14기 이건일

1-1 $f(x,y,z) = x^2+y^2+1.5z^2+xy+yz+zx$

(1) $\frac{\partial f}{\partial x} = 2x + y + z$

$\frac{\partial f}{\partial y} = 2y + x + z$

$\frac{\partial f}{\partial z} = 3z + y + x$

$\frac{\partial^2 f}{\partial x^2} = 2$, $\frac{\partial^2 f}{\partial y^2} = 2$, $\frac{\partial^2 f}{\partial z^2} = 3$

$\frac{\partial^2 f}{\partial x \partial y} = 1$ $\frac{\partial^2 f}{\partial y \partial z} = 1$ $\frac{\partial^2 f}{\partial z \partial x} = 1$

$A = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 3 \end{bmatrix}$

주대각선 소행렬식 방법

$D_1 = 2 > 0$

$D_2 = 4-1 = 3 > 0$

$D_3 = \det(A) = 2(6-1) - 1(3-1) + 1(1-2)$

$= 7 > 0$

D_1, D_2, D_3 모두 양수이므로 A는 PD.

(2)

$\mathrm{adj}(A) = \begin{bmatrix} 5 & -2 & -1 \\ -2 & 5 & -1 \\ -1 & -1 & 3 \end{bmatrix}$

$A^{-1} = \frac{1}{\det(A)} \begin{bmatrix} 5 & -2 & -1 \\ -2 & 5 & -1 \\ -1 & -1 & 3 \end{bmatrix} = \frac{1}{7} \begin{bmatrix} 5 & -2 & -1 \\ -2 & 5 & -1 \\ -1 & -1 & 3 \end{bmatrix}$

(3) xy 항이 kxy가 되어 $\det(A')=0$ 인 상황

$$A' = \begin{bmatrix} 2 & k & 1 \\ k & 2 & 1 \\ 1 & 1 & 3 \end{bmatrix} \qquad \det(A') = -3k^2 + 2k + 8 = 0$$

$\det(A')=0$ 이면 A'는 non-invertible.

Non-invertible 한 행렬 A는 $Ax=b$ 에서 어떤 b에 대해서 해가 존재하지 않거나 무수히 많이 존재함.

따라서 모든 b에 대해 유일한 해라는 조건이 만족하지 않음.

2

$$B = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$M = B^T B = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\det(M - \lambda I) = \begin{bmatrix} 1-\lambda & 1 & 0 \\ 1 & 1-\lambda & 0 \\ 0 & 0 & 1-\lambda \end{bmatrix}$$

$$(1-\lambda)\{(1-\lambda)^2 - 1\}$$
$$= (1-\lambda)\{(1-\lambda+1)(1-\lambda-1)\}$$
$$= (1-\lambda)\{(2-\lambda)(-\lambda)\}$$
$$= -\lambda(1-\lambda)(2-\lambda)$$

$\lambda_1 = 2,\ \lambda_2 = 1,\ \lambda_3 = 0$

$\sigma_1 = \sqrt{2},\ \sigma_2 = 1,\ \sigma_3 = 0$

i) $\lambda_1 = 2 \quad (M - 2I)w = 0$

$$M - 2I = \begin{bmatrix} -1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}\begin{matrix} w_1 \\ w_2 \\ w_3 \end{matrix} \qquad v_1 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$

$-w_1 + w_2 = 0, \quad w_3 = 0$

$w_1 = w_2$

ii) $\lambda_2 = 1 \quad (M - I)w = 0$

$$M - I = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}^{w} \qquad v_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \qquad V = \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \end{bmatrix}$$

$w_2 = 0, \ w_1 = 0, \ w_3$ 자유

iii) $\lambda_3 = 0 \quad Mw = 0$

$$M = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}^{w} \qquad v_3 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$$

$w_1 + w_2 = 0, \ w_3 = 0$

$w_1 = -w_2, \ w_3 = 0$

$$BB^T = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$$

$$u_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad u_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \qquad U = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\Sigma = \begin{bmatrix} \sqrt{2} & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

$$B = U \Sigma V^T$$

$$U = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \Sigma = \begin{bmatrix} \sqrt{2} & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \quad V^T = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \end{bmatrix}$$

