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

Introduction to Data Science

Lecture 19 Optimization (Review)

Zicheng Wang

Continued From Last Lecture

Example 6

For $\boldsymbol{x}, \boldsymbol{y} \in R^n$, we denote by $||\boldsymbol{x} - \boldsymbol{y}||$ the distance between $\boldsymbol{x}$ and $\boldsymbol{y}$, i.e.,

$$||\boldsymbol{x} - \boldsymbol{y}|| = \sqrt{\sum_{\mathrm{i}=1}^{\mathrm{n}} (x_i - y_i)^2}.$$

$f(\boldsymbol{x}, \boldsymbol{y}) = ||\boldsymbol{x} - \boldsymbol{y}||$ is a also a convex function.

Proof not required.

Minimization

- We already know that the maximum or supremum of a set of convex functions is convex

- Some <u>special</u> forms (NOT ALL) of minimization also preserve convexity

- If f is convex in (x, y), and C is **a convex nonempty set**, then the function

$$g(x) = \inf_{y \in C} f(x, y)$$

is convex.

Proof not required.

Example 7

Give a point $\boldsymbol{x} \in R^n$, and a convex set $\mathrm{C} \in R^n$. Let

$$f(\boldsymbol{x}) = inf_{\boldsymbol{y} \in C} ||\boldsymbol{x} - \boldsymbol{y}||,$$

i.e., the distance between $\boldsymbol{x}$ and the nearest point of C.

Show that $f(\boldsymbol{x})$ is a convex function.

Example 7

We know that $h(\boldsymbol{x}, \boldsymbol{y}) = ||\boldsymbol{x} - \boldsymbol{y}||$ is a convex function

By the Minimization result, we can obtain that $f(\boldsymbol{x}) = inf_{\boldsymbol{y} \in C}||\boldsymbol{x} - \boldsymbol{y}||$ is a convex function

Example 8

What about the distance between $\boldsymbol{x}$ and the farthest point of C?

$$f(\boldsymbol{x}) = sup_{\boldsymbol{y} \in C} ||\boldsymbol{x} - \boldsymbol{y}||,$$

Do we need the condition that C is a convex set?

The maximum of a set of convex functions is convex.

We impose no condition on the set. In addition, we only require that $f(\boldsymbol{x}) = ||\boldsymbol{x} - \boldsymbol{y}||$ is convex for fixed $\boldsymbol{y}$.

