# Lecture 1-1 : Fundamental Math Related to Optimization 

노트 제목

## Optimization problem

⇒ Formulated as "Nonlinear Programming"

- min $f(x)$
- Subject to $g_i(\underset{\sim}{x}) \le 0 \quad (i=1,\cdots,m)$

  and $h_j(\underset{\sim}{x}) = 0 \quad (j=1,\cdots,l)$

( design variable $\underset{\sim}{x} = \{x_1, x_2, \cdots, x_n\}^T$ )

Will Begin with

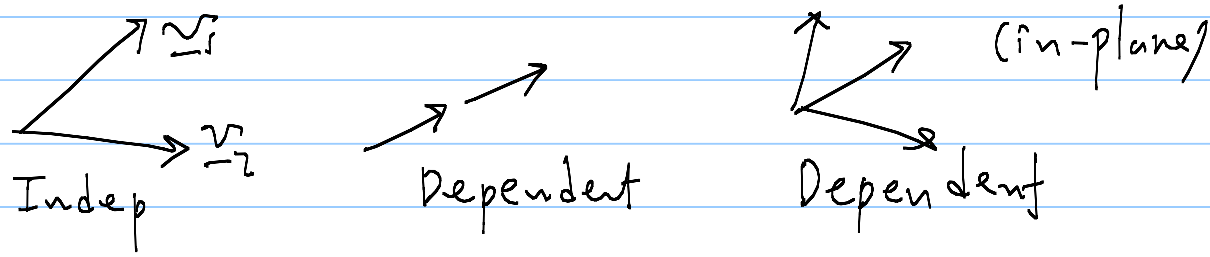
> * Mathematical Notation
> * "Directional Derivative": $D_\alpha f$
> * Taylor expansion, Hessian

- Design (or decision) Variables : $\underset{\sim}{x}$

$$\underset{\sim}{x} = \begin{Bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{Bmatrix} \in \mathbb{R}^n$$

Real

## ☑ Linear Independence of vectors




$\Rightarrow$ Vectors are linearly independent

if $\sum_{k=1}^{m} \alpha_k \underset{-}{v}_k = 0$ implies

all $\alpha_k \equiv 0$

## ☑ Quadratic Form / Positive-definite

Quadratic form :

$$\begin{cases} f = x_1^2 - 6x_1x_2 + 9x_2^2 \\ f = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j \end{cases}$$

Quadratic form includes

$$f = \underbrace{c + \sum b_i x_i}_{} + \frac{1}{2} \sum_i \sum_j a_{ij} x_i x_j$$

$$= c + \underset{\sim}{b}^T \underset{\sim}{x} + \frac{1}{2} \underset{\sim}{x}^T \underset{\sim}{A} \underset{\sim}{x}$$

where

$$\underset{\sim}{A} = [a_{ij}] = A^T$$

Example :

$$f = x_1^2 - 6x_1 x_2 + 9x_2$$

$$= (x_1, x_2)\begin{bmatrix} 1 & -3 \\ -3 & 9 \end{bmatrix}\begin{Bmatrix} x_1 \\ x_2 \end{Bmatrix}$$

$$= \underset{\sim}{x}^T \underline{A}\, \underset{\sim}{x}$$

$\uparrow$ symmetric $\underline{A} = \underline{A}^T$

* Positive-Definiteness of Q-Form

* Q-form is P-D (positive-definite) if $\underset{\sim}{x}^T \underline{A}\, \underset{\sim}{x} > 0$ for every non-zero $\underset{\sim}{x}$

and $\underset{\sim}{x}^T \underline{A}\, \underset{\sim}{x} = 0$ iff $\underset{\sim}{x} = 0$

In this case, we can also say that

$\underline{A}$ = positive-definite matrix

* Positive Semi-definite

If $\underset{\sim}{x}^T \underset{\sim}{A} \underset{\sim}{x} \geq 0$

for every nonzero $\underset{\sim}{x}$

Example:

