Introduction to Convex Optimization

Lê Xuân Thanh

Contents

Contents

Convex sets in $\mathbb{R}^n$

Definition

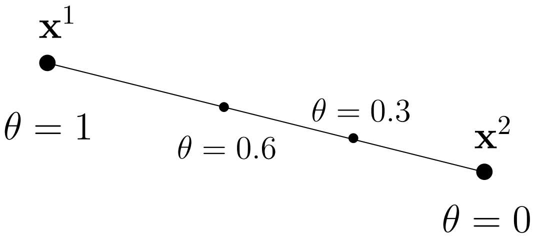
A set $C \subset \mathbb{R}^n$ is *convex* if for any $\mathbf{x}^1, \mathbf{x}^2 \in C$ and $\theta \in [0,1]$ we have $\theta \mathbf{x}^1 + (1-\theta)\mathbf{x}^2 \in C$

Examples:

- Let $\mathbf{x}^1, \mathbf{x}^2$ be two distinct points in $\mathbb{R}^n$. The set

  $$[\mathbf{x}^1, \mathbf{x}^2] := \{\mathbf{y} \in \mathbb{R}^n \mid \mathbf{y} = \theta \mathbf{x}^1 + (1-\theta)\mathbf{x}^2 \text{ for some } \theta \in [0,1]\}$$

  forms the *line segment* between $\mathbf{x}^1$ and $\mathbf{x}^2$



Convex sets in $\mathbb{R}^n$

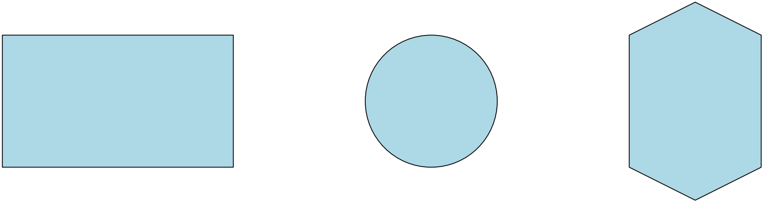
Definition

A set $C \subset \mathbb{R}^n$ is *convex* if for any $\mathbf{x}^1, \mathbf{x}^2 \in C$ and $\theta \in [0,1]$ we have $\theta \mathbf{x}^1 + (1-\theta)\mathbf{x}^2 \in C$

Intuition: C is convex if $[\mathbf{x}^1, \mathbf{x}^2] \subset C$ for any $\mathbf{x}^1, \mathbf{x}^2 \in C$

Examples:

- Convex sets in $\mathbb{R}^2$



Convex sets in $\mathbb{R}^n$

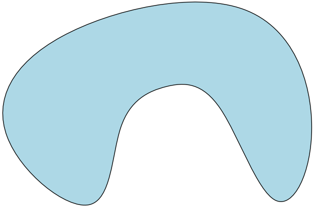
Definition

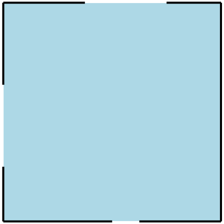
A set $C \subset \mathbb{R}^n$ is *convex* if for any $\mathbf{x}^1, \mathbf{x}^2 \in C$ and $\theta \in [0,1]$ we have $\theta\mathbf{x}^1 + (1-\theta)\mathbf{x}^2 \in C$

Intuition: C is convex if $[\mathbf{x}^1, \mathbf{x}^2] \subset C$ for any $\mathbf{x}^1, \mathbf{x}^2 \in C$

Examples:

- Nonconvex sets in $\mathbb{R}^2$





Important examples of convex sets in $\mathbb{R}^n$

- *Affine set:*
  A set $A \subset \mathbb{R}^n$ is *affine* if for any $\mathbf{x}^1, \mathbf{x}^2 \in A$ and $\theta \in \mathbb{R}$ we have $\theta\mathbf{x}^1 + (1-\theta)\mathbf{x}^2 \in A$
- *Half space:*
  $$\mathbf{a}^t\mathbf{x} \leq b \quad \text{for } \mathbf{a} \in \mathbb{R}^n, b \in \mathbb{R}$$
- *Solution set of system of linear inequalities:*
  $$A\mathbf{x} \leq \mathbf{b} \quad \text{for } A \in \mathbb{R}^{m\times n}, \mathbf{b} \in \mathbb{R}^m$$
- *Ball:*
  $$\|\mathbf{x} - \mathbf{x}_c\| \leq r \quad \text{for given } \mathbf{x}_c \in \mathbb{R}^n \text{ and } r > 0$$

