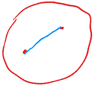
MAT 3007 Optimization: Tutorial 12

Guxin DU

The Chinese University of Hong Kong, Shenzhen

July 10, 2025

Recap: Convex Problems



Definition 1 (Convex set).

The set $S \subset \mathbb{R}^n$ is convex if for $\forall\ \mathbf{x}, \mathbf{y} \in S$ and $\forall\ \lambda \in [0, 1]$, we have $\lambda\mathbf{x} + (1 - \lambda)\mathbf{y} \in S$.

Definition 2 (Convex function).

A function $f : \mathbb{R}^n \to \mathbb{R}$ is convex if

(1) its domain Ω is convex and

(2) $\forall\ \mathbf{x}_1, \mathbf{x}_2 \in \Omega$ and $\forall\ \alpha \in [0, 1]$ satisfy

$$f(\alpha\mathbf{x}_1 + (1 - \alpha)\mathbf{x}_2) \leq \alpha f(\mathbf{x}_1) + (1 - \alpha)f(\mathbf{x}_2).$$

Definition 3 (Concave function).

A function g is concave if $-g$ is convex.

Recap: Convex Problems

Theorem 4 (Characterization of convex differentiable functions).

Suppose a function $f : \mathbb{R}^n \to \mathbb{R}$ is twice differentiable on Ω, then the following are equivalent:

(1) *f is convex*

(2) $f(\mathbf{x}_2) \geq f(\mathbf{x}_1) + \nabla f(\mathbf{x}_1)^T(\mathbf{x}_2 - \mathbf{x}_1)$ *for* $\forall\ \mathbf{x}_1, \mathbf{x}_2 \in \Omega$

(3) $\nabla^2 f(\mathbf{x}) \succeq 0,\ \forall\ \mathbf{x} \in \Omega$

Remark: First order characterization of convexity implies that the stationary point is global minimal.

e.g.1 $f(\mathbf{x}) = a^T\mathbf{x} + b$ is convex and concave.

e.g.2 $f(\mathbf{x}) = \mathbf{x}^T Q\mathbf{x} + c^T\mathbf{x} + d$ is convex if and only if $Q \succeq 0$.

$$\forall x, y \in \Omega \quad f(y) \geq f(x) + \nabla f^T(x)(y-x)$$

$2 \Rightarrow 1$

Any $z = \alpha x + (1-\alpha) y$ for $\alpha \in [0,1]$

We want to show that

$$f(z) \leq \alpha f(x) + (1-\alpha) f(y)$$

$$1-\alpha \times \quad \underline{f(y) \geq f(z) + \nabla f^T(z)(y-z)}$$

$+$

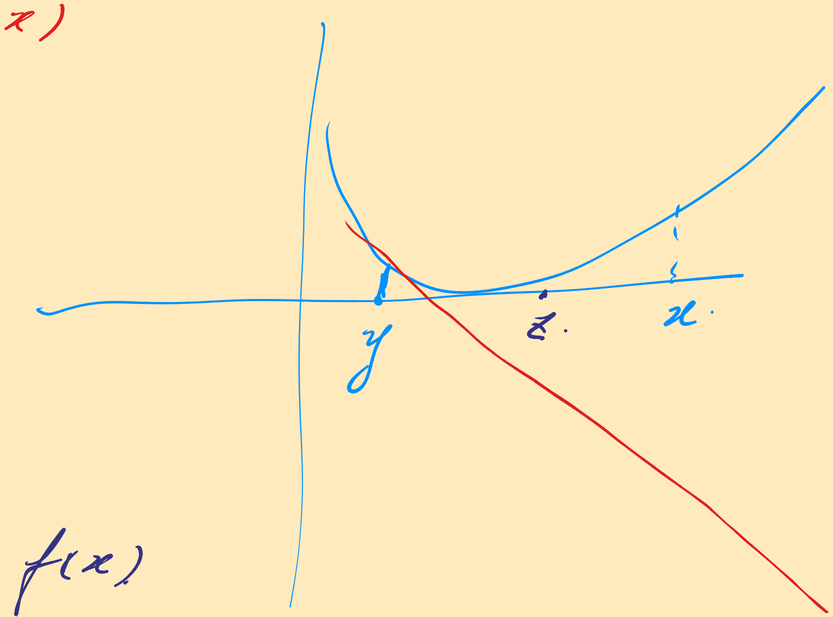
$$\alpha \times \quad f(x) \geq f(z) + \nabla f^T(z)(x-z)$$