$f = \frac{1}{3} x^2$ : P-D

( $f > 0$ for $x \neq 0$
$f = 0$ iff $x = 0$ )

$f = \frac{1}{3} x_1^2 + \frac{3}{2} x_2^2$ : P-D

( $f > 0$ for $x_1 = x_2 \neq 0$
$f = 0$ iff $x_1 = x_2 = 0$ )

$f = x_1^2 - 2x_1 x_2 + 2x_2^2$ : P-D

( $f = (x_1 - x_2)^2 + x_2^2$ :

$f = 0$ iff $x_1 = x_2$ & $x_2 = 0$

$\rightarrow x_1 = x_2 = 0$ )

$f = x_1^2 - 2x_1 x_2 + x_2^2$

$= (x_1 - x_2)^2$ : P-SD

( $f$ can be zero for $x_1 = x_2 \neq 0$ )

## How to check p-p of $\underline{x}^T \underline{A}\, \underline{x}$ ?

i) Check eigenvalues of $\underline{A}$.

Let eigen value and vector of $\underline{A}$ as $\lambda_i$, $\underline{\xi}_i$ $(i=1,\cdots,n)$

$$\begin{cases} \underset{\sim}{A}\, \underset{\sim}{\xi}_i = \lambda_i\, \underline{\xi}_i & (i=1,\cdots,n) \\ \underset{\sim}{\xi}_j^T \underset{\sim}{\xi}_i = \delta_{ij} = \begin{matrix} 1 & \text{if } i=j \\ 0 & \text{else} \end{matrix} \end{cases}$$

Then $\underset{\sim}{x} = \sum_{i=1}^{n} \alpha_i \underline{\xi}_i$

$$\begin{aligned} Q &= \underline{x}^T \underline{A}\, \underline{x} \\ &= \left(\sum_{i=1}^{n} \alpha_i \underline{\xi}_i\right)^T \underline{A} \left(\sum_{j=1}^{n} \alpha_j \underset{\sim}{\xi}_j\right) \\ &= \left(\sum_{i=1}^{n} \alpha_i \underline{\xi}_i\right)^T \left(\sum_{j=1}^{n} \alpha_j \underbrace{\underline{A}\, \underline{\xi}_j}_{\lambda_j \underset{\sim}{\xi}_j}\right) \\ &= \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j \lambda_j \underbrace{\underline{\xi}_i^T \underset{\sim}{\xi}_j}_{\delta_{ij}} \\ &= \sum_{i=1}^{n} \alpha_i^2 \lambda_i \end{aligned}$$

Therefore,

all $\lambda_i$ (eigenvalues of $\underset{\sim}{A}$) $> 0$

then $Q$ = P-D.

## ii) Sylvester's test

$$\underset{\sim}{A} = \begin{bmatrix} \boxed{a_{11}}^{\,A^1} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & & \\ \vdots & & \ddots & \\ a_{n1} & & & a_{1n} \end{bmatrix}$$

$A^1$, $A^2$

① $A$ = P-D

if $\det A^i > 0$ for $i = 1, \cdots, n$

② $A$ = P-SD

if $\det A^i \geq 0$

③ $A$ = Negative-Definite

if $(-1)^i \det A^i > 0$

for $i = 1, \cdots, n$

## 4 Sets

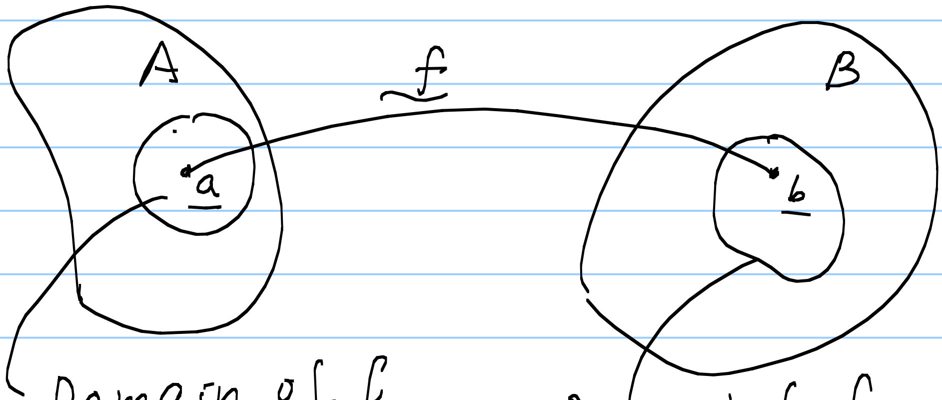
- Closed set : $S = \{ \underline{x} \mid |\underline{x}| \le 1 \}$
  Open set : $S = \{ \underline{x} \mid |\underline{x}| < 1 \}$

