Differenzierbarkeit

Wir betrachten zuerst die Differenzierbarkeit reellwertiger Funktionen.

Definition. Sei $f : \mathbb{R}^n \to \mathbb{R}$ und $x^0 \in D(f)$ ein innerer Punkt.

Dann heißt f **differenzierbar** an x^0 , wenn es einen Vektor $\vec{c} = (c_1, ..., c_n)$ und eine auf einer Umgebung $U(x^0)$ definierte Funktion $f_0(x)$ gibt, sodass

- $f(x) = f(x^0) + \vec{c}(\vec{x} - \vec{x}^0) + \|x - x^0\| f_0(x)$
- $\lim\limits_{x \to x^0} f_0(x) = 0$.

Die lineare Approximation $\widetilde{f}(x) = f(x^0) + \vec{c}(\vec{x} - \vec{x}^0)$ stellt eine Hyperebene im $\mathbb{R}^{n+1}$ dar. Für $n = 1$ ist dies eine Gerade (die Tangente), für $n = 2$ eine Ebene (die Tangentialebene).

Wesentlich ist dabei, dass die Abweichung der linearen Approximation $\widetilde{f}$ von f bei Annäherung an x^0 von höherer als erster Ordnung verschwindet.

Bemerkung. Man kann zeigen, dass der Vektor $\vec{c}$ durch f eindeutig bestimmt ist.

Satz. Sei $f : \mathbb{R}^n \to \mathbb{R}$ an x^0 differenzierbar. Dann existieren alle partiellen Ableitungen, und es gilt

$\frac{\partial f}{\partial x_k}(x^0) = c_k$, i.e. $\vec{c} = \mathrm{grad} f(x^0)$.

Beweis. Laut Voraussetzung gilt

$f(x) = f(x^0) + \vec{c}(\vec{x} - \vec{x}^0) + \|x - x^0\| f_0(x)$.

Setzen wir $x = x^0 + h\vec{e}_k$, dann ist $\vec{c}(\vec{x} - \vec{x}^0) = hc_k$ und $\|x - x^0\| = |h|$.

Weiters ist $\left| \frac{f(x^0 + h\vec{e}_k) - f(x^0)}{h} - c_k \right| = \left| f_0(x^0 + h\vec{e}_k) \right|$. Mit $h \to 0$ erhalten wir daraus $\frac{\partial f}{\partial x_k}(x^0) = c_k$. $\square$

Bemerkung. Aus der Existenz aller partiellen Ableitungen folgt i.a. **nicht** die Differenzierbarkeit (d.h. obiger Satz ist i.a. nicht umkehrbar).

Durch den Gradienten von f wird eine Richtung im $\mathbb{R}^n$ festgelegt. In welcher Beziehung diese Richtung zur Änderung von f steht, zeigt folgender

Satz. Sei $f : \mathbb{R}^n \to \mathbb{R}$ an x^0 differenzierbar. Dann ist f an x^0 in jeder Richtung $\vec{a}$ (mit $\|\vec{a}\| = 1$) differenzierbar und es gilt

(i) $\frac{\partial f}{\partial \vec{a}}(x^0) = (\mathrm{grad} f(x^0)) \cdot \vec{a}$,

(ii) $\left|\frac{\partial f}{\partial \vec{a}}(x^0)\right| \leq \|\mathrm{grad} f(x^0)\|$.

Beweis. Laut Voraussetzung gilt auf einer Umgebung $U(x^0)$

$f(x) = f(x^0) + \mathrm{grad} f(x^0)(\vec{x} - \vec{x}^0) + \|x - x^0\| f_0(x)$.

Mit $x = x^0 + h\vec{a}$ und folglich $\|x - x^0\| = |h|$ gilt dann

$\left|\frac{f(x^0+h\vec{a})-f(x^0)}{h} - \mathrm{grad} f(x^0) \cdot \vec{a}\right| = \left|f_0(x^0 + h\vec{a})\right|$.

Mit $h \to 0$ folgt die erste Behauptung. Die zweite Behauptung ergibt sich durch Anwendung der Cauchy-Schwarzschen Ungleichung. □

Bemerkung. Aus der zweiten Aussage des obigen Satzes wird ersichtlich, dass der Gradient von f jene Richtung angibt, in der die Änderung von f maximal ist. Dies ist etwa insofern von Bedeutung, wenn f ein Temperaturfeld beschreibt. Dann gibt der Gradient die Richtung des Wärmestromes an, da Wärme stets in Richtung des größten Temperaturgefälles fließt.

Beispiel. Bestimme die Richtungsableitung von $f(x, y) = x^2 + xy$ im Punkt $(2, 1)$ in Richtung $45°$.

