

Aufzeichnungen zur Mathematik

August 2019

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:

```
(all IR
  (lambda (x)
    (=
      (f x)
      (f (+ x T))))))
```

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 dasselbe 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

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

Inhaltsverzeichnis

1 Grundlagen

- 1.1 Zerlegung des Allquantors 1
- 1.2 Allquantifizierung über Produktmengen . . . 2
- 1.3 Zahlen als Mengen 2

2 Analysis

- 2.1 Halley-Verfahren 3
- 2.2 Integral der Potenzfunktion 3
- 2.3 Zum Basler Problem 4
- 2.4 Integraldarstellung der Zeta-Funktion 4
- 2.5 Metamorphose 5
- 2.6 Grünwald-Letnikov-Ableitung 6
- 2.7 Eine Reihe von Kehrwerten 6

3 Lineare Algebra

- 3.1 Lineare Abbildungen 6
- 3.2 Hodge-Stern-Operator 7
- 3.3 Skalarprodukträume 7

4 Wahrscheinlichkeitsrechnung

- 4.1 Bivariate Verteilungen 8
- 4.2 Zufallszahlengeneratoren 8
- 4.3 Gesetz der totalen Wahrscheinlichkeit 8

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):

```
(=
  (f x)
  (f (+ x T)))
```

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

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

$$\begin{aligned} R[1, 1] &= \frac{a_1(a_0 + a_1 x) - a_0 a_2 x}{a_1 - a_2 x} \\ &= \frac{a_0 a_1 + (a_1^2 - a_0 a_2) x}{a_1 - a_2 x}. \end{aligned} \quad (2.3)$$

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

$$0 = a_0 a_1 + (a_1^2 - a_0 a_2) x. \quad (2.4)$$

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

$$x = x_0 - \frac{a_0 a_1}{a_1^2 - a_0 a_2}. \quad (2.5)$$

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 $u = \ln x$.

2.3 Zum Basler Problem

Betrachte die Summenformel

$$\sum_{k=m}^{n-1} q^k = \frac{q^n - q^m}{q - 1}. \quad (q \neq 1) \quad (2.14)$$

Mit Differentialrechnung lässt sich daraus die Formel

$$\sum_{k=m}^{n-1} k^p q^k = \left(q \frac{d}{dq} \right)^p \frac{q^n - q^m}{q - 1} \quad (2.15)$$

gewinnen, das ist eine bekannte Technik. Man beachte dabei

$$\left(q \frac{d}{dq} \right)^p \neq q^p \left(\frac{d}{dq} \right)^p. \quad (p \geq 2) \quad (2.16)$$

Die Vorgehensweise wird nun umgedreht, so dass integriert wird, anstelle zu differenzieren:

$$\int_0^q \frac{1}{q} \sum_{k=m}^{n-1} q^k dq = \int_0^q \left(\frac{1}{q} \right) \frac{q^n - q^m}{q - 1} dq. \quad (2.17)$$

Es ergibt sich

$$\sum_{k=m}^{n-1} \frac{q^k}{k^2} = \int_0^q \frac{1}{q} \int_0^q \frac{q^n - q^m}{q^2 - q} dq dq. \quad (2.18)$$

Man setzt nun $m = 1$ und lässt $n \rightarrow \infty$ gehen. Das bringt

$$\sum_{k=1}^{\infty} \frac{q^k}{k^2} = \int_0^q \frac{1}{q} \int_0^q \frac{1}{1 - q} dq dq. \quad (2.19)$$

Um die Polstelle lässt sich in der komplexen Ebene herum integrieren. Dabei ergibt sich

$$\sum_{k=1}^{\infty} \frac{q^k}{k^2} = \int_q^0 \frac{\ln(1 - q)}{q} dq, \quad (2.20)$$

und somit

$$\sum_{k=1}^{\infty} \frac{1}{k^2} = \lim_{q \rightarrow 1} \int_q^0 \frac{\ln(1 - z)}{z} dz. \quad (2.21)$$

Die Definitionslücke bei $z = 0$ lässt sich stetig beheben. Dann muss aber auch das Integral existieren. Mittels Substitution $1 - z := e^x$ ergibt sich noch

$$\sum_{k=1}^{\infty} \frac{1}{k^2} = \lim_{q \rightarrow 1} \int_{\ln(1-q)}^0 \frac{x e^x}{e^x - 1} dx. \quad (2.22)$$

