EECS 127/227AT Optimization Models in Engineering
Spring 2020

Discussion 5

1. Simple constrained optimization problem

Consider the optimization problem

$$\begin{aligned} \min_{x_1, x_2 \in \mathbb{R}} \ & f(x_1, x_2) \\ \text{subject to } & 2x_1 + x_2 \geq 1 \\ & x_1 + 3x_2 \geq 1 \\ & x_1 \geq 0, \ x_2 \geq 0 \end{aligned}$$

(a) Make a sketch of the feasible set.

Solution: See figure 1.

For each of the following objective functions, give the optimal set or the optimal value.

(b) $f(x_1, x_2) = x_1 + x_2$

Solution: Using the drawing (figure 2) it seems that the solution is such that $x_1^\star = \frac{2}{5}$ and $x_2^\star = \frac{1}{5}$.

One can verify the optimality of such point using the first order convexity condition:

$$\nabla f(\frac{2}{5}, \frac{1}{5})^\top ((\frac{2}{5}, \frac{1}{5}) - (x_1, x_2)) \geq 0, \quad \forall (x_1, x_2) \in \mathcal{X}$$

Where $\mathcal{X}$ is the feasible set.

It can also be derived using strong duality (see next lecture on duality).

(c) $f(x_1, x_2) = -x_1 - x_2$

Solution: Here (figure 2) the problem is unbounded below as if $(x_1, x_2) = t(1, 1)$ with $t \geq 0$ then (x_1, x_2) is always feasible and $-2t \to -\infty$ when $t \to \infty$.

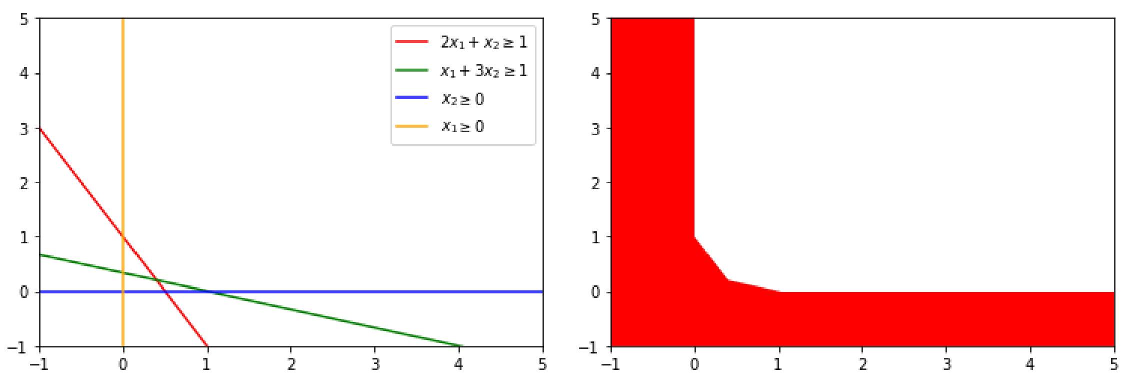

Figure 1: The feasible set is in white on the right figure.

