Analysis II

gehört bei Prof. Dr. M. Griebel

Sommersemester 2001

Inhaltsverzeichnis

Einführung

Die Betrachtungen von ANALYSIS I blieben im wesentlichen auf *Funktionen* f von $\mathbb{R}$ nach $\mathbb{R}$ beschränkt.

Die reellen Zahlen erfüllen die Axiome eines vollständigen, geordneten Körpers. Als Körper ist $\mathbb{R}$ Vektorraum über sich selbst: eindimensionaler $\mathbb{R}$-Vektorraum, üblicherweise dargestellt als «Zahlengerade». Dies rechtfertigt die gleichwertigen Schreibweisen:

$$f : \mathbb{R} \to \mathbb{R} \quad \text{und} \quad f : \mathbb{R}^1 \to \mathbb{R}^1.$$

Ziel der Vorlesung

- Differential- und Integralrechnung in mehreren Variablen
- Einführung in gewöhnliche und partielle Differentialgleichungen
- Variationsrechnung (Ausblick)

Hierzu werden die kommenden Betrachtungen auf *Abbildungen* f von endlichdimensionalen $\mathbb{R}$-Vektorräumen $\mathbb{R}^n$ in endlichdimensionale $\mathbb{R}$-Vektorräume $\mathbb{R}^m$ erweitert, in Zeichen (i.Z.)

$$f : \mathbb{R}^n \to \mathbb{R}^m \quad \text{mit} \quad 1 \leq n < \infty, 1 \leq m < \infty, \quad n, m \in \mathbb{N}.$$

Es zeigt sich, daß die im Eindimensionalen gewonnenen Ergebnisse vielfach auf n-dimensionale Betrachtungen «analog» übertragen und verallgemeinert werden können. In vielen Bereichen kommt es jedoch zu neuartigen Fragestellungen, gegenüber deren Antworten bisherige Ergebnisse zu Spezialfällen werden oder sich als Vereinfachungen erweisen.

Es ist zweckmäßig, zu *n-Tupeln* (Vektoren) und im weiteren auch zu *Matrizen* überzugehen, was u.a. die Einführung von *Normen* erfordert, die an die Stelle des *Betrages* $|x|$ einer Zahl x treten. Letzterer wird zum Spezialfall.

Wo möglich und geboten, werden Vektorräume über dem Körper $\mathbb{C}$ in die Betrachtungen einbezogen.

Literatur zur Vorlesung

- Barner M.; Flohr F.: *Analysis I*, *Analysis II*. De Gruyter, 2000, 1996
- Erwe F.: *Differential- und Integralrechnung I, II*. Bibliographisches Institut, 1968
- Forster O.: *Analysis 1*, *Analysis 2*. Vieweg, 1999
- Heuser H.: *Lehrbuch der Analysis Teil 1*, *Teil 2*. Teubner, 1994, 1995
- Königsberger K.: *Analysis 1*, *Analysis 2*. Springer, 1993
- Zeidler E. (Hrsg.): *Teubner-Taschenbuch der Mathematik Teil 1, Teil 2*. Teubner, 1996, 1995

Bemerkungen zur Notation

Im Schriftbild werden *Abbildungen* und *Funktionen* (als Bezeichnung vorrangig gebräuchlich für Abbildungen in die Zielmenge $\mathbb{R}$) oder *n-Tupel* (Vektoren) und *Zahlen* (Skalare) nicht unterschieden; Ausnahme: Kapitel 3.8 Nablakalkül.

Erinnerung

Eine nichtleere Menge V heißt *Vektorraum* oder *linearer Raum* über dem Körper K, wenn *auf* V die beiden Verknüpfungen (Operationen)

1. der *Addition* der Elemente $x \in V$: für je zwei Elemente $x, y \in V$ gibt es ein Element $z \in V$ mit $z = x + y$, ihre *Summe*,
2. der *Multiplikation* mit Elementen $c \in K$: für jedes $x \in V$ und jedes $c \in K$ gibt es ein $y \in V$ mit $y = cx$, das *Produkt* aus x und dem Skalar c,

erklärt und dabei die *Vektorraumaxiome* (vgl. Vorlesung LINEARE ALGEBRA I) erfüllt sind. V heißt *reeller* bzw. *komplexer* Vektorraum, wenn K der Körper $\mathbb{R}$ der reellen bzw. der Körper $\mathbb{C}$ der komplexen Zahlen ist. Die Elemente von V heißen *Punkte* (Vektoren). Sie werden als n-Tupel $x \in V$ mit

$$x := (x_1, \ldots, x_n)\,,\ x_i \in K,\ i, n \in \mathbb{N}$$

dargestellt.

Sei $V := \mathbb{R}^n$, $K := \mathbb{R}$, $x \in V$ mit $x := (x_1, \ldots, x_n)\,, x_i \in \mathbb{R}$ und $c \in \mathbb{R}$. Gelten

1. die *Vektoraddition* $x + y = (x_1 + y_1, \ldots, x_n + y_n)$ und
2. die *Skalarmultiplikation* $cx = (cx_1, \ldots, cx_n)\,,$

so wird $\mathbb{R}^n$ zum $\mathbb{R}$- oder *reellen* Vektorraum.

Mit $0 := (0, \ldots, 0)$ ist - wie üblich - der *Nullvektor* bezeichnet.

1 Topologie des $\mathbb{R}^n$

1.1 Skalarprodukt. Norm. Abstand

Sei $\mathbb{R}^n$ Vektorraum über $\mathbb{R}$, $x, y \in \mathbb{R}^n$ und $c \in \mathbb{R}$.

Definition 1.1 (*Skalarprodukt*)

Die *Funktion* $\langle\,,\,\rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, definiert durch

$$(x, y) \mapsto \langle x, y \rangle := \sum_{i=1}^{n} x_i y_i,$$

heißt *Skalarprodukt* von x und y, i.Z. auch $\langle\,|\,\rangle\,, \boldsymbol{x} \cdot \boldsymbol{y}, \boldsymbol{xy}$.

Bemerkung

x, y heißen *orthogonal* (i.Z. $x \perp y$), falls $\langle x, y \rangle = 0$.

Definition 1.2 (*Norm auf* $\mathbb{R}^n$)

Eine *Funktion* $N : \mathbb{R}^n \to \mathbb{R}$ mit den Eigenschaften

$$\begin{array}{ll} (1) & N(x) > 0, \text{ falls } x \neq 0 \\ (2) & N(cx) = |c|\, N(x) \\ (3) & N(x + y) \leq N(x) + N(y) \end{array}$$

heißt *Norm auf* $\mathbb{R}^n$, i.Z. auch $\|x\|$ oder $\|\cdot\|$.

Bemerkung

Für $n = 1$ sind *alle* Normen mit $N(x) = c\,|x|$ mit beliebigem $c > 0$ beschrieben.

Für $n > 1$ gibt es viele Möglichkeiten, Def. 1.2 zu erfüllen.

Da Normen im Endlichdimensionalen einander äquivalent sind (vgl. Def. 1.6 und Satz 1.9), kommt es auf eine geschickte Auswahl für die jeweilige Fragestellung an.

Beispiele

$\|x\|_2 := \sqrt{\langle x, x \rangle}$	EUKLIDische Norm
$\|x\|_p := (\|x_1\|^p + \cdots + \|x_n\|^p)^{\frac{1}{p}}$	p – Norm
$\|x\|_\infty := \max\{\|x_1\|, \ldots, \|x_n\|\}$	Maximumsnorm

Bemerkung

$$\frac{1}{\sqrt{n}}\,\|x\|_2 \leq \|x\|_\infty \leq \|x\|_2 \quad \text{bzw.} \quad \|x\|_\infty \leq \|x\|_2 \leq \sqrt{n}\,\|x\|_\infty\,.$$

Definition 1.3 (EUKLID*ischer Abstand*)

Sei $\mathbb{R}^n$ versehen mit der EUKLIDischen Norm.

$$d(x, y) := \|x - y\|_2$$

heißt EUKLID*ischer Abstand*, i.Z. auch $\operatorname{dist}(x, y)$, i.W. auch *Distanz*, Entfernung.

Bemerkung

Zur Verallgemeinerung von *Abstand* zu *Metrik auf einer Menge* X siehe Def. nach Satz 1.4.

Satz 1.1 (*Eigenschaften des Skalarprodukts*)

Das Skalarprodukt ist eine *positiv definite symmetrische* Bilinearform:

$$\begin{array}{rcll}
\langle x,x\rangle & \geq & 0 & \text{positiv definit} \\
\langle x,x\rangle & = & 0 \;\Leftrightarrow\; x=0 & \\
\langle x,y\rangle & = & \langle y,x\rangle & \text{symmetrisch} \\
\langle x+y,z\rangle & = & \langle x,z\rangle + \langle y,z\rangle & \\
\langle z,x+y\rangle & = & \langle z,x\rangle + \langle z,y\rangle & \\
\langle cx,y\rangle & = & c\,\langle x,y\rangle & \\
\langle x,cy\rangle & = & c\,\langle x,y\rangle &
\end{array}$$

Satz 1.2 (*Eigenschaften der* EUKLID*ischen Norm*)

Die EUKLID*ische Norm* $\|x\|_2 := \sqrt{\langle x,x\rangle}$ hat die Eigenschaften

$$\begin{array}{rcll}
\|x\|_2 & \geq & 0, & \|x\|_2 = 0 \;\Leftrightarrow\; x=0 \\
\|cx\|_2 & = & |c|\,\|x\|_2 & \\
\|x+y\|_2 & \leq & \|x\|_2 + \|y\|_2 & \text{Dreiecksungleichung}
\end{array}$$

Satz 1.3 (SCHWARZ*sche Ungleichung. Lineare Abhängigkeit*)

Mit der Notation für die EUKLIDische Norm gilt

$$\langle x,y\rangle^2 \quad\leq\quad \|x\|_2^2\,\|y\|_2^2$$

sowie

$$\langle x,y\rangle = \|x\|_2\,\|y\|_2 \quad\Leftrightarrow\quad x \text{ und } y \text{ sind } \textit{linear abhängig}.$$

Beweis von Satz 1.3 (SCHWARZ*sche Ungleichung*)

Die SCHWARZsche Ungleichung ist für $x=0$ oder $y=0$ offensichtlich richtig.

Sei jetzt $y \neq 0, c \in \mathbb{R}$. Wegen der Eigenschaften des Skalarproduktes ist

$$\begin{array}{rcl}
0 & \leq & \langle x+cy, x+cy\rangle \\
 & = & \langle x, x+cy\rangle + \langle cy, x+cy\rangle \\
 & = & \langle x,x\rangle + \langle x,cy\rangle + \langle cy,x\rangle + \langle cy,cy\rangle \\
 & = & \langle x,x\rangle + 2c\,\langle x,y\rangle + c^2\,\langle y,y\rangle \\
 & = & \|x\|_2^2 + 2c\,\langle x,y\rangle + c^2\,\|y\|_2^2\,.
\end{array}$$

Mit $c = -\frac{\langle x,y\rangle}{\|y\|_2^2}$ wird

$$\begin{array}{rcl}
0 & \leq & \|x\|_2^2 - \dfrac{\langle x,y\rangle^2}{\|y\|_2^2}, \\
0 & \leq & \|x\|_2^2\,\|y\|_2^2 - \langle x,y\rangle^2\,, \\
\langle x,y\rangle^2 & \leq & \|x\|_2^2\,\|y\|_2^2\,.
\end{array}$$

$\square$

Beweis von Satz 1.3 (*Lineare Abhängigkeit*)

Lineare Abhängigkeit $\Rightarrow$ *Gleichheit.* Sei $x = cy$. Dann gilt

$$\begin{aligned} \langle x, y\rangle^2 &= \langle cy, y\rangle^2 \\ &= c^2 \|y\|_2^4 \\ &= \|cy\|_2^2 \|y\|_2^2 \\ &= \|x\|_2^2 \|y\|_2^2 . \end{aligned}$$

Gleichheit $\Rightarrow$ *lineare Abhängigkeit.* Sei $\langle x, y\rangle^2 = \|x\|_2^2 \|y\|_2^2$ und $y \neq 0$. Dann gibt es ein $c = \frac{\langle x,y\rangle}{\|y\|_2^2}$, so daß $x = cy$. Denn

$$\begin{aligned} \langle x - cy, x - cy\rangle &= \left\langle x - \frac{\langle x, y\rangle}{\|y\|_2^2} y, x - \frac{\langle x, y\rangle}{\|y\|_2^2} y \right\rangle \\ &= \frac{\|x\|_2^2 \|y\|_2^2 - |\langle x, y\rangle|^2}{\|y\|_2^2} \\ &= 0. \end{aligned}$$

$\square$

Beweis der Dreiecksungleichung mit Hilfe der SCHWARZ*schen Ungleichung*

Für die EUKLIDischen Norm gilt

$$\begin{aligned} \|x + y\|_2^2 &= \langle x + y, x + y\rangle \\ &= \|x\|_2^2 + 2\langle x, y\rangle + \|y\|_2^2 \\ &\leq \|x\|_2^2 + 2\|x\|_2 \|y\|_2 + \|y\|_2^2 \\ &= |\|x\|_2 + \|y\|_2|^2 . \end{aligned}$$

$\square$

Bemerkung

Lineare Abhängigkeit, Gleichheit in der SCHWARZschen Ungleichung und Gleichheit in der Dreiecksungleichung sind äquivalent.

$$\begin{aligned} x = cy &\Leftrightarrow \langle x, y\rangle^2 = \|x\|_2^2 \|y\|_2^2 \\ &\Leftrightarrow \|x + y\|_2 = \|x\|_2 + \|y\|_2 . \end{aligned}$$

Satz 1.4 (*Eigenschaften des Abstandes*)

$$\begin{aligned} d(x, y) &\geq 0 \\ d(x, y) &= 0 \Leftrightarrow x = y \\ d(x, y) &= d(y, x) \\ d(x, y) &\leq d(x, z) + d(z, y) \end{aligned}$$

Metrik

Sei X eine Menge. Eine Abbildung $d : X \times X \to \mathbb{R}$ mit den Eigenschaften des Abstandes nach Satz 1.4 heißt *Metrik auf X*.

Bezeichnungen

- $(\mathbb{R}^n; \langle\, , \,\rangle)$ heißt EUKLID*ischer* Vektorraum
- $(\mathbb{R}^n; \|\cdot\|)$ heißt *normierter* Vektorraum
- $(\mathbb{R}^n; d\,(\, , \,))$ heißt *metrischer* Vektorraum

1.2 Folgen. Reihen. Konvergenz in $\mathbb{R}^n$

Für Folgen, Reihen und Konvergenz lassen sich die in $\mathbb{R}^1$ notierten Definitionen im wesentlichen durch Ersetzen des absoluten Betrages $|\cdot|$ durch die Norm $\|\cdot\|$, hier zunächst die Maximumsnorm $\|\cdot\|_\infty$, auf $\mathbb{R}^n$ übertragen.

Definition 1.4 (*konvergente Folge*)

Sei $x_k = (x_{1,k}, \ldots, x_{n,k}) \in \mathbb{R}^n$ *Folgenglied.* Die *Folge* (x_k) heißt *konvergent* gegen $a = (a_1, \ldots, a_n)$, wenn

$$\lim_k \|x_k - a\|_\infty = 0, \text{ i.Z. } \lim x_k = a.$$

Satz 1.5

Die *Folge* (x_k) konvergiert $\Leftrightarrow$ Jede *Komponentenfolge* $(x_{i,k})$, $i = 1, \ldots, n$ konvergiert.

Bemerkung (Beweisansatz) $\forall i = 1, \ldots, n$:

$$\begin{aligned} |x_{i,k} - a_i| &\leq \|x_k - a\|_\infty \\ \|x_k - a\|_\infty &\leq |x_{1,k} - a_1| + \ldots + |x_{n,k} - a_n| \end{aligned}$$

Satz 1.6 (CAUCHY*sches Konvergenzkriterium in* $\mathbb{R}^n$).

Die *Folge* (x_k) konvergiert $\Leftrightarrow$ Es gilt das CAUCHY*sches Konvergenzkriterium*

$$\forall \varepsilon > 0\ \exists k_0\ \forall k \geq k_0\ \forall p > 0 \quad \|x_{k+p} - x_k\|_\infty < \varepsilon.$$

Beweis

analog Beweis zu ANALYSIS I Satz 2.10 für erste Komponentenfolge mit $i = 1$ in $\mathbb{R}^1$ und sukzessiver Hinzunahme der Komponentenfolgen mit $i = 2, \ldots, n$. $\square$

Bemerkung

Die Rechenregeln für konvergente Folgen in $\mathbb{R}^1$ können auf $\mathbb{R}^n$ übertragen werden.

Definition 1.5 (*beschränkte Folge*)

Die *Folge* (x_k) heißt *beschränkt*, wenn

$$\exists c \in \mathbb{R} \, \forall k \in \mathbb{N} \quad \|x_k\|_\infty < c.$$

Satz 1.7 (*Satz von* BOLZANO-WEIERSTRASS *in* $\mathbb{R}^n$)

Jede beschränkte Folge (x_k) besitzt eine konvergente Teilfolge.

Beweis

Satz 1.7 gelte in $\mathbb{R}^1$ als bewiesen. Betrachte Komponentenfolgen $(x_{i,k})$. Diese sind für alle $i = 1, \ldots, n$ beschränkt.

1.*Schritt*: Für $i = 1$ besitzt $(x_{1,k})$ eine konvergente Teilfolge $(x_{1,k'})$. Gehe für (x_k) zu dieser bzgl. k über: $(x_{k'})$.

2.*Schritt*: Für $i = 2$ besitzt $(x_{2,k'})$ eine konvergente Teilfolge $(x_{2,k''})$. Gehe für (x_k) zu dieser bzgl. k über: $(x_{k''})$.

Nach n *Schritten* ergibt sich $(x_{k^{(n)}})$ mit der Eigenschaft, daß jede Komponentenfolge konvergiert $\Rightarrow$ $(x_{k^{(n)}})$ konvergiert. $\square$

Mit der Verwendung der Maximumsnorm $\|\cdot\|_\infty$ ist die Reduktion auf Eigenschaften der Komponentenfolgen einfach möglich. Dies bedeutet aber keine Auszeichnung dieser Norm:

Satz 1.8

Sei N eine Norm auf $\mathbb{R}^n$. Dann gilt $\forall N \, \exists a, b > 0$:

$$\begin{array}{ll} \text{i)} & N(x) \leq a \|x\|_\infty \\ \text{ii)} & \|x\|_\infty \leq bN(x) \end{array}$$

Beweis zu i)

Sei $x = (x_1, \ldots, x_n)$ dargestellt als Linearkombination $x = x_1 e_1 + \ldots + x_n e_n$.

$$\begin{array}{rcl} N(x) & = & N(x_1 e_1 + \ldots + x_n e_n) \\ & \leq_{(3)} & N(x_1 e_1) + \ldots + N(x_n e_n) \\ & =_{(2)} & |x_1| \, N(e_1) + \ldots + |x_n| \, N(e_n) \\ & \leq & \|x\|_\infty \left(N(e_1) + \ldots + N(e_n)\right). \end{array}$$

Dabei bedeutet (3) die Eigenschaft (3) der Norm nach Def. 1.2, (2) entsprechend. Mit $a := N(e_1) + \ldots + N(e_n)$ folgt die Aussage i). $\square$

Beweis zu ii) (indirekt): Annahme $\neg \exists b > 0$.

Dann kann man zu jedem $b \in \mathbb{N}$ jeweils ein x_k finden mit $\|x_k\|_\infty < kN(x_k)$ und wegen der Eigenschaft (2) der Norm $\Rightarrow N\left(\frac{x_k}{\|x_k\|_\infty}\right) < \frac{1}{k}$.

Mit $y_k = \frac{x_k}{\|x_k\|_\infty}$ und $N(y_k) < \frac{1}{k}$ ergibt sich daraus ein Widerspruch.

Denn wegen $\|y_k\|_\infty = 1$ ist (y_k) beschränkt $\Rightarrow \exists$ eine konvergente Teilfolge (z_k) mit Grenzwert z bzw. $\lim_k \|z_k - z\|_\infty = 0$.

Mit i) $N(z_k - z) \leq a \|z_k - z\|_\infty$ ergibt sich

$$\begin{aligned} N(z) &= N(z - z_k + z_k) \\ &\leq N(z - z_k) + N(z_k) \\ &\leq a \|z - z_k\|_\infty + N(z_k). \end{aligned}$$

Zudem folgt aus $N(y_k) < \frac{1}{k}$, daß für die Teilfolge (z_k) gilt:

$\lim_k N(z_k) = 0 \Rightarrow N(z) = 0 \Rightarrow z = 0$.

Andererseits gilt: $z_k = z + (z_k - z)$,

$\|z_k\|_\infty = 1 \Rightarrow 1 \leq \|z\|_\infty + \|z_k - z\|_\infty \Rightarrow \|z\| \geq 1 \Rightarrow z \neq 0$. Widerspruch $\square$

Definition 1.6 (*Äquivalenz von Normen*)

Eine Norm N' heißt *äquivalent* zu einer Norm N (i. Z. $N' \sim N$), falls

$$\exists a, b > 0 \, \forall x \in \mathbb{R}^n \quad aN(x) \leq N'(x) \leq bN(x)$$

Beachte: a, b dieser Def. sind unabhängig von den a, b in Satz 1.8!

Satz 1.9 (*Äquivalenz von Normen*)

Je zwei Normen auf $\mathbb{R}^n$ mit $n < \infty$ sind äquivalent.

Beweis

Nach Def. 1.6 und Satz 1.8 gilt: $N \sim \|\cdot\|_\infty \wedge N' \sim \|\cdot\|_\infty \Rightarrow N \sim N'$.

Noch zu zeigen, daß Äquivalenzrelation gegeben, d.h. Reflexivität, Symmetrie und Transitivität:

- Jede Norm ist zu sich selbst äquivalent.
- $aN(x) \leq N'(x) \leq bN(x) \Rightarrow \frac{1}{b}N'(x) \leq N(x) \leq \frac{1}{a}N'(x)$.
- $(aN(x) \leq N'(x) \leq bN(x)) \wedge (a'N'(x) \leq N''(x) \leq b'N'(x))$
  $\Rightarrow aa'N(x) \leq N''(x) \leq bb'N(x)$. $\square$

Definition 1.7 (*Reihe. Reihensumme*)

Eine Folge von Teilsummen

$$(s_k) = \left(\sum_{j=1}^{k} x_j \right)$$

heißt *(unendliche) Reihe.* Wenn existent, wird der Grenzwert

$$\lim_{k \to \infty} s_k = s = \sum_{k=1}^{\infty} x_k$$

als *Reihensumme* bezeichnet.

Satz 1.10

Die Reihe $\sum_{j=1}^{\infty} x_j$ konvergiert

$\Leftrightarrow$ Jede ihrer Komponentenreihen $\sum_{i=1}^{\infty} x_{ij}, i = 1, \ldots, n$ konvergiert.

Definition 1.8 (*absolute Konvergenz. Norm-Konvergenz*)

Die Reihe $\sum_{j=1}^{\infty} x_j$ heißt *absolut konvergent* (*Norm-konvergent*), wenn die Reihe $\sum_{j=1}^{\infty} \|x_j\|$ konvergiert.

Bemerkung

Wegen

$$\left\| \sum_{j=k+1}^{k+p} x_j \right\| \leq \sum_{j=k+1}^{k+p} \|x_j\|$$

und CAUCHYschem Konvergenzkriterium gilt:

absolut konvergent $\Rightarrow$ konvergent.

Bemerkung

absolut konvergente Reihe $\Leftrightarrow$ n Komponentenreihen absolut konvergent.

1.3 Offene Mengen

Definition 1.9 (*ε-Umgebung*)

Sei $\varepsilon > 0, x_0 \in \mathbb{R}^n$.

$$U_\varepsilon(x_0) := \{x \mid \|x - x_0\| < \varepsilon\}$$

heißt *ε-Umgebung* von x_0.

Beispiele

- $U_\varepsilon(x_0) := \{x \mid \|x - x_0\|_2 < \varepsilon\}$ *ε-Umgebung* von x_0 bzgl. der EUKLIDischen Norm: «Kugel». Diese Norm bewirkt eine Umgebung, die nach allen Richtungen hin «isotrop» ist.
- $W_\varepsilon(x_0) := \{x \mid \|x - x_0\|_\infty < \varepsilon\}$ *ε-Umgebung* von x_0 bzgl. der Maximumsnorm: «Würfel».

Bemerkung

$$W_{\frac{\varepsilon}{\sqrt{n}}}(x_0) \subset U_\varepsilon(x_0) \subset W_\varepsilon(x_0)$$

Bemerkung

Norm in Umgebungsdefinition beliebig wählbar: Umgebung ist immer *konvex.*

Definition 1.10 (*Umgebung*)

$U = U(x_0)$ heißt *Umgebung* von x_0 $\Leftrightarrow$ Es gibt eine ε-Umgebung mit $U_\varepsilon(x_0) \subset U(x_0)$.

Definition 1.11 (*offene Umgebung. abgeschlossene Umgebung*)

$U \subset \mathbb{R}^n$ heißt *offen* $\Leftrightarrow U$ ist Umgebung eines jeden Punktes $x \in U$.

$M \subset \mathbb{R}^n$ heißt *abgeschlossen* $\Leftrightarrow M^c = \mathbb{R}^n \setminus M$ ist *offen*.

Beispiele

- $U_\varepsilon(x_0), W_\varepsilon(x_0)$ sind offen
- $\{x \mid \|x - x_0\| \leq \varepsilon\}$ ist nicht offen, aber abgeschlossen

Satz 1.11

$\emptyset$ ist offen *und* abgeschlossen.

$\mathbb{R}^n$ ist offen *und* abgeschlossen.

Sei $U_i, i \in I$ Familie von offenen Mengen, dann ist $U = \bigcup_{i \in I} U_i$ offen.

Sind U, V offen, dann ist $U \cap V$ offen.

(Gilt analog für *beliebige* Schnitte und für *endliche* Vereinigungen *abgeschlossener* Mengen.)

Topologie. Topologischer Raum

Sei X eine Menge, $\mathcal{T}$ eine Menge von Teilmengen von X mit den Eigenschaften von Satz 1.11., d.h.

- $\emptyset, X \in \mathcal{T}$,
- Vereinigung *beliebig* vieler Mengen von $\mathcal{T}$ gehören zu $\mathcal{T}$,
- Durchschnitt *endlich* vieler Mengen von $\mathcal{T}$ gehören zu $\mathcal{T}$.

Dann heißt $(X; \mathcal{T})$ *topologischer Raum* und $\mathcal{T}$ als Menge der offenen Mengen von X heißt *Topologie*.

Das System der offenen Mengen des metrischen Vektorraumes $\mathbb{R}^n$ ist eine Topologie. Da jeder metrische Raum ein HAUSDORFF-Raum ist, gilt der nachstehende

Satz 1.12 (*Trennungssatz*)

Sei $x \neq y$, dann existieren stets Umgebungen $U(x)$, $V(y)$ mit $U \cap V = \emptyset$.

Folgerung

$\{x\}$ ist abgeschlossen.

Definition 1.12 (*abgeschlossene Hülle. Inneres. Rand*)

Sei $M \subset \mathbb{R}^n$. Dann heißt

$\bar{M} = \bigcap \{M' \mid M' \, \text{abgeschlossen}, M' \supset M\}$ *abgeschlossene Hülle* von M,

$M^o = \bigcup \{U \mid U \, \text{offen}, U \subset M\}$ *Inneres* von M,

$\bar{M} \setminus M^o = \mathrm{b}M = \partial M$ *Rand* von M (b wie «boundary»).

Bemerkung

Es gilt $M^o \subset M \subset \bar{M}$ mit M^o offen, $\bar{M}$ abgeschlossen.

Beispiel 1 ε-Umgebung $U_\varepsilon(x_0) = \{x \mid \|x - x_0\|_2 < \varepsilon\}$

- $U_\varepsilon^o(x_0) = U_\varepsilon(x_0)$,
- $\overline{U_\varepsilon(x_0)} = \{x \mid \|x - x_0\|_2 \leq \varepsilon\}$,
- $\partial U_\varepsilon(x_0) = \{x \mid \|x - x_0\|_2 = \varepsilon\}$.

Beispiel 2 x-Achse $M = \{(x,0) \mid x \in \mathbb{R}\} \subset \mathbb{R}^2$

- $M^o = \emptyset$,
- $\bar{M} = M$,
- $\partial M = M$.

Beispiel 3 $M = \mathbb{Q}^n$

- $M^o = \emptyset$,
- $\bar{M} = \mathbb{R}^n$,
- $\partial M = \mathbb{R}^n$.

Definition 1.13 (*Häufungspunkt. Limes*)

i) x_0 heißt *Häufungspunkt* der Folge (x_j) :$\Leftrightarrow$ In *jeder* Umgebung von x_0 liegen *unendlich* viele x_j.

ii) x_0 heißt *Limes* der Folge (x_j), i.Z. $\lim x_j = x_0$:$\Leftrightarrow$ In *jeder* Umgebung von x_0 liegen *fast alle* x_j.

Bemerkung

- Limites eindeutig bestimmt
- (x_j) konvergiert $\Leftrightarrow$ (x_j) ist beschränkt und hat genau *einen* Häufungspunkt (vgl. Cauchy-Folgen)
- Rechenregeln

$$\begin{aligned} \lim(x_j + y_j) &= \lim x_j + \lim y_j \\ \lim cx_j &= c \lim x_j \\ \lim \langle x_j, y_j \rangle &= \langle \lim x_j, \lim y_j \rangle \end{aligned}$$

Satz 1.13

Folgende Aussagen sind äquivalent:

1. M abgeschlossen
2. $M = \bar{M}$
3. $\partial M \subset M$
4. Jeder Häufungspunkt von M liegt in M.

Satz 1.14 (*Abgeschlossene Hülle*)

$\bar{M} = M^o \cup \partial M = M \cup \{\text{Häufungspunkte von } M\}$.

Beweis (zu Satz 1.13)

1. $\Rightarrow$ 2.: $M \subset \bar{M} = \bigcap \left\{ \bar{M} \subset M \mid \bar{M} \text{ abgeschlossen} \right\} \subset M$.

2. $\Rightarrow$ 3.: $\partial M \subset \bar{M} = M$.

3. $\Rightarrow$ 4.: $\partial M \subset M \Rightarrow$ jeder Häufungspunkt (HP) von M liegt in M.

Beweis in zwei Schritten.

a) (indirekt) Sei x_0 HP von M. Z.z. $x_0 \in \bar{M}$.

Annahme: $x_0 \notin \bar{M} \Rightarrow x_0 \in \bar{M}^c$.

$\bar{M}$ abgeschlossen $\Rightarrow \bar{M}^c$ offen $\Rightarrow \exists U(x_0) \subset \bar{M}^c$.

Wegen $M \subset \bar{M}$ folgt $U(x_0) \subset \bar{M}^c$.

$U(x_0)$ enthält also keine Punkte aus M, also kann x_0 kein HP sein, d.h. $x_0 \in \bar{M}$.

b) Sei $\partial M \subset M$ und x_0 HP von M.

Wegen a) $x_0 \in \bar{M}$ ist $M^o \cup \partial M = M \cup \partial M = M$.

4. $\Rightarrow$ 1.: Sei $x_0 \in M^c$.

Dann ist wegen 4. x_0 kein HP.

Also $\exists U\,(x_0)$ mit maximal endlich vielen Punkten $\bar{x} \in M$.

Verkleinere U auf $U' := U \cap B_\varepsilon\,(x_0)$ mit $\varepsilon := \min_{\bar{x} \in U(x_0) \cap M} \|x_0 - \bar{x}\|$.

$\varepsilon > 0$, da maximal endlich viele $\bar{x}$ und $x_0 \notin M$.

Also ist U' Umgebung von x_0.

U' offen $\Rightarrow M^c$ offen $\Rightarrow M$ abgeschlossen. $\square$

Definition 1.14 (*innerer Punkt*)

x_0 heißt *innerer Punkt* von $M \Leftrightarrow x_0 \in M^o$.

Definition 1.15 (dichte Menge. überall dichte Menge)

$N \subset M$ heißt *dicht* in $M \Leftrightarrow \bar{N} \supset M$.

N *überall* dicht $\Leftrightarrow N$ dicht in $\mathbb{R}^n$.

Satz 1.15

$\mathbb{Q}^n$ *überall* dicht (bzw. $\mathbb{Q}^n$ dicht in $\mathbb{R}^n$).

Bemerkung

$\mathbb{Q}^n$ ist abzählbar.

1.4 Kompaktheit

Definition 1.16 (*Überdeckungen*)

Sei M eine Menge, I, J Indexmengen.

Die Familie von Mengen $\mathfrak{M}$ von M heißt *Überdeckung* von M, falls

$$\mathfrak{M} = \{M_i, i \in I\} \text{ mit } \bigcup_{i \in I} M_i \supset M.$$

Die Familie von Mengen $\mathfrak{M}$ von M heißt *offene Überdeckung* von M, falls

$$\mathfrak{M} = \{M_i, i \in I\} \text{ mit } \bigcup_{i \in I} M_i \supset M, \; M_i \text{ offen } \forall i \in I.$$

Die Familie von Mengen $\mathfrak{M}'$ von M heißt *Teilüberdeckung* von M, falls

$$\mathfrak{M}' = \{M_i, i \in J, J \subset I\}.$$

Die Familie von Mengen $\mathfrak{M}$ von M heißt *endliche Überdeckung* von M, falls I endlich ist.

Beispiel Sei $M \subset \mathbb{R}^n$

- mit $M_1 = \mathbb{R}^n$ wird $\mathfrak{M} = \{M_1\}$ eine *endliche* Überdeckung.
- mit $M_i = \mathbb{R}^n, i = 1, 2, 3, \ldots$ wird $\mathfrak{M} = \{M_i, i \in \mathbb{N}\}$ eine *unendliche* Überdeckung.

Definition 1.17 (*beschränkte Menge*)

Sei $M \subset \mathbb{R}^n$. M heißt *beschränkt* $\Leftrightarrow \exists R > 0 \; \forall x \in M \quad \|x\| \leq R$.

Definition 1.18 (*folgenkompakt*)

Sei $K \subset \mathbb{R}^n$ eine Menge. K heißt *folgenkompakt*, wenn jede Folge in K eine konvergente Teilfolge mit Grenzwert in K besitzt:

$$\forall (x_k) \; \exists (x_{j_k}) \quad \lim (x_{j_k}) \in K \text{ mit } x_k \in K.$$

Definition 1.19 (*überdeckungskompakt*)

Sei $K \subset \mathbb{R}^n$ eine Menge. K heißt *überdeckungskompakt*, wenn *jede* offene Überdeckung von K eine endliche Teilüberdeckung besitzt.

Der nun folgende Satz 1.16 spielt eine ganz wesentliche Rolle bei der Topologie metrischer Räume.

Satz 1.16

Sei $K \subset \mathbb{R}^n$ eine Menge. Dann sind äquivalent

1. K ist folgenkompakt.

2. K ist überdeckungskompakt.
3. K ist beschränkt und abgeschlossen.
4. Jede unendliche Menge $M \subset K$ hat mindestens *einen* Häufungspunkt in K.

Beweis

Der Beweis von Satz 1.16 wurde über folgenden Ringschluß: (2) $\Rightarrow$ (4) $\Rightarrow$ (1) $\Rightarrow$ (3) $\Rightarrow$ (2) durchgeführt.

(2) $\Rightarrow$ (4): (indirekt) Annahme: $M \subset K$ hat *keinen* HP $\Rightarrow$ M endlich. Widerspruch.

(4) $\Rightarrow$ (1): Beweis über Fallunterscheidungen $\{x_n \mid n \in \mathbb{N}\}$ endlich bzw. unendlich.

(1) $\Rightarrow$ (3): Beweis durch Kontraposition. K nicht beschränkt $\wedge$ K nicht abgeschlossen $\Rightarrow$ K nicht folgenkompakt.

