VII Differentialrechnung mehrerer Veränderlicher

In diesem Kapitel erweitern wir die Differentialrechnung von Funktionen einer Variablen auf solche mit mehreren Veränderlichen. Wiederum lassen wir uns von der zentralen Idee der linearen Approximierbarkeit leiten. Im Vergleich zu unseren bisherigen Untersuchungen ist die Situation im Falle von Funktionen mehrerer Variablen nun hingegen deutlich komplizierter, da die linearen Abbildungen in der mehrdimensionalen Situation eine deutlich reichhalterigere Struktur besitzen als dies bei nur einer Variablen der Fall ist.

Wir beginnen in Abschnitt 1 mit dem Begriff der Differenzierbarkeit.

1 Differenzierbare Abbildungen

In diesem Abschnitt betrachten wir $\mathbb{R}^n$ und $\mathbb{R}^m$ als normierte Vektoräume und $f : U \to \mathbb{R}^m$ eine Abbildung definiert auf einer offenen Menge $U \subset \mathbb{R}^n$. Von besonderer Wichtigkeit wird dann der Vektorraum $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$, der Raum aller linearen Abbildungen von $\mathbb{R}^n$ in $\mathbb{R}^m$, sein, welchen wir mit der in Abschnitt VI.2 definierten Operatornorm versehen. Die Betrachtungen in Abschnitt VI.2 implizieren, aufgrund der Endlichkeit der Dimensionen, dass jede lineare Abbildung $\mathbb{R}^n \to \mathbb{R}^m$ stetig und dass $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ ein Banachraum ist.

Im Folgenden definieren wir die Differenzierbarkeit einer Funktion in einem Punkt als Approximierbarkeit durch eine lineare Abbildung.

1.1 Definition. Es seien $U \subset \mathbb{R}^n$, $x_0 \in U$ ein innerer Punkt von U und $f : U \to \mathbb{R}^m$ eine Funktion. Dann heißt f *differenzierbar* in $x_0 \in U$, falls Abbildungen $A \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ und $r : U \to \mathbb{R}^m$ existieren mit

$$f(x) = f(x_0) + A(x - x_0) + r(x), \quad x \in U,$$

und

$$\lim_{x \to x_0} \frac{r(x)}{\|x - x_0\|} = 0$$

gilt.

1.2 Bemerkungen. a) Als Norm auf $\mathbb{R}^n$ wählen wir die Euklidische Norm. Wir notieren jedoch, dass die obige Definition nach Theorem VI.1.10 unabhängig von der gewählten Norm ist.
b) Gilt $n = m = 1$, so ist die obige Definition konsistent mit derjenigen aus Kapitel IV.1.
c) Ist $f : U \subset \mathbb{R}^n \to \mathbb{R}^m$ in $x_0 \in U$ differenzierbar, so ist die lineare Abbildung A in Definition 1.1 eindeutig bestimmt, vgl. die Übungen.
d) Die eindeutig bestimmte Abbildung A aus Defintion 1.1 heißt die *Ableitung* oder das *Differential* von f in x_0. Wir schreiben

$$A = f'(x_0) \quad \text{ oder } A = Df(x_0).$$

e) Wählt man in $\mathbb{R}^n$ und $\mathbb{R}^m$ kanonische Basen, so wird $A \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ durch eine Matrix $(a_{ij})_{m\times n}$ beschrieben. Für $x \in \mathbb{R}^n$ gilt dann

$$Ax = \begin{pmatrix} a_{11} & \cdot \ \cdot & a_{1n} \\ \cdot & & \\ \cdot & & \\ \cdot & & \\ a_{m1} & \cdot \ \cdot & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{pmatrix}.$$

Im Folgenden identifizieren wir stillschweigend eine lineare Abbildung A mit der sie darstellenden Matrix (a_{ij}).

