RORON

Gabriel

HW 2

Problem 1

1) The Lagrangian of (P) can be written as

$$L : \begin{cases} \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^n & \longrightarrow \mathbb{R} \\ (x, \lambda, \nu) & \longrightarrow c^T x - \lambda^T x + \nu^T (Ax - b) \end{cases}$$

Let $(\lambda, \nu) \in \mathbb{R}^d \times \mathbb{R}^n$ be fixed.

As a function of x, L is affine, hence convex, with gradient

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

thus $$g(\lambda, \nu) := \inf_{x \in \mathbb{R}^d} L(x, \lambda, \nu) = \begin{cases} -\nu^T b & \text{if } c - \lambda + A^T \nu = 0 \\ -\infty & \text{otherwise} \end{cases}$$

Hence the dual of (P) is

$$\max_{(\lambda, \nu)} g(\lambda, \nu) \quad \text{s.t} \quad \lambda \geq 0, \ (\lambda, \nu) \in \operatorname{dom} g$$

$$\iff \max_{(\lambda, \nu)} -\nu^T b \quad \text{s.t} \quad \lambda \geq 0, \ c - \lambda + A^T \nu = 0$$

$$\iff \boxed{\max_{\nu} -\nu^T b \quad \text{s.t} \quad c + A^T \nu \geq 0}$$

With the change of variable $\nu \longleftarrow -\nu$, the dual may be rewritten as

$$\max_{\nu} \nu^T b \quad \text{s.t} \quad c \geq A^T \nu$$

Hence $\boxed{\text{the dual of (P) is (D)}}$

2) (D) may be **rewritten** as $\quad - \min_y -b^T y \quad$ s.t. $A^T y \leq c$

Let (D') be the problem $\quad \min_y -b^T y \quad$ st $A^T y \leq c$

The Lagrangian of (D') is

$$L: \begin{cases} \mathbb{R}^n \times \mathbb{R}^d & \longrightarrow \mathbb{R} \\ (y, \lambda) & \longrightarrow -b^T y + \lambda^T (A^T y - c) \end{cases}$$

which is affine in y, with gradient $\nabla_y L(y,\lambda) = -b + A\lambda$

Hence $g(\lambda) = \inf_{y \in \mathbb{R}^n} L(y,\lambda) = \begin{cases} -\lambda^T c & \text{if } -b + A\lambda = 0 \\ -\infty & \text{otherwise} \end{cases}$

The dual of (D') is $\quad \max_{\lambda \in \mathbb{R}^d} g(\lambda) \quad$ s.t $\lambda \geq 0$, $\lambda \in \operatorname{dom} g$

$$\Longleftrightarrow \quad \max_\lambda -\lambda^T c \quad \text{s.t} \quad \lambda \geq 0, \ A\lambda = b$$

The dual of (D) is $\quad - \max_\lambda -\lambda^T c \quad$ st $\lambda \geq 0$, $A\lambda = b$

$$\Longleftrightarrow \quad \min_\lambda \lambda^T c \quad \text{st} \quad \lambda \geq 0, \ A\lambda = b$$

Hence <u>the dual of (D) is (P)</u>

3) The Lagrangian of (Self-Dual) is

$$L: \begin{cases} \mathbb{R}^{d+n} \times \mathbb{R}^{n+d} \times \mathbb{R}^n & \longrightarrow \mathbb{R} \\ \left(\begin{pmatrix} x \\ y \end{pmatrix}, \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix}, \nu\right) & \longrightarrow \begin{pmatrix} c - \lambda_1 + A^T \nu \\ -b + A\lambda_2 \end{pmatrix}^T \begin{pmatrix} x \\ y \end{pmatrix} - \lambda_2^T c - \nu^T b \end{cases}$$

which is affine in $\begin{pmatrix} x \\ y \end{pmatrix}$, with gradient $\nabla_{\begin{pmatrix} x \\ y \end{pmatrix}} L\left(\begin{pmatrix} x \\ y \end{pmatrix}, \lambda, \nu\right) = \begin{pmatrix} c - \lambda_1 + A^T \nu \\ -b + A\lambda_2 \end{pmatrix}$

Thus $g(\lambda, \nu) = \inf_{\begin{pmatrix} x \\ y \end{pmatrix}} L\left(\begin{pmatrix} x \\ y \end{pmatrix}, \lambda, \nu\right) = \begin{cases} -\lambda_2^T c - \nu^T b & \text{if } c - \lambda_1 + A^T \nu = 0, \ -b + A\lambda_2 = 0 \\ -\infty & \text{otherwise} \end{cases}$

