Calculus IB: Lecture 14

Luo Luo

Department of Mathematics, HKUST

http://luoluo.people.ust.hk/

Outline

Outline

Optimization Problems

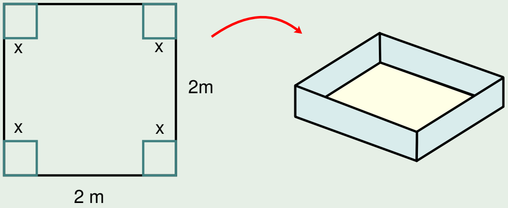
Mathematical optimization (alternatively spelled optimisation) is the selection of a best element (with regard to some criterion) from some set of available alternatives.

Optimization problems of sorts arise in all quantitative disciplines from computer science and engineering to operations research and economics, and the development of solution methods has been of interest in mathematics for centuries.

In the simplest case, an optimization problem consists of finding the maximum or minimum of a certain function to meet certain purpose, depending on the requirement of the problem.

Example (volume of the box)

If an open-top box is to be made by cutting four square corners from a 2m by 2m sheet of tin, what is the largest possible volume of the box?



We write the volume of the box as a function of x:

$$V(x) = x(2-2x)^2 = 4x - 8x^2 + 4x^3$$

where $0 \leq x \leq 1$. We want to find the maximum of V.

Optimization Problems

Example (volume of the box)

We first find the critical point of $V(x) = 4x - 8x^2 + 4x^3$:

$$\frac{dV}{dx} = 4 - 16x + 12x^2 = 0 \iff x = 1 \quad \text{or} \quad x = \frac{1}{3}.$$

By comparing the function values at critical points and endpoints

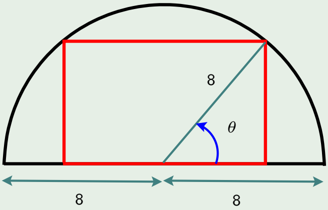
$$V(1) = 0, \quad V(0) = 0, \quad V\left(\frac{1}{3}\right) = \frac{16}{27},$$

the maximum volume is $\frac{16}{27}\text{m}^3$, which reached by cutting $\frac{1}{3}$m.

Optimization Problems

Example (rectangle in semi-circle)

What is the largest area of rectangle inscribed in a semi-circle of radius 8?



The area of the inscribed rectangle as a function of the angle θ, where $0 \leq \theta \leq \dfrac{\pi}{2}$ is:

$$A(\theta) = 2(8\cos\theta)(8\sin\theta) = 128\cos\theta\sin\theta = 64\sin 2\theta,$$

where we use the identity $\sin 2\theta = 2\sin\theta\cos\theta$.

Optimization Problems

Example (rectangle in semi-circle)

We first find the critical point of $A(\theta) = 64 \sin 2\theta$:

$$A'(\theta) = 128 \cos 2\theta = 0 \Longleftrightarrow 2\theta = \frac{\pi}{2} \Longleftrightarrow \theta = \frac{\pi}{4}.$$

By comparing the function values at critical points and endpoints

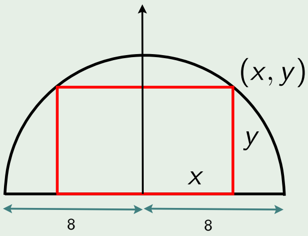
$$A(0) = 0 = A\left(\frac{\pi}{2}\right), \qquad A\left(\frac{\pi}{4}\right) = 64,$$

the maximum area is 64, which reached when $\theta = \dfrac{\pi}{4}$ and the rectangle is a square of base length $2 \cdot \left(8 \cos \frac{\pi}{4}\right) = 8\sqrt{2}$.

Example (rectangle in semi-circle)

We can also formulate the problem with

$$A(x) = 2xy = 2x\sqrt{8^2 - x^2}, \text{ where } 0 \le x \le 8.$$



We find the critical point by computing the derivative

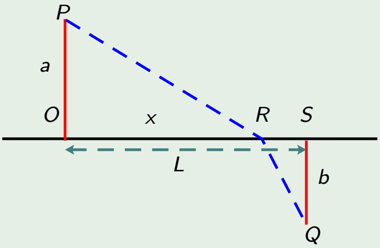
$$\frac{dA}{dx} = 2\sqrt{64 - x^2} - \frac{2x^2}{\sqrt{64 - x^2}} = \frac{128 - 4x^2}{\sqrt{64 - x^2}} = 0 \Longrightarrow x = \sqrt{32} = 4\sqrt{2}.$$

The maximum area is reached at $x = \sqrt{32} = 4\sqrt{2}$, i.e., $A(4\sqrt{2}) = 64$.