$\vec{a} = (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$, $\mathrm{grad} f = (2x + y, x)$, $\mathrm{grad} f|_{(2,1)} = (5, 2)$

$\frac{\partial f}{\partial \vec{a}}\Big|_{(2,1)} = \mathrm{grad} f|_{(2,1)} \cdot \vec{a} = (5, 2) \cdot (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}) = \frac{7}{\sqrt{2}}$

Satz. (ohne Beweis)
Sei $f:\mathbb{R}^n \to \mathbb{R}$ an x^0 differenzierbar. Dann gilt

(i) $\exists\ M > 0$ und eine Umgebung $U(x^0)$, auf der gilt

$|f(x) - f(x^0)| \leq M\|x - x^0\|$,

(ii) f ist stetig an x^0 .

Wie wir bereits wissen, ist die bloße Existenz der partiellen Ableitungen nicht hinreichend fr die Stetigkeit. Erst die zusätzliche Voraussetzung der Beschränktheit der partiellen Ableitungen (auf einer Umgebung von x^0) garantiert die Stetigkeit.

Welche zusätzliche Forderung an die partiellen Ableitungen garantiert die Differenzierbarkeit ?

Satz. (ohne Beweis)
Existieren für $f:\mathbb{R}^n \to \mathbb{R}$ auf einer Umgebung $U(x^0)$ die partiellen Ableitungen $f_{x_1}, f_{x_2}, ..., f_{x_n}$ und sind diese **stetig** an x^0 , dann ist f differenzierbar an x^0 .

Beispiel. Sei $f:\mathbb{R}^3 \to \mathbb{R}$ mit $f(x,y,z) = \sin(xyz)$. Dann sind $f_x = yz\cos(xyz)$, $f_y = xz\cos(xyz)$, $f_z = xy\cos(xyz)$ stetig in jedem (x,y,z) . Somit ist f differenzierbar in jedem x^0 .

Nun zur Differenzierbarkeit vektorwertiger Funktionen.

Definition. Sei $f:\mathbb{R}^n \to \mathbb{R}^m$ und $x^0 \in D(f)$ ein innerer Punkt.

Dann heißt $f = (f_1, ..., f_m)$ **differenzierbar** an x^0 , wenn es eine $m \times n$ Matrix $C = (c_{\mu\nu})$ und eine auf einer Umgebung $U(x^0)$ definierte Funktion $f_0(x)$ (mit Werten in $\mathbb{R}^m$) gibt, sodass

- $f(x) = f(x^0) + C(\vec{x} - \vec{x}^0) + \|x - x^0\| f_0(x)$
- $\lim\limits_{x \to x^0} f_0(x) = \vec{0}$.

Bemerkung.

(i) Differenzierbarkeit heißt also, dass

$$\frac{1}{\|x-x^0\|}\left(f(x)-f(x^0)-C(\vec{x}-\vec{x}^0)\right)\to\vec{0} \quad \text{für} \quad x\to x^0 \ .$$

(ii) $f:\mathbb{R}^n\to\mathbb{R}^m$ ist genau dann differenzierbar an x^0 , wenn jede Koordinatenfunktion $f_i:\mathbb{R}^n\to\mathbb{R}$ differenzierbar an x^0 ist.

(iii) Falls $f:\mathbb{R}^n\to\mathbb{R}^m$ differenzierbar an x^0 ist, dann ist

$$c_{\mu\nu}=\frac{\partial f_\mu}{\partial x_\nu}(x^0) \ .$$

Definition. Existieren für $f:\mathbb{R}^n\to\mathbb{R}^m$ an einer Stelle x^0 die partiellen Ableitungen $\frac{\partial f_\mu}{\partial x_\nu}(x^0)$, dann heißt die Matrix

$$J_f(x^0)=\left(\frac{\partial f_\mu}{\partial x_\nu}(x^0)\right)=\frac{\partial(f_1,f_2,\ldots,f_m)}{\partial(x_1,x_2,\ldots,x_n)}=\frac{df}{dx}(x^0)=f'(x^0)$$

die **Jacobi-Matrix**, **Funktionalmatrix** oder **Ableitung** von f .

Die Zeilenvektoren der Jacobi-Matrix sind also die Gradienten der jeweiligen Koordinatenfunktionen. Weiters gilt

Satz. (Lipschitz-Bedingung für differenzierbare Funktionen)

Sei $f:\mathbb{R}^n\to\mathbb{R}^m$. Existiert für $x^0\in D(f)$ die Jacobi-Matrix, dann gibt es ein $\delta>0$ und eine Schranke M , sodass für alle x mit $\|x-x^0\|<\delta$ gilt, dass $\|f(x)-f(x^0)\|\le M\|x-x^0\|$.

