

Aufzeichnungen zur Mathematik

John Smith, 2016
Lizenz: Creative Commons CC0

Inhaltsverzeichnis

1 Grundlagen	1
1.1 Zerlegung des Allquantors	1
1.2 Allquantifizierung über Produktmengen	2
1.3 Zahlen als Mengen	2
2 Analysis	3
2.1 Halley-Verfahren	3
2.2 Integral der Potenzfunktion	4
3 Lineare Algebra	4
3.1 Lineare Abbildungen	4
3.2 Hodge-Stern-Operator	5
3.2.1 Orientierung	5
3.3 Skalarprodukträume	5

1 Grundlagen

1.1 Zerlegung des Allquantors

Eine periodische Funktion f erfüllt die Gleichung

$$f(x) = f(x + T) \quad (1.1)$$

definiert. Macht man eine Substitution $x := u - T$, so ergibt sich

$$f(u - T) = f((u - T) + T) = f(u). \quad (1.2)$$

Somit gilt auch $f(x) = f(x - T)$. Diese Folgerung war sehr kurz, eigentlich zu kurz um wissenschaftlich zu sein.

Eine Funktion $f: \mathbb{R} \rightarrow \mathbb{R}$ heißt periodisch mit Periode T , falls die Funktionalgleichung

$$\forall x \in \mathbb{R}: f(x) = f(x + T) \quad (1.3)$$

erfüllt ist. Für die Gleichung ergibt sich zunächst der folgende abstrakte Syntaxbaum (AST):

$$\begin{aligned} (= & \\ & (\text{f } x) \\ & (\text{f } (+ \ x \ T))) \end{aligned}$$

Aber was ist mit dem Allquantor? Der Allquantor ist eigentlich eine Funktion in zwei Variablen. Das erste Argument ist eine Menge M , das zweite ein Prädikat p . Der Allquantor überprüft nun, ob das Prädikat p für alle Elemente von M gültig ist. Wir wandeln nun die Schreibweise durch

$$\forall x \in M: p(x) \longrightarrow \text{all}(M, p) \quad (1.4)$$

um. Die Funktionalgleichung wird als nun in der Form

$$\text{all}(\mathbb{R}, \lambda x. f(x) = f(x + T)) \quad (1.5)$$

dargestellt. Dazu gehört folgender AST:

$$\begin{aligned} & (\text{all } \mathbb{R} \\ & \quad (\text{lambda } (x) \\ & \quad \quad (= \\ & \quad \quad \quad (\text{f } x) \\ & \quad \quad \quad (\text{f } (+ \ x \ T)))))) \end{aligned}$$

Unter einer *freien Substitution* versteht man nun den Austausch jedes Vorkommens einer *freien Variable* durch einen AST. Durch den λ -Term wird die Variable x nun gebunden, ist also nicht mehr frei. Daher können wir nicht einfach eine freie Substitution $[x:=u-T]$ durchführen.

Viel mehr gilt die folgende Regel. Ist g eine passende Bijektion, so gilt

$$\text{all}(M, p) = \text{all}(g^{-1}(M), p \circ g). \quad (1.6)$$

Z.B. wählt man $M := \{1, 2, 3\}$ und $p(x) := (x < 4)$ sowie $g(x) := x/2$. Somit ergibt sich

$$\begin{aligned} & (\forall x \in \{1, 2, 3\}: x < 4) \\ & = (\forall x \in \{2, 4, 6\}: x/2 < 4). \end{aligned}$$

Sei nun $g(x) := x - T$. Nun ist $g^{-1}(\mathbb{R}) = \mathbb{R}$. Wendet man das nun auf (1.5) an, so ergibt sich

$$\text{all}(\mathbb{R}, \lambda x. f(x - T) = f(x)). \quad (1.7)$$

Es handelt sich also eigentlich nicht um eine Substitution, sondern um die Verkettung des Prädikats mit einer Funktion. Sicherlich lässt sich eine solche durch eine Substitution darstellen, aber das ist sehr schwammig. Und diese Schwammigkeit fällt erst auf, wenn der Allquantor auch ausgeschrieben wird. Im Allquantor ist ja, wie wir gesehen haben, eine Variablenbindung enthalten.