1.3 Beispiel. Es sei $B = (b_{ij}) \in M_n(\mathbb{R})$ eine symmetrische Matrix und $f : \mathbb{R}^n \to \mathbb{R}$ definiert durch

$$f(x) := (x|Bx) := \sum_{i,j=1}^n b_{ij}x_ix_j,$$

wobei $(\cdot|\cdot)$ das Skalarprodukt auf $\mathbb{R}^n$ bezeichnet. Dann gilt für $x_0, h \in \mathbb{R}^n$ und $x = x_0 + h \in \mathbb{R}^n$

$$\begin{aligned} f(x) = f(x_0+h) &= (x_0+h|Bx_0+Bh) = (x_0|Bx_0) + \underbrace{(x_0|Bh) + (h|Bx_0)}_{=2(Bx_0|h)} + (h|Bh) \\ &= f(x_0) + 2(Bx_0|x-x_0) + \underbrace{(x-x_0|B(x-x_0))}_{=r(x)} \\ &= f(x_0) + (2Bx_0)^T \cdot (x-x_0) + r(x). \end{aligned}$$

Mit der Cauchy-Schwarzschen Ungleichung folgt

$$|r(x)| \leq \|x-x_0\| \cdot \|B(x-x_0)\|$$

und es gilt
$$\frac{|r(x)|}{\|x-x_0\|} \leq \|B(x-x_0)\| \overset{x\to x_0}{\longrightarrow} 0,$$
da $B:\mathbb{R}^n \to \mathbb{R}^n$ stetig ist; vgl. die Übungen. Daher ist f in $x_0 \in \mathbb{R}^n$ differenzierbar und es gilt
$$f'(x_0) = (2Bx_0)^T \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}) = (\mathbb{R}^n)'.$$

1.4 Satz. *Ist $U \subset \mathbb{R}^n$ offen und $f: U \to \mathbb{R}^m$ eine in $x_0 \in U$ differenzierbare Funktion, so ist f in x_0 stetig.*

Beweis. Für $x \in U$ gilt
$$f(x) = f(x_0) + f'(x_0)\cdot(x-x_0) + r(x)$$
Da die lienare Abbildung $f'(x_0)$ nach Satz VI.2.6 stetig ist und da $\lim_{x\to x_0} r(x) = 0$ gilt, folgt $\lim_{x\to x_0} f(x) = f(x_0)$, welches nach Satz VI.2.4 bedeutet, dass f in x_0 stetig ist.

□

Betrachtet man eine in $x_0 \in U$ differenzierbare Funktion $f: U \to \mathbb{R}^m$, so interessiert man sich natürlich für die Frage, wie man die Ableitung $f'(x_0) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ konkret berechnen kann. Wir verfolgen bei der Beantwortung dieser Frage die folgende Idee: Da $f'(x_0)$ linear ist, genügt es $f'(x_0)$ auf einer Basis $\{v_1, \ldots, v_n\}$ des $\mathbb{R}^n$ zu kennen. Wir berechnen daher zunächst $f'(x_0)v$ für ein $v \in \mathbb{R}^n$ mit $v \neq 0$.
Hierzu setzen wir $x = x_0 + tv$ mit $t \in \mathbb{R}$. Da U offen ist, existiert ein $\delta > 0$, so dass $x \in U$ ist für alle t mit $|t| < \delta$. Daher gilt mit den Bezeichnungen aus Definition 1.1
$$f'(x_0)v = \frac{f(x_0+tv) - f(x_0)}{t} - \frac{r(x)}{t}.$$
Da $\lim_{t\to 0} \frac{r(x)}{t} = 0$ gilt, erhalten wir
$$f'(x_0)v = \lim_{t\to 0} \frac{f(x_0+tv) - f(x_0)}{t}.$$
Dies motiviert die folgende Definition.

1.5 Definition. Es sei $U \subset \mathbb{R}^n$ offen, $f: U \to \mathbb{R}^m$ eine Funktion, $x_0 \in U$ und $v \in \mathbb{R}^n \setminus \{0\}$. Existiert
$$D_v f(x_0) := \lim_{t\to 0} \frac{f(x_0+tv) - f(x_0)}{t} \in \mathbb{R}^m,$$
so heißt $D_v f(x_0)$ die *Richtungsableitung* von f in x_0 in Richtung v.

