# Convexe Optimization

ESSAKINE Amer

DM2

## Ex 1:

1) The lagrangian is given by:

$$L(x,\lambda,\nu) = c^T x - \lambda^T x + \nu^T (b - Ax)$$
$$= [c - \lambda - A^T\nu]^T x + \nu^T b$$

Hence the dual function is

$$g(\lambda,\nu) = \begin{cases} \nu^T b & \text{if } c - \lambda - A^T\nu = 0 \\ -\infty & \text{otherwise} \end{cases}$$

and the dual problem is

$$\begin{aligned} \max \quad & \nu^T b \\ \text{s.t} \quad & c - \lambda - A^T\nu = 0 \\ & \lambda \geq 0 \end{aligned}$$

which is equivalent to

$$\begin{aligned} \max \quad & \cancel{\nu^T b}\ b^T\nu \\ \text{s.t.} \quad & A^T\nu \leq c \end{aligned}$$

Hence the dual of (P) is (D)

**2.**

$$\max_y \; b^T y \qquad (D)$$
$$\text{s.t. } A^T y \le c$$

equivalent to
$$\min_x \; -b^T y$$
$$\text{s.t } A^T y \le c$$

The Lagrangian is given by:

$$L(y, \lambda) = -b^T y + \lambda^T (A^T y - c)$$
$$= [-b + A\lambda]^T y - \lambda^T c$$

Hence the dual function is

$$g(\lambda) = \begin{cases} \cancel{-b+A} \; -\lambda^T c & \text{if } -b + A\lambda = 0 \\ -\infty & \text{otherwise} \end{cases}$$

and the dual problem is

$$\max \; -\lambda^T c$$
$$\text{s.t. } A\lambda = b$$
$$\lambda \ge 0$$

which is equivalent to
$$\min \; c^T \lambda$$
$$\text{s.t. } A\lambda = b$$
$$\lambda \ge 0$$

Hence (P) is equivalent to the dual of (D)

(2

3 The Lagrangian is given by:

$$L(x, y, \lambda_1, \lambda_2, v) = c^T x - b^T y + \lambda_1^T [A^T y - c] - \lambda_2^T x + v^T (b - Ax)$$

$$L(x, y, \lambda_1, \lambda_2, v) = [c - \lambda_2 - A^T v] x + [-b + A\lambda_1]^T y - c^T \lambda_1 + b^T v$$

Hence the dual function is

$$g(\lambda_1, \lambda_2, v) = \begin{cases} -c^T \lambda_1 + b^T v & \text{if } c - \lambda_2 - A^T v = 0 \\ & \quad -b + A\lambda_1 = 0 \\ -\infty & \text{otherwise} \end{cases}$$

and the dual problem is

$$\begin{aligned} \max_{\lambda_1, \lambda_2, v} \quad & -c^T \lambda_1 + b^T v \\ \text{s.t.} \quad & c - \lambda_2 - A^T v = 0 \\ & -b + A\lambda_1 = 0 \\ & \lambda_1 \geq 0 \\ & \lambda_2 \geq 0 \end{aligned}$$

which is equivalent to

$$\begin{aligned} \min_{\lambda_1, v} \quad & c^T \lambda_1 - b^T v \\ \text{s.t.} \quad & A\lambda_1 = b \\ & \lambda_1 \geq 0 \\ & A^T v \leq c \qquad (\text{since } \lambda_2 = -A^T v + c) \end{aligned}$$

~~Hence the dual of~~

Hence (Self dual) is a ~~dual~~ self dual problem.

