

- Der Satz gilt analog für  $k$ -te gemischte partielle Ableitungen, falls  $f \in C^k(D, \mathbb{R})$ .

**2.11 Definition.** Für zwei Funktionen  $f, g : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$  und  $\mathbf{x}_0 \in D$ ,  $k \in \mathbb{N}_0$  schreibt man

$$f(\mathbf{x}) = g(\mathbf{x}) + o(\|\mathbf{x} - \mathbf{x}_0\|^k) \quad \text{für } \mathbf{x} \rightarrow \mathbf{x}_0$$

falls

$$\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} \frac{f(\mathbf{x}) - g(\mathbf{x})}{\|\mathbf{x} - \mathbf{x}_0\|^k} = 0.$$

- Sei  $f : \mathbb{R} \rightarrow \mathbb{R}$  differenzierbar an der Stelle  $x_0$ . Nach Kapitel 3 1.6 ist die beste lineare Approximation von  $f$  an der Stelle  $x_0$  gegeben durch  $g(x) = f(x_0) + f'(x_0)(x - x_0)$ , d.h.

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + o(|x - x_0|).$$

Für  $f : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$  reicht partielle Differenzierbarkeit für das Analogon obiger Formel nicht.

**2.12 Definition.** Sei  $D \subseteq \mathbb{R}^n$  offen. Eine Funktion  $f : D \rightarrow \mathbb{R}$  heißt in  $\mathbf{x}_0 \in D$  **total differenzierbar**, wenn es einen Vektor  $\mathbf{a} \in \mathbb{R}^n$  mit

$$f(\mathbf{x}) = f(\mathbf{x}_0) + \mathbf{a} \cdot (\mathbf{x} - \mathbf{x}_0) + o(\|\mathbf{x} - \mathbf{x}_0\|), \quad (2.13)$$

für  $\mathbf{x}$  nahe  $\mathbf{x}_0$ , gibt.

**2.14 Satz.** Ist  $f$  in  $\mathbf{x}_0 \in D$  total differenzierbar mit  $f(\mathbf{x}) = f(\mathbf{x}_0) + \mathbf{a} \cdot (\mathbf{x} - \mathbf{x}_0) + o(\|\mathbf{x} - \mathbf{x}_0\|)$ , dann gilt:

a)  $f$  ist in  $\mathbf{x}_0$  stetig,

$$b) \lim_{t \rightarrow 0} \frac{1}{t} (f(\mathbf{x}_0 + t\mathbf{v}) - f(\mathbf{x}_0)) = \mathbf{a} \cdot \mathbf{v}, \quad \forall \mathbf{v} \in \mathbb{R}^n, \mathbf{v} \neq 0,$$

c)  $f$  ist partiell differenzierbar und  $\mathbf{a}$  ist eindeutig bestimmt als  $\mathbf{a} = \text{grad } f(\mathbf{x}_0)$ .

- Wenn  $f$  total differenzierbar ist, so ist

$$f(\mathbf{x}) = f(\mathbf{x}_0) + \text{grad } f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0) + o(\|\mathbf{x} - \mathbf{x}_0\|) \quad (2.15)$$

die beste lineare Approximation von  $f$  nahe  $\mathbf{x}_0$ .

**2.16 Satz.** Jede  $C^1$ -Funktion  $f : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ , mit  $D$  offen, ist auf  $D$  total differenzierbar, d.h. jede stetig partiell differenzierbare Funktion ist total differenzierbar.

**2.17 Newton - Verfahren.** Zur näherungsweise Lösung eines nichtlinearen Gleichungssystems

$$f(x, y) = 0, \quad g(x, y) = 0, \quad (*)$$

für  $f, g \in C^1$  ersetzt man (\*) durch die lineare Approximation

$$\begin{aligned} f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0) &= 0, \\ g(x_0, y_0) + g_x(x_0, y_0)(x - x_0) + g_y(x_0, y_0)(y - y_0) &= 0. \end{aligned}$$

