Zum Nachweis müssen wir uns ein wenig mit den Eigenschaften des totalen Differentials beschäftigen.

Satz: Ist $f: U \to \mathbb{R}$, $U \subset \mathbb{R}^n$ offen, $\xi \in U$ in alle Richtungen stetig partiell differenzierbar so gilt

$$\frac{f(x) - f(\xi) - D_\xi f\,(x-\xi)}{|x-\xi|} \to 0$$

für $x \to \xi$.

⌈ Wir müssen zeigen, dass

$$\frac{f^i(x) - f^i(\xi) - D_\xi f^i\,(x-\xi)}{|x-\xi|} \to 0$$

also $D_\xi f^i\,(x-\xi) = (\nabla_\xi f^i)(x-\xi)$

$$= \sum_{j=1}^{n} \partial_j f^i(\xi)\cdot(x_j-\xi_j)$$

und für $\Theta_j = (x_1,\dots,x_j,\xi_{j+1},\dots,\xi_n)$ gilt

$$f^i(x) - f^i(\xi) = \left(f^i(\Theta_n) - f^i(\Theta_{n-1})\right) + \left(f^i(\Theta_{n-1}) - f^i(\Theta_{n-2})\right) + \dots + \left(f^i(\Theta_1) - f^i(\Theta_0)\right)$$

$$= \sum_{j=1}^{n} f^i(\Theta_j) - f^i(\Theta_{j-1})$$

"Teleskopsumme"

$$\rightsquigarrow \left| f^i(x) - f^i(\xi) - D_\xi f^i (x-\xi) \right|$$

$$= \left| \sum_{j=1}^{n} -\partial_j f^i(\xi) \cdot (x_j - \xi_j) + f^i(\Theta_j) - f^i(\Theta_{j-1}) \right|$$

Ich erinnere nun an den Zwischenwertsatz aus der Differentialrechnung in einer Variablen:

Satz: Ist $f : [a,b] \to \mathbb{R}$ stetig differenzierbar so existiert ein $t \in [a,b]$ mit

$$\frac{f(a) - f(b)}{a-b} = f'(t).$$

Das sagt uns

$$\leq \sum_{j=1}^{n} \max_{t \in [\Theta_j, \Theta_{j-1}]} \left| \partial_j f^i(\xi) - \partial_j f^i(t) \right| \cdot |x_j - \xi_j|$$

where the maximum is formed over all points of the line segment connecting Θ_j and Θ_{j-1}. Since $|x_j - \xi_j| \leq |x - \xi|$ we arrive at

$$\frac{\left| f^i(x) - f^i(\xi) - D_\xi f^i (x-\xi) \right|}{|x-\xi|} \leq \frac{n \cdot \cancel{|x-\xi|} \cdot \sum_{j=1}^{n} \max_{t \in [\Theta_j, \Theta_{j-1}]} \left| \partial_j f^i(\xi) - \partial_j f^i(t) \right|}{\cancel{|x-\xi|}}$$

und da $\Theta_j \to \xi$ bei $x \to \xi$ geht dieser Ausdruck gegen 0 aufgrund der Stetigkeit von ∂f. $\lrcorner$ (3)

Anders gesagt

$$|f(\xi)-f(x)-D_\xi f(\xi-x)| \leq r(x)\cdot|\xi-x|$$

mit $r(x)\to 0$ bei $x\to\xi$.

Der Zwischenwertsatz liefert auch direkt die

<u>Zwischenwertabschätzung</u>

$$|f(\xi)-f(x)| \leq \max_{t\in[x,\xi]}\left|\frac{\partial f}{\partial(x-\xi)}(t)\right|\cdot|\xi-x|$$

indem man ihn anwendet auf die Komponenten von

$$t\longrightarrow f(\xi+t(x-\xi))$$

<u>Beobachtung</u>

Die Eigenschaft, dass

$$|f(x)-f(\xi)-L(x-\xi)| \leq r(x)(x-\xi)$$

und $r(x)\to 0$ bei $x\to\xi$, bestimmt ein $L\in \mathrm{Mat}(n,m)$ eindeutig.

