Prof. Dr. H. Brenner **Osnabrück SS 2014**

Analysis II

Vorlesung 49

Die Taylor-Formel - Vorbereitungen

Ein Polynom in n Variablen,

$$f(x_1, \ldots, x_n) = \sum_{(r_1,\ldots,r_n) \in \mathbb{N}^n} a_{(r_1,\ldots,r_n)} x_1^{r_1} x_2^{r_2} \cdots x_n^{r_n}$$

(wobei die Summe endlich ist) lässt sich entlang des Grades des Exponententupels, also

$$\mid r \mid = \mid (r_1, \ldots, r_n) \mid = \sum_{j=1}^{n} r_j$$

anordnen, also

$$f(x_1, \ldots, x_n) = \sum_{d=0}^{e} \left(\sum_{(r_1,\ldots,r_n) \in \mathbb{N}^n, \, |r|=d} a_{(r_1,\ldots,r_n)} x_1^{r_1} x_2^{r_2} \cdots x_n^{r_n} \right).$$

Für jedes $k \in \mathbb{N}$ kann man dies auch schreiben als

$$f(x_1, \ldots, x_n) = T_k(x_1, \ldots, x_n) + R_k(x_1, \ldots, x_n)$$

mit $(x = (x_1, \ldots, x_n))$

$$T_k(x) = \sum_{d=0}^{k} \left(\sum_{(r_1,\ldots,r_n) \in \mathbb{N}^n, \, |r|=d} a_{(r_1,\ldots,r_n)} x_1^{r_1} x_2^{r_2} \cdots x_n^{r_n} \right)$$

und

$$R_k(x) = \sum_{d=k+1}^{e} \left(\sum_{(r_1,\ldots,r_n) \in \mathbb{N}^n, \, |r|=d} a_{(r_1,\ldots,r_n)} x_1^{r_1} x_2^{r_2} \cdots x_n^{r_n} \right).$$

Für R_k gilt dabei

$$\lim_{x \to 0} \frac{|| R_k(x) ||}{|| x ||^k} = 0 \, .$$

Bei $k = 1$ ist

$$T_1(x) = a_{(0,\ldots,0)} + a_{(1,0,\ldots,0)} x_1 + \cdots + a_{(0,\ldots,0,1)} x_n$$

die lineare Approximation von f im Punkt $0 = (0, \ldots, 0)$, und dabei gilt für die Abweichung in der linearen Approximation die Beziehung $r(x) = \frac{R_1(x)}{||x||}$. Im Allgemeinen liefern die Polynome $T_k(x)$ bessere Approximationen im Nullpunkt als die lineare Approxmation, und mit $R_k(x)$ kann man die Abweichung kontrollieren. Entscheidend für uns ist, dass man nicht nur für Polynomfunktionen, sondern generell für hinreichend oft differenzierbare Funktionen f approximierende Polynome finden und die Abweichung gut kontrollieren kann. Dies ist der Inhalt der *Taylor-Formel für Funktionen in mehreren Variablen.*

Zu einem Vektor $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ und einem Tupel $r = (r_1, \ldots, r_n)$ aus natürlichen Zahlen setzt man abkürzend

$$x^r := x_1^{r_1} \cdots x_n^{r_n}.$$

Entsprechend schreibt man für eine Polynomfunktion abkürzend

$$f(x_1, \ldots, x_n) = \sum_{(r_1,\ldots,r_n)\in\mathbb{N}^n} a_{(r_1,\ldots,r_n)} x_1^{r_1} x_2^{r_2} \cdots x_n^{r_n} = \sum_{r\in\mathbb{N}^n} a_r x^r.$$