Dieses kann man als

$$\mathbf{A} \begin{pmatrix} x \\ y \end{pmatrix} = \mathbf{g}$$

schreiben, wobei

$$\begin{aligned} \mathbf{A} &= \begin{pmatrix} f_x(x_0, y_0) & f_y(x_0, y_0) \\ g_x(x_0, y_0) & g_y(x_0, y_0) \end{pmatrix}, \\ \mathbf{g} &= - \begin{pmatrix} f(x_0, y_0) \\ g(x_0, y_0) \end{pmatrix} + \mathbf{A} \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}. \end{aligned}$$

Falls  $\det \mathbf{A} \neq 0$  erhält man eine verbesserte Lösung zum Startwert  $(x_0, y_0)^T$  durch

$$\begin{aligned} x_1 &= x_0 + \frac{1}{\det \mathbf{A}} \det \begin{pmatrix} -f(x_0, y_0) & f_y(x_0, y_0) \\ -g(x_0, y_0) & g_y(x_0, y_0) \end{pmatrix}, \\ y_1 &= y_0 + \frac{1}{\det \mathbf{A}} \det \begin{pmatrix} f_x(x_0, y_0) & -f(x_0, y_0) \\ g_x(x_0, y_0) & -g(x_0, y_0) \end{pmatrix}. \end{aligned}$$

Dieses Vorgehen kann wiederholt werden und liefert oft gute Ergebnisse bei entsprechender Wahl des Startpunktes.

**2.18 Definition.** Sei  $f : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$  und  $\mathbf{x}_0 \in D$  und  $\mathbf{v} \in \mathbb{R}^n$  ein Einheitsvektor, d.h.  $\|\mathbf{v}\| = 1$ . Falls der Grenzwert

$$\lim_{t \rightarrow 0} \frac{1}{t} (f(\mathbf{x}_0 + t\mathbf{v}) - f(\mathbf{x}_0))$$

existiert wird er mit  $\partial_{\mathbf{v}} f(\mathbf{x}_0)$  oder  $\frac{\partial f}{\partial \mathbf{v}}(\mathbf{x}_0)$  bezeichnet und heißt **Richtungsableitung** von  $f$  an der Stelle  $\mathbf{x}_0$  in Richtung  $\mathbf{v}$ .

- Dies ist eine Verallgemeinerung von partiellen Ableitungen welche der Wahl  $\mathbf{v} = \mathbf{e}_i$  entspricht.
- $\partial_{\mathbf{v}} f(\mathbf{x}_0)$  beschreibt das Verhalten von  $f$  längs der Geraden  $\mathbf{x}_0 + t\mathbf{v}$ .

**2.19 Satz.** Für jede auf der offenen Menge  $D \subseteq \mathbb{R}^n$  total differenzierbare Funktion  $f$  und alle  $\mathbf{v} \in \mathbb{R}^n$ ,  $\|\mathbf{v}\| = 1$  gilt:

$$\frac{\partial f}{\partial \mathbf{v}}(\mathbf{x}_0) = \text{grad } f(\mathbf{x}_0) \cdot \mathbf{v} = \sum_{i=1}^n f_{x_i}(\mathbf{x}_0) v_i$$

- Wir hatten bereits gesehen, dass sich Funktionen längs verschiedener Kurven verschieden verhalten können (siehe Beispiel stetiger „partieller“ Funktionen). Deshalb ist es sinnvoll auch Ableitungen längs von Kurven und nicht nur längs von Geraden (Richtungsableitungen) zu betrachten.

**2.20 Satz (Kettenregel).** Für jede  $C^1$ -Funktion  $f : D \rightarrow \mathbb{R}$ ,  $D \subseteq \mathbb{R}^n$  offen, und für jedes Kurvenstück  $\mathbf{x} : [a, b] \rightarrow D$  gilt:

$$\frac{d}{dt} f(\mathbf{x}(t)) = \text{grad } f(\mathbf{x}(t)) \cdot \dot{\mathbf{x}}(t). \quad (2.21)$$