- Bounded set : For $\underset{\sim}{a} \in S$,

  $$\underset{\sim}{a}^T \underset{\sim}{a} \le \infty$$

- Compact set : Closed and bounded

## 1 Definition of a function $\underline{f}$




Domain of $f$
$(A \subset R^n)$

Range of $f$
$B \subset R^m$

## Gradient

① For a scalar ftn $f(x) \in C^1$

← a set of continuous ftn having conti first-derivative

$$\underline{\nabla} f \triangleq \begin{Bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{Bmatrix} = \left\{ \frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \cdots, \frac{\partial f}{\partial x_n} \right\}^T$$

② For a vector ftn $\underline{g}(\underline{x})$

$$\underline{g}(x) = (\underline{g}_1, \underline{g}_2, \underline{g}_3, \cdots \underline{g}_m)$$

$$g_i(\underline{x}) \in C^1$$

$$\underline{\nabla}\, \underline{g} \triangleq \begin{bmatrix} \frac{\partial g_1}{\partial x_1} & \cdots & \frac{\partial g_m}{\partial x_1} \\ \vdots & \ddots & \\ \frac{\partial g_1}{\partial x_n} & & \frac{\partial g_m}{\partial x_n} \end{bmatrix}$$

## ③ Directional derivative of $f$ in the direction of $\underset{\sim}{d} \neq 0$ at $\underset{\sim}{x}$

$$D_d f \triangleq \lim_{a \to 0} \frac{f(\underset{\sim}{x} + a\underline{d}) - f(\underset{\sim}{x})}{a}$$

$$= \frac{d}{da} f(\underline{x} + a\underline{d})\Big|_{a=0}$$

$$= \underbrace{(\underline{\nabla f})^T \underline{d}}_{\text{matrix notation}} \quad \Big(\text{or } \underbrace{\underline{\nabla f} \cdot \underline{d}}_{\text{vector notation}}\Big)$$

why? $\frac{d}{da} f(\underline{x} + \underline{a}\underline{d})\Big|_{a=0}$

$$\underset{\text{Chain Rule}}{=} \sum_{i=1}^{n} \left[ \frac{\partial f}{\partial(x_i + a d_i)} \frac{\partial}{\partial a}(x_i + a d_i) \right]_{a=0}$$

$$= \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} d_i$$

$$= \left( \frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \cdots \frac{\partial f}{\partial x_n} \right) \begin{pmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \end{pmatrix}$$

$$= (\underline{\nabla f})^T \underline{d}$$

## Taylor Expansion

$$f(x + ad) = f(x) + \frac{\partial f}{\partial x}(x)\,(ad) + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(ad)^2 + \cdots$$

$a > 0$

$$f(\underset{\sim}{x} + a\underset{\sim}{d}) = f(\underset{\sim}{x}) + a \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} d_i$$

$$= \frac{1}{2} a \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j} d_i d_j + \cdots$$

$$= f(\underset{\sim}{x}) + a(\nabla f)^T \underline{d} + \frac{1}{2} a^2 d^T \underline{H}\, \underline{d} + \cdots$$

where

$$\underset{\sim}{H} = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & & & \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \cdots & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_n} \end{bmatrix}$$

Hessian (or Hessian matrix)

$(\underset{\sim}{H} = \underset{\sim}{H}^T \leftarrow$ always symmetric$)$