CSE203B Convex Optimization: Chapter 5 Duality

CK Cheng
Dept. of Computer Science and Engineering
University of California, San Diego

Chapter 5 Duality

- Primal and Dual Problem Mechanism
  - Primal Problem
  - Lagrangian Function
  - Lagrange Dual Problem
- Examples (Primal Dual Conversion Procedure) Examples
  - Linear Programming
  - Quadratic Programming
  - Conjugate Functions (Duality)
  - Entropy Maximization
- Interpretation (Duality) Theory
  - Saddle-Point Interpretation
  - Geometric Interpretation
  - Slater's Condition
  - Shadow-Price Interpretation
- KKT Conditions (Optimality Conditions) Optimality
- Sensitivity (Shadow-Price)
- Generalized Inequalities

Duality

Primal Problem (Feasible Solution)

$$\begin{aligned} &\min\ f_0(x) && x \in R^n \\ &s.t.\ f_i(x) \le 0 && i = 1,\dots,m \\ &\quad\ \ h_i(x) = 0 && i = 1,\dots,p \end{aligned}$$

domain D $= dom\ f_0 \cap_i\ dom\ f_i\ \cap_i dom\ h_i$

feasible set $E = \{x \mid f_i(x) \le 0, h_i(x) = 0\ \forall i\ \forall x \in D\}$

$E \subset D$

Opt: $x^*, p^* = f_0(x^*)$

Lagrangian: $L: R^n \times R^m \times R^p \to R$

$$L(x,\lambda,v) = f_0(x) + \sum_{i=1}^m \lambda_i f_i(x) + \sum_{i=1}^p v_i h_i(x)$$

λ_i, v_i: *Lagrange multiplier*, $\lambda_i, \in R_+, v_i \in R$.

Lagrange dual function

$g(\lambda, v) = \inf_{x \in D} L(x,\lambda,v)$ (x may not be feasible)

$$g(\lambda,v) = \min_{x \in D} L(x,\lambda,v)$$

$$= -\max_{x \in D} -L(x,\lambda,v)$$

$$= -\max_{x \in D} \underbrace{\underbrace{\underbrace{-f_0(x) - \sum \lambda_i f_i(x) - \sum v_i h_i(x)}_{\text{convex w.r.t } \lambda_i, v_i}}_{\text{convex w.r.t } \lambda_i, v_i.}}_{\text{concave w.r.t } \lambda_i, v_i}$$

Duality

Dual Problem (Infeasible Solution)

$max_{\lambda,v}\ g(\lambda,v)\quad s.t.\ \lambda \succcurlyeq 0$

1. $g(\lambda,v)$ *is concave*
2. $g(\lambda,v) \le p^*$ an optimal value where $\lambda \succcurlyeq 0$

Proof 1: By definition of $g(\lambda,v)$ and the convexity of pointwise max operation on convex functions.

Proof 2: For any feasible $\tilde{x}$ and $\lambda \succcurlyeq 0$

$f_0(\tilde{x}) \ge L(\tilde{x},\lambda,v)$ (Because $\sum \lambda_i f_i(\tilde{x}) + \sum v_i h_i(\tilde{x}) \le 0$)

$L(\tilde{x},\lambda,v) \ge g(\lambda,v)$ by definition of $g(\lambda,v)$

Thus $p^* = f_0(x^*) \ge g(\lambda,v)$

Example: Linear Programming

Prime: $\quad min\ c^T x$

$$subject\ to \begin{cases} Ax \preccurlyeq b \\ x \succcurlyeq 0 \end{cases} \Rightarrow \begin{cases} Ax - b \preccurlyeq 0 \\ -x \preccurlyeq 0 \end{cases}$$