**2.22 Polarkoordinaten.** Im  $\mathbb{R}^2$  besteht folgende Beziehung zwischen den Polarkoordinaten  $(r, \varphi)$  und den kartesischen Koordinaten  $(x, y)$

$$x = r \cos \varphi, \quad y = r \sin \varphi.$$

Eine Funktion  $f$  von  $(x, y)$  kann also durch

$$F(r, \varphi) := f(r \cos \varphi, r \sin \varphi)$$

als Funktion von  $(r, \varphi)$  betrachtet werden. Die partiellen Ableitungen von  $F$  berechnen sich durch die Kettenregel als

$$\begin{aligned} F_r &= f_x \cos \varphi + f_y \sin \varphi, \\ F_\varphi &= -f_x r \sin \varphi + f_y r \cos \varphi. \end{aligned} \quad (2.23)$$

Dieses Gleichungssystem kann man nach  $f_x$  und  $f_y$  auflösen und erhält für  $r \neq 0$

$$\begin{aligned} f_x &= F_r \cos \varphi - \frac{1}{r} F_\varphi \sin \varphi, \\ f_y &= F_r \sin \varphi + \frac{1}{r} F_\varphi \cos \varphi. \end{aligned} \quad (2.24)$$

Es gibt Formeln auch für höhere partielle Ableitungen. Insbesondere haben wir

$$\Delta f := f_{xx} + f_{yy} = F_{rr} + \frac{1}{r}F_r + \frac{1}{r^2}F_{\varphi\varphi}, \quad (2.25)$$

welches die Formeln für den **Laplace - Operator**  $\Delta$  in kartesischen Koordinaten und in Polarkoordinaten sind. Anloge Formeln gelten für die Polar- oder Kugelkoordinaten im  $\mathbb{R}^3$ .

## 7.3 Anwendungen

**3.1 Richtung des stärksten Anstiegs.** Sei  $f : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$  eine  $C^1$ -Funktion und  $\mathbf{x}_0 \in D$ . Für die Richtungsableitung in Richtung  $\mathbf{v}$  gilt

$$\begin{aligned} \partial_{\mathbf{v}} f(\mathbf{x}_0) &= \text{grad } f(\mathbf{x}_0) \cdot \mathbf{v} \\ &= \|\mathbf{v}\| \|\text{grad } f(\mathbf{x}_0)\| \cos \alpha, \end{aligned}$$

wobei  $\alpha$  der Winkel zwischen  $\mathbf{v}$  und  $\text{grad } f(\mathbf{x}_0)$  ist. Somit ist der Wert der Richtungsableitung maximal wenn  $\cos \alpha = 1$ , d.h.  $\mathbf{v} = \text{grad } f(\mathbf{x}_0)$ . Somit haben wir:

Richtung von  $\text{grad } f(\mathbf{x}_0)$  = Richtung des maximalen Anstiegs von  $f$  in  $\mathbf{x}_0$ .

**3.2 Gradientenverfahren.** Zur Bestimmung des Maximums einer Funktion  $f : D \rightarrow \mathbb{R}$  wählt man einen Startwert  $\mathbf{x}_0$ , bestimmt dann die Richtung des maximalen Anstiegs  $\text{grad } f(\mathbf{x}_0)$  und berechnet einen neuen Punkt

$$\mathbf{x}_1 = \mathbf{x}_0 + h \text{grad } f(\mathbf{x}_0)$$

mit einer Schrittweite  $h$ . Falls  $f(\mathbf{x}_1) < f(\mathbf{x}_0)$  ist man zu weit gegangen und versucht es mit halber Schrittweite  $\frac{h}{2}$ . Falls  $f(\mathbf{x}_1) > f(\mathbf{x}_0)$  wiederholt man das Verfahren mit  $\mathbf{x}_1$  als Startwert. Auf diese Weise nähert man sich einer Maximalstelle von  $f$  in  $D$ , welche vom Startwert abhängt und nicht das globale Maximum sein muß. Analog für Minimum.