(3) $\Rightarrow$ (2): (indirekt) Annahme: Zur offenen Überdeckung von K gibt es keine endliche Teilüberdeckung. Mit Hilfe einer Intervallschachtelung durch sukzessive Konstruktion von Würfeln W_k in $\mathbb{R}^n$ gemäß Annahme gelangt man über $W_1 \supset W_2 \supset \ldots$ zu $\bigcap_{k \in \mathbb{N}} W_k = \{x_0\}$ mit $\lim_{k \to \infty} = x_0$. Als Folgerung hieraus ergibt sich der Widerspruch zur Annahme, also deren Negation. □

Bemerkung

Für andere Räume als $\mathbb{R}^n$ kann es durchaus sein, daß zwischen Folgenkompaktheit und Überdeckungskompaktheit keine Äquivalenz besteht.

Satz 1.17 (Heine-Borel-*Eigenschaft*)

Sei $M \subset \mathbb{R}^n$. M ist kompakt $\Leftrightarrow$ M ist beschränkt und abgeschlossen.

Bemerkung

Satz 1.17 entspricht der Äquivalenz 1. $\Leftrightarrow$ 3. aus Satz 1.16.

Satz 1.18

Sei $K, M \subset \mathbb{R}^n$. K ist kompakt $\wedge$ M ist abgeschlossen $\Rightarrow$ $K \cap M$ ist kompakt.

Beweis

i) K ist kompakt $\Rightarrow$ K ist beschränkt $\Rightarrow$ $K \cap M$ ist beschränkt.

ii) K ist kompakt $\Rightarrow$ K ist abgeschlossen $\Rightarrow$ $K \cap M$ ist abgeschlossen.

iii) $K \cap M$ ist beschränkt $\wedge$ $K \cap M$ ist abgeschlossen $\Rightarrow$ $K \cap M$ ist kompakt. □

Satz 1.19

Sei $K_i \subset \mathbb{R}^n$ eine Familie von Mengen, I eine beliebige Indexmenge. Dann gilt

1. $\forall i \in I$ mit I beliebig
   K_i ist kompakt $\Rightarrow$ $\bigcap_{i \in I} K_i$ ist kompakt,

2. $\forall i \in I$ mit I endlich
   K_i ist kompakt $\Rightarrow \bigcup_{i \in I} K_i$ ist kompakt.

Bemerkung

Satz 1.18 und Satz 1.19 beschreiben die Vererbbarkeit der Kompaktheit; sie erschaffen damit ein starkes Instrument für die Konstruktion kompakter Mengen von höherer Komplexität.

1.5 Konvexe Mengen

Beispiele für konvexe Teilmengen des $\mathbb{R}^n$ sind

- die *Gerade* $x = p + ta$ mit $t \in \mathbb{R}, a \neq 0$
- die *Hyperebene* $x = b + t_1 a_1 + \ldots + t_{n-1} a_{n-1}$ mit $t_1, \ldots, t_{n-1} \in \mathbb{R}$ und linear unabhängigen Vektoren $a_1, \ldots, a_{n-1}$.

Definition 1.20 (*konvexe Menge*)

Sei $M \subset \mathbb{R}^n$, $t \in \mathbb{R}$. M heißt *konvex*, wenn

$$\forall x, y \in M \; \forall t \in [0,1] \quad tx + (1-t)\, y \in M.$$

Bemerkung vgl. Def. 3.10 (*konvexe Funktion*).

Satz 1.20

Sei $K_i \subset \mathbb{R}^n$ eine Familie konvexer Mengen, I eine beliebige Indexmenge.
K_i ist konvex $\Rightarrow \bigcap_{i \in I} K_i$ ist konvex.

Beispiele

- Alle Hyperebenen sind konvex.
- Alle affinen Unterräume in $\mathbb{R}^n$ sind konvex (Schnitte von Hyperebenen).
- Halbräume des $\mathbb{R}^n$ sind konvex, z.B. $\{x \mid c_1 x_1 + \ldots + c_n x_n \geq b\}$.
- «Normkugeln» $\{x \mid N(x) \leq r\}$ mit $r \geq 0$ sind konvex.

Satz 1.21 (*Normfunktionen und kompakte konvexe Mengen*)

1. Für beliebige Norm N ist die durch $N(x-a) \leq r$, $r \geq 0$ gegebene Menge von Punkten eine kompakte konvexe Menge, die symmetrisch bzgl. a ist. Im Fall $r > 0$ hat die Menge innere Punkte.
2. Zu jeder beliebigen kompakten konvexen und bzgl. 0 symmetrischen Teilmenge $K \subset \mathbb{R}^n$ mit inneren Punkten existiert genau eine Norm N mit

$$K = \{x \mid N(x) \leq 1\}.$$

Beweis

Betrachte Halbgerade H von 0 durch beliebigen Punkt $x \neq 0$.

$H \cap \partial K = \{x'\}$ ist der H und K gemeinsame Randpunkt.

Für diesen gilt $x' = f(x)x$, wenn $f(x) = \sup\left\{\lambda \in \mathbb{R}_0^+ \mid \lambda x \in K\right\}$ gesetzt und damit zu dem Parameterwert von H wird, der zum Randpunkt x' gehört.

$\forall x \neq 0 \quad f(x) > 0$, da 0 innerer Punkt von K.

Nun soll $N(x)$ auf den Randpunkten x' von K den Wert 1 annehmen, d.h. $1 = N(x') = f(x)N(x)$ bzw. für $x \neq 0$ $N(x) = \frac{1}{f(x)}$.

Es gibt somit zu K höchstens *eine* Norm, für die K Einheitskugel ist.

Bleibt zu zeigen, daß $N(x) = \frac{1}{f(x)}$ für $x \neq 0$ bzw. $= 0$ für $x = 0$ wirklich eine Norm ist, d.h. die Eigenschaften nach Def. 1.2 erfüllt. $\square$

Die durch K bestimmte Norm $N(x)$ muß nicht analytisch darstellbar sein; den Rand von K nennt man auch *Eichfläche* bzw. *Eichkurve*, da durch sie die Vektornorm und damit die Länge von Strecken festgelegt ist.

2 Stetige Abbildungen $\mathbb{R}^n \to \mathbb{R}^m$

Stetigkeit in mehrdimensionalen Räumen bedeutet, die Definitionen für Stetigkeit, die den Begriff der offenen Menge voraussetzen, auf die Dimensionen n bzw. m zu verallgemeinern. Dabei ist es gleich, ob man mit Hilfe des Abstands in metrischen Räumen ($\mathbb{R}$ ist ein metrischer Raum) über ε-Umgebungen zum Begriff der offenen Menge gelangt oder ob man den Begriff der offenen Menge bzw. den der Umgebung axiomatisch in topologischen Räumen erklärt ($\mathbb{R}$ ist ein topologischer Raum).

Wir kennen schon als Beispiele für Abbildungen zwischen Räumen verschiedener Dimension

- die *Norm* $N : \mathbb{R}^n \to \mathbb{R}$ definiert durch $x \mapsto \|x\|$ oder kurz: die Norm $N(x)$
- die *Folge* $f : \mathbb{N} \setminus \{0\} \to \mathbb{R}^n$ definiert durch $n \mapsto f(n) = x_n$ oder kurz: die Folge (x_n)

Es wird sich als zweckmäßig erweisen, für eine Abbildung allgemeiner von einer *Definitionsmenge* $D \subset \mathbb{R}^n$ auszugehen. Von dieser werden häufig die Eigenschaften einer offenen oder einer kompakten Menge verlangt werden. Damit ist folgende Darstellungsweise gegeben:

$$f : D \subset \mathbb{R}^n \to \mathbb{R}^m$$

$$x \mapsto f(x) = \begin{pmatrix} f_1(x) \\ \vdots \\ f_m(x) \end{pmatrix} \in \mathbb{R}^m$$

und die Komponentenschreibweise

$$\begin{array}{rcl} y_1 & = & f_1(x_1, \ldots, x_n) \\ \vdots & & \vdots \\ y_m & = & f_m(x_1, \ldots, x_n) \end{array}$$

Spezialfall 1

Lineare Abbildungen mit den Abbildungsgleichungen

$$\begin{array}{rcccl} y_1 & = & a_{11}x_1 + \cdots + a_{1n}x_n & = & f_1(x_1, \ldots, x_n) \\ \vdots & & \vdots & & \vdots \\ y_m & = & a_{m1}x_1 + \cdots + a_{mn}x_n & = & f_m(x_1, \ldots, x_n) \end{array}$$

lassen sich auch schreiben als $y = Ax$ mit der $m \times n$-Matrix A, die selbst als Abbildung mit dem Definitionsbereich $\mathbb{R}^n$ aufgefaßt werden kann. A ist *injektiv* nur dann, wenn $\dim(\text{Bild}(A)) = n$, hierzu muß die notwendige Bedingung $n \leq m$ erfüllt sein. Zur *Bijektivität* von A ist mindestens notwendig $n = m$, notwendig und hinreichend ist $\det A \neq 0$.

Spezialfall 2

Mit $n \geq 1, m = 1$ führt die vorstehende allgemeine Darstellung zu

$$f : D \subset \mathbb{R}^n \to \mathbb{R}$$
$$x \mapsto y = f(x) = f(x_1, \ldots, x_n) \in \mathbb{R}$$

und damit zu reellwertigen Funktionen mit n reellwertigen Argumenten. Zu deren graphischen Darstellung wird $\mathbb{R}^{n+1}$ verwendet, in welchem die Funktionswerte «über» den entsprechenden Stellen $x \in D \subset \mathbb{R}^n$ «aufgetragen» werden. Beispiel: Darstellung der EUKLIDischen Einheitskugel (Halbkugel) mit $x_1^2 + x_2^2 \leq 1$ durch die Gleichung $y = \sqrt{1 - x_1^2 - x_2^2}$ im $\mathbb{R}^3$.

Definition 2.1 (*Niveaumenge*)

Sei $c \in \mathbb{R}$. Die Menge $\{x \mid f(x) = c\}$ heißt *Niveaumenge*.

Spezialfall 3

Mit $n = 1, m \geq 1$ führt die vorstehende allgemeine Darstellung zu

$$f : D \subset \mathbb{R} \to \mathbb{R}^m$$
$$x \mapsto f(x) = \begin{pmatrix} f_1(x) \\ \vdots \\ f_m(x) \end{pmatrix} \in \mathbb{R}^m$$

und die Komponentenschreibweise

$$\begin{array}{rcl} y_1 & = & f_1(x) \\ \vdots & & \vdots \\ y_m & = & f_m(x) \end{array}$$

Dies entspricht einer «Parameterdarstellung» $(f_1(x), \ldots, f_m(x))$ einer Kurve im $\mathbb{R}^m$.

2.1 Stetigkeit

Definition 2.2 (*stetig im Punkt*)

Sei $D \subset \mathbb{R}^n$, $f : D \to \mathbb{R}^m$.

- f heißt ***stetig im Punkt*** $a \in D$, wenn $\forall \varepsilon > 0\, \exists \delta > 0\, \forall x \in D$ mit $\|x - a\| < \delta$ gilt
$$\|f(x) - f(a)\| < \varepsilon.$$
(wegen Normäquivalenz ist *Norm* beliebig)
- f heißt ***stetig im Punkt*** $a \in D$, wenn es zu jeder *Umgebung* V des Punktes $f(a)$ eine *Umgebung* U des Punktes a gibt mit $f(U \cap D) \subset V$. (Formulierung mit *Umgebungsbegriff*.)
- f heißt ***stetig im Punkt*** $a \in D$, wenn $\lim_{x \to a} f(x) = f(a)$. (Formulierung mit *Grenzwertbegriff*.)

Alle drei Definitionen sind äquivalent.

Definition 2.3 (*stetige Abbildung*)

$\forall x \in D \quad f$ stetig $\Leftrightarrow$ f heißt ***stetig in*** D oder: ***f ist stetig***.

Beispiele

- $f : \mathbb{R}^n \setminus \{0\} \to \mathbb{R}^n \setminus \{0\}$, erklärt durch $f(x) = \frac{1}{\|x\|^2} \cdot x$ mit $x \neq 0$ ist stetig.
- Obere Halbkugel $y = \sqrt{1 - x_1^2 - x_2^2}$ mit $x_1^2 + x_2^2 \leq 1$ ist stetig.
- $f(x) = \begin{cases} 0 & \text{falls } x \in \mathbb{Q}^n \\ 1 & \text{sonst} \end{cases}$ ist nicht stetig.
- $f(x, y) = \begin{cases} \frac{2xy}{x^2+y^2} & \text{falls } x^2 + y^2 > 0 \\ 0 & \text{falls } x = 0, y = 0 \end{cases}$ ist stetig in allen $(x, y) \neq (0, 0)$, nicht stetig in $(0, 0)$, obwohl eine Prüfung auf Stetigkeit in Form einer «partiellen» Stetigkeit über die Geraden $x = 0$ bzw. $y = 0$ letzteres nicht zeigt. Eine Näherung an $(0, 0)$ über die Diagonale ergibt aber, daß $f(x, x) = 1$ für $x \neq 0$. Daher:

Warnung! partielle Stetigkeit $\overset{\not\Rightarrow}{\Leftarrow}$ Stetigkeit

Satz 2.1

$f : D \to \mathbb{R}^m$ stetig, $M \subset D$ beliebig $\Rightarrow$ $f|_M$ stetig.

Satz 2.2

$f = (f_1, \ldots, f_m)^\top$ stetig in $a \in D \Leftrightarrow f_1, \ldots, f_m$ in a stetig.

Beweis

Mit Abschätzung

$$\begin{aligned} |f_i(x) - f_i(a)| &\leq \|f(x) - f(a)\|_\infty \\ &\leq |f_1(x) - f_1(a)| + \ldots + |f_m(x) - f_m(a)| \end{aligned}$$

$\square$

Satz 2.3

Sei $D \subset \mathbb{R}^n$ offen.

f stetig $\Leftrightarrow$ Das Urbild jeder offenen Teilmenge des $\mathbb{R}^m$ ist eine offene Teilmenge des $\mathbb{R}^n$.

Beweis

«$\Rightarrow$» Betrachte Urbild $A = f^{-1}(B)$ irgendeiner offenen Teilmenge B von $\mathbb{R}^m$.

Für jedes $a \in A$ ist f in a stetig.

Nach Def. 2.2 (Formulierung mit Umgebungsbegriff) gibt es zu jeder Umgebung V von $f(a)$ eine offene Umgebung U von a mit $f(U \cap D) \subset V$.

Wegen $a \in U \cap D$ und $U \cap D$ offen ist A offen.

«$\Leftarrow$» Betrachte beliebige offene Umgebung V eines beliebigen Punktes $b = f(a)$.

Dann ist $U = f^{-1}(V)$ offene Teilmenge von $\mathbb{R}^n$.

Für diese Umgebung U von a gilt $U \subset D$ und $f(U) \subset V$, d.h. f ist stetig in a.

Weil $a \in D$ beliebig war, ist f stetig in D. □

2.2 Rechenregeln

Satz 2.4

Sei $f : D \to \mathbb{R}^m$, $g : D \to \mathbb{R}^m$, f, g stetig und $c \in \mathbb{R}$.

$$f + g \quad \text{und} \quad cf \quad \text{sind stetig}$$

Beweis

Punktweise für beliebige $a \in D$.

$$\begin{aligned} \|(f+g)(x) - (f+g)(a)\| &= \|f(x) - f(a) + g(x) - g(a)\| \\ &\leq \|f(x) - f(a)\| + \|g(x) - g(a)\|\,. \\ \|(cf)(x) - (cf)(a)\| &= |c|\,\|f(x) - f(a)\|\,. \end{aligned}$$

Dies entspricht jeweils der Definition der Stetigkeit für f und g. □

Satz 2.5

Sei $f : D \subset \mathbb{R}^n \to \mathbb{R}$, $g : D \subset \mathbb{R}^n \to \mathbb{R}$, f, g stetig.

Dann sind fg stetig, $\frac{f}{g}$ stetig in Punkten a mit $g\,(a) \neq 0$.

Satz 2.6

Sei $\langle\, , \rangle$ Skalarprodukt in $\mathbb{R}^m$, $D \subset \mathbb{R}^n$, $f, g : D \to \mathbb{R}^m$ stetig.

Dann ist $\langle f, g\rangle : D \to \mathbb{R}$ stetig.

Beweis

Mit Satz 2.2 und Satz 2.4 ergibt sich die Stetigkeit unter Benutzung der Einheitsvektoren aus

$$\langle f, g \rangle := \left\langle \sum_i^m f_i e_i, \sum_j^m g_j e_j \right\rangle = \sum_{i,j} \langle e_i, e_j \rangle f_i g_j.$$

□

Satz 2.7

Sei $D \subset \mathbb{R}^n$, $E \subset \mathbb{R}^m$, $f : D \to E$ stetig, $g : E \to \mathbb{R}^k$ stetig.

Dann ist $g \circ f : D \to \mathbb{R}^k$ stetig.

Beweis (mit Folgenkriterium) Sei $x_n \to x_0$.

$$\begin{aligned}(g \circ f)(x_0) &= g(f(\lim_{n\to\infty} x_n)) \\ &= g(\lim_{n\to\infty} f(x_n)) \\ &= \lim_{n\to\infty} g(f(x_n)) \\ &= \lim_{n\to\infty} (g \circ f)(x_n).\end{aligned}$$

□

2.3 Stetige Abbildungen auf kompakten Mengen

Definition 2.4 (*gleichmäßig stetig*)

Sei $D \subset \mathbb{R}^n$. $f : D \to \mathbb{R}^m$ heißt *gleichmäßig stetig* in D, wenn

$$\forall \varepsilon > 0 \, \exists \delta > 0 \, \forall x, y \in D \quad \|x - y\| < \delta \Rightarrow \|f(x) - f(y)\| < \varepsilon.$$

Bemerkung

Unterschied zur normalen Stetigkeit: δ von x *unabhängig*.

Satz 2.8

Sei $D \subset \mathbb{R}^n$ kompakt, $f : D \to \mathbb{R}^m$ stetig.

Dann ist f gleichmäßig stetig.

Beweis

1. Beweismöglichkeit wie zu Satz 3.11 aus Vorlesung ANALYSIS I.

2. Beweismöglichkeit: Sei $\varepsilon > 0$.

Wähle zu $x \in D$ ein $\delta_x > 0$ so, daß $\forall y \in D \ \|x - y\| < \delta_x \Rightarrow \|f(x) - f(y)\| < \frac{\varepsilon}{2}$.

Setze für D als offene Überdeckung $\mathfrak{M} := \left\{ B_{\frac{\delta_x}{2}}(x) \mid x \in D \right\}$. Da D kompakt, existiert endliche Teilüberdeckung $\mathfrak{M}' := \left\{ B_{\frac{\delta_{x'}}{2}}(x') \mid x' \in D' \right\}$ mit $D' \subset D$ endlich.

Setze $\delta := \min\left\{\delta_{\frac{x'}{2}} \mid x' \in D'\right\} > 0$. Seien *nun* $x, y \in D$ mit $\|x-y\| < \delta$.

$\mathfrak{M}'$ überdeckt $D \Rightarrow \exists x' \in D' \ \ \|x-x'\| < \delta_{\frac{x'}{2}}$.

Wegen $\|x'-y\| \leq \|x'-x\| + \|x-y\| < \delta_{\frac{x'}{2}} + \delta \leq \delta_{x'}$ folgt

$\|f(x)-f(y)\| \leq \|f(x)-f(x')\| + \|f(x')-f(y)\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$. □

Satz 2.9

Sei $D \subset \mathbb{R}^n$ kompakt, $f : D \to \mathbb{R}^m$ stetig.

Dann ist $f(D)$ kompakt.

Beweis:

1. Beweismöglichkeit wie Satz 3.10 und Satz 3.15 aus Vorlesung Analysis I.

2. Beweismöglichkeit: Sei $\mathfrak{M} = \{M_i \mid i \in I\}$ offene Überdeckung von $f(D)$.

Zu $i \in I, x \in f^{-1}(M_i)$ existiert eine offene Menge $U_{x,i}$ mit $x \in U_{x,i}$ so, daß $f(U_{x,i} \cap D) \subset M_i$.

Setze $\tilde{M}_i := \bigcup_{x \in f^{-1}(M_i)} U_{x,i}$ und somit $f(\tilde{M}_i \cap D) \subset M_i$.

$\tilde{\mathfrak{M}} := \{\tilde{M}_i \mid i \in I\}$ ist dann offene Überdeckung von D.

Wegen der Kompaktheit von D gibt es eine endliche Teilüberdeckung

$\tilde{\mathfrak{M}}' = \{\tilde{M}_j \mid j \in J\}$ mit $J \subset I$ endlich und hierzu die endliche Teilüberdeckung $\mathfrak{M}' = \{M_j \mid j \in J\}$, d.h. $f(D)$ ist kompakt. □

Korollar 2.10

Stetige reellwertige Funktionen $f : D \to \mathbb{R}$ auf kompakten Mengen sind beschränkt und nehmen Minimum und Maximum an.

2.4 Gleichmäßige Konvergenz

Definition 2.5 (*punktweise Konvergenz. gleichmäßige Konvergenz*)

Sei $D \subset \mathbb{R}^n$, $f, f_n : D \to \mathbb{R}^m$, (f_n) Folge von Abbildungen.

- (f_n) *konvergiert* in D *punktweise* gegen f, i.Z. $f_n \to f, \lim f_n = f$
  $\Leftrightarrow \forall \varepsilon > 0 \, \forall x \in D \, \exists N \, \forall n \geq N \quad \|f_n(x) - f(x)\| < \varepsilon$
- (f_n) *konvergiert* in D *gleichmäßig* gegen f, i.Z. $f_n \rightrightarrows f$
  $\Leftrightarrow \forall \varepsilon > 0 \, \exists N \, \forall x \in D \, \forall n \geq N \quad \|f_n(x) - f(x)\| < \varepsilon$

Bemerkung

Bei punktweiser Konvergenz ist N abhängig von x, bei gleichmäßiger Konvergenz nicht.

Definition (*Menge aller beschränkten Abbildungen*)

$$\mathcal{B}(D, \mathbb{R}^m) := \{f : D \to \mathbb{R}^m \mid f \text{ beschränkt}\}$$

Bemerkung

$\mathcal{B}(D,\mathbb{R}^m)$ ist mit der auf ihr erklärten (Abbildungs-) *Norm*

$$\|f\|_\infty := \sup\{\|f(x)\| \mid x \in D\}$$

ein *normierter* $\mathbb{R}$-*Vektorraum.*

Satz 2.11

Sei $f, f_n \in \mathcal{B}(D,\mathbb{R}^m)$. Dann gilt

$$f_n \rightrightarrows f \Leftrightarrow \|f_n - f\|_\infty \to 0.$$

f_n heißt auch *normkonvergent* gegen f.

Satz 2.12 (CAUCHY*sches Konvergenzkriterium*)

Sei $f_n \in \mathcal{B}(D,\mathbb{R}^m)$. Dann

$$\begin{gathered} \exists f \in \mathcal{B}(D,\mathbb{R}^m) \quad f_n \rightrightarrows f \\ \Leftrightarrow \\ \forall \varepsilon > 0\, \exists N \in \mathbb{N}\, \forall m,n \geq N \quad \|f_n - f_m\|_\infty < \varepsilon. \end{gathered}$$

In $\mathcal{B}(D,\mathbb{R}^m)$ konvergiert jede CAUCHY-Folge.

Beweis:

«⇒» Nach Satz 2.11 gilt $\|f_n - f\|_\infty \to 0$. Mit hinreichend großen m, n gilt die Abschätzung

$\|f_n - f_m\|_\infty \leq \|f_n - f\|_\infty + \|f - f_m\|_\infty < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon > 0.$

«⇐» $\forall x \in D\ \|f_n(x) - f_m(x)\| \leq \|f_n - f_m\|_\infty$ ergibt die in x gleichmäßige Stetigkeit der CAUCHY-Folge $(f_n(x))_{n\in\mathbb{N}}$ im $\mathbb{R}^m$, d.h. $\exists f\ (f_n(x) \to f)$ gleichmäßig in x, d.h. $f_n \rightrightarrows f$. □

Satz 2.13

Sei $f, f_n \in \mathcal{B}(D,\mathbb{R}^m)$, f_n stetig, $f_n \rightrightarrows f$.

Dann ist f *stetig in* D.

Beweis

wie in Vorlesung ANALYSIS I

Bemerkung

Entsprechend für Reihen, WEIERSTRASSsches Konvergenzkriterium o.ä.

2.5 Allgemeinere Konvergenz und Stetigkeit

Konvergenz wurde bisher für *Folgen* (x_n) in $\mathbb{R}^n$ über die Norm erklärt

$$x_n \to x_0 \Leftrightarrow \|x_n - x_0\| \to 0$$

bzw. für *Abbildungsfolgen* (f_n) als punktweise bzw. gleichmäßige Konvergenz mit Hilfe der Maximumsnorm.

Verallgemeinerung der «Konvergenz»

Der Begriff der Konvergenz läßt sich verallgemeinern, wenn ein *normierter* $\mathbb{R}$-*Vektorraum* $(V;N)$ *versehen mit der Normfunktion* $N : V \to \mathbb{R}$ betrachtet wird.

Definition 2.6 (*Normkonvergenz*)

Sei $(V;N)$ ein $\mathbb{R}$-Vektorraum mit der Norm $N : V \to \mathbb{R}$, (v_n) mit $n \in \mathbb{N}$ eine Folge aus V.

$$v_n \to_N v_0 :\Leftrightarrow N(v_n - v_0) \to 0$$

wird als *Normkonvergenz* der Folge (v_n) *bezüglich* N bezeichnet und eine *bezüglich* N normkonvergente CAUCHY-Folge wird erklärt mit

$$\forall \varepsilon > 0\, \exists N \in \mathbb{N}\, \forall m, n \geq N \quad N(v_n - v_m) < \varepsilon.$$

Beispiel

Sei $V = \mathcal{C}^0([0,1])$, $N(v) = \int_0^1 |v(t)|\, dt$,

$$v_n(t) := \begin{cases} 0 & \text{falls} \quad t \leq \frac{1}{2} \\ n\left(t - \frac{1}{2}\right) & \text{falls} \quad \frac{1}{2} < t < \frac{1}{2} + \frac{1}{n} \\ 1 & \text{falls} \quad t \geq \frac{1}{2} + \frac{1}{n} \end{cases}.$$

(v_n) ist CAUCHY-Folge *bezüglich* N, denn $N(v_m - v_n) \leq \frac{1}{\min(m,n)}$.

Aber (v_n) konvergiert nicht, d.h. $\neg\exists v \in \mathcal{C}^0([0,1])$ mit $N(v_n - v) \to 0$, denn es müßten $v|_{[0,\frac{1}{2}]} = 0$ und $v|_{[\frac{1}{2},1]} = 1$ sein.

Es handelt sich also um eine CAUCHY-Folge, die *nicht* konvergiert!

Definition 2.7 (BANACH-*Raum*)

$(V;N)$ heißt *vollständig*, falls jede CAUCHY-Folge in V bzgl. N in V konvergiert.

Ein *vollständiger* normierter Raum heißt BANACH-*Raum*.

Beispiel 1

$(\mathcal{B}(D,\mathbb{R}^m), \|\cdot\|_\infty)$ ist BANACH-Raum.

Beispiel 2

$(\mathbb{R}^n, \|\cdot\|_{\text{beliebig}})$ ist BANACH-Raum.

Weitere Verallgemeinerung der «Konvergenz»

Der Begriff der Konvergenz läßt sich weiter verallgemeinern, wenn z.B. ein *metrischer* $\mathbb{R}$-*Vektorraum* $(V;d)$ *versehen mit der Metrik* $d(v,w)$ anstelle der Norm $N(v-w)$ betrachtet wird.

Allgemein versteht man unter einem *metrischen Raum* $(M; d)$ eine beliebige Menge M zusammen mit einer *Distanzfunktion* $d : M \times M \to \mathbb{R}$, die die Eigenschaften des Abstandes nach Satz 1.4 hat. Jede beliebige Menge M kann zu einem metrischen Raum gemacht werden, z.B. mit der «trivialen» Distanzfunktion

$$d(x, y) = \begin{cases} 0 & \text{falls} \quad x = y \\ 1 & \text{falls} \quad x \neq y \end{cases} .$$

Ein metrischer Raum ist *vollständig*, falls jede CAUCHY-Folge in ihm konvergiert.

Bemerkung

Ein Skalarprodukt $\langle\,,\,\rangle$ induziert eine Norm N über $N(v) := \sqrt{\langle v, v\rangle}$. Eine Norm N induziert eine Metrik d über $d(v, w) := N(v - w)$. Das Umgekehrte gilt nicht.

Noch weitere Verallgemeinerung der «Konvergenz»

Der Begriff der Konvergenz läßt sich noch weiter verallgemeinern, wenn ein *topologischer Raum* $(X; \mathcal{T})$ *versehen mit der Topologie* $\mathcal{T}$ betrachtet wird:

Definition

Sei (x_n) Folge in $(X; \mathcal{T})$. Dann wird erklärt

$$\begin{gathered} x_n \to x :\Leftrightarrow \forall U \in \mathcal{T} \\ \text{mit} \\ x \in U \; \exists N \in \mathbb{N} \, \forall n \geq N \quad x_n \in U. \end{gathered}$$

Bemerkung

CAUCHY-*Kriterium* ist nicht mehr formulierbar.

Warnung!

Im allgemeinen ist *Grenzwert* nicht eindeutig.

Beispiel

Sei X beliebig mit mindestens 2 Elementen, $\mathcal{T} := \{X, \emptyset\}$. Jede Folge (x_n) konvergiert gegen jedes $x \in X$.

Entsprechende Verallgemeinerungen der «Stetigkeit»

- Zwischen *metrischen* Räumen $(V_1; d_1), (V_2; d_2)$ durch das $\varepsilon - \delta$-Kriterium: $f : V_1 \to V_2$ ist stetig in $v \in V_1$ genau dann, wenn

$$\forall \varepsilon > 0 \, \exists \delta > 0 \, \forall w \in V_1 \quad d_1(v, w) < \delta \Rightarrow d_2(f(v), f(w)) < \varepsilon.$$

- Zwischen *topologischen* Räumen $(X_1; \mathcal{T}_1), (X_2; \mathcal{T}_2)$: $f : X_1 \to X_2$ ist stetig in $x \in X_1$ genau dann, wenn

$$\forall U_2 \in \mathcal{T}_2 \text{ mit } f(x) \in U_2 \; \exists U_1 \in \mathcal{T}_1 \text{ mit } x \in U_1 \quad f(U_1) \subset U_2.$$

- oder äquivalent hierzu über das Folgenkriterium unter Nutzung des Konvergenzbegriffs von Folgen für konvergente Folgen (x_n)

$$f\left(\lim_{n\to\infty} x_n\right) \in \lim_{n\to\infty} f(x_n).$$

ÜBERSICHT

$$\begin{array}{cccccc}
 & & \text{Raum} & \mathbb{R}^n & \text{Raum} & \\
V & \text{mit} & \swarrow & & \searrow & \text{mit} \\
 & N+\text{CF} & \text{BANACH} & \leftarrow & \text{HILBERT} & \langle\, ,\, \rangle + \text{CF} \\
 & & \downarrow & & \downarrow & \\
 & N & \text{normierter} & \leftarrow & \text{prä} - \text{HILBERT} & \langle\, ,\, \rangle \\
 & & \downarrow & & & \\
 & d & \text{metrischer} & \searrow & \downarrow & \\
 & & \downarrow & & & \\
 & \mathcal{T} & \text{topologischer} & & \text{linearer} &
\end{array}$$

mit «es gelten» CF Cauchykriterium, $\langle\, ,\, \rangle$ Skalarprodukt, N Normfunktion, d Distanzfunktion, $\mathcal{T}$ Topologie (offene Mengen). Die Pfeile sind als Implikationen zu verstehen.

2.6 PEANO-HILBERT-Kurve

Konstruktion einer stetigen und surjektiven, aber nicht injektiven Abbildung $f : [0,1] \to [0,1] \times [0,1] = [0,1]^2$ bzw. $f : I \to Q$. GIUSEPPE PEANO hat im Jahr 1890 einen Existenzbeweis veröffentlicht: *Mathematische Annalen 36* (1890) S.157-160, DAVID HILBERT diesen durch die Angabe eines Verfahrens ergänzt: *Mathematische Annalen 38* (1891) S.455-460.

Vorgehensweise

- Schritt 1: Unterteile $I = [0,1]$ in 4 gleichgroße Intervalle $I_1, \dots, I_4$. Unterteile $Q = [0,1]^2$ in 4 gleichgroße Quadrate $Q_1, \dots, Q_4$. Definiere die stetige Abbildung $f_1 : I \to Q$ mit $f_1(I_k) \subset Q_k$.
- Schritt 2: Unterteile jedes I_k in 4 gleichgroße Intervalle $I_{k_1}, \dots, I_{k_4}$. Unterteile jedes Q_k in 4 gleichgroße Quadrate $Q_{k_1}, \dots, Q_{k_4}$. Definiere die stetige Abbildung $f_2 : I \to Q$ mit $f_2(I_{k_1 k_2}) \subset Q_{k_1 k_2}$.
- ...
- Schritt n: Unterteile jedes $I_{k_1 \dots k_{(n-1)}}$ in 4 gleichgroße Intervalle $I_{k_1 \dots k_{(n-1)_1}}$, $\dots$, $I_{k_1 \dots k_{(n-1)_4}}$. Unterteile jedes $Q_{k_1 \dots k_{(n-1)}}$ in 4 gleichgroße Quadrate $Q_{k_1 \dots k_{(n-1)_1}}, \dots, Q_{k_1 \dots k_{(n-1)_4}}$. Definiere die stetige Abbildung $f_n : I \to Q$ mit $f_n(I_{k_1 \dots k_n}) \subset Q_{k_1 \dots k_n}$.

Dies führt zur Folge von Funktionen (f_n), d.h. Zerlegungen von I und Q.

Satz 2.14

$f := \lim f_n : I \to Q$ existiert und ist stetig.

Beweis

1. *Stetigkeit* von f.

Zu $\varepsilon > 0$ wähle $n \in \mathbb{N}$ so, daß $2^{-n} < \varepsilon$.

Zu $x \in I$ existiert ein Quadrat $Q_{k_1 \ldots k_n} \ni f_n(x)$.

Wegen der Schachtelung gilt $\forall m \in \mathbb{N}\ f_{n+m}(x) \in Q_{k_1 \ldots k_n}$.

$Q_{k_1 \ldots k_n}$ hat Kantenlänge 2^{-n}, d.h. $\|f_{n+m}(x) - f_n(x)\|_\infty \leq 2^{-n} < \varepsilon$.

$(f_n(x))$ ist daher CAUCHY-Folge.

Da n unabhängig von x ist, ist $(f_n(x))$ gleichmäßig konvergent und nach Satz 2.13 ist $f = \lim f_n$ stetig.

2. *Surjektivität* von f.

Sei $y \in Q$, dann $\exists k_1 k_2 k_3 \ldots$ so, daß $\forall n \in \mathbb{N}\ y \in Q_{k_1 \ldots k_n}$.

Die Intervallfolge $(I_{k_1 \ldots k_n})_{n \in \mathbb{N}}$ bildet eine Intervallschachtelung und legt so $x \in I$ fest.

Da f stetig, gilt $f(x) = y$. □

Bemerkung

Es gibt auch bijektive, allerdings keine stetige *und* bijektive Abbildungen $I \to Q$.

Anwendungen

Existenz und Verfahren zur Anwendung solcher Abbildungen erlaubt eine «Dimensionsreduktion». Anwendungsbereiche sind

- Datenbanken
- Geometrisches Suchen und Sortieren
- Lastbalancierung beim parallelen Behandeln von FE-Netzen.

3 Differentiation von $f : \mathbb{R}^n \to \mathbb{R}^m$

Die Idee der Differenzierbarkeit im Eindimensionalen besteht in der Approximation einer Funktion in einer Umgebung eines festen Punktes durch eine *lineare* Funktion.