Demnach gilt

$$\sum_{k=1}^{\infty} \frac{1}{k^2} = \int_{-\infty}^0 \frac{x e^x}{e^x - 1} dx. \quad (2.23)$$

Substitution $x := -x$ ergibt nun

$$\sum_{k=1}^{\infty} \frac{1}{k^2} = \int_0^{\infty} \frac{x e^{-x}}{1 - e^{-x}} dx. \quad (2.24)$$

Erweitert man den Bruch nun mit e^x , was wegen $e^x \neq 0$ gefahrlos getan werden kann, dann ergibt sich

$$\sum_{k=1}^{\infty} \frac{1}{k^2} = \int_0^{\infty} \frac{x}{e^x - 1} dx. \quad (2.25)$$

Schaut man sich den Graph des Integranden an, dann bemerkt man, dass sich dieses Integral besonders gut numerisch integrieren lässt.

Die Bernoulli-Zahlen \bar{B}_k mit $\bar{B}_1 = -1/2$ sind implizit definiert durch

$$\frac{x}{e^x - 1} = \sum_{k=0}^{\infty} \bar{B}_k \frac{x^k}{k!}. \quad (2.26)$$

Demnach gilt:

$$\sum_{k=1}^{\infty} \frac{1}{k^2} = \lim_{x \rightarrow \infty} \sum_{k=0}^{\infty} \bar{B}_k \frac{x^{k+1}}{(k+1)!}. \quad (2.27)$$

2.4 Integraldarstellung der Zeta-Funktion

2.4.1 Vorbetrachtung

Wir wollen zunächst eine Formel für das wiederholte Integral

$$(I_a^n f)(x) := \left(\int_a^x \right)^n f(x) (dx)^n \quad (2.28)$$

$$= \int_a^x \int_a^x \dots \int_a^x f(x) dx dx \dots dx \quad (2.29)$$

herleiten.

Aus der Formel für Parameterintegrale erhält man zunächst den Spezialfall

$$\int_a^x \frac{\partial \varphi(x, t)}{\partial x} dt + \varphi(x, x) = \frac{d}{dx} \int_a^x \varphi(x, t) dt. \quad (2.30)$$

Einsetzen von $\varphi(x, t) := (x - t)^n f(t)$ bringt für $n > 0$ nun

$$n \int_a^x (x - t)^{n-1} f(t) dt = \frac{d}{dx} \int_a^x (x - t)^n f(t) dt. \quad (2.31)$$

Wiederholt man das noch n mal, dann findet man

$$n! f(x) = \left(\frac{d}{dx} \right)^{n+1} \int_a^x (x - t)^n f(t) dt, \quad (2.32)$$

wobei $n! = \Gamma(n + 1)$ gilt. Man definiert nun noch

$$(I_a f)(x) := \int_a^x f(t) dt. \quad (2.33)$$

Da sich Integral und Ableitung gegenseitig aufheben, ergibt sich

$$\Gamma(n + 1) (I_a^{n+1} f)(x) = \int_a^x (x - t)^n f(t) dt. \quad (2.34)$$

Umformen bringt die folgende Formel:

Riemann-Liouville-Integral.

$$(I_a^n f)(x) = \frac{1}{\Gamma(n)} \int_a^x (x-t)^{n-1} f(t) dt. \quad (2.35)$$

Für n kann nun sogar eine komplexe Zahl eingesetzt werden, wobei man hofft, dass das Integral konvergiert.

2.4.2 Integraldarstellung

Satz und Definition. Zeta-Funktion.

Für $\operatorname{Re}(s) > 1$ konvergiert die Reihe:

$$\zeta(s) := \sum_{k=1}^{\infty} \frac{1}{k^s}. \quad (2.36)$$

Ohne Beweis. Betrachte nun die Summenformel

$$\sum_{k=m}^{n-1} q^k = \frac{q^n - q^m}{q - 1}. \quad (q \neq 1) \quad (2.37)$$

Mit Differentialrechnung lässt sich daraus die Formel

$$\sum_{k=m}^{n-1} k^p q^k = \left(q \frac{d}{dq} \right)^p \frac{q^n - q^m}{q - 1} \quad (2.38)$$

gewinnen, das ist eine bekannte Technik. Man beachte dabei

$$\left(q \frac{d}{dq} \right)^p \neq q^p \left(\frac{d}{dq} \right)^p. \quad (p \geq 2) \quad (2.39)$$