Die gleiche Abkürzungsphilosophie übernimmt man für Richtungsableitungen. Wenn V ein $\mathbb{R}$-Vektorraum ist mit einer Basis $w_1, \ldots, w_n$, so setzt man $D_i := D_{w_i}$,und für $r = (r_1, \ldots, r_n)$ setzt man

$$D^r := D_1^{r_1} \circ D_2^{r_2} \circ \cdots \circ D_n^{r_n}.$$

Diese Bezeichnung verwendet man insbesondere im $\mathbb{R}^n$, versehen mit der Standardbasis und den partiellen Ableitungen. Man beachte, dass man aufgrund des Satzes von Schwarz unter gewissen Differenzierbarkeitsvoraussetzungen sämtliche Reihenfolgen von Richtungsableitungen in dieser Weise ausdrücken kann. Des weiteren definieren wir für ein Tupel $r = (r_1, \ldots, r_n)$ die *Fakultät* durch

$$r! := r_1! \cdots r_n!$$

und bei $\sum_{j=1}^n r_j = k$ die *Polynomialkoeffizienten* durch

$$\binom{k}{r} := \frac{k!}{r!} = \frac{k!}{r_1! r_2! \cdots r_n!}.$$

Bevor wir die Taylor-Formel beweisen, die das lokale Verhalten einer Funktion in einer „kleinen“ offenen Ballumgebung eines Punktes beschreibt, wenden wir uns dem lokalen Verhalten in dem Punkt längs einer fixierten Richtung zu, wofür wir die Taylor-Formel in einer Variablen zur Verfügung haben. Zu einer Funktion ($G \subseteq V$, V endlichdimensionaler reeller Vektorraum)

$$f\colon G \longrightarrow \mathbb{R}$$

ist die Differenzierbarkeit im Punkt $P \in G$ in Richtung $v \in V$ äquivalent zur Differenzierbarkeit der Funktion

$$h\colon I \longrightarrow \mathbb{R},\ t \longmapsto h(t) = f(P + tv),$$

für $t = 0$, wobei I ein geeignetes reelles Intervall ist. Wir werden zunächst zeigen, dass eine entsprechende Beziehung auch für höhere Ableitungen gilt.

Satz 49.1. *Sei $G \subseteq \mathbb{R}^n$ offen,*

$$f\colon G \longrightarrow \mathbb{R}$$

eine k-mal stetig differenzierbare Funktion, $P \in G$ ein Punkt und $v = (v_1, \ldots, v_n) \in \mathbb{R}^n$ eine fixierte Richtung. Es sei

$$h\colon I \longrightarrow \mathbb{R},\, t \longmapsto h(t) = f(P + tv),$$

wobei I ein offenes Intervall um 0 sei mit $P + tv \in G$ für alle $t \in I$. Dann ist h ebenfalls k-mal stetig differenzierbar, und es gilt

$$h^{(k)}(t) \;=\; \sum_{|r|=k} \frac{k!}{r!} D^r f(P + tv) \cdot v^r$$

für alle $t \in I$.

Beweis. Wir zeigen durch Induktion, dass

$$h^{(k)}(t) \;=\; \sum_{(i_1,\ldots,i_k)\in\{1,\ldots,n\}^k} D_{i_k} \cdots D_{i_1} f(P + tv) \cdot v_{i_1} \cdots v_{i_k}$$

gilt. Hier wird also über jede Richtungsreihenfolge der Länge k aufsummiert, später werden wir unter Verwendung des Satzes von Schwarz gleiche Summanden zusammenfassen. Für $k = 1$ ist

$$\begin{aligned}
h'(t) &= (D_v f)(P + tv) \\
&= (Df)_{P+tv}(v) \\
&= (Df)_{P+tv}\left(\sum_{j=1}^{n} v_j e_j\right) \\
&= \sum_{j=1}^{n} v_j \cdot (Df)_{P+tv}(e_j) \\
&= \sum_{j=1}^{n} v_j \cdot D_j f(P + tv).
\end{aligned}$$

