7b. Nichtlineare Optimierung
Optimalitätsbedingungen

Optimierung SoSe 2020

Dr. Alexey Agaltsov

Plan

- Optimalitätsbedingungen
- Regularitätsvoraussetzungen
- Dualität für konvexe Programme

Mathematisches Programm

$$\begin{aligned} \text{Minimiere} \quad & f(x) \ \text{ über } x \in \mathbb{R}^n \\ \text{u.d.N.} \quad & g_i(x) \leq 0, \ i = 1, \dots, m_U \\ & h_i(x) = 0, \ i = 1, \dots, m_G \end{aligned}$$

$$f, g_i, h_i \in C^1(\mathbb{R}^n)$$

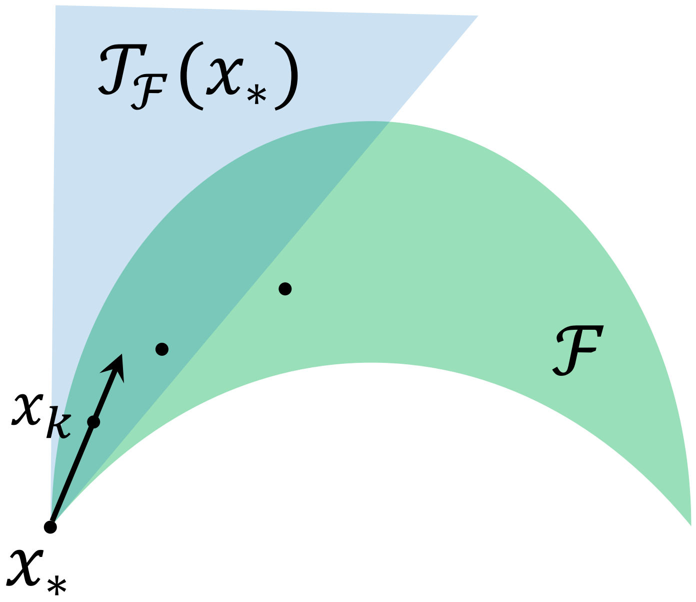
- Setze $\mathcal{F} = \{x \in \mathbb{R}^n : g_i(x) \leq 0 \ i = 1, \dots, m_U, h_i(x) = 0 \ i = 1, \dots, m_G\}$
- Angenommen, x_* ist eine optimale Lösung
- Wie kann man x_* charakterisieren und bestimmen?

Tangentialkegel

- $d \in \mathbb{R}^n$ heißt Tangentialvektor zu $\mathcal{F}$ in x_* falls:

$$\exists (x_k) \subseteq \mathcal{F} \text{ mit } x_k = x_* + d/k + o(1/k),\ k \to \infty$$

- Tangentialvektoren bilden Tangentialkegel $\mathcal{T}_{\mathcal{F}}(x_*)$



Lemma 7.4. Optimalitätsbedingung

Sei x_* ein lokales Minimum von f auf $\mathcal{F}$. Dann $-\nabla f(x_*)^T d \leq 0 \;\; \forall d \in \mathcal{T}_\mathcal{F}(x_*)$

Beweis.

Sei $d \in \mathcal{T}_\mathcal{F}(x_*)$ und $x_k = x_* + d/k + o(1/k) \in \mathcal{F}$

$$f(x_*) \leq f(x_k) \qquad \textit{für alle } k \geq K$$

$$= f(x_* + d/k + o(1/k))$$

$$= f(x_*) + \nabla f(x_*)^T d/k + o(1/k)$$

subtrahiere $f(x_*)$
multipliziere mit k
$k \to \infty$

$$-\nabla f(x_*)^T d \leq 0$$

$\square$

Linearisierter Tangentialkegel

$$\mathcal{A}(x_*) = \{i = 1, \dots, m_U\colon\ g_i(x_*) = 0\} \qquad \textit{Aktive Menge bei } x_*$$

- Sei $d \in \mathcal{T}_{\mathcal{F}}(x_*)$ sodass $\exists (x_k) \in \mathcal{F}$ mit $x_k = x_* + d/k + o(1/k)$

$$\underbrace{h_i(x_k)}_{=0} = \underbrace{h_i(x_*)}_{=0} + \nabla h_i(x_*)^T d/k + o(1/k),\ i = 1, \dots, m_G \ \Rightarrow\ \nabla h_i(x_*)^T d = 0$$