The dual of (Self-dual) is therefore

$$\max_{(\lambda, \nu)} \; -\lambda_2^T c - \nu^T b \quad \text{s.t} \quad \begin{cases} \lambda_1 \geq 0 \\ \lambda_2 \geq 0 \\ c - \lambda_1 + A^T \nu = 0 \\ -b + A\lambda_2 = 0 \end{cases}$$

$$\Leftrightarrow \quad \max_{(\lambda_2, \nu)} \; -\lambda_2^T c - \nu^T b \quad \text{s.t} \quad \begin{cases} c + A^T \nu \geq 0 \\ A\lambda_2 = b \\ \lambda_2 \geq 0 \end{cases}$$

$$\Leftrightarrow \quad \min_{(\lambda_2, \nu)} \; \lambda_2^T c + \nu^T b \quad \text{s.t} \quad \begin{cases} c + A^T \nu \geq 0 \\ A\lambda_2 = b \\ \lambda_2 \geq 0 \end{cases}$$

With the change of variable $\nu \leftarrow -\nu$, this rewrites as

$$\min_{(\lambda_2, \nu)} \; \lambda_2^T c - \nu^T b \quad \text{s.t} \quad \begin{cases} c \geq A^T \nu \\ A\lambda_2 = b \\ \lambda_2 \geq 0 \end{cases}$$

Hence <u>the dual of (Self-dual) is itself</u>

4) • Let us prove that x^* is optimal for (P).

Consider x such that $Ax = b$ and $x \geq 0$

Then (x, y^*) is feasible for (Self-dual), hence

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

$$\Rightarrow \quad c^T x^* \leq c^T x$$

Thus x^* solves (P)

• Let us prove that y^* solves (D)

Consider y such that $A^T y \leq c$

Then (x^*, y) is feasible for (Self-dual), hence

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

$$\Rightarrow \quad -b^T y^* \leq -b^T y$$

$$\Rightarrow \quad b^T y^* \geq b^T y$$

Hence y^* solves (D).

- Since (P) is a linear program, strong duality holds.
Hence (P) and its dual have the same optimal value.
But the dual of (P) is (D), thus $c^T x^* = b^T y^*$
That is, $c^T x^* - b^T y^* = 0$
Since (x^*, y^*) is optimal for (Self-Dual), we deduce that the optimal value of (Self-Dual) is 0.

Problem 2

1) - Let f denote the L_1-norm.
Let y be such that $x \to y^T x - f(x)$ is bounded above.
Let $x_i = t$ and $x_k = 0$ for $k \neq i$
then $y^T x - f(x) = y_i t - |t|$
Since this must remain bounded above when $t \to \pm \infty$,
we must have $|y_i| \leq 1$, hence $\|y\|_\infty \leq 1$.

- Consider y such that $\|y\|_\infty \leq 1$
Then $y^T x - f(x) = \sum_{i=1}^{d} (y_i x_i - |x_i|)$

$$= \sum_{i=1}^{d} \underbrace{\underbrace{x_i^+}_{\geq 0} \underbrace{(y_i - 1)}_{\leq 0} - \underbrace{x_i^-}_{\geq 0} \underbrace{(y_i + 1)}_{\geq 0}}_{\leq 0}$$

≤ 0

The bound is attained for $x_1 = \dots = x_d = 0$

Hence $\boxed{\begin{array}{l} \operatorname{dom} f^* = \{ y \,/\, \|y\|_\infty \leq 1 \} \\ f^*(y) = 0 \end{array}}$

2) <u>Lemma</u> : let $f : X \to \mathbb{R}$

$g : Y \to \mathbb{R}$

Then $\inf\limits_{x,y} \left(f(x) + g(y) \right) = \inf\limits_{y} \left(g(y) + \inf\limits_{x} f(x) \right)$

Proof : $\boxed{\leq}$ Let $x, y \in X \times Y$

Then $\inf\limits_{x,y} f(x) + g(y) \leq f(x) + g(y)$

Hence $\inf\limits_{x,y} \left[f(x) + g(y) \right] - g(y) \leq f(x)$ for all x

Hence $\inf\limits_{x,y} \left[f(x) + g(y) \right] - g(y) \leq \inf\limits_{x} f(x)$

Hence $\inf\limits_{x,y} \left[f(x) + g(y) \right] \leq g(y) + \inf\limits_{x} f(x)$ for all y

