Lecture 7

Convex duality

Emil Gustavsson
Department of Mathematical Sciences
Chalmers University of Technology and Göteborg University
November 19, 2012

CHALMERS | GÖTEBORGS UNIVERSITET

Definition

Consider the problem

$$
\begin{aligned}
f^* = \text{infimum} \quad & f(\mathbf{x}), && \text{(1a)} \\
\text{subject to} \quad & \mathbf{x} \in S. && \text{(1b)}
\end{aligned}
$$

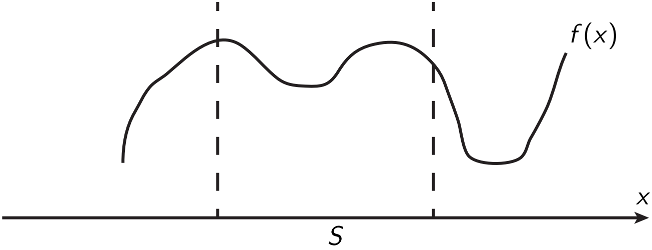
We say that a **relaxation** to (1) is a problem on the following form:

$$
\begin{aligned}
f_R^* = \text{infimum} \quad & f_R(\mathbf{x}), && \text{(2a)} \\
\text{subject to} \quad & \mathbf{x} \in S_R, && \text{(2b)}
\end{aligned}
$$

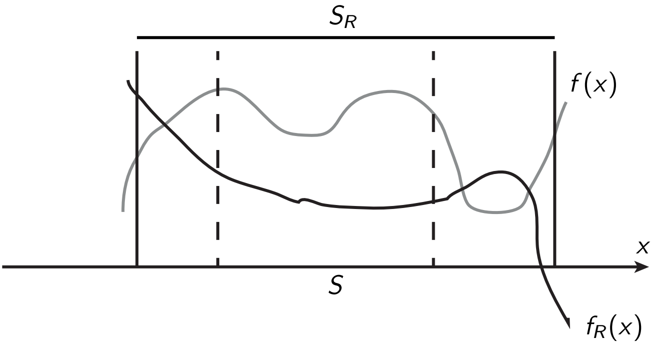
where

- $f_R(\mathbf{x}) \leq f(\mathbf{x})$ for all $\mathbf{x} \in S$ (the relaxed function is lower on S)
- $S_R \supseteq S$ (the relaxed feasible set is larger)

Example



Example



Relaxation theorem

$$f^* = \text{infinum} \quad f(\mathbf{x}), \qquad \text{subject to} \quad \mathbf{x} \in S. \qquad (1)$$

$$f_R^* = \text{infinum} \quad f_R(\mathbf{x}), \qquad \text{subject to} \quad \mathbf{x} \in S_R. \qquad (2)$$

The relaxation theorem

a) $f_R^* \leq f^*$

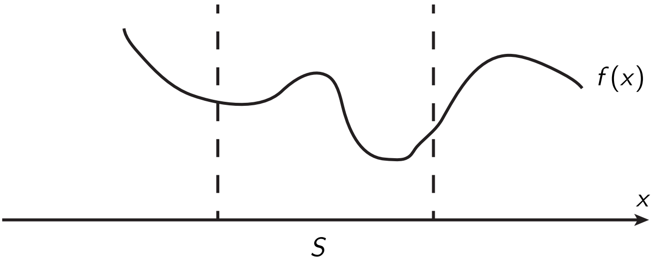
b) If (2) is infeasible, then so is (1)

c) If (2) has an optimal solution, $\mathbf{x}_R^*$, for which

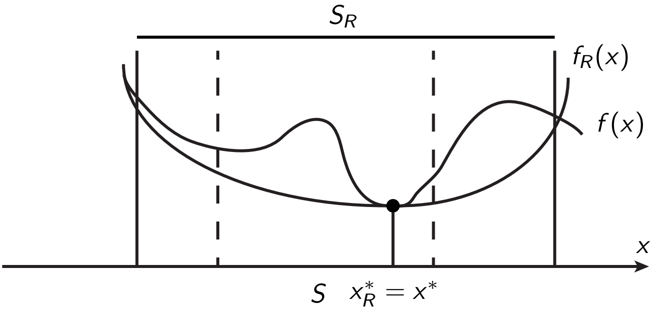
$$\mathbf{x}_R^* \in S \quad \text{and} \quad f_R(\mathbf{x}_R^*) = f(\mathbf{x}_R^*),$$

then $\mathbf{x}_R^*$ is an optimal solution to (1) as well

Example



Example



Relaxation theorem

The relaxation theorem

a) $f_R^* \leq f^*$

b) If (2) is infeasible, then so is (1)

c) If (2) has an optimal solution, $\mathbf{x}_R^*$, for which

$$\mathbf{x}_R^* \in S \quad \text{and} \quad f_R(\mathbf{x}_R^*) = f(\mathbf{x}_R^*),$$

then $\mathbf{x}_R^*$ is an optimal solution to (1) as well

Proof.