Lagrangian

$$\begin{aligned} L(x,\lambda) &= c^T x + \lambda_I^T (Ax - b) - \lambda_{II}^T x \\ &= -\lambda_I^T b + (A^T \lambda_I - \lambda_{II} + c)^T x, \quad \lambda_I, \lambda_{II} \succcurlyeq 0 \\ g(\lambda) &= \inf_x L(x,\lambda) \end{aligned}$$

$$g(\lambda) = \begin{cases} -b^T \lambda_I, & A^T \lambda_I + c \geq 0\ (A^T \lambda_I - \lambda_{II} + c = 0) \rightarrow A^T\lambda_I + c = \lambda_{II} \geq 0 \\ -\infty, & otherwise \quad (A^T \lambda_I - \lambda_{II} + c \neq 0) \end{cases}$$

Dual:

$$\begin{aligned} &\max\ -\mathrm{b}^T \lambda_I\ (min\ b^T \lambda_I) \\ &subject\ to\ A^T \lambda_I + c \geq 0 \\ &\qquad \lambda_I \geq 0 \end{aligned}$$

Example: Linear Programming

Prime: $\quad \min[-1\ \ -1] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

$$\text{subject to } \begin{bmatrix} 1 & 3 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \preccurlyeq \begin{bmatrix} 3 \\ 2 \end{bmatrix}, x_1, x_2 \geq 0.$$

Dual: $\quad max -[3\ 2] \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix}$

$$\text{subject to } \begin{bmatrix} 1 & 1 \\ 3 & 0 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} + \begin{bmatrix} -1 \\ -1 \end{bmatrix} \succcurlyeq \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\lambda_1, \lambda_2 \geq 0$$

Results: $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 1/3 \end{bmatrix}, p^* = -\frac{7}{3}$

$$\begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} = \begin{bmatrix} 1/3 \\ 2/3 \end{bmatrix}, d^* = -\frac{7}{3}$$

Example: Linear Programming

$$\min c^T x$$
$$\text{subject to } Ax = b,\ x \geq 0,\ (or - x \leq 0)$$

Lagrangian: $L(x, \lambda, v) = c^T x + \lambda^T(-x) + v^T(Ax - b)$
$$= -b^T v + (c + A^T v - \lambda)^T x$$

Lagrange Dual: $g(\lambda, v) = \inf_x L(x, \lambda, v)$

1. If $A^T v - \lambda + c = 0 \quad \rightarrow \quad g(\lambda, v) = -b^T v$
2. Else $\quad \rightarrow \quad g(\lambda, v) = -\infty$

Properties:

1. g is linear on affine domain $\{(\lambda, v) | A^T v - \lambda + c = 0\}$, hence concave.

2. If $\lambda \succcurlyeq 0 \ \Rightarrow A^T v + c \succcurlyeq 0$

$$p^* \geq -b^T v \quad if \ A^T v + c \succcurlyeq 0$$

$$\boxed{\begin{aligned} &\max -b^T v \\ &\qquad A^T v + c \succcurlyeq 0 \end{aligned}} \quad \textbf{or} \quad \boxed{\begin{aligned} &\max\ b^T v \\ &\qquad A^T v \preccurlyeq c \end{aligned}}$$

Example: Quadratic Programming

$$\min x^T x$$
$$\text{subject to } Ax = b$$

Lagrangian:

$$L(x, v) = x^T x + v^T(Ax - b)$$

To minimize L over x, we set $\nabla_x L(x, v) = 2x + A^T v = 0$

$\Rightarrow x = -\frac{1}{2} A^T v$ (1)

Lagrange Dual Function:

$$g(v) = L\left(x = -\frac{1}{2} A^T v, v\right) = -\frac{1}{4} v^T A A^T v - b^T v$$

A concave function of v

Lower Bound Property: $p^* \geq -\frac{1}{4} v^T A A^T v - b^T v, \ \forall v$

To maximize $g(v)$, we set $\quad v = -2(AA^T)^{-1} b$

Thus, we have $g(v) = -\frac{1}{4} v^T A A^T v - b^T v = b^T (AA^T)^{-1} b$ (2)

(3) From (1) and (2), we have $\begin{cases} x^* = A^T (AA^T)^{-1} b \\ p^* = x^{*T} x^* = b^T (AA^T)^{-1} b \end{cases}$