Dualität bei konvexer Optimierung

Laslo Hunhold

Seminar zur Numerik I im SS 2016, Universität zu Köln
2016-05-02

Ziel dieses Vortrags: Theorie und Anwendung dualer Probleme

- Duales Lagrange-Problem
- Duales Wolfe-Problem
- Schwache und starke Lagrange- und Wolfe-Dualität
- Beispiele für duale Probleme

Verwendete Literatur:

[KRT] E. de Klerk, C. Roos, T. Terlaky, Nonlinear Optimization, Lecture Notes, University of Waterloo, 2006

Konvexes Optimierungsproblem und Lagrange-Funktion

Betrachte (NLOP) mit $I = \emptyset$ für die Zielfunktion $f : \mathcal{C} \to \mathbb{R}$:

$$\begin{aligned} \inf_x \quad & f(x) \\ \text{gemäß} \quad & g_j(x) \leq 0, \quad g_j : \mathcal{C} \to \mathbb{R}, \quad j \in J = \{1, \ldots, m\} \\ & x \in \mathcal{C} \subseteq \mathbb{R}^n. \end{aligned}$$

Konvexes Optimierungsproblem (KOP), falls $\mathcal{C}$ konvex und $f, g_1, \ldots, g_m$ differenzierbar konvex.

Definiere die Lagrange-Funktion als

$$L(x, y) := f(x) + \sum_{j=1}^{m} y_j g_j(x), \quad L : \mathcal{C} \times \mathbb{R}_+^m \to \mathbb{R},$$

die für ein (KOP) in x konvex ist.

Ein Sattelpunkt ist ein Tupel $(\bar{x}, \bar{y}) \in \mathcal{C} \times \mathbb{R}_+^m$ mit der Eigenschaft

$$\forall x \in \mathcal{C} : \forall y \in \mathbb{R}_+^m : L(\bar{x}, y) \leq L(\bar{x}, \bar{y}) \leq L(x, \bar{y}).$$

Karush-Kuhn-Tucker-Theorem

Sei (KOP) Slater-regulär, $\bar{x} \in \mathcal{F}_{KOP}$. Dann gilt

$$\left\{ \bar{x} \text{ optimale Lösung für (KOP)} \right\} \Leftrightarrow \left\{ \exists \bar{y} \in \mathbb{R}_+^m : (\bar{x}, \bar{y}) \text{ Sattelpunkt von } L(x, y) \right\}$$

Duales Lagrange-Problem

Sei ein (KOP) gegeben, definiere die duale Lagrange-Funktion $\psi(y) := \inf_{x \in \mathcal{C}}(L(x, y))$ und betrachte das zum (KOP) duale Lagrange-Problem (DLP)

$$\begin{aligned} \sup_{y} \quad & \psi(y) \\ \text{gemäß} \quad & y \in \mathbb{R}^m_+ . \end{aligned} \tag{DLP}$$

(DLP) ist konvex, selbst wenn $f, g_1, \ldots, g_m$ nicht konvex sind.

Definiere für $(\bar{x}, \bar{y}) \in \mathcal{F}_{KOP} \times \mathbb{R}^m_+$ die Dualitätslücke als $f(\bar{x}) - \psi(\bar{y})$. Es gilt

$$\psi(\bar{y}) = \inf_{x \in \mathcal{C}}(L(x, \bar{y})) \leq L(\bar{x}, \bar{y}) = f(\bar{x}) + \sum_{j=1}^{m} \bar{y}_j g_j(\bar{x}) \leq f(\bar{x}),$$

woraus folgt, daß die Dualitätslücke positiv ist.

Motivation

- $\psi(y)$ ist untere Schranke der Zielfunktion $f(x)$ des (KOP)
- Suche nach der größten untere Schranke der Zielfunktion $f(x)$ des (KOP)

Fragen

- Wann ist $\psi(\bar{y}) = f(\bar{x})$?
- $\bar{x}$ dann eine optimale Lösung des (KOP)?
- Zusammenhang zwischen dem (DLP) und der KKT-Theorie?
- Was gilt falls (KOP) Slater-regulär?

Theoreme über Lagrange-Dualität

Theorem über schwache Lagrange-Dualität

Sei $(\bar{x}, \bar{y}) \in \mathcal{F}_{KOP} \times \mathbb{R}^m_+$. Dann gilt

1. $\psi(\bar{y}) \leq f(\bar{x})$
2. $\psi(\bar{y}) = f(\bar{x}) \Leftrightarrow \inf_{x \in \mathcal{C}}(L(x, \bar{y})) = f(\bar{x})$.

