香港中文大學(深圳)
The Chinese University of Hong Kong, Shenzhen

Introduction to Data Science

Lecture 16 Optimization: Convex Function

Zicheng Wang

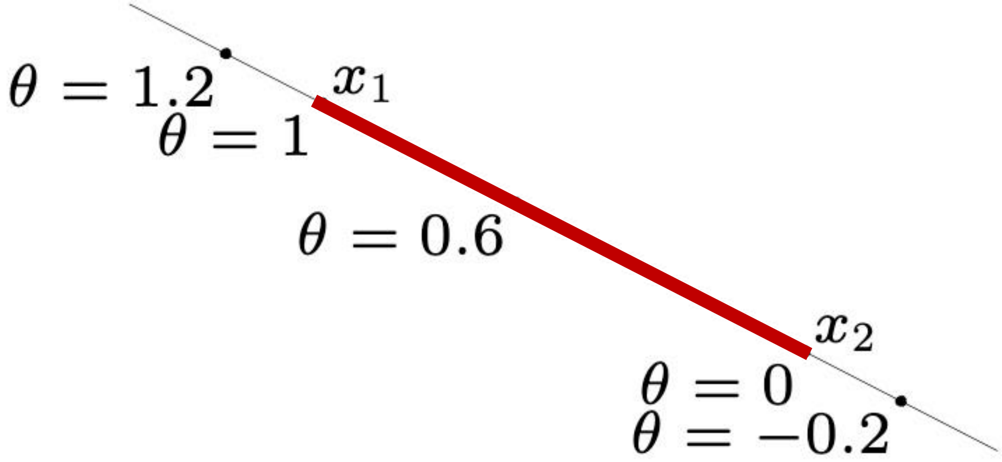
Review: Convex Set

Line Segment

- Let $\boldsymbol{x_1} \neq \boldsymbol{x_2}$ be two points in $\mathbb{R}^n$. Points of the form

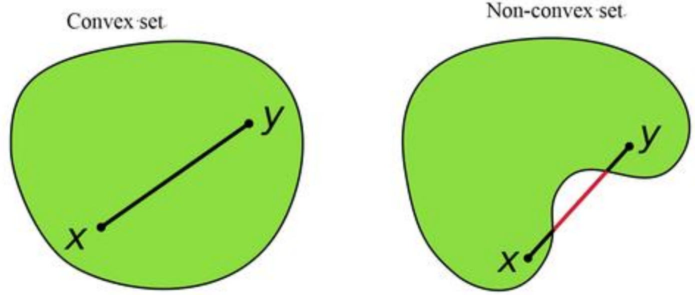
$$\boldsymbol{x} = \theta \boldsymbol{x}_1 + (1-\theta)\boldsymbol{x}_2$$

where $\boldsymbol{\theta} \in [\mathbf{0}, \mathbf{1}]$, form the **line segment** between $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$.



Convex Set

- Set C is a **convex set** if the line segment between any two points in C lies in C.



- Formal definition: A set C is convex if $\forall\, \boldsymbol{x_1}, \boldsymbol{x_2} \in C, \forall \theta \in [0,1]$
$$\theta \boldsymbol{x_1} + (1-\theta)\boldsymbol{x_2} \in C.$$

Remark: In this lecture, I will use **bold** form to represent a high dimension point. Without bold form, it represents a scalar

Convex Set Examples

- The empty set $\emptyset$, the singleton set $\{\boldsymbol{x_0}\}$, and the complete space R are convex sets.

- An interval of $[a, b] \subset R$ is a convex set

- In R^n the set $\mathrm{H} \coloneqq \{\boldsymbol{x} \in R^n : a_1 x_1 + \cdots + a_n x_n = c\}$ is a convex set

- Half spaces, e.g., $\mathrm{H} \coloneqq \{(x, y) : y \leq ax + b\}$ are convex sets

- A disk with center $(0,0)$ and radius c is a convex subset of R^2

Remark: In this lecture, I will use **bold** form to represent a high dimension point. Without bold form, it represents a scalar

Steps for Showing the Convexity of a Set

Prove $\mathrm{H} := \{(x, y) : y = ax + b\}$ is a convex set

For any (x_1, y_1) and (x_2, y_2) in H,

- $y_1 = ax_1 + b$
- $y_2 = ax_2 + b$

1. Use the assumption that $(x_1, y_1), (x_2, y_2) \in H$

- $\theta(x_1, y_1) + (1-\theta)(x_2, y_2) = (\theta x_1 + (1-\theta)x_2), \theta y_1 + (1-\theta)y_2)$
- Then for any $\theta \in [0,1]$