Die Vorgehensweise wird nun umgedreht, so dass integriert wird, anstelle zu differenzieren. Es ergibt sich

$$\sum_{k=m}^{n-1} \frac{q^k}{k^p} = \left(\int_0^q \frac{1}{q} \right)^p \frac{q^n - q^m}{q - 1} (dq)^p. \quad (2.40)$$

Damit ist folgendes gemeint:

$$\sum_{k=m}^{n-1} \frac{q^k}{k^p} = \int_0^q \frac{1}{q} \int_0^q \frac{1}{q} \dots \int_0^q \frac{1}{q} \left(\frac{q^n - q^m}{q - 1} \right) dq dq \dots dq. \quad (2.41)$$

Wir setzen nun $m = 1$. Verlangen wir $q < 1$, dann können wir $n \rightarrow \infty$ auf beiden Seiten der Gleichung anwenden und erhalten:

$$\sum_{k=1}^{\infty} \frac{q^k}{k^p} = \left(\int_0^q \frac{1}{q} \right)^p \frac{q}{1 - q} (dq)^p = \left(\int_0^q dq \frac{1}{q} \right)^p \frac{q}{1 - q}. \quad (2.42)$$

Beachtet man nun $\frac{d \ln q}{dq} = \frac{1}{q}$ und stellt dies zu $d \ln q = \frac{dq}{q}$ um, dann wird klar, dass wir hier mit der Substitution $x = \ln q$ weiterarbeiten können. Es ergibt sich:

$$\sum_{k=1}^{\infty} \frac{q^k}{k^p} = \left(\int_{e^x=0}^{e^x=q} dx \right)^p \frac{e^x}{1 - e^x}. \quad (2.43)$$

Nun werden aber auch alle q in den Grenzen bis auf das letzte von innen nach außen gegen e^x ersetzt. Dann ergibt sich immer $e^x = e^x$, also $x = x$. Somit gilt:

$$\sum_{k=1}^{\infty} \frac{q^k}{k^p} = [x := \ln q] \left(\int_{-\infty}^x dx \right)^p \frac{e^x}{1 - e^x}. \quad (2.44)$$

Wendet man nun Formel (2.35) mit $n = p$ an, dann ergibt sich

$$\sum_{k=1}^{\infty} \frac{q^k}{k^p} = \frac{1}{\Gamma(p)} \int_{-\infty}^{\ln q} (\ln q - t)^{p-1} \frac{e^t}{1 - e^t} dt. \quad (2.45)$$

Wir substituieren nun $x = -t$ und erhalten

$$\sum_{k=1}^{\infty} \frac{q^k}{k^p} = \frac{1}{\Gamma(p)} \int_{\infty}^{-\ln q} (\ln q + x)^{p-1} \frac{e^{-x}}{e^{-x} - 1} dx. \quad (2.46)$$

Den Integrand können wir wegen $e^x \neq 0$ gefahrlos mit e^x erweitern und erhalten

$$\frac{e^{-x}}{e^{-x} - 1} = \frac{1}{1 - e^x}. \quad (2.47)$$

Wir können nun auf beiden Seiten $q \rightarrow 1$ anwenden, wobei das Integral dann natürlich für beide Grenzen als uneigentliches zu verstehen ist. Wegen $\ln q \rightarrow 0$ ergibt sich

$$\sum_{k=1}^{\infty} \frac{1}{k^p} = \frac{1}{\Gamma(p)} \int_{\infty}^0 x^{p-1} \frac{1}{1 - e^x} dx = \frac{1}{\Gamma(p)} \int_0^{\infty} \frac{x^{p-1}}{e^x - 1} dx. \quad (2.48)$$

Nimmt man jetzt an, dass Formel (2.35) auch für gewisse komplexe Zahlen $n = s$ gültig ist, dann ergibt sich für die Zeta-Funktion die Integraldarstellung:

Integraldarstellung der Zeta-Funktion.

$$\zeta(s) = \frac{1}{\Gamma(s)} \int_0^{\infty} \frac{x^{s-1}}{e^x - 1} dx. \quad (2.49)$$

Tatsächlich gilt auch diese Formel für $\operatorname{Re}(s) > 1$, was an dieser Stelle nicht weiter gezeigt werden soll. Stattdessen erhalten wir daraus die Einsicht, dass das Riemann-Liouville-Integral offenbar konsistente Resultate liefern kann. Die genauen Voraussetzungen dafür stehen momentan im Dunklen und müssten herausgearbeitet werden.

