Exercises

## Exercise 1

1. The associated Lagrangian of (P) is : $L(x,\lambda,\nu) = c^T x + \nu^T(Ax-b) - \lambda^T x$
$(x,\lambda \in \mathbb{R}^d,\ \nu \in \mathbb{R}^n)$

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

which is affine in $x$

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

The dual problem is: maximize $g(\lambda,\nu)$
s.t. $\lambda \succeq 0$

$\Leftrightarrow$ maximize $-\nu^T b$
s.t. $c + A^T\nu - \lambda = 0$
$\lambda \succeq 0$

$\Leftrightarrow$ maximize $-\nu^T b$
s.t. $c + A^T\nu \succeq 0$

and by using $-\nu$ instead of $\nu$ we obtain <u>(D) as the dual of (P)</u>

2. (D) can be rewritten into min $-b^T y$
s.t. $A^T y - c \preceq 0$

The Lagrangian is $L(y,\lambda,\nu) = -b^T y + \lambda^T(A^T y - c)$
by the same reasonning as before

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

So the dual problem is maximize $-\lambda^T c$
s.t. $A\lambda = b$
$\lambda \succeq 0$

$\Leftrightarrow$ minimize $\lambda^T c$ $\Leftrightarrow$ <u>(P)</u>
s.t. $A\lambda = b$
$\lambda \succeq 0$

3. The Lagrangian associated to the problem is

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

which is affine in $x$ and in $y$, so

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

The dual problem is

$$\max_{\lambda,\nu} \; -\nu^T b - \lambda^T c$$
$$\text{s.t.} \quad c + A^T\nu - \lambda = 0$$
$$-b + A\lambda = 0$$
$$\lambda \succeq 0$$

$$\iff \min_{\lambda,\nu} \; c^T\lambda + b^T\nu$$
$$\text{s.t.} \quad A\lambda = b$$
$$A^T\nu + c \succeq 0$$

by writting $x = \lambda$ and $y = -\nu$, the dual problem is equivalent to:

$$\min_{x,y} \; c^T x - b^T y$$
$$\text{s.t.} \quad Ax = b$$
$$x \geq 0$$
$$A^T y \leq c$$

Therefore <u>the problem is self-dual</u>

4. If $[x^*, y^*]$ is solution of the above problem, then $Ax^* = b$

$$x^* \geq 0$$
$$A^T y^* \leq c$$

hence $x^*$ is a feasible point for (P) and $y^*$ for (D), so <u>strong duality holds</u> and we have $p^* = d^*$

Let $x, y$ such that $Ax = b$, $x \geq 0$, $A^T y \leq c$

$$c^T x - b^T y \leqslant c^T x - \max_{\substack{y \\ \text{s.t. } A^T y \leqslant c}} b^T y = \underbrace{c^T x - d^*}_{\text{independant of } y}$$

$$\Rightarrow \min_{\substack{y \\ \text{s.t. } A^T y \leq c}} (c^T x - b^T y) \leqslant c^T x - d^*$$

It is true for all $x$ feasible, therefore

$$\min_{\substack{x,y \\ \text{s.t. } Ax = b,\, x \geq 0,\, A^T y \leq c}} (c^T x - b^T y) \leqslant \min_{\substack{x \\ \text{s.t. } Ax = b,\, x \geq 0}} (c^T x - d^*) = p^* - d^*$$

In the same way, we can prove that $\min_{\substack{x,y \\ \text{s.t. } Ax = b,\, x \geq 0,\, A^T y \leq c}} c^T x - b^T y \geqslant p^* - d^*$

so <u>the optimal value of (Self-Dual) is $p^* - d^* = 0$</u>

Moreover, if we note $x'$ the optimal solution of (P) and $y'$ the optimal solution of (D) then we have $Ax' = b$, $x' \geq 0$, $A^T y' \leq c$ so $[x', y']$ is a feasible point for (Self-Dual), and we have $c^T x' - b^T y' = p^* - d^* = 0 = \min_{x,y} (c^T x - b^T y)$

so <u>$[x', y']$ is the optimal solution of (Self-Deal)</u>

Therefore, we have proven that:

- $[x^*, y^*]$ can be obtained by solving (P) and (D)
- the optimal value of (Self-Dual) is exactly 0

<u>Exercise 2</u>

1. First, we note $\|.\|_*$ the dual norm of $\|.\|$ a norm on $\mathbb{R}^\ell$ ($\ell \in \mathbb{N}$)

Recall that the conjugate of $f_0 = \|.\|$ is given by

$$f_0^*(y) = \begin{cases} 0 & \text{if } \|y\|_* \leq 1 \\ \infty & \end{cases} \quad \text{(example 3.26 of the book)}$$

and that the dual norm of $\|.\|_1$ is $\|.\|_\infty$ (*)

Then, the conjugate of $f_0(x) = \|x\|_1$ is given by

$$\boxed{f_0^*(y) = \begin{cases} 0 & \text{if } \|y\|_\infty \leq 1 \\ \infty & \text{otherwise} \end{cases}}$$

(*) Let $u \in \mathbb{R}^l$ ($l \in \mathbb{N}$). $\|u\|_{1*} = \sup_{\|v\|_1 = 1} v^T u$

~~let~~ ~~$v \in \mathbb{R}^l$ s.t. $\|v\|_1 = 1$~~

- we have $v^T u = \sum_{i=1}^{l} v_i u_i \leq \sum_{i=1}^{l} |v_i||u_i| \leq \|u\|_\infty \underbrace{\sum_{i=1}^{l} v_i}_{=1} = \|u\|_\infty$
- if we note $k \in [\![1;l]\!]$ such that $u_k = \|u\|_\infty$

  let $v = (0, \dots, 0, \underset{k^{th} \text{ coordinate}}{\text{sign}(u_k)}, 0, \dots, 0)$, $v^T u = \|u\|_\infty$ and $\|v\|_1 = 1$

  so $\sup_{\|v\|_1 = 1} v^T u \geq \|u\|_\infty$

- So $\underline{\|u\|_{1*} = \|u\|_\infty}$

2. We have $\min_x \|Ax - b\|_2^2 + \|x\|_1 \iff \min_{x,y} \|y\|_2^2 + \|x\|_1$ s.t. $Ax - b = y$

The associated Lagrangian is:

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

So $g(\nu) = \inf_{x,y} L(x,y,\nu) = -2 \sup_y \left( \frac{1}{2} \nu^T y - \frac{1}{2} \|y\|_2^2 \right) - \sup_x \left( \nu^T A x - \|x\|_1 \right) - \nu^T b$

- if $\|\nu^T A\|_\infty > 1$, $g(\nu) = -\infty$ according to 1.

• otherwise, recall that the conjugate of $f(x) = \frac{1}{2}\|x\|^2$ is $f^*(y) = \frac{1}{2}\|y\|_*^2$ (example 3.27 from the book)

and that the dual norm of $\|.\|_2$ is $\|.\|_2$ (**).
Then, $g(\nu) = -2 \times \frac{1}{2}\|\frac{1}{2}\nu\|_2^2 - \nu^T b = -\frac{1}{4}\|\nu\|_2^2 - \nu^T b$

• Therefore the dual of (RLS) is: $\min \ \frac{1}{4}\|\nu\|_2^2 + \nu^T b$
s.t. $\|\nu^T A\|_\infty \leq 1$

(**) Let $u \in \mathbb{R}^l$ ($l \in \mathbb{N}$).
• let $v \in \mathbb{R}^l$ such that $\|v\|_2 = 1$
$v^T u \leq \|v\|_2 \|u\|_2$ by Cauchy-Schwarz inequality
$\Rightarrow \|u\|_{2*} \leq \|u\|_2$
• $u^T u = \|u\|_2$ so $\|u\|_{2*} \geq \|u\|_2$
$\Rightarrow \underline{\|u\|_{2*} = \|u\|_2}$

