6. Extrema unter Nebenbedingungen

Betrachte $f(x_1,x_2) := x_1 x_2$ unter der Nebenbedingung $g(x) = x_1 + x_2 - 2 = 0$ (Fläche des Rechtecks mit Seitenlänge 4). Angenommen, es gibt $\varphi_1, \varphi_2 : [a,b] \to \mathbb{R}$ mit

$$g(\varphi_1(t), \varphi_2(t)) = 0$$

(z.B., $\varphi_1(t) = t$, $\varphi_2(t) = 2-t$). Dann suchen wir das Extremum von f unter der Nebenbedingung $g = 0$ durch Ableiten von

$$h(t) = f(\varphi_1(t), \varphi_2(t)) = f(\varphi(t)).$$

Suche also t_0 mit

$$0 = h'(t_0) = \partial_1 f(\varphi(t_0))\, \varphi_1'(t_0) + \partial_2 f(\varphi(t_0))\, \varphi_2'(t_0).$$

Wegen $g(\varphi(t)) = 0$ gilt auch

$$0 = \partial_1 g(\varphi(t_0))\, \varphi_1'(t_0) + \partial_2 g(\varphi(t_0))\, \varphi_2'(t_0).$$

Also folgt für $x_0 := \varphi(t_0)$

$$\begin{pmatrix} \partial_1 f(x_0) & \partial_2 f(x_0) \\ \partial_1 g(x_0) & \partial_2 g(x_0) \end{pmatrix} \begin{pmatrix} \varphi_1'(t_0) \\ \varphi_2'(t_0) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Falls $\varphi'(t_0) \neq 0$ ist, so sind $\nabla f(x_0)$ und $\nabla g(x_0)$ linear abhängig. Ist $\nabla g(x_0) \neq 0$, so erhalten wir

$$\exists \lambda_0 \in \mathbb{R} : \nabla f(x_0) + \lambda_0 \nabla g(x_0) = 0.$$

Zusammen mit $g(x_0) = 0$ lässt sich also x_0 bestimmen (ohne nach t_0 aufzulösen).

Im Beispiel: $f(x) = x_1 x_2$, $g(x) = x_1 + x_2 - 2$, $\nabla f(x) = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, $\nabla g(x) = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \neq 0$

$$\left.\begin{aligned} x_2 + \lambda_0 \cdot 1 &= 0 \\ x_1 + \lambda_0 \cdot 1 &= 0 \end{aligned}\right\} x_1 = x_2 = -\lambda_0$$

$$x_1 + x_2 - 2 = 0 \quad \Rightarrow \quad x_1 = x_2 = 1.$$

$\measuredangle_1$

Satz 6.1. (Extrema unter Nebenbedingungen). Seien $U \subset \mathbb{R}^n$ offen, $f \in C^1(U;\mathbb{R})$, $g \in C^1(U;\mathbb{R}^m)$, $m < n$. Es gelte $\operatorname{rk} g'(x) = m$ (voller Rang) für alle $x \in U$. Sei $M := \{x \in U : g(x) = 0\}$ (Nebenbedingung). Falls $f|_M$ an der Stelle $x_0 \in M$ ein lokales Extremum besitzt, so gibt es ein $\lambda^{(0)} \in \mathbb{R}^m$ mit

$$f'(x_0) + (\lambda^{(0)})^T g'(x_0) = 0$$

d.h.,

$$\nabla f(x_0) + \sum_{j=1}^{m} \lambda_j^{(0)} \nabla g_j(x_0) = 0.$$

Der Vektor $\lambda^{(0)} \in \mathbb{R}^m$ heißt Lagrange-Multiplikator der Extremalaufgabe.

