DAVID SOTO

CONVEX OPTIMIZATION

HOMEWORK 2

Exercise 1. For a given $c \in \mathbb{R}^d$, $b \in \mathbb{R}^n$ and $A \in \mathbb{R}^{n \times d}$ consider the two following linear optimization problems:

(P) $\displaystyle \min_x \; c^T x$ s.t. $Ax = b$, $x \geq 0$

(D) $\displaystyle \max_y \; b^T y$ s.t. $A^T y \leq c$

1. Let us compute the dual problem of (P).
The Lagrangian is:

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

The Lagrangian is an affine function of x.
It follows that the dual function g is given by

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

By definition, the dual problem is given by

$$\max \; g(\lambda,\nu) \quad \text{s.t.} \quad \lambda \geq 0$$

Hence, after making the implicit constraints explicit, we obtain:

(Dual Problem) $\max \; \nu^T b$ s.t. $\boxed{\begin{matrix} c - \lambda - A^T\nu = 0 \\ \lambda \geq 0 \end{matrix}}$ (*)

(*) $\left.\begin{matrix} A^T\nu = c - \lambda \\ \lambda \geq 0 \end{matrix}\right\} \iff A^T \nu \leq c$

which gives us:

(Dual Problem) $\max \; b^T \nu$ s.t. $A^T \nu \leq c$ (D)

(D) is the dual problem of (P)!

2. Let us compute the dual problem of (Δ)
First of all, we need to rewrite (Δ) in the "standard form"

$$(\Delta) \quad \boxed{\begin{array}{ll} \min\limits_{y} & -b^T y \\ \text{s.t.} & A^T y \leq C \end{array}}$$

The Lagrangian is:

$$\begin{aligned} \mathcal{L}(y,\lambda) &= -b^T y + \lambda^T (A^T y - C) \\ &= -b^T y + \lambda^T A^T y - \lambda^T C \\ &= (\lambda^T A^T - b^T)\, y - \lambda^T C \end{aligned}$$

The Lagrangian is an affine function of y.
It follows that the dual function g is given by:

$$g(\lambda) = \inf_{y} \mathcal{L}(y,\lambda) = \begin{cases} -\lambda^T C & \text{if } A\lambda - b = 0 \\ -\infty & \text{otherwise.} \end{cases}$$

By definition the dual problem is:

$$\text{Dual} \quad \begin{array}{ll} \text{Max} & g(\lambda) \\ \text{s.t.} & \lambda \geq 0 \end{array}$$

Hence, after making the implicit constraints explicit, we obtain:

$$(\text{Dual Problem}) \quad \boxed{\begin{array}{ll} \text{Max} & -\lambda^T C \\ \text{s.t.} & A\lambda = b \\ & \lambda \geq 0 \end{array}}$$

$$(\text{Dual Problem}) \iff \boxed{\begin{array}{ll} \text{Min} & C^T \lambda \\ \text{s.t.} & A\lambda = b \\ & \lambda \geq 0 \end{array}}$$ we recognize problem (P).

Hence, we can deduce that the dual of (Δ) is (P).

3. (Self-Dual)

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

Let us prove that the problem (Self-Dual) is actually a Self-Dual problem!

The dual function of this problem is by definition:

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

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

→ * We first minimize over y by setting the gradient with respect to y equal to zero: this yields:

$$-b^T + \lambda_2^T A^T = 0 \iff A\lambda_2 = b$$

→ * We can minimize over x by setting the gradient with respect to x equal to zero. We obtain:

$$c^T - \nu^T A - \lambda_1^T = 0 \iff A^T \nu + \lambda_1 = c$$

* N.B: we have minimized over x and y this way because the lagrangian is linear in x and y!

