

2.7 Die Ableitung einer Funktion im Mehrdimensionalen

Bemerkung: Im Eindimensionalen wird die Ableitung einer Funktion üblicherweise mit Hilfe des Differenzenquotienten $\frac{f(x)-f(a)}{x-a}$ definiert. Das kann für eine Funktion $f: D \rightarrow \mathbb{R}^l$ und $D \subseteq \mathbb{R}^k$ und $k, l \geq 1$ beliebig nicht funktionieren, da im „Quotienten“ ein Vektor aus \mathbb{R}^l durch einen Vektor aus \mathbb{R}^k dividiert würde. Man verwendet darum für die Verallgemeinerung die folgende Charakterisierung:

Satz 108 Es sei $I \subseteq \mathbb{R}$ ein offenes Intervall, $f: I \rightarrow \mathbb{R}$ und $a \in I$. Dann sind äquivalent:

(i) f ist bei a differenzierbar,

(ii) Es gibt ein $\alpha \in \mathbb{R}$, sodass $\lim_{x \rightarrow a} \frac{f(x)-f(a)-\alpha(x-a)}{x-a} = 0$.

Beweis: (i) \Rightarrow (ii) Es sei $\alpha := f'(a)$. Dann ist

$$\frac{f(x)-f(a)-\alpha(x-a)}{x-a} = \underbrace{\frac{f(x)-f(a)}{x-a}}_{\rightarrow f'(a)} - \alpha \xrightarrow{x \rightarrow a} 0.$$

(ii) Wegen

$$0 = \lim_{x \rightarrow a} \frac{f(x)-f(a)-\alpha(x-a)}{x-a} = \lim_{x \rightarrow a} \left(\frac{f(x)-f(a)}{x-a} - \alpha \right) = \lim_{x \rightarrow a} \frac{f(x)-f(a)}{x-a} - \alpha$$

$$\text{existiert } f'(a) = \lim_{x \rightarrow a} \frac{f(x)-f(a)}{x-a} = \alpha.$$

Definition Es sei $U \subseteq \mathbb{R}^k$ offen, $f: U \rightarrow \mathbb{R}^l$ und $a \in U$. Dann ist f bei a differenzierbar, wenn es eine lineare Abbildung $L: \mathbb{R}^k \rightarrow \mathbb{R}^l$ gibt, derart dass

$$\lim_{x \rightarrow a} \frac{1}{\|x-a\|} (f(x)-f(a)-L(x-a)) = 0.$$

Bemerkungen: 1) Satz 108 bzw. die Definition besagen, dass f bei a differenzierbar ist, wenn es sich in der Nähe von a sehr gut durch die „Tangentialabbildung“ $x \mapsto f(a) + L(x-a)$ approximieren lässt. (Dabei heißt „sehr gut“, dass der Fehler $f(x)-f(a)-L(x-a)$ rascher gegen 0 geht als der Abstand $x-a$, wenn man $x \rightarrow a$ gehen lässt.)

2) Im Eindimensionalen (oder wenn $k=l=1$) ist die lineare Abbildung L gerade $L(x) = f'(a) \cdot x$.

3) Aus der Definition ist zunächst nicht ersichtlich, dass die lineare

Abbildung L (wenn sie überhaupt existiert) eindeutig bestimmt ist. Der es wäre denkbar, dass es mehrere solche lineare Abbildungen gibt. Der nachfolgende Satz zeigt aber, dass das unmöglich ist.

Satz 109 Es sei $U \subseteq \mathbb{R}^k$ offen, $f: U \rightarrow \mathbb{R}^l$ und $\underline{a} \in U$. Wenn f bei \underline{a} differenzierbar ist, dann existieren alle partiellen Ableitungen $\frac{\partial f_i}{\partial x_j}(\underline{a})$ (mit $1 \leq i \leq l, 1 \leq j \leq k$) und jede nach der Definition der Differenzierbarkeit existierende lineare Abbildung L besitzt die (nach Satz 99(iii) existierende) Darstellung $L: \mathbb{R}^k \rightarrow \mathbb{R}^l, L(\underline{x}) = A \cdot \underline{x}$ mit

