

0.1 Extremalaufgaben mit Nebenbedingungen

Problemstellung: „Restringierte“ Optimierungsaufgabe mit Gleichungsnebenbedingungen.

Sei $f: D \rightarrow \mathbb{R}$ und $g: D \rightarrow \mathbb{R}^k$, $D \subseteq \mathbb{R}^n$. Wir suchen einen Punkt $\hat{x} \in D$, s.d. $\hat{x} \in S := \{x \in D \mid g(x) = 0\}$ und $\exists U(\hat{x})$ s.d. $f(\hat{x}) \leq f(x)$, $\forall x \in U(\hat{x}) \cap S$.

Dann heißt $\hat{x}$ lokales Minimum unter Nebenbedingung $g(x) = 0$. Analog: lokales Maximum unter Nebenbedingung $\hat{x} \in S$, s.d. $\exists U(\hat{x})$ mit $f(\hat{x}) \geq f(x) \forall x \in U(\hat{x}) \cap S$.

Satz 0.1 (Multiplikatorregel von Lagrange: Notwendige Bed. 1. Ordnung für lokales Minimum unter Nebenbedingungen). Sei $D \subseteq \mathbb{R}^n$ offen, $f: D \rightarrow \mathbb{R}$ und $g: D \rightarrow \mathbb{R}^k$ partiell stetig differenzierbar. Sei $\hat{x} \in D$ ein Extremum unter der Nebenbedingung $g(x) = 0$ und die Gradienten $\nabla g_1(\hat{x}), \dots, \nabla g_k(\hat{x})$ seien linear unabhängig in $\mathbb{R}^n$. Dann gilt

$$\exists \hat{\lambda} = \begin{pmatrix} \hat{\lambda}_1 \\ \vdots \\ \hat{\lambda}_k \end{pmatrix} \in \mathbb{R}^k \text{ mit } \sum_{i=1}^k \hat{\lambda}_i \nabla g_i(\hat{x}) = \nabla f(\hat{x}) \iff \nabla g(\hat{x}) \hat{\lambda} = \nabla f(\hat{x}).$$

Die Zahlen $\hat{\lambda}_1, \dots, \hat{\lambda}_k$ heißen Lagrange-Multiplikatoren.

Beweis. Nach Voraussetzungen gilt

$$\frac{\partial g_i(\hat{x})}{\partial x} = \underbrace{\left(\frac{\partial g_i(\hat{x})}{\partial x_1} \dots \frac{\partial g_i(\hat{x})}{\partial x_n} \right)}_{i=1 \dots k, \text{ linear unabhängige Vektoren}}.$$

Also hat $\frac{\partial g}{\partial x}(\hat{x}) \in \mathbb{R}^{k \times n}$ Rang k . O.B.d.A. die ersten k Spalten von $\frac{\partial g}{\partial x}(\hat{x})$ bilden eine quadratische invertierbare Matrix. Dann lassen sich x und $\frac{\partial g}{\partial x}(\hat{x})$ aufspalten:

$$x = \begin{pmatrix} y \\ z \end{pmatrix} \quad \underbrace{\frac{\partial g}{\partial x}(\hat{x})}_{\in \mathbb{R}^{k \times n}} = \left(\underbrace{\frac{\partial g}{\partial y}(\hat{x})}_{\in \mathbb{R}^{k \times k}}; \underbrace{\frac{\partial g}{\partial z}(\hat{x})}_{\in \mathbb{R}^{k \times (n-k)}} \right).$$

mit $y \in \mathbb{R}^k$, $z \in \mathbb{R}^{n-k}$ und $\frac{\partial g}{\partial y}(\hat{x}) \in \mathbb{R}^{k \times k}$ regulär.