Optimization Problems

Example (Snell's Law)

The ground surfaces on two sides of a certain boundary line are made of different materials.
A man can run with speed v_1 m/s on the side of location P which a m from the line, and v_2 m/s on the side of location Q which is b m from the line.
Running along a straight path on either side, how should he choose the cross-boundary point R in order to reach location Q as soon as possible?

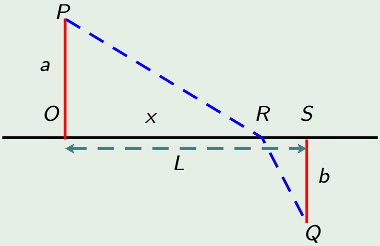


Example (Snell's Law)

Let $OR = x$, $OS = L$, then $RS = L - x$. The time to reach Q is the function as follows:

$$T(x) = \frac{\sqrt{x^2 + a^2}}{v_1} + \frac{\sqrt{(L-x)^2 + b^2}}{v_2}, \qquad \text{where } 0 \le x \le L.$$

We want to select x in $[0, L]$ to minimize $T(x)$.



Optimization Problems

Example (Snell's Law)

We first consider the critical points of $T(x)$:

$$T = \frac{\sqrt{x^2 + a^2}}{v_1} + \frac{\sqrt{(L-x)^2 + b^2}}{v_2} \qquad \text{where } 0 \leq x \leq L$$

$$\frac{dT}{dx} = \frac{x}{v_1\sqrt{x^2 + a^2}} + \frac{-(L-x)}{v_2\sqrt{(L-x)^2 + b^2}}$$

$$\frac{dT}{dx} = 0 \iff \frac{x}{v_1\sqrt{x^2 + a^2}} = \frac{L-x}{v_2\sqrt{(L-x)^2 + b^2}}$$

Solving the equation by squaring it will lead to a 4-th degree equation which is so complicated.

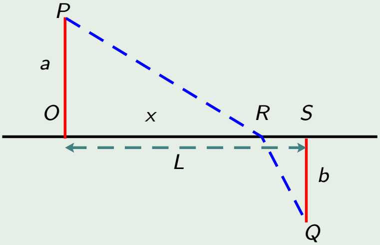
We can also look at its geometric meaning and construct a simple relationship.

Example (Snell's Law)

Let $\angle OPR = \alpha$, $\angle RQS = \beta$. Then we can simplify the condition of $T'(x) = 0$ as follows:

$$\frac{x}{v_1\sqrt{x^2+a^2}} = \frac{L-x}{v_2\sqrt{(L-x)^2+b^2}} \iff \frac{1}{v_1}\cdot\frac{OR}{PR} = \frac{1}{v_2}\cdot\frac{RS}{QR}$$

$$\iff \frac{1}{v_1}\cdot\sin\alpha = \frac{1}{v_2}\cdot\sin\beta \iff \frac{\sin\alpha}{\sin\beta} = \frac{v_1}{v_2}$$

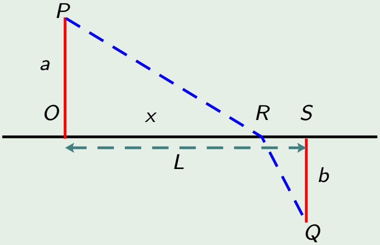


Optimization Problems

Example (Snell's Law)

Let $\angle OPR = \alpha$, $\angle RQS = \beta$. Hence, the condition of $T'(x) = 0$ can be written as

$$\frac{\sin\alpha}{\sin\beta} = \frac{v_1}{v_2}.$$



Can we conclude the x leads to $T'(x) = 0$ is a minimizer?

Example (Snell's Law)

We should check the sign of $T''(x)$:

$$T''(x) = \frac{a^2}{v_1 \left(x^2 + a^2\right)^{\frac{3}{2}}} + \frac{b^2}{v_2 \left((L - x)^2 + a^2\right)^{\frac{3}{2}}} > 0.$$

Hence, the function $T'(x)$ is increasing on $(-\infty, +\infty)$.

