Höheren Ableitungen und Richtungsableitungen

Erinnerung:
Sind V, W endlich-dimensionale, normierte $\mathbb{R}$-Vektorräume, $U \subseteq V$ offen, $n \in \mathbb{N}$, und ist $f : U \to W$ in $p \in U$ mindestens n-mal differenzierbar, dann ist die n-fache Ableitung im Punkt p eine multilineare Abbildung

$$d^n f(p) : \underbrace{V \times ... \times V}_{n\text{-mal}} \longrightarrow W.$$

Die Werte dieser multilinearen Abbildung erhält man im Fall einer reellwertigen Funktion durch Richtungsableitungen.

Proposition (8.3)

Sei $U \subseteq V$ offen, $p \in U$ und $f : U \to \mathbb{R}$ eine n-mal in p differenzierbare Funktion. Dann gilt für alle $v_1, ..., v_n \in V$ jeweils

$$d^n f(p)(v_1, ..., v_n) \quad = \quad \partial_{v_1} \cdots \partial_{v_n} f(p).$$

Beweis von Proposition 8.3 (Skizze)

- Der Beweis erfolgt durch vollständige Induktion über $n \in \mathbb{N}_0$. Den Induktionsanfang erhält man durch die triviale Gleichung $f(p) = f(p)$. Der Induktionsschritt basiert, wie wir sehen werden, auf der bereits bekannten Gleichung

$$df(p)(v) = \partial_v f(p)$$

für alle $v \in V$ (Proposition 6.4).

- Sei nun $n \in \mathbb{N}$ und die Aussage für $n-1$ vorausgesetzt. Die Voraussetzung liefert

$$d^{n-1}f(p+h)(v_2, ..., v_n) \quad = \quad \partial_{v_2} \cdots \partial_{v_n} f(p+h)$$

für $v_2, ..., v_n \in V$ und hinreichend kleines $h \in V$.

Beweis von Proposition 8.3 (Skizze)

- Die Differenzierbarkeit von $g = d^{n-1}f$ im Punkt p liefert die Gleichung $g(p+h) = g(p) + dg(p)(h) + \psi(h)$ für hinreichend kleines h, wobei $\lim_{h\to 0} \|h\|^{-1}\psi(h) = 0$ gilt.
- Für die Hilfsfunktion $\tilde{f} = \partial_{v_2} \cdots \partial_{v_n} f$ erhalten wir auf Grund der Gleichung auf der letzten Seite

$$\tilde{f}(p+h) \quad = \quad g(p)(v_2, ..., v_n) + dg(p)(h, v_2, ..., v_n) + \varphi(h)$$

wobei $\lim_{h\to 0} \|h\|^{-1}\varphi(h) = 0$ gilt.

- Die Hilfunktion $\tilde{f}$ ist somit in p total differenzierbar. Mit Proposition 6.4 erhalten wir

$$\begin{aligned} \partial_{v_1}\partial_{v_2} \cdots \partial_{v_n} f(p) \quad &= \quad \partial_{v_1}\tilde{f}(p) \quad = \quad d\tilde{f}(p)(v_1) \\ &= \quad dg(p)(v_1, ..., v_n) \quad = \quad d^n f(p)(v_1, ..., v_n). \end{aligned}$$

Definition der Hessematrix einer Funktion

Definition (8.4)

Sei $U \subseteq \mathbb{R}^n$ offen, $p \in \mathbb{R}^n$ und $f : \mathbb{R}^n \to \mathbb{R}$ eine in p mindestens zweimal differenzierbare Funktion. Sei $\mathscr{E} = (e_1, ..., e_n)$ die Einheitsbasis des $\mathbb{R}^n$. Dann wird die Darstellungsmatrix $\mathscr{H}(f)(p) = \mathscr{M}_{\mathscr{E}}(d^2 f(p))$ mit den Einträgen

$$a_{ij} \quad = \quad d^2 f(p)(e_i, e_j) \quad = \quad \partial_{ij} f(p)$$

die Hesse-Matrix von f an der Stelle p genannt.

