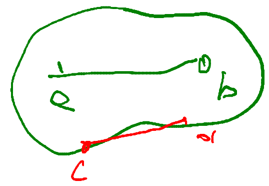
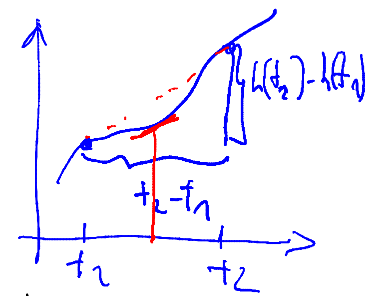
1.3 Mittelwertsätze und Taylor–Entwicklungen

Satz (Mittelwertsatz): Sei $f : D \to \mathbb{R}$ eine auf einer offenen Menge $D \subset \mathbb{R}^n$ differenzierbare, skalare Funktion. Weiterhin seien $\mathbf{a}, \mathbf{b} \in D$ Punkte in D, so dass die Verbindungsstrecke



$$[\mathbf{a}, \mathbf{b}] := \{\underbrace{\mathbf{a} + t(\mathbf{b} - \mathbf{a})} \mid t \in [0, 1]\}$$



ganz in D liegt. Dann gibt es eine Zahl $\theta \in (0, 1)$ mit

$$\boxed{f(\mathbf{b}) - f(\mathbf{a}) = \operatorname{grad} f(\mathbf{a} + \theta(\mathbf{b} - \mathbf{a})) \cdot (\mathbf{b} - \mathbf{a})}$$

Beweis: Wir setzen MWS 1d $h(t_2) - h(t_1) = h'(t_1+\theta(t_2-t_1))(t_2-t_1)$

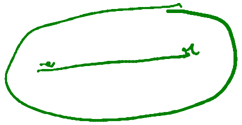
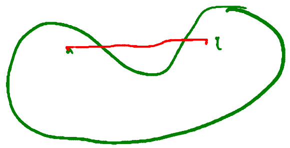
$$h(t) := f(\mathbf{a} + t(\mathbf{b} - \mathbf{a})) \qquad f(a) = h(0) \qquad f(b) = h(1)$$

Aus dem Mittelwertsatz für eine Veränderliche folgt dann mit der Kettenregel

$$\begin{aligned} f(\mathbf{b}) - f(\mathbf{a}) &= h(1) - h(0) \overset{\text{1d. MWS}}{=} h'(\theta) \cdot (1 - 0) \\ &= \operatorname{grad} f(\mathbf{a} + \theta(\mathbf{b} - \mathbf{a})) \cdot (\mathbf{b} - \mathbf{a}) \end{aligned}$$

Definition und Beispiel.

Definition: Gilt die Bedingung $[\mathbf{a}, \mathbf{b}] \subset D$ für **alle** Punkte $\mathbf{a}, \mathbf{b} \in D$, so heißt die Menge D konvex.

konvex   nicht konvex

Beispiel zum Mittelwertsatz: Gegeben sei die skalare Funktion

$$f(x, y) := \cos x + \sin y \qquad a = (0,0) \qquad b = \left(\frac{\pi}{2}, \frac{\pi}{2}\right)$$

Offensichtlich gilt

$$\underbrace{f(0,0)}_{f(a)} = \underbrace{f(\pi/2, \pi/2)}_{f(b)} = 1 \quad \Rightarrow \quad f(\pi/2, \pi/2) - f(0,0) = 0 \;= f(b) - f(a)$$

