Mathematik für Naturwissenschaftler II

Jörg Eschmeier

Universität des Saarlandes

Sommersemester 2019

Inhaltsverzeichnis

1 Matrizen, Vektorräume, lineare Gleichungssysteme

1.1 Matrizen

Eine $(m \times n)$-Matrix über $\mathbb{R}$ (oder $\mathbb{C}$) ist ein rechteckiges Schema

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ a_{21} & \dots & a_{2n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}, \quad a_{ij} \in \mathbb{R} \text{ (oder } \mathbb{C})$$

mit m *Zeilen* und n *Spalten*. Zwei Matrizen $A = (a_{ij}), B = (b_{ij})$ heißen gleich, wenn $a_{ij} = b_{ij}$ ist für alle $i \in \{1, \dots, m\}$ und $j \in \{1, \dots, n\}$. Ist $m = n$, so nennt man A eine *quadratische Matrix*. Wir schreiben

$$\mathbb{R}^{m,n} \text{ bzw. } \mathbb{C}^{m,n}$$

für die Menge aller $(m \times n)$-Matrizen mit Koeffizienten in $\mathbb{R}$ bzw. $\mathbb{C}$. Man nennt:

- eine $(1 \times n)$-Matrix $(a_1, \dots, a_n)$ auch *Zeilenvektor*,
- eine $(m \times 1)$-Matrix

  $$\begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix}$$

  auch *Spaltenvektor*,
- eine quadratische Matrix $A = (a_{ij})_{1 \le i,j \le n}$ *symmetrisch*, falls

  $$a_{ij} = a_{ji} \text{ für alle } i, j \in \{1, \dots, n\} \text{ gilt.}$$

Für eine $(m \times n)$-Matrix

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}$$

heißt die $(n \times m)$-Matrix

$$A^t = \begin{pmatrix} a_{11} & \dots & a_{m1} \\ \vdots & & \vdots \\ a_{1n} & \dots & a_{mn} \end{pmatrix}$$

die *transponierte Matrix*. Es ist

$$\begin{aligned} i\text{-te Zeile von } A^t &= i\text{-te Spalte von } A \qquad (i = 1, \dots, n) \\ j\text{-te Spalte von } A^t &= j\text{-te Zeile von } A \qquad (j = 1, \dots, m). \end{aligned}$$

Eine quadratische Matrix der Form

$$A = \begin{pmatrix} a_{11} & & 0 \\ & \ddots & \\ 0 & & a_{nn} \end{pmatrix} \text{(d.h. } a_{ij} = 0 \text{ für } i \neq j)$$

heißt *Diagonalmatrix*. Die Diagonalmatrix

$$I = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}$$

heißt *Einheitsmatrix*.

1.2 Rechenoperationen für Matrizen

Für $(m \times n)$-Matrizen $A = (a_{ij}), B = (b_{ij})$ und $\alpha \in \mathbb{C}$ setzt man

$$A + B = (a_{ij} + b_{ij})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}, \ \alpha A = (\alpha a_{ij})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}.$$

Man nennt die $(m \times n)$-Matrix

$$0_{mn} = \begin{pmatrix} 0 & \dots & 0 \\ \vdots & & \vdots \\ 0 & \dots & 0 \end{pmatrix}$$

die $(m \times n)$-*Nullmatrix*.

1.3 Rechenregeln

Für $(m \times n)$-Matrizen A, B, C und Zahlen α, β gilt:

(1) $(A + B) + C = A + (B + C)$ (Assoziativität der Addition)

(2) $A + B = B + A$ (Kommutativität der Addition)

(3) Es existiert ein *neutrales Element* 0 bzgl. der Addition, d.h. ein Element mit

$$A + 0 = 0 + A = A$$

für alle $(m \times n)$-Matrizen A, nämlich $0 = 0_{mn}$.

(4) Zu jeder $(m \times n)$-Matrix $A \,(= (a_{ij}))$ existiert genau eine $(m \times n)$-Matrix A' mit

$$A + A' = A' + A = 0_{mn} \qquad (\text{nämlich } A' = (-a_{ij})).$$

(5) $\alpha(A+B) = \alpha A + \alpha B,\ (\alpha+\beta)A = \alpha A + \beta A$ (*Distributivgesetze*)

(6) $(\alpha\beta)A = \alpha(\beta A)$ (*gemischtes Assoziativgesetz*)

(7) $1_{\mathbb{R}}\, A = A$ und $0_{\mathbb{R}}\, A = 0$, wobei $1_{\mathbb{R}}$ und $0_{\mathbb{R}}$ die Eins und Null in $\mathbb{R}$ sind und 0 die Nullmatrix bezeichnet.

(8) $(A+B)^t = A^t + B^t,\ (\alpha A)^t = \alpha A^t$.

Das neutrale Element bezüglich der Addition in (3) ist eindeutig bestimmt, denn ist auch $0'$ ein solches Element, so ist $0' = 0' + 0 = 0$.

Definition 1.1. (*Vektorräume*) Eine Menge V zusammen mit zwei Abbildungen

$$\begin{gathered} + : V \times V \to V,\ (x,y) \mapsto x+y \\ \cdot : \mathbb{R} \times V \to V(\text{ bzw. } \mathbb{C} \times V \to V),\ (\alpha, x) \mapsto \alpha x \end{gathered}$$

heißt ein $\mathbb{R}$-*Vektorraum* (bzw. $\mathbb{C}$-*Vektorraum*), falls “ $+$ “ und “ $\cdot$ “ die Eigenschaften (1)-(7) besitzen.

Beispiele 1.2. (i) $\mathbb{R}^{m,n}$ (bzw. $\mathbb{C}^{m,n}$) ist ein $\mathbb{R}$-Vektorraum (bzw. $\mathbb{C}$-Vektorraum).

(ii) $\mathbb{R}$ mit “ $+$ “ und “ $\cdot$ “ ist ein $\mathbb{R}$-Vektorraum (entsprechend $\mathbb{C}$ ein $\mathbb{C}$-Vektorraum).

(iii)

$$\mathbb{R}^n = \mathbb{R}^{n,1} = \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mid x_i \in \mathbb{R} \text{ für } i = 1, \dots, n \right\} \text{ mit komponentenweiser Addition}$$

und skalarer Multiplikation ist ein $\mathbb{R}$-Vektorraum (entsprechend $\mathbb{C}^n$ ein $\mathbb{C}$-Vektorraum). Man schreibt die Elemente von $\mathbb{R}^n$ (bzw. $\mathbb{C}^n$) oft auch als Zeilenvektoren und unterscheidet zwischen Spalten- und Zeilenvektoren nur, wenn nötig.

(iv) Die Menge aller $\mathbb{R}$-wertigen ($\mathbb{C}$-wertigen) Funktionen auf einem beliebigen Definitionsbereich D

$$\mathrm{Abb}(D, \mathbb{R}) = \{f | f : D \to \mathbb{R} \text{ Funktion}\} \text{ (entsprechend } \mathrm{Abb}(D, \mathbb{C}))$$

wird zu einem $\mathbb{R}$- (bzw. $\mathbb{C}$-)Vektorraum bzgl. den punktweise definierten Verknüpfungen

$$\begin{gathered} f + g : D \to \mathbb{R},\ (f+g)(x) = f(x) + g(x), \\ \alpha f : D \to \mathbb{R},\ (\alpha f)(x) = \alpha f(x). \end{gathered}$$

Mit den gleichen Verknüpfungen wird auch die Menge aller stetigen (oder differenzierbaren) Funktionen oder die Menge aller Polynome mit Koeffizienten in $\mathbb{R}$ bzw. $\mathbb{C}$ zu einem $\mathbb{R}$- (bzw. $\mathbb{C}$-) Vektorraum.

1.4 Erzeugendensysteme und Teilvektorräume

Sei V ein Vektorraum über $K = \mathbb{R}$ oder $K = \mathbb{C}$ und seien $x_1, \ldots, x_r \in V$. Ein Vektor $x \in V$ heißt *Linearkombination* von $x_1, \ldots, x_r$, falls

$$\alpha_1, \ldots, \alpha_r \in K \text{ existieren mit } x = \sum_{i=1}^{r} \alpha_i x_i.$$

Man nennt $\{x_1, \ldots, x_r\}$ ein *Erzeugendensystem* von V, falls jeder Vektor $x \in V$ Linearkombination von $x_1, \ldots, x_r$ ist.

Beispiele 1.3. (1) Für alle $\binom{x}{y} \in \mathbb{R}^2$ ist.

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ y \end{pmatrix} = x \begin{pmatrix} 1 \\ 0 \end{pmatrix} + y \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Also ist $\{\binom{1}{0}, \binom{0}{1}\}$ ein Erzeugendensystem von $\mathbb{R}^2$ (entsprechend von $\mathbb{C}^2$). Allgemeiner bilden die Vektoren

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ldots, e_n = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

ein Erzeugendensystem des $\mathbb{R}$-Vektorraumes $\mathbb{R}^n$ (bzw. $\mathbb{C}$-Vektorraumes $\mathbb{C}^n$).

(2) Aber auch die Vektoren $\binom{1}{0}, \binom{1}{1}$ bilden ein Erzeugendensystem von $\mathbb{R}^2$, denn

$$\begin{aligned} \begin{pmatrix} x \\ y \end{pmatrix} &= \alpha \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \beta \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\ &\Leftrightarrow x = \alpha + \beta \text{ und } y = \beta \\ &\Leftrightarrow \beta = y \text{ und } \alpha = x - y. \end{aligned}$$

Ist V ein K-Vektorraum ($K = \mathbb{R}$ oder $K = \mathbb{C}$) und $\varnothing \neq U \subset V$ eine Teilmenge mit

(i) $x, y \in U \Rightarrow x + y \in U$,

(ii) $x \in U, \alpha \in K \Rightarrow \alpha x \in U$,

so bildet U mit den induzierten Verknüpfungen

$$U \times U \to U, (x, y) \mapsto x + y, \;\; K \times U \to U, (\alpha, x) \mapsto \alpha x$$

wieder einen K-Vektorraum. Eine solche Teilmenge $U \subset V$ heißt *Teilvektorraum* von V.

Beispiele 1.4. (1) Die Menge

$$\left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 | \; x+y+z=0 \right\} \subset \mathbb{R}^3$$

ist ein Teilvektorraum.

(2) $\{f | \; f : \mathbb{R} \to \mathbb{R}$ ist stetig $\} \subset \; \mathrm{Abb}(\mathbb{R}, \mathbb{R})$ ist ein Teilvektorraum.

(3) Sind $x_1, \ldots, x_r \in V$ beliebige Elemente in einem Vektorraum, so ist

$$U = \{x \in V | \; x \text{ ist Linearkombination von } x_1, \ldots, x_r\} \subset V$$

ein Teilvektorraum. Man schreibt $\mathrm{LH}\{x_1, \ldots, x_r\}$ für diesen Teilvektorraum und nennt ihn die *lineare Hülle* von $x_1, \ldots, x_r$.

1.5 Lineare Unabhängigkeit und Basen

Definition 1.5. Sei V ein K-Vektorraum und seien $x_1, \ldots, x_r \in V$. Man nennt $\{x_1, \ldots, x_r\}$ eine *Basis* von V, falls

(i) $\{x_1, \ldots, x_r\}$ ein Erzeugendensystem von V ist,

(ii) aus $\sum_{i=1}^r \alpha_i x_i = \sum_{i=1}^r \beta_i x_i$ folgt, dass $\alpha_i = \beta_i$ für $i = 1, \ldots, r$ ist.

Beachte: (i) und (ii) sind q̈uivalent zu (i) und

$$\text{(ii)'} \qquad \text{aus } \sum_{i=1}^r \alpha_i x_i = 0 \text{ folgt } \alpha_i = 0 \text{ für } i = 1, \ldots, r.$$

Denn ist $\sum_{i=1}^r \alpha_i x_i = \sum_{i=1}^r \beta_i x_i$, so folgt

$$\sum_{i=1}^r (\alpha_i - \beta_i) x_i = \left(\sum_{i=1}^r \alpha_i x_i \right) - \left(\sum_{i=1}^r \beta_i x_i \right) = 0.$$

Beispiele 1.6. (1) $\binom{1}{0}, \binom{0}{1}$ bilden eine Basis von $\mathbb{R}^2$. Allgemeiner bilden

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ldots, e_n = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

eine Basis von $\mathbb{R}^n$ (und auch von $\mathbb{C}^n$, denn sie bilden Beispiel 1.3(1) ein Erzeugendensystem und wegen

$$\sum_{i=1}^n \alpha_i e_i = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}$$

folgt aus $\sum_{i=1}^n \alpha_i e_i = 0$, dass $\alpha_i = 0$ ist für $i = 1, \ldots, n$.

(2) Die Vektoren $(1,1), (1,0), (0,1)$ bilden eine Erzeugendensystem des $\mathbb{R}^2$, aber keine Basis, da etwa

$$\alpha(1,0) + \beta(0,1) + \gamma(1,1) = 0 \text{ ist mit } (\alpha, \beta, \gamma) = (1,1,-1) \neq (0,0,0).$$

Definition 1.7. (Lineare Unabhängigkeit) Sei V ein K-Vektorraum. Vektoren $x_1, \ldots, x_r \in V$ heißen *linear unabhängig*, falls aus

$$\sum_{i=1}^r \alpha_i x_i = 0$$

folgt, dass $\alpha_1 = \alpha_2 = \ldots = \alpha_r = 0$. *ist.*

Also bilden Vektoren $x_1, \ldots, x_r \in V$ eine Basis von V genau dann, wenn $x_1, \ldots, x_r$ ein linear unabhängiges Erzeugendensystem bilden.

Regeln 1.8. Für Vektoren $x_1, \ldots, x_r, x \in V$ gilt

(1) $x_1, \ldots, x_r$ bilden ein Erzeugendensystem von $V \Leftrightarrow \mathrm{LH}\{x_1, \ldots, x_r\} = V$.

(2) Ist $\{x_1, \ldots, x_r\}$ ein Erzeugendensystem von V und ist $U \subset V$ ein Teilvektorraum mit $x_1, \ldots, x_r \in U$, so ist $U = V$.

(3) die Vektoren $x_1, \ldots, x_r$ sind linear unabhängig genau dann, wenn keiner der Vektoren x_j $(j = 1, \ldots, r)$ Linearkombination der übrigen Vektoren x_i mit $i \neq j$ ist.

(4) Sind $x_1, \ldots, x_r$ linear unabhängig und ist $x \notin \mathrm{LH}\{x_1, \ldots, x_r\}$, so sind auch $x_1, \ldots, x_r, x$ linear unabhängig.

Satz 1.9. *Ist* $\{x_1, \ldots, x_m\}$ *ein Erzeugendensystem von* V *und sind* $y_1, \ldots, y_n \in V$ *linear unabhängig, so ist* $n \leq m$.

Idee: Wir nehmen an, dass $n > m$ ist. Wähle $\alpha_1, \ldots, \alpha_m \in K$ mit $y_1 = \sum_{i=1}^m \alpha_i x_i$. Dann gibt es ein $j \in \{1, \ldots, m\}$ mit

$$\alpha_j \neq 0.$$

Für den so gewählten Index j gilt

$$x_j = \frac{1}{\alpha_j}(y_1 - \sum_{\substack{i=1 \\ i \neq j}}^m \alpha_i x_i) \in \mathrm{LH}\{x_1, \ldots, x_{j-1}, y_1, x_{j+1}, \ldots, x_m\}.$$

Nach Regel 1.8(2) bilden dann aber auch die Vektoren

$$x_1, \ldots, x_{j-1}, y_1, x_{j+1}, \ldots, x_m$$

ein Erzeugendensystem von V.

Nacheinander kann man alle $x_1, \ldots, x_m$ durch die Vektoren $y_1, \ldots, y_m$ ersetzen und erhält

$$V = \text{ LH}\{y_1, \ldots, y_m\}.$$

Dann wäre aber y_{m+1} Linearkombination von $y_1, \ldots, y_m$ im Widerspruch zur Regel 1.8(3).

1.6 Dimension von Vektorräumen

Sei V ein Vektorraum über $K = \mathbb{R}$ oder $K = \mathbb{C}$.

Definition 1.10. Gibt es eine Maximalzahl n linear unabhängiger Vektoren in V, so heißt n die *Dimension* von V (geschrieben: $\dim V = n$). Gibt es keine solche Maximalzahl, so heißt V *unendlich dimensional* ($\dim V = \infty$).

Folgerung 1.11. (1) *Ist $\{x_1, \ldots, x_n\}$ eine Basis von V, so ist $\dim V = n$.*

(Sonst gäbe es linear unabhängige Vektoren $y_1, \ldots, y_m \in V$ mit $m > n$ im Widerspruch zu Satz 1.9.)

(2) *Sei $\dim V = n$. Sind $x_1, \ldots, x_n \in V$ linear unabhängig, so bilden $x_1, \ldots, x_n$ eine Basis von V.*

(Sonst gäbe es $x_{n+1} \in V$ mit $x_{n+1} \notin \text{LH}\{x_1, \ldots, x_n\}$. Nach Regel 1.8(4) wären $x_1, \ldots, x_{n+1}$ linear unabhängig im Widerspruch zu $\dim V = n$.) Jede Basis von V besteht aus n Vektoren.

(3) $\dim_{\mathbb{R}}(\mathbb{R}^n) = \mathbb{R}$ *und* $\dim_{\mathbb{C}}(\mathbb{C}^n) = n$ *(folgt direkt aus (1)).*

(4) *Ist $\{x_1, \ldots, x_r\}$ ein Erzeugendensystem von V, so gibt es $1 \leq i_1 < \ldots < i_n \leq r$ so, dass die Vektoren $x_{i_1}, \ldots, x_{i_n}$ eine Basis von V bilden.*

(Ist n die Maximalzahl linear unabhängiger Vektoren in $\{x_1, \ldots, x_r\}$ und sind $x_{i_1}, \ldots, x_{i_n}$ linear unabhängig, so bilden diese Vektoren nach Regel 1.8(4) ein Erzeugendensystem von V.)

Merkregel: Jedes Erzeugendensystem von V enthält eine Basis von V.

Ist $\dim V = n$, so kann man jedes System linear unabhängiger Vektoren $y_1, \ldots, y_r$ in V zu einer Basis $y_1, \ldots, y_n$ von V ergänzen. Dies zeigt der nächste Satz, den wir im Folgenden nicht weiter benutzen werden.

Satz (Austauschsatz von Steinitz)

Sei $x_1, \ldots, x_n$ eine Basis von V und seien $y_1, \ldots, y_r \in V$ linear unabhängig. Dann ist $r \leq n$ und es gibt $x_{i_1}, \ldots, x_{i_{n-r}}$ so, dass

$$y_1, \ldots, y_r, x_{i_1}, \ldots, x_{i_{n-r}}$$

eine Basis von V bilden.

Idee: Nach Satz 1.9 ist $r \leq n$.

Ist $r = n$, so bilden $y_1, \ldots, y_r$ eine Basis von V nach (1) und (2) von Folgerung 1.11.

Ist $r < n$, so gibt es ein $i \in \{1, \ldots, n\}$ so, dass $y_1, \ldots, y_r, x_i$ linear unabhängig sind, denn sonst wären $x_1, \ldots, x_n \in \text{ LH}\{y_1, \ldots, y_r\} = V$. Ist $r = n-1$, so ist der Beweis beendet. Sonst mache weiter so.

1.7 Lineare Gleichungssysteme

Seien $A = (a_{ij}) \in \mathbb{R}^{m,n}$ eine Matrix und $b = (b_i) \in \mathbb{R}^m$ ein Spaltenvektor. Dann heißt

$$\begin{aligned} a_{11}x_1 + \ldots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + \ldots + a_{2n}x_n &= b_2 \\ &\ldots \\ a_{m1}x_1 + \ldots + a_{mn}x_n &= b_m \end{aligned}$$

(abgekürzt: $Ax = b$) ein *lineares Gleichungssystem* aus m Gleichungen für $x_1, \ldots, x_n$. Man nennt:

- das lineare Gleichungssystem *homogen*, falls $b = 0$ ist,
- $A = (a_{ij})$ die *Koeffizientenmatrix*,
- das lineare Gleichungssystem *lösbar*, wenn die *Lösungsmenge*

  $$L = \{x = (x_i) \in \mathbb{R}^n |\ Ax = b\}$$

  nicht-leer ist.

Die Matrix

$$(A|b) := \left(\begin{array}{ccc|c} a_{11} & \ldots & a_{1n} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{m1} & \ldots & a_{mn} & b_m \end{array}\right) \in \mathbb{R}^{m,n+1}$$

bezeichnet man als die *erweiterte Koeffizientenmatrix* des linearen Gleichungssystems. Ziel im Folgenden ist es, ein Verfahren zur Bestimmung der Lösungsmenge herzuleiten.

Beispiel 1.12.

$$\begin{aligned} x_1 + 2x_2 - 3x_3 &= 4 \\ x_1 + 3x_2 + x_3 &= 11 \\ 2x_1 + 5x_2 + 4x_3 &= 13 \\ 2x_1 + 6x_2 + 2x_3 &= 22 \end{aligned}$$

Wir eliminieren x_1 aus den Gleichungen (2)-(4), indem wir

- Gleichung (2) ersetzen durch Gleichung (2) minus Gleichung (1)
- Gleichung (3) ersetzen durch Gleichung (3) minus 2 mal Gleichung (1)
- Gleichung (4) ersetzen durch Gleichung (4) minus 2 mal Gleichung (1).

In Kurzschreibweise für die erweiterte Koeffizientenmatrix bedeutet dies

$$\left(\begin{array}{ccc|c} 1 & 2 & -3 & 4 \\ 1 & 3 & 1 & 11 \\ 2 & 5 & 4 & 13 \\ 2 & 6 & 2 & 22 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 2 & -3 & 4 \\ 0 & 1 & 4 & 7 \\ 0 & 1 & 10 & 5 \\ 0 & 2 & 8 & 14 \end{array}\right).$$

Im nächsten Schritt eliminieren wir x_2 aus den Gleichungen (3) und (4):

$$\to \left(\begin{array}{ccc|c} 1 & 2 & -3 & 4 \\ 0 & 1 & 4 & 7 \\ 0 & 0 & 6 & -2 \\ 0 & 0 & 0 & 0 \end{array}\right).$$

Da sich bei diesen Zeilenumformungen die Lösungsmenge nicht verändert, hat das ursprüngliche Gleichungssystem die eindeutige Lösung

$$x_3 = -\frac{1}{3}, x_2 = 7 - 4x_3 = \frac{25}{3},\ x_1 = 4 - 2x_2 + 3x_3 = -\frac{41}{3}.$$

Definition 1.13. Unter einer *elementaren Zeilenumformung* einer Matrix $A = (a_{ij}) \in \mathbb{R}^{m,n}$ versteht man:

(1) die Vertauschung zweier Zeilen von A,

(2) die Multiplikation einer Zeile mit einer Zahl $\lambda \neq 0$,

(3) die Addition eines beliebigen Vielfachen einer Zeile zu einer **anderen** Zeile.

Merkregel: Elementare Zeilenumformungen ändern die Lösungsmenge nicht.

1.8 Der Gauß-Algorithmus

Sei das lineare Gleichungssystem

$$(*) \quad Ax = b \qquad (A \in \mathbb{R}^{m,n}, b \in \mathbb{R}^m)$$

gegeben. Wir nehmen an, dass ein $i \in \{1, \ldots, m\}$ existiert mit $a_{i1} \neq 0$. Sonst kommt x_1 gar nicht vor.

1. Schritt: Falls nötig, vertausche man zwei Gleichungen so, dass x_1 in der 1. Gleichung mit einem Koeffizienten $a_{11} \neq 0$ vorkommt.

2. Schritt: Ersetze für jedes $i = 2, \dots, m$ die i-te Zeile L_i durch

$$L_i + (-\frac{a_{i1}}{a_{11}})L_1.$$

Man erhält ein lineares Gleichungssystem mit derselben Lösungsmenge der Form

$$a'_{11}x_1 + a'_{12}x_2 + \ldots + a'_{1n}x_n = b'_1$$

$$(**) \begin{cases} a'_{2,j_2}x_{j_2} + & \ldots + & a'_{2n}x_n = b'_2 \\ \vdots & \vdots & \vdots \\ a'_{m,j_2}x_{j_2} + & \ldots + & a'_{mn}x_n = b'_m \end{cases}$$

wobei x_{j_2} die erste Unbekannte ist, die in einer der neuen Gleichungen $(2) - (m)$ mit einem Koeffizienten $\neq 0$ vorkommt.

Wende wieder Schritt 1 und 2 an auf das lineare Gleichungssystem $(**)$. Tritt eine Gleichung der Form

$$0x_1 + \ldots + 0x_n = b \neq 0$$

auf, so ist das ursprüngliche lineare Gleichungssystem $(*)$ nicht lösbar. Gleichungen der Form

$$0x_1 + \ldots + 0x_n = 0$$

kann man weglassen. Nach endlich vielen Schritten erhält man entweder ein unlösbares System oder ein lineares Gleichungssystem der Form $(***)$ (*Normalform*)

$$\begin{array}{lllr} c_{11}x_1 + c_{12}x_2 + & & \ldots & +c_{1n}x_n = d_1 \\ & c_{2,j_2}x_{j_2} + c_{2,j_2+1}x_{j_2+1} + & \ldots & +c_{2n}x_n = d_2 \\ & \ldots & & \\ & c_{r,j_r}x_{j_r} + c_{r,j_r+1}x_{j_r+1} + & \ldots & +c_{rn}x_n = d_r \end{array}$$

mit $1 < j_2 < \ldots < j_r \leq n, \ 1 \leq r \leq \min(m, n)$ und

$$c_{11} \neq 0, \ c_{2,j_2} \neq 0, \ldots, c_{r,j_r} \neq 0.$$

Man nennt die Variablen x_i mit $i \in \{1, \ldots, n\} \setminus \{1, j_2, \ldots, j_r\}$ die *freien Variablen* von $(***)$. Für ein lineares Gleichungssystem in Normalform $(***)$ gibt es zwei Fälle:

(i) $r = n$, das heißt $j_i = i$ für $i = 2, \dots, n$.

Dies ist nur möglich, wenn $m \geq n$ ist. In diesem Fall gibt es keine freien Variablen und $(***)$ hat eine eindeutige Lösung.

