# 24.6 Extremwertaufgaben

Auch für Funktionen $f : \mathbb{R}^n \to \mathbb{R}$ ist es ein wesentliches Anliegen, Maxima und Minima aufzufinden. Wie im Eindimensionalen nennt man einen Punkt $\boldsymbol{p}$ ein Maximum von $f$, wenn es eine Umgebung $U(\boldsymbol{p})$ gibt, sodass $f(\boldsymbol{x}) \leq f(\boldsymbol{p})$ für alle $\boldsymbol{p} \in U(\boldsymbol{p})$. Analog gilt für ein Minimum, dass in einer Umgebung $U(\boldsymbol{p})$ immer $f(\boldsymbol{x}) \geq f(\boldsymbol{p})$ ist. Wie gewohnt fassen wir Minima und Maxima unter dem Sammelbegriff *Extrema* zusammen.

## Kandidaten für Extremstellen lassen sich durch Nullsetzen des Gradienten finden

Wie im Eindimensionalen ist die Differenzialrechnung ein wesentliches Mittel, um Extremwerte von Funktionen zu ermitteln. Analog zum Fall $\mathbb{R} \to \mathbb{R}$ gibt es aber gewisse Einschränkungen. So werden mit Mitteln der Differenzialrechnung an Stellen, an denen die fragliche Funktion nicht differenzierbar ist, sicher keine Extremwerte gefunden; auch Randextrema werden so nicht berücksichtigt und müssen separat überprüft werden.

Klammern wir diese Probleme aber einmal vorläufig aus und betrachten eine am besten überall differenzierbare Funktion $f$: $\mathbb{R}^n \to \mathbb{R}$. Durch diese Funktion wird ja eine Fläche im $\mathbb{R}^{n+1}$ beschrieben. Überall dort, wo ein Extremum vorliegt, wird die Tangentialebene an die Fläche „waagrecht", also parallel zur $(x_1 - \ldots - x_n)$-Hyperebene sein, siehe Abb. 24.19. Damit das erfüllt ist, müssen alle partiellen Ableitungen $\frac{\partial f}{\partial x_i}$ verschwinden.

**Notwendige Bedingung für Extrema**

Ist eine Funktion $f$: $\mathbb{R}^n \to \mathbb{R}$ in $\boldsymbol{p}$ differenzierbar und hat dort ein relatives Extremum, so gilt

$$\frac{\partial f}{\partial x_i}(\boldsymbol{p}) = 0 \quad \text{für } i = 1, \ldots, n\,.$$

Es ist allerdings klar, dass nicht an allen *kritischen Punkten*, wo $\frac{\partial f}{\partial x_i} = 0$ für alle $i = 1, \ldots, n$ ist, auch wirklich ein Extremwert vorliegen muss. So ist etwa für $f(x_1, x_2) = x_1 x_2$ im Ursprung $\mathbf{0} = (0, 0)^\mathrm{T}$ ja $\frac{\partial f}{\partial x_1}(\mathbf{0}) = x_2|_{x_1=x_2=0} = 0$ und

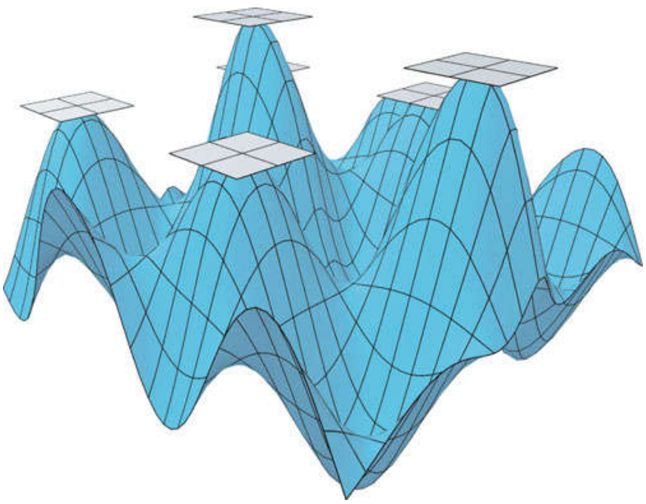

**Abb. 24.19** An Extremstellen ist die Tangentialebene parallel zur $x_1$-$x_2$-Ebene. Daher lassen sich Kandidaten für Extrema durch Nullsetzen des Gradienten finden

$\frac{\partial f}{\partial x_2}(\mathbf{0}) = x_1|_{x_1=x_2=0} = 0$. In jeder Umgebung von $\mathbf{0}$ gibt es aber größere und kleinere Funktionswerte als $f(0,0) = 0$. Es liegt, wie in Abb. 24.20 dargestellt, ein Sattelpunkt vor.