(3)

4) • Suppose that we solve (P), (D) which gives $\bar{x}$, $\bar{y}$

Then $c^T \bar{x} \leq b^T x^*$

and $-b^T \bar{y} \leq -b^T x^*$

Hence $c^T \bar{x} - b^T \bar{y} \leq c^T x^* - b^T x^*$

and since $(\bar{x}, \bar{y})$ is a feasible point for (Self dual) $\begin{cases} A\bar{x} = b \\ \bar{x} \geq 0 \\ A^T \bar{y} \leq c \end{cases}$

then $[x^*, y^*] = [\bar{x}, \bar{y}]$

• By the strong duality property of linear programs, both (P) and (D) have strong duality ((P) and (D) are feasible by feasibility of (Self dual))

→ By strong duality of (P).

$$c^T x^* \geq d^* = b^T y^* , \text{~~Hence~~}$$

since the dual of (P) is (D) whose optimal value si $b^T y^*$

Hence $c^T x^* - b^T y^* \geq 0$

→ By strong duality of (D):

$$b^T y^* \geq d^* = c^T x^*$$

Since (P) is the dual of (D)

Hence $c^T x^* - b^T y^* \leq 0$

We conclude that the optimal value of (Self Dual) is exactly 0 $(c^T x^* - b^T y^* = 0)$

(4)

## Ex2:

1) Let $f : x \longmapsto \|x\|_1$

Let $y \in \mathbb{R}^d$, $x \in \mathbb{R}^d$

$$y^T x - \|x\|_1 = \sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} |x_i|$$

$$= \sum_{i=1}^{n} x_i (y_i - \varepsilon_i^x)$$

where $\varepsilon_i^x = \begin{cases} 1 & \text{if } x_i \geq 0 \\ -1 & \text{otherwise} \end{cases}$

If there exists $i = 1, \dots, n$ $y_i > 1$.

the we can take $x_p = (0, \dots, \underset{\substack{\downarrow \\ i^{eme} \text{ position}}}{p}, 0, \dots, 0)$ $p \in \mathbb{N}$.

Hence $y^T x_p - f(x_p) = p \times (y_i - 1) \xrightarrow[p \mapsto +\infty]{} +\infty$

Similary, if there exists $i = 1, \dots, n$, $y_i < -1$.

$\cancel{y^T x_p}$ define $x_p = (0, \dots, 0, \underset{\substack{\downarrow \\ i^{eme} \text{ position}}}{-p}, 0, \dots, 0)$, $p \in \mathbb{N}$

$$y^T x_p - f(x_p) = -p(y_i + 1) \xrightarrow[p \mapsto +\infty]{} +\infty$$

Hence if $\|y\|_\infty > 1$, $f^*(y) = +\infty$

5

~~And if $\|y$~~

If $\|y\|_\infty \leq 1$.

$$y^T x - f(x) = \sum_{i=1}^{n} x_i ( y_i - \varepsilon_i^x )$$

If $x_i \geq 0$, $\varepsilon_i^x = 1$

and $x_i ( y_i - 1 ) \leq 0$

If $x_i \leq 0$, $\varepsilon_i^x = -1$

and $x_i ( y_i + 1 ) \leq 0$

Hence $y^T x - f(x) \leq 0$

and if we take $x = 0$ for example, we find equality

Hence $f^*(y) = \begin{cases} 0 & \text{if } \|y\|_1 \leq 1 \\ +\infty & \text{otherwise} \end{cases}$

K

## Ex3:

1) Suppose that we solved (Sep 2.) and let $[w^*, z^*]$ be an optimal solution,

denote $f(z,w) = \frac{1}{mT} 1^T z + \frac{1}{2}\|w\|_2^2$

$z \in \mathbb{R}^d$

Let $w$ be any vector in $\mathbb{R}^d$, and define $z$ by:

$i = 1, \dots, d \quad z_i = \max(0, 1 - y_i(w^T x_i))$

Then $\quad f(z,w) \geq f(z^*, w^*)$

but
$$f(z,w) = \frac{1}{mT} 1^T z + \frac{1}{2}\|w\|_2^2$$
$$= \frac{1}{T}\left[\frac{1}{m}\sum_{i=1}^{m} \ell(w, x_i, y_i) + \frac{T}{2}\|w\|_2^2\right]$$

and
$$f(z^*, w^*) = \frac{1}{T}\left[\frac{1}{m}\sum_{i=1}^{m} z_i + \frac{T}{2}\|w^*\|_2^2\right]$$
$$\geq \frac{1}{T}\left[\sum_{i=1}^{m} \max(1 - y_i(w^T x_i)) + \frac{T}{2}\|w^*\|_2^2\right]$$

since $z_i \geq 1 - y_i(w^T x_i)$ and $z_i \geq 0$
$$\geq \frac{1}{T}\left[\sum_{i=1}^{m} \ell(w^*, x_i, y_i) + \frac{T}{2}\|w^*\|_2^2\right]$$