Linearisierter Tangentialkegel

$$\mathcal{A}(x_*) = \{i = 1, \dots, m_U\colon\ g_i(x_*) = 0\} \qquad \textit{Aktive Menge bei } x_*$$

- Sei $d \in \mathcal{T}_{\mathcal{F}}(x_*)$ sodass $\exists (x_k) \in \mathcal{F}$ mit $x_k = x_* + d/k + o(1/k)$

$$\underbrace{h_i(x_k)}_{=0} = \underbrace{h_i(x_*)}_{=0} + \nabla h_i(x_*)^T d/k + o(1/k),\ i = 1, \dots, m_G \quad \Rightarrow \quad \nabla h_i(x_*)^T d = 0$$

$$\underbrace{g_i(x_k)}_{\leq 0} = \underbrace{g_i(x_*)}_{=0} + \nabla g_i(x_*)^T d/k + o(1/k),\ i \in \mathcal{A}(x_*) \quad \Rightarrow \quad \nabla g_i(x_*)^T d \leq 0$$

- Also $\mathcal{T}_{\mathcal{F}}(x_*) \subseteq \mathcal{T}_{\mathcal{F}}^{lin}(x_*)$, wobei $\mathcal{T}_{\mathcal{F}}^{lin}(x_*)$ ist der linearisierte Tangentialkegel:

$$\mathcal{T}_{\mathcal{F}}^{lin}(x_*) = \{d\colon\ \nabla g_i(x_*)^T d \leq 0\ \ i \in \mathcal{A}(x_*),\ \nabla h_i(x_*)^T d = 0\ \ i = 1, \dots, m_G\}$$

Abadie Constraint Qualification (Abadie CQ)

- Sei die *Abadie Constraint Qualification* $\mathcal{T}_{\mathcal{F}}(x_*) = \mathcal{T}_{\mathcal{F}}^{lin}(x_*)$ erfüllt

Lemma 7.4

$$-\nabla f(x_*)^T d \leq 0 \text{ für alle } d \text{ mit } \nabla g_i(x_*)^T d \leq 0, i \in \mathcal{A}(x_*), \nabla h_i(x_*)^T d = 0, i = 1, \dots, m_G$$

Farkas Lemma

$$-\nabla f(x_*) = \sum_{i \in \mathcal{A}(x_*)} \underbrace{\mu_i^*}_{\geq 0} \nabla g_i(x_*) + \sum_{i=1}^{m_G} \underbrace{\eta_i^*}_{\geq 0} \nabla h_i(x_*) + \sum_{i=1}^{m_G} \underbrace{\theta_i^*}_{\geq 0} \big(-\nabla h_i(x_*)\big)$$

Abadie Constraint Qualification (Abadie CQ)

- Sei die *Abadie Constraint Qualification* $\mathcal{T}_{\mathcal{F}}(x_*) = \mathcal{T}_{\mathcal{F}}^{lin}(x_*)$ erfüllt

Lemma 7.4

$$-\nabla f(x_*)^T d \leq 0 \text{ für alle } d \text{ mit } \nabla g_i(x_*)^T d \leq 0, i \in \mathcal{A}(x_*), \nabla h_i(x_*)^T d = 0, i = 1, \ldots, m_G$$

Farkas Lemma

$$-\nabla f(x_*) = \sum_{i \in \mathcal{A}(x_*)} \underset{\geq 0}{\mu_i^*} \nabla g_i(x_*) + \sum_{i=1}^{m_G} \underset{\geq 0}{\eta_i^*} \nabla h_i(x_*) + \sum_{i=1}^{m_G} \underset{\geq 0}{\theta_i^*} \left(-\nabla h_i(x_*)\right)$$

