

-74-

Und $\partial_x^m f := \frac{\partial^{m_1} f}{\partial x_1^{m_1} \dots \partial x_n^{m_n}}$

Dann ist:

$$T_k f(y; x_0) = \sum_{|m| \leq k} \frac{1}{m!} \partial_x^m f(x_0) y^m$$

$k=2$:

$$T_2 f(y; x_0) = f(x_0) + \sum_{i=1}^n \frac{\partial f}{\partial x_i}(x_0) y_i$$

$$+ \sum_{m_1 + \dots + m_n = 2} \frac{1}{m_1! \dots m_n!} \frac{\partial^2 f}{\partial x_1^{m_1} \dots \partial x_n^{m_n}}(x_0) y_1^{m_1} \dots y_n^{m_n}$$

Multindices $(m_1, \dots, m_n)$ mit

$$m_1 + \dots + m_n = 2.$$

-75-

Für jedes $1 \leq i \leq n$:

$$\begin{array}{ccc} (0, \dots, 2, \dots, 0) & \frac{1}{2!} \frac{\partial^2 f}{\partial x_i^2}(x_0) y_i^2 \\ \uparrow & \\ i & \end{array}$$

Für jedes $1 \leq i < j \leq n$:

$$\begin{array}{ccc} (0, \dots, 1, 0, \dots, 0, 1, 0, \dots, 0) & \frac{1}{1!1!} \frac{\partial^2 f}{\partial x_i \partial x_j}(x_0) y_i y_j \\ \uparrow \quad \quad \uparrow & \\ i \quad \quad j & \end{array}$$

Quadratische Summand:

$$\frac{1}{2} \sum_{i=1}^n \frac{\partial^2 f}{\partial x_i^2}(x_0) y_i^2 + \sum_{i < j} \frac{\partial^2 f}{\partial x_i \partial x_j}(x_0) y_i y_j$$

$$= \frac{1}{2} \left\{ \sum_{i=1}^n \frac{\partial^2 f}{\partial x_i^2}(x_0) y_i^2 + 2 \sum_{i < j} \frac{\partial^2 f}{\partial x_i \partial x_j}(x_0) y_i y_j \right\}$$

- 76 -

$$= \frac{1}{2} \sum_{1 \leq i, j \leq n} \frac{\partial^2 f}{\partial x_i \partial x_j}(x_0) y_i y_j$$

$$= \frac{1}{2} {}^t y H_{\text{ess}_f}(x_0) y$$

wobei $y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$, ${}^t y = (y_1, \dots, y_n)$

Also:

$$T_2 f(y; x_0) = f(x_0) + \langle \nabla f(x_0), y \rangle + \frac{1}{2} {}^t y H_{\text{ess}_f}(x_0) y$$

Satz 2.48 Sei $f \in C^k(X, \mathbb{R})$:

Dann: $f(x) = T_k f(x - x_0, x_0) + E_k f(x, x_0)$

wobei $\lim_{\substack{x \rightarrow x_0 \\ \neq}} \frac{E_k f(x, x_0)}{\|x - x_0\|^k} = 0.$

- 77 - .

Fall $n = 2$: $f \in C^2(X, \mathbb{R})$

$$f(x_0 + y) = f(x_0) + \langle \nabla f(x_0), y \rangle + \frac{1}{2} {}^t y H_{\frac{f}{f}}(x_0) y + E_2(x_0 + y, x_0)$$

$$\text{Mit } \lim_{\substack{y \rightarrow 0 \\ y \neq 0}} \frac{E(x_0 + y, x_0)}{\|y\|^2} = 0.$$

Beispiel: $f(x_1, x_2) = x_1^3 + x_2^3 - 3x_1x_2$

$$x_0 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

$$\nabla f(x_1, x_2) = \begin{pmatrix} 3x_1^2 - 3x_2 \\ 3x_2^2 - 3x_1 \end{pmatrix}$$

$$\nabla f(1, 1) = 0.$$

Wie kommt man auf das
Taylorpolynom?

Sei $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ glatt.

Seien $x_0, y \in \mathbb{R}^2$ und:

$$h(t) := f(x_0 + ty) \quad t \in \mathbb{R}.$$

Dann ist $h: \mathbb{R} \rightarrow \mathbb{R}$ eine glatte

Funktion und aus Satz 4.43,

Analysis I folgt, dass gegeben

$k \geq 1$ es für jedes $t \in \mathbb{R}_{\geq 0}$ ein

$\xi \in]0, t[$ gibt mit:

$$\cancel{h(t)} \quad h(t) = \sum_{j=0}^k \frac{h^{(j)}(0)}{j!} t^j + \frac{h^{(k+1)}(\xi)}{(k+1)!} t^{k+1}$$

- 75 -

Insbesondere für $t=1$:

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

Es sind dann die Ableitungen

$$h(0), h'(0), h''(0), \dots$$

die zu den Taylorpolynomen führen.