Hence, the dual function is given by:

$$g(\lambda_1, \lambda_2, \nu) = \begin{cases} \nu^T b - \lambda_2^T c & \text{if } A\lambda_2 = b \text{ and } A^T \nu + \lambda_1 = c \\ -\infty & \text{otherwise} \end{cases}$$

The dual problem of (Self-Dual) is therefore given by:

(Dual)
$$\begin{aligned} \max_{\nu, \lambda_2} \quad & \nu^T b - \lambda_2^T c \\ \text{s.t.} \quad & A\lambda_2 = b \\ & \lambda_2 \geq 0 \\ & \left.\begin{aligned} A^T\nu + \lambda_1 = c \\ \lambda_1 \geq 0 \end{aligned}\right\} \end{aligned}$$

However
$$\left.\begin{aligned} A^T\nu + \lambda_1 = c \\ \lambda_1 \geq 0 \end{aligned}\right\} \iff A^T\nu \leq c$$
Therefore, the dual can be written as follows:

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

This is equal to (Self-Dual) !

(here $x = \lambda_2$ & $y = \nu$)

Indeed, by changing the notations $\lambda_2 = x$ and $\nu = y$ we obtain:

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

Hence, we can conclude that (Self-Dual) is a self-dual problem.

(1)

4. • The constraints of the problem (Self-Dual) are independent, that is only depends on the variable x or the variable y and the constraints of (Self-Dual) are the constraints of (P) and (D) accumulated.
If we first optimize the function with respect to y and then with respect to x, we can write that:

$$\min_{x,y} c^T x - b^T y = \min_x c^T x + \min_y (-b^T y)$$

$$= \min_x c^T x - \max_y b^T y$$

under the constraints of (Self-Dual), that is the constraints of (P) and the constraints of (D).
We recognize that it is the problem (P) minus the problem (D)!
Therefore, we can conclude that by solving (P) and (D), which gives us the optimal points x^* and y^*, we obtain the optimal point (x^*, y^*) of the problem (Self-Dual)

• From what we have seen above, we have:

(Self-Dual)	=	(P)	−	(D)
$\min_{x,y} c^T x - b^T y$ s.t. $Ax = b$, $x \geq 0$, $A^T y \leq c$	=	$\min_x c^T x$ s.t. $Ax = b$, $x \geq 0$	−	$\max b^T y$ s.t. $A^T y \leq c$

From question 1 we know that problem (D) is the dual problem of (P)

The constraints in (P) are a linear equality and a linear inequality. $\Rightarrow$ the constraints are convex.
The objective function is linear therefore it is convex and the domain of the objective function ($\mathbb{R}^d$) is open.
The problem (P) is convex !
In this question we assume the problem (Self-Dual) feasible $\Rightarrow$ the problem (P) is also feasible !.
that is $\exists\, x \in \text{relint}\,\mathcal{D}$ (where $\mathcal{D} = \{x,\ Ax = b\ \&\ x \geq 0\}$) such that $x > 0$ $\Rightarrow$ Slater's condition is satisfied !
By Slater's condition we know that strong duality holds for problem (P), that is $p^* = d^*$ where p^* is the optimal value of (P) and d^* is the optimal value of (D) the dual of (P)
Hence, from what we have seen previously, we can deduce that:

(Self-Dual)

$$\boxed{\begin{aligned} \min_{x,y}\ & c^T x - b^T y \\ \text{s.t.}\ & Ax = b \\ & x \geq 0 \\ & A^T y \leq c \end{aligned}} = p^* - d^* = 0$$

Hence, we can conclude that the optimal value of (Self-Dual) is 0.

Exercise 2: For a given $A \in \mathbb{R}^{n \times d}$ and $b \in \mathbb{R}^n$, we consider the following optimization problem:

$$\underset{x}{\text{Min}} \ \|Ax - b\|_2^2 + \|x\|_1 \qquad \text{(RLS)}$$

1. Let us compute the conjugate of the function $f(x) = \|x\|_1$.
By definition, the conjugate of f is given by:

$$f^*(y) = \sup_x \left(y^T x - f(x) \right)$$

$$f^*(y) = \sup_x \left(y^T x - \|x\|_1 \right)$$

By Cauchy-Shwarz inequality, we know that

$$\langle y, x \rangle = y^T x \leq \|x\|_1 \cdot \|y\|_1 .$$

* If $\|y\|_1 \leq 1 \Rightarrow y^T x \leq \|x\|_1$ (Cauchy-Schwarz)

$$\Rightarrow y^T x - \|x\|_1 \leq 0$$

By definition of the supremum we obtain:

$$\sup_x \left(y^T x - \|x\|_1 \right) \leq 0$$

For $x = 0$ we have $y^T x - \|x\|_1 = 0$. We can conclude that if $\|y\|_1 \leq 1$ then

$$f^*(y) = \sup_x \left(y^T x - f(x) \right) = 0 .$$

* If $\|y\|_1 > 1$. By choosing $x = ty$ we obtain:

$$y^T x - \|x\|_1 = t\, y^T y - |t|\, \|y\|_1$$

$$= t \|y\|_1^2 - t \|y\|_1 \qquad (\text{if } t > 0)$$

$$y^T x - \|x\|_1 = t \left(\|y\|_1^2 - \|y\|_1 \right)$$

However, $\|y\|_1 > 1 \Rightarrow \|y\|_1^2 - \|y\|_1 > 0$

Therefore by letting $t \longrightarrow +\infty$ we obtain

$$y^T x - \|x\|_1 \longrightarrow +\infty \quad (\text{where } x = ty)$$

$$\Rightarrow \sup_x \left(y^T x - \|x\|_1 \right) = +\infty .$$

Hence, we can conclude that the conjugate of $\|x\|_1$ is given by:

$$\boxed{f^*(y) = \begin{cases} 0 & \text{if } \|y\|_1 \leq 1 \\ +\infty & \text{otherwise} \end{cases}}$$

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

2. We start by introducing a new variable and equality constraint and reformulate (RLS):

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

The dual function of (RLS) is therefore given by:

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

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

$\rightarrow$ We first minimize over x:

$$\inf_x \left\{ \|x\|_1 - \nu^T A x \right\} = -\sup_x \left\{ \nu^T A x - \|x\|_1 \right\}$$

$$= -\|\nu^T A\|_1^* \quad \text{(where } \|x\|_1^* \text{ is}$$

the conjugate function of $\|x\|_1$). Using the result of question 1, we obtain:

$$\inf_x \left\{ \|x\|_1 - \nu^T A x \right\} = \begin{cases} 0 & \text{if } \|\nu^T A\|_1 \leq 1 \\ -\infty & \text{otherwise} \end{cases}$$

$$\Rightarrow g(\nu) = \begin{cases} \inf_y \left\{ \|y\|_2^2 + \nu^T y + \nu^T b \right\} & \text{if } \|\nu^T A\|_1 \leq 1 \\ -\infty & \text{otherwise} \end{cases}$$

$\rightarrow$ We can minimize over y by setting the gradient with respect to y equal to zero, that is:

$$2y + \nu = 0 \iff y = -\frac{1}{2}\nu$$

$$\Rightarrow g(\nu) = \begin{cases} \nu^T b + \|-\frac{1}{2}\nu\|_2^2 + \left(-\frac{1}{2}\right)\nu^T \nu & \text{if } \|\nu^T A\|_1 \leq 1 \\ -\infty & \text{otherwise.} \end{cases}$$

$$\Rightarrow g(\nu) = \begin{cases} \nu^T b + \frac{1}{4}\|\nu\|_2^2 - \frac{1}{2}\|\nu\|_2^2 & \text{if } \|\nu^T A\|_1 \leq 1 \\ -\infty & \text{otherwise.} \end{cases}$$

③

$$\Rightarrow g(\nu) = \begin{cases} \nu^T b - \frac{1}{4} \|\nu\|_2^2 & \text{if } \|\nu^T A\|_1 \leq 1 \\ -\infty & \text{otherwise} \end{cases}$$

Hence, the dual of (RLS) is:

(Dual)
$$\boxed{\begin{array}{ll} \underset{\nu}{\text{Max}} & \nu^T b - \frac{1}{4} \|\nu\|_2^2 \\ \text{s.t.} & \|\nu^T A\|_1 \leq 1 \end{array}}$$

4

(Sep.2) $\boxed{\begin{aligned} &\underset{w,z}{\text{Min}} \quad \frac{1}{n\mathcal{C}} \mathbb{1}^T z + \frac{1}{2}\|w\|_2^2 \\ &\text{s.t.} \quad z_i \geq 1 - y_i (w^T x_i) \quad \forall i = 1, \dots, n. \\ &\qquad\;\; z \geq 0 \end{aligned}}$