$$A = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\underline{a}) & \frac{\partial f_1}{\partial x_2}(\underline{a}) & \dots & \frac{\partial f_1}{\partial x_k}(\underline{a}) \\ \frac{\partial f_2}{\partial x_1}(\underline{a}) & \frac{\partial f_2}{\partial x_2}(\underline{a}) & \dots & \frac{\partial f_2}{\partial x_k}(\underline{a}) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_l}{\partial x_1}(\underline{a}) & \frac{\partial f_l}{\partial x_2}(\underline{a}) & \dots & \frac{\partial f_l}{\partial x_k}(\underline{a}) \end{pmatrix} = \left(\frac{\partial f_i}{\partial x_j}(\underline{a}) \right)_{\substack{1 \leq i \leq l \\ 1 \leq j \leq k}} \in M_{l,k}(\mathbb{R}).$$

Insbesondere ist die lineare Abbildung L eindeutig bestimmt.

Beweis: Eine nach der Definition der Differenzierbarkeit existierende lineare Abbildung $L: \mathbb{R}^k \rightarrow \mathbb{R}^l$ besitzt nach Satz 99(iii) eine Matrixdarstellung, dh $L(\underline{x}) = A \cdot \underline{x}$ für ein $A = (\alpha_{ij})_{\substack{1 \leq i \leq l \\ 1 \leq j \leq k}} \in M_{l,k}(\mathbb{R})$. Wir wollen zeigen, dass $\alpha_{ij} = \frac{\partial f_i}{\partial x_j}(\underline{a})$. Die Definition der Differenzierbarkeit besagt gerade

$$\frac{1}{\|\underline{x} - \underline{a}\|} (f(\underline{x}) - f(\underline{a}) - L(\underline{x} - \underline{a})) \xrightarrow{\underline{x} \rightarrow \underline{a}} 0.$$

Gibt man (mit Hilfe von Satz 95) zu den einzelnen Komponenten über, so erhält man daraus

$$\frac{1}{\|\underline{x} - \underline{a}\|} (f_i(\underline{x}) - f_i(\underline{a}) - \sum_{s=1}^k \alpha_{is} (x_s - a_s)) \xrightarrow{\underline{x} \rightarrow \underline{a}} 0 \quad \text{für } 1 \leq i \leq l.$$

Das gilt insbesondere wenn $x_s = a_s$ für $1 \leq s \leq k, s \neq j$. Dann sind

$$\sum_{s=1}^k \alpha_{is} (x_s - a_s) = \alpha_{ij} (x_j - a_j) \quad \text{und} \quad \|\underline{x} - \underline{a}\| = \sqrt{\sum_{s=1}^k (x_s - a_s)^2} = |x_j - a_j|.$$

Für derartige \underline{x} gilt also

$$\frac{f_i(a_1, \dots, a_{j-1}, x_j, a_{j+1}, \dots, a_k) - f_i(a_1, \dots, a_k) - \alpha_{ij} (x_j - a_j)}{|x_j - a_j|} \xrightarrow{\underline{x} \rightarrow \underline{a}} 0$$

für $1 \leq i \leq l, 1 \leq j \leq k$ oder, was dasselbe besagt,

$$\left| \frac{f_i(a_1, \dots, a_{j-1}, x_j, a_{j+1}, \dots, a_k) - f_i(a_1, \dots, a_k)}{x_j - a_j} - \alpha_{ij} \right| \xrightarrow{x_j \rightarrow a_j} 0$$

für $1 \leq i \leq l$, $1 \leq j \leq k$. Also ist

$$\frac{\partial f_i}{\partial x_j}(\underline{a}) = \lim_{x_j \rightarrow a_j} \frac{f_i(a_1, \dots, a_{j-1}, x_j, a_{j+1}, \dots, a_k) - f_i(a_1, \dots, a_k)}{x_j - a_j} = \alpha_{ij}$$

für $1 \leq i \leq l$, $1 \leq j \leq k$.

Definition: Es sei $U \subseteq \mathbb{R}^k$ offen, $f: U \rightarrow \mathbb{R}^l$, $\underline{a} \in U$ und f bei \underline{a} differenzierbar.

