Tutorium Analysis II FSS 2009: Differentialrechnung im $\mathbb{R}^n$

zusammengestellt von Markus Huggenberger

Vorlesung von
Prof. Dr. H. J. Bartels
Universität Mannheim

14. Mai 2009

Agenda

Outline

1. Elementare Strukturen
   - Normierter Vektorraum
   - Konvergenz: Definition
   - Weitere grundlegende Begriffe

Strukturen im $\mathbb{R}^n$

Was ist der $\mathbb{R}^n$?

$$\mathbb{R}^n = \{(x_1, \ldots, x_n) : x_i \in \mathbb{R},\ i = 1, \ldots, n\}$$

Was brauchen wir zum „Rechnen“ im $\mathbb{R}^n$?

1. **Vektorraumstruktur**:
   - *Addition* von $x, y \in \mathbb{R}^n$: $x + y := (x_1 + y_1, \ldots, x_n + y_n)$
   - *Skalarmultiplikation* $\lambda \in \mathbb{R}$ und $c \in \mathbb{R}^n$: $\lambda \cdot x := (\lambda x_1, \ldots, \lambda x_n)$

Strukturen im $\mathbb{R}^n$

2. **Normierung**: Der Betrag bzw. die Länge von $x \in \mathbb{R}^n$ wird wie folgt definiert:

$$\|x\| = |x| = \sqrt{x_1^2 + \ldots + x_n^2}$$

Es gelten die folgenden (allgemeinen) Normeigenschaften:

1) $\|x\| \geq 0$ für alle $x \in \mathbb{R}^n$ und aus $\|x\| = 0$ folgt $x = 0$
2) $\|cx\| = |c| \, \|x\|$ für alle $x \in \mathbb{R}^n$ und $c \in \mathbb{R}$
3) $\|x + y\| \leq \|x\| + \|y\|$ für alle $x, y \in \mathbb{R}^n$

Ableitung einer Topologie

Aus dem durch diese Norm induzierten Abstandsbegriff (Den Abstand von x und y kann man als $|y-x|$ definieren.) werden die folgenden Umgebungsbegriffe abgeleitet:

- *offene Kugel*: $K(a,r) := \{x \in \mathbb{R}^n : |x-a| < r\}$
- *abgeschlossene Kugel*: $\bar{K}(a,r) := \{x \in \mathbb{R}^n : |x-a| \leq r\}$

a und r bezeichnen Mittelpunkt und Radius der Kugel.

Diese Kugelbegriffe werden im Folgenden dazu verwendet Teilmengen des $\mathbb{R}^n$ und deren Punkte zu charakterisieren.

Topologie: Innere, äußere und Randpunkte

Sei $A \subset \mathbb{R}^n$. Beliebige Punkte $a \in \mathbb{R}^n$ können mittels der zuletzt definierten Kugelumgebungen bzgl. der Teilmenge A charakterisiert werden.

Betrachte die folgende Begriffsbildung:

- a ist *innerer Punkt* von A: $\exists\, r > 0 : K(a,r) \subset A$
- a ist *äußerer Punkt* von A: $\exists\, r > 0 : K(a,r) \subset \mathbb{R}^n \setminus A$
- a ist *Randpunkt* von A: $\forall\, r > 0$ gilt
  - (i) $\exists\, x \in K(a,r) : x \in A$
  - (ii) $\exists\, y \in K(a,r) : y \in \mathbb{R}^n \setminus A$
- *Berührpunkt*: $\forall\, r > 0$ gilt: $A \cap K(a,r) \neq \emptyset$.

Topologie: Innere, äußere und Randpunkte

Es gelten:

- x ist Berührpunkt $\Leftrightarrow$ x ist Randpunkt oder x ist innerer Punkt.
- x ist Randpunkt $\Leftrightarrow$ x ist weder innerer noch äußerer Punkt.

Charakterisierung von Berührpunkten durch Konvergenz:

$a \in \mathbb{R}^n$ ist genau dann Berührpunkt von $A \subset \mathbb{R}^n$, wenn es eine Folge $x^{(m)} \in A$ gibt, so dass

$$\lim_{m \to \infty} x^{(m)} = a.$$

Topologie: Offene und abgeschlossene Mengen

Zuletzt wurde die Lage von Punkten bzgl. einer Teilmenge beschreiben, nun soll aufbauend auf diesen Begriffen, die *Teilmenge selbst* charakterisiert werden.

Eine Teilmenge des $A \subset \mathbb{R}^n$ heißt

- *abgeschlossen*, wenn A alle seine Randpunkte enthält.
- *offen*, wenn A keinen Randpunkt enthält.

Wichtige Beispiele zu offenen und abgeschlossenen Mengen:

- $\emptyset$ ist sowohl abgeschlossen als auch offen
- $\mathbb{R}^n$ ist sowohl abgeschlossen als auch offen
- $\mathbb{Q} \subset \mathbb{R}$ ist weder offen noch abgeschlossen

Topologie: Offene und abgeschlossene Mengen

Es gelten:

- A ist abgeschlossen $\Leftrightarrow \mathbb{R}^n \setminus A$ ist offen.
- A ist offen $\Leftrightarrow \mathbb{R}^n \setminus A$ ist abgeschlossen.

