genug, diese Grundlagen hier kurz zu wiederholen.

## Vektor- und Matrizenrechnung

In diesem Buch werden überwiegend die beiden reellen Vektorräume  $\mathbb{R}^2$  und  $\mathbb{R}^3$  auftreten, die gleichwertig entweder als zwei- beziehungsweise dreidimensionale Spaltenvektorräume oder als zwei- oder dreidimensionale Zeilenvektorräume aufgefasst werden.

Zwischen Spaltenvektoren

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

und Zeilenvektoren  $(x_1, x_2, \dots, x_n)$  können Sie mit Hilfe der *Transposition* umschalten. Die mit einem hochgestellten  $\top$  bezeichnete Transposition vertauscht Zeilen mit Spalten:

$$(x_1, x_2, \dots, x_n)^\top := \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}^\top = (x_1, x_2, \dots, x_n)$$



Spaltenvektoren (und analog Zeilenvektoren) werden komponentenweise addiert:

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} := \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{pmatrix}$$

Die Vektoraddition ist kommutativ und assoziativ.

Jeden reellen Vektor können Sie mit einer beliebigen reellen Zahl  $\lambda \in \mathbb{R}$  *skalar multiplizieren*. Die skalare Multiplikation wird ebenfalls komponentenweise durchgeführt:

$$\lambda \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} := \begin{pmatrix} \lambda x_1 \\ \lambda x_2 \\ \vdots \\ \lambda x_n \end{pmatrix}$$

Vektoraddition und skalare Multiplikation wird auch für reelle Zeilenvektoren komponentenweise definiert.

Mit Hilfe des *Standardskalarprodukts* können Sie Winkel zwischen zwei Vektoren

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}; \quad x, y \in \mathbb{R}^n$$

definieren und die *Euklidische Norm* oder Länge eines Vektors  $x \in \mathbb{R}^n$  bestimmen.



Für beliebige Vektoren  $x, y \in \mathbb{R}^n$  heißt die reelle Zahl

$$\langle x, y \rangle := \sum_{i=1}^n x_i y_i$$

das *Standardskalarprodukt* der Vektoren  $x$  und  $y$ . Eine häufig verwendete Kurzschreibweise für das Skalarprodukt ist

$$x \cdot y := \langle x, y \rangle.$$

Die *Euklidische Norm* eines Vektors  $x \in \mathbb{R}^n$  ist durch

$$\|x\| := \sqrt{x \cdot x} = \sqrt{\sum_{i=1}^n x_i^2}$$

definiert.

Den Winkel zwischen den beiden Vektoren  $x, y \in \mathbb{R}^n$  bestimmen Sie aus dem Skalarprodukt der beiden Vektoren mit einer einfachen Formel.



Der *Winkel*  $\alpha$  zwischen den beiden Vektoren  $x$  und  $y$  aus  $\mathbb{R}^n$  lässt sich nach der Formel

$$\cos(\alpha) = \frac{x \cdot y}{\|x\| \|y\|}$$

bestimmen.

## Das Vektorprodukt

Für die Darstellung von Flächenintegralen werden in Kapitel 8 einige Eigenschaften des *Vektor- oder Kreuzprodukts* zweier Vektoren benötigt.



Für zwei Vektoren

$$x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \quad \text{und} \quad y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \quad \text{mit} \quad x, y \in \mathbb{R}^3$$

heißt der durch

$$x \times y := \begin{pmatrix} x_2 y_3 - x_3 y_2 \\ x_3 y_1 - x_1 y_3 \\ x_1 y_2 - x_2 y_1 \end{pmatrix}$$

definierte Vektor das *Vektor- oder Kreuzprodukt von  $x$  und  $y$* .

Für Kreuzprodukte gelten die folgenden Rechenregeln:

- ✓  $x \times x = \vec{0}$
- ✓  $(\lambda x) \times y = x \times (\lambda y) = \lambda(x \times y)$  für beliebige Zahlen  $\lambda \in \mathbb{R}$
- ✓ Zwei Vektoren  $x$  und  $y$  sind genau dann linear abhängig, wenn das Kreuzprodukt  $x \times y = \vec{0}$  ergibt.

✓ Für drei Vektoren  $x, y, z \in \mathbb{R}^3$  gelten:

$$(x + y) \times z = x \times z + y \times z$$

$$x \times (y + z) = x \times y + x \times z$$

✓ Das Kreuzprodukt ist *antisymmetrisch*, das heißt:  $x \times y = -y \times x$

✓ Das Assoziativgesetz ist für das Kreuzprodukt im allgemeinen **nicht** erfüllt!

✓ Das Kreuzprodukt  $x \times y$  ist orthogonal zu  $x$  und zu  $y$ .

✓ Sind  $x$  und  $y$  linear unabhängig, so bilden  $x, y$  und  $x \times y$  eine positiv orientierte Basis.

✓  $\|x \times y\|^2 = \|x\|^2 \|y\|^2 - (x \cdot y)^2$

Ist  $\alpha \neq 0$  der Winkel zwischen den beiden linear unabhängigen Vektoren  $x, y \in \mathbb{R}^3$ , dann gilt:

$$\|x \times y\| = \|x\| \cdot \|y\| \cdot |\sin(\alpha)| = \|x\| \cdot \|y\| \cdot \sin(\alpha)$$

Die Länge  $\|x \times y\|$  des Kreuzprodukts entspricht also dem Flächeninhalt des durch  $x$  und  $y$  aufgespannten Parallelogramms.

### Das Spatprodukt

Drei beliebige Vektoren  $a, b, c$  des  $\mathbb{R}^3$  spannen einen *Spat* auf. Dieser Körper ähnelt einem schiefen Backstein, ein Beispiel ist in Abbildung 1.1 dargestellt. Für die drei Vektoren  $a, b$  und  $c$  heißt die durch  $a \bullet (b \times c)$  definierte reelle Zahl das *Spatprodukt* aus  $a, b$  und  $c$ .

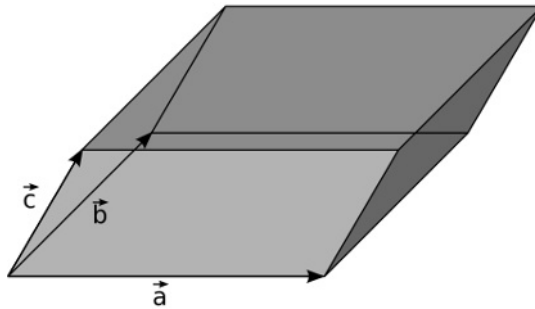


Abbildung 1.1: Ein Spat oder Parallelepiped, das durch die drei Vektoren  $a, b$  und  $c$  aufgespannt wird.

Der Absolutbetrag  $|a \bullet (b \times c)|$  entspricht dem Volumen des Spats.

Neben den Vektoren sind reelle  $(m, n)$ -Matrizen  $A \in \mathbb{R}^{m \times n}$  mit  $m$  Zeilen und  $n$  Spalten weitere mathematische Objekte aus der Linearen Algebra, die bei der mehrdimensionalen Analysis eine wichtige Rolle spielen.

Matrizen können Sie als Vektoren auffassen und genau wie bei Spalten- oder Zeilenvektoren können Sie auch Matrizen derselben Dimension komponentenweise addieren oder mit einer reellen Zahl multiplizieren. Zwei Matrizen, bei denen die Spaltenzahl der ersten Matrix mit der Zeilenzahl der zweiten Matrix übereinstimmt, können Sie nach der folgenden Definition auch miteinander multiplizieren.



Ist  $A = (a_{il})$  eine reelle  $(m, n)$ -Matrix und  $B = (b_{lk})$  eine reelle  $(n, p)$ -Matrix dann heißt die  $(m, p)$ -Matrix  $C = (c_{ik}) \in \mathbb{R}^{m \times p}$  mit den Komponenten

$$c_{ik} := \sum_{l=1}^n a_{il} b_{lk} \quad i = 1, \dots, m, \quad k = 1, \dots, p$$

das *Produkt von A und B*.

Das Standardskalarprodukt der Spaltenvektoren  $x, y \in \mathbb{R}^n$  können Sie als ein besonderes Matrizenprodukt interpretieren: Betrachten Sie die beiden Spaltenvektoren

$$x := \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad y := \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

als  $n \times 1$ -Matrizen und definieren Sie zum Spaltenvektor  $x$  den Zeilenvektor  $x^\top = (x_1, x_2, \dots, x_n)$ , eine einzeilige und  $n$ -spaltigen Matrix. Das Skalarprodukt  $\langle x, y \rangle = x^\top y$  ist dann das Matrizenprodukt aus der  $1 \times n$ -Matrix  $x^\top$  und der  $n \times 1$ -Matrix  $y$ .

