MVA : Convex Optimization

Homework 2

Borachhun YOU

Exercise 1 : (LP Duality)

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

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

$$(D) : \max_y b^T y \quad \text{s.t. } A^T y \leq c$$

① Compute the dual of (P) :

Standard form of (P) : $\min_x c^T x$ s.t. $Ax - b = 0$, $-x \leq 0$

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

Lagrange dual function :
$$g(\lambda, \nu) = \inf_x L(x, \lambda, \nu) = \inf_x \left((c + A^T\nu - \lambda)^T x - b^T\nu \right) = \begin{cases} -b^T\nu & \text{if } A^T\nu - \lambda + c = 0 \\ -\infty & \text{otherwise} \end{cases}$$

Lagrange dual problem :
$$\begin{cases} \max_{\lambda, \nu} -b^T\nu \\ \text{s.t. } A^T\nu - \lambda + c = 0 \\ \lambda \geq 0 \end{cases} \iff \begin{cases} \max_\nu -b^T\nu \\ \text{s.t. } A^T\nu + c \geq 0 \end{cases} \iff \begin{cases} \max_\nu b^T\nu \\ \text{s.t. } A^T\nu \leq c \end{cases}$$

② Compute the dual of (D) :

Standard form of (D) : $\min_y -b^T y$ s.t. $A^T y - c \leq 0$

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

Lagrange dual function :
$$g(\lambda, \nu) = \inf_y L(y, \lambda, \nu) = \begin{cases} -c^T\lambda & \text{if } A\lambda - b = 0 \\ -\infty & \text{otherwise} \end{cases}$$

Lagrange dual problem :
$$\begin{cases} \max_\lambda -c^T\lambda \\ \text{s.t. } A\lambda - b = 0 \\ \lambda \geq 0 \end{cases} \iff \begin{cases} \min_\lambda c^T\lambda \\ \text{s.t. } A\lambda = b \\ \lambda \geq 0 \end{cases}$$

③ Prove that the problem is self-dual:

$$\begin{aligned} \min_{x,y} \quad & c^T x - b^T y \\ \text{s.t.} \quad & Ax = b \\ & x \geq 0 \qquad \text{(Self-dual)} \\ & A^T y \leq c \end{aligned}$$

Standard form:
$$\begin{aligned} \min_{x,y} \quad & c^T x - b^T y \\ \text{s.t.} \quad & Ax - b = 0 \\ & -x \leq 0 \\ & A^T y - c \leq 0 \end{aligned}$$

~~Lagrangian~~ $\cancel{L(x,y,\lambda_1,\lambda_2,\nu) = c^T x - b^T y + \lambda_1^T(Ax-b) - \lambda_2^T x + \nu^T(A^T y - c)}$

Lagrangian
$$\begin{aligned} L(x,y,\lambda_1,\lambda_2,\nu) &= c^T x - b^T y - \lambda_1^T x + \lambda_2^T(A^T y - c) + \nu^T(Ax - b) \\ &= (c^T - \lambda_1^T + \nu^T A)x + (-b^T + \lambda_2^T A^T)y - \lambda_2^T c - \nu^T b \\ &= (c - \lambda_1 + A^T \nu)^T x + (A\lambda_2 - b)^T y - c^T \lambda_2 - b^T \nu \end{aligned}$$

Lagrange dual function:
$$\begin{aligned} g(\lambda_1,\lambda_2,\nu) &= \inf_{x,y} L(x,y,\lambda_1,\lambda_2,\nu) \\ &= \begin{cases} \inf_y (A\lambda_2 - b)^T y - c^T\lambda_2 - b^T\nu & \text{if } A^T\nu - \lambda_1 + c = 0 \\ -\infty & \text{otherwise} \end{cases} \\ &= \begin{cases} -c^T\lambda_2 - b^T\nu & \text{if } A^T\nu - \lambda_1 + c = 0 \text{ and } A\lambda_2 - b = 0 \\ -\infty & \text{otherwise} \end{cases} \end{aligned}$$