Als Ableitung $f'(\underline{a})$ von f bei \underline{a} bezeichnen wir die sogenannte jacobische

Matrix

$$f'(\underline{a}) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\underline{a}) & \dots & \frac{\partial f_1}{\partial x_k}(\underline{a}) \\ \vdots & & \vdots \\ \frac{\partial f_l}{\partial x_1}(\underline{a}) & \dots & \frac{\partial f_l}{\partial x_k}(\underline{a}) \end{pmatrix},$$

die wir mit der linearen Abbildung $\mathbb{R}^k \rightarrow \mathbb{R}^l$, $x \mapsto f'(\underline{a}) \cdot x$ identifizieren.

Bemerkung: Statt $f'(\underline{a})$ wird als Notation auch $Df(\underline{a})$ verwendet.

Satz 110 Es sei $U \subseteq \mathbb{R}^k$ offen, $f: U \rightarrow \mathbb{R}^l$ und $\underline{a} \in U$. Wenn f bei \underline{a} differenzierbar ist, ist f bei \underline{a} stetig.

Beweis: Für $x \in U \setminus \{\underline{a}\}$ ist

$$f(x) = f(\underline{a}) + \underbrace{f'(\underline{a}) \cdot (x - \underline{a})}_{\rightarrow \underline{a}} + \underbrace{\|x - \underline{a}\|}_{\rightarrow 0} \cdot \underbrace{\frac{1}{\|x - \underline{a}\|} (f(x) - f(\underline{a}) - f'(\underline{a}) \cdot (x - \underline{a}))}_{\rightarrow 0} \xrightarrow{x \rightarrow \underline{a}} f(\underline{a}),$$

also ist f bei \underline{a} stetig. (Bei $f'(\underline{a}) \cdot (x - \underline{a}) \rightarrow \underline{a}$ wurde verwendet, dass

lineare Abbildungen stetig sind, siehe Bsp. 6) auf Seite 103.)

Bemerkung: Aus der Existenz aller partiellen Ableitungen $\frac{\partial f_i}{\partial x_j}(\underline{a})$ an einer

Stelle \underline{a} folgt nicht, dass f bei \underline{a} differenzierbar ist. (Oh man kann

Satz 109 nicht „umkehren.“) Aus der Existenz aller partiellen Ableitungen

bei \underline{a} folgt ja nicht einmal die Stetigkeit von f bei \underline{a} , wie wir in Bsp. 2)

auf Seite 110 gesehen haben. Die Differenzierbarkeit von f bei \underline{a} folgt

aber, wenn die partiellen Ableitungen $\frac{\partial f_i}{\partial x_j}$ in einer Umgebung von \underline{a} nicht nur

existieren, sondern dort auch stetig sind, wie der folgende Satz 111 zeigt:

Satz 11.1 Es sei $U \subseteq \mathbb{R}^k$ offen, $f: U \rightarrow \mathbb{R}^l$ und $\underline{a} \in U$. Die partiellen Ableitungen $\frac{\partial f_i}{\partial x_j}(\underline{x})$ mögen für alle $\underline{x} \in U$ existieren und seien bei \underline{a} stetig. Dann ist f bei \underline{a} differenzierbar.

Beweis: Für $1 \leq i \leq l$ ist

$$\begin{aligned} f_i(x_1, \dots, x_k) - f_i(a_1, \dots, a_k) &= \sum_{j=1}^k \left(f_i(a_1, \dots, a_{j-1}, x_j, a_{j+1}, \dots, a_k) - f_i(a_1, \dots, a_{j-1}, a_j, a_{j+1}, \dots, a_k) \right) \\ &= f_i(x_1, a_2, \dots, a_k) - f_i(a_1, a_2, \dots, a_k) \\ &\quad + f_i(a_1, x_2, \dots, a_k) - f_i(a_1, a_2, x_3, \dots, a_k) \\ &\quad + f_i(a_1, a_2, x_3, \dots, a_k) - f_i(a_1, a_2, a_3, x_4, \dots, a_k) \\ &\quad + \dots \\ &\quad + f_i(a_1, \dots, a_{k-1}, x_k) - f_i(a_1, \dots, a_k) \end{aligned}$$

$$= \sum_{j=1}^k (x_j - a_j) \frac{\partial f_i}{\partial x_j}(a_1, \dots, a_{j-1}, \xi_j, a_{j+1}, \dots, a_k) \quad \text{für ein } \xi_j \text{ zwischen } a_j \text{ und } x_j$$

