Convex Optimization - Homework 2

David FAGET CAÑO

Exercise 1 (LP Duality) For 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,

$$\begin{aligned} &\min_x c^T x \\ \text{s.t.} \quad & Ax = b \\ & x \geq 0 \end{aligned} \tag{P}$$

and

$$\begin{aligned} &\max_y b^T y \\ \text{s.t.} \quad & A^T y \leq c \end{aligned} \tag{D}$$

1. Compute the dual of problem (P) and simplify it if possible.

Using the notation introduced in class, we can write inequality constraints functions as $f_i(x) = -x_i$, $i = 1, \dots, d$, and the first constraint functions can be written as $h_i(x) = [Ax - b]_i$, $i = 1, \dots, n$

the Lagrangian associated to the problem is:

$$\begin{aligned} L(x, \lambda, \nu) &= c^T x - \sum_{i=1}^{d} \lambda_i x_i + \sum_{i=1}^{n} \nu_i [Ax - b]_i \\ &= c^T x - \lambda^T x + \nu^T (Ax - b) \\ &= (c^T - \lambda^T + \nu^T A) x - b^T \nu \\ &= (c - \lambda + A^T \nu)^T x - b^T \nu \end{aligned}$$

Now, we have to find the Lagrange dual function g:

$$\begin{aligned} g(\lambda, \nu) &= \inf_{x \in \mathcal{D}} L(x, \lambda, \nu) \\ &= \inf_{x \in \mathcal{D}} \left[(c - \lambda + A^T \nu)^T x - b^T \nu \right] \\ &= -b^T \nu + \inf_{x \in \mathcal{D}} (c - \lambda + A^T \nu)^T x \end{aligned}$$

But $(c - \lambda + A^T \nu)^T x$ is bounded (as a function of x) iff $c - \lambda + A^T \nu = 0$. As a consequence,

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

Therefore, the Lagrange dual problem can be written as follows:

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

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

However, it can be more simplified. We remark that we can regroup the constraint $\lambda \geq 0$ and the condition $c - \lambda + A^T\nu = 0$ in order to obtain $c + A^T\nu \geq 0$. Finally, the Lagrange dual problem can be expressed as:

$$\begin{aligned} \max_{\nu} \quad & -\nu^T b \\ \text{s.t.} \quad & c + A^T\nu \geq 0 \end{aligned} \qquad \Longleftrightarrow \quad (D)$$

→ 2. Compute the dual of problem (D).

This is a maximization problem. The first step is turning it into a minimization problem:

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

We are now able to compute the Lagrangian:

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

Now, we have:

$$\begin{aligned} g(\lambda,\nu) &= -\lambda^T c + \inf_{y \in \mathcal{D}} (A\lambda - b)y \\ &= \begin{cases} -\lambda^T c & \text{if} \quad A\lambda - b = 0 \\ -\infty & \text{otherwise}. \end{cases} \end{aligned}$$

In the last step, we employed the same reasoning as for the problem (P). Hence, the dual problem is:

$$\begin{aligned} \max_{\lambda} \quad & -\lambda^T c \\ \text{s.t.} \quad & \lambda \geq 0 \\ & A\lambda = b \end{aligned} \qquad \Longleftrightarrow \quad (P)$$

3. A problem is called *self-dual* if its dual is the problem itself. Prove that the following problem is self-dual.

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

Let's prove that the given problem is self-dual.

We remark that it can be written as:

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

The next step is to compute the Lagrangian:

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

Then,

$$\begin{aligned} g((x,y),\lambda,\nu) &= \inf_{x,y \in D} \mathcal{L}((x,y),\lambda,\nu) \\ &= -\lambda^T c - \nu^T b + \inf_{x,y \in D} \left[(c - \lambda + A\nu^T)x + (-b + A\lambda)y\right] \\ &= \begin{cases} -\lambda^T c - \nu^T b & \text{if } (c - \lambda + A\nu^T) = (-b + A\lambda) = 0 \\ -\infty & \text{otherwise.} \end{cases} \end{aligned}$$

As a consequence, the dual problem can be written as:

$$\begin{aligned} &\max_{\lambda,\nu} \quad -\lambda^T c - \nu^T b \\ &\text{s.t.} \quad A\lambda - b = 0 \\ &\qquad\quad A\nu^T + c \geq 0 \\ &\qquad\quad \lambda \geq 0 \end{aligned}$$

Now, we can re-write it as a minimization problem and we can also change ν by $-\nu$ to obtain:

$$\min_{\lambda, \nu} \quad -\nu^T b - \lambda^T c$$

$$\text{s.t.} \quad A\lambda = b$$
$$\lambda \geq 0$$
$$A^T \nu \leq c$$

which is the original problem. As a conclusion, it is self-dual.

