

Optimierung

Vorlesung 5 Optimierung mit Nebenbedingungen

Gegeben: Funktion $f, h_1, h_2, \dots, h_m: \mathbb{R}^n \rightarrow \mathbb{R}$

Ziel:

$$\min_{x \in \mathbb{R}^n} f(x)$$

$$\text{wobei } h_1(x) = 0, h_2(x) = 0, \dots, h_m(x) = 0$$

Unrestringierter Fall: Notwendige Bedingung für lokales Minimum x^*

- keine Abstiegsrichtung bei x^* : $\nabla f(x^*) = 0$

Restringierter Fall: Notwendige Bedingung für lokales Minimum x^*

- x^* ist **zulässig**: $\forall i: h_i(x^*) = 0$
- keine **zulässige** Abstiegsrichtung s bei x^*

Restringierter Fall: Notwendige Bedingung für lokales Minimum x^*

- x^* ist **zulässig**: $\forall i: h_i(x^*) = 0$
- keine **zulässige** Abstiegsrichtung s

Zulässiger inkrementeller Schritt δ :

$$h_i(x^* + \delta) = h_i(x^*) = 0$$

Taylor-Entwicklung bei x^* :

$$h_i(x^* + \delta) = h_i(x^*) + \delta^\top \nabla h_i(x^*) + o(\|\delta\|)$$

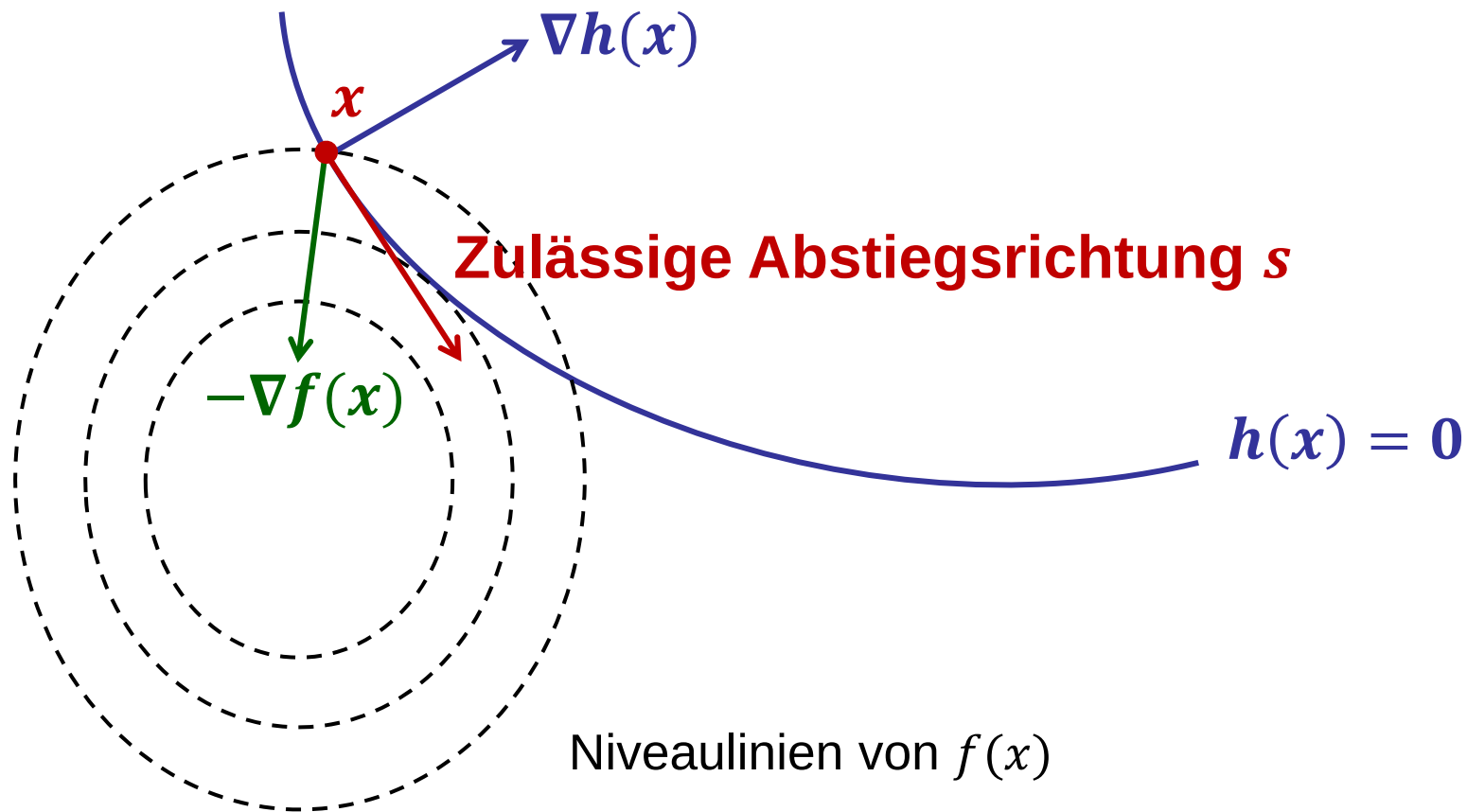
Zulässige Richtungen s sind orthogonal zu $\nabla h_i(x^*)$: $s^\top \nabla h_i(x^*) = 0$

