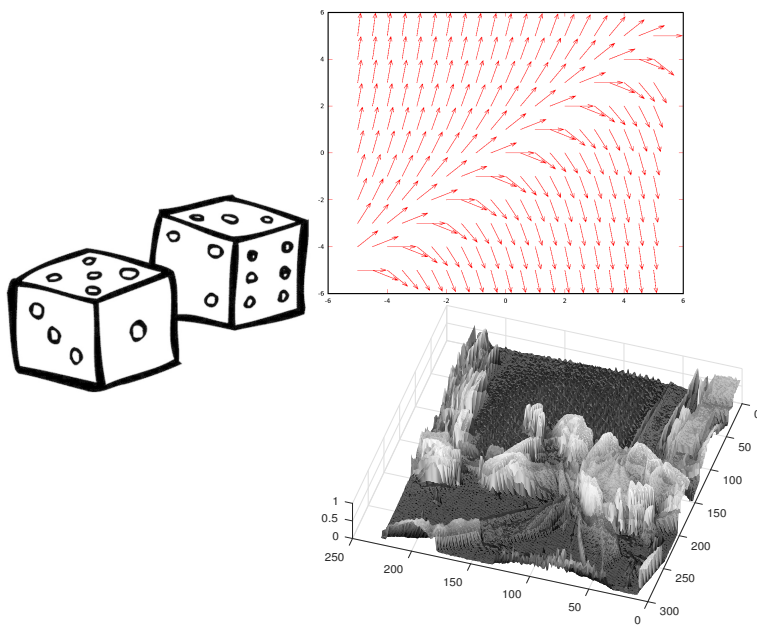


Analysis für Ingenieure

Teil III

Skript

(R. Axthelm)



Inhaltsverzeichnis

1	Gewöhnliche Differentialgleichungen	1
1.1	Gewöhnliche Differentialgleichungen erster Ordnung	2
1.2	Gewöhnliche Differentialgleichungen n -ter Ordnung mit konstanten Koeffizienten	16
1.3	Systeme von Differentialgleichungen an ausgewählten Beispielen	23
1.4	Die Laplace-Transformation	28
2	Elementare Wahrscheinlichkeitsrechnung	34
2.1	Kombinatorik	34
2.2	Grundlagen der Wahrscheinlichkeitstheorie	37
3	Differentiation in höherdimensionalen Räumen	51
3.1	Ableitungen von Funktionen mehrerer Veränderlicher	53
3.2	Extremwertberechnung	60
3.3	Optimierung mit Nebenbedingung	77
3.4	weitere Ableitungsbegriffe und ausgewählte Anwendungsbeispiele	82
A	Tabellen zur Laplace-Transformation	96
B	Matlab-Skripte	102
	Index	105

INHALTSVERZEICHNIS

Gewöhnliche Differentialgleichungen_____

Denn wer da hat, dem wird gegeben, daß er die Fülle habe; wer da nicht hat, von dem wird auch genommen, was er hat.

Matthäus-Evangelium, Kapitel 13, Vers 12

“Differentialgleichungen werden oft benötigt, um Vorgänge in der Natur zu beschreiben, bei denen das Änderungsverhalten von Größen verglichen wird.

Die ersten Differentialgleichungen waren die der gleichmäßigen und gleichmäßig beschleunigten Bewegung. Im Jahr 1590 erkannte Galileo Galilei den Zusammenhang zwischen der Fallzeit eines Körpers und seiner Fallgeschwindigkeit sowie dem Fallweg und formulierte (noch) mit geometrischen Mitteln das Gesetz des freien Falles.

Als Isaac Newton auch Bewegungen unter zum Betrag oder Quadrat der Geschwindigkeit proportionaler Reibung betrachtete, war er genötigt, die Differentialrechnung und den heute geläufigen Formalismus der Differentialgleichungen einzuführen.

Durch die exakte Formulierung des Grenzwertbegriffes, der Ableitung und des Integrals stellte schließlich Augustin Louis Cauchy im 19. Jahrhundert die Theorie der gewöhnlichen Differentialgleichungen auf ein festes Fundament und machte sie somit vielen Wissenschaften zugänglich.

Das wissenschaftliche Interesse an Differentialgleichungen ist im Wesentlichen darin begründet, dass mit ihnen auf Grund vergleichsweise einfacher Beobachtungen und Experimente vollständige Modelle geschaffen werden können.”¹

(Weiteres aus [Walter \(1972\)](#), [Heuser \(2009⁶\)](#) und [Rießinger \(2011⁸\)](#))

Definition 1.1 (Gewöhnliche Differentialgleichung (DGL/ODE)). *Unter einer gewöhnlichen Differentialgleichung (ODE = Ordinary Differential Equation) verstehen wir eine Gleichung, in welcher unabhängige Variable, Funktionen und Ableitungen von Funktionen nach genau einer Variable auftreten. Eine ODE der Form*

$$f(x, y(x), y'(x), \dots, y^{(n)}(x)) = 0$$

heißt ODE n-ter Ordnung. Wird diese Gleichung nach $y^{(n)}$ aufgelöst zu einer ODE der Form

$$y^{(n)} = f(x, y(x), y'(x), \dots, y^{(n-1)}(x))$$

so sprechen wir von einer ODE in expliziter Form; andernfalls von einer ODE in impliziter Form.

¹aus Wikipedia

1.1 Gewöhnliche Differentialgleichungen erster Ordnung

Beispiel 1 Die einfachste ODE

erster Ordnung lautet

$$y'(x) = 0.$$

Diese Gleichung lässt sich durch Integration leicht lösen:

$$y(x) = C$$

Die Integrationskonstante gibt schon den Hinweis darauf, dass es unendlich viele Lösungen gibt.

Beispiel 2 Die zweiteinfachste ODE

$$y'(x) = x$$

hat die Lösungen

$$y(x) = \frac{1}{2}x^2 + C.$$

Wir betrachten die explizite Differentialgleichung erster Ordnung

$$y'(t) = f(t, y(t)). \quad (1)$$

Dabei gelte für die rechte Seite $f : G \subset \mathbb{R}^2 \rightarrow \mathbb{R}$.

Die Differentialgleichung (1) gestattet eine einfache geometrische Interpretation. Verläuft eine Integralkurve $y(x)$ durch den Punkt $(x_0, y_0) \in G$, d.h. ist $y(x_0) = y_0$, so beträgt ihre Steigung an dieser Stelle $y'(x_0) = f(x_0, y_0)$. Durch die Differentialgleichung (1) wird also die Steigung der durch den Punkt (x_0, y_0) gehenden Lösungskurve vorgeschrieben. Diese Betrachtung lässt sich für jeden Punkt aus G anstellen. Sie führt auf die Begriffe *Linienelement* und *Richtungsfeld*. Siehe dazu Beispiel (3)².

Beispiel 3 Richtungsfeld (siehe Abb. 1)

Jedes Linienelement besteht aus dem Tripel $(x, y, f(x, y))$. Alle Linienelemente bilden in ihrer Gesamtheit, d.h. für alle $(x, y) \in G$ das Richtungsfeld. Eine Lösung der Differentialgleichung "passt" auf das Richtungsfeld. Man kann sich einen guten Überblick über den Verlauf der Lösungen verschaffen, indem man sich das Richtungsfeld aufzeichnet und dann versucht, Kurven einzuzichnen, welche auf das Richtungsfeld passen. Dieses Vorgehen legt die Vermutung nahe, dass es zu jedem Punkt $(x_0, y_0) \in G$ genau eine Lösungskurve y gibt, welche durch diesen Punkt verläuft.

²aus: Wikipedia; <http://de.wikipedia.org/wiki/Richtungsfeld>

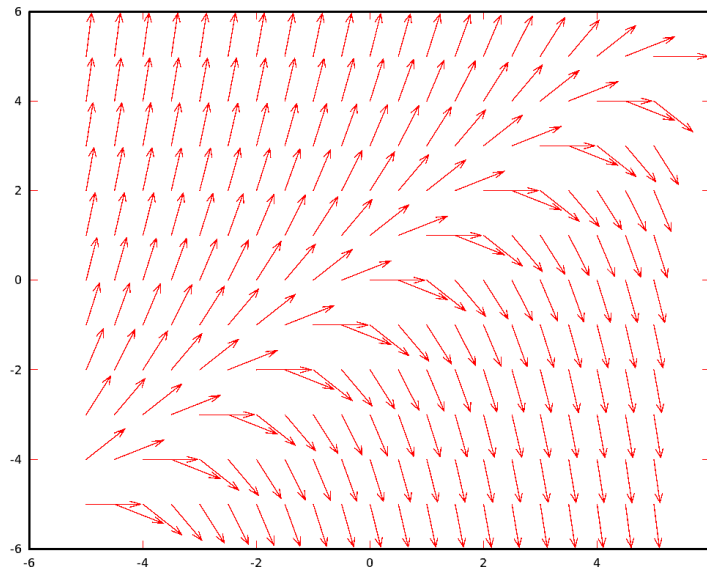


Abbildung 1: Richtungsfeld für y zur Gleichung $y' = y - x$. (src: s. Abb 14)

Definition 1.2 (Anfangswertproblem (AWP)). Gegeben ist eine auf einer Menge $G \subset \mathbb{R}^2$ erklärte Funktion $f(x, y)$ und ein fester Punkt $(x_0, y_0) \in G$. Gesucht ist eine Funktion $y \in C^1(J)$, $x_0 \in J$, für welche

$$\begin{aligned} y'(x) &= f(x, y(x)) & \text{in } J \\ y(x_0) &= y_0 & x_0 \in J \end{aligned} \quad (2)$$

gilt. Eine solche Problemstellung nennen wir Anfangswertproblem und die zweite Gleichung in (2) heißt Anfangsbedingung.

Beispiel 4 Matthäus Prinzip

$$y' = \alpha y$$



Die Nomenklaturen *Anfangsbedingung*, bzw. *Anfangswert* suggerieren, dass es sich hierbei um Werte am Anfang handelt. Das muss nicht sein. Es kann sich um irgendwelche bekannten Werte handeln. Meist ist es nur so, dass man eben die Situation am Anfang kennt. Weit philosophischer ist da noch die Fragestellung nach dem Anfang.... =-|

Definition 1.3 (Trennung der Variablen). Es seien $I, J \subset \mathbb{R}$ offene Intervalle, $f : I \rightarrow \mathbb{R}$, $g : J \rightarrow \mathbb{R}$ stetige Funktionen mit $g(y) \neq 0$ für alle $y \in J$. Dann heißt die Differentialgleichung

$$y'(x) = f(x) g(y)$$

eine Differentialgleichung mit getrennten Variablen.

Beispiel 5 "Gleiches zu Gleichem gesellt sich gern"

$$y'(x) = x y$$

Die allgemeine Vorgehensweise ist immer die Gleiche, nämlich so:

$$\begin{aligned} & \frac{dy}{dx} = f(x) g(y) \\ \Leftrightarrow & \frac{dy}{g(y)} = f(x) dx \\ \Leftrightarrow & G(y(x)) := \int \frac{dy}{g(y)} = \int f(x) dx =: F(x) \\ \Rightarrow & y(x) = G^{-1}(F(x)) \end{aligned}$$

Sind wir in der Lage die Stammfunktionen F und G zu bestimmen und dann noch G zu invertieren, so haben wir die Möglichkeit eine analytische Darstellung der Lösungen zu berechnen. Im Moment sprechen wir noch von einer Schar von Lösungen. Wollen wir eine eindeutige Lösung, also eine, die ihren Anfangswert $y(x_0) = y_0$ beinhaltet so ersetzen wir das unbestimmte Integral durch das bestimmte und stecken die geforderte Bedingung in die Integrationsgrenzen:

Formel 1: Trennung der Variablen:

Für

$$F(x) = \int_{x_0}^x f(s) ds \quad \text{und} \quad G(y) = \int_{y_0}^y \frac{1}{g(t)} dt$$

ist

$$y(x) = G^{-1}(F(x))$$

die Lösung des AWP

$$y'(x) = f(x) g(y), \quad y(x_0) = y_0.$$

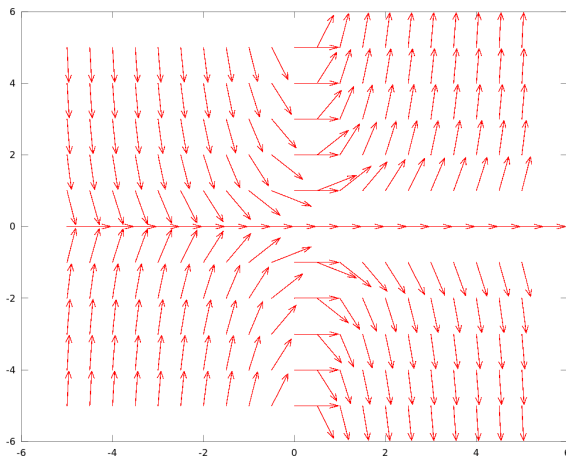
Beispiel 5 nun mit der Lösungsformel:

$$F(x) = \int_{x_0}^x s \, ds = \frac{1}{2}x^2 - \frac{1}{2}x_0^2$$

$$G(y) = \int_{y_0}^y \frac{1}{t} \, dt = \ln y - \ln y_0$$

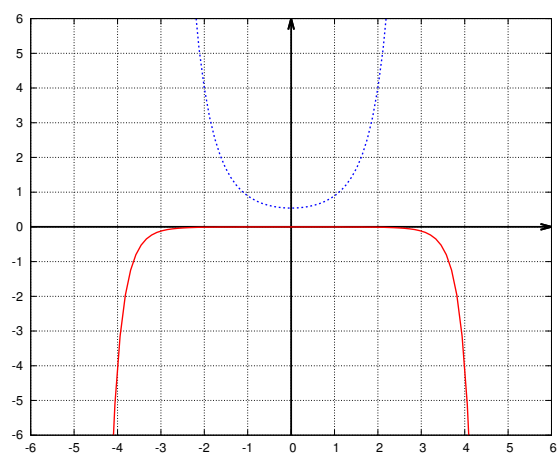
$$G^{-1}(y) = y_0 e^y$$

$$\begin{aligned} y(x) &= G^{-1}(F(x)) = y_0 e^{F(x)} \\ &= y_0 e^{\frac{1}{2}(x^2 - x_0^2)} \end{aligned}$$



Richtungsfeld für

$$y'(x) = xy.$$



$$x_0 = -4 \quad y_0 = -4 \quad y(x) = -4e^{\frac{1}{2}(x^2-16)}$$

$$x_0 = -2 \quad y_0 = 4 \quad y(x) = 4e^{\frac{1}{2}(x^2-4)}$$



Für jede auftretende Ableitung wird ein Mal (formal) integriert und es entsteht eine noch zu bestimmende Integrationskonstante. Um diese zu belegen benötigen wir eine zusätzliche Bedingung. Bei ODEs erster Ordnung kommt eine Ableitung vor, wir integrieren also ein Mal und erhalten eine Konstante, zur Bestimmung derer wir dann eine Bedingung - hier Anfangswert - benötigen. Klar?

Es lassen sich allerdings die Variablen nicht immer trennen wie zum

Beispiel 6

$$y' + \frac{y}{x} = \cos x. \quad (3)$$

Grundsätzlich handelt es sich um einen bestimmten Typ von ODE:

Definition 1.4. Es sei $I \subset \mathbb{R}$ und $a_0, \dots, a_{n-1} : I \rightarrow \mathbb{R}$ stetige Funktionen. Die Gleichung

$$y^{(n)}(x) + a_{n-1}(x)y^{(n-1)}(x) + \dots + a_0(x)y^{(0)}(x) = 0$$

heißt *homogene, lineare Differentialgleichung n -ter Ordnung mit nicht-konstanten Koeffizienten*.
Ist weiterhin $g : I \rightarrow \mathbb{R}$ eine stetige Funktion, so heißt

$$y^{(n)}(x) + a_{n-1}(x) y^{(n-1)}(x) + \cdots + a_0(x) y^{(0)}(x) = g(x)$$

inhomogene, lineare Differentialgleichung n -ter Ordnung mit nicht-konstanten Koeffizienten.

Beispiel 7

1. $y'' + 3x^2 y' + 7y = 0$ ist eine lineare, homogene Differentialgleichung zweiter Ordnung mit nicht-konstanten Koeffizienten.
2. $y' - 3 \sin x y = \cos x$ ist eine lineare, inhomogene Differentialgleichung erster Ordnung mit nicht-konstanten Koeffizienten.
3. $(y')^2 = y$ ist eine nicht-lineare, homogene Differentialgleichung erster Ordnung (mit konstanten Koeffizienten).

Der inhomogene Anteil in Beispiel 6, nämlich $\cos x$ stört bei der Separation der Variablen. Der Trick, den man hier anwendet ist wieder so eine "schöner-wär's-Regel": Schöner wär's, wenn die Gleichung homogen wäre, etwa so aussähe:

$$\tilde{y}' + \frac{\tilde{y}}{x} = 0$$

Mit Trennung der Variablen erhalten wir nun Lösungen der Form

$$\tilde{y}(x) = \frac{\tilde{C}}{x}$$

Diese Lösungen passen aber nun in die Gleichung (3) nicht rein. Wir ersetzen nun $\tilde{C}$ durch $C(x)$ und machen den Ansatz

$$y(x) = \frac{C(x)}{x}.$$

Die Idee ist nun, $C(x)$ so zu bestimmen, dass Gleichung (3) erfüllt ist.

$$y'(x) = \frac{C'(x)}{x} - \frac{C(x)}{x^2} \stackrel{!}{=} \cos x - \frac{y}{x} = \cos x - \frac{C(x)}{x^2}$$

$$\Leftrightarrow C'(x) = x \cos x$$

$$\begin{aligned} \Leftrightarrow C(x) &= \int x \cos x \, dx + \bar{C} \\ &= x \sin x + \cos x + \bar{C} \end{aligned}$$

$$\Rightarrow y(x) = \sin x + \frac{\cos x}{x} + \frac{\bar{C}}{x}$$

Nun können wir die Konstante (!!!) $\bar{C}$ bestimmen, so dass $y(x_0) = y_0$ erfüllt ist:

$$y_0 = \sin x_0 + \frac{\cos x_0}{x_0} + \frac{\bar{C}}{x_0}$$

$$\Leftrightarrow \bar{C} = y_0 x_0 - x_0 \sin x_0 - \cos x_0$$

und dann haben wir unsere Lösung beisammen! Voilà:

$$\Rightarrow y(x) = \sin x + \frac{\cos x}{x} + \frac{1}{x} (y_0 x_0 - x_0 \sin x_0 - \cos x_0) \quad (4)$$

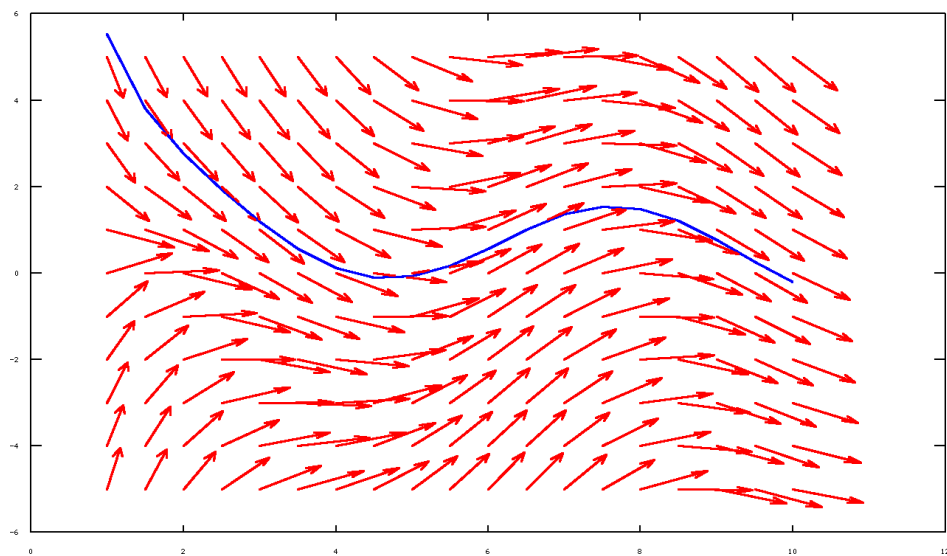


Abbildung 2: Richtungsfeld der Gleichung (3) und Lösung (4) mit $x_0 = \pi$ und $y_0 = 1$.

Formel 2: Variation der Konstanten: Gegeben sei eine inhomogene Differentialgleichung

$$y'(x) = y f(x) + g(x). \quad (5)$$

Die Variation der Konstanten führen wir nach dem folgenden Schema durch:

- Wir lösen die homogene Gleichung

$$\tilde{y}'(x) = \tilde{y} f(x)$$

mit Trennung der Variablen und erhalten das Ergebnis

$$\tilde{y}(x) = \tilde{C} e^{\int f(x) dx}.$$

- Wir variieren die Konstante $\tilde{C}$ zu eine x -Abhängigen Größe $C(x)$ und definieren die Ansatzfunktion

$$y(x) = C(x) e^{\int f(x) dx}.$$

- Wir wählen $C(x)$ nun so, dass die Gleichung (5) erfüllt ist. Das führt auf

$$C(x) = \int g(x) e^{-\int f(x) dx} dx$$

Beispiel 8 Partikulärlösung

$$y' - 2y = 4x - 2, \quad y(0) = 1$$

Störfunktion:

$$g(x) = 4x - 2 \in \mathbb{P}_1$$

Ansatz der partikulären Lösung:

$$y_p = a_1 x + a_0 \in \mathbb{P}_1$$

$$\begin{aligned} y_p' - 2y_p &= (-2a_1)x + (2a_0 + a_1) \\ &= 4x - 2 \end{aligned}$$

$\Rightarrow$

$$a_1 = -2$$

$$a_0 = 0$$

$\Rightarrow$

$$y_p = -2x$$

allg. Lsg der hom. DGL:

$$y_0 = C e^{2x}$$

allg. Lsg der inhom. DGL:

$$y = y_0 + y_p = C e^{2x} - 2x$$

Anfangsbedingung in C stecken:

$$y = e^{2x} - 2x$$

Beispiel 9 Variation der Konstanten

$$y' = \sin x y + \sin x, y(0) = 0$$

Die allgemeine Lösung der homogenen Gleichung lautet

$$y_0 = e^{\int \sin x dx} = e^{-\cos x + \tilde{C}} = \underbrace{e^{\tilde{C}}}_{=:C} e^{-\cos x} = C e^{-\cos x}.$$

Variation der Konstanten führt auf folgenden Rechenweg:

$$\begin{aligned} y &= C(x) e^{-\cos x} \\ \Rightarrow y' &= (C'(x) + \sin x C(x)) e^{-\cos x} \end{aligned}$$

eingesetzt in die inhom. ODE führt auf

$$\begin{aligned} \Leftrightarrow (C'(x) + \sin x C(x)) e^{-\cos x} &= \sin x C(x) e^{-\cos x} + \sin x \\ \Leftrightarrow C'(x) e^{-\cos x} &= \sin x \\ \Leftrightarrow C'(x) &= \sin x e^{\cos x} \\ &= -\frac{d}{dx} e^{\cos x} \\ \Rightarrow C(x) &= -e^{\cos x} + \tilde{C} \\ \Rightarrow y(x) &= (-e^{\cos x} + \tilde{C}) e^{-\cos x} \\ \Rightarrow y(x) &= \tilde{C} e^{-\cos x} - 1 \end{aligned} \tag{6}$$

Nun noch die Anfangswerte einbauen:

$$\begin{aligned} y(0) &= \tilde{C} e^{-1} - 1 \stackrel{!}{=} 0 \\ \Rightarrow \tilde{C} &= 1 \\ \Rightarrow & \end{aligned}$$

$$y(x) = e^{-\cos x} - 1$$

Mit der Variation der Konstanten erhielten wir im Zwischenschritt Gleichung (6) die allgemeine Lösung der inhomogenen Gleichung in der Form

$$y = C e^{-\cos x} - 1 = y_0 + y_p.$$

y_p ist eine partikuläre Lösung der inhomogenen Differentialgleichung, das heißt eine der unendlich vielen, möglichen Lösungen. Sie dazu Abbildung 3.

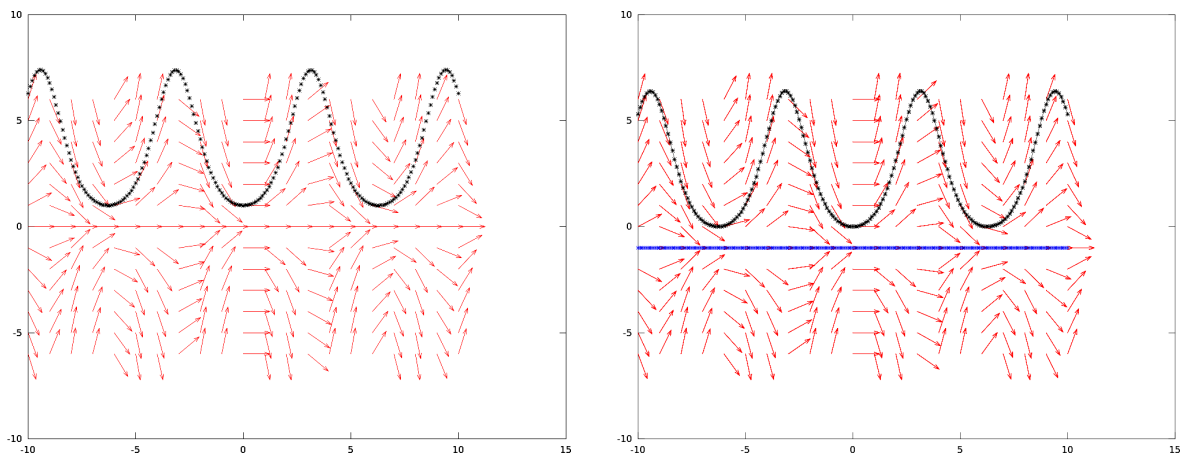


Abbildung 3: links: Richtungsfeld und allgemeine ($\star$) Lösung y_0 der homogenen Differentialgleichung; rechts: allgemeine ($\star$) und partikuläre ($\star$) Lösung y und y_p der inhomogenen Differentialgleichung



Die partikuläre Lösung ist eine Funktion, die in die inhomogene DGL passt. Sehr wahrscheinlich ist sie aber nicht diejenige, die die Anfangsbedingung erfüllt. Das genügt also nicht. Zusammen ($y = y_0 + y_p$) mit der allgemeinen Lösung der homogenen Gleichung, die ja eine zu bestimmende Konstante enthält, kann y dann sowohl die inhomogene DGL als auch die Anfangsbedingung erfüllen.

Formel 3: Superpositionsprinzip: Zur Differentialgleichung mit konstanten Koeffizienten

$$y'(x) = \alpha y(x) + g(x) \quad (7)$$

ist $y_0 = C e^{\alpha x}$ die allgemeine Lösung der homogenen Differentialgleichung

$$y_0'(x) = \alpha y_0(x). \quad (8)$$

Die allgemeine Lösung der inhomogenen Differentialgleichung (7) ergibt sich aus der Summe der allgemeinen Lösung der homogenen Differentialgleichung (8) und einer partikulären Lösung y_p :

$$y = y_0 + y_p$$

Eine partikuläre Lösung erhalten wir, indem wir eine Ansatzfunktion gemäß der nachstehenden Tabelle wählen und die auftretenden Konstanten so bestimmen, dass y_p die inhomogenen Gleichung (7) erfüllt.

Störfunktion $g(x)$	Lösungsansatz $y_p(x)$
$\sum_{k=0}^n b_k x^k$	$\sum_{k=0}^n A_k x^k$, falls $\alpha \neq 0$
$b e^{ax}$	$A e^{ax}$, falls $\alpha \neq a$ $x A e^{ax}$, falls $\alpha = a$
$b \cos(cx)$ oder $b \sin(cx)$	$A \cos(cx) + B \sin(cx)$

Besteht die Störfunktion aus einer Summe oder einem Produkt der angegebenen Funktionstypen so bildet man die Ansatzfunktion ebenfalls als Summe oder Produkt der entsprechenden Funktionstypen.

Beispiel 10 Partikulärlösung

(a)

$$y' = 2y + 4e^{3x}$$

$$\begin{array}{ll}
\text{Lsg der hom. ODE} & y_0 = C e^{2x} \\
\text{Ansatz der Partikulärlösung} & y_p = A e^{3x} \quad (a \neq \alpha) \\
\Rightarrow & y'_p = 3 A e^{3x} \\
\Rightarrow & 3 A e^{3x} = 2 A e^{3x} + 4 e^{3x} \\
\Leftrightarrow & A e^{3x} = 4 e^{3x} \\
\Leftrightarrow & A = 4 \\
\Rightarrow & y_p = 4 e^{3x} \\
\Rightarrow &
\end{array}$$

$$y(x) = C e^{2x} + 4 e^{3x}$$

(b)

$$y' = 2y + 4e^{2x}$$

$$\begin{array}{ll}
\text{Lsg der hom. ODE} & y_0 = C e^{2x} \\
\text{Ansatz der Partikulärlösung} & y_p = x A e^{2x} \quad (a = \alpha) \\
\Rightarrow & y'_p = A(1 + 2x) e^{2x} \\
\Rightarrow & A(1 + 2x) e^{2x} = 2x A e^{2x} + 4 e^{2x} \\
\Leftrightarrow & A e^{2x} = 4 e^{2x} \\
\Leftrightarrow & A = 4 \\
\Rightarrow & y_p = 4x e^{2x} \\
\Rightarrow &
\end{array}$$

$$y(x) = (C + 4x) e^{2x}$$

Beispiel 11 Ein AWP und Partikulärlösung

 Gesucht ist die Funktion $y : \mathbb{R}_0^+ \Rightarrow \mathbb{R}$ mit

$$y' + 3y = \sin(2x) e^x, \quad y(0) = 0.$$

$$\text{Lsg der hom. ODE} \quad y_0 = C e^{-3x}$$

Beim Ansatz für die partikuläre Lösung beachten wir nun, dass die Störfunktion aus einem Produkt einer Sinus- und einer e-Funktion besteht. Die Ansatzfunktion y_p müssen wir nun entsprechend ausrichten:

$$\text{Ansatz der Partikulärlösung} \quad y_p = (a \cos(2x) + b \sin(2x)) e^x$$

Wir setzen y_p in die inhom. Gln ein und erhalten

$$0 = \cos(2x) \underbrace{(2a + 4b)}_{\stackrel{!}{=} 0} + \sin(2x) \underbrace{(4a - 2b - 1)}_{\stackrel{!}{=} 0}$$

$$\Rightarrow a = \frac{1}{5}, b = \frac{-1}{10}$$

$$\Rightarrow y_p = \frac{e^x}{10} (2 \sin(2x) - \cos(2x))$$

$$\Rightarrow y = C e^{-3x} + \frac{e^x}{10} (2 \sin(2x) - \cos(2x))$$

Nun noch C so wählen, dass der Anfangswert $y(0) = 0$ erfüllt ist

$$y(0) = C - \frac{1}{10} = 0$$

$$\Rightarrow C = \frac{1}{10}$$

$\Rightarrow$

$$y(x) = \frac{1}{10} (e^{-3x} + e^x (2 \sin(2x) - \cos(2x)))$$

Gelegentlich gibt es Gleichungen, die sich durch Trennung der Variablen lösen lassen auch wenn sie zunächst gar nicht danach aussehen. Das ist dann der Fall, wenn man eine Gleichung mit einer geeigneten Substitution auf eine Differentialgleichung mit getrennten Variablen zurückführen kann.

