

Rezepte zur Mathematik

Inhaltsverzeichnis

1 Grundlagen	2
1.1 Logik	2
2 Lineare Algebra	3
2.1 Lineare Abbildungen	3
3 Analysis	6
3.1 Geometrische Bezüge	6
3.2 Rekurrenzen und Summen	7
3.3 Grenzwerte	10
3.4 Differentialrechnung	12
3.5 Zur Lösung von Gleichungen	13
3.6 Differentialgleichungen	16
4 Wahrscheinlichkeitsrechnung	20
4.1 Zufallsgrößen	20

1 Grundlagen

1.1 Logik

Boolesche Algebra für die Aussagenlogik

Man nutze die Umformung $(A \Rightarrow B) \equiv (\neg A \vee B)$, um Tautologien der klassischen Aussagenlogik mittels boolescher Algebra zu bestätigen.

Betrachten wir $(A \Rightarrow B) \wedge A \Rightarrow B$, die Formel zum Modus ponens. Mit den demorganschen Gesetzen findet sich

$$\begin{aligned} (A \Rightarrow B) \wedge A \Rightarrow B &\equiv \neg((\neg A \vee B) \wedge A) \vee B \equiv \neg(\neg A \vee B) \vee \neg A \vee B \\ &\equiv \neg C \vee C \equiv 1, \end{aligned}$$

wobei $C := \neg A \vee B$ abgekürzt wurde. Die Formel ist folglich stets erfüllt, ganz gleich mit welchen Wahrheitswerten die Variablen A, B belegt werden.

Natürliches Schließen

Natürliches Schließen systematisiert gewohnte Beweisführung.

Beispiel 1. Wir wollen zeigen, dass das Quadrat einer geraden Zahl ebenfalls gerade ist. Die Definition

$$2 \mid a \iff \exists n \in \mathbb{Z}: a = 2n.$$

bringt die Behauptung zunächst in die Form

$$(\exists n \in \mathbb{Z}: a = 2n) \implies (\exists n' \in \mathbb{Z}: a^2 = 2n').$$

Nun führen wir den Beweis mittels natürlichem Schließen. Aufgrund der Prämisse haben wir einen Zeugen $n \in \mathbb{Z}$, über den nichts weiter bekannt ist, außer dass $a = 2n$. Sei $n' := 2n^2$. Wir dürfen rechnen

$$a^2 = (2n)^2 = 4n^2 = 2n'.$$

Mit n' wurde ein Zeuge für die rechte Existenzaussage konstruiert. Somit wurde die Konklusion aus der Prämisse abgeleitet. Der Beweis ist erbracht.

2 Lineare Algebra

2.1 Lineare Abbildungen

Inverse Matrix

Die a_{ij} seien Elemente eines Körpers, bspw. reelle oder komplexe Zahlen. Zur Invertierung einer 2×2 -Matrix existiert die Formel

$$A^{-1} = \frac{1}{\det A} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}, \quad A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \quad \det A = a_{11}a_{22} - a_{12}a_{21}.$$

Betrachten wir die Rotationsmatrix

$$R(\varphi) = \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix}.$$

Zunächst erhält man $\det R(\varphi) = 1$ aufgrund des trigonometrischen Pythagoras

$$(\cos \varphi)^2 + (\sin \varphi)^2 = 1.$$

Mit den Symmetrien $\cos(-\varphi) = \cos \varphi$ und $\sin(-\varphi) = -\sin \varphi$ findet sich nun

$$R(-\varphi) = \begin{pmatrix} \cos(-\varphi) & -\sin(-\varphi) \\ \sin(-\varphi) & \cos(-\varphi) \end{pmatrix} = \begin{pmatrix} \cos \varphi & \sin \varphi \\ -\sin \varphi & \cos \varphi \end{pmatrix} = R(\varphi)^{-1}.$$

Lineare Koordinatenabbildungen

Die darstellende Matrix A einer linearen Abbildung $L: \mathbb{R}^n \rightarrow \mathbb{R}^m$ lässt sich sofort gemäß $A\mathbf{e}_k = L(\mathbf{e}_k)$ bestimmen.

Der Raum der linearen Abbildungen $\mathbb{R}^n \rightarrow \mathbb{R}^m$ ist kanonisch isomorph zum Matrizenraum $\mathbb{R}^{m \times n}$. Zu jeder Matrix A gehört somit genau eine lineare Abbildung, und zwar $L(\mathbf{v}) := A\mathbf{v}$. Hat man nun umgekehrt L gegeben, lässt sich damit sofort die zugehörige Matrix A bestimmen. Und zwar gilt $A\mathbf{e}_k = L(\mathbf{e}_k)$. Die $A\mathbf{e}_k$ sind hierbei die Spaltenvektoren der Matrix. Darin ist enthalten, dass eine lineare Abbildung bereits durch die Bilder der Basisvektoren eindeutig bestimmt ist.

Basiswechsel

Umständliche Rechnungen zum Basiswechsel kann man mit der Gleichung $B\mathbf{v}_B = A\mathbf{v}_A$ zu einer kurzen Matrizenrechnung reduzieren.

Sei K ein Körper und V ein Vektorraum über K . Eine Basis $B = (\mathbf{b}_1, \dots, \mathbf{b}_n)$ von V induziert die bijektive lineare Abbildung

$$\Phi_B: K^n \rightarrow V, \quad \Phi_B(\mathbf{e}_k) := \mathbf{b}_k,$$

die ihrerseits die Basis eindeutig charakterisiert. Mit $E = (\mathbf{e}_1, \dots, \mathbf{e}_n)$ ist hierbei die Standardbasis gemeint. Man bezeichnet Φ_B als das durch B induzierte Koordinatensystem. Es ordnet jedem Koordinatenvektor \mathbf{v}_B eins-zu-eins einen Vektor $\mathbf{v} = \Phi_B(\mathbf{v}_B)$ zu.

