Summary of Lagrangian Duality (Part I)

Large Scale Optimization

Esteve Codina

Master's Degree in Statistics and Operational Research

UNIVERSITAT POLITÈCNICA
DE CATALUNYA
BARCELONATECH

February 2018

Contents

1. Basic Definitions and Properties

Basic Definitions

Reference Problem and Lagrangian

$$\begin{array}{ll} \text{Min}_x & f(x) \\ \text{s.t. :} & h(x) = 0 \mid \lambda \in \mathbb{R}^p \\ & g(x) \geq 0 \mid \mu \in \mathbb{R}^q_+ \\ & x \in X \end{array}$$

Associated Lagrangian function:

$$\mathcal{L}(x, \lambda, \mu) = f(x) - \lambda^\top h(x) - \mu^\top g(x)$$

Dual Lagrangian Function; Dual Lagrangian Problem

$$\begin{array}{lll} w(\lambda, \mu) \triangleq \text{Min}_x & \mathcal{L}(x, \lambda, \mu) \\ (\text{LD}_0) & x \in X \\ & \downarrow \end{array}$$

($S_{LD}(\lambda, \mu)$ the solution set)

$$\begin{array}{ll} \text{Max}_x & w(\lambda, \mu) \\ (\text{LD}) & (\lambda, \mu) \in D \end{array}$$

$$D \triangleq \{ (\lambda, \mu) \in \mathbb{R}^p \times \mathbb{R}^q_+ \mid w(\lambda, \mu) > -\infty \}$$

Usually, $D \equiv \mathbb{R}^p \times \mathbb{R}^q_+$

Properties of $w(\lambda, \mu)$

Weak Duality Theorem

Let $(\lambda, \mu) \in D$, x feasible primal (of problem (P)). Then $w(\lambda, \mu) \le f(x)$

- $\text{gap}(x, \lambda, \mu) \triangleq f(x) - w(\lambda, \mu) =$
$f(x) - f(\bar{x}) + \lambda^\top h(\bar{x}) + \mu^\top g(\bar{x}) \ge 0,\ \bar{x} \in \mathsf{S}_{LD}(\lambda, \mu)$.
- If (λ^*, μ^*) solves (LD) and $\bar{x} \in \mathsf{S}_{LD}(\lambda^*, \mu^*)$, then
$\text{gap}(\bar{x}, \lambda^*, \mu^*) = {\lambda^*}^\top h(\bar{x}) + {\mu^*}^\top g(\bar{x})$
- If $\text{gap}(\bar{x}, \lambda^*, \mu^*) = {\lambda^*}^\top h(\bar{x}) + {\mu^*}^\top g(\bar{x}) = 0$, then $\bar{x}$ solves (P)

Other properties. (See Bazaraa's textbook, Chapter 6)

- $w(\lambda, \mu)$ is concave on D
- The subgradient set of $w(\lambda, \mu)$ at a point $(\bar{\lambda}, \bar{\mu})$ is given by:

$$\partial w(\lambda, \mu) = \text{Hull}\left(\left\{ \begin{pmatrix} -h(x^*) \\ -g(x^*) \end{pmatrix} \,\middle|\, x^* \in \mathsf{S}_{LD}(\lambda, \mu) \right\}\right)$$

2. Types of Problems

Types of Problems

- **Convex Problems**. (including LP's) $f(x)$ convex, h affine, $g(x)$ locally convex at the solution (active constraints). **Have null duality gap**.
- Problems with $w(\lambda, \mu)$ **differentiable** also **have null duality gap**.
- **IP's, MILP's may have non-null duality gap**. Additional Lagrangian heuristics are needed to obtain a good solution.

