Lecture Note Optimization for Robotics I

What is the goal of optimization?

→ Find the best solution to a problem subject to certain constraints.

decision variable. objective function

Derivative for vector Function.

ความชัน

→ Gradient : $\nabla_x f \in \mathbb{R}^{n \times 1}$; $\nabla_x f$ is column vector. $\begin{bmatrix} \cdots \\ \cdots \\ \cdots \end{bmatrix}$

→ Partial Derivative : $\dfrac{\partial f}{\partial x}$; $\dfrac{\partial f}{\partial x}$ is row vector $[\ \cdots\ \cdots\ \cdots\ ]$

relation between $\nabla_x f = \left(\dfrac{\partial f}{\partial x}\right)^T$

ความโค้ง

→ Hessian : $\nabla_x^2 f$ is the matrix of second-order partial derivative.

$$\min_x f(x)$$

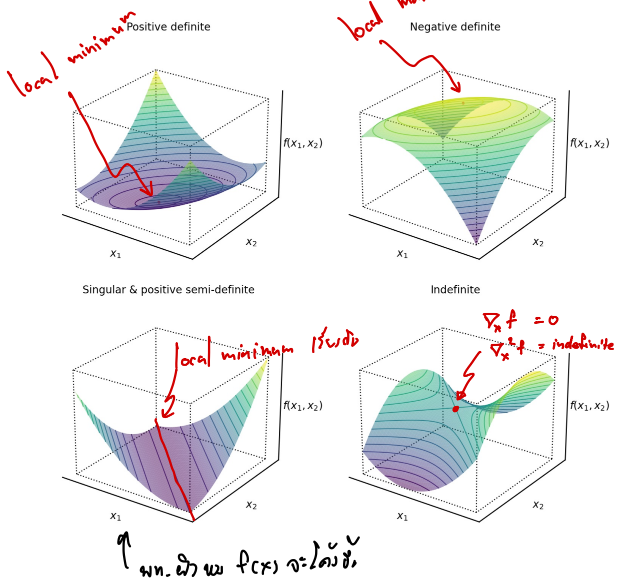
$x \in \mathbb{R}^N$: variable

$f(x) : \mathbb{R}^N \to \mathbb{R}$: objective function.

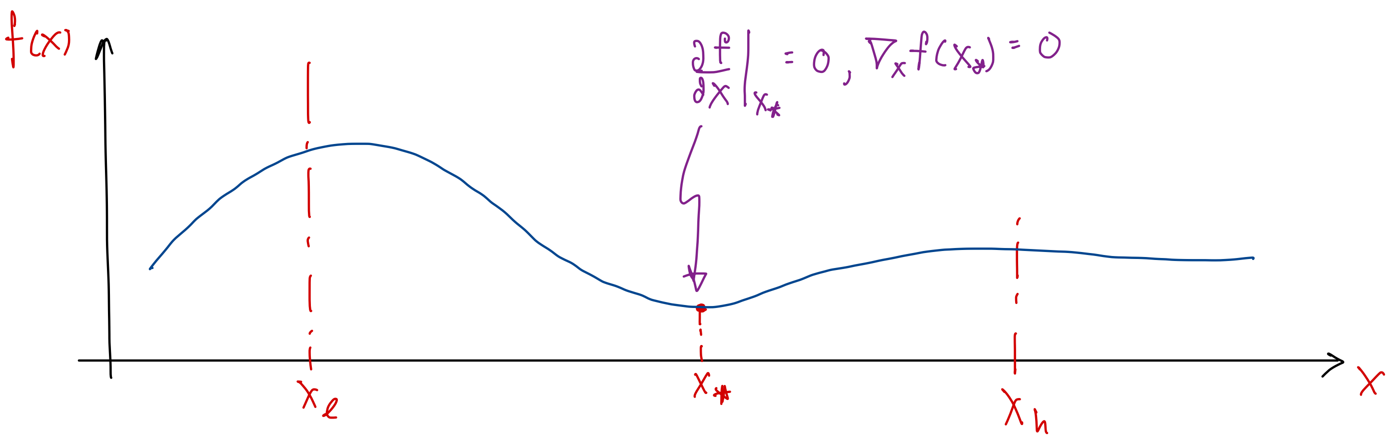
x_* : is the optimal solution

→ global optimum

→ local optimum



$$\left.\frac{\partial f}{\partial x}\right|_{x_*} = 0, \quad \nabla_x f(x_*) = 0$$



Problems related to optimization.

