Analysis II

Apl. Prof. Dr. Axel Grünrock

Heinrich-Heine-Universität Düsseldorf

Sommersemester 2020

Fortsetzung: Totale Differenzierbarkeit

Wir befinden uns immer noch im Abschnitt 2.2 über die totale Differenzierbarkeit und haben die Jacobimatrix als Verallgemeinerung der Ableitung für vektorwertige Funktionen mehrerer Veränderlicher kennengelernt. Es stellt sich die Frage nach entsprechenden Verallgemeinerungen der Ableitungsregeln. Was die Produktregel anbelangt, so werden wir dieser Frage in den Übungen nachgehen für den Fall, dass ein Produkt aus einem Skalarfeld und einem Vektorfeld gebildet wird. Hier in der "Vorlesung" soll es nun um die Kettenregel im allgemeinen Rahmen gehen:

Die Kettenregel

Satz 3: Es seien $\Omega \subset \mathbb{R}^n$, $\Omega' \subset \mathbb{R}^k$ offen und

$$f : \Omega \to \mathbb{R}^m \qquad \text{sowie} \qquad g : \Omega' \to \Omega$$

Funktionen, so dass g in $x_0 \in \Omega'$ und f in $g(x_0) \in \Omega$ total differenzierbar sind mit Funktionalmatrizen $Dg(x_0)$ bzw. $Df(g(x_0))$. Dann ist auch $f \circ g : \Omega' \to \mathbb{R}^m$ in $x_0 \in \Omega'$ total differenzierbar und es gilt

$$D(f \circ g)(x_0) = Df(g(x_0)) \cdot Dg(x_0).$$

(Der $\cdot$ deutet hier das Produkt zweier Matrizen an.)

Beweis der Kettenregel (I)

Wir haben

$$f(y+h) = f(y) + Df(y)h + \varphi(h), \qquad \text{wobei}$$

$$\lim_{h\to 0} \frac{\varphi(h)}{|h|} = 0, \qquad \text{und}$$

$$g(x_0+k) = g(x_0) + Dg(x_0)k + \psi(k), \qquad \text{wobei}$$

$$\lim_{k\to 0} \frac{\psi(k)}{|k|} = 0.$$

Hieraus folgt mit $y = g(x_0)$ und $h = Dg(x_0)k + \psi(k)$, dass

Beweis der Kettenregel (II)

$$f(g(x_0+k)) - f(g(x_0)) = f(g(x_0) + \underbrace{Dg(x_0)k + \psi(k)}_{=:h}) - f(g(x_0))$$

$$= Df(g(x_0))Dg(x_0)k + r_1(k) + r_2(k) \qquad \text{wobei}$$

$$r_1(k) = Df(g(x_0))\psi(k) \qquad \text{und} \qquad r_2(k) = \varphi(Dg(x_0)k + \psi(k)).$$

Nun ist

$$\frac{|r_1(k)|}{|k|} = \frac{|Df(g(x_0))\psi(k)|}{|k|} \le \|Df(g(x_0))\| \frac{|\psi(k)|}{|k|} \to 0 \quad (k \to 0)$$

und - für hinreichend kleines k in der dritten Zeile -

Beweis der Kettenregel (III)

$$\frac{|r_2(k)|}{|k|} = \frac{|\varphi(Dg(x_0)k + \psi(k))|}{|k|}$$

$$= \frac{|\varphi(Dg(x_0)k + \psi(k))|}{|Dg(x_0)k + \psi(k))|} \frac{|Dg(x_0)k + \psi(k)|}{|k|}$$

$$\leq (\|Dg(x_0)\| + 1) \frac{|\varphi(Dg(x_0)k + \psi(k))|}{|Dg(x_0)k + \psi(k))|} \to 0 \quad (k \to 0).$$

Insgesamt also $\lim\limits_{k \to 0} \dfrac{r_1(k) + r_2(k)}{|k|} = 0$ und damit die Behauptung.

