Discussion 10

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

$$f_0(x) \doteq \begin{cases} x^3 - 3x^2 + 4, & x \geq 0 \\ -x^3 - 3x^2 + 4, & x < 0. \end{cases} \tag{1}$$

(a) Consider the minimization problem

$$p^* = \inf_{x \in \mathbb{R}} \quad f_0(x) \tag{2}$$

$$\text{s.t.} \quad -1 \leq x, \quad x \leq 1. \tag{3}$$

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

$$\lim_{h \to 0} \frac{f_0(0 + h) - f_0(0)}{h} = \lim_{h \to 0} \frac{h^3 - 3h^2 + 4 - 4}{h} \tag{4}$$

$$= 0. \tag{5}$$

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

$$\lim_{h \to 0} \frac{f_0(0) - f_0(0 - h)}{h} = \lim_{h \to 0} \frac{4 - h^3 + 3h^2 - 4}{h} \tag{6}$$

$$= 0. \tag{7}$$

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:

$$\nabla_x f_0(x) = \begin{cases} 3x^2 - 6x, & x \geq 0 \\ -3x^2 - 6x, & x < 0 \end{cases} = 0 \Rightarrow x \in \{0, \pm 2\} \tag{8}$$

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:

$$f_0(1) = f_0(-1) = 2 \tag{9}$$

$$f_0(0) = 4. \tag{10}$$

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

ii. Show that the dual problem can be represented as

$$d^* = \sup_{\lambda_1, \lambda_2 \geq 0} g(\vec{\lambda}), \tag{11}$$

where

$$g(\vec{\lambda}) = \min\left\{g_1(\vec{\lambda}), g_2(\vec{\lambda})\right\}, \tag{12}$$

with

$$g_1(\vec{\lambda}) = \inf_{x \geq 0} x^3 - 3x^2 + 4 - \lambda_1(x+1) + \lambda_2(x-1) \tag{13}$$

$$g_2(\vec{\lambda}) = \inf_{x < 0} -x^3 - 3x^2 + 4 - \lambda_1(x+1) + \lambda_2(x-1). \tag{14}$$

Solution: The Lagrangian is given by

$$\mathcal{L}(x, \vec{\lambda}) = f_0(x) + \lambda_1(-x-1) + \lambda_2(x-1). \tag{15}$$

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

$$g(\vec{\lambda}) = \inf_x \mathcal{L}(x, \vec{\lambda}) \tag{16}$$

$$= \min\left\{\inf_{x \geq 0} \mathcal{L}(x, \vec{\lambda}), \inf_{x < 0} \mathcal{L}(x, \vec{\lambda})\right\} \tag{17}$$

$$= \min\left\{g_1(\vec{\lambda}), g_2(\vec{\lambda})\right\} \tag{18}$$

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

iii. Next, show that

$$g_1(\vec{\lambda}) \leq -3\lambda_1 + \lambda_2 \tag{19}$$

$$g_2(\vec{\lambda}) \leq \lambda_1 - 3\lambda_2. \tag{20}$$

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

$$g_1(\vec{\lambda}) = \inf_{x \geq 0} \mathcal{L}(x, \vec{\lambda}) \tag{21}$$

$$\leq \mathcal{L}(2, \vec{\lambda}) \tag{22}$$

$$= -3\lambda_1 + \lambda_2 \tag{23}$$

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

$$g_2(\vec{\lambda}) = \inf_{x < 0} \mathcal{L}(x, \vec{\lambda}) \tag{24}$$

$$\leq \mathcal{L}(-2, \vec{\lambda}) \tag{25}$$

$$= \lambda_1 - 3\lambda_2, \tag{26}$$

giving us the two desired inequalities.

We now use these inequalities to show that $g(\vec{\lambda}) \le 0$ for all $\lambda_1, \lambda_2 \ge 0$. 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

$$g(\vec{\lambda}) = \min\left\{g_1(\vec{\lambda}), g_2(\vec{\lambda})\right\} \tag{27}$$
$$\le \min\left\{-3\lambda_1 + \lambda_2, \lambda_1 - 3\lambda_2\right\} \tag{28}$$
$$\le 0. \tag{29}$$

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 $-3\lambda_1 + \lambda_2 \ge 0$, i.e., $\lambda_2 \ge 3\lambda_1$; this implies that the second expression $\lambda_1 - 3\lambda_2 \le 0$, so $g(\vec{\lambda}) \le 0$ 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 $g(\vec{\lambda}) \le 0$ holds. Thus, as these cases are exhaustive, $g(\vec{\lambda}) \le 0$ for all $\lambda_1, \lambda_2 \ge 0$ as desired.