3.1 Differenzierbarkeit

Definition 3.1 (*Differenzierbarkeit*)

Sei $U \subset \mathbb{R}^n$ offen. Die Abbildung $f : U \to \mathbb{R}^m$ heißt *differenzierbar in* $a \in U$, wenn es eine *lineare* Abbildung $T : \mathbb{R}^n \to \mathbb{R}^m$ und eine *in a stetige* Abbildung $r : U \to \mathbb{R}^m$ gibt derart, daß

$$f(x) = f(a) + T(x - a) + r(x)\|x - a\|, \quad \forall x \in U$$

und $r(a) = 0$.

Bemerkung

Nach dieser Definition ist f in a differenzierbar, wenn eine solche lineare Abbildung T angegeben werden kann, daß mit ihr für die Abbildung

$$r(x) := \frac{f(x) - f(a) - T(x-a)}{\|x-a\|}$$

mit $x \neq a$

$$\lim_{x \to a} r(x) = 0.$$

Hierbei ist $r(x)$ eindeutig bzgl T und der Limes hängt nicht von der Wahl der Norm ab.

Satz 3.1

Es kann höchstens eine solche lineare Abbildung T geben.

Beweis

(indirekt) Mit zwei *linearen* Abbildungen T_1 und T_2 und den ihnen zugehörigen *stetigen* Abbildungen r_1 und r_2 gilt wegen

$f(x) = f(a) + T_1(x-a) + r_1(x)\|x-a\|$ und

$f(x) = f(a) + T_2(x-a) + r_2(x)\|x-a\|$ die Gleichung

$(T_1 - T_2)(x-a) = (r_2 - r_1)(x)\|x-a\|$ sowie

$\lim_{x \to a} r_1(x) = 0$ und $\lim_{x \to a} r_2(x) = 0$.

Mit beliebigem h und genügend kleinem $t > 0$ kann gesetzt werden: $x - a =: th$.

Hieraus ergeben sich

$(T_1 - T_2)h = (r_2(a+th) - r_1(a+th))\|h\|$ und wegen

$\lim_{t \to 0}(r_2(a+th) - r_1(a+th)) = 0$

$(T_1 - T_2)h = 0$ für $h \in \mathbb{R}^n$, d.h. $T_1 = T_2$. $\square$

Definition 3.2 (*Ableitung*)

Sei f in $a \in U \subset \mathbb{R}^n$ differenzierbar. Die *lineare* Abbildung T heißt dann *Ableitung von f an der Stelle a*, i.Z. $T = f'(a)$.

Mit $h \in \mathbb{R}^n$ wird der Vektor $f'(a)h \in \mathbb{R}^m$ auch mit $f'(a,h)$ bzw. Th bezeichnet.

Definition 3.3 (*Gradient*)

Sei $U \subset \mathbb{R}^n$, $f : U \to \mathbb{R}$ eine reellwertige differenzierbare Funktion.

$T = f'(a) = \mathrm{grad} f(a)$ heißt *Gradient* von f an der Stelle a und es gilt

$$f'(a,h) = \langle \mathrm{grad} f(a), h \rangle\,, \quad \forall h \in \mathbb{R}^n$$

Bem.: Weitere Schreibweise für grad ist ∇.

Bemerkung

Nach Def. gilt $f(a+h) - f(a) = \langle \mathrm{grad} f(a), h \rangle + r(a+h) \|h\|$ mit $\lim_{h \to 0} r(a+h) = 0$.

Geometrische Interpretation:

- Für hinreichend kleines $\|h\|$ ist der Zuwachs $f(a+h) - f(a)$ durch das Skalarprodukt $\langle \mathrm{grad} f(a), h \rangle$ gut wiedergegeben. Geometrisch gedeutet zeigt $\mathrm{grad} f(a)$ in die Richtung des größten Zuwachses der Funktion f an der Stelle a, da $\langle \mathrm{grad} f(a), h \rangle$ seinen größten Wert als Funktion von h annimmt, wenn h und $\mathrm{grad} f(a)$ gleichgerichtet sind, d.h. h ein positives Vielfaches von $\mathrm{grad} f(a)$ ist.
- Betrachtet man die *Niveaumengen* von f, d.h. $f(x) = c$ (Niveauhyperflächen), so schmiegen sich die h, wenn $h \perp \mathrm{grad} f(a)$, an die Niveaulinien an. Sie spannen die Tangentialhyperebene an die Niveauhyperfläche durch den Punkt a auf, der Vektor $\mathrm{grad} f(a) \in \mathbb{R}^n$ steht senkrecht auf dieser und damit senkrecht auf der Niveauhyperfläche.
- Betrachtet man den *Graph* von f als Hyperfläche im Raum $(x_1, \ldots, x_n, y)$ mit $y = f(x_1, \ldots, x_n)$, dann wird die Hyperebene (Tangentialhyperebene) approximiert mit $y = f(a) + \langle \mathrm{grad} f(a), x - a \rangle$. $\begin{pmatrix} -\mathrm{grad} f(a) \\ 1 \end{pmatrix} \in \mathbb{R}^{n+1}$ ist Normalenvektor der Tangentialhyperebene und damit der Normalenvektor der Hyperfläche in $(a, f(a))$.

Bestimmung von $T = f'(a)$

Es gilt nach Def. 3.1

$$f(x) = f(a) + T(x-a) + r(x) \|x - a\|, \quad \forall x \in U.$$

Sei $v_1, \ldots, v_n$ sei eine Basis des $\mathbb{R}^n$, zu der die Bildvektoren der linearen Abbildung $f'(a)$ gesucht werden. Bestimmt man zunächst für einen beliebigen Vektor v den Vektor $f'(a, v)$ und setzt hierzu $x = a + tv$ mit $t \in \mathbb{R}, t \neq 0$ in die Def. ein, so wird

$$\begin{array}{lcl} f(a+tv) & = & f(a) + f'(a, tv) + r(a+tv) \, |t| \, \|v\| \\ f'(a, v) & = & \frac{f(a+tv) - f(a)}{t} - r(a+tv) \frac{|t|}{t} \|v\| \\ \text{also} & & \\ f'(a, v) & = & \lim_{t \to \infty} \frac{f(a+tv) - f(a)}{t} \end{array}$$

Definition 3.4 (*Richtungsableitung*)

$D_v f(a)$ heißt *Richtungsableitung* von f an der Stelle a bezüglich eines Basisvektors v. Es gilt

$$D_v f(a) = \lim_{t \to 0} \frac{f(a+tv) - f(a)}{t}.$$

Bemerkung

$f'(a)$ ist durch die *Richtungsableitungen* bzgl. der Basisvektoren $v_1, \ldots, v_n$ des $\mathbb{R}^n$ festgelegt.

Definition 3.5 (*partielle Ableitungen*)

$D_i f(a) := D_{e_i} f(a)$ sind die Richtungsableitungen von f an der Stelle a bzgl. der *Standardbasisvektoren* e_i des $\mathbb{R}^n$. Sie heißen *partielle Ableitungen*. Es gilt

$$D_i f(a) = \lim_{x_i \to a_i} \frac{f(a_1, \ldots, x_i, \ldots, a_n) - f(a_1, \ldots, a_i, \ldots, a_n)}{x_i - a_i}.$$

Andere Schreibweisen für partielle Ableitungen sind $\frac{\partial f}{\partial x_i}$ bzw. f_{x_i}

Bemerkung

Damit ist die Bestimmung der Ableitung auf die Differentialrechnung im Eindimensionalen zurückgeführt.

Bemerkung

Limites nach Def. 3.5 können existieren, ohne daß f insgesamt differenzierbar ist:

partiell differenzierbar $\overset{\not\Rightarrow}{\Leftarrow}$ differenzierbar.

Ist f differenzierbar, so existieren alle partiellen Ableitungen und es gilt mit $h = \sum_{i=1}^n h_i e_i$

$$f'(a,h) = \sum_{i=1}^n f'(a, e_i) h_i = \sum_{i=1}^n D_i f(a) h_i.$$

Die weitere Darstellung der $D_i f(a)$ mittels der Standardbasis des $\mathbb{R}^m$, d.h. die weitere Zerlegung von f in die Komponenten $f_1, \ldots, f_m$ führt zur Darstellung von $f'(a)$ als $m \times n$-Matrix

Definition 3.6 (Jacobi-*Matrix*)

Die Darstellung der partiellen Ableitungen als $m \times n$-Matrix

$$f'(a) := \begin{pmatrix} D_1 f_1(a) & D_2 f_1(a) & \ldots & D_n f_1(a) \\ D_1 f_2(a) & D_2 f_2(a) & \ldots & D_n f_2(a) \\ \vdots & \vdots & & \vdots \\ D_1 f_m(a) & D_2 f_m(a) & \ldots & D_n f_m(a) \end{pmatrix}$$

heißt Jacobi-*Matrix* der differenzierbaren Abbildung f an der Stelle a.

Prüfung auf Differenzierbarkeit

$$(\star)\ f(x) = f(a) + T(x-a) + r(x)\|x-a\|$$

Vorgehensweise

- Untersuche, ob die $D_i f_j(a)$ existieren. Existenz ist notwendige Bedingung.
- Prüfe mit der Matrix $T = f'(a) = (D_i f_j(a))$, ob

$$\lim_{h \to 0} r(h) = \lim_{h \to 0} \frac{f(a+h) - f(a) - Th}{\|h\|} = 0.$$

Beispiel 1

Sei $f : \mathbb{R}^2 \to \mathbb{R}^2$ gegeben durch $f_1(x_1, x_2) = x_1^2 - x_2^2$ und $f_2(x_1, x_2) = 2x_1x_2$.

Ist f an der Stelle a differenzierbar, dann existiert die Jacobi-Matrix, hier $\begin{pmatrix} 2a_1 & -2a_2 \\ 2a_2 & 2a_1 \end{pmatrix}$.

Mit

$$r_1(h) = \frac{(a_1+h_1)^2 - (a_2+h_2)^2 - (a_1^2 - a_2^2) - (2a_1h_1 - 2a_2h_2)}{\|h\|} = \frac{h_1^2 - h_2^2}{\|h\|}, \text{ und}$$

$$r_2(h) = \frac{2(a_1+h_1)(a_2+h_2) - 2a_1a_2 - (2a_2h_1 + 2a_1h_2)}{\|h\|} = \frac{2h_1h_2}{\|h\|}$$

werden unter Verwendung der Maximumsnorm $\|h\|_\infty = \max\{|h_1|, |h_2|\}$ die Abschätzungen $|r_1(h)| \leq 2\|h\|_\infty$ und $|r_2(h)| \leq 2\|h\|_\infty$ erhalten. Somit ist

$$\lim_{h \to 0} r_1(h) = \lim_{h \to 0} r_2(h) = 0.$$

f ist also in a differenzierbar mit $f'(a) =$ der vorstehenden Jacobi-Matrix.

Beispiel 2

Für eine affine Abbildung $f : \mathbb{R}^n \to \mathbb{R}^m$ gilt $f(x) = Ax + c$ mit $A : \mathbb{R}^n \to \mathbb{R}^m$ linear und $c \in \mathbb{R}^m$ konstant. Dann ist

$$f(a+h) = A(a+h) + c = Aa + Ah + c = Aa + c + Ah = f(a) + Ah.$$

Daraus folgt $r(h) = 0$ für $\forall h \in \mathbb{R}^n$, d.h. f ist differenzierbar an allen Stellen $a \in \mathbb{R}^n$ und $f'(a) = A$.

Beispiel 3

$f : \mathbb{R}^2 \to \mathbb{R}$ sei erklärt durch

$$f(x_1, x_2) = \begin{cases} 0 & \text{für } (x_1, x_2) = (0,0) \\ \frac{x_1x_2}{\sqrt{x_1^2 + x_2^2}} & \text{für } (x_1, x_2) \neq (0,0) \end{cases} .$$

Obgleich an $a = (0,0)$ die partiellen Ableitungen existieren, ist f an dieser Stelle *nicht* differenzierbar.

Wegen $\forall x_1 \in \mathbb{R}\ f(x_1, 0) = 0$ und $\forall x_2 \in \mathbb{R}\ f(0, x_2) = 0$ existieren $D_1 f(0) = 0$ und $D_2 f(0) = 0$.

Unter der Annahme der Differenzierbarkeit von f an $a = 0$ ist T die Nullabbildung und man erhält mit $f(a) = 0$ und $Th = 0$

$$r(h) = \frac{h_1h_2}{\sqrt{h_1^2 + h_2^2}} \frac{1}{\|h\|}$$

mit der Bedingung $\lim_{h\to 0} r(h) = 0$.

Ohne Einschränkung kann $\|h\| = \|h\|_2$, der EUKLIDischen Norm, gesetzt werden. Dann wird $r(h) = \frac{h_1 h_2}{h_1^2+h_2^2}$ und mit $h_1 = h_2$ zu $r(h) = \frac{1}{2}$, d.h. $\lim_{h\to 0} r(h) \neq 0$.

Beispiel 4

$f : \mathbb{R}^2 \to \mathbb{R}$ sei erklärt durch

$$f(x_1, x_2) = \begin{cases} 0 & \text{für } (x_1, x_2) = (0,0) \\ \frac{x_1 x_2^3}{x_1^2+x_2^6} & \text{sonst} \end{cases} .$$

Es existieren an der Stelle $a = 0$ alle Richtungsableitungen (mit dem Wert 0); trotzdem ist f in a nicht differenzierbar (nicht einmal stetig).

Unter Verwendung der EUKLIDischen Norm wird $r(h) = \frac{h_1 h_2^3}{h_1^2+h_2^6} \frac{1}{\sqrt{h_1^2+h_2^2}}$ und mit z.B. $h_1 = t^3, h_2 = t$ erhält man

$$r(t^3, t) = \frac{t^6}{2t^6\sqrt{t^6+t^2}} = \frac{1}{2\,|t|} \frac{1}{\sqrt{1+t^4}}.$$

Somit erweist sich r in keiner Umgebung der Stelle a als beschränkt.

Bemerkung

Die Existenz von Richtungsableitungen ist weder für Stetigkeit, noch für Differenzierbarkeit hinreichend.

Satz 3.2

f ist in (an der Stelle) a differenzierbar $\Rightarrow$ f ist in a stetig.

Beweis

$$\begin{aligned} f(x) &= f(a) + T(x-a) + r(x)\,\|x-a\| \\ &= f(a) + f'(a, x-a) + r(x)\,\|x-a\|. \end{aligned}$$

Es gilt daher die Abschätzung

$$\|f(x) - f(a)\| \leq \|f'(a)\|\,\|x-a\| + \|r(x)\|\,\|x-a\|.$$

Wegen Differenzierbarkeit gilt $\lim_{x\to a} r(x) = 0$, und es wird

$$\|f(x) - f(a)\| \leq \|f'(a)\|\,\|x-a\|.$$

Bei geeigneter Wahl einer Konstanten c mit $c\,\|x-a\|$ entspricht die vorstehende Ungleichung der Definition für Stetigkeit nach Def. 2.2. $\square$

Bemerkung zum Gradienten

$$f'(a,h) = \langle \mathrm{grad} f(a), h\rangle\,, \quad \forall h \in \mathbb{R}^n.$$

Der Vektor $\mathrm{grad}f(a)$ ist dadurch eindeutig in $\mathbb{R}^n$.

Bisher wurde das Skalarprodukt in $\mathbb{R}^n$ als positiv definite symmetrische Bilinearform (vgl. Satz 1.1) betrachtet, was an (kanonischen) Schreibweisen erlaubt:

$$\mathrm{grad} f(a) = \begin{pmatrix} D_1 f(a) \\ \vdots \\ D_n f(a) \end{pmatrix} = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{pmatrix} = \begin{pmatrix} f_{x_1} \\ \vdots \\ f_{x_n} \end{pmatrix} := \nabla f(a).$$

Dies führt zur Anwendung des sog. Nabla-Operators ∇ als «formalen» Vektor (vektorwertigen Differentialoperator) im Nabla-Kalkül (vgl. Kapitel 3.8). Für die Einheitsvektoren der Standardbasis, z.B. des $\mathbb{R}^3$, gilt die Matrix

$$(\langle e_i, e_j\rangle) = \mathrm{Diag}(1,1,1).$$

Dagegen wird ein «Skalarprodukt» als *nicht ausgeartete (nicht entartete)* symmetrische Bilinearform als MINKOWSKI-Form (MINKOWSKImetrik) in der allgemeinen Relativitätstheorie verwendet. Hier gilt für die entsprechende Basis des $\mathbb{R}^4$ die Matrix

$$\left(\langle e_i, e_j\rangle_M\right) = \mathrm{Diag}(1,1,1,-1)$$

und für den Gradienten

$$\mathrm{grad}_M f(a) = \begin{pmatrix} f_{x_1} \\ f_{x_2} \\ f_{x_3} \\ -f_{x_4} \end{pmatrix}.$$

3.2 Rechenregeln

Um Beweisführungen zugänglicher zu machen, ist es zweckmäßig, die Differenzierbarkeitsdefinition aus Def. 3.1

$$(\star)\ f(x) = f(a) + T(x-a) + r(x)\,\|x-a\|$$

$$\text{mit } \lim_{x\to a} r(x) = 0$$

in die Form

$$(\star\star)\ f(a+h) = f(a) + f'(a,h) + r(a+h)\,\|h\|$$

$$\text{mit } \lim_{h\to 0} r(a+h) = 0$$

umzuschreiben.

Satz 3.3

Sei $f : U \to \mathbb{R}^m, g : U \to \mathbb{R}^m$ mit $U \subset \mathbb{R}^n$ offen in $a \in U$ differenzierbar, $c \in \mathbb{R}$.

Dann sind auch $f + g$ und cf differenzierbar in a und es gilt

$$\begin{aligned} (f+g)'(a) &= f'(a) + g'(a) \\ (cf)'(a) &= cf'(a) \end{aligned}$$

Satz 3.4 (*Produktregel für reellwertige Funktion*)

Sei $f : U \to \mathbb{R}, g : U \to \mathbb{R}$ mit $U \subset \mathbb{R}^n$ offen in $a \in U$ differenzierbar.

Dann ist auch fg differenzierbar in a und es gilt

$$(fg)'(a,h) = f(a)\, g'(a,h) + f'(a,h)\, g(a).$$

Beweis

Multiplizieren von f und g in der Form $(\star\star)$ ergibt

$$\begin{aligned} (fg)(a+h) = & \quad f(a)\, g(a) + f(a)\, g'(a,h) + g(a)\, f'(a,h) \\ & + f'(a,h)\, g'(a,h) \\ & + (g(a) + g'(a,h))\, r_f(a,h)\, \|h\| \\ & + (f(a) + f'(a,h))\, r_g(a,h)\, \|h\| \\ & + r_f(a,h)\, r_g(a,h)\, \|h\|^2 \end{aligned}$$

Abschätzung von $|f'(a,h)\, g'(a,h)| \leq c\|h\|^2$, Zusammenfassung der letzten sechs Summanden (letzten vier Zeilen der Gleichung) zu $r(h)\, h$ mit $\lim_{h\to 0} r(h) = 0$ und Übernahme der Schreibweise aus dem Satz ergibt

$$(fg)(a+h) = (fg)(a,h) + (fg)'(a,h) + r(h)\, h.$$

$\square$

Bemerkung

Die Produktregel für *reellwertige* Funktion läßt sich auch schreiben als

$$\mathrm{grad} fg(a) = f(a)\, \mathrm{grad} g(a) + g(a)\, \mathrm{grad} f(a).$$

Satz 3.4 bezieht sich auf rellwertige Funktionen und es ist möglich, dies zu modifizieren; dazu zwei Fälle:

Fall 1

Produkt *reellwertiger* Funktion $f : U \subset \mathbb{R}^n \to \mathbb{R}$ mit *vektorwertiger* Abbildung $g : U \subset \mathbb{R}^n \to \mathbb{R}^m$.

Bei komponentenweiser Multiplikation ergibt sich die Regel

$$(fg)'(a,h) = f(a)\, g'(a,h) + f'(a,h)\, g(a).$$

Fall 2

Produkt *vektorwertiger* Abbildungen $f : U \subset \mathbb{R}^n \to \mathbb{R}^p$ und $g : U \subset \mathbb{R}^n \to \mathbb{R}^q$.

An die Stelle des Produkts zweier reellwertiger Funktionen tritt als bilineare Abbildung $[\ ,\ ] : \mathbb{R}^p \times \mathbb{R}^q \to \mathbb{R}^m$ als Produkt

$$[f, g]'(a, h) = [f(a), g'(a, h)] + [f'(a, h), g(a)].$$

Es ist auf die Reihenfolge zu achten, da die bilineare Abbildung nicht unbedingt symmetrisch ist.

Satz 3.5

Sei $f : U \subset \mathbb{R}^n \to \mathbb{R}$ mit U offen in $a \in U$ differenzierbar mit $f(a) \neq 0$.

Dann ist

$$\left(\frac{1}{f}\right)'(a, h) = -\left(\frac{1}{(f(a))^2}\right) f'(a, h).$$

Bemerkung

Diese Regel läßt sich auch schreiben als

$$\operatorname{grad}\left(\frac{1}{f}\right)(a) = -\left(\frac{1}{(f(a))^2}\right) \operatorname{grad} f(a).$$

Bemerkung

Die Regel für $\left(\frac{g}{f}\right)'(a, h)$ erhält man über den Ansatz $g(a)\left(\frac{1}{f}\right)(a)$ und Anwendung der Produktregel.

Satz 3.6 (*Kettenregel*)

Sei $f : U \subset \mathbb{R}^n \to \mathbb{R}^m$, U offen, $g : V \subset \mathbb{R}^m \to \mathbb{R}^p$, V offen, in $a \in U$ bzw. in $b \in V$ differenzierbare Abbildungen und $f(U) \subset V$ und $b = f(a)$.

Dann gilt

1. $g \circ f : U \to \mathbb{R}^p$ ist differenzierbar in a
2. $(g \circ f)'(a) = g'(f(a)) f'(a)$ bzw. $(g \circ f)'(a, h) = g'(f(a), f'(a, h))$

Für die Jacobi-Matrizen von Abbildungen f, g und ihrer Verkettung $H = g \circ f$ gilt

$$\begin{pmatrix} D_1 H_1(a) & \dots & D_n H_1(a) \\ \vdots & & \vdots \\ D_1 H_p(a) & \dots & D_n H_p(a) \end{pmatrix} =$$

$$= \begin{pmatrix} D_1 g_1(b) & \dots & D_m g_1(b) \\ \vdots & & \vdots \\ D_1 g_p(b) & \dots & D_m g_p(b) \end{pmatrix} \begin{pmatrix} D_1 f_1(a) & \dots & D_n f_1(a) \\ \vdots & & \vdots \\ D_1 f_m(a) & \dots & D_n f_m(a) \end{pmatrix},$$

somit wird

$$D_i H_j(a) = \sum_{k=1}^{m} D_k g_j(b) \cdot D_i f_k(a).$$

Ist g *reellwertig*, d.h. $p = 1$, wird auch H reellwertig und

$$\begin{aligned} D_i H(a) &= \sum_{k=1}^{m} D_k g(b) D_i f_k(a) \\ &= \sum_{k=1}^{m} D_i f_k(a) D_k g(b), \quad \text{d.h.} \\ \operatorname{grad} H &= (D_i f_k(b))^\top \operatorname{grad} g. \end{aligned}$$

Für die *Ableitung der Umkehrabbildung einer bijektiven Abbildung* gilt wegen der Kettenregel

$$\begin{aligned} g \circ f = \mathrm{id}_{\mathbb{R}^n} &\Rightarrow g'(b)\, f'(a) = \mathrm{id}_{\mathbb{R}^{n\times n}}, \quad \text{d.h.} \\ g'(f(a)) &= (f'(a))^{-1}. \end{aligned}$$

Vorausgesetzt ist, daß $f : U \to \mathbb{R}^n$ eine Umkehrabbildung $g : V \to \mathbb{R}^n$, V offen, besitzt und daß f in $a \in U$ und g in $b = f(a)$ *differenzierbar* sind.

Das Ergebnis der Anwendung der Kettenregel auf die Ableitung der Umkehrabbildung besteht schon, wenn g in $b = f(a)$ *lediglich stetig* ist, weil daraus die Differenzierbarkeit folgt:

Satz 3.7

Sei $U, V \subset \mathbb{R}^n$ offen, $f : U \to \mathbb{R}^n$ eine in $a \in U$ *differenzierbare* Abbildung, $g : V \to \mathbb{R}^n$ ihre in $b = f(a)$ *stetige* Umkehrabbildung und $f'(a)$ die Jacobi-Matrix.

Ist die Jacobi-Determinante $\det f'(a) \neq 0$, so ist g in b *differenzierbar* und $g'(b) = (f'(a))^{-1}$.

Beweis

O.B.d.A. gelte $a = 0, f(a) = 0$.

Nach Voraussetzung ist f in 0 differenzierbar.

Es gibt dann (analog der Form ($\star\star$) für die Darstellung der Differenzierbarkeitsdefinition) für f eine Darstellung

$$f(x) = F(x)\, x \quad \text{mit} \quad \lim\nolimits_{x\to 0} F(x) = f'(0) \ ,$$

wobei $F(x)$ linear und in $x = 0$ stetig ist, d.h. $F(0) = f'(0)$.

Wegen der Voraussetzung $\det F(0) = \det f'(0) \neq 0$ gibt es in einer geeigneten Umgebung $W \subset V$ von 0 die inversen linearen Abbildungen $(F(x))^{-1}$.

Die Berechnung der inversen Matrix mit Determinanten (Cramersche Regel) zeigt, daß $\lim_{x\to 0} (F(x))^{-1} = (f'(0))^{-1}$.

Aus $y = F(x)\, x$ mit $x \in W$ gewinnt man die äquivalente Gleichung

$$x = (F(x))^{-1}\, y.$$

Durch Einsetzen von $x = g(y)$ und Erklärung der dadurch entstehenden

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

erhält man $g(y) = G(y)\, y$ mit $\lim_{y\to 0} G(y) = (f'(0))^{-1}$ als Darstellung für die Umkehrabbildung g für $y \in W$, wobei die Stetigkeit von g in 0 verwendet ist.

Somit ist g in $y = 0$ differenzierbar und es ist $g'(0) = (f'(0))^{-1}$. □

Bemerkung (andere Formulierung von Satz 3.7)

f^{-1} (total) diff'bar $\wedge \left(f^{-1}\right)' = (f')^{-1} \circ f^{-1} \Rightarrow f^{-1}$ stetig diff'bar.

Dabei sind $(f')^{-1}$ und f^{-1} jeweils stetig

3.3 Mittelwertsatz

Im Falle des Eindimensionalen postuliert der Mittelwertsatz zu gegebenen Punkten a und b die Existenz eines zwischen a und b liegenden Punktes c mit gewissen Eigenschaften. Setzt man im Mehrdimensionalen voraus, daß die Definitionsmenge D einer Abbildung *konvex* ist, dann ist sichergestellt, daß falls ein c auf der Verbindungsstrecke von a und b liegt, $c \in D$ gilt. Der Mittelwertsatz lautet dann:

Satz 3.8 (*Mittelwertsatz)*

Sei $U \subset \mathbb{R}^n$ konvex, $f : U \to \mathbb{R}$ differenzierbar.

$$\exists c \in \overline{a, b} \quad f(b) - f(a) = f'(c, b-a).$$

Beweis

Die Verbindungsstrecke wird durch $\gamma(t) = a + t(b-a)$ mit $0 \le t \le 1$ beschrieben.

Diese differenzierbare Abbildung hat die konstante Ableitung $\gamma'(t) = b - a$.

Für die Funktion $F : [0,1] \to \mathbb{R}$, definiert durch $F(t) = f(\gamma(t))$ gibt es bei Anwendung des «gewöhnlichen» Mittelwertsatzes eine Stelle $\vartheta \in]0,1[$ so, daß $F(1) - F(0) = F'(\vartheta)$.

Mit der Kettenregel erhält man $F'(t) = f'(\gamma(t), \gamma'(t))$.

Wegen $F(0) = f(a)$, $F(1) = f(b)$ ist $F'(\vartheta) = f(b) - f(a)$ die linke Seite,

mit $t := \vartheta$ und $\gamma(\vartheta) =: c$ als Punkt der Verbindungsstrecke ist $\gamma'(\vartheta) = b - a$ und $f'(\gamma(t), \gamma'(t))$ wird zur rechten Seite der Behauptung. □

Bemerkung

Ersetzt man die Verbindungsstrecke durch einen beliebigen in U verlaufenden differenzierbaren Weg, so erhält man $f(b) - f(a) = f'(\gamma(\vartheta), \gamma'(\vartheta))$ mit $0 < \vartheta < 1$. $b - a$ ist also durch $\gamma'(\vartheta)$ ersetzt.

Bemerkung

Allgemeiner bedeutet das Betrachten einer in $U \subset \mathbb{R}^n$ erklärten reellwertigen Funktion f längs eines in U verlaufenden Weges γ das Betrachten der Verkettung $F := f \circ \gamma$ von $\gamma : [a, b] \to U$ mit $f : U \to \mathbb{R}$. Sind γ und f differenzierbar, so gilt

$$F'(t) = f'(\gamma(t), \gamma'(t)) = \langle \mathrm{grad} f(\gamma(t)), \gamma'(t) \rangle .$$

Bemerkung

Umbenennungen führen zu einer anderen Darstellung des Mittelwertsatzes:

$$f(x+h) - f(x) = f'(x + \vartheta h, h) \quad \text{mit} \quad 0 < \vartheta < 1.$$

Für *Abschätzungen* des Zuwachses von $f(x+h) - f(x)$ ist z.B. die Kenntnis eines Zwischenpunktes nicht unbedingt erforderlich. Eine solche Abschätzung erhält man unter den Bedingungen des folgenden Satzes:

Satz 3.9 (*Schrankensatz*)

Sei $U \subset \mathbb{R}^n$ konvex, $f : U \to \mathbb{R}$ differenzierbar. Ist $\mathrm{grad} f$ auf der Strecke von x nach $x + h$ beschränkt, dann gilt

$$|f(x+h) - f(x)| \le S \|h\| ,$$

mit $S \le \|\mathrm{grad} f(x + th)\|$ für alle $0 \le t \le 1$.

Beweis

Aus $f(x+h) - f(x) = f'(x + \vartheta h, h) = \langle \mathrm{grad} f(x + \vartheta h), h \rangle$ mit $0 < \vartheta < 1$ folgt wegen $|f(x+h) - f(x)| \le \|\mathrm{grad} f(x + \vartheta h)\| \, \|h\|$ das Behauptete. $\square$

Bemerkung vgl. Satz 3.29

Satz 3.10 (*Charakterisierung der konstanten Funktion*)

Sei $U \subset \mathbb{R}^n$ offen und wegzusammenhängend, $f : U \to \mathbb{R}$ differenzierbar.

Die Bedingung

$$\forall x \in U \, \forall h \in \mathbb{R}^n \quad f'(x, h) = 0$$

ist notwendig und hinreichend dafür, daß f konstant ist, d.h.

$$f'(x, h) = 0 \Leftrightarrow f(x) = c.$$

Beweis:

«$\Leftarrow$» (notwendig) offensichtlich.

«$\Rightarrow$» (hinreichend):

Seien zwei Punkte a, b durch einen Streckenzug $\overline{a = a_0, a_1, \ldots, a_{k-1}, a_k = b}$ verbunden. Anwendung des Mittelwertsatzes auf Strecken $\overline{a_{j+1}, a_j}$ führt unter Beachtung von $f'(c_j, a_{j+1} - a_j) = 0$ zu den Gleichungen $f(a_{j+1}) = f(a_j)$ für $j = 0, 1, \ldots, k-1$, d.h. $f(b) = f(a)$. $\square$

Bemerkung

$$f'(x,h)=0 \Leftrightarrow D_1 f(x)=0,\dots,D_n f(x)=0 \Leftrightarrow \mathrm{grad} f(x)=0$$

Bekanntlich ist es für die Differenzierbarkeit einer Abbildung f an einer Stelle a nicht hinreichend, daß in a alle partiellen Ableitungen existieren. Sind diese jedoch stetig, dann folgt daraus die Differenzierbarkeit von f in a. Da sich die Eigenschaft der Stetigkeit relativ leicht prüfen läßt, ist eine solche Aussage über die Differenzierbarkeit für Anwendungen wichtig.

Satz 3.11 (*Stetigkeit und Differenzierbarkeit*)

Sei $U \subset \mathbb{R}^n$, $f: U \to \mathbb{R}$ besitze sämtliche partiellen Ableitungen $D_1 f,\dots,D_n f$.

Sämtliche $D_i f$ sind in a stetig $\Rightarrow$ f ist in a differenzierbar.

Beweis

Zu zeigen ist

$$\lim_{h\to 0}\frac{f(a+h)-f(a)-f'(a,h)}{\|h\|}=0.$$

Setze für die Ableitung $f'(a,h)=D_1 f(a)\,h_1+\cdots+D_n f(a)\,h_n$ die Unterteilung von $a, a+h$ an mit

$$\begin{array}{rcl} a_0 &=& a,\\ a_1-a_0 &=& h_1 e_1,\\ a_2-a_1 &=& h_2 e_2,\\ \vdots && \vdots\\ a_n-a_{n-1} &=& h_n e_n. \end{array}$$

Entwickle $f(a+h)-f(a)$ als Teleskopsumme und wende jeweils den Mittelwertsatz auf die «Einzelsummanden» an, d.h. zu jeder «Strecke» $\overline{a_{j-1},a_j}$ gehöre der Zwischenwert c_j:

$$\begin{array}{ccccccccccc} f(a+h) & - & f(a_{n-1}) &&&&&&&& \\ & + & f(a_{n-1}) & - & f(a_{n-2}) &&&&&& \\ &&& + & f(a_{n-2}) & - & f(a_{n-3}) &&&& \\ && \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \\ &&&&&&& + & f(a_1) & - & f(a_0) \\ & c_n && c_{n-1} && c_{n-2} && \cdots && c_1 & \end{array}$$

Dann ist

$$\begin{array}{rcl} f(a+h)-f(a) &=& \sum_{j=1}^n D_j f(c_j)\,h_j\\ f(a+h)-f(a)-f'(a,h) &=& \sum_{j=1}^n D_j f(c_j)\,h_j - f'(a,h)\\ &=& \sum_{j=1}^n \left(D_j f(c_j)\,h_j - D_j f(a)\,h_j\right)\\ |f(a+h)-f(a)-f'(a,h)| &\le& \sum_{j=1}^n |D_j f(c_j)\,h_j - D_j f(a)\,h_j|\\ &\le& \|h\| \sum_{j=1}^n |D_j f(c_j) - D_j f(a)|\\ \frac{|f(a+h)-f(a)-f'(a,h)|}{\|h\|} &\le& \sum_{j=1}^n |D_j f(c_j) - D_j f(a)| \end{array}$$

Wegen der Stetigkeit von $D_1 f, ..., D_n f$ folgt die Behauptung. $\square$

Bemerkung

Die Sätze 3.8 bis 3.11 sind von *reellwertigen* Funktionen $f : U \to \mathbb{R}$ auf *vektorwertige* Abbildungen $f : U \to \mathbb{R}^m$ übertragbar, da sie für jede Einzelkomponente f_i von f gelten. Als lediglich *technische* Schwierigkeit erweist sich, daß bezüglich des Mittelwertsatzes die Zwischenstellen für die Einzelkomponenten im allgemeinen nicht dieselben sind. Ist der Schrankensatz anwendbar, läßt sich ein S z.B. als $S = \max\{S_1, \ldots, S_m\}$ oder als $S = S_1 + \ldots + S_m$ bilden, wenn man in $\mathbb{R}^m$ die Maximumsnorm oder die l_1-Norm verwendet. Man bekommt als Normabschätzung für den Zuwachsvektor:

$$\|f(x+h) - f(x)\| \le S\,\|h\|\,.$$

Folgerung (vgl. Satz 3.10)

Ist $U \subset \mathbb{R}^n$ der differenzierbaren $f : U \to \mathbb{R}^m$ wegzusammenhängend, dann gilt

$$\forall x \in U\,\forall h \in \mathbb{R}^n \quad f'(x,h) = 0 \Rightarrow f \text{ konstant.}$$

Satz 3.12

Sei $U \subset \mathbb{R}^n$, $f : U \to \mathbb{R}$ stetig differenzierbar, $K \subset U$ kompakt und konvex.

$$\forall \varepsilon > 0\,\exists \delta > 0\,\forall x, y \in K \text{ mit } \|y - x\| \le \delta \text{ gilt}$$
$$|f(y) - f(x) - \langle \mathrm{grad} f(x), y - x\rangle| < \varepsilon\,\|y - x\|\,.$$

Beweis

Nach dem Mittelwertsatz, der SCHWARZschen Ungleichung und weil grad$f$ auf der kompakten Menge K gleichmäßig stetig ist, gilt

$$\begin{aligned} |f(y) - f(x) - \langle \mathrm{grad} f(x), y - x\rangle| &= |\langle \mathrm{grad} f(z), y - x\rangle - \langle \mathrm{grad} f(x), y - x\rangle| \\ &\le \|\mathrm{grad} f(z) - \mathrm{grad} f(x)\|\,\|y - x\|\,, \end{aligned}$$

wobei z ein Punkt der Verbindungsstrecke $\overline{x,y}$ ist. □

3.4 Höhere Ableitungen. TAYLORscher Satz

Die Ableitung f' einer differenzierbaren Abbildung f kann in zweierlei Weisen aufgefaßt werden:

1. Ist f an allen Stellen $x \in U \subset \mathbb{R}^n$ differenzierbar, so erhält man die Ableitung f' von f als Abbildung von U

$$\begin{aligned} f' : U &\to \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m) \quad \text{definiert durch} \\ x &\mapsto f'(x). \end{aligned}$$

Dabei ist $\mathcal{L}$ die Menge der linearen Abbildungen $L : \mathbb{R}^n \to \mathbb{R}^m$. Eine direkte Einführung einer zweiten Ableitung f'' bietet sich unmittelbar an mit

$$\begin{aligned} f'' &:= (f')' \qquad \text{und daher} \\ f'' : U &\to \mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)) \end{aligned}$$

(Dieser Weg wird zunächst nicht weiter verfolgt).