(ii) $r < n$, das heißt $(***)$ hat weniger Gleichungen als Unbekannte. Man erhält alle Lösungen von $(***)$, indem man für die freien Variablen beliebige reelle Zahlen einsetzt und für $x_1, x_{j_2}, \dots, x_{j_r}$ die zugehörigen eindeutigen Lösungen nimmt. In diesem Fall besitzt $(***)$ unendlich viele Lösungen.

Ein lineares Gleichungssystem hat also entweder keine, genau eine oder unendliche viele Lösungen.

Schritt 2 kann man auch ersetzen durch:

Schritt 2' Ersetze für jedes $i = 2, \dots, m$ die i-te Zeile L_i durch

$$-a_{i1}L_1 + a_{11}L_i.$$

Beispiel 1.14. (a)

$$\begin{aligned} 2x_1 + x_2 - 2x_3 + 3x_4 &= 1 \\ 3x_1 + 2x_2 - x_3 + 2x_4 &= 4 \\ 3x_1 + 3x_2 + 3x_3 - 3x_4 &= 5 \end{aligned}$$

umformen der erweiterten Koeffizientenmatrix mit dem Gauß-Algorithmus liefert:

$$\left(\begin{array}{cccc|c} 2 & 1 & -2 & 3 & 1 \\ 3 & 2 & -1 & 2 & 4 \\ 3 & 3 & 3 & -3 & 5 \end{array}\right) \to \left(\begin{array}{cccc|c} 2 & 1 & -2 & 3 & 1 \\ 0 & 1 & 4 & -5 & 5 \\ 0 & 3 & 12 & -15 & 7 \end{array}\right) \to \left(\begin{array}{cccc|c} 2 & 1 & -2 & 3 & 1 \\ 0 & 1 & 4 & -5 & 5 \\ 0 & 0 & 0 & 0 & -8 \end{array}\right).$$

Also ist das Gleichungssystem nicht lösbar.

(b)

$$\left(\begin{array}{ccc|c} 1 & 2 & -3 & 4 \\ 1 & 3 & 1 & 11 \\ 2 & 5 & -4 & 13 \\ 2 & 6 & 2 & 22 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 2 & -3 & 4 \\ 0 & 1 & 4 & 7 \\ 0 & 1 & 2 & 5 \\ 0 & 2 & 8 & 14 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 2 & -3 & 4 \\ 0 & 1 & 4 & 7 \\ 0 & 0 & -2 & -2 \\ 0 & 0 & 0 & 0 \end{array}\right).$$

Die letzte Matrix entspricht einem linearen Gleichungssystem in Normalform mit $r = 3 = n$. Also ist das Gleichungssystem eindeutig lösbar mit Lösungsmenge

$$L = \{(1, 3, 1)\}.$$

(c)

$$\left(\begin{array}{cccc|c} 1 & 2 & -2 & 3 & 2 \\ 2 & 4 & -3 & 4 & 5 \\ 5 & 10 & -8 & 11 & 12 \end{array}\right) \to \left(\begin{array}{cccc|c} 1 & 2 & -2 & 3 & 2 \\ 0 & 0 & 1 & -2 & 1 \\ 0 & 0 & 2 & -4 & 2 \end{array}\right) \to \left(\begin{array}{cccc|c} 1 & 2 & -2 & 3 & 2 \\ 0 & 0 & 1 & -2 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

Die letzte Matrix entspricht einem Gleichungssystem in Normalform mit $r = 2 < 4 = n$. Freie Variablen sind x_2 und x_4. Wählt man $x_4 = t, x_2 = s$ in $\mathbb{R}$ beliebig, so folgt

$$\begin{aligned} x_3 &= 1 + 2x_4 = 1 + 2t, \\ x_1 &= 2 - 2x_2 + 2x_3 - 3x_4 = 2 - 2s + 2(1 + 2t) - 3t \\ &= 4 - 2s + t. \end{aligned}$$

Also ist das Gleichungssystem lösbar mit Lösungsmenge

$$L = \{(4 - 2s + t, s, 1 + 2t, t) |\, s, t \in \mathbb{R}\}.$$

Spezialfall: Homogene Systeme

Homogene Systeme besitzen immer die triviale Lösung $(x_i)_{i=1}^n = 0$. Bringt man ein homogenes System in Normalform $(***)$, so ist

$$d_1 = \ldots = d_r = 0$$

und für

(i) $r = n$ hat man nur die triviale Lösung,

(ii) $r < n$ hat man auch nicht-triviale Lösungen.

Folgerung

Ein homogenes lineares Gleichungssystem mit weniger Gleichungen als Unbekannten hat immer nicht-triviale Lösungen.

Satz 1.15. *Seien $A = (a_{ij}) \in \mathbb{R}^{m,n}, b = (b_i) \in \mathbb{R}^m$. Die Lösungsmenge*

$$L_{\text{homogen}} = \{x = (x_i) \in \mathbb{R}^n |\, Ax = 0\}$$

des homogenen linearen Gleichungssystems $Ax = 0$ ist ein Teilvektorraum des $\mathbb{R}^n$. Ist $z = (z_i) \in \mathbb{R}^n$ Lösung des inhomogenen Systems

$$Ax = b,$$

so erhält man die gesamte Lösungsmenge von $Ax = b$ als

$$L = \{z + x |\, x \in L_{\text{homogen}}\} \ (=: z + L_{\text{homogen}}).$$

Idee: Offensichtlich ist $L_{\text{homogen}} \subset \mathbb{R}^n$ ein Teilvektorraum und

$$z + L_{\text{homogen}} \subset L.$$

Ist $\tilde{z} \in L$ beliebig, so ist $\tilde{z} - z \in L_{\text{homogen}}$ und

$$\tilde{z} = z + (\tilde{z} - z) \subset z + L_{\text{homogen}}.$$

Merkregel: {Lösungen von $Ax = b$} = spezielle Lösung + {Lösungen von $Ax = 0$}.

2 Matrixmultiplikation und inverse Matrizen

2.1 Matrixmultiplikation

Definition 2.1. (Matrixprodukt) Für Matrizen $A = (a_{ij}) \in \mathbb{R}^{\ell,m}$, $B = (b_{ij}) \in \mathbb{R}^{m,n}$ definiert man das Produkt $AB \in \mathbb{R}^{\ell,n}$ der Matrizen A und B als die $(\ell \times n)$-Matrix

$$AB = \left(\sum_{k=1}^{m} a_{ik} b_{kj} \right)_{\substack{1 \le i \le \ell \\ 1 \le j \le n}} \in \mathbb{R}^{\ell,n}.$$

Genauso definiert man auch das Matrixprodukt $AB \in \mathbb{C}^{\ell,n}$ von komplexen Matrizen $A \in \mathbb{C}^{\ell,m}, B \in \mathbb{C}^{m,n}$.

Beachte: Das Produkt AB zweier Matrizen A und B ist nur definiert, wenn die Spaltenzahl der links stehenden Matrix A gleich der Zeilenzahl der rechts stehenden Matrix B ist.

Rechenregeln 2.2. Für Matrizen $A, B, C, \ldots$ mit passenden Zeilen- und Spaltenzahlen gilt:

(1) **Assoziativität:** $(AB)C = A(BC)$.

(2) **Distributivität:**

$$(A + \tilde{A})B = AB + \tilde{A}B,$$
$$A(B + \tilde{B}) = AB + A\tilde{B}.$$

(3) **Gemischte Assoziativität:** Für $\lambda \in \mathbb{R}$ (bzw. $\mathbb{C}$) ist

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

(4) **Neutrales Element der Multiplikation:** Für $A \in \mathbb{R}^{n,n}$ (bzw. $\mathbb{C}^{n,n}$) und die Einheitsmatrix

$$I = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix} \in \mathbb{R}^{n,n}$$

gilt $IA = A = AI$.

(5) **Verträglichkeit mit der Transposition:**

$$(AB)^t = B^t A^t,$$

denn definitionsgemäß ist

$$(AB)^t_{ij} = (AB)_{ji} = \sum_{k=1}^{m} b_{ki} a_{jk} = \sum_{k=1}^{m} (B^t)_{ik} (A^t)_{kj} = (B^t A^t)_{ij}.$$

(6) Das Matrixprodukt ist im Allgemeinen **nicht kommutativ:**

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

Problem: Hat jede quadratische Matrix $0 \neq A \in \mathbb{R}^{n,n}$ ein multiplikatives Inverses, das heißt gibt es ein $A' \in \mathbb{R}^{n,n}$ mit $AA' = A'A = I$?. Im Allgemeinen nicht:

$$\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \left(\begin{array}{c|c} a+c & b+d \\ \hline 0 & 0 \end{array}\right) \neq I \text{ für alle } a,b,c,d \in \mathbb{R}.$$

Definition 2.3. Eine Matrix $A \in \mathbb{R}^{n,n}$ heißt *invertierbar* oder *regulär*, falls eine Matrix $A^{-1} \in \mathbb{R}^{n,n}$ existiert mit

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

Man überlegt sich leicht, dass die Matrix A^{-1} in diesem Fall eindeutig bestimmt ist und dass für die Inversenbildung von Matrizen die folgenden Regeln gelten.

Rechenregeln 2.4. (1) A, B invertierbar $\Rightarrow AB$ invertierbar und $(AB)^{-1} = B^{-1}A^{-1}$.

Denn: $(B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}A)B = B^{-1}B = I = \ldots = (AB)(B^{-1}A^{-1})$.

Hierbei ist die Reihenfolge der Faktoren wichtig.

(2) A invertierbar $\Rightarrow A^t$ invertierbar und $(A^t)^{-1} = (A^{-1})^t$.

Denn: $(A^{-1})^t A^t = (AA^{-1})^t = I^t = I = \ldots = A^t(A^{-1})^t$.

(3) Für $A \in \mathbb{R}^{n,n}$ definiert die Linksmultiplikation mit A eine Abbildung

$$L_A : \mathbb{R}^n \to \mathbb{R}^n, x \mapsto Ax.$$

Ist A invertierbar, so ist die Abbildung L_A bijektiv.

Injektivität: Aus $Ax = Ay$ folgt $A(x-y) = Ax - Ay = 0$ und damit

$$x - y = A^{-1}A(x-y) = A^{-1}0 = 0.$$

Surjektivität: Für $y \in \mathbb{R}^n$ ist $y = A(A^{-1}y) = L_A(A^{-1}y)$.

(4) Seien A und L_A wie in (3). Dann gilt:

$$L_A \text{ ist injektiv } \Leftrightarrow \text{ Die Spalten von } A \text{ sind linear unabhängig in } \mathbb{R}^n.$$

$$\text{Denn ist } A = (A_1, \ldots, A_n) \text{ mit Spalten } A_j, \text{ so ist } \sum_{j=1}^{n} x_j A_j = A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

$$L_A \text{ ist surjektiv} \Leftrightarrow \text{Die Spalten von } A \text{ bilden ein Erzeugendensystem von } \mathbb{R}^n.$$

(5) Nach Folgerung 1.11 (2) und (4) gilt:

L_A ist injektiv $\Leftrightarrow L_A$ ist surjektiv $\Leftrightarrow L_A$ ist bijektiv.

2.2 Berechnung der inversen Matrix

Sei $A = (a_{ij}) \in \mathbb{R}^{n,n}$ invertierbar. Man kann den Gauß-Algorithmus (Abschnitt 1.8) benutzen, um die inverse Matrix $A^{-1} \in \mathbb{R}^{n,n}$ zu berechnen. Sei

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \text{ die } k\text{-te Spalte von } A^{-1}.$$

Dann ist

$$A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = k\text{-te Spalte von } AA^{-1} = e_k$$

oder äquivalent: Der Vektor

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$$

löst das lineare Gleichungssystem

$$(*) \begin{cases} a_{11}x_1 & + \ldots + & a_{1n}x_n & = 0 \\ \vdots & & & \vdots \\ a_{k1}x_1 & + \ldots + & a_{kn}x_n & = 1 \\ \vdots & & & \vdots \\ a_{n1}x_1 & + \ldots + & a_{nn}x_n & = 0 \end{cases}$$

Nach Regel 2.4(3) hat dieses lineare Gleichungssystem eine eindeutige Lösung. Also gibt es im Gauß-Algorithmus (Abschnitt1.8) keine freien Variablen. Durch elementare Zeilenumformungen von Typ (1),(2) oder (3) aus Definition 1.13 bringt man das Gleichungssystem $(*)$ in die Normalform

$$(***) \begin{cases} c_{11}x_1 + c_{12}x_2 + \ldots + c_{1n}x_n = d_1 \\ \qquad\qquad c_{22}x_2 + \ldots + c_{2n}x_n = d_2 \\ \qquad\qquad\qquad \cdots \\ \qquad\qquad\qquad\qquad c_{nn}x_n = d_n \end{cases}$$

mit $c_{ii} \neq 0$ für $i = 1, \ldots, n$. Durch elementare Zeilenumformungen vom Typ (2) und (3) (Definition 1.13) bringt man $(***)$ in die Form

$$\begin{array}{rcrcl} x_1+0+0+ & \dots & +0 & = & b_1 \\ x_2+0+ & \dots & +0 & = & b_2 \\ & \dots & & & \\ & & x_n & = & b_n \end{array}$$

Merkregel: Ist $A = (a_{ij}) \in \mathbb{R}^{n,n}$ invertierbar, so bringt man die Matrix

$$\left(\begin{array}{ccc|ccc} a_{11} & \dots & a_{1n} & 1\,0\,0 & \dots & 0 \\ a_{21} & \dots & a_{2n} & 0\,1\,0 & \dots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} & 0\,0\,0 & \dots & 1 \end{array}\right)$$

durch elementare Zeilenumformungen vom Typ (1), (2) oder (3) aus Definition 1.13 wie oben beschrieben in die Form

$$\left(\begin{array}{ccc|ccc} 1\,0\,0 & \dots & 0 & b_{11} & \dots & b_{1n} \\ 0\,1\,0 & \dots & 0 & & \dots & \\ \vdots\,\vdots\,\vdots & \ddots & \vdots & \vdots & & \vdots \\ 0\,0\,0 & \dots & 1 & b_{n1} & \dots & b_{nn} \end{array}\right).$$

Dann ist $B = (b_{ij}) = A^{-1}$. Man kann sich zusätzlich überlegen:
Wenn dieses Verfahren für eine Matrix $A \in \mathbb{R}^{n,n}$ funktioniert, dann ist A invertierbar.

Beispiel 2.5. Finde die Inverse der Matrix

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

Durch elementare Zeilenumformungen erhält man

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 2 & -1 & 3 & 0 & 1 & 0 \\ 4 & 1 & 8 & 0 & 0 & 1 \end{array}\right) \to \left(\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & -1 & -1 & -2 & 1 & 0 \\ 0 & 1 & 0 & -4 & 0 & 1 \end{array}\right)$$

$$\to \left(\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & 1 & 1 & 2 & -1 & 0 \\ 0 & 0 & -1 & -6 & 1 & 1 \end{array}\right) \to \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -11 & 2 & 2 \\ 0 & 1 & 0 & -4 & 0 & 1 \\ 0 & 0 & 1 & 6 & -1 & -1 \end{array}\right)$$

$$\text{Also ist } A^{-1} = \begin{pmatrix} -11 & 2 & 2 \\ -4 & 0 & 1 \\ 6 & -1 & -1 \end{pmatrix}.$$

Bemerkung 2.6 (Umkehrung von Regel 2.4(3)): Ist für $A \in \mathbb{R}^{n,n}$ die Linksmultiplikation L_A bijektiv, so ist A invertierbar. Zur Begründung wähle man $b_1, \ldots b_n \in \mathbb{R}^n$ mit $Ab_k = e_k$ und beachte, dass die Matrix $B = (b_1, \ldots, b_n)$ mit den Spaltenvektoren b_k die Identität $AB = (Ab_1, \ldots, Ab_n) = I$ erfüllt. Wegen $L_A L_B = L_{AB} = id$ ist dann $L_B = L_A^{-1}$ und damit auch $BA = L_B L_A(I) = I$.

3 Determinanten

3.1 Definition und Eigenschaften der Determinante

In der Mathematik zeigt man, dass für jedes $n \geq 1$ eine Abbildung

$$\mathbb{R}^{n,n} \to \mathbb{R}, A \mapsto \det A \quad \text{(manchmal auch als } |A| \text{ geschrieben)}$$

existiert, die jeder $(n \times n)$-Matrix $A = (a_{ij}) \in \mathbb{R}^{n,n}$ eine reelle Zahl $\det A \in \mathbb{R}$ zuordnet (die *Determinante* der Matrix A) und die die folgenden Eigenschaften besitzt:
Schreibt man die $(n \times n)$-Matrix $A = (a_{ij}) \in \mathbb{R}^{n,n}$ in der Form

$$A = \begin{pmatrix} A_1 \\ \vdots \\ A_n \end{pmatrix},$$

wobei $A_i = (a_{i1}, \ldots, a_{in})$ für $i = 1, \ldots, n$ die i-Zeile von A ist, so gilt

(1)

$$\det \begin{pmatrix} A_1 \\ \vdots \\ A_i + A_i' \\ \vdots \\ A_n \end{pmatrix} = \det \begin{pmatrix} A_1 \\ \vdots \\ A_i \\ \vdots \\ A_n \end{pmatrix} + \det \begin{pmatrix} A_1 \\ \vdots \\ A_i' \\ \vdots \\ A_n \end{pmatrix},$$

$$\det \begin{pmatrix} A_1 \\ \vdots \\ \lambda A_i \\ \vdots \\ A_n \end{pmatrix} = \lambda \det \begin{pmatrix} A_1 \\ \vdots \\ A_i \\ \vdots \\ A_n \end{pmatrix} \qquad (\lambda \in \mathbb{R}).$$

Insbesondere ist $\det A = 0$, falls eine Zeile 0 ist.

(2) Ensteht A' aus A durch Vertauschen zweier Zeilen, so ist

$$\det A' = -\det A.$$

(3) Die Determinante der Einheitsmatrix ist

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

(4) Sind zwei Zeilen von A gleich, so ist $\det A = 0$ (folgt aus (2)).

(5) Addiert man ein Vielfaches einer Zeile zu einer anderen Zeile, so ändert sich die Determinante nicht (folgt aus (1) und (4)).

(6) Sind die Zeilen von A linear abhängig, so ist $\det A = 0$ (folgt aus (5) und der letzten Zeile von (1)).

(7) $\det(A^t) = \det A$.

(8) Alle Regeln bleiben richtig, wenn man überall Zeilen durch Spalten ersetzt.

(9) Für $A, B \in \mathbb{R}^{n,n}$ ist $\det(AB) = \det(A)\det(B)$.

Man kann sogar zeigen (siehe Abschnitt 3.2):

Satz 3.1. *Die Abbildung*

$$\mathbb{R}^{n,n} \to \mathbb{R}, A \mapsto \det A$$

ist die einzige Abbildung, die als Funktion der Zeilen die Eigenschaften (1), (2) und (3) besitzt.

Entsprechend kann man die Determinante $\det A \in \mathbb{C}$ komplexer $n \times n$-Matrizen $A \in \mathbb{C}^{n,n}$ definieren. Alle bisher und im Folgenden angegebenen Regeln und Sätze bleiben richtig für die Determinanten komplexer Matrizen.

Beispiele 3.2. $n = 1$: $\quad A = (a_{11}) \quad \Rightarrow \det A = a_{11}$

$n = 2$:

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

$n = 3$:

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \quad \Rightarrow$$

$$\begin{aligned} \det A &= a_{11} \det \begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix} - a_{21} \det \begin{pmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{pmatrix} + a_{31} \det \begin{pmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{pmatrix} \\ &= a_{11}(a_{22}a_{33} - a_{32}a_{23}) - a_{21}(a_{12}a_{33} - a_{32}a_{13}) + a_{31}(a_{12}a_{23} - a_{22}a_{13}) \end{aligned}$$

Diese Rechenregel nennt man Entwicklung nach der 1. Spalte.

Allgemein gilt für $n = 3$: Entwicklung nach der j-ten Spalte ($j = 1, 2, 3$)

$$\det(A) = (-1)^{1+j} a_{1j} \det(A_{1j}) + (-1)^{2+j} a_{2j} \det(A_{2j}) + (-1)^{3+j} a_{3j} \det(A_{3j}).$$

Entwicklung nach der i-ten Zeile ($i = 1, 2, 3$)

$$\det(A) = (-1)^{i+1} a_{i1} \det(A_{i1}) + (-1)^{i+2} a_{i2} \det(A_{i2}) + (-1)^{i+3} a_{i3} \det(A_{i3}).$$

Hierbei bezeichnet $A_{ij} \in \mathbb{R}^{2,2}$ die Matrix, die aus A durch Streichen der i-ten Zeile und j-ten Spalte entsteht.

Die Vorzeichen bei a_{ij} ergeben sich nach dem Schema

$$\begin{pmatrix} + & - & + \\ - & + & - \\ + & - & + \end{pmatrix}.$$

3.2 Berechnung von Determinanten (noch einmal Gauß)

Sei $A = (a_{ij}) \in \mathbb{R}^{n,n}$ und sei

$$C = \begin{pmatrix} c_{11} & c_{12} & \dots & \dots & \dots & c_{1n} \\ 0 & \dots & c_{2,j_2} & \dots & \dots & c_{2n} \\ \vdots & & \vdots & & & \vdots \\ 0 & \dots & \dots & c_{r,j_r} & \dots & c_{rn} \\ 0 & \dots & \dots & \dots & \dots & 0 \\ \vdots & & & & & \vdots \\ 0 & \dots & \dots & \dots & \dots & 0 \end{pmatrix}$$

mit $c_{11} \neq 0, c_{2,j_2} \neq 0, \dots, c_{r,j_r} \neq 0$ die Matrix, die durch Anwenden des Gauß-Algorithmus (Abschnitt1.8) aus A hervorgeht. Hierbei werden nur Zeilenumformungen vom Typ (1) oder (3) aus Definition 1.13 angewendet. Also ist:

(i) $\det A = (-1)^k \det C$, falls k die Anzahl der ausgeführten Zeilenvertauschungen ist.

(ii) Gibt es freie Variablen ($\Leftrightarrow r < n$), so ist $\det A = \det C = 0$.

(iii) Gibt es keine freien Variablen ($\Leftrightarrow r = n$), so ist

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

eine obere Dreiecksmatrix mit $c_{ii} \neq 0$ für $i = 1, \ldots, n$ und

$$\det A = (-1)^k \det C = (-1)^k c_{11} \cdots c_{nn} \neq 0,$$

denn nach den Regeln (1), (3) und (5) aus Abschnitt 3.1 ist

$$\det C = c_{11} \cdots c_{nn} \det \begin{pmatrix} 1 & & * \\ & \ddots & \\ 0 & & 1 \end{pmatrix} = c_{11} \cdots c_{nn} \det \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix} = c_{11} \cdots c_{nn}$$

Insbesondere folgt hieraus, dass die Determinatenfunktion $\mathbb{R}^{n,n} \to \mathbb{R}, A \mapsto \det\, A$ eindeutig bestimmt ist durch die Eigenschaften (1), (2) und (3) aus Abschnitt 3.1.

3.3 Dreiecks- und Diagonalmatrizen

Für obere bzw. untere Dreiecksmatrizen

$$A = \begin{pmatrix} a_{11} & & * \\ & \ddots & \\ 0 & & a_{nn} \end{pmatrix} \text{(das heißt } a_{ij} = 0 \text{ für } i > j)$$

bzw.

$$A = \begin{pmatrix} a_{11} & & 0 \\ & \ddots & \\ * & & a_{nn} \end{pmatrix} \text{(das heißt } a_{ij} = 0 \text{ für } j > i)$$

gilt:

$$\det A = a_{11} \cdots a_{nn}.$$

Merkregel: Die Determinante einer oberen (unteren) Dreiecksmatrix oder einer Diagonalmatrix ist das Produkt der Diagonalelemente.

3.4 Invertierbare Matrizen

Für $A \in \mathbb{R}^{n,n}$(oder $\mathbb{C}^{n,n}$) gilt:

$$A \text{ ist invertierbar} \Leftrightarrow \det(A) \neq 0$$

($\Leftrightarrow$ Das homogene lineare Gleichungssystem $Ax = 0$ hat nur die triviale Lösung).

Idee: “ $\Rightarrow$ “ Ist A invertierbar, so gibt es keine freien Variablen und wie in Abschnitt 3.2 folgt

$$\det A = (-1)^k \det C \neq 0.$$

" $\Leftarrow$ " Ist $\det A \neq 0$, so ist $r = n$ und nach Abschnitt 1.8 ist das lineare Gleichungssystem $Ax = b$ für jedes $b \in \mathbb{R}^n$ (bzw. $b \in \mathbb{C}^n$) eindeutig lösbar. Nach Bemerkung 2.6 ist A invertierbar.

Satz 3.3 (Zusammenfassung). *Für $A \in \mathbb{R}^{n,n}$ (oder $\mathbb{C}^{n,n}$) sind äquivalent:*

(i) *A ist invertierbar,*

(ii) $\det A \neq 0$,

(iii) *die Zeilen von A sind linear unabhängig,*

(iv) *die Spalten von A sind linear unabhängig,*

(v) *das lineare Gleichungssystem $Ax = b$ hat für jedes $b \in \mathbb{R}^n$ (bzw. $b \in \mathbb{C}^n$) mindestens eine Lösung,*

(vi) *das lineare Gleichungssystem $Ax = b$ hat für jedes $b \in \mathbb{R}^n$ (bzw. $b \in \mathbb{C}^n$) höchstens eine Lösung,*

(vii) *das homogene lineare Gleichungssystem $Ax = 0$ hat hat nur die triviale Lösung $x = 0$,*

(viii) *das lineare Gleichungssystem $Ax = b$ hat für jedes $b \in \mathbb{R}^n$ (bzw. $b \in \mathbb{C}^n$) genau eine Lösung,*

(ix) *die Abbildung $L_A : \mathbb{R}^n \to \mathbb{R}^n$ (bzw. $\mathbb{C}^n \to \mathbb{C}^n$), $x \mapsto Ax$ ist bijektiv.*

Insbesondere folgt, dass eine quadratische Matrix $A \in \mathbb{R}^{n,n}$ (oder $\mathbb{C}^{n,n}$) genau dann invertierbar ist, wenn ihre Zeilenvektoren (oder äquivalent Spaltenvektoren) eine Basis vom $\mathbb{R}^n$ (bzw. $\mathbb{C}^n$) bilden.