Menge der inneren Punkte $\mathring{A}$ ist die größte offene Menge, die A enthält. $\mathring{A}$ heißt der *offene Kern* von A.

Menge aller Berührpunkte $\bar{A}$ ist die kleinste abgeschlossene Menge, in der A enthalten ist. $\bar{A}$ heißt die *abgeschlossene Hülle* von A.

Stabilität topologischer Begriffe unter Mengenoperationen

Man kann zeigen:

- Der Durchschnitt beliebig vieler abgeschlossener Mengen ist abgeschlossen.
- Die Vereinigung endlich vieler abgeschlossener Mengen ist abgeschlossen.
- Der Durchschnitt endlich vieler offener Mengen ist offen.
- Die Vereinigung beliebig vieler offener Mengen ist offen.

Konvergenz: Definition

Definition

Eine Folge $x^{(m)}$ von Vektoren des $\mathbb{R}^n$ konvergiert gegen den Grenzwert y genau dann, wenn es zu jedem $\varepsilon > 0$ ein N gibt mit

$$\left| x^{(m)} - y \right| \leq \varepsilon \quad \text{für alle } m \geq N$$

Äquivalente Bedingungen

Äquivalente Bedingung

Sei $x^{(m)} = (x_{m1}, \ldots, x_{mn})$ eine Punktfolge im $\mathbb{R}^n$. Die Folge $\{x^{(m)}\}$ konvergiert genau dann, wenn gilt: Jede der n Zahlenfolgen $\{x_{mk}\}$ $(k = 1, \ldots, n)$ konvergiert und zwar gegen y_k.

Cauchyches Konvergenzkriterium

Die Folge $x^{(m)} \in \mathbb{R}^n$ konvergiert genau dann, wenn es zu jedem $\varepsilon > 0$ ein N gibt mit $\left|x^{(p)} - x^{(q)}\right| \leq \varepsilon$ für alle $p, q \geq N$.

Beschränktheit

Definition

Eine Menge $M \subset \mathbb{R}^n$ heißt *beschränkt* genau dann, wenn es ein $r > 0$ gibt, so dass $|x| \leq r$ für alle $x \in M$.

Dies ist äquivalent zur Bedingung, dass ein r existiert mit $M \subset \bar{K}(0, r)$.

Für beschränkte Folgen gilt der *Satz von Bolzano-Weierstraß*: Jede beschränkte Punktfolge enthält eine konvergente Teilfolge.

Kompaktheit

Definition

Eine Teilmenge $A \subset \mathbb{R}^n$ heißt *kompakt*, wenn es zu *jeder* (!!) offenen Überdeckung

$$\mathcal{U} = \{U_i, i \in I\}$$

endlich viele Indizes $i_1, \ldots, i_k \in I$ gibt, so dass gilt

$$A \subset \bigcup_{\nu=1}^{k} U_{i_\nu}.$$

Kompaktheit

Was ist nochmal eine offene Überdeckung?

Sei $A \subset \mathbb{R}^n$. Eine Familie

$$\mathcal{U} = \{U_i, i \in I\}$$

heißt *Überdeckung* von A, wenn gilt:

$$A \subset \bigcup_{i \in I} U_i.$$

Eine Überdeckung heißt „offen“, wenn alle U_i offen sind.

Kompaktheit im $\mathbb{R}^n$

Theorem (Überdeckungssatz von Heine-Borel)

Sei $A \subset \mathbb{R}^n$ beschränkt und abgeschlossen. Ist $\mathcal{U}$ (wie oben) eine offene Überdeckung von A, so wird A bereits durch endlich viele der U_i überdeckt.

Also gilt im $\mathbb{R}^n$: abgeschlossen und beschränkt $\Rightarrow$ kompakt

Man kann weiter zeigen:
nicht abgeschlossen oder nicht beschränkt $\Rightarrow$ nicht kompakt

Somit gilt:

Theorem

Eine Teilmenge des $A \subset \mathbb{R}^n$ ist genau dann kompakt, wenn A beschränkt und abgeschlossen ist.

Outline

Grundbegriffe

- *Funktion von n Veränderlichen*:
  Abbildung einer Teilmenge $D \subset \mathbb{R}^n$ nach $\mathbb{R}$ oder $\mathbb{C}$
- *Abbildungen in den $\mathbb{R}^m$*:
  Abbildung einer Teilmenge $D \subset \mathbb{R}^n$ in den $\mathbb{R}^m$
- *Graph* einer Funktion $f : \mathbb{R}^n \to \mathbb{R}$ ist die Menge

$$\{(x, y) : x \in \mathbb{R}^n, y \in \mathbb{R} \text{ mit } y = f(x)\} \subset \mathbb{R}^{n+1}$$

- *Höhenlinien bzw. Niveaulinien* helfen bei der Veranschaulichung eines Graphen. Es gilt:

$$N_z = \{x : f(x) = z\}$$

Konvergenz von Funktionen

Definition (Konvergenz von Funktionen)

