Lecture #3 (Optimization)

- We tend to optimize our decision making in day-to-day lives. We plan our activities for each day at the beginning of the day, execute decisions, account for uncertainties in our planning and re-optimize

(Real-time optimization)

Concerns with minimizing or maximizing certain objective subject to some constraints.

- Knapsack problem (One of the most basic optimization problems)

Knapsack
(Total capacity : C)

n -items.
Value of i^{th}-item: v_i
Capacity of i^{th}-item: C_i

Goal: Maximize total value of the knapsack, i.e. pick and chose items that maximize the total value of the knapsack, subject to capacity constraints.

Q: How do we mathematically formulate this objective?

A: We introduce optimization variables $\{x_i\}_{i=1}^{n}$

Each optimization variable is binary, i.e., $x_i \in \{0, 1\}$

$x_i = 0 \Rightarrow$ Product is not selected
$x_i = 1 \Rightarrow$ Item is selected.

Objective function: $\sum_{i=1}^{n} x_i v_i$

Optimization Problem:

$$\max_{\{x_i\}} \sum_{i=1}^{n} x_i v_i$$

$$\text{s.t.} \quad \sum_{i=1}^{n} x_i c_i \leq C \quad \text{[Capacity constraints]}$$

$$x_i \in \{0,1\} \quad \text{for all } i \in \{1,2,\dots,n\}$$

→ Optimization variables or Decision Variables are binary (integeral) in nature

↳ Example of Integer Linear Program (ILP)

- Integer: because decision variables are integers
- Linear: Because objective function and constraints are linear in decision variables.

↔ Parallels with Portfolio Maximization:

↳ You have a limited budget (\$ C)

↳ You have 'n' assets to choose from.

↳ Each asset 'i' is characterized by the cost of the asset c_i and expected return 'v_i'

↳ Objective is to maximize total return, subject to budget constraints.

Ex: Power System Economics

You have a network of 'n' generators. The cost associated with i^{th}-generator to produce a total power P_i is modeled using a quadratic function:

$$C_i(P_i) = \underbrace{a_i P_i^2 + b_i P_i + c_i}_{\text{cost of producing } P_i \text{ power by } i^{th} \text{ generator.}}$$

Total load demand is P_{tot}.

Objective: Minimize the total cost of generation

Additional constraints: Each generator can not produce power beyond a certain limit and obviously cannot produce negative power.

Mathematical Formulation:

$$\min_{\{P_i\}} \sum_{i=1}^{n} a_i P_i^2 + b_i P_i + c_i$$

$$\text{subject to} \quad \sum_{i=1}^{n} P_i = P_{tot}$$

$$0 \leq P_i \leq P_{i,max} \quad \text{for all } i \in \{1, \dots, n\}$$

Here, $\{P_i\}$ decision variables are continuous and not integers.

Ex: Least-Squares Regression:

Suppose price of a house is modeled using

$$\hat{y} = a_0 + a_1 x_1 + a_2 x_2 + \dots + a_n x_n$$

Here, $(a_0, a_1, \dots, a_n)$ are coefficients to be determined

$\rightarrow$ Optimization variables

$(x_1, \dots, x_n)$ are various features, like, the number of rooms, area, locality, etc...

How do we find $(a_0, a_1, \dots, a_n)$?

Suppose we are given prices and specification of 'm' houses (they form our training sets). We want to ensure that our model predicts price $\hat{y}_j$ for the j^{th}-house on the training set, that is very close to true price $(y_{true,j})$ of the j^{th}-house.

Objective f^n: $$\sum_{j=1}^{m} \left(\hat{y}_j - y_{true,j}\right)^2$$

Optimization Problem:

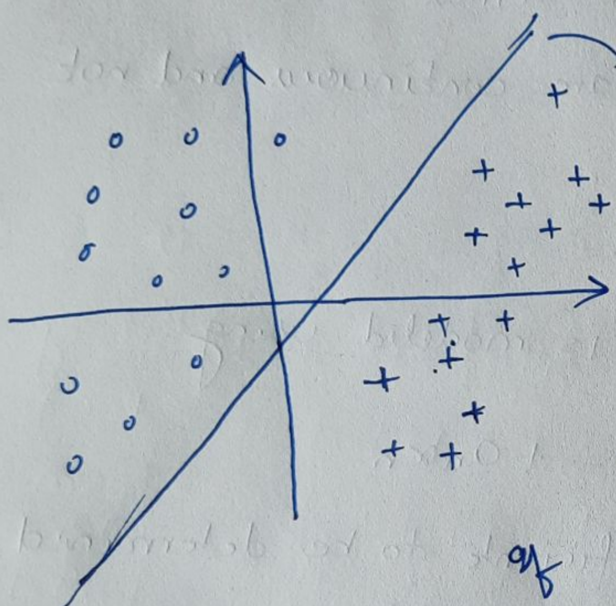
$$\min_{\{a_0, a_1, \ldots, a_n\}} \sum_{j=1}^{m} \left(\hat{y}_j - y_{true,j}\right)^2$$

s.t. $\hat{y}_j = a_0 + a_1 x_1^{(j)} + a_2 x_2^{(j)} + \ldots + a_n x_n^{(j)}$

for all $j \in \{1, 2, \ldots, m\}$

Ex: Classification Problem (Support Vector Machines)

SVM



Find a separating hyperplane $w^T x + b = 0$ such that the plane (line in 2D) separates the two classes with maximum margin.

