IX. Funktionen mehrerer Veränderlicher

Lineare Abbildungen

Definition. *Seien X, Y Vektorräume über $\mathbb{R}$. Mit $L(X,Y)$ bezeichnen wir den Vektorraum aller linearen Abbildungen von X nach Y.*
Ist $A \in L(\mathbb{R}^n, \mathbb{R}^m)$, so definiert man die Norm von A durch

$$\|A\| = \sup\{|A(x)| \ : \ |x| \leq 1\}$$

.
Es gelten die folgenden Aussagen für $A \in L(\mathbb{R}^n, \mathbb{R}^m)$:

$$\|A\| = \max\{|A(x)| \ : \ |x| = 1\},$$

ist $|A(x)| \leq \lambda |x|$ für alle $x \in \mathbb{R}^n$, so folgt $\|A\| \leq \lambda$.

Satz. *(a) Ist $A \in L(\mathbb{R}^n, \mathbb{R}^m)$, dann gilt $\|A\| < \infty$ und A ist eine gleichmäßig stetige Abbildung von $\mathbb{R}^n$ nach $\mathbb{R}^m$.*
(b) Sind $A, B \in L(\mathbb{R}^n, \mathbb{R}^m)$ und ist c ein Skalar, so gilt

$$\|A + B\| \leq \|A\| + \|B\| \quad \text{und} \quad \|cA\| = |c|\,\|A\|.$$

$L(\mathbb{R}^n, \mathbb{R}^m)$ ist ein metrischer Raum mit der Metrik

$$d(A,B) = \|A - B\|.$$

(c) Für $A \in L(\mathbb{R}^n, \mathbb{R}^m)$ und $B \in L(\mathbb{R}^m, \mathbb{R}^k)$ gilt:

$$\|BA\| \leq \|B\|\,\|A\|.$$

Satz. *Sei Ω die Menge aller invertierbaren linearen Abblidungen auf $\mathbb{R}^n$.*
(a) Gilt für $A \in \Omega$ und $B \in L(\mathbb{R}^n)$ die Abschätzung

$$\|B - A\|\,\|A^{-1}\| < 1,$$

so ist $B \in \Omega$.
(b) Ω ist eine offene Teilmenge von $L(\mathbb{R}^n)$ und die Abbildung $A \to A^{-1}$ ist eine stetige, injektive, selbstinverse Abbildung auf Ω.

Bemerkung. *Ist $A \in L(\mathbb{R}^n, \mathbb{R}^m)$ und $A = [a_{i,j}]$ die Matrixdarstellung von A in Bezug auf die Standardbasen, dann gilt*

$$\|A\| \leq \left(\sum_{i,j} a_{ij}^2 \right)^{1/2} .$$

Sind $a_{11}, \ldots, a_{mn}$ stetige, reellwertige Funktionen auf einem metrischen Raum S, und ist A_p jene lineare Abbildung von $\mathbb{R}^n$ nach $\mathbb{R}^m$, deren Matrix die Elemente $a_{ij}(p) \quad (p \in S)$ hat, dann ist $p \to A_p$ eine stetige Abbildung von S nach $L(\mathbb{R}^n, \mathbb{R}^m)$.
Abbildungen dieser Art sind die totalen Ableitungen im nächsten Abschnitt.

Differentiation

Bemerkung. *Ist $f : (a, b) \longrightarrow \mathbb{R}^m$ eine differenzierbare Funktion, so bezeichnet man mit $f'(x)$ denjenigen Vektor in $\mathbb{R}^m$, für den gilt*

$$\lim_{h\to 0}\left(\frac{f(x+h)-f(x)}{h} - f'(x)\right) = 0.$$

Die lineare Abbildung $h \to hf'(x)$ von $\mathbb{R}$ nach $\mathbb{R}^m$ kann mit $f'(x)$ identifiziert werden und somit ist $f'(x) \in L(\mathbb{R}, \mathbb{R}^m)$ diejenige lineare Abbildung von $\mathbb{R}$ nach $\mathbb{R}^m$, für die

$$\lim_{h\to 0}\frac{f(x+h)-f(x)-f'(x)(h)}{h} = 0$$

gilt.

Diese Interpretation der Ableitung führt nun zum Begriff der Ableitung von Funktionen von $\mathbb{R}^n$ nach $\mathbb{R}^m$.

