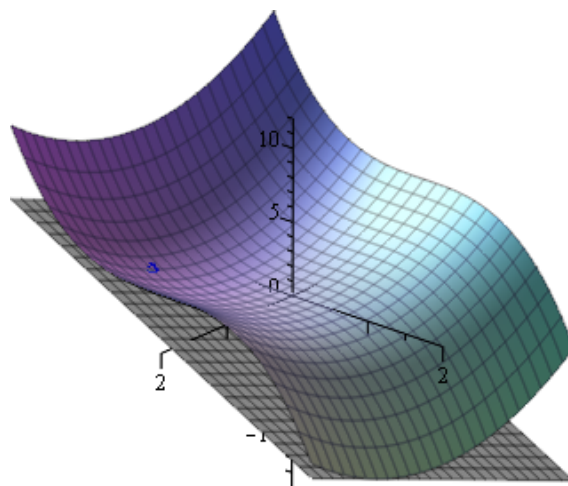


Abbildung 19:  $f(x, y) = x^2 - y^3$  mit Tangentialebene in  $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ 

### 3.4 Extremwerte

Bei Funktionen  $f : \mathbb{R} \rightarrow \mathbb{R}$  stellt man sich die Ableitung in einem Punkt gerne als die Steigung des Funktionsgraphen in diesem Punkt vor. Bei Funktionen mit mehrdimensionalen Bildern macht der Begriff der Steigung keinen Sinn mehr, da dort keine „größer“- oder „kleiner“-Relation existiert.

Auch bei Funktionen  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  (deren Graph eine Art „Berg- und Tallandschaft“ darstellt) können wir nicht von „der“ Steigung in einem Punkt sprechen, da die Steigung in einem Punkt von der Richtung abhängt, in der sie gemessen wird. Statt dessen interessiert man sich eher für die maximale Steigung in einem Punkt.

**3.4.1 Definition** Ist  $f : D \rightarrow \mathbb{R}$  mit  $D \subset \mathbb{R}^n$ , in  $\vec{a} \in D$  partiell differenzierbar, dann ist der Vektor der partiellen Ableitungen an der Stelle  $\vec{a} \in D$  der *Gradient* von  $f$  in  $\vec{a}$ :

$$\text{grad } f(\vec{a}) = \left( \frac{\partial}{\partial x_1} f(\vec{a}), \frac{\partial}{\partial x_2} f(\vec{a}), \dots, \frac{\partial}{\partial x_n} f(\vec{a}) \right)$$

**3.4.2 Bemerkung** Der Gradient entspricht der Ableitung von  $f$  bei Skalarwertigen Funktionen, die in diesem Fall eine  $(1, n)$ -Matrix ist. Es ist also  $\text{grad } f(\vec{a}) = f'(\vec{a})$ .

**3.4.3 Satz** Der Gradient von  $f$  zeigt in die Richtung des stärksten Anstiegs von  $f$ .

**3.4.4 Beispiele** (1)  $f(x, y) = x^2 - y^3$  in  $\vec{a} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$  (vgl. Abbildung 19).

Wir haben  $\text{grad } f(x, y) = (2x, -3y^2)$ , also  $\text{grad } f(1, -1) = (2, -3)$ . Im Punkt  $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$  ist der stärkste Anstieg also in Richtung  $\begin{pmatrix} 2 \\ -3 \end{pmatrix}$  zu finden.

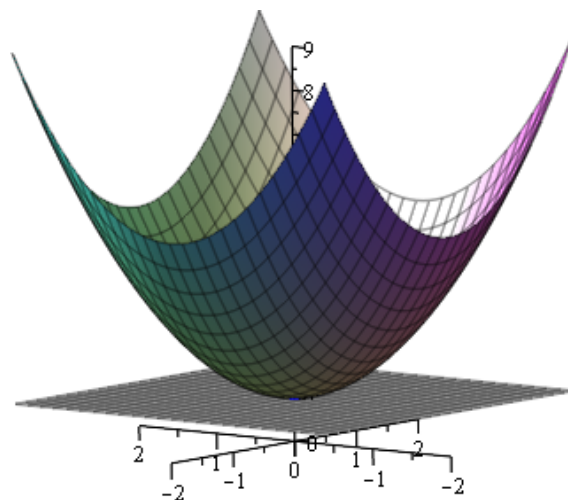
(2)  $f(x, y) = x^2 + y^2 + 1$ .