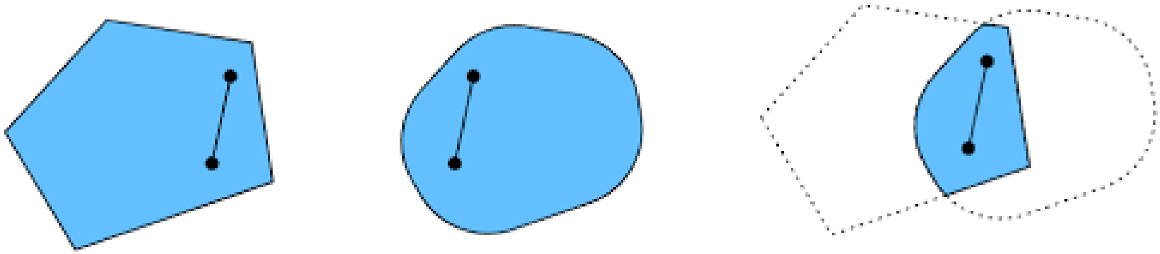
2. Characterize the new point within the line segment

  - $\theta y_1 + (1-\theta)y_2 = a(\theta x_1 + (1-\theta)x_2) + b$

3. Use (1) and (2) to show that the new point is in H

Properties of convex sets.

Lemma: If both S_1 and S_2 are convex sets, then $S_1 \cap S_2$ is also a convex set.



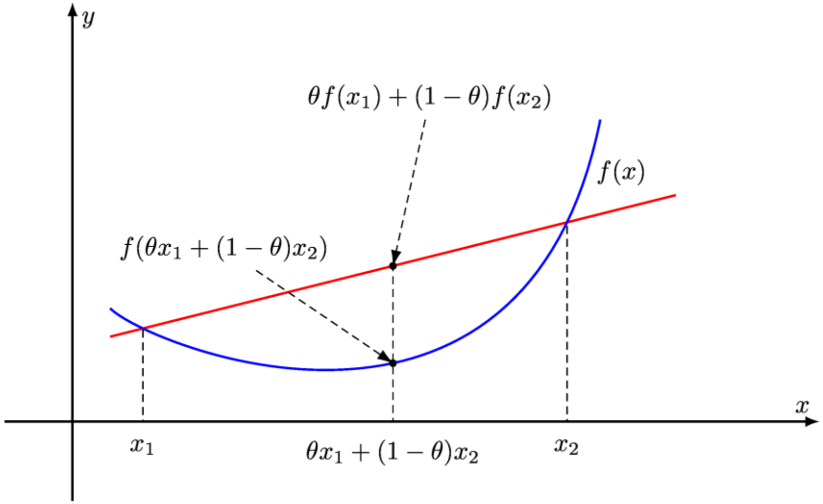
Convex Function

Convex Function

Definition: A function $f(\boldsymbol{x}): R^n \rightarrow R$ is **convex** if (1) its domain is a convex set, and (2) for any $\boldsymbol{x_1}, \boldsymbol{x_2} \in dom(f)$ and any $0 \leq \lambda \leq 1$, we have

$$f(\boldsymbol{z}) \leq \lambda f(\boldsymbol{x_1}) + (1-\lambda) f(\boldsymbol{x_2})$$

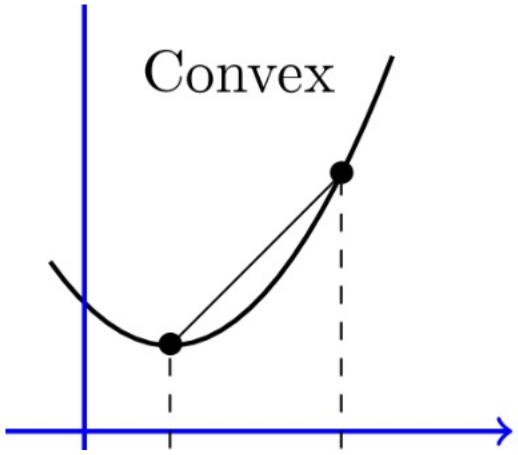
where $\boldsymbol{z} = \lambda \boldsymbol{x_1} + (1-\lambda)\boldsymbol{x_2}$.

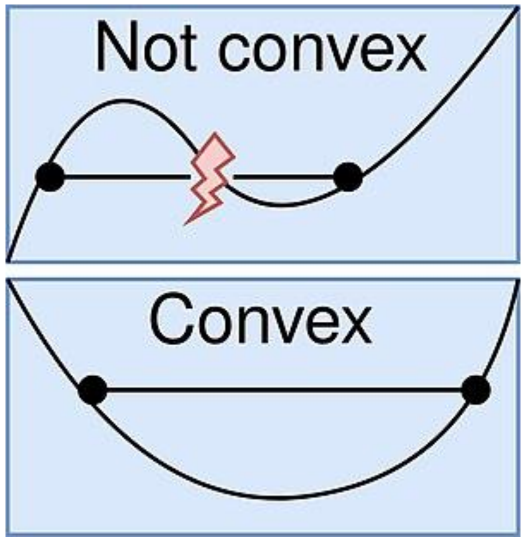


Function f evaluated at the combination of two points $\boldsymbol{x_1}, \boldsymbol{x_2}$ is no larger than the **same combination of $f(\boldsymbol{x_1})$** and $\boldsymbol{f(x_2)}$

Other Definition

- Pick any two points on the function curve.
- The line segment between these two points is above the function curve.