Ex 3 :

1. (Sep 2) is the ~~[illegible]~~ minimization problem of the sum of 2 terms, each of which depends only on 1 variable (the left term depends on z and the right term depends on w).
Since (Sep 2) is a minimization w.r.t. w and z, we can deduce that (Sep 2) is equivalent to:

$$\underbrace{\underset{w}{\text{Min}} \ \frac{1}{2}\|w\|_2^2}_{(L)} + \underbrace{\boxed{\begin{aligned} &\underset{z}{\text{Min}} \quad \frac{1}{n\mathcal{C}} \mathbb{1}^T z \\ &\text{s.t.} \quad z_i \geq 1 - y_i (w^T x_i) \quad \forall i = 1, \dots, n \\ &\qquad\;\; z \geq 0 \end{aligned}}}_{(R)}$$

The minimization with respect to z of (R) is immediate as the objective function is linear in z and each component of z has a lower bound, we have:

(R)

$$\boxed{\begin{aligned} &\text{Min} \quad \frac{1}{n\mathcal{C}} \mathbb{1}^T z \\ &\text{s.t.} \quad z_i \geq 1 - y_i (w^T x_i) \quad \forall i = 1, \dots, n \\ &\qquad\;\; z \geq 0 \end{aligned}} = \frac{1}{n\mathcal{C}} \sum_{i=1}^{n} \max\{0, 1 - y_i (w^T x_i)\}$$

Hence, we can deduce that (Sep 2) is equivalent to:

$$\underset{w}{\text{Min}} \quad \frac{1}{n\mathcal{C}} \sum_{i=1}^{n} \max\{0, 1 - y_i (w^T x_i)\} + \frac{1}{2}\|w\|_2^2$$

i.e. $$\underset{w}{\text{Min}} \quad \frac{1}{n\mathcal{C}} \sum_{i=1}^{n} \mathcal{L}(w, x_i, y_i) + \frac{1}{2}\|w\|_2^2 \quad \text{(Sep 2)}$$

Since $\mathcal{C} > 0$, it is patent that solving (Sep 2) is equivalent to solving the following problem:

$$\min_{\omega} \, \mathcal{C}\left(\frac{1}{n\mathcal{C}} \sum_{i=1}^{n} \mathcal{L}(\omega, x_i, y_i) + \frac{1}{2}\|\omega\|_2^2\right) = (\text{Sep } 1)$$

↳ this problem and (Sep 2) give the same optimal point!

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

→ Hence, we can conclude that (Sep 2) solves problem (Sep 1)

2. We will denote $\lambda = (\lambda_1, \dots, \lambda_n)^T \in \mathbb{R}^n$ in this part.
The lagrangian function of (Sep 2) is given by:

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

$$= \frac{1}{n\mathcal{C}} \mathbb{1}^T z + \frac{1}{2}\|\omega\|_2^2 + \mathbb{1}^T \lambda - \lambda^T z - \sum_{i=1}^{n} \lambda_i y_i (\omega^T x_i) - \pi^T z$$

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

The dual function is therefore given by:

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

We first minimize over z by setting the gradient of the lagrangian with respect to z equal to zero (since the lagrangian is affine in z). We obtain:

$$\frac{1}{n\mathcal{C}} \mathbb{1} - \lambda - \pi = 0$$

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

Now, let us compute $\inf_{w} \left\{ \frac{1}{2}\|w\|_2^2 - \sum_{i=1}^{m} \lambda_i y_i (w^T x_i) \right\} = B$

The function $h : \left[w \longmapsto \frac{1}{2}\|w\|_2^2 - \sum_{i=1}^{m} \lambda_i y_i (w^T x_i) \right]$ is convex and differentiable.
Therefore, we minimize h over w by setting the gradient with respect to w equal to zero. This yields:

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

$$\Longleftrightarrow \quad w = \sum_{i=1}^{m} \lambda_i y_i x_i .$$

$$\Rightarrow B = \frac{1}{2}\Big\|\sum_{i=1}^{m} \lambda_i y_i x_i\Big\|_2^2 - \sum_{i=1}^{m} \lambda_i y_i \left(\Big[\sum_{j=1}^{m} \lambda_j y_j x_j\Big]^T x_i \right)$$