Outline

We will today

- Consider a specific relaxation technique called **Lagrangian relaxation**.
- Complicated constraints will be added to the objective function with a penalty.
- This gives us a relaxation of the original problem.
- The goal is to find the best possible penalties such that the relaxed problem has the same optimal value as the original one.

A simple relaxation

Consider the problem to find

$$
\begin{aligned}
f^* = \text{infinum} \quad & f(\mathbf{x}), && \text{(3a)} \\
\text{subject to} \quad & g_i(\mathbf{x}) \leq 0, \quad i = 1, \dots, m, && \text{(3b)} \\
& \mathbf{x} \in X. && \text{(3c)}
\end{aligned}
$$

- Assume the constraints (3b) are complicated
- If we removing them, the resulting problem

$$
\begin{aligned}
f^* = \text{infinum} \quad & f(\mathbf{x}), \\
\text{subject to} \quad & \mathbf{x} \in X,
\end{aligned}
$$

  is easy to solve
- Clearly a relaxation of the original problem

Lagrangian relaxation, I

$$
\begin{aligned}
f^* = \text{infinum} \quad & f(\mathbf{x}), && (4a) \\
\text{subject to} \quad & g_i(\mathbf{x}) \leq 0, \quad i = 1, \ldots, m, && (4b) \\
& \mathbf{x} \in X. && (4c)
\end{aligned}
$$

So instead of just removing the constraints, we add each constraint g_i to the objective function with a **multiplier** μ_i.

We define the **Lagrange function** as

$$
L(\mathbf{x}, \boldsymbol{\mu}) := f(\mathbf{x}) + \sum_{i=1}^{m} \mu_i g_i(\mathbf{x}) = f(\mathbf{x}) + \boldsymbol{\mu}^T \mathbf{g}(\mathbf{x})
$$

Given a multiplier vector $\boldsymbol{\mu} \in \mathbb{R}^m$, we define

the **Lagrangian dual function** evaluated at the point $\boldsymbol{\mu} \in \mathbb{R}^m$ as

$$q(\boldsymbol{\mu}) = \begin{array}{ll} \text{infinum} & L(\mathbf{x}, \boldsymbol{\mu}) \\ \text{subject to} & \mathbf{x} \in X \end{array} \quad = \quad \begin{array}{ll} \text{infinum} & f(\mathbf{x}) + \boldsymbol{\mu}^T \mathbf{g}(\mathbf{x}), \\ \text{subject to} & \mathbf{x} \in X \end{array}$$

We will later show that this is in fact a relaxation of the original problem

Lagrange multipliers

We call the vector $\boldsymbol{\mu}^* \in \mathbb{R}^m_+$ (nonnegative) a **Lagrange multiplier** if

$$f^* = q(\boldsymbol{\mu}^*) = \begin{array}{ll} \text{infinum} & f(\mathbf{x}) + {\boldsymbol{\mu}^*}^T \mathbf{g}(\mathbf{x}), \\ \text{subject to} & \mathbf{x} \in X \end{array}$$

- If we minimize the Lagrange function $L(\mathbf{x}, \boldsymbol{\mu}^*)$ with $\mathbf{x} \in X$, we obtain the same value as
- minimizing $f(\mathbf{x})$ with $\{\mathbf{x} \in X,\ \mathbf{g}(\mathbf{x}) \leq \mathbf{0}\}$
- The multipliers, μ_i must be nonnegative in order not to favour $g_(\mathbf{x}) > 0$.

Optimality conditions

Now assume that we have a Lagrange multiplier, $\boldsymbol{\mu}^*$. How do we obtain the optimal solution $\mathbf{x}^*$?