Definition. *Sei $E \subseteq \mathbb{R}^n$ eine offene Menge und $f : E \longrightarrow \mathbb{R}^m$ eine Funktion, sowie $x \in E$. Existiert eine lineare Abbildung $A \in L(\mathbb{R}^n, \mathbb{R}^m)$, derart dass*

$$\lim_{h\to 0}\frac{|f(x+h)-f(x)-A(h)|}{|h|} = 0 \qquad (*)$$

gilt, so sagt man, f ist an der Stelle x differenzierbar und man schreibt $f'(x) = A$. Ist f an jeder Stelle $x \in E$ differenzierbar, so sagt man f ist differenzierbar in E.
Die lineare Abbildung $f'(x)$ nennt man auch Differential an der Stelle x oder auch totale Ableitung.
Die Abbildung $x \to f'(x)$ ordnet jedem $x \in E$ ein Element in $L(\mathbb{R}^n, \mathbb{R}^m)$ zu, also

$$f' : E \longrightarrow L(\mathbb{R}^n, \mathbb{R}^m).$$

Satz. *Gilt die Formel (*) für zwei lineare Abbildungen A_1 und A_2, dann ist $A_1 = A_2$.*

Bemerkung. *Ist f an der Stelle x differenzierbar, so ist f in x auch stetig.*
Für eine lineare Abbildung $A \in L(\mathbb{R}^n, \mathbb{R}^m)$ gilt $A'(x) = A$ für jedes $x \in \mathbb{R}^n$.

Satz (Kettenregel)**.** *Sei E eine offene Teilmenge des $\mathbb{R}^n$ und $f : E \longrightarrow \mathbb{R}^m$ eine Funktion, die an der Stelle $x_0 \in E$ differenzierbar ist. Ferner bilde die Funktion g eine offene Menge, die das Bild $f(E)$ enthält, nach $\mathbb{R}^k$ ab, und g sei an der Stelle $f(x_0)$ differenzierbar. Sei $F(x) = g(f(x))$. Dann ist $F : E \longrightarrow \mathbb{R}^k$ an der Stelle x_0 differenzierbar und es gilt*

$$F'(x_0) = g'(f(x_0)) \circ f'(x_0),$$

dabei ist $g'(f(x_0))$ eine lineare Abbildung von $\mathbb{R}^m$ nach $\mathbb{R}^k$ und $f'(x_0)$ eine lineare Abbildung von $\mathbb{R}^n$ nach $\mathbb{R}^m$ und die Verknüpfung der beiden die Ableitung von F an der Stelle x_0, also eine lineare Abbildung von $\mathbb{R}^n$ nach $\mathbb{R}^k$.

Um die Ableitung einer Funktion mehrerer Veränderlicher in Matrixform darstellen zu können, führen wir nun den Begriff der partiellen Ableitung ein.

Definition. *Sei $\{e_1, \ldots, e_n\}$ die Standardbasis im $\mathbb{R}^n$ und $\{u_1, \ldots, u_m\}$ die Standardbasis im $\mathbb{R}^m$. Ist E eine offene Menge im $\mathbb{R}^n$ und $f : E \longrightarrow R^m$ eine Funktion, so sind die Komponenten $f_i : E \longrightarrow \mathbb{R}$ $(i = 1, \ldots, m)$ definiert durch*

$$f(x) = \sum_{i=1}^{m} f_i(x) u_i$$

und es gilt $f_i(x) = f(x).u_i$ (inneres Produkt in $\mathbb{R}^m$).
Für $x \in E$ und $1 \leq i \leq m$, sowie $1 \leq j \leq n$ definieren wir

$$(D_j f_i)(x) = \lim_{t \to 0} \frac{f_i(x + te_j) - f_i(x)}{t},$$

vorausgesetzt der Grenzwert existiert.
Schreibt man $f_i(x) = f_i(x_1, \ldots, x_n)$, so ist $D_j f_i$ die Ableitung von f_i bezüglich x_j, wobei die anderen Variablen fest bleiben, man sagt auch $D_j f_i$ ist die partielle Ableitung von f_i nach x_j.
Notation: $D_j f_i = \frac{\partial f_i}{\partial x_j}$.

Bemerkung. *Für die Funktion $f(x, y) = \frac{xy}{x^2+y^2}$, $x^2+y^2 \neq 0$; $f(x, y) = 0$, $x = y = 0$ existieren überall die partiellen Ableitungen $\frac{\partial f}{\partial x}$ und $\frac{\partial f}{\partial y}$, f ist jedoch an der Stelle $(0, 0)$ nicht differenzierbar, ja sogar unstetig.*

Satz. *Sei E eine offene Teilmenge von $\mathbb{R}^n$ und $f : E \longrightarrow \mathbb{R}^m$ eine an der Stelle $x \in E$ differenzierbare Funktion. Dann existieren die partiellen Ableitungen $(D_j f_i)(x)$ und es gilt*

$$f'(x)(e_j) = \sum_{i=1}^{m} (D_j f_i)(x) u_i \ , \ (1 \leq j \leq n).$$

Die Matrixdarstellung der linearen Abbildung $f'(x)$ bezüglich der Standardbasen hat die Gestalt

$$[f'(x)] = \begin{bmatrix} (D_1 f_1)(x) & \ldots & (D_n f_1)(x) \\ \ldots & \ldots & \ldots \\ (D_1 f_m)(x) & \ldots & (D_n f_m)(x) \end{bmatrix},$$

ist $h = \sum_{j=1}^{n} h_j e_j$ so folgt

$$f'(x)(h) = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} (D_j f_i)(x)\, h_j \right) u_i.$$

