IV.5. Extrema unter Nebenbedingungen

Aufgabe: Bestimme Extrema einer Funktion f auf einer UMF $M \subseteq \mathbb{R}^n$, d.h.

$$\sup_{x\in M} f(x) \quad \text{oder} \quad \inf_{x\in M} f(x)$$

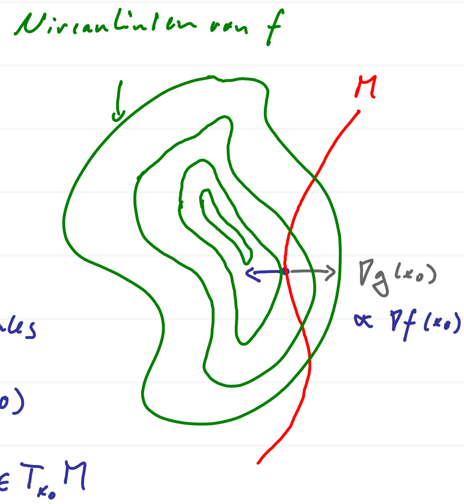
Hierbei ist $f \in C^1(U \subseteq \mathbb{R}^n, \mathbb{R})$, $g \in C^1(U, \mathbb{R}^k)$, $M := \{x \in \mathbb{R}^n \mid g(x) = y\}$ und y regulärer Wert von g.

Wenn M eine offene Teilmenge von $\mathbb{R}^n$ ist, müssen wir außer dem Rand ∂M dazu nur die stationären Punkte $\{x \mid \nabla f(x) = 0\}$ untersuchen.

Allgemein gilt:

Satz: Sei y ein regulärer Wert von g (d.h., $g(x) = y \Rightarrow \operatorname{Rang} g'(x) = k$). Ist x_0 Extremalpunkt von f auf M, dann ist $\nabla f(x_0)$ eine Linearkombination von $\{\nabla g_i(x_0)\}_{i=1}^k$. D.h. es gibt $\lambda_1, \dots, \lambda_k \in \mathbb{R}$ (sog. „Lagrangemultiplikatoren"), so dass

$$\nabla f(x_0) = \sum_{i=1}^k \lambda_i \nabla g_i(x_0).$$



Beweis: Für jedes $v \in T_{x_0}M$ gibt es eine C^1-Kurve $\gamma: (-1,1) \to M$ mit $\gamma(0) = x_0$ und $v = \dot\gamma(0)$. Ist x_0 Extremalpunkt, dann hat $t \mapsto (f \circ \gamma)(t)$ bei $t = 0$ ein lokales Extremum, d.h.

$$0 \overset{!}{=} (f\circ\gamma)'(0) = f'(x_0)\dot\gamma(0) = \langle \nabla f(x_0), v\rangle \quad \forall v \in T_{x_0}M$$

Also ist $\nabla f(x_0) \perp T_{x_0}M$, & wegen $T_{x_0}M = \{\nabla g_i(x_0)\}_{i=1..k}^{\perp}$ ist

$$\nabla f(x_0) = \operatorname{span}\{\nabla g_i(x_0)\}_{i=1}^k$$

$\square$

Bemerkungen zu den Lagrange-Multiplikatoren:

- Regularität bedeutet, dass $\{\nabla g_i(x_0)\}_{i=1}^{k}$ lin. unabh. sind für alle x mit $g(x)=y$.
- Regularität impliziert Eindeutigkeit der Lagrange-Multiplikatoren:

  angenommen $\nabla f(x_0) = \sum_i \lambda_i \nabla g_i(x_0) = \sum_i \tilde{\lambda}_i \nabla g_i(x_0)$, dann ist

  $\sum_i (\lambda_i - \tilde{\lambda}_i) \nabla g_i(x_0) = 0$ und wegen Lin. Unabh. $\lambda_i = \tilde{\lambda}_i \ \forall i$.

„Lösungsrezept" für das Auffinden von Extrema unter Nebenbedingungen:

(i) Löse die $k+n$ Gleichungen
$$\nabla f(x) = \sum_{i=1}^{k} \lambda_i \nabla g_i(x)$$
$$g(x) = y$$
nach den $k+n$ Variablen (λ, x) auf.

(keine Garantie, dass dies klappt)

$\longrightarrow$ Kandidaten für Extrempunkte

(ii) Betrachte folgende Punkte separat:

- kritische Punkte (d.h. $x \in M$: Rang $g'(x) < k$)
- Punkte an denen keine Diff.barkeit vorliegt
- Punkte auf dem Rand

(iii) Bestimme Extrema unter den Kandidaten

Rechtfertige Existenz von Min./Max. wo möglich durch Kompaktheitsargument (ggfs. durch Einschränken / Erweitern des Def.bereichs)

Bsp. 1: Sei $A = A^T \in \mathbb{R}^{n \times n}$ und $f(x) := \langle x, Ax \rangle$, $x \in \mathbb{R}^n$.