Beispiel 12 Substitution

$$y' = \frac{2y - x}{x}$$

lässt sich zunächst besser lesen als

$$y' = 2 \frac{y}{x} - 1.$$

Die Methode durch Trennung der Variablen ist direkt nicht möglich. Wir könnten hier mit Variation der Konstanten ran oder aber mit einer geeigneten Substitution: Wähle $u = \frac{y}{x}$ und wir erhalten eine neue Differentialgleichung, nämlich

$$u' = \frac{u - 1}{x}.$$

Hier können wir direkt mit Trennung der Variablen weiterrechnen und erhalten die Lösung

$$u = Cx - 1.$$

Wir machen die Rücktransformation $y = ux$ und erhalten die Lösungen der ursprünglichen Gleichung

$$y = (Cx + 1)x.$$

Formel 4: Substitution: Gegeben sei die lineare DGL erster Ordnung mit

$$y' = f(x, y)$$

Je nach Form von f kann eine geschickte Substitution die DGL in eine, leichter zu lösende umgewandelt werden.

$y' = f(ax + by + c), b \neq 0$	Substitution: $u := ax + by + c$
---------------------------------	-------------------------------------

Löse $u' = bf(u) + a$	Rücktransformation: $y = \frac{1}{b}(u - ax - c)$
--------------------------	--

$y' = f\left(\frac{y}{x}\right)$	Substitution: $u := \frac{y}{x}$
----------------------------------	-------------------------------------

Löse $u' = \frac{1}{x}(f(u) - u)$	Rücktransformation: $y = ux$
--------------------------------------	---------------------------------

Das sind natürlich nicht alle Substitutionsmöglichkeiten aber die typischen. In anderen Fällen lohnt es sich aber sicher, über Möglichkeiten nachzudenken.

Bisher haben wir uns mit homogenen und inhomogenen ODE erster Ordnung mit konstanten und nicht konstanten Koeffizienten beschäftigt. Alle waren linear, das heißt, dass y in den verschiedenen Termen nie im Produkt mit sich oder einer seiner Ableitungen auftrat. Nur für ODE erster Ordnung, da wir gerade die Methode der Substitution kennengelernt haben, betrachten wir ein paar Beispiele von nichtlinearen ODE erster Ordnung.

Beispiel 13 nichtlineare ODE mit Substitution lösen

$$y' = (x + y + 1)^3 - 1$$

Substitution: $z = x + y + 1$

Dann gilt $z' = y' + 1$ bzw. $y' = z' - 1$ und das führt auf

$$z' = z^3.$$

Mit Trennung der Variablen gelöst ergibt das

$$z = \frac{1}{\sqrt{-2(x+C)}}$$

und nach Rücktransformation erhalten wir die Lösung

$$y = \frac{1}{\sqrt{-2(x+C)}} - x - 1.$$

Beispiel 14

$$y' = 1 + \frac{y}{x} + \frac{y^2}{x^2}$$

Substitution: $z = \frac{y}{x}$

Dann gilt $y = x z$ und $\frac{dy}{dx} = z + x z'$, was auf

$$z' = \frac{1 + z^2}{x}$$

führt. Durch Trennung der Variablen erhalten wir die Lösung $z = \tan(\ln |x| + C)$ und insgesamt nach der Rücksubstitution die Gesamtlösung

$$y = \tan(\ln |x| + C) x.$$

Eine nicht lineare DGL ist dann gegeben, wenn der Differentialoperator $L(y)$ einer DGL

$$L(y) = g$$

nicht linear ist, das heißt

$$L(\alpha y_1 + \beta y_2) \neq \alpha L(y_1) + \beta L(y_2)$$

gilt. Zum Beispiel ist das bei Beispiel 7.3, Beispiel 13 und Beispiel 14 jeweils der Fall.

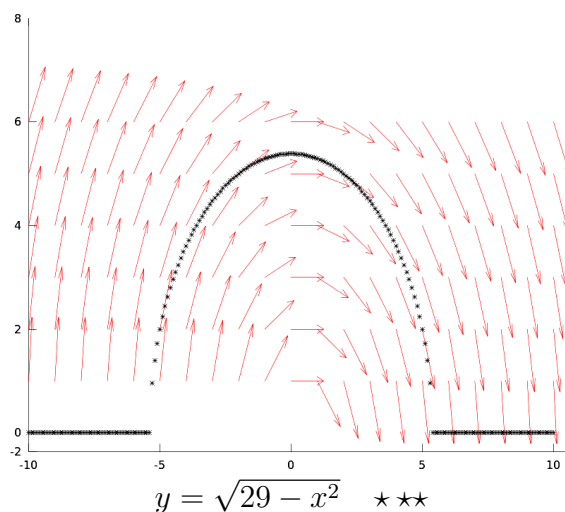


Beispiel 15 Ein nicht lineares Beispiel Hier wird mit bekannten Tricks gearbeitet.

$$\begin{aligned}
 y' &= -\frac{x}{y} \\
 \Leftrightarrow y'y &= -x \\
 \Leftrightarrow \frac{1}{2} (y^2)' &= -x \\
 \Leftrightarrow y &= \pm \sqrt{C - x^2}
 \end{aligned}$$

Mit $y(-5) = 2$ etwa

$$\Rightarrow y = \sqrt{29 - x^2}$$



Die Lösbarkeit von AWP's der Form

$$y' = f(x, y), \quad y(x_0) = y_0$$



und die Eindeutigkeit der Lösung wird bereits durch die bloße und stetige Differenzierbarkeit der rechten Seite f sichergestellt. Da wir bisher nicht genügend mathematische Begriffe und Grundlagen beisammen haben, um, den an dieser Stelle sehr wichtigen Satz von **Picard-Lindelöf** zitieren zu können, begnügen wir uns damit, dass Existenz und Eindeutigkeit der Lösung sichergestellt ist, so lange f genügend schön ist. Was schön genau bedeutet möge der geneigte Leser in [Heuser \(2009⁶\)](#), S. 139 nachschlagen

1.2 Gewöhnliche Differentialgleichungen n -ter Ordnung mit konstanten Koeffizienten

Eine inhomogene, lineare Differentialgleichung n -ter Ordnung die Form

$$y^{(n)}(x) + a_{n-1}(x) y^{(n-1)}(x) + \cdots + a_0(x) y^{(0)}(x) = g(x).$$

Zu einer ODE n -ter Ordnung mit stetigen Koeffizienten a_i , $i = 0, \dots, n-1$ (konstante Koeffizienten sind insbesondere stetig!) können wir n sogenannte Grundlösungen $y_j(x)$, $j = 1, \dots, n$ berechnen. Deren Linearkombination wird uns die gesuchten Lösungen

$$y(x) = C_1 y_1(x) + \cdots + C_n y_n(x)$$

des Gleichungssystems liefern. Diese Linearkombination enthält also demzufolge n noch zu bestimmende Konstanten C_i , die eine eindeutige Lösung "fixieren". Wir benötigen zu einer ODE n -ter Ordnung n Bedingungen, um alle freien Konstanten zu berechnen.

Definition 1.5 (lineare (un-)abhängig). Eine Menge von Funktionen $\{f_1, \dots, f_m\}$ heißt linear unabhängig, wenn man keine der beteiligten Funktionen aus den anderen linear

kombinieren kann, das heißt es gilt

$$\sum_{j=1}^m \alpha_j f_j(x) = 0 \quad \Rightarrow \quad \alpha_1 = \dots = \alpha_m = 0, \quad \forall x.$$

Kann man eine Funktion f_i aus den anderen linear kombinieren so heißt f_i linear abhängig.

Beispiel 16

1.

$$p_1(x) = 1, \quad p_2(x) = x$$

sind linear unabhängig, denn $1 \neq c x$.

2.

$$f_1(x) = 2 \cos x, \quad f_2(x) = \sin\left(x + \frac{\pi}{2}\right)$$

sind linear abhängig, denn es gilt $f_1(x) = 2 f_2(x)$.

Satz 1.6. Eine lineare, homogene ODE n -ter Ordnung besitzt n linear unabhängige Lösungen

$$y_1(x), \dots, y_n(x),$$

die sogenannten Grundlösungen der ODE.

Es ist nicht immer so einfach wie in Beispiel 16 die lineare Unabhängigkeit zu prüfen. Die Wronski-Determinante aus Definition 1.7 und der darauffolgende Satz 1.8 werden dabei eine Hilfe sein.

Definition 1.7 (Wronski-Determinante). Es seien Funktionen $y_1, \dots, y_n$ gegeben. Dann heißt die Funktion

$$W(x) = \det \begin{pmatrix} y_1(x) & \dots & y_n(x) \\ \vdots & & \vdots \\ y_1^{(n-1)}(x) & \dots & y_n^{(n-1)}(x) \end{pmatrix}$$

Wronski-Determinante.

Satz 1.8. Funktionen $y_1, \dots, y_n$ sind genau dann linear unabhängig, wenn die entsprechende Wronski-Determinante nicht verschwindet, also $W(x) \neq 0$ für ein x .

Beweis Satz 1.8:

für $n = 2$:

$$\begin{aligned}
 W(x) &= \det \begin{pmatrix} e^{\lambda_1 x} & e^{\lambda_2 x} \\ \lambda_1 e^{\lambda_1 x} & \lambda_2 e^{\lambda_2 x} \end{pmatrix} \\
 &= (\lambda_2 - \lambda_1) e^{(\lambda_1 + \lambda_2)x} = 0 \\
 \Leftrightarrow \quad \lambda_1 &= \lambda_2
 \end{aligned}$$

□

Beispiel 17

$$y'' - \frac{1}{2x}y' + \frac{1}{2x^2}y = 0$$

ist eine ODE 2-ter Ordnung und besitzt, woher auch immer wir das wissen, die Grundlösungen

$$y_1(x) = x \quad \text{und} \quad y_2(x) = \sqrt{x}.$$

Diese Funktionen und deren jeweilige erster Ableitung

$$y_1'(x) = 1 \quad \text{und} \quad y_2'(x) = \frac{1}{2\sqrt{x}}$$

bilden die Wronski-Determinante

$$W(x) = \det \begin{pmatrix} x & \sqrt{x} \\ 1 & \frac{1}{2\sqrt{x}} \end{pmatrix} = \frac{-1}{2}\sqrt{x} \neq 0.$$

Dann erhalten wir damit die Lösungsschar

$$y(x) = C_1 x + C_2 \sqrt{x}.$$

Jede weitere Lösung wird linear abhängig von den beiden gegebenen sein und ist folglich irrelevant.



Wenn Sie also zu einer linearen, homogenen Differentialgleichung n -ter Ordnung n linear unabhängige Lösungen gefunden haben wissen Sie, dass Sie die weitere Suche einstellen können. Sie werden mehr nicht finden. Die lineare Unabhängigkeit lässt sich über die Wronski-Determinante bestimmen.

Schön und gut, aber was sollen wir damit anfangen, wenn das Auffinden der Grundlösungen noch nicht klar ist? Die muss man ja auch erst "entdecken"....

Idee: Jede **homogene, lineare ODE n -ter Ordnung mit konstanten Koeffizienten** kann einem sogenannten charakteristischen Polynom zugeordnet werden. Die Nullstellen eines solchen Polynoms definieren dann eine Lösung der ODE.

Definition 1.9 (Charakteristisches Polynom). *Es sei*

$$\sum_{k=0}^n a_k y^{(k)}(x) = 0$$

eine homogene, lineare ODE n -ter Ordnung mit konstanten Koeffizienten, dann heißt

$$p(x) = \sum_{k=0}^n a_k x^k$$

das charakteristische Polynom der Differentialgleichung.

Satz 1.10. *Es sei*

$$\sum_{k=0}^n a_k y^{(k)}(x) = 0$$

eine homogene, lineare ODE n -ter Ordnung mit mit konstanten Koeffizienten und $\lambda \in \mathbb{R}$ eine Nullstelle des charakteristischen Polynoms der ODE. Dann ist

$$y(x) = e^{\lambda x}$$

eine Lösung der ODE.

Beweis 1.10:

$$\begin{aligned} & \sum_{k=0}^n a_k y^{(k)}(x) = 0 \\ \Leftrightarrow & \sum_{k=0}^n a_k (e^{\lambda x})^{(k)} = 0 \\ \Leftrightarrow & \sum_{k=0}^n a_k \lambda^k e^{\lambda x} = 0 \\ \Leftrightarrow & \sum_{k=0}^n a_k \lambda^k = 0 \end{aligned}$$

Damit ist λ Nullstelle des charakteristischen Polynoms

$$p(x) = \sum_{k=0}^n a_k x^k$$

□

Das charakteristische Polynom zu

$$y'' - 3y' + 2y = 0$$

lautet

$$p(x) = x^2 - 3x + 2$$

und hat die Nullstellen

$$\lambda_{1,2} = \frac{3}{2} \pm \frac{1}{2}$$

und damit sind

$$y_1(x) = e^x, y_2(x) = e^{2x}$$

Lösungen der obigen ODE.

Beispiel 19

Das charakteristische Polynom zu

$$y''' - y'' - 2y' = 0$$

hat die Nullstellen

$$\lambda_1 = 0, \lambda_{2,3} = \frac{1}{2} \pm \frac{3}{2}i,$$

was auf die Lösungen

$$y_1(x) = 1, y_2(x) = e^{\frac{1}{2}x}, y_3(x) = e^{-\frac{1}{2}x}$$

führt.

Satz 1.11. *Es sei*

$$\sum_{k=0}^n a_k y^{(k)}(x) = 0$$

eine homogene, lineare ODE n-ter Ordnung mit konstanten Koeffizienten und $\lambda_i \in \mathbb{R}$, $i = 1, \dots, n$ Nullstellen des charakteristischen Polynoms der ODE. Dann bilden die Funktionen

$$y_i(x) = e^{\lambda_i x}$$

ein Fundamentalsystem der Differentialgleichung. Jede Lösung der ODE hat deshalb die Form

$$y(x) = \sum_{i=1}^n C_i e^{\lambda_i x}.$$

Beispiel 20

Die Lösung der Differentialgleichung in Beispiel 18 lautet

$$y(x) = C_1 e^x + C_2 e^{2x}.$$

Die Lösung der Differentialgleichung in Beispiel 19 lautet

$$y(x) = C_1 + C_2 e^{2x} + C_3 e^{-x}.$$

Beispiel 21

Das charakteristische Polynom zu

$$y'' - 2y' + y = 0$$

lautet $p(x) = (x - 1)^2$ hat die Nullstelle $\lambda = 1$ mit algebraischer Vielfachheit 2. Das zugehörige Fundamentalsystem lautet dann

$$y_1(x) = e^x, \quad y_2(x) = x e^x$$

und die Lösung der ODE demzufolge

$$y(x) = (C_1 + C_2 x) e^x.$$

Beispiel 22

Das charakteristische Polynom zu

$$y'' - 4y' + 4y = 0$$

lautet $p(x) = (x - 2)^2$. Daraus ergibt sich die Lösungsmenge zu

$$y(x) = C_1 e^{2x} + C_2 x e^{2x}.$$

Beispiel 23

Das charakteristische Polynom zu

$$y'' + y = 0$$

lautet $p(x) = x^2 + 1$ und hat die komplexen Nullstellen $\lambda_{1/2} = \pm i$. Wir wählen als Fundamentalsystem Real- und Imaginärteil von $\varphi(x) = e^{ix}$, also

$$y_1(x) = \sin x, \quad y_2(x) = \cos x.$$

Die allgemeine Lösung lautet dann

$$y(x) = C_1 \cos x + C_2 \sin x.$$

Formel 5: Lösungen über das charakteristische Polynom bestimmen:

Zu einer homogenen, linearen ODE n -ter Ordnung mit konstanten Koeffizienten

$$y^{(n)} + a_{n-1} y^{(n-1)} + \dots + a_1 y' + a_0 y = 0$$

berechnen Sie die Linearfaktorzerlegung des charakteristischen Polynoms

$$p(x) = (x - \lambda_1)^{m_1} \dots (x - \lambda_s)^{m_s}, \quad \lambda_i \in \mathbb{R}, \quad i = 1, \dots, s \leq n.$$

Dann bilden die Funktionen

$$\begin{aligned} &e^{\lambda_1 x}, x e^{\lambda_1 x}, \dots, x^{(m_1-1)} e^{\lambda_1 x}, \\ &\vdots \\ &e^{\lambda_s x}, x e^{\lambda_s x}, \dots, x^{(m_s-1)} e^{\lambda_s x} \end{aligned}$$

ein Fundamentalsystem der Differentialgleichung und wir erhalten als Lösungsmenge mit noch zu bestimmenden $C^{kl} \in \mathbb{R}$

$$y(x) = \sum_{k=1}^s \sum_{l=0}^{m_k-1} C^{kl} x^l e^{\lambda_k x},$$

wobei $\sum_{l=1}^s m_l = n$ gilt.

Formel 6: komplexe Nullstellen für das charakteristische Polynom:

Bei komplexen Nullstellen: Es sei

$$y'' + a_1 y' + a_0 = 0$$

eine ODE, deren charakteristisches Polynom die komplexen Nullstellen $\lambda_{1,2} = a \pm i b$ besitzt. Dann ist

$$\varphi(x) = e^{(a+ib)x}$$

eine komplexe Lösung der ODE. Ein reelles Fundamentalsystem erhalten wir durch

$$y_1(x) = \operatorname{Re}(\varphi(x)) = e^{ax} \cos(bx), \quad y_2(x) = \operatorname{Im}(\varphi(x)) = e^{ax} \sin(bx).$$



Haben wir eine inhomogene ODE vorliegen so gehen wir vor wie im ersten Fall, das heißt, dass wir entweder die Methode Variation der Konstanten wählen oder das Superpositionsprinzip. Sie dürfen sich das aussuchen.

Beispiel 24

Zur ODE

$$y^{(3)}(x) + y''(x) + y'(x) + y(x) = 2x + 5$$

betrachten wir zunächst das homogene Pendant:

$$y^{(3)}(x) + y''(x) + y'(x) + y(x) = 0$$

Das entsprechende charakteristische Polynom lautet

$$p(x) = (x + 1)(x + i)(x - i).$$

Daraus ergeben sich die drei Grundlösungen

$$y_1(x) = e^{-x}, \quad y_2(x) = \cos x \quad \text{und} \quad y_3(x) = \sin x.$$

Wir erhalten die allgemeine Lösung der homogenen ODE

$$y_0(x) = C_1 e^{-x} + C_2 \cos x + C_3 \sin x.$$

Zusammen mit der partikulären Lösung

$$y_p(x) = 2x + 3$$

erhalten wir die allgemeine Lösung der inhomogenen ODE als

$$y(x) = C_1 e^{-x} + C_2 \cos x + C_3 \sin x + 2x + 3.$$

1.3 Systeme von Differentialgleichungen an ausgewählten Beispielen

Formel 7: Auflösung des homogenen Systems: Die Lösung zu

$$\mathbf{w}' = \mathbf{A} \mathbf{w}, \quad \mathbf{A} \in \mathbb{R}^{n \times n}, \quad \mathbf{w} \in \mathbb{R}^n$$

ist gegeben durch

$$\mathbf{w} = e^{\mathbf{A}x} \mathbf{C}.$$

$\mathbf{w}$ lässt sich zerlegen in

$$\mathbf{w} = \sum_{i=1}^n C_i v_i e^{\lambda_i x},$$

wobei $v_i \in \mathbb{R}^n, i = 1, \dots, n$ die Eigenvektoren von A zu den n verschiedenen Eigenwerten λ_i sind.

Beispiel 25

$$\begin{aligned} u' &= u + v & u(0) &= 0 \\ v' &= 4u - 2v & v(0) &= 5 \end{aligned}$$

Eigenwerte:

$$\lambda_1 = 2, \lambda_2 = -3$$

Eigenvektoren:

$$v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, v_2 = \begin{pmatrix} 1 \\ -4 \end{pmatrix}$$

Allgemeine Lösung:

$$\mathbf{w} = C_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} e^{2x} + C_2 \begin{pmatrix} 1 \\ -4 \end{pmatrix} e^{-3x}$$

Anfangswerte einbringen liefert:

$$C_1 = 1, C_2 = -1$$

Lösung des AWP:

$$\begin{aligned} u(x) &= e^{2x} - e^{-3x} \\ v(x) &= e^{2x} + 4e^{-3x} \end{aligned}$$

Heuristik: Wie kann man dieses Vorgehen einsehen?

$$\mathbf{w}' = \mathbf{A} \mathbf{w} \Rightarrow \mathbf{w} = e^{\mathbf{A}x} \tilde{\mathbf{C}}$$

Aus der Linearen Algebra wissen wir

$$\mathbf{A} = \mathbf{T}^{-1} \mathbf{D} \mathbf{T}$$

$$\mathbb{R}^n \xrightarrow{\mathbf{A}} \mathbb{R}^n$$

und

$$\mathbf{T} \downarrow$$

$$e^{\mathbf{A}x} = e^{\mathbf{T}^{-1} \mathbf{D} \mathbf{T} x} = \mathbf{T}^{-1} e^{\mathbf{D}x} \mathbf{T}$$

$$\mathbb{R}^n \xrightarrow{\mathbf{D}} \mathbb{R}^n$$

gilt.

Damit erhalten wir insgesamt sowas wie

$$\begin{aligned} \mathbf{w} &= e^{\mathbf{A}x} \tilde{\mathbf{C}} = \mathbf{T}^{-1} e^{\mathbf{D}x} \mathbf{T} \tilde{\mathbf{C}} \\ &= (v_1, \dots, v_n) \begin{pmatrix} e^{\lambda_1 x} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & e^{\lambda_n x} \end{pmatrix} \underbrace{\mathbf{T} \tilde{\mathbf{C}}}_{=: \mathbf{C}}, \quad \mathbf{C} = \begin{pmatrix} C_1 \\ \vdots \\ C_n \end{pmatrix} \\ &= C_1 v_1 e^{\lambda_1 x} + \dots + C_n v_n e^{\lambda_n x} = \sum_{i=1}^n C_i v_i e^{\lambda_i x} \end{aligned}$$

Beispiel 26 stark vereinfachtes Räuber-Beute-Modell (homogen)

$$\begin{aligned} u'(t) &= u(t) - \frac{1}{10} v(t), & u(0) &= 10 \\ v'(t) &= -v(t) + u(t), & v(0) &= 50 \end{aligned}$$

$$w = (u, v)^T:$$

$$w'(t) = \begin{pmatrix} 1 & -\frac{1}{10} \\ 1 & -1 \end{pmatrix} w$$

$$w(t) = e^{At} w_0, \quad A = \begin{pmatrix} 1 & -\frac{1}{10} \\ 1 & -1 \end{pmatrix}, \quad w_0 = \begin{pmatrix} 10 \\ 50 \end{pmatrix}$$

Ist A diagonalisierbar so gibt es Transformationsmatrizen T mit $A = T^{-1} D T$ und

$$D = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix},$$

wobei λ_1, λ_2 Eigenwerte der Matrix A sind und $T^{-1} = (v_1, v_2)$ mit v_i ist Eigenvektor von A zum Eigenwert $\lambda_i, i = 1, 2$.

$$\begin{aligned} \lambda_1 &= -\frac{3}{10} \sqrt{10} & v_1 &= \begin{pmatrix} 1 - \frac{3}{10} \sqrt{10} \\ 1 \end{pmatrix} \\ \lambda_2 &= \frac{3}{10} \sqrt{10} & v_2 &= \begin{pmatrix} 1 + \frac{3}{10} \sqrt{10} \\ 1 \end{pmatrix} \end{aligned}$$

Daraus ergibt sich die Transformationsmatrix

$$T = \begin{pmatrix} -1 & \frac{3}{10} \sqrt{10} + 1 \\ 1 & \frac{3}{10} \sqrt{10} - 1 \end{pmatrix} \frac{\sqrt{10}}{6}$$

und die Lösung

$$\begin{aligned}w(t) &= T^{-1} e^{Dt} T w_0 \\&= \begin{pmatrix} (3\sqrt{10} - 5) e^{-\frac{3}{\sqrt{10}}t} + (3\sqrt{10} + 5) e^{\frac{3}{\sqrt{10}}t} \\ (40 + 15\sqrt{10}) e^{-\frac{3}{\sqrt{10}}t} (15\sqrt{10} - 40) e^{\frac{3}{\sqrt{10}}t} \end{pmatrix} \frac{\sqrt{10}}{6} \\&= \begin{pmatrix} (3\sqrt{10} - 5) e^{\lambda_1 t} + (3\sqrt{10} + 5) e^{\lambda_2 t} \\ (40 + 15\sqrt{10}) e^{-\lambda_1 t} (15\sqrt{10} - 40) e^{\lambda_2 t} \end{pmatrix} \frac{\sqrt{10}}{6} \\&= \frac{1}{6} \begin{pmatrix} 30 - 5\sqrt{10} \\ 40\sqrt{10} + 150 \end{pmatrix} e^{\lambda_1 t} + \frac{1}{6} \begin{pmatrix} 30 + 5\sqrt{10} \\ 150 - 40\sqrt{10} \end{pmatrix} e^{\lambda_2 t} \\&= \frac{40\sqrt{10} + 150}{6} \left(1 - \frac{3}{10} \sqrt{10} \right) e^{\lambda_1 t} + \frac{150 - 40\sqrt{10}}{6} \left(1 + \frac{3}{10} \sqrt{10} \right) e^{\lambda_2 t} \\&= \frac{40\sqrt{10} + 150}{6} v_1 e^{\lambda_1 t} + \frac{150 - 40\sqrt{10}}{6} v_2 e^{\lambda_2 t} \\&= \sum_{j=1}^2 C_j v_j e^{\lambda_j t}\end{aligned}$$

Beispiel 27 algebraische Vielfachheit 2

$$\begin{aligned}u' &= 3u - 4v & u(0) &= 3 \\v' &= u - v & v(0) &= 1\end{aligned} \tag{9}$$

Eigenwerte:

$$\lambda = 1$$

Ansatz für die Allgemeine Lösung:

$$\begin{aligned}u(x) &= (a + bx) e^x \\v(x) &= (c + dx) e^x\end{aligned}$$

Einsetzen in (9) und Koeffizientenvergleich:

$$\left. \begin{aligned}a + b &= 3a - 4c \\b &= 3b - 4d \\c + d &= a - c \\d &= b - d\end{aligned} \right\} \Rightarrow c = \frac{a}{2} - \frac{b}{4}, d = \frac{b}{2}$$

Allgemeine Lösung:

$$\begin{aligned}u(x) &= (a + bx) e^x \\v(x) &= \left(\frac{a}{2} - \frac{b}{4} - \frac{b}{2}x \right) e^x\end{aligned}$$

Anfangswerte einbringen liefert:

$$a = 3, b = 2$$

Lösung des AWP:

$$u(x) = (3 + 2x) e^x$$

$$v(x) = (1 + x) e^x$$



Für jede höhere Vielfachheit erweitert man den Lösungsansatz um entsprechend viele Polynomgrade. Beim Auftreten von komplexen Nullstellen verfährt man analog zum entsprechenden Fall in Kapitel 1.2, S. 22. (Siehe dazu auch Heuser (2009⁶), S. 470, Bsp 51.4)

Die inhomogenen DGL behandelt man mit der Methode der Variation der Konstanten:

Beispiel 28

$$u' = 2u - v + 1 \quad u(0) = 0$$

$$v' = -u + 2v - 1 \quad v(0) = 1$$

$$w' = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} w + \begin{pmatrix} 1 \\ -1 \end{pmatrix} =: Aw + b$$

EWe von A :

$$\lambda_1 = 1, \lambda_2 = 3$$

EVec von A :

$$v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, v_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

Allgemeine Lösung des homogenen Systems:

$$w_0 = C_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} e^x + C_2 \begin{pmatrix} 1 \\ -1 \end{pmatrix} e^{3x}$$

Ansatz der Lösung des inhomogenen Systems:

$$w = C_1(x) \begin{pmatrix} 1 \\ 1 \end{pmatrix} e^x + C_2(x) \begin{pmatrix} 1 \\ -1 \end{pmatrix} e^{3x}$$

Einsetzen in die DGLs liefert:

1.4 Die Laplace-Transformation

Die Idee der Laplace-Transformation besteht darin, lineare Differentialgleichungen mit konstanten Koeffizienten als algebraische Gleichungen darzustellen, deren Lösung zurücktransformiert dann eine potentielle Lösung der Dgl darstellt; ein Verfahren, das sich gerade unter Ingenieuren und Physikern großer Beliebtheit erfreut. Es werden dabei nur ODEs mit konstanten Koeffizienten und zu denen AWP's behandelt.

Definition 1.12 (Laplace-Integral). Sei $f : \mathbb{R}_0^+ \rightarrow \mathbb{R}$, dann heißt

$$F(s) := \int_0^{\infty} e^{-st} f(t) dt \quad (10)$$

Laplace-Integral.

Satz 1.13. Wenn für eine Funktion f gilt, dass

$$\int_0^{\infty} f(t) dt < \infty,$$

dann gibt es eine Zahl σ mit

$$\int_0^{\infty} e^{-st} f(t) dt \quad \begin{array}{ll} \text{konvergiert für} & s > \sigma \\ \text{divergiert für} & s < \sigma \end{array}$$

Definition 1.14 (Konvergenzbereich). σ heißt Konvergenzabszisse von (10). Der Konvergenzbereich ist gegeben durch

$$\mathbb{K}_f := \mathbb{D}_F.$$

Definition 1.15 (Laplace-Transformation).

$$\mathcal{L} : f \rightarrow F, \quad \mathcal{L}\{f\}(s) = F(s)$$

heißt Laplace-Transformation. Dabei ist f die Originalfunktion und F ist die Bildfunktion.