Sei $D \subset \mathbb{R}^n$ und sei $f : D \longrightarrow \mathbb{R}^m$. Ist a ein Berührpunkt von D so sagen wir: Der Grenzwert

$$b = \lim_{x \longrightarrow a} f(x)$$

existiert, falls zu jedem $\varepsilon > 0$ ein $\delta > 0$ existiert mit

$$|b - f(x)| \leq \varepsilon \quad \forall\, x \in D \cap \bar{K}(a, \delta).$$

Äquivalente Bedingung Ist $f(x) = (f_1(x), \ldots, f_m(x))$ so ist obige Bedingung gleichbedeutend mit

$$\lim_{x \longrightarrow a} f_k = b_k \quad \forall\, k = 1, \ldots, m.$$

Konvergenz von Funktionen

Vorsicht bei iterierten Grenzwerten:

Es gilt **nicht** immer

$$\lim_{x\to x_1}\left(\lim_{y\to y_1} f(x,y)\right) = \lim_{y\to y_1}\left(\lim_{x\to x_1} f(x,y)\right)$$

Gegenbeispiel:

$$\lim_{x\to 0}\left(\lim_{y\to 0}\frac{x^2}{x^2+y^2}\right) = 1 \quad \text{und} \quad \lim_{y\to 0}\left(\lim_{x\to 0}\frac{x^2}{x^2+y^2}\right) = 0$$

Falls *jedoch der Grenzwert* $\lim_{(x,y)\to(x_1,y_1)} f(x,y)$ *exisitiert*, dann stimmt er mit den beiden iterierten Grenzwerten überein!

Stetigkeit: Definition

Definition (Stetigkeit)

Eine Funktion $f : D \to \mathbb{R}^m$ heißt *stetig* im Punkt $a \in D$, wenn

$$\lim_{x \to a} f(x) = f(a)$$

gilt, d.h. wenn es zu jedem $\varepsilon > 0$ ein $\delta > 0$ gibt mit

$$|f(x) - f(a)| \leq \varepsilon \qquad \forall x \in D \cap \bar{K}(a, \delta)$$

f heißt stetig, wenn f in jedem Punkt von D stetig ist.

Bemerkung Ist f durch m Koordinatenfunktionen $f_i, \dots, f_m$ gegeben, dann heißt f genau dann stetig, wenn alle f_i $i = 1, \dots, m$ stetig sind.

Stetigkeit: „Baukastenprinzip“

Summe, Produkt, Quotient stetiger Funktionen Sind f und g stetig in a, so sind auch die Funktionen $f \pm g$, fg an der Stelle a stetig. Ebenso ist $\frac{f}{g}$ in a stetig, sofern $g(a) \neq 0$.

Komposition stetiger Funktionen Sei $D \subset \mathbb{R}^n$ und seien $f : D \to \mathbb{R}^m$, $g : E \to \mathbb{R}^l$ mit $f(D) \subset E \subset \mathbb{R}^m$ weiter sei

$$h = g \circ f : D \to \mathbb{R}^l$$

die Komposition von f und g.
Ist $\lim_{x \to a} f(x) = b$ und $\lim_{y \to b} g(x) = c$ so gilt

$$\lim_{x \to a} h(x) = c.$$

Ist also f stetig an der Stelle a und g stetig an der Stelle $f(a)$, so ist auch h stetig in a.

Weitere zentrale Begriffe und Zusammenhänge

Folgende Zusammenhänge und Begriffsbildungen solltet ihr *unbedingt* im Skript nachlesen:

- Ur-/Bilder offener, abgeschlossener, kompakter Mengen unter stetigen Abbildungen?
- Maxima, Minima stetiger Funktionen auf kompakten Mengen?
- gleichmäßig stetig?
- gleichmäßig stetig auf kompakten Mengen?

Outline

Definition von Kurven

Kurven im allgemeinen sind Abbildungen der Form

$$f : D \to \mathbb{R}^n \qquad \text{mit } D \subset \mathbb{R}$$

Es geht also um Abbildungen aus dem $\mathbb{R}$ in einen höherdimensionalen Raum.

Definition

Stetige Kurven sind stetige Abbildungen von $D \subset \mathbb{R}$ in den $\mathbb{R}^n$. Dabei ist D ein eigentliches oder uneigentliches Intervall.

Eine Kurve ist genau dann stetig, wenn alle Komponentenfunktionen f_i stetig sind.

Differenzierbarkeit von Kurven

Definition

Eine Kurve heißt k*-mal (stetig) differenzierbar*, wenn alle Funktionen f_i, $i = 1, \ldots, n$ k-mal (stetig) differenzierbar sind.

Definition

$f'(t) = (f_1'(t), \ldots, f_n'(t))$ heißt Tangentialvektor an die Kurve.

Für $t_0 \in D$ heißt

$$g_{t_0}(t) = f(t_0) + f'(t_0)(t - t_0)$$

die Tangente an die Kurve f.

Singuläre vs. reguläre Kurven

Definition

Die Kurve $f : D \to \mathbb{R}^n$ heißt *glatt* oder *regulär*, wenn sie stetig differenzierbar ist und überall $f'(t) \neq 0$ gilt. Ein Parameterwert t mit $f'(t) = 0$ heißt *singulär*.

Weitere wichtige Begriffe

- Bogenlänge
- Rektifizierbar

Outline

Gegenstand

Im Gegensatz zum vorherigen Abschnitt betrachten wir nun Funktionen vom Typ:

$$f : D \to \mathbb{R} \qquad \text{mit } D \subset \mathbb{R}^n$$

Es geht also zunächst um die Differenzierbarkeit von (reellwertigen) Funktionen in n Veränderlichen.