Contents

Convex combination

Definition

A *convex combination* of the points $\mathbf{x}^1, \ldots, \mathbf{x}^k \in \mathbb{R}^n$ is a point of the form

$$\theta_1 \mathbf{x}^1 + \ldots + \theta_k \mathbf{x}^k$$

where $\theta_1, \ldots, \theta_k \in [0, 1]$ satisfying $\theta_1 + \ldots + \theta_k = 1$

Examples:

- Let $\mathbf{x}^1, \mathbf{x}^2$ be two distinct points in $\mathbb{R}^n$. Any point on the line segment between $\mathbf{x}^1, \mathbf{x}^2$ is a convex combination of $\mathbf{x}^1, \mathbf{x}^2$.
- Let $\mathbf{x}^1, \mathbf{x}^2, \mathbf{x}^3$ be three distinct points in $\mathbb{R}^2$ that are not on the same line. Any point $\mathbf{x}$ inside the triangle whose vertices are $\mathbf{x}^1, \mathbf{x}^2, \mathbf{x}^3$ is a convex combination of these three vertices.

Convex hull

Definition

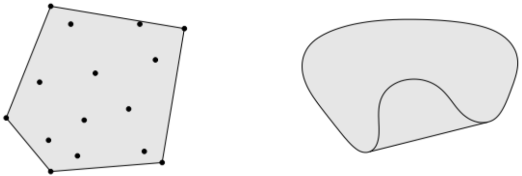
The *convex hull* of a set $C \subset \mathbb{R}^n$ is

$$\begin{aligned}\operatorname{conv}(C) = \{\theta_1 \mathbf{x}^1 + \ldots + \theta_k \mathbf{x}^k \mid \mathbf{x}^i \in C, \theta_i \geq 0, i = 1, \ldots, k,\\ \theta_1 + \ldots + \theta_k = 1\}\end{aligned}$$

Proposition

The convex hull of a set $C \subset \mathbb{R}^n$ is the smallest convex set containing C (in sense of set inclusion)

Corollary: $C \subset \mathbb{R}^n$ is convex if and only if $C = \operatorname{conv}(C)$



Contents

- Definition
- Convex hull
- Extreme point

Extreme points of convex sets

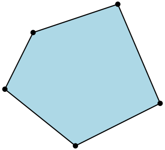
Definition

A point $\mathbf{x}$ is an *extreme point* of a convex set $C \subset \mathbb{R}^n$ if

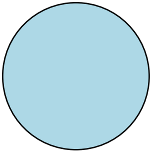
- $\mathbf{x} \in C$, and
- for any $\mathbf{y}, \mathbf{z} \in C$ together with $\lambda \in (0, 1)$, the equality $\mathbf{x} = \lambda \mathbf{y} + (1 - \lambda)\mathbf{z}$ implies that $\mathbf{x} = \mathbf{y} = \mathbf{z}$

In other words: an extreme point of C is a point in C that cannot be an interior point of a line segment lying in C

No extreme point



Finite extreme points



Infinite extreme points

Extreme points of convex sets

Definition

A point $\mathbf{x}$ is an *extreme point* of a convex set $C \subset \mathbb{R}^n$ if

- $\mathbf{x} \in C$, and
- for any $\mathbf{y}, \mathbf{z} \in C$ together with $\lambda \in (0, 1)$, the inequality $\mathbf{x} = \lambda \mathbf{y} + (1 - \lambda)\mathbf{z}$ implies that $\mathbf{x} = \mathbf{y} = \mathbf{z}$

Remarks:

- Extreme point is boundary point
- Boundary point may not be extreme point

Krein[1]-Milman[2] theorem

Every nonempty compact convex set $C \subset \mathbb{R}^n$ is the convex hull of its extreme points

---

[2]Mark Grigorievich Krein (03.04.1907-17.10.1989): a Soviet mathematician

[2]David Pinhusovich Milman (15.01.1912-12.07.1982): a Soviet and later Israeli mathematician