Doch warum darf man (1.6) eigentlich anwenden? Nun, hierzu zerlegen wir den Allquantor weiter. Es ist nämlich

$$\text{all}(M, p) = (p(M) = \{\text{true}\}). \quad (1.8)$$

Sei id die identische Funktion. Nun ist aber p das selbe wie $p \circ \text{id}$. Außerdem gilt ja $\text{id} = g \circ g^{-1}$. Somit ergibt sich

$$\begin{aligned} p(M) &= (p \circ g \circ g^{-1})(M) \\ &= (p \circ g)(g^{-1}(M)). \end{aligned} \quad (1.9)$$

In der letzten Gleichung wurde $(f \circ g)(M) = f(g(M))$ verwendet. Das soll noch schnell gezeigt werden. Es gilt

$$\begin{aligned} & (f \circ g)(M) \\ &= \{(f \circ g)(x) \mid x \in M\} \\ &= \{f(y) \mid y = g(x) \wedge x \in M\} \\ &= \{f(y) \mid y \in g(M)\} = f(g(M)). \end{aligned} \quad (1.10)$$

Nun gut, das ist auch ein Pseudobeweis. Wir können hier noch

$$\bigcup_{i \in I} f(A_i) = f\left(\bigcup_{i \in I} A_i\right) \quad (1.11)$$

verwenden. Dann gilt aber

$$\begin{aligned} (f \circ g)(M) &= \bigcup_{x \in M} \{f(g(x))\} = \bigcup_{x \in M} f(\{g(x)\}) \\ &= f\left(\bigcup_{x \in M} \{g(x)\}\right) = f(g(M)). \end{aligned}$$

Für den Existenzquantor gilt analog

$$\exists x \in A: p(x) \quad \text{gdw.} \quad \text{true} \in p(A). \quad (1.12)$$

Für die Beschreibung von Mengen gilt außerdem

$$\begin{aligned} \{x \in A \mid p(x)\} &= \text{filter}(A, p) \\ &= p^{-1}(\{\text{true}\}). \end{aligned} \quad (1.13)$$

Die Beschreibung von Mengen lässt sich also als Urbildmenge darstellen.

Für den Anzahlquantor gilt außerdem

$$\exists^n x \in A: p(x) \quad \text{gdw.} \quad \#\text{filter}(A, p) = n. \quad (1.14)$$

1.2 Allquantifizierung über Produktmengen

Sei $M := A \times B$. Sei $M_x := \pi_1^{-1}(\{x\})$, das ist die Faser von x für die Projektion $\pi_1(x, y) := x$. Es gilt nun

$$M = \bigcup_{x \in A} M_x. \quad (1.15)$$

Es gilt außerdem (was zu zeigen ist)

$$\forall t \in \bigcup_{i \in I} A_i [P(t)] \iff \forall i \in I \forall t \in A_i [P(t)]. \quad (1.16)$$

Daher ist

$$\begin{aligned} \forall (x, y) \in M [P(x, y)] \\ \iff \forall x \in A \forall (x, y) \in M_x [P(x, y)]. \end{aligned} \quad (1.17)$$

Wegen

$$\begin{aligned} M_x &= \{x\} \times B \\ &= \{(x, y) \mid x \in \{x\} \wedge y \in B\} \\ &= \{(x, y) \mid y \in B\} \end{aligned} \quad (1.18)$$

gilt nun

$$(x, y) \in M_x \iff y \in B. \quad (1.19)$$

Somit gilt

$$\begin{aligned} \forall (x, y) \in M [P(x, y)] \\ \iff \forall x \in A \forall y \in B [P(x, y)]. \end{aligned} \quad (1.20)$$

Verwendet man π_2 anstelle von π_1 , so gelangt man zu der Erkenntnis $\forall x \forall y \iff \forall y \forall x$.