1.6 Satz. *Es sei $f : U \subset \mathbb{R}^n \to \mathbb{R}^m$ eine differenzierbare Funktion in $x_0 \in U$. Dann existiert $D_v f(x_0)$ für alle $v \in \mathbb{R}^n \setminus \{0\}$ und es gilt*

$$D_v f(x_0) = Df(x_0) \cdot v.$$

Beweis. Für $v \neq 0$ gilt

$$f(x_0 + tv) = f(x_0) + Df(x_0)(tv) + r(x_0 + tv)$$

mit $\lim_{t\to 0} \frac{r(x_0+tv)}{||tv||} = 0$. Daher gilt

$$\frac{f(x_0 + tv) - f(x_0)}{t} = Df(x_0)v + \frac{r(x_0 + tv)}{t},$$

welches für $t \to 0$ die Behaupting impliziert.

□

Es sei an dieser Stelle angemerkt, dass die Umkehrung des obigen Satzes im Allgemeinen falsch ist. Betrachte hierzu zum Beispiel die Funktion $f : \mathbb{R}^2 \to \mathbb{R}$ gegeben durch

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

Wir haben in Definition 1.5 die Richtungsableitung einer Funktion bezüglich einer beliebigen Richtung definiert. Die Ableitung in Richtung der Koordinatenachsen ist von besonderer Wichtigkeit.

1.7 Definition. a) Für die Ableitung in Richtung der Koordinatenachsen e_j für $j = 1, \ldots, n$ schreibt man

$$\partial_j f(x_0) := \frac{\partial f}{\partial x_j}(x_0) := D_{e_j} f(x_0) = \lim_{t\to 0} \frac{f(x_0 + te_j) - f(x_0)}{t}, \quad 1 \leq j \leq n,$$

und nennt $\partial_j f(x_0)$ die *partielle Ableitung* von f in x_0 bezüglich e_j.
b) Eine Funktion $f : U \subset \mathbb{R}^n \to \mathbb{R}^m$ heißt in x_0 *partiell differenzierbar*, wenn alle partiellen Ableitungen $\partial_1 f(x_0), \ldots, \partial_n f(x_0)$ in x_0 existieren; analog heißt f in x_0 *stetig partiell differenzierbar*, falls alle partiellen Ableitungen $\partial_j f$ in x_0 existieren und stetig sind.

Existiert $\partial_j f(x_0)$ für ein $j \in 1, \ldots, n$, so gilt

$$\partial_j f(x_0) = \lim_{h\to 0} \frac{1}{h}[f(x_0^1, \ldots, x_0^{j-1}, x_0^j + h, x_0^{j+1}, \ldots, x_0^n) - f(x_0)]$$

d.h. f ist genau dann in x_0 partiell bezüglich x_j differenzierbar, wenn die Abbildung

$$t \mapsto f(x_0^1, \ldots, x_0^{j-1}, t, x_0^{j+1}, \ldots, x_0^n)$$

als Funktion einer Variablen in x_0^j differenzierbar ist.

Wählt man in $\mathbb{R}^n$ und $\mathbb{R}^m$ die kanonischen Basen und identifiziert $A \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ mit der Matrix $(a_{ij})_{m\times n}$, so erhält man die folgende Darstellung der Ableitung.

1.8 Satz. *Ist $f = (f_1, \ldots, f_m) : U \subset \mathbb{R}^n \to \mathbb{R}^m$ eine in $x_0 \in U$ differenzierbare Funktion, so gilt*

$$f'(x_0) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x_0) & \cdots & \frac{\partial f_1}{\partial x_n}(x_0) \\ \vdots & & \\ \vdots & & \\ \frac{\partial f_m}{\partial x_1}(x_0) & \cdots & \frac{\partial f_m}{\partial x_n}(x_0) \end{pmatrix}_{m\times n}.$$

Beweis. a) Für $f_1 : U \subset \mathbb{R}^n \to \mathbb{R}$ und $h = (h_1, \ldots, h_n) \in \mathbb{R}^n$ mit $h = \sum_{j=1}^n h_j e_j$ gilt die Darstellung

$$Df_1(x_0)h = \sum_{j=1}^n h_j Df_1(x_0)e_j \overset{1.6}{=} \sum_{j=1}^n \partial_j f_1(x_0)h_j,$$

und ebenso gilt eine analoge Darstellung für $f_2, \ldots, f_m$.
b) Die obige Darstellung folgt nun, da die Funktion $f = (f_1, \ldots, f_m)$ in x_0 genau dann differenzierbar ist, wenn jede Koordinatenfunktion f_j in x_0 differenzierbar ist.