## Lineare Gleichungssysteme und das Gauß-Verfahren

Lineare Gleichungssysteme, kurz LGS, treten bei vielen mathematischen Fragestellungen auf. Beispielsweise müssen Sie beim in Kapitel 4 behandelten Newtonverfahren zur Approximation von Nullstellen einer mehrdimensionalen Funktion in jedem Schritt ein LGS lösen. Ein LGS können Sie formal als

$$Ax = b$$

mit der Systemmatrix  $A \in \mathbb{R}^{n \times m}$ , dem gesuchten Vektor  $x \in \mathbb{R}^m$  und dem Vektor der rechten Seite  $b \in \mathbb{R}^n$  schreiben. Die meisten LGS, die in diesem Buch vorkommen werden, sind *quadratische LGS*, das heißt  $m = n$ : die Systemmatrix  $A \in \mathbb{R}^{n \times n}$  ist also quadratisch und die beiden Vektoren  $x$  und  $b$  gehören dem selben Vektorraum  $\mathbb{R}^n$  an.

## Determinanten

Manchmal empfiehlt es sich, ein LGS zuerst auf Lösbarkeit zu untersuchen, bevor Sie versuchen, die Lösung zu berechnen. Dies können Sie für quadratische LGS beispielsweise mit Hilfe der sogenannten *Determinante*  $\det(A)$  der Systemmatrix  $A$  machen. Die eigentliche mathematische Definition der Determinanten ist etwas kompliziert, allerdings reicht es aus, eine Berechnungsvorschrift für Determinanten anzugeben.



Für eine quadratische Matrix  $A \in \mathbb{R}^{2 \times 2}$  mit

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

heißt die Zahl

$$\det(A) := \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} := a_{11}a_{22} - a_{12}a_{21}$$

die *Determinante* von  $A$ . Für eine Matrix  $A \in \mathbb{R}^{3 \times 3}$  ist

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

Determinanten für Matrizen mit mehr als drei Zeilen und Spalten berechnen Sie mit Hilfe des sogenannten *Laplace'schen Entwicklungssatzes* über die *Kofaktoren* bestimmter Matrixkomponenten.



Ist  $A \in \mathbb{R}^{n \times n}$  eine quadratische Matrix mit  $n$  Zeilen und Spalten, dann heißt die quadratische Matrix mit  $(n-1)$  Zeilen und Spalten, die aus  $A$  durch Streichen der  $i$ -ten Zeile und der  $k$ -ten Spalte entsteht, die *Untermatrix*  $A_{ik}$ .

$$\alpha_{ik} := (-1)^{i+k} \det(A_{ik})$$

heißt der *Kofaktor* zum Element  $a_{ik}$  der Matrix  $A$ .

Mit Hilfe von Untermatrizen lassen sich ganz allgemein Determinanten auf zwei Weisen rekursiv berechnen.

- ✓ Die eine Möglichkeit ist die *Entwicklung nach der  $k$ -ten Spalte*: für  $k = 1, 2, \dots, n$  ist:

$$\det(A) = \sum_{i=1}^n a_{ik} \alpha_{ik}$$

- ✓ Die andere Variante ist die *Entwicklung nach der  $i$ -ten Zeile*: für  $i = 1, 2, \dots, n$  ist:

$$\det(A) = \sum_{k=1}^n a_{ik} \alpha_{ik}$$

Mit diesen beiden Methoden können Sie die Berechnung der Determinante jeder quadratischen Matrix  $A \in \mathbb{R}^{n \times n}$  auf die Berechnung von Determinanten quadratischer Matrizen mit nur zwei Zeilen und Spalten zurückführen.

Determinanten werden zum Beispiel bei der weiter hinten in diesem Kapitel beschriebenen Eigenwertberechnung und bei der Transformation verschiedener Koordinatensysteme zur mehrdimensionalen Integration, die in Kapitel 6 behandelt wird, benötigt.

Außerdem können Sie mit Hilfe der Determinante der Systemmatrix die eindeutige Lösbarkeit eines quadratischen LGS feststellen.



Ist für die Systemmatrix  $A \in \mathbb{R}^{n \times n}$  eines quadratischen LGS

$$Ax = b$$

die Determinante  $\det(A) \neq 0$  von null verschieden, dann existiert ein einziger Lösungsvektor  $x \in \mathbb{R}^n$  für dieses LGS. Ist dagegen  $\det(A) = 0$ , dann gibt es entweder keine oder unendlich viele Lösungen des LGS.

## Das Gauß-Verfahren

Es gibt sehr viele verschiedene Verfahren zur exakten oder näherungsweise Lösung eines LGS. Die meisten dieser Verfahren eignen sich hauptsächlich für die computergestützte Lösung sehr großer LGS, das heißt für eine sehr große Zahl von Unbekannten und Gleichungen.

Einige Verfahren können Sie aber durchaus auch zur Berechnung der Lösung kleinerer LGS von Hand einsetzen. Das bekannteste und wichtigste dieser Verfahren ist der Gauß-Algorithmus. Die zugrundeliegende Idee besteht darin, das zu lösende Gleichungssystem auf *obere Dreiecksgestalt* zu bringen und dann die Lösung durch *Rückwärtslösen* zu bestimmen.

Ist beispielsweise  $R \in \mathbb{R}^{n \times n}$  eine *obere Dreiecksmatrix*:

$$R = \begin{pmatrix} r_{11} & r_{12} & \dots & r_{1n} \\ 0 & r_{22} & \dots & r_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & r_{nn} \end{pmatrix}$$

mit

$$\det(R) = r_{11}r_{22} \cdots r_{nn} \neq 0,$$

dann ist die Lösung des LGS  $Rx = c$  eindeutig bestimmt und das *Rückwärtslösen* können Sie nach dem folgenden Algorithmus vornehmen:

### 1. Starten Sie in der $n$ -ten Zeile:

$$r_{nn}x_n = c_n \implies x_n = \frac{c_n}{r_{nn}}$$

### 2. Weiter geht es mit der $(n-1)$ -ten Zeile:

$$r_{(n-1)(n-1)}x_{n-1} + r_{(n-1)n}x_n = c_{n-1} \implies x_{n-1} = \frac{1}{r_{(n-1)(n-1)}}(c_{n-1} - r_{(n-1)n}x_n)$$

Diese können Sie nach  $x_{n-1}$  auflösen, da Sie schon  $x_n$  aus dem letzten Schritt kennen.

### 3. Auf diese Weise arbeiten Sie sich rückwärts von unten nach oben durch die einzelnen Zeilen des gestaffelten Systems.

In der  $k$ -ten Zeile sieht das so aus:

$$r_{kk}x_k + r_{k(k+1)}x_{k+1} + \dots + r_{kn}x_n = c_k \implies x_k = \frac{1}{r_{kk}}(c_k - \sum_{j=k+1}^n r_{kj}x_j)$$

Zur Lösung eines beliebigen LGS  $Ax = b$  mit  $(m, n)$ -Matrix  $A \in \mathbb{R}^{m \times n}$  und rechter Seite  $b \in \mathbb{R}^m$  wird, wenn möglich, beim *Gauß-Algorithmus* das LGS durch geeignete *elementare Zeilenumformungen* zu einem LGS  $Rx = c$  mit Dreiecksmatrix  $R$  und mit neuer rechter Seite  $c$  umgeformt, welches dieselbe Lösung wie das ursprüngliche System besitzt.

Die erlaubten *elementaren Zeilenumformungen* sind dabei:

- ✓ Multiplikation einer Zeile mit einem Skalar  $\lambda \neq 0$ :

$$\sum_{k=1}^n a_{ik}x_k = b_i \longleftrightarrow \sum_{k=1}^n (\lambda a_{ik})x_k = \lambda b_i$$

- ✓ Vertauschung zweier Zeilen

- ✓ Addition eines Vielfachen einer Zeile zu einer anderen Zeile:

$$\sum_{k=1}^n (a_{ik} + \lambda a_{jk})x_k = \sum_{k=1}^n a_{ik}x_k + \lambda \sum_{k=1}^n a_{jk}x_k = b_i + \lambda b_j$$

Mit Hilfe der obigen Zeilenumformungen führen Sie die Gauß-Elimination so durch:

1. **Bilden Sie die *erweiterte Systemmatrix*  $(A; b)$  indem Sie die Matrix  $A$  rechts um eine Spalte mit dem Rechte-Seite-Vektor  $b$  erweitern. Setzen Sie  $i = 1$ .**

Nach jedem der folgenden Schritte erhalten Sie eine neue erweiterte Systemmatrix, die Sie aber der Übersichtlichkeit halber wieder mit  $(A; b)$  bezeichnen. Genauso bezeichnen Sie auch die neuen Komponenten wieder mit  $a_{ij}$ .

