Vorlesung
Einführung in die Mathematische Optimierung

(Wintersemester 2016/17)

Kapitel 4: Dualität und Konische Optimierung

Volker Kaibel

Otto-von-Guericke Universität Magdeburg

(Version vom 22. November 2016)

Gliederung

Das Setup

(Beliebiges) Minimierungsproblem

$$\min\{f(x) \mid x \in X\}$$

$$X = \{x \in X_0 \mid g_i(x) \leq 0 \text{ für alle } i \in [m], h_i(x) = 0 \text{ für alle } i \in [p]\}$$

Daten

- $\varnothing \neq X_0 \subseteq \mathbb{R}^n$ beliebig
- $f, g_1, \ldots, g_m, h_1, \ldots, h_p : \mathbb{R}^n \to \mathbb{R}$ beliebig

Lagrange-Funktion

Definition 4.1

Die Abbildung $L : X_0 \times \mathbb{R}_+^m \times \mathbb{R}^p \to \mathbb{R}$ mit

$$L(x, \lambda, \mu) = f(x) + \sum_{i=1}^{m} \lambda_i g^{(i)}(x) + \sum_{i=1}^{p} \mu_i h^{(i)}(x)$$

heißt die **Lagrangefunktion** zu $(f, X_0, (g_i)_{i \in [m]}, (h_i)_{i \in [p]})$.

Lagrange-primales Problem

Definition 4.2

Die **Lagrange-primale Funktion** ist $L_P : X_0 \to \mathbb{R} \cup \{\infty\}$ mit

$$L_P(x) = \sup\{L(x, \lambda, \mu) \,|\, (\lambda, \mu) \in \mathbb{R}_+^m \times \mathbb{R}^p\}\,.$$

Bemerkung 4.3

Für alle $x \in X_0$ *gilt* $L_P(x) = \begin{cases} f(x) & \textit{falls } x \in X \\ \infty & \textit{sonst} \end{cases}.$

Also ist $\min\{f(x) \,|\, x \in X\}$ *äquivalent zum* **Lagrange-primalen Problem**

$$\min\{L_P(x) \,|\, x \in X_0\}\,.$$

Lagrange-duales Problem

Definition 4.4

Die **Lagrange-duale Funktion** ist $L_D : \mathbb{R}^m_+ \times \mathbb{R} \to \mathbb{R} \cup \{-\infty\}$ mit

$$L_D(\lambda, \mu) = \inf\{L(x, \lambda, \mu) \,|\, x \in X_0\}\,.$$

Das **Lagrange-duale Problem** ist

$$\max\{L_D(\lambda, \mu) \,|\, (\lambda, \mu) \in \mathbb{R}^m_+ \times \mathbb{R}^p\}\,.$$

Bemerkung 4.5

Die Lagrange-duale Funktion $L_D : \mathbb{R}^m_+ \times \mathbb{R}^p \to \mathbb{R}$ *ist konkav.*

Schwache Lagrange-Dualität

Satz 4.6

Der Optimalwert des Lagrange-primalen Problems ist nach unten durch den Optimalwert des Lagrange-dualen Problems beschränkt. Gilt

$$L_D(\lambda^\star, \mu^\star) = L_P(x^\star)\,,$$

für $x^\star \in X_0$ *und* $(\lambda^\star, \mu^\star) \in \mathbb{R}^m_+ \times \mathbb{R}^p$*, so sind*

- $x^\star$ *Optimallösung des Lagrange-primalen Problems und*
- $(\lambda^\star, \mu^\star)$ *Optimallösung des Lagrange-dualen Problems.*