Theorem

Let $\boldsymbol{\mu}^* \in \mathbb{R}^m_+$ be a Lagrange multiplier. Then, $\mathbf{x}^*$ is optimal in the original problem if and only if

a) $\mathbf{x}^* \in \operatorname*{argmin}_{\mathbf{x} \in X} L(\mathbf{x}, \boldsymbol{\mu}^*)$ ($\mathbf{x}^*$ is one of the solutions)

b) $\mathbf{x}^* \in X,\ \mathbf{g}(\mathbf{x}^*) \leq \mathbf{0}$, ($\mathbf{x}^*$ is feasible)

c) $\mu_i g_i(\mathbf{x}^*) = 0,\ i = 1, \dots, m$ (complementarity)

Proof:

Example 1

Consider the problem

$$
\begin{aligned}
f^* = \text{minimum} \;\; f(\mathbf{x}) &= x_1^2 + x_2^2, && \text{(5a)}\\
\text{subject to} \qquad x_1 + x_2 &\geq 4, && \text{(5b)}\\
x_1, x_2 &\geq 0. && \text{(5c)}
\end{aligned}
$$

We decide to Lagrangian relax constraint (5b). Find the optimal solution if you know that $\mu^* = 4$ is a Lagrange multiplier.

Solution:

Connections with KKT

These optimality conditions extend the KKT-conditions. Assume

- The Lagrange function is not globally minimized, stationary points are sufficient
- $X = \mathbb{R}^n$
- $f, g_i \in C^1$

$$
\begin{aligned}
\mathbf{x}^* \in \operatorname*{argmin}_{\mathbf{x} \in X} L(\mathbf{x}, \boldsymbol{\mu}^*) \ &\Rightarrow \quad \nabla f(\mathbf{x}^*) + \sum_{i=1}^{m} \mu_i^* \nabla g_i(\mathbf{x}^*) = 0 \\
\mathbf{x}^* \in X,\ \mathbf{g}(\mathbf{x}^*) \leq \mathbf{0} \ &\Rightarrow \quad g_i(\mathbf{x}^*) \leq 0,\ i = 1, \dots, m \\
\mu_i^* g_i(\mathbf{x}^*) = 0, i = 1, \dots, m \ &\Rightarrow \quad \mu_i^* g_i(\mathbf{x}^*) = 0,\ i = 1, \dots, m
\end{aligned}
$$

How do we find Lagrange multipliers?

$$q(\boldsymbol{\mu}) = \underset{\mathbf{x}\in X}{\text{infinum}}\, L(\mathbf{x}, \boldsymbol{\mu}) = \underset{\mathbf{x}\in X}{\text{infinum}}\, \left\{ f(\mathbf{x}) + \boldsymbol{\mu}^T \mathbf{g}(\mathbf{x}) \right\}.$$

Take $\boldsymbol{\mu} \geq \mathbf{0}$, and a feasible $\mathbf{x}$ (i.e. $\mathbf{x} \in X$, $\mathbf{g}(\mathbf{x}) \leq \mathbf{0}$) then

$$q(\boldsymbol{\mu}) \leq f(\mathbf{x})$$

- This means that the problem $\inf_{\mathbf{x}\in X} L(\mathbf{x}, \boldsymbol{\mu})$ is a relaxation to the original problem whenever $\boldsymbol{\mu} \geq \mathbf{0}$.
- Since $q(\boldsymbol{\mu})$ is always smaller than $f(\mathbf{x})$, we would like to find a $\boldsymbol{\mu}$ such that $q(\boldsymbol{\mu})$ is as large as possible.

Lagrangian dual problem, I

$$q(\boldsymbol{\mu}) = \operatorname*{infimum}_{\mathbf{x} \in X} L(\mathbf{x}, \boldsymbol{\mu}) = \operatorname*{infimum}_{\mathbf{x} \in X} \left\{ f(\mathbf{x}) + \boldsymbol{\mu}^T \mathbf{g}(\mathbf{x}) \right\}.$$

We define the **Lagrangian dual problem** as the problem to find

$$\begin{aligned} q^* = \ & \text{supremum } q(\boldsymbol{\mu}), \\ & \text{subject to} \quad \boldsymbol{\mu} \geq \mathbf{0} \end{aligned}$$

- For some $\boldsymbol{\mu}$, $q(\boldsymbol{\mu}) = -\infty$ is possible; if this is true for all $\boldsymbol{\mu} \geq \mathbf{0}$, we say that $q^* = -\infty$.
- The **effective domain** of q is $D_q = \{\boldsymbol{\mu} \in \mathbb{R}^m \mid q(\boldsymbol{\mu}) > -\infty\}$

Lagrangian dual problem, II

Theorem
The effective domain D_q of q is a convex set, and q is a concave function on D_q.