**3.3 Tangentialebene.** Sei  $D \subseteq \mathbb{R}^n$  offen und  $f : D \rightarrow \mathbb{R}$  eine  $C^1$  Funktion. Die Niveaumenge  $N_c = \{\mathbf{x} : f(\mathbf{x}) = c\}$  ist eine Hyperfläche. Für jede Kurve  $\mathbf{x} : [a, b] \rightarrow N_c$  in  $N_c$  gilt also  $f(\mathbf{x}(t)) = c$  und somit

$$\text{grad } f(\mathbf{x}(t)) \cdot \dot{\mathbf{x}}(t) = 0, \quad (3.4)$$

d.h. der Gradient  $\text{grad } f$  steht orthogonal zu allen Kurventangenten  $\dot{\mathbf{x}}$ .  
Für  $n = 2$  ist also die Normalengleichung der Tangente an die Niveaukurve  $\mathbf{N}_c$  durch den Punkt  $(x_0, y_0) \in \mathbf{N}_c$  gegeben durch

$$f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0) = 0.$$

Für  $n = 3$  spannen die Tangentialvektoren aller Niveaukurven durch einen Punkt  $(x_0, y_0, z_0) \in \mathbf{N}_c$  eine Ebene auf, die **Tangentialebene**. Die Normalengleichung der Tangentialebene ist

$$\text{grad } f(x_0, y_0, z_0) \cdot (x - x_0, y - y_0, z - z_0) = 0.$$

- Ein Gebiet  $D$  heißt **konvex**, wenn  $D$  offen ist und gilt:

$$\mathbf{x}, \mathbf{y} \in D \quad \Rightarrow \quad t\mathbf{y} + (1 - t)\mathbf{x} \in D \quad \forall t \in [0, 1].$$

**3.5 Satz (Taylor-Formel).** Ist  $D \in \mathbb{R}^n$  ein konvexes Gebiet,  $f \in C^{k+1}(D)$ ,  $\mathbf{x} \in D$ , dann gilt mit  $\mathbf{x} + \mathbf{v} \in D$  die Taylorformel

$$f(\mathbf{x} + \mathbf{v}) = f(\mathbf{x}) + \partial_{\mathbf{v}} f(\mathbf{x}) + \frac{1}{2!} \partial_{\mathbf{v}}^2 f(\mathbf{x}) + \cdots + \frac{1}{k!} \partial_{\mathbf{v}}^k f(\mathbf{x}) + R_k(\mathbf{x}, \mathbf{v})$$

mit dem Restglied

$$R_k(\mathbf{x}, \mathbf{v}) = \frac{1}{(k+1)!} \partial_{\mathbf{v}}^{k+1} f(\mathbf{x} + \xi_k \mathbf{v}),$$

wobei  $\xi_k \in (0, 1)$  und  $\partial_{\mathbf{v}} f = \sum_{i=0}^n v_i \frac{\partial f}{\partial x_i}$ .

- Sei  $f \in C^{k+1}(D)$ , dann ist für  $\mathbf{v} = \mathbf{x} - \mathbf{x}_0$

$$p(\mathbf{x}) := \sum_{i=0}^k \partial_{\mathbf{v}}^i f(\mathbf{x}_0) \tag{3.6}$$

das **Taylor-Polynom** von  $f$  an der Stelle  $\mathbf{x}_0$ . Es gilt:

$$f(\mathbf{x}) = p(\mathbf{x}) + o(\|\mathbf{x} - \mathbf{x}_0\|^k). \tag{3.7}$$

**3.8 Definition.** Für eine  $C^2$ -Funktion  $f$  heißt die symmetrische Matrix

$$H_f(\mathbf{x}) := \left( \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}) \right)_{i,j=1,\dots,n}$$

die **Hesse-Matrix** von  $f$  an der Stelle  $\mathbf{x}$ .