Sattelpunkte der Lagrange-Funktion

Bemerkung 4.7

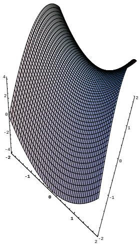
Ein Punkt $(x^\star, \lambda^\star, \mu^\star) \in X_0 \times \mathbb{R}^m_+ \times \mathbb{R}^p$ *erfüllt genau dann*

$$L_D(\lambda^\star, \mu^\star) = L_P(x^\star)\,,$$

wenn er ein **Sattelpunkt** *der Lagrangefunktion* L *ist, d.h., wenn*

$$L(x, \lambda^\star, \mu^\star) \geq L(x^\star, \lambda^\star, \mu^\star) \geq L(x^\star, \lambda, \mu)$$

für alle $(x, \lambda, \mu) \in X_0 \times \mathbb{R}^m_+ \times \mathbb{R}^p$ *gilt.*



Sattelpunkte und KKT

Lemma 4.8

Ist $\varnothing \neq X_0 \subseteq \mathbb{R}^n$ *konvex, und sind die Funktionen* $f, g_1, \ldots, g_m$ *differenzierbar und konvex und* $h_1, \ldots, h_p$ *affin, so ist* $(x^\star, \lambda^\star, \mu^\star) \in X_0 \times \mathbb{R}^m_+ \times \mathbb{R}^p$ *genau dann ein Sattelpunkt der Lagrangefunktion, wenn* $x^\star \in X$ *ist und die KKT-Bedingungen*

$$\operatorname{grad}_{x^\star} f + \sum_{i=1}^{m} \lambda_i^\star \operatorname{grad}_{x^\star} g_i + \sum_{i=1}^{p} \mu_i^\star \operatorname{grad}_{x^\star} h_i \in -\mathrm{N}_{x^\star}(X_0)$$

und

$$\lambda_i^\star = 0 \quad \textit{für alle } i \in [m] \textit{ mit } g_i(x^\star) < 0$$

erfüllt sind.

Starke Lagrange-Dualität

Satz 4.9

Seien $f, g_1, \ldots, g_m : \mathbb{R}^n \to \mathbb{R}$ *konvex und differenzierbar,* $h_1, \ldots, h_p : \mathbb{R}^n \to \mathbb{R}$ *affin und* $\emptyset \neq X_0 \subseteq \mathbb{R}^n$ *eine konvexe Menge, so dass* $(X_0, (g_i)_{i\in[m]}, (h_i)_{i\in[p]})$ *ein reguläres Tripel bilden, und*

$$X = \{x \in X_0 \mid g_i(x) \leq 0 \text{ für alle } i \in [m], h_i(x) = 0 \text{ für alle } i \in [p]\} .$$

Ein Punkt $x^\star \in X$ *ist genau dann Optimallösung von* $\min\{f(x) \mid x \in X\}$, *wenn es* $(\lambda^\star, \mu^\star) \in \mathbb{R}^m_+ \times \mathbb{R}^p$ *gibt, so dass*

$$L_D(\lambda^\star, \mu^\star) = f(x^\star)$$

für die zugehörige Lagrange-duale Funktion L_D *gilt.*

Konische Optimierungsprobleme

Konische Minimierungsprobleme

$$\min\{\langle c,x\rangle \mid Ax \le b, Cx = d, x \in K\},$$

- $X_0 = K \subseteq \mathbb{R}^n$ abgeschlossener konvexer Kegel
- $f(x) = \langle c,x\rangle$ linear
- $g_i(x) = \langle A_{i,\star},x\rangle - b_i$ affin ($i \in [m]$)
- $h_i(x) = \langle C_{i,\star},x\rangle - d_i$ affin ($i \in [p]$)

Konische Maximierungsprobleme

$$\max\{\langle c,x\rangle \mid Ax \le b, Cx = d, x \in K\}$$

Primal-duale Paare konischer Optimierungsprobleme

Definition 4.10

Die zu den konischen Optimierungsproblemen

$$\min\{\langle c, x\rangle \,|\, Ax \le b, Cx = d, x \in K\}$$

und

$$\max\{\langle c, x\rangle \,|\, Ax \le b, Cx = d, x \in K\}$$

dualen konischen Optimierungsprobleme sind

$$\max\{\langle -b, \lambda\rangle + \langle -d, \mu\rangle \,|\, A^\mathsf{T}\lambda + C^\mathsf{T}\mu + u = -c, (\lambda, \mu, u) \in \tilde{K}\}$$

bzw.