2. Ist f an allen Stellen $x \in U \subset \mathbb{R}^n$ differenzierbar, so erhält man die Ableitung f' von f als Abbildung von $U \times \mathbb{R}^n$

$$\begin{aligned} f' : U \times \mathbb{R}^n &\to \mathbb{R}^m \quad \text{definiert durch} \\ (x,h) &\mapsto f'(x)h. \end{aligned}$$

Dabei ist f' linear in der Variablen h. Ergebnis ist ein «Ableitungsvektor» $f'(x)h$ des $\mathbb{R}^m$, der auch (wie bisher) mit $f'(x,h)$ bezeichnet wird. Eine Einführung einer zweiten Ableitung f'' kann geschehen mit

$$\begin{aligned} f''(x,h,k) &:= g_h'(x,k) \qquad \text{und daher} \\ f'' : U \times \mathbb{R}^n &\times \mathbb{R}^n \to \mathbb{R}^m \end{aligned}$$

Dazu wird für festes $h \in \mathbb{R}^n$ die Abbildung $g_h : U \to \mathbb{R}^m$ mit $g_h(x) = f'(x,h)$ erklärt. Ist dieses g_h für jedes $h \in \mathbb{R}^n$ differenzierbar, so heißt f zweimal differenzierbar. Diese Betrachtungsweise führt Ableitungen als multilineare Abbildungen ein, d.h. linear, bilinear usw.

Bemerkung

f'' ist für jedes $x \in U$ eine bilineare Abbildung, d.h. insbesondere

$$\begin{aligned} f''(x, h_1 + h_2, k) &= f''(x, h_1, k) + f''(x, h_2, k), \\ f''(x, ch, k) &= cf''(x, h, k). \end{aligned}$$

Bei Verwendung der Standardbasis von $\mathbb{R}^n$ mit h und k als Richtungsvektoren gilt

$$\begin{aligned} f'(x,h) &= \textstyle\sum_{i=1}^n D_i f(x) h_i, \\ f''(x,h,k) &= \textstyle\sum_{j=1}^n D_j \left(\sum_{i=1}^n D_i f(x) h_i\right) k_j \\ &= \textstyle\sum_{j=1}^n \sum_{i=1}^n D_j D_i f(x) h_i k_j. \end{aligned}$$

Bemerkung

Existiert f'', dann existieren auch die partiellen Ableitungen zweiter Ordnung $D_j D_i f(x)$. Schreibweisen sind

$$D_j D_i f = \frac{\partial \left(\frac{\partial f}{\partial x_i}\right)}{\partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i} = f_{x_i x_j} = f_{ij}$$

Betrachtet man h und k als Parameter, dann wird in Anwendung der Def. 3.1 auf die Ableitung von g_h die *definierende Bedingung für die zweite Ableitung* von f zu

$$f'(x+k, h) - f'(x,h) = f''(x,h,k) + R_x(h,k)\|k\|$$

mit $\lim_{k \to 0} R_x(h,k) = 0$ und R_x linear in h.

Satz 3.13 (*Satz von* SCHWARZ)

Sei $U \subset \mathbb{R}^n$, $f : U \to \mathbb{R}^m$ zweimal differenzierbar. Dann ist

$$\forall k, h \in \mathbb{R}^n \quad f''(x,h,k) = f''(x,k,h),$$

d.h. daß die bilineare Abbildung $(h,k) \mapsto f''(x,h,k)$ symmetrisch ist.

Beweis

Beweisidee: $f(x+h+k) - f(x+h) - f(x+k) + f(x)$ ist ein Näherungswert für $f''(x,h,k)$. Zu zeigen ist

$$\lim_{s\searrow 0} \frac{f(x+sh+sk) - f(x+sh) - f(x+sk) + f(x)}{s^2} = f''(x,h,k)\,.$$

Da die linke Seite bei Vertauschung von h und k sich nicht ändert, folgt die Behauptung der Symmetrie unmittelbar.

O.B.d.A. genügt es, den Beweis für reellwertiges f zu führen.

Schritt 1:

Es gilt mit $0 < \vartheta < 1$

$$\begin{array}{rl} (\star) & f(x+h+k) - f(x+h) - f(x+k) + f(x) \\ = & f''(x,h,k) + R_x(h, \vartheta h + k)\,\|\vartheta h + k\| - R_x(h, \vartheta h)\,\|\vartheta h\|\,. \end{array}$$

Mit $0 \le t \le 1$ für die Hilfsfunktion

$$F(t) := f(x+th+k) - f(x+th)$$

für die beiden ersten Glieder der linken Seite von $(\star)$ gelten

$$\begin{aligned} F'(t) &= f'(x+th+k,h) - f'(x+th,h)\,, \\ F(1) - F(0) &= f(x+h+k) - f(x+h) - f(x+k) + f(x)\,. \end{aligned}$$

Nach dem Mittelwertsatz für Funktionen einer reellen Variablen gilt

$$F(1) - F(0) = F'(\vartheta) \quad \text{mit} \quad 0 < \vartheta < 1.$$

Mit «Einschieben» von $f'(x,h)$ wird dann

$$\begin{array}{rl} & f(x+h+k) - f(x+h) - f(x+k) + f(x) \\ = & (f'(x+\vartheta h+k,h) - f'(x,h)) - (f'(x+\vartheta h,h) - f'(x,h))\,. \end{array}$$

Mit Hilfe der *definierenden* Bedingung für die zweite Ableitung (vgl. Bemerkung vor Satz) können die Differenzen der rechten Seite umgeformt werden zu

$$\begin{aligned} f'(x+\vartheta h+k,h) - f'(x,h) &= f''(x,h,\vartheta h+k) + R_x(h,\vartheta h+k)\,\|\vartheta h+k\|\,, \\ f'(x+\vartheta h,h) - f'(x,h) &= f''(x,h,\vartheta h) + R_x(h,\vartheta h)\,\|\vartheta h\|\,. \end{aligned}$$

Wegen $f''(x,h,\vartheta h+k) - f''(x,h,\vartheta h) = f''(x,h,k)$ folgt bei Ausführung der Differenz der rechten Seite die Gleichung $(\star)$.

Schritt 2:

Mit Substitution von h durch sh, von k durch sk mit $s > 0$ und da R_x nach Voraussetzung in h linear ist wird der Zähler des lim-Ausdrucks für reellwertiges f zu

$$\begin{array}{rl} & f(x+sh+sk) - f(x+sh) - f(x+sk) + f(x) \\ = & s^2\left(f''(x,h,k) + R_x(h, s(\vartheta h+k))\,\|\vartheta h+k\| - R_x(h, s\vartheta h)\,\|\vartheta h\|\right). \end{array}$$

Wegen $\lim_{s\to 0} R_x(h, s(\vartheta h+k)) = 0$ und $\lim_{s\to 0} R_x(h, s\vartheta h) = 0$ folgt die Behauptung. $\square$

Folgerung

Ist f zweimal differenzierbar, dann gilt $D_i D_j f(x) = D_j D_i f(x)$.

Bemerkung

Entsprechend Satz 3.11 bzw. seiner Verallgemeinerung auf $\mathbb{R}^m$ kann ausgesagt werden: Existieren alle zweiten partiellen Ableitungen von f und sind sie stetig, dann ist f zweimal differenzierbar.

Höhere oder *p-te Ableitungen* werden in Verfolg der zweiten Auffassung gesehen als

$$f^{(p)} : U \times \underbrace{\mathbb{R}^n \times \cdots \times \mathbb{R}^n}_{p-\text{mal}} \to \mathbb{R}^m$$

$$\left(x, h^{(1)}, \dots, h^{(p)}\right) \mapsto f^{(p)} \left(x, h^{(1)}, \dots, h^{(p)}\right)$$

Bei festem x besteht eine totalsymmetrische p-fach lineare Abbildung.

Bemerkung

Ist $f^{(p)}$ stetig, dann heißt f *p-mal stetig differenzierbar*. Gilt $\forall p \in \mathbb{N} \quad \exists f^{(p)}$, so heißt f *unendlich oft differenzierbar*.

Satz 3.14 (TAYLOR*sche Formel*)

Sei $U \subset \mathbb{R}^n$, $f : U \to \mathbb{R}$ $(p+1)$-mal differenzierbar, die Verbindungsstrecke $\overline{x, x+h} \subset U$.

Dann gilt

$$f(x+h) = f(x) + f'(x,h) + \frac{1}{2!} f''(x,h,h) +$$

$$+ \ldots + \frac{1}{p!} f^{(p)}(x, \underbrace{h, \ldots, h}_{p-\text{mal}}) + R_p(x,h)$$

wobei das Restglied R_p mit $0 < \vartheta < 1$ geschrieben werden kann als

$$R_p(x,h) = \frac{1}{(p+1)!} f^{(p+1)}(x + \vartheta h, \underbrace{h, \ldots, h}_{(p+1)-\text{mal}}).$$

Beweis

Beweisidee analog der des Mittelwertsatzes. Setze $\gamma(t) = x + th$ mit $0 \le t \le 1$.

Wende auf $F = f \circ \gamma$ den (eindimensionalen) TAYLORschen Satz an:

$$F(1) = F(0) + F'(0) + \frac{1}{2!} F''(0) + \ldots + \frac{1}{p!} F^{(p)}(0) + \frac{1}{(p+1)!} F^{(p+1)}(\vartheta).$$

Die Ableitungen von F sind

$$\begin{array}{lcl cl}
F'(t) & = & f'(\gamma(t), \gamma'(t)) & = & f'(\gamma(t), h), \\
F''(t) & = & f''(\gamma(t), h, \gamma'(t)) & = & f''(\gamma(t), h, h), \\
\vdots & & \vdots & & \vdots \\
F^{(p)}(t) & = & \dots & = & f^{(p)}(\gamma(t), \underbrace{h, \dots, h}_{p-\text{mal}}), \\
F^{(p+1)}(t) & = & \dots & = & f^{(p+1)}(\gamma(t), \underbrace{h, \dots, h}_{(p+1)-\text{mal}}).
\end{array}$$

Mit $\gamma(0) = x$ und $\gamma(1) = x + h$ folgt das Behauptete. □

Bemerkung

Unter der Voraussetzung, daß f $(p+1)$-mal stetig differenzierbar ist, erhält man eine *Integraldarstellung* des Restgliedes

$$\begin{aligned}
R_p(x, h) &= \frac{1}{p!} \int_0^1 (1-t)^p F^{(p+1)}(t) dt \\
&= \frac{1}{p!} \int_0^1 (1-t)^p f^{(p+1)}(x + th, \underbrace{h, \dots, h}_{(p+1)-\text{mal}}) dt.
\end{aligned}$$

Bemerkung

Entsprechend dem Mittelwertsatz ist auch für die TAYLOR-Formel die Übertragung auf vektorwertige Abbildungen möglich.

3.5 Zweite Ableitungen

Die zweite Ableitung n-dimensionaler reellwertiger Funktionen ist sehr wichtig wegen ihrer Rolle in physikalisch-mathematischen Modellen, wie Wellenausbreitung, Wärmediffusion, Strömungen u.a.

Die zweite Ableitung ist eine symmetrische Bilinearform. $f''(x, h, h)$ mit $h \in \mathbb{R}^n$ ist die zugehörige quadratische Form und heißt auch HESSE-Form von f an der Stelle x.

Definition 3.7 (HESSE-*Form bzgl. Standardbasis*)

Die HESSE-*Form* hat bezüglich der Standardbasis die Darstellung

$$f''(x, h, h) = \sum_{i=1}^{n} \sum_{j=1}^{n} f_{ij}(x) h_i h_j = h^\top H h$$

mit $f_{ij}(x) = D_i D_j f(x)$, $h = \sum h_i e_i$ und der HESSE-*Matrix* $H_{ij} = (f_{ij})$.

Erinnerung (Lineare Algebra)

Eine *quadratische Form* $Q : \mathbb{R}^n \to \mathbb{R}$ in x hat die Gestalt

$$Q(x) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j = x^\top A x$$

mit $x = (x_1, x_2, \ldots, x_n)^\top$ und der (reellen quadratischen) symmetrischen Matrix $A = (a_{ij})$. Sie läßt sich nach den Eigenwerten λ, die sie annehmen kann, klassifizieren (hier: Eigenwert(e) λ der Matrix A abgekürzt mit EWe)

i)	positiv definit	$\forall x \neq 0\ Q(x) > 0$	$\Leftrightarrow$	alle EWe > 0
ii)	negativ definit	$\forall x \neq 0\ Q(x) < 0$	$\Leftrightarrow$	alle EWe < 0
iii)	positiv semidefinit	$\forall x \neq 0\ Q(x) \geq 0$	$\Leftrightarrow$	alle EWe ≥ 0
iv)	negativ semidefinit	$\forall x \neq 0\ Q(x) \leq 0$	$\Leftrightarrow$	alle EWe ≤ 0
v)	indefinit	$\exists x_1\ Q(x_1) > 0$	$\Leftrightarrow$	$\exists$EWe > 0
		und $\exists x_2\ Q(x_2) < 0$		und $\exists$EWe < 0

Im Falle $n = 2$ erhält man als einfaches Kriterium mit $A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$

i)	positiv definit	$\det A > 0 \wedge a > 0$
ii)	negativ definit	$\det A > 0 \wedge a < 0$
iii)	positiv semidefinit	$\det A \geq 0 \wedge a \geq 0$
iv)	negativ semidefinit	$\det A \geq 0 \wedge a \leq 0$
v)	indefinit	$\det A < 0$

Quadriken

Die Funktion $f : U \to \mathbb{R}$ sei in einer Umgebung von $a \in U$ zweimal differenzierbar, $f''(a, h, k)$ sei *nicht* die Nullmatrix.

Definition 3.8 (*Quadrik*)

Die quadratische Gleichung

$$x_{n+1} = f(a) + f'(a)(x - a) + \frac{1}{2}(x - a)^\top f''(a)(x - a)$$

beschreibt eine *Quadrik* (Fläche 2.Ordnung) im $\mathbb{R}^{n+1}$.

Bemerkung

Man spricht in diesem Falle auch von der Schmiegequadrik an den Graphen von f in $(a, f(a))$. Im Falle $n = 2$ kann man durch geeignete affine Koordinatentransformation jede beliebige Schmiegequadrik auf eine der folgenden Normalformen bringen

(E)	$z = \pm(x^2 + y^2)$	elliptisches Paraboloid
(H)	$z = x^2 - y^2$	hyperbolisches Paraboloid
(P)	$z = \pm x^2$	parabolischer Zylinder

Bemerkung

Zweimal differenzierbare Abbildungen f können klassifiziert werden:

Der Graph von f heißt in einem Punkt $(a, f(a))$

elliptisch	$\Leftrightarrow$	$f''(a)$ ist positiv oder negativ definit
hyperbolisch	$\Leftrightarrow$	$f''(a)$ ist nicht singulär und indefinit
parabolisch	$\Leftrightarrow$	$f''(a)$ singulär und ungleich der Nullmatrix
flach	$\Leftrightarrow$	$f''(a)$ ist gleich der Nullmatrix

Ein *hyperbolischer* Punkt heißt *Sattelpunkt*.

Beispiel

Die Funktion $f(x,y) = 3x^2y - y^3$ hat die HESSE-Matrix

$$f''(x,y) = \begin{pmatrix} D_{xx} & D_{xy} \\ D_{yx} & D_{yy} \end{pmatrix} f(x,y) = 6 \begin{pmatrix} y & x \\ x & -y \end{pmatrix}.$$

Sie ist an der Stelle $(x,y) = (0,0)$, d.h. im Nullpunkt, wegen

$$f''(0,0) = 6 \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

flach, ansonsten wegen

$$\det f''(x,y) = 36 \begin{vmatrix} y & x \\ x & -y \end{vmatrix} = -36(y^2 + x^2) < 0 \quad \forall (x,y) \neq (0,0)$$

indefinit und damit hyperbolisch.

Lokale Minima und Maxima

Zur Charakterisierung lokaler Extrema bei n-dimensionalen Funktionen bedarf es der Entwicklung entsprechender Bedingungen, denen $f : U \to \mathbb{R}$ genügen muß.

Definition 3.9 (*kritische Stelle*)

Sei $U \subset \mathbb{R}^n$ offen, $f : U \to \mathbb{R}$ differenzierbar, $a \in U$.

a heißt *kritische Stelle* von f, wenn $\forall h \in \mathbb{R}^n \quad f'(a,h) = 0$.

Bemerkung

$\forall h \in \mathbb{R}^n \quad f'(a,h) = 0 \Leftrightarrow D_1 f(a) = \ldots = D_n f(a) = 0 \Leftrightarrow \mathrm{grad} f(a) = 0.$

Satz 3.15

Sei $U \subset \mathbb{R}^n$ offen, $f : U \to \mathbb{R}$ differenzierbar, $a \in U$.

a ist lokales Extremum einer differenzierbaren Funktion $f \Rightarrow a$ ist kritische Stelle von f.

Beweis (für Maximum, Minimum analog)

Voraussetzung mit V Umgebung: $\exists V\, \forall x \in V \quad f(x) \leq f(a)$.

Wähle $h \neq 0$ so, daß $a+h, a-h \in V$.

Betrachte f auf der Verbindungsstrecke $\overline{a-h, a+h}$, d.h. die Funktion $F(t) := f(a+th)$ mit $-1 \leq t \leq 1$.

Da $F(t) \leq F(0)$ und F differenzierbar, folgt $F'(0) = 0 = f'(a,h)$.

Wegen der Linearität von f' bezüglich h gilt dies nicht nur für ein $h = 0 \in \mathbb{R}^n$, sondern für alle $h \in \mathbb{R}^n$. $\square$

Die Vorgehensweise zur Bestimmung von Extremalstellen einer differenzierbaren rellwertigen Funktion f von n Variablen besteht daher in zwei Schritten

1. Ermittle zunächst alle kritischen Stellen von f
2. Prüfe die kritischen Stellen einzeln auf ihre Extremaleigenschaft.

Satz 3.16

Sei $U \subset \mathbb{R}^n$, $f : U \to \mathbb{R}$ zweimal stetig differenzierbar, $a \in U$ kritische Stelle.

Ist die quadratische Form $f''(a,h,h)$ *positiv definit*, so ist a *Minimum* von f, ist sie *negativ definit*, so ist a *Maximum* von f, und ist sie *indefinit*, so ist a *keine* Extremalstelle von f , sondern Sattelpunkt.

Beweis

$f(x) = f(a) + f'(a, x-a) + \frac{1}{2} f''(a + \vartheta(x-a), x-a, x-a)$ mit $0 < \vartheta < 1$ (Taylor-Formel). $f'(a, x-a) = 0$ nach Voraussetzung. Das weitere Verhalten von f hängt daher von f'', genauer von dem (eigentlichen) Restglied ab. Wegen der Stetigkeit der zweiten Ableitung ist

$$f''(a + \vartheta(x-a), x-a, x-a) = f''(a, x-a, x-a) + r(x)\,\|x-a\|^2$$

mit $\lim_{x \to a} r(x) = 0$.

Fall 1: $f''(a,h,h)$ *positiv definit.*

$\exists c > 0\ f(a,h,h) \geq c\,\|h\|^2$, denn das Minimum c der quadratischen Form ist auf der Einheitskugel $\|h\| = 1$ positiv nach Voraussetzung. Bei Wahl von $\|x-a\|$ so klein, daß $|r(x)| < \frac{c}{2}$, wird

$$f(x) - f(a) = \frac{1}{2}\left(f''(a, x-a, x-a) + r(x)\,\|x-a\|^2\right)$$

und daraus die Abschätzung

$$f(x) - f(a) \geq \frac{c}{4}\,\|x-a\|^2\,,$$

d.h. a ist lokales Minimum für f.

Fall 2: $f''(a,h,h)$ *negativ definit.*

Beweisführung analog.

Fall 3: $f''(a,h,h)$ *indefinit.*

$\exists h_0 \quad \|h_0\| = 1, f''(a,h_0,h_0) > 0$ und $\exists k_0 \quad \|k_0\| = 1, f''(a,k_0,k_0) < 0$.

Betrachte die beiden Geraden durch a mit den Richtungen h_0 und k_0: für ein genügend kleines $\|x-a\|$ wird $f(x) - f(a)$ positiv bzw. negativ. D.h. a ist kein lokales Extremum. $\square$

Bemerkung

Wenn die quadratische Form $f''(a,h,h)$ *positiv semidefinit* bzw. *negativ semidefinit* ist, so ist dies für ein lokales Extremum *notwendig, aber nicht hinreichend.*

Beispiel

$f(x,y) = x^2 + y^3$. Setze $\operatorname{grad} f(x,y) = (2x, 3y^2) = (0,0)$. Dann ist $(0,0)$ der einzige kritische Punkt.

$$f''(0,0) = \begin{pmatrix} 2 & 0 \\ 0 & 6y \end{pmatrix}_{(0,0)} = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$$

ist positiv semidefinit, aber $(0,0)$ ist kein Extremum.

ÜBERSICHT

Ist (x_0, y_0) kritische Stelle von $f : U \subset \mathbb{R}^2 \to \mathbb{R}$ und notiert man

$$\begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix} = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$$

- Zitat: SCHWARZ bringt Symmetrie! - dann ist

$a > 0$	$ac - b^2 > 0$	hinreichend für Minimum
$a < 0$	$ac - b^2 > 0$	hinreichend für Maximum
$a \geq 0$	$ac - b^2 \geq 0$	notwendig für Minimum
$a \leq 0$	$ac - b^2 \geq 0$	notwendig für Maximum
	$ac - b^2 < 0$	hinreichend für kein Extremum

Konvexe Funktionen

Definition 3.10 (*konvexe Funktion*)

Sei $U \subset \mathbb{R}^n$ offen und konvex. $f : U \to \mathbb{R}$ heißt *konvex*, wenn

$$\forall x_1, x_2 \in U \; x_1 \neq x_2 \wedge \forall t \in [0,1]$$

$$f\left((1-t)\,x_1 + tx_2\right) \leq (1-t)\,f(x_1) + tf(x_2)\,.$$

Für $t \in \,]0,1[$ und mit «$<$» heißt f *streng konvex.*

Konkav und *streng konkav* analog $\Leftrightarrow -f$ konvex bzw. streng konvex.

Bemerkung

vgl. Def. 1.20 (*konvexe Menge*)

Satz 3.17

Sei $U \subset \mathbb{R}^n$ konvex, $f : U \to \mathbb{R}$ differenzierbar.

f ist *streng* konvex genau dann, wenn

$$\forall a, x \in U\ x \neq a \quad f(x) > f(a) + f'(a, x-a).$$

Satz 3.18 (*Konvexitätskriterium*)

Sei $U \subset \mathbb{R}^n$ konvex, $f : U \to \mathbb{R}$ *zweimal* differenzierbar.

Ist f'' überall in U positiv definit, dann ist f streng konvex.

Beweis

Mit der Voraussetzung, daß f in U zweimal differenzierbar ist, gilt mit der TAYLOR-Formel

$$f(x) = f(a) + f'(a, x-a) + \frac{1}{2} f''(a + \vartheta(x-a), x-a, x-a)$$

mit $0 < \vartheta < 1$. Nach Satz 3.17 gilt für strenge Konvexität

$$f(x) - f(a) - f'(a, x-a) > 0.$$

Dies bedeutet, daß für $x \neq a$

$$f''(a + \vartheta(x-a), x-a, x-a) > 0.$$

Hieraus folgt als hinreichende Bedingung für die strenge Konvexität:

$$\forall y \in U\ \forall h \in \mathbb{R}^n \quad f''(y, h, h) > 0.$$

□

Bemerkung

Hinreichende Bedingung für Konvexität zweimal differenzierbarer Funktionen ist, daß alle HESSE-Formen positiv semidefinit sind. Aber

$$\begin{array}{lcl} \text{streng konvex} & \Rightarrow & H_f \text{ positiv semidefinit} \\ \text{streng konvex} & \not\Rightarrow & H_f \text{ positiv definit} \end{array}$$

3.6 Lineare Differentialoperatoren

Sofern nicht anders angegeben, erfolgen die nachstehenden Notationen auf der Grundlage eines $\mathbb{R}$-Vektorraumes mit den entsprechenden Einheitsvektoren, die die Standardbasis bilden.

Definition 3.11 (*linearer Differentialoperator*)

Werden in dem multivariaten Polynom

$$P(x) = \sum a_{k_1 \cdots k_n} x_1^{k_1} \cdots x_n^{k_n}$$

die x_i durch die (partiellen) Ableitungsoperatoren D_i ersetzt, so entsteht ein sog. *linearer Differentialoperator* $P(D)$ mit konstanten Koeffizienten.

Hat P den Grad k mit $k = \max\{k_1 + \cdots + k_n\}$, so ist $P(D)$ die lineare Abbildung

$$P(D) : \mathcal{C}^k(U) \to \mathcal{C}(U)$$

und

$$P(D) f := \sum a_{k_1 \ldots k_n} D_1^{k_1} \cdots D_n^{k_n} f,$$

wobei D_i^p bedeutet, D_i p-mal anzuwenden.

Laplace-Operator

Setzt man die Koeffizienten $a_{k_1 \ldots k_n}$ gleich Null mit Ausnahme von

$$a_{2,0,\ldots,0} = a_{0,2,\ldots,0} = \cdots = a_{0,0,\ldots,2} = 1,$$

dann erhält man für $P(D)$ den Laplace-*Operator*:

$$\Delta := D_1^2 + D_2^2 + \cdots + D_n^2,$$

d.h. mit der Euklidischen Metrik auf $\mathbb{R}^n$ wird über die zugehörige quadratische Form

$$Q(x) = x_1^2 + x_2^2 + \cdots + x_n^2$$

dieser Differentialoperator induziert.

Alternativ wird der Laplace-Operator Δ über die Spur (=Summe der Diagonaleinträge) der Hesse-Matrix einer zweimal differenzierbaren Funktion

$$\mathrm{spur} H_f(x) = \sum_{i=1}^{n} D_i^2 f = \Delta f(x)$$

hergeleitet.

Die Bedeutung dieses Operators - insbesondere für die theoretische Physik - ergibt sich aus seiner zentralen Eigenschaft der *Drehinvarianz*.

Satz 3.19

Für jede Orthonormalbasis (ONB) $v_1, \ldots, v_n$ des euklidischen $\mathbb{R}^n$ gilt

$$\Delta f = D_{v_1}^2 f + \ldots + D_{v_n}^2 f.$$

Beweis

Allgemein ist

$$D_{v_i} D_{v_j} f(x) = v_i^\top H_f(x) v_j.$$

Mit der Orthogonalmatrix $V := (v_n)$ wird

$$D_{v_i} D_{v_j} f(x) = e_i^\top \tilde{H}_f(x) e_i$$

mit $\tilde{H}_f(x) = V^\top H_f V$ (Transformation auf Standardbasis).

Da V orthogonal, wird

$$\sum_{i=1}^{n} D_{v_i}^2 f = \mathrm{spur}\tilde{H},$$

d.h. H und $\tilde{H}$ haben gleiche Spur. □

Bemerkung

Sei $D \subset \mathbb{R}^n$ offen, $f : D \to \mathbb{R}$. Der LAPLACE-Operator Δ mit

$$\Delta f := \sum_{i=1}^{n} \frac{\partial^2 f}{\partial x_i^2} = \mathrm{spur} H_f$$

ist invariant unter orthogonalen Koordinatentransformationen, einer wichtigen Eigenschaft zur Modellierung physikalischer Phänomene in isotropen Medien. Unter allen *linearen Differentialoperatoren zweiter Ordnung mit konstanten Koeffizienten* sind $c\Delta + d$ mit $c, d \in \mathbb{R}$ die einzigen Differentialoperatoren mit dieser Eigenschaft. (Zweiter Ordnung heißt: das zugehörige Polynom hat den Grad 2.)

Beispiel 1 Potentialgleichung

- homogene Potentialgleichung $\Delta u = 0$, wichtig z.B. in RIEMANN-Theorie der komplex differenzierbaren Funktionen. So kann man zeigen, daß jede komplex differenzierbare Funktion (reell) $\Delta f = 0$ erfüllt.
- inhomogene Potentialgleichung $\Delta u = f$ mit f als fest gegebener Funktion, wichtig z.B. in Gravitationstheorie, allgemeiner «Potentialtheorie».

Definition 3.12 (*harmonische Funktion*)

Die Lösungen u der homogenen Potentialgleichung $\Delta u = 0$ heißen *harmonische Funktionen.*

Beispiel 2 Wellen- und Schwingungsgleichung

$$\Delta_x \psi = \frac{1}{c^2} \psi_{tt}$$

mit $c \neq 0$ für Funktionen $\psi(x,t)$ mit «Ort» $x \in \mathbb{R}^n$, «Zeit» $t \in \mathbb{R}$ und

$$\Delta_x = \sum_{i=1}^{n} \frac{\partial^2}{\partial x_i^2}$$

beschreibt Dynamik elektrischer oder magnetischer Felder, ...

Beispiel 3 Wärmeleitungsgleichung

$$\Delta_x \psi = \frac{1}{k} \psi_t$$

mit $k > 0$ beschreibt Dynamik bei der Wärmeausbreitung.

3.7 Maximumprinzipien

Satz 3.20

Sei $D \subset \mathbb{R}^n$ offen und beschränkt, $f : \bar{D} \to \mathbb{R}$ zweimal differenzierbar in D und stetig auf $\bar{D}$. f genüge der Differentialgleichung $\Delta f \geq 0$ in D.

Dann nimmt f ihr Maximum auf dem Rand ∂D an.

Beweis

Definiere $M := \max\left\{f(x), x \in \bar{D}\right\}, m := \max\left\{f(x), x \in \partial D\right\}$. Diese Maxima existieren, da $\bar{D}$ und ∂D kompakt. Immer gilt $m \leq M$. Zu zeigen: $m = M$.

Annahme: $m < M$.

Definiere $g(x) := \sum_{i=1}^n x_i^2$. Dann «Laplace-$g$» $\Delta g = 2n > 0$.

Setze $\varepsilon := \frac{M-m}{2\max\{g(x), x \in \partial D\}} > 0$.

Definiere $h := f + \varepsilon g$. Dann gelten $h \geq f$ und $\Delta h = \Delta f + \varepsilon \Delta g > 0$ sowie

$$\max\{h(x), x \in \partial D\} \leq m + \varepsilon \max\{g(x), x \in \partial D\} = m + \frac{M-m}{2} < M.$$

Daraus folgt: h nimmt Maximum in $x_0 \in D$ an. Notwendig für Maximum ist $H_h(x_0)$ negativ semidefinit.

Wegen $\Delta h(x_0) > 0 \Rightarrow \exists k \in \{1, \ldots, n\}$ mit $\frac{\partial^2 h}{\partial x_k^2}(x_n) > 0$.

Aber $\frac{\partial^2 h}{\partial x_k^2}(x_0) = e_k^\top H_h(x_0) e_k \leq 0$. Widerspruch, d.h. $m = M$. □

Korollar 3.21

Sei $D \subset \mathbb{R}^n$ offen und beschränkt, $f : \bar{D} \to \mathbb{R}$ zweimal differenzierbar in D und stetig auf $\bar{D}$. f sei in D harmonisch, d.h. $\Delta f = 0$.

Dann nimmt f ihr Minimum und ihr Maximum auf dem Rand ∂D an. («Strikt»-werden der Differentialgleichung.)

Beweis

$\Delta f \geq 0, \Delta(-f) \geq 0 \Rightarrow_{\text{wegen } 3.20} f, (-f)$ nehmen Maximum auf ∂D an. □

Anwendung auf Dirichlet-Problem für inhomogene Potentialgleichungen:

Finde $u : \bar{D} \to \mathbb{R}$, D offen, beschränkt, u in D zweimal differenzierbar, auf $\bar{D}$ stetig mit $\Delta u = f$ in D, $u|_{\partial D} = g$.

(*Anm.*: f ist vorgegeben, $u|_{\partial D}$ bedeutet «eingeschränkt auf Rand».)

Satz 3.22

Falls eine Lösung u des DIRICHLETschen Problems existiert, dann ist sie eindeutig.

Beweis

Seien u_1, u_2 Lösungen. Bilde $u := u_1 - u_2$. Dann gilt in D:

$$\Delta u = \Delta u_1 - \Delta u_2 = f - f = 0,$$

ebenso

$$u|_{\partial D} = u_1|_{\partial D} - u_2|_{\partial D} = g - g = 0.$$

Mit Korollar 3.21 (u anstelle von f) nimmt u Minimum und Maximum auf Rand an, d.h. $u|_{\partial D} = 0 \Rightarrow u = 0 \Rightarrow u_1 = u_2$. $\square$

Bemerkung

Existenzaussagen sind wesentlich schwieriger.

Literatur

Protter; Weinberger: *Maximum Principles in Differential Equations.*

3.8 Nablakalkül

Rechnungen und Schreibweisen gestalten sich oft leichter mit Hilfe des Nablakalküls unter Benutzung des Nablaoperators ∇ (auch $\frac{\partial}{\partial \boldsymbol{r}}$) als symbolischen Vektor.

Sei $a, b : \mathbb{R}^n \to \mathbb{R}$, $\boldsymbol{f}, \boldsymbol{g} : \mathbb{R}^n \to \mathbb{R}^m$,

$\boldsymbol{v}, \boldsymbol{w} : \mathbb{R}^n \to \mathbb{R}^n$ («Vektorfeld»).

Gradient (Symbol $\mathrm{grad} a = \nabla a$, Verknüpfung ∇a)

Vektorgradient (Symbol $\mathrm{grad}\boldsymbol{f}$, Verknüpfung $\nabla \boldsymbol{f}$)

$$\nabla \boldsymbol{f} := \left(\frac{\partial f^i}{\partial x_j}\right)_{ij} = \begin{pmatrix} \frac{\partial f^1}{\partial x_1} & \cdots & \frac{\partial f^1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f^m}{\partial x_1} & \cdots & \frac{\partial f^m}{\partial x_n} \end{pmatrix} = \begin{pmatrix} \mathrm{grad} f^1 \\ \vdots \\ \mathrm{grad} f^m \end{pmatrix},$$

was der JACOBI-Matrix von $\boldsymbol{f}$ entspricht.

