تمرین تئوری سری سوم

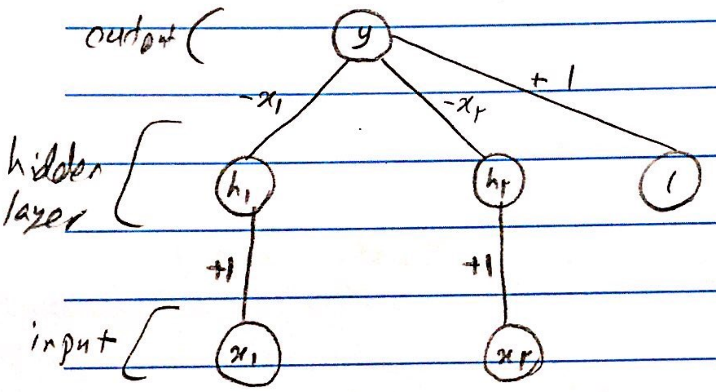
98107077 سید [illegible]

**سوال ۱:**

$\underline{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{R}^2$

در ناحیه آبی رنگ (+1) : $|x_1| + |x_2| \leq 1 \longleftarrow -|x_1| - |x_2| + 1 \geq 0$

در باقی نواحی : (-1) : $|x_1| + |x_2| > 1 \longleftarrow -|x_1| - |x_2| + 1 < 0$




$h_1 = \text{sign}(x_1)$

$h_2 = \text{sign}(x_2)$

$y = h_1(-x_1) + h_2(-x_2) + 1$

$y = -x_1\,\text{sign}(x_1) - x_2\,\text{sign}(x_2) + 1$

$y = -|x_1| - |x_2| + 1$

$$\text{sign}(y) \begin{cases} \to +1 & (-|x_1| - |x_2| + 1 > 0) \\ \to -1 & (-|x_1| - |x_2| + 1 < 0) \end{cases}$$

Year:............ Month:............ Day:............

سوال 2 (مشتق برداری)

$f = f_2 \circ f_1 \qquad f : \mathbb{R}^2 \to \mathbb{R}^2$

$$\begin{cases} f_1 : \mathbb{R}^2 \to \mathbb{R}^3 & \underline{x}_2 = f_1(\underline{x}) \\ f_2 : \mathbb{R}^3 \to \mathbb{R}^2 & \underbrace{\underline{o}}_{\text{output}} = \underline{x}_3 = f_2(\underline{x}_2) \end{cases}$$

$$f_1\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} \frac{1}{\pi} x_1 \sin(\pi x_2) \\ e^{x_1 - 1} x_2^2 \\ x_1 x_2 \end{bmatrix}$$

$$f_2\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} x_1 + x_2 + x_3 \\ x_1^2 + x_2^2 + x_3^2 \end{bmatrix}$$

$$\frac{\partial \underline{f}}{\partial \underline{x}} = \frac{\partial \underline{o}}{\partial \underline{x}} = \frac{\partial \underline{o}}{\partial \underline{x}_2} \cdot \frac{\partial \underline{x}_2}{\partial \underline{x}_1} = J_f(\underline{x}) = J_{f_2}(\underline{x}_2)\, J_{f_1}(\underline{x}) \in \mathbb{R}^{2\times 2}$$

FMD

می‌خواهیم $J_f(\underline{x}_q)$ را بیابیم به ازای: $\underline{x}_q = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$

$\underline{x}_1 = \underline{x}_q = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \qquad v_1 = e_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \qquad v_2 = e_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

$$\underset{3\times 2}{J_{f_1}}\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} \frac{\sin(\pi x_2)}{\pi} & x_1 \cos(\pi x_2) \\ e^{x_1 - 1} x_2^2 & 2 e^{x_1 - 1} x_2 \\ x_2 & x_1 \end{bmatrix}$$

$$\underset{2\times 3}{J_{f_2}}\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} 1 & 1 & 1 \\ 2x_1 & 2x_2 & 2x_3 \end{bmatrix}$$

Scanned with CamScanner

Year:.............. Month:.............. Day:..............

ادامه سوال 2

$k=1 \qquad x_1 = x_2$

$$x_2 = f_1(x_1) = f_1\left(\begin{bmatrix}1\\2\end{bmatrix}\right) = \begin{bmatrix}0\\4\\2\end{bmatrix} \qquad J_{f_1}\left(\begin{bmatrix}1\\2\end{bmatrix}\right) = \begin{bmatrix}0 & 1\\4 & 4\\2 & 1\end{bmatrix}$$

$$v_1^{new} = J_{f_1}\left(\begin{bmatrix}1\\2\end{bmatrix}\right) v_1^{old} = \begin{bmatrix}0 & 1\\4 & 4\\2 & 1\end{bmatrix}\begin{bmatrix}1\\0\end{bmatrix} = \begin{bmatrix}0\\4\\2\end{bmatrix}$$

$$v_2^{new} = J_{f_1}\left(\begin{bmatrix}1\\2\end{bmatrix}\right) v_2^{old} = \begin{bmatrix}0 & 1\\4 & 4\\2 & 1\end{bmatrix}\begin{bmatrix}0\\1\end{bmatrix} = \begin{bmatrix}1\\4\\1\end{bmatrix}$$

