

Inhaltsverzeichnis

1	Eigenwerte und Eigenvektoren	1
1.1	Komplexe Matrizen	1
1.2	Einführendes Beispiel	4
1.3	Eigenwerte und Eigenvektoren	5
1.3.1	Berechnung von Eigenwerten und Eigenvektoren	5
1.3.2	Eigenschaften von Eigenwerten und Eigenvektoren	7
1.3.3	Diagonalähnliche Matrizen	8
2	Ebene Kurven, Raumkurven und Anwendungen	16
2.1	Raumkurven	16
2.2	Bogenlänge und Krümmung ebener Kurven	19
2.3	Bogenlänge und Krümmung einer Raumkurve	25
2.4	Skalar- und Vektorfelder	26
2.4.1	Divergenz und Rotation eines Vektorfeldes	27
2.5	Linien- oder Kurvenintegrale	29
3	Grundlagen der Statistik	33
3.1	Vorbemerkungen	33
3.2	Zufallsexperimente - grundlegende Begriffe und Eigenschaften	34
3.3	Wahrscheinlichkeitsaxiome	36
3.4	Laplace-Experimente	38
3.5	Hilfsmittel aus der Kombinatorik	39
3.6	Bedingte Wahrscheinlichkeiten und Unabhängigkeit von Ereignissen	42
3.7	Zufallsvariable und Verteilungsfunktion	48
3.8	Diskrete Zufallsvariablen	50
3.9	Stetige Zufallsvariable	60
3.10	Funktionen von einer oder mehreren Zufallsvariablen	69
3.11	Grenzwertsätze	76

Kapitel 1

Eigenwerte und Eigenvektoren

Es empfiehlt sich, die Kapitel zur Linearen Algebra aus dem ersten Semester zu wiederholen.

1.1 Komplexe Matrizen

Bisher haben wir nur Matrizen mit reellen Einträgen betrachtet. Im Folgenden werden wir auch komplexe Einträge zulassen, d. h. $A \in \mathbb{C}^{m \times n}$ mit $a_{ik} \in \mathbb{C}$,

$$A = (a_{ik}), i = 1, 2, \dots, m; k = 1, 2, \dots, n.$$

Solche Matrizen treten z. B. bei der Behandlung linearer Netzwerke auf.

1.1.1 Beispiel

Bei einem sogenannten Vierpol (Zweitor) mit den komplexen Eingangs- bzw. Ausgangsspannungen U_1 bzw. U_2 , den komplexen Eingangs- bzw. Ausgangsströmen I_1 bzw. I_2 und dem komplexen Querwiderstand $Z = R + j\omega L$ erhält man durch Anwendung der Maschenregel das lineare Gleichungssystem

$$\begin{aligned} -U_1 + ZI_1 - ZI_2 &= 0 &\iff U_1 &= ZI_1 - ZI_2 \\ U_2 - ZI_1 + ZI_2 &= 0 &\iff U_2 &= ZI_1 - ZI_2. \end{aligned}$$

Die Maschenregel besagt, dass in jeder Masche die Summe der Spannungen gleich Null ist. Das lineare Gleichungssystem lässt sich in Matrizenform schreiben (Ohmsches Gesetz).

$$\begin{pmatrix} Z & -Z \\ Z & -Z \end{pmatrix} \begin{pmatrix} I_1 \\ I_2 \end{pmatrix} = \begin{pmatrix} U_1 \\ U_2 \end{pmatrix}$$

1.1.2 Bemerkung

Die Rechenoperationen wie Addition, Subtraktion, Multiplikation und Multiplikation mit einer skalaren Größe gelten analog zu den Festlegungen für reelle Matrizen.

Für Matrizen aus $\mathbb{C}^{m \times n}$ benötigen wir noch einige weitere Definitionen.

1.1.3 Definition

Sei $A \in \mathbb{C}^{m \times n}$. Dann heißt

- $A^* = (a_{ik}^*) \in \mathbb{C}^{m \times n}$ die zu A konjugierte Matrix,
- $\bar{A} = (A^*)^T = (a_{ki}^*) \in \mathbb{C}^{m \times n}$ die konjugiert transponierte Matrix

1.1.4 Beispiel

$$\begin{aligned} A &= \begin{pmatrix} 1+2j & 4+5j & j \\ 2-3j & 2 & 3-j \end{pmatrix} \\ A^* &= \begin{pmatrix} 1-2j & 4-5j & -j \\ 2+3j & 2 & 3+j \end{pmatrix} \\ \bar{A} &= \begin{pmatrix} 1-2j & 2+3j \\ 4-5j & 2 \\ -j & 3+j \end{pmatrix} \end{aligned}$$

1.1.5 Bemerkung

Offenbar gelten folgende Rechengesetze:

$$\begin{aligned} (A^*)^* &= A \\ (A+B)^* &= A^*+B^* \\ (A \cdot B)^* &= A^* \cdot B^* \\ \bar{\bar{A}} &= (A^T)^* \\ \overline{(\bar{A})} &= A \\ \overline{A+B} &= \bar{A}+\bar{B} \\ \overline{A \cdot B} &= \bar{B} \cdot \bar{A} \end{aligned}$$

Wir beschäftigen uns zunächst mit einigen speziellen quadratischen Matrizen, die für technische Anwendungen besonders wichtig sind. Im komplexen Fall sind dies die hermiteschen und unitären Matrizen, was im reellen Fall den symmetrischen und orthogonalen Matrizen entspricht.

1.1.6 Definition

Eine Matrix $A \in \mathbb{C}^{n \times n}$ heißt hermitesch, wenn $A = \bar{A}^T$ gilt, d. h.

$$a_{ik} = a_{ki}^* \text{ für alle } i, k = 1, 2, \dots, n.$$

1.1.7 Bemerkung

Ist $A \in \mathbb{C}^{n \times n}$ hermitesch, und zerlegt man A in $A = B + jC$ mit reellen Matrizen $B, C \in \mathbb{R}^{n \times n}$, so gilt:

- Alle Elemente $a_{ii}, i = 1, 2, \dots, n$, der Hauptdiagonalen sind reell, da nach Definition $a_{ii} = a_{ii}^*$ gelten muss.
- B ist eine symmetrische und C eine schiefsymmetrische Matrix, d. h.

$$B = B^T \text{ und } C = -C^T \text{ bzw.}$$

$$b_{ik} = b_{ki} \text{ und } c_{ik} = -c_{ki} \text{ für alle } i, k = 1, 2, \dots, n$$

- $\det A \in \mathbb{R}$.
- Für reelle Matrizen fallen die Begriffe hermitesch und symmetrisch zusammen.

1.1.8 Beispiel

Die Matrix $A \in \mathbb{C}^{3 \times 3}$,

$$A = \begin{pmatrix} 1 & 1-j & 3 \\ 1+j & 2 & 2j \\ 3 & -2j & -1 \end{pmatrix}$$

ist hermitesch.

Es gilt $A = B + jC$ mit

$$B = \begin{pmatrix} 1 & 1 & 3 \\ 1 & 2 & 0 \\ 3 & 0 & -1 \end{pmatrix} \text{ und } C = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 2 \\ 0 & -2 & 0 \end{pmatrix}$$

$$\begin{aligned} \det A &= \begin{vmatrix} 1 & 1-j & 3 \\ 1+j & 2 & 2j \\ 3 & -2j & -1 \end{vmatrix} \\ &= 1 \cdot 2 \cdot (-1) + (1-j) \cdot 2j \cdot 3 + 3 \cdot (1+j)(-2j) - 3 \cdot 2 \cdot 3 - (-2j) \cdot 2j \cdot 1 - (-1)(1+j)(1-j) \\ &= -10 \end{aligned}$$

1.1.9 Definition

Eine Matrix $A \in \mathbb{C}^{n \times n}$ heißt unitär, wenn gilt:

$$A \cdot \bar{A} = E.$$

Ist speziell $A \in \mathbb{R}^{n \times n}$ mit $A \cdot A^T = E$, so nennt man A orthogonal.

1.1.10 Bemerkung

Sei $A \in \mathbb{C}^{n \times n}$ unitär. Dann ist auch \bar{A} unitär und es gilt:

- $A \cdot \bar{A} = \bar{A} \cdot A = E$
- $\bar{A} = A^{-1}$
- $|\det A| = 1$
- Das Produkt unitärer Matrizen ist wieder unitär.

1.1.11 Beispiel

Sei $A = \frac{1}{\sqrt{3}} \begin{pmatrix} 1+j & 1 \\ j & -1-j \end{pmatrix}$. Dann ist

$$A^* = \frac{1}{\sqrt{3}} \begin{pmatrix} 1-j & 1 \\ -j & -1+j \end{pmatrix}, \quad \bar{A} = \frac{1}{\sqrt{3}} \begin{pmatrix} 1-j & -j \\ 1 & -1+j \end{pmatrix}$$

Wir prüfen, ob A unitär ist, indem wir $A \cdot \bar{A}$ berechnen.

$$\begin{aligned} A \cdot \bar{A} &= \frac{1}{3} \begin{pmatrix} 1+j & 1 \\ j & -1-j \end{pmatrix} \begin{pmatrix} 1-j & -j \\ 1 & -1+j \end{pmatrix} \\ &= \frac{1}{3} \begin{pmatrix} (1+j)(1-j)+1 & (1+j)(-j)+(-1+j) \\ j(1-j)+(-1-j) & -j^2+(-1-j)(-1+j) \end{pmatrix} \\ &= \frac{1}{3} \begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \end{aligned}$$

Wir erhalten also $A \cdot \bar{A} = E$, d. h. A ist unitär.

1.2 Einführendes Beispiel

In vielen technischen Anwendungen hat man sogenannte Eigenwertprobleme zu behandeln. Wir betrachten zunächst ein sehr einfaches geometrisches Beispiel. Als Anwendung werden wir zum Schluss das Verhalten eines gekoppelten mechanischen Problems betrachten. Wir betrachten die Spiegelung eines Punktes $P(x_1, x_2)$ in der x_1x_2 -Ebene an der x_1 -Achse. Die Spiegelung liefert einen Bildpunkt $P'(u_1, u_2)$, wobei $u_1 = x_1$, $u_2 = -x_2$ bzw. "ausführlicher"

$$\begin{aligned}u_1 &= 1 \cdot x_1 + 0 \cdot x_2 \\u_2 &= 0 \cdot x_1 - 1 \cdot x_2\end{aligned}$$

In Matrix-Vektorform bedeutet dies:

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}. \quad (1.1)$$

Die Fragestellung ist nun folgende:

Welche vom Nullvektor verschiedenen Vektoren gehen bei der Spiegelung in einen Vektor gleicher oder entgegengesetzter Richtung über?

Wann gilt also $u = \lambda x$?

Eingesetzt in (1.1) soll also gelten:

$$\begin{aligned}Ax &= \lambda x \\ \Leftrightarrow Ax &= \lambda Ex \\ \Leftrightarrow (A - \lambda E)x &= \vec{0}\end{aligned}$$

Für unser Beispiel also

$$\begin{pmatrix} 1 - \lambda & 0 \\ 0 & -1 - \lambda \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Dies ist ein homogenes lineares Gleichungssystem. Es besitzt genau dann eine vom Nullvektor verschiedene Lösung, wenn $\det(A - \lambda E) = 0$ gilt, d. h. in unserem Beispiel, wenn

$$\begin{aligned}\begin{vmatrix} 1 - \lambda & 0 \\ 0 & -1 - \lambda \end{vmatrix} &= 0 \\ \Leftrightarrow (1 - \lambda)(-1 - \lambda) &= 0 \\ \Leftrightarrow \lambda_1 = 1, \lambda_2 = -1\end{aligned}$$

Die Lösungen heißen Eigenwerte von A . Zu diesen bestimmt man nun die sogenannten Eigenvektoren, d. h. Vektoren, für die $A\vec{x} = \lambda\vec{x}$ gilt.

- $\lambda_1 = 1$: Dann gilt

$$\begin{aligned}(A - 1 \cdot E)\vec{x} = \vec{0} &\Leftrightarrow \begin{cases} 0 \cdot x_1 + 0 \cdot x_2 = 0 \\ 0 \cdot x_1 - 2 \cdot x_2 = 0 \end{cases} \\ &\Leftrightarrow x_2 = 0, x_1 = \alpha\end{aligned}$$

Somit ist $\vec{x}_1 = \alpha \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ Eigenvektor zum Eigenwert $\lambda_1 = 1$. Dieser ist bis auf einen konstanten Faktor eindeutig bestimmt.

Nach Normierung auf Betrag 1 erhält man den normierten Eigenvektor $\vec{v}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ zum Eigenwert $\lambda_1 = 1$.

- $\lambda_2 = -1$:

$$\begin{aligned}(A + 1 \cdot E)\vec{x} = \vec{0} &\Leftrightarrow \begin{cases} 2 \cdot x_1 + 0 \cdot x_2 = 0 \\ 0 \cdot x_1 + 0 \cdot x_2 = 0 \end{cases} \\ &\Leftrightarrow x_1 = 0, x_2 = \beta\end{aligned}$$

Somit ist $\vec{x}_2 = \beta \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ Eigenvektor zum Eigenwert $\lambda_2 = -1$ und $\vec{v}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ normierter Eigenvektor zum Eigenwert $\lambda_2 = -1$.

1.2.1 Bemerkung

Geometrisch bedeutet das Ergebnis des Beispiels, dass Punkte auf der x_1 -Achse erhalten bleiben. Länge und Richtung der Ortsvektoren $\begin{pmatrix} \alpha \\ 0 \end{pmatrix}$ bleiben erhalten.

Für die Ortsvektoren $\begin{pmatrix} 0 \\ \beta \end{pmatrix}$ bleibt die Länge erhalten, die Richtung wird umgekehrt.

1.3 Eigenwerte und Eigenvektoren

A bezeichne im Folgenden stets eine komplexe $(n \times n)$ -Matrix. Wir wollen Vektoren $\vec{x} \in \mathbb{C}^n$ bestimmen, die bei Anwendung von A nur auf ein skalares Vielfaches ihrer selbst abgebildet werden.

1.3.1 Definition

Eine Zahl $\lambda \in \mathbb{C}$ heißt Eigenwert der Matrix A , falls ein Vektor $\vec{x} \in \mathbb{C}^n \setminus \{0\}$ existiert, so dass gilt:

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

Jeder solche Vektor \vec{x} heißt dann Eigenvektor der Matrix A zum Eigenwert λ . Die Menge aller Eigenwerte von A nennt man das Spektrum von A .

Im Folgenden suchen wir nun Antworten zu:

- Wie berechnet man Eigenwerte und Eigenvektoren?
- Wie viele Eigenwerte und zugehörige Eigenvektoren gibt es und welche Eigenschaften besitzen sie?
- Was sind besondere Eigenschaften von Eigenwerten und Eigenvektoren spezieller Matrizen?

1.3.1 Berechnung von Eigenwerten und Eigenvektoren

Sei $A \in \mathbb{C}^{n \times n}$ und $\vec{x} \neq \vec{0}$ ein Eigenvektor zum Eigenwert λ von A , d. h.

$$\begin{aligned} A\vec{x} &= \lambda\vec{x} \\ \Leftrightarrow A\vec{x} &= \lambda E\vec{x} \\ \Leftrightarrow (A - \lambda E)\vec{x} &= \vec{0} \end{aligned}$$

Ausführlicher geschrieben:

$$\begin{pmatrix} a_{11} - \lambda & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} - \lambda & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & a_{n-1n} \\ a_{n1} & a_{n2} & \cdots & a_{nn-1} & a_{nn} - \lambda \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Dies ist ein homogenes lineares Gleichungssystem. Wenn $\vec{x} \neq \vec{0}$ Lösung ist, muss die Koeffizientendeterminante verschwinden, d. h. es muss gelten

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

1.3.2 Definition (Charakteristisches Polynom)

Es sei $A \in \mathbb{C}^{n \times n}$ eine $n \times n$ -Matrix. Die Funktion p ,

$$\begin{aligned} p: \mathbb{C} &\rightarrow \mathbb{C}, \\ \lambda &\mapsto \det(A - \lambda E), \end{aligned}$$

heißt *charakteristisches Polynom von A*. p ist aufgrund der Definition der Determinante ein algebraisches Polynom genau n -ten Grades in λ .

1.3.3 Satz

Es sei $A \in \mathbb{C}^{n \times n}$ eine $n \times n$ -Matrix.

- a) Eine Zahl $\lambda \in \mathbb{C}$ ist genau dann ein Eigenwert von A , wenn λ eine Nullstelle des charakteristischen Polynoms von A ist, also

$$p(\lambda) = \det(A - \lambda E) = 0$$

gilt.

- b) Ein Vektor $\vec{x} \in \mathbb{C}^n \setminus \{\vec{0}\}$ ist genau dann Eigenvektor von A zum Eigenwert λ , wenn \vec{x} eine Lösung des homogenen, linearen Gleichungssystems

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

ist.

Aus den obigen Überlegungen leitet man die folgende Berechnungsstrategie für Eigenwerte und Eigenvektoren ab:

- a) Berechne – als Funktion von λ – das charakteristische Polynom p von A gemäß $p(\lambda) = \det(A - \lambda E)$ sowie dessen (unter Umständen auch mehrfach vorkommende) Nullstellen $\lambda_1, \dots, \lambda_n \in \mathbb{C}$. Dies sind die Eigenwerte von A .
- b) Berechne jeweils eine vom Nullvektor verschiedene Lösung $\vec{x}_1, \dots, \vec{x}_n$ der homogenen linearen Gleichungssysteme

$$\begin{aligned} (A - \lambda_1 E)\vec{x}_1 &= \vec{0}, \\ (A - \lambda_2 E)\vec{x}_2 &= \vec{0}, \\ &\vdots \\ (A - \lambda_n E)\vec{x}_n &= \vec{0}. \end{aligned}$$

Dann ist \vec{x}_1 Eigenvektor zu λ_1 , \vec{x}_2 Eigenvektor zu λ_2 , \dots , \vec{x}_n Eigenvektor zu λ_n .

1.3.4 Beispiel

Wir berechnen die Eigenwerte und Eigenvektoren der Matrix

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

1. Schritt:

$$\begin{aligned} p(\lambda) &= \det(A - \lambda E) = \begin{vmatrix} 1 - \lambda & 3 \\ 1 & -1 - \lambda \end{vmatrix} \\ &= (1 - \lambda)(-1 - \lambda) - 3 \stackrel{!}{=} 0 \end{aligned}$$

Also:

$$\lambda_1 = 2, \lambda_2 = -2$$

Die Eigenwerte von A lauten $\lambda_1 = 2$ und $\lambda_2 = -2$.

2. Schritt

$$(A - \lambda_1 E)\vec{x}_1 = \vec{0} \Leftrightarrow \begin{pmatrix} 1-2 & 3 \\ 1 & -1-2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Eine Lösung ist z. B. $\vec{x}_1 = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$ bzw. normiert $\vec{v}_1 = \frac{1}{\sqrt{10}} \begin{pmatrix} 3 \\ 1 \end{pmatrix}$.

$$(A - \lambda_2 E)\vec{x}_2 = \vec{0} \Leftrightarrow \begin{pmatrix} 1-(-2) & 3 \\ 1 & -1-(-2) \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Eine Lösung ist z. B. $\vec{x}_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$ bzw. normiert $\vec{v}_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \end{pmatrix}$.

1.3.2 Eigenschaften von Eigenwerten und Eigenvektoren

1.3.5 Satz (Eigenschaften von Eigenwerten und Eigenvektoren)

Es sei $A \in \mathbb{C}^{n \times n}$ eine $n \times n$ -Matrix. Dann gilt:

- A hat höchstens n verschiedene Eigenwerte.
- Zu jedem Eigenwert λ von A gibt es genau $(n - \text{Rg}(A - \lambda E))$ linear unabhängige Eigenvektoren.
- Sind $\lambda_1, \lambda_2, \dots, \lambda_m$ paarweise verschiedene Eigenwerte von A und $\vec{x}_1, \vec{x}_2, \dots, \vec{x}_m$ zugehörige Eigenvektoren, dann sind die Eigenvektoren $\vec{x}_1, \vec{x}_2, \dots, \vec{x}_m$ linear unabhängig.

1.3.6 Beispiel (Mehrfacher Eigenwert)

Gegeben sei die Matrix A ,

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

Das charakteristische Polynom von A ergibt sich zu

$$p(\lambda) = \det(A - \lambda E) = \det \begin{pmatrix} 1-\lambda & 0 & 0 \\ 0 & 1-\lambda & 0 \\ 0 & 0 & 1-\lambda \end{pmatrix} = (1-\lambda)^3 \stackrel{!}{=} 0.$$

Dieses Polynom hat offenbar die dreifache Nullstelle

$$\lambda_1 = \lambda_2 = \lambda_3 = 1.$$

Also ist $\lambda = 1$ ein dreifacher Eigenwert A . Wegen

$$(A - \lambda E)\vec{x} = \begin{pmatrix} 1-1 & 0 & 0 \\ 0 & 1-1 & 0 \\ 0 & 0 & 1-1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

existieren zu $\lambda = 1$ genau 3 linear unabhängige Eigenvektoren, z. B.

$$\vec{x}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \vec{x}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \vec{x}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Dieses Beispiel zeigt, dass zu einem k -fachen Eigenwert einer Matrix (bzw. einer k -fachen Nullstelle des charakteristischen Polynoms) k linear unabhängige Eigenvektoren existieren **können**.

1.3.7 Beispiel (Mehrfacher Eigenwert)

Gegeben sei die Matrix A ,

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

Das charakteristische Polynom von A ergibt sich zu

$$p(\lambda) = \det(A - \lambda E) = \det \begin{pmatrix} 1 - \lambda & 1 \\ 0 & 1 - \lambda \end{pmatrix} = (1 - \lambda)^2 \stackrel{!}{=} 0$$

Dieses Polynom hat offenbar die doppelte Nullstelle

$$\lambda_1 = \lambda_2 = 1.$$

Also ist $\lambda_1 = \lambda_2 = 1$ ein doppelter Eigenwert von A . Wegen

$$(A - \lambda_{1/2}E)\vec{x} = \vec{0} \Leftrightarrow \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

ergeben sich alle zu λ gehörenden Eigenvektoren zu

$$\vec{x} = \begin{pmatrix} \alpha \\ 0 \end{pmatrix} = \alpha \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \alpha \in \mathbb{R} \setminus \{0\}.$$

Insbesondere gibt es also keine zwei linear unabhängigen Eigenvektoren zum doppelten Eigenwert $\lambda_1 = \lambda_2 = 1$.

Dieses Beispiel zeigt, dass zu einem k -fachen Eigenwert einer Matrix A (bzw. einer k -fachen Nullstelle des charakteristischen Polynoms) durchaus weniger als k linear unabhängige Eigenvektoren existieren können.

1.3.3 Diagonalähnliche Matrizen

1.3.8 Definition (Ähnliche Matrizen)

Zwei $n \times n$ Matrizen A und B heißen *ähnlich*, falls es eine reguläre $n \times n$ -Matrix C gibt, so dass

$$B = C^{-1}AC$$

gilt. Die durch die reguläre $n \times n$ -Matrix C gegebene Transformation T_C ,

$$T_C : A \mapsto C^{-1}AC,$$

heißt *Ähnlichkeitstransformation*, und die Matrix C wird *Transformationsmatrix* genannt.

1.3.9 Satz (Eigenwerte, Eigenvektoren ähnlicher Matrizen)

Es seien A und B zwei ähnliche $n \times n$ Matrizen und C die zugehörige reguläre $n \times n$ Transformationsmatrix. Ferner sei λ ein Eigenwert von A und \vec{x} ein zugehöriger Eigenvektor, d. h. es gelte $A\vec{x} = \lambda\vec{x}$. Dann ist λ auch ein Eigenwert von B und $\vec{y} = C^{-1}\vec{x}$ ein zugehöriger Eigenvektor.

Beweis. Mit $B = C^{-1}AC$ und $y = C^{-1}\vec{x}$ gilt:

$$B\vec{y} = (C^{-1}AC)(C^{-1}\vec{x}) = C^{-1}A(\underbrace{CC^{-1}}_E)\vec{x} = C^{-1}\underbrace{A\vec{x}}_{\lambda\vec{x}} = C^{-1}\lambda\vec{x} = \lambda C^{-1}\vec{x} = \lambda\vec{y}.$$

□

1.3.10 Definition (Diagonalmatrix)

Eine $n \times n$ -Matrix $D = (d_{ij})$ heißt *Diagonalmatrix*, falls $d_{ij} = 0$ für $i \neq j$ gilt, kurz:

$$D = \text{diag}(d_{11}, d_{22}, \dots, d_{nn}).$$

1.3.11 Definition (Diagonalähnliche Matrizen)

Eine $n \times n$ -Matrix heißt *diagonalähnlich* (oder *diagonalisierbar*), falls es eine Diagonalmatrix $D = \text{diag}(d_{11}, d_{22}, \dots, d_{nn})$ und eine reguläre $n \times n$ -Matrix C gibt, so dass

$$D = C^{-1}AC$$

gilt, also A und D ähnlich sind.

1.3.12 Satz (Charakterisierung diagonalähnlicher Matrizen)

Eine $n \times n$ -Matrix A ist genau dann diagonalähnlich, wenn sie n linear unabhängige Eigenvektoren besitzt. Für die Transformationsmatrix C gilt dann

$$C = (\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n),$$

wobei die \vec{x}_i die n linear unabhängigen Eigenvektoren bezeichnen und für die Diagonalmatrix gilt

$$D = \text{diag}(\lambda_1, \dots, \lambda_n).$$

1.3.13 Beispiel

Wir prüfen, ob die Matrix

$$A = \begin{pmatrix} \frac{52}{25} & \frac{36}{25} & 0 \\ \frac{36}{25} & \frac{73}{25} & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

diagonalisierbar ist, und berechnen ihre Eigenwerte und Eigenvektoren. Wenn möglich bestimmen wir eine reguläre Matrix C mit $C^{-1}AC = \text{diag}(d_{11}, d_{22}, d_{33})$.

$$\begin{aligned} \det(A - \lambda E) &= \begin{vmatrix} \frac{52}{25} - \lambda & \frac{36}{25} & 0 \\ \frac{36}{25} & \frac{73}{25} - \lambda & 0 \\ 0 & 0 & 3 - \lambda \end{vmatrix} \\ &= -\lambda^3 + 8\lambda^2 - 19\lambda + 12 \end{aligned}$$

Die Nullstellen des charakteristischen Polynoms, d. h. die Eigenwerte von A sind $\lambda_1 = 1, \lambda_2 = 4, \lambda_3 = 3$.

Bestimmung zugehöriger Eigenvektoren:

$$\lambda_1 = 1: (A - E)\vec{x} = \vec{0}$$

z. B. mit dem Gauß-Algorithmus berechnet man

$$\vec{x}_1 = \alpha \begin{pmatrix} -\frac{4}{3} \\ 1 \\ 0 \end{pmatrix} \text{ bzw. } \vec{v}_1 = \begin{pmatrix} -\frac{4}{5} \\ \frac{3}{5} \\ 0 \end{pmatrix}$$

als zugehörigen normierten Eigenvektor.

$$\lambda_2 = 4: (A - 4E)\vec{x} = \vec{0}$$

z. B. mit dem Gauß-Algorithmus berechnet man

$$\vec{x}_2 = \beta \begin{pmatrix} \frac{3}{4} \\ 1 \\ 0 \end{pmatrix} \text{ bzw. } \vec{v}_2 = \begin{pmatrix} \frac{3}{5} \\ \frac{4}{5} \\ 0 \end{pmatrix}$$

als zugehörigen normierten Eigenvektor.

$$\lambda_3 = 3: (A - 3E)\vec{x} = \vec{0}$$

z. B. mit dem Gauß-Algorithmus berechnet man

$$\vec{x}_3 = \gamma \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \text{ bzw. } \vec{v}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

als zugehörigen normierten Eigenvektor.