3.5 Entwicklung nach Zeilen und Spalten

Sei $A = (a_{ij}) \in \mathbb{R}^{n,n}$ (oder $\mathbb{C}^{n,n}$). Für $i, j \in \{1, \dots, n\}$ sei

$$A_{ij} = \begin{pmatrix} a_{11} & \dots & a_{1j} & \dots & a_{1n} \\ & & | & & \\ a_{i1} & --- & a_{ij} & --- & a_{in} \\ & & | & & \\ a_{n1} & \dots & a_{nj} & \dots & a_{nn} \end{pmatrix} \in \mathbb{R}^{n-1,n-1} \text{ (oder } \mathbb{C}^{n-1,n-1})$$

die Matrix, die aus A durch Streichen der i-ten Zeile und j-ten Spalte entsteht.

Satz 3.4. (Laplacescher Entwicklungssatz) *Für $i, j \in \{1, \dots, n\}$ gilt:*

$$\begin{aligned} \det A &= \sum_{k=1}^{n} (-1)^{i+k} a_{ik} \det(A_{ik}) \text{ (Entwicklung nach der } i\text{-ten Zeile)} \\ &= \sum_{k=1}^{n} (-1)^{k+j} a_{kj} \det(A_{kj}) \text{ (Entwicklung nach der } j\text{-ten Spalte)} \end{aligned}$$

Idee: Zeige, dass die beiden rechten Seiten jeweils eine Funktion

$$\mathbb{R}^{n,n} \to \mathbb{R} \quad (\text{bzw. } \mathbb{C}^{n,n} \to \mathbb{C})$$

mit den Eigenschaften (1)-(3) aus Abschnitt 3.1 definieren und wende den Eindeutigkeitssatz (Satz 3.1) an.

3.6 Cramersche Regel und Berechnung der inversen Matrix

Sei $A = (a_{ij}) \in \mathbb{R}^{n,n}$ (oder $\mathbb{C}^{n,n}$) eine Matrix mit Spaltenvektoren $A_1, \ldots, A_n \in \mathbb{R}^n$ (bzw. $\mathbb{C}^n$). Für $b = (b_i), x = (x_i) \in \mathbb{R}^n$ (bzw. $\mathbb{C}^n$) gilt

$$\begin{aligned} a_{11}x_1 + \ldots + a_{1n}x_n &= b_1 \\ \vdots \qquad\quad & \quad \vdots \\ a_{n1}x_1 + \ldots + a_{nn}x_n &= b_n \end{aligned}$$

genau dann, wenn $Ax = b$ ist bzw. wenn

$$x_1 A_1 + \ldots + x_n A_n = b$$

gilt. Ersetzt man den i-ten Spaltenvektor der Matrix A durch den Spaltenvektor $b \in \mathbb{R}^n$, so folgt mit den Regeln (1), (4) und (7) aus Abschnitt 3.1

$$\begin{aligned} &\det(A_1, \ldots, b, \ldots, A_n) = \det(A_1, \ldots, x_1 A_1 + \ldots + x_n A_n, \ldots, A_n) \\ &= \sum_{\substack{j=1 \\ j \neq i}}^{n} x_j \det(A_1, \ldots, A_j, \ldots, A_n) + x_i \det A = x_i \det A \end{aligned}$$

Satz 3.5. (Cramersche Regel) *Sei $A \in \mathbb{R}^{n,n}$ (oder $\mathbb{C}^{n,n}$) invertierbar und $b \in \mathbb{R}^n$ (bzw. $\mathbb{C}^n$). Dann ist die eindeutige Lösung $x = (x_i) \in \mathbb{R}^n$ (bzw. $\mathbb{C}^n$) des linearen Gleichungssystems $Ax = b$ gegeben durch*

$$x_i = \frac{\det(A_1, \ldots, A_{i-1}, b, A_{i+1}, \ldots, A_n)}{\det(A)} \qquad (i = 1, \ldots, n).$$

Folgerung 3.6. Sei $A \in \mathbb{R}^{n,n}$ (oder $\mathbb{C}^{n,n}$) invertierbar mit Spaltenvektoren $A_1, \ldots, A_n$. Die Spalten $x^{(j)}$ $(j = 1, \ldots, n)$ der inversen Matrix A^{-1} sind die eindeutigen Lösungen der linearen Gleichungssysteme $Ax^{(j)} = e_j$. Nach Satz 3.5 sind die Komponenten von $x^{(j)}$ gegeben durch

$$x_i^{(j)} = \frac{\det(A_1, \ldots, A_{i-1}, e_j, A_{i+1}, \ldots, A_n)}{\det(A)} \qquad (i = 1, \ldots, n).$$

Durch Entwickeln nach der i-ten Spalte erhält man

$$x_i^{(j)} = (-1)^{i+j} \frac{\det(A_{ji})}{\det(A)} \qquad (i, j = 1, \ldots, n),$$

wobei A_{ji} die $(n-1) \times (n-1)$-Matrix bezeichnet, die aus A durch Streichen der j-ten Zeile und i-ten Spalte hervorgeht. Also ist die inverse Matrix zu A gegeben durch

$$A^{-1} = (x_i^{(j)})_{i,j=1}^n = \left((-1)^{i+j} \frac{\det(A_{ji})}{\det(A)} \right)_{i,j=1}^n .$$

4 Skalarprodukt und Vektorprodukt

4.1 Skalarprodukt

Sei V ein Vektorraum über $\mathbb{R}$. Eine Abbildung

$$\langle \cdot, \cdot \rangle : V \times V \to \mathbb{R}, (x, y) \mapsto \langle x, y \rangle$$

heißt *Skalarprodukt*, falls für alle $u_1, u_2, u, v \in V$ und $\lambda \in \mathbb{R}$ gilt

(i) $\langle u_1 + u_2, v \rangle = \langle u_1, v \rangle + \langle u_2, v \rangle, \langle \lambda u, v \rangle = \lambda \langle u, v \rangle$,

(ii) $\langle u, v \rangle = \langle v, u \rangle$,

(iii) $\langle u, u \rangle \geq 0$ und $\langle u, u \rangle = 0 \Leftrightarrow u = 0$.

Beachte: Aus (i) und (ii) folgt auch, dass für alle $u, v_1, v_2 \in V$ und $\lambda \in \mathbb{R}$ gilt

$$\langle u, v_1 + v_2 \rangle = \langle u, v_1 \rangle + \langle u, v_2 \rangle, \langle u, \lambda v \rangle = \lambda \langle u, v \rangle.$$

Insbesondere ist $\langle u, 0 \rangle = 0 = \langle 0, u \rangle$ für alle $u \in V$.

Beispiele 4.1. (1) Auf dem $\mathbb{R}$-Vektorraum $\mathbb{R}^n$ definiert

$$\left\langle \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \right\rangle = \sum_{i=1}^n x_i y_i$$

ein Skalarprodukt (das *kanonische Skalarprodukt* auf $\mathbb{R}^n$). Für $x = (x_i), y = (y_i) \in \mathbb{R}^n$ nennt man

$$\|x\| = (\sum_{i=1}^n x_i^2)^{1/2}$$

die *Länge* oder *euklidische Norm* des Vektors x und

$$\|x - y\| = (\sum_{i=1}^n (x_i - y_i)^2)^{1/2}$$

den *euklidischen Abstand* von x und y.

(2) Auf dem $\mathbb{R}$-Vektorraum $V = \{f | f : [a, b] \to \mathbb{R} \text{ stetig}\}$ wird ein Skalarprodukt definiert durch

$$\langle f, g \rangle = \int_a^b f(t) g(t) dt.$$

4.2 Cauchy-Schwarzsche Ungleichung

Sei V ein $\mathbb{R}$-Vektorraum und $\langle \cdot, \cdot \rangle$ ein Skalarprodukt auf V. In der Mathematik zeigt man, dass für alle $u, v \in V$ gilt:

$$|\langle u, v \rangle| \leq \sqrt{\langle u, u \rangle}\sqrt{\langle v, v \rangle} \text{ (Cauchy-Schwarzsche Ungleichung).}$$

Es gilt " = " statt " $\leq$ " in dieser Ungleichung genau dann, wenn u und v linear abhängig sind. Insbesondere gilt für das kanonische Skalarprodukt auf dem $\mathbb{R}^n$

$$|\langle u, v \rangle| \leq \|u\| \; \|v\|$$

für alle Vektoren $u, v \in \mathbb{R}^n$. Für $u, v \in \mathbb{R}^n \setminus \{0\}$ nennt man die eindeutige Zahl $\varphi \in [0, \pi]$ mit

$$\cos\varphi = \frac{\langle u, v \rangle}{\|u\| \; \|v\|}$$

den *Winkel* zwischen u und v.

4.3 Orthogonalität

Zwei Vektoren $u, v \in \mathbb{R}^n$ heißen *orthogonal* (geschrieben $u \perp v$), falls

$$\langle u, v \rangle = 0.$$

Für $u \neq 0 \neq v$ ist dies äquivalent dazu, dass der Winkel zwischen u und v gleich $\frac{\pi}{2}$ ist. Für $u, v \in \mathbb{R}^n$ mit $u \perp v$ gilt

$$\langle u + v, u + v \rangle = \langle u, u \rangle + \langle v, v \rangle$$

oder äquivalent

$$\|u + v\|^2 = \|u\|^2 + \|v\|^2 \qquad (\textit{Pythagoras}).$$

Sei $V \subset \mathbb{R}^n$ ein Teilvektorraum mit $\dim v = r$. Eine Basis $v_1, \ldots, v_r$ von V heißt

- *Orthogonalbasis*, falls $\langle v_i, v_j \rangle = 0$ ist für alle $i \neq j$,
- *Orthonormalbasis*, falls $\langle v_i, v_j \rangle = \begin{cases} 0, & i \neq j \\ 1, & i = j \end{cases}$.

Satz 4.2. (Gram-Schmidt) *Sei $V \subset \mathbb{R}^n$ ein Teilvektorraum und seien $v_1, \ldots, v_r \in V$ Vektoren, die eine Basis von V bilden. Dann bilden die Vektoren*

$$\begin{aligned}
u_1 &:= v_1, \\
u_2 &:= v_2 - \frac{\langle v_2, u_1 \rangle}{\langle u_1, u_1 \rangle} u_1, \\
u_3 &:= v_3 - \frac{\langle v_3, u_2 \rangle}{\langle u_2, u_2 \rangle} u_2 - \frac{\langle v_3, u_1 \rangle}{\langle u_1, u_1 \rangle} u_1, \\
&\ldots \\
u_r &:= v_r - \sum_{i=1}^{r-1} \frac{\langle v_r, u_i \rangle}{\langle u_i, u_i \rangle} u_i
\end{aligned}$$

eine Orthogonalbasis von V, und die Vektoren

$$\frac{u_1}{\|u_1\|}, \ldots, \frac{u_r}{\|u_r\|}$$

bilden eine Orthonormalbasis von V.

Nützliche Beobachtung 4.3: Von 0 verschiedene Vektoren $u_1, \ldots, u_r \in \mathbb{R}^n$ mit $\langle u_i, u_j \rangle = 0$ für $i \neq j$ sind immer linear unabhängig, denn aus $\sum_{i=1}^r \alpha_i u_i = 0$, folgt

$$0 = \langle \sum_{i=1}^r \alpha_i u_i, u_j \rangle = \sum_{i=1}^r \alpha_i \langle u_i, u_j \rangle = \alpha_j \langle u_j, u_j \rangle$$

und damit $\alpha_j = 0$ für $j = 1, \ldots, r$.

4.4 Das Vektorprodukt

Das Vektorprodukt ist eine Abbildung

$$\mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3, (a, b) \mapsto a \times b,$$

die folgendermaßen definiert ist:

(i) Sind a, b linear abhängig, so ist $a \times b = 0$.

(ii) Sind a, b linear unabhängig und ist $\varphi \in]0, \pi[$ der Winkel zwischen a und b, so ist $a \times b \in \mathbb{R}^3$ definiert durch

- $a \times b$ steht senkrecht auf der durch a und b gegebenen Ebene,
- $a, b, a \times b$ bilden in dieser Reihenfolge ein Rechtssystem (das heißt $a \times b$ zeigt in die Richtung, in die sich eine Schraube bewegt bei einer Drehung, die a auf kürzerem Wege in b überführt).
- $\|a \times b\| = \|a\| \; \|b\| \sin \varphi$

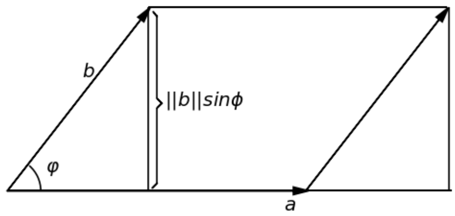


Abbildung 1: Norm des Vektorproduktes

Beachte: Das Produkt $\|a\| \, \|b\| \sin \varphi$ ist die Fläche des Parallelogramms

Man kann sich folgende Rechenregeln für das Vektorprodukt überlegen. Es gilt:

(i) $(a_1 + a_2) \times b = a_1 \times b + a_2 \times b, \; a \times (b_1 + b_2) = a \times b_1 + ax \times b_2$,

(ii) $(\lambda a) \times b = \lambda(a \times b) = a \times (\lambda b)$,

(iii) $b \times a = -(a \times b)$, insbesondere: $a \times a = 0$,

(iv) $a \perp a \times b$ und $b \perp a \times b$

für alle $a, b, a_1, a_2, b_1, b_2 \in \mathbb{R}^3$ und $\lambda \in \mathbb{R}$.

4.5 Das Vektorprodukt in kartesischen Koordinaten

Seien

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

die kanonischen Basisvektoren in $\mathbb{R}^3$. Dann ist

$$\begin{aligned} e_1 \times e_2 &= e_3, \\ e_1 \times e_3 &= -e_2, \\ e_2 \times e_3 &= e_1. \end{aligned}$$

Damit folgt für beliebige Vektoren $a = (a_i), b = (b_i) \in \mathbb{R}^3$

$$\begin{aligned} a \times b &= (\sum_{i=1}^{3} a_i e_i) \times (\sum_{i=1}^{3} b_i e_i) \\ &= (a_2 b_3 - a_3 b_2) e_1 - (a_1 b_3 - a_3 b_1) e_2 + (a_1 b_2 - a_2 b_1) e_3. \end{aligned}$$

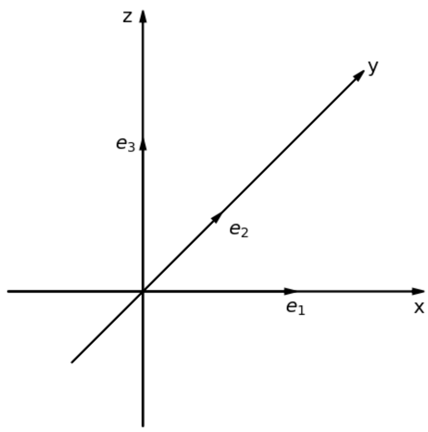


Abbildung 2: Einheitsvektoren

Merkregel: Formal erhält man diese Darstellung von $a \times b$, wenn man

$$\begin{pmatrix} a_1 & b_1 & e_1 \\ a_2 & b_2 & e_2 \\ a_3 & b_3 & e_3 \end{pmatrix}$$

wie eine übliche Matrix behandelt und "ihre Determinante" durch Entwickeln nach der 3. Spalte berechnet.

Beispiele 4.3. (a) Sei $M \subset \mathbb{R}^n$ ein Teilvektorraum mit Orthonormalbasis $\{u_1, \dots, u_r\}$ und sei $x \in \mathbb{R}^n$. Dann ist

$$p = \sum_{i=1}^{r} \langle x, u_i \rangle u_i \in M$$

der eindeutige Punkt in M mit minimalem Abstand zu x. Denn wegen

$$\langle x - p, u_j \rangle = \langle x, u_j \rangle - \langle \sum_{i=1}^{r} \langle x, u_i \rangle u_i, u_j \rangle = 0$$

für $j = 1, \dots, r$ ist $x - p \perp M$, das heißt es ist $x - p \perp m$ für alle $m \in M$. Für $m \in M$ ist dann $p - m \in M$ und mit Pythagoras folgt

$$\|x - m\|^2 = \|(x - p) + (p - m)\|^2 = \|x - p\|^2 + \|p - m\|^2 \geq \|x - p\|^2.$$

(b) Greift eine Kraft $F \in \mathbb{R}^3$ im Punkt $r \in \mathbb{R}^3$ an einem starren Körper an, der drehbar ist um den Koordinatenursprung $0 \in \mathbb{R}^3$, so wirkt definitionsgemäß das Drehmoment

$$D = r \times F \in \mathbb{R}^3.$$

Die Norm von D

$$\|D\| = \|r\| \|F\| \sin \varphi$$

ist maximal, wenn der Winkel φ zwischen dem Ortsvektor r des Angriffspunktes und der wirkenden Kraft F gleich $\frac{\pi}{2}$ ist und minimal (nämlich $= 0$) für $\varphi \in \{0, \pi\}$.

(c) Mit Teil (a) kann man zeigen, dass der Abstand d eines Punktes $p \in \mathbb{R}^3$ zu der Geraden

$$G = \{a + tu \mid t \in \mathbb{R}\} \qquad (u \in \mathbb{R}^3 \setminus \{0\}, a \in \mathbb{R}^3)$$

gegeben ist durch

$$d = \frac{\|(p-a) \times u\|}{\|u\|}$$

und dass der Abstand d eines Punktes $p \in \mathbb{R}^3$ zu der Ebene

$$E = \{a + (t_1 u_1 + t_2 u_2) \mid t_1, t_2 \in \mathbb{R}\} \qquad (u_1, u_2 \in \mathbb{R}^3 \text{ linear unabhängig und } a \in \mathbb{R}^3 \text{ beliebig})$$

gegeben ist durch

$$d = \frac{|\langle p-a, u_1 \times u_2 \rangle|}{\|u_1 \times u_2\|}$$

5 Eigenwerte und Eigenvektoren

5.1 Eigenwerte

Eine Abbildung $\varphi : \mathbb{R}^n \to \mathbb{R}^m$ heißt *linear*, falls für alle $x, y \in \mathbb{R}^n$ und $\lambda \in \mathbb{R}$ gilt

$$\varphi(x+y) = \varphi(x) + \varphi(y), \ \varphi(\lambda x) = \lambda \varphi(x).$$

Beispiele 5.1. (a) Die Abbildungen $\varphi_i : \mathbb{R}^2 \to \mathbb{R}^2$,

$$\begin{aligned}
\varphi_1(x) &= -x \qquad \text{(Spiegelung am Nullpunkt)}\\
\varphi_2\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} &= \begin{pmatrix} x_1 + a x_2 \\ x_2 \end{pmatrix}, \ a \in \mathbb{R} \qquad \text{(Scherung)}\\
\varphi_3\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} &= e^{i\varphi}(x_1 + i x_2) = \begin{pmatrix} \cos(\varphi)x_1 - \sin(\varphi)x_2 \\ \cos(\varphi)x_2 + \sin(\varphi)x_1 \end{pmatrix} \qquad \text{(Drehung um den Winkel } \varphi\text{)}
\end{aligned}$$

sind linear.

(b) Für $A \in \mathbb{R}^{m,n}$ ist die Linksmultiplikation

$$L_A : \mathbb{R}^n \to \mathbb{R}^m, \ x \mapsto Ax$$

linear.

(c) Ist $\varphi : \mathbb{R}^n \to \mathbb{R}^m$ linear und ist

$$A = (\varphi(e_1), \ldots, \varphi(e_n)) \in \mathbb{R}^{m,n}$$

die Matrix mit den Spaltenvektoren $\varphi(e_j)$ $(j = 1, \ldots, n)$, so gilt

$$\varphi = L_A.$$

Definition 5.2. Sei $\varphi : \mathbb{R}^n \to \mathbb{R}^n$ linear ($A \in \mathbb{R}^{n,n}$ eine Matrix). Eine Zahl $\lambda \in \mathbb{R}$ heißt *Eigenwert* von φ (von A), falls ein Vektor $x \in \mathbb{R}^n \setminus \{0\}$ existiert mit

$$\varphi(x) = \lambda x \qquad (Ax = \lambda x).$$

In diesem Fall heißt x *Eigenvektor* zum Eigenwert λ.

Beispiele 5.3. (1) Sei

$$A = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \in \mathbb{R}^{2,2}.$$

Dann ist

- $\binom{1}{1}$ Eigenvektor zum Eigenwert 0, denn $A\binom{1}{1} = \binom{0}{0} = 0\binom{1}{1}$,
- $\binom{2}{-2}$ Eigenvektor zum Eigenwert 2, denn $A\binom{2}{-2} = \binom{4}{-4} = 2\binom{2}{-2}$.

Gibt es weitere Eigenwerte? Eine Antwort auf diese Frage geben wir in Beispiel 5.5(1).

(2) Gibt es immer Eigenwerte? Nein, denn ist etwa

$$A = A_\alpha = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \in \mathbb{R}^{2,2},$$

so ist $\varphi = L_A$ die Drehung um den Winkel α. Dann ist anschaulich klar und auch nicht schwierig zu zeigen, dass φ genau dann einen Eigenwert hat, wenn $\alpha \in 2\mathbb{Z}\pi (= \{2k\pi | k \in \mathbb{Z}\})$ oder $\alpha \in (2\mathbb{Z}+1)\pi (= \{(2k+1)\pi | k \in \mathbb{Z}\})$ ist. Im ersten Fall ist $\varphi = \mathrm{id}$ und der einzige Eigenwert ist 1, im zweiten Fall ist $\varphi = -\mathrm{id}$ und der einzige Eigenwert ist -1.

5.2 Charakteristisches Polynom

Seien $A = (a_{ij}) \in \mathbb{R}^{n,n}$ und $x = (x_i)_{1 \le i \le n} \in \mathbb{R}^n \setminus \{0\}, \lambda \in \mathbb{R}$. Dann gilt

$$Ax = \lambda x$$

genau dann, wenn

$$\begin{array}{lllll} (a_{11} - \lambda)x_1 & +a_{12}x_2 & +\ldots+ & a_{1n}x_n & = 0, \\ a_{21}x_1 & +(a_{22} - \lambda)x_2 & +\ldots+ & a_{2n}x_n & = 0, \\ & \ldots & & & \\ a_{n1}x_1 & +a_{n2}x_2 & +\ldots+ & (a_{nn} - \lambda)x_n & = 0 \end{array}$$

Dies ist ein homogenes lineares Gleichungssystem mít Koeffizientenmatrix

$$A - \lambda I = \begin{pmatrix} (a_{11} - \lambda) & a_{12} & \ldots & a_{1n} \\ a_{21} & (a_{22} - \lambda) & \ldots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \ldots & (a_{nn} - \lambda) \end{pmatrix} \in \mathbb{R}^{n,n}.$$

Es gibt nicht triviale Lösungen genau dann, wenn (Zusammenfassung 3.3)

$$A - \lambda I \text{ nicht invertierbar ist } \Leftrightarrow \det(A - \lambda I) = 0.$$

Dies ist genau der Fall $r < n$ im Gauß-Algorithmus (Abschnitte 1.8 und 3.2).
Merkregel: $\lambda \in \mathbb{R}$ ist Eigenwert von $A \in \mathbb{R}^{n,n} \Leftrightarrow \det(A - \lambda I) = 0$.

Satz 5.4. *Sei* $A = (a_{ij}) \in \mathbb{R}^{n,n}$.

(a) Die Abbildung

$$p_A : \mathbb{R} \to \mathbb{R}, p_A(\lambda) = \det(A - \lambda I)$$

ist ein Polynom n-ten Grades (das charakteristische Polynom *von A) der Form*

$$p_A(\lambda) = (-1)^n \lambda^n + (-1)^{n-1}(a_{11} + \ldots + a_{nn})\lambda^{n-1} + \ldots + \det(A).$$

(b) Die Eigenwerte von A sind genau die Nullstellen des charakteristischen Polynoms p_A.

(c) Für jeden Eigenwert λ von A ist

$$\mathrm{Eig}(A, \lambda) := \{x \in \mathbb{R}^n | (A - \lambda I)x = 0\} \subset \mathbb{R}^n$$

ein Teilvektorraum (der Eigenraum *von A zu λ).*

Für $A \in \mathbb{C}^{n,n}$ und lineare Abbildungen $\varphi : \mathbb{C}^n \to \mathbb{C}^n$ definiert man Eigenwerte $\lambda \in \mathbb{C}$ und Eigenvektoren völlig analog. In diesem Fall ist

$$p_A : \mathbb{C} \to \mathbb{C}, p_A(\lambda) = \det(A - \lambda I)$$

ein komplexes Polynom und alles bleibt richtig in Satz 5.4 wichtig, wenn man überall $\mathbb{R}$ durch $\mathbb{C}$ ersetzt.

Beispiele 5.5. (1) Sei

$$A = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \in \mathbb{R}^{2,2}.$$

Das charakteristische Polynom von A

$$p_A(\lambda) = \det \left(\begin{array}{c|c} 1-\lambda & -1 \\ \hline -1 & 1-\lambda \end{array}\right) = (1-\lambda)^2 - 1 = \lambda^2 - 2\lambda = \lambda(\lambda - 2)$$

besitzt die Nullstellen $\lambda = 0$ und $\lambda = 2$. Dies sind genau die Eigenwerte von A.

(2) Sei

$$A = A_{\pi/2} = \left(\begin{array}{c|c} 0 & -1 \\ \hline 1 & 0 \end{array}\right) \in \mathbb{R}^{2,2} \quad \text{(Drehung um den Winkel } \frac{\pi}{2}\text{)}.$$

Das charakteristische Polynom von A

$$p_A(\lambda) = \det \left(\begin{array}{c|c} -\lambda & -1 \\ \hline 1 & -\lambda \end{array} \right) = \lambda^2 + 1$$

hat keine Nullstellen in $\mathbb{R}$. Also hat A keine Eigenwerte in $\mathbb{R}$. Wegen $p_A(\lambda) = (\lambda - i)(\lambda + i)$ hat $A \in \mathbb{C}^{2,2}$ als komplexe Matrix genau die Eigenwerte $\lambda = i$ und $\lambda = -i$.