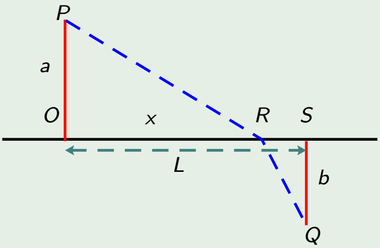
Let x_0 be the critical point of T such that $T'(x_0) = 0$. Then, we have $T'(x) < T(x_0) = 0$ for all $x < x_0$ and $T'(x) > T(x_0) = 0$ for all $x > x_0$, which implies $f(x_0)$ is the unique global minimum.

Optimization Problems

Example (Snell's Law)

Let $\angle OPR = \alpha$, $\angle RQS = \beta$. The point R should be the one satisfies

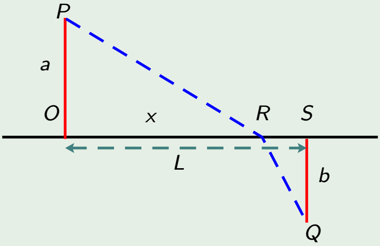
$$\frac{\sin\alpha}{\sin\beta} = \frac{v_1}{v_2}.$$



Example (Snell's Law)

Let $\angle OPR = \alpha$, $\angle RQS = \beta$. In optics, Snell's law says if a light passing through boundary OS between two different isotropic media (such as water, glass, or air) from P with directions PR, its trace satisfies

$$\frac{\sin\alpha}{\sin\beta} = \frac{v_1}{v_2}.$$

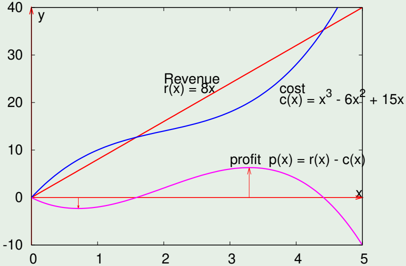


Optimization Problems

Example (maximizing the profit)

Suppose $c(x) = x^3 - 6x^2 + 15x$ is the cost of producing x thousands of units of a product, and $r(x) = 8x$ the revenue of selling x thousands of units of the product.

Then $p(x) = r(x) - c(x)$ is the profit of selling x thousands of units of the product. Find the production level which maximizes the profit.



Optimization Problems

Example (maximizing the profit)

We find the critical points

$$p'(x) = r'(x) - c'(x) = 8 - (3x^2 - 12x + 15) = 0 \Longrightarrow x = 2 \pm \frac{\sqrt{15}}{3}$$

It is easy to check that $p'(x) < 0$ for $x < 2 - \frac{\sqrt{15}}{3}$, or $x > 2 + \frac{\sqrt{15}}{3}$, and $p'(x) > 0$ when $2 - \frac{\sqrt{15}}{3} < x < 2 + \frac{\sqrt{15}}{3}$.

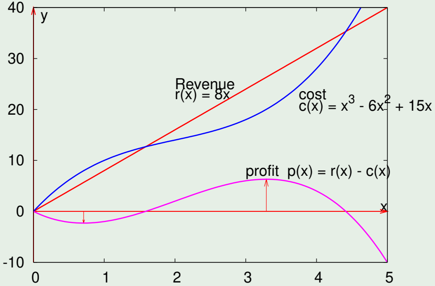
Hence, the maximum profit is reached at $x = 2 + \frac{\sqrt{15}}{3}$.

Optimization Problems

Example (maximizing the profit)

At $x = 2 - \frac{\sqrt{15}}{3}$, the profit is a local minimum (local maximum loss).

Hence, you will be fired if you select $x = 2 - \frac{\sqrt{15}}{3}$.



'

Closed Form Expression

For all above examples, we can provide the closed from expression for the solution of the optimization problems.

A closed-form expression is an expression expressed using a finite number of standard operations. It may contain

- constants, variables
- certain "well-known" operations (e.g., $+ - \times \div$)
- elementary functions (e.g., n-th root, exponent, logarithm, trigonometric functions, and inverse trigonometric functions)

It usually does not contain limit, differentiation, or integration (will be introduced in last 2 or 3 weeks).

The set of operations and functions admitted in a closed-form expression may vary with author and context.

Closed Form Expression

Unfortunately, the most of optimization problems in real-world application have no closed from expression.