Für $V = K^n$ kann man $\Phi_B = B$ setzen, indem B als die Matrix betrachtet wird, deren Spalten die Basisvektoren \mathbf{b}_k sind, und zwar genau auf die Weise, wie B als geordnete Basis angegeben wurde.

Seien nun zwei Basen A, B des K^n gegeben. Aufgabe ist es, den Koordinatenvektor \mathbf{v}_A bezüglich Basis B darzustellen. Man erhält die Lösung gemäß

$$\mathbf{v} = \Phi_B(\mathbf{v}_B) = \Phi_A(\mathbf{v}_A) \iff \mathbf{v}_B = \Phi_B^{-1}(\Phi_A(\mathbf{v}_A)),$$

bzw. als Matrizenrechnung

$$\mathbf{v} = B\mathbf{v}_B = A\mathbf{v}_A \iff \mathbf{v}_B = B^{-1}A\mathbf{v}_A.$$

Die Verkettung $T_B^A := B^{-1}A = \Phi_B^{-1} \circ \Phi_A$ ist immer eine lineare Abbildung zwischen Koordinatenräumen, und somit unter allen Umständen als Matrix darstellbar. Man bezeichnet sie als *Transformationsmatrix*.

Mit diesen Überlegungen ist eigentlich auch bereits fast alles Wichtige gesagt, um Darstellungsmatrizen und ihre Transformation bei Basiswechsel verstehen zu können.

Darstellungsmatrizen

Beim Basiswechsel ist die Darstellungsmatrix zu transformieren, was man mittels Matrizenrechnung kurzfassen kann.

Gegeben seien Basen A, B von V und C, D von W . Es gilt

$$\begin{aligned} \mathbf{v} &= A\mathbf{v}_A = B\mathbf{v}_B, \\ \mathbf{w} &= C\mathbf{w}_C = D\mathbf{w}_D. \end{aligned}$$

Für eine lineare Abbildung $f: V \rightarrow W$ gilt

$$\begin{aligned} \mathbf{w} = f(\mathbf{v}) &\iff C\mathbf{w}_C = f(A\mathbf{v}_A) \\ &\iff \mathbf{w}_C = (C^{-1} \circ f \circ A)\mathbf{v}_A. \end{aligned}$$

Die Darstellungsmatrix definiert man daher als

$$M_C^A(f) := C^{-1} \circ f \circ A.$$

Man betrachte nun die Kette

$$\mathbf{v}_B \xrightarrow{A^{-1}B} \mathbf{v}_A \xrightarrow{M_C^A(f)} \mathbf{w}_C \xrightarrow{D^{-1}C} \mathbf{w}_D.$$

Demzufolge gilt

$$M_D^B(f) = D^{-1}C M_C^A(f)A^{-1}B = T_D^C M_C^A(f)T_A^B.$$

Lineare Unabhängigkeit

Zum Vektorraum K^n ist $B = (\mathbf{b}_1, \dots, \mathbf{b}_n)$ genau dann eine Basis, wenn $\det(B) \neq 0$ gilt.

Das System $B = (\mathbf{b}_1, \dots, \mathbf{b}_n)$ ist gemäß Definition genau dann linear unabhängig, wenn

$$x_1\mathbf{b}_1 + \dots + x_n\mathbf{b}_n = 0 \implies \forall k: x_k = 0.$$

Kompakt geschrieben lautet diese Aussage

$$B\mathbf{x} = 0 \implies \mathbf{x} = 0$$

mit $\mathbf{x} = (x_1, \dots, x_n)$. Äquivalent zu ihr ist

$$\text{Kern}(B) = \{0\}.$$

Wir haben andererseits

$$\text{Kern}(B) = \{0\} \iff B \text{ bijektiv} \iff \det(B) \neq 0.$$

Der Test auf lineare Unabhängigkeit ist damit auf die Berechnung einer Determinante zurückgeführt, was für ein konkretes B automatisch durchgeführt werden kann.

3 Analysis

3.1 Geometrische Bezüge

Polarkoordinaten

Bei der Umrechnung in Polarkoordinaten wird das Auftreten umständlicher Fallunterscheidungen verhindert, indem man $x = r \cos \varphi$ nach φ umstellt und das Vorzeichen $s(y)$ hinzufügt.

Die Umrechnung von Polarkoordinaten in kartesische Koordinaten ist mit

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} r \cos \varphi \\ r \sin \varphi \end{pmatrix}$$

müheles durchführbar. Zur händischen Umrechnung von kartesischen Koordinaten in Polarkoordinaten nutzt man am besten die Formel

$$\varphi = s(y) \arccos\left(\frac{x}{r}\right), \quad r = \sqrt{x^2 + y^2},$$

wobei $s(0) = 1$ ist und sonst $s(y) = \operatorname{sgn}(y)$ gilt. Auf diese Weise entgeht man komplizierten Fallunterscheidungen.

Rotationsmatrizen

Die Multiplikation von Rotationsmatrizen bietet eine Merkhilfe für die Additionstheoreme.