(3) Das charakteristische Polynom von

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

ist

$$p_A(\lambda) = \det \left(\begin{array}{c|c} 1-\lambda & 4 \\ \hline 2 & 3-\lambda \end{array} \right) = (1-\lambda)(3-\lambda) - 8 = \lambda^2 - 4\lambda - 5 = (\lambda+1)(\lambda-5).$$

Also besitzt A genau die Eigenwerte $\lambda = -1$ und $\lambda = 5$. Die Eigenvektoren zu λ sind die Lösungen $x \neq 0$ des linearen Gleichungssystems

$$(A - \lambda I)x = 0.$$

Mit dem Gauß-Algorithmus ergibt sich für $\lambda = -1$:

$$\left(\begin{array}{c|c} 1-(-1) & 4 \\ \hline 2 & 3-(-1) \end{array} \right) = \begin{pmatrix} 2 & 4 \\ 2 & 4 \end{pmatrix} \rightarrow \begin{pmatrix} 2 & 4 \\ 0 & 0 \end{pmatrix}.$$

Also ist $\mathrm{Eig}(A, -1) = \{(-2t, t) | t \in \mathbb{R}\}$. Eigenvektoren zum Eigenwert $\lambda = -1$ sind die Vektoren $t\binom{-2}{1}$ mit $t \in \mathbb{R} \setminus \{0\}$.

Für $\lambda = 5$ erhält man entsprechend:

$$\left(\begin{array}{c|c} 1-5 & 4 \\ \hline 2 & 3-5 \end{array} \right) = \begin{pmatrix} -4 & 4 \\ 2 & -2 \end{pmatrix} \rightarrow \begin{pmatrix} -1 & 1 \\ 0 & 0 \end{pmatrix}.$$

Also ist $\mathrm{Eig}(A, 5) = \{(t, t) | t \in \mathbb{R}\}$. Eigenvektoren zum Eigenwert $\lambda = 5$ sind die Vektoren $t\binom{1}{1}$ mit $t \in \mathbb{R} \setminus \{0\}$.

5.3 Symmetrische Matrizen

Reelle Matrizen brauchen keine Eigenwerte zu besitzen. Für symmetrische Matrizen ist die Situation besser. Man zeigt in der Mathematik:

Satz 5.6. *Ist $A = (a_{ij}) \in \mathbb{R}^{n,n}$* symmetrisch *(d.h. $A = A^t$) oder ist $A = (a_{ij}) \in \mathbb{C}^{n,n}$* hermitesch *(d.h. $a_{ji} = \overline{a}_{ij}$ für alle i, j), so gibt es $\lambda_1, \ldots, \lambda_k \in \mathbb{R}$ paarweise verschieden und $n_1, \ldots, n_k \in \mathbb{N} \setminus \{0\}$ mit*

(i) $p_A(\lambda) = (-1)^n(\lambda - \lambda_1)^{n_1}(\lambda - \lambda_2)^{n_2} \cdots (\lambda - \lambda_k)^{n_k} \qquad (\Rightarrow n_1 + \ldots + n_k = n)$,

(ii) $\dim \operatorname{Eig}(A, \lambda_i) = n_i$ *für* $i = 1, \ldots, k$ *(über* $\mathbb{C}$ *bildet man den komplexen Eigenraum*

$$\operatorname{Eig}(A, \lambda_i) = \{x \in \mathbb{C}^n | (A - \lambda_i I)x = 0\}),$$

(iii) Eigenvektoren zu verschiedenen Eigenwerten sind orthogonal zueinander.

5.4 Diagonalisierung

Sei $A \in \mathbb{R}^{n,n}$ symmetrisch (oder $A \in \mathbb{C}^{n,n}$ hermitesch). Seien $\lambda_1, \ldots, \lambda_k$ die verschiedenen Eigenwerte von A und

$$n_i = \dim \operatorname{Eig}(A, \lambda_i) \qquad (i = 1, \ldots, k).$$

Mit dem Gram-Schmidt Orthonormalisierungsverfahren (Satz 4.2) erhält man für $i = 1, \ldots, k$ Orthonormalbasen

$$x_{i1}, \ldots, x_{in_i} \quad \text{von} \quad \operatorname{Eig}(A, \lambda_i).$$

Die Vektoren (beachte, dass $n_1 + \ldots + n_k = n$ ist)

$$\{u_1, \ldots, u_n\} = \{x_{11}, \ldots, x_{1n_1}, \ldots, x_{k1}, \ldots, x_{kn_k}\}$$

bilden eine Orthonormalbasis von $\mathbb{R}^n$ (bzw. $\mathbb{C}^n$). Sei

$$U = (u_1, \ldots, u_n) \in \mathbb{R}^{n,n} \quad (\text{ bzw. } \mathbb{C}^{n,n})$$

die Matrix mit den Spaltenvektoren $u_1, \ldots, u_n$. Dann gilt:

$$\begin{aligned}
AU &= (Au_1, \ldots, Au_n) \\
&= (\lambda_1 u_1, \ldots, \lambda_1 u_{n_1}, \lambda_2 u_{n_1+1}, \ldots, \lambda_2 u_{n_1+n_2}, \ldots, \lambda_k u_n) \\
&= U \begin{pmatrix}
\lambda_1 & & & & & & \\
& \ddots & & & & & \\
& & \lambda_1 & & & & \\
& 0 & & \ddots & & 0 & \\
& & & & \lambda_k & & \\
& & & & & \ddots & \\
& & & & & & \lambda_k
\end{pmatrix}.
\end{aligned}$$

Da die Spalten von $U = (u_1, \ldots, u_n)$ eine Basis von $\mathbb{R}^n$ (bzw. $\mathbb{C}^n$) bilden, ist die Matrix U invertierbar (Zusammenfassung 3.3)

$$\Rightarrow U^{-1}AU = \begin{pmatrix} \lambda_1 & & & & & & \\ & \ddots & & & & & \\ & & \lambda_1 & & & & \\ & 0 & & \ddots & & 0 & \\ & & & & \lambda_k & & \\ & & & & & \ddots & \\ & & & & & & \lambda_k \end{pmatrix}.$$

Definition 5.7. Eine Matrix $U \in \mathbb{R}^{n,n}$ (bzw. $U \in \mathbb{C}^{n,n}$) heißt *orthogonal* (bzw. *unitär* im komplexen Fall), falls die Spalten von U eine Orthonormalbasis von $\mathbb{R}^n$ (bzw. $\mathbb{C}^n$) bilden.

Bemerkung 5.8. Nach Satz 3.3 (Zusammenfassung) sind orthogonale und unitäre Matrizen invertierbar. Man kann zeigen: Eine Matrix $U \in \mathbb{R}^{n,n}$ ist orthogonal genau dann, wenn sie invertierbar ist und $U^{-1} = U^t$ gilt. Eine Matrix $U = (u_{ij}) \in \mathbb{C}^{n,n}$ ist unitär genau dann, wenn sie invertierbar ist und $U^{-1} = (\overline{u}_{ji})_{i,j=1}^n$ ist.

Satz 5.9. *Ist $A \in \mathbb{R}^{n,n}$ symmetrisch (bzw. $A \in \mathbb{C}^{n,n}$ hermitesch), so gibt es eine orthogonale Matrix $U \in \mathbb{R}^{n,n}$ (bzw. eine unitäre Matrix $U \in \mathbb{C}^{n,n}$) so, dass*

$$U^{-1}AU = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_k \end{pmatrix}$$

Diagonalgestalt hat. Dabei sind die Zahlen λ_i genau die Eigenwerte von A entsprechend der Dimension $n_i = \dim \mathrm{Eig}(A, \lambda_i)$ wiederholt.

Bemerkung 5.10. *Das zu Beginn von Abschnitt 5.4 erklärte Diagonalisierungsverfahren funktioniert nicht nur für symmetrische bzw. hermitesche Matrizen, sondern allgemeiner für jede Matrix $A \in \mathbb{R}^{n,n}$ (bzw. $A \in \mathbb{C}^{n,n}$), für die eine Basis*

$$\{u_1, \ldots, u_n\} = \{x_{11}, \ldots, x_{1n_1}, \ldots, x_{k1}, \ldots, x_{kn_k}\}$$

von $\mathbb{R}^n$ (bzw. $\mathbb{C}^n$) aus Eigenvektoren

$$x_{i1}, \ldots, x_{in_i} \in \mathrm{Eig}(A, \lambda_i) \quad (i = 1, \ldots, k)$$

zu paarweise verschiedenen Eigenwerten $\lambda_1, \ldots, \lambda_k$ existiert. Ist $U = (u_1, \ldots, u_n) \in \mathbb{R}^{n,n}$ (bzw. $\mathbb{C}^{n,n}$)

die Matrix mit den Spaltenvektoren $u_1, \ldots, u_n$, so folgt genau wie im Abschnitt vor Definition 5.7, dass

$$U^{-1}AU = \begin{pmatrix} \lambda_1 & & & & & & \\ & \ddots & & & & & \\ & & \lambda_1 & & & & \\ & 0 & & \ddots & & 0 & \\ & & & & \lambda_k & & \\ & & & & & \ddots & \\ & & & & & & \lambda_k \end{pmatrix}.$$

6 Differentialrechnung mehrerer Veränderlicher

6.1 Euklidische Norm und Stetigkeit

Viele Größen in der Chemie, Physik,..... hängen von mehreren Variablen ab.

Beispiele 6.1. (1) Zustandsgleichung idealer Gase:

$$p = R\frac{T}{V} \qquad (R \text{ Gaskonstante})$$

(2) Wellenfunktion eines freien Elektrons: $\psi: \mathbb{R} \times \mathbb{R}^3 \to \mathbb{C}$,

$$\psi(t,x) = A \exp(i(\langle k, x\rangle - \omega t)) \qquad (k \in \mathbb{R}^3, \omega \in \mathbb{R})$$

(3) Arithmetisches Mittel:

$$f : \mathbb{R}^n \to \mathbb{R},\ f(x_1, \ldots, x_n) = \frac{1}{n}\sum_{i=1}^{n} x_i.$$

Satz 6.2. *Die euklidische Norm*

$$\mathbb{R}^n \to \mathbb{R},\ (x_1, \ldots, x_n) \mapsto \|(x_1, \ldots, x_n)\| = \sqrt{\langle x, x\rangle} = \sqrt{\sum_{i=1}^{n} x_i^2}$$

hat die Eigenschaften

(i) $\|x\| \geq 0$ *und* $\|x\| = 0 \Leftrightarrow x = 0$,

(ii) $\|\lambda x\| = |\lambda|\ \|x\| \qquad (\lambda \in \mathbb{R}, x \in \mathbb{R}^n)$,

(iii) $\|x + y\| \leq \|x\| + \|y\|,\ |\ \|x\| - \|y\|\ | \leq \|x - y\|$ (Dreiecksungleichung),

(iv) $\max_{i=1,\ldots,n} |x_i| \leq \|(x_i)\| \leq \sqrt{n} \max_{i=1,\ldots,n} |x_i|$.

Definition 6.3. (a) Eine Folge $(x_k)_{k\geq 0}$ von Vektoren $x_k \in \mathbb{R}^n$ heißt *konvergent* gegen einen Vektor $x \in \mathbb{R}^n$ (abgekürzt als $\lim_{k\to\infty} x_k = x$), falls zu jedem $\epsilon > 0$ ein $k_\epsilon \in \mathbb{N}$ existiert mit $\|x_k - x\| < \epsilon$ für alle $k \geq k_\epsilon$.

(b) Sei $D \subset \mathbb{R}^n$. Eine Funktion $f : D \to \mathbb{R}$ heißt *stetig* in $a \in D, \lim_{k\to\infty} f(x_k) = f(a)$ ist für jede Folge $(x_k)_{k\geq 0}$ in D mit $\lim_{k\to\infty} x_k = a$.

Für $x_k = (x_{k1}, \ldots, x_{kn}), x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ gilt $\lim_{k\to\infty} x_k = x$ in $\mathbb{R}^n$ wegen Satz 6.2(iv) genau dann, wenn für jedes $i = 1, \ldots, n$

$$\lim_{k\to\infty} x_{ki} = x_i \text{ in } \mathbb{R} \text{ ist.}$$

Merkregel: Konvergenz von Folgen in $\mathbb{R}^n$ bedeutet komponentenweise Konvergenz in $\mathbb{R}$.

Aus der Definition der Stetigkeit und den Grenzwertsätzen für Folgen in $\mathbb{R}$ folgt:

Satz 6.4. *(a) Sind $f, g : D \to \mathbb{R}$ stetig in $a \in D$, so sind auch*

$$f + g,\ f \cdot g,\ \alpha f\ (\alpha \in \mathbb{R}), \frac{f}{g}\ (\textit{falls } g \textit{ keine Nullstelle hat})$$

stetig in a.

(b) Kompositionen stetiger Funktionen sind stetig.

Beispiele 6.5. (1) Die Projektionen

$$\pi_i :\ \mathbb{R}^n \to \mathbb{R},\ (x_1, \ldots, x_n) \mapsto x_i \qquad (i = 1, \ldots, n)$$

sind stetig, da Konvergenz in $\mathbb{R}^n$ komponentenweise Konvergenz bedeutet.

(2) Nach Satz 6.4(a) sind alle Polynomfunktionen $p :\ \mathbb{R}^n \to, \mathbb{R}$,d

$$p(x_1, \ldots, x_n) = \sum_{i_1=0}^{N} \cdots \sum_{i_n=0}^{N} a_{i_1\ldots i_n} x_1^{i_1} \cdots x_n^{i_n}$$

stetig.

(3) Nach Satz 6.4(b) sind

$$f : \mathbb{R}^n \to \mathbb{R},\ x \mapsto \|x\| = \sqrt{x_1^2 + \ldots + x_n^2} \qquad (\text{oder } x \mapsto e^{x_1^2+\ldots+x_n^2})$$

stetig.

6.2 Partielle Differenzierbarkeit

Man nennt eine Funktion $f : \mathbb{R}^n \to \mathbb{R}$ *partiell differenzierbar* nach x_i, wenn sie als Funktion der i-ten Variablen bei festgehaltenen übrigen Variablen differenzierbar ist.

Definition 6.6. Seien $U \subset \mathbb{R}^n$ und $a = (a_1, \ldots, a_n) \in U$.

(a) $f : U \to \mathbb{R}$ heißt in a *partiell differenzierbar* nach x_i, falls

$$f_i(x) = f(a_1, \ldots, a_{i-1}, x, a_{i+1}, \ldots, a_n)$$

differenzierbar in $x = a_i$ ist. In diesem Fall nennt man

$$f_{x_i}(a) \text{ (oder } \frac{\partial f}{\partial x_i}(a)) := f_i'(a_i)$$

die *i-te partielle Ableitung* von f im Punkt a.

(b) Ist $f : U \to \mathbb{R}$ partiell differenzierbar nach x_i auf ganz U, so heißt die Funktion

$$f_{x_i} \text{ (oder } \frac{\partial f}{\partial x_i}) : U \to \mathbb{R},\ a \mapsto f_{x_i}(a)$$

die i-te *partielle Ableitung* von f.

(c) $f : U \to \mathbb{R}$ heißt *partiell differenzierbar*, falls f nach jedem x_i ($i = 1, \ldots, n$) auf ganz U partiell differenzierbar ist.

Die üblichen 1-dimensionalen Rechenregeln (Summen-, Produkt-, Quotientenregel) übertragen sich direkt auf partielle Ableitungen.

Beispiele 6.7. (a) $r : \mathbb{R}^n \to \mathbb{R}, r(x) = \|x\| = \sqrt{x_1^2 + \ldots + x_n^2}$ ist partiell differenzierbar auf $\mathbb{R}^n \setminus \{0\}$ mit

$$\begin{aligned} \frac{\partial r}{\partial x_i}(x) &= \frac{d}{dt}(x_1^2 + \ldots + x_{i-1}^2 + t^2 + x_{i+1}^2 + \ldots + x_n^2)^{1/2}|_{t=x_i} \\ &= \frac{1}{2}\frac{2x_i}{r(x)} = \frac{x_i}{\|x\|}. \end{aligned}$$

(b) Ist $f : (0, \infty) \to \mathbb{R}$ differenzierbar, so ist

$$f \circ r : \mathbb{R}^n \setminus \{0\} \to \mathbb{R},\ x \mapsto f(r(x)) = f(\|x\|)$$

partiell differenzierbar. Nach der 1-dimensionalen Kettenregel (Satz 9.5 in [E]) ist

$$\begin{aligned} \frac{\partial (f \circ r)}{\partial x_i}(x) &= (f \circ r)_i'(x_i) = (f \circ r_i)'(x_i) = f'(r_i(x_i))r_i'(x_i) \\ &= f'(r(x))\frac{\partial r}{\partial x_i}(x) \overset{(a)}{=} f'(\|x\|)\frac{x_i}{\|x\|}. \end{aligned}$$

(c) Konkretes Beispiel: Wähle $f(x) = \frac{1}{x}$ in (b). Dann ist

$$\frac{\partial}{\partial x_i}(\frac{1}{\|x\|}) = -\frac{1}{\|x\|^2}\frac{x_i}{\|x\|} = -\frac{x_i}{\|x\|^3}.$$

Wenn $f : U \to \mathbb{R}$ partiell differenzierbar ist, nennt man die $\mathbb{R}^n$-wertige Funktion

$$\mathrm{grad} f \text{ (oder } \nabla f \text{ gelesen „ Nabla } f\text{“)} : U \to \mathbb{R}^n, \ \mathrm{grad} f(x) = (\frac{\partial f}{\partial x_1}(x), \dots, \frac{\partial f}{\partial x_n}(x))$$

den *Gradienten* von f.

Beispiel 6.8. Für $f(x) = \frac{1}{\|x\|}$ ist

$$\mathrm{grad} f(x) = (\frac{\partial f}{\partial x_1}(x), \dots, \frac{\partial f}{\partial x_n}(x)) = -\frac{x}{\|x\|^3} \in \mathbb{R}^n.$$

6.3 Totale Differenzierbarkeit

Sei $U \subset \mathbb{R}^n$ und sei $f : U \to \mathbb{R}$ eine Funktion.

Satz 6.9. *Ist $f : U \to \mathbb{R}$ partiell differenzierbar auf U und sind $\frac{\partial f}{\partial x_i} : U \to \mathbb{R}$ $(i = 1, \dots, n)$ stetig in einem Punkt $x^0 \in U$, so gibt es in x^0 stetige Funktionen $\Delta_1, \dots, \Delta_n : U \to \mathbb{R}$ mit*

$$f(x) - f(x^0) = \sum_{\nu=1}^{n} (x_\nu - x_\nu^0)\Delta_\nu(x) \textit{ für alle } x \in U.$$

Idee für $n = 2$: Nach dem MWS aus der Vorlesung MfN I (Satz 10.4 in [E]) ist

$$\begin{aligned} f(x_1, x_2) - f(x_1^0, x_2^0) &= (f(x_1, x_2) - f(x_1^0, x_2)) + (f(x_1^0, x_2) - f(x_1^0, x_2^0)) \\ &= (x_1 - x_1^0)\frac{\partial f}{\partial x_1}(\xi_{x_1,x_2}, x_2) + (x_2 - x_2^0)\frac{\partial f}{\partial x_2}(x_1^0, \eta_{x_2}) \end{aligned}$$

mit Stellen ξ_{x_1,x_2} zwischen x_1 und x_1^0 sowie η_{x_2} zwischen x_2 und x_2^0. Für $x \to x^0$ gilt

$$\begin{array}{rccccl} \Delta_1(x_1, x_2) := & \frac{\partial f}{\partial x_1}(\xi_{x_1,x_2}, x_2) & \longrightarrow & \frac{\partial f}{\partial x_1}(x^0) = & \Delta_1(x^0) \\ \Delta_2(x_1, x_2) := & \frac{\partial f}{\partial x_2}(x_1^0, \eta_{x_2}) & \longrightarrow & \frac{\partial f}{\partial x_2}(x^0) = & \Delta_2(x^0). \end{array}$$

Definition 6.10. (Totale Differenzierbarkeit)
Eine Funktion $f : U \to \mathbb{R}$ heißt *total differenzierbar* in $x^0 \in U$, falls in x^0 stetige funktionen $\Delta_1, \dots, \Delta_n : U \to \mathbb{R}$ existieren mit

$$f(x) - f(x^0) = \sum_{\nu=1}^{n} (x_\nu - x_\nu^0)\Delta_\nu(x) \text{ für alle } x \in U.$$

Folgerungen 6.11. *(a) Ist f total differenzierbar in x^0, so ist f stetig in x^0 (Vorsicht: Aus der partiellen Differenzierbarkeit kann man nicht auf Stetigkeit schließen).*

(b) Ist f total differenzierbar in $x^0 \in U$, so ist f partiell differenzierbar in x^0 und

$$\frac{\partial f}{\partial x_i}(x^0) = \Delta_i(x^0) \quad (i = 1, \dots, n).$$

(c) Ist $f : U \to \mathbb{R}$ stetig partiell differenzierbar, das heißt ist f partiell differenzierbar und sind alle partiellen Ableitungen $\frac{\partial f}{\partial x_i} : U \to \mathbb{R}$ $(i = 1, \dots, n)$ stetig, so ist f total differenzierbar in jedem Punkt $x^0 \in U$.

Merkregel: Es gelten die folgenden Implikationen:

stetig partielle Differenzierbarkeit $\Rightarrow$ totale Differenzierbarkeit $\Rightarrow$ partielle Differenzierbarkeit.

Die Umkehrungen sind am Allgemeinen falsch.

Satz 6.12. (Kettenregel)
Ist $f : U \to \mathbb{R}$ total differenzierbar in $x^0 \in U$ und ist $\varphi : [a,b] \to \mathbb{R}^n, \varphi(t) = (\varphi_1(t), \dots, \varphi_n(t))$ eine Funktion mit

$$\varphi([a,b]) \subset U \text{ und } \varphi(t_0) = x^0$$

für ein $t_0 \in [a,b]$ so, dass φ differenzierbar ist in t_0 (:$\Leftrightarrow$ alle φ_ν sind differenzierbar in t_0), so ist $f \circ \varphi : [a,b] \to \mathbb{R}$ differenzierbar in t_0 mit

$$(f \circ \varphi)'(t_0) = \sum_{\nu=1}^{n} \frac{\partial f}{\partial x_\nu}(\varphi(t_0))\varphi_\nu'(t_0).$$

Idee: Seien $\Delta_1, \dots, \Delta_n : U \to \mathbb{R}$ in $x^0 = \varphi(t_0)$ stetige Funktionen wie in Definition 6.10. Dann konvergiert der Differenzenquotient

$$\frac{f(\varphi(t)) - f(\varphi(t_0))}{t - t_0} = \sum_{\nu=1}^{n} \frac{\varphi_\nu(t) - \varphi_\nu(t_0)}{t - t_0} \quad \Delta_\nu(\varphi(t))$$

für $t \to t_0$ gegen

$$\sum_{\nu=1}^{n} \varphi_\nu'(t_0) \frac{\partial f}{\partial x_\nu}(\varphi(t_0)).$$

Folgerung 6.13. *Sei $f : U \to \mathbb{R}$ total differenzierbar auf $U \subset \mathbb{R}^n$. Sei $V \subset \mathbb{R}^m$ und sei*

$$\varphi : V \to \mathbb{R}^n, \ t \mapsto (\varphi_1(t), \dots, \varphi_n(t))$$

eine Funktion mit partiell differenzierbaren Koordinatenfunktionen $\varphi_i : V \to \mathbb{R}$ und $\varphi(V) \subset U$. Dann ist $f \circ \varphi : V \to \mathbb{R}$ partiell differenzierbar mit

$$\frac{\partial (f \circ \varphi)}{\partial t_\mu}(t) = \sum_{\nu=1}^{n} \frac{\partial f}{\partial x_\nu}(\varphi(t)) \frac{\partial \varphi_\nu}{\partial t_\mu}(t) \qquad \text{für } t \in V.$$

Beispiel 6.14. (Polarkoordinaten in $\mathbb{R}^2$) Schreibe $(x,y) \in \mathbb{R}^2$ in Polarkoordinaten

$$x = r\cos\varphi, \ y = r\sin\varphi.$$

Ist $f : \mathbb{R}^2 \to \mathbb{R}$ total differenzierbar, so ist

$$F : (0,\infty) \times \mathbb{R} \to \mathbb{R}, \ F(r,\varphi) = f(r\cos\varphi, r\sin\varphi)$$

partiell differenzierbar mit

$$\begin{aligned}
\tfrac{\partial F}{\partial r}(r,\varphi) &= \tfrac{\partial f}{\partial x}(r\cos\varphi, r\sin\varphi)\tfrac{\partial(r\cos\varphi)}{\partial r} &&+ \tfrac{\partial f}{\partial y}(r\cos\varphi, r\sin\varphi)\tfrac{\partial(r\sin\varphi)}{\partial r}\\
&= \tfrac{\partial f}{\partial x}(r\cos\varphi, r\sin\varphi)\cos\varphi &&+ \tfrac{\partial f}{\partial y}(r\cos\varphi, r\sin\varphi)\sin\varphi\\
\tfrac{\partial F}{\partial \varphi}(r,\varphi) &= \tfrac{\partial f}{\partial x}(r\cos\varphi, r\sin\varphi)\tfrac{\partial(r\cos\varphi)}{\partial \varphi} &&+ \tfrac{\partial f}{\partial y}(r\cos\varphi, r\sin\varphi)\tfrac{\partial(r\sin\varphi)}{\partial \varphi}\\
&= -\tfrac{\partial f}{\partial x}(r\cos\varphi, r\sin\varphi)r\sin\varphi &&+ \tfrac{\partial}{\partial y}(r\cos\varphi, r\sin\varphi)r\cos\varphi.
\end{aligned}$$

Satz 6.15. (Mittelwertsatz) *Sei $f : U \to \mathbb{R}$ total differenzierbar. Sind $x, x^0 \in U$ so, dass die ganze Verbindungsstrecke zu U gehört, das heißt mit*

$$\{x^0 + t(x - x^0) |\ 0 \le t \le 1\} \subset U,$$

dann existiert ein $\theta \in [0,1]$ mit

$$f(x) - f(x^0) = \sum_{\nu=1}^{n} \frac{\partial f}{\partial x_\nu}(x^0 + \theta(x - x^0))(x_\nu - x_\nu^0).$$

Idee: Nach der mehrdimensionalen Kettenregel (Satz 6.12) ist $F : [0,1] \to \mathbb{R}, F(t) = f(x^0 + t(x - x^0))$ differenzierbar mit

$$F'(t) = \sum_{\nu=1}^{n} \frac{\partial f}{\partial x_\nu}(x^0 + t(x - x^0))(x_\nu - x_\nu^0) \qquad (t \in [0,1]).$$

Nach dem MWS der 1-dimensionalen Differentialrechnung (Satz 10.4 in [E]) gibt es ein $\theta \in [0,1]$ mit

$$f(x) - f(x^0) = \frac{F(1) - F(0)}{1 - 0} = F'(\theta).$$

6.4 Höhrere partielle Ableitungen

Ist $f : U \to \mathbb{R}$ partiell differenzierbar, so sind die partiellen Ableitungen wieder Funktionen

$$\frac{\partial f}{\partial x_i} : U \to \mathbb{R},\ x \mapsto \frac{\partial f}{\partial x_i}(x) \qquad (i = 1, \ldots, n)$$

und man kann nach deren partiellen Ableitungen fragen. Wenn sie alle existieren, heißt f 2-*mal partiell differenzierbar* und man schreibt

$$\frac{\partial^2 f}{\partial x_j \partial x_i}(\text{oder auch } \ f_{x_i,x_j}) := \frac{\partial}{\partial x_j}(\frac{\partial f}{\partial x_i}).$$

Entsprechend definiert man *partielle Ableitungen höherer Ordnung*, wobei die Ordnung die Anzahl der ausgeführen partiellen Ableitungen bezeichnet.

