CALLARD Baptiste Homework 2 - Convex Optimization.

exercice 1: let $c \in \mathbb{R}^d$, $b \in \mathbb{R}^n$ and $A \in \mathbb{R}^{n \times d}$

1) Let's compute the dual of (P):

$$(P) \begin{cases} \min_x & c^T x \\ \text{st.} & Ax = b \\ & x \geq 0 \end{cases} \iff \begin{cases} \min_x & c^T x \\ \text{st.} & Ax = b \\ & -x \leq 0 \end{cases} \text{ (standart form).}$$

let $x \in \mathbb{R}^d$, $\lambda \in \mathbb{R}^d$, $\nu \in \mathbb{R}^n$, the lagrangian function of (P) is:

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

The Lagrangian dual function is

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

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

In fact, the lagrangian is an affine function of x.

Hence, if $(c^T + \nu^T A - \lambda^T) \neq 0$ it goes to $-\infty$:

* $c^T + \nu^T A - \lambda^T < 0$ it goes to $-\infty$ when $x \to +\infty$
* $c^T + \nu^T A - \lambda^T > 0$ it goes to $-\infty$ when $x \to -\infty$

Therefore:

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

Finally the dual of (P) is

$$\begin{cases} \max_{\lambda, \nu} -\nu^T b \\ \text{st } c^T + \nu^T A - \lambda^T = 0 \\ \lambda \geq 0 \end{cases}$$

As λ doesn't appear in objective, we can simplify as: $c^T + \nu^T A - \lambda^T = 0$ and $\lambda \geq 0$
$\Leftrightarrow A^T \nu + c^T \geq 0$

$$\boxed{\begin{cases} \max_{\nu} -b^T \nu \\ \text{st } A^T\nu + c \geq 0 \end{cases}} \underset{\Leftrightarrow}{\nu \to -\nu} \boxed{\begin{array}{l} \max_{\nu} b^T \nu \\ \text{st } A^T \nu \leq c \end{array}}$$

The dual of (P) is (D).

2) Let's derive the dual of (D):

$$(D) : \begin{cases} \max_y b^T y \\ \text{st. } A^T y \leq c \end{cases} \Leftrightarrow \begin{cases} -\min_y -b^T y \\ \text{st } A^T y \leq c \end{cases} \quad \text{(standart form)}$$

Let $y \in \mathbb{R}^n$, $\lambda \in \mathbb{R}^d$. The Lagrangian of (D') $\begin{cases} \min_y -b^T y \\ \text{st } A^T y \leq c \end{cases}$ is:

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

The lagrangian dual is

$$g(\lambda) = \inf_y -b^T y + \lambda^T (A^T y - c) = \inf_y (-b^T + \lambda^T A^T) y - \lambda^T c$$

Similarly the Lagrangian is an affine function in y.

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

Therefore, the dual of (D') $\begin{cases} \min\limits_{y} -b^Ty \\ \text{st: } A^Ty \leq c \end{cases}$ is

$$\begin{cases} \max\limits_{\lambda} -\lambda^Tc \\ \text{st } -b^T + \lambda^TA = 0 \\ \lambda \geq 0 \end{cases} \iff \begin{cases} \max\limits_{\lambda} -\lambda^Tc \\ \text{st: } A\lambda = b \\ \lambda \geq 0 \end{cases}$$

; Finally the dual of the original problem (D) $\begin{cases} -\min\limits_{y} -b^Ty \\ \text{st } A^Ty \leq c \end{cases}$ is $\begin{cases} -\max\limits_{\lambda} -\lambda^Tc \\ A\lambda = b \\ \lambda \geq 0 \end{cases}$

This can be re-written as $\boxed{\begin{cases} \min\limits_{\lambda} c^T\lambda \\ A\lambda = b \\ \lambda \geq 0 \end{cases}}$. We can notice that it is exactly the problem (P).

3) Let's derive the dual of (Self-Dual) which is supposed to be itself.

Let $x \in \mathbb{R}^d$, $y \in \mathbb{R}^d$, $\lambda_1 \in \mathbb{R}^d$, $\lambda_2 \in \mathbb{R}^d$, $\nu \in \mathbb{R}^n$

(Self Dual) $\begin{cases} \min\limits_{x,y} c^Tx - b^Ty \\ Ax = b \ (\nu) \\ -x \leq 0 \ (\lambda_1) \\ A^Ty \leq c \ (\lambda_2) \end{cases}$ in standart form.