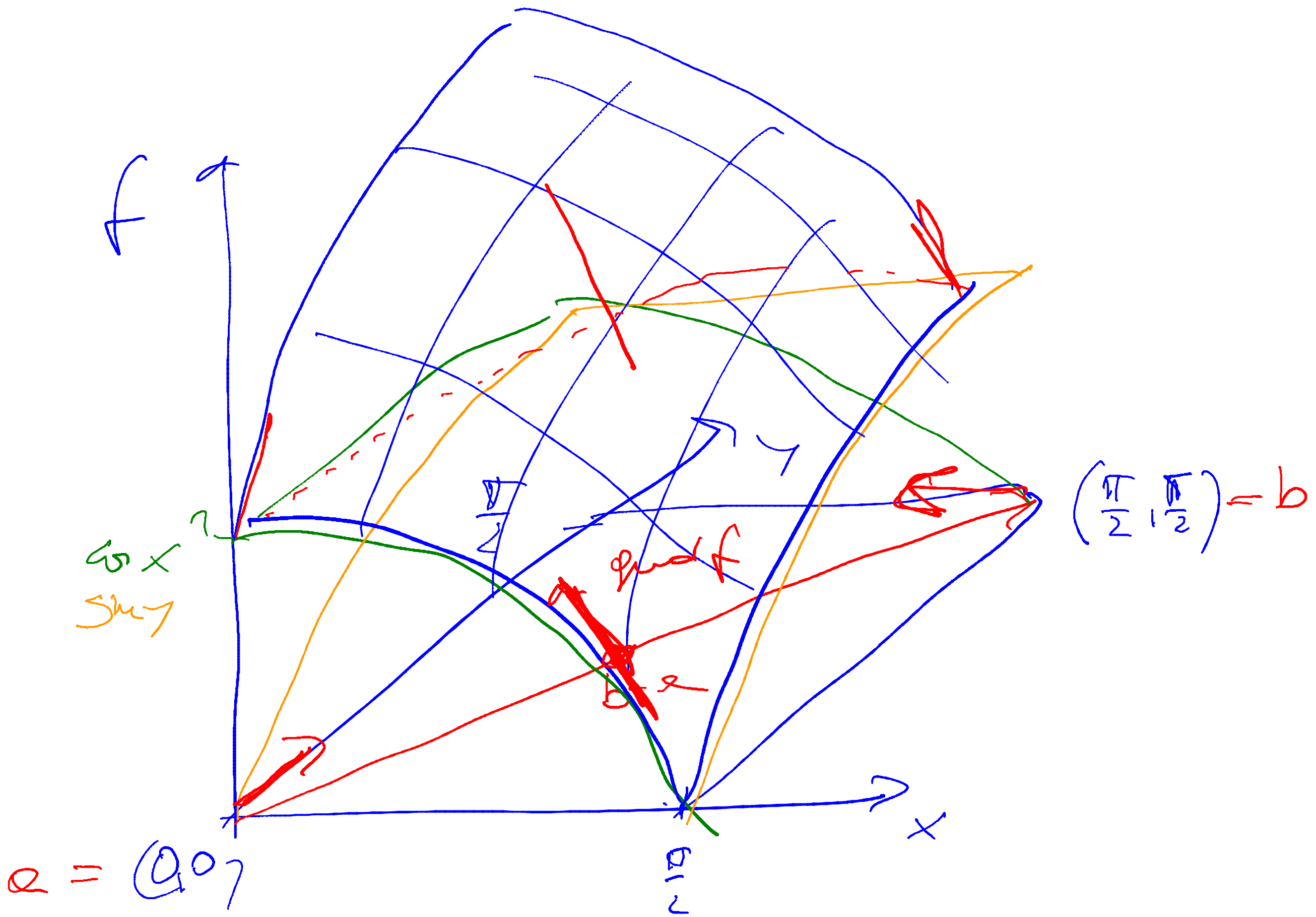
Nach dem Mittelwertsatz existiert ein $\theta \in (0, 1)$ mit $\quad \frac{\pi}{2}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$

$$\operatorname{grad} f\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix} + \theta \begin{pmatrix} \frac{\pi}{2} - 0 \\ \frac{\pi}{2} - 0 \end{pmatrix}\right) \cdot \begin{pmatrix} \frac{\pi}{2} - 0 \\ \frac{\pi}{2} - 0 \end{pmatrix} = \operatorname{grad} f\left(\theta \begin{pmatrix} \pi/2 \\ \pi/2 \end{pmatrix}\right) \cdot \begin{pmatrix} \pi/2 \\ \pi/2 \end{pmatrix} = 0$$

In der Tat gilt diese Beziehung für $\theta = \frac{1}{2}$.

$$\operatorname{grad} f = \begin{pmatrix} -\sin x \\ \cos y \end{pmatrix}; \quad \operatorname{grad} f\left(\frac{\pi}{4}, \frac{\pi}{4}\right) = \begin{pmatrix} -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{pmatrix} = \frac{\sqrt{2}}{2}\begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

$f(x,y) = \cos x + \sin y$



$\operatorname{grad} f = \begin{pmatrix} -\sin x \\ \cos y \end{pmatrix}$

$\operatorname{grad} f \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$

$\operatorname{grad} f \begin{pmatrix} \frac{\pi}{2} \\ \frac{\pi}{2} \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$

Mittelwertsatz gilt nur für skalare Funktionen.

