



universität
wien

DIPLOMARBEIT / DIPLOMA THESIS

Titel der Diplomarbeit / Title of the Diploma Thesis

Überblick über die Ein- und Mehrdimensionale Differenzialrechnung

verfasst von / submitted by
Stefan SCHOBER

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of
Magister der Naturwissenschaften (Mag. rer. nat.)

Wien, 2020 / Vienna, 2020

Studienkennzahl lt. Studienblatt /
degree programme code as it appears on
the student record sheet:

UA 190 406 412

Studienrichtung lt. Studienblatt /
degree programme as it appears on
the student record sheet:

Lehramtsstudium UF Mathematik UF Physik

Betreut von / Supervisor:

Ao. Univ.-Prof. Mag. Dr. Peter Raith

Danksagung

Als erstes möchte ich mich bei Herrn Ao. Univ.-Prof. Mag. Dr. Peter Raith bedanken, der mich als verständnisvoller und motivierender Betreuer bei der Entstehung der Diplomarbeit begleitet hat.

Ebenso darf ich den vielen Professoren und Professorinnen der drei Institute (Mathematik, Physik und Lehrerinnenbildung) für die vielen Eindrücke und das für meinen zukünftigen Beruf nötige Wissen danken.

Abschließend möchte ich mich bei meiner Familie bedanken, die mir mein Lehramtsstudium ermöglicht und mich dabei tatkräftig unterstützt hat.

Besonders meiner Mutter und meiner Lebensgefährtin Patricia danke ich für die vielen motivierenden Gespräche und ihr Interesse an der Diplomarbeit.

Inhaltsverzeichnis

Inhaltsverzeichnis	5
0. Grundlagen	7
1. Die Metrik	7
2. Mengen, Folgen und Grenzwerte	10
2.1. Mengen	10
2.2. Folgen und Grenzwerte	15
3. Normen und Räume	17
4. Intervallschachtelung	22
5. Matrizen	23
A. Differenzialrechnung im Eindimensionalen	27
1. Grundbegriffe	27
1.1. Stetigkeit	27
1.2. Monotonie und Beschränktheit	30
1.3. Nullstellensatz und Zwischenwertsatz	31
2. Umkehrsatz	33
3. Differenzierbarkeit im Eindimensionalen	34
4. Extremwerte	37
4.1. Lokale Extrema	37
4.2. Globale Extrema	40
5. Mittelwertsatz der Differentialrechnung	41
6. Differentiationsregeln für Eindimensionale Funktionen	44
6.1. Faktorregel	44
6.2. Summenregel	44
6.3. Produktregel	45
6.4. Quotientenregel	45
6.5. Kettenregel	46
6.6. Potenzfunktion	48
6.7. Polynomfunktionen	48
6.8. Exponentialfunktion	49
6.9. Logarithmus Funktion	49
6.10. Trigonometrische Funktionen	50
B. Differentialrechnung im Mehrdimensionalen	53
1. Grundbegriffe	54
1.1. Banachscher Fixpunktsatz	58
2. Differenzierbarkeit in $\mathbb{R}^p$	63
2.1. Differenzierbarkeit in einem Punkt	63
2.2. Partielle Ableitung	63

INHALTSVERZEICHNIS

2.3.	Mittelwertsatz (verallgem.)	67
2.4.	Differentiationsregeln für Mehrdimensionale Funktionen	69
2.5.	Richtungsableitung	72
3.	Extremwerte	74
3.1.	lokale Extrema	74
3.2.	Beispiele	79
4.	Hauptsatz über Implizite Funktionen	81
5.	Umkehrsatz	87
6.	Das Lagrange-Verfahren - Eine Methode zur Lösung von Optimierungsproblemen	90
6.1.	Modifizierte Hesse-Matrix	95
7.	Weitere Beispiele zur Differentialrechnung im $\mathbb{R}^p$	98
Abbildungsverzeichnis		101
Symbolverzeichnis		103
Literaturverzeichnis		105
Anhang		107

0. Grundlagen

« quod erat demonstrandum »

« was zu beweisen war »

Euklid von Alexandria, griechischer Mathematiker (um 300 v. Chr.)

Worum geht es?

Im ersten Abschnitt werden grundlegende Begriffe, die in den späteren Sätzen, Beweisen und Beispielen für das Verständnis nötig sind, beschrieben.

Zunächst werden Mengen, in denen Abstandsfunktionen definiert sind, die sogenannten Metrische Räume, vorgestellt. Darauf folgen typische Arten von Mengen, das Prinzip der Intervallschachtelung, sowie weiteren Definitionen von Folgen, Grenzwerten, Normen, Räumen und speziellen Matrizen.

Dieser Abschnitt orientiert sich hierbei an [LA1], [LA2], [AGB], [GM2] und [A12].

Bemerkung.

Der griechische Mathematiker Euklid von Alexandria schloss seine Beweise mit den Worten "*hóper édei déixai*". In das Lateinische übersetzt bedeutet das "*quod erat demonstrandum*", bzw. im Deutschen sinngemäß "*was zu beweisen war*". Im Folgenden werden Beweise stets mittels des Symbol $\square$, mit eben dieser Bedeutung, abgeschlossen.

1. Die Metrik

Um mit den reellen Zahlen $\mathbb{R}$ arbeiten zu können, werden der Abstandsbegriff und somit Metriken benötigt. Dabei ist die Metrik (auch Abstandsfunktion) eine Funktion, die je zwei Elementen des Raums einen nicht negativen reellen Wert zuordnet, der als Abstand der beiden Elemente voneinander aufgefasst werden kann. Im $\mathbb{R}^2$ ist dieser die *Euklidische Norm* (0.19, S. 19) und kann als Betrag des Abstandsvektors der beiden Punkte gesehen werden. Ein Raum, in dem eine Abstandsfunktion definiert ist, wird im Folgenden als metrischer Raum bezeichnet.

0. Grundlagen

Der *Abstand* zweier reeller Zahlen ergibt sich aus der Differenz von der größeren weniger der kleineren Zahl und ist stets positiv (≥ 0):

DEFINITION 0.1: Abstand

vgl. A12, S.37

Seien $a, b \in \mathbb{R}$. Der *Abstand* zwischen den Elementen a und b ist definiert durch

$$d(a, b) := \max(a, b) - \min(a, b) = \begin{cases} a - b, & \text{falls } a \geq b, \\ b - a, & \text{falls } a < b. \end{cases}$$

Die folgenden *metrischen Axiome* werden zur Definition der *Metrik* und des *metrischen Raumes* verwendet. Sie sagen aus, dass der Abstand zweier Elemente stets positiv oder null ist (*Definitheit*) und sich der Wert nicht ändert, wenn die Elemente vertauscht werden (*Symmetrie*). Verkürzen lässt er sich auch nicht, wenn der Gesamtabstand zweier Elemente als Summe der Abstände der Elemente zu einem dritten ermittelt wird (*Dreiecksungleichung*).

SATZ 0.1: Die metrischen Axiome

vgl. A12, S.37

- (M1) $d(a, b) \geq 0$, wobei $d(a, b) = 0 \Leftrightarrow a = b$ (Definitheit)
- (M2) $d(a, b) = d(b, a)$ (Symmetrie)
- (M3) $d(a, b) \leq d(a, c) + d(c, b)$ (Dreiecksungleichung)

Beweis.

Bei näherer Betrachtung der Bedeutung von (M1) und (M2) erscheinen die beiden Aussagen trivial. Für das Axiom (M3) genügt es, die Fallunterscheidungen durchzuführen (siehe nachfolgende Abbildungen)



In den Abbildungen ist zu erkennen, dass der Abstand zwischen a und b stets kleiner (oder gleich) als die Summe der Abstände von a zu c und c zu b ist. $\square$

Bemerkung.

(M1) Der Abstand zweier verschiedener Punkte ist positiv. Der Abstand eines Punktes zu sich selbst ist Null.

(M2) Der Abstand eines Punktes a zu einem Punkt b ist gleich dem Abstand von b zu a .

DEFINITION 0.2: Metrik, metrischer Raum

vgl. A12, S.37

Ist je zwei Elementen a, b einer beliebigen nichtleeren Menge A eine reelle Zahl $d(a, b)$ zugeordnet, sodass die metrischen Axiome (M1), (M2) und (M3) erfüllt sind, so heißt $d(a, b)$ die *Metrik* auf A . A sei somit ein *metrischer Raum*.

Bemerkung.

Die Elemente eines metrischen Raumes werden häufig *Punkte* genannt.

Die Zahl $d(a, b)$ heißt *Abstand* zwischen den Punkten a und b .

2. Mengen, Folgen und Grenzwerte

2.1. Mengen

Für eine Vielzahl von Sätzen und Beweisen ist es nötig, sich eine *Umgebung* um einen Punkt zu definieren. Häufig wird diese ϵ -*Umgebung* genannt, die meist ein möglichst minimaler Abstandsbereich sein soll, und häufig mit ϵ bezeichnet wird.

DEFINITION 0.3: Umgebung, ϵ -Umgebung

vgl. A12, S.166

Sei (X, d_X) ein metrischer Raum und $U_\epsilon(a) = \{x \in X : d(x, a) < \epsilon\}$. Eine Teilmenge U von X heißt *Umgebung von $a \in X$* bzw. ϵ -*Umgebung von a* , falls ein $\epsilon > 0$ existiert, sodass $U_\epsilon(a) \subseteq U$.

Oftmals ist für die Definitionsmenge einer Funktion eine offene Menge zweckmäßig, da der sog. Rand nicht für die Definition nicht sinnvoll ist.

DEFINITION 0.4: Offene Menge

vgl. LA1, S.217

Eine Teilmenge G von einem metrischen Raum X heißt *offen*, wenn jedes $\xi \in G$ eine ϵ -Umgebung besitzt, die noch ganz in G enthalten ist.

Bemerkung.

Der Rand einer Menge (0.6, S. 11) ist folglich nicht Teil der offenen Menge, da dort keine Umgebung existiert, die zur Gänze in der Menge liegt.

Beispiel.

Ist die leere Menge $\emptyset$ offen?

Gäbe es einen Punkt $\xi_0 \in \emptyset$, der eine ϵ -Umgebung $U_\epsilon(\xi_0)$ besitzt, die nicht ganz in $\emptyset$ enthalten ist, so wäre diese Menge nicht offen. Da $\emptyset$ jedoch keine Punkte besitzt, kann solch ein Punkt nicht existieren. $\diamond$

Beispiel.

Ist die Menge der reellen Zahlen $\mathbb{R}$ offen?

Beh.: Jedes offene Intervall (a, b) ist eine offene Menge:

Ist nämlich $\xi \in (a, b)$ und $\epsilon := \min(\xi - a, b - \xi)$, folglich ϵ die kleinste Zahl des Intervalls, so

2. Mengen, Folgen und Grenzwerte

liegt $U_\epsilon \in (a, b)$. Entsprechend ist ersichtlich, dass alle Intervalle der Form $(-\infty, a)$, $(a, +\infty)$ offen sind. $\diamond$

Der *Rand* einer Menge sind jene Elemente der Menge die wegfallen, wenn aus ihr eine offene Menge gebildet wird. Daher kann mit der *Differenzmenge* aus der Menge und ihrer offenen Menge der *Rand* gebildet werden.

DEFINITION 0.5: Differenzmenge

vgl. LA1, S.20

Die *Differenzmenge* $M \setminus N$ ist die Menge aller Elemente von M , die nicht in N liegen.

DEFINITION 0.6: Randpunkt; Rand

vgl. GM2, S.14

Sei X ein metrischer Raum und $Y \subset X$ eine Teilmenge und $x \in X$. Der Punkt x heißt *Randpunkt* von Y , wenn in jeder Umgebung von x sowohl ein Punkt von Y als auch ein Punkt von $X \setminus Y$ liegt. Die Menge aller Randpunkte von Y heißt *Rand*.

Weitere nützliche Begriffe sind beschränkte, abgeschlossene und kompakte Menge.

DEFINITION 0.7: Beschränkte Menge

vgl. LA1, S.70

Eine Menge $M \subseteq \mathbb{R}$ heißt *nach unten beschränkt*, wenn es ein $a \in \mathbb{R}$ gibt, so dass $a \leq x$ für alle $x \in M$ ist. a heißt dann *untere Schranke* von M .

Hingegen heißt M *nach oben beschränkt*, wenn es ein $b \in \mathbb{R}$ gibt, so dass $b \geq x$ für alle $x \in M$ ist. b heißt dann *obere Schranke* von M .

Eine Menge M heißt *beschränkt*, wenn sie nach unten und nach oben beschränkt ist.

Beispiel.

Die Intervalle $[0, 1]$, $[-3, 3]$ und $[-\pi, \pi]$ sind beschränkt.

0. Grundlagen

DEFINITION 0.8: Beschränkte Mengen in metrischen Räumen

vgl. LA1, S.85

Eine Teilmenge M eines metrischen Raumes X heißt *beschränkt*, wenn es ein $R > 0$ und ein $a \in X$ gibt, sodass $M \subseteq U_R(a)$ gilt.

Bemerkung.

Für Teilmengen von $\mathbb{R}$ stimmen die Definitionen (0.7, S. 11) und (0.8, S. 12) überein.

DEFINITION 0.9: Abgeschlossene Menge

vgl. A12, S.166

Eine Teilmenge G eines metrischen Raumes X heißt *abgeschlossen*, wenn $X \setminus G$, die Differenzmenge von G , offen ist. Zudem ist G genau dann *abgeschlossen*, wenn der Grenzwert jeder konvergenten Folge in G wieder in G liegt.

Beispiel.

Die leere Menge $\emptyset$, die reelle Zahlenmenge $\mathbb{R}$, sowie zum Beispiel das Intervall $[0, 1]$ sind abgeschlossen.

Als nächstes wird die *kompakte Mengen* definiert, allerdings wird diese Eigenschaft in diesem Zusammenhang üblicherweise *folgenkompakt* genannt. Für metrische Räume kann bewiesen werden, dass das äquivalent zur üblichen Definition von kompakt ist:

DEFINITION 0.10: Kompakte Menge

vgl. LA1, S.225

Eine Menge M heißt *kompakt*, wenn jede Folge aus M eine konvergente Teilfolge besitzt, deren Grenzwert wieder in M liegt.

Für den Beweis des anschließenden Satzes, der Auskunft darüber gibt, wann eine Menge kompakt ist (0.3, S. 13), ist nun folgender fundamentaler Satz für die Beschränktheit einer Folge nötig:

SATZ 0.2: Satz von Bolzano-Weierstraß

vgl. LA1, S.156

Jede beschränkte Folge (a_n) enthält eine konvergente Teilfolge.



Abbildung 0.1.: *Bernhard Bolzano, tschechischer Mathematiker (1781 - 1848)*

Quelle (abgerufen am 22.08.2019): de.wikipedia.org/wiki/Bernhard_Bolzano



Abbildung 0.2.: *Karl Weierstraß, deutscher Mathematiker (1815 - 1897)*

Quelle (abgerufen am 22.08.2019): de.wikipedia.org/wiki/Karl_Weierstraß

SATZ 0.3: Wann ist eine Menge kompakt?

vgl. LA1, S.225

Eine Teilmenge $M \subseteq \mathbb{R}^n$ ist genau dann kompakt, wenn sie beschränkt und abgeschlossen ist.

Beispiel.

Das Intervall $[0, 1]$ ist abgeschlossen und beschränkt und somit *kompakt*.

Bemerkung.

Halboffene und Offene Intervalle der Formen $[a, b[$, $]a, b]$, $]a, b[$ sind nicht abgeschlossen und daher nicht kompakt.

0. Grundlagen

Die reellen Zahlen sind zwar abgeschlossen, aber nicht beschränkt und daher auch nicht kompakt.

Beweis.

Sei M eine kompakte Menge. Wäre sie unbeschränkt, so würde sie eine Folge (x_n) mit $|x_n| > n$ enthalten und jede Teilfolge derselben wäre unbeschränkt und somit divergent. WID!

Die Menge M ist abgeschlossen, denn wenn die Folge (x_n) aus M gegen x strebt, strebt auch eine Teilfolge von (x_n) gegen ein Element aus M , aber auch, wie jede Teilfolge von (x_n) , gegen x . Folglich liegt x in M .

Sei nun umgekehrt M eine abgeschlossene und beschränkte Menge, dann ist jede Folge in M ebenfalls beschränkt und besitzt nach dem Satz von Bolzano-Weierstraß (0.2, S. 13) eine konvergente Teilfolge. Weil M abgeschlossen ist, gehört der Grenzwert der Folge zu M . Somit ist M kompakt. $\square$

2.2. Folgen und Grenzwerte

Für den Grenzwert einer Funktion wird zunächst der *Häufungspunkt einer Menge* definiert, jedoch ist dazu der Begriff der *konvergenten Folge* nötig.

DEFINITION 0.11: Konvergenz einer Folge

vgl. LA1, S.144

Eine Zahlenfolge (a_n) *konvergiert* gegen a , wenn es zu jeder positiven Zahl ε einen Index $n_0(\varepsilon)$ gibt, sodass

$$\text{für alle } n > n_0(\varepsilon) \text{ stets } d(a_n, a) < \varepsilon$$

ist. a heißt *Grenzwert der Folge* (a_n) und wird wie folgt ausgedrückt:

$$\lim_{n \rightarrow \infty} a_n = a$$

Beispiel. Nullfolge

Für gegebenes $\epsilon > 0$ gibt es ein n_0 mit $\frac{1}{n_0} < \epsilon$. Sei jetzt $n \geq n_0$, dann ist $|\frac{1}{n} - 0| \leq \frac{1}{n_0} < \epsilon$. Deshalb konvergiert die Folge gegen 0. (Solche Folgen werden *Nullfolgen* genannt.)

DEFINITION 0.12: Häufungspunkt einer Menge

vgl. LA1, S.234

Sei (x_n) aus M eine Teilmenge eines metrischen Raum X eine gegen ξ konvergente Folge, deren Glieder alle $\neq \xi$ sind. So heißt ξ *Häufungspunkt* der Menge M .

Bemerkung.

Der *Häufungspunkt einer Menge* sollte nicht mit dem *Häufungspunkt einer Folge* verwechselt werden. Obwohl die beiden Begriffe eng verwoben sind, haben sie dennoch andere Bedeutungen. So versteht sich unter ersterem ein Punkt, der von unendlich vielen Punkten der Menge umgeben ist. Ein Häufungspunkt einer Folge (auch *Häufungswert*) ist hingegen ein Punkt, der der Grenzwert einer unendlichen Teilfolge der Folge ist.

DEFINITION 0.13: Grenzwert einer Funktion

vgl. LA1, S.235

Sei f eine auf X definierte Funktion und ξ ein Häufungspunkt von X . Die Funktion f besitzt den Grenzwert η für $x \rightarrow \xi$, wenn für jede Folge (x_n) aus X , die gegen ξ strebt und deren Glieder alle $\neq \xi$ sind, stets $f(x_n) \rightarrow \eta$ konvergiert.

$$f(x) \rightarrow \eta \text{ für } x \rightarrow \xi \text{ bzw. } \lim_{x \rightarrow \xi} f(x) = \eta$$

Eine unverzichtbare Art von Folgen sind die *Cauchy-Folgen*, die eine zentrale Rolle bei Konvergenzeigenschaften spielen:

DEFINITION 0.14: Cauchyfolge

vgl. LA1, S.156

Die Folge (a_n) heißt *Cauchyfolge*, wenn es zu jedem $\epsilon > 0$ einen Index $n_0 = n_0(\epsilon)$ gibt, so dass für alle $m, n > n_0$ stets $|a_m - a_n| < \epsilon$ bleibt.

Bemerkung.

Hierbei werden im Verlauf der Folge die Abstände der Folgenglieder beliebig klein.

3. Normen und Räume

Zuvor wurde der sog. *metrische Raum* vorgestellt (0.2, S. 9), der als Menge, in der eine Abstandsfunktion (Metrik) definiert ist, verstanden wird.

Sind zusätzlich die Summe und das Vielfache festgelegt und die nachfolgend definierten *Vektorraumaxiome* gelten, so handelt es sich um einen *linearen Raum* bzw. *Vektorraum*.

DEFINITION 0.15: Linearer Raum, Vektorraum

vgl. LA1, S.118

Sei E eine beliebige nichtleere Menge. Diese Menge heißt *linearer Raum (Vektorraum)*, wenn für zwei Elemente a, b aus E und jede Zahl α (Skalar) eine Summe $a + b$ und ein Vielfaches αa so definiert sind, dass $a + b$ und αa wieder in E liegen und folgende *Vektorraumaxiome* erfüllt sind:

- (i) **Assoziativgesetz:** $a + (b + c) = (a + b) + c$
- (ii) **Kommutativgesetz:** $a + b = b + a$
- (iii) **Neutrales Element der Addition:** $\exists 0 \in E$ mit $a + 0 = a \quad \forall a \in E$
- (iv) **Additives Inverses:** $\forall a \in E \exists -a \in E$, sodass $a + (-a) = 0$
- (v) **Vektoriell distributivgesetz:** $\alpha(a + b) = \alpha a + \alpha b$
- (vi) **Skalares distributivgesetz:** $(\alpha + \beta)a = \alpha a + \beta a$
- (vii) **Skalares Assoziativgesetz:** $(\alpha\beta)a = \alpha(\beta a)$
- (viii) $1a = a$

Beispiel.

Die zweidimensionale euklidische Ebene ($\mathbb{R}^2$) und der Raum ($\mathbb{R}^3$) sind Vektorräume, hingegen sind die natürlichen Zahlen $\mathbb{N}$ kein Vektorraum, da in ihnen kein additives Inverses Element existiert.

0. Grundlagen

Um in der mehrdimensionalen Analysis besser mit Funktionen arbeiten zu können, werden Normen benötigt. Diese sind Abbildungen, die einem Element eines Vektorraums eine positive reelle Zahl zuordnen und dabei zusätzlich bestimmte Eigenschaften aufweisen müssen:

DEFINITION 0.16: Norm

vgl. AGB, S.452; vgl. LA2, S.11

Sei X ein Vektorraum über $\mathbb{R}$ mit Elementen $f, g, \dots$. Eine Abbildung $\|\cdot\| : X \rightarrow [0; \infty)$ heißt *Norm* auf X , wenn die folgenden drei Bedingungen (Normaxiome) gelten:

1. **Positivität:** $\|f\| \geq 0$ und $\|f\| = 0 \Leftrightarrow f = 0$
2. **Homogenität:** $\|\alpha f\| = |\alpha| \|f\|$
3. **Dreiecksungl.:** $\|f + g\| \leq \|f\| + \|g\|$

Konkrete Normen sind zum Beispiel Vektornormen wie die *p-Normen*, die für reelle Zahlen $p \geq 1$ definiert sind. Dazu zählen unter anderem die *Summennorm* mit $p = 1$, die *euklidische Norm* mit $p = 2$ und die *Maximumsnorm* mit $p \rightarrow \infty$:

DEFINITION 0.17: p-Normen

vgl. LA2, S.12

Sei $X \in \mathbb{R}^n$ und $x \in X$ ein Vektor.

So wird $\|x\|_p = (\sum_{i=1}^n ((x_i)^p)^{\frac{1}{p}}$ als *p-Norm* bezeichnet.

DEFINITION 0.18: 1-Norm; Summennorm

vgl. AGB, S.452

Sei $X \in \mathbb{R}^n$ und $x \in X$ ein Vektor.

So wird $\|x\|_1 := \sum_{i=1}^n |x_i|$ als *Summennorm* bezeichnet.

DEFINITION 0.19: 2-Norm; Euklidische Norm

vgl. AGB, S.295

Sei $X \in \mathbb{R}^n$ und $x \in X$ ein Vektor.

So wird $\|x\| = \sqrt{\sum_{i=1}^n (x_i)^2}$ als *euklidische Norm* bezeichnet.

Bemerkung.