Its lagrangian is:

$$\mathcal{L}: \begin{array}{ccc} \mathbb{R}^d \times \mathbb{R}^n \times \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^n & \longrightarrow & \mathbb{R} \\ x, y, \lambda_1, \lambda_2, \nu & \longrightarrow & c^Tx - b^Ty + \nu^T(Ax-b) - \lambda_1^Tx + \lambda_2^T(A^Ty - c) \end{array}$$

The lagrangian dual function is:

$$g(\lambda_1, \lambda_2, \nu) = \inf_{x,y} (c^Tx - b^Ty + \nu^T(Ax - b) - \lambda_1^Tx + \lambda_2^T(A^Ty - c))$$

It can be dissociated in an affine function of x and an affine function of y. We can the apply the same argument as previously:

$$g(\lambda_1, \lambda_2, \nu) = \inf_{x} (c^T + \nu^TA - \lambda_1^T)x + \inf_{y} (-b^T + \lambda_2^TA^T)y - \nu^Tb - \lambda_2^Tc$$

$$= \begin{cases} -\nu^Tb - \lambda_2^Tc \text{ if } (c^T + \nu^TA - \lambda_1^T) = (-b^T + \lambda_2^TA^T) = 0 \\ -\infty \qquad \text{otherwise.} \end{cases}$$

Finally we can get the dual of (Self Dual).

$$\begin{cases} \max\limits_{\lambda_1,\lambda_2,\upsilon} -\upsilon^T b - \lambda_2^T c \\ \quad c^T + \upsilon^T A - \lambda_1^T = 0 \\ \quad -b^T + \lambda_2^T A^T = 0 \\ \quad \lambda_1 \geq 0 \\ \quad \lambda_2 \geq 0 \end{cases} \iff \begin{cases} -\min\limits_{\lambda_2,\upsilon} \upsilon^T b + \lambda_2^T c \\ \quad A^T\upsilon + c \geq 0 \\ \quad A\lambda_2 = b \\ \quad \lambda_2 \geq 0 \end{cases}$$

We can define $\mu^T = -\upsilon^T$ Then, dual of (Self Dual) is equivalent to:

$$\begin{cases} -\min\limits_{\mu,\lambda_2} -b^T\mu + c^T\lambda_2 \\ \quad c \geq A^T\mu \\ \quad A\lambda_2 = b \\ \quad \lambda_2 \geq 0 \end{cases} \iff \begin{cases} \min\limits_{\mu,\lambda_2} -b^T\mu + c^T\lambda_2 \\ \quad c \geq A^T\mu \\ \quad A\lambda_2 = b \\ \quad \lambda_2 \geq 0 \end{cases}$$

Solving, the two problems above is equivalent as we will get the same optimal values λ_2^* and μ^* (Just objective will be opposite).

Finally:

This is exactly the problem (Self-Dual) if we rename (λ_2, μ) by (x, y). Thus this problem is self-dual

4) If we look more closely to the problem (Self-Dual) it is clear that constraints in x can be separated of those in y. Thus this join minimization can be optimized independantly (ie first in x and then in y or inversely). Hence, using $\min(-a) = -\max(a)$ we get:

$$\begin{array}{l} \min\limits_{x,y} c^Tx - b^Ty \\ \text{st. } Ax = b \\ \quad x \geq 0 \\ \quad A^Ty \leq c \end{array} = \underbrace{\begin{array}{l} \min\limits_{x} c^Tx \\ \text{st } Ax = b \\ \quad x \geq 0 \end{array}}_{(P)} - \underbrace{\begin{array}{l} \max\limits_{y} b^Ty \\ \text{st } A^Ty \leq c \end{array}}_{(D)} \qquad (*)$$

So (Self Dual) is the substraction of (P) by (D).

We know that (Self Dual) is feasible and bounded. Let (x^*, y^*) its optimal solution.

If we solve (P) and (D) we can optain respectively x^* and y^* optimal solutions. In the same time, we also solve (Self Dual) so (x^*, y^*) can also be obtain by solving (P) and (D).

Moreover, the dual of (P) is in fact exactly (D).

This can be seen with question 1 and 2.

We have strong duality as (P) is a linear problem (linear objective and linear constraints) which is feasible.

So if we note p^* the objective of (P) in x^* and d^* the objective of (D). Strong duality gives:

$p^* = d^*$.