3-1

(1). $C_1 = \{ (x,y) \in \mathbb{R}^2 : x+y \leq 1 \}$

Suppose that $(x_1, y_1), (x_2, y_2) \in C_1$

$x_1 + y_1 \leq 1, \quad x_2 + y_2 \leq 1$

$\theta \in [0,1]$

$$(x_\theta, y_\theta) = \theta(x_1, y_1) + (1-\theta)(x_2, y_2)$$

$$x_\theta = \theta x_1 + (1-\theta) x_2$$
$$y_\theta = \theta y_1 + (1-\theta) y_2$$

$$x_\theta + y_\theta = \theta(x_1 + y_1) + (1-\theta)(x_2 + y_2) \leq \theta \cdot 1 + (1-\theta) \cdot 1 = 1$$

$\therefore (x_\theta, y_\theta) \in C_1$, C_1 is convex.

$C_2 = \{ x \in \mathbb{R}^n : \|x\|_1 \le 1 \}$

Suppose that $x_1, x_2 \in C_2$

$\|x_1\|_1 \le 1 \quad , \quad \|x_2\|_1 \le 1$

$\theta \in [0,1]$, $x_\theta = \theta x_1 + (1-\theta) x_2$ 로 두자.

l_1-norm의 삼각부등식과 동차성으로

$$\|x_\theta\|_1 = \|\theta x_1 + (1-\theta)x_2\|_1 \le \|\theta x_1\|_1 + \|(1-\theta)x_2\|_1$$
$$= \theta\|x_1\|_1 + (1-\theta)\|x_2\|_1$$

유한히 상계 적용.

$$\|x_\theta\|_1 \le \theta \cdot 1 + (1-\theta)\cdot 1 = 1$$

$\therefore$ $x_\theta \in C_2$, C_2는 convex

$C_3 = \{ (x,y) \in \mathbb{R}^2 : y \ge e^x \}$

Suppose that $(x_1, y_1), (x_2, y_2) \in C_3$

$$y_1 \ge e^{x_1} \quad , \quad y_2 \ge e^{x_2}$$

$\theta \in [0,1]$

$(x_\theta, y_\theta) = \theta(x_1, y_1) + (1-\theta)(x_2, y_2)$ 라고 두자.

$x_\theta = \theta x_1 + (1-\theta)x_2$, $y_\theta = \theta y_1 + (1-\theta) y_2$

지수함수 e^x는 convex,

$$e^{x_\theta} = e^{\theta x_1 + (1-\theta)x_2} \le \theta e^{x_1} + (1-\theta) e^{x_2}$$

$y_1 \ge e^{x_1}$, $y_2 \ge e^{x_2}$

$$y_\theta = \theta y_1 + (1-\theta) y_2 \ge \theta e^{x_1} + (1-\theta) e^{x_2} \ge e^{x_\theta}$$

$\therefore$ $(x_\theta, y_\theta) \in C_3$, C_3는 convex.

(2)

$$S = \{(x,t) \in \mathbb{R}^n \times \mathbb{R} : t \geq f(x)\}$$

Suppose that $(x_1, t_1), (x_2, t_2) \in S$

$t_1 \geq f(x_1)$, $t_2 \geq f(x_2)$

$\theta \in [0,1]$에 대해 $(x_\theta, t_\theta) = \theta(x_1, t_1) + (1-\theta)(x_2, t_2)$ 라 두자.

$\rightarrow \quad x_\theta = \theta x_1 + (1-\theta)x_2$, $t_\theta = \theta t_1 + (1-\theta)t_2$

f가 convex 이므로

$$f(x_\theta) = f(\theta x_1 + (1-\theta)x_2) \leq \theta f(x_1) + (1-\theta)f(x_2)$$

위의 t_1, t_2 부등식을 고려

$$t_\theta = \theta t_1 + (1-\theta)t_2 \geq \theta f(x_1) + (1-\theta)f(x_2) \geq f(x_\theta)$$

$\therefore (x_\theta, t_\theta) \in S$. S는 convex set

3-2

(1) -a.

