Convex Optimization and applications in Machine Learning **MVA Fall 2019**

Homework 2 — November 4

Instructor: Alexandre d'Aspremont *Student: Antoine Moulin*

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,

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

and

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

1. Compute the dual of problem (P) and simplify it if possible.
2. Compute the dual of problem (D).
3. A problem is called ***self-dual*** if its dual is the problem itself. Prove that the following problem is self-dual.

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

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.

Solution.

1. Let $x \in \mathbb{R}^d, \lambda \in \mathbb{R}^d, \nu \in \mathbb{R}^n$. The Lagrangian function of (P) is given by

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

As the Lagrangian is linear in x, the Lagrange dual function is

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

Hence, the dual of problem (P) is

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

which can be simplified in

$$\boxed{\begin{aligned} &\max_{\nu} -b^T\nu \\ \text{s.t. } & A^T\nu + c \succeq 0 \end{aligned}}$$

2. The problem (D) is equivalent to:

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

Let $y \in \mathbb{R}^n, \lambda \in \mathbb{R}^d$. Its Lagrangian function is given by

$$\mathcal{L}(y, \lambda) = -b^T y + \lambda^T \left(A^T y - c\right) = (A\lambda - b)^T y - c^T \lambda$$

As the Lagrangian is linear in y, the Lagrange dual function is

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

Hence, the dual of problem (D) is

$$\boxed{\begin{aligned} &\max_{\lambda} -c^T\lambda \\ \text{s.t. } & A\lambda = b \\ & \lambda \succeq 0 \end{aligned}}$$

3. Let $x \in \mathbb{R}^d, y \in \mathbb{R}^n, \lambda_1 \in \mathbb{R}^d, \lambda_2 \in \mathbb{R}^d, \nu \in \mathbb{R}^n$. The Lagrangian function of (Self-Dual) is given by

$$\begin{aligned} \mathcal{L}(x, y, \lambda_1, \lambda_2, \nu) &= c^T x - b^T y - \lambda_1^T x + \lambda_2^T \left(A^T y - c\right) + \nu^T (b - Ax) \\ &= \left(c - \lambda_1 - A^T\nu\right)^T x + (A\lambda_2 - b)^T y - c^T\lambda_2 + b^T\nu \end{aligned}$$

As the Lagrangian is linear in x and y, the Lagrange dual function is

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

Hence, the dual of problem (Self-Dual) is

$$\begin{aligned} &\max_{\lambda_1, \lambda_2, \nu} -c^T\lambda_2 + b^T\nu \\ \text{s.t. } & A\lambda_2 - b = 0 \\ & c - \lambda_1 - A^T\nu = 0 \\ & \lambda_1 \succeq 0 \\ & \lambda_2 \succeq 0 \end{aligned}$$

which is equivalent to

$$\begin{aligned} &\max_{\lambda_2,\nu} -c^T\lambda_2 + b^T\nu \\ \text{s.t. } & A\lambda_2 = b \\ & \lambda_2 \succeq 0 \\ & A^T\nu \preceq c \end{aligned}$$

Changing the notations, $\lambda_2 \leftrightarrow x$ and $\nu \leftrightarrow y$, and changing the max in min, we obtain:

$$\boxed{\begin{aligned} &\min_{x,y} c^Tx - b^Ty \\ \text{s.t. } & Ax = b \\ & x \succeq 0 \\ & A^Ty \preceq c \end{aligned}}$$

which is exactly the problem (Self-Dual).

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

$$\min_{x,y} c^Tx - b^Ty = \min_x c^Tx + \min_y -b^Ty = \min_x c^Tx - \max_y b^Ty$$

under the associated constraints. We recognize that it is the problem (P) minus the problem (D).

$$\boxed{\text{By solving (P) and (D), which give us the optimal points } x^* \text{ and } y^* \text{ respectively, we also solve the problem (Self-Dual) and } [x^*, y^*] \text{ is an optimal point.}}$$

• If we write precisely the above equality, we have:

$$\begin{aligned} &\min_{x,y} c^Tx - b^Ty \\ \text{s.t. } & Ax = b \\ & x \succeq 0 \\ & A^Ty \preceq c \end{aligned} \quad = \quad \begin{aligned} &\min_{x} c^Tx \\ \text{s.t. } & Ax = b \\ & x \succeq 0 \end{aligned} \quad - \quad \begin{aligned} &\max_{y} b^Ty \\ \text{s.t. } & A^Ty \preceq c \end{aligned}$$

On the left, we have the (Self-Dual) problem. The first term on the right is the problem (P) and the second term is the problem (D). If we make a change of variable in (D) ($y \leftrightarrow -\nu$), we see that it is the dual problem of (P) (see question 1).