Im folgenden werden einige Anwendungen der Kettenregel beschrieben:

(a) Sei $\gamma : (a,b) \longrightarrow E \subset \mathbb{R}^n$ eine differenzierbare Kurve in E und $f : E \longrightarrow \mathbb{R}$ eine differenzierbare Funktion. Wir definieren $g(t) = f(\gamma(t))$ für $t \in (a,b)$. Dann gilt nach der Kettenregel

$$g'(t) = f'(\gamma(t)) \circ \gamma'(t),$$

dabei ist $\gamma'(t) \in L(\mathbb{R}, \mathbb{R}^n)$, also in der Matrixdarstellung

$$[\gamma'(t)] = \begin{bmatrix} \gamma_1'(t) \\ \cdot \\ \cdot \\ \gamma_n'(t) \end{bmatrix}$$

eine Spaltenmatrix und $f'(\gamma(t)) \in L(\mathbb{R}^n, \mathbb{R})$, also in der Matrixdarstellung

$$[f'(\gamma(t))] = [(D_1 f)(\gamma(t)) \ldots (D_n f)(\gamma(t))]$$

eine Zeilenmatrix und es gilt

$$g'(t) = \sum_{i=1}^{n} (D_i f)(\gamma(t))\, \gamma_i'(t).$$

Mit $(\nabla f)(x) = \sum_{i=1}^{n} (D_i f)(x)\, e_i$ bezeichnet man den Gradienten von f an der Stelle x. Die obige Formel für $g'(t)$ lässt sich dann in der Form

$$g'(t) = (\nabla f)(\gamma(t)).\gamma'(t)$$

schreiben, wobei wir hier das innere Produkt im $\mathbb{R}^n$ verwenden.

Spezialfall : sei $x \in E$ fix und $u \in \mathbb{R}^n$ ein Einheitsvektor ($|u| = 1$) und sei $\gamma(t) = x + tu$ für $t \in \mathbb{R}$. Dann ist $\gamma'(t) = u$ für jedes t und $g'(0) = (\nabla f)(x).u$.
Andererseits ist

$$g'(0) = \lim_{t \to 0} \frac{g(t) - g(0)}{t} = \lim_{t \to 0} \frac{f(x + tu) - f(x)}{t},$$

was als Ableitung von f in Richtung u gedeutet werden kann. Man schreibt dafür

$$(D_u f)(x) = (\nabla f)(x).u = \sum_{i=1}^{n} (D_i f)(x)\, u_i$$

und nennt $(D_u f)(x)$ die Richtungsableitung von f an der Stelle x in Richtung des Einheitsvektors u.
Bei fixem x ist $(D_u f)(x)$ maximal, wenn u ein positives skalares Vielfaches von $(\nabla f)(x)$ ist.

(b) Sei $f : \mathbb{R}^q \longrightarrow \mathbb{R}$ eine differenzierbare Funktion der Variablen $u_1, \ldots, u_q$. Diese Variablen werden als Funtionen $u_j : \mathbb{R}^p \longrightarrow \mathbb{R}$, $j =$

$1, \ldots, q$ aufgefasst. Man betrachtet nun die Funktion $\varphi : \mathbb{R}^p \longrightarrow \mathbb{R}$ definiert durch

$$\varphi(x_1, \ldots, x_p) = f(u_1(x_1, \ldots, x_p), \ldots, u_q(x_1, \ldots, x_p))$$

und erhält aus der Kettenregel

$$\begin{aligned} \frac{\partial \varphi}{\partial x_1} &= \frac{\partial f}{\partial u_1}\frac{\partial u_1}{\partial x_1} + \frac{\partial f}{\partial u_2}\frac{\partial u_2}{\partial x_1} + \cdots + \frac{\partial f}{\partial u_q}\frac{\partial u_q}{\partial x_1} \\ &\;\;\vdots \\ \frac{\partial \varphi}{\partial x_p} &= \frac{\partial f}{\partial u_1}\frac{\partial u_1}{\partial x_p} + \frac{\partial f}{\partial u_2}\frac{\partial u_2}{\partial x_p} + \cdots + \frac{\partial f}{\partial u_q}\frac{\partial u_q}{\partial x_p}. \end{aligned}$$

(c) Transformation von Differentialausdrücken auf neue Koordinaten: Sei $U : \mathbb{R}^2 \longrightarrow \mathbb{R}$ eine differenzierbare Funktion in kartesischen Koordinaten (x, y). Nach Übergang zu Polarkoordinaten (r, φ), wobei $x = r\cos\varphi$ und $y = r\sin\varphi$ ist, schreiben wir $U(x, y) = u(r(x, y), \varphi(x, y))$ und erhalten aus der Kettenregel

$$\begin{aligned} \frac{\partial U}{\partial x} &= \frac{\partial u}{\partial r}\frac{\partial r}{\partial x} + \frac{\partial u}{\partial \varphi}\frac{\partial \varphi}{\partial x} \\ \frac{\partial U}{\partial y} &= \frac{\partial u}{\partial r}\frac{\partial r}{\partial y} + \frac{\partial u}{\partial \varphi}\frac{\partial \varphi}{\partial y}. \end{aligned}$$