Beispiel 6.16. Sei $f(x,y) = 3x^2y - 8x\cos y + x$. Dann gilt

$$f_x(x,y) = 6xy - 8\cos y + 1,\ f_y(x,y) = 3x^2 + 8x\sin y$$

und

$$f_{x,x}(x,y) = 6y,\ f_{xy}(x,y) = 6x + 8\sin y,\ f_{yx}(x,y) = 6x + 8\sin y,\ f_{yy}(x,y) = 8x\cos y.$$

Gilt immer $f_{xy}(x,y) = f_{yx}(x,y)$? Nein, aber:

Satz 6.17. (Satz von Schwarz) *Sei $f : U \to \mathbb{R}$ eine Funktion auf $U \subset \mathbb{R}^n$ und seien $i, j \in \{1, \ldots, n\}$. Existieren*

$$f_{x_i,x_j} \text{ und } f_{x_j,x_i}$$

auf U und sind beide Funktionen stetig auf U, so sind sie gleich.

Allgemein gilt: Existieren für $f : U \to \mathbb{R}$ alle partiellen Ableitungen der Ordnung $\leq k$ und sind sie stetig auf U ($:\Leftrightarrow$: *f ist k-mal stetig differenzierbar*), so sind diese partiellen Ableitungen unabhängig von der Reihenfolge.

In diesem Fall schreibt man für $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$ mit $\alpha_1 + \ldots + \alpha_n \leq k$

$$\frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \ldots \partial x_n^{\alpha_n}} = \frac{\partial}{\partial x_1}(\ldots \frac{\partial}{\partial x_1}(\ldots (\frac{\partial}{\partial x_n} \ldots (\frac{\partial}{\partial x_n} f) \ldots) \ldots) \ldots).$$

Weitere übliche Abkürzungen sind:

$$\frac{\partial^{\alpha|} f}{\partial x^\alpha} = \frac{\partial^\alpha f}{\partial x_1^{\alpha_1} \ldots \partial x_n^{\alpha_n}},\ |\alpha| = \alpha_1 + \ldots + \alpha_n,\ \alpha! = \alpha_1! \cdots \alpha_n!.$$

Satz 6.18. (Taylorformel) *Sei $f : U \to \mathbb{R}$ $(k+1)$-mal stetig differenzierbar und seien $x, x^0 \in U$ mit $\{x^0 + t(x - x^0) |\ t \in [0,1]\} \subset U$. Dann gibt es ein $\theta \in [0,1]$ mit*

$$\begin{aligned} f(x) &= \sum_{\substack{\alpha \in \mathbb{N}^n \\ |\alpha| \leq k}} \tfrac{1}{\alpha!} \tfrac{\partial^{|\alpha|} f}{\partial x^\alpha}(x^0)(x - x^0)^\alpha \\ &+ \sum_{\substack{\alpha \in \mathbb{N}^n \\ |\alpha| = k+1}} \tfrac{1}{\alpha!} \tfrac{\partial^{|\alpha|} f}{\partial x^\alpha}(x^0 + \theta(x - x^0))(x - x^0)^\alpha, \end{aligned}$$

wobei $(x - x^0)^\alpha = (x_1 - x_1^0)^{\alpha_1} \cdots (x_n - x_n^0)^{\alpha_n}$ sei.

In der Situation von Satz 6.18 nennt man die erste Summe

$$f(x) = \sum_{\substack{\alpha \in \mathbb{N}^n \\ |\alpha| \leq k}} \frac{1}{\alpha!} \frac{\partial^{|\alpha|} f}{\partial x^\alpha}(x^0)(x - x^0)^\alpha$$

das k-te *Taylorpolynom* von f im Punkt x^0 und die zweite Summe

$$\sum_{\substack{\alpha \in \mathbb{N}^n \\ |\alpha| = k+1}} \frac{1}{\alpha!} \frac{\partial^{|\alpha|} f}{\partial x^\alpha}(x^0 + \theta(x - x^0))(x - x^0)^\alpha$$

das *Restglied* $(k+1)$-*Ordnung* von f bei Entwicklung um den Punkt x^0.

Beispiel 6.19. Im Spezialfall $k = 1, n = 2$ erhält man mit $x_\theta = x^0 + \theta(x - x^0)$

$$\begin{aligned} f(x) &= f(x^0) + \tfrac{\partial f}{\partial x_1}(x^0)(x_1 - x_1^0) + \tfrac{\partial f}{\partial x_2}(x^0)(x_2 - x_2^0) \\ &+ \tfrac{1}{2} \tfrac{\partial^2 f}{\partial x_1^2}(x_\theta)(x_1 - x_1^0)^2 + \tfrac{\partial^2 f}{\partial x_1 \partial x_2}(x_\theta)(x_1 - x_1^0)(x_2 - x_2^0) + \tfrac{1}{2} \tfrac{\partial^2 f}{\partial x_2^2}(x_\theta)(x_2 - x_2^0)^2 \\ &= f(x^0) + \textstyle\sum_{i=1}^2 \tfrac{\partial f}{\partial x_i}(x^0)(x_i - x_i^0) + \tfrac{1}{2} \textstyle\sum_{i,j=1}^2 \tfrac{\partial^2 f}{\partial x_i \partial x_j}(x_\theta)(x_i - x_i^0)(x_j - x_j^0). \end{aligned}$$

Vergisst man das Restglied, so erhält man für beliebiges $n \geq 1$ mit $a = x^0$:

$k = 1: \quad f(x) \approx f(a) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a)(x_i - a_i)$ (*lineare Approximation*)
$k = 2: \quad f(x) \approx f(a) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a)(x_i - a_i) + \frac{1}{2}\sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(a)(x_i - a_i)(x_j - a_j)$ (*quadratische Approximation*)

6.5 Lokale Extrema: Notwendige Bedingung

Für Funktionen $f : [a, b] \to \mathbb{R}$ wurde gezeigt (Satz 10.2 und Satz 10.8 in [E]):

- Ist f differenzierbar in $x_0 \in]a, b[$ und hat f ein lokales Extremum in x_0, so ist $f'(x_0) = 0$.
- Ist f 2-mal differenzierbar auf $[a, b]$ und ist $x_0 \in]a, b[$ mit $f'(x_0) = 0$, dann gilt:

$$\begin{aligned} f''(x_0) > 0 \quad &\Rightarrow \quad f \text{ hat ein lokales Minimum in } x_0 \\ f''(x_0) < 0 \quad &\Rightarrow \quad f \text{ hat ein lokales Maximum in } x_0. \end{aligned}$$

Gibt es ähnliche Ergebnisse für Funktionen mehrerer Veränderlicher?
Für $x_0 \in \mathbb{R}^n$ und $\epsilon > 0$ sei

$$B_\epsilon(x_0) = \{x \in \mathbb{R}^n | \; \|x - x_0\| < \epsilon\}$$

die *Kugel* mit Radius ϵ und Mittelpunkt x_0. Eine Menge $U \subset \mathbb{R}^n$ heißt *offen*, falls

$$\text{für jedes } x_0 \in U \text{ ein } \epsilon > 0 \text{ existiert mit } B_\epsilon(x_0) \subset U.$$

Sei $U \subset \mathbb{R}^n$ im Folgenden eine offene Menge.

Definition 6.20. Sei $f : U \to \mathbb{R}$ eine Funktion und sei $x_0 \in U$. Man sagt:

(i) f hat in x_0 ein *lokales Maximum*, falls ein $\epsilon > 0$ existiert mit

$$f(x) \leq f(x_0) \text{ für alle } x \in B_\epsilon(x_0).$$

Entsprechend definiert man *lokale Minima.*

(ii) f hat ein *lokales Extremum* in x_0, falls f ein lokales Maximum oder Minimum in x_0 besitzt.

(iii) f hat ein *absolutes Maximum* in x_0, falls $f(x) \leq f(x_0)$ für alle $x \in U$ gilt. Entsprechend definiert man *absolute Minima* und *absolute Extrema.*

Satz 6.21. *Sei $f : U \to \mathbb{R}$ partiell differenzierbar. Hat f in x^0 ein lokales Extremum, so ist*

$$\text{grad} f(x^0) = 0$$

Idee: Nach dem entsprechenden Satz aus der 1-dimensionalen Differentialrechnung (Satz 10.2 in [E]) gilt

$$\frac{\partial f}{\partial x_i}(x^0) = \frac{d}{dx} f(x_1^0, \ldots, x_{i-1}^0, x, x_{i+1}^0, \ldots, x_n^0)|_{x = x_i^0} = 0 \;\; \text{für } i = 1, \ldots, n.$$

6.6 Lokale Extrema: Hinreichende Bedingung

Eine symmetrische Matrix $A = (a_{ij}) \in \mathbb{R}^{n,n}$ heißt

- *positiv* (bzw. *negativ*) *definit*, falls

$$\langle Ax, x\rangle > 0 \text{ (bzw. } \langle Ax, x\rangle < 0) \text{ für alle } x \in \mathbb{R}^n \setminus \{0\} \text{ gilt,}$$

- *indefinit*, falls $x, y \in \mathbb{R}^n$ existieren mit

$$\langle Ax, x\rangle < 0 < \langle Ay, y\rangle.$$

Bemerkung 6.22. In der Mathematik zeigt man:
(i) Eine symmetrische Matrix $A \in \mathbb{R}^{n,n}$ ist positiv (bzw. negativ) definit

$$\Leftrightarrow \text{alle Eigenwerte von } A \text{ sind } > 0 \text{ (bzw. } < 0).$$

Sie ist indefinit genau dann, wenn mindestens ein Eigenwert $\lambda > 0$ existiert und mindestens ein Eigenwert $\mu < 0$.
(ii) Eine symmetrische Matrix

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

ist

- positiv definit $\Leftrightarrow a > 0$ und $ad - b^2 > 0$
- negativ definit $\Leftrightarrow a < 0$ und $ad - b^2 > 0$
- indefinit $\Leftrightarrow ad - b^2 < 0$.

Definition 6.23. Sei $f : U \to \mathbb{R}$ 2-mal stetig differenzierbar. Dann heißt

$$\mathrm{Hess}f(x) = (\frac{\partial^2 f}{\partial x_i \partial x_j}(x))_{i,j=1,\ldots,n} \in \mathbb{R}^{n,n}$$

die *Hesse-Matrix* von f im Punkt $x \in U$.

Beachte: Nach dem Satz von Schwarz (Satz 6.17) ist $\mathrm{Hess}f(x) \in \mathbb{R}^{n,n}$ eine symmetrische Matrix.

Analog zum 1-dimensionalen Fall (Satz 10.8 in [E]) gilt:

Satz 6.24. *Sei $f : U \to \mathbb{R}$ 2-mal stetig differenzierbar auf einer offenen Menge $U \subset \mathbb{R}^n$ und sei $x^0 \in U$ mit $\mathrm{grad}f(x^0) = 0$. Dann gilt:*

(i) $\mathrm{Hess}f(x^0)$ positiv (negativ) definit $\Rightarrow f$ hat in x^0 ein lokales Minimum (Maximum).

(ii) Hess$f(x^0)$ *indefinit* $\Rightarrow f$ *hat kein lokales Extremum in* x^0.

Für $n = 1$ ist grad$f(x^0) = f'(x^0)$, (Hess$f$)$(x^0) = f''(x_0)$.

Beispiele 6.25. (i) Sei $f(x,y) = x^2 - 2x + y^2 + 1$. Dann gilt

$$\text{grad}f(x,y) = (f_x(x,y), f_y(x,y) = (2x-2, 2y) = 0 \Leftrightarrow (x,y) = (1,0).$$

Nur hier kann ein lokales Extremum vorliegen. Da

$$\text{Hess}f(0,1) = \left(\begin{array}{c|c} \frac{\partial^2 f}{\partial x^2}(x,y) & \frac{\partial^2 f}{\partial x \partial y}(x,y) \\ \hline \frac{\partial^2 f}{\partial y \partial x}(x,y) & \frac{\partial^2 f}{\partial y^2}(x,y) \end{array}\right)|_{(x,y)=(1,0)} = \left(\begin{array}{c|c} 2 & 0 \\ \hline 0 & 2 \end{array}\right)$$

positiv definit ist (Bemerkung 6.22), liegt hier ein lokales Minimum vor.

(ii) Sei

$$f(x,y) = c + x^2 - y^2 \quad (c \in \mathbb{R} \text{ fest}).$$

Dann gilt

$$\text{grad}f(x,y) = (2x, -2y) = 0 \Leftrightarrow (x,y) = (0,0).$$

Nur hier kann ein lokales Extremum vorliegen. Da

$$\text{Hess}f(0,0) = \left(\begin{array}{c|c} 2 & 0 \\ \hline 0 & -2 \end{array}\right)$$

indefinit ist (Bemerkung 6.22), liegt hier kein lokales Extremum vor. Also hat f überhaupt kein lokales Extremum.

7 Kurvenintegrale

7.1 Länge von Kurven

Eine *stetig differenzierbare Kurve* in $\mathbb{R}^n$ ist eine Abbildung $\varphi : I = [a,b] \to \mathbb{R}^n, t \mapsto \varphi(t) = (\varphi_1(t), \ldots, \varphi_n(t))$ mit stetig differenzierbaren Koordinatenfunktionen $\varphi_i : [a,b] \to \mathbb{R}$. Für $t \in I$ heißt

$$\varphi'(t) = (\varphi_1'(t), \ldots, \varphi_n'(t))$$

der *Tangentialvektor* von φ in t.

Hilfssatz (Kapitel 4 in [Fo]) Ist $\varphi : I \to \mathbb{R}^n$ stetig differenzierbar, so gibt es zu jedem $\epsilon > 0$ ein $\delta > 0$ mit

$$\|\frac{\varphi(t) - \varphi(s)}{t - s} - \varphi'(t)\| < \epsilon$$

für alle $s, t \in I$ mit $0 < |s-t| < \delta$.

Wir benutzen dieses Resultat, um die Bogenlänge von Kurven definieren. Ist $T = (t_j)_{j=0}^n$ mit $t_j = t_j^{(n)} = a + j\frac{b-a}{n}$ und ist n so groß, dass $\frac{b-a}{n} < \delta$ ist, so gilt (nach Satz 6.2(iii))

$$\left|\sum_{j=1}^n \|\varphi(t_j) - \varphi(t_{j-1})\| - \sum_{j=1}^n \|\varphi'(t_j)\|(t_j - t_{j-1})\right|$$

$$\leq \sum_{j=1}^n \|\frac{\varphi(t_j) - \varphi(t_{j-1})}{t_j - t_{j-1}} - \varphi'(t_j)\|(t_j - t_{j-1}) \leq \epsilon(b-a).$$

Für $n \to \infty$ konvergiert nach Satz 11.1 in [E]

$$\sum_{j=1}^n \|\varphi'(t_j^{(n)})\|(t_j^{(n)} - t_{j-1}^{(n)}) \xrightarrow{n} \int_a^b \|\varphi'(t)\| dt.$$

Also konvergieren auch die Längen der Polygonzüge (siehe Abbildung 3)

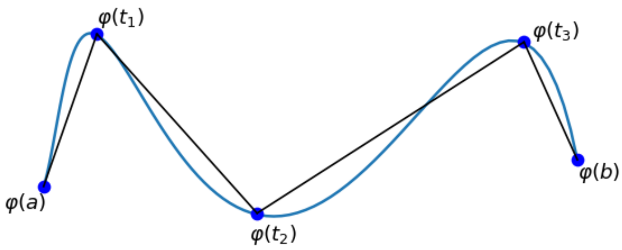


Abbildung 3: Länge eines Polygonzuges

$$\sum_{j=1}^n \|\varphi(t_j^{(n)}) - \varphi(t_{j-1}^{(n)})\| \xrightarrow{n} \int_a^b \|\varphi'(t)\| dt.$$

Definition 7.1. Für eine stetig differenzierbare Kurve $\varphi : [a,b] \to \mathbb{R}^n$ heißt

$$L(\varphi) = \int_a^b \|\varphi'(t)\| dt$$

die *Länge* oder *Bogenlänge* von φ.

7.2 Arbeit als Kurvenintegral

Gegeben seien ein stetiges Kraftfeld $f : \mathbb{R}^3 \to \mathbb{R}^3$ und eine stetig differenzierbare Kurve $\varphi : I = [a,b] \to \mathbb{R}^3$. Gesucht ist die Größe $A(f, \varphi)$ der Arbeit, die zu verrichtet wird, wenn ein Massenpunkt im Kraftfeld f

längs der Kurve φ bewegt wird.

Physikalische Definition: Die geleistete Arbeit in einem konstanten Kraftfeld bei Verschiebung längs einer Geraden ist definitionsgemäß gleich der Komponente der Kraft in Wegrichtung multipliziert mit der Weglänge.

Für ein beliebiges stetiges Kraftfeld $f : \mathbb{R}^3 \to \mathbb{R}^3$ approximieren wir die Kurve φ durch einen Polygonzug $P = P(T)$ gegeben durch eine Teilung $T = (t_j)_{j=0}^n$ des Parameterintervalles $[a, b]$

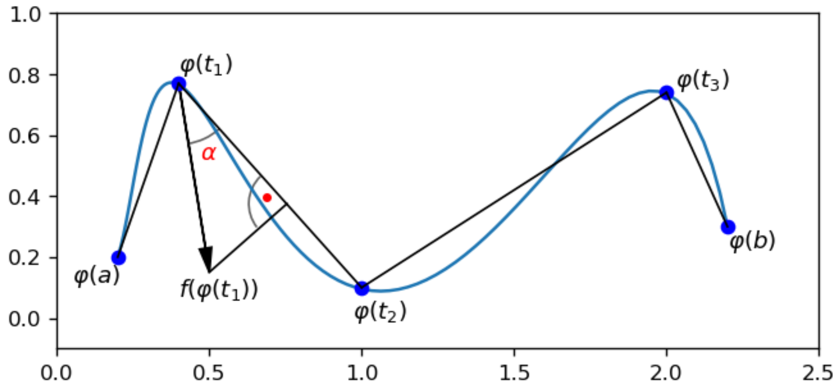


Abbildung 4: Arbeit längs eines Polygonzuges

und ersetzen f auf der Strecke von $\varphi(t_{j-1})$ nach $\varphi(t_j)$ durch $f(\varphi(t_{j-1}))$. Die längs der Verbindungsstrecke von $\varphi(t_{j-1})$ nach $\varphi(t_j)$ bei konstantem Kraftfeld $f \equiv f(\varphi(t_{j-1}))$ geleistete Arbeit ist gleich der

$$\begin{aligned}
&\text{Kraft in Wegrichtung im Punkt } f(\varphi(t_{j-1})) \times \|\varphi(t_j) - \varphi(t_{j-1})\| \\
&= \cos(\alpha)\|f(\varphi(t_{j-1}))\| \, \|\varphi(t_j) - \varphi(t_{j-1})\| \\
&= (\text{Abschnitt 4.2})\langle f(\varphi(t_{j-1})), \ \varphi(t_j) - \varphi(t_{j-1})\rangle.
\end{aligned}$$

Sei $\epsilon > 0$ und sei $M = \max\{\|f \circ \varphi(t)\| \mid t \in [a, b]\}$. Wählt man $\delta > 0$ zu ϵ genau wie im Hilfssatz aus Abschnitt 7.1, so folgt für jede Teilung $T = (t_j)_{j=0}^n$ von $[a, b]$ mit $t_j = t_j^{(n)} = a + j\frac{b-a}{n}$ und $\frac{b-a}{n} < \delta$

$$\begin{aligned}
&\left|\sum_{j=1}^{n}\langle f(\varphi(t_{j-1})), \varphi(t_j) - \varphi(t_{j-1})\rangle - \sum_{j=1}^{n}\langle f(\varphi(t_{j-1})), \varphi'(t_{j-1})\rangle(t_j - t_{j-1})\right| \\
&\le \sum_{j=1}^{n}\left|\langle f(\varphi(t_{j-1})), \frac{\varphi(t_j) - \varphi(t_{j-1})}{t_j - t_{j-1}} - \varphi'(t_{j-1})\rangle\right|(t_j - t_{j-1}) \\
&\le \sum_{j=1}^{n} M\epsilon(t_j - t_{j-1}) = M\epsilon(b - a).
\end{aligned}$$

Hierbei haben wir im letzten Schritt die Cauchy-Schwarzsche Ungleichung aus Abschnitt 4.2 benutzt. Für $n \to \infty$ konvergiert wieder nach Satz 11.1 in [E]

$$\sum_{j=1}^{n} \langle f(\varphi(t_{j-1})), \varphi'(t_{j-1})\rangle (t_j - t_{j-1}) \xrightarrow{n} \int_a^b \langle f(\varphi(t)), \varphi'(t)\rangle dt.$$

Also konvergiert auch die längs des Polygonzuges geleistete Arbeit

$$\sum_{j=1}^{n} \langle f(\varphi(t_{j-1})), \varphi(t_j) - \varphi(t_{j-1})\rangle \xrightarrow{n} \int_a^b \langle f(\varphi(t)), \varphi'(t)\rangle dt.$$

In der Physik definiert man daher

$$A(f, \varphi) := \int_a^b \langle f(\varphi(t)), \varphi'(t)\rangle dt$$

als die Arbeit, die bei Verschiebung eines Massenpunktes längs des Weges φ im Kraftfeld $f : \mathbb{R}^3 \to \mathbb{R}^3$ geleistet wird.

Ist $V : \mathbb{R}^3 \to \mathbb{R}$ stetig differenzierbar mit

$$\operatorname{grad} V = f \text{ (das heißt mit } \frac{\partial V}{\partial x_i} = f_i \text{ für } i = 1, 2, 3),$$

so folgt mit der Kettenregel (Satz 6.12) und dem Hauptsatz der Differential- und Integralrechnung (Satz 11.7 in [E])

$$\begin{aligned} A(f, \varphi) &= \int_a^b \sum_{i=1}^{3} \frac{\partial V}{\partial x_i}(\varphi(t))\varphi_i'(t) dt \\ &= \int_a^b (V \circ \varphi)'(t) dt = V(\varphi(b)) - V(\varphi(a)). \end{aligned}$$

Gibt es eine solche Abbildung V, so ist die auf einem Weg von $\varphi(a)$ nach $\varphi(b)$ geleistete Arbeit $A(f, \varphi)$ unabhängig von der Wahl des Weges φ.

Definition 7.2. Sei $U \subset \mathbb{R}^n$ offen.

(a) Ein *Vektorfeld* auf U ist eine Abbildung $f : U \to \mathbb{R}^n$.

(b) Sei $f = (f_1, \dots, f_n) : U \to \mathbb{R}^n$ ein stetiges Vektorfeld. Eine stetig differenzierbare Abbildung $V : U \to \mathbb{R}$ mit

$$\operatorname{grad} V = f \quad (\Leftrightarrow \frac{\partial V}{\partial x_i} = f_i \text{ für } i = 1, \dots, n)$$

heißt *Potential* (oder *Stammfunktion* von f).

(c) Sei $f : U \to \mathbb{R}^n$ ein stetiges Vektorfeld und $\varphi : [a,b] \to U$ stetig differenzierbar. Dann heißt

$$\int_\varphi f \cdot dx \quad \left(\text{oder } \int_\varphi \sum_{\nu=1}^n f_\nu dx_\nu\right) := \int_a^b \langle f(\varphi(t)), \varphi'(t)\rangle dt$$

das *Kurvenintegral* von f längs φ.

Bemerkung 7.3. Sei $f : U \to \mathbb{R}^n$ ein stetiges Vektorfeld.

(i) Hat f ein Potential $V : U \to \mathbb{R}$, so gilt:

(a) Für jede stetig differenzierbare Kurve $\varphi : [a,b] \to U$ gilt

$$\int_\varphi f \cdot dx = V(\varphi(b)) - V(\varphi(a)).$$

(b) Ist φ (wie in (a)) zusätzlich *geschlossen*, das heißt ist $\varphi(b) = \varphi(a)$, so ist $\int_\varphi f \cdot dx = 0$. Für Kurvenintegrale längs geschlossener Kurven φ schreibt man auch $\oint_\varphi f \cdot dx$ statt $\int_\varphi f \cdot dx$.

(ii) Lassen sich je zwei Punkte von U durch eine stetig differenzierbare Kurve in U verbinden und sind $V_1, V_2 : U \to \mathbb{R}$ zwei Potentiale von f auf U, so ist die Differenz $V_1 - V_2 : U \to \mathbb{R}$ konstant auf U.