Betrachten wir nun (1.16) für endliche Zerlegungen. Sei dazu $I(n) := \{1, \dots, n\}$ und $U_n := U_{n-1} \cup A_n$ mit Anfang $U_1 := A_1$. Nun gilt $U_n = \bigcup_{i \in I(n)} A_i$.

Es ist nun

$$\begin{aligned} \forall t \in U_n [P(t)] \\ \iff \forall t \in U_{n-1} \cup A_n [P(t)] \\ \iff \forall t \in U_{n-1} [P(t)] \wedge \forall t \in A_n [P(t)] \end{aligned} \quad (1.21)$$

Falls (1.16) für $n-1$ gilt (Induktionsvoraussetzung), so ist

$$\begin{aligned} \forall t \in U_{n-1} [P(t)] \\ \iff \forall i \in I(n-1) \forall t \in A_i [P(t)]. \end{aligned} \quad (1.22)$$

Zusammenführen, d. h.

$$\begin{aligned} \forall t \in A_n [P(t)] \wedge \forall i \in I(n-1) \forall t \in A_i [P(t)] \\ \iff \forall i \in I(n) \forall t \in A_i [P(t)], \end{aligned} \quad (1.23)$$

bringt (1.16) für n . Die Induktion beginnen wir zunächst bei $n=1$, um möglichen Problemen mit der leeren Menge aus dem Weg zu gehen.

Im allgemeinen Fall verwenden wir zunächst die prädikatenlogische Definition der Vereinigungsmenge:

$$\bigcup_{i \in I} A_i := \{x \mid \exists i \in I [x \in A_i]\}. \quad (1.24)$$

Nun gilt

$$\begin{aligned} \forall t \in \bigcup_{i \in I} A_i [P(t)] \\ \iff \forall t [t \in \bigcup_{i \in I} A_i \implies P(t)] \\ \iff \forall t [\exists i [t \in A_i] \implies P(t)] \\ \iff \forall t [\forall i [t \in A_i \implies P(t)]] \\ \stackrel{?!}{\iff} \forall i \forall t [t \in A_i \implies P(t)] \\ \iff \forall i \forall t \in A_i [P(t)]. \end{aligned} \quad (1.25)$$

Was noch zu zeigen ist, ist die Regel

$$\forall x \forall y [P(x, y)] \iff \forall y \forall x [P(x, y)]. \quad (1.26)$$

1.3 Zahlen als Mengen

Die Potenzmenge einer Menge M wird auch als 2^M notiert, weil $|2^M| = 2^{|M|}$ gilt falls M endlich ist. Nach dem Von-Neumann-Modell ist aber $2 := \{0, 1\}$. Daher entspricht 2^M der Menge

$$\{f \mid f: M \rightarrow \{0, 1\}\}. \quad (1.27)$$

Jeder Teilmenge entspricht eine Funktion f , welche für jedes $x \in M$ auswählt, ob x vorkommen soll oder nicht.

Die Potenzmenge kann daher auch als Menge aller charakteristischen Funktionen von M gesehen werden.

Die bisherige Betrachtung ermöglicht nun in Erfahrung zu bringen, was 3^M , 4^M , usw. bedeutet. Bei n^M handelt es sich um die Menge der Funktionen

$$\{f \mid f: M \rightarrow \{0, 1, \dots, n-1\}\}. \quad (1.28)$$

Eine Funktion f zählt also für jedes $x \in M$ wie oft es vorkommen soll, wobei $n-1$ die Maximale Anzahl ist. Somit kann n^M als Menge von Multimengen interpretiert werden.

