1.3 Mittelwertsätze und Taylor–Entwicklungen

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

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

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

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

Beweis: Wir setzen

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

Aus dem MWS für eine Veränderliche folgt dann mit der Kettenregel

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

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

Beispiel: (zum Mittelwertsatz)

Gegeben sei die skalare Funktion

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

Offensichtlich gilt

$$f(0, 0) = f(\pi/2, \pi/2) = 1 \quad \Rightarrow \quad f(\pi/2, \pi/2) - f(0, 0) = 0$$

Nach dem Mittelwertsatz existiert ein $\theta \in (0, 1)$ mit

$$\text{grad}\, f\left(\theta \begin{pmatrix} \pi/2 \\ \pi/2 \end{pmatrix}\right) \cdot \begin{pmatrix} \pi/2 \\ \pi/2 \end{pmatrix} = 0$$

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

ACHTUNG:

Der Mittelwertsatz für mehrere Variablen gilt nur für **skalare** Funktionen!!!

Beispiel: Betrachte die **vektorwertige** Funktion

$$\mathbf{f}(t) := \begin{pmatrix} \cos t \\ \sin t \end{pmatrix}, \qquad t \in [0, \pi/2]$$

Nun gilt

$$\mathbf{f}(\pi/2) - \mathbf{f}(0) = \begin{pmatrix} 0 \\ 1 \end{pmatrix} - \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

und

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

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

Satz: **(Mittelwert–Abschätzungssatz)**

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

$$\|\mathbf{f}(\mathbf{b}) - \mathbf{f}(\mathbf{a})\|_2 \leq \|\mathbf{J}\,\mathbf{f}(\mathbf{a} + \theta(\mathbf{b} - \mathbf{a})) \cdot (\mathbf{b} - \mathbf{a})\|_2$$

Beweis:

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

$$g(\mathbf{x}) := (\mathbf{f}(\mathbf{b}) - \mathbf{f}(\mathbf{a}))^T \mathbf{f}(\mathbf{x})$$

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.

Taylor–Entwicklungen für skalare Funktionen mehrerer Variablen

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

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

Weiter sei

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

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

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

wobei $D_i^{\alpha_i} = \underbrace{D_i \dots D_i}_{\alpha_i\text{–mal}}$ und schreiben für $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$

$$\mathbf{x}^\alpha := x_1^{\alpha_1} x_2^{\alpha_2} \dots x_n^{\alpha_n}$$

Satz von Taylor:

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 alle $\mathbf{x} \in D$ die folgenden Entwicklung nach Taylor:

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

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

Beachte: Summation über $|\alpha| \le m$ und $|\alpha| = m + 1$.

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

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

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

$$g(1) = g(0) + g'(0) \cdot (1-0) + \frac{1}{2} g''(\xi) \cdot (1-0)^2, \qquad \xi \in (0,1)$$

Berechnung von $g'(0)$:

$$\begin{aligned} g'(0) &= \frac{d}{dt} f(x_1^0 + t(x_1 - x_1^0), x_2^0 + t(x_2 - x_2^0), \ldots, x_n^0 + t(x_n - x_n^0))|_{t=0} \\ &= D_1 f(\mathbf{x}_0)(x_1 - x_1^0) + \ldots + D_n f(\mathbf{x}_0)(x_n - x_n^0) \\ &= \sum_{|\alpha|=1} \frac{D^\alpha f(\mathbf{x}_0)}{\alpha!} (\mathbf{x} - \mathbf{x}_0)^\alpha \end{aligned}$$

Taylor–Entwicklung um $t = 0$

$$g(1) = g(0) + g'(0) + \frac{1}{2} g''(0) + \frac{1}{6} g^{(3)}(\xi), \qquad \xi \in (0,1)$$

Berechnung von $g''(0)$:

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

Beweis der Taylor–Formel mittels vollständiger Induktion!

Beispiel:

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

Zur Berechnung von $T_2(\mathbf{x};\mathbf{x}_0)$ benötigen wir die partiellen Ableitungen bis zur zweiten Ordnung.