Notation (üblich unter den Elektrotechnikern):

$$f(t) \circ \bullet F(s) \quad \text{und} \quad F(s) \bullet \circ f(t)$$



Die Linearitätseigenschaft des uneigentlichen Integrals überträgt sich auf die Laplacetransformation. Demnach gilt

$$\mathcal{L}\{\alpha f(t) + \beta g(t)\} = \alpha \mathcal{L}\{f(t)\} + \beta \mathcal{L}\{g(t)\}.$$

Beispiel 29

Es sei $f(t) = 1$.

$$\begin{aligned}\mathcal{L}\{1\}(s) &= \lim_{b \rightarrow \infty} \int_0^b e^{-st} dt = \frac{-1}{s} e^{-st} \Big|_0^b \\ &= \lim_{b \rightarrow \infty} \frac{-1}{s} (e^{-bs} - e^{-0}) = \frac{1}{s}\end{aligned}$$

Es ist also

$$\mathcal{L}\{1\}(s) = \frac{1}{s}, \quad \text{für } s > 0.$$

Beispiel 30

$$\begin{aligned}\mathcal{L}\{e^{at}\}(s) &= \lim_{b \rightarrow \infty} \int_0^b e^{(a-s)t} dt = \frac{1}{s-a}, \quad a < s \\ &= \frac{1}{s-a}, \quad a < s\end{aligned}$$

Beispiel 31 Übung Für $s > |a|$ gilt

$$\mathcal{L}\{\cosh(at)\}(s) = \mathcal{L}\left\{\frac{e^{at} + e^{-at}}{2}\right\}(s) = \frac{1}{2}\mathcal{L}\{e^{at}\}(s) + \frac{1}{2}\mathcal{L}\{e^{-at}\}(s) = \frac{s}{s^2 - a^2}.$$

Satz 1.16 (Differentiationssatz). *Es gilt*

$$\mathcal{L}\{f^{(k)}(t)\}(s) = s^k \mathcal{L}\{f(t)\}(s) - \sum_{l=1}^k s^{k-l} f^{(l-1)}(0).$$

Beispiel 32 Übung

$$u'' - 6u' + 9u = 0, \quad u(0) = 1, \quad u'(0) = 1$$

Es ist

$$\begin{aligned}\mathcal{L}\{u''\}(s) &= s^2 \mathcal{L}\{u\}(s) - 1 - s \\ \mathcal{L}\{u'\}(s) &= s \mathcal{L}\{u\}(s) - 1 \\ \mathcal{L}\{u'' - 6u' + 9u\}(s) &= s^2 \mathcal{L}\{u\}(s) - 1 - s - 6s \mathcal{L}\{u\}(s) + 6 + 9 \mathcal{L}\{u\}(s) \\ &= \mathcal{L}\{u\}(s)(s^2 - 6s + 9) + 5 - s = 0 \\ \Leftrightarrow \mathcal{L}\{u\}(s) &= \frac{s-5}{(s-3)^2} = \frac{s}{(s-3)^2} - 5 \frac{1}{(s-3)^2} \\ &=: \mathcal{L}\{u_1\}(s) - 5 \mathcal{L}\{u_2\}(s)\end{aligned}$$

Es ist

$$u_1 = (1 + 3t) e^{3t} \quad \text{aus Heuser (2009⁶)}$$

und

$$u_2 = t e^{3t} \quad \text{aus Anhang A}$$

woraus folgt, dass

$$u = u_1 - 5 u_2 = (1 + 3t) e^{3t} - 5 t e^{3t} = (1 - 2t) e^{3t}$$



Die Laplace-Transformation ist nur Heuristik, das heißt, dass jede Funktion, die wir als Ergebnis erhalten zunächst nur eine potentielle Lösung ist. Wir müssen jedes Mal prüfen, ob sie der entsprechenden Differentialgleichung tatsächlich genügt.

Allgemein gilt Folgendes. Es sei die ODE zweiter Ordnung

$$u'' + a u' + b u = f, \quad u(0) = u_0, \quad u'(0) = d_0$$

gegeben. Erst mal jeden Term einzeln behandeln, dann zusammenaddieren:

$$\mathcal{L}\{u''\}(s) = s^2 \mathcal{L}\{u\}(s) - s u(0) - u'(0)$$

$$\mathcal{L}\{u'\}(s) = s \mathcal{L}\{u\}(s) - u(0)$$

Damit folgt

$$u'' + a u' + b u = f$$

$$\Leftrightarrow \mathcal{L}\{u'' + a u' + b u\}(s) = \mathcal{L}\{f\}(s)$$

$$\Leftrightarrow \mathcal{L}\{u''\}(s) + a \mathcal{L}\{u'\}(s) + b \mathcal{L}\{u\}(s) = \mathcal{L}\{f\}(s)$$

$$\Leftrightarrow s^2 \mathcal{L}\{u\}(s) - s u_0 - d_0 + a s \mathcal{L}\{u\}(s) - a u_0 + b \mathcal{L}\{u\}(s) = \mathcal{L}\{f\}(s)$$

$$\Leftrightarrow \mathcal{L}\{u\}(s)(s^2 + a s + b) - (s + a) u_0 - d_0 = \mathcal{L}\{f\}(s)$$

$\Leftrightarrow$

$$\mathcal{L}\{u\}(s) = \frac{F(s) + (s + a) u_0 + d_0}{s^2 + a s + b}$$

Formel 8: Laplace-Transformation fuer eine lineare ODE zweiter Ordnung mit konstanten Koeffizienten:

Für

$$u'' + a u' + b u = f, \quad u(0) = u_0, \quad u'(0) = d_0$$

ist

$$\mathcal{L}\{u\}(s) = \frac{F(s) + (s + a) u_0 + d_0}{s^2 + a s + b}$$

Beispiel 33 Übung

$$u'' + 4u' = \cos(2t), \quad u(0) = 0, \quad u'(0) = 1$$

$$\begin{aligned} F(s) &= \lim_{b \rightarrow \infty} \int_0^b e^{-st} \cos(2t) dt \\ &= \lim_{b \rightarrow \infty} \left(\frac{-1}{s} e^{-st} \cos(2t) \Big|_0^b - \int_0^b \frac{2}{s} e^{-st} \sin(2t) dt \right) \\ &= \frac{1}{s} - \frac{2}{s} \lim_{b \rightarrow \infty} \int_0^b e^{-st} \sin(2t) dt \\ &= \frac{1}{s} - \frac{2}{s} \lim_{b \rightarrow \infty} \left(\frac{-1}{s} e^{-st} \sin(2t) \Big|_0^b - \int_0^b \frac{-2}{s} e^{-st} \cos(2t) dt \right) \\ &= \frac{1}{s} - \frac{4}{s^2} \lim_{b \rightarrow \infty} \int_0^b e^{-st} \cos(2t) dt \end{aligned}$$

$$\Leftrightarrow \lim_{b \rightarrow \infty} \int_0^b e^{-st} \cos(2t) dt = \frac{s}{s^2 + 4} = F(s)$$

Mit der Formel auf Seite 30 gilt dann:

$$\begin{aligned} \mathcal{L}\{u\}(s) &= \frac{F(s) + (s+a)u_0 + d_0}{s^2 + as + b} \\ &= \frac{s}{s(s+4)(s^2+4)} + \frac{1}{s(s+4)} \\ &= \frac{1}{20} \frac{1}{s+4} - \frac{1}{20} \frac{s}{s^2+4} + \frac{1}{5} \frac{1}{s^2+4} + \frac{1}{s(s+4)} \end{aligned}$$

Aus der Tabelle lesen wir:

$$\begin{array}{ll} \frac{1}{s+4} & \bullet \longrightarrow e^{-4t} \\ \frac{s}{s^2+4} & \bullet \longrightarrow \cos(2t) \\ \frac{1}{s^2+4} & \bullet \longrightarrow \frac{1}{2} \sin(2t) \\ \frac{1}{s(s+4)} & \bullet \longrightarrow \frac{1}{4} (1 - e^{-4t}) \end{array}$$

Daraus folgt dann

$$u(t) = -\frac{1}{5} e^{-4t} - \frac{1}{20} \cos(2t) + \frac{1}{10} \sin(2t) + \frac{1}{4}$$

lineare ODE erster Ordnung $y'(x) = \alpha(x) y(x) + g(x)$, AWP: $y(0) = y_a$		
homogen: $g(x) = 0$	inhomogen: $g(x) \neq 0$	
Methode: Trennung der Variablen (TV)	Methode: Zunächst die homogene ODE lösen $\rightarrow y_0(x)$ und dann für	
Lösung: $y(x) = C e^{\int \alpha(x) dx}$	konstante Koeff. $\alpha \neq \alpha(x)$	nicht konstante Koeff. $\alpha = \alpha(x)$
	Methode: Partikulärlsg. ³ (PL) $\rightarrow y_p(x)$ Lösung: $y(x) = y_0(0) + y_p(x)$ [auch VK mgl.]	Methode: Variation der Konst. $\rightarrow y(x) = C(x) e^{\int \alpha(x) dx}$ Lösung: in $y(x)$ $C(x) = \int g(x) e^{-\int \alpha(x) dx} dx$
Bei AWP's Konstante so bestimmen, dass $y(0) = y_a$ erfüllt ist		
nicht lineare ODEs substituieren		
homogen: $g(x) = 0$	inhomogen: $g(x) \neq 0$	
Je nach Form geeignete Subst. wählen Je nach Form der neuen ODE Methode wählen Rücksubstitution	hom. ODE lösen $\rightarrow y_0(x)$ dann je nach Form inhom. ODE lösen $\rightarrow y(x)$	
lineare ODEs höherer Ordnung mit konstanten Koeffizienten		
homogen: $g(x) = 0$	inhomogen: $g(x) \neq 0$	
Nullstellen des charakt. Polynoms (CP) λ_1 Nullst. mit Vfh 1, λ_k Nullst. mit Vfh s+1 Lösung: $y(x) = C_1 e^{\lambda_1 x} + \dots + C_{k_{s-1}} x^{s-1} e^{\lambda_k x} + C_{k_s} x^s e^{\lambda_k x}$	hom. ODE lösen $\rightarrow y_0(x)$ dann inhom. mit PL $\rightarrow y_p(x)$ Lösung: $y(x) = y_0(x) + y_p(x)$	
Oder: Bei AWP auch Laplace-Transformation möglich		
AWP und RWP bei ODEs zweiter Ordnung		
AWP: $y''(x) + \alpha y'(x) + \beta y(x) = g(x)$ auf $\mathbb{R}_0^+$ mit $y(0) = y_a, y'(0) = y_b$ RWP: $y''(x) + \alpha y'(x) + \beta y(x) = g(x)$ auf $[a,b]$ mit $y(a) = y_a, y(b) = y_b$		
Unterscheidung AWP-RWP zur Best. von C, Lösungsmethode wie oben		
Systeme von ODEs erster Ordnung mit konstanten Koeffizienten als AWP $a u' + b v = g, u(0) = u_0$ und $c v' + d u = f, v(0) = v_0$		
homogen: $g(x) = 0$	inhomogen: $g(x) \neq 0$	
Eigenwerte λ_i der Matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ bestimmen λ_1 Nullst. mit Vfh 1, λ_k Nullst. mit Vfh s+1 Lösung: $y_0(x) = C_1 e^{\lambda_1 x} + \dots + C_{k_{s-1}} x^{s-1} e^{\lambda_k x} + C_{k_s} x^s e^{\lambda_k x}$	PL mit Ansatz $y_p(x) = y_{p,g}(x) + y_{p,f}(x)$	
Bei $\lambda_j = a_j + i b_j \in \mathbb{C}$: $y_0(x) = \dots e^{a_j x} (C_{j,1} \cos(b_j x) + C_{j,2} \sin(b_j x)) \dots$	Lösung: $y(x) = y_0(x) + y_p(x)$	

Abbildung 4: Übersicht über die hier behandelten Lösungsmethoden von ODEs

richtig
rechnen
reicht

2.1 Kombinatorik

Aus Elementen einer Menge lassen sich häufig auf eine bestimmte Weise neue Mengen zusammenstellen. Die Art und Weise einer solchen Zusammenstellung führt auf den Begriff *Permutation* (Anordnung), *Kombination* (Auswahl) und *Variation*. Beim Begriff der Variation werden Anordnung und Auswahl vereinigt, indem bei der Auswahl von Elementen auf deren Reihenfolge Rücksicht genommen wird. Die Grundaufgabe der Kombinatorik besteht darin, die Anzahl der Auswahl- oder Anordnungsmöglichkeiten zu ermitteln.

Definition 2.1 (Fakultät). Die Abbildung $! : \mathbb{N} \rightarrow \mathbb{N}$ mit

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

$$0! := 1$$

heißt Fakultät.

Definition 2.2 (Binomialkoeffizient). Die Binomialkoeffizienten sind definiert als

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

Wir sagen “ n über k ”.

Definition 2.3 (Permutation). Eine Permutation P_n nennt man eine Anordnung von n Elementen in einer bestimmten Reihenfolge.

Satz 2.4 (Anordnungsmöglichkeiten).

1. Anzahl der Permutationen ohne Wiederholung

n verschiedene Dinge lassen sich unter Berücksichtigung der Reihenfolge auf

$$P_n := n! = n \cdot (n-1) \cdots 1$$

verschiedene Arten anordnen. Das sind gerade die Anzahl der Permutationen von n verschiedenen Elementen.

2. Anzahl der Permutationen mit Wiederholung

n Dinge, von denen k gleich sind lassen sich auf

$$P_n^{(k)} := \frac{n!}{k!}$$

verschiedene Arten anordnen.

Beispiel 34

1. In einem Hörsaal wurde eine Reihe mit 13 Sitzplätzen von genau 13 Studenten besetzt. Es gibt $13! = 6'227'020'800$ Möglichkeiten für die Sitzordnung.
2. Eine Reihe von 13 Sitzplätzen im Hörsaal wird von 13 Studenten mit ihren Taschen belegt. Unter den 13 Taschen befinden sich 4 gleiche. Dann gibt es $\frac{13!}{4!} = 259'459'200$ Möglichkeiten zur Anordnung der Taschen.

Satz 2.5 (Verallgemeinerung). Für die Anzahl $P_n^{(k_1, \dots, k_m)}$ der Permutationen von n Elementen, eingeteilt in m Gruppen mit jeweils $k_1, \dots, k_m$ gleichen Elementen ($k_1 + \dots + k_m = n$), gilt

$$P_n^{(k_1, \dots, k_m)} = \frac{n!}{k_1! \cdots k_m!}.$$

Beispiel 35

Aus den fünf Ziffern 4, 4, 5, 5, 5 können $P_5^{(2,3)} = \frac{5!}{2!3!} = 10$ verschiedene fünfstelligen Zahlen gebildet werden.

Satz 2.6 (Auswahlmöglichkeiten ohne Berücksichtigung der Reihenfolge).

Aus n verschiedenen Dingen werden k Stück ohne Berücksichtigung der Reihenfolge ausgewählt. Dann beträgt die Anzahl der verschiedenen Auswahlmöglichkeiten:

1. beim Ziehen ohne Zurücklegen

$$C_n^{(k)} := \binom{n}{k} = \frac{n!}{k!(n-k)!}, \quad k = 1, \dots, n$$

2. beim Ziehen mit Zurücklegen

$$C_n^{(k)} := \binom{n+k-1}{k} = \frac{(n+k-1)!}{k!(n-k)!}, \quad k \in \mathbb{N}$$

Beispiel 36

1. Es gibt $\binom{30}{4} = 27'405$ Möglichkeiten, aus 30 Teilnehmern einer Wahlversammlung einen 4-köpfigen Wahlvorstand ohne Zuordnung der Funktionen zusammenzustellen.
2. Mit k Würfeln sind $C_6^{(k)} = \binom{k+6-1}{k}$ verschiedene Würfe möglich. Für zwei Würfel gilt demzufolge

$$C_6^{(2)} = \binom{7}{2} = \frac{7!}{2!5!} = 21.$$

Definition 2.7 (Variation). *Variation nennt man eine Auswahl von k Elementen aus n verschiedenen Elementen unter Beachtung der Reihenfolge. Das bedeutet: Variationen sind Kombinationen mit Berücksichtigung der Reihenfolge. Deshalb ist auch bei den Variationen zwischen Variation mit und ohne Wiederholung zu unterscheiden.*

Satz 2.8 (Auswahlmöglichkeiten mit Berücksichtigung der Reihenfolge).

Aus n verschiedenen Dingen werden k Stück mit Berücksichtigung der Reihenfolge ausgewählt. Dann beträgt die Anzahl der verschiedenen Auswahlmöglichkeiten:

1. beim Ziehen ohne Zurücklegen

$$V_n^{(k)} := k! \binom{n}{k} = \frac{n!}{(n-k)!} = n \cdot (n-1) \cdots (n-k+1), \quad k \leq n$$

2. beim Ziehen mit Zurücklegen

$$V_n^{(k)} := n^k, \quad k \in \mathbb{N}$$

Beispiel 37

1. Wieviele Möglichkeiten gibt es, in einer Wahlversammlung mit 30 Teilnehmern einen 4-köpfigen Wahlvorstand, bestehend aus dem Vorsitzenden, seinem Stellvertreter und dem 1. und 2. Wahlhelfer zusammenzusetzen? Die Antwort lautet

$$\binom{30}{4} 4! = 657'720$$

2. (a) Beim Fußball-Toto sind für 12 Spiele 3^{12} verschiedene Tipps möglich.
 (b) Mit der digitalen Einheit Byte, die aus 8 Bits besteht, können $2^8 = 256$ verschiedene Zeichen dargestellt werden.

Definition 2.9 (Urnenmodelle). *Ein Urnenmodell ist ein Gedankenexperiment, das in der Wahrscheinlichkeitstheorie verwendet wird, um verschiedene Zufallsexperimente auf einheitliche und anschauliche Weise zu modellieren. Dazu wird ein fiktives Gefäß, eine Urne, mit einer Anzahl Kugeln gefüllt, welche anschließend zufällig gezogen werden.*

“ k aus n ”	Anzahl der Möglichkeiten	
	ohne Wiederholungen ($k \leq n$)	mit Wiederholungen ($k \leq n$)
Permutationen	$P_n = n!, (n = k)$	$P_n^{(k)} = \frac{n!}{k!}$
Kombinationen	$C_n^{(k)} = \binom{n}{k}$	$C_n^{(k)} = \binom{n+k-1}{k}$
Variationen	$V_n^{(k)} = k! \binom{n}{k}$	$V_n^{(k)} = n^k$

Abbildung 5: Zusammenstellung der Formeln der Kombinatorik

Damit ist gemeint, dass bei jedem Zug alle in der Urne befindlichen Kugeln die gleiche Wahrscheinlichkeit haben, ausgewählt zu werden. Dadurch kann die Bestimmung interessierender Wahrscheinlichkeiten auf die Lösung kombinatorischer Abzählprobleme zurückgeführt werden. ⁴



Beim Urnenmodell gibt es also vier Situationen: Man zieht und behält oder man zieht und legt wieder zurück. Beide Situationen teilen sich nochmal in die Situation, dass die Reihenfolge der Ziehung eine Rolle spielt oder nicht. Alle reale Situationen versucht man nun derart zu abstrahieren, dass sie auf eine der vier Situationen des Urnenmodells passen. Dann weiss man welche Formel angewendet werden muss.

2.2 Grundlagen der Wahrscheinlichkeitstheorie

Definition 2.10 (Laplace-Wahrscheinlichkeit). *Gibt es bei einem Zufallsexperiment m mögliche Fälle, die alle gleichermaßen möglich sind, sich aber gegenseitig ausschließen, und gibt es g Fälle, bei denen das Ereignis A zutrifft, so definieren wir die Wahrscheinlichkeit für das Ereignis A als*

$$P(A) = \frac{\text{günstige Fälle}}{\text{mögliche Fälle}} = \frac{g}{m}.$$

Da stets $0 \leq g \leq m$ gilt, folgt der

Satz 2.11. *Für ein beliebiges Ereignis A gilt*

$$0 \leq P(A) \leq 1.$$

Beispiel 38 Zahlenlotto Beim Zahlenlotto werden 6 aus 45 gezogen.

Frage: Wie groß ist die Wahrscheinlichkeit für A: “6 Richtige”?

⁴aus Wikipedia

Für das Urnenmodell k aus n gilt, $n = 45$ und $k = 6$. Wir ziehen Zahlen aus der Urne ohne Wdh und betrachten die Auswahl mit Rf. Dann erhalten wir die Anzahl aller möglichen Fälle

$$m = V_{45}^{(6)} = 6! \binom{45}{6}$$

und die Anzahl günstiger Fälle ist dann genau ein Zahlensortiment, dass aber in $6!$ verschiedenen Anordnungen auftreten kann, also gilt

$$g = 6!.$$

Daraus ergibt sich die Wahrscheinlichkeit

$$P(A) = \frac{6!}{6! \binom{45}{6}} \approx 1.2277e - 07.$$

Dieses Beispiel könnten wir aber auch getrost ohne Berücksichtigung der Reihenfolge ansetzen, denn dann erhielten wir als Anzahl möglicher Fälle

$$m = C_{45}^{(6)} = \binom{45}{6}$$

und als Anzahl günstiger Fälle

$$g = 1,$$

da ja nun die Reihenfolge keine Rolle spielt betrachten wir auch nicht verschiedene Permutationen als verschiedene Fälle. Wir erhalten dann

$$P(A) = \frac{1}{\binom{45}{6}} \approx 1.2277e - 07.$$

Beispiel 39 Münzexperiment	Eine Laplace-Münze wird 6 Mal geworfen. Für das Urnenmodell halten wir fest: $n = 2$, $k = 6$, mit mit Wdh, mit Rf. Das führt insgesamt auf
----------------------------	---

$$m = n^k = 2^6 = 64$$

mögliche Fälle.

Frage: Wie groß ist die Wahrscheinlichkeit folgender Ereignisse:

(a) A: "Der dritte Wurf zeigt Zahl."

Bei jedem einzelnen Wurf gibt es zwei mögliche und einen günstigen Fall. Das führt auf

$$P(A) = \frac{1}{2}.$$

(b) B: "Genau drei Mal erscheint Zahl."

Das Ereignis beinhaltet einen Wurf mit (zzzkkk) und allen möglichen Permutationen davon, also $g = 6!$ günstige Fälle. Das führt dann auf

$$P(B) = \frac{6!}{2^6}.$$

(c) C: "Mindestens viermal erscheint Zahl."

Ergebnis	Anzahl Permutationen
(zzzzkk)	$6!/4!/2!=15$
(zzzzzk)	$6!/5!=6$
(zzzzzz)	1

Insgesamt erhalten wir $g = 22$ günstige Fälle und damit die Wahrscheinlichkeit

$$P(C) = \frac{22}{64} = \frac{11}{32}.$$

Definition 2.12. Die möglichen Ausgänge des Experiments werden *Ergebnisse* genannt, und die Menge aller möglichen Ergebnisse nennt man den *Wahrscheinlichkeits- oder Ergebnisraum* Ω . Die Ergebnisse ω sind Elemente von Ω . Als ein *Ereignis* bezeichnen wir eine beliebige Ergebniskonstellaton, bzw. eine beliebige Teilmenge $A \subseteq \Omega$.

Beispiel 40

Beim einmaligen Würfeln ist

$$\Omega = \{1, 2, \underbrace{3}_{\text{Ergebnis}}, 4, 5, 6\}$$

der Wk-raum. Ein Ereignis ist zum Beispiel "würfle eine gerade Zahl", d.h. $A = \{2, 4, 6\}$.

Beispiel 41

Eine Münze werde drei Mal geworfen. Die Ergebnisse sind alle möglichen Wurfsequenzen, welche zusammengefasst den Wk-Raum

$$\Omega = \{kkk, kkz, kzk, kzz, zkk, z kz, z zk, zzz\}$$

bilden. Ein Ereignis ist z.B. "mindestens zweimal Kopf", d.h. $A = \{kkk, kkz, kzk, zkk\} \subseteq \Omega$.

Definition 2.13 (Spezielle Ereignisse).

- Die Ereignisse $\{\omega_1\}, \{\omega_2\}, \dots \in \Omega$ enthalten genau ein Ergebnis und werden *Grundereignisse* genannt. Ergebnisse sind also Grundereignisse, aus denen alle anderen Ereignisse zusammengesetzt sind.
- Das Komplement $\bar{A}$ eines Ereignisses A , d.h. $\bar{A} = \Omega \setminus A$ heißt *Gegenereignis*. Es gilt $\bar{A} \cup A = \Omega$ und $\bar{A} \cap A = \emptyset$.

- Ereignisse A und B mit $A \cap B = \emptyset$ heißen *unvereinbar* oder auch *disjunkt*.
- Als *sicheres Ereignis* gilt $A = \Omega$, und als *unmögliches Ereignis* ist $A = \emptyset$. Es ist $P(\Omega) = 1$ und $P(\emptyset) = 0$.



Es seien A und B Ereignisse, dann bedeutet $A \cap B$, dass A und B eintreten und $A \cup B$, dass A oder B eintreten, wobei betont sein soll, dass es sich um ein UND-eingeschlossenes ODER im Sinne der Aussagenlogik ahndelt.

Beispiel 42

Beim einmaligen Würfeln ist $\Omega = \{1, 2, 3, 4, 5, 6\}$. Zu C : "Augenzahl durch drei teilbar" und B : "Augenzahl durch zwei teilbar" ist $C \cap B = \{6\}$, bzw. "Augenzahl durch drei teilbar UND Augenzahl durch zwei teilbar" und $C \cup B = \{2, 3, 4, 6\}$, bzw. "Augenzahl durch drei teilbar ODER Augenzahl durch zwei teilbar":

Den Wk-Begriff der Laplace-Wk, den wir in Definition 2.10 kennengelernt haben ist intuitiv und bietet einen leichtfüßigen Zugang in dieses Thema, da er Situationen behandelt, die uns aus dem Alltg bekannt sind. Sobald wir aber die Situation "alle Möglichkeiten treten gleichwahrscheinlich auf" verlassen, wie es etwa bei einem gezinkten Würfel der Fall ist, brauchen wir einen übergeordneten Wk-Begriff. Eine abstraktere, allgemeingültige Definition erhalten wir durch den axiomatischen Ansatz in

Definition 2.14 (Wahrscheinlichkeit von Kolmogorow). *Die Wahrscheinlichkeit auf einem Ergebnisraum Ω ist eine Abbildung*

$$P : \{A \mid A \subseteq \Omega\} \rightarrow [0, 1]$$

mit folgenden sogenannten Axiome von Kolmogorow erfüllen:

[Ko 1] *Für jedes Element A gilt*

$$0 \leq P(A) \leq 1.$$

[Ko 2] *Das sichere Ereignis hat die Wahrscheinlichkeit 1:*

$$P(\Omega) = 1$$

[Ko 3] *Sind A und B unvereinbare Ereignisse, d.h. $A \cap B = \emptyset$, so gilt*

$$P(A \cup B) = P(A) + P(B).$$

Beispiel 43

(1) Die Wahrscheinlichkeit beim Würfeln ist eine Laplace-Wahrscheinlichkeit, denn alle möglichen Fälle treten mit gleicher Wk auf. Obendrein genügt sie den Axiomen von Kolmogorow.

- (2) Anders verhält es sich beim gezinkten Würfel: Es sei etwa $P(1) = \frac{1}{2}$ und $P(i) = \frac{1}{10}$ ($i = 2, \dots, 6$). Die Wk, eine 1 zu würfeln ist größer, als die eine 2 zu würfeln. Um die Nicht Laplace-Wk hier auch rechnerisch einsehen zu können, benötigen wir zunächst noch ein paar Rechenregeln.

Satz 2.15.

1. *Wahrscheinlichkeit des komplementären Ereignisses: Für ein beliebiges Ereignis $A \subseteq \Omega$ gilt*

$$P(\bar{A}) = 1 - P(A).$$

2. *Wahrscheinlichkeit von Teilmengen: Für zwei Ereignisse A, B mit $A \subseteq B$ gilt*

$$P(A) \leq P(B).$$

3. *Additionssatz für beliebige Ereignisse: Für beliebige Ereignisse $A, B \subseteq \Omega$ gilt*

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

Beweis Satz 2.15:

- (1) *Es ist $\bar{A} \cap A = \emptyset$ und $\bar{A} \cup A = \Omega$. Damit gilt*

$$P(\bar{A} \cup A) = P(\Omega) \stackrel{(Ko_2)}{=} 1.$$

Andererseits gilt auch

$$P(\bar{A} \cup A) \stackrel{(Ko_3)}{=} P(\bar{A}) + P(A).$$

Zusammen führt das auf die Behauptung

$$P(\bar{A}) + P(A) = 1 \quad \Leftrightarrow \quad P(A) = 1 - P(\bar{A}).$$

- (2) *Da $A \subset B$ gilt für $\bar{A} := B \setminus A$, dass $A \cup \bar{A} = B$ und $A \cap \bar{A} = \emptyset$ ist. Mit (1) gilt:*

$$P(A \cup \bar{A}) = P(A) + \underbrace{P(\bar{A})}_{\geq 0} = P(B) \quad \Rightarrow \quad P(A) \leq P(B)$$

- (3) *Wir stellen zunächst für Ereignisse A und B mit $A \subset B$ folgende Überlegung an:*

$$P(A) = P((A \setminus B) \cup B) \stackrel{(Ko_3)}{=} P(A \setminus B) + P(B)$$

⇒

$$P(A \setminus B) = P(A) - P(A \cap B)$$

Das können wir nun verwenden für

$$\begin{aligned} P(A \cup B) &= P((A \setminus (A \cap B)) \cup (A \cap B) \cup (B \setminus (A \cap B))) \\ &= P(A \setminus (A \cap B)) + P(A \cap B) + P(B \setminus (A \cap B)) \\ &= P(A) - P(A \cap B) + P(A \cap B) + P(B) - P(A \cap B) \\ &= P(A) + P(B) - P(A \cap B) \end{aligned}$$

□

Beispiel 44 zu Punkt (1) in Satz 2.15 Manchmal ist es einfacher die Wk des Gegenereignisses zu berechnen. Die Differenz zu 1 liefert dann den gesuchten Wert. Es sei nach der Wk des Ereignisses A : "der Wert ist mindestens 3" im Würfelexperiment gefragt. Das Gegenereignis $\bar{A}$ beinhaltet die Negation von A , also $\neg$ ("der Wert ist mindestens 3") = "der Wert ist nicht mindestens 3" oder anders formuliert "der Wert ist 1 oder 2". Es ist also

$$\bar{A} : \text{"der Wert ist genau 1"} \vee \text{"der Wert ist genau 2"} =: \bar{A}_1 \cup \bar{A}_2$$

und dann gilt

$$P(A) = 1 - P(\bar{A}) = 1 - P(\bar{A}_1) - P(\bar{A}_2) = 1 - \frac{1}{6} - \frac{1}{6} = \frac{2}{3}.$$

Die Idee besteht darin, die Ereignisse so zu formulieren, dass man leicht berechenbare kleine Terme hat deren Wahrscheinlichkeitswerte dann einfach nur noch verrechnet werden müssen.

Wir haben bisher Rechenregeln für die Wk von Vereinigung von Ereignissen und Differenzen von Ereignissen. Was uns noch fehlt ist eine Regel für den Schnitt, also die logische UND-Verknüpfung von Ereignissen. Das funktioniert an dieser Stelle nicht allgemeingültig. Wir führen dazu folgende Definition ein:

Definition 2.16 ((un-) abhängige Ereignisse). Zwei Ereignisse A und B heißen *unabhängig*, wenn

$$P(A \cap B) = P(A) \cdot P(B)$$

gilt. Sonst heißen Sie abhängig.

Es stellt sich hier die Frage, ob sich zwei Ereignisse gegenseitig beeinflussen, d.h. wirkt sich das Eintreten des einen Ereignisses auf das Eintreten des anderen aus oder nicht?

Beispiel 45 Ein Würfel wird ein Mal geworfen.

