Rolf Haftmann

Gradient, Hessematrix, Definitheit, Taylorentwicklung und Extremwertaufgaben von Funktionen mehrerer Variabler

Gradient: Spaltenvektor der partiellen Ableitungen

$$\nabla f = \operatorname{grad} f = \begin{pmatrix} \dfrac{\partial f}{\partial x_1} \\ \dfrac{\partial f}{\partial x_2} \\ \cdots \\ \dfrac{\partial f}{\partial x_n} \end{pmatrix} = \begin{pmatrix} f_{x_1} \\ f_{x_2} \\ \cdot\cdot \\ f_{x_n} \end{pmatrix}$$

Hessematrix: Matrix der zweiten partiellen Ableitungen

$$\boldsymbol{H}_f = \begin{pmatrix} \dfrac{\partial^2 f}{\partial x_1^2} & \dfrac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \dfrac{\partial^2 f}{\partial x_1 \partial x_n} \\ \dfrac{\partial^2 f}{\partial x_2 \partial x_1} & \dfrac{\partial^2 f}{\partial x_2^2} & \cdots & \dfrac{\partial^2 f}{\partial x_2 \partial x_n} \\ \cdots & \cdots & \cdots & \cdots \\ \dfrac{\partial^2 f}{\partial x_n \partial x_1} & \dfrac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \dfrac{\partial^2 f}{\partial x_n^2} \end{pmatrix} = \begin{pmatrix} f_{x_1 x_1} & f_{x_1 x_2} & \cdots & f_{x_1 x_n} \\ f_{x_2 x_1} & f_{x_2 x_2} & \cdots & f_{x_2 x_n} \\ \cdots & \cdots & \cdots & \cdots \\ f_{x_n x_1} & f_{x_n x_2} & \cdots & f_{x_n x_n} \end{pmatrix}$$

Falls die zweiten partiellen Ableitungen stetig sind, so sind die gemischten partiellen Ableitungen gleich: $f_{x_i x_j} = f_{x_j x_i}$ (Satz von Schwarz). Die Hessematrix ist dann symmetrisch.

Definitheit symmetrischer Matrizen

Für Extremwertuntersuchungen von Funktionen mehrerer Veränderlicher wird der Definitheitsbegriff benötigt:

Eine symmetrische Matrix $\boldsymbol{A}$ heißt **positiv definit**, wenn $\vec{r}^{\mathsf{T}}\boldsymbol{A}\vec{r}>0$ für alle $\vec{r}\neq\vec{0}$ ist, **negativ definit**, wenn $\vec{r}^{\mathsf{T}}\boldsymbol{A}\vec{r}<0$ für alle $\vec{r}\neq\vec{0}$ ist, und **indefinit**, wenn es sowohl Vektoren mit $\vec{r}^{\mathsf{T}}\boldsymbol{A}\vec{r}>0$ als auch solche mit $\vec{r}^{\mathsf{T}}\boldsymbol{A}\vec{r}<0$ gibt.

Es gibt allerdings Matrizen, die weder positiv oder negativ definit noch indefint sind. Das ist dann der Fall, wenn $\vec{r}^{\mathsf{T}}\boldsymbol{A}\vec{r}$ für manche Richtungen $\vec{r}\neq\vec{0}$ gleich 0 wird, ansonsten aber einheitliches Vorzeichen hat.

Ob eine symmetrische Matrix definit ist, kann mit Hilfe ihrer **Hauptminoren** bestimmt werden, das sind die Determinanten aller der Matrizen, die entstehen, wenn man immer ausgehend von der linken oberen Ecke quadratische Matrizen bildet: $\begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} & \cdots \\ a_{12} & a_{22} & a_{23} & a_{24} & \cdots \\ a_{13} & a_{23} & a_{33} & a_{34} & \cdots \\ a_{14} & a_{24} & a_{34} & a_{44} & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{pmatrix}$.

Bei dieser Darstellung wurde berücksichtigt, dass wegen der Symmetrie $a_{12}=a_{21}$ usw. gilt. Hauptminoren sind also a_{11}, $\begin{vmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{vmatrix}$, $\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{vmatrix}$ usw.