Contents

Contents

Convex function

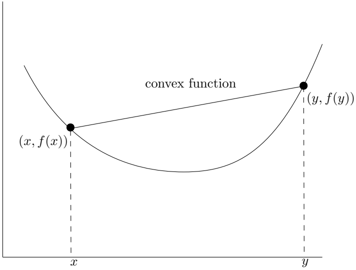
Definition

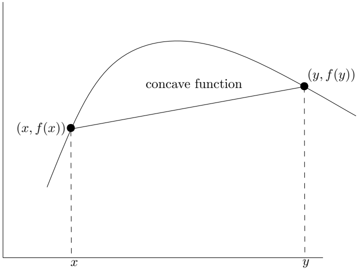
A function $f : \mathbb{R}^n \to \mathbb{R}$ is *convex* if

- $\mathrm{dom}(f)$ is a convex set, and
- for all $\mathbf{x}, \mathbf{y} \in \mathrm{dom}(f)$ and $\theta \in [0, 1]$ we have

$$f(\theta\mathbf{x} + (1-\theta)\mathbf{y}) \leq \theta f(\mathbf{x}) + (1-\theta) f(\mathbf{y})$$

Intuition: Graph of f is under the secant line





Strictly convex function

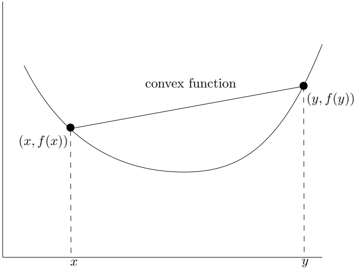
Definition

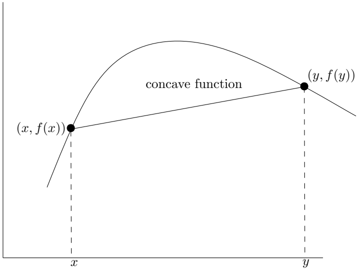
A function $f: \mathbb{R}^n \to \mathbb{R}$ is *strictly convex* if

- $\mathrm{dom}(f)$ is a convex set, and
- for all $\mathbf{x}, \mathbf{y} \in \mathrm{dom}(f)$ and $\theta \in [0,1]$ we have

$$f(\theta \mathbf{x} + (1-\theta)\mathbf{y}) < \theta f(\mathbf{x}) + (1-\theta) f(\mathbf{y})$$

Intuition: Graph of f is strictly under the secant line



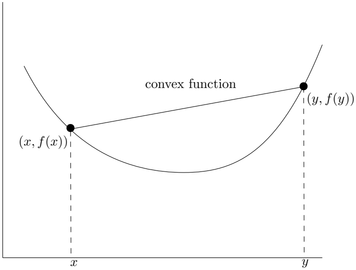


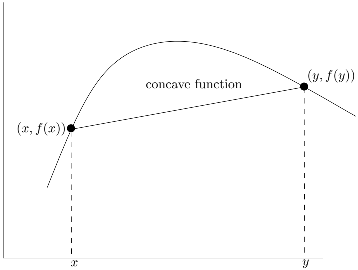
Concave and strictly concave functions

Definition

- A function $f: \mathbb{R}^n \to \mathbb{R}$ is *concave* if $-f$ is convex.
- A function $f: \mathbb{R}^n \to \mathbb{R}$ is *strictly concave* if $-f$ is strictly convex.

Intuition: Graph of f is above the secant line





Contents

- Definition
- Recognize convexity
  - Via epigraph
  - First-order condition
  - Second-order condition

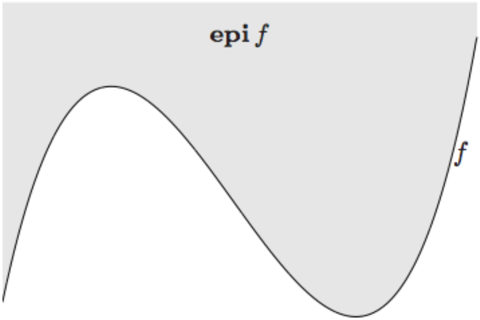
Epigraph

Definition

The *epigraph* of a function $f: \mathbb{R}^n \to \mathbb{R}$ is defined as

$$\mathrm{epi}(f) = \{(\mathbf{x}, t) \in \mathbb{R}^{n+1} \mid \mathbf{x} \in \mathrm{dom}(f), f(\mathbf{x}) \leq t\}.$$

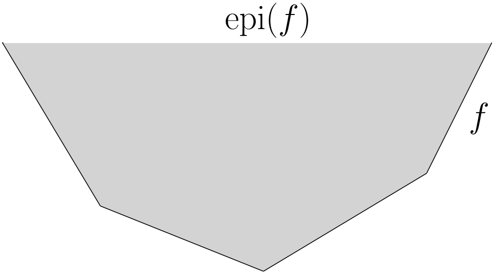
Intuition: Epigraph is the area above the graph.