Idee zu (ii): Ist $\varphi : [a,b] \to \mathbb{R}^n$ eine stetig differenzierbare Kurve mit $\varphi([a,b]) \subset U$, so gilt

$$V_1(\varphi(b)) - V_1(\varphi(a)) = \int_\varphi f \cdot dx = V_2(\varphi(b)) - V_2(\varphi(a)).$$

Also ist

$$(V_1 - V_2)(\varphi(b)) = (V_1 - V_2)(\varphi(a)).$$

Beispiel 7.4. Die Funktion $f : \mathbb{R}^2 \to \mathbb{R}^2, f(x,y) = (-y, x)$ hat kein Potential, denn für die geschlossene Kurve $\varphi : [0, 2\pi] \to \mathbb{R}^2, \varphi(t) = (r \cos t, r \sin t)$ ist

$$\begin{aligned} \int_\varphi f \cdot dx &= \int_0^{2\pi} \langle f(\varphi(t)), \varphi'(t)\rangle dt = \int_0^{2\pi} \langle \begin{pmatrix} -r \sin t \\ r \cos t \end{pmatrix}, \begin{pmatrix} -r \sin t \\ r \cos t \end{pmatrix} \rangle dt \\ &= \int_0^{2\pi} r^2 dt = 2\pi r^2 \neq 0 \text{ für } r \neq 0. \end{aligned}$$

Eine notwendige Bedingung für die Existenz eines Potentials liefert:

Satz 7.5. *Ist $f : U \to \mathbb{R}^n, x \mapsto (f_1(x), \ldots, f_n(x))$ ein stetig differenzierbares Vektorfeld (das heißt alle $f_i : U \to \mathbb{R}$ seien stetig differenzierbar) und hat f ein Potential $V : U \to \mathbb{R}$, so gilt*

$$\frac{\partial f_j}{\partial x_i} = \frac{\partial f_i}{\partial x_j} \text{ für } i,j = 1, \ldots, n.$$

Ein stetig differenzierbares Vektorfeld, das diese Gleichungen erfüllt, nennt man *geschlossen*.

Idee: Mit dem Satz von Schwarz (Satz 6.17) folgt

$$\frac{\partial f_j}{\partial x_i} = \frac{\partial}{\partial x_i}\frac{\partial}{\partial x_j}V = \frac{\partial}{\partial x_j}\frac{\partial}{\partial x_i}V = \frac{\partial f_i}{\partial x_j}$$

für $i, j = 1, \ldots, n$.

Nützliche notwendige und hinreichende Bedingungen für die Existenz von Potentialen liefert der folgende Satz (§18 in [Fo3]).

Satz 7.6. *Seien $U \subset \mathbb{R}^n$ offen und $f : U \to \mathbb{R}^n$ ein stetiges Vektorfeld.*

(a) Dann hat f ein Potential genau dann, wenn $\oint_\varphi f \cdot dx = 0$ ist für jede stückweise stetig differenzierbare geschlossene Kurve φ in U oder äquivalent, wenn für je zwei stückweise stetig differenzierbare Kurven $\varphi_i : [a_i, b_i] \to U$ mit denselben Anfangspunkten $\varphi_1(a_1) = \varphi_2(a_2)$ und denselben Endpunkten $\varphi_1(b_1) = \varphi_2(b_2)$ gilt $\int_{\varphi_1} f \cdot dx = \int_{\varphi_2} f \cdot dx$.

(b) Ist f stetig differenzierbar und ist $U \subset \mathbb{R}^n$ konvex (U heißt konvex, wenn für je zwei Punkte $x, y \in U$ und jedes $t \in [0, 1]$ auch $x + t(y - x) \in U$ gehört), so gilt

$$f \text{ hat ein Potential } \Leftrightarrow \tfrac{\partial f_j}{\partial x_i} = \tfrac{\partial f_i}{\partial x_j} \text{ für } i, j = 1, \ldots, n.$$

Eine stückweise stetig differenzierbare Kurve in $\mathbb{R}^n$ ist eine stetige Abbildung

$$\varphi : [a, b] \to \mathbb{R}^n$$

derart, dass $a = t_0 < \ldots < t_r = b$ existieren so, dass $\varphi|_{[t_{j-1}, t_j]}$ stetig differenzierbar ist für $j = 1, \ldots, r$.

Definition 7.7. Ein stetiges Vektorfeld $f : U \to \mathbb{R}^n$ auf einer offenen Menge $U \subset \mathbb{R}^n$ heißt *konservativ*, falls $\oint_\varphi f \cdot dx = 0$ ist für jede stückweise stetig differenzierbare geschlossene Kurve φ in U.

7.3 Differentialformen

Ein Vektorfeld $f = (f_1, \ldots, f_n) : U \to \mathbb{R}^n$ über einer offenen Menge $U \subset \mathbb{R}^n$ schreibt man oft auch in der Form

$$f_1 dx_1 + \ldots + f_n dx_n$$

und nennt dies eine *Differentialform* (vom Grad 1) über U. Die Differentialform $f_1 dx_1 + \ldots + f_n dx_n$ heißt

- *stetig* oder *stetig differenzierbar*
- *geschlossen*,

wenn das Vektorfeld $f = (f_1, \ldots, f_n) : U \to \mathbb{R}^n$ diese Eigenschaft hat. Man nennt die Differentialform $f_1 dx_1 + \ldots + f_n dx_n$

- *exakt*,

wenn $f = (f_1, \ldots, f_n) : U \to \mathbb{R}^n$ ein Potential $V : U \to \mathbb{R}$ hat. In diesem Fall schreibt man

$$f_1 dx_1 + \ldots + f_n dx_n = dV$$

und nennt dV das *totale Differential* von V sowie V eine *Stammfunktion* der Differentialform $f_1 dx_1 + \ldots + f_n dx_n$. Ist $f_1 dx_1 + \ldots + f_n dx_n$ eine stetige Differentialform auf U und ist $\varphi : [a, b] \to U$ eine stetig differenzierbare Kurve, so heißt

$$\int_\varphi f_1 dx_1 + \ldots + f_n dx_n = \int_\varphi f \cdot dx$$

das Integral *der Differentialform* $f_1 dx_1 + \ldots + f_n dx_n$ längs der Kurve φ.

Beispiel 7.8. Ist $f : \mathbb{R}^3 \to \mathbb{R}^3$ stetig und ist $V : \mathbb{R}^3 \to \mathbb{R}$ eine Stammfunktion von f, so gilt nach dem Hauptsatz der Differential- und Integralrechnung (Satz 11.7 in [E])

$$\begin{aligned} \int_0^{x_1} f_1(t, x_2, x_3) dt &= V(x_1, x_2, x_3) + \alpha_{x_2, x_3}, \\ \int_0^{x_2} f_2(x_1, t, x_3) dt &= V(x_1, x_2, x_3) + \beta_{x_1, x_3}, \\ \int_0^{x_3} f_3(x_1, x_2, t) dt &= V(x_1, x_2, x_3) + \gamma_{x_1, x_2}, \end{aligned}$$

wobei $\alpha_{x_2,x_3}, \beta_{x_1,x_3}$ und γ_{x_1,x_2} nur von den angegebenen Variablen abhängen. Oft genügt dies, um Stammfunktionen zu raten. So hat etwa die Funktion

$$f(x_1, x_2, x_3) = (x_2 x_3, x_3(x_1 - 2x_2), x_2(x_1 - x_2))$$

eine Stammfunktion $V : \mathbb{R}^3 \to \mathbb{R}$ nach Satz 7.6(b), denn

$$\frac{\partial f_i}{\partial x_j} = \frac{\partial f_j}{\partial x_i} \quad \text{für } i, j = 1, 2, 3.$$

Zur Bestimmung einer Stammfunktion V beachte man:

$$\begin{aligned} \tfrac{\partial V}{\partial x_1}(x_1, x_2, x_3) &= x_2 x_3 & &\Leftrightarrow V(x_1, x_2, x_3) = x_1 x_2 x_3 + \alpha_{x_2, x_3}, \\ \tfrac{\partial V}{\partial x_2}(x_1, x_2, x_3) &= x_3 x_1 - 2x_3 x_2 & &\Leftrightarrow V(x_1, x_2, x_3) = x_2 x_3 x_1 - x_3 x_2^2 + \beta_{x_1, x_3}, \\ \tfrac{\partial V}{\partial x_3}(x_1, x_2, x_3) &= x_2 x_1 - x_2^2 & &\Leftrightarrow V(x_1, x_2, x_3) = x_3 x_2 x_1 - x_3 x_2^2 + \gamma_{x_1, \gamma_2}. \end{aligned}$$

Also ist $V : \mathbb{R}^3 \to \mathbb{R}$,

$$V(x_1, x_2, x_3) = x_1 x_2 x_3 - x_2^2 x_3$$

eine Stammfunktion von f.

8 Bereichsintegrale

8.1 Iterierte Riemannintegrale

Sei $Q \subset \mathbb{R}^2$ ein abgeschlossener Quader, d. h.

$$Q = [a_1, b_1] \times [a_2, b_2] \qquad (a_1 < b_1, a_2 < b_2)$$

und sei $f : Q \to \mathbb{R}$ stetig. Man kann zeigen ([Fo], Satz 9.1), dass

$$[a_1, b_1] \to \mathbb{R}, \qquad x_1 \mapsto \int_{a_2}^{b_2} f(x_1, x_2)dx_2,$$

$$[a_2, b_2] \to \mathbb{R}, \qquad x_2 \mapsto \int_{a_1}^{b_1} f(x_1, x_2)dx_1$$

stetig sind und dass

$$\int_{a_1}^{b_1} \left(\int_{a_2}^{b_2} f(x_1, x_2)dx_2 \right) dx_1 = \int_{a_2}^{b_2} \left(\int_{a_1}^{b_1} f(x_1, x_2)dx_1 \right) dx_2.$$

Ein entsprechender Satz gilt für stetige Funktionen

$$f : Q = [a_1, b_1] \times \ldots \times [a_n, b_n] \to \mathbb{R}.$$

Satz 8.1 (und Definition)**.** *Das Integral*

$$\int_Q f dx_1 \ldots dx_n := \int_{a_1}^{b_1} \left(\ldots \int_{a_{n-1}}^{b_{n-1}} \left(\int_{a_n}^{b_n} f(x_1, \ldots, x_n)dx_n \right) dx_{n-1} \ldots \right) dx_1$$

ist wohldefiniert und unabhängig von der Integrationsreihenfolge. Man nennt es das Volumen- *oder* Bereichsintegral *von f über Q.*

8.2 Bereichsintegrale

Eine Menge $M \subset \mathbb{R}^n$ heißt beschränkt, wenn ein $R > 0$ existiert mit

$$\|x\| \leq R \text{ für alle } x \in M.$$

Man kann allgemeiner für jede beschränkte stetige Funktion $f : M \to \mathbb{R}$ sinnvoll ein Integral (*Lebesgue-Integral*)

$$\int_M f dx_1 \ldots dx_n \text{ (oder } \int_M f dx) \in \mathbb{R}$$

definieren, falls $M \subset \mathbb{R}^n$ beschränkt ist und

- $M \subset \mathbb{R}^n$ offen ist (Abschnitt 6.5) oder
- $M = \mathbb{R}^n \setminus U$ das Komplement einer offenen Menge $U \subset \mathbb{R}^n$ ist oder
- M endliche Vereinigung oder endlicher Durchschnitt (abzählbar unendlich genügt auch) von Mengen dieses Typs ist.

Wir nennen eine solche Menge *zulässig*.
Für $n = 2$ ist

$$\int\limits_M f dx$$

die Größe des Volumens zwischen der Grundfläche M und dem Graphen von f.

Eine Menge $M \subset \mathbb{R}^n$ heißt *abgeschlossen*, falls $\mathbb{R}^n \setminus M$ offen ist. Abgeschlossene beschränkte Mengen sind Beispiele zulässiger Mengen. Man kann zeigen:

Bemerkung 8.2. Ist $M \subset \mathbb{R}^n$ abgeschlossen und beschränkt und ist $f : M \to \mathbb{R}$ stetig, so ist

- f beschränkt und
- es gibt $x_0, x_1 \in M$ mit $f(x_0) = \min_{x \in M} f(x)$ und $f(y_0) = \max_{x \in M} f(x)$.

Beispiel 8.3. Seien $\varphi, \psi : [a, b] \to \mathbb{R}$ stetig mit $\varphi \leq \psi$ und

$$M = \{(x, y) \in \mathbb{R}^2 | \quad a \leq x \leq b \text{ und } \varphi(x) \leq y \leq \psi(x)\}.$$

Man kann zeigen: M ist zulässig (sogar abgeschlossen und beschränkt) und für jede stetige Funktion $f : M \to \mathbb{R}$ gilt

$$\int\limits_M f dxdy = \int\limits_a^b \left(\int\limits_{\varphi(x)}^{\psi(x)} f(x, y) dy \right) dx.$$

Etwa: Für $\varphi(x) \equiv 0, \psi(x) = \sqrt{r^2 - x^2} \qquad (0 \leq x \leq r)$ entspricht M dem Viertelkreis
Sei $f(x, y) = \sqrt{r^2 - x^2 - y^2}$ $((x, y) \in M)$ und V das Volumen der Kugel

$$B = \{(x, y, z) \in \mathbb{R}^3 | \; x^2 + y^2 + z^2 \leq r^2\} \subset \mathbb{R}^3$$

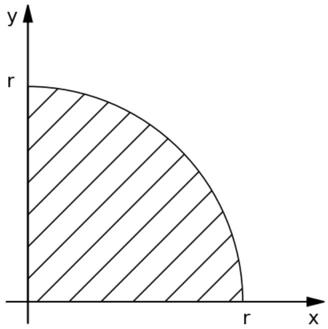


Abbildung 5: Viertelkreis

mit Radius r und Mittelpunkt $0 \in \mathbb{R}^3$. Dann ist

$$\begin{aligned}
\frac{1}{8}V = \int\limits_M f dx dy = &\int\limits_0^r \left(\int\limits_0^{\psi(x)} \sqrt{r^2 - x^2 - y^2}\ dy \right) dx \\
=&(\text{bei festem } x \text{ substituiere } y \text{ durch } y\sqrt{r^2 - x^2}) \\
&\int\limits_0^r (r^2 - x^2) \left(\int\limits_0^1 \sqrt{1 - y^2} dy \right) dx \overset{11.3.(d)}{=} \frac{\pi}{4} \int\limits_0^r (r^2 - x^2) dx \\
=&\frac{\pi}{4}(r^3 - \frac{1}{3}r^3) = \frac{\pi}{6}r^3.
\end{aligned}$$

Also ist $V = \frac{4}{3}\pi r^3$.

8.3 Mehrdimensionale Substitutionsregel (Transformationsformel)

Satz 8.4. *Seien $U, V \subset \mathbb{R}^n$ offen, $f : U \to V$ bijektiv und stetig differenzierbar so, dass*

$$\det\left((\frac{\partial f_i}{\partial x_j}(x))_{1 \leq i,j \leq n} \right) \neq 0 \qquad \text{(Funktionaldeterminante von } f \text{ in } x)$$

für alle $x \in U$ ist. Sei $M \subset U$ eine zulässige Menge und

$$g : f(M) \to \mathbb{R}$$

stetig und beschränkt. Dann ist $f(M) \subset \mathbb{R}^n$ zulässig und

$$\int\limits_{f(M)} g(x) dx = \int\limits_M g(f(x))\ |\det \left(\frac{\partial f_i(x)}{\partial x_j}_{1 \leq i,j \leq n} \right) |\ dx$$

Wenn $N \subset M \subset \mathbb{R}^n$ zwei zulässige Mengen sind, die sich nur um eine niederdimensionale Menge $M \setminus N$ unterscheiden, so kann man zeigen, dass

$$\int\limits_M f dx = \int\limits_N f dx$$

für jede beschränkte stetige Funktion $f : M \to \mathbb{R}$ gilt. Dies gilt zum Beispiel, wenn $M = \overline{B}_r(a) = \{x \in \mathbb{R}^n | \|x-a\| \le r\}$ eine abeschlossene Kugel und $N = B_r(a) = \{x \in \mathbb{R}^n | \|x-a\| < r\}$ die offene Kugel mit Radius $r > 0$ um einen Punkt $a \in \mathbb{R}^n$ ist oder im $\mathbb{R}^2$, wenn $M \setminus N$ enthalten ist in einer Geraden, im Rand eines Rechtecks, eines Kreises.

Ist $M = M_1 \cup \ldots \cup M_r$ Vereinigung von zulässigen Mengen $M_i \subset \mathbb{R}^n$ so, dass je zwei dieser Mengen höchstens eine niederdimensionale Menge gemeinsam haben, dann gilt für jede stetige beschränkte Funktion $f : M \to \mathbb{R}$

$$\int_M f\,dx = \sum_{i=1}^r \int_{M_i} f\,dx.$$

8.4 Anwendungen der Transformationsformel

(a) **Affine Transformationen** Seien $M \subset \mathbb{R}^n$ zulässig, $\alpha \in \mathbb{R} \setminus \{0\}, \beta \in \mathbb{R}^n$ und

$$g : \alpha M + \beta \ (= \{\alpha x + \beta |\ x \in M\}) \to \mathbb{R}$$

stetig und beschränkt. Dann ist

$$\int\limits_{\alpha M + \beta} g(x)dx = \int\limits_M g(\alpha x + \beta)|\alpha|^n dx,$$

denn

$$\det((\frac{\partial(\alpha x_i + \beta_i)}{\partial x_j})_{1 \le i,j \le n}) = \det \begin{pmatrix} \alpha & & 0 \\ & \ddots & \\ 0 & & \alpha \end{pmatrix} = \alpha^n.$$

(b) **Ebene Polarkoordinaten** Für $R > 0$ sei $B_R = \{(x,y) \in \mathbb{R}^2 |\ x^2 + y^2 \le R^2\} \subset \mathbb{R}^2$. Sei $g : B_R \to \mathbb{R}$ stetig und beschränkt. Dann gilt

$$\int\limits_{B_R} g\,dxdy = \int\limits_0^{2\pi} \left(\int\limits_0^R g(r\cos\varphi, r\sin\varphi) r dr \right) d\varphi.$$

Beachte, dass für $f(r,\varphi) = (r\cos\varphi, r\sin\varphi)$ gilt

$$|\det \begin{pmatrix} \frac{\partial f_1}{\partial r}(r,\varphi) & \frac{\partial f_1}{\partial \varphi}(r,\varphi) \\ \frac{\partial f_2}{\partial r}(r,\varphi) & \frac{\partial f_2}{\partial \varphi}(r,\varphi) \end{pmatrix}| = |\det \left(\begin{array}{c|c} \cos\varphi & -r\sin\varphi \\ \hline \sin\varphi & r\cos\varphi \end{array} \right)| = r.$$

Integriert man nicht über ganz B_R, sondern nur über

(i) einen Sektor $S = S(R,\alpha,\beta) = \{(r\cos\varphi, r\sin\varphi) |\ 0 \le r \le R, \alpha \le \varphi \le \beta\}$,

$(0 \le \alpha < \beta \le 2\pi)$, so folgt

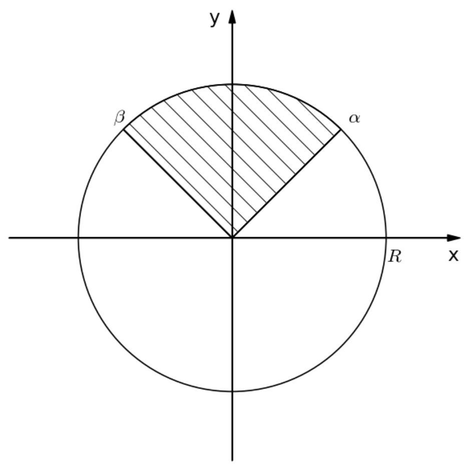


Abbildung 6: Sektor

$$\int\limits_S g\,dxdy = \int\limits_\alpha^\beta \int\limits_0^R g(r\cos\varphi, r\sin\varphi) r dr\ d\varphi.$$

(ii) einen Kreisring $K = K_{R_1,R_2} = \{(x,y) \in \mathbb{R}^2 |\ R_1^2 \leq x^2 + y^2 \leq R_2^2\}$

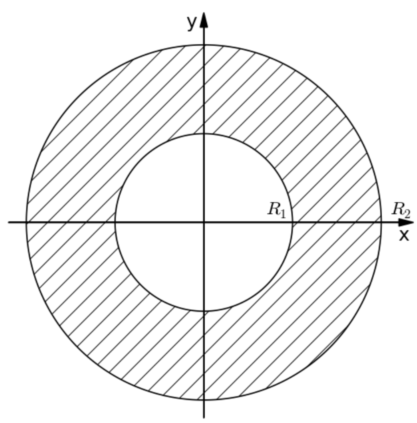


Abbildung 7: Kreisringe

so folgt

$$\int\limits_K g(x,y) dx\, dy = \int\limits_0^{2\pi} \int\limits_{R_1}^{R_2} g(r\cos\varphi, r\sin\varphi) dr\, d\varphi.$$

(c) **Räumliche Polarkoordinaten** Ersetzt man $n = 2$ durch $n = 3$, so gilt entsprechend mit

$$B_R = \{(x, y, z) \in \mathbb{R}^3 |\ x^2 + y^2 + z^2 \leq R^2\},\ \ g : B_R \to \mathbb{R} \text{ stetig beschränkt}$$

für die Transformation in Kugelkoordinaten

$$\int\limits_{B_R} g(x, y, z) dx dy dz = \int\limits_0^{2\pi} \int\limits_0^{\pi} \int\limits_0^{R} g(r \sin\theta \cos\varphi, r \sin\theta \sin\varphi, r \cos\theta) r^2 \sin\theta dr d\theta d\varphi$$

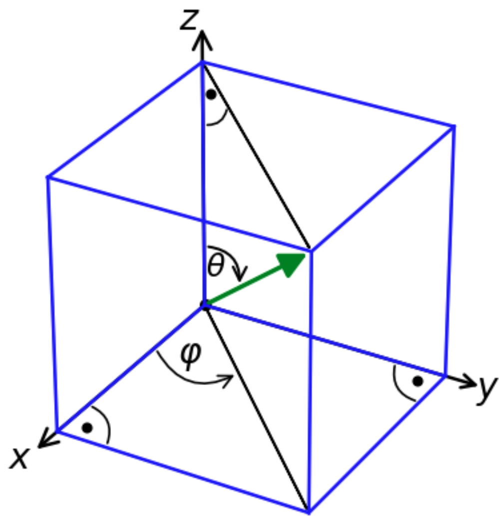


Abbildung 8: Räumliche Polarkoordinaten

8.5 Beispiele

(a) Berechne $I(a) = \int_0^{\infty} \exp(-ax^2) dx$ für $a > 0$. Es ist

$$\int\limits_0^R e^{-ax^2} dx \overset{11.3}{=} (x \to -x) \int\limits_0^{-R} -e^{-a(-x)^2} dx = \int\limits_{-R}^0 e^{-ax^2} dx.$$

Also ist

$$\int\limits_0^R e^{-ax^2} dx = \frac{1}{2} \int\limits_{-R}^R e^{-ax^2} dx$$

und

$$2I(a) = \lim_{R \to \infty} \int\limits_{-R}^R e^{-ax^2} dx = \int\limits_{-\infty}^{+\infty} e^{-ax^2} dx =: J(a).$$

Unter Benutzung von Polarkoordinaten folgt

$$J(a)^2 = \left(\int\limits_{-\infty}^{\infty} e^{-ax^2} dx \right) \left(\int\limits_{-\infty}^{\infty} e^{-ay^2} dy \right) = \int\limits_{-\infty}^{\infty} \left(\int\limits_{-\infty}^{\infty} e^{-a(x^2+y^2)} dy \right) dx$$

$$= \int_0^{2\pi} \int_0^{\infty} e^{-a((r\cos\varphi)^2+(r\sin\varphi)^2)} r\, dr d\varphi$$

$$= \int_0^{2\pi} \int_0^{\infty} e^{-ar^2} r\, dr d\varphi$$

$$= 2\pi \lim_{R\to\infty} \int_0^{R} e^{-ar^2} r\, dr$$

$$= 2\pi \lim_{R\to\infty} \frac{1}{-2a} e^{-ar^2}|_0^R = \frac{\pi}{a}$$

Insgesamt folgt, dass $I(a) = \frac{1}{2}J(a) = \frac{1}{2}\sqrt{\frac{\pi}{a}}$.

(b) Für $B_R = \{(x,y) \in \mathbb{R}^2 |\ x^2 + y^2 \le R^2\} \subset \mathbb{R}^2$ gilt:
$\int_{B_R}(x^2+y^2)dxdy = \int_0^{2\pi}\int_0^R((r\cos\varphi)^2 + (r\sin\varphi)^2)r\, drd\varphi$

$$= 2\pi \int_0^R r^3 dr = 2\pi \frac{1}{4} r^4|_0^R = \frac{\pi}{2}R^4.$$

(c) **Uneigentliches Bereichsintegral** Mit B_R wie in (b) und den Kreisringen $K_{s,R} = \{(x,y) \in \mathbb{R}^2 |\ s^2 \le x^2 + y^2 \le R^2\}$ folgt

$$\int_{B_R} \frac{y^2}{(x^2+y^2)^{3/2}} dxdy = \lim_{s\downarrow 0} \int_{K_{s,R}} \frac{y^2}{(x^2+y^2)^{3/2}} dxdy$$

$$= \lim_{s\downarrow 0} \int_0^{2\pi} \int_s^R \frac{r^2\sin^2\varphi}{r^3} r\, drd\varphi$$

$$= \lim_{s\downarrow 0} \int_0^{2\pi} \int_s^R \sin^2\varphi\, drd\varphi =$$

$$= \lim_{s\downarrow 0} (R-s) \int_0^{2\pi} \sin^2\varphi\, d\varphi$$

$$= \text{(Partielle Integration) } R\pi.$$

9 Gewöhnliche Differentialgleichungen

9.1 Problemstellung

Betrachte Funktionen $f(t), y(t), \ldots$, die von der Zeit t abhängen.

Häufige Situation: Man kennt weder die Funktionen noch ihre Ableitungen, wohl aber Beziehungen zwischen ihnen.

Beispiele 9.1. (a) $y'(t) = cy(t)$ (c Konstante)

(b) $m\,y''(t) = a\,y'(t) + Dy(t)$ (m, a, D Konstanten)

(c) $y'(t) = y(t)^2 - t + 1$.

Definition 9.2. Eine Gleichung der Form

$$(*) \qquad y^{(n)}(t) = f(t, y(t), y^{(1)}(t), \ldots, y^{(n-1)}(t)),$$

wobei f eine Funktion von $(n+1)$-Variablen ist, heißt *gewöhnliche Differentialgleichung n-ter Ordnung.* Unter einer Lösung von $(*)$ versteht man eine n-mal differenzierbare Funktion $y(t)$, die $(*)$ erfüllt. Hängt y von mehreren Variablen ab und steht rechts eine Funktion, die von t und den partiellen Ableitungen von y abhängt, so spricht man von einer *partiellen Differentialgleichung.*
Wir betrachten zunächst nur gewöhnliche Differentialgleichungen 1. Ordnung.