Die Eigenvektoren sind linear unabhängig, da die aus den Eigenvektoren gebildete Determinante von Null verschieden ist.

$$\begin{vmatrix} -\frac{4}{5} & \frac{3}{5} & 0 \\ \frac{3}{5} & \frac{4}{5} & 0 \\ 0 & 0 & 1 \end{vmatrix} = -1$$

Damit ist A diagonalisierbar. Die Transformationsmatrix C ist

$$C = \begin{pmatrix} -\frac{4}{5} & \frac{3}{5} & 0 \\ \frac{3}{5} & \frac{4}{5} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

In diesem speziellen Fall gilt $C^{-1} = C$ und

$$C^{-1}AC = D = \text{diag}(1, 4, 3).$$

1.3.14 Satz (Eigenwerte, Eigenvektoren hermitescher Matrizen)

Es sei eine hermitesche $n \times n$ -Matrix. Dann gilt:

- Jeder Eigenwert λ von A ist reell.
- Zwei zu verschiedenen Eigenwerten λ und μ von A gehörende Eigenvektoren \vec{x} und \vec{y} sind unitär, d. h. für das (komplexe) Skalarprodukt gilt:

$$(\vec{x}^*)^T \vec{y} = \sum_{i=1}^n x_i^* y_i = 0.$$

- A ist diagonalähnlich.

1.3.15 Beispiel

Wir betrachten erneut die Matrix A ,

$$A = \begin{pmatrix} \frac{52}{25} & \frac{36}{25} & 0 \\ \frac{36}{25} & \frac{73}{25} & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$

Die Matrix ist reell und symmetrisch (im reellen Fall gleichbedeutend mit hermitesch). Aus Beispiel 1.3.13 wissen wir bereits, dass die Eigenwerte reell sind, und A diagonalisierbar ist. Für die Eigenvektoren gilt:

$$(\vec{v}_1^*)^T \vec{v}_2 = (\vec{v}_1^*)^T \vec{v}_3 = (\vec{v}_2^*)^T \vec{v}_3 = \vec{0},$$

- die Eigenvektoren sind unitär.

1.3.16 Definition (Positive (Semi)Definitheit)

Eine reelle symmetrische $n \times n$ -Matrix A heißt *positiv definit* (*positiv semidefinit*), falls für alle reellen Vektoren $\vec{x} \in \mathbb{R}^n \setminus \{\vec{0}\}$ gilt:

$$\vec{x}^T A \vec{x} > 0 \quad (\vec{x}^T A \vec{x} \geq 0).$$

Eine hermitesche $n \times n$ -Matrix A heißt *(komplex) positiv definit*, (*(komplex) positiv semidefinit*), falls für alle $\vec{x} \in \mathbb{C}^n \setminus \{\vec{0}\}$ gilt:

$$(\vec{x}^*)^T A \vec{x} > 0 \quad ((\vec{x}^*)^T A \vec{x} \geq 0).$$

1.3.17 Beispiel

(1) Gegeben sei die Matrix A ,

$$A = \begin{pmatrix} 2 & 0 \\ 0 & a \end{pmatrix}, a \in \mathbb{R}.$$

Für jeden beliebigen reellen Vektor $\vec{x} \neq (0, 0)^T$ gilt

$$\begin{aligned} \vec{x}^T A \vec{x} &= (x_1, x_2) \begin{pmatrix} 2 & 0 \\ 0 & a \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = (x_1, x_2) \begin{pmatrix} 2x_1 \\ ax_2 \end{pmatrix} \\ &= 2x_1^2 + ax_2^2 \begin{cases} > 0 \text{ falls } a > 0, \\ \geq 0 \text{ falls } a = 0. \end{cases} \end{aligned}$$

Also ist A positiv definit für $a > 0$ und positiv semidefinit für $a=0$.

(2) Es sei A wie oben und $\vec{x} \neq (0, 0)^T$ ein beliebiger komplexer Vektor. Dann gilt

$$\begin{aligned} (\vec{x}^*)^T A \vec{x} &= (x_1^*, x_2^*) \begin{pmatrix} 2 & 0 \\ 0 & a \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = (x_1^*, x_2^*) \begin{pmatrix} 2x_1 \\ ax_2 \end{pmatrix} \\ &= 2x_1^* x_1 + ax_2^* x_2 = 2|x_1|^2 + a|x_2|^2 \begin{cases} > 0 \text{ falls } a > 0, \\ \geq 0 \text{ falls } a = 0. \end{cases} \end{aligned}$$

Also ist A auch im komplexen erweiterten Sinne positiv definit für $a > 0$ und positiv semidefinit für $a = 0$.

An dieser Stelle lässt sich eine Verbindung zur Analysis für Funktionen in mehreren Variablen herstellen. Man kann zeigen, dass eine hinreichend glatte Funktion $f : \mathbb{R}^n \rightarrow \mathbb{R}$ an einer Stelle $\vec{x} \in \mathbb{R}^n$ ein lokales Minimum (Maximum) besitzt, wenn gilt:

- $\text{grad } f(\vec{x}) = \vec{0}$
- Die Hesse-Matrix an der Stelle \vec{x} ist positiv definit (negativ definit).

Dabei bezeichnet die Hesse-Matrix die Matrix mit den partiellen Ableitungen zweiter Ordnung

$$\begin{pmatrix} \frac{\partial^2 f(\vec{x})}{\partial x_1^2} & \frac{\partial^2 f(\vec{x})}{\partial x_2 \partial x_1} & \cdots & \frac{\partial^2 f(\vec{x})}{\partial x_n \partial x_1} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f(\vec{x})}{\partial x_1 \partial x_n} & \frac{\partial^2 f(\vec{x})}{\partial x_2 \partial x_n} & \cdots & \frac{\partial^2 f(\vec{x})}{\partial x_n^2} \end{pmatrix}$$

1.3.18 Satz (Charakterisierung positiv (semi)definiter Matrizen)

- Eine reelle symmetrische $n \times n$ -Matrix A ist genau dann positiv (semi)definit, wenn alle Eigenwerte von A positiv (nicht negativ) sind.
Eine hermitesche $n \times n$ -Matrix A ist genau dann positiv (semi)definit, wenn alle Eigenwerte von A positiv (nicht negativ) sind.

b) Sei A eine reelle symmetrische $n \times n$ -Matrix. Die k -te Abschnittsdeterminante, $k = 1, 2, \dots, n$ sei gegeben durch

$$\Delta_k = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ \vdots & \vdots & & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kk} \end{vmatrix}.$$

Dann gilt: A ist genau dann positiv definit, wenn $\Delta_k > 0$ für alle $k = 1, 2, \dots, n$. A ist genau dann negativ definit, wenn $(-1)^k \Delta_k > 0$ für alle $k = 1, 2, \dots, n$.

1.3.19 Beispiel

Wir zeigen, dass die Matrix A ,

$$A = \begin{pmatrix} 2 & \sqrt{2} + j \\ \sqrt{2} - j & 4 \end{pmatrix},$$

(komplex) positiv definit ist.

$$\begin{aligned} \det(A - \lambda E) &= \begin{vmatrix} 2 - \lambda & \sqrt{2} + j \\ \sqrt{2} - j & 4 - \lambda \end{vmatrix} \\ &= (2 - \lambda)(4 - \lambda) - (\sqrt{2} + j)(\sqrt{2} - j) \\ &= \lambda^2 - 6\lambda + 5 \\ \lambda^2 - 6\lambda + 5 = 0 &\Leftrightarrow \lambda_1 = 1 \text{ oder } \lambda_2 = 5. \end{aligned}$$

Da die Matrix A hermitesch ist und alle Eigenwerte positiv, ist A positiv definit.

Im folgenden Satz geben wir ein einfach zu überprüfendes, hinreichendes Kriterium für positive (Semi)Definitheiten.

1.3.20 Satz ((Schwach) Zeilensummenkriterium)

Es sei $A \in \mathbb{C}^{n \times n}$ eine beliebige hermitesche $n \times n$ -Matrix.

Wenn A das sogenannte *Zeilensummenkriterium*

$$|a_{ii}| > \sum_{\substack{k=1 \\ k \neq i}}^n |a_{ik}|, \text{ für } 1 \leq i \leq n,$$

erfüllt und alle Diagonalelemente positiv sind, dann sind alle Eigenwerte von A positiv. Matrizen, die das Zeilensummenkriterium erfüllen, nennt man auch *diagonaldominante Matrizen*. Diagonaldominante Matrizen mit positiven Diagonalelementen sind also insbesondere positiv definit.

Wenn A das sogenannte *schwache Zeilensummenkriterium*

$$|a_{ii}| \geq \sum_{\substack{k=1 \\ k \neq i}}^n |a_{ik}|, \text{ für } 1 \leq i \leq n,$$

erfüllt und alle Diagonalelemente nichtnegativ sind, dann sind alle Eigenwerte von A nicht negativ. Matrizen, die das schwache Zeilensummenkriterium erfüllen, nennt man auch *schwach diagonaldominante Matrizen*. Schwach diagonaldominante Matrizen mit nichtnegativen Diagonalelementen sind also insbesondere positiv semidefinit.

1.3.21 Beispiel

Die Matrix A aus Beispiel 1.3.19 erfüllt offensichtlich das starke Zeilensummenkriterium, woraus die positive Definitheit unmittelbar folgt.

Wie das folgende einfache Beispiel zeigt, ist das starke Zeilensummenkriterium nur hinreichend.

Sei

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

Es gilt

$$\begin{aligned}\det(A - \lambda E) &= (1 - \lambda)(2 - \lambda) - 1 \\ &= \lambda^2 - 3\lambda + 1\end{aligned}$$

und

$$\begin{aligned}\lambda^2 - 3\lambda + 1 &= 0 \\ \Leftrightarrow \lambda_1 &= \frac{3}{2} + \frac{1}{2}\sqrt{5}, \lambda_2 = \frac{3}{2} - \frac{1}{2}\sqrt{5}\end{aligned}$$

Die Matrix ist also positiv definit. Das starke Zeilensummenkriterium ist offensichtlich nicht erfüllt.

1.3.22 Satz (Eigenschaften unitärer Matrizen)

Es sei A eine unitäre $n \times n$ -Matrix.

- Jeder Eigenwert λ von A hat den Betrag 1, $|\lambda| = 1$.
- Zwei zu verschiedenen Eigenwerten λ und μ von A gehörende Eigenvektoren \vec{x} und \vec{y} sind unitär, d. h. für das (komplexe) Skalarprodukt gilt:

$$(\vec{x}^*)^T \vec{y} = \sum_{i=1}^n x_i^* y_i = 0$$

- A ist diagonalähnlich.

Der folgende Satz zeigt einige Besonderheiten orthogonaler Matrizen auf.

1.3.23 Satz (Längen und Winkelinvarianz)

Es sei $A \in \mathbb{R}^{n \times n}$ eine orthogonale Matrix und $\vec{x} \in \mathbb{R}^n$ beliebig. Dann gilt:

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

Ist ferner $\vec{x} \neq \vec{0}$ und $\vec{y} \in \mathbb{R}^n \setminus \{\vec{0}\}$ ein weiterer Vektor, so gilt:

$$\sphericalangle(\vec{x}, \vec{y}) = \sphericalangle(A\vec{x}, A\vec{y}).$$

Beweis. Da $A^T A = E$, ergibt sich die Längeninvarianz wie folgt:

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

Die Winkelinvarianz ergibt sich entsprechend:

$$\cos \sphericalangle(A\vec{x}, A\vec{y}) = \frac{(A\vec{x})^T \cdot A\vec{y}}{|A\vec{x}||A\vec{y}|} = \frac{\vec{x}^T (A^T \cdot A) \vec{y}}{|\vec{x}||\vec{y}|} = \frac{\vec{x}^T \vec{y}}{|\vec{x}||\vec{y}|} = \cos \sphericalangle(\vec{x}, \vec{y}).$$

□

1.3.24 Beispiel

Gegeben sei die Matrix

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

- Die Matrix A ist orthogonal, da

$$A \cdot A^T = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = E.$$

b) Wir bestätigen die Eigenschaft der Längeninvarianz. Es ist für $\vec{x} \in \mathbb{R}^2$:

$$A\vec{x} = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}}x_1 - \frac{1}{\sqrt{2}}x_2 \\ \frac{1}{\sqrt{2}}x_1 + \frac{1}{\sqrt{2}}x_2 \end{pmatrix}$$

$$|A\vec{x}| = \left\{ \left(\frac{1}{\sqrt{2}}x_1 - \frac{1}{\sqrt{2}}x_2 \right)^2 + \left(\frac{1}{\sqrt{2}}x_1 + \frac{1}{\sqrt{2}}x_2 \right)^2 \right\}^{1/2} = \{x_1^2 + x_2^2\}^{1/2} = |\vec{x}|.$$

c) Für die gegebene Matrix A rechnen wir die Winkelinvarianz noch einmal explizit nach.
Es gilt für $\vec{x}, \vec{y} \in \mathbb{R}^2 \setminus \{\vec{0}\}$:

$$\cos \sphericalangle(\vec{x}, \vec{y}) = \frac{\langle \vec{x}, \vec{y} \rangle}{|\vec{x}||\vec{y}|} = \frac{x_1y_1 + x_2y_2}{\sqrt{x_1^2 + x_2^2} \cdot \sqrt{y_1^2 + y_2^2}}$$

$$\cos \sphericalangle(A\vec{x}, A\vec{y}) = \frac{\langle A\vec{x}, A\vec{y} \rangle}{|A\vec{x}||A\vec{y}|}$$

$$= \frac{(\frac{1}{2}x_1y_1 - \frac{1}{2}x_1y_2 - \frac{1}{2}x_2y_1 + \frac{1}{2}x_2y_2) + (\frac{1}{2}x_1y_1 + \frac{1}{2}x_1y_2 + \frac{1}{2}x_2y_1 + \frac{1}{2}x_2y_2)}{\sqrt{\frac{1}{2}(x_1 - x_2)^2 + \frac{1}{2}(x_1 + x_2)^2} \sqrt{\frac{1}{2}(y_1 - y_2)^2 + \frac{1}{2}(y_1 + y_2)^2}}$$

$$= \frac{x_1y_1 + x_2y_2}{\sqrt{x_1^2 + x_2^2} \sqrt{y_1^2 + y_2^2}}$$

1.3.25 Definition

Eine Matrix $A \in \mathbb{C}^{n \times n}$ heißt *normal*, falls $\bar{A}A = A\bar{A}$ gilt.

1.3.26 Satz

- Ist $A \in \mathbb{C}^{n \times n}$ hermitesch oder unitär, dann ist A normal.
- Sei $A \in \mathbb{C}^{n \times n}$ normal. Dann sind zwei zu verschiedenen Eigenwerten λ und μ von A gehörende Eigenvektoren \vec{x} und \vec{y} unitär.
- Ist $A \in \mathbb{C}^{n \times n}$ normal, dann ist A diagonalähnlich.

1.3.27 Beispiel

Wir untersuchen die Matrix

$$A = \begin{pmatrix} 1 & \frac{\sqrt{3}}{2} + \frac{1}{2}j \\ j & 1 \end{pmatrix}.$$

Es gilt

$$A \cdot \bar{A} = \begin{pmatrix} 1 & \frac{\sqrt{3}}{2} + \frac{1}{2}j \\ j & 1 \end{pmatrix} \begin{pmatrix} 1 & -j \\ \frac{\sqrt{3}}{2} - \frac{1}{2}j & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 2 & \frac{\sqrt{3}}{2} - \frac{1}{2}j \\ \frac{\sqrt{3}}{2} + \frac{1}{2}j & 2 \end{pmatrix}$$

$$\bar{A} \cdot A = \begin{pmatrix} 1 & -j \\ \frac{\sqrt{3}}{2} - \frac{1}{2}j & 1 \end{pmatrix} \begin{pmatrix} 1 & \frac{\sqrt{3}}{2} + \frac{1}{2}j \\ j & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 2 & \frac{\sqrt{3}}{2} - \frac{1}{2}j \\ \frac{\sqrt{3}}{2} + \frac{1}{2}j & 2 \end{pmatrix}$$

Somit ist A normal, aber nicht hermitesch!

1.3.28 Bemerkung (Hauptachsentransformation)

Auf Basis des obigen Satzes zusammen mit der Charakterisierung diagonalähnlicher Matrizen (vgl. Satz 1.3.12) lässt sich zeigen, dass normale Matrizen (und damit insbesondere hermitesche und unitäre Matrizen) durch unitäre Ähnlichkeitstransformationen auf Diagonalgestalt gebracht werden können. Diesen Prozess nennt man *Hauptachsentransformation*.

Kapitel 2

Ebene Kurven, Raumkurven und Anwendungen

2.1 Raumkurven

Es empfiehlt sich, das Kapitel Anwendungen der Differentialrechnung aus dem 2. Semester zu wiederholen.

In Verallgemeinerung von ebenen Kurven betrachten wir in diesem Abschnitt Kurven im Raum.

Auch hier kann man sich anschaulich vorstellen, dass eine Kurve im Raum die Bahn eines Massenpunktes beschreibt, der zum Zeitpunkt t den Ortsvektor

$$\vec{r}(t) = \begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix}$$

besitzt.

2.1.1 Definition (Parameterdarstellung)

Die vektorwertige Funktion

$$\vec{r}(t) = \begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix}, a \leq t \leq b,$$

(bzw. die drei Gleichungen $x = x(t)$, $y = y(t)$, $z = z(t)$, $a \leq t \leq b$)

heißt eine Parameterdarstellung der Kurve K

$$K = \{(x(t), y(t), z(t)) : t \in [a, b]\}.$$

t heißt Parameter, $[a, b]$ Parameterintervall.

2.1.2 Beispiel

Die Schraubenlinie lässt sich durch die Parameterdarstellung

$$\vec{r}(t) = \begin{pmatrix} R \cdot \cos(\omega t) \\ R \cdot \sin(\omega t) \\ c \cdot t \end{pmatrix}, t \geq 0,$$

R , ω , c Konstanten beschreiben.

Betrachtet man nur $x(t) = R \cos(\omega t)$, $y(t) = R \sin(\omega t)$, so erhält man einen Kreis mit Radius R in der xy -Ebene. Durch $z(t) = c \cdot t$ ändert sich die "Höhe" in Abhängigkeit von t .

Anwendung: Elektronen, die schief in ein homogenes Magnetfeld eingeschossen werden, bewegen sich auf einer Schraubenlinie um die Feldrichtung (z -Achse). Dann bedeutet ω die Winkelgeschwindigkeit, mit der sich die Elektronen auf einer Kreisbahn bewegen.

Grenzwerte und Ableitungen vektorwertiger Funktionen nach dem Parameter werden analog zu den Definitionen für ebene Kurven komponentenweise erklärt, d. h. ist

$$\vec{r}(t) = \begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix}$$

die Parameterdarstellung einer Kurve in \mathbb{R}^3 , so ist

$$\lim_{t \rightarrow t_0} \vec{r}(t) = \begin{pmatrix} \lim_{t \rightarrow t_0} x(t) \\ \lim_{t \rightarrow t_0} y(t) \\ \lim_{t \rightarrow t_0} z(t) \end{pmatrix}$$

und

$$\dot{\vec{r}}(t) = \frac{d}{dt} \vec{r}(t) = \begin{pmatrix} \dot{x}(t) \\ \dot{y}(t) \\ \dot{z}(t) \end{pmatrix},$$

wobei wir wieder wie üblich die Ableitung nach dem Kurvenparameter mit einem Punkt kennzeichnen.

2.1.3 Bemerkung

- (1) Analog lassen sich unter entsprechenden Differenzierbarkeitsvoraussetzungen an $x(t), y(t), z(t)$ auch höhere Ableitungen bilden.
- (2) Der Vektor $\dot{\vec{r}}(t)$ zeigt in Richtung der Kurventangente.
- (3) Beschreibt z. B. $\vec{r}(t)$ die Bahnkurve eines Massenpunktes in Abhängigkeit von der Zeit t , so ist $\vec{v}(t) = \dot{\vec{r}}(t)$ der Geschwindigkeits- und $\vec{a}(t) = \dot{\vec{v}}(t) = \ddot{\vec{r}}(t)$ der Beschleunigungsvektor.

2.1.4 Beispiel

Für die schraubenlinienförmige Bahnkurve eines Elektrons in einem Magnetfeld (siehe Beispiel 2.1.2) gilt:

$$\vec{v}(t) = \dot{\vec{r}}(t) = \begin{pmatrix} -R\omega \sin(\omega t) \\ R\omega \cos(\omega t) \\ c \end{pmatrix}$$

$$\vec{a}(t) = \dot{\vec{v}}(t) = \ddot{\vec{r}}(t) = \begin{pmatrix} -R\omega^2 \cos(\omega t) \\ -R\omega^2 \sin(\omega t) \\ 0 \end{pmatrix}$$

2.1.5 Definition

Eine Parameterdarstellung $\vec{r}(t)$, $a \leq t \leq b$, einer Kurve im \mathbb{R}^3 heißt ebenso wie im \mathbb{R}^2 regulär, wenn

$$\dot{\vec{r}}(t) \neq \vec{0} \text{ für alle } t \in [a, b].$$

2.1.6 Beispiel

Die schraubenlinienförmige Bahnkurve eines Elektrons in einem Magnetfeld ist überall regulär.

2.1.7 Satz (Rechenregeln)

Seien $\vec{a}, \vec{b}: \mathbb{R} \rightarrow \mathbb{R}^3$; $\varphi: \mathbb{R} \rightarrow \mathbb{R}$ differenzierbar. Dann gilt:

a)

$$\frac{d}{dt}(\vec{a}(t) + \vec{b}(t)) = \dot{\vec{a}}(t) + \dot{\vec{b}}(t)$$

b)

$$\frac{d}{dt} \langle \vec{a}(t), \vec{b}(t) \rangle = \langle \dot{\vec{a}}(t), \vec{b}(t) \rangle + \langle \vec{a}(t), \dot{\vec{b}}(t) \rangle$$

c)

$$\frac{d}{dt} \langle \vec{a}(t) \times \vec{b}(t) \rangle = \dot{\vec{a}}(t) \times \vec{b}(t) + \vec{a}(t) \times \dot{\vec{b}}(t)$$

d)

$$\frac{d}{dt}(\varphi(t) \cdot \vec{a}(t)) = \dot{\varphi}(t)\vec{a}(t) + \varphi(t)\dot{\vec{a}}(t)$$

Beweis: Übungsaufgabe

2.1.8 Definition

(1) Für $\dot{\vec{r}}(t) \neq \vec{0}$ ist

$$\vec{T}(t) = \frac{1}{|\dot{\vec{r}}(t)|} \dot{\vec{r}}(t) = \frac{1}{|\dot{\vec{r}}(t)|} \begin{pmatrix} \dot{x}(t) \\ \dot{y}(t) \\ \dot{z}(t) \end{pmatrix}$$

der Tangenteneinheitsvektor (in positiver Richtung) im Kurvenpunkt $(x(t), y(t), z(t))$.

(2) Für $\dot{\vec{T}}(t) \neq \vec{0}$ ist

$$\vec{N}(t) = \frac{1}{|\dot{\vec{T}}(t)|} \cdot \dot{\vec{T}}(t)$$

der Hauptnormaleneinheitsvektor.

2.1.9 Bemerkung

(1) Eine Parameterdarstellung für die Tangente an die Kurve $\vec{r}(t)$ im Punkt $(x(t_0), y(t_0), z(t_0))$ lautet analog zur Tangentengleichung an eine ebene Kurve

$$\begin{aligned} \vec{T}_{t_0}(\lambda) &= \vec{r}(t_0) + \lambda \dot{\vec{r}}(t_0) \\ &= \begin{pmatrix} x(t_0) \\ y(t_0) \\ z(t_0) \end{pmatrix} + \lambda \begin{pmatrix} \dot{x}(t_0) \\ \dot{y}(t_0) \\ \dot{z}(t_0) \end{pmatrix}, \lambda \in \mathbb{R}. \end{aligned}$$

(2) Ausgehend von der Gleichung

$$\langle \vec{T}(t), \vec{T}(t) \rangle = 1$$

erhalten wir mit Satz 2.1.7 b), indem wir die Gleichung nach dem Parameter differenzieren:

$$\begin{aligned} \frac{d}{dt} \langle \vec{T}(t), \vec{T}(t) \rangle &= \langle \dot{\vec{T}}(t), \vec{T}(t) \rangle + \langle \vec{T}(t), \dot{\vec{T}}(t) \rangle \\ &= 2 \langle \vec{T}(t), \dot{\vec{T}}(t) \rangle \\ &= \frac{d}{dt}(1) \\ &= 0 \end{aligned}$$

Also ist $\langle \vec{T}(t), \dot{\vec{T}}(t) \rangle = 0$, d. h. $\dot{\vec{T}} \perp \vec{T}$ und somit

$$\vec{N} = \frac{1}{|\dot{\vec{T}}|} \dot{\vec{T}} \perp \vec{T}.$$

Der Hauptnormaleneinheitsvektor ist also senkrecht zum Tangenteneinheitsvektor. Er zeigt in Richtung der Krümmung.

2.1.10 Beispiel

Wir betrachten wieder die Schraubenlinie.

$$\begin{aligned} \dot{\vec{r}}(t) &= \begin{pmatrix} -R\omega \sin(\omega t) \\ R\omega \cos(\omega t) \\ c \end{pmatrix} \\ |\dot{\vec{r}}(t)| &= \sqrt{R^2\omega^2(\sin^2(\omega t) + \cos^2(\omega t)) + c^2} \\ &= \sqrt{R^2\omega^2 + c^2} \\ \vec{T}(t) &= \frac{1}{\sqrt{R^2\omega^2 + c^2}} \begin{pmatrix} -R\omega \sin(\omega t) \\ R\omega \cos(\omega t) \\ c \end{pmatrix} \\ \ddot{\vec{r}}(t) &= \begin{pmatrix} -R\omega^2 \cos(\omega t) \\ -R\omega^2 \sin(\omega t) \\ 0 \end{pmatrix} \\ |\ddot{\vec{r}}(t)| &= \sqrt{R^2\omega^4} = R\omega^2 \\ \vec{N}(t) &= \frac{1}{R\omega^2} \begin{pmatrix} -R\omega^2 \cos(\omega t) \\ -R\omega^2 \sin(\omega t) \\ 0 \end{pmatrix} \end{aligned}$$

2.2 Bogenlänge und Krümmung ebener Kurven

2.2.1 Satz

a) Die Länge eines Kurvenstücks mit regulärer Parameterdarstellung $\vec{r}(t) = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix}$, $a \leq t \leq b$ beträgt

$$L = \int_a^b \sqrt{\dot{x}(t)^2 + \dot{y}(t)^2} dt = \int_a^b |\dot{\vec{x}}(t)| dt$$

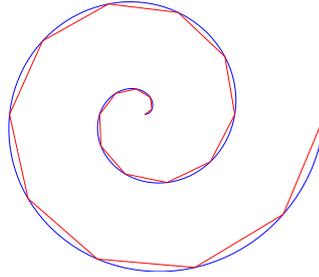
b) Ist $f \in C^1[a, b]$, so hat der Graph von f die Länge

$$L = \int_a^b \sqrt{1 + [f'(x)]^2} dx$$

c) Die Länge eines Kurvenstücks in Polarkoordinatendarstellung $R(\varphi)$, $\alpha \leq \varphi \leq \beta$, beträgt

$$L = \int_\alpha^\beta \sqrt{R(\varphi)^2 + \left(\frac{dR(\varphi)}{d\varphi}\right)^2} d\varphi$$

zu a): Man zerlegt das Parameterintervall $[a, b]$ in n äquidistante Teilintervalle $[t_{i-1}, t_i]$, $i = 1, 2, \dots, n$ mit $a = t_0 < t_1 < \dots < t_n = b$, $t_i - t_{i-1} = \Delta t$.



