

Skript zum Mathematik-Studium

7. August 2017

Inhaltsverzeichnis

1 Grundlagen	3
1.1 Prädikatenlogik	3
1.2 Mengenlehre	3
1.3 Funktionen	4
2 Lineare Algebra	5
2.1 Vektorräume	5
2.2 Lineare Abhängigkeit	5
2.3 Lineare Abbildungen	6
2.4 Darstellungsmatrizen	8
2.5 Gruppen	9
2.6 Quadratische Matrizen	10
3 Algebra	11
3.1 Gruppen	11
3.2 Zyklische Gruppen	12
4 Analysis	13
4.1 Grenzwert einer Funktion	13
4.2 Stetige Funktionen	13
4.3 Die Richtungsableitung	14
5 Dynamische Systeme	16
5.1 Fixpunkt-Methode	16
6 Tricks	18
6.1 Matrizen vom Typ 2×2	18

Formeln

Additionstheoreme

$$\sin(x + y) = \sin x \cos y + \cos x \sin y,$$

$$\cos(x + y) = \cos x \cos y - \sin x \sin y$$

Polarkoordinaten

$$x = r \cos \varphi,$$

$$y = r \sin \varphi,$$

$$r = \sqrt{x^2 + y^2},$$

$$\varphi = \operatorname{sgn}(y) \arccos\left(\frac{x}{r}\right) \quad (\varphi \neq \pi)$$

Inverse Matrix

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

1 Grundlagen

1.1 Prädikatenlogik

Sei M eine endliche Menge und $(x_k)_{k=1}^n$ eine bijektive Abzählung von M , die jedem k ein Element von M zuordnet. Man definiert nun

$$\forall x \in M [P(x)] := \bigwedge_{k=1}^n P(x_k) := x_1 \wedge x_2 \wedge \dots \wedge x_n \quad (1.1)$$

und

$$\exists x \in M [P(x)] := \bigvee_{k=1}^n P(x_k) := x_1 \vee x_2 \vee \dots \vee x_n. \quad (1.2)$$

Formale Beweise lassen sich induktiv führen. Hierfür ist es notwendig, die rechten Seiten wie beim Summenzeichen rekursiv zu formalisieren:

$$\bigwedge_{k=1}^1 P(x_k) := P(x_1), \quad \bigwedge_{k=1}^n P(x_k) := P(x_n) \wedge \bigwedge_{k=1}^{n-1} P(x_k). \quad (1.3)$$

Nun lässt sich die Regel $C \vee (A \wedge B) = (C \vee A) \wedge (C \vee B)$ auf den Allquantor übertragen. Der Induktionsanfang ist

$$C \vee \bigwedge_{k=1}^1 P(x_k) = C \vee P(x_1) = \bigwedge_{k=1}^1 (C \vee P(x_1)). \quad (1.4)$$

Der Induktionsschritt ist

$$C \vee \bigwedge_{k=1}^n P(x_k) = C \vee \left(P(x_n) \wedge \bigwedge_{k=1}^{n-1} P(x_k) \right) = (C \vee P(x_n)) \wedge \left(C \vee \bigwedge_{k=1}^{n-1} P(x_k) \right) \quad (1.5)$$

$$= (C \vee P(x_n)) \wedge \bigwedge_{k=1}^{n-1} (C \vee P(x_k)) = \bigwedge_{k=1}^n (C \vee P(x_k)). \quad (1.6)$$

1.2 Mengenlehre

Regeln zur Mengenbau-Schreibweise:

$$x \in \{u \mid P(u)\} \iff P(x), \quad (1.7)$$

$$\{f(x) \mid P(x)\} := \{y \mid y = f(x) \wedge P(x)\}, \quad (1.8)$$

$$\{x \in M \mid P(x)\} := \{x \mid x \in M \wedge P(x)\}. \quad (1.9)$$

Definition. Gleichheit und Teilmengenrelation von Mengen:

$$A = B \iff \forall x [x \in A \iff x \in B], \quad (1.10)$$

$$A \subseteq B \iff \forall x [x \in A \implies x \in B]. \quad (1.11)$$

Definition. Grundlegende Operationen:

$$A \cup B := \{x \mid x \in A \vee x \in B\}, \quad (\text{Vereinigung}) \quad (1.12)$$

$$A \cap B := \{x \mid x \in A \wedge x \in B\}, \quad (\text{Schnitt}) \quad (1.13)$$

$$A \setminus B := \{x \mid x \in A \wedge x \notin B\}, \quad (\text{Differenz}) \quad (1.14)$$

$$A \triangle B := \{x \mid x \in A \oplus x \in B\}. \quad (\text{symmetrische Differenz}) \quad (1.15)$$

1.3 Funktionen

Definition. *Bildmenge* von $M \subseteq A$ unter $f: A \rightarrow B$:

$$f(M) := \{f(x) \mid x \in M\}. \quad (1.16)$$

Die maximale Bildmenge $f(A)$ wird einfach *Bildmenge* von f genannt.

Die Regel $f(M \cup N) = f(M) \cup f(N)$ ist allgemeingültig, denn:

$$f(M \cup N) = \{f(x) \mid x \in M \vee x \in N\} = \{y \mid y = f(x) \wedge (x \in M \vee x \in N)\} \quad (1.17)$$

$$= \{y \mid y = f(x) \wedge x \in M \vee y = f(x) \wedge x \in N\} \quad (1.18)$$

$$= \{y \mid y \in \{f(x) \mid x \in M\} \vee y \in \{f(x) \mid x \in N\}\} \quad (1.19)$$

$$= \{y \mid y \in f(M) \vee y \in f(N)\} = \{y \mid y \in f(M)\} \cup \{y \mid y \in f(N)\} \quad (1.20)$$

$$= f(M) \cup f(N). \quad (1.21)$$

Auch $f(M \cap N) = f(M) \cap f(N)$ gilt:

$$f(M \cap N) = \{f(x) \mid x \in M \wedge x \in N\} = \{y \mid y = f(x) \wedge x \in M \wedge x \in N\} \quad (1.22)$$

$$= \{y \mid y = f(x) \wedge y = f(x) \wedge x \in M \wedge x \in N\} \quad (1.23)$$

$$= \{y \mid y \in \{f(x) \mid x \in M\} \wedge y \in \{f(x) \mid x \in N\}\} \quad (1.24)$$

$$= \{y \mid y \in f(M) \wedge y \in f(N)\} = \{y \mid y \in f(M)\} \cap \{y \mid y \in f(N)\} \quad (1.25)$$

$$= f(M) \cap f(N). \quad (1.26)$$

Hierbei wurde $A = A \wedge A$ sowie das Kommutativ- und das Assoziativgesetz zunutze gemacht.

Es gilt sogar die allgemeine Regel

$$f\left(\bigcup_{i \in I} M_i\right) = \bigcup_{i \in I} f(M_i). \quad (1.27)$$

Man rechnet nach:

$$f\left(\bigcup_{i \in I} M_i\right) = \{y \mid y = f(x) \wedge x \in \bigcup_{i \in I} M_i\} = \{y \mid y = f(x) \wedge \exists i \in I [x \in M_i]\} \quad (1.28)$$

$$= \{y \mid \exists i \in I \underbrace{[y = f(x) \wedge x \in M_i]}_{y \in f(M_i)}\} = \bigcup_{i \in I} f(M_i). \quad (1.29)$$

Definition. Eine Funktion $f: A \rightarrow B$ heißt *surjektiv*, wenn $f(A) = B$ ist.

Definition. Eine Funktion $f: A \rightarrow B$ heißt *injektiv*, wenn

$$\forall x_1, x_2 \in A [f(x_1) = f(x_2) \implies x_1 = x_2] \quad (1.30)$$

gilt.

