Discussion 11

1. An optimization problem

Consider the primal optimization problem,

$$p^\star = \min_{\vec{x} \in \mathbb{R}^2} \frac{1}{2}x_1^2 + \frac{1}{2}x_2^2 \tag{1}$$

$$\text{s.t.} \quad x_1 \geq 0, \tag{2}$$

$$x_1 + x_2 \geq 2. \tag{3}$$

First we solve the primal problem directly.

(a) Sketch the feasible region and argue that $\vec{x}^\star = (1, 1)$ and $p^\star = 1$.

Solution:

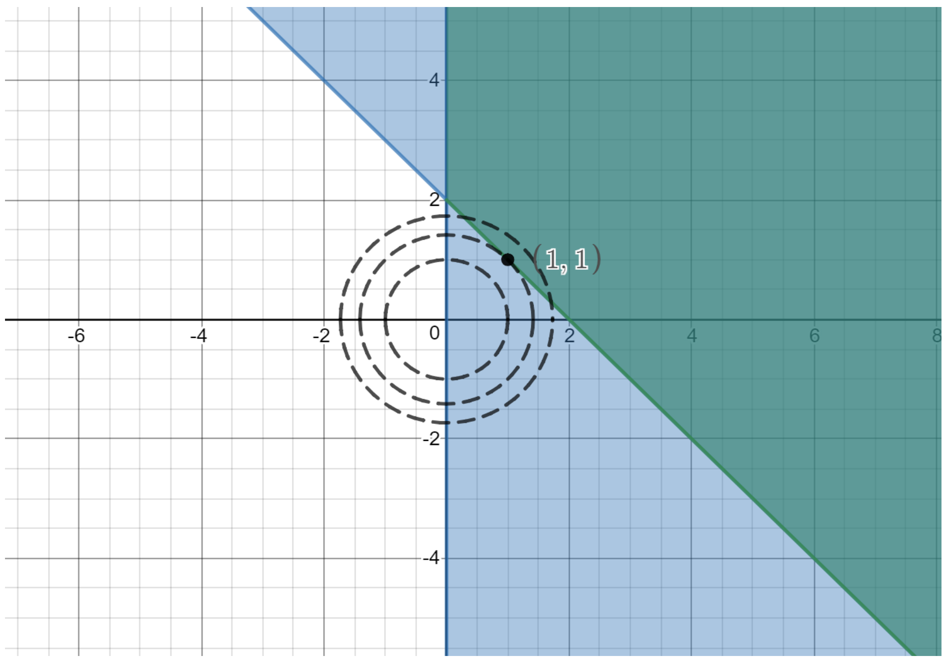

Figure 1

The feasible region is shown in Figure 1 as the green shaded region.

The level curves for the function are concentric circles centered around the origin and the objective value increases as we go away from the origin. Thus the optimal point is the point in the feasible region that is closest to the origin. This point is $\vec{x}^\star = (1, 1)$. Thus, $p^\star = 1$.

(b) Next we solve the problem with the help of the dual. First, find the Lagrangian $\mathcal{L}(\vec{x}, \vec{\lambda})$.

Solution: The Lagrangian $\mathcal{L} : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ is given by

$$\mathcal{L}(\vec{x}, \vec{\lambda}) = \frac{1}{2}x_1^2 + \frac{1}{2}x_2^2 + \lambda_1(-x_1) + \lambda_2(-x_1 - x_2 + 2).$$

(c) Formulate the dual problem.

Solution: Define the dual objective function $g : \mathbb{R}^2 \to \mathbb{R}$ by:

$$g(\vec{\lambda}) := \min_{\vec{x} \in \mathbb{R}^2} \mathcal{L}(\vec{x}, \vec{\lambda}).$$

For each fixed $\vec{\lambda} \in \mathbb{R}^2$, let $\vec{x}^\star(\vec{\lambda})$ denote the corresponding minimizer of $\mathcal{L}(\vec{x}, \vec{\lambda})$. Then

$$0 = \nabla \mathcal{L}(\vec{x}^\star(\vec{\lambda}), \vec{\lambda}) = \begin{bmatrix} x_1^\star(\vec{\lambda}) - \lambda_1 - \lambda_2 \\ x_2^\star(\vec{\lambda}) - \lambda_2 \end{bmatrix},$$

so $\vec{x}^\star(\vec{\lambda}) = (\lambda_1 + \lambda_2, \lambda_2)$. The dual objective function is thus given by

$$\begin{aligned} g(\vec{\lambda}) &= \min_{\vec{x} \in \mathbb{R}^2} \mathcal{L}(\vec{x}, \vec{\lambda}) \\ &= \mathcal{L}(\vec{x}^\star(\vec{\lambda}), \vec{\lambda}) \\ &= \frac{1}{2}(\lambda_1 + \lambda_2)^2 + \frac{1}{2}\lambda_2^2 + \lambda_1(-\lambda_1 - \lambda_2) + \lambda_2(-\lambda_1 - 2\lambda_2 + 2) \\ &= -\frac{1}{2}\lambda_1^2 - \lambda_1\lambda_2 - \lambda_2^2 + 2\lambda_2. \end{aligned} \tag{4}$$