Beachte: Der Mittelwertsatz für mehrere Variablen gilt nur für skalare Funktionen, aber i.A. nicht für vektorwertige Funktionen!

Beispiel: Betrachte die **vektorwertige** Funktion $f: D \to \mathbb{R}^2$

$$\mathbf{f}(t) := \begin{pmatrix} \cos t \\ \sin t \end{pmatrix}, \qquad t \in [0, \pi/2] = D \subseteq \mathbb{R}^1 \qquad f'(t) = \begin{pmatrix} -\sin t \\ \cos t \end{pmatrix}$$

Nun gilt

$$f(b) - f(a) = \mathbf{f}(\underbrace{\pi/2}_{b}) - \mathbf{f}(\underbrace{0}_{a}) = \begin{pmatrix} 0 \\ 1 \end{pmatrix} - \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix} \qquad \left\| \begin{pmatrix} -1 \\ 1 \end{pmatrix} \right\|_2 = \sqrt{2}$$

? $\neq$

und

$$f'(a + \theta(b-a)) \cdot (b-a) = \mathbf{f}'\left(\theta \frac{\pi}{2}\right) \cdot \left(\frac{\pi}{2} - 0\right) = \frac{\pi}{2} \begin{pmatrix} -\sin(\theta\pi/2) \\ \cos(\theta\pi/2) \end{pmatrix} \qquad \left\| \frac{\pi}{2} \begin{pmatrix} -\sin(\cdot) \\ \cos(\cdot) \end{pmatrix} \right\|_2 = \frac{\pi}{2}$$

unabhängig von θ

ABER: Die Vektoren auf der rechten Seite haben die Längen $\sqrt{2}$ bzw. $\pi/2$!

$$1{,}41\ldots = \sqrt{2} = \|f(b) - f(a)\|_2 \leq \Big\| \underbrace{f'\left(\theta \tfrac{\pi}{2}\right)}_{Jf\ 2\times 1} \cdot \underbrace{\tfrac{\pi}{2}}_{b-a} \Big\|_2 = \frac{\pi}{2} = 1{,}5\ldots$$

Der Mittelwert–Abschätzungssatz für vektorwertige Funktionen.

vektorwertig

Satz: Die Funktion $\mathbf{f} : D \to \mathbb{R}^m$ sei differenzierbar auf der offenen Menge $D \subset \mathbb{R}^n$. Weiterhin seien $\mathbf{a}, \mathbf{b}$ Punkte in D mit $[\mathbf{a}, \mathbf{b}] \subset D$. Dann gibt es ein $\theta \in (0, 1)$ mit

$$\|\mathbf{f}(\mathbf{b}) - \mathbf{f}(\mathbf{a})\|_2 \leq \|\underbrace{\mathbf{J}\,\mathbf{f}(\mathbf{a} + \theta(\mathbf{b} - \mathbf{a}))}_{m \times n} \cdot \underbrace{(\mathbf{b} - \mathbf{a})}_{n\ \text{Vektor}}\|_2$$

Beweisidee: Anwendung des Mittelwertsatzes auf die skalare Funktion $g(\mathbf{x})$ definiert durch

$g : D \to \mathbb{R}$

$$\boxed{g(\mathbf{x}) := (\mathbf{f}(\mathbf{b}) - \mathbf{f}(\mathbf{a}))^T \mathbf{f}(\mathbf{x})} \qquad \text{(Skalarprodukt!)}$$

Bemerkung: Eine andere (abgeschwächte) Form der Mittelwert–Abschätzung ist

$$\|\mathbf{f}(\mathbf{b}) - \mathbf{f}(\mathbf{a})\| \leq \sup_{\xi \in [\mathbf{a}, \mathbf{b}]} \|\mathbf{J}\,\mathbf{f}(\xi))\| \cdot \|(\mathbf{b} - \mathbf{a})\|$$

wobei $\|\cdot\|$ eine beliebige Vektor– bzw. zugehörige Matrixnorm ist.