Definition. Für zwei Funktionen $f: A \rightarrow B$ und $g: B \rightarrow C$ wird die Funktion

$$(g \circ f): A \rightarrow C, \quad (g \circ f)(x) := g(f(x)) \quad (1.31)$$

als *Verkettung* oder *Komposition* von g und f bezeichnet. Man spricht *g nach f*.

Zu beachten ist, dass die Zielmenge von f mit der Definitionsmenge von g übereinstimmen muss. Für die Situation $f: A \rightarrow B$ und $g: D \rightarrow C$ mit $B \subseteq D$ kann g natürlich auf B eingeschränkt werden. Die korrekte Notation lautet somit $g|_B \circ f$.

Die Verkettung von zwei Surjektionen ist wieder surjektiv und die Verkettung von zwei Injektionen ist wieder injektiv. Somit ist die Verkettung von zwei Bijektionen wieder bijektiv.

Die Surjektivität können wir für $f: A \rightarrow B$ und $g: B \rightarrow C$ einfach nachrechnen:

$$(g \circ f)(A) = \{g(f(x)) \mid x \in A\} = \{g(y) \mid y = f(x) \wedge x \in A\} \quad (1.32)$$

$$= \{g(y) \mid y \in \{f(x) \mid x \in A\}\} = \{g(y) \mid y \in f(A)\} \quad (1.33)$$

$$= \{g(y) \mid y \in B\} = g(B) = C. \quad (1.34)$$

Die Injektivität lässt sich auch einfach nachrechnen:

$$(g \circ f)(x_1) = (g \circ f)(x_2) \iff g(f(x_1)) = g(f(x_2)) \quad (1.35)$$

$$\implies f(x_1) = f(x_2) \implies x_1 = x_2. \quad (1.36)$$

2 Lineare Algebra

2.1 Vektorräume

Untervektorraum-Kriterium. Ist V ein Vektorraum über dem Körper K und $U \subseteq V$, so ist U auch ein Vektorraum, falls

$$v, w \in U \implies v + w \in U \quad \text{und} \quad v \in U \implies \lambda v \in U \quad (2.1)$$

für $\lambda \in K$ gilt. Die beiden Bedingungen sind die Bausteine von Linearkombinationen. Das Kriterium sagt also aus, dass U abgeschlossen bezüglich jeder Linearkombination sein muss. Das heißt, dass sich das Ergebnis jeder Linearkombination wieder in U befinden muss.

Wenn K ein Körper und M eine nichtleere Menge ist, so ist K^M ein Vektorraum bezüglich punktweiser Addition als Vektoraddition und punktweiser Multiplikation mit einer Konstanten als Skalarmultiplikation. Ein Spezialfall davon ist also $\mathbb{R}^{\mathbb{R}}$. Es ist nun so, dass $C^1(\mathbb{R} \rightarrow \mathbb{R})$ ein Untervektorraum von $C(\mathbb{R} \rightarrow \mathbb{R})$ und dieser ein Untervektorraum von $\mathbb{R}^{\mathbb{R}}$ ist.

Betrachte dazu das Untervektorraum-Kriterium. Wenn zwei Funktionen stetig sind, so ist auch deren Summe und Produkt stetig. Die Skalare können als konstante Funktionen betrachtet werden. Wenn zwei Funktionen stetig differenzierbar sind, so auch deren Summe und Produkt. Wenn eine Funktion stetig differenzierbar ist, so ist sie erst recht stetig.

Bemerkung: Die Menge M darf nicht leer sein, weil K^M mindestens die Nullabbildung enthalten muss, welche als Nullvektor verwendet wird. Wäre M leer, so wäre $|K^M| = |K|^0 = 0$, wobei $|K| \neq 0$ vorausgesetzt werden kann, weil ein Körper nicht leer sein darf. Nun würde K^M keinen einzigen Vektor enthalten, was nicht sein darf, weil jeder Vektorraum mindestens den Nullvektor enthalten muss.

2.2 Lineare Abhängigkeit

Definition. Eine Menge $\{v_k\}_{k=1}^n$ von Vektoren $v_k \in V$ heißt *linear unabhängig*, wenn $\sum_{k=1}^n \lambda_k v_k = 0$ nur gilt, falls alle $\lambda_k = 0$ sind. Andernfalls heißen die Vektoren *linear abhängig*.

Das Definition ist äquivalent dazu, dass sich keiner der Vektoren als Linearkombination der anderen darstellen lässt.

Sei nun zunächst $n = 2$. Sind v_1, v_2 linear abhängig, so gilt

$$\lambda_1 v_1 + \lambda_2 v_2 = 0 \quad (2.2)$$

und nach Umformung gilt $v_2 = \lambda v_1$ mit $\lambda = -\lambda_1/\lambda_2$.

Da die Vektoren dem gleichen Vektorraum entstammen, kann ihr äußeres Produkt gebildet werden. Es gilt nun

$$v_1 \wedge v_2 = v_1 \wedge (\lambda v_1) = \lambda v_1 \wedge v_1 = 0. \quad (2.3)$$

Tatsächlich ist die Menge $\{v_k\}_{k=1}^n$ genau dann linear abhängig, wenn

$$v_1 \wedge v_2 \wedge \dots \wedge v_n = 0 \quad (2.4)$$

gilt. Es genügt, dies am Beispiel $n = 3$ zu verifizieren. Sei also ohne Beschränkung der Allgemeinheit $v_3 = \lambda_1 v_1 + \lambda_2 v_2$. Nun ist:

$$v_1 \wedge v_2 \wedge v_3 = v_1 \wedge v_2 \wedge (\lambda_1 v_1 + \lambda_2 v_2) = \lambda_1 v_1 \wedge v_2 \wedge v_1 + \lambda_2 v_1 \wedge v_2 \wedge v_2 \quad (2.5)$$

$$= -\lambda_1 \underbrace{(v_1 \wedge v_1)}_{=0} \wedge v_2 + \lambda_2 v_1 \wedge \underbrace{(v_2 \wedge v_2)}_{=0} = 0. \quad (2.6)$$

Das Prinzip ist klar: Jeder Vektor der Linearkombination steckt im Faktor $A = v_1 \wedge v_2$. Weil sowohl $A \wedge v_1 = 0$ als auch $A \wedge v_2 = 0$ gilt, verschwindet auch die gesamte Linearkombination.

2.3 Lineare Abbildungen

Definition. Sind V und W Vektorräume über dem selben Körper K , so heißt eine Abbildung $f: V \rightarrow W$ *linear*, wenn sie additiv und homogen ist. *Additiv* bedeutet, dass für alle $v, w \in V$ gilt:

$$f(v + w) = f(v) + f(w) \quad (2.7)$$

und *homogen* bedeutet, dass für alle $\lambda \in K$ und $v \in V$ gilt:

$$f(\lambda v) = \lambda f(v). \quad (2.8)$$

Beispiel. $f(x) := 2x$, $f: \mathbb{R} \rightarrow \mathbb{R}$. Bei \mathbb{R} handelt es sich natürlich um einen Vektorraum, denn \mathbb{R}^n ist ja ein Vektorraum und $\mathbb{R}^1 = \mathbb{R}$.

Beispiel. $D: C^1(\mathbb{R} \rightarrow \mathbb{R}) \rightarrow C(\mathbb{R} \rightarrow \mathbb{R})$ mit

$$D(f) := x \mapsto \frac{df(x)}{dx}. \quad (2.9)$$

Beispiel. $I: C([a, b] \rightarrow \mathbb{R}) \rightarrow C^1([a, b] \rightarrow \mathbb{R})$ mit

$$I(f) := x \mapsto \int_a^x f(x) dx, \quad (2.10)$$

wobei a fest aber beliebig ist.

