Lecture 01 : Basics of Optimization

	variable	Solution space	Solution complexity
Continuous optimization	take any real value / continuous	infinite	polynomial in the dim of problem
Discrete Optimization	discrete	finite	(mostly) exponential

Trend in AI : formulate the problems into optimization problem

What is optimization?

$$\min_{x} \; f(x) \quad \text{s.t.} \quad x \in \varphi$$

$f(x)$ ← objective function

$x \in \varphi$ ← constraint set

Ex 1

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

$$\varphi = [0,1] \cup [4,7]$$

$$\left.\begin{array}{l} \min f(x) \\ \text{s.t. } x \in \varphi \end{array}\right\} \quad x^* = 1$$

Ex. 2 Location of cities

$$\min_{x} \sum_{i=1}^{m} \|x - y_i\|_2 \quad \text{s.t.} \quad x \in \varphi$$

Ex 2. Image denoising :

$$\min_{X} \sum_{i,j} \left| y_{ij} - (\underbrace{K * X}_{\text{convolution}})_{ij} \right|^2 + \lambda \sum \left(X_{ij} - X_{i,j+1} \right)^2 + \left(X_{i+1,j} - X_{i,j} \right)^2$$

$$\text{s.t.} \quad X \in [0,1]^{m \times n}$$

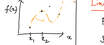
$Y \in \mathbb{R}^{m \times n}$, $\|Y\|_F = \sqrt{y_1^2 + y_2^2 + \cdots}$ ← L_2 norm

Ex 3 Machine learning

$(x_i, y_i) \to$ inputs $(i = 1, \dots, m)$

$$\min_{\theta} \sum_{i=1}^{m} \ell\left(h_\theta(x_i), y_i\right)$$

θ ← parameter; $h_\theta(x_i)$ ← hypothesis function; y_i ← real output



Curve fitting

$$\boxed{h_\theta(x) = w_1 + w_2 x + w_3 x^2}$$

$$\ell(y_1, y_2) = (y_1 - y_2)^2 \quad \frac{\partial \ell}{\partial \theta} = (h_\theta - y_i)\, x_i$$

k, λ hyperparameter

Linear Program : An optimization type

Political winning : investing money to win elections

Issues	Class 1	Class 2	Class 3
$x_1 \to$ A	−2	5	3
x_2 B	8	2	−5
x_3 C	0	0	10
x_4 D	10	0	2
Population	100,000	200,000	50,000
majority	50,000	100000	25000

$$\min \; x_1 + x_2 + x_3 + x_4$$

$$\text{s.t.} \quad -2x_1 + 5x_2 + 0 \cdot x_3 + 10x_4 \geq 50000$$
$$8x_1 + 2x_2 \geq 100000$$
$$3x_1 - 5x_2 + 10x_3 + 2x_4 \geq 25000$$
$$x_1, x_2, x_3, x_4 \geq 0$$

linear program

class

$$x_1 + \frac{25}{22} x_2 + \frac{10}{22} x_3 + \frac{2}{22} x_4 \geq \frac{31\,00000}{111}$$

$$x_1 + x_2 + x_3 + x_4 \geq \frac{31\,00000}{111}$$

linear → f ←

An Optimal solution →

$$\begin{aligned} x_1^* &= \frac{2017000}{111} \to (A) \\ x_2^* &= \frac{452000}{111} \to (B) \\ x_3^* &= 0 \to (C) \\ x_4^* &= \frac{620000}{111} \to (D) \end{aligned}$$

Optimal value

$$x_1^* + x_2^* + x_3^* + x_4^* = \frac{31\,00000}{111}$$

Standard form of linear program

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n \quad \text{variables to optimize} \quad c = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} \in \mathbb{R}^n$$

$$\max \; c^T x = c_1 x_1 + c_2 x_2 + \cdots + c_n x_n$$

$$\text{s.t.} \quad A_{m \times n} \, x_{n \times 1} \leq b_{m \times 1} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$$

$$x \geq 0 \equiv \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \geq \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

Weak duality

feasible x and y

primal dual

$$b^T y \geq x^T A^T y \geq x^T c$$

$$y \geq 0 \qquad x \geq 0$$

$$y^T b \geq y^T A x$$

$$b^T y \geq c^T x$$

min max

$$Ax \leq b$$

$$a_1^T x \leq b_1 \quad \times y_1$$

$$\vdots$$

$$A = \begin{bmatrix} - a_1^T - \\ \vdots \\ - a_m^T - \end{bmatrix}$$

$$b^T y^* = c^T x^*$$

dual optimal primal optimal