Die *Euklidische Norm*, oder *2-Norm*, kann im zwei- und dreidimensionalen euklidischen Raum als Länge von Vektoren verstanden werden. Zur Berechnung wird der Satz des Pythagoras verwendet (siehe Beispiel).

Beispiel.

Die Länge des Vektors (x, y) berechnet sich durch $\|(x, y)\| = \sqrt{x^2 + y^2}$.

DEFINITION 0.20: Maximumsnorm

vgl. LA2, S.11

Sei $X \in \mathbb{R}^n$ und $x \in X$ ein Vektor.

So wird $\|x\|_\infty := \max(|x_1|, \dots, |x_n|)$ als *Maximumsnorm* bezeichnet.

Beispiel.

$$\|(1, -2, 0, -5)\|_{(max)} = \max(|1|, |-2|, |0|, |-5|) = 5$$

Besitzt der zuvor vorgestellte lineare Raum (0.15, S. 17) eine ihm zugeordnete Norm, so heißt dieser *Normierter Raum*. Als Beispiele für Normierte Räume sollen der *Normierte Raum* $\mathbb{R}^p$, der *Euklidische Raum* und der *Banachraum* dienen:

DEFINITION 0.21: Normierter Raum

vgl. LA2, S.11

Ein linearer Raum E mit Elementen $f, g, \dots$ heißt *normierter Raum*, wenn jedem $f \in E$ eine reelle Zahl $\|f\|$ (die *Norm von f*), so zugeordnet ist, sodass die Normaxiome (0.16, S.18) gelten.

0. Grundlagen

Bemerkung.

Ein *normierter Raum* ist auch ein *metrischer Raum*, die Umkehrung gilt nicht!
(Eine Norm kann eine Metrik induzieren.)

DEFINITION 0.22: Der Normierter Raum $\mathbb{R}^p$

vgl. LA2, S.12

Sei $r \geq 1$ eine reelle Zahl und $x := (x_1, \dots, x_p)$.

Dann ist $\mathbb{R}^p$ mit $\|x\| = (\sum_{k=1}^p |x_k|^r)^{\frac{1}{r}}$ ein normierter Raum.

DEFINITION 0.23: Skalarprodukt / Inneres Produkt

vgl. LAG, S.274

Seien x, y zwei Elemente des $\mathbb{R}^p$ mit $x = (x_1, \dots, x_p)$ und $y = (y_1, \dots, y_p)$.
Das (*kanonische*) *Skalarprodukt* ist definiert durch

$$\langle x, y \rangle = x_1 y_1 + \dots + x_p y_p$$

wobei $\langle \cdot, \cdot \rangle : \mathbb{R}^p \times \mathbb{R}^p \rightarrow \mathbb{R}$ mit $(x, y) \mapsto \langle x, y \rangle$ ist.

DEFINITION 0.24: Euklidischer Raum/ reeller Skalarproduktraum

vgl. LAG, S.274

Ein reeller Vektorraum E , in dem das *Skalarprodukt* (*inneres Produkt*) (A.9, S. 37) definiert ist, heißt *euklidischer Raum*.

Beispiel.

Der Vektorraum der reellen Zahlen $\mathbb{R}$ bildet mit dem Skalarprodukt $\langle x, y \rangle = xy$ einen euklidischen Raum.

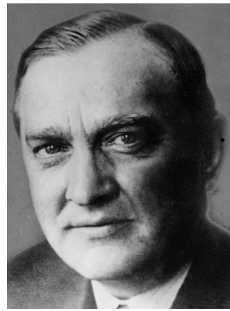


Abbildung 0.3.: *Stefan Banach, polnischer Mathematiker (1892 - 1945)*

Quelle (abgerufen am 22.08.2019): de.wikipedia.org/wiki/Stefan_Banach



Abbildung 0.4.: *Augustin-Louis Cauchy, französischer Mathematiker (1789 - 1857)*

Quelle (abgerufen am 22.08.2019): de.wikipedia.org/wiki/Augustin-Louis_Cauchy

Für die Definition des Banachraumes wird der Begriff der Cauchyfolge (0.14, S. 16) benötigt:

DEFINITION 0.25: Banachraum

vgl. LA2, S.16

Ein normierter Raum E heißt *vollständig* oder *Banachraum*, wenn jede Cauchyfolge aus E gegen ein Element aus E konvergiert.

4. Intervallschachtelung

Ein bekanntes Hilfsmittel für Beweise ist das nun folgende Prinzip der Intervallschachtelung:

DEFINITION 0.26: Intervallschachtelung

vgl. LA1, S.157

Eine Folge abgeschlossener Intervalle $I_n := [a_n, b_n]$ heißt *Intervallschachtelung*, wenn $I_1 \supset I_2 \supset I_3 \supset \dots$ ist und weiters die Folge der Intervalllängen $b_n - a_n$ gegen 0 strebt.

SATZ 0.4: Prinzip der Intervallschachtelung

vgl. LA1, S.158

Für jede reelle Intervallschachtelung gibt es ein eindeutig bestimmtes $a \in \mathbb{R}$ mit $a \in [a_n, b_n]$ für alle $n \in \mathbb{N}$.

Beweis.

Da die beiden Folgen (a_n) und (b_n) monoton und beschränkt sind, gilt $a := \lim a_n$ und $b := \lim b_n$. Da weiters $b_n - a_n$ gegen $b - a$ bzw. 0 strebt, muss $a = b$ sein. Und da für alle n nun $a_n \leq \sup a_k = a = b = \inf b_k \leq b_n$ ist, liegt a in jedem Intervall $[a_n, b_n]$. Ist dies auch für die Zahl c der Fall und $a_n \leq c \leq b_n$ für alle n , so muss auch $a \leq c \leq b$ und daher $c = a$ sein. $\square$

5. Matrizen

Matrizen finden sich in vielen Teilgebieten der Mathematik, so auch in der ein- und mehrdimensionalen Analysis. Bekannte Beispiele hierfür sind die später folgende *Jacobi-Matrix* (B.2, S. 64) und die *Hesse-Matrix* (B.15, S. 75).

Darüber hinaus werden für einige der später folgenden Beweise die Begriffe *Einheitsmatrix* und *Abbildungsnorm einer Matrix*, sowie die *Definitheit einer Matrix* benötigt.

DEFINITION 0.27: Einheitsmatrix

vgl. LA2, S.54

Die (q, q) -Einheitsmatrix ist eine Matrix mit q Spalten und q Zeilen, folglich eine quadratische Matrix, deren Hauptdiagonalelemente 1 und sonst 0 aufweist:

$$I := \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & 0 & \ddots & \vdots \\ 0 & \dots & 0 & 1 \end{pmatrix}$$

DEFINITION 0.28: Unterdeterminante

vgl. LA2, S.309ff

Werden aus einer (q, p) -Matrix

$$A := \begin{pmatrix} a_{11} & \dots & a_{1p} \\ \vdots & & \\ a_{q1} & \dots & a_{qp} \end{pmatrix}$$

$r < q$ Zeilen und $s < p$ Spalten gestrichen, wobei r und s so gewählt sind, dass die verbleibende Matrix $\mathbf{B}$ quadratisch ist ($q-r = q-s$), so wird $\det \mathbf{B}$ als $(q-r)$ -reihige *Unterdeterminante* von A bezeichnet.

Beispiel. Zweireihige Unterdeterminante

$\begin{vmatrix} 1 & 3 \\ 0 & 5 \end{vmatrix}$ ist eine zweireihige Unterdeterminante von $\begin{pmatrix} 1 & 2 & 3 & 3 \\ 7 & 6 & 5 & 4 \\ 0 & 8 & 4 & 5 \end{pmatrix}$, die durch Streichung der zweiten Zeile und zweiter und dritter Spalte entsteht.

DEFINITION 0.29: Abbildungsnorm einer Matrix

vgl. LA2, S.55ff

Sei $(a_{j,k})$ eine Matrix aus der Menge aller q, p -Matrizen $\mathfrak{M}(q, p)$, $\mathbb{R}^p$ und $\mathbb{R}^q$ mit beliebigen (festen) Normen ausgestattet. Es gelte $A : \mathbb{R}^p \rightarrow \mathbb{R}^q$ und es sei $(a_{j,k}) = [A]$ die Darstellungsmatrix von A . Die Abbildung A besitze eine Norm $\|A\|$, sodass die *Abbildungsnorm* der Matrix definiert ist durch: $\|A\| := \sup \|Ax\|$ für alle x mit $\|x\| = 1$

Beispiel. Abbildungsnorm

vgl. LA2, S.58, Aufgabe + 4a

Im folgenden sei $A \in \mathfrak{S}(\mathbb{R}^p)$, $p \geq 2$.

Zeige: Wird $\mathbb{R}^p$ mit der l^∞ - Norm versehen, so ist $\|A\| = |[A]|_\infty$

Sei $[A] = (a_{jk})$ und $x := \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix}$, dann ist

$$\|Ax\|_\infty = \max_{j=1}^p \left| \sum_{k=1}^p a_{jk} x_k \right| \leq \left(\max_{j=1}^p \sum_{k=1}^p |a_{jk}| \right) \|x\|_\infty$$

somit gilt

$$\|A\| \leq \max_{j=1}^p \sum_{k=1}^p |a_{jk}| = |[A]|_\infty.$$

Es gibt einen Index m mit $|[A]|_\infty = \sum_{k=1}^p |a_{mk}|$.

Werden die Komponenten $z_1, \dots, z_p$ des Vektors z durch

$$z_k := \begin{cases} 1, & \text{falls } a_{mk} \geq 0, \\ -1, & \text{falls } a_{mk} < 0 \end{cases}$$

bezeichnet, so ist $\|z\|_\infty = 1$ und $\sum_{k=1}^p a_{mk} z_k = \sum_{k=1}^p |a_{mk}| = |[A]|_\infty$. Daher gilt

$$\|A\| = \sup_{x \neq 0} \frac{\|Ax\|_\infty}{\|x\|_\infty} \geq \frac{\|Az\|_\infty}{\|z\|_\infty} \geq |[A]|_\infty.$$

Daraus folgt mit der obigen Abschätzung von $\|A\|$ die Behauptung. $\diamond$

Bei späteren Untersuchungen von Matrizen wird die Invertierbarkeit einer Matrix nötig sein:

DEFINITION 0.30: Inverse Matrix

vgl. LA2, S.306

Sei A eine quadratische (p, p) - Matrix. Die Matrix A^{-1} heißt *Inverse Matrix* von A , wenn gilt:

$$AA^{-1} = A^{-1}A = I$$

Bemerkung.

Die Gleichung in obiger Definition gilt nur mit Matrizenmultiplikation.
 I ist die Einheitsmatrix.

Um einfach überprüfen zu können, ob eine Matrix eine Inverse besitzt ist folgendes Kriterium hilfreich:

SATZ 0.5: Invertierbarkeitskriterium

vgl. LA2, S.306

Eine quadratische Matrix ist genau dann invertierbar, wenn ihre Determinante nicht verschwindet ($\neq 0$ ist).

A. Differenzialrechnung im Eindimensionalen

« Von allen, die bis jetzt nach Wahrheit forschten, haben die Mathematiker allein eine Anzahl Beweise finden können, woraus folgt, dass ihr Gegenstand der allerleichteste gewesen sein müsse. »

René Descartes, französischer Mathematiker (1596-1650)

Worum geht es?

Im Kapitel **A** wird die Differentialrechnung im Eindimensionalen veranschaulicht. Dabei werden Begriffe wie *Stetigkeit*, *Monotonie* und einige besondere Punkte von Funktionen erläutert, die nötig sind, um diese untersuchen und beschreiben zu können. Ebenso werden die Voraussetzungen für differenzierbare Funktionen gezeigt. Abschließend werden einige *Differentiationsregeln*, die für den praktischen Gebrauch sinnvoll sind, beschrieben.

1. Grundbegriffe

Zunächst werden die Begriffe *Stetigkeit*, *Differenzierbarkeit* und die wesentlichen Sätze der Differentialrechnung vorgestellt.

1.1. Stetigkeit

Eine grundlegende Eigenschaft von Funktionen in metrischen Räumen (0.2, S. 9) ist die *Stetigkeit*. Dazu werden außerdem die verschiedenen Stetigkeitsdefinitionen betrachtet und bewiesen. Das *Folgenkriterium* soll dabei als praktisches Hilfsmittel zur Überprüfung auf Stetigkeit vorgestellt werden.

Dieser Abschnitt orientiert sich hierbei an [LA1] und [A12].

DEFINITION A.1: $\epsilon\delta$ -Definition der Stetigkeit in $\mathbb{R}^p$

vgl. GM2, S.27ff

Eine Funktion $f : X \rightarrow Y$ mit $X \subseteq \mathbb{R}$, $Y \subseteq \mathbb{R}^p$ heißt *stetig* in $\xi \in \mathbb{R}$, genau dann wenn es für jedes $\epsilon > 0$ ein $\delta_\epsilon > 0$ gibt, so dass $\|f(\xi) - f(x)\| < \epsilon$ für alle $x \in X$ mit $\|\xi - x\| < \delta_\epsilon$.

Ist die Funktion f stetig in allen Punkten $\xi \in X$, so heißt f *stetig*.

A. Differenzialrechnung im Eindimensionalen

Bemerkung.

Alternativ kann der Satz auch mit folgenden Metriken definiert werden:

$$d_Y(f(x), f(\xi)) < \epsilon \text{ für alle } x, \xi \in X \text{ mit } d_X(x, \xi) < \delta$$

Eine alternative Formulierung der obigen Definition anhand von Umgebungen liefert der nun folgende Satz über die *Umgebungs-Definition der Stetigkeit*:

DEFINITION A.2: Umgebungs-Definition der Stetigkeit

vgl. LA2, S.32

Eine Funktion $f : X \rightarrow \mathbb{R}^p$ mit $X \subseteq \mathbb{R}$ heißt *stetig* in $\xi \in \mathbb{R}$, genau dann wenn es zu jeder ϵ -Umgebung U_ϵ von $f(\xi)$ stets eine δ -Umgebung V von x gibt, sodass $f(V \cap X) \subset U_\epsilon$ ist.

SATZ A.1: Folgenkriterium der Stetigkeit

vgl. A12, S.81

Sei $f : X \rightarrow \mathbb{R}^q$ mit $X \subseteq \mathbb{R}^p$. Die Abbildung f ist genau dann in ξ stetig, wenn gilt:
Für jede gegen ξ konvergente Folge $(x_n)_{n \in \mathbb{N}}$ in X konvergiert die Bildfolge $(f(x_n))_{n \in \mathbb{N}}$ in Y gegen $f(\xi)$.

Beweis.

Seien f stetig in $\xi \in X$ und $(x_n)_{n \in \mathbb{N}}$ eine Folge in X mit $\xi = \lim_{n \rightarrow \infty} x_n$.

Zu $\epsilon > 0$ gibt es dann ein $\delta > 0$, sodass $d_Y(f(x), f(\xi)) < \epsilon$ für alle $x \in X$ mit $d_X(x, \xi) < \delta$ gilt. Weiter existiert zu δ ein $N \in \mathbb{N}$ mit $d_X(x_n, \xi) < \delta$ für alle $n \geq N$.

Damit folgt $d_Y(f(x_n), f(\xi)) < \epsilon$ für alle $n \geq N$, d.h. $\lim_{n \rightarrow \infty} f(x_n) = f(\xi)$.

Für die umgekehrte Implikation sei nun vorausgesetzt, dass f nicht stetig in ξ ist.

Das heißt, es gibt ein $\epsilon > 0$, sodass für jedes $\delta > 0$ ein Punkt $x \in X$ existiert mit sowohl $d_X(x, \xi) < \delta$ als auch $d_Y(f(x), f(\xi)) \geq \epsilon$. Insbesondere gibt es dann zu jedem $n \in \mathbb{N}$ einen Punkt $x_n \in X$ mit $d_X(x_n, \xi) < \frac{1}{n}$ und $d_Y(f(x_n), f(\xi)) \geq \epsilon$.

Damit ist $(x_n)_{n \in \mathbb{N}}$ konvergent gegen ξ in X , aber $(f(x_n))_{n \in \mathbb{N}}$ konvergiert nicht gegen $f(\xi)$ in Y . $\square$

DEFINITION A.3: Gleichmäßige Stetigkeit

vgl. LA1, S.226

Eine Funktion $f : X \rightarrow Y$ mit $X \in \mathbb{R}$, $Y \in \mathbb{R}^p$ heißt *gleichmäßig stetig* auf X , wenn es zu jedem $\epsilon > 0$ ein $\delta > 0$ gibt, so dass für alle $x_1, x_2 \in X$ mit $d_X(x_1, x_2) < \delta$ immer $d_Y(f(x_1), f(x_2)) < \epsilon$ ist.

Bemerkung.

Eine gleichmäßig stetige Funktion ist stetig, umgekehrt nicht!

1.2. Monotonie und Beschränktheit

Um Funktionen besser beschreiben zu können, werden die Begriffe für das Wachstum (*Monotonie*), die Beschränktheit, sowie für die kleinsten bzw. größten Werte (*Extrema*) benötigt. Dieser Abschnitt orientiert sich hierbei an [LA1], [ABG] und [TA1].

Die *Monotonie* gibt Auskunft über den prinzipiellen Verlauf einer Funktion als Ganzes oder abschnittsweise:

DEFINITION A.4: Monotonie

vgl. LA1, S.115

Sei $X \subseteq \mathbb{R}$. Eine reelle Funktion $f : X \rightarrow \mathbb{R}$ heißt

monoton wachsend auf X , wenn für $x_1 < x_2$, stets $f(x_1) \leq f(x_2)$

streng monoton wachsend auf X , wenn für $x_1 < x_2$, stets $f(x_1) < f(x_2)$,

monoton fallend auf X , wenn für $x_1 < x_2$, stets $f(x_1) \geq f(x_2)$,

streng monoton fallend auf X , wenn für $x_1 < x_2$, stets $f(x_1) > f(x_2)$ ist.

Wenn eine Funktion einen Zahlenbereich nicht verlässt, folglich *beschränkt* ist, so muss sie die Randwerte nicht zwingend annehmen:

DEFINITION A.5: Beschränktheit

vgl. LA1, S.116

Sei $X \subseteq \mathbb{R}$ und $f : X \rightarrow \mathbb{R}$ eine reellwertige Funktion.

Die f heißt nach *oben (unten) beschränkt*, wenn ihr Wertebereich $f(X)$ nach oben (unten) beschränkt ist. Das heißt, es gibt Zahlen α bzw. β , sodass $f(x) \leq \alpha$ bzw. $f(x) \geq \beta$ für alle $x \in X$ ist. Somit heißen

$\sup_{x \in X} f(x)$ das *Supremum* von f

$\inf_{x \in X} f(x)$ das *Infimum* von f

Die Funktion muss hierbei Supremum und Infimum nicht annehmen.

1.3. Nullstellensatz und Zwischenwertsatz

Diesem Abschnitt beschäftigt sich mit weiteren besonderen Punkten des Funktionsgraphen, nämlich jenen, die die x-Achse schneiden (Nullstellen). Daraus wird dann der sogenannte Zwischenwertsatz hergeleitet.

Dieser Abschnitt orientiert sich hierbei an [LA1] und [A12].

Zunächst wird der Spezialfall des *Zwischenwertsatzes*, der *Nullstellensatz*, betrachtet:

SATZ A.2: Nullstellensatz (Bolzano)

vgl. LA1, S.223

Ist die Funktion f auf dem Intervall $[a, b]$ stetig und ist $f(a) < 0$ und $f(b) > 0$ (oder umgekehrt), so besitzt die Funktion mindestens eine Nullstelle.

Beweis.

Seien $[a, b] = I_0$, und oBdA $f(a) < 0, f(b) > 0$. Die Menge $A := \{x \in [a, b] : f(x) \leq 0\}$ ist nichtleer und beschränkt. Mit Hilfe des Intervallschachtelungsprinzips wird wie folgt vorgegangen:

Sei $(I_n)_{n \in \mathbb{N}_0}$ mit $I_n \in [a_n, b_n]$, $f(a_n) < 0$, $f(b_n) \geq 0$ und $|I_n| = \frac{1}{2^n} |I_0|$. Sei I_n mit obigen Eigenschaften bereits gefunden. Nun wird $m := \frac{1}{2}(a_n + b_n)$ betrachtet und folgendes definiert:

$$[a_{n+1}, b_{n+1}] := \begin{cases} [a_n, m] & \text{falls } f(m) \geq 0, \\ [m, b_n] & \text{falls } f(m) < 0 \end{cases}$$

Somit gibt es Dank der Intervallschachtelung ein x_0 mit $\{x_0\} = \bigcap_{n \in \mathbb{N}_0} I_n$.

Nun gilt $\lim_{n \rightarrow \infty} a_n = x_0 = \lim_{n \rightarrow \infty} b_n$ und da f stetig ist, folgt nach dem Folgenkriterium für Stetigkeit:

$$0 \geq \lim_{n \rightarrow \infty} f(a_n) = f(x_0) = \lim_{n \rightarrow \infty} f(b_n) \geq 0$$

somit ist $f(x_0) = 0$

□

A. Differenzialrechnung im Eindimensionalen

Ein weiterer hilfreicher Satz von Bolzano ist der *Zwischenwertsatz*:

SATZ A.3: Zwischenwertsatz (Bolzano)

vgl. LA1, S.223

Eine stetige Funktion $f : [a, b] \rightarrow \mathbb{R}$ nimmt jeden Wert zwischen $f(a)$ und $f(b)$ an.

Beweis.

Gilt $f(a) = f(b)$ gibt es nichts zu beweisen.

Andernfalls sei oBdA $f(a) < f(b)$ und η ein beliebiger Punkt aus $]f(a), f(b)[$. Dann folgt die Behauptung, indem der Nullstellensatz auf die Funktion $g(x) := f(x) - \eta$ angewendet wird. ($f(a) > f(b)$ analog). $\square$

2. Umkehrsatz

SATZ A.4: Umkehrsatz für streng monotone Funktionen

vgl. LA1, S.231

Sei f eine auf dem Intervall I definierte, streng monoton wachsende Funktion, dann ist ihre Umkehrfunktion f^{-1} auf $f(I)$ vorhanden, streng monoton wachsend und stetig. Für streng monoton fallende Funktionen analog.

Beweis.

Da f streng monoton wächst, gilt für $x_1, x_2 \in I : x_1 < x_2 \Leftrightarrow f(x_1) < f(x_2)$. Daraus folgt, dass f injektiv und die Umkehrfunktion f^{-1} streng monoton wachsend auf ihrem Definitionsbereich $f(I)$ ist. Weiter wird $\xi := f^{-1}(\eta)$ definiert und gezeigt, dass f^{-1} an einer beliebigen Stelle $\eta \in f^{-1}$ stetig ist:

1.Fall:

Ist $\xi \in I$ und ξ kein Randpunkt von I , dann existiert ein abgeschlossenes Intervall $[\xi - r, \xi + r]$ mit $r > 0$, das noch ganz in I liegt.

Für $0 < \epsilon \leq r$ gehören die Punkte $\xi - \epsilon$ und $\xi + \epsilon$ zu I , und wegen des strengen Wachsens von f gilt $f(\xi - \epsilon) < \eta < f(\xi + \epsilon)$. Infolge dessen gibt es ein $\delta > 0$, so dass auch $f(\xi - \epsilon) < \eta - \delta < \eta + \delta < f(\xi + \epsilon)$ ist.

Für jedes $y \in U_\delta(\eta) \cap f(I)$ gilt erst recht $f(\xi - \epsilon) < y < f(\xi + \epsilon)$, und wegen des strengen Wachsens von f^{-1} folgt daraus sofort $\xi - \epsilon < f^{-1}(y) < \xi + \epsilon$ und somit $|f^{-1}(y) - \xi| < \epsilon$.