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

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

$$= -\frac{1}{2}\sum_{j=1}^{m}\sum_{i=1}^{m} \lambda_i \lambda_j y_i y_j x_j^T x_i$$

Hence, the dual function is given by:

$$g(\lambda,\pi) = \begin{cases} \mathbb{1}^T\lambda - \frac{1}{2}\Big(\sum_{j=1}^{m}\sum_{i=1}^{m} \lambda_i \lambda_j y_i y_j x_j^T x_i\Big) & \text{if } \frac{1}{m\mathcal{C}}\mathbb{1} - \lambda - \pi = 0 \\ -\infty & \text{otherwise.} \end{cases}$$

The dual problem of (Sep 2) is therefore given by:

$$\text{(Dual)} \quad \begin{aligned} \max_{\lambda,\pi} \quad & \sum_{i=1}^{m} \lambda_i - \frac{1}{2}\Big(\sum_{j=1}^{m}\sum_{i=1}^{m} \lambda_i \lambda_j y_i y_j x_j^T x_i\Big) \\ \text{s.t.} \quad & \frac{1}{m\mathcal{C}}\mathbb{1} - \lambda - \pi = 0 \\ & \lambda \geq 0 \\ & \pi \geq 0 \end{aligned} \quad (*)$$

Notice that

$$(*) \quad \left.\begin{aligned} & \frac{1}{m\mathcal{C}}\mathbb{1} = \lambda + \pi \\ & \lambda_i \geq 0 \quad \forall i=1,\dots,m \\ & \pi \geq 0 \end{aligned}\right\} \Longleftrightarrow \left\{ \begin{aligned} & \forall i = 1,\dots,m \\ & 0 \leq \lambda_i \leq \frac{1}{m\mathcal{C}} \end{aligned}\right.$$

Therefore the dual of (Sep 2) can be simplified as

$$\text{(Dual)} \quad \boxed{\begin{array}{ll} \max\limits_{\lambda} & \sum\limits_{i=1}^{m} \lambda_i - \frac{1}{2}\left(\sum\limits_{j=1}^{m}\sum\limits_{i=1}^{m} \lambda_i \lambda_j y_i y_j x_j^T x_i\right) \\ \text{s.t.} & 0 \leq \lambda_i \leq \frac{1}{m\mathcal{C}} \quad \forall i = 1, \dots, m. \end{array}}$$

END

Ex 5 : (Boolean LP)

$$\begin{aligned} \min \quad & c^T x \\ \text{s.t.} \quad & Ax \leq b \\ & x_i(1-x_i) = 0 \quad \forall i = 1, \dots, n. \end{aligned}$$

The Lagrangian is:

$$\begin{aligned} L(x, \lambda, \nu_1, \dots, \nu_n) &= c^T x + \lambda^T (Ax - b) + \sum_{i=1}^{n} \nu_i (x_i^2 - x_i) \\ &= c^T x + \lambda^T (Ax - b) + \sum_{i=1}^{n} \nu_i x_i^2 - \sum_{i=1}^{n} \nu_i x_i . \end{aligned}$$

If we denote $\nu^T = (\nu_1, \dots, \nu_n)$ and $\operatorname{diag}(\nu) = \begin{pmatrix} \nu_1 & & 0 \\ & \nu_2 & \\ & & \ddots \\ 0 & & \nu_n \end{pmatrix}$
the diagonal matrix with values $\nu_1, \dots, \nu_n$ on its diagonal,
we can rewrite the Lagrangian as:

$$\begin{aligned} L(x, \lambda, \nu) &= c^T x + \lambda^T (Ax - b) + x^T \operatorname{diag}(\nu) x - \nu^T x \\ &= x^T \operatorname{diag}(\nu) x + c^T x + \lambda^T A x - \lambda^T b - \nu^T x \\ &= x^T \operatorname{diag}(\nu) x + (c^T + \lambda^T A - \nu^T) x - \lambda^T b . \end{aligned}$$