Die Menge der linearen Abbildungen:

$$\text{Hom}_K(V, W) := \{f \mid f: V \rightarrow W, f \text{ ist linear}\}. \quad (2.11)$$

Bei $\text{Hom}_K(V, W)$ handelt es sich selbst wieder um einen Vektorraum über dem Körper K . Man definiert dafür

$$(f + g)(v) := f(v) + g(v), \quad (\lambda f)(v) := \lambda f(v). \quad (2.12)$$

Die Beweisskizze hierfür:

$$\begin{aligned} (f + g)(v + w) &= f(v + w) + g(v + w) = f(v) + f(w) + g(v) + g(w) \\ &= f(v) + g(v) + f(w) + g(w) = (f + g)(v) + (f + g)(w). \end{aligned}$$

Die restlichen Axiome können als kleine Übung vom Leser überprüft werden.

Bemerkung: Wegen $f(v) \in W$ muss λ bei $\lambda f(v)$ aus dem Körper von W sein. Somit sind alle betrachteten Vektorräume über dem Körper K .

Frage. Ist eine additive Abbildung auch homogen?

Man bemerkt zunächst

$$f(2v) = f(v + v) = f(v) + f(v) = 2f(v). \quad (2.13)$$

Allgemein ergibt sich bei dieser Betrachtung $f(nv) = nf(v)$ für jede natürliche Zahl $n \geq 1$. Dies soll zur Übung noch mal formal verifiziert werden. Man definiert dazu für eine gegebene Folge von Vektoren $(v_k)_{k=1}^n$ das Summenzeichen rekursiv:

$$\sum_{k=1}^1 v_k := v_1, \quad \sum_{k=1}^n v_k := \sum_{k=1}^{n-1} v_k + v_n. \quad (2.14)$$

Der Induktionsanfang ist einfach

$$f\left(\sum_{k=1}^1 v_k\right) = f(v_1). \quad (2.15)$$

Der Induktionsschritt ist

$$\begin{aligned} f\left(\sum_{k=1}^n v_k\right) &= f\left(\sum_{k=1}^{n-1} v_k\right) + f(v_n) \\ &= \sum_{k=1}^{n-1} f(v_k) + f(v_n) = \sum_{k=1}^n f(v_k). \end{aligned} \quad (2.16)$$

Setzt man nun $v_k = v$ für alle k , so ergibt sich

$$f(nv) = f\left(\sum_{k=1}^n v\right) = \sum_{k=1}^n f(v) = nf(v). \quad (2.17)$$

Weiterhin gilt

$$f(0v) = f(0) = f(0 + 0) = f(0) + f(0). \quad (2.18)$$

Aus $f(0) = f(0) + f(0)$ folgt $f(0) = 0 = 0f(v)$.

Beachte nun

$$0 = f(\underbrace{-v + v}_{=0}) = f(-v) + f(v). \quad (2.19)$$

Daraus folgt $f(-v) = -f(v)$. Nach den bisherigen Ausführungen ergibt sich $f(nv) = nf(v)$ für alle $n \in \mathbb{Z}$.

Die Fragestellung lässt sich auch für rationale Zahlen bejahen. Die grundlegende Feststellung dazu ist

$$f(v) = f(\tfrac{1}{2}v + \tfrac{1}{2}v) = f(\tfrac{1}{2}v) + f(\tfrac{1}{2}v) = 2f(\tfrac{1}{2}v). \quad (2.20)$$

Division durch zwei bringt $\tfrac{1}{2}f(v) = f(\tfrac{1}{2}v)$. Allgemein gilt wieder $\tfrac{1}{n}f(v) = f(\tfrac{1}{n}v)$ für $n \in \mathbb{N}$ mit $n \geq 1$. Ist q nun eine rationale Zahl, so gibt es die Darstellung $q = \frac{m}{n}$ mit $m \in \mathbb{Z}$ und $n \in \mathbb{N}, n \geq 1$. Es gilt nun

$$f(qv) = f(m \cdot \tfrac{1}{n} \cdot v) = m \cdot f(\tfrac{1}{n} \cdot v) = m \cdot \tfrac{1}{n} \cdot f(v) = qf(v). \quad (2.21)$$

Was ist nun mit reellen Zahlen? Sei dazu $s_n := \sum_{k=0}^n q_k$ eine Reihe von rationalen Zahlen q_k , welche gegen eine reelle Zahl r konvergiert. Am besten $q_k = \frac{d_k}{10^k}$ mit $q_0 \in \mathbb{Z}$ und $q_{k \neq 0} \in \{0 \dots 9\}$.

Nun gilt

$$\begin{aligned} f(rv) &= f\left(\left(\lim_{n \rightarrow \infty} \sum_{k=0}^n q_k\right)v\right) \stackrel{?}{=} f\left(\lim_{n \rightarrow \infty} \left(\sum_{k=0}^n q_k\right)v\right) = f\left(\lim_{n \rightarrow \infty} \sum_{k=0}^n q_k v\right) \\ &\stackrel{??}{=} \lim_{n \rightarrow \infty} f\left(\sum_{k=0}^n q_k v\right) = \lim_{n \rightarrow \infty} \sum_{k=0}^n q_k f(v) = \lim_{n \rightarrow \infty} \left(\sum_{k=0}^n q_k\right) f(v) \\ &\stackrel{?}{=} \left(\lim_{n \rightarrow \infty} \sum_{k=0}^n q_k\right) f(v) = rf(v). \end{aligned} \quad (2.22)$$

Um die Fragezeichen zu klären, nimmt man nun an, dass V und W mit einer Norm ausgestattet und somit metrische Räume sind. Der folgende elementare Satz ist jetzt aufschlussgebend.

Satz. Eine Abbildung $f: X \rightarrow Y$ zwischen metrischen Räumen (X, d_x) und (Y, d_y) ist genau dann stetig, wenn für jede konvergente Folge (x_n) mit $x_n \in X$ die Eigenschaft

$$f\left(\lim_{n \rightarrow \infty} x_n\right) = \lim_{n \rightarrow \infty} f(x_n) \quad (2.23)$$

gilt.

2.4 Darstellungsmatrizen

Wird in V die Basis B ausgewählt und in W die Basis B' , so ist dadurch der eindeutige Isomorphismus

$$M_{B'}^B: \text{Hom}_K(V, W) \rightarrow K^{\dim W \times \dim V} \quad (2.24)$$

bestimmt. Man nennt $M_{B'}^B(f)$ die *Darstellungsmatrix* von f . Bei $M_{B'}^B$ handelt es sich als Isomorphismus zwischen Vektorräumen um eine bijektive lineare Abbildung.

Das Problem mit dem Isomorphismus $M_{B'}^B$ ist, dass dieser nicht kanonisch ist, weil er von den Basen B und B' abhängt, die willkürlich gewählt werden können. Nun gibt es aber den Koordinatenraum K^n , welcher die Standardbasis als kanonische Basis besitzt.

Die Tatsache, dass die Verkettung zweier linearer Abbildungen selbst wieder linear ist, kann dazu verwendet werden, eine lineare Abbildung in eine Verkettung zu zerlegen. Das veranlasst uns dazu, bei der Bildung der Darstellungsmatrix einen Zwischenschritt einzufügen. Somit ist

$$\text{Hom}_K(V, W) \xrightarrow{F_{B'}^B} \text{Hom}_K(K^{\dim V}, K^{\dim W}) \xrightarrow{\varphi} K^{\dim W \times \dim V}, \quad (2.25)$$

sodass $M_{B'}^B = \varphi \circ F_{B'}^B$ gilt. Bei φ handelt es sich nun um einen kanonischen Isomorphismus, sodass eine lineare Abbildung zwischen Koordinatenräumen mit ihrer Darstellungsmatrix identifiziert werden kann. Man kann also bedenkenlos

$$\text{Hom}_K(K^n, K^m) = K^{m \times n} \quad (2.26)$$

setzen.