**3.9 Folgerung.** Sei  $f \in C^{k+1}(D)$ ,  $k = 0, 1, 2$ , wobei  $D$  ein konvexes Gebiet ist und  $\mathbf{x}, \mathbf{x}_0 \in D$ . Dann gilt mit Punkten  $\tilde{\mathbf{x}}$  und  $\mathbf{x}^*$  zwischen  $\mathbf{x}$  und  $\mathbf{x}_0$ :

$$\begin{aligned} f(\mathbf{x}) &= f(\mathbf{x}_0) + \text{grad } f(\tilde{\mathbf{x}}) \cdot (\mathbf{x} - \mathbf{x}_0), \\ f(\mathbf{x}) &= f(\mathbf{x}_0) + \text{grad } f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0) \\ &\quad + \frac{1}{2}(\mathbf{x} - \mathbf{x}_0) H_f(\mathbf{x}^*)(\mathbf{x} - \mathbf{x}_0), \\ f(\mathbf{x}) &= f(\mathbf{x}_0) + \text{grad } f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0) \\ &\quad + \frac{1}{2}(\mathbf{x} - \mathbf{x}_0) H_f(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) + o(\|\mathbf{x} - \mathbf{x}_0\|^2). \end{aligned}$$

- Bisher hatten wir es mit Funktionen der Form

$$y = g(x) \tag{*}$$

zu tun. Aber man kann auch eine Abhängigkeit zwischen  $x$  und  $y$  implizit durch eine Gleichung

$$f(x, y) = 0$$

vorgeben. Wie kann man daraus eine Funktion der Form (\*) gewinnen?

**3.10 Definition.** Sei  $f : D \subseteq \mathbb{R}^2 \rightarrow \mathbb{R}$ . Man sagt, durch  $f(x, y) = 0$  ist auf dem Intervall  $I \subseteq \mathbb{R}$  eine **implizite Funktion**  $g : I \rightarrow K$  mit Werten in  $K \subseteq \mathbb{R}$  erklärt, wenn es zu jedem  $x \in I$  genau ein  $y \in K$  gibt mit  $(x, y) \in D$  und  $f(x, y) = 0$ . Dieses  $y$  wird als  $g(x)$  bezeichnet.

**Beispiele:**

- Sei  $f(x, y) = 3x + 2y - 1 = 0$ . Für alle  $x$  ist diese Gleichung eindeutig nach  $y$  auflösbar. Wir erhalten eine explizite Darstellung  $y = g(x) = \frac{1 - 3x}{2}$ .

- Sei  $f(x, y) = e^y + y^3 + x^3 + x^2 - 1 = 0$ . Für alle  $x$  hat  $e^y + y^3 = 1 - x^2 - x^3$  genau eine Lösung  $g(x)$ , denn  $h(y) = e^y + y^3$  ist strikt monoton steigend. Aber  $g$  ist nicht durch elementare Umformungen darstellbar. Daher ist nur die Existenz einer Funktion  $g(x)$  mit

$$e^{g(x)} + g(x)^3 = 1 - x^2 - x^3$$

gesichert.

**3.11 Satz (über implizite Funktionen).** Sei  $D \in \mathbb{R}^2$  offen und  $f : D \rightarrow \mathbb{R}$  eine  $C^1$ -Funktion. Ist  $(x_0, y_0) \in D$  ein Punkt der Niveaumenge  $f(x, y) = 0$  mit  $f_y(x_0, y_0) \neq 0$ , dann gibt es Intervalle  $I \subseteq \mathbb{R}$  und  $K \subseteq \mathbb{R}$  mit den Mittelpunkten  $x_0$  bzw.  $y_0$  so das gilt:

- $R = \{(x, y) : x \in I, y \in K\} \subseteq D$ ,  $f_y(x, y) \neq 0$  für alle  $(x, y) \in R$ .
- Durch  $f(x, y) = 0$  ist auf  $I$  eindeutig eine differenzierbare implizite Funktion  $g : I \rightarrow K$  erklärt, für deren Ableitung gilt:

$$g'(x) = -\frac{f_x(x, g(x))}{f_y(x, g(x))} = -\frac{f_x(x, y)}{f_y(x, y)} \quad \forall x \in I.$$

- Nachdem wir wissen, dass  $g$  differenzierbar ist, kann man die Formel aus Satz 3.11 b) leichter einsehen durch die Kettenregel.

$$0 = f(x, g(x)) \quad \Rightarrow \quad 0 = f_x(x, g(x)) + f_y(x, g(x))g'(x)$$

Man kann auch höhere Ableitungen berechnen; z.B.