2. **Ist die Komponente  $a_{ii} = 0$ , dann tauschen Sie eine Zeile  $j > i$  mit  $a_{ji} \neq 0$  der erweiterten Systemmatrix  $(A; b)$  mit der Zeile  $i$ . Falls es unterhalb der  $i$ -ten Zeile keine Zeile  $j > i$  mit  $a_{ji} \neq 0$  gibt, sind Sie fertig und gehen zu Schritt 7.**

Prinzipiell ist es gleichgültig, durch welche Zeile Sie dabei die  $i$ -te Zeile ersetzen, solange nach dem Tausch die Komponente  $a_{ii} \neq 0$  ist. Bei der praktischen Rechnung wählen Sie hier möglichst eine Zeile  $j$  mit Komponente  $a_{ij} = 1$ .

3. **Putzen Sie unterhalb der  $i$ -ten Zeile die  $i$ -e Spalte, indem Sie zu jeder Zeile  $j$  mit  $i < j \leq m$  das  $-\frac{a_{ji}}{a_{ii}}$ -fache der  $i$ -ten Zeile addieren.**

Sie bearbeiten also nur die Zeilen unterhalb der  $i$ -ten Zeile und erhalten eine neue erweiterte Systemmatrix, bei der in der  $i$ -ten Spalte unterhalb der Diagonalkomponente  $a_{ii}$  nur noch Nullen stehen.

4. **Setzen Sie  $i = i + 1$ .**

5. **Ist  $i = n$ , dann sind Sie fertig und gehen zu Schritt 7.**

6. **Gehen Sie zu Schritt 2 und fahren Sie dort fort.**

Falls Sie nicht vorher zu Schritt 7 springen, dann wiederholen Sie diesen Algorithmus Zeile für Zeile bis Sie in der letzten Zeile  $m$  angekommen sind.

7. **Beenden Sie die Elimination.**



Mit diesem Eliminationsalgorithmus bringen Sie die erweiterte Matrix  $(A; b)$  auf die folgende Gestalt:

$$\left( \begin{array}{cccc|cccc|c} 1 & \alpha_{12} & \alpha_{13} & \cdots & \alpha_{1r} & \alpha_{1,(r+1)} & \cdots & \alpha_{1n} & \beta_1 \\ 0 & 1 & \alpha_{23} & \cdots & \alpha_{2r} & \alpha_{2,(r+1)} & \cdots & \alpha_{2n} & \beta_2 \\ \vdots & 0 & 1 & \ddots & \alpha_{3r} & \alpha_{3,(r+1)} & \cdots & \alpha_{3n} & \beta_3 \\ \vdots & & & \ddots & \vdots & \vdots & & \vdots & \vdots \\ 0 & \dots & \dots & 0 & 1 & \alpha_{r,(r+1)} & \cdots & \alpha_m & \beta_r \\ \hline 0 & \dots & \dots & & & & & & \beta_{r+1} \\ \vdots & & & & & & & & \vdots \\ 0 & & & & & & & & \beta_m \end{array} \right) =: (\tilde{A}, \tilde{b})$$



Die Zahl  $r$  wird *Rang der Matrix A* genannt.

In den ersten  $r$  Zeilen ist dies ein gestaffeltes Gleichungssystem. Hier stehen »Leerzeichen« für die Null und  $\alpha_{ij}$  und  $\beta_i$  für irgendwelche Zahlen. Ist  $r < m$ , dann treten in der neu erhaltenen Matrix  $A$  reine Nullzeilen auf, während auf der neuen rechten Seite  $b$  in diesen Zeilen eventuell auch andere Zahlen auftauchen können.

Für die Lösbarkeit des LGS  $Ax = b$  gilt:

- ✓ **Nicht lösbar**, falls die Zahlen  $\beta_{r+1}, \dots, \beta_m$  nicht alle gleich null sind.
- ✓ **Nicht eindeutig lösbar**, falls  $r = m < n$  ist, das heißt, falls alle  $\beta_{r+1} = \beta_{r+2} = \dots = \beta_m = 0$  sind. Die Unbekannten  $x_{r+1}, \dots, x_n$  können frei gewählt werden. Die restlichen Unbekannten  $x_1, \dots, x_r$  ergeben sich dann eindeutig aus den frei gewählten Unbekannten.
- ✓ **Eindeutig lösbar**, falls  $r = m = n$  ist. Die Lösung  $x$  erhalten Sie durch Rückwärtslösen.

## Eigenwerte, Eigenvektoren und die Definitheit von Matrizen

Wie in Abschnitt »Ganz sicher: hinreichende Optimalitätsbedingung« aus Kapitel 5 dargestellt wird, spielt bei der Untersuchung auf Extremstellen mehrdimensionaler Funktionen die *Definitheit* von symmetrischen Matrizen eine ähnlich wesentliche Rolle, wie das Vorzeichen der zweiten Ableitung für eindimensionale Funktionen.



Eine symmetrische  $(n, n)$ -Matrix  $A$  heißt:

- ✓ **Positiv semidefinit**, falls  $x^T Ax \geq 0$  für alle  $x \in \mathbb{R}^n$
- ✓ **Positiv definit**, falls  $x^T Ax > 0$  für alle  $\|x\| \neq 0$
- ✓ **Negativ semidefinit**, falls  $x^T Ax \leq 0$  für alle  $x \in \mathbb{R}^n$
- ✓ **Negativ definit**, falls  $x^T Ax < 0$  für alle  $\|x\| \neq 0$
- ✓ **Indefinit** in allen anderen Fällen

Anhand der Definition ist die Definitheit einer gegebenen Matrix meistens nicht so einfach zu erkennen. Es gibt aber einige Berechnungsmethoden dazu, mit deren Hilfe Sie eine gegebene Matrix auf Definitheit prüfen können. Die zuverlässigste Methode verwendet dazu die Eigenwerte der Matrix.



Für eine gegebene  $n \times n$ -Matrix  $A$  wird beim *Eigenwertproblem* ein Vektor  $x \neq \vec{0}$  und ein  $\lambda \in \mathbb{K}$  gesucht, so, dass

$$Ax = \lambda x$$

ist. Ein solches  $\lambda$  heißt *Eigenwert* von  $A$ . Jeder Vektor  $x$ , der die obige Gleichung mit  $\lambda$  erfüllt, heißt *Eigenvektor von  $A$  zum Eigenwert  $\lambda$* .



Die Gleichung  $\det(A - \lambda E) = 0$  heißt die *charakteristische Gleichung* von  $A$ .

Das Polynom  $P(\lambda) = \det(A - \lambda E)$  heißt das *charakteristische Polynom der Matrix  $A$* .

Die Lösungen der charakteristischen Gleichung  $\det(A - \lambda E) = 0$  einer  $n \times n$ -Matrix  $A$  und damit die Eigenwerte von  $A$  sind nichts anderes als die Nullstellen des charakteristischen Polynoms.



Zur Bestimmung der Definitheit einer symmetrischen Matrix betrachten Sie die Eigenwerte dieser Matrix. Ist  $A$  eine symmetrische Matrix mit den Eigenwerten  $\lambda_1, \lambda_2, \dots, \lambda_n$ . Dann ist  $A$ :

- ✓ Positiv semidefinit genau dann, wenn  $\lambda_i \geq 0$  für  $i = 1, \dots, n$
- ✓ Positiv definit genau dann, wenn  $\lambda_i > 0$  für  $i = 1, \dots, n$
- ✓ Negativ semidefinit genau dann, wenn  $\lambda_i \leq 0$  für  $i = 1, \dots, n$
- ✓ Negativ definit genau dann, wenn  $\lambda_i < 0$  für  $i = 1, \dots, n$

Die Eigenvektoren  $v$  zu einem gegebenen Eigenwert  $\lambda$  einer Matrix  $A$  berechnen Sie mit Hilfe des Gauß-Verfahrens aus Abschnitt »Lineare Gleichungssysteme und das Gauß-Verfahren« aus dem linearen Gleichungssystem

$$(A - \lambda E)v = 0.$$

Dabei entsteht die Matrix dieses Gleichungssystems einfach dadurch, dass Sie von den Diagonaleinträgen der Matrix  $A$  den Eigenwert  $\lambda$  abziehen.