Die Länge einer Sehne beträgt

$$\Delta s_i = \sqrt{(\Delta x_i)^2 + (\Delta y_i)^2}$$

mit

$$\Delta x_i = x(t_i) - x(t_{i-1}), \quad \Delta y_i = y(t_i) - y(t_{i-1})$$

Nach dem Mittelwertsatz der Differentialrechnung existieren $\xi_i, \eta_i \in [t_{i-1}, t_i]$ mit

$$\Delta x_i = \dot{x}(\xi_i) \Delta t, \quad \Delta y_i = \dot{y}(\eta_i) \Delta t$$

d.h.

$$\Delta s_i = \sqrt{\dot{x}(\xi_i)^2 + \dot{y}(\eta_i)^2} \Delta t$$

Approximation der Kurvenlänge durch die Summe über die Sehnenlängen liefert

$$L \approx \sum_{i=0}^n \sqrt{\dot{x}(\xi_i)^2 + \dot{y}(\eta_i)^2} \Delta t$$

Der Grenzübergang $\Delta t \rightarrow 0$ liefert dann die gewünschte Kurvenlänge

$$L = \int_a^b \sqrt{\dot{x}(t)^2 + \dot{y}(t)^2} dt$$

zu b): folgt aus a), da $x(t) = t, y(t) = f(t)$ reguläre Parametrisierung der Kurve $y = f(x)$ ist, und dann ist

$$\dot{x}(t) = 1, \dot{y}(t) = f'(t).$$

zu c): folgt aus a), da die zugehörige Parameterdarstellung durch $\vec{r}(\varphi) = \begin{pmatrix} R(\varphi) \cos \varphi \\ R(\varphi) \sin \varphi \end{pmatrix}$ gegeben ist. Denn dann ist

$$\dot{x}(\varphi) = \cos \varphi \frac{dR(\varphi)}{d\varphi} - R(\varphi) \sin \varphi, \quad \dot{y}(\varphi) = \sin \varphi \frac{dR(\varphi)}{d\varphi} + R(\varphi) \cos \varphi$$

also

$$\begin{aligned} (\dot{x}(\varphi))^2 + (\dot{y}(\varphi))^2 &= \cos^2 \varphi \left(\frac{dR(\varphi)}{d\varphi} \right)^2 - 2 \sin \varphi \cos \varphi R(\varphi) \frac{dR(\varphi)}{d\varphi} + \sin^2 \varphi (R(\varphi))^2 \\ &\quad + \sin^2 \varphi \left(\frac{dR(\varphi)}{d\varphi} \right)^2 + 2 \sin \varphi \cos \varphi R(\varphi) \frac{dR(\varphi)}{d\varphi} + \cos^2 \varphi (R(\varphi))^2 \\ &= (R(\varphi))^2 (\sin^2 \varphi + \cos^2 \varphi) + \left(\frac{dR(\varphi)}{d\varphi} \right)^2 (\sin^2 \varphi + \cos^2 \varphi) \\ &= (R(\varphi))^2 + \left(\frac{dR(\varphi)}{d\varphi} \right)^2. \end{aligned}$$

2.2.2 Beispiel

a) Cardioide: Sei

$$x(t) = 2 \sin t - \sin(2t), \quad y(t) = 2 \cos t - \cos(2t), \quad 0 \leq t \leq 2\pi.$$

Dann gilt

$$\dot{x}(t) = 2 \cos t - 2 \cos(2t), \quad \dot{y}(t) = -2 \sin t + 2 \sin(2t)$$

und somit

$$\begin{aligned} (\dot{x}(t))^2 + (\dot{y}(t))^2 &= 4 (\cos^2 t - 2 \cos t \cos(2t) + \cos^2(2t)) + 4 (\sin^2 t - 2 \sin t \sin(2t) + \sin^2(2t)) \\ &= 8 - 8 \cos t \cos(2t) - 8 \sin t \sin(2t) \\ &\quad \text{(Trig. Pythagoras: } \sin^2 \alpha + \cos^2 \alpha = 1) \\ &= 8 - 8 \cos t (1 - 2 \sin^2 t) - 8 \sin t (2 \sin t \cos t) \\ &\quad \text{(Additionstheoreme für } \sin(2t), \cos(2t)) \\ &= 8 - 8 \cos t \\ &= 16(1 - \cos^2(t/2)) \\ &\quad \text{(Additionstheorem: } \cos(t) = 2 \cos^2(t/2) - 1). \end{aligned}$$

Somit ist die Bogenlänge

$$\begin{aligned} L &= \int_0^{2\pi} \sqrt{\dot{x}(t)^2 + \dot{y}(t)^2} dt \\ &= 4 \int_0^{2\pi} \sqrt{1 - \cos^2(t/2)} dt \\ &= 4 \int_0^{2\pi} |\sin(t/2)| dt \\ &= 4 \int_0^{2\pi} \sin(t/2) dt \\ &= -8 \cos(t/2) \Big|_0^{2\pi} \\ &= 16 \end{aligned}$$

b) Der Bogen der Normalparabel $y = x^2$ über dem Intervall $[0, x_0]$, $x_0 > 0$, hat die Länge

$$L = \int_0^{x_0} \sqrt{1 + (2x)^2} dx = \int_0^{x_0} \sqrt{1 + 4x^2} dx$$

Berechnung der Stammfunktion von $\sqrt{1 + 4x^2}$.

Nötige Substitutionen:

$$t = 2x, \quad x = t/2, \quad dt = 2dx$$

$$v = \operatorname{arsinh} t = \ln(t + \sqrt{t^2 + 1}),$$

$$t = \sinh v, \quad \sqrt{t^2 + 1} = \cosh v$$

$$dt = \cosh v dv$$

$$\begin{aligned} \int \sqrt{1 + 4x^2} dx &= \frac{1}{2} \int \sqrt{1 + t^2} dt \\ &= \frac{1}{2} \int (\cosh v)^2 dv \\ &= \frac{1}{2} \frac{1}{4} \int (e^v + e^{-v})^2 dv \\ &= \frac{1}{8} \int (e^{2v} + 2 + e^{-2v}) dv \\ &= \frac{1}{8} \left(\frac{1}{2} e^{2v} + 2v - \frac{1}{2} e^{-2v} \right) + C \\ &= \frac{1}{4} \frac{1}{2} (e^v + e^{-v})(e^v - e^{-v}) \frac{1}{2} + \frac{1}{4} v + C \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{4} \cosh v \sinh v + \frac{1}{4} v + C \\
&= \frac{1}{4} t \sqrt{1+t^2} + \frac{1}{4} \ln(t + \sqrt{t^2+1}) + C \\
&= \frac{1}{2} x \sqrt{1+4x^2} + \frac{1}{4} \ln(2x + \sqrt{1+4x^2}) + C
\end{aligned}$$

Damit folgt

$$\begin{aligned}
L &= \left. \frac{1}{2} x \sqrt{1+4x^2} + \ln(2x + \sqrt{1+4x^2}) \right|_0^{x_0} \\
&= \frac{1}{2} x_0 \sqrt{1+4x_0^2} + \ln(2x_0 + \sqrt{1+4x_0^2})
\end{aligned}$$

- c) Archimedische Spirale mit $R(\varphi) = \varphi$, $\frac{dR(\varphi)}{d\varphi} = 1$. Die Bogenlänge nach k Umläufen, d.h. $0 \leq \varphi \leq k \cdot 2\pi$ beträgt

$$\begin{aligned}
L &= \int_0^{k \cdot 2\pi} \sqrt{\varphi^2 + 1} d\varphi \\
&= \left. \frac{1}{2} \varphi \sqrt{1+\varphi^2} + \frac{1}{2} \ln(\varphi + \sqrt{1+\varphi^2}) \right|_0^{2k\pi} \\
&= k\pi \sqrt{1+(2k\pi)^2} + \frac{1}{2} \ln(2k\pi + \sqrt{1+(2k\pi)^2})
\end{aligned}$$

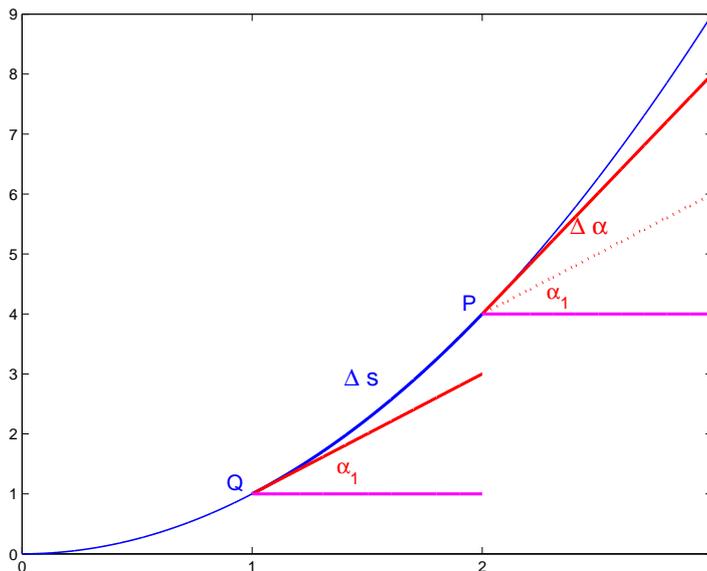
Anschaulich: Fährt man einen Weg entlang, so merkt man am Einschlag des Lenkrads, ob eine Kurve stärker oder schwächer gekrümmt ist, und ob es sich um eine Links- oder Rechtskurve handelt.

2.2.3 Definition

Die Krümmung κ in einem Kurvenpunkt (x, y) ist die Änderung des Neigungswinkels der Tangente bezogen auf die Bogenlänge des zugehörigen Kurvenstücks:

$$\kappa = \frac{d\alpha}{ds} = \lim_{\Delta s \rightarrow 0} \frac{\Delta \alpha}{\Delta s},$$

(wenn der Grenzwert existiert).



2.2.4 Satz (Krümmung bei Parameterdarstellung)

Für eine Kurve in Parameterdarstellung gilt für die Krümmung im Punkt $(x(t), y(t))$:

$$\kappa = \kappa(t) = \frac{\dot{x}(t)\ddot{y}(t) - \dot{y}(t)\ddot{x}(t)}{(\dot{x}(t)^2 + \dot{y}(t)^2)^{3/2}}$$

Zum *Beweis* berechnen wir $\frac{d\alpha}{ds}$ als $\frac{d\alpha}{dt} \cdot \frac{dt}{ds}$. Dabei ist $\alpha = \arctan(\dot{y}(t)/\dot{x}(t))$, also

$$\frac{d\alpha}{dt} = \frac{1}{1 + (\dot{y}(t)/\dot{x}(t))^2} \cdot \frac{\dot{x}(t)\ddot{y}(t) - \dot{y}(t)\ddot{x}(t)}{\dot{x}(t)^2} = \frac{\dot{x}(t)\ddot{y}(t) - \dot{y}(t)\ddot{x}(t)}{\dot{x}(t)^2 + \dot{y}(t)^2}.$$

Für die Bogenlänge ist $s(t) = \int_a^t \sqrt{\dot{x}(u)^2 + \dot{y}(u)^2} du$, also nach dem Hauptsatz der Differential- und Integralrechnung

$$\begin{aligned} \frac{ds}{dt} &= \sqrt{\dot{x}(t)^2 + \dot{y}(t)^2} \\ \implies \frac{dt}{ds} &= \frac{1}{\sqrt{\dot{x}(t)^2 + \dot{y}(t)^2}}. \end{aligned}$$

Damit ergibt sich die behauptete Formel für $\frac{d\alpha}{ds} = \frac{d\alpha}{dt} \cdot \frac{dt}{ds}$.

2.2.5 Definition

Punkte einer Kurve, in denen die Krümmung ein relatives Extremum hat, heißen Scheitel.

2.2.6 Beispiel (Ellipse)

Für die Parametrisierung $x(t) = a \cos t$, $y(t) = b \sin t$ gilt

$$\begin{aligned} \dot{x}(t) &= -a \sin t, & \dot{y}(t) &= b \cos t \\ \ddot{x}(t) &= -a \cos t, & \ddot{y}(t) &= -b \sin t \end{aligned}$$

Damit gilt für die Krümmung

$$\begin{aligned} \kappa(t) &= \frac{(-a \sin t)(-b \sin t) - b \cos t(-a \cos t)}{(a^2 \sin^2 t + b^2 \cos^2 t)^{3/2}} \\ &= \frac{ab(\sin^2 t + \cos^2 t)}{(a^2 \sin^2 t + b^2 \cos^2 t)^{3/2}} \\ &= \frac{ab}{(a^2 \sin^2 t + b^2 \cos^2 t)^{3/2}} \end{aligned}$$

Ist speziell $a = b = R$ (Kreis mit Radius R), dann ist

$$\kappa(t) = \frac{R^2}{(R^2 \sin^2 t + R^2 \cos^2 t)^{3/2}} = \frac{1}{R}.$$

In einem Kreis ist die Krümmung also in jedem Punkt gleich groß. Mit wachsendem Radius wird die Krümmung kleiner. Beides entspricht der Anschauung.

Wir berechnen noch die Scheitel der Ellipse:

$$\begin{aligned} \dot{\kappa}(t) &= \frac{-3ab(a^2 - b^2) \sin t \cos t}{(a^2 \sin^2 t + b^2 \cos^2 t)^{5/2}} \\ \ddot{\kappa}(t) &= \frac{3ab(a^2 - b^2)[a^2 \sin^2 t(1 + 3 \cos^2 t) - b^2 \cos^2 t(1 + 3 \sin^2 t)]}{(a^2 \sin^2 t + b^2 \cos^2 t)^{7/2}} \end{aligned}$$

Da

$$\dot{\kappa}(t) = 0 \iff t = k \cdot \frac{\pi}{2}, \quad k \in \mathbb{Z} \text{ und } \ddot{\kappa}\left(k \cdot \frac{\pi}{2}\right) \neq 0$$

liegen die Scheitel bei $t = k \cdot \frac{\pi}{2}$, $k \in \mathbb{Z}$. Die Krümmungen dort betragen

$$\kappa(k\pi) = \frac{a}{b^2} \quad \kappa\left(\frac{\pi}{2} + k\pi\right) = \frac{b}{a^2}, \quad k \in \mathbb{Z}$$

2.2.7 Satz (Krümmung bei kartesischen Koordinaten)

Sei $y = f(x)$, $f \in C^2[a, b]$. Dann ist die Krümmung im Punkt $(x, f(x))$ des Graphen von f gegeben durch

$$\kappa(x) = \frac{y''}{(1 + (y')^2)^{3/2}}$$

Außerdem gilt:

$\kappa(x) > 0 \Leftrightarrow f''(x) > 0$ -Linkskrümmung (zunehmende α -Werte)

$\kappa(x) < 0 \Leftrightarrow f''(x) < 0$ -Rechtskrümmung (abnehmende α -Werte)

Der *Beweis* ergibt sich aus Satz 2.2.4 unter Verwendung der Parameterdarstellung $x = t, y = f(t)$ für den Graphen von f .

2.2.8 Beispiel (Krümmung der Sinuskurve)

$$\begin{aligned} f(x) &= \sin x, & f'(x) &= \cos x, & f''(x) &= -\sin x \\ \Rightarrow \kappa(x) &= \frac{-\sin x}{(1 + \cos^2 x)^{3/2}} \end{aligned}$$

Für $x \in]0, \pi[$ ist $\kappa(x) < 0$, also Rechtskrümmung, für $x \in]\pi, 2\pi[$ ist $\kappa(x) > 0$, also Linkskrümmung.

Scheitel:

$$\frac{d\kappa}{dx} = \frac{2 \cos x (1 + \sin^2 x)}{(1 + \cos^2 x)^3}$$

Somit ist

$$\frac{d\kappa(x)}{dx} = 0 \iff \cos x = 0 \iff x = (2k + 1)\frac{\pi}{2}, \quad k \in \mathbb{Z}.$$

Da $\frac{d^2\kappa(x)}{dx^2}((2k + 1)\frac{\pi}{2}) \neq 0$, handelt es sich um relative Extremalstellen von $\kappa(x)$. Die Scheitel befinden sich also an den Stellen

$$x = (2k + 1)\frac{\pi}{2}, \quad k \in \mathbb{Z}.$$

2.2.9 Satz (Krümmung bei Polardarstellung)

Für eine Kurve in Polardarstellung $R(\varphi)$ gilt für die Krümmung im Kurvenpunkt $(\varphi, R(\varphi))$:

$$\kappa(\varphi) = \frac{R(\varphi)^2 + 2\dot{R}(\varphi)^2 - R(\varphi)\ddot{R}(\varphi)}{(\dot{R}(\varphi)^2 + R(\varphi)^2)^{3/2}}.$$

Der *Beweis* geht wieder durch Rückführung auf Satz 2.2.4, diesmal mit $t = \varphi$, $x(t) = R(t) \cdot \cos t$, $y(t) = R(t) \cdot \sin t$.

2.2.10 Beispiel (Archimedische Spirale)

Es sei $R(\varphi) = \varphi$, $\varphi \geq 0$, die Polardarstellung der archimedischen Spirale. Dann gilt für die Krümmung

$$\kappa(\varphi) = \frac{\varphi^2 + 2}{(1 + \varphi^2)^{3/2}}$$

Aus

$$\dot{\kappa}(\varphi) = \frac{-\varphi(4 + \varphi^2)}{(1 + \varphi^2)^{5/2}} < 0 \text{ falls } \varphi > 0,$$

folgt, dass die Krümmung mit wachsendem Winkel abnimmt. Es gibt also keine Scheitel.

2.3 Bogenlänge und Krümmung einer Raumkurve

In Verallgemeinerung der entsprechenden Resultate für ebene Kurven halten wir fest:

2.3.1 Satz

Ist $\vec{r}(t)$, $a \leq t \leq b$, die reguläre Darstellung einer Raumkurve, so beträgt die Bogenlänge

$$L = \int_a^b \sqrt{\dot{x}(t)^2 + \dot{y}(t)^2 + \dot{z}(t)^2} dt = \int_a^b |\dot{\vec{r}}(t)| dt.$$

2.3.2 Beispiel

Wir betrachten noch einmal die Schraubenlinie aus Beispiel 2.1.2, d. h.

$$\vec{r}(t) = \begin{pmatrix} R \cos(\omega t) \\ R \sin(\omega t) \\ ct \end{pmatrix}$$

und berechnen die Länge der Kurve nach n Umläufen, d. h. $a = 0$, $b = n \cdot \frac{2\pi}{\omega}$. Nach Beispiel 2.1.10 ist

$$|\dot{\vec{r}}(t)| = \sqrt{R^2\omega^2 + c^2},$$

also gilt:

$$\begin{aligned} L &= \int_0^{n \cdot \frac{2\pi}{\omega}} |\dot{\vec{r}}(t)| dt = \sqrt{R^2\omega^2 + c^2} \int_0^{n \cdot \frac{2\pi}{\omega}} 1 dt \\ &= \frac{n \cdot 2\pi}{\omega} \sqrt{R^2\omega^2 + c^2} \end{aligned}$$

Auch bei einer Raumkurve misst man mit der Krümmung κ die Abweichung von einer Geraden, d. h. die Richtungsänderung der Kurventangente pro Bogenlängenänderung. Die Krümmung ändert sich i. a. in jedem Kurvenpunkt und ist somit eine Funktion des Kurvenparameters. Es gilt:

$$\kappa(t) = \frac{|\dot{\vec{r}}(t) \times \ddot{\vec{r}}(t)|}{|\dot{\vec{r}}(t)|^3}.$$

2.3.3 Bemerkung

Ist $\vec{r}(t)$ speziell von der Gestalt $\vec{r}(t) = \begin{pmatrix} x(t) \\ y(t) \\ 0 \end{pmatrix}$, so gilt

$$\begin{aligned} \kappa(t) &= \frac{|\dot{\vec{r}}(t) \times \ddot{\vec{r}}(t)|}{|\dot{\vec{r}}(t)|^3} \\ &= \frac{\left| \begin{pmatrix} 0 \\ 0 \\ \dot{x}\dot{y} - \dot{y}\dot{x} \end{pmatrix} \right|}{(\dot{x}(t)^2 + \dot{y}(t)^2)^{3/2}} \\ &= \frac{|\dot{x}\dot{y} - \dot{y}\dot{x}|}{(\dot{x}(t)^2 + \dot{y}(t)^2)^{3/2}} \end{aligned}$$

d. h. der Betrag der Krümmung einer ebenen Kurve.

Bei einer ebenen Kurve unterscheidet man noch zwischen Rechts- und Linkskrümmung in Abhängigkeit von κ . Bei einer Raumkurve ist dies nicht möglich.

2.3.4 Beispiel (Krümmung einer Schraubenlinie)

Für

$$\vec{r}(t) = \begin{pmatrix} R \cos(\omega t) \\ R \sin(\omega t) \\ ct \end{pmatrix}$$

gilt

$$\begin{aligned} \dot{\vec{r}}(t) \times \ddot{\vec{r}}(t) &= \begin{pmatrix} -R\omega \sin(\omega t) \\ R\omega \cos(\omega t) \\ c \end{pmatrix} \times \begin{pmatrix} -R\omega^2 \cos(\omega t) \\ -R\omega^2 \sin(\omega t) \\ 0 \end{pmatrix} \\ &= \begin{pmatrix} cR\omega^2 \sin(\omega t) \\ -cR\omega^2 \cos(\omega t) \\ R^2\omega^3 \sin^2(\omega t) + R^2\omega^3 \cos^2(\omega t) \end{pmatrix} \\ &= \begin{pmatrix} cR\omega^2 \sin(\omega t) \\ -cR\omega^2 \cos(\omega t) \\ R^2\omega^3 \end{pmatrix} \\ &= R\omega^2 \begin{pmatrix} c \sin(\omega t) \\ -c \cos(\omega t) \\ R\omega \end{pmatrix} \end{aligned}$$

$$\begin{aligned} |\dot{\vec{r}}(t) \times \ddot{\vec{r}}(t)| &= R\omega^2 \sqrt{c^2 (\underbrace{\sin^2(\omega t) + \cos^2(\omega t)}_{=1}) + R^2\omega^2} \\ &= R\omega^2 \sqrt{R^2\omega^2 + c^2} \end{aligned}$$

Bereits in Beispiel 2.1.10 hatten wir berechnet,

$$|\dot{\vec{r}}(t)| = \sqrt{R^2\omega^2 + c^2}$$

Somit erhalten wir für die Krümmung

$$\begin{aligned} \kappa(t) &= \frac{R\omega^2 \sqrt{R^2\omega^2 + c^2}}{(R^2\omega^2 + c^2) \sqrt{R^2\omega^2 + c^2}} \\ &= \frac{R\omega^2}{R^2\omega^2 + c^2} \end{aligned}$$

Die Schraubenlinie besitzt also für jeden Parameter, d. h. auch für jeden Kurvenpunkt, die gleiche konstante Krümmung.

2.4 Skalar- und Vektorfelder

In Anwendungen unterscheidet man die Begriffe Skalar- und Vektorfelder.

Wird jedem Punkt der Ebene oder des Raumes eine skalare Größe zugeordnet, d. h.

$$\Phi : \mathbb{R}^2 \rightarrow \mathbb{R} \text{ bzw. } \Phi : \mathbb{R}^3 \rightarrow \mathbb{R},$$

so spricht man von einem ebenen bzw. räumlichen Skalarfeld.

Beispiele für Skalarfelder sind:

- Dichteverteilung im Inneren der Erdkugel
- Temperaturverteilung in einem Raum

Skalare Felder können sich z. B. in Abhängigkeit von der Zeit verändern. Wir betrachten hier jedoch nur stationäre, d. h. zeitunabhängige Skalarfelder.

Wird jedem Punkt der Ebene oder des Raumes eine vektorielle Größe zugeordnet, d. h.

$$\vec{F} : \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad \vec{F}(x, y) = \begin{pmatrix} F_1(x, y) \\ F_2(x, y) \end{pmatrix} \text{ bzw.}$$

$$\vec{F} : \mathbb{R}^3 \rightarrow \mathbb{R}^3, \quad \vec{F}(x, y, z) = \begin{pmatrix} F_1(x, y, z) \\ F_2(x, y, z) \\ F_3(x, y, z) \end{pmatrix}$$

so spricht man von einem ebenen bzw. räumlichen Vektorfeld.

Beispiele für Vektorfelder sind

- Magnetfeld in der Umgebung eines stromdurchflossenen Leiters
- Elektrisches Feld in einem geladenen Plattenkondensator
- Strömungsgeschwindigkeiten

Auch hier betrachten wir nur den stationären, d. h. zeitunabhängigen Fall.

2.4.1 Divergenz und Rotation eines Vektorfeldes

Wir erläutern den Begriff der Divergenz eines Vektorfeldes zunächst am Modell einer strömenden Flüssigkeit.

Stellt man sich den Vektor $\vec{v}(x, y, z) = \begin{pmatrix} v_1(x, y, z) \\ v_2(x, y, z) \\ v_3(x, y, z) \end{pmatrix}$ als die pro Zeiteinheit durch eine senkrecht zu \vec{v}

gelegte Fläche strömende Flüssigkeitsmenge vor, so ist $\frac{\partial v_1}{\partial x}$ die Änderung der x -Komponente in x -Richtung, die Summe $\frac{\partial v_1}{\partial x} + \frac{\partial v_2}{\partial y} + \frac{\partial v_3}{\partial z}$ der drei Ableitungen also die Änderung der aus einem Volumenelement herausströmenden Flüssigkeitsmenge gegenüber der einströmenden Flüssigkeitsmenge.

2.4.1 Definition

Die Divergenz eines Vektorfeldes $\vec{F}(x, y, z)$ ist das skalare Feld

$$\operatorname{div} \vec{F} = \frac{\partial F_1}{\partial x} + \frac{\partial F_2}{\partial y} + \frac{\partial F_3}{\partial z}.$$

2.4.2 Bemerkung

Ist $\operatorname{div} \vec{F} > 0$, so enthält das Feld Quellen, ist $\operatorname{div} \vec{F} < 0$, so enthält es Senken. Ein Feld mit $\operatorname{div} \vec{F} = 0$ heißt quellen und senkenfrei.

2.4.3 Satz (Rechenregeln für Divergenz)

Seien \vec{A} und \vec{B} Vektorfelder, Φ ein Skalarfeld, \vec{a} ein konstanter Vektor und c eine Konstante, Dann gilt:

- (1) $\operatorname{div} \vec{a} = 0$
- (2) $\operatorname{div}(\Phi \cdot \vec{A}) = \langle \operatorname{grad} \Phi, \vec{A} \rangle + \Phi \cdot (\operatorname{div} \vec{A})$
- (3) $\operatorname{div}(c \cdot \vec{A}) = c \cdot (\operatorname{div} \vec{A})$
- (4) $\operatorname{div}(\vec{A} + \vec{B}) = \operatorname{div} \vec{A} + \operatorname{div} \vec{B}$
- (5) $\operatorname{div}(\vec{A} + \vec{a}) = \operatorname{div} \vec{A}$

In der Strömung eines Flusses beobachtet man z. B. hinter einem Brückenpfeiler Wirbelbildungen des Wassers. Quer zur Strömungsrichtung treten Geschwindigkeitskomponenten auf, die sich mit der Entfernung vom Brückenpfeiler ändern. Solche Wirbelbildungen, die durch Änderung der Komponenten eines Vektors in den senkrecht zu ihm zeigenden Richtungen entstehen, werden durch die Rotation beschrieben.

2.4.4 Definition

Die Rotation eines Vektorfeldes $\vec{F}(x, y, z)$ ist das Vektorfeld

$$\operatorname{rot}\vec{F} = \begin{pmatrix} \frac{\partial F_3}{\partial y} - \frac{\partial F_2}{\partial z} \\ \frac{\partial F_1}{\partial z} - \frac{\partial F_3}{\partial x} \\ \frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y} \end{pmatrix}.$$

2.4.5 Bemerkung

- (1) Ein Vektorfeld \vec{F} mit $\operatorname{rot}\vec{F} = \vec{0}$ heißt wirbelfrei.
- (2) Führt man für den Vektor der partiellen Ableitungen die Bezeichnung des Nabla-Operators

$$\vec{\nabla} = \begin{pmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \\ \frac{\partial}{\partial z} \end{pmatrix}$$

ein, so kann man die Rotation eines Vektorfeldes \vec{F} auch formal als äußeres Produkt

$$\operatorname{rot}\vec{F} = \vec{\nabla} \times \vec{F}$$

darstellen.

2.4.6 Satz (Rechenregeln für Rotation:)

Seien \vec{A} und \vec{B} Vektorfelder, Φ ein Skalarfeld, \vec{a} ein konstanter Vektor und c eine Konstante, Dann gilt:

- (1) $\operatorname{rot} \vec{a} = \vec{0}$
- (2) $\operatorname{rot}(\Phi \cdot \vec{A}) = (\operatorname{grad}\Phi) \times \vec{A} + \Phi \cdot (\operatorname{rot}\vec{A})$
- (3) $\operatorname{rot}(c \cdot \vec{A}) = c \cdot (\operatorname{rot}\vec{A})$
- (4) $\operatorname{rot}(\vec{A} + \vec{B}) = \operatorname{rot}\vec{A} + \operatorname{rot}\vec{B}$
- (5) $\operatorname{rot}(\vec{A} + \vec{a}) = \operatorname{rot}\vec{A}$

Wir halten noch fest.