Man definiert normalerweise $A^1 := A$ und $A^{n+1} := A \times A^n$. Jetzt stellt sich aber die Frage, was denn A^0 ist. Weiterhin gültig sein sollte

$$|A^0| = |A|^0 = 1. \quad (1.29)$$

Daher müsste sich eine einelementige Menge ergeben. Die einfachste einelementige Menge ist $\{0\}$. Andererseits ist A^n eine Menge von n -Tupeln. Also müsste A^0 das leere Tupel enthalten. Somit müsste $0 = \{\} = ()$ sein. Das kartesische Produkt ist eigentlich nicht assoziativ. Aus diesem Grund definiert man explizit

$$(x, y, z) := (x, (y, z)) \quad (1.30)$$

und so weiter. Beginnt man mit der leeren Menge, so ergeben sich die alternativen Darstellungen:

$$() := \{\}, \quad (1.31)$$

$$(a) := (a, \{\}), \quad (1.32)$$

$$(a, b) := (a, (b, \{\})), \quad (1.33)$$

$$(a, b, c) := (a, (b, (c, \{\}))), \quad (1.34)$$

und so weiter. Dies entspricht der Darstellung von Listen in Lisp, also einfach verketteten Listen.

Bei $f: A^0 \rightarrow B$ handelt es sich somit um eine Funktion, welche dem leeren Tupel einen Wert zuordnet. Eine solche nullstellige Funktion entspricht einer Konstante. In der Informatik muss die »Funktion« natürlich zusätzlich referenziell transparent sein.

2 Analysis

2.1 Halley-Verfahren

Im Folgenden wird die Herleitung des Halley-Verfahrens unter Verwendung einer Padé-Approximation dargestellt. Die Padé-Approximation $R[m, 0]$ ist die Taylorreihe. Sei

$$a_k = \frac{1}{k!} f^{(k)}(x_0). \quad (2.1)$$

Die Taylorreihe ist

$$R[m, 0] = \sum_{k=0}^m a_k (x - x_0)^k. \quad (2.2)$$

Sei

$$A = \begin{vmatrix} a_{m-n+1} & a_{m-n+2} & \dots & a_{m+1} \\ \dots & \dots & \dots & \dots \\ a_m & a_{m+1} & \dots & a_{m+n} \\ \sum_{i=n}^m a_{i-n} x^i & \sum_{i=n-1}^m a_{i-n+1} x^i & \dots & \sum_{i=0}^m a_i x^i \end{vmatrix}$$

und

$$B = \begin{vmatrix} a_{m-n+1} & a_{m-n+2} & \dots & a_{m+1} \\ \dots & \dots & \dots & \dots \\ a_m & a_{m+1} & \dots & a_{m+n} \\ x^n & x^{n-1} & \dots & x^0 \end{vmatrix}.$$

Es ist $R[m, n] = A/B$. Ersetze dann x gegen $x - x_0$. Es ist

$$\begin{aligned} R[1, 1] &= \frac{a_1(a_0 + a_1x) - a_0a_2x}{a_1 - a_2x} \\ &= \frac{a_0a_1 + (a_1^2 - a_0a_2)x}{a_1 - a_2x}. \end{aligned} \quad (2.3)$$

Die Nullstelle von f ist gesucht, und die Approximation muss dort auch ungefähr null sein. Setzt man also $R[1, 1](x) = 0$, so folgt

$$0 = a_0a_1 + (a_1^2 - a_0a_2)x. \quad (2.4)$$

Ersetzt man noch x gegen $x - x_0$ und formt danach um, so erhält man

$$x = x_0 - \frac{a_0a_1}{a_1^2 - a_0a_2}. \quad (2.5)$$

Ausgeschrieben bekommt man also $x = \varphi(x_0)$ mit

$$\varphi(x) = x - \frac{2f(x)f'(x)}{2f'(x)^2 - f(x)f''(x)}. \quad (2.6)$$

Das Halley-Verfahren ist dann die Fixpunktiteration $x_{n+1} = \varphi(x_n)$. Man sieht, dass das Halley-Verfahren in das Newton-Verfahren übergeht, wenn man $f(x)f''(x)$ verschwinden lässt.