Beim Lösen des Gleichungssystems für die Eigenvektoren dürfen Sie die *triviale* Lösung, den Nullvektor, nicht als Eigenvektor wählen! Mit dem Gauß-Verfahren müssen Sie allerdings mindestens eine ganze Nullzeile erzeugen können. Das heißt: es muss mindestens eine eindimensionale Lösungsmenge für dieses System geben. Sie haben also unendlich viele Lösungen als Eigenvektoren zur Wahl.

## Eindimensionale Analysis

Die eindimensionale Analysis beschäftigt sich mit reellwertigen Funktionen einer reellen Variablen  $f: \mathbb{R} \rightarrow \mathbb{R}$ . Im Wesentlichen wird das Änderungsverhalten solcher Funktionen untersucht: Wie ändern sich die Funktionswerte  $f(x)$ , wenn das Argument  $x$  geändert wird? Solche Untersuchungen werden in der Mathematik mit Hilfe geeigneter Folgen und ihrer Grenzwerte durchgeführt. Dies liefert Begriffe wie *Stetigkeit* und *Differenzierbarkeit* und gibt Ihnen geeignete Methoden zur Extremstellensuche. Selbst der *Integralbegriff* und Näherungsmethoden wie die *Taylorreihenentwicklung* beruhen auf Grenzwerten. *Folgen* und ihr Verhalten bilden die Grundlage der ganzen Analysis und sind damit auch die Grundlage für die erfolgreiche mathematische Beschreibung der Welt. Dies ist auch in der mehrdimensionalen Analysis der Fall, die in den folgenden Kapiteln näher beschrieben wird. Grund genug, hier einen kurzen Überblick über Folgen und Grenzwerte zu geben.

### Folgen, Häufungspunkte und Grenzwerte

Viele wichtige analytische Begriffe wie *Stetigkeit* und *Differenzierbarkeit* sind über Grenzwertseigenschaften bestimmter Folgen definiert.



Eine *Folge von Elementen einer Menge  $V$*  ist eine Abbildung

$$f: \mathbb{Z} \rightarrow V$$

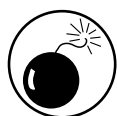
der Menge  $\mathbb{Z}$  der ganzen Zahlen in den Raum  $V$ , wobei der Definitionsbereich  $D(f) =: I$  nach unten beschränkt ist.

$I$  heißt der *Indexbereich* der Folge. Mit  $a_n = f(n)$  schreibt man für die Folge auch  $(a_n)_{n \in I}$ .

Die Indexmenge  $I$  einer Folge kann bei einer beliebigen ganzen Zahl  $n_0 \in \mathbb{Z}$  beginnen:

$$a_{n_0}, a_{n_0+1}, a_{n_0+2}, \dots$$

In Kurzschreibweise wird dies zu  $(a_n)_{n=n_0}^\infty$ . Das erlaubt oft, ein einfaches *Bildungsgesetz* für das allgemeine Folgenglied  $a_n$  mit  $n \geq n_0$  anzugeben.



Folgenglieder können aus beliebigen Mengen  $V$  stammen. Beispielsweise aus einem der Spaltenvektorräume  $\mathbb{R}^n$  oder einer endlichen Menge  $M$  mit irgendwelchen Elementen. Oder einfach aus den reellen oder komplexen Zahlen. In diesem Buch werden Sie es mit *reellen* und *komplexen Zahlenfolgen*, deren Folgengliedern  $a_n \in \mathbb{R}$  oder  $a_n \in \mathbb{C}$  reelle beziehungsweise komplexe Zahlen sind und mit *Punktfolgen* von Spaltenvektoren  $a_n \in \mathbb{R}^n$  zu tun haben.

In diesem Buch werden die meisten auftretenden Folgen  $(a_n)_{n \in I}$  aus  $\mathbb{R}^2$  oder  $\mathbb{R}^3$  stammen. Die behandelten Begriffe lassen sich aber fast immer auch auf beliebige endlichdimensionale Räume  $\mathbb{R}^n$  übertragen.



Ich werde im folgenden *Punkte* aus  $\mathbb{R}^n$  meist ohne Pfeile darstellen:

$$a_n \in \mathbb{R}^m, \quad x \in \mathbb{R}^n, \dots$$

Allerdings wird ab und zu zur besonderen Verdeutlichung der Vektoreigenschaften auch die Pfeilschreibweise in diesem Buch verwendet werden.

Grundlegend und sehr wichtig sind die Begriffe *Umgebung* und  $\varepsilon$ -*Umgebung* eines Punktes  $x \in \mathbb{R}^n$ .



Die Menge

$$U_\varepsilon(x) := \{y \in V \mid \|y - x\| < \varepsilon\}$$

mit  $\varepsilon > 0$  heißt  $\varepsilon$ -*Umgebung* von  $x$ . Dabei heißt

$$\|x\| := \sqrt{x_1^2 + x_2^2 + \dots + x_n^2}$$

die *Euklidische Norm*.



Eine Teilmenge  $U(x) \subseteq V$  heißt eine *Umgebung* von  $x$ , falls es eine reelle Zahl  $\varepsilon > 0$  gibt, so dass  $U_\varepsilon(x) \subseteq U(x)$  ist.

Der Begriff einer Umgebung erlaubt die Definition einiger weiterer wichtiger mathematischer Begriffe.



Eine Menge  $M \subseteq V$  heißt *offen*, falls es für jeden Punkt  $x \in M$  eine Umgebung  $U(x)$  gibt, die ganz in der Menge  $M$  enthalten ist:  $U(x) \subset M$ .

Der *Rand*  $\partial M$  einer Menge  $M \subset V$  ist die Menge aller Punkte  $x \in V$  für die jede Umgebung  $U(x) \subseteq V$  sowohl Punkte  $y \in M$  aus der Menge  $M$ , als auch Punkte  $y' \in V \setminus M$  aus dem Komplement  $V \setminus M := \{y' \in V \mid y' \notin M\}$  der Menge  $M$  enthält.

Der *Abschluss*  $\bar{M}$  einer Menge  $M \subseteq V$  bezeichnet die Menge

$$\bar{M} := M \cup \partial M \subseteq V,$$

das *Innere*  $\mathring{M} \subseteq M \subseteq V$  der Menge  $M$  ist

$$\mathring{M} := M \setminus \partial M \subseteq V.$$

Die Elemente  $x \in \mathring{M}$  heißen *innere Punkte* von  $M$ .

Mit Hilfe der beiden Begriffe Umgebung und  $\varepsilon$ -Umgebung können Sie von *konvergenten Folgen* und ihren *Grenzwerten* sprechen.



Eine Folge  $(a_n)_{n \in I}$  mit Folgengliedern  $a_n \in V$  heißt *konvergent*, falls es ein  $a \in V$  gibt, so dass es zu jeder reellen Zahl  $\varepsilon > 0$  eine Zahl  $N(\varepsilon) \in I$  mit der Eigenschaft

$$a_n \in U_\varepsilon(a) \quad \forall \quad n \in I \quad \text{mit} \quad n \geq N(\varepsilon)$$

gibt.  $a$  heißt *Grenzwert* oder *Limes* der Folge  $(a_n)_{n \in I}$ .

Eine nicht konvergente Folge heißt *divergent*.

Eine Folge  $(x^{(k)})_{k=1,2,3,\dots}$  von Punkten  $x^{(k)} = (x_1^{(k)}, x_2^{(k)}, \dots, x_n^{(k)}) \in \mathbb{R}^n$  konvergiert also gegen den Grenzwert  $\bar{x} = (\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n)$ , in Formeln

$$\lim_{k \rightarrow \infty} x^{(k)} = \bar{x}$$

falls eine der folgenden beiden Bedingungen erfüllt ist:

- ✓  $\lim_{k \rightarrow \infty} \|\bar{x} - x^{(k)}\| = 0$
- ✓  $\lim_{k \rightarrow \infty} x_j^{(k)} = \bar{x}_j$  für alle  $j = 1, 2, \dots, n$

Diese Bedingungen sind äquivalent: falls eine davon erfüllt ist, ist es automatisch auch die andere.

Grenzwertuntersuchungen sind im allgemeinen nicht einfach. Verschiedene Rechenregeln für Grenzwerte können Ihnen das Leben dabei etwas leichter machen.