$\lambda_i^* := \eta_i^* - \theta_i^*$

$\mu_i^* := 0$ für $i \in \mathcal{A}(x_*)^c$

$$\nabla f(x_*) + \sum_{i=1}^{m_U} \overset{\geq 0}{\mu_i^*} \nabla g_i(x_*) + \sum_{i=1}^{m_G} \overset{\in \mathbb{R}}{\lambda_i^*} \nabla h_i(x_*),$$

$$\mu_i^* g_i(x_*) = 0,\ i = 1, \ldots, m_U$$ *Komplementaritätsbedingung*

KKT-Punkt

- Wir bezeichnen $(x, \mu, \lambda) \in \mathbb{R}^{n \times m_U \times m_G}$ als KKT-Punkt (nach Karush-Kuhn-Tucker) falls die KKT-Bedingungen gelten:

$$\nabla f(x) + \sum_{i=1}^{m_U} \mu_i \nabla g_i(x) + \sum_{i=1}^{m_G} \lambda_i \nabla h_i(x) = 0$$

$$g_i(x) \leq 0,\ h_i(x) = 0\ \ \forall i$$

$$\mu \geq 0$$

$$\mu_i g_i(x) = 0,\ i = 1, \dots, m_U$$

$$\nabla_x \mathcal{L}(x, \mu, \lambda) = 0$$

$$\mathcal{L}(x, \mu, \lambda) = f(x) + \sum_{i=1}^{m_U} \mu_i g_i(x) + \sum_{i=1}^{m_G} \lambda_i h_i\,(x)$$

Lagrange-Funktion

Satz 7.5. Optimalitätsbedingungen

$$\begin{aligned} \text{Minimiere } & f(x) \text{ über } x \in \mathbb{R}^n \\ \text{u.d.N. } & g_i(x) \leq 0,\ i = 1, \dots, m_U \\ & h_i(x) = 0,\ i = 1, \dots, m_G \end{aligned} \qquad f, g_i, h_i \in C^1(\mathbb{R}^n)$$

- Sei x_* ein lokales Minimum und sei die Abadie CQ bei x_* erfüllt:

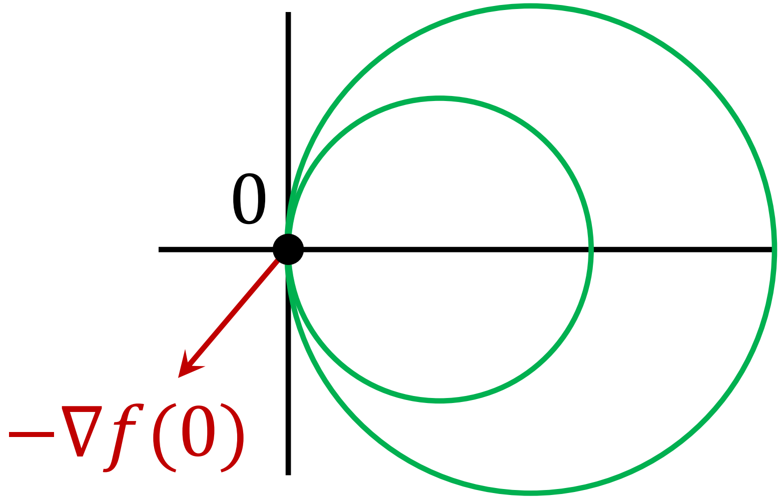
$$\mathcal{T}_{\mathcal{F}}(x_*) = \mathcal{T}_{\mathcal{F}}^{lin}(x_*)$$

- Dann gibt es (μ_*, λ_*) sodass (x_*, μ_*, λ_*) ein KKT-Punkt ist

Lagrange-Multiplikatoren

Beispiel 7.6. Problem ohne Lagrange-Multiplikatoren

$$\begin{aligned} \text{Minimiere} \quad & f(x) = x_1 + x_2 \\ \text{u.d.N.} \quad & h_1(x) = (x_1 - 1)^2 + x_2^2 - 1 = 0 \\ & h_2(x) = (x_1 - 2)^2 + x_2^2 - 4 = 0 \end{aligned}$$



- Es gibt keine λ_1, λ_2 so dass:

$$\nabla f(0) + \lambda_1 \nabla h_1(0) + \lambda_2 \nabla h_2(0) = 0$$

- $\mathcal{T}_{\mathcal{F}}(0) = \{0\} \neq \mathcal{T}_{\mathcal{F}}^{lin}(0) = \{0\} \times \mathbb{R}$