Finally,

$$\begin{cases} \min\limits_{x,y} c^T x - b^T y = p^* - d^* = 0 \\ \text{st } Ax = b \\ \quad x \geq 0 \\ \quad A^T y \leq c \end{cases}$$

exercice 2 :

1) Let's compute the conjugate of $\|x\|_1$.

We note f^* the conjugate of $\|\cdot\|_1$, let $x \in \mathbb{R}^d$, $y \in \mathbb{R}^d$.

$$f^*(y) = \sup_{x \in \mathbb{R}^d} \sum_{i=1}^{m} x_i y_i - |x_i|$$

We will proceed by disjunction of cases :

* if there exists $i \in \{1, \dots, d\}$ such that $y_i > 1$ then let's consider $x = [0 \dots 0 \underset{\substack{\nearrow \\ ith}}{t} 0 \dots 0]$ with t in ith position.
with $t > 0$.

$$\sum_{i=1}^{m} x_i y_i - |x_i| = t y_i - t = t \underbrace{(y_i - 1)}_{>0} \xrightarrow[t \to +\infty]{} +\infty$$

* if there exists $i \in \{1, \dots, d\}$ such that $y_i < -1$ then let's consider $x = [0 \dots 0 \underset{\substack{\uparrow \\ ith}}{t} 0 \dots 0]$ with t in the ith position
with $t < 0$

$$\sum_{i=1}^{m} x_i y_i - |x_i| = t y_i + t = t \underbrace{(y_i + 1)}_{<0} \xrightarrow[t \to -\infty]{} +\infty$$

* if $-1 \leq y \leq 1$ then we can see that

$$\sum_{i=1}^{m} x_i y_i - |x_i| \leq \sum_{i=1}^{m} \underbrace{|x_i|}_{\geq 0} \underbrace{(|y_i| - 1)}_{\leq 0} \leq 0 \quad \text{then}$$

$\sup_{x} \sum_{i=1}^{m} x_i y_i - |x_i| \leq 0$ but for $x = (0 \dots\dots 0)$ the bound is reached.

We can conclude that:

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

We can note that $[-1 \leq y \leq 1] \iff \|y\|_\infty \leq 1$.

2) Let's compute the dual of the (RLS).

The (RLS) is equivalent to a new problem, let $x \in \mathbb{R}^d$, $y \in \mathbb{R}^d$, $\nu \in \mathbb{R}^d$:

$$\begin{cases} \min_{x,y} \|y\|_2^2 + \|x\|_1^1 \\ \text{st} \quad y = Ax - b \end{cases}$$

The Lagrangian function is

$$\mathcal{L} : \begin{array}{l} \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d \longrightarrow \mathbb{R} \\ x, y, \nu \longrightarrow \|y\|_2^2 + \|x\|_1^1 + \nu^T(y - Ax + b) \end{array}$$

The Lagrangian dual is

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

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

$$= -\sup_x ((A^T\nu)^T x - \|x\|_1) + \inf_y (\|y\|_2^2 + \nu^T y) + \nu^T b$$

$$= -f^*(A^T\nu) + \inf_y (\|y\|_2^2 + \nu^T y) + \nu^T b \quad \text{(from question 2.1)}$$

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

Let's focus on $\inf_y \|y\|_2^2 + \nu^T y$. We note $h : y \mapsto \|y\|_2^2 + \nu^T y$

h is differentiable and convex thus

$\nabla h(y) = 2y + \nu$ so $\nabla h(\tilde{y}) = 0 \iff \tilde{y} = -\frac{1}{2}\nu$ is the minimum.

$h(\tilde{y}) = \|\frac{1}{2}\nu\|_2^2 - \frac{1}{2}\nu^T\nu = -\frac{1}{4}\|\nu\|_2^2$

Finally:

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

Hence the dual of (RLS) is

$$\boxed{\begin{cases} \max_\nu & -\frac{1}{4}\|\nu\|_2^2 + b^T\nu \\ \text{st} & \|A^T\nu\|_\infty \leq 1 \end{cases}}$$

exercice 3:

1) We have n data points $x_i \in \mathbb{R}^d$, $y_i \in \{-1, 1\}$, $w \in \mathbb{R}^d$

Let's define $\mathcal{L}(w, x_i, y_i) = \max(0, 1 - y_i(w^T x_i))$ and

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

(Sep 2) $\left\{ \begin{array}{l} \min \frac{1}{n\tau} \mathbb{1}^T z + \frac{1}{2} \|w\|_2^2 \\ \text{st } \forall i \in [\![1, m]\!],\ z_i \geq 1 - y_i(w^T x_i) \\ z \geq 0 \end{array} \right.$

