

# EECS 127/227AT Optimization Models in Engineering

## Spring 2020

## Homework 10

**This homework is due Friday, April 10, 2020 at 23:00 (11pm).**

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

This version was compiled on 2020-04-12 23:08.

**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. (Optional) Strong Duality but no KKT

In this question, we will see an example of a problem where strong duality holds but the KKT conditions don't hold. Consider the following problem:

$$\begin{aligned} p^* = \min_{x_1, x_2 \in \mathbb{R}} \quad & x_1^2 + x_2^2 \\ \text{s.t.} \quad & (x_1 - 1)^2 + (x_2 - 1)^2 \leq 1 \\ & (x_1 - 1)^2 + (x_2 + 1)^2 \leq 1. \end{aligned}$$

- (a) Sketch the feasible set. Find the optimal solution  $x^*$ .

**Solution:** From Fig. 1, we can see that the feasible set contains only one point,  $x = (1, 0)$

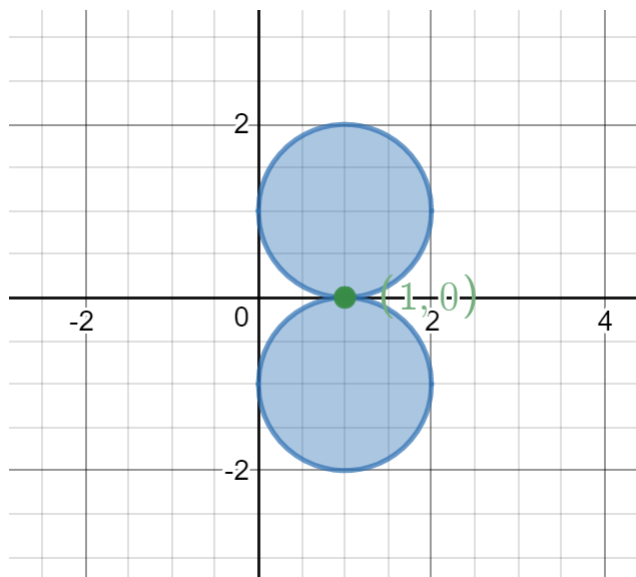


Figure 1: Feasible set

and thus  $x^* = (1, 0)$  and  $p^* = 1$ .

- (b) Write the KKT conditions and solve them.

**Solution:** The KKT conditions are,

$$2x_1 + 2\lambda_1(x_1 - 1) + 2\lambda_2(x_1 - 1) = 0 \quad (1)$$

$$2x_2 + 2\lambda_1(x_2 - 1) + 2\lambda_2(x_2 + 1) = 0$$

$$(x_1 - 1)^2 + (x_2 - 1)^2 \leq 1 \quad (2)$$

$$(x_1 - 1)^2 + (x_2 + 1)^2 \leq 1 \quad (3)$$

$$\lambda_1 \geq 0$$

$$\lambda_2 \geq 0$$

$$\lambda_1((x_1 - 1)^2 + (x_2 - 1)^2 - 1) = 0$$

$$\lambda_2((x_1 - 1)^2 + (x_2 + 1)^2 - 1) = 0.$$

From Equations 2 and 3, we have  $x_1 = 1, x_2 = 0$ . Substituting in 1 we get,

$$2 = 0,$$

which is clearly impossible so the KKT conditions have no solution.

- (c) Formulate and solve the Lagrange dual problem. Does strong duality hold?

**Solution:** The Lagrangian is given by,

$$\begin{aligned} \mathcal{L}(x_1, x_2, \lambda_1, \lambda_2) &= x_1^2 + x_2^2 + \lambda_1((x_1 - 1)^2 + (x_2 - 1)^2 - 1) + \lambda_2((x_1 - 1)^2 + (x_2 + 1)^2 - 1) \\ &= (1 + \lambda_1 + \lambda_2)x_1^2 + (1 + \lambda_1 + \lambda_2)x_2^2 - 2(\lambda_1 + \lambda_2)x_1 - 2(\lambda_1 - \lambda_2)x_2 + \lambda_1 + \lambda_2. \end{aligned}$$