Wir illustrieren dies im Fall

$$h'(0) \text{ und } h''(0).$$

Um die Ableitungen von h zu

berechnen bemerken wir, dass h

Verketzung von zwei Abbildungen ist:

$$\mathbb{R} \xrightarrow{c} \mathbb{R}^2 \xrightarrow{f} \mathbb{R}$$

$$t \mapsto x_0 + ty$$

Es folgt dann aus dem auf Seite 51 behandelten Spezialfall der Kettenregel: (Satz 2.35)

$$h'(t) = (f \circ c)'(t)$$

$$= \langle \nabla f(c(t)), c'(t) \rangle$$

$$= \frac{\partial f}{\partial x_1}(c(t)) y_1 + \frac{\partial f}{\partial x_2}(c(t)) y_2$$

Woraus mit zweifacher Anwendung:

$$h''(t) = \frac{\partial^2 f}{\partial x_1^2}(c(t)) y_1^2 + \frac{\partial^2 f}{\partial x_2 \partial x_1}(c(t)) y_2 y_1$$

$$+ \frac{\partial^2 f}{\partial x_1 \partial x_2}(c(t)) y_1 y_2 + \frac{\partial^2 f}{\partial x_2^2}(c(t)) y_2^2$$

$$S_{a+3} \quad 2.43 \Rightarrow$$

$$= \frac{\partial^2 f}{\partial x_1^2} (c(t)) y_1^2 + 2 \frac{\partial^2 f}{\partial x_1 \partial x_2} (c(t)) y_1 y_2 + \frac{\partial^2 f}{\partial x_2^2} (c(t)) y_2^2.$$

Als-

$$h(0) = f(x_0)$$

$$h'(0) = \frac{\partial f}{\partial x_1} (x_0) y_1 + \frac{\partial f}{\partial x_2} (x_0) y_2$$

$$= \langle \nabla f(x_0), y \rangle$$

$$h''(0) = \frac{\partial^2 f}{\partial x_1^2} (x_0) y_1^2 + 2 \frac{\partial^2 f}{\partial x_1 \partial x_2} (x_0) y_1 y_2$$

$$+ \frac{\partial^2 f}{\partial x_2^2} (x_0) y_2^2$$

$$= {}^t y \operatorname{Hess}_f(x_0) y.$$

Also: folgt aus

$$h(1) = h(0) + \cancel{h'(0)} + \frac{h''(0)}{2!} + \frac{h^{(3)}(\xi)}{3!}$$

$$f(x_0 + y) = f(x_0) + \langle \nabla f(x_0), y \rangle + \frac{1}{2} y^T \text{Hess}_f(x_0) y + \frac{h^{(3)}(\xi)}{3!}$$

wobei $0 < \xi < 1$.

In dieser Formel erkennen wir
das Taylorpolynom:

$$T_2 f(y; x_0) = f(x_0) + \langle \nabla f(x_0), y \rangle + \frac{1}{2} y^T \text{Hess}_f(x_0) y$$

Mithals Binomialsatz kann man
allgemein

die Formel

$$T_k f(y; x_0) = \sum_{j=0}^k \frac{h^{(j)}(x_0)}{j!}$$

herleiten.

2.8. Kritische Punkte.

Sei $X \subset \mathbb{R}^n$ offen: wir erinnern an Definition 2.18, nämlich,

$$\forall x = (x_1, \dots, x_n) \in X \quad \exists \delta > 0$$

so dass der Hypercubus $C(x, \delta) :=$

$$]x_1 - \delta, x_1 + \delta[\times \dots \times]x_n - \delta, x_n + \delta[\subset X.$$

Sei also $X \subset \mathbb{R}^n$ und $f: X \rightarrow \mathbb{R}$.
offen

Def. 2.49

(a) $x_0 \in X$ ist lokale Maximale
und $f(x_0)$ lokales Maximum falls
 $\exists \delta > 0$ mit $C(x_0, \delta) \subset X$ und
 $f(y) \leq f(x_0) \quad \forall y \in C(x_0, \delta).$

(b) Analog mit lokaler Minimalestelle
wobei dann $f(y) \geq f(x_0) \quad \forall y \in C(x_0, \delta)$.

Falls eine der beiden Fälle zutrifft
nennen wir $x_0 \in X$ eine lokale
Extremalstelle.