Der Gradient dieser Funktion ist  $\text{grad } f(x, y) = (2x, 2y)$ , also z.B.

$$\text{grad } f(0, 0) = \vec{0}.$$

In  $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$  haben hier also keine Richtung des stärksten Anstiegs!

Grund hierfür ist, dass die Funktion an der Stelle  $\vec{0}$  ein Minimum hat und die Tangentialebene in diesem Punkt daher waagerecht ist (vgl. Abbildung 20). Dies entspricht bei eindimensionalen Funktionen einer waagerechten Tangente mit Steigung 0.

Abbildung 20:  $f(x, y) = x^2 + y^2 + 1$  mit Tangentialebene in  $\vec{0}$ 

Bei Funktionen  $f : \mathbb{R} \rightarrow \mathbb{R}$  wissen wir, dass in einem lokalen Extremum stets  $f'(x) = 0$  gilt (das *notwendige Kriterium* für Extrema). Im Mehrdimensionalen ist die Situation ganz ähnlich:

**3.4.5 Satz (Notwendiges Kriterium für Extremwerte)** Eine Funktion  $f : D \rightarrow \mathbb{R}$  mit  $D \subset \mathbb{R}^n$  sei in  $\vec{a} \in D$  partiell differenzierbar. Hat  $f$  in  $\vec{a}$  ein lokales Extremum, dann gilt

$$\text{grad } f(\vec{a}) = \vec{0}.$$

**3.4.6 Bemerkung** Wie im Eindimensionalen können wir dies nutzen, um *mögliche* Stellen für Extrema zu finden. Es ist aber auch hier unklar, ob an diesen Stellen wirklich Extremwerte vorliegen. Betrachten wir beispielsweise die Sattelfläche  $f(x, y) = x^2 - y^2$  (Abbildung 12). Der Gradient ist  $\text{grad } f(x, y) = (2x, -2y)$ .

In  $\vec{0}$  ist also  $\text{grad } f(\vec{0}) = \vec{0}$ , das notwendige Kriterium ist somit erfüllt. Tatsächlich gibt es in  $\vec{0}$  aber keinen Extremwert!

Im Eindimensionalen konnten wir die Entscheidung, ob ein Extremum vorliegt, meistens treffen, indem wir das Vorzeichen der zweiten Ableitung prüften: ist  $f'(a) = 0$  und  $f''(a) > 0$ , haben wir ein Minimum in  $a$ , ist  $f'(a) = 0$  und  $f''(a) < 0$ , haben wir ein Maximum.

Im Mehrdimensionalen ist dies leider nicht mehr so einfach, da wir nicht *eine* zweite Ableitung haben, sondern mehrere partielle Ableitungen zweiter Ordnung:

**3.4.7 Definition** Sei  $f : D \rightarrow \mathbb{R}$ ,  $D \subset \mathbb{R}^n$ , eine zweimal partiell differenzierbare Funktion. Die  $(n, n)$ -Matrix der partiellen Ableitungen zweiter Ordnung in  $\vec{a} \in D$  heißt *Hesse-Matrix* von  $f$  in  $\vec{a}$

$$H_f(\vec{a}) = \begin{pmatrix} f_{x_1, x_1}(\vec{a}) & f_{x_1, x_2}(\vec{a}) & \cdots & f_{x_1, x_n}(\vec{a}) \\ f_{x_2, x_1}(\vec{a}) & f_{x_2, x_2}(\vec{a}) & \cdots & f_{x_2, x_n}(\vec{a}) \\ \vdots & \vdots & \ddots & \vdots \\ f_{x_n, x_1}(\vec{a}) & f_{x_n, x_2}(\vec{a}) & \cdots & f_{x_n, x_n}(\vec{a}) \end{pmatrix}.$$

**3.4.8 Bemerkung** Gibt es alle partiellen Ableitungen zweiter Ordnung und sind diese stetig, dann kann nach Satz 3.3.3 die Reihenfolge der Differenziation vertauscht werden:  $f_{x_i, x_k}(\vec{a}) = f_{x_k, x_i}(\vec{a})$ .  $H_f(\vec{a})$  ist in diesem Fall „symmetrisch“.