Der Induktionsschluss ergibt sich aus

$$\begin{aligned}
h^{(k+1)}(t) &= (h^{(k)})'(t) \\
&= \left(\sum_{(i_1,\ldots,i_k)\in\{1,\ldots,n\}^k} D_{i_k} \cdots D_{i_1} f(P + tv) \cdot v_{i_1} \cdots v_{i_k}\right)'(t) \\
&= \sum_{j=1}^{n} D_j \left(\sum_{(i_1,\ldots,i_k)\in\{1,\ldots,n\}^k} D_{i_k} \cdots D_{i_1} f(P + tv) \cdot v_{i_1} \cdots v_{i_k}\right) v_j \\
&= \sum_{j=1}^{n} \left(\sum_{(i_1,\ldots,i_k)\in\{1,\ldots,n\}^k} D_j D_{i_k} \cdots D_{i_1} f(P + tv) \cdot v_{i_1} \cdots v_{i_k} v_j\right) \\
&= \sum_{(i_1,\ldots,i_k,j)\in\{1,\ldots,n\}^{k+1}} D_j D_{i_k} \cdots D_{i_1} f(P + tv) \cdot v_{i_1} \cdots v_{i_k} v_j.
\end{aligned}$$

Aufgrund des Satzes von Schwarz kommt es nicht auf die Reihenfolge der Richtungableitungen an, d.h. zwei Summanden in der obigen Summe stimmen überein, wenn darin die jeweiligen Richtungsableitungen gleichhäufig vorkommen. Die Anzahl der Tupel $(i_1, \ldots, i_k)$ in $\{1, \ldots, n\}^k$, bei denen die Zahl i genau r_i-mal vorkommt, wird durch die Polynomialkoeffizienten

$$\binom{k}{r} = \frac{k!}{r!} = \frac{k!}{r_1! \cdots r_n!}$$

beschrieben. Daraus ergibt sich die Behauptung. □

Definition 49.2. Es sei $G \subseteq \mathbb{R}^n$ eine offene Teilmenge,

$$f \colon G \longrightarrow \mathbb{R}$$

eine k-mal stetig-differenzierbare Funktion und $P \in G$. Dann heißt

$$\sum_{r=(r_1,\ldots,r_n),\, |r| \leq k} \frac{1}{r!} D^r f(P) \cdot v^r$$

das *Taylor-Polynom vom Grad* $\leq k$ zu f in P.

Es liegt also ein Polynom in den (verschobenen) Variablen $v_1, \ldots, v_n$ vor. Wenn $P = (a_1, \ldots, a_n)$ ist, so schreibt man meistens $x_i - a_i$ statt v_i, wobei die x_i die Standardkoordinaten des $\mathbb{R}^n$ bezeichnen. Man spricht auch vom *Taylor-Polynom der Ordnung* k oder einfach vom k-ten Taylor-Polynom.

Das 0-teTaylor-Polynom ist das konstante Polynom, das durch den Funktionswert $f(P)$ gegeben ist, das 1-teTaylor-Polynom ist die lineare Approximation von f in P und das 2-teTaylor-Polynom ist die *quadratische Approximation* von f in P.

Bemerkung 49.3. Ein Polynom f vom Grad d stimmt mit seinem Taylor-Polynom vom Grad $k \geq d$ im Nullpunkt $0 = (0, \ldots, 0)$ überein. Wegen der Additivität der Richtungsableitungen muss man dies nur für $f = a x_1^{r_1} \cdots x_n^{r_n}$ überprüfen. Es ist aber

$$D^r f(0) = D_1^{r_1} \cdots D_n^{r_n} f(0) = (r_1!) \cdots (r_n!)\, a = r! a$$

und

$$D^s f(0) = 0$$

für jedes n-Tupel $s = (s_1, \ldots, s_n) \neq r$.