Consider the simple polynomial function

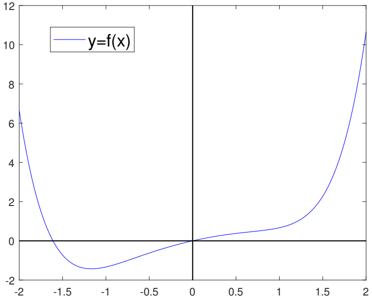
$$y = f(x) = \frac{1}{6}x^6 - \frac{1}{2}x^2 + x \text{ whose derivative is } f'(x) = x^5 - x + 1.$$

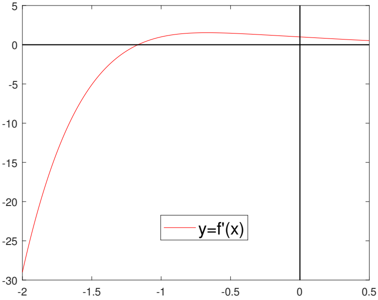
We want to find its local minimum/maximum.

Closed Form Expression

Consider that $f'(-1) = 1 > 0$ and $f'(-2) = -29 < 0$, then intermediate value theorem tell us there exists some c in $(-2, -1)$ such that $f'(c) = 0$.

We also have $f''(x) = 5x^4 - 1$. Hence $f''(c) > 0$ since $-2 < c < -1$, which means $f(c)$ is a local minimum.





Closed Form Expression

The Abel-Ruffini theorem (first asserted in 1799 and completely proved in 1824) shows that there is no closed form expression for the solution of

$$x^5 - x + 1 = 0.$$

We can use iterative methods to generate a sequence of improving approximate solutions.

For optimization problems, one class of popular functions we are interested in is convex/concave function.

Outline

Convex/Concave Functions

Let f is a real valued function defined on interval I. We call f is convex if for any x, y in I and t in $[0, 1]$, it holds that

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y).$$

We call f is strictly convex if for any x, y in I ($x \neq y$) and t in $(0, 1)$, we have

$$f(tx + (1-t)y) < tf(x) + (1-t)f(y).$$

We call f is strongly convex if for any x, y in I ($x \neq y$) and t in $(0, 1)$, we have

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y) - \frac{1}{2}ct(1-t)(x-y)^2$$

for some constant $c > 0$ (c is independent to x, y, t and I).

The definitions of concave, strictly concave, strongly concave are similar.

1st/2nd Order Condition

Theorem (1st/2nd order condition)

Suppose function f is twice differentiable over an open interval I. Then, the following statements are equivalent:

(a) *f is convex.*

(b) $f(y) \geq f(x) + f'(x)(y - x)$, *for all x and y in I.*

(c) $f''(x) \geq 0$, *for all x in I.*

We say (b) is first-order condition and (b) is second order condition.

We prove this theorem by showing:

- $(a) \implies (b)$
- $(b) \implies (a)$
- $(b) \implies (c)$
- $(c) \implies (b)$

f is convex $\Longrightarrow f(y) \geq f(x) + f'(x)(y-x)$

For any $t \in [0,1]$ and x, y in I, the convexity of f means

$$f(ty + (1-t)x) \leq tf(y) + (1-t)f(x)$$

Since $ty + (1-t)x = x + t(y-x)$, we have

$$\begin{aligned} & f(x + t(y-x)) \leq tf(y) + (1-t)f(x) \\ \Longrightarrow & f(x + t(y-x)) - f(x) \leq t(f(y) - f(x)). \end{aligned}$$

If $x = y$, then

$$f(y) \geq f(x) + f'(x)(y-x) \Longleftrightarrow f(x) \geq f(x).$$

If $t = 0$, then

$$f(ty + (1-t)x) \leq tf(y) + (1-t)f(x) \Longrightarrow f(x) \leq f(x).$$

If $t = 1$, then

$$f(ty + (1-t)x) \leq tf(y) + (1-t)f(x) \Longrightarrow f(y) \leq f(y).$$

f is convex $\implies f(y) \geq f(x) + f'(x)(y-x)$

If $x \neq y$ and $0 < t < 1$, then

$$f(x + t(y-x)) - f(x) \leq t(f(y) - f(x))$$
$$\implies f(y) - f(x) \geq \frac{f(x+t(y-x)) - f(x)}{t(y-x)} \cdot (y-x)$$

We let $h = t(y-x)$ and take $t \to 0^+$. If $y > x$, then $h \to 0^+$ and

$$\lim_{t \to 0^+} \frac{f(x+t(y-x)) - f(x)}{t(y-x)} = \lim_{h \to 0^+} \frac{f(x+h) - f(x)}{h} = f'(x).$$

If $y < x$, then $h \to 0^-$ and

$$\lim_{t \to 0^+} \frac{f(x+t(y-x)) - f(x)}{t(y-x)} = \lim_{h \to 0^-} \frac{f(x+h) - f(x)}{h} = f'(x).$$

Hence, we have

$$f(y) - f(x) \geq f'(x)(y-x).$$

f is convex $\Longrightarrow f(y) \geq f(x) + f'(x)(y-x)$

Combing all above analysis, we have proved $(a) \Longrightarrow (b)$ in the theorem.