Proof:

Lagrangian dual problem, III

Since q is a concave function on the convex set D_q, the optimization problem

$$\begin{aligned} q^* = \text{supremum}\ & q(\boldsymbol{\mu}), \\ \text{subject to}\quad & \boldsymbol{\mu} \geq \mathbf{0} \end{aligned}$$

is a convex problem.

Weak duality, I

Theorem Let $\mathbf{x} \in X$, $\mathbf{g}(\mathbf{x}) \leq \mathbf{0}$ and $\boldsymbol{\mu} \geq \mathbf{0}$ (both feasible in their respective problems). Then

$$q(\boldsymbol{\mu}) \leq f(\mathbf{x}).$$

In particular

$$q^* \leq f^*$$

holds.
If $q(\boldsymbol{\mu}) = f(\mathbf{x})$, then the pair $(\mathbf{x}, \boldsymbol{\mu})$ is optimal in its respecitve problem.

Proof:

Weak duality, II

- **Note 1:**

  The reverse is not always true. Meaning that $\mathbf{x}$ and $\boldsymbol{\mu}$ could be optimal in their respective problems and $q(\boldsymbol{\mu}) < f(\mathbf{x})$.

- **Note 2:**

  The week duality theorem is a consequence of the relaxation theorem. For any $\boldsymbol{\mu} \geq \mathbf{0}$

$$
\begin{aligned}
f &:= f \\
f_R &:= L(\cdot, \boldsymbol{\mu}) \\
S &:= X \cap \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{g}(\mathbf{x}) \leq \mathbf{0}\} \\
S_R &:= X
\end{aligned}
$$

Weak duality, III

- **Note 3:**

  Powerful tool in practice when solving optimization problems.

  Example.

  We are solving some complicated optimization problem and have obtained an $\mathbf{x}$ with a cost $f(\mathbf{x}) = 1,000,000$ Skr. We also have found an $\boldsymbol{\mu}$ with $q(\boldsymbol{\mu}) = 999,999$ Skr. Then we know that we can not improve our solution with more than 1 Skr, meaning that its hard to justify searching for better solutions.

- **Note 4:**

  If we have an $\mathbf{x}$ and an $\boldsymbol{\mu}$, we get a quality measure of the solution $\mathbf{x}$ by $\frac{f(\mathbf{x})-q(\boldsymbol{\mu})}{q(\boldsymbol{\mu})}$

Weak duality, IV

$$\begin{aligned} f^* = \text{infinum} \quad & f(\mathbf{x}), \\ \text{subject to} \quad & \mathbf{g}(\mathbf{x}) \leq \mathbf{0}, \\ & \mathbf{x} \in X. \end{aligned} \tag{P}$$

$$\begin{aligned} q^* = \text{supremum} \quad & q(\boldsymbol{\mu}), \\ \text{subject to} \quad & \boldsymbol{\mu} \geq \mathbf{0} \end{aligned} \tag{D}$$

- From weak duality we have that $q^* \leq f^*$
- If $q^* = f^*$, we say that there is no duality gap.
- As we will see, this occurs when (P) is convex and some suitable CQ holds.

Example 2

Consider the problem

$$
\begin{aligned}
f^* = \text{minimize} \quad & f(\mathbf{x}) = x_1^2 + x_2^2, && \text{(6a)}\\
\text{subject to} \quad & x_1 + x_2 \geq 4, && \text{(6b)}\\
& x_1, x_2 \geq 0. && \text{(6c)}
\end{aligned}
$$

We decide to Lagrangian relax constraint (6b). Formulate the dual function $q(\boldsymbol{\mu})$ explicitly, find q^* and f^*.

Solution:

Note on how to construct $L(\mathbf{x}, \boldsymbol{\mu})$

When considering the problem to find

$$\begin{aligned} f^* = \text{infinum} \quad & f(\mathbf{x}), \\ \text{subject to} \quad & \mathbf{g}(\mathbf{x}) \leq \mathbf{0}, \\ & \mathbf{x} \in X, \end{aligned} \tag{P}$$

we can decide which constraints to relax.