Mit Hilfe der Transformationsformel zwischen kartesischen und Polarkoordinaten erhalten wir schließlich

$$\begin{aligned} \frac{\partial U}{\partial x} &= \frac{\partial u}{\partial r}\cos\varphi - \frac{\partial u}{\partial \varphi}\frac{\sin\varphi}{r} \\ \frac{\partial U}{\partial y} &= \frac{\partial u}{\partial r}\sin\varphi + \frac{\partial u}{\partial \varphi}\frac{\cos\varphi}{r}. \end{aligned}$$

Satz. *Sei E eine konvexe, offene Teilmenge des $\mathbb{R}^n$ und $f : E \longrightarrow \mathbb{R}^m$ eine auf E differenzierbare Funktion. Es existiere ferner eine Konstante $M > 0$ mit $\|f'(x)\| \leq M$ für jedes $x \in E$. Dann gilt*

$$|f(b) - f(a)| \leq M\,|b - a|$$

für alle $a, b \in E$.

Korollar. *Ist E und f wie oben und gilt zusätzlich $f'(x) = 0$ für alle $x \in E$, dann ist f konstant auf E.*

Wir wissen bisher, dass die Existenz der Ableitung $f'(x)$ bewirkt, dass alle partiellen Ableitungen $(D_i f_j)(x)$ an der Stelle x existieren. Andererseits folgt aus der Existenz der partiellen Ableitungen in einem Punkt im allgemeinen nicht die Existenz von f' in diesem Punkt. Welche zusätzliche Bedingung an die partiellen Ableitungen bewirkt, dass die Ableitung f' auf E existiert?

Definition. *Sei E eine offene Teilmenge des $\mathbb{R}^n$. Die Funktion $f : E \longrightarrow \mathbb{R}^m$ heißt stetig differenzierbar auf E, wenn f auf E differenzierbar ist und f' eine stetige Abbildung von E nach $L(\mathbb{R}^n, \mathbb{R}^m)$ ist, d.h. für jedes $x \in E$ und für jedes $\epsilon > 0$ existiert ein $\delta > 0$ mit*

$$\|f'(x) - f'(y)\| < \epsilon,$$

falls $y \in E$ und $|x - y| < \delta$ ist.
Wir sagen dann : f ist eine $\mathcal{C}^1$-Abbildung und schreiben $f \in \mathcal{C}^1(E)$.

Satz. *Sei E eine offene Teilmenge des $\mathbb{R}^n$ und $f : E \longrightarrow \mathbb{R}^m$ eine Funktion. Dann gilt : f ist genau dann stetig differenzierbar auf E, wenn alle partiellen Ableitungen $D_j f_i$, $1 \leq j \leq n$, $1 \leq i \leq m$ auf E existieren und auf E stetig sind.*
Insbesondere gilt : existieren alle partiellen Ableitungen $D_j f_i$ auf E und sind sie auf E stetig, dann existiert die Ableitung f' auf E und ist auf E eine stetige Funktion im Sinne der obigen Definition.

Ableitungen höherer Ordnung

Definition. *Sei $f : E \longrightarrow \mathbb{R}$ eine differenzierbare Funktion mit den partiellen Ableitungen $D_1 f, \ldots, D_n f$. Sind die Funktionen $D_j f$ selbst differenzierbar, dann sind die partiellen Ableitungen 2. Ordnung von f definiert durch*

$$D_{ij} f = D_i D_j f \quad (i, j = 1, \ldots, n);$$

wir schreiben auch

$$D_i D_j f = \frac{\partial^2 f}{\partial x_i \, \partial x_j}.$$

Sind alle diese Funktionen $D_{ij} f$ stetig auf E, so sagt man $f \in \mathcal{C}^2(E)$.
Eine Abbildung $f : E \longrightarrow \mathbb{R}^m$ gehört zu $\mathcal{C}^2(E)$, wenn jede Komponente von f zu $\mathcal{C}^2(E)$ gehört.

Bemerkung. *Es kann vorkommen, dass die Ableitungen $D_{ij} f$ und $D_{ji} f$ in einem Punkt $x_0 \in E$ beide existieren und $D_{ij} f(x_0) \neq D_{ji} f(x_0)$ gilt.*

Satz. *Sei E eine offene Teilmenge des $\mathbb{R}^2$ und $f : E \longrightarrow \mathbb{R}$, ferner mögen die partiellen Ableitungen $D_1 f$ und $D_{21} f$ auf ganz E existieren. Sei $Q \subset E$ ein abgeschlossenes Rechteck, dessen Seiten parallel zu den Koordinatenachsen verläuft mit (a, b) und $(a + h, b + k)$ als gegenüberliegenden Ecken ($h \neq 0, k \neq 0$).*

Man setze

$$\Delta(f,Q) = f(a+h,b+k) - f(a+h,b) - f(a,b+k) + f(a,b).$$

Dann existiert eine Punkt (x,y) im Innern von Q mit

$$\Delta(f,Q) = h\,k(D_{21}f)(x,y).$$

(Vergleiche mit dem Mittelwertsatz aus Kapitel V !)

