EECS 127/227AT Optimization Models in Engineering
Spring 2020 Homework 7 - PRACTICE

This homework is NEVER DUE. All problems are intended as practice for the midterm exam, and problems and solutions have been released simultaneously.

This version was compiled on 2020-03-07 11:06.

1. 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

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

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

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

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

2. (Sp '19 Midterm 2 #7) Gradient descent algorithm

Consider $g : \mathbb{R}^n \to \mathbb{R}$, $g(\vec{x}) = \frac{1}{2}\vec{x}^\top Q\vec{x} - \vec{x}^\top \vec{b}$, where Q is a symmetric positive definite matrix, i.e., $Q \in \mathbb{S}^n_{++}$.

(a) Write the update rule for the gradient descent algorithm

$$\vec{x}_{k+1} = \vec{x}_k - \eta\nabla g(\vec{x}_k),$$

where η is the step size of the algorithm, and bring it into the form

$$(\vec{x}_{k+1} - \vec{x}^*) = P_\eta(\vec{x}_k - \vec{x}^*),$$

where $P_\eta \in \mathbb{R}^{n\times n}$ is a matrix that depends on η. Find $\vec{x}^*$ and P_η in terms of Q, $\vec{b}$ and η. *Note:* $\vec{x}^*$ is a minimizer of g.

(b) Write a condition on the step size η and the matrix Q that ensures convergence of $\vec{x}_k$ to $\vec{x}^*$ for every initialization of $\vec{x}_0$.

(c) Assume all eigenvalues of Q are distinct. Let η_m denote the largest stepsize that ensures convergence for all initializations $\vec{x}_0$, based on the condition computed in part (b). Does there exist an initialization $\vec{x}_0 \neq \vec{x}^*$ for which the algorithm converges to the minimum value of g for certain values of the step size η that are larger than η_m? Justify your answer.

Hint: The question asks if such initializations exist; not whether it is practical to find them.

3. (Sp '19 Midterm 2 #3) Convexity of sets

Determine if each set C given below is convex. Prove that each set is convex or provide an example to show that it is not convex. You may use any techniques used in class or discussion to demonstrate or disprove convexity.

(a) $C = \{\vec{x} \in \mathbb{R}^2 \mid x_1 x_2 \geq 0\}$, where $\vec{x} = \begin{bmatrix} x_1 & x_2 \end{bmatrix}^\top$.

(b) $C = \{X \in \mathbb{S}^n \mid \lambda_{\min}(X) \geq 2\}$, where $\mathbb{S}^n$ is the set of symmetric matrices in $\mathbb{R}^{n \times n}$ and $\lambda_{\min}(X)$ is the minimum eigenvalue of X.

(c) Let $\mathcal{H}(\vec{w})$ denote the hyperplane with normal direction $\vec{w} \in \mathbb{R}^n$, i.e.,

$$\mathcal{H}(\vec{w}) = \{\vec{x} \in \mathbb{R}^n \mid \vec{x}^\top \vec{w} = 0\}.$$

Let $P : \mathbb{R}^n \to \mathbb{R}^n$ be given by

$$P(\vec{x}) = \underset{\vec{y} \in \mathcal{H}(\vec{w})}{\operatorname{argmin}} \|\vec{y} - \vec{x}\|_2 .$$

Let

$$C = \{P(\vec{x}) \mid \vec{x} \in \mathcal{B}\}$$

where $\mathcal{B} = \{\vec{x} \in \mathbb{R}^n \mid \|\vec{x}\|_2 \leq 1\}$.

4. Minimizing a sum of logarithms

Consider the following problem:

$$\begin{aligned} p^* = \max_{x \in \mathbb{R}^n} \quad & \textstyle\sum_{i=1}^n \alpha_i \ln x_i \\ \text{s.t.} \quad & x \geq 0, \quad \mathbf{1}^\top x = c, \end{aligned}$$

where $c > 0$ and $\alpha_i > 0$, $i = 1, \ldots, n$. Problems of this form arise, for instance, in maximum-likelihood estimation of the transition probabilities of a discrete-time Markov chain.

Determine in closed-form a minimizer, and show that the optimal objective value of this problem is

$$p^* = \alpha \ln(c/\alpha) + \sum_{i=1}^n \alpha_i \ln \alpha_i,$$

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. Then, can you relax the equality constraint to an inequality constraint while preserving the set of solutions?

(b) After relaxing the equality constraint to an inequality constraint, form the Lagrangian $\mathcal{L}(x, \mu)$ for this problem, where μ is the dual variable corresponding to the inequality constraint containing c.

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

(d) Assume strong duality holds, so $p^* = d^*$. (We will prove why this holds later). From the μ^* obtained in the previous part, how do we obtain the optimal primal variable x^*? And finally, what is the optimal objective function value p^*?

5. Homework process

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