$$f_{xx} + f_{xy}g' + g_{yx}g' + f_{yy}(g')^2 + f_y g'' = 0.$$

Also gilt:

$$\begin{aligned} g'(x) &= \frac{-f_x}{f_y} \\ g''(x) &= -\frac{1}{f_y} (f_{xx} + f_{xy}g' + g_{yx}g' + f_{yy}(g')^2) \\ &= -\frac{1}{(f_y)^3} (f_{xx} (f_y)^2 - 2f_{xy}f_x f_y + f_{yy} (f_x)^2). \end{aligned}$$

Somit kann man Extremstellen der impliziten Funktion  $g(x)$  bestimmen. Seien  $x$  und  $y = g(x)$  derart, dass

$$f(x, y) = 0 \quad f_x(x, y) = 0 \quad f_y(x, y) \neq 0$$

dann ist  $x$  ein lokales Maximum der Funktion  $g(x)$  falls

$$\frac{f_{xx}}{f_y} > 0$$

und  $x$  ist ein lokales Minimum falls  $\frac{f_{xx}}{f_y} < 0$ .

**Beispiel:** Sei  $f(x, y) = e^y + y^3 + x^3 + x^2 - 1$ . Dann gilt:

$$f_x = 3x^2 + 2x, \quad f_y = e^y + 3y^2, \quad f_{xx} = 6x + 2.$$

Sei  $g$  die in Satz 3.11 b) definierte Funktion, so gilt:

$$g'(x) = -\frac{f_x}{f_y} = -\frac{3x^2 + 2x}{e^y + 3y^2}.$$

Hat  $g$  an der Stelle  $x_0$  eine Extremstelle, d.h.  $g'(x_0) = 0$ , so gilt:

$$0 = \frac{f_x(x_0, g(x_0))}{f_y(x_0, g(x_0))} = \frac{3x_0^2 + 2x_0}{e^{y_0} + 3y_0^2}.$$

Dies ergibt  $x_0 = 0$  oder  $x_0 = -\frac{2}{3}$ . Wir haben

$$g''(x_0) = -\frac{f_{xx}(x_0, y_0)}{f_y(x_0, y_0)} = -\frac{6x_0 + 2}{e^{y_0} + 3y_0^2}$$

und also ist  $g''(0) = -2$  also ist 0 ein Maximum. Weiter gilt  $g''(-\frac{2}{3}) = 2$  und also ist  $-\frac{2}{3}$  ein Minimum.

**3.12 Definition.** Sei  $f : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ . Ein Punkt  $\mathbf{a} \in D$  heißt **lokales Maximum** (bzw. **Minimum**) von  $f$ , wenn es eine  $r$ -Umgebung  $B_r(\mathbf{a})$  gibt so, dass für alle  $\mathbf{x} \in B_r(\mathbf{a}) \cap D$  gilt

$$f(\mathbf{x}) \leq f(\mathbf{a}) \quad (\text{ bzw. } f(\mathbf{a}) \leq f(\mathbf{x})).$$

Falls diese Ungleichungen für alle  $\mathbf{x} \in D$  gelten, heißt  $\mathbf{a}$  **globales Maximum** (bzw. **Minimum**). Ein **Extremum** ist ein Minimum oder ein Maximum.



**3.13 Satz (Lokale Extrema im Innern).** Ist  $f$  auf  $B_r(\mathbf{a})$ ,  $\mathbf{a} \in \mathbb{R}^n$ , eine  $C^1$ -Funktion, so gilt:

$$\mathbf{a} \text{ ist lokale Extremstelle} \quad \Rightarrow \quad \text{grad } f(\mathbf{a}) = 0.$$

- Ein Punkt  $\mathbf{a} \in \mathbb{R}^n$  für den  $\text{grad } f(\mathbf{a}) = 0$  gilt, heißt **stationärer Punkt**. Ein stationärer Punkt, der kein Extrempunkt ist, heißt **Sattelpunkt**.
- Satz 3.13 besagt, dass man Extrempunkte im Innern zwischen den stationären Punkten suchen muss.

