EECS 127/227AT Optimization Models in Engineering
Spring 2020 Discussion 9

1. Magic with constraints

In this question, we will represent a problem in two different ways and show that strong duality holds in one case but doesn't hold in the other.

Let

$$f_0(x) \doteq \begin{cases} x^3 - 3x^2 + 4, & x \geq 0 \\ -x^3 - 3x^2 + 4, & x < 0 \end{cases}.$$

(a) Consider the minimization problem

$$\begin{aligned} p^* = \inf_{x \in \mathbb{R}} \; & f_0(x) \\ \text{s.t. } & -1 \leq x, \;\; x \leq 1. \end{aligned} \tag{1}$$

i. Show that $p^* = 2$ and the the set of optimizers $x \in \mathcal{X}^*$ is $\mathcal{X}^* = \{-1, 1\}$ by examining the "critical" points, i.e., points where the gradient is zero, points on the boundaries, and $\pm\infty$.

Solution: First, let us establish the critical points at which the derivative of $f_0(x) = 0$. Note that by definition, $f_0(x)$ is differentiable everywhere except possibly at $x = 0$. We first show that $f_0(x)$ is in fact differentiable everywhere by taking the right and left derivatives at $x = 0$ and showing that they are equivalent.

We can calculate these right and left derivatives as follows. For $h > 0$, the right derivative at $x = 0$ is given by

$$\begin{aligned} & \lim_{h \to 0} \frac{f_0(0+h) - f_0(0)}{h} \\ & = \lim_{h \to 0} \frac{h^3 - 3h^2 + 4 - 4}{h} \\ & = 0. \end{aligned}$$

Similarly, for $h > 0$, the left derivative at $x = 0$ is given by

$$\begin{aligned} & \lim_{h \to 0} \frac{f_0(0) - f_0(0-h)}{h} \\ & = \lim_{h \to 0} \frac{4 - h^3 + 3h^2 - 4}{h} \\ & = 0. \end{aligned}$$

Thus, f_0 is differentiable everywhere, and $x = 0$ is a critical point since its derivative is zero.

Next, we calculate all critical points at which the derivative is zero:

$$\nabla_x f_0(x) = \begin{cases} 3x^2 - 6x, & x \geq 0 \\ -3x^2 - 6x, & x < 0 \end{cases} = 0 \Rightarrow x \in \{0, \pm 2\}$$

We now have a list of all critical points to test: $x \in \{0, \pm 2\}$ (where the derivative is 0), $x = \pm 1$ (constraint boundaries), and $x = \pm\infty$. The only critical points that fall within our constraints are $x \in \{0, \pm 1\}$, so we examine the function at these 3 points:

$$f_0(1) = f_0(-1) = 2$$
$$f_0(0) = 4.$$

Thus, $p^* = 2$ and $\mathcal{X}^* = \{-1, 1\}$.

Teaching Notes: Skip most of this explanation for time and just point out the critical points on a plot. Note that even though we often solve problems in this class via the dual, it is often possible to reason about the primal problem and solve it directly.

ii. Show that the dual problem can be represented as

$$d^* = \sup_{\lambda_1, \lambda_2 \geq 0} g(\vec{\lambda}),$$

where

$$g(\vec{\lambda}) = \min\left\{g_1(\vec{\lambda}), g_2(\vec{\lambda})\right\},$$

with

$$g_1(\vec{\lambda}) = \inf_{x \geq 0} x^3 - 3x^2 + 4 - \lambda_1(x+1) + \lambda_2(x-1)$$
$$g_2(\vec{\lambda}) = \inf_{x < 0} -x^3 - 3x^2 + 4 - \lambda_1(x+1) + \lambda_2(x-1).$$

Solution: The Lagrangian is given by

$$\mathcal{L}(x, \vec{\lambda}) = f_0(x) + \lambda_1(-x-1) + \lambda_2(x-1).$$

The dual function $g(\vec{\lambda})$ is then given by

$$\begin{aligned} g(\vec{\lambda}) &= \inf_x \mathcal{L}(x, \vec{\lambda}) \\ &= \min\left\{\inf_{x \geq 0} \mathcal{L}(x, \vec{\lambda}),\ \inf_{x<0} \mathcal{L}(x, \vec{\lambda})\right\} \\ &= \min\left\{g_1(\vec{\lambda}), g_2(\vec{\lambda})\right\} \end{aligned}$$

for the given $g_1(\vec{\lambda})$ and $g_2(\vec{\lambda})$, as desired.

iii. Next, show that

$$g_1(\vec{\lambda}) \leq -3\lambda_1 + \lambda_2$$
$$g_2(\vec{\lambda}) \leq \lambda_1 - 3\lambda_2.$$

Use this to show that $g(\vec{\lambda}) \leq 0$ for all $\lambda_1, \lambda_2 \geq 0$.