(Hier wurde der Mittelwertsatz der Differentialrechnung k -mal angewendet)

$$= \sum_{j=1}^k \frac{\partial f_i}{\partial x_j}(\underline{a}) \cdot (x_j - a_j) + \sum_{j=1}^k \left(\frac{\partial f_i}{\partial x_j}(a_1, \dots, a_{j-1}, \xi_j, a_{j+1}, \dots, a_k) - \frac{\partial f_i}{\partial x_j}(\underline{a}) \right) (x_j - a_j)$$

Es folgt, dass

$$\begin{aligned} \frac{1}{\|\underline{x} - \underline{a}\|} \left(f_i(x_1, \dots, x_k) - f_i(a_1, \dots, a_k) - \sum_{j=1}^k \frac{\partial f_i}{\partial x_j}(\underline{a}) \cdot (x_j - a_j) \right) \\ = \sum_{j=1}^k \left(\frac{\partial f_i}{\partial x_j}(a_1, \dots, a_{j-1}, \xi_j, a_{j+1}, \dots, a_k) - \frac{\partial f_i}{\partial x_j}(\underline{a}) \right) \cdot \frac{x_j - a_j}{\|\underline{x} - \underline{a}\|} \end{aligned}$$

und daher

$$\begin{aligned} \frac{1}{\|\underline{x} - \underline{a}\|} \left| f_i(x_1, \dots, x_k) - f_i(a_1, \dots, a_k) - \sum_{j=1}^k \frac{\partial f_i}{\partial x_j}(\underline{a}) \cdot (x_j - a_j) \right| \\ \leq \sum_{j=1}^k \underbrace{\left| \frac{\partial f_i}{\partial x_j}(a_1, \dots, a_{j-1}, \xi_j, a_{j+1}, \dots, a_k) - \frac{\partial f_i}{\partial x_j}(\underline{a}) \right|}_{\rightarrow 0} \cdot \underbrace{\frac{|x_j - a_j|}{\|\underline{x} - \underline{a}\|}}_{\leq 1} \xrightarrow{\underline{x} \rightarrow \underline{a}} 0 \end{aligned}$$

für $1 \leq i \leq l$. Dabei wurde verwendet: Wenn $\underline{x} \rightarrow \underline{a}$ gilt, dann gehen auch $x_{j+1} \rightarrow a_{j+1}, \dots, x_k \rightarrow a_k$ und $\xi_j \rightarrow a_j$, da ξ_j zwischen a_j und x_j liegt. Die Differenzierbarkeit von f bei \underline{a} folgt nun aus Satz 9.5.

Bemerkungen: 1) Als Kurzschreibweise für die Jacobische Matrix wird auch

$$\frac{\partial (f_1, \dots, f_\ell)}{\partial (x_1, \dots, x_k)} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \dots & \frac{\partial f_1}{\partial x_k} \\ \vdots & & \vdots \\ \frac{\partial f_\ell}{\partial x_1} & \dots & \frac{\partial f_\ell}{\partial x_k} \end{pmatrix}$$

verwendet. (Dabei ist $f: U \rightarrow \mathbb{R}^\ell$ (mit $U \subseteq \mathbb{R}^k$) und $f(x) = (f_1(x), \dots, f_\ell(x))$.)

2) Aus Satz 111 folgt sofort: Ist $f \in C^1(U, \mathbb{R}^\ell)$, so ist f auf U differenzierbar.

3) Es ist möglich, dass eine Funktion f in einem Punkt a differenzierbar ist, die partiellen Ableitungen $\frac{\partial f_i}{\partial x_j}$ bei a aber nicht stetig sind. (Das ist ja schon im Eindimensionalen, d.h. im Fall $k=l=1$, so.)