Epigraph of convex function

Proposition

A function $f\colon \mathbb{R}^n \to \mathbb{R}$ is convex if and only if $\mathrm{epi}(f)$ is convex.



Contents

- Definition
- Recognize convexity
  - Via epigraph
  - First-order condition
  - Second-order condition

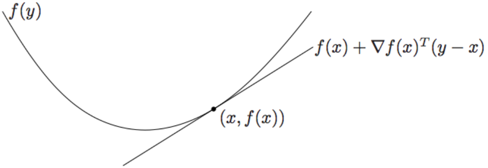
First-order condition for convexity of function

Theorem

Suppose $f : \mathbb{R}^n \to \mathbb{R}$ is differentiable. Then f is convex if and only if $\mathrm{dom}(f)$ is convex and

$$f(\mathbf{y}) \geq f(\mathbf{x}) + \nabla f(\mathbf{x})^T(\mathbf{y} - \mathbf{x}) \qquad \forall \mathbf{x}, \mathbf{y} \in \mathrm{dom}(f).$$

Intuition: Graph of f is always above its tangent hyperplanes



Contents

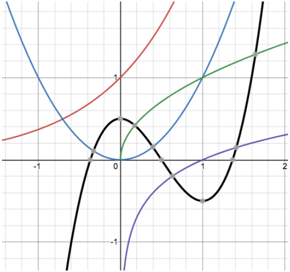
Second-order condition for convexity of function

Theorem (1-dimensional case)

Suppose $f : \mathbb{R} \to \mathbb{R}$ is twice differentiable. Then f is convex if and only if $\mathrm{dom}(f)$ is convex and $f''(x) \geq 0$ for all $x \in \mathrm{dom}(f)$.

Examples:

- e^{ax} is convex on $\mathbb{R}$ for any $a \in \mathbb{R}$
- x^a is convex on $\mathbb{R}_{++}$ when $a \geq 1$ or $a \leq 0$
- $\log x$ is concave on $\mathbb{R}_{++}$



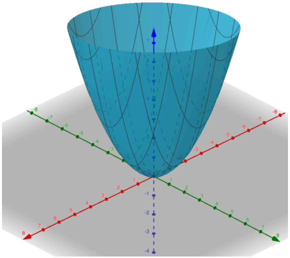
Second-order condition for convexity of function

Theorem

Suppose $f : \mathbb{R}^n \to \mathbb{R}$ is twice differentiable. Then f is convex if and only if $\mathrm{dom}(f)$ is convex and its Hessian matrix $\nabla^2 f(\mathbf{x})$ is positive semidefinite for all $\mathbf{x} \in \mathrm{dom}(f)$.

Examples: $f(x, y) = x^2 + xy + y^2$ is convex on $\mathbb{R}^2$ as

$$\nabla^2 f(x, y) = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \succcurlyeq 0$$



Contents

Formulation

General form:

$$\min \quad f(\mathbf{x}) \quad \text{s.t.} \quad \mathbf{x} \in C \subset \mathrm{dom}(f)$$

in which $f : \mathbb{R}^n \to \mathbb{R}$ is a convex function and C is a convex set.

Explicit form of feasible set:

$$C = \{\mathbf{x} \in \mathrm{dom}(f) \mid g_i(\mathbf{x}) \leq 0, i = 1, \ldots, k; \mathbf{a}_i^t \mathbf{x} + b_j = 0, j = 1, \ldots, m\}$$

in which $g_i (i = 1, \ldots, k)$ are convex functions.