Theorem (1st/2nd order condition)

Suppose function f is twice differentiable over an open interval I. Then, the following statements are equivalent:

- (a) *f is convex.*
- (b) $f(y) \geq f(x) + f'(x)(y-x)$, *for all x and y in I.*
- (c) $f''(x) \geq 0$, *for all x in I.*

Then we want to show that $(b) \Longrightarrow (a)$.

$f(y) \geq f(x) + f'(x)(y-x) \implies f$ is convex

The statement (b) means for any x, y and z in I, we have

$$
\begin{aligned}
f(x) \geq & f(z) + f'(z)(x-z) \\
f(y) \geq & f(z) + f'(z)(y-z)
\end{aligned}
$$

Let $z = tx + (1-t)y$, then z is in I for any $0 \leq t \leq 1$. We have

$$
\begin{aligned}
& tf(x) + (1-t)f(y) \\
\geq & tf(z) + tf'(z)(x-z) + (1-t)f(z) + (1-t)f'(z)(y-z).
\end{aligned}
$$

On the other hand

$$
\begin{aligned}
t(x-z) + (1-t)(y-z) = & tx - tz + (1-t)y - (1-t)z \\
= & tx + (1-t)y - z = 0.
\end{aligned}
$$

Hence, $tf(x) + (1-t)f(y) \geq f(z) = f(tx + (1-t)y)$, which means (a).

We have proved $(a) \iff (b)$ in the theorem.

Theorem (1st/2nd order condition)

Suppose function f is twice differentiable over an open interval I. Then, the following statements are equivalent:

- (a) *f is convex.*
- (b) $f(y) \geq f(x) + f'(x)(y-x)$, *for all x and y in I.*
- (c) $f''(x) \geq 0$, *for all x in I.*

Then we want to show that $(b) \implies (c)$ and $(c) \implies (b)$.

$f(y) \geq f(x) + f'(x)(y-x) \implies f''(x) \geq 0$

The first order condition (b) means for any x and y in I, we have

$$f(y) \geq f(x) + f'(x)(y-x) \implies f'(x)(y-x) \leq f(y) - f(x).$$
$$f(x) \geq f(y) + f'(y)(x-y) \implies f(y) - f(x) \leq f'(y)(y-x).$$

Then $f'(x)(y-x) \leq f(y) - f(x) \leq f'(y)(y-x)$.

For any $x \neq y$, we divide $(y-x)^2$ on both sides and obtain

$$\frac{f'(y) - f'(x)}{y-x} \geq 0.$$

Let $y = x + h$ and take $h \to 0$, then we have

$$f''(x) = \lim_{h \to 0} \frac{f'(x+h) - f'(x)}{h} = \lim_{h \to 0} \frac{f'(y) - f'(x)}{(y-x)^2} \geq 0.$$

$f''(x) \geq 0 \implies f(y) \geq f(x) + f'(x)(y - x)$

Exercise

If f is twice differentiable on (a, b) and continuous on $[a, b]$, then

$$f(b) - f(a) - f'(a)(b - a) = \frac{f''(c)}{2}(b - a)^2$$

for some $c \in (a, b)$.

For any x and y in I, we have

$$f(y) - f(x) - f'(x)(y - x) = \frac{f''(z)}{2}(y - x)^2$$

for some z in $[x, y]$ (which means z is in I and $f''(z) \geq 0$). Then

$$f(y) = f(x) + f'(x)(y - x) + \frac{f''(z)}{2}(y - x)^2 \geq f(x) + f'(x)(y - x).$$

1st/2nd Conditions

Now we have presented all of proof of the theorem.

Theorem (1st/2nd order condition)

Suppose function f is twice differentiable over an open interval I. Then, the following statements are equivalent:

(a) f *is convex.*

(b) $f(y) \geq f(x) + f'(x)(y - x)$, *for all* x *and* y *in* I.

(c) $f''(x) \geq 0$, *for all* x *in* I.

We can also give 1st/2nd conditions for strictly/strongly convex function.

1st/2nd Conditions

Now we have presented all of proof of the theorem.