$$\min\{\langle b, \lambda\rangle + \langle d, \mu\rangle \,|\, A^\mathsf{T}\lambda + C^\mathsf{T}\mu + u = c, (\lambda, \mu, u) \in \tilde{K}\}$$

mit

$$\tilde{K} = \mathbb{R}^m_+ \times \mathbb{R}^p \times K^\circ.$$

Lagrange-Dual

Das duale konische Optimierungsproblem ist jeweils äquivalent zum Lagrange-dualen Problem.

Das Duale des Dualen

Bemerkung 4.11

Die zu

$$\max\{\langle -b,\lambda\rangle + \langle -d,\mu\rangle \mid A^{\mathsf{T}}\lambda + C^{\mathsf{T}}\mu + u = -c, (\lambda,\mu,u) \in \tilde{K}\}$$

und

$$\min\{\langle b,\lambda\rangle + \langle d,\mu\rangle \mid A^{\mathsf{T}}\lambda + C^{\mathsf{T}}\mu + u = c, (\lambda,\mu,u) \in \tilde{K}\}$$

(mit $\tilde{K} = \mathbb{R}^m_+ \times \mathbb{R}^p \times K^\circ$) dualen konischen Optimierungsprobleme sind äquivalent zu

$$\min\{\langle c,x\rangle \mid Ax \le b, Cx = d, x \in K\}$$

bzw.

$$\max\{\langle c,x\rangle \mid Ax \le b, Cx = d, x \in K\}.$$

Schwache konische Dualität

Satz 4.12

Bei zueinander dualen konischen Optimierungsproblemen ist der Optimalwert des Maximierungsproblems höchstens so groß wie der Optimalwert des Minimierungsproblems. Stimmen für zwei zulässige Lösungen der beiden Probleme die Zielfunktionswerte überein, so sind diese Lösungen also jeweils Optimallösungen.

Reguläre konische Optimierungsprobleme

Definition 4.13

Ein konisches Optimierungsproblem, dessen zulässige Menge durch $Ax \leq b$, $Cx = d$ und $x \in K$ (mit einem abgeschlossenen konvexen Kegel K, $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, $C \in \mathbb{R}^{p \times n}$ und $d \in \mathbb{R}^p$) beschrieben wird, ist **regulär**, wenn

1. K polyedrisch ist oder
2. ein $x^{(s)} \in \operatorname{int}(K)$ existiert mit $Ax^{(s)} \leq b$ und $Cx^{(s)} = d$ oder
3. $p = 0$ ist und ein $x^{(s)} \in K$ mit $Ax^{(s)} < b$ existiert.

Starke konische Dualität

Satz 4.14

Für jedes Paar von zueinander dualen konischen Optimierungsproblemen gilt:

1. *Ist eins der beiden konischen Optimierungsprobleme regulär, und besitzt es eine Optimallösung, so hat auch das andere eine Optimallösung und die Optimalwerte der beiden Probleme stimmen überein.*
2. *Sind beide konischen Optimierungsprobleme regulär, so besitzt das eine genau dann eine Optimallösung, wenn das andere eine Optimallösung hat, und in dem Fall stimmen die Optimalwerte der beiden Probleme überein.*

Möglichkeiten für reguläre Paare

Status primal vs. Status dual

	Optimallösung	unbeschränkt	unzulässig
Optimallösung	**Werte gleich**	unmöglich	unmöglich
unbeschränkt	unmöglich	unmöglich	**möglich**
unzulässig	unmöglich	**möglich**	**möglich**

Vorsicht

Ein allgemeines (reguläres) konisches Optimierungsproblem kann zulässig und beschränkt sein und trotzdem keine Optimallösung haben.
Bei *linearen* Optimierungsproblemen kann das aber nicht auftreten.