4. Assume the above problem feasible and bounded, and let $[x^*, y^*]$ be its optimal solution. Using the strong duality property of linear programs, show that

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

⊙ The (Self-Dual) problem can be decomposed into two different problems, since the constraints are independent. In a schematic form, we have:

$$(\text{Self-Dual}) \iff \begin{cases} \min_{x} & c^T x \\ \text{s.t.} & Ax = b \\ & x \geq 0 \end{cases} + \begin{cases} \min_{y} & -b^T y \\ \text{s.t.} & A^T y \leq c \end{cases}$$

$$\iff \begin{cases} \min_{x} & c^T x \\ \text{s.t.} & Ax = b \\ & x \geq 0 \end{cases} + \begin{cases} \max_{y} & b^T y \\ \text{s.t.} & A^T y \leq c \end{cases}$$

$$\iff (P) + (D)$$

As a consequence, we can say that x^* is an optimal solution for (P) and y^* is an optimal solution for (D).

⊙ As we assume the (Self-Dual) problem to be feasible and bounded, problem (P) is also feasible and bounded.

As seen in class, in order to say that a problem is strictly feasible, linear inequalities do not need to hold with strict inequality. As a consequence, the problem (P) is strictly feasible. Since it is a convex problem (the objective function and the constraints are linear), it satisfies Slater's constraint qualification. Hence, strong duality holds for problem (P).

As we have seen before, (D) is the dual problem of (P). Thus, we have that $p^* = d^*$ (using class notation).

But $\quad p^* = d^* \iff c^T x^* = b^T y^*$

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

In conclusion, the optimal value of (Self-Dual) is exactly 0.

Exercise 2 (Regularized Least-Square) For given $A \in \mathbb{R}^{n\times d}$ and $b \in \mathbb{R}^n$, consider the following optimization problem,

$$\min_x \|Ax - b\|_2^2 + \|x\|_1. \tag{RLS}$$

1. Compute the conjugate of $\|x\|_1$.

We have to compute $f^*(y) = \sup_{x\in\mathbb{R}^d} \left(y^T x - \|x\|_1 \right)$

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

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

We are going to examine the different possible cases :

- If it exist an $i \in \{1,\dots,d\}$ such as $y_i < -1$,

  we can take x such as

$$x = \begin{cases} x_i = t < 0 \\ x_j = 0 & \forall j \neq i \end{cases}$$

Under these circumstances,

$$\sum_{i=1}^d x_i y_i - \sum_{i=1}^d |x_i| = (t y_i - t) = t(y_i - 1) \xrightarrow[t\to-\infty]{} +\infty$$

- Similarly, if it $\exists\, i \in \{1,\dots,d\}$ / $y_i > 1$, by choosing

$$x = \begin{cases} x_i = t > 0 \\ x_j = 0 & \forall j \neq i \end{cases}$$

  we have

$$\sum_{i=1}^d x_i y_i - \sum_{i=1}^d |x_i| = t(y_i - 1) \xrightarrow[t\to+\infty]{} +\infty$$

- Last, if $|y_i| \leq 1$ $\forall i = 1,\dots,d$, we have :

$$-x_i \leq y_i x_i \leq x_i$$

$$\Leftrightarrow \quad y_i x_i \leq |x_i|$$

$$\Leftrightarrow \quad \sum_{i=1}^d y_i x_i \leq \sum_{i=1}^d |x_i|$$

$$\Leftrightarrow \quad y^T x - \|x\|_1 \leq 0$$

As a consequence, $f^*(y) = 0$ in this case

In light of the above, we have:

$$f^*(y) = \begin{cases} 0 & \text{if } |y_i| \leq 1 \quad \forall i = 1, \dots, d \\ +\infty & \text{otherwise.} \end{cases}$$

→ 2. Compute the dual of (RLS).