Zulässige Abstiegsrichtung s :

$$\forall i: s^\top \nabla h_i(x^*) = 0, \quad s^\top \nabla f(x^*) < 0$$

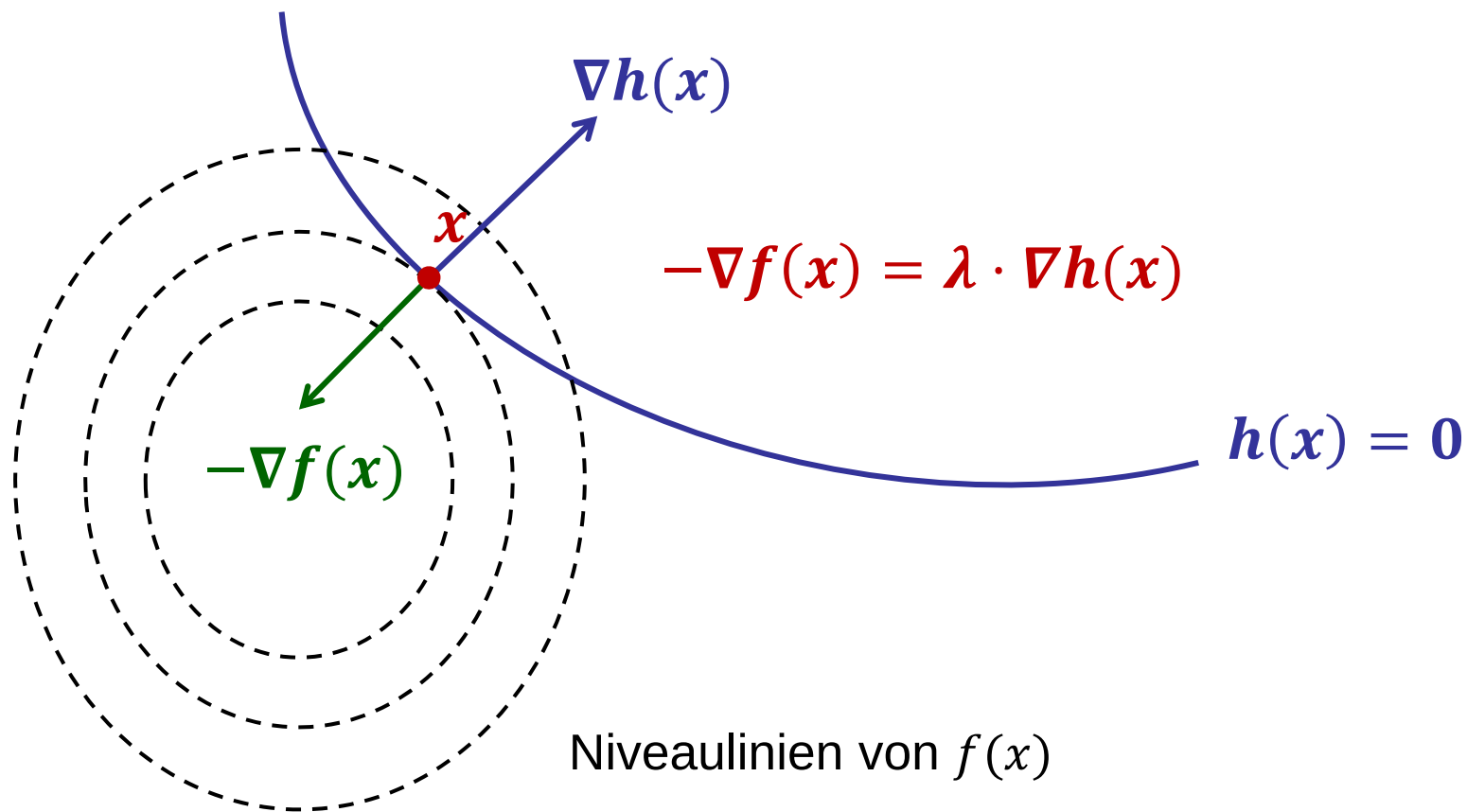
Zulässige Abstiegsrichtung s :

$$\forall i: s^\top \nabla h_i(x^*) = 0, \quad s^\top \nabla f(x^*) < 0$$



Keine zulässige Abstiegsrichtung s :

$$\forall s: \text{ Falls } \forall i: s^\top \nabla h_i(x^*) = 0, \quad \text{dann } s^\top \nabla f(x^*) = 0$$