The dual problem is given by

$$\begin{aligned} d^\star = \max_{\vec{\lambda} \in \mathbb{R}^2} \quad & g(\vec{\lambda}), \\ \text{s.t.} \quad & \lambda_1 \geq 0, \\ & \lambda_2 \geq 0. \end{aligned}$$

(d) Solve the dual problem to find $\vec{\lambda}^\star$ and $d^\star$.

Solution: First, we show that $\{g(\vec{\lambda}) | \vec{\lambda} \in \mathbb{R}^2_+\}$ has a finite upper bound. From (4), for each $\vec{\lambda} \in \mathbb{R}^2_+$:

$$\begin{aligned} g(\vec{\lambda}) &= -\frac{1}{2}\lambda_1^2 - \lambda_1\lambda_2 - \lambda_2^2 + 2\lambda_2 \\ &\leq -\lambda_2^2 + 2\lambda_2 \\ &= -(\lambda_2 - 1)^2 + 1 \\ &\leq 1. \end{aligned}$$

Note that equality is attained if and only if $\vec{\lambda} = (0, 1)$. Thus, we have $\vec{\lambda}^\star = (0, 1)$ and $d^\star = g(\vec{\lambda}^\star) = 1$.

(e) Does strong duality hold?

Solution: Yes, since $p^\star = d^\star = 1$; also, note that the primal optimization problem is convex and Slater's condition holds. To show that Slater's condition holds it suffices to exhibit a point in the interior of the feasible set. For instance, $(x_1, x_2) = (2, 2)$ works.

(f) Finally we use the KKT conditions to find $\vec{x}^\star, \vec{\lambda}^\star$. First we write down the KKT conditions and then we find $\tilde{\vec{x}}$ and $\tilde{\vec{\lambda}}$ that satisfy them.

Solution: The KKT conditions are

i. (Lagrangian Stationarity) $\nabla_x \mathcal{L}(\vec{x}, \vec{\lambda}) = 0 \Rightarrow \vec{x} + (-\lambda_1, 0)^\top + (-\lambda_2, -\lambda_2)^\top = 0$.

ii. (Dual Feasibility) $\lambda_1, \lambda_2 \geq 0$.

iii. (Primal Feasibility) $x_1 \geq 0$ and $x_1 + x_2 \geq 2$.

iv. (Complementary Slackness) $\lambda_1 x_1 = 0$ and $\lambda_2(2 - x_1 - x_2) = 0$.

From the first condition we have that

$$\tilde{x}_1 = \tilde{\lambda}_1 + \tilde{\lambda}_2, \quad \tilde{x}_2 = \tilde{\lambda}_2.$$

From the fourth condition, this implies that $\tilde{\lambda}_1(\tilde{\lambda}_1 + \tilde{\lambda}_2) = 0$. This, in conjunction with the second condition, implies $\tilde{\lambda}_1 = 0$. It follows that $\tilde{x}_1 = \tilde{x}_2 = \tilde{\lambda}_2 = 1$ since $\tilde{x}_1 = \tilde{x}_2 = \tilde{\lambda}_2 = 0$ does not satisfy the third condition.

(g) Argue why the optimal primal and dual solutions are given by $\vec{x}^\star = \vec{\tilde{x}}$ and $\vec{\lambda}^\star = \vec{\tilde{\lambda}}$.

Solution: The primal problem is convex and satisfies strong duality. Thus, the KKT conditions are both necessary and sufficient for optimality of the proposed primal point and dual point.

2. Complementary Slackness

Consider the problem:

$$p^\star = \min_{x \in \mathbb{R}} \quad x^2 \tag{5}$$

$$\text{s.t.} \quad x \geq 1,\ x \leq 2. \tag{6}$$

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

Solution: We have a strictly feasible point, e.g., $x = 1.5$, that lies in the relative interior of the domain of the problem; 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,

$$d^\star = \max_{\lambda_1, \lambda_2 \in \mathbb{R}_+} g(\lambda_1, \lambda_2). \tag{7}$$

Solution:

$$g(\lambda_1, \lambda_2) = \inf_x \mathcal{L}(x, \lambda_1, \lambda_2). \tag{8}$$

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,

$$g(\lambda_1, \lambda_2) = -\frac{(\lambda_2 - \lambda_1)^2}{4} + \lambda_1 - 2\lambda_2. \tag{9}$$

(d) Solve the dual problem in (7) for $d^\star$.

