7.2 Höhere Ableitungen

Vektorwertige Funktionen sind genau dann differenzierbar, wenn ihre Koordinatenfunktionen differenzierbar sind. Es ist also keine wesentliche Einschränkung, wenn wir in diesem Abschnitt nur reellwertige Funktionen betrachten.

7.2.1 Zweite partielle Ableitungen

Sei $G \subset \mathbb{R}^n$ offen, $\xi \in G$ und $f\colon G \to \mathbb{R}$ in G nach x_k partiell differenzierbar. Ist die partielle Ableitung $D_k f$ in $\xi \in G$ nach x_j differenzierbar, so heißt $D_j D_k f(\xi) := D_j\,(D_k f)\,(\xi)$ eine zweite partielle Ableitung von f in ξ. Andere Schreibweisen sind

$$D_j D_k f(\xi) \;=\; f_{x_k x_j}(\xi) \;=\; \frac{\partial^2 f}{\partial x_j \partial x_k}(\xi) \;=\; \frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_k}\right)(\xi)\ .$$

Für $j = k$ schreibt man auch $D_k^2 f := D_k D_k f$ bzw $\dfrac{\partial^2 f}{\partial x_k^2} = \dfrac{\partial}{\partial x_k}\left(\dfrac{\partial f}{\partial x_k}\right)$.

Im allgemeinen kommt es bei den zweiten Ableitungen auf die Reihenfolge an (Beispiel siehe Aufgabe 7.5.1.A). Es gilt aber

> **Satz von H.A.Schwarz:**
> Sei $G \subset \mathbb{R}^n$ offen und $f\colon G \to \mathbb{R}$ in G nach x_k und x_j differenzierbar.
> f und die partiellen Ableitungen $D_j f$ und $D_k f$ seien in G stetig.
> $D_j f$ sei in G nach x_k differenzierbar und die zweite partielle Ableitung $D_k D_j f$ sei in $\xi \in G$ stetig.
> Dann ist auch $D_k f$ in ξ nach x_j differenzierbar und es gilt

$$D_k D_j f\,(\xi) \;=\; D_j D_k f\,(\xi) \quad \text{bzw} \quad \frac{\partial^2 f}{\partial x_k \partial x_j}(\xi) \;=\; \frac{\partial^2 f}{\partial x_j \partial x_k}(\xi)\ .$$

Etwas schwächer ist die folgende Formulierung:

> Sind f und alle partiellen Ableitungen 1. und 2. Ordnung in G stetig, so darf die Reihenfolge der Differentiation vertauscht werden.

Die folgende Aussage benutzt die 2. (totale) Ableitung (7.2.3):

> Ist $f\colon G \to \mathbb{R}$ in G einmal und an der Stelle $\xi \in G$ zweimal differenzierbar, so sind die zweiten partiellen Ableitungen an der Stelle ξ unabhängig von der Reihenfolge. (Beweis siehe [BF2, S. 125f])

$f\colon G \to \mathbb{R}$ besitze in $\xi \in G$ sämtliche partiellen Ableitungen zweiter Ordnung. Dann heißt

$$\mathbf{H}_f(\xi) \;:=\; \begin{pmatrix} D_1 D_1 f\,(\xi) & \dots & D_1 D_n f\,(\xi) \\ \vdots & \ddots & \vdots \\ D_n D_1 f\,(\xi) & \dots & D_n D_n f\,(\xi) \end{pmatrix}$$

Hesse-Matrix von f an der Stelle ξ. Sie ist symmetrisch, falls die zweiten partiellen Ableitungen unabhängig von der Reihenfolge sind.