Setting the derivative with respect to  $x_1, x_2$  to 0 we get that  $\mathcal{L}$  achieves minimum at,

$$\begin{aligned} x_1 &= \frac{\lambda_1 + \lambda_2}{1 + \lambda_1 + \lambda_2}, \\ x_2 &= \frac{\lambda_1 - \lambda_2}{1 + \lambda_1 + \lambda_2}. \end{aligned}$$

The dual function is given by,

$$g(\lambda_1, \lambda_2) = \begin{cases} -\frac{(\lambda_1 + \lambda_2)^2 + (\lambda_1 - \lambda_2)^2}{1 + \lambda_1 + \lambda_2} + \lambda_1 + \lambda_2, & 1 + \lambda_1 + \lambda_2 \geq 0 \\ -\infty, & \text{otherwise.} \end{cases}$$

The dual problem writes,

$$d^* = \sup_{\lambda_1 \geq 0, \lambda_2 \geq 0} \frac{\lambda_1 + \lambda_2 - (\lambda_1 - \lambda_2)^2}{1 + \lambda_1 + \lambda_2}.$$

Since  $g$  is symmetric in  $\lambda_1, \lambda_2$ , the optimum if it exists must have  $\lambda_1^* = \lambda_2^*$ . Using this the dual function simplifies to  $g(\lambda) = \frac{2\lambda}{1+2\lambda}$ . As  $\lambda_1 \rightarrow \infty$  we see that  $g(\lambda) \rightarrow 1$  and  $g(\lambda) \leq 1$ . This gives  $d^* = 1$  but there does not exist dual optimal solution. Since  $d^* = p^* = 1$ , strong duality holds. The KKT conditions do not have a solution here because dual optimal solution is not attained.

This problem highlights the difference between using sup vs max for the dual problem. Note that the supremum is well defined while the maximum is not attained.

## 2. The Duality of the $\ell_1$ and $\ell_\infty$ norms

For this problem, we will prove the duality of the  $\ell_1$  and  $\ell_\infty$  norms. Recall that the  $\ell_1$  and  $\ell_\infty$  norms, denoted by  $\|\cdot\|_1$  and  $\|\cdot\|_\infty$  respectively, are defined as follows for  $\vec{x} \in \mathbb{R}^n$ :

$$\|\vec{x}\|_1 = \sum_{i=1}^n |x_i|, \quad \|\vec{x}\|_\infty = \max_{i \in [n]} |x_i|.$$

We will show that the  $\ell_1$  and  $\ell_\infty$  norms are duals of each other; that is, we will show that:

$$\|\vec{x}\|_1 = \max_{\|\vec{y}\|_\infty=1} \vec{y}^\top \vec{x} \text{ and } \|\vec{x}\|_\infty = \max_{\|\vec{y}\|_1=1} \vec{y}^\top \vec{x}.$$

(a) To start, we will first prove the following inequality for all  $\vec{x}, \vec{y} \in \mathbb{R}^n$ :

$$\vec{x}^\top \vec{y} \leq \|\vec{x}\|_1 \|\vec{y}\|_\infty.$$

**Solution:** Note that when  $\vec{x} = 0$  or  $\vec{y} = 0$ , the left hand side of the inequality is 0 while the right hand side is non-negative. This proves the inequality when either  $\vec{x}$  or  $\vec{y}$  is 0. Now, assume that  $\vec{x}, \vec{y} \neq 0$ . We now show the inequality by the following series of inequalities:

$$\vec{x}^\top \vec{y} = \sum_{i=1}^n x_i y_i \leq \sum_{i=1}^n |x_i| |y_i| \leq \sum_{i=1}^n |x_i| \|\vec{y}\|_\infty \leq \|\vec{y}\|_\infty \sum_{i=1}^n |x_i| = \|\vec{x}\|_1 \|\vec{y}\|_\infty.$$

(b) Now, show that:

$$\max_{\|\vec{y}\|_\infty=1} \vec{y}^\top \vec{x} \geq \|\vec{x}\|_1$$

and using (a), conclude that  $\|\vec{x}\|_1 = \max_{\|\vec{y}\|_\infty=1} \vec{y}^\top \vec{x}$ .