2.4.7 Satz (Eigenschaften eines wirbelfreien Vektorfeldes)

- (1) Sei \vec{F} ein wirbelfreies Vektorfeld, d. h. $\operatorname{rot}\vec{F} = \vec{0}$. Dann existiert ein Skalarfeld Φ mit $\vec{F} = \operatorname{grad}\Phi$.
- (2) Ist Φ ein Skalarfeld, so ist der zugehörige Gradient wirbelfrei, d. h. $\operatorname{rot}(\operatorname{grad}\Phi) = \vec{0}$.

2.4.8 Bemerkung

Ist $\operatorname{rot}\vec{F} = \vec{0}$, dann heißt das nach Satz 2.4.7 zugehörige Skalarfeld die zu \vec{F} gehörende Potentialfunktion.

2.4.9 Beispiel

Sei

$$\vec{F}(x, y, z) = (x^2 + y^2 + z^2) \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

Dann gilt

$$\operatorname{rot}\vec{F} = \begin{pmatrix} \frac{\partial F_3}{\partial y} - \frac{\partial F_2}{\partial z} \\ \frac{\partial F_1}{\partial z} - \frac{\partial F_3}{\partial x} \\ \frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y} \end{pmatrix} = \begin{pmatrix} 2yz - 2yz \\ 2xz - 2xz \\ 2xy - 2xy \end{pmatrix} = \vec{0}$$

Demnach besitzt \vec{F} eine Potentialfunktion Φ .

Bestimmung von Φ . Wegen $\text{grad}\Phi = \vec{F}$ muss gelten:

$$\begin{aligned}\frac{\partial\Phi}{\partial x} &= (x^2 + y^2 + z^2) \cdot x \Rightarrow \Phi = \frac{1}{4}x^4 + \frac{1}{2}x^2y^2 + \frac{1}{2}x^2z^2 + \psi_1(y, z) \\ \frac{\partial\Phi}{\partial y} &= (x^2 + y^2 + z^2) \cdot y \Rightarrow \Phi = \frac{1}{2}x^2y^2 + \frac{1}{4}y^4 + \frac{1}{2}y^2z^2 + \psi_2(x, z) \\ \frac{\partial\Phi}{\partial z} &= (x^2 + y^2 + z^2) \cdot z \Rightarrow \Phi = \frac{1}{2}x^2z^2 + \frac{1}{2}y^2z^2 + \frac{1}{4}z^4 + \psi_3(x, y)\end{aligned}$$

Vergleicht man diese drei Darstellungen miteinander, so erhält man

$$\Phi = \frac{1}{4}(x^4 + y^4 + z^4) + \frac{1}{2}(x^2y^2 + x^2z^2 + y^2z^2).$$

2.5 Linien- oder Kurvenintegrale

Linien- oder Kurvenintegrale werden z. B. benutzt, um die physikalische Arbeit, die ein ebenes bzw. räumliches Kraftfeld $\vec{F}(x, y)$ bzw. $\vec{F}(x, y, z)$ an einem Massenpunkt verrichtet, zu beschreiben, wenn dieser unter dem Einfluss des Feldes von einem Punkt P_1 aus längs einer Kurve C in einen Punkt P_2 verschoben wird. Wir beziehen uns hier auf den räumlichen Fall, da der ebene als Spezialfall darin enthalten ist.

2.5.1 Definition

Sei $\vec{F}(x, y, z)$ ein räumliches Vektorfeld, $\vec{r}(t)$ der Ortsvektor einer von P_1 nach P_2 verlaufenden Raumkurve C mit $t_1 \leq t \leq t_2$, und $\dot{\vec{r}}(t)$ der zugehörige Tangentenvektor der Kurve. Dann heißt das Integral

$$\int_C \langle \vec{F}, d\vec{r} \rangle = \int_{t_1}^{t_2} \langle \vec{F}, \dot{\vec{r}} \rangle dt$$

Linien- oder Kurvenintegral des Vektorfeldes $\vec{F}(x, y, z)$ längs der Raumkurve C .

2.5.2 Bemerkung

a) In ausführlicher Schreibweise lautet das Linienintegral:

$$\begin{aligned}& \int_C \{F_1(x, y, z)dx + F_2(x, y, z)dy + F_3(x, y, z)dz\} \\ &= \int_{t_1}^{t_2} \{F_1(x(t), y(t), z(t))\dot{x}(t) + F_2(x(t), y(t), z(t))\dot{y}(t) + F_3(x(t), y(t), z(t))\dot{z}(t)\} dt\end{aligned}$$

b) Im Allgemeinen hängt der Wert eines Linienintegrals sowohl vom Anfangs- und Endpunkt des Integrationsweges als auch vom Verlauf des Verbindungsweges ab.

c) Für ein Linienintegral längs einer geschlossenen Kurve C verwenden wir die Notation

$$\oint_C \langle \vec{F}, d\vec{r} \rangle$$

In physikalischen Anwendungen spricht man von Zirkulation des Vektorfeldes \vec{F} längs der geschlossenen Kurve C .

d) Die konkrete Berechnung eines Linienintegrals erfolgt in 3 Schritten:

- Ersetzung der Koordinaten x, y, z im Feldvektor durch die parameterabhängigen Koordinaten $x(t), y(t), z(t)$ der Raumkurve C .
- Differentiation des Ortsvektor $\vec{r}(t)$ nach dem Parameter t liefert den Tangentenvektor $\dot{\vec{r}}(t)$.

- Das Skalarprodukt $\langle \vec{F}, \dot{\vec{r}} \rangle$ hängt jetzt nur noch vom Parameter t ab und wird nach der Variablen t in den Grenzen von t_1 bis t_2 integriert.

2.5.3 Beispiel

a) Gegeben sei das räumliche Vektorfeld $\vec{F}(x, y, z) = \begin{pmatrix} 2x + y^2 \\ x^2yz \\ x + z \end{pmatrix}$ und die Kurve C , die durch den

Ortsvektor $\vec{r}(t) = \begin{pmatrix} t \\ t^2 \\ t \end{pmatrix}$, $0 \leq t \leq 1$, beschrieben wird. Wir wollen das zugehörige Linienintegral

$$\int_C \langle \vec{F}, d\vec{r} \rangle$$

berechnen.

- Ersetzung der Koordinaten im Feldvektor durch die parameterabhängigen Koordinaten $x(t) = t$, $y(t) = t^2$, $z(t) = t$.

$$\vec{F}(x(t), y(t), z(t)) = \begin{pmatrix} 2t + t^4 \\ t^5 \\ 2t \end{pmatrix}$$

- Bestimmung des Tangentenvektors

$$\dot{\vec{r}}(t) = \begin{pmatrix} 1 \\ 2t \\ 1 \end{pmatrix}$$

- Berechnung des Linienintegrals

$$\begin{aligned} \int_C \langle \vec{F}, d\vec{r} \rangle &= \int_0^1 \left\langle \begin{pmatrix} 2t + t^4 \\ t^5 \\ 2t \end{pmatrix}, \begin{pmatrix} 1 \\ 2t \\ 1 \end{pmatrix} \right\rangle dt \\ &= \int_0^1 (2t + t^4 + 2t^6 + 2t) dt \\ &= \int_0^1 (2t^6 + t^4 + 4t) dt \\ &= \left[\frac{2}{7}t^7 + \frac{1}{5}t^5 + 2t^2 \right]_0^1 \\ &= \frac{87}{35} \end{aligned}$$

b) Wir berechnen das Linienintegral für dasselbe Vektorfeld und die Kurve \tilde{C} , die durch

$\vec{r}(t) = \begin{pmatrix} t \\ t \\ t \end{pmatrix}$, $0 \leq t \leq 1$, gegeben ist.

- Ersetzung der Koordinaten im Feldvektor durch die parameterabhängigen Koordinaten $x(t) = t$, $y(t) = t$, $z(t) = t$.

$$\vec{F}(x(t), y(t), z(t)) = \begin{pmatrix} 2t + t^2 \\ t^4 \\ 2t \end{pmatrix}$$

- Bestimmung des Tangentenvektors

$$\dot{\vec{r}}(t) = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

- Berechnung des Linienintegrals

$$\begin{aligned} \int_{\tilde{C}} \langle \vec{F}, d\vec{r} \rangle &= \int_0^1 \left\langle \begin{pmatrix} 2t + t^2 \\ t^4 \\ 2t \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\rangle dt \\ &= \int_0^1 (t^4 + t^2 + 4t) dt \\ &= \left[\frac{1}{5}t^5 + \frac{1}{3}t^3 + 2t^2 \right]_0^1 \\ &= \frac{38}{15} \end{aligned}$$

2.5.4 Bemerkung

In dem Beispiel haben wir gesehen, dass wir, obwohl Anfangs- und Endpunkt der Kurven C und \tilde{C} übereinstimmen, für die verschiedenen Verbindungswege unterschiedliche Werte für die Linienintegrale erhalten. Unter gewissen Voraussetzungen, die wir im Folgenden untersuchen wollen, hängt der Wert eines Linienintegrals nur vom Anfangs- und Endpunkt, nicht aber vom eingeschlagenen Verbindungsweg ab. Dies führt zunächst auf folgende Definition.

2.5.5 Definition

Ein Vektorfeld \vec{F} heißt konservativ oder Potentialfeld, wenn das Linienintegral $\int_C \langle \vec{F}, d\vec{r} \rangle$ nur vom Anfangs- und Endpunkt, nicht aber vom Verbindungsweg C der beiden Punkte abhängt. Man spricht dann auch von der Wegunabhängigkeit des Linienintegrals.

Um nun Bedingungen formulieren zu können, unter denen ein Vektorfeld konservativ ist, benötigen wir den Begriff des einfach-zusammenhängenden Bereichs, den wir anschaulich erläutern:

Ein Bereich heißt einfach-zusammenhängend, wenn sich jede im Bereich gelegene geschlossene Kurve auf einen Punkt „zusammenziehen“ lässt.

2.5.6 Satz (Wegunabhängigkeit eines Linienintegrals)

Ein Linienintegral $\int_C \langle \vec{F}, d\vec{r} \rangle$ ist genau dann wegunabhängig, wenn in einem einfach-zusammenhängenden Bereich, der den Integrationsweg C enthält, gilt, dass

$$\operatorname{rot} \vec{F} = \vec{0}$$

ist.

Dies wiederum ist äquivalent dazu, dass das Linienintegral längs einer im Bereich liegenden geschlossenen Kurve C den Wert Null hat:

$$\oint_C \langle \vec{F}, d\vec{r} \rangle = 0.$$

2.5.7 Beispiel

Gegeben sei das Vektorfeld $\vec{F}(x, y, z) = \begin{pmatrix} 2xy + z^3 \\ x^2 \\ 3xz^2 \end{pmatrix}$.

Wir untersuchen, ob die Linienintegrale über \vec{F} wegunabhängig sind.

Es gilt

$$\operatorname{rot} \vec{F}(x, y, z) = \begin{pmatrix} \frac{\partial F_3}{\partial y} - \frac{\partial F_2}{\partial z} \\ \frac{\partial F_1}{\partial z} - \frac{\partial F_3}{\partial x} \\ \frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y} \end{pmatrix}$$

$$\begin{aligned}
&= \begin{pmatrix} 0 \\ 3z^2 - 3z^2 \\ 2x - 2x \end{pmatrix} \\
&= \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}
\end{aligned}$$

d. h. $\operatorname{rot} \vec{F} = \vec{0}$ im gesamten \mathbb{R}^3 , also einem einfach zusammenhängenden Bereich. Insbesondere ist somit die Zirkulation, d. h.

$$\oint \langle \vec{F}, d\vec{r} \rangle$$

stets Null. Dies bestätigen wir noch einmal am Beispiel des geschlossenen Weges C , der durch

$$\vec{r}(t) = \begin{pmatrix} t(1-t) \\ t^2(1-t) \\ t(1-t)^2 \end{pmatrix}, 0 \leq t \leq 1,$$

gegeben ist.

- Ersetzung der Koordinaten im Feldvektor durch die parameterabhängigen Koordinaten $x(t) = t(1-t)$, $y(t) = t^2(1-t)$, $z(t) = t(1-t)^2$.

$$\vec{F}(x(t), y(t), z(t)) = \begin{pmatrix} 2t^3(1-t)^2 + t^3(1-t)^6 \\ t^2(1-t)^2 \\ 3t^3(1-t)^5 \end{pmatrix}$$

- Bestimmung des Tangentenvektors

$$\dot{\vec{r}}(t) = \begin{pmatrix} 1-2t \\ 2t-3t^2 \\ 1-4t+3t^2 \end{pmatrix}$$

- Berechnung des Linienintegrals

$$\begin{aligned}
\oint_C \langle \vec{F}, d\vec{r} \rangle &= \int_0^1 \left\langle \begin{pmatrix} 2t^3(1-t)^2 + t^3(1-t)^6 \\ t^2(1-t)^2 \\ 3t^3(1-t)^5 \end{pmatrix}, \begin{pmatrix} 1-2t \\ 2t-3t^2 \\ 1-4t+3t^2 \end{pmatrix} \right\rangle dt \\
&= \int_0^1 (8t^3 - 50t^4 + 144t^5 - 252t^6 + 280t^7 - 189t^8 + 70t^9 - 11t^{10}) dt \\
&= 0
\end{aligned}$$

Kapitel 3

Grundlagen der Statistik

3.1 Vorbemerkungen

Wahrscheinlichkeitsrechnung und Statistik sind Teile der Angewandten Mathematik, die sich u. a. zum Ziel setzt, auf reale Probleme durch die Anwendung mathematischer Methoden eine Antwort zu geben.

Schema:

Reales Problem (*Modellbildung*) \longrightarrow
Mathematisches Problem (*Anwendung mathematischer Methoden*) \longrightarrow
Lösung des mathematischen Problems \longrightarrow
Antwort

3.1.1 Bemerkung

Die Modellbildung ist oftmals sehr schwierig bzw. oft nur unter stark vereinfachenden Annahmen möglich. Die Qualität und Brauchbarkeit der Antwort sind von der Güte des Modells abhängig.

3.1.2 Beispiel

Beispiele für Fragestellungen, die mit Hilfe der Wahrscheinlichkeitsrechnung und Statistik beantwortet werden können, sind:

Wie groß ist der wahrscheinliche Anteil

- defekter Werkstücke an der Gesamtproduktion eines Betriebes,
- von Glühlampen, deren Lebensdauer größer als T ist,
- von Wählern, die am nächsten Sonntag die Partei XY wählen würden,
- weißer Kugeln, die aus einer Urne mit einer bestimmten Anzahl weißer und schwarzer Kugeln gezogen werden?

Die Beispiele weisen bereits in eine Richtung, die wir einschlagen werden. Man verwendet Stichprobenverfahren, d. h. man überprüft **nicht alle** betreffenden Objekte (was oft auch gar nicht möglich wäre), sondern wählt **zufällig** welche aus und schließt aus dem Ergebnis auf die **Grundgesamtheit** (Schließende Statistik).

Um diese Verfahren verstehen und anwenden zu können, benötigt man die Wahrscheinlichkeitsrechnung. Sie beantwortet z. B. Fragen der folgenden Art.

In einer Urne befinden sich 1000 weiße und 1000 schwarze Kugeln. Welches Ergebnis ist zu erwarten, wenn man 100 Kugeln zufällig zieht?

3.2 Zufallsexperimente - grundlegende Begriffe und Eigenschaften

Zufallsexperiment – Ergebnis hängt vom Zufall ab im Gegensatz zu deterministischen Experimenten. Die mathematische Beschreibung der (möglichen) Ergebnisse erfolgt mit Hilfe von Mengen.

3.2.1 Definition

- Ergebnismenge Ω : Menge aller möglichen Ergebnisse ω eines Zufallsexperimentes
- Elementarereignis $A = \{\omega\}$: Konkretes mögliches Ergebnis eines Zufallsexperimentes
- Ereignis A : Teilmenge von Ω , d. h. eine Menge von Elementarereignissen. A tritt ein, wenn das konkrete Ergebnis ω des Zufallsexperimentes in A ist.
- Ereignisraum \mathcal{A} : Menge aller Ereignisse, die sich aus der Ergebnismenge eines Zufallsexperimentes bilden lassen. Ist Ω endlich oder abzählbar, so ist der Ereignisraum $\mathcal{P}(\Omega) = \{A : A \subset \Omega\}$

3.2.2 Beispiel (Würfel mit einem idealen Würfel)

Ergebnismenge: $\Omega = \{1, 2, 3, 4, 5, 6\}$

Elementarereignisse: $\{1\}, \{2\}, \{3\}, \{4\}, \{5\}, \{6\}$

Ereignisse z.B.: $A = \{1\}, B = \{2, 4, 6\}, \Omega, \phi$

3.2.3 Bemerkung

Insbesondere sind ϕ und Ω Ereignisse mit folgender spezieller Bedeutung:

ϕ – unmögliches Ereignis, enthält kein mögliches Ergebnis,

Ω – sicheres Ereignis, enthält alle möglichen Ergebnisse.

3.2.4 Beispiel (Alkoholgehalt einer Flüssigkeit in Prozent)

Ergebnismenge: $\Omega = [0, 100]$

Elementarereignisse z. B.: $\{7\}, \{10, 0384\}$

Ereignisse z. B.: $A = [0, 7], B = [7, 100]$ – Alkoholgehalt von höchstens bzw. mindestens 7%.

3.2.5 Bemerkung (Verknüpfung von Ereignissen)

Seien Ω die Ergebnismenge eines Zufallsexperimentes und $A, B \subset \Omega$ beliebige Ereignisse.

- $A \cup B$: Mindestens eines der Ereignisse A oder B tritt ein.
- $A \cap B$: A und B treten beide gleichzeitig ein. Ist $A \cap B = \phi$, so schließen sich die Ereignisse A und B gegenseitig aus (unvereinbare Ereignisse).
- $A \dot{\cup} B$: Vereinigung unvereinbarer Ereignisse, d. h. A und B sind disjunkt ($A \cap B = \phi$).
- $\bar{A} = \Omega \setminus A$: Komplementärereignis, A tritt nicht ein.
- $A \setminus B = A \cap \bar{B}$: A tritt ein, B aber nicht.

3.2.6 Beispiel (Würfeln mit einem Würfel)

$\Omega = \{1, 2, 3, 4, 5, 6\}$

$A = \{1, 3, 5\}$ – Würfeln einer ungeraden Zahl.

$B = \{2, 4, 6\}$ – Würfeln einer geraden Zahl.

Es gilt: $\bar{A} = B, \bar{B} = A, A \cup B = \Omega$ (sicheres Ereignis), $A \cap B = \phi$ (unmögliches Ereignis)

3.2.7 Bemerkung (De Morgansche Regeln für Mengen (Ereignisse))

a) $\overline{A \cup B} = \bar{A} \cap \bar{B}$

b) $\overline{A \cap B} = \bar{A} \cup \bar{B}$

3.2.8 Beispiel (Würfeln mit einem Würfel)

$A = \{1, 2, 3, 4\}$ – Würfeln einer Zahl kleiner als 5.

$B = \{2, 4, 6\}$ – Würfeln einer geraden Zahl.

$\overline{A} = \{5, 6\}$, $\overline{B} = \{1, 3, 5\}$,

$\overline{A \cup B} = \{5\} = \overline{A} \cap \overline{B}$,

$\overline{A \cap B} = \{1, 3, 5, 6\} = \overline{A} \cup \overline{B}$.

3.2.9 Bemerkung (Mehrstufige Zufallsexperimente)

Werden n Zufallsexperimente nacheinander oder gleichzeitig durchgeführt und ist Ω_i die Ergebnismenge des i -ten Experimentes, so ist die Ergebnismenge des Gesamtexperimentes

$$\Omega = \{(\omega_1, \omega_2, \dots, \omega_n) : \omega_i \in \Omega_i\} = \Omega_1 \times \Omega_2 \times \dots \times \Omega_n,$$

d. h. die Menge aller geordneter n -Tupel.

Sind alle Ω_i identisch ($\Omega_i = \Omega$), so schreibt man auch Ω^n statt $\Omega_1 \times \Omega_2 \times \dots \times \Omega_n$.

3.3 Wahrscheinlichkeitsaxiome

Wir betrachten zunächst sehr allgemein ein Axiomensystem, dem eine reelle Funktion genügen muß, damit wir sie eine Wahrscheinlichkeit nennen. Mit der Formulierung der Axiome ist es Kolmogorov 1933 gelungen, den Wahrscheinlichkeitsbegriff widerspruchsfrei zu formulieren. Die Folgerungen aus diesen Axiomen können dann allgemein für die anschließend betrachteten konkreten Fälle benutzt werden.

3.3.1 Axiom (Axiome einer Wahrscheinlichkeit von Kolmogorov)

Eine Funktion P , die jedem Ereignis eine reelle Zahl zuordnet, heißt eine Wahrscheinlichkeit, wenn sie folgende Axiome erfüllt.

- a) $0 \leq P(A) \leq 1$ für jedes Ereignis A (Nichtnegativität),
- b) $P(\Omega) = 1$ für das sichere Ereignis (Normierung),
- c) $P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$ für Ereignisse, die paarweise unvereinbar sind, d. h. $A_j \cap A_k = \emptyset$ für alle j, k mit $j \neq k$ (Additivität).

3.3.2 Bemerkung

Hat man in c) $A_i = \emptyset$ für alle $i \geq n + 1$, so erhält man speziell

$$c^*) \quad P\left(\bigcup_{i=1}^n A_i\right) = \sum_{i=1}^n P(A_i) \text{ für paarweise unvereinbare Ereignisse (endliche Additivität).}$$

Aus den Axiomen ergeben sich unmittelbar einige allgemeine Folgerungen, die für die konkrete Rechnung sehr nützlich sind.

3.3.3 Folgerung

- a) $P(\emptyset) = 0$
- b) $P(\bar{A}) = 1 - P(A)$
- c) $A \subset B \implies P(A) \leq P(B)$ (Monotonie)
- d) $P(B \setminus A) = P(B) - P(A \cap B)$ für beliebige Ereignisse
- e) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ Additionssatz für beliebige Ereignisse

denn:

zu a):

$$\begin{aligned} \emptyset &= \emptyset \cup \emptyset, \quad \emptyset = \emptyset \cap \emptyset \implies P(\emptyset) \stackrel{c^*}{=} P(\emptyset) + P(\emptyset) \\ &\implies P(\emptyset) = 0 \end{aligned}$$

zu c):

$$B = B \cap \Omega = B \cap (A \cup \bar{A}) = (B \cap A) \cup (B \cap \bar{A})$$

Da $B \cap A = A$ wegen $A \subset B$ und $B \cap \bar{A} = B \setminus A$, gilt $B = A \cup (B \setminus A)$ mit $A \cap (B \setminus A) = \emptyset$. Daraus folgt

$$P(B) = P(A) + \underbrace{P(B \setminus A)}_{\geq 0} \implies P(A) \leq P(B).$$

zu b), d), e) vgl. Übung.

3.3.4 Bemerkung

Bei endlichen Ergebnismengen $\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$ bzw. abzählbar unendlichen Ergebnismengen $\Omega = \{\omega_1, \omega_2, \dots\}$ genügt für die Bestimmung der Wahrscheinlichkeiten die Angabe der Wahrscheinlichkeiten für die Elementarereignisse $p_i = P(\{\omega_i\}) \geq 0$, wobei $\sum_{i=1}^n p_i = 1$ bzw. $\sum_{i=1}^{\infty} p_i = 1$ gelten muß. Die Wahrscheinlichkeit für ein Ereignis A ist dann $P(A) = \sum_{\{i: \omega_i \in A\}} p_i$.

Bei überabzählbarer Ergebnismenge ist der Sachverhalt komplizierter. Dies wird später behandelt.

3.3.5 Bemerkung (Interpretation des Wahrscheinlichkeitsbegriffes)

- a) Man denkt an die häufige Wiederholung eines Zufallsexperimentes (z. B. Würfeln). Bei n -maliger Durchführung und m -maligem Eintreten des Ereignisses A kann man davon ausgehen, daß für die relative Häufigkeit m/n mit der A eingetreten ist, gilt: $m/n \approx P(A)$. Man kann die Wahrscheinlichkeit also als relative Häufigkeit in einer großen Zahl von Versuchen interpretieren.
- b) Ist $P(A)$ sehr klein, so kann man davon ausgehen, daß das Ereignis A bei einem Versuch **nicht** eintritt. (Beachte: $P(A) = 0$ heißt nicht, daß A nie eintreten kann. A unmögliches Ereignis $\Leftrightarrow A = \phi \Rightarrow P(A) = 0$, die Umkehrung gilt nicht.)

3.4 Laplace-Experimente

3.4.1 Definition (Laplace-Experiment)

Ein Zufallsexperiment heißt Laplace-Experiment, wenn gilt:

- a) Die Ergebnismenge $\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$ ist endlich.
- b) Alle Elementarereignisse sind gleichwahrscheinlich.

3.4.2 Bemerkung

Mit dieser Definition und der Bemerkung 3.3.4 ist die Laplace-Wahrscheinlichkeit gegeben durch

$$p_i = P(\{\omega_i\}) = 1/n$$

mit $n = \#\Omega =$ Anzahl der Elemente von Ω . Die Wahrscheinlichkeit für das Eintreten eines Ereignisses A bei einem Laplace-Experiment ist somit

$$P(A) = \sum_{\{i:\omega_i \in A\}} \frac{1}{n} = \frac{\#A}{\#\Omega} = \frac{\text{Anzahl der für } A \text{ günstigen Fälle}}{\text{Anzahl der insgesamt möglichen Fälle}}$$

3.4.3 Beispiel (Roulette)

$\Omega = \{0, 1, 2, 3, \dots, 36\}$, $p_i = 1/37$.

$$P(\text{ungerade Zahl}) = P(\text{rot}) = P(\text{schwarz}) = 18/37 < 1/2$$

Für die Bestimmung der Anzahlen günstiger und möglicher Fälle bei Laplace-Experimenten benötigen wir Abzählmethoden, die im folgenden Abschnitt behandelt werden.

Im Hinblick auf die Ausgangsfrage nach Stichprobenverfahren interessieren uns dabei besonders n -stufige Zufallsexperimente.

3.5 Hilfsmittel aus der Kombinatorik

Wie bereits in Kapitel 1 (vgl. Bemerkung 3.2.9) kurz angesprochen, versteht man unter einem n -stufigen Zufallsexperiment folgendes. n Zufallsexperimente werden nacheinander oder gleichzeitig durchgeführt. Die Ergebnismenge für das i -te Experiment sei Ω_i . Die Ergebnismenge für das Gesamtexperiment ist dann die folgende Menge geordneter n -Tupel

$$\Omega = \{(\omega_1, \omega_2, \dots, \omega_n) : \omega_i \in \Omega_i\} = \Omega_1 \times \Omega_2 \times \dots \times \Omega_n,$$

(bzw. Ω^n , falls $\Omega_1 = \Omega_2 = \dots = \Omega_n = \Omega$).

3.5.1 Satz (Produktregel der Kombinatorik)

Ist $\#\Omega_i = n_i$, dann ist $\#\Omega = \prod_{i=1}^n n_i$.

denn: Für ω_1 gibt es n_1 Möglichkeiten, zu jeder dieser Möglichkeiten gibt es für ω_2 n_2 Möglichkeiten, etc.

3.5.2 Beispiel (Würfeln mit unterscheidbaren Würfeln)

Für jeden Würfel ist die Ergebnismenge $\Omega_i = \{1, 2, 3, 4, 5, 6\}$ mit $n_i = 6$. Bei zwei Würfeln ist also $\Omega = \{(\omega_1, \omega_2) : \omega_i \in \Omega_i\}$, $\#\Omega = 6 \cdot 6$.

Achtung: Es ist z. B. $(1, 2) \neq (2, 1)$. Es wird also hierbei unterschieden, mit welchem Würfel die 1 bzw. die 2 gewürfelt wird.

3.5.3 Satz (Anordnungsmöglichkeiten verschiedener Elemente)

Für n verschiedene Elemente (Dinge), $n \in \mathbb{N}$, gibt es

$$n! = \prod_{i=1}^n i$$

verschiedene Anordnungsmöglichkeiten (Permutationen).

denn: Für das erste Element gibt es n Möglichkeiten, zu jeder dieser Möglichkeiten für das zweite Element $n-1$ Möglichkeiten, etc., d. h. mit der Produktregel (vgl. Satz 3.5.1) und $n_1 = n$, $n_2 = n-1$, $n_3 = n-2$, \dots , $n_n = n - (n-1) = 1$ also insgesamt $n!$ Möglichkeiten.