Keine zulässige Abstiegsrichtung s :

$$\text{Falls } \forall i: s^\top \nabla h_i(x^*) = 0, \quad \text{dann } s^\top \nabla f(x^*) = 0$$

Satz (Lagrange Multiplikatoren): Unter Annahme von Regularität, gibt es bei x^* keine zulässige Abstiegsrichtung genau dann, wenn λ_i^* existieren, so dass

$$\nabla f(x^*) + \sum_{i=1}^m \lambda_i^* \cdot \nabla h_i(x^*) = 0.$$

Beweis:

Satz (Lagrange Multiplikatoren): Unter Annahme von Regularität, gibt es bei x^* keine zulässige Abstiegsrichtung genau dann, wenn λ_i^* existieren, so dass

$$\nabla f(x^*) + \sum_{i=1}^m \lambda_i^* \cdot \nabla h_i(x^*) = 0.$$

Regularität: Insbesondere müssen $\nabla h_i(x^*)$ linear unabhängig sein

Methode von Lagrange:

Finde Vektoren x^* und λ^* , für welche

$$\begin{aligned} \nabla f(x^*) + \sum_{i=1}^m \lambda_i^* \cdot \nabla h_i(x^*) &= 0 \\ h_i(x^*) &= 0, \quad i = 1, \dots, m \end{aligned}$$

- $n + m$ Unbekannte, $n + m$ Gleichungen

Lagrange: $\nabla f(x^*) + \sum_i \lambda_i^* \cdot \nabla h_i(x^*) = 0, \quad h_i(x^*) = 0 \ (i = 1, \dots, m)$

Beispiel:

$$\begin{aligned} \min f(x) &= x_1 + x_2 \\ \text{unter } h(x) &= x_1^2 - x_2 = 0 \end{aligned}$$

Lagrange Funktion:

$$L(x, \lambda) := f(x) + \sum_{i=1}^m \lambda_i \cdot h_i(x)$$

Gradient:

$$\frac{\partial}{\partial x_j} L(x, \lambda) = \frac{\partial}{\partial x_j} f(x) + \sum_{i=1}^m \lambda_i \cdot \frac{\partial}{\partial x_j} h_i(x)$$

$$\frac{\partial}{\partial \lambda_j} L(x, \lambda) = h_j(x)$$

Lagrange Methode:

$$\nabla f(x^*) + \sum_{i=1}^m \lambda_i^* \cdot \nabla h_i(x^*) = 0, \quad h_i(x^*) = 0 \quad (i = 1, \dots, m)$$

Lagrange Funktion:

$$L(x, \lambda) := f(x) + \sum_{i=1}^m \lambda_i \cdot h_i(x)$$

Lagrange Methode:

Notwendige Bedingung für lokales Minimum bei regulärem Punkt x^* :

$$\nabla L(x^*, \lambda^*) = 0$$

- Reduziert das Problem auf das Finden eines stationären Punktes eines unrestringierten Problems!
- Methoden: (konj.) Gradientenmethode, Newton, Quasi-Newton, ...

Gegeben: Funktion $f, h_1, h_2, \dots, h_m: \mathbb{R}^n \rightarrow \mathbb{R}$

Ziel:

$$\min_{x \in \mathbb{R}^n} f(x)$$

$$\text{wobei } g_1(x) \leq 0, g_2(x) \leq 0, \dots, g_p(x) \leq 0$$

Notwendige Bedingung für lokales Minimum x^* :

- x^* ist **zulässig**: $\forall i: g_i(x^*) \leq 0$
- keine **zulässige** Abstiegsrichtung s bei x^*

Aktive Nebenbedingungen:

$$I^* := \{i \mid g_i(x^*) = 0\}$$

- Nur aktive Nebenbedingungen beschränken die gültigen Richtungen

Aktive Nebenbedingung g_i an Punkt x^* :

$$g_i(x^*) = 0$$

Zulässige Richtung s and Punkt x^* :