Mit Hilfe des letzten Satzes können wir die Frage beantworten, unter welchen Bedingungen die Reihenfolge der Differentiation bei partiellen Ableitungen höherer Ordnung vertauscht werden kann.

Satz. *Sei E eine offene Teilmenge des $\mathbb{R}^2$ und $f : E \longrightarrow \mathbb{R}$ eine Funktion, für die die partiellen Ableitungen $D_1f, D_{21}f$ und D_2f auf ganz E existieren und für die gilt, dass $D_{21}f$ an der Stelle $(a,b) \in E$ stetig ist. Dann existiert die partielle Ableitung $D_{12}f$ an der Stelle (a,b) und es gilt*

$$(D_{12}f)(a,b) = (D_{21}f)(a,b).$$

Der Taylor'sche Lehrsatz und Extremwerte

Satz. *Sei G eine offene, konvexe Teilmenge des $\mathbb{R}^p$ und $f \in \mathcal{C}^{n+1}(G)$, ferner seien $x_0 \in G$ und $x_0 + h \in G$. Dann existiert ein $\theta \in (0,1)$, sodass*

$$f(x_0+h) = f(x_0) + \sum_{k=1}^{n} \frac{1}{k!}\,(h_1D_1 + \cdots + h_pD_p)^k f(x_0)$$

$$+\frac{1}{(n+1)!}\,(h_1D_1 + \cdots + h_pD_p)^{n+1} f(x_0+\theta h),$$

dabei ist $(h_1D_1 + \cdots + h_pD_p)^k$ die k-fache Hintereinanderausführung des Operators $(h_1D_1 + \cdots + h_pD_p)$.
Spezialfall: $n = 1$

$$f(x_0+h) = f(x_0) + f'(x_0).h + \frac{1}{2}\sum_{j,k=1}^{p} \frac{\partial^2 f(x_0+\theta h)}{\partial x_j\,\partial x_k}\,h_jh_k.$$

Um Extremwerte für Funktionen mehrerer Veränderlicher behandeln zu können, benötigen wir die folgende Aussage:

Satz. *Sei G eine offene, konvexe Teilmenge des $\mathbb{R}^p$ und $f \in \mathcal{C}^2(G)$, ferner seien $x_0 \in G$ und $x_0 + h \in G$. Dann gilt*

$$f(x_0+h) = f(x_0) + f'(x_0).h + \frac{1}{2}\sum_{j,k=1}^{p} \frac{\partial^2 f(x_0)}{\partial x_j\,\partial x_k}\,h_jh_k + |h|^2\rho(h),$$

wobei $\lim_{h\to 0}\rho(h) = 0$.

Satz. *Sei G eine offene Teilmenge des $\mathbb{R}^p$ und $f \in \mathcal{C}^1(G)$, ferner habe f in $\xi \in G$ ein lokales Extremum. Dann gilt:*

$$(\nabla f)(\xi) = f'(\xi) = \left(\frac{\partial f(\xi)}{\partial x_1}, \ldots, \frac{\partial f(\xi)}{\partial x_p}\right) = 0.$$

Definition. *Ein Punkt $\xi \in G$ heißt kritischer Punkt für f, wenn $(\nabla f)(\xi) = 0$.*

Wir wollen nun auch eine hinreichende Bedingung für die Existenz eines lokalen Extremums angeben.

Definition. *Sei $A = (a_{jk})$ eine $(p \times p)$-Matrix und $x = (x_1, \ldots, x_p)$. Sei*

$$Q_A(x) = \sum_{j,k=1}^{p} a_{jk} x_j x_k.$$

Q_A heißt die zu A gehörige quadratische Form. A heißt positiv definit, wenn $Q_A(x) > 0$ für alle $x \neq 0$. A heißt negativ definit, wenn $Q_A(x) < 0$ für alle $x \neq 0$. A heißt indefinit, wenn Vektoren $h_0, k_0 \neq 0$ existieren mit $Q_A(h_0) > 0$ und $Q_A(k_0) < 0$.

