رادمهر کرمیان ۹۸۱۰۴۵۵۶

T1)

a)

$$w^T x + b = 0 \iff \langle w^T, x \rangle + b = 0$$

اگر u و v ۲ بردار دلخواه باشند،

$$\Rightarrow w^T u + b = 0 \;,\; w^T v + b = 0 \Rightarrow w^T (u - v) = 0$$

$$\Rightarrow \langle w^T, u - v \rangle = 0 \Rightarrow |w^T| \, |u - v| \cos\theta = 0 \Rightarrow \cos\theta = 0 \Rightarrow \theta = \pi/2$$

بنابراین بردار با هر بردار از ۲ نقطه صفحه عمود است. چون زاویه $\frac{\pi}{2}$ می‌سازد.

b)

$$w^T x + b = 0 \,,\; u = x + \alpha w$$

$$\rightarrow w^T u + b = w^T (x + \alpha w) + b = w^T x + b + w^T \alpha w = 0 + \alpha w^T w$$

$$= \alpha |w|^2 \Rightarrow > 0 \checkmark$$

$\alpha > 0$

c)

x روی صفحه است:

$$w^T x + b = 0 \iff \alpha (w^T x + b) = 0 \iff \alpha w^T x + \alpha b = 0$$

$$\iff \underbrace{(\alpha w^T)}_{w^{T'}} x + \underbrace{(\alpha b)}_{b'} = 0 \quad w' x + b' = 0$$

۲

الف

$$\alpha>0 \; ; \; \begin{cases} w^Tx+b>0 \Rightarrow \alpha(w^Tx+b)>0 \Rightarrow w'^Tx+b'>0 \\ w^Tx+b<0 \Rightarrow \alpha(w^Tx+b)<0 \Rightarrow w'^Tx+b'<0 \end{cases}$$

$$\alpha<0 \; \begin{cases} w^Tx+b>0 \Rightarrow \alpha(w^Tx+b)<0 \Rightarrow w'^Tx+b'<0 \\ w^Tx+b<0 \Rightarrow \alpha(w^Tx+b)>0 \Rightarrow w'^Tx+b'>0 \end{cases}$$

به ازای هر $\alpha\neq0$ ⟸ اگر نقطه‌ای در $w^Tx+b=0$ صدق کند در $w'^Tx+b'=0$ نیز صدق می‌کند.

اگر $\alpha>0$ ⟸ نیم‌فضای بالا در نیم‌فضای بالای ماند

اگر $\alpha<0$ ⟸ جای نیم‌فضاها عوض می‌شود.

⟸ معادله‌ی hyperplane یکسان است.

d_1

هر نقطه از فضا را می‌توان : $u=x+\alpha w$ بیان کرد

$w^Tx+b=0$ ⟸ چون w بر صفحه عمود است ⟸ x تصویر u بر صفحه ⟸ $D(u,\text{hyperplane})=|\alpha w|$

$$\Rightarrow w^Tx+b=0 \Rightarrow w^Tu+b = w^T(x+\alpha w)+b = w^Tx+\alpha w^Tw+b = 0+\alpha|w|^2 = \alpha|w|^2 \Rightarrow |\alpha w|\times|w| = |\langle w,u\rangle+b|$$

$$\Rightarrow |\alpha w| = \frac{|\langle w,u\rangle+b|}{|w|}$$

T_2/

$$S=\{(۰,۱),(۱,۰),(۲,۴)\}$$

$$L_S(h)=\frac{1}{m}\sum_{i=1}^{m}(h(x_i)-y_i)^۲=\frac{1}{m}\sum(a_0+a_1x_i+a_2x_i^۲-y_i)^2$$

$$=\frac{1}{3}\left((a_0-1)^۲+(a_0+a_1+a_2-۰)^۲+(a_0+۲a_1+۴a_2-۴)^۲\right)$$

$$\rightarrow \frac{\partial L}{\partial a_0}=\frac{1}{۳}\left(۲(a_0-1)+۲(a_0+a_1+a_2)+۲(a_0+۲a_1+۴a_2-۴)\right)$$