We first remark that

$$(RLS) \iff \begin{aligned} &\min_{x,z} \ \|z\|_2^2 + \|x\|_1 \\ &\text{s.t.} \quad z = Ax + b \end{aligned}$$

$$\iff \begin{aligned} &\min_{x,z} \ \|z\|_2^2 + \|x\|_1 \\ &\text{s.t.} \quad Ax - z - b = 0 \end{aligned}$$

Now, the Lagrangian is:

$$L((x,z), \lambda, v) = \|z\|_2^2 + \|x\|_1 + \sum_{i=1}^{n} v_i (Ax_i - z_i - b_i)$$
$$= \|z\|_2^2 + \|x\|_1 + v^T (Ax - z - b)$$

We compute the Lagrange dual function:

$$g(\lambda, v) = \inf_{(x,z) \in D} \left(\|z\|_2^2 + \|x\|_1 + v^T (Ax - z - b) \right)$$

$$= \inf_{(x,z) \in D} \left(z^T z + \|x\|_1 + v^T A x - v^T z - v^T b \right)$$

$$= \inf_x \left(\|x\|_1 + v^T A x \right) + \inf_z \left(z^T z - v^T z \right) - v^T b$$

Let's calculate both infimums.

⊙ $\inf_x \left(\|x\|_1 + v^T A x \right)$

Using the previous question, we have that

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

$$= \begin{cases} 0 & \text{if } |(A^T v)_i| \leq 1, \ \forall i = 1, \dots, n \\ -\infty & \text{otherwise.} \end{cases}$$

$\odot \inf_z (z^T z - v^T z)$

Let $h : z \mapsto z^T z - v^T z$ be a quadratic function.
We have that

$$\nabla h(z) = 0 \iff 2z^T - v^T = 0$$
$$\iff z = \frac{1}{2} v$$

Thus,

$$\inf_z (z^T z - v^T z) = \frac{1}{4} v^T v - \frac{1}{2} v^T v$$
$$= -\frac{1}{4} v^T v$$

As a consequence,

$$g((x,z), \lambda, v) = \begin{cases} -\frac{1}{4} v^T v - b^T v & \text{if } |(A^T v)_i| \leq 1 \quad \forall i = 1, \dots, n \\ -\infty & \text{otherwise.} \end{cases}$$

We can write the dual problem of (RLS) as follows:

$$\max_v \; -\frac{1}{4} v^T v - b^T v$$
$$\text{s.t.} \quad |(A^T v)_i| \leq 1 \quad \forall i = 1, \dots, n$$

Which is equivalent to:

$$\min_v \; \frac{1}{4} v^T v + b^T v$$
$$\text{s.t.} \quad |(A^T v)_i| \leq 1 \quad \forall i = 1, \dots, n$$

Exercise 3 (Data Separation) Assume we have n data points $x_i \in \mathbb{R}^d$, with label $y_i \in \{-1, 1\}$. We are searching for an hyper-plane defined by its normal ω, which separates the points according to their label. Ideally, we would like to have

$$\omega^T x_i \leq -1 \Rightarrow y_i = -1 \quad \text{and} \quad \omega^T x_i \geq 1 \Rightarrow y_i = 1.$$

Unfortunately, this condition is rarely met with real-life problems. Instead, we solve an optimization problem which minimizes the gap between the hyper-plane and the miss-classified points. To do so, we will use a specific *loss function*

$$\mathcal{L}(\omega, x_i, y_i) = \max\left\{0\ ;\ 1 - y_i(\omega^T x_i)\right\}, \tag{1}$$

which is equal to 0 when the point x_i is well-classified (the sign of $\omega^T x_i$ and y_i is the same), but is strictly positive when the sign of $\omega^T x_i$ and y_i is different. To improve the performances, instead of minimizing the loss function alone, we also use a quadratic regularizer as follow,

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

where τ is the regularization parameter.

1. Consider the following quadratic optimization problem ($\mathbf{1}$ is a vector full of ones),

$$\begin{aligned} &\min_{\omega, z} \frac{1}{n\tau}\mathbf{1}^T z + \frac{1}{2}\|\omega\|_2^2 \\ \text{s.t.}\quad & z_i \geq 1 - y_i(\omega^T x_i) \qquad \forall i = 1 \dots n \qquad (\lambda_i) \\ & z \geq 0 \qquad (\pi) \end{aligned} \tag{Sep. 2}$$

Explain why problem (Sep. 2) solves problem (Sep. 1).