Theorem (1st/2nd order condition)

Suppose function f is twice differentiable over an open interval I. Then, the following statements are equivalent:

(a) f *is convex.*

(b) $f(y) \geq f(x) + f'(x)(y - x)$, *for all* x *and* y *in* I.

(c) $f''(x) \geq 0$, *for all* x *in* I.

We can also give 1st/2nd conditions for strictly/strongly convex function.

1st/2nd Conditions

Now we have presented all of proof of the theorem.

Theorem (1st/2nd order condition)

Suppose function f is twice differentiable over an open interval I. Then, the following statements are equivalent:

(a) f *is convex.*

(b) $f(y) \geq f(x) + f'(x)(y - x)$, *for all* x *and* y *in* I.

(c) $f''(x) \geq 0$, *for all* x *in* I.

We can also give 1st/2nd conditions for strictly/strongly convex function.

Outline

Global/Local Minimum of Convex Function

Any local minimum of convex function is also a global minimum.

Theorem (also holds for non-differentiable function)

Suppose function f is convex on interval I. If x^* is is a local minimum over I, then x^* is also a global minimum of f over a convex set I.

Proof.

Since $f(x^*)$ is a local minimum, for any y in I, we can choose a sufficient small $t < 1$, such that $ty + (1-t)x^*$ in I and $f(x^*) \leq f(x^* + t(y - x^*))$. The convexity of f implies

$$f(x^*) \leq f(x^* + t(y - x^*)) = f(ty + (1-t)x^*) \leq tf(y) + (1-t)f(x^*)$$

$$\Longrightarrow f(x^*) \leq tf(y) + (1-t)f(x^*) \Longrightarrow f(x^*) \leq f(y)$$

□

First-Order Optimal Condition

Theorem

If function f is convex and differentiable over an interval I. Then any point x^* that satisfies $f'(x^*) = 0$ holds that $f(x^*)$ is a global minimum.

Proof.

The 1-st order condition of convex and differentiable function means

$$f(y) \geq f(x^*) + f'(x^*)(y - x^*) = f(x^*)$$

for all y in I. □

Consider that the convex and differentiable function $f(x) = e^x$ with domain $[1, \infty)$. The minimum is $f(1) = e$ but $f'(1) = e \neq 0$.

First-Order Optimal Condition

We desire to establish an equivalent condition for global minimum of convex and differentiable function.

A good strategy is relaxing the condition of $f'(x^*) = 0$ to

$$f'(x^*)(y - x^*) \geq 0$$

holds for all y in I.

First-Order Optimal Condition

Theorem (sufficient condition)

If function f is convex and differentiable over an interval I. Then any point x^* that satisfies

$$f'(x^*)(y - x^*) \geq 0$$

for all y in I holds that $f(x^*)$ is a global minimum.

Theorem (necessary condition)

If function f is convex and differentiable over an interval I. Then for any point x^* such that $f(x^*)$ is a global minimum, we have

$$f'(x^*)(y - x^*) \geq 0$$

for all y in I.

Proof (necessary condition).

Let x^* in I and $f(x^*)$ is a global minimum. Suppose y in I such that

$$f'(x^*)(y - x^*) < 0.$$

There must hold that $y \neq x^*$. Let $t > 0$, $h = t(y - x^*)$. Taking $t \to 0^+$, then

$$\begin{aligned}
&\lim_{t\to 0^+} \frac{f(x^* + t(y - x^*)) - f(x^*)}{t} \\
=&(y - x^*) \cdot \lim_{t\to 0^+} \frac{f(x^* + t(y - x^*)) - f(x^*)}{t(y - x^*)} \\
=&(y - x^*) \cdot \lim_{h\to 0} \frac{f(x^* + h) - f(x^*)}{h} = f'(x^*)(y - x^*) < 0.
\end{aligned}$$

For sufficient small t, we have $f(x^* + t(y - x^*) - f(x^*) < 0$ for $x^* + t(y - x^*)$ in I, which contradicts to x^* is global minimum. Hence, we must have

$$f'(x^*)(y - x^*) \geq 0$$

and the convexity means $f(y) \geq f(x^*) + f'(x^*)(y - x^*) \geq f(x^*)$. □

First-Order Optimal Condition

Note that the equivalent condition

$$f'(x^*)(y - x^*) \geq 0,$$

only depends on the function value and the derivative of f.

Hence, it works for any convex and differentiable even if f' is not differentiable (f'' may not exist).