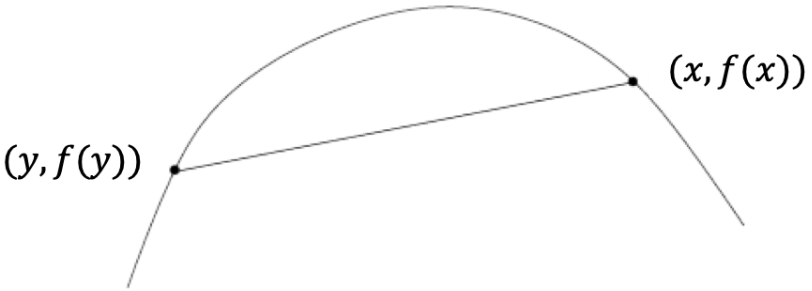


Concave Function

Definition: A function $f(\boldsymbol{x}): R^n \rightarrow R$ is **concave** if (1) the domain of f is a convex set, and (2) for any $\boldsymbol{x}, \boldsymbol{y} \in dom(f)$ and any $0 \leq \lambda \leq 1$, we have

$$f(\boldsymbol{z}) \geq \lambda f(\boldsymbol{x}) + (1-\lambda) f(\boldsymbol{y})$$

where $\boldsymbol{z} = \lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}$.

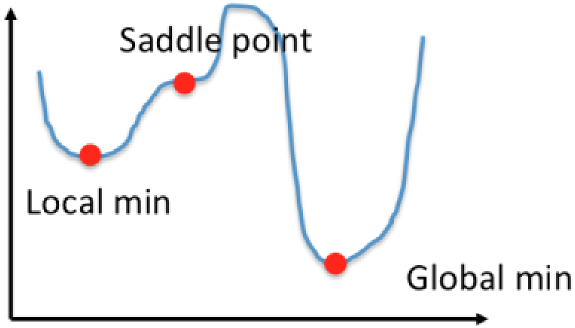


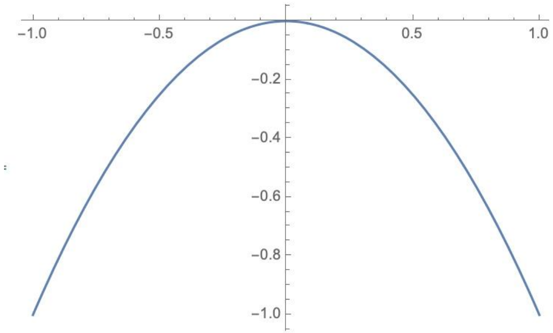
the region below a concave function is a convex set.

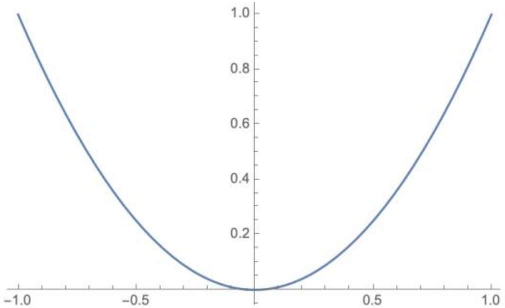
If f is concave, then -f is convex!

If f is convex, then -f is concave!

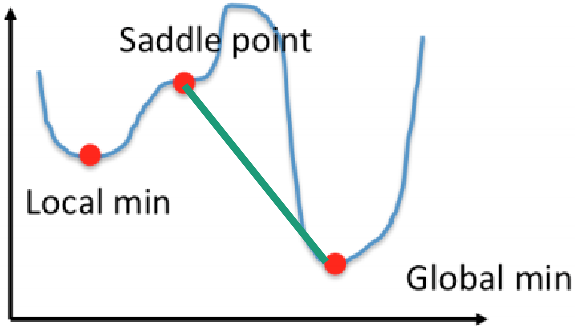
Convex function?

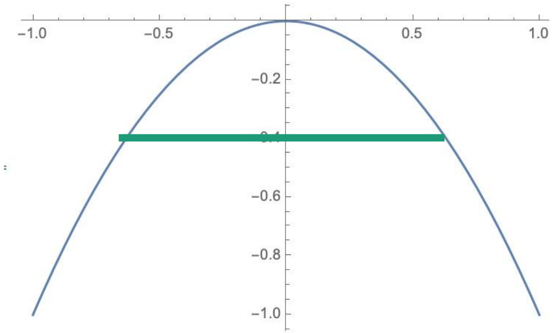


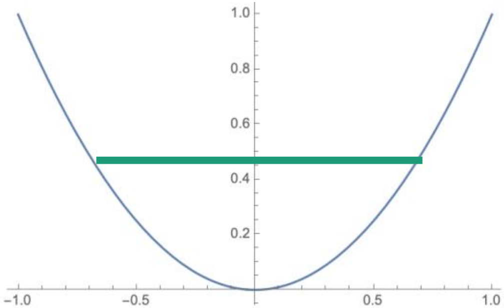




Convex function?







