香港中文大學（深圳）
The Chinese University of Hong Kong, Shenzhen

Introduction to Data Science

Lecture 15 Optimization: Convex Set

Zicheng Wang

Review: Interval Estimation

Interval Estimation

- A random variable: X with variance $\boldsymbol{\sigma^2}$
- Data: $\mathbf{X_1, X_2, \cdots, X_n}$
- Target: estimate the mean (μ) of the random variable
- Point estimate: $\bar{X} = \frac{X_1 + \cdots + X_n}{n}$
- What's the probability that μ lies in $T = [\bar{X} - \frac{b\sigma}{\sqrt{n}}, \bar{X} + \frac{a\sigma}{\sqrt{n}}]$, where $a, b > 0$ are constants?

Interval Estimation

- Compute $P(\bar{X} - \frac{b\sigma}{\sqrt{n}} \leq \mu \leq \bar{X} + \frac{a\sigma}{\sqrt{n}})$

- $P\left(\bar{X} - \frac{b\sigma}{\sqrt{n}} \leq \mu \leq \bar{X} + \frac{a\sigma}{\sqrt{n}}\right) = P(-a \leq \frac{\sqrt{n}(\bar{X}-\mu)}{\boldsymbol{\sigma}} \leq b)$

- By the Central Limit Theorem, $\frac{\sqrt{\mathrm{n}}\,(\bar{\mathrm{X}}-\mu)}{\boldsymbol{\sigma}} \sim N(0,1)$

- P($-a \leq \frac{\sqrt{n}(\bar{X}-\mu)}{\boldsymbol{\sigma}} \leq b$) = $\Phi(b) - \Phi(-a)$, where $\Phi(x)$ is the CDF of a standard normal distribution

Formally, we say $[\bar{X} - \frac{b\sigma}{\sqrt{n}}, \bar{X} + \frac{a\sigma}{\sqrt{n}}]$ is a $\Phi(b) - \Phi(-a)$

Confidence Interval

Optimization

Optimization

Objective Function

Constraints

Given: a function f: $A \to \mathbb{R}$ from some set A to the real numbers

Sought: an element $\mathbf{x}_0 \in A$ such that $f(\mathbf{x}_0) \leq f(\mathbf{x})$ for all $\mathbf{x} \in A$ ("minimization") or such that $f(\mathbf{x}_0) \geq f(\mathbf{x})$ for all $\mathbf{x} \in A$ ("maximization").

Decision Variables

Optimization problem in standard form

Objective Function

$$\begin{array}{lll} \text{minimize} & \boxed{f_0(x)} & \\ \text{subject to} & f_i(x) \leq 0, & i = 1, \ldots, m \\ & h_i(x) = 0, & i = 1, \ldots, p \end{array}$$

Decision Variables

Constraints

- $\boxed{x} \in \mathbf{R}^n$ is the optimization variable
- $f_0 : \mathbf{R}^n \to \mathbf{R}$ is the objective or cost function
- $f_i : \mathbf{R}^n \to \mathbf{R}$, $i = 1, \ldots, m$, are the inequality constraint functions
- $h_i : \mathbf{R}^n \to \mathbf{R}$ are the equality constraint functions

Motivating Example

- Suppose you want to start your own blind box business.

- Let D denote the one season (three months) random demand, with CDF $F(\cdot)$, and mean $\mu = E[D]$.

- At the beginning of each season, you place an order Q to Pop Mart, with a cost c for each blind box.

- Each blind box can be sold at a price of $p > c$.

- At the end of each season, unsold blind boxes are salvaged, and you get $s < c$ for each salvaged box.

Motivating Example

- For simplicity, let's assume that D is a continuous random variable and you can also place a continuous order Q.

- You want to choose the optimal order quantity Q so as to maximize your expected profit.

- How should you formulate the problem?

Problem Formulation

- First we observe that if the realized demand $D > Q$, then your profit is $(p-c)Q$. Otherwise, your profit is $(p-c)D + (s-c)(Q-D)$.

- Let's define $(Q-D)^+ = \max(Q-D, 0)$.

- Given D, your profit is $p \cdot \min(Q, D) + s(Q-D)^+ - cQ$.

- The objective function (i.e., the expected profit) is then given by $f(Q) = pE[\min(Q, D)] + sE[(Q-D)^+] - cQ$

Problem Formulation

- For this problem, we have one inequality constraint: $Q \geq 0$.