The constraints in (P) are a linear equality and a linear inequality. Besides, the domain of the objective function (which is $\mathbb{R}^d$) is open. Hence, as the problem is supposed feasible, we know by the Slater's condition that strong duality holds. If we denote by p^* the optimal value of (P) and by d^* the optimal value of (D), we have:

$$\boxed{\begin{aligned} &\min_{x,y} c^Tx - b^Ty \quad = \quad p^* - d^* = 0 \\ \text{s.t. } & Ax = b \\ & x \succeq 0 \\ & A^Ty \preceq c \end{aligned}}$$

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 $||x||_1$.
2. Compute the dual of (RLS).

Solution.

1. Let $x, y \in \mathbb{R}^d$, $f = ||\cdot||_1$. We have

$$y^T x - f(x) = \sum_{i=1}^d y_i x_i - \sum_{i=1}^d |x_i|$$

Suppose there exists a $j \in [\![1, d]\!]$ such that $y_j > 1$. We choose x such that $x_j = t > 0$ and $x_i = 0$ for $i \neq j$. We have $y^T x - f(x) = (y_j - 1)\, t \underset{t\to+\infty}{\longrightarrow} +\infty$, hence $f^*(y) = +\infty$.

Similarly, if $y_j < -1$, we take $t < 0$ and $y^T x - f(x) = (y_j + 1)\, t \underset{t\to-\infty}{\longrightarrow} +\infty$, *i.e.* $f^*(y) = +\infty$.

Let suppose $||y||_\infty \leq 1$. We have

$$\begin{aligned} y^T x - f(x) &\leq \sum_{i=1}^d |y_i x_i| - \sum_{i=1}^d |x_i| \\ &\leq \sum_{i=1}^d \underbrace{(|y_i| - 1)}_{\leq 0} |x_i| \\ &\leq 0 \end{aligned}$$

with equality if $x = 0$. Hence, $f^*(y) = 0$. Finally,

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

2. The problem (RLS) is equivalent to

$$\begin{aligned} &\min_x ||y||_2^2 + ||x||_1 \\ &\text{s.t. } y = Ax - b \end{aligned} \tag{1}$$

Let $x \in \mathbb{R}^d, y \in \mathbb{R}^n, \nu \in \mathbb{R}^n$. The Lagrangian function of the new problem is given by

$$\begin{aligned} \mathcal{L}(x, y, \nu) &= ||y||_2^2 + ||x||_1 + \nu^T (y - Ax + b) \\ &= ||y||_2^2 + \nu^T y + ||x||_1 - \left(A^T \nu\right)^T x + \nu^T b \end{aligned}$$

and the dual function is

$$\begin{aligned} g(\nu) &= \inf_{x,y} \mathcal{L}(x, y, \nu) \\ &= b^T \nu + \inf_y \left(||y||_2^2 + \nu^T y\right) + \inf_x \left(||x||_1 - \left(A^T\nu\right)^T x\right) \end{aligned}$$

The function $h : y \mapsto ||y||_2^2 + \nu^T y$ is convex and differentiable. Its gradient is given by $\nabla h(y) = 2y + \nu$ and we have $\nabla h(y) = 0$ iff $y = -\frac{1}{2}\nu$. Hence, the minimum of h is $\frac{1}{4}||\nu||_2^2 - \frac{1}{2}||\nu||_2^2 = -\frac{1}{4}||\nu||_2^2$.

We can express the last term using the conjugate of $||\cdot||_1$:

$$\inf_x ||x||_1 - \left(A^T\nu\right)^T x = \sup_x \left(A^T\nu\right)^T x - ||x||_1 = ||\cdot||_1^* \left(A^T\nu\right)$$

Thus,

$$g(\nu) = b^T\nu - \frac{1}{4}||\nu||_2^2 + ||\cdot||_1^*\left(A^T\nu\right)$$

$||\cdot||_1^*$ is the indicator function of the $||\cdot||_\infty$-unit ball. Hence, the dual problem of (RLS) is

$$\boxed{\begin{aligned} &\max_\nu b^T\nu - \frac{1}{4}||\nu||_2^2 \\ \text{s.t. } &\left|\left|A^T\nu\right|\right|_\infty \leq 1 \end{aligned}}$$

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

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

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

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

which is equal to 0 when the point x_i is well-classified (the sign of $\omega^T x_i$ and y_i are the same), but is strictly positive when the sign of $w^T x_i$ and y_i are 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),

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

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

Solution.