Wenn man zu einem Polynom f die Taylor-Polynome in einem Punkt $P = (a_1, \ldots, a_n)$ berechnen möchte, so kann man (neben der Berechnung der Ableitungen) auch folgendermaßen vorgehen: Man schreibt das Polynom f in den Variablen $y_i = x_i - a_i$. Dazu ersetzt man in f die Variablen x_i durch

$$x_i = x_i - a_i + a_i = y_i + a_i$$

und rechnet dies aus, bis ein Polynom in y_i dasteht. Aus diesem Polynom sind die Taylor-Polynome im Entwicklungspunkt P direkt ablesbar.

Beispiel 49.4. Wir betrachten die Funktion

$$f\colon \mathbb{R}^2 \longrightarrow \mathbb{R},\ (x,y) \longmapsto e^y \sin x - 3xy,$$

und wollen die Taylor-Polynome bis zur Ordnung 3 dazu im Nullpunkt berechnen. Das Taylor-Polynom der Ordnung 0 ist das konstante Nullpolynom, da $f(0,0) = 0$ ist. Für das Taylor-Polynom der Ordnung 1 müssen wir die beiden partiellen Ableitungen ausrechnen. Diese sind

$$\frac{\partial f}{\partial x} = e^y \cos x - 3y \text{ und } \frac{\partial f}{\partial y} = e^y \sin x - 3x$$

mit den Werten 1 und 0. Daher ist x die lineare Approximation zu f, also das Taylor-Polynom der Ordnung 1. Für das Taylor-Polynom der Ordnung 2 berechnen wir die zweiten Ableitungen, diese sind

$$\frac{\partial}{\partial x}\frac{\partial f}{\partial x} = -e^y \sin x\ ,$$

$$\frac{\partial}{\partial y}\frac{\partial f}{\partial x} = e^y \cos x - 3$$

und

$$\frac{\partial}{\partial y}\frac{\partial f}{\partial y} = e^y \sin x\ .$$

Die Werte dieser zweiten partiellen Ableitungen sind $0, -2, 0$, sodass das zweite Taylor-Polynom (also die quadratische Approximation) gleich

$$x - 2xy$$

ist. Für das Taylor-Polynom der Ordnung 3 berechnen wir die dritten Ableitungen, diese sind

$$\frac{\partial}{\partial x}\frac{\partial}{\partial x}\frac{\partial f}{\partial x} = -e^y \cos x\ ,$$

$$\frac{\partial}{\partial y}\frac{\partial}{\partial x}\frac{\partial f}{\partial x} = -e^y \sin x\ ,$$

$$\frac{\partial}{\partial y}\frac{\partial}{\partial y}\frac{\partial f}{\partial x} = e^y \cos x\ ,$$

und

$$\frac{\partial}{\partial y}\frac{\partial}{\partial y}\frac{\partial f}{\partial y} = e^y \sin x\ .$$

Die Werte dieser dritten partiellen Ableitungen sind $-1, 0, 1, 0$, sodass (wegen $(3,0)! = 6$ und $(1,2)! = 2$) das dritte Taylor-Polynom gleich

$$x - 2xy - \frac{1}{6}x^3 + \frac{1}{2}xy^2$$

ist.

Satz 49.5. *Sei $G \subseteq \mathbb{R}^n$ offen,*

$$f\colon\, G \longrightarrow \mathbb{R}$$

eine $(k+1)$-mal stetig differenzierbare Funktion, $P \in G$ ein Punkt und $v \in \mathbb{R}^n$ derart, dass die Strecke von P nach $P+v$ ganz in G liegt. Dann gibt es ein $c \in [0,1]$ mit

$$f(P+v) \;=\; \sum_{|r|\leq k} \frac{1}{r!} D^r f(P) \cdot v^r + \sum_{|r|=k+1} \frac{1}{r!} D^r f(P+cv) \cdot v^r .$$

Beweis. Die Funktion

$$h\colon\,]-\delta, 1+\delta[\longrightarrow \mathbb{R},\, t \longmapsto h(t) = f(P+tv),$$

