Question.2-01

최적화의 대상이 되는 함수 $f(x)$가 다음과 주어졌다고 하자.

$$f(x) = 3(x+2)^2 - 1$$

위의 $f(x)$에 대하여 optimal point는 $x_o = \underset{x}{\operatorname{argmin}} f(x)$에 대하여 $(x_o, f(x_o))$로 정의될때, 다음 문제들의 답을 구하시오.

1) derivative $\frac{df(x)}{dx}$를 구한 뒤, $f(x)$와 $\frac{df(x)}{dx}$의 개형을 각각 그리시오.

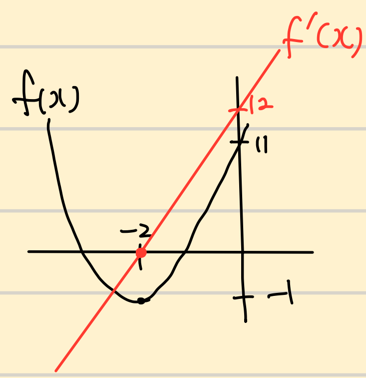
2) $f(x)$의 값이 최소가 되는 optimal point에서의 x값을 구하시오.

3) differential coefficient를 이용하여 $x = -4, -3, -2, -1, 0$에서 각각 x가 optimal point로 이동해야 하는 방향을 $+, -$, 변화없음 중 하나로 나타내시오.

4) 3)의 풀이를 바탕으로 $\frac{df(x)}{dx}$와, 임의의 x가 optimal point로 이동하기 위한 방향과의 관계를 설명하시오.

1) $f(x) = 3(x^2 + 4x + 4) - 1 = 3x^2 + 12x + 11$

$\frac{df}{dx} = 6x + 12$



2) optimal point 에서의 x값 (x_o)

$f'(x_o) = 0, \quad x_o = -2$

3) differential coefficient $= 6x + 12$

$f'(-4) < 0, \quad f'(-3) < 0, \quad f'(-2) = 0, \quad f'(-1) > 0, \quad f'(0) > 0$

이동해야 하는 방향 : $+$ ⇓ 변화없음 이동해야 하는 방향 : $-$

4) $\frac{df}{dx}$ 는 현재 위치에서 $f(x)$ 값이 증가하기 위한 x의 이동방향을 나타낸다.

$\therefore$ optimal point로 가기위한 방향은 $\frac{df}{dx}$ 와 반대방향이다.

Question.2-02

최적화의 대상이 되는 함수 $f(x,y)$가 다음과 주어졌다고 하자.

$$f(x,y) = x^2 + y^2 - xy$$

이때 주어진 f의 contour plot은 오른쪽 그림과 같고,

optimal point는 $\underset{(x,y)}{\operatorname{argmin}} f(x,y)$을 만족시키는 (x,y)로 정의할 때,

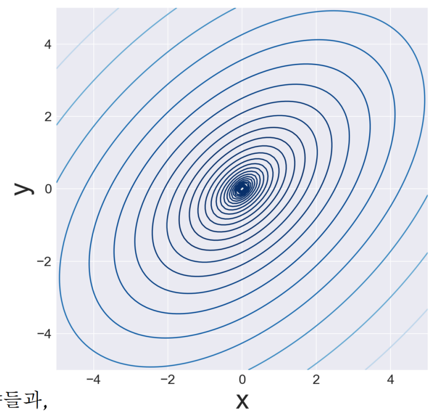
다음 물제들에 답하시오.

1) f에 대해 임의의 (x,y)에서의 gradient $\nabla_{(x,y)} f(x,y)$를 구하시오.

2) 다음의 (x,y) pair들에 대해 gradient들을 구하고,
contour plot 위에 gradient들을 표시하시오

$$(x,y) = (1,1), (2,1)$$

3) 위 (x,y)들에 대해 (x,y)들에서 함수값 f를 가장 크게 증가시키는 방향들과,
가장 크게 감소시키는 방향들을 구하시오.