Es sei $c_x := \cos x$ und $s_x := \sin x$. Es stellt $R_x := \begin{pmatrix} c_x & -s_x \\ s_x & c_x \end{pmatrix}$ die Rotationsmatrix für die Rotation um den Winkel x gegen den Uhrzeigersinn dar. Die Additionstheoreme lassen sich nun via $R_{x+y} = R_x R_y$ ablesen an der Matrizenmultiplikation

$$\begin{pmatrix} c_{x+y} & -s_{x+y} \\ s_{x+y} & c_{x+y} \end{pmatrix} = \begin{pmatrix} c_x & -s_x \\ s_x & c_x \end{pmatrix} \begin{pmatrix} c_y & -s_y \\ s_y & c_y \end{pmatrix} = \begin{pmatrix} c_x c_y - s_x s_y & -c_x s_y - c_y s_x \\ c_x s_y + c_y s_x & c_x c_y - s_x s_y \end{pmatrix}.$$

Konstruktion konformer Abbildungen

Zur Konstruktion konformer Abbildungen der Ebene nutze man holomorphe Funktionen.

Jeder komplexen Zahl ist gemäß

$$\Phi(a + bi) = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

oder äquivalent

$$\Phi(re^{i\varphi}) = r \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix}$$

genau eine Matrix zugeordnet. Die Abbildung Φ ist ein Isomorphismus vom Körper der komplexen Zahlen in einen Körper, der eine Unterstruktur des Matrizenrings darstellt.

Diese Beziehung schafft eine Verbindung zwischen dem Rechnen mit komplexen Zahlen und Konzepten der linearen Algebra. Sie sagt aus, dass eine komplexe Zahl als eine als *Drehskalierung* bezeichnete lineare Abbildung betrachtet werden kann. Die Multiplikation von komplexen Zahlen entspricht der Verkettung der ihnen entsprechenden Abbildungen.

Man nennt $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ eine *konforme* Abbildung, wenn sie differenzierbar ist und ihre Ableitung Df an jeder Stelle eine nichtverschwindende Drehskalierung ist. Nun ist die Ableitung von $f(x, y) = (u(x, y), v(x, y))$ ja genau dann eine Drehskalierung, wenn

$$Df = \begin{pmatrix} \partial_x u & \partial_y u \\ \partial_x v & \partial_y v \end{pmatrix} = \begin{pmatrix} \partial_x u & -\partial_x v \\ \partial_x v & \partial_x u \end{pmatrix}.$$

Diese Gleichung zwischen Matrizen beschreibt aber nichts anderes als die Cauchy-Riemann-Gleichungen. Wir setzen daher $z := x + yi$ und $f(z) := u + vi$. Das heißt, f ist genau dann eine konforme Abbildung, wenn f holomorph ist und ihre Ableitung keine Nullstellen besitzt.

3.2 Rekurrenzen und Summen

Vollständige Induktion

Eine Aussage $A(n)$ für $n \in \mathbb{Z}$, $n \geq n_0$ sei offensiv hergeleitet worden unter Nutzung von höheren oder heuristischen Methoden, deren Gültigkeit nicht zweifelsfrei klar ist. Gelingt anschließend ein Beweis durch vollständige Induktion, braucht man sich darum nicht weiter kümmern und kann voranschreiten.

Das Verfahren der vollständigen Induktion verläuft folgendermaßen. Im *Induktionsschritt* schlussfolgert man, dass die *Induktionsvoraussetzung* $A(n)$ die Aussage $A(n+1)$ nach sich zieht. Kann man zudem einen *Induktionsanfang* $A(n_0)$ bestätigen, was oft trivial ist, gilt die Aussage $A(n)$ für alle $n \geq n_0$ als bewiesen.

Beispiel 1. Die Summation kann man definieren durch

$$\sum_{k=0}^0 a_k := a_0, \quad \sum_{k=0}^{n+1} a_k := a_{n+1} + \sum_{k=0}^n a_k.$$

Zu beweisen sei die Aussage $\sum_{k=0}^n ca_k = c \sum_{k=0}^n a_k$ für alle $n \geq 0$, wobei c eine Konstante ist. Der Induktionsanfang ist mit $\sum_{k=0}^0 ca_k = ca_0 = c \sum_{k=0}^0 a_k$ trivial erbracht. Im Schritt rechnet man

$$\sum_{k=0}^{n+1} ca_k = ca_{n+1} + \sum_{k=0}^n ca_k \stackrel{(IV)}{=} c(a_{n+1} + \sum_{k=0}^n a_k) = c \sum_{k=0}^{n+1} a_k,$$

wobei (IV) gemäß Induktionsvoraussetzung gilt.

Beispiel 2. Zu beweisen sei $\sum_{k=0}^n k = \frac{n}{2}(n+1)$. Der Anfang ist trivial. Im Schritt rechnet man

$$\begin{aligned} \sum_{k=0}^{n+1} k &= n+1 + \sum_{k=0}^n k \stackrel{(IV)}{=} n+1 + \frac{n}{2}(n+1) \\ &= (n+1)\left(1 + \frac{n}{2}\right) = \frac{n+1}{2}(n+2). \end{aligned}$$

Zur Auffindung von Formeln kann man alle erdenklichen Methoden zum Einsatz bringen, obgleich diese unzuverlässig oder heuristisch sind. Man darf Umformungen benutzen, bei denen nicht klar ist, unter welchen Prämissen diese gültig sind. Sogar vage und skizzenhafte Argumentationen sind nicht unbedingt tabu.