Second Order Condition (SOC)

Suppose f is a twice continuously differentiable function. Then f is convex **if and only if**

(1) dom(f) is a convex set

(2) for any $\boldsymbol{x} \in$ dom(f), any unit vector $\boldsymbol{e}$ and any θ,

$$\frac{d^2 f(\boldsymbol{x} + \theta \boldsymbol{e})}{d\theta^2} \geq 0$$

One dimension: $\boldsymbol{f}''(\boldsymbol{x}) \geq \mathbf{0}$ **for all** $\boldsymbol{x} \in \boldsymbol{dom}(\boldsymbol{f})$.

Second Order Condition (SOC)

Suppose f is a twice continuously differentiable function. Then f is concave

if and only if

(1) dom(f) is a convex set

(2) for any $\boldsymbol{x} \in$ dom(f), any unit vector $\boldsymbol{e}$ and any θ,

$$\frac{d^2 f(\boldsymbol{x} + \theta \boldsymbol{e})}{d\theta^2} \leq 0$$

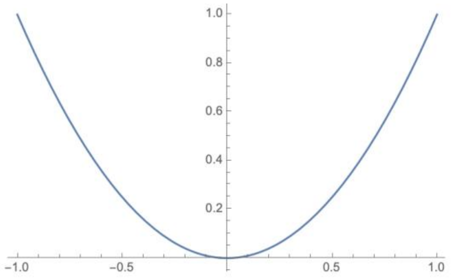
One dimension: $\boldsymbol{f''(x) \leq 0}$ **for all** $\boldsymbol{x \in dom(f)}$.

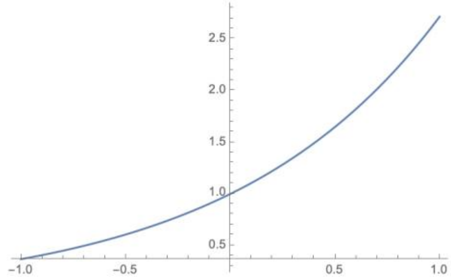
Examples of Convex/Concave Functions

Convex

- $f(x) = ax + b\ (also\ concave)$
- $f(x) = x^2$
- $f(x) = e^x$

Verify using the second order condition

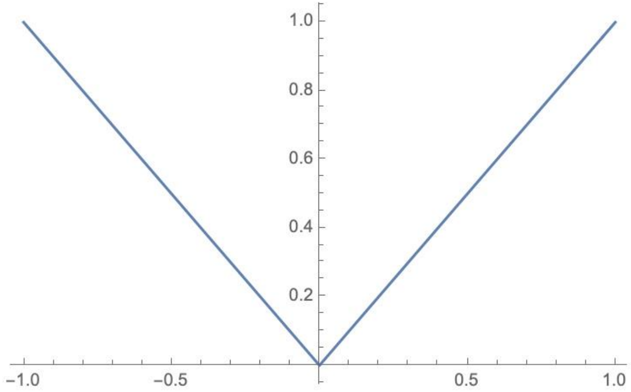




Examples of Convex/Concave Functions

Convex

- $f(x) = ax + b \; (also\ concave)$
- $f(x) = x^2$
- $f(x) = e^x$
- $f(x) = |x|$



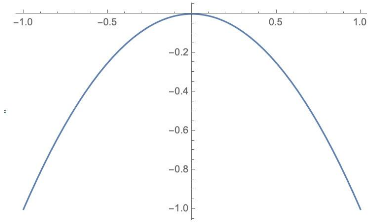
Examples of Convex/Concave Functions

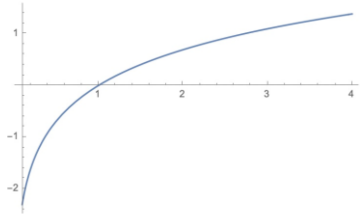
Convex

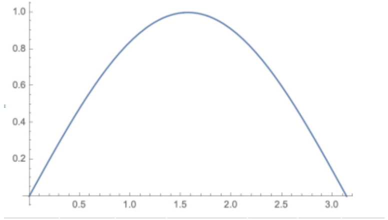
- $f(x) = ax + b\ (also\ concave)$
- $f(x) = x^2$
- $f(x) = e^x$
- $f(x) = |x|$

Concave

- $f(x) = -x^2$
- $f(x) = \log(x)$ on $(0, +\infty)$
- $f(x) = \sin(x)$ on $[0, \pi]$







High dimensional Convex functions

General procedure

$$\frac{d^2 f(\boldsymbol{x} + \theta \boldsymbol{e})}{d\theta^2} \geq 0$$

1. Choose any $\mathbf{y}$, any unit vector $\boldsymbol{e}$ and any θ
2. Define g(θ) = $f(\boldsymbol{y} + \theta \boldsymbol{e})$
3. Prove g'' (θ) ≥ 0

The Blind Box Problem

$$\text{minimize} \quad \overbrace{-(pE[\min(Q,D)] + sE[(Q-D)^+] - cQ)}^{\text{convex function of Q?}}$$

$$\text{subject to} \quad \underbrace{Q \geq 0}_{\text{Convex set}}$$

In order to show the convexity of the objective function, it suffices to show that $pE[\min(Q,D)] + sE[(Q-D)^+] - cQ$ is concave in Q.

