HW2

Ex. 1

1. Lagrangian function:

For $x \in \mathbb{R}^d$, $\lambda \in \mathbb{R}^d$, $\nu \in \mathbb{R}^m$:

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

Dual function:

$$g(\lambda, \nu) = \inf_{x \in \mathbb{R}^d} \left(b^T \nu + (c - A^T \nu - \lambda)^T x \right)$$

If $c - A^T \nu - \lambda \neq 0$, then g is unbounded from below, hence the infimum becomes $-\infty$. So:

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

Dual problem:

$$\begin{array}{lll} \text{maximize} & b^T \nu & \text{in } \lambda \in \mathbb{R}^d, \nu \in \mathbb{R}^m \\ \text{subject to} & \underbrace{A^T \nu + \lambda = c \;,\; \lambda \geq 0}_{A^T \nu \leq c} & \end{array}$$

2. First, notice that

$$\max_y (b^T y) = \min_y (-b^T y)$$

Thus, we will proceed by considering the problem (D) as a minimization problem.

Lagrangian function:

For $y \in \mathbb{R}^m$, $\lambda \in \mathbb{R}^d$:

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

Dual function:

$$g(\lambda) = \inf_{y \in \mathbb{R}^m} \left((-b^T + \lambda^T A^T) y - \lambda^T c \right)$$

If $A\lambda \neq b$, then g is unbounded from below, hence the infimum becomes $-\infty$. So:

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

Dual problem:

maximize $-\lambda^T c$ in $\lambda \in \mathbb{R}^d$

subject to $A\lambda = b, \quad \lambda \geq 0$

3. Lagrangian function:

For $x, \lambda, \mu \in \mathbb{R}^d$, $y, \nu \in \mathbb{R}^m$:

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

Similarly as before, we then get:

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

Dual problem:

maximize $b^T \nu - c^T \mu$ in $\nu \in \mathbb{R}^m$, $\lambda, \mu \in \mathbb{R}^d$

subject to $c - \lambda - A^T \nu = 0$

$A\mu = b$

$\lambda, \mu \geq 0$

which is equivalent to:

minimize $c^T \mu - b^T \nu$ in $\nu \in \mathbb{R}^m$, $\mu \in \mathbb{R}^d$

subject to $A^T \nu \leq c$

$A\mu = b$

$\mu \geq 0$

Thus, we can conclude that this problem is self-dual

4. First of all, one can notice that (P) and (D) are the dual of one another. Thus, by strong duality of linear programs, we have that:

$c^T x' = b^T y'$

where x' is the optimal solution of (P) and y' is the optimal solution of (D).

Furthermore, since by combining the primal and dual constraints, we get the self-dual constraints, we can be certain that x' and y' satisfy these constraints.

By strong duality for the self-dual problem, we know that:

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

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

Since the optimal solutions x' and y' of (P) and (D) satisfy this equation, they are also optimal solutions for the self-dual problem.

Furthermore, as shown above, the optimal value of the self-dual is 0.

.

Ex. 2

1. By the definition of the conjugate, for $y \in \mathbb{R}^d$:

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

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

This is equivalent to searching for

$$\sup_{x_i \in \mathbb{R}} (y_i x_i - |x_i|)$$

$\forall i \in \{1, \dots, d\}$.

If $\exists i$ s.t. $|y_i| > 1$, it suffices to choose x_i s.t. $\mathrm{sign}(y_i) = \mathrm{sign}(x_i)$. Then:

$$y_i x_i - |x_i| = |x_i| \underbrace{(|y_i| - 1)}_{>0} \xrightarrow[|x_i| \to \infty]{} \infty$$

If $\forall i$ $|y_i| \le 1$, then:

$y_i x_i \leq |x_i|$

$\Rightarrow y_i x_i - |x_i| \leq 0$

By taking $x_i = 0$, we find the supremum.