Setze $\hat{x} = \begin{pmatrix} \hat{y} \\ \hat{z} \end{pmatrix}$. Wende nun Satz ?? auf $g(x) = g(y, z) = 0$ an. Dann existieren Umgebungen $U(\hat{z}) \subseteq \mathbb{R}^{n-k}$, $U(\hat{y}) \subseteq \mathbb{R}^k$ und eine eindeutige Abbildung

$$\begin{aligned} \varphi: U(\hat{z}) &\rightarrow U(\hat{y}) \\ z &\mapsto \varphi(z) = y, \end{aligned}$$

s.d. φ folgende Eigenschaften erfüllt

- (1) $g(\varphi(z), z) = 0 \forall z \in U(\hat{z})$
- (2) $\hat{y} = \varphi(\hat{z})$
- (3) $\varphi \in C^1(U(\hat{z}), \mathbb{R}^k)$ stetig differenzierbar.
- (4) $\underbrace{\varphi'(\hat{x})}_{D_x \varphi(\hat{x})} = - \left(\frac{\partial g}{\partial y}(\hat{x}) \right)^{-1} \cdot \left(\frac{\partial g}{\partial z}(\hat{x}) \right)$

Betrachte $\tilde{f}(z) = f(\varphi(z), z)$, $\tilde{f}(z): U(\hat{z}) \rightarrow \mathbb{R}$. Da $\hat{x}$ Extremum von $f(x)$ unter $g(x) = 0$, ist $\hat{z}$ lokales Extremum von $\tilde{f}(z)$ in $U(\hat{z})$. Mit ?? folgt also $\forall i = 1 \dots n-l$:

$$0 = \frac{\partial \tilde{f}(\hat{z})}{\partial z_i} \stackrel{\text{Kettenregel}}{=} \frac{\partial f(\hat{x})}{\partial y} \cdot \frac{\partial \varphi(\hat{z})}{z_i} + \frac{\partial f(\hat{x})}{z_i}$$

$\tilde{f} = f(\varphi(z), z)$

Damit folgt

$$0 = \frac{\partial f(\hat{x})}{\partial y} \cdot \frac{\partial \varphi(\hat{z})}{\partial z} + \frac{\partial f(\hat{x})}{\partial z} \quad (*)$$

Definiere

$$\hat{\lambda}^T = \underbrace{\frac{\partial f(\hat{x})}{\partial y}}_{\left(\frac{\partial f}{\partial y_1} \dots \frac{\partial f}{\partial y_k}\right)} \cdot \left(\frac{\partial g}{\partial y}(\hat{x})\right)^{-1}$$

Damit folgt

$$\frac{\partial f(\hat{x})}{\partial y} = \hat{\lambda}^T \left(\frac{\partial g}{\partial y}(\hat{x})\right).$$

Mit (*) folgt

$$\frac{\partial f(\hat{x})}{\partial y} \left(-\left(\frac{\partial g(\hat{x})}{\partial y}\right)^{-1} \frac{\partial g(\hat{x})}{\partial z}\right) + \frac{\partial f(\hat{x})}{\partial z} = -\hat{\lambda}^T \frac{\partial g(\hat{x})}{\partial z} + \frac{\partial f(\hat{x})}{\partial z} = 0.$$

Insgesamt folgt

$$\left. \begin{aligned} \frac{\partial f(\hat{x})}{\partial y} &= \hat{\lambda}^T \frac{\partial g}{\partial y}(\hat{x}) \\ \frac{\partial f(\hat{x})}{\partial z} &= \hat{\lambda}^T \frac{\partial g}{\partial z}(\hat{x}) \end{aligned} \right\} \implies \frac{\partial f(\hat{x})}{\partial x} = \hat{\lambda}^T \frac{\partial g(\hat{x})}{\partial x}.$$

□

Bemerkung 0.2 (Interpretation von Satz 0.1). Definiere Lagrange-Funktion

$$\mathcal{L}(x, \lambda) := f(x) - \lambda^T g(x), \quad (x, \lambda) \in D \times \mathbb{R}^k.$$