1) $\nabla f(x,y) = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix} = \begin{bmatrix} 2x-1 \\ 2y-1 \end{bmatrix}$

2) $\nabla f(x,y)\big|_{(1,1)} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

$\nabla f(x,y)\big|_{(2,1)} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$

3) (1,1) 에 대해 f는 가장 크게 증가시키는 방향 $= \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

// 감소 // $= \begin{bmatrix} -1 \\ -1 \end{bmatrix}$

(2,1) // 증가 // $= \begin{bmatrix} 3 \\ 1 \end{bmatrix}$

// 감소 // $= \begin{bmatrix} -3 \\ -1 \end{bmatrix}$

Question.2-03

함수 f가 $\vec{\theta}$에 대해 다음과 같이 주어졌을 때, Jacobian matrix $\dfrac{\partial f(\vec{\theta})}{\partial \vec{\theta}}$를 구하시오.

1) $$\vec{\theta} = \begin{pmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \end{pmatrix} \qquad f(\vec{\theta}) = (\theta_1)^3 - 2(\theta_2)^2 + \theta_1\theta_3$$

2) $$\vec{\theta} = \begin{pmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \end{pmatrix} \qquad f(\vec{\theta}) = ln^2(\theta_1) - 3e^{\theta_2\theta_3} + tan(\theta_3)$$

$$\begin{array}{c|cc} & x & \vec{x} \\ \hline f & & \boxed{\dfrac{\partial f}{\partial \vec{x}}} \text{ gradient} \\ \vec{f} & \boxed{\dfrac{\partial \vec{f}}{\partial x}} & \boxed{\dfrac{\partial \vec{f}}{\partial \vec{x}}} \text{ Jacobian matrix} \end{array}$$

Jacobian matrix $= \dfrac{\partial \vec{f}}{\partial \vec{x}}$

$$= \begin{bmatrix} \nabla f_1(\vec{x}) \\ \nabla f_2(\vec{x}) \\ \vdots \\ \nabla f_n(\vec{x}) \end{bmatrix}$$ gradient를 row-wise로 쌓은 matrix

1) gradient $= \begin{bmatrix} \dfrac{\partial f}{\partial \theta_1} \\ \dfrac{\partial f}{\partial \theta_2} \\ \dfrac{\partial f}{\partial \theta_3} \end{bmatrix} = \begin{bmatrix} 3\theta_1^2 + \theta_3 \\ -4\theta_2 \\ \theta_1 \end{bmatrix}$

or $[3\theta_1^2 + \theta_3, \ -4\theta_2, \ \theta_1]$

Question.2-04

함수 $\vec{f}$가 θ에 대해 다음과 같이 주어졌을 때, Jacobian matrix $\dfrac{\partial \vec{f}(\theta)}{\partial \theta}$를 구하시오.

$$\vec{f}(\theta) = \begin{pmatrix} f_1(\theta) \\ f_2(\theta) \\ f_3(\theta) \end{pmatrix} = \begin{pmatrix} \theta^3 - 2\theta^2 + 10 \\ ln(\theta) - sin(\theta)cos(\theta) \\ e^{\theta+10} - e^{2\theta} \end{pmatrix}$$

Jacobian matrix $\dfrac{\partial \vec{f}}{\partial \vec{\theta}} = \begin{bmatrix} \dfrac{\partial f_1}{\partial \theta} \\ \dfrac{\partial f_2}{\partial \theta} \\ \vdots \\ \dfrac{\partial f_n}{\partial \theta} \end{bmatrix}$

Question.2-05

함수 $\vec{f}$가 $\vec{\theta}$에 대해 다음과 같이 주어졌을 때, Jacobian matrix $\dfrac{\partial \vec{f}(\vec{\theta})}{\partial \vec{\theta}}$를 구하시오.

$$\vec{\theta} = \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix} \qquad \vec{f}(\vec{\theta}) = \begin{pmatrix} f_1(\vec{\theta}) \\ f_2(\vec{\theta}) \\ f_3(\vec{\theta}) \end{pmatrix} = \begin{pmatrix} (\theta_1)^3 - 2(\theta_2)^2 + 10 \\ ln(\theta_1) - sin(\theta_1)cos(\theta_2) \\ e^{\theta_1+10} - e^{2\theta_2} \end{pmatrix}$$