Kriterium von Sylvester: Eine symmetrische Matrix ist genau dann positiv definit, wenn alle Hauptminoren positiv sind. Sie ist genau dann negativ definit, wenn alle Hauptminoren ungerader Ordnung negativ und alle Hauptminoren gerader Ordnung positiv sind. Sie ist indefinit, wenn ein Hauptminor gerader Ordnung negativ ist und/oder es Hauptminoren ungerader Ordnung mit entgegengesetztem Vorzeichen gibt. (Bei der Indefinitheit steht bewusst nicht „genau dann“, da es auch indefinite Matrizen gibt, für die diese Bedingung nicht erfüllt ist.)

Eine einreihige Matrix (Zahl) ist positiv bzw. negativ definit genau dann, wenn sie positiv bzw. negativ ist. Sie kann nicht indefinit sein. Ist sie gleich 0, so ist sie weder definit noch indefinit.

Eine zweireihige Matrix $\boldsymbol{A}$ ist genau dann positiv oder negativ definit, wenn ihr zweiter Hauptminor, d.h. $\det\boldsymbol{A}$ selbst positiv ist. Ist dies der Fall, so ist sie positiv bzw. negativ definit je nachdem, ob ihr erster Hauptminor, d.h. das Element a_{11}, positiv bzw. negativ ist. (Wegen $\det\boldsymbol{A}=a_{11}a_{22}-a_{12}^2$ kann a_{11} im Falle $\det\boldsymbol{A}>0$ nicht gleich 0 sein.) Weiter ist die zweireihige Matrix $\boldsymbol{A}$ indefinit (im zweireihigen Fall sogar genau dann), wenn $\det\boldsymbol{A}$ negativ ist. Ist $\det\boldsymbol{A}=0$, so ist sie weder definit noch indefinit.

Funktionen einer Veränderlichen

Taylorentwicklung (unter Differenzierbarkeitsvoraussetzungen):

$$\boxed{f(x)=f(x_0)+f'(x_0)\,(x-x_0)+\frac{1}{2}\,f''(x_0)\,(x-x_0)^2+\frac{1}{6}\,f'''(x_0)\,(x-x_0)^3+\ldots}$$

In der Umgebung von x_0 gilt $f(x)\approx T_1(x)=f(x_0)+f'(x_0)\,(x-x_0)$.

Der Faktor $(x-x_0)$ wechselt an der Stelle $x=x_0$ sein Vorzeichen. Da dieser Faktor noch mit $f'(x_0)$ multipliziert wird, ist $f(x)$ in x_0 für $f'(x_0)>0$ streng monoton wachsend und für $f'(x_0)<0$ streng monoton fallend, ein Extremwert kann also nur dann vorliegen, wenn $f'(x_0)=0$ ist.

Notwendige Extremwertbedingung: $f'(x_0)=0$ (Tangente waagerecht).

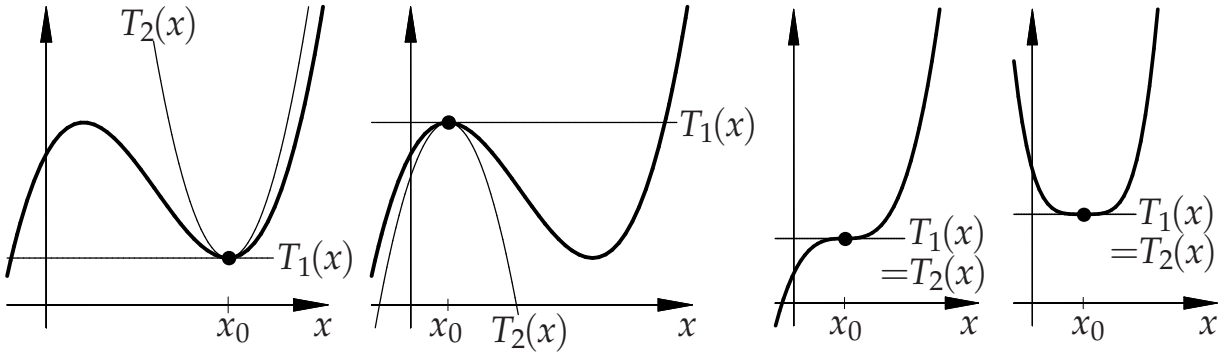
Ist dies der Fall, so gilt $f(x)\approx T_2(x)=f(x_0)+\frac{1}{2}\,f''(x_0)\,(x-x_0)^2$.