Satz. *Sei G eine offene Teilmenge von $\mathbb{R}^p$ und $f \in \mathcal{C}^2(G)$, ferner sei $f'(\xi) = 0$ für ein $\xi \in G$, also ξ ein kritischer Punkt für f. Sei*

$$H_f(\xi) = \begin{pmatrix} \frac{\partial^2 f(\xi)}{\partial x_1^2} & \ldots & \frac{\partial^2 f(\xi)}{\partial x_p \, \partial x_1} \\ \ldots & \ldots & \ldots \\ \frac{\partial^2 f(\xi)}{\partial x_1 \, \partial x_p} & \ldots & \frac{\partial^2 f(\xi)}{\partial x_p^2} \end{pmatrix}.$$

Die symmetrische Matrix $H_f(\xi)$ heißt Hesse'sche Matrix von f an der Stelle ξ. Dann gilt: ist $H_f(\xi)$ positiv definit, so besitzt f an der Stelle ξ ein lokales Minimum; ist $H_f(\xi)$ negativ definit, so besitzt f an der Stelle ξ ein lokales Maximum; ist $H_f(\xi)$ indefinit, so hat f in ξ kein lokales Extremum.

Bemerkung. *Für den Fall $p = 2$ lassen sich die hinreichenden Bedingungen für die Existenz lokaler Extrema wie folgt formulieren:*
sei

$$\Delta = \frac{\partial^2 f(\xi)}{\partial x^2} \frac{\partial^2 f(\xi)}{\partial y^2} - \left(\frac{\partial^2 f(\xi)}{\partial x \, \partial y}\right)^2;$$

wenn $\frac{\partial^2 f(\xi)}{\partial x^2} > 0$ und $\Delta > 0$, dann ist $f(\xi)$ ein lokales Minimum; wenn $\frac{\partial^2 f(\xi)}{\partial x^2} < 0$ und $\Delta > 0$, so ist $f(\xi)$ ein lokales Maximum. Wenn $\Delta < 0$ ist, so liegt kein lokales Extremum vor.
Für den allgemeinen Fall kann man die Theorie der quadratischen Formen verwenden, insbesondere Eigenschaften der sogenannten Hauptminoren der Hesse'schen Matrix (siehe Lineare Algebra).

Kontraktionen und der Satz über inverse Abbildungen

Definition. *Sei* (X, d) *ein metrischer Raum und* $\varphi : X \longrightarrow X$ *eine Abbildung. Existiert ein* $c < 1$ *mit* $d(\varphi(x), \varphi(y)) \leq c\, d(x, y)$ *für jedes* $x, y \in X$, *dann bezeichnet man* φ *als eine Kontraktion von* X *nach* X.

Satz. *Sei* (X, d) *ein volständiger metrischer Raum und* $\varphi : X \longrightarrow X$ *eine Kontraktion. Dann besitzt* φ *genau einen Fixpunkt, d.h. es gibt genau ein* $x \in X$ *mit* $\varphi(x) = x$.

Diese Aussage wird für den Beweis des Satzes über inverse Funktionen verwendet.

Satz. *Sei* E *eine offene Teilmenge des* $\mathbb{R}^n$ *und* $f : E \longrightarrow \mathbb{R}^n$ *eine stetig differenzierbare Funktion,* $f \in \mathcal{C}^1(E)$. *Für ein* $a \in E$ *sei* $f'(a) \in L(\mathbb{R}^n, \mathbb{R}^n)$ *invertierbar und es sei* $b = f(a)$. *(Die Invertierbarkeit von* $f'(a)$ *kann auch dadurch ausgedrückt werden, dass die Determinante der Matrix (Jacobi-Determinante)*

$$[f'(a)] = \begin{bmatrix} \frac{\partial f_1(a)}{\partial x_1} & \dots & \frac{\partial f_1(a)}{\partial x_n} \\ \dots & \dots & \dots \\ \frac{\partial f_n(a)}{\partial x_1} & \dots & \frac{\partial f_n(a)}{\partial x_n} \end{bmatrix}$$

ungleich Null ist.)
Dann gilt:
(i) es existieren offene Mengen U, V *in* $\mathbb{R}^n$, $a \in U$, $b \in V$ *derart, dass* f *auf* U *injektiv ist und* $f(U) = V$ *ist;*
(ii) ist g *die Inverse von* f *definiert auf* V *durch* $g(f(x)) = x$ *für* $x \in U$, *dann ist* $g \in \mathcal{C}^1(V)$ *und es gilt* $g'(y) = [f'(g(y))]^{-1}$ *für* $y \in V$.

Bemerkung. *Schreibt man* $y = f(x)$, *so ist das System*

$$y_i = f_i(x_1, \dots, x_n) \ , \ i = 1, \dots, n$$

nach $x_1, \dots, x_n$ *auflösbar, wenn man* x *und* y *auf hinreichend kleine Umgebungen von* a *und* b *beschränkt. Die Lösungen sind eindeutig bestimmt und stetig differenzierbar.*

Satz. *Ist* $f : E \longrightarrow \mathbb{R}^n$ *eine* $\mathcal{C}^1$*- Abbildung auf der offene Menge* $E \subset \mathbb{R}^n$ *und ist* $f'(x)$ *invertierbar für jedes* $x \in E$, *dann ist* $f(W)$ *eine offene Teilmenge von* $\mathbb{R}^n$ *für jede offene Menge* $W \subset E$. *(*f *ist eine offene Abbildung).*

Bemerkung. *Im letzten Satz besitzt jeder Punkt* $x \in E$ *eine Umgebung, in der* f *injektiv ist;* f *ist lokal injektiv. Die Funktion* $f = (f_1, f_2)$ *mit* $f_1(x, y) = e^x \cos y$ *und* $f_2(x, y) = e^x \sin y$ *ist lokal injektiv, aber nicht global injektiv auf* $\mathbb{R}^2$.