Bestimme $\sup_{x \in S^{n-1}} f(x) \underset{\uparrow}{=} \max_{x \in S^{n-1}} f(x)$

S^{n-1} kompakt & f stetig

$$\left.\begin{array}{l} g(x) := \langle x, x \rangle \\ y = 1 \end{array}\right\} \Rightarrow \nabla g(x) = 2x \neq 0 \quad \forall x \in S^{n-1} \Rightarrow \text{Regularität } \checkmark$$

$$\left.\begin{array}{ccc} & \nabla f(x) = & \lambda \nabla g(x) \\ A=A^T \to & \| & \| \\ & 2Ax & 2\lambda x \end{array}\right\} \Rightarrow$$

d.h. die Kandidaten sind wegen $Ax = \lambda x$ die Eigenvektoren von A & wegen

$$f(x) = \langle x, Ax \rangle = \lambda \langle x, x \rangle \underset{x \in S^{n-1}}{\overset{\uparrow}{=}} \lambda$$

ist $\max_{x \in S^{n-1}} f(x) = \lambda_{max}$ größter Eigenwert von A.

Bsp. ②: Auf $\Sigma_n := \{ p \in \mathbb{R}^n \mid \sum_{i=1}^n p_i = 1 \wedge \forall i : p_i \geq 0 \}$ ist die „<u>Entropie</u>" definiert durch $S(p) := -\sum_{i=1}^n p_i \log p_i$ (mit $0 \log 0 := 0$)

Bestimme $\sup_{p \in \Sigma_n} S(p) = ?$

Betrachte dazu zunächst das Innere Σ_n° und $S|_{\Sigma_n^\circ}$

Die Nebenbedingung ist gegeben durch $g(p) := \sum_{i=1}^n p_i \overset{!}{=} 1$.

Daraus folgt wegen $\nabla g_i(x) = 1$ Regularität.

$\tilde{p} \in \Sigma_n^\circ$ extremal $\Rightarrow$ $\nabla S(\tilde{p}) = \lambda \nabla g(\tilde{p})$ für geeignetes $\lambda \in \mathbb{R}$.

Mit $(\nabla S(\tilde{p}))_i = -(\log \tilde{p}_i) - 1$ bedeutet dies

$$\lambda + 1 + \log \tilde{p}_i = 0 \quad \forall i \quad \text{also} \quad \tilde{p}_i = \text{Konstant}$$

$$\underset{\sum_i \tilde{p}_i = 1}{\overset{\uparrow}{\Rightarrow}} \quad \tilde{p}_i = \frac{1}{n}$$

Da Σ_n kompakt ist & S stetig, existiert ein Maximum.

Angenommen dies ist bei $\hat{p} = (\underbrace{\hat{p}_1, \dots, \hat{p}_m}_{\in \Sigma_m^\circ}, 0, \dots, 0)$, $m \leq n$, dann ist

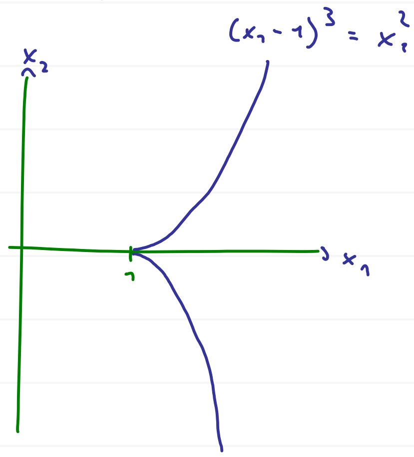
$$S(\hat{p}) = \max_{p \in \Sigma_n} S(p) = \max_{p \in \Sigma_m^\circ} S(p) = -\sum_{i=1}^m \frac{1}{m} \log \frac{1}{m} = \log m$$

Da $\log n \geq \log m$, muss also $n = m$ und daher $\tilde{p} = \hat{p}$.

$$\Rightarrow \quad \max_{p \in \Sigma_n} S(p) = \log n .$$

<u>Bsp. 3:</u> Bestimme $\inf_{x : g(x) = 0} f(x)$ für $f : \mathbb{R}^2 \to \mathbb{R}$ $f(x) := \|x\|^2$.

mit $g(x) := (x_1 - 1)^3 - x_2^2$



$\lambda \nabla g(x) = \nabla f(x)$ hat <u>keine</u> Lösung auf

$M := \{ x \mid g(x) = 0 \}$

Dennoch gibt es ein Minimum, und zwar bei $(1,0)$. Dies ist ein <u>kritischer</u> Punkt, für den $\nabla g(x) = \left(3(x_1 - 1)^2 , -2x_2 \right) = 0$ ist.