□

1.9 Definition. a) Die in Satz 1.8 definierte Matrix heißt *Jacobimatrix* oder *Funktionalmatrix* von f in x_0.
b) Gilt $m = 1$, so heißt

$$\mathrm{grad} f(x_0) := \left(\frac{\partial f}{\partial x_1}(x_0), \ldots, \frac{\partial f}{\partial x_n}(x_0)\right),$$

der *Gradient* von f in x_0. Letzterer wird auch mit $\nabla f(x_0) := \mathrm{grad}\, f(x_0)$ bezeichnet.

1.10 Bemerkung. Ist $f : U \subset \mathbb{R}^n \to \mathbb{R}$ in $x_0 \in U$ differenzierbar, so zeigt $\nabla f(x_0)$ in Richtung des steilsten Anstiegs und $-\nabla f(x_0)$ in Richtung des steilsten Abfalls von f. Dies folgt aus der Cauchy-Schwarzschen Ungleichung, denn für $v \in \mathbb{R}^n \backslash \{0\}$ gilt

$$D_v f(x_0) = Df(x_0) \cdot v = (\mathrm{grad} f(x_0) | v) \leq \|\mathrm{grad} f(x_0)\| \|v\|$$

und das Gleichheitszeichen gilt genau dann, wenn $\mathrm{grad} f(x_0) = \lambda v$ für ein $\lambda \geq 0$ ist.

Betrachtet man zum Beispiel die Funktion $f : \mathbb{R}^3 \to \mathbb{R}$ gegeben durch

$$f(x, y, z) := x^2 \sin \frac{y}{2} + e^{3z},$$

so ist der Gradient von f gegeben durch

$$\operatorname{grad} f(x, y, z) = (2x \sin \frac{y}{2}, \frac{x^2}{2} \cos \frac{y}{2}, 3e^{3z}).$$

Im Folgenden wollen wir Kriterien für die Differenzierbarkeit von Funktionen $f = (f_1, \ldots, f_m) : U \subset \mathbb{R}^n \to \mathbb{R}^m$ im Punkte $x_0 \in U$ entwickeln, die einfacher handzuhaben sind als Definition 1.1. Notwendigerweise müssen zunächst alle partiellen Ableitungen

$$\frac{\partial f_k}{\partial x_j}(x_0), \quad j = 1, \ldots, n; k = 1, \ldots, m,$$

existieren, ansonsten wäre f in x_0 nicht differenzierbar. Ferner ist es für die Differenzierbarkeit einer Funktion f ebenfalls notwendig, dass

$$\lim_{x \to x_0} \frac{r(x)}{\|x - x_0\|} = 0 \quad \text{für } r(x) := f(x) - f(x_0) - A(x - x_0)$$

gilt, wobei $A = f'(x_0)$ gesetzt ist.
Es ist interessant zu bemerken, dass die Existenz aller Richtungsableitungen $D_v f(x_0)$ für $v \in \mathbb{R}^n \backslash \{0\}$ noch nicht einmal die Stetigkeit von f in x_0 impliziert. Ein Gegenbeispiel hierfür ist durch die Funktion

$$f : \mathbb{R}^2 \to \mathbb{R}, f(x, y) = \begin{cases} 1, & \text{für } 0 < y < x^2 \\ 0, & \text{sonst} \end{cases}$$

gegeben. Dann existiert $D_v f(0, 0)$ für alle $v \in \mathbb{R}^2$ mit $v \neq 0$, aber f ist nicht stetig in $(0, 0)$; vgl. auch die Übungen.
Es gilt jedoch das folgende Kriterium. Hierbei nennen wir eine differenzierbare Funktion $f : U \to \mathbb{R}^m$ auf einer offenen Menge $U \subset \mathbb{R}^n$ *stetig differenzierbar* in $x_0 \in U$, falls ihre Ableitung $Df : U \to \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m), x \mapsto Df(x)$ in x_0 stetig ist.