- Hence, the optimization problem is as follows

$$
\begin{aligned}
&\text{maximize} && pE[\min(Q, D)] + sE[(Q - D)^+] - cQ \\
&\text{subject to} && Q \geq 0
\end{aligned}
$$

- In standard form, we have

$$
\begin{aligned}
&\text{minimize} && -(pE[\min(Q, D)] + sE[(Q - D)^+] - cQ) \\
&\text{subject to} && -Q \leq 0
\end{aligned}
$$

Breakout Room Question

- Suppose you want to start your own blind box business.
- Let D denote the one season (three months) random demand, which follows a uniform distribution in $[10,100]$.
- At the beginning of each season, you place an order Q to Pop Mart, with a cost 10 Yuan for each blind box.
- Each blind box can be sold at a price of 20 Yuan.
- At the end of each season, unsold blind boxes are salvaged, and you get 3 Yuan for each salvaged box.

- How many blind boxes should you order to maximize your expected profit?

- *Convex Optimization*，*Stephen Boyd and Lieven Vandenberghe*
- https://web.stanford.edu/~boyd/cvxbook/
- Convex sets: (No need to read the part about matrix analysis)
  - 2.1.1, 2.1.4, 2.1.5
  - 2.2.1, 2.2.2, 2.2.4
  - 2.3.1
- Exercises: (whenever multiple-dimension, prove it in **two-dimension**)
  - 2.1, 2.2, 2.3, 2.4, 2.7, 2.11, 2.12, 2.15, 2.16, 2.17
  - The solution is available online.

Terminologies

$$
\begin{aligned}
\text{minimize} \quad & f(\boldsymbol{x}) \\
\text{subject to} \quad & g_i(\boldsymbol{x}) \leq 0, i = 1, \dots, m \\
& h_i(\boldsymbol{x}) = 0, i = 1, \dots, n
\end{aligned}
$$

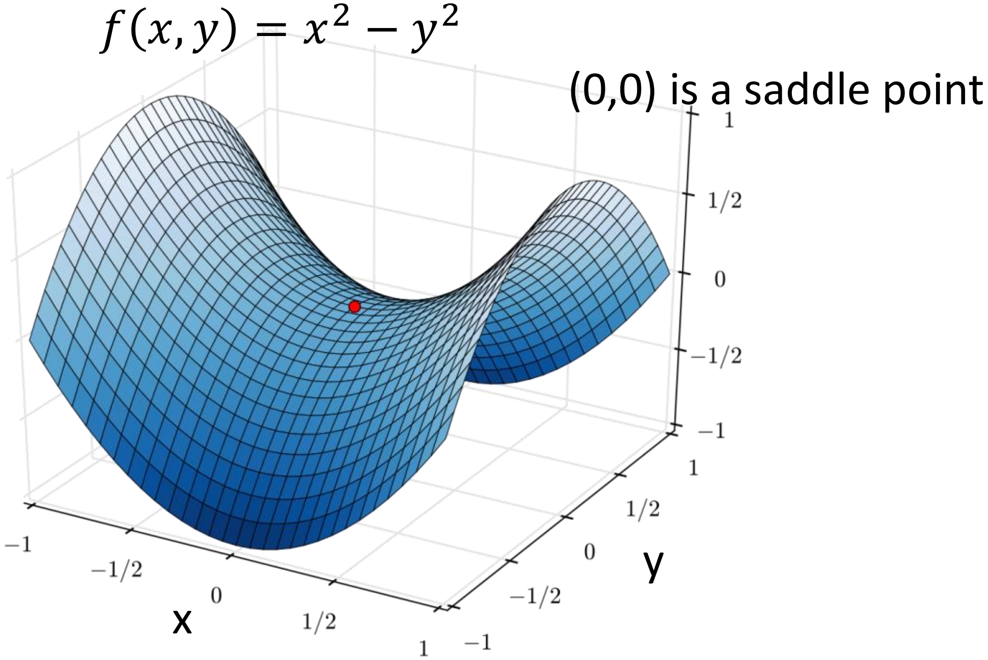
- **Feasible Set**: the set of all points such that the constraints can **all** be satisfied.

- $\boldsymbol{x}^*$ is the **global minimizer**, if $f(\boldsymbol{x}^*) \leq f(\boldsymbol{y})$ for any $\boldsymbol{y}$ in the feasible set.

