Convex Functions
Optimization Methods (CS1.404), Spring 2024

Naresh Manwani

Machine Learning Lab, IIIT-H

February 5th, 2024

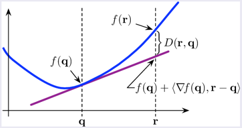
INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD

First Order Characterization of Convex Functions

The Gradient Inequality

Let $f : S \rightarrow \mathbb{R}$ be a continuously differentiable function defined on a convex set $S \subseteq \mathbb{R}^n$ Then, f is convex over S if and only if

$$f(\mathbf{y}) \geq f(\mathbf{x}) + \nabla f(\mathbf{x})^T(\mathbf{y} - \mathbf{x}), \ \forall \mathbf{x}, \mathbf{y} \in S$$



The Gradient Inequality for Strictly Convex Function

Let $f : S \rightarrow \mathbb{R}$ be a continuously differentiable function defined on a convex set $S \subseteq \mathbb{R}^n$ Then, f is strictly convex over S if and only if

$$f(\mathbf{y}) > f(\mathbf{x}) + \nabla f(\mathbf{x})^T(\mathbf{y} - \mathbf{x}), \ \forall \mathbf{x}, \mathbf{y} \in S \ \ (\mathbf{x} \neq \mathbf{y})$$

Sufficiency of Stationarity Under Convexity

Proposition

Let f be a continuously differentiable function which is convex over a convex set $S \subseteq \mathbb{R}^n$ Suppose that $\nabla f(\mathbf{x}^*) = \mathbf{0}$ for some $\mathbf{x}^* \in S$, then $\mathbf{x}^*$ is global minimizer of f over S.

Monotonicity of the Gradient of Convex Functions

Theorem

Suppose that f is a continuously differentiable function over a convex set $S \subseteq \mathbb{R}^n$. Then, f is convex over S if and only if

$$(\nabla f(\mathbf{x}) - \nabla f(\mathbf{y}))^T(\mathbf{x} - \mathbf{y}) \geq 0, \; \forall \mathbf{x}, \mathbf{y} \in S$$

Second Order Characterization of Convex Function

Theorem

Let f be a twice continuously differentiable function over an open convex set $S \subseteq \mathbb{R}^n$. Then, f is convex over S if and only if $\nabla^2 f(\mathbf{x})$ is positive semi-definite for any $\mathbf{x} \in S$.

Sufficient Second Order Characterization for Strict Convexity

Let f be a twice continuously differentiable function over a convex set $S \subseteq \mathbb{R}^n$, and suppose that $\nabla^2 f(\mathbf{x})$ is strictly positive definite for all $\mathbf{x} \in S$. Then, f is strictly convex over S.

Examples of Convex Functions:

1. log-sum-exponential function: $f(\mathbf{x}) = \ln(e^{x_1} + e^{x_2} + \ldots + e^{x_n})$
2. quadratic over linear: $f(x_1, x_2) = \frac{x_1^2}{x_2}$
3. $f(x) = x \log x$ where f is defined over $S = \{x \in \mathbb{R} \mid x > 0\}$
4. $f(\mathbf{x}) = \|A\mathbf{x} - \mathbf{b}\|^2$

Jensen's Inequality

Theorem

Let $f : S \to \mathbb{R}$ be a convex function defined on the convex set $S \subseteq \mathbb{R}^n$. Then for any $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k \in S$ and $\boldsymbol{\lambda} \in \Delta_k$, the following inequality holds:

$$f(\sum_{j=1}^{k} \lambda_k \mathbf{x}_k) \leq \sum_{j=1}^{k} \lambda_k f(\mathbf{x}_k)$$

where Δ_k is k-dimensional probability simplex.

INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY HYDERABAD