Die folgenden Regeln gelten dann, wenn die beteiligten Folgen  $(a_n)$  und  $(b_n)$  konvergent sind:

- ✓ Konstante Faktoren  $\lambda \in \mathbb{R}$  können Sie herausziehen:

$$\lim_{n \rightarrow \infty} (\lambda a_n) = \lambda \lim_{n \rightarrow \infty} a_n$$

- ✓ Die Grenzwerte von Summen- beziehungsweise Differenzfolgen dürfen addiert (beziehungsweise subtrahiert) werden:

$$\lim_{n \rightarrow \infty} (a_n + b_n) = \lim_{n \rightarrow \infty} a_n + \lim_{n \rightarrow \infty} b_n$$

$$\lim_{n \rightarrow \infty} (a_n - b_n) = \lim_{n \rightarrow \infty} a_n - \lim_{n \rightarrow \infty} b_n$$

- ✓ Der Grenzwert der Produktfolge  $(a_n b_n)$  ist Produkt der Grenzwerte:

$$\lim_{n \rightarrow \infty} (a_n b_n) = \left( \lim_{n \rightarrow \infty} a_n \right) \left( \lim_{n \rightarrow \infty} b_n \right)$$

- ✓ Falls die Folgenglieder  $b_n \neq 0$  und der Grenzwert  $\lim_{n \rightarrow \infty} b_n \neq 0$  sind, dann ist der Grenzwert der Quotientenfolge der Quotient der Grenzwerte:

$$\lim_{n \rightarrow \infty} \left( \frac{a_n}{b_n} \right) = \frac{\lim_{n \rightarrow \infty} a_n}{\lim_{n \rightarrow \infty} b_n}$$

- ✓ Beträge können Sie herausziehen:

$$\lim_{n \rightarrow \infty} |a_n| = \left| \lim_{n \rightarrow \infty} a_n \right|$$

Wenn Sie die Elemente einer Folge oder einer nicht leeren (unendlich großen) Teilmenge  $M \subseteq V$  eines normierten Raums  $(V, \|\cdot\|)$  betrachten, ist es manchmal wichtig zu wissen, wie diese Elemente verteilt sind. Beispielsweise können alle Elemente gleich weit voneinander entfernt sein, wie die natürlichen Zahlen  $\mathbb{N} \subset \mathbb{R}$ . Oder sich um einen bestimmten Punkt ansammeln: Denken Sie an die Menge  $M := \{\frac{1}{n} \mid n \in \mathbb{N}\} \subset \mathbb{R}$ , deren Elemente sich um den Punkt 0 häufen. Solche Punkte werden *Häufungspunkte* oder *Häufungswerte* genannt.



Ein Punkt  $x \in V$  heißt *Häufungspunkt der Teilmenge*  $M \subset V$ , falls in jeder beliebigen noch so kleinen Umgebung  $U(x)$  mindestens noch ein weiterer, von  $x$  verschiedener Punkt  $y \in M$  liegt.



Ein Häufungspunkt  $x$  einer Menge  $M$  kann, aber muss nicht Element von  $M$  sein. Außerdem kann eine Menge  $M$  mehrere verschiedene Häufungspunkte haben.

Aus der Definition des Begriffs Häufungspunkt folgt sofort, dass in jeder beliebigen Umgebung  $U(x)$  eines Häufungspunkts  $x \in V$  von  $M$  unendlich viele Punkte aus  $M$  liegen: Sie können um den Häufungspunkt  $x$  unendlich viele konzentrische geschachtelte  $\varepsilon$ -Umgebungen  $U_i \subseteq M$  mit  $i \in \mathbb{N}$  mit immer kleinerem Radius  $\varepsilon$  so anordnen, dass bei  $i_2 > i_1$  für die Umgebungen  $U_{i_2} \subset U_{i_1}$  gilt. Betrachten Sie dazu Abbildung 1.2. In jeder Umgebung  $U_i$  muss mindestens ein  $y_i \in M$  liegen. Diesen Punkt können Sie so wählen, dass  $y_i \in U_i$  und  $y_i \notin U_j$  für alle  $j > i$  ist. Da innerhalb von  $U_i$  unendlich viele weitere Umgebungen  $U_j$  liegen, müssen wirklich unendlich viele Punkte in jeder dieser Umgebungen von  $x$  enthalten sein. Die Elemente von  $M$  häufen sich im wahrsten Sinne des Wortes.

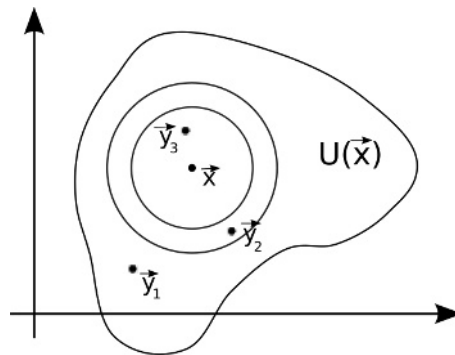


Abbildung 1.2: In jeder Umgebung des Häufungspunkts  $x$  liegen unendlich viele Punkte  $y_i$

## Grenzwerte reellwertiger Funktionen und Stetigkeit

Für die eindimensionale Analysis sind oft Grenzwerte von Funktionen  $f: \mathbb{R} \rightarrow \mathbb{R}$  interessant. Vergleichen Sie beispielsweise die beiden Graphen in Abbildung 1.3, dann stellen Sie sofort einen prinzipiellen Unterschied fest. An der Stelle  $x_0 = 0$  springt die linke Funktion, während der Graph der rechten Funktion dort durchgehend weiterverläuft.

Etwas mathematischer formuliert geht es hier um die Frage, wie sich die Folge der Funktionswerte  $f(x_n)$  verhält, wenn  $(x_n)$  eine Folge reeller Zahlen mit Grenzwert  $\lim_{n \rightarrow \infty} x_n = x_0$  ist. Dies führt zum Begriff der *Stetigkeit* eindimensionaler Funktionen.

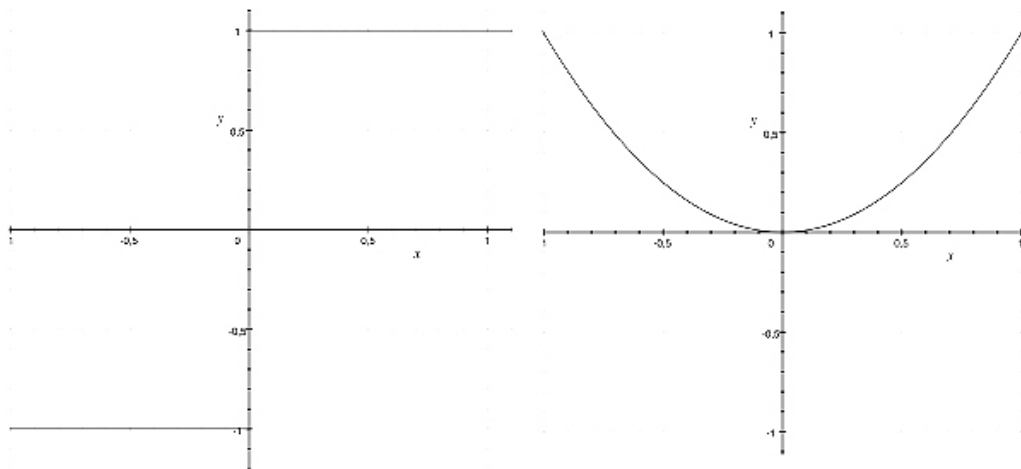


Abbildung 1.3: Eine unstetige (links) und eine stetige Funktion (rechts)



Eine Funktion  $f: \mathbb{R} \rightarrow \mathbb{R}$  mit Definitionsbereich  $D(f) \subseteq \mathbb{R}$  *konvergiert an einer Stelle*  $x_0 \in \mathbb{R}$  gegen den Wert  $c \in \mathbb{R}$ , falls für jede Folge  $(x_n)$  reeller Zahlen  $x_n \in D(f)$  aus dem Definitionsbereich von  $f$  mit  $\lim_{n \rightarrow \infty} x_n = x_0$  die Folge der Funktionswerte  $(f(x_n))$  gegen  $c$  konvergiert:

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

Für den *Grenzwert*  $c$  der Funktion  $f$  schreibt man:

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

In dieser Definition sind auch die Spezialfälle  $x_0 = \pm\infty$  beziehungsweise  $\lim_{n \rightarrow \infty} f(x_n) = \pm\infty$  eingeschlossen.

Die Stelle  $x_0$  muss nicht im Definitionsbereich  $D(f)$  von  $f$  liegen. Allerdings muss  $x_0$  ein Häufungspunkt von  $D(f)$  sein, damit überhaupt Folgen  $(x_n)$  von Elementen  $x_n \in D(f)$  aus dem Definitionsbereich von  $f$  existieren, die gegen die Zahl  $x_0$  konvergieren. Das bedeutet, dass  $x_0$  entweder ein *innerer* oder ein *Randpunkt* von  $D(f)$  sein muss.