Start from (Sep 1)

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

τ is a regularization parameter and makes sens only if $\tau > 0$, we can devide by τ. We add also the variable $z_i = \max(0, 1 - y_i(w^T x_i))$

(Sep 1) $\Leftrightarrow \left\{ \begin{array}{l} \min_{z,w} \frac{1}{n\tau} \mathbb{1}^T z + \frac{1}{2} \|w\|_2^2 \\ \text{st } \forall i \in [\![1, m]\!]\ z_i = \max(0, 1 - y_i(w^T x_i)) \end{array} \right.$

It is almost finished, we just have to show that:

$$\left[\begin{array}{l} \forall i \in [\![1, m]\!],\ z_i \geq 1 - y_i(w^T x_i) \\ z \geq 0 \end{array} \right] \Leftrightarrow \left[\forall i \in [\![1, m]\!]\ z_i = \max(0, 1 - y_i(w^T x_i)) \right]$$

($\Leftarrow$) By definition of the max.

($\Rightarrow$) We have $\forall i \in [\![1, m]\!]$ $z_i \geq \max(0, 1 - y_i(w^T x_i))$. Let's suppose that at optimality we have $z_i^* > \max(0, 1 - y_i(w^{*T} x_i))$. It means that there exists $\varepsilon > 0$ such that $\tilde{z} = z^* - \varepsilon \geq \max(0, 1 - y_i(w^{*T} x_i))$. Thus the objective function would decrease in $\tilde{z} = z^* - \varepsilon \geq 0$ Contradition. And $z_i^* = \max(0, 1 - y_i(w^{*T} x_i))$.

Therefore we have equivalence at optimality.

Finally: (sep 1) $\Leftrightarrow \left\{ \begin{array}{l} \min_{z,w} \frac{1}{n\tau} \mathbb{1}^T z + \frac{1}{2} \|w\|_2^2 \\ \forall i \in [\![1, m]\!]\ z_i \geq 1 - y_i(w^T x_i) \\ z \geq 0 \end{array} \right.$

This is exactly (sep 2). There fore we have showed that (Sep 2) solves problem (Sep 1).

2) Let's compute the dual of (Sep 2).

Let $\lambda \in \mathbb{R}^n$, $\pi \in \mathbb{R}^n$

The lagrangian is

$$\mathcal{L} \quad \begin{aligned} \mathbb{R}^d \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n &\longrightarrow \mathbb{R} \\ w, z, \lambda, \pi &\longrightarrow \frac{1}{n\tau} \mathbb{1}^T z + \frac{1}{2}\|w\|_2^2 + \left[\sum_{i=1}^m \lambda_i (1 - y_i(w^T x_i) - z_i)\right] - \pi^T z \end{aligned}$$

The lagrangian dual is:

$$g(\lambda, \pi) = \inf_{w,z} \mathcal{L}(w, z, \lambda, \pi)$$

$$= \inf_w \left[\frac{1}{2}\|w\|_2^2 - \sum_{i=1}^m \lambda_i y_i w^T x_i\right] + \inf_z \left[\frac{1}{n\tau}\mathbb{1}^T z - \sum_{i=1}^m \lambda_i z_i - \pi^T z\right] + \sum_{i=1}^m \lambda_i$$

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

the minimization over z is a minimization of an linear function then the infimum is $-\infty$ except when $\frac{1}{n\tau}\mathbb{1}^T - \lambda^T - \pi^T = 0$

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

We define $h(w) = \frac{1}{2}\|w\|_2^2 - \sum_{i=1}^m \lambda_i y_i w^T x_i$ which is a twice differentiable function

$\nabla h(w) = w - \sum_{i=1}^m \lambda_i y_i x_i$ and convex $\nabla^2 h(w) = \mathbb{1} > 0$

$\nabla h(\tilde{w}) = 0 \iff \tilde{w} = \sum_{i=1}^m \lambda_i y_i x_i$ is the minimum

The minimum is reached for :

$$h(\tilde{w}) = \frac{1}{2}\left\|\sum_{i=1}^m \lambda_i y_i x_i\right\|^2 - \sum_{i=1}^m \lambda_i y_i \left(\sum_{j=1}^m \lambda_j y_j x_j\right)^T x_i$$

$$h(\tilde{\omega}) = \frac{1}{2} \| \sum_{i=1}^{m} \lambda_i y_i x_i \|_2^2 - \| \sum_{i=1}^{m} \lambda_i y_i x_i \|_2^2$$