We can minimize over x by setting the gradient with respect to x equal to zero, that is:

$$\begin{aligned} \nabla_x L(x, \lambda, \nu) = 2 \operatorname{diag}(\nu) x + c + A^T \lambda - \nu &= 0 \\ \iff \quad 2 \operatorname{diag}(\nu) x &= \nu - c - A^T \lambda \\ \iff \quad \operatorname{diag}(\nu) x &= \frac{1}{2} (\nu - c - A^T \lambda) \end{aligned}$$

* $\operatorname{diag}(\nu)$ is diagonal, therefore it is invertible iff $\forall i, \ \nu_i > 0$

Assuming that $\nu \succ 0$, we obtain:

$$\hat{x} = \frac{1}{2} \operatorname{diag}(\nu)^{-1} (\nu - c - A^T \lambda)$$ minimizes the Lagrangian.

The dual function is therefore given by:

$$\begin{aligned} \underset{\nu \succ 0}{g(\lambda, \nu)} = & \left(\frac{1}{2} \operatorname{diag}(\nu)^{-1} (\nu - c - A^T \lambda) \right)^T \operatorname{diag}(\nu) \left(\frac{1}{2} \operatorname{diag}(\nu)^{-1} (\nu - c - A^T \lambda) \right) \\ & + (c^T + \lambda^T A - \nu^T) \cdot \left(\frac{1}{2} \operatorname{diag}(\nu)^{-1} (\nu - c - A^T \lambda) \right) - \lambda^T b \end{aligned}$$

Since $\operatorname{diag}(\nu)^{-1}$ is diagonal $\Rightarrow \left(\operatorname{diag}(\nu)^{-1} \right)^T = \operatorname{diag}(\nu)^{-1}$

$$\Rightarrow \underset{\nu \geq 0}{g(\lambda,\nu)} = \frac{1}{4}\left((\nu - c - A^T\lambda)^T \cancel{\mathrm{diag}}(\nu)^{-1} \cancel{\mathrm{diag}}(\nu)\, \mathrm{diag}(\nu)^{-1} (\nu - c - A^T\lambda)\right)$$
$$- \frac{1}{2}(\nu - c - A^T\lambda)^T \mathrm{diag}(\nu)^{-1} (\nu - c - A^T\lambda)$$
$$- \lambda^T b$$

$$\Rightarrow \underset{\nu \geq 0}{g(\lambda,\nu)} = \frac{1}{4}(\nu - c - A^T\lambda)^T \mathrm{diag}(\nu)^{-1}(\nu - c - A^T\lambda)$$
$$- \frac{1}{2}(\nu - c - A^T\lambda)^T \mathrm{diag}(\nu)^{-1}(\nu - c - A^T\lambda)$$
$$- \lambda^T b$$

$$\underset{\nu \geq 0}{g(\lambda,\nu)} = -b^T\lambda - \frac{1}{4}(c + A^T\lambda - \nu)^T \mathrm{diag}(\nu)^{-1}(c + A^T\lambda - \nu)$$

where $\mathrm{diag}(\nu)^{-1} = \begin{pmatrix} \frac{1}{\nu_1} & & 0 \\ & \ddots & \\ 0 & & \frac{1}{\nu_n} \end{pmatrix}$.

If we denote A_i the ith column of A, we obtain the dual function:

$$\boxed{g(\lambda,\nu) = \begin{cases} -b^T\lambda - \frac{1}{4}\sum_{i=1}^{n} \frac{(c_i + A_i^T\lambda - \nu_i)^2}{\nu_i} & \text{if } \nu \geq 0 \\ -\infty & \text{otherwise} \end{cases}}$$

Hence, the dual problem is:

$$\text{(Dual)} \quad \boxed{\begin{aligned} &\mathrm{Max} \quad -b^T\lambda - \frac{1}{4}\sum_{i=1}^{n} \frac{(c_i + A_i^T\lambda - \nu_i)^2}{\nu_i} \\ &\text{s.t.} \qquad \nu \geq 0, \quad \lambda \geq 0 \end{aligned}}$$