$\square$

Beispiel zur Anwendung der Kettenregel

Wir betrachten die Funktionen

$$f : \mathbb{R}^3 \to \mathbb{R}^2,\ y = (y_1, y_2, y_3)^\top \mapsto f(y) = \begin{pmatrix} y_1^2 + y_2^2 + y_3^2 \\ y_1 y_2 y_3 \end{pmatrix}$$

aus dem vorigen Beispiel und

$$g : \mathbb{R}^2 \to \mathbb{R}^3,\ x = (x_1, x_2)^\top \mapsto g(x) := \begin{pmatrix} x_1^2 \\ x_2^2 \\ x_1 x_2 \end{pmatrix}.$$

Bereits berechnet haben wir

$$Df(y) = \begin{pmatrix} 2y_1 & 2y_2 & 2y_3 \\ y_2 y_3 & y_1 y_3 & y_1 y_2 \end{pmatrix},$$

Fortsetzung des Beispiels

für g ergibt sich

$$Dg(x) = \begin{pmatrix} 2x_1 & 0 \\ 0 & 2x_2 \\ x_2 & x_1 \end{pmatrix}$$

und damit $D(f \circ g)(x) = Df(g(x))Dg(x)$

$$= \begin{pmatrix} 2x_1^2 & 2x_2^2 & 2x_3^2 \\ x_1 x_2^3 & x_1^3 x_2 & x_1^2 x_2^2 \end{pmatrix} \begin{pmatrix} 2x_1 & 0 \\ 0 & 2x_2 \\ x_2 & x_1 \end{pmatrix}$$

$$= \begin{pmatrix} 4x_1^3 + 2x_1 x_2^2 & 4x_2^3 + 2x_1^2 x_2 \\ 3x_1^2 x_2^3 & 3x_1^3 x_2^2 \end{pmatrix}.$$

Aufgabe: Zur Überprüfung berechne man einen expliziten Ausdruck für die zusammengesetzte Abbildung $f \circ g : \mathbb{R}^? \to \mathbb{R}^?$ und bestimme daraus direkt die Jacobi-Matrix $D(f \circ g)(x)$. Dieser Weg sollte ja zum selben Ergebnis führen.

Ein Spezialfall der Kettenregel

Betrachten wir für $\Omega \subset \mathbb{R}^n$ den Fall $m = 1$, also die folgende Situation

$$\mathbb{R}^k \supset \Omega' \quad \xrightarrow{g} \quad \Omega \quad \xrightarrow{f} \quad \mathbb{R},$$

dabei sei die totale Differenzierbarkeit von f und g vorausgesetzt. Dann ist $f \circ g$ eine reellwertige Funktion und also

$$D(f \circ g)(x) = \nabla(f \circ g)(x) = \left(\frac{\partial(f \circ g)}{\partial x_1}(x), \ldots, \frac{\partial(f \circ g)}{\partial x_k}(x) \right).$$

Andererseits ergibt die Kettenregel für $f : \Omega \to \mathbb{R}$

$$\begin{aligned} D(f \circ g)(x) &= Df(g(x))Dg(x) = \nabla f(g(x))Dg(x) \\ &= \left(\frac{\partial f}{\partial y_i}(g(x)) \right)_{1 \leq i \leq n} \left(\frac{\partial g_i}{\partial x_j}(x) \right)_{\substack{1 \leq i \leq n \\ 1 \leq j \leq k}}. \end{aligned}$$

Fortsetzung des Spezialfalls

In Komponenten: Für $1 \leq j \leq k$ ist dann

$$\frac{\partial(f \circ g)}{\partial x_j}(x) = \sum_{i=1}^{n} \frac{\partial f}{\partial y_i}(g(x)) \frac{\partial g_i}{\partial x_j}(x).$$

Spezialisieren wir jetzt noch einen Schritt weiter, indem wir auch $k = 1$ annehmen, so wird $\dfrac{\partial}{\partial x} = \dfrac{d}{dx}$ die (totale) eindimensionale Ableitung und

$$\frac{d(f \circ g)}{dx}(x) = \sum_{i=1}^{n} \frac{\partial f}{\partial y_i}(g(x)) \frac{\partial g_i}{dx}(x) = \underbrace{\nabla f(g(x))}_{\text{Zeile}} \underbrace{g'(x)}_{\text{Spalte}}.$$