- (a) Wir betrachten die Ereignisse A : "Die Zahl ist eine 1 oder eine 2", sowie B : "Die Zahl ist gerade". Die Ereignisse A und B sind unabhängig. Es gilt hier

$$P(A \cap B) = P(\{2\}) = \frac{1}{6}$$

und

$$P(A) \cdot P(B) = \frac{2}{6} \cdot \frac{3}{6} = \frac{1}{6}.$$

Wir sehen hier aus rein rechnerischen Gründen, dass die Bedingung für Unabhängigkeit erfüllt ist. Aber auch anschaulich ist klar, dass das Eintreten von B in keinem Zusammenhang vom Eintreten von A steht.

- (b) Anders verhält es sich in diesem Fall: Wir betrachten die Ereignisse B wie vorher und C : "Die Zahl ist ungerade". Die Ereignisse B und C sind nicht unabhängig. Wenn B eintritt ist klar, dass C nicht eintritt und umgekehrt. Außerdem ist der Schnitt leer und damit

$$P(C \cap B) = 0,$$

was bedeutet, dass per Definition die beiden Ereignisse abhängig sind.

Satz 2.17. *Sind die Ereignisse A und B unabhängig, so sind auch die Ereignisse $\bar{A}$ und B , die Ereignisse A und $\bar{B}$ sowie die Ereignisse $\bar{A}$ und $\bar{B}$ unabhängig.*

Beweis Satz 2.17:

Aus

$$\bar{A} \cap B = B \setminus A = B \setminus (A \cap B)$$

folgt direkt

$$\begin{aligned} P(\bar{A} \cap B) &= P(B \setminus (A \cap B)) \\ &= P(B) - P(A \cap B) \\ &= P(B) - P(A) \cdot P(B) \\ &= P(B) (1 - P(A)) \\ &= P(B) \cdot P(\bar{A}). \end{aligned}$$

Der Fall $\bar{B} \cap A$ wird in analoger Weise bewiesen.

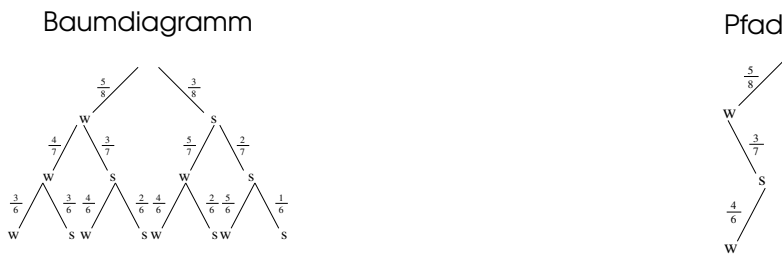
Dann noch den letzten Fall. Dazu sehen wir ein, dass $\bar{A} \cap \bar{B} = \Omega \setminus (A \cup B)$ gilt. Im Zweifel malen Sie sich ein Venn-Diagramm dazu. Damit gilt jedenfalls

$$\begin{aligned} P(\bar{A} \cap \bar{B}) &= P(\Omega \setminus (A \cup B)) \\ &= P(\Omega) - P(A \cup B) \\ &= 1 - P(A) - P(B) + P(A \cap B) \\ &= 1 - P(A) - P(B) + P(A) \cdot P(B) \\ &= (1 - P(B)) - P(A)(1 - P(B)) \\ &= P(\bar{B}) \cdot P(\bar{A}). \end{aligned}$$

□

Wahrscheinlichkeiten, die über logische UND-Verknüpfungen definiert sind nennt man auch *mehrstufige Versuche*. Mehrstufige Versuche haben wir schon kennengelernt bei der Betrachtung des Würfellexperiments, wenn ein Würfel mehrfach geworfen wird. Auch ein einmaliger Wurf mehrerer Würfel, kann als mehrstufiger Prozess verstanden werden, da es keine Rolle spielt ob Würfel gleichzeitig oder hintereinander gewürfelt werden. Jeder Wurf bringt eine Wahrscheinlichkeit zu Tage und stellt einen Sachverhalt dar. Die einzelnen Sachverhalte werden multipliziert. Auf diese Weise ergibt sich das sogenannte *Baumdiagramm* eines mehrstufigen Versuchs. Es erlaubt auch bei nicht Laplace-Wk ein einfaches Aufmultiplizieren der Wahrscheinlichkeiten der einzelnen Stufen. Wir betrachten am besten ein Beispiel dazu:

Beispiel 46 Baumdiagramm In einer Urne befinden sich 5 weiße und 3 schwarze Kugeln. Wir ziehen 3 Kugeln hintereinander ohne die gezogenen zurückzulegen. Gesucht sind die Wahrscheinlichkeiten aller acht möglicher Ereignisse in $\Omega = \{www, wws, wsw, wss, sww, sws, ssw, sss\}$. Alle möglichen Ergebnisse veranschaulichen wir nun in einem Diagramm folgender Art:



Wir interessieren uns für das Ereignis $A = \{wsw\}$. Die Wahrscheinlichkeiten eines Pfades werden multipliziert:

$$P(\{wsw\}) = \frac{5}{8} \cdot \frac{3}{7} \cdot \frac{4}{6} = \frac{5}{28} \approx 0.178571429$$

Die Wahrscheinlichkeiten mehrerer Pfade werden addiert. Sei $B =$ "mindestens zwei schwarze Kugeln":

$$\begin{aligned} P(\{(wss, sws, ssw, sss)\}) &= P(\{wss\}) + P(\{sws\}) + P(\{ssw\}) + P(\{sss\}) \\ &= \frac{5}{8} \cdot \frac{3}{7} \cdot \frac{2}{6} + \frac{3}{8} \cdot \frac{5}{7} \cdot \frac{2}{6} + \frac{3}{8} \cdot \frac{2}{7} \cdot \frac{5}{6} + \frac{3}{8} \cdot \frac{2}{7} \cdot \frac{1}{6} \\ &= \frac{30 + 30 + 30 + 6}{336} = \frac{96}{336} = \frac{2}{7} \approx 0.285714286 \end{aligned}$$

Das Ereignis $C =$ "höchstens zwei weiße Kugeln" ist ja gleich zu $\Omega \setminus \bar{C}$, wobei $\bar{C} = \{www\}$. Die Wahrscheinlichkeit von C erhalten wir über die Wahrscheinlichkeit des Komplementes $\bar{C}$:

$$P(C) = 1 - P(\bar{C}) = 1 - P(\{www\}) = 1 - \frac{5}{8} \cdot \frac{4}{7} \cdot \frac{3}{6} = \frac{336 - 60}{336} \approx 0.821428571$$

Allgemein können wir für den Umgang mit derartigen Ereignisbäumen folgende Regeln aufstellen:

Formel 9: Baumdiagramm:

- Entlang eines Pfades werden die Stufenwahrscheinlichkeiten multipliziert, dies ergibt die Pfadwahrscheinlichkeit.
- Die Pfadwahrscheinlichkeiten verschiedener Pfade, die zu einem Ereignis gehören, werden addiert.



Der Vorteil eines Baumdiagramms ist, dass auf übersichtliche Art Wahrscheinlichkeiten berechnet werden können auch wenn es sich nicht um eine Laplace-Wk handelt, so wie es im Beispiel 46 der Fall ist. Eine schwarze Kugel zu ziehen ist wahrscheinlicher als eine weisse Kugel zu ziehen!

Wir sehen im oben abgebildeten Baumdiagramm, dass die Wk, eine schwarze Kugel zu ziehen variiert, je nachdem wieviele Kugeln, schwarze oder weisse, bereits gezogen wurden. Wie sich die Wk also verhält wird durch andere Umstände bedingt. Als bedingte Wahrscheinlichkeiten bezeichnen wir Wahrscheinlichkeiten von Ereignissen, die sich auf bereits eingetretene Ereignisse beziehen. Betrachten wir dazu einfach mal ein Beispiel:

Beispiel 47

Die folgende Tabelle beschreibt die Verbreitung der Rot-Grün-Blindheit bei $N = 965$ Personen.

	Männer (M)	Frauen (F)	total
normal (N)	485	436	921
rgb (B)	38	6	44
	523	442	965

- (a) Die Frage nach dem Anteil der Männer die Rot-Grün-blind sind kann auch aufgefasst werden als die Wahrscheinlichkeit, dass ein Mann Rot-Grün-blind ist:

$$\frac{P(B \cap M)}{P(M)} = \frac{38}{523} \approx 0.073$$

- (b) Welcher Anteil der Frauen ist gesund?

$$\frac{P(N \cap F)}{P(F)} = \frac{436}{442} \approx 0.986$$

- (c) Welcher Anteil der Rot-Grün-Blinden ist weiblich?

$$\frac{P(F \cap B)}{P(B)} = \frac{6}{44}$$

allgemein: Welcher Anteil von bla ist gaga?

$$\frac{P(\text{bla} \cap \text{gaga})}{P(\text{bla})}$$

Definition 2.18 (bedingte Wahrscheinlichkeit). Seien A und B zwei Ereignisse mit $P(B) > 0$. Dann ist die bedingte Wahrscheinlichkeit $P(A|B)$, d.h. die Wahrscheinlichkeit von A unter der Bedingung B , gegeben durch

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

Falls die Wahrscheinlichkeiten von A und B positiv sind, also nicht Null, können wir $P(A|B)$ durch $P(B|A)$ ausdrücken und umgekehrt: Es gilt nämlich

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

Daraus erhalten wir dann

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}. \quad (11)$$

Satz 2.19 (totale Wahrscheinlichkeit). Seien A, B Ereignisse mit $0 < P(A) < 1$. Dann lässt sich $P(B)$ ausdrücken als

$$P(B) = P(B|A) \cdot P(A) + P(B|\bar{A}) \cdot P(\bar{A}) \quad (12)$$

und wird mit totaler Wk des Ereignisses B bezeichnet.

Die totale Wk sammelt alle Wken zu allen Bedingungen. Quasi.

Beweis Satz 2.19:

$$\begin{aligned} P(B|A) \cdot P(A) + P(B|\bar{A}) \cdot P(\bar{A}) &= \frac{P(B \cap A)}{P(A)} P(A) + \frac{P(B \cap \bar{A})}{P(\bar{A})} P(\bar{A}) \\ &= P(B \cap A) + P(B \cap \bar{A}) \\ &= P((B \cap A) \cup (B \cap \bar{A})) \\ &= P(B \cap (A \cup \bar{A})) \\ &= P(B \cap \Omega) \\ &= P(B) \end{aligned}$$

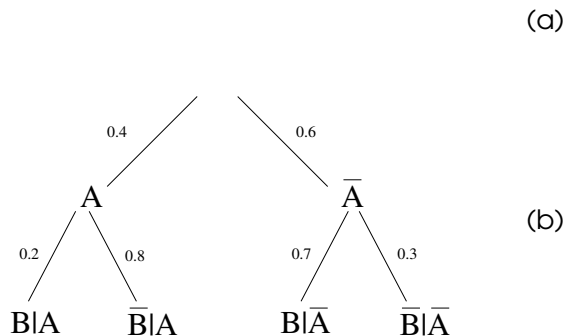
□

Die Gleichungen (11) und (12) führen direkt zum

Satz 2.20 (Satz von Bayes). *Es gilt die Formel*

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B|A) \cdot P(A) + P(B|\bar{A}) \cdot P(\bar{A})}.$$

Beispiel 48 Wir betrachten mal folgenden Ereignisbaum:



$$\begin{aligned} P(B) &= P(B|A) \cdot P(A) + P(B|\bar{A}) \cdot P(\bar{A}) \\ &= 0.2 \cdot 0.4 + 0.7 \cdot 0.6 = 0.5 \end{aligned}$$

$$\begin{aligned} P(A|B) &= \frac{P(B|A) \cdot P(A)}{P(B|A) \cdot P(A) + P(B|\bar{A}) \cdot P(\bar{A})} \\ &= \frac{0.2 \cdot 0.4}{0.2 \cdot 0.4 + 0.7 \cdot 0.6} = 0.16 \end{aligned}$$

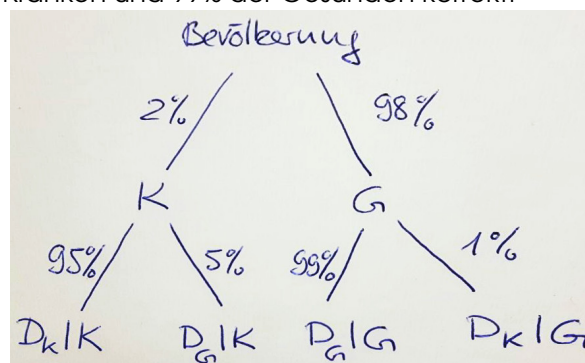
Es gibt eine enge Beziehung zwischen den Begriffen "Unabhängigkeit" und "bedingt Wahrscheinlichkeit": Falls A und B unabhängige Ereignisse sind mit $P(B) > 0$, so vereinfacht sich die Formel für $P(A|B)$ zu



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

Das liegt daran, dass bei einer Unabhängigkeit der Ereignisse das Ereignis B keinen Einfluss auf die Wahrscheinlichkeit vom Ereignis A hat.

Beispiel 49 Von einer Krankheit weiß man, dass 2% der Bevölkerung davon befallen sind. Ein Test diagnostiziert 95% der Kranken und 99% der Gesunden korrekt.



(a) Wie wahrscheinlich zeigt der Test Diagnose "krank"?

Gefragt ist hier nach der totalen W_K

$$\begin{aligned} P(D_K) &= P(D_K|K) \cdot P(K) + P(D_K|G) \cdot P(G) \\ &= 95\% \cdot 2\% + 1\% \cdot 98\% = 2,88\% \end{aligned}$$

- (b) Wie wahrscheinlich ist man tatsächlich krank, wenn der Test "krank" diagnostiziert? Hier behelfen wir uns mit dem Satz von Bayes.

$$P(K|D_K) = \frac{P(D_K|K) \cdot P(K)}{P(D_K)} = \frac{95\% \cdot 2\%}{2,88\%} \approx 65,97\%$$

- (c) Wie wahrscheinlich ist man tatsächlich gesund, wenn der Test "gesund" diagnostiziert?

$$P(G|D_G) = \frac{P(D_G|G) \cdot P(G)}{P(D_G)} = \frac{98\% \cdot 99\%}{2\% \cdot 5\% + 98\% \cdot 99\%} \approx 99,9\%$$

Beispiel 50

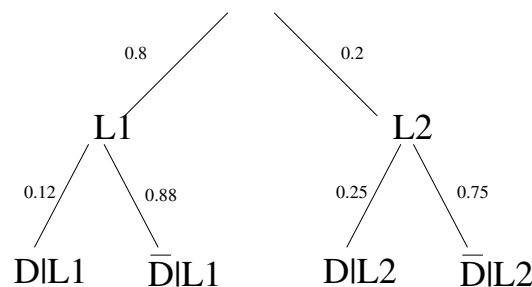
Eine Fabrik bezieht 80% eines Bestandteils vom Lieferanten L_1 und 20% vom Lieferanten L_2 . Es ist bekannt, dass 12% der Teile von L_1 und 25% der Teile von L_2 defekt sind. Bei der Lieferung werden die Teile durchmischt.

- (a) Mit welcher Wahrscheinlichkeit ist ein zufällig bestimmtes Teil verwendbar?

$$P(\bar{D}) = P(\bar{D}|L_1) \cdot P(L_1) + P(\bar{D}|L_2) \cdot P(L_2) = 0.8 \cdot 0.88 + 0.2 \cdot 0.75 = 0.854$$

- (b) Ein zufällig bestimmtes Teil ist defekt. Mit welcher Wahrscheinlichkeit stammt es von L_2 ?

$$P(L_2|D) = \frac{P(D|L_2) \cdot P(L_2)}{P(D|L_2) \cdot P(L_2) + P(D|L_1) \cdot P(L_1)} = \frac{0.2 \cdot 0.25}{0.2 \cdot 0.25 + 0.12 \cdot 0.8} \approx 0.342$$



Permutationen sind Variationen der Anordnung von Elementen. P_n beschreibt die Anzahl möglicher Anordnungen von n Elementen.

n verschiedenen Elemente: $P_n = n!$
 Unter den n Elementen gibt es I Gruppen mit
 jeweils $k_i, i = 1, \dots, I$ gleichen Elementen: $P_n^{(k_1, \dots, k_I)} = \frac{n!}{\prod_{i=1}^I k_i!}$

Kombinationen und **Variationen** beschreiben die Anzahl Möglichkeiten, die sich aus **Urnenmodellen** " k aus n " ergeben

	ohne Wiederholungen	mit Wiederholungen
ohne Reihenfolge Kombinationen	$C_n^{(k)} = \binom{n}{k}$	$C_n^{(k)} = \binom{n+k-1}{k}$
mit Reihenfolge Variationen	$V_n^{(k)} = k! \binom{n}{k}$	$V_n^{(k)} = n^k$

Wahrscheinlichkeitsrechnung und ihre Rechenregeln

Axiome von Kolmogorow

(Ko₁) Wertebereich $0 \leq P(A) \leq 1$
 (Ko₂) sicheres Ereignis $P(\Omega) = 1$
 (Ko₃) A, B unvereinbar, d.h. $A \cap B = \emptyset$
 $P(A \cup B) = P(A) + P(B)$

Folgerungen daraus

$P(A) = 1 - P(\bar{A})$
 allgemein für $B \subseteq A$ $P(A \setminus B) = P(A) - P(B) \geq 0$
 $\Leftrightarrow P(A) \geq P(B)$
 Additionssatz $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
 $\Rightarrow P(\emptyset) = 0$
 A und B unabhängig: $\Leftrightarrow P(A \cap B) = P(A) \cdot P(B)$

bedingte Wk

"wie wahrscheinlich ist A
 unter der Bedingung von B ?" $P(A|B) = P(A \cap B) / P(B)$
 totale Wk $P(B) = P(B|A) P(A) + P(B|\bar{A}) P(\bar{A})$
 Satz von Bayes $P(A|B) = P(B|A) P(A) / P(B)$

Beispiel einer Umformulierungsmöglichkeit:

Die Wk, ein defektes Teil (D) von der Firma (L1) zu erhalten ist: $P(D|L_1)$

Die Wk, dass Firma (L1) das defekte Teil lieferte: $P(L_1|D)$

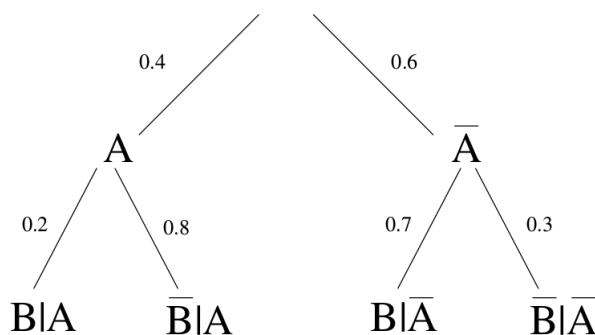
Beispiel einer Wk:

Laplace-Wk

$$P(A) = \frac{\text{Anzahl günstiger Fälle}}{\text{Anzahl möglicher Fälle}}$$

Lässt sich dann auf diese Weise aufstellen, wenn die möglichen Fälle gleichermaßen möglich sind (kein gezinkter Würfel hier!)

Baumdiagramm



Multiplikation entlang der Pfade

$$P(B|A) \cdot P(A) = 0.4 \cdot 0.2 = 0.08$$

$$P(B|\bar{A}) \cdot P(\bar{A}) = 0.6 \cdot 0.7 = 0.42$$

Addition der Pfadrereignisse

$$\begin{aligned} P(B) &= P(B|A) \cdot P(A) + P(B|\bar{A}) \cdot P(\bar{A}) \\ &= 0.08 + 0.42 = 0.5 \end{aligned}$$

Differentiation

in höherdimensionalen Räumen

Am Ende ist es schon so,
dass ich mir das zutraue.

Wir haben Ableitungen von Funktionen bisher als momentane Änderungsrate einer Funktion $F : \mathbb{R} \rightarrow \mathbb{R}$ an einem Punkt $x \in \mathbb{R}$ kennengelernt. In diesem Kapitel wollen wir uns dem Ableitungsbegriff für Funktionen in höheren Raumdimensionen befassen. Wir können uns etwa eine Temperaturverteilung in einem Zimmer vorstellen. Jeder Punkt im Raum ist ein $x \in \mathbb{R}^3$ und eine Funktion ordnet diesem Punkt eine Temperatur $T(x) \in \mathbb{R}$ zu. Wir haben es ja dann mit einer Funktion $T : \mathbb{R}^3 \rightarrow \mathbb{R}$ zu tun. Von einer Ableitung der Temperatur erwarten wir nun eine Auskunft darüber, wie sich die Temperatur ändert, wenn wir den Ort im Raum wechseln. Nun gibt es da ja jede Menge Richtungen. Im Eindimensionalen war das nicht der Fall. Sie sehen, dass sich da schon mehr Möglichkeiten auftun werden.

Generell bedeutet Funktion im Höherdimensionalen alle Varianten von

$$u : \mathbb{R}^n \rightarrow \mathbb{R}^m$$

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto \begin{pmatrix} u^1(x_1, \dots, x_n) \\ \vdots \\ u^m(x_1, \dots, x_n) \end{pmatrix}, \quad u^i : \mathbb{R}^n \rightarrow \mathbb{R}.$$

In Kapitel 3.1 führen wir den grundlegendsten Ableitungsbegriff, nämlich die partielle Ableitung, ein, die eine tragende Rolle spielen wird bei den meisten weiteren Begriffen. Wir betrachten dabei zunächst nur skalare Funktionen ($m=1$) mehrerer Veränderlicher ($n>1$) und dazu die Anwendungsbereiche Optimierung/Extremalstellen (Kapitel 3.2) und Optimierung mit Nebenbedingung (Kapitel 3.3). Im letzten Unterkapitel, Kapitel 3.4, werden wir uns ein paar ausgewählte Anwendungen verschiedener Ableitungsbegriffe zuwenden.

Zur Einstimmung betrachten bildliche Darstellungen von den verschiedenen Funktionsarten.

Beispiel 51 $m = n = 1$: Graph

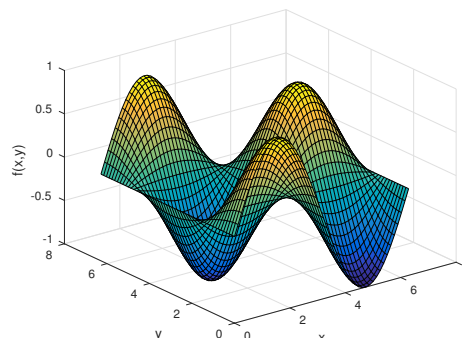
$$u : \mathbb{R} \rightarrow \mathbb{R}$$

$$x \mapsto x^2$$

Beispiel 52 $m = 2, n = 1$: Fläche

$$u : [0, 2\pi) \times [0, 2\pi) \subset \mathbb{R}^2 \rightarrow \mathbb{R}$$

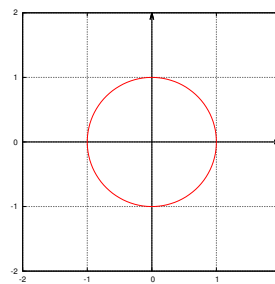
$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \sin x_1 \cos x_2$$



Beispiel 53 $m = 1, n = 2$: Kurve in der Ebene

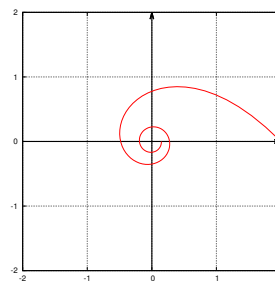
(a) $u : [0, 2\pi) \subset \mathbb{R} \rightarrow \mathbb{R}^2$

$$x \mapsto \begin{pmatrix} \cos x \\ \sin x \end{pmatrix}$$



(b) $u : [0, 4\pi) \subset \mathbb{R} \rightarrow \mathbb{R}^2$

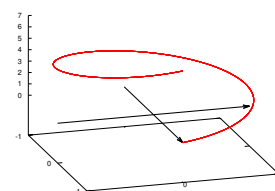
$$x \mapsto \frac{2}{x+1} \begin{pmatrix} \cos x \\ \sin x \end{pmatrix}$$



Beispiel 54 $m = 1, n = 3$: Kurve im Raum

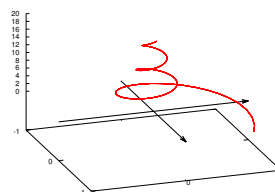
(a) $u : [0, 2\pi) \subset \mathbb{R} \rightarrow \mathbb{R}^3$

$$x \mapsto \begin{pmatrix} \cos x \\ \sin x \\ x \end{pmatrix}$$



(b) $u : [0, 6\pi) \subset \mathbb{R} \rightarrow \mathbb{R}^3$

$$x \mapsto \frac{2}{x+1} \begin{pmatrix} \cos x \\ \sin x \\ x \end{pmatrix}$$



Beispiel 55 $m = 2, n = 3$: Fläche im Raum (Kleinsche Flasche)

$$u : \mathbb{R}^2 \rightarrow \mathbb{R}^3$$

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \begin{pmatrix} u_1(x_1, x_2) \\ u_2(x_1, x_2) \\ u_3(x_1, x_2) \end{pmatrix}$$

mit

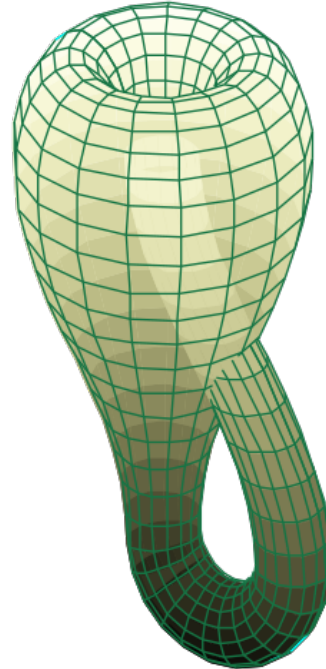
$$u_1(x_1, x_2) = 2(1 - \sin x_1) \cos x_1$$

$$+ (2 - \cos x_1) \cos x_2 (2e^{-(\frac{x_1}{2} - \pi)^2} - 1)$$

$$u_2(x_1, x_2) = (2 - \cos x_1) \sin x_2$$

$$u_3(x_1, x_2) = 6 \sin x_1 + \frac{1}{2} (2$$

$$- \cos x_1) \sin x_1 \cos x_2 e^{-(x_1 - \frac{3}{2}\pi)^2}$$



Quelle: Wikipedia

Skalarwertige Funktionen mehrerer Veränderlicher finden Sie hauptsächlich in Mathematikbüchern aus der Rubrik "Analysis II". Parametrisierungen sind Abbildungen $\mathbb{R} \rightarrow \mathbb{R}^{2,3}$ und $\mathbb{R}^2 \rightarrow \mathbb{R}^3$. Einen wertvollen Lektürtipp erhalten Sie mit [do Carmo \(1993\)](#).

3.1 Ableitungen von Funktionen mehrerer Veränderlicher

Wir betrachten zunächst skalare Funktionen

$$u : \mathbb{R}^n \rightarrow \mathbb{R}.$$

Solche Funktionen beschreiben "Flächen" im $\mathbb{R}^{n+1}$. Der Wert einer Ableitung oder Steigung in einem Punkt einer solchen Fläche hängt davon ab in welche Richtung man "schaut". Infolgedessen haben wir es nun mit sogenannten Richtungsableitungen zu tun.

Definition 3.1 (Richtungsableitung). Die Richtungsableitung von u am Punkt $x \in \mathbb{R}^n$ in Richtung $y \in \mathbb{R}^n$ ist gegeben durch den Grenzwert (sofern existent)

$$\frac{\partial}{\partial y} u(x) := \lim_{h \rightarrow 0} \frac{u(x + h y) - u(x)}{h}.$$

Wir schreiben auch u_y oder $D_y u$ statt $\frac{\partial}{\partial y} u$.



Wenn man eine Fläche mit der Ebene schneidet, die die betrachtete Richtung y enthält so reduziert sich der Ableitungsbegriff auf genau den, den wir schon im Zusammenhang mit Funktionen vom Typ $f : \mathbb{R} \rightarrow \mathbb{R}$ kennen.

Ein Spezialfall einer Richtungsableitung ist die in Richtung der Koordinatenachsen:

Definition 3.2 (partielle Ableitung). Die partiellen Ableitungen von u am Punkt $x \in \mathbb{R}^n$ sind gegeben durch

$$\begin{aligned} \frac{\partial}{\partial x_i} u(x) &:= \lim_{h \rightarrow 0} \frac{u(x + h e_i) - u(x)}{h} \\ &= \lim_{h \rightarrow 0} \frac{u(x_1, \dots, x_{i-1}, x_i + h, x_{i+1}, \dots, x_n) - u(x_1, \dots, x_n)}{h}, \quad i = 1, \dots, n \end{aligned}$$

Beispiel 56

Die partiellen Ableitungen von $u(x_1, x_2) = x_1^2 + x_1 \sin(x_2)$ sind gegeben durch

$$\begin{aligned} \frac{\partial}{\partial x_1} u(x_1, x_2) &= 2x_1 + \sin(x_2) \\ \frac{\partial}{\partial x_2} u(x_1, x_2) &= x_1 \cos(x_2). \end{aligned}$$

Die partiellen Ableitungen von u bei $x = (1, 2)^T$ sind dann gegeben durch

$$\begin{aligned} \frac{\partial}{\partial x_1} u(1, 2) &= 2 + \sin 2 \approx 2.9093 \\ \frac{\partial}{\partial x_2} u(1, 2) &= \cos 2 \approx -0.41615. \end{aligned}$$

Definition 3.3 (Gradient). Der Gradient einer Funktion u ist gegeben durch

$$\text{grad}(u)^T = \nabla u := \begin{pmatrix} u_{x_1} \\ \vdots \\ u_{x_n} \end{pmatrix}.$$

∇ heißt Nabla-Operator.



Der Nabla-Operator liefert immer einen Spaltenvektor, während hingegen der

$$\text{grad} f := (f_{x_1}, \dots, f_{x_n})$$

einen Zeilenvektor beschreibt. Die Komponenten sind aber die Gleichen.

Beispiel 57

Der Gradient von $u(x_1, x_2) = x_1^2 + x_1 \sin(x_2)$ ist gegeben durch

$$\nabla u = \begin{pmatrix} u_{x_1} \\ u_{x_2} \end{pmatrix} = \begin{pmatrix} 2x_1 + \sin(x_2) \\ x_1 \cos(x_2) \end{pmatrix}.$$

Der Gradient von $u(x_1, x_2) = x_1^2 + x_1 \sin(x_2)$ am Punkt $x = (1, 0)$ ist gegeben durch

$$(\nabla u)(x)|_{x=\begin{pmatrix} 1 \\ 0 \end{pmatrix}} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$

Der Gradient ist ein Vektor. Er hat eine Richtung und eine Länge



$$\nabla u(x) = \underbrace{\|\nabla u(x)\|}_{\text{Länge}} \underbrace{\frac{\nabla u(x)}{\|\nabla u(x)\|}}_{\text{Richtung}}$$

Der Betrag $\|\nabla u\|$ beinhaltet den Wert der Steigung von u am Punkt x in Richtung $\frac{\nabla u}{\|\nabla u\|}$.