Da $\xi = f^{-1}(\eta)$ ist, läuft dies auf die Abschätzung $|f^{-1}(y) - f^{-1}(\eta)| < \epsilon$ hinaus, die gerade besagt, dass f^{-1} an der Stelle η stetig ist.

2.Fall:

Ist $\xi \in I$ und ξ ein Randpunkt von I , oBdA der linke Randpunkt, so ist I links abgeschlossen und es existiert ein Intervall $[\xi, \xi + r] \subset I$ mit $r > 0$.

Ist $0 < \epsilon \leq r$, so ist $(\xi + \epsilon) \in I$ und es gilt $f(\xi) = \eta < f(\xi + \epsilon)$.

Infolgedessen gibt es ein $\delta > 0$, so dass $f(\xi) < \eta + \delta < f(\xi + \epsilon)$ ist.

Für jedes $y \in U_\delta(\eta) \cap f(I)$ gilt erst recht $f(\xi) \leq y < f(\xi + \epsilon)$, und somit, dass f^{-1} in ξ stetig ist. (Für ξ ist rechter Randpunkt von I analog) $\square$

3. Differenzierbarkeit im Eindimensionalen

Worum geht es?

Bei der Suche nach der *Ableitung* (geometrisch gesehen, die Tangentensteigung in einem Punkt) werden die Bedingungen für die *Differenzierbarkeit* geklärt.

DEFINITION A.6: Differenzierbarkeit in einem Punkt

vgl. LA1, S.261

Sei f eine auf dem (völlig beliebigen) offenen Intervall I definierte reellwertige Funktion. f ist *differenzierbar im Punkt* $\xi \in I$, wenn der Grenzwert $\lim_{x \rightarrow \xi} \frac{f(x) - f(\xi)}{x - \xi}$ bzw. $\lim_{h \rightarrow 0} \frac{f(\xi+h) - f(\xi)}{h}$ existiert und endlich ist; dieser Limes wird mit $f'(\xi)$ bezeichnet und die *Ableitung* von f an der Stelle ξ genannt.
Ist die Funktion in jedem Punkt $\xi \in I$ differenzierbar, so heißt f *differenzierbar auf* I .

Bemerkung.

Weitere Bezeichnungen für $f'(\xi)$ sind $\frac{df(\xi)}{dx}$ bzw. $\frac{d}{dx}f(\xi)$

DEFINITION A.7: stetig differenzierbar

vgl. LA1, S.285

Sei f eine auf dem Intervall I definierte Funktion. f heißt *stetig differenzierbar*, wenn ihre Ableitung f' auf I vorhanden und stetig ist.
 f heißt *n-mal stetig differenzierbar* auf I , wenn die n-te Ableitung $f^{(n)}$ auf I existiert und stetig ist.

Bemerkung.

Die Menge aller auf I stetig differenzierbaren Funktionen werden oft als $C^1(I)$ bezeichnet. Höhere Ableitungen analog.

DEFINITION A.8: Ableitungsfunktion

vgl. LA1, S.261

Sei f eine auf dem (völlig beliebigen) offenen Intervall I definierte reellwertige Funktion. Ist weiter $f \in C^1$, dann heißt die Funktion $f' : x \mapsto f'(x)$ Ableitungsfunktion (od. Ableitung von f).

SATZ A.5: Satz über die Umkehrfunktion

vgl. LA1, S.272

Sei f eine stetige, streng monotone Funktion auf dem Intervall I , sodass ihre Umkehrfunktion f^{-1} auf dem Intervall $f(I)$ existiert.

Ist in $\xi \in I$ die Ableitung $f'(\xi)$ vorhanden und $\neq 0$, so kann f^{-1} in $\eta := f(\xi)$ differenziert werden und es gilt:

$$(f^{-1})'(\eta) = \frac{1}{f'(\xi)} = \frac{1}{f'(f^{-1}(\eta))}$$

Beweis.

Ist $(y_n)_{n \in \mathbb{N}}$ eine beliebige Folge aus $f(I)$, die gegen η strebt und deren Glieder alle $\neq \eta$ sind, so liegen die $x_n := f^{-1}(y_n)$ alle in I , sind $\neq \xi$ und wegen der Stetigkeit von f^{-1} strebt $x_n \rightarrow f^{-1}(\eta) = \xi$. Daher konvergiert

$$\frac{f^{-1}(y_n) - f^{-1}(\eta)}{y_n - \eta} = \frac{x_n - \eta}{f(x_n) - f(\xi)} = \frac{1}{\frac{f(x_n) - f(\xi)}{x_n - \xi}} \rightarrow \frac{1}{f'(\xi)} \quad \square$$

SATZ A.6: Aus Differenzierbarkeit folgt Stetigkeit

vgl. LA1, S.262

Sei f eine auf dem (völlig beliebigen) offenen Intervall I definierte reellwertige Funktion, die in einem (beliebigen) Punkt $\xi \in I$ differenzierbar ist. Dann ist sie dort auch *stetig*. Die Umkehrung des Satzes gilt nicht!

A. Differenzialrechnung im Eindimensionalen

Beweis.

Zu zeigen: $\lim_{x \rightarrow \xi} f(x) = f(\xi)$ bzw. $\lim_{x \rightarrow \xi} (f(x) - f(\xi)) = 0$

Somit $\lim_{x \rightarrow \xi} (f(x) - f(\xi)) = \lim_{x \rightarrow \xi} \frac{f(x) - f(\xi)}{x - \xi} (x - \xi) = f'(\xi) * 0 = 0$ □

Beispiel.

Sei $f : \mathbb{R} \rightarrow \mathbb{R}^+$ mit $f(x) = |x|$. f ist stetig im Nullpunkt, dort aber nicht differenzierbar.

z.B.: $\frac{|\frac{1}{n}| - 0}{\frac{1}{n} - 0} = \frac{\frac{1}{n}}{\frac{1}{n}} \rightarrow 1$, aber $\frac{|\frac{-1}{n}| - 0}{\frac{-1}{n} - 0} = \frac{\frac{1}{n}}{\frac{-1}{n}} \rightarrow -1$ ◇

4. Extremwerte

Besondere Punkte der Funktion sollten bei ihrer Beschreibung nicht fehlen. Solche sind zum Beispiel *lokale* und *globale Extrema*.

DEFINITION A.9: Kritischer Punkt, stationärer Punkt

vgl. TA1, S.175

Sei $f : D \rightarrow \mathbb{R}$ differenzierbar, so heißt die Stelle $\xi \in D$ mit $f'(\xi) = 0$ *kritischer* bzw. *stationärer Punkt*.

4.1. Lokale Extrema

DEFINITION A.10: Lokale Extrema

vgl. LA1, S.266

Sei $f : X \rightarrow \mathbb{R}$ mit $X \subseteq \mathbb{R}$ eine reelle Funktion. Eine Stelle $\xi \in X$ heißt *lokales Maximum* bzw. *Minimum*, wenn es eine δ -Umgebung U von ξ gibt, sodass für alle $x \in U \cap X$ stets $f(x) \leq f(\xi)$ bzw. $f(x) \geq f(\xi)$ ist. Lokale Maxima und Minima heißen auch *lokale Extrema*.

Notwendige Bedingung

SATZ A.7: Notwendige Bedingung für lokale Extrema

vgl. LA1, S.266

Die Funktion $f : X \rightarrow \mathbb{R}$ besitze im inneren Punkt ξ von X ein lokales Extremum und sei überdies in ξ differenzierbar. Dann ist notwendig $f'(\xi) = 0$.

Bemerkung.

Der Satz sagt aus, dass die Funktion an jeder lokalen Extremstelle eine horizontale Tangente besitzt. **Die Umkehrung des Satzes gilt nicht!**

A. Differenzialrechnung im Eindimensionalen

Beweis.

Obiger Satz wird für Maxima gezeigt (Minima analog):

f besitze in ξ ein lokales Maximum und sei dort auch differenzierbar. Zu zeigen ist, dass $f'(\xi) = \lim_{h \rightarrow 0} \frac{f(\xi+h)-f(\xi)}{h} = 0$ ist.

Da f in ξ differenzierbar ist, gilt, dass linksseitiger und rechtsseitiger Grenzwert des Differenzialquotienten, den selben Wert ergeben:

$$\lim_{h \rightarrow 0^-} \frac{f(\xi+h)-f(\xi)}{h} = f'(\xi) = \lim_{h \rightarrow 0^+} \frac{f(\xi+h)-f(\xi)}{h}.$$

Weiter besitzt f in ξ laut Voraussetzung ein lokales Maximum, daher gilt für $\delta_r > 0$ mit $h \in (0, \delta_r)$:

$$f(\xi+h) \leq f(\xi) \Leftrightarrow f(\xi+h) - f(\xi) \leq 0 \quad \text{und somit ist auch} \quad \frac{f(\xi+h)-f(\xi)}{h} \leq 0$$

$$\text{Für den rechtsseitigen Grenzwert gilt demnach: } f'(\xi) = \lim_{h \rightarrow 0^+} \frac{f(\xi+h)-f(\xi)}{h} \leq 0$$

Andererseits gilt auch für $\delta_l > 0$ mit $h \in (-\delta_l, 0)$:

$$f(\xi+h) \leq f(\xi) \Leftrightarrow f(\xi+h) - f(\xi) \leq 0$$

Es ist somit $f'(\xi) = \lim_{h \rightarrow 0^-} \frac{f(\xi+h)-f(\xi)}{h} \geq 0$. Daher ist $f'(\xi) \leq 0 \leq f'(\xi)$ und somit ist bewiesen, dass $f'(\xi) = 0$ ist. $\square$

Hinreichende Bedingung

SATZ A.8: Hinreichende Bedingung für lokale Extrema (1)

vgl. ABG, S.323

Sei f an der Stelle ξ und in einer Umgebung $U(\xi)$ differenzierbar.

Wenn $f'(\xi) = 0$ ist und f' an der Stelle ξ einen Vorzeichenwechsel von $+$ nach $-$ hat, dann ist $f'(\xi)$ ein lokales Maximum.

Wenn $f'(\xi) = 0$ ist und f' an der Stelle ξ einen Vorzeichenwechsel von $-$ nach $+$ hat, dann ist $f'(\xi)$ ein lokales Minimum.

SATZ A.9: Hinreichende Bedingung für lokale Extrema (2)

vgl. LA1, S.281ff

Die Funktion f sei an der Stelle ξ zweimal differenzierbar.

Wenn $f'(\xi) = 0$ und $f''(\xi) < 0$ ist, so liegt ein lokales Maximum vor.

Wenn $f'(\xi) = 0$ und $f''(\xi) > 0$ ist, so liegt ein lokales Minimum vor.

Bemerkung. Da für den Beweis des Satzes unter anderem der Mittelwertsatz nötig ist, folgt dieser erst später. (5, S. 43)

Der Extremalsatz

SATZ A.10: Satz von Maxima und Minima / Extremalsatz

vgl. LA1, S.225

Sei $f : [a, b] \rightarrow \mathbb{R}$ eine stetige Funktion mit kompakten Definitionsbereich $X = [a, b]$. Dann besitzt f ein *Minimum* und ein *Maximum*.
Es gibt somit in X eine Minimalstelle x_1 und eine Maximalstelle x_2 , so dass

$$f(x_1) \leq f(x) \leq f(x_2) \text{ für alle } x \in X \text{ ist.}$$

Bemerkung.

Der Satz gibt nur Auskunft über die Existenz von Extrema, liefert aber kein Verfahren zum bestimmen.

Beweis.

Angenommen f ist nach oben beschränkt. Wäre das nicht der Fall, dann würde für jedes $n \in \mathbb{N}$ ein $z_n \in [a, b]$ existieren, sodass $f(z_n) \geq n$.

Nach dem Satz von Bolzano-Weierstraß (vgl. 0.2, S. 13) hat die Folge $(z_n)_{n \in \mathbb{N}}$ auf $[a, b]$ eine konvergente Teilfolge $(z_{n_k})_{k \in \mathbb{N}}$, sodass der Grenzwert $s = \lim_{k \rightarrow \infty} (z_{n_k}) \in [a, b]$ liegt.

Da f stetig ist, folgt $\lim_{k \rightarrow \infty} f(z_{n_k}) = f(s)$.

Das ist ein Widerspruch dazu, dass $f(x_n) \geq n$. Somit muss f nach oben beschränkt sein.

Weiters sei zu zeigen, dass $\sup(f)$ ein Funktionswert ist. Dann existiert eine Folge $(y_n)_{n \in \mathbb{N}}$ mit $y_n \in f[a, b]$ für alle n , die gegen $\sup(f)$ konvergiert. Da $y_n \in f[a, b]$ liegt, ist jedes y_n ein Funktionswert der Form $y_n = f(x_n)$ für $x_n \in [a, b]$. Deshalb gibt es nun eine Folge $(x_n)_{n \in \mathbb{N}}$ mit $x_n \in [a, b]$ für alle n , sodass die Folge der Funktionswerte $(y_n)_{n \in \mathbb{N}}$ gegen $\sup(f)$ konvergiert.

Nach dem Satz von Bolzano Weierstraß (vgl. 0.2, S. 13) hat die Folge $(x_n)_{n \in \mathbb{N}}$ eine konvergente Teilfolge $(x_{n_k})_{k \in \mathbb{N}}$, sodass der Grenzwert $s = \lim_{k \rightarrow \infty} (x_{n_k})$ existiert. Da f stetig ist gilt, $\lim_{k \rightarrow \infty} f(x_{n_k}) = f(s)$. Da nun jede konvergente Teilfolge einer Folge gegen den selben Grenzwert konvergiert, gilt, $\lim_{k \rightarrow \infty} f(x_{n_k}) = \sup(f)$. Somit gilt nun $f(s) =$

A. Differenzialrechnung im Eindimensionalen

$\sup(f)$ und $\sup(f)$ ist ein Funktionswert. Da nun gezeigt ist, dass $f(x) \leq \sup(f)$ für alle $x \in [a, b]$, sind Supremum und Maximum identisch.
(der Beweis für das Minimum erfolgt analog mit dem Infimum) $\square$

4.2. Globale Extrema

DEFINITION A.11: Globale Extrema

vgl. LA1, S.116

Wird das Supremum angenommen, dann gibt es ein $x_0 \in X$ mit $f(x_0) = \sup_{x \in X} f(x)$ und $\max_{x \in X} f(x)$ wird *Maximum* von f auf X genannt. x_0 heißt dann *Maximalstelle*.

Wird das Infimum angenommen, dann gibt es ein $x_1 \in X$ mit $f(x_1) = \inf_{x \in X} f(x)$ und $\min_{x \in X} f(x)$ wird *Minimum* von f auf X genannt. x_1 heißt dann *Minimalstelle*.

Minimum und Maximum heißen Extremwerte (Extrema),
Minimalstelle und Maximalstelle heißen Extremstellen.

Bemerkung.

Um festzustellen, ob es sich bei *lokalen Extrema* sogar um *globale Extrema* handelt, müssen auch die Definitionsbereichsgrenzen, Polstellen und Unstetigkeitsstellen untersucht werden.

5. Mittelwertsatz der Differentialrechnung

Zunächst wird einen Spezialfall, der Satz von Rolle, definiert und damit einen der zentralen Sätze der Differenzialrechnung, der Mittelwertsatz beweisen.
Dieser Abschnitt orientiert sich hierbei an [LA1].

SATZ A.11: Satz von Rolle

vgl. LA1, S.279

Ist die Funktion φ auf dem kompakten Intervall $[a, b]$ stetig, auf (a, b) differenzierbar und die Funktionswerte der Intervallendpunkte stimmen überein, somit $\varphi(a) = \varphi(b)$, so verschwindet ihre Ableitung an mindestens einer Stelle $\xi \in (a, b)$.

Michel Rolle, französischer Mathematiker (1652 - 1719)

Beweis.

Für konstantes φ ist der Satz trivial. Ist φ nicht konstant, dann existiert ein $x_0 \in (a, b)$ mit $\varphi(x_0) > \varphi(a)$ oder $\varphi(x_0) < \varphi(a)$. Dann existiert eine Maximalstelle (bzw. Minimalstelle) der Funktion in einem Punkt $\xi \in (a, b)$. Laut der notwendigen Bedingung für lokale Extrema (A.7, S. 37) gilt $\varphi'(\xi) = 0$ $\square$

SATZ A.12: Mittelwertsatz der Differentialrechnung

vgl. LA1, S.279

Ist die Funktion f auf dem kompakten Intervall $[a, b]$ stetig und im inneren des Intervalls differenzierbar, so gibt es mindestens einen Punkt ξ in (a, b) , an dem $f'(\xi) = \frac{f(b) - f(a)}{b - a}$ bzw. $f(b) - f(a) = f'(\xi)(b - a)$ ist.

A. Differenzialrechnung im Eindimensionalen

Beweis.

Der Mittelwertsatz folgt, indem der Satz von Rolle auf folgende Funktion angewendet wird:

$$F : [a, b] \rightarrow \mathbb{R} \quad \text{mit} \quad F(x) := f(x) - \frac{f(b) - f(a)}{b - a}(x - a)$$

F ist stetig in $[a, b]$ und differenzierbar in $]a, b[$. Da $F(a) = f(a) = F(b)$ ist, existiert nach dem Satz von Rolle ein $\xi \in]a, b[$ mit $F'(\xi) = 0$.

Es folgt die Behauptung, da

$$F'(\xi) = f'(\xi) - \frac{f(b) - f(a)}{b - a} = 0$$

□

Da nun der Mittelwertsatz bewiesen ist, kann nun der Beweis für den Satz Hinreichende *Bedingung für Lokale Extrema* (2) (A.9, S. 38) geführt werden. Dazu wird jedoch noch folgender Satz benötigt:

SATZ A.13: Monotonie

vgl. LA1, S.281

Sei $f : I \rightarrow \mathbb{R}$ auf einem Intervall I stetig und im Inneren Punkt ξ des Intervalls auch differenzierbar, dann ist f
streng monoton steigend, falls $f'(\xi) > 0$ für alle ξ .
streng monoton fallend, falls $f'(\xi) < 0$ für alle ξ .

Beweis.

Für zwei beliebige Punkte $x_1 < x_2$ aus I ist bei geeignetem $\xi \in (x_1, x_2)$ stets $f(x_2) - f(x_1) = f'(\xi)(x_2 - x_1)$. Daraus folgen obige Behauptungen.

□

Nun folgt der Beweis für den Satz Hinreichende *Bedingung für Lokale Extrema* (2):

Beweis.

Es wird gezeigt, ob ein lokales Maximum vorliegt (Minimum analog).

Seien $f : [a, b] \rightarrow \mathbb{R}$, $f''(\xi) < 0$ und $f'(\xi) = 0$. Dann folgt, dass $0 > f''(\xi) = \lim_{h \rightarrow 0} \frac{f'(\xi+h) - f'(\xi)}{h} = \lim_{h \rightarrow 0} \frac{f'(\xi+h)}{h}$.

Es existiert ein $\delta > 0$, sodass $\frac{f'(\xi+h)}{h} < 0$ für alle $x \in (a, b)$ mit $|h| < \delta$.

Daher folgt für den linksseitigen Grenzwert $f'(\xi) > 0$ und für den rechtsseitigen $f'(\xi) < 0$ und f ist auf $(\xi - \delta, \xi)$ streng monoton wachsend, sowie auf $(\xi, \xi + \delta)$ streng monoton fallend. Daraus folgt, dass in ξ ein lokales Maximum vorliegen muss. $\square$

6. Differentiationsregeln für Eindimensionale Funktionen

Für den Praktiker sind folgende Differentiationsregeln ein notwendiges Hilfsmittel, um schnell an die Ableitung(-sfunktion) zu kommen. Daher werden sie hier kurz definiert und mittels Differentialquotienten und den Rechenregeln für Grenzwerte beweisen.

6.1. Faktorregel

vgl. GM1, S.170

Sei $f : V \rightarrow \mathbb{R}$ in $x \in V$ differenzierbare Funktion und $\lambda \in \mathbb{R}$, dann ist $\lambda f : V \rightarrow \mathbb{R}$ differenzierbar und es gilt:

$$(\lambda f)'(x) = \lambda f'(x)$$

Beweis.

Für $\xi \in V$ gilt:

$$(\lambda f)'(\xi) = \lim_{x \rightarrow \xi} \frac{\lambda f(x) - \lambda f(\xi)}{x - \xi} = \lambda \lim_{x \rightarrow \xi} \frac{f(x) - f(\xi)}{x - \xi} = \lambda f'(\xi)$$

□

6.2. Summenregel

GM1, S.170

Seien $f, g : V \rightarrow \mathbb{R}$ in $x \in V$ differenzierbare Funktionen, dann ist $f+g : V \rightarrow \mathbb{R}$ differenzierbar und es gilt:

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

Beweis.

Für $\xi \in V$ gilt:

$$\begin{aligned} (f+g)'(\xi) &= \lim_{x \rightarrow \xi} \frac{[f(x) + g(x)] - [f(\xi) + g(\xi)]}{x - \xi} = \lim_{x \rightarrow \xi} \frac{f(x) - f(\xi)}{x - \xi} + \lim_{x \rightarrow \xi} \frac{g(x) - g(\xi)}{x - \xi} \\ &= f'(\xi) + g'(\xi) \end{aligned}$$

□

6.3. Produktregel

GM1; S.170

Seien $f, g : V \rightarrow \mathbb{R}$ in $x \in V$ differenzierbare Funktionen, dann ist $fg : V \rightarrow \mathbb{R}$ differenzierbar und es gilt:

$$(fg)'(x) = f'(x)g(x) + f(x)g'(x)$$

Beweis.

Für $\xi \in V$ gilt:

$$\begin{aligned}(fg)'(\xi) &= \lim_{x \rightarrow \xi} \frac{(f \cdot g)(x) - (f \cdot g)(\xi)}{x - \xi} \\ &= \lim_{x \rightarrow \xi} \frac{f(x)g(x) - f(\xi)g(\xi)}{x - \xi} = \dots\end{aligned}$$

(Hinzufügen des folgenden Termes: $-f(\xi)g(x) + f(\xi)g(x) = 0$)

$$\begin{aligned}\dots &= \lim_{x \rightarrow \xi} \frac{f(x)g(x) - f(\xi)g(x) + f(\xi)g(x) - f(\xi)g(\xi)}{x - \xi} \\ &= \lim_{x \rightarrow \xi} \left(\frac{f(x)g(x) - f(\xi)g(x)}{x - \xi} + \frac{f(\xi)g(x) - f(\xi)g(\xi)}{x - \xi} \right) \\ &= \lim_{x \rightarrow \xi} \frac{f(x)g(x) - f(\xi)g(x)}{x - \xi} + \lim_{x \rightarrow \xi} \frac{f(\xi)g(x) - f(\xi)g(\xi)}{x - \xi} \\ &= \lim_{x \rightarrow \xi} g(x) \frac{f(x) - f(\xi)}{x - \xi} + \lim_{x \rightarrow \xi} f(\xi) \frac{g(x) - g(\xi)}{x - \xi}\end{aligned}$$

Da für $\lim_{x \rightarrow \xi} g(x) \rightarrow g(\xi)$ und die Grenzwerte der Quotienten jeweils die Ableitungen an der Stelle ξ sind, folgt obige Differentiationsregel. $\square$

6.4. Quotientenregel

GM1; S.170

Seien $f, g : V \rightarrow \mathbb{R}$ in $x \in V$ differenzierbare Funktionen, und ist $g(\xi) \neq 0$ für alle $\xi \in V$, so ist auch die Funktion $(f/g) : V \rightarrow \mathbb{R}$ differenzierbar und es gilt:

$$\left(\frac{f}{g} \right)'(x) = \frac{f'(x)g(x) - f(x)g'(x)}{g(x)^2}$$

A. Differenzialrechnung im Eindimensionalen

Beweis.