$\pi(Q) \stackrel{\text{def}}{=} pE[\min(Q, D)] + sE[(Q - D)^+] - cQ$

Let the pdf of D be h(x) on [0, ∞)

$$\pi(Q) = \mathrm{p}\int_0^Q x\, h(x)dx + \mathrm{p}\int_Q^\infty Q\, h(x)dx + s\int_0^Q (Q - x)\, h(x)dx - cQ.$$

$\pi(Q) \stackrel{\text{def}}{=} pE[\min(Q, D)] + sE[(Q - D)^+] - cQ$

Let the pdf of D be h(x) on [0, ∞)

$$\pi(Q) = \mathrm{p}\int_0^Q x\, h(x)dx + \mathrm{p}\int_Q^\infty Q\, h(x)dx + s\int_0^Q (Q - x)\, h(x)dx - cQ.$$

$$\pi'(Q) = \mathrm{p}\, Q\, h(Q) - \mathrm{p}\, Q\, h(Q) + \mathrm{p}\int_Q^\infty h(x)dx + (Q - Q)h(Q) + s\int_0^Q h(x)dx - c$$

$$= \mathrm{p}\int_Q^\infty h(x)dx + s\int_0^Q h(x)dx - c.$$

$\pi(Q) \stackrel{\text{def}}{=} pE[\min(Q, D)] + sE[(Q-D)^+] - cQ$

Let the pdf of D be h(x) on [0, ∞)

$$\pi(Q) = \mathrm{p}\int_0^Q x\, h(x)dx + \mathrm{p}\int_Q^\infty Q\, h(x)dx + s\int_0^Q (Q-x)\, h(x)dx - cQ.$$

$$\pi'(Q) = \mathrm{p}\, Q\, h(Q) - \mathrm{p}\, Q\, h(Q) + \mathrm{p}\int_Q^\infty h(x)dx + (Q-Q)h(Q) + s\int_0^Q h(x)dx - c$$

$$=\mathrm{p}\int_Q^\infty h(x)dx + s\int_0^Q h(x)dx - c\,.$$

$$\pi''(Q) = -\mathrm{p}\, h(Q) + \mathrm{s}\;\, h(Q) = (s-p)h(Q).$$

As the salvage value is no larger than the price, $\pi''(Q) \le 0$

Convex Function vs. Convex Set

Convex function *vs.* Convex set (Part I)

- $C = \{\boldsymbol{x}: f(\boldsymbol{x}) \leq r\}$ is a convex set if $f(\boldsymbol{x})$ is a convex function

1. Choose any two points satisfying the conditions

$$\boldsymbol{x_1}: f(\boldsymbol{x_1}) \leq r \text{ and } \boldsymbol{x_2}: f(\boldsymbol{x_2}) \leq r$$

2. Introduce a new point

$$\lambda \boldsymbol{x_1} + (1-\lambda)\, \boldsymbol{x_2}$$

3. Use (1) to show that the point introduced in (2) satisfies the conditions as well

$$f(\lambda\, \boldsymbol{x_1} + (1-\lambda)\, \boldsymbol{x_2}) \leq \lambda f(\boldsymbol{x_1}) + (1-\lambda) f(\boldsymbol{x_2}) \leq r$$

Convex function *vs.* Convex set (Part I)

- $C = \{\boldsymbol{x}: f(\boldsymbol{x}) \leq r\}$ is a convex set if $f(\boldsymbol{x})$ is a convex function

1. Choose any two points satisfying the conditions

$$\boldsymbol{x_1}: f(\boldsymbol{x_1}) \leq r \;\text{ and }\; \boldsymbol{x_2}: f(\boldsymbol{x_2}) \leq r$$

2. Introduce a new point

$$\lambda \boldsymbol{x_1} + (1-\lambda)\, \boldsymbol{x_2}$$

3. Use (1) to show that the point introduced in (2) satisfies the conditions as well

$$\underbrace{f(\lambda\, \boldsymbol{x_1} + (1-\lambda)\, \boldsymbol{x_2}) \leq \lambda f(\boldsymbol{x_1}) + (1-\lambda) f(\boldsymbol{x_2})}_{f \text{ is a convex function}} \leq r$$