Der Satz über implizite Funktionen

Wann lässt sich die Gleichung $f(x,y)=0$ in einer Umgebung von (a,b) wo $f(a,b)=0$ ist nach y oder nach x auflösen?
Wir behandeln zunächst den Spezialfall einer linearen Abbildung: sei $x=(x_1,\ldots,x_n)\in\mathbb{R}^n$ und $y=(y_1,\ldots,x_m)\in\mathbb{R}^m$, wir schreiben

$$(x,y)=(x_1,\ldots x_n,y_1,\ldots,y_m)\in\mathbb{R}^{n+m}.$$

Sei $A\in L(\mathbb{R}^{n+m},\mathbb{R}^n)$. Dann lässt sich A in zwei lineare Abbidungen $A_x\in L(\mathbb{R}^n,\mathbb{R}^n)$ und $A_y\in L(\mathbb{R}^m,\mathbb{R}^n)$ zerlegen, wobei $A_x(h):=A(h,0)$ für $h\in\mathbb{R}^n$ und $A_y(k):=A(0,k)$ für $k\in\mathbb{R}^m$ ist. Es gilt dann

$$A(h,k)=A_x(h)+A_y(k).$$

Satz. *Ist $A\in L(\mathbb{R}^{n+m},\mathbb{R}^n)$ und ist A_x invertierbar, dann gibt es zu jedem $k\in\mathbb{R}^m$ genau ein $h\in\mathbb{R}^n$ mit $A(h,k)=0$. Dieses h kann mit Hilfe der Formel*

$$h=-(A_x)^{-1}A_y(k)$$

berechnet werden.

Die Gleichung $A(h,k)=0$ kann bei vorgegebenem k auf genau eine Weise nach h aufgelöst werden, und die Lösung h ist eine lineare Funktion von k.

Satz. *Sei E eine offene Teilmenge des $\mathbb{R}^{n+m}$ und $f:E\longrightarrow\mathbb{R}^n$ eine stetig differenzierbare Funktion derart, dass $f(a,b)=0$ für ein $(a,b)\in E$ gilt. Sei $A=f'(a,b)$ und A_x invertierbar. Dann gibt es offene Mengen $U\subset\mathbb{R}^{n+m}$ und $W\subset\mathbb{R}^m$ mit $(a,b)\in U$ und $b\in W$, die die folgende Eigenschaft haben : jedem $y\in W$ entspricht genau ein x mit $(x,y)\in U$ und $f(x,y)=0$. Wird dieses x als $g(y)$ definiert, so ist g eine $\mathcal{C}^1$-Abbildung von W nach $\mathbb{R}^n$ mit $g(b)=a$ und $f(g(y),y)=0$ für $y\in W$, und es gilt*

$$g'(b)=-(A_x)^{-1}\circ A_y.$$

(Die Funktion g wird durch $f(g(y),y)=0$ implizit definiert.)

Bemerkung. *$f(x,y)=0$ ist ein System von n Gleichungen in $n+m$ Variablen*

$$\begin{aligned} f_1(x_1,\ldots,x_n,y_1,\ldots,y_m) &= 0\\ &\ldots\\ f_n(x_1,\ldots,x_n,y_1,\ldots,y_m) &= 0 \end{aligned}$$

Die Invertierbarkeit von A_x bedeutet, dass die $(n\times n)$-Matrix

$$\begin{bmatrix} D_1f_1 & \ldots & D_nf_1\\ \ldots & \ldots & \ldots\\ D_1f_n & \ldots & D_nf_n \end{bmatrix}$$

an der Stelle (a,b) invertierbar ist (Determinante $\neq 0$ hat).

Gilt das obige Gleichungssystem für $x = a$ und $y = b$, so besagt der Satz über implizite Funktionen, dass die $x_1, \dots, x_n$ für jedes y nahe genug bei b durch $y_1, \dots, y_m$ ausgedrückt werden können, und dass die Lösung eine stetig differenzierbare Funktion von y wird.