Kleine Bauchschmerzen bereitet (2.26) aber doch, denn es können ja trotzdem andere Basen als die Standardbasen gewählt werden. Es genügt hierfür aber, (2.25) zu spezialisieren. Man erhält

$$\text{Hom}_K(K^n, K^m) \xrightarrow{F_{B'}^B} \text{Hom}_K(K^n, K^m) \xrightarrow{\varphi} K^{m \times n}. \quad (2.27)$$

Bei $F_{B'}^B$ handelt es sich jetzt um einen Automorphismus. Bei einem Automorphismus muss es sich aber zwangsweise um einen Basiswechsel handeln. Hierbei werden B und B' zu einer gemeinsamen Basis (B, B') zusammengefasst. Von dieser gemeinsamen Basis wird nun in die gemeinsame Standardbasis (E, E') gewechselt.

Zuweilen findet man oft Angaben wie

$$f \begin{bmatrix} x \\ y \end{bmatrix} := \begin{bmatrix} 2x + 4y \\ 4x \end{bmatrix} \quad (2.28)$$

vor. Solche Angaben sind eigentlich recht sinnfrei, weil f wegen (2.26) auch gleich als Matrix beschrieben werden kann:

$$f \begin{bmatrix} x \\ y \end{bmatrix} := \begin{bmatrix} 2x + 4y \\ 4x \end{bmatrix} = \begin{bmatrix} 2 & 4 \\ 4 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}. \quad (2.29)$$

Die Tupelschreibweise und die Koordinaten (x, y) suggerieren, dass wir es hier mit dem Koordinatenraum und seiner Standardbasis zu tun haben. Man wollte eine lineare Abbildung direkt, aber ohne Festlegung auf Basen beschreiben. Das ist jedoch gar nicht möglich, weil die direkte Beschreibung immer die rechte Seite von (2.26) bedingt.

Die Festlegung auf andere Basen bedeutet nun, dass die Darstellungsmatrix transformiert werden muss. Viele Aufgaben der linearen Algebra mit konkreten linearen Abbildungen können somit vollständig durch Matrizenrechnung gelöst werden.

2.5 Gruppen

Betrachte die Menge der Selbstabbildungen $X^X := \{f \mid f: X \rightarrow X\}$. Die Menge der bijektiven Selbstabbildungen wird *symmetrische Gruppe* $S(X)$ genannt und bildet bezüglich der Verkettung von Abbildungen die Gruppe $(S(X), \circ)$.

(Axiom E) Abgeschlossenheit gilt, weil die Verkettung zweier bijektiver Abbildungen auch wieder bijektiv ist und $S(X)$ alle bijektiven Selbstabbildungen umfassen soll.

(Axiom A) Das Assoziativgesetz $(h \circ g) \circ f = h \circ (g \circ f)$ gilt für *alle* Abbildungen.

(Axiom N) Auf *jeder* Menge X gibt es die identische Abbildung id_X , so dass $f \circ \text{id}_X = f$ und $\text{id}_X \circ f = f$.

(Axiom I) Da die Abbildungen als bijektiv vorausgesetzt sind, gibt es zu jeder Abbildung f ein g , so dass $f \circ g = \text{id}_X$ und $g \circ f = \text{id}_X$.

Betrachte nun $K^{n \times n}$, den Matrizenraum der quadratischen Matrizen mit Einträgen aus dem Körper K . Man definiert nun die Menge der regulären quadratischen Matrizen:

$$\text{GL}(n, K) := \{A \mid A \in K^{n \times n} \wedge \det(A) \neq 0\}. \quad (2.30)$$

Bei $\text{GL}(n, K)$ handelt es sich auch um eine Gruppe. Das neutrale Element ist die Einheitsmatrix E_n , die durch

$$E_2 := \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad E_3 := \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \text{u.s.w.} \quad (2.31)$$

definiert ist.

Nun lässt sich aber ein Vektorraum der Dimension n über dem Körper K betrachten. Nach Wahl einer Basis B gibt es nun

$$M_B^B: \text{Hom}_K(V, V) \rightarrow K^{\dim V \times \dim V}. \quad (2.32)$$

Eine lineare Selbstabbildung wird als *Endomorphismus* bezeichnet. Eine bijektive lineare Selbstabbildung wird *Automorphismus* genannt.

Wenn M_B^B auch ein Gruppenisomorphismus bezüglich der multiplikativen Struktur ist, so ergibt sich

$$M_B^B: \text{Aut}_K(V) \rightarrow \text{GL}(\dim V, K). \quad (2.33)$$

Dann ist $\text{Aut}_K(V)$ aber eine Gruppe, und zwar eine Untergruppe von $S(V)$. Man spricht bei $\text{Aut}_K(V)$ von der *Automorphismengruppe*. Zur Übersicht ergibt sich folgendes Diagramm:

$$\begin{array}{ccc} \text{Aut}_K(V) & \subseteq & \text{Hom}_K(V, V) \\ M_B^B \downarrow & & M_B^B \downarrow \\ \text{GL}(\dim V, K) & \subseteq & K^{\dim V \times \dim V} \end{array} \quad (2.34)$$

2.6 Quadratische Matrizen

Die Multiplikation von Matrizen kann nach dem *falkschen Schema* ausgeführt werden:

		b_{11}	b_{12}
		b_{21}	b_{22}
a_{11}	a_{12}	\sum	\sum
a_{21}	a_{22}	\sum	\sum

Problem. Zu einer gegebenen regulären Matrix A soll die Matrix B gefunden werden, sodass $AB = E_2$ gilt.

Es gibt drei Ansätze zur Lösung.

Der **erste Ansatz** ist die Trennung der Matrixgleichung in vier Gleichungen, wodurch sich zwei LGS ergeben. Zunächst lautet die Matrixgleichung:

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \quad (2.35)$$

Hierdurch ergeben sich die beiden LGS:

$$\left| \begin{array}{cc|c} a_{11}b_{11} + a_{12}b_{21} & = & 1 \\ a_{21}b_{11} + a_{22}b_{21} & = & 0 \end{array} \right|, \quad \left| \begin{array}{cc|c} a_{11}b_{12} + a_{12}b_{22} & = & 0 \\ a_{21}b_{12} + a_{22}b_{22} & = & 1 \end{array} \right|. \quad (2.36)$$

Der **zweite Ansatz** ist die Verwendung des Gauß-Jordan-Verfahrens um $[E_2|B]$ aus $[A|E_2]$ zu gewinnen.

Der **dritte Ansatz** ist die Anwendung der Lösungsformel

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}^{-1} = \frac{1}{a_{11}a_{22} - a_{12}a_{21}} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}. \quad (2.37)$$

Bei Matrizen vom Typ 3×3 oder 4×4 ist nur noch das Gauß-Jordan-Verfahren sinnvoll. Ab 5×5 oder bei $\mathbb{C}^{n \times n}$ anstelle von $\mathbb{R}^{n \times n}$ wird man Matlab/Octave verwenden.