Der Faktor $(x-x_0)^2$ ist für $x\neq x_0$ positiv. Da dieser Faktor noch mit $f''(x_0)$ multipliziert wird, hat $f(x)$ in x_0 für $f''(x_0)>0$ ein Minimum und für $f''(x_0)<0$ ein Maximum. Ist auch $f''(x_0)=0$, so hängt das Monotonie- und Extremwertverhalten von höheren Gliedern der Taylorentwicklung ab.

Zusammengefasst lauten die vorgestellten **Extremwertbedingungen**:

Für Extremum notwendig: $f'(x_0)=0$, wenn ja (stationärer, d.h. extremwertverdächtiger Punkt),

$$\text{für}\begin{cases} f''(x_0)>0 & \text{lokales Minimum,}\\ f''(x_0)<0 & \text{lokales Maximum,}\\ f''(x_0)=0 & \text{so nicht entscheidbar.}\end{cases}$$



$f''(x_0)>0$: Minimum $\quad$ $f''(x_0)<0$: Maximum $\quad$ $f''(x_0)=0$: kein Extremwert bzw. Extremwert

Funktionen mehrerer Veränderlicher

Taylorentwicklung (unter Differenzierbarkeitsvoraussetzungen) für zwei Veränderliche:

$$\begin{aligned} f(x,y) &= f(x_0,y_0) + f_x(x_0,y_0)\,(x-x_0) + f_y(x_0,y_0)\,(y-y_0) \\ &\quad + \frac{1}{2}\Big(f_{xx}(x_0,y_0)\,(x-x_0)^2 + f_{xy}(x_0,y_0)\,(x-x_0)(y-y_0) + f_{yy}(x_0,y_0)\,(y-y_0)^2 \Big) + \dots \\ &= f(x_0,y_0) + \begin{pmatrix} f_x(x_0,y_0) \\ f_y(x_0,y_0) \end{pmatrix} \cdot \begin{pmatrix} x-x_0 \\ y-y_0 \end{pmatrix} + \frac{1}{2}(x-x_0 \quad y-y_0) \begin{pmatrix} f_{xx}(x_0,y_0) & f_{xy}(x_0,y_0) \\ f_{yx}(x_0,y_0) & f_{yy}(x_0,y_0) \end{pmatrix} \begin{pmatrix} x-x_0 \\ y-y_0 \end{pmatrix} + \dots \end{aligned}$$

Mit $\vec{x} = \begin{pmatrix} x \\ y \end{pmatrix}$, $\vec{x}_0 = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$, $\vec{x}-\vec{x}_0 = \begin{pmatrix} x-x_0 \\ y-y_0 \end{pmatrix}$ kann man auch schreiben

$$\boxed{f(\vec{x}) = f(\vec{x}_0) + \nabla f(\vec{x}_0) \cdot (\vec{x}-\vec{x}_0) + \frac{1}{2}(\vec{x}-\vec{x}_0)^{\mathsf{T}} \boldsymbol{H}_f(\vec{x}_0)\,(\vec{x}-\vec{x}_0) + \dots}$$

Dabei steht der $\cdot$ beim Gradienten jeweils für das Skalarprodukt, während der Term mit der Hessematrix im Sinne der Matrizenmultiplikation zu verstehen ist. Man beachte die Analogie zu der Darstellung für eine Veränderliche: $f(x) = f(x_0) + f'(x_0)\cdot(x-x_0) + \frac{1}{2}(x-x_0)\,f''(x_0)(x-x_0) + \dots$.

In der Umgebung von $\vec{x}_0$ gilt $f(\vec{x}) \approx T_1(\vec{x}) = f(\vec{x}_0) + \nabla f(\vec{x}_0) \cdot (x-x_0)$.

