**EECS 127/227AT** **Fall 2020**

**Homework 7**

Homework 7 is due on Gradescope by Friday 10/30 at 11.59 p.m.

## 1 Formulating problems as LPs or QPs

This problem explores what kinds of problems can be formulated as LPs or QPs. The relevant portions of the textbooks are Secs. 9.1–9.5 of the textbook of Calfiore and El Ghaoui and Secs. 4.3–4.4 of the textbook of Boyd and Vandenberghe.

Formulate the problem

$$p_j^* := \min_x \ f_j(x),$$

for different functions $f_j$, $j = 1, \ldots, 4$, as QPs or LPs, or, if you cannot, explain why.

In our formulations, we always use $x \in \mathbb{R}^n$ as the variable, and assume that $A \in \mathbb{R}^{m \times n}$, $y \in \mathbb{R}^m$. If you obtain an LP or QP formulation, make sure to put the problem in standard form, stating precisely what the variables, objective, and constraints are.

(a) $f_1(x) = \|Ax - y\|_\infty + \|x\|_1$.

(b) $f_2(x) = \|Ax - y\|_2^2 + \|x\|_1$.

(c) $f_3(x) = \|Ax - y\|_2^2 - \|x\|_1$.

(d) $f_4(x) = \|Ax - y\|_2^2 + \|x\|_1^2$.

## 2 Bounds on a linear fractional function

While this problem discusses an SOCP, it is basically about ways to reformulate one optimization problem into another. The relevant sections of the textbooks are Secs. 8.3–8.4 of the textbook of Calafiore and El Ghaoui and Secs. 5.1–5.2 of the textbook of Boyd and Vandenberghe.

Consider the function $f : \mathbb{R}^n \to \mathbb{R}$, with values on its domain $\mathrm{dom}(f) := \{x \in \mathbb{R}^n : \|x\|_2 \leq 1\}$ given by

$$f(x) = \frac{b^T x}{1 - a^T x}.$$

Here, $a, b \in \mathbb{R}^n$ are given, with $a, b \neq 0$ and $\|a\|_2 < 1$. In this exercise, we consider the problem

$$p^* := \max_{\|x\|_2 \leq 1} f(x).$$

(a) Show that indeed $f$ is well-defined everywhere on $\{x \in \mathbb{R}^n : \|x\|_2 \leq 1\}$.

(b) Is the above problem convex, as stated? Provide a proof or a counterexample.

(c) Show that the problem can be expressed as an SOCP in one variable, namely

$$p^* = \min_t \ t \ : \ \|at + b\|_2 \leq t.$$

**Hint**: For given $t \geq 0$, express the condition $f(x) \leq t$ for every $x \in \mathbb{R}^n$ such that $\|x\|_2 \leq 1$ in simple terms.

(d) Show that $p^* = t^*$, where $t^*$ is the (unique) positive solution to the equation $t = \|at + b\|_2$. Express the optimal value $p^*$ in closed-form.

(e) Show that $x^* := a + (1/t^*)b$ is an optimal point for the original problem.

(f) What about a lower bound on $f$? Derive an explicit expression for two scalars $\hat{p}, \sigma > 0$ such that $f(x) \in [\hat{p} - \sigma, \hat{p} + \sigma]$ on $\{x \in \mathbb{R}^n : \|x\|_2 \leq 1\}$.

## 3 A Slalom Problem

A skier must slide from left to right by going through $n$ parallel gates of known position $(x_i, y_i)$ and width $c_i$, $i = 1, \ldots, n$. The initial position $(x_0, y_0)$ is given, as well as the final one, $(x_{n+1}, y_{n+1})$. Before reaching the final position, the skier must go through gate $i$ by passing between the points $(x_i, y_i - c_i/2)$ and $(x_i, y_i + c_i/2)$ for each $i \in \{1, \ldots, n\}$. Figure 1 is a representation. Use values for $(x_i, y_i, c_i)$ from Table 1.

