Convex Optimization HW2

Tong ZHAO

tong.zhao@eleves.enpc.fr

Exercise 1 (LP Duality)

1. We introduce two lagrange multiplier : λ associated with $x \geq 0$ and ν associated with $Ax = b$. Then the Lagrangian is

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

L is affine in x, hence

$$g(\lambda,\nu) = \inf_x L(x,\lambda,\nu) = \begin{cases} -\nu^T b & A^T\nu - \lambda + c = 0 \\ -\infty & \text{otherwise} \end{cases} \quad \xRightarrow{\text{dual}}_{\text{problem}} \quad \begin{array}{ll} \max & -\nu^T b \\ \text{s.t.} & A^T \nu + c \geq 0 \end{array}$$

2. We introduce the lagrange multiplier λ $(\lambda \geq 0)$ associated with $A^T y \leq c$, then the lagrangian is

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

Hence

$$g(\lambda) = \inf_\lambda L(x,\lambda) = \begin{cases} -\lambda^T c & A\lambda - b = 0 \\ -\infty & \text{otherwise} \end{cases} \quad \xRightarrow{\text{dual}}_{\text{problem}} \quad \begin{array}{ll} \max & -\lambda^T c \\ \text{s.t.} & A\lambda = b \\ & \lambda \geq 0 \end{array}$$

3. The problem can be seen as $(P) + (D)$, hence its dual problem can be seen as $(DP) + (DD)$, namely.

$$\max -\nu^T b - \lambda^T c = \min \nu^T b + \lambda^T c = \min b^T \nu + c^T \lambda$$
$$\text{s.t.} \quad A^T \nu + c \geq 0$$
$$A\lambda = b$$
$$\lambda \geq 0$$

by taken $\nu = -\nu$, we have

$$\min c^T \lambda - b^T \nu$$
$$\text{s.t.} \quad A^T \nu \leq c$$
$$\lambda \geq 0$$
$$A\lambda = b$$

4. Suppose that x^* is the optimal solution of (P), according to the LP strong duality, its dual problem (D) finds its maximum value $c^T x^* = b^T y^*$. So $\min -b^T y$ s.t. $A^T y \leq c$ find its optimal value at y^*. So $\min (P) + (D) = c^T x^* - b^T y^* = 0$

So the optimal value of Self - Dual is exactly 0.

Exercise 2 (Regularized Least-Square)

1. The conjugate of $\|x\|_1$ can be written as:

$$f^*(y) = \sup \{ y^T x - \|x\|_1 \} = \sup \{ y^T x - \sum_{i=1}^{N} |x_i| \}$$
$$= \sup \{ \sum_{i=1}^{N} (|y_i| - 1) |x_i| \}$$

If $\|y\|_\infty = y_k > 1$, we let $|x_i| \to \infty$, then $f^*(y) \to \infty$, so we have

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

2. we take $z = Ax$, the RLS can be reformulated as

$$\begin{aligned} \min \quad & \|z - b\|^2 + \|x\|_1 \\ \text{s.t.} \quad & z = Ax \end{aligned}$$

so the dual function is

$$g(u) = \min_{x, z} \|z - b\|^2 + \|x\|_1 + u^T(z - Ax)$$

$$\frac{\partial g(u)}{\partial x} = \frac{\partial \|x\|_1}{\partial x} - A^T u = 0 \quad \Rightarrow \quad \|A^T u\|_\infty \leq 1$$ by conjugate function $A^T u = \text{sign}(x)$

$$\frac{\partial g(u)}{\partial z} = 2(z - b) + u = 0 \quad \Rightarrow \quad z = b - \frac{u}{2}$$

So $g(u) = -\frac{\|u\|^2}{4} + u^T b \quad \text{s.t.} \quad \|A^T u\|_\infty \leq 1$

so the dual problem is

$$\begin{aligned} \min \quad & \frac{\|u\|^2}{4} - u^T b \\ \text{s.t.} \quad & \|A^T u\|_\infty \leq 1 \end{aligned}$$

Exercise 3 (Data Seperation)

1. Sep. 1 can be expressed as:

$$\min_w \frac{1}{n} \sum_{i=1}^{n} \max\{0; 1 - y_i(w^T x_i)\} + \frac{\tau}{2} \|w\|_2^2$$

$$= \min_w \tau \left(\frac{1}{n\tau} \sum_{i=1}^{n} \max\{0; 1 - y_i(w^T x_i)\} + \frac{1}{2} \|w\|_2^2 \right)$$

We know that $\max\{0; 1 - y_i(w^T x_i)\} \geq 0$, $\max\{0; 1 - y_i(w^T x_i)\} \geq 1 - y_i(w^T x_i)$

So if we have $z_i \geq 1 - y_i(w^T x_i)$ and $z_i \geq 0$, z_i satisfaits Sep. 1, since we want to minimize the function, minimize Sep. 2 also minimizes Sep. 1

So the problem

$$\begin{aligned} \min \quad & \frac{1}{n\tau} \mathbb{1}^T z + \frac{1}{2} \|w\|_2^2 \\ \text{s.t.} \quad & z_i \geq 1 - y_i (w^T x_i) \quad \forall i = 1, \dots, n \\ & z \geq 0 \end{aligned}$$

solve problem (Sep. 1)

2. We take the lagrangian in the usual manner

$$L(w, z, \lambda, \pi) = \frac{1}{2} w^T w + \sum_{i=1}^{n} \lambda_i [1 - y_i (w^T x_i) - z_i] + \frac{1}{n\tau} \sum_{i=1}^{n} z_i - \sum_{i=1}^{n} \pi_i z_i$$

we first minimize w, z for fixed λ, z

$$\frac{\partial L(w, z, \lambda, \pi)}{\partial w_i} = w_i - y_i \lambda_i x_i = 0 \qquad \frac{\partial L(w, z, \lambda, \pi)}{\partial z_i} = \frac{1}{n\tau} - \lambda_i - \pi_i = 0 \Rightarrow 0 \leq \lambda_i \leq \frac{1}{n\tau}$$

so the lagrangian can be expressed as:

$$L(w, \xi, \lambda, \pi) = \frac{1}{2}w^T w + \sum_{i=1}^{n} \lambda_i [1 - y_i(w^T x_i) - \xi_i] + \frac{1}{n\tau}\sum_{i=1}^{n} \xi_i - \sum_{i=1}^{n} \pi_i \xi_i \quad \cancel{\sum_{i=1}^{n} \xi_i \lambda_i}$$

$$= \frac{1}{2}\sum_{i,j=1}^{N} \lambda_i y_i x_i^T x_j y_j \lambda_j + \sum_{i=1}^{n} \lambda_i - \sum_{i,j=1}^{N} \lambda_i y_i x_i^T x_j y_j \lambda_j + \cancel{\frac{1}{n\tau}\sum_{i=1}^{N} \xi_i} - \cancel{\sum_{i=1}^{n}\left(\frac{1}{n\tau} - \lambda_i\right)\xi_i}$$

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

$$\text{s.t.} \quad 0 \le \lambda_i \le \frac{1}{n\tau} \quad \text{for } i \in \{1, \cdots, N\}$$

so the dual problem is

$$\max_{\lambda} \sum_{i=1}^{n} \lambda_i - \frac{1}{2}\sum_{i,j=1}^{N} \lambda_i y_i x_i^T x_j y_j \lambda_j$$

$$\text{s.t.} \quad 0 \le \lambda_i \le \frac{1}{n\tau} \quad \text{for } i \in \{1, \cdots, N\}$$