Discussion 10

1. Simple Constrained Optimization Problem with Duality

Consider the optimization problem

$$\begin{aligned} \min_{x_1, x_2 \in \mathbb{R}} \quad & f(x_1, x_2) && (1) \\ \text{s.t.} \quad & 2x_1 + x_2 \geq 1 && (2) \\ & x_1 + 3x_2 \geq 1 && (3) \\ & x_1 \geq 0, && (4) \\ & x_2 \geq 0 && (5) \end{aligned}$$

(a) Express the Lagrangian of the problem $\mathcal{L}(x_1, x_2, \lambda_1, \lambda_2, \lambda_3, \lambda_4)$.

Solution: The Lagrangian is

$$\mathcal{L}(x_1, x_2, \lambda_1, \lambda_2, \lambda_3, \lambda_4) = f(x_1, x_2) + \lambda_1(-2x_1 - x_2 + 1) + \lambda_2(1 - x_1 - 3x_2) - \lambda_3 x_1 - \lambda_4 x_2. \tag{6}$$

Note that the domain of the Lagrangian is considered to be $\mathbb{R}^2 \times \mathbb{R}^4$ here.

(b) Show that $\mathcal{L}$ is concave in $(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$ for each fixed (x_1, x_2).

Solution: For each fixed (x_1, x_2), $-\mathcal{L}$ is convex in $(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$ as a affine function of $(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$. So $\mathcal{L}$ is concave in $(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$.

(c) Express the dual function of the problem, and show that it is concave.

Solution: The dual function is $g(\lambda_1, \lambda_2, \lambda_3, \lambda_4) = \min_{x_1, x_2} \mathcal{L}(x_1, x_2, \lambda_1, \lambda_2, \lambda_3, \lambda_4)$.

We can show that by showing that $-g$ is convex.

$$\begin{aligned} -g(\vec{\lambda}) &= -\min_{x_1, x_2} \mathcal{L}(x_1, x_2, \lambda_1, \lambda_2, \lambda_3, \lambda_4) && (7) \\ &= \max_{x_1, x_2} -\mathcal{L}(x_1, x_2, \lambda_1, \lambda_2, \lambda_3, \lambda_4). && (8) \end{aligned}$$

When (x_1, x_2) is fixed, the function $-\mathcal{L}$ is affine in $(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$.

Because the max of convex functions is convex, $-g$ is convex. Therefore g is concave.

(d) Assume f is convex. Show that $\mathcal{L}$ is convex in (x_1, x_2) for each fixed $(\lambda_1, \lambda_2, \lambda_3, \lambda_4) \in \mathbb{R}^4_+$.

Solution: For each fixed $(\lambda_1, \lambda_2, \lambda_3, \lambda_4) \in \mathbb{R}^4_+$, $\mathcal{L}$ is convex in (x_1, x_2) because it is a linear combination of functions that are convex in (x_1, x_2), with non-negative coefficients.

(e) Denoting $\mathcal{X} = \{(x_1, x_2) \mid 2x_1 + x_2 \geq 1,\ x_1 + 3x_2 \geq 1,\ x_1 \geq 0,\ x_2 \geq 0\}$, show that

$$\max_{\lambda_1 \geq 0, \lambda_2 \geq 0, \lambda_3 \geq 0, \lambda_4 \geq 0} \mathcal{L}(x_1, x_2, \lambda_1, \lambda_2, \lambda_3, \lambda_4) = \begin{cases} f(x_1, x_2) & \text{if } (x_1, x_2) \in \mathcal{X} \\ +\infty & \text{otherwise} \end{cases} \tag{9}$$