Now we want to simplify it to have only one dual variable: we optimize analytically over ν: thanks to hint we have:

$$\forall i = 1,\dots,n, \quad \sup_{\nu_i \geq 0}\left(-\frac{(c_i + A_i^T\lambda - \nu_i)^2}{\nu_i}\right) = \begin{cases} 4(c + A_i^T\lambda) & \text{if } c + A_i^T\lambda \leq 0 \\ 0 & \text{if } c + A_i^T\lambda \geq 0 \end{cases}$$

that is $\forall i = 1, \dots, n$:

$$\sup_{\nu_i \geq 0} \left(- \frac{(c_i + A_i^T\lambda - \nu_i)^2}{\nu_i} \right) = 4 \min\{0, (c_i + A_i^T\lambda)\}$$

This allows us to get rid of ν in the dual problem.
The dual problem is therefore given by:

(Dual)

$$\begin{array}{ll} \text{Max} & -b^T\lambda + \sum_{i=1}^{n} \min\{0, c_i + A_i^T\lambda\} \\ \text{s.t.} & \lambda \geq 0 \end{array}$$

2. (LP relaxation) is (2)

$$\begin{array}{ll} \text{Min} & c^Tx \\ \text{s.t.} & Ax \leq b \\ & 0 \leq x_i \leq 1 \quad \forall i = 1, \dots, n. \end{array}$$

$\iff$ (2)

$$\begin{array}{ll} \text{Min} & c^Tx \\ \text{s.t.} & Ax \leq b \\ & -x \leq 0 \\ & x - 1 \leq 0 \end{array}$$

where $1 = (1, 1, \dots, 1)^T$

The Lagrangian is:

$$\begin{aligned} L(x, \lambda_1, \lambda_2, \lambda_3) &= c^Tx + \lambda_1^T(Ax - b) - \lambda_2^T x + \lambda_3^T(x - 1) \\ &= c^Tx + \lambda_1^T Ax - \lambda_1^T b - \lambda_2^T x + \lambda_3^T x - \lambda_3^T 1 \\ &= (c^T + \lambda_1^T A - \lambda_2^T + \lambda_3^T)x - \lambda_1^T b - \lambda_3^T 1 \end{aligned}$$

The Lagrangian is an affine function of x
It follows that the dual function is given by:

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

Hence, the dual of the LP relaxation is:

(Dual LP Relaxation)

$$\begin{array}{ll} \text{Max} & -\lambda_1^T b - \lambda_3^T 1 \\ \text{s.t.} & A^T\lambda_1 - \lambda_2 + \lambda_3 + c = 0 \\ & \lambda_1 \geq 0, \lambda_2 \geq 0, \lambda_3 \geq 0 \end{array}$$

(Dual LP Relaxation)

$$\boxed{\begin{aligned} &\text{Max} \quad -\lambda_1^T b - \mathbb{1}^T \lambda_3 \\ &\text{s.t.} \quad \lambda_3 = -A^T\lambda_1 - c + \lambda_2 \\ &\qquad \lambda_1 \geq 0,\ \lambda_2 \geq 0,\ \lambda_3 \geq 0 . \end{aligned}}$$

$$\Leftrightarrow \quad \boxed{\begin{aligned} &\text{Max} \quad -\lambda_1^T b - \mathbb{1}^T \lambda_3 \\ &\text{s.t.} \quad \lambda_3 \geq -A^T\lambda_1 - c \\ &\qquad \lambda_1 \geq 0,\ \lambda_3 \geq 0 \end{aligned}}$$

$$\Leftrightarrow \quad \boxed{\begin{aligned} &\text{Max} \quad -\lambda_1^T b - \mathbb{1}^T \lambda_3 \\ &\text{s.t.} \quad \boxed{\lambda_3 \geq \max\{0, -A^T\lambda_1 - c\}} \\ &\qquad \lambda_1 \geq 0 \end{aligned}} \Leftrightarrow \lambda_3 \geq -\min\{0, A^T\lambda_1 + c\}$$

Dual LP Relaxation $\Leftrightarrow$

$$\begin{aligned} &\text{Max} \quad -b^T\lambda_1 + \sum_{i=1}^{m} \min\{0, A_i^T\lambda_1 + c_i\} \\ &\text{s.t.} \quad \lambda_1 \geq 0 . \end{aligned}$$

which is equivalent to the Lagrangian relaxation problem derived in question 1.
We conclude that the lower bound obtained via Lagrangian relaxation and via LP relaxation are the same.