$$=۲a_0+۲a_1+۱۰/۳\,a_2-۱۰/۳$$

$$\rightarrow \frac{\partial L}{\partial a_1}=\frac{1}{۳}\left(۲(a_0+a_1+a_2)+۲\times۲\times(a_0+۲a_1+۴a_2-۴)\right)$$

$$=۲a_0+۱۰/۳\,a_1+۶a_2-۱۶/۳$$

$$\rightarrow \frac{\partial L}{\partial a_2}=\frac{1}{۳}\left(۰+۲(a_0+a_1+a_2)+۲\times۴\times(a_0+۲a_1+۴a_2-۴)\right)$$

$$=۱۰/۳\,a_0+۶a_1+۳۴/۳\,a_2-۳۲/۳$$

$$\rightarrow \left.\frac{\partial L}{\partial a_i}\right|_{i\in\{۱,۲,۳\}}=۰ \xrightarrow{\text{دستگاه}} \begin{pmatrix}a_0\\a_1\\a_2\end{pmatrix}=\begin{pmatrix}۱\\-۲٫۵\\۲٫۵\end{pmatrix}$$

$$X = (\hat{x}_۱, \hat{x}_۲, \hat{x}_۳) = \begin{pmatrix} ۱ & ۱ & ۱ \\ ۰ & ۱ & ۲ \\ ۰ & ۱ & ۴ \end{pmatrix}$$

$$\to A = XX^T = \begin{pmatrix} ۱ & ۱ & ۱ \\ ۰ & ۱ & ۲ \\ ۰ & ۱ & ۴ \end{pmatrix} * \begin{pmatrix} ۱ & ۰ & ۰ \\ ۱ & ۱ & ۱ \\ ۱ & ۲ & ۴ \end{pmatrix} = \begin{pmatrix} ۳ & ۳ & ۵ \\ ۳ & ۵ & ۹ \\ ۵ & ۹ & ۱۷ \end{pmatrix}$$

$$\to A^{-1} = \begin{pmatrix} ۱ & -۱,۵ & ۰,۵ \\ -۳,۵ & ۶,۵ & -۳ \\ ۲,۵ & -۳ & ۱,۵ \end{pmatrix}, \quad B = Xy = \begin{pmatrix} ۱ & ۱ & ۱ \\ ۰ & ۱ & ۲ \\ ۰ & ۱ & ۴ \end{pmatrix} \begin{pmatrix} ۱ \\ ۰ \\ ۴ \end{pmatrix} = \begin{pmatrix} ۵ \\ ۸ \\ ۱۶ \end{pmatrix}$$

$$\to w = A^{-1}B = \begin{pmatrix} ۱ \\ -۳,۵ \\ ۲,۵ \end{pmatrix} \checkmark$$

T۳ //

$$\to x = \begin{pmatrix} x_۱ \\ x_۲ \end{pmatrix}, \quad x^T = (۱, x^۰, x^۱, x^۲, x^۳, x^۴, x^۵)$$

$$\to w^T = (a^۰, a^۱, a^۲, a^۳, a^۴, a^۵) \Rightarrow h(x) = \langle x, w \rangle$$

$$= (a^۰ + ax + a^۲x^۲ + a^۳x^۳ + a^۴x^۴ + a^۵x^۵)$$

$$\to x(x_۱)^T = (۱,۰,۱,۰,۱,۰) \qquad y_۱ = -۱$$

$$x(x_۲)^T = (۱,۱,۰,۱,۱,۰) \qquad y_۲ = -۱$$

$$x(x_۳)^T = (۱,۶,۶,۳۶,۳۶,۳۶) \Rightarrow \qquad y_۳ = -۱$$

$$x(x_۴)^T = (۱,۲,۱,۴,۱,۲) \qquad y_۴ = ۱$$

$$x(x_۵)^T = (۱,۱,۲,۱,۴,۲) \qquad y_۵ = ۱$$

$$x(x_۶)^T = (۱,۵,۵,۲۵,۲۵,۲۵) \qquad y_۶ = ۱$$

b)