Teleskopsumme

Bezüglich $(\Delta s)_k := s_{k+1} - s_k$ gilt die Teleskopsumme $\sum_{k=m}^{n-1} (\Delta s)_k = s_n - s_m$. Um beispielsweise eine Identität $\sum_{k=0}^{n-1} a_k = s_n$ zu beweisen, genügt es, $s_0 = 0$ und $(\Delta s)_k = a_k$ zu bestätigen.

Für die Summe der Dreieckszahlen gilt bekanntlich die Identität

$$\sum_{k=0}^{n-1} k = \frac{n}{2}(n-1).$$

Es sei $s_n := \frac{n}{2}(n-1)$. Man erhält direkt ersichtlich $s_0 = 0$, und kurzerhand

$$(\Delta s)_k = \frac{k+1}{2}k - \frac{k}{2}(k-1) = \frac{1}{2}(k^2 + k - k^2 + k) = k.$$

Wir haben die Formel somit bereits bewiesen, ohne eine vollständige Induktion zu bemühen.

Lineare Rekurrenzen

Jede lineare Rekurrenz lässt sich durch Hilfsvariablen und homogene Koordinaten in die Form eines homogenen linearen Systems erster Ordnung bringen, das in vektorieller Form als $\mathbf{v}_{n+1} = A\mathbf{v}_n$ ausgedrückt werden kann. Die Lösung ist daher $\mathbf{v}_n = A^n\mathbf{v}_0$, wobei die Matrixpotenz A^n mittels Eigenwerttheorie bzw. Theorie der Matrixfunktionen bestimmt werden kann.

Die lineare Rekurrenz $x_{n+1} = ax_n + b$ wird mittels homogener Koordinaten zum System

$$\begin{pmatrix} x_{n+1} \\ 1 \end{pmatrix} = \begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_n \\ 1 \end{pmatrix}.$$

Bei der linearen Rekurrenz $x_{n+1} = ax_n + bx_{n-1} + c$ setzen wir $y_n = x_{n-1}$ und erhalten somit das System

$$\begin{pmatrix} x_{n+1} \\ y_{n+1} \\ 1 \end{pmatrix} = \begin{pmatrix} a & b \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x_n \\ y_n \end{pmatrix} + \begin{pmatrix} c \\ 0 \end{pmatrix}.$$

Mittels homogener Koordinaten wird dieses zu

$$\begin{pmatrix} x_{n+1} \\ y_{n+1} \\ 1 \end{pmatrix} = \begin{pmatrix} a & b & c \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_n \\ y_n \\ 1 \end{pmatrix}.$$

3.3 Grenzwerte

Fixpunktiteration

Wann immer eine Rekurrenz der Form $x_{n+1} = \varphi(x_n)$ auftritt, betrachte man die Fixpunktgleichung $x = \varphi(x)$. Ob x_n gegen den Fixpunkt x konvergiert, hängt allerdings von den Eigenschaften von φ ab.

Man kann sich die Frage stellen, welchen Wert der endlose Kettenbruch

$$a + \frac{b}{a + \frac{b}{a + \frac{b}{a + \dots}}}$$

für feste Zahlen a, b besitzt. Zunächst muss geklärt sein, was damit gemeint sein soll. Man beginnt bei einem Startwert x_0 und betrachtet die Rekurrenz

$$x_{n+1} = a + \frac{b}{x_n}.$$

Wir abstrahieren durch Einführung der Funktion φ , so dass

$$x_{n+1} = \varphi(x_n)$$

gilt. Wie man mühelos bestätigen kann, handelt es sich bei φ für $a > 0$ und $b > 0$ um eine stetige Selbstabbildung auf $\mathbb{R}_{>0}$. Nehmen wir nun an, die Folge (x_n) konvergiert gegen eine Zahl x . Dann muss gelten

$$x = \lim_{n \rightarrow \infty} x_n = \lim_{n \rightarrow \infty} x_{n+1} = \lim_{n \rightarrow \infty} \varphi(x_n) = \varphi(\lim_{n \rightarrow \infty} x_n) = \varphi(x),$$

wobei φ aufgrund der Stetigkeit aus der Grenzwertbildung herausgezogen werden durfte. Die Gleichung $x = \varphi(x)$ bezeichnet man als *Fixpunktgleichung*. Ihre Lösungen bezeichnet man als *Fixpunkte* von φ . Speziell für den Kettenbruch ergibt sich die quadratische Gleichung

$$x^2 - ax - b = 0.$$

So erhält man $1 + \sqrt{2}$ für $a = 2$ und $b = 1$. Für $a = 1$ und $b = 1$ erhält man den goldenen Schnitt.

Folgen als Reihen

Jede Folge kann durch Bildung der Differenzenfolge als Reihe aufgefasst werden, womit sämtliche Konvergenzkriterien für Reihen zur Verfügung stehen.

Zu jeder Folge (a_n) kann man die durch

$$(\Delta a)_k := a_{k+1} - a_k$$

definierte Differenzenfolge bilden. Umgekehrt lässt sich a_n über die Teleskopsumme

$$a_n - a_0 = \sum_{k=0}^{n-1} (\Delta a)_k$$

zurückgewinnen. Hiermit ist (a_n) als eine Folge von Partialsummen dargestellt, kann also als Reihe aufgefasst werden.

Majorantenkriterium

Kann man zu einer Reihe von Elementen eines Banachraums eine konvergente Majorante finden, dann muss die Reihe absolut konvergent sein.