Für $\xi \in V$ gilt:

$$\begin{aligned}\left(\frac{f}{g}\right)'(\xi) &= \lim_{x \rightarrow \xi} \frac{\left(\frac{f(x)}{g(x)}\right) - \left(\frac{f(\xi)}{g(\xi)}\right)}{x - \xi} \\ &= \lim_{x \rightarrow \xi} \left[\frac{1}{g(x)g(\xi)} \left(g(\xi) \frac{f(x) - f(\xi)}{x - \xi} - f(\xi) \frac{g(x) - g(\xi)}{x - \xi} \right) \right] \\ &= \frac{f'(\xi)g(\xi) - f(\xi)g'(\xi)}{g(\xi)^2}\end{aligned}$$

□

6.5. Kettenregel

vgl. LA1, S.271ff

Vorgelegt seien die Funktionen $g : I_g \subset \mathbb{R} \rightarrow \mathbb{R}$ und $f : I_f \subset \mathbb{R} \rightarrow \mathbb{R}$ mit $I_f \supset g(I_g)$. g sei in ξ , f in $g(\xi)$ differenzierbar. Dann ist $f \circ g$ in ξ differenzierbar, und es gilt

$$(f \circ g)'(\xi) = f'(g(\xi)) g'(\xi)$$

Für den Beweis der Kettenregel wird der nun folgende Satz verwendet:

SATZ A.14: Alternative Darstellung differenzierbarer Funktionen

vgl. LA1, S.268

Die Funktion $f : I \rightarrow \mathbb{R}$ ist genau dann in $\xi \in I$ differenzierbar, wenn das Inkrement $f(\xi + h) - f(\xi)$ in folgender Form dargestellt werden kann:

$$f(\xi + h) - f(\xi) = ah + r(h) \quad \text{mit} \quad \lim_{h \rightarrow 0} \frac{r(h)}{h} = 0$$

Bemerkung.

In diesem Fall ist $a = f'(\xi)$. r ist der Fehler der Approximation.

Beweis.

vgl. LA1, S.267ff

Ist die Funktion $f : I \rightarrow \mathbb{R}$ in $\xi \in I$ differenzierbar, so strebt

$$\rho(h) := \frac{f(\xi+h)-f(\xi)}{h} - f'(\xi) \rightarrow 0 \quad \text{für } h \rightarrow 0.$$

Wird $r(h) := \rho(h)h$ gesetzt, gilt die Gleichung $f(\xi + h) - f(\xi) = f'(\xi)h + r(h)$ mit $\lim_{h \rightarrow 0} \frac{r(h)}{h} = 0$.

Wenn weiters $f(\xi + h) - f(\xi) = ah + r(h)$ zu $\frac{f(\xi+h)-f(\xi)}{h} = a + \frac{r(h)}{h}$ umgeformt und davon nun der Grenzwert gebildet wird, folgt, das $\lim_{h \rightarrow 0} \frac{f(\xi+h)-f(\xi)}{h} = a + \lim_{h \rightarrow 0} \frac{r(h)}{h} = a + 0 = a$. Somit ist gezeigt, dass a die Ableitung an der Stelle ξ ist (vgl. A.6, S. 34). $\square$

Nun folgt der Beweis der Kettenregel:

Beweis.

Für $\xi \in V$ wird $\eta := g(\xi)$ gesetzt und gemäß der **Differenzierbarkeit in einem Punkt** (A.6, S. 34) gilt nun:

$$g(\xi + h) - g(\xi) = g'(\xi)h + r_1(h), \quad \frac{r_1(h)}{h} =: \varrho_1(h) \rightarrow 0 \quad \text{für } h \rightarrow 0,$$

$$f(\eta + k) - f(\eta) = f'(\eta)k + r_2(k), \quad \frac{r_2(k)}{k} =: \varrho_2(k) \rightarrow 0 \quad \text{für } k \rightarrow 0.$$

Aus diesen Darstellungen folgt

$$\begin{aligned} (f \circ g)(\xi + h) - (f \circ g)(\xi) &= f(g(\xi + h)) - f(g(\xi)) \\ &= f'(\eta)[g(\xi + h) - g(\xi)] + r_2(g(\xi + h) - g(\xi)) \\ &= f'(\eta)[g'(\xi)h + r_1(h)] + r_2(g'(\xi)h + r_1(h)) \\ &= f'(\eta)g'(\xi)h + r(h) \end{aligned}$$

$$\begin{aligned} \text{mit} \quad r(h) &:= f'(\eta)r_1(h) + r_2(g'(\xi)h + r_1(h)) \\ &= f'(\eta)h\varrho_1(h) + [g'(\xi)h + r_1(h)]\varrho_2(g'(\xi)h + r_1(h)). \end{aligned}$$

Offenbar strebt $\frac{r(h)}{h} \rightarrow 0$ für $h \rightarrow 0$, und daraus folgt mit der **Alternativen Darstellung differenzierbarer Funktionen** (A.14, S. 46) die Behauptung. $\square$

6.6. Potenzfunktionen

vgl. LA1, S.273

Sei $f : V \rightarrow \mathbb{R}$ eine in $x \in V$ differenzierbare Funktion, dann gilt für $n \in \mathbb{N}$ und alle $x \neq 0$:

$$f(x)' = (x^n)' = n x^{n-1}$$

Insbesondere gilt diese Ableitungsregel auch für:

$$(x^{-n})' = -n x^{-n-1} \quad \text{für } n \in \mathbb{N} \text{ und alle } x \neq 0 \text{ und } \left(\frac{1}{x}\right)' = (x^{-1})' = -1 x^{-2} = -\frac{1}{x^2}$$

Beweis.

Für $\xi \in V$ gilt:

$$\begin{aligned} f'(\xi) &= \lim_{x \rightarrow \xi} \frac{x^n - \xi^n}{x - \xi} = \lim_{x \rightarrow \xi} \frac{(x - \xi)(x^{n-1} + x^{n-2}\xi + \dots + \xi^{n-1})}{x - \xi} \\ &= \lim_{x \rightarrow \xi} x^{n-1} + x^{n-2}\xi + \dots + \xi^{n-1} = \xi^{n-1} + \xi^{n-2}\xi + \dots = n\xi^{n-1} \end{aligned}$$

□

Bemerkung.

Für obigen Beweis wird verwendet, dass $x^n - \xi^n = (x - \xi)(x^{n-1} + x^{n-2}\xi + \dots)$ ist. Dieser Zusammenhang lässt sich mittels Polynomdivision finden.

6.7. Polynomfunktionen

vgl. LA1, S.273

Bei Anwendung von Kettenregel, Faktorregel und der Ableitungsregel für Potenzfunktionen, lässt sich die Ableitung eines beliebigen Polynoms bestimmen:

$$(a_0 + a_1x + a_2x^2 + \dots + a_nx^n)' = a_1 + 2a_2x + \dots na_nx^{n-1} \quad \text{für alle } x$$

Bemerkung.

Für den Beweis, siehe Summenregel und Ableitungsregel für Potenzfunktionen.

6.8. Exponentialfunktion

vgl. LA1, S.274

Sei $f : \mathbb{R} \rightarrow \mathbb{R}$ mit $x \mapsto a^x$, $a \in \mathbb{R}^+$ eine Feste Zahl und f eine in x differenzierbare Funktion, dann gilt für alle x :

$$(a^x)' = a^x \ln(a) \quad \text{bzw. wenn } a = e \quad (e^x)' = e^x$$

Beweis.

$$\lim_{x \rightarrow \xi} \frac{a^x - a^\xi}{x - \xi} = \lim_{h \rightarrow 0} \frac{a^{x+h} - a^x}{h} = a^x \lim_{h \rightarrow 0} \frac{a^h - 1}{h} = a^x \ln(a)$$

□

Bemerkung.

Hierbei wurde verwendet, dass $\lim_{h \rightarrow 0} \frac{a^h - 1}{h} = \ln(a)$ ist. $\ln(e) = 1$

6.9. Logarithmus Funktion

vgl. LA1, S.274

Sei $f : V \rightarrow \mathbb{R}$, wobei $V \subset \mathbb{R}^+$, mit $x \mapsto \ln(x)$, eine in x differenzierbare Funktion, dann gilt:

$$(\ln(x))' = \frac{1}{x} \quad \text{für alle } x > 0.$$

Bemerkung.

Die natürliche Logarithmusfunktion $\ln(x)$ ist die Umkehrfunktion der e-Funktion.

Beweis.

Da e^x auf dem gesamten Definitionsbereich positiv ist, folgt aus dem Umkehrsatz für streng monotone Funktionen (A.4, S. 33) und der Ableitungsregel der Exponentialfunktion (6.8, S. 49), dass $(\ln x)' = \frac{1}{e^{\ln(x)}}$ ist. □

6.10. Trigonometrische Funktionen

vgl. GM1, S.273

Für die nun folgenden Winkelfunktionen sei $x \in \mathbb{R}$ im Bogenmaß.
Weiter werden bei der Beweisführung folgende Aussagen verwendet:

$\sin(x)$ und $\cos(x)$ sind stetige Funktionen auf $\mathbb{R}$ $\lim_{a \rightarrow 0} \frac{\sin(a)}{a} = 1$

Additionstheoreme:

$$\begin{aligned}\sin(x+y) &= \sin(x)\cos(y) + \cos(x)\sin(y) \\ \cos(x+y) &= \cos(x)\cos(y) - \sin(x)\sin(y)\end{aligned}$$

Da $\sin(x)$ eine ungerade und $\cos(x)$ eine gerade Funktion sind, folgt

$$\sin(-x) = -\sin(x) \text{ und } \cos(-x) = \cos(x).$$

Daraus folgen weiter

$$\begin{aligned}\sin(x-y) &= \sin(x)\cos(y) - \cos(x)\sin(y) \\ \cos(x-y) &= \sin(x)\sin(y) + \cos(x)\cos(y)\end{aligned}$$

Und daher kann folgendermaßen umgeformt werden:

$$\begin{aligned}\sin(x) &= \sin\left(\frac{x+y}{2} + \frac{x-y}{2}\right) = \sin\left(\frac{x+y}{2}\right)\cos\left(\frac{x-y}{2}\right) + \cos\left(\frac{x+y}{2}\right)\sin\left(\frac{x-y}{2}\right) \\ \sin(y) &= \sin\left(\frac{x+y}{2} - \frac{x-y}{2}\right) = \sin\left(\frac{x+y}{2}\right)\cos\left(\frac{x-y}{2}\right) - \cos\left(\frac{x+y}{2}\right)\sin\left(\frac{x-y}{2}\right)\end{aligned}$$

Werden die beiden Gleichungen subtrahiert, ergibt sich folgender Zusammenhang:

$$\sin(x) - \sin(y) = 2 \cos\left(\frac{x+y}{2}\right) \sin\left(\frac{x-y}{2}\right)$$

$$\text{Und analog dazu: } \cos(x) - \cos(y) = -2 \sin\left(\frac{x+y}{2}\right) \sin\left(\frac{x-y}{2}\right)$$

6. Differentiationsregeln für Eindimensionale Funktionen

Die Sinusfunktion:

Sei $f : V \rightarrow \mathbb{R}$ die in $x \in V$ differenzierbare Sinusfunktion mit $f : x \mapsto \sin(x)$, dann gilt:

$$\sin(x)' = \cos(x)$$

Beweis.

$$\begin{aligned} \lim_{x \rightarrow \xi} \frac{\sin(x) - \sin(\xi)}{x - \xi} &= \lim_{x \rightarrow \xi} \frac{2 \cos\left(\frac{x+\xi}{2}\right) \sin\left(\frac{x-\xi}{2}\right)}{x - \xi} = \\ &= \lim_{x \rightarrow \xi} \cos\left(\frac{x+\xi}{2}\right) \frac{\sin\left(\frac{x-\xi}{2}\right)}{\frac{x-\xi}{2}} = \cos(\xi) \end{aligned}$$

Hierbei wurde verwendet, dass für $x \rightarrow \xi$ der Ausdruck $\frac{x-\xi}{2}$ gegen 0 geht und daher $\lim_{a \rightarrow 0} \frac{\sin(a)}{a} = 1$ ist. $\square$

Die Cosinusfunktion:

Sei $f : V \rightarrow \mathbb{R}$ die in $x \in V$ differenzierbare Cosinusfunktion mit $f : x \mapsto \cos(x)$, dann gilt:

$$\cos(x)' = -\sin(x)$$

Beweis.

$$\lim_{x \rightarrow \xi} \frac{\cos(x) - \cos(\xi)}{x - \xi} = \lim_{x \rightarrow \xi} \frac{-2 \sin\left(\frac{x+\xi}{2}\right) \sin\left(\frac{x-\xi}{2}\right)}{x - \xi} =$$

$$\lim_{x \rightarrow \xi} -\sin\left(\frac{x+\xi}{2}\right) \frac{\sin\left(\frac{x-\xi}{2}\right)}{\frac{x-\xi}{2}} = -\sin(\xi)$$

$\square$

A. Differenzialrechnung im Eindimensionalen

Die Tangensfunktion:

Sei $f : V \rightarrow \mathbb{R}$ die in $x \in V$ differenzierbare Tangensfunktion mit $f : x \mapsto \tan(x)$, dann gilt:

$$\tan(x)' = 1 + \tan^2(x)$$

Beweis.

Da $\tan(\xi) = \frac{\sin(\xi)}{\cos(\xi)}$ folgt die Behauptung mittels Anwendung der Quotientenregel. $\square$

B. Differentialrechnung im Mehrdimensionalen

« Es gibt Dinge, die den meisten Menschen unglaublich erscheinen,
die nicht Mathematik studiert haben. »

Archimedes von Syrakus, griechischer Mathematiker (um 287 v. Chr.)

Worum geht es?

In Kapitel **B** werden, wie auch schon im Kapitel **A**, die *Differenzierbarkeit* (hier mehrdimensional) und charakteristische Punkte von Funktionen erläutert, jedoch nun allgemein gültig und somit auch für mehrere Dimensionen. Der *Umkehrsatz*, der *Hauptsatz über Implizite Funktionen*, sowie das *Lagrange-Verfahren* werden ebenso vorgestellt.

Was ist mehrdimensionale Analysis?

In der *mehrdimensionalen Analysis* werden Abbildungen zwischen höherdimensionalen euklidischen Räumen (0.24, S. 20) betrachtet. Solche Funktionen, die von mehreren Variablen abhängen, können wie folgt allgemein dargestellt werden:

$$f : \mathbb{R}^p \supseteq D \rightarrow \mathbb{R}^q \quad x = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix} \mapsto f(x_1, \dots, x_p) = \begin{pmatrix} f_1(x) \\ \vdots \\ f_q(x) \end{pmatrix}$$

Beispiel.

An diesem Beispiel ist ersichtlich, wie der Ebene ($\mathbb{R}^2$) reelle Funktionswerte ($\mathbb{R}$) zugeordnet werden:

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f(x, y) = x^2 + y^2$$

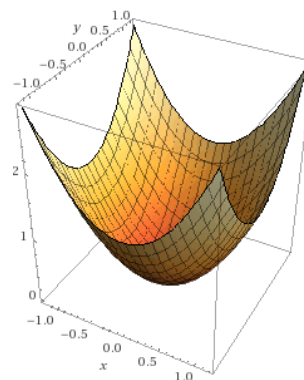


Abbildung B.1.: 3D-Grafik der Funktion

Bemerkung.

Abbildungen von $\mathbb{R}^2$ nach $\mathbb{R}$ sind noch dreidimensional vorstellbar.

1. Grundbegriffe

Wie können die verschiedenen Funktionsarten eingeteilt werden?

Grundsätzlich können Funktionen darin unterschieden werden, von und in welche Räume sie abbilden.

Eine *reellwertige Funktion*, ein- oder mehrdimensional, ordnet einer beliebigen Menge reelle Zahlen zu. Handelt es sich bei der Ausgangsmenge um $\mathbb{R}^p$, wird diese *Skalarfeld* genannt.

Wenn die Funktion in einen Vektorraum abbildet, heißt diese *vektorwertige Funktion*, bzw. ist von einem *Vektorfeld* die Rede, wenn es sich um den Vektorraum $\mathbb{R}^p$ handelt.

DEFINITION B.1: Reellwertige Funktionen

vgl. GM1, S.104

Sei D eine Teilmenge von $\mathbb{R}^p$.

Eine Abbildung $f : D \rightarrow \mathbb{R}$ heißt *reellwertige Funktion* auf D .

Die Menge D heißt Definitionsbereich von f .

Beispiele für reellwertige Funktionen...

$$\begin{array}{lll} \text{in einer Variable (p=1):} & f(x) = 2x & f(x) = x + 1 & f(x) = x^2 \\ \text{mehrerer Variablen (p>1):} & f(x_1, x_2) = x_1 + x_2 & f(x_1, x_2, x_3) = 2x_1 + \frac{x_2}{x_3} \end{array}$$

Bemerkung.

Reellwertige Funktionen werden häufig als *reelle Funktionen* bezeichnet.

Beispiel.

Sei $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ mit $f(x_1, x_2) = x_1 \cdot x_2$

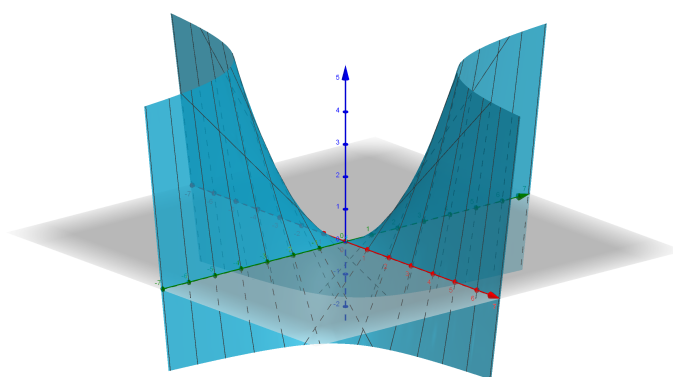


Abbildung B.2.: Beispiel einer reellwertigen Funktion

DEFINITION B.2: Skalarfeld

vgl. LA2; S.380

Sei X eine nichtleere Teilmenge des $\mathbb{R}^p$. Eine Funktion $f : X \rightarrow \mathbb{R}$ heißt *Skalarfeld*.

Bemerkung.

Ein Skalarfeld ordnet jedem Punkt des Raumes einen Wert (Skalar) zu.

Beispiel.

Die Temperaturverteilung im Raum als Skalarfeld, kann mit Farben dargestellt werden, die für die entsprechenden Werte stehen:

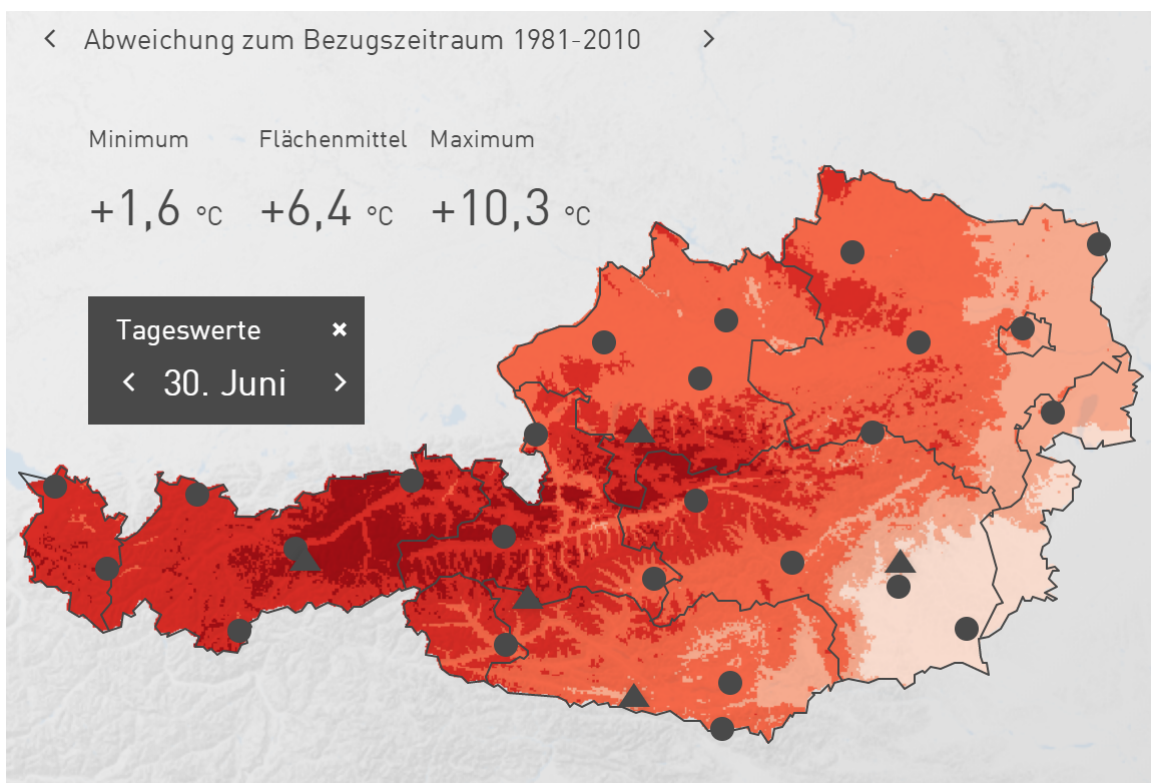


Abbildung B.3.: Tagesmittel der Lufttemperatur für den 30. Juni 2019 (ZAMG)

Quelle (abgerufen am 01.07.2019):

zamg.ac.at/cms/de/klima/klima-aktuell/klimamonitoring/?param=t&period=period-ymd-2019-06-30&ref=3

DEFINITION B.3: Vektorwertige Funktionen

vgl. GM2, S.85

Sei D eine Teilmenge von $\mathbb{R}$. Eine Abbildung $f : D \rightarrow V$ bei der die Zielmenge V ein Vektorraum ist, heißt *vektorwertigen Funktion* auf D . Oftmals ist V der Vektorraum $\mathbb{R}^p$.

Beispiel. Einheitskreis in der Ebene

Sei $f : [0, 2\pi[\rightarrow \mathbb{R}^2$ eine Funktion, die x auf den Vektor der Form $\begin{pmatrix} \cos(x) \\ \sin(x) \end{pmatrix}$ abbildet.

Beispiel (Schraubenfeder).

Sei $f : [0, 4.5\pi] \rightarrow \mathbb{R}^3$ mit $f(x) := \begin{pmatrix} \cos(x) \\ \sin(x) \\ \frac{x}{10} \end{pmatrix}$

Nachstehende Abbildung zeigt die hier definierte vektorwertige Funktion, die an eine Schraubenfeder erinnert:

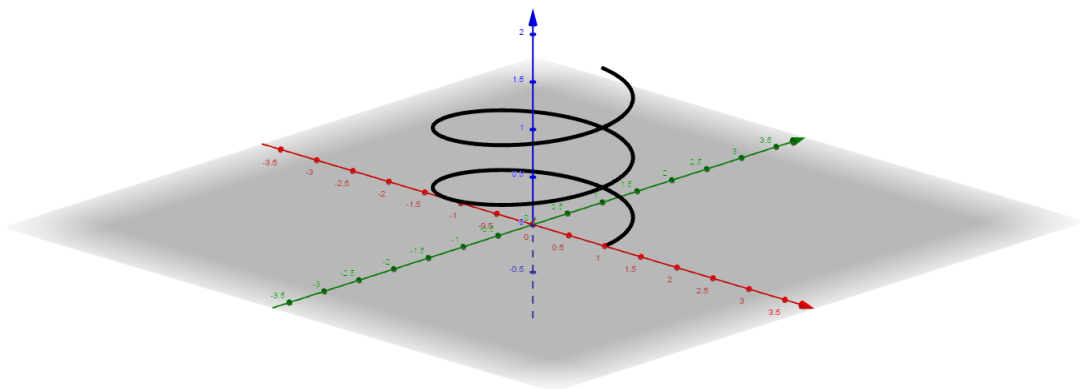


Abbildung B.4.: Schraubenfeder

DEFINITION B.4: Vektorfeld

vgl. LA2; S.380

Sei X eine nichtleere Teilmenge des $\mathbb{R}^p$. Eine Funktion $f : X \rightarrow \mathbb{R}^q$ heißt *Vektorfeld*.