$$g(x) = (f(b)-f(a))^T f(x)$$

$$g(b)-g(a) = (f(b)-f(a))^T(f(b)-f(a)) = \|f(b)-f(a)\|_2^2$$

((MWS skalare Fkt

$$\operatorname{grad} g(a+\theta(b-a)) \cdot (b-a) = \operatorname{grad}\left(\sum_{j=1}^{n}(f_j(b)-f_j(a))f_j(a+\theta(b-a))\right)(b-a)$$

$$= \sum_{i=1}^{n} \frac{\partial}{\partial x_i}\left(\qquad \right)\cdot(b_i-a_i)$$

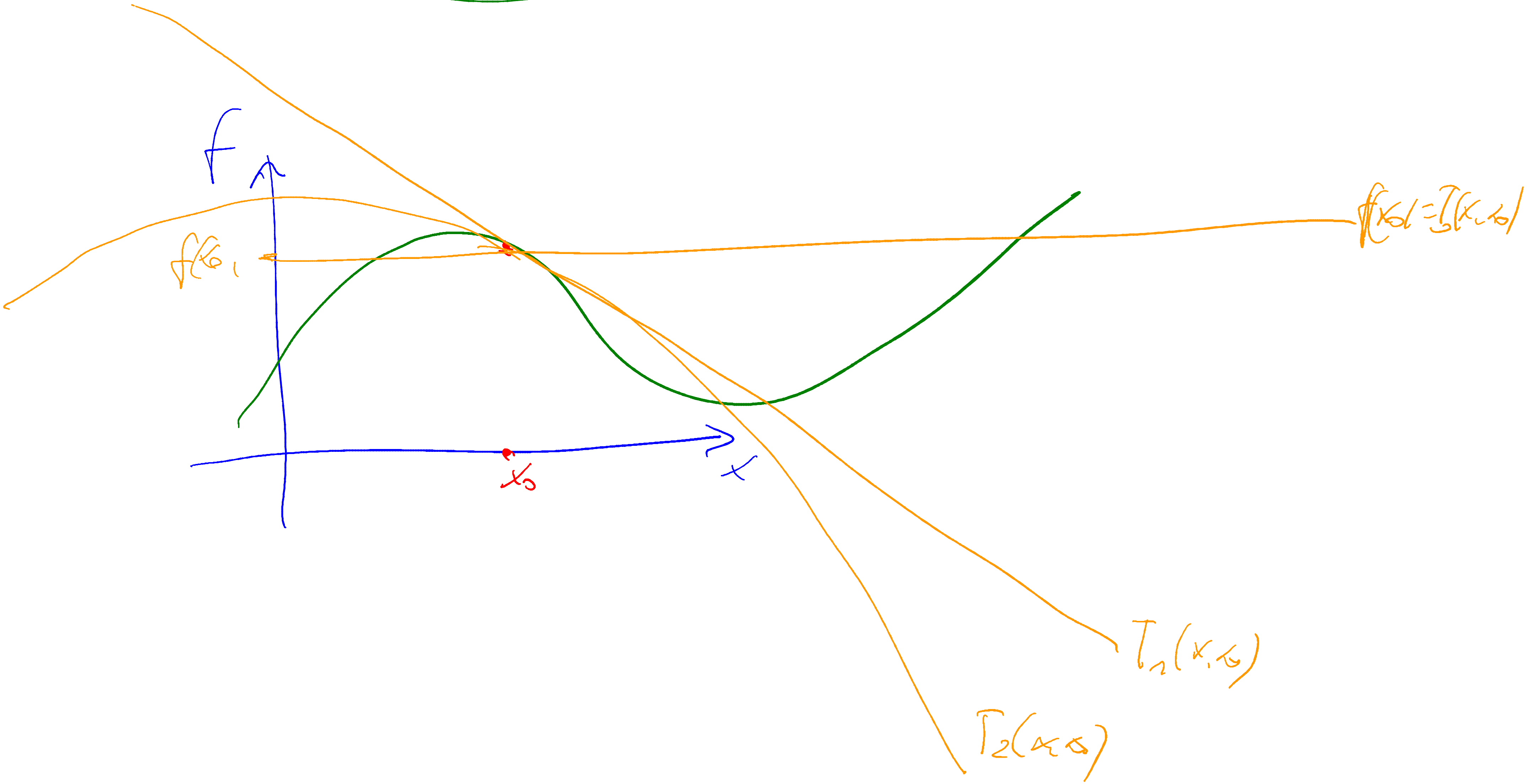
$$= \sum_{i=1}^{n}\sum_{j=1}^{n}(f_j(b)-f_j(a))\underbrace{\frac{\partial}{\partial x_i} f_j(a+\theta(b-a))}_{(Jf)_{ji}}(b_i-a_i) =$$