Hence $\inf\limits_{x,y} \left[f(x) + g(y) \right] \leq \inf\limits_{y} \left[g(y) + \inf\limits_{x} f(x) \right]$

$\boxed{\geq}$ Let $x, y \in X \times Y$

Since $\inf\limits_{x} f(x) \leq f(x)$, $g(y) + \inf\limits_{x} f(x) \leq g(y) + f(x)$

Hence $\inf\limits_{y} \left(g(y) + \inf\limits_{x} f(x) \right) \leq g(y) + f(x)$ for all x, y

Thus $\inf\limits_{y} \left(g(y) + \inf\limits_{x} f(x) \right) \leq \inf\limits_{x,y} \left(g(y) + f(x) \right)$

Note that (RLS) can be rewritten as

$$\min_{x,y} \|y - b\|_2^2 + \|x\|_1 \quad \text{s.t} \quad y = Ax$$

The Lagrangian may be written as

$$L\left(\begin{pmatrix} x \\ y \end{pmatrix}, v \right) = \|y - b\|_2^2 + \|x\|_1 + v^T (y - Ax)$$

For fixed v, $\inf\limits_{x,y} L\left(\begin{pmatrix} x \\ y \end{pmatrix}, v \right) = \inf\limits_{x,y} \left[\|y - b\|_2^2 + v^T y + \|x\|_1 - v^T A x \right]$

$$\underset{\text{(Lemma)}}{=} \inf_{y} \left[\|y - b\|_2^2 + v^T y + \inf_{x} \left(\|x\|_1 - v^T A x \right) \right]$$

$$= \inf_{y} \left[\|y - b\|_2^2 + v^T y - \sup_{x} \left((A^T v)^T x - \|x\|_1 \right) \right]$$

$$= \inf_y \left[\|y-b\|_2^2 + v^T y - f^*(A^T v) \right]$$

$$= \begin{cases} \inf_y \left(\|y-b\|_2^2 + v^T y \right) & \text{if } \|A^T v\|_\infty \leq 1 \\ -\infty & \text{otherwise} \end{cases}$$

Since $\|y-b\|_2^2 + v^T y = \|y\|_2^2 - 2(b - \frac{v}{2})^T y + \|b - \frac{v}{2}\|_2^2 + \|b\|_2^2 - \|b - \frac{v}{2}\|_2^2$

$$= \|y - (b - \tfrac{v}{2})\|_2^2 + \|b\|_2^2 - \|b - \tfrac{v}{2}\|_2^2$$

$$\inf_y \left(\|y-b\|_2^2 + v^T y \right) = \|b\|_2^2 - \|b - \tfrac{v}{2}\|_2^2$$

Hence $g(v) = \inf_{x,y} L\left(\binom{x}{y}, v\right) = \begin{cases} \|b\|_2^2 - \|b - \frac{v}{2}\|_2^2 & \text{if } \|A^T v\|_\infty \leq 1 \\ -\infty & \text{otherwise} \end{cases}$

The dual of (RLS) is $\max_v \; g(v)$ s.t $v \in \text{dom } g$.

$$\boxed{\max_v \; \|b\|_2^2 - \|b - \tfrac{v}{2}\|_2^2 \quad \text{s.t } \|A^T v\|_\infty \leq 1}$$

It suffices to solve $\min_v \; \|b - \frac{v}{2}\|_2^2$ st $\|A^T v\|_\infty \leq 1$

Problem 3

1) . First, note that $\begin{cases} z_i \geq 1 - y_i \omega^T x_i \\ z_i \geq 0 \end{cases} \iff z_i \geq \max(0, 1 - y_i \omega^T x_i)$

$\iff z_i \geq \mathcal{L}(\omega, x_i, y_i)$

. Suppose there exists optimal ω^*, z^* for (Sep 2)

Then the inequality constraints in (Sep 2) are binding at (ω^*, z^*)

Indeed, if this were not the case, we would have

for some i $z_i^* > \max(0, 1 - y_i \omega^T x_i)$

But then, by changing z_i^* into $z_i^* - \varepsilon$ for some sufficiently small $\varepsilon >$

the objective function in (Sep 2) would decrease strictly, and the

inequality constraint would still hold, contradicting the optimality of (w^*, z^*). Hence $\forall i \in [\![1,n]\!]$, $z_i^* = \max(0, 1 - y_i w^{*T} x_i)$
$= \mathcal{L}(w^*, x_i, y_i)$

- Let us prove that (w^*, z^*) is optimal for (Sep 1).