$k=2$:

$$x_3 = f_2(x_2) = f_2\left(\begin{bmatrix}0\\4\\2\end{bmatrix}\right) = \begin{bmatrix}6\\20\end{bmatrix} \qquad J_{f_2}\left(\begin{bmatrix}0\\4\\2\end{bmatrix}\right) = \begin{bmatrix}1 & 1 & 1\\0 & 8 & 4\end{bmatrix}$$

$$v_1^{new} = J_{f_2}\left(\begin{bmatrix}0\\4\\2\end{bmatrix}\right) v_1^{old} = \begin{bmatrix}1 & 1 & 1\\0 & 8 & 4\end{bmatrix}\begin{bmatrix}0\\4\\2\end{bmatrix} = \begin{bmatrix}6\\40\end{bmatrix}$$

$$v_2^{new} = J_{f_2}\left(\begin{bmatrix}0\\4\\2\end{bmatrix}\right) v_2^{old} = \begin{bmatrix}1 & 1 & 1\\0 & 8 & 4\end{bmatrix}\begin{bmatrix}1\\4\\1\end{bmatrix} = \begin{bmatrix}6\\36\end{bmatrix}$$

$$\frac{\partial f}{\partial x} = J_f(x_1) = \begin{bmatrix}6 & 6\\40 & 36\end{bmatrix}$$

Year:.............. Month:.............. Day:..............

سوال 3 (BP)

$x \longrightarrow z_1 = W_1 x \longrightarrow \underline{y}_1 = f(z_1) \longrightarrow z_2 = \underline{w}_2^T \underline{y}_1 \longrightarrow y = f(z_2)$

$$\frac{\partial y}{\partial \underline{w}_2} = \frac{\partial y}{\partial z_2} \cdot \frac{\partial z_2}{\partial \underline{w}_2} = f'(z_2) \cdot \frac{\partial \overbrace{\underline{w}_2^T \underline{y}_1}^{\underline{y}_1}}{\partial \underline{w}_2} = f'(z_2)\, \underline{y}_1$$

$$= f'(z_2) f(\underline{W}_1 \underline{x})$$

$$\frac{\partial y}{\partial W_1} = \underbrace{\frac{\partial y}{\partial z_2}}_{f'(z_2)} \cdot \underbrace{\frac{\partial z_2}{\partial \underline{y}_1}}_{\underline{w}_2} \cdot \overbrace{\frac{\partial \underline{y}_1}{\partial z_1}}^{1} \cdot \overbrace{\frac{\partial z_1}{\partial W_1}}^{2}$$

$$\frac{\partial \underline{w}_2^T \underline{y}_1}{\partial \underline{y}_1} = \underline{w}_2$$

1 $\frac{\partial \underline{y}_1}{\partial \underline{z}_1}$ : $z_1 \in R^{m_1}$ $\quad z_1 \rightarrow f \xrightarrow{} \underline{y}_1, \ \underline{y}_1 \in R^{m_1} \Rightarrow \frac{\partial \underline{y}_1}{\partial \underline{z}_1} \in R^{m_1 \times m_1}$ $\quad y_i = f(z_i)$

elementwise

$$J = \begin{bmatrix} \frac{\partial f(z_1)}{\partial \underline{z}_1} & \frac{\partial f(z_1)}{\partial \underline{z}_2} & \cdots & \frac{\partial f(z_n)}{\partial \underline{z}_{m_1}} \\ \frac{\partial f(z_2)}{\partial \underline{z}_1} & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ \frac{\partial f(z_{m_1})}{\partial \underline{z}_1} & \cdots & \cdots & \frac{\partial f(z_{m_1})}{\partial \underline{z}_{m_1}} \end{bmatrix} = \begin{bmatrix} f'(z_1) & 0 & 0 \cdots & 0 \\ 0 & f'(z_2) & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & \cdots & f'(z_{m_1}) \end{bmatrix}_{m_1 \times m_1} = diag\big(f'(z_i)\big)_{(i=1 \ldots m_1)}$$

Year:............ Month:............ Day:............

② : $\dfrac{\partial z_1}{\partial W_1} = \dfrac{\partial W_1 x}{\partial W_1}$

$W_1$ را وکتورایز می‌کنیم و در نهایت مشتق بردار $z_1$ نسبت به بردار $vec(W_1)$ را محاسبه می‌کنیم.