Beispiel: Es sei $f: \mathbb{R}^3 \rightarrow \mathbb{R}^2$, $f(x, y, z) = (xy \sin z, x^2 - y^2 - \cos z)$. Die partiellen Ableitungen

$$\frac{\partial f_1}{\partial x}(x, y, z) = y \sin z, \quad \frac{\partial f_1}{\partial y}(x, y, z) = x \sin z, \quad \frac{\partial f_1}{\partial z}(x, y, z) = xy \cos z$$

$$\frac{\partial f_2}{\partial x}(x, y, z) = 2x, \quad \frac{\partial f_2}{\partial y}(x, y, z) = -2y, \quad \frac{\partial f_2}{\partial z}(x, y, z) = \sin z$$

existieren auf ganz \mathbb{R}^3 und sind dort stetig, d.h. $f \in C^1(\mathbb{R}^3, \mathbb{R}^2)$. Daher ist f in jedem Punkt $(x, y, z) \in \mathbb{R}^3$ differenzierbar und

$$f'(x, y, z) = \begin{pmatrix} y \sin z & x \sin z & xy \cos z \\ 2x & -2y & \sin z \end{pmatrix} \quad \forall (x, y, z) \in \mathbb{R}^3.$$

Satz 112 Es sei $U \subseteq \mathbb{R}^k$ offen, $f, g: U \rightarrow \mathbb{R}^\ell$, $a \in U$ und f und g seien bei a differenzierbar. Dann gelten

(i) $f+g: U \rightarrow \mathbb{R}^\ell$ ist bei a differenzierbar und $(f+g)'(a) = f'(a) + g'(a)$,

(ii) $\alpha f: U \rightarrow \mathbb{R}^\ell$ ist bei a differenzierbar (mit $\alpha \in \mathbb{R}$) und $(\alpha f)'(a) = \alpha f'(a)$.

Beweis: Nach Voraussetzung gelten

$$\lim_{x \rightarrow a} \frac{1}{\|x-a\|} (f(x) - f(a) - f'(a) \cdot (x-a)) = 0$$

und

$$\lim_{x \rightarrow a} \frac{1}{\|x-a\|} (g(x) - g(a) - g'(a) \cdot (x-a)) = 0$$

Dabei ist

$$\lim_{x \rightarrow a} \frac{1}{\|x-a\|} ((f+g)(x) - (f+g)(a) - (f'(a) + g'(a)) \cdot (x-a))$$

$$= \lim_{x \rightarrow a} \frac{1}{\|x-a\|} (f(x) - f(a) - f'(a) \cdot (x-a)) + \lim_{x \rightarrow a} \frac{1}{\|x-a\|} (g(x) - g(a) - g'(a) \cdot (x-a)) = 0 + 0 = 0,$$

da $f+g$ ist bei \underline{a} differenzierbar und $(f+g)'(\underline{a}) = f'(\underline{a}) + g'(\underline{a})$, sowie

$$\lim_{\underline{x} \rightarrow \underline{a}} \frac{1}{\|\underline{x} - \underline{a}\|} \left((\alpha f)(\underline{x}) - (\alpha f)(\underline{a}) - (\alpha f'(\underline{a})) \cdot (\underline{x} - \underline{a}) \right) \\ = \alpha \lim_{\underline{x} \rightarrow \underline{a}} \frac{1}{\|\underline{x} - \underline{a}\|} \left(f(\underline{x}) - f(\underline{a}) - f'(\underline{a}) \cdot (\underline{x} - \underline{a}) \right) = \alpha \cdot 0 = 0,$$

da αf ist bei \underline{a} differenzierbar und $(\alpha f)'(\underline{a}) = \alpha \cdot f'(\underline{a})$.

(Dabei wurden Satz 96 (i) bzw (iv) verwendet.)

Satz 113 (Kettenregel) Es seien $U \subseteq \mathbb{R}^k$ und $V \subseteq \mathbb{R}^l$ offen, $f: U \rightarrow V$, $g: V \rightarrow \mathbb{R}^m$, $\underline{a} \in U$ und f bei \underline{a} und g bei $f(\underline{a})$ differenzierbar. Dann ist auch $g \circ f: U \rightarrow \mathbb{R}^m$ bei \underline{a} differenzierbar und es gilt $(g \circ f)'(\underline{a}) = g'(f(\underline{a})) \cdot f'(\underline{a})$.

