The University of Texas at Austin
Department of Electrical and Computer Engineering

EE381K: Convex Optimization — Fall 2019

LECTURE 14

Aryan Mokhtari Wednesday, October 16, 2019.

Goal: In this lecture, we talk about optimality conditions for constrained and unconstrained convex optimization problems.

1 Convex Optimization

Now let us revisit the basic general optimization problem we saw at the beginning of the course:

$$\begin{aligned} \min: \quad & f_0(\mathbf{x}) \\ \text{s.t.}: \quad & f_i(\mathbf{x}) \leq 0, \qquad i = 1, \ldots, m. \\ & h_i(\mathbf{x}) = 0, \qquad i = 1, \ldots, p. \end{aligned}$$

$\mathbf{x}$ is **feasible** if $\mathbf{x} \in \mathbf{dom} f_0$ and $\mathbf{x}$ satisfies the constraints.

Optimal value $p^* = \inf\{f_0(\mathbf{x}) \mid \mathbf{x} \in \mathbf{dom} f_0,\ f_i(\mathbf{x}) \leq 0, i = 1, \ldots, m,\ h_i(\mathbf{x}) = 0, i = 1, \ldots, p\}$

$\mathbf{x}^*$ is **(globally) optimal** if $\mathbf{x}^*$ is feasible and $f(\mathbf{x}^*) = p^*$.

X_{opt} is the **set of optimal solutions**

A feasible point $\tilde{\mathbf{x}}$ is **locally optimal** if there is an $R > 0$ such that

$$f_0(\tilde{\mathbf{x}}) = \inf\{f_0(\mathbf{x}) \mid \mathbf{x} \in \mathbf{dom} f_0,\ f_i(\mathbf{x}) \leq 0, i = 1, \ldots, m,\ h_i(\mathbf{x}) = 0, i = 1, \ldots, p,\ \|\mathbf{x} - \tilde{\mathbf{x}}\|_2 \leq R\}$$

Standard form convex optimization problem

$$\begin{aligned} \min: \quad & f_0(\mathbf{x}) \\ \text{s.t.}: \quad & f_i(\mathbf{x}) \leq 0, \qquad i = 1, \ldots, m. \\ & \mathbf{a}_i^\top \mathbf{x} + b_i = 0, \qquad i = 1, \ldots, p. \end{aligned}$$

$f_0, f_1, \ldots, f_m$ are convex, equality constraints are affine. The constraint set is a convex set.

Theorem 1. *Any locally optimal point of a convex optimization problem is also (globally) optimal.*

Proof: Suppose $\tilde{\mathbf{x}}$ is a locally optimal point. Then, $\tilde{\mathbf{x}}$ is feasible and

$$f_0(\tilde{\mathbf{x}}) = \inf\{f_0(\mathbf{x}) \mid \mathbf{x} \text{ is feasible },\ \|\mathbf{x} - \tilde{\mathbf{x}}\|_2 \leq R\}$$

Now suppose that x is not globally optimal, i.e., there exists a feasible point $\mathbf{y}$ such that $f_0(\mathbf{y}) < f_0(\tilde{\mathbf{x}})$. Evidently, $\|\mathbf{y} - \tilde{\mathbf{x}}\| > R$. Now consider the point $\mathbf{z}$ defined as $\mathbf{z} = \theta \mathbf{y} + (1 - \theta)\tilde{\mathbf{x}}$ where $\theta = R/(2\|\mathbf{y} - \tilde{\mathbf{x}}\|)$. Thus, $\|\mathbf{z} - \mathbf{x}\| = R/2 < R$ and by convexity of the feasible set $\mathbf{z}$ is feasible. By convexity of f_0 we have

$$f_0(\mathbf{z}) \leq \theta f_0(\mathbf{y}) + (1 - \theta) f_0(\tilde{\mathbf{x}}) < f_0(\tilde{\mathbf{x}})$$

which contradicts with the fact that $\tilde{\mathbf{x}}$ is a local minimum. Hence, $\tilde{\mathbf{x}}$ is also global minimum.

2 Optimality conditions

We now give optimality conditions for unconstrained optimization. These are the familiar first-order conditions of optimality for convex optimization. While simple, the condition is extremely useful both algorithmically and for analysis.