- $\boldsymbol{x}^*$ is called a **saddle point**, if $\frac{df(\boldsymbol{x}^*+t\boldsymbol{e})}{dt}|_{t=0}= 0$ for any $\boldsymbol{e}$. (First-order derivative in any direction is 0 at point $\boldsymbol{x}^*$), but the function attains neither a local maximum value nor a local minimum value

Remark: In this lecture, I will use **bold** form to represent a high dimension point. Without bold form, it represents a scalar

Graphical illustration of a saddle point

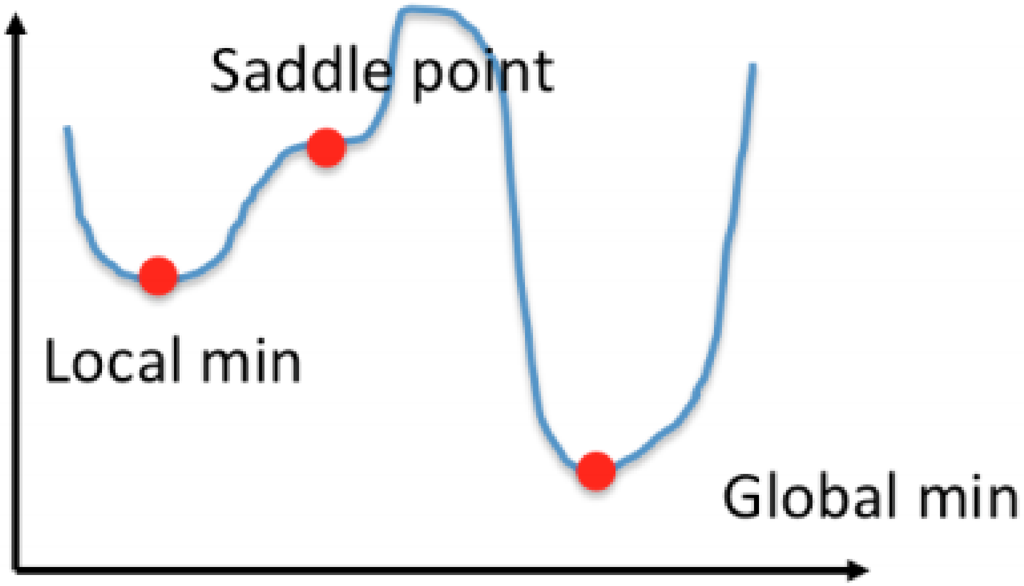


Terminologies

$$\begin{aligned} \text{minimize} \quad & f(\boldsymbol{x}) \\ \text{subject to} \quad & g_i(\boldsymbol{x}) \leq 0, i = 1, \dots, m \\ & h_i(\boldsymbol{x}) = 0, i = 1, \dots, n \end{aligned}$$

- **Local minimizer**
  - Denote S as the feasible set
  - Denote B(**x**,ε) = {**y**: $\|\boldsymbol{y} - \boldsymbol{x}\| \leq \varepsilon$} as the set of all points such that the distance from x and each point in the set is smaller than ε.
  - If there exists an $\varepsilon > 0$ such that for any $\boldsymbol{y} \in S \cap$ B($\boldsymbol{x}^*$,ε), f($\boldsymbol{x}^*$)$\leq f(\boldsymbol{y})$. Then $\boldsymbol{x}^*$ is called a local minimizer of the optimization problem.

Graphical illustration



Convex Optimization

- Verifying if a point is a local minimum can be easy.
- However, how can we ensure that a local minimum is indeed a global minimum?
- A category of problems known as convex optimization has property that any local minimum is also a global minimum.

This will be proved in subsequent lectures.

Convex vs. Non-Convex

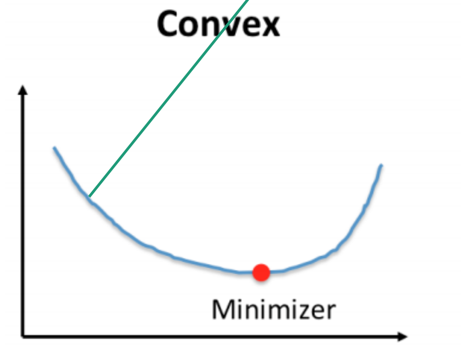
This is not a local minimizer, as the function value is smaller when the variable is just slightly higher.