## Exercise 3

1. We have (Sep. 1) $\Leftrightarrow \min_w \frac{1}{m\tau} \sum_{i=1}^m \max(0, 1 - y_i(w^T x_i)) + \frac{1}{2}\|w\|_2^2$

Now, we can note that for $i = 1, \dots, n$, $\min_{z_i} z_i$ s.t. $z_i \geq \max(0, 1 - y_i(w^T x_i))$

is equal to $\max(0, 1 - y_i(w^T x_i))$

and that $z_i \geq \max(0, 1 - y_i(w^T x_i)) \Leftrightarrow z_i \geq 1 - y_i(w^T x_i)$, $z_i \geq 0$

Therefore (Sep. 1) $\Leftrightarrow \min_w \frac{1}{m\tau} \sum_{i=1}^m \min_{z_i} z_i$ s.t. $z_i \geq 0, z_i \geq 1 - y_i(w^T x_i)$ $+ \frac{1}{2}\|w\|_2^2$

$$\Leftrightarrow \min_{\omega, z} \quad \frac{1}{m\tau} \mathbb{1}^T z + \frac{1}{2} \|\omega\|_2^2 \qquad (\text{Sep. } 2)$$

$$\text{s.t.} \quad z_i \geq 1 - y_i (\omega^T x_i) \quad \forall i = 1, \dots, n$$

$$z \geq 0$$

2. The associated Lagrangian of (Sep. 2) is

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

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

Recall that the conjugate of $f(x) = \frac{1}{2}\|x\|_2^2$ is $f^*(y) = \frac{1}{2}\|y\|_2^2$ (see previous exercise)

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

Therefore the dual problem of (Sep 2) is equivalent to

$$\max \quad \mathbb{1}^T \lambda - \frac{1}{2} \| \sum_{i=1}^m \lambda_i y_i x_i \|_2^2$$

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

$$\pi \geq 0$$

$$\frac{1}{m\tau}\mathbb{1} - \lambda - \pi = 0$$

$$\Leftrightarrow \max \quad \mathbb{1}^T \lambda - \frac{1}{2} \| \sum_{i=1}^m \lambda_i y_i x_i \|_2^2$$

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

$$\lambda \leq 1/n\tau$$

## Exercise 4

Let (P) the problem $\max \; a^T x$ s.t. $Cx \leq d$ $\left( \Leftrightarrow \min \; -a^T x \text{ s.t. } Cx \leq d \right)$

The associated Lagrangian is $L(a,z) = -a^T x + z^T(Ca - d)$
$$= (-x + C^T z)^T a - z^T d$$
$$\Rightarrow g(z) = \inf_a L(a,z) = \begin{cases} -z^T d & \text{if } C^T z = x \\ -\infty & \text{otherwise} \end{cases}$$

The dual problem of (P) is

$$(D): \begin{array}{ll} \max & -d^T z \\ \text{s.t.} & z \geq 0 \\ & C^T z = x \end{array} \Leftrightarrow \begin{array}{ll} \min & d^T z \\ \text{s.t.} & z \geq 0 \\ & C^T z = x \end{array}$$

Now, we will prove that

$$\overset{(*_1)}{\begin{array}{ll} \min_x & c^T x \\ \text{s.t.} & \left[\begin{array}{ll} \min & d^T z \\ \text{s.t.} & z \geq 0 \\ & C^T z = x \end{array}\right] \leq b \end{array}} \Leftrightarrow \overset{(*_2)}{\begin{array}{ll} \min_{x,z} & c^T x \\ \text{s.t.} & d^T z \leq b \\ & C^T z = x \\ & z \geq 0 \end{array}}$$

Indeed, • suppose $x^*, z^*$ are optimal for $(*_2)$, then $x^*$ is also feasible for $(*_1)$ and the objective values are the same.