2.2 Integral der Potenzfunktion

Ich will eine ungewöhnliche Methode zur Berechnung des Integrals

$$\int x^n dx \quad (2.7)$$

vorführen. Wähle die Substitution $x = e^u$. Es gilt nun

$$\frac{dx}{du} = \frac{d}{dx} e^u = e^u \quad (2.8)$$

und somit $dx = e^u du$. Man hätte auch

$$x^n = e^{\ln(x)n} = e^{un} \quad (2.9)$$

rechnen können, was vielleicht eher

$$\frac{du}{dx} = \frac{d}{dx} \ln(x) = \frac{1}{x} \quad (2.10)$$

suggeriert hätte. Hier muss man bedenken, dass die Substitution nochmals ausgeführt werden kann, dass also

$$\frac{1}{x} = \frac{1}{e^u} \quad (2.11)$$

gilt. Nach Substitutionsregel ergibt sich nun

$$\begin{aligned} \int x^n dx &= \int e^{nu} e^u du = \int e^{(n+1)u} du \\ &= \frac{1}{n+1} e^{(n+1)u} = \frac{x^{n+1}}{n+1}. \end{aligned} \quad (2.12)$$

Es geht noch weiter. Im Fall $n = -1$ gilt

$$\int \frac{1}{x} dx = \int \frac{1}{e^u} e^u du = \int du = u. \quad (2.13)$$

Aber Umformen der Substitution ergibt ja $u = \ln x$.

3 Lineare Algebra

3.1 Lineare Abbildungen

Seien V, W zwei Vektorräume über dem Körper K . Eine Abbildung $f: V \rightarrow W$ heißt *additiv*, wenn für alle $v, w \in V$ gilt:

$$f(v + w) = f(v) + f(w) \quad (3.1)$$

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

$$f(\lambda v) = \lambda f(v). \quad (3.2)$$

Ist eine additive Abbildung auch homogen? Man bemerkt zunächst

$$f(2v) = f(v + v) = f(v) + f(v) = 2f(v). \quad (3.3)$$

Allgemein ergibt sich bei dieser Betrachtung $f(nv) = nf(v)$ für jede natürliche Zahl $n \geq 1$. Weiterhin gilt

$$f(0v) = f(0) = f(0 + 0) = f(0) + f(0). \quad (3.4)$$

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

Beachte nun

$$0 = f(\underbrace{-v + v}_{=0}) = f(-v) + f(v). \quad (3.5)$$

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

$$f(v) = f(\tfrac{1}{2}v + \tfrac{1}{2}v) = f(\tfrac{1}{2}v) + f(\tfrac{1}{2}v) = 2f(\tfrac{1}{2}v). \quad (3.6)$$

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

$$\begin{aligned} f(qv) &= f(m \cdot \tfrac{1}{n} \cdot v) = m \cdot f(\tfrac{1}{n} \cdot v) \\ &= m \cdot \tfrac{1}{n} \cdot f(v) = qf(v). \end{aligned} \quad (3.7)$$

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

$$\begin{aligned} f(rv) &= f\left(\left(\lim_{n \rightarrow \infty} \sum_{k=0}^n q_k\right)v\right) \stackrel{?}{=} f\left(\lim_{n \rightarrow \infty} \left(\sum_{k=0}^n q_k\right)v\right) \\ &= f\left(\lim_{n \rightarrow \infty} \sum_{k=0}^n q_k v\right) \stackrel{??}{=} \lim_{n \rightarrow \infty} f\left(\sum_{k=0}^n q_k v\right) = \lim_{n \rightarrow \infty} \sum_{k=0}^n q_k f(v) \\ &= \lim_{n \rightarrow \infty} \left(\sum_{k=0}^n q_k\right) f(v) \stackrel{?}{=} \left(\lim_{n \rightarrow \infty} \sum_{k=0}^n q_k\right) f(v) = rf(v). \end{aligned} \quad (3.8)$$

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

$$f\left(\lim_{n \rightarrow \infty} x_n\right) = \lim_{n \rightarrow \infty} f(x_n) \quad (3.9)$$

gilt.