Bemerkung.

Ein Vektorfeld ordnet jedem Punkt des Raumes einen Vektor zu.

Beispiel.

Folgende Abbildung zeigt die grafische Darstellung eines Vektorfeldes, welches mit Hilfe des Programmes Mathematica (Wolfram) erstellt wurde.

Es wurde nachstehender Befehl verwendet, um somit ein Vektorfeld von $\mathbb{R}^2$ nach $\mathbb{R}^2$ darzustellen.

```
VektorPlot[x,y,x,-3,3,y,-3,3]
```

Vergleiche hierzu die Dokumentation (aufgerufen am 15.06.2019):

<https://reference.wolfram.com/language/howto/PlotAVectorField.html>

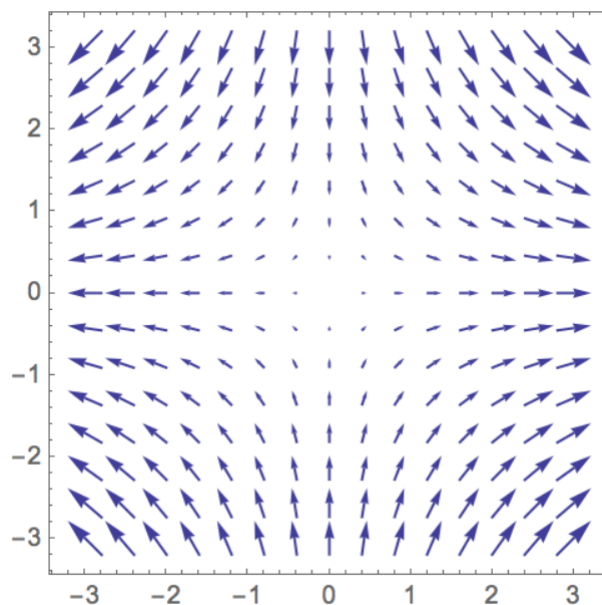


Abbildung B.5.: VektorPlot eines Vektorfeldes

1.1. Banachscher Fixpunktsatz

Der nun folgende *Banachsche Fixpunktsatz* kann dazu verwendet werden, um die Existenz und Eindeutigkeit eines Fixpunktes zu zeigen. Überdies ermöglicht er mittels Iteration, einem numerischen Verfahren, die näherungsweise Bestimmung von Fixpunkten.

Damit der Satz gültig ist und damit ein Fixpunkt mit $\varphi(\xi) = \xi$ gegeben ist, müssen eine Teilmenge des Banachraums (0.25, S.21) und eine kontrahierende Selbstabbildung (B.8, S. 60) vorliegen.

Dieser Abschnitt orientiert sich hierbei an [LA1], [LA2] und [GM2].

Zunächst werden die Voraussetzungen und der Begriff Fixpunkt definiert:

DEFINITION B.5: Fixpunkt

vgl- LA1, S.159

Sei $f : X \rightarrow X$ eine Funktion wobei X eine nichtleere Menge ist.
Ein Punkt $\xi \in X$ heißt *Fixpunkt von f* , wenn $f(\xi) = \xi$.

DEFINITION B.6: Selbstabbildung

vgl- LA1, S.106

Sei $f : X \rightarrow X$ wobei X eine nichtleere Menge ist.
Die Funktion f heißt *Selbstabbildung auf X* , wenn f die Menge X in sich selbst abbildet und daher $f(X) \subset X$ gilt.

Beispiel. Identische Abbildung, Identität

Die *Identische Abbildung* bildet jedes Element der Definitionsmenge auf sich selbst ab:

$$id_X : X \subseteq \mathbb{R} \rightarrow X \text{ ist definiert durch } id_X := x \quad \text{für alle } x \in X.$$

Beispiel. Das Zählen

Beim Zählen einer natürlichen Zahl ($n \in \mathbb{N}$) ihr Nachfolger zugeordnet:

$$f : \mathbb{N} \rightarrow \mathbb{N} \text{ ist definiert durch } f(x) := x + 1 \quad \text{für alle } x \in \mathbb{N}.$$

Bemerkung. Selbstabbildungseigenschaft monotoner Funktionen

Sei $f : [a, b] \rightarrow \mathbb{R}$ eine differenzierbare Funktion.

- $f'(x) \neq 0$ für alle $x \in [a, b]$
Sind nun die Randpunkte $f(a), f(b) \in [a, b]$, so ist die Funktion eine Selbstabbildung.
- $f(x) = 0$ für ein oder mehrere $x \in [a, b]$
Die Funktion hat ein oder mehrere Extrema, deren Funktionswerte neben denen der Randpunkte in $[a, b]$ liegen müssen, damit die Funktion selbstabbildend ist.

Kontrahierende Funktionen, sind eine besondere Form der *dehnungsbeschränkten Funktionen* mit einer Dehnungsschranke, die kleiner 1 ist:

DEFINITION B.7: Dehnungsbeschränkte Funktion

vgl. LA1, S.136

Sei $f : X \rightarrow \mathbb{R}$. Die Funktion f heißt *dehnungsbeschränkte Funktion* auf $X \subset \mathbb{R}$, wenn es eine Dehnungsschranke q gibt, sodass

$$|f(x) - f(y)| \leq q |x - y| \quad \text{für alle } x, y \in X$$

Da diese Art von Funktionen auch stetig ist, werden sie auch *Lipschitz-stetige Funktionen* genannt und dazu die Dehnungsschranke q als Lipschitz-Konstante bezeichnet.



Abbildung B.6.: *Rudolf Otto Sigismund Lipschitz, deutscher Mathematiker (1832 - 1903)*

Quelle (abgerufen am 22.08.2019): de.wikipedia.org/wiki/Rudolf_Lipschitz

Beispiel.

Sei $f : X \rightarrow \mathbb{R}$ mit $f(x) = x$. Die Funktion f ist Lipschitz-stetig mit $q = 1$, denn es gilt

$$|f(x) - f(y)| = |x - y| \leq 1|x - y| \quad \text{für alle } x, y \in X$$

B. Differentialrechnung im Mehrdimensionalen

Beispiel.

Sei $f : [a, b] \rightarrow \mathbb{R}$ mit $f(x) = x^2$ und $a, b \in \mathbb{R}$ mit $a < b$. Dann ist die Funktion f Lipschitz-stetig für alle $x, y \in [a, b]$, denn

$$\begin{aligned} |f(x) - f(y)| &= |x^2 - y^2| = |x + y||x - y| \leq (|x| + |y|) \cdot |x - y| \leq \\ &2 \max\{|x|, |y|\} \cdot |x - y| \leq 2 \max\{|a|, |b|\} \cdot |x - y| \end{aligned}$$

Als Lipschitz-Konstante kann $q = 2 \max\{|a|, |b|\}$ gewählt werden.

Beispiel.

Sei $f : \mathbb{R} \rightarrow [0, 1]$ mit $f(x) := \frac{x+2}{x+1}$. Die Lipschitz-Konstante ist zu berechnen.

Umformen ergibt: $f(x) = 1 + \frac{1}{x+1}$

Mittels der Definition (B.7, S. 59) folgt: $|f(x) - f(y)| = \left| \frac{1}{x+1} - \frac{1}{y+1} \right| = \frac{|x-y|}{|(x+1)(y+1)|} \leq q |x-y|$

und somit ist die Funktion *dehnungsbeschränkt* mit $q = \sup_{x, y \in [0, 1]} \frac{1}{|(x+1)(y+1)|} = 1$

Bemerkung.

Wäre im vorigen Beispiel das Intervall $[1, 2]$ die Zielmenge, dann wäre die Funktion *kontrahierend* mit $q = \frac{1}{2}$

DEFINITION B.8: Kontrahierende Funktion

vgl. LA1, S.221

Sei $f : X \rightarrow \mathbb{R}$.

Eine dehnungsbeschränkte Funktion mit einer Dehnungsschranke $q < 1$ wird *kontrahierend* genannt.

Bemerkung. Überprüfung der Kontraktionseigenschaft

Um zu überprüfen, ob es sich bei der differenzierbaren Funktion $f : [a, b] \rightarrow \mathbb{R}$ um eine kontrahierende Funktion handelt, muss lediglich gezeigt werden, dass die Dehnungsschranke $q := \max\{|f'(x)| : x \in [a, b]\} < 1$ ist.

Beispiel.

Sei $f : [0, 1] \rightarrow \mathbb{R}$ mit $f(x) = x^2$.

Die Funktion f ist selbstabbildend auf $[0, 1]$, da f steigend ist und

$$\min(f(x)) = f(0) = 0 \quad \text{und} \quad \max(f(x)) = f(1) = 1.$$

Mit Hilfe der Differentialgleichung kann die Kontraktionseigenschaft überprüft werden.
Ist die Funktion einmal stetig differenzierbar auf der Definitionsmenge und

$$|f'(x)| < 1 \quad \text{für alle } x \in [0, 1],$$

so handelt es sich um eine kontrahierende Funktion.

Somit ergibt sich für $f'(1) = 2$, weshalb die Funktion f keine kontrahierende Funktion ist.

Da nun alle Voraussetzungen bekannt sind, folgt der *Banachsche Fixpunktsatz* und die dazu gehörige Beweisführung:

SATZ B.1: Banachscher Fixpunktsatz

vgl. LA2, S.35

Sei X eine (nichtleere) abgeschlossene Teilmenge eines Banachraumes und f eine kontrahierende Selbstabbildung von X , für alle $x, y \in X$ gelte somit

$$\|f(x) - f(y)\| \leq q \|x - y\| \quad \text{mit einem festen positiven } q < 1,$$

dann besitzt f genau einen Fixpunkt x in X .

Für differenzierbare Funktionen gilt insbesondere:

$$|f'(x)| \leq q \quad \text{für } q < 1.$$

Das selbe kann iterativ gewonnen werden:

Wird ein beliebigen Startpunkt $x_0 \in X$ gewählt und $x_{n+1} := f(x_n)$ ($n = 0, 1, 2, \dots$) gesetzt, so konvergiert $x_n \rightarrow x$.

Überdies gilt die Fehlerabschätzung

$$\|x_n - x\| \leq \frac{q^n}{1 - q} \|x_1 - x_0\|$$

B. Differentialrechnung im Mehrdimensionalen

Die Beweisführung des Satzes unterteilt sich in den Nachweis der Eindeutigkeit und der Existenz, sowie der Konvergenz der Iteration:

Beweis.

Eindeutigkeit:

Seien $x, y \in X$ zwei Fixpunkte von f und $q < 1$. Dann gilt

$$\|x - y\| = \|f(x) - f(y)\| \leq q\|x - y\|$$

woraus folgt: $\|x - y\| = 0$, d.h. $x = y$.

Existenz:

Sei $x_0 \in X$ beliebig und $x_k := Ax_{k-1}$ für alle $k \geq 1$.

Es gilt $\|x_{k+1} - x_k\| = \|f(x_k) - f(x_{k-1})\| \leq q\|x_k - x_{k-1}\| \leq q^2\|x_{k-1} - x_{k-2}\| \leq \dots \leq q^k\|x_1 - x_0\| =: q^k c$.

Daraus folgt für $m > k$

$$\|x_m - x_k\| = \left\| \sum_{i=k}^{m-1} (x_{i+1} - x_i) \right\| \leq \sum_{i=k}^{m-1} \|x_{i+1} - x_i\| \leq \sum_{i=k}^{m-1} q^i c \leq \frac{q^k c}{1-q}$$

Dies zeigt, dass die Folge $(x_k)_{k \in \mathbb{N}}$ eine Cauchy-Folge in X ist und wegen der Vollständigkeit von X gegen ein Element x konvergiert. Weil X abgeschlossen ist, liegt x sogar in X .

$x_{k+1} = f(x_k)$ folgt durch Grenzübergang $k \rightarrow \infty$

$x = f(x)$

□

Beispiel.

Sei $f: [0, 1] \rightarrow \mathbb{R}$ mit $f(x) = \frac{x}{5}$.

Die Funktion f ist selbstabbildend auf $[0, 1]$, da die Funktion f steigend ist, $\min(f(x)) = 0$ und $\max(f(x)) = \frac{1}{5}$.

Hier ist für $f'(1) = \frac{1}{5} < 1$, weshalb die Funktion f eine kontrahierende Funktion ist.

Der Fixpunkt lässt sich durch Iteration finden:

Beliebiger Startpunkt $x_0 = \frac{1}{2}$ und somit $x_1 = f(x_0) = \frac{1}{10}$, $x_2 = \frac{1}{50}$ usw.

Der gesuchte Fixpunkt ist $x = 0$, da $f(0) = 0$.

2. Differenzierbarkeit in $\mathbb{R}^p$

Eine Funktion, die einen Vektor auf eine reelle Zahl (reellwertige Funktion) bzw. auf einen Vektor (vektorielle Funktion) zuordnet, kann nicht mehr wie in eindimensionalen Fällen abgeleitet werden. Hier kann nur jeweils nach einer Veränderlichen abgeleitet werden, indem die Ableitung einer Funktion mehrerer Veränderlicher mit Hilfe von *Richtungsableitungen* bzw. *partieller Ableitungen* durchgeführt wird.

2.1. Differenzierbarkeit in einem Punkt

DEFINITION B.9: Differenzierbarkeit in einem Punkt

vgl. LA2, S.259

Eine Funktion $f : I \rightarrow \mathbb{R}^q$ heißt *differenzierbar* im Punkt $\xi \in I$, wenn es eine (q, p) -Matrix A gibt, so dass für alle $\xi + h$ aus einer ϵ -Umgebung $U \subset G$ von ξ das Inkrement $f(\xi + h) - f(\xi)$ die folgende Darstellung gestattet

$$f(\xi + h) - f(\xi) = Ah + r(h) \quad \text{mit} \quad \lim_{h \rightarrow 0} \frac{r(h)}{\|h\|} = 0$$

2.2. Partielle Ableitung

Wie bereits zuvor erwähnt, sind *partielle Ableitungen* eine besondere Form der Richtungsableitungen, nämlich in Richtung der Koordinatenachsen. Partielle Ableitungen sind somit entsprechend Ableitungen eindimensionaler Funktionen nur nach einer Variable durchzuführen, wobei die jeweils anderen Variablen als Konstanten aufgefasst werden:

DEFINITION B.10: Partielle Ableitung

vgl. LA2, S.247ff

Sei X die offene Definitionsmenge der reellwertigen Funktionen f von mehreren Veränderlichen: $f : X \subset \mathbb{R}^p \rightarrow \mathbb{R}$.

Ist $(\xi_1, \dots, \xi_p)$ ein Punkt aus X , und besitzt die *partielle Funktion* $x_1 \mapsto f(x_1, \xi_1, \dots, \xi_p)$ in ξ_1 eine Ableitung, deren Wert α_1 ist, so ist die Funktion f in $(\xi_1, \dots, \xi_p)$ *partiell nach x_1 differenzierbar*. Ihre *partielle Ableitung nach x_1* im Punkt $(\xi_1, \dots, \xi_p)$ ist α_1 :

$$D_1 f(\xi_1, \dots, \xi_p) = \alpha_1 \quad \text{bzw.} \quad \frac{\partial f(\xi_1, \dots, \xi_p)}{\partial x_1} = \alpha_1$$

Partielle Ableitungen nach $x_2, \dots, x_p$ analog.

B. Differentialrechnung im Mehrdimensionalen

Beispiel.

Sei $f(x, y, z) = xy + z$ eine differenzierbare Funktion.

Es sind die partiellen Ableitungen zu bestimmen:

$$\frac{\partial f}{\partial x} = y \quad \frac{\partial f}{\partial y} = x \quad \frac{\partial f}{\partial z} = 1$$

DEFINITION B.11: Partielle Ableitung höherer Ordnung

vgl. LA2, S.248ff

Sei X die offene Definitionsmenge der reellwertigen Funktionen f von mehreren Veränderlichen: $f : X \subset \mathbb{R}^p \rightarrow \mathbb{R}$.

Existiert überdies die partielle Ableitung $D_k f$ an der Stelle $(\xi_1, \dots, \xi_p)$ und es sei möglich, an der Stelle $(\xi_1, \dots, \xi_p)$ nach x_j partiell zu differenzieren, so folgt für die *partielle Ableitung zweiter Ordnung* von f an der Stelle $(\xi_1, \dots, \xi_p)$:

$$D_j D_k f(\xi_1, \dots, \xi_p) = \frac{\partial^2 f(\xi_1, \dots, \xi_p)}{\partial x_j \partial x_k}$$

Partielle Ableitungen höherer Ordnung nach $x_2, \dots, x_p$ analog.

Jacobi-Matrix

SATZ B.2: Die Jacobi-Matrix

vgl. LA2; S.260

Die Funktion $f : G \subset \mathbb{R}^p \rightarrow \mathbb{R}^q$ (G offen) sei im Punkt ξ differenzierbar und habe die Komponenten $f_1, \dots, f_q$. Es sei somit

$$f = \begin{pmatrix} f_1 \\ \vdots \\ f_q \end{pmatrix} \quad \text{mit Funktionen } f_j : G \rightarrow \mathbb{R}$$

Dann ist jedes f_j an der Stelle ξ nach allen p Veränderlichen $x_1, \dots, x_p$ partiell differenzierbar, und die (q, p) -Matrix A aus Definition (B.9, S. 63) ist eindeutig bestimmt. A

ist die sogenannte *Jacobi-Matrix* (*Funktionalmatrix*) von f an der Stelle ξ mit:

$$J_f(\xi) := \begin{pmatrix} \frac{\partial f_1(\xi)}{\partial x_1} & \cdots & \frac{\partial f_1(\xi)}{\partial x_p} \\ \vdots & & \vdots \\ \frac{\partial f_q(\xi)}{\partial x_1} & \cdots & \frac{\partial f_q(\xi)}{\partial x_p} \end{pmatrix}$$

Bemerkung.

Die Jacobi-Matrix J_f wird auch *Ableitung* von f an der Stelle ξ genannt und oftmals mit

$$f'(\xi) \quad \text{bzw.} \quad Df(\xi) \quad \text{bezeichnet.}$$

Beweis.

Da f im Punkt ξ differenzierbar ist, gilt laut Definition $f(\xi + h) - f(\xi) = Ah + r(h)$ mit $\lim_{h \rightarrow 0} \frac{r(h)}{\|h\|} = 0$ und

$$\xi = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_p \end{pmatrix}, \quad h = \begin{pmatrix} h_1 \\ \vdots \\ h_p \end{pmatrix}, \quad A = \begin{pmatrix} a_{11} & \cdots & a_{1p} \\ \vdots & & \vdots \\ a_{q1} & \cdots & a_{qp} \end{pmatrix}, \quad r(h) = \begin{pmatrix} r_1(h) \\ \vdots \\ r_q(h) \end{pmatrix}$$

Dann ist für $j = 1, \dots, q$

$$f_j(\xi + h) - f_j(\xi) = a_{j1}h_1 + \dots + a_{jk}h_k + \dots + a_{jp}h_p + r_j(h) \quad \text{mit} \quad \lim_{h \rightarrow 0} \frac{r_j(h)}{\|h\|} = 0.$$

Nun wird die j -te Komponente von $f(\xi + h)$ betrachtet:

$$f_j(\xi + h) = f_j(\xi) + \sum_{k=1}^p a_{jk}h_k + r_j(h)$$

Wird nun h mit dem Einheitsvektor e_k multipliziert, sodass $h_k e_k$, dann ist nur die j -te Komponente ungleich Null und es ergibt sich: $f_j(\xi + h_k e_k) = f_j(\xi) + a_{jk}h_k + r_j(h_k e_k)$

Mit $\lim_{h \rightarrow 0} \frac{r(h)}{h} = 0$ lässt sich nun die partielle Ableitung der j -ten Komponente berechnen:

$$\frac{\partial f_j}{\partial x_k}(\xi) = \lim_{h \rightarrow 0} \frac{f_j(\xi + h_k e_k) - f_j(\xi)}{h_k} = a_{jk} + \lim_{h \rightarrow 0} \frac{r_j(h_k e_k)}{h_k} = a_{jk} \quad \square$$

Gradient

Der Spaltenvektor, der alle partiellen Ableitungen der Funktion enthält, wird *Gradient* genannt:

DEFINITION B.12: Gradient

vgl. LA2, S.273

Sei f eine reellwertige Funktion mit $f; G \subset \mathbb{R}^p \rightarrow \mathbb{R}$ (G offen).
Weiters besitze f an der Stelle ξ partielle Ableitungen nach allen ihren Veränderlichen $x_1, \dots, x_p$ (ohne notwendigerweise in ξ differenzierbar zu sein), dann heißt folgender Vektor der *Gradient* von f an der Stelle ξ :

$$\text{grad } f(\xi) = \begin{pmatrix} \frac{\partial f(\xi)}{\partial x_1} \\ \vdots \\ \frac{\partial f(\xi)}{\partial x_p} \end{pmatrix}$$

Bemerkung.

Werden kartesische Koordinate verwendet, so bezeichnet ∇f den Gradienten.

Beispiel.

Sei $f(x, y, z) = xy + z$ eine differenzierbare Funktion.

Es ist der Gradient zu bestimmen:

$$\text{grad } f = \nabla f = \begin{pmatrix} y \\ x \\ 1 \end{pmatrix}$$

2.3. Mittelwertsatz (verallgem.)

SATZ B.3: Mittelwertsatz für vektorwertige Funktionen

vgl. LA2, S.278ff

Die Funktion $f : G \subset \mathbb{R}^p \rightarrow \mathbb{R}^q$ (G offen) sei stetig differenzierbar, und $x_0, x_0 + h$ seien Punkte, die mitsamt ihrer Verbindungsstrecke S in G liegen. Dann gilt

$$f(x_0 + h) - f(x_0) = \left(\int_0^1 f'(x_0 + th) dx \right) h$$

Hierbei ist das Integral in der Klammer eine Matrix, die auf h angewendet werden soll.

Ist $|\cdot|$ irgend eine mit den Normen von $\mathbb{R}^p$ und $\mathbb{R}^q$ verträgliche Matrixnorm auf $\mathfrak{M}(q, p)$, dann kann folgendermaßen abgeschätzt werden:

$$\|f(x_0 + h) - f(x_0)\| \leq M \|h\| \quad M := \max_{x \in S} |f'(x)|(*)$$

Beweis.

Sei $f = \begin{pmatrix} f_1 \\ \vdots \\ f_q \end{pmatrix}$, $\varphi_j(t) := f_j(x_0 + th)$ für $0 \leq t \leq 1$ und $h = \begin{pmatrix} h_1 \\ \vdots \\ h_p \end{pmatrix}$.

Da es sich hier um eine C^1 -Funktion handelt, sind die partiellen Ableitungen aller f_j auf G stetig, und es ist

$$f_j(x_0 + h) - f_j(x_0) = \varphi_j(1) - \varphi_j(0) = \int_0^1 \varphi_j'(t) dt = \int_0^1 f_j'(x_0 + th) h dt.$$

Infolgedessen gilt

$$\begin{aligned} f(x_0 + h) - f(x_0) &= \begin{pmatrix} \int_0^1 f_1'(x_0 + th) h dt \\ \vdots \\ \int_0^1 f_q'(x_0 + th) h dt \end{pmatrix} = \int_0^1 \begin{pmatrix} f_1'(x_0 + th) h dt \\ \vdots \\ f_q'(x_0 + th) h dt \end{pmatrix} = \\ &= \int_0^1 f'(x_0 + th) h dt = \left(\int_0^1 f'(x_0 + th) dt \right) h. \end{aligned}$$

Da h ein fester Vektor ist, kann dieser hier aus den Riemannschen Summen des Integrals herausgezogen werden.