7.2.2 Höhere partielle Ableitungen

Höhere partielle Ableitungen werden entsprechend definiert:

$$\begin{aligned} \frac{\partial^k f}{\partial x_{j_k} \dots \partial x_{j_1}}(\xi) &:= \frac{\partial}{\partial x_{j_k}} \left(\frac{\partial^{k-1} f}{\partial x_{j_{k-1}} \dots \partial x_{j_1}} \right)(\xi) \quad \text{bzw} \\ D_{j_k} \dots D_{j_1} f(\xi) &:= D_{j_k} \left(D_{j_{k-1}} \dots D_{j_1} f \right)(\xi) \,. \end{aligned}$$

Man sagt, f ist in $G \subset \mathbb{R}^n$ k-mal *stetig differenzierbar*, bzw f gehört zur Klasse $C^k(G)$, wenn sämtliche partiellen Ableitungen von f bis zur k-ten Ordnung in G existieren und stetig sind. In diesem Fall kommt es bei ihnen nicht auf die Reihenfolge an.

Man sagt, f gehört zur Klasse $C^\infty(G)$, wenn f zur Klasse $C^k(G)$ für alle $k \in \mathbb{N}$ gehört. $C^0(G) := C(G)$ ist die Klasse der stetigen Funktionen auf G.

Multiindexschreibweise

Es ist häufig zweckmäßig, mehrere Indizes $\alpha_j \in \mathbb{N}_0$ zu einem *Multiindex* $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}_0^n$ zusammenzufassen und abkürzend $b_\alpha := b_{\alpha_1,\dots,\alpha_n}$ zu schreiben. In Abschnitt 6.3.7.a hatten wir für Polynome von mehreren Variablen bereits die Schreibweise $x^\alpha := x_1^{\alpha_1} \cdot \ldots \cdot x_n^{\alpha_n}$ eingeführt.

$|\alpha| := \alpha_1 + \ldots + \alpha_n$ heißt *Ordnung* des Multiindex $\alpha \in \mathbb{N}_0^n$. Ferner setzt man

$$\alpha! := \alpha_1! \cdot \ldots \cdot \alpha_n! \qquad \text{und} \qquad \binom{|\alpha|}{\alpha} := \frac{|\alpha|!}{\alpha!} = \frac{(\alpha_1 + \ldots + \alpha_n)!}{\alpha_1! \cdot \ldots \cdot \alpha_n!} \,.$$

$\binom{|\alpha|}{\alpha}$ heißt auch *Polynomialkoeffizient*. Für $n = 2$ sind dies die bekannten Binomialkoeffizienten (1.3.4).

Sei $\alpha = (\alpha_1, \dots, \alpha_n)$ ein Multiindex und A eine Menge mit $|\alpha|$ Elementen. Dann gibt es $\binom{|\alpha|}{\alpha}$ verschiedene Abbildungen $f\colon A \to \{1, \dots, n\}$, die genau α_j Elemente von A auf j abbilden (j=1,. . . ,n) . Außerdem gilt für $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ und $p \in \mathbb{N}$ die sog. *Polynomialformel* :

$$(x_1 + \ldots + x_n)^p = \sum_{|\alpha|=p} \binom{p}{\alpha} x^\alpha = \sum_{|\alpha|=p} \frac{p!}{\alpha_1! \cdot \ldots \cdot \alpha_n!} x_1^{\alpha_1} \cdot \ldots \cdot x_n^{\alpha_n} \,.$$

Für $n = 2$ ist dies die bekannte Binomialformel (1.3.5.4) .

Für Funktionen $f \in C^{|\alpha|}(G)$ werden die partiellen Ableitungen in der Form

$$D^\alpha f := D_1^{\alpha_1} D_2^{\alpha_2} \ldots D_n^{\alpha_n} f := \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \dots \partial x_n^{\alpha_n}} = \frac{\partial^{\alpha_1 + \ldots + \alpha_n} f}{\partial x_1^{\alpha_1} \dots \partial x_n^{\alpha_n}}$$

geschrieben. Nach dem Satz von H.A.Schwarz kommt es dabei nicht auf die Reihenfolge der Ableitungen an.