Wenn kein metrischer Raum vorliegt, lassen sich die Begriffe Cauchy-Folge und Vollständigkeit nicht mehr formulieren und es liegt eventuell nicht einmal mehr ein Hausdorff-Raum vor, was eine sehr komische Situation darstellt. Sind die Vektorräume nicht topologisch, so lässt sich nicht einmal mehr der Begriff des Grenzwertes formulieren.

3.2 Hodge-Stern-Operator

3.2.1 Orientierung

Sei V ein \mathbb{R} -Vektorraum und sei $B = (b_k)_{k=1}^n$ eine Basis von V , die wir positiv orientiert nennen. Ist B' eine andere Basis, so gibt es eine Basiswechselmatrix $T_{B'}^B$. Sei nun

$$s := \text{sgn}(\det(T_{B'}^B)). \quad (3.10)$$

Die Basis B' heißt nun *positiv orientiert*, wenn $s > 0$ ist und *negativ orientiert*, wenn $s < 0$ ist.

Im Koordinatenraum soll natürlich die Standardbasis positiv orientiert sein.

Beachte, dass für das äußere Produkt gilt:

$$\bigwedge_{k=1}^n b'_k = \det(T_{B'}^B) \bigwedge_{k=1}^n b_k. \quad (3.11)$$

Angenommen, die Basis B' ist nun eine Permutation von $B = (b_k)$, also die Festlegung

$$B' := (b_{\sigma(k)})_{k=1}^n, \quad \sigma \in \text{Sym}(n). \quad (3.12)$$

Die Basiswechselmatrix ist nun die Permutationsmatrix und somit gilt:

$$\det(T_{B'}^B) = \text{sgn}(\sigma). \quad (3.13)$$

3.3 Skalarprodukträume

Ist V ein Skalarproduktraum und ist $B = (e_k)_{k=1}^n$ eine Orthonormalbasis von V , so ist der Hodge-Stern-Operator eine lineare Abbildung, definiert durch

$$\begin{aligned} & * (e_{\sigma(1)} \wedge \dots \wedge e_{\sigma(k)}) \\ & := \text{sgn}(\sigma) s(B) e_{\sigma(k+1)} \wedge \dots \wedge e_{\sigma(n)} \end{aligned} \quad (3.14)$$

mit $*$: $\Lambda^k(V) \rightarrow \Lambda^{n-k}(V)$, wobei $0 \leq k \leq n$ gilt.

Dabei ist $s(B) = +1$, wenn B positiv orientiert ist, und $s(B) = -1$, wenn B negativ orientiert ist.

Sei $B = (b_k)$ nun eine Orthogonalbasis. Man definiert nun $g_{ij} := \langle b_i, b_j \rangle$. Jetzt lassen sich den Basisvektoren durch

$$e_k := \frac{b_k}{\|b_k\|} = \frac{b_k}{\sqrt{g_{kk}}} \quad (3.15)$$

normierte Basisvektoren zuordnen. Damit ergibt sich

$$\begin{aligned} & * (b_{\sigma(1)} \wedge \dots \wedge b_{\sigma(k)}) = \\ & = r \text{sgn}(\sigma) s(B) b_{\sigma(k+1)} \wedge \dots \wedge b_{\sigma(n)} \end{aligned} \quad (3.16)$$

mit

$$\begin{aligned} r &= \frac{\sqrt{g_{\sigma(1)\sigma(1)} \dots g_{\sigma(k)\sigma(k)}}}{\sqrt{g_{\sigma(k+1)\sigma(k+1)} \dots g_{\sigma(n)\sigma(n)}}} \\ &= \frac{g_{\sigma(1)\sigma(1)} \dots g_{\sigma(k)\sigma(k)}}{\sqrt{\det g}}. \end{aligned} \quad (3.17)$$