$\ulcorner$ Es folgt für $L, L' \in \mathrm{Mat}(n,m)$, dass

$$\frac{|L(x-\xi)-L'(x-\xi)|}{|x-\xi|} \leq r(x)+r'(x)$$

aber für $x=\xi+ty$ ist die linke Seite unabhängig von t und falls $L(y)\neq L'(y)$ geht sie dann bei $t\to 0$ nicht gegen 0, im Widerspruch zu obiger Ungleichung. $\lrcorner$

Diese Eindeutigkeit liefert

<u>Corollar</u> (Kettenregel)

$f: U \to \mathbb{R}^m$, $g: V \to \mathbb{R}^k$ mit $f(U) \subseteq V$

$$\Rightarrow \quad D_x(g \circ f) = D_{f(x)} g \cdot D_x f$$

Beobachte für $n = m = k = 1$ ist das genau die bekannte Kettenregel $(g \circ f)'(x) = g'(f(x)) \cdot f'(x)$

$$|g \circ f(x) - g \circ f(z) - D_{f(x)} g \cdot D_x f \cdot (x-z)|$$

$$\leq |g \circ f(x) - g \circ f(z) - D_{f(x)} g (f(x) - f(z))| + C' |r_f(z)| \cdot |x-z|$$

for $C \in \mathbb{R}$ with $|D_{f(x)} g(\ )| \leq C|y|$ for all $y \in \mathbb{R}^m$ ($C' = |$largest or smallest coefficient of $D_{f(x)} g|$ works).

$$\leq |r_g(f(z))| \cdot |f(x) - f(z)| + C' |r_f(z)| \cdot |x-z|$$

$$\leq |r_g(f(z))| \cdot C \cdot |x-z| + C' |r_f(z)| \cdot |x-z|$$

wegen der Zwischenwertabschätzung, also

$$= \left(C \cdot r_g(f(z)) + C' r_f(z) \right) \cdot |x-z| \to 0$$

Hieraus folgt nun direkt der erste der noch unerklärten Sätze: Nämlich, dass

$$D_x f \cdot v = \frac{\partial f}{\partial v}(x) \qquad \forall x \in U,\ v \in \mathbb{R}^n$$

⌜ Man wende die Kettenregel auf

$$f_{v,x} : \left]-\varepsilon/|v|, \varepsilon/|v|\right[\longrightarrow \mathbb{R}^m$$
$$t \longmapsto f(x+tv)$$

an: Man erhält mit $g(t) = x+tv$, $f_{x,v} = f \circ g$ und

$$\frac{\partial f}{\partial v}(x) := f'_{v,x}(0) = D_0 f_{v,x}(1) = D_x f \cdot D_0 g(1)$$
$$= D_x f \cdot v$$

da $D_0 g(\lambda) = \lambda v$ für alle $\lambda \in \mathbb{R}$. ⌟

Hier ist noch eine Anwendung: Gegeben $f, g : \mathbb{R} \to \mathbb{R}$ so gilt

$$D_t(f \cdot g) = D_t\left(\begin{array}{ccccccc} \mathbb{R} & \xrightarrow{\Delta} & \mathbb{R}\times\mathbb{R} & \xrightarrow{f\times g} & \mathbb{R}\times\mathbb{R} & \xrightarrow{\mu} & \mathbb{R} \\ x & \longmapsto & (x,x) & & (x,y) & \longmapsto & x\cdot y \end{array} \right)$$
$$= D_{(f(t),g(t))}\mu \cdot D_{(t,t)}(f\times g) \cdot D_t \Delta$$

Also

$$D_t \Delta = \begin{pmatrix} 1 \\ 1 \end{pmatrix},\ D_{(t,t)}(f \times g) = \begin{pmatrix} f'(t) & 0 \\ 0 & g'(t) \end{pmatrix} \quad D_{(a,b)}\mu = \begin{pmatrix} b & a \end{pmatrix}$$

$$\rightsquigarrow D_t(f \cdot g) = \begin{pmatrix} g(t) & f(t) \end{pmatrix} \cdot \begin{pmatrix} f'(t) & 0 \\ 0 & g'(t) \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$
$$= \left(g(t) \cdot f'(t) + f(t) \cdot g'(t) \right) \qquad \text{„Produktregel"}$$

Etwas Allgemeiner $D_t \langle f, g \rangle = \langle \nabla_t f, g \rangle + \langle f, \nabla_t g \rangle$ für $f, g : \mathbb{R} \to \mathbb{R}^n$.