Convex function *vs.* Convex set (Part I)

- $C = \{\boldsymbol{x}: f(\boldsymbol{x}) \le r\}$ is a convex set if $f(\boldsymbol{x})$ is a convex function

1. Choose any two points satisfying the conditions

$$\boldsymbol{x_1}: f(\boldsymbol{x_1}) \le r \;\text{ and }\; \boldsymbol{x_2}: f(\boldsymbol{x_2}) \le r$$

2. Introduce a new point

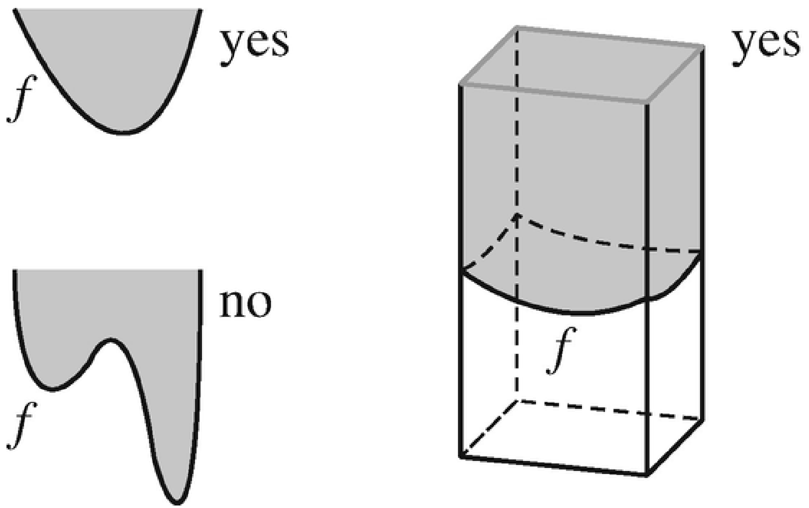
$$\lambda \boldsymbol{x_1} + (1-\lambda)\,\boldsymbol{x_2}$$

3. Use (1) to show that the point introduced in (2) satisfies the conditions as well

$$f(\lambda\,\boldsymbol{x_1} + (1-\lambda)\,\boldsymbol{x_2}) \le \underbrace{\lambda f(\boldsymbol{x_1}) + (1-\lambda) f(\boldsymbol{x_2}) \le r}_{\text{Step (1)}}$$

Convex function *vs.* Convex set (Part II)

$C = \{(\boldsymbol{x}, y): y \geq f(\boldsymbol{x})\}$ is a convex set if $f(\boldsymbol{x})$ is a convex function.



Geometrically, this means that **the region above a convex function is a convex set.**

Convex function *vs.* Convex set (Part II)

$C = \{(\boldsymbol{x}, y): y \geq f(\boldsymbol{x})\}$ is a convex set if $f(\boldsymbol{x})$ is a convex function.

1. Choose any two points satisfying the conditions

$$(\boldsymbol{x_1}, y_1): y_1 \geq f(\boldsymbol{x_1}) \text{ and } (\boldsymbol{x_2}, y_2): y_2 \geq f(\boldsymbol{x_2})$$

2. Introduce a new point

$$\lambda(\boldsymbol{x_1}, y_1) + (1-\lambda)(\boldsymbol{x_2}, y_2) = (\lambda \boldsymbol{x_1} + (1-\lambda)\, \boldsymbol{x_2}, \lambda y_1 + (1-\lambda)\, y_2)$$

3. Use (1) to show that the point introduced in (2) satisfies the conditions as well

$$\lambda y_1 + (1-\lambda)\, y_2 \geq \lambda f(\boldsymbol{x_1}) + (1-\lambda) f(\boldsymbol{x_2}) \geq f(\lambda \boldsymbol{x_1} + (1-\lambda)\, \boldsymbol{x_2})$$

Convex function *vs.* Convex set (Part II)

$C = \{(\boldsymbol{x}, y) : y \geq f(\boldsymbol{x})\}$ is a convex set if $f(\boldsymbol{x})$ is a convex function.

1. Choose any two points satisfying the conditions

$$(\boldsymbol{x_1}, y_1) : y_1 \geq f(\boldsymbol{x_1}) \text{ and } (\boldsymbol{x_2}, y_2) : y_2 \geq f(\boldsymbol{x_2})$$

2. Introduce a new point

$$\lambda(\boldsymbol{x_1}, y_1) + (1 - \lambda)(\boldsymbol{x_2}, y_2) = (\lambda \boldsymbol{x_1} + (1 - \lambda)\, \boldsymbol{x_2}, \lambda y_1 + (1 - \lambda)\, y_2)$$

3. Use (1) to show that the point introduced in (2) satisfies the conditions as well

$$\underbrace{\lambda y_1 + (1 - \lambda)\, y_2 \geq \lambda f(\boldsymbol{x_1}) + (1 - \lambda) f(\boldsymbol{x_2})}_{\text{Step (1)}} \geq f(\lambda \boldsymbol{x_1} + (1 - \lambda)\, \boldsymbol{x_2})$$

Convex function *vs.* Convex set (Part II)

$C = \{(\boldsymbol{x}, y): y \geq f(\boldsymbol{x})\}$ is a convex set if $f(\boldsymbol{x})$ is a convex function.