Eindimensionale Taylorpolynome

Erinnerung:
Sei $p \in \mathbb{N}_0$, sei $I \subseteq \mathbb{R}$ ein offenes Intervall, $a \in I$ und $f : I \to \mathbb{R}$ eine in a mindestens m-mal differenzierbare Funktion. Dann ist das m-te Taylorpolynom

$$\tau_m(f,a)(x) \quad = \quad \sum_{k=0}^{m} \frac{1}{k!} f^{(k)}(a)(x-a)^k.$$

von f an der Stelle a definiert.

Satz (8.5)

Sei $m \in \mathbb{N}$, $I \subseteq \mathbb{R}$ ein offenes Intervall und $f : I \to \mathbb{R}$ eine mindestens m-mal differenzierbare Funktion. Seien $a, x \in I$ mit $x > a$. Dann gibt es einen Punkt $\xi \in \,]a, x[$ mit

$$f(x) \quad = \quad \tau_{m-1}(f,a)(x) + \frac{1}{m!} f^{(m)}(\xi)(x-a)^m.$$

Beweis von Satz 8.5

z.zg.: Es gibt ein $\xi \in]a, x[$ mit

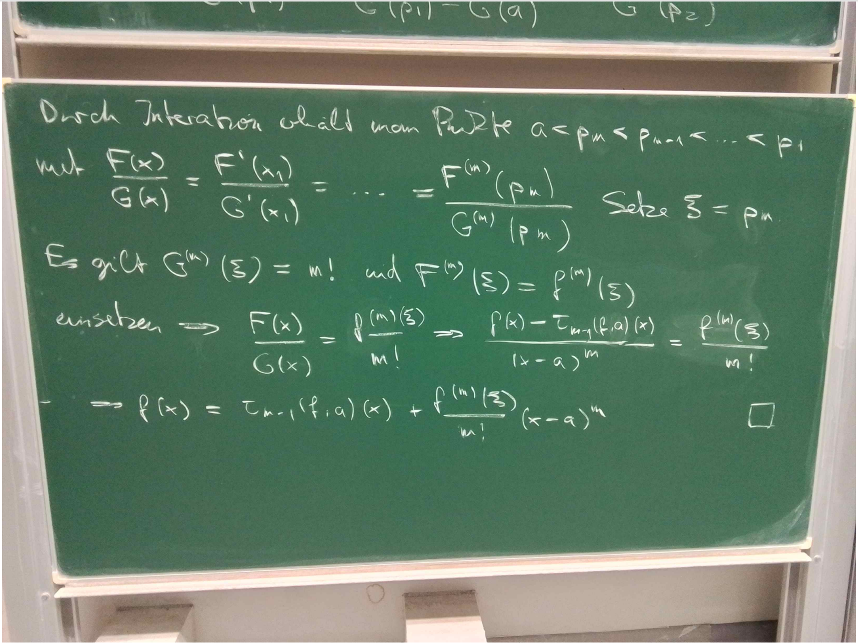
$$f(x) = T_{m-1}(f, a)(x) + \frac{1}{m!} f^{(m)}(\xi)(x-a)^m$$

Definiere die Hilfsfunktionen $F(t) = f(t) - T_{m-1}(f, a)(t)$ und $G(t) = (t-a)^m$. verallgemeinerter Mittelwertsatz

$$\Rightarrow \exists p_1 \in]a, x[\text{ mit } \frac{F(x)}{G(x)} = \frac{F(x) - F(a)}{G(x) - G(a)} = \frac{F'(p_1)}{G'(p_1)}$$

verallg. MWS $\Rightarrow \exists p_2 \in]a, p_1[$ mit

$$\frac{F'(p_1)}{G'(p_1)} = \frac{F'(p_1) - F'(a)}{G'(p_1) - G'(a)} = \frac{F''(p_2)}{G''(p_2)}$$



Korrektur 1. Zeile: Durch Iteration erhält man...

Definition (8.6)

Sei $n \in \mathbb{N}$, $U \subseteq V$ offen, $p \in U$ und $f : U \to \mathbb{R}$ eine im Punkt p mindestens m-mal differenzierbare Funktion. Dann bezeichnet man die Funktion für $x \in V$ gegeben

$$\tau_m(f,p)(x) \quad = \quad f(p) + \sum_{k=1}^{m} \frac{1}{k!} d^k f(p)(\underbrace{x-p, ..., x-p}_{k\text{-mal}})$$

als Taylorpolynom m-ten Grades von f an der Stelle p.