B. Differentialrechnung im Mehrdimensionalen

$\max_{x \in S} |f'(x)|$ existiert, da $x \mapsto |f'(x)|$ eine reellwertige Funktion auf der kompakten Menge S ist. Da die Matrixnorm $\|\cdot\|$ mit den Normen von $\mathbb{R}^p$ und $\mathbb{R}^q$ verträglich sein sollte, ergibt sich mit der Gleichung (*) in Verbindung mit der Dreiecksungleichung die Abschätzung

$$\begin{aligned} \|f(x_0 + h) - f(x_0)\| &= \left\| \left(\int_0^1 f'(x_0 + th) dt \right) h \right\| \leq \left\| \int_0^1 f'(x_0 + th) dt \right\| \|h\| \leq \\ &\leq \left(\int_0^1 |f'(x_0 + th)| dt \right) \|h\| \leq M \|h\| \end{aligned}$$

□

2.4. Differentiationsregeln für Mehrdimensionale Funktionen

Faktorregel

vgl. LA2, S.266

Sei $f : G \subset \mathbb{R}^p \rightarrow \mathbb{R}^q$ (G offen) eine in $\xi \in G$ differenzierbare Funktion und $\lambda \in \mathbb{R}$, dann ist $\lambda f : G \subset \mathbb{R}^p \rightarrow \mathbb{R}^q$ differenzierbar und es gilt:

$$(\lambda f)'(\xi) = \lambda f'(\xi)$$

Beweis.

Nach Voraussetzung ist $f(\xi + h) - f(\xi) = f'(\xi)h + r(h)$ mit $\lim_{h \rightarrow 0} \frac{r(h)}{\|h\|} = 0$

daher gilt $\lambda f(\xi + h) - \lambda f(\xi) = \lambda (f'(\xi)h + r(h))$

da $\frac{r(h)}{\|h\|} \rightarrow 0$ strebt für $h \rightarrow 0$ existiert die Ableitung und $= \lambda f'(\xi)$ □

Summenregel

LA2, S.266

Seien $f, g : G \subset \mathbb{R}^p \rightarrow \mathbb{R}^q$ (G offen) in $\xi \in G$ differenzierbare Funktionen, dann ist $f + g : G \subset \mathbb{R}^p \rightarrow \mathbb{R}^q$ differenzierbar und es gilt:

$$(f+g)'(\xi) = f'(\xi) + g'(\xi)$$

Beweis.

Nach Voraussetzung ist

$$f(\xi + h) - f(\xi) = f'(\xi)h + r_1(h) \quad \text{mit} \quad \lim_{h \rightarrow 0} \frac{r_1(h)}{\|h\|} = 0,$$

$$g(\xi + h) - g(\xi) = g'(\xi)h + r_2(h) \quad \text{mit} \quad \lim_{h \rightarrow 0} \frac{r_2(h)}{\|h\|} = 0$$

Somit folgt

$$(f + g)(\xi + h) - (f + g)(\xi) = [f'(\xi) + g'(\xi)]h + r(h), \quad r(h) = r_1(h) + r_2(h)$$

B. Differentialrechnung im Mehrdimensionalen

Und da $r(h)/\|h\| \rightarrow 0$ strebt für $h \rightarrow 0$, ist die Ableitung $(f + g)(\xi)$ tatsächlich vorhanden und $= f'(\xi) + g'(\xi)$. $\square$

Produktregel

LA2: S.269

Seien $f, g : G \subset \mathbb{R}^p \rightarrow \mathbb{R}$ (G offen) in $\xi \in G$ differenzierbare Funktionen, dann ist fg in ξ differenzierbar und es gilt:

$$(fg)'(\xi) = f'(\xi)g(\xi) + f(\xi)g'(\xi)$$

Beweis.

Für hinreichend kleine $\|h\| \neq 0$ ist

$$\begin{aligned}(fg)(\xi + h) - (fg)(\xi) &= f(\xi + h)g(\xi + h) - f(\xi)g(\xi) \\ &= [f(\xi) + f'(\xi)h + r_1(h)][g(\xi) + g'(\xi)h + r_2(h)] - f(\xi)g(\xi) \\ &= [f(\xi)g'(\xi) + g(\xi)f'(\xi)]h + r(h)\end{aligned}$$

$$\begin{aligned}\text{mit } r(h) &:= [f(\xi) + f'(\xi)h] r_2(h) + [g(\xi) + g'(\xi)h] r_1(h) \\ &\quad + r_1(h) r_2(h) + [f'(\xi)h] [g'(\xi)h]\end{aligned}$$

Jeder Summand auf der rechten Seite strebt bei Division durch $\|h\|$ gegen 0, wenn $h \rightarrow 0$ geht.

Wird folgende Gleichung betrachtet, wird das beim letzten Summanden ebenso klar ersichtlich:

$$\frac{[f'(\xi)h][g'(\xi)h]}{\|h\|} = g'(\xi) \left[(f'(\xi)h) \frac{h}{\|h\|} \right]$$

Infolgedessen strebt $r(h)/\|h\| \rightarrow 0$ für $h \rightarrow 0$, womit die Behauptung über das Produkt bewiesen ist. $\square$

Quotientenregel

LA2; S.269

Seien $f, g : G \subset \mathbb{R}^p \rightarrow \mathbb{R}$ (G offen) in $\xi \in G$ differenzierbare Funktionen, und ist $g(\xi) \neq 0$ für alle $\xi \in G$, so ist auch die Funktion (f/g) an der Stelle ξ differenzierbar und es gilt:

$$\left(\frac{f}{g}\right)'(x) = \frac{f'(x)g(x) - f(x)g'(x)}{g(x)^2}$$

Beweis.

Mit Hilfe der Produktregel (2.4, S. 70) lässt sich die Quotientenregel herleiten. Dazu wird der Quotient $\frac{f}{g}$ zu $f \cdot g^{-1}$ umgeschrieben und dann der Beweis ähnlich dem der Produktregel durchgeführt:

$$\begin{aligned} \left(\frac{f}{g}\right)'(x) &= (fg^{-1})'(x) = f'(x)\frac{1}{g(x)} + f(x)\left(\frac{1}{g(x)}\right)' \\ &= \frac{f'(x)}{g(x)} + \frac{f(x)(-g'(x))}{g^2(x)} = \frac{f'(x)g(x) - f(x)g'(x)}{g^2(x)} \end{aligned}$$

□

Kettenregel

vgl. LA2, S.267

Vorgelegt seien die Funktionen

$$\begin{aligned} g : G \subset \mathbb{R}^p &\rightarrow \mathbb{R}^q \text{ mit offenem } G \\ f : U \subset \mathbb{R}^q &\rightarrow \mathbb{R}^r \text{ mit offenem } U \supset g(G) \end{aligned}$$

g sei in ξ , f in $g(\xi)$ differenzierbar. Dann ist $f \circ g$ in ξ differenzierbar, und es gilt

$$(f \circ g)'(\xi) = f'(g(\xi)) g'(\xi)$$

Bemerkung.

Der Beweis der Kettenregel im Merdimensionalen erfolgt analog zu dem aus der eindimensionalen Analysis, jedoch unter Verwendung von $\|h\|$ und $\|k\|$.

2.5. Richtungsableitung

DEFINITION B.13: Richtungsableitung

vgl. LA2, S.272ff

Sei $f(x, y)$ eine reellwertige, mehrdimensionale Funktion auf einer Menge $G \subseteq \mathbb{R}^p$ und $|v| = 1$ (*Richtungsvektor mit Betrag 1*).

Wenn der Grenzwert $\frac{\partial f(\xi)}{\partial v} := \lim_{t \rightarrow 0} \frac{f(\xi + tv) - f(\xi)}{t}$ mit $\xi \in G$ existiert, so bezeichnet dieser die *Richtungsableitung* von f im Punkt ξ in Richtung v .

Bemerkung.

Weitere Schreibweisen für $\frac{\partial f(\xi)}{\partial v}$ sind: $D_v f(\xi)$ bzw. $\frac{\partial f}{\partial v}(\xi)$

Bemerkung.

Werden Richtungsableitungen in Richtung der Koordinatenachsen gebildet, so werden diese *partielle Ableitungen* genannt (siehe B.10, S. 63).

Richtungsableitungen von $f(\xi)$ für den Richtungsvektor v werden oft auch mit $\partial_v f(\xi)$ bezeichnet.

Beispiel.

Sei $f : [0, 5] \times [0, 5] \rightarrow \mathbb{R}$ mit $f(x, y) = x + y$.

Unter der Annahme, dass sie existiert, ist die Richtungsableitung im Punkt $P(1, 2)$

in Richtung $v = \begin{pmatrix} 3 \\ 4 \end{pmatrix}$ zu bestimmen:

$$\frac{\partial f(P)}{\partial v} = \lim_{t \rightarrow 0} \frac{f((1, 2) + t(3, 4)) - f(1, 2)}{t} = \lim_{t \rightarrow 0} \frac{(1 + 3t) + (2 + 4t) - (1 + 2)}{t} = \lim_{t \rightarrow 0} \frac{7t}{t} = 7$$

Die Ableitung der Funktion im Punkt P in Richtung v ist 7.

SATZ B.4: Existenz der Richtungsableitung

vgl. LA2, S.272ff

Die Funktion $f : G \subset \mathbb{R}^p \rightarrow \mathbb{R}$ (G offen) sei in $\xi \in G$ differenzierbar. Dann existiert die Richtungsableitung $\frac{\partial f(\xi)}{\partial v}$ für jeden Richtungsvektor v , und es ist

$$\frac{\partial f(\xi)}{\partial v} = f'(\xi)v = \frac{\partial f(\xi)}{\partial x_1}v_1 + \dots + \frac{\partial f(\xi)}{\partial x_p}v_p \quad \text{für } v := \begin{pmatrix} v_1 \\ \vdots \\ v_p \end{pmatrix}.$$

Beweis.

Da G offen ist gibt es ein $\delta > 0$ mit $U_\delta(\xi) \subseteq G$. Sei nun $u : (-\delta, \delta) \rightarrow G$ durch $u(t) = \xi + tv$.

Dann ist nach der Kettenregel (2.4, S. 71) $(f \circ u)'(0) = f'(\xi)v$.

Andererseits ist $\frac{\partial f(\xi)}{\partial v} = \lim_{t \rightarrow 0} \frac{f(\xi + tv) - f(\xi)}{t} = (f \circ u)'(0) = f'(\xi)v$. □

3. Extremwerte

Lokale und *globale Extrema* sind auch bei mehrdimensionalen Funktionen Punkte, die es wert sind, untersucht zu werden.

3.1. lokale Extrema

Bei der Suche nach einem lokalen Extremum einer mehrdimensionalen reellwertigen Funktion, ist es sinnvoll, so wie schon beim eindimensionalen Fall (siehe Abschnitt A.11, S. 40), notwendige und hinreichende Bedingungen für Extremwerte zu definieren. Dabei ist es nicht zwingend, dass ein solches Extremum auch existiert:

DEFINITION B.14: Lokale Extrema

vgl. ABG; S.318ff

Sei $X \subset \mathbb{R}^p$ und $f : X \rightarrow \mathbb{R}$. Dann heißt $(x_0, y_0) \in X$ *lokales Maximum* (bzw. *lokales Minimum*) von f , falls ein $\delta > 0$ existiert, sodass $f(x, y) \leq f(x_0, y_0)$ (bzw. $f(x, y) \geq f(x_0, y_0)$) für alle $(x, y) \in U_\delta(x_0, y_0)$.

Bemerkung.

Ein lokales Maximum (bzw. lokales Minimum) von f liegt dann vor, wenn eine Umgebung um (x_0, y_0) existiert, in der kein Punkt einen größeren (bzw. kleineren) Funktionswert annimmt.

Notwendige Bedingung

Analog zum eindimensionalen Fall, bei dem die erste Ableitung der Funktion $=0$ als *notwendige Bedingung* (A.7, S. 37) für das Vorliegen eines lokalen Extremums sein muss, wird im mehrdimensionalen der *Gradient* (B.12, S.66) der Funktion untersucht:

SATZ B.5: Notwendige Bedingung für lokale Extrema

vgl. LA2, S.311

Die Funktion $f : X \subset \mathbb{R}^p \rightarrow \mathbb{R}$ besitze in dem inneren Punkt ξ von X ein *lokales Extremum* und sei überdies in ξ nach allen Veränderlichen partiell differenzierbar. Dann ist notwendig $\text{grad } f(\xi) = 0$.
Ist f in ξ sogar differenzierbar, so muss $f'(\xi) = 0$ sein.

Beweis.

Ist $\xi = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_p \end{pmatrix}$ und $g_1(x_1) := f(x_1, \xi_2, \dots, \xi_p); \dots; g_p(x_p) := f(x_1, \xi_2, \dots, \xi_p)$,

so besitzt die partielle Funktion g_k im Punkt ξ_k ein lokales Extremum. Und da ξ_k ein innerer Punkt des Definitionsbereichs von g_k und die Ableitung $g'_k(\xi_k)$ vorhanden und $= \partial f(\xi)/\partial x_k$ ist, ergibt sich nun die Behauptung ohne weitere Umstände aus Satz (A.7, S. 37). $\square$

Hinreichende Bedingung

Als hinreichende Bedingung für das Vorliegen eines lokalen Extremums einer mehrdimensionalen reellwertigen Funktion wird die sogenannte *Hesse-Matrix* auf ihre *Definitheit* untersucht:

DEFINITION B.15: Hesse-Matrix

vgl. LA2, S.312

$$H_f(x) := \begin{pmatrix} \frac{\partial^2 f(x)}{\partial x_1^2} & \frac{\partial^2 f(x)}{\partial x_2 \partial x_1} & \cdots & \frac{\partial^2 f(x)}{\partial x_p \partial x_1} \\ \frac{\partial^2 f(x)}{\partial x_1 \partial x_2} & \frac{\partial^2 f(x)}{\partial x_2^2} & \cdots & \frac{\partial^2 f(x)}{\partial x_p \partial x_2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f(x)}{\partial x_1 \partial x_p} & \frac{\partial^2 f(x)}{\partial x_2 \partial x_p} & \cdots & \frac{\partial^2 f(x)}{\partial x_p^2} \end{pmatrix}$$



Abbildung B.7.: *Ludwig Otto Hesse, deutscher Mathematiker (1811 - 1874)*

Quelle (abgerufen am 22.08.2019): de.wikipedia.org/wiki/Otto_Hesse

DEFINITION B.16: Definitheit einer Matrix

vgl. A12, S.207

Sei A eine symmetrische (p, p) -Matrix. Die Matrix heißt

1. positiv definit, wenn $x^t Ax > 0$,
2. negativ definit, wenn $x^t Ax < 0$

für alle p -zeiligen Spaltenvektoren mit $x \in \mathbb{R}^p$ und $x \neq 0$. x^t ist der Zeilenvektor, der durch transponieren (somit werden Spalten Zeilen) aus dem Spaltenvektor x entsteht.

Damit nun die *Definitheit* einer Matrix mit Hilfe dem *Kriterium von Sylvester* untersucht werden kann, muss eine *symmetrische Matrix* vorliegen:

DEFINITION B.17: Symmetrische Matrix

vgl. LA2, S.308

Eine quadratische Matrix (a_{jk}) heißt *symmetrisch*, wenn $a_{jk} = a_{kj}$ für alle j, k gilt.

Folgendes Kriterium bietet eine Möglichkeit, eine Matrix auf ihre *Definitheit* zu untersuchen:

SATZ B.6: Kriterium von Sylvester

vgl. LA2, S.309

Sei $A := (a_{jk})$ eine symmetrische (p, p) -Matrix und

$$\Delta_k := \begin{vmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & & \vdots \\ a_{k1} & \cdots & a_{kk} \end{vmatrix}$$

ihr k -ter Hauptminor ($k = 1, \dots, p$). Dann gilt:

- A ist *positiv definit* $\Leftrightarrow$ alle Δ_k sind positiv
- A ist *negativ definit* $\Leftrightarrow$ die Vorzeichen der Δ_k alternieren, beginnend mit negativem Δ_1
- A ist die Nullmatrix (bei linearen Funktionen) $\Leftrightarrow$ alle Δ_k sind gleich Null

Bemerkung.

Eine symmetrische Matrix ist genau dann positiv definit, wenn ihre führenden Hauptminoren positiv sind.



Abbildung B.8.: *James Joseph Sylvester, britischer Mathematiker (1814 - 1897)*

Quelle (abgerufen am 22.08.2019): de.wikipedia.org/wiki/James_Joseph_Sylvester

SATZ B.7: Hinreichende Bedingungen für lokale Extrema

vgl. GM2; S.95

Sei $U \subset \mathbb{R}^q$ offen, $f : U \rightarrow \mathbb{R}$ eine zweimal stetig differenzierbare Funktion und $x \in U$ ein Punkt mit $\text{grad } f(x) = 0$.

1. Ist $H_f(x)$ positiv definit, so besitzt f in x ein striktes lokales Minimum.
2. Ist $H_f(x)$ negativ definit, so besitzt f in x ein striktes lokales Maximum.
3. Ist $H_f(x)$ indefinit, so besitzt f in x kein lokales Extremum.

Beweis.

vgl. LA2; S.313

Sei $Q : \mathbb{R}^p \rightarrow \mathbb{R}$ die quadratische Form der Hessischen Matrix $H_f(\xi)$. da $f'(\xi)$ verschwindet, ist für alle hinreichend kleinen $h \in \mathbb{R}^p$

$$\frac{f(\xi + h) - f(\xi)}{\|h\|} = \frac{1}{2}Q\left(\frac{h}{\|h\|}\right) + \rho(h) \quad \text{mit} \quad \lim_{h \rightarrow 0} \rho(h) = 0.$$

Ist nun $H_f(\xi)$ positiv definit, so gibt es ein $\alpha > 0$ mit $Q(h/\|h\|) \geq \alpha$ für alle

B. Differentialrechnung im Mehrdimensionalen

$h \neq 0$. Zu diesem α existiert wegen $\lim_{h \rightarrow 0} \rho(h) = 0$ ein $\delta > 0$, so dass für $\|h\| < \delta$ stets $|\rho| < \frac{\alpha}{2}$ bleibt. δ darf so klein gewählt werden, dass die δ -Umgebung U von ξ noch ganz in G liegt. Da jedes $x \in U$ in der Form $\xi + h$ mit $\|h\| < \delta$ geschrieben werden kann, ergibt sich aus obiger Gleichung sofort, dass für alle $x \in U$ stets $f(x) - f(\xi) > 0$ ist. $f(\xi)$ erweist sich so als lokales Minimum in engeren Sinne von f .

Ganz entsprechend ergibt sich, dass $f(\xi)$ ein lokales Maximum ist, wenn $H_f(\xi)$ negativ definit ist.

Nun sei $H_f(\xi)$ indefinit. Dann gibt es von 0 verschiedene Vektoren h_0, k_0 , sodass

$$Q(h_0) > 0 \quad \text{und} \quad Q(k_0) < 0$$

ist. Für alle Zahlen $\lambda \neq 0$ gilt somit

$$Q\left(\frac{\lambda h_0}{\|\lambda h_0\|}\right) = Q\left(\frac{\lambda h_0}{|\lambda| \|h_0\|}\right) = \frac{\lambda^2}{\lambda^2 \|h_0\|^2} Q(h_0) = \frac{Q(h_0)}{\|h_0\|^2} =: \varepsilon_1 > 0$$

und ganz entsprechend

$$Q\left(\frac{\lambda k_0}{\|\lambda k_0\|}\right) = \frac{Q(k_0)}{\|k_0\|^2} =: \varepsilon_2 < 0.$$

Wird in obiger Gleichung nun h durch λh_0 bzw. durch λk_0 ersetzt, so ergibt sich für hinreichend kleines $\lambda \neq 0$, dass $f(\xi + \lambda h_0) > f(\xi)$, aber $f(\xi + \lambda k_0) < f(\xi)$. $f(\xi)$ kann daher kein lokales Extremum sein. $\square$

3.2. Beispiele

Beispiel.

Es ist zu untersuchen, ob die folgende Funktion lokale Extrema besitzt:

$$f(x, y) := -x^2 + x + xy - y^2 + 1$$

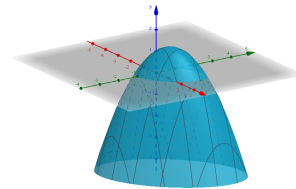


Abbildung B.9.: Extrema-BSP-1

1.Schritt: Die Funktion f partiell ableiten

$$\frac{\partial f(x, y)}{\partial x} = -2x + y + 1$$

$$\frac{\partial f(x, y)}{\partial y} = -2y + x + 1$$

2.Schritt: Partielle Ableitungen gleich Null setzen und das Gleichungssystem lösen

$$x = 1$$

$$y = 1$$

und somit den stationären Punkt $P(1, 1)$

3.Schritt: Hesse-Matrix bestimmen

Dazu werden die 2. partiellen Ableitungen gebildet:

$$\frac{\partial^2 f(x, y)}{\partial x^2} = -2$$

$$\frac{\partial^2 f(x, y)}{\partial x \partial y} = 1$$

$$\frac{\partial^2 f(x, y)}{\partial y \partial x} = 1$$

$$\frac{\partial^2 f(x, y)}{\partial y^2} = -2$$

Daraus ergibt sich die folgende Hesse-Matrix:

$$H_f(x, y) = \begin{pmatrix} -2 & 1 \\ 1 & -2 \end{pmatrix}$$

4.Schritt: Definitheit der Hesse-Matrix bestimmen

Laut dem Kriterium von Sylvester (vgl. B.6, S. 76) ist die Matrix negativ definit, da der erste Hauptminor negativ und der zweite positiv ist: $\Delta_1 = -2 < 0$, $\Delta_2 = 3 > 0$

Die Funktion besitzt daher ein lokales Maximum in $(1, 1)$.

B. Differentialrechnung im Mehrdimensionalen

Beispiel.

Es ist zu untersuchen, ob die folgende Funktion lokale Extrema besitzt:

$$f(x, y) := x^2 + y^2$$

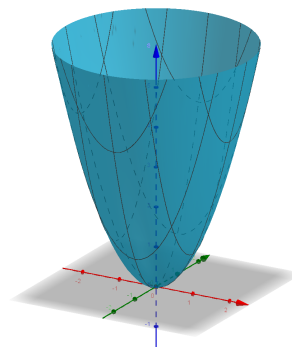


Abbildung B.10.: Extrema-BSP-2

1.Schritt: Die Funktion f partiell ableiten

$$\frac{\partial f(x, y)}{\partial x} = 2x$$

$$\frac{\partial f(x, y)}{\partial y} = 2y$$

2.Schritt: Partielle Ableitungen gleich Null setzen und das Gleichungssystem lösen

$$x = 0$$

$$y = 0$$

und somit den stationären Punkt $P(0, 0)$

3.Schritt: Hesse-Matrix bestimmen

Dazu werden die 2. partiellen Ableitungen gebildet:

$$\frac{\partial^2 f(x, y)}{\partial x^2} = 2$$

$$\frac{\partial^2 f(x, y)}{\partial x \partial y} = 0$$

$$\frac{\partial^2 f(x, y)}{\partial y \partial x} = 0$$

$$\frac{\partial^2 f(x, y)}{\partial y^2} = 2$$

Daraus ergibt sich die folgende Hesse-Matrix:

$$H_f(x, y) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

4.Schritt: Definitheit der Hesse-Matrix bestimmen

Die Matrix ist positiv definit, da die Hauptminoren positiv sind: $\Delta_1 = 2 > 0$, $\Delta_2 = 4 > 0$
Die Funktion besitzt daher ein lokales Minimum in $(0, 0)$.