1. If we first minimize with respect to ω and then with respect to z, (Sep. 2) is equivalent to

$$\min_\omega \frac{1}{2}||\omega||_2^2 + \quad \begin{aligned} &\min_z \frac{1}{n\tau}\mathbf{1}^T z \\ \text{s.t. } &\forall i \in [\![1, n]\!], z_i \geq 1 - y_i\left(\omega^T x_i\right) \\ &z \succeq 0 \end{aligned}$$

The minimization problem with respect to z is immediate as the objective function is linear in z and we have a lower bound on each component of z. Hence, we have

$$\begin{aligned} \min_{z} \ & \frac{1}{n\tau}\mathbf{1}^T z & = \frac{1}{n\tau}\sum_{i=1}^{n}\max\left\{0, 1 - y_i\left(\omega^T x_i\right)\right\} \\ \text{s.t. } & \forall i \in [\![1, n]\!], z_i \geq 1 - y_i\left(\omega^T x_i\right) & \\ & z \succeq 0 & \end{aligned}$$

and the problem is equivalent to

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

Multiplying by $\tau > 0$ and using the definition of the loss function, we see that

$$\boxed{\text{(Sep. 2) is equivalent to (Sep. 1)}}$$

2. Let $z \in \mathbb{R}^n, \omega \in \mathbb{R}^d, \lambda \in \mathbb{R}^n, \pi \in \mathbb{R}^n$. The Lagrangian function of (Sep. 2) is given by

$$\begin{aligned} \mathcal{L}\left(z, \omega, \lambda, \pi\right) &= \frac{1}{n\tau}\mathbf{1}^T z + \frac{1}{2}\left|\left|\omega\right|\right|_2^2 + \sum_{i=1}^{n}\lambda_i\left[1 - y_i\left(\omega^T x_i\right) - z_i\right] - \pi^T z \\ &= \left(\frac{1}{n\tau}\mathbf{1} - \lambda - \pi\right)^T z + \frac{1}{2}\left|\left|\omega\right|\right|_2^2 - \sum_{i=1}^{n}\lambda_i y_i\left(\omega^T x_i\right) + \mathbf{1}^T\lambda \end{aligned}$$

Let's compute the dual function.

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

Let $h : \omega \mapsto \frac{1}{2}\left|\left|\omega\right|\right|_2^2 - \sum_{i=1}^{n}\lambda_i y_i\left(\omega^T x_i\right)$. h is differentiable and convex. Its gradient is given by $\nabla h(\omega) = \omega - \sum_{i=1}^{n}\lambda_i y_i x_i$. We have $\nabla h(\omega) = 0$ iff $\omega_{min} = \sum_{i=1}^{n}\lambda_i y_i x_i$, and at this point:

$$\begin{aligned} h\left(\omega_{min}\right) &= \frac{1}{2}\left(\sum_{i=1}^{n}\lambda_i y_i x_i\right)^T\left(\sum_{i=1}^{n}\lambda_i y_i x_i\right) - \sum_{i=1}^{n}\lambda_i y_i\left(\sum_{j=1}^{n}\lambda_j y_j x_j\right)^T x_i \\ &= \frac{1}{2}\sum_{1\leq i,j\leq n}\lambda_i\lambda_j y_i y_j x_i^T x_j - \sum_{1\leq i,j\leq n}\lambda_i\lambda_j y_i y_j x_i^T x_j \\ &= -\frac{1}{2}\sum_{1\leq i,j\leq n}\lambda_i\lambda_j y_i y_j x_i^T x_j \end{aligned}$$

Hence,

$$g\left(\lambda, \pi\right) = \begin{cases} \sum_{i=1}^{n}\lambda_i - \frac{1}{2}\sum_{1\leq i,j\leq n}\lambda_i\lambda_j y_i y_j x_i^T x_j & \text{if } \frac{1}{n\tau}\mathbf{1} - \lambda - \pi = 0 \\ -\infty & \text{otherwise} \end{cases}$$

Finally, the dual of (Sep. 2) is

$$\begin{aligned}
&\max_{\lambda,\pi} \sum_{i=1}^{n} \lambda_i - \frac{1}{2} \sum_{1\le i,j\le n} \lambda_i\lambda_j y_i y_j x_i^T x_j \\
\text{s.t. } & \frac{1}{n\tau}\mathbf{1} - \lambda - \pi = 0 \\
& \lambda \succeq 0 \\
& \pi \succeq 0
\end{aligned}$$

which can be simplified as

$$\boxed{\begin{aligned}
&\max_{\lambda} \sum_{i=1}^{n} \lambda_i - \frac{1}{2} \sum_{1\le i,j\le n} \lambda_i\lambda_j y_i y_j x_i^T x_j \\
\text{s.t. } & 0 \preceq \lambda \preceq \frac{1}{n\tau}\mathbf{1}
\end{aligned}}$$

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

$$\begin{aligned}
&\min_{x} c^T x \\
\text{s.t. } & \sup_{a\in\mathcal{P}} a^T x \le b
\end{aligned} \tag{R1}$$

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

$$\begin{aligned}
&\min_{x} c^T x \\
\text{s.t. } & d^T z \le b \\
& C^T z = x \\
& z \succeq 0
\end{aligned} \tag{R2}$$

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

Solution.