Review

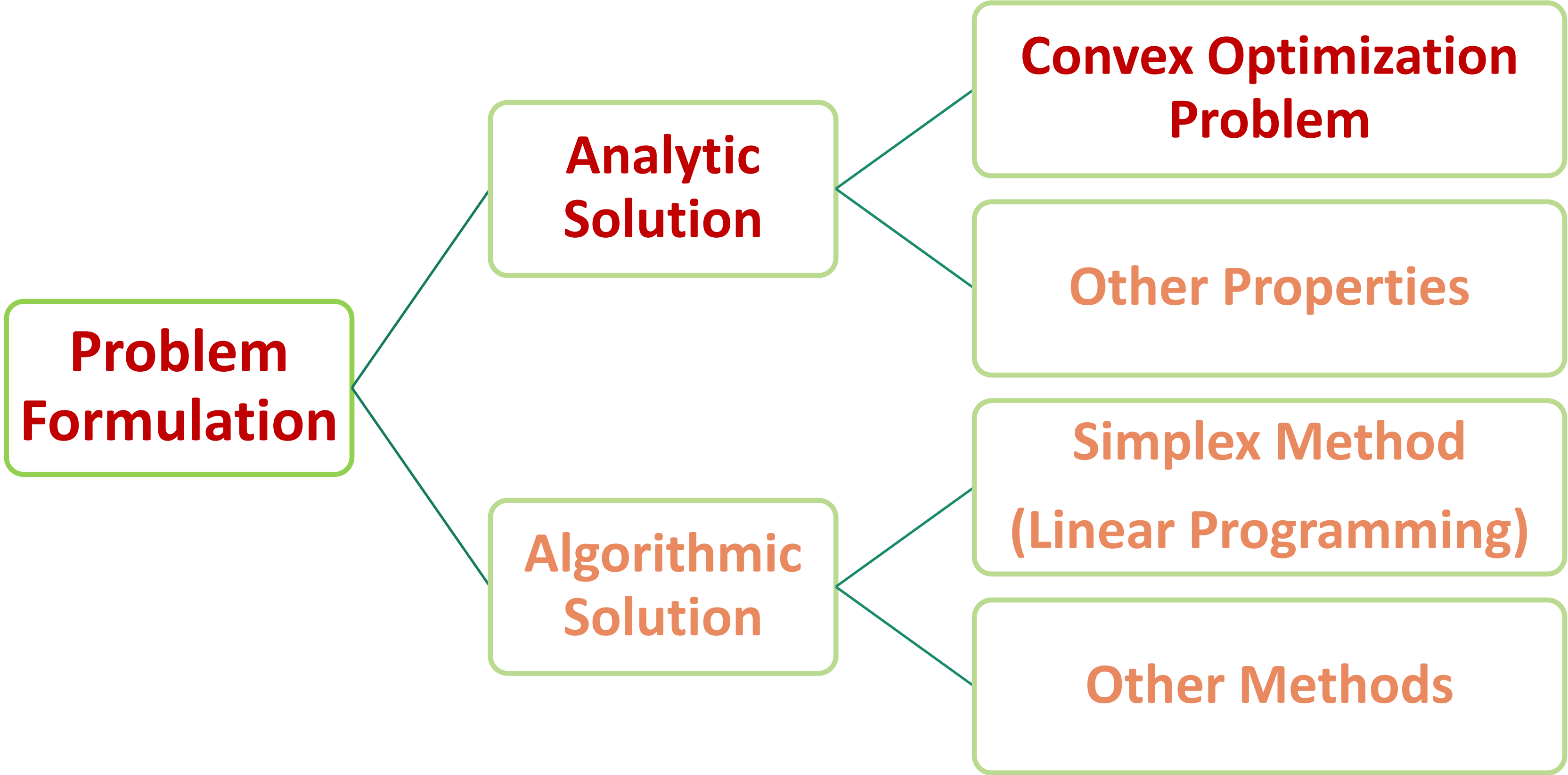
Mathematical optimization (alternatively spelled *optimisation*) or **mathematical programming** is the selection of a best element, with regard to some criterion, from some set of available alternatives.

Objective Function

Constraints

Given: a function f: $A \to \mathbb{R}$ from some set A to the real numbers
Sought: an element $\mathbf{x}_0 \in A$ such that $f(\mathbf{x}_0) \leq f(\mathbf{x})$ for all $\mathbf{x} \in A$ ("minimization") or such that $f(\mathbf{x}_0) \geq f(\mathbf{x})$ for all $\mathbf{x} \in A$ ("maximization").

Decision Variables


Problem Formulation
Analytic Solution
Algorithmic Solution
Convex Optimization Problem
Other Properties
Simplex Method (Linear Programming)
Other Methods

Problem Formulation

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \leq 0, \quad i = 1, \ldots, m \\ & h_i(x) = 0, \quad i = 1, \ldots, p \end{array}$$

- $x \in \mathbf{R}^n$ is the optimization variable
- $f_0 : \mathbf{R}^n \to \mathbf{R}$ is the objective or cost function
- $f_i : \mathbf{R}^n \to \mathbf{R}$, $i = 1, \ldots, m$, are the inequality constraint functions
- $h_i : \mathbf{R}^n \to \mathbf{R}$ are the equality constraint functions

Terminologies

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

- **Feasible Set**: the set of all points such that the constraints can **all** be satisfied.

- $\boldsymbol{x}^*$ is the **global minimizer**, if $f(\boldsymbol{x}^*) \leq f(\boldsymbol{y})$ for any $\boldsymbol{y}$ in the feasible set.