3.5.4 Beispiel

Für die Anordnung von fünf verschiedenen Gegenständen auf einem Regal gibt es $5! = 120$ Möglichkeiten.

3.5.5 Satz (Anordnungsmöglichkeiten von Gruppen gleicher Elemente)

Von den n Dingen seien jeweils n_1, n_2, \dots, n_r gleich, wobei $\sum_{i=1}^r n_i = n$. Dann gibt es $\frac{n!}{n_1!n_2!\dots n_r!}$ verschiedene Anordnungsmöglichkeiten.

denn: Macht man zunächst die gleichen Dinge unterscheidbar, dann (vgl. Satz 3.5.3) gibt es $n!$ Möglichkeiten. In jeder der x möglichen Anordnungen bei gleichen Dingen können jeweils n_k gleiche Dinge auf $n_k!$ Arten permutiert werden, d. h. $x \cdot n_1!n_2!\dots n_r! = n!$, d. h. nach Umstellung die gesuchte Formel.

3.5.6 Beispiel

Von den fünf Gegenständen seien drei optisch nicht unterscheidbare Kerzenleuchter und zwei Blumensträuße. Dann hat man $\frac{5!}{3!2!} = 10$ optisch unterscheidbare Möglichkeiten.

Modellvorstellung für die folgenden Überlegungen ist das sogenannte Urnenmodell:

Ziehung von k Kugeln aus einer mit n verschiedenen Kugeln gefüllten Urne mit bzw. ohne Berücksichtigung der Reihenfolge der gezogenen Kugeln und ohne bzw. mit Zurücklegen bereits gezogener Kugeln.

3.5.7 Satz (Auswahlmögl. mit Berücksichtigung der Reihenfolge)

Werden von n verschiedenen Dingen k Stück mit Berücksichtigung der Reihenfolge ausgewählt, dann ist die Anzahl der Auswahlmöglichkeiten

- a) ohne Zurücklegen (ohne Wiederholung)

$$n(n-1)(n-2) \cdot \dots \cdot (n-k+1) = \frac{n!}{(n-k)!} \text{ für } k \leq n,$$

- b) mit Zurücklegen (mit Wiederholung)

n^k für beliebiges k .

denn:

zu a) Für den ersten Zug gibt es n Möglichkeiten, zu jeder dieser Möglichkeiten für den zweiten Zug $n-1$ Möglichkeiten, etc. beim k -ten Zug $(n-(k-1)) = (n-k+1)$ Möglichkeiten.

zu b) Da nach jedem Zug zurückgelegt wird, stehen für jeden Zug n Möglichkeiten zur Verfügung, d.h. $\underbrace{n \cdot n \cdot n \cdot \dots \cdot n}_{k\text{-mal}} = n^k$ Möglichkeiten.

3.5.8 Beispiel

Wie groß ist die Wahrscheinlichkeit dafür, daß von $r \leq 365$ Personen mindestens zwei am gleichen Tag Geburtstag haben, wenn wir jeden Tag als Geburtstag gleichwahrscheinlich annehmen. (Ohne 29. Februar)

Da jede Person an jedem Tag Geburtstag haben kann (mit Reihenfolge, mit Wiederholung), gibt es 365^r Möglichkeiten für die Geburtstage von r Personen.

Ereignis A : Mindestens zwei Personen haben am gleichen Tag Geburtstag.

Komplementärereignis \bar{A} : Alle haben an verschiedenen Tagen Geburtstag.

$$\#\bar{A} = \frac{365!}{(365-r)!} \text{ (mit Reihenfolge, ohne Wiederholung)}$$

$$\text{Also } P(\bar{A}) = \frac{\text{Anzahl der für } \bar{A} \text{ günstigen Fälle}}{\text{Anzahl der insgesamt möglichen Fälle}} = \frac{365!}{(365-r)! 365^r} \text{ und } P(A) = 1 - P(\bar{A}).$$

$$r = 23: P(A) \approx 0,51 = 51\%$$

$$r = 30: P(A) \approx 0,73 = 73\%$$

$$r = 50: P(A) \approx 0,97 = 97\%$$

3.5.9 Satz (Auswahlmögl. ohne Berücksichtigung der Reihenfolge)

Werden von n verschiedenen Dingen k Stück ohne Berücksichtigung der Reihenfolge ausgewählt, dann ist die Anzahl der Auswahlmöglichkeiten

- a) ohne Zurücklegen (ohne Wiederholung)

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} \text{ für } k \leq n,$$

- b) mit Zurücklegen (mit Wiederholung)

$$\binom{n+k-1}{k} = \frac{(n+k-1)!}{k!(n-1)!} \text{ für beliebiges } k.$$

denn:

zu a) Aus jeder Möglichkeit ohne Berücksichtigung der Reihenfolge erhält man durch Permutation der k Elemente $k!$ Möglichkeiten mit Berücksichtigung der Reihenfolge, d. h. $x \cdot k! = n(n-1) \cdot \dots \cdot (n-k+1)$,

also $x = \binom{n}{k}$ Möglichkeiten.

zu b) (etwas komplizierter)

3.5.10 Beispiel (Lotto)

Ziehung der Lottozahlen 6 aus 49, d. h. $n = 49$, $k = 6$.

Die Kugeln werden nicht zurückgelegt. Es gibt $\binom{49}{6} = 13.983.816$ Möglichkeiten.

Mit welcher Wahrscheinlichkeit erhält man mit einem Tip genau 4 Richtige?

Es gibt $\binom{6}{4} = 15$ Möglichkeiten, aus den 6 Richtigen 4 auszuwählen und $\binom{43}{2} = 903$ Möglichkeiten, aus den 43 Falschen 2 auszuwählen.

Somit $\binom{6}{4} \binom{43}{2} = 13.545$ günstige Fälle für genau 4 Richtige, d. h.

$$P(\text{genau 4 Richtige}) = \frac{13.545}{13.983.816} \approx 9,69 \cdot 10^{-4}.$$

Mit welcher Wahrscheinlichkeit erhält man mit einem Tip mindestens 4 Richtige?

Ereignis A_i – genau i Richtige, Ereignis A – mindestens 4 Richtige.

Es gilt: $A = A_4 \cup A_5 \cup A_6$ mit paarweise disjunkten Ereignissen A_4, A_5, A_6 . Daher (Additivität, vgl. Bemerkung 3.3.2) $P(A) = P(A_4) + P(A_5) + P(A_6)$.

$$P(A_5) = \frac{\binom{6}{5} \binom{43}{1}}{13.983.816} = \frac{258}{13.983.816}, \quad P(A_6) = \frac{\binom{6}{6} \binom{43}{0}}{13.983.816} = \frac{1}{13.983.816}.$$

Somit

$$P(A) = \frac{13.545 + 258 + 1}{13.983.816} \approx 9,87 \cdot 10^{-4}.$$

Würden die Kugeln bei der Ziehung der Lottozahlen nach jedem Zug zurückgelegt, so gäbe es $\binom{49+6-1}{6} = \binom{54}{6} = 25.827.165$ verschiedene Möglichkeiten.

Tabellarische Zusammenstellung

	mit Berücksichtigung der Reihenfolge geordnete Stichprobe vom Umfang k	ohne Berücksichtigung der Reihenfolge ungeordnete Stichprobe vom Umfang k
ohne Wiederholung ohne Zurücklegen	$\frac{n!}{(n-k)!}$	$\binom{n}{k}$
mit Wiederholung mit Zurücklegen	n^k	$\binom{n+k-1}{k}$

3.6 Bedingte Wahrscheinlichkeiten und Unabhängigkeit von Ereignissen

In der Praxis interessiert man sich häufig für die Wahrscheinlichkeit für das Eintreten eines Ereignisses A unter der Bedingung, daß ein anderes Ereignis B bereits eingetreten ist.

Man spricht dann von bedingter Wahrscheinlichkeit von A unter der Bedingung B und kennzeichnet sie mit dem Symbol $P(A|B)$.

Wir machen uns die Zusammenhänge zunächst an einem Beispiel klar.

3.6.1 Beispiel

Die Personalzusammensetzung eines Betriebes sei in folgender Tabelle zusammengestellt.

	Angestellte (A)	Arbeiter (\bar{A})	Summe
weiblich (W)	400	100	500
männlich (\bar{W})	200	800	1000
Summe	600	900	1500

Unter den Mitarbeitern wird ein Preis ausgelost. Die Wahrscheinlichkeit für

- Angestellter Mitarbeiter erhält Preis:

$$P(A) = \frac{\#A}{\#\text{Mitarbeiter}} = \frac{600}{1500} = \frac{2}{5}$$

- Weibliche Person erhält Preis:

$$P(W) = \frac{\#W}{\#\text{Mitarbeiter}} = \frac{500}{1500} = \frac{1}{3}$$

- Weibliche Angestellte erhält Preis:

$$P(A \cap W) = \frac{\#(A \cap W)}{\#\text{Mitarbeiter}} = \frac{400}{1500} = \frac{4}{15}$$

Durch Indiskretion wird bekannt, daß eine weibliche Person ausgelost wurde. Unter dieser Bedingung (mit dieser Information) ist dann die Wahrscheinlichkeit für Angestellte erhält Preis:

$$P(A|W) = \frac{\#(A \cap W)}{\#W} = \frac{400}{500} = \frac{\frac{400}{1500}}{\frac{500}{1500}} = \frac{P(A \cap W)}{P(W)}$$

A , unter der Bedingung,
daß W erfüllt ist

Ist umgekehrt bekannt, daß ein Mitarbeiter aus der Gruppe der Angestellten ausgelost wurde, so ist die Wahrscheinlichkeit für weiblichen Mitarbeiter:

$$P(W|A) = \frac{\#(A \cap W)}{\#A} = \frac{400}{600} = \frac{\frac{400}{1500}}{\frac{600}{1500}} = \frac{P(A \cap W)}{P(A)}$$

W , unter der Bedingung,
daß A erfüllt ist

3.6.2 Definition (Bedingte Wahrscheinlichkeit)

Sei $P(B) > 0$.

- Die bedingte Wahrscheinlichkeit für das Eintreten eines Ereignisses A unter der Bedingung, daß ein Ereignis B eintritt bzw. eintreten wird, beträgt

$$P(A|B) = \frac{P(A \cap B)}{P(B)}.$$

- b) Ändert die Information über das Eintreten von B nichts an der Wahrscheinlichkeit für das Eintreten von A ($P(A|B) = P(A)$), d. h. $P(A \cap B) = P(A)P(B)$, so heißen A und B (stochastisch) unabhängig voneinander.
- c) n Ereignisse A_1, A_2, \dots, A_n heißen (vollständig) unabhängig, wenn für jedes k -Tupel $\{i_1, i_2, \dots, i_k\}$, $k \in \{2, \dots, n\}$, verschiedener Zahlen $i_j \in \{1, 2, \dots, n\}$ gilt:

$$P\left(\bigcap_{j=1}^k A_{i_j}\right) = \prod_{j=1}^k P(A_{i_j})$$

Den letzten Punkt der vorigen Definition wollen wir noch an einem Beispiel erläutern.

3.6.3 Beispiel

Drei Ereignisse A_1, A_2, A_3 sind vollständig unabhängig, wenn gilt:

$$\begin{aligned} P(A_1 \cap A_2) &= P(A_1)P(A_2) \\ P(A_1 \cap A_3) &= P(A_1)P(A_3) \\ P(A_2 \cap A_3) &= P(A_2)P(A_3) \\ P(A_1 \cap A_2 \cap A_3) &= P(A_1)P(A_2)P(A_3) \end{aligned}$$

Wie das folgende einfache Beispiel belegt, genügt es nicht, nur die letzte der genannten Bedingungen nachzuprüfen.

Beim Würfeln mit einem Würfel ist $\Omega = \{1, 2, 3, 4, 5, 6\}$. Wir betrachten die Ereignisse $A = \{1, 3, 5\}$ (Würfeln einer ungeraden Zahl), $B = \{2, 3, 4, 6\}$ (Augenzahl durch 2 oder 3 teilbar) und $C = \{1, 2, 3\}$ (Augenzahl kleiner als 4). Es ist

$$P(A) = \frac{1}{2}, \quad P(B) = \frac{2}{3}, \quad P(C) = \frac{1}{2}.$$

Die Schnittmengen sind

$$\begin{aligned} A \cap B &= \{3\}, \quad A \cap C = \{1, 3\}, \quad B \cap C = \{2, 3\}, \\ A \cap B \cap C &= \{3\}. \end{aligned}$$

Somit erhält man

$$\begin{aligned} P(A \cap B) &= \frac{1}{6} \neq P(A)P(B), \\ P(A \cap C) &= \frac{1}{3} \neq P(A)P(C), \\ P(B \cap C) &= \frac{1}{3} = P(B)P(C), \\ P(A \cap B \cap C) &= \frac{1}{6} = P(A)P(B)P(C). \end{aligned}$$

Insgesamt sind also nur die Ereignisse B und C stochastisch unabhängig. Die Ereignisse A , B und C sind **nicht** vollständig unabhängig.

Wir betrachten noch weitere Beispiele.

3.6.4 Beispiel

Jemand hat er erfahren, daß er beim Lotto (6 aus 49) mindestens vier Richtige hat. Wie groß ist unter dieser Bedingung die Wahrscheinlichkeit, sogar sechs Richtige zu haben?

- A : sechs Richtige

- B : mindestens vier Richtige

$$\begin{aligned}
 P(A|B) &= \frac{P(A \cap B)}{P(B)} \\
 &= \frac{P(A)}{P(B)} \\
 &= \frac{1}{\frac{13.983.816}{13.804}} = \frac{1}{13.804} \approx 7.24 \cdot 10^{-5}
 \end{aligned}$$

3.6.5 Beispiel

Eine Münze wird zehnmal geworfen. 500 Personen geben unabhängig voneinander einen Tip über das Ergebnis der Münzwurfserie ab. Wie groß ist die Wahrscheinlichkeit, daß mindestens eine Person höchstens einen Fehler macht?

$$\Omega = \{(i_1, i_2, \dots, i_{10}) : i_j \in \{\text{Kopf, Zahl}\}\}, \#\Omega = 2^{10} = 1024$$

A_i : Die i -te Person macht höchstens einen Fehler, $\#A_i = 11$, $P(A_i) = 11/1024$.

\bar{A}_i : Die i -te Person macht mehr als einen Fehler, $P(\bar{A}_i) = 1 - P(A_i) = 1 - 11/1024$.

$\bar{A} = \bigcap_{i=1}^{500} \bar{A}_i$: Alle Personen machen mehr als einen Fehler.

Da die Tips unabhängig erfolgen, ist

$$P(\bar{A}) = \prod_{i=1}^{500} P(\bar{A}_i) = \left(1 - \frac{11}{1024}\right)^{500}.$$

A : Mindestens eine Person macht höchstens einen Fehler mit

$$P(A) = 1 - P(\bar{A}) = 1 - \left(1 - \frac{11}{1024}\right)^{500} \approx 0.995 = 99.5\%$$

Durch Umstellen der Definitionsgleichung für bedingte Wahrscheinlichkeiten erhält man den

3.6.6 Satz (Multiplikationssatz für bedingte Wahrscheinlichkeiten)

Sei $P(B) > 0$. Dann gilt:

$$P(A \cap B) = P(A|B)P(B).$$

3.6.7 Bemerkung

Durch wiederholtes Anwenden von Satz 3.6.6 erhält man den allgemeinen Multiplikationssatz für bedingte Wahrscheinlichkeiten:

$$\begin{aligned}
 &P\left(\bigcap_{i=1}^n A_i\right) \\
 &= P\left(A_n \cap \left[\bigcap_{i=1}^{n-1} A_i\right]\right) \\
 &= P\left(A_n \left| \bigcap_{i=1}^{n-1} A_i \right.\right) P\left(A_{n-1} \cap \left[\bigcap_{i=1}^{n-2} A_i\right]\right) \\
 &= \dots \\
 &= P\left(A_n \left| \bigcap_{i=1}^{n-1} A_i \right.\right) P\left(A_{n-1} \left| \bigcap_{i=1}^{n-2} A_i \right.\right) \cdot \dots \cdot P(A_3|A_2 \cap A_1)P(A_2|A_1)P(A_1)
 \end{aligned}$$

3.6.8 Beispiel

Bei einer Lieferung von 20 Leuchtstoffröhren befinden sich 3 defekte. Der Händler entnimmt 3 Röhren als Stichprobe.

Wie groß ist die Wahrscheinlichkeit, daß er nur intakte Röhren zieht (Ereignis A)?

Die Entnahme von 3 verschiedenen Röhren entspricht dem dreimaligen nacheinander Ziehen ohne Zurücklegen.

Ereignis A_i : Ziehung einer intakten Röhre bei der i -ten Ziehung, $i = 1, 2, 3$.

- a) Ziehung:
20 Röhren, 17 intakte, $P(A_1) = 17/20$
- b) Ziehung unter der Bedingung, daß A_1 eingetreten ist:
19 Röhren, 16 intakte, $P(A_2|A_1) = 16/19$
- c) Ziehung unter der Bedingung, daß A_1 und A_2 eingetreten sind:
18 Röhren, 15 intakte, $P(A_3|A_2 \cap A_1) = 15/18$

Mit dem Multiplikationssatz folgt:

$$\begin{aligned} P(A) &= P(A_1 \cap A_2 \cap A_3) \\ &= P(A_3|A_2 \cap A_1)P(A_2|A_1)P(A_1) \\ &= \frac{17}{20} \cdot \frac{16}{19} \cdot \frac{15}{18} \\ &= \frac{34}{57} \approx 0.60 = 60\% \end{aligned}$$

3.6.9 Definition (Vollständige Ereignisdisjunktion)

Sei Ω Ergebnismenge. Die Ereignisse B_1, B_2, \dots, B_n bilden eine vollständige (totale) Ereignisdisjunktion, wenn gilt:

- a) $B_i \cap B_j = \emptyset$ für alle $i \neq j$, d. h. paarweise disjunkte, sich gegenseitig ausschließende Ereignisse.
- b) $P(B_i) > 0$ für alle i .
- c) $\bigcup_{i=1}^n B_i = \Omega$.

3.6.10 Bemerkung

Eine vollständige Ereignisdisjunktion zerlegt den Ergebnisraum in sich gegenseitig ausschließende Ereignisse. Jedes Elementarereignis ist in **genau** einem B_i enthalten.

3.6.11 Satz (Satz von der totalen Wahrscheinlichkeit)

Sei Ω Ergebnismenge, B_1, B_2, \dots, B_n vollständige Ereignisdisjunktion und A ein beliebiges Ereignis. Dann gilt:

$$P(A) = \sum_{i=1}^n P(A|B_i)P(B_i)$$

denn:

$$\sum_{i=1}^n P(A|B_i)P(B_i) = \sum_{i=1}^n \frac{P(A \cap B_i)}{P(B_i)} P(B_i)$$

$$\begin{aligned}
&= \sum_{i=1}^n P(A \cap B_i) \\
&= P\left(\bigcup_{i=1}^n (A \cap B_i)\right) \\
&= P(A)
\end{aligned}$$

Dabei gilt die vorletzte Gleichung mit Hilfe der Additionsformel für paarweise unvereinbare Ereignisse (vgl. Bemerkung 3.3.2), da $B_i \cap B_k = \emptyset$ für $j \neq k$ und daraus $(A \cap B_i) \cap (A \cap B_k) = \emptyset$ für $j \neq k$. Die letzte Gleichung gilt, weil die B_i nach Voraussetzung eine vollständige Ereignisdisjunktion bilden.

3.6.12 Beispiel

3 Lieferanten beliefern ein Geschäft mit 200, 300, 500 Leuchtstoffröhren, wobei 2%, 1%, 4% der Röhren defekt sind. Eine beliebige Röhre wird der Lieferung entnommen. Wie groß ist die Wahrscheinlichkeit dafür, dass diese defekt ist?

Ereignis A : Röhre defekt

Ereignis B_i : Röhre stammt vom Lieferanten i .

$$\begin{aligned}
P(A) &= P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + P(A|B_3)P(B_3) \\
&= \frac{2}{100} \cdot \frac{200}{1000} + \frac{1}{100} \cdot \frac{300}{1000} + \frac{4}{100} \cdot \frac{500}{1000} \\
&= 0.027
\end{aligned}$$

3.6.13 Beispiel

Im Betriebszeitraum eines technischen Systems liege mit den Wahrscheinlichkeiten 0.1, 0.3 und 0.6 einer der drei möglichen Belastungsgrade 1, 2 oder 3 vor. Beim Vorliegen des Belastungsgrades 1, 2 bzw. 3 fällt das System im Betriebszeitraum mit den Wahrscheinlichkeiten 0.4, 0.2 bzw. 0.1 aus.

Wie groß ist die Wahrscheinlichkeit für einen Ausfall des Systems im Betriebszeitraum?

Wir betrachten die Ereignisse:

A : Das System fällt im Betriebszeitraum aus

B_i : Es liegt Belastungsgrad i vor ($i = 1, 2, 3$)

Die Ereignisse B_i , $i = 1, 2, 3$, bilden eine vollständige Ereignisdisjunktion. Mit dem Satz von der totalen Wahrscheinlichkeit folgt also:

$$\begin{aligned}
P(A) &= P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + P(A|B_3)P(B_3) \\
&= 0.4 \cdot 0.1 + 0.2 \cdot 0.3 + 0.1 \cdot 0.6 \\
&= 0.16
\end{aligned}$$

Ohne Vorabinformation über den Belastungsgrad beträgt die Ausfallwahrscheinlichkeit des Systems im Betriebszeitraum 16%.

Durch Kombination der Definition für bedingte Wahrscheinlichkeiten, dem Multiplikationssatz für bedingte Wahrscheinlichkeiten und dem Satz von der totalen Wahrscheinlichkeit erhält man:

3.6.14 Satz (Bayessche Formel)

Sei Ω Ergebnismenge, B_1, B_2, \dots, B_n vollständige Ereignisdisjunktion und A ein beliebiges Ereignis mit $P(A) > 0$. Dann gilt

$$P(B_k|A) = \frac{P(A|B_k)P(B_k)}{\sum_{i=1}^n P(A|B_i)P(B_i)}.$$

denn:

$$\begin{aligned}P(B_k|A) &= \frac{P(B_k \cap A)}{P(A)} \\ &= \frac{P(A|B_k)P(B_k)}{\sum_{i=1}^n P(A|B_i)P(B_i)}\end{aligned}$$

3.6.15 Beispiel

Von den Röhren aus Beispiel 3.6.12 wurde eine defekte ausgewählt. Wie groß ist die Wahrscheinlichkeit, dass diese vom Lieferanten i , $i = 1, 2, 3$, stammt?

$$P(B_1|A) = \frac{P(A|B_1)P(B_1)}{P(A)} = \frac{\frac{4}{200} \cdot \frac{200}{1000}}{0.027} \approx 0.148 = 14.8\%$$

$$P(B_2|A) = \frac{P(A|B_2)P(B_2)}{P(A)} = \frac{\frac{3}{300} \cdot \frac{300}{1000}}{0.027} \approx 0.111 = 11.1\%$$

$$P(B_3|A) = \frac{P(A|B_3)P(B_3)}{P(A)} = \frac{\frac{20}{500} \cdot \frac{500}{1000}}{0.027} \approx 0.741 = 74.1\%$$

3.6.16 Beispiel

Bezüglich Beispiel 3.6.13 kann man auch nach der Wahrscheinlichkeit dafür fragen, dass zum Zeitpunkt eines Systemausfalls der Belastungsgrad i , $i = 1, 2, 3$, vorliegt.

$$P(B_1|A) = \frac{P(A|B_1)P(B_1)}{P(A)} = \frac{0.4 \cdot 0.1}{0.16} = 0.25 = 25\%$$

$$P(B_2|A) = \frac{P(A|B_2)P(B_2)}{P(A)} = \frac{0.23 \cdot 0.3}{0.16} = 0.375 = 37.5\%$$

$$P(B_3|A) = \frac{P(A|B_3)P(B_3)}{P(A)} = \frac{0.1 \cdot 0.6}{0.16} = 0.375 = 37.5\%$$

3.7 Zufallsvariable und Verteilungsfunktion

Die Definition von Zufallsvariablen dient dazu, jedem Elementarereignis bzw. jedem möglichen Ergebnis eines Zufallsexperimentes eine reelle Zahl zuzuordnen, d. h. eine Zufallsvariable X ist eine Abbildung $X : \Omega \rightarrow W \subset \mathbb{R}$.

3.7.1 Beispiel

- Würfeln mit einem homogenen Würfel. Die Zufallsvariable $X(\omega)$ beschreibt z. B. die Augenzahl.
- Würfeln mit zwei homogenen Würfeln. Die Zufallsvariable $X(\omega_1, \omega_2)$ beschreibt z. B. die Augensumme oder auch die Augenzahl des ersten Würfels etc.
- Ein Punkt mit den Koordinaten (ω_1, ω_2) wird zufällig aus einem Kreis mit Mittelpunkt 0 und Radius 1 markiert. Interessiert man sich für die Wahrscheinlichkeit, mit der dieser Punkt vom Mittelpunkt den Abstand $r \in [0, 1]$ hat, so definiert man als Zufallsvariable X den Abstand vom Mittelpunkt, d. h. $X(\omega_1, \omega_2) = \sqrt{\omega_1^2 + \omega_2^2}$.
- n -facher Münzwurf. Die Zufallsvariable $X(\omega_1, \dots, \omega_n)$ beschreibt z. B. die Anzahl der geworfenen Wappen.

3.7.2 Definition (Zufallsvariable)

Sei Ω die Ergebnismenge. Eine Zufallsvariable X ist eine Funktion

$$X : \Omega \rightarrow W \subset \mathbb{R} \quad \omega \rightarrow X(\omega) = x \in \mathbb{R},$$

wobei $P(\{\omega \in \Omega : X(\omega) \leq a\})$ für alle $a \in \mathbb{R}$ existieren muss.

3.7.3 Bemerkung

Man unterscheidet zwischen diskreten und stetigen Zufallsvariablen.

diskret: Es können nur endlich viele oder abzählbar unendlich viele Werte angenommen werden, z.B. Anzahl beobachteter Sternschnuppen pro Jahr, Anzahl der Störfälle eines technischen Systems pro Jahr, Anzahl der Wappen bei n -fachem Münzwurf.

stetig: Jeder beliebige Wert aus einem endlichen oder unendlichen Intervall kann angenommen werden, z.B. Lebensdauer eines technischen Systems, Niederschlagsmenge pro Monat.

Bei einer Zufallsvariablen sind von Interesse

- Der Wertebereich (Augenzahl, Abstand vom Mittelpunkt, etc.)
- Die Wahrscheinlichkeit P , dass die Zufallsvariable X einen bestimmten Wert annimmt, bzw. der Wert in einem bestimmten Intervall liegt.
- Im Mittel angenommene Werte und ein Maß für die Streuung um diese.

Dies führt auf folgende Definition.

3.7.4 Definition (Verteilungsfunktion)

Sei X eine Zufallsvariable über der Ergebnismenge Ω . Dann heißt die Abbildung

$$F : \mathbb{R} \rightarrow [0, 1]$$

$$x \rightarrow F(x) = P(\{\omega \in \Omega : X(\omega) \leq x\}) = P(X \leq x)$$

Verteilungsfunktion der Zufallsvariablen X .

3.7.5 Bemerkung

$F(x)$ gibt die Wahrscheinlichkeit dafür an, dass X höchstens gleich einer vorgegebenen Zahl x ist und besitzt folgende Eigenschaften:

- a) F ist monoton wachsend mit $0 \leq F(x) \leq 1$,
- b) $\lim_{x \rightarrow -\infty} F(x) = 0$,
- c) $\lim_{x \rightarrow \infty} F(x) = 1$,
- d) $P(a < X \leq b) = P(X \leq b) - P(X \leq a) = F(b) - F(a)$

3.8 Diskrete Zufallsvariablen

Beachten Sie das entsprechende MathePrisma Modul:

<http://www.MathePrisma.uni-wuppertal.de/Module/Verteilu>

Sei X eine diskrete Zufallsvariable mit Wertebereich (Menge der Realisierungen) $W = \{x_0, x_1, x_2, \dots\}$, wobei $x_i < x_{i+1}$ gelten soll. Die Wahrscheinlichkeitsverteilung sei $\{p_0, p_1, p_2, \dots\}$ mit $p_i = P(X = x_i) = P(\{\omega \in \Omega : X(\omega) = x_i\})$, $i = 0, 1, 2, \dots$ (Wahrscheinlichkeit der Realisierung x_i , vgl. auch Abschnitt 3).