Duale Paare linearer Optimierungsproblemen

Maximierungsproblem

$$
\begin{array}{lccccccc}
\max & \langle c_1, x_1\rangle & + & \langle c_2, x_2\rangle & + & \langle c_3, x_3\rangle & & \\
 & A_{1,1}x_1 & + & A_{1,2}x_2 & + & A_{1,3}x_3 & \leq & b_1 \\
 & A_{2,1}x_1 & + & A_{2,2}x_2 & + & A_{2,3}x_3 & \geq & b_2 \\
 & A_{3,1}x_1 & + & A_{3,2}x_2 & + & A_{3,3}x_3 & = & b_3 \\
 & & & & & (x_1, x_2, x_3) & \in & \mathbb{R}_+^{n_1} \times \mathbb{R}_-^{n_2} \times \mathbb{R}^{n_3}
\end{array}
$$

Minimierungsproblem

$$
\begin{array}{lccccccc}
\min & \langle b_1, y_1\rangle & + & \langle b_2, y_2\rangle & + & \langle b_3, y_3\rangle & & \\
 & A_{1,1}^T y_1 & + & A_{2,1}^T y_2 & + & A_{3,1}^T y_3 & \geq & c_1 \\
 & A_{1,2}^T y_1 & + & A_{2,2}^T y_2 & + & A_{3,2}^T y_3 & \leq & c_2 \\
 & A_{1,3}^T y_1 & + & A_{2,3}^T y_2 & + & A_{3,3}^T y_3 & = & c_3 \\
 & & & & & (y_1, y_2, y_3) & \in & \mathbb{R}_+^{m_1} \times \mathbb{R}_-^{m_2} \times \mathbb{R}^{m_3}
\end{array}
$$

Dualisieren von LPs

Regeln

Bedingung	Variable
Gleichung	nicht vorzeichenbeschränkte Variable
$\leq$-Ungleichung (MAX)	nicht-negative Variable (MIN)
$\leq$-Ungleichung (MIN)	nicht-positive Variable (MAX)
$\geq$-Ungleichung (MAX)	nicht-positive Variable (MIN)
$\geq$-Ungleichung (MIN)	nicht-negative Variable (MAX)

LP-Dualität

Satz 4.15 (Schwacher Dualitätssatz)

Bei zueinander dualen linearen Optimierungsproblemen ist der Optimalwert des Maximierungsproblems höchstens so groß wie der Optimalwert des Minimierungsproblems. Stimmen für zwei zulässige Lösungen der beiden Probleme die Zielfunktionswerte überein, so sind diese Lösungen also jeweils Optimallösungen.

Satz 4.16 (Starker Dualitätssatz)

Für jedes Paar von zueinander dualen linearen Optimierungsproblemen besitzt das eine genau dann eine Optimallösung, wenn das andere eine Optimallösung hat, und in dem Fall stimmen die Optimalwerte der beiden Probleme überein.

Ein Paar zueinander dualer SDPs

Semidefinites Optimierungsproblem:

Für $c \in \mathbb{S}^k$, $c^{(i)} \in \mathbb{S}^k$ für alle $i \in [m]$ und $d \in \mathbb{R}^p$:

$$\max\{\langle c, x\rangle \mid \langle c^{(i)}, x\rangle = d_i \text{ für alle } i \in [p], x \in \mathbb{S}^k_+\}$$

(Regulär, falls eine *positiv-definite* zulässige Matrix $x^{(s)}$ existiert.)

Dazu duales SDP:

$$\min\{\langle d, \mu\rangle \mid \sum_{i=1}^{p} \mu_i c^{(i)} - c \in \mathbb{S}^k_+, \mu \in \mathbb{R}^p\}.$$