Jacobian $\dfrac{\partial \vec{f}}{\partial \vec{\theta}} = \begin{bmatrix} \dfrac{\partial f_1}{\partial \vec{\theta}} \cdots \\ \vdots \\ \dfrac{\partial f_n}{\partial \vec{\theta}} \cdots \end{bmatrix} = n \overbrace{\begin{bmatrix} \dfrac{\partial f_1}{\partial \theta_1} & \dfrac{\partial f_1}{\partial \theta_2} & \cdots & \dfrac{\partial f_1}{\partial \theta_m} \\ \vdots & & & \\ \dfrac{\partial f_n}{\partial \theta_1} & \dfrac{\partial f_n}{\partial \theta_2} & \cdots & \dfrac{\partial f_n}{\partial \theta_m} \end{bmatrix}}^{m}$

Question.2-06

$\vec{\theta}, \vec{f}$가 다음과 같이 주어지고,

$$\vec{\theta} = \begin{pmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \end{pmatrix} \qquad \vec{f}(\vec{\theta}) = \begin{pmatrix} f_1(\vec{\theta}) \\ f_2(\vec{\theta}) \\ f_3(\vec{\theta}) \end{pmatrix}$$

⇒ paired vector function.

$\vec{\theta}$에 대한 함수 $\vec{f}$는 다음과 같을 때,

$$f_i(\vec{\theta}) = sin(\theta_i)$$

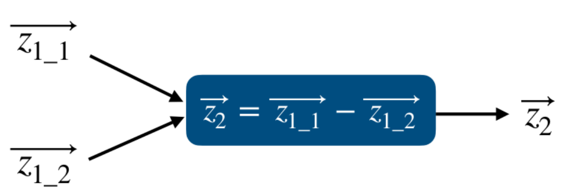
= paired function

Jacobian matrix $\dfrac{\partial \vec{f}(\vec{\theta})}{\partial \vec{\theta}}$를 구하시오.

- $$\frac{\partial \vec{f}(\vec{\theta})}{\partial \vec{\theta}} = \begin{bmatrix} \frac{\partial \sin(\theta_1)}{\partial \theta_1} & \cancel{\frac{\partial \sin(\theta_1)}{\partial \theta_2}} & \cancel{\frac{\partial \sin(\theta_1)}{\partial \theta_3}} \\ \cancel{\frac{\partial \sin\theta_2}{\partial \theta_1}} & \frac{\partial \sin\theta_2}{\partial \theta_2} & \cancel{\frac{\partial \sin\theta_2}{\partial \theta_3}} \\ \cancel{\frac{\partial \sin\theta_3}{\partial \theta_1}} & \cancel{\frac{\partial \sin\theta_3}{\partial \theta_2}} & \frac{\partial \sin\theta_3}{\partial \theta_3} \end{bmatrix} = \begin{bmatrix} \cos\theta_1 & 0 & 0 \\ 0 & \cos\theta_2 & 0 \\ 0 & 0 & \cos\theta_3 \end{bmatrix}$$

Q.2-07.

다음 연산에서 $\dfrac{\partial \vec{z_2}}{\partial \vec{z_{1\_1}}}, \dfrac{\partial \vec{z_2}}{\partial \vec{z_{1\_2}}}$, 를 각각 구하시오.



- $$\frac{\partial \vec{z_2}}{\partial \vec{z_{1\text{-}1}}} = \text{identity matrix}$$

- $$\frac{\partial \vec{z_2}}{\partial \vec{z_{1\text{-}2}}} = -1 \cdot \text{identity matrix}$$

Question.2-08

다음 연산에서 $\frac{\partial \vec{z_2}}{\partial \vec{z_1}}$를 구하시오.

$$\vec{z_1} \longrightarrow \boxed{\vec{z_2} = \vec{z_1} \bigcirc \vec{z_1}} \longrightarrow \vec{z_2}$$

이때 $\bigcirc$은 Hadamard product을 의미하며,

$$\vec{z_1} \bigcirc \vec{z_1} = \begin{pmatrix} z_1^{(1)} * z_1^{(1)} \\ z_1^{(2)} * z_1^{(2)} \\ \vdots \\ z_1^{(n)} * z_1^{(n)} \end{pmatrix}$$ 로 계산된다.