Divergenz (Symbol $\mathrm{div}\boldsymbol{v}$, Verknüpfung $\nabla \cdot \boldsymbol{v} = \langle \nabla, \boldsymbol{v} \rangle$)

$$\nabla \cdot \boldsymbol{v} := \sum_{i=1}^{n} \frac{\partial v^i}{\partial x_i} = \begin{pmatrix} \frac{\partial}{\partial x_1} \\ \vdots \\ \frac{\partial}{\partial x_n} \end{pmatrix} \cdot \begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix}.$$

Rotation (nur für $n=3$) (Symbol $\mathrm{rot}\boldsymbol{v} = \mathrm{curl}\boldsymbol{v}$, Verknüpfung $\nabla \times \boldsymbol{v}$)

$$\nabla \times \boldsymbol{v} := \begin{pmatrix} \frac{\partial v^3}{\partial x_2} & -\frac{\partial v^2}{\partial x_3} \\ \frac{\partial v^1}{\partial x_3} & -\frac{\partial v^3}{\partial x_1} \\ \frac{\partial v^2}{\partial x_1} & -\frac{\partial v^1}{\partial x_2} \end{pmatrix} =$$

$$= \begin{pmatrix} \frac{\partial}{\partial x_1} \\ \frac{\partial}{\partial x_2} \\ \frac{\partial}{\partial x_3} \end{pmatrix} \times \begin{pmatrix} v^1 \\ v^2 \\ v^3 \end{pmatrix} = \det \begin{vmatrix} e_1 & \frac{\partial}{\partial x_1} & v^1 \\ e_2 & \frac{\partial}{\partial x_2} & v^2 \\ e_3 & \frac{\partial}{\partial x_3} & v^3 \end{vmatrix}.$$

Es ist zumeist praktisch, Vektoren als Spaltenvektoren, Gradienten als Zeilenvektoren zu notieren. Dann läßt sich z.B. die TAYLOR-Entwicklung schreiben

$$a(x) = a(x_0) + \underbrace{\nabla a(x_0)}_{\text{Zeilenvektor}} \underbrace{(x - x_0)}_{\text{Spaltenvektor}} + \frac{1}{2}(x - x_0)^\top \boldsymbol{H}_a(x)(x - a) + Rest,$$

manchmal mit $\boldsymbol{H}_a = \nabla^2 a = \nabla\nabla a$, eigentlich $\nabla(\nabla_a)^\top$.

Es gilt für die Beziehung zwischen Nablaoperator und LAPLACE-Operator

$$\nabla \cdot \nabla a = \nabla \cdot \left(\frac{\partial a}{\partial x_i}\right)_i = \sum_{i=1}^{n} \frac{\partial}{\partial x_i} \frac{\partial a}{\partial x_i} = \sum_{i=1}^{n} \frac{\partial^2 a}{\partial x_i^2} = \Delta a$$

Weitere *Regeln*

- $\nabla\nabla \cdot a$ ist nicht definiert
- $\nabla(a\boldsymbol{f}) = a\nabla\boldsymbol{f} + \boldsymbol{f}\nabla a$
- $\nabla \cdot (a\boldsymbol{v}) = a\nabla \cdot \boldsymbol{v} + (\nabla a)\boldsymbol{v}$
- $\nabla \cdot (\boldsymbol{v} \times \boldsymbol{w}) = (\nabla \times \boldsymbol{v})^\top \boldsymbol{w} - (\nabla \times \boldsymbol{w})^\top \boldsymbol{v}$
- $\nabla \cdot \nabla \times \boldsymbol{v} = 0$
- $\nabla \times (\nabla \times \boldsymbol{v}) = \mathrm{rot}\,\mathrm{rot}\boldsymbol{v} = \mathrm{grad}\,\mathrm{div}\boldsymbol{v} - \Delta\boldsymbol{v}$

Rechenregeln für den Nablaoperator (Vektordifferentiation)

Der Nablakalkül übersetzt Vektoranalysis in Vektoralgebra und umgekehrt.

Nutzung der Schreibweisen und Regeln aus der *Vektoralgebra* mit den Vektoren $\boldsymbol{a}$, $\boldsymbol{b}$ und $\boldsymbol{c}$ sowie $\alpha \in \mathbb{R}$

- $\boldsymbol{ab}$ für das *Skalarprodukt* mit

$$\begin{aligned} \boldsymbol{ab} &= \boldsymbol{ba}, \\ \alpha(\boldsymbol{ab}) &= (\alpha\boldsymbol{a})\boldsymbol{b}, \\ \boldsymbol{a}(\boldsymbol{b} + \boldsymbol{c}) &= \boldsymbol{ab} + \boldsymbol{ac}, \\ \boldsymbol{a}^2 &:= \boldsymbol{aa} = |\boldsymbol{a}|^2 \end{aligned}$$

- $\boldsymbol{a} \times \boldsymbol{b}$ für das *Vektorprodukt* mit

$$\begin{aligned} \boldsymbol{a} \times \boldsymbol{b} &= -(\boldsymbol{b} \times \boldsymbol{a}), \\ \alpha(\boldsymbol{a} \times \boldsymbol{b}) &= (\alpha \boldsymbol{a}) \times \boldsymbol{b}, \\ \boldsymbol{a} \times (\boldsymbol{b} + \boldsymbol{c}) &= (\boldsymbol{a} \times \boldsymbol{b}) + (\boldsymbol{a} \times \boldsymbol{c}), \\ \boldsymbol{a} \times \boldsymbol{a} &= 0 \end{aligned}$$

- $(\boldsymbol{abc})$ für das *Spatprodukt* mit

$$(\boldsymbol{abc}) := (\boldsymbol{a} \times \boldsymbol{b})\boldsymbol{c}$$

- *Entwicklungssatz*

$$\boldsymbol{a} \times (\boldsymbol{b} \times \boldsymbol{c}) = \boldsymbol{b}(\boldsymbol{ac}) - \boldsymbol{c}(\boldsymbol{ab})$$

- *Identität von* Lagrange

$$(\boldsymbol{a} \times \boldsymbol{b})(\boldsymbol{c} \times \boldsymbol{d}) = (\boldsymbol{ac})(\boldsymbol{bd}) - (\boldsymbol{bc})(\boldsymbol{ad})$$

und Nutzung der Basis $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ ergibt als Schreibweisen für Nablaoperator, Gradient, Divergenz und Rotation

$$\nabla := \mathbf{e}_1 \frac{\partial}{\partial x_1} + \mathbf{e}_2 \frac{\partial}{\partial x_2} + \mathbf{e}_3 \frac{\partial}{\partial x_3}$$

$$\mathbf{grad} f = \nabla f, \quad \mathrm{div} \boldsymbol{v} = \nabla \boldsymbol{v}, \quad \mathbf{rot} \boldsymbol{v} = \nabla \times \boldsymbol{v}$$

«Textersetzungsregeln des Nablakalküls»

1. Schreibe den gewünschten Ausdruck als «formales» Produkt mittels ∇.

2. Multipliziere eine Linearkombinationen von Funktionen oder Vektoren X, Y mit $\alpha, \beta \in \mathbb{R}$ «formal» aus

   $$\nabla(\alpha X + \beta Y) = \alpha \nabla X + \beta \nabla Y.$$

3. Schreibe ein Produkt $\nabla(XY)$ mit der «Produktregel» als

   $$\nabla(XY) = \nabla(\bar{X}Y) + \nabla(X\bar{Y}),$$

   wobei der Querstrich jeweils den *zu differenzierenden* Faktor bezeichnet.

4. Forme $\nabla(\bar{X}Y)$ und $\nabla(X\bar{Y})$ *streng* nach den Regeln der Vektoralgebra um und zwar so, daß alle Größen *ohne* Querstrich *links* von ∇ und alle Größen *mit* Querstrich *unmittelbar rechts* von ∇ stehen. Behandle dabei ∇ als *Vektor*. Stehen als Ergebnis alle zu differenzierenden Faktoren rechts von ∇, dann kann dieser wieder zum *Differentialoperator* werden:

5. Schreibe die so gewonnenen «formalen» Produkte mit ∇ wieder als Ausdrücke der Vektoranalysis und lasse den Querstrich, weil jetzt entbehrliche Hilfskennzeichnung, weg, z.B.

   $$\begin{aligned} U(\nabla \bar{V}) &= U \, \mathbf{grad} \, V, \\ V \times (\nabla \times \bar{W}) &= V \times \mathbf{rot} \, W. \end{aligned}$$

Beispiel 1

$$\begin{aligned}\mathbf{grad}\,(UV) &= \nabla(UV)\\ &= \nabla(\bar{U}V) + \nabla(U\bar{V})\\ &= V(\nabla\bar{U}) + U(\nabla\bar{V})\\ &= V\,\mathbf{grad}\,U + U\,\mathbf{grad}\,V\end{aligned}$$

Beispiel 2

$$\begin{aligned}\mathrm{div}\,(\boldsymbol{v}\times\boldsymbol{w}) &= \nabla(\boldsymbol{v}\times\boldsymbol{w})\\ &= \nabla(\bar{\boldsymbol{v}}\times\boldsymbol{w}) + \nabla(\boldsymbol{v}\times\bar{\boldsymbol{w}})\\ &= \boldsymbol{w}(\nabla\times\bar{\boldsymbol{v}}) - \boldsymbol{v}(\nabla\times\bar{\boldsymbol{w}})\\ &= \boldsymbol{w}\,\mathbf{rot}\,\boldsymbol{v} - \boldsymbol{v}\,\mathbf{rot}\,\boldsymbol{w}\end{aligned}$$

Beispiel 3

$$\begin{aligned}\mathbf{grad}\,(\boldsymbol{vw}) &= \nabla(\boldsymbol{vw})\\ &= \nabla(\bar{\boldsymbol{v}}\boldsymbol{w}) + \nabla(\boldsymbol{v}\bar{\boldsymbol{w}})\\ &= (\boldsymbol{w}\nabla)\bar{\boldsymbol{v}} + \boldsymbol{w}\times(\nabla\times\bar{\boldsymbol{v}})\\ &\quad +(\boldsymbol{v}\nabla)\bar{\boldsymbol{w}} + \boldsymbol{v}\times(\nabla\times\bar{\boldsymbol{w}})\\ &= (\boldsymbol{w}\,\mathbf{grad})\,\boldsymbol{v} + \boldsymbol{w}\times\mathbf{rot}\,\boldsymbol{v}\\ &\quad +(\boldsymbol{v}\,\mathbf{grad})\,\boldsymbol{w} + \boldsymbol{v}\times\mathbf{rot}\,\boldsymbol{w}\end{aligned}$$

3.9 Fixpunktsatz von Banach

Definition (*Kontraktion*)

Ein Operator (eine Abbildung) $K : B \to B$ heißt *kontrahierend* oder eine *Kontraktion*, wenn

$$\exists c \in [0,1[\ \forall x,y \in B \quad d\left(K\left(x\right),K\left(y\right)\right) \leq cd\left(x,y\right).$$

Schreibweise als Operator Kx, als Abbildung $K(x)$.

Erinnerung

Banach-Raum heißt ein vollständiger, normierter Raum, d.h. ein normierter Raum, in dem jede Fundamentalfolge (Cauchy-Folge) konvergiert.

Satz 3.23 (*Fixpunktsatz von* Banach)

Sei B ein Banach-Raum mit Metrik $d(\ ,\ )$, $K : B \to B$ eine Kontraktion.

Dann existiert *genau ein* Fixpunkt z von K, d.h. $K\left(z\right) = z$.

Beweis

Schreibe $K^n x := \underbrace{K \circ K \circ \cdots \circ K}_{\text{n−mal}} x$. Dann gilt mit $c < 1$ die Abschätzung (n-malige Kontraktion)

$$
\begin{aligned}
d(K^n x, K^{n+k} x) &\leq cd(K^{n-1} x, K^{n+k-1} x) \\
&\leq \ldots \\
&\leq c^n d(x, K^k x) \\
&\leq c^n \left(d(x, Kx) + d(Kx, K^k x)\right) \\
&\leq \ldots \\
&\leq c^n (d(x, Kx) + d(Kx, K^2 x) + \cdots + d(K^{k-1} x, K^k x)) \\
&\leq c^n \left(1 + c + c^2 + \ldots + c^{k-1}\right) d(x, Kx) \\
&\leq c^n \left(\sum_{i=0}^{\infty} c^i d(x, Kx)\right) \\
&= \frac{c^n}{1-c} d(x, Kx).
\end{aligned}
$$

Daraus folgt, daß $(K^n x)_{n \in \mathbb{N}}$ für jeden beliebigen Startwert $x \in B$ eine CAUCHY-Folge ist. Da B als BANACH-Raum vollständig ist, konvergiert sie gegen ein $z \in B$. Wegen der Stetigkeit von K gilt insgesamt

$$K(z) = K(\lim K^n x) = \lim K^{n+1}(x) = z,$$

d.h. z ist Fixpunkt.

Die Eindeutigkeit wird geprüft durch Annahme $\exists x, x'$ mit $Kx = x$ und $Kx' = x'$. Dann gilt

$$d(x, x') = d(Kx, Kx') \leq cd(x, x') \Rightarrow x = x'$$

wegen $c < 1$. □

Bemerkung Vorsicht

- $d(Kx, Ky) < d(x, y)$ ist nicht hinreichend für Existenz, nur für Eindeutigkeit, z.B.

$$x \mapsto x + \frac{1}{x}.$$

- $d(Kx, Ky) \leq d(x, y)$ ist nicht hinreichend für Eindeutigkeit, z.B.

$$x \mapsto x + 1.$$

Bemerkung

Der Fixpunktsatz gilt allgemein in vollständigen metrischen Räumen. Die Vektorraum-Struktur war zur Beweisführung nicht benötigt worden.

Anwendungsbereich

- als Methode für Beweis von Existenz und Eindeutigkeit in der Analysis: « *vor* Lösung deren Existenz und Eindeutigkeit nachweisen»

- als numerisches Verfahren, denn aus

$$d\left(K^n x, K^{n+k} x\right) \leq \frac{c^n}{1-c} d(x, Kx)$$

folgt mit

$$\lim_k d\left(K^n x, K^{n+k} x\right) = d\left(K^n x, z\right)$$

die apriori-Abschätzung für Fehler nach n Iterationen

$$d\left(K^n x, z\right) \leq \frac{c^n}{1-c} d(x, Kx).$$

Es handelt sich also um eine «saubere» Konvergenz, die die Fehlerabschätzung mitliefert. Das Verfahren kann mit beliebigem x starten, doch ist es in der Praxis sinnvoll, ein x möglichst nahe dem Fixpunkt zu wählen. Die Folge $(K^n x)$ konvergiert mindestens so schnell gegen den Fixpunkt z wie die geometrische Folge $\gamma c^n \to 0$ mit $\gamma = \frac{d(x,Kx)}{1-c}$, $c < 1$.

Nullstellenbestimmung

Eine wichtige Anwendung des Fixpunktsatzes von Banach besteht in der Bestimmung von *Nullstellen* einer Abbildung:

Finde für $f : U \to \mathbb{R}^n, U \subset \mathbb{R}^n$ diejenigen $x \in U$, für die $f(x) = 0$ gilt.

Hierzu ist es notwendig, die Suche nach *Nullstellen*

- Nullstellengleichung der Form $f(x) = 0$

auf äquivalente Weise als Suche nach *Fixpunkten*

- Fixpunktgleichung der Form $x = \varphi(x)$

zu formulieren.

Bestimme die Fixpunkte der Abbildung

$$K : U \to \mathbb{R}^n$$

erklärt durch

$$K(x) = x - A(x) f(x).$$

Hierbei ist $A(x)$ eine jedem $x \in U$ zugeordnete umkehrbare lineare Abbildung, die freigewählt werden kann: «frei» bis auf eine Kontraktionszahl c, damit $K(x)$ eine Kontraktion ist.

In der Praxis wählt man A unabhängig von x. Ist f differenzierbar, so auch K und man erhält als Ableitung

$$K'(x) = \mathrm{id}_{\mathbb{R}^n} - A \circ f'(x).$$

Mittels der Norm $\|f\| = \sup_{\|x\|=1} \|f(x)\|$ läßt sich die Kontraktionsforderung für K ausdrücken als

$$\|\mathrm{id}_{\mathbb{R}^n} - A \circ f'(x)\| \leq c < 1.$$

Wird U als konvex vorausgesetzt, ergibt sich der *Beweis* hierfür direkt aus dem *Schrankensatz* (Satz 3.9):

$$\|K(x) - K(y)\| \leq \sup_{z \in U} \|K'(z)\| \, \|x - y\| \, .$$

Finde nun die kompakte Teilmenge M von U mit $K : M \to M$.

Ist ein a schon beinahe Nullstelle, d.h. sein $\|f(a)\|$ klein, dann ist er auch beinahe ein Fixpunkt von K, da aus $K(x) - x = -A(x)f(x)$ die Abschätzung

$$\|K(a) - a\| \leq \|A\| \, \|f(a)\|$$

folgt.

Eine kleine Kugel um a sollte durch K in sich abgebildet werden, da K kontrahierend ist.

Dies ist der Fall, wenn als Radius der Kugel

$$r = \frac{\|A\| \, \|f(a)\|}{1 - c}$$

gewählt wird, denn dann gilt für $\|x - a\| \leq r$ die Abschätzung

$$\begin{array}{rcl} \|K(x) - a\| & \leq & \|K(x) - K(a)\| + \|K(a) - a\| \\ & \leq & c\,\|x - a\| + \|A\| \, \|f(a)\| \\ & \leq & cr + (1 - c)r = r \end{array} \quad .$$

Somit ist der Fixpunktsatz von BANACH anwendbar: in der Kugel $\|x - a\| \leq r$ hat K genau einen *Fixpunkt* und daher f genau eine *Nullstelle*.

Diese Überlegungen zur Nullstellenbestimmung werden zusammengefaßt in

Satz 3.24

Sei $U \subset \mathbb{R}^n$, $f : U \to \mathbb{R}^n$ differenzierbar, A eine umkehrbare lineare Abbildung. Es gelte für $\|x - a\| \leq r$ die Kontraktionseigenschaft (Abschätzung)

$$\|\mathrm{id}_{\mathbb{R}^n} - A \circ f'(x)\| \leq c < 1.$$

Gilt nun

$$\|f(a)\| \leq \frac{r(1 - c)}{\|A\|}$$

so hat f *genau eine Nullstelle* z in der Kugel $\|x - a\| \leq r$. Diese Nullstelle z ist Grenzwert jeder Folge (x_k), die bei beliebig gewähltem Startwert x_0 mit $\|x_0 - a\| \leq r$ *rekursiv* durch $x_{k+1} = x_k - Af(x_k)$ gegeben ist. Es gilt die Fehlerabschätzung

$$\|z - x_k\| \leq \frac{c^k}{1 - c} \|A\| \, \|f(x_0\| \, .$$

Bemerkung

Spielraum für die Wahl der umkehrbaren linearen Abbildung A

- $A = (f'(a))^{-1}$ im modifiziertem NEWTON-Verfahren mit linearer Konvergenz, wobei A unabhängig von x ist
- $A = (f'(x))^{-1}$ im allgemeinen NEWTON-Verfahren mit (lokal) quadratischer Konvergenz.

Vorgehensweise

1. Wähle a nahe bei der (vermuteten) Nullstelle.
2. Wähle A, z.B. $A = (f'(a))^{-1}$, doch mit einem a so, daß $\det f'(a) \neq 0$, da sonst $\|\mathrm{id}_{\mathbb{R}^n} - A \circ f'(x)\| = 0$.
3. Finde für stetiges f' eine Kugelumgebung $\|x - a\| \leq r$ so, daß
$$\|\mathrm{id}_{\mathbb{R}^n} - A \circ f'(x)\| \leq c < 1$$
mit für eine schnelle Konvergenz hinreichend kleinem c, z.B. $c \leq \frac{1}{2}$.
4. Prüfe, ob
$$\|f(a)\| \leq \frac{r(1-c)}{\|A\|}$$
erfüllt ist. Wenn nicht, wähle besseres a.

Beispiel (nach Barner;Flohr: *Analysis II.* S.151 ff.)

$f = \begin{pmatrix} f \\ g \end{pmatrix}$ mit $f(x,y) = x^2 - y^2 - 1 = 0$ und $g(x,y) = 2xy - 1 = 0$

(Anmerkung: Die expliziten exakten Lösungen sind $x_0 = \sqrt{\frac{\sqrt{2}+1}{2}}, y_0 = \sqrt{\frac{\sqrt{2}-1}{2}}$.)

1. Wähle als Ausgangspunkt $a = (1, \frac{1}{2})$. Es ist dann $\|f(a)\|_2 = \frac{1}{4}$ und die JACOBI-Matrix an der Stelle $(1, \frac{1}{2})$ ist $\begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix}$.
2. Als A bietet sich die Inverse zur JACOBI-Matrix an: $\begin{pmatrix} \frac{2}{5} & \frac{1}{5} \\ -\frac{1}{5} & \frac{2}{5} \end{pmatrix}$ mit $\det A = \frac{1}{5}$.
3. Die Prüfung der Bedingung $\|\mathrm{id}_{\mathbb{R}^n} - A \circ f'(x)\| \leq c < 1$ führt auf
$$\left\| \begin{pmatrix} (1 - \frac{4}{5}x - \frac{2}{5}y) & (-\frac{2}{5}x + \frac{4}{5}y) \\ (\frac{2}{5}x - \frac{4}{5}y) & (1 - \frac{4}{5}x - \frac{2}{5}y) \end{pmatrix} \right\|_2 \leq c$$
Der rechnerische Aufwand für die EUKLIDische Norm wird verringert bei Verwendung einer Abschätzung
$$\left\| \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \right\|_2 \leq \sqrt{\alpha^2 + \beta^2 + \gamma^2 + \delta^2},$$
die auf die Bedingung $(x-1)^2 + (y - \frac{1}{2})^2 \leq \frac{5}{8}c^2$ führt, aus der $r = c\sqrt{\frac{5}{8}}$ als Radius um $a = (1, \frac{1}{2})$ folgt.

4. Aus der noch zu erfüllenden Bedingung $\|f(a)\|\,\|A\| \le r(1-c)$ wird mit $\|f(a)\|_2 = \frac{1}{4}$ und $\|A\|_2 = \sqrt{\frac{1}{5}}$ für die Kontraktionskonstante $\frac{\sqrt{2}}{10} \le c(1-c)$. Dem wird z.B. mit $c = \frac{1}{5}$ genügt, woraus sich ein Kreis um $(1, \frac{1}{2})$ mit $r = c\sqrt{\frac{5}{8}} = \frac{1}{\sqrt{40}}$ ergibt.

Dieser Kreis wird mit der Kontraktionskonstante $\frac{1}{5}$ durch die kontrahierende Abbildung $K = \begin{pmatrix} K_1 \\ K_2 \end{pmatrix}$ mit

$$\begin{aligned} K_1(x,y) &= x - \tfrac{2}{5}x^2 - \tfrac{2}{5}xy + \tfrac{2}{5}y^2 + \tfrac{3}{5} \\ K_2(x,y) &= y + \tfrac{1}{5}x^2 - \tfrac{4}{5}xy - \tfrac{1}{5}y^2 + \tfrac{1}{5} \end{aligned}$$

in sich abgebildet.

Die a-priori-Abschätzung für Fehler in der n-ten Iteration ergibt mit dem Term $\frac{c^k}{1-c}\,\|A\|\,\|f(x_0\|$ nach Satz 3.24 $\frac{c^n}{1-c}\,\|A\|_2\,\|f(a\|_2 = \frac{\sqrt{5}}{16}\left(\frac{1}{5}\right)^n$.

Normen für Abbildungen

Die anläßlich der Definitionen der punktweisen und der gleichmäßigen Konvergenz von *Abbildungsfolgen* (vgl. Def. 2.5) sowie der *Normkonvergenz* (vgl. Satz 2.11) gebrauchte *Abbildungsnorm* bezieht sich auf die Menge der *beschränkten* Abbildungen

$$\mathcal{B}(D, \mathbb{R}^m) := \{f \mid f : D \to \mathbb{R}^m \wedge f \text{ beschränkt}\},$$

die mit der auf ihr erklärten Abbildungsnorm

$$\|f\|_\infty := \sup\{\|f(x)\| \mid x \in D\} \quad \text{auch geschrieben als} \quad \|f\| := \sup_{x\in D}\|f(x)\|$$

ein *normierter* $\mathbb{R}$-Vektorraum ist. Diese Norm ist insbesondere für alle *stetigen* Abbildungen mit *kompakter* Definitionsmenge D erklärt, in welchem Falle auch

$$\|f\| := \max_{x\in D}\|f(x)\|$$

gesetzt werden kann.

Der in der Definition erkennbare Ansatz lautet: Die Vektornorm $\|f(x)\|$ induziert die Abbildungsnorm (Operatornorm) $\|f\|$.

Rechenregeln für die Norm in $\mathcal{B}(D, \mathbb{R}^m)$

(1) $\|f\| = 0 \Rightarrow f(x) = 0$ für alle $x \in D$

(2) $\|cf\| = |c|\,\|f\|$

(3) $\|f+g\| = \|f\| + \|g\|$

Problem

Da die lineare Abbildung $f : \mathbb{R}^n \to \mathbb{R}^m$, sofern sie nicht die Nullabbildung ist, nicht beschränkt ist, ist für sie eine Norm, wie vorstehend definiert, nicht erklärt.

Trotzdem ist es zweckmäßig, auch hierfür eine Norm einzuführen. Wegen der *Homogenität* von f, d.h.

$$f(\lambda x_1, \lambda x_2, \dots, \lambda x_n) = \lambda f(x_1, x_2, \dots, x_n)$$

für beliebige λ, ist f durch die Werte auf der Einheitskugel $\|x\| = 1$ schon festgelegt. Die *Einschränkung* auf diese kompakte Menge ist stetig, sodaß hierfür eine Norm im obigen Sinne existiert:

$$\|f\| := \sup_{\|x\|=1} \|f(x)\| = \sup_{x\neq 0} \frac{\|f(x)\|}{\|x\|}$$

mit $f_{| \|x\|=1}$. Mit $c := \lambda$ kann man allen Abbildungen, für die mit $x \neq 0, c \neq 0$ die Funktionalgleichung

$$\|f(cx)\| = |c|^s \, \|f(x)\|$$

gilt, eine Norm zuordnen mit

$$\|f\| := \sup_{\|x\|=1} \|f(x)\| = \sup_{x\neq 0} \frac{\|f(x)\|}{\|x\|^s},$$

und für alle $x \neq 0$ gilt die (zentrale) Abschätzung

$$\|f(x)\| \leq \underbrace{\|f\|}_{\text{Operatornorm}} \; \underbrace{\|x\|^s}_{\text{Vektornorm}} .$$

Spezialfall

Homogene Funktionen, die aus *total-symmetrischen p-fach linearen* Funktionen $g(x_1, x_2, \dots, x_p)$ dadurch entstehen, daß $x_1 = x_2 = \cdots x_p = x$. Dann heißt

$$f(x) := g(\underbrace{x, \dots, x}_{p-\text{mal}})$$

eine *Form p-ter Ordnung* und es gilt für sie mit $c \in \mathbb{R}$ die Funktionalgleichung

$$f(cx) = c^p f(x),$$

z.B. kubische ($p = 3$) und quadratische ($p = 2$) Formen (vgl. z.B. Def. 3.8 Quadriken).

Satz 3.25

Sei $f, g \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$, d.h. lineare Abbildungen des $\mathbb{R}^n$ in den $\mathbb{R}^m$. Dann gilt

$$\begin{array}{lrcll}
(1) & \|f\| & > & 0 & \text{für } f \neq 0 \\
(2) & \|cf\| & = & |c| \, \|f\| & \text{mit } c \in \mathbb{R} \\
(3) & \|f+g\| & \leq & \|f\| + \|g\| & \\
(4) & \|f(x)\| & \leq & \|f\| \, \|x\| & \text{mit } x \in \mathbb{R}^n \\
(5) & \|f \circ g\| & \leq & \|f\| \, \|g\| & \text{für } \mathbb{R}^n = \mathbb{R}^m
\end{array}$$

Bemerkung

- in (4) ist $\|f\|$ Operatornorm, $\|x\|$ Vektornorm.

- Der Vektorraum $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ aller linearen Abbildungen von $\mathbb{R}^n$ in $\mathbb{R}^m$ ist nach (1), (2) und (3) ein normierter Vektorraum.
- Der Vektorraum $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ aller linearen Abbildungen von $\mathbb{R}^n$ in $\mathbb{R}^m$ ist nach (1), (2), (3) und (5) eine *normierte Algebra*.

Beweis (zu Satz 3.25)

der Teilsätze (1), (2) und (4) direkt mit der Definition der Norm,

für (3) wird die für beliebiges $x \in \mathbb{R}^n$ geltende Abschätzung

$$\|(f+g)(x)\| = \|f(x) + g(x)\| \le \|f(x)\| + \|g(x)\| \le (\|f\| + \|g\|)\,\|x\|\,.$$

$\square$

Beispiel

Sei $f : \mathbb{R}^n \to \mathbb{R}^n$ eine lineare Abbildung, A die zugehörige Matrix.

Dann gilt für die EUKLIDische Abbildungsnorm $\|f\|_2$: $\|x\|_2 = \sqrt{\sum_i x_i^2}$.

Man kommt zu einer dieser entsprechenden *Matrixnorm* auf folgende Weise:

$A = (a_{ij}) \in \mathbb{R}^{n \times n}$ sei symmetrisch.

Dann gibt es n Eigenvektoren v_i mit Eigenwerten $\lambda_i \in \mathbb{R}$, die orthogonal sind, d.h. es gilt $Av_i = \lambda_i v_i$ mit $v_i^\top v_j = 0$ für $i \ne j$.

In dem durch die Eigenvektoren aufgespannten $\mathbb{R}^n$ ist x darstellbar als $x = \sum c_i v_i$ mit $c_i \in \mathbb{R}$.

Dann ist die Operatornorm (hier: *Spektralnorm*)

$$\begin{array}{rcll}
\|A\|_2 & = & \sup_{\|x\|_2=1} \|Ax\|_2 & \\
 & = & \sup \|A \sum c_i v_i\|_2 & \\
 & = & \sup \|\sum c_i \lambda_i v_i\|_2 & \\
 & = & \sup \sqrt{\left(\sum c_i \lambda_i v_i\right)^\top \left(\sum c_j \lambda_j v_j\right)} & \\
 & = & \sup \sqrt{\sum c_i^2 \lambda_i^2 v_i^\top v_i} & \text{Orthogonalitt} \\
 & \le & \sup \sqrt{\sum c_i^2 \lambda_{\max}^2 v_i^\top v_i} & \\
 & = & \sup |\lambda_{\max}| \sqrt{\sum c_i^2 v_i^\top v_i} & \text{mit } \sqrt{\sum c_i^2 v_i^\top v_i} = \|x\|_2 = 1 \\
 & = & |\lambda_{\max}| & \text{wobei } |\lambda_{\max}| = \max |\lambda_i|
\end{array}$$

Also ist $\|A\|_2 \le |\lambda_{\max}|$. Zur Abschätzung nach unten wird der spezielle, zu $\lambda_{\max}$ gehörende Eigenvektor $x = v_{\max}$ eingesetzt

$$\|A\|_2 \ge \frac{\|Ax\|_2}{\|x\|_2} = \frac{\|\lambda_{\max} x\|_2}{\|x\|_2} = |\lambda_{\max}|\,,\ \text{d.h.}$$

$$\|A\|_2 = |\lambda_{\max}|\,.$$

Entsprechend werden gewonnen (und gelten auch mit $A \in \mathbb{R}^n \times \mathbb{R}^m$):

- die *Spaltensummennorm* $\|A\|_1 = \sup_{x \ne 0} \frac{\|Ax\|_1}{\|x\|_1} = \max_k \sum_i |a_{ik}|$

- die *Zeilensummennorm* $\|A\|_\infty = \sup_{x\neq 0} \frac{\|Ax\|_\infty}{\|x\|_\infty} = \max_i \sum_k |a_{ik}|$

Praxis

Im Vektorraum $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ aller linearen Abbildungen von $\mathbb{R}^n$ in $\mathbb{R}^m$ sind alle Normen äquivalent, man verwendet deshalb möglichst einfache Normen obiger Art oder auch

$$\begin{array}{ll} \|A\|_{\text{abs}} := \sum_{i=1}^n \sum_{j=1}^n |a_{ij}| & \\ \|A\|_{\text{F}} := \sqrt{\sum_{i=1}^n \sum_{j=1}^n a_{ij}^2} & \text{nach FROBENIUS} \end{array}$$

Weitere Beziehungen sind

$$\max_{ik} |a_{ik}| \leq \|A\| \leq \sqrt{nm} \max_{ik} |a_{ik}|$$

sowie

$$\begin{array}{rcl} \|AB\|_{\text{abs}} & \leq & \|A\|_{\text{abs}} \|B\|_{\text{abs}} \\ \|AB\|_{\text{F}} & \leq & \|A\|_{\text{F}} \|B\|_{\text{F}} \\ \|Ax\|_\infty & \leq & \|A\|_{\text{abs}} \|x\|_\infty \\ \|Ax\|_2 & \leq & \|A\|_{\text{F}} \|x\|_2 \end{array}$$

Definition 3.13 (*Homogenität*)

Sei V, W $\mathbb{R}$-Vektorräume. $f : V \to W$ heißt

- *homogen vom Grad* 1, falls

$$\forall v \in V \; \forall c \in \mathbb{R} \quad f(cv) = cf(v)$$

- *positiv homogen vom Grad* $s \in \mathbb{R}$, falls

$$\forall v \in V \; \forall c > 0 \quad f(cv) = c^s f(v)$$

Fazit

Sei V, W normierte Räume und $f : V \to W$ positiv homogen vom Grad s.

Definiere Operatornorm

$$\|f\|_{V\to W} := \sup_{v\in V} \frac{\|f(v)\|_W}{\|v\|_V^s} = \sup_{\|v\|_V = 1} \|f(v)\|_W$$

Besonders wichtig: $l : V \to W$ linear, also positiv homogen vom Grad 1 mit Norm

$$\|l\|_{V\to W} := \sup_{v\in V} \frac{\|l(v)\|_W}{\|v\|_V^1} = \sup_{\|v\|_V = 1} \|l(v)\|_W$$

und der wichtigen Abschätzung

$$\begin{array}{rcll} \|l(v)\|_W & \leq & \|l\|_{V\to W} \|v\|_V & \\ \|l(v_1) - l(v_2)\|_W & \leq & \|l\|_{V\to W} \|v_1 - v_2\|_V & \text{Linearität} \end{array}$$

Verallgemeinerung auf nichtlineare Abbildungen

Definition 3.14 (LIPSCHITZ-*stetig*)

Sei V, W *metrische* Räume, $V' \subset V$.

$f : V' \to W$ heißt LIPSCHITZ-*stetig,* falls

$$\exists L \in \mathbb{R}\, \forall v_1, v_2 \in V' \quad \|f(v_1) - f(v_2)\|_W \leq L \|v_1 - v_2\|_V \text{ bzw.}$$

$$\exists L \in \mathbb{R}\, \forall v_1, v_2 \in V' \quad d_W(f(v_1), f(v_2)) \leq L d_V(v_1, v_2).$$

Bemerkung

L heißt LIPSCHITZ-*Konstante*. In geometrischer Interpretation ist L größter Dehnungskoeffizient.

Beispiel 1

Gegeben sei $f : \mathbb{R} \to \mathbb{R}$ mit $f(x) = 2x + 3$.

Dann ist $|f(x) - f(y)| = |(2x+3) - (2y+3)| \leq L|x - y|$.

Somit ist f LIPSCHITZ-stetig mit $L = 2$.

Beispiel 2

Gegeben sei $f : V \to W$ mit f Kontraktion, d.h. $d_W(f(v_1), f(v_2)) \leq c d_V(v_1, v_2)$ mit $c < 1$.

Der kontrahierende Operator f ist LIPSCHITZ-stetig mit $L = c$.

Beispiel 3

Gegeben sei $l : V \to W$ mit l linear.

Dann gilt: l ist LIPSCHITZ-stetig mit $L = \|l\|_{V \to W}$ genau dann, wenn die Operatornorm $\|l\|_{V \to W}$ endlich ist.