Solution: Let's just do it for λ_4:

$$\begin{aligned} \max_{\lambda_4 \geq 0} \mathcal{L}(x_1, x_2, \lambda_1, \lambda_2, \lambda_3, \lambda_4) &= \max_{\lambda_4 \geq 0} \left(f(x_1, x_2) + \lambda_1(-2x_1 - x_2 + 1) + \lambda_2(1 - x_1 - 3x_2) - \lambda_3 x_1 - \lambda_4 x_2\right) & (10) \\ &= f(x_1, x_2) + \lambda_1(-2x_1 - x_2 + 1) + \lambda_2(1 - x_1 - 3x_2) - \lambda_3 x_1 + \max_{\lambda_4 \geq 0} -\lambda_4 x_2 & (11) \end{aligned}$$

$$\max_{\lambda_4 \geq 0} -\lambda_4 x_2 = \begin{cases} 0 & \text{if } x_2 \geq 0 \\ +\infty & \text{otherwise} \end{cases} \tag{12}$$

One can show the same results for λ_1, λ_2 and λ_3, resulting in:

$$\max_{\lambda_1 \geq 0, \lambda_2 \geq 0, \lambda_3 \geq 0, \lambda_4 \geq 0} \mathcal{L}(x_1, x_2, \lambda_1, \lambda_2, \lambda_3, \lambda_4) = \begin{cases} f(x_1, x_2) & \text{if } (x_1, x_2) \in \mathcal{X} \\ +\infty & \text{otherwise.} \end{cases} \tag{13}$$

(f) Conclude that $\min_{(x_1, x_2) \in \mathcal{X}} \max_{\lambda_1 \geq 0, \lambda_2 \geq 0, \lambda_3 \geq 0, \lambda_4 \geq 0} \mathcal{L}(x_1, x_2, \lambda_1, \lambda_2, \lambda_3, \lambda_4) = \min_{(x_1, x_2) \in \mathcal{X}} f(x_1, x_2)$.

Solution: From part (e):

$$\max_{\lambda_1 \geq 0, \lambda_2 \geq 0, \lambda_3 \geq 0, \lambda_4 \geq 0} \mathcal{L}(x_1, x_2, \lambda_1, \lambda_2, \lambda_3, \lambda_4) = \begin{cases} f(x_1, x_2) & \text{if } (x_1, x_2) \in \mathcal{X}, \\ +\infty & \text{otherwise.} \end{cases} \tag{14}$$

Thus:

$$\min_{(x_1, x_2) \in \mathcal{X}} \max_{\lambda_1 \geq 0, \lambda_2 \geq 0, \lambda_3 \geq 0, \lambda_4 \geq 0} \mathcal{L}(x_1, x_2, \lambda_1, \lambda_2, \lambda_3, \lambda_4) = \min_{(x_1, x_2) \in \mathcal{X}} f(x_1, x_2). \tag{15}$$

(g) Assuming f is convex, formulate the first order condition on $\mathcal{L}$ as a function of ∇f and $(\lambda_1, \lambda_2, \lambda_3, \lambda_4) \in \mathbb{R}^4_+$ to solve:

$$\min_{x_1, x_2} \mathcal{L}(x_1, x_2, \lambda_1, \lambda_2, \lambda_3, \lambda_4) \tag{16}$$

Solution: Defining $\vec{\lambda} := (\lambda_1, \lambda_2, \lambda_3, \lambda_4) \in \mathbb{R}^4$ as the vector of dual variables, the first order condition is

$$0 = \nabla_{x_1, x_2} \mathcal{L}(x_1^\star(\vec{\lambda}), x_2^\star(\vec{\lambda}), \lambda_1, \lambda_2, \lambda_3, \lambda_4) = \nabla_{x_1, x_2} f(x_1^\star(\vec{\lambda}), x_2^\star(\vec{\lambda})) + \begin{bmatrix} -2\lambda_1 - \lambda_2 - \lambda_3 \\ -\lambda_1 - 3\lambda_2 - \lambda_4 \end{bmatrix}.$$

NOTE: Note here that $x_1^\star(\vec{\lambda})$ and $x_2^\star(\vec{\lambda})$ are functions of $\vec{\lambda}$. Also, because the inequalities are defined by affine functions in this problem, the role of the nonnegativity of the $\lambda_i, 1 \leq i \leq 4$ is obscured. However, in general, it is important.