Folgerungen

- $\left\{\psi(\bar{y}) = f(\bar{x})\right\} \Rightarrow \left\{ \begin{array}{l} \bar{x} \text{ optimale Lösung für (KOP)} \\ \bar{y} \text{ optimale Lösung für (DLP)} \end{array} \right\}$
- $\left\{\mathcal{C} = \mathbb{R}^n\right\} \wedge \left\{ \{f, g_1, \dots, g_m\} \subset \mathcal{C}^1(\mathbb{R}^n)\right\} \Rightarrow \left\{(\bar{x}, \bar{y}) \text{ ist ein KKT-Punkt für (KOP)}\right\}$

Theorem über starke Lagrange-Dualität

Sei (KOP) Slater-regulär, $\bar{x} \in \mathcal{F}_{KOP}$. Dann gilt

$$\left\{\bar{x} \text{ optimale Lösung für (KOP)}\right\} \Leftrightarrow \left\{\exists \bar{y} \in \mathbb{R}^m_+ : \begin{array}{l} \bar{y} \text{ optimale Lösung für (DLP)} \\ \psi(\bar{y}) = f(\bar{x}) \end{array} \right\}.$$

Dies folgt direkt aus dem Karush–Kuhn–Tucker-Theorem mit Korollar 2.31 in [KRT].

Bemerkungen

- $\left\{\text{(KOP) nicht Slater-regulär}\right\} \not\Rightarrow \left\{\psi(\bar{y}) < f(\bar{x})\right\}$
- (KOP) hat nicht immer eine optimale Lösung $\bar{x}$
- Große Ähnlichkeit zu Voraussetzungen von Karush-Kuhn-Tucker-Korollaren

Beispiel (Slater-Regularität und Dualitätslücke)

$$\begin{aligned} \inf_x \quad & f(x) = x \\ \text{gemäß} \quad & g(x) = x^2 \leq 0 \\ & x \in \mathbb{R} \end{aligned} \tag{KOP1}$$

Slater-Regularität

$$\left\{ \nexists \tilde{x} \in \mathrm{ri}(\mathbb{R}) = \mathbb{R} : g(\tilde{x}) = \tilde{x}^2 < 0 \right\} \Rightarrow \left\{ \text{(KOP1) nicht Slater-regulär} \right\}$$

Dualitätslücke

Betrachte die duale Lagrange-Funktion $\psi : \mathbb{R}_+ \to \mathbb{R}$ von (KOP1)

$$\psi(y) = \inf_{x \in \mathbb{R}} (x + yx^2) = \begin{cases} \psi(-\frac{1}{2y}) = -\frac{1}{4y} & y > 0 \\ \psi(-\infty) = -\infty & y = 0. \end{cases}$$

Das duale Lagrange-Problem lautet

$$\begin{aligned} \sup_y \quad & \psi(y) \\ \text{gemäß} \quad & y \in \mathbb{R}_+ \end{aligned}$$

Es gilt

$$\psi(\bar{y}) = \sup_{y \in \mathbb{R}_+} \psi(y) = \psi(\infty) = 0 = f(0) = f\left(-\frac{1}{2\bar{y}}\right) = f(\bar{x}),$$

womit die Dualitätslücke verschwindet.

Sei ein (KOP) gegeben mit $\mathcal{C} = \mathbb{R}^n$ und $\{f', g_1', \ldots, g_m'\} \subset \mathcal{C}^0(\mathcal{C})$ und betrachte das zum (KOP) duale Wolfe-Problem (DWP)

$$\begin{aligned} \sup_{x,y} \quad & L(x,y) \\ \text{gemäß} \quad & \nabla_x L(x,y) = 0_n \\ & (x,y) \in \mathcal{F}_{KOP} \times \mathbb{R}^m_+. \end{aligned} \tag{DWP}$$

Bemerkungen:

- Einschränkungen sind genau die Karush-Kuhn-Tucker-Bedingungen
- Falls $L(x,y)$ einen Sattelpunkt hat, sind (DLP) und (DWP) äquivalent:

Äquivalenz zu (DLP)

$$\left\{\forall y \in \mathbb{R}^m_+ : L(\cdot, y) \in \mathcal{C}^1(\mathbb{R}^n) \text{ konvex}\right\} \underset{\text{Fermat}}{\overset{\text{konvexer}}{\Rightarrow}} \left\{\forall (\bar{x}, y) \in \mathbb{R}^n \times \mathbb{R}^m_+ : \{\nabla_x L(\bar{x}, y) = 0\} \Leftrightarrow \{L(\bar{x}, y) \text{ min}\}\right\}$$

Duale Lagrange-Funktion des (DWP) im Sattelpunkt $(\bar{x}, y)$ ist

$$\psi(y) = \inf_{x \in \mathbb{R}^n} (L(x,y)) = L(\bar{x}, y)$$

(DWP) kann umgeschrieben werden in

$$\begin{aligned} \sup_{y} \quad & \psi(y) \\ \text{gemäß} \quad & y \in \mathbb{R}^m_+ \end{aligned}$$

Dies ist genau das (DLP).

Theoreme über Wolfe-Dualität

Theorem über schwache Wolfe-Dualität

Sei $\hat{x} \in \mathcal{F}_{KOP}$, $(\bar{x}, \bar{y}) \in \mathcal{F}_{DWP}$. Dann gilt

$$L(\bar{x}, \bar{y}) \leq f(\hat{x})$$

Beweis

$$\left\{\forall y \in \mathbb{R}_+^m : L(\cdot, y) \in \mathcal{C}^1(\mathbb{R}^n) \text{ konvex}\right\} \wedge \left\{\forall (\bar{x}, \bar{y}) \in \mathcal{F}_{DWP} : \nabla_x L(\bar{x}, \bar{y}) = 0\right\} \Rightarrow$$

$$\left\{\inf_{x \in \mathbb{R}^n} L(x, \bar{y}) = L(\bar{x}, \bar{y})\right\} \Leftrightarrow$$

$$\left\{\forall x \in \mathbb{R}^n : L(\bar{x}, \bar{y}) \leq L(x, \bar{y})\right\} \Rightarrow$$

$$\left\{\forall \hat{x} \in \mathcal{F}_{KOP} : L(\bar{x}, \bar{y}) \leq L(\hat{x}, \bar{y}) = f(\hat{x}) + \sum_{j=1}^m \bar{y}_j g_j(\hat{x}) \leq f(\hat{x})\right\}$$

Theorem über starke Wolfe-Dualität

Sei (KOP) Slater-regulär, $\bar{x} \in \mathcal{F}_{KOP}$. Dann gilt

$$\left\{\bar{x} \text{ optimale Lösung für (KOP)}\right\} \Leftrightarrow \left\{\exists \bar{y} \in \mathbb{R}_+^m : (\bar{x}, \bar{y}) \text{ optimale Lösung für (DWP)}\right\}.$$

Dies folgt direkt aus dem Karush–Kuhn–Tucker-Theorem mit Korollar 2.33 in [KRT].

Beispiel (Wolfe-Lösungsverfahren)

$$\begin{aligned}
\inf_{x} \quad & f(x) = x_1 + \exp(x_2) \\
\text{gemäß} \quad & 3x_1 - 2\exp(x_2) \geq 10 \\
& x_2 \geq 0 \\
& x \in \mathbb{R}^2
\end{aligned} \tag{KOP2}$$

Normalisierte Nebenbedingungen

$$\begin{aligned}
g_1(x) &= 10 - 3x_1 + 2\exp(x_2) && \leq 0 \\
g_2(x) &= -x_2 && \leq 0
\end{aligned}$$

Slater-Regularität

$$\left\{ \tilde{x} = (5,0) \in \mathbb{R}^2 : g_1(\tilde{x}) < 0 \wedge g_2(\tilde{x}) \leq 0 \right\} \Rightarrow \left\{ \text{(KOP2) Slater-regulär} \right\}$$

Lösung über duales Wolfe-Problem

$$\begin{aligned}
\sup_{x,y} \quad & L(x,y) = x_1 + \exp(x_2) + y_1(10 - 3x_1 + 2\exp(x_2)) - y_2 x_2 \\
\text{gemäß} \quad & \frac{\partial L(x,y)}{\partial x_1} = 1 - 3y_1 = 0 \\
& \frac{\partial L(x,y)}{\partial x_2} = \exp(x_2) + 2\exp(x_2)y_1 - y_2 = 0 \\
& (x,y) \in \mathcal{F}_{KOP2} \times \mathbb{R}^2_+
\end{aligned}$$

$$\to \left\{ y_1 = \frac{1}{3} \right\} \to \left\{ \frac{5}{3}\exp(x_2) - y_2 = 0 \right\} \to \left\{ \tilde{f}(x_2) = \frac{5}{3}\exp(x_2) - \frac{5}{3}x_2\exp(x_2) + \frac{10}{3} \right\} \to \left\{ \tilde{f}(\bar{x} = 0) = 5 \right\}$$

