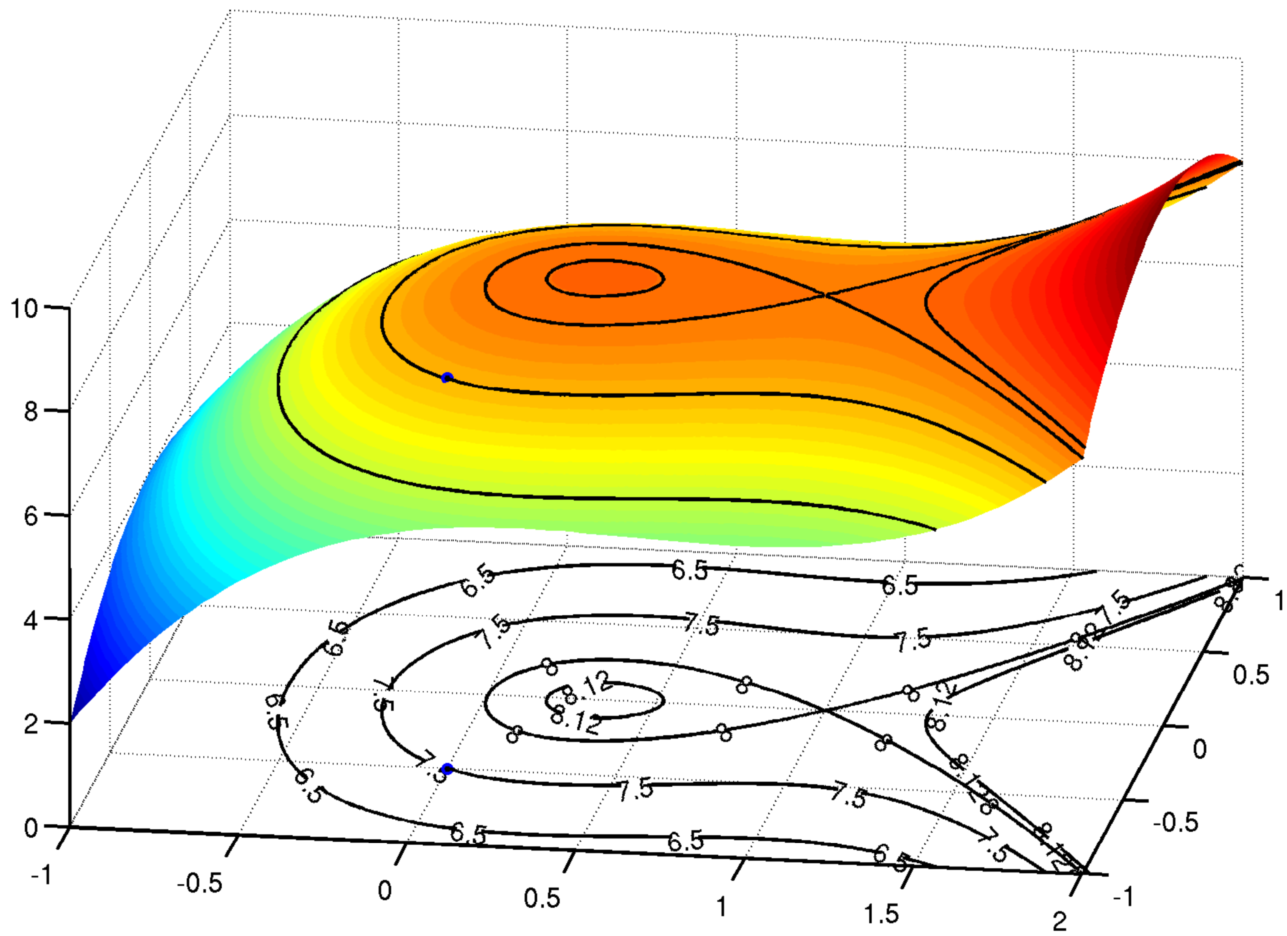
Vorlesung: Analysis II für Ingenieure

Wintersemester 09/10

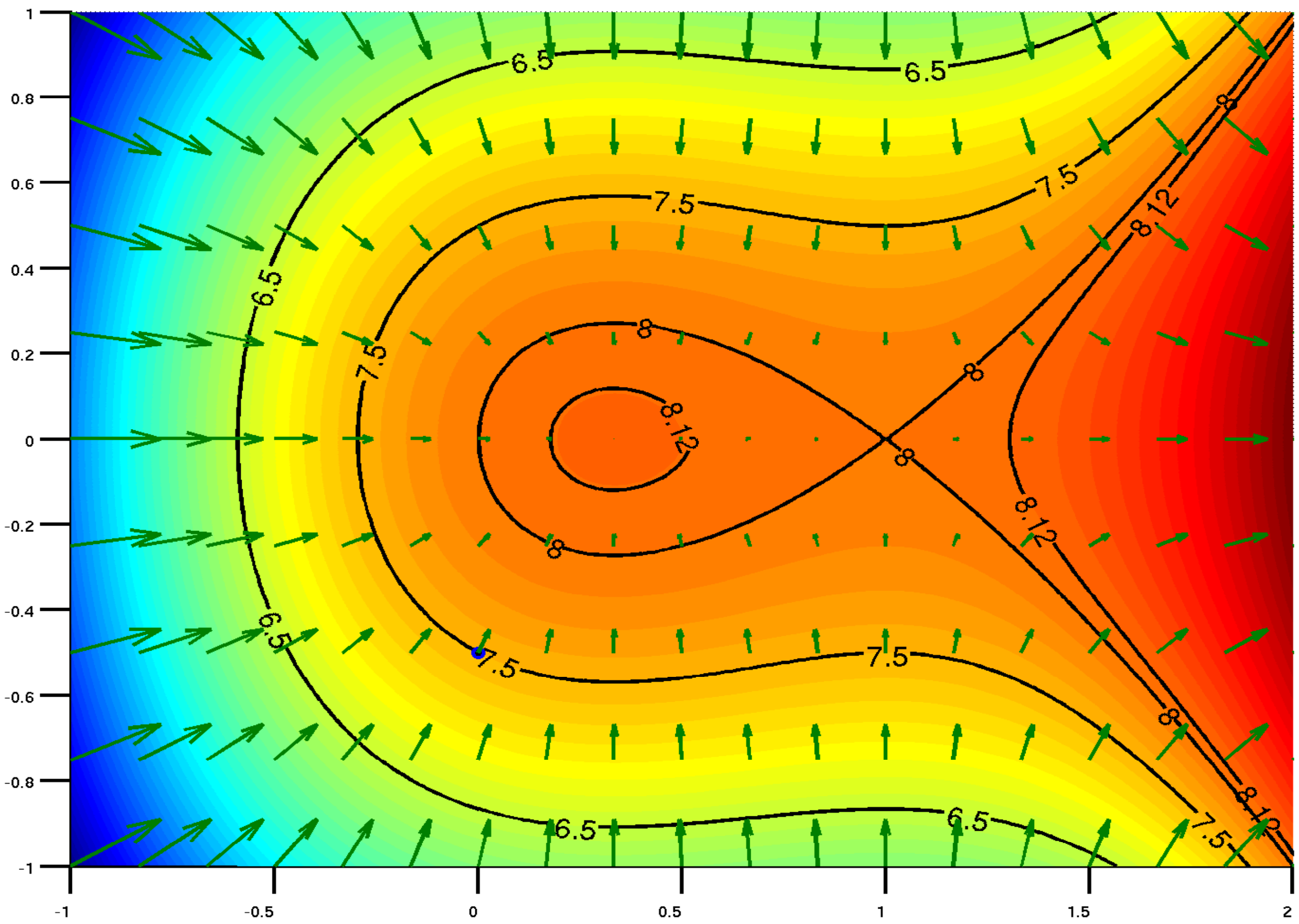
Michael Karow

Themen: Taylor-Entwicklung und lokale Extrema

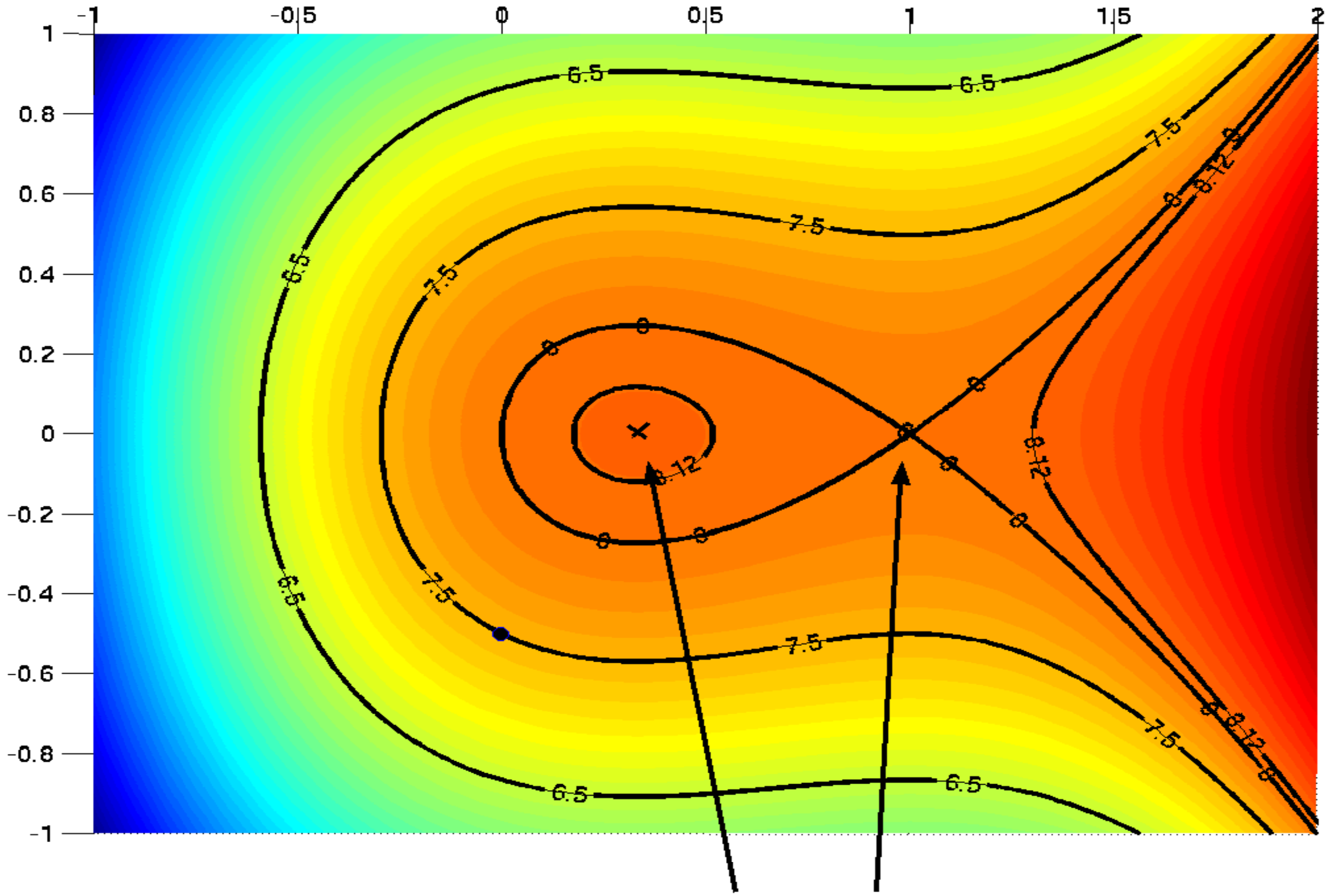
Motivierendes Beispiel: die Funktion $f(x,y) = x(x-1)^2 - 2\,y^2$.



Dieselbe Funktion von oben betrachtet mit Gradientenfeld



Nullstellen des Gradientenfeldes



An der linken Nullstelle hat f ein lokales Maximum, an der rechten einen Sattelpunkt. **Frage:** Wie kann man das rechnerisch herausfinden?

Antwort: durch Taylorentwicklung.

Taylor-Entwicklung für eine Funktion von einer Variablen.

