

EECS 127/227AT Optimization Models in Engineering

Spring 2019

Homework 10

Release date: 11/15/19.

Due date: 11/21/19, 23:00 (11 pm). Please L^AT_EX or handwrite your homework solution and submit an electronic version.

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 saved as a PDF.

If you do not attach a PDF “printout” of your IPython notebook, you will not receive credit for problems that involve coding. Make sure that your results and your plots are visible. Assign the IPython printout to the correct problem(s) on Gradescope.

1. Geometric interpretation of duality

This problem has three types of questions. To make the student less confused between the questions we will write for each question if it is a question:

- (L): related to the lectures
- (P): specific to the problem
- (G): involving geometrical interpretation

In this problem, we look into the interpretation of duality using geometry. We will consider the following problem:

$$\begin{aligned} \min_x \quad & x^\top A x + b^\top x \\ \text{subject to} \quad & c^\top x + d \leq 0 \\ & -1 \leq x_1 \leq 1 \\ & -1 \leq x_2 \leq 1 \end{aligned} \tag{1}$$

Where $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, $b = \begin{bmatrix} 0 \\ -0.3 \end{bmatrix}$ and $c = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. During the problem, the value of A and d will change.

We can rewrite (1) as a general optimization problem:

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & f_0(x) \\ \text{subject to} \quad & f_i(x) \leq 0, \quad i = 1, 2, 3, 4, 5 \end{aligned}$$

With

$$\begin{aligned} f_0(x) &= x^\top A x + b^\top x \\ f_1(x) &= c^\top x + d \\ f_2(x) &= \begin{bmatrix} 0 & 1 \end{bmatrix} x - 1 \\ f_3(x) &= \begin{bmatrix} 0 & -1 \end{bmatrix} x - 1 \\ f_4(x) &= \begin{bmatrix} 1 & 0 \end{bmatrix} x - 1 \\ f_5(x) &= \begin{bmatrix} -1 & 0 \end{bmatrix} x - 1 \end{aligned}$$

Notation: In the remaining of the exercise, we will gather the box constraints $-1 \leq x_1 \leq 1$ and $-1 \leq x_2 \leq 1$ in one notation: $-1 \leq x \leq 1$.

1.1. Convex Problem

First, let's consider $A = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$ and $d = -1$.

- (a) (P) [A few sentences] Is problem (1) convex? Justify your answer.
- (b) (P) [Class and short justification] What is the class (LP, QP, QCQP, SOCP) of the problem?
- (c) (P) [Algebraic justification] Solve the problem.

1.2. Non-Convex Problem

Now, let's consider $A = \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix}$ and $d = 0.3$.

- (a) (P) [Short justification] Is problem (1) convex?

1.3. iPython Set-up and First Geometric Interpretation

Let try to understand duality using geometry. We can define the following set in \mathbb{R}^2 :

$$\mathcal{G} = \{(f_1(x), f_0(x)) \in \mathbb{R} \times \mathbb{R} \mid x \in \mathbb{R}^2, -1 \leq x \leq 1\}$$

- (a) (G) [Jupyter Notebook] Answer questions 1, 2, 3 and 4 on the iPython notebook.
Note: In the written exercises, we use $d = -1$ in the convex problem for computational simplicity. However, in the notebook, we will use $d = 0.3$ for both the convex and non-convex problems.
- (b) (L) [Algebraic justification] Show that the set \mathcal{A} is convex if the problem (1) is convex. The set \mathcal{A} is defined by:

$$\mathcal{A} = \{(u, t) \mid \exists x, f_1(x) \leq u, f_0(x) \leq t, -1 \leq x \leq 1\}$$

- (c) (L) [Algebraic justification] Show that any $\mathbf{z} \in \mathcal{A}$ can be expressed as $\hat{\mathbf{z}} + [\alpha_1, \alpha_2]^\top$ with $\hat{\mathbf{z}} \in \mathcal{G}$ and $\alpha_1, \alpha_2 > 0$.

We can write this condition as $\mathcal{A} = \mathcal{G} + \mathbb{R}_+ \times \mathbb{R}_+$

Remark: The convexity of the set \mathcal{A} is important for the equivalence between the primal problem and the dual problem. Basically, to have strong duality, we need to be able to define a supporting hyperplane (tangent in higher dimension than \mathbb{R}^2) of \mathcal{A} (this is equivalent to saying that $\{f_1(x) \mid x \in \mathbb{R}^2, -1 \leq x \leq 1\}$ is not a point) and to have \mathcal{A} be a convex set. If the problem is convex and the set $\{f_1(x) \mid x \in \mathbb{R}^2, -1 \leq x \leq 1\}$ is not a point (c.f. Slater's conditions), then \mathcal{A} is convex, and admits a supporting hyperplane, and therefore strong duality holds. The next section of the exercise will give you a taste of the proof.

1.4. Duality

Lagrangian

Recall the **Lagrangian** for the optimization problem is,

$$\mathcal{L}(x, \lambda_1, \lambda_2, \lambda_3) = f_0(x) + \sum_{i=1}^3 \lambda_i f_i(x)$$

where λ_i are the **Lagrange multipliers or dual variables**.