4. Hauptsatz über Implizite Funktionen

Der *Hauptsatz über implizite Funktionen* dient dazu, festzustellen, ob und wo ein Gleichungssystem nach einer Veränderlichen eindeutig auflösbar ist:

SATZ B.8: Hauptsatz über Implizite Funktionen

vgl. LA2, S.292ff

Die Mengen $G \subset \mathbb{R}^p$ und $H \subset \mathbb{R}^q$ seien nichtleer und offen und die Funktion $F : G \times H \rightarrow \mathbb{R}^q$ sei stetig differenzierbar. Ferner seien $\xi \in G$ und $\eta \in H$ Punkte, für die $F(\xi, \eta) = 0$ und $\frac{\partial F(\xi, \eta)}{\partial y}$ invertierbar ist.

Dann gibt es eine δ -Umgebung $U = U_\delta(\xi) \subset G$, eine ε -Umgebung $V = V_\varepsilon(\eta) \subset H$ und genau eine stetig differenzierbare Funktion $f : U \rightarrow V$ mit $f(\xi) = \eta$ und $F(x, f(x)) = 0$ für alle $x \in U$.

Bemerkung.

Für jedes feste $x \in U$ ist $f(x)$ sogar die einzige in V liegende Lösung der Gleichung $F(x, y) = 0$.

Bemerkung.

Im Fall $p = q = 1$ ist $\frac{\partial F}{\partial y}$ die "gewöhnliche" partielle Ableitung.

Bemerkung.

Hierbei sind folgende Bezeichnungen verwendet worden:

$$x := \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix}, \quad y := \begin{pmatrix} y_1 \\ \vdots \\ y_q \end{pmatrix}, \quad G \times H = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} : x \in G, y \in H \right\} \subset \mathbb{R}^{p+q} \quad \text{mit } G \subset \mathbb{R}^p, H \subset \mathbb{R}^q$$

$$F \text{ ist eine } \mathbb{R}^q\text{-wertige Funktion der Form } F = \begin{pmatrix} F_1 \\ \vdots \\ F_q \end{pmatrix} \text{ mit reellwertigen Funktionen } F_1, \dots, F_q$$

B. Differentialrechnung im Mehrdimensionalen

Beweis.

Setze $D := \frac{\partial F(\xi, \eta)}{\partial y}$.

Falls eine Funktion f mit $f : U \rightarrow V$ mit $f(\xi) = \eta$ existiert, verschwindet nicht nur $F(x, f(x))$, sondern auch $D^{-1}F(x, f(x))$ auf U , und damit ist trivialerweise $f(x) - D^{-1}F(x, f(x)) = f(x)$ für alle $x \in U$.

Genügt umgekehrt eine Funktion f dieser Gleichung - einer *Fixpunktgleichung* -, so ist für alle einschlägigen x auch $F(x, f(x)) = 0$.

Daher liegt es nahe, die Existenzaussage des Satzes wie folgt durchzuführen:

Es werden zunächst Fixpunkte der Abbildung A gesucht, die auf einer noch geeigneten festzulegenden Menge M von $\mathbb{R}^q$ -wertigen Funktionen der Veränderlichen $x \in \mathbb{R}^p$ durch

$(Ag)(x) := g(x) - D^{-1}F(x, g(x))$ definiert wird.

Da F stetig differenzierbar ist, sind F und alle in $\partial F/\partial y$ auftretenden partiellen Ableitungen stetig, ferner ist gewiss $D^{-1} \frac{\partial F(\xi, \eta)}{\partial y} = I := (q, q)$ -Einheitsmatrix.

Nun ist ersichtlich, dass es aus Stetigkeitsgründen eine δ -Umgebung $U \subset G$ von ξ und eine ϵ -Umgebung von $V \subset H$ von η mit folgenden Eigenschaften gibt:

$$\begin{aligned} \|I - D^{-1} \frac{\partial F(x, y)}{\partial y}\| &\leq \frac{1}{2} \quad \text{für } x \in U, y \in V, \\ \|D^{-1} F(x, \eta)\| &\leq \frac{1}{4} \epsilon \quad \text{für } x \in U; \end{aligned} \tag{B.1}$$

dabei ist die in der ersten Abschätzung auftretende Norm die Abbildungsnorm auf $\mathfrak{M}(q, q)$. Nachdem die Umgebungen U, V festgelegt sind, wird die Menge M erklärt: Sie bestehe aus allen stetigen Funktionen $g : U \rightarrow \mathbb{R}^q$ mit

$$g(\xi) = \eta \quad \text{und} \quad \|g(x) - \eta\| \leq \frac{1}{2} \epsilon \quad \text{für jedes } x \in U. \tag{B.2}$$

Die Werte eines jeden $g \in M$ liegen alle in V .

M ist nicht leer, da die konstante Funktion $\eta : x \mapsto \eta$ ($x \in U$) zu M gehört.

Wegen $g(U) \subset V$ ist jedes $g \in M$ auf U beschränkt, liegt in dem Banachraum $B(U, \mathbb{R}^q)$, der mit der Supremumsnorm $|\varphi| := \sup_{x \in U} \|\varphi(x)\|$ versehen ist.

Darüber hinaus ist M eine abgeschlossene Teilmenge von $B(U, \mathbb{R}^q)$.

Für jedes $g \in M$ ist

$$Ag \quad \text{stetig und} \quad (Ag)(\xi) = \eta. \tag{B.3}$$

Ferner gilt wegen der Abschätzung $\|D^{-1}F(x, \eta)\| \leq \frac{1}{4}\epsilon$ für $x \in U$:

$$\|(A\eta)(x) - \eta\| = \|D^{-1}F(x, \eta)\| \leq \frac{1}{4}\epsilon \quad \text{für jedes } x \in U. \quad (\text{B.4})$$

Für beliebiges festes $x \in U$ wird die Funktion $\Phi : V \rightarrow \mathbb{R}^q$ erklärt durch

$$\Phi(y) := y - D^{-1} F(x, y), \quad (\text{B.5})$$

so ist $\Phi'(y) = I - D^{-1} \frac{\partial F(x, y)}{\partial y}$.

Aus der Abschätzung in (B.1) folgt nun mit der Abschätzung aus dem *Mittelwertsatz für vektorwertige Funktionen* (siehe B.3, S. 67) die Ungleichung:

$$\|\Phi(y) - \Phi(z)\| \leq \frac{1}{2}\|y - z\|, \quad \text{für beliebige } y, z \in V. \quad (\text{B.6})$$

Insbesondere gilt

$$\|\Phi(g_1(x)) - \Phi(g_2(x))\| \leq \frac{1}{2}\|g_1(x) - g_2(x)\| \quad \text{für alle } g_1, g_2 \in M. \quad (\text{B.7})$$

Wegen $\Phi(g_j(x)) = g_j(x) - D^{-1}F(x, g_j(x)) = (Ag_j)(x)$ läuft dies auf die Abschätzung

$$\|(Ag_1)(x) - (Ag_2)(x)\| \leq \frac{1}{2}\|g_1(x) - g_2(x)\| \quad \text{für } g_1, g_2 \in M \text{ und alle } x \in U, \quad (\text{B.8})$$

somit auf

$$\|Ag_1 - Ag_2\| \leq \frac{1}{2}\|g_1 - g_2\| \quad \text{für beliebige } g_1, g_2 \in M \quad (\text{B.9})$$

hinaus. (B.8) liefert zusammen mit (B.4), dass für jedes $g \in M$ und $x \in U$ gilt:

$$\|(Ag)(x) - \xi\| \leq \|(Ag)(x) - (A\xi)(x)\| + \|(A\xi)(x) - \xi\| \leq \frac{1}{2}\|g(x) - \xi\| + \frac{1}{4}\epsilon \leq \frac{1}{4}\epsilon + \frac{1}{4}\epsilon = \frac{1}{2}\epsilon.$$

Diese Abschätzung besagt in Verbindung mit (B.3), dass A die (abgeschlossene) Menge $M \subset B(U, \mathbb{R}^q)$ in sich abbildet, und (B.9) dass A darüber hinaus sogar eine Kontraktion ist. Nach dem Banachschen Fixpunktsatz (siehe B.1, S. 61) besitzt folglich A einen Fixpunkt f in M , es gibt somit eine stetige Funktion

$$f : U \rightarrow V \quad \text{mit} \quad f(\xi) = \eta \quad \text{und} \quad Af = f, \quad \text{somit} \quad F(c, f(x)) = 0 \quad \text{für alle } x \in U.$$

B. Differentialrechnung im Mehrdimensionalen

Bei vorgegebenen $x \in U$ sei nun $y \in V$ eine Lösung der Gleichung $F(x, y) = 0$.

Aus (B.6) folgt dann

$$\|y - f(x)\| = \|\Phi(y) - \Phi(f(x))\| \leq \frac{1}{2}\|y - f(x)\|$$

Diese Ungleichung kann aber nur bestehen, wenn $\|y - f(x)\| = 0$ und daher $y = f(x)$ gilt. Damit ist sowohl die Schlussbehauptung des Satzes als auch die eindeutige Bestimmtheit von f bewiesen. $\square$

Beispiele

Diese Beispiele zeigen, wie der *Hauptsatz über Implizite Funktionen* zum Auflösen von konkreten Funktionen nach y verwendet werden kann:

Beispiel.

vgl. ABG, 5.67, S.350

Sei $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ mit $f(x, y) := xye^{-(x+y)}$.

Wird die partielle Ableitung nach y mit $f_y(x, y) = x(1 - y)e^{-(x+y)}$ gebildet, so ist ersichtlich, dass für die Werte $x = 0$ bzw. $y = 1$, keine Invertierung der partiellen Ableitung der Funktion möglich ist.

Wird $(\xi, \eta) \in \mathbb{R}$ mit $\xi \neq 0$, $\eta \neq 1$ gewählt, so lässt sich die Gleichung $xye^{-(x+y)} = c$ für jedes $c \in \mathbb{R}$ lokal nach y in der Form $xf(x)e^{-(x+f(x))} = c$ auflösen.

Beispiel.

vgl. LA2, S.295, A1

Für hinreichend kleine x, y kann $e^{\sin(xy)} + x^2 - 2y - 1 = 0$ nach y aufgelöst werden.

Beispiel.

vgl. LA2, S.295, A2

Für hinreichend kleine x, y und genügend nahe bei -1 liegende z kann $x^2 + y^2 + z^2 + \cosh(xyz) = 0$ nach z aufgelöst werden.

B. Differentialrechnung im Mehrdimensionalen

Beispiel.

vgl. LA2, S.295, A3

Für genügend nahe bei 1 liegende x, y, z kann das Gleichungssystem $-2x^2 + y^2 + z^2 = 0$, $x^2 + e^{y-1} - 2y = 0$ durch stetige Funktionen $y = \varphi(x), z = \psi(x)$ aufgelöst werden.

Das folgenden Beispiel soll nun zeigen, dass eine globale Auflösung der Funktion oftmals nicht möglich ist, aber dennoch lokale Lösungen gefunden werden kann:

Beispiel.

vgl. ABG, Bsp.5.65, S.349

Sei $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ mit $f(x, y) := x^2 + y^2 - 1$ der Einheitskreis im $\mathbb{R}^2$ um den Koordinatenursprung. Wird versucht die Funktion $f(x, y)$ global nach x , bzw. y aufzulösen, d.h. sie explizit durch $f(x)$ bzw. $g(y)$ darzustellen, wird ersichtlich, dass es zu jedem x bzw. y je zwei Werte gibt.

Somit ist eine globale Auflösung nicht möglich!

Lokal lässt sich jedoch für $y > 0$ die Funktion $f_1 : (-1, 1) \rightarrow \mathbb{R}$ mit $f_1(x) = \sqrt{1 - x^2}$ und für $y < 0$ die Funktion $f_2 : (-1, 1) \rightarrow \mathbb{R}$ mit $f_2(x) = -\sqrt{1 - x^2}$ finden. Für $y = 0$ existiert jedoch keine entsprechende Funktion. (Überlegungen für Funktionen $g(y)$ analog).

5. Umkehrsatz

Der Umkehrsatz beschreibt welche Bedingungen eine Funktion $f : \mathbb{R}^p \rightarrow \mathbb{R}^p$ erfüllen muss um umkehrbar zu sein und die Eigenschaften der Umkehrfunktion φ .

SATZ B.9: Umkehrsatz

vgl. LA2, A1, S.300

Sei $f : G \subset \mathbb{R}^p \rightarrow \mathbb{R}^p$ (G offen) eine stetig differenzierbare Funktion, und in einem gewissen Punkt $\xi \in G$ sei $f'(\xi)$ invertierbar. Dann gibt es eine offene Umgebung $W \subset G$ von ξ und eine ε -Umgebung V von $\eta := f(\xi)$, so dass f die Umgebung W bijektiv auf V abbildet. Die Umkehrung

$$\varphi : V \rightarrow W \quad \text{von } f|_W$$

ist stetig differenzierbar, und ihre Abbildung berechnet sich nach der Formel

$$\varphi'(y) = (f'(x))^{-1} \quad \text{mit } y = f(x), \quad x \in W.$$

Beweis.

vgl. GM2; S.111

Der Umkehrsatz wird auf den Satz über Implizite Funktionen zurückgeführt.

Dazu wird folgende Funktion definiert:

$$F : G \times \mathbb{R}^p \rightarrow \mathbb{R}^p \quad \text{durch} \quad F(x, y) := f(x) - y$$

Es gilt $F(\xi, \eta) = 0$. Da $\frac{\partial F}{\partial y}(x, y) = -f'(y)$ und $f'(\xi)$ invertierbar ist, kann der Satz über Implizite Funktionen angewendet werden.

Es gibt folglich eine offene Umgebung V von η , eine Umgebung $U \subset G$ von ξ und eine stetig differenzierbare Abbildung $\varphi : V \rightarrow U$ mit folgenden Eigenschaften:

- Sei $0 = F(\varphi(x), y) = f(\varphi(x)) - y$, d.h. $f(\varphi(x)) = y$ für alle $y \in V$
- Ist $(x, y) \in V \times U$ mit $F(x, y) = 0$, d.h. $y = f(x)$, so folgt $x = \varphi(y)$.

Auf Grund der Stetigkeit von f gibt es eine offene Umgebung $f^{-1}(V) \subset U$ von ξ mit $f(W) \subset V$. Wegen der zweiten Eigenschaft gilt sogar $V := f(W)$.

Da φ stetig ist, ist V eine offene Umgebung von η . Daher ist $f : U \rightarrow V$ bijektiv mit der Umkehrung φ . □

Bemerkung. Der Satz macht nur Aussagen über die **lokale** Umkehrbarkeit der Funktion.

SATZ B.10: Satz über die Umkehrfunktion (alternativ)

vgl. GM2, S.111

Sei $U \subset \mathbb{R}^n$ offen und $f : U \rightarrow \mathbb{R}^n$ eine stetig differenzierbare Funktion. Sei $a \in U$ und $b := f(a)$. Die Jacobi-Matrix $J_f(a)$ sei invertierbar (d.h. $\det J_f(a) \neq 0$).

Dann gibt es eine offene Umgebung $U_a \subset U$ von a und eine offene Umgebung V_b von b , sodass f die Menge U_a bijektiv auf V_b abbildet und die Umkehrfunktion

$$g = f^{-1} : V_b \rightarrow U_a$$

stetig differenzierbar ist. Es gilt $J_g(b) = (J_f(a))^{-1}$.

Beweis.

Betrachte $F(x, y) = f(x) - y = 0$. Dies stellt ein Gleichungssystem in $\mathbb{R}^{2n}$ dar:

$$\begin{aligned} F_1(x_1, \dots, x_n, y_1, \dots, y_n) &= f_1(x_1, \dots, x_n) - y_1 \\ F_2(x_1, \dots, x_n, y_1, \dots, y_n) &= f_2(x_1, \dots, x_n) - y_2 \\ &\vdots \\ F_n(x_1, \dots, x_n, y_1, \dots, y_n) &= f_n(x_1, \dots, x_n) - y_n \end{aligned}$$

Laut Voraussetzung gilt $F(a, b) = 0$ und $F(x, y)$ ist stetig differenzierbar auf einer Umgebung von (a, b) .

$$\text{Wegen } \begin{vmatrix} \frac{\partial F_1}{\partial x_1} & \dots & \frac{\partial F_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial F_n}{\partial x_1} & \dots & \frac{\partial F_n}{\partial x_n} \end{vmatrix} = \begin{vmatrix} \frac{\partial f_1}{\partial x_1} & \dots & \frac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1} & \dots & \frac{\partial f_n}{\partial x_n} \end{vmatrix} \neq 0 \in a$$

lässt sich das Gleichungssystem $F(x, y) = 0$ lokal nach x auflösen, d.h. es existiert eine offene Umgebung V_b von b und eine stetig differenzierbare Funktion g auf V_b mit $x = g(y)$ auf V_b und $y = f(g(y))$. Dies bedeutet aber, dass $g = f^{-1}$.

Mit der Kettenregel und $f^{-1} \circ f = \text{id}_{U_a}$ (vgl. 1.1, S. 58) folgt sofort, dass

$$J_f^{-1}(y) = (J_f(f^{-1}(y)))^{-1} \text{ für } y \in V(b)$$

□

Beispiele

Beispiel.

vgl. GM2, S.114, Aufg. 8.1

Sei $F : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ eine Funktion mit $F(X, y) := (x^2 - y^2, 2xy)$.

Es ist die Jacobi-Matrix von F und, wo sie existiert, ihre Inverse zu berechnen:

Die zu F zugehörige Jacobi-Matrix lautet:

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

Da, wenn die Jacobi-Matrix invertierbar ist, gilt, dass ihre Determinante ungleich 0 ist, folgt für

$$\det(J_F(x, y)) = 4x^2 + 4y^2$$

die inverse Jacobi-Matrix mit

$$J_F(x, y)^{-1} = \begin{pmatrix} \frac{x}{2(x^2+y^2)} & \frac{y}{2(x^2+y^2)} \\ \frac{-y}{2(x^2+y^2)} & \frac{x}{2(x^2+y^2)} \end{pmatrix}$$

Beispiel.

Sei $f(x, y) := \begin{pmatrix} e^x \cos(y) \\ e^x \sin(y) \end{pmatrix}$ eine auf $\mathbb{R}^2$ stetig differenzierbare Funktion

mit der Ableitung $\begin{pmatrix} e^x \cos(y) & -e^x \sin(y) \\ e^x \sin(y) & e^x \cos(y) \end{pmatrix}$, die überall invertierbar ist.

Dann ist die Inverse Ableitungsfunktion für alle $x, y \in \mathbb{R}$

$$f'(x, y)^{-1} = \begin{pmatrix} e^x \cos(y) & e^x \sin(y) \\ -e^x \sin(y) & e^x \cos(y) \end{pmatrix}$$

6. Das Lagrange-Verfahren - Eine Methode zur Lösung von Optimierungsproblemen

Wenn Extremstellen von mehrdimensionalen Funktionen mit einer Nebenbedingung gesucht sind, ist das *Lagrange-Verfahren* ein praktisches Hilfsmittel, obwohl es nur notwendige Bedingung für das Vorliegen eines Extremums ist:

SATZ B.11: Lagrange-Multiplikatorenregel

vgl. LA2, S.320ff

X sei eine offene Teilmenge des $\mathbb{R}^p$, die Funktionen

$$f : X \rightarrow \mathbb{R} \quad \text{und} \quad g := \begin{pmatrix} g_1 \\ \vdots \\ g_q \end{pmatrix} : X \rightarrow \mathbb{R}^q \quad (q < p)$$

seien stetig differenzierbar, und f besitze in ξ ein lokales Extremum unter der Nebenbedingung $g(x) = 0$. Ferner gebe es in

$$g'(\xi) = \begin{pmatrix} \frac{\partial g_1(\xi)}{\partial x_1} & \cdots & \frac{\partial g_1(\xi)}{\partial x_p} \\ \vdots & & \vdots \\ \frac{\partial g_q(\xi)}{\partial x_1} & \cdots & \frac{\partial g_q(\xi)}{\partial x_p} \end{pmatrix}$$

eine q -reihige Unterdeterminante, die nicht verschwindet. Dann existieren q Zahlen $\lambda_1, \dots, \lambda_q$ (**Lagrange-Multiplikatoren**), mit denen die folgende Gleichung gilt:

$$f'(\xi) + \sum_{j=1}^q \lambda_j g'_j(\xi) = 0$$

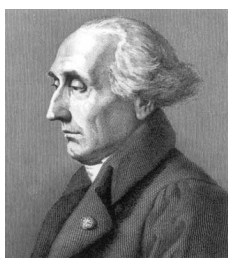


Abbildung B.11.: *Joseph-Louis de Lagrange, italienischer Mathematiker (1736 - 1813)*

Quelle (abgerufen am 22.08.2019): de.wikipedia.org/wiki/Joseph-Louis_Lagrange

Beweis.

Nach Voraussetzung gibt es in der (q, p) -Matrix $g'(\xi)$ eine q -reihige Unterdeterminante $\neq 0$. O.B.d.A kann angenommen werden, dass bereits

$$\begin{vmatrix} \frac{\partial g_1(\xi)}{\partial x_1} & \cdots & \frac{\partial g_1(\xi)}{\partial x_q} \\ \vdots & & \vdots \\ \frac{\partial g_q(\xi)}{\partial x_1} & \cdots & \frac{\partial g_q(\xi)}{\partial x_q} \end{vmatrix}$$

ist. Nun werden

$$y := \begin{pmatrix} x_1 \\ \vdots \\ x_q \end{pmatrix}, \quad z := \begin{pmatrix} x_{q+1} \\ \vdots \\ x_p \end{pmatrix}, \quad \xi := \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_p \end{pmatrix}, \quad \eta := \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_q \end{pmatrix}, \quad \zeta := \begin{pmatrix} \xi_{q+1} \\ \vdots \\ \xi_p \end{pmatrix}$$

gesetzt und

$f(x, y)$ bzw. $g(x, y)$ an Stelle von $f(x)$ bzw. $g(x)$ geschrieben.

Da f in ξ ein lokales Extremum unter der Nebenbedingung $g(x) = 0$ besitzt, ist $g(\eta, \zeta) = 0$.

Ferner folgt aus der oben erwähnten Unterdeterminante und dem Invertierbarkeitskriterium (siehe 0.30, S. 25), dass

$$\frac{\partial g}{\partial y}(\eta, \zeta) = \begin{pmatrix} \frac{\partial g_1(\eta, \zeta)}{\partial x_1} & \cdots & \frac{\partial g_1(\eta, \zeta)}{\partial x_q} \\ \vdots & & \vdots \\ \frac{\partial g_q(\eta, \zeta)}{\partial x_1} & \cdots & \frac{\partial g_q(\eta, \zeta)}{\partial x_q} \end{pmatrix} \quad \text{invertierbar ist.}$$

Nach dem Hauptsatz über Implizite Funktionen (siehe B.8, S. 81) und dem Satz über die stetige Differenzierbarkeit folgt nun, dass $g(y, z) = 0$ stetig differenzierbar nach y aufgelöst werden kann.