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

Thus :

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

The dual of (Sep 2) is

$$\begin{cases} \max_{\lambda, \pi} -\frac{1}{2} \| \sum_{i=1}^{m} \lambda_i y_i x_i \|_2^2 + \mathbb{1}^T \lambda \\ \frac{1}{m\tau} \mathbb{1} - \pi - \lambda = 0 \\ \lambda \geq 0 \\ \pi \geq 0 \end{cases}$$

This can be simplify as

$$\boxed{\begin{cases} \max_{\lambda} -\frac{1}{2} \| \sum_{i=1}^{m} \lambda_i y_i x_i \|_2^2 + \mathbb{1}^T \lambda \\ \frac{1}{m\tau} \mathbb{1} \geq \lambda \\ \lambda \geq 0 \end{cases}}$$

exercice 4:

let (R1)
$$\begin{cases} \min\limits_{x} c^T x \\ \text{st } \sup\limits_{a \in P} a^T x \leqslant b \end{cases}$$

• First we use the hint and derive the dual of

$$(P) \begin{cases} \max\limits_{a} a^T x \\ c^T a \leqslant d \end{cases} \iff \begin{cases} -\min\limits_{a} -a^T x \\ \text{st: } c^T a \leqslant d \end{cases}$$, we compute the dual of $$\begin{cases} \min\limits_{a} -a^T x \\ \text{st: } c^T a \leqslant d \end{cases}$$

• let $\lambda \in \mathbb{R}^m$

The lagrangian is

$$\mathcal{L}(x, \lambda) = -a^T x + \lambda^T (c^T a - d)$$

and the lagrangian dual is

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

Hence the dual is

$$\begin{cases} \max\limits_{\lambda} -\lambda^T d \\ \text{st } c\lambda = x \\ \lambda \geqslant 0 \end{cases}$$ and the dual of (P) is $$\begin{cases} -\max\limits_{\lambda} -\lambda^T d \\ \text{st } c\lambda = x \\ \lambda \geqslant 0 \end{cases} \iff \begin{cases} \min\limits_{\lambda} \lambda^T d \\ \text{st } c\lambda = x \\ \lambda \geqslant 0 \end{cases}$$

• This is a linear program, hence strong duality holds thus $(\sup\limits_{a \in P} a^T x \leqslant b) = (\min\limits_{\lambda} \lambda^T d$ st $c\lambda = x$ and $\lambda \geqslant 0)$.

$$(R1) \iff \begin{cases} \min\limits_{x} c^T x \\ \text{st } \begin{cases} \min\limits_{\lambda} \lambda^T d \\ \text{st } C\lambda = x \\ \lambda \geqslant 0 \end{cases} \leqslant b \quad (*) \end{cases}$$ by strong duality of (LP)

But to show that $(*)$ is true it is sufficient to find $\tilde{\lambda} \in \mathbb{R}^m$ such that $\tilde{\lambda}^T d \leqslant b$ and $C\tilde{\lambda} = x$. Because if one $\tilde{\lambda}$ verifies these conditions, then in particular it implies that the minimum verifies it too.

There fore, it is equivalent to solve (R1) or (R2) as:

$$\begin{cases} \min\limits_{x} c^T x \\ \text{st} \begin{cases} \min\limits_{\lambda} \lambda^T d \\ \text{st } C\lambda = x \\ \lambda \geq 0 \end{cases} \leq b \end{cases} \iff \begin{cases} \min\limits_{x} c^T x \\ \text{st } \lambda^T d \leq b \\ C\lambda = x \\ \lambda \geq 0 \end{cases} \iff (R2)$$

exercice 5 :

1) Let's derive the dual of the boolean LP (BLP) :

$$\begin{cases} \min_x & c^T x \\ & Ax \leq b \\ & x_i(1-x_i) = 0 \quad \forall i \in \{1, \dots, m\} \end{cases}$$

The lagrangian is :

$$\mathcal{L}(x, \lambda, \nu) = c^T x + \lambda^T (Ax - b) + \sum_{i=1}^{m} \nu_i x_i (1 - x_i)$$

$$= \sum_{i=1}^{m} (c_i + \lambda_i a_i^T + \nu_i) x_i - \nu_i x_i^2 - \lambda^T b$$

with a_i the ith column of A.

Let note $h_i(x_i) = (c_i + \lambda_i a_i^T + \nu_i) x_i - \nu_i x_i^2$, h_i is is twice derivable for $i = 1, \dots, m$.

