# EECS 127/227AT Optimization Models in Engineering
Spring 2020

# Homework 11

**This homework is due Friday, April 17, 2020 at 23:00 (11pm).**
**Self grades are due Friday, April 24, 2020 at 23:00 (11pm).**

This version was compiled on 2020-04-11 20:09.

**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 your IPython notebook with solutions saved as a PDF.

**1. Median versus average**

For a given vector $\vec{v} \in \mathbb{R}^n$, the average can be found as the solution to the optimization problem

$$\min_{x \in \mathbb{R}} \|\vec{v} - x\vec{\mathbf{1}}\|_2^2, \tag{1}$$

where $\vec{\mathbf{1}}$ is the vector of ones in $\mathbb{R}^n$. Similarly, the median (any value $x$ such that there is an equal number of values in $v$ above or below $x$) can be found via

$$\min_{x \in \mathbb{R}} \|\vec{v} - x\vec{\mathbf{1}}\|_1. \tag{2}$$

We consider a robust version of the average problem (1):

$$\min_{x} \max_{\vec{u}\,:\,\|\vec{u}\|_\infty \le \lambda} \|\vec{v} + \vec{u} - x\vec{\mathbf{1}}\|_2^2, \tag{3}$$

in which we assume that the components of $v$ can be independently perturbed by a vector $u$ whose magnitude is bounded by a given number $\lambda \ge 0$.

(a) Is the robust problem (3) convex? Justify your answer precisely, based on expression (3), and without further manipulation.

(b) Show that problem (3) can be expressed as

$$\min_{x \in \mathbb{R}} \sum_{i=1}^{n} \left(|v_i - x| + \lambda\right)^2.$$

(c) Express the problem as a QP. State precisely the variables, and constraints if any.

(d) Show that when $\lambda$ is large, the solution set approaches that of the median problem (2).

(e) It is often said that the median is a more robust notion of “middle” value than the average, when noise is present in $\vec{v}$. Based on the previous part, justify this statement.

**2. LASSO vs. Ridge**

Say that we have the data set $\{(\vec{x}^{(i)}, y^{(i)})\}_{i=1,\dots,n}$ of features $\vec{x}^{(i)} \in \mathbb{R}^d$ and values $y^{(i)} \in \mathbb{R}$. Define $X = \begin{bmatrix}\vec{x}^{(1)} & \dots & \vec{x}^{(n)}\end{bmatrix}^\top$ and $y = \begin{bmatrix}y^{(1)} & \dots & y^{(n)}\end{bmatrix}^\top$. For the sake of simplicity, assume that the data has been centered and whitened so that each feature has mean 0 and variance 1 and the features are uncorrelated, i.e. $X^\top X = nI$.

Consider the linear least squares regression with regularization in the $\ell_1$-norm, also known as LASSO:

$$\vec{w}^* = \arg\min_{\vec{w} \in \mathbb{R}^d} \|X\vec{w} - \vec{y}\|_2^2 + \lambda \|\vec{w}\|_1.$$

This problem will compare $\ell_1$-regularization with $\ell_2$-regularization (ridge regression) to understand their similarities and differences. We will do this by looking at the elements of $\vec{w}^*$ in the solution to each problem.

(a) First, we decompose this optimization problem into $d$ univariate optimization problems over each element of $\vec{w}$. Let $X = \begin{bmatrix} \vec{x}_1 & \dots & \vec{x}_d \end{bmatrix}$ and recall that $X^\top X = nI$.

(b) If $w_i^* > 0$, then what is the value of $w_i^*$? What is the condition on $\vec{y}^\top \vec{x}_i$ for this to be possible?

(c) If $w_i^* < 0$, then what is the value of $w_i^*$? What is the condition on $\vec{y}^\top \vec{x}_i$ for this to be possible?

(d) What can we conclude about $w_i^*$ if $|\vec{y}^\top \vec{x}_i| \leq \frac{\lambda}{2}$? How does the value of $\lambda$ impact the individual entries $w_i^*$?