h convex $\Longleftrightarrow \forall x,y, \forall \theta \in [0,1], h(\theta x + (1-\theta)y) \leq \theta h(x) + (1-\theta)h(y)$

$h(x) = f(x) + g(x)$

$h(\theta x + (1-\theta)y) = f(\theta x + (1-\theta)y) + g(\theta x + (1-\theta)y)$

f, g가 convex 이므로

$$f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta)f(y)$$
$$g(\theta x + (1-\theta)y) \leq \theta g(x) + (1-\theta)g(y)$$

$$h(\theta x + (1-\theta)y) \leq \theta(f(x) + g(x)) + (1-\theta)(f(y) + g(y))$$

$$h(\theta x + (1-\theta)y) \leq \theta h(x) + (1-\theta)h(y)$$

$\therefore f(x) + g(x)$는 convex

(1)-b.

$h(x) = f(Ax+b)$로 두자.

임의의 x, y, $\theta \in [0,1]$에 대해

$$h(\theta x+(1-\theta)y) = f(A(\theta x+(1-\theta)y)+b)$$

행렬의 선형성으로

$$A(\theta x+(1-\theta)y)+b = \theta(Ax+b)+(1-\theta)(Ay+b)$$

따라서

$$h(\theta x+(1-\theta)y)+b = f(\theta(Ax+b)+(1-\theta)(Ay+b))$$

f가 convex 이므로

$$f(\theta u+(1-\theta)v) \le \theta f(u)+(1-\theta)f(v)$$

$u = Ax+b$, $v = Ay+b$로 놓으면

$$h(\theta x+(1-\theta)y) \le \theta f(Ax+b)+(1-\theta)f(Ay+b)$$

$$h(\theta x+(1-\theta)y) \le \theta h(x)+(1-\theta)h(y)$$

$\therefore$ $f(Ax+b)$는 convex.

(2) $f(x) = \|Ax-b\|^2$, 임의의 $x, y \in \mathbb{R}^d$, $\theta \in [0,1]$

$u := Ax-b$, $v := Ay-b$

$f(x) = \|u\|^2$, $f(y) = \|v\|^2$

$$f(\theta x+(1-\theta)y) = \|A(\theta x+(1-\theta)y-b\|^2$$

선형성으로

$$A(\theta x+(1-\theta)y)-b = \theta(Ax-b)+(1-\theta)(Ay-b) = \theta u+(1-\theta)v$$

$$f(\theta x+(1-\theta)y) = \|\theta u+(1-\theta)v\|^2$$

$$\|\theta u + (1-\theta)v\|^2 = (\theta u + (1-\theta)v)^T(\theta u + (1-\theta)v)$$
$$= \theta^2\|u\|^2 + (1-\theta)^2\|v\|^2 + 2\theta(1-\theta)u^Tv$$

차이 계산

$$D := \theta\|u\|^2 + (1-\theta)\|v\|^2 - \|\theta u + (1-\theta)v\|^2$$
$$= \theta\|u\|^2 + (1-\theta)\|v\|^2 - \{\theta^2\|u\|^2 + (1-\theta)^2\|v\|^2 + 2\theta(1-\theta)u^Tv\}$$
$$= \theta(1-\theta)\|u\|^2 + \theta(1-\theta)\|v\|^2 - 2\theta(1-\theta)u^Tv$$
$$= \theta(1-\theta)(\|u\|^2 + \|v\|^2 - 2u^Tv)$$
$$= \theta(1-\theta)(u-v)^T(u-v)$$
$$= \theta(1-\theta)\|u-v\|^2$$

$$\theta \in [0,1] \Rightarrow \theta(1-\theta) \geq 0, \quad \|u-v\|^2 \geq 0$$

$$D = \theta(1-\theta)\|u-v\|^2 \geq 0$$

$$\|\theta u + (1-\theta)v\|^2 \leq \theta\|u\|^2 + (1-\theta)\|v\|^2$$

원래 변수로 돌아가면

$$f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta)f(y)$$

$\therefore$ $f(x) = \|Ax-b\|^2$ 는 convex.