Beispiel 4

Gegeben sei $\|\cdot\| : V \to \mathbb{R}$ mit $\|\cdot\|$ Normfunktion.

Dann ist $\|v_1\|_V = \|v_1 - v_2 + v_2\| \leq \|v_1 - v_2\| + \|v_2\|$.

Hieraus folgen $\|v_1\| - \|v_2\| \leq \|v_1 - v_2\|$ (und umgekehrt)

sowie $|\|v_1\| - \|v_2\|| \leq \|v_1 - v_2\|$ (und umgekehrt).

D.h. Normfunktionen (Normen) sind LIPSCHITZ-stetig (L-stetig) mit $L = 1$.

Beispiel 5

Gegeben sei $f : \mathbb{R} \to \mathbb{R}$ mit $f(x) = x^2$.

Dann ist $|f(x) - f(y)| = \left|x^2 - y^2\right| = |x + y|\,|x - y|$.

Nach Definition müßte gelten $|x + y| \leq L$.

Dies ist wegen $x, y \in \mathbb{R}$ für beliebige x, y nicht erfüllbar:

f ist nicht LIPSCHITZ-stetig.

Aber:

Definition 3.15 (*lokal* LIPSCHITZ-*stetig*)

Sei V, W *metrische* Räume, $V' \subset V$.

$f : V' \to W$ heißt *lokal* LIPSCHITZ-*stetig*, falls $\forall v \in V'$ eine Umgebung V'' von v mit $V'' \subset V'$ existiert, so daß $f|_{V''}$ L-stetig. Dabei hängt die LIPSCHITZ-Konstante i.allg. von der Umgebung V'' ab.

Beispiel 5 (nochmals)

Zu $v \in V' \subset \mathbb{R}$ wähle $V'' =]-|v|-1, |v|+1[$.

Dann ist $\forall x, y \in V''$

$$\begin{aligned} |f(x) - f(y)| &= \left|x^2 - y^2\right| \\ &= |x+y|\,|x-y| \\ &\leq (|x| + |y|)\,|x-y| \\ &\leq 2\,(|v|+1)\,|x-y|\,. \end{aligned}$$

Daraus ergibt sich die *lokale* LIPSCHITZ-Stetigkeit mit $L = 2\,(|v|+1)$.

Bemerkung

Globale L-Stetigkeit impliziert lokale L-Stetigkeit, aber nicht umgekehrt.

Beispiel 6

Gegeben sei $f : \mathbb{R} \to \mathbb{R}$ mit $f(x) = \sqrt{|x|}$.

Dann ist $|f(x) - f(0)| = \left|\sqrt{|x|}\right| = \frac{1}{\left|\sqrt{|x|}\right|}\,|x - 0|$.

Nach Definition müßte gelten $\frac{1}{\left|\sqrt{|x|}\right|} \leq L$, was nicht erfüllbar ist, d.h., f ist nicht *lokal* L-*stetig*.

Bemerkung

Stetigkeit allein impliziert nicht lokale L-Stetigkeit.

Satz 3.26

Sei V, W *metrische* Räume, $V' \subset V$.

Ist $f : V' \to W$ lokal LIPSCHITZ-stetig, dann ist f stetig.

Beweis

Sei $v \in V'$.

Wähle V'' so, daß $f|_{V''}$ L-stetig ist mit der LIPSCHITZ-Konstanten L.

Zu einem gegebenen $\varepsilon > 0$ wähle $\delta = \frac{\varepsilon}{L+1} > 0$.

Dann ist $\forall \tilde{v} \in V''$ mit $d(v, \tilde{v}) < \delta$

$$d_W(f(v), v(\tilde{v})) \leq L d_V(v, \tilde{v}) \leq L\delta < (L+1)\delta = \varepsilon$$

$\square$

Satz 3.27

Sei V, W *normierte* Räume.

Die lineare Abbildung $l : V \to W$ ist stetig $\Leftrightarrow$ l ist LIPSCHITZ-stetig.

Beweis

«$\Leftarrow$» L-Stetigkeit $\Rightarrow$ lokale L-Stetigkeit $\Rightarrow$ Stetigkeit.

«$\Rightarrow$» l ist stetig in 0.

Zu $\varepsilon = 1$ $\exists \delta > 0$, sodaß $\forall v \in V$ mit $\|v\| < \delta$ gilt: $\|l(v)\| < \varepsilon = 1$.

Setze $L := \frac{2}{\delta}$. Für $\forall v_1, v_2 \in V$ ist dann

$$\begin{aligned}
\|l(v_1) - l(v_2)\| &= \|l(v_1 - v_2)\| \\
&= \frac{2\,\|v_1 - v_2\|}{\delta} \left\| l\left(\frac{\delta}{2\,\|v_1 - v_2\|} (v_1 - v_2) \right) \right\| \\
&\leq \frac{2\,\|v_1 - v_2\|}{\delta} \cdot 1 \\
&= L\,\|v_1 - v_2\|\,.
\end{aligned}$$

$\square$

Beispiel (Gegenbeispiel)

Sei $V = \mathcal{C}^1([0,1]) = \{f \mid f : [0,1] \to \mathbb{R}\,\text{stetig differenzierbar}\}$, $\|f\| := \max |f|$, $W = \mathbb{R}$, $l : V \to \mathbb{R}$, $l(f) = f'(0)$ linear.

Betrachte Funktionenfolge (f_n) mit $f_n(x) = \frac{1}{n}\sin(nx)$, $f'(0) = 1$.

Es gilt $\|f_n\| \to 0$, d.h. $f_n \to 0$.

Andererseits ist $l(f_n) = 1$ aber $l(0) = 0$.

Es ergibt sich Unstetigkeit nach dem Folgenkriterium.

Satz 3.28

Seien V, W *normierte* Räume, $l : V \to W$ linear.

Ist $\dim V < \infty$, d.h. V endlichdimensional, dann ist l stetig.

Beweis

Setze $n := \dim V < \infty$. Wähle Basis $b_1, \ldots, b_n \in V$.

$$\forall v \in V\ \exists c_i \in \mathbb{R} \quad v = \sum_{i=1}^{n} c_i b_i.$$

Definiere zweite Norm auf V: $\|v\|' := \sum_{i=1}^{n} |c_i|$.

Wegen Normäquivalenz in endlichdimensionalen Räumen (vgl. Satz 1.9) gilt

$$\exists k \in \mathbb{R} \quad \|v\|' \leq k\,\|v\|\,.$$

Setze $M := \max \|l(b_i)\|_W$. Dann gilt

$$\begin{aligned}\|l(v)\| &\leq \sum_{i=1}^{n} |c_i| \, \|l(b_i)\|_W \\ &\leq M \sum_{i=1}^{n} |c_i| \\ &= M \|v\|' \\ &\leq Mk \|v\|,\end{aligned}$$

d.h. l ist stetig. □

Satz 3.29

Sei $D \subset \mathbb{R}^n$ konvex, $f : D \to \mathbb{R}^m$ differenzierbar, $\|\cdot\|_a$ eine Norm auf $\mathbb{R}^n$, $\|\cdot\|_b$ eine Norm auf $\mathbb{R}^m$.

Dann gilt

$$\exists k \in \mathbb{R} \, ((\forall \xi \in D \;\; \|f'(\xi)\|_{a \to b} \leq k) \Rightarrow (\forall x, y \in D \;\; \|f(x) - f(y)\|_b \leq k \|x - y\|_a)) \, .$$

Bemerkung vgl. Satz 3.9 (*Schrankensatz*)

Beweis

Nach dem Mittelwertsatz (Satz 3.8)

$$\exists t \in \,]0,1[\quad f(x) - f(y) = f'(x + t(y-x), (y-x)),$$

d.h.

$$\begin{aligned}\|f(y) - f(x)\|_b &= \|f'(x + t(y-x), (y-x))\|_b \\ &\leq \|f'(x + t(y-x))\|_{a \to b} \|y - x\| \\ &\leq k \|y - x\| \, .\end{aligned}$$

□

Korollar 3.30

Sei $D \subset \mathbb{R}^n$ offen, $f : D \to \mathbb{R}^m$ stetig differenzierbar.

Dann ist f lokal Lipschitz-stetig.

3.10 Umkehrsatz. Implizite Funktionen

Gegeben seien die Matrix $A \in \mathbb{R}^{n \times n}$, die lineare Abbildung $l : \mathbb{R}^n \to \mathbb{R}^n$ mit $l(x) = Ax$. Umkehrbarkeit von l ist äquivalent der Regularität von A, d.h. $\det A \neq 0$.

Bemerkung

$l' = A.$

Ist $f : \mathbb{R}^n \to \mathbb{R}^n$ in $a \in \mathbb{R}^n$ differenzierbar, dann läßt sich f lokal um a «gut» durch lineare Funktionen approximieren. Dies verbindet sich mit der Hoffnung, daß $\det f'(a) \neq 0$ hinreichend für die lokale Umkehrbarkeit von f ist. Globale Umkehrbarkeit ist nicht zu erwarten.

Satz 3.31 (*Umkehrsatz*)

Sei $D \subset \mathbb{R}^n$ eine offene Menge, $f : D \to \mathbb{R}^n$ eine in $a \in D$ mit $\det f'(a) \neq 0$ stetig differenzierbare Abbildung.

Dann existieren eine offene Menge $D' \subset D$ mit $a \in D'$ und eine offene Menge $E \subset \mathbb{R}^n$ mit $f(a) \in E$, sodaß $f|_{D'} : D' \to E$ bijektiv und die Umkehrabbildung $f^{-1} : E \to D'$ stetig differenzierbar ist.

Bemerkung vgl. Satz 3.7.

Danach ist $(f^{-1})'(f(x)) = (f')^{-1}(x)$ mit $x \in D$.

Beweis

O.B.d.A. sei $a = 0, f(a) = 0, f'(a) = \mathrm{id}_{\mathbb{R}^n}$. Ansonsten Übergang von $f(x)$ zu $f'^{-1}(a)(f(x+a) - f(a))$.

Definiere zu $y \in \mathbb{R}^n$ die Abbildung $K_y : D \to \mathbb{R}^n$ mit $x \mapsto x - f(x) + y$.

Es gilt $K_y'(0) = \mathrm{id}_{\mathbb{R}^n} - f'(0) = 0$.

f stetig differenzierbar $\Rightarrow K_y'$ stetig $\Rightarrow$ Operatornorm hängt stetig von x ab $\Rightarrow$ $\exists \varepsilon > 0 \quad \|x\| \leq 2\varepsilon \Rightarrow \|K_0'(x)\| \leq \frac{1}{2}$.

Außerdem sei ε so klein, daß $\bar{B}_{2\varepsilon}(0) \subset D$. Für $x \in \bar{B}_{2\varepsilon}(0)$ und $y \in \bar{B}_{\varepsilon}(0)$ gilt

$$\begin{aligned}
\|K_y(x)\| &= \|x - f(x) + y\| \\
&\leq \|x - f(x)\| + \|y\| \\
&= \|K_0(x) - K_0(0)\| + \|y\| \\
&\leq \frac{1}{2}\|x - 0\| + \|y\| \\
&\leq \frac{1}{2} 2\varepsilon + \varepsilon \\
&= 2\varepsilon,
\end{aligned}$$

womit die Abbildung $K_y : \bar{B}_{2\varepsilon}(0) \to \bar{B}_{2\varepsilon}(0)$ zu einer Abbildung in sich wird.

Wegen $K_y' = K_0'$ gilt $\left\|K_y'\right\| = \|K_0'\| \leq \frac{1}{2} \Rightarrow K_y$ ist eine Kontraktion mit Kontraktionszahl $\frac{1}{2}$.

$\bar{B}_{2\varepsilon}(0)$ ist vollständiger metrischer Raum.

Mit Fixpunktsatz von Banach (vgl. Satz 3.23) folgt $\forall y \in \bar{B}_{\varepsilon}(0)\ \exists_1 x \in \bar{B}_{2\varepsilon}(0)$

$$\begin{aligned}
x = K_y(x) = x - f(x) + y &\Rightarrow f(x) = y \\
&\Rightarrow \exists f^{-1} : \bar{B}_{\varepsilon}(0) \to \bar{B}_{2\varepsilon}(0).
\end{aligned}$$

Zu zeigen: *Stetigkeit* von f^{-1}

Sei $y, \tilde{y} \in \bar{B}_\varepsilon(0)$, $x := f^{-1}(y)$, $\tilde{x} := f^{-1}(\tilde{y})$.

Dann gilt

$$\begin{aligned}
\|x - \tilde{x}\| &= \|K_0(x) + f(x) - K_0(\tilde{x}) - f(\tilde{x})\| \\
&\leq \|K_0(x) - K_0(\tilde{x})\| + \|f(x) - f(\tilde{x})\| \\
&\leq \frac{1}{2}\|x - \tilde{x}\| + \|f(x) - f(\tilde{x})\|
\end{aligned}$$

Somit gilt

$$\begin{aligned}
\|x - \tilde{x}\| \leq 2\|f(x) - f(\tilde{x})\| \quad &\Rightarrow \quad \left\|f^{-1}(y) - f^{-1}(\tilde{y})\right\| \leq 2\|y - \tilde{y}\| \\
&\Rightarrow \quad f^{-1} \text{ ist stetig.}
\end{aligned}$$

Zu zeigen: *Differenzierbarkeit* von f^{-1}

Sei f stetig $\Rightarrow \exists \delta > 0 \quad x \in B_\delta(0) \Rightarrow f(x) \in B_\varepsilon(0)$,

$D' := B_\delta(0) \subset \bar{B}_{2\varepsilon}(0) \subset D$, $E := f(D') \Rightarrow f : D' \to E$ bijektiv.

E ist offen, denn E ist Urbild der offenen Menge D' unter der stetigen Abbildung f^{-1}. D.h. f^{-1} ist stetig differenzierbar.

(vgl. hierzu auch Satz 2.3 und Satz 3.7) $\square$

Satz 3.32 (*Implizite Funktionen*)

Sei $D \subset \mathbb{R}^p \times \mathbb{R}^n$ offen, $f : D \to \mathbb{R}^n$ stetig differenzierbar, $f(u_0, x_0) = 0$ mit $(u_0, x_0) \in D$, es gelte

$$\det \frac{\partial f}{\partial x}(u_0, x_0) = \det \left(\frac{\partial f_i}{\partial x_i}\right)_{(u_0, x_0)} \neq 0.$$

Dann existieren offene Mengen $D_u \subset \mathbb{R}^p$ mit $u_0 \in D_u$, $D_x \subset \mathbb{R}^n$ mit $x_0 \in D_x$ sowie eine stetig differenzierbare Funktion $g : D_u \to D_x$, sodaß

$$\forall u \in D_u, x \in D_x \quad f(u, x) = 0 \Leftrightarrow x = g(u).$$

Beweis

Definiere $F : D \to \mathbb{R}^p \times \mathbb{R}^n$, $F(u, x) = (u, f(u, x))$. Dann gilt

$$F' = \begin{pmatrix} \mathrm{id}_{\mathbb{R}^p} & 0 \\ \frac{\partial f}{\partial u} & \frac{\partial f}{\partial x} \end{pmatrix}, \det F'(u_0, x_0) = \det \frac{\partial f}{\partial x}(u_0, x_0) \neq 0.$$

Nach Umkehrsatz $\exists D' \subset \mathbb{R}^p \times \mathbb{R}^n$ mit $(u_0, x_0) \in D'$ offen, $\exists E \subset \mathbb{R}^p \times \mathbb{R}^n$ offen, sodaß $F : D' \to E$ bijektiv und $F^{-1} : E \to D'$ stetig differenzierbar.

Wähle eine offene Produktmenge $E_u \times E_y \subset E$ mit $F(u_0, x_0) \in E_u \times E_y$.

Setze $D'' := F^{-1}(E_u \times E_y)$. Da F stetig ist, ist nach Satz 2.3 D'' offen.

Schreibe F^{-1} block-komponentenweise $F^{-1} = \left(F_u^{-1}, F_x^{-1}\right)$.

Setze $D''_u := \{u \in \mathbb{R}^p \mid \exists x \in \mathbb{R}^n \ (u,x) \in D''\}$.

Definiere $g := D''_u \to \mathbb{R}^n, g(u) = F_x^{-1}(u,0)$. Dies ist wohldefiniert, denn

$$\begin{aligned} u \in D''_u &\Rightarrow \exists x \ (u,x) \in D'' \\ &\Rightarrow F(u,x) = (u,y) \in E_u \times E_y \\ &\Rightarrow (u,0) \in E_u \times E_y, \end{aligned}$$

also ist F_x^{-1} anwendbar.

Schränke D'' weiter ein und wähle eine offene Produktmenge $D_u \times D_x \subset D''$ mit $u_0 \in D_u, x_0 \in D_x$.

Setze $E' := F(D_u \times D_x)$. E' ist offen, da F^{-1} stetig.

Dann gilt

$$\begin{aligned} \forall u \in D_u, x \in D_x \ f(u,x) = 0 &\Leftrightarrow F(u,x) = (u,0) \\ &\Leftrightarrow F^{-1}(u,0) = (u,x) \\ &\Leftrightarrow x = F_x^{-1}(u,0) = g(u). \end{aligned}$$

□

Beispiel 1

Ebene Kurven in $\mathbb{R}^2$ definiert durch $f(x,y) = 0$. Gilt nun z.B. $f_y(x_0,y_0) \neq 0$ und $f(x_0,y_0) = 0$, dann läßt sich die Kurve lokal durch $y = \varphi(x)$ beschreiben. Es gilt

$$\varphi'(x) = -\frac{f_x(x,y)}{f_y(x,y)}.$$

Die Tangentengleichung im Punkt (x,y) mit eigenen lokalen Koordinaten ξ, η bestimmt sich dann zu

$$\eta - y = -\frac{f_x(x,y)}{f_y(x,y)}(\xi - x),$$

bzw. in symmetrischer Form zu $f_x(x,y)(\xi - x) + f_y(x,y)(\eta - y) = 0$.

Entsprechend für *Hyperflächen* mit $f(x_1,\ldots,x_n) = 0$. Mit $\mathrm{grad} f \neq 0$ ist

$$\langle \mathrm{grad} f(x), \xi - x \rangle = 0$$

die Gleichung der Tangentialhyperebene mit den Koordinaten $\xi = (\xi_1,\ldots,\xi_n)$.

Beispiel 2

Eine *Raumkurve in* $\mathbb{R}^3$ läßt sich i.allg. als Durchschnitt zweier Flächen $f(x,y,z) = 0$ und $g(x,y,z) = 0$ bestimmen. Dies läßt sich lokal mittels Satz 3.32 (Implizite Funktionen) zeigen durch die Jacobi-Matrix

$$\begin{pmatrix} f_x & f_y & f_z \\ g_x & g_y & g_z \end{pmatrix}.$$

Ist $f_x g_y - f_y g_x$ an einer Stelle mit $f = g = 0$ von 0 verschieden, dann ist in einer Umgebung dieser Stelle (x, y) eine Darstellung von z als stetig differenzierbare Funktion mit $x = \varphi(z), y = \psi(z)$ möglich. Damit erhält man

$$\begin{pmatrix} \varphi' \\ \psi' \end{pmatrix} = -\begin{pmatrix} f_x & f_y \\ g_x & g_y \end{pmatrix}^{-1} \begin{pmatrix} f_z \\ g_z \end{pmatrix} = \begin{pmatrix} \frac{f_y g_z - f_z g_y}{f_x g_y - f_y g_x} \\ \frac{f_z g_x - f_x g_z}{f_x g_y - f_y g_x} \end{pmatrix}$$

Analoges Resultat erzielbar unter Anwendung der Kettenregel aus

$$\left.\begin{array}{r} f(\varphi(z), \psi(z), z) = 0 \\ g(\varphi(z), \psi(z), z) = 0 \end{array}\right\} \Rightarrow \left\{\begin{array}{l} f_x\varphi' + f_y\psi' + f_z = 0 \\ g_x\varphi' + g_y\psi' + g_z = 0 \end{array}\right.$$

und Auflösen wie oben.

Extrema unter Nebenbedingungen

Beispiel

Betrachtet wird die *Ellipse* $ax^2 + 2bxy + cy^2 = 1$ bzw.

$$g(x, y) = ax^2 + 2bxy + cy^2 - 1 = 0$$

mit $a > 0, ac - b^2 > 0$. Finde Punkte der Ellipse, die von $(0, 0)$ maximales bzw. minimales *Abstandsquadrat* (Quadrat des EUKLIDischen Abstandes) haben, d.h.

- finde Extrema «Abstand»,
  hier von $f(x, y) = x^2 + y^2$,
- unter der Nebenbedingung «Ellipse»,
  hier von $g(x.y) = ax^2 + 2bxy + cy^2 - 1 = 0$.

Bemerkung

Das Minimum von f in $(0, 0)$ erfüllt nicht die Nebenbedingung.

Naives, nichtdestotrotz richtiges, wenngleich aufwendiges Vorgehen:

Erfülle $g(x, y) = 0$ durch Parameterdarstellung $x = \varphi(t), y = \psi(t)$ und bestimme die Extremalstellen von $F(t) = f(\varphi(t), \psi(t))$. Erweist sich als mühsam.

Einfacher: *Kritische* Stellen von F genügen (vgl. Def. 3.9, Bem. zu Def. 3.9 und Satz 3.15) bei Anwendung der Kettenregel bei parametrischer Darstellung der Gleichung

$$F'(t) = D_1 f(\varphi(t), \psi(t))\varphi'(t) + D_2 f(\varphi(t), \psi(t))\psi'(t) = 0$$

und wegen Nebenbedingung $g(\varphi(t), \psi(t)) = 0$ gilt $\forall t$

$$D_1 g(\varphi(t), \psi(t))\varphi'(t) + D_2 g(\varphi(t), \psi(t))\psi'(t) = 0.$$

Für jede kritische Stelle von F sind die Vektoren $(D_1 f, D_2 f)$ und $(D_1 g, D_2 g)$ linear abhängig. D.h.

$\exists\lambda\ D_1 f(x, y) - \lambda D_1 g(x, y) = 0$ und $D_2 f(x, y) - \lambda D_2 g(x, y) = 0$ und $g(x, y) = 0$.

Parameterdarstellung ist verschwunden.

Fazit

> *Kritische* Stellen von $F \Leftrightarrow$ *kritische* Stellen von $G = f - \lambda g$ in $\mathbb{R}^2$ mit $g(x, y) = 0$.

Im Beispiel der Ellipse ist $D_1 f(x, y) - \lambda D_1 g(x, y) = 2x - 2\lambda(ax + by) = 0$ bzw. $D_2 f(x, y) - \lambda D_2 g(x, y) = 2y - 2\lambda(bx + cy) = 0$, d.h. es ist das lineare Gleichungssystem zu lösen mit $(1 - \lambda a)x - \lambda by = 0$ und $-\lambda bx + (1 - \lambda c)y = 0$ sowie $g(0, 0) = -1$. Nur nichttriviale Lösungen sind zu betrachten.

λ muß genügen

$$\begin{vmatrix} 1 - \lambda a & -\lambda b \\ -\lambda b & 1 - \lambda c \end{vmatrix} = 0 \Leftrightarrow b^2\lambda^2 - (1 - \lambda a)(1 - \lambda c) = 0.$$

Die Aufgabe, Extremwertprobleme unter Nebenbedingungen zu lösen, kann in allgemeiner Form mit Hilfe der LAGRANGEschen Multiplikatoren λ angegeben werden:

Satz 3.33

> Sei $U \subset \mathbb{R}^n$, $f : U \to \mathbb{R}$ stetig differenzierbar, $g = (g_1, \ldots, g_r)$ mit (mindestens) in U erklärten stetig differenzierbaren reellwertigen Funktionen $g_1, \ldots, g_r$, $M \subset U$ erklärt durch $g_1(x) = 0, \ldots, g_r(x) = 0$, die Ableitung von g von U in $\mathbb{R}^r$ habe überall den Rang r, d.h. M sei eine $(n - r)$-dimensionale Untermannigfaltigkeit des $\mathbb{R}^n$.
>
> Hat dann die Einschränkung $f|_M$ in $a \in M$ ein lokales Extremum, dann existieren $\lambda_1, \ldots, \lambda_r$ so, daß a eine kritische Stelle ist für die in U erklärte Funktion
> $$G := f - \lambda_1 g_1 - \ldots - \lambda_r g_r,$$
> $$\text{d.h.} \quad f'(a) = \sum_{i=1}^r \lambda_i g_i'(a) \quad \text{bzw.} \quad \mathrm{grad} f(a) = \sum_{i=1}^r \lambda_i \mathrm{grad} g_i(a).$$
> in $\mathbb{R}^n$ bzw. unter Zugrundelegung der Standardbasis in $\mathbb{R}^n$.

Beweis

Die Untermannigfaltigkeit M kann in einer Umgebung von a mittels Parameterdarstellung beschrieben werden als $\varphi : V \to \mathbb{R}^n$ mit $V \subset \mathbb{R}^p$ und $p = n - r$.

Die Verkettung $F = f \circ \varphi$ muß im Urbild $b \in V$ von a eine kritische Stelle haben, da F in b ein lokales Extremum hat, wie auch $f|_M$.

Demzufolge hat man

$$(\star) \quad F'(b, k) = f'(a, \varphi'(b, k)) = 0$$

für alle $k \in \mathbb{R}^p$, und wegen $g_i(\varphi(u)) = 0$ für alle $u \in V$ hat man

$$(\star\star) \quad g_i'(a, \varphi'(b, k)) = 0$$

für alle $k \in \mathbb{R}^p$.

Zu zeigen ist, daß es $\lambda_1, \dots \lambda_r$ gibt, für die

$$(\star\star\star) \quad f'(a,h) - \sum_{i=1}^{r} \lambda_i g_i'(a,h) = 0$$

für alle $h \in \mathbb{R}^n$ ist.

Aus den Bedingungen $(\star)$ und $(\star\star)$ folgt, daß für die h, die im Tangentialraum in a an M liegen, d.h. für $\varphi'(b,k)$ mit beliebigem $k \in \mathbb{R}^p$, die Gleichung $(\star\star\star)$ für alle $\lambda_1, \dots, \lambda_r$ erfüllt ist.

Die Gleichung $(\star\star\star)$ gilt sogar für alle $h \in \mathbb{R}^n$. Denn bei Wahl des Komplementärraumes W zum Tangentialraum in a mit $\dim(W) = r$, wird dieser von $w_1, \dots, w_r$ aufgespannt. Dann lassen sich die $\lambda_1, \dots, \lambda_r$, weil nach Voraussetzung für beliebiges $h \in \mathbb{R}^n$

$$\operatorname{rang} \begin{pmatrix} g_1'(a,w_1) & \dots & g_r'(a,w_1) \\ \vdots & & \vdots \\ g_1'(a,w_r) & \dots & g_r'(a,w_r) \end{pmatrix} = r$$

ist, bestimmen aus dem linearen Gleichungssystem

$$\begin{array}{ll} f'(a,w_1) - \lambda_1 g_1'(a,w_1) - \dots - \lambda_r g_r'(a,w_1) & = 0 \\ \vdots & \vdots \\ f'(a,w_r) - \lambda_1 g_1'(a,w_r) - \dots - \lambda_r g_r'(a,w_r) & = 0. \end{array}$$

Da jedes beliebige $h \in \mathbb{R}^n$ dargestellt werden kann als $h = \varphi'(b,k) + c_1 w_1 + \dots + c_r w_r$, ist die Gleichung $(\star\star\star)$ für alle $h \in \mathbb{R}^n$ erfüllt.

Also hat $G := f - \lambda_1 g_1 - \dots - \lambda_r g_r$ für geeignete λ_i in a eine kritische Stelle. $\Box$

Vorgehensweise zur Bestimmung lokaler Extrema von f unter den Nebenbedingungen $g_1 = 0, ..., g_r = 0$

1. Wähle den Ansatz $G := f - \lambda_1 g_1 - \dots - \lambda_r g_r$ und bestimme hierfür die kritischen Stellen von G, die die Nebenbedingungen erfüllen.

2. Ermittle die $x_1, \dots, x_n, \lambda_1, \dots, \lambda_r$ aus den $n + r$ Gleichungen

$$\begin{array}{rcl} D_1 f(x) & = & \lambda_1 D_1 g_1(x) + \dots + \lambda_r D_1 g_r(x) \\ \vdots & & \vdots \\ D_n f(x) & = & \lambda_1 D_n g_1(x) + \dots + \lambda_r D_n g_r(x) \\ g_1(x) & = & 0 \\ \vdots & & \vdots \\ g_r(x) & = & 0 \end{array}$$

3. Prüfe auf Vorliegen von wirklichen Extrema (Minima, Maxima) oder Sattelpunkten.

Anwendung

Bestimmung der Extrema einer quadratischen Form als Bestimmung der Eigenwerte

Sei A eine symmetrische $n \times n$-Matrix.

Es ist $f(x) = x^\top Ax$ mit $x \in \mathbb{R}^n$.

Finde das Maximum von f auf Einheitssphäre S^{n-1}, d.h. unter den Nebenbedingungen $g(x) = \langle x, x\rangle - 1 = x^\top x - 1 = 0$.

Da S^{n-1} kompakt ist, hat f ihr Maximum m an einer Stelle $v \in S^{n-1}$.

Mit $g'(x) = 2x^\top \neq 0$ für alle $x \in S^{n-1}$ folgt $\exists\lambda\ f'(x) = \lambda g'(x)$, d.h. mit $2v^\top A = 2\lambda v^\top$ das Eigenwertproblem $Av = \lambda v$.

Weiterhin gilt für $v^\top v = 1 : \lambda = \lambda v^\top v = v^\top Av =_{\text{def}} f(v) =_{\text{def}} m$.

Fazit

Jede Maximalstelle v von f auf S^{n-1} ist ein Eigenvektor von A und das Maximum $m = f(v)$ ist der Eigenwert zu v.

Verallgemeinerung:

Satz 3.34 (*Hauptachsentransformation*)

Die symmetrische reelle $n \times n$-Matrix A besitzt Eigenvektoren $v_1, \ldots, v_n$ und Eigenwerte $\lambda_1, \ldots, \lambda_n$ mit folgenden Eigenschaften:

- $v_1, \ldots, v_n$ sind paarweise senkrecht,
- die Eigenwerte λ_k zu v_k sind Maxima von f auf $S^{n-1} \cap H_{k-1}$ mit der orthogonalen Zerlegung $H_0 := \mathbb{R}^n, H_k = [v_1, \ldots, v_k]^\perp$.

4 Gewöhnliche Differentialgleichungen

4.1 Problemstellung. Klassifikation

Differentialgleichung (DGL) wird eine Gleichung genannt, in der neben einer oder mehreren unabhängigen Veränderlichen und einer oder mehreren Funktionen dieser Veränderlichen auch noch die Ableitungen dieser Funktionen nach den unabhängigen Veränderlichen auftreten.

Klassifikation

Nach der Anzahl der in ihnen enthaltenen unabhängigen Veränderlichen (Variablen) werden unterschieden

- *gewöhnliche Differentialgleichungen* (GDGL) als Funktionen *einer* Variablen,

- *partielle Differentialgleichungen* als Funktionen *mehrerer* Variablen.

Im folgenden werden GDGL betrachtet. Mit G ist ein *Gebiet* bezeichnet, also eine *offene* und *zusammenhängende* Teilmenge des $\mathbb{R}^n$, im Falle des $\mathbb{R}$ sind die zusammenhängenden Teilräume genau die Intervalle I.

Definition 4.1 (*explizite DGL 1.Ordnung. Lösung*)

Sei $G \subset \mathbb{R} \times \mathbb{R}$, $f : G \to \mathbb{R}$ stetig, $I \subset \mathbb{R}$ ein Intervall.

Dann

- heißt $y' = f(x, y)$ *(explizite) Differentialgleichung erster Ordnung*,
- heißt eine differenzierbare *Funktion* $\varphi : I \to \mathbb{R}$ *Lösung in* I mit *Graph*

$$\Gamma_\varphi := \{(x, y) \in I \times \mathbb{R} \mid y = \varphi(x)\} \subset G$$

und mit

$$\varphi'(x) = f(x, \varphi(x)), \quad \forall x \in I.$$

Beispiel (trivial)

$y' = f(x)$ führt nach direkter Integration zu $\varphi(x) = c + \int_{x_0}^{x} f(t)dt$.

Geometrische Interpretation

Die Differentialgleichung $y' = f(x, y)$ wird geometrisch interpretiert als *Richtungsfeld* auf $G \subset \mathbb{R}^2$. In jedem Punkt $(x, y) \in G$ wird die *Steigung* gegeben mit $y' = f(x, y)$. Die Aufgabe lautet dann:

Finde eine differenzierbare Funktion $\varphi(x)$, deren Graph in jedem Punkt in f die vorgegebene Steigung hat (Integralkurve des Richtungsfeldes).

Definition 4.2 (*System von n DGL 1.Ordnung. Lösung*)

Sei $G \subset \mathbb{R} \times \mathbb{R}^n$, $f : G \to \mathbb{R}^n$ stetig, $I \subset \mathbb{R}$ ein Intervall.

Dann

- heißt $y' = f(x, y)$ *System von n Differentialgleichungen erster Ordnung*,
- heißt eine differenzierbare *Abbildung* $\varphi : I \to \mathbb{R}^n$ *Lösung in* I mit *Graph*

$$\Gamma_\varphi := \{(x, y) \in I \times \mathbb{R}^n \mid y = \varphi(x)\} \subset G$$

und mit

$$\varphi'(x) = f(x, \varphi(x)), \quad \forall x \in I.$$

Das Gleichungssystem $y' = f(x, y)$ in Komponentenschreibweise:

$$\begin{aligned} y_1' &= f_1(x, y_1, \ldots, y_n) \\ y_2' &= f_2(x, y_1, \ldots, y_n) \\ &\vdots \\ y_n' &= f_n(x, y_1, \ldots, y_n) \end{aligned}$$

Definition 4.3 (*explizite DGL n-ter Ordnung. Lösung*)

Sei $G \subset \mathbb{R} \times \mathbb{R}^n$, $f : G \to \mathbb{R}^n$ stetig, $I \subset \mathbb{R}$ ein Intervall.

Dann

- heißt $y^{(n)} = f\left(x, y, y', \ldots, y^{(n-1)}\right)$ *explizite Differentialgleichung n-ter Ordnung*,
- heißt eine n-mal differenzierbare *Funktion* $\varphi : I \to \mathbb{R}$ *Lösung in* I mit *Graph*

  $$\Gamma_\varphi := \left\{ \left(x, y_0, y_1, \ldots, y_{(n-1)}\right) \in I \times \mathbb{R}^n \mid y_i = \varphi^{(i)}(x), 0 \leq i \leq n-1 \right\} \subset G$$

  und mit

  $$\varphi^{(n)}(x) = f\Big(x, \varphi(x), \varphi'(x), \ldots, \varphi^{(n-1)}(x)\Big), \quad \forall x \in I.$$

Bemerkung

Eine implizite Darstellung ist $\tilde{f}(x, y, y', \ldots, y^{(n)}) = 0$ für $G \subset \mathbb{R}^{n+2}$ und mit $f : G \to \mathbb{R}$.

Eine speziellere, abstrakte Sichtweise ergibt sich, wenn *Integralkurven auf Vektorfeldern* betrachtet werden:

Definition 4.4 (*Vektorfeld*)

Ein *Vektorfeld* v auf $G \subset \mathbb{R}^n$ ist eine Abbildung $v : G \to \mathbb{R}^n$, die jedem Punkt $x \in G$ einen Vektor $v(x) \in \mathbb{R}^n$ zuordnet.

Interpretation

Geschwindigkeitsfeld einer stationären Strömung.

Definition 4.5 (I*ntegralkurve auf einem Vektorfeld*)

Eine *Integralkurve auf einem Vektorfeld* $v : G \to \mathbb{R}^n$ ist eine differenzierbare Kurve $\varphi : I \to G$ mit $\dot{\varphi}(t) = v(\varphi(t))$.

Interpretation

Mit der Zeit t versehene Bahnkurve eines in der Strömung v mitgeführten Partikels.

Beispiel

Rotationsfeld $v(x, y) = (-y, x)$.