Nun zurück zu Minima und Maxima: Ich erinnere an die Taylorentwicklung: Für $U \subseteq \mathbb{R}$ offen, $f: U \to \mathbb{R}$, beliebig oft differenzierbar, und $\xi \in U$ setze:

$$T_{k,\xi} f : U \longrightarrow \mathbb{R}$$

$$x \longmapsto f(\xi) + f'(\xi)(\xi - x) + \frac{1}{2} f''(\xi)(\xi - x)^2 + \ldots + \frac{1}{k!} f^{(k)}(\xi)(\xi - x)^k.$$

"k-te Taylor-Approximation"

und $R_{k,\xi} f : U \longrightarrow \mathbb{R}$ als $f - T_{k,\xi} f$.

"k-tes Restglied".

Dann ist $T_{k,\xi} f$ das eindeutig bestimmte Polynom P mit $P(\xi) = f(\xi)$, $P'(\xi) = f'(\xi), \ldots, P^{(k)}(\xi) = f^{(k)}(\xi)$.

Proposition Für alle $x, \xi \in U$ gibt es t zwischen x und ξ mit

$$(R_{k,\xi} f)(x) = \frac{1}{k!} f^{(k+1)}(t)(x-t)^k (x - \xi)$$

'Cauchy'sches Restglied'

⌜ Wende den Zwischenwertsatz auf

$$F(t) = f(t) + f'(t)(x-t) + \frac{1}{2} f''(t)(x-t)^2 + \ldots + \frac{1}{k!} f^{(k)}(t)(x-t)^k$$

an: Das liefert t zwischen x und ξ mit

$$\frac{F(\xi) - F(x)}{\xi - x} = F'(t) \quad \text{oder} \quad F(\xi) = F(x) + F'(t)(\xi - x)$$

Also $F(\xi) = T_{k,\xi} f(x)$ und

$$F'(t) = \frac{1}{k!} f^{(k+1)}(t)\,(x-t)^k \quad \lrcorner$$

Für $f : U \to \mathbb{R}$, $U \subseteq \mathbb{R}^n$ offen, $\xi \in U$ definieren wir nun

$$(T_{0,\xi} f)(u) = f(\xi)$$

$$(T_{1,\xi} f)(u) = f(\xi) + \langle \nabla_\xi f, u - \xi \rangle$$

$$(T_{2,\xi} f)(u) = f(\xi) + \langle \nabla_\xi f, u-\xi \rangle + \tfrac{1}{2} \langle u-\xi, H_\xi f \cdot (u-\xi) \rangle$$

(die höheren Taylor-Approximation bedürfen entweder langer Formeln, oder mehrdimensionaler Matrizen und wir brauchen sie nicht).

<u>Satz</u>: Zu jedem $\xi, u \in U$ mit $[\xi, z] \subseteq U$ gibt es ein $z \in [\xi, u]$, $z = \xi + t(u-\xi)$ mit $0 \leq t \leq 1$, so dass

$$f(u) = f(\xi) + \langle \nabla_\xi f, u-\xi \rangle + \tfrac{1}{2}(1-t) \langle u-\xi, H_z f \cdot (u-\xi) \rangle$$

$\ulcorner$ Man wende die Cauchy'sche Restgliedformel für $k = 1$ auf die Funktion

$$f_{\xi, u-\xi}(t) = f(\xi + t(u-\xi))$$

mit Entwicklungspunkt 0 an. Mit den Abkürzungen $g(t) = \xi + t(u-\xi)$, $\hat{f} = f_{\xi, u-\xi}$ rechnet man

(8)

$$\hat{f}(1) = \hat{f}(0) + \hat{f}'(0)\cdot(1-0) + \hat{f}''(t)\cdot(1-0)(1-t)$$