3 Algebra

3.1 Gruppen

Man betrachte die Menge $M = \{A, B, C\}$, wobei $\triangle ABC$ ein gleichseitiges Dreieck sein soll. Außerdem definiert man

$$S(M) := \{f \mid f: M \rightarrow M, f \text{ ist bijektiv}\}. \quad (3.1)$$

In Abschnitt 2.5 wurde schon festgestellt, dass $(S(M), \circ)$ eine Gruppe ist. Außerdem muss $S(M)$ eine Anzahl von $3! = 6$ Elementen besitzen, wie aus der elementaren Kombinatorik bekannt ist. Nun definieren wir

$$d_1 := (A \mapsto B \mapsto C \mapsto A), \quad d_2 := (A \mapsto C \mapsto B \mapsto A). \quad (3.2)$$

Wir finden nun folgende Gesetzmäßigkeiten heraus:

$$d_1^2 = d_1 \circ d_1 = \text{id}, \quad d_2^2 = d_2 \circ d_2 = \text{id}, \quad d_1^{-1} = d_2, \quad d_2^{-1} = d_1. \quad (3.3)$$

Weiteres Herumspielen führt uns aber nicht zu den drei anderen Elementen in $S(M)$, egal wie sehr wir uns bemühen. Trotzdem haben wir zu jedem Element ein inverses, wobei man beachtet, dass id zu sich selbst invers ist. Auch id ist enthalten, und die restlichen Gruppenaxiome gelten auch. Somit handelt es sich bei

$$H := \{\text{id}, d_1, d_2\} \quad (3.4)$$

um eine Untergruppe von $S(X)$.

Definition. Ist $(G, *)$ eine Gruppe und $H \subseteq G$, so heißt H *Untergruppe* von G , wenn $(H, *)$ auch wieder eine Gruppe bildet. Man schreibt dann $H \leq G$.

Drei Elemente haben wir aus dem Teich $S(M)$ gefischt, drei stecken noch im Schlamm. Aber der Teich ist nicht besonders tief, mehr eine Pfütze als ein Teich. Was aus dem Wasser gezogen wird, möchten wir aber sogleich anschauen.

Korollar aus den Gruppenaxiomen. *Jede Gruppe besitzt genau ein neutrales Element. Zu jedem Element gibt es genau ein inverses.*

Nähere Überlegung führt zu der Erkenntnis, dass es ausgeschlossen ist, dass zwei unterschiedliche Elemente das selbe inverse besitzen. Wenn b nämlich zu a_1 und zu a_2 invers ist, so sind a_1 und a_2 auch invers zu b . Weil das inverse Element zu b aber eindeutig bestimmt sein muss, gilt zwingend $a_1 = a_2$.

Jedes Element besitzt also seinen inversen Partner, wobei ein Element aber auch zu sich selbst invers sein darf.

Die restlichen Elemente von $S(M)$ müssen sich zu solchen Partnerschaften gruppieren lassen. Dann müssen zwei Elemente ein Paar bilden und eines ist zu sich selbst invers? Falsch. Alle drei sind zu sich selbst invers. Das sind

$$s_1 := (A)(BC) = (A \mapsto A)(B \mapsto C \mapsto B), \quad (3.5)$$

$$s_2 := (B)(AC) = (B \mapsto B)(A \mapsto C \mapsto A), \quad (3.6)$$

$$s_3 := (C)(AB) = (C \mapsto C)(A \mapsto B \mapsto A). \quad (3.7)$$

Alle Operation lassen sich bei $\triangle ABC$ deuten. Es handelt sich um Kongruenzabbildungen, welche das Dreieck auf sich selbst abbilden, so dass man keinen Unterschied bemerkt, außer dass zwei oder alle drei Ecken vertauscht sind. Das Element d_1 ist eine Drehung um 120° und d_2 ist eine Drehung um $240^\circ = -120^\circ$. Das Element id ist eine Drehung um $0^\circ = 360^\circ$. Bei s_1, s_2, s_3 handelt es sich um Spiegelungen mit den Winkelhalbierenden als Spiegelachsen.

Man könnte nun auf die Idee kommen, dass auch bei einem Quadrat die Symmetrien genau die Elemente der symmetrischen Gruppe sind. Die Gruppe hat aber $4! = 24$ Elemente, eine solche Anzahl von Symmetrien scheint fragwürdig. Ein Gegenbeispiel ist sofort ersichtlich:

$$f := (A)(B)(CD) = (A \mapsto A)(B \mapsto B)(C \mapsto D \mapsto C). \quad (3.8)$$

Da müsste das Quadrat auseinandergeschnitten, eines der Stücke umgedreht, und dann beide Stücke wieder zusammengeklebt werden.

Tatsächlich bilden die Symmetrien des Quadrates aber eine Gruppe, die aus diesem Grund als *Symmetriegruppe* bezeichnet wird. Die Übereinstimmung der Symmetriegruppe des Dreiecks mit der symmetrischen Gruppe der Ecken stellt sich als Koinzidenz heraus.

3.2 Zyklische Gruppen

Die Gruppe (3.4) ist von besonderer Gestalt. Sie lässt sich von nur einem einzigen Element erzeugen.

Definition. Eine Gruppe heißt *zyklisch*, wenn sie die Form

$$\langle g \rangle := \{g^n \mid n \in \mathbb{Z}\} \quad (3.9)$$

besitzt. Man nennt g den *Erzeuger* und die Gruppe $\langle g \rangle$ das *Erzeugnis* von g .

Es ergibt sich nun:

n	-5	-4	-3	-2	-1	0	1	2	3	4	5
d_1^n	d_1	d_2	id	d_1	d_2	id	d_1	d_2	id	d_1	d_2

(3.10)

Die Elemente wiederholen sich immer. Es gibt einen Zyklus der Länge drei. Es ist

$$\langle d_1 \rangle = \{\text{id}, d_1, d_2\}, \quad \langle d_2 \rangle = \{\text{id}, d_1, d_2\}. \quad (3.11)$$

Die Regel $\langle g^{-1} \rangle = \langle g \rangle$ gilt natürlich immer. Für viele Gruppenelemente ist aber $\langle g^2 \rangle \neq \langle g \rangle$. Als Beispiel lässt sich die Spiegelung s_1 , die an der Stelle (3.5) definiert wurde, heranziehen. Es gilt $s_1^2 = \text{id}$ und daher:

$$\langle s_1^2 \rangle = \langle \text{id} \rangle = \{\text{id}\} \neq \langle s_1 \rangle = \{\text{id}, s_1\}. \quad (3.12)$$

4 Analysis

4.1 Grenzwert einer Funktion

Definition. Sei $D \subseteq \mathbb{R}$ und a ein Berührungspunkt von D . Sei $f: D \rightarrow \mathbb{R}$. Es sei $\lim_{x \rightarrow a} f(x) = L$ genau dann, wenn für jede Folge (x_n) , $x_n \in D$ mit $x_n \rightarrow a$ die Bilderfolge $f(x_n)$ gegen L geht. Wie bei Folgen wird L als *Grenzwert* bezeichnet.

Diese Definition wirkt zunächst außerordentlich unhandlich, da zunächst nicht klar ist, wie man etwas für alle beliebigen Folgen überprüfen soll. Dies ist für einfache Beispiele aber nicht sonderlich schwer, da man einfach eine Variable verwenden kann, welche für alle möglichen Folgen steht.

Ein einführendes Beispiel. Betrachtet wird die Funktion $f(x) := x^2$ mit $f: \mathbb{R} \rightarrow \mathbb{R}$. Wir würden gerne $L = \lim_{x \rightarrow 2} f(x)$ berechnen. Nun sei (x_n) eine beliebige Folge mit $x_n \rightarrow 2$. Es gilt nun

$$\lim_{n \rightarrow \infty} f(x_n) = \lim_{n \rightarrow \infty} x_n^2 = \lim_{n \rightarrow \infty} (x_n \cdot x_n) = \left(\lim_{n \rightarrow \infty} x_n \right) \cdot \left(\lim_{n \rightarrow \infty} x_n \right) = 2 \cdot 2 = 4. \quad (4.1)$$

Da die Folge (x_n) konvergent ist, konnten wir den Grenzwertsatz für Summen anwenden. Somit muss $L = 4$ sein.