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

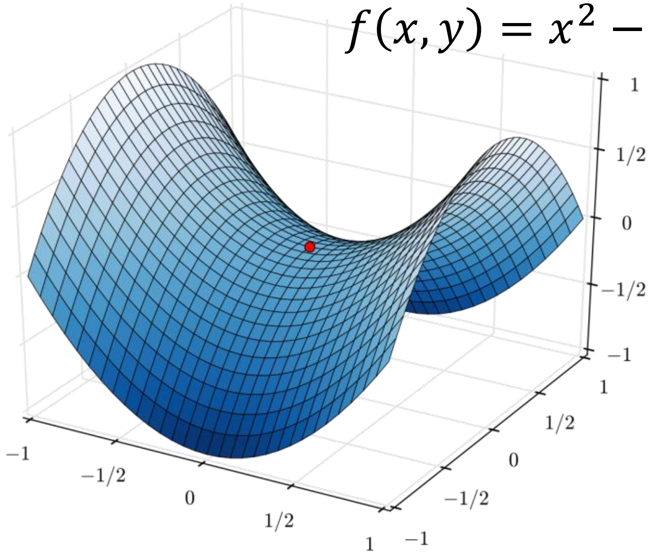
Terminologies

$$\begin{aligned} \text{minimize} \quad & f(\boldsymbol{x}) \\ \text{subject to} \quad & g_i(\boldsymbol{x}) \leq 0, i = 1, \ldots, m \\ & h_i(\boldsymbol{x}) = 0, i = 1, \ldots, n \end{aligned}$$

- **Local minimizer**
  - Denote S as the feasible set
  - Denote B(**x**,ε) = {**y**: $\|\boldsymbol{y} - \boldsymbol{x}\| \leq \varepsilon\}$ as the set of all points such that the distance from x and each point in the set is smaller than ε.
  - If there exists an $\varepsilon > 0$ such that for any $\boldsymbol{y} \in S \cap$ B($\boldsymbol{x}^*$,ε), f($\boldsymbol{x}^*$)$\leq f(\boldsymbol{y})$. Then $\boldsymbol{x}^*$ is called a local minimizer of the optimization problem.

Remark: In general, the first order condition you learned from Calculus I, $\frac{df(\boldsymbol{x}^*+t\boldsymbol{e})}{dt}\Big|_{t=0} = 0$ for any $\boldsymbol{e}$, does not guarantee a local minimizer or maximizer.

$$f(x, y) = x^2 - y^2$$



(0,0) is a saddle point

Why Convex Optimization Problem?

- Any local minimum is also a global minimum.
- Any interior local minimum satisfies the first order condition.

$$\nabla f(p) = \begin{bmatrix} \frac{\partial f}{\partial x_1}(p) \\ \vdots \\ \frac{\partial f}{\partial x_n}(p) \end{bmatrix} \longrightarrow \nabla f(\boldsymbol{x}^*) = \boldsymbol{0}$$

Convex Optimization Problem

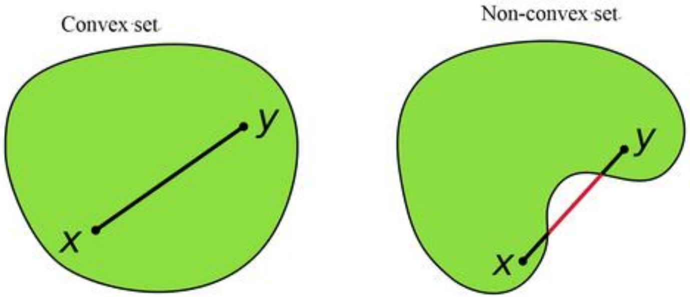
$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & g_i(x) \le 0, i = 1, \dots, m \\ & h_i(x) = 0, i = 1, \dots, n \end{array}$$

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**.

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 $\theta \in [0,1]$

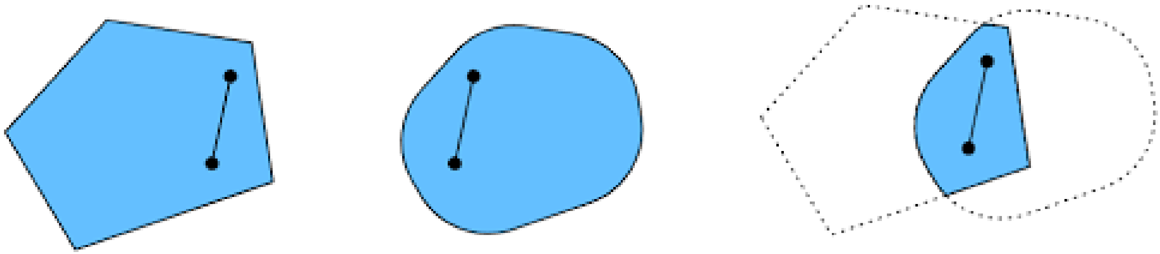
2. Characterize the new point within the line segment

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

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.