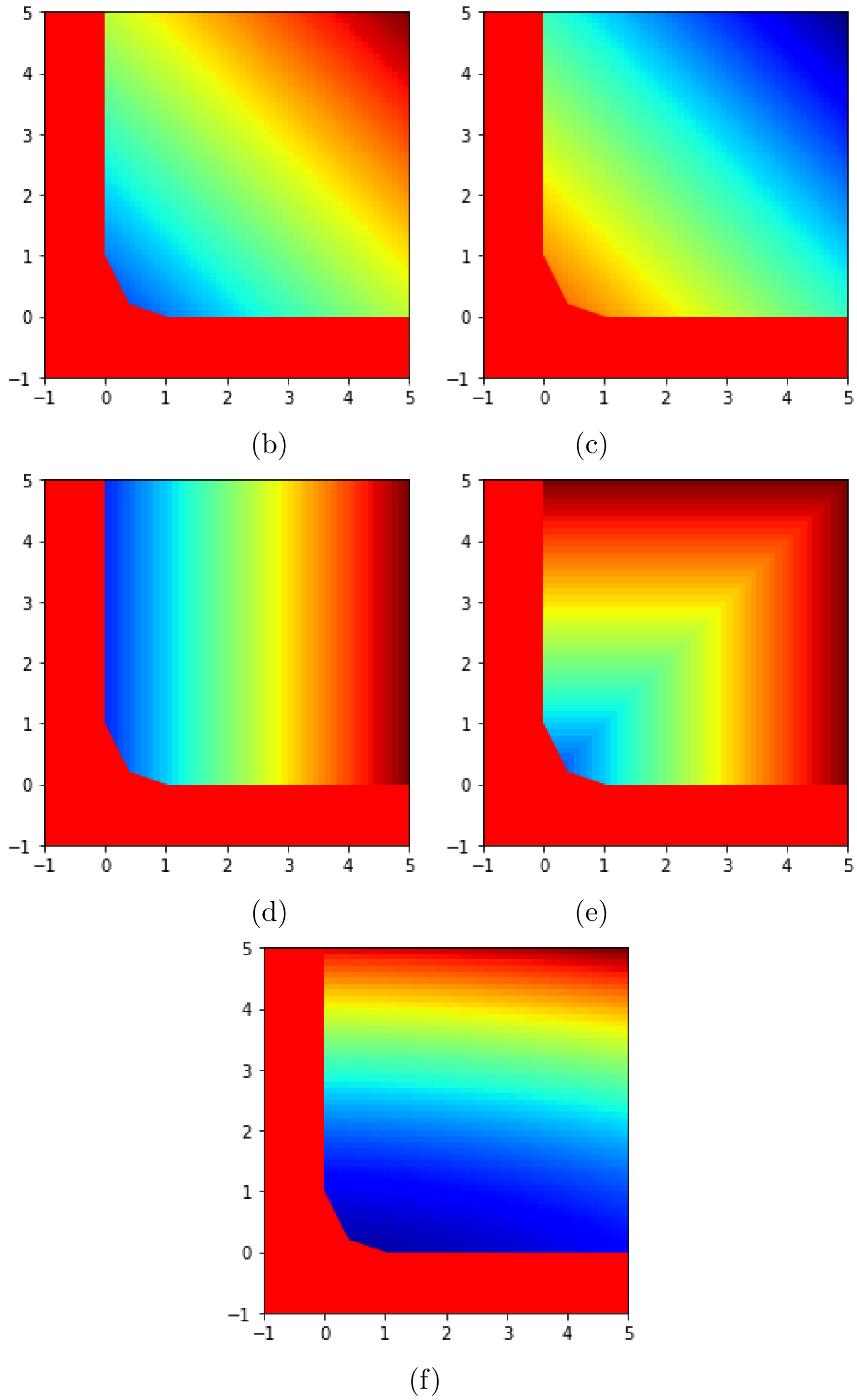

Figure 2: Solution of 2(b): $\vec{x}^\star = (\frac{2}{5}, \frac{1}{5})$, 2(c) is unbounded below, solutions of 2(d): $\vec{x}^\star = \{(0, x_2) \mid x_2 \geq 1\}$, solution of 2(e): $\vec{x}^\star = (\frac{1}{4}, \frac{1}{4})$, solution of 2(f): $\vec{x}^\star = (\frac{2}{5}, \frac{1}{5})$. In red is the unfeasible points, then the level sets are shown with colors; blue points are points (x_1, x_2) with the lowest value $f(x_1, x_2)$, red points are the ones with highest value.

(d) $f(x_1, x_2) = x_1$

Solution: The set of solutions is $S = \{\vec{x}, x_1 = 0 \text{ and } x_2 \geq 1\}$ (see figure 2).

(e) $f(x_1, x_2) = \max\{x_1, x_2\}$

Solution: Using the drawing (see figure 2) it seems that the solution is such that:

$$x_1^\star = x_2^\star = \frac{1}{3}$$

.

Here, it might be hard to use the first order convexity condition, as the objective function is not differentiable (you can use sub-gradients, but it is beyond the scope of class).

Another technique is to use a slack variable. The problem is equivalent to

$$\begin{aligned} &\min_{x_1, x_2, t} t \\ \text{subject to } & t \geq x_1, t \geq x_2 \\ & 2x_1 + x_2 \geq 1 \\ & x_1 + 3x_2 \geq 1 \\ & x_1 \geq 0, \ x_2 \geq 0 \end{aligned}$$

Here we can use the first order condition to check optimality conditions.