$\nabla h_i(x_i) = c_i + \lambda_i a_i^T + \nu_i - 2\nu_i x_i$

$\nabla^2 h_i(x_i) = -2\nu_i$

So $\tilde{x}_i$ is the minimum if $\nabla h_i(\tilde{x}_i) = 0$ and $\nabla^2 h_i(\tilde{x}_i) \geq 0$

Thus $\tilde{x}_i$ is the minimum if $\begin{cases} \tilde{x}_i = \dfrac{c_i + \lambda_i a_i^T + \nu_i}{2\nu_i} \\ \nu_i \leq 0 \end{cases}$

Hence $g(\lambda, \nu) = \inf_x \mathcal{L}(x, \lambda, \nu)$

$$= \sum_{i=1}^{m} \frac{(c_i + \lambda_i a_i^T + \nu_i)^2}{2\nu_i} - \frac{(c_i + \lambda_i a_i^T + \nu_i)^2}{4\nu_i} - \lambda^T b$$

$$= \sum_{i=1}^{m} \frac{(c_i + \lambda_i a_i^T + \nu_i)^2}{4\nu_i} - \lambda^T b$$

So the dual is :

$$\begin{cases} \max\limits_{\substack{\nu_i \leq 0 \\ \lambda}} \sum_{i=1}^{m} \frac{(c_i + \lambda_i a_i^T + \nu_i)^2}{4\nu_i} - \lambda^T b \\ \lambda \geq 0 \end{cases}$$

$$\Leftrightarrow \begin{cases} \max\limits_{\substack{\nu_i \geq 0 \\ \lambda}} \sum_{i=1}^{m} -\frac{(c_i + \lambda_i a_i^T - \nu_i)^2}{4\nu_i} - \lambda^T b \\ \lambda \geq 0 \end{cases}$$ change of variable $\nu_i \to -\nu_i$

The variables $(\nu_1, \dots, \nu_n, \lambda)$ are not dependent. Thus we can optimize first with respect to ν_i. We the Hint, we can get :

$$\begin{cases} \max\limits_{\lambda} \sum_{i=1}^{m} \min(0, c_i + a_i^T \lambda_i) - \lambda^T b \\ \lambda \geq 0 \end{cases}$$

$$\Leftrightarrow \begin{cases} \max\limits_{z, \lambda} \mathbb{1}^T z - \lambda^T b \\ z \leq 0 \\ z \leq c + A^T \lambda \quad \forall i = 1, \dots, m \\ \lambda \geq 0 \end{cases}$$

↳ we use the same trick than exercice 3. As we are in maximization only inequalities are necessary.

2) The LP relaxation is

$$\begin{cases} \min c^T x \\ \text{st} \quad Ax \leq b \\ \quad -x \leq 0 \\ \quad x \leq \mathbb{1} \end{cases}$$

As it is a LP, strong duality holds. Therefore, we can get much information for it. The lagrangian is:

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

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

it is affine in x:

$$g(\lambda_1, \lambda_2, \lambda_3) = \inf_x \mathcal{L}(x, \lambda_1, \lambda_2, \lambda_3)$$

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

The dual is

$$\begin{cases} \max\limits_{\lambda_1, \lambda_2, \lambda_3} -\lambda_1^T b - \lambda_3^T \mathbb{1} \\ \quad c^T + \lambda_1^T A - \lambda_2^T + \lambda_3^T = 0 \\ \quad \lambda_1 \geq 0 \\ \quad \lambda_2 \geq 0 \\ \quad \lambda_3 \geq 0 \end{cases}$$

it is equivalent to

$$\begin{cases} \max\limits_{\lambda_1, \lambda_3} & -\lambda_1^T b - \lambda_3^T \mathbb{1} \\ & c + A^T\lambda_1 + \lambda_3 \geq 0 \\ & \lambda_1 \geq 0 \\ & \lambda_3 \geq 0 \end{cases}$$

if we note $z = -\lambda_3$ and $\lambda_1 = \lambda$

$$\begin{cases} \max\limits_{z, \lambda} & -\lambda^T b + \mathbb{1}^T z \\ & c + A^T\lambda \geq z \\ & z \leq 0 \\ & \lambda \geq 0 \end{cases}$$

we get exactly the same problem. Recall that we have strong duality. Thus the lower bound of the LP relaxation or its dual are the same.

Moreover the dual of its relaxation is the same as the dual of Boolean LP. Thus lower bounds are the same.