**3.4.9 Beispiel** Betrachten wir  $f(x, y) = x^3 + y^3$ .

Es ist  $f_x(x, y) = 3x^2$  und  $f_y(x, y) = 3y^2$  und

$$\begin{aligned} f_{xx}(x, y) &= 6x & f_{xy}(x, y) &= 0 \\ f_{yx}(x, y) &= 0 & f_{yy}(x, y) &= 6y \end{aligned}$$

und die Hesse-Matrix lautet daher

$$H_f(x, y) = \begin{pmatrix} 6x & 0 \\ 0 & 6y \end{pmatrix}.$$

Beispielsweise ist also  $H_f(1, 2) = \begin{pmatrix} 6 & 0 \\ 0 & 12 \end{pmatrix}$ .

Die Hesse-Matrix wird nun also die Rolle der zweiten Ableitung übernehmen, wir benötigen aber noch ein Kriterium, das statt  $f''(x) > 0$  bzw.  $f''(x) < 0$  verwendet werden kann, da wir bei einer Matrix natürlich nicht einfach nach dem Vorzeichen fragen können:

**3.4.10 Definition** Eine symmetrische Matrix  $A$  habe die Eigenwerte  $\lambda_1, \dots, \lambda_n$ .

$A$  ist *positiv definit*, wenn alle  $\lambda_i > 0$  sind,

$A$  ist *negativ definit*, wenn alle  $\lambda_i < 0$  sind,

$A$  ist *positiv semidefinit*, wenn alle  $\lambda_i \geq 0$  sind,

$A$  ist *negativ semidefinit*, wenn alle  $\lambda_i \leq 0$  sind und

$A$  ist *indefinit*, wenn es positive und negative Eigenwerte gibt.

Damit haben wir nun alle Werkzeuge zusammen, um die Extrema einer Funktion finden zu können:

**3.4.11 Satz (Hinreichendes Kriterium für Extremwerte)** Sei  $f : D \rightarrow \mathbb{R}$ ,  $D \subset \mathbb{R}^n$ , eine Funktion mit stetigen partiellen Ableitungen erster und zweiter Ordnung.

In  $\vec{a} \in D$  gelte  $\text{grad } f(\vec{a}) = \vec{0}$ .  $f$  hat dann in  $\vec{a}$

... ein Maximum, wenn  $H_f(\vec{a})$  negativ definit ist,

... ein Minimum, wenn  $H_f(\vec{a})$  positiv definit ist und

... kein lokales Extremum, wenn  $H_f(\vec{a})$  indefinit ist.

Ist  $H_f(\vec{a})$  semidefinit, können wir keine Aussage über ein Extremum in diesem Punkt machen.

**3.4.12 Beispiele (1)**  $f(x, y) = x^2 + y^2 + 1$  (vgl. Abbildung 20)

Es ist  $f_x(x, y) = 2x$  und  $f_y(x, y) = 2y$ , es ist also

$$\text{grad } f(x, y) = (2x, 2y) = (0, 0) \Leftrightarrow \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

An der Stelle  $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$  ist also ein Extremum möglich.

Die Hesse-Matrix von  $f$  lautet

$$H_f(\vec{a}) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \quad \text{für alle } \vec{a}.$$

Bei einer solchen Dreiecksmatrix sind die Eigenwerte nach Satz 2.11.6 besonders leicht zu ermitteln: sie stehen auf der Diagonalen von links oben nach rechts unten. In diesem Fall gibt es also nur einen Eigenwert, nämlich 2.  $H_f(\vec{0})$  ist folglich positiv definit und  $f$  hat daher ein Minimum in  $\vec{0}$ .