Plan

- Optimalitätsbedingungen
- Regularitätsvoraussetzungen
- Dualität für konvexe Programme

Regularitätsvoraussetzungen

$$
\begin{aligned}
\text{Minimiere} \quad & f(x) \quad \text{über } x \in \mathbb{R}^n \\
\text{u.d.N.} \quad & g_i(x) \le 0, \ i = 1, \dots, m_U \\
& h_i(x) = 0, \ i = 1, \dots, m_G
\end{aligned}
\qquad f, g_i, h_i \in C^1(\mathbb{R}^n)
$$

- Sei x_* eine optimale Lösung

Engl. Constraint Qualifications (CQ)

- Regularitätsvoraussetzungen garantieren, dass es Lagrange-Multiplikatoren (μ_*, λ_*) gibt, sodass (x_*, μ_*, λ_*) ein KKT-Punkt ist

Lineare NB (Satz 5.12)

Abadie CQ (Satz 7.5)

LICQ (Satz 7.7)

Slater-Bedingung (Satz 7.8)

Satz 7.7. LICQ (Linear Independence CQ)

- Sei $x_* \in \mathcal{F}$ und $\mathcal{A}(x_*)$ die aktive Menge bei x_*:

$$\mathcal{A}(x_*) = \{i = 1, \dots, m_U : g_i(x_*) = 0\}$$
$$= \{i_1, \dots, i_k\}$$

- Angenommen, x_* erfüllt die LICQ:

$\nabla g_{i_1}(x_*), \dots, \nabla g_{i_k}(x_*), \nabla h_1(x_*), \dots, \nabla h_{m_G}(x_*)$ sind linear unabhängig

- Dann erfüllt x_* die Abadie CQ:

$$\mathcal{T}_{\mathcal{F}}(x_*) = \mathcal{T}_{\mathcal{F}}^{lin}(x_*)$$

Beweis

$$h(x) = \begin{bmatrix} h_1(x) \\ \vdots \\ h_{m_G}(x) \end{bmatrix}$$

$$g_{\mathcal{A}(x_*)}(x) := \begin{bmatrix} g_{i_1}(x) \\ \vdots \\ g_{i_k}(x) \end{bmatrix}, \quad Dg_{\mathcal{A}(x_*)}(x) = \begin{bmatrix} \nabla g_{i_1}(x)^T \\ \vdots \\ \nabla g_{i_k}(x)^T \end{bmatrix}, \quad Dh(x) = \begin{bmatrix} \nabla h_1(x)^T \\ \vdots \\ \nabla h_{m_G}(x)^T \end{bmatrix}$$

$$F(x) := \begin{bmatrix} g_{\mathcal{A}(x_*)}(x) \\ h(x) \end{bmatrix}$$

x_* erfüllt LICQ

$$DF(x_*) = \begin{bmatrix} Dg_{\mathcal{A}(x_*)}(x_*) \\ Dh(x_*) \end{bmatrix}$$ hat vollen Zeilenrang

Beweis

Sei $d \in \mathcal{T}_{\mathcal{F}}^{lin}(x_*)$ und $R(x,t) \coloneqq F(x) - tDF(x_*)d$

$D_x R(x_*, 0) = DF(x_*)$ hat vollen Zeilenrang und $R(x_*, 0) = 0$

Satz von der impliziten Funktion

$\exists \varepsilon > 0, \exists y(t) \in C^1([0,\varepsilon), \mathbb{R}^n)$: $\quad R(y(t), t) \equiv 0$ und $y(0) = x_*$

Beweis

Sei $d \in \mathcal{T}_{\mathcal{F}}^{lin}(x_*)$ und $R(x,t) := F(x) - tDF(x_*)d$

$D_x R(x_*,0) = DF(x_*)$ hat vollen Zeilenrang und $R(x_*,0) = 0$

Satz von der impliziten Funktion

$\exists \varepsilon > 0, \exists y(t) \in C^1([0,\varepsilon),\mathbb{R}^n)$: $R(y(t),t) \equiv 0$ und $y(0) = x_*$

$g_{\mathcal{A}(x_*)^c}(x_*) < 0$

Ableitung in $t = 0$

$g_{\mathcal{A}(x_*)}(y(t)) = tDg_{\mathcal{A}(x_*)}(x_*)^T d \leq 0$

$h(y(t)) = tDh(x_*)^T d = 0$

$d \in \mathcal{T}_{\mathcal{F}}^{lin}(x_*)$

$y(t) = x_* + td + o(t)$

$\exists \eta \in (0,\varepsilon)$: $y(t) \in \mathcal{F}\ \forall t \in [0,\eta]$

Def.