Der Term $\nabla f(\vec{x}_0) \cdot (x-x_0) = f_x(x_0,y_0)\,(x-x_0) + f_y(x_0,y_0)\,(y-y_0)$ kann in der Umgebung von $\vec{x}_0$ sowohl positiv als auch negativ werden, sobald eine der Komponenten des Gradienten ungleich 0 ist. Ein Extremwert kann also nur dann vorliegen, wenn alle partiellen Ableitungen gleich 0 sind.

Notwendige Extremwertbedingung: $\nabla f(\vec{x}_0) = \vec{0}$ (Tangentialebene waagerecht).

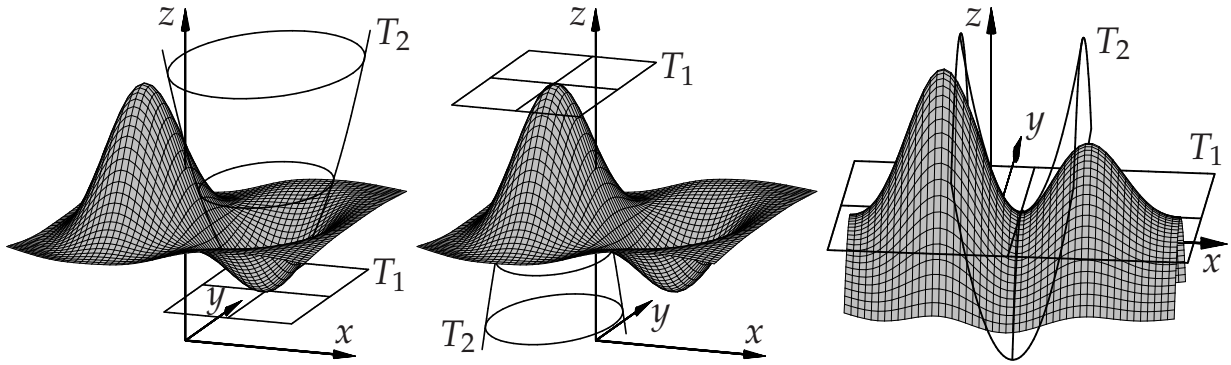
Ist dies der Fall, so gilt $f(\vec{x}) \approx T_2(\vec{x}) = f(\vec{x}_0) + \dfrac{1}{2}(\vec{x}-\vec{x}_0)^{\mathsf{T}} \boldsymbol{H}_f(\vec{x}_0)\,(\vec{x}-\vec{x}_0)$.

Ist der Term $(\vec{x}-\vec{x}_0)^{\mathsf{T}} \boldsymbol{H}_f(\vec{x}_0)\,(\vec{x}-\vec{x}_0)$ für alle $\vec{x} \neq \vec{x}_0$ positiv, so liegt ein Minimum vor, ist er für alle $\vec{x} \neq \vec{x}_0$ negativ, so liegt ein Maximum vor. Nimmt der Term sowohl positive als auch negative Werte an, so hat die Funktion $f(\vec{x})$ in $\vec{x}_0$ aus manchen Richtungen ein Minimum und aus manchen Richtungen ein Maximum, so dass dort insgesamt kein Extremwert, sondern ein Sattelpunkt vorliegt. Ist keine der drei genannten Alternativen richtig, was passiert, wenn der Term auch für manche $\vec{x} \neq \vec{x}_0$ gleich 0 ist, ansonsten aber einheitliches Vorzeichen hat, so hängt das Monotonie- und Extremwertverhalten von höheren Gliedern der Taylorentwicklung ab.

Zusammengefasst lauten die **Extremwertbedingungen** unter Verwendung des Definitheitsbegriffs:

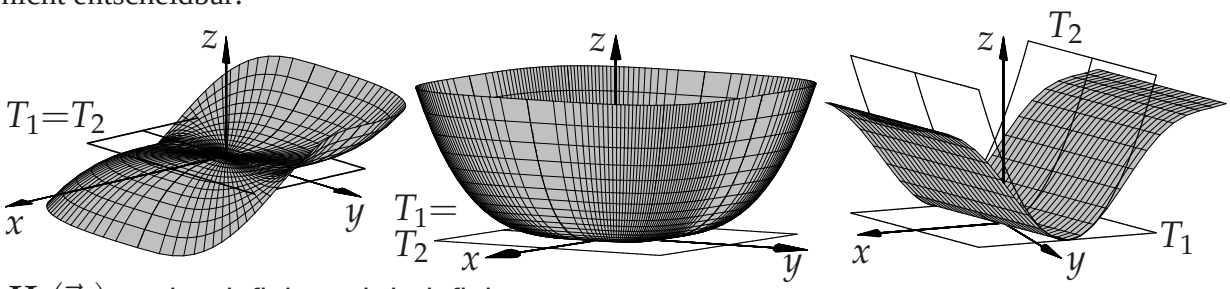
Für Extremum notwendig: $\nabla f(\vec{x}_0) = \vec{0}$, wenn ja (stationärer, d.h. extremwertverdächtiger Punkt),

$$\text{für} \begin{cases} \boldsymbol{H}_f(\vec{x}_0) \text{ positiv definit} & \text{lokales Minimum,} \\ \boldsymbol{H}_f(\vec{x}_0) \text{ negativ definit} & \text{lokales Maximum,} \\ \boldsymbol{H}_f(\vec{x}_0) \text{ indefinit} & \text{kein Extremwert („Sattelpunkt“),} \\ \boldsymbol{H}_f(\vec{x}_0) \text{ weder definit noch indefinit} & \text{so nicht entscheidbar.} \end{cases}$$



$\boldsymbol{H}_f(\vec{x}_0)$ pos. def.: Minimum $\quad$ $\boldsymbol{H}_f(\vec{x}_0)$ neg. def.: Maximum $\quad$ $\boldsymbol{H}_f(\vec{x}_0)$ indefinit: Sattelpunkt

Ist die Hessematrix weder definit noch indefinit, so ist die Extremwerteigenschaft mit ihrer Hilfe nicht entscheidbar:



$\boldsymbol{H}_f(\vec{x}_0)$ weder definit noch indefinit:

kein Extremwert, Extremwert bzw. nichtstrenger Extremwert (in der Umgebung Punkte mit gleichem Funktionswert)

Für Funktionen einer Veränderlichen entspricht dies dem Fall $f''(x_0)=0$.

Die Darstellung der Taylorentwicklung in der Form

$$\boxed{f(\vec{x}) = f(\vec{x}_0) + \nabla f(\vec{x}_0)\cdot(\vec{x}-\vec{x}_0) + \frac{1}{2}(\vec{x}-\vec{x}_0)^{\mathsf{T}}\boldsymbol{H}_f(\vec{x}_0)\,(\vec{x}-\vec{x}_0) + \ldots}$$

und die angegebenen Extremwertbedingungen gelten auch für beliebig viele Veränderliche.

Für zwei Veränderliche können die Extremwertbedingungen wegen des Definitheitskriteriums für zweireihige Matrizen speziell in folgender Form notiert werden:

Für Extremum notwendig: $\nabla f(\vec{x}_0)=\vec{0}$, wenn ja (stationärer Punkt),

$$\text{für}\begin{cases} \det\boldsymbol{H}_f>0 & \text{lokales Extremum: für } \begin{cases} f_{xx}>0 & \text{Minimum,}\\ f_{xx}<0 & \text{Maximum,}\end{cases}\\ \det\boldsymbol{H}_f<0 & \text{Sattelpunkt,}\\ \det\boldsymbol{H}_f=0 & \text{so nicht entscheidbar.}\end{cases}$$

Für eine Veränderliche reduzieren sie sich auf die bekannten Bedingungen für $f'(x_0)$ und $f''(x_0)$. Der Fall der Indefinitheit kann allerdings bei Funktionen einer Veränderlichen nicht auftreten. Da es nur eine Richtung gibt, kann es für Funktionen einer Veränderlichen keine Sattelpunkte geben, bei denen die Funktion aus manchen Richtungen ein Maximum und aus manchen ein Minimum hat.

Bilder aus:
Haftmann, R.: EAGLE-GUIDE Differenzialrechnung. Vom Ein- zum Mehrdimensionalen.
Leipzig: Edition am Gutenbergplatz 2009