Durch Angabe der p_i mit $p_i \geq 0$, $\sum_{i=0}^{\infty} p_i = 1$ (vgl. Bemerkung 3.3.4) ist die Verteilungsfunktion festgelegt.

Die Definition führt im diskreten Fall auf die folgenden konkreten Darstellungen.

a) Für $W = \{x_0, x_1, x_2, \dots\}$ ist

$$F(x) = P(X \leq x) = \begin{cases} 0 & x < x_0 \\ \sum_{i=0}^k p_i & x_k \leq x < x_{k+1}, \quad k = 0, 1, 2, \dots \end{cases}$$

b) Für $W = \{x_0, x_1, x_2, \dots, x_n\}$ ist

$$F(x) = P(X \leq x) = \begin{cases} 0 & x < x_0 \\ \sum_{i=0}^k p_i & x_k \leq x < x_{k+1}, \quad k = 0, 1, 2, \dots, n-1 \\ 1 & x_n \leq x \end{cases}$$

Im diskreten Fall sind die Verteilungsfunktionen also Treppenfunktionen. Es genügt daher $F(x_k) = P(X \leq x_k)$ für $x_k \in W$ anzugeben.

Bevor wir uns **speziellen** diskreten Verteilungsfunktionen zuwenden, definieren wir zunächst noch einige wichtige sogenannte parametrische Kenngrößen. Dies sind Parameter, die die jeweilige Verteilung charakterisieren.

3.8.1 Definition (Parametrische Kenngrößen für diskrete Zufallsvariablen)

Sei X eine diskrete Zufallsvariable mit $W = \{x_0, x_1, x_2, \dots\}$ und $p_i = P(X = x_i)$, $i = 0, 1, 2, \dots$

a) **Erwartungswert (Mittelwert):**

$$\mu(X) = E(X) = \sum_{i=0}^{\infty} p_i x_i$$

b) **Varianz:** Mittlere quadratische Abweichung einer Zufallsvariablen vom Erwartungswert, d. h.

$$\sigma^2(X) = \text{Var}(X) = E(X - E(X))^2 = \sum_{i=0}^{\infty} p_i (x_i - E(X))^2$$

Es gilt $\text{Var}(X) = E(X^2) - (E(X))^2$, denn:

$$\begin{aligned} \text{Var}(X) &= \sum_{i=0}^{\infty} p_i (x_i - E(X))^2 \\ &= \sum_{i=0}^{\infty} p_i (x_i^2 - 2x_i E(X) + (E(X))^2) \end{aligned}$$

$$\begin{aligned}
&= \sum_{i=0}^{\infty} p_i x_i^2 - 2E(X) \underbrace{\sum_{i=0}^{\infty} p_i x_i}_{=E(X)} + (E(X))^2 \underbrace{\sum_{i=0}^{\infty} p_i}_{=1} \\
&= E(X^2) - (E(X))^2
\end{aligned}$$

Diese Formel erleichtert oft die konkrete Berechnung der Varianz.

c) **Standardabweichung (Streuung):**

$$\sigma(X) = \sqrt{\text{Var}(X)}$$

d) **Variationskoeffizient:** Relatives Maß für die Streuung,

$$V(X) = \left| \frac{\sqrt{\text{Var}(X)}}{E(X)} \right|$$

3.8.2 Beispiel

In einem Rechnernetzwerk werden bis zu 6 Störungen am Tag beobachtet. Die Wahrscheinlichkeit p_i für i Störungen am Tag sei

i	0	1	2	3	4	5	6
p_i	0.02	0.1	0.15	0.25	0.35	0.12	0.01

X bezeichne die zufällige Anzahl der Störungen am Tag, $W = \{0, 1, 2, 3, 4, 5, 6\}$.

$E(X) = \sum_{i=0}^6 p_i \cdot i = 3.21$ ist die mittlere Anzahl der Störungen pro Tag.

$\text{Var}(X) = \sum_{i=0}^6 p_i (i - 3.21)^2 = 1.399818$, $\sqrt{\text{Var}(X)} \approx 1.2672$, $V(X) \approx 0.3948$.

Im folgenden betrachten wir spezielle Wahrscheinlichkeitsverteilungen mit ihren Interpretationen und parametrischen Kenngrößen.

a) **Gleichverteilung (Laplace-Verteilung)**

Eine Zufallsvariable X mit $W = \{x_1, x_2, \dots, x_n\}$ heißt gleichverteilt, wenn

$$p_i = P(X = x_i) = 1/n, \quad i = 1, 2, \dots, n.$$

Da $p_i \geq 0$ und $\sum_{i=1}^n p_i = 1$, sind (vgl. Abschnitt 2) die Axiome für eine Wahrscheinlichkeit erfüllt.

Die Verteilungsfunktion (gehörend zu den bereits definierten Laplace-Experimenten) ist gegeben durch:

$$F(x) = P(X \leq x) = \begin{cases} 0 & x < x_1 \\ k/n & x_k \leq x < x_{k+1}, \quad k = 1, 2, \dots, n-1 \\ 1 & x_n \leq x \end{cases}.$$

Im Spezialfall $x_i = i$ erhält man für die parametrischen Kenngrößen:

$$\begin{aligned}
E(X) &= \frac{1}{n} \sum_{i=1}^n i = \frac{1}{n} \frac{n(n+1)}{2} = \frac{n+1}{2}, \\
\text{Var}(X) &= \frac{1}{n} \sum_{i=1}^n \left(i - \frac{n+1}{2}\right)^2 = \dots = \frac{(n-1)(n+1)}{12}.
\end{aligned}$$

3.8.3 Beispiel (Würfeln mit einem homogenen Würfel)

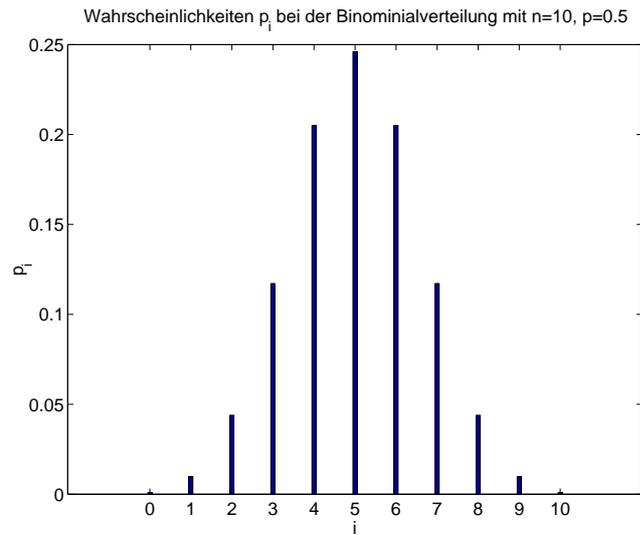
$W = \{1, 2, 3, 4, 5, 6\}$, $p_i = 1/6$, $i = 1, 2, 3, 4, 5, 6$.
Es gilt: $E(X) = 3.5$, $Var(X) = 35/12 = 2.91\bar{6}$

b) Binomialverteilung

Eine Zufallsvariable X mit $W = \{0, 1, 2, \dots, n\}$ heißt binomialverteilt, wenn

$$P(X = i) = p_i = \binom{n}{i} p^i (1-p)^{n-i} = b(i, n, p), \quad i = 0, 1, 2, \dots, n,$$

mit einem festen p , $0 < p < 1$.

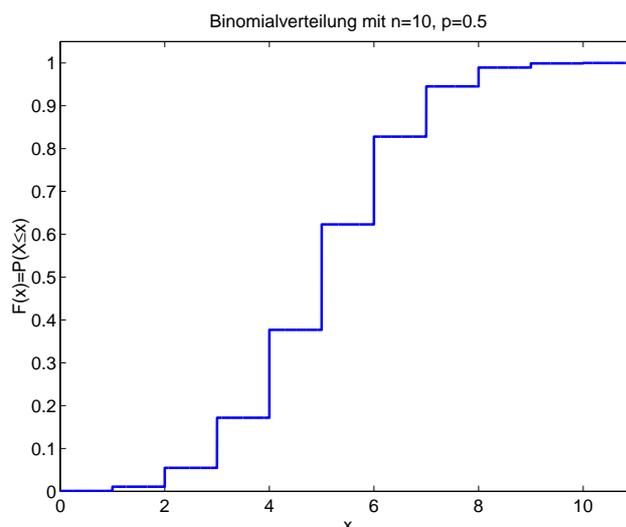


Da $p_i \geq 0$ und nach dem binomischen Lehrsatz

$$\sum_{i=0}^n p_i = \sum_{i=0}^n \binom{n}{i} p^i (1-p)^{n-i} = 1,$$

sind (vgl. Abschnitt 3) die Axiome für eine Wahrscheinlichkeit erfüllt.
Die zugehörige Verteilungsfunktion ist gegeben durch:

$$F(k) = P(X \leq k) = \sum_{i=0}^k \binom{n}{i} p^i (1-p)^{n-i} \quad 0 \leq k \leq n.$$



Für die parametrischen Kenngrößen gilt (ohne Beweis):

$$E(X) = np \text{ und } Var(X) = np(1 - p).$$

Interpretation: Wir erläutern zunächst das **Bernoullische Versuchsschema**. Es werden n voneinander unabhängige Versuche durchgeführt. Dabei interessieren nur die Versuchsausgänge **Erfolg** oder **Misserfolg** (sogenannte 0 – 1 Experimente). Bei jedem einzelnen Versuch ist p die Wahrscheinlichkeit für Erfolg und somit $1 - p$ die Wahrscheinlichkeit für das komplementäre Ereignis Misserfolg (Ziehen mit Zurücklegen).

3.8.4 Beispiel

Urne mit 5 weißen und 3 schwarzen Kugeln. Die Ziehung einer weißen Kugel sei das Ereignis Erfolg, die Ziehung einer schwarzen Kugel das komplementäre Ereignis Misserfolg. Somit ist $p = 5/8$ und $1 - p = 3/8$.

Wir führen das Experiment nun 5-mal unter jeweils gleichen Bedingungen durch, d. h. die gezogene Kugel wird wieder zurückgelegt.

Wie groß ist die Wahrscheinlichkeit, genau dreimal Erfolg zu haben?

Wir untersuchen im folgenden diese Fragestellung allgemein. Wir führen das Bernoulli-Experiment nun n mal durch. Wie groß ist die Wahrscheinlichkeit $P(X = k)$ genau k -mal, $0 \leq k \leq n$, Erfolg zu haben? $\Omega = \{(i_1, i_2, \dots, i_n) : i_j \in \{0, 1\}, j = 1, 2, \dots, n\}$, wobei 1 für Erfolg und 0 für Mißerfolg stehen soll.

$$A_k = \{(i_1, i_2, \dots, i_n) \in \Omega : \text{genau } k \text{ der } i_j \text{ sind gleich } 1\}$$

$$X : \underbrace{\Omega}_{n} \longrightarrow \{0, 1, 2, \dots, n\}, A_k \longrightarrow k \\ = \bigcup_{i=0}^n A_k$$

Es ist z. B. $\underbrace{(1, 1, \dots, 1)}_{k\text{-mal}}, \underbrace{(0, 0, \dots, 0)}_{(n-k)\text{-mal}} \in A_k$

Die Wahrscheinlichkeit für das Eintreten genau dieser Konstellation ist

$$\underbrace{\overbrace{p}^{1.\text{Platz}} \cdot \overbrace{p}^{2.\text{Platz}} \cdot \dots \cdot \overbrace{p}^{k\text{-ter Platz}}}_{k\text{-mal}} \cdot \underbrace{\overbrace{(1-p)}^{(k+1)\text{-ter Platz}} \cdot \overbrace{(1-p)}^{(k+2)\text{-ter Platz}} \cdot \dots \cdot \overbrace{(1-p)}^{n\text{-ter Platz}}}_{(n-k)\text{-mal}} = p^k (1-p)^{n-k}$$

Alle weiteren Elemente von A_k entstehen durch Permutationen der 2 Gruppen gleicher Elemente von k Einsen und $n - k$ Nullen, d. h.

$$\#A_k = \frac{n!}{k!(n-k)!} = \binom{n}{k}.$$

Für jede einzelne dieser Möglichkeiten ist die Wahrscheinlichkeit $p^k(1-p)^{n-k}$, also ist die Wahrscheinlichkeit für genau k -mal Erfolg:

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

Die oben genannte Verteilungsfunktion $F(k)$ gibt somit die Wahrscheinlichkeit an, höchstens k -mal Erfolg zu haben.

3.8.5 Beispiel

20 mal Würfeln mit einem Würfel. Wie groß ist die Wahrscheinlichkeit, mindestens dreimal eine 6 zu würfeln?

A_k : genau k -mal die 6,

$$X : \underbrace{\Omega}_{=\cup_{i=0}^{20} A_k} \longrightarrow \{0, 1, 2, \dots, 20\}, A_k \longrightarrow k.$$

$p = 1/6$ ist die Wahrscheinlichkeit, bei einem Wurf eine 6 zu würfeln. Die Wahrscheinlichkeit, bei 20 Würfeln genau k -mal eine 6 zu würfeln ist

$$P(X = k) = \binom{20}{k} \left(\frac{1}{6}\right)^k \left(1 - \frac{1}{6}\right)^{20-k}, \quad 0 \leq k \leq 20.$$

Die Wahrscheinlichkeit, bei 20 Würfeln mindestens dreimal die 6 zu würfeln, beträgt somit:

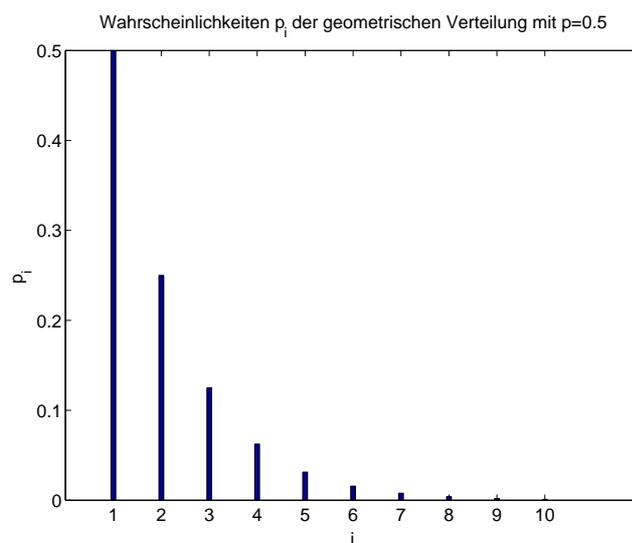
$$\begin{aligned} P(X \geq 3) &= 1 - P(X \leq 2) \\ &= 1 - \sum_{k=0}^2 \binom{20}{k} \left(\frac{1}{6}\right)^k \left(1 - \frac{1}{6}\right)^{20-k} \\ &= 1 - \left[\left(\frac{5}{6}\right)^{20} + 20 \cdot \frac{1}{6} \left(\frac{5}{6}\right)^{19} + 190 \left(\frac{1}{6}\right)^2 \left(\frac{5}{6}\right)^{18} \right] \approx 0.67134 \end{aligned}$$

c) Geometrische Verteilung

Eine Zufallsvariable X mit $W = \{1, 2, 3, \dots\}$ heißt geometrisch verteilt, wenn

$$p_i = p(1-p)^{i-1}, \quad i = 1, 2, \dots,$$

mit einem festen p , $0 < p < 1$.



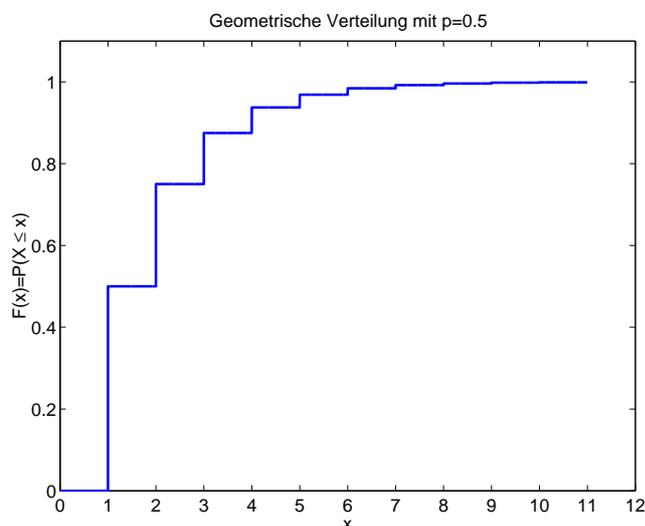
Da $p_i \geq 0$ und (vgl. Summenwert der geometrischen Reihe)

$$\sum_{i=1}^{\infty} p_i = \sum_{i=1}^{\infty} p(1-p)^{i-1} = p \sum_{j=0}^{\infty} (1-p)^j = p \frac{1}{1-(1-p)} = 1,$$

handelt es sich (vgl. Abschnitt 3) um eine Wahrscheinlichkeit.

Die zugehörige Verteilungsfunktion ist gegeben durch:

$$F(k) = P(X \leq k) = \sum_{i=1}^k p(1-p)^{i-1}.$$



Für die parametrischen Kenngrößen gilt (ohne Beweis):

$$E(X) = \frac{1}{p} \text{ und } Var(X) = \frac{1-p}{p^2}.$$

Interpretation: p bedeutet wie bei der Binomialverteilung die Wahrscheinlichkeit für Erfolg und $1-p$ die Wahrscheinlichkeit für Misserfolg bei einmaliger Durchführung eines 0 – 1 Experimentes. p_i ist die

Wahrscheinlichkeit dafür, genau bei der i -ten Durchführung, $i \geq 1$, zum ersten Mal Erfolg zu haben, denn:

$(i - 1)$ -mal Mißerfolg hat die Wahrscheinlichkeit $(1 - p)^{i-1}$, dann einmal Erfolg mit der Wahrscheinlichkeit p , also insgesamt $p(1 - p)^{i-1}$.

Die oben genannte Verteilungsfunktion $F(k)$ gibt somit die Wahrscheinlichkeit an, spätestens im k -ten Versuch zum ersten Mal Erfolg zu haben.

3.8.6 Beispiel (Würfeln mit einem Würfel)

Wie groß ist die Wahrscheinlichkeit, bis zum einschließlich vierten Versuch noch keine 6 gewürfelt zu haben?

A_k : Erste 6 genau im k -ten Versuch,

$$X : \underbrace{\Omega}_{=\bigcup_{k=1}^{\infty} A_k} \longrightarrow \{1, 2, 3, \dots\}, A_k \longrightarrow k.$$

$p = 1/6$ ist die Wahrscheinlichkeit, bei einem Wurf eine 6 zu würfeln, somit

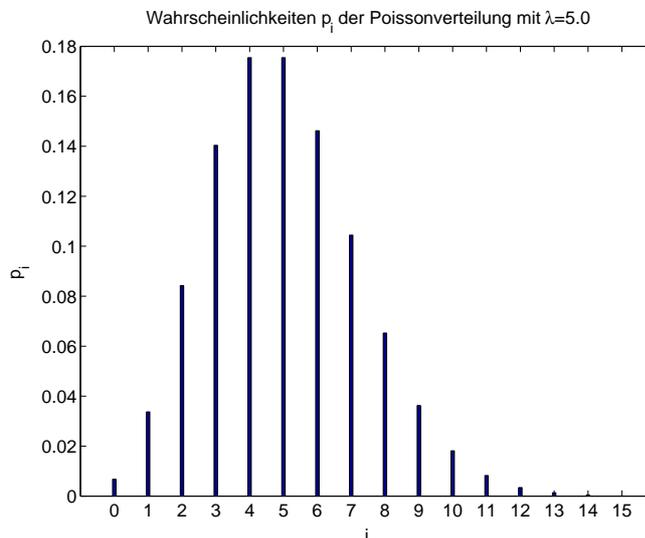
$$\begin{aligned} P(X \geq 5) = 1 - P(X \leq 4) &= 1 - \sum_{i=1}^4 \frac{1}{6} \left(1 - \frac{1}{6}\right)^{i-1} \\ &= 1 - \frac{1}{6} \left(1 + \frac{5}{6} + \frac{25}{36} + \frac{125}{216}\right) \approx 0.4823 \end{aligned}$$

d) Poissonverteilung

Eine Zufallsvariable X mit $W = \{0, 1, 2, 3, \dots\}$ heißt poissonverteilt, wenn

$$p_i = \frac{\lambda^i}{i!} e^{-\lambda} = p_{\lambda}(i), \quad i = 0, 1, 2, \dots,$$

mit einem festen $\lambda > 0$.

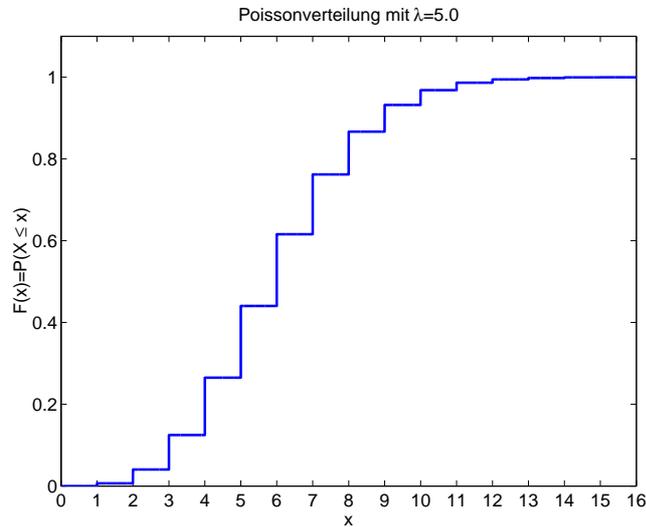


Da $p_i \geq 0$ und (vgl. Potenzreihenentwicklung der Exponentialfunktion)

$$\sum_{i=0}^{\infty} p_i = e^{-\lambda} \sum_{i=0}^{\infty} \frac{\lambda^i}{i!} = e^{-\lambda} e^{\lambda} = 1,$$

handelt es sich (vgl. Bemerkung 3.3.4) um eine Wahrscheinlichkeit.
Die zugehörige Verteilungsfunktion ist gegeben durch:

$$F(k) = P(X \leq k) = \sum_{i=0}^k \frac{\lambda^i}{i!} e^{-\lambda}, \quad k = 0, 1, 2, \dots$$



Für die parametrischen Kenngrößen gilt (ohne Beweis):

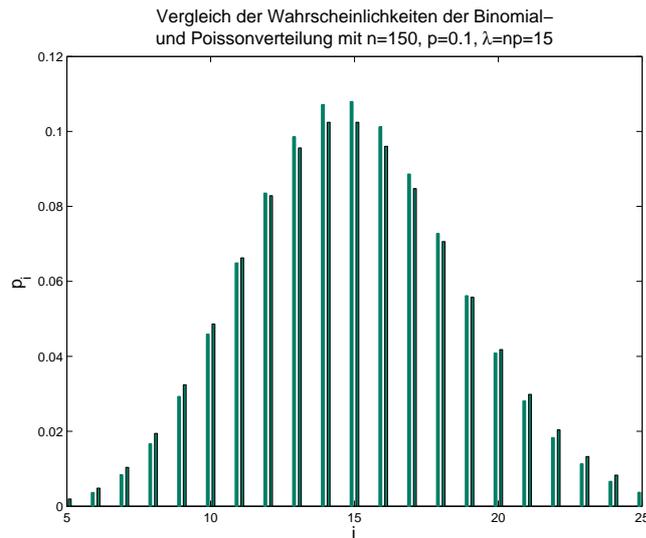
$$E(X) = \lambda \text{ und } Var(X) = \lambda.$$

Interpretation: Anwendbar bei sogenannten **seltene[n] Ereignissen**, d. h. man hat ein Bernoullisches Versuchsschema mit sehr geringer Erfolgswahrscheinlichkeit.

Für $np = \lambda$ konstant gilt

$$\lim_{n \rightarrow \infty, p \rightarrow 0} b(i, n, p) = p_\lambda(i).$$

Faustregel: Für $np < 10$ und $n > 1500p$ darf die Binomialverteilung näherungsweise durch die Poissonverteilung ersetzt werden.



3.8.7 Beispiel

Sei X die zufällige Anzahl von Anrufen pro fünf Minuten in einer Notrufzentrale. Aus Beobachtungen weiß man, dass X poissonverteilt ist mit einem Erwartungswert von $E(X) = \lambda = 10$. Mit welcher Wahrscheinlichkeit gibt es innerhalb von fünf Minuten mindestens einen Anruf?

Ereignis A_i : genau i Anrufe in 5 Minuten.

Ereignis A : mindestens ein Anruf in 5 Minuten

Ereignis \bar{A} : kein Anruf in 5 Minuten

$$p_i = P(A_i) = \frac{10^i}{i!} e^{-10}, \quad i = 0, 1, \dots$$

$$P(A) = \sum_{i=1}^{\infty} p_i = 1 - P(\bar{A}) = 1 - p_0 = 1 - e^{-10} \approx 0.99995$$

Es ist also ziemlich sicher, dass innerhalb von 5 Minuten ein Anruf erfolgt.

e) Hypergeometrische Verteilung

Seien $n, M, N \in \mathbb{N}_0$. Eine Zufallsvariable X mit $W = \{0, 1, 2, \dots, n\}$ heißt hypergeometrisch verteilt, wenn

$$p_i = \frac{\binom{M}{i} \binom{N-M}{n-i}}{\binom{N}{n}}, \quad i = 0, 1, 2, \dots, \min(n, M), \quad n - i \leq N - M, \quad p_i = 0 \text{ sonst,}$$

mit festen N, M , wobei $M \leq N, n \leq N$.

Da $p_i \geq 0$ und (ohne Beweis) $\sum_{i=0}^{\min(n, M)} p_i = 1$, handelt es sich (vgl. Bemerkung 3.3.4) um eine Wahrscheinlichkeit.

Die zugehörige Verteilungsfunktion ist gegeben durch:

$$F(k) = P(X \leq k) = \sum_{i=0}^k \frac{\binom{M}{i} \binom{N-M}{n-i}}{\binom{N}{n}}.$$

Für die parametrischen Kenngrößen gilt (ohne Beweis):

$$E(X) = n \frac{M}{N} \quad \text{und} \quad \text{Var}(X) = n \frac{M}{N} \left(1 - \frac{M}{N}\right) \frac{N-n}{N-1}.$$

Interpretation: Wir erläutern die Bedeutung der hypergeometrischen Verteilung an einem Beispiel.

3.8.8 Beispiel

Vorgegeben ist eine Urne mit insgesamt N Kugeln. Davon sind M Kugeln schwarz und $N - M$ weiß. Wir wählen $n \leq N$ Kugeln ohne Zurücklegen aus.

Wie groß ist die Wahrscheinlichkeit, dass genau i schwarze Kugeln, $i \leq n, i \leq M$, gezogen werden?

$\Omega = \{\{i_1, i_2, \dots, i_n\} : i_j \text{ seien } n \text{ verschiedene Kugeln aus der Urne}\}$, $\#\Omega = \binom{N}{n}$ (ohne Wiederholung, ohne Reihenfolge)

Die Anzahl der Möglichkeiten, aus M schwarzen Kugeln i auszuwählen beträgt $\binom{M}{i}$. Die Anzahl der Möglichkeiten, aus $N - M$ weißen Kugeln $n - i$ auszuwählen beträgt $\binom{N-M}{n-i}$. Somit ist

$$P(X = i) = \frac{\binom{M}{i} \binom{N-M}{n-i}}{\binom{N}{n}}.$$

Z. B. $N = 5, M = 3, n = 4, i = 2$ ergibt

$$P(X = 2) = \frac{\binom{3}{2} \binom{5-3}{4-2}}{\binom{5}{4}} = \frac{3}{5}.$$

3.8.9 Bemerkung

Für $M \gg n$ gilt die Näherung

$$\frac{\binom{M}{i} \binom{N-M}{n-i}}{\binom{N}{n}} \approx \binom{n}{i} p^i (1-p)^{n-i} \text{ mit } p = \frac{M}{N}.$$

Zusammen mit der Näherung durch die Poissonverteilung gilt

$$\frac{\binom{M}{i} \binom{N-M}{n-i}}{\binom{N}{n}} \approx \frac{\lambda^i}{i!} e^{-\lambda} \text{ für } n \text{ groß, } \lambda = np, p = \frac{M}{N} \text{ klein.}$$

3.9 Stetige Zufallsvariable

Ist die Menge W der Realisierungen der Zufallsvariable X ein endliches oder unendliches Intervall, dann lässt sich die Verteilung nicht mehr durch diskrete Werte $p_i = P(X = x_i)$, $x_i \in W$ beschreiben. Es ist die Angabe einer Verteilungsfunktion $F : \mathbb{R} \rightarrow \mathbb{R}$, so dass $F(x) = P(X \leq x)$ nötig. Zu einer Verteilungsfunktion gehört eine sogenannte Dichte, die im folgenden definiert wird.