Sei I ein Intervall und $f : I \to \mathbb{R}$ an der Stelle x m-mal diff'bar. Dann gilt für $x + \Delta x \in I$,

$$f(x + \Delta x) = f(x) + \sum_{k=1}^{m} \frac{f^{(k)}(x)}{k!} (\Delta x)^k + R_m(\Delta x),$$

wobei

$$\lim_{\Delta x \to 0} \frac{R_m(\Delta x)}{(\Delta x)^m} = 0.$$

Falls f in einer Umgebung von x sogar $(m+1)$-mal stetig differenzierbar ist, hat man die Restglieddarstellung

$$R_m(\Delta x) = \frac{f^{(m+1)}(x + t\,\Delta x)}{(m+1)!} (\Delta x)^{m+1}$$

für ein $t \in]0, 1[$, das von Δx abhängt.

Taylor-Entwicklungen bis zur dritten Ordnung

Sei $f : I \to \mathbb{R}$ an der Stelle x mindestens 3-mal diff'bar. Dann hat man

$$\begin{aligned} f(x+\Delta x) &= f(x) + f'(x)\,\Delta x + R_1(\Delta x) \\ &= f(x) + f'(x)\,\Delta x + \frac{1}{2} f''(x)\,(\Delta x)^2 + R_2(\Delta x) \\ &= f(x) + f'(x)\,\Delta x + \frac{1}{2} f''(x)\,(\Delta x)^2 + \frac{1}{6} f'''(x)\,(\Delta x)^3 + R_3(\Delta x) \end{aligned}$$

mit

$$\lim_{\Delta x \to 0} \frac{R_1(\Delta x)}{\Delta x} = 0, \qquad R_1(\Delta x) = \tfrac{1}{2} f''(x + t\,\Delta x)\,(\Delta x)^2, \quad t \in]0,1[$$

$$\lim_{\Delta x \to 0} \frac{R_2(\Delta x)}{(\Delta x)^2} = 0, \qquad R_2(\Delta x) = \tfrac{1}{6} f'''(x + t\,\Delta x)\,(\Delta x)^3, \quad t \in]0,1[$$

$$\lim_{\Delta x \to 0} \frac{R_3(\Delta x)}{(\Delta x)^3} = 0.$$

Anwendung:

Berechnung einer Näherung für $\sqrt{103}$ mittels Taylor-Entwicklung.