→ Root Finding : Find x^* such that $f(x^*) = 0$

→ Fixed Point : Find x^* such that $f(x^*) = x^*$

Example

Root Finding

$$f(x) = x^2 - 4 = 0$$

$$x = \pm 2$$

Fixed Point

$$g(x) = \cos(x)$$

Find x^* that make $g(x^*) = x^*$

Iteration .... $x_0 = 1$

$x_1 = \cos(x_0) = \cos(1) \approx 0.54$

$x_2 = \cos(x_1) = \cos(0.54) \approx 0.856$

$\vdots$

$x_n = 0.739$

Optimization via Root Finding

- Newton's Method

→ First-Order Taylor Expansion

$$f(x_k + \Delta x) \approx f(x_k) + \nabla f(x_k)^T \Delta x = 0$$

$$\Delta x = -\nabla f(x_k)^{-1} f(x_k)$$

→ Second-Order Taylor Expansion

$$f(x_k + \Delta x) \approx \underbrace{f(x_k)}_{\text{ค่าของ } f(n) \text{ ปัจจุบัน}} + \underbrace{\nabla f(x_k)^T \Delta x}_{\text{ความชัน (แนวโน้ม)}} + \underbrace{\frac{1}{2} \Delta x^T \nabla^2 f(x_k) \Delta x}_{\text{ความโค้ง}}$$

$$\frac{\partial f(x_k + \Delta x)}{\partial \Delta x} \approx \frac{\partial}{\partial \Delta x}\left(f(x_k) + \nabla f(x_k)^T \Delta x + \frac{1}{2} \Delta x^T \nabla^2 f(x_k) \Delta x\right) = 0$$

$$\nabla f(x_k) + \nabla^2 f(x_k) \Delta x = 0$$

$$\Delta x = \underbrace{-}_{\text{descent}} \underbrace{\nabla^2 f(x_k)^{-1}}_{\text{learning rate}} \underbrace{\nabla f(x_k)}_{\text{gradient}}$$

$$x_{k+1} = x_k + \Delta x$$

Rastrigin Function

$$f(x) = Ad + \sum_{i=0}^{d} x^T x - A \sum_{i=0}^{d} \cos(2\pi x)$$

$d = 2$

$$f(x) = A(2) + x_1^T x_1 + x_2^T x_2 - A(\cos(2\pi x_1) + \cos(2\pi x_2))$$

$$\frac{\partial f(x)}{\partial x_1} = 2x_1 + A\sin(2\pi x_1)2\pi$$

$$\frac{\partial f(x)}{\partial x_2} = 2x_2 + A\sin(2\pi x_2)2\pi \quad ; X = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$\therefore \frac{\partial f(x)}{\partial x} = 2X^T + 2\pi A\sin(2\pi X^T)$$

$$\frac{\partial^2 f(x)}{\partial x^2} = 2 + 2\pi A(2\pi)\cos(2\pi X^T)$$

$$= 2 + 4\pi^2 A\cos(2\pi X^T)$$

"Damped" Newton's Method.

$$H = \nabla_x^2 f(x_k)$$

while H is not positive definite (PSD)

while $H \preceq 0$

$$H = H + \beta I$$

$$\Delta x = -H^{-1} \nabla_x f(x_k)$$

Armijo Line Search.

→ Find the direction : Damped Newton method.

$$\Delta x = -(\nabla^2 f(x_k) + \beta I)^{-1} \nabla f(x_k)$$

ขยายจนกว่าจะ: positive definite.

→ Find the step size : Armijo line search

$$\alpha = 1$$

$$f(x_k + \alpha \Delta x) > f(x_k) + b \alpha \nabla f(x_k)^T \Delta x$$

$$\alpha = c \alpha$$

$$x_{k+1} = x_k + \alpha \Delta x$$