(e) Now consider the case of ridge regression, which uses the the $\ell_2$ regularization $\lambda\|\vec{w}\|_2^2$.

$$\vec{w}^* = \arg\min_{\vec{w} \in \mathbb{R}^d} \|X\vec{w} - \vec{y}\|_2^2 + \lambda \|\vec{w}\|_2^2.$$

Write down the new condition for $\vec{w}_i^*$ to be 0. How does this differ from the condition obtained in part (4) and what does this suggest about LASSO?

### 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 $c_i$, $i = 1, \dots, n$. 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 $i \in \{1, \dots, n\}$. See Figure 1.

Table 1: Problem data for Problem 2.

| $i$ | $x_i$ | $y_i$ | $c_i$ |
|---|---|---|---|
| 0 | 0 | 4 | $N/A$ |
| 1 | 4 | 5 | 3 |
| 2 | 8 | 4 | 2 |
| 3 | 12 | 6 | 2 |
| 4 | 16 | 5 | 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.

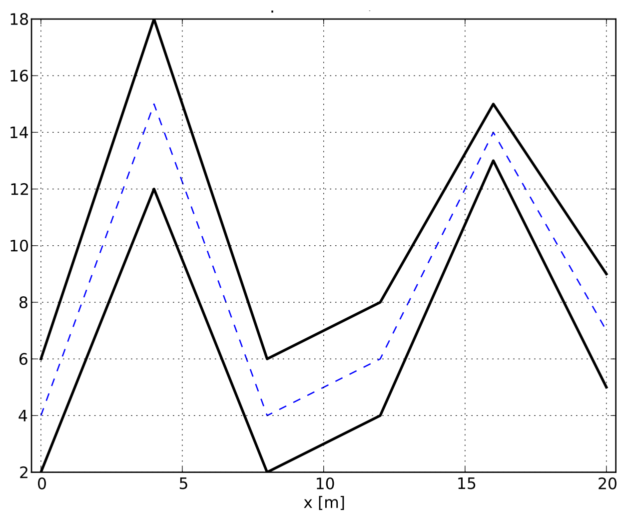


Figure 1: Slalom problem with $n = 6$ gates. The initial and final positions are fixed and not included in the figure. The skier slides from left to right. The middle path is dashed and connects the center points of gates.

(b) Solve the problem numerically with the data given in Table 1.
*Hint:* You should be able to use packages such as `cvxpy` and `numpy`.

**4. Robust linear programming**

In this problem we will consider a version of linear programming under uncertainty.

(a) Consider vector $\vec{x} \in \mathbb{R}^n$. Recall that $\vec{x}^\top \vec{y} \leq \|\vec{x}\|_1$ for all $\vec{y}$ such that $||\vec{y}||_\infty \leq 1$. Further this inequality is tight, since it holds with equality for $\vec{y} = \text{sgn}(\vec{x})$. You saw this in the previous homework, just remind your self of the solution, no need to turn anything in.

Let us focus now on a LP in standard form:

$$\begin{aligned} &\min_{\vec{x}} \vec{c}^\top \vec{x} \\ \text{s.t. } &\vec{a}_i^\top \vec{x} \leq b_i, \quad i = 1, ..., m. \end{aligned} \tag{4}$$

Consider the set of linear inequalities in (4). Suppose you don't know the vectors $\vec{a}_i$ exactly. Instead you are given nominal values $\vec{\overline{a}}_i$, and you know that the actual vectors satisfy $\|\vec{a}_i - \vec{\overline{a}}_i\|_\infty \leq \rho$ for a given $\rho > 0$. In other words, the actual components $a_{ij}$ can be anywhere in the intervals $[\overline{a}_{ij} - \rho, \overline{a}_{ij} + \rho]$. Or equivalently, each vector $\vec{a}_i$ can lie anywhere in a hypercube with corners $\vec{\overline{a}}_i + \vec{v}$ where $\vec{v} \in \{-\rho, \rho\}^n$. We desire that the set of inequalities that constrain problem (4) be satisfied for all possible values of $\vec{a}_i$; i.e., we replace these with the constraints

$$\vec{a}_i^\top \vec{x} \leq b_i \ \ \forall \vec{a}_i \in \{\vec{\overline{a}}_i + \vec{v} \ \mid \|\vec{v}\|_\infty \leq \rho\} \ \ i = 1, ..., m. \tag{5}$$