- **Many constraints are included in X**
  - The problem $q(\boldsymbol{\mu})$ is harder to solve
  - The bounds are better ($q(\boldsymbol{\mu})$ will be closer to $f(\mathbf{x})$)
- **Few constraints are included in X**
  - The problem $q(\boldsymbol{\mu})$ is easier to solve
  - The bounds are worse ($q(\boldsymbol{\mu})$ will be further away from $f(\mathbf{x})$)

Global optimality, I

We will now

- Characterize every optimal primal and dual solution
- when there exist a Lagrange multiplier and
- no duality gap

Global optimality, II

Theorem

The vector $(\mathbf{x}^*, \boldsymbol{\mu}^*)$ is a pair of optimal primal solution and Lagrange multiplier if and only if

$$\mathbf{x} \in X, \quad \mathbf{g}(\mathbf{x}^*) \leq \mathbf{0}, \quad \text{(primal feasibility)} \tag{7a}$$

$$\boldsymbol{\mu}^* \geq \mathbf{0}, \quad \text{(dual feasibility)} \tag{7b}$$

$$\mathbf{x}^* \in \operatorname*{argmin}_{\mathbf{x} \in X} L(\mathbf{x}, \boldsymbol{\mu}^*), \quad \text{(Lagrangian optimality)} \tag{7c}$$

$$\mu_i^* g_i(\mathbf{x}^*) = 0, \quad \text{(complementary slackness)} \tag{7d}$$

Proof:

Global optimality, III

Theorem
The vector $(\mathbf{x}^*, \boldsymbol{\mu}^*)$ is a pair of optimal primal solution and Lagrange multiplier if and only if $\mathbf{x} \in X$ and $\boldsymbol{\mu} \geq \mathbf{0}$, and $(\mathbf{x}^*, \boldsymbol{\mu}^*)$ is a **saddle point** of the Lagrangian function on $X \times \mathbb{R}^m_+$, that is,

$$L(\mathbf{x}^*, \boldsymbol{\mu}) \leq L(\mathbf{x}^*, \boldsymbol{\mu}^*) \leq L(\mathbf{x}, \boldsymbol{\mu}^*), \quad (\mathbf{x}, \boldsymbol{\mu}) \in X \times \mathbb{R}^m_+.$$

Strong duality

- So far, we have not assumed any properties of the problem

$$
\begin{aligned}
f^* = \text{infinum} \quad & f(\mathbf{x}), \\
\text{subject to} \quad & \mathbf{g}(\mathbf{x}) \leq \mathbf{0}, \\
& \mathbf{x} \in X,
\end{aligned}
\tag{P}
$$

- However, the characterization of primal-dual optimal solutions depends on the fact that
  - There exist a Lagrange multiplier, $\boldsymbol{\mu}^*$
  - The duality gap is zero ($f^* = q^*$)
- In order to establish **strong duality**, that is, to establish sufficient conditions under which there is no duality gap, we need convexity.

Convexity and Slater

- Consider the problem

$$\begin{aligned} f^* = \text{infinum} \quad & f(\mathbf{x}), \\ \text{subject to} \quad & \mathbf{g}(\mathbf{x}) \leq \mathbf{0}, \\ & \mathbf{x} \in X, \end{aligned} \tag{P}$$

  where f and g_i, $i = 1, \ldots, m$ are convex functions and X is a convex set.

- Assume the Slater condition:

$$\exists\, \mathbf{x} \in X \text{ with } \mathbf{g}(\mathbf{x}) < \mathbf{0}$$

Strong duality Theorem, I

Theorem

Suppose (P) is convex, and Slater's CQ holds, then

a) There is no duality gap ($f^* = q^*$) and there exist at least one Lagrange multiplier $\boldsymbol{\mu}^*$. Moreover, the set of Lagrange multiplier is bounded and convex.

b) If the infinum in (P) is attained at some $\mathbf{x}^*$, then the pair $(\mathbf{x}^*, \boldsymbol{\mu}^*)$ satisfies the global optimality conditions.

c) If the functions f and g_i, $i = 1, \ldots, m$ are in C^1 and X is open, then the global optimality conditions are equivalent with the KKT-conditions.