Man wird also noch zusätzliche Bedingungen brauchen, um abzuklären, ob an einem *kritischen Punkt* mit $\frac{\partial f}{\partial x_1} = \ldots = \frac{\partial f}{\partial x_n} = 0$ auch wirklich ein Extremum vorliegt, und wenn ja, welches. Ähnlich wie im Eindimensionalen kann dabei die Aussage oft mithilfe der zweiten Ableitungen getroffen werden.

Dabei berufen wir uns wieder auf den Satz von Taylor, diesmal in der mehrdimensionalen Fassung von Abschn. 24.3. Dieser stellt fest, dass eine zweimal differenzierbare Funktion an einem Extremum durch ein Paraboloid angenähert werden kann. Öffnet sich dieses Paraboloid nach unten, so liegt ein Maximum vor, öffnet es sich nach oben, so haben wir ein Minimum. Sind manche der Parabeln nach oben, andere nach unten geöffnet, so liegt kein Extremum vor.

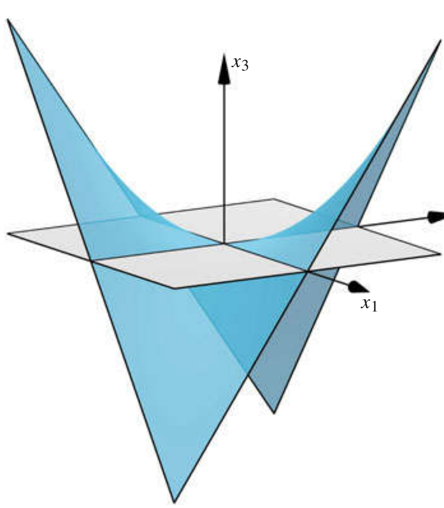


**Abb. 24.20** Für die Funktion $x_3 = x_1 x_2$ ist $\mathbf{grad}\, f|_{\mathbf{0}} = \mathbf{0}$, es liegt jedoch keine Extremstelle, sondern ein Sattelpunkt vor

Für die Approximation durch ein Paraboloid benötigen wir die zweiten Ableitungen, die in der **Hesse-Matrix** zusammengefasst werden,

$$\boldsymbol{H} = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{pmatrix}.$$

Im Fall $f \in C^2$ ist $\boldsymbol{H}$ nach dem Satz von Schwarz natürlich symmetrisch. Zu dieser Matrix gehört, wie im Kap. 21 ausführlich diskutiert, eine quadratische Form, eben genau das entsprechende Paraboloid. Die Form dieses Paraboloids hängt eng mit der Definitheit von $\boldsymbol{H}$ zusammen.

Mit den Ergebnissen aus Kap. 21, natürlich immer unter der Voraussetzung, dass $\boldsymbol{p}$ ein innerer Punkt von $D(f)$ und $\mathbf{grad}\, f(\boldsymbol{p}) = \mathbf{0}$ ist, folgt:

- Ist die Hesse-Matrix $\boldsymbol{H}(\boldsymbol{p})$ positiv definit, so hat $f$ an $\boldsymbol{p}$ ein relatives Minimum.
- Ist die Hesse-Matrix $\boldsymbol{H}(\boldsymbol{p})$ negativ definit, so hat $f$ an $\boldsymbol{p}$ ein relatives Maximum.
- Ist die Hesse-Matrix $\boldsymbol{H}(\boldsymbol{p})$ indefinit, so liegt kein Extremum vor.

Die Definitheit kann man immer anhand der Eigenwerte überprüfen. In vielen Fällen gibt es jedoch ein bequemeres Kriterium. Dazu definieren wir die Unterdeterminanten:

$$\Delta_1 = \frac{\partial^2 f}{\partial x_1^2}$$

$$\Delta_2 = \begin{vmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_2 \partial x_1} \\ \frac{\partial^2 f}{\partial x_1 \partial x_2} & \frac{\partial^2 f}{\partial x_2^2} \end{vmatrix}$$

$$\vdots$$

$$\Delta_n = \begin{vmatrix} \frac{\partial^2 f}{\partial x_1^2} & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_1} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_1 \partial x_n} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{vmatrix}$$

Die Hesse-Matrix ist positiv definit (und $f$ hat an $\boldsymbol{p}$ ein lokales Minimum), wenn alle $\Delta_i > 0$ sind. Sie ist negativ definit (und $f$ hat an $\boldsymbol{p}$ ein relatives Maximum), wenn die Vorzeichen gemäß

$$\Delta_1 < 0\,, \quad \Delta_2 > 0\,, \quad \Delta_3 < 0\,, \ldots$$

alternieren. Im Fall von $[2 \times 2]$-Matrizen gibt es zudem ein sehr einfaches Kriterium für die Indefinitheit von Matrizen: $\boldsymbol{H}|_{\boldsymbol{p}}$ ist indefinit (und $f$ hat an $\boldsymbol{p}$ kein Extremum), wenn $\Delta_2 < 0$ ist.