Falls $\hat{x}$ lokales Minimum von f unter Nebenbedingung $g(x) = 0$ und $\text{Rg}\left(\frac{\partial g(\hat{x})}{\partial x}\right) = k$. Dann ex. genau ein $\hat{\lambda} \in \mathbb{R}^k$ s.d. $(\hat{x}, \hat{\lambda})$ ein stationärer Punkt der Lagrange Funktion ist:

$$\begin{aligned} \nabla_x \mathcal{L}(\hat{x}, \hat{\lambda}) &= \nabla f(\hat{x}) - \nabla g(\hat{x}) \hat{\lambda} = 0 \\ \nabla_\lambda \mathcal{L}(\hat{x}, \hat{\lambda}) &= g(\hat{x}) = 0. \end{aligned}$$

Beispiel 0.3 (Anwendung von 0.1). Sei $A = (a_{ij})_{i,j=1}^n \in \mathbb{R}^{n \times n}$ eine symmetrische Matrix. Dann betrachte

$$f(x) := (x, Ax)_2 = \sum_{i,j=1}^n a_{ij} x_i x_j.$$

Bestimme Extrema von $f(x)$ unter Nebenbedingungen $\|x\| = 1$.

Definiere $g(x) = \|x\|_2^2 - 1$ und $S := \{x \in \mathbb{R}^n \mid g(x) = 0\}$. Dann gilt für $x \in S$: $\nabla g(x) = 2x \neq 0$, da $\|x\|_2^2 = 1$. Für $f(x) = \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j$ gilt für $k = 1 \dots n$:

$$\begin{aligned} \frac{\partial f}{\partial x_k} &= \sum_{i=1}^n \sum_{j=1}^n a_{ij} \delta_{ik} x_j + \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i \delta_{jk} \\ &= \sum_{j=1}^n a_{kj} x_j + \sum_{i=1}^n a_{ik} x_i \\ &\stackrel{a_{ij}=a_{ji}}{=} 2 \sum_{i=1}^n a_{ki} x_i \end{aligned}$$

Also folgt

$$\nabla f(x) = 2Ax.$$

Existiert ein $\hat{x}$? Da S kompakt und f stetig, nimmt f (auf S) ein Maximum und Minimum an. Nach Satz 0.1 ex. ein $\hat{\lambda} \in \mathbb{R}$, s.d.

$$\begin{aligned} \nabla f(\hat{x}) &= \hat{\lambda} \nabla g(\hat{x}) \\ \implies 2A\hat{x} &= \hat{\lambda} 2\hat{x} \\ \implies A\hat{x} &= \hat{\lambda} \hat{x}. \end{aligned}$$

Also ist $\hat{\lambda}$ Eigenwert von A zum Eigenvektor $\hat{x}$. Damit folgt

$$f(\hat{x}) = (\hat{x}, A\hat{x})_2 = (\hat{x}, \hat{\lambda}\hat{x})_2 = \hat{\lambda} \underbrace{\|\hat{x}\|_2^2}_{=1} = \hat{\lambda}.$$

Das bedeutet, dass

$$\inf\{(x, Ax)_2 \mid \|x\|_2 = 1\} = f(\hat{x}) = \hat{\lambda} = \lambda_{\min}.$$

Also folgt

$$\begin{aligned} \lambda_{\min} &= \min_{\|x\|_2=1} \underbrace{x^T Ax}_{(x, Ax)_2} = \min_{x \in \mathbb{R}^n \setminus \{0\}} \underbrace{\frac{x^T Ax}{\|x\|_2^2}}_{\text{Rayley-Quotient}} \\ \lambda_{\max} &= \max_{x \in \mathbb{R}^n \setminus \{0\}} \frac{x^T Ax}{\|x\|_2^2}. \end{aligned}$$

$\lambda_{\min}$ bzw. $\lambda_{\max}$ sind der kleinste bzw. größte Eigenwert von A .