- Für genug kleines $\delta > 0$ gilt:

$$g_i(x^* + \delta \cdot s) \leq 0$$

- Taylor-Approximation:

$$g_i(x^* + \delta \cdot s) \approx g_i(x^*) + \delta \cdot s^\top \nabla g_i(x^*)$$

- Zulässige Richtung s bei x^* (notwendige Bedingung):

$$s^\top \nabla g_i(x^*) \leq 0$$

Zulässige Abstiegsrichtung s bei x^* :

$$s^\top \nabla f(x^*) < 0 \quad \text{und} \quad \forall i \in I^*: s^\top \nabla g_i(x^*) \leq 0$$

Bei **regulärem lokalem Minimum** x^* hat folgendes System keine Lösung:

$$\begin{aligned} s^\top \nabla g_i(x^*) &\leq 0, & \forall i \in I^* \\ s^\top \nabla f(x^*) &< 0 \end{aligned} \quad (A)$$

Hinreichende Bedingung, damit (A) keine Lösung hat:

$$-\nabla f(x^*) = \sum_{i \in I^*} \lambda_i \cdot \nabla g_i(x^*), \quad \text{mit } \lambda_i \geq 0 \text{ für } i \in I^* \quad (B)$$

Lemma von Farkas:

- (B) ist auch eine notwendige Bedingung, damit (A) keine Lösung hat

Satz (Lemma von Farkas): Für beliebige Vektoren $\mathbf{a}_1, \dots, \mathbf{a}_k$ und $\mathbf{g}$ ist die Menge

$$S = \{\mathbf{s} \mid \mathbf{s}^\top \mathbf{g} < 0 \text{ und } \mathbf{s}^\top \mathbf{a}_i \leq 0 \text{ for all } i = 1, \dots, k\}$$

genau dann leer, falls $\lambda_i > 0$ existieren, so dass

$$-\mathbf{g} = \sum_{i=1}^k \lambda_i \mathbf{a}_i .$$

- Äquivalent zum **Dualitätssatz** der **linearen Programmierung**

Anwendung hier: System

$$\mathbf{s}^\top \nabla f(\mathbf{x}^*) < 0, \quad \mathbf{s}^\top \nabla g_i(\mathbf{x}^*) \leq 0, \quad \forall i \in I^*$$

hat genau dann keine Lösung, wenn

$$-\nabla f(\mathbf{x}^*) = \sum_{i \in I^*} \lambda_i \cdot \nabla g_i(\mathbf{x}^*), \quad \text{mit } \lambda_i \geq 0 \text{ für } i \in I^*$$

Notwendige Bedingung für (reguläres) lokales Minimum bei x^* :

$$-\nabla f(x^*) = \sum_{i \in I^*} \lambda_i \cdot \nabla g_i(x^*), \quad \text{mit } \lambda_i \geq 0 \text{ für } i \in I^*$$

Karush-Kuhn-Tucker Satz: Sei x^* ein reguläres, lokales Minimum. Dann existieren Multiplikatoren λ_i^* , derart dass

$$\nabla f(x^*) + \sum_{i=1}^p \lambda_i^* \nabla g_i(x^*) = 0$$

$$\lambda_i^* \geq 0, \quad i = 1, \dots, p$$

$$\lambda_i^* g_i(x^*) = 0, \quad i = 1, \dots, p$$

Zur Erinnerung:

- Funktion $f: \mathbb{R}^n \rightarrow \mathbb{R}$ heisst konvex, falls für alle $x, y \in \mathbb{R}^n$ und $\lambda \in [0,1]$:

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y)$$

- Eine Menge $S \subseteq \mathbb{R}^n$ heisst, falls für alle $x, y \in S$ und $\lambda \in [0,1]$:

$$\lambda x + (1 - \lambda)y \in S$$

Der Schnitt zweier konvexer Mengen ist konvex:

$$S_1, S_2 \text{ konvex} \implies S_1 \cap S_2 \text{ konvex}$$

Konvexe Nebenbedingung:

$$\text{Funktion } g_i(x) \text{ konvex} \implies \{x : g_i(x) \leq 0\} \text{ ist konvex}$$

Konvexe Minimierung:

$$\min_{x \in S} f(x), \quad \text{wobei Fkt. } f(x) \text{ und Menge } S \text{ konvex sind}$$

Konvexe Minimierung: **lokales Minimum $x' \in S$ ist auch globales Minimum:**

- Globales Minimum $x^* \in S$
- Konvexität von S : Verbindungslinie in S
- Konvexität von f : Betrachte $z = \lambda x' + (1 - \lambda)x^*$ auf Verbindungslinie:

$$f(z) = f(\lambda x' + (1 - \lambda)x^*) \leq \lambda f(x') + (1 - \lambda)f(x^*) < f(x')$$

