Applied Linear Algebra in Data Analysis

Introduction to Optimization

Sivakumar Balasubramanian

Department of Bioengineering
Christian Medical College, Bagayam
Vellore 632002

Optimization

- Optimization is the process of finding the best solution to a problem from a set of possible solutions.
- Optimization problems come up in many applications in engineering, science, economics, biology, medicine, operations research, etc.
- Optimization problems can be classified in different ways, but one major classification gives us: **unconstrained** and **constrained** optimization problems.

A general optimization problem

- A general optimization problem can be fomulated as the following,

$$\begin{aligned} \min_{\mathbf{x}\in\mathcal{X}} \;\; & f\left(\mathbf{x}\right) \\ \text{subject to} \;\; & \mathbf{g}\left(\mathbf{x}\right) \leq \mathbf{0},\;\; \mathbf{g}\left(\mathbf{x}\right) = \begin{bmatrix} g_1\left(\mathbf{x}\right) & g_2\left(\mathbf{x}\right) & \cdots & g_p\left(\mathbf{x}\right)\end{bmatrix}^\top \\ & \mathbf{h}\left(\mathbf{x}\right) = \mathbf{0},\;\; \mathbf{h}\left(\mathbf{x}\right) = \begin{bmatrix} h_1\left(\mathbf{x}\right) & h_2\left(\mathbf{x}\right) & \cdots & h_q\left(\mathbf{x}\right)\end{bmatrix}^\top \end{aligned}$$

  where, $f\left(\mathbf{x}\right)$ is the **objective function** and $\mathbf{g}\left(\mathbf{x}\right)$ represents the set of **inquality constaints** and $\mathbf{h}\left(\mathbf{x}\right)$ represents the set of **equality constraints**.

- In this course, we will only focus on optimization problems over $\mathbb{R}^n$, and mostly problems where the objective function and the constraints are differentiable.

A general optimization problem

- Most optimization problems of practical significance cannot be solved analytically, and we must resort to numerical iterative methods to find a solution.

- We can never solve these problems exactly through numerical means, and must content outselves with finding an approximate "good enough" solution.

Mathematical preliminaries: Sequences and Limits

We first review the notions of continuity and differentiability of functions of single and multiple variables, since we will be dealing with differentiable functions in optimization problems.

Sequences and Limits:

- A sequence of real numbers is a function whose domain is a set of natural numbers $1, 2, \ldots, k, \ldots$ and whose range is a set of real numbers. The sequence is denoted by $\{x_k\}_{k=1}^{\infty}$ or $\{x_k\}$.

- A number x^* is said to be the **limit** of the sequence $\{x_k\}$ if for every $\epsilon > 0$, there exists an integer K such that for all $k > K$, we have $|x_k - x^*| < \epsilon$.

$$\lim_{k \to \infty} x_k = x^* \quad \text{or} \quad x_k \to x^*$$

  A sequence that has a limit is called a **convergent sequence**.

Sequences and Limits

We can extend these ideas to $\mathbb{R}^n$.

- A sequence in $\mathbb{R}^n$ is a function whose domain is a set of natural numbers $1, 2, \ldots, k, \ldots$ and whose range is $\mathbb{R}^n$. The sequence is denoted by $\{\mathbf{x}_k\}_{k=1}^{\infty}$ or $\{\mathbf{x}_k\}$.
- $\mathbf{x}^*$ is said to be the **limit** of the sequence $\{\mathbf{x}_k\}$ if for every $\epsilon > 0$, there exists an integer K such that for all $k > K$, we have $\|\mathbf{x}_k - \mathbf{x}^*\| < \epsilon$.

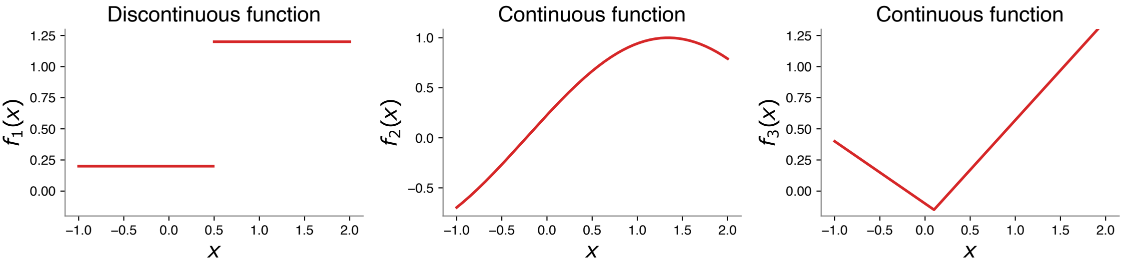
$$\lim_{k\to\infty} \mathbf{x}_k = \mathbf{x}^* \quad \text{or} \quad \mathbf{x}_k \to \mathbf{x}^*$$

- The limit of a convergent sequence is unique.

Continuity

Consider the function $f : \Omega \to \mathbb{R}$, where $\Omega \subseteq \mathbb{R}^n$. This function is continuous at the point $\mathbf{x}_0 \in \Omega$, if and only if,

$$\lim_{\mathbf{x} \to \mathbf{x}_0} f\left(\mathbf{x}\right) = f\left(\mathbf{x}_0\right)$$



Some definitions

Consider the problem of finding the minimum of single variable function $f(x)$.

$$x^* = \arg\min_{x \in \mathbb{R}} f(x)$$

- *Necessary Condition:*

Steepest descent algorithm

- Consider the experiment tossing a dice, and we observe the count of the dots that turn on the top face of the dice.
  - Observed outcome is an even number. $A = \{2, 4, 6\} \subset S$
  - Observed outcome is a positive number. $A = S \implies$ **Sure event**
  - Observed outcome is 0. $A = \{\} \implies$ **Impossible event**
- For discrete sample spaces and **elementary event** is an event with just single sample point.
- We can combine events to produce other events that might be of interest to us. Set operations can be used to perform algebra on events.