Beispiel: Für Multiindizes α, β und das Monom $x^\alpha = x_1^{\alpha_1} \cdot \ldots \cdot x_n^{\alpha_n}$ gilt

$$D^\beta x^\alpha \;=\; D_1^{\beta_1} D_2^{\beta_2} \ldots D_n^{\beta_n} (x_1^{\alpha_1} x_2^{\alpha_2} \cdot \ldots \cdot x_n^{\alpha_n}) \;=\; \begin{cases} \dfrac{\alpha!}{(\alpha-\beta)!}\, x^{\alpha-\beta} & \text{für } \beta \le \alpha \\ 0 & \text{sonst .} \end{cases}$$

Dabei wurde die durch $\beta = (\beta_1, \ldots, \beta_n) \le \alpha = (\alpha_1, \ldots, \alpha_n) \;:\Longleftrightarrow\; \beta_j \le \alpha_j$ für $j = 1, \ldots, n$ definierte Halbordnung zwischen den Multiindizes verwendet.

7.2.3 Höhere Ableitungen

Ist $f\colon G \to \mathbb{R}$ in G differenzierbar, so ist die Ableitung f' von f eine Abbildung $f'\colon G \to \mathrm{Hom}\,(\mathbb{R}^n, \mathbb{R})$ von G in den Raum der Linearformen auf dem $\mathbb{R}^n$. Ihre Koordinatenfunktionen bzgl der kanonischen dualen Basis sind die partiellen Ableitungen $D_k f$.

Die erste Ableitung f' wiederum ist differenzierbar in einem Punkt $\xi \in G$, wenn sie (bzw ihre Koordinatenfunktionen) bei ξ linear approximierbar ist. In diesem Fall schreibt man $f''(\xi) := (f')'(\xi)$ und nennt $f''(\xi)$ die zweite Ableitung von f im Punkt ξ. f ist also in ξ genau dann zweimal (total) differenzierbar, wenn sie in einer Umgebung von ξ einmal (total) differenzierbar ist, und ihre partiellen Ableitungen in ξ noch mal (total) differenzierbar sind.

Ist f in G zweimal differenzierbar, so ist die zweite Ableitung f'' von f eine Abbildung $f''\colon G \to \mathrm{Hom}\,(\mathbb{R}^n, \mathrm{Hom}\,(\mathbb{R}^n, \mathbb{R}))$.
Der Raum $\mathrm{Hom}\,(\mathbb{R}^n, \mathrm{Hom}\,(\mathbb{R}^n, \mathbb{R}))$ wird durch die Zuordnung

$$\varphi \mapsto \widehat{\varphi} \quad , \qquad \widehat{\varphi}(x, y) \;:=\; \varphi(x)(y)$$

isomorph auf den Raum $\mathrm{Bil}\,(\mathbb{R}^n)$ der Bilinearformen $\widehat{\varphi}\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ abgebildet. Man interpretiert daher die zweite Ableitung als Abbildung
$f''\colon G \to \mathrm{Bil}\,(\mathbb{R}^n)$ und schreibt übersichtlicher
$f''(\xi, h, k) := f''(\xi)(h, k) := (f''(\xi)(h))(k)$. Es gilt:

> Ist $f\colon G \to \mathbb{R}$ in G einmal und an der Stelle $\xi \in G$ zweimal differenzierbar, so ist $f''(\xi)$ eine symmetrische Bilinearform. (Beweis siehe [BF2, S. 125f])

$\mathrm{Bil}\,(\mathbb{R}^n)$ ist wiederum auf kanonische Weise isomorph zum Raum $\mathbb{R}^{n\times n}$ der reellen $(n\times n)$-Matrizen. Bzgl der kanonischen Basen entspricht $f''(\xi)$ dabei die Hesse-Matrix $\mathbf{H}_f(\xi)$. Es ist

$$f''(\xi, h, k) \;=\; f''(\xi)(h, k) \;=\; h^\top \left(\mathbf{H}_f(\xi)\right) k \;.$$