Solution: Because $g_1(\vec{\lambda})$ is the infimum over all $x \geq 0$ of $\mathcal{L}(x, \vec{\lambda})$, it is less than or equal to any instantiation of $\mathcal{L}(x, \vec{\lambda})$ at a particular value of $x \geq 0$. Thus, for instantiation $x = 2$, we can write

$$\begin{aligned} g_1(\vec{\lambda}) &= \inf_{x \geq 0} \mathcal{L}(x, \vec{\lambda}) \\ &\leq \mathcal{L}(2, \vec{\lambda}) \\ &= \lambda_1 - 3\lambda_2 \end{aligned}$$

as desired. Analogously, we can instantiate $g_2(\vec{\lambda})$ at $x = -2$ and write

$$\begin{aligned} g_2(\vec{\lambda}) &= \inf_{x<0} \mathcal{L}(x, \vec{\lambda}) \\ &\leq \mathcal{L}(-2, \vec{\lambda}) \\ &= \lambda_1 - 3\lambda_2, \end{aligned}$$

giving us the two desired inequalities.

We now use these inequalities to show that $g(\vec{\lambda}) \leq 0$ for all $\lambda_1, \lambda_2 \geq 0$. Since $g(\vec{\lambda})$ is the minimization of $g_1(\vec{\lambda})$ and $g_2(\vec{\lambda})$, we can use the upper bounds we just established to write

$$\begin{aligned} g(\vec{\lambda}) &= \min\left\{g_1(\vec{\lambda}), g_2(\vec{\lambda})\right\} \\ &\leq \min\left\{-3\lambda_1 + \lambda_2, \lambda_1 - 3\lambda_2\right\} \\ &\leq 0. \end{aligned}$$

The last inequality follows from a subtle relationship between the two expressions over which we are minimizing. First, note that it is sufficient to show that either $-3\lambda_1 + \lambda_2$ or $\lambda_1 - 3\lambda_2$ must be negative, since $g(\vec{\lambda})$ is determined by the minimum of the two values. Consider the case in which $-3\lambda_1 + \lambda_2 \geq 0$, i.e., $\lambda_2 \geq 3\lambda_1$; this implies that the second expression $\lambda_1 - 3\lambda_2 \leq 0$, so $g(\vec{\lambda}) \leq 0$ holds. Alternatively, if $\lambda_1 - 3\lambda_2 > 0$, i.e., $\lambda_1 > 3\lambda_2$, then the first expression $-3\lambda_1 + \lambda_2 < 0$, so $g(\vec{\lambda}) \leq 0$ holds. Thus, as these cases are exhaustive, $g(\vec{\lambda}) \leq 0$ for all $\lambda_1, \lambda_2 \geq 0$ as desired.

iv. Show that $g(\vec{0}) = 0$ and conclude that $d^* = 0$.

Solution: In part (a)iii., we proved that $g(\vec{\lambda}) \leq 0$ for all $\lambda_1, \lambda_2 \geq 0$. Since d^* is the supremum over all feasible values of $g(\vec{\lambda})$, it is sufficient to show that there exists a $\vec{\lambda}$ for which this upper bound is attained.

Toward this objective, consider g at $\vec{\lambda} = 0$:

$$\begin{aligned} g(\vec{0}) &= \min\left\{g_1(\vec{0}), g_2(\vec{0})\right\} \\ &= \min\left\{\inf_{x\geq 0} x^3 - 3x^2 + 4,\ \inf_{x<0} -x^3 - 3x^2 + 4\right\} \\ &= \min\{0, 0\} \\ &= 0. \end{aligned}$$

Note that the third equality can be shown by examining the critical points of each objective function, which are the same as those of the unconstrained primal function in part (a)i.; this minimum is achieved at $x = \pm 2$.

We can now conclude that the maximum possible value of the dual (i.e., zero) is attained for $\vec{\lambda} = \vec{0}$, and thus $d^* = 0$ as desired.

v. Does strong duality hold?

Solution: Since $d^* = 0 < 2 = p^*$, strong duality does not hold. This is not surprising, since the objective function $f_0(x)$ is non-convex.