$$= \underbrace{(f(b)-f(a))^T}_{m}\underbrace{(Jf)(a+\theta(b-a))}_{m\times n}\cdot\underbrace{(b-a)}_{n}$$

$$\|f(b)-f(a)\|_2^{\cancel{2}} \leq \|(f(b)-f(a))^T Jf(\qquad)\cdot(b-a)\|_2 \leq \cancel{\|f(b)-f(a)\|_2}\,\|Jf(\quad)\cdot(b-a)\|_2$$

$n=1$

$$f(x) = \underbrace{\underbrace{\underbrace{f(x_0)}_{T_0(x,x_0)} + f'(x_0)(x-x_0)}_{T_1(x,x_0)} + \frac{1}{2} f''(x_0)(x-x_0)^2}_{T_2(x,x_0)} + \dots$$



Taylor–Entwicklungen: Notationen.

Zunächst definieren wir einen Multiindex $\alpha \in \mathbb{N}_0^n$ als

$$\alpha := (\alpha_1, \dots, \alpha_n) \in \mathbb{N}_0^n$$

Weiterhin sei

$$|\alpha| := \alpha_1 + \cdots + \alpha_n \qquad \alpha! := \alpha_1! \cdot \cdots \cdot \alpha_n!$$

Ist $f : D \to \mathbb{R}$ $\;|\alpha|$–mal stetig differenzierbar, so setzen wir

$$D^\alpha f = D_1^{\alpha_1} D_2^{\alpha_2} \dots D_n^{\alpha_n} f = \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \dots \partial x_n^{\alpha_n}},$$

wobei $D_i^{\alpha_i} = \underbrace{D_i \dots D_i}_{\alpha_i\text{–mal}}$ und wir schreiben

$$\mathbf{x}^\alpha := x_1^{\alpha_1} x_2^{\alpha_2} \dots x_n^{\alpha_n} \qquad \text{für } \mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n.$$

$n=2$

$|\alpha| = 0 \qquad (\alpha_1, \alpha_2) = (0,0) \qquad D^{(0,0)} f = f$

$|\alpha| = 1 \qquad (\alpha_1, \alpha_2) = \begin{cases} (1,0) \\ (0,1) \end{cases} \qquad D^{(1,0)} f = \frac{\partial}{\partial x_1} f \qquad D^{(0,1)} f = \frac{\partial}{\partial x_2} f$

$|\alpha| = 2 \qquad (\alpha_1, \alpha_2) = \begin{cases} (2,0) \\ (1,1) \\ (0,2) \end{cases} \qquad D^{(1,1)} f = \frac{\partial^2}{\partial x_1 \partial x_2} f \qquad D^{(0,2)} f = \frac{\partial^2}{\partial x_2^2} f$

Satz: (Satz von Taylor)

skalar

Sei $D \subset \mathbb{R}^n$ offen und konvex, $f : D \to \mathbb{R}$ eine $\mathcal{C}^{m+1}$–Funktion und sei $\mathbf{x}_0 \in D$. Dann gilt für $\mathbf{x} \in D$ die Taylor–Entwicklung

$$\boxed{f(\mathbf{x}) = T_m(\mathbf{x};\mathbf{x}_0) + R_m(\mathbf{x};\mathbf{x}_0)}$$

$$T_m(\mathbf{x};\mathbf{x}_0) = \sum_{|\alpha|\leq m} \frac{D^\alpha f(\mathbf{x}_0)}{\alpha!}(\mathbf{x}-\mathbf{x}_0)^\alpha$$

$$R_m(\mathbf{x};\mathbf{x}_0) = \sum_{|\alpha|=m+1} \frac{D^\alpha f(\mathbf{x}_0+\theta(\mathbf{x}-\mathbf{x}_0))}{\alpha!}(\mathbf{x}-\mathbf{x}_0)^\alpha$$

mit einem geeigneten $\theta \in (0,1)$.