The following problem is equivalent to (Sep. 1):

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

$$\text{s.t.} \quad z_i = \mathcal{L}(w, x_i, y_i), \quad \forall i = 1, \dots, n.$$

Note that we have divided by τ the objective function. We can rewrite this last problem as follows:

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

$$\text{s.t.} \quad z_i = \max\{0\ ;\ 1 - y_i(w^T x_i)\} \quad \forall i = 1, \dots, n$$

But this constraint implies $z \geq 0$ and $z_i \geq 1 - y_i(w^T x_i)$, $\forall i = 1, \dots, n$; which are the constraints of (Sep. 2). Now, note that in the objective function, the goal is in part to minimize $\sum_{i=1}^{n} z_i$. But the two inequalities present in (Sep. 2) give the same lower bounds than $\mathcal{L}(w, x_i, y_i)$, $\forall i = 1, \dots, n$. From this, we can conclude that (Sep. 2) solves (Sep. 1).

2. Compute the dual of (Sep. 2), and try to reduce the number of variables. Use the notations λ_i and π for the dual variables.

As usual, the first step is to compute the Lagrangian. We have that:

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

The Lagrange dual function is:

$$g(\lambda, \pi) = \inf_{w,z \in D} \left(\frac{1}{n\tau} \mathbb{1}^T z + \frac{1}{2} \|w\|_2^2 + \sum_{i=1}^{n} \lambda_i (-z_i + 1 - y_i(w^T x_i)) - \pi^T z \right)$$

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

$$+ \inf_{z} \left(\frac{1}{n\tau} \mathbb{1}^T z - \sum_{i=1}^{n} \lambda_i z_i - \pi^T z \right)$$

$$+ \sum_{i=1}^{n} \lambda_i$$

As we did in Exercise 2 - 2), let's examine both infimums separately.

$$\odot \quad \inf_{w} \left(\frac{1}{2} \|w\|_2^2 - \sum_{i=1}^{n} \lambda_i y_i (w^T x_i) \right)$$

Let h_w be defined as $h_w(w) = \frac{1}{2} \|w\|_2^2 - \sum_{i=1}^{n} \lambda_i y_i (w^T x_i)$

We have that:

$$\nabla h_w(w) = 0$$

$$\Leftrightarrow \nabla \left(\frac{1}{2} w^T w - \sum_{i=1}^{n} \lambda_i y_i (w^T x_i) \right) = 0$$

$$\iff \quad w = \sum_{i=1}^{n} \lambda_i y_i x_i$$

Hence,

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

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

$$= -\frac{1}{2} \sum_{i=1}^{n} \lambda_i^2 y_i^2 \|x_i\|_2^2 \qquad \text{(Do not forget that } x_i \in \mathbb{R}^d\text{).}$$

$$\odot \quad \inf_{z} \left(\frac{1}{n\pi} \mathbb{1}^T z - \sum_{i=1}^{n} \lambda_i z_i - \pi^T z \right)$$

Now, $h_z(z) = \frac{1}{n\pi} \mathbb{1}^T z - \sum_{i=1}^{n} \lambda_i z_i - \pi^T z$

$$= \frac{1}{n\pi} \mathbb{1}^T z - \lambda^T z - \pi^T z$$

$$= \left(\frac{1}{n\pi} \mathbb{1} - \lambda - \pi \right)^T z \quad .$$

We remark that h_z is in fact a linear function.

Hence,

$$\inf_{z} h_z(z) = \begin{cases} 0 , & \text{if } \left(\frac{1}{n\pi}\mathbb{1} - \lambda - \pi \right) = 0 \\ -\infty & \text{otherwise} . \end{cases}$$