(mit einem geeigneten $\delta > 0$) ist nach Satz 49.1 $(k+1)$-mal differenzierbar. Aufgrund der Taylor-Formel für eine Variable gibt es ein $c \in [0,1]$[1] mit

$$h(1) \;=\; \sum_{j=0}^{k} \frac{h^{(j)}(0)}{j!} + \frac{h^{(k+1)}(c)}{(k+1)!} .$$

Wir drücken die einzelnen Summanden mit Hilfe von Satz 49.1 aus und erhalten

$$\begin{aligned} f(P+v) &= h(1) \\ &= \sum_{j=0}^{k} \frac{h^{(j)}(0)}{j!} + \frac{h^{(k+1)}(c)}{(k+1)!} \\ &= \sum_{|r|\leq k} \frac{1}{r!} D^r f(P) \cdot v^r + \sum_{|r|=k+1} \frac{1}{r!} D^r f(P+cv) \cdot v^r . \end{aligned}$$

□

Die Taylor-Formel

Satz 49.6. *Sei $G \subseteq \mathbb{R}^n$ offen,*

$$f\colon\, G \longrightarrow \mathbb{R}$$

eine k-mal stetig differenzierbare Funktion, $P \in G$ ein Punkt und $\epsilon > 0$ derart, dass $U(P,\epsilon) \subseteq G$ ist. Dann gilt für alle v mit $P+v \in U(P,\epsilon)$ die Beziehung

$$f(P+v) \;=\; \sum_{|r|\leq k} \frac{1}{r!} D^r f(P) \cdot v^r + R_k(v),$$

wobei

$$\lim_{v \to 0} \frac{||\, R_k(v)\, ||}{||\, v\, ||^k} = 0$$

ist.

[1] Der Beweis der Taylor-Formel für eine Variable zeigt, dass c zwischen den beiden beteiligten Punkten gewählt werden kann.

Beweis. Nach Satz 49.3 gibt es zu jedem $v \in U(0, \epsilon)$ ein (von v abhängiges) $c \in [0, 1]$ mit

$$\begin{aligned} f(P+v) &= \sum_{|r| \leq k-1} \frac{1}{r!} D^r f(P) \cdot v^r + \sum_{|r|=k} \frac{1}{r!} D^r f(P+cv) \cdot v^r \\ &= \sum_{|r| \leq k} \frac{1}{r!} D^r f(P) \cdot v^r + \sum_{|r|=k} \frac{1}{r!} \left(D^r f(P+cv) - D^r f(P) \right) v^r. \end{aligned}$$

Die rechte Summe ist also die Abweichungsfunktion R_k, die wir abschätzen müssen. Wegen

$$\begin{aligned} \|\, R_k(v) \,\| &\leq \sum_{|r|=k} \frac{1}{r!} \,\|\, D^r f(P+cv) - D^r f(P) \,\| \cdot \|\, v^r \,\| \\ &= \sum_{|r|=k} \frac{1}{r!} \,\|\, D^r f(P+cv) - D^r f(P) \,\| \cdot |v_1^{r_1}| \cdots |v_n^{r_n}| \\ &\leq \sum_{|r|=k} \frac{1}{r!} \,\|\, D^r f(P+cv) - D^r f(P) \,\| \cdot \|\, v \,\|^{r_1} \cdots \|\, v \,\|^{r_n} \\ &= \sum_{|r|=k} \frac{1}{r!} \,\|\, D^r f(P+cv) - D^r f(P) \,\| \cdot \|\, v \,\|^k \end{aligned}$$

ist

$$\frac{\|\, R_k(v) \,\|}{\|\, v \,\|^k} \leq \sum_{|r|=k} \frac{1}{r!} \,\|\, D^r f(P+cv) - D^r f(P) \,\| \, .$$

Da nach Voraussetzung die k-ten Richtungsableitungen stetig sind, existiert für jede einzelne Funktion $D^r f(P+cv) - D^r f(P)$ der Limes für $v \to 0$ und ist gleich 0. Daher gilt dies auch für die Summe rechts und damit auch für den Ausdruck links. $\square$