2.5 Metamorphose

Seien $f(x)$ und $g(x)$ zwei reelle Funktionen und sei r aus dem Intervall $[0, 1]$. Eine Metamorphose von f zu g lässt sich durch die Funktion

$$h(x) = (1 - r)f(x) + rg(x) \quad (2.50)$$

modellieren. Dabei handelt es sich um das punktweise gewichtete arithmetische Mittel der beiden Funktionen. Man bezeichnet dies auch als Konvexkombination.

Allgemeiner muss es sich bei f und g natürlich nicht um reelle Funktionen handeln. Für zwei Parameterkurven oder

zwei Vektorfelder lässt sich so auch eine Metamorphose bilden.

Sind $F(x, y) = 0$ und $G(x, y) = 0$ zwei implizite Funktionen, so lässt sich die Metamorphose durch

$$H(x, y) = (1 - r)F(x, y) + rG(x, y) = 0. \quad (2.51)$$

modellieren. Setzt man $F(x, y) = y - f(x)$ und $G(x, y) = y - g(x)$, so erhält man das zuvor beschriebene Modell als Spezialfall.

Sind allgemeiner F und G Funktionen mit mehreren Variablen, so kann die Metamorphose der impliziten Funktionen analog durch

$$(1 - r)F(x_1, \dots, x_n) + rG(x_1, \dots, x_n) = 0. \quad (2.52)$$

beschrieben werden.

2.6 Grünwald-Letnikow-Ableitung

Den Translationsoperator T_h definiert man durch

$$(T_h f)(x) := f(x + h). \quad (2.53)$$

Für den Differenzoperator gilt nun:

$$(\Delta_h f)(x) := f(x + h) - f(x) = (T_h - 1)f(x). \quad (2.54)$$

Setzt man nun die binomische Reihe ein, dann ergibt sich

$$\Delta_h^n = (T_h - 1)^n = (-1)^n \sum_{k=0}^n \binom{n}{k} (-1)^k T_h^k. \quad (2.55)$$

Wegen $T_h^k = T_{hk}$ ergibt sich nun

$$(\Delta_h^n f)(x) = (-1)^n \sum_{k=0}^n \binom{n}{k} (-1)^k f(x + kh). \quad (2.56)$$

Für eine natürliche Zahl n geht die Reihe nur bis $n = k$. Wir wollen jetzt aber auch komplexe Zahlen für n einsetzen.

Nun ergibt sich heuristisch

$$(D^n f)(x) = \lim_{h \rightarrow 0} \left(\frac{\Delta_h}{h} \right)^n f(x) = \lim_{h \rightarrow 0} \frac{1}{h^n} (\Delta_h^n f)(x). \quad (2.57)$$

Das führt uns zu folgender Definition:

Umgekehrte Grünwald-Letnikow-Ableitung.

$$(D^n f)(x) := \lim_{h \rightarrow 0} \frac{(-1)^n}{h^n} \sum_{k=0}^{\infty} \binom{n}{k} (-1)^k f(x + kh). \quad (2.58)$$

Substituiert man nun $h := -h$ dann ergibt sich:

Direkte Grünwald-Letnikow-Ableitung.

$$(D^n f)(x) := \lim_{h \rightarrow 0} \frac{1}{h^n} \sum_{k=0}^{\infty} \binom{n}{k} (-1)^k f(x - kh). \quad (2.59)$$

Man beachte nun aber auch

$$(a + b)^n = a^n \left(1 + \frac{b}{a} \right)^n \quad (2.60)$$

$$= a^n \sum_{k=0}^{\infty} \binom{n}{k} \left(\frac{b}{a} \right)^k = \sum_{k=0}^{\infty} \binom{n}{k} a^{n-k} b^k, \quad (2.61)$$

falls die binomische Reihe konvergiert. Demnach ergibt sich

$$\Delta_h^n = (T_h - 1)^n = \sum_{k=0}^{\infty} \binom{n}{k} (-1)^k T_h^{n-k}. \quad (2.62)$$

Das führt nun zu:

Definition. Grünwald-Letnikow-Ableitung.

$$(D^n f)(x) := \lim_{h \rightarrow 0} \frac{1}{h^n} \sum_{k=0}^{\infty} \binom{n}{k} (-1)^k f(x + (n - k)h). \quad (2.63)$$