Lagrange dual problem:
$$\left\{\begin{aligned} \max_{\lambda_1,\lambda_2,\nu} \quad & -c^T\lambda_2 - b^T\nu \\ \text{s.t.} \quad & A^T\nu - \lambda_1 + c = 0 \\ & A\lambda_2 - b = 0 \\ & \lambda_1 \geq 0, \ \lambda_2 \geq 0 \end{aligned}\right. \iff \left\{\begin{aligned} \max_{\lambda_1,\lambda_2,\nu} \quad & -c^T\lambda_2 - b^T\nu \\ \text{s.t.} \quad & A^T\nu + c \geq 0 \\ & A\lambda_2 - b = 0 \\ & \lambda_2 \geq 0 \end{aligned}\right.$$

$$\iff \left\{\begin{aligned} \max_{\lambda_2,\nu} \quad & -c^T\lambda_2 + b^T\nu \\ \text{s.t.} \quad & A^T\nu \leq c \\ & A\lambda_2 = b \\ & \lambda_2 \geq 0 \end{aligned}\right. \iff \left\{\begin{aligned} \min_{\lambda_2,\nu} \quad & c^T\lambda_2 - b^T\nu \\ \text{s.t.} \quad & A\lambda_2 = b \\ & \lambda_2 \geq 0 \\ & A^T\nu \leq c \end{aligned}\right.$$

The same as the initial problem

Thus, the problem is self-dual.

④ • Show that the optimal solution $[x^*, y^*]$ of (Self-Dual) can be obtained by solving (P) and (D):

Since the constraints of (Self-Dual) are independent, we can write the problem as the sum of (P) and (D):

$$\begin{array}{llll} \min\limits_{x,y} & c^T x - b^T y & = \min\limits_{x} \; c^T x & + \min\limits_{y} \; -b^T y \\ \text{s.t.} & Ax = b & \text{s.t. } Ax = b & \text{s.t. } A^T y \leq c \\ & x \geq 0 & \quad x \geq 0 & \\ & A^T y \leq c & & \end{array}$$

$$= \begin{array}{l} \min\limits_{x} \; c^T x \\ \text{s.t. } Ax = b \\ \quad x \geq 0 \end{array} + \begin{array}{l} \max\limits_{y} \; b^T y \\ \text{s.t. } A^T y \leq c \end{array}$$

Thus, solving (P) and (D) gets the optimal solution x^* and y^* respectively, and $[x^*, y^*]$ the optimal solution of (Self-Dual).

• Show that the optimal value of (Self-Dual) is exactly 0:

(P) and (D) are convex and feasible since (Self-Dual) is feasible.

Then the strong duality holds: $c^T x^* = b^T y^*$

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

Therefore, the optimal value of (Self-Dual) is 0.

Exercise 2 : (Regularized Least-Square)

$A \in \mathbb{R}^{n \times d}$, $b \in \mathbb{R}^n$

(RLS) : $\min_x \|Ax - b\|_2^2 + \|x\|_1$

① Compute the conjugate of $\|x\|_1$:

Let $f(x) = \|x\|_1$

$$f^*(y) = \sup_x (y^T x - \|x\|_1) = \sup_x \left(\sum_{i=1}^d y_i x_i - \sum_{i=1}^d |x_i| \right)$$

- $\forall i, -1 \leq y_i \leq 1$:

$$y_i x_i \leq |y_i x_i| \leq |x_i|$$

$$\sum_{i=1}^d y_i x_i \leq \sum_{i=1}^d |x_i|$$

$$\sum_{i=1}^d y_i x_i - \sum_{i=1}^d |x_i| \leq 0$$

$$\Rightarrow \sup_x \left(\sum_{i=1}^d y_i x_i - \sum_{i=1}^d |x_i| \right) = 0$$

- $\exists i, y_i > 1$:

Take x such that $x_i = t > 0$ and $x_j = 0$ for $j \neq i$

$$\sum_{i=1}^d y_i x_i - \sum_{i=1}^d |x_i| = y_i t - t = t(y_i - 1) \xrightarrow[t \to +\infty]{} +\infty$$

$$\Rightarrow \sup_x \left(\sum_{i=1}^d y_i x_i - \sum_{i=1}^d |x_i| \right) = +\infty$$

- $\exists i, y_i < -1$:

Take x such that $x_i = t < 0$ and $x_j = 0$ for $j \neq i$

$$\sum_{i=1}^d y_i x_i - \sum_{i=1}^d |x_i| = y_i t - t = t(y_i - 1) \xrightarrow[t \to -\infty]{} +\infty$$

$$\Rightarrow \sup_x \left(\sum_{i=1}^d y_i x_i - \sum_{i=1}^d |x_i| \right) = +\infty$$

Thus, $f^*(y) = \begin{cases} 0 & \text{if } \forall i, |y_i| \leq 1 \\ +\infty & \text{otherwise} \end{cases}$ or $f^*(y) = \begin{cases} 0 & \text{if } \|y\|_\infty \leq 1 \\ +\infty & \text{otherwise} \end{cases}$

② Compute the dual of (RLS):

We can rewrite (RLS) into: $\min_{x,y} \|y\|_2^2 + \|x\|_1$ with $y \in \mathbb{R}^n$
s.t. $y = Ax - b$

Lagrangian of the new problem:

$$L(x,y,\nu) = \|y\|_2^2 + \|x\|_1 + \nu^T(y - Ax + b)$$
$$= \|y\|_2^2 + \nu^T y + \|x\|_1 - \nu^T Ax + \nu^T b$$

Lagrange dual function:

$$g(\nu) = \inf_{x,y} L(x,y,\nu) = \inf_y \left(\|y\|_2^2 + \nu^T y\right) + \inf_x \left(\|x\|_1 - \nu^T Ax\right) + \nu^T b$$

$$= \inf_y \left(y^T y + \nu^T y\right) + \inf_x \left(-(A^T\nu)^T x + \|x\|_1\right) + b^T \nu$$

Consider $f : y \mapsto y^T y + \nu^T y$ a quadratic function

$\nabla f = 2y + \nu = 0$

$y = -\frac{\nu}{2}$

$$\inf_y \left(y^T y + \nu^T y\right) = \frac{\nu^T \nu}{4} - \frac{\nu^T \nu}{2} = -\frac{\nu^T \nu}{4}$$

Also, $\inf_x \left(-(A^T\nu)^T x + \|x\|_1\right) = \inf_x -\left((A^T\nu)^T x - \|x\|_1\right)$

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

$$= \begin{cases} 0 & \text{if } \|A^T\nu\|_\infty \leq 1 \\ -\infty & \text{otherwise} \end{cases}$$ (from question ①)

$$\Rightarrow g(\nu) = \begin{cases} -\frac{\nu^T \nu}{4} + b^T \nu & \text{if } \|A^T\nu\|_\infty \leq 1 \\ -\infty & \text{otherwise} \end{cases}$$

Lagrange dual problem: $\max_\nu \; b^T \nu - \frac{1}{4}\|\nu\|_2^2$
s.t. $\|A^T\nu\|_\infty \leq 1$

Exercise 3 : (Data Separation)

Loss function : $\mathcal{L}(w, x_i, y_i) = \max\{0, 1 - y_i(w^T x_i)\}$

$$(\text{Sep. }1) : \min_{w} \frac{1}{n} \sum_{i=1}^{n} \mathcal{L}(w, x_i, y_i) + \frac{\tau}{2} \|w\|_2^2$$

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

① Show that (Sep. 2) solves (Sep. 1) :