Strong duality Theorem, II

Consequences of (P) being convex is thus:

- If we can solve the dual problem to optimality (i.e. find $\boldsymbol{\mu}^*$ and q^*), we also know the optimal primal value f^*.
- But then we have solved

$$q^* = q(\boldsymbol{\mu}^*) = \min_{\mathbf{x} \in X} L(\mathbf{x}, \boldsymbol{\mu}^*).$$

- We can find the optimal solution $\mathbf{x}^*$ by letting it fulfill the optimality conditions

$$\begin{aligned} \mathbf{x}^* \in \underset{\mathbf{x} \in X}{\operatorname{argmin}} L(\mathbf{x}, \boldsymbol{\mu}^*) \\ \mathbf{g}(\mathbf{x}^*) \leq \mathbf{0}, \\ \mu_i^* g_i(\mathbf{x}^*) = 0, \quad i = 1, \ldots, m \end{aligned}$$

Note: The last step is not always trivial.

Consider the linear program

$$
\begin{aligned}
f^* = \text{infinum} \quad & \mathbf{c}^T\mathbf{x}, && \text{(8a)}\\
\text{subject to} \quad & \mathbf{A}\mathbf{x} = \mathbf{b}, && \text{(8b)}\\
& \mathbf{x} \geq \mathbf{0}, && \text{(8c)}
\end{aligned}
$$

where $\mathbf{c} \in \mathbb{R}^n$, $\mathbf{A} \in \mathbb{R}^{m\times n}$ and $\mathbf{b} \in \mathbb{R}^m$. If we let $X = \mathbb{R}^n_+$, the Lagrangian dual problem is

$$
\begin{aligned}
q^* = \text{supremum} \quad & \mathbf{b}^T\boldsymbol{\mu}, && \text{(9a)}\\
\text{subject to} \quad & \mathbf{A}^T\boldsymbol{\mu} \leq \mathbf{c}, && \text{(9b)}\\
& \boldsymbol{\mu} \in \mathbb{R}^m && \text{(9c)}
\end{aligned}
$$

Why?

$$\begin{aligned} f^* = \text{infinum} \quad & \mathbf{c}^T\mathbf{x}, && \text{(10a)}\\ \text{subject to} \quad & \mathbf{A}\mathbf{x} = \mathbf{b}, && \text{(10b)}\\ & \mathbf{x} \geq \mathbf{0}, && \text{(10c)} \end{aligned}$$

Lagrangian relax constraint (10b) and get $L(\mathbf{x}, \boldsymbol{\mu}) = \mathbf{c}^T\mathbf{x} + \boldsymbol{\mu}^T(\mathbf{b} - \mathbf{A}\mathbf{x})$.

$$\begin{aligned} q(\boldsymbol{\mu}) &= \inf_{\mathbf{x} \geq \mathbf{0}} \left\{\mathbf{c}^T\mathbf{x} + \boldsymbol{\mu}^T(\mathbf{b} - \mathbf{A}\mathbf{x})\right\} = \mathbf{b}^T\boldsymbol{\mu} + \inf_{\mathbf{x} \geq \mathbf{0}} \left\{(\mathbf{c} - \mathbf{A}^T\boldsymbol{\mu})^T\mathbf{x}\right\}\\ &= \begin{cases} \mathbf{b}^T\boldsymbol{\mu} & \text{if } \mathbf{A}^T\boldsymbol{\mu} \leq \mathbf{c},\\ -\infty & \text{otherwise} \end{cases} \end{aligned}$$

$$\Longrightarrow \qquad \begin{aligned} q^* = \text{supremum} \quad & \mathbf{b}^T\boldsymbol{\mu}, && \text{(11a)}\\ \text{subject to} \quad & \mathbf{A}^T\boldsymbol{\mu} \leq \mathbf{c}, && \text{(11b)}\\ & \boldsymbol{\mu} \in \mathbb{R}^m && \text{(11c)} \end{aligned}$$

Questions for the break

Find f^* and q^* in the following problems when Lagrangian relaxing the blue constraint. If the problem if infeasible, by definition $f^* = \infty$

$$\begin{aligned} f^* = \min \ & 1/x, \\ \text{s.t.} \quad & x \leq 0, \\ & x > 0. \end{aligned} \tag{A}$$

$$\begin{aligned} f^* = \min \ & x, \\ \text{s.t.} \quad & x^2 \leq 0, \\ & x > 0. \end{aligned} \tag{B}$$

$$\begin{aligned} f^* = \min \ & x_1 + x_2, \\ \text{s.t.} \quad & x_1 \leq 0, \\ & x_1 > 0. \end{aligned} \tag{C}$$