For a point x_i with label y_i → $+1$ or -1

if $w^T x_i + b > 0 \Rightarrow$ Predicted label is $+1$

$w^T x_i + b < 0 \Rightarrow$ Predicted label is -1.

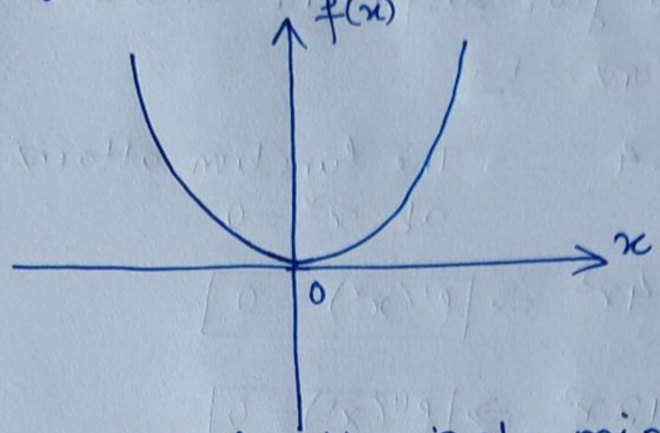
Optimization Problem:

$$\min_{\{w, b\}} \frac{1}{2} \|w\|^2$$

s.t. $y_i(w^T x_i + b) \geq 1$ for all $i = 1, 2, \ldots, m$.

Continuous Optimization: Convexity

↳ Consider the most simple looking function
$f(x) = x^2$



We know that the function gets minimized at $x^* = 0$ with a minimum value of $f(x^*) = 0$.

What else happens at $x^* = 0$?

$f(x) = x^2$

Derivative of $f(x)$, $f'(x) = 2x \Rightarrow \boxed{f'(x^*) = 0}$

Second-derivative of $f(x)$, $\boxed{f''(x) = 2 > 0} \Rightarrow \boxed{f''(x^*) > 0}$

Points at which the derivatives vanish are known as <u>stationary points</u>.

What about the function $f(x) = -x^2$? The function gets maximized at $x^* = 0$ with a maximum value of $f(x^*) = 0$.

$f(x) = -x^2$

$f'(x) = -2x \Rightarrow \boxed{f'(x^*) = 0}$

$f''(x) = -2 < 0 \Rightarrow \boxed{f''(x^*) < 0}$

This exercise suggests that if $\boxed{f'(x^*) = 0 \text{ and } f''(x^*) > 0}$, then x^* is a <u>minimizer</u>.

Likewise, if $\boxed{f'(x^*) = 0 \text{ and } f''(x^*) < 0}$, then x^* is a maximizer.

+ However, these are just sufficient conditions and not necessary conditions. Eg:

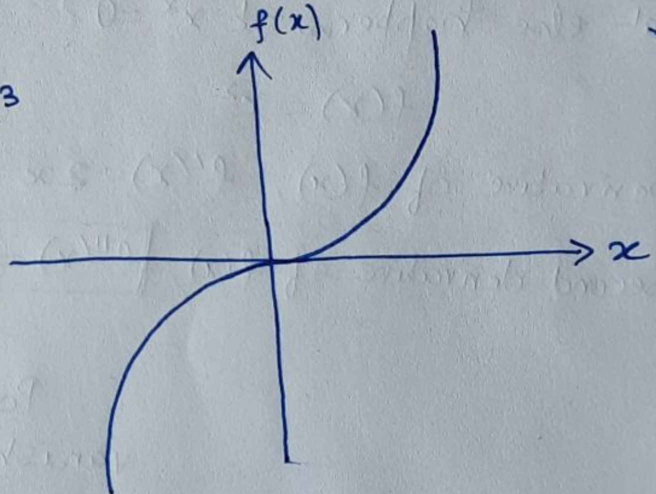
Consider, $f(x) = x^4$ → The function attains a minimum at $x^* = 0$

$\Rightarrow f'(x) = 4x^3 \Rightarrow \boxed{f'(x^*) = 0}$

$f''(x) = 12x^2 \Rightarrow \boxed{f''(x^*) = 0}$

Even though $f''(x^*)$ is not positive, x^* is still a minimizer.

Another example: $f(x) = x^3$



At $x=0$, function attains neither a minimum or a maximum, but

$f'(x) = 3x^2$

$\Rightarrow f'(0) = 0 \Rightarrow x^* = 0$ is a stationary point.

Such stationary points at which function does not attain either maxima or minima are called points of inflexion.