**3.14 Satz.** Ist  $U \subseteq \mathbb{R}^n$  offen,  $\mathbf{a} \in U$  ein stationärer Punkt einer Funktion  $f \in C^2(U, \mathbb{R})$  und  $H_f(\mathbf{a}) = (f_{x_i x_j}(\mathbf{a}))$  die Hesse-Matrix von  $f$  in  $\mathbf{a}$ . Dann gilt:

- Ist  $H_f(\mathbf{a})$  positiv definit, so ist  $\mathbf{a}$  ein lokales Minimum.
- Ist  $H_f(\mathbf{a})$  negativ definit, so ist  $\mathbf{a}$  ein lokales Maximum.
- Ist  $H_f(\mathbf{a})$  indefinit, so ist  $\mathbf{a}$  ein Sattelpunkt.

**3.15 Folgerung.** Sei  $f : U \subseteq \mathbb{R}^2 \rightarrow \mathbb{R}$  ein  $C^2$ -Funktion und  $\mathbf{a}$  ein stationärer Punkt. Dann gilt:

- $f_{xx}(\mathbf{a}) > 0$ , und  $\det H_f(\mathbf{a}) > 0$ , so ist  $\mathbf{a}$  lokales Minimum,
  - $f_{xx}(\mathbf{a}) < 0$ , und  $\det H_f(\mathbf{a}) > 0$ , so ist  $\mathbf{a}$  lokales Maximum,
  - $\det H_f(\mathbf{a}) < 0$ , so ist  $\mathbf{a}$  Sattelpunkt.
- Sei  $\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$  negativ definit. Da  $\mathbf{A}$  ähnlich zu  $\mathbf{B} = \begin{pmatrix} a & 0 \\ 0 & d - \frac{b^2}{a} \end{pmatrix}$  ist, muß auch  $\mathbf{B}$  negativ definit sein. Dies ist äquivalent zu

$$a < 0, \quad ad - b^2 > 0.$$

**Beispiel:** Sei  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  gegeben durch:

$$f(x, y) = 6xy - 3y^2 - 2x^3$$

Wir versuchen die Extremstellen zu ermitteln:

$$\begin{aligned} f_x &= 6y - 6x^2 \\ f_y &= 6x - 6y \\ \text{grad } f = 0 &\Leftrightarrow x = y \text{ und } x^2 = y \end{aligned}$$

Mögliche Kandidaten sind also  $(0, 0)$  und  $(1, 1)$ . Zur genauen Bestimmung berechnen wir die Hessematrix

$$\begin{aligned} H_f(x, y) &= \begin{pmatrix} -12x & 6 \\ 6 & -6 \end{pmatrix} \\ H_f(0, 0) &= \begin{pmatrix} 0 & 6 \\ 0 & -6 \end{pmatrix} && \text{indefinit} \\ H_f(1, 1) &= \begin{pmatrix} -12 & 6 \\ 6 & -6 \end{pmatrix} && \text{negativ definit} \end{aligned}$$

Damit liegt bei  $(0, 0)$  ein Sattelpunkt und bei  $(1, 1)$  ein Maximum vor.

**3.16 Methode der kleinsten Quadrate.** *Aus experimentellen Daten sind die Werte gewisser Konstanten  $x_1, \dots, x_n$  zu bestimmen. Da Messungen mit Fehlern behaftet sind, werden mehr Messdaten  $y_1, \dots, y_m$  ermittelt als nötig wären um die  $x_i$  zu bestimmen. Oft sind die experimentellen Daten  $y_1, \dots, y_m$  nur indirekt mit den Konstanten  $x_i$  verbunden. Also ergibt sich ein überbestimmtes nichtlineares Gleichungssystem*

$$y_i = f_i(x_1, \dots, x_n) \quad i = 1, \dots, m > n. \quad (3.17)$$

Um die Messfehler zu berücksichtigen wird eine Lösung  $\mathbf{x}$  gesucht so, dass  $y_i - f_i(\mathbf{x})$  minimal ist. Dazu wird die **Methode der kleinsten Quadrate** benutzt, d.h. finde  $\mathbf{x} \in \mathbb{R}^n$  mit

$$F(\mathbf{x}) := \sum_{i=1}^m (y_i - f_i(\mathbf{x}))^2 = \text{Min} ! \quad (3.18)$$

