Die Tangentialebene

Der Graph der linearen Approximation ist Tangentialebene an den Graph der Funktion. In Symbolen: Es sei $D \subset \mathbb{R}^2$. Es sei $f : D \to \mathbb{R}$, $(x, y) \mapsto f(x, y)$ differenzierbar. Dann ist

$$\{(x, y, z) \mid z = f(x_0, y_0) + \frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0)\}$$

Tangentialebene an den Graphen $\{(x, y, z) \mid z = f(x, y)\}$ im Punkt $(x_0, y_0, f(x_0, y_0))$.

Die Kettenregel

Es seien $f : D \subset \mathbb{R}^2 \to \mathbb{R}$, $(x,y) \mapsto f(x,y)$, $g_1 : D_1 \subset \mathbb{R} \to \mathbb{R}$, und $D_2 \subset \mathbb{R} \to \mathbb{R}$ differenzierbar. Dann ist $h := f \circ (g_1, g_2) : t \mapsto f(g_1(t), g_2(t))$ differenzierbar in ihrem Definitionsbereich, und es gilt

$$h'(t) = \frac{\partial f}{\partial x}(g_1(t), g_2(t))g_1'(t) + \frac{\partial f}{\partial y}(g_1(t), g_2(t))g_2'(t)$$

Beispiel: die Quotientenregel

Wir setzen $F(x, y) = \frac{x}{y}$. Die partiellen Ableitungen sind

$$\frac{\partial f}{\partial x} = \frac{1}{y}, \frac{\partial f}{\partial y} = -\frac{x}{y^2}.$$

Nach der Kettenregel gilt

$$\frac{d\frac{g_1(t)}{g_2(t)}}{dt} = \frac{1}{g_2(t)} g_1'(t) - \frac{g_1(t)}{g_2(t)^2} g_2'(t)$$

Die Richtungsableitung

Es sei $f : D \subset \mathbb{R}^2 \to \mathbb{R}$ differenzierbar. Es sei (x, y) ein innerer Punkt von D. Es sei $v = \begin{bmatrix} p \\ q \end{bmatrix} \in \mathbb{R}^2$ ein Vektor der Länge 1, i.e. $p^2 + q^2 = 1$. Die Richtungsableitung von f bei (x, y) in Richtung v ist definiert als die Ableitung der Funktion $t \mapsto f(x+tp, y+tq)$ an der Stelle $t = 0$.

Die Richtungsableitung

Nach der Kettenregel ist die Ableitung gleich

$$\frac{\partial f}{\partial x}(x+tp, y+tq)p + \frac{\partial f}{\partial y}(x+tp, y+tq)q,$$

und die Richtungsableitung ist

$$\partial_v f = \frac{\partial f}{\partial x}p + \frac{\partial f}{\partial y}q.$$

Der Gradient

Den Vektor $\begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix}$ nennen wir Gradient von f bei (x, y), geschrieben $\nabla f(x, y)$. Damit läßt sich die Richtungsableitung als Skalarprodukt

$$\partial_v f = \langle \nabla f, v \rangle$$

Der Gradient ist Normalvektor der Niveaulinie durch (x, y). Wenn v tangential zur Niveaulinie ist, ist die Richtungsableitung 0. Die größte Richtungsableitung ist die in Richtung des Gradienten; diese Beobachtung wird verwendet bei numerischen Verfahren zur Optimierung.

Optimierung

Wenn f an der Stelle (x, y) ein lokales Maximum/Minimum hat, dann muß der Gradient notwendigerweise 0 sein. Die Umkehrung gilt nicht, wie das Beispiel $f(x, y) = xy$ zeigt.

Für Funktionen in einer Variablen ist

$$f'(x) = 0, f''(x) < 0$$

eine notwendige Bedingung für ein Maximum. Ein ähnliches Kriterium für lokale Extreme in zwei Variablen wäre nützlich!

Optimierung

Vielleicht:

$$\frac{\partial f(x,y)}{\partial x} = \frac{\partial f(x,y)}{\partial y} = 0, \frac{\partial^2 f(x,y)}{\partial x^2} < 0, \frac{\partial^2 f(x,y)}{\partial y^2} < 0?$$

Dann hätten die partiellen Funktionen auf jeden Fall ein Maximum.

Optimierung

Gegenbeispiel: $f(x,y) = x^2 + y^2 - 3xy$.

Man muß auch die gemischte partielle Ableitung $\frac{\partial^2 f(x,y)}{\partial x \partial y} = \frac{\partial^2 f(x,y)}{\partial y \partial x}$ verwenden.