~~$1-\alpha \quad z \quad \alpha z$~~

$$(1-\alpha) f(y) + \alpha f(x) \geq f(z) + \cancel{\nabla f^T(z)\Big[\underbrace{(1-\alpha)(y-z) + \alpha(x-z)}_{-z + (1-\alpha)y + \alpha x \,=\, 0}\Big]}$$

(1) $\Rightarrow$ (2) We are given $f(x)$ convex.

Goal: Prove $f(y) \geq f(x) + \nabla f^T(x)(y-x)$



$z = y + \alpha(x-y), \quad \alpha \in [0,1]$

$f(y+\alpha(x-y)) \leq \alpha f(x) + (1-\alpha) f(y)$

$\alpha f(y) + [f(y+\alpha(x-y)) - f(y)] \leq \alpha f(x)$

$f(y) + \dfrac{f(y+\alpha(x-y)) - f(y)}{\alpha} \leq f(x)$ holds for all $\alpha \in (0,1]$

$\downarrow \alpha \to 0$

$f(y) + \nabla f^T(y)(x-y) \leq f(x)$

Proof of the First Order Characterization

Proof.

$\Leftarrow$ let we set $z = \lambda x + (1-\lambda)y$, then we want to prove

$$\lambda f(x) + (1-\lambda)f(y) \geq f(\lambda x + (1-\lambda)y) = f(z).$$

We have

$$f(x) \geq f(z) + \nabla f(z)^T(x-z)$$
$$f(y) \geq f(z) + \nabla f(z)^T(y-z)$$

Let the first inequality times λ and the second one times $1-\lambda$, we will get the ideal result. $\square$

Proof of the First Order Characterization

$\Rightarrow$ let we assume f is convex and for any $x \neq z$, we define the following function $g : (0,1] \to \mathbb{R}$.

$$g(\alpha) = \frac{f(x + \alpha(z-x)) - f(x)}{\alpha}, \quad \alpha \in (0,1]$$

If we can prove $g(\alpha)$ is monotonically increasing, then

$$g(1) = f(z) - f(x) \geq g(0) = \nabla f(x)^T (z - x).$$

Suppose $0 < \alpha_1 < \alpha_2$, let $\bar{\alpha} = \frac{\alpha_1}{\alpha_2}$, $\bar{z} = x + \alpha_2(z - x)$.Then

$$f(x + \bar{\alpha}(\bar{z} - x)) \leq \bar{\alpha} f(\bar{z}) + (1 - \bar{\alpha}) f(x)$$

$$\text{i.e.} \quad \frac{f(x + \bar{\alpha}(\bar{z} - x)) - f(x)}{\bar{\alpha}} \leq f(\bar{z}) - f(x)$$

This equals to $g(\alpha_1) \leq g(\alpha_2)$.

Theorem 5.

As a proposition, a convex differentiable function f has an optimal point at x^* on convex set Ω if and only if

$$\nabla f(x^*)^T(z - x^*) \geq 0, \forall z \in \Omega$$

Sufficiency: Directly from the first order chracterization.
Necessity: FONC for constrained problems:

$$S_\Omega(x^*) \cap S_D(x^*) = \emptyset.$$

Review: prove by contradiction, suppose for some direction z, we have

$$\lim_{\alpha \downarrow 0} \frac{f(x^* + \alpha(z - x^*)) - f(x^*)}{\alpha} = \nabla f(x^*)^T(z - x^*) < 0.$$

By the continuity of $g(\alpha)$, ... (finish the proof by yourself)

Recap on properties

Theorem 6 (Composition with ~~linear~~ affine function).

Suppose a function f is convex, then $f(A\mathbf{x}+b)$ is a convex function. (Similiar version for concave functions)

Theorem 7 (max of convex function is convex).

Suppose functions $(f_i)_{i\in I}$ is a set of convex functions where I is a finite index set, then $f(x) = max\{f_i(x)|i \in I\}$ is a convex function. (Note: it takes max over I pointwisely) (it can be extended to uncountably many set I)

Theorem 8 (min of concave function is concave).