Extrema mit Nebenbedingungen

Definition. *Sei E eine offene Teilmenge von $\mathbb{R}^p$, und $f : E \longrightarrow \mathbb{R}$ eine reellwertige Funktion, sowie $g : E \longrightarrow \mathbb{R}^q$, $q < p$. Wir sagen f besitzt in $\xi \in E$ ein lokales Maximum bzw. Minimum unter der Nebenbedingung $g(x) = 0$, wenn $\xi \in N = \{x \in E \; : \; g(x) = 0\}$ und eine Umgebung U von ξ existiert, so dass $f(x) \leq f(\xi)$ für alle $x \in U \cap N$ bzw. $f(x) \geq f(\xi)$ für alle $x \in U \cap N$ gilt.*

Der folgende Satz liefert eine notwendige Bedingung für ein Extremum mit Nebenbedingungen:

Satz. *Sei E eine offene Teilmenge von $\mathbb{R}^p$, und $f : E \longrightarrow \mathbb{R}$ eine reellwertige Funktion, sowie $g : E \longrightarrow \mathbb{R}^q$, $q < p$. Beide Funktionen seien stetig differenzierbar auf E und f besitze in $\xi \in E$ ein lokales Extremum unter der Nebenbedingnung $g(x) = 0$. Ferner gebe es in der Matrix*

$$[g'(\xi)] = \begin{bmatrix} \frac{\partial g_1(\xi)}{\partial x_1} & \dots & \frac{\partial g_1(\xi)}{\partial x_p} \\ \dots & \dots & \dots \\ \frac{\partial g_q(\xi)}{\partial x_1} & \dots & \frac{\partial g_q(\xi)}{\partial x_p} \end{bmatrix}$$

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

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

besteht.

Setzt man

$$F(x, \lambda_1, \dots, \lambda_q) = f(x) + \lambda_1 \, g_1(x) + \dots + \lambda_q \, g_q(x)$$

und bildet die partiellen Ableitungen von F nach $x_1, \dots, x_p, \lambda_1, \dots, \lambda_q$ und setzt diese Null, so erhält man die Lagrange'schen Bedingungen des obigen Satzes. Also:

$$\frac{\partial f(x)}{\partial x_k} + \lambda_1 \, \frac{\partial g_1(x)}{\partial x_k} + \dots + \lambda_q \, \frac{\partial g_q(x)}{\partial x_k} = 0 \; , \; k = 1, \dots, p$$

und $g_j(x) = 0$ für $j = 1, \dots, q$. Man löst nun dieses System. Jeder Punkt ξ, dessen Koordinaten $\xi_1, \dots, \xi_p$ den "Anfang" einer Lösung

$x_1, \ldots, x_p, \lambda_1, \ldots, \lambda_q$ bilden und für den mindestens eine q-reihige Unterdeterminante der Matrix $g'(\xi)$ nicht verschwindet, könnte dann Stelle eines lokalen Extremums von f unter der Nebenbedingung $g(x) = 0$ sein. Ob in einem solchen Punkt ξ aber tatsächlich $f(\xi)$ lokal extremal ist, muss in jedem Fall gesondert geprüft werden.

Differentiation von Integralen

Satz. *Sei*
(a) $\varphi : [a,b] \times [c,d] \longrightarrow \mathbb{R}$,
(b) α *monoton wachsend auf* $[a,b]$,
(c) $\varphi^t \in \mathcal{R}(\alpha)$ *auf* $[a,b]$ *für jedes* $t \in [c,d]$, *wobei* $\varphi^t(x) := \varphi(x,t)$,
(d) zu $s \in (c,d)$ *und zu jedem* $\epsilon > 0$ *existiere ein* $\delta > 0$ *derart, dass*

$$|(D_2\varphi)(x,t) - (D_2\varphi)(x,s)| < \epsilon,$$

für alle $x \in [a,b]$ *und für alle* $t \in (s-\delta, s+\delta)$.
Man definiere

$$f(t) = \int_a^b \varphi(x,t)\, d\alpha(x)\ ,\ t \in [c,d].$$

Dann ist $(D_2\varphi)^s \in \mathcal{R}(\alpha)$, *weiters existiert* $f'(s)$, *und es gilt*

$$f'(s) = \int_a^b (D_2\varphi)(x,s)\, d\alpha(x).$$

Bemerkung. *(d) gilt sicher dann, wenn* $D_2\varphi$ *stetig auf dem Rechteck* $[a,b] \times [c,d]$ *ist.*

Satz. *Sei* $Q_0 = (a_0, b_0) \times (\alpha_0, \beta_0) \subset \mathbb{R}^2$ *ein (eventuell unbeschränktes) Rechteck und* $f : Q_0 \longrightarrow \mathbb{R}\ ,\ (x,\lambda) \mapsto f(x,\lambda)$ *eine stetige und nach* λ *stetig differenzierbare Funktion. Seien* $a, b : (\alpha_0, \beta_0) \longrightarrow (a_0, b_0)$ *differenzierbare Funktionen. Dann ist die Funktion*

$$F(\lambda) := \int_{a(\lambda)}^{b(\lambda)} f(x,\lambda)\, dx$$

auf (α_0, β_0) *nach* λ *differenzierbar und es gilt*

$$F'(\lambda) = f(b(\lambda),\lambda)\, b'(\lambda) - f(a(\lambda),\lambda)\, a'(\lambda) + \int_{a(\lambda)}^{b(\lambda)} f_\lambda(x,\lambda)\, dx.$$