1. Choose any two points satisfying the conditions

$$(\boldsymbol{x_1}, y_1): y_1 \geq f(\boldsymbol{x_1}) \text{ and } (\boldsymbol{x_2}, y_2): y_2 \geq f(\boldsymbol{x_2})$$

2. Introduce a new point

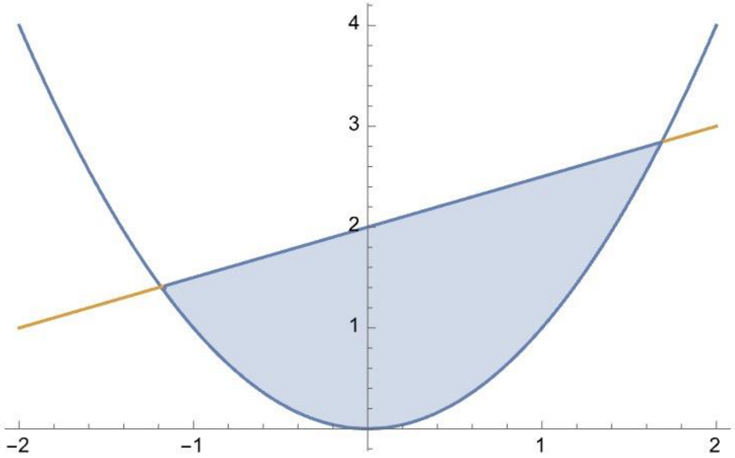
$$\lambda(\boldsymbol{x_1}, y_1) + (1-\lambda)(\boldsymbol{x_2}, y_2) = (\lambda \boldsymbol{x_1} + (1-\lambda)\, \boldsymbol{x_2}, \lambda y_1 + (1-\lambda)\, y_2)$$

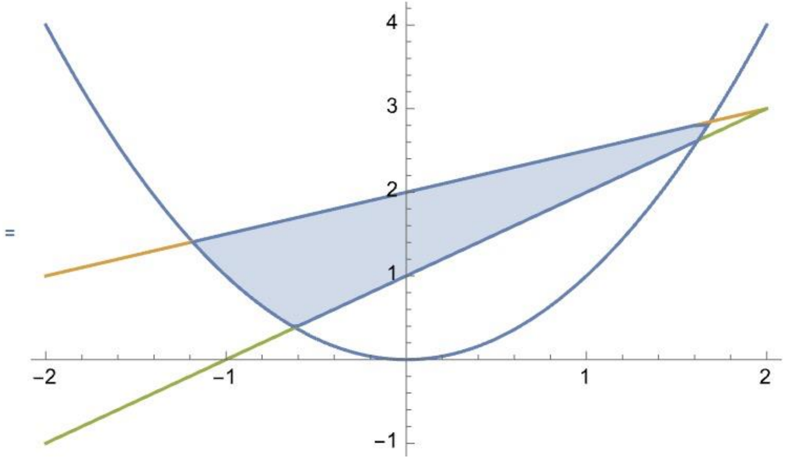
3. Use (1) to show that the point introduced in (2) satisfies the conditions as well

$$\lambda y_1 + (1-\lambda)\, y_2 \geq \underbrace{\lambda f(\boldsymbol{x_1}) + (1-\lambda) f(\boldsymbol{x_2}) \geq f(\lambda \boldsymbol{x_1} + (1-\lambda)\, \boldsymbol{x_2})}_{f \text{ is a convex function}}$$

Application 1

- if $f(x)$ is a convex function, is the following region a convex set?





Application 2

Prove a unit disk, e.g., $\mathrm{H} \coloneqq \{(x, y): x^2 + y^2 \leq 1\}$ is a convex set

Proof: Using the three steps arguments could be complicate.

For any (x_1, y_1) and (x_2, y_2) in H,

- $x_1^2 + y_1^2 \leq 1$
- $x_2^2 + y_2^2 \leq 1$
- $\theta(x_1, y_1) + (1-\theta)(x_2, y_2) = (\theta x_1 + (1-\theta)x_2), \theta y_1 + (1-\theta)y_2)$
- Then for any $\theta \in [0,1]$
  - $\underbrace{(\theta x_1 + (1-\theta)x_2)^2 + (\theta y_1 + (1-\theta)y_2)^2 \leq 1}_{\text{Cauchy-Schwarz Inequality}}$???

$$\left(\sum_{k=1}^{n} a_k\, b_k\right)^2 \leq \left(\sum_{k=1}^{n} a_k^2\right)\left(\sum_{k=1}^{n} b_k^2\right),$$