Let denote (R1) the initial problem. Let $x \in \mathbb{R}^n$. The quantity $\sup_{a\in\mathcal{P}} a^T x$ in the constraint can be considered as the optimal value of a maximization problem (P_x) of the variable a defined by:

$$\begin{aligned}
&\max_{a\in\mathbb{R}^n} a^T x \\
\text{s.t. } & C^T a \preceq d
\end{aligned} \tag{P_x}$$

Note that the supremum is reached because the points x that are considered are such that the convex problem (P_x) is bounded. This problem is equivalent to

$$\begin{aligned}
&\min_{a\in\mathbb{R}^n} (-x)^T a \\
\text{s.t. } & C^T a \preceq d
\end{aligned}$$

which is an LP problem. For $a \in \mathbb{R}^n, z \in \mathbb{R}^m$, the Lagrangian is given by

$$\begin{aligned}
\mathcal{L}(a, z) &= (-x)^T a + z^T \left(C^T a - d\right) \\
&= (Cz - x)^T a - d^T z
\end{aligned}$$

and the dual problem is:

$$\begin{aligned} &\min_z d^T z \\ \text{s.t. } &Cz = x \\ &z \succeq 0 \end{aligned} \tag{D_x}$$

By strong duality, the optimal value of (D_x) is also reached and is equal to the optimal value of (P_x). Let denote by Δ_x the set of the feasible points of (D_x). The initial problem (R1) is equivalent to:

$$\begin{aligned} &\min_x c^T x \\ \text{s.t. } &\inf_{z \in \Delta_x} d^T z \leq b \end{aligned} \tag{T}$$

Let F be the set of feasible points of this new problem (T). We have:

$$\begin{aligned} F &= \left\{ x \in \mathbb{R}^n \mid \inf_{z \in \Delta_x} d^T z \leq b \right\} \\ &= \left\{ x \in \mathbb{R}^n \mid \exists z \in \Delta_x, \left(d^T z = \inf_{t \in \Delta_x} d^T t \right) \wedge \left(d^T z \leq b \right) \right\} \\ &= \left\{ x \in \mathbb{R}^n \mid \exists z \in \mathbb{R}^m, \left(d^T z = \inf_{t \in \Delta_x} d^T t \right) \wedge \left(d^T z \leq b \right) \wedge (Cz = x) \wedge (z \succeq 0) \right\} \end{aligned}$$

Let $x \in F$ and $z \in \mathbb{R}^m$ that satisfies the conditions. We have:

$$\begin{aligned} d^T z = \inf_{t \in \Delta_x} d^T t &\Leftrightarrow d^T z = \sup_{a \in \mathcal{P}} a^T x && \text{(strong duality)} \\ &\Leftrightarrow d^T z = \sup_{a \in \mathcal{P}} \left(C^T a \right) z && \text{(cf } Cz = x\text{)} \\ &\Leftrightarrow d^T z = d^T z && \text{(cf } a \in \mathcal{P} \text{ and } z \succeq 0\text{)} \end{aligned}$$

The last assertion is always true. Hence, we can write:

$$F = \left\{ x \in \mathbb{R}^n \mid \exists z \in \mathbb{R}^m, \left(d^T z \leq b \right) \wedge (Cz = x) \wedge (z \succeq 0) \right\}$$

We have just shown that (T) is equivalent to:

$$\begin{aligned} &\min_x c^T x \\ \text{s.t. } &d^T z \leq b \\ &C^T z = x \\ &z \succeq 0 \end{aligned}$$

which is exactly the problem (R2). Finally,

$$\boxed{\text{(R1) is equivalent to (R2)}}$$

Optional Exercise 5 (Boolean LP)

A ***Boolean LP*** is an optimization problem of the form

$$\begin{aligned} &\min_x c^T x \\ \text{s.t. } &Ax \leq b \\ &\forall i \in [\![1, n]\!], x_i \in \{0, 1\} \end{aligned}$$

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

$$\begin{aligned} &\min_x c^T x \\ \text{s.t. } &Ax \leq b \\ &\forall i \in [\![1, n]\!], 0 \leq x_i \leq 1 \end{aligned}$$

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

$$\begin{aligned} &\min_x c^T x \\ \text{s.t. } &Ax \leq b \\ &\forall i \in [\![1, n]\!], x_i (1 - x_i) = 0 \end{aligned}$$

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

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

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***.

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

Solution.