Solche Rechnungen müssen für einfache Beispiele nicht weiter durchgeführt werden, denn die Zusammenhänge sind hier so direkt, dass sich die Grenzwertsätze für Folgen gleich auf Grenzwerte für Funktionen übertragen lassen.

Angenommen es gilt $\lim_{x \rightarrow a} f(x) = A$ und $\lim_{x \rightarrow a} g(x) = B$. Sei $x_n \rightarrow a$. Dann konvergieren $f(x_n) \rightarrow A$ und $g(x_n) \rightarrow B$. Somit ist

$$\lim_{n \rightarrow \infty} (f(x_n) \cdot g(x_n)) = \left(\lim_{n \rightarrow \infty} f(x_n) \right) \cdot \left(\lim_{n \rightarrow \infty} g(x_n) \right) = AB. \quad (4.2)$$

Es ergibt sich

$$\lim_{x \rightarrow a} (f(x) \cdot g(x)) = AB. \quad (4.3)$$

Die anderen Grenzwertsätze lassen sich auf die gleiche Art verifizieren.

Jetzt können wir einfach

$$\lim_{x \rightarrow a} x^2 = \lim_{x \rightarrow a} (x \cdot x) = \left(\lim_{x \rightarrow a} x \right) \cdot \left(\lim_{x \rightarrow a} x \right) = a^2 \quad (4.4)$$

rechnen, ohne den Umweg über eine Folge gehen zu müssen.

Diese Zusammenhänge ermöglichen es, Grenzwerte für eine große Klasse von Funktionen sofort in die Hand geschenkt zu bekommen.

4.2 Stetige Funktionen

Definition. Eine Funktion $f: D \rightarrow \mathbb{R}$ heißt *stetig* an der Stelle $a \in D$, wenn $\lim_{x \rightarrow a} f(x) = f(a)$ gilt. Man nennt f stetig in D , wenn f an jeder Stelle von D stetig ist.

Beispiel. Die Funktion $f(x) := x^2$ mit $f: \mathbb{R} \rightarrow \mathbb{R}$ ist stetig. Denn es gilt

$$\lim_{x \rightarrow a} f(x) = \lim_{x \rightarrow a} x^2 \stackrel{(4.4)}{=} \left(\lim_{x \rightarrow a} x \right)^2 = a^2 = f(a) \quad (4.5)$$

für jedes $a \in \mathbb{R}$.

Ist $f: D \rightarrow \mathbb{R}$ eine stetige Funktion, so folgt sofort

$$\lim_{n \rightarrow \infty} f(x_n) = f\left(\lim_{n \rightarrow \infty} x_n\right) \quad (4.6)$$

für jede konvergente Folge (x_n) , $x_n \in D$ mit $x_n \rightarrow a \in D$.

4.3 Die Richtungsableitung

Man stelle sich einen affinen Raum A vor, auf welchem eine Funktion $f_A: A \rightarrow \mathbb{R}$ definiert ist. Um eine solche Funktion konkret angeben zu können, benötigen wir eine *Koordinatendarstellung* von f_A . Dazu wählt man eine affine Basis (p_0, B) , die ein Koordinatensystem

$$\varphi: \mathbb{R}^n \rightarrow A, \quad \varphi(x) := p_0 + \sum_{k=1}^n x_k b_k \quad (4.7)$$

mit $x = (x_k)_{k=1}^n$ und $B = (b_k)_{k=1}^n$ induziert. Man beachte, dass φ eine invertierbare affine Abbildung ist. Ein Koordinatensystem ist also eine invertierbare affine Abbildung zwischen dem Koordinatenraum und einem affinen Raum.

Wählt man speziell $A = \mathbb{R}^n$, so kann B als Matrix aus Spaltenvektoren $b_k = (b_{ik})$ dargestellt werden. Dann ist

$$y = \varphi(x) = p_0 + Bx. \quad (4.8)$$

Die Matrix B muss nun regulär sein, damit die Umformung

$$x = B^{-1}(y - p_0) = \varphi^{-1}(y) \quad (4.9)$$

formuliert werden kann. Betrachtet man einen affinen Raum A als abstraktes Objekt, so ist dieser nicht über ein Koordinatensystem φ erfassbar, weil A selbst nicht erfassbar ist. Denkt man sich aber zwei Koordinatensysteme φ, ψ , so ist der Koordinatensystemwechsel $T = \varphi^{-1} \circ \psi$ sehr wohl erfassbar. Den Raum A können wir nun völlig außer acht lassen. Alle Darstellungen von Funktionen können sich somit auf den \mathbb{R}^n beziehen und über

$$T: \mathbb{R}^n \rightarrow \mathbb{R}^n, \quad T(x) := p + Ax \quad (4.10)$$

mit einer Matrix $A \in \text{GL}(\mathbb{R}, n)$ transformiert werden. Man beachte hierzu, dass die Inverse einer affinen Abbildung und die Verkettung von zwei affinen Abbildungen wieder eine affine Abbildung ist. Somit muss $T = \varphi^{-1} \circ \psi$ auch wieder eine affine Abbildung sein.

Anstelle von f_A betrachtet man stattdessen die konkrete Funktion

$$f: \mathbb{R}^n \rightarrow \mathbb{R}, \quad f := f_A \circ \varphi. \quad (4.11)$$

Bei einer Koordinatentransformation ergibt sich nun

$$f \circ T = f_A \circ \varphi \circ \varphi^{-1} \circ \psi = f_A \circ \psi, \quad (4.12)$$

was als gleichberechtigte Darstellung von f_A angesehen werden darf.

Bei f handelt es sich um eine Funktion in mehreren Variablen. Wir würden gerne auch für f so etwas wie eine Ableitung definieren, um auch für solche Funktionen eine Differentialrechnung entwickeln zu können.

Betrachte dazu eine Gerade $g \in \mathbb{R}^n$ und einen Punkt $p \in \mathbb{R}^n$. Für die Anschauung sei $n = 2$. Man definiert nun die Projektion

$$\pi: \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}^n, \quad \pi(x, y) := x, \quad (4.13)$$

die einen Punkt (x, y) senkrecht auf die Stelle x herunterprojiziert.

Zu beachten ist, dass $x = (x_k)_{k=1}^n$ ein Tupel, y aber nur eine reelle Zahl ist.

Man betrachte nun die Ebene $\pi^{-1}(g)$. Die Schnittmenge

$$G = \text{Graph}(f) \cap \pi^{-1}(g) \quad (4.14)$$

ist eine Kurve, die als Graph einer Funktion $s: \mathbb{R} \rightarrow \mathbb{R}$ dargestellt werden kann. Aber s lässt sich mit dem herkömmlichen Differentialquotient ableiten. Wählt man auf g nun eine affine Basis (p, v) , so wird dadurch das Koordinatensystem

$$\varphi: \mathbb{R} \rightarrow \mathbb{R}^n, \quad \varphi(t) := p + tv \quad (4.15)$$

induziert. Ein solches Koordinatensystem für einen eindimensionalen Raum wird auch als *Lineal* bezeichnet.

Nun gibt es eine Darstellung $s = f \circ \varphi$.