**3.19 Satz.** Sei  $\mathbf{A} \in \mathbb{R}^{m \times n}, m > n$ .

a) Das Minimalproblem

$$F(\mathbf{x}) = \|\mathbf{Ax} - \mathbf{y}\|^2 = \text{Min} ! \quad (3.20)$$

ist immer lösbar.

b) Lösungen von (3.20) und von

$$\mathbf{A}^T \mathbf{A} \mathbf{x} = \mathbf{A}^T \mathbf{y} \quad (3.21)$$

stimmen immer überein.

c) Für zwei Lösungen  $\mathbf{x}, \mathbf{x}_0$  von (3.20) gilt stets  $\mathbf{A} \mathbf{x} = \mathbf{A} \mathbf{x}_0$ . Ist  $\text{Rang } \mathbf{A} = n$ , so ist die Lösung von (3.20) eindeutig bestimmt.

- Satz 3.19 liefert eine praktische Methode zur Lösung von

$$\|\mathbf{A} \mathbf{x} - \mathbf{y}\|^2 = \text{Min} !$$

Man berechnet  $\mathbf{B} := \mathbf{A}^T \mathbf{A}$  und  $\mathbf{z} := \mathbf{A}^T \mathbf{y}$  und löst mittels Gaußelimination das System  $\mathbf{B} \mathbf{x} = \mathbf{z}$ , welches nach Satz 3.19 immer lösbar ist.

- Falls die Funktionen  $f_i$  in (3.18) nichtlinear sind kann (3.18) oft nur approximativ gelöst werden. Sukzessive Näherungslösungen werden mit Hilfe einer Variante des Gradientenverfahrens 3.2 berechnet. Sei  $\mathbf{x}_0$  ein Startwert. Dann wird anstelle von (3.18) das lineare Problem

$$\|\mathbf{y} - f(\mathbf{x}_0) - J_f(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0)\|^2 = \text{Min} ! \quad (*)$$

gelöst, wobei  $J_f = \begin{pmatrix} \text{grad } f_1(\mathbf{x}_0)^T \\ \vdots \\ \text{grad } f_m(\mathbf{x}_0)^T \end{pmatrix}$  die **Jakobimatrix** der vektorwertigen Funktion  $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$  ist.

Sei  $\mathbf{x}_1$  eine Lösung von (\*). Dann wird wie beim Gradientenverfahren durch  $\mathbf{x}_0 + h(\mathbf{x}_1 - \mathbf{x}_0)$  eine verbesserte Lösung konstruiert.

- In Extremwertaufgaben  $f(\mathbf{x}) = \text{Extr} !$  ist oft die Menge der zulässigen Punkte  $\mathbf{x}$  durch Nebenbedingungen der Form

$$g_1(\mathbf{x}) = 0, \dots, g_k(\mathbf{x}) = 0$$

eingeschränkt. Eine solche Aufgabe bezeichnen wir mit

$$f(\mathbf{x}) = \text{Extr} ! \quad \text{NB } g_i(\mathbf{x}) = 0 \quad i = 1, \dots, k.$$

**3.22 Definition.** Sei  $M = \{\mathbf{x} \in \mathbb{R}^n; g(\mathbf{x}) = 0\}$ . Man sagt,  $\mathbf{a} \in M$  ist ein **Maximum (Minimum) von  $f$  unter der Nebenbedingung  $g(\mathbf{x}) = 0$** , wenn es eine Umgebung  $B_r(\mathbf{a})$  gibt so, dass  $f(\mathbf{x}) \leq f(\mathbf{a})$ , ( $f(\mathbf{a}) \leq f(\mathbf{x})$ ) gilt für alle  $\mathbf{x} \in B_r(\mathbf{a}) \cap M$ .