Bezeichnung: In der obigen Taylor–Entwicklung heißt $T_m(\mathbf{x};\mathbf{x}_0)$ Taylor–Polynom m–ten Grades und $R_m(\mathbf{x};\mathbf{x}_0)$ wird als Lagrange–Restglied bezeichnet.

Wir definieren eine skalare Funktion einer Variablen $t \in [0,1]$ als

$$g(t) := f(\mathbf{x}_0 + t(\mathbf{x} - \mathbf{x}_0))$$

$g(1) = f(x)$ $g(0) = f(x_0)$

und berechnen die Taylor–Entwicklung um $t = 0$. Es gilt:

$T_1(x_0)$

$T_0(x_0)$

$$f(x) = g(1) = \underbrace{g(0)}_{f(x_0)} + g'(0) \cdot (1-0) + \frac{1}{2} g''(\xi) \cdot (1-0)^2 \quad \text{für ein } \xi \in (0,1).$$

Die Berechnung von $g'(0)$ liefert mit der Kettenregel

$$\begin{aligned} g'(0) &= \frac{d}{dt} f(x_1^0 + t(x_1 - x_1^0), x_2^0 + t(x_2 - x_2^0), \ldots, x_n^0 + t(x_n - x_n^0))\Big|_{t=0} \\ &= D_1 f(\mathbf{x}_0) \cdot (x_1 - x_1^0) + \ldots + D_n f(\mathbf{x}_0) \cdot (x_n - x_n^0) = \operatorname{grad} f(x_0) \cdot (x - x_0) \\ &= \sum_{|\alpha|=1} \frac{D^\alpha f(\mathbf{x}_0)}{\alpha!} \cdot (\mathbf{x} - \mathbf{x}_0)^\alpha = \sum_{i=1}^{n} \frac{\partial}{\partial x_i} f(x_0)(x_i - x_{0i}) \end{aligned}$$

Fortsetzung der Herleitung.

Berechnung von $g''(0)$ liefert

$$
\begin{aligned}
g''(0) &= \frac{d^2}{dt^2} f(\mathbf{x}_0 + t(\mathbf{x} - \mathbf{x}_0))\Big|_{t=0} = \frac{d}{dt} \sum_{k=1}^{n} D_k f(\mathbf{x}^0 + t(\mathbf{x} - \mathbf{x}^0))(x_k - x_k^0)\Big|_{t=0} \\
&= D_{11} f(\mathbf{x}_0)(x_1 - x_1^0)^2 + D_{21} f(\mathbf{x}_0)(x_1 - x_1^0)(x_2 - x_2^0) \\
&\quad + \ldots + D_{ij} f(\mathbf{x}_0)(x_i - x_i^0)(x_j - x_j^0) + \ldots + \\
&\quad + D_{n-1,n} f(\mathbf{x}_0)(x_{n-1} - x_{n-1}^0)(x_n - x_n^0) + D_{nn} f(\mathbf{x}_0)(x_n - x_n^0)^2) \\
&= \sum_{|\alpha|=2} \frac{D^\alpha f(\mathbf{x}_0)}{\alpha!}(\mathbf{x} - \mathbf{x}_0)^\alpha \qquad \text{(Vertauschungssatz von Schwarz!)}
\end{aligned}
$$

Nun: Beweis der Taylor–Formel mittels vollständiger Induktion!