Ein Punkt  $x_0 \in \partial M$  aus dem Rand einer Menge  $M \subseteq V$  heißt ein *Randpunkt* von  $M$ . Es gilt daher, daß in jeder Umgebung  $U(x_0)$  mindestens ein Punkt  $y \in M$  und ein Punkt  $z \in V \setminus M$  liegen. Dabei muss der Punkt  $x_0$  nicht unbedingt selbst aus der Menge  $M$  sein.

Bei der Grenzwertdefinition für Funktionen ist es wichtig, dass alle möglichen Folgen  $x_n \in D(f)$  mit Grenzwert  $x_0$  untersucht werden. Bekommen Sie für verschiedene solcher Folgen verschiedene Grenzwerte der Funktionswertfolgen  $f(x_n)$ , so existiert der Funktionsgrenzwert nicht. Dieselben Rechenregeln wie für die Berechnung von Folgengrenzwerten helfen Ihnen auch bei der Berechnung von Grenzwerten von Funktionen.



Konvergieren die beiden reellwertigen Funktionen  $f$  und  $g$  mit gemeinsamem Definitionsbereich  $D \subseteq D(f) \cap D(g)$  an der Stelle  $x_0$  gegen die Grenzwerte  $\lim_{x \rightarrow x_0} f(x) =: f_0$  beziehungsweise  $\lim_{x \rightarrow x_0} g(x) =: g_0$ , dann gelten die folgenden Rechenregeln:

- ✓ Jede Linearkombination  $af(x) + bg(x)$  mit  $a, b \in \mathbb{R}$  konvergiert gegen die Linearkombination der Grenzwerte:

$$\lim_{x \rightarrow x_0} (af(x) + bg(x)) = a \lim_{x \rightarrow x_0} f(x) + b \lim_{x \rightarrow x_0} g(x)$$

- ✓ Die Produktfunktion  $f(x) \cdot g(x)$  konvergiert gegen das Produkt der Grenzwerte:

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

- ✓ Falls  $\lim_{x \rightarrow x_0} g(x) \neq 0$  ist und  $g(x) \neq 0$  auf  $D$ , konvergiert die Quotientenfunktion  $\frac{f(x)}{g(x)}$  gegen den Quotienten der Grenzwerte:

$$\lim_{x \rightarrow x_0} \left( \frac{f(x)}{g(x)} \right) = \frac{\lim_{x \rightarrow x_0} f(x)}{\lim_{x \rightarrow x_0} g(x)}$$

Mit Hilfe von Grenzwerten für Funktionen wird der wichtige Begriff *Stetigkeit* definiert.



Eine Funktion  $f: \mathbb{R} \rightarrow \mathbb{R}$  mit Definitionsbereich  $D(f)$  heißt *stetig an der Stelle*  $x_0 \in D(f)$ , falls

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

ist.

Eine Funktion  $f: \mathbb{R} \rightarrow \mathbb{R}$  mit Definitionsbereich  $D(f)$  heißt *stetig auf der Menge*  $M \subseteq D(f)$ , falls  $f$  an jeder Stelle  $x_0 \in M$  stetig ist.

## Differenzierbarkeit und Kurvendiskussion

Die Frage nach dem Änderungsverhalten einer Funktion  $f: I \rightarrow \mathbb{R}$  führt über eine weitere Grenzwertbetrachtung zum Begriff der Differenzierbarkeit. Ist  $f: I \rightarrow \mathbb{R}$  eine auf einem Intervall  $I \subseteq \mathbb{R}$  definierte Funktion, dann nennt man den Quotienten

$$\frac{f(x) - f(x_0)}{x - x_0}$$

den *Differenzenquotienten von  $f$  an der Stelle  $x_0 \in I$* . Er beschreibt die Änderung der Funktionswerte in Abhängigkeit von der Änderung der Argumente und entspricht anschaulich der Steigung einer Geraden durch die beiden Punkte  $(x, f(x))$  und  $(x_0, f(x_0))$ . Der Grenzwert des Differenzenquotienten für  $x \rightarrow x_0$  entspricht dann der Tangentensteigung an die Funktion  $f$  im Punkt  $x_0$ , das heißt: der *Ableitung*.





Die Funktion  $f$  heißt *an der Stelle  $x_0$  differenzierbar*, falls für  $x$  gegen  $x_0$  der Grenzwert des Differenzenquotienten

$$\lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{h \rightarrow 0} \frac{f(x_0 + h) - f(x_0)}{h}$$

*eigentlich existiert*, das heißt nicht  $\pm\infty$  ist.

In diesem Fall wird der Grenzwert mit  $f'(x_0)$  bezeichnet und heißt *die Ableitung der Funktion  $f$  an der Stelle  $x_0$* .

Eine häufig verwendete Schreibweise für  $f'(x_0)$  ist der *Differenzialquotient*:

$$\frac{df}{dx}(x_0) := \frac{d}{dx}f(x_0) := f'(x_0)$$

Nicht alle Funktionen besitzen an jeder beliebigen Stelle  $x_0 \in I$  eine Ableitung, manche Funktionen sind sogar nirgends differenzierbar. Andere Funktionen sind dagegen auf ihrem ganzen Definitionsbereich differenzierbar und definieren damit eine weitere Funktion.



Eine Funktion  $f$  heißt auf  $I$  *differenzierbar*, falls  $f$  an jeder Stelle  $x \in I$  differenzierbar ist. Die Funktion  $f' : I \rightarrow \mathbb{R}$  mit

$$f'(x) := \lim_{h \rightarrow 0} \frac{f(x + h) - f(x)}{h}$$

heißt *Ableitung der Funktion  $f$* .

Die *Ableitung einer Funktion* hat eine anschauliche geometrische Bedeutung: Betrachten Sie die in Abbildung 1.4 dargestellte Gerade  $g_h(x) = mx + b$  durch die beiden Punkte  $(x_0 + h, f(x_0 + h))$  und  $(x_0, f(x_0))$ .

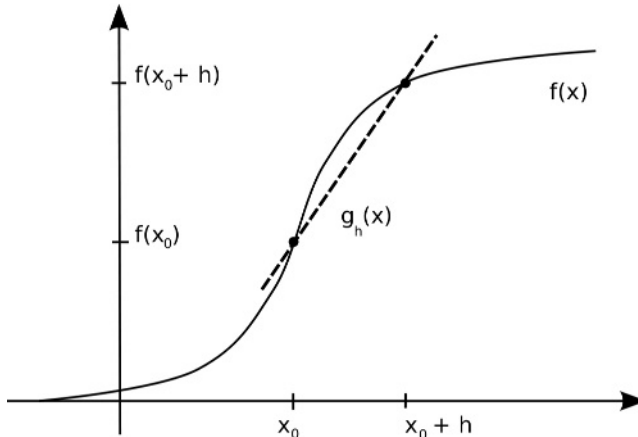


Abbildung 1.4: Die Gerade  $g_h(x)$  durch die beiden Punkte  $(x_0 + h, f(x_0 + h))$  und  $(x_0, f(x_0))$

Der Differenzenquotient  $\frac{f(x_0+h)-f(x_0)}{h}$  entspricht gerade der Steigung der Geraden  $g_h(x)$ . Wenn Sie den Grenzübergang  $h \rightarrow 0$  durchführen, erhalten Sie die Gerade  $g_0(x)$  mit:

$$g_0(x) := \lim_{h \rightarrow 0} g_h(x) = f'(x_0)(x - x_0) + f(x_0)$$

Das ist die Tangente zum Graphen der Funktion  $f$  an der Stelle  $x_0$ .



Die Ableitung  $f'(x_0)$  entspricht also der Steigung der Tangente an den Graphen der Funktion  $f$  im Punkt  $(x_0, f(x_0))$  und wird daher auch *Steigung der Funktion  $f$  an der Stelle  $x_0$*  genannt.

Natürlich können Sie auch bei einer Ableitung  $f'(x)$  fragen, ob diese wieder differenzierbar ist. Und falls dies so ist, auch für die Ableitung der Ableitung, und so weiter. Das rechtfertigt die folgende Definition der » $n$ -ten Ableitung«.



Unter der Voraussetzung, dass die Funktion  $f$  und die entsprechenden Ableitungsfunktionen von  $f$  jeweils differenzierbar sind, bezeichnet man:

- ✓  $f^{(0)}(x) := f(x)$
- ✓  $f^{(1)}(x) := f'(x) = \frac{d}{dx}f(x)$
- ✓  $f^{(2)}(x) := f''(x) = \frac{d^2}{dx^2}f(x)$
- ✓ Allgemein:  $f^{(n)}(x) := \frac{d}{dx}f^{(n-1)}(x) = \frac{d^n}{dx^n}f(x)$

Dabei heißt  $f^{(n)}(x)$  die  *$n$ -te Ableitung von  $f$* .