Im folgenden Satz klären wir Rechenregeln des Gradienten, hinter denen letztlich Rechenregeln der partiellen Ableitungen stecken. Es wird nichts Überraschendes dabei herauskommen, da die Rechenregeln für die partiellen Ableitungen wie zu erwarten, sich wie diejenigen im eindimensionalen Fall verhalten. Der Vollständigkeit halber und auch um Sicher zu gehen werfen wir dennoch einen Blick darauf:

Satz 3.4. Für alle Konstanten $c \in \mathbb{R}$ und Skalarfelder $u, v : \mathbb{R}^n \rightarrow \mathbb{R}$ gilt:

$$\nabla c = \vec{0}$$

Linearität

$$\begin{aligned} \nabla (c \cdot u) &= c \cdot \nabla u \\ \nabla (u + v) &= \nabla u + \nabla v \end{aligned}$$

Produktregel

$$\begin{aligned} \nabla (u v) &= u \nabla v + v \nabla u \\ \nabla (u^n) &= n u^{n-1} \nabla u \quad \text{für } n \neq 0 \end{aligned}$$

Beweis Satz 3.4:

ÜA

□

Im Folgenden wollen wir einen Zusammenhang zwischen partieller und Richtungsableitung herstellen. Bisher haben wir für die Richtungsableitung ja lediglich den Differenzenquotient notiert, so ist sie definiert. Für die partielle Ableitung haben wir aber schon eine bequeme Methode entdeckt, wie wir sie jeweils berechnen können. Wir werden gleich in Satz 3.6 sehen,

dass wir die Richtungsableitung über die Bestimmung der partiellen Ableitungen berechnen können. Das ist natürlich sehr viel einfacher, als jeweils einen Grenzwert zu berechnen. Wir benötigen nur eben noch zwei Kenntnisse, um den entsprechenden Satz schnell einzusehen.

Wenn wir eine Parametrisierung $g : \mathbb{R} \rightarrow \mathbb{R}^m, t \mapsto g(t)$, hier also eine Kurve im Raum, ableiten wollen haben wir es ja nur mit einer Variablen $t \in \mathbb{R}$ zu tun. Da sich die Ableitungen auf jede der m Komponenten von $g(t) \in \mathbb{R}^m$ beziehen erhalten wir dort unseren altbekannten Ableitungsbegriff für Funktionen im Eindimensionalen:

$$g'(t) = \begin{pmatrix} g'_1(t) \\ \vdots \\ g'_M(t) \end{pmatrix}$$

Beispiel 58 Erste Ableitung einer Parametrisierung

$$g(t) = \begin{pmatrix} t \\ t \sin t \\ t^2 \end{pmatrix} \Rightarrow g'(t) = \begin{pmatrix} 1 \\ \sin t + t \cos t \\ 2t \end{pmatrix}$$

$$\text{und } g''(t) = \begin{pmatrix} 0 \\ 2 \cos t - t \sin t \\ 2 \end{pmatrix}, \text{ etc.}$$

Das ist auch schon alles an Eigenschaften, was wir für unsere Zwecke von Parametrisierungen benötigen. Damit können wir nun folgende Definition aussprechen.

Satz 3.5 (Kettenregel). Sei $f : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ und $g : I \subseteq \mathbb{R} \rightarrow \mathbb{R}^n$. Dann gilt

$$\frac{d}{dt} (f \circ g)(t) = \sum_{k=1}^n \frac{\partial}{\partial g_k} f(g(t)) g'_k(t).$$

Beweis Satz 3.5:

Wir beschränken den Beweis auf $n = 2$. Er lässt sich sukzessive auf höhere Raumdimensionen fortführen. Der Beweis selbst ist ein reines durchrechnen des Differenzenquotienten wir die Ableitung:

$$\frac{d}{dt} f(g_1(t), g_2(t)) = \lim_{h \rightarrow 0} \frac{f(g_1(t+h), g_2(t+h)) - f(g_1(t), g_2(t))}{h}$$

Wir addieren eine **Null**

$$= \lim_{h \rightarrow 0} \left(\frac{f(g_1(t+h), g_2(t)) - f(g_1(t), g_2(t))}{h} \right. \\ \left. + \frac{f(g_1(t+h), g_2(t+h)) - f(g_1(t+h), g_2(t))}{h} \right)$$

und multiplizieren mit *Eins*

$$= \lim_{h \rightarrow 0} \left(\frac{f(g_1(t+h), g_2(t)) - f(g_1(t), g_2(t))}{g_1(t+h) - g_1(t)} \cdot \frac{g_1(t+h) - g_1(t)}{h} + \frac{f(g_1(t+h), g_2(t+h)) - f(g_1(t+h), g_2(t))}{g_2(t+h) - g_2(t)} \cdot \frac{g_2(t+h) - g_2(t)}{h} \right)$$

In jedem Summand stecken Faktoren, die Differenzenquotienten darstellen. Mit $\tilde{h} := g_i(t+h) - g_i(t)$ für $i \in \{1, 2\}$ gilt auch $g_i(t+h) = g_i(t) + \tilde{h}$ und damit:

$$= \lim_{h, \tilde{h} \rightarrow 0} \left(\frac{f(g_1(t) + \tilde{h}, g_2(t)) - f(g_1(t), g_2(t))}{\tilde{h}} \cdot \frac{g_1(t+h) - g_1(t)}{h} + \frac{f(g_1(t+h), g_2(t) + \tilde{h}) - f(g_1(t+h), g_2(t))}{\tilde{h}} \cdot \frac{g_2(t+h) - g_2(t)}{h} \right) \\ = f_{g_1}(g)g'_1(t) + f_{g_2}(g)g'_2(t)$$

□

Beispiel 59

Erst mal ein einfaches Beispiel: Es sei $f(x) = x_1 \cdot x_2$ und $g(t) = (t, t)$. Dann ist

$$(f \circ g)(t) = g_1(t) \cdot g_2(t) = t \cdot t = t^2.$$

Direkt nach t abgeleitet ergibt

$$(f \circ g)'(t) = 2t$$

und der Weg über Kettenregel sieht so aus:

$$(f \circ g)'(t) = f_{g_1}(g)g'_1(t) + f_{g_2}(g)g'_2(t) = g_2(t) \cdot 1 + g_1(t) \cdot 1 = t + t = 2t$$

Jetzt etwas schwieriger

Beispiel 60

Es sei $f(x) = f(x_1, x_2) = x_1 + x_2 \cos x_1$ und $g(t) = \begin{pmatrix} \cos t \\ \sin t \end{pmatrix}$. Dann lautet die Verkettung

$$(f \circ g)(t) = f(g_1(t), g_2(t)) = g_1(t) + g_2(t) \cos g_1(t) = \cos t + \sin t \cos(\cos t).$$

Leiten wir diese Funktion nach t ab, so erhalten wir

$$(f \circ g)'(t) = (\cos t + \sin t \cos(\cos t))' = -\sin t + \cos t \cos(\cos t) + \sin^2 t \sin(\cos t).$$

Das gleiche machen wir nun über die Kettenregel und verwenden dazu die Ableitungen

$$g'_1(t) = -\sin t \quad \text{und} \quad g'_2(t) = \cos t$$

und

$$\begin{aligned} f_{x_1} = 1 - x_2 \sin x_1 &\Rightarrow f_{g_1}(g) = 1 - g_2(t) \sin g_1(t) \\ f_{x_2} = \cos x_1 x_1 &\Rightarrow f_{g_2}(g) = \cos g_1(t) . \end{aligned}$$

Jetzt die Kettenregel:

$$\begin{aligned} \frac{d}{dt}(f(g(t))) &= f_{g_1} g_1'(t) + f_{g_2} g_2'(t) \\ &= (1 - g_2(t) \sin g_1(t))(-\sin t) + \cos g_1(t) \cos t \\ &= -(1 - \sin t \sin(\cos t)) \sin t + \cos t \cos(\cos t) \end{aligned}$$

Voilà!

Das war die Königsetappe, aber die mühevollen Arbeit wird belohnt werden. Wir erhalten jetzt ein Werkzeug, mit dessen Hilfe wir bequem jede Richtungsableitung berechnen können:

Satz 3.6 (Zusammenhang zwischen Nabla-Operator und Richtungsableitung). Sei $\nu \in \mathbb{R}^n$ mit $\|\nu\| = 1$. Dann gilt

$$\nabla u \cdot \nu = u_\nu .$$

Beweis Satz 3.6:

Bei der Richtungsableitung betrachten wir ja die Änderungsrate von u am Punkt x in Richtung ν . Das reduziert unseren Ausdruck auf eine eindimensionale Ableitung und es gilt:

$$u_\nu = \frac{d}{dt} u(x + t\nu) \Big|_{t=0}$$

für $g(t) = x + t\nu = (x_1 + t\nu_1, x_2 + t\nu_2)^T = (g_1(t), g_2(t))^T$ schreiben wir

$$\begin{aligned} &= \frac{d}{dt} u(g(t)) \Big|_{t=0} \\ &= \frac{d}{dt} u(g_1(t), g_2(t)) \Big|_{t=0} \end{aligned}$$

und mit der Kettenregel erhalten wir

$$\begin{aligned} &= (u_{g_1} g_1'(t) + u_{g_2} g_2'(t)) \Big|_{t=0} \\ &= u_{g_1} \nu_1 + u_{g_2} \nu_2 \\ &= \nabla u(x) \cdot \nu \end{aligned}$$

□

Folgerung: Aus Satz 3.6 können wir direkt folgern, dass der Gradient einer Funktion $u : \mathbb{R}^n \rightarrow \mathbb{R}$ immer in Richtung des steilsten Anstiegs zeigt, denn es gilt $\forall \nu \in \mathbb{R}^n$ mit $\|\nu\| = 1$

$$|u_\nu| = |\nabla u \cdot \nu| = \|\nabla u\| \|\nu\| \underbrace{\cos \angle(\nabla u, \nu)}_{\leq 1} \leq \|\nabla u\| \|\nu\| = \|\nabla u\|.$$



Da der Gradient immer in Richtung des steilsten Anstiegs Am Punkt x zeigt ist entsprechend $-\nabla u$ die Richtung des steilsten Abstiegs. Man sagt auch *Flussrichtung* zu $-\nabla u$. In diese Richtung würde an diesem Punkt Wasser fließen.

Diese Eigenschaft macht den Gradienten zu einem zentralen Element bei der Differentiation in höheren Raumdimensionen. Er findet Einsatz bei der Modellierung, sobald etwas fließt; ob das ein Temperatur- oder Druckausgleich ist, Wasser, Lawinen, eine Konzentration einer chemischen Substanz in einer Flüssigkeit, ein Elektronenfluss oder Ähnliches.

Beispiel 61

Der Wert des steilsten Anstiegs von $u(x_1, x_2) = x_1^2 + x_1 \sin(x_2)$ am Punkt $x = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ ist gegeben durch:

$$\|\nabla u(1, 0)\| = \left\| \begin{pmatrix} u_{x_1}(1, 0) \\ u_{x_2}(1, 0) \end{pmatrix} \right\| = \left\| \begin{pmatrix} 2 \\ 1 \end{pmatrix} \right\| = \sqrt{5}.$$

Eine beliebige Richtungsableitung erhalten wir durch

$$\nabla u \cdot \nu = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} \nu_1 \\ \nu_2 \end{pmatrix} = 2\nu_1 + \nu_2.$$

Eine kleine, handliche Notation, um kurzerhand die Stetigkeit und Differenzierbarkeit einer Funktion zu beschreiben bietet die folgende

Definition 3.7. Es sei $D \subseteq \mathbb{R}^n$. Die Funktion $f : D \rightarrow \mathbb{R}$ heißt auf D k -mal stetig differenzierbar, falls alle partiellen Ableitungen bis zur Ordnung k existieren und stetig sind. Wir schreiben

$$f \in C^k(D) := \{f : D \rightarrow \mathbb{R} \mid f^{(m)} \in C^0(D), m \leq k\}$$

Mit Ableitungen bis zur Ordnung (m) meint dabei m partielle Ableitungen in alle möglichen Achsenrichtungen. Also bei $m = 1$ bezeichnet $f^{(1)}$ erste Ableitung in x_1 - und erste in x_2 -Richtung und für $m = 2$ meint $f^{(2)}$ zwei Ableitungen in Richtung x_1 , also $f_{x_1x_1}$, zwei in Richtung x_2 , also $f_{x_2x_2}$ und Mischformen, d.h. erste in x_1 , zweite in x_2 und umgekehrt, also $f_{x_1x_2}$ und $f_{x_2x_1}$.

Werfen wir also noch einen Blick auf Ableitungen höherer Ordnung grad von ∇ :

$$\text{grad } \nabla f = \text{grad} \begin{pmatrix} f_x \\ f_y \end{pmatrix} = \begin{pmatrix} \text{grad } f_x \\ \text{grad } f_y \end{pmatrix} = \begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix}$$

Diese Matrix, die für diese Art zweiter Ableitung steht hat einen Namen:

Definition 3.8 (Hessematrix). Sei $f \in C^k(D)$, $D \subseteq \mathbb{R}^n$. Dann heißt die Matrix

$$Hf = \begin{pmatrix} f_{x_1 x_1} & \cdots & f_{x_1 x_n} \\ \vdots & & \vdots \\ f_{x_n x_1} & \cdots & f_{x_n x_n} \end{pmatrix}$$

Hessematrix der Funktion f .

Dabei meint

$$f_{x_i x_j} = \frac{\partial^2 f}{\partial x_i \partial x_j}$$

erste partielle Ableitung in Richtung x_i und zweite partielle Ableitung in Richtung x_j .

Beispiel 62 Es sei $f(x) = 2x_1 + x_2^3$, dann gilt

$$\begin{aligned} f_{x_1} &= 2, & f_{x_2} &= 3x_2^2, \\ f_{x_1 x_1} &= 0, & f_{x_1 x_2} &= 0, \\ f_{x_2 x_1} &= 0, & f_{x_2 x_2} &= 6x_2. \end{aligned}$$

Die Hessematrix lautet dann

$$Hf = \begin{pmatrix} 0 & 0 \\ 0 & 6x_2 \end{pmatrix}.$$

Satz 3.9. Sei $D \subseteq \mathbb{R}^n$ und $f \in C^2(D)$, dann gilt

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i} \quad \forall 1 \leq i, j \leq n.$$

3.2 Extremwertberechnung

Wie im eindimensionalen Fall ist man bei Funktionen mehrerer Veränderlicher an Extremwerten interessiert.

Definition 3.10 (lokales Extremum). Sei $f \in C^0(U(x))$, $x \in \mathbb{R}^n$. Ein lokales Extremum von f ist gegeben, wenn

$$\begin{aligned} f(x) &\geq f(s) & \forall s \in U(x) \\ \text{bzw} & & \\ f(x) &\leq f(s) & \forall s \in U(x) \end{aligned}$$

Definition 3.11 (stationärer Punkt). Es sei $D \subseteq \mathbb{R}^n$ und $f \in C^1(D)$. Die Stelle $s \in D$ mit

$$(\nabla f)(s) = 0$$

heißt kritische Stelle oder stationärer Punkt von f .

Bei einem lokalen Extremum gilt



$$\nabla f = 0,$$

falls $f \in C^1(U(x))$. Aber Achtung: Wir werden sehen, dass dies kein hinreichendes Kriterium ist. Stationäre Punkte sind lediglich Kandidaten für Extremstellen; mehr nicht.

Beispiel 63 Extremstelle Für $f(x, y) = x^2 + y^2$ gilt

$$\nabla f = \begin{pmatrix} 2x \\ 2y \end{pmatrix} = 0 \quad \Leftrightarrow \quad \begin{pmatrix} x \\ y \end{pmatrix} = 0.$$

f hat bei $(0, 0)$ ein Extremum, was man leicht einsieht, wenn man sich den Graphen dazu anschaut.

Beispiel 64 Sattelpunkt Für $f(x, y) = x^2 - y^2$ gilt

$$\nabla f = \begin{pmatrix} 2x \\ -2y \end{pmatrix} = 0 \quad \Leftrightarrow \quad \begin{pmatrix} x \\ y \end{pmatrix} = 0.$$

f hat aber bei $(0, 0)$ kein Extremum, sondern einen Sattelpunkt.

Die Frage, die sich nun stellt ist "wann ist ein Extremum ein Maximum und wann ein Minimum?" Wir brauchen etwas Analoges zum eindimensionalen Fall, da galt ja für Funktionen $f : \mathbb{R} \rightarrow \mathbb{R}$ bei $f'(s) = 0$

$f'(s) = 0$	$\wedge$	$f''(s) < 0$	Maximum
$f'(s) = 0$	$\wedge$	$f''(s) > 0$	Minimum

Das Pendant zur zweiten Ableitung im eindimensionalen Fall, welche das Krümmungsverhalten - wohlbermerkt das Verhalten nicht die Krümmung selbst! - eine Kurve beschreibt, je nach Vorzeichen, ist im höherdimensionalen durch die Definitheit der Hessematrix gegeben. Wir erinnern uns an Inhalte der Linearen Algebra:

Definition 3.12 (Definitheit). Eine quadratische Matrix $A \in \mathbb{R}^{n \times n}$ heißt positiv (semi) definit, falls

$$\langle Ax, x \rangle \underset{(\geq)}{>} 0 \quad \forall x \in \mathbb{R}^n \setminus \{0\},$$

negativ (semi) definit, falls

$$\langle Ax, x \rangle \underset{(\leq)}{<} 0 \quad \forall x \in \mathbb{R}^n \setminus \{0\}$$

und heißt indefinit, falls es Vektoren $x, y \in \mathbb{R}^n \setminus \{0\}$ gibt mit

$$\langle Ax, x \rangle < 0 < \langle Ay, y \rangle.$$

Eines ist schnell klar: Wenn die Matrix positiv definit ist, so müssen die Diagonalelemente positiv sein, denn $\langle Ax, x \rangle > 0 \forall x$ bedeutet insbesondere auch für die Einheitsvektoren e_i , und zwar alle. Und dann gilt nämlich $\langle Ae_i, e_i \rangle = A_{ii} > 0$. Die Umkehrung gilt allerdings nicht. Wenn die Diagonalelemente einer Matrix alle positiv sind, heisst das noch lange nicht, dass die Matrix positiv definit ist. Dazu ein Beispiel:

Beispiel 65

$$A = \begin{pmatrix} 1 & 0 \\ -3 & 1 \end{pmatrix}$$

ist nicht positiv definit. Berechnen Sie $\langle Ax_i, x_i \rangle$ mit $x_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ und $x_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. Sie erhalten dann

$$\langle Ax_1, x_1 \rangle = -1 \quad \text{und} \quad \langle Ax_2, x_2 \rangle = 1.$$

Da stellt sich natürlich die Frage, wie man generell diese Begingung prüfen kann. Wir können ja nicht alle $x \in \mathbb{R}^n$ ausprobieren! Wir werden einige Methoden kennenlernen, die sich alle auf symmetrische Matrizen beziehen. Das macht aber nichts, denn mit dem folgenden Satz können wir klären, wie man sich geschickt in die Situation manövriert, bezüglich Definitheit immer eine symmetrische Matrix vorliegen zu haben.

Satz 3.13. *Eine quadratische Matrix A , die nicht notwendigerweise symmetrisch ist, ist genau dann positiv definit, wenn ihr symmetrischer Anteil*

$$\frac{1}{2}(A + A^T)$$

positiv definit ist. Entsprechendes gilt für "negativ Definitheit" und "Semidefinitheit".

Beweis Satz 3.13:

$$\begin{aligned} 0 \leq \langle (A + A^T)x, x \rangle &= \langle Ax, x \rangle + \langle A^T x, x \rangle \\ &= \langle Ax, x \rangle + \langle x, Ax \rangle = 2 \langle Ax, x \rangle \end{aligned}$$

□

Beispiel 66

Der symmetrische Anteil der Matrix A in Beispiel 65 lautet

$$\frac{1}{2}(A + A^T) = \frac{1}{2} \left(\begin{pmatrix} 1 & 0 \\ -3 & 1 \end{pmatrix} + \begin{pmatrix} 1 & -3 \\ 0 & 1 \end{pmatrix} \right) = \frac{1}{2} \begin{pmatrix} 2 & -3 \\ -3 & 2 \end{pmatrix}.$$

Der Test mit den gleichen x_1 , und x_2 liefert

$$\langle Ax_1, x_1 \rangle = -1 \quad \text{und} \quad \langle Ax_2, x_2 \rangle = 1$$

und somit Indefinitheit der Matrix $\frac{1}{2}(A + A^T)$ und mit Satz 3.13 folglich auch Indefinitheit der Matrix A .

Wir können uns fortan auf symmetrische Matrizen beschränken, wenn das von Nöten ist.

Satz 3.14 (Kriterium von Sylvester / Hurwitz-Kriterium). *Es sei $A \in \mathbb{R}^{n \times n}$ eine symmetrische Matrix und $M_k \in \mathbb{R}^{k \times k}$ bezeichnen die n führenden Hauptminoren von A :*

$$M_k = (A_{ij})_{1 \leq i, j \leq k}$$

Dann gilt:

- A ist positiv definit, wenn gilt

$$\det M_k > 0 \quad \forall k,$$

das heißt, wenn die Determinanten aller n führender Hauptminoren positiv sind.

- A ist negativ definit, wenn gilt

$$\det M_k = (-1)^k C_k, \quad C_k > 0.$$

das heißt, wenn die Determinanten der geradzahligen Hauptminoren positiv und die anderen negativ sind.

Sind keine der beiden Eigenschaften gegeben, so liefert dieses Kriterium keine Aussage.

(ohne Beweis)

Beispiel 67

(a) Die Matrix

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix}$$

besitzt zwei führende Hauptminoren, nämlich

$$M_1 = (1) \quad \text{und} \quad M_2 = \begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix} = A.$$

Die entsprechenden Determinanten

$$\det M_1 = 1 \quad \text{und} \quad \det M_2 = 1$$

sind alle positiv und damit ist A positiv definit.

(b) Die Matrix

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

besitzt drei führende Hauptminoren, nämlich

$$M_1 = (-3), \quad M_2 = \begin{pmatrix} -3 & -3 \\ -3 & -5 \end{pmatrix} \quad \text{und} \quad M_3 = \begin{pmatrix} -3 & -3 & 1 \\ -3 & -5 & 0 \\ 1 & 0 & -4 \end{pmatrix} = B.$$

Die entsprechenden Determinanten

$$\det M_1 = -3, \quad \det M_2 = 6 \quad \text{und} \quad \det M_3 = -19$$

zeigen ein alternierendes Vorzeichen, jeweil negativ bei ungeradzahligen Minoren. Infolge dessen ist A negativ definit.

(c) Die Matrix

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

ist nicht symmetrisch. Wir betrachten also ihren symmetrischen Anteil

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

Es ist

$$\det M_1, \det M_2 > 0$$

und daher sind A_S und auch A positiv definite Matrizen.



Den Faktor $\frac{1}{2}$ beim symmetrischen Anteil, kann man bei der Untersuchung auf Definitheit auch ignorieren. Es geht ja hier nur um das Vorzeichen.

Wir können bereits jetzt schon einige Fälle durch bloßes Hinschauen und ein wenig den Kopf neigen herausfinden. Überzeugen Sie sich:

Beispiel 68

$$\underbrace{\begin{pmatrix} 1 & 2 \\ 2 & -4 \end{pmatrix}}_{\text{indefinit}} \quad \underbrace{\begin{pmatrix} -2 & 1 \\ 1 & -2 \end{pmatrix}}_{\text{neg. def.}} \quad \underbrace{\begin{pmatrix} 4 & 2 \\ 2 & 1 \end{pmatrix}}_{?} \quad \underbrace{\begin{pmatrix} 1 & 1 \\ -2 & 1 \end{pmatrix}}_{\text{pos. definit}}$$

Das linke Beispiel klärt sich direkt durch das wechselnde Vorzeichen auf der Diagonalen. Bei der zweiten Matrix erhalten wir als Determinante des zweiten führenden Hauptminors den

Wert Null, so dass dieses Kriterium keine Aussage liefert. Bei der dritten Matrix können wir das Hurwitz-Kriterium im Kopf durchrechnen. Zur Bestimmung der Definitheit der Matrix ganz rechts, betrachten wir den symmetrischen Anteil und erhalten über das Hurwitz-Kriterium, dass dieser positiv definit ist, also gilt das auch für die ursprüngliche Matrix.

Das Hurwitz-Kriterium liefert zwar eine Möglichkeit, wie man relativ einfach nachrechnen kann, wie es um die Definitheit einer Matrix bestellt ist, allerdings deckt es nicht alle Fälle ab (siehe Beispiel 68). Meistens sind die Matrizen, deren Definitheit nicht mit dem Hurwitz-Kriterium geklärt werden können indefinit. Allerdings kann man sich dessen nicht sicher sein.

Ein Kriterium, das auf jeden Fall Auskunft liefert, aber etwas rechenaufwendiger ist beinhaltet der folgende

Satz 3.15. *Es seien $\lambda_1, \dots, \lambda_l \in \mathbb{R}$, $l \leq n$ Eigenwerte der symmetrischen Matrix $A \in \mathbb{R}^{n \times n}$. Dann gilt*

$$\begin{aligned} \lambda_1, \dots, \lambda_l &> (\geq) 0 &\Rightarrow A \text{ ist positiv (semi-) definit} \\ \lambda_1, \dots, \lambda_l &< (\leq) 0 &\Rightarrow A \text{ ist negativ (semi-) definit} \\ \exists \lambda_i, \lambda_k \mid \lambda_i \lambda_k &< 0 &\Rightarrow A \text{ ist indefinit} \end{aligned}$$

Beweis Satz 3.15:

Da A symmetrisch ist gibt es n Eigenvektoren $v_1, \dots, v_n$, die eine Orthonormalbasis (ONB) des $\mathbb{R}^n$ bilden (Lineare Algebra), das heißt es gilt

$$\langle v_i, v_j \rangle = \delta_{ij}.$$

Dabei beschreibe

$$\delta_{ij} = \begin{cases} 1 & \text{für } i = j \\ 0 & \text{für } i \neq j \end{cases}$$

das Kronecker-Symbol. Es kann also jedes $x \in \mathbb{R}^n$ als Linearkombination der ONB dargestellt werden

$$x = \sum_{i=1}^n \alpha_i v_i$$

Damit gilt

$$\begin{aligned} \langle Ax, x \rangle &= \left\langle A \sum_{i=1}^n \alpha_i v_i, \sum_{j=1}^n \alpha_j v_j \right\rangle = \left\langle \sum_{i=1}^n \alpha_i A v_i, \sum_{j=1}^n \alpha_j v_j \right\rangle \\ &= \left\langle \sum_{i=1}^n \alpha_i \lambda_i v_i, \sum_{j=1}^n \alpha_j v_j \right\rangle = \sum_{i=1}^n \alpha_i \lambda_i \sum_{j=1}^n \alpha_j \langle v_i, v_j \rangle \\ &= \sum_{i=1}^n \alpha_i \lambda_i \sum_{j=1}^n \alpha_j \delta_{ij} = \sum_{i=1}^n \alpha_i \lambda_i \alpha_i = \sum_{i=1}^n \alpha_i^2 \lambda_i \end{aligned}$$

Daraus ergibt sich direkt, dass die Vorzeichen der λ_i die Definitheit der Matrix A bestimmen.

□

Beispiel 69 Betrachten wir die dritte Matrix aus Beispiel 68, deren Definitheit noch ungeklärt ist.

$$A = \begin{pmatrix} 4 & 2 \\ 2 & 1 \end{pmatrix}$$

hat die Eigenwerte

$$\lambda_1 = 0 \quad \text{und} \quad \lambda_2 = 5$$

und ist folglich positiv semidefinit.

Es gibt noch ein weiteres Kriterium, dass leicht überprüfbar ist, welches sich auf stark diagonaldominante Matrizen bezieht.

Definition 3.16 (diagonaldominante Matrizen). Falls für eine Matrix $A \in \mathbb{R}^{n \times n}$

$$\sum_{\substack{i=1 \\ i \neq j}}^n |a_{ij}| < (\leq) |a_{jj}| \quad \forall j = 1, \dots, n$$

gilt, so heißt sie stark (schwach) diagonaldominant.

In Beispiel 68 sind die zweite und die dritte Matrix stark diagonaldominant, die jeweils anderen nicht. Wenn eine Matrix diese Eigenschaft hat, dann wirkt sich das derart stark auf die Größe der Eigenwerte aus, dass sie entsprechend alle ein positives Vorzeichen erhalten.

Satz 3.17. Eine reelle, symmetrische, stark/schwach diagonaldominante Matrix mit nichtnegativen (nichtpositiven) Diagonaleinträgen ist positiv (negativ) definit/semidefinit.

(ohne Beweis)



Die Aussage des Satzes 3.17 ist eine Implikation. Die Umkehrung gilt nicht. Es kann durchaus positiv definite Matrizen geben, die nicht diagonaldominant sind.

Beispiel 70

1. Die symmetrische Matrix

$$A = \begin{pmatrix} 4 & -2 & 1 \\ -2 & 6 & -3 \\ 1 & -3 & 5 \end{pmatrix}$$

ist stark diagonaldominant mit positiven Diagonalelementen und somit positiv definit. Dieses Ergebnis wird von Hurwitz bestätigt:

$$\det M_1 = 4 > 0, \det M_2 = 20 > 0 \quad \text{und} \quad \det M_3 = 70 > 0$$

2. Die symmetrische Matrix

$$A = \begin{pmatrix} -4 & -2 & 1 \\ -2 & -6 & -3 \\ 1 & -3 & -5 \end{pmatrix}$$

ist stark diagonaldominant mit negativen Diagonalelementen und somit negativ definit. Dieses Ergebnis wird von Hurwitz bestätigt:

$$\det M_1 = -4 < 0, \det M_2 = 20 > 0 \quad \text{und} \quad \det M_3 = -46 < 0$$

3. Die symmetrische Matrix

$$A = \begin{pmatrix} 4 & -2 & 2 \\ -2 & 6 & -3 \\ 2 & -3 & 5 \end{pmatrix}$$

ist nur schwach diagonaldominant mit positiven Diagonalelementen. Sie ist demnach zunächst positiv semidefinit. Semidefinitheit beinhaltet Definitheit. A könnte also durchaus auch positiv definit sein. Was sagt Hurwitz dazu?

$$\det M_1 = 4 > 0, \det M_2 = 20 > 0 \quad \text{und} \quad \det M_3 = 64 > 0$$

Die Matrix A ist tatsächlich positiv definit.

4. Die symmetrische Matrix

$$A = \begin{pmatrix} 4 & -2 & 2 \\ -2 & 6 & -3 \\ 2 & -3 & 1 \end{pmatrix}$$

ist gar nicht diagonaldominant. Was sagt Hurwitz dazu?

$$\det M_1 = 4 > 0, \det M_2 = 20 > 0 \quad \text{und} \quad \det M_3 = -16 < 0$$

Der weiß das also auch nicht.

Formel 10: Vorgehensweise bei der Untersuchung der Definitheit einer Matrix:

1. Diagonale Gibt es wechselnde Vorzeichen auf der Diagonale?
Wenn ja ist die Matrix indefinit
2. Symmetrie Ist die Matrix symmetrisch?
Wenn nicht betrachte weiterhin ihren symmetrischen Anteil $A + A^T$
3. Diagonaldominanz
Ist die Matrix diagonaldominant?
Achtung: Liefert dieses Kriterium Semidefinitheit ist bei Bedarf
immer noch auf Definitheit zu prüfen
4. Hurwitz/Sylvester
Liefert dieses Kriterium eine Aussage?
5. Eigenwerte Hat die Matrix nur positive/negative Eigenwerte?