Höhere Ableitungen werden rekursiv analog wie die zweite definiert:
Die k-te Ableitung von f ist eine Abbildung von G in den Raum der symmetrischen k-Linearformen $\varphi\colon (\mathbb{R}^n)^k \to \mathbb{R}$. Diese Ableitung ist differenzierbar

in $\xi \in G$, wenn sie bei ξ linear approximierbar ist. In diesem Fall schreibt man $f^{(k+1)}(\xi) := (f^{(k)})'(\xi)$ und faßt $f^{(k+1)}(\xi)$ vermöge

$$f^{(k+1)}(\xi)(h_1, \ldots, h_{k+1}) \;:=\; \left[f^{(k+1)}(\xi)(h_{k+1})\right](h_1, \ldots, h_k)$$

als $(k+1)$-Linearform auf.

f ist genau dann in ξ $(k+1)$-mal (total) differenzierbar, wenn f in einer Umgebung von ξ k-mal (total) differenzierbar ist, und ihre k-ten partiellen Ableitungen in ξ noch einmal (total) differenzierbar sind.

f heißt *(k-mal) stetig differenzierbar* auf G, wenn f auf G k-mal differenzierbar ist, und die k-te Ableitung $f^{(k)}$ als Funktion von G in den Raum der k-Linearformen stetig ist. Dies ist genau dann der Fall, wenn alle ihre partiellen Ableitungen bis zur k-ten Ordnung auf G existieren und stetig sind.

Die Koordinaten der k-ten Ableitung bzgl der kanonischen Basis des Raums der k-Linearformen sind die k-ten partiellen Ableitungen. Anders ausgedrückt: Für Vektoren $h^{(j)} = (h_1^{(j)}, \ldots, h_n^{(j)}) \in \mathrm{I\!R}^n$ gilt

$$f^{(k)}(\xi)(h^{(1)}, \ldots, h^{(k)}) \;=\; \sum_{j_1=1}^{n} \cdots \sum_{j_k=1}^{n} \frac{\partial^k f\,(\xi)}{\partial x_{j_1} ... \partial x_{j_k}}\, h_{j_1}^{(1)} \cdot \ldots \cdot h_{j_k}^{(k)}\ .$$

Sind alle Argumente gleich, etwa $h = (h_1, \ldots, h_n) = h^{(1)} = \ldots = h^{(k)}$, so kann man wegen der Symmetrie der k-Linearform $f^{(k)}(\xi)$ einige dieser n^k Summanden zusammenfassen. Für $|\alpha| = k$ liefern insgesamt $\dfrac{|\alpha|!}{\alpha!} = \dfrac{k!}{\alpha_1! \cdot \ldots \cdot \alpha_n!}$ Summanden denselben Beitrag $D_1^{\alpha_1} \ldots D_n^{\alpha_n} f(\xi) h_1^{\alpha_1} \cdot \ldots \cdot h_n^{\alpha_n}$. Unter Verwendung der Multiindexschreibweise erhält man daher

$$\frac{1}{k!} f^{(k)}(\xi)(h, \ldots, h) \;=\; \sum_{|\alpha|=k} \frac{D^\alpha f(\xi)}{\alpha!}\, h^\alpha\ .$$

7.2.4 Taylorformel

Die Schwierigkeiten der mehrdimensionalen Taylorformel liegen mehr in der Schreibweise, als im Inhalt. Um zu einigermaßen übersichtlichen Formeln zu kommen, kann man mit der Multiindexschreibweise oder mit den höheren (totalen) Ableitungen arbeiten (7.2.3). Wir geben beide Varianten an.

a) Taylorpolynom

Seien $G \subset \mathrm{I\!R}^n$ offen, $x, \xi \in G$, $h := x - \xi$ und $f\colon G \to \mathrm{I\!R}$ m-mal stetig differenzierbar. Dann heißt das Polynom

$$T_m(x) \;=\; T_m(\xi + h) \;=\; \sum_{|\alpha| \le m} \frac{D^\alpha f(\xi)}{\alpha!} (x - \xi)^\alpha \;=\; \sum_{k=0}^{m} \frac{1}{k!}\, f^{(k)}(\xi, \underbrace{h, \ldots, h}_{k-mal})$$

das m-te *Taylor-Polynom* von f mit Entwicklungspunkt ξ. Die Differenz

$$R_m(x) \;:=\; R_{f,\xi,m}(x) \;:=\; f(x) - T_m(x)$$

heißt das m-te *Restglied* der Taylorentwicklung von f um ξ.