$d \in \mathcal{T}_{\mathcal{F}}(x_*)$

$\square$

Satz 7.8. Slater-Bedingung

$$\begin{aligned} \text{Minimiere } \; & f(x) \; \text{ über } x \in \mathbb{R}^n \\ \text{u.d.N. } \; & g_i(x) \leq 0, i = 1, \dots, m_U \\ & Ax = b \end{aligned}$$

$f, g_i \in C^1(\mathbb{R}^n)$ konvex

- Angenommen, die Slater-Bedingung erfüllt ist:

Es gibt $\bar{x} \in \mathbb{R}^n$ *so, dass* $g_i(\bar{x}) < 0, i = 1, \dots, m_U$ *und* $A\bar{x} = b$

Slater-Punkt

- Dann ist für alle $x_* \in \mathcal{F}$ die Abadie CQ erfüllt:

$$\mathcal{T}_{\mathcal{F}}(x_*) = \mathcal{T}_{\mathcal{F}}^{lin}(x_*)$$

Beweis

$$\mathcal{T}_{\mathcal{F}}(x_*) \subseteq \mathcal{T}_{\mathcal{F}}^{lin}(x_*) \text{ im Allgemeinen}$$

$$\mathcal{T}_{\mathcal{F}}^{<}(x_*) \coloneqq \{d : Dg_{\mathcal{A}(x_*)}(x_*)d < 0, Dh(x_*)^T d = 0\}$$

Wir beweisen, dass unter der Slater-Bedingung:

1. $\mathcal{T}_{\mathcal{F}}^{lin}(x_*) \subseteq \overline{\mathcal{T}_{\mathcal{F}}^{<}(x_*)}$
2. $\mathcal{T}_{\mathcal{F}}^{<}(x_*) \subseteq \mathcal{T}_{\mathcal{F}}(x_*)$

Somit erhalten wir, dass $\mathcal{T}_{\mathcal{F}}(x_*) = \mathcal{T}_{\mathcal{F}}^{lin}(x_*)$

Beweis

Slater-Punkt $\bar{x}$: $g(\bar{x}) < 0$ *und* $A\bar{x} = b$

Behauptung: $\mathcal{T}_{\mathcal{F}}^{lin}(x_*) \subseteq \overline{\mathcal{T}_{\mathcal{F}}^{<}(x_*)}$ *für alle* $x_* \in \mathcal{F}$

Sei $d \in \mathcal{T}_{\mathcal{F}}^{lin}(x_*)$ und setze:

$$\bar{d}(t) = d + t(\bar{x} - x_*),\ \ t > 0$$

$$A\bar{d}(t) = Ad + t(A\bar{x} - Ax_*) = 0$$

Beweis

Slater-Punkt $\bar{x}$: $g(\bar{x}) < 0$ *und* $A\bar{x} = b$

Behauptung: $\mathcal{T}_{\mathcal{F}}^{lin}(x_*) \subseteq \overline{\mathcal{T}_{\mathcal{F}}^{<}(x_*)}$ *für alle* $x_* \in \mathcal{F}$

Sei $d \in \mathcal{T}_{\mathcal{F}}^{lin}(x_*)$ und setze:

$$\bar{d}(t) = d + t(\bar{x} - x_*),\ \ t > 0$$

$$A\bar{d}(t) = Ad + t(A\bar{x} - Ax_*) = 0$$

$$i \in \mathcal{A}(x_*):\ \nabla g_i(x_*)^T \bar{d}(t) = \nabla g_i(x_*)^T d + t\nabla g_i(x_*)^T(\bar{x} - x_*)$$