• suppose $x^*$ an optimal solution of $(*_1)$, then $\min_{z \geq 0,\, C^T z = x^*} d^T z \leq b$

so there exists $z^*$ such that $\begin{cases} d^T z^* \leq b \\ C^T z^* = x^* \\ z^* \geq 0 \end{cases}$

Hence $(x^*, z^*)$ is feasible with the same objective value as $(*_1)$

• Thus, if we note $p_1^*$ and $p_2^*$ the optimal values of $(*_1)$ and $(*_2)$ we have proven that $p_1^* \leq p_2^*$ and $p_2^* \leq p_1^*$ so $p_1^* = p_2^*$

Finally, the problem (P) is linear program that is feasible since $\mathcal{P}$ is non empty, therefore strong duality holds for (P), so:

$$\sup_{a \in \mathcal{P}} a^T x = \min_{z \geq 0,\, C^T z = x} d^T z \Rightarrow \begin{array}{ll} \min_x & c^T x \\ \text{s.t.} & \sup_{a \in \mathcal{P}} a^T x \leq b \end{array} \Leftrightarrow (*_1)$$

We have proven that:

$$\begin{array}{ll} \min_x & c^T x \\ \text{s.t.} & \sup_{a \in \mathcal{P}} a^T x \leq b \end{array} \iff \begin{array}{ll} \min_{x,z} & c^T x \\ \text{s.t.} & d^T z \leq b \\ & C^T z = x \\ & z \geq 0 \end{array}$$

## Exercise 5

1. Let the problem (3) $\min \; c^T x$

s.t. $Ax \leq b$ , $x_i(1-x_i) = 0 \quad i = 1, \dots, n$

The associated Lagrangian is $\underline{L(x, \lambda, \nu)} = c^T x + \lambda^T (Ax - b) + \nu^T (x(1-x))$

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

$$= -x^T \operatorname{diag}(\nu) x + (c + A^T \lambda + \nu)^T x - b^T \lambda$$

- if $\exists k \in [\![1, \dots, n]\!]$ such that $\nu_k > 0$, we have $\inf_x L(x, \lambda, \nu) = -\infty$, (choosing $x_t = (0, \dots, 0, t, 0, \dots, 0)$ ($k^{th}$ coordinate) and letting $t$ grow to infinity we have $L(x, \lambda, \nu) \sim -t^2 \nu_k \to -\infty$)
- if $\exists k$ that $\nu_k = 0$ and $(c + A^T \lambda)_k \neq 0$ we have $\inf_x L(x, \lambda, \nu) = \infty$ (choosing $x_t = (0, \dots, 0, t, 0, \dots, 0)$ and letting $t$ grow to $\pm$ infinity $L(x, \lambda, \nu) \xrightarrow[t \to \pm\infty]{} -\infty$
- otherwise : $x \mapsto L(x, \lambda, \nu)$ is coercive, so it reaches its minimum on $x$ such that $\nabla_x L = 0$, if $\nu \neq 0$ (if $\nu = 0$, with the assumption made, we have $L(x, \lambda, \nu) = -b^T \lambda$ so it reaches its minimum everywhere)

We have $\nabla_x L(x, \lambda, \nu) = -2 \operatorname{diag}(\nu)\, x + (c + A^T \lambda + \nu)$

so $\nabla_x L = 0 \iff 2 \operatorname{diag}(\nu)\, x = c + A^T \lambda + \nu$

$\iff \nu_i x_i = 1/2\, (c + A^T \lambda + \nu)_i$

We have $L(x,\lambda,\nu) = -b^T\lambda + \sum_{i=1}^{m} -\nu_i x_i^2 + \underbrace{(c+A^T\lambda+\nu)_i x_i}_{\nu_i x_i + (c+A^T\lambda)_i x_i}$

so for $x$ such that $\nabla_x L = 0$, we have

$$L(x,\lambda,\nu) = -b^T\lambda + \sum_{\substack{i=1\\ \nu_i\neq 0}}^{m}\left(-(\nu_i x_i)^2/\nu_i + (c+A^T\lambda+\nu)_i/\nu_i\,(\nu_i x_i)\right) + \underbrace{\sum_{\substack{i=1\\ \nu_i=0}}^{m}(c+A^T\lambda)_i x_i}_{=0 \text{ by assumption}}$$

