Homework 9

This homework is due at 11 PM on March 23, 2023.

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. **Quadratic inequalities**

   Consider the set S defined by the following inequalities:

$$(x_1 \geq -x_2 + 1 \text{ and } x_1 \leq 0) \text{ or } (x_1 \leq -x_2 + 1 \text{ and } x_1 \geq 0)\,. \tag{1}$$

   To be more precise,

$$S_1 = \{\vec{x} \in \mathbb{R}^2 \mid x_1 \geq -x_2 + 1, x_1 \leq 0\} \tag{2}$$
$$S_2 = \{\vec{x} \in \mathbb{R}^2 \mid x_1 \leq -x_2 + 1, x_1 \geq 0\} \tag{3}$$
$$S = S_1 \cup S_2. \tag{4}$$

   (a) Draw the set S. Is it convex?

   (b) Show that the set S, can be described as a single quadratic inequality of the form $q(\vec{x}) = \vec{x}^\top A\vec{x} + 2\vec{b}^\top \vec{x} + c \leq 0$, for matrix $A = A^\top \in \mathbb{R}^{2\times 2}$, $\vec{b} \in \mathbb{R}^2$ and $c \in \mathbb{R}$ i.e S can be written as $S = \{\vec{x} \in \mathbb{R}^2 \mid q(\vec{x}) \leq 0\}$). Find $A, \vec{b}, c$.
   Hint: Can you combine the constraints to make one quadratic constraint?

   (c) Give the definition of the convex hull of a set. What is the convex hull of this set, i.e., S?

   (d) We will now consider some convex optimization problems over S_1 that illustrate the role of the constraints in the optimization problem. For each of the following optimization problems find the optimal point, $\vec{x}^*$. Describe the constraints that are active in attaining the optimal value. *Hint: Suppose that there exists a point $\vec{x}$ such that $\nabla f(\vec{x}) = 0$. From the first order characterization of a convex function $\vec{x}$ would be an optimum value for f subject to no constraints. If $\vec{x}$ is not in the constraint set S_1, then the optimum point must be on the boundary of the set, i.e. it satisfies at least one of the constraints defining S_1 with equality.*

      i. Minimize $f(\vec{x}) = (x_1 + 1)^2 + (x_2 - 3)^2$ subject to $\vec{x} \in S_1$.
      ii. Minimize $f(\vec{x}) = (x_1 + 2)^2 + (x_2 - 2)^2$ subject to $\vec{x} \in S_1$.
      iii. Minimize $f(\vec{x}) = x_1^2 + x_2^2$ subject to $\vec{x} \in S_1$.

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 $\mathcal{X}$ and $\mathcal{Y}$, respectively, which may or may not be convex.

(a) Show that

$$\min_{\vec{x} \in \mathcal{X}} \min_{\vec{y} \in \mathcal{Y}} F_0(\vec{x}, \vec{y}) = \min_{\vec{y} \in \mathcal{Y}} \min_{\vec{x} \in \mathcal{X}} F_0(\vec{x}, \vec{y}), \tag{5}$$

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 $p^\star \geq d^\star$, where

$$p^\star \doteq \min_{\vec{x} \in \mathcal{X}} \max_{\vec{y} \in \mathcal{Y}} F_0(\vec{x}, \vec{y}) \tag{6}$$

$$d^\star \doteq \max_{\vec{y} \in \mathcal{Y}} \min_{\vec{x} \in \mathcal{X}} F_0(\vec{x}, \vec{y}). \tag{7}$$

This statement is referred to as the *min-max theorem*.

3. 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.)

4. Dual of the dual of a linear program

Consider a standard linear program P:

$$\min_{\vec{x}} \quad \vec{c}^\top \vec{x} \tag{8}$$
$$\text{s.t.} \quad A\vec{x} = \vec{b} \tag{9}$$
$$x \geq 0. \tag{10}$$

where $\vec{x}, \vec{c} \in \mathbb{R}^n, \vec{b} \in \mathbb{R}^m, A \in \mathbb{R}^{m \times n}$.

