Optimization Methods (CS1.404), Spring 2024
Lecture 17

Naresh Manwani

Machine Learning Lab, IIIT-H

March 14th, 2024

Constrained Optimization Problem

$$\begin{aligned} \min_{\mathbf{x}} \quad & f(\mathbf{x}) \\ s.t. \quad & h_j(\mathbf{x}) \leq 0, \; j = 1, \ldots, l \\ & e_i(\mathbf{x}) = 0, \; i = 1, \ldots, m \end{aligned}$$

where

- $h_j : \mathbb{R}^n \to \mathbb{R}, \; j = 1, \ldots, l$
- $e_i : \mathbb{R}^n \to \mathbb{R}, \; i = 1, \ldots, m$
- Assume that all h_j and e_i are sufficiently smooth functions.
- **Feasible set:** Any point that satisfies constraints is called feasible point. Set of all feasible points is called feasible set and is described as $\mathcal{X} = \{\mathbf{x} \in \mathbb{R}^n \mid h_j(\mathbf{x}) \leq 0, \; e_i(\mathbf{x}) = 0, \; j = 1 \ldots l, \; i = 1 \ldots m\}$.

Local and Global Minima

Definition: Global Minima

A point $\mathbf{x}^* \in \mathcal{X}$ is said to be global minimum point of f over $\mathcal{X}$ if $f(\mathbf{x}) \geq f(\mathbf{x}^*)$, $\forall \mathbf{x} \in \mathcal{X}$. If $f(\mathbf{x}) > f(\mathbf{x}^*)$, $\forall \mathbf{x} \in \mathcal{X}$, $\mathbf{x} \neq \mathbf{x}^*$, then $\mathbf{x}^*$ is called strict global minima.

Definition: Local Minima

A point $\mathbf{x}^* \in \mathcal{X}$ is said to be local minimum point of f over $\mathcal{X}$ if there exists $\epsilon > 0$ such that $f(\mathbf{x}) \geq f(\mathbf{x}^*)$, $\forall \mathbf{x} \in \mathcal{X} \cap B(\mathbf{x}^*, \epsilon)$. If $f(\mathbf{x}) > f(\mathbf{x}^*)$, $\forall \mathbf{x} \in \mathcal{X} \cap B(\mathbf{x}^*, \epsilon)$, $\mathbf{x} \neq \mathbf{x}^*$, then $\mathbf{x}^*$ is called strict local minima.

Constrained Convex Optimization Problem

$$\begin{aligned} \min_{\mathbf{x}} \quad & f(\mathbf{x}) \\ s.t. \quad & h_j(\mathbf{x}) \leq 0, \ j = 1, \ldots, l \\ & e_i(\mathbf{x}) = 0, \ i = 1, \ldots, m \end{aligned}$$

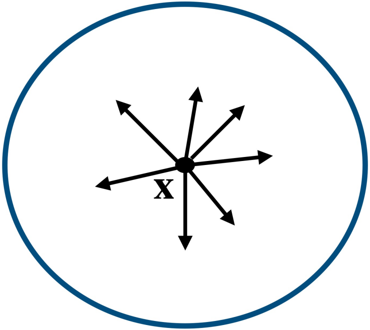
where

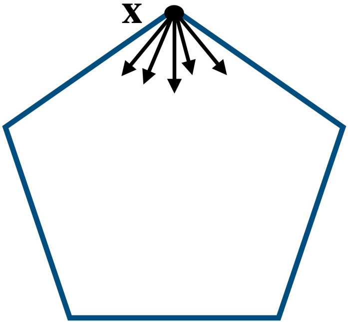
- $f(\mathbf{x})$ is convex.
- $h_j : \mathbb{R}^n \to \mathbb{R}, \ j = 1, \ldots, l$ are convex functions.
- $e_i : \mathbb{R}^n \to \mathbb{R}, \ i = 1, \ldots, m$ are affine functions.
- Any local minima is a global minima.

Feasible Direction

Definition

A vector $\mathbf{d} \in \mathbb{R}^n$, $\mathbf{d} \neq \mathbf{0}$ is said to be a feasible direction at $\mathbf{x} \in \mathcal{X}$ if there exist $\delta_1 > 0$ such that $\mathbf{x} + \alpha\mathbf{d} \in \mathcal{X}$, $\forall \alpha \in (0, \delta_1)$.





Let $\mathcal{F}(\mathbf{x})$ represents the set of feasible directions at $\mathbf{x} \in \mathcal{X}$.

Descent Direction

Definition

A vector $\mathbf{d} \in \mathbb{R}^n$, $\mathbf{d} \neq \mathbf{0}$ is said to be a descent direction at $\mathbf{x} \in \mathcal{X}$ if there exists $\delta_2 > 0$ such that $f(\mathbf{x} + \alpha\mathbf{d}) < f(\mathbf{x})$, $\forall \alpha \in (0, \delta_2)$.

Let $\mathcal{D}(\mathbf{x})$ represents the set of descent directions at $\mathbf{x} \in \mathcal{X}$.

Characterization of Local Minima in Constrained Optimization

Theorem

Let $\mathcal{X}$ be a nonempty set in $\mathbb{R}^n$ and $\mathbf{x}^* \in \mathcal{X}$ be a local minimum of f over $\mathcal{X}$. Then $\mathcal{F}(\mathbf{x}^*) \cap \mathcal{D}(\mathbf{x}^*) = \phi$.

- $\mathbf{x}^* \in \mathcal{X}$ is a local minima if $\Rightarrow \mathcal{F}(\mathbf{x}^*) \cap \mathcal{D}(\mathbf{x}^*) = \phi$.
- Consider any $\mathbf{x} \in \mathcal{X}$ and assume $f \in \mathbb{C}^2$. Then $\nabla f(\mathbf{x})^T \mathbf{d} < 0$ implies there exists $\delta > 0$ such that $f(\mathbf{x} + \alpha \mathbf{d}) < f(\mathbf{x})$, $\forall \alpha \in (0, \delta)$. Such $\mathbf{d}$ is a descent direction ($\mathbf{d} \in \mathcal{D}(\mathbf{x})$).
- Let $\tilde{\mathcal{D}}(\mathbf{x}) = \{\mathbf{d} \in \mathbb{R}^n \mid \nabla f(\mathbf{x})^T \mathbf{d} < 0\}$, then $\tilde{\mathcal{D}}(\mathbf{x}) \subseteq \mathcal{D}(\mathbf{x})$.