Sie können immer eine Aussage über die Beschaffenheit der Eigenwerte erhalten. Die Schritte sind nach Schwierigkeitsgrad geordnet. Viele Fällen lösen Sie schon durch geschultes Hingucken bei Schritt 1 bis 4, manchmal auch 5.

Der folgende Satz liefert uns nun den Zusammenhang von Definitheit und unserem eigentlichen Anliegen, nämlich entscheiden zu können, ob ein stationärer Punkt, ein Kandidat also für eine Extremstelle, tatsächlich ein solcher ist und wenn ja, ob es sich um ein lokales Maximum oder Minimum handelt.

Satz 3.18. Sei $f \in C^2(D)$ und $s \in D \subseteq \mathbb{R}^n$ mit

$$\nabla f(s) = 0.$$

Dann ist

s lokales Minimum, wenn $Hf(s)$ positiv definit ist,
 s lokales Maximum, wenn $Hf(s)$ negativ definit ist,
 s kein lokales Extremum, wenn $Hf(s)$ indefinit ist.

Bei Semidefinitheit lässt sich das Kriterium nicht anwenden.

(ohne Beweis)

Beispiel 71 Wir suchen lokale Extremstellen der Funktion

$$f(x, y) = 2x^2 + xy^2 - \frac{1}{3}y^3 - 8x. \quad (13)$$

Schauen wir zunächst nach stationären Punkten. Der Gradient

$$\nabla f = \begin{pmatrix} y^2 + 4x - 8 \\ 2xy - y^2 \end{pmatrix}$$

verschwindet an den Stellen

$$K_1 = \begin{pmatrix} -2 \\ -4 \end{pmatrix}, K_2 = \begin{pmatrix} 2 \\ 0 \end{pmatrix}, K_3 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

Links in Abbildung 6 sind die entsprechenden Stellen geplottet. Selbst mit der Graphik ist nicht klar auszumachen, welche dieser Stellen tatsächlich Extremalstellen darstellen.

Sehen wir uns also die Hessematrix an. Diese lautet zunächst

$$Hf(x, y) = \begin{pmatrix} 4 & 2y \\ 2y & 2x - 2y \end{pmatrix}.$$

Diese werten wir für die drei Kandidaten aus und untersuchen die Definitheit an den entsprechenden Stellen:

$Hf(K_1) = \begin{pmatrix} 4 & -8 \\ -4 & 4 \end{pmatrix}$	indefinit	$\Rightarrow K_1$ ist kein Extremum
$Hf(K_2) = \begin{pmatrix} 4 & 0 \\ 0 & 4 \end{pmatrix}$	pos. definit	$\Rightarrow K_2$ ist ein lok. Minimum
$Hf(K_3) = \begin{pmatrix} 4 & 4 \\ 4 & -2 \end{pmatrix}$	indefinit	$\Rightarrow K_3$ ist kein Extremum

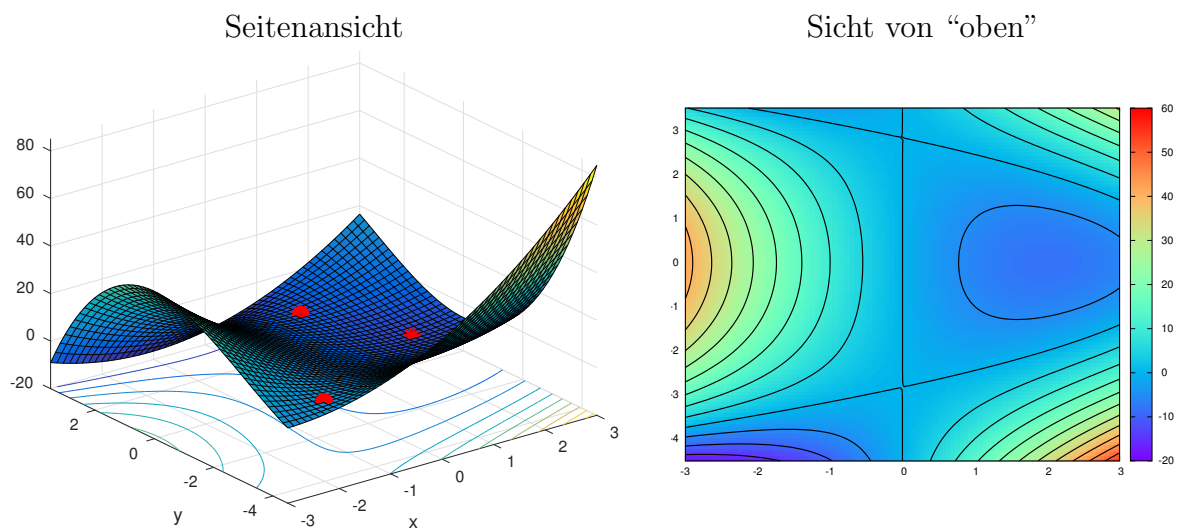


Abbildung 6: Graph der Funktion (13)

Demnach hat $f(x)$ ein einziges lokales Extremum und zwar ein lokales Minimum bei $K_2 = (2, 0)$.

Genau wie im eindimensionalen untersucht man die Funktion am Rand des Gebietes auf lokale Extrema. Die kann es ja geben auch wenn kein stationärer Punkt vorhanden ist. Im Höherdimensionalen haben wir es dann mit Richtungsableitungen zu tun. Aus diesem Grund schauen wir uns die Richtungsableitungen für sich nochmal kurz an. Dann haben wir es später leichter. Es ist ja für eine Funktion $f : \mathbb{R}^n \rightarrow \mathbb{R}$ die Richtungsableitung von f in Richtung $v \in \mathbb{R}^n$ mit $\|v\| = 1$ gegeben durch

$$f_v = \nabla f \cdot v = f_{x_1}v_1 + \cdots + f_{x_n}v_n = \sum_{i=1}^n f_{x_i}v_i.$$

Genau so machen wir es mit der zweiten Richtungsableitung. Wir wollen die Richtungsablei-

tung von f_v in Richtung $w \in \mathbb{R}^n$, $\|w\| = 1$, dann ist das doch

$$\begin{aligned}
 f_{vw} &= \nabla f_v \cdot w && (\text{VW haha :)}) \\
 &= \nabla (\nabla f \cdot v) \cdot w \\
 &= \nabla (f_{x_1} v_1 + \cdots + f_{x_n} v_n) \cdot w \\
 &= \begin{pmatrix} (f_{x_1} v_1 + \cdots + f_{x_n} v_n)_{x_1} \\ \vdots \\ (f_{x_1} v_1 + \cdots + f_{x_n} v_n)_{x_n} \end{pmatrix} \cdot \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} \\
 &= \begin{pmatrix} f_{x_1 x_1} v_1 + \cdots + f_{x_n x_1} v_n \\ \vdots \\ f_{x_1 x_n} v_1 + \cdots + f_{x_n x_n} v_n \end{pmatrix} \cdot \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} \\
 &= \underbrace{\begin{pmatrix} f_{x_1 x_1} & \cdots & f_{x_n x_1} \\ \vdots & & \vdots \\ f_{x_1 x_n} & \cdots & f_{x_n x_n} \end{pmatrix}}_{=Hf^T} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \cdot \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} \\
 &= v^T H f \cdot w = \langle v, H f w \rangle
 \end{aligned}$$

Wir haben also insgesamt eine bestechend übersichtliche Form der Richtungsableitungen. Die erste wird über den Gradienten geliefert und die zweite über die Hessematrix:

Formel 11: erste und zweite Richtungsableitung: Es sei $f, : \mathbb{R}^n \rightarrow \mathbb{R}, f \in C^2(\mathbb{R}^n)$ und $v, w \in \mathbb{R}^n$ mit $\|v\| = \|w\| = 1$. Dann gilt

$$\begin{aligned}
 f_v(x) &= \nabla f(x) \cdot v \\
 f_{vw}(x) &= \langle v, H f(x) w \rangle
 \end{aligned}$$



Beachten Sie, dass die Richtungen der Richtungsableitung in Formel 11 konstante Richtungen sind, die nicht von $x \in \mathbb{R}^n$ abhängen. Wir werden später (Kapitel 3.4) auch Ableitungen in Richtung ortsabhängiger Richtungen berechnen.

Beispiel 72

Die Funktion (13) in Beispiel 71 leiten wir einmal in Richtung $v = (1, -2) \frac{1}{\sqrt{5}}$ und ein zweites Mal in Richtung $w = (-2, -1) \frac{1}{\sqrt{5}}$ ab:

$$\begin{aligned}
 f_{vw} &= \langle v, Hf w \rangle \\
 &= (1, -2) \frac{1}{\sqrt{5}} \begin{pmatrix} 4 & 2y \\ 2y & 2x - 2y \end{pmatrix} \begin{pmatrix} -2 \\ -1 \end{pmatrix} \frac{1}{\sqrt{5}} \\
 &= \frac{1}{5} (1, -2) \begin{pmatrix} -8 - 2y \\ -4y - 2x + 2y \end{pmatrix} \\
 &= \frac{1}{5} (-8 + 2y + 4x)
 \end{aligned}$$

Beispiel 73 Es sei die Funktion

$$f(x) = x_1^3 - 18x_1^2 + 81x_1 + 12x_2^2 - 144x_2 + 24x_1x_2 \quad (14)$$

definiert auf dem Dreieck

$$D = \{x \in \mathbb{R}^2 \mid x_1 \geq 0, x_2 \geq 0, x_1 + x_2 \leq 10\}.$$

Wir suchen Extremwerte von f auf D . Kandidaten sind nun nicht nur stationäre Punkte im Inneren des Dreiecks $\overset{\circ}{D}$, sondern auch stationäre Punkte auf dem Rand ∂D des Dreiecks und Werte in den Eckpunkten des Dreiecks.

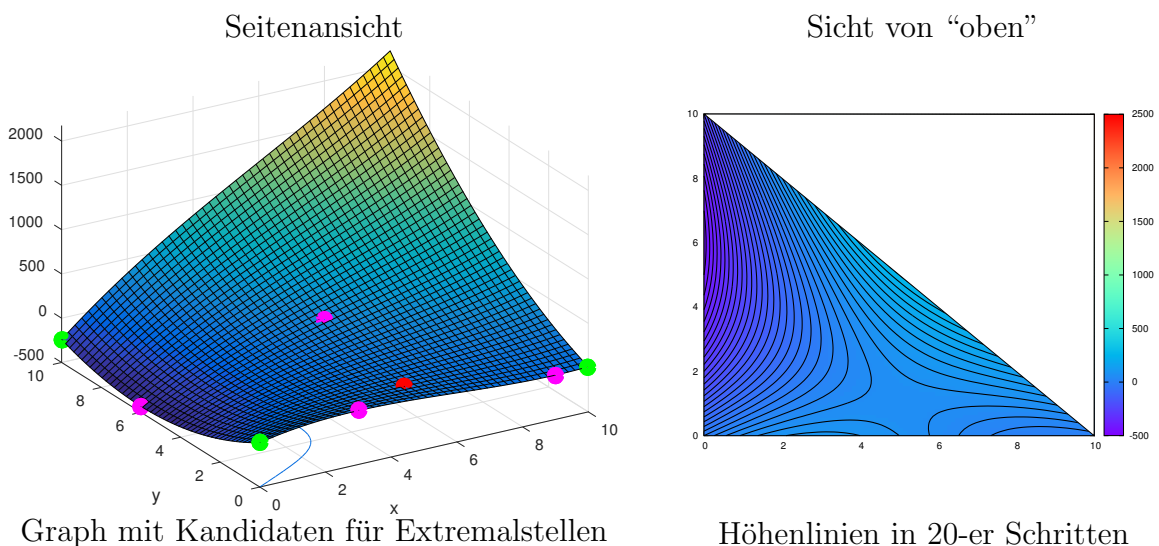


Abbildung 7: Graph der Funktion (14)

Wir gehen die einzelnen Punkte nacheinander durch:

1. Im Innern $\overset{\circ}{D}$:

$$\nabla f = \begin{pmatrix} 3x_1^2 - 36x_1 + 81 + 24x_2 \\ 24x_2 - 144 + 24x_1 \end{pmatrix} = 0 \quad \Leftrightarrow \quad x = \begin{pmatrix} 5 \\ 1 \end{pmatrix} =: K_0$$

Die Hessematrix von f lautet

$$Hf(x) = \begin{pmatrix} 6x_1 - 36 & 24 \\ 24 & 24 \end{pmatrix}$$

und man sieht direkt, dass

$$Hf(K_0) = \begin{pmatrix} -6 & 24 \\ 24 & 24 \end{pmatrix}$$

eine indefinite Matrix ist. K_0 ist also kein Extremum. Schade eigentlich.

2. Auf den Seitenkanten:

Es sei $\partial D = S_1 \cup S_2 \cup S_3$ mit

$$S_1 = \{x \in \partial D \mid 0 \leq x_1 \leq 10, x_2 = 0\}$$

$$S_2 = \{x \in \partial D \mid x_1 = 0, 0 \leq x_2 \leq 10\}$$

$$S_3 = \{x \in \partial D \mid 0 \leq x_1, x_2 \leq 10, x_1 + x_2 = 10\}$$

Wir haben nun zwei Möglichkeiten: Entweder wir schränken die Funktion jeweils auf eine Kante ein und erhalten dadurch eine neue Funktion, die nur noch von einer Variablen abhängt

$$f_{v_1} = \nabla f \cdot v_1 = \begin{pmatrix} 3x_1^2 - 36x_1 + 81 + 24x_2 \\ 24x_2 - 144 + 24x_1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 3x_1^2 - 36x_1 + 81 + 24x_2.$$

Diese Ableitung ist aber ja noch auf ganz D gegeben. Wir interessieren uns nur für Punkte auf S_1 . Es gilt also $x_2 = 0$ und damit erhalten wir

$$f_{v_1}|_{S_1} = 3x_1^2 - 36x_1 + 81.$$

Wir erhalten unüberraschenderweise das gleiche Ergebnis wie vorher, also auch die gleichen Kandidaten K_1 und K_2 . Um auf Extrema zu untersuchen - und zwar nur entlang der Kante S_1 - berechnen wir nun die zweite Ableitung in Richtung v_1 , also

$$f_{v_1 v_1}(x) = \nabla f_{v_1}(x) \cdot v_1 = f_{v_1, x_1} = 6x_1 - 36.$$

Es ist dann

$$f_{v_1 v_1}(K_1) < 0 \quad \text{und} \quad f_{v_1 v_1}(K_2) > 0$$

und damit ist K_1 ein Maximum und K_2 ein Minimum aus S_1 , aber zunächst nur dort. Wenn bei einem Maximum/Minimum auf der Kante die Funktion in Richtung Inneres fällt/steigt, dann ist es tatsächlich ein lokales Maximum/Minimum, sonst nicht. In der Graphik, links in der Abbildung 7 kann man die Situationen an den beiden Kandidaten K_1 und K_2 gut erkennen. K_1 sollte rein anschaulich ein lokales Maximum sein und K_2 ein lokales Minimum. Schauen wir mal. Wir betrachten die Richtungsableitung von f jeweils in den Punkten K_1 und K_2 in Richtung Gebietsinneres. Der entsprechend richtungsweisende, normierte (!) Vektor ist

$$v_1^\perp = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Es ist dann

$$f_{v_1^\perp} = \nabla f \cdot v_1^\perp = 24x_2 - 144 + 24x_1.$$

Und demnach ist

$$f_{v_1^\perp}(K_1) = 24 \cdot 0 - 144 + 24 \cdot 3 < 0$$

und

$$f_{v_1^\perp}(K_2) = 24 \cdot 0 - 144 + 24 \cdot 9 > 0$$

und damit ist erwiesenermaßen K_1 ein lokales Maximum und K_2 lokales Minimum auf $\overline{D}$.

So verfahren wir weiter und überprüfen, ob die Funktion f ausgehend von Extrema auf den Kanten, sich in's Innere entsprechend steigend oder fallend fortsetzt.

$$\begin{aligned} f_{v_2}|_{S_2} &= \nabla f|_{S_2} \cdot v_2 & \text{mit } v_2 &= \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ &= (24x_2 - 144 + 24x_1)|_{S_2, \text{ d.h. } x_1=0} \\ &= 24x_2 - 144 = 0 \\ \Leftrightarrow \quad x_2 &= 6 \end{aligned}$$

Damit ist

$$K_3 = \begin{pmatrix} 0 \\ 6 \end{pmatrix}$$

ein weiterer Kandidat. Wir prüfen das. Zunächst auf S_2

$$\begin{aligned} f_{v_2 v_2}|_{S_2} &= \nabla f_{v_2} \cdot v_2|_{S_2} \\ &= \begin{pmatrix} 24 \\ 24 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix}|_{S_2} = 24 \end{aligned}$$

Damit ist K_3 ein lokales Minimum auf S_2 . Was passiert in Richtung Inneres von D , also von S_2 aus in Richtung

$$v_2^\perp = \begin{pmatrix} 1 \\ 0 \end{pmatrix} ?$$

$$\nabla f \cdot v_2^\perp(K_3) = \begin{pmatrix} 3x_1^2 - 36x_1 + 81 + 24x_2 \\ 24x_2 - 144 + 24x_1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} \Big|_{x=K_3} = 81 + 24 \cdot 6 > 0$$

Ein lokales Minimum auf der Kante, bei dem die Funktion zum Innern hin steigt ist dann auch ein lokales Minimum auf $\overline{D}$.

Nun noch Kante S_3 . Wir untersuchen die Richtungsableitung von f in Richtung

$$v_3 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

$$\begin{aligned} f_{v_3}|_{S_3} &= \nabla f \cdot v_3|_{S_3} \\ &= \frac{1}{\sqrt{2}} \begin{pmatrix} 3x_1^2 - 36x_1 + 81 + 24x_2 \\ 24x_2 - 144 + 24x_1 \end{pmatrix} \Big|_{S_3, \text{ d.h. } x_2=10-x_1} \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix} \\ &= \frac{1}{\sqrt{2}} \begin{pmatrix} 3x_1^2 - 36x_1 + 81 + 24(10-x_1) \\ 24(10-x_1) - 144 + 24x_1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix} \\ &= \frac{3}{\sqrt{2}} (x_1^2 - 20x_1 + 75) = 0 \\ \Leftrightarrow \quad x_1 &\in \{5, 15\}, \end{aligned}$$

wobei Punkte mit $x_1 = 15$ nicht im Definitionsgebiet enthalten sind; dieser Wert also ignoriert werden kann. Wir erhalten also einen neuen Kandidaten

$$K_4 = (5, 5).$$

Es ist nun wieder durch

$$\begin{aligned} f_{v_3 v_3} &= \nabla f_{v_3} \cdot v_3 = \frac{1}{\sqrt{2}} \begin{pmatrix} 6x_1 - 60 \\ 0 \end{pmatrix} \cdot v_3 = \frac{1}{2} (6x_1 - 60) \\ \Rightarrow \quad f_{v_3 v_3}(K_4) &= \frac{1}{2} (6 \cdot 5 - 60) < 0 \end{aligned}$$

K_4 ist also auf S_3 ein lokales Maximum. Wie sieht's Richtung "Landesinnere", also Richtung

$$v_3^\perp = \begin{pmatrix} -1 \\ -1 \end{pmatrix} \frac{1}{\sqrt{2}}$$

aus?

$$f_{v_3^\perp}(K_4) = \nabla f \cdot v_3^\perp|_{x=K_4} = \frac{-1}{\sqrt{2}} (3x_1^2 - 12x_1 + 48x_2 - 63)|_{x=K_4} \frac{-192}{\sqrt{2}} < 0$$

Wenn die Funktion, die am Rand ein lokales Maximum hat in Richtung Inneres abfällt, dann handelt es sich auch um ein lokales Maximum auf $\overline{D}$.

3. Ecken eines Gebietes sind grundsätzlich mal Kandidaten und müssen untersucht werden. Wir haben also

$$K_5 := \begin{pmatrix} 0 \\ 0 \end{pmatrix}, K_6 := \begin{pmatrix} 10 \\ 0 \end{pmatrix}, K_7 := \begin{pmatrix} 0 \\ 10 \end{pmatrix}.$$

Schauen wir nach der Ecke K_5 : Wenn

$$\nabla f \cdot \begin{pmatrix} -1 \\ 0 \end{pmatrix} \quad \text{und} \quad \nabla f \cdot \begin{pmatrix} 0 \\ -1 \end{pmatrix}$$

im Punkt K_5 gleiches Vorzeichen haben, dann handelt es sich um ein Extremum, andernfalls nicht. D'accord? Prüfen wir das:

$$\nabla f(K_5) \cdot \begin{pmatrix} -1 \\ 0 \end{pmatrix} = -81 < 0 \quad \text{und} \quad \nabla f(K_5) \cdot \begin{pmatrix} 0 \\ -1 \end{pmatrix} = 144 > 0$$

belegt, dass es sich bei K_5 nicht um ein Extremum handelt. Nächste Ecke:

$$\nabla f(K_6) \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 300 - 360 + 81 + 24 > 0$$

und

$$\nabla f(K_6) \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix} \frac{1}{\sqrt{2}} = (300 - 360 + 81 + 144 - 240) \frac{1}{\sqrt{2}} < 0$$

zeigt, dass es sich bei K_6 auch nicht um ein Extremum handelt. Letzte Ecke: Es ist

$$\nabla f(K_7) \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} = 240 - 144 > 0$$

und

$$\nabla f(K_7) \cdot \begin{pmatrix} -1 \\ 1 \end{pmatrix} \frac{1}{\sqrt{2}} = (-81 - 240 + 240) \frac{1}{\sqrt{2}} < 0.$$

Demnach liegt an keiner der drei Ecken ein lokales Extremum vor.

Zusammengefasst gibt es lokale Maxima bei K_1 und K_4 und lokale Minima bei K_2 und K_3 . Stellen wir die Funktionswerte gegenüber so erhalten wir mit

$$f(K_1) = 108, f(K_4) = 260, f(K_2) = 0, f(K_3) = -432$$

ein globales Maximum bei $K_4 = (5, 5)$ und ein globales Minimum bei $K_3 = (0, 6)$.

puh.

Bemerkung 3.19. Wenn bei der Prüfung eines Kandidaten Unklarheit bleibt so kann man immer die Erfüllung von Definition ?? auch auf einem direkteren, aber nicht immer bequemeren Weg prüfen: Bei der Funktion $f(x, y) = 3xy - x^3 - y^3$ ist

$$f(1, 1) = 1$$

und jeder kleinste ϵ -Abstand vom Punkt $(1, 1)$ führt auf

$$f(1 \pm \epsilon, 1 \pm \epsilon) = 3(\pm\epsilon + 1)^2 - 2(\pm\epsilon + 1)^3 < 1,$$

womit f bei $(1, 1)$ ein lokales Maximum hat.

3.3 Optimierung mit Nebenbedingung

Betrachten wir einmal eine Menge Punkte, die so gestaltet ist, dass die Funktionswerte an diesen Punkten alle gleich sind. Oder anders formuliert: Wir machen einen Schnitt vom Graphen f mit einer horizontalen Ebene und fassen alle betroffenen Urbildpunkte zu einer Menge M zusammen:

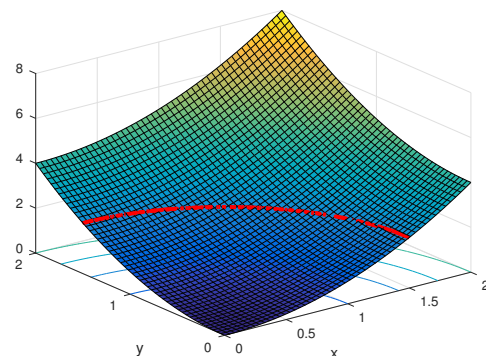
$$M_c := \{x \in \mathbb{R}^2 \mid f(x) = c\} \quad \text{Levelset}$$

Entlang einer Höhenlinie ist die Steigung Null. Wenn also $\tau = \tau(x)$ in Richtung des Höhenlinienverlaufs zeigt, dann gilt

$$f_\tau(x) = 0, \quad x \in M.$$

Gleichzeitig gilt auch

$$f_\tau = \nabla f \cdot \tau.$$



Höhenlinie eines Graphen

Wenn dieses Skalarprodukt verschwindet, und zwar ohne, dass ∇f selbst verschwindet, wir uns also nicht an einem stationären Punkt befinden, folgt doch, dass

$$\nabla f \perp \tau$$

ist. Es ist dann also der steilste Anstieg bei

$$f_{\tau^\perp} \parallel \nabla f$$

zu suchen. Diese kleine Erkenntnis führt uns direkt zum nächsten Thema.

Wir betrachten zwei Funktionen $f, g, : \mathbb{R}^2 \Rightarrow \mathbb{R}$ und suchen ein Extremum von f unter der Bedingung g . Anders formuliert: Wir suchen ein $(x, y) \in \mathbb{R}^2$, für das f extremal ist und für das die Nebenbedingung $g(x, y) = c$ erfüllt ist.

Veranschaulichen kann man sich das auf folgende Weise: Wir markieren alle $(x, y) \in \mathbb{R}^2$, für die die Funktion g einen bestimmten konstanten Wert $c \in \mathbb{R}$ annimmt:

$$M_c = \{(x, y) \in \mathbb{R}^2 \mid g(x, y) = c\}$$

Dann betrachten wir die Kurve, die wir erhalten, wenn wir f nur an Werten aus M_c anwenden. In Abbildung 8 beschreibt die rote Kurve in der linken Graphik gerade die Menge M_c , also diejenigen Punkte, die die Nebenbedingung erfüllen. Rechts beschreibt die rote Kurve das Bild von M_c unter f , also $f(M_c)$. Was man also sucht ist

$$\max_{(x,y) \in M_c} f(x, y).$$

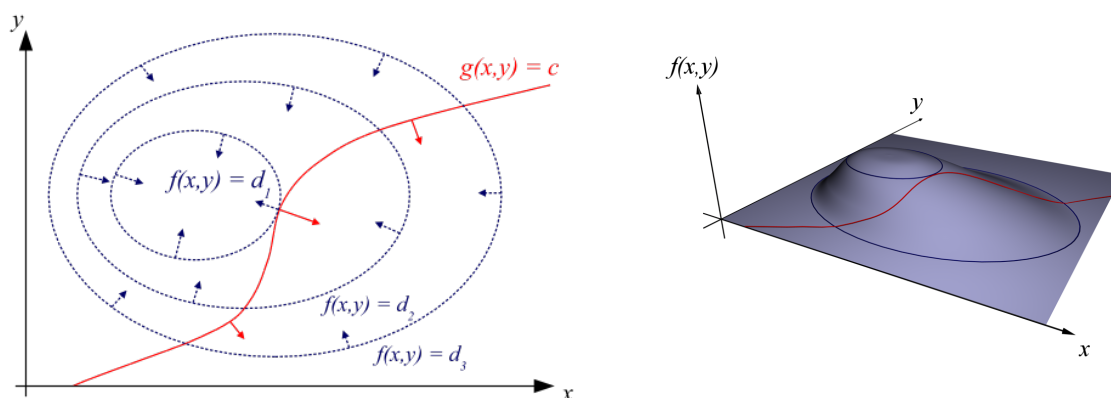


Abbildung 8: Extrema von f mit Nebenbedingung g . (Quelle: [Wikipedia:Lagrange-Multiplikator](#))

Wir betrachten nun die Höhenlinie von f , die die Linie $f(M_c)$ am Extremum $(\bar{x}, \bar{y})$ berührt. An dieser Stelle sind die Gradienten von f und g jeweils richtungsgleich zueinander. Warum ist das so? Die Ableitungen in Richtung der Höhenlinien sind jeweils Null:

$$f_v = 0 \quad \text{und} \quad g_w = 0$$

Da die Höhenlinien sich am Punkt $(\bar{x}, \bar{y})$ berühren, haben Sie an dieser Stelle die gleiche Richtung:

$$v(\bar{x}, \bar{y}) = w(\bar{x}, \bar{y})$$

Der Gradient steht jeweils senkrecht auf die Richtung der Höhenlinie und damit gilt

$$\nabla f(\bar{x}, \bar{y}) \parallel \nabla g(\bar{x}, \bar{y}).$$

Wir suchen also Punkte, für die es ein $\lambda \in \mathbb{R} \setminus \{0\}$ gibt mit

$$\nabla f = -\lambda \nabla g. \quad (15)$$

Definition 3.20 (Lagrange-Funktion). Für Funktionen $f, g : \mathbb{R}^2 \rightarrow \mathbb{R}$ heißt die Funktion $\Lambda : \mathbb{R}^2 \rightarrow \mathbb{R}$ definiert als

$$\Lambda(x, y, \lambda) := f(x, y) + \lambda (g(x, y) - c)$$

Lagrange-Funktion und $\lambda \in \mathbb{R} \setminus \{0\}$ heißt Lagrange-Multiplikator.

Stationäre Punkte mit $\nabla g \neq 0$ der Lagrange-Funktion erfüllen gerade die Gleichung (15). Sehen Sie das?

$$\nabla \Lambda = \begin{pmatrix} \Lambda_x \\ \Lambda_y \\ \Lambda_\lambda \end{pmatrix} = \begin{pmatrix} f_x(x, y) + \lambda g_x(x, y) \\ f_y(x, y) + \lambda g_y(x, y) \\ g(x, y) - c \end{pmatrix} = 0 \Leftrightarrow \nabla f = -\lambda \nabla g \wedge g(x, y) = c$$

Dabei muss allerdings zunächst $\nabla g \neq 0$ gelten.

Beispiel 74 " $\nabla g \neq 0$ " Wir suchen Extrema von $f(x, y) = x + y$, für die $x^2 + y^2 = 1$ ist.

Es ist also die Nebenbedingung

$$g(x, y) = x^2 + y^2$$

und $c = 1$. Der Gradient Lagrange-Funktion

$$\Lambda(x, y, \lambda) = x + y + \lambda (x^2 + y^2 - 1)$$

lautet

$$\nabla \Lambda = \begin{pmatrix} 1 + 2x\lambda \\ 1 + 2y\lambda \\ x^2 + y^2 - 1 \end{pmatrix} = 0 \Leftrightarrow x = y = -\frac{1}{2\lambda} \wedge \lambda = \pm \frac{1}{\sqrt{2}}$$

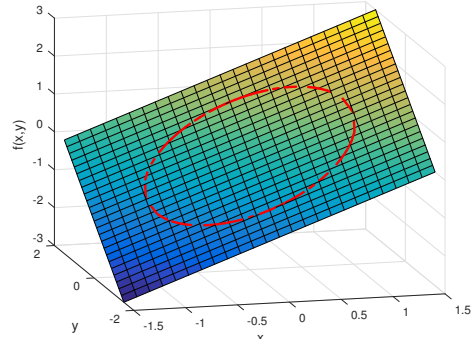
Die zu optimierende Funktion f hat unter der Nebenbedingung $x^2 + y^2 = 1$ an den Stellen

$$\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right) \quad \text{und} \quad \left(\frac{-1}{\sqrt{2}}, \frac{-1}{\sqrt{2}}\right)$$

die Werte $\sqrt{2}$ und $-\sqrt{2}$.