Sei (a_k) eine Folge von Elementen eines Banachraums. Eine Reihe $\sum_{k=0}^n b_k$ reeller Zahlen $b_k \geq 0$ mit $\|a_k\| \leq b_k$ für alle $k \geq k_0$ nennt man eine *Majorante* der Reihe $\sum_{k=0}^n a_k$.

Beispiel. Die absolute Konvergenz der Reihe

$$\sum_{k=1}^n \frac{1}{4k^2 - 2k}$$

bestätigt die Majorante mit $b_k = 1/k^2$, denn

$$\begin{aligned} \left| \frac{1}{4k^2 - 2k} \right| \leq \frac{1}{k^2} &\iff k^2 \leq |4k^2 - 2k| \\ \iff k \leq |4k - 2| &\iff k^2 \leq (4k - 2)^2 \\ \iff 0 \leq 15k^2 - 16k + 4. \end{aligned}$$

Diese Ungleichung ist für alle $k \in \mathbb{Z}$ erfüllt, und somit erst recht für $k \geq 1$.

Ausklammern

Manchmal ist es zielführend, einen Term auszuklammern.

Ein elementares Beispiel bietet die Rechnung

$$\frac{x}{x+1} = \frac{x}{x} \frac{1}{1 + \frac{1}{x}} = \frac{1}{1 + \frac{1}{x}} \xrightarrow{x \rightarrow \infty} \frac{1}{1+0} = 1.$$

Etwas schwieriger ist es, die Konvergenz $(x^n + 1)^{1/n} - x \rightarrow 0$ für $n \geq 2$ bezüglich $x \rightarrow \infty$ zu bestätigen. Die binomische Reihe ist allerdings nicht direkt anzuwenden, da das Argument bei großen x außerhalb des Konvergenzbereichs liegt. Klammert man aber x aus, löst sich diese Komplikation magisch in Luft auf. Es findet sich

$$\begin{aligned} (x^n + 1)^{1/n} - x &= x \left(\left(1 + \frac{1}{x^n} \right)^{1/n} - 1 \right) \\ &= x \left(\sum_{k=0}^{\infty} \binom{1/n}{k} x^{-nk} - 1 \right) = \sum_{k=1}^{\infty} \binom{1/n}{k} \frac{1}{x^{nk-1}} \xrightarrow{x \rightarrow \infty} 0, \end{aligned}$$

denn wegen $nk > 1$ konvergiert jeder Summand gegen null. Im letzten Schritt ist streng genommen noch zu bestätigen, dass die Grenzwerte vertauscht werden dürfen. Bezüglich der Substitution $\frac{1}{x} = u$ ist gemeint

$$\lim_{u \rightarrow 0} \lim_{m \rightarrow \infty} \sum_{k=1}^m \binom{1/n}{k} u^{nk-1} = \lim_{m \rightarrow \infty} \lim_{u \rightarrow 0} \sum_{k=1}^m \binom{1/n}{k} u^{nk-1}.$$

Abgesichert wird dies durch die gleichmäßige Konvergenz der Partialsummenfolge der Potenzreihe auf den kompakten Teilmengen ihres Konvergenzintervalls. Dieser Umstand muss nicht extra nachgewiesen werden, er gilt für jede Potenzreihe.

3.4 Differentialrechnung**Ableitungen verifizieren**

Es ist zuträglich, nach dem Bestimmen der Ableitung im Anschluss mit einem Plotter oder einem CAS die Probe durchzuführen.

Die Ableitung einer reellen Funktion f ist definiert als

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

Möchte man erfahren, ob sie unter Anwendung der Ableitungsregeln richtig ermittelt wurde, kann man die Probe machen, indem der Differenzenquotient

$$D_h f(x) = \frac{f(x+h) - f(x)}{h}$$

an einer konkreten Stelle x für ein kleines h numerisch berechnet und mit $f'(x)$ verglichen wird.

Da dies recht umständlich ist, ist es sinnvoll, diese Aufgabe einem Funktionsplotter zu überlassen. Numerisch günstiger ist es, den Differenzenquotient nicht naiv gemäß der Definition zu berechnen, sondern als

$$D_h f(x) = \frac{f(x+h) - f(x-h)}{2h}.$$

Mit dem Plotter kann man schließlich

$$10^n (D_h f(x) - f'(x)) \approx 0$$

für $n \in \{0, 1, 2, \dots\}$ prüfen.

3.5 Zur Lösung von Gleichungen

Umformungen verifizieren

Mit einem Plotter für implizite Kurven lassen sich Gleichungsumformungen verifizieren.

Die Gleichung $y = \frac{x}{\sqrt{x^2+1}}$ ist nach x umzustellen. Wir beschränken uns hierbei auf $x \geq 0$ und $y \geq 0$, da y als Funktion von x punktsymmetrisch zum Koordinatenursprung ist. Unter dieser Einschränkung gilt

$$\begin{aligned} y = \frac{x}{\sqrt{x^2+1}} &\iff y^2 = \frac{x^2}{x^2+1} \iff y^2(x^2+1) = x^2 \\ &\iff y^2x^2 + y^2 = x^2 \iff y^2x^2 - x^2 = -y^2 \\ &\iff x^2(y^2 - 1) = -y^2 \iff x^2(1 - y^2) = y^2 \\ &\iff x^2 = \frac{y^2}{1-y^2} \iff x = \frac{y}{\sqrt{1-y^2}}. \end{aligned}$$

Lässt man die Gleichung mit einem Funktionsplotter als implizite Kurve zeichnen, darf dem Graph durch die jeweilige Äquivalenzumformung keine Veränderung widerfahren. Auf diese Weise lässt sich die Richtigkeit jedes Umformungsschritts grafisch prüfen.