In light of the above,

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

We obtain the following dual problem:

$$\max_{\lambda, \pi} \; -\frac{1}{2} \sum_{i=1}^{n} \lambda_i^2 y_i^2 \|x_i\|_2^2 + \mathbb{1}^T \lambda$$

$$\text{s.t.} \quad \frac{1}{n\pi}\mathbb{1} - \lambda - \pi = 0$$

$$\lambda \geqslant 0$$

Since the objective function does not depend on π, we can simplify the dual problem and rewrite it as follows:

$$\max_{\lambda} \quad -\frac{1}{2} \sum_{i=1}^{n} \lambda_i^2 \, y_i^2 \, \|x_i\|_2^2 \; + \; \mathbb{1}^T \lambda$$

$$\text{s.t.} \quad \frac{1}{nT} \mathbb{1} \geqslant \lambda$$

$$\lambda \geqslant 0$$

Optional Exercise 4 (Robust linear programming) Sometimes, it is possible to encounter problems with some uncertainty in the constraints. One way to deal with them is to solve their worst-case scenario, and this can be achieved by using robust programming. Consider the following robust LP

$$\begin{array}{rl} & \min_x c^T x \\ \text{s.t.} & \sup_{a \in \mathcal{P}} a^T x \leq b, \end{array}$$

with variable $x \in \mathbb{R}^n$, where $\mathcal{P} = \{a \mid C^T a \preceq d\}$ is a nonempty polyhedra. The supremum represents the worst-case scenario for the constraint. Show that this problem is equivalent to the following LP.

$$\begin{array}{rl} & \min c^T x \\ \text{s.t.} & d^T z \leq b \\ & C^T z = x \\ & z \geq 0 \end{array}$$

Hint. Find the dual of the problem of maximizing $a^T x$ over $a \in \mathcal{P}$ (with variable a).

First of all, we are going to use the hint and find the dual problem of :

$$\begin{array}{rl} \max_a & a^T x \\ \text{s.t.} & a \in P \end{array}$$

This problem is equivalent to the following :

$$\begin{array}{rl} \min_a & -a^T x \\ \text{s.t.} & C^T a \leq d \end{array}$$

We remark that the dual of this problem has already been established in Exercise 1 ((P) is the dual of (D)). Hence, the dual problem is :

$$\begin{array}{rl} \min_z & d^T z \\ \text{s.t.} & C^T z = x \\ & z \geqslant 0 \end{array} \qquad (*)$$

We now return to our original problem. The robust LP can be expressed as :

$$\min_x \quad c^T x$$
$$\text{s.t.} \quad f(x) \leq b$$

where $f(x) = \sup_{a \in P} a^T x \leq b$. We notice that $f(x)$ is the optimal value of:

$$\max_x \quad x^T a$$
$$\text{s.t.} \quad a \in P$$

which is exactly the problem whose dual is $(*)$. Hence, $f(x)$ is also the optimal value of $(*)$ (strong duality holds here as seen in Exercise 1), so we have $f(x) \leq b \iff \exists z \,/ \begin{cases} d^T z \leq b \\ Cz = x \\ z \geq 0 \end{cases}$

which are exactly the constraints of the second LP given in the exercise.

Optional Exercise 5 (Boolean LP) A *Boolean LP* is an optimization problem of the form

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax \preceq b \\ & x_i \in \{0,1\}, \quad i = 1, \ldots, n, \end{array}$$

and is, in general, very difficult to solve. Consider the LP relaxation of this problem,

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax \preceq b \\ & 0 \leq x_i \leq 1, \quad i = 1, \ldots, n, \end{array} \tag{2}$$

which is far easier to solve, and gives a lower bound on the optimal value of the Boolean LP. In this problem we derive another lower bound for the Boolean LP, and work out the relation between the two lower bounds.

1. *Lagrangian relaxation.* The Boolean LP can be reformulated as the problem

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax \preceq b \\ & x_i(1 - x_i) = 0, \quad i = 1, \ldots, n, \end{array}$$

which has quadratic equality constraints. Find the Lagrange dual of this problem and simplify it to have only one dual variable. *Hint.* You can use that

$$\begin{aligned}\sup_{y\geq 0}\left(-\frac{(b+a^Tx-y)^2}{y}\right) &= \begin{cases} 4(b+a^Tx) & b+a^Tx\leq 0\\ 0 & b+a^Tx\geq 0\end{cases}\\ &= 4\min\{0,(b+a^Tx)\}.\end{aligned}$$

The optimal value of the dual problem (which is convex) gives a lower bound on the optimal value of the Boolean LP. This method of finding a lower bound on the optimal value is called *Lagrangian relaxation*.