Gleichheit gilt nach dem Satz von Schwarz, unter der Voraussetzung daß alle zweiten Ableitungen stetig sind.

Optimierung für quadratische Funktionen

Es sei $f(x,y) = ax^2 + bxy + cy^2$. Dann gilt

$$\frac{\partial f}{\partial x}(0,0) = \frac{\partial f}{\partial y}(0,0) = 0,$$

$$\frac{\partial^2 f}{\partial x^2}(0,0) = 2a, \frac{\partial^2 f}{\partial x \partial y}(0,0) = b, \frac{\partial^2 f}{\partial y^2}(0,0) = 2c.$$

Optimierung für quadratische Funktionen

Der Nullpunkt ist ein striktes Minimum, wenn $f(x,y) > 0$ ist für $(x,y) \neq (0,0)$. Das geht nur für $a > 0$. Unter Verwendung von $a > 0$ folgt

$$0 < 4a(ax^2 + bxy + cy^2) = (2ax + by)^2 + (4ac - b^2)y^2,$$

und das gilt für alle $(x,y) \neq (0,0)$ dann und nur dann wenn $4ac - b^2 > 0$ ist. Dieser Ausdruck ist gleich der Determinante der Matrix

$$\begin{bmatrix} 2a & b \\ b & 2c \end{bmatrix} = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2}(0,0) & \frac{\partial^2 f}{\partial x \partial y}(0,0) \\ \frac{\partial^2 f}{\partial x \partial y}(0,0) & \frac{\partial^2 f}{\partial y^2}(0,0) \end{bmatrix}$$

Die Hessesche Matrix

Es sei $f : D \subset \mathbb{R}^2$ zweimal stetig differenzierbar, und (x, y) ein innerer Punkt von D. Die Hessesche Matrix von f bei (x, y) ist definiert als

$$H_f(x,y) = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2}(0,0) & \frac{\partial^2 f}{\partial x \partial y}(0,0) \\ \frac{\partial^2 f}{\partial x \partial y}(0,0) & \frac{\partial^2 f}{\partial y^2}(0,0) \end{bmatrix}$$

Die Hessesche Matrix

Die quadratische Approximation (auch: das Taylorpolynom zweiter Ordnung) von f bei (x, y) ist definiert als

$$f(x+s, y+t) \approx f(x, y) + \frac{\partial f(x, y)}{\partial x} s + \frac{\partial f(x, y)}{\partial y} t$$

$$+ \frac{1}{2} \frac{\partial^2 f(x, y)}{\partial x^2} s^2 + \frac{\partial^2 f(x, y)}{\partial x \partial y} st + \frac{1}{2} \frac{\partial^2 f(x, y)}{\partial y^2} t^2$$

$$= f(x, y) + f'(x, y) \begin{bmatrix} s \\ t \end{bmatrix} + \frac{1}{2} \left\langle \begin{bmatrix} s \\ t \end{bmatrix}, H_f \begin{bmatrix} s \\ t \end{bmatrix} \right\rangle .$$

Ein hinreichendes Kriterium für Minima

Wenn $\nabla f(x, y) = \vec{0}$ und $\frac{\partial^2 f(x,y)}{\partial x^2} > 0$ und $\frac{\partial^2 f(x,y)}{\partial y^2} > 0$ und $\det(H_f(x, y)) > 0$, dann hat f bei (x, y) ein lokales Minimum.

Die Bedingung an die zweiten partiellen Ableitungen ist equivalent dazu, daß $H_f(x, y)$ positiv definit ist.

Der Beweis geht ähnlich wie im univariaten Fall. Die wesentliche Idee ist, daß das Vorzeichen von $f(u, v) - f(x, y)$ nur von der quadratischen Approximation abhängt, wenn die Determinante der Hesseschen Matrix ungleich Null ist.

Ein hinreichendes Kriterium für Maxima

Wenn $\nabla f(x,y) = \vec{0}$ und $\frac{\partial^2 f(x,y)}{\partial x^2} < 0$ und $\frac{\partial^2 f(x,y)}{\partial y^2} < 0$ und $\det(H_f(x,y)) > 0$, dann hat f bei (x,y) ein lokales Maximum.

Die Bedingung an die zweiten partiellen Ableitungen ist equivalent dazu, daß $H_f(x,y)$ negativ definit ist.

Sattelpunkte

Wenn $\nabla f(x, y) = \vec{0}$ und $\det(H_f(x, y)) < 0$ gilt, liegt ein Sattelpunkt vor, und kein lokales Optimum.

Wenn $\det(H_f(x, y)) = 0$ ist, dann reicht die quadratische Approximation nicht aus, um festzustellen, ob ein lokales Optimum vorliegt.