Beispiele.

1) Betrachte $f:\mathbb{R}^2\to\mathbb{R}^2$ mit $f_1(x,y)=x(1-y)$, $f_2(x,y)=xy$ besitzt die Jacobi-Matrix

$$J_f=\begin{pmatrix}1-y & -x\\ y & x\end{pmatrix} .$$

2) Die Polarkoordinatenabbildung $P:\mathbb{R}^2\to\mathbb{R}^2$ mit $P_1(r,\varphi)=r\cos\varphi$, $P_2(r,\varphi)=r\sin\varphi$ besitzt die Jacobi-Matrix

$$J_P = \begin{pmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{pmatrix} .$$

Satz. (Ableitungsregeln)

1) Seien $f, g : \mathbb{R}^n \to \mathbb{R}^m$ an einem inneren Punkt x^0 von $D(f)$ bzw. $D(g)$ differenzierbar, und seien $\lambda, \mu \in \mathbb{R}$.

Dann ist auch die Funktion $\lambda f + \mu g$ an x^0 differenzierbar und es gilt

$$J_{\lambda f+\mu g}(x^0) = \lambda J_f(x^0) + \mu J_g(x^0) .$$

2) Seien $f, g : \mathbb{R}^n \to \mathbb{R}^m$ an einem inneren Punkt x^0 von $D(f)$ bzw. $D(g)$ differenzierbar.

Dann ist auch die (reellwertige) Funktion $f * g = \sum\limits_{i=1}^{m} f_i g_i$ an x^0 differenzierbar und es gilt

$$(f * g)' = \mathrm{grad}(f * g) = {}^t(f') \cdot g + f \cdot {}^t(g') = {}^t(J_f) \cdot g + f \cdot {}^t(J_g) =$$
$$\sum_{i=1}^{m} ((\mathrm{grad} f_i) g_i + f(\mathrm{grad} g_i)) .$$

Satz. (**Kettenregel**)

Sei $f : \mathbb{R}^n \to \mathbb{R}^m$ an einem inneren Punkt $x^0 \in D(f)$ differenzierbar und $g : \mathbb{R}^m \to \mathbb{R}^l$ an $y^0 = f(x^0) \in D(g)$ differenzierbar (wobei y^0 ein innerer Punkt von $D(g)$ sei).

Dann ist $h = g \circ f : \mathbb{R}^n \to \mathbb{R}^l$ an x^0 differenzierbar und es gilt

$$J_h(x^0) = J_g(y^0) J_f(x^0) \quad \text{bzw.} \quad \tfrac{dh}{dx}(x^0) = \tfrac{dg}{dy}(y^0) \tfrac{df}{dx}(x^0) .$$

Beispiel. Sei $f : \mathbb{R}^n \to \mathbb{R}$ und $\psi : \mathbb{R} \to \mathbb{R}^n$.

Dann ist $\varphi = f \circ \psi : \mathbb{R} \to \mathbb{R}$, $J_f(x) = \left(\frac{\partial f}{\partial x_1}(x), \frac{\partial f}{\partial x_2}(x), ..., \frac{\partial f}{\partial x_n}(x)\right)$, und

$J_\psi(t) = \begin{pmatrix} \psi_1'(t) \\ \psi_2'(t) \\ .... \\ \psi_n'(t) \end{pmatrix}$. Nach der Kettenregel ist dann

$$J_\varphi(t) = \varphi'(t) = J_f(\psi(t)) J_\psi(t) = \sum_{i=1}^{n} \tfrac{\partial f}{\partial x_i}\left(\psi(t)\right) \psi_i'(t) =$$

$$= \tfrac{\partial f}{\partial x_1}\left(\psi(t)\right) \psi_1'(t) + \tfrac{\partial f}{\partial x_2}\left(\psi(t)\right) \psi_2'(t) + ... + \tfrac{\partial f}{\partial x_n}\left(\psi(t)\right) \psi_n'(t)$$

Beispiel. Sei $f : \mathbb{R}^2 \to \mathbb{R}^2$ mit $f(x,y) = (x^2+y^2, xy)$, und weiters $g : \mathbb{R}^2 \to \mathbb{R}$ mit $g(u,v) = e^{uv}$.

Bestimme die Jacobi-Matrix von $g \circ f$ im Punkt $(1,1)$.