Die Approximationseigenschaft der Taylorpolynome

Satz (8.7)

Sei $U \subseteq \mathbb{R}^n$ offen und $f : U \to \mathbb{R}$ eine m-mal stetig differenzierbare Funktion. Dann gibt es zu jedem Punkt $p \in U$ eine Umgebung U_0 des Nullpunkts und eine Funktion $\psi : U_0 \to \mathbb{R}$ mit

$$f(p+h) \quad = \quad \tau_m(f,p)(p+h) + \psi(h)$$

für alle $h \in U_0$ und

$$\lim_{h \to 0} \frac{\psi(h)}{\|h\|^m} = 0.$$

Beweis von Satz 8.7 (Skizze)

- Der Beweis erfolgt durch Anwendung von Satz 8.5 auf die eindimensionale Hilfsfunktion $g(t) = f(p + th)$.
- Mit Hilfe der mehrdimensionalen Kettenregel werden die höheren Ableitungen $g^{(k)}(t)$ durch die Richtungsableitungen von f ausgedrückt.
- Nach Proposition 8.3 hängen die Richtungsableitungen mit den höheren totalen Ableitungen von f zusammen.
- Damit wiederum stellt man eine Verbindung zwischen den Taylorpolynomen von g und f her. Die Darstellung von $g(1)$ durch das $(m-1)$-te Taylorpolynom von g plus „Fehlerterm" nach Satz 8.5 liefert die gewünschte Approximationseigenschaft von $\tau_m(f, p)$.

Approximation zweiter Ordnung

Folgerung (8.8)

Sei $U \subseteq \mathbb{R}^n$ offen, $f : U \to \mathbb{R}$ eine Funktion und $a \in U$ ein Punkt, in dem f zweimal differenzierbar ist. Dann gibt es eine Funktion $\psi : U_a \to \mathbb{R}$ mit

$$f(a+h) = f(a) + df(a) \cdot h + \frac{1}{2}\, {}^{\mathrm{t}}h \cdot \mathscr{H}(f)(a) \cdot h + \psi(h)$$

und $\lim_{h \to 0} \|h\|^{-2}\, \psi(h) = 0$.

Lokale Extrema im Mehrdimensionalen

Definition (8.9)

Sei $U \subseteq \mathbb{R}^n$ eine beliebige Teilmenge, $a \in U$ und $f : U \to \mathbb{R}$ eine reellwertige Funktion. Man sagt, f hat im Punkt a ein

(i) lokales Maximum, wenn eine Umgebung $U' \subseteq \mathbb{R}^n$ von a existiert, so dass $f(a) \geq f(x)$ für alle $x \in U \cap U'$ gilt,

(ii) isoliertes lokales Maximum, wenn U so gewählt werden kann, dass sogar $f(a) > f(x)$ für alle $x \in (U \cap U') \setminus \{a\}$ erfüllt ist.

Entsprechend definiert man lokale Minima und isolierte lokale Minima. Wie in der Analysis einer Variablen verwenden wir den Begriff Extremum als Oberbegriff für Minima und Maxima.

Notwendiges Kriterium für Extrema

Satz (8.10)

Sei $U \subseteq \mathbb{R}^n$ offen, $f : U \to \mathbb{R}$ differenzierbar und $a \in U$ ein lokales Extremum von f. Dann gilt $df(a) = 0$.

Einen Punkt, in dem die erste Ableitung einer Funktion f verschwindet, bezeichnet man als kritische Stelle von f.