Dabei werden grundsätzlich 2 Formen von Differenzierbarkeit unterschieden (die JEDER verstehen sollte!):

- partielle Differenzierbarkeit ($\frac{\partial f}{\partial \cdot}$)
- totale Differenzierbarkeit ($\frac{df}{d\cdot}$)

Partielle Differenzierbarkeit I

Notwendige Begriffe:

- $f : D \to \mathbb{R}, \quad D \subset \mathbb{R}^n$: reellwertige Funktion in n Veränderlichen
- $y \in D$: innerer Punkt (d.h. existiert offene Umgebung um y)
- e_k: $k - ter$ Einheitsvektor (überall Null, Eins an k-ter Stelle)

Die erste Form der Differenzierbarkeit entspricht im Wesentlichen dem, was wir bereits in Analysis 1 kennengelernt haben. Es wird nur nach einer Veränderlichen abgeleitet...

Partielle Differenzierbarkeit II

Definition

f heißt *partiell nach x_k differenzierbar*, falls der Limes

$$\lim_{h \to 0} \frac{f(y + he_k) - f(y)}{h}$$

existiert. Dieser Limes wird dann wie folgt bezeichnet

$$\frac{\partial f}{\partial x_k}(y) = f_{x_k}(y) = D_{x_k} f(y)$$

Bemerkung:
In Definition muss man sich auf solche h beschränken für die gilt

$$h \neq 0 \quad \text{und} \quad y + he_k \in D$$

Partielle Differenzierbarkeit III

Verallgemeinerung 1:

Man nennt f *partiell differenzierbar* in $y \in D$ (ohne Einschränkung auf x_k), wenn *jede* partielle Ableitung $\frac{\partial f}{\partial x_k}$, $k = 1, \ldots, n$ existiert.

Verallgemeinerung 2:

Betrachte nun Funktionen mit mehrdimensionalem Bildbereich, d.h. Funktionen der Form

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

In diesem Fall wird die partielle Ableitung in y komponentenweise definiert durch den Vektor

$$\frac{\partial f}{\partial x_k}(x) = \left(\frac{\partial f_1}{\partial x_k}, \ldots, \frac{\partial f_n}{\partial x_k} \right)$$

Totale Differenzierbarkeit

Nun wird die zweite Form der Differenzierbarkeit definiert:

Definition

Die Funktion f mit dem Definitionsbereich D heißt *total differenzierbar* in dem inneren Punkt $x \in D$, wenn es lineare Funktion $L : \mathbb{R}^n \to \mathbb{R}$ gibt, etwa

$$L(z) = \sum_{i=1}^{n} a_i z_i$$

mit

$$f(x) - f(y) = L(x - y) + R(x)$$

und

$$\lim_{x \to y} \frac{R(x)}{|x - y|} = 0$$

Zusammenhänge

Es gelten die folgenden Zusammenhänge:

i) f in y total diff'bar
$\Rightarrow$ f in y stetig
$\Rightarrow$ f in y partiell diff'bar

ii) f in y partiell diff'bar und partielle Ableitungen stetig in y
$\Rightarrow$ f in y total diff'bar

iii) f partiell diff'bar
folgt *weder stetig noch total diff'bar*

Allgemeine totale Differenzierbarkeit

Sei nun f eine Abbildung mit mehrdimensionalen Definitions- und Bildbereich, d.h. f habe die Form

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

Dann heißt f *total differenzierbar* in y genau dann, wenn eine lineare Abbildung $L : \mathbb{R}^n \to \mathbb{R}^n$ existiert, so dass auf auf einer kleinen Kugel um y gilt

$$f(x) - f(y) = L(x - y) + R(x) \qquad \text{mit } \lim_{x \to y} \frac{R(x)}{|y - x|} = 0$$

$R(x)$ ist eine Abbildung vom $R : D \to \mathbb{R}^m$.

Allgemeine totale Differenzierbarkeit

$f : D \to \mathbb{R}^m$ ist genau dann total differenzierbar in y, wenn die Funktionen $f_1, \ldots, f_m$ total differenzierbar in y sind.

Für differenzierbare Abbildungen f ergibt sich L wie folgt:

$$L(z) = \left(\frac{\partial f_i}{\partial x_j}\right)_{i=1,\ldots,m;\, j=1,\ldots,n} \cdot z$$

Ausgeschrieben bedeutet dies

$$\left(\frac{\partial f_i}{\partial x_j}\right)_{i=1,\ldots,m;\, j=1,\ldots,n} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{pmatrix}$$

Diese Matrix wird auch *Fundamentalmatrix* von f genannt.