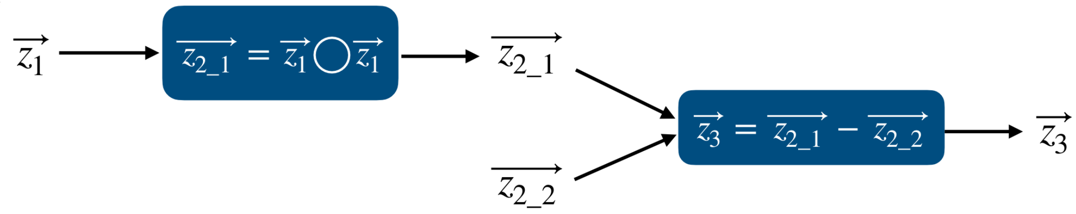
- $\frac{\partial \vec{z_2}}{\partial \vec{z_1}}$ = Jacobians for Hadamard product node.

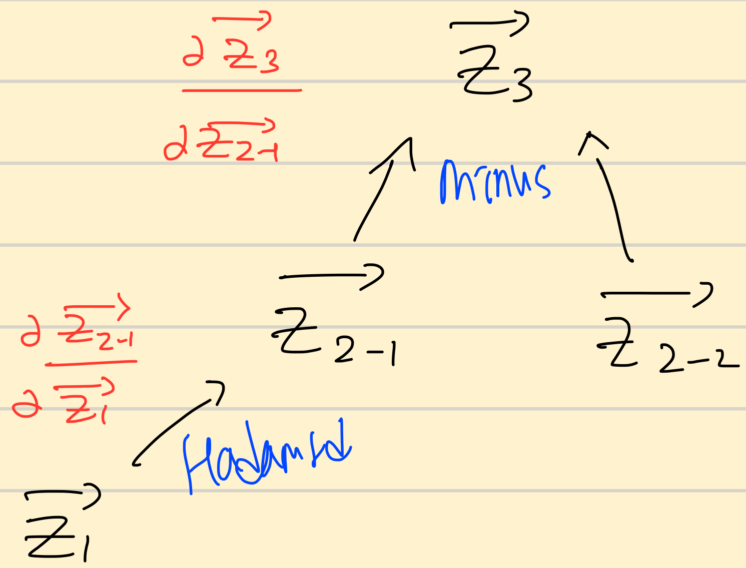
$$= \begin{pmatrix} \frac{\partial (z_1^{(1)} \times z_1^{(1)})}{\partial z_1^{(1)}} & \frac{\partial (z_1^{(1)} \times z_1^{(1)})}{\partial z_1^{(2)}} & \cdots \\ & & \\ & & \end{pmatrix}$$

$$= \begin{pmatrix} 2 \cdot z_1^{(1)} & 0 & \cdots & 0 \\ 0 & 2 \cdot z_1^{(2)} & 0 & 0 \\ \vdots & 0 & \ddots & 0 \\ 0 & 0 & 0 & 2 \cdot z_1^{(n)} \end{pmatrix}$$

Question.2-09

다음 연산에서 $\frac{\partial \vec{z_3}}{\partial \vec{z_1}}$를 구하시오.





$$\frac{\partial \vec{z_3}}{\partial \vec{z_1}} = \frac{\partial \vec{z_3}}{\partial \vec{z_{2-1}}} \times \frac{\partial \vec{z_{2-1}}}{\partial \vec{z_1}}$$

$$* \ \frac{\partial \vec{z_{2-1}}}{\partial \vec{z_1}} = \begin{pmatrix} 2 \cdot z_1^{(1)} & & \\ & \ddots & \\ & & 2 \cdot z_1^{(n)} \end{pmatrix}$$

$$* \ \frac{\partial \vec{z_3}}{\partial \vec{z_{2-1}}} = \text{identity matrix}$$

$$\therefore \ \frac{\partial \vec{z_3}}{\partial \vec{z_1}} = \begin{pmatrix} 2 \cdot z_1^{(1)} & \cdots & 0 \\ 0 & \ddots & \vdots \\ 0 & \cdots & 2 \cdot z_1^{(n)} \end{pmatrix}$$