Notes:

- g_i is convex $\Rightarrow [g_i \leq 0]$ is a convex set $\Rightarrow \cap_{i=1}^m [g_i \leq 0]$ is convex
- Another form:

$$\max \quad h(\mathbf{x}) \quad \text{s.t.} \quad \mathbf{x} \in C \subset \mathrm{dom}(f)$$

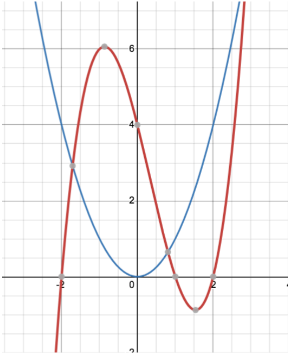
in which $h : \mathbb{R}^n \to \mathbb{R}$ is a concave function and C is a convex set.

Contents

Solution existence

Important theorem

- If x^* is a local minimizer of a convex optimization problem, it is also a global minimizer.
- In addition, if the objective function is strictly convex, then x^* is the unique global minimizer.



Corollary

Consider the convex optimization problem (P):

$$\begin{array}{lll} \min & f(\mathbf{x}) & \\ \text{s.t.} & g_i(\mathbf{x}) \leq 0 & i = 1, \ldots, k \\ & h_j(\mathbf{x}) = \mathbf{a}_j^t \mathbf{x} + b_j = 0 & j = 1, \ldots, m \end{array}$$

Lagrangian function: $\mathcal{L}(\mathbf{x}, \lambda, \mu) = f(\mathbf{x}) + \sum_{i=1}^{k} \lambda_i g_i(\mathbf{x}) + \sum_{j=1}^{m} \mu_j h_j(\mathbf{x})$

KKT system:

1. $\nabla_{\mathbf{x}} \mathcal{L}(\mathbf{x}, \lambda, \mu) = \mathbf{0}$
2. $g_i(\mathbf{x}) \leq 0$ for $i = 1, \ldots, k$
3. $\lambda_i \geq 0$ for $i = 1, \ldots, k$
4. $\lambda_i g_i(\mathbf{x}) = 0$ for $i = 1, \ldots, k$
5. $h_j(\mathbf{x}) = 0$ for $j = 1, \ldots, m$

Theorem

If (x^*, λ^*, μ^*) solves the above KKT system, then x^* is a global minimizer of (P).

Example 1

Consider the quadratic programming problem

$$\min_{\mathbf{x}\in\mathbb{R}^n} \quad \frac{1}{2}\mathbf{x}^t Q\mathbf{x} + \mathbf{c}^t\mathbf{x}$$

where Q is a positive semidefinite matrix.

Lagrange function

$$\mathcal{L}(\mathbf{x}) = \frac{1}{2}\mathbf{x}^t Q\mathbf{x} + \mathbf{c}^t\mathbf{x}$$

KKT system:

$$\nabla\mathcal{L}(\mathbf{x}) = \mathbf{0} \quad \Leftrightarrow \quad Q\mathbf{x} + \mathbf{c} = \mathbf{0}$$

Theorem

x^* is global optimal of the above quadratic programming problem if and only if $Q\mathbf{x} + \mathbf{c} = \mathbf{0}$.

Example 2

Consider the quadratic programming problem

$$\min \quad \frac{1}{2}\mathbf{x}^t Q\mathbf{x} + \mathbf{c}^t\mathbf{x} \qquad \text{s.t.} \qquad A\mathbf{x} = \mathbf{0}$$

where Q is a positive semidefinite matrix.

Lagrange function

$$\mathcal{L}(\mathbf{x}, \mu) = \frac{1}{2}\mathbf{x}^t Q\mathbf{x} + \mathbf{c}^t\mathbf{x} + \mu^t A\mathbf{x}$$

KKT system:

$$\begin{cases} \nabla_{\mathbf{x}}\mathcal{L}(\mathbf{x}) & = \mathbf{0} \\ \nabla_{\mu}\mathcal{L}(\mathbf{x}) & = \mathbf{0} \end{cases} \Leftrightarrow \begin{cases} Q\mathbf{x} + \mathbf{c} + A^t\mu & = \mathbf{0} \\ A\mathbf{x} & = \mathbf{0} \end{cases} \Leftrightarrow \begin{bmatrix} Q & A^t \\ A & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mu \end{bmatrix} = \begin{bmatrix} -\mathbf{c} \\ \mathbf{0} \end{bmatrix}$$

Thank you for your attention!