This homework is due at 11 PM on October 28, 2022.

Submission Format: Your homework submission should consist of a single PDF file that contains all of your answers (any handwritten answers should be scanned), as well as a printout of your completed Jupyter notebook(s).

1. Lagrangian Dual of a QP

   Consider the standard form of a convex quadratic program, with \( Q \succ 0 \):

   \[
   \begin{aligned}
   \min_{\bar{x}} & \quad \frac{1}{2}\bar{x}^\top Q\bar{x} \\
   \text{s.t.} & \quad A\bar{x} \leq \bar{b}
   \end{aligned}
   \]  
   \hspace{1in} (1)
   \hspace{1in} (2)

   (a) Write the Lagrangian function \( \mathcal{L}(\bar{x}, \bar{\lambda}) \).

   (b) Write the Lagrangian dual function, \( g(\bar{\lambda}) \).

   (c) Show that the Lagrangian dual problem is convex by writing it in “standard QP form” – that is, in a form similar to the original problem. Is the Lagrangian dual problem convex in general?
2. Optimizing Over Multiple Variables

In this exercise, we consider several problems in which we optimize over two variables, \( \vec{x} \in \mathbb{R}^n \) and \( \vec{y} \in \mathbb{R}^m \), and a general (possibly nonconvex) objective function, \( F_0(\vec{x}, \vec{y}) \). Suppose also that \( \vec{x} \) and \( \vec{y} \) are constrained to different feasible sets \( X \) and \( Y \), respectively, which may or may not be convex.

(a) Show that

\[
\min_{\vec{x} \in X} \min_{\vec{y} \in Y} F_0(\vec{x}, \vec{y}) = \min_{\vec{y} \in Y} \min_{\vec{x} \in X} F_0(\vec{x}, \vec{y}),
\]

i.e., if we minimize over both \( \vec{x} \) and \( \vec{y} \), then we can exchange the minimization order without altering the optimal value.

(b) Show that \( p^* \geq d^* \), where

\[
p^* = \min_{\vec{x} \in X} \max_{\vec{y} \in Y} F_0(\vec{x}, \vec{y}),
\]

\[
d^* = \max_{\vec{y} \in Y} \min_{\vec{x} \in X} F_0(\vec{x}, \vec{y}).
\]

This statement is referred to as the min-max theorem.
3. Minimizing a Sum of Logarithms

NOTE: This problem had some edits to problems/solutions for clarity. The problems should be the same, but introduce a little more notation for organizational purposes – you are fine (in terms of self-grades, etc) if you did not use this notation. The problems also have modified wording to accommodate this new notation. The solutions are the same, except: part (a) has a more complete solution, and part (c) has a revised solution. Part (c) had some ambiguity (whether to show strong duality holds for the relaxed or original problem) and any solution which answered either question correctly should get full marks.

Consider the following problem:

\[ p^* = \max_{\vec{x} \in \mathbb{R}^n} \sum_{i=1}^{n} \alpha_i \log(x_i) \]  
\[ \text{s.t. } \vec{x} \geq 0, \quad \vec{1}^\top \vec{x} = c, \]

where \( c > 0 \) and \( \alpha_i > 0, \ i = 1, \ldots, n \). (Recall that if \( \vec{x} \) is a vector then by “\( \vec{x} \geq 0 \)” we mean “\( x_i \geq 0 \) for each \( i \).”) Problems of this form arise, for instance, in maximum-likelihood estimation of the transition probabilities of a discrete-time Markov chain. we will determine in closed-form a minimizer, and show that the optimal objective value of this problem is

\[ p^* = \alpha \log(c/\alpha) + \sum_{i=1}^{n} \alpha_i \log(\alpha_i), \]

where \( \alpha = \sum_{i=1}^{n} \alpha_i \). We will show this in a series of steps.

(a) First, express the problem as a minimization problem which has optimal value \( p^*_{\text{min}} \). Then, show that the same problem with a relaxed equality constraint, i.e., \( \vec{1}^\top \vec{x} \leq c \) instead of \( \vec{1}^\top \vec{x} = c \), has the same optimal value \( p^* = p^*_{\text{min}} \), and the same solutions.