So:

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

2. Let us reformulate our problem slightly:

$$\underset{x,y}{\text{minimize}} \quad \|y\|_2^2 + \|x\|_1$$

$$\text{subject to} \quad Ax - b = y$$

Then the Lagrangian is given by:

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

where $x \in \mathbb{R}^d$, $y, \nu \in \mathbb{R}^n$

Dual function:

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

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

$$+ \inf_y \left(\|y\|_2^2 + \nu^T y \right) + \nu^T b$$

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

$$= \begin{cases} -0 & \text{if } \|A^T\nu\|_\infty \leq 1 \\ -\infty & \text{else} \end{cases}$$

- Set $f : \mathbb{R}^n \longrightarrow \mathbb{R}_+$

$$y \longmapsto \|y\|_2^2 + \nu^T y$$

Then:

$(\nabla_y f(y))_i = 0$

$\Leftrightarrow 2y_i + \nu_i = 0$

$\Leftrightarrow y_i = -\frac{\nu_i}{2}$

Thus:

$$\inf_y \left(\|y\|_2^2 + \nu^T y \right) = \left\| -\frac{\nu}{2} \right\|_2^2 - \frac{1}{2}\|\nu\|_2^2$$

$$= -\frac{1}{4}\|\nu\|_2^2$$

So:

$$g(\nu) = \begin{cases} -\frac{1}{4}\|\nu\|_2^2 + b^T\nu & \text{if } \|A^T\nu\|_\infty \leq 1 \\ -\infty & \text{else} \end{cases}$$

Dual problem:

maximize $\quad -\frac{1}{4}\|\nu\|_2^2 + b^T\nu$

subject to $\quad \|A^T\nu\|_\infty \leq 1$

Ex. 3

1. First of all we have:

$$\min_{w,z} \frac{1}{n\tau} 1^T z + \frac{1}{2}\|w\|_2^2$$

$$= \frac{1}{\tau} \min_{w,z} \frac{1}{n}\sum_{i=1}^n z_i + \frac{\tau}{2}\|w\|_2^2$$

s.t. $z_i \geq 1 - y_i(w^T x_i)$

$\quad z \geq 0$

Since $\frac{1}{n}\sum_{i=1}^n z_i$ is the sum of positive numbers, it can be minimized by minimizing the individual $z_i \geq 0$.

Consider the following 2 cases:

- $1 - y_i(w^T x_i) \leq 0$:

  z_i is minimized by taking $z_i = 0$

- $1 - y_i(w^T x_i) > 0$:

  z_i is minimized by taking $z_i = 1 - y_i(w^T x_i)$

$\Rightarrow z_i = \max(0, 1 - y_i(w^T x_i))$ minimizes the sum

Hence, (Sep. 2) can be written as:

$$\frac{1}{\tau} \min_w \frac{1}{n}\sum_{i=1}^n L(w, x_i, y_i) + \frac{\tau}{2}\|w\|_2^2$$

which solves (Sep. 1) up to a constant.

2. Lagrangian function:

For $w \in \mathbb{R}^d$, $z, \pi \in \mathbb{R}^n$, $\lambda_i \in \mathbb{R}$ $\forall i \in \{1, \dots, n\}$:

$$L(w, z, \pi, \lambda_1, \dots, \lambda_n)$$

$$= \frac{1}{n\tau} \sum_{i=1}^{n} z_i + \frac{1}{2} \sum_{i=1}^{d} w_i^2$$

$$+ \sum_{i=1}^{n} \lambda_i (1 - y_i(w^T x_i) - z_i)$$

$$- \sum_{i=1}^{n} \pi_i z_i$$

Let us first minimize wrt. z:

$$(\nabla_z L(w, z, \pi, \lambda_1, \dots, \lambda_n))_i = 0$$

$$\Leftrightarrow \frac{1}{n\tau} - \lambda_i - \pi_i = 0$$

$$\Leftrightarrow \frac{1}{n\tau} = \lambda_i + \pi_i$$

Thus, L is minimal wrt. z if $\frac{1}{n\tau} = \lambda_i + \pi_i$.

Minimization wrt. w:

$$\nabla_w L(w, z, \pi, \lambda_1, \dots, \lambda_n) = 0$$

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

$$\Leftrightarrow w = \sum_{i=1}^{n} \lambda_i y_i x_i \qquad (*)$$

Thus, L is minimal wrt w if $(*)$ holds $\forall i \in \{1, \dots, d\}$.

Finally, the dual problem is given by:

maximize $\displaystyle \sum_{i=1}^{n} \lambda_i - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \lambda_i \lambda_j y_i y_j (x_i^T x_j)$

subject to
$$\left.\begin{array}{l} \frac{1}{n\tau} = \lambda_i + \pi_i \\ \lambda_i \geq 0 \quad \forall i \in \{1, \dots, n\} \\ \pi \geq 0 \end{array}\right\} \begin{array}{l} 0 \leq \lambda_i \leq \frac{1}{n} \\ \forall i \in \{1, \dots, n\} \end{array}$$