Integer Programming Problems

Reference Problem and Lagrangian Dual

$$\begin{array}{rll} z^* = & \text{Min}_x & c^\top x \\ & \text{s.t.}: & Ax = b \mid \lambda \\ & & Bx = d \\ \text{(P)} & & x \in \mathbb{Z}^n \end{array}$$

$$\begin{array}{rll} z^*_{\text{LR}} = & \text{Min}_x & c^\top x \\ & \text{s.t.}: & Ax = b \mid \lambda \\ \text{(LR)} & & x \in X_{\text{LR}} \end{array}$$

$$\begin{array}{rll} w(\lambda) = & \text{Min}_x & c^\top x - \lambda^\top (Ax - b) \\ (\text{LD}_0) & & x \in X \end{array}$$

$$\begin{array}{rll} z^*_{\text{LD}} = & \text{Max}_\lambda & w(\lambda) \\ \text{(LD)} & & \lambda \in D \end{array}$$

$$X = \{\, x \in \mathbb{Z}^n \mid Bx = d \,\}$$
$$X_{\text{LR}} = \{\, x \in \mathbb{R}^n \mid Bx = d \,\}$$

Theorem

$$\begin{array}{llll} z^*_{\mathsf{LD}} = \mathsf{Max}_\lambda & w(\lambda) & = \ \mathsf{Min}_x & c^\top x \\ \text{(LD)} & \lambda \in D & \text{s.t.}: & Ax = b \\ & & & x \in \mathsf{Hull}(X) \end{array}$$

Consequences

$$z^*_{\mathsf{LR}} \overset{(a)}{\leq} z^*_{\mathsf{LD}} \overset{(b)}{\leq} z^*$$

- if $\mathsf{Hull}(X) = X_{\mathsf{RL}}$, then (a) becomes $=$
- if $\mathsf{Hull}(\{x \in X | Ax = b\}) = \mathsf{Hull}(X) \cap \{x \in \mathbb{R}^n | Ax = b\}$, then (b) becomes $=$
- In general $\mathsf{Hull}(X) \subset X_{\mathsf{LR}}$ and $z^*_{\mathsf{LR}} < z^*_{\mathsf{LD}} \leq z^*$.

  If $X = \{x \in \mathbb{Z}^2_+ | 2x_1 + 2x_2 \leq 5\}$, then
  $\mathsf{Hull}(X) = \{x \in \mathbb{R}^2_+ | x_1 + x_2 \leq 2\}$ and
  $X_{\mathsf{LR}} = \{x \in \mathbb{R}^2_+ | 2x_1 + 2x_2 \leq 5\}$

3. Solving (LD)

1st Order Optimality Conditions. (1st OOC)

Assume that for the problem (P), problem (LD) reduces to:

$$\begin{array}{ll} \text{Max}_x & w(\lambda,\mu) \\ \text{(LD)} & \mu \geq 0 \quad | \, \theta \geq 0 \end{array}$$

1st OOC. Assume w differentiable

$$\begin{array}{r} \nabla_\lambda w(\lambda^*,\mu^*) = 0 \\ -\nabla_\mu w(\lambda^*,\mu^*) = \theta \geq 0, \ \theta^\top \mu^* = 0, \ \mu^* \geq 0 \end{array}$$

1st OOC. Assume w non-differentiable

$$\begin{array}{l} \exists r \in \partial w(\lambda^*,\mu^*) \, \text{such that :} \\ \ r_\lambda = 0 \\ -r_\mu = \theta \geq 0, \ \theta^\top \mu^* = 0, \ \mu^* \geq 0 \end{array}$$

In variational form:

$$\partial w(\lambda^*,\mu^*) \in \mathsf{N}_{\mathbb{R}^p \times \mathbb{R}^q_+}(\lambda^*,\mu^*) \equiv \begin{cases} r_\lambda = 0 \\ r_\mu \leq 0, \ r_\mu^\top \mu^* = 0 \end{cases}$$

Dual Search Methods

Methods for solving the Lagrangian Dual Problem (LD)

$$\begin{array}{ll} \text{Max}_x & w(\lambda, \mu) \\ \text{(LD)} & (\lambda, \mu) \in D \end{array}$$

- Subgradient method

  (see `subgradient-algorithm.pdf`)
- Outer approximation method (Dantzig's cutting plane)

  (see `DANTZIG-CUTTING-PLANE1-AP8.pdf`)
- Bundle Methods (Lemarechal)
- Non-differentiable optimization methods

  (see `wolfe12.pdf`, `wolfe13.pdf`)