Convex optimization III

Source: Ahmadi, ORF 363 slides.

Convex functions

Epigraph and sublevel sets

Convex optimization problems

Overview

- In a convex problem, every local minimum is automatically a global minimum

Overview

- In a convex problem, every local minimum is automatically a global minimum

- In the unconstrained case, every stationary point (i.e., zero of the gradient) is automatically a global minimum.

Overview

- In a convex problem, every local minimum is automatically a global minimum
- In the unconstrained case, every stationary point (i.e., zero of the gradient) is automatically a global minimum.
- We will also see new characterizations for convex functions that make the task of checking convexity somewhat easier

Review

Convex optimization problem:

$$\min\ f(x)$$

subject to $x \in \Omega$

where

$$x \in \mathbb{R}^n, \quad f : \mathbb{R}^n \to \mathbb{R} \quad \Omega = \{x \mid g_i(x) \leq 0, h_j(x) = 0\}$$

Review

Consider a function $f : \mathbb{R}^n \rightarrow \mathbb{R}$
A point $x^* \in \mathbb{R}^n$ is said to be a:

- **feasible** if $x^* \in \Omega$ i.e. $g_i(x^*) \leq 0, \forall i, h_j(x^*) \leq 0, \forall j$

Review

Consider a function $f : \mathbb{R}^n \to \mathbb{R}$
A point $x^* \in \mathbb{R}^n$ is said to be a:

- **feasible** if $x^* \in \Omega$ i.e. $g_i(x^*) \leq 0, \forall i, h_j(x^*) \leq 0, \forall j$
- Local minimum: $f(x^*) \leq f(x)$ $\forall x$ near x^* ($\|x - x^*\|$ small)
- Strict local minimum: $f(x^*) < f(x)$ $\forall x$ near x^* ($\|x - x^*\|$ small)
- Global minimum: $f(x^*) \leq f(x)$ $\forall x \in \mathbb{R}^n$
- Strict global minimum: $f(x^*) < f(x)$ $\forall x \in \mathbb{R}^n$

Local implies global minimum

Theorem. Consider an optimization problem

$$\min \ f(x)$$
$$\text{s.t. } x \in \Omega$$

where $f(x)$ is a convex function and Ω is a convex set. Then, every local minimum is also a global minimum.

Characterization of convex functions

Theorem. Suppose $f : \mathbb{R}^n \to \mathbb{R}$ is twice differentiable over its domain. Then, the following are equivalent:

1. f is convex
2. $f(y) \geq f(x) + \nabla f(x)^T (y - x), \quad \forall x, y \in \text{dom}(f)$
3. $\nabla^2 f(x) \succeq 0, \quad \forall x \in \text{dom}(f)$

Characterization of convex functions

Theorem. Suppose $f : \mathbb{R}^n \to \mathbb{R}$ is twice differentiable over its domain. Then, the following are equivalent:

1. f is convex
2. $f(y) \geq f(x) + \nabla f(x)^T (y - x), \quad \forall x, y \in \text{dom}(f)$
3. $\nabla^2 f(x) \succeq 0, \quad \forall x \in \text{dom}(f)$

Unconstrained convex optimization

Corollary. Consider an unconstrained optimization problem:

$$\min f(x)$$
$$\text{s.t. } x \in \Omega$$

where f is convex and differentiable. Then, any point x_0 that satisfies $\nabla f(x_0) = 0$ is a global minimum

Remarks

- $\nabla f(x_0) = 0$ implies local and **global** optimality

Remarks

- $\nabla f(x_0) = 0$ implies local and **global** optimality
- In absence of convexity, $\nabla f(x_0) = 0$ is not sufficient even for local optimality (e.g., think of $f(x) = x^3$ and $x_0 = 0$)

Remarks

- $\nabla f(x_0) = 0$ implies local and **global** optimality
- In absence of convexity, $\nabla f(x_0) = 0$ is not sufficient even for local optimality (e.g., think of $f(x) = x^3$ and $x_0 = 0$)
- Necessary condition for (unconstrained) local optimality of a point x_0 was $\nabla^2 f(x) \succeq 0$
  - A convex function automatically passes this test

Quadratic functions revisited

Let $f(x) = x^T Ax + bx + c$ (A is symmetric)

When is f convex?