$J_f = \begin{pmatrix} 2x & 2y \\ y & x \end{pmatrix}$, $J_f|_{(1,1)} = \begin{pmatrix} 2 & 2 \\ 1 & 1 \end{pmatrix}$

$f(1,1) = (2,1)$, $J_g = \begin{pmatrix} ve^{uv} & ue^{uv} \end{pmatrix}$, $J_g|_{(2,1)} = \begin{pmatrix} e^2 & 2e^2 \end{pmatrix}$

$$J_{g \circ f}|_{(1,1)} = \begin{pmatrix} e^2 & 2e^2 \end{pmatrix} \begin{pmatrix} 2 & 2 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 4e^2 & 4e^2 \end{pmatrix}$$

Beachte auch, dass $(g \circ f)(x,y) = e^{xy(x^2+y^2)}$.

Sei $f : \mathbb{R}^n \to \mathbb{R}$. Die partiellen Ableitungen $\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, ..., \frac{\partial f}{\partial x_n}$ können ihrerseits wieder (auf gewissen Teilmengen ihres Definitionsbereiches) differenzierbar sein.

Wir nennen dann die Funktionen $\frac{\partial^2 f}{\partial x_i \partial x_k} = \frac{\partial}{\partial x_k}\left(\frac{\partial f}{\partial x_i}\right)$ **partielle Ableitungen zweiter Ordnung** . Dabei wird auch die Schreibweise $f_{x_i x_k}$ verwendet.

Rekursiv können dann (falls existent) die Ableitungen höherer Ordnung definiert werden.

Wenn für $l = 0, 1, 2, ..., k$ auf einer Menge $X \subseteq \mathbb{R}^n$ alle partiellen Ableitungen l-ter Ordnung von f existieren und dort stetig sind, heißt

f **k-mal stetig differenzierbar auf** X und man schreibt $f \in C^k(X)$.

Beispiele.

(i) $f(x,y) = x^y$ ist auf $\{(x,y) \in \mathbb{R}^2 \ : \ x > 0\}$ definiert und besitzt dort die partiellen Ableitungen

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

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

Man beachte, dass hier $\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}$ gilt.

(ii) Sei $f(x,y) = \begin{cases} xy\frac{x^2-y^2}{x^2+y^2} & \text{wenn } (x,y) \neq (0,0) \\ 0 & \text{wenn } (x,y) = (0,0) \end{cases}$.

Dann gilt $f_x(0,0) = \lim\limits_{h \to 0} \frac{f(h,0)-f(0,0)}{h} = 0$ und $f_y(0,0) = 0$.

Für $(x,y) \neq (0,0)$ gilt $f_x(x,y) = \frac{y(x^4+4x^2y^2-y^4)}{(x^2+y^2)^2}$, $f_y(x,y) = \frac{x(x^4-4x^2y^2-y^4)}{(x^2+y^2)^2}$.

Daraus erhalten wir $f_{xy}(0,0) = \lim\limits_{h \to 0} \frac{f_x(0,h)-f_x(0,0)}{h} = -1$ und $f_{yx}(0,0) = \lim\limits_{h \to 0} \frac{f_y(h,0)-f_y(0,0)}{h} = +1$.

Also existieren zwar $f_{xy}(0,0)$, $f_{yx}(0,0)$, aber $f_{xy}(0,0) \neq f_{yx}(0,0)$.

Für $(x,y) \neq (0,0)$ gilt hingegen

$$f_{xy} = f_{yx} = \frac{(x^2-y^2)(x^4+10x^2y^2+y^4)}{(x^2+y^2)^3} \ .$$

Bei Annäherung an $(0,0)$ entlang der Geraden $y = x$ erhalten wir als Richtungsgrenzwert für f_{xy} und f_{yx} den Wert 0, im Gegensatz zu oben ermittelten Werten $f_{xy}(0,0) = -1$ bzw. $f_{yx}(0,0) = +1$, d.h. f_{xy} und f_{yx} sind im Ursprung **nicht** stetig.

Satz. (ohne Beweis)

Für $f : \mathbb{R}^n \to \mathbb{R}$ seien $\frac{\partial^2 f}{\partial x_i \partial x_k}$ und $\frac{\partial^2 f}{\partial x_k \partial x_i}$ stetig an x^0 .

Dann gilt $\frac{\partial^2 f}{\partial x_i \partial x_k}(x^0) = \frac{\partial^2 f}{\partial x_k \partial x_i}(x^0)$.

Bemerkung. Für C^k-Funktionen sind die Ableitungen bis zur k-ten Ordnung vertauschbar, weil jede Permutation der dabei beteiligten Variablen durch endlich viele spezielle Permutationen, bei denen nur zwei aufeinanderfolgende Elemente vertauscht werden, darstellbar ist.