Das 0. Taylorpolynom $T_0(x)$ ist konstant gleich dem Funktionswert $f(\xi)$ im Entwicklungspunkt.

Das 1. Taylorpolynom ist die *lineare Näherung* von f. Es ist

$$\begin{aligned} T_1(x) \;&:=\; T_0(x) + \sum_{|\alpha|=1} \frac{D^\alpha f(\xi)}{\alpha!}(x-\xi)^\alpha \;=\; f(\xi) + \langle \operatorname{grad} f(\xi), (x-\xi)\rangle \\ &=\; f(\xi) + [D_1 f(\xi)(x_1-\xi_1) + \ldots + D_n f(\xi)(x_n-\xi_n)] \end{aligned}$$

Das 2. Taylorpolynom ist die *quadratische Näherung* von f. Es ist

$$\begin{aligned} T_2(x) \;&:=\; T_1(x) + \sum_{|\alpha|=2} \frac{D^\alpha f(\xi)}{\alpha!}(x-\xi)^\alpha \\ &=\; T_1(x) + \frac{1}{2}\sum_{j,k=1}^{n} D_j D_k f(\xi)(x_j-\xi_j)(x_k-\xi_k) \\ &=\; T_1(x) + \frac{1}{2}\,(x-\xi)^\top\,\mathbf{H}_f(\xi)\,(x-\xi)\;. \end{aligned}$$

Dabei ist $\mathbf{H}_f(\xi)$ die aus den zweiten partiellen Ableitungen gebildete (symmetrische) Hesse-Matrix von f an der Stelle ξ (siehe 7.2.1).

Das 3. Taylorpolynom T_3 einer Funktion $f\colon \mathbb{R}^2 \to \mathbb{R}$ mit dem Entwicklungspunkt $\xi = 0 = (0,0)$ ist z.B. ($\vec{x} := (x,y)$, Funktionswert und Ableitungen ($f_x := D_1 f$ usw) sind jeweils an der Stelle 0 zu berechnen) :

$$\begin{aligned} T_3(x,y) \;&=\; T_3(\vec{x}) \;=\; f + \langle \operatorname{grad} f, \vec{x}\rangle + \frac{1}{2}\,\vec{x}^\top\,\mathbf{H}_f\,\vec{x} + \sum_{|\alpha|=3}\frac{D^\alpha f}{\alpha!}\,\vec{x}^\alpha \\ &=\; f + f_x x + f_y y + \frac{1}{2}\left[f_{xx}x^2 + f_{yy}y^2\right] + f_{xy}xy \\ &\qquad + \frac{1}{6}\left[f_{xxx}x^3 + f_{yyy}y^3\right] + \frac{1}{2}\left[f_{xxy}x^2y + f_{yyx}y^2x\right]\;. \end{aligned}$$

Das m-te Taylorpolynom ist dasjenige eindeutig bestimmte Polynom vom Grad $\leq m$, das an der Stelle ξ mit f einschließlich aller partiellen Ableitungen bis zur m-ten Ordnung übereinstimmt. Es gilt $\lim\limits_{x\to\xi} \frac{f(x)-T_m(x)}{\|x-\xi\|^m} = 0$.