Hence,

$\forall w \quad \frac{1}{m}\sum_{i=1}^{m} \ell(w, x_i, y_i) + \frac{T}{2}\|w\|_2^2 \geq \frac{1}{m}\sum_{i=1}^{m} \ell(w^*, x_i, y_i) + \frac{T}{2}\|w^*\|$

and $w^*$ is an optimal solution for (Sep 1.)

$\ominus$

The Lagrangian of (sep $\alpha$) is given by

$$\mathcal{L}(w, z, \lambda, \pi) = \frac{1}{nT} 1^T z + \frac{1}{2}\|w\|_2^2 + \sum_{i=1}^{m} \lambda_i [1 - y_i (w^T x_i) - z_i]$$

$$- \pi^T z$$

$$\mathcal{L}(w, z, \lambda, \pi) = \left[\frac{1}{nT} 1 - \lambda - \pi\right] z + \left[\frac{1}{2}\|w\|_2^2 - \sum_{i=1}^{m} \lambda_i y_i (w^T x_i)\right]$$

$$+ 1^T \lambda \cancel{- \pi^T z}$$

~~Hence~~ denote ~~$g(w)$~~

$$h(w) = \frac{1}{2}\|w\|_2^2 - \sum_{i=1}^{m} \lambda_i y_i (w^T x_i)$$

$h$ is differentiable and convex, and the gradient is:

$$\nabla h(w) = w - \sum_{i=1}^{m} \lambda_i y_i x_i$$

hence

$$w^* = \sum_{i=1}^{m} \lambda_i y_i x_i$$

and

$$h(w^*) = \frac{1}{2} \sum_{1 \le i, j \le m} \lambda_i \lambda_j y_i y_j x_i^T x_j$$

$$- \sum_{i=1}^{m} \lambda_i y_i \left(\sum_{j=1}^{m} \lambda_j y_j x_j\right)^T x_i$$

$$= \frac{1}{2} \sum_{1 \le i, j \le m} \lambda_i \lambda_j y_i y_j x_i^T x_j$$

$$- \sum_{1 \le i, j \le m} \lambda_i \lambda_j y_i y_j x_i^T x_j$$

$$= - \frac{1}{2} \sum_{1 \le i, j \le m} \lambda_i \lambda_j y_i y_j x_i^T x_j$$

(8

Hence the dual function is

$$g(\lambda,\pi) = \begin{cases} -\frac{1}{2}\sum_{1\leq i,j\leq m} \lambda_i \lambda_j y_i y_j x_i^T x_j + 1^T\lambda \cancel{+\pi} & \text{if } \frac{1}{mT}1 - \lambda - \pi = 0 \\ -\infty & \text{otherwise} \end{cases}$$

and the dual problem is

$$\max \quad -\frac{1}{2}\sum_{1\leq i,j\leq m} \lambda_i \lambda_j y_i y_j x_i^T x_j + 1^T\lambda$$

$$\text{s.t.} \quad \frac{1}{mT}1 - \lambda - \pi = 0$$

$$\lambda \geqslant 0$$

$$\pi \geqslant 0$$

which simplifies to:

$$\max \quad \sum_{i=1}^{m}\lambda_i - \frac{1}{2}\sum_{1\leq i,j\leq m} \lambda_i \lambda_j y_i y_j x_i^T x_j$$

$$\text{s.t.} \quad \lambda \geqslant 0$$

$$\frac{1}{mT}1 \geqslant \lambda$$

(9