$n=2$ (stolen)

$$f(x) = f(x_0) + \frac{D^{(1,0)}}{(1,0)!} f(x_0)(x-x_0)^{(1,0)} + \frac{D^{(0,1)} f(x_0)}{(0,1)!}(x-x_0)^{(0,1)}$$

$$\frac{1}{(2,0)!} D^{(2,0)} f(x_0)(x-x_0)^{(2,0)} + \frac{1}{(1,1)!} D^{(1,1)} f(x_0)(x-x_0)^{(1,1)} + \frac{1}{(0,2)!} D^{(0,2)} f(x_0)(x-x_0)^{(0,2)} + \dots$$

$$= f(x_0) \underbrace{+ \frac{1}{1}\frac{\partial}{\partial x_1} f(x_0)(x-x_0)_1 + \frac{1}{1}\frac{\partial}{\partial x_2} f(x_0)(x_2-x_{02})}$$

$$\underbrace{+ \frac{1}{2}\frac{\partial^2}{\partial x_1^2} f(x_0)(x_1-x_{01})^2 + \frac{1}{1}\frac{\partial^2}{\partial x_1 \partial x_2} f(x_0)(x_1-x_{01})(x_2-x_{02}) + \frac{1}{2!}\frac{\partial^2}{\partial x_2^2} f(x_0)(x_2-x_{02})^2} + \dots$$

$$= f(x_0) + \operatorname{grad} f(x_0) \cdot (x-x_0)$$

$$\frac{1}{2}(x-x_0)^T \underbrace{\begin{pmatrix} \frac{\partial^2}{\partial x_1^2} f & \frac{\partial^2}{\partial x_2 \partial x_1} f \\ \frac{\partial^2}{\partial x_1 \partial x_2} f & \frac{\partial^2}{\partial x_2^2} \end{pmatrix}}_{\text{Hesse Matrix: } H(x)} (x_0)(x-x_0)$$

(2×2 weil $n=2$)

Sym. falls $f \in C^2$

Beweis des Satzes von Taylor.

Die Funktion

$$g(t) := f(\mathbf{x}^0 + t(\mathbf{x} - \mathbf{x}^0))$$

ist $(m+1)$–mal stetig differenzierbar, und es gilt

$$g(1) = \sum_{k=0}^{m} \frac{g^{(k)}(0)}{k!} + \frac{g^{(m+1)}(\theta)}{(m+1)!} \quad \text{für ein } \theta \in [0,1].$$

Weiterhin gilt (per Induktion über k)

$$\frac{g^{(k)}(0)}{k!} = \sum_{|\alpha|=k} \frac{D^\alpha f(\mathbf{x}^0)}{\alpha!} (\mathbf{x} - \mathbf{x}^0)^\alpha$$

und

$$\frac{g^{(m+1)}(\theta)}{(m+1)!} = \sum_{|\alpha|=m+1} \frac{D^\alpha f(\mathbf{x}^0 + \theta(\mathbf{x} - \mathbf{x}^0))}{\alpha!} (\mathbf{x} - \mathbf{x}^0)^\alpha$$

Beispiel zur Taylor–Entwicklung.

1. Berechne das Taylor–Polynom $T_2(\mathbf{x};\mathbf{x}_0)$ zweiten Grades der Funktion

$$f(x,y,z) = x\,y^2\,\sin z$$

   zum Entwicklungspunkt $(x,y,z) = (1,2,0)^T$.

2. Die Berechnung von $T_2(\mathbf{x};\mathbf{x}_0)$ benötigt die partiellen Ableitungen bis zur zweiten Ordnung.

3. Diese Ableitungen müssen am Punkt $(x,y,z) = (1,2,0)^T$ ausgewertet werden.

4. Als Ergebnis erhält man $T_2(\mathbf{x};\mathbf{x}_0)$ in der Form

$$T_2(\mathbf{x};\mathbf{x}_0) = 4z(x+y-2)$$

5. Berechnung auf Folie.

$$f(x,y,z) = xy^2 \sin z \qquad (x_0,y_0,z_0) = (1,2,0)$$

$$f(x_0,y_0,z_0) = 0$$

$$\operatorname{grad} f = \left(y^2 \sin z,\ 2xy \sin z,\ xy^2 \cos z\right)$$

$$\operatorname{grad} f(x_0,y_0,z_0) = (0,0,4)$$

$$Hf = \begin{pmatrix} 0 & 2y \sin z & y^2 \cos z \\ & 2x \sin z & 2xy \cos z \\ & & -xy^2 \sin z \end{pmatrix} \qquad Hf(x_0,y_0,z_0) = \begin{pmatrix} 0 & 0 & 4 \\ 0 & 0 & 4 \\ 4 & 4 & 0 \end{pmatrix}$$