Quadratic functions revisited

$$\min_x x^T A x + bx + c \quad A \text{ is symmetric}$$

$A \not\succeq 0$ (f is not convex) $\Rightarrow$ unbounded below

Quadratic functions revisited

$$\min_x x^T A x + bx + c \quad A \text{ is symmetric}$$

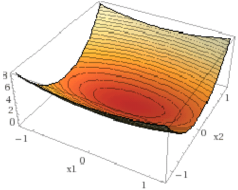
$A \succ 0 \Rightarrow f$ convex

Quadratic functions revisited

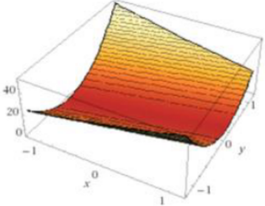
$$\min_x x^T A x + bx + c \quad A \text{ is symmetric}$$

$A \succ 0 \Rightarrow f$ convex

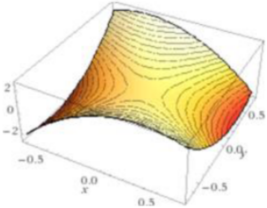
$A \succeq 0 \Rightarrow f$ convex, but there might be many optimal solutions



$A \succ 0$



$A \succcurlyeq 0$
(but $A \nsucc 0$)



$A \nsucccurlyeq 0$

Least squares, revisited

$$\min_x \|Ax - b\|^2$$

$A \in \mathbb{R}^{m \times n}$ (columns of A are linearly independent)
$b \in \mathbb{R}^m$

$$\nabla f(x) = 0 \Rightarrow x = (A^T A)^{-1} A^T b$$

Strict Convexity

Recall that $f : \mathbb{R}^n \to \mathbb{R}$ is strictly convex if $\forall x, y, x \neq y, \forall \lambda \in (0, 1)$,

$$f(\lambda x + (1 - \lambda)y) < \lambda f(x) + (1 - \lambda)f(y)$$

- f is strictly convex $\Rightarrow$ f is convex

Strict Convexity

Recall that $f : \mathbb{R}^n \to \mathbb{R}$ is strictly convex if
$\forall x, y, x \neq y, \forall \lambda \in (0, 1)$,

$$f(\lambda x + (1 - \lambda)y) < \lambda f(x) + (1 - \lambda)f(y)$$

- f is strictly convex $\Rightarrow$ f is convex

- f is convex $\not\Rightarrow$ f is strictly convex

e.g., $\quad f(x) = x$

Strict convexity

- $\nabla^2 f(x) \succ 0, \forall x \in \Omega \Rightarrow f$ strictly convex on Ω

Strict convexity

- $\nabla^2 f(x) \succ 0, \forall x \in \Omega \Rightarrow f$ strictly convex on Ω

- Converse not true:

$$f(x) = x^4$$

f is strictly convex but $f''(0) = 0$

Characterize of strictly convex function

Theorem. Suppose $f : \mathbb{R}^n \to \mathbb{R}$ is differentiable over its domain. Then, the following are equivalent:

1. f is convex
2. $f(y) > f(x) + \nabla f(x)^T (y - x), \quad \forall x, y \in \text{dom}(f)$

Uniqueness of optimal solution

Theorem. Consider an optimization problem

$$\min f(x)$$
$$\text{s.t. } x \in \Omega$$

where $f : \mathbb{R}^n \to \mathbb{R}$ is strictly convex on Ω and Ω is a convex set. Then, the optimal solution is unique (assuming it exists).

Example. Is the function $f(x) = (x_1 - 3x_2)^2$ convex, strictly convex, or neither?

Optimality Condition for convex problem

Theorem. Consider an optimization problem

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

where $f : \mathbb{R}^n \to \mathbb{R}$ is convex and differentiable and Ω is a convex set. Then, a point x is optimal iff $x \in \Omega$ and

$$\nabla f(x)^T (y - x) \geq 0, \quad \forall y \in \Omega$$

Intuition (1)

$$x \text{ is optimal } \Leftrightarrow \nabla f(x)^T(y - x) \geq 0, \quad \forall y \in \Omega$$

Intuition (2)

$$x \text{ is optimal } \Leftrightarrow \nabla f(x)^T (y - x) \geq 0, \quad \forall y \in \Omega$$

Optimality Condition for convex problem

f is convex.

$$x \text{ is optimal } \Leftarrow \nabla f(x)^T (y - x) \geq 0, \quad \forall y \in \Omega$$

Proof.