Beispiel (quadratisches konvexes Optimierungsproblem) (1)

Sei $Q \in \mathbb{R}^{n \times n}$ symmetrisch positiv semidefinit, $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$

$$\begin{aligned} \inf_{x} \quad & f(x) = \frac{1}{2} x^T Q x + c^T x \\ \text{gemäß} \quad & Ax \geq b \Leftrightarrow Ax - b \in \mathbb{R}^m_+ \\ & x \in \mathbb{R}^n_+ \end{aligned} \tag{QKOP}$$

Normalisierte Nebenbedingungen

$$g_j(x) = \begin{cases} (-a^j)^T x + b_j & j = 1, \dots, m \\ -x_{j-m} & j = m+1, \dots, m+n \end{cases} \leq 0$$

Lösung über duales Wolfe-Problem

$$\begin{aligned} \inf_{x,y,z} \quad & L(x,y,z) = \frac{1}{2} x^T Q x + c^T x + y^T(-Ax + b) + z^T(-x) \\ \text{gemäß} \quad & \nabla_x L(x,y,z) = \frac{1}{2}(Q^T + Q)x + c - A^T y - z \overset{Q^T = Q}{=} Qx + c - A^T y - z = 0_n \\ & (x,y,z) \in \mathcal{F}_{QKOP} \times \mathbb{R}^m_+ \times \mathbb{R}^n_+ \end{aligned}$$

Substituiere $c = -Qx + A^T y + z$ in $L(x,y,z)$:

$$\begin{aligned} \inf_{x,y,z} \quad & -\frac{1}{2} x^T Q x + b^T y \\ \text{gemäß} \quad & Qx + c - A^T y - z = 0_n \\ & (x,y,z) \in \mathcal{F}_{QKOP} \times \mathbb{R}^m_+ \times \mathbb{R}^n_+ \end{aligned}$$

Beispiel (quadratisches konvexes Optimierungsproblem) (2)

Eliminiere x mit Cholesky-Zerlegung von $Q = D^T D$ (Q symmetrisch positiv definit) und $d = Dx$

$$\begin{aligned} \inf_{y,d} \quad & -\frac{1}{2} d^T d + b^T y \\ \text{gemäß} \quad & D^T z + c - A^T y - z = 0_n \\ & (y, z, d) \in \mathbb{R}_+^m \times \mathbb{R}_+^n \times \mathbb{R}_+^n \end{aligned}$$

Die Optimalitätsbedingungen sind

- $y^T (Ax - b) = 0$
- $z^T x = 0$
- $d = Dx$

Konklusion

Klärung der Ausgangsfragen

Lagrange

- $\psi(\bar{y}) = f(\bar{x})$ bei Slater-Regularität und Existenz von optimaler Lösung für (KOP) $\bar{x}$
- $\bar{x}$ eine optimale Lösung des (KOP) genau bei Slater-Regularität und $\psi(\bar{y}) = f(\bar{x})$
- $(\bar{x}, \bar{y})$ KKT-Punkt falls $\{f, g_1, \dots, g_m\} \subset \mathcal{C}^1(\mathcal{C})$

Wolfe

- Keine Aussage zu $L(\bar{x}, \bar{y}) = f(\hat{x})$ möglich, da nichtkonvexes Problem
- $\bar{x}$ eine optimale Lösung des (KOP) genau bei Slater-Regularität und Existenz von $\bar{y}$, sodaß $(\bar{x}, \bar{y})$ optimale Lösung für (DWP)
- Einschränkungen sind genau die KKT-Bedingungen

Ausblick

- 2016-05-23: Kaczmarz-Algorithmus zur Lösung von Least-Squares-Problemen
  Kevin Scislak, Hamed Rafei
- 2016-05-30: Algorithmen für nichtrestringierte Optimierungsprobleme
  Ferhat Dogan, Felix Leikert, Frederic Tillmanns
- 2016-06-27: Algorithmen für restringierte Optimierungsprobleme
  Gina Göbel, Tobias Wagner und Marcel Freyschmidt