(f) $f(x_1, x_2) = x_1^2 + 9x_2^2$

Solution: Using the drawing (see figure 2) it seems that the solution is such that $x_1^\star = \frac{1}{2}$ and $x_2^\star = \frac{1}{6}$.

It can be verified using the first order convexity condition:

$$\nabla f(\frac{1}{2}, \frac{1}{6})^\top ((\frac{1}{2}, \frac{1}{6}) - (x_1, x_2)) \geq 0, \quad \forall (x_1, x_2) \in \mathcal{X}$$

Where $\mathcal{X}$ is the feasible set.

2. About general optimization

In this exercise, we test your understanding of the general framework of optimization and its language. We consider an optimization problem in standard form:

$$p^* = \min_{\vec{x} \in \mathbb{R}^n} f_0(\vec{x}) \ : \ f_i(\vec{x}) \leq 0, \ \ i = 1, \ldots, m.$$

In the following we denote by $\mathcal{X}$ the feasible set. Note that the feasible set is a subset of R^n that satisfies the inequalities $f_i(\vec{x}) \leq 0$, i.e $\mathcal{X} = \{\vec{x} \in \mathbb{R}^n \mid f_i(\vec{x}) \leq 0, i = 1, \ldots, m\}$. We make no assumption about the convexity of $f_0(\vec{x})$ and $f_i(\vec{x})$, $i = 1, \ldots, m$. For the following statements, provide a proof or counter-example.

(a) A general optimization problem can be expressed as one with a linear objective.

Solution: The statement is true:

$$p^* = \min_{\vec{x} \in \mathcal{X}, t} t \ : \ t \geq f_0(\vec{x}).$$

t is called a slack variable.

(b) A general optimization problem can be expressed as one without any constraints.

Solution: Again the statement is true: let us define

$$g(\vec{x}) := \begin{cases} f_0(\vec{x}) & \text{if } \vec{x} \in \mathcal{X}, \\ +\infty & \text{otherwise.} \end{cases}$$

Then

$$p^* = \min_{\vec{x}} \ g(\vec{x}).$$

Remark that if you define the indicator function of the complement of the set $\mathcal{X}$ as:

$$1_{\Omega \setminus \mathcal{X}}(\vec{x}) = \begin{cases} 0 & \text{if } \vec{x} \in \mathcal{X}, \\ 1 & \text{otherwise.} \end{cases}$$

Then, you can write g as:

$$g(\vec{x}) = \max_{\mu} f_0(\vec{x}) + \mu 1_{\Omega \setminus \mathcal{X}}(\vec{x})$$

And

$$p^* = \min_{\vec{x}} \ \max_{\mu} f_0(\vec{x}) + \mu 1_{\Omega \setminus \mathcal{X}}(\vec{x}).$$

(c) If at the optimal point $\vec{x}^\star$, one constraint is not active (i.e. $f_i(\vec{x}^\star) < 0$), then we can remove the constraint from the original problem and obtain the same optimum value.

Solution: This is *not* true in general. Consider the problem

$$p^* := \min_{x} \ f_0(x) \ : \ |x| \leq 1,$$

where

$$f_0(x) = \begin{cases} x^2 & \text{if } |x| \leq 1, \\ -1 & \text{otherwise.} \end{cases}$$

The constraint $|x| \leq 1$ is not active at the optimum $x^* = 0$, and $p^* = 0$. However, if we remove it, the new optimal value becomes -1.

(d) If the problem is convex, and at the optimal point $\vec{x}^\star$, one constraint is not active ($f_i(\vec{x}^\star) < 0$), then we can remove the constraint from the original problem and obtain the same optimum value.

Assume that the minimum is attained for some $\vec{x}^* \in \mathbb{R}^n$.

Solution: The statement is true. Without loss of generality, we may assume we have no inequality constraints that are active at optimum.