Question.2-10

$\alpha, \vec{\beta}, \vec{\gamma}, \vec{\delta}$가 다음과 같이 주어졌다.

$$\vec{\beta} = \begin{pmatrix} \beta_1(\vec{\gamma}) \\ \beta_2(\vec{\gamma}) \\ \beta_3(\vec{\gamma}) \end{pmatrix} \qquad \vec{\gamma} = \begin{pmatrix} r_1(\vec{\delta}) \\ r_2(\vec{\delta}) \end{pmatrix} \qquad \underset{\theta}{\vec{\delta}} = \begin{pmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \end{pmatrix}$$

$$\alpha = \frac{1}{3}\sum_{i=1}^{3}\beta_i \qquad \vec{\beta}(\vec{\gamma}) = \begin{pmatrix} \beta_1(\vec{\gamma}) \\ \beta_2(\vec{\gamma}) \\ \beta_3(\vec{\gamma}) \end{pmatrix} = \begin{pmatrix} (r_1)^2 + 2r_2 \\ 2(r_1)^2 - 4(r_2)^2 \\ r_1 + 3(r_2)^2 \end{pmatrix} \qquad \vec{\gamma}(\vec{\delta}) = \begin{pmatrix} \gamma_1(\vec{\delta}) \\ \gamma_2(\vec{\delta}) \end{pmatrix} = \begin{pmatrix} sin(\delta_1) + cos(\delta_2) + tan(\delta_3) \\ e^{\delta_1} - e^{2\delta_2} + ln(\delta_3) \end{pmatrix}$$

이때 $\dfrac{\partial \alpha}{\partial \vec{\delta}}$를 구하시오.

$$\vec{\theta} \to \vec{r} \to \vec{\beta} \to \alpha$$

gradient

$$\underset{(1\times 3)}{\frac{\partial \alpha}{\partial \vec{\theta}}} = \underset{(1\times 3)}{\frac{\partial \alpha}{\partial \vec{\beta}}} \underset{(3\times 2)}{\frac{\partial \vec{\beta}}{\partial \vec{r}}} \underset{(2\times 3)}{\frac{\partial \vec{r}}{\partial \vec{\theta}}}$$

$$* \ \frac{\partial \vec{r}}{\partial \vec{\theta}} = \begin{pmatrix} \frac{\partial r_1}{\partial \theta_1} & \frac{\partial r_1}{\partial \theta_2} & \frac{\partial r_1}{\partial \theta_3} \\ \frac{\partial r_2}{\partial \theta_1} & \cdots & \end{pmatrix}$$

$$= \begin{pmatrix} \cos\theta_1 & -\sin\theta_2 & \frac{1}{\cos\theta_3{}^2} \\ e^{\theta_1} & -2e^{2\theta_2} & \frac{1}{\theta_3} \end{pmatrix}$$

$$* \ \frac{\partial \vec{\beta}}{\partial \vec{r}} = \begin{pmatrix} 2r_1 & 2 \\ 4r_1 & -8r_2 \\ 1 & 6r_2 \end{pmatrix} \qquad * \ \frac{\partial \alpha}{\partial \vec{\beta}} = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}$$

gradient

$$\underset{(1\times 3)}{\frac{\partial \alpha}{\partial \vec{\theta}}} = \underset{(1\times 3)}{\frac{\partial \alpha}{\partial \vec{\beta}}} \underset{(3\times 2)}{\frac{\partial \vec{\beta}}{\partial \vec{r}}} \underset{(2\times 3)}{\frac{\partial \vec{r}}{\partial \vec{\theta}}} = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix} \begin{pmatrix} 2r_1 & 2 \\ 4r_1 & -8r_2 \\ 1 & 6r_2 \end{pmatrix} \begin{pmatrix} \cos\theta_1 & -\sin\theta_2 & \frac{1}{\cos\theta_3{}^2} \\ e^{\theta_1} & -2e^{2\theta_2} & \frac{1}{\theta_3} \end{pmatrix}$$