Suppose functions $(f_i)_{i\in I}$ is a set of concave functions where I is a finite index set, then $f(x) = min\{f_i(x)|i \in I\}$ is a concave function. (Note: it takes min over I pointwisely)

Some Proof

1. Linear Composition:

$$\begin{aligned}
& f\left(A(\lambda x+(1-\lambda)y)+b\right) \\
=& f\left(\lambda(Ax+b)+(1-\lambda)(Ay+b)\right) \\
\leq& \lambda f(Ax+b)+(1-\lambda)f(Ay+b)
\end{aligned}$$

$\sup_i (a_i + b_i) \leq \sup_i a_i + \sup_j b_j$

2. Taking maximum:

$$\begin{aligned}
& \sup_i f_i(\lambda x+(1-\lambda)y) \\
\leq& \sup_i \lambda f_i(x)+\sup_i (1-\lambda)f_i(y) \\
=& \lambda \sup_i f_i(x)+(1-\lambda)\sup_i f_i(y)
\end{aligned}$$

$p^*_{\tilde{c}} \underset{\geq}{\leq} \alpha p^*_{c_1} + (1-\alpha) p^*_{c_2}$?

Exercise 1

$c_1 \quad c_2 \quad \alpha \in [0,1]$

$p^*_{c_1} \quad p^*_{c_2} \quad p^*_{\tilde{c}} \quad \tilde{c} = \alpha c_1 + (1-\alpha) c_2$

Consider the following linear program

$$
p^* = \begin{array}{rl} \min\limits_{x} & \mathbf{c}^\top \mathbf{x} \\ \text{s.t.} & A\mathbf{x} \leq b. \end{array} \qquad = \begin{array}{rl} \max\limits_{y} & b^T y \\ & A^T y = c \\ & y \leq 0 \end{array}
$$

Let p^* denote its optimal value.

- Is p^* convex or concave with c ?
- Is p^* convex or concave with b ?

$$
p^*_{\tilde{c}} = \min_x (\alpha c_1 + (1-\alpha) c_2)^T x \geq \min_x \alpha c_1^T x + \min_x (1-\alpha) c_2^T x
$$

$$
= \alpha \underbrace{\min_x c_1^T x}_{p^*_{c_1}} + (1-\alpha) \underbrace{\min_x c_2^T x}_{p^*_{c_2}}
$$

Exercise 1: Solution

$$p^* = \min_{\{\mathbf{x}|A\mathbf{x}\le b\}} \mathbf{c}^T\mathbf{x} \qquad \text{denote } \Omega := \{\mathbf{x}|A\mathbf{x}\le b\}$$

(1) We pick $\mathbf{c}_1$ $\mathbf{c}_2$. Then, for any $\mathbf{y}\in\Omega$, we have

$$\min_{\Omega}\ \mathbf{c}_1^T\mathbf{x} \le \mathbf{c}_1^T\mathbf{y} \qquad \min_{\Omega}\ \mathbf{c}_2^T\mathbf{x} \le \mathbf{c}_2^T\mathbf{y}$$

Thus, for $\alpha\in[0,1]$ and $\forall\,\mathbf{y}\in\Omega$,

$$\alpha(\min_{\Omega}\ \mathbf{c}_1^T\mathbf{x}) + (1-\alpha)(\min_{\Omega}\ \mathbf{c}_2^T\mathbf{x}) \le \alpha(\mathbf{c}_1^T\mathbf{y}) + (1-\alpha)(\mathbf{c}_2^T\mathbf{y}).$$

So,

$$\alpha(\min_{\Omega}\ \mathbf{c}_1^T\mathbf{x}) + (1-\alpha)(\min_{\Omega}\ \mathbf{c}_2^T\mathbf{x}) \le \min_{\Omega}\ (\alpha\mathbf{c}_1^T\mathbf{y} + (1-\alpha)\mathbf{c}_2^T\mathbf{y}).$$

Hence, p^* is concave w.r.t. $\mathbf{c}$.

(2) By considering dual problem, we obtain convexity of p^* w.r.t. b by using same techniques.

Thanks for coming!