9.2 Existenz und Eindeutigkeit der Lösungen

Numerische Lösungsverfahren: Man geht von einem bekannten Anfangswert zur Zeit $t = 0$ aus

$$y_0 = y(0).$$

Dann kennt man nach $(*)$ auch $y'(0) = f(0, y_0)$. Näherungsweise gilt für kleine $\tau > 0$

$$y'(0) \approx \frac{y(\tau) - y(0)}{\tau}, \text{ d.h. } y(\tau) \approx \tau y'(0) + y(0).$$

Mit $y(\tau)$ kennt man nach $(*)$ auch

$$y'(\tau) = f(\tau, y(\tau)).$$

Wie zuvor erhält man

$$y'(\tau) \approx \frac{y(2\tau) - y(\tau)}{\tau}, \text{ d.h. } y(2\tau) \approx \tau y'(\tau) + y(\tau).$$

Nacheinander erhält man näherungsweise

$$y(0), y(\tau), y(2\tau), \ldots, y((n+1)\tau) \approx \tau y'(n\tau) + y(n\tau).$$

Geometrische Bedeutung: Sei $D \subset \mathbb{R}^2, f : D \to \mathbb{R}, (t, y) \mapsto f(t, y)$.

In jedem Punkt (t, y) ist die Steigung bzw. Tangentenrichtung der Lösungskurve durch die Differentialgleichung $y'(t) = f(t, y)$ gegeben.

Nährungsweise kann man durch jeden vorgegebenen Anfangspunkt (t_0, y_0) eine Lösungskurve zeichnen.
In der Mathematik zeigt man (Forster II, §10, Satz 2 und 3):

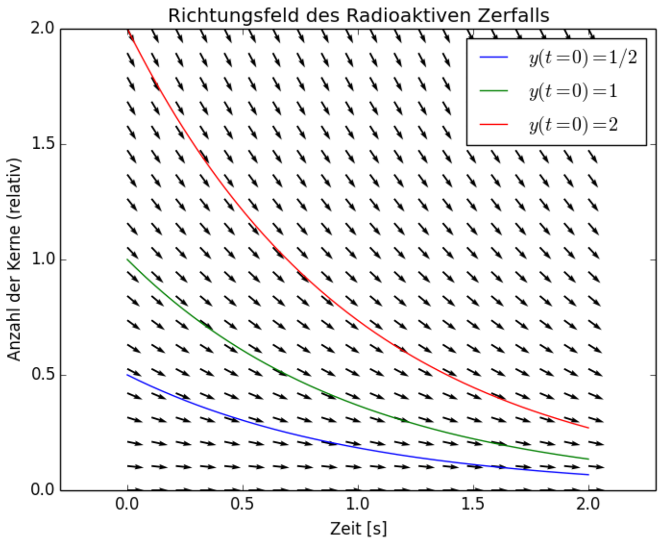


Abbildung 9: Richtungsfeld

Satz 9.3. *Ist $D \subset \mathbb{R}^2$ offen und $f : D \to \mathbb{R}$ stetig mit stetiger partieller Ableitung*

$$D \to \mathbb{R}, (t, y) \mapsto \frac{\partial f}{\partial y}(t, y)$$

nach y, so gilt

(i) (Eindeutigkeit) *Sind $y_1, y_2 : [\alpha, \beta] \to \mathbb{R}$ Lösungen von*

$$(*) \qquad y' = f(t, y)$$

mit $y_1(t_0) = y_2(t_0)$ für ein $t_0 \in [\alpha, \beta]$, so gilt $y_1(t) = y_2(t)$ für alle $t \in [\alpha, \beta]$.

(ii) (Existenz) *Für jedes Paar $(t_0, y_0) \in D$ existiert ein $\epsilon > 0$ und eine Lösung*

$$y :]t_0 - \epsilon, t_0 + \epsilon[\to \mathbb{R}$$

von $(*)$ mit Anfangswert $y(t_0) = y_0$.

Wie findet man konkrete Lösungen? Wir betrachten wichtige Spezialfälle.

9.3 Trennung der Variablen

Seien $I, J \subset \mathbb{R}$ im Folgenden beliebige Intervalle (alles ist zugelassen: offen, abgeschlossen, halboffen, beschränkt oder unberschränkt). Seien $t_0 \in I$, $y_0 \in J$ und $f : I \times J \to \mathbb{R}$ eine Abbildung der Form

$$f(t, y) = g(t)h(y)$$

mit stetigen Funktionen $g : I \to \mathbb{R}, h : J \to \mathbb{R} \setminus \{0\}$. Sei $y : I \to \mathbb{R}$ eine Lösung von

$$(*) \quad y'(t) = g(t)h(y(t)),\ y(t_0) = y_0.$$

Seien $G : I \to \mathbb{R}$, $H : J \to \mathbb{R}$ die Stammfunktionen von g und $1/h$ mit $G(t_0) = 0 = H(y_0)$, das heißt

$$G(t) = \int_{t_0}^{t} g(s)ds,\ H(y) = \int_{y_0}^{y} \frac{1}{h(\xi)} d\xi.$$

Dann ist $H \circ y : I \to \mathbb{R}$ differenzierbar mit

$$(H \circ y)'(t) = \frac{y'(t)}{h(y(t))} = g(t) \text{ für } t \in I.$$

Wegen $(H \circ y)(t_0) = 0 = G(t_0)$ folgt (Satz 10.5 in [SP])

$$G(t) = H(y(t)) \text{ für alle } t \in I.$$

Da $H'(y) = \frac{1}{h(y)}$ stetig ohne Nullstellen ist, ist nach dem Zwischenwertsatz (Satz 6.3 in [E]) $H' > 0$ oder $H' < 0$ auf J. Also ist $H : J \to \mathbb{R}$ streng monoton und hat eine differenzierbare Umkehrfunktionen $H^{-1} : H(J) \to \mathbb{R}$. Es folgt

$$y(t) = H^{-1}(G(t)) \text{ für alle } t \in I.$$

Zusammenfassend haben wir gezeigt: Wenn die Differentialgleichung $(*)$ mit $g : I \to \mathbb{R}$, $h : J \to \mathbb{R}$ wie oben eine Lösung $y : I \to \mathbb{R}$ auf dem Intervall I hat, dann ist $G(I) \subset H(J)$ und für $t \in I$ gilt

$$y(t) = H^{-1}(G(t)) \quad \text{ für alle } t \in I.$$

Die Umkehrung gilt auch (Satz 11.1 in [Fo]).

Satz 9.4. *Seien $g : I \to \mathbb{R}, h : J \to \mathbb{R} \setminus \{0\}$ stetig. Die Differentialgleichung*

$$y'(t) = g(t)h(y(t)),\ y(t_0) = y_0$$

hat eine Lösung $y : I' \to \mathbb{R}$ auf dem Intervall $I' \subset I$ dann und nur dann, wenn für die Funktionen

$$G(t) = \int_{t_0}^{t} g(s)ds,\ H(y) = \int_{y_0}^{y} \frac{1}{h(\xi)} d\xi$$

die Inklusion $G(I') \subset H(J)$ gilt. In diesem Fall ist die eindeutige Lösung gegeben durch

$$y(t) = H^{-1}(G(t)) \qquad (t \in I')$$

Merkregel Schreibe $y'(t) = g(t)h(y), y(t_0) = y_0$ formal als

$$\frac{dy}{dt} = g(t)h(y),\ y(t_0) = y_0$$

forme um zu

$$\frac{dy}{h(y)} = g(t)dt,$$

setze die integrierten Seiten gleich

$$H(y) = \int_{y_0}^{y} \frac{d\xi}{h(\xi)} = \int_{t_0}^{t} g(s)ds = G(t)$$

und löse die entstehende Gleichung auf nach y.

Beispiele 9.5. (a) **Chemische Reaktion 1. Ordnung:** Seien $a, k > 0$. Die Differentialgleichung

$$(*) \qquad \frac{dy}{dt} = y' = k(a - y), \qquad y(0) = 0$$

hat die gewünschte Form mit $g(t) \equiv k$, $h(y) = a - y$ für $t \in I = \mathbb{R}$, $y \in J = (-\infty, a)$. Schreibe die Differentialgleichung als

$$\frac{dy}{a - y} = k\ dt.$$

Gleichsetzen der integrierten Seiten führt zu

$$H(y) = \int_0^y \frac{d\xi}{a - \xi} = \int_0^t k\ ds = G(t).$$

Auflösen der Gleichung

$$H(y) = \ln \frac{a}{a - y} = -\ln(a - \xi)|_0^y = kt = G(t)$$

nach y ergibt

$$\frac{a - y}{a} = e^{-kt} \text{ bzw. } y(t) = a - ae^{-kt} = a(1 - e^{-kt}).$$

Wegen $H(J) = H((-\infty, a)) = \ln(0, \infty) = \mathbb{R} = G(\mathbb{R})$ ist

$$y : \mathbb{R} \to \mathbb{R},\ y(t) = a(1 - e^{-kt})$$

nach Satz 9.4 die eindeutige Lösung von $(*)$ auf $\mathbb{R}$.

(b) **Chemische Reaktion 2. Ordnung:** Seien $a, b, k > 0$.Die Differentialgleichungen

(i) $\frac{dy}{dt} = k(a - y)^2,\ y \in J = (-\infty, a), I = \mathbb{R}$

oder für $0 < a < b$

(ii) $\frac{dy}{dt} = k(a - y)(b - y),\ y \in J = (-\infty, a), I = \mathbb{R}$

mit Anfangswert $y(0) = 0$ liefern die Gleichungen

$$\text{(i)} \int_0^y \frac{d\xi}{(a - \xi)^2} = \int_0^t k\ ds \qquad \text{(ii)} \int_0^y \frac{d\xi}{(a - \xi)(b - \xi)} = \int_0^t k\ ds.$$

Im Fall (i) folgt:

$$H(y) = \frac{1}{a-y} - \frac{1}{a} = \frac{1}{a-\xi}|_0^y = kt = G(t)$$

$$\Rightarrow y(t) = a - \frac{a}{akt+1} = \frac{a^2kt}{akt+1} \text{ löst (i) für } t \in (-1/ka, \infty).$$

Beachte: $H(J) = (-1/a, \infty) = G((-1/ka, \infty))$.

Im Fall (ii) folgt mit $\frac{1}{(a-\xi)(b-\xi)} = (\frac{1}{b-\xi} - \frac{1}{a-\xi})\frac{1}{a-b}$

$$\begin{aligned} H(y) &= \frac{1}{a-b} \ln \frac{a-\xi}{b-\xi}|_0^y = kt = G(t) \\ \Rightarrow \ln(\frac{b}{a}\frac{a-y}{b-y}) &= (a-b)kt \\ \Rightarrow \frac{y-a}{y-b} &= \frac{a}{b}e^{(a-b)kt} \\ \Rightarrow y(t) &= \frac{a - a\ e^{(a-b)kt}}{1 - \frac{a}{b}\ e^{(a-b)kt}} = \frac{ab(1-e^{(a-b)kt})}{b - a\ e^{(a-b)kt}} \text{ löst (ii) für } t \in [0, \infty). \end{aligned}$$

Streng genommen müsste man noch ergänzen: Wegen $H(J) \supset [0, \infty)$ gilt $G([0, \infty)) \subset H(J)$.

9.4 Lineare Differentialgleichungen 1. Ordnung

Definition 9.6. Seien $a, b : I \to \mathbb{R}$ $(I \subset \mathbb{R}$ Intervall) stetig. Dann heißt

$$y' = a(t)y + b(t) \qquad (f(t, y) = a(t)y + b(t))$$

eine *lineare Differentialgleichung 1. Ordnung*. Sie heißt *homogen*, falls $b \equiv 0$ ist, sonst *inhomogen*.

Eindeutige Lösung der homogenen linearen Differentialgleichung

$$(*) \qquad y' = a(t)y, \qquad y(t_0) = c$$

ist

$$\varphi(t) = c \exp\left(\int_{t_0}^{t} a(s)ds\right) \qquad (t \in I).$$

Nach der Kettenregel löst φ die Gleichung $(*)$. Die Eindeutigkeit folgt aus Satz 9.3 (i). Dies folgt auch aus Satz 9.4 mit $g(t) = a(t)$ und $h(y) = y$.

Eindeutige Lösung der inhomogenen Gleichung

$$(**) \qquad y' = a(t)y + b(t),\ y(t_0) = c$$

ist

$$\psi(t) = \varphi(t)\left(c + \int_{t_0}^{t} \frac{b(s)}{\varphi(s)}ds\right) \text{ mit } \varphi(t) = \exp\left(\int_{t_0}^{t} a(s)ds\right).$$

Nach der Produktregel löst ψ die Gleichung $(**)$. Die Eindeutigkeit folgt wieder aus Satz 9.3.

Beispiele 9.7. (1) $y' = ky + b,\ y(0) = 0 \qquad (k, b \in \mathbb{R})$

Lösung: Mit $a(t) \equiv k,\ b(t) \equiv b$ und $t_0 = 0 = c$ folgt

$$\begin{aligned}\psi(t) &= \exp\left(\int_0^t k\ ds\right)\left(\int_0^t \frac{b}{\varphi(s)} ds\right) = be^{kt}\int_0^t e^{-ks} ds \\ &= \frac{b}{k}e^{kt}(1 - e^{-kt}) = \frac{b}{k}(e^{kt} - 1).\end{aligned}$$

(2) $y' = \frac{y}{t} - \sqrt{t},\ y(1) = 3$

Lösung: Mit $a(t) = \frac{1}{t},\ b(t) = -\sqrt{t}$ und $t_0 = 1, c = 3$ folgt

$$\varphi(t) = \exp\left(\int_1^t \frac{1}{s} ds\right) = t$$

und

$$\begin{aligned}\psi(t) &= t\left(3 - \int_1^t \frac{\sqrt{s}}{s} ds\right) = t(3 - 2s^{\frac{1}{2}}|_1^t) \\ &= t(5 - 2t^{\frac{1}{2}}).\end{aligned}$$

9.5 Exakte Differentialgleichungen

Sei $D \subset \mathbb{R}^2$ offen und seien $P, Q : D \to \mathbb{R}$ stetige Funktionen so, dass $0 \notin Q(D)$. Die Differentialgleichung

$$(*) \qquad y' = -\frac{P(t,y)}{Q(t,y)}$$

heißt *exakt*, wenn die Differentialform

$$P(t,y)dt + Q(t,y)dy$$

exakt ist (Abschnitt 7.3), d.h. wenn eine stetig partiell differenzierbare Funktion $F : D \to \mathbb{R}$ existiert mit

$$\frac{\partial F}{\partial t}(t,y) = P(t,y),\ \frac{\partial F}{\partial y}(t,y) = Q(t,y) \qquad \text{für alle } (t,y) \in D.$$

In diesem Fall heißt F *Stammfunktion* der Differentialgleichung $(*)$.

Satz 9.8. *Ist $F : D \to \mathbb{R}$ Stammfunktion der Differentialgleichung*

$$(*) \qquad y' = -\frac{P(t,y)}{Q(t,y)} \quad (P, Q : D \to \mathbb{R} \text{ stetig }, 0 \notin Q(D))$$

und ist $y : I \to \mathbb{R}$ differenzierbar mit

$$F(t, y(t)) = F(t_0, y_0) \text{ für alle } t \in I,$$

so löst y die Differentialgleichung $(*)$.

Idee: Nach der Kettenregel (Satz 6.12) gilt für $t \in I$

$$\begin{aligned} 0 = \frac{d}{dt}F(t,y(t)) &= \frac{\partial F}{\partial t}(t,y(t)) + \frac{\partial F}{\partial y}(t,y(t))y'(t) \\ &= P(t,y(t)) + Q(t,y(t))y'(t). \end{aligned}$$

Man beachte, dass aus der Gültigkeit der Gleichung $F(t,y(t)) = F(t_0,y_0)$ für alle $t \in I$, nicht immer automatisch folgt, dass $y(t_0) = y_0$ ist. Man muss also versuchen, diese Gleichung so nach $y(t)$ aufzulösen, dass die Lösung auch noch den richtigen Anfangswert an der Stelle t_0 hat.

Beispiel 9.9. Seien $P(t,y) = -(y - \frac{1}{t-1}), Q(t,y) = 2y - t$ definiert auf $D = \{(t,y) \in \mathbb{R}^2 | \, t < 1 \text{ und } y \neq \frac{t}{2}\}$. Wir wollen das Anfangswertproblem

$$(*) \qquad y' = \frac{y - \frac{1}{t-1}}{2y - t} = \frac{y(t-1) - 1}{(2y-t)(t-1)}, \quad y(0) = -1.$$

lösen. Gesucht ist eine stetig differenzierbare Funktion $F : D \to \mathbb{R}$ mit

$$(**) \qquad \frac{\partial F}{\partial t}(t,y) = -y + \frac{1}{t-1}, \; \frac{\partial F}{\partial y}(t,y) = 2y - t.$$

Intergrieren nach t bzw. y liefert

$$\begin{aligned} F(t,y) &= -yt + \ln(1-t) + \alpha_y, \\ F(t,y) &= \quad y^2 - yt + \beta_t \end{aligned}$$

mit Konstanten α_y, β_t, die nur von y bzw. t abhängen. Eine Lösung von $(**)$ ist daher

$$F(t,y) = y^2 - yt + \ln(1-t).$$

Auflösen der Gleichung $F(t,y) = F(0,-1) = 1$ nach y liefert

$$\begin{aligned} (y - \frac{t}{2})^2 &= 1 + \frac{t^2}{4} - \ln(1-t) \\ \Leftrightarrow y &= \frac{t}{2} \pm \sqrt{1 + \frac{t^2}{4} - \ln(1-t)}. \end{aligned}$$

Beachte, dass wegen der Logarithmus-Ungleichung

$$\ln x < x - 1 \text{ für } x \in (0,\infty) \setminus \{1\}$$

die rechte Seite wohldefiniert und differenzierbar in t ist für $t < 1$, denn für $t < 1$, $t \neq 0$ ist

$$1 + \frac{t^2}{4} - \ln(1-t) > \frac{1}{4}(t^2 + 4t + 4) = \frac{1}{4}(t+2)^2 \geq 0.$$

Also löst

$$y(t) = \frac{t}{2} - \sqrt{1 + \frac{t^2}{4} - \ln(1-t)} \qquad (t < 1)$$

das Anfangswertproblem $(*)$.

9.6 Integrierende Faktoren

Seien $P, Q : D \to \mathbb{R}$ wie in Abschnitt 9.5. Wenn

$$(*) \qquad y' = -\frac{P(t,y)}{Q(t,y)}$$

nicht exakt ist, hilft manchmal folgende Beobachtung weiter. Für jede stetige Funktion $R : D \to \mathbb{R} \setminus \{0\}$ hat $(*)$ dieselben Lösungen wie

$$(**) \qquad y' = -\frac{P(t,y)R(t,y)}{Q(t,y)R(t,y)}.$$

Es genügt also, eine Funktion R (einen "integrierenden Faktor") zu finden, für den $(**)$ exakt ist, und dann wie in Abschnitt 9.5 fortzufahren.

Einfaches Beispiel: Die Differentialgleichung

$$y' = g(t)h(y), \; y(t_0) = y_0$$

mit getrennten Variablen ist im Allgemeinen nicht exakt in der Form

$$y' = -\frac{P(t,y)}{Q(t,y)} \text{ mit } P(t,y) = g(t)h(y), \; Q(t,y) = -1,$$

aber mit dem integrierenden Faktor $R(t,y) = \frac{1}{h(y)}$ erhält man die exakte Form

$$y' = -\frac{g(t)}{-\frac{1}{h}(y)}.$$

Eine Stammfunktion der zugehörigen Differentialform

$$g(t)dt - \frac{1}{h(y)}dy$$

ist

$$F(t,y) = \int_{t_0}^{t} g(s)ds - \int_{y_0}^{y} \frac{d\xi}{h(\xi)}.$$

Auflösen der Gleichung

$$\int_{t_0}^{t} g(s)ds - \int_{y_0}^{y} \frac{d\xi}{h(\xi)} = F(t_0, y_0) = 0$$

nach y führt wieder zu dem Lösungsverfahren aus Abschnitt 9.3.

9.7 Lineare Differentialgleichungen 2. Ordnung

Eine *gewöhnliche Differentialgleichung 2. Ordnung* ist eine Gleichung der Form

$$y''(t) = f(t, y(t), y'(t)),$$

wobei f eine Funktion von drei Variablen ist. Einfachster Fall:

$$(*) \qquad y'' = f(t),\ f: I \to \mathbb{R} \text{ stetig auf einem Intervall } I \subset \mathbb{R}.$$

Für $a \in I$ beliebig und jede Lösung $y: I \to \mathbb{R}$ ist

$$y'(t) = y'(a) + \int_a^t f(s)ds$$

$$\Rightarrow y(t) = y(a) + \int_a^t y'(r)dr = y(a) + y'(a)(t-a) + \int_a^t \left(\int_a^r f(s)ds \right) dr.$$

Für $\alpha, \beta \in \mathbb{R}$ ist

$$y(t) = \alpha + \beta(t-a) + \int_a^t \left(\int_a^r f(s)ds \right) dr$$

die eindeutige Lösung von $(*)$ mit

$$y(a) = \alpha,\ y'(a) = \beta.$$

Eine *lineare Diffentialgleichung 2. Ordnung* ist eine Gleichung der Form

$$y'' + a_1(t)y'(t) + a_0(t)y(t) = b(t),$$

wobei $a_1, a_0, b: I \to \mathbb{R}$ stetige Funktionen auf einem Intervall $I \subset \mathbb{R}$ sind. Die Differentialgleichung heißt *homogen*, falls $b(t) \equiv 0$, sonst *inhomogen*.

Einfaches Beispiel: $m\,y'' = -ky$ (Federpendel)
oder äquivalent $y'' + \frac{k}{m}y = 0$. Mit $w = \sqrt{\frac{k}{m}}$ erhält man Lösungen $y_1(t) = A_1 \sin(\omega t), y_2(t) = A_2 \cos(\omega t)$ oder allgemeiner

$$y(t) = A_1 \sin(\omega t) + A_2 \cos(\omega t)$$

mit $A_1, A_2 \in \mathbb{R}$ beliebig. Hierbei sind die

$$\begin{array}{lccl} \text{Anfangsauslenkung} & A_2 & = & y(0), \\ \text{Anfangsgeschwindigkeit} & A_1\omega & = & y'(0) \end{array}$$

beliebig vorgebbar. Gibt es weitere Lösungen?
Nein, denn man kann zeigen (Forster II, §12):

Satz 9.10. *Sei $I \subset \mathbb{R}$ ein Intervall, sei $t_0 \in I$ und seien $a_0, a_1, b: I \to \mathbb{R}$ stetig. Dann existiert zu jedem Paar $(c_0, c_1) \in \mathbb{R}^2$ genau eine Lösung $\varphi: I \to \mathbb{R}$ des* Anfangswertproblems (AWP)

$$y'' + a_1(t)y' + a_0(t)y = b(t),\ y(t_0) = c_0, y'(t_0) = c_1.$$

Geometrisch: *Durch jeden Punkt $(t_0, c_0) \in I \times \mathbb{R}$ gibt es genau eine Lösungskurve $\varphi: I \to \mathbb{R}$, die durch den Punkt geht und Steigung c_1 in t_0 hat.*

Satz 9.11. *Seien $a_0, a_1, b : I \to \mathbb{R}$ stetig. Sei*

$$(*) \qquad y'' + a_1(t)y' + a_0(t)y = 0$$

die durch a_0, a_1, b gegebene homogene lineare Differentialgleichung 2. Ordnung und

$$(**) \qquad y'' + a_1(t)y' + a_0(t)y = b(t)$$

die entsprechende inhomogene Differentialgleichung mit rechter Seite b.

(a) Sind φ_1, φ_2 Lösungen von $(*)$ und ist $\alpha \in \mathbb{R}$, so sind auch

$$\varphi_1 + \varphi_2,\ \alpha\,\varphi_1$$

Lösungen von $(*)$.

(b) Ist φ_0 irgendeine Lösung von $(**)$, so erhält man alle Lösungen von $(**)$ als

$$\{\varphi_0 + \varphi |\ \varphi \textit{ löst } (*)\}.$$

(c) Die Menge

$$L_h = \{\varphi : I \to \mathbb{R} |\ \varphi \textit{ löst } (*)\}$$

ist bezüglich punktweiser Addition $(\varphi_1 + \varphi_2)(t) = \varphi_1(t) + \varphi_2(t)$ und skalarer Multiplikation $(\alpha\varphi)(t) = \alpha\,\varphi(t)$ ein 2-dimensionaler $\mathbb{R}$-Vektorraum.

Ideen:

(a) Seien $\varphi_1, \varphi_2 \in L_h$ und $\alpha \in \mathbb{R}$. Dann gilt für alle $t \in I$

$$(\varphi_1 + \varphi_2)''(t) + a_1(t)(\varphi_1 + \varphi_2)'(t) + a_0(t)(\varphi_1 + \varphi_2)(t) = 0 + 0 = 0,$$
$$(\alpha\varphi_1)''(t) + a_1(t)(\alpha\varphi_1)'(t) + a_0(t)(\alpha\varphi_1)(t) = \alpha 0 = 0.$$

(b) Sei φ_0 eine spezielle Lösung von $(**)$. Ist ψ eine weitere Lösung, so gilt

$$\psi = \varphi_0 + (\psi - \varphi_0) \text{ und } \psi - \varphi_0 \in L_h.$$

Umgekehrt ist für jedes $\varphi \in L_h$ die Funktion $\varphi_0 + \varphi$ eine Lösung von $(**)$.

(c) Nach (a) ist $L_h \subset \mathrm{Abb}(I, \mathbb{R})$ ein Teilvektorraum und nach Satz 9.10 ist für beliebiges $t_0 \in I$ die Abbildung

$$L_h \to \mathbb{R}^2,\ \varphi \mapsto (\varphi(t_0), \varphi'(t_0))$$

ein Vektorraumisomorphismus (d.h. eine bijektive lineare Abbildung zwischen Vektorräumen). Solche Abbildungen erhalten die Dimension.

Fazit: Man erhält alle Lösungen einer inhomogenen linearen Differentialgleichung 2. Ordnung, indem man eine spezielle Lösung der inhomogenen Differentialgleichung wählt und alle Lösungen der homogenen Gleichung addiert.