Ohne Beweis

Bemerkung: Sind f und g wie in Satz 113, $h = g \circ f$ und f, g und h haben Komponenten

$$f(\underline{x}) = (f_1(\underline{x}), \dots, f_l(\underline{x})), \quad g(\underline{y}) = (g_1(\underline{y}), \dots, g_m(\underline{y})) \quad \text{und} \quad h(\underline{x}) = (h_1(\underline{x}), \dots, h_m(\underline{x})),$$

so besagt Satz 113, dass

$$\underbrace{\begin{pmatrix} \frac{\partial h_1}{\partial x_1}(\underline{a}) & \dots & \frac{\partial h_1}{\partial x_k}(\underline{a}) \\ \vdots & & \vdots \\ \frac{\partial h_m}{\partial x_1}(\underline{a}) & \dots & \frac{\partial h_m}{\partial x_k}(\underline{a}) \end{pmatrix}}_{n \times k \text{- Matrix}} = \underbrace{\begin{pmatrix} \frac{\partial g_1}{\partial y_1}(f(\underline{a})) & \dots & \frac{\partial g_1}{\partial y_l}(f(\underline{a})) \\ \vdots & & \vdots \\ \frac{\partial g_m}{\partial y_1}(f(\underline{a})) & \dots & \frac{\partial g_m}{\partial y_l}(f(\underline{a})) \end{pmatrix}}_{m \times l \text{- Matrix}} \cdot \underbrace{\begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\underline{a}) & \dots & \frac{\partial f_1}{\partial x_k}(\underline{a}) \\ \vdots & & \vdots \\ \frac{\partial f_l}{\partial x_1}(\underline{a}) & \dots & \frac{\partial f_l}{\partial x_k}(\underline{a}) \end{pmatrix}}_{l \times k \text{- Matrix}}$$

Für die Eintragungen von $h'(\underline{a})$ gilt also

$$\frac{\partial h_i}{\partial x_j}(\underline{a}) = \frac{\partial g_i}{\partial y_1}(f(\underline{a})) \cdot \frac{\partial f_1}{\partial x_j}(\underline{a}) + \frac{\partial g_i}{\partial y_2}(f(\underline{a})) \cdot \frac{\partial f_2}{\partial x_j}(\underline{a}) + \dots + \frac{\partial g_i}{\partial y_l}(f(\underline{a})) \cdot \frac{\partial f_l}{\partial x_j}(\underline{a}) \\ = \sum_{s=1}^l \frac{\partial g_i}{\partial y_s}(f(\underline{a})) \cdot \frac{\partial f_s}{\partial x_j}(\underline{a})$$

Beispiel Sind $f: \mathbb{R} \rightarrow \mathbb{R}^2$, $f(x) = (x^3, 1+x^2)$ und $g: \mathbb{R}^2 \rightarrow \mathbb{R}$, $g(y_1, y_2) = e^{y_1} \sin y_2$, so ist

$h: \mathbb{R} \rightarrow \mathbb{R}$, $h(x) = g(f(x)) = e^{x^3} \sin(1+x^2)$ und daher $h'(x) = 3x^2 e^{x^3} \sin(1+x^2) + 2x e^{x^3} \cos(1+x^2)$

Das kann man auch mittels Satz 113 berechnen, nämlich

$$h'(x) \quad \text{,,} = \frac{\partial h}{\partial x}(x) = \frac{\partial g}{\partial y_1}(f(x)) \cdot \frac{\partial f_1}{\partial x}(x) + \frac{\partial g}{\partial y_2}(f(x)) \cdot \frac{\partial f_2}{\partial x}(x) \\ = e^{y_1} \sin y_2 \Big|_{(y_1, y_2) = (x^3, 1+x^2)} \cdot 3x^2 + e^{y_1} \cos y_2 \Big|_{(y_1, y_2) = (x^3, 1+x^2)} \cdot 2x \\ = 3x^2 e^{x^3} \sin(1+x^2) + 2x e^{x^3} \cos(1+x^2).$$