Sei $f(x) = \sqrt{x}$, $x = 100$. Dann $f'(x) = \frac{1}{2\sqrt{x}}$, $f''(x) = -\frac{1}{4\sqrt{x^3}}$ und

$$\begin{aligned}\sqrt{100 + \Delta x} &= f(x + \Delta x)\\ &= f(x) + f'(x)\,\Delta x + R_1(\Delta x)\\ &= 10 + \frac{1}{20}\Delta x + R_1(\Delta x)\end{aligned}$$

mit

$$R_1(\Delta x) = \frac{1}{2} f''(x + t\,\Delta x)\,(\Delta x)^2 = \frac{1}{2}\left(-\frac{1}{4\sqrt{100 + t\,\Delta x}^{\,3}}\right)(\Delta x)^2$$

Für $\Delta x > 0$ ist

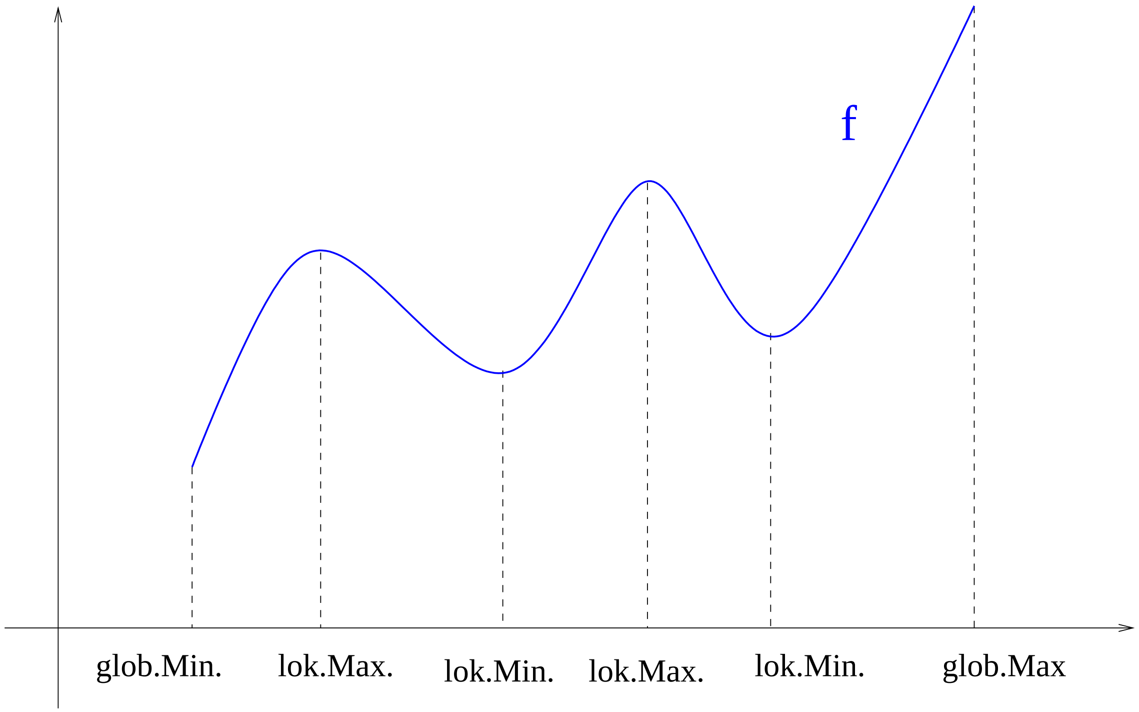
$$|R_1(\Delta x)| \le |R_1(0)| = \frac{(\Delta x)^2}{8000}$$

Für $\Delta x = 3$ bekommt man so

$$\sqrt{103} = \underbrace{10 + \frac{3}{20}}_{=10.15} + R_1(3), \qquad |R_1(3)| \le \frac{9}{8000} = 0.001125.$$

$\sqrt{103}$ auf 16 Stellen genau: 10.14889156509222

Lokale Extrema im 1-dimensionalen Fall



Eine diff'bare Funktion $f : [a, b] \longrightarrow \mathbb{R}$ mit lokalen Extrema im inneren des Definitionsbereichs und globalen Extrema am Rand.