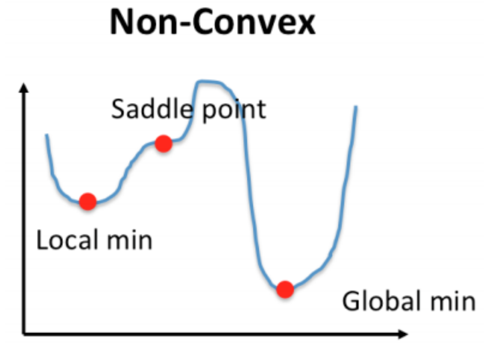
Convex problem

- Local minimizer is the global minimizer.
- Can be solved efficiently
- Gradient descent (or many other acceleration methods) converges to the global solution

Non-Convex problem

- It's not easy to find the global solution
- Use heuristics to find local optimal solutions





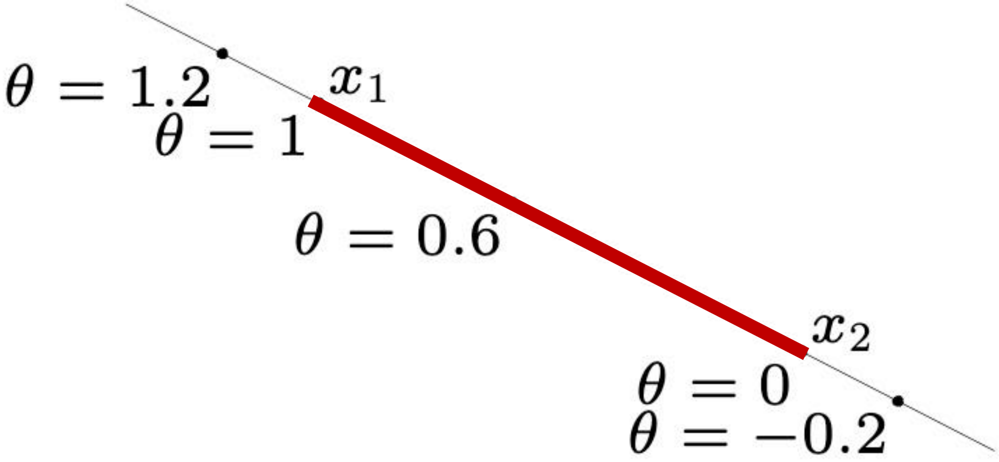
Introduction to Convex Optimization

Line Segment

- Let $\boldsymbol{x_1} \neq \boldsymbol{x_2}$ be two points in $\mathbb{R}^n$. Points of the form

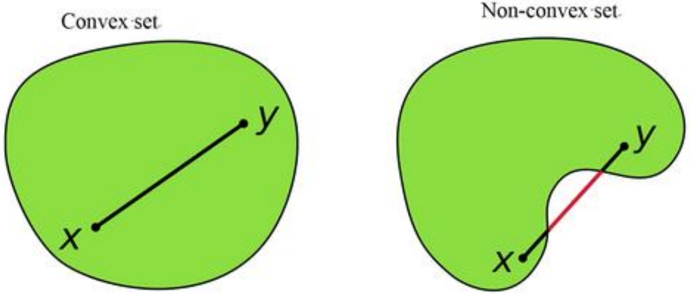
$$\boldsymbol{x} = \theta \boldsymbol{x}_1 + (1 - \theta)\boldsymbol{x}_2$$

where $\boldsymbol{\theta} \in [\mathbf{0}, \mathbf{1}]$, form the **line segment** between $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$.



Convex Set

- Set C is a **convex set** if the line segment between any two points in C lies in C.

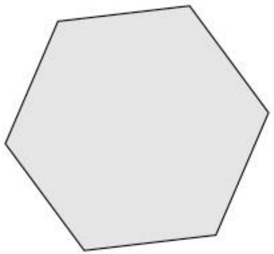


- Formal definition: A set C is convex if $\forall\ \boldsymbol{x_1}, \boldsymbol{x_2} \in C, \forall \theta \in [0,1]$
$$\theta \boldsymbol{x_1} + (1-\theta)\boldsymbol{x_2} \in C.$$

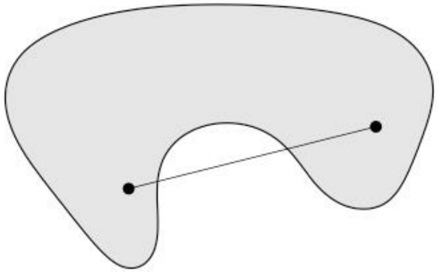
Remark: In this lecture, I will use **bold** form to represent a high dimension point. Without bold form, it represents a scalar

Examples

- Which set is a convex set?