Bemerkung 6.2: Sei in der obigen Situation $\mathcal{L}: U \times \mathbb{R}^m \to \mathbb{R}, (x,\lambda) \mapsto f(x) + \sum_{j=1}^m \lambda_j g_j(x)$. Wir haben gesehen, dass $(x_0, \lambda^{(0)})$ ein kritischer Punkt von $\mathcal{L}$ ist:

$$\mathcal{L}'(x,\lambda) = \left(f'(x) + \sum_{j=1}^{m} \lambda_j g_j'(x), g_1(x), \dots, g_m(x) \right) = 0$$

für $x = x_0$, $\lambda = \lambda^{(0)}$. Falls

$$\mathcal{L}_{xx}(x_0, \lambda^{(0)}; h, h) > 0 \text{ für alle } h \in \mathbb{R}^n \setminus \{0\}, \qquad (*)$$

so folgt insbesondere

$$\mathcal{L}(x_0, \lambda^{(0)}) \leq \mathcal{L}(x, \lambda^{(0)}) \qquad \text{für alle } x \in U(x_0)$$

und damit auch

$$f(x_0) \leq f(x) \qquad \text{für alle } x \in M \cap U(x_0),$$

da $g_j(x) = 0$ für $x \in M$. Wir haben also ein lokales Minimum von f an der Stelle x_0 unter der Nebenbedingung $g(x) = 0$. Die Funktion $\mathcal{L}$ heißt Lagrangefunktion. Bedingung $(*)$ lässt sich noch abschwächen. ∮

Beispiel 6.3 (Volumen des Quaders). Seien $f: (0,3) \to \mathbb{R}$, $x \mapsto x_1 x_2 x_3$, $g: \mathbb{R}^3 \to \mathbb{R}$, $x \mapsto x_1 + x_2 + x_3 - 3$, $\nabla f(x) = (x_2 x_3, x_1 x_3, x_1 x_2)^T$, $\nabla g(x) = (1,1,1)^T \neq 0$, $\operatorname{rk} g(x) = 1$ für alle $x \in \mathbb{R}^3$. Dann ergibt

$$\nabla f(x) + \lambda \nabla g(x) = 0$$

das System

$$\left.\begin{array}{l} x_2 x_3 + \lambda = 0 \\ x_1 x_3 + \lambda = 0 \end{array}\right\} x_3 \neq 0: x_1 = x_2$$

$$\left.\begin{array}{l} x_1 x_3 + \lambda = 0 \\ x_1 x_2 + \lambda = 0 \end{array}\right\} x_1 \neq 0: x_3 = x_2$$

Wegen $g(x) = x_1 + x_2 + x_3 - 3 = 0$ folgt mit $x_1 = x_2 = x_3$ die Beziehung

$x_i = 1$, $1 \le i \le 3$. Also, $\lambda = -1$. Mit $\mathcal{L}(x,\lambda) = f(x) + \lambda g(x)$ erhalten wir an $x^* = (1,1,1)^T$ die Matrix

$$\mathcal{L}_{xx}(x^*,\lambda) = \begin{pmatrix} 0 & x_3^* & x_2^* \\ x_3^* & 0 & x_1^* \\ x_2^* & x_1^* & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}.$$

Offenbar ist $\mathcal{L}_{xx}(x^*,\lambda)$ nicht positiv definit. Aber für alle $h \in \ker g'(x^*) = \ker(1,1,1)$ folgt

$$g'(x^*)h = 0 \iff h_1 + h_2 + h_3 = 0 \quad \text{für } h = (h_1, h_2, h_3)^T.$$

Also

$$\begin{aligned} \mathcal{L}_{xx}(x^*,\lambda) &= h_1(h_2+h_3) + h_2(h_1+h_3) + h_3(h_1+h_2) \\ &= h_1(-h_1) + h_2(-h_2) + h_3(-h_3) \\ &= -h_1^2 - h_2^2 - h_3^2 = -\|h\|^2 \end{aligned}$$

für alle $h \in \ker g'(x^*)$. Das ist hinreichend dafür, dass x^* Lösung von

$$\max f(x) \quad \text{u.d.N.} \quad g(x) = 0$$

ist. ↯