Kettenregel

- $D \subset \mathbb{R}^l$, $E \subset \mathbb{R}^m$
- $g : D \to \mathbb{R}^m$ und $f : E \to \mathbb{R}^n$ wobei $g(D) \subset E$
- g total diff'bar in $x \in D$, f total diff'bar in $g(x) \in E$

Dann ist auch $h = f \circ g$ total diff'bar in x und es gilt

$$h'(x) = f'(g(x)) \cdot g'(x)$$

d.h. in Matrixschreibweise (Das ist also auch eine Matrixmultiplikation.)

$$\underbrace{\left(\frac{\partial h_i}{\partial x_k}\right)}_{\mathbb{R}^{m\times l}} = \underbrace{\left(\frac{\partial f_i}{\partial y_j}\right)}_{\mathbb{R}^{m\times n}} \cdot \underbrace{\left(\frac{\partial g_j}{\partial x_k}\right)}_{\mathbb{R}^{n\times l}}$$

Kettenregel: Spezialfall

Sei

1. $g : D \to \mathbb{R}^m$ eine Kurve, d.h. $l = 1$ also $D \subset \mathbb{R}$
2. $f : E \to \mathbb{R}$ eine reellwertige Funktion

dann ist

$$f \circ g : \mathbb{R} \to \mathbb{R} \qquad \text{mit} \qquad x \mapsto f(g(x))$$

und die Kettenregel lautet in diesem Fall

$$\frac{df}{dt} = \frac{d}{dt} f(g(t_0)) = \sum_{i=1}^{m} \frac{\partial f}{\partial y_i}(g_i(t_0)) \frac{dg_i}{dt}(t_0)$$

oder in Vektorschreibweise

$$\frac{df}{dt} = \frac{dg}{dt}(t_0) \cdot \operatorname{grad} f(g(t_0)) \quad (\cdot \text{ Skalarprodukt})$$

Richtungsableitung

Allgemein wird

$$e \cdot \operatorname{grad} f(x)$$

als Richtungsableitung in Richtung des Vektors e bezeichnet. Die Länge des Vektors e ist dabei auf eins zu normieren.

Aus der Cauchy-Schwarze Ungleichung folgt unmittelbar, dass

$$|e \cdot \operatorname{grad} f| \leq |\operatorname{grad} f|$$

Gleichheit gilt, falls e in Richtung des Gradienten zeigt. Also gibt der Gradient die Richtung des stärksten Wachsens bzw. Fallens der Funkion an.

Außerdem: Gradient vs. Tangentialvektor.

Outline

Vorbemerkungen

Notation Sei $i = (i_1, \ldots, i_n)$ ein n-Tupel natürlicher Zahlen. Dann sind

$$|i| := i_1 + i_2 + \ldots + i_n$$

und

$$i! := i_1! \cdot i_2! \cdot \ldots \cdot i_n!$$

Und für $x \in \mathbb{R}^n$ gilt

$$x^i = x_1^{i_1} \cdot \ldots \cdot x_n^{i_n}$$

Seien: $D \subset \mathbb{R}^n$, $f : D \to \mathbb{R}$, a innerer Punkt von D

Definition (k-mal diff'bar)

f heißt k-mal diff'bar in a, wenn f in einer ganzen Umgebung von a $k-1$-mal diff'bar ist und sämtliche partiellen Ableitungen sind in a total diffbar.

Definition (k-mal stetig diff'bar)

f heißt k-mal stetig diff'bar in a, wenn f in einer Umgebung von a k-mal diff'bar ist und alle k-ten-partiellen Ableitungen von f in a stetig sind.

Die partiellen Ableitungen werden wie folgt bezeichnet:

$$\frac{\partial^k}{\partial x_{i_k} \dots \partial x_{i_1}} = \frac{\partial}{\partial x_{i_k}} \cdots \frac{\partial}{\partial x_{i_1}} f$$

Satz/Lemma von Schwarz

Theorem

Ist f im Punkt a k-mal stetig differenzierbar, so ist

$$\frac{\partial^k f}{\partial x_{i_1} \dots \partial x_{i_k}}(a)$$

unabhängig von der Reihenfolge der x_i.

Schwächere Voraussetzungen Satz 5.4 im Skript.

Die Taylorformel

Erinnerung: Die Taylorreihe im Eindimensionalen (Ana 1)

Definition

Sei $f : I \to \mathbb{R}$, wobei $I \subset \mathbb{R}$. f unendlich oft differenzierbar und a in I. Dann heißt

$$T_f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!}(x-a)^k$$

die *Taylor-Reihe* von f im Punkt a.

Dies ist zunächst lediglich eine Definition und enthält insb. keine Aussage über Konvergenzradius oder Limes von T_f.

Taylor mehrdimensional: Vorbemerkung I

Situation:

- $U \subset \mathbb{R}^n$
- $f : U \to \mathbb{R}$ k-mal stetig diff'bar
- $x \in U$, $y \in \mathbb{R}^n$
- $x + ty \in U$ für alle $t \in [0, 1]$

Um das mehrdimensionale Problem $U \subset \mathbb{R}^n$ auf die bekannte Taylorreihe zurückzuführen betrachten wir folgende Funktion:

$$g := \begin{cases} [0,1] & \to \mathbb{R} \\ t & \mapsto g(t) := f(x+ty) \end{cases}$$

Taylor mehrdimensional: Vorbemerkung II

Frage Wie bekommen wir die $k-te$ Ableitung von g?