and $-(\nu_i x_i)^2/\nu_i + (c+A^T\lambda+\nu)_i/\nu_i\,(\nu_i x_i)$

$$= \frac{1}{\nu_i}\times\left[-\frac{1}{4}(c+A^T\lambda+\nu)_i^2 + \frac{1}{2}(c+A^T\lambda+\nu)_i^2\right]$$

$$= 1/4\,(c+A^T\lambda+\nu)_i^2/\nu_i$$

Finally $$g(\lambda,\nu) = \begin{cases} -b^T\lambda + \sum_{i=1}^{m} 1/4\,(c+A^T\lambda+\nu)_i^2/\nu_i\,\mathbb{1}_{\nu_i\neq 0} \\ \quad \text{if } \nu\succeq 0 \text{ and if } \nu_k = 0 \Rightarrow (c+A^T\lambda)_k = 0 \\ -\infty \text{ otherwise} \end{cases}$$

(we note $\mathcal{H} = \{\nu,\lambda\in\mathbb{R}^n \mid \nu\succeq 0,\ \nu_k = 0 \Rightarrow (c+A^T\lambda)_k = 0\}$

The dual problem is $$\max_{\lambda,\nu}\ -b^T\lambda + \sum_{i=1}^{m} 1/4\,(c+A^T\lambda+\nu)_i^2/\nu_i\,\mathbb{1}_{\nu_i\neq 0}$$
$$\text{s.t. } (\nu,\lambda)\in\mathcal{H}$$
$$\lambda\succeq 0$$

Using the hint, the problem can be reduced to:

$$\max_{\lambda}\ -b^T\lambda + \sum_{i=1}^{m}\min(0;\ c_i + A_i^T\lambda) \qquad (D)$$
$$\text{s.t. } \lambda\succeq 0$$

(where $A_i$ is the $i^{th}$ ligne of $A$)

by optimizing over $\nu$

2. The Lagrangian of (2) is :

$$L(x, \lambda_1, \lambda_2, \lambda_3) = c^T x + \lambda_1^T (Ax - b) + \lambda_2^T(-x) + \lambda_3^T(x-1)$$
$$= (c + A^T\lambda_1 - \lambda_2 + \lambda_3)^T x - b^T\lambda_1 - \mathbb{1}^T\lambda_3$$

So $g(\lambda_1, \lambda_2, \lambda_3) = \begin{cases} -b^T\lambda_1 - \mathbb{1}^T\lambda_3 & \text{if } c + A^T\lambda - \lambda_2 + \lambda_3 = 0 \\ -\infty & \text{otherwise} \end{cases}$

The dual problem is:
$$\boxed{\begin{aligned} \max_{\lambda_1, \lambda_2, \lambda_3} \quad & -b^T\lambda_1 - \mathbb{1}^T\lambda_3 \\ \text{s.t.} \quad & \lambda_1, \lambda_2, \lambda_3 \geq 0 \\ & c + A^T\lambda_1 - \lambda_2 + \lambda_3 = 0 \end{aligned}}$$

$$\Leftrightarrow \max_{\lambda, \lambda_3} \ -b^T\lambda - \sum_{i=1}^m \lambda_{3_i} \quad \Leftrightarrow \max_{\lambda, \lambda'} \ -b^T\lambda + \sum_{i=1}^m \lambda'_i$$
$$\text{s.t.} \quad \lambda \geq 0, \ \lambda_3 \geq 0 \qquad \text{s.t.} \quad \lambda \geq 0, \ \lambda' \leq 0$$
$$c + A^T\lambda + \lambda_3 \geq 0 \qquad c + A^T\lambda \geq \lambda'$$

maximizing over $\lambda'$ such that $\lambda'_i \leq 0$ and $\lambda'_i \leq c + A^T\lambda$

we have that $\max \sum_{i=1}^m \lambda_i' = \sum_{i=1}^m \min(0; (c + A^T\lambda)_i)$

So the dual problem is equivalent to the problem (D) obtained in 1.

Therefore, the lower bound obtained via Lagrangian relaxation and via the LP relaxation are the same