2.2 Gradient und Richtungsableitung

Def.: Ist $f: \mathbb{R}^n \supset \Omega \to \mathbb{R}$ nach jeder Variable x_k partiell diff'bar, so nennen wir den Vektor

$$\nabla f(x) := \left(\frac{\partial f}{\partial x_1}(x), \ldots, \frac{\partial f}{\partial x_n}(x)\right)$$

den <u>Gradienten</u> von f in $x \in \Omega$.

(Der Operator ∇ wird mitunter als "Nabla-Operator" bezeichnet.)

Der Gradient spielt für Funktionen mehrerer Veränderlicher in verschiedener Hinsicht dieselbe Rolle, wie die erste Ableitung für Funktionen einer Variablen. Z.B. gilt die folgende Verallgemeinerung des Mittelwertsatzes (zur Erinnerung: $f(b)-f(a) = f'(\xi)(b-a)$).

<u>Satz</u> (MWS): Es sei $\Omega \subset \mathbb{R}^n$ offen und konvex(*), $f: \Omega \to \mathbb{R}$ stetig partiell diff'bar, $x \in \Omega$ und $x+h \in \Omega$. Dann gibt es ein $\xi \in [x, x+h]$, so daß

$$f(x+h) - f(x) = \langle \nabla f(\xi), h \rangle$$

($\langle\ \rangle$ das $\mathbb{R}^n$-Skalarprodukt, also $\langle \nabla f(\xi), h\rangle = \sum_{k=1}^{n} \frac{\partial f}{\partial x_k}(\xi) \cdot h_k$.)

Ähnlich wie im eindimensionalen ergeben sich:

(*) d.h.: Zu jedem Punktepaar $(x,y) \in \Omega^2$ liegt die gesamte Verbindungsstrecke $\underset{[x,x+h]}{\cancel{\mathcal{S}}} := \{\lambda x + (1-\lambda) y : 0 \le \lambda \le 1\}$ in Ω.

Folgerungen (für Ω und f wie im MWS):

(1) Ist $\nabla f(x) = 0$ für alle $x \in \Omega$, so ist f konstant.

(2) Ist $\nabla f(x) = (c_1, \dots, c_n)$ ein konstanter Vektor, so ist f eine affin-lineare Funktion.

Für das Auffinden einer Extremalstelle spielt der Gradient ebenfalls eine wichtige Rolle. Im Hinblick auf den eindimensionalen Fall ist das folgende notwendige Kriterium nicht überraschend.

<u>Satz</u>: Es sei $\Omega \subset \mathbb{R}^n$ offen und $f : \Omega \to \mathbb{R}$ partiell d'bar nach allen Variablen. In $x_0 \in \Omega$ besitze f ein lokales Extremum. Dann ist $\nabla f(x_0) = 0$, also $\frac{\partial f}{\partial x_k}(x_0) = 0$ für alle $k \in \{1, \dots, n\}$.

Begründung: Ist f in $x_0 = (x_0^{(1)}, \dots, x_0^{(n)})$ extremal, so gilt dies auch für die Hilfsfunktion

$$\varphi_k : t \mapsto \varphi_k(t) = f(x_0^{(1)}, \dots, x_0^{(k-1)}, x_0^{(k)} + t, x_0^{(k+1)}, \dots, x_0^{(n)})$$

in $t = 0$. Also ist $\varphi_k'(0) = 0$, wie wir aus der Differentialrechnung einer Veränderlicher wissen. Dies bedeutet aber gerade $\frac{\partial f}{\partial x_k}(x_0) = 0$, und das gilt für alle $k \in \{1, \dots, n\}$. Also: $\nabla f(x_0) = 0$.

Def.: Ist $\nabla f(x_0)=0$, so heißt x_0 eine <u>kritische Stelle</u> von f.

Bsp.: (a) $f:\mathbb{R}^2\to\mathbb{R}$, $(x,y)\mapsto f(x,y)=x^2+y^2$.

Dann ist $\nabla f(x,y) = 2(x,y)$ und wir haben

$$\nabla f(x,y)=0 \;(=(0,0)) \iff (x,y)=(0,0).$$

Der Nullpunkt ist die einzige kritische Stelle von f. In diesem Fall liegt in $(0,0)$ tatsächlich ein globales Minimum vor, denn es ist $f(0,0)=0$ und $f(x,y)>0$ für alle $(x,y)\in\mathbb{R}^2\setminus\{(0,0)\}$.

(b) Die Bedingung $\nabla f(x,y)=0$ ist im allgemeinen nicht hinreichend für das Vorliegen eines Extremums.

z.B.: $f:\mathbb{R}^2\to\mathbb{R}$, $(x,y)\mapsto f(x,y)=x^2-y^2$.

Hier ist $\nabla f(x,y)=2(x,-y)$, also ebenfalls $(x_0,y_0)=(0,0)$ die (einzige) kritische Stelle von f. Es liegt im Nullpunkt aber kein lokales Extremum vor, denn in jeder Kugel um $(0,0)$ (egal wie klein!) liegen sowohl positive als auch negative Funktionswerte von f (bzw.: werden dort angenommen.).