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

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

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

Berechnung auf Folie

Beispiel:

Man berechne das Taylor–Polynom $T_3(\mathbf{x};\mathbf{x}_0)$ dritten Grades der Funktion

$$f(x,y) = e^y\ \cos x$$

zum Entwicklungpunkt $\mathbf{x}_0 = (0,0)^T$:

1) unter Verwendung des Taylorschen Satzes,
2) unter Verwendung bekannter Reihenentwicklungen.

Berechnung auf Folie

Bemerkung:

1) 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 $\mathbf{x}_0$ beschränkt durch C, so gilt die **Restgliedabschätzung**

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

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

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

Bemerkung: (Fortsetzung)

2) 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 $\mathbf{x}_0$.

Hesse–Matrix = Jacobi–Matrix des Gradienten ∇f

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

$$\begin{aligned} f(\mathbf{x}) &= f(\mathbf{x}_0) + \text{grad}\, f(\mathbf{x}_0)(\mathbf{x}-\mathbf{x}_0) + \frac{1}{2}(\mathbf{x}-\mathbf{x}_0)^T \mathbf{H}f(\mathbf{x}_0)(\mathbf{x}-\mathbf{x}_0) \\ &\quad + O(\|\mathbf{x}-\mathbf{x}_0\|^3) \end{aligned}$$

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

Kapitel 2: Anwendung der Differentialrechnung mehrerer Variablen

2.1 Extrema von Funktionen mehrerer Variablen

Definition: Sei $f : D \to \mathbb{R}$, $D \subset \mathbb{R}^n$, $\mathbf{x}^0 \in D$.

1) $f(\mathbf{x})$ hat in $\mathbf{x}^0$ ein **globales Maximum**, falls gilt:

$$\forall\, \mathbf{x} \in D \;:\quad f(\mathbf{x}) \le f(\mathbf{x}^0)$$

2) $f(\mathbf{x})$ hat in $\mathbf{x}^0$ ein **strenges globales Maximum**, falls gilt:

$$\forall\, \mathbf{x} \in D \setminus \{\mathbf{x}^0\} \;:\quad f(\mathbf{x}) < f(\mathbf{x}^0)$$

3) $f(\mathbf{x})$ hat in $\mathbf{x}^0$ ein **lokales Maximum**, falls es ein ε gibt mit:

$$\forall\, \mathbf{x} \in D \;:\quad \|\mathbf{x} - \mathbf{x}^0\| < \varepsilon \;\Rightarrow\; f(\mathbf{x}) \le f(\mathbf{x}^0)$$

4) $f(\mathbf{x})$ hat in $\mathbf{x}^0$ ein **strenges lokales Maximum**, falls es ein ε gibt mit:

$$\forall\, \mathbf{x} \in D \;:\quad 0 < \|\mathbf{x} - \mathbf{x}^0\| < \varepsilon \;\Rightarrow\; f(\mathbf{x}) < f(\mathbf{x}^0)$$

Analoge Definitionen für minimale Werte (siehe **Analysis I**)

Satz: (Notwendige Bedingung I)

Besitzt $f(\mathbf{x})$ mit $f \in \mathcal{C}^2$ in einem Punkt $\mathbf{x}^0 \in D^0$ ein lokales Extremum (Minimum oder Maximum), so gilt grad $f(\mathbf{x}^0) = (0, \ldots, 0)^T$.

Beweis:

Für ein beliebiges $\mathbf{v} \in \mathbb{R}^n$, $\mathbf{v} \neq 0$ ist die Funktion

$$\varphi(t) := f(\mathbf{x}^0 + t\mathbf{v})$$

in einer Umgebung von $t^0 = 0$ stetig differenzierbar.
Gleichzeitig hat $\varphi(t)$ bei $t^0 = 0$ ein lokales Extremum. Damit folgt:

$$\varphi'(0) = \text{grad}\, f(\mathbf{x}^0)\, \mathbf{v} = 0$$

Da dies für alle $\mathbf{v} \neq 0$ gilt, folgt die Bedingung:

$$\text{grad}\, f(\mathbf{x}^0) = (0, \ldots, 0)^T$$