Erste Ableitung nach Kettenregel:

$$\begin{aligned}\frac{dg}{dt}(t) &= \frac{d}{dt}f(x_1+ty_1;x_2+ty_2;\dots;x_n+ty_n)\\ &= \sum_{i=1}^{n}\frac{\partial}{\partial x_i}f(x+ty)\frac{d}{dt}(x_i+ty_i)\\ &= \sum_{i=1}^{n}\frac{\partial}{\partial x_i}f(x+ty)y_i\end{aligned}$$

Taylor mehrdimensional: Vorbemerkung III

Was passiert, wenn wir nochmals nach t ableiten? (Wer das langweilig findet überlege sich unmittelbar, wie man von der k-ten zur $k+1$.-ten Ableitung kommt!)

$$\begin{aligned}
\frac{d^2 g}{dt^2}(t) &= \frac{d}{dt}\left[\sum_{i=1}^{n} \frac{\partial}{\partial x_i} f(x+ty) y_i\right] \\
&= \sum_{j=1}^{n} \frac{\partial}{\partial x_j}\left[\sum_{i=1}^{n} \frac{\partial}{\partial x_i} f(x+ty) y_i\right] \cdot \frac{d}{dt}(x_j + t\cdot y_j) \\
&= \sum_{j=1}^{n}\sum_{i=1}^{n} \frac{\partial}{\partial x_j}\frac{\partial}{\partial x_i} f(x+ty) y_i y_j
\end{aligned}$$

Taylor mehrdimensional: Vorbemerkung IV

Diese Doppelsumme können wir auch in ein Summenzeichen packen, nämlich:

$$\frac{d^2g}{dt^2} = \sum_{i,j=1}^{n} \frac{\partial}{\partial x_j} \frac{\partial}{\partial x_i} f(x+ty) y_i y_j$$

Jetzt kann man anfangen zu schauen welche Ableitungen gleich sind und diese zusammenfassen. Außerdem lässt sich dieses Vorgehen für höhere Ableitungen verallgemeinern. Beides findet ihr auf S. 42/43 im Skript.

Letztlich wurde in der Vorlesung per vollständiger Induktion gezeigt, dass

$$\frac{d^mg}{dt^m} = \sum_{i_1,\ldots,i_m=1}^{n} \frac{\partial}{\partial x_{i_m}} \cdots \frac{\partial}{\partial x_{i_1}} f(x+ty) y_{i_1} \cdot \ldots \cdot y_{i_m}$$

und dies lässt sich kompakter (Zusammenfassen und Notation!) schreiben als

$$\frac{d^mg}{dt^m} = \sum_{|\alpha|=m} \frac{m!}{\alpha!} D^\alpha f(x+ty) y^\alpha$$

Die Taylorformel im Mehrdimensionalen

Theorem (Taylorformel)

Sei $U \subset \mathbb{R}^n$ offen, $x \in U$ und $y \in \mathbb{R}^n$, so dass $x + ty \in U$ für alle $t \in [0,1]$ gilt. Die Funktion $f : U \to \mathbb{R}$ sei $(m+1)$-mal stetig diff'bar, dann gilt:

$$\begin{aligned} f(x+y) &= \sum_{i=0}^{m} \frac{1}{i!} \left(y_1 \frac{\partial}{\partial x_1} + \ldots + y_n \frac{\partial}{\partial x_n} \right)^i f(x) + R_m \\ &= \sum_{|\alpha| \leq m} \frac{D^\alpha f(x)}{\alpha!} y^\alpha + R_m \end{aligned}$$

Die Taylorformel im Mehrdimensionalen

Für das Restglied gilt:

$$R_m = \frac{1}{m!}\int_0^1 (1-t)^m \left(y_1 \frac{\partial}{\partial x_1} + \ldots + y_n \frac{\partial}{\partial x_n}\right)^{m+1} f(x+ty)dt$$

Eine der Vorlesung sehr ähnliche, aber ausführlichere, Darstellung der Taylorformel findet ihr in Forster - Analysis 2.

Absolute und relative Minima/Maxima

Definition (absolutes, relatives Maximum)

Die Funktion $f : D \to \mathbb{R}$, $D \subset \mathbb{R}^n$ hat an der Stelle x ein

a) *absolutes Maximum*, falls $f(y) \leq f(x)$ für alle $y \in D$ gilt.

b) *relatives Maximum*, wenn es ein $\delta > 0$ gibt, so dass $f(y) \leq f(x)$ für alle $y \in D$ mit $|y - x| \leq \delta$.

Analoge Formulierungen – natürlich mit umgekehrten Vorzeichen – gelten für absolute und relative Minima.

Stationäre Punkte

Definition (Stationärer Punkt)

Ein innerer Punkt x des Definitionsbereichs von f heißt *stationärer Punkt* von f, wenn

$$\frac{\partial f}{\partial x_i} = 0 \qquad i = 1, \dots, n$$

Es stellt sich die Frage, wann stationäre Punkte Maxima bzw. Minima sind. Diese Frage wird (teilweise) vom folgenden Satz beantwortet.

Maxima und Minima

Betrachte die quadratische Form

$$Q(y_1, \ldots, y_n) = \sum_{i,k=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_k}(x) y_i y_k$$

Theorem

Sei f an der Stelle $x \in \mathbb{R}^n$ zweimal stetig diff'bar. Verschwinden alle ersten Ableitungen in x und ist

1. *$Q < 0$ für alle y (d.h. Q ist negativ definit): relatives Maximum.*
2. *$Q > 0$ für alle y (d.h. Q ist positiv definit): relatives Minimum.*
3. *Q indefinit: kein Extremum*

Für Q semidefinit (d.h. $\leq / \geq 0$) liefert der Satz keine Entscheidung.