Definition. *Richtungsableitung* von f in Richtung v an der Stelle p :

$$(D_v f)(p) := (f \circ \varphi)'(0). \quad (4.16)$$

Aus der Definition können wir sofort eine Analogiebetrachtung bei den Differentialquotienten extrahieren:

$$f'(x) = \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h}, \quad (D_v f)(p) = \lim_{h \rightarrow 0} \frac{f(p+hv) - f(p)}{h}. \quad (4.17)$$

Der Koordinatenraum \mathbb{R}^n besitzt die kanonische Basis $(e_k)_{k=1}^n$. Die Basisvektoren lassen sich natürlich als Richtungsvektoren verwenden, und diesen Richtungsableitungen kommt eine besondere Bedeutung zu.

Definition. Die Richtungsableitungen bezüglich der kanonischen Basis werden als *partielle Ableitungen* bezeichnet:

$$(D_k f)(p) := (D_{e_k} f)(p). \quad (4.18)$$

Bemerkung: Es gibt die alternativen Schreibweisen

$$(D_v f)(p) \equiv \frac{\partial f(x)}{\partial v} \Big|_{x=p}, \quad (D_k f)(p) \equiv \frac{\partial f(x)}{\partial x_k} \Big|_{x=p}. \quad (4.19)$$

Außerdem gibt es eine zweite Vorstellung von Differenzierbarkeit. Diese Vorstellung ist, dass an der Stelle p eine Tangentialebene an den Graphen von f gelegt werden kann. Eine solche Tangentialebene muss die Funktionsgleichung

$$T(x_1, x_2) = f(p) + a_1(x_1 - p_1) + a_2(x_2 - p_2) \quad (4.20)$$

besitzen, wobei a_1, a_2 zwei zu bestimmende Parameter sind. Die Tangentialebene soll die Funktion f an der Stelle p nun möglichst gut approximieren. Für ein kleines h soll also

$$f(p+h) - T(p+h) \approx 0 \quad (4.21)$$

sein. Schärfer soll sogar

$$\lim_{h \rightarrow 0} \frac{f(p+h) - T(p+h)}{\|h\|} = 0 \quad (4.22)$$

gelten.

Definition. Eine Funktion $f: \mathbb{R}^n \rightarrow \mathbb{R}$ heißt *total differenzierbar* an der Stelle p , wenn

$$\lim_{h \rightarrow 0} \frac{f(p+h) - f(p) - \langle a, x-p \rangle}{\|h\|} = 0 \quad (4.23)$$

ist. Der Koordinatenvektor a ist hierdurch eindeutig bestimmt und wird Ableitung von f genannt. Man nennt

$$(df)(p) = \sum_{k=1}^n a_k dx^k(p) \quad (4.24)$$

das *totale Differenzial* von f . Hiernach ergibt sich

$$\langle a, x-p \rangle = (df)(p)(x-p) \quad (4.25)$$

über duale Paarung, wobei die linke Seite bezüglich Koordinatentupeln gemeint ist und die rechte bezüglich tatsächlichen Vektoren.

Das wirkt zunächst wie eine zweite unnötige Schreibweise. Betrachtet man aber f_A anstelle von f , so muss es nicht unbedingt ein Skalarprodukt $\langle a, x-p \rangle$ geben, weil $x-p$ als Differenz zweier Punkte $x, p \in A$ im Verschiebungsvektorraum V liegt. Und auf V braucht kein Skalarprodukt definiert zu sein. Stattdessen entstammt $(df)(p)$ aus dem Dualraum V^* und ist bezüglich der dualen Basis $(dx^k)_{k=1}^n$ definiert, so dass $(df)(p)$ und $x-p$ dual gepaart werden können.

Noch eine Bemerkung zum Definitionsbereich von f : Es reicht aus, wenn f auf einer offenen Umgebung des betrachteten Punktes p definiert ist. Die Umgebung sollte aber offen sein, man sich bei Grenzwertbildungen aus jeder Richtung dem Punkt p nähern kann. Sonst würde man einen Ableitungsbegriff benötigen, der bestimmte Richtungen ausschließt, was unnötigen technischen Ballast darstellt.

5 Dynamische Systeme

5.1 Fixpunkt-Methode

Betrachte eine Iteration

$$x_{n+1} = \varphi(x_n) \quad (5.1)$$

mit einem Startwert x_0 . Z. B.

$$\varphi(x) = \frac{1}{a+x}, \quad a \in \mathbb{N} \quad (5.2)$$

mit beliebigem Startwert. Man stellt numerisch fest, dass die Folge (x_n) bei diesem Beispiel (es handelt sich um Kettenbrüche) offenbar immer gegen eine Zahl konvergiert. Außerdem scheint der Startwert dafür unbedeutend zu sein. Bei $a = 2$ erkennt ein Rechenmeister mit geschultem Blick noch, dass die Folge gegen $\sqrt{2} - 1$ konvergiert, bei $a = 1$ gegen den goldenen Schnitt. Aber bei $a = 3$ ist der Wert zunächst von einem dunklen Schleier umgeben.

Eine weitere Beobachtung ist nun, dass sich die Werte der Folge immer weniger voneinander unterscheiden, wenn n wächst. Es gilt also $x_n \approx x_{n+1}$ bzw. $x_n \approx \varphi(x_n)$ für große n . Für den Grenzwert $x = \lim_{n \rightarrow \infty} x_n$ müsste dann

$$x = \varphi(x) \quad (5.3)$$

gelten. Eine solche Gleichung wird als *Fixpunktgleichung* bezeichnet, die dazu gehörige Iteration (5.1) als *Fixpunkt-Iteration*.

Setzt man nun das Beispiel mit $a = 2$ ein, so ergibt sich

$$x = \frac{1}{2+x}, \quad (5.4)$$

was zur quadratischen Gleichung

$$x^2 + 2x - 1 = 0 \quad (5.5)$$

führt, deren eine Lösung $\sqrt{2} - 1$ ist. Warum es eine zweite Lösung gibt, und diese nicht Ergebnis der Iteration ist, ist zunächst unklar.

Die Beispiele

$$\varphi(x) = \frac{1}{1-bx}, \quad b \in \{1, 2\} \quad (5.6)$$

scheinen niemals zu konvergieren. Für $b = 1$ ergibt sich eine periodische Folge, und bei $b = 2$ scheint die Folge recht chaotisch zu sein. Die zugehörigen Fixpunkt-Gleichungen führen auf quadratische Gleichungen, welche keine reellen Lösungen besitzen.

Wenn also der Fixpunkt eine komplexe Zahl ist, so kann die Folge niemals gegen den Fixpunkt konvergieren, weil der Wert von φ niemals eine komplexe Zahl sein wird, wenn als Argument eine reelle Zahl gegeben wurde. Würden wir mit einem komplexen Startwert beginnen, so hätten wir eine Chance auf Konvergenz. Ein Plotter für die komplexe Ebene zeigt aber, dass für fast alle Zahlen bis auf einzelne offenbar keine Konvergenz vorliegt. Das Beispiel $b = 1$ ist für fast alle komplexen Startwerte periodisch mit Periodenlänge drei.

Beim Beispiel $a = 2$ scheinen die Folgenwerte von der zweiten Lösung der quadratischen Gleichung abgestoßen zu werden und zur ersten Lösung zu wandern, was klar wird wenn man die Iteration ganz in der Nähe der zweiten Lösung beginnt. Man hat es hier mit einem abstoßenden und einem anziehenden Fixpunkt zu tun. Das Abstoßen geschieht um so schneller, je weiter der der Iterationswert vom abstoßenden Fixpunkt entfernt ist.