Analog zum Fall $n=1$ gilt jetzt
allgemein:

Satz 2.50 ([K.] Prop. 3.8.1)

Sei $f : X \rightarrow \mathbb{R}$ differenzierbar
in x_0 und X offen
in $\mathbb{R}^n$.

Falls x_0 lokale Extremalstelle f. $f \in$
 $df(x_0) = 0 \quad (\Leftrightarrow \nabla f(x_0) = 0)$.

Beweis: Nun sei $\delta > 0$ mit $C(x_0, \delta) \subset X$
und OEDA:

$$f(y) \leq f(x_0) \quad \forall y \in C(x_0, \delta).$$

Insbesondere folgt für jede Koordinate j :

$$f(x_0 + te_j) \leq f(x_0) \quad \forall t \in]-\delta, \delta[$$

d.h. für $g_j(t) := f(x_0 + te_j)$

folgt: $g_j:]-\delta, \delta[\rightarrow \mathbb{R}$ ist

differenzierbar und $0 = |0|_{\mathbb{R}} = 0$

Maximalestelle. Es folgt aus Satz 4.15

(3) in Analysis I:

$$g_j'(0) = 0$$

d.h.
$$\frac{\partial f}{\partial x_j}(x_0) = 0$$

Woraus
$$\begin{aligned} J_f(x_0) &= \left(\frac{\partial f}{\partial x_1}(x_0), \dots, \frac{\partial f}{\partial x_n}(x_0) \right) \\ &= (0, \dots, 0) \end{aligned}$$

und somit $df(x_0) = 0$, da nach Voraussetzung f in X differenzierbar ist. $\square$

Def. 2.51. ([K0] Def. 3.8.2)

Sei $f: X \rightarrow \mathbb{R}$ differenzierbar.
offen
 $\mathbb{R}^n$

Ein kritischer Punkt ist ein $x_0 \in X$
für welches $df(x_0) = 0 \iff \underbrace{Df(x_0)}_{=0}$

Problem: finde Maxima und Minima
einer Funktion $f: X \rightarrow \mathbb{R}$.
 $X \subset \mathbb{R}^n$

Falls X kompakt (Def. 2.12 (3))
und f stetig existieren solche
immer (Satz 2.17).

$n=1$, $X=[a,b]$, $f:[a,b] \rightarrow \mathbb{R}$
stetig in $]a,b[$ differenzierbar:
Dann ist der Algorithmus:

- (1) Finde alle kritischen Punkte
in $]a,b[$
- (2) Vergleiche deren Werte mit f
mit $f(a)$ und $f(b)$.

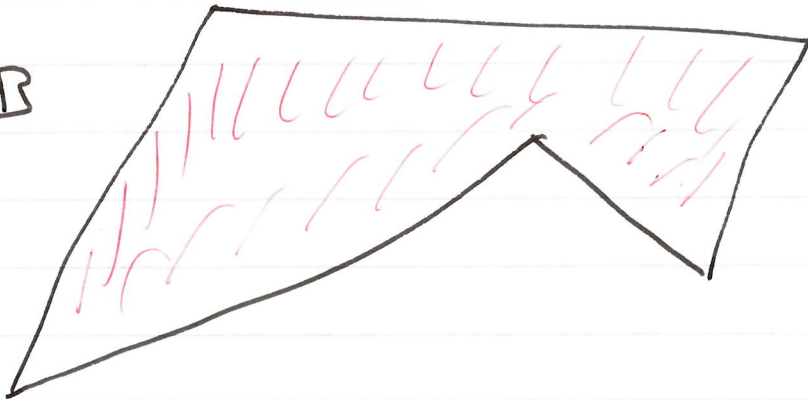
Äquivalent: Zerlege $[a, b] =]a, b[\cup \{a, b\}$

Wobei $]a, b[$ offen ist und $\{a, b\}$ der "Rand" ist von $[a, b]$.