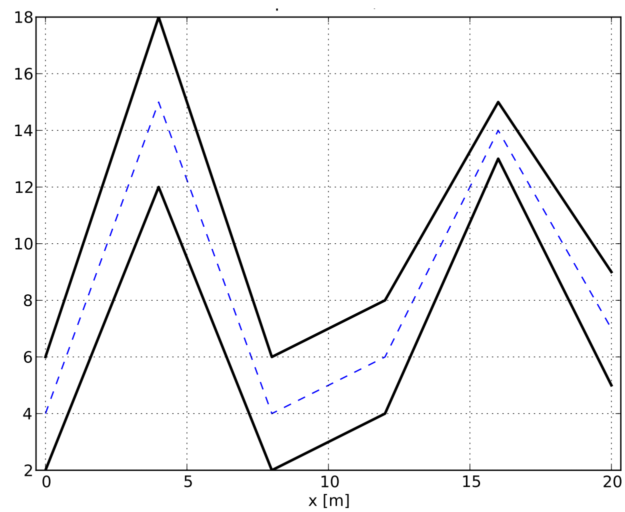


Figure 1: Slalom problem with $n = 6$ gates. The initial and final positions are fixed and not included in the figure. The skier slides from left to right. The middle path is dashed and connects the center points of gates.

Table 1: Problem data for Problem 2.

| $i$ | $x_i$ | $y_i$ | $c_i$ |
|---|---|---|---|
| 0 | 0 | 4 | $N/A$ |
| 1 | 4 | 15 | 3 |
| 2 | 8 | 4 | 2 |
| 3 | 12 | 6 | 2 |
| 4 | 16 | 14 | 1 |
| 5 | 20 | 7 | 2 |
| 6 | 24 | 4 | $N/A$ |

(a) Given the data $\{(x_i, y_i, c_i)\}_{i=0}^{n+1}$, write an optimization problem that minimizes the total length of the path. Your answer should come in the form of an SOCP.

(b) Solve the problem numerically with the data given in Table 1. You may use this starter code.

*Hint:* You should be able to use packages such as `cvxpy` and `numpy`.

## 4 Prefix-free codes and entropy

This problem discusses an important application of the convex optimization in communication theory, and how the concept of entropy can be discovered via the KKT conditions. The relevant sections of the textbooks are Secs. 8.4–8.5 of the textbook of Calafiore and El Ghaoui and Secs. 5.3–5.5 of the textbook of Boyd and Vandenberghe.

Let $(p_1, \ldots, p_n)$ be a probability distribution on the set $\{1, \ldots, n\}$, i.e. $p_i \geq 0$ for $1 \leq i \leq n$ and $\sum_{i=1}^{n} p_i = 1$. The unit simplex in $\mathbb{R}^n$, denoted $\Delta_n$ is, by definition

$$\{x \in \mathbb{R}^n : x_i \geq 0 \text{ for all } 1 \leq i \leq n \text{ and } \textstyle\sum_{i=1}^{n} x_i = 1\}.$$

Therefore the set of probability distributions on $\{1, \ldots, n\}$ can be identified with $\Delta_n$.

A *code* for the set $\{1, \ldots, n\}$ is a mapping from $\{1, \ldots, n\}$ to $\{0,1\}^* - \emptyset$. Here $\{0,1\}^* - \emptyset := \bigcup_{n \geq 1} \{0,1\}^n$ denotes the set of all binary strings other than the empty string. The image of $i$ under the code is called the *codeword* associated to $i$. We are interested in the length sequence of the codewords, which we denote as $(l_1, \ldots, l_n)$ or, more compactly, as $l \in \mathbb{N}^n$.