Was aber tun, wenn der Gradient der Nebenbedingung dann per Zufall an der gesuchten Stelle doch verschwindet?

Beispiel 75 " $\nabla g = 0$ "

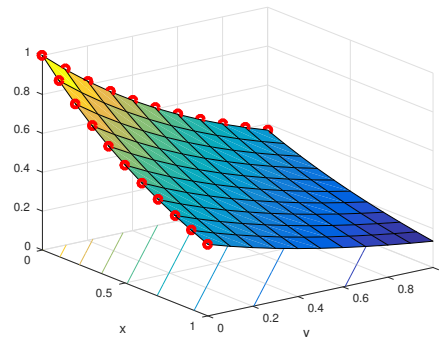


f sei die zu optimierende Funktion:

$$f(x) = e^{-(x_1+x_2)} \quad f : \mathbb{R}_0^+ \times \mathbb{R}_0^+ \rightarrow \mathbb{R}$$

und g beschreibe die Nebenbedingung:

$$g(x) = x_1 x_2 = 0 \quad g : \mathbb{R}_0^+ \times \mathbb{R}_0^+ \rightarrow \mathbb{R}$$



Die Lagrange-Funktion

$$\Lambda(x, \lambda) = \begin{pmatrix} \lambda x_2 - e^{-\|x\|_1} \\ \lambda x_1 - e^{-\|x\|_1} \\ x_1 x_2 \end{pmatrix}$$

liefert keine kritischen Punkte. Wir schauen noch an welchen Stellen der Gradient der Nebenbedingung verschwindet:

$$\nabla g = 0 \quad \Leftrightarrow \quad x = 0$$

Das ist ein Randpunkt, der die Nebenbedingung erfüllt. Der Rand und auch die Ecken müssen bei der Suche nach Extrema separat betrachtet werden. In diesem Fall muss man das eben noch tun.

$$f(0) = e^0 = 1 \quad \text{und} \quad f(x) = e^{-\|x\|_1} < 1 \quad \forall x \neq 0.$$

Der Ursprung ist also tatsächlich ein lokales Maximum.

Bei mehreren Nebenbedingungen erweitern wir einfach um entsprechend viele Lagrange-Multiplikatoren.

Formel 12: Optimierung mit Nebenbedingungen: Es $f : \mathbb{R}^n \rightarrow \mathbb{R}$ eine zu optimierende Funktion unter den s Nebenbedingungen $g_k(x) = 0, k = 1, \dots, s$. Wir suchen dann stationäre Punkte der Lagrange-Funktion $\Lambda : \mathbb{R}^n \times \mathbb{R}^s \rightarrow \mathbb{R}$

$$\Lambda(x, \lambda) = f(x) + \sum_{k=1}^L \lambda_k g_k(x)$$

mit $x = (x_1, \dots, x_n), \lambda = (\lambda_1, \dots, \lambda_s)$.



Wir suchen nur stationäre Punkte von Λ , keine Extrema! Wir müssen daher auch nicht nach der Definitheit der Hessematrix Ausschau halten. Ob es sich bei den stationären Punkte um speziellen Fall dann tatsächlich um Maxima oder Minima handelt, muss man von Fall zu Fall prüfen.

Beispiel 76

Beispiele mit nichtlinearen Ausdrücken werden dann aber auch schnell zu kompliziert, um sie analytisch zu lösen. Sie können sich gerne mal am folgenden Beispiel versuchen: Gesucht ist ein Extremum von f unter den Nebenbedingungen g_1 und g_2 mit:

$$f(x, y) = x^2 - y^2 x, \quad g_1(x, y) = x^2 + y^2 - 1, \quad g_2(x, y) = x - 2y + 1$$

Die Lagrange-Funktion und deren Gradient lautet dann:

$$\Rightarrow \quad \Lambda(x, y, \lambda_1, \lambda_2) = x^2 - y^2 x + \lambda_1 (x^2 + y^2 - 1) + \lambda_2 (x - 2y + 1)$$

$$\Rightarrow \quad \nabla \Lambda = \begin{pmatrix} 2x - y^2 + 2x\lambda_1 + \lambda_2 \\ -2yx + 2y\lambda_1 - 2\lambda_2 \\ x^2 + y^2 - 1 \\ x - 2y + 1 \end{pmatrix}$$

Nullstelle des Gradienten mit Newton berechnet:

$$\nabla \Lambda = 0 \quad \Leftrightarrow \quad (x, y) = (0.6, 0.8), \quad \lambda = (-0.04, -0.512)$$

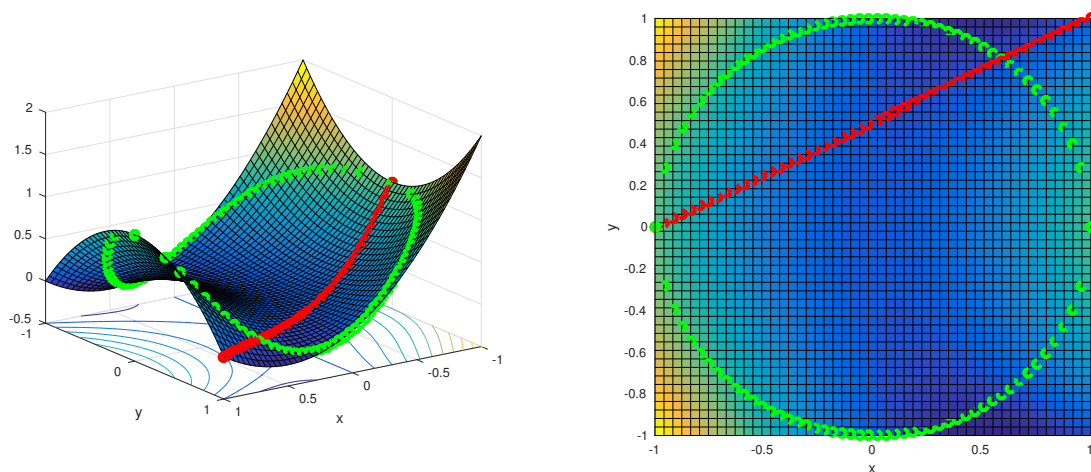


Abbildung 9: Funktion mit Nebenbedingung

3.4 weitere Ableitungsbegriffe und ausgewählte Anwendungsbeispiele

Wir betrachten zunächst nicht nur Funktionen mehrerer Veränderlicher sondern vektorwertige Funktionen mehrerer Veränderlicher, nämlich Funktionen der Form

$$f : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$$

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto \begin{pmatrix} f_1(x_1, \dots, x_n) \\ \vdots \\ f_m(x_1, \dots, x_n) \end{pmatrix}$$

Dazu gleich mal ein

Beispiel 77

$$f(x_1, x_2) = \begin{pmatrix} x_1 \cdot x_2 \\ x_1 \cos x_2 + x_2 \cos x_1 \end{pmatrix}$$

Eine Sorte erste Ableitung ist der Gradient

$$\text{grad} f = \begin{pmatrix} \text{grad} f_1 \\ \vdots \\ \text{grad} f_m \end{pmatrix},$$

der einfach auf jede Komponente angewandt wird. Wir erhalten damit die sogenannte

Definition 3.21 (Jacobi-Matrix). Sei $f : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$ mit $f \in C^1(D)^m$, dann heißt

$$Df(x) = \begin{pmatrix} \frac{\partial}{\partial x_1} f_1(x) & \cdots & \frac{\partial}{\partial x_n} f_1(x) \\ \vdots & \ddots & \vdots \\ \frac{\partial}{\partial x_1} f_m(x) & \cdots & \frac{\partial}{\partial x_n} f_m(x) \end{pmatrix}, \quad x = (x_1, \dots, x_n)$$

Jacobi-Matrix von f . Es ist

$$Df_{ij}(x) = \frac{\partial}{\partial x_j} f_i(x).$$

Da es sich ja im Grunde um einen Gradienten von f handelt schreibt man auch

$$\nabla f, \text{ grad} f$$

stattdessen.

Beispiel 78 Jacobi-Matrix Für

$$f(x, y) = \begin{pmatrix} x^2 + y \\ x y \end{pmatrix}$$

ist die Jacobi-Matrix

$$Df(x, y) = \begin{pmatrix} 2x & 1 \\ y & x \end{pmatrix}.$$

Bei der vektorwertigen Funktion $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$, $n \in \{2, 3\}$ könnte es sich um ein sogenanntes Vektorfeld handeln. Vektorfeldern im Speziellen werden Sie noch beim Thema Vektoranalysis behandeln. Wir werfen hier schon mal einen Blick darauf, als Anwendungsbeispiel zweier weiterer Ableitungsbegriffe, die interessante Eigenschaften von f aufzeigen. In Worten gesprochen ist ein Vektorfeld zu f eine graphische Darstellung der Abbildung. Siehe dazu Abbildung 10. Es zeigt in jedem Punkt der Ebene ($n = 2$) oder auch im Raum ($n = 3$) den Wert der Funktion, die dann bei (x, y) ausgewertet als Richtungsvektor dargestellt werden kann. Anders als bei Richtungsfeldern von ODEs in Kapitel 1.1 spielt bei den Vektorfeldern die Länge der Vektoren eine Rolle. Sie beschreiben an jedem Ort den jeweiligen Betrag des Funktionswertes.

Neben der Jacobi-Matrix gibt es als Ableitung erster Ordnung für vektorwertige Funktionen noch die Divergenz.

Definition 3.22 (Divergenz). Sei $f : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^n$. Dann heißt

$$\nabla \cdot f = \sum_{i=1}^n \frac{\partial}{\partial x_i} f_i$$

Divergenz von f ; auch $\operatorname{div} f$. Der Operator $\nabla \cdot$ für sich heißt Divergenzoperator.



Der Punkt in $\nabla \cdot f$ ist ein wesentlicher Bestandteil der Definition. Bedenken Sie, dass mit ∇f durchaus die Jacobi-Matrix gemeint sein kann. Zudem handelt es sich bei der Definition der Divergenz auch tatsächlich um ein Skalarprodukt, nämlich in folgendem Sinne:

$$\nabla \cdot f = \begin{pmatrix} \frac{\partial}{\partial x_1} \\ \vdots \\ \frac{\partial}{\partial x_n} \end{pmatrix} \cdot \begin{pmatrix} f_1 \\ \vdots \\ f_n \end{pmatrix} = \frac{\partial}{\partial x_1} f_1 + \cdots + \frac{\partial}{\partial x_n} f_n$$

Bemerkung 3.23. Die Divergenz einer vektorwertigen Funktion ist gerade die Spur der Jacobi-Matrix:

$$\nabla \cdot f = \operatorname{Spur}(Df)$$

Beispiel 79 Divergenz

(a) Die Divergenz der Funktion aus Beispiel 78 lautet

$$\nabla \cdot f = \begin{pmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \end{pmatrix} \cdot \begin{pmatrix} f_1 \\ f_2 \end{pmatrix} = \begin{pmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \end{pmatrix} \cdot \begin{pmatrix} x^2 + y \\ x y \end{pmatrix} = f_{1,x} + f_{2,y} = 2x + x = 3x$$

(b)

$$f(x, y) = \begin{pmatrix} x^2 y \\ 3 - \sqrt{y} \end{pmatrix} \Rightarrow \nabla \cdot f = f_x^1 + f_y^2 = 2xy - \frac{1}{2\sqrt{y}}$$

Für den Divergenzoperator gibt es eine schöne Anschauung. Da Vektorfelder in der "Anschauung" meist Strömungen, etwa von irgendwelchen Flüssigkeiten beschreiben, wird die Divergenz in diesen Fällen oft als "Quellendichte" interpretiert. Ist die Divergenz in einem Punkt nämlich positiv, so fließt die Flüssigkeit in der Umgebung des Punktes heraus (Divergenz). Im negativen Fall fließt die Flüssigkeit folglich in die Umgebung des Punktes hinein (Konvergenz).

Beispiel 80

(a)

$$f(x, y) = \begin{pmatrix} x \\ y \end{pmatrix} \Rightarrow \nabla \cdot f = 2$$

(b)

$$f(x, y) = \begin{pmatrix} -x \\ -y \end{pmatrix} \Rightarrow \nabla \cdot f = -2$$

(s. Abb. 10)

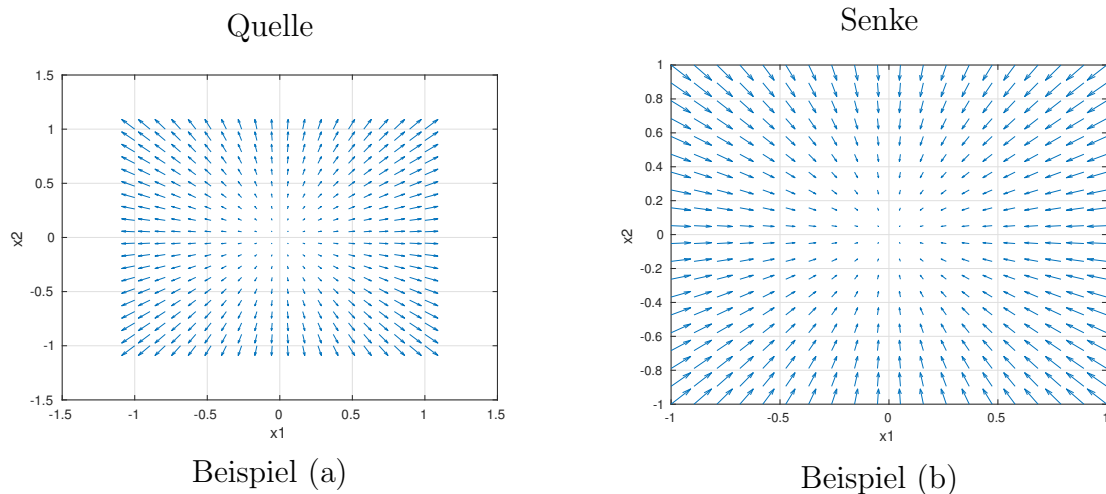


Abbildung 10: Vektorfelder aus Beispiel 80

Beispiel 81

$$f(x, y) = \begin{pmatrix} 2x^2 y \\ 3x - y \end{pmatrix} \Rightarrow \nabla \cdot f = 4xy - 1$$

(s. Abb. 11)

Eine weitere interessante Ableitung für Vektorfelder ist die

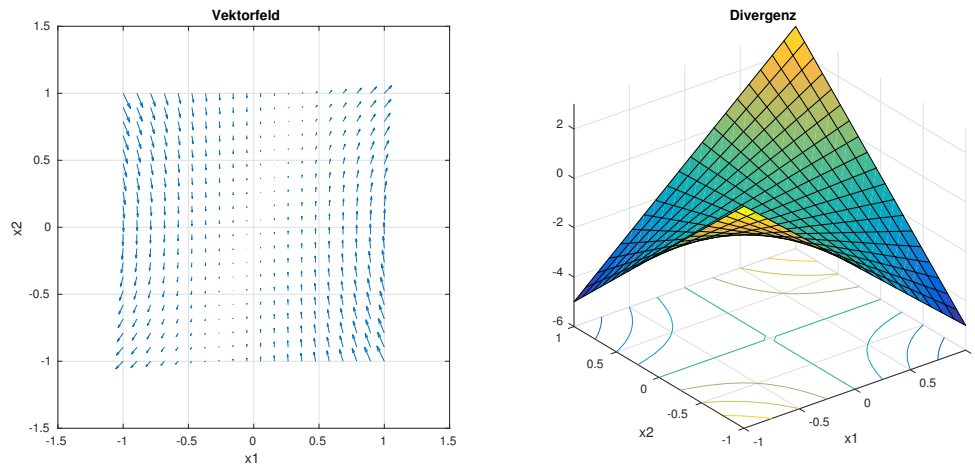


Abbildung 11: links: Vektorfeld u , rechts: Divergenz des Vektorfelds $\nabla \cdot u$ zu Beispiel 81

Definition 3.24 (Rotation). Sei $f : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^n$, $n \in \{2, 3\}$ ein Vektorfeld. Dann heißt

$$\nabla \times f = \begin{pmatrix} \frac{\partial}{\partial x_1} \\ \vdots \\ \frac{\partial}{\partial x_n} \end{pmatrix} \times \begin{pmatrix} f_1 \\ \vdots \\ f_n \end{pmatrix}, \quad n = 2, 3$$

Rotation von f ; auch $\text{rot } f$. Der Operator $\nabla \times$ heißt Rotationsoperator

Beispiel 82

(a) Für Beispiel 81 ist die Rotation

$$\nabla \times f = \begin{pmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \end{pmatrix} \times \begin{pmatrix} 2x^2y \\ 3x - y \end{pmatrix} = (3x - y)_x - (2x^2y)_y = 3 - 2x^2.$$

Abbildung 12 zeigt, wie ein Würfel in der Mitte der Strömung stärker rotiert würde als am Rand.

(b) Im $\mathbb{R}^3$:

$$\begin{pmatrix} \frac{\partial}{\partial x_1} \\ \frac{\partial}{\partial x_2} \\ \frac{\partial}{\partial x_3} \end{pmatrix} \times \begin{pmatrix} x_1 x_2 x_3 \\ x_1 + x_2 + x_3 \\ x_2^2 \end{pmatrix} = \begin{pmatrix} (x_2^2)_{x_2} - (x_1 + x_2 + x_3)_{x_3} \\ (x_1 x_2 x_3)_{x_3} - (x_2^2)_{x_1} \\ (x_1 + x_2 + x_3)_{x_1} - (x_1 x_2 x_3)_{x_2} \end{pmatrix} = \begin{pmatrix} 2x_2 - 1 \\ x_1 x_2 \\ 1 - x_1 x_3 \end{pmatrix}$$

Während bei der Rotation in der Ebene die Drehachse immer senkrecht zur Ebene steht ist hier nur noch die Drehgeschwindigkeit interessant. Wir erhalten hier eine Zahl. Die Rotation im Raum beinhaltet hingegen nicht nur eine Drehgeschwindigkeit (Betrag des Vektors) sondern auch die Richtung der Drehachse,

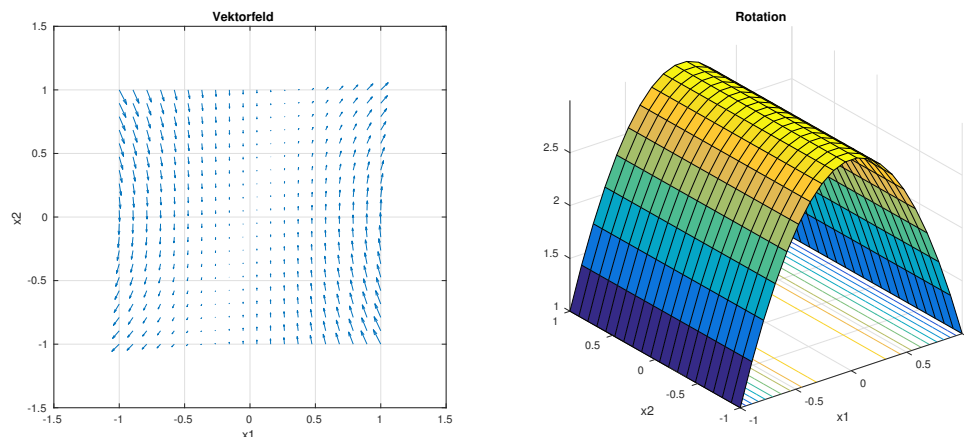


Abbildung 12: Rotation eines Vektorfeldes, Beispiel 82 (a), resp. 81

Wir schlagen nun eine Brücke zwischen Ableitungsausdrücken von vektorwertigen und skalarwertigen Funktionen. Zu einer skalarwertigen Funktion mehrerer Veränderlicher $u : \mathbb{R}^n \rightarrow \mathbb{R}$ ist ja der Gradient

$$\nabla u : \mathbb{R}^n \rightarrow \mathbb{R}^n$$

eine vektorwertige Funktion. Wenn wir nun auf dieses Vektorfeld die Divergenz anwenden passiert Folgendes:

$$\nabla \cdot \nabla u = \nabla \cdot \begin{pmatrix} u_{x_1} \\ \vdots \\ u_{x_n} \end{pmatrix} = u_{x_1 x_1} + \cdots + u_{x_n x_n} = \sum_{i=1}^n u_{x_i x_i}$$

Dies ist eine besondere Form der zweiten Ableitung, die demzufolge auch einen eigenen Namen trägt, nämlich:

Definition 3.25 (Laplace-Operator). Der Operator $\Delta : C^2(D \subseteq \mathbb{R}^n) \rightarrow \mathbb{R}$ mit

$$\Delta := \sum_{i=1}^n \frac{\partial^2}{\partial x_i^2}$$

heißt Laplace-Operator. Und zu einer skalaren Funktion $u : \mathbb{R}^n \rightarrow \mathbb{R}$ heißt die Abbildung

$$\Delta u = u_{x_1 x_1} + \cdots + u_{x_n x_n} = \sum_{i=1}^n u_{x_i x_i}$$

Laplace von u .

???

Beispiel 83 Laplace Sei

$$f : \mathbb{R}^2 \rightarrow \mathbb{R} \\ x \mapsto x_1^2 + \cos x_2$$

Dann gilt

$$\Delta f = 2 - \cos x_2 .$$

Bemerkung 3.26. *Der Laplace einer Funktion u ist die Spur der Hessematrix von u , denn es gilt*

$$\text{Spur} H u = \text{Spur} \begin{pmatrix} u_{x_1 x_1} & \cdots & u_{x_1 x_n} \\ \vdots & \ddots & \vdots \\ u_{x_n x_1} & \cdots & u_{x_n x_n} \end{pmatrix} = u_{x_1 x_1} + \cdots + u_{x_n x_n} = \Delta u .$$

Auf den Laplace-Operator kommen wir gleich noch einmal zurück. Mit ihm kann man viele Prozesse in der Natur beschreiben. Wir wollen uns das mit Anwendungsbeispielen noch genauer anschauen. Da aber Prozesse in der Natur meist zeitabhängige sind, wollen wir eine Variable in der Zeit noch als abhängige Größe mit aufnehmen.

Insbesondere in diesem Zusammenhang, aber nicht nur, sind wir an einem Ableitungsbegriff interessiert, der alle partiellen Ableitungen beinhaltet. Das sogenannte totale Differential erfüllt diesen Wunsch.

Definition 3.27 (totales Differential). *Sei $f \in C^k(D)$. Dann bezeichnet*

$$df = \sum_{i=1}^n \frac{\partial f}{\partial x_i} dx_i$$

das totale Differential von f .

Beispiel 84 totales Differential Das totale Differential der Funktion $f(x, y) = x \sin(y)$ ist gegeben durch

$$df = f_x dx + f_y dy = \sin y dx + x \cos y dy$$

Im Grunde ist das totale Differential ganz was Ähnliches wie der Gradient. Da stecken ja auch alle partiellen Ableitungen drin. So gesehen ist der Sinn dieser Definition erst mal nicht so einsichtig.

Richtig sinnvoll wird dieser Ausdruck, wenn wir eine Zeitabhängigkeit noch mit in's Spiel bringen. Am besten wird das am folgenden Beispiel deutlich:

Beispiel 85 aus der Fluidmechanik Es sei $f(x, y)$ die Temperatur des Wasserpartikels am Ort (x, y) . Das Wassermolekül bewegt sich im Laufe der Zeit, also gilt $(x, y) = (x(t), y(t))$. Ein mitbewegter Beobachter, der das Partikel (x, y) verfolgt und dabei die zeitliche Änderung der Temperatur misst nimmt Folgendes wahr:

$$\frac{d}{dt}f(x(t), y(t)) = \frac{\partial}{\partial x}f(x(t), y(t))\frac{d}{dt}x(t) + \frac{\partial}{\partial y}f(x(t), y(t))\frac{d}{dt}y(t)$$

Das führt dann auf

$$df = \frac{\partial}{\partial x}f dx + \frac{\partial}{\partial y}f dy.$$

In diesem Beispiel hängt f nur implizit von der Zeit ab, nämlich deshalb, weil die Partikelposition von der Zeit abhängt. Eine explizite Zeitabhängigkeit herrscht dann, wenn eine der Veränderlichen, also eines der Argumente von f eine Variable in der Zeit ist. So wie im folgenden Beispiel.

Beispiel 86

Im Fall von Beispiel 85 ändert sich die Temperatur an einer Stelle nicht. Eine zeitliche Änderungsrate gibt es nur deshalb, weil der Ort sich ändert. Herrscht auch eine Temperaturabhängigkeit von der Zeit an einem festen Ort so schreiben wir

$$f = f(t, x(t)), \quad x(t) = (x_1(t), x_2(t)).$$

Wir könnten es auch so formulieren:

$$f = f(t, x_1(t), x_2(t)).$$

Es meint das Gleiche.

In diesem Fall unterscheiden wir zwischen totaler Zeitableitung und partieller Zeitableitung. Nehmen wir einmal

$$f(t, x(t)) = t^3 + x(t)^2.$$

Dann ist die partielle Zeitableitung

$$\frac{\partial}{\partial t}f(t, x(t)) = 3t^2$$

die partielle Ortsableitung

$$\frac{\partial}{\partial x}f(t, x(t)) = 2x(t)$$

und die totale Zeitableitung (Kettenregel)

$$\begin{aligned} \frac{d}{dt}f(t, x(t)) &= \frac{\partial}{\partial t}f(t, x(t)) \cdot \frac{d}{dt}t + \frac{\partial}{\partial x}f(t, x(t)) \cdot \frac{d}{dt}x(t) \\ &= 3t^2 + 2x(t) \cdot x'(t) \end{aligned}$$



Es besteht also ein enormer Unterschied zwischen

$$\frac{\partial}{\partial t}f = f_t \quad \text{und} \quad \frac{d}{dt}f = \dot{f}!$$

Definition 3.28 (materielle/substantielle Ableitung). Es sei auf dem Gebiet $D \subseteq \mathbb{R}^n$ eine Funktion $f : [0, T] \times D \rightarrow \mathbb{R}$ und ein Vektorfeld $v \in \mathbb{R}^n$ gegeben. Dann heißt

$$\dot{f}(t, x) = f_t + v \cdot \nabla f$$

materielle oder auch substantielle Ableitung von f . f_t nennen wir auch lokale Änderung und $v \cdot \nabla f$ konvektive Änderung.

“Im mathematischen Sinne handelt es sich um die totale Ableitung einer Größe, die von der Zeit und dem Ort eines materiellen Punktes abhängt. Sie beschreibt nun die Änderung der Größe für einen Beobachter, der sich mit der “Substanz bzw. Materie” bewegt. Der lokale Anteil, $\frac{\partial f}{\partial t}$, gibt an wie sich die Größe an dem festen Ort $\vec{x}$, d.h. lokal, verändert. Der konvektive Anteil, $\vec{v} \cdot \nabla f$, beinhaltet die Änderung, welche sich zusätzlich durch die Bewegung des materiellen Punktes einstellt. Zusammen ergibt sich die vollständige substantielle Änderung für den betrachteten materiellen Punkt. Verwendet wird hierbei das Modell des mitbewegten Beobachters, welches auch als Lagrangesche Betrachtungsweise bekannt ist. Daneben existiert die Eulersche Betrachtungsweise, welche einen feststehenden Beobachter nutzt und mit der lokalen Änderung verknüpft ist (der konvektive Term fällt ohne Bewegung heraus). Genau genommen ist die Größe f selbst Euler’sch definiert, ihre materielle Ableitung hingegen Lagrange’sch.”⁵

Beispiel 87 Schoggi

Es sei $C : \mathbb{R}^3 \rightarrow \mathbb{R}$ die Konzentration von Schoggipulver in einem Milchglas. Es beschreibe das Vektorfeld $v \in \mathbb{R}^3$ die Bewegung der Milch. Die Konzentration C genügt - stark idealiert! - der Gleichung

$$C_t + \nabla C \cdot v = 0,$$

wobei wir davon ausgehen, dass kein Pulver von Außen dazukommt und auch nichts abgesogen wird.



Beispiel 88 Kontinuitätsgleichung

Wenn man von einem divergenzfreien Fluid ausgeht, dann gilt ja $\nabla \cdot v = 0$ und dann kann man über die materielle Ableitung direkt auf die sogenannte Kontinuitätsgleichung

$$C_t + \nabla \cdot (Cv) = f,$$

schließen. Bei der allgemeinen Herleitung stößt man auf diese Gleichung, wenn man Masseerhaltung fordert. Das aber nur am Rande bemerkt.

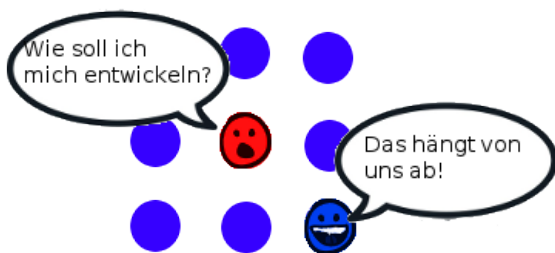
Wenn nun ein physikalisches Gesetz besagt, dass

$$f_t + \nabla f \cdot v = 0$$

⁵Physikalische Interpretation (Quelle: Wikipedia)

gilt, dann würde das ja im Falle, dass f eine Temperatur beschreibt zum Beispiel, bei windstillem Gebiet $v = 0$ bedeuten, dass es keine Temperaturänderung gibt. Wir wissen aber aus Erfahrung, dass sich verschiedene Temperaturen in einem Medium oder auch Schoggipulver in Milch bewegt, auch wenn man nicht rührt und auch wenn man sich vorstellt es gäbe die Erdanziehungskraft nicht ... oke, lassen wir das an dieser Stelle mit der eigenen Erfahrung eher weg :)

Die Diffusionsgleichung oder auch Wärmeleitungsgleichung beschreibt den Zusammenhang zwischen der zeitlichen Änderung und der räumlichen Änderung einer Konzentration an einem Ort in einem Körper und eignet sich zur Berechnung instationärer Temperaturfelder.

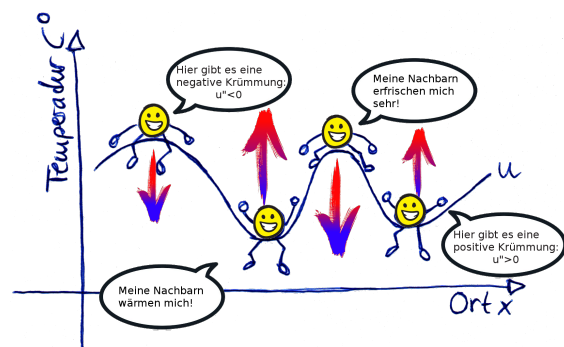


Die Wärmeleitungsgleichung ist eine partielle Differentialgleichung, die besagt, dass die zeitliche Entwicklung von Wärme an einem Ort, von der Wärmeverteilung der unmittelbaren Umgebung abhängt.