We can choose to “dualize” only some constraints. Define the Lagrangian without the box constraints on x as:

$$\hat{\mathcal{L}}(x, \lambda_1) = f_0(x) + \lambda_1 f_1(x) \quad \forall x, \text{ s.t. } -1 \leq x_1 \leq 1, -1 \leq x_2 \leq 1$$

Given an x in the box constraints and a $\lambda_1 \geq 0$ we can define the following function:

$$h_{x, \lambda_1} : z \rightarrow h_{x, \lambda_1}(z) = f_0(x) + \lambda_1(f_1(x) - z)$$

Then we can use $h_{x, \lambda_1}(z)$ to define the following set:

$$H_{x, \lambda_1} = \{(z, h_{x, \lambda_1}(z)), z \in \mathbb{R}\}$$

- (a) (G) [Short justification] Show that given an x and a λ_1 , the set H_{x, λ_1} is a line in \mathbb{R}^2 .

- (b) (G) [No justification needed] State the slope of the line H_{x, λ_1}

Define the vertical line when the x-coordinate is 0 as

$$L = \{(0, y) \mid y \in \mathbb{R}\}$$

- (c) (G) [Short justification] Explain how you can find $\hat{\mathcal{L}}(x, \lambda_1)$ with the set H_{x, λ_1} and L . Run the section “Geometric interpretation of the Lagrangian” on the iPython notebook.

Primal problem

We can now define the primal problem of (1) as:

$$p^* = \min_{\substack{x \\ -1 \leq x \leq 1}} \max_{\lambda_1 \geq 0} \hat{\mathcal{L}}(x, \lambda_1) \quad (2)$$

- (d) (G) [Jupyter notebook] Answer questions 5, 6, 7, 8, 9 and 10 in the Jupyter notebook.
- (e) (L) [Algebraic justification] Show mathematically that problem (2) is equivalent to (1).

We can also define the dual problem of (1) as:

$$d^* = \max_{\lambda_1 \geq 0} \min_{\substack{x \\ -1 \leq x \leq 1}} \hat{\mathcal{L}}(x, \lambda_1)$$

- (f) (G) [Jupyter Notebook] Run the cells corresponding to the section “Solving the dual problem” and answer questions 11, 12 and 13 on the iPython notebook.
- (g) (L) [Algebraic justification] Show mathematically that $d^* \leq p^*$

Define the Lagrange dual function as:

$$g(\lambda_1) = \min_{\substack{x \\ -1 \leq x \leq 1}} \hat{\mathcal{L}}(x, \lambda_1)$$

- (h) (L) [Short justification] Show that g is a concave function.
- (i) (L) [Short justification] Show that the dual problem is a convex problem with respect to λ_1 . Express it in the standard form of a convex problem
- (j) (L) [Algebraic justification] Recall \mathcal{A} from 1.3 b). Given $(u, t) \in \mathcal{A}$, show that $\forall \lambda_1 \geq 0$, $(u, t)^\top (\lambda_1, 1) \geq g(\lambda_1)$.

Remark: Given λ_1 and x^* such that x^* minimizes $\hat{\mathcal{L}}(x^*, \lambda_1)$ for λ_1 fixed, H_{x^*, λ_1} defines a supportive hyperplane of \mathcal{G} and \mathcal{A} . It can be written as $H_{x^*, \lambda_1} = \{(u, t) \mid (u, t)^\top (\lambda_1, 1) = g(\lambda_1)\}$. This remark is the mathematical explanation of the intuition behind question 11 of the iPython notebook.

1.5. Convex Relaxation of the Problem Using Duality

In this section, we will consider only the NON-convex problem.

We can write $g(\lambda_1) = \min_{\substack{x \\ -1 \leq x \leq 1}} x^\top A x + b^\top x + \lambda_1 (c^\top x + d)$.

Recall that, in this case, $A = \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix}$ and $d = 0.3$. Recall also that we only consider $\lambda_1 \geq 0$.

- (a) (P) [Algebraic justification] Express the Lagrange dual function for the non convex problem as

$$g(\lambda_1) = \min_{\substack{x \\ -1 \leq x \leq 1}} g_1(x_1, \lambda_1) - g_2(x_2, \lambda_1) + g_3(\lambda_1)$$

where you will find the functions g_1, g_2 and g_3 .

Note that the function $g_1(x_1, \lambda_1)$ is convex in x_1 and the function $g_2(x_2, \lambda_1)$ is convex in x_2 . To find $g(\lambda_1)$ we will minimize the Lagrangian. Minimizing the Lagrangian can be interpreted as minimizing $g_1(x_1, \lambda_1)$ in x_1 and maximizing $g_2(x_2, \lambda_1)$ in x_2 .

- (b) (P) [Algebraic justification] Find the expression for $\min_{\substack{x_1 \\ -1 \leq x_1 \leq 1}} g_1(x_1, \lambda_1)$. Make sure you consider the box-constraints for x_1 .
- (c) (P) [Algebraic justification] Find the expression for $\max_{\substack{x_2 \\ -1 \leq x_2 \leq 1}} g_2(x_2, \lambda_1)$. Make sure you consider the box-constraints for x_2 .
- (d) (P) [Algebraic justification] Combine the expressions you got in part (b) and (c) and explicitly state the Lagrange dual function $g(\lambda_1)$.
- (e) (P) [Algebraic justification] Find the dual optimal solution (d^*).
- (f) (G) [Jupyter notebook] Answer questions 14 and 15 of the iPython notebook.