Note that the above defines an *infinite* number of constraints (of the form $\vec{\overline{a}}_i^\top \vec{x} + \vec{v}^\top \vec{x} \leq b_i, \ \ \forall \vec{v}$ satisfying $\|\vec{v}\|_\infty \leq \rho$, $i = 1, 2, \ldots, m$).

(b) Argue why for our LP we can replace the infinite set of constraints as above to a finite set of $2^n m$ constraints of the form,

$$\vec{\bar{a}}_i^\top \vec{x} + \vec{v}^\top \vec{x} \leq b_i \quad \forall \vec{v} \in \{-\rho, \rho\}^n \quad i = 1, ..., m.$$

*Hint: What do you know about the optimal solutions of LPs?*

(c) Use result from part (a) to show that the constraint set in Equation (5) is in fact equivalent to the much more compact set of $m$ nonlinear inequalities

$$\vec{\bar{a}}_i^\top \vec{x} + \rho \|\vec{x}\|_1 \leq b_i, \quad i = 1, ..., m.$$

We now would like to formulate the LP with uncertainty introduced. We are therefore interested in situations where the vectors $\vec{a}_i$ are uncertain, but satisfy bounds $\|\vec{a}_i - \vec{\bar{a}}_i\|_\infty \leq \rho$ for given $\vec{\bar{a}}_i$ and $\rho$. We want to minimize $\vec{c}^\top \vec{x}$ subject to the constraint that the inequalities $\vec{a}_i^\top \vec{x} \leq b_i$ are satisfied for *all* possible values of $\vec{a}_i$.

We call this a *robust LP*:

$$\begin{aligned} &\min_{\vec{x}} \vec{c}^\top \vec{x} \\ &\text{s.t. } \vec{a}_i^\top \vec{x} \leq b_i, \quad \forall \vec{a}_i \in \{\vec{\bar{a}}_i + \vec{v} \mid \|\vec{v}\|_\infty \leq \rho\} \quad i = 1, ..., m. \end{aligned} \tag{6}$$

(d) Using the result from part (c), express the above optimization problem as an LP.

**5. Formulating problems as LPs or QPs**

Formulate the problem

$$p_j^* \doteq \min_{\vec{x}} \; f_j(\vec{x}),$$

for different functions $f_j$, $j = 1, \ldots, 4$, as convex QPs or LPs, or, if you cannot, explain why. In our formulations, we always use $\vec{x} \in \mathbb{R}^n$ as the variable, and assume that $A \in \mathbb{R}^{m \times n}$, $\vec{y} \in \mathbb{R}^m$. If you obtain a convex LP or QP formulation, state precisely what the variables, objective, and constraints are.

(a) $f_1(\vec{x}) = \|A\vec{x} - \vec{y}\|_\infty + \|\vec{x}\|_1$

(b) $f_2(\vec{x}) = \|A\vec{x} - \vec{y}\|_2^2 + \|\vec{x}\|_1$

(c) $f_3(\vec{x}) = \|A\vec{x} - \vec{y}\|_2^2 - \|\vec{x}\|_1$

(d) $f_4(\vec{x}) = \|A\vec{x} - \vec{y}\|_2^2 + \|\vec{x}\|_1^2$

**6. Sphere enclosure**

Let $B_i$, $i = 1, \ldots, m$, be $m$ Euclidean balls in $\mathbb{R}^n$, with centers $\vec{x}_i$, and radii $\rho_i \geq 0$. We wish to find a ball $B$ of minimum radius that contains all the $B_i$, $i = 1, \ldots, m$. Cast this problem as an SOCP.

**7. Homework process**

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