Aus Analysis I bekannt: Sei x ein innerer Punkt des Intervalls I und sei $f : I \to \mathbb{R}$ differenzierbar an der Stelle x. Dann gelten folgende Aussgen.

1. Wenn f an der Stelle x ein lokales Minimum oder ein lokales Maximum hat, dann ist $f'(x) = 0$.

2. Wenn $f'(x) = 0$ und $f'(x) > 0$, dann hat f an der Stelle x ein lokales Minimum.

3. Wenn $f'(x) = 0$ und $f'(x) < 0$, dann hat f an der Stelle x ein lokales Maximum.

Diese Aussagen werden auf den folgenden Seite mittels Taylor-Entwicklung bewiesen.

Taylor-Entwicklung bis zur 1. Ordnung:

$$f(x + \Delta x) = f(x) + f'(x)\,\Delta x + R_1(\Delta x)$$

Angenommen, $f'(x) > 0$.

Wegen $\lim_{\Delta x \to 0} \frac{R_1(\Delta x)}{\Delta x} = 0$ hat man $\left|\frac{R_1(\Delta x)}{\Delta x}\right| < f'(x)$,

wenn $|\Delta x|$ hinreichend klein ist. Für solche Δx folgt

$$|R_1(\Delta x)| < f'(x)\,|\Delta x|$$

Konsequenz:

$f'(x)\,\Delta x$ und $f'(x)\,\Delta x + R_1(\Delta x)$ haben dasselbe Vorzeichen. Somit

$$f(x + \Delta x) = f(x) + f'(x)\,\Delta x + R_1(\Delta x) \begin{cases} > f(x) & \text{wenn } \Delta x > 0 \\ < f(x) & \text{wenn } \Delta x < 0. \end{cases}$$

Also hat f kein lokales Extremum an der Stelle x.

Analog argumentiert man, wenn $f'(x) < 0$.

Fazit: Damit f an der Stelle x ein lokales Extremum haben kann, muss $f'(x) = 0$ sein.

Taylor-Entwicklung bis zur 2. Ordnung:

$$f(x + \Delta x) = f(x) + f'(x)\,\Delta x + \frac{1}{2} f''(x)\,(\Delta x)^2 + R_2(\Delta x)$$

Angenommen, $f'(x) = 0$, $f''(x) > 0$.

Wegen $\lim_{\Delta x \to 0} \frac{R_2(\Delta x)}{(\Delta x)^2} = 0$ hat man $\left|\frac{R_2(\Delta x)}{(\Delta x)^2}\right| < \frac{1}{2} f''(x)$,

wenn $|\Delta x|$ hinreichend klein ist. Für solche Δx folgt

$$|R_2(\Delta x)| < \frac{1}{2} f''(x)\,(\Delta x)^2$$

Konsequenz:

$\frac{1}{2} f''(x)\,(\Delta x)^2$ und $\frac{1}{2} f''(x)\,(\Delta x)^2 + R_2(\Delta x)$ haben dasselbe Vorzeichen. Somit

$$f(x + \Delta x) = f(x) + \frac{1}{2} f''(x)\,(\Delta x)^2 + R_2(\Delta x) > f(x) \text{ wenn } \Delta x \neq 0.$$

Also hat f ein lokales Minimum an der Stelle x.

Fazit: Wenn $f'(x) = 0$, $f''(x) > 0$, dann hat f an der Stelle x ein lokales Minimum.

Der n-dimensionale Fall

(Wir betrachten hier aus didaktischen Gründen $n = 2$)

Höhere partielle Ableitungen

Beispiel:

$$\frac{\partial^2 f}{\partial x\,\partial y}(x,y) := \frac{\partial}{\partial x}\,\frac{\partial f}{\partial y}(x,y)$$

Man leitet die partielle Ableitung nach y einfach noch einmal nach x ab.