HINT: Argue by contradiction. Let \( \vec{x}^* \) be a solution to the “relaxed” minimization problem (i.e., the version of the problem with inequality constraints) which is not feasible for the original minimization problem. Consider the solution vector given by

\[ \vec{x}^* \triangleq \begin{bmatrix} c - \vec{1}^\top \vec{x}^r + x_1^r \\ x_2^r \\ \vdots \\ x_n^r \end{bmatrix} \]

(b) After relaxing the equality constraint to an inequality constraint, form the Lagrangian \( \mathcal{L}(\vec{x}, \mu) \) for the relaxed minimization problem, where \( \mu \) is the dual variable corresponding to the inequality constraint \( \vec{1}^\top \vec{x} \leq c \). (You do not need to “dualize” the constraint \( \vec{x} \geq 0 \), meaning you do not need to add a term to the Lagrangian or a dual variable for it.)

(c) Now derive the dual function \( g(\mu) \) for the relaxed minimization problem, and solve the dual problem \( d^*_\mu = \max_{\mu \geq 0} g(\mu) \). What is the optimal dual variable \( \mu^* \)?

(d) Show that strong duality holds for the relaxed problem, so \( p^*_r = d^*_\mu \).

(e) From the \( \mu^* \) obtained in the previous part, how do we obtain the optimal primal variable \( x^* \)?

What is the optimal objective function value \( p^*_r \)? Finally, what is \( p^* \)?
4. A Linear Program

Let \( A \in \mathbb{R}^{m \times n} \), \( \gamma \in \mathbb{R}^m \) and \( \mu > 0 \). First consider the following problem:

\[
p^* = \min_{\tilde{x}} \| A\tilde{x} - \gamma \|_1.
\]  \hspace{1cm} (10)

For \( j \in \{1, \ldots, n\} \), we denote by \( \tilde{a}_j \) the \( j \)-th column of \( A \), so that \( A = [\tilde{a}_1 \ldots \tilde{a}_n] \).

(a) Express the problem as an LP.

(b) Show that a dual to the problem can be written as

\[
d^* = \max_{\tilde{u}} -\tilde{u}^T \gamma : A^T \tilde{u} = 0, \|\tilde{u}\|_\infty \leq 1.
\]  \hspace{1cm} (11)

**HINT:** Use the fact that, for any vector \( z \):

\[
\max_{\tilde{u} : \|\tilde{u}\|_1 \leq 1} \tilde{u}^T \tilde{z} = \|\tilde{z}\|_\infty, \quad \max_{\tilde{u} : \|\tilde{u}\|_\infty \leq 1} \tilde{u}^T \tilde{z} = \|\tilde{z}\|_1.
\]  \hspace{1cm} (12)

Now, consider the following more complicated problem involving both the \( \ell_1 \) and \( \ell_\infty \) norms:

\[
p^* = \min_{\tilde{x}} \| A\tilde{x} - \gamma \|_1 + \mu \|\tilde{x}\|_\infty.
\]  \hspace{1cm} (13)

(c) Express the problem as an LP.

(d) Show that a dual to the problem can be written as

\[
d^* = \max_{\tilde{u}} -\tilde{u}^T \gamma : \|\tilde{u}\|_\infty \leq 1, \|A^T \tilde{u}\|_1 \leq \mu.
\]  \hspace{1cm} (14)

**HINT:** Use the fact that, for any vector \( z \):

\[
\max_{\tilde{u} : \|\tilde{u}\|_1 \leq 1} \tilde{u}^T \tilde{z} = \|\tilde{z}\|_\infty, \quad \max_{\tilde{u} : \|\tilde{u}\|_\infty \leq 1} \tilde{u}^T \tilde{z} = \|\tilde{z}\|_1.
\]  \hspace{1cm} (15)
5. Visualizing the Dual Problem

Download the Jupyter notebook `dual_visualize.ipynb`; complete the code where designated and answer the questions given in the space provided. (If you prefer, for questions that do not involve writing code, you can write solutions on separate paper or \LaTeX\ PDF, just make sure to correctly mark the relevant pages when uploading to Gradescope.)
6. Homework Process

With whom did you work on this homework? List the names and SIDs of your group members.

*NOTE:* If you didn’t work with anyone, you can put “none” as your answer.