Eine Anwendung in der Physik (I)

Es sei f eine Komponente eines (z. B. Kraft-)feldes, das auf einen Massepunkt am Ort $(x, y, z) \in \mathbb{R}^3$ wirkt. f sei zeit- und ortsabhängig, also

$$f : (t, x, y, z) \mapsto f(t, x, y, z) \in \mathbb{R}.$$

Soweit für einen festen Massepunkt am Ort (x, y, z). Interessanter wird die Situation mit der weitergehenden Annahme, dass sich der Massepunkt (Elementarteilchen o. ä.) auf einer Bahn

$$\gamma : \mathbb{R} \to \mathbb{R}^3, \quad t \mapsto \gamma(t) = (x(t), y(t), z(t))$$

bewegt. Dann wird nämlich die Zeitabhängigkeit der Komponente des Feldes, die auf den Massepunkt wirkt, beschrieben durch die Funktion

Eine Anwendung in der Physik (II)

$$\varphi : \mathbb{R} \to \mathbb{R}, \quad t \mapsto \varphi(t) := f(t, x(t), y(t), z(t)).$$

Die Kettenregel in der zuletzt genannten Form ergibt hier

$$\frac{d\varphi(t)}{dt} = \frac{\partial f}{\partial t}(t, x(t), y(t), z(t)) \cdot 1 +$$

$$\frac{\partial f}{\partial x}(t, x(t), y(t), z(t))x'(t) + \frac{\partial f}{\partial y}(t, x(t), y(t), z(t))y'(t) +$$

$$\frac{\partial f}{\partial z}(t, x(t), y(t), z(t))z'(t).$$

In dieser Situation wird in der physikalischen Literatur oft nicht unterschieden zwischen den Funktionen f und φ, was mitunter erhebliche Verwirrung stiften kann.

Der Begriff der Richtungsableitung ...

... liefert uns eine weitere Anwendung der Kettenregel.

Definition: Es sei $\Omega \subset \mathbb{R}^n$ offen, $x_0 \in \Omega$ und $f : \Omega \to \mathbb{R}^m$ eine Funktion. $\xi \in \mathbb{R}^n$ sei ein fester Vektor mit $|\xi| = 1$. Falls der Grenzwert existiert, nennen wir

$$\frac{\partial f}{\partial \xi}(x_0) := \lim_{t \to 0} \frac{1}{t}(f(x_0 + t\xi) - f(x_0))$$

die *Richtungsableitung* von f nach ξ in x_0.

Bemerkungen

(1) Wir nennen f in diesem Fall auch in x_0 nach ξ oder in Richtung von ξ differenzierbar.

Bemerkungen

(1) Wir nennen f in diesem Fall auch in x_0 nach ξ oder in Richtung von ξ differenzierbar.

(2) Für $\xi = e_j$ ergibt die Definition $\dfrac{\partial f}{\partial \xi}(x_0) = \dfrac{\partial f}{\partial x_j}(x_0)$. Die partiellen Ableitungen erweisen sich also als die Richtungsableitungen nach den kanonischen Basisvektoren.

Bemerkungen

(1) Wir nennen f in diesem Fall auch in x_0 nach ξ oder in Richtung von ξ differenzierbar.

(2) Für $\xi = e_j$ ergibt die Definition $\frac{\partial f}{\partial \xi}(x_0) = \frac{\partial f}{\partial x_j}(x_0)$. Die partiellen Ableitungen erweisen sich also als die Richtungsableitungen nach den kanonischen Basisvektoren.

(3) Im Fall total differenzierbarer, reellwertiger Funktionen kann man Richtungsableitungen mit Hilfe des Gradienten recht einfach berechnen, wie der nächste Satz zeigen wird.