Satz von Schwarz (Vertauschbarkeit der Ableitungsreihenfolge) :

Wenn $\frac{\partial f}{\partial x}(x,y)$ und $\frac{\partial^2 f}{\partial x\,\partial y}(x,y)$ existieren und stetig sind, dann existiert auch $\frac{\partial^2 f}{\partial y\,\partial x}(x,y)$ und es ist

$$\frac{\partial^2 f}{\partial y\,\partial x}(x,y) = \frac{\partial^2 f}{\partial x\,\partial y}(x,y).$$

Taylor-Entwicklung für eine Funktion von 2 Variablen

$$f(x + \Delta x, y + \Delta y) = \underbrace{f(x, y) + \sum_{k=1}^{m} \frac{1}{k!} P_k(\Delta x, \Delta y)}_{\text{Taylorpolynom}} + R_m(\Delta x, \Delta y),$$

wobei

$$P_k(\Delta x, \Delta y) = \text{Summe über alle Produkte der Form}$$

$$\underbrace{\frac{\partial^k f}{\partial u_1 \partial u_2 \ldots \partial u_k}(x, y)}_{\text{partielle Ableitung}} \cdot \underbrace{(\Delta u_1)(\Delta u_2) \ldots (\Delta u_k)}_{k \text{ Faktoren}}$$

wobei $u_j = x$ oder $u_j = y$.

und

$$\lim_{(\Delta x, \Delta y) \to (0,0)} \frac{R_m(\Delta x, \Delta y)}{\left(\sqrt{(\Delta x)^2 + (\Delta y)^2}\right)^m} = 0$$

Bemerkung: $\sqrt{(\Delta x)^2 + (\Delta y)^2}$ ist die Länge des Vektors $(\Delta x, \Delta y)$.

Beispiel: der Summand P_3 des Taylor-Polynoms

$$\begin{aligned}
P_3(\Delta x, \Delta y) \quad = \quad & \frac{\partial^3 f}{\partial x\,\partial x\,\partial x}(x,y)\,(\Delta x)(\Delta x)(\Delta x) + \frac{\partial^3 f}{\partial y\,\partial y\,\partial y}(x,y)\,(\Delta y)(\Delta y)(\Delta y) \\
& + \frac{\partial^3 f}{\partial y\,\partial x\,\partial x}(x,y)\,(\Delta y)(\Delta x)(\Delta x) + \frac{\partial^3 f}{\partial x\,\partial y\,\partial y}(x,y)\,(\Delta x)(\Delta y)(\Delta y) \\
& + \frac{\partial^3 f}{\partial x\,\partial y\,\partial x}(x,y)\,(\Delta x)(\Delta y)(\Delta x) + \frac{\partial^3 f}{\partial y\,\partial x\,\partial y}(x,y)\,(\Delta y)(\Delta x)(\Delta y) \\
& + \frac{\partial^3 f}{\partial x\,\partial x\,\partial y}(x,y)\,(\Delta x)(\Delta x)(\Delta y) + \frac{\partial^3 f}{\partial y\,\partial y\,\partial x}(x,y)\,(\Delta y)(\Delta y)(\Delta x)
\end{aligned}$$

Einige Terme haben denselben Wert (Satz von Schwarz).
Man kann P_3 deshalb kürzer schreiben:

$$\begin{aligned}
P_3(\Delta x, \Delta y) \quad &= \quad \frac{\partial^3 f}{\partial x^3}(x,y)\,(\Delta x)^3 + \frac{\partial^3 f}{\partial y^3}(x,y)\,(\Delta y)^3 \\
&= \quad 3\,\frac{\partial^3 f}{(\partial x)^2\,\partial y}(x,y)\,(\Delta x)^2(\Delta y) + 3\,\frac{\partial^3 f}{\partial x\,(\partial y)^2}(x,y)\,(\Delta x)(\Delta y)^2.
\end{aligned}$$

Zur Beruhigung: Es kommt in der Praxis selten vor, dass man Taylor-Entwicklung bis zu 3. Ordnung machen muss.

Die Summanden P_1 und P_2

$$P_1(\Delta x, \Delta y) = \frac{\partial f}{\partial x}(x,y)\,\Delta x + \frac{\partial f}{\partial y}(x,y)\,\Delta y.$$

$$\begin{aligned} P_2(\Delta x, \Delta y) &= \frac{\partial^2 f}{\partial x\,\partial x}(x,y)\,(\Delta x)\,(\Delta x) + \frac{\partial^2 f}{\partial y\,\partial y}(x,y)\,(\Delta y)\,(\Delta y) \\ &\quad + \frac{\partial^2 f}{\partial x\,\partial y}(x,y)\,(\Delta x)\,(\Delta y) + \frac{\partial^2 f}{\partial y\,\partial x}(x,y)\,(\Delta y)\,(\Delta x) \\ &= \frac{\partial^2 f}{\partial x^2}(x,y)\,(\Delta x)^2 + \frac{\partial^2 f}{\partial y^2}(x,y)\,(\Delta y)^2 + 2\,\frac{\partial^2 f}{\partial x\,\partial y}(x,y)\,(\Delta x)\,(\Delta y). \end{aligned}$$

Der Summand P_1 in Vektorschreibweise

$$
\begin{aligned}
P_1(\Delta x, \Delta y) &= \frac{\partial f}{\partial x}(x,y)\,\Delta x + \frac{\partial f}{\partial y}(x,y)\,\Delta y \\
&= \begin{bmatrix} \frac{\partial f}{\partial x}(x,y) \\ \frac{\partial f}{\partial y}(x,y) \end{bmatrix} \cdot \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} \\
&= \nabla f(x,y) \underset{\substack{\uparrow \\ \text{Skalarprodukt}}}{\cdot} \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix}
\end{aligned}
$$

Der Summand P_1 ist nicht anderes als das Differential von f an der Stelle (x, y).