Die Bedingung für die Integralkurve lautet $\dot{\varphi}_1 = -\varphi_2$ und $\dot{\varphi}_2 = \varphi_1$, d.h. $\dot{\varphi} = \mathrm{i}\varphi$ für $\varphi = \varphi_1 + \mathrm{i}\varphi_2$. Die Lösung lautet: $\varphi(t) = c\mathrm{e}^{\mathrm{i}t}$ mit $c \in \mathbb{C}, t \in \mathbb{R}$.

Eine spezielle Interpretation eines *Systems von n Differentialgleichungen erster Ordnung* nach Def. 4.2 ergibt sich - in der Sache veranlasst durch physikalische Fragestellungen - in der Betrachtung zeitabhängiger Vektorfelder, was zum Begriff des dynamischen Systems führt.

Definition 4.6 (*dynamisches System. Lösung*)

Sei $G \subset \mathbb{R} \times \mathbb{R}^n$ offen, $(t, y) \in G$, t «Zeit».

Dann

- heißt die stetige Abbildung $F : G \to \mathbb{R}^n$ *dynamisches System*,
- heißt die differenzierbare Kurve $\varphi : I \to \mathbb{R}^n$ *Lösung* bzw. *Integralkurve* des dynamischen Systems F mit $(t, \varphi(t)) \in G$ und

$$\dot{\varphi}(t) = F(t, \varphi(t)) \Leftrightarrow \dot{y} = F(t, y), \quad \forall t \in I.$$

Bemerkung

Die Definition gilt analog für $\mathbb{C}$ mit $G \subset \mathbb{R} \times \mathbb{C}^n$, $F : G \to \mathbb{C}^n$, $\varphi : I \to \mathbb{C}^n$ mit $I \subset \mathbb{R}$.

Bisher wurden nur allgemeine Lösungen oder allgemeine Integrale betrachtet, die noch willkürliche Konstanten $c_1, \ldots, c_n$ enthalten, die durch zusätzliche Bedingungen bestimmt werden können. Man kommt so zu den partikulären Lösungen bzw. Integralen. Betreffen die Vorgaben nur eine Stelle des Definitionsbereiches, spricht man von einer Anfangswertaufgabe, sonst von einer Randwertaufgabe.

Definition 4.7 (*Anfangswertproblem*)

Sei $(t_0, y_0) \in G$ als Anfangswert gegeben.

Als *Anfangswertproblem* (Anfangswertaufgabe) wird bezeichnet die Bestimmung der *Lösung* φ mit

1. $\varphi(t_0) = y_0$,
2. $\dot{y} = F(t, y)$ und
3. $y(t_0) = y_0$.

Die Untersuchung des Anfangswertproblems bezieht sich zunächst im wesentlichen auf die Prüfung von Existenz und Eindeutigkeit von Lösungen. Bevor hierzu etwas ausgesagt wird, wird gezeigt, daß man sich bei der Betrachtung dieser oder ähnlicher Fragestellungen auf die Untersuchung von Systemen von DGL 1.Ordnung beschränken kann, denn

Satz 4.1 (*Reduktion*)

Jede Differentialgleichung n-ter Ordung läßt sich auf ein System von Differentialgleichungen 1.Ordnung zurückführen.

Beweis

Es ist zu zeigen, daß man zu *einer* beliebigen DGL n-ter Ordnung ein *System* von DGL 1.Ordnung aufstellen kann und eine *Lösung des Systems* von DGL 1.Ordnung auch *eine* Lösung der DGL n-ter Ordnung ist.

1. Sei $G \subset \mathbb{R}^{n+1}, f : G \to \mathbb{R}, y^{(n)} = f\left(x, y, y', y'', \ldots, y^{(n-1)}\right)$.

2. Schreibe die DGL n-ter Ordnung $y^{(n)}$ um als System von DGL 1.Ordnung

$$Y = \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_{n-2} \\ y_{n-1} \end{pmatrix}, F(x, Y) = \begin{pmatrix} y_0' \\ y_1' \\ \vdots \\ y_{n-2}' \\ y_{n-1}' \end{pmatrix} = Y' = \begin{pmatrix} y_0' & = & y_1 \\ y_1' & = & y_2 \\ \vdots & & \vdots \\ y_{n-2}' & = & y_{n-1} \\ y_{n-1}' & = & y_n \end{pmatrix}.$$

3. Ist $\varphi : I \to \mathbb{R}$ eine Lösung von $y^{(n)}$, dann gilt

$$\varphi^{(n)}(x) = f\left(x, \varphi(x), \varphi'(x), \ldots, \varphi^{(n-1)}(x)\right).$$

Setzt man

$$\Phi = \begin{pmatrix} \varphi_0 \\ \varphi_1 \\ \vdots \\ \varphi_{n-2} \\ \varphi_{n-1} \end{pmatrix} \text{ mit } \begin{pmatrix} \varphi_0(x) & = & \varphi(x) \\ \varphi_1(x) & = & \varphi'(x) \\ \vdots & & \vdots \\ \varphi_{n-2}(x) & = & \varphi^{(n-2)}(x) \\ \varphi_{n-1}(x) & = & \varphi^{(n-1)}(x) \end{pmatrix},$$

dann ist Φ eine Lösung von $Y' = F(x, Y)$.

4. Sei umgekehrt $\Phi : I \to \mathbb{R}^n$ eine Lösung des Systems von DGL 1.Ordnung $Y' = F(x, Y)$. Dann ist die erste Komponente von Φ, d.h. $\varphi := \varphi_0 : I \to \mathbb{R}$ Lösung der DGL n-ter Ordnung. Ist nämlich φ_{n-1} einmal differenzierbar, so folgt aus dem «Rückwärtsgang» durch die $n-1$ DGL 1.Ordnung, daß φ selbst n-mal differenzierbar ist und man bekommt als n-te Gleichung $\varphi^{(n)}(x) = f\left(x, \varphi(x), \varphi'(x), \ldots, \varphi^{(n-1)}(x)\right)$.

□

Bemerkung

Somit bedarf es auch nur einer Theorie für Systeme von DGL 1.Ordnung.

Beispiel

Bewegung eines Massenpunktes nach dem Zweiten NEWTONschen Gesetz.

Sei $x := x(t)$ der Ort, $\dot{x} := \frac{dx}{dt}$ die Geschwindigkeit und $\ddot{x} := \frac{d^2x}{dt^2}$ die Beschleunigung eines Massenpunktes.

Mit m als Masse und mit F als Kraft gilt $m\ddot{x} = F(t, x, \dot{x})$, d.h. ein System von drei DGL 2.Ordnung mit $x = (x_1, x_2, x_3)$.

Mit der Geschwindigkeit $v = (v_1, v_2, v_3)$ gewinnt man das System von sechs DGL 1.Ordnung durch $\dot{x} := v$ und $\dot{v} = \frac{1}{m} F(t, x, v)$.

4.2 Existenz. Eindeutigkeit

Betrachtet wird das Anfangswertproblem für ein System von gewöhnlichen DGL 1.Ordnung

$$y' = F(x, y) \quad \text{und} \quad y(x_0) = y_0.$$

Definition 4.8 (*Vektorraum* $\mathcal{C}([x_0 - \delta, x_0 + \delta])$)

Vektorraum oder *Raum* $\mathcal{C}([x_0 - \delta, x_0 + \delta])$ ist die Bezeichnung für die Menge der in dem Intervall $[x_0 - \delta, x_0 + \delta]$ stetigen vektorwertigen Abbildungen mit der *Operatornorm*

$$\|g\|_{\mathcal{C}} = \max_{x \in [x_0 - \delta, x_0 + \delta]} \|g(x)\|_\infty .$$

Bemerkung

$\mathcal{C}([x_0 - \delta, x_0 + \delta])$ ist normierter vollständiger Vektorraum oder *Funktionenraum.*

Satz 4.2 (*Satz von* Picard-Lindelöf)

Sei $x_0 \in I$ mit Intervall $I \subset \mathbb{R}, F : I \times \mathbb{R}^n \to \mathbb{R}$ stetig und erfülle die Lipschitz-Bedingung

$$\forall x \in I \, \forall y, z \in \mathbb{R}^n \quad \|F(x, y) - F(x, z)\|_\infty \leq L \|y - z\|_\infty .$$

Dann gibt es zu jedem Anfangswert $y_0 \in \mathbb{R}^n$ *genau eine* stetig differenzierbare Abbildung $\varphi : I \to \mathbb{R}^n$ mit $\varphi'(x) = F(x, \varphi(x))$ und $\varphi(x_0) = y_0$.

Beweis

Angenommen, das Anfangswertproblem habe die Lösung φ. Gemäß dem Hauptsatz der Differential- und Integralrechnung gilt dann die Fixpunktgleichung (Integralgleichung)

$$(\star) \quad \varphi(x) = y_0 + \int_{x_0}^{x} F(t, \varphi(t))\, dt = (\Phi(\varphi))(x),$$

d.h. φ ist Fixelement der Abbildung Φ des Raumes aller in I stetigen vektorwertigen Funktionen in sich, die durch $(\star)$ beschrieben ist.

Wir zeigen zunächst, daß für $x \in [x_0 - \delta, x_0 + \delta]$ die Abbildung Φ *erstens* Abbildung von $\mathcal{C}([x_0 - \delta, x_0 + \delta])$ in sich (Selbstabbildung) ist, was unmittelbar aus $(\star)$ folgt, und *zweitens* kontrahierend ist. Dies wird erreicht durch Fixieren von δ so, daß Φ kontrahierend wird. Hierzu wird betrachtet

$$(\Phi(\psi))(x) - (\Phi(\varphi))(x) = \int_{x_0}^{x} (F(t, \psi(t)) - F(t, \varphi(t)))\, dt.$$

Mit der Abschätzung

$$\|(\Phi(\psi))(x) - (\Phi(\varphi))(x)\|_\infty \leq \operatorname{sign}(x - x_0) \int_{x_0}^{x} \|F(t, \psi(t)) - F(t, \varphi(t))\|_\infty\, dt$$

erhält man unter Beachtung der LIPSCHITZ-Bedingung

$$\|F(t,\psi(t)) - F(t,\varphi(t))\|_\infty \leq L\,\|\psi(t) - \varphi(t)\|_\infty$$

und durch Übergang von Vektornorm auf Abbildungsnorm (Operatornorm)

$$\|\Phi(\psi) - \Phi(\varphi)\|_{\mathcal{C}} \leq \int_{x_0-\delta}^{x_0+\delta} L\,\|\psi - \varphi\|_{\mathcal{C}}\,dt = 2\delta L\,\|\psi - \varphi\|_{\mathcal{C}}\,.$$

Wählt man $\delta < \frac{1}{2L}$, so ist die Abbildung Φ kontrahierend.

Gemäß dem Fixpunktsatz von BANACH gilt (wegen Selbstabbildung und Kontraktion)

$$\exists_1 \varphi \in \mathcal{C}([x_0-\delta, x_0+\delta]) \quad \varphi(x) = y_0 + \int_{x_0}^{x} F(t,\varphi(t))\,dt.$$

Da F stetig ist, gilt $\varphi'(x) = F(x,\varphi(x))$, d.h. die Anfangsbedingung $\varphi(x_0) = y_0$ ist erfüllt. Somit gibt es genau eine Lösung φ für das AWP im Intervall $|x - x_0| \leq \delta$.

Dieses Ergebnis kann nun iterativ auf das AWP an den Stellen (Randpunkten) $x_0 - \delta$ und $x_0 + \delta$ mit den AW $\varphi(x_0 - \delta)$ und $\varphi(x_0 + \delta)$ angewandt werden und führt zu einer eindeutigen Lösung im doppelt so großen Intervall $|x - x_0| \leq 2\delta$. Durch Fortsetzen dieses Verfahrens auf die entsprechend gewählten Randpunkte der jeweils betrachteten Intervalle kann die zunächst für das Intervall $|x - x_0| \leq \delta$ gültige Existenz- und Eindeutigkeitsaussage auf das gesamte Intervall $I \subset \mathbb{R}$ erweitert werden. $\square$

Bemerkung

Die an den Beweisanfang gestellte Integralversion mittels des Hauptsatzes der Differential- und Integralrechnung ist eine äquivalente Formulierung zur DGL. D.h.

- $F : G \to \mathbb{R}^n$ stetig auf $G \subset \mathbb{R} \times \mathbb{R}^n$ offen,
- $\varphi : I \to \mathbb{R}^n$ stetig auf I mit $(x, \varphi(x)) \in G, \quad \forall x,$

löst das AWP

$$\left.\begin{array}{l} y' = F(x,y) \\ y(x_0) = y_0 \end{array}\right\} \Leftrightarrow \varphi(x) = \varphi(x_0) + \int_{x_0}^{x} F(t,\varphi(t))\,dt \quad \forall x \in I.$$

Dies ist von Vorteil bei der Betrachtung von Grenzprozessen, da die Integralversion es erlaubt, mit stetigen statt mit differenzierbaren Funktionen φ zu arbeiten.

Bemerkung

Existieren alle partiellen Ableitungen $D_1 f(x,y), \ldots, D_n f(x,y)$ und sind diese beschränkt, so erfüllt F die LIPSCHITZ-Bedingung des Satzes von PICARD-LINDELÖF.

Bemerkung

Im Satz von PICARD-LINDELÖF wird vorausgesetzt, daß f bzgl. aller Variablen stetig ist und bzgl. der y-Variablen der LIPSCHITZ-Bedingung genügt. Aus der Gültigkeit der LIPSCHITZ-Bedingung folgt, daß f bzgl. der y-Variablen *gleichmäßig* stetig ist.

(*Anm.*: Aus der LIPSCHITZ-Bedingung allein folgt nicht, daß f insgesamt stetig ist.)

Beweis

Gemäß der Dreiecksungleichung und der LIPSCHITZ-Bedingung hat man

$$\begin{aligned}\|f(x,y)-f(x_0,y_0)\|_\infty &\leq \|f(x,y)-f(x,y_0)\|_\infty+\|f(x,y_0)-f(x_0,y_0)\|_\infty\\ &\leq L\|y-y_0\|_\infty+\|f(x,y_0)-f(x_0,y_0)\|_\infty .\end{aligned}$$

Sei $\varepsilon>0$ vorgegeben. Sei $\delta>0$ so bestimmt, daß für $|x-x_0|<\delta$ gilt

$$\|f(x,y_0)-f(x_0,y_0)\|_\infty<\varepsilon.$$

Dann gilt $\forall(x,y)$ mit $|x-x_0|<\delta$ bzw. $\|y-y_0\|_\infty<\frac{\varepsilon}{L}$

$$\|f(x,y)-f(x_0,y_0)\|_\infty<L\frac{\varepsilon}{L}+\varepsilon=2\varepsilon,$$

d.h. f ist in (x_0,y_0) stetig. □

4.3 Lösungsmethoden für GDGL

Die Lösung $\varphi(x)$ einer gegebenen allgemeinen DGL ist nicht in jedem Falle explizit angebbar. Allerdings gibt es für bestimmte Klassen gewöhnlicher DGL Lösungstechniken. Solche vorzustellen, ist Aufgabe des gegenwärtigen Abschnitts.

Analytische Lösungsverfahren

- Iterationsverfahren nach PICARD-LINDELÖF
- Methode der Trennung der Variablen $y'=f(x,y))=g(x)h(x)$
- Methode der Variation der Konstanten (für homogene bzw. nichthomogene lineare DGL)
- für DGL höherer Ordnung sog. «Tricks», abhängig vom Typ
  1. Energietrick
  2. Trick der inversen Funktion
  3. EULER-LAGRANGEsche Gleichung
  4. Berührungstransformation (LEGENDRE, LIE et al.)
  5. LAPLACE-Transformation

6. u.a.m.

Führen analytische Lösungsverfahren nicht zum Ziele, verbleiben die Verfahren numerischer Approximation. Hierzu wird eine Theorie der Numerik gewöhnlicher DGL benötigt. Vorstehender Sachverhalt gilt grundsätzlich auch für partielle DGL.

Iterationsverfahren nach Picard-Lindelöf

Dieses Iterationsverfahren (Fixpunktiteration) beruht auf der Integraldarstellung des AWP. Hierzu definiert man die Abbildungen $\varphi_k : I \to \mathbb{R}^n, k \in \mathbb{N}$ durch die Rekursion

$$\begin{aligned}\varphi_0(x) &:= y_0\\ \varphi_{k+1}(x) &:= y_0 + \int_{x_0}^{x} f(t, \varphi_k(t))\, dt.\end{aligned}$$

Durch vollständige Induktion und mit Anwendung des Majorantenkriteriums läßt sich zeigen, daß die Folge (φ_k) gleichmäßig gegen die Lösung des AWP

$$\varphi(x) := y_0 + \int_{x_0}^{x} f(t, \varphi(t))\, dt.$$

konvergiert. Manchmal kann so die Lösung eines AWP direkt berechnet werden.

Beispiel

In $\mathbb{R} \times \mathbb{R}$ wird zu der DGL $y' = 2xy$ die Lösung $\varphi(x)$ gesucht, die der Bedingung $\varphi(0) = y_0$ genügt. Dann lautet die Rekursionsgleichung

$$\varphi_{k+1}(x) = y_0 + 2\int_{x_0}^{x} t\varphi_k(t)\, dt.$$

Die ersten drei Glieder der Funktionenfolge sind

$$\begin{aligned}\varphi_0(x) &= y_0 & &\\ \varphi_1(x) &= y_0 + 2y_0\textstyle\int_0^x t dt &&= y_0\left(1+x^2\right)\\ \varphi_2(x) &= y_0 + 2y_0\textstyle\int_0^x t\left(1+t^2\right)dt &&= y_0\left(1+x^2+\frac{x^4}{2}\right)\end{aligned}$$

Durch vollständige Induktion kann man zeigen

$$\varphi_k(x) = y_0\left(1 + x^2 + \frac{x^4}{2!} + \frac{x^6}{3!} + \ldots + \frac{x^{2k}}{k!}\right).$$

Hieraus erhält man

$$\varphi(x) = \lim_{k\to\infty} \varphi_k(x) = y_0 \sum_{k=0}^{\infty} \frac{x^{2k}}{k!} = y_0 e^{x^2}.$$

Die Probe durch Einsetzen ergibt

$$\begin{array}{lccc} 1. & y' = 2xy & : & y_0 e^{x^2} 2x = 2xy_0 e^{x^2}\\ 2. & y(0) = \varphi(0) & : & y_0 e^0 = y_0\end{array}$$

Methode der Trennung der Variablen

Die Methode der Trennung der Variablen beruht auf nachstehenden Definition und Satz.

Definition 4.9 (*DGL mit getrennten Variablen*)

Seien $I, J \subset \mathbb{R}$ offen, $f : I \to \mathbb{R}, g : J \to \mathbb{R}$ stetig, $g(y) \neq 0 \, \forall y \in J$.

Dann ist $y' = f(x)\, g(y)$ eine *Differentialgleichung mit getrennten Variablen.*

Satz 4.3 (*Lösung von DGL mit getrennten Variablen*)

Seien $I, J \subset \mathbb{R}$ offen, $(x_0, y_0) \in I \times J$, $F(x) = \int_{x_0}^{x} f(t)\, dt$, $G(y) = \int_{y_0}^{y} \frac{1}{g(t)} dt$.

Ist für $I' \subset I$ mit $x_0 \in I'$, $F(I') \subset G(J)$, dann existiert *genau eine* Lösung $\varphi : I' \to \mathbb{R}$ von $y' = f(x)\, g(y)$ mit $\varphi(x_0) = y_0$. Die Lösung erfüllt $G(\varphi(x)) = F(x) \; \forall x \in I'$.

Beweis

«Eindeutigkeit» *Zu zeigen*:

Ist $\varphi : I' \to \mathbb{R}$ irgendeine Lösung von $y' = f(x)\, g(y)$ mit $\varphi(x_0) = y_0$, dann gilt

$$G(\varphi(x)) = F(x) \quad \forall x \in I.$$

Denn

$$\begin{aligned} \varphi'(x) = f(x)\, g(\varphi(x)) \qquad & \Rightarrow \qquad \int_{x_0}^{x} \frac{\varphi'(t)}{g(\varphi(t))} dt = \int_{x_0}^{x} f(t)\, dt \\ & \Rightarrow_{u=\varphi(t)} \quad \int_{y_o}^{\varphi(x)} \frac{du}{g(u)} = \int_{x_0}^{x} f(t)\, dt \\ & \Rightarrow \qquad G(\varphi(x)) = F(x). \end{aligned}$$

Wegen $G'(y) = \frac{1}{g(y)} \neq 0$ ist G streng monoton. Somit existiert eine stetig differenzierbare Umkehrfunktion $H : G(y) \to \mathbb{R}$ und damit

$$\varphi(x) = H(F(x)) \quad \forall x \in I',$$

d.h. die Lösung, falls sie überhaupt existiert, ist durch die Anfangsbedingung eindeutig bestimmt.

«Existenz»

$\varphi : I' \to \mathbb{R}$ mit $\varphi(x) := H(F(x))$ ist stetig differenzierbar.

Wegen $F(x_0) = 0 = G(y_0)$ ist

$$\varphi(x_0) = H(F(x_0)) = H(0) = y_0.$$

Weiterhin gilt $G(\varphi(x)) = F(x)$ und somit

$$G'(\varphi(x))\,\varphi'(x) = \frac{\varphi'(x)}{g(\varphi(x))} = F'(x) = f(x).$$

Daraus folgt $\varphi'(x) = f(x)\,g(\varphi(x))$ und φ existiert.

Es gibt somit genau eine Lösung. □

Vorgehensweise (vereinfacht)

$$\begin{aligned} y' &= f(x)\,g(x) \\ \frac{dy}{dx} &= f(x)\,g(x) \\ \frac{dy}{g(y)} &= f(x)\,dx \\ \int \frac{dy}{g(y)} &= \int f(x)\,dx + \text{const.} \end{aligned}$$

Auflösung der Schlußgleichung nach $y = \varphi(x)$.

Beispiel

In $\mathbb{R}^2$ wird zu der DGL $y' = y^2$ die Lösung $\varphi(x)$ gesucht, die der Bedingung $\varphi(0) = y_0$ genügt. Da die Lipschitz-Bedingung erfüllt ist, existiert eine eindeutige Lösung.

Fall 1 $y_0 = 0$

Ansatz $\varphi(x) = 0,\ \forall x \in \mathbb{R}$. Probe durch Einsetzen in DGL und Bedingung ergibt, daß diese erfüllt sind. D.h. $\varphi(x) = 0$ ist einzige Lösung.

Fall 2 $y_0 > 0$

Sei $\varphi : I \to \mathbb{R}$ irgendeine Lösung des AWP über I, das den Nullpunkt enthält. Dann ist $\varphi(x) > 0,\ \forall x \in I$. Denn sonst gäbe es ein $x_1 \in I$ mit $\varphi(x_1) = 0$ und daher $\varphi(x) = 0,\ \forall x \in I$, was im Widerspruch zur Annahme steht. Man kann also die Betrachtung auf $y' = y^2$ in $\mathbb{R} \times \{x \in \mathbb{R} \mid x > 0\}$ beschränken.

Es wird $y' = y^2$ dann zur DGL mit getrennten Variablen $y' = y^2 \cdot 1$ durch die Notationen

$$\begin{aligned} f : \mathbb{R} \to \mathbb{R} \quad &\text{mit} \quad f(x) = 1 \\ g : \{x \in \mathbb{R} \mid x > 0\} \to \mathbb{R} \quad &\text{mit} \quad g(y) = y^2 \end{aligned}$$

sowie

$$\begin{aligned} F(x) &= \int_0^x 1dx &&= x \\ G(y) &= \int_{y_0}^y \tfrac{1}{t^2}dt &&= \tfrac{1}{y_0} - \tfrac{1}{y}. \end{aligned}$$

Wegen $G(\{x \in \mathbb{R} \mid x > 0\}) = \left]-\infty, \frac{1}{y_0}\right[$ ist $I' := \left]-\infty, \frac{1}{y_0}\right[$ das maximale Intervall mit $F(I') \subset G(\{x \in \mathbb{R} \mid x > 0\})$. Mit Satz 4.3 wird dann

$$\frac{1}{y_0} - \frac{1}{\varphi(x)} = x \Rightarrow \varphi(x) = \frac{y_0}{1 - y_0 x}, \quad \left(x < \frac{1}{y_0}\right).$$

Fall 3 $y_0 < 0$ analog mit

$$\varphi : \left]\tfrac{1}{y_0}, \infty\right[\to \mathbb{R} \quad \text{und} \quad \varphi(x) = \tfrac{y_0}{1 - y_0 x}, \ \left(x > \tfrac{1}{y_0}\right).$$

Methode der Variation der Konstanten

Die mit «Variation der Konstanten» bezeichnete Lösungstechnik erstreckt sich auf homogene bzw. nichthomogene lineare DGL.

Definition 4.10 (*homogene, inhomogene lineare DGL 1.Ordnung*)

Sei $I \subset \mathbb{R}$ Intervall, $a, b : I \to \mathbb{R}$ stetige Funktionen.

$$y' = a(x)\, y + b(x)$$

heißt im Falle $b = 0$ *homogene*, im Falle $b \neq 0$ *inhomogene lineare* Differentialgleichung 1.Ordnung.

Satz 4.4

Sei $I \subset \mathbb{R}$ Intervall, $x_0 \in I, y_0 \in \mathbb{R}$, $\varphi : I \to \mathbb{R}$ Lösung der *homogenen* linearen Differentialgleichung 1.Ordnung $y' = a(x)\, y$ mit Anfangswert $\varphi(x_0) = y_0$.

Es gibt *genau eine* Lösung $\varphi : I \to \mathbb{R}$ mit

$$\varphi(x) = y_0 \exp\left(\int_{x_0}^{x} a(t)\, dt\right).$$

Beweis

1. Die Lösung $\varphi(x) = y_0 \exp\left(\int_{x_0}^{x} a(t)\, dt\right)$ erfüllt $y' = a(x)\, y$ mit $\varphi'(x) = a(x)\, \varphi(x)$ und $\varphi(x_0) = y_0$.
2. Da die Funktion $a(x)\, y$ in $I \times \mathbb{R}$ stetig partiell nach y differenzierbar ist, ist die LIPSCHITZ-Bedingung lokal erfüllt. Mit Satz 4.2 folgt daraus die Eindeutigkeit.

$\square$

Im Falle der *inhomogenen* linearen DGL 1.Ordnung $y' = a(x)\, y + b(x)$ mit Anfangswert $\psi(x_0) = y_0$ existiert unter Voraussetzung der LIPSCHITZ-Stetigkeit nach dem Satz von PICARD-LINDELÖF ebenfalls genau eine Lösung $\psi(x)$.

Löse zuerst die *homogene* lineare DGL 1.Ordnung $y' = a(x)\, y$ zum Anfangswert $\varphi(x_0) = 1$, d.h.

$$\varphi(x) = 1 \cdot \exp\left(\int_{x_0}^{x} a(t)\, dt\right).$$

Es gilt $\forall x \quad \varphi(x) \neq 0$. Definiere

$$u(x) := \frac{\psi(x)}{\varphi(x)}.$$

Dann

$$
\begin{array}{lrcl}
 & a(x)\varphi(x)u(x)+b(x) & = & a(x)\psi(x)+b(x) \\
 & & = & \psi'(x) \\
 & & = & (\varphi(x)u(x))' \\
 & & = & \varphi'(x)u(x)+\varphi(x)u'(x) \\
 & & = & a(x)\varphi(x)u(x)+\varphi(x)u'(x) \\
\Rightarrow & b(x) & = & \varphi(x)u'(x) \\
\Rightarrow & u'(x) & = & \frac{b(x)}{\varphi(x)} \\
\Rightarrow & u(x) & = & u(x_0)+\int_{x_0}^{x} u(t)\,dt \\
 & & = & y_0+\int_{x_0}^{x}\frac{b(t)}{\varphi(t)}dt \\
\Rightarrow & \psi(x) & = & \varphi(x)u(x) \\
 & & = & \varphi(x)\left(y_0+\int_{x_0}^{x}\frac{b(t)}{\varphi(t)}dt\right) \\
 & & = & 1\cdot\exp\left(\int_{x_0}^{x}a(t)\,dt\right)\left(y_0+\int_{x_0}^{x}\frac{b(t)}{\varphi(t)}dt\right).
\end{array}
$$

Beispiel

Sei $y'=2xy+x^3$ und $y(0)=y_0$. Dann ist

$$
\begin{aligned}
\varphi(x) &= 1\cdot\exp\left(\int_0^x 2t dt\right) \\
&= e^{x^2} \\
\psi(x) &= \varphi(x)u(x) \\
&= e^{x^2}\left(y_0+\int_0^x e^{-t^2}t^3 dt\right)
\end{aligned}
$$

Mit Substitution $s=t^2$ und partieller Integration erhält man

$$
\begin{aligned}
\psi(x) &= e^{x^2}\left(y_0+\frac{1}{2}-\frac{1}{2}\left(x^2+1\right)e^{-x^2}\right) \\
&= \left(y_0+\frac{1}{2}\right)e^{x^2}-\frac{1}{2}\left(x^2+1\right).
\end{aligned}
$$

Satz 4.5

Sei $I\subset\mathbb{R}$ Intervall, $a,b:I\to\mathbb{R}$ stetige Funktionen, $x_0\in I, y_0\in\mathbb{R}$ beliebig.

Es gibt *genau eine* Lösung $\psi:I\to\mathbb{R}$ der *inhomogenen* linearen Differentialgleichung 1.Ordnung

$$y'=a(x)y+b(x)$$

mit $\psi(x_0)=y_0$:

$$\psi(x)=\varphi(x)\left(y_0+\int_{x_0}^{x}\varphi(t)^{-1}b(t)\,dt\right) \text{ mit } \varphi(x)=\exp\left(\int_{x_0}^{x}a(t)\,dt\right).$$

Lineare DGL n-ter Ordnung mit konstanten Koeffizienten

Hier ist die Lösung äquivalent zu der Berechnung der Nullstellen eines Polynoms n-ten Grades. Die homogene lineare DGL n-ter Ordnung mit konstanten Koeffizienten $a_i \in \mathbb{R}$ bzw. $a_i \in \mathbb{C}$ kann geschrieben werden

$$\begin{gathered}-a_n y^{(n)} = a_0 y + a_1 y' + a_2 y'' + \ldots + a_{n-1} y^{(n-1)} \\ \Leftrightarrow \\ a_0 y + a_1 y' + a_2 y'' + \ldots + a_{n-1} y^{(n-1)} + a_n y^{(n)} = 0 \\ \Leftrightarrow \\ a_0 y^{(0)} + a_1 y^{(1)} + a_2 y^{(2)} + \ldots + a_{n-1} y^{(n-1)} + a_n y^{(n)} = 0\end{gathered}$$

Wählt man für die $y^{(i)} = \frac{d^i y}{dx^i} = \frac{d^i}{dx^i} y$ den Operator T^i anstelle $\frac{d^i}{dx^i}$, so wird der Term

$$\begin{gathered}a_0 y + a_1 \tfrac{d}{dx} y + a_2 \tfrac{d^2}{dx^2} y + \ldots + a_{n-1} \tfrac{d^{n-1}}{dx^{n-1}} y + a_n \tfrac{d^n}{dx^n} y \\ =: \\ \left(a_0 + a_1 \tfrac{d}{dx} + a_2 \tfrac{d^2}{dx^2} + \ldots + a_{n-1} \tfrac{d^{n-1}}{dx^{n-1}} + a_n \tfrac{d^n}{dx^n}\right) y \\ =: \\ \left(a_0 T^0 + a_1 T^1 + a_2 T^2 + \ldots + a_{n-1} T^{n-1} + a_n T^n\right) y \\ =: \\ P(T)\, y\end{gathered}$$

Dann kann die homogene lineare DGL n-ter Ordnung mit konstanten Koeffizienten $a_i \in \mathbb{R}$ bzw. $a_i \in \mathbb{C}$ geschrieben werden als

$$\begin{gathered}a_0 y + a_1 y' + a_2 y'' + \ldots + a_{n-1} y^{(n-1)} + a_n y^{(n)} = 0 \\ \Leftrightarrow \\ P(T)\, y = P\left(\tfrac{d}{dx}\right) y = \textstyle\sum_{i=0}^{n} a_i y^{(i)} = 0.\end{gathered}$$

Fazit

Der lineare Differentialoperator n-ter Ordnung $P(T)$ wird betrachtet als Polynom n-ten Grades eines Differentialoperators $\frac{d}{dx}$, d.h. in der Unbekannten $T = \frac{d}{dx}$.

Die lineare DGL n-ter Ordnung wird betrachtet als die Anwendung von $P(T)$ auf $y(x)$.

Rechenregeln für Polynome von Differentialoperatoren

Sei $P_1(T), P_2(T) \in \mathbb{P}(T)$.

Addition

$$P(T) := P_1(T) + P_2(T)$$

Sei $y : I \to \mathbb{C}$ genügend oft differenzierbar, dann

$$P\left(\frac{d}{dx}\right) y = P_1\left(\frac{d}{dx}\right) y + P_2\left(\frac{d}{dx}\right) y.$$

Multiplikation

$$Q(T) := P_1(T)\, P_2(T)$$

Sei $y : I \to \mathbb{C}$ genügend oft differenzierbar, dann

$$Q\left(\frac{d}{dx}\right) y = P_1\left(\frac{d}{dx}\right)\left(P_2\left(\frac{d}{dx}\right) y\right).$$

Bemerkung

Mit *konstanten* Koeffizienten $a_i \in \mathbb{R}$ bzw. $a_i \in \mathbb{C}$ gilt *Kommutativität*

$$P_1\left(\frac{d}{dx}\right) P_2\left(\frac{d}{dx}\right) = P_2\left(\frac{d}{dx}\right) P_1\left(\frac{d}{dx}\right).$$

Wegen der Produktregel für Ableitungen gilt dies *nicht* im Falle nichtkonstanter Koeffizienten $a_i(x)$.

Betrachtet sei nun die Anwendung $P(T) = P\left(\frac{d}{dx}\right)$ auf

$$y(x) := e^{\lambda x} \quad \text{bzw.} \quad \tfrac{dy}{dx} = \lambda e^{\lambda x}.$$

Lemma 4.6

$$\forall P(T) \in \mathbb{P}(T)\ \forall \lambda \in \mathbb{C} \quad P\left(\frac{d}{dx}\right) e^{\lambda x} = P(\lambda)\, e^{\lambda x}.$$

Beweis

Sei $P(T) \in \mathbb{P}(T)$ mit $P(T) = \sum_{i=0}^{n} a_i T^i$. Dann

$$\begin{aligned} P\left(\frac{d}{dx}\right) e^{\lambda x} &= \sum_{i=0}^{n} a_i \frac{d^i}{dx^i} e^{\lambda x} \\ &= \sum_{i=0}^{n} a_i \lambda^i e^{\lambda x} \\ &= P(\lambda)\, e^{\lambda x}. \end{aligned}$$

□

Folgerung

Ist λ eine Nullstelle $P(\lambda) = 0$, dann ist $\varphi(x) = e^{\lambda x}$ eine Lösung von $P\left(\frac{d}{dx}\right) y = 0$.

Satz 4.7

Sei $P(T) \in \mathbb{P}(T)$ mit $P(T) = \sum_{i=0}^{n} a_i T^i$ und $a_n = 1$, $P(T)$ besitze n paarweise verschiedene Nullstellen $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$.

Dann hat die Differentialgleichung

$$P\left(\frac{d}{dx}\right)y = 0$$

ein *Fundamentalsystem von Lösungen* $\varphi_k : I \to \mathbb{C}$

$$\varphi_k(x) = e^{\lambda_k x},$$

mit $k = 1, \ldots, n$.