3.9.1 Definition (Verteilungsdichte)

Eine Zufallsgröße mit der Verteilungsfunktion F heißt stetig (verteilt), wenn eine stückweise stetige Funktion f existiert, so dass für alle $x \in \mathbb{R}$ die Gleichung

$$P(X \leq x) = F(x) = \int_{-\infty}^x f(u) du$$

gilt. Die Funktion f heißt Dichte (Verteilungsdichte) der Verteilungsfunktion F .

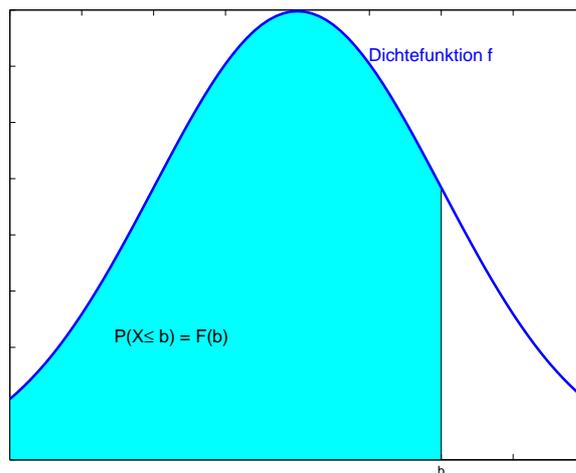
3.9.2 Bemerkung

Sei X stetig verteilt.

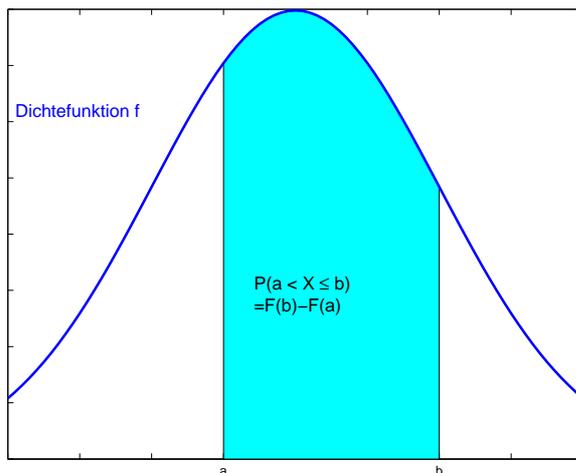
- F ist stetig. Ist f in u_0 stetig, so gilt dort $F'(u_0) = f(u_0)$, d. h. für stetiges f ist die Dichte die erste Ableitung der Verteilungsfunktion.
- Da F monoton wachsend ist (vgl. Bemerkung 3.7.5), gilt $F'(x) = f(x) \geq 0$. Somit ist $F(x)$ anschaulich der Inhalt der Fläche, die von f und der u -Achse zwischen $-\infty$ und x eingeschlossen wird.
- Es gilt $P(a < X \leq b) = F(b) - F(a) = \int_a^b f(u) du$.
- Wegen Bemerkung 3.7.5 ist

$$1 = \lim_{x \rightarrow \infty} F(x) = \lim_{x \rightarrow \infty} \int_{-\infty}^x f(u) du = \int_{-\infty}^{\infty} f(u) du.$$

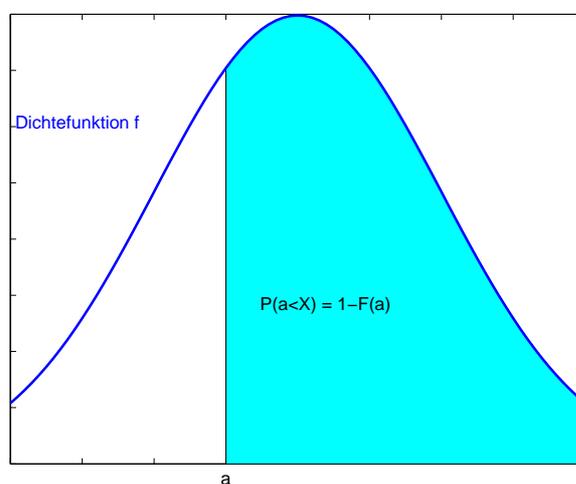
- Aus jeder stückweise stetigen Funktion f mit $f(u) \geq 0$ und $\int_{-\infty}^{\infty} f(u) du = 1$ erhält man mit $F(x) = \int_{-\infty}^x f(u) du$ eine Verteilungsfunktion.
- Ist die Zufallsvariable X stetig verteilt mit der Dichte f , so kann man die Wahrscheinlichkeiten, mit denen X in Intervallen liegt, wie folgt veranschaulichen:
 - $P(X \leq b)$



ii) $P(a < X \leq b)$



iii) $P(a \leq X)$



g) Es ist $P(X = a) = \int_a^a f(u) du = 0$. Das bedeutet aber nicht, dass der Wert a für die Zufallsvariable X unmöglich ist. Dies zeigt, warum eine Charakterisierung der Verteilung bei stetigen Zufallsvariablen nicht wie im diskreten Fall durch Angabe von Werten $P(X = x_i)$ möglich ist.

3.9.3 Definition (Parametrische Kenngrößen)

Sei X stetige Zufallsvariable mit der Verteilungsfunktion F und der Dichte f .

a) Erwartungswert (Mittelwert):

$$\mu(X) = E(X) = \int_{-\infty}^{\infty} f(x)x dx$$

b) Varianz: Mittlere quadratische Abweichung einer Zufallsvariablen vom Erwartungswert, d.h.

$$\sigma^2(X) = Var(X) = E(X - E(X))^2 = \int_{-\infty}^{\infty} f(x)(x - E(X))^2 dx$$

Wie im diskreten Fall gilt auch hier die Beziehung $Var(X) = E(X^2) - (E(X))^2$, denn:

$$Var(X) = \int_{-\infty}^{\infty} f(x)(x - E(X))^2 dx$$

$$\begin{aligned}
&= \int_{-\infty}^{\infty} f(x)(x^2 - 2E(X)x + (E(X))^2) dx \\
&= \underbrace{\int_{-\infty}^{\infty} f(x)x^2 dx}_{=E(X^2)} - 2E(X) \underbrace{\int_{-\infty}^{\infty} f(x)x dx}_{=E(X)} + (E(X))^2 \underbrace{\int_{-\infty}^{\infty} f(x) dx}_{=1} \\
&= E(X^2) - (E(X))^2
\end{aligned}$$

c) Standardabweichung (Streuung):

$$\sigma(X) = \sqrt{\text{Var}(X)}$$

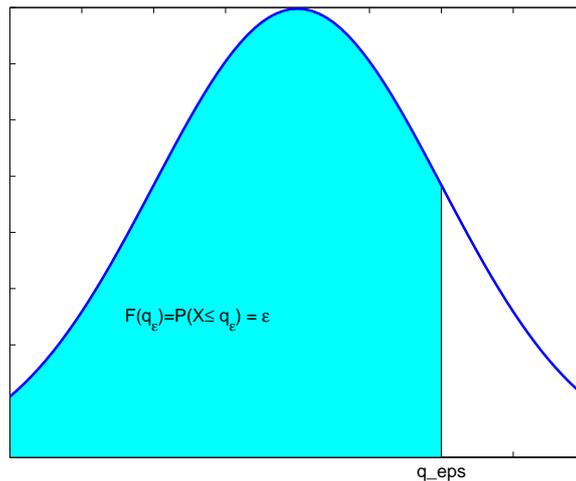
d) Variationskoeffizient: Relatives Maß für die Streuung,

$$V(X) = \left| \frac{\sqrt{\text{Var}(X)}}{E(X)} \right|$$

e) ε -Quantil: Jede Zahl q_ε , $0 < \varepsilon < 1$, für die

$$F(q_\varepsilon) = P(X \leq q_\varepsilon) = \varepsilon$$

gilt, heißt ε -Quantil der Zufallsvariablen X .



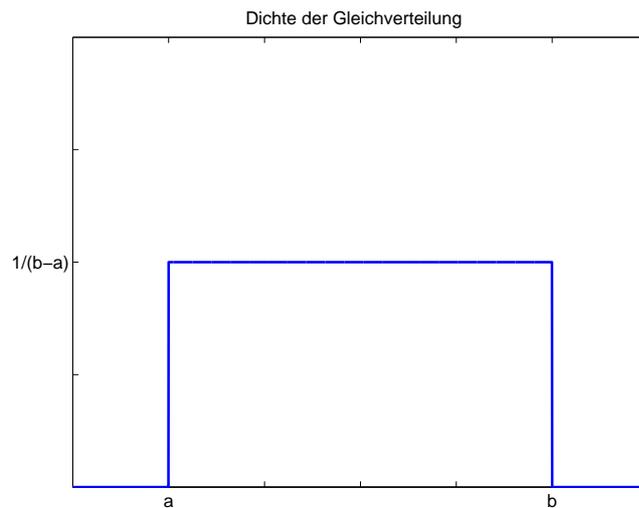
Speziell heißen $q_{0.5}$ auch Median (Zentrum der Wahrscheinlichkeit), $q_{0.25}$ unteres Quartil und $q_{0.75}$ oberes Quartil.

Im folgenden betrachten wir spezielle Verteilungen, die in der Praxis wichtig sind, und beschäftigen uns mit ihren Interpretationen und parametrischen Kenngrößen.

a) **Stetige Gleichverteilung**

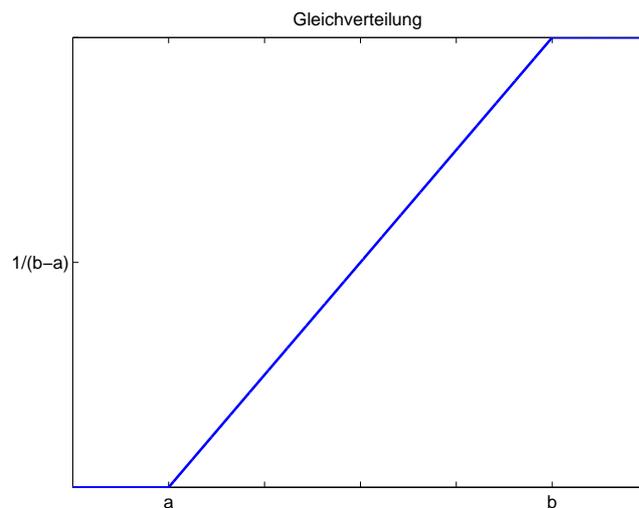
Eine Zufallsvariable X heißt gleichverteilt (gleichmäßig verteilt), wenn sie die Dichte

$$f(x) = \begin{cases} \frac{1}{b-a}, & a \leq x \leq b \\ 0, & \text{sonst} \end{cases}$$



hat. Als Verteilungsfunktion ergibt sich

$$F(x) = \int_{-\infty}^x f(u) du = \begin{cases} 0, & x \leq a \\ \frac{x-a}{b-a}, & a < x \leq b \\ 1, & x > b \end{cases} .$$



Eine gleichmäßig verteilte Zufallsvariable hat als parametrische Kenngrößen:

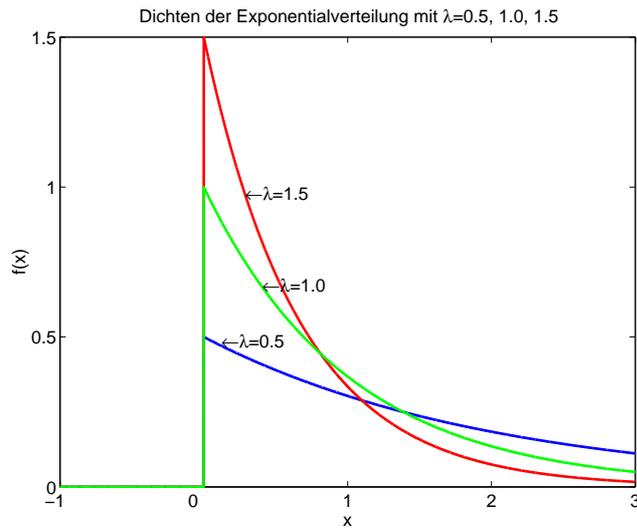
$$E(X) = \frac{b+a}{2}, \quad Var(X) = \frac{(b-a)^2}{12} .$$

b) **Exponentialverteilung**

Eine Zufallsvariable X heißt exponentialverteilt mit dem Parameter $\lambda > 0$, wenn sie die Dichte

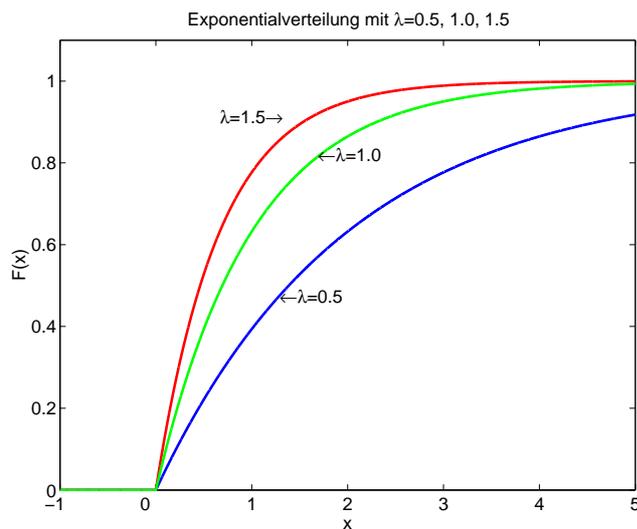
$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0 & x < 0 \end{cases}$$

hat.



Als Verteilungsfunktion ergibt sich

$$F(x) = \int_{-\infty}^x f(u) du = \begin{cases} 1 - e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases} .$$



Eine exponentialverteilte Zufallsvariable hat als parametrische Kenngrößen:

$$E(X) = \frac{1}{\lambda}, \quad \text{Var}(X) = \frac{1}{\lambda^2} \quad \text{und} \quad V(x) = 1 .$$

Interpretation: Exponentialverteilungen werden oft verwendet als Modell für die Lebensdauer technischer Geräte.

3.9.4 Beispiel

Die Lebensdauer T eines elektronischen Bauelementes sei eine exponentialverteilte Zufallsvariable mit $\lambda = 1/10$ (d. h. $E(X) = 10$). Wie groß ist der wahrscheinliche Anteil an Bauelementen mit einer Lebensdauer von mehr als 10.

Wir haben

$$F(t) = \begin{cases} 1 - e^{-\frac{1}{10}t}, & t \geq 0 \\ 0, & t < 0 \end{cases}$$

und es ergibt sich

$$\begin{aligned} P(T > 10) &= 1 - P(T \leq 10) \\ &= 1 - F(10) = 1 - (1 - e^{-1}) \\ &= e^{-1} \approx 0.3679. \end{aligned}$$

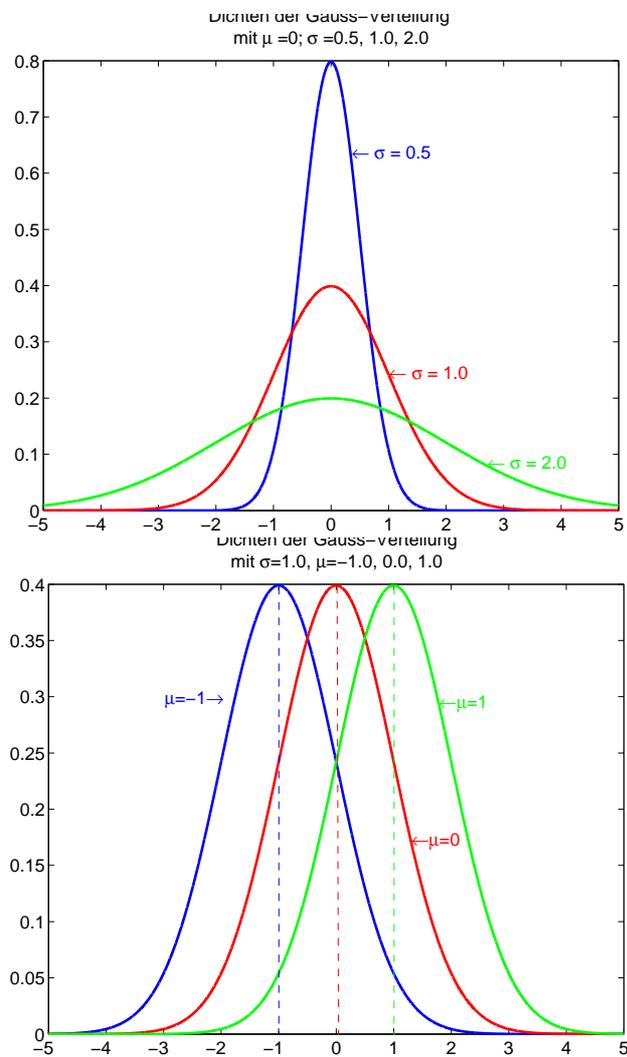
Der Anteil beträgt somit circa 36.79%.

c) **Normalverteilung (Gaußsche Verteilung)**

Eine Zufallsvariable X heißt normalverteilt mit den Parametern $\mu \in \mathbb{R}$ und $\sigma > 0$, wenn sie die Dichte

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

hat.



Als Verteilungsfunktion ergibt sich

$$F(x) = \int_{-\infty}^x f(u) du = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^x e^{-\frac{1}{2}\left(\frac{u-\mu}{\sigma}\right)^2} du.$$

Dieses Integral ist **nicht elementar berechenbar!**

Eine normalverteilte Zufallsvariable hat als parametrische Kenngrößen:

$$E(X) = \mu, \quad Var(X) = \sigma^2 \quad \text{und} \quad V(x) = \frac{\sigma}{|\mu|}.$$

Interpretation: Normalverteilungen werden benutzt, wenn man davon ausgehen kann, dass die Werte einer Zufallsvariablen um einen Mittelwert streuen, wie z. B. bei zufälligen Messfehlern und Fertigungstoleranzen. Ein weiteres wesentliches Stichwort ist der zentrale Grenzwertsatz (vgl. Kapitel 10), der besagt, dass viele Zufallsvariablen näherungsweise normalverteilt sind.

3.9.5 Bemerkung

- i) Ist die Zufallsvariable X normalverteilt mit den Parametern μ und σ , so sagt man, sie sei $N(\mu, \sigma^2)$ -verteilt. Die Bedeutung der Parameter ergibt sich aus $E(X) = \mu$ und $Var(X) = \sigma^2$.
- ii) Wie oben bereits bemerkt, ist das Integral in der Verteilungsfunktion nicht elementar berechenbar. Man arbeitet daher mit tabellierten Werten (s. u.).
- iii) Die Dichte ist symmetrisch bzgl. μ , besitzt an der Stelle μ ihr absolutes Maximum und hat an den Stellen $\mu - \sigma$, $\mu + \sigma$ ihre Wendepunkte.
- iv) Die zu $\mu = 0$ und $\sigma = 1$ gehörende Verteilung heißt Standardnormalverteilung bzw. $N(0, 1)$ -Verteilung. Ihre Dichte ist

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$$

und ihre Verteilungsfunktion

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{1}{2}u^2} du.$$

Wegen der Symmetrieeigenschaften von φ gilt:

$$\Phi(0) = 1/2 \quad \text{und} \quad \Phi(-x) = 1 - \Phi(x).$$

Die Werte der $N(0, 1)$ -Verteilung liegen in Tabellenform vor. Da man es aber allgemein mit $N(\mu, \sigma^2)$ -verteilten Zufallsgrößen zu tun hat, muss man diese in $N(0, 1)$ -verteilte Zufallsgrößen umrechnen (vgl. Satz 3.9.6).

3.9.6 Satz

Sei X $N(\mu, \sigma^2)$ -verteilt. Dann ist $Y = \frac{X - \mu}{\sigma}$ $N(0, 1)$ -verteilt.

denn: Zu jedem $x \in \mathbb{R}$ existiert ein $y \in \mathbb{R}$, so dass $x = \sigma y + \mu$ ($\Leftrightarrow y = \frac{x - \mu}{\sigma}$). Also gilt:

$$\begin{aligned} \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^x e^{-\frac{1}{2}\left(\frac{u-\mu}{\sigma}\right)^2} du &= \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\sigma y + \mu} e^{-\frac{1}{2}\left(\frac{u-\mu}{\sigma}\right)^2} du \\ &= \underbrace{\frac{1}{\sqrt{2\pi}} \int_{-\infty}^y e^{-\frac{1}{2}v^2} dv}_{\substack{\text{mit der Substitution} \\ v = \frac{u-\mu}{\sigma}, \quad dv = \frac{1}{\sigma} du, \\ u_1 = -\infty \rightarrow v_1 = -\infty, \\ u_2 = \sigma y + \mu \rightarrow v_2 = y}} \\ &= \Phi(y) \end{aligned}$$

(Der Graph wird durch diese Transformation verschoben und gedehnt bzw. gestaucht.)

3.9.7 Beispiel

Ein Kunde kauft Zylinder mit 20mm Durchmesser und akzeptiert Abweichungen von weniger als $\pm 0.5\text{mm}$. Der Lieferant produziert Zylinder, die $N(20, \sigma^2)$ -verteilt sind.

- i) Es sei $\sigma = 0.8\text{mm}$. Wieviel Prozent der gelieferten Zylinder werden abgelehnt? Wir berechnen zunächst, wieviel Prozent akzeptiert werden und verwenden dabei Satz 3.9.6.

$$\begin{aligned}
 P(19.5 < X < 20.5) &= P(\underbrace{-0.5 < X - 20 < 0.5}_{\text{Subtraktion von } \mu}) \\
 &= P\left(\underbrace{\frac{-0.5}{0.8} < \frac{X - 20}{0.8} < \frac{0.5}{0.8}}_{\text{Division durch } \sigma}\right) \\
 &\quad Y = \frac{X-20}{0.8} \text{ ist nun nach Satz 3.9.6 } N(0,1)\text{-verteilt, also} \\
 &= \Phi\left(\frac{0.5}{0.8}\right) - \underbrace{\Phi\left(\frac{-0.5}{0.8}\right)}_{=1 - \Phi\left(\frac{0.5}{0.8}\right)} \\
 &= 2\Phi\left(\frac{0.5}{0.8}\right) - 1 \\
 &\approx 2 \cdot 0.734 - 1 \\
 &= 0.468
 \end{aligned}$$

Insgesamt werden also 46.8% akzeptiert, d. h. 53.2% werden abgelehnt.

- ii) Wie groß ist σ , wenn 20% abgelehnt werden?

Da 80% akzeptiert werden, ergibt sich:

$$\begin{aligned}
 0.8 &= P\left(\frac{-0.5}{\sigma} < \frac{X - 20}{\sigma} < \frac{0.5}{\sigma}\right) \\
 &= 2\Phi\left(\frac{0.5}{\sigma}\right) - 1
 \end{aligned}$$

Durch Umstellung der Gleichung erhält man $\Phi\left(\frac{0.5}{\sigma}\right) = 0.9$. Aus der Tabelle liest man $\Phi(1.28) \approx 0.9$ ab, d. h. $0.5/\sigma \approx 1.28$ und somit $\sigma \approx 0.39$.

3.9.8 Bemerkung

Aus dem Beispiel lässt sich ablesen, dass für $a > 0$ für ein um den Erwartungswert symmetrisches Intervall

$$P(\mu - a\sigma < X < \mu + a\sigma) = P\left(-a < \frac{X - \mu}{\sigma} < a\right) = 2\Phi(a) - 1$$

gilt, wenn X $N(\mu, \sigma^2)$ -verteilt ist.

Einige markante Werte sind:

$$\begin{aligned}
 P(\mu - \sigma < X < \mu + \sigma) &= 2\Phi(1) - 1 \\
 &\approx 2 \cdot 0.841 - 1 = 0.682 \\
 P(\mu - 1.96\sigma < X < \mu + 1.96\sigma) &= 2\Phi(1.96) - 1 \\
 &\approx 2 \cdot 0.975 - 1 = 0.95 \\
 P(\mu - 2\sigma < X < \mu + 2\sigma) &= 2\Phi(2) - 1 \\
 &\approx 2 \cdot 0.977 - 1 = 0.954 \\
 P(\mu - 3\sigma < X < \mu + 3\sigma) &= 2\Phi(3) - 1 \\
 &\approx 2 \cdot 0.999 - 1 = 0.998
 \end{aligned}$$

Anders ausgedrückt liegen bei $N(\mu, \sigma^2)$ -verteilten Zufallsgrößen

68.2%	aller Werte im Intervall	$[\mu - \sigma, \mu + \sigma]$
95%	aller Werte im Intervall	$[\mu - 1.96\sigma, \mu + 1.96\sigma]$
95.4%	aller Werte im Intervall	$[\mu - 2\sigma, \mu + 2\sigma]$
99.8%	aller Werte im Intervall	$[\mu - 3\sigma, \mu + 3\sigma]$

Einen sehr wichtigen Hinweis auf die Bedeutung von Normalverteilungen zeigt uns im folgenden ein Spezialfall des Zentralen Grenzwertsatzes, der für geeignete Parameterwerte eine Näherung der Binomialverteilung durch die Normalverteilung begründet.

3.9.9 Satz (Grenzwertsatz von De Moivre und Laplace)

Sei X eine binomialverteilte Zufallsvariable mit den Parametern n und p , d. h. (vgl. Kapitel 8) $E(X) = np$ und $Var(X) = np(1-p)$. Dann ist für hinreichend großes n X näherungsweise $N(\mu, \sigma^2)$ -verteilt mit den Parametern $\mu = np$ und $\sigma^2 = np(1-p)$, d. h.

$$F(k) \approx \Phi\left(\frac{k - np}{\sqrt{np(1-p)}}\right), \quad k = 0, 1, \dots, n,$$

sowie

$$P(k_1 \leq X \leq k_2) \approx \Phi\left(\frac{k_2 - np}{\sqrt{np(1-p)}}\right) - \Phi\left(\frac{k_1 - np}{\sqrt{np(1-p)}}\right)$$

mit $k_1, k_2 = 0, 1, \dots, n$, $k_1 < k_2$.

Stetigkeitskorrektur: Da die Verteilungsfunktion der Binomialverteilung stückweise konstant ist, gilt $F(k) = F(k+1/2)$. Die Näherung lässt sich durch eine sogenannte Stetigkeitskorrektur verbessern, d. h. man verwendet

$$F(k) \approx \Phi\left(\frac{k + 1/2 - np}{\sqrt{np(1-p)}}\right), \quad k = 0, 1, \dots, n,$$

sowie

$$P(k_1 \leq X \leq k_2) \approx \Phi\left(\frac{k_2 + 1/2 - np}{\sqrt{np(1-p)}}\right) - \Phi\left(\frac{k_1 - 1/2 - np}{\sqrt{np(1-p)}}\right)$$

mit $k_1, k_2 = 0, 1, \dots, n$, $k_1 \leq k_2$.

3.9.10 Beispiel

Wir werfen 1000 mal eine faire Münze und zählen die Anzahl der Wappen. Dann ist $\mu = 1000 \cdot 0.5 = 500$ und $\sigma^2 = 1000 \cdot 0.5(1-0.5) = 250$. Die Wahrscheinlichkeit dafür, dass die Anzahl der tatsächlich geworfenen Wappen zwischen $\mu - 1.96\sigma \approx 469$ und $\mu + 1.96\sigma \approx 531$ liegt, beträgt circa 95%.

Faustregel: Satz 3.9.9 liefert eine brauchbare Näherung, wenn $np(1-p) = \sigma^2 > 9$ ist. (In diesem Fall ist dann $0 \leq \mu - 3\sigma < \mu + 3\sigma \leq n$).

Es gibt natürlich noch eine ganze Menge weiterer Verteilungen. Z. B. χ^2 -Verteilung, Student- t -Verteilung, Weibull-Verteilung. Zur Erlangverteilung vgl. Übungen.