**Beispiel** Wir betrachten die Funktion $f\colon \mathbb{R}^2 \to \mathbb{R}$,

$$f(x_1, x_2) = \mathrm{e}^{-x_1^2/2} \cosh x_2 .$$

Gradientenbildung liefert:

$$\begin{aligned}
\frac{\partial f}{\partial x_1} &= -x_1\, \mathrm{e}^{-x_1^2/2} \cosh x_2 \\
\frac{\partial f}{\partial x_2} &= \mathrm{e}^{-x_1^2/2} \sinh x_2
\end{aligned}$$

Aus der ersten Gleichung sehen wir, dass ein kritischer Punkt $x_1 = 0$ erfüllen muss, aus der zweiten, dass entsprechend $x_2 = 0$ sein muss. Die zweiten Ableitungen erhalten wir zu

$$\begin{aligned}
\frac{\partial^2 f}{\partial x_1^2} &= (x_1^2 - 1)\, \mathrm{e}^{-x_1^2/2} \cosh x_2 \\
\frac{\partial^2 f}{\partial x_1\, \partial x_2} &= -x_1\, \mathrm{e}^{-x_1^2/2} \sinh x_2 \\
\frac{\partial^2 f}{\partial x_2^2} &= \mathrm{e}^{-x_1^2/2} \cosh x_2 .
\end{aligned}$$

Die Hesse-Matrix im kritischen Punkt ist damit

$$\boldsymbol{H}\Big|_{\boldsymbol{0}} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} .$$

Mit

$$\Delta_2 = \begin{vmatrix} -1 & 0 \\ 0 & 1 \end{vmatrix} = -1 < 0$$

sehen wir sofort, dass $\boldsymbol{H}$ indefinit ist und an $\boldsymbol{x} = \boldsymbol{0}$ ein Sattelpunkt liegt. ◀

Im Höherdimensionalen lässt sich leider kein ganz so einfaches Kriterium für die Indefinitheit angeben, hier kann man einer Diagonalisierung der Hesse-Matrix und einer Untersuchung der Eigenwerte meist nicht ausweichen.

Allerdings wird es oft vorkommen, dass sich mit der Hesse-Matrix keine Aussagen treffen lassen, nämlich dann, wenn Eigenwerte von $\boldsymbol{H}$ null sind. Dann muss man mehr oder weniger „trickreich" (und von Beispiel zu Beispiel verschieden) vorgehen, um kritische Punkte doch noch zu klassifizieren, wie etwa auf S. 906 gezeigt.

**Beispiel** Wir bestimmen die Extrema der Funktion $f\colon M \to \mathbb{R}$ $f(x, y) = x^3 + y^3 + 3xy$ auf $M = \mathbb{R}^2$. Dazu bilden wir die partiellen Ableitungen und setzen sie null:

$$\begin{aligned}
f_x(x, y) &= 3x^2 + 3y \stackrel{!}{=} 0 \quad x^2 + y = 0 \\
f_y(x, y) &= 3y^2 + 3x \stackrel{!}{=} 0 \quad y^2 + x = 0
\end{aligned}$$

Mit $y = -x^2$ aus der ersten Gleichung erhält man in der zweiten $x^4 + x = x\,(x^3 + 1) = 0$, also die beiden Lösungen $x = 0$ und $x = -1$ und mit $y = -x^2$ die beiden Punkte $\boldsymbol{p}_1 = (0, 0)^{\mathrm{T}}$ und $\boldsymbol{p}_2 = (-1, -1)$. Für die Determinante der Hesse-Matrix erhält man

$$\Delta_2 = \begin{vmatrix} 6x & 3 \\ 3 & 6y \end{vmatrix} = 36xy - 9,$$

also $\Delta_2|_{\boldsymbol{p}_1} = -9 < 0$ und $\Delta_2|_{p_2} = 27 > 0$. An $\boldsymbol{p}_1$ liegt ein Sattelpunkt, $\boldsymbol{p}_2$ ist wegen $f_{xx}|_{\boldsymbol{p}_2} = -6 < 0$ ein lokales Maximum. Globale Extrema auf $M$ besitzt diese Funktion nicht, da z. B. $f(x, 0) = x^3$ beliebig große und kleine Werte annimmt. ◀

Teil IV