Mit Lemma 4.6 ist klar, daß die φ_k Lösungen sind. Als Elemente eines Fundamentalsystems müssen sie auch *linear unabhängig* sein, d.h. ihre WRONSKI-*Determinante*

$$W(x) = \begin{vmatrix} \varphi_1(x) & \varphi_2(x) & \ldots & \varphi_n(x) \\ \varphi_1'(x) & \varphi_2'(x) & \ldots & \varphi_n'(x) \\ \vdots & \vdots & & \vdots \\ \varphi_1^{(n-1)}(x) & \varphi_2^{(n-1)}(x) & \ldots & \varphi_n^{(n-1)}(x) \end{vmatrix} \neq 0.$$

Wegen

$$\varphi_k^{(j)}(x) = \lambda_k^j e^{\lambda_k x}$$

wird

$$W(0) = \begin{vmatrix} \varphi_1(0) & \varphi_2(0) & \ldots & \varphi_n(0) \\ \varphi_1'(0) & \varphi_2'(0) & \ldots & \varphi_n'(0) \\ \vdots & \vdots & & \vdots \\ \varphi_1^{(n-1)}(0) & \varphi_2^{(n-1)}(0) & \ldots & \varphi_n^{(n-1)}(0) \end{vmatrix} = \begin{vmatrix} 1 & 1 & \ldots & 1 \\ \lambda_1 & \lambda_2 & \ldots & \lambda_n \\ \vdots & \vdots & & \vdots \\ \lambda_1^{n-1} & \lambda_2^{n-1} & \ldots & \lambda_n^{n-1} \end{vmatrix},$$

zur VANDERMONDE-Determinante, und damit

$$W(0) = \prod_{p>q} (\lambda_p - \lambda_q) \neq 0,$$

da λ_p, λ_q paarweise verschieden sind. Dies bedeutet die lineare Unabhängigkeit der φ_k.

Beispiel

Für die DGL $y''' - y'' - 2y' = 0$ wird das Polynom zu $T^3 - T^2 - 2T^1 = 0$ mit den Nullstellen $\lambda_1 = 0, \lambda_2 = -1, \lambda_3 = 2$.

$$\begin{aligned} \varphi_1(x) &= 1 \\ \varphi_2(x) &= e^{-x} \\ \varphi_3(x) &= e^{2x} \end{aligned}$$

bilden das Fundamentalsystem von Lösungen; die Lösung lautet mit $\varphi(x) = \alpha\varphi_1(x) + \beta\varphi_2(x) + \gamma\varphi_3(x)$

$$y = \alpha + \beta e^{-x} + \gamma e^{2x},$$

was sich durch Probe bestätigen läßt.

Bemerkung

Die Funktionen φ_k spannen einen Lösungsraum auf, in dem jede Lösung eine Linearkombination der φ_k darstellt. Dies verschafft z.B. die Möglichkeit eines Basiswechsels, falls ein Fundamentalsystem aus komplexen Funktionen besteht und ein reelles Fundamentalsystem gewonnen werden soll.

Beispiel

Für die DGL $y'' + c^2 y = 0$ wird das Polynom zu $T^2 + c^2 = 0$ mit den Nullstellen $\lambda_1 = ic, \lambda_2 = -ic$, d.h. $\varphi_1(x) = e^{icx}, \varphi_2(x) = e^{-icx}$. Letztere sind eine komplexe Basis des Fundamentalsystems. Dieses kann z.B. transformiert werden in eine reelle Basis des Fundamentalsystems gemäß

$$\begin{array}{rclcl} \psi_1(x) & = & \frac{1}{2}(\varphi_1(x) + \varphi_2(x)) & = & \cos(cx) \\ \psi_2(x) & = & \frac{1}{2i}(\varphi_1(x) - \varphi_2(x)) & = & \sin(cx) \end{array}$$

Bemerkung

Bei mehrfachen Nullstellen ist das Polynom $P(T) \in \mathbb{P}$

$$P(T) = T^n + a_{n-1}T^{n-1} + \ldots + a_1 T + a_0$$

zerlegbar in Linearfaktoren

$$P(T) = (T - \lambda_1)^{k_1} (T - \lambda_2)^{k_2} \ldots (T - \lambda_r)^{k_r}$$

mit $\lambda_j \in \mathbb{C}$ paarweise verschieden, $k_j \geq 1$ Vielfachheit der Nullstellen und $\sum_j k_j = n$.

Lemma 4.8

Sei $\lambda \in \mathbb{C}, k \in \mathbb{N}, I \in \mathbb{R}$ k-mal differenzierbar.

Dann gilt
$$\left(\frac{d}{dx} - \lambda\right)^k \left(y(x)\, e^{\lambda x}\right) = y^{(k)}(x)\, e^{\lambda x}.$$

Lemma 4.9.

Sei $P(T) \in \mathbb{P}(T), \lambda \in \mathbb{C}, P(\lambda) \neq 0$.

Ist $g : \mathbb{R} \to \mathbb{C}$ Polynomfunktion vom Grad k, so gilt

$$P\left(\frac{d}{dx}\right)\left(g(x)\, e^{\lambda x}\right) = h(x)\, e^{\lambda x}$$

mit $h : \mathbb{R} \to \mathbb{C}$, und h ist wieder Polynom vom Grad k.

Satz 4.10

Sei $P(T) \in \mathbb{P}(T)$ mit $P(T) = \sum_{i=0}^{n} a_i T^i$ und $a_n = 1$, $P(T)$ besitze paarweise verschiedene Nullstellen $\lambda_i \in \mathbb{C}$ mit Vielfachheiten $k_j, 1 \leq j \leq r$.

Dann hat die Differentialgleichung

$$P\left(\frac{d}{dx}\right)y = 0$$

ein *Fundamentalsystem von Lösungen* («Basis») $\varphi_{jm} : I \to \mathbb{C}$ mit

$$\varphi_{jm}(x) = x^m e^{\lambda_j x}$$

mit $1 \leq j \leq r$ und $0 \leq m \leq k_j - 1$.

Prinzipielle Vorgehensweise zur Gewinnung von Lösungen

Sei $P(T) = \sum_{i=0}^{n} a_i T^i$, $a_i \in \mathbb{C}, a_n = 1$, $b : I \to \mathbb{C}$ stetig.
Betrachte die *inhomogene* DGL

$$P\left(\frac{d}{dx}\right)y = b(x).$$

Dann ist die Vorgehensweise

1. Bestimme gemäß Satz 4.10 ein Fundamentalsystem von Lösungen der *homogenen* DGL
$$P\left(\frac{d}{dx}\right)y = 0.$$
2. Führe die *inhomogene* DGL zurück auf ein System von DGL 1.Ordnung.
3. Bestimme eine Lösung der *inhomogenen* DGL durch Variation der Konstanten gemäß Satz 4.5

Das Vorgehen vereinfacht sich erheblich im Spezialfall

$$P\left(\frac{d}{dx}\right)y = e^{\mu x}, \mu \in \mathbb{C}, P(\mu) \neq 0.$$

Dann ist

$$\psi(x) = \frac{1}{P(\mu)}e^{\mu x}$$

Lösung, denn

$$P\left(\frac{d}{dx}\right)\psi(x) = P\left(\frac{d}{dx}\right)\frac{1}{P(\mu)}e^{\mu x} = \frac{1}{P(\mu)}P\left(\frac{d}{dx}\right)e^{\mu x} =_{4.6} \frac{1}{P(\mu)}P(\mu)e^{\mu x} = e^{\mu x}.$$

Satz 4.11

Sei $P(T) \in \mathbb{P}(T)$ mit $P(T) = \sum_{i=0}^{n} a_i T^i$ und $a_n = 1$, $b : \mathbb{R} \to \mathbb{C}$ mit $b(x) = f(x)e^{\mu x}$ mit f Polynom in x vom Grad m, komplex, $\mu \in \mathbb{C}$ Nullstelle k-ter Ordnung ($k \geq 0$) von $P(T)$.

Dann hat die Differentialgleichung

$$P\left(\frac{d}{dx}\right) y = f(x)\, e^{\mu x}$$

eine spezielle Lösung $\varphi : \mathbb{R} \to \mathbb{C}$ mit

$$\varphi(x) = h(x)\, e^{\mu x}$$

mit h Polynom vom Grad $m + k$.

Beweis

Nach Voraussetzung ist

$$P\left(\frac{d}{dx}\right) = Q\left(\frac{d}{dx}\right)\left(\frac{d}{dx} - \mu\right)^k$$

mit Q Polynom mit $Q(\mu) \neq 0$. Der Beweis wird durch vollständige Induktion nach dem Grad m des Polynoms $f(x)$ geführt.

Zwischenbemerkung

- gemäß Satz 4.6 gilt $P\left(\frac{d}{dx}\right) e^{\lambda x} = P(\lambda)\, e^{\lambda x}$
- gemäß Satz 4.8 gilt $\left(\frac{d}{dx} - \lambda\right)^k \left(y(x)\, e^{\lambda x}\right) = y^{(k)}(x)\, e^{\lambda x}$
- gemäß Satz 4.9 gilt $P\left(\frac{d}{dx}\right)\left(g(x)\, e^{\lambda x}\right) = h(x)\, e^{\lambda x}$ mit g, h Polynomfunktionen vom Grad k.

Induktionsanfang $m = 0$. Löse $P\left(\frac{d}{dx}\right) y = ce^{\mu x}$.

Spezielle Lösung ist

$$\varphi(x) = \frac{c}{k!Q(\mu)} x^k e^{\mu x},$$

da mit Satz 4.6 und Satz 4.8

$$\begin{aligned}
P\left(\frac{d}{dx}\right)\left(x^k e^{\mu x}\right) &= Q\left(\frac{d}{dx}\right)\left(\frac{d}{dx} - \mu\right)^k \left(x^k e^{\mu k}\right) \\
&= Q\left(\frac{d}{dx}\right)\left(k! e^{\mu x}\right) \\
&= k! Q(\mu)\, e^{\mu x}.
\end{aligned}$$

Induktionsschritt $(m-1) \to m$. Mit Satz 4.8 und Satz 4.9 gilt

$$\begin{aligned}
P\left(\frac{d}{dx}\right)\left(x^{m+k} e^{\mu x}\right) &= Q\left(\frac{d}{dx}\right)\left(\frac{d}{dx} - \mu\right)^k \left(x^{m+k} e^{\mu k}\right) \\
&= Q\left(\frac{d}{dx}\right)\left(\frac{(m+k)!}{m!} x^m e^{\mu x}\right) \\
&= g(x)\, e^{\mu x}
\end{aligned}$$

mit $g(x)$ Polynom vom Grad m. Für ein geeignetes $c \in \mathbb{C}$ ist $f_1(x) := f(x) - cg(x)$ ein Polynom vom Grad $m-1$. Nach Induktionsvoraussetzung gibt es dann ein Polynom $h_1(x)$ vom Grad $m-1+k$ mit

$$P\left(\frac{d}{dx}\right)\left(h_1(x)\, e^{\mu x}\right) = f_1(x)\, e^{\mu x}.$$

Mit $h(x) := h_1(x) + cx^{m+k}$ gilt

$$\begin{aligned} P\left(\frac{d}{dx}\right)\left(h(x)\, e^{\mu x}\right) &= \left(f_1(x) + cg(x)\right) e^{\mu x} \\ &= f(x)\, e^{\mu x}. \end{aligned}$$

□

Systeme von linearen DGL mit konstanten Koeffizienten

Die Betrachtung von Systemen von linearen DGL mit konstanten Koeffizienten ist eng verknüpft mit der Eigenwerttheorie von Matrizen. Infolgedessen besteht die explizite Bestimmung eines Fundamentalsystems von Lösungen in der Transformation der Matrix des Systems auf die Normalform.

Satz 4.12

Sei $A \in \mathbb{C}^{n\times n}, a \in \mathbb{C}^n$ Eigenvektor von A zum Eigenwert $\lambda \in \mathbb{C}$.

Die *Abbildung* $\varphi : \mathbb{R} \to \mathbb{C}^n$, definiert durch $\varphi(x) := ae^{\lambda x}$ ist *Lösung* der Differentialgleichung $y' = Ay$.

Beweis

$\varphi'(x) = a\lambda e^{\lambda x} = A\left(ae^{\lambda x}\right) = A\varphi(x)$. □

Satz 4.13

Sei $A \in \mathbb{C}^{n\times n}, a_1, \ldots, a_n \in \mathbb{C}^n$ eine Basis von Eigenvektoren von A zu den Eigenwerten $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$.

Die *Abbildungen* $\varphi_k : \mathbb{R} \to \mathbb{C}^n$, definiert durch $\varphi_k(x) := a_k e^{\lambda_k x}$ mit $k = 1, \ldots, n$ bilden ein *Fundamentalsystem von Lösungen* der Differentialgleichung $y' = Ay$.

Transformationen

Eine Matrix $A \in \mathbb{C}^{n\times n}$ besitzt genau dann eine Basis aus Eigenvektoren, wenn es eine invertierbare Matrix S gibt, sodaß die Matrix

$$B = S^{-1}AS$$

Diagonalgestalt hat. Die Spalten von S sind dann die Eigenvektoren mit den Diagonalelementen von B als den zugehörigen Eigenwerten. Allerdings ist es nicht immer möglich, eine Matrix A in dieser Weise auf Diagonalgestalt zu bringen.

In diesem Falle hilft der Umstand weiter, daß es stets eine invertierbare Matrix S gibt, sodaß die Matrix

$$B = S^{-1}AS$$

Dreiecksgestalt hat.

Lemma 4.14

Sei $A \in \mathbb{C}^{n\times n}$, $S \in GL(n, \mathbb{C})$. Dann gilt

$$\begin{array}{rcc} \varphi : \mathbb{R} \to \mathbb{C}^n & \text{Lösung der DGL} & y' = Ay \\ & \Leftrightarrow & \\ \psi := S^{-1}\varphi : \mathbb{R} \to \mathbb{C}^n & \text{Lösung der transformierten DGL} & z' = S^{-1}ASz. \end{array}$$

Beweis

$$\begin{aligned} y' = Ay \quad & \Leftrightarrow \quad & S^{-1}y' \\ & & = S^{-1}Ay \\ & & = S^{-1}ASS^{-1}y \\ & \Rightarrow_{z=S^{-1}y} & = S^{-1}ASz = z' \end{aligned}$$

□

Gemäß Lemma 4.14 läßt sich das System von linearen DGL $y' = Ay$ durch eine geeignete Transformation auf die Form $z' = Bz$ mit Dreiecksgestalt von B bringen

$$\begin{bmatrix} z_1' \\ z_2' \\ \vdots \\ z_{n-1}' \\ z_n' \end{bmatrix} = \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1(n-1)} & b_{1n} \\ & b_{22} & \dots & b_{2(n-1)} & b_{2n} \\ & & & \vdots & \vdots \\ & & & b_{(n-1)(n-1)} & b_{(n-1)n} \\ & & & & b_{nn} \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_{n-1} \\ z_n \end{bmatrix}$$

Zur vorgegebenen Anfangsbedingung $\psi(x_0) = y_0$ läßt sich durch «Rückwärtssubstitution» analog dem GAUSS-Verfahren die vollständige Lösung ermitteln:

Nach der letzten Zeile führt $z_n' = b_{nn}z_n$ zu einer Lösung der Form

$$\psi_n(x) = \alpha_n e^{b_{nn}x}$$

mit α_n so, daß $\psi_n(x_0) = \alpha_n e^{b_{nn}x_0} = y_{0n}$. Für die vorletze Zeile wird dann

$$z_{n-1}' - b_{(n-1)(n-1)}z_{n-1} = b_{(n-1)n}\psi_n(x),$$

deren Lösung mit Lemma 4.11 zu ermiteln ist. So fortsetzend gelangt man schließlich zur vollständigen Lösung.

5 Variationsrechnung (Ausblick)

Aufgabenstellung

Erinnerung

Extremstellen von $f \in \mathcal{C}^1(\Omega)$ mit Ω offene Teilmengen des $\mathbb{R}^n$.

Notwendig: x_0 kritischer Punkt von f, d.h.

$$\nabla f(x_0) = 0 \Leftrightarrow_{f \in \mathcal{C}^1(\Omega)} \frac{\partial f}{\partial a}(x) = 0, \quad \forall a \in S^{n-1} \subset \mathbb{R}^n.$$

Nun: Analoge Bedingungen für die stationären Stellen eines *Funktionals*

$$\mathfrak{F} : M \to \mathbb{R}$$

mit M Teilmenge eines Funktionenraumes. $\mathfrak{F}$ ist Integral, das auf den Funktionenraum gebildet wird.

Beispiel

Sei M die Menge aller differenzierbaren Kurven $X : I \to \mathbb{R}^3$, deren Spur auf einer Fläche $\mathfrak{S}$ in $\mathbb{R}^3$ liegt und die zwei gegebene Punkte $P, Q \in \mathfrak{S}$ verbinden. Die Frage nach der kürzesten Verbindung von P und Q unter allen Kurven $X \in M$ führt auf die notwendige Bedingung, der ein Minimierer der Bogenlänge in M genügen muß

$$L(X) = \int_I \left|\dot{X}(t)\right| dt \quad \text{für z.B.} \quad X(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix}.$$

Beispiel

Schwingungen einer Saite. Minimiere die elastische Energie

$$D(X) = \frac{1}{2} \int_I \left|\dot{X}\right|^2 dt,$$

d.h. bestimme eine Kurve $X : I \to \mathbb{R}^3$ aller $\mathcal{C}^1$-Verbindungskurven zweier Punkte $P, Q \in \mathbb{R}^3$.

Formalisierung der Aufgabenstellung

- Wähle «*erweiterten Phasenraum*» U als Gebiet in $\mathbb{R}^{2n+1} := \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n$ mit der Lagrange-Funktion $F : U \to \mathbb{R}, F \in \mathcal{C}^1$.
- Betrachte die Abbildungen $u \in \mathcal{C}^1(I, \mathbb{R}^n), I = [a, b]$, deren 1-graph in U liegt. Dabei ist
$$1 - \text{graph} u := \{(x, u(x), u'(x))\}$$
Punkt in $x \in I$. $F(x, u(x), u'(x))$ ist dann wohldefiniert und stetig auf I.

- Bilde *Funktional* (Variationsintegral)

$$\mathfrak{F}(u) := \int_a^b F(x, u(x), u'(x))\, dx$$

auf der Menge

$$\mathcal{C}_1(I, U) := \left\{ u \in \mathcal{C}^1(I, \mathbb{R}^n) \mid 1 - \text{graph} u \subset U \right\}.$$

- Wähle zwei Punkte $P, Q \in \mathbb{R}^n$ verschieden und definiere Menge

$$M := \{ u \in \mathcal{C}_1(I, U) \mid u(a) = P, u(b) = Q \}.$$

- Minimiere $\mathfrak{F}$ auf M: $\mathfrak{F}(u) \to \min \text{ in } M$.

Zur Lösung dieses Minimierungsproblems ist zu untersuchen, wie sich $\mathfrak{F}(u)$ ändert, wenn die Abbildung u variiert wird:

- Wähle Vektorfeld

$$\varphi \in \mathcal{C}^1(I, \mathbb{R}^n).$$

- Bilde Kurvenschar $h(\cdot, \varepsilon)$ mit Scharparameter $\varepsilon \in \left]-\varepsilon_0, \varepsilon_0\right[, \varepsilon_0 > 0$ und mit

$$h(x, \varepsilon) ;= u(x) + \varepsilon\varphi(x), x \in I, |\varepsilon| < \varepsilon_0.$$

Die Kurve u ist in die Kurvenschar h *eingebettet*.

- Betrachte Funktion

$$\Phi : \left]-\varepsilon_0, \varepsilon_0\right[\to \mathbb{R} \quad \text{definiert als} \quad \Phi(\varepsilon) := \mathfrak{F}(u + \varepsilon\varphi).$$

- Bilde Ableitung an Stelle $\varepsilon = 0$

$$\dot{\Phi}(0) = \frac{d}{d\varepsilon} \mathfrak{F}(u + \varepsilon\varphi)\,|_{\varepsilon=0}.$$

Definition 5.1 (*Erste Variation*)

Als *erste Variation* von $\mathfrak{F}$ an der Stelle u in Richtung φ wird der Ausdruck

$$\delta\mathfrak{F}(u, \varphi) := \dot{\Phi}(0)$$

bezeichnet.

Bemerkung

$\delta\mathfrak{F}(u, \varphi)$ ist für Variationsintegral $\mathfrak{F} : M \to \mathbb{R}$ das Analogon zur Richtungsableitung der Funktion $f : \Omega \to \mathbb{R}$

$$\frac{\partial}{\partial a} f(x_0) = \frac{d}{d\varepsilon} f(x_0 + \varepsilon a) = \nabla f(x_0)\, a.$$

Die Einschränkung $a \in S^{n-1}$ ist irrelevant, da $\nabla f(x_0)\, a$ linear in a ist.

Weiterhin gilt

$$\dot{\Phi}(\varepsilon) = \int_a^b \frac{d}{d\varepsilon} F\left(x, u(x) + \varepsilon\varphi(x), u'(x) + \varepsilon\varphi'(x)\right) dx.$$

Betrachtet man die LAGRANGE-Funktion in der Form $F(x, z, p)$ mit $x \in \mathbb{R}, z \in \mathbb{R}^n, p \in \mathbb{R}^n$, dann ergibt sich nach der Kettenregel

$$\begin{aligned} &\tfrac{d}{d\varepsilon} F\left(x, u(x) + \varepsilon\varphi(x), u'(x) + \varepsilon\varphi'(x)\right) dx \\ = \quad & F_z\left(x, u(x) + \varepsilon\varphi(x), u'(x) + \varepsilon\varphi'(x)\right) \varphi(x) \\ + \quad & F_p\left(x, u(x) + \varepsilon\varphi(x), u'(x) + \varepsilon\varphi'(x)\right) \varphi'(x). \end{aligned}$$

Damit wird

$$\delta\mathfrak{F}(u, \varphi) \quad = \quad \textstyle\int_a^b \left(F_z(x, u(x), u'(x)) \varphi(x) + F_p(x, u(x), u'(x)) \varphi'(x)\right) dx$$

bzw. in Koordinatenschreibweise

$$\delta\mathfrak{F}(u, \varphi) \quad = \quad \textstyle\int_a^b \sum_{j=1}^n \left(F_{z_j}(x, u(x), u'(x)) \varphi_j(x) + F_{p_j}(x, u(x), u'(x)) \varphi_j'(x)\right) dx$$

Lemma 5.1

Sei $F \in \mathcal{C}^1(U), F_p \in \mathcal{C}^1(u, \mathbb{R}^n)$.

Dann läßt sich die erste Variation

$$\delta\mathfrak{F}(u, \varphi), \quad \forall u \in \mathcal{C}^2(I, \mathbb{R}^n)$$

schreiben als

$$\begin{aligned} \delta\mathfrak{F}(u, \varphi) \quad = \quad & \int_a^b \sum_{j=1}^n \left(F_{z_j}(x, u(x), u'(x)) - \frac{d}{dx} F_{p_j}(x, u(x), u'(x))\right) \varphi_j(x)\, dx \\ & + \sum_{j=1}^n F_{p_j}(x, u(x), u'(x)) \varphi_j(x) \Big|_{x=a}^{x=b}. \end{aligned}$$

Definition 5.2 (*Funktionenklasse*)

Mit $\mathcal{C}_c^1(I^o, \mathbb{R}^n)$ wird die Klasse der Funktionen $\varphi \in \mathcal{C}^1(I, \mathbb{R}^n)$, die «nahe am Rand von I verschwinden», bezeichnet. Ist also $\varphi \in \mathcal{C}_c^1(I^o, \mathbb{R}^n)$ und $I = [a, b]$, so gibt es Zahlen α, β mit $a < \alpha < \beta < b$, sodaß

$$\varphi(x) = 0, \quad \forall x \notin\,]\alpha, \beta[\,.$$

Satz 5.2 (*Fundamentalsatz der Variationsrechnung*)

Sei $f \in \mathcal{C}^0(I, \mathbb{R}^n), I = [a, b]$. Dann gilt

$$\int_a^b f(x) \varphi(x)\, dx = 0 \quad \forall \varphi \in \mathcal{C}_c^1(I^o, \mathbb{R}^n) \quad \Rightarrow \quad f(x) = 0 \quad \forall x \in I.$$

Beweis

Literatur

Satz 5.3 (*System der* Euler-Lagrange-*Gleichungen auf* I)

Sei $F, F_{p_1}, \ldots, F_{p_n} \in \mathcal{C}^1(U), u \in \mathcal{C}^2(I, \mathbb{R}^n)$.

1. Dann gilt

$$\begin{aligned} & \delta\mathfrak{F}(u, \varphi) = 0, \quad \forall \varphi \in \mathcal{C}_c^1(I^o, \mathbb{R}^n) \\ \Rightarrow \quad & \tfrac{d}{dx} F_{p_j}(x, u(x), u'(x)) = F_{z_j}(x, u(x), u'(x)), \quad i \leq j \leq n. \end{aligned}$$

2. Dann gilt

$$\begin{aligned} & \delta\mathfrak{F}(u, \varphi) = 0, \quad \forall \varphi \in \mathcal{C}^1(I, \mathbb{R}^n) \\ \Rightarrow \quad & \tfrac{d}{dx} F_{p_j}(x, u(x), u'(x)) = F_{z_j}(x, u(x), u'(x)), \quad i \leq j \leq n. \end{aligned}$$

   und *zusätzlich* die sog. «natürlichen Randbedingungen» für $x = a, b$

$$F_{p_j}(x, u(x), u'(x)) = 0, \quad 1 \leq j \leq n.$$

Definition 5.3 (Euler*sches Vektorfeld.* Euler*scher Operator*)

$L_F(u) : I \to \mathbb{R}^n$ mit

$$L_F(u) := F_z(\cdot, u, u') - \frac{d}{dx} F_p(\cdot, u, u')$$

heißt Euler*sches Vektorfeld* oder Euler*scher Operator*.

Anwendung in der Physik

Variationelle Ableitung von $\mathfrak{F}$ nach u:

$$\frac{\delta\mathfrak{F}}{\delta u}(u) = L_F(u).$$

Mit Lemma 5.1 wird

$$\begin{aligned} \delta\mathfrak{F}(u, \varphi) &= \int_a^b \frac{\delta\mathfrak{F}}{\delta u}(u)\, \varphi dx \quad \forall \varphi \in \mathcal{C}_c^1(I, \mathbb{R}^n) \\ &= \left\langle \frac{\delta\mathfrak{F}}{\delta u}(u), \varphi \right\rangle \end{aligned}$$

bei Auffassung des Integrals als Skalarprodukt im Funktionenraum. Dies erklärt die Analogie zur Richtungsableitung

$$\frac{\partial}{\partial a} f(x_0) = \nabla f(x_0)\, a.$$

Abkürzungen für Euler-Lagrange-Gleichungssystem von n gewöhnlichen DGL 2.Ordnung für u :

$$L_F(u) = 0 \quad \text{bw.} \quad \tfrac{\delta\mathfrak{F}}{\delta u}(u) = 0$$

Definition 5.4 (*Extremale*)

Jede Lösung $u \in \mathcal{C}(I, \mathbb{R}^n)$ der EULER-LAGRANGE-Gleichung L_F mit

$$L_F = 0$$

heißt *Extremale von F* oder *F-Extremale*.

Definition 5.5 (LAGRANGE-*Funktion*)

$F(x, y, p)$ mit $F, F_p \in \mathcal{C}^1$ heißt *Null*-LAGRANGE*sche*, wenn jede Abbildung $u : I \to \mathbb{R}^n \in \mathcal{C}^2$ eine *F-Extremale* ist, d.h. $L_F(u) = 0$.

Satz 5.4

Sei $F(x, y, p)$ Null-LAGRANGEsche und $h \in \mathcal{C}^2$ mit $h : [a, b] \times [0, 1] \to \mathbb{R}^n$ mit festen Endpunkten P und Q, d.h.

$$h(a, \varepsilon) \equiv P, h(b, \varepsilon) \equiv Q, \quad \forall \varepsilon \in [0, 1],$$

dann gilt für das F zugeordnete Funktional

$$\mathfrak{F}(h(\cdot, \varepsilon)) \equiv \text{const.}, \quad \forall \varepsilon \in [0, 1].$$

Bemerkung

Voraussetzungen sind abschwächbar

- $F \in \mathcal{C}^1$ für $F(x, z, p)$ linear in p.
- h muß nur aus $\mathcal{C}^1$ sein.

Bemerkung

Für jede Funktion $S(x, z)$ ist die LAGRANGE-Funktion

$$G(x, z, p) := S_x(x, z) + S_z(x, z)\, p$$

eine Null-LAGRANGEsche.

Bemerkung

Umeichen (umdefinieren) eines Variationsintegrals:

Addieren einer Null-LAGRANGEschen zum Integranden ändert die Extremale nicht.

Die EULER-LAGRANGE-DGL $L_F(u) = 0$ ist notwendige Bedingung dafür, daß ein $u \in \mathcal{C}^2$ ein lokales Minimum des Funktionals

$$\mathfrak{F}(u) := \int_I F(x, u(x), u'(x))\, dx, \quad I = [a, b]$$

bei festen (wählbaren) oder freien (parametrisierbaren) Randbedingungen liefert.

Setze voraus: $F(x, z, p), F_p(x, z, p) \in \mathcal{C}^1$,

Sei $M := \{u \in \mathcal{C}_1(I, U) \mid u(a) = P, u(b) = Q\}$,

Satz 5.5

Sei $u \in M$ lokaler Minimierer (entspr. Minimum) von $\mathfrak{F}$ in M.
D.h. $\exists r > 0$, sodaß

$$\begin{aligned} &\mathfrak{F}(u) \leq \mathfrak{F}(v), \quad \forall v \in M \text{ mit } \sup_I |u - v| < r \\ \Rightarrow \quad &\delta\mathfrak{F}(u, \varphi) = 0, \quad \forall \varphi \in \mathcal{C}^1(I, \mathbb{R}^n) \text{ mit } \varphi(a) = \varphi(b) = 0. \end{aligned}$$

Ist zudem $u \in \mathcal{C}^2$, so ist u eine Extremale. Es gilt für $a < x < b$

$$F_z(x, u(x), u'(x)) - \frac{d}{dx} F_p(x, u(x), u'(x)) = 0.$$

Satz 5.6

Sei $u \in \mathcal{C}^1(I, \mathbb{R}^n)$ lokaler Minimierer von $\mathfrak{F}$ in $\mathcal{C}^1(I, \mathbb{R}^n)$.
D.h. $\exists r > 0$, sodaß

$$\begin{aligned} &\mathfrak{F}(u) \leq \mathfrak{F}(v), \quad \forall v \in \mathcal{C}^1(I, \mathbb{R}^n) \text{ mit } \sup_I |u - v| < r \\ \Rightarrow \quad &\delta\mathfrak{F}(u, \varphi) = 0, \quad \forall \varphi \in \mathcal{C}^1(I, \mathbb{R}^n). \end{aligned}$$

Ist zudem $u \in \mathcal{C}^2$ auf $]a, b[$, so gelten

$$F_z(x, u(x), u'(x)) - \frac{d}{dx} F_p(x, u(x), u'(x)) = 0$$

und zusätzlich die natürlichen Randbedingungen für $x = a, x = b$

$$F_p(x, u(x), u'(x)) = 0.$$

Hamilton-Prinzip

Betrachte Funktional $\mathfrak{A} := \int_{t_1}^{t_2} L(x(t), \dot{x}(t))\, dt$ mit Lagrange-Funktion $L(x, v) = T(x, v) - V(x)$, wobei

$$T(x, v) = \frac{1}{2} \langle v, A(x) v \rangle = \frac{1}{2} \sum_{j,k=1}^{n} a_{jk}(x) v_j v_k$$

eine positiv definite quadratische Form in $v = (v_1, \ldots, v_n)$ ist mit Koeffizienten $a_{jk}(x)$, die von $x = (x_1, \ldots, x_n)$ abhängen.

- $A(x)$ positiv definit und symmetrisch
- $T(x, v)$ kinetische Energie im Phasenraum, (x, v)-Raum
- $V(x)$ potentielle Energie
- $E(x, v) = T(x, v) + V(x)$ Gesamtenergie

Dies entspricht einem System von N Massenpunkten mit

- Bewegung der Teilchen $x(t) = (x_1(t), \ldots, x_N(t)) \in \mathbb{R}^{3N=n}$
- Geschwindigkeit der Teilchen $\dot{x}(t) = (\dot{x}_1(t), \ldots, \dot{x}_N(t)) \in \mathbb{R}^{3N=n}$

Das HAMILTON*sche Prinzip* besagt:

Dem System von N Massenpunkten ist die LAGRANGE-Funktion $L(x,v)$ zugeordnet mit

$$L(x,v) = \frac{1}{2}\langle v, A(x)v\rangle - V(x).$$

Interpretation

Unter allen denkbaren Bewegungen $x(t)$ des Systems aus $\mathcal{C}^1$ sind die *physikalisch* möglichen daduch gekennzeichnet, daß sie das «Wirkintegral»

$$\mathfrak{A}(x) = \int_{t_1}^{t_2} L(x,\dot{x})\,dt$$

stationär machen, d.h.

$$\delta\mathfrak{A}(x,\varphi) = 0$$

$\forall\varphi : [t_1,t_2] \to \mathbb{R}^n, \varphi \in \mathcal{C}^1$, die nahe an t_1, t_2 verschwinden.

Wegen $L(x,v) = \frac{1}{2}\langle v, A(x)v\rangle - V(x)$ gibt es sogar Lösungen aus $\mathcal{C}^2$, d.h. die EULER-LAGRANGE-Form

$$\frac{d}{dt}L_v(x,\dot{x}) = L_x(x,\dot{x}) \Leftrightarrow \frac{d}{dt}[A(x)\dot{x}] = -V_x(x).$$

Mit Kraftfeld $K := -\mathrm{grad}_x V = -V_x$ entspricht dies den NEWTONschen Bewegungsgleichungen für ein konservatives Kraftfeld

$$A(x) \mathrel{\hat{=}} \mathrm{diag}(m_1, \ldots, m_N)$$

mit $m_j\ddot{x} = K_j(x), j = 1, \ldots, N$.

Mit der EULER-Beziehung (Vektorfeld) ergibt sich

$$vT_v(x,v) = 2T(x,v)$$

und mit

$$L_v = T_v$$

folgt

$$\varphi(x,v) := vL_v(x,v) - L(x,v) = T(x,v) + V(x),$$

d.h. $\varphi \sim$ Gesamtenergie längs jeder Extremalen (Lösung) $x(t)$ des Funktionals $\mathfrak{A}$ konstant: Energieerhaltung.

Anwendung LEGENDRE-Transformation, die durch L erzeugt wird:

Kanonische Impulse y zu einem Punkt x

$$y := L_v(x,v) = A(x)v.$$

Mit $A(x) > 0 \Rightarrow \exists A^{-1}(x) =: B(x) \Rightarrow v = B(x)\,y$.

Die zugehörige HAMILTON-Funktion $H(x,y)$ mit x Ort und y Impuls ist

$$\begin{aligned} H(x,y) &= [yv - L(x,v)]_{v=B(x)y} \\ &= [vL_v(x,v) - L(x,v)]_{v=B(x)y} \\ &= \Phi(x,v)\,|_{v=B(x)y} \\ &= \left[\frac{1}{2}\langle v, A(x)\,v\rangle + V(x)\right]_{v=B(x)y} \\ &= \frac{1}{2}\langle y, B(x)\,y\rangle + V(x), \end{aligned}$$

wobei sich die letzte Gleichung aus A symmetrisch, dh. $A = A^\top, B = B^\top = A^{-1}$, ergibt.

Fazit der LEGENDRE-Transformation

mit $y = A(x)\,v$ und $v = B(x)\,y$ mit v Geschwindigkeit und y Impuls gilt

$$L(x,v) + H(x,y) = yv$$

mit

$$\begin{aligned} \text{LAGRANGE} - \text{Funktion} \qquad & L(x,v) = \frac{1}{2}\langle v, A(x)\,v\rangle - V(x) \\ \text{HAMILTON} - \text{Funktion} \qquad & H(x,y) = \frac{1}{2}\langle y, B(x)\,y\rangle + V(x). \end{aligned}$$

$$(t,x,v) \longleftrightarrow (t,x,y)$$

$$\begin{aligned} \dot{x} &= H_y(x,y) \\ \dot{y} &= -H_x(x,y) \end{aligned} \quad \text{oder} \begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \nabla H(x,y).$$