Let $w \in \mathbb{R}^d$, and define z by $\forall i$, $z_i = \max(0, 1 - y_i w^T x_i)$

Then (w, z) is feasible for (Sep 2), hence

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

$$\iff \frac{1}{n}\sum_{i=1}^n z_i^* + \frac{\tau}{2}\|w^*\|_2^2 \leq \frac{1}{n}\sum_{i=1}^n z_i + \frac{\tau}{2}\|w\|_2^2$$

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

Hence (w^*, z^*) solves (Sep 1)

Thus, (Sep 2) solves (Sep 1)

2) The Lagrangian of (Sep 2) is

$$L\left(\begin{pmatrix} w \\ z \end{pmatrix}, \begin{pmatrix} \lambda \\ \pi \end{pmatrix}\right) = \frac{1}{n\tau} 1^T z + \frac{1}{2}\|w\|_2^2 + \lambda^T \begin{pmatrix} \vdots \\ 1 - y_i w^T x_i - z_i \\ \vdots \end{pmatrix} - \pi^T z$$

As a function of $\begin{pmatrix} w \\ z \end{pmatrix}$, it is the sum of an affine function and a convex function. It is thus convex in $\begin{pmatrix} w \\ z \end{pmatrix}$.

Besides, $\frac{\partial L}{\partial w_i} = w_i + \sum_{j=1}^n -\lambda_j y_j (x_j)_i$

$$\frac{\partial L}{\partial z_i} = \frac{1}{n\tau} - \lambda_i - \pi_i$$

- If $\forall i \in [\![1,n]\!]$, $\frac{1}{n\tau} = \lambda_i + \pi_i$, then $w^* = \sum_{j=1}^n \lambda_j y_j x_j$

$z^* = z$

is a critical point (for any z), hence yields a minimum.

- If there exists some i with $\frac{1}{n\tau} \neq \lambda_i + \pi_i$, by letting $z_i \to \pm\infty$

and $z_k = 0$ for $k \neq i$, one easily proves that the infimum is $-\infty$.

• Let us compute the value of the minimum when $\forall i,\ \frac{1}{n\tau} = \lambda_i + \pi_i$.

For any z, $L\left(\begin{pmatrix} w^* \\ z \end{pmatrix}, \begin{pmatrix} \lambda \\ \pi \end{pmatrix}\right) = \frac{1}{2} \left\| \sum_{j=1}^{n} \lambda_j y_j x_j \right\|_2^2 + \sum_{i=1}^{n} \lambda_i$

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

$$= \frac{1}{2} \left\| \sum_{j=1}^{n} \lambda_j y_j x_j \right\|_2^2 + \sum_{i=1}^{n} \lambda_i - \left\| \sum_{j=1}^{n} \lambda_j y_j x_j \right\|_2^2$$

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

Therefore,

$$g(\lambda, \pi) = \inf_{w, z} L\left(\begin{pmatrix} w \\ z \end{pmatrix}, \begin{pmatrix} \lambda \\ \pi \end{pmatrix}\right)$$

$$= \begin{cases} \sum_{i=1}^{n} \lambda_i - \frac{1}{2} \left\| \sum_{i=1}^{n} \lambda_i y_i x_i \right\|_2^2 & \text{if } \forall i \in [\![1, n]\!],\ \lambda_i + \pi_i = \frac{1}{n\tau} \\ -\infty & \text{otherwise} \end{cases}$$

The dual of (Sep 2) can then be written as

$$\max_{(\lambda, \pi)} g(\lambda, \pi) \quad \text{s.t.} \quad \begin{aligned} &\lambda \geq 0 \\ &\pi \geq 0 \\ &(\lambda, \pi) \in \operatorname{dom} g \end{aligned}$$

$$\Longleftrightarrow \max_{(\lambda, \pi)} \sum_{i=1}^{n} \lambda_i - \frac{1}{2} \left\| \sum_{i=1}^{n} \lambda_i y_i x_i \right\|_2^2 \quad \text{s.t.} \quad \begin{aligned} &\pi = \frac{1}{n\tau} \mathbb{1} - \lambda \\ &\lambda \geq 0 \\ &\pi \geq 0 \end{aligned}$$

Therefore, π may be dropped and we get

$$\boxed{\max_{\lambda} \sum_{i=1}^{n} \lambda_i - \frac{1}{2} \left\| \sum_{i=1}^{n} \lambda_i y_i x_i \right\|_2^2 \quad \text{s.t.} \quad \frac{1}{n\tau} \mathbb{1} \geq \lambda \geq 0}$$