Calculus IB: Lecture 15

Luo Luo

Department of Mathematics, HKUST

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

Outline

Outline

Geometric View of Convex Function

Definition (convex function)

Let f is a real valued function defined on interval I. We call f is convex if for any x_1, x_2 in I and t in $[0, 1]$, it holds that

$$f(tx_1 + (1-t)x_2) \leq tf(x_1) + (1-t)f(x_2).$$

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) *$f(x) \geq f(x_0) + f'(x_0)(x - x_0)$, for all x and x_0 in I.*

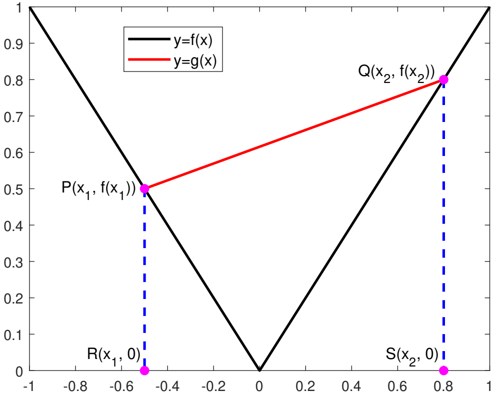
(c) *$f''(x) \geq 0$, for all x in I.*

Geometric View of Convex Function

Let $f(x)$ be a convex function and define the linear function $(x_1 \neq x_2)$

$$g(x) = \frac{f(x_2) - f(x_1)}{x_2 - x_1}(x - x_1) + f(x_1).$$

Then we have $g(x_1) = f(x_1)$, $g(x_2) = f(x_2)$ and $g(x) \geq f(x)$ for any $x_1 \leq x \leq x_2$.



Geometric View of Convex Function

We can prove $g(x) \geq f(x)$ for any $x_1 < x < x_2$ by the convexity of f.

Proof.

Since $x_1 < x < x_2$, there exists $0 < t < 1$ such that $x = tx_1 + (1-t)x_2$. Using the definition of convexity, we have

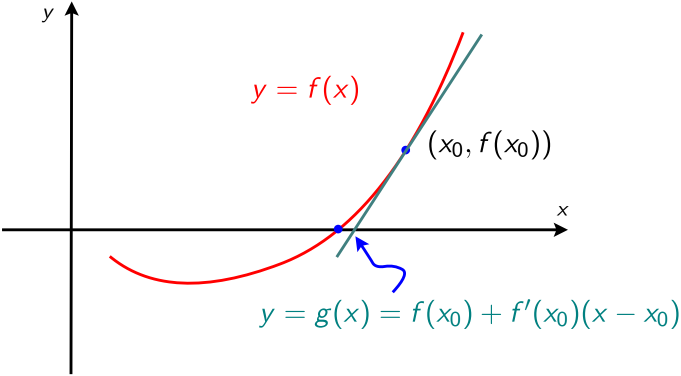
$$
\begin{aligned}
g(x) =& \frac{f(x_2) - f(x_1)}{x_2 - x_1}(x - x_1) + f(x_1) \\
=& \frac{f(x_2) - f(x_1)}{x_2 - x_1}(tx_1 + (1-t)x_2 - x_1) + f(x_1) \\
=& \frac{f(x_2) - f(x_1)}{x_2 - x_1}(1-t)(x_2 - x_1) + f(x_1) \\
=& (1-t)(f(x_2) - f(x_1)) + f(x_1) \\
=& (1-t)f(x_2) + tf(x_1) \\
\geq& f(tx_1 + (1-t)x_2) = f(x)
\end{aligned}
$$

□

Geometric View of Convex Function

Theorem (1st order condition)

Suppose f is differentiable over an open interval I. Then f is convex is equivalent to $f(x) \geq f(x_0) + f'(x_0)(x - x_0)$ holds for all x and x_0 in I.



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 a local minimum over I, then x^* is also a global minimum of f over 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

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

□

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 optimal condition

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

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

Hence, it works even if f' is not differentiable (f'' does not exist).

First-Order Optimal Condition

Exercise

Provide examples of $f(x)$ for the following cases respectively

- $f(x)$ is convex, but not differentiable
- $f(x)$ is convex and differentiable, but $f'(x)$ is non-differentiable
- $f(x)$ is convex and differentiable, and $f^{(n)}(x)$ is differentiable for all positive integer n.

Outline

Linear Approximation

Recall that the slope of the tangent line to the graph of $y = f(x)$ at $x = a$ is the derivative

$$f'(a) = \lim_{h \to 0} \frac{f(a+h) - f(a)}{h}$$

Therefore the equation of the tangent line to the graph of $y = f(x)$ at the point $(a, f(a))$ is determined by the slope condition

$$y = f(a) + f'(a)(x - a).$$

Letting $x = a + h$, we have

$$f'(a) \approx \frac{f(a+h) - f(a)}{h} = \frac{f(x) - f(a)}{x - a} \qquad \text{when } h = x - a \approx 0$$

i.e., $f(x) \approx f(a) + f'(a)(x - a)$, when $x \approx a$. In other words,

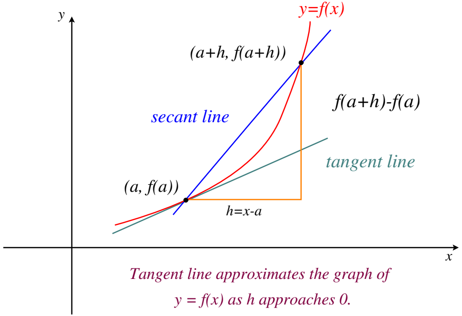
$$\text{tangent line} \xrightarrow{\text{approximates}} \text{graph of } y = f(x) \text{ near the point } (a, f(a))$$

Linear Approximation

The *tangent line approximation at $x = a$*, or *linear approximation at $x = a$*, or *linearization at $x = a$*, of a function $y = f(x)$ (differentiable at $x = a$) is that we are using the tangent line equation (or the corresponding linear function) to approximate the given function.

$$y = f(x) \xleftarrow{\approx} \text{Tangent Line Equation}: \; y = f(a) + f'(a)(x - a)$$

$$\Longrightarrow f(x) \approx f(a) + f'(a)(x - a) \qquad \text{for } x - a \approx 0$$



Tangent line approximates the graph of $y = f(x)$ as h approaches 0.