In einer Raumdimension kann man sich das so vorstellen. Die Wärmeentwicklung in der Zeit bezogen auf Zu- und Abnahme der Energie wird vom Krümmungsverhalten im Ort der Temperaturkurve bestimmt. Es gilt in 1d

$$u_t \sim u''$$

$$u_t \sim C \Delta u = C (u_{xx} + u_{yy}).$$



Definition 3.29 (Diffusionsgleichung). Es sei F ein Quellterm (etwa eine Hitzequelle, die von außen auf das System wirkt), $C \in \mathbb{R}$ der Diffusionskoeffizient. Dann heißt die Gleichung

$$u_t - \nabla \cdot (C \nabla u) = F$$

Diffusions- oder auch Wärmeleitungsgleichung.



Mathematisch sind Wärmeleitungsgleichung und Diffusionsgleichung identisch, statt Temperatur und Temperaturleitfähigkeit treten hier Konzentration und Diffusionskoeffizient auf.

Beispiel 89 Beispiel aus der Bildverarbeitung

Satz 3.30. Für $u \in C^2(D \subset \mathbb{R}^2)$ gilt

$$u_{\nu\nu} + u_{\tau\tau} = u_{xx} + u_{yy} = \Delta u$$

Beweis Satz 3.30:

$\ddot{U}A$

□

Die Diffusionsgleichung in Definition 3.29 beschreibt eine isotrope Diffusion. Das bedeutet, dass die Ausbreitung der Substanzen keine Vorzugsrichtungen haben. Wählen wir für den Laplace

$$\Delta u = u_{\nu\nu} + u_{\tau\tau}$$

und unterdrücken eine Richtung, sagen wir τ dann erhalten wir einen diffusiven Prozess, der eine Vorzugsrichtung hat. Wir sprechen dann von einer anisotropen Diffusion.

Beispiel 90 anisotrope Diffusion für Bilder

Wenn wir die Anisotropie so wählen, dass die Vorzugsrichtung ortsabhängig ist, haben wir direkt nochmal was zu rechnen. Bisher haben wir uns Richtungsableitungen zweiter Ordnung nur mit konstanten Richtungen angeschaut. Wie sieht das aber aus wenn wir an

$$u_{v(x)w(x)} = ?$$

interessiert sind? Rechnen wir das doch mal durch.

Es sei $u : \mathbb{R}^n \rightarrow \mathbb{R}$, $v, w : \mathbb{R}^n \rightarrow \mathbb{R}^n$ mit $\|v\| = 1$ und $\|w\| = 1$. Dann gilt

$$\begin{aligned} u_{v(x)}(x) &= \nabla u(x) \cdot v(x) \\ \Rightarrow u_{v(x)w(x)}(x) &= \nabla u_{v(x)}(x) \cdot w(x) \\ &= \sum_{i=1}^n \sum_{j=1}^n (u_{x_j}(x) v_j(x))_{x_i} w_i(x) \\ &= \sum_{i=1}^n \sum_{j=1}^n u_{x_j x_i}(x) v_j(x) w_i(x) + u_{x_j} \underbrace{v_{j,x_i}}_{=Dv_{ji}(x)} w_i(x) \\ &= v^T(x) H u(x) w(x) + \nabla u^T(x) Dv(x) w(x) \end{aligned}$$

Formel 13: Die **Richtungsableitungen zweiter Ordnung** in Richtung vektorwertiger Funktionen $v, w : \mathbb{R}^n \rightarrow \mathbb{R}^n$ mit $\|v\| = \|w\| = 1$ in jedem Punkt lautet

$$u_{v(x)w(x)}(x) = v^T(x) H u(x) w(x) + \nabla u^T(x) Dv(x) w(x).$$

Wenn

$$v = \frac{\nabla u}{\|\nabla u\|}$$

ist, dann gilt

$$\nabla u^T(x) Dv(x) = 0$$

und damit folgt dann direkt für diese spezielle Situation

Formel 14: Für

$$v = \frac{\nabla u}{\|\nabla u\|}$$

und $w : \mathbb{R}^n \rightarrow \mathbb{R}^n$ beliebig mit $\|w\| = 1$ in jedem Punkt lautet

$$u_{vw}(x) = \frac{\nabla u^T}{\|\nabla u\|} H u w .$$

Speziell ist dann

$$u_{\frac{\nabla u}{\|\nabla u\|} \frac{\nabla u}{\|\nabla u\|}}(x) = \frac{\nabla u^T}{\|\nabla u\|} H u \frac{\nabla u}{\|\nabla u\|} ,$$

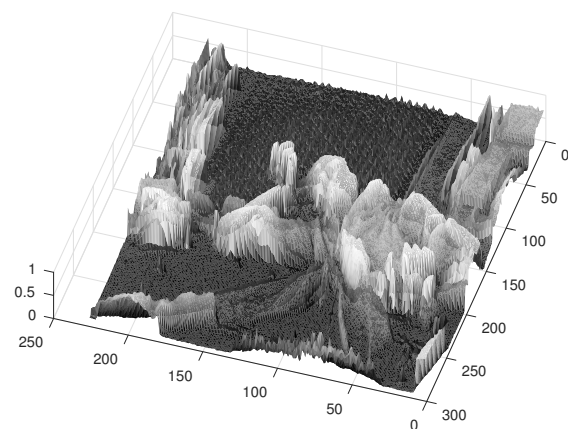
was gerade wieder der Form der zweiten Richtungsableitung für konstante Richtungen entspricht.

Nun ein paar Bildbeispiele, wie sich die verschiedenen Diffusionen auf digitale Bilder auswirken, je nachdem ob es sich um eine isotrope oder anisotrope Diffusion handelt. Man kann an diesen Beispielen sehr schön die gewählten Richtungen sehen.

Beispiel 91 iso- und anisotrope Diffusion



Original



Das Bild als Graph
(spiegelverkehrt)



$$u_t - \Delta u = 0$$



$$u_t - u_{vv} = 0, \quad v = (1, 1)^T$$



$$u_t - u_{\frac{\nabla u}{\|\nabla u\|} \frac{\nabla u}{\|\nabla u\|}} = 0$$

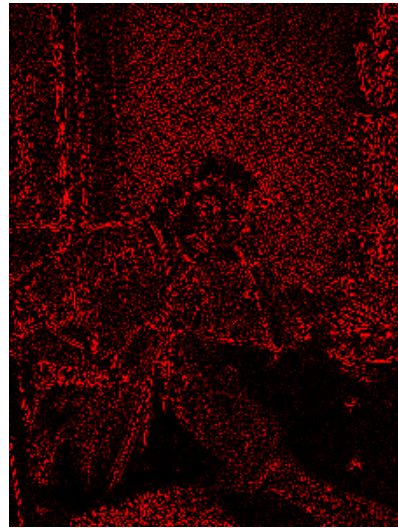


$$u_t - u_{\frac{\nabla u^\perp}{\|\nabla u^\perp\|} \frac{\nabla u^\perp}{\|\nabla u^\perp\|}} = 0$$

Beispiel 92 Texturerkennung

Sehr schön ist auch die Überlegung wie sich mittel Richtungsableitungen Texturen im Bild detektieren lassen. Die Idee ist dabei, dass sich bei sehr rauen, unregelmäßigen Bereichen die Ableitung der normierten Gradienten in tangentialer Richtung an Texturen stark ändern und bei glatten Kanten eher nicht. Die Änderung der Richtungsableitung $u_{\frac{\nabla}{\|\nabla u\|}}$ in tangentialer Richtung ist ja

3 Differentiation in höherdimensionalen Räumen



$$u \frac{\nabla u}{\|\nabla u\|} \frac{\nabla u^\perp}{\|\nabla u^\perp\|} = \frac{\nabla u^T H u \nabla u^\perp}{\|\nabla u\|^2}$$

Abbildung	1. Ableitung	2. Ableitung
$f : \mathbb{R}^n \rightarrow \mathbb{R}$ $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto f(x_1, \dots, x_n)$	Gradient $\nabla f = \begin{pmatrix} f_{x_1} \\ \vdots \\ f_{x_n} \end{pmatrix}$	Hessematrix $Hf =$ $\begin{pmatrix} f_{x_1 x_1} & \cdots & f_{x_1 x_n} \\ \vdots & & \vdots \\ f_{x_n x_1} & \cdots & f_{x_n x_n} \end{pmatrix}$
		Laplace $\Delta f = \sum_{i=1}^n f_{x_i x_i}$
Bemerkung: Der Gradient zeigt in Richtung des steilsten Anstiegs Der Laplace ist die Divergenz des Gradienten: $\Delta f = \nabla \cdot \nabla f$ und die Spur der Hessematrix: $\Delta f = \text{Spur} Hf$		
	Richtungsableitungen	
$v, w \in \mathbb{R}^n, \ w\ = \ v\ = 1$	$f_v = \nabla f \cdot v$	$f_{vw} = v^T Hf w$
$u, g : \mathbb{R}^n \rightarrow \mathbb{R}^n$ $\ u(x)\ = \ g(x)\ = 1 \quad \forall x \in \mathbb{R}^n$	$f_{u(x)}(x) = \nabla f(x) \cdot u(x)$	$f_{u(x)g(x)} =$ $u(x)^T Hf(x) g(x)$ $+ \nabla f(x)^T Du g(x)$
$g : \mathbb{R} \rightarrow \mathbb{R}^m$ $t \mapsto \begin{pmatrix} g_1(t) \\ \vdots \\ g_m(t) \end{pmatrix}$	$g'(t) = \begin{pmatrix} g'_1(t) \\ \vdots \\ g'_m(t) \end{pmatrix}$	$g''(t) = \begin{pmatrix} g''_1(t) \\ \vdots \\ g''_m(t) \end{pmatrix}$
Bemerkung: Alle Ableitungen werden kptw durchgeführt		
$f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto \begin{pmatrix} f_1(x_1, \dots, x_n) \\ \vdots \\ f_m(x_1, \dots, x_n) \end{pmatrix}$	Jacobi-Matrix $Df =$ $\begin{pmatrix} f_{1,x_1} & \cdots & f_{1,x_n} \\ \vdots & & \vdots \\ f_{m,x_1} & \cdots & f_{m,x_n} \end{pmatrix}$	
Bemerkung: Die Jacobi-Matrix des Gradienten ist die Hessematrix: $D\nabla f = Hf$		
$f : \mathbb{R}^n \rightarrow \mathbb{R}^m$	Divergenz $\nabla \cdot f = \sum_{i=1}^n f_{i,x_i}$	
Bemerkung: Die Divergenz ist die Spur der Jacobi-Matrix		
	Rotation $\text{rot} f = \nabla \times f$	
Kettenregel:	$f : \mathbb{R}^n \rightarrow \mathbb{R}, g : \mathbb{R} \rightarrow \mathbb{R}^n$ $\frac{d}{dt}(f \circ g)(t) = \sum_{i=1}^n f_{g_i}(g) g'_i(t) = \nabla f(g) \cdot g'(t)$	

Abbildung 13: Übersicht der verschiedenen Differentialoperatoren

A Tabellen zur Laplace-Transformation

Die nachfolgenden Tabellen zur Laplace-Transformation sind [Bronstein \(2008⁷\)](#) entnommen.

$$F(p) = \int_0^{\infty} e^{-pt} f(t) dt, \quad f(t) = 0 \text{ für } t < 0$$

Die in der Tabelle auftretende Konstante C ist die EULERSche Konstante $C = 0,577216$

$F(p)$	$f(t)$
0	0
$\frac{1}{p}$	1
$\frac{1}{p^n}$	$\frac{t^{n-1}}{(n-1)!}$
$\frac{1}{(p-\alpha)^n}$	$\frac{t^{n-1}}{(n-1)!} e^{\alpha t}$
$\frac{1}{(p-\alpha)(p-\beta)}$	$\frac{e^{\beta t} - e^{\alpha t}}{\beta - \alpha}$
$\frac{p}{(p-\alpha)(p-\beta)}$	$\frac{\beta e^{\beta t} - \alpha e^{\alpha t}}{\beta - \alpha}$
$\frac{1}{p^2 + 2\alpha p + \beta^2}$	$\frac{e^{-\alpha t}}{\sqrt{\beta^2 - \alpha^2}} \sin \sqrt{\beta^2 - \alpha^2} t$
$\frac{\alpha}{p^2 + \alpha^2}$	$\sin \alpha t$
$\frac{\alpha \cos \beta + p \sin \beta}{p^2 + \alpha^2}$	$\sin(\alpha t + \beta)$
$\frac{p}{p^2 + 2\alpha p + \beta^2}$	$\left(\cos \sqrt{\beta^2 - \alpha^2} t - \frac{\alpha}{\sqrt{\beta^2 - \alpha^2}} \sin \sqrt{\beta^2 - \alpha^2} t \right) e^{-\alpha t}$
$\frac{p}{p^2 + \alpha^2}$	$\cos \alpha t$
$\frac{p \cos \beta - \alpha \sin \beta}{p^2 + \alpha^2}$	$\cos(\alpha t + \beta)$
$\frac{\alpha}{p^2 - \alpha^2}$	$\sinh \alpha t$
$\frac{p}{p^2 - \alpha^2}$	$\cosh \alpha t$

$F(p)$	$f(t)$
$\frac{1}{(p-\alpha)(p-\beta)(p-\gamma)}$	$-\frac{(\beta-\gamma)e^{\alpha t} + (\gamma-\alpha)e^{\beta t} + (\alpha-\beta)e^{\gamma t}}{(\alpha-\beta)(\beta-\gamma)(\gamma-\alpha)}$
$\frac{1}{(p-\alpha)(p-\beta)^2}$	$\frac{e^{\alpha t} - [1 + (\alpha-\beta)t]e^{\beta t}}{(\alpha-\beta)^2}$
$\frac{p}{(p-\alpha)(p-\beta)^2}$	$\frac{\alpha e^{\alpha t} - [\alpha + \beta(\alpha-\beta)t]e^{\beta t}}{(\alpha-\beta)^2}$
$\frac{p^2}{(p-\alpha)(p-\beta)^2}$	$\frac{\alpha^2 e^{\alpha t} - [2\alpha - \beta + \beta(\alpha-\beta)t]\beta e^{\beta t}}{(\alpha-\beta)^2}$
$\frac{1}{(p^2 + \alpha^2)(p^2 + \beta^2)}$	$\frac{\alpha \sin \beta t - \beta \sin \alpha t}{\alpha\beta(\alpha^2 - \beta^2)}$
$\frac{p}{(p^2 + \alpha^2)(p^2 + \beta^2)}$	$\frac{\cos \beta t - \cos \alpha t}{(\alpha^2 - \beta^2)}$
$\frac{p^2 + 2\alpha^2}{p(p^2 + 4\alpha^2)}$	$\cos^2 \alpha t$
$\frac{2\alpha^2}{p(p^2 + 4\alpha^2)}$	$\sin^2 \alpha t$
$\frac{p^2 - 2\alpha^2}{p(p^2 - 4\alpha^2)}$	$\cosh^2 \alpha t$
$\frac{2\alpha^2}{p(p^2 - 4\alpha^2)}$	$\sinh^2 \alpha t$
$\frac{2\alpha^2 p}{p^4 + 4\alpha^4}$	$\sin \alpha t \cdot \sinh \alpha t$
$\frac{\alpha(p^2 + 2\alpha^2)}{p^4 + 4\alpha^4}$	$\sin \alpha t \cdot \cosh \alpha t$
$\frac{\alpha(p^2 - 2\alpha^2)}{p^4 + 4\alpha^4}$	$\cos \alpha t \cdot \sinh \alpha t$
$\frac{p^3}{p^4 + 4\alpha^4}$	$\cos \alpha t \cdot \cosh \alpha t$
$\frac{\alpha p}{(p^2 + \alpha^2)^2}$	$\frac{t}{2} \sin \alpha t$

A TABELLEN ZUR LAPLACE-TRANSFORMATION

$F(p)$	$f(t)$
$\frac{\alpha p}{(p^2 - \alpha^2)^2}$	$\frac{t}{2} \sinh \alpha t$
$\frac{\alpha \beta}{(p^2 - \alpha^2)(p^2 - \beta^2)}$	$\frac{\beta \sinh \alpha t - \alpha \sinh \beta t}{\alpha^2 - \beta^2}$
$\frac{p}{(p^2 - \alpha^2)(p^2 - \beta^2)}$	$\frac{\cosh \alpha t - \cosh \beta t}{\alpha^2 - \beta^2}$
$\frac{1}{\sqrt{p}}$	$\frac{1}{\sqrt{\pi t}}$
$\frac{1}{p\sqrt{p}}$	$2\sqrt{\frac{t}{\pi}}$
$\frac{1}{p^n \sqrt{p}}$	$\frac{n!}{(2n)!} \frac{4^n}{\sqrt{\pi}} t^{n-\frac{1}{2}} \quad (n > 0, \text{ ganz})$
$\frac{1}{\sqrt{p+\alpha}}$	$\frac{1}{\sqrt{\pi t}} e^{-\alpha t}$
$\sqrt{p+\alpha} - \sqrt{p+\beta}$	$\frac{1}{2t\sqrt{\pi t}} (e^{-\beta t} - e^{-\alpha t})$
$\sqrt{\sqrt{p^2 + \alpha^2} - p}$	$\frac{\sin \alpha t}{t\sqrt{2\pi t}}$
$\sqrt{\frac{\sqrt{p^2 + \alpha^2} - p}{p^2 + \alpha^2}}$	$\sqrt{\frac{2}{\pi t}} \sin \alpha t$
$\sqrt{\frac{\sqrt{p^2 + \alpha^2} + p}{p^2 + \alpha^2}}$	$\sqrt{\frac{2}{\pi t}} \cos \alpha t$
$\sqrt{\frac{\sqrt{p^2 - \alpha^2} - p}{p^2 - \alpha^2}}$	$\sqrt{\frac{2}{\pi t}} \sinh \alpha t$
$\sqrt{\frac{\sqrt{p^2 - \alpha^2} + p}{p^2 - \alpha^2}}$	$\sqrt{\frac{2}{\pi t}} \cosh \alpha t$
$\frac{1}{p\sqrt{p+\alpha}}$	$\frac{2}{\sqrt{\alpha\pi}} \cdot \int_0^{\sqrt{\alpha t}} e^{-\tau^2} d\tau$
$\frac{1}{(p+\alpha)\sqrt{p+\beta}}$	$\frac{2e^{-\alpha t}}{\sqrt{\pi(\beta-\alpha)}} \cdot \int_0^{\sqrt{(\beta-\alpha)t}} e^{-\tau^2} d\tau$

$F(p)$	$f(t)$
$\frac{\sqrt{p+\alpha}}{p}$	$\frac{e^{-\alpha t}}{\sqrt{\pi t}} + 2\sqrt{\frac{\alpha}{\pi}} \cdot \int_0^{\sqrt{\alpha t}} e^{-\tau^2} d\tau$
$\frac{1}{\sqrt{p^2+\alpha^2}}$	$J_0(\alpha t)$ (Besselsche Funktion der Ordnung 0)
$\frac{1}{\sqrt{p^2-\alpha^2}}$	$I_0(\alpha t)$ (modifizierte Besselsche Funktion der Ordnung 0)
$\frac{1}{\sqrt{(p+\alpha)(p+\beta)}}$	$e^{-\frac{\alpha+\beta}{2}t} \cdot I_0\left(\frac{\alpha-\beta}{2}t\right)$
$\frac{1}{\sqrt{p^2+2\alpha p+\beta^2}}$	$e^{-\alpha t} \cdot J_0\left(\sqrt{\alpha^2-\beta^2}t\right)$
$\frac{e^{1/p}}{p\sqrt{p}}$	$\frac{\sinh 2\sqrt{t}}{\sqrt{\pi}}$
$\arctan \frac{\alpha}{p}$	$\frac{\sin \alpha t}{t}$
$\arctan \frac{2\alpha p}{p^2-\alpha^2+\beta^2}$	$\frac{2}{t} \sin \alpha t \cdot \cos \beta t$
$\arctan \frac{p^2-\alpha p+\beta}{\alpha\beta}$	$\frac{e^{\alpha t}-1}{t} \sin \beta t$
$\frac{\ln p}{p}$	$-C - \ln t$
$\frac{\ln p}{p^{n+1}}$	$\frac{t^n}{n!}[\psi(n) - \ln t], \quad \psi(n) = 1 + \frac{1}{2} + \dots + \frac{1}{n} - C$
$\frac{(\ln p)^2}{p}$	$(\ln t + C)^2 - \frac{\pi^2}{6}$
$\ln \frac{p-\alpha}{p-\beta}$	$\frac{1}{t} (e^{\beta t} - e^{\alpha t})$
$\ln \frac{p+\alpha}{p-\alpha} = 2 \operatorname{artanh} \frac{\alpha}{p}$	$\frac{2}{t} \sinh \alpha t$
$\ln \frac{p^2+\alpha^2}{p^2+\beta^2}$	$2 \cdot \frac{\cos \beta t - \cos \alpha t}{t}$
$\ln \frac{p^2-\alpha^2}{p^2-\beta^2}$	$2 \cdot \frac{\cosh \beta t - \cosh \alpha t}{t}$

A TABELLEN ZUR LAPLACE-TRANSFORMATION

$F(p)$	$f(t)$
$e^{-\alpha\sqrt{p}}, \quad \operatorname{Re} \alpha > 0$	$\frac{\alpha}{2\sqrt{\pi}} \frac{e^{-\alpha^2/4t}}{t\sqrt{t}}$
$\frac{1}{\sqrt{p}} e^{-\alpha\sqrt{p}}, \quad \operatorname{Re} \alpha \geq 0$	$\frac{e^{-\alpha^2/4t}}{\sqrt{\pi t}}$
$\frac{(\sqrt{p^2 + \alpha^2} - p)^\nu}{\sqrt{p^2 + \alpha^2}}, \quad \operatorname{Re} \nu > -1$	$\alpha^\nu J_\nu(\alpha t) \quad (\text{s. S. 398})$
$\frac{(p - \sqrt{p^2 - \alpha^2})^\nu}{\sqrt{p^2 - \alpha^2}}, \quad \operatorname{Re} \nu > -1$	$\alpha^\nu I_\nu(\alpha t) \quad (\text{s. S. 399})$
$\frac{1}{p} e^{-\beta p} \quad (\beta > 0, \text{ reell})$	$\begin{cases} 0 & \text{für } t < \beta \\ 1 & \text{für } t > \beta \end{cases}$
$\frac{e^{-\beta\sqrt{p^2 + \alpha^2}}}{\sqrt{p^2 + \alpha^2}}$	$\begin{cases} 0 & \text{für } t < \beta \\ J_0(\alpha\sqrt{t^2 - \beta^2}) & \text{für } t > \beta \end{cases}$
$\frac{e^{-\beta\sqrt{p^2 - \alpha^2}}}{\sqrt{p^2 - \alpha^2}}$	$\begin{cases} 0 & \text{für } t < \beta \\ I_0(\alpha\sqrt{t^2 - \beta^2}) & \text{für } t > \beta \end{cases}$
$\frac{e^{-\beta\sqrt{(p+\alpha)(p+\beta)}}}{\sqrt{(p+\alpha)(p+\beta)}}$	$\begin{cases} 0 & \text{für } t < \beta \\ e^{-(\alpha+\beta)\frac{t}{2}} I_0\left(\frac{\alpha-\beta}{2}\sqrt{t^2 - \beta^2}\right) & \text{für } t > \beta \end{cases}$
$\frac{e^{-\beta\sqrt{p^2 + \alpha^2}}}{p^2 + \alpha^2} \left(\beta + \frac{1}{\sqrt{p^2 + \alpha^2}} \right)$	$\begin{cases} 0 & \text{für } t < \beta \\ \frac{\sqrt{t^2 - \beta^2}}{\alpha} J_1(\alpha\sqrt{t^2 - \beta^2}) & \text{für } t > \beta \end{cases}$
$\frac{e^{-\beta\sqrt{p^2 - \alpha^2}}}{p^2 - \alpha^2} \left(\beta + \frac{1}{\sqrt{p^2 - \alpha^2}} \right)$	$\begin{cases} 0 & \text{für } t < \beta \\ \frac{\sqrt{t^2 - \beta^2}}{\alpha} I_1(\alpha\sqrt{t^2 - \beta^2}) & \text{für } t > \beta \end{cases}$
$e^{-\beta p} - e^{-\beta\sqrt{p^2 + \alpha^2}}$	$\begin{cases} 0 & \text{für } t < \beta \\ \frac{\beta\alpha}{\sqrt{t^2 - \beta^2}} J_1(\alpha\sqrt{t^2 - \beta^2}) & \text{für } t > \beta \end{cases}$
$e^{-\beta\sqrt{p^2 - \alpha^2}} - e^{-\beta p}$	$\begin{cases} 0 & \text{für } t < \beta \\ \frac{\beta\alpha}{\sqrt{t^2 - \beta^2}} I_1(\alpha\sqrt{t^2 - \beta^2}) & \text{für } t > \beta \end{cases}$

$F(p)$	$f(t)$
$\frac{1 - e^{-\alpha p}}{p}$	$\begin{cases} 0 & \text{für } t > \alpha \\ 1 & \text{für } 0 < t < \alpha \end{cases}$
$\frac{e^{-\alpha p} - e^{-\beta p}}{p}$	$\begin{cases} 0 & \text{für } 0 < t < \alpha \\ 1 & \text{für } \alpha < t < \beta \\ 0 & \text{für } t > \beta \end{cases}$

B Matlab-Skripte

```
function Richtungsfeld
% dgl ist die erste Ableitung von y nach x, entspricht
% also dem Ausdruck y'(x), bzw y'(y,x)
close all

eps=1.0e-08;
Ix = [-4, 4];
Iy = [-4, 4];

y = Iy(1):1:Iy(2);
x = Ix(1):1:Ix(2);

for yn = 1:length(y)
    for xn = 1:length(x)
        % Laenge des Vektors fuer Normierung
        len = sqrt( dgl(y(yn), x(xn))^2 + 1 );
        % Laenge des Vektors entlang der Abszisse
        dx(yn,xn) = 1 / sqrt(len^2+eps^2);
        % Laenge des Vektors entlang der Ordinate
        dy(yn,xn) = dgl(y(yn), x(xn)) / sqrt(len^2+eps^2);
    end
end

% Vektoren zeichnen
quiver(x, y, dx, dy, '-r');
grid on
end

function dy=dgl(y,x)

dy = (1+x.^2)./(1+y.^2);

end
```

Abbildung 14: Matlab-Skript zur graphischen Darstellung von Richtungsfeldern; hier zur Gleichung $y' = (x + y)/(x - y)$

```

function PlotFlaeche()

DIM = [50 50];
Ix=[-1 1]; hx=(Ix(2)-Ix(1))/(DIM(1)-1);
Iy=[-1 1]; hy=(Iy(2)-Iy(1))/(DIM(2)-1);

% zu optimierende Funktion -----
[xx,yy] = meshgrid(Ix(1):hx:Ix(2),Iy(1):hy:Iy(2));

f = @(x,y) x.^2-y.^2.*x;

zz=f(xx,yy);
surfc(xx,yy,zz);
hold on

% Nebenbedingungen -----
x1 = Ix(1):hx:Ix(2);
y1 = sqrt(1-x1.^2);
y2 = (x1+1)/2;

plot3(x1,y1,f(x1,y1),'go','LineWidth',3);
plot3(x1,-y1,f(x1,-y1),'go','LineWidth',3);
plot3(x1,y2,f(x1,y2),'ro','LineWidth',3);
xlabel('x'); ylabel('y');
end

```

Abbildung 15: Matlab-Skript zur Darstellung der Graphiken in Abbildung 9

Vektorfeld.m

```
Ix = [-1 1];
Iy = [-1 1];
DIM = [20 20];
hx = (Ix(2)-Ix(1))/(DIM(1)-1);
hy = (Iy(2)-Iy(1))/(DIM(2)-1);

[x1, x2] = meshgrid(Ix(1):hx:Ix(2), Iy(1):hy:Iy(2));

u1 = 2*x1.^2.*x2;
u2 = 3*x1-x2;
div = @(x1,x2) 4*x1.*x2-1;

subplot(1,2,1)
plot vector field
quiver(x1, x2, u1, u2)
grid
xlabel('x1');
ylabel('x2');

x3=div(x1,x2);
xx = x1(1,:);
yy = 1./4./sqrt(1.0e-08+xx.^2);
subplot(1,2,2)
surfc(x1,x2,x3);
xlabel('x1');
ylabel('x2');
```

Abbildung 16: Matlab-Skript zur Darstellung von Vektorfeldern und Divergenzen

Index

$C^k(D)$, 59

Ω , 39

(un-) abhängige Ereignisse, 42

Divergenzoperator, 83

Lagrange-Multiplikator, 79

Rotationsoperator, 85

Additionssatz, 41

Anfangswertproblem (AWP), 3

Axiome von Kolmogorow, 40

Baumdiagramm, 44

bedingte Wahrscheinlichkeit, 46

Charakteristisches Polynom, 19

Definitheit, 61

diagonaldominante Matrizen, 66

Diffusionsgleichung, 90

Divergenz, 83

Ereignis, 39

Ergebnisraum, 39

Ergebnisse, 39

Eulersche Betrachtungsweise, 89

Flussrichtung, 59

Gegenereignis, 39

Gewöhnliche Differentialgleichung (DGL/O-DE), 1

Gradient, 54

Grundereignisse, 39

Hessematrix, 60

Hurwitz-Kriterium, 63

indefinit, 61

Jacobi-Matrix, 82

k-mal stetig differenzierbar, 59

Kettenregel, 56

Konvergenzbereich, 28

Kriterium von Sylvester, 63

kritische Stelle, 60

Kronecker-Symbol, 65

Lagrange-Funktion, 79

Lagrangesche Betrachtungsweise, 89

Laplace-Integral, 28

Laplace-Operator, 86

Laplace-Transformation, 28

Laplace-Wahrscheinlichkeit, 37

lineare (un-)abhängig, 16

lokales Extremum, 60

materielle/substantielle Ableitung, 89

mehrstufige Versuche, 44

Nabla-Operator, 54

negativ (semi) definit, 61

partielle Ableitung, 54

Pfad, 44

Pfadwahrscheinlichkeiten, 45

positiv (semi) definit, 61

Richtungsableitung, 53

Rotation, 85

Satz von Bayes, 47

sicheres Ereignis, 40

stationärer Punkt, 60

Stufenwahrscheinlichkeiten, 45

symmetrischer Anteil einer Matrix, 62

totale Wahrscheinlichkeit, 46

totales Differential, 87

Trennung der Variablen, 3

unmögliches Ereignis, 40

unvereinbar, 40

Wahrscheinlichkeit, 40

Wahrscheinlichkeit von Kolmogorow, 40

Wronski-Determinante, 17

Literatur

Bronstein, I. *Taschenbuch der Mathematik* (Verlag Harri Deutsch, 2008⁷).

do Carmo, M. *Differentialgeometrie von Kurven und Flächen* (1993).

Heuser, H. *Gewöhnliche Differentialgleichungen*, Band 1 (Vieweg+Teubner Stuttgart, 2009⁶).

Rießinger, T. *Mathematik für Ingenieure* (Springer-Verlag, 2011⁸).

Walter, W. *Gewöhnliche Differentialgleichungen* (Springer-Verlag, 1972).