A



B

C

Convex Set Examples

- The empty set $\emptyset$, the singleton set $\{\boldsymbol{x_0}\}$, and the complete space R are convex sets.
- An interval of $[a, b] \subset R$ is a convex set
- In R^n the set $\mathrm{H} \coloneqq \{\boldsymbol{x} \in R^n : a_1 x_1 + \cdots + a_n x_n = c\}$ is a convex set
- Half spaces, e.g., $\mathrm{H} \coloneqq \{(x, y) : y \leq ax + b\}$ are convex sets
- A disk with center $(0,0)$ and radius c is a convex subset of R^2

Remark: In this lecture, I will use **bold** form to represent a high dimension point. Without bold form, it represents a scalar

Steps for Showing the Convexity of a Set

Prove $\mathrm{H} := \{(x, y) : y = ax + b\}$ is a convex set

For any (x_1, y_1) and (x_2, y_2) in H,

- $y_1 = ax_1 + b$
- $y_2 = ax_2 + b$
- $\theta(x_1, y_1) + (1 - \theta)(x_2, y_2) = (\theta x_1 + (1 - \theta)x_2), \theta y_1 + (1 - \theta)y_2)$
- Then for any $\theta \in [0,1]$
  - $\theta y_1 + (1 - \theta)y_2 = a(\theta x_1 + (1 - \theta)x_2) + b$

Steps for Showing the Convexity of a Set

Prove $\mathrm{H} := \{(x, y): y = ax + b\}$ is a convex set

For any (x_1, y_1) and (x_2, y_2) in H,

- $y_1 = ax_1 + b$
- $y_2 = ax_2 + b$

1. Use the assumption that $(x_1, y_1), (x_2, y_2) \in H$

- $\theta(x_1, y_1) + (1-\theta)(x_2, y_2) = \underbrace{(\theta x_1 + (1-\theta)x_2), \theta y_1 + (1-\theta)y_2)}$

2. Characterize the new point within the line segment

- Then for any $\theta \in [0,1]$
  - $\underbrace{\theta y_1 + (1-\theta)y_2 = a(\theta x_1 + (1-\theta)x_2) + b}$

3. Use (1) and (2) to show that the new point is in H

Prove Half spaces, e.g., $\mathrm{H} \coloneqq \{(x, y): y \leq ax + b\}$ are convex sets

Proof:

For any (x_1, y_1) and (x_2, y_2) in H,

- $y_1 \leq ax_1 + b$
- $y_2 \leq ax_2 + b$
- $\theta(x_1, y_1) + (1-\theta)(x_2, y_2) = (\theta x_1 + (1-\theta)x_2), \theta y_1 + (1-\theta)y_2)$
- Then for any $\theta \in [0,1]$
  - $\theta y_1 \leq a\theta x_1 + \theta b$
  - $(1-\theta)y_2 \leq a(1-\theta)x_2 + (1-\theta)b$
  - $\theta y_1 + (1-\theta)y_2 \leq a(\theta x_1 + (1-\theta)x_2) + b$

Properties of convex sets.

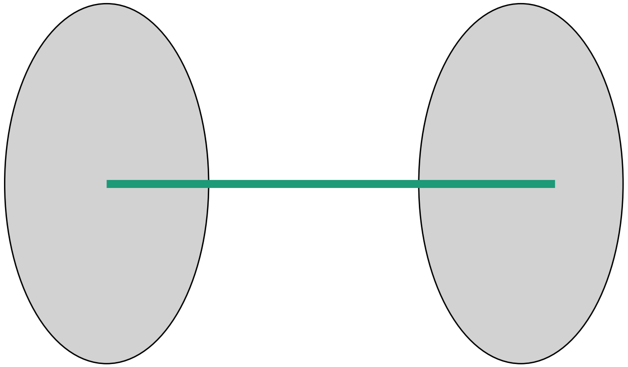
Lemma: If both $\boldsymbol{S_1}$ and $\boldsymbol{S_2}$ are convex sets, then $\boldsymbol{S_1 \cap S_2}$ is also a convex set.

Proof

- Given any two points $\mathbf{x_1}$ and $\mathbf{x_2}$ in $S_1 \cap S_2$,
- Let $\mathbf{x}$ be a point on the line segment between $\boldsymbol{x_1}$ and $\boldsymbol{x_2}$.
- As S_1 is convex set, $\mathbf{x}$ is within S_1
- As S_2 is convex set, $\mathbf{x}$ is within S_2
- Thus, $\mathbf{x}$ is within $S_1 \cap S_2$

Question:

Is the union of two convex sets a convex set?



Blind Box Problem

- In standard form, we have

$$\begin{aligned} &\text{minimize} && \overbrace{-(pE[\min(Q,D)] + sE[(Q-D)^+] - cQ)}^{\text{Next lecture: covex function}} \\ &\text{subject to} && \underbrace{-Q \leq 0}_{\text{It is a half space, and hence a convex set}} \end{aligned}$$