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

- $\forall i$, $z_i \geq 0$ because $\max\{0, \dots\} \geq 0$
- If $1 - y_i(w^T x_i) \geq 0 \Rightarrow 1 - y_i(w^T x_i) = z_i$
  If $1 - y_i(w^T x_i) < 0 \Rightarrow 1 - y_i(w^T x_i) < z_i$
  $\Rightarrow 1 - y_i(w^T x_i) \leq z_i$

$$(\text{Sep. }1) \Leftrightarrow \begin{array}{ll} \min\limits_{w, z} & \frac{1}{n} \sum_{i=1}^{n} z_i + \frac{\tau}{2}\|w\|_2^2 \\ \text{s.t.} & z_i \geq 1 - y_i(w^T x_i) \\ & z \geq 0 \end{array}$$

$$\Leftrightarrow \begin{array}{ll} \min\limits_{w, z} & \frac{1}{n\tau} \mathbb{1}^T z + \frac{1}{2}\|w\|_2^2 \\ \text{s.t.} & z_i \geq 1 - y_i(w^T x_i) \\ & z \geq 0 \end{array} : (\text{Sep. }2)$$

② Compute the dual of (Sep. 2) :

Standard form of (Sep. 2) :
$$\begin{array}{ll} \min\limits_{w, z} & \frac{1}{n\tau} \mathbb{1}^T z + \frac{1}{2}\|w\|_2^2 \\ \text{s.t.} & 1 - y_i(w^T x_i) - z_i \leq 0 \\ & -z \leq 0 \end{array}$$

Lagrangian : $L(w, z, \lambda_i, \pi) = \frac{1}{n\tau} \mathbb{1}^T z + \frac{1}{2}\|w\|_2^2 + \sum_{i=1}^{n} \lambda_i (1 - y_i(w^T x_i) - z_i) - \pi^T z$

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

Lagrange dual function :

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

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

Consider $h : w \longmapsto \frac{1}{2} w^T w - \sum_{i=1}^{n} \lambda_i y_i (x_i^T w)$ a quadratic equation

$$\nabla h = w - \sum_{i=1}^{n} \lambda_i y_i x_i = 0$$

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

$$\inf_{w} \left(\frac{1}{2} \|w\|_2^2 - \sum_{i=1}^{n} \lambda_i y_i (w^T x_i) \right) = \frac{1}{2} \left\| \sum_{i=1}^{n} \lambda_i y_i x_i \right\|_2^2 - \left\| \sum_{i=1}^{n} \lambda_i y_i x_i \right\|_2^2$$

$$= -\frac{1}{2} \left\| \sum_{i=1}^{n} \lambda_i y_i x_i \right\|_2^2$$

$$\inf_{z} \left(\left(\frac{1}{n\tau} \mathbb{1} - \lambda - \pi \right)^T z \right) = \begin{cases} 0 & \text{if } \frac{1}{n\tau} \mathbb{1} - \lambda - \pi = 0 \\ -\infty & \text{otherwise} \end{cases}$$

$$g(\lambda, \pi) = \begin{cases} -\frac{1}{2} \left\| \sum_{i=1}^{n} \lambda_i y_i x_i \right\|_2^2 + \mathbb{1}^T \lambda & \text{if } \frac{1}{n\tau} \mathbb{1} - \lambda - \pi = 0 \\ 0 & \text{otherwise} \end{cases}$$

Lagrange dual problem :

$$\begin{cases} \max\limits_{\lambda, \pi} & -\frac{1}{2} \left\| \sum_{i=1}^{n} \lambda_i y_i x_i \right\|_2^2 + \mathbb{1}^T \lambda \\ \text{s.t.} & \frac{1}{n\tau} \mathbb{1} - \lambda - \pi = 0 \\ & \lambda \geq 0 \ , \ \pi \geq 0 \end{cases} \iff \begin{cases} \max\limits_{\lambda} & -\frac{1}{2} \left\| \sum_{i=1}^{n} \lambda_i y_i x_i \right\|_2^2 + \mathbb{1}^T \lambda \\ \text{s.t.} & \frac{1}{n\tau} \mathbb{1} - \lambda \geq 0 \\ & \lambda \geq 0 \end{cases}$$