**Solution:** Let  $\vec{y}$  be defined as  $y_i = \text{sgn}(x_i)$ . Note that  $\|\vec{y}\|_\infty = 1$ . Therefore, we have:

$$\vec{y}^\top \vec{x} = \sum_{i=1}^n y_i x_i = \sum_{i=1}^n |x_i| = \|\vec{x}\|_1.$$

Therefore, we can conclude that  $\max_{\|\vec{y}\|_\infty=1} \vec{y}^\top \vec{x} \geq \|\vec{x}\|_1$ . Note that from (a), we also get  $\max_{\|\vec{y}\|_\infty=1} \vec{y}^\top \vec{x} \leq \|\vec{x}\|_1$ . From both of these inequalities, we have:

$$\|\vec{x}\|_1 = \max_{\|\vec{y}\|_\infty=1} \vec{y}^\top \vec{x}.$$

(c) Finally, show the following inequality:

$$\max_{\|\vec{y}\|_1=1} \vec{y}^\top \vec{x} \geq \|\vec{x}\|_\infty$$

and prove the second equality.

**Solution:** Let  $i^* = \text{argmax}_{i \in [n]} |x_i|$ . Note that  $|x_{i^*}| = \|\vec{x}\|_\infty$ . Now, define  $\vec{y}$  as:

$$y_i = \begin{cases} 0 & i \neq i^* \\ \text{sgn}(x_{i^*}) & \text{o.w} \end{cases}.$$

Notice that  $\|\vec{y}\|_1 = 1$ . With  $\vec{y}$ , we have:

$$\vec{y}^\top \vec{x} = |x_{i^*}| = \|\vec{x}\|_\infty.$$

From this, we get that  $\max_{\|\vec{y}\|_1=1} \vec{y}^\top \vec{x} \geq \|\vec{x}\|_\infty$ . From (a), we again obtain  $\max_{\|\vec{y}\|_1=1} \vec{y}^\top \vec{x} \leq \|\vec{x}\|_\infty$ . The previous two inequalities allow us to conclude:

$$\max_{\|\vec{y}\|_1=1} \vec{y}^\top \vec{x} = \|\vec{x}\|_\infty.$$

### 3. A Linear Program

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

$$p^* = \min_x \|Ax - y\|_1.$$

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

(a) Express the problem as an LP.

**Solution:** The problem writes

$$\min_{\vec{x}, \vec{z}} z^T \mathbf{1} : \quad z_i \geq |(Ax - y)_i|, \quad i = 1, \dots, m.$$

which is an LP:

$$\min_{\vec{x}, \vec{z}} z^T \mathbf{1} : \quad z_i \geq (Ax - y)_i, \quad z_i \geq -(Ax - y)_i, \quad i = 1, \dots, m. \quad (4)$$

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

$$d^* = \max_{\vec{u}} -\vec{u}^\top \vec{y} : \quad A^\top \vec{u} = 0, \quad \|\vec{u}\|_\infty \leq 1.$$

*Hint:* use the fact that, for any vector  $z$ :

$$\max_{u : \|u\|_1 \leq 1} u^T z = \|z\|_\infty, \quad \max_{u : \|u\|_\infty \leq 1} u^T z = \|z\|_1.$$

**Solution:**

Based on the hint, we use the Lagrangian

$$\mathcal{L}(\vec{x}, \vec{u}) = \vec{u}^\top (A\vec{x} - \vec{y}),$$

which is such that

$$p^* = \min_{\vec{x}} \max_{\vec{u}} \{ \mathcal{L}(\vec{x}, \vec{u}) : \|\vec{u}\|_\infty \leq 1 \}. \quad (5)$$