$z_1 \in R^{m_1\times 1} \quad W_1 \in R^{m_1\times n} \quad vec(W_1) \in R^{(m_1\times n)\times 1} \quad \underline{x} \in R^{n} \quad \underline{w_2} \in R^{m_1}$

ماتریس J بدین صورت خواهد بود:

$$J = \begin{bmatrix} \dfrac{\partial z_1}{\partial w_{11}} & & \dfrac{\partial z_1}{\partial w_{m_1 n}} \\ \dfrac{\partial z_2}{\partial w_{11}} & & \vdots \\ \vdots & & \vdots \\ \dfrac{\partial z_{m_1}}{\partial w_{11}} & & \dfrac{\partial z_{m_1}}{\partial w_{m_1 n}} \end{bmatrix}_{m_1\times(m_1\times n)}$$

ابتدا یک ستون از ماتریس J را بررسی می‌کنیم

ستون اول به طور مثال:

$$\begin{bmatrix} \dfrac{\partial z_1}{\partial w_{11}} \\ \vdots \\ \dfrac{\partial z_m}{\partial w_{11}} \end{bmatrix}$$

به طور کلی عبارت مقابل را محاسبه می‌کنیم:

$$\frac{\partial \underline{z}_1}{\partial w_{ij}} = \begin{bmatrix} \dfrac{\partial z_1}{\partial w_{ij}} \\ \dfrac{\partial z_2}{\partial w_{ij}} \\ \vdots \\ \dfrac{\partial z_{m_1}}{\partial w_{ij}} \end{bmatrix}$$

مولفه i ام بردار $z_1$

$$z_k = \sum_{l=1}^{n} w_{kl} x_l \Rightarrow \frac{\partial z_k}{\partial w_{ij}} = \sum_{l=1}^{n} x_l \frac{\partial w_{kl}}{\partial w_{ij}} = \sum_{l=1}^{n} x_l \, \mathbb{I}(i=k , j=l)$$

$$\frac{\partial \underline{z}_1}{\partial w_{ij}} = x_j \underline{e}_i = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ x_j \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}_{m_1\times 1}$$

مولفه i ام

به طور مثال ستون اول ماتریس J به صورت $\begin{bmatrix} x_1 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}_{m_1\times 1}$ است.

$$\frac{\partial y}{\partial W_1} = \underbrace{\overbrace{f'(z_2)}^{\text{scalar}}}_{\frac{\partial y}{\partial z_2}} \; \underbrace{\overbrace{\underline{w}_2^T}^{1\times m_1}}_{\frac{\partial z_2}{\partial \underline{y}_1}} \cdot \underbrace{\overbrace{diag(f'(z_1))}^{m_1\times m_1}}_{\frac{\partial \underline{y}_1}{\partial \underline{z}_1}} \cdot \underbrace{\overbrace{J}^{m_1\times(m_1\times n)}}_{\frac{\partial \underline{z}_1}{\partial W_1}} \Longleftarrow ①,②$$

Scanned with CamScanner

Year:.............. Month:.............. Day:..............

(سوال 4)

$$\underline{\phi}(x) = \begin{bmatrix} e^{-\frac{x^2}{2}} \\ \frac{1}{\sqrt{1}} x^1 e^{-\frac{x^2}{2}} \\ \frac{1}{\sqrt{2!}} x^2 e^{-\frac{x^2}{2}} \\ \vdots \\ \frac{1}{\sqrt{\infty!}} x^\infty e^{-\frac{x^2}{2}} \end{bmatrix} \qquad \underline{\phi}(x') = \begin{bmatrix} e^{-\frac{x'^2}{2}} \\ \frac{1}{\sqrt{1}} x'^1 e^{-\frac{x'^2}{2}} \\ \frac{1}{\sqrt{2!}} x'^2 e^{-\frac{x'^2}{2}} \\ \vdots \\ \frac{1}{\sqrt{\infty!}} x'^\infty e^{-\frac{x'^2}{2}} \end{bmatrix}$$

$$K(x, x') = \langle \underline{\phi}(x), \underline{\phi}(x') \rangle = \underline{\phi}(x)^T \underline{\phi}(x')$$

$$= e^{\frac{-x^2-x'^2}{2}} + \frac{1}{\sqrt{1}} \frac{1}{\sqrt{1}} x^1 x'^1 e^{\frac{-x^2-x'^2}{2}} + \frac{1}{\sqrt{2!}} \frac{1}{\sqrt{2!}} x^2 x'^2 e^{\frac{-x^2-x'^2}{2}} + \cdots + \frac{1}{\sqrt{\infty!}} \frac{1}{\sqrt{\infty!}} x^\infty x'^\infty e^{\frac{-x^2-x'^2}{2}}$$

$$= e^{-\frac{x^2+x'^2}{2}} \underbrace{\left[ 1 + xx' + \frac{1}{2!} x^2 x'^2 + \frac{1}{3!} x^3 x'^3 + \cdots \right]}_{\text{Taylor series expansion of } e^{xx'}}$$

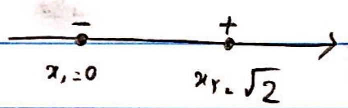
$$= e^{\frac{-x^2-x'^2}{2}} e^{xx'} = e^{\frac{-x^2-x'^2+2xx'}{2}} = \boxed{e^{-\frac{|x-x'|^2}{2}}}$$

Year:............ Month:............ Day:............