Satz 2.17

$$\leq \underbrace{\nabla g_i(x_*)^T d}_{\leq 0} + t\big(\underbrace{g_i(\bar{x})}_{< 0} - \underbrace{g_i(x_*)}_{= 0}\big) < 0$$

$$\bar{d}(t) \in \mathcal{T}_{\mathcal{F}}^{>}(x_*) \text{ und } \bar{d}(t) \to d \text{ für } t \to +0$$

$$\Rightarrow\ d \in \overline{\mathcal{T}_{\mathcal{F}}^{<}(x_*)}$$

Beweis

Behauptung: $\mathcal{T}_{\mathcal{F}}^{<}(x_*) \subseteq \mathcal{T}_{\mathcal{F}}(x_*)$ *für alle* $x_* \in \mathcal{F}$

Sei $d \in \mathcal{T}_{\mathcal{F}}^{<}(x_*)$ und setze $x(t) = x_* + td$

$$g_i\big(x(t)\big) = \overbrace{g_i(x_*)}^{=0} + t\nabla g_i(x_*)^T d + o(t) \qquad i \in \mathcal{A}(x_*)$$

$$= t\big(\underbrace{\nabla g_i(x_*)^T d}_{<0} + o(1)\big)$$

$$\exists \varepsilon > 0\colon\ g_i\big(x(t)\big) < 0\ \forall t \in (0, \varepsilon)$$

$$Ax(t) = Ax_* = b$$

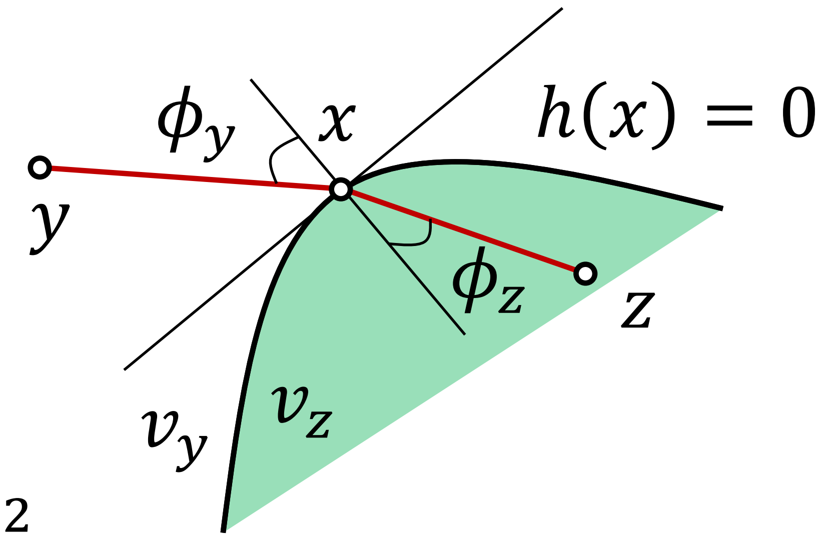
$x(t)$ ist zulässig $\forall t \in (0, \varepsilon)$

$$\Rightarrow d \in \mathcal{T}_{\mathcal{F}}(x_*)$$

$\square$

Aufgabe 7.9. Snelliussches Brechungsgesetz

- Die Kurve $h(x) = 0$ mit $h \in C^1(\mathbb{R}^2)$ trennt zwei Medien mit Lichtgeschwind. v_y bzw. v_z



- Ein Lichtstrahl reist von y nach z entlang des Pfads der minimalen Laufzeit

$$\text{Minimiere} \quad \frac{\|x-y\|_2}{v_y} + \frac{\|x-z\|_2}{v_z} \quad \text{über } x \in \mathbb{R}^2$$

$$\text{u.d.N.} \quad h(x) = 0$$

- Beweisen Sie die Formel: $\dfrac{\sin(\phi_y)}{v_y} = \dfrac{\sin(\phi_z)}{v_z}$

Plan

- KKT-Bedingungen
- Regularitätsvoraussetzungen
- Dualität für konvexe Programme

Konvexes Programm

Minimiere $f(x)$ über $x \in \mathbb{R}^n$ $\quad f \in C^1(\mathbb{R}^n)$ konvex

u.d.N. $a_i^T x = b_i,\ i = 1, \dots, m_G$ $\quad a_i \in \mathbb{R}^n, b_i \in \mathbb{R}$

