Convex Optimization
Homework 2

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.
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.

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

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.

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).

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*

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

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),

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

   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.

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

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

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.

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

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

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

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

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

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

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

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

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

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

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 (2), are the same. *Hint.* Derive the dual of the LP relaxation (2) and simplify it.