1.6. Convex Relaxation Using Eigenvalue Shifting

In this section, we will consider only the NON-convex problem.

The duality gap gives a way to compute a lower-bound on the solution of the non convex problem with the dual problem which is convex. To get a sense of the gap between the solution and the estimate done with the dual problem, it can be nice to get an upper-bound on the non convex problem.

To do so, we can relax the problem into a convex problem using eigenvalue shifting. Basically, the non convexity of the problem is due to the eigenvalue -1 of A that makes the matrix non convex.

To make the problem convex, we can set the -1 eigenvalue of A to 0 by considering $\hat{A} = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}$.

- (a) (P) [Short algebraic justification] Show that $\forall x, x^\top A x \leq x^\top \hat{A} x$ (with A the matrix that defines the non convex problem: $A = \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix}$).
- (b) (P) [Short algebraic and conceptual justification] Show that solving (1) with $A = \hat{A} = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}$ will provide an upperbound on the problem (1) with $A = \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix}$.

Another way to compute an upperbound on the non convex problem is to perform a random search on the domain of the problem.

- (c) (G) [Jupyter notebook] Run the iPython notebook section “Upperbound on the non convex problem” and answer question 16 of the iPython notebook.

1.7. Sensitivity Analysis

In this section, we will consider only the CONVEX problem.

To be able to see the sensitivity analysis, we will consider $d = .3$ and $A = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$.

We have already shown that the problem is convex. It might be harder to solve the problem here, because the constraints are active at the optimal point. We can solve the problem with the help of the geometrical interpretation of duality (i.e., using the KKT conditions). Basically, when the problem is convex, the optimal point is reached either on the right side of the red line or on the red line. This means that either $\lambda_1^* = 0$ or $f_1(x^*) = 0$. This can be written as $\lambda_1^* f_1(x^*) = 0$ (it is called **complementary slackness**) and $\lambda_1^* \geq 0$ (**dual feasibility**; see question 5 of the iPython notebook) and $f_1(x^*) \leq 0$ (**primal feasibility**; see question 3 of the iPython notebook).

If $\lambda_1^* = 0$, then the optimal point is on the left side of the red line and the problem can be solved by taking the gradient of f_0 and setting it to 0.

If $\lambda_1^* \neq 0$, the problem is harder to solve. In this case the optimal point is on the red line. Question 11 of the iPython notebook shows that the optimal point is such that H_{x^*, λ_1^*} is tangent to \mathcal{G} (more generally, H_{x^*, λ_1^*} is a supporting hyperplane of \mathcal{G} or \mathcal{A}). Therefore, in this case, we can find the solution by looking at the tangent line (or supporting hyperplane) to the set \mathcal{G} (or \mathcal{A}) at the red line (or for the hyperplane $\{(f_0(x), 0) \mid \exists x, f_1(x) = 0\}$). This is called the **stationarity condition** and can be mathematically expressed as

$$\nabla_x \hat{\mathcal{L}}(x^*, \lambda_1^*) = 0$$

Together, the **primal feasibility**, the **dual feasibility**, the **complementary slackness**, and the **stationarity condition** are known as the Karush-Kuhn-Tucker (KKT) conditions and are necessary conditions of any minimization problem. They are sufficient conditions for convex problems. This means that if you find x^* and λ_1^* that satisfy the KKT conditions for the convex problem, then they are the optimal solution of the problem.

- (a) (P) [Algebraic justification] Use the conditions that $\nabla_x \hat{\mathcal{L}}(x^*, \lambda_1^*) = 0$ and $c^\top x^* + d = 0$ to find the optimal value of x^* and λ_1^* for the problem.
- (b) (G) [Jupyter notebook] Answer question 17 of the iPython notebook.

Now, you should know everything to be able to understand how dual variable can be used to do sensitivity analysis of an optimization problem.

Consider perturbing d by a small amount such that the constraint $f_1(x) \leq 0$ is moved slightly. This is equivalent to the following problem:

$$\begin{aligned} p^*(u) = \min_x \quad & f_0(x) \\ \text{subject to} \quad & f_1(x) \leq u \\ & -1 \leq x \leq 1 \end{aligned}$$

As you can see, problem (1) is equivalent to this problem when $u = 0$ (i.e., $p^* = p^*(0)$).

- (c) (G) [Jupyter notebook] Answer question 18 of the iPython notebook.
- (d) (L) [Algebraic justification] Show that

$$p^*(u) \geq p^*(0) - \lambda_1^* u$$

where λ_1^* is the optimal dual variable computed in question (a).

Hint: Show first that for any x which is feasible, $g(\lambda_1^*) \leq f_0(x) + \lambda_1^* u$.

To go further, feel free to read p.232 to p.241 and p.249 to p.253 of the Boyd's book.