$g_i(x) \leq 0,\ i = 1, \dots, m_U$ $\quad g_i \in C^1(\mathbb{R}^n)$ konvex

Lagrange-Funktion

$$\mathcal{L}(x, \mu, \lambda) = f(x) + \sum_{i=1}^{m_U} \underbrace{\mu_i}_{\geq 0} g_i(x) + \sum_{i=1}^{m_G} \lambda_i (b_i - a_i^T x)$$

Das duale Problem

$$G(\mu, \lambda) := \inf_{x \in \mathbb{R}^n} \mathcal{L}(x, \mu, \lambda)$$

Die duale Funktion

Maximiere $G(\mu, \lambda)$ über $\mu \in \mathbb{R}^{m_U}, \lambda \in \mathbb{R}^{m_G}$

u.d.N. $\mu \geq 0$

Satz 7.10. Dualität für konvexe Programme

Minimiere $f(x)$ über $x \in \mathbb{R}^n$	$f \in C^1(\mathbb{R}^n)$ konvex
u.d.N. $a_i^T x = b_i,\ i = 1, \dots, m_G$	$a_i \in \mathbb{R}^n, b_i \in \mathbb{R}$
$g_i(x) \leq 0,\ i = 1, \dots, m_U$	$g_i \in C^1(\mathbb{R}^n)$ konvex

- Angenommen, die Slater-Bedingung ist erfüllt
- Sei x_* eine optimale Lösung und μ_*, λ_* die zugehörigen Lagrange-Multiplikatoren, sodass (x_*, μ_*, λ_*) ein KKT-Punkt ist (Sätze 8.2, 8.4)
- Dann ist μ_*, λ_* eine optimale Lösung des dualen Problems und die optimalen Werte im Primalen und im Dualen übereinstimmen

Beweis

Sei (x_*, μ_*, λ_*) ein KKT-Punkt

$$\nabla_x \mathcal{L}(x_*, \mu_*, \lambda_*) = 0$$

$\mathcal{L}(\cdot, \mu_*, \lambda_*)$ ist konvex

$$\mathcal{L}(x_*, \mu_*, \lambda_*) \leq \mathcal{L}(x, \mu_*, \lambda_*) \quad \forall x \in \mathbb{R}^n$$

$$\mathcal{L}(x_*, \mu_*, \lambda_*) = f(x_*) + \sum_{i=1}^{m_U} \overbrace{\mu_{*,i} g_i(x_*)}^{=0} + \sum_{i=1}^{m_G} \overbrace{\lambda_{*,i}(b_i - a_i^T x_*)}^{=0}$$

Beweis

Sei (x_*, μ_*, λ_*) ein KKT-Punkt

$$\nabla_x \mathcal{L}(x_*, \mu_*, \lambda_*) = 0$$

$\mathcal{L}(\cdot, \mu_*, \lambda_*)$ ist konvex

$$\mathcal{L}(x_*, \mu_*, \lambda_*) \leq \mathcal{L}(x, \mu_*, \lambda_*) \;\; \forall x \in \mathbb{R}^n$$

$$\begin{aligned}
\mathcal{L}(x_*, \mu_*, \lambda_*) &= f(x_*) + \sum_{i=1}^{m_U} \overbrace{\mu_{*,i} g_i(x_*)}^{=0} + \sum_{i=1}^{m_G} \overbrace{\lambda_{*,i}(b_i - a_i^T x_*)}^{=0} \\
&\geq f(x_*) + \sum_{i=1}^{m_U} \underbrace{\mu_i g_i(x_*)}_{\leq 0} + \sum_{i=1}^{m_G} \underbrace{\lambda_i (b_i - a_i^T x_*)}_{=0} \quad \forall \mu \geq 0, \forall \lambda \\
&= \mathcal{L}(x_*, \mu, \lambda)
\end{aligned}$$

$\Rightarrow (x_*, \mu_*, \lambda_*)$ ist ein Sattelpunkt von $\mathcal{L}$

$\Rightarrow (\mu_*, \lambda_*)$ ist dual optimal und optimale Werte übereinstimmen (Satz 5.14) $\square$