2.1 Constrained optimization

First we consider a general constrained convex optimization problem.

$$\begin{aligned} &\min f(\mathbf{x}) \\ &\text{s.t } \mathbf{x} \in \mathcal{X}, \end{aligned}$$

where $f : \mathcal{X} \to \mathbb{R}$ is convex and differentiable and the set $\mathcal{X}$ is convex. Then a point $\mathbf{x}^*$ is optimal if and only if $\mathbf{x} \in \mathcal{X}$ and

$$\nabla f(\mathbf{x}^*)^\top (\mathbf{y} - \mathbf{x}^*) \geq 0, \qquad \text{for all } \mathbf{y} \in \mathcal{X}$$

What does it mean?
- If you move from $\mathbf{x}^*$ towards any feasible $\mathbf{y}$, you will increase f locally
- It means that $-\nabla f(\mathbf{x}^*)$ defines a supporting hyperplane to the feasible set at $\mathbf{x}^*$

Proof:
(i) Suppose $\mathbf{x}^*$ satisfies the condition

$$\nabla f(\mathbf{x}^*)^\top (\mathbf{y} - \mathbf{x}^*) \geq 0, \qquad \text{for all } \mathbf{y} \in \mathcal{X}.$$

Further, by convexity of function f we have

$$f(\mathbf{y}) \geq f(\mathbf{x}^*) + \nabla f(\mathbf{x}^*)^\top (\mathbf{y} - \mathbf{x}^*) \qquad \text{for all } \mathbf{y} \in \mathcal{X}$$

Combining these two results we can write that

$$f(\mathbf{y}) \geq f(\mathbf{x}^*) \qquad \text{for all } \mathbf{y} \in \mathcal{X}$$

and therefore $\mathbf{x}^*$ is an optimal point.

(ii) Now suppose $\mathbf{x}^*$ is an optimal solution. We prove the optimality condition with contradiction. Suppose there exists a feasible $\mathbf{y}$ such that

$$\nabla f(\mathbf{x}^*)^\top (\mathbf{y} - \mathbf{x}^*) < 0$$

Now consider the feasible point $\mathbf{x}^* + \alpha(\mathbf{y} - \mathbf{x}^*)$ for $\alpha \in [0, 1]$ and define the function $g : \mathbb{R} \to \mathbb{R}$ as $g(\alpha) = f(\mathbf{x}^* + \alpha(\mathbf{y} - \mathbf{x}^*))$. Note that the first derivative of g is given by

$$g'(\alpha) = \nabla f(\mathbf{x}^* + \alpha(\mathbf{y} - \mathbf{x}^*))^\top (\mathbf{y} - \mathbf{x}^*)$$

Hence, we have

$$g'(0) = \nabla f(\mathbf{x}^*)^\top (\mathbf{y} - \mathbf{x}^*) < 0.$$

If $g(0) = f(\mathbf{x}^*)$ and $g'(0) < 0$ then for small positive α we have

$$f(\mathbf{x}^* + \alpha(\mathbf{y} - \mathbf{x}^*)) = g(\alpha) < g(0) = f(\mathbf{x}^*).$$

But this contradicts the optimality of $\mathbf{x}^*$.

2.2 Unconstrained optimization

The formulation of an unconstrained optimization problem is as follows

$$\min : f(\mathbf{x})$$

where $f : \mathbb{R}^n \to \mathbb{R}$ is differentiable and convex. In these problems, the necessary and sufficient condition for the optimal solution $\mathbf{x}^*$ is

$$\nabla f(\mathbf{x}^*) = \mathbf{0} \quad \text{at } \mathbf{x} = \mathbf{x}^*. \tag{1}$$

Proof. (i) Suppose $\mathbf{x}^*$ satisfies $\nabla f(\mathbf{x}^*) = 0$. Recall that if f is convex, we have

$$f(\mathbf{y}) \geq f(\mathbf{x}^*) + \nabla f(\mathbf{x}^*)^T (\mathbf{y} - \mathbf{x}^*) \text{ for all } \mathbf{y} \in dom(f).$$

At $\mathbf{x}^*$, $\nabla f(\mathbf{x}^*) = \mathbf{0}$ and hence this reduces to

$$f(\mathbf{y}) \geq f(\mathbf{x}^*) \text{ for all } \mathbf{y} \in dom(f),$$

which indeed is the statement that $\mathbf{x}^*$ is a global optimum.
(ii) Now suppose $\mathbf{x}^*$ is an optimal solution. From previous theorem $\nabla f(\mathbf{x}^*)^\top (\mathbf{y} - \mathbf{x}^*) \geq 0$ for any $\mathbf{y} \in \mathbb{R}^n$. Since with $\mathbf{y} - \mathbf{x}^*$ we can make any direction, $\nabla f(\mathbf{x}^*) = \mathbf{0}$. (Take $\mathbf{y} = \mathbf{x}^* - \alpha \nabla f(\mathbf{x}^*)$). □

Note of course that this need not be unique. For uniqueness, we need a stronger version of convexity. This will be addressed later.