2.7 Eine Reihe von Kehrwerten

Im Folgenden wird eine Herleitung für die Formel

$$\sum_{n=1}^{\infty} \frac{1}{4n^2 - 2n} = \ln 2 \quad (2.64)$$

beschrieben. Man spürt zunächst das Muster

$$\sum_{n=1}^{\infty} \frac{1}{4n^2 - 2n} = \sum_{n=1}^{\infty} \frac{1}{2n(2n - 1)} = \sum_{n=1}^{\infty} \frac{1}{(2n)^2}$$

auf. Man gelangt zu

$$\frac{1}{(2n)^2} = \frac{(2n - 2)!}{(2n)!} = \frac{\Gamma(2n - 1)}{\Gamma(2n + 1)} = B(2, 2n - 1),$$

wobei B die Betafunktion ist. Vermittels ihrer Definition als Parameterintegral findet man

$$\begin{aligned} \sum_{n=1}^{\infty} \frac{1}{(2n)^2} &= \sum_{n=1}^{\infty} \int_0^1 (1 - t)(t^2)^{n-1} dt \\ &= \int_0^1 (1 - t) \sum_{k=0}^{\infty} (t^2)^k dt = \int_0^1 \frac{1 - t}{1 - t^2} dt = \ln 2. \end{aligned}$$

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

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

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 behandeln. Die grundlegende Feststellung dazu ist

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

Division durch zwei bringt $\frac{1}{2}f(v) = f(\frac{1}{2}v)$. Allgemein gilt wieder $\frac{1}{n}f(v) = f(\frac{1}{n}v)$ 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

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

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

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

4 Wahrscheinlichkeitsrechnung

4.1 Bivariate Verteilungen

Gegeben ist eine bivariate Verteilung $\mathbf{X} = (X_1, X_2)$ mit Dichte $f(x_1, x_2)$. Zu berechnen ist die Kovarianzmatrix zu dieser Verteilung. Die Matrix ist definiert gemäß

$$\text{Cov}(\mathbf{X}) = E((\mathbf{X} - \mu)(\mathbf{X} - \mu)^T) \quad (4.1)$$

mit

$$\mu = E(\mathbf{X}) = E \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} = \begin{pmatrix} E(X_1) \\ E(X_2) \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}. \quad (4.2)$$

Aufgrund der Linearität von E gilt

$$\begin{aligned} \text{Cov}(\mathbf{X})_{ij} &= E((X_i - \mu_i)(X_j - \mu_j)) \\ &= E(X_i X_j - \mu_j X_i - \mu_i X_j + \mu_i \mu_j) \\ &= E(X_i X_j) - \mu_j E(X_i) - \mu_i E(X_j) + \mu_i \mu_j \\ &= E(X_i X_j) - \mu_i \mu_j - \mu_i \mu_j + \mu_i \mu_j \\ &= E(X_i X_j) - \mu_i \mu_j = E(X_i X_j) - E(X_i)E(X_j). \end{aligned}$$

Die Berechnung von $E(X_i X_j)$, $E(X_i)$, $E(X_j)$ gestaltet sich etwas komplizierter als bei univariaten Verteilungen. Der Erwartungswert einer Abbildung $g(X, Y)$ wird berechnet gemäß

$$E(g(X, Y)) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) f(x, y) dx dy, \quad (4.3)$$

wobei $f(x, y)$ die Dichte von (X, Y) ist.

Man definiert die Randdichten:

$$f_1(x) := \int_{-\infty}^{\infty} f(x, y) dy, \quad (4.4)$$

$$f_2(y) := \int_{-\infty}^{\infty} f(x, y) dx. \quad (4.5)$$

Anwendung von (4.3) ergibt nun

$$E(X) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x f(x, y) dx dy \quad (4.6)$$

$$= \int_{-\infty}^{\infty} x \left(\int_{-\infty}^{\infty} f(x, y) dy \right) dx = \int_{-\infty}^{\infty} x f_1(x) dx. \quad (4.7)$$

und

$$E(Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y f(x, y) dx dy \quad (4.8)$$

$$= \int_{-\infty}^{\infty} y \left(\int_{-\infty}^{\infty} f(x, y) dx \right) dy = \int_{-\infty}^{\infty} y f_2(y) dy. \quad (4.9)$$

Für das Produkt wird auch (4.3) angewendet:

$$E(XY) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy f(x, y) dx dy. \quad (4.10)$$

4.2 Zufallszahlengeneratoren

Ein RNG (random number generator) lässt sich modellieren als eine Folge (X_k) , wobei die Zufallsgrößen X_k identisch

verteilt und unabhängig sind. Ein RNG entsteht auch, wenn dasselbe Zufallsexperiment mit Verteilung X immer wiederholt wird, die Verteilung der X_k ist dann die Verteilung von X .