Exchanging min and max leads to the dual;

$$p^* \geq d^* = \max_{\vec{u}} g(\vec{u}),$$

with  $g$  the dual function

$$g(\vec{u}) = \min_{\vec{x}} \mathcal{L}(\vec{x}, \vec{u}) = \begin{cases} -\vec{u}^\top \vec{y} & \text{if } A^\top \vec{u} = 0, \\ -\infty & \text{otherwise.} \end{cases}$$

The dual problem finally writes

$$d^* = \max_{\vec{u}} -\vec{u}^\top \vec{y} : A^\top \vec{u} = 0, \quad \|\vec{u}\|_\infty \leq 1.$$

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

$$p^* = \min_x \|Ax - y\|_1 + \mu \|x\|_\infty.$$

(c) Express the problem as an LP.

**Solution:** The problem writes

$$\min_{x, z, t} z^T \mathbf{1} + \mu t : t \geq \|x\|_\infty, \quad z_i \geq |(Ax - y)_i|, \quad i = 1, \dots, m.$$

which is an LP:

$$\begin{aligned} \min_{x, z, t} z^T \mathbf{1} + \mu t : \quad & t \geq x_j, \quad t \geq -x_j, \quad j = 1, \dots, n \\ & z_i \geq (Ax - y)_i, \quad z_i \geq -(Ax - y)_i, \quad i = 1, \dots, m. \end{aligned} \tag{6}$$

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

$$d^* = \max_u -u^T y : \|u\|_\infty \leq 1, \quad \|A^T u\|_1 \leq \mu.$$

*Hint:* use the fact that, for any vector  $z$ :

$$\max_{u : \|u\|_1 \leq 1} u^T z = \|z\|_\infty, \quad \max_{u : \|u\|_\infty \leq 1} u^T z = \|z\|_1.$$

**Solution:**

Based on the hint, we use the Lagrangian

$$\mathcal{L}(x, u, v) = u^T (Ax - y) + v^T x,$$

which is such that

$$p^* = \min_x \max_{u, v} \{\mathcal{L}(x, u, v) : \|u\|_\infty \leq 1, \quad \|v\|_1 \leq \mu\}. \tag{7}$$

Exchanging min and max leads to the dual;

$$p^* \geq d^* = \max_{u, v} g(u, v),$$

with  $g$  the dual function

$$g(u, v) = \min_x \mathcal{L}(x, u, v) = \begin{cases} -u^T y & \text{if } A^T u + v = 0, \\ -\infty & \text{otherwise.} \end{cases}$$

The dual problem writes

$$d^* = \max_u -u^T y : A^T u + v = 0, \|u\|_\infty \leq 1, \|v\|_1 \leq \mu.$$

We can eliminate  $v$ :

$$d^* = \max_u -u^T y : \|u\|_\infty \leq 1, \|A^T u\|_1 \leq \mu.$$

#### 4. Fenchel conjugate

Given a function  $f : \mathbb{R} \rightarrow \mathbb{R}$ , its Fenchel conjugate  $f^* : \mathbb{R} \rightarrow \mathbb{R} \cup \{\pm\infty\}$  is defined as

$$f^*(x) = \sup_{y \in \mathbb{R}} \{\langle y, x \rangle - f(y)\}.$$

Note that this transform is always well-defined when  $f(x)$  is convex.

- (a) Show that for every function  $f$ , its Fenchel conjugate  $f^*$  is convex.

**Solution:** Note that  $f^*(x)$  is the pointwise maximum of the functions  $\langle y, x \rangle - f(y)$ , each of which is an affine function of  $x$ . Consequently,  $f^*$  is convex.

- (b) Given  $f(x) = (x - 2)^2$ , find its Fenchel conjugate  $f^*$ .

**Solution:**

$$f^*(x) = \sup_y xy - (y - 2)^2 = x \left( \frac{x}{2} + 2 \right) - \left( \frac{x}{2} \right)^2 = \frac{x^2}{4} + 2x.$$

- (c) Given  $f(x) = (x - 2)^2$ , find the Fenchel conjugate of its Fenchel conjugate,  $(f^*)^*(x)$ .