Beweis von Satz 8.10

geg: $f: U \to \mathbb{R}$, total diff'bar in a

Setze voraus, dass a ein lokales Maximum von f ist.
(Der Fall eines Minimums läuft analog.)
Verkleinere die offene Umg. U von a so, dass $f(x) \leq f(a)$ (*)
für alle $x \in U$ gilt. z.zg: $df(a) = 0$

gleichbedeutend: $\partial_k f(a) = 0$ für $1 \leq k \leq n$.
Sei $k \in \{1, \dots, n\}$. Betrachte die Hilfsfkt. $g:]-\varepsilon, \varepsilon[\to \mathbb{R}$,
$t \mapsto p + t e_k$, wobei $\varepsilon \in \mathbb{R}^+$ so gewählt ist, dass

$\dots, \text{wobei } \varepsilon \in \mathbb{R}$ so gewählt ist, dass

$p + te_k \in U$ für $|t| < \varepsilon$ gilt. $\Rightarrow g'(0) = \partial_k f(a)$

Wegen $(*)$ gilt $g(0) \geq g(t)$ für alle $t \in \,]-\varepsilon, \varepsilon[$

$\Rightarrow$ 0 ist lokales Maximum von g $\Rightarrow \partial_k f(a) = g'(0) = 0$. $\square$

hinreichende Kriterien für Extrema

Satz (8.11)

Sei $U \subseteq \mathbb{R}^n$ offen, $f : U \to \mathbb{R}$ eine zweimal stetig differenzierbare Funktion und $a \in U$ eine kritische Stelle von f.

(i) Ist $\mathscr{H}(f)(a)$ positiv definit, dann besitzt f in a ein isoliertes lokales Minimum.

(ii) Ist $\mathscr{H}(f)(a)$ negativ definit, dann besitzt f in a ein isoliertes lokales Maximum.

(iii) Ist $\mathscr{H}(f)(a)$ indefinit, dann hat f in a kein lokales Extremum.

Beweis von Satz 8.11

zu i) geg: zweimal stetig diff'bare Fkt.

$f: U \to \mathbb{R}$, $a \in U$ kritische Stelle

(d.h. $df(a) = 0$)

Vor.: $\mathcal{H}(f)(a)$ ist positiv definit

Folgerung 8.8 $\Rightarrow$ Nach hinreichender Verkleinerung der Umgebung U von a gilt

$$f(a+h) = f(a) + \frac{1}{2} {}^t h \, \mathcal{H}(f)(a) \, h + \varphi(a)$$

für alle $h \in \mathbb{R}^n$ mit $a+h \in U$, wobei die

Funktion ψ die Eigenschaft

$$\lim_{h \to 0} \frac{1}{\|h\|^2} \psi(h) = 0 \quad \text{besitzt}$$

z.zg.: Für hinreichend kleines $\|h\|$ gilt

$$^t h \mathcal{H}(f)(a) h + \psi(h) > 0$$

Betrachte $S = \{ h \in \mathbb{R}^n \mid \|h\| = 1 \}$

$h \mapsto {}^t h \mathcal{H}(f)(a) h$ nimmt auf S nur positive Werte an (weil die Hessematrix positiv definit ist) $\Rightarrow$ Die Funktion nimmt auf S ein lokales Minimum $m > 0$ an

$\lim_{h \to 0} \frac{1}{\|h\|^2} \psi(h) = 0 \Rightarrow$ Nach Verkleinerung von U

...ositive
...definit
...n
...g von U

können wir $\frac{1}{\|h\|}|\varphi(h)| < \frac{1}{4} m \quad \forall h \in \mathbb{R}^n$ mit $a+h \in U$
und $h \neq 0_{\mathbb{R}^n}$ annehmen $\Rightarrow |\varphi(h)| < \frac{1}{4} m \|h\|^2$
für diese h.

Sei nun ein solches h vorgeg. und $h_0 = \frac{1}{\|h\|} h$

$$\Rightarrow h_0 \in S \Rightarrow {}^t h\, \mathcal{H}(f)(a)\, h = \|h\|^2\, {}^t h_0\, \mathcal{H}(f,a)\, h_0$$

$$\geq m \|h\|^2$$ $\Rightarrow$ erhalte

$$\frac{1}{2}\, {}^t h\, \mathcal{H}(f)(a)\, h + \varphi(h) > \frac{1}{2} m \|h\|^2 - \frac{1}{4} m \|h\|^2$$

$$= \frac{1}{4} m \|h\|^2 > 0$$

<u>zu (ii)</u> Wende Teil (i) auf $-f$ an.

zu (iii) Vor.: $\mathcal{H}(f)(a)$ ist indefinit

$\Rightarrow$ Es gibt einen Vektor v mit ${}^t v\, \mathcal{H}(f)(a) v > 0$

und einen Vektor w mit ${}^t w\, \mathcal{H}(f)(a) w < 0$.