Um den Begriff der <u>Richtungsableitung</u> einer Funktion

$$f : \mathbb{R}^n \supset \Omega \to \mathbb{R}$$

zu erklären, gehen wir noch einmal zurück zum Begriff der partiellen Ableitung: Mit $x = (x_1, \dots, x_n)$

$$\frac{\partial f}{\partial x_k}(x) = \frac{d}{dt} f(x_1, \dots, x_{k-1}, x_k + t, x_{k+1}, \dots, x_n)\Big|_{t=0}$$

Bezeichnen wir die kanonischen Basisvektoren des $\mathbb{R}^n$, also

$$e_1 = (1, 0, \dots, 0),\ e_2 = (0, 1, 0, \dots, 0), \dots, e_n = (0, \dots, 0, 1),$$

können wir dies auch schreiben als

$$\frac{\partial f}{\partial x_k}(x) = \frac{d}{dt} f(x + t e_k)\Big|_{t=0}.$$

Die partielle Ableitung nach x_k gibt also die Steigung des Graphen von f in Richtung des Einheitsvektors e_k an. Hierbei muß man sich natürlich nicht auf die Vektoren e_k beschränken, sondern man kann beliebige Richtungen betrachten.

<u>Def.</u>: Es sei $\Omega \subset \mathbb{R}^n$ offen und $f : \Omega \to \mathbb{R}$ eine Funktion sowie $\xi \in \mathbb{R}^n$ ein fester Vektor mit $|\xi| = 1$. Dann heißt $\frac{\partial f}{\partial \xi}(x) := \frac{d}{dt} f(x + t\xi)\Big|_{t=0}$ (falls existent!) die Richtungsableitung von f nach ξ im Punkt $x \in \Omega$.

Bem. • Die partiellen Ableitungen $\frac{\partial f}{\partial x_k}$ sind dann gerade die Richtungsableitungen von f nach den kanonischen Einheitsvektoren e_k.

• Verwenden wir die Def. der Ableitung (nach t), erhalten wir

$$\frac{\partial f}{\partial \xi}(x) = \lim_{t \to 0} \frac{1}{t}\left(f(x+t\xi) - f(x)\right).$$

Bsp. Wir berechnen die Richtungsableitung von

$$f: \mathbb{R}^2 \to \mathbb{R}, \quad (x,y) \longmapsto f(x,y) = x^2 y$$

in Richtung $\xi = \left(\frac{3}{5}, \frac{4}{5}\right) \left(= (\xi_1, \xi_2)\right)$. (Beachte: $|\xi| = 1$!)

Wir haben

$$\frac{\partial f}{\partial \xi}(x,y) = \frac{d}{dt} f(x+t\xi_1, y+t\xi_2)\Big|_{t=0} = \frac{d}{dt}(x+t\xi_1)^2(y+t\xi_2)\Big|_{t=0}$$

$$= 2(x+t\xi_1)(y+t\xi_2)\cdot\xi_1 + (x+t\xi_1)^2\cdot\xi_2\Big|_{t=0}$$

$$= 2xy\xi_1 + x^2\xi_2 = \frac{6}{5}xy + \frac{4}{5}x^2 = \frac{x}{5}(6y+4x)$$

Ist die Funktion f partiell diff'bar nach allen Variablen und sind die partiellen Ableitungen alle stetig (man sagt dann, f sei stetig partiell diff'bar), so kann man die Berechnung der Richtungsableitungen mit Hilfe des Gradienten

$$\nabla f(x) = \left(\frac{\partial f}{\partial x_1}(x), \dots, \frac{\partial f}{\partial x_n}(x)\right)$$

deutlich vereinfachen. Genauer gilt der folgende

Satz: Ist $f : \mathbb{R}^n \supset \Omega \to \mathbb{R}$ stetig partiell diff'bar, so existiert zu jedem $\xi \in \mathbb{R}^n$ mit $|\xi| = 1$ die Richtungsableitung $\frac{\partial f}{\partial \xi}(x)$ und es gilt die Identität

$$\frac{\partial f}{\partial \xi}(x) = \langle \nabla f(x), \xi \rangle = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x)\, \xi_i \,.$$

Fortsetzung des Bsp.: Für $f(x,y) = x^2 y$ haben wir

$\nabla f(x,y) = (2xy, x^2)$. Es folgt

$$\frac{\partial f}{\partial \xi}(x,y) = 2xy\,\xi_1 + x^2 \xi_2 \underset{\xi = \frac{1}{5}(3,4)}{=} \frac{1}{5}(6xy + 4x^2)$$

in Übereinstimmung mit dem Ergebnis oben. Durch die Anwendbarkeit der Ableitungsregeln wird die Rechnung abgekürzt.

Für die Richtungsableitungen und den Gradienten werden nun auch die entsprechenden Elastizitäten definiert.

Def.: Es sei $f : (0,\infty)^n \supset \Omega \to (0,\infty)$ stetig partiell diff'bar. Dann heißen

(a) $\vec{\mathcal{E}}_f(x) := (\mathcal{E}_{f,x_1}(x), \dots, \mathcal{E}_{f,x_n}(x))$

der Elastizitätsgradient von f in x und

(b) $\mathcal{E}_{f,\xi}(x) := \langle \xi, \vec{\mathcal{E}}_f(x) \rangle = \sum_{i=1}^{n} \xi_i\, \mathcal{E}_{f,x_i}(x)$

die Richtungselastizität von f nach ξ in x.

Bsp.: Für die Cobb-Douglas-Funktion

$$f(x) = x_1^{s_1} x_2^{s_2} \cdots x_n^{s_n} = \prod_{k=1}^{n} x_k^{s_k} \qquad (x_k > 0)$$

hatten wir bereits $\varepsilon_{f,x_k}(x) = s_k$ ausgerechnet (unabhängig von x!). Hieraus ergeben sich

$$\vec{\varepsilon_f}(x) = (s_1, \ldots, s_n) \quad \text{und} \quad \varepsilon_{f,\xi}(x) = \sum_{k=1}^{n} s_k \xi_k .$$