(a) Formulate the Lagrangian of the problem P, and write the dual problem.

Note: The dual problem should not have the variable $\vec{x}$.

(b) Express the dual problem as an equivalent minimization problem. Find the dual of this minimization problem, i.e., the dual of the dual. Compare it to the original linear program formulation.

5. **Minimizing a Sum of Logarithms**

Consider the following problem, which arises in estimation of transition probabilities of a discrete-time Markov chain:

$$p^\star = \max_{\vec{x} \in \mathbb{R}^n} \quad \sum_{i=1}^{n} \alpha_i \log(x_i) \tag{11}$$

$$\text{s.t.} \quad \vec{x} \geq 0, \quad \vec{1}^\top \vec{x} = c, \tag{12}$$

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.") We will determine in closed-form a minimizer, and show that the optimal objective value of this problem is

$$p^\star = \alpha \log(c/\alpha) + \sum_{i=1}^{n} \alpha_i \log(\alpha_i), \tag{13}$$

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

(a) First, express the problem as a minimization problem which has optimal value $p^\star_{\text{min}}$.

(b) In optimization, we often "relax" problems of the form $p^\star_{\text{min}} = \min_{\vec{x} \in \mathcal{X}} f_0(\vec{x})$, i.e., replacing the constraint set $\mathcal{X}$ with a larger constraint set $\mathcal{X}_r$, and instead solving $p^\star_r = \min_{\vec{x} \in \mathcal{X}_r} f_0(\vec{x})$, then showing a connection between $p^\star_{\text{min}}$ and $p^\star_r$. In this problem, a particular relaxation we will use is to replace the equality constraint $\vec{1}^\top \vec{x} = c$ with an inequality constraint $\vec{1}^\top \vec{x} \leq c$.

Show that the relaxed problem has the same optimal value as the original problem, i.e., $p^\star_r = p^\star_{\text{min}}$, and the two problems have the same solutions.

HINT: First argue that $p^\star_r \leq p^\star_{\text{min}}$. Then, suppose for the sake of contradiction that $p^\star_r < p^\star_{\text{min}}$. Let $\vec{x}^r$ be a solution to the relaxed minimization problem which has objective value $p^\star_r$. Consider the vector $\vec{x}$ given by

$$\vec{x} \doteq \begin{bmatrix} c - 1^\top \vec{x}^r + x_1^r \\ x_2^r \\ \vdots \\ x_n^r \end{bmatrix}. \tag{14}$$

Show that $\vec{x}$ is feasible for the original problem and has objective value $< p^\star_r$. Argue that this implies $p^\star_{\text{min}} < p^\star_r$ and derive a contradiction. Finally, argue that any solution to the relaxed problem is a solution to the original problem, and vice-versa — you might need to use a construction similar to $\vec{x}$.

(c) After relaxing the equality constraint to an inequality constraint, form the Lagrangian $\mathcal{L}(\vec{x}, \vec{\lambda}, \mu)$ for the relaxed minimization problem, where λ_i is the dual variable corresponding to the inequality $x_i \geq 0$, and μ is the dual variable corresponding to the inequality constraint $\vec{1}^\top \vec{x} \leq c$.

(d) Now derive the dual function $g(\vec{\lambda}, \mu)$ for the relaxed minimization problem, and solve the dual problem $d^\star_r = \max\limits_{\substack{\vec{\lambda} \geq \vec{0} \\ \mu \geq 0}} g(\vec{\lambda}, \mu)$. What are the optimal dual variables $\vec{\lambda}^\star, \mu^\star$?

(e) Show that strong duality holds for the relaxed problem, so $p^\star_r = d^\star_r$.

(f) From the $\vec{\lambda}^\star, \mu^\star$ obtained in the previous part, how do we obtain the optimal primal variable $x^\star$? What is the optimal objective function value $p^\star_r$? Finally, what is $p^\star$?

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.