Allgemein: falls $K \subset \mathbb{R}^n$

kompakte Teilmenge und $f: K \rightarrow \mathbb{R}$

$$K = X \cup B$$



stetig. Zerlege $K = X \cup B$

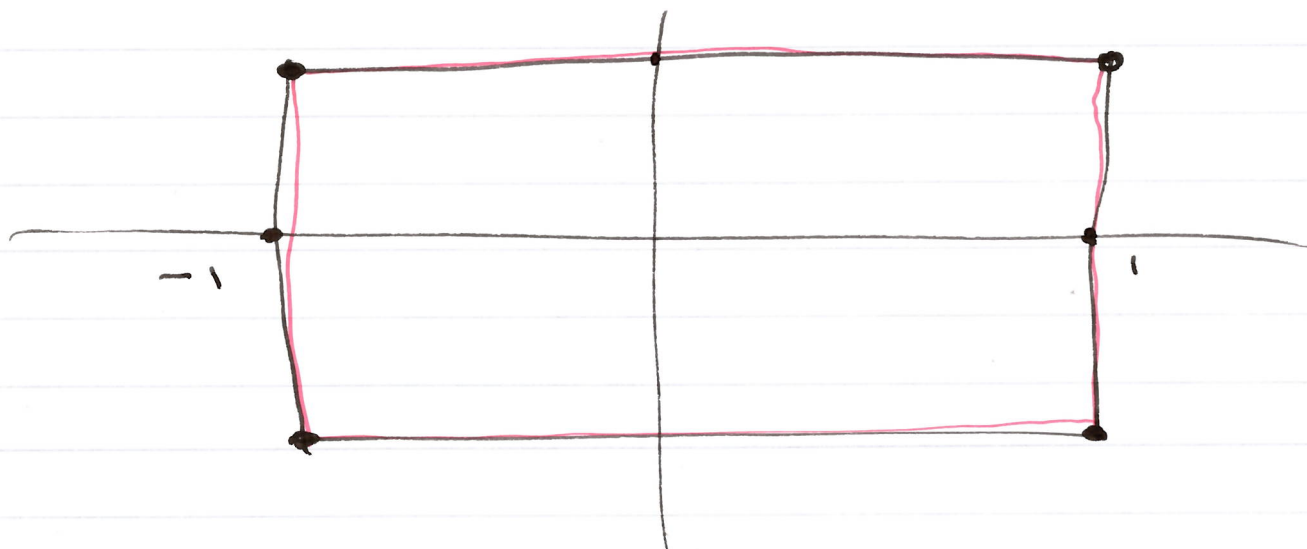
wobei X offen ist, das "Innere" von K und $B \subset K$ der "Rand" von K ist.

Algorithmus: $f \in \mathcal{B}$ gegeben f auf X
differenzierbar so:

(1) bestimme die kritischen Punkte
von f in X

(2) bestimme Maxima und Minima
von $f|_B$.

Beispiel: $f(x, y) = x^2 - 2y^2$ auf
 $K = [-1, 1]^2 =]-1, 1[\cup B$



$$\nabla f(x, y) = \begin{pmatrix} 2x \\ -4y \end{pmatrix}$$

Kritischer Punkt: $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$.

Jetzt ist

$$B = \{ (1, y) : -1 \leq y \leq 1 \} \cup$$

$$\cup \{ (x, 1) : -1 \leq x \leq 1 \}$$

$$\cup \{ (-1, y) : -1 \leq y \leq 1 \}$$

$$\cup \{ (x, -1) : -1 \leq x \leq 1 \}.$$

Die Maxima und Minima von $f|_B$

ist dann ein 1-D. Problem:

$$(1) \quad [-1, 1] \rightarrow \mathbb{R}$$

$$y \mapsto 1 - 2y^2$$

kritischer Punkt in $] -1, 1[$: $y = 0$.

$$(2) \quad [-1, 1] \rightarrow \mathbb{R}$$

$$x \mapsto x^2 - 2$$

kritischer Punkt in $] -1, 1[$: $x = 0$.

$$(3) \quad [-1, 1] \rightarrow (-1)^2 - 2y^2 = 1 - 2y^2$$

kritischer Punkt in $] -1, 1[$: $y = 0$.

$$(3) \quad [-1, 1] \rightarrow \mathbb{R}$$

$$x \mapsto x^2 - 2(-1)^2 = x^2 - 2$$

kritischer Punkt in $] -1, 1[$: $x = 0$.

Die möglichen Extrempunkten sind:

$$(0,0), (1,0), (0,1), (-1,0), (0,-1)$$

$$(1,1), (-1,1), (-1,-1), (1,-1).$$

Die entsprechenden Werte sind:

$$0, 1, -2, 1, -2, \\ -1, -1, -1, -1.$$

Also: • $(1,0)$ und $(-1,0)$ Maximal
stellen

• $(0,1)$ und $(0,-1)$ Minimal
stellen.

Als nächstes befassen wir uns mit dem Problem der Verallgemeinerung auf $\mathbb{R}^n$ von folgenden Kriterien:

Satz 2.52 (Siehe Analysis I)
Korollar 4.4c

Sei $f: [a, b] \rightarrow \mathbb{R}$ stetig und in $]a, b[$ zweimal differenzierbar.