Bisektion

Bisektion ist ein allgemeines und automatisierbares Verfahren zur Ermittlung von Nullstellen.

Gegeben ist eine Bestimmungsgleichung $T_1(x) = T_2(x)$ in der Variable x , deren Lösungen zu ermitteln sind. Die Problemstellung ist gleichbedeutend damit, die Nullstellen der reellen Funktion $f(x) = T_2(x) - T_1(x)$ zu ermitteln. Angenommen, f ist stetig und für $a < b$ ist $f(a)f(b) < 0$, das heißt, $f(a)$ und $f(b)$ haben unterschiedliches Vorzeichen. Dann muss f gemäß dem Zwischenwertsatz mindestens eine Nullstelle im Intervall $[a, b]$ besitzen.

Wir betrachten nun den folgenden Vorgang. Der Mittelwert $c := (a+b)/2$ wird berechnet. Sollte $f(c) = 0$ sein, endet der Vorgang mit der gefundenen Nullstelle c . Ist $f(a)f(c) < 0$, muss eine Nullstelle in $[a, c]$ liegen und es wird $b := c$ gesetzt, andernfalls in $[c, b]$ und es wird $a := c$ gesetzt. Auf das neue Intervall $[a, b]$ wird der Vorgang abermals angewendet, solange bis der Abstand $b - a$ hinreichend klein geworden ist. Nun ist $a \approx b$ eine Näherung für die Nullstelle.

Was ist nun, wenn die Funktion mehrere Nullstellen besitzt? Dann betrachtet man eine hinreichend feine Aufteilung von $[a, b]$ in die n Teilintervalle $[a_k, b_k]$ mit $a_k = a + hk$ und $b_k = a_k + h$, wobei $h = (b - a)/n$. Ist $f(a_k)f(b_k) < 0$, kommt Bisektion zur Anwendung.

Bisher erfuhren solcherlei Funktionen keine Beachtung, die eine Nullstelle besitzen, obwohl kein Vorzeichenwechsel vorliegt. Setzen wir dafür voraus, dass f einmal differenzierbar ist und eine stetige Ableitung besitzt. Wenn kein Vorzeichenwechsel vorhanden ist, schneidet f die waagerechte Achse nicht, sondern berührt sie nur, womit f bei der Nullstelle einen Extrempunkt besitzt. An einem solchen muss aber zwingend $f'(x) = 0$ gelten, und zwar mit Vorzeichenwechsel. Ergo kann man wie folgt vorgehen. Man ermittelt zusätzlich die Nullstellen der Ableitung f' mittels Bisektion und schöpft davon alle die ab, die ebenfalls Nullstellen von f sind.

Wir beuten das Verfahren nun aus, indem wir es zum Zeichnen einer durch $f(x, y) = 0$ beschriebenen Kurve nutzen. Dazu wird die y -Achse durch die $y_k = y_0 + kd$ fein abgetastet. Zu jedem y_k bestimmt man Lösungen x der Gleichung $f(x, y_k) = 0$ wie beschrieben mittels Bisektion. Zur Verbesserung überlagert man die Zeichnung mit der analogen Zeichnung bezüglich Abtastung der x -Achse mit jeweils Lösungen y der Gleichung $f(x_k, y) = 0$.

Abschätzung der Lage von Lösungen

Auch wenn bei einer Gleichung kein Lösungsansatz zu finden ist, lassen sich oft Abschätzungen angeben, wo die Lösungen befindlich sein müssen.

Zum Beispiel wollen wir uns Klarheit über die reellen Lösungen des Gleichungssystems

$$x^2 + 1 = y,$$

$$xy^3 = 1$$

verschaffen. Wir eliminieren y in der zweiten Gleichung mittels der ersten Gleichung und erhalten

$$0 = P(x) := x(x^2 + 1)^3 - 1 = x^7 + 3x^5 + 3x^3 + x - 1.$$

Die Koeffizientenfolge des Polynoms $P(x)$ besitzt nur einen Vorzeichenwechsel, womit laut der descartesschen Regel [1][2] genau eine positive Lösung existieren muss. Die Koeffizientenfolge des Polynoms $P(-x)$ besitzt überhaupt keinen Vorzeichenwechsel, womit es keine negativen Lösungen geben kann.

Betrachten wir nun $x > 1$. Dann ist $x^2 > 1$ und damit $y = x^2 + 1 > 2$. Somit ist $1/x = y^3 > 8$. Man erhält $x < 1/8$ im Widerspruch zu $x > 1$, womit die Lösung im Intervall $[0, 1]$ liegen muss.

Man könnte sich nun um bessere Abschätzungen bemühen, erhält eine Folge von Abschätzungen jedoch blitzartig mit dem Bisektionsverfahren, weil $P(0) = -1$ negativ und $P(1) = 7$ positiv ist.

3.6 Differentialgleichungen

Euler-Verfahren

Das Euler-Verfahren ist eine unkompliziert zu handhabende und automatisierbare Methode zur numerischen Lösung eines Anfangswertproblems.

Das Auffinden symbolischer Lösungen von Anfangswertproblemen kann schwierig sein. Manchmal muss man dafür neue Funktionen definieren, die zu den speziellen Funktionen gezählt werden. Bei manchen Gleichungen stellt sich die Frage, ob es denn überhaupt möglich ist, eine symbolische Lösung zu gewinnen.