Bemerkungen

(1) Wir nennen f in diesem Fall auch in x_0 nach ξ oder in Richtung von ξ differenzierbar.

(2) Für $\xi = e_j$ ergibt die Definition $\dfrac{\partial f}{\partial \xi}(x_0) = \dfrac{\partial f}{\partial x_j}(x_0)$. Die partiellen Ableitungen erweisen sich also als die Richtungsableitungen nach den kanonischen Basisvektoren.

(3) Im Fall total differenzierbarer, reellwertiger Funktionen kann man Richtungsableitungen mit Hilfe des Gradienten recht einfach berechnen, wie der nächste Satz zeigen wird.

(4) Die Richtungsableitung wird nur definiert für Vektoren ξ mit $|\xi| = 1$, da wir Informationen über die Änderung von f in Richtung von ξ haben wollen, die unabhängig von der Länge von ξ sind.

Satz 4

Es sei $\Omega \subset \mathbb{R}^n$ offen, $f : \Omega \to \mathbb{R}$ total differenzierbar in $x_0 \in \Omega$ und $\xi \in \mathbb{R}^n$ mit $|\xi| = 1$. Dann existiert die Richtungsableitung von f nach ξ in x_0 und es gilt

$$\frac{\partial f}{\partial \xi}(x_0) = \nabla f(x_0)\xi = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(x_0)\xi_j \,.$$

Zum Beweis wendet man die Kettenregel an auf die Verknüpfung $f \circ \gamma$, wobei die innere Funktion

$$\gamma : (-\varepsilon, \varepsilon) \to \Omega, \quad t \mapsto \gamma(t) = x_0 + t\xi$$

affin-linear (und also insbesondere total differenzierbar) ist. Versuchen Sie es selbst, ggf. hilft ein Blick ins Skript weiter.

Bemerkungen

(1) Ohne die Voraussetzung der totalen Differenzierbarkeit wird die Aussage des Satzes falsch.

Bemerkungen

(1) Ohne die Voraussetzung der totalen Differenzierbarkeit wird die Aussage des Satzes falsch.

(2) Umgekehrt ist die Existenz *aller* Richtungsableitungen nicht hinreichend für die totale Differenzierbarkeit, wie wir in den Übungen anhand eines Beispiels sehen werden.

Bemerkungen

(1) Ohne die Voraussetzung der totalen Differenzierbarkeit wird die Aussage des Satzes falsch.

(2) Umgekehrt ist die Existenz *aller* Richtungsableitungen nicht hinreichend für die totale Differenzierbarkeit, wie wir in den Übungen anhand eines Beispiels sehen werden.

(3) Wenn die Formel aus Satz 4 gilt, haben wir

$$\frac{\partial f}{\partial \xi}(x_0) = \nabla f(x_0)\xi = |\nabla f(x_0)| \cos(\alpha),$$

wobei α der Winkel zwischen ξ und $\nabla f(x_0)$ ist. Dessen Cosinus wird maximal für $\xi = \frac{\nabla f(x_0)}{|\nabla f(x_0)|}$, was bedeutet: *Der Gradient $\nabla f(x_0)$ weist in die Richtung des stärksten Anstiegs der Funktion f.*

Niveaulinien und -flächen

Die geometrische Interpretation des Gradienten in Bemerkung (3) wird ergänzt durch eine weitere Anwendung der Kettenregel. Gegeben sei eine differenzierbare Kurve

$$\gamma : \mathbb{R} \supset I \to N_f(c) \subset \mathbb{R}^n,$$

wobei $N_f(c) = \{x \in \mathbb{R}^n : f(x) = c\}$ die Niveaufläche (-linie im Fall $n = 2$) von f zum Wert $c \in \mathbb{R}$ ist. Dann ist $c = f \circ \gamma(t)$, daraus folgt $0 = \dfrac{d}{dt} f \circ \gamma(t) = \nabla f(\gamma(t))\gamma'(t)$, dabei ist $\gamma'(t)$ eine Tangente an $N_f(c)$. Also: *Der Gradient steht senkrecht auf den Niveauflächen von* f.