Interessant ist die Frage, ob die Taylorpolynome in irgendeinem Sinne die Ausgangsfunktion f approximieren. Trivialerweise gilt (für $f \in C^\infty(G)$)

$$f(x) \;=\; \lim_{m\to\infty} T_m(x) \quad\Longleftrightarrow\quad R_m(x) \;\to\; 0 \quad \text{für}\;\; m\to\infty\;.$$

Um das Restglied R_m abzuschätzen braucht man sogenannte

b) Restglieddarstellungen

Seien $G \subset \mathbb{R}^n$ offen, $x, \xi \in G$, $h = x - \xi$ und $f\colon G \to \mathbb{R}$ $(m+1)$-mal stetig differenzierbar, also aus der Klasse $C^{m+1}(G)$. Außerdem liege die Strecke $\overline{x,\xi}$ in G. Dann gilt für das m-te Restglied $R_m = f - T_{f,\xi,m}$:

1) *Restglieddarstellung von Lagrange:* Es gibt ein $\vartheta \in]0,1[$ mit

$$\begin{aligned} R_m(x) &= R_m(\xi + h) = \sum_{|\alpha|=m+1} \frac{D^\alpha f(\xi+\vartheta h)}{\alpha!}\, h^\alpha \\ &= \frac{1}{(m+1)!} f^{(m+1)}(\,\xi + \vartheta h\,, \underbrace{h\,, \ldots\,, h}_{(m+1)\text{-mal}}) \,. \end{aligned}$$

2) *Integraldarstellung des Restglieds:*

$$\begin{aligned} R_m(x) &= R_m(\xi + h) = (m+1) \int_0^1 (1-t)^m \left(\sum_{|\alpha|=m+1} \frac{D^\alpha f(\xi+th)}{\alpha!}\, h^\alpha \right) dt \\ &= \frac{1}{m!} \int_0^1 (1-t)^m f^{(m+1)}(\,\xi + th\,, \underbrace{h\,, \ldots\,, h}_{(m+1)\text{-mal}})\, dt \,. \end{aligned}$$

c) Taylorreihe

Ist f in G unendlich oft differenzierbar, so heißt

$$T_f(x) := T_f(\xi + h) := \sum_{|\alpha|=0}^{\infty} \frac{D^\alpha f(\xi)}{\alpha!}(x-\xi)^\alpha = \sum_{k=0}^{\infty} \frac{1}{k!} f^{(k)}(\xi, \underbrace{h, \ldots, h}_{k-mal})$$

die *Taylorreihe* von f um ξ.

Taylorreihen sind Potenzreihen in mehreren Variablen (siehe Abschnitt 6.3.7.b). Wie bei einer Variablen ist das zentrale Problem:

Wann konvergiert die Taylorreihe einer Funktion gegen diese Funktion?

Die Antwort ist einfach, wenn die Funktion durch eine Potenzreihe beschrieben wird (wir setzen o.B.d.A. $\xi := 0$, $h := x - \xi = x$) :

> Konvergiert z.B. die Potenzreihe $f(x) := \sum_{\alpha \in \mathbb{N}_0^n} a_\alpha x^\alpha$ im offenen Würfel $W_r^0 :=]-r, r[^n$ absolut, so ist die dargestellte Funktion $f(x)$ dort beliebig oft stetig differenzierbar. Ihre Ableitungen erhält man durch gliedweise Differentiation und die definierende Reihe ist die Taylorreihe von f um 0.
> *Kurz:* Jede konvergente Potenzreihe ist die Taylorreihe der dargestellten Funktion im Entwicklungspunkt.

Manchmal kann man eine Potenzreihenentwicklung von f durch Umformungen und Zurückführen auf bekannte Reihen erhalten. Daß diese dann die Taylorreihe von f ist, garantiert der obige Satz.

Achtung: Nicht jede unendlich oft differenzierbare Funktion wird durch ihre Taylorreihe dargestellt. Ein Beispiel finden Sie in Abschnitt 4.4.3 . Bei mehreren Variablen spielt evt auch noch die Reihenfolge der Summation eine Rolle.

Näheres zu diesem Thema siehe z.B. [WA2, S.96] . Beispiele finden Sie in Abschnitt 7.5.3 .