#4-1

X \ Y	0	1
0	$\frac{1}{3}$	$\frac{1}{3}$
1	0	$\frac{1}{3}$

$P(X=0) = \frac{2}{3}$ $\quad P(Y=0) = \frac{1}{3}$

$P(X=1) = \frac{1}{3}$ $\quad P(Y=1) = \frac{2}{3}$

(a) $$H(X) = -\frac{2}{3}\log_2\frac{2}{3} - \frac{1}{3}\log_2\frac{1}{3}$$
$$= -\frac{2}{3}(1 - \log_2 3) + \frac{1}{3}\log_2 3$$
$$= \log_2 3 - \frac{2}{3} \approx 0.9183$$

$H(Y)$: 대칭

$H(Y) \approx 0.9183$

(b) $H(X|Y)$

i) $Y=0 \longrightarrow X=0$ 확정. $H(X|Y=0)=0$

ii) $Y=1$ $P(X=0|Y=1) = \frac{1}{3}/\frac{2}{3} = \frac{1}{2}$, $H(X|Y=1)=1$
$P(X=1|Y=1) = \frac{1}{3}/\frac{2}{3} = \frac{1}{2}$

$$H(X|Y) = P(Y=0)\cdot 0 + P(Y=1)\cdot 1 = \frac{2}{3}$$

$H(Y|X)$ 대칭.

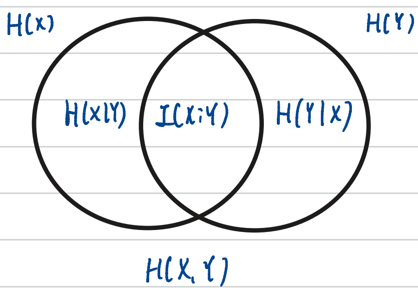
$$H(Y|X) = \frac{2}{3}$$

(c) $H(X,Y) = H(Y) + H(X|Y)$
$= (\log_2 3 - \frac{2}{3}) + \frac{2}{3}$
$= \log_2 3 \approx 1.5850$

(d) $H(Y) - H(Y|X) = (\log_2 3 - \frac{2}{3}) - \frac{2}{3}$
$= \log_2 3 - \frac{4}{3} \approx 0.2516$

(e) $I(X;Y) = H(X) - H(X|Y)$
$= (\log_2 3 - \frac{2}{3}) - \frac{2}{3}$
$= \log_2 3 - \frac{4}{3} \approx 0.2516.$

(f)



4-2

(a) $p = (0.9,\ 0.1)$, $q = (0.5,\ 0.5)$

$$D(q\|p) = \sum_i q_i \ln \frac{q_i}{p_i} = 0.5 \ln \frac{0.5}{0.9} + 0.5 \ln \frac{0.5}{0.1} \approx 0.5108$$

$$D(p\|q) = \sum_i p_i \ln \frac{p_i}{q_i} = 0.9 \ln \frac{0.9}{0.5} + 0.1 \ln \frac{0.1}{0.5} \approx 0.3681$$

(b)

$$D(p\|q) = \sum_x p(x) \ln \frac{p(x)}{q(x)} = -\sum_x p(x) \ln \frac{q(x)}{p(x)} = -\mathbb{E}_p\left[\ln\left(\frac{q(x)}{p(x)}\right)\right]$$

$\ln(\cdot)$는 concave 이므로 Jensen.

$$\mathbb{E}_p[\ln Z] \le \ln(\mathbb{E}_p[Z])$$

$Z = \frac{q(x)}{p(x)}$ 로 두자.

$$\mathbb{E}_p\left[\ln\left(\frac{q(x)}{p(x)}\right)\right] \le \ln\left(\mathbb{E}_p\left[\frac{q(x)}{p(x)}\right]\right) = \ln\left(\sum_x p(x)\frac{q(x)}{p(x)}\right)$$

$$= \ln\left(\sum_x q(x)\right)$$

$$= \ln 1 = 0$$

$$\therefore\ D(p\|q) = -\mathbb{E}_p\left[\ln\left(\frac{q(x)}{p(x)}\right)\right] \ge 0$$