(b) Now, consider a problem equivalent to the minimization in (1):

$$\begin{aligned} p^* = \inf_{x\in\mathbb{R}} \ & f_0(x) \\ \text{s.t.}\ & x^2 \leq 1. \end{aligned} \tag{2}$$

Observe that $p^* = 2$ and the set of optimizers $x \in \mathcal{X}^*$ is $\mathcal{X}^* = \{-1, 1\}$, since this problem is equivalent to the one in part (a).

i. Show that the dual problem can be represented as

$$d^* = \sup_{\lambda \geq 0} g(\lambda),$$

where

$$g(\lambda) = \min(g_1(\lambda), g_2(\lambda)),$$

with

$$\begin{aligned} g_1(\lambda) &= \inf_{x \geq 0} x^3 - 3x^2 + 4 + \lambda(x^2 - 1) \\ g_2(\lambda) &= \inf_{x < 0} -x^3 - 3x^2 + 4 + \lambda(x^2 - 1). \end{aligned}$$

Solution: This solution is identical in strategy to that in part (a)ii. The Lagrangian is given by

$$\mathcal{L}(x, \lambda) = f_0(x) + \lambda(x^2 - 1).$$

The dual function $g(\lambda)$ is then given by

$$\begin{aligned} g(\lambda) &= \inf_x \mathcal{L}(x, \lambda) \\ &= \min \left\{ \inf_{x \geq 0} \mathcal{L}(x, \lambda), \inf_{x < 0} \mathcal{L}(x, \lambda) \right\} \\ &= \min \{g_1(\lambda), g_2(\lambda)\} \end{aligned}$$

for the given $g_1(\lambda)$ and $g_2(\lambda)$, as desired.

ii. Show that $g_1(\lambda) = g_2(\lambda) = \begin{cases} 4 - \lambda, & \lambda \geq 3 \\ -\frac{4}{27}(3 - \lambda)^3 + 4 - \lambda, & 0 \leq \lambda < 3. \end{cases}$

Solution: We first show that $g_2(\vec{\lambda}) = g_1(\vec{\lambda})$:

$$\begin{aligned} g_2(\lambda) &= \inf_{x < 0} -x^3 - 3x^2 + 4 + \lambda(x^2 - 1) \\ &= \inf_{-x > 0} (-x)^3 - 3(-x)^2 + 4 + \lambda((-x)^2 - 1) \\ &= \inf_{x \geq 0} x^3 - 3x^2 + 4 + \lambda(x^2 - 1) \\ &= g_1(\lambda). \end{aligned}$$

The last equality follows from a change in the variable over which we compute the infimum ($-x$ to x), which does not affect the value of the infimum. Note also that we have added the point $x = 0$ as a feasible point by amending our constraint from $-x > 0$ to $-x \geq 0$; this does not affect the value of the infimum either, since we do not require it to be attained as we do when minimizing.

Next, let us compute $g_1(\lambda)$ directly. Setting the derivative of g_1's objective function with respect to x to zero, we have

$$3x^2 - 2(3 - \lambda)x = 0 \implies x = 0 \text{ or } x = \frac{2}{3}(3 - \lambda).$$

We now consider all critical points of g_1's objective function: $x \in \{0, \frac{2}{3}(3-\lambda\}$ (where the derivative is 0) and $x \in \{0, \infty\}$ (boundary points).

First, suppose $\lambda \geq 3$. In this case, $x = \frac{2}{3}(3-\lambda)$ is no longer in the range $x \geq 0$, so we need only check boundary points $x = 0$ and $x = \infty$. As $x \to \infty$, the function value also approaches infinity, so the infimum is attained at $x = 0$, and thus $g_1(\lambda) = 4 - \lambda$.

Next, assume $0 \leq \lambda < 3$. In this case, we must check the function value at $x = 0$, $x = \frac{2}{3}(3-\lambda)$, and $x = \infty$ to determine where the infimum is attained. As previously established, the function approaches infinity as $x \to \infty$, so we need only compare the values $4 - \lambda$ (at $x = 0$) and $-\frac{4}{27}(3-\lambda)^3 + 4 - \lambda$ at $x = \frac{2}{3}(3-\lambda)$. Since $3 - \lambda > 0$, we know that $-\frac{4}{27}(3-\lambda)^3$ is always negative, and thus the infimum is $-\frac{4}{27}(3-\lambda)^3 + 4 - \lambda$.

Combining the two cases above yields the desired expression for $g_1(\lambda) = g_2(\lambda)$.

iii. Conclude that $d^* = 2$ and the optimal $\lambda = \frac{3}{2}$.