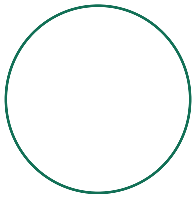
Instead, we consider $f(\boldsymbol{x}) = \sum_i a_i x_i^2$ with a_i>0

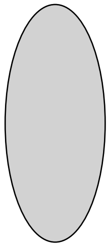
Given any point $\mathbf{y}$, any unit vector $\boldsymbol{e}$ and any θ

$$g(\theta) = f(\boldsymbol{y} + \theta \boldsymbol{e}) = \sum_i a_i (y_i + \theta e_i)^2$$

Then $g''(\theta) = \sum_i a_i e_i^2 \geq 0$. So $f(\boldsymbol{x})$ is a convex function.

$\{\boldsymbol{x}: f(\boldsymbol{x}) \leq r\}$ forms a ball/disk or an ellipsoid, so it is a convex set.





1. Choose any $\mathbf{y}$, any unit vector $\boldsymbol{e}$ and any θ
2. Define $g(\theta) = f(\boldsymbol{y} + \theta \boldsymbol{e})$
3. Prove $g''(\theta) \geq 0$

Convex Optimization Problem

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & g_i(x) \leq 0, i = 1, \dots, m \\ & h_i(x) = 0, i = 1, \dots, n \end{array}$$

A convex optimization problem needs to satisfy the following two conditions:

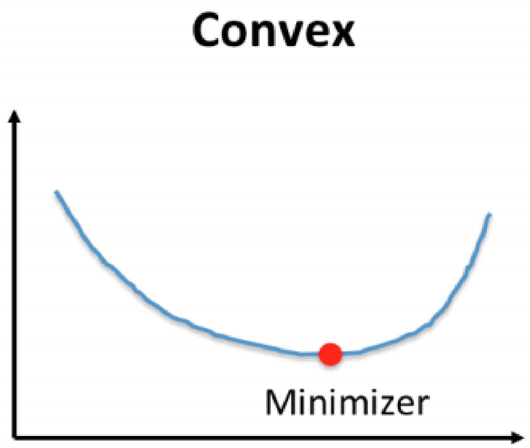
- Its feasible set is a **convex set**.
- Its objective function is a **convex function**.

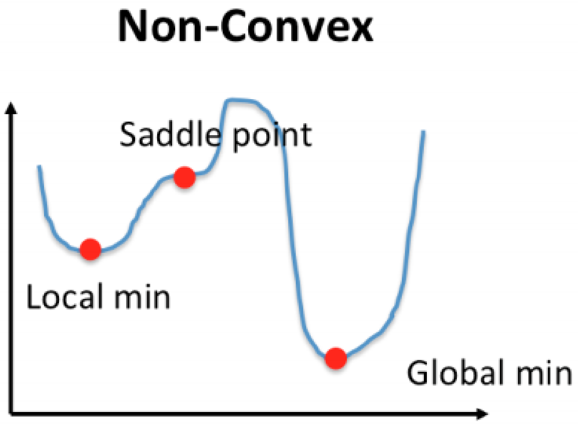
Theorem: For **convex optimization** problems

local minimizer = global minimizer

In other words, any local optimum is also a global optimum.

This theorem, as simple as it is, is one of the most important theorems in convex programming!!





Theorem 1. *Consider an optimization problem*

$$
\begin{aligned}
&min.\ f(x)\\
&s.t.\ \ x \in \Omega,
\end{aligned}
$$

where f is a convex function and Ω is a convex set. Then, any local minimum is also a global minimum.

Proof by contradiction

Proof: Let $\bar{x}$ be a local minimum.

$$\Rightarrow \bar{x} \in \Omega \text{ and } \exists \epsilon > 0 \text{ s.t. } f(\bar{x}) \leq f(x)\ \forall x \in \underbrace{B(\bar{x}, \epsilon)}.$$

B($\mathbf{x}$,ε) = {$\mathbf{y}$: $\|\boldsymbol{y} - \boldsymbol{x}\| \leq \varepsilon$}

Suppose for the sake of contradiction that $\exists z \in \Omega$ with

$$f(z) < f(\bar{x}).$$

Because of convexity of Ω, we have

$$\lambda\bar{x} + (1-\lambda)z \in \Omega, \ \forall \lambda \in [0,1].$$

By convexity of f, we have

$$\begin{aligned} f(\lambda\bar{x} + (1-\lambda)z) &\leq \lambda f(\bar{x}) + (1-\lambda)f(z) \\ &< \lambda f(\bar{x}) + (1-\lambda)f(\bar{x}) = f(\bar{x}). \end{aligned}$$

But, as $\lambda \to 1$, $\boxed{(\lambda\bar{x} + (1-\lambda)z)} \to \bar{x}$ and the previous inequality contradicts local optimality of $\bar{x}$. $\square$

This point will eventually move into the ball $B(\bar{x}, \varepsilon)$