**Solution:**

$$\begin{aligned} (f^*)^*(x) &= \sup_y xy - f^*(y) \\ &= \sup_y xy - \frac{y^2}{4} - 2y \\ &= x(2x - 4) - (x - 2)^2 - 4(x - 2) \\ &= (x - 2)^2 = f(x) \end{aligned}$$

*The next two parts of this question are optional and you will not be required to submit a solution. However, understanding them and their solutions may be useful for the final project.*

- (d) **(Optional)** Consider the following piecewise linear function:

$$f(x) = \begin{cases} -x & x < 0, \\ x & 0 \leq x \leq 0.5, \\ 1 - x & 0.5 \leq x \leq 1, \\ x - 1 & x \geq 1. \end{cases}$$

Verify that  $f^*$  and  $(f^*)^*$  are given by the following:

$$f^*(x) = \begin{cases} \infty & x > 1 \\ x & 0 \leq x \leq 1 \\ 0 & -1 \leq x \leq 0 \\ \infty & x < -1 \end{cases} \quad \text{and} \quad (f^*)^*(x) = \begin{cases} x-1 & x \geq 1 \\ 0 & 0 \leq x \leq 1 \\ -x & x < 0 \end{cases}.$$

*Hint: Using the plot of  $f$  might be helpful.  $f^*(y)$  can be interpreted as the maximum distance between the function  $f$  and the line passing through the origin with slope  $y$ .*

**Solution:** We first verify the definition of  $f^*$ . From the definition of the Fenchel conjugate:

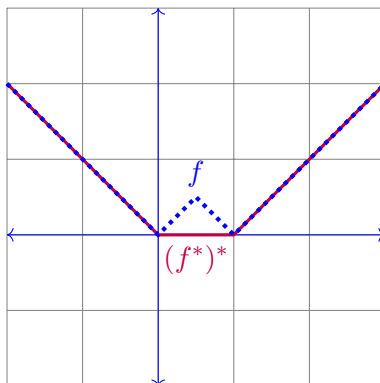
$$f^*(x) = \max_z xz - f(z).$$

When  $x > 1$ , we have  $\lim_{z \rightarrow \infty} xz - f(z) = \infty$  and when  $x < -1$ ,  $\lim_{z \rightarrow -\infty} xz - f(z) = \infty$ . Considering the case  $0 \leq x \leq 1$ ,  $\max_z xz - f(z) = x$  at  $z = 1$  as the function  $xz - f(z) \leq 0$  for all  $z \leq 0$  and  $g(z) = xz - f(z)$  is non-increasing in  $[0, 0.5] \cup [1, \infty)$  and increasing in the interval  $[0.5, 1]$ . Finally, when  $-1 \leq x \leq 0$ , the function  $g(z) = xz - f(z)$  is increasing in the interval  $(-\infty, 0]$  and decreasing in the interval  $[0, \infty)$ . Therefore,  $\max_z xz - f(z) = 0$  when  $x \in [-1, 0]$ .

Now, consider  $(f^*)^*(x) = \max_z xz - f^*(z)$ . For any  $x$ ,  $g(z) = xz - f^*(z)$  is maximized in the interval  $[-1, 1]$  as  $f^*(z) = \infty$  when  $z \in (-\infty, -1) \cup (1, \infty)$ . Since,  $f^*$  is piecewise linear in the interval  $[-1, 1]$ ,  $g(z)$  is maximized at one of the points  $\{-1, 0, 1\}$ . When  $x > 1$ ,  $g(z)$  is maximized at  $z = 1$ , when  $0 \leq x \leq 1$ ,  $z = 0$  is the maximizer and finally when  $x < 0$ ,  $g(z)$  is maximized at  $z = -1$ .

- (e) **(Optional)** Compare the epigraph of  $(f^*)^*$  with the convex hull of the epigraph of  $f$ .

**Solution:**



The epigraph of  $(f^*)^*$  is the same as the *convex hull* of the epigraph of  $f$ .

**5. Jupyter Notebook** Coming soon!

**6. Homework process**

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