Genauer: Es gibt eine δ -Umgebung $U \subset \mathbb{R}^{p-q}$ von ζ und eine stetig differenzierbare Funktion

$$k : U \rightarrow \mathbb{R}^q \quad \text{mit} \quad g(h(z), z) = 0 \quad \text{für alle} \quad z \in U.$$

Auf einer hinreichend kleinen δ_1 -Umgebung $W \subset \mathbb{R}^{p-q}$ von ζ ist dann die reellwertige Funktion

$$\phi(z) := f(h(z), z) \quad \text{differenzierbar.}$$

Ihre partiellen Ableitungen an der Stelle ζ lassen sich gemäß der mehrdimensionalen

B. Differentialrechnung im Mehrdimensionalen

Kettenregel (siehe 2.4, S. 71) berechnen. Mit

$$\frac{\partial f(\xi)}{\partial y} := \left(\frac{\partial f(\xi)}{\partial x_1}, \dots, \frac{\partial f(\xi)}{\partial x_q} \right) \quad \text{und} \quad \frac{\partial f(\xi)}{\partial z} := \left(\frac{\partial f(\xi)}{\partial x_{q+1}}, \dots, \frac{\partial f(\xi)}{\partial x_p} \right)$$

folgt somit

$$\phi'(\zeta) = \frac{\partial f(\xi)}{\partial y} h'(\zeta) + \frac{\partial f(\xi)}{\partial z}.$$

Aus den Voraussetzungen über das Extremalverhalten von f an der Stelle ξ ergibt sich nun sofort, dass ϕ in ζ ein lokales Extremum **ohne** Nebenbedingungen besitzt. Nach dem Satz über lokale Extrema (siehe B.14, S. 74) ist somit $\phi'(\zeta) = 0$ und daher

$$\frac{\partial f(\xi)}{\partial y} h'(\zeta) + \frac{\partial f(\xi)}{\partial z} = 0.$$

Weiters gilt

$$h'(\zeta) = - \left(\frac{\partial g(\xi)}{\partial y} \right)^{-1} \frac{\partial g(\xi)}{\partial z}.$$

Wird dies in obige Gleichung eingesetzt, so folgt

$$-\frac{\partial f(\xi)}{\partial y} \left(\frac{\partial g(\xi)}{\partial y} \right)^{-1} \frac{\partial g(\xi)}{\partial z} + \frac{\partial f(\xi)}{\partial z} = 0$$

Für die $(1, q)$ -Matrix

$$L = (\lambda_1, \dots, \lambda_q) = -\frac{\partial f(\xi)}{\partial y} \left(\frac{\partial g(\xi)}{\partial y} \right)^{-1}$$

ist trivialerweise

$$\frac{\partial f(\xi)}{\partial y} + L \frac{\partial g(\xi)}{\partial y} = 0.$$

Andererseits lässt sich dies auch in der Form

$$\frac{\partial f(\xi)}{\partial z} + L \frac{\partial g(\xi)}{\partial z} = 0 \quad \text{schreiben.}$$

Aus den beiden letzten Gleichungen folgt die Gleichung im obigen Satz. □

Beispiele

Beispiel.

LA2, A1, S.329

Es sind die Extremwerte der Funktion $f(x, y) := xy$ auf der Geraden $x + y = 1$ zu berechnen.

Diese Aufgabe könnte zum Beispiel ein möglichst großes Rechteck mit Fläche $x \cdot y$ sein, dessen halber Umfang $x + y = 1$ ist. Dann gilt $0 < x < 1$ und $0 < y < 1$, da es keine negativen Längen gibt und die Nebenbedingung erfüllt sein soll.

Nun kann das Lagrange-Verfahren genutzt werden um mögliche Kandidaten für das Maximum zu finden:

Dazu ist zunächst die Nebenbedingung auf die Form $g(x, y) = 0$ zu bringen:

$$g(x, y) = x + y - 1 = 0$$

Das Lagrange-Gleichungssystem lautet nun:

$$L(x, y, \lambda) = xy + \lambda(x + y - 1)$$

Die partiellen Ableitungen des Gleichungssystems werden gebildet und $= 0$ gesetzt:

$$L_x = \frac{\partial L}{\partial x} = y + \lambda = 0 \quad L_y = \frac{\partial L}{\partial y} = x + \lambda = 0 \quad L_\lambda = x + y - 1 = 0$$

Es wird die erste Gleichung nach x bzw. die zweite nach y umgeformt und in die dritte Gleichung eingesetzt. Daraus folgt, dass $\lambda = -\frac{1}{2}$ ist und somit ist mit $(\frac{1}{2}, \frac{1}{2})$ ein stationärer Punkt gefunden, der das Gleichungssystem löst.

Abschließend kann nun die notwendige Bedingung für ein Extremum überprüft werden (vgl. B.5, S. 74):

Dazu wird nun die erste Ableitung der mit Hilfe der Nebenbedingung auf eine Variable reduzierten Funktion $f(x) = x(1 - x)$ gebildet und gleich 0 gesetzt. Dies liefert $f'(x) = -2x + 1 = 0$, woraus $x = \frac{1}{2}$ folgt.

Nun sind noch die Randpunkte zu prüfen, die jedoch für $x = 1$ mit $y = 0$ bzw. $x = 0$ mit $y = 1$ keine Maxima der Funktion f liefern.

Somit ist obiger stationäre Punkt das Maximum.

B. Differentialrechnung im Mehrdimensionalen

Beispiel.

LA2, A7, S.329

Es sind die Extremwerte der Funktion $f(x, y) := x + y$ unter der Nebenbedingung $g(x, y) := x^2 + y^2 - 1 = 0$ zu berechnen.

Das Lagrange-Gleichungssystem lautet nun:

$$L(x, y, \lambda) = x + y + \lambda(x^2 + y^2 - 1)$$

Die partiellen Ableitungen des Gleichungssystems werden gebildet und $= 0$ gesetzt:

$$\begin{aligned} L_x = \frac{\partial L}{\partial x} &= 1 + 2\lambda x = 0 \\ L_y = \frac{\partial L}{\partial y} &= 1 + 2\lambda y = 0 \\ L_\lambda &= x^2 + y^2 - 1 = 0 \end{aligned}$$

Es wird die erste Gleichung nach x bzw. die zweite nach y umgeformt und in die dritte Gleichung eingesetzt. Daraus folgt, dass $\lambda_{1,2} = \pm \frac{1}{\sqrt{2}}$ ist und somit sind mit

$\left(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)$ und $\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$ zwei Punkte gefunden, die das Gleichungssystem lösen.

Da es sich bei der Nebenbedingung um den Einheitskreis (kompakte Menge) handelt und somit Extrema existieren müssen, sind die gesuchten Extremwerte das lokale Minimum mit $f\left(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right) = -\sqrt{2}$ und das lokale Maximum mit $f\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right) = \sqrt{2}$

6.1. Modifizierte Hesse-Matrix

Da bei der Suche nach Extremstellen von mehrdimensionalen Funktionen mit einer Nebenbedingung und unter Anwendung des *Lagrange-Verfahrens* die übliche Hesse-Matrix oftmals nicht angewendet werden kann, gibt es die sog.

DEFINITION B.18: Modifizierte Hesse-Matrix

vgl. MMO, S.221

$$\tilde{H}_f(\lambda, x) = \begin{bmatrix} 0 & \cdots & 0 & \frac{\partial g_1}{\partial x_1} & \cdots & \frac{\partial g_1}{\partial x_n} \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & \frac{\partial g_m}{\partial x_1} & \cdots & \frac{\partial g_m}{\partial x_n} \\ -\frac{\partial g_1}{\partial x_1} & \cdots & -\frac{\partial g_m}{\partial x_1} & \frac{\partial^2 L}{\partial x_1^2} & \cdots & \frac{\partial^2 L}{\partial x_1 \partial x_n} \\ \vdots & & \vdots & \vdots & & \vdots \\ -\frac{\partial g_1}{\partial x_n} & \cdots & -\frac{\partial g_m}{\partial x_n} & \frac{\partial^2 L}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 L}{\partial x_n^2} \end{bmatrix}$$

Mit ihr kann nun eine *hinreichende Bedingung für Extrema unter Nebenbedingungen* formuliert werden:

SATZ B.12: Extrema unter Nebenbedingungen

vgl. MMO, S.221

Sei (λ^*, ξ) ein stationärer Punkt (A.9, S. 37) der Lagrangefunktion. Dann hat die Funktion f unter den Nebenbedingungen $g_1(x) = 0, \dots, g_m(x) = 0$ im Punkt ξ ein

1. lokales Minimum, wenn $\tilde{\Delta}_i(\lambda^*, \xi) > 0$ für $i = 2m + 1, \dots, m + n$,
2. lokales Maximum, wenn $(-1)^i \tilde{\Delta}_i(\lambda^*, \xi) > 0$ für $i = 2m + 1, \dots, m + n$

B. Differentialrechnung im Mehrdimensionalen

Beispiel.

Sei $f(x_1, x_2, x_3) = x_1^2 + x_2 - 2x_3$ eine Funktion in *drei Variablen* mit den *beiden Nebenbedingungen* $g_1(x_1, x_2, x_3) = x_1 + x_2 = 0$ und $g_2(x_1, x_2, x_3) = 2x_2 + x_3 = 0$. Es ist die Funktion nach Extrema zu untersuchen! ($n=3$; $m=2$)

Zunächst wird das Lagrange-Gleichung aufgestellt:

$$L(x_1, x_2, x_3, \lambda_1, \lambda_2) = x_1^2 + x_2 - 2x_3 + \lambda_1 (x_1 + x_2) + \lambda_2 (2x_2 + x_3)$$

Mit Hilfe der partiellen Ableitungen ($= 0$) wird das Lagrange-Gleichungssystem aufgestellt:

$$\begin{aligned} L_{x_1} &= \frac{\partial L}{\partial x_1} = 2x_1 + \lambda_1 = 0 \\ L_{x_2} &= \frac{\partial L}{\partial x_2} = 1 + \lambda_1 + \lambda_2 = 0 \\ L_{x_3} &= \frac{\partial L}{\partial x_3} = -2 + \lambda_2 = 0 \\ L_{\lambda_1} &= \frac{\partial L}{\partial \lambda_1} = x_1 + x_2 = 0 \\ L_{\lambda_2} &= \frac{\partial L}{\partial \lambda_2} = 2x_2 + x_3 = 0 \end{aligned}$$

Daraus können nun ein Stationärer Punkt in $(x_1, x_2, x_3) = (\frac{5}{2}, -\frac{5}{2}, 5)$ und die Lagrang Multiplikatoren $\lambda_1 = -5$ und $\lambda_2 = 2$ gefunden werden.

Somit lautet die Hesse-Matrix:

$$H_L(\lambda, x) = \begin{pmatrix} 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 2 & 1 \\ 1 & 0 & 2 & 0 & 0 \\ 1 & 2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{pmatrix}$$

Wegen der zwei Nebenbedingungen werden die Vorzeichen der beiden ersten Spalten geändert, um die modifizierte Hesse-Matrix zu erhalten:

6. Das Lagrange-Verfahren - Eine Methode zur Lösung von Optimierungsproblemen

$$\tilde{H}_L(\lambda, x) = \begin{pmatrix} 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 2 & 1 \\ -1 & 0 & 2 & 0 & 0 \\ -1 & -2 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 \end{pmatrix}$$

Da die Anzahl der Variablen $n = 3$ und die Anzahl der Nebenbedingungen $m = 2$ ist, sind somit $n - m = 1$ Minoren zu überprüfen:

$$\det \tilde{H}_L(\lambda, x) = 2$$

Da alle relevanten Hauptminoren positiv sind ist der oben gefundene Stationäre Punkt ein **lokales Minimum**.

7. Weitere Beispiele zur Differentialrechnung im $\mathbb{R}^p$

Hinweis: Die bei den folgenden Beispielen verwendeten Grafiken und Diagramme wurden mit der Internetseite www.wolframalpha.com/widgets erstellt.

BEISPIEL B.1.

vgl. LA2, Aufgabe +2, S.265

$$f(x, y) := \begin{cases} \frac{xy}{x^2+y^2}, & \text{falls } x \neq 0, \\ 0, & \text{falls } x = 0 \end{cases}$$

Zu zeigen ist, dass f im Nullpunkt unstetig ist, alle partiellen Ableitungen besitzt, aber dennoch nicht differenzierbar ist.

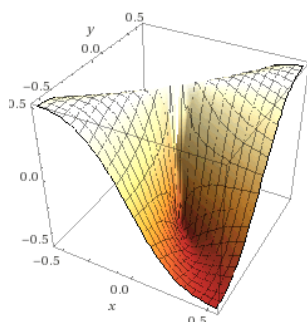


Abbildung B.12.: 3D-Grafik der Funktion

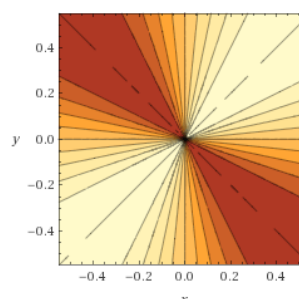


Abbildung B.13.: Schnitt-Grafik der Funktion

Existieren die partiellen Ableitungen? $\Rightarrow$ Ist f partiell diffbar in $(0,0)$?

$$\frac{\partial f(0,0)}{\partial x} = \lim_{h \rightarrow 0} \frac{f(0+h,0) - f(0,0)}{h} = \lim_{h \rightarrow 0} \frac{0 - 0}{h} = 0$$

Der Grenzwert existiert, daher existiert die partielle Ableitung von x im Punkt $(0,0)$ (y analog)

Ist f im Nullpunkt stetig? $\Rightarrow$ Existiert somit $\lim_{(x,y) \rightarrow 0} f(x,y)$?

Das ist anhand des Folgenkriterium für Stetigkeit zeigbar:

Setze $x_n := (\frac{1}{n})$, $y_n := (\frac{1}{n})$:

$$\lim_{n \rightarrow \infty} f\left(\frac{1}{n}, \frac{1}{n}\right) = \lim_{n \rightarrow \infty} \frac{\frac{1}{n} \cdot \frac{1}{n}}{\frac{1}{n^2} + \frac{1}{n^2}} = \frac{1}{2} \quad f\left(\lim_{n \rightarrow \infty} \left(\frac{1}{n}, \frac{1}{n}\right)\right) = f(0,0) = 0$$

Daher ist f in $(0,0)$ nicht stetig.

Aus Unstetigkeit folgt Undifferenzierbarkeit:

Da f in $(0,0)$ nicht stetig ist, ist f dort auch nicht differenzierbar und somit nicht diffbar.

7. Weitere Beispiele zur Differentialrechnung im $\mathbb{R}^p$

BEISPIEL B.2.

vgl. LA2, Aufgabe 3, S.275

$$f(x, y) := \begin{cases} \frac{xy^2}{x^2+y^4}, & \text{falls } x \neq 0, \\ 0, & \text{falls } x = 0 \end{cases}$$

Zu zeigen ist, dass f im Nullpunkt nicht stetig ist und weiters nicht differenzierbar ist, aber dennoch Ableitungen in alle Richtungen besitzt.

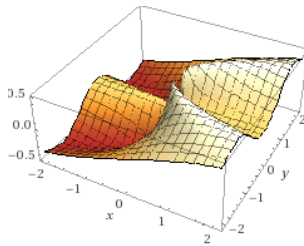


Abbildung B.14.: 3D-Grafik der Funktion

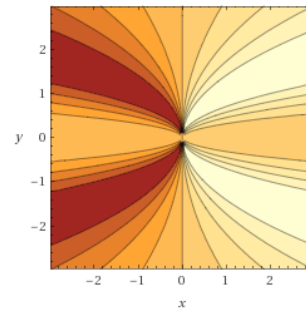


Abbildung B.15.: Schnitt-Grafik der Funktion

Existieren Ableitungen in alle Richtungen? $\Rightarrow$ Ist f partiell diffbar in alle Richtungen?

$f(x, y)$ ist eine reellwertige, mehrdim. Fkt auf einer Menge $G \in \mathbb{R}^2$ und $0 \neq v = (r, s) \in \mathbb{R}^2$:

$$\frac{f((0,0)+tv) - f(0,0)}{t} = \frac{f(tr, ts) - 0}{t} = \frac{1}{t} \frac{t^3 r s^2}{t^2 r^2 + t^4 s^4} = \frac{r s^2}{r^2 + t^2 s^4}$$

zwei Mögliche Fälle:

- $r \neq 0$, dann gilt: $\frac{r s^2}{r^2 + t^2 s^4} \rightarrow \frac{s^2}{r}$ für $t \rightarrow 0$
- $r = 0$, dann gilt: $\frac{r s^2}{r^2 + t^2 s^4} = \frac{0 \cdot s^2}{0^2 + t^2 s^4} = 0 \rightarrow 0$ für $t \rightarrow 0$

Da es für jedes v eine Richtungsableitung gibt, ist f partiell differenzierbar in alle Richtungen.

Ist f im Nullpunkt unstetig? $\Rightarrow$ Existiert daher $\lim_{(x,y) \rightarrow 0} f(x, y)$ nicht?

Wird $x_n := (\frac{1}{n^2})$, $y_n := (\frac{1}{n})$ gesetzt, ist anhand des Folgenkriterium die Stetigkeit zeigbar:

$$\lim_{n \rightarrow \infty} f\left(\frac{1}{n^2}, \frac{1}{n}\right) = \lim_{n \rightarrow \infty} \frac{\frac{1}{n^2} \cdot \frac{1}{n^2}}{\frac{1}{n^4} + \frac{1}{n^4}} = 1 \quad f\left(\lim_{n \rightarrow \infty} \left(\frac{1}{n^2}, \frac{1}{n}\right)\right) = f(0, 0) = 0$$

Daher ist f in $(0, 0)$ nicht stetig.

Aus Unstetigkeit folgt Undifferenzierbarkeit:

Da f in $(0, 0)$ unstetig ist, ist f dort nicht differenzierbar. Somit ist f nicht differenzierbar.

Abbildungsverzeichnis

0.1. Bernhard Bolzano	13
0.2. Karl Weierstraß	13
0.3. Stefan Banach	21
0.4. Augustin-Louis Cauchy	21
B.1. 3D-Grafik der Funktion	53
B.2. Beispiel einer reelwertigen Funktion	54
B.3. Tagesmittel der Lufttemperatur für 30. Juni 2019	55
B.4. Schraubenfeder	56
B.5. VektorPlot eines Vektorfeldes	57
B.6. Rudolf Otto Sigismund Lipschitz	59
B.7. Ludwig Otto Hesse	75
B.8. James Joseph Sylvester	77
B.9. Extrema-BSP-1	79
B.10. Extrema-BSP-2	80
B.11. Joseph-Louis de Lagrange	90
B.12. 3D-Grafik der Funktion	98
B.13. Schnitt-Grafik der Funktion	98
B.14. 3D-Grafik der Funktion	99
B.15. Schnitt-Grafik der Funktion	99

Für die Erstellung folgender Grafiken wurde **wolframalpha.com/widgets** verwendet:
B.1, B.5, B.12, B.13, B.14, B.15

Für die Erstellung folgender Grafiken wurde **GeoGebra Classic** (5.0.295.0) verwendet:
B.2, B.4, B.9, B.10

Symbolverzeichnis

$d(a, b)$	Abstand zwischen a und b
$\max(a, b)$	größere Zahl von a, b
$\min(a, b)$	kleinere Zahl von a, b
$\mathfrak{M}(q, p)$	Menge aller (q, p) -Matrizen
H_f	Hesse-Matrix
$\tilde{H}_f$	Modifizierte Hesse-Matrix
I	Einheitsmatrix
id_X	Identische Abbildung (Identität)
J_f	Jacobi-Matrix
$[A]$	Darstellungsmatrix von A
Δ_i	Hauptminor
L	Lagrange-Gleichungssystem
λ_i	Lagrange-Multiplikator
$\mathfrak{S}$	Menge aller linearen Abbildungen des linearen Raumes
∇	Nabla-Operator, Gradientsymbol bei kartesischen Koordinaten
(X, d_X)	metrischer Raum mit Abstandsfunktion d_X
$\subseteq$	die erste Menge ist Untermenge der zweiten
$\supseteq$	die erste Menge ist Obermenge der zweiten
(a_n)	Folge
$grad$	Gradient
$ \cdot $	Betrag
$\ \cdot\ $	Norm
$\langle \cdot, \cdot \rangle$	Skalarprodukt (inneres Produkt)
$\square$	q.e.d., Ende des Beweises

Literaturverzeichnis

LA1: Harro Heuser:

Lehrbuch der Analysis Teil 1

Vieweg+Teubner, Wiesbaden, 2009 (17.Aufl.); ISBN 978-3-8348-0777-9

LA2: Harro Heuser:

Lehrbuch der Analysis Teil 2

Vieweg+Teubner, Wiesbaden, 2008 (14.Aufl.); ISBN 978-3-8351-0208-8

A12: Rupert Lasser, Frank Hofmaier:

Analysis 1+2

Springer-Verlag, Berlin Heidelberg, 2012; ISBN 978-3-642-28643-8

GM1: Otto Forster:

Grundkurs Mathematik - Analysis 1

Springer-Verlag, Wiesbaden, 2016 (12.Aufl.); ISBN 978-3-658-19410-9

GM2: Otto Forster:

Grundkurs Mathematik - Analysis 2

Springer-Verlag, Wiesbaden, 2017 (11.Aufl.) ;ISBN 978-3-658-19410-9

AGB: Jürgen Appell:

Analysis in Beispielen und Gegenbeispielen

Springer-Verlag, Berlin Heidelberg, 2009 ; ISBN 978-3-540-88902-1

TA1: Florian Modler, Martin Kreh:

Tutorium Analysis 1 und Lineare Algebra 1

Spektrum Akademischer Verlag, Heidelberg, 2010; ISBN 978-3-8274-2345-0

LAG: Gerd Fischer:

Lineare Algebra

Vieweg+Teubner, Wiesbaden, 2008 (16.Aufl.); ISBN 978-3-8348-0428-0

MMO: Karl Mosler, Rainer Dykerhoff, Christoph Scheicher :

Mathematische Methoden für Ökonomen

Springer-Verlag, Berlin Heidelberg, 2018 (3.Aufl.); ISBN 978-3-662-54245-3

Anhang

Abstract (in English)

This diploma thesis deals with topics from the lectures *analysis in one variable* (« *Analysis in einer Variablen* ») and *analysis in several variables* (« *Analysis in mehreren Variablen* ») from the diploma program for teaching candidates. The basic theorems and definitions for multi-dimensional metric spaces are generalized, but since these concepts are well known, they can only be discussed in more detail and not newly developed. With the help of extensive literature research this diploma thesis tries to give a common thread through the topic and to provide a good overview by using motivating examples.

This academic paper is split up into the following 3 chapters. First the necessary basics for analysis (see Chapter 0) and the properties of one-dimensional functions (see Chapter A) are discussed. The concepts for one-dimensional functions are generalized for multidimensional functions in the last chapter (see Chapter B).

Furthermore the implicit function theorem, the inverse function theorem, and the method of Lagrange multipliers are presented.

The literature used for the preparation of this diploma thesis is documented by shortcuts (see « Literaturverzeichnis » = bibliography).

Many of the graphics used are created by using WolframAlpha and GeoGebra.

Abstract (in German)

Die vorliegende Diplomarbeit behandelt unter anderem Themengebiete aus den Vorlesungen *Analysis in einer Variablen* und *Analysis in mehreren Variablen* aus dem Diplomstudium für LehramtskandidatInnen. Dabei werden die grundlegenden Sätze und Definitionen für mehrdimensionale metrische Räume verallgemeinert und anhand von Beispielen untersucht.

Da diese Konzepte jedoch wohlbekannt sind, können sie nur näher besprochen, jedoch nicht neu entwickelt werden. Dennoch versucht diese Diplomarbeit mit Hilfe einer umfangreichen Literaturrecherche einen roten Faden durch die Thematik zu geben und mit motivierenden Beispielen einen guten Überblick zu verschaffen.

Die Arbeit ist in 3 Kapitel unterteilt. Zunächst wird auf die für Analysis nötigen *Grundlagen* (siehe Kapitel 0) und die Eigenschaften von eindimensionalen Funktionen (siehe Kapitel A) eingegangen. Im darauffolgenden Kapitel *Differenzialrechnung im Mehrdimensionalen* (siehe Kapitel B) werden die Konzepte aus dem Eindimensionalen für mehrdimensionale Funktionen verallgemeinert.

Darüber hinaus werden der *Hauptsatz über Implizite Funktionen*, der *Umkehrsatz*, sowie eine Methode zur Lösung von Optimierungsaufgaben (*Lagrange-Verfahren*) mit dazu passenden Beispielen vorgestellt.

Die für die Erstellung dieser Diplomarbeit verwendete Literatur ist anhand der Kürzel (siehe Literaturverzeichnis) nachzuvollziehen.

Viele der verwendeten Grafiken wurden mit *WolframAlpha* und *GeoGebra* erstellt.