13 Differentiation in mehreren Variablen

Satz: Gilt für $f : [a, b] \to \mathbb{R}$, dass $f'(x) \geq 0 \, \forall x \in [a, b]$ (kurz: $f' \geq 0$), so ist f monoton wachsend. Entsprechend $f' > 0 \Rightarrow f$ streng monoton wachsend; $f' \leq 0 \Rightarrow f$ monoton fallend; $f' < 0 \Rightarrow f$ streng monoton fallend.

Von Differentiation und Integration in mehreren Variablen werde ich nicht die exakte mathematische Definition angeben, weil das zu lange dauern würde; ich werde aber erklären, wie diese Operationen zu verstehen und handzuhaben sind.

Partielle Ableitungen. Halten wir in der Funktion $f(x, y)$, d.h. $f : \mathbb{R}^2 \to \mathbb{R}$, die Variable y fest, so erhalten wir eine Fkt $g(x) = f(x, y)$ von einer Variablen. Dies entspricht dem Einsetzen eines Zahlenwertes für y, aber oftmals würde man diesen Wert weiter einfach y nennen. (Man nennt eine solche Größe einen *Parameter*: er tritt in einer Funktion auf, seinen Wert verändern wir aber nicht; wir betrachten den Parameter nicht als Argument der Funktion; z.B. in der Funktion $h(x) = e^{-\lambda x}$ ist λ ein Parameter.) Die Ableitung von g bezeichnet man als *partielle Ableitung von f nach x* und schreibt

$$g' = \frac{\partial f}{\partial x}.$$

Wir können nun $\partial f / \partial x$ wieder als Fkt von x und y betrachten, $\partial f / \partial x : \mathbb{R}^2 \to \mathbb{R}$. Entsprechend bezeichnet man für jeden festen Wert von x die Ableitung von $h(y) = f(x, y)$ als partielle Ableitung von f nach y und schreibt

$$h' = \frac{\partial f}{\partial y}.$$

Bsp:

$$f(s,t) = se^t + \sin(st)\,, \quad \frac{\partial f}{\partial s} = e^t + t\cos(st)\,, \quad \frac{\partial f}{\partial t} = se^t + s\cos(st)\,.$$

Die partiellen Ableitungen lassen sich als *Richtungsableitungen* veranschaulichen: Graph von f als “Gebirge”, für festes y bewegt man sich nur in x-Richtung, Steigung $= \partial f / \partial x$. Die Richtungsableitung in Richtung des Vektors $e \in \mathbb{R}^2$ mit $\|e\| = 1$ im Punkt $u \in \mathbb{R}^2$ lässt sich definieren als die Ableitung der Funktion $g(s) = f(u + se)$. Man findet dafür manchmal die Schreibweise

$$\frac{dg}{ds} = \frac{\partial f}{\partial e}.$$

Die Richtungsableitung lässt sich aus den partiellen Ableitungen berechnen gemäß

$$\frac{\partial f}{\partial e} = e_1 \frac{\partial f}{\partial x} + e_2 \frac{\partial f}{\partial y}.$$

Anschauliche Bedeutung: Wenn Sie sich vom Punkt u aus um (die infinitesimale Strecke) ds in Richtung e bewegen, können Sie sich ebenso zuerst um $e_1\, ds$ in x-Richtung bewegen und dann um $e_2\, ds$ in y-Richtung – Sie kommen am selben Punkt $u + e\, ds$ an. Dabei nimmt der Wert der Funktion f erst um $e_1\, ds$ und dann um $e_2\, ds$ zu. Also $f(u+e\, ds) = f(u) + e_1 \frac{\partial f}{\partial x} + e_2 \frac{\partial f}{\partial y}$. Die allgemeine Fassung dieser Regel ist die

Kettenregel in mehreren Variablen: Sei $\gamma : [a,b] \to \mathbb{R}^d$ eine Kurve in $\mathbb{R}^d$ und $f : \mathbb{R}^d \to \mathbb{R}$. Dann ist

$$\frac{d}{dt} f(\gamma(t)) = \sum_{i=1}^{d} \frac{\partial f}{\partial x_i} \frac{d\gamma_i}{dt}.$$

Bsp. Eine Fliege befindet sich zur Zeit t am Ort $r(t) \in \mathbb{R}^3$, und am Ort $x \in \mathbb{R}^3$ herrscht die Temperatur $T(x)$. Dann spürt die Fliege zur Zeit t die Temperatur $T(r(t))$.