Sei $x_0 \in]a, b[$ ein kritischer Punkt von f .

(1) Falls $f^{(2)}(x_0) > 0$ ist x_0 eine strikte lokale Minimalstelle.

(2) Falls $f^{(2)}(x_0) < 0$ ist x_0 eine strikte lokale Maximalstelle.

Im allgemeinen ist $\text{Hess}_f(x_0)$ eine symmetrische Matrix. Falls $\nabla f(x_0) = 0$ und $\det(\text{Hess}_f(x_0)) \neq 0$ nennen wir x_0 einen nicht degenerierten kritischen Punkt.

Aus Lin. Alg. wissen wir, dass falls A eine symmetrische Matrix mit $\det A \neq 0$, dann ist A orthogonal diagonalisierbar und besitzt

p Eigenwerte > 0
 q Eigenwerte < 0

mit $p + q = n$.

Wir bemerken:

A positiv definit d.h. $\langle y, Ay \rangle > 0 \quad \forall y \neq 0$
 $\Leftrightarrow q = 0$

- 96 -

A negativ definit d.h. $\forall y \quad Ay < 0$
 $\forall y \neq 0$

$$\Leftrightarrow p = 0.$$

Satz 2.52 hat dann folgende

Verallgemeinerung:

Satz 2.53 ([Ko] 3.8.7).

Sei $f: X \rightarrow \mathbb{R}$ eine auf
offen $\cap$
 $\mathbb{R}^n$

X C^2 -Funktion und $x_0 \in X$ ein
nicht degenerierter kritischer Punkt.

Seien p resp. q die Anzahl positi-
ver resp. negativer Eigenwerte von
 $Hess_f(x_0)$.

- 97 -

(1) Falls $p = n$ ist x_0 eine lokale
Minimalstelle.

(2) Falls $q = n$ ist x_0 eine lokale
Maximalstelle.

(3) Falls $p, q \neq n$ ist x_0 kein
lokales Extremum von f .

Lemma 2.54. Sei $A \in M_{n,n}(\mathbb{R})$.

(1) Falls A positiv definit gibt
es $c > 0$ mit $\forall y \quad Ay \geq c \|y\|^2$
 $\forall y \in \mathbb{R}^n$.

(2) Falls A negativ definit gibt es $c > 0$
mit $\forall y \quad Ay \leq -c \|y\|^2 \quad \forall y \in \mathbb{R}^n$.

Beweis: offensichtlich folgt (2) aus (1).

Zu (1), Sei $v_1, \dots, v_n$ eine Orthonormalbasis von $\mathbb{R}^n$ die aus Eigenvektoren von A besteht. Sei

$$A v_i = \lambda_i v_i.$$

OEDA können wir annehmen:

$$\lambda_1 \geq \dots \geq \lambda_n > 0.$$

Sei $y \in \mathbb{R}^n$ und $y = \sum_{i=1}^n y_i v_i$.

Dann folgt:

$$\langle y, Ay \rangle = \langle Ay, y \rangle$$

$$= \left\langle \sum_i \lambda_i y_i v_i, \sum_j y_j v_j \right\rangle$$

$$= \sum_{i=1}^n \lambda_i y_i^2 \geq \lambda_n \sum_{i=1}^n y_i^2$$

$$\text{Ab.} \quad {}^t y A y \geq \lambda_n \|y\|^2. \quad \square$$

Beweis Satz 2.53 :

(1) Aus Satz 2.48 folgt:

$$f(x_0 + y) = f(x_0) + {}^t y \operatorname{Hess}_f(x_0) y + E_2 f(x_0 + y, x_0)$$

Und aus Lemma 2.54 :

$$\begin{aligned} f(x_0 + y) &\geq f(x_0) + c \|y\|^2 + E_2 f \\ &= f(x_0) + \left(c + \frac{E_2 f(x_0 + y, x_0)}{\|y\|^2} \right) \|y\|^2. \end{aligned}$$

$$\text{Da} \quad \lim_{\substack{y \rightarrow 0 \\ y \neq 0}} \frac{E_2 f(x_0 + y, x_0)}{\|y\|^2} = 0$$

folgt dass es ein $\delta > 0$ gibt

so dass:

$$\frac{|E_2 f(x_0 + y, x_0)|}{\|y\|^2} \leq \frac{c}{2}$$

$$\forall y \in C(0, \delta)$$

Woraus folgt:

$$f(x_0 + y) \geq f(x_0) + \frac{c}{2} \|y\|^2 \quad \forall y \in C(0, \delta)$$

also ist x_0 eine lokale Minimalstelle.