für ein $0 \leq t \leq 1$. Nun gilt

$$\hat{f}(1) = f(\xi + (u-\xi)) = f(u)$$

$$\hat{f}(0) = f(\xi)$$

$$\hat{f}'(0) = (D_0 \hat{f})(1) = (D_0(f\circ g))(1)$$

$$= D_{g(0)}f \cdot \underbrace{D_0 g(1)}_{u-\xi} = D_\xi f\cdot(u-\xi)$$

$$= \langle \nabla_\xi f, (u-\xi)\rangle$$

$$\hat{f}''(t) = D_t\left(s \longmapsto \hat{f}'(s)\right)(1)$$

$$= D_t\left(s \longmapsto (D_s\hat{f})(1)\right)(1)$$

$$= D_t\left(s \longmapsto D_s f\circ g(1)\right)(1)$$

$$= D_t\left(s \longmapsto D_{g(s)}f\, \underbrace{D_s g(1)}_{u-\xi}\right)(1)$$

$$= D_t\left(s \longmapsto \langle (\nabla f)(g(s)), u-\xi\rangle\right)(1)$$

$$= \langle D_t\left(s \longmapsto (\nabla f)(g(s))\right)(1), u-\xi\rangle$$

$$= \langle D_{g(t)}(\nabla f)\cdot D_t g(1), u-\xi\rangle$$

$$= \langle H_{g(t)} f \cdot (u-\xi), u-\xi\rangle \quad \checkmark$$

Gilt also $\nabla f(\xi) = 0$, etwa weil f in ξ ein Extremum hat, so folgt

$$f(u) = f(\xi) + \tfrac{1}{2}(1-t)\langle u-\xi, H_z \cdot (u-\xi)\rangle$$

für ein $z \in [\xi, u]$ mit $z = \xi + t(u-\xi)$.

Aber H_z ist nach dem Satz von Schwarz symmetrisch, also orthogonal diagonalisierbar

<u>Beobachtung</u>: Betrachte die Funktion
$q: \mathbb{R}^n \to \mathbb{R}$, $v \longmapsto \langle v, Av\rangle$
für $A \in \mathrm{Mat}(n,n)$ orthogonal diagonalisierbar.

i) $q(v) > 0 \quad \forall v \in V$
$\iff$ A hat nur positive EW's

ii) $q(v) < 0 \quad \forall v \in V$
$\iff$ A hat nur neg. EW's

iii) q nimmt sowohl positive als auch negative Werte an
$\iff$ A hat sowohl pos. als neg. EW's.

⌜ Klar für Diagonales A, aber $\langle v, Av\rangle = \langle W_B v, W_B \cdot A \cdot W_B^{-1} W_B v\rangle$
für jede ONB B. $= \langle v_B, A_{BB} v_B\rangle$ ⌟

Das Extremwertkriterium lautete

<u>Theorem:</u> Ist $f : U \to \mathbb{R}^n$, $U \subseteq \mathbb{R}^n$ offen, zweimal stetig in alle Richtungen diff'bar und $\xi \in \mathbb{R}^n$ mit $\nabla f(\xi) = 0$. Dann gilt

i) sind alle Eigenwerte von $H_\xi f$ positiv so hat f bei x ein lokales Minimum

ii) ——————— negativ
——————— Maximum

iii) hat $H_x f$ sowohl positive als auch negative Eigenwerte so ist x ein <u>Sattelpunkt</u> von f.

Wir beobachten uns noch, dass die Positivität/Negativität von $v \mapsto \langle v, H_\xi f \cdot v \rangle$ sich auf $v \longmapsto \langle v, H_z f\, v \rangle$ überträgt für z nah genug an ξ.

Auch wenn $\langle v, H_\xi f\, v \rangle \geq 0 \ \forall v \in \mathbb{R}^n$ (das passiert genau, wenn alle EW's von $H_\xi f \geq 0$ sind), muss das gleiche nicht für $\langle v, H_x f\, v \rangle$ gelten!

Um Extrema auf Randpunkten zu erkennen gibt es folgendes nützliches Kriterium:

<u>Satz</u> Ist $f : U \to \mathbb{R}$, $U \subset \mathbb{R}^n$, und $g : V \to \mathbb{R}^n$, $V \subset \mathbb{R}^m$ offen mit

i) $g(V) = U$,

ii) $\ker(D_x g) = 0 \quad \forall x \in V$

so stimmen die Extrema von f mit denen von $f \circ g$ überein, im Sinne, dass $g(v)$ Extremum von f ist genau dann wenn v Extremum von $f \circ g$ ist.

Ein g wie in diesem Satz nennt man eine <u>Parametrisierung</u> von U, und sagt zu Bedingung ii) g habe vollen Rang.
Die Zahl m ist wegen der Kettenregel und den Sätzen von Steinitz eindeutig bestimmt und heißt <u>Dimension von U</u>.

Ist $f \circ g$ hinreichend differenzierbar so kann man also f auf Extrema überprüfen ohne dass U offen ist.