Für eine Folge von Ergebnissen (ω_k) mit $\omega_k \in \Omega$ haben die $X_k: \Omega \rightarrow \mathbb{R}$ dann die Werte $x_k = X(\omega_k)$. Die Folge (ω_k) ist aber allgemein unbekannt, nur die Folge (x_k) der Realisierungen ist gegeben, bzw. diese wird durch den RNG oder das wiederholte Experiment erzeugt.

Die uniforme Verteilung hat die Verteilungsfunktion

$$U: \mathbb{R} \rightarrow [0, 1], \quad U(x) := \begin{cases} 0 & \text{wenn } x < 0, \\ x & \text{wenn } x \in [0, 1], \\ 1 & \text{wenn } x > 1. \end{cases} \quad (4.11)$$

Möchte man eine Verteilung mit streng monoton steigender Verteilungsfunktion $F: \mathbb{R} \rightarrow [0, 1]$ simulieren, hat aber zunächst nur einen RNG mit uniformer Verteilung zur Verfügung, dann kann man wie folgt vorgehen.

Sei X die zu simulierende Zufallsgröße mit Verteilung F und X_u die Zufallsgröße mit uniformer Verteilung U . Der RNG nimmt den Wert $u = X_u(\omega)$ an. Den Wert $x = X(\omega)$ berechnet man gemäß $x = F^{-1}(u)$. Dieses Verfahren nennt man *Inversionsmethode*.

Demnach gilt $X = F^{-1} \circ X_u$. Zu zeigen ist, dass X auch die gewünschte Verteilung F besitzt. Man erhält

$$X(\omega) \leq x \iff F^{-1}(X_u(\omega)) \leq x \quad (4.12)$$

$$\iff X_u(\omega) \leq F(x). \quad (4.13)$$

Damit ergibt sich

$$P(X \leq x) = P(X_u \leq F(x)) = U(F(x)) = F(x). \quad (4.14)$$

Somit hat X tatsächlich die gewünschte Verteilung.

4.3 Gesetz der totalen Wahrscheinlichkeit

Sei Ω eine Ergebnismenge, sei $A \subseteq \Omega$ ein Ereignis und $(B_k)_{k=1}^n$ eine Zerlegung von Ω in paarweise disjunkte Ereignisse. Dann können wir rechnen

$$\begin{aligned} P(A) &= P(A \cap \bigcup_{k=1}^n B_k) = P(\bigcup_{k=1}^n (A \cap B_k)) \\ &= \sum_{k=1}^n P(A \cap B_k) = \sum_{k=1}^n P(A | B_k) P(B_k). \end{aligned}$$

Dies ist das *Gesetz der totalen Wahrscheinlichkeit*, eine der wichtigen Rechenregeln in den Grundlagen der Wahrscheinlichkeitsrechnung.

Sei $X: \Omega \rightarrow [a, b]$ eine Zufallsgröße mit einer Dichte $f: [a, b] \rightarrow \mathbb{R}$. Sei $(I_k)_{k=0}^{n-1}$ eine Zerlegung von $[a, b]$ in n gleichgroße disjunkte Intervalle. Sei $h = (b-a)/n$. Für großes n ist nach dem Gesetz der totalen Wahrscheinlichkeit die Rechnung

$$\begin{aligned} P(A) &= \sum_{k=0}^{n-1} P(A | X^{-1}(I_k)) P(X^{-1}(I_k)) \\ &\approx \sum_{k=0}^{n-1} P(A | X \in I_k) f(a + kh) h. \end{aligned}$$

gültig. Im Übergang $n \rightarrow \infty$ ist

$$P(A) = \int_a^b P(A \mid X = x) f(x) \, dx. \quad (4.15)$$

Diese Formel ist die kontinuierliche Form des Gesetzes.

Literatur

- [1] Joakim Munkhammar: »Riemann-Liouville Fractional Derivatives and the Taylor-Riemann Series«, U.U.D.M. Project Report 2004:7. Department of Mathematics, Uppsala University.