Unter einem *Fundamentalsystem* der Differentialgleichung

$$(*) \qquad y'' + a_1(t)y' + a_0(t)y = 0$$

verseht man eine Basis y_1, y_2 des Lösungsraumes L_h.

Satz 9.12. *Seien y_1, y_2 Lösungen von $(*)$. Dann gilt:*

$$y_1, y_2 \text{ ist ein Fundamentalsystem von } (*)$$

$$\Leftrightarrow \det \left(\begin{array}{c|c} y_1(t) & y_2(t) \\ \hline y_1'(t) & y_2'(t) \end{array}\right) \neq 0 \text{ für mindestens ein } t \in I$$

$$\Leftrightarrow \det \left(\begin{array}{c|c} y_1(t) & y_2(t) \\ \hline y_1'(t) & y_2'(t) \end{array}\right) \neq 0 \text{ für jedes } t \in I$$

Idee: Definiere die *Wronski-Determinante* von $(y_1.y_2)$ als

$$W(t) := W_{(y_1,y_2}(t) := \det \left(\begin{array}{c|c} y_1(t) & y_2(t) \\ \hline y_1'(t) & y_2'(t) \end{array}\right).$$

Ist y_1, y_2 ein Fundamentalsystem, so existieren nach Satz 9.10 für beliebiges $t_0 \in I$ und jedes Paar $(y_0, y_0') \in \mathbb{R}^2$ eindeutige $A, B \in \mathbb{R}$ mit

$$\begin{aligned} (A\ y_1 + B\ y_2)(t_0) &= y_0, \\ (A\ y_1 + B\ y_2)'(t_0) &= y_0', \end{aligned}$$

d.h. das LGS

$$\begin{aligned} &A\ y_1(t_0) + B\ y_2(t_0) = y_0, \\ &A\ y_1'(t_0) + B\ y_2'(t_0) = y_0' \\ \Leftrightarrow &\left(\begin{array}{c|c} y_1(t_0) & y_2(t_0) \\ \hline y_1'(t_0) & y_2'(t_0) \end{array}\right) \begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} y_0 \\ y_0' \end{pmatrix} \end{aligned}$$

ist für jedes $t_0 \in I$ und jede rechte Seite $(y_0, y_0') \in \mathbb{R}^2$ eindeutig lösbar. Nach Satz 3.3 ist dies äquivalent zu

$$W(t_0) \neq 0 \text{ für alle } t_0 \in I.$$

Ist $W(t) \neq 0$ für ein t, so sind y_1, y_2 linear unabhängig, denn sonst wäre

$$
\begin{aligned}
& y_1 = c\ y_2 \text{ oder } y_2 = c\ y_1 \\
\Rightarrow & \begin{pmatrix} y_1(t) \\ y_1'(t) \end{pmatrix} = c \begin{pmatrix} y_2(t) \\ y_2'(t) \end{pmatrix} \text{ oder } \begin{pmatrix} y_2(t) \\ y_2'(t) \end{pmatrix} = c \begin{pmatrix} y_1(t) \\ y_1'(t) \end{pmatrix}
\end{aligned}
$$

und wieder nach Satz 3.3 wäre $W(t) = 0$.

Bemerkung 9.13. Ist y_1, y_2 Fundamentalsystem der homogenen Differentialgleichung $(*)$, so ist für $t_0 \in I$ und $(y_0, y_0') \in \mathbb{R}^2$ die Funktion

$$y(t) = A\ y_1(t) + B\ y_2(t) \text{ mit}$$

$$\begin{pmatrix} A \\ B \end{pmatrix} = \left(\begin{array}{c|c} y_1(t_0) & y_2(t_0) \\ \hline y_1'(t_0) & y_2'(t_0) \end{array} \right)^{-1} \begin{pmatrix} y_0 \\ y_0' \end{pmatrix}$$

die Lösung von $(*)$ mit Anfangswerten $y(t_0) = y_0, y'(t_0) = y_0'$.

Beispiel 9.14. Die Differentialgleichung $y'' + y = 0$ hat die Lösungen

$$\varphi_1(t) = \cos t, \varphi_2(t) = \sin t.$$

Da ihre Wronski-Determinante

$$W_{(\varphi_1, \varphi_2)}(t) = \det \left(\begin{array}{c|c} \cos t & \sin t \\ \hline -\sin t & \cos t \end{array} \right) = 1 \neq 0$$

ist, ist φ_1, φ_2 ein Fundamentalsystem und die Lösungen sind genau die Funktionen

$$\varphi(t) = A\cos t + B\sin t \qquad (A, B \in \mathbb{R}).$$

Sei

$$M(t) = \left(\begin{array}{c|c} \cos t & \sin t \\ \hline -\sin t & \cos t \end{array} \right).$$

Für $y_0, y_0', t_0 \in \mathbb{R}$ ist die Lösung φ mit $\varphi(t_0) = y_0, \varphi'(t_0) = y_0'$ gegeben durch $\varphi = A\cos t + B\sin t$ mit Koeffizienten

$$\begin{pmatrix} A \\ B \end{pmatrix} = M(t_0)^{-1} \begin{pmatrix} y_0 \\ y_0' \end{pmatrix} = \left(\begin{array}{c|c} \cos t_0 & -\sin t_0 \\ \hline \sin t_0 & \cos t_0 \end{array} \right) \begin{pmatrix} y_0 \\ y_0' \end{pmatrix}.$$

9.8 Lineare Differentialgleichungen 2. Ordnung mit konstanten Koeffizienten

Für $a_0, a_1, b \in \mathbb{R}$ nennt man

$$(*) \qquad y'' + a_1 y' + a_0 y = b$$

eine *linare Differentialgleichung 2. Ordnung mit konstanten Koeffizienten*. Eine spezielle Lösung für $a_0 \neq 0$ ist

$$y \equiv \frac{b}{a_0}.$$

Nach Satz 9.11 braucht man nur noch die homogene Differentialgleichung

$$(**) \qquad y'' + a_1 y' + a_0 y = 0$$

zu lösen.

Ansatz Sei $y(t) = e^{\lambda t}$. Dann gilt

$$y''(t) + a_1 y'(t) + a_0 y(t) = (\lambda^2 + a_1\lambda + a_0)e^{\lambda t}$$

und y löst die homogene Gleichung $(**)$ genau dann, wenn λ die Gleichung

$$\lambda^2 + a_1\lambda + a_0 = 0 \quad (\textit{Charakteristische Gleichung})$$

erfüllt. Quadratische Ergänzung führt zu

$$\lambda^2 + a_1\lambda + a_0 = (\lambda + \frac{a_1}{2})^2 - (\frac{a_1^2 - 4a_0}{4}).$$

1. Fall: $a_1^2 - 4a_0 > 0$

$$\Rightarrow \lambda_\pm = \frac{1}{2}(-a_1 \pm \sqrt{a_1^2 - 4a_0}) \text{ lösen die charakteristische Gleichung.}$$

Da

$$W(t) = \det \left(\begin{array}{c|c} e^{e^{\lambda_+ t}} & e^{\lambda_- t} \\ \hline \lambda_+ e^{\lambda_+ t} & \lambda_- e^{\lambda_- t} \end{array} \right) = (\lambda_- - \lambda_+)e^{(\lambda_+ + \lambda_-)t} \neq 0$$

ist, ist $e^{\lambda_+ t}, e^{\lambda_- t}$ ein Fundamentalsystem von $(**)$ und die Lösungen sind genau die Funktionen

$$y(t) = A\ e^{\lambda_+ t} + B\ e^{\lambda_- t} \qquad (A, B \in \mathbb{R}).$$

2. Fall: $a_1^2 - 4a_0 = 0$

Die charakteristische Gleichung hat genau eine reelle Lösung $\lambda = -\frac{a_1}{2}$. Also ist

$$y_1(t) = e^{-\frac{a_1}{2}t}$$

Lösung von $(**)$. Eine zweite Lösung ist $y_2(t) = t\ e^{-\frac{a_1}{2}t}$, denn

$$y_2'(t) = (1 - \frac{a_1}{2}t)e^{-\frac{a_1}{2}t}, y_2''(t) = (\frac{a_1^2}{4}t - a_1)e^{-\frac{a_1}{2}t}$$
$$\Rightarrow y_2'' + a_1 y_2' + a_0 y_2 = 0.$$

Wegen

$$W(t) = \left[\det\left(\begin{array}{c|c} 1 & t \\ \hline -\frac{a_1}{2} & 1-\frac{a_1}{2}t \end{array}\right)\right] (e^{-\frac{a_1}{2}t})^2 = e^{-a_1 t} \neq 0$$

bilden $e^{-\frac{a_1}{2}t}, t\, e^{-\frac{a_1}{2}t}$ ein Fundamentalsystem von $(**)$ und die Lösungen sind genau die Funktionen

$$y(t) = A\; e^{-\frac{a_1}{2}t} + B\; t\; e^{-\frac{a_1}{2}t} \qquad (A, B \in \mathbb{R}).$$

3. Fall: $\quad a_1^2 - 4a_0 < 0$

Die charakteristische Gleichung hat zwei verschiedene nicht reelle Lösungen λ_1, λ_2. Mit

$$\sigma := -\frac{a_1}{2},\; \tau = \frac{1}{2}\sqrt{4a_0 - a_1^2} > 0$$

folgt $\lambda_1 = \sigma + i\tau,\; \lambda_2 = \sigma - i\tau$. Dies gibt die komplexen Lösungen

$$\psi_1(t) = e^{\sigma t}e^{i\tau t},\; \psi_2(t) = e^{\sigma t}e^{-i\tau t}.$$

Durch Bilden von Linearkombinationen erhält man die reellen Lösungen

$$\begin{aligned} \varphi_1(t) &= \frac{1}{2}(\psi_1(t) + \psi_2(t)) = e^{\sigma t}\cos(\tau t), \\ \varphi_2(t) &= \frac{1}{2i}(\psi_1(t) - \psi_2(t)) = e^{\sigma t}\sin(\tau t). \end{aligned}$$

Man prüft leicht nach, dass dies Lösungen der homogenen Gleichung $(**)$ sind. Wegen

$$W(t) = \left[\det\left(\begin{array}{c|c} \cos(\tau t) & \sin(\tau t) \\ \hline \sigma\cos(\tau t) - \tau\sin(\tau t) & \sigma\sin(\tau t) + \tau\cos(\tau t) \end{array}\right)\right] e^{2\sigma t} = \tau\; e^{2\sigma t} \neq 0$$

bilden φ_1, φ_2 ein Fundamentalsystem, und die Lösungen von $(**)$ sind genau die Funktionen der Form

$$y(t) = A\; e^{\sigma t}\cos(\tau t) + B\; e^{\sigma t}\sin(\tau t) \qquad (A, B \in \mathbb{R}).$$

Beispiel 9.15. Die charakteristische Gleichung von

$$y'' + 2y' + 2y = 0 \qquad (\text{ d.h. } a_1 = 2 = a_0 \Rightarrow a_1^2 - 4a_0 = -4 < 0)$$

hat die Lösungen

$$\lambda_1 = -1 + i,\; \lambda_2 = -1 - i \qquad (\sigma = -1,\; \tau = 1).$$

Also lautet die allgemeine Lösung $(\varphi_1(t) = e^{-t}\cos t,\; \varphi_2(t) = e^{-t}\sin t)$

$$\varphi(t) = A\; e^{-t}\cos t + B\; e^{-t}\sin t \qquad (A, B \in \mathbb{R}).$$

Die Lösung mit Anfangswerten $y(0) = 1 = y'(0)$ hat die Koeffizienten (siehe Bemerkung 9.13)

$$\begin{aligned} \begin{pmatrix} A \\ B \end{pmatrix} &= \left(\begin{array}{c|c} \varphi_1(0) & \varphi_2(0) \\ \hline \varphi_1'(0) & \varphi_2'(0) \end{array}\right)^{-1} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \left(\begin{array}{c|c} 1 & 0 \\ \hline -1 & 1 \end{array}\right)^{-1} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\ &= \left(\begin{array}{c|c} 1 & 0 \\ \hline 1 & 1 \end{array}\right) \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}. \end{aligned}$$

9.9 Der gedämpfte harmonische Oszillator

Die Bewegungsgleichung eines 1-dimsensionalen Federpendels gedämpft durch eine Reibungskraft proportinal zu $y'(t)$ und angetrieben durch die Federkraft proportional zur Auslenkung lautet

$$m\, y''(t) = -ay'(t) - D\, y(t) \qquad (a, D > 0).$$

Dies ist eine lineare Differentialgleichung 2.Ordnung

$$y''(t) + a_1 y'(t) + a_0 y(t) = 0$$

mit konstanten Koeffizienten $a_1 = \frac{a}{m}$, $a_0 = \frac{D}{m}$. Man unterscheidet die Fälle:

(a) **Starke Dämpfung:** $\quad a_1^2 > 4a_0 \Leftrightarrow a^2 > 4mD$

Die Lösungen (1. Fall in §9.8)

$$y(t) = A\, e^{\lambda_+ t} + B\, e^{\lambda_- t}$$

mit

$$\lambda_+ = \frac{1}{2m}(-a + \sqrt{a^2 - 4mD}),\ \lambda_- = \frac{1}{2m}(-a - \sqrt{a^2 - 4mD})$$

klingen wegen $\lambda_- < \lambda_+ < 0$ exponentiell ab, ohne zu schwingen.

(b) **Aperiodischer Grenzfall:** $\quad a_1^2 = 4a_0 \Leftrightarrow a^2 = 4mD$

Die Lösungen (2.Fall in §9.8)

$$y(t) = (A + Bt)e^{-\frac{a}{2m}t}$$

ergeben immer noch keine Schwingungen.

(c) **Schwache Dämpfung:** $\quad a_1^2 < 4a_0 \Leftrightarrow a^2 < 4mD$

Die Lösungen (3.Fall in §9.8)

$$y(t) = (A\, \cos(\tau t) + B\, \sin(\tau t))e^{\sigma t}$$

mit

$$\sigma = -\frac{a}{2m},\ \tau = \frac{1}{2m}\sqrt{4mD - a^2}$$

beschreiben Schwingungen mit Periode $\frac{2\pi}{\tau}$, deren Amplituden exponentiell abklingen.

Indem man die Additionstheoreme für die Winkelfunktionen benutzt, sieht man, dass reelle Zahlen $C \geq 0$ und $\varphi \in [0, 2\pi]$ existieren mit

$$A \cos(\tau t) + B \sin(\tau t) = C \cos(\tau t - \varphi) = C(\cos\varphi \cos(\tau t) + \sin\varphi \sin(\tau t))$$

für alle $t \in \mathbb{R}$. Offensichtlich genügt es, C und φ so zu wählen, dass

$$(A, B) = (C\cos\varphi, C\sin\varphi)$$

gilt. Wegen

$$\{(\cos\varphi, \sin\varphi) |\ \varphi \in [0, 2\pi]\} = \{(x, y) \in \mathbb{R}^2 |\ x^2 + y^2 = 1\}$$

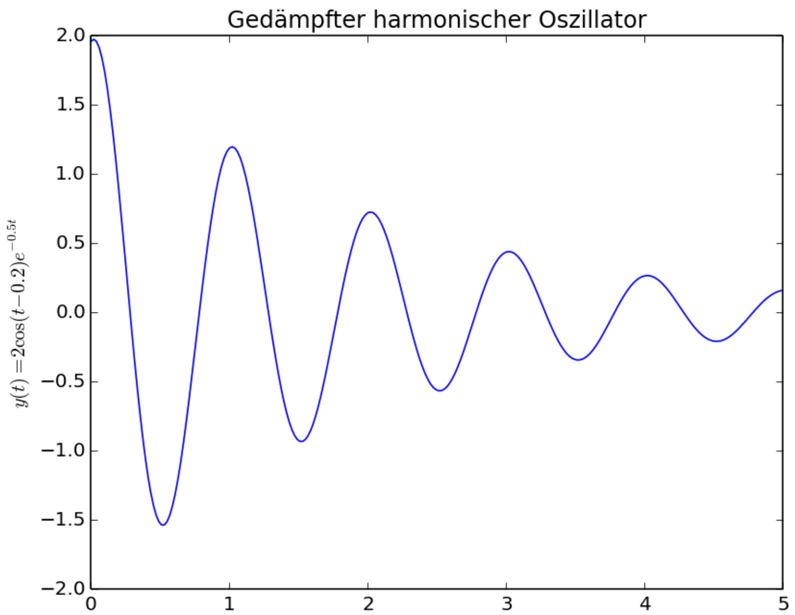
erhält man Lösungen, indem man $C = \sqrt{A^2 + B^2}$ setzt und ein $\varphi \in [0, 2\pi]$ wählt mit

$$(\cos\varphi, \sin\varphi) = (\frac{A}{C}, \frac{B}{C}).$$

Also besteht die Lösungsmenge im schwach gedämpften Fall genau aus allen phasenverschobenen Schwingungen

$$y(t) = C\cos(\tau t - \varphi)e^{\sigma t}$$

mit Frequenz τ und exponentiell abklingenden Amplituden $Ce^{\sigma t}$.



10 Anhang: Existenz von Determinanten

Sei $J_n = \{1, \ldots, n\}$. Unter einer *Permutation* der Menge J_n versteht man eine bijektive Abbildung $\pi : \{1, \ldots, n\} \to \{1, \ldots, n\}$. Sei

$$\mathcal{S}_n = \{\pi | \pi : \{1, \ldots, n\} \to \{1, \ldots, n\} \text{ ist bijektiv}\}.$$

Eine Transposition von J_n ist eine Permutation, die zwei Elemente vertauscht und alle anderen festhält. Für Transpositionen $\tau \in \mathcal{S}_n$ ist

$$\tau \circ \tau = \mathrm{id}.$$

Satz 10.1. *Jede Permutation $\pi \in \mathcal{S}_n$ $(n \geq 2)$ ist Komposition endlich vieler Transpositionen.*

Idee: Für $n = 2$ besteht $\mathcal{S}_2$ nur aus der identischen Abbildung und der Transposition von 1 und 2. Für beliebiges $n \geq 2$ folgt die Behauptung induktiv. Sei die Aussage des Satzes für ein $n \geq 2$ gezeigt und sei $\pi \in \mathcal{S}_{n+1}$. Setze $k := \pi(n+1)$ und definiere $\tau \in \mathcal{S}_{n+1}$ durch

$$\tau = \begin{cases} \text{Transposition von } k \text{ und } n+1, & \text{falls } k \neq n+1 \\ \qquad \mathrm{id} & \text{, falls } k = n+1. \end{cases}$$

Dann bildet die Permutation $\tau \circ \pi \in \mathcal{S}_{n+1}$ das Element $n+1$ auf sich ab. Nach Induktionsvoraussetzung existieren Transpositionen $\tau_1, \ldots, \tau_r \in \mathcal{S}_n$ mit

$$\tau \circ \pi(j) = \tau_1 \circ \ldots \circ \tau_r(j) \text{ für } j = 1, \ldots, n.$$

Indem man $\tau_j(n+1) := n+1$ definiert, kann man $\tau_1, \ldots, \tau_r$ zu Transpositionen von $\mathcal{S}_{n+1}$ machen mit

$$\tau \circ \pi = \tau_1 \circ \ldots \circ \tau_r.$$

Dann ist aber $\pi = \tau \circ \tau_1 \circ \ldots \circ \tau_r$.

Man kann zeigen: (Lingenberg, Lineare Algebra, § 9.2)

Kein Produkt einer geraden Anzahl von Transpositionen ist gleich einem Produkt einer ungeraden Anzahl von Transpositionen. Mit Produkt ist hier die Komposition von Abbildungen gemeint.

Definition 10.2. Für $\pi \in \mathcal{S}_n$ heißt

$$\epsilon(\pi) = \begin{cases} 1, & \pi \text{ ist Produkt einer geraden Anzahl von Transpositionen} \\ -1, & \pi \text{ ist Produkt einer ungeraden Anzahl von Transpositionen} \end{cases}$$

das *Vorzeichen* von π.

Rechenregeln: Es gilt:

(1) $\epsilon(\mathrm{id}) = 1, \epsilon(\tau) = -1$ für jede Transposition τ,

(2) $\epsilon(\sigma \circ \tau) = \epsilon(\sigma) \cdot \epsilon(\tau)$ für alle $\sigma, \tau \in \mathcal{S}_n$,

(3) $\epsilon(\pi) = (-1)^r$, falls π Produkt von r Transpositionen ist.

Definition 10.3. Für $A = (a_{ij}) \in \mathbb{R}^{n,n}$ (oder $\mathbb{C}^{n,n}$) heißt die Zahl

$$\det(A) = \sum_{\pi \in \mathcal{S}_n} \epsilon(\pi) a_{1\pi(1)} a_{2\pi(2)} \dots a_{n\pi(n)}$$

die *Determinante* von A.

Beispiele 10.4. Für $n = 1$ besteht die Matrix A aus nur einem Koeffizienten und die Permutationsgruppe $\mathcal{S}_1$ nur aus der Identität. Es gilt

$$A = (a_{11}) \Rightarrow \det(A) = a_{11}.$$

Sei $n = 2$ und

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

Da $\mathcal{S}_2 = \{\mathrm{id}, \tau\}$ nur aus der Identität und der Transposition τ mit $\tau(1) = 2,\ \tau(2) = 1$ besteht, folgt

$$\begin{aligned} \det(A) &= \epsilon(\mathrm{id}) a_{11} a_{22} + \epsilon(\tau) a_{1\tau(1)} a_{2\tau(2)} \\ &= a_{11} a_{22} - a_{12} a_{21}. \end{aligned}$$

Sei $n = 3$ und

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}.$$

Die Permutationsgruppe $\mathcal{S}_3$ besteht aus $3! = 6$ Permutationen. Auflisten dieser Permutationen und Anwenden der Definition liefert

$$\begin{aligned} \det(A) &= (a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32}) - (a_{12}a_{21}a_{33} - a_{13}a_{21}a_{32}) + (a_{12}a_{23}a_{31} - a_{13}a_{22}a_{31}) \\ &= a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{21} \begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + a_{31} \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} \\ &= a_{11} \det(A_{11}) - a_{21} \det(A_{21}) + a_{31} \det(A_{31}), \end{aligned}$$

wobei $A_{ij} \in \mathbb{R}^{2,2}$ die (2×2)-Matrix ist, die aus A durch Streichen der i-ten Zeile und j-ten Spalte entsteht. Diese Rechenregel nennt man Entwicklung nach der 1. Spalte.

Wie in Abschnitt 3.2 gezeigt, ist die Determinantenfunktion

$$\mathbb{R}^{n,n} \to \mathbb{R}, A \mapsto \det A$$

eindeutig bestimmt durch die Eigenschaften (1),(2) und (3) aus Abschnitt 3.1. Wir zeigen im Folgenden, dass diese Eigenschaften direkt aus der oben gegebenen Definition der Determinante folgen.

(1) Besitzt $A \in \mathbb{R}^{n,n}$ die Zeilenvektoren

$$A_\mu = (a_{\mu 1}, \ldots, a_{\mu n}) \qquad (\mu = 1, \ldots, n)$$

und ersetzt man den i-ten Zeilenvektor durch $A_i + A_i'$ mit

$$A_i' = (a_{i1}', \ldots, a_{in}'),$$

so folgt

$$\det \begin{pmatrix} A_1 \\ \vdots \\ A_i + A_i' \\ \vdots \\ A_n \end{pmatrix} = \sum_{\pi \in \mathcal{S}_n} \epsilon(\pi) a_{1\pi(1)} \cdots (a_{i\pi(i)} + a_{i\pi(i)}') \cdots a_{n\pi(n)}$$

$$= \det \begin{pmatrix} A_1 \\ \vdots \\ A_i \\ \vdots \\ A_n \end{pmatrix} + \det \begin{pmatrix} A_1 \\ \vdots \\ A_i' \\ \vdots \\ A_n \end{pmatrix}.$$

Für $\lambda \in \mathbb{R}$ gilt

$$\det \begin{pmatrix} A_1 \\ \vdots \\ \lambda A_i \\ \vdots \\ A_n \end{pmatrix} = \sum_{\pi \in \mathcal{S}_n} \epsilon(\pi) a_{1\pi(1)} \cdots (\lambda a_{i\pi(i)}) \cdots a_{n\pi(n)} = \lambda \det A.$$

(2) Entsteht A' aus A durch Vertauschen der i-ten und j-ten Zeile mit $1 \leq i < j \leq n$ und bezeichnet τ die Transpositionen von i und j, so folgt

$$\begin{aligned}
&\det A' = \textstyle\sum_{\pi \in \mathcal{S}_n} \epsilon(\pi) a_{1\pi(i)} \cdots a_{i\pi(i)}' \cdots a_{j\pi(j)}' \cdots a_{n\pi(n)} \\
&= -\textstyle\sum_{\pi \in \mathcal{S}_n} \epsilon(\pi \circ \tau) a_{1\pi\circ\tau(1)} \cdots a_{j\pi\circ\tau(j)} \cdots a_{i\pi\circ\tau(ii)} \cdots a_{n\pi\circ\tau(n)} \\
&= -\textstyle\sum_{\pi \in \mathcal{S}_n} \epsilon(\pi) a_{1\pi(1)} \cdots a_{j\pi(j)} \cdots a_{i\pi(i)} \cdots a_{n\pi(n)} \\
&= -\det A.
\end{aligned}$$

(3) Für die Determinante der Einheitsmatrix $I \in \mathbb{R}^{n,n}$ ergibt sich

$$\det I = \sum_{\pi \in \mathcal{S}_n} \epsilon(\pi)\delta_{1\pi(1)}\delta_{2\pi(2)} \cdots \delta_{n\pi(n)} = \epsilon(\mathrm{id}) = 1.$$

Hierbei ist definitionsgemäß $\delta_{ij} = 0$ für $i \neq j$ und $\delta_{ij} = 1$ für $i = j$ (*Kronecker-Symbol*).

Literatur

[E] Eschmeier, J., Mathematik für Naturwissenschaftler I, Vorlesungsskript, Universität des Saarlandes, 2018/19.

[Fo] Forster, O., Analysis 2, Springer Spektrum, 2013.

[Fo3] Forster, O., Analyisis 3, Springer Spektrum, 2017.

[SP] Schulze-Pillot, R., Mathematik für Naturwissenschaftlier, Vorlesungsskript, Universität des Saarlandes, 2017/18.