Der Summand P_2 in Matrix-Vektorschreibweise

$$\begin{aligned} P_2(\Delta x, \Delta y) &= \frac{\partial^2 f}{\partial x^2}(x,y)\,(\Delta x)^2 + \frac{\partial^2 f}{\partial y \partial x}(x,y)\,(\Delta y)(\Delta x) \\ &\quad + \frac{\partial^2 f}{\partial x \partial y}(x,y)\,(\Delta x)(\Delta y) + \frac{\partial^2 f}{\partial y^2}(x,y)\,(\Delta y)^2 \\ &= \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} \cdot \underbrace{\begin{bmatrix} \frac{\partial^2 f}{\partial x^2}(x,y) & \frac{\partial^2 f}{\partial x \partial y}(x,y) \\ \frac{\partial^2 f}{\partial y \partial x}(x,y) & \frac{\partial^2 f}{\partial y^2}(x,y) \end{bmatrix}}_{(\text{Hess } f)(x,y)} \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} \end{aligned}$$

Die Matrix der zweiten partiellen Ableitungen heisst **Hesse-Matrix**.

Taylor-Entwicklung bis zur zweiten Ordnung für eine Funktion von 2 Variablen

$$f(x + \Delta x, y + \Delta y)$$

$$= \quad f(x,y) + P_1(\Delta x, \Delta y) + \frac{1}{2} P_2(\Delta x, \Delta y) + R_2(\Delta x, \Delta y)$$

$$= \quad f(x,y) + \nabla f(x,y) \cdot \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} + \frac{1}{2} \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} \cdot (\text{Hess } f)(x,y) \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} + R_2(\Delta x, \Delta y)$$

Lokale Extrema

Wie im 1-dimensionalen Fall argumentiert man:
Wenn f an der Stelle (x, y) ein lokales Maximum oder ein lokales Minimum hat, dann müssen an der Stelle (x, y) alle partiellen Ableitungen 1. Ordnung 0 sein.

Notwendige Bedingung für ein lokales Extremum ist also

$$(\nabla f)(x, y) = (0, 0) \qquad (*)$$

Terminologie: Wenn die Bedingung $(*)$ erfüllt ist, dann nennt man (x, y) einen **kritischen Punkt**.

Lokale Extrema II

Angenommen, die notwendige Bedingung $(\nabla f)(x,y) = (0,0)$ ist erfüllt. Dann sieht die Taylor-Entwicklung so aus:

$$f(x+\Delta x, y+\Delta y) \;=\; f(x,y) + \underbrace{\frac{1}{2}\begin{bmatrix}\Delta x\\ \Delta y\end{bmatrix} \cdot (\text{Hess } f)(x,y)\begin{bmatrix}\Delta x\\ \Delta y\end{bmatrix}}_{=P_2(\Delta x,\Delta y)} + R_2(\Delta x, \Delta y)$$

Es gilt: Wenn $P_2(\Delta x, \Delta y) \neq 0$ und $(\Delta x, \Delta y)$ hinreichend klein, dann ist

$$|R_2(\Delta x, \Delta y)| << |P_2(\Delta x, \Delta y)|$$

Unter dieser Voraussetzung folgt:

$$f(x+\Delta x, y+\Delta y) > f(x,y) \quad \Leftrightarrow \quad P_2(\Delta x, \Delta y) > 0$$

$$f(x+\Delta x, y+\Delta y) < f(x,y) \quad \Leftrightarrow \quad P_2(\Delta x, \Delta y) < 0$$

Zitat von der vorherigen Seite:

Wenn $\nabla f(x,y) = (0,0)$ und $(\Delta x, \Delta y)$ hinreichend klein, dann gilt

$$f(x+\Delta x, y+\Delta y) > f(x,y) \quad \Leftrightarrow \quad P_2(\Delta x, \Delta y) > 0,$$

$$f(x+\Delta x, y+\Delta y) < f(x,y) \quad \Leftrightarrow \quad P_2(\Delta x, \Delta y) < 0.$$

Folgerungen:

1. Wenn $P_2(\Delta x, \Delta y) > 0$ für **alle** $(\Delta x, \Delta y) \neq (0,0)$,
   dann hat f an der Stelle (x,y) ein lokales Minimum.

