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)
  - 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)
- Sensitivity (Shadow-Price) (Perturbation)
- Generalized Inequalities

Duality

Primal Problem (Solution x is feasible)

$\min\ f_0(x) \qquad x \in R^n$

$s.t.\ f_i(x) \le 0 \quad i = 1, \dots, m \qquad$ *domain D*

$\quad h_i(x) = 0 \quad i = 1, \dots, p \qquad = dom\ f_0 \cap_i dom\ f_i \cap_i dom\ h_i$

Notation: 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 (Solution x may not be feasible)

$$g(\lambda, v) = \inf_{x \in D} L(x, \lambda, v) = \min_x f_0(x) + \sum_i \lambda_i f_i(x) + \sum_i v_i h_i(x)$$

Dual Problem

$$max_{\lambda, v}\ g(\lambda, v) \quad s.t. \lambda \in R_+^m, v \in R^p$$

$\uparrow\uparrow$ Shadow Price.

$$-g(\lambda, v) = \max_x -\sum \lambda_i f_i(x) - \sum v_i h_i(x) - f_0(x)$$

$$= \max_x \underbrace{-\sum \lambda_i f_i(x) - \sum v_i h_i(x) - f_0(x)}_{\text{convex } (\lambda_i, v_i)}$$

convex(λ_i, v_i)

Duality

Dual Problem (Solution x may not be feasible)

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

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 $\sum \lambda_i f_i(\tilde{x}) + \sum v_i h_i(\tilde{x}) \le 0$)

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

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

Example: Linear Programming

$x \in R^n$.

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

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

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

$A^T\lambda_1 + c = \lambda_{II} \geq 0$

Dual:

$$\max\ -b^T\lambda_I\ (min\ b^T\lambda_I)$$
$$subject\ to\ A^T\lambda_I + c \geq 0$$
$$\lambda_I \geq 0$$

$-A^T\lambda_1 + \lambda_{II} = +c = +\nabla f_0(x)$

5

$K^* = \left\{ [-A^T\ +I]\begin{bmatrix}\lambda_1 \\ \lambda_{II}\end{bmatrix} \middle| \begin{bmatrix}\lambda_1 \\ \lambda_{II}\end{bmatrix} \geq 0 \right\}$

~~$K = \{x \mid A^T x \geq 0, +Ix \leq 0\}$~~

$K = \{x \mid -Ax \geq 0, Ix \geq 0\}$

Example: Linear Programming

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

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

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

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

Dual

$$\begin{cases} \min \ b^T\lambda \\ \text{s.t} \ -A^T\lambda - c \leq 0 \\ \quad -\lambda \leq 0 \end{cases}$$

$$L(\lambda, x) = b^T\lambda + x_I^T(-A^T\lambda - c) + x_{II}^T(-\lambda) \qquad x_I, x_{II} \geq 0$$

$$= (b - Ax_I - x_{II})\lambda - x_1^T c$$

$$g(x) = \min_\lambda L(\lambda, x) = \min (b - Ax_1 - x_{II})\lambda - x_1^T c$$

① $-Ax_1 - x_{II} + b = 0 \Rightarrow g(x) = -x_1^T c$

② $\neq 0 \Rightarrow g(x) \to -\infty$

Dual of Dual

Thus, we have

$$\max \ -c^T x_1$$

$$Ax_1 + x_{II} = b \qquad x_1, x_{II} \geq 0$$

is the primal

or $\min \ c^T x$

$$Ax \leq b$$

$$x \geq 0$$

Example: Linear Programming

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

$\min c^T x$
s.t. $Ax - b = 0, -x \leq 0$

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

$A^T v + c = \lambda \geq 0$

max $-b^T v$ $A^T v + c \succcurlyeq 0$	or	max $b^T v$ $A^T v \preccurlyeq c$

Example: Quadratic Programming

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

$x \in R^n \quad A \in R^{m \times n}$

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

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

$\longrightarrow \frac{1}{4} v^T A A^T v + v^T(A \frac{-1}{2} A^T v - b)$

$= \frac{-1}{4} v^T A A^T v - v^T b$

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

$\nabla_v g(v) = -\frac{1}{2} A A^T v - b$

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 $v = -2(AA^T)^{-1} b$

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 $\begin{cases} x^* = A^T (AA^T)^{-1} b \\ p^* = x^{*T} x^* = b^T (AA^T)^{-1} b \end{cases}$

Example: Quadratic Program

Quadratic Program

$$\min x^T P x \qquad P \in S_{++}^n$$
$$s.t.\ Ax \preccurlyeq b$$

Lagrange Dual Function:

$$g(\lambda) = \min_x x^T P x + \lambda^T(Ax - b)$$
$$= -\frac{1}{4}\lambda^T A P^{-1} A^T \lambda - b^T \lambda$$

Dual Problem:

$$\max -\frac{1}{4}\lambda^T A P^{-1} A^T \lambda - b^T \lambda$$
$$s.t.\ \lambda \succcurlyeq 0$$

Example: Quadratic Program (nonconvex prob.)

$$\min x^T A x + 2b^T x \qquad L(x,\lambda) = x^TAx + 2b^Tx + \lambda(x^Tx - 1)$$
$$s.t.\ x^T x \leq 1 \qquad A \in S^n, A \nsucccurlyeq 0$$

$$x^TAx = x^TA^Tx$$

Dual Function:

$$g(\lambda) = \min_x x^T(A + \lambda I)x + 2b^T x - \lambda$$

Unbounded below if $A + \lambda I \nsucccurlyeq 0$ or if $A + \lambda I \succcurlyeq 0$ & $b \notin R(A + \lambda I)$

Minimized by $x = -(A + \lambda I)^+ b$

Otherwise $g(\lambda) = -b^T(A + \lambda I)^+ b - \lambda$

$$x^TAx = \frac{1}{2}x^T(A+A^T)x$$

Dual Problem:

$$\begin{aligned} \max\ & -b^T(A + \lambda I)^+ b - \lambda \\ s.t.\ & A + \lambda I \succcurlyeq 0 \\ & b \in R(A + \lambda I) \end{aligned} \qquad \Bigg| \qquad \begin{aligned} \max\ & -t - \lambda \\ s.t.\ & \begin{bmatrix} A + \lambda I & b \\ b^T & t \end{bmatrix} \succcurlyeq 0 \end{aligned}$$

$$\begin{bmatrix} I & 0 \\ -((A + \lambda I)^+ b)^T & I \end{bmatrix} \begin{bmatrix} A + \lambda I & b \\ b^T & t \end{bmatrix} \begin{bmatrix} I & -(A + \lambda I)^+ b \\ 0 & I \end{bmatrix} \succcurlyeq 0$$

$$\begin{bmatrix} A + \lambda I & 0 \\ 0 & -b^T(A + \lambda I)^+ b + t \end{bmatrix} \succcurlyeq 0$$