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

Solution: In part **1.**((a))iii, we proved that $g(\vec{\lambda}) \le 0$ for all $\lambda_1, \lambda_2 \ge 0$. 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 $\vec{\lambda} = 0$:

$$g(\vec{0}) = \min\left\{g_1(\vec{0}), g_2(\vec{0})\right\} \tag{30}$$
$$= \min\left\{\inf_{x \ge 0} x^3 - 3x^2 + 4,\ \inf_{x<0} -x^3 - 3x^2 + 4\right\} \tag{31}$$
$$= \min\{0, 0\} \tag{32}$$
$$= 0. \tag{33}$$

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 **1.**((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 (2):

$$p^* = \inf_{x \in \mathbb{R}} \quad f_0(x) \tag{34}$$
$$\text{s.t.} \quad x^2 \le 1. \tag{35}$$

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

$$d^* = \sup_{\lambda \geq 0} g(\lambda), \tag{36}$$

where

$$g(\lambda) = \min\{g_1(\lambda), g_2(\lambda)\}, \tag{37}$$

with

$$g_1(\lambda) = \inf_{x \geq 0} x^3 - 3x^2 + 4 + \lambda(x^2 - 1) \tag{38}$$

$$g_2(\lambda) = \inf_{x < 0} -x^3 - 3x^2 + 4 + \lambda(x^2 - 1). \tag{39}$$

Solution: This solution is identical in strategy to that in part **1.**((a))ii The Lagrangian is given by

$$\mathcal{L}(x, \lambda) = f_0(x) + \lambda(x^2 - 1). \tag{40}$$

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

$$\begin{aligned} g(\lambda) &= \inf_x \mathcal{L}(x, \lambda) && (41)\\ &= \min\left\{\inf_{x \geq 0} \mathcal{L}(x, \lambda), \inf_{x < 0} \mathcal{L}(x, \lambda)\right\} && (42)\\ &= \min\{g_1(\lambda), g_2(\lambda)\} && (43) \end{aligned}$$

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 $g_2(\vec{\lambda}) = g_1(\vec{\lambda})$:

$$\begin{aligned} g_2(\lambda) &= \inf_{x < 0} -x^3 - 3x^2 + 4 + \lambda(x^2 - 1) && (44)\\ &= \inf_{-x > 0} (-x)^3 - 3(-x)^2 + 4 + \lambda((-x)^2 - 1) && (45)\\ &= \inf_{x \geq 0} x^3 - 3x^2 + 4 + \lambda(x^2 - 1) && (46)\\ &= g_1(\lambda). && (47) \end{aligned}$$

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

$$3x^2 - 2(3 - \lambda)x = 0 \implies x = 0 \text{ or } x = \frac{2}{3}(3 - \lambda). \tag{48}$$

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:

$$\frac{12}{27}(3-\lambda)^2 - 1 = 0 \implies (3-\lambda)^2 = \frac{9}{4} \implies \lambda = \frac{3}{2} \text{ or } \lambda = \frac{9}{2}. \tag{49}$$

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.

2. Complementary Slackness

Consider the problem:

$$p^\star = \min_{x \in \mathbb{R}} \quad x^2 \tag{50}$$
$$\text{s.t.} \quad x \geq 1, \; x \leq 2. \tag{51}$$

(a) Does Slater's condition hold? Is the problem convex? Does strong duality hold?

Solution: We have a strictly feasible point $x = 1.5$ that lies in the relative interior of the domain of the objective function; thus, Slater's condition holds. The objective function x^2 is convex and the inequality constraints are affine and thus convex, so the problem is convex. Since Slater's condition holds and the problem is convex, strong duality holds.

(b) Find the Lagrangian $\mathcal{L}(x, \lambda_1, \lambda_2)$.

Solution: $\mathcal{L}(x, \lambda_1, \lambda_2) = x^2 + \lambda_1(-x+1) + \lambda_2(x-2)$.

(c) Find the dual function $g(\lambda_1, \lambda_2)$ so that the dual problem is given by,

$$d^\star = \max_{\lambda_1, \lambda_2 \in \mathbb{R}_+} g(\lambda_1, \lambda_2). \tag{52}$$

Solution:

$$g(\lambda_1, \lambda_2) = \inf_x \mathcal{L}(x, \lambda_1, \lambda_2). \tag{53}$$

Note that $\mathcal{L}$ is convex with respect to x, thus setting the gradient with respect to x to 0 we obtain, $x = \frac{\lambda_1 - \lambda_2}{2}$. Thus,

$$g(\lambda_1, \lambda_2) = -\frac{(\lambda_2 - \lambda_1)^2}{4} + \lambda_1 - 2\lambda_2. \tag{54}$$

(d) Solve the dual problem in (52) for $d^\star$.

Solution: Let us first try setting gradient with respect to λ_1 and λ_2 to 0. This gives us,

$$\frac{\lambda_2 - \lambda_1}{2} + 1 = 0 \tag{55}$$

$$-\frac{\lambda_2 - \lambda_1}{2} - 2 = 0. \tag{56}$$

This has no solution. We can see that a quadratic objective function could be unbounded even if it was convex. To get meaningful solutions we must check for optimal values at the boundaries. Checking at boundary $\lambda_1 = 0$.

$$g(0, \lambda_2) = -\frac{\lambda_2^2}{4} - 2\lambda_2. \tag{57}$$

This is a concave function so taking gradient with respect to λ_2 and setting it to zero we obtain,

$$-\frac{\lambda_2}{2} - 2 = 0 \tag{58}$$

$$\implies \lambda_2 = -4. \tag{59}$$

This is not feasible so we must check value at $\lambda_2 = 0$. We have $g(0, 0) = 0$. Finally let us check at the other boundary $\lambda_2 = 0$.

$$g(\lambda_1, 0) = -\frac{\lambda_1^2}{4} + \lambda_1. \tag{60}$$

Again this is a concave function so taking gradient with respect to λ_1 and setting it to zero we obtain,

$$-\frac{\lambda_1}{2} + 1 = 0 \tag{61}$$

$$\implies \lambda_1 = 2. \tag{62}$$

We have $g(2, 0) = -1 + 2 = 1$. Thus $d^\star = 1$.

(e) Solve for $x^\star, \lambda_1^\star, \lambda_2^\star$ that satisfy KKT conditions.

Solution: We have: From stationarity,

$$\nabla_x \mathcal{L}(x, \lambda_1, \lambda_2) = 0 \tag{63}$$

$$\implies 2x - \lambda_1 + \lambda_2 = 0. \tag{64}$$

From primal feasibility,

$$x \geq 1 \tag{65}$$

$$x \leq 2. \tag{66}$$

From dual feasibility,

$$\lambda_1 \geq 0 \tag{67}$$

$$\lambda_2 \geq 0. \tag{68}$$

Finally from complementary slackness,

$$\lambda_1(-x+1) = 0 \tag{69}$$

$$\lambda_2(x-2) = 0. \tag{70}$$

First observe that we cannot have $\lambda_1 \neq 0$ and $\lambda_2 \neq 0$ since in this case complementary slackness would not have any feasible solutions for x. Next assume that $\lambda_1 = 0, \lambda_2 \neq 0$. Then from complementary slackness, $x = 2$. Substituting this in equation 64, we get $\lambda_2 = -4$ which violates dual feasibility. Next assume that $\lambda_1 = 0, \lambda_2 = 0$. Then from 64 we have $x = 0$ which violates primal feasibility. Finally assume that $\lambda_1 \neq 0, \lambda_2 = 0$. From complementary slackness we have $x = 1$ and from 64 we have $\lambda_1 = 2$ which satisfies dual feasibility.

Thus $x^\star = 1, \lambda_1^\star = 2, \lambda_2^\star = 0$ satisfy KKT conditions.

(f) Can you spot a connection between the values of $\lambda_1^\star, \lambda_2^\star$ in relation to whether the corresponding inequality constraints are strict or not at the optimal $x^\star$?

Solution: We have $\lambda_1 \neq 0$ and the corresponding inequality $x \geq 1$ is satisfied with equality (and hence is not strict) at $x^\star = 1$.

We have $\lambda_2 = 0$ and the corresponding inequality is strict at $x^\star = 1$. The non-zero λ_1 tells us that if we relax the constraint $x \geq 1$ (for example, to $x \geq 0.9$) we can reduce the objective function further.