Solution: Since $g_1(\lambda) = g_2(\lambda)$, we have $g(\lambda) = g_1(\lambda) = g_2(\lambda)$. We examine each range of possible λ values in turn to determine the supremum.

For $\lambda \geq 3$, the supremum value of $g(\lambda) = 1$ is achieved at $\lambda = 3$.

For $0 \leq \lambda < 3$, the supremum of $g(\lambda)$ is computed as follows. First, we set the derivative of $g(\lambda)$ with respect to λ to 0:

$$\frac{12}{27}(3-\lambda)^2 - 1 = 0 \implies (3-\lambda)^2 = \frac{9}{4} \implies \lambda = \frac{3}{2} \text{ or } \lambda = \frac{9}{2}.$$

Since the expression is valid only for $0 \leq \lambda < 3$, we examine values at $\lambda \in \{0, 3\}$ (boundary points) and at the computed $\lambda = \frac{3}{2}$. We observe that the supremum is achieved at $\lambda = \frac{3}{2}$ with $g(\frac{3}{2}) = 2$.

Finally, we note that the overall supremum occurs in the second case, at $\lambda = \frac{3}{2}$, and thus $d^* = 2$ as desired.

iv. Does strong duality hold?

Solution: In this case, $p^* = 2 = d^*$, so strong duality holds.

Teaching Notes: This question shows that if the problem is non-convex, the strategy of forming a Langrange dual and solving the dual problem in order to find the optimal value of the primal problem may or may not work depending on the form of constraints we consider while forming the dual. For some form of the constraints, strong duality may hold, while for others strong duality may not hold.

2. Linear programming

Express the following problems as LPs.

(a)

$$\begin{aligned} \min_{\vec{x} \in \mathbb{R}^k} & \left[\max_{i=1,\ldots,k} x_i - \min_{j=1,\ldots,k} x_j \right] \\ \text{s.t.} \quad & A\vec{x} = \vec{b} \end{aligned}$$

Solution: We can reformulate this problem by introducing slack variables t and u as follows:

$$\begin{aligned} \min_{\vec{x}, t, u} \; & t - u \\ \text{s.t.} \quad & A\vec{x} = \vec{b} \\ & t \geq x_i, \; i = 1, \ldots, k \\ & u \leq x_i, \; i = 1, \ldots, k \end{aligned}$$

where $\vec{x} \in \mathbb{R}^k$, $t \in \mathbb{R}$, and $u \in \mathbb{R}$. Note that both formulations are equivalent to finding the vector $\vec{x}$ with the minimum difference between the highest and lowest entries, subject to dynamics $A\vec{x} = \vec{b}$; slack variables t and u form upper and lower bounds on these values.

Teaching Notes: Draw a picture of lines t and u bounding the x_i values.

(b)

$$\begin{aligned} &\min_{\vec{x} \in \mathbb{R}^k} \sum_{i=1}^{k} |x_i| \\ &\text{s.t.} \quad A\vec{x} = \vec{b} \end{aligned}$$

Solution: We again solve this problem using slack variables. First, consider the simple, related scalar minimization

$$\min_{x \in \mathbb{R}} |x|.$$

We can reformulate this problem using slack variable t as follows:

$$\begin{aligned} &\min_{x,t \in \mathbb{R}} \; t \\ &\text{s.t.} \quad |x| \le t \end{aligned}$$

and further linearize the constraint to get LP

$$\begin{aligned} &\min_{x,t \in \mathbb{R}} \; t \\ &\text{s.t.} \quad x \le t, \; -x \le t. \end{aligned}$$

You can read more about this absolute value constraint trick **here**.

Returning to our original minimization, we can replace each absolute value in the summation with a slack variable, resulting in final LP

$$\begin{aligned} &\min_{\vec{x}, \vec{t} \in \mathbb{R}^k} \sum_{i=1}^{k} t_i \\ &\text{s.t.} \quad A\vec{x} = \vec{b} \\ &\qquad\; x_i \le t_i, \; -x \le t_i, \; i = 1, \ldots, k \end{aligned}$$

as desired.

Teaching Notes: Note that this trick is nice/useful because the problem we're solving is convex but not differentiable. If students are confused, draw the absolute value constraint $|x| \le t$ on a number line and show graphically that it's equivalent to the two linear constraints. Note also that this trick doesn't work when you're maximizing over absolute values — it's not a magic bullet to turn nonconvex problems into convex ones! The graphical representation might be helpful in conveying this too.