3.10 Funktionen von einer oder mehreren Zufallsvariablen

Wir behandeln im folgenden das sogenannte Gesetz vom unbewussten Statistiker. Dies heißt so, weil die Formeln oft einfach als selbstverständlich angenommen werden.

3.10.1 Satz (Gesetz vom unbewussten Statistiker)

Sei $X : \Omega \rightarrow \mathbb{R}$ Zufallsvariable, $h : \mathbb{R} \rightarrow \mathbb{R}$ eine reelle Funktion und $Y = h \circ X$ wieder Zufallsvariable. Dann gilt

a) bei stetigen Zufallsvariablen

$$E(Y) = \int_{-\infty}^{\infty} h(x)f(x) dx ,$$

wobei f die Wahrscheinlichkeitsdichte von X bezeichnet und

b) bei diskreten Zufallsvariablen

$$E(Y) = \sum_{i=0}^{\infty} h(x_i)p_i ,$$

wobei $p_i = P(X = x_i)$.

$E(Y)$ kann also ohne Kenntnis der Dichte von Y berechnet werden.

Denn: Wir betrachten nur den Fall, dass h streng monoton wachsend ist, sonst ist der Beweis recht schwierig. Außerdem beschränken wir uns auf stetige Zufallsvariablen, der diskrete Fall geht analog.

Sei G die Verteilungsfunktion und g die Dichte von Y und F die Verteilungsfunktion und f die Dichte von X . Als weitere Einschränkung betrachten wir hier nur den Fall, dass f und g stetig sind. Wegen der vorausgesetzten strengen Monotonie von h existiert die Umkehrfunktion, die wir mit h^{-1} bezeichnen. Dann gilt:

$$G(y) = P(Y \leq y) = P(h(X) \leq y) = P(X \leq h^{-1}(y)) = F(h^{-1}(y)) .$$

Differenziert man beide Seiten nach y , so erhält man durch Anwendung der Kettenregel und der Regel für die Differentiation von Umkehrfunktionen:

$$g(y) = f(h^{-1}(y)) \cdot \frac{1}{h'(h^{-1}(y))}$$

Somit:

$$\begin{aligned} E(Y) &= \int_{-\infty}^{\infty} yg(y) dy \\ &= \int_{-\infty}^{\infty} yf(h^{-1}(y)) \frac{1}{h'(h^{-1}(y))} dy \\ &= \int_{-\infty}^{\infty} h(x)f(x) \frac{h'(x)}{h'(x)} dx \\ &= \int_{-\infty}^{\infty} h(x)f(x) dx \end{aligned}$$

Für die vorletzte Umformung wurde die Substitution $y = h(x)$, d. h. $h^{-1}(y) = x$ und $dy = h'(x) dx$ verwendet.

3.10.2 Beispiel

X sei gleichverteilt im Intervall $[0, 3]$, Y sei das Volumen einer Kugel mit Radius X , d. h. $Y = \frac{4}{3}\pi X^3$. Berechne den Erwartungswert für das Kugelvolumen.

Es ist

$$f(x) = \begin{cases} 1/3, & 0 \leq x \leq 3 \\ 0, & \text{sonst} \end{cases}, \quad h(x) = \frac{4}{3}\pi x^3.$$

Somit ist der Erwartungswert für Y

$$E(Y) = \int_0^3 \frac{4}{3}\pi x^3 \cdot \frac{1}{3} dx = \frac{\pi}{9} x^4 \Big|_0^3 = 9\pi.$$

Dies entspricht einer Kugel vom Radius $(27/4)^{1/3} \approx 1.89$.

In den vorausgegangenen Kapiteln traten bereits mehrere Zufallsvariable gemeinsam auf, ohne dass wir dies explizit erwähnt haben.

Man erhält z. B. beim gleichzeitigen Werfen zweier Würfel das zufällige Ergebnis (X, Y) , wobei X das zufällige Ergebnis des einen, Y das des anderen Würfels bezeichnet. Bei der Binomialverteilung mit den Parametern n und p hat man es mit n Einzelversuchen (Bernoulli-Experimenten) mit den zufälligen Ergebnissen X_1, X_2, \dots, X_n zu tun, wobei jedes X_i die Werte 1 oder 0 annehmen kann. Die Zufallsvariablen X_1, X_2, \dots, X_n bilden hierbei eine n -dimensionale Zufallsvariable (X_1, X_2, \dots, X_n) .

Im folgenden betrachten wir Zufallsvariablen X und Y als Komponenten des zufälligen Vektors (X, Y) bzw. manchmal auch allgemeiner Zufallsvariablen X_1, X_2, \dots, X_n als Komponenten des zufälligen Vektors (X_1, X_2, \dots, X_n) .

3.10.3 Definition (Gemeinsame Verteilungsfunktion)

Unter der gemeinsamen Verteilungsfunktion des zufälligen Vektors (X, Y) versteht man die Wahrscheinlichkeit

$$F(x, y) = P(\{X \leq x\} \cap \{Y \leq y\}) \quad \text{kurz } F(x, y) = P(X \leq x, Y \leq y)$$

mit $x, y \in \mathbb{R}$.

Entsprechend ist die gemeinsame Verteilungsfunktion des zufälligen Vektors (X_1, X_2, \dots, X_n)

$$F(x_1, x_2, \dots, x_n) = P(X_1 \leq x_1, X_2 \leq x_2, \dots, X_n \leq x_n)$$

für $x_1, x_2, \dots, x_n \in \mathbb{R}$.

Im folgenden bezeichnen wir mit $F_X(x)$ bzw. $F_Y(y)$ die jeweiligen Verteilungsfunktionen der (eindimensionalen) Zufallsvariablen X bzw. Y (sogenannte Randverteilungen).

3.10.4 Satz (Eigenschaften von $F(x, y)$)

- $\lim_{x \rightarrow -\infty, y \rightarrow -\infty} F(x, y) = 0$ und $\lim_{x \rightarrow \infty, y \rightarrow \infty} F(x, y) = 1$,
- $0 \leq F(x, y) \leq 1$ für alle $x, y \in \mathbb{R}$,
- $\lim_{y \rightarrow \infty} F(x, y) = F_X(x)$ und $\lim_{x \rightarrow \infty} F(x, y) = F_Y(y)$,
- für $x_1 \leq x_2$ und $y_1 \leq y_2$ gelten die Ungleichungen
 $F(x_1, y_1) \leq F(x_2, y_1) \leq F(x_2, y_2)$ und $F(x_1, y_1) \leq F(x_1, y_2) \leq F(x_2, y_2)$,
- $P(X > x, Y \leq y) = F_Y(y) - F(x, y)$ und $P(X \leq x, Y > y) = F_X(x) - F(x, y)$,
- $P(a < X \leq b, c < Y \leq d) = [F(b, d) - F(b, c)] - [F(a, d) - F(a, c)]$.

3.10.5 Definition (Gemeinsame Verteilungsdichte)

Existiert die gemischte partielle Ableitung $\frac{\partial^2 F(x, y)}{\partial x \partial y}$, so ist

$$f(x, y) = \frac{\partial^2 F(x, y)}{\partial x \partial y}$$

die gemeinsame Verteilungsdichte des zufälligen Vektors (X, Y) .

Im folgenden bezeichnen wir mit $f_X(x)$ bzw. $f_Y(y)$ die jeweiligen Dichten der (eindimensionalen) Zufallsvariablen X bzw. Y (sogenannte Randverteilungsdichten).

3.10.6 Satz (Eigenschaften von $f(x, y)$)

- a) $F(x, y) = \int_{-\infty}^y \int_{-\infty}^x f(u, v) du dv,$
- b) $f(x, y) \geq 0,$
- c) $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(u, v) dv du = 1,$
- d) $f_X(x) = \int_{-\infty}^{\infty} f(x, v) dv$ und $f_Y(y) = \int_{-\infty}^{\infty} f(u, y) du,$
- e) $P((X, Y) \in G) = \iint_G f(u, v) dG.$

In Übereinstimmung mit der Definition der Unabhängigkeit von Ereignissen halten wir fest:

3.10.7 Definition (Unabhängigkeit von Zufallsvariablen)

Zwei Zufallsvariablen X und Y mit der gemeinsamen Verteilungsfunktion $F(X, Y)$ heißen voneinander unabhängig, wenn

$$P(X \leq x, Y \leq y) = P(X \leq x) \cdot P(Y \leq y)$$

für alle $x, y \in \mathbb{R}$ gilt.

Dies ist gleichbedeutend damit, dass für alle $x, y \in \mathbb{R}$ $F(x, y) = F_X(x) \cdot F_Y(y)$ (bzw. $f(x, y) = f_X(x) \cdot f_Y(y)$, falls die gemeinsame Verteilungsdichte existiert) erfüllt ist. Die Zufallsvariablen X_1, X_2, \dots, X_n heißen unabhängig, wenn für alle x_1, x_2, \dots, x_n gilt

$$P(X_1 \leq x_1, X_2 \leq x_2, \dots, X_n \leq x_n) = P(X_1 \leq x_1) \cdot P(X_2 \leq x_2) \cdot \dots \cdot P(X_n \leq x_n).$$

3.10.8 Beispiel

Gegeben sei die gemeinsame Verteilungsdichte

$$f(x, y) = \begin{cases} \frac{1}{2}xy & \text{falls } 0 \leq x \leq y \leq 2 \\ 0 & \text{sonst} \end{cases}$$

des zufälligen Vektors (X, Y) . Wir berechnen die Wahrscheinlichkeit $P(X^2 \geq Y)$, d. h.

$$G = \{(x, y) : 0 \leq x \leq y \leq 2, y \leq x^2\} \cup \{(x, y) : 1 \leq y \leq 2, \sqrt{y} \leq x \leq y\}$$

$$\begin{aligned} P(X^2 \geq Y) &= \iint_G f(x, y) dG \\ &= \int_1^2 \int_{\sqrt{y}}^y \frac{1}{2}xy dx dy \\ &= \int_1^2 \frac{1}{4}x^2y \Big|_{x=\sqrt{y}}^y dy \\ &= \frac{1}{4} \int_1^2 (y^3 - y^2) dy \\ &= \frac{1}{4} \left(\frac{1}{4}y^4 - \frac{1}{3}y^3 \right) \Big|_1^2 \\ &= \frac{1}{4} \left(4 - \frac{8}{3} - \frac{1}{4} + \frac{1}{3} \right) \\ &= \frac{17}{48} \end{aligned}$$

3.10.9 Beispiel

Gegeben sei die gemeinsame Verteilungsdichte

$$f(x, y) = \begin{cases} a(x + y) & \text{falls } 0 \leq x, y \leq 1 \\ 0 & \text{sonst} \end{cases}$$

des zufälligen Vektors (X, Y) . Welchen Wert muss a haben? Sind X und Y unabhängig?

$$\begin{aligned} 1 &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dx dy \\ &= a \int_0^1 \int_0^1 (x + y) dx dy \\ &= a \int_0^1 \left(\frac{1}{2}x^2 + yx \right) \Big|_0^1 dy \\ &= a \int_0^1 \left(\frac{1}{2} + y \right) dy \\ &= a \cdot \frac{1}{2} (y + y^2) \Big|_0^1 \\ &= a \end{aligned}$$

Somit muss a den Wert 1 haben.

$$f_X(x) = \int_0^1 (x + y) dy = xy + \frac{1}{2}y^2 \Big|_0^1 = x + \frac{1}{2}, \quad 0 \leq x \leq 1$$

$$f_Y(y) = \int_0^1 (x + y) dx = \frac{1}{2}x^2 + xy \Big|_0^1 = y + \frac{1}{2}, \quad 0 \leq y \leq 1$$

d. h. $f_X(x)f_Y(y) = xy + \frac{1}{2}(x + y) + \frac{1}{4} \neq f(x, y)$. Somit sind die Zufallsvariablen X und Y **nicht** unabhängig.

Im folgenden betrachten wir die gemeinsame Verteilung diskreter Zufallsvariablen.

Es seien $W_X = \{x_0, x_1, \dots\}, W_Y = \{y_0, y_1, \dots\}$ die Mengen der Realisierungen der diskreten Zufallsvariablen X und Y . Mit $p_i = P(X = x_i)$ und $q_j = P(Y = y_j)$ bezeichnen wir die zugehörigen Einzelwahrscheinlichkeiten. Mit $r_{ij} = P(X = x_i, Y = y_j)$ bildet die Menge $\{r_{ij} : i, j = 0, 1, \dots\}$ die gemeinsame Verteilung des Vektors (X, Y) . Die Mengen $\{p_i : i = 0, 1, \dots\}$ und $\{q_j : j = 0, 1, \dots\}$ bilden die Randverteilungen von (X, Y) . Die Randverteilungen lassen sich aus der gemeinsamen Verteilung durch

$$p_i = \sum_{j=0}^{\infty} r_{ij}, \quad i = 0, 1, \dots \quad q_j = \sum_{i=0}^{\infty} r_{ij}, \quad j = 0, 1, \dots$$

berechnen. X und Y sind unabhängig, wenn

$$P(X = x_i, Y = y_j) = P(X = x_i) \cdot P(Y = y_j) \text{ bzw. } r_{ij} = p_i q_j$$

für alle $i, j = 0, 1, \dots$ erfüllt ist.

3.10.10 Beispiel

Wir nehmen zwei unterscheidbare Würfel mit zufälligen Augenzahlen X_1 und X_2 . Weiter sei $X = \max\{X_1, X_2\}$ und Y die Anzahl der geraden Augenzahlen, d. h. $W_X = \{1, 2, 3, 4, 5, 6\}$ und $W_Y = \{0, 1, 2\}$. X_1, X_2 sind unabhängig, da gilt:

$$P(X_1 = i, X_2 = j) = P(X_1 = i) \cdot P(X_2 = j) = \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36}$$

X, Y sind abhängig, da z.B. für $X = 1$ zwangsläufig $Y = 0$ sein muss, d. h.

$$P(X = 1) = 1/36, P(Y = 0) = 9/36$$

aber

$$P(X = 1, Y = 0) = 1/36$$

Wir betrachten nun folgende Verknüpfungen von Zufallsvariablen.

Produkt $Z = X \cdot Y$:

W_Z besteht aus allen Produkten $x \cdot y$ mit x, y Realisierungen von X bzw. Y , bzw. allgemeiner $Z = X_1 \cdot X_2 \cdot \dots \cdot X_n$, wobei W_Z aus allen Produkten $x_1 \cdot x_2 \cdot \dots \cdot x_n$ mit x_i Realisierung von X_i besteht.

Summe $Z = X + Y$:

W_Z besteht aus allen Summen $x + y$ mit Realisierungen x, y von X bzw. Y , bzw. allgemeiner $Z = X_1 + X_2 + \dots + X_n$, wobei W_Z aus allen Summen $x_1 + x_2 + \dots + x_n$ mit x_i Realisierung von X_i besteht.

Im folgenden Satz werden einige Eigenschaften für die parametrischen Kenngrößen für Summen und Produkte von Zufallsvariablen angegeben. Dabei ist es wichtig, genau darauf zu achten, ob die Unabhängigkeit der Zufallsvariablen vorausgesetzt wird.

3.10.11 Satz

Sind X_1, X_2, \dots, X_n Zufallsvariable, $a_1, a_2, \dots \in \mathbb{R}$, so gilt:

a)

$$E\left(\sum_{i=1}^n a_i X_i\right) = \sum_{i=1}^n a_i E(X_i)$$

b)

$$\text{Var}\left(\sum_{i=1}^n a_i X_i\right) = \sum_{i=1}^n \text{Var}(a_i X_i) = \sum_{i=1}^n a_i^2 \text{Var}(X_i),$$

falls die Zufallsvariablen unabhängig sind.

c)

$$E\left(\prod_{i=1}^n X_i\right) = \prod_{i=1}^n E(X_i),$$

falls die Zufallsvariablen unabhängig sind.

3.10.12 Beispiel (Fortsetzung von Beispiel 3.10.10)

$$\begin{aligned} E(X_1 \cdot X_2) &= E(X_1) \cdot E(X_2) \\ &= 3.5 \cdot 3.5 = 12.25 \\ E(X_1 + X_2) &= E(X_1) + E(X_2) \\ &= 3.5 + 3.5 \\ &= 7 \end{aligned}$$

Wir befassen uns nun noch damit, wie man für $Z = X + Y$ die Verteilung aus den Verteilungen von X und Y bestimmen kann, wenn X und Y unabhängig sind.

3.10.13 Satz

Sind X, Y unabhängige diskrete Zufallsvariable mit $W_X = \{x_1, x_2, \dots\}$ und $W_Y = \{y_1, y_2, \dots\}$, dann gilt

$$P(X + Y = z) = \sum_{i=1}^{\infty} P(X = x_i) \cdot P(Y = z - x_i).$$

3.10.14 Satz

Sind X, Y stetige unabhängige Zufallsvariable, so ist die Verteilungsdichte h von $X + Y$ die sogenannte Faltung der Verteilungsdichten f und g von X bzw. Y , d. h.

$$h(x) = (f \star g)(x) = \int_{-\infty}^{\infty} g(u)f(x-u)du \left(= \int_{-\infty}^{\infty} f(u)g(x-u)du \right)$$

3.10.15 Beispiel

Wir betrachten wieder das Werfen zweier Würfel mit den zufälligen Ergebnissen X und Y , d. h. $W_X = W_Y = \{1, 2, 3, 4, 5, 6\}$, $Z = X + Y$ mit $W_Z = \{2, 3, \dots, 12\}$.

$$\begin{aligned} P(X + Y = 6) &= P(X = 1) \cdot P(Y = 5) + P(X = 2) \cdot P(Y = 4) + P(X = 3) \cdot P(Y = 3) \\ &\quad + P(X = 4) \cdot P(Y = 2) + P(X = 5) \cdot P(Y = 1) \\ &= 5 \cdot \frac{1}{36} \\ &= \frac{5}{36} \end{aligned}$$

3.10.16 Beispiel

Die Zufallsvariablen X und Y seien unabhängig und beide gleichverteilt mit den Verteilungsdichten $f = g$:

$$f(x) = \begin{cases} \frac{1}{T} & \text{falls } 0 \leq x \leq T, T > 0 \\ 0 & \text{sonst} \end{cases}$$

Dann ist

$$\begin{aligned} h(x) &= \int_{-\infty}^{\infty} \underbrace{f(u)}_{(\neq 0 \text{ für } 0 \leq u \leq T)} \underbrace{f(x-u)}_{(\neq 0 \text{ für } 0 \leq x-u \leq T \Leftrightarrow x-T \leq u \leq x)} du \\ &= \begin{cases} \int_0^x \frac{1}{T^2} du & \text{falls } 0 \leq x \leq T \\ \int_{x-T}^T \frac{1}{T^2} du & \text{falls } T \leq x \leq 2T \\ 0 & \text{sonst} \end{cases} \\ &= \begin{cases} \frac{x}{T^2} & \text{falls } 0 \leq x \leq T \\ \frac{2T-x}{T^2} & \text{falls } T \leq x \leq 2T \\ 0 & \text{sonst} \end{cases} \end{aligned}$$

3.10.17 Beispiel

X und Y seien unabhängige Zufallsvariable und beide exponentialverteilt mit dem Parameter λ , d.h. $f = g$ mit

$$f(x) = \begin{cases} \frac{1}{\lambda} e^{-\frac{x}{\lambda}} & \text{falls } x \geq 0 \\ 0 & \text{sonst} \end{cases}$$

Dann ist die Verteilungsdichte von $X + Y$

$$\begin{aligned} h(x) &= \int_{-\infty}^{\infty} \underbrace{f(u)}_{(\neq 0 \text{ für } 0 \leq u)} \underbrace{f(x-u)}_{(\neq 0 \text{ für } 0 \leq x-u \Leftrightarrow u \leq x)} du \\ &= \begin{cases} \int_0^x \frac{1}{\lambda} e^{-\frac{u}{\lambda}} \cdot \frac{1}{\lambda} e^{-\frac{x-u}{\lambda}} du & \text{falls } x \geq 0 \\ 0 & \text{sonst} \end{cases} \end{aligned}$$

$$\begin{aligned} &= \begin{cases} \frac{1}{\lambda^2} \int_0^x e^{-\frac{u}{\lambda}} du & \text{falls } x \geq 0 \\ 0 & \text{sonst} \end{cases} \\ &= \begin{cases} \frac{1}{\lambda^2} \cdot x \cdot e^{-\frac{x}{\lambda}} & \text{falls } x \geq 0 \\ 0 & \text{sonst} \end{cases} \end{aligned}$$

3.11 Grenzwertsätze

Bei vielen in der Praxis vorkommenden Zufallsvariablen kann man davon ausgehen, dass sie additiv aus vielen unabhängigen nichtdominierenden Einzeleinflüssen zusammengesetzt sind. In diesem Zusammenhang interessiert die Wahrscheinlichkeitsverteilung der Summe $Z_n = X_1 + X_2 + \dots + X_n$. Gilt $E(X_i) = \mu$ und $Var(X_i) = \sigma^2$ für alle $i = 1, 2, \dots, n$, so gilt $E(Z_n) = n\mu$ und $Var(Z_n) = n\sigma^2$. Für $n \rightarrow \infty$ sind dies unendliche Größen (falls $\mu \neq 0$ bzw. $\sigma \neq 0$). Man betrachtet daher die standardisierte Zufallsvariable $Y_n = \frac{Z_n - E(Z_n)}{\sqrt{Var(Z_n)}} = \frac{Z_n - n\mu}{\sigma\sqrt{n}}$, womit dann $E(Y_n) = 0$ und $Var(Y_n) = 1$ gilt.

Der folgende zentrale Grenzwertsatz besagt, dass Y_n näherungsweise normalverteilt ist. Die Normalverteilung spielt daher in der Statistik eine besondere Rolle.

3.11.1 Satz (Zentraler Grenzwertsatz)

Die Zufallsvariablen X_1, X_2, \dots, X_n seien für jedes $n \in \mathbb{N}$ unabhängig mit derselben Verteilung, d. h. insbesondere $E(X_i) = \mu$ und $Var(X_i) = \sigma^2$ für alle i . Dann gilt für die Verteilungsfunktion der sogenannten standardisierten Summen

$$Y_n = \frac{\sum_{i=1}^n X_i - n\mu}{\sigma\sqrt{n}} = \frac{Z_n - n\mu}{\sigma\sqrt{n}}$$

$$\lim_{n \rightarrow \infty} P(Y_n \leq x) = \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{u^2}{2}} du$$

3.11.2 Bemerkung

Für große n (Faustregel $n \geq 30$) ist Y_n ungefähr $N(0, 1)$ -verteilt, die Summe $Z_n = \sum_{i=1}^n X_i$ ungefähr $N(n\mu, n\sigma^2)$ -verteilt und das arithmetische Mittel $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ ungefähr $N(\mu, \frac{\sigma^2}{n})$ -verteilt.

3.11.3 Beispiel

Ein Autohaus hat ermittelt, dass bei einer Standardabweichung von $\sigma = 1.6$ durchschnittlich alle 2.4 Werkzeuge ein PKW des Typs Fiasko verkauft wird.

Mit welcher Wahrscheinlichkeit werden im Laufe eines Quartals (75 Werkzeuge) mindestens 35 Fiaskos verkauft?

Sei $X_i, i = 1, 2, \dots, n$, die zufällige Zeitspanne zwischen dem Verkauf des $(i-1)$ -ten und des i -ten Fiaskos. Dann ist $Z_n = \sum_{i=1}^n X_i$ die Zeitspanne bis zum Verkauf des n -ten Fiaskos. Gesucht ist also $P(Z_{35} \leq 75)$.

Wir gehen davon aus, dass die X_i unabhängig sind. Es gilt $E(Z_{35}) = 35 \cdot 2.4 = 84$ und $Var(Z_{35}) = 35 \cdot 1.6^2 = 89.6$.

Wegen $n = 35$ können wir nach Bemerkung 3.11.2 annehmen, dass Z_{35} näherungsweise $N(84; 89.6)$ -verteilt ist. Somit ist

$$P(Z_{35} \leq 75) = \Phi\left(\frac{75 - 84}{\sqrt{89.6}}\right) \approx \Phi(-0.95) \approx 0.171$$

Für grobe Abschätzungen ohne große Vorinformationen dient die folgende Tschebyscheffsche Ungleichung.

3.11.4 Satz (Tschebyscheffsche Ungleichung)

Sei X Zufallsvariable mit $E(X) = \mu$ und $Var(X) = \sigma^2$. Dann lässt sich für jedes $c > 0$ die Wahrscheinlichkeit für die Abweichung vom Erwartungswert abschätzen durch

$$P(|X - \mu| \geq c) \leq \frac{\sigma^2}{c^2}$$

denn: Für eine stetige Zufallsvariable X mit der Dichte $f(x)$ gilt:

$$\begin{aligned}
 \text{Var}(X) &= \sigma^2 \\
 &= \int_{-\infty}^{\infty} \underbrace{(x - \mu)^2 f(x)}_{\geq 0} dx \\
 &\geq \int_{\{|x - \mu| \geq c\}} \underbrace{(x - \mu)^2 f(x)}_{\geq c^2} dx \\
 &\geq c^2 \int_{\{|x - \mu| \geq c\}} f(x) dx \\
 &= c^2 P(|X - \mu| \geq c)
 \end{aligned}$$

Der Beweis für den diskreten Fall geht analog.

3.11.5 Beispiel

Der Nennwert der Kapazität von Kondensatoren sei $300\mu F$. Die tatsächlichen Kapazitäten X streuen jedoch mit einer Standardabweichung von $\sigma = 12\mu F$ um den Nennwert.

Eine obere Schranke für die Wahrscheinlichkeit, dass die Kapazität um mehr als 5% vom Nennwert abweicht ist $P(|X - 300| \geq 15) \leq \frac{12^2}{15^2} = 0.64$.

Wissen wir, dass die Kapazität $N(300, 12^2)$ -verteilt ist, so erhalten wir für die Wahrscheinlichkeit einer mehr als 5% Abweichung

$$\begin{aligned}
 P(|X - 300| \geq 15) &= 1 - P(-15 \leq X - 300 \leq 15) \\
 &= 1 - P\left(-\frac{15}{12} \leq \frac{X - 300}{12} \leq \frac{15}{12}\right) \\
 &= 1 - (\Phi(1.25) - \Phi(-1.25)) \\
 &= 2 - 2\Phi(1.25) \\
 &\approx 0.2113
 \end{aligned}$$

Man sieht, dass die zusätzliche Kenntnis der Wahrscheinlichkeitsverteilung mit erheblichem Informationsgewinn verbunden ist.

Wegen des folgenden schwachen Gesetzes der großen Zahlen liegt für große n das arithmetische Mittel (aus Messungen, Stichproben) in der Nähe des Erwartungswertes. Man nimmt daher das arithmetische Mittel auch als Schätzwert für den Erwartungswert.

3.11.6 Satz (Schwaches Gesetz der großen Zahlen)

Die Zufallsvariablen X_1, X_2, \dots, X_n seien alle untereinander unabhängig mit gleichem Erwartungswert μ und gleicher Varianz σ^2 . Dann gilt für jedes $\varepsilon > 0$

$$P\left(\left|\frac{1}{n} \sum_{i=1}^n X_i - \mu\right| \geq \varepsilon\right) \leq \frac{\sigma^2}{n \cdot \varepsilon^2}$$

und damit

$$\lim_{n \rightarrow \infty} P\left(\left|\frac{1}{n} \sum_{i=1}^n X_i - \mu\right| \geq \varepsilon\right) = 0$$

3.11.7 Beispiel

Jemand geht an 250 Tagen stets zu einem zufällig gewählten Zeitpunkt zur Bushaltestelle. Die zufällige

Wartezeit X_i am i -ten Tag bis zum Eintreffen des nächsten Busses habe den Erwartungswert $\mu = 3$ Minuten und eine Varianz von $\sigma^2 = 1.5$ Minuten².

Gesucht ist eine untere Schranke für die Wahrscheinlichkeit, dass die mittlere Wartezeit zwischen 2.5 und 3.5 Minuten liegt.

$$\begin{aligned} P\left(\left|\frac{1}{250} \sum_{i=1}^{250} X_i - 3\right| < 0.5\right) &= 1 - P\left(\left|\frac{1}{250} \sum_{i=1}^{250} X_i - 3\right| \geq 0.5\right) \\ &\geq 1 - \frac{1.5}{250 \cdot 0.5^2} = 0.976 \end{aligned}$$