NAIM Mouna HM2

Exercice 1: $c \in \mathbb{R}^d$, $b \in \mathbb{R}^n$, $A \in \mathbb{R}^{n \times d}$

$$(P) \quad \begin{cases} \min_x c^T x \\ \text{s.t } Ax = b \\ \quad x \geq 0 \end{cases}$$

$$(D) \quad \begin{cases} \max b^T y \\ \text{s.t } A^T y \leq c \end{cases}$$

1) Let's calculate the Lagrangian of (P):

$$\mathcal{L}(x, \lambda, \nu) = c^T x - \lambda^T x + \nu^T (Ax - b)$$

$$= -\nu^T b + (c - \lambda + A^T \nu) x$$

Then, the dual function g is:

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

Thus, the dual problem of (P) is:

$$\begin{cases} \max_{\lambda, \nu} -\nu^T b \, , \quad \text{s.t } \lambda \geq 0 \\ A^T \nu + c = \lambda \end{cases}$$

Let's $\mu = -y$ and eliminate λ:

$\mu = -y \Rightarrow A^T\mu + c = -A^Ty + c$ and $\lambda \geq 0$

$$\Rightarrow -A^Ty + c \geq 0$$

$$\Leftrightarrow A^Ty \leq c$$

Thus, $\begin{cases} \max\limits_y b^Ty \\ \text{s.t } A^Ty \leq c \end{cases}$

Conclusion: The dual problem of (P) is (D)

2 Let's calculate the Lagragian of (D):

$$\mathcal{L}(y, v) = -b^Ty + v^T(A^Ty - c)$$

$$= -v^Tc + (Av - b)^Ty$$

The dual function g is:

$$g(v) = \begin{cases} -v^Tc & \text{if } Av - b = 0 \\ -\infty & \text{otherwise} \end{cases}$$

$\Rightarrow$ The dual of the (D) problem is:

$$\begin{cases} \max\limits_v -v^Tc & \text{st } v \geq 0 \\ Av = b \end{cases}$$

3 Let's calculate the Lagrangian of (Self-Dual):

$$\mathcal{L}(x, y, \lambda, \mu, v) = c^T x - b^T y - \lambda^T x + \mu^T (Ax - b) + v^T (A^T y - c)$$

$$= -\mu^T b - v^T c + (c - \lambda + A^T \mu)^T x + (Av - b)^T y$$

The dual function is:

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

So the dual problem is:

$$\begin{cases} \max\limits_{\mu, v} -b^T \mu - c^T v \\ \text{s.t } \lambda, \mu \geq 0 \\ c + A^T \mu = \lambda \\ Av = b \end{cases}$$

We eliminate λ as previously and we change variables $x = v$ and $y = -\mu$

Then
$$\begin{cases} \max_{x,y} b^T y - c^T x \\ \text{s.t } x \geq 0 \\ A^T y \leq c \\ Ax = b \end{cases}$$

We have then:

$$\begin{cases} \min_{x,y} c^T x - b^T y \quad \text{st} \quad x \geq 0 \\ A^T y \leq 0 \quad \text{and} \quad Ax = b \end{cases}$$

We find then the original problem

4

Let's x', y' be the respective optimal solutions of (P) and (D)

Then for any feasible point x, y for (P) and (D), we have those inequalities:

$$\begin{cases} c^T x' \leq c^T x \\ b^T y' \geq b^T y \end{cases} \iff c^T x' - b^T y' \geq c^T x - b^T y$$

So $[x', y']$ is an optimal solution for Self-Dual.

Thus $[x', y'] = [x^*, y^*]$

By solving (P) and (D), we have the solution of Self-dual.

We have (D) is the dual of (P), and duality holds for linear programs:

$$p^* = d^* \iff c^T x^* = b^T y^*$$

$$\iff c^T x^* - b^T y^* = 0$$

So the optimal value of self-Dual is 0

Exercise 2:

1. By definition, we have:

$$f_*(y) = \sup_{x \in \mathbb{R}^d} (y^T x - \|x\|_1)$$

$$= \sup_{x \in \mathbb{R}^d} \sum_{i=1}^{d} y_i x_i - |x_i|$$

$$= \sum_{i=1}^{d} \sup_{x_i \in \mathbb{R}} (y_i x_i - |x_i|)$$

Let $i \in (1, \dots, d)$ and let's determine x_i maximizing each term of the sum

We separate two cases:

- If $x_i \geq 0$, $|x_i| = x_i$ and $y_i x_i - |x_i| = x_i (y_i - 1)$

$$\begin{cases} \text{If } y_i > 1 \text{ then } x_i = +\infty \\ \text{If } y_i \leq 1 \text{ then } x_i = 0 \end{cases}$$

- If $x_i < 0$: $y_i x_i - |x_i| = x_i (y_i + 1)$

$$\begin{cases} \text{If } y_i < -1 \text{ then } x_i = +\infty \\ \text{If } y_i \geq -1 \text{ then } x_i = 0 \end{cases}$$