Karush-Kuhn-Tucker Satz: Sei x^* ein reguläres, lokales Minimum. Dann existieren Multiplikatoren λ_i^* , derart dass

$$\nabla f(x^*) + \sum_{i=1}^p \lambda_i^* \nabla g_i(x^*) = 0$$

$$\lambda_i^* \geq 0, \quad i = 1, \dots, p$$

$$\lambda_i^* g_i(x^*) = 0, \quad i = 1, \dots, p$$

Hinreichende Karush-Kuhn-Tucker Bedingungen

Satz: Falls $f(x), g_1(x), \dots, g_p(x)$ **konvex** sind, dann ist jeder Punkt x^* , welcher die Karush-Kuhn-Tucker Bedingungen erfüllt, ein globales Minimum.

Satz: Falls $f(x), g_1(x), \dots, g_p(x)$ **konvex** sind, dann ist jeder Punkt x^* , welcher die Karush-Kuhn-Tucker Bedingungen erfüllt, ein globales Minimum.

Eigenschaft konvexer Funktionen:

$$\forall x', x^*: f(x') \geq f(x^*) + \nabla f(x^*)^\top (x' - x^*)$$

Satz: Falls $f(x), g_1(x), \dots, g_p(x)$ **konvex** sind, dann ist jeder Punkt x^* , welcher die Karush-Kuhn-Tucker Bedingungen erfüllt, ein globales Minimum.

Betrachte Karush-Kuhn-Tucker Punkt x^* und ein beliebiges $x' \neq x^*$:

Konvexität:

$$\begin{aligned} f(x') &\geq f(x^*) + \nabla f(x^*)^\top (x' - x^*) \\ \forall i: g_i(x') &\geq g_i(x^*) + \nabla g_i(x^*)^\top (x' - x^*) \end{aligned}$$

$$f(x') \geq f(x') + \sum_{i=1}^p \lambda_i^* g_i(x')$$

Satz: Falls $f(x), g_1(x), \dots, g_p(x)$ **konvex** sind, dann ist jeder Punkt x^* , welcher die Karush-Kuhn-Tucker Bedingungen erfüllt, ein globales Minimum.

Betrachte Karush-Kuhn-Tucker Punkt x^* und ein beliebiges $x' \neq x^*$:

$$f(x') \geq f(x^*) + \left[\nabla f(x^*) + \sum_{i=1}^p \lambda_i^* \nabla g_i(x^*) \right]^\top (x' - x^*) + \sum_{i=1}^p \lambda_i^* g_i(x^*)$$

Punkt x^* erfüllt Karush-Kuhn-Tucker Bedingungen:

$$\nabla f(x^*) + \sum_{i=1}^p \lambda_i^* \nabla g_i(x^*) = 0$$

$$\lambda_i^* \geq 0 \text{ und } \lambda_i^* g_i(x^*) = 0, \quad i = 1, \dots, p$$

Ist $x^* = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}$ Optimallösung des folgenden Minimierungsproblems?

$$\min f(x) = -x_1 - x_2$$

$$g_1(x) = x_1^2 - x_2 \leq 0$$

$$g_2(x) = x_1^2 + x_2^2 - 1 \leq 0$$

Gegeben: Funktion $f, h_1, h_2, \dots, h_m: \mathbb{R}^n \rightarrow \mathbb{R}$

Ziel:

$$\min_{x \in \mathbb{R}^n} f(x)$$

$$\text{wobei } h_1(x) = 0, h_2(x) = 0, \dots, h_m(x) = 0 \\ g_1(x) \leq 0, g_2(x) \leq 0, \dots, g_p(x) \leq 0$$

Notwendige Bedingung für (reguläres) lokales Minimum x^* :

Falls $s^\top \nabla h_i(x^*) = 0$ für $i = 1, \dots, m$ und $s^\top \nabla g_i(x^*) \leq 0$ für $i \in I^*$, dann gilt
 $s^\top \nabla f(x^*) \geq 0$.

Farkas-Lemma: Es gibt Multiplikatoren λ_i^* ($i = 1, \dots, m$) und $\mu_i^* \geq 0$ ($i \in I^*$),

$$\nabla f(x^*) + \sum_i^m \lambda_i^* \nabla h_i(x^*) + \sum_{i \in I^*} \mu_i^* \nabla g_i(x^*) = 0.$$