Solution: The gradient of g is:

$$\nabla g(\vec{\lambda}) = \begin{bmatrix} -\frac{1}{2}\lambda_1 + \frac{1}{2}\lambda_2 + 1 \\ \frac{1}{2}\lambda_1 - \frac{1}{2}\lambda_2 - 2 \end{bmatrix}.$$

We claim that for any interior point of $\mathbb{R}^2_+$, there exists a boundary point of $\mathbb{R}^2_+$ that attains a strictly larger value of g; in other words, the set of maximizers of g, if non-empty, must be a subset of the boundary of $\mathbb{R}^2_+$. To see this, let $\vec{\lambda} \in \mathbb{R}^2_+$ be an interior point of $\mathbb{R}^2_+$, i.e., $\lambda_1 > 0$, $\lambda_2 > 0$. If $\lambda_1 \geq \lambda_2$, then

$$g(\lambda_1 - \lambda_2, 0) = -\frac{1}{4}(\lambda_2 - \lambda_1)^2 + \lambda_1 - \lambda_2$$

$$\begin{aligned} &= g(\lambda_1, \lambda_2) + \lambda_2 \\ &> g(\lambda_1, \lambda_2). \end{aligned}$$

Otherwise, $\lambda_1 < \lambda_2$, in which case

$$\begin{aligned} g(0, \lambda_2 - \lambda_1) &= -\frac{1}{4}(\lambda_2 - \lambda_1)^2 - 2\lambda_2 + 2\lambda_1 \\ &= g(\lambda_1, \lambda_2) + \lambda_1 \\ &> g(\lambda_1, \lambda_2). \end{aligned}$$

In words, for any given interior point $\vec{\lambda}$ of $\mathbb{R}^2_+$, there exists a boundary point of $\mathbb{R}^2_+$ at which g attains a strictly larger value. Thus, when maximizing g over $\mathbb{R}^2_+$, it suffices to consider the boundary points of $\mathbb{R}^2_+$. We do so below in a piecewise manner. Among all vectors $\vec{\lambda} \in \mathbb{R}^2_+$ satisfying $\lambda_1 = 0$, the vector that maximizes $g(\vec{\lambda}) = g(0, \lambda_2) = -\frac{1}{4}\lambda_2^2 - 2\lambda_2$ is $(0, 0)$, with corresponding function value $g(0, 0) = 0$. Meanwhile, among all vectors $\vec{\lambda} \in \mathbb{R}^2_+$ satisfying $\lambda_2 = 0$, the vector that maximizes $g(\vec{\lambda}) = g(\lambda_1, 0) = -\frac{1}{4}\lambda_1^2 + \lambda_1$ is $(2, 0)$, with corresponding value $g(2, 0) = 1$. We thus conclude that $\lambda^\star = (2, 0)$ and $d^\star = g(\lambda^\star) = 1$.

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

Solution: From Lagrangian stationarity,

$$\nabla_x \mathcal{L}(x, \lambda_1, \lambda_2) = 0 \tag{10}$$

$$\implies 2x - \lambda_1 + \lambda_2 = 0. \tag{11}$$

From primal feasibility,

$$x \geq 1 \tag{12}$$

$$x \leq 2. \tag{13}$$

From dual feasibility,

$$\lambda_1 \geq 0 \tag{14}$$

$$\lambda_2 \geq 0. \tag{15}$$

Finally, from complementary slackness,

$$\lambda_1(-x + 1) = 0 \tag{16}$$

$$\lambda_2(x - 2) = 0. \tag{17}$$

First observe that we cannot have $\lambda_1 \neq 0$ and $\lambda_2 \neq 0$ since in this case complementary slackness lead to not having 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 (11), we get $\lambda_2 = -4$ which violates dual feasibility. Next assume that $\lambda_1 = 0, \lambda_2 = 0$. Then from (11) 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 (11) we have $\lambda_1 = 2$, which satisfies dual feasibility.

Thus $x^\star = 1, \lambda_1^\star = 2, \lambda_2^\star = 0$ satisfy the 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 active 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 this constraint is active) at $x^\star = 1$. We have $\lambda_2 = 0$ and the corresponding inequality is strict (and hence this constraint is inactive) 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. Note that in general, it is possible for an inequality constraint to be active even when the corresponding Lagrange multiplier is 0.