1.11 Satz. *Es seien $U \subset \mathbb{R}^n$ offen, $x_0 \in U$ und $f : U \to \mathbb{R}^m$ eine Funktion. Die folgenden Aussagen sind äquivalent:*

i) f ist stetig differenzierbar in x_0.

ii) Alle partiellen Ableitungen $\frac{\partial f_j}{\partial x_i}(x_0)$ $j = 1, \ldots, m,\ i = 1, \ldots, n$ existieren und sind stetig in x_0.

Beweis.
i) $\Rightarrow$ ii): folgt direkt aus Satz 1.6.
ii) $\Rightarrow$ i): Wir bemerken zunächst, dass f genauu dann in x_0 differenzierbar ist, falls alle Funktionen $f_1, \ldots, f_m$ in x_0 differenzierbar sind. Wir betrachten daher oBdA Funktionen $f : U \subset \mathbb{R}^n \to \mathbb{R}$.
Für $h = (h_1, \ldots, h_n) \in \mathbb{R}^n$ setzen wir

$$z_0 := x_0, z_1 := z_0 + h_1 e_1, z_2 := z_1 + h_2 e_2, \ldots, z_n := z_{n-1} + h_n e_n = x_0 + h.$$

Dann ist $\|x_0 - z_j\| \leq \|h\|$ für $j = 0, \ldots, n$. Es gilt also $z_j \in U$ für alle $j = 0, \ldots, n$, falls h nur genügend klein ist und der Mittelwertsatz aus Kapitel IV.2 impliziert

$$\begin{aligned} f(x_0+h) - f(x_0) &= f(z_n) - f(z_{n-1}) + f(z_{n-1}) - f(z_{n-2}) + \ldots + f(z_1) - f(z_0) \\ &= \frac{\partial f}{\partial x_n}(\xi_n) \cdot h_n + \frac{\partial f}{\partial x_{n-1}}(\xi_{n-1}) h_{n-1} + \ldots + \frac{\partial f}{\partial x_1}(\xi_1) h_1 \qquad (1.1) \end{aligned}$$

für geeignete $\xi_j \in (z_{j-1}, z_j)$.
Deswegen gilt

$$|f(x_0+h) - f(x_0) - \text{grad } f(x_0) \cdot h| \leq \sum_{j=1}^{n} |\frac{\partial f}{\partial x_j}(\xi_j) - \frac{\partial f}{\partial x_j}(x_0)| \cdot |h_j|$$

und somit

$$\frac{1}{\|h\|} |f(x_0+h) - f(x_0) - \text{grad } f(x_0) \cdot h| \leq \sum_{j=1}^{n} \left| \frac{\partial f(\xi_j)}{\partial x_j} - \frac{\partial f}{\partial x_j}(x_0) \right| \overset{h \to 0}{\longrightarrow} 0,$$

da $\frac{\partial f}{\partial x_j}$ für alle $j = 1, \ldots, n$ in x_0 stetig ist.

□

1.12 Bemerkung. Die Gleichung (1.1) impliziert unmittelbar das folgende Resultat: Sind alle partiellen Ableitungen $\frac{\partial f_k}{\partial x_j} : U \to \mathbb{R}$ für alle $j = 1, \ldots, n$, $k = 1, \ldots, m$ beschränkt in einer Umgebung von x_0, so ist f stetig in x_0.

Zum Abschluss dieses Abschnitts fassen wir die bisher diskutierten Ergebnisse graphisch zusammen. Für eine Funktion $f : U \subset \mathbb{R}^n \to \mathbb{R}^m$ gilt:

$$\begin{array}{ccc}
f \text{ stetig differenzierbar} & \Leftrightarrow & f \text{ stetig partiell differenzierbar} \\
\Downarrow & & \Downarrow \\
f \text{ differenzierbar} & & f \text{ part. diff'bar mit lokal beschr. part. Abl.} \\
\Downarrow & & \Downarrow \\
\text{alle Richtungsableitungen } D_v f \text{ existieren} & & f \text{ ist stetig} \\
\Downarrow & & \Downarrow \\
f \text{ partiell differenzierbar} & \Rightarrow & f \text{ stetig bzgl. jeder Koordinate}
\end{array}$$