Satz 7.11. Hinreichende Optimalitätsbedingung

Minimiere $f(x)$ über $x \in \mathbb{R}^n$	$f \in C^1(\mathbb{R}^n)$ konvex
u.d.N. $a_i^T x = b_i, \ i = 1, \dots, m_G$	$a_i \in \mathbb{R}^n, b_i \in \mathbb{R}$
$g_i(x) \leq 0, \ i = 1, \dots, m_U$	$g_i \in C^1(\mathbb{R}^n)$ konvex

- Angenommen, die Slater-Bedingung ist erfüllt
- Sei (x_*, μ_*, λ_*) ein KKT-Punkt
- Dann ist x_* eine optimale Lösung des Primalen und (μ_*, λ_*) eine optimale Lösung des Dualen. Außerdem ist die starke Dualität erfüllt

Beweis

Sei (x_*, μ_*, λ_*) ein KKT-Punkt und sei x zulässig

$$f(x) \geq f(x_*) + \nabla f(x_*)^T (x - x_*)$$

$$= f(x_*) - \sum_{i=1}^{m_U} \mu_{*,i} \nabla g_i(x_*)^T (x - x_*) + \sum_{i=1}^{m_G} \overbrace{\lambda_{*,i}\, a_i^T (x - x_*)}^{=0}$$

Beweis

Sei (x_*, μ_*, λ_*) ein KKT-Punkt und sei x zulässig

$$f(x) \geq f(x_*) + \nabla f(x_*)^T (x - x_*)$$

$$= f(x_*) - \sum_{i=1}^{m_U} \underbrace{\mu_{*,i} \nabla g_i(x_*)^T (x - x_*)}_{\leq \mu_{*,i}(g_i(x) - g_i(x_*)) = \mu_{*,i} g_i(x) \leq 0} + \sum_{i=1}^{m_G} \overbrace{\lambda_{*,i}\, a_i^T (x - x_*)}^{=0}$$

Komplementaritätsbedingung

$$\Rightarrow \; f(x) \geq f(x_*)$$

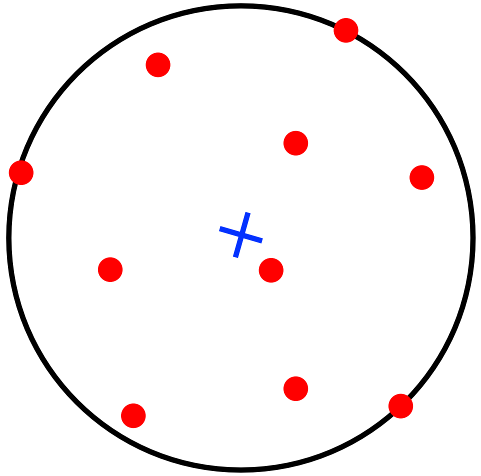
$\Rightarrow x_*$ ist eine optimale Lösung des primalen Problems

μ_*, λ_* ist eine optimale Lösung des Dualen nach Satz 7.10

$\square$

Kleinste Umschließende Kugel

- Seien $x_1, \ldots, x_m \in \mathbb{R}^n$ gegeben
- Finde die kleinste Kugel $B_{R_*}(x_*) = \{x\colon\ \|x - x_*\|_2 \leq R_*\}$, die alle Punkte $x_1, \ldots, x_m$ enthält
- Optimaler Standort einer Notfall-Einrichtung



Aufgabe 7.12. Kleinste Umschließende Kugel

$$\text{Minimiere } \gamma \text{ über } \gamma > 0, x \in \mathbb{R}^n$$
$$\text{u.d.N. } \|x_i - x\|_2^2 \leq \gamma, \; i = 1, \dots, m$$

$$\gamma = R^2$$

- Formulieren Sie das duale Problem und klassifizieren Sie sie
- Zeigen Sie, dass die Slater-Bedingung erfüllt ist
- Sei μ_* eine optimale Lösung des Dualen. Zeigen Sie, dass die optimale Lösung des primalen ist gegeben durch:

$$x_* = \sum_{i=1}^{m} \mu_{*,i} x_i$$

ähnlich zu SVN

Zusammenfassung

- Optimalitätsbedingungen
- Regularitätsvoraussetzungen
- Dualität für konvexe Programme

Nächstes Video

- 7c. Nichtlineare Optimierung: Algorithmen