The sum is finite if $|y_i| < 1$

then $\|y\|_\infty < 1$

Thus: $$f_*(y) = \begin{cases} 0 & \text{if } \|y\|_\infty \leq 1 \\ +\infty & \text{otherwise} \end{cases}$$

2 The Lagrangian of the problem is:

Let's $y = Ax - b$

$$\mathcal{L}(x, y, \lambda) = y^T y + \|x\|_1 + \lambda^T (Ax - b - y)$$

$$= y^T I y - \lambda^T y + \|x\|_1 + \lambda^T A x - \lambda^T b$$

•) On the one hand, $\mathcal{L}$ is a quadratic form with respect to y. It is minimized by

$y = \frac{1}{2}\lambda$

•) On the other hand,

$$\inf_x \left(\|x\|_1 - (-A^T\lambda)^T x\right) = -\sup_x \left((-A^T\lambda)^T x - \|x\|_1\right)$$

$$= -f_*(-A^T\lambda)$$

So, we have:

$$g(\lambda) = \inf_{x,y} \left(y^T I y - \lambda^T y + \|x\|_1 + \lambda^T A x - \lambda^T b \right)$$

$$= \frac{1}{4} \lambda^T \lambda - \frac{1}{2} \lambda^T \lambda - \lambda^T b + \inf_x \left(\|x\|_1 - (-A^T \lambda)^T x \right)$$

$$= -\frac{1}{4} \|\lambda\|_2^2 - \lambda^T b - f_* (-A^T \lambda)$$

$$= \begin{cases} -\frac{1}{4} \|\lambda\|_2^2 - \lambda^T b & \text{if } \|A^T \lambda\|_\infty \leq 1 \\ -\infty & \text{otherwise} \end{cases}$$

Finally, the dual of (RLS) is:

$$\begin{cases} \max_{\lambda} -\frac{1}{4} \|\lambda\|_2^2 - \lambda^T b \\ \text{s.t } \|A^T \lambda\|_\infty \leq 1 \end{cases}$$

Exercise 3:

1) $\mathcal{L}(\omega, x_i, y_i) = \max\{0, 1 - y_i(\omega^T x_i)\}$

If a data-point x_i is misclassified we have $1 - y_i(\omega^T x_i) < 0$. The constraint $z \geq 0$ ensures that if x_i is misclassified, then $z_i = 0$

- On the other hand, if x_i is well classified $1 - y_i(\omega^T x_i) > 0$ and according to the constraints on z_i, we must have

  $z_i = 1 - y_i(\omega^T x_i)$ for the i-th to be minimal

Conclusion:

By choosing $z_i = \max\{0, 1 - y_i(\omega^T x_i)\}$ we fall back to (svp1). As τ is a constant, dividing svp1 by τ, the optimal value of ω doesn't change.

②

The Lagrangian of (sep 2) is:

$$\mathcal{L}(\omega, z, \lambda, \pi) = \frac{1}{n\tau} \mathbb{1}^T z + \frac{1}{2} \|\omega\|_2^2$$

$$+ \sum_{i=1}^{n} \lambda_i \left(1 - y_i (\omega^T x_i) - z_i\right) - \pi^T z$$

$$= \left(\frac{1}{n\tau} \mathbb{1} - \lambda - \pi\right)^T z + \frac{1}{2} \|\omega\|_2^2 - \omega^T \sum_{i=1}^{n} \lambda_i x_i y_i + \mathbb{1}^T \lambda$$

•) On the one hand, $\mathcal{L}$ is a quadratic form bounded with respect to ω.

The gradient with respect to ω set to 0 gives:

$$\min_{\omega} \mathcal{L} = \left(\frac{1}{n\tau} \mathbb{1} - \lambda - \pi\right)^T z - \frac{1}{2} \left\| \sum_{i=1}^{n} \lambda_i x_i y_i \right\|_2^2 + \mathbb{1}^T \lambda$$

•) On the other hand, $\mathcal{L}$ is linear with respect to z, so:

$$\min_{z, \omega} \mathcal{L} = \begin{cases} -\frac{1}{2} \left\| \sum_{i=1}^{n} \lambda_i x_i y_i \right\|^2 + \mathbb{1}^T \lambda & \text{if } \frac{1}{n\tau} \mathbb{1} = \lambda + \pi \\ -\infty & \text{otherwise} \end{cases}$$

then, the dual is:

$$\max_{\lambda, \pi} -\frac{1}{2}\left\|\sum_{i=1}^{n} \lambda_i x_i y_i\right\|_2^2 + 1^T \lambda$$

$$\text{s.t} \quad \lambda, \pi \geq 0$$

$$\frac{1}{n\tau} 1 = \lambda + \pi$$

We eliminate π. we have:

$$\lambda, \pi \geq 0 \Rightarrow \begin{cases} \frac{1}{n\pi} \geq 1 \\ \lambda \geq 0 \end{cases}$$

Conclusion:

$$\max_{\lambda, \pi} -\frac{1}{2}\left\|\sum_{i=1}^{n} \lambda_i x_i y_i\right\|_2^2 + 1^T \lambda$$

$$\text{st} \quad \lambda \geq 0$$

$$0 \leq \lambda \leq \frac{1}{n\tau}$$