Theorem 2. *If there exists an interior point* $\mathbf{x}$ *of the feasible set* $\mathcal{X}$ *such that* $\nabla f(\mathbf{x}) = 0$*, then* $\mathbf{x}$ *is the optimal point of the constrained problem.*

For a constrained problem, though, this is not typically the case, as it means the constraints are not *binding*. Indeed, typically there is no feasible point $\mathbf{x}$ satisfying $\nabla f(\mathbf{x}) = 0$. The generalization of the local optimality condition to the constrained case is natural. The condition $\nabla f(\mathbf{x}) = 0$, recall, is really saying that there are no descent directions from the point $\mathbf{x}$, i.e., $\langle \nabla f(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle \geq 0$, for all$\mathbf{y}$ (which as noted above, can only happen when $\nabla f(\mathbf{x}) = 0$). In the constrained case, we need to check that this inequality holds for all feasible directions.

2.3 Examples

Now we study a few examples of convex optimization and use optimality condition to find their solution.

- **Unconstrained quadratic optimization.** Consider the problem of minimizing the quadratic function

$$f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^\top \mathbf{A}\mathbf{x} + \mathbf{b}^\top \mathbf{x} + c$$

where $\mathbf{A} \in \mathbf{S}_+^n$ which makes f convex. The necessary and sufficient condition for $\mathbf{x}^*$ to be a minimizer of f is

$$\nabla f(\mathbf{x}^*) = \mathbf{A}\mathbf{x}^* + \mathbf{b} = \mathbf{0}$$

If $\mathbf{A} \succ \mathbf{0}$ then we have $\mathbf{x}^* = -\mathbf{A}^{-1}\mathbf{b}$
for singular $\mathbf{A}$, if $\mathbf{b} \in \mathcal{R}(\mathbf{A})$, then the set of optimal points is the (affine) set $X_{opt} = -\mathbf{A}^\dagger \mathbf{b} + \mathcal{N}(\mathbf{A})$
for singular $\mathbf{A}$, if $\mathbf{b} \notin \mathcal{R}(\mathbf{A})$, then the set of optimal points is empty and f is unbounded below.

- **Problems with equality constraints only.** Consider the problem

$$\min f(\mathbf{x}) \qquad s.t.\ \mathbf{A}\mathbf{x} = \mathbf{b}$$

Then optimality condition for a feasible $\mathbf{x}$ is that

$$\nabla f(\mathbf{x})^\top (\mathbf{y} - \mathbf{x}) \geq 0 \qquad \text{for all } \mathbf{A}\mathbf{y} = \mathbf{b}$$

Since $\mathbf{x}$ is feasible, every feasible $\mathbf{y}$ has the form $\mathbf{y} = \mathbf{x} + \mathbf{v}$ for some $\mathbf{v} \in \mathcal{N}(\mathbf{A})$. Hence, we can rewrite this condition as

$$\nabla f(\mathbf{x})^\top \mathbf{v} \geq 0 \qquad \text{for all } \mathbf{v} \in \mathcal{N}(\mathbf{A})$$

If a linear function is nonnegative on a subspace, then it must be zero on the subspace, so

$$\nabla f(\mathbf{x})^\top \mathbf{v} = 0 \qquad \text{for all } \mathbf{v} \in \mathcal{N}(\mathbf{A})$$

Hence, $\nabla f(\mathbf{x}) \perp \mathcal{N}(\mathbf{A})$. Using the fact that $\mathcal{N}(\mathbf{A})^\perp = \mathcal{R}(\mathbf{A}^\top)$ we can write the optimality condition as

$$\nabla f(\mathbf{x}) \in \mathcal{R}(\mathbf{A}^\top)$$

i.e., there exists a $\mathbf{v}$ such that

$$\nabla f(\mathbf{x}) + \mathbf{A}^\top \mathbf{v} = \mathbf{0}$$