Allerdings existieren allgemeine Verfahren zur numerischen Lösung von Anfangswertproblemen am Computer. Hiermit lässt sich ggf. auch eine numerische Probe bereits gefundener symbolischer Lösungen vornehmen. Das einfachste wie auch prototypische dieser Verfahren ist das Euler-Verfahren, welches im Folgenden erläutert werden soll.

Betrachten wir

$$x' = f(t, x), \quad x'(t_0) = x_0.$$

Man spricht vom allgemeinen Anfangswertproblem erster Ordnung. Den Differentialquotient approximieren wir nun durch einen Differenzenquotient:

$$\frac{x(t+h) - x(t)}{h} \approx f(t, x).$$

Umformung dieser Gleichung führt zu

$$x(t+h) \approx x(t) + hf(t, x).$$

Demnach ergibt sich eine näherungsweise Lösung des Anfangswertproblems nun als iterative Lösung der Differenzengleichung

$$\begin{aligned} t_k &= t_0 + hk, \\ x_{k+1} &= x_k + hf(t_k, x_k). \end{aligned}$$

Diese Iteration wird als *Euler-Verfahren* bezeichnet.

Nun interessieren uns auch Dgln. höherer Ordnung, und ganz besonders die physikalisch wichtigen Dgln. zweiter Ordnung. Jede Dgl. höherer Ordnung ist aber als System erster Ordnung formulierbar. Vielen wichtigen Modellen liegt

außerdem ein System erster Ordnung zugrunde. Es ist nun so, dass das Euler-Verfahren ohne besondere Mühe auf ein System erster Ordnung übertragbar ist.

Das Anfangswertproblem zu einem System von zwei Dgln. hat z. B. die allgemeine Form

$$\begin{aligned}x' &= f_1(t, x, y), & x'(t_0) &= x_0, \\y' &= f_2(t, x, y), & y'(t_0) &= y_0.\end{aligned}$$

Fasst man die beiden Variablen x, y zu einem Vektor $\mathbf{x} = (x, y)$ zusammen und definiert

$$f(t, \mathbf{x}) := \begin{pmatrix} f_1(t, \mathbf{x}) \\ f_2(t, \mathbf{x}) \end{pmatrix},$$

bekommt das Anfangswertproblem die kompakte vektorielle Form

$$\mathbf{x}' = f(t, \mathbf{x}), \quad \mathbf{x}'(t_0) = \mathbf{x}_0.$$

Diese Gleichung ist von der gleichartigen Form wie die anfangs beschriebene. Man kann hier im Weiteren analog vorgehen und findet

$$\begin{aligned}t_k &= t_0 + hk, \\ \mathbf{x}_{k+1} &= \mathbf{x}_k + hf(t_k, \mathbf{x}_k).\end{aligned}$$

So unscheinbar diese Iteration auch aussieht, ermöglicht sie die numerische Lösung einer außerordentlichen Vielfalt an Gleichungen.

Es folgt noch eine alternative Herleitung des Euler-Verfahrens. Beide Seiten der Dgl. $x' = f(x, t)$ werden von a bis b integriert. Das macht

$$x(b) - x(a) = \int_a^b f(t, x) dt.$$

Die Differentialgleichung ist hiermit zu einer äquivalenten Integralgleichung umgeformt.

Ist der Abstand zwischen den Stellen a und b klein genug, erhält man die genügend gute Näherung

$$\int_a^b f(t, x) dt \approx (b - a)f(a, x).$$

Setzen wir $a := t$ und $b := t + h$ ein, finden wir zum zweiten Mal die näherungsweise Gleichung.

Fundamentallösung

Ein Weg zur Lösung einer inhomogenen linearen gewöhnlichen Differentialgleichung mit konstanten Koeffizienten besteht in der Auffindung einer Fundamentallösung.

Die Faltung zweier Funktionen f, g ist definiert als

$$(f * g)(x) := \int_{\mathbb{R}} f(x - \xi)g(\xi) \, d\xi.$$

Die Faltung verhält sich, falls sie existiert, wie eine bilineare Abbildung. Das heißt, für Funktionen f, f_1, f_2, g und eine Konstante c gilt

$$\begin{aligned}(f_1 + f_2) * g &= f_1 * g + f_2 * g, \\ (cf) * g &= c(f * g).\end{aligned}$$

Außerdem gilt die Ableitungsregel

$$(f * g)' = f' * g = f * g'.$$

Sei nun L ein linearer Differentialoperator mit konstanten Koeffizienten. Unter Nutzung der genannten Regeln findet sich die Regel

$$L(g * f) = (Lg) * f.$$

Betrachten wir die Aufgabe, die inhomogene Dgl. $y' - y = f$ zu lösen, wobei f eine hinreichend gutartige gegebene Funktion sein soll. Zur Abstraktion notiert man die linke Seite mit dem linearen Differentialoperator $L = D^1 - D^0$, so dass $Ly = f$ gilt.