$$f(x,y,z) = 0 + (0,0,4)\begin{pmatrix} x-1 \\ y-2 \\ z \end{pmatrix} + \frac{1}{2}(x-1,y-2,z)\begin{pmatrix} 0 & 0 & 4 \\ 0 & 0 & 4 \\ 4 & 4 & 0 \end{pmatrix}\begin{pmatrix} x-1 \\ y-2 \\ z \end{pmatrix} =$$

$$= 4z + \frac{1}{2}(x-1,y-2,z)\begin{pmatrix} 4z \\ 4z \\ 4(x-1)+4(y-2) \end{pmatrix} = 4z + \frac{1}{2}\left(8(x-1) + 8(y-2)\right)z =$$

$$= 4z + 4(\cancel{x}-1)z + 4(y-2)z =$$

$$= \underline{4(\cancel{x}+y-2)z}$$

$$f(x,y,z) = xy^2 \sin z \underset{(1,2,0)}{=} (1+(x-1))(2+(y-2))^2 \sin(z) =$$

$$= (1+x-1)(4+4(y-2)+\cancel{(y-2)^2})(z - \cancel{\frac{z^3}{6}} + \dots)$$

$$= 4z + 4(x-1)z + 4(y-2)z = \dots$$

$$\sin z = z + $$

x Taylor in $x_0 = 1$

y^2 Taylor in $y_0 = 2$

$$x = 1 + (x-1)$$

$$y^2 = 4 + 4(y-2) + \dots$$

Bemerkung zum Restglied eines Taylor–Polynoms.

Bemerkung: Das Restglied eine Taylor–Polynoms enthält **alle** partiellen Ableitungen der Ordnung $(m+1)$:

$$R_m(\mathbf{x};\mathbf{x}_0) = \sum_{|\alpha|=m+1} \frac{D^\alpha f(\mathbf{x}_0 + \theta(\mathbf{x}-\mathbf{x}_0))}{\alpha!}(\mathbf{x}-\mathbf{x}_0)^\alpha$$

Sind all diese Ableitungen in der Nähe von $\mathbf{x}_0$ durch eine Konstante C beschränkt, so gilt die Restgliedabschätzung

Anzahl der Terme

$$|R_m(\mathbf{x};\mathbf{x}_0)| \le \frac{n^{m+1}}{(m+1)!}\, C\, \|\mathbf{x}-\mathbf{x}_0\|_\infty^{m+1}$$

Für die Approximationsgüte des Taylor–Polynoms einer $\mathcal{C}^{m+1}$–Funktion folgt daher

$$f(\mathbf{x}) = T_m(\mathbf{x};\mathbf{x}_0) + O\left(\|\mathbf{x}-\mathbf{x}_0\|^{m+1}\right)$$

Spezialfall $m=1$: Für eine $\mathcal{C}^2$–Funktion $f(\mathbf{x})$ bekommt man

$$f(\mathbf{x}) = f(\mathbf{x}^0) + \operatorname{grad} f(\mathbf{x}^0)\cdot(\mathbf{x}-\mathbf{x}^0) + O(\|\mathbf{x}-\mathbf{x}^0\|^2).$$

Die Hesse–Matrix.

Man nennt die Matrix

$$\mathbf{H}f(\mathbf{x}_0) := \begin{pmatrix} f_{x_1x_1}(\mathbf{x}_0) & \dots & f_{x_1x_n}(\mathbf{x}_0) \\ \vdots & & \vdots \\ f_{x_nx_1}(\mathbf{x}_0) & \dots & f_{x_nx_n}(\mathbf{x}_0) \end{pmatrix}$$

die Hesse–Matrix von $f(\mathbf{x})$ im Punkt $\mathbf{x}_0$.

$$\text{Hesse–Matrix} = \text{Jacobi–Matrix des Gradienten } \nabla f$$

Die Taylor–Entwicklung einer $\mathcal{C}^3$–Funktion lautet daher

$$f(\mathbf{x}) = f(\mathbf{x}_0) + \operatorname{grad} f(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) + \frac{1}{2}(\mathbf{x} - \mathbf{x}_0)^T \mathbf{H}f(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) + O(\|\mathbf{x} - \mathbf{x}_0\|^3)$$

Die Hesse–Matrix einer $\mathcal{C}^2$–Funktion ist symmetrisch.