Verkleinert man U, dann erhält wie unter (i) Darstellungen

$$f(a+tv) = f(a) + t^2\, {}^t v\, \mathcal{H}(f)(a)\, v + \psi(tv)$$

$$f(a+tw) = f(a) + t^2\, {}^t w\, \mathcal{H}(f)(a)\, w + \psi(tw)$$

Wie unter (i) zeigt man, dass ein $\varepsilon \in \mathbb{R}^+$ existiert mit $|\psi(tv)|, |\psi(tw)| < t^2 \min \underbrace{\{ {}^t v\, \mathcal{H}(f)(a) v, -{}^t w\, \mathcal{H}(f)(a) w \}}_{> 0}$

Dann gilt $f(a+tv) > 0$ und $f(a+tw) < 0$ für alle t mit $0 < |t| < \varepsilon$, $\Rightarrow$ a kein lokales Extremum. $\square$

<u>Anwendung</u>: Bestimmung der lokalen Extrema von

$$f(x,y) = y^2(x-1) + x^2(x+1)$$

1. Schritt: Bestimmung der kritischen Stellen

$$df(x,y) = \left(\partial_1 f(x,y) \quad \partial_2 f(x,y) \right)$$

$$\partial_1 f(x,y) = y^2 + 3x^2 + 2x, \quad \partial_2 f(x,y) = 2y(x-1)$$

(x,y) ist kritische Stelle von $f \iff df(x,y) = 0$

$\iff \partial_1 f(x,y) = 0$ <u>und</u> $\partial_2 f(x,y) = 0$

$\partial_2 f(x,y) = 0 \Rightarrow 2y(x-1) = 0 \Rightarrow x = 1$ oder $y = 0$

<u>1. Fall</u>: $x = 1$ $\quad \partial_1 f(x,y) = 0 \iff y^2 + 3 \cdot 1^2 + 2 \cdot 1 = 0$

$\partial_2 f(x,y) = 0 \Rightarrow 2y(x-1) = 0 \Rightarrow x = 1$ oder $y = 0$

$\Leftrightarrow y^2 + 5 = 0$ ↯ Ergebnis: Es gibt keine kritische Stelle (x,y) mit $x = 1$

2. __Fall__: $y = 0$ $\quad \partial_1 f(x,y) = 0 \Leftrightarrow 3x^2 + 2x = 0 \Leftrightarrow$

$$3x\left(x + \frac{2}{3}\right) = 0 \Leftrightarrow x = 0 \vee x = -\frac{2}{3}$$

Ergebnis: f hat genau zwei kritische Stellen, nämlich $(0,0)$, $(-\frac{2}{3}, 0)$

$$\mathcal{H}(f)(x,y) = \begin{pmatrix} \partial_{11} f(x,y) & \partial_{12} f(x,y) \\ \partial_{21} f(x,y) & \partial_{22} f(x,y) \end{pmatrix} = \begin{pmatrix} 6x+2 & 2y \\ 2y & 2(x-1) \end{pmatrix}$$

$$\mathcal{H}(f)(0,0) = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}$$

${}^t e_1 \mathcal{H}(f)(0,0) e_1 = 2 > 0$

${}^t e_2 \mathcal{H}(f)(0,0) e_2 = -2 < 0$

$\Rightarrow$ Matrix ist indefinit $\Rightarrow$ __kein__ lokales Extremum im Punkt $(0,0)$

$$\mathcal{H}(f)\left(-\frac{2}{3}, 0\right) = \begin{pmatrix} -2 & 0 \\ 0 & -\frac{10}{3} \end{pmatrix}$$

$\mathcal{H}(f)(-\frac{2}{3}, 0)$ ist negativ def. $\Rightarrow$ isoliertes lokales Max