Linear Approximation

Example ($f(x) = \sqrt{x+1}$)

Find the linear approximation of $f(x) = \sqrt{x+1}$ at $x = 0$.

We have

$$f'(x) = \frac{1}{2}(x+1)^{-1/2}, \text{ i.e. } f'(0) = \frac{1}{2}$$

and the equation of the tangent line at $(0, 1)$ is

$$y = 1 + \frac{1}{2}(x - 0) = 1 + \frac{1}{2}x$$

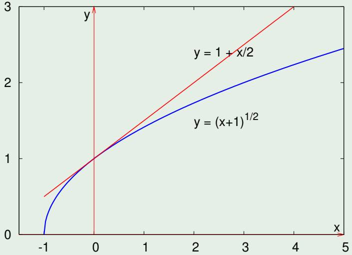
which is also called the linear approximation of $f(x) = \sqrt{x+1}$ at $x = 0$.

Linear Approximation

Example ($f(x) = \sqrt{x+1}$)

Let absolute error be $\sqrt{x+1} - 1 + \frac{x}{2}$.

x	$y = \sqrt{x+1}$	$y = 1 + \frac{x}{2}$	absolute error
0.200	1.095445	1.1000	$< 10^{-2}$
0.050	1.024695	1.0250	$< 10^{-3}$
0.005	1.002497	1.0025	$< 10^{-5}$



Linear Approximation

Example ($\sqrt[3]{8.5}$)

Find an approximate value of $\sqrt[3]{8.5}$ by the linear approximation of a suitable function.

Let $f(x) = \sqrt[3]{x} = x^{1/3}$, with $f'(x) = \frac{1}{3}x^{-2/3}$.

The linear approximation at $x = 8$ is:

$$f(x) \approx f(8) + f'(8)(x-8) = 2 + \frac{1}{12}(x-8) = \frac{x}{12} + \frac{4}{3}$$

Thus

$$f(8.5) \approx \frac{8.5}{12} + \frac{4}{3} = \frac{245}{120} = 2.04167.$$

Note that $\sqrt[3]{8.5} = 2.04093$ from a calculator.

Differential of the Function

The tangent line approximation at x is

$$f(x + \Delta x) \approx f(x) + f'(x)\Delta x,$$

where Δx denotes some increment in x (which could be negative).

Then we use Δy or Δf to denote the change in the function values

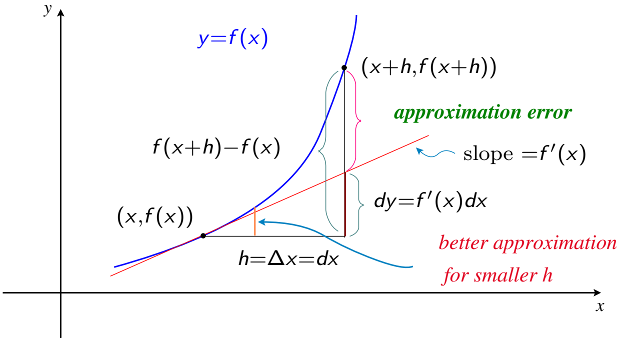
$$\Delta y = \Delta f = f(x + \Delta x) - f(x).$$

and the linear approximation be expressed as

$$\Delta f \approx f'(x)\Delta x.$$

Note that $f'(x)\Delta x$ is the change of y-value along the tangent line!

Differential of the Function



The notation of differentials $df = f'(x)dx$ is obtained by expressing Δx as dx, and $dy = df = f'(x)dx$ can be used as an approximation of

$$\Delta y = f(x + \Delta x) - f(x).$$

Example (area of a circle)

Approximate the increase in the area of a circle when the radius is increased from 10m to 10.1m.

The area of the circle is $A(r) = \pi r^2$, then $dA = A'(r)dr = 2\pi r dr$.

For $r = 10$m and $dr = 0.1$m, we have $dA = 2\pi(10)0.1 = 2\pi\text{m}^2$, which approximates the change in area ΔA.

The approximate area at $r = 10.1$ m is:

$$A \approx A(10) + dA = 100\pi + 2\pi = 102\pi\text{m}^2.$$

The exact area at $r = 10.1$m is $A = \pi(10.1)^2 = \pi(10)^2 + \Delta A$. The absolute error of the estimate is

$$|\pi(10.1)^2 - 102\pi| = 102.01\pi - 102\pi = 0.01 = |\Delta A - dy|\text{m}^2$$

Linear Approximation of Convex Function

Given a differentiable function $y = f(x)$ defined on an open interval I, its linear approximation is

$$f(x) \approx f(a) + f'(a)(x - a).$$

We additionally suppose f is convex, then the first-order condition means

$$f(x) \geq f(a) + f'(a)(x - a).$$

Hence, the linear approximation provides a lower bound of convex function.