بایدداشته‌باشیم: $y_i\langle w^T, X(x_i)\rangle > 0$

$\to$ ۱) $y_1\langle w, X(x_1)\rangle = a^0+a^2+a^4 \not> 0$

$\to$ ۲) $y_2\langle w_2, X(x_2)\rangle = a^0+a^1+a^3 \not> 0$

$\to$ ۳) $y_3\langle w, X(x_3)\rangle = a^0+4a^1+4a^2+34a^3+34a^4+34a^5 < 0$

$\to$ ۴) $y_4\langle w, X(x_4)\rangle = a^0+2a^1+a^2+4a^3+a^4+2a^5 > 0$

$\to$ ۵) $y_5\langle w, X(x_5)\rangle = a^0+a^1+2a^2+a^3+4a^4+2a^5 > 0$

$\to$ ۶) $y_6\langle w, X(x_6)\rangle = a^0+5a^1+5a^2+25a^3+25a^4+25a^5 > 0$

c)

$w^{t+1} = w^t + y_i\langle w, X(x_i)\rangle$

$\to w^{0T} = (0,0,0,0,0,0) \to y_1\langle w, X(x_1)\rangle < 0 \Rightarrow w^1 = w^0 + y_1 X(x_1) = (-1,0,-1,0,-1,0)^T$

$\to y_4\langle w, X(x_4)\rangle \not> 0 \to w^2 = w^1 + y_4\langle w, X(x_4)\rangle = (0,2,0,4,0,2)^T$

T4)

a)

~~برهان خلف: ... $y_i(\tilde w^{*T}x_i+b^*) = 1$~~

برهان خلف: فرض شود که نقطهٔ بهینه برای مسئله $y_i(w^{*T}x_i+b) \geq 1$ باشد، جواب دیگری با ... متناظر متفاوتی وجود دارد.

$\forall i \in \{1,\dots,m\}: y_i(w^{*T}x_i+b^*) \geq \alpha \geq 1$

$\to \frac{1}{\alpha} y_i(w^{*T}x_i+b^*) \geq 1 \Rightarrow \tilde w^T = \frac{w^{*T}}{\alpha}, \tilde b = \frac{b^*}{\alpha} \Rightarrow \|(\tilde w, \tilde b)\| < \|(w^*, b^*)\|$

$\Rightarrow$ ✗

$\Rightarrow \forall i \in \{1, \dots, m\} : y_i(w^{*T} x_i + b^*) = 1$

b/

$y_i(w^{*T} + b^*) \geq 1 \Rightarrow y_i(w^{*T} + b^*) > 0$

c/

$|y_i| = 1 \Rightarrow |w^{*T} x_i + b^*| \geq 1 , \quad \exists i \in \{1, \dots, m\} \ni |w^{*T} x_i + b^*| = 1$

$\Rightarrow \frac{1}{|w^*|} |w^{*T} x_i + b^*| \geq \frac{1}{|w^*|}$ and $\exists i \in \{1, \dots, m\} \ni \frac{|w^{*T} x_i + b^*|}{|w^*|} = \frac{1}{|w^*|}$

$\xrightarrow{\max} d = \min\left(\frac{1}{|w^*|} |w^{*T} x_i + b^*|\right) = \frac{1}{|w^*|}$ is maximized. ✓

d/

$\max d = \tilde{d} = \min \frac{1}{|\tilde{w}|} |\tilde{w}^T x_i + \tilde{b}| , \quad \forall i \in \{1, \dots, m\} \; y_i(\tilde{w}^T x_i + \tilde{b}) > 0$

برهان خلف فرض کنیم که $\tilde