A code is said to be *prefix-free* if no codeword is a prefix of any other codeword. A well-known and easily proved inequality for the length sequence of a prefix-free code is *Kraft's inequality*, which states that

$$\sum_{i=1}^{n} 2^{-l_i} \leq 1$$

for the length sequence of any prefix-free codeword. To see why this is true, think of the binary strings as the nodes of a binary tree, associate the weight $2^k$ to a node at depth $k$ (i.e. a node corresponding to a binary string of length $k$) and add up the weights of all the nodes in the tree that correspond to codewords. The prefix-free condition boils down to there being at most one codeword on any path in the binary tree starting from the root and this observation will lead you to a proof of Kraft's inequality after some thought.

An interesting question in coding theory, with applications in digital communications, is to ask what the least expected length is among prefix-free codes. We will study this problem in this question.

Let $p \in \Delta_n$ be a probability distribution. Our goal is to solve the integer program

$$\min_{l \in \mathbb{N}^n} \; p^T l \; : \; \sum_{i=1}^{n} 2^{-l_i} \leq 1.$$

Here we will study an approximation of this problem, which we get by relaxing the integrality constraints and considering the relaxed optimization problem

$$\min_{l \in \mathbb{R}^n_+} \; p^T l \; : \; \sum_{i=1}^{n} 2^{-l_i} \leq 1.$$

We also observe that if any $l_i < 0$ then the constraint is violated, so this optimization problem is equivalent to

$$\min_{l \in \mathbb{R}^n} \; p^T l \; : \; \sum_{i=1}^{n} 2^{-l_i} \leq 1. \tag{1}$$

(a) Find the optimal solution of the optimization problem (1). Your answer will involve the so-called *entropy* of the probability distribution $p$, defined as

$$H(p) := -\sum_{i=1}^{n} p_i \log_2 p_i.$$

(b) Suggest a rounding scheme for the optimal solution of the convex relaxation and give an approximation guarantee.

## 5 A matrix problem with strong duality

This problem discusses a convex optimization problem arising from the perturbation analysis of dynamical systems. The relevant portions of the textbooks are Secs. 8.3–8.5 of the textbook of Calafiore and El Ghaoui and Secs. 5.5–5.5 of the textbook of Boyd and Vandenberghe.

Consider the problem

$$p^* \doteq \min_{\Delta} \; c^\top (A+\Delta)^{-1} b \;:\; \|\Delta\| \leq 1,$$

where $A \in \mathbb{R}^{n\times n}$, with smallest singular value $\sigma_{\min}(A)$ strictly greater than one, and $b, c \in \mathbb{R}^n$ with $b, c \neq 0$. Here, $\|\cdot\|$ stands for the largest singular value norm, i.e. the spectral norm. This problem arises in the study of equilibrium states of a dynamical system subject to perturbations.

(a) Show that the objective function is well-defined everywhere on the feasible set.

**Hint:** You can show that a square matrix is invertible if it has no singular values equal to $0$.

(b) Is the problem, as stated, convex? Give a proof or a counter-example.

(c) Show that the problem can be expressed as

$$\min_x \; c^\top x \;:\; \|Ax - b\|_2^2 \leq \|x\|_2^2.$$

(d) Let $K := A^\top A - I$. Since $\sigma_{\min}(A) > 1$, we know that $K$ is invertible. Prove that

$$AK^{-1}A^\top - I = (AA^\top - I)^{-1}.$$

(e) Show that the feasible set of the formulation in (c) is an ellipsoid, expressing it in terms of the matrix $K := A^\top A - I$, the vector $x_0 := K^{-1}A^\top b$, and the scalar $\gamma := x_0^\top K x_0 - b^\top b$. Explain why the above problem (which we called the new problem) is convex.

(f) Form a Lagrange dual to the problem. Does strong duality hold?

(g) Show that the optimal value can be written

$$p^* = c^\top (A^\top A - I)^{-1} A^\top b - \|(AA^\top - I)^{-1/2} b\|_2 \cdot \|(A^\top A - I)^{-1/2} c\|_2.$$