Remark: The union of two convex sets is in general not a convex set

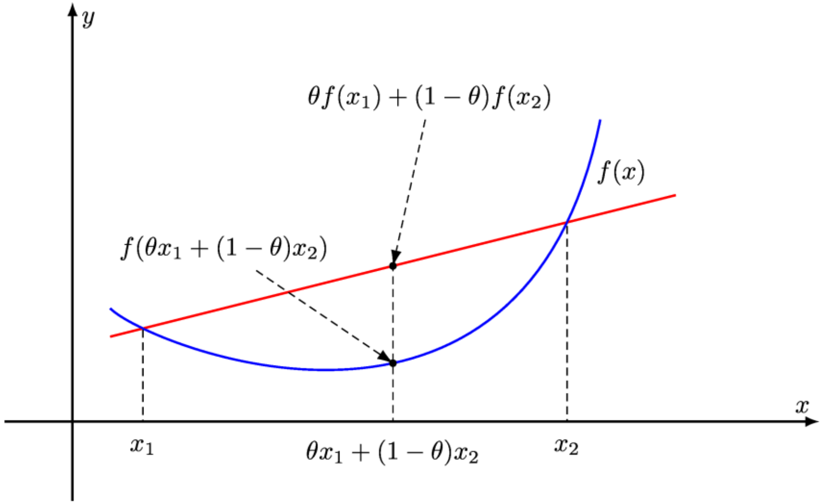
Convex Function

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 $\boldsymbol{f(x_2)}$

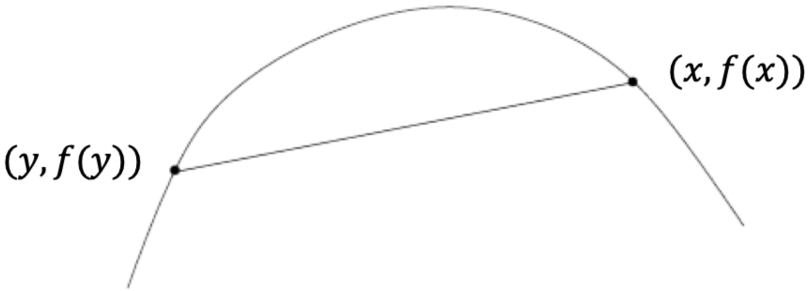


Concave Function

Definition: A function $f(\boldsymbol{x}): R^n \to R$ 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 $\boldsymbol{z} = \lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}$.



If f is concave, then -f is convex!

If f is convex, then -f is concave!

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 $\boldsymbol{e}$ satisfying that there exists $\epsilon > 0$ such that $\boldsymbol{x} + \epsilon\boldsymbol{e} \in$ dom(f),

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

One dimension: $\boldsymbol{f''(x) \geq 0}$

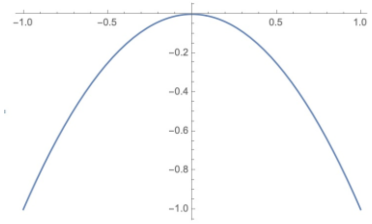
Examples of Convex/Concave Functions

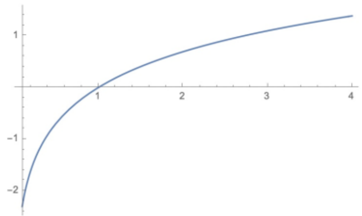
Convex

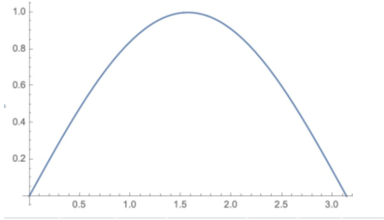
- $f(x) = ax + b$ $(also\ concave)$
- $f(x) = x^2$
- $f(x) = e^x$