2 (praktischere) Fragen

Warum heißt das Ding quadratische Form? Betrachte $n = 2$. Seien

$$A = \begin{pmatrix} a & b \\ b & c \end{pmatrix} \quad \text{und} \quad v = \begin{pmatrix} x \\ y \end{pmatrix}$$

Dann gilt

$$\begin{aligned} Q(v) = Q(x,y) &= v^t \cdot A \cdot v = \\ &= \begin{pmatrix} x & y \end{pmatrix} \cdot \begin{pmatrix} a & b \\ b & c \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} \\ &= ax^2 + 2bxy + cy^2 \end{aligned}$$

2 (praktischere) Fragen

Also

$$Q(x,y) = ax^2 + 2bxy + cy^2$$

Wie entscheiden wir nun ob Q positiv oder negativ definit ist? Man kann zeigen, dass Q bzw. A genau dann

1. indefinit, wenn $ac - b^2 < 0$
2. positiv definit, wenn $a > 0$ und $ac - b^2 > 0$
3. negativ definit, wenn $a < 0$ und $ac - b^2 > 0$

Für den allgemeinen Fall: vgl. Satz 1.69 auf Seite 47 im (neuen) Skript.

Implizite Funktionen: Gegenstand (1)

Es geht um die Auflösung von Gleichungen der Form

$$F(x, y) = 0,$$

dabei ist im einfachsten Fall $D \subset \mathbb{R}^2$ und $F : D \to \mathbb{R}$, also $x \in \mathbb{R}$ und $y \in \mathbb{R}$. Ziel ist es y als Funktion von x darzustellen, etwa durch

$$y = f(x).$$

Dies ist natürlich nur dann möglich, wenn es zu jedem x genau ein y mit $F(x, y) = 0$ und $(x, y) \in D$.

f ist in dann die durch $F(x, y) = 0$ *implizit definierte Funktion*.

Implizite Funktionen: Gegenstand (2)

Einfache Beispiele – ähnlich wie im Skript – lauten

- $F(x, y) = x + y - 2 = 0$
  Dies lässt sich mehr als leicht nach y auflösen. Es gilt: $y = 2 - x$
- $F(x, y) = e^y + y + x = 0$
  Diese Gleichung ist zwar streng monoton wachsend in y lässt sich jedoch (algebraisch) nicht nach y auflösen.

Was kann man da tun? Man kann versuchen die Gleichung „lokal" nach y aufzulösen.

Implizite Funktionen: Gegenstand (3)

Einen Hinweis, wie das funktionieren kann, gibt die Kettenregel: Wir nehmen dazu an, dass $f(x) = y$ existiert. Dann können wir $F(x, y) = 0$ schreiben als

$$F(x, y) = F(x, f(x)) = 0$$

Aus der Kettenregel folgt:

$$\frac{\partial F}{\partial x}(x, f(x)) + \frac{\partial F}{\partial y}(x, f(x))f'(x) = 0 \tag{1}$$

Für $\frac{\partial F}{\partial y}(x, f(x)) \neq 0$ ist dies leicht nach $f'(x)$ auflösbar.

- Der Fall $\frac{\partial F}{\partial y}(x, y) = 0$ wird nicht weiter betrachtet.
- $\frac{\partial F}{\partial y}(x, y) \neq 0$ garantiert nicht die Lösbarkeit $F(x, y) = 0$ (vgl. bspw. $F(x, y) = x^2 + e^y$)

Sätze über implizite Funktionen

Die Frage nach (eindeutiger) Lösbarkeit des dargstellten Problems soll in einem allgemeineren Kontext betrachtet werden und zwar für Systeme von mehreren Gleichungen der Form

$$F_i(x_1, \ldots, x_m; y_1, \ldots, y_n) = 0 \qquad \text{für } i = 1, \ldots, n. \tag{2}$$

Dies sind n Gleichungen mit n Unbekannten $y_1, \ldots, y_n$. Wir setzen alle F_i als stetig differenzierbar voraus.

Wie übertragen sich die obigen Bemerkungen auf diesen allgemeineren Fall?

Sätze über implizite Funktionen

Analoge Bedingungen zu (1) ergeben sich durch Ableiten nach allen x_k $k = 1, \ldots, m$ und zwar folgt für die Ableitungen aus der Kettenregel

$$\frac{\partial F_i}{\partial x_k} + \sum_{j=1}^{n} \frac{\partial F_i}{\partial y_j} \frac{\partial f_j}{\partial x_k} = 0,$$

Mit den Abkürzungen aus der Vorlesung, nämlich

$$\frac{\partial F}{\partial x} := \left(\frac{\partial F_i}{\partial x_k}\right), \quad \frac{\partial F}{\partial y} := \left(\frac{\partial F_i}{\partial y_j}\right), \quad \frac{\partial f}{\partial x} := \left(\frac{\partial f_j}{\partial x_k}\right)$$

lassen sich die resultierenden Gleichungen in Matrixform wie folgt zusammenfassen:

$$\frac{\partial F}{\partial x} + \frac{\partial F}{\partial y} \cdot \frac{\partial f}{\partial x} = 0$$

Für $\det\left(\frac{\partial F}{\partial y}\right) \neq 0$ lässt sich diese Gleichung nach $\frac{\partial f}{\partial x}$ auflösen.