Bei den Beispielen $b = 1$ und $b = 2$ ergibt sich jedoch, dass sich die Folgenwerte um die Fixpunkte herum bewegen, ohne sie in absehbarer Zeit zu erreichen. Bei $b = 2$ liegen die Werte von (x_n) auf einem Kreis um einen der Fixpunkte, dessen Mittelpunkt aber nicht der Fixpunkt sein muss. Für $\text{Im}(x_0) = 0$ entartet der Radius des Kreises ins unendliche. Die beiden Fixpunkte sind zueinander komplex konjugiert, und so kann sich die Folge bei $\text{Im}(x_0) = 0$ nicht für die Umkreisung von einem der beiden Fixpunkte entscheiden.

Bei $b = 2 + i/2$ ergibt sich noch ein viel seltsameres Gebilde. Die Folgenwerte werden wieder von einem der beiden Fixpunkte angezogen und von dem anderen abgestoßen. Dabei bilden sich jedoch Strudel mit fünf Armen aus, wobei die Arme des abstoßenden Fixpunktes mit denen des anziehenden Fixpunktes verbunden sind.

Doch da ist noch viel mehr. Nehmen wir das Beispiel $\varphi(x) = x^2 - 1$. Damit Konvergenz vorliegt, dürfen die Folgenwerte notwendigerweise einen bestimmten Betrag nicht überschreiten, denn nach der umgekehrten Dreiecksungleichung gilt:

$$|\varphi(x)| = |x^2 - 1| \geq ||x^2| - |1|| = ||x|^2 - 1|. \quad (5.7)$$

Für $|x| \geq 2$ ergibt sich $|x|^2 \geq 4$ und $|x|^2 - 1 \geq 3$. Somit ist

$$||x|^2 - 1| = |x|^2 - 1 \quad (5.8)$$

und $|\varphi(x)| \geq 3$. Man beachtet jetzt

$$|\varphi(x)| \geq |x|^2 - 1 \implies \sqrt{|\varphi(x)| + 1} \geq |x|. \quad (5.9)$$

Nun ist aber $|\varphi(x)| > \sqrt{|\varphi(x)| + 1}$ für $|\varphi(x)| \geq (\sqrt{5} + 1)/2$, was bei $|x| \geq 2$ erfüllt ist. Insgesamt ergibt sich $|\varphi(x)| > |x|$, was folglich die Divergenz der Folge zur Folge hat.

Die Menge der Startwerte $x_0 \in \mathbb{C}$, so dass (x_n) mit $x_{n+1} = x_n^2 + c$ beschränkt bleibt, heißt *Julia-Menge* zur Konstanten c . Die Menge zu $c = -1$, die innerhalb des Kreises $|x| = 2$ liegt, ist ein außerordentlich seltsames Gebilde, welches dem Augenschein nach fraktale Strukturen besitzt.

Für einen Startwert x_0 wird die Folge (x_n) nach (5.1) auch als *Orbit* von x_0 bezeichnet. Im Fall, dass negative n nicht zugelassen sind, spricht man auch von einem *Semiorbit*. Dieser Begriff ist mit dem Gruppentheoretischen Begriff *Orbit* deckungsgleich, denn die Iterierten φ^n bilden bezüglich Verkettung eine Gruppe und erwirken eine Gruppenaktion $\varphi^n(x_0)$. Sind negative n nicht zugelassen, so handelt es sich nicht um eine Gruppe, sondern um ein Monoid, dessen Elemente eine Monoidaktion anstelle einer Gruppenaktion erwirken.

Definition. Die Struktur (T, X, Φ) mit

$$\Phi: T \times X \rightarrow X, \quad \Phi(t, x) := \varphi^t(x) \quad (5.10)$$

wird als *dynamisches System* bezeichnet. Das System heißt *diskret*, wenn $t \in \mathbb{N}$ bzw. $t \in \mathbb{Z}$ ist. Die Fixpunkte von φ werden auch *Ruhelagen* von Φ genannt.

6 Tricks

6.1 Matrizen vom Typ 2×2

Matrizen der Form

$$\begin{bmatrix} a & -b \\ b & a \end{bmatrix} = r \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \quad (6.1)$$

sind isomorph zu den komplexen Zahlen $a + bi = re^{i\theta}$. Das erscheint zunächst wie eine unwesentliche Tatsache, aber es ermöglicht, Matrizenrechnungen gegen Rechnungen mit komplexen Zahlen auszutauschen. Der Isomorphismus ist

$$\varphi: M \rightarrow \mathbb{C}, \quad \varphi\left(\begin{bmatrix} a & -b \\ b & a \end{bmatrix}\right) := a + bi. \quad (6.2)$$

Zunächst verifiziert man, dass M und \mathbb{C} Körper sind, indem die Gültigkeit der Körperaxiome nachgewiesen wird. Hierbei ergibt sich, dass $M \setminus \{0\}$ eine abelsche Gruppe ist. Es gilt insbesondere $M \setminus \{0\} \leq \text{GL}(2, \mathbb{R})$.

Definition. Sind $(K, +, \bullet)$ und $(K', +', \bullet')$ zwei Körper, so wird $\varphi: K \rightarrow K'$ als *Körperhomomorphismus* bezeichnet, wenn

$$\varphi(a + b) = \varphi(a) +' \varphi(b), \quad (6.3)$$

$$\varphi(a \bullet b) = \varphi(a) \bullet' \varphi(b) \quad (6.4)$$

für alle $a, b \in K$ gilt und $\varphi(1) = 1$ ist.

Ein bijektiver Körperhomomorphismus ist ein *Körperisomorphismus*.
Nun kommutiert das folgende Diagramm:

$$\begin{array}{ccc} M & \xrightarrow{\text{Berechnung}} & M \\ \varphi \downarrow & & \uparrow \varphi^{-1} \\ \mathbb{C} & \xrightarrow{\text{Berechnung}} & \mathbb{C} \end{array} \quad (6.5)$$

Zunächst ergeben sich ein paar Zusammenhänge, die ein wenig verblüffend erscheinen:

$$\varphi(A^T) = \overline{\varphi(A)}, \quad |\varphi(A)| = \sqrt{\det A}, \quad \text{Spur}(A) = 2 \operatorname{Re}(\varphi(A)). \quad (6.6)$$

Wichtiger sind Regeln wie

$$A^{-1} = \varphi^{-1}\left(\frac{1}{\varphi(A)}\right), \quad A^n = \varphi^{-1}(\varphi(A)^n). \quad (6.7)$$

So lassen sich Matrizen einfach invertieren und große Matrixpotenzen berechnen. Aber das bringt uns auf eine seltsame Idee. Wir definieren nun

$$f(A) := \varphi^{-1}(f(\varphi(A))) = (\varphi^{-1} \circ f \circ \varphi)(A) \quad (6.8)$$

für jede Funktion $f: G \rightarrow \mathbb{C}$ mit $\varphi(A) \in G$. Diese Definition erweitert (6.7).

Für eine Matrix $A \in M$ lassen sich jetzt eine Reihe scheinbar absurder Dinge berechnen:

$$\sqrt{A}, \quad A^{3/2}, \quad A^i, \quad A^A, \quad \frac{1}{A+1}, \quad e^A, \quad \ln(A), \quad \sin(A). \quad (6.9)$$

Einige würden jetzt den Nutzen solcher Operationen, ja sogar die Operationen selbst in Frage stellen. Aber das hat man bei den komplexen Zahlen zunächst auch getan. Tatsächlich ist \sqrt{A} eine Teilantwort auf das Problem, dass zu einer Matrix A die Matrix X mit $X^2 = A$ gefunden werden soll.

Die nächste Frage besteht nun darin, ob sich Operationen wie $f(x) = e^x$ auch auf Matrizen übertragen lassen, die nicht in der Form (6.1) vorliegen.