2. Wenn $P_2(\Delta x, \Delta y) < 0$ für **alle** $(\Delta x, \Delta y) \neq (0,0)$,
   dann hat f an der Stelle (x,y) ein lokales Maximum.

3. Wenn $P_2(\Delta x, \Delta y) < 0$ für **einige** $(\Delta x, \Delta y)$ und $P_2(\Delta x, \Delta y) > 0$ für **einige** $(\Delta x, \Delta y)$,
   dann hat f an der Stelle (x,y) weder ein lok. Maximum noch ein lok. Minimum,
   sondern einen **Sattelpunkt**.

Terminologie:
Im 1. Fall nennt man P_2 und ebenso die zugehörige Hesse-Matrix **positiv definit**.
Im 2. Fall nennt man P_2 und ebenso die zugehörige Hesse-Matrix **negativ definit**.
Im 3. Fall nennt man P_2 und ebenso die zugehörige Hesse-Matrix **indefinit**.

Frage: Wie prüft man Definitheit? (Antwort auf der nächsten Seite)

Allgem. Definitheitsdefinition und Definitheitstest für 2×2-Matrizen

Allgemeine Definitheitsdefinition: Eine symmetrische Matrix $A \in \mathbb{R}^{n\times n}$ heisst

1. positiv definit, wenn $\vec{x} \cdot A\vec{x} > 0$ für alle $\vec{x} \neq 0$,
2. negativ definit, wenn $\vec{x} \cdot A\vec{x} < 0$ für alle $\vec{x} \neq 0$,
3. indefinit, wenn $\vec{x} \cdot A\vec{x} < 0$ für einige $\vec{x}$ und $\vec{x} \cdot A\vec{x} > 0$ für einige $\vec{x}$.

Bemerkung: Es gibt noch die semidefiniten Fälle. Siehe lineare Algebra.

Definitheitstest im 2×2-Fall: Die symmetrische Matrix $A = \begin{bmatrix} a & b \\ b & c \end{bmatrix} \in \mathbb{R}^{2\times 2}$ ist

1. positiv definit, wenn $\det(A) > 0$ und $a > 0$,
2. negativ definit, wenn $\det(A) > 0$ und $a < 0$,
3. indefinit, wenn $\det(A) < 0$.

Wenn $\det(A) = 0$, dann ist A semidefinit.

Der Grund, warum wir den semidefiniten Fall nicht näher betrachten, ist folgender:

Wenn die Hesse-Matrix Hess $f(x,y)$ **nur semidefinit ist, dann hat man keine Aussage über lokale Maxima/Minima. Man muss dann höhere Ableitungen betrachten**

Einfaches Rechenbeispiel

Sei $f(x,y) = x(x-1)^2 - 2\,y^2 + 8$ (Beispielfunktion aus der Einleitung). Dann

$$\nabla f(x,y) \;=\; \begin{bmatrix} \frac{\partial f}{\partial x}(x,y) \\ \frac{\partial f}{\partial y}(x,y) \end{bmatrix} = \begin{bmatrix} (x-1)(3x-1) \\ -4\,y \end{bmatrix} = \begin{bmatrix} 3\,x^2 - 4\,x + 1 \\ -4\,y \end{bmatrix}$$

$$(\text{Hess } f)(x,y) \;=\; \begin{bmatrix} \frac{\partial^2 f}{\partial x^2}(x,y) & \frac{\partial^2 f}{\partial x \partial y}(x,y) \\ \frac{\partial^2 f}{\partial y \partial x}(x,y) & \frac{\partial^2 f}{\partial y^2}(x,y) \end{bmatrix} = \begin{bmatrix} 6x-4 & 0 \\ 0 & -4 \end{bmatrix}$$

Nullstellen des Gradienten (=kritische Punkte, Kandidaten für ein lokales Extremum):

$$\nabla f(x,y) \;\Leftrightarrow\; (x,y) = (\frac{1}{3}, 0) \quad \text{oder} \quad (x,y) = (1,0)$$

Einsetzen dieser Punkte in die Hesse-Matrix:

$$(\text{Hess } f)(\frac{1}{3}, 0) = \underbrace{\begin{bmatrix} -2 & 0 \\ 0 & -4 \end{bmatrix}}_{\text{negativ definit}} \qquad (\text{Hess } f)(1,0) = \underbrace{\begin{bmatrix} 2 & 0 \\ 0 & -4 \end{bmatrix}}_{\text{indefinit}}$$

$\Rightarrow f$ hat ein lok. Maximum an der Stelle $(\frac{1}{3}, 0)$ und einen Sattelpunkt an der Stelle $(1,0)$.