As usual, we first compute the Lagrangian:

$$\begin{aligned}L(x,\lambda,\nu) &= c^Tx + \lambda^T(Ax-b) + \sum_{i=1}^n \nu_i\, x_i(1-x_i)\\ &= c^Tx + \lambda^T(Ax-b) + \nu^Tx - \sum_{i=1}^n \nu_i x_i^2\\ &= (c + A^T\lambda + \nu)^Tx - \lambda^Tb - \sum_{i=1}^n \nu_i x_i^2\end{aligned}$$

Hence,

$$\begin{aligned}g(\lambda,\nu) &= \inf_{x} L(x,\lambda,\nu)\\ &= -\lambda^Tb + \inf_x\left((c+A^T\lambda+\nu)^Tx - \sum_{i=1}^n \nu_i x_i^2\right)\end{aligned}$$

We notice that if $\nu>0$, then $\inf_x\left((c+A^T\lambda+\nu)^Tx - \sum_{i=1}^n \nu_i x_i^2\right) = -\infty$.

On the other hand, if $\nu\leq 0$, we have:

$$\inf_x\left((c+A^T\lambda+\nu)^Tx - \sum_{i=1}^n \nu_i x_i^2\right) = \frac{1}{4}\sum_{i=1}^n \frac{(c_i + a_i^T\lambda + \nu_i)^2}{\nu_i}$$

where a_i denotes the i th column of A.

As a consequence,

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

Now, we remark that by changing y for $-y$ in the given hint, we obtain that:

$$\sup_{y\leq 0}\frac{(b+a^Tx+y)^2}{y} = 4\min\{0,(b+a^Tx)\}$$

Hence,

$$\sup_{\nu\leq 0}\frac{(c_i + a_i^T\lambda + \nu_i)^2}{\nu_i} = 4\min\{0,(c_i + a_i^T\lambda)\}$$

Since the goal of the dual problem is to maximize the objective function, this last identity allows us to eliminate ν from the dual problem, and to write it as follows:

$$\begin{aligned} \max \quad & -\lambda^T b + \sum_{i=1}^{n} \min\{0, c_i + a_i^T \lambda\} \\ \text{s.t.} \quad & \lambda \geq 0 \end{aligned}$$

→ 2. Show that the lower bound obtained via Lagrangian relaxation, and via the LP relaxation (2), are the same. *Hint.* Derive the dual of the LP relaxation (2) and simplify it.

As told in the hint, we are going to derive the dual of the LP relaxation (2) and to simplify it.
This problem can be written as:

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

Its Lagrangian is:

$$\begin{aligned} L(x, w, y, z) &= c^T x + w^T (Ax - b) - y^T x + z^T (x - \mathbb{1}) \\ &= (c + A^T w - y + z)^T x - w^T b - z^T \mathbb{1}. \end{aligned}$$

where $\mathbb{1} = (1, \dots, 1)^T$.
Hence, the dual function is:

$$g(w, y, z) = \inf_x L(x, w, y, z)$$

Since $L(x, w, y, z)$ is a linear function in x, we have that:

$$g(w, y, z) = \begin{cases} -w^T b - z^T \mathbb{1} & \text{if } c + A^T w - y + z = 0 \\ -\infty & \text{otherwise.} \end{cases}$$

As a consequence, the dual problem is:

$$\begin{aligned} \max. \quad & -w^T b - z^T \mathbb{1} \\ \text{s.t.} \quad & c + A^T w - y + z = 0 \\ & w \geq 0, \; y \geq 0, \; z \geq 0 \end{aligned}$$

We notice that from the constraints we have :

$$\begin{cases} z = y - c - A^T w \\ z \geq 0 \end{cases}$$

Since $z^T \mathbb{1} = \sum_{i=1}^{n} z_i$, we can write the dual problem as :

$$\begin{aligned} \max. \quad & -w^T b + \sum_{i=1}^{n} \min \{ 0, c_i + a_i^T w - y \} \\ \text{s.t.} \quad & w \geq 0, \ y \geq 0 \end{aligned}$$

We note that we have to maximize the objective function and that $y \geq 0$ is multiplied by (-1). Hence, we must have $y = 0$. Then the problem is equivalent to

$$\begin{aligned} \max \quad & -w^T b + \sum_{i=1}^{n} \min \{ 0, c_i + a_i^T w \} \\ \text{s.t.} \quad & w \geq 0 \end{aligned}$$

which is the same dual problem as in part 1. As a consequence, we conclude that both relaxations give the same lower bound.