Satz über implizite Funktionen

Die letzte Beobachtung wird von Satz 1.77 aus der Vorlesung formalisiert.

Jedoch reicht die Bedingung $\frac{\partial F}{\partial y} \neq 0$ nicht aus um die lokale Existenz und Eindeutigkeit einer Lösung zu garantieren. Setzt man die Lösbarkeit von (2) in einem Punkt $\bar{x}$ und $\bar{y}$ voraus, so ist existiert ein Bereich um diesen Punkt in dem Gleichungssystem eindeutig lösbar ist. Dies beschreibt der folgende Satz:

Satz über implizite Funktionen

Theorem

Sind die Funktionen $F_i(x_1, \ldots, x_m; y_1, \ldots, y_n)$ *in einem offenen Bereich des* $\mathbb{R}^{m+n}$ r*-mal stetig differenzierbar und ist* $\overline{x} = (\overline{x_1}, \ldots, \overline{x_m}),\ \overline{y} = (\overline{y_1}, \ldots, \overline{y_n})$ *Lösung des Gleichungssystems (2) mit* $\frac{\partial(F_1,\ldots,F_n)}{\partial y_1,\ldots,y_n}(\overline{x}, \overline{y}) \neq 0$, *so existieren reelle Zahlen* $a > 0,\ b > 0$ *mit folgender Eigenschaft: Das Gleichungssystem*

$$F_i(x_1, \ldots, x_m, y_1, \ldots, y_n) = 0$$

hat für jedes x *mit* $|x - \overline{x}| \leq a$ *genau eine Lösung* $y = f(x)$ *mit* $|y - \overline{y}| \leq b$. *Die Abbildung* $f : K(\overline{x}, a) \to \mathbb{R}^n$ *ist* r*-mal stetig differenzierbar.*

Satz über die lokale Umkehrbarkeit einer Abbildung

Aus dem Satz über implizite Funktionen lässt sich lokale Existenz einer Umkehrfunktion folgern. Dies beschreibt der folgende Satz:

Theorem

Sind die Funktionen $h_i(y_1, \ldots, y_n)$ *für* $i = 1, \ldots, n$ *in einer offenen Teilmenge* $D \subseteq \mathbb{R}^n$ *stetig differenzierbar und ist dort die Funktionaldeterminante* $\frac{\partial(h_1,\ldots,h_n)}{\partial(y_1,\ldots,y_n)} \neq 0$, *so ist das Bild* $h(U)$ *jeder offenen Teilmenge* $U \subseteq D$ *offen in* $\mathbb{R}^n$, *und zu jedem* $\overline{y} \in D$ *gibt es eine offene Menge* $V \subseteq D$, $\overline{y} \in V$, *so dass* $h : V \to h(V)$ *eine stetig differenzierbare Umkehrabbildung besitzt.*

Die Umkehrabbildung ist im Allgemeinen nicht auf ganz D definiert.

Extremwerte unter Nebenbedingungen

Eine weitere zentrale Anwendung der Sätze über implizite Funktionen ist die Bestimmung von lokalen Extremwerten unter Nebenbedingungen. Dazu dient das Verfahren der Lagrangeschen Multiplikatoren, das nun skizziert werden soll:

„Schreibweise“des Optimierungsproblems

Es werden Extremwerte einer Zielfunktion

$$f : U \to \mathbb{R} \qquad \text{mit } U \subset \mathbb{R}^n \text{ offen}$$

gesucht. (Für den folgenden Satz muss f stetig differenzierbar sein.)

Es werden jedoch noch r weitere Bedingungen gestellt. Diese werden durch r stetig differenzierbare Funktionen $f_i : U \to \mathbb{R}$ und die r Gleichungen

$$f_i(x) = 0 \qquad \text{für } i = 1, \ldots, r$$

beschrieben. $M := \{x \in U : f_1(x) = \ldots = f_r(x) = 0\}$ enthält also alle Punkte von U, die den Nebenbedingungen genügen.

Der folgende Satz über Lagrangesche Multiplikatoren gibt eine notwendige Bedingung für das Vorliegen relativer Extremwerte in diesem Fall vor:

Satz über Lagrangesche Multiplikatoren

Theorem

Seien f, f_i und M wie zuletzt beschrieben, dann gilt:
Ist $a \in M$ ein Punkt, an dem die Funktion f einen relativen Extremwert annimmt, so ist entweder der Rang der Funktionalmatrix

$$\left(\frac{\partial f_i}{\partial x_k}(a)\right)$$

kleiner als r oder: es gibt Konstanten λ_i $(i = 1, \ldots, r)$ mit der Eigenschaft, dass für die Funktion

$$F = f - \sum_{i=1}^{r} \lambda_i f_i$$

die Gleichungen $\frac{\partial F}{\partial x_k}(a) = 0$ gelten $(k = 1, \ldots, n)$.