Ein G , das die Gleichung $LG = \delta$ erfüllt, bezeichnet man als *Fundamentallösung*. Mit δ ist hierbei die Delta-Distribution gemeint. Die Auffindung des richtigen G gelingt durch Beachtung von Zusatzbedingungen, wobei man dann von einer *greenschen Funktion* spricht, obwohl es sich im Allgemeinen auch um eine Distribution handeln kann. Mit ihr lässt sich die Lösung der Differentialgleichung als Faltung $y = G * f$ darstellen, denn es gilt

$$Ly = L(G * f) = (LG) * f = \delta * f = f.$$

An Stellen $x \neq 0$ muss $LG(x) = 0$ gelten, womit G die Lösung der homogenen Gleichung sein muss, das ist $G(x) = Ce^x$. Die Konstante C ist hierbei allerdings

für $x < 0$ und für $x > 0$ separat zu ermitteln. Außerdem muss die Anfangsbedingung $y_0 = y(x_0)$ als Randbedingung betrachtet werden. Man gelangt zu

$$G(x - \xi) = C_1 e^{x-\xi} \text{ für } x_0 \leq x < \xi,$$

$$G(x - \xi) = C_2 e^{x-\xi} \text{ für } \xi < x < \infty.$$

Um $LG(x - \xi) = \delta(x - \xi)$ zu erhalten, muss an der Stelle $x = \xi$ ein Einheitsprung befindlich sein, was $C_2 = C_1 + 1$ aufgrund $e^0 = 1$ nach sich zieht. Mit $f(x) = 0$ für $x < x_0$ findet man nun

$$\begin{aligned} y &= \int_{x_0}^x G(x - \xi) f(\xi) d\xi + \int_x^{\infty} G(x - \xi) f(\xi) d\xi \\ &= C_2 e^x \int_{x_0}^x e^{-\xi} f(\xi) d\xi + C_1 e^x \int_x^{\infty} e^{-\xi} f(\xi) d\xi \\ &= C_1 e^x \int_{x_0}^{\infty} e^{-\xi} f(\xi) d\xi + e^x \int_{x_0}^x e^{-\xi} f(\xi) d\xi. \end{aligned}$$

Aus der Anfangsbedingung $y_0 = y(x_0)$ erhält man

$$y_0 e^{-x_0} = C_1 \int_{x_0}^{\infty} e^{-\xi} f(\xi) d\xi.$$

Es findet sich schließlich

$$y = y_0 e^{x-x_0} + e^x \int_{x_0}^x e^{-\xi} f(\xi) d\xi.$$

Probe durch Einsetzen in die Dgl. bestätigt, dass die richtige Lösung gefunden ist. Wir haben dieselbe Lösung gefunden, die der Ansatz *Variation der Konstanten* zum Resultat hat.

4 Wahrscheinlichkeitsrechnung

4.1 Zufallsgrößen

Zufallsgröße

Eine Zufallsgröße $X: \Omega \rightarrow \Omega'$ schafft eine kausale Verbindung, dergestalt dass ein Ergebnis $\omega \in \Omega$ zum Ergebnis $X(\omega)$ wird. Ein Ereignis $A' \subseteq \Omega'$ tritt daher genau dann ein, wenn das Urbild $X^{-1}(A')$ eintritt. Dieser Umstand induziert die Definition

$$P_X(A') := P(X^{-1}(A')) = P(X \in A').$$

Eine Zufallsgröße, was ist das? Eine Zufallsgröße kann man sich zunächst einfach als eine Abbildung $X: \Omega \rightarrow \Omega'$ zwischen Ergebnismengen vorstellen. Sei bspw.

$$\Omega := \{(w_1, w_2) \mid w_1, w_2 \in \{1, \dots, 6\}\}$$

die Menge der Ergebnisse des Wurfs zweier gewöhnlicher Würfel. Das heißt, wurde mit dem ersten Würfel eine Drei gewürfelt, und mit dem zweiten eine Fünf, ist das Ergebnis $(3, 5)$. Jedes elementare Ereignis $\{(w_1, w_2)\}$ besitzt offenbar die gleiche Wahrscheinlichkeit

$$P(\{(w_1, w_2)\}) = \frac{1}{|\Omega|} = \frac{1}{36}.$$

Für ein beliebiges Ereignis A gilt daher

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

Ein gutes Beispiel für eine Zufallsgröße ist die Summe der Augenzahlen, also

$$X((w_1, w_2)) := w_1 + w_2.$$

Des Pudels Kern liegt nun in der Beantwortung der Frage, wie wahrscheinlich ein aus Funktionswerten von X bestehendes Ereignis ist.

Ein elementares Ereignis $\{x\}$ tritt doch genau dann ein, wenn x der Funktionswert $x = X(\omega)$ zum Ergebnis ω ist. Wurde bspw. das Ergebnis $\omega = (3, 5)$ gewürfelt, ist das elementare Ereignis

$$\{X(\omega)\} = \{3 + 5\} = \{8\}$$

eingetreten. Ein elementares Ereignis $\{x\}$ tritt also genau dann ein, wenn das Ergebnis ω im Urbild $X^{-1}(x)$ liegt, für das sich die Schreibweise

$$X^{-1}(x) = \{X = x\}$$

eingebürgert hat. Demnach stimmt die Wahrscheinlichkeit von $\{x\}$ mit der des Urbildereignisses $\{X = x\}$ überein. Das heißt, es gilt

$$P_X(\{x\}) = P(X^{-1}(x)) = P(\{X = x\}) = P(X = x).$$

Beispielsweise ist

$$\{X = 8\} = \{(2, 6), (6, 2), (3, 5), (5, 3), (4, 4)\}.$$

Damit ergibt sich

$$P(X = 8) = \frac{|\{X = 8\}|}{|\Omega|} = \frac{5}{36}$$

als Wahrscheinlichkeit der Augensumme acht.

Literatur

- [1] »Vorzeichenregel von Descartes«. In: Wikipedia.
- [2] Ilja Bronstein u. a.: »Taschenbuch der Mathematik«. Harri Deutsch, 7. Auflage. Abschnitt 1.6.3.2, S. 44–45, »Gleichungen mit reellen Koeffizienten«.