Bsp. Wie vorher, aber diesmal ist T eine Fkt von $x \in \mathbb{R}^3$ und $t \in \mathbb{R}$: sie ändert sich mit der Zeit. Dann spürt die Fliege zur Zeit t die Temperatur $T(r(t), t)$. Die Änderungsrate (= Zeitableitung) der gespürten Temperatur erhält man, indem man $\gamma(t) = (r(t), t) \in \mathbb{R}^4$ in die Kettenregel einsetzt:

$$\frac{dT(r(t),t)}{dt} = \sum_{i=1}^{3} \frac{\partial T}{\partial x_i}(r(t),t)\, \dot{r}(t) + \frac{\partial T}{\partial t}(r(t),t).$$

Der Vektor mit den Komponenten $\partial f/\partial x_i$ heißt *Gradient* von f,

$$\operatorname{grad} f = \nabla f = \Big(\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_d} \Big).$$

Das Symbol ∇ heißt “Nabla” (nach einem nahöstlichen Saiteninstrument dieser Form). Der Gradient von f ist ein *Vektorfeld*, d.h. eine Funktion $\nabla f : \mathbb{R}^d \to \mathbb{R}^d$. An jedem Punkt x gibt der Vektor $\nabla f(x)$ folgendes an: die Richtung von $\nabla f(x)$ in $\mathbb{R}^d$ ist die *Richtung des steilsten Anstiegs*; der Betrag $\|\nabla f(x)\|$ ist die Steigung (Richtungsableitung) in dieser Richtung; die Tangentialebene T an den Graphen von f im Punkt ξ hat die Gleichung

$$T(x) = f(\xi) + \sum_{i=1}^{d} \frac{\partial f}{\partial x_i}(\xi)\,(x_i - \xi_i) = f(\xi) + \nabla f(\xi) \cdot (x - \xi).$$

Dabei haben wir im letzten Term die Schreibweise für das *Skalarprodukt* verwendet,

$$u \cdot v = \sum_{i=1}^{d} u_i v_i$$

Das Skalarprodukt ist eine Fkt $\mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$; das Skalarprodukt zweier Vektoren ist also eine Zahl (ein Skalar; nicht etwa ein Vektor).

Zweite partielle Ableitungen. Eine Fkt $f : \mathbb{R}^d \to \mathbb{R}$ hat d erste partielle Ableitungen und d^2 zweite partielle Ableitungen

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial}{\partial x_i} \frac{\partial}{\partial x_j} f\,,$$

die eine Matrix bilden, die so genannte *Hesse-Matrix* H. Bsp: $f(x,y)$

$$\begin{pmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{pmatrix}$$

Entsprechend hat $f : \mathbb{R}^d \to \mathbb{R}$ d^n n-te partielle Ableitungen. Allerdings sind manche dieser Ableitungen gleich:

Satz. Wenn die Funktion $f : \mathbb{R}^2 \to \mathbb{R}$ zweimal partiell differenzierbar ist und die zweiten partiellen Ableitungen stetig sind, dann

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

Man erhält also dasselbe Ergebnis, ob man erst nach x und dann nach y ableitet oder umgekehrt. Infolgedessen ist H eine symmetrische Matrix, $H^T = H$.

Satz. Hat die differenzierbare Funktion $f : \mathbb{R}^d \to \mathbb{R}$ in $x \in \mathbb{R}^d$ ein (lokales) Minimum oder Maximum, so ist $\nabla f(x) = 0$. Ist umgekehrt $\nabla f(x) = 0$ und die Hesse-Matrix $H(x)$ positiv-definit, so hat f in x ein lokales Minimum; ist $\nabla f(x) = 0$ und die Hesse-Matrix $H(x)$ negativ-definit, so hat f in x ein lokales Maximum.

Definition. Eine Matrix $A \in \mathcal{M}(n,n)$ heißt positiv-definit, wenn für alle $u \in \mathbb{R}^n$ mit $u \neq 0$ gilt $u^T A u > 0$. Eine Matrix $A \in \mathcal{M}(n,n)$ heißt negativ-definit, wenn für alle $u \in \mathbb{R}^n$ mit $u \neq 0$ gilt $u^T A u < 0$.