Concave

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







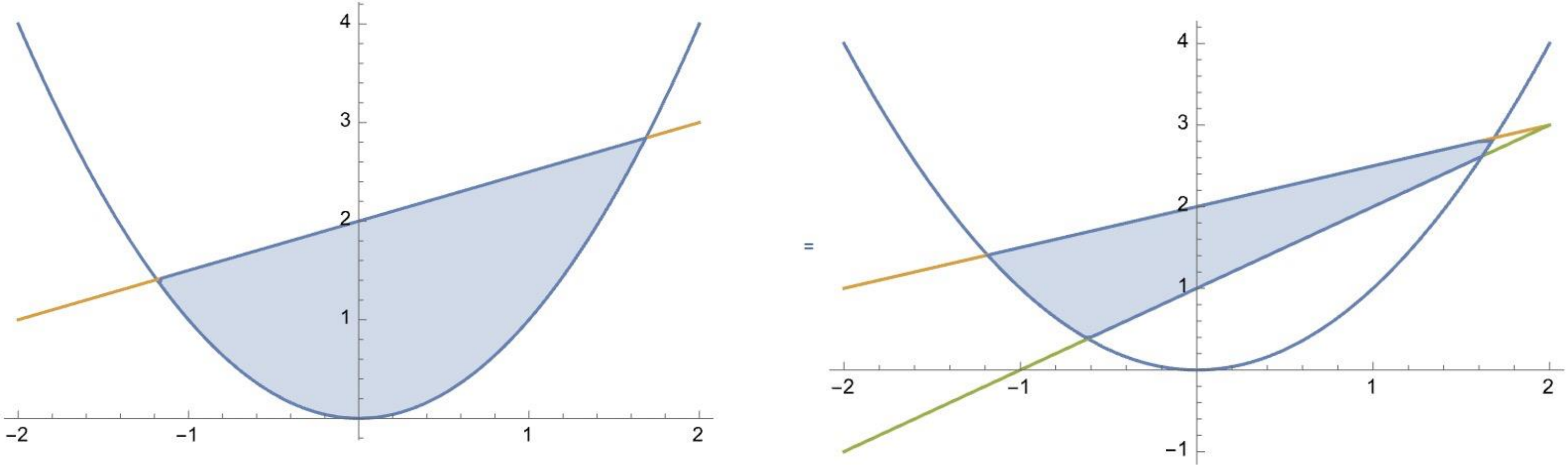
Convex Function VS. Convex Set

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

- $C = \{(\boldsymbol{x}, y): y \geq f(\boldsymbol{x})\}$ is a convex set if and only if $f(\boldsymbol{x})$ is a convex function.
  - C is also called the epigraph of $f(\boldsymbol{x})$

Application 1 (convex function ⇒ convex set)

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



Intersection of the epigraph of a convex function and a convex set

Application 2 (convex function ⇒ convex set)

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

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

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

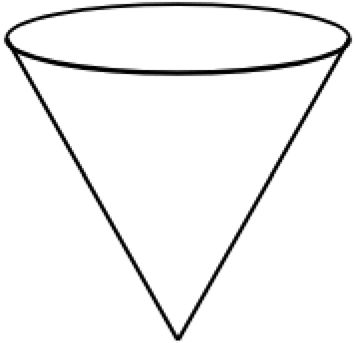
Then g''$(0) = \sum_i 2 a_i e_i^2 \geq 0$. So $f(\boldsymbol{x})$ is a convex function.

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

Application 3 (convex set ⇒ convex function)

Prove $\mathrm{f}(\mathbf{x}) = \sqrt{\sum_{\mathrm{i}=1}^{\mathrm{n}} x_i^2}$ is a convex function

Its epigraph is a cone (and thus a convex set), which implies that $\mathrm{f}(\mathbf{x})$ is a convex function.



Operations Preserving Convexity

If $f_1, \dots, f_m$ are convex functions, then $f(\boldsymbol{x}) = \max\{f_1(\boldsymbol{x}), \dots, f_m(\boldsymbol{x})\}$ is also convex.

- **Maximum of a set of convex functions**

If $f_1, f_2, \dots, f_n$ are convex functions, and $w_1, w_2, \dots, w_n \geq 0$, then $f = w_1 f_1 + w_2 f_2 + \cdots + w_n f_n$ is also a convex function.

- **Nonnegative weighted sums of convex functions**
- **This result can be generalized to integration**

These properties extend to infinite sums and integrals. For example if $f(x, y)$ is convex in x for each $y \in \mathcal{A}$, and $w(y) \geq 0$ for each $y \in \mathcal{A}$, then the function g defined as

$$g(x) = \int_{\mathcal{A}} w(y) f(x, y)\, dy$$

is convex in x (provided the integral exists).

Operations Preserving Convexity

If f is convex, then $g(\boldsymbol{x}) = f(a\boldsymbol{x} + \boldsymbol{b}) - c$ is also convex.

- **Composition with affine function**

- If f is convex in (x, y), and C is a convex nonempty set, then the function

$$g(x) = \inf_{y \in C} f(x, y)$$

  is convex.

  - **Minimization of jointly convex function over a convex nonempty set**