Note the optimal solution $\vec{x}_1^*$. To see this, suppose

$$\begin{aligned} \vec{x}_1^* = \arg\min_{\vec{x}} \ & f_0(\vec{x}) \\ \text{s.t.} \ & f_i(\vec{x}) \leq 0, \ i = 1, ..., m \end{aligned}$$

and that at optimum, $\mathcal{I} \subseteq \{1, ..., m\}$ is an index set such that for each $i \in \mathcal{I}$ we have $f_i(\vec{x}_1^*) < 0$. Then define

$$\tilde{f}_0(\vec{x}) = \begin{cases} f_0(\vec{x}) & \text{if } \vec{x} \in \mathcal{X}, \\ +\infty & \text{otherwise.} \end{cases}$$

where $\mathcal{X} = \cap_{i \in \mathcal{I}^c} \{\vec{x} \mid f_i(\vec{x}) \le 0\}$. Note that $\tilde{f}_0(\vec{x})$ is convex. Hence the original optimization writes

$$\begin{aligned} p_1^* = \min_{\vec{x}} \ & \tilde{f}_0(\vec{x}) \\ \text{s.t. } & f_i(\vec{x}) \le 0, \ i \in \mathcal{I} \end{aligned}$$

where all the inequality constraints are not active at optimum.

Now we claim the above is equivalent to $p_2^* = \min_{\vec{x}} \tilde{f}_0(\vec{x})$. To show this, we use the fact that the problem defining p_2^* is convex which implies that all local minima are global minima. Adding a constraint that is inactive at optimum implies that the global minima of problem defining p_2^* are unaffected and hence is still optimum for the constrained problem (i.e., the problem defining p_1^*). Note this is not the case with non-convex optimization – if you remove an inactive constraint, you may have $p_2^* < p_1^*$.

To rigorously show this, for contradiction suppose that $\vec{x}_1$ is optimal for the problem defining p_1^* and that $\vec{x}_2$ is optimal for the problem defining p_2^* and suppose further that $\tilde{f}(\vec{x}_2) = p_2^* < p_1^* = \tilde{f}(\vec{x}_1)$ (note we automatically have $p_2^* \le p_1^*$). Then since $\tilde{f}_0(\vec{x})$ is convex, we have that

$$\begin{aligned} \tilde{f}(\lambda\vec{x}_1 + (1-\lambda)\vec{x}_2) &\le \lambda\tilde{f}(\vec{x}_1) + (1-\lambda)\tilde{f}(\vec{x}_2) && \text{(Convexity of } \tilde{f}) \\ &< \lambda\tilde{f}(\vec{x}_1) + (1-\lambda)\tilde{f}(\vec{x}_1) && (\tilde{f}(\vec{x}_2) < \tilde{f}(\vec{x}_1)) \\ &= \tilde{f}(\vec{x}_1) \\ &= f(\vec{x}_1) && \text{(assuming optimal } \vec{x}_1 \in \mathcal{X}) \end{aligned}$$

which holds for all $\lambda \neq 1$. We can choose λ such that the point $\lambda\vec{x}_1 + (1-\lambda)\vec{x}_2$ lies in the local neighbourhood of $\vec{x}_1$ to get a contradiction since $\vec{x}_1$ is a local minima.

3. Convexity and composition of functions

Let $f : \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R}^n \to \mathbb{R}$. Define the composition of f with g as $h = f \circ g : \mathbb{R}^n \to \mathbb{R}$ such that $h(\vec{x}) = f(g(\vec{x}))$.

(a) Show that if f is convex and non decreasing and g is convex, then h is convex.

Solution:

$$\begin{aligned} h(\lambda\vec{x} + (1-\lambda)\vec{y}) &= f(g(\lambda\vec{x} + (1-\lambda)\vec{y})) \\ &\le f(\lambda g(\vec{x}) + (1-\lambda)(g(\vec{y}))) && (g \text{ convex and } f \text{ nondecreasing}) \\ &\le \lambda f(g(\vec{x})) + (1-\lambda)f(g(\vec{y})) && (f \text{ convex}) \\ &= \lambda h(\vec{x}) + (1-\lambda)h(\vec{y}) \end{aligned}$$

So h is convex.

(b) Show that there exists f non decreasing and g convex, such that $h = f \circ g$ is not convex.

Solution: Take $n = 1$, $f(x) = \log(x)$ and $g(x) = x$. Then $h(x) = \log(x)$ is not convex.

(c) Show that there exists f convex and g convex such that $h = f \circ g$ is not convex.

Solution: Take $n = 1$, $f(x) = -x$ and $g(x) = x^2$, then $h(x) = -x^2$ is not convex.