سوال 5 (مسئله SVM):

$$D = \left\{ (x_1 = 0, y_1 = -1), (x_2 = \sqrt{2}, y_2 = 1) \right\}$$

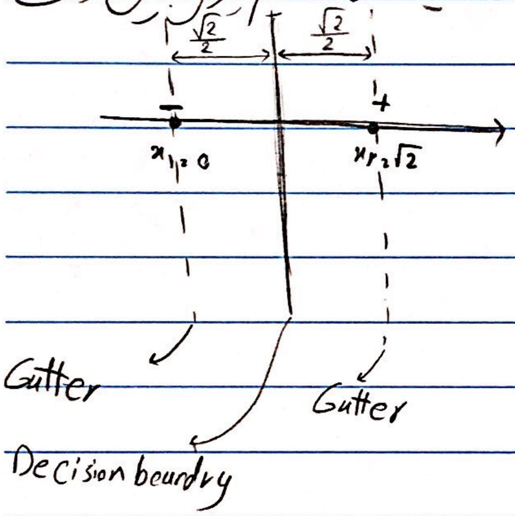
$$\underline{\varphi}(x) = \begin{bmatrix} 1 \\ \sqrt{2}\,x \\ x^2 \end{bmatrix}$$




$$\underline{\varphi}(x_1) = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

$$\underline{\varphi}(x_2) = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$$

با توجه به نوع قرارگیری داده‌ها در فضای یک بعدی در نهایت مرز تصمیم‌گیری بدین صورت خواهد بود: (اما در این مسئله به فضای ۳ بعدی می‌رویم)




الف: در فضای سه بعدی مرز تصمیم‌گیری یک صفحه است

بردار واصل دو دیتاپوینت ($\underline{\varphi}(x_2) - \varphi(x_1)$)، بردار عمود بر مرز تصمیم‌گیری است که با بردار $\underline{w}$ نیز موازی می‌باشد.

$$\underline{\varphi}(x_2) - \underline{\varphi}(x_1) = \begin{bmatrix} 0 \\ 2 \\ 2 \end{bmatrix}$$

ب) فاصله دو نقطه برابر $\|\underline{\varphi}(x_2) - \underline{\varphi}(x_1)\| = 2\sqrt{2}$ می‌باشد

پس فاصله هر نقطه تا مرز تصمیم‌گیری $\sqrt{2}$ می‌باشد.

ج:

$$\text{margin} = \frac{1}{\|w\|} = \sqrt{2} \rightarrow \|w\| = \frac{\sqrt{2}}{2}$$

$$\underline{w} = \|w\| \frac{\underline{\varphi}(x_2) - \underline{\varphi}(x_1)}{\|\varphi(x_2) - \varphi(x_1)\|} = \frac{\sqrt{2}}{2} \times \frac{\begin{bmatrix} 0 \\ 2 \\ 2 \end{bmatrix}}{2\sqrt{2}} = \frac{1}{4}\begin{bmatrix} 0 \\ 2 \\ 2 \end{bmatrix}$$

Year:.............. Month:.............. Day:..............

د)

$$y_i\left(w^T\phi(x_i)+w_0\right)=1 \xrightarrow{\times y_i}$$

$$w^T\phi(x_i)+w_0=y_i$$ برای هر دو دیتا باید به یک مقدار $w_0$ برسیم (هر دو روی Gutter هستند.)

$$w^T\phi(x_1)+w_0=-1 \rightarrow w_0=-1-w^T\underline{\phi}(x_1)$$

$$w^T\phi(x_2)+w_0=1 \rightarrow w_0=1-w^T\underline{\phi}(x_2)$$

$$\underline{w}=\begin{bmatrix}0\\0.5\\0.5\end{bmatrix}$$

$$\Rightarrow\begin{cases}w_0=-1-(0)=-1\\ w_0=1-(1+1)=-1\end{cases}$$

ه)

$$f(x)=w_0+w^T\phi(x)$$

$$=-1+[0,0.5,0.5]^T\phi(x) \qquad \underline{\phi}(x)=\begin{bmatrix}1\\ \sqrt{2}\,x\\ x^2\end{bmatrix}$$

Scanned with CamScanner