2. Lagrangian Dual of a QP

Consider the general form of a convex quadratic program, where $Q \succ 0$ is a positive definite $n \times n$ matrix, $A \in \mathbb{R}^{m \times n}$, and $\vec{b} \in \mathbb{R}^m$:

$$\min_{\vec{x}} \quad \frac{1}{2}\vec{x}^\top Q\vec{x} \tag{17}$$
$$\text{s.t.} \quad A\vec{x} \leq \vec{b} \tag{18}$$

(a) Write the Lagrangian function $\mathcal{L}(\vec{x}, \vec{\lambda})$.

Solution:

$$\mathcal{L}(\vec{x}, \vec{\lambda}) = \frac{1}{2}\vec{x}^\top Q\vec{x} + \vec{\lambda}^\top (A\vec{x} - \vec{b}), \tag{19}$$

where $\vec{\lambda} \in \mathbb{R}^m$. The domain of the Lagrangian is $\mathbb{R}^n \times \mathbb{R}^m$.

(b) Write the Lagrangian dual function, $g(\vec{\lambda})$.

Solution: The dual function is

$$g(\vec{\lambda}) = \inf_{\vec{x}} \ \mathcal{L}(\vec{x}, \vec{\lambda}). \tag{20}$$

Since here also the inequality constraints are defined by affine functions, it turns out that $\mathcal{L}(\vec{x}, \vec{\lambda})$ is convex in $\vec{x}$ for each $\vec{\lambda}$, so we can find this infimum by setting $\nabla_{\vec{x}}\mathcal{L}(\vec{x}^*, \vec{\lambda}) = 0$:

$$Q\vec{x}^* + A^\top\vec{\lambda} = 0 \implies \vec{x}^* = -Q^{-1}A^\top\vec{\lambda}. \tag{21}$$

Substituting, we get

$$g(\vec{\lambda}) = \mathcal{L}(\vec{x}^*, \vec{\lambda}) \tag{22}$$
$$= \frac{1}{2}\vec{\lambda}^\top AQ^{-\top}A^\top\vec{\lambda} - \vec{\lambda}^\top AQ^{-1}A^\top\vec{\lambda} - \vec{\lambda}^\top\vec{b} \tag{23}$$
$$= -\frac{1}{2}\vec{\lambda}^\top AQ^{-1}A^\top\vec{\lambda} - \vec{\lambda}^\top\vec{b}. \tag{24}$$

(c) Show that the Lagrangian dual problem is convex by writing it in standard QP form. Is the Lagrangian dual problem convex in general?

Solution: The Lagrangian dual problem can be written as

$$\max_{\vec{\lambda} \geq 0} \ g(\vec{\lambda}) = \max_{\vec{\lambda} \geq 0} \ -\frac{1}{2}\vec{\lambda}^\top AQ^{-1}A^\top\vec{\lambda} - \vec{\lambda}^\top\vec{b}, \tag{25}$$

which is the maximization of a concave function of $\vec{\lambda}$ over the convex region given by the non-negative orthant $\vec{\lambda} \geq 0$. The dual problem is therefore convex.

While in this problem, the primal problem was convex, it turns out that the Lagrangian dual problem is a convex problem even when the primal is not. To see this, examine its general form:

$$\max_{\vec{\lambda} \geq 0} \min_{\vec{x}} \mathcal{L}(\vec{x}, \vec{\lambda}) = \max_{\vec{\lambda} \geq 0} \min_{\vec{x}} \left[f_0(\vec{x}) + \sum_{i=1}^{n} \lambda_i f_i(\vec{x}) \right]. \tag{26}$$

The inner minimum represents the pointwise minimum of affine functions of $\vec{\lambda}$, which we know to be concave in $\vec{\lambda}$. The resulting maximization problem of a concave objective in $\vec{\lambda}$ over the convex region $\vec{\lambda} \geq 0$ is then a convex optimization problem!

Note, however, that for fixed $\vec{x}$ the function $f_0(\vec{x}) + \sum_{i=1}^{n} \lambda_i f_i(\vec{x})$, as a function of $\vec{x}$, will not in general be convex unless $\vec{\lambda}$ has nonnegative entries. This feature is obscured in both of the examples covered in this discussion set.