Eine Funktion  $f$  heißt  *$n$ -mal differenzierbar*, falls ihre  $n$ -te Ableitung  $f^{(n)}$  existiert.

Eine Funktion  $f$  heißt  *$n$ -mal stetig differenzierbar*, wenn  $f^{(n)}$  existiert und stetig ist.



Genau wie für die erste Ableitung  $f'(x)$  gilt auch für die zweite und jede weitere Ableitung einer entsprechend oft differenzierbaren Funktion  $f$ , dass diese nicht unbedingt stetig sein müssen.

## **Globale und lokale Extremstellen**

Bei der Untersuchung auf Extremstellen wird zwischen *globalen* oder *lokalen Extremstellen* unterschieden. Als globale Extremstellen der Funktion  $f$  werden die Stellen im Definitionsbereich von  $f$  bezeichnet, an denen die größten (beziehungsweise kleinsten) Funktionswerte angenommen werden.



Ist  $f: D(f) \subseteq \mathbb{R} \rightarrow \mathbb{R}$  eine reelle Funktion, dann heißt  $x_0 \in D(f)$  eine *globale Maximalstelle* von  $f$ , falls

$$f(x_0) \geq f(x) \quad \text{für alle} \quad x \in D(f)$$

ist. Gilt für eine Stelle  $x_0 \in D(f)$ :

$$f(x_0) \leq f(x) \quad \text{für alle} \quad x \in D(f)$$

dann heißt  $x_0$  eine *globale Minimalstelle* von  $f$ .

Der Funktionswert  $f(x_0)$  an einer globalen Extremstelle  $x_0$  heißt *globales Minimum* beziehungsweise *globales Maximum* von  $f$ .

Ein  $x_0 \in D(f)$  heißt *globale Extremalstelle* von  $f$ , wenn es eine globale Minimal- oder Maximalstelle ist.

*Lokale Extremstellen* sind solche, bei denen der Funktionswert zwar größer (oder kleiner) ist als alle anderen Funktionswerte in der Nähe, aber nicht unbedingt der größte (oder kleinste) Funktionswert, der überhaupt auftritt. Sozusagen ein Berggipfel, der nicht unbedingt der höchste Berg des Gebirges sein muss.



Ist  $f: \mathbb{R} \rightarrow \mathbb{R}$  eine reelle Funktion mit Definitionsbereich  $D(f) \subseteq \mathbb{R}$ , dann heißt  $x_0 \in D(f)$  *lokale Maximalstelle* von  $f$ , falls es ein  $\varepsilon > 0$  gibt, so dass für alle  $x \in U_\varepsilon(x_0) \cap D(f)$  gilt:  $f(x_0) \geq f(x)$ .

Ein Punkt  $x_0 \in D(f)$  heißt *lokale Minimalstelle* von  $f$ , falls es ein  $\varepsilon > 0$  gibt, so dass für alle  $x \in U_\varepsilon(x_0) \cap D(f)$  gilt:  $f(x_0) \leq f(x)$ .

Der Funktionswert  $f(x_0)$  heißt *lokales Maximum* beziehungsweise *lokales Minimum* von  $f$ .

Ein  $x_0 \in D(f)$  heißt *lokale Extremalstelle* von  $f$ , wenn es eine lokale Minimal- oder Maximalstelle ist.

Die Bestimmung von Extremstellen beruht auf einigen grundlegenden Eigenschaften lokaler Maxima beziehungsweise lokaler Minima. Dabei reicht es zunächst, wenn Sie sich die Eigenschaften lokaler Extrema ansehen, denn jedes globale Extremum ist auch ein lokales Extremum.



Ist die Funktion  $f: \mathbb{R} \rightarrow \mathbb{R}$  auf dem offenen Intervall  $D(f) = (a, b)$  definiert und an der Stelle  $x_0 \in (a, b)$  differenzierbar, dann gilt:

Ist  $x_0$  eine lokale Extremalstelle, dann ist  $f'(x_0) = 0$ .

Tatsächlich muss die Ableitung nicht einmal stetig sein, es reicht schon, dass die Ableitung  $f'(x_0)$  an einer Extremalstelle  $x_0$  von  $f$  existiert. Dann muss  $f'(x_0) = 0$  gelten.



Ein Punkt  $x_0$  des Definitionsbereichs von  $f$  mit verschwindender Ableitung  $f'(x_0) = 0$  heißt ein *stationärer* oder *kritischer Punkt*.



Ein stationärer Punkt von  $f$  ist nicht notwendig eine Extremalstelle.

Es gilt nur: Ist  $x_0$  eine lokale Extremalstelle,  $\implies f'(x_0) = 0$ , nicht die Umkehrung.

Außerdem kann eine Funktion eine lokale Extremstelle in einem Punkt  $x_0$  ihres Definitionsbereichs haben, ohne dort differenzierbar zu sein.

Bei der Suche nach den lokalen oder globalen Maxima und Minima einer Funktion auf einem Intervall sind die stationären Punkte mögliche Kandidaten. Da aber nicht jeder stationäre Punkt auch eine Extremstelle ist, müssen diese Punkte weiter untersucht werden.

Die zweite Ableitung ist ein Hilfsmittel, mit dem Sie an kritischen Punkten entscheiden können, ob ein Minimum oder ein Maximum vorliegt.



Ist die Funktion  $f: \mathbb{R} \rightarrow \mathbb{R}$  auf dem offenen Intervall  $(a, b)$  zweimal stetig differenzierbar, dann gilt für Punkte  $x_0 \in (a, b)$ :

- ✓ Ist  $f'(x_0) = 0$  und  $f''(x_0) < 0$ , so ist  $x_0$  eine lokale Maximalstelle.
- ✓ Ist  $f'(x_0) = 0$  und  $f''(x_0) > 0$ , so ist  $x_0$  eine lokale Minimalstelle.

Ist  $f''(x_0) = 0$ , dann können Sie nicht direkt entscheiden, ob ein Minimum, ein Maximum oder ein *Sattelpunkt* vorliegt.



Ein kritischer Punkt  $x_0 \in D(f)$  der Definitionsmenge einer differenzierbaren Funktion  $f$ , der weder eine Minimalstelle noch eine Maximalstelle von  $f$  ist, heißt *Sattelpunkt* von  $f$ .

Sattelpunkte sind Spezialfälle von *Wendepunkten*, die gleichzeitig kritische Punkte sind. An Wendepunkten ändert sich das Krümmungsverhalten des Graphen von einer Linkskurve in eine Rechtskurve oder umgekehrt.



Ist  $x_0 \in D(f)$  ein Punkt der Definitionsmenge einer zweimal stetig differenzierbaren Funktion  $f$  mit  $f''(x_0) = 0$  und hat die zweite Ableitung links von  $x_0$  ein anderes Vorzeichen als rechts von  $x_0$ , das heißt gilt entweder

$$f''(x) < 0 \text{ für } x < x_0 \quad \text{und} \quad f''(x) > 0 \text{ für } x > x_0$$

oder

$$f''(x) > 0 \text{ für } x < x_0 \quad \text{und} \quad f''(x) < 0 \text{ für } x > x_0$$

dann heißt  $x_0$  *Wendepunkt* von  $f$ . Dabei werden nur  $x$  in einer kleinen Umgebung von  $x_0$  betrachtet.

Ärgerlicherweise reicht für einen Sattelpunkt die Bedingung  $f''(x_0) = 0$  nicht aus – es könnte sich immer noch um eine Extremstelle handeln. Um das zu klären, müssen Sie die Funktion  $f$  weiter untersuchen.



Ist  $f$  genügend oft stetig differenzierbar und verschwinden die ersten  $n$  Ableitungen  $f'(x_0) = f''(x_0) = \dots = f^{(n)}(x_0) = 0$ , aber die  $n + 1$ -te Ableitung  $f^{(n+1)}(x_0) \neq 0$  nicht, dann besitzt  $f$  an der Stelle  $x_0$  einen Sattelpunkt, falls  $n$  eine gerade Zahl ist. Ist  $n$  eine ungerade Zahl, also  $n + 1$  gerade, dann hat  $f$  an der Stelle  $x_0$  ein Extremum und das Vorzeichen von  $f^{(n+1)}(x_0)$  gibt an, ob es sich um ein Maximum ( $f^{(n+1)}(x_0) < 0$ ) oder ein Minimum ( $f^{(n+1)}(x_0) > 0$ ) handelt.

Dabei ist in diesem Zusammenhang die Funktion  $f$  *genügend oft stetig differenzierbar*, wenn sie mindestens  $n + 1$ -mal stetig differenzierbar ist.

## Integration

Neben der Differentialrechnung ist die Integralrechnung der zweite große und sehr wichtige Themenbereich in der Analysis. In einer Hinsicht ist die Integralrechnung die Umkehrung der Differentialrechnung: Für eine gegebene Funktion  $f(x)$  suchen Sie dabei eine Funktion  $F(x)$ , eine sogenannte *Stammfunktion*, die als Ableitung die Funktion  $f$  besitzt. Diese Form der Integration wird auch *unbestimmte Integration* genannt.



Sind  $F: \mathbb{R} \rightarrow \mathbb{R}$  und  $f: \mathbb{R} \rightarrow \mathbb{R}$  auf dem Intervall  $[a, b]$  definierte Funktionen, so heißt  $F$  eine *Stammfunktion* zu  $f$ , wenn für alle  $x \in [a, b]$

$$F'(x) = f(x)$$

gilt.

Wahrscheinlich ist allerdings die andere Sichtweise der Integralrechnung die, die Sie zuerst kennengelernt haben: die Flächenberechnung unterhalb des Graphen einer gegebenen Funktion. Diese Sichtweise liegt der Integraldefinition als Grenzwert *Riemannscher Summen* zugrunde.



Eine *Riemannsche Summe* zu einer gegebenen Funktion  $f: I \rightarrow \mathbb{R}$  auf dem Intervall  $I = [a, b] \subset \mathbb{R}$  und einer Unterteilung in Teilintervalle  $[x_{i-1}, x_i]$  für  $1 \leq i \leq n$  durch gegebene Punkte

$$a = x_0 < x_1 < \dots < x_n = b$$

ist eine Summe

$$S_Z = \sum_{i=1}^n f(\bar{x}_i)(x_i - x_{i-1})$$

mit Zwischenpunkten  $\bar{x}_i \in [x_{i-1}, x_i]$ . Eine solche Unterteilung des Intervalls  $I$  in lauter Teilintervalle  $[x_{i-1}, x_i]$  wird *Zerlegung* von  $I$  genannt. Die *Feinheit* einer Zerlegung ist die maximale Länge  $x_i - x_{i-1}$  eines Teilintervalls der Zerlegung.

Eine Riemannsche Summe ist also nichts anderes, als eine Näherung an die Fläche unter dem Graphen von  $f$  durch viele kleine Rechtecke der Breite  $x_i - x_{i-1}$  und Höhe  $f(\bar{x}_i)$ , wie sie in Abbildung 1.5 dargestellt ist.

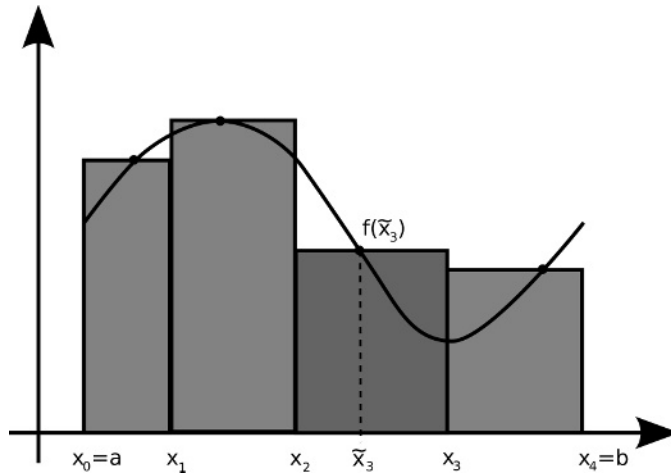


Abbildung 1.5: Approximation der Fläche  $F$  durch Rechtecksflächen

Falls alle möglichen Riemannschen Summen für immer kleinere Feinheiten der zugrundeliegenden Zerlegung gegen ein und denselben festen Wert konvergieren, dann erhalten Sie damit das Integral der Funktion  $f$ .



Es sei  $f: \mathbb{R} \rightarrow \mathbb{R}$  eine auf dem Intervall  $[a, b]$  beschränkte Funktion mit  $|f(x)| \leq M$  für alle  $a \leq x \leq b$ .

Konvergieren für alle Folgen von Zerlegungen  $(Z_n)$  mit  $\lim_{n \rightarrow \infty} \|Z_n\| = 0$  und beliebige Auswahl von Zwischenpunkten die Riemannschen Summen  $S_{Z_n}$  gegen einen gemeinsamen Grenzwert  $S$ , dann heißt  $f$  *integrierbar auf  $[a, b]$*  und die Zahl  $S$  heißt *bestimmtes Integral über  $f$  im Intervall  $[a, b]$* . In Formeln:

$$\int_a^b f(x) \, dx := S = \lim_{\|Z\| \rightarrow 0} \sum_{i=1}^n f(\bar{x}_i)(x_i - x_{i-1}).$$

Überraschend ist der Zusammenhang zwischen der Differentialrechnung und der zunächst rein geometrisch definierten Integralrechnung. Sie erhalten mit der Flächenberechnung gleichzeitig eine Umkehrung der Differentiation.

Diese Tatsache heißt in der reellen Analysis der *Hauptsatz der Differential- und Integralrechnung*.



Der *Hauptsatz der Differential- und Integralrechnung* erlaubt Ihnen, Integrale mit Hilfe von Stammfunktionen und umgekehrt Stammfunktionen durch Integrale zu berechnen:

Für jede auf einem Intervall  $I$  stetige Funktion  $f$  und jedes  $a \in I$  gilt:

✓ **Existenz von Stammfunktionen:** Die Funktion

$$F_a(x) := \int_a^x f(t) \, dt$$

mit  $x \in I$  ist eine Stammfunktion zu  $f$  auf  $I$ .

✓ **Eindeutigkeit:** Jede andere Stammfunktion zu  $f$  hat die Form  $F(x) = F_a(x) + c$  mit einer Konstanten  $c \in \mathbb{R}$ .

✓ **Integralberechnung:** Ist  $F$  eine beliebige Stammfunktion zu  $f$ , so gilt für  $a, b \in I$ :

$$\int_a^b f(x) \, dx = F(b) - F(a)$$

Mit dem Hauptsatz der Differential- und Integralrechnung haben Sie also eine Methode, um entweder Flächen unterhalb des Graphen einer gegebenen Funktion  $f$  oder Stammfunktionen von  $f$  zu bestimmen, vorausgesetzt, Sie können das Integral über die Funktion  $f$  berechnen.

Anstatt die Grenzwertdefinition des Integrals direkt zu verwenden, werden Sie üblicherweise zur Integration einer gegebenen Funktion  $f$  eine andere Integrationsmethode verwenden. Die beiden wichtigsten Methoden zur Integration sind die *partielle Integration* und die *Substitutionsformel*.

Die *partielle Integration* wird aus der Produktregel der Differentiation abgeleitet.



Gilt für die differenzierbaren Funktionen  $f(x)$  und  $g(x)$ , dass das Produkt  $f(x)g'(x)$  integrierbar ist, dann ist auch das Produkt  $f'(x)g(x)$  integrierbar und es gilt die Formel:

$$\int f'(x)g(x) \, dx = f(x)g(x) - \int f(x)g'(x) \, dx$$

Bei der *Substitution* hilft die Kettenregel der Differentiation. Für die *bestimmte Integration*, die Flächenberechnung, müssen Sie dabei auch die Grenzen der Integrale beachten.



Die Substitutionsformel der bestimmten Integration lautet:

$$\int_a^b f(g(x))g'(x) \, dx = F(g(x)) \Big|_a^b$$

Eine zweite Form der Substitutionsregel ist ebenfalls oft nützlich.



Eine wichtige Variante der *Substitutionsformel*:

Es sei  $f: [\alpha, \beta] \rightarrow \mathbb{R}$  und die umkehrbare Funktion  $g: [a, b] \rightarrow [\alpha, \beta]$  mit nicht verschwindender Ableitung  $g'$  und Umkehrfunktion  $g^{-1}$  und  $g(a) = \alpha$  und  $g(b) = \beta$  gegeben.

Hat  $f(g(t)) g'(t)$  auf  $[a, b]$  eine Stammfunktion  $H(t)$ , dann ist  $H(g^{-1}(x))$  auf  $[\alpha, \beta]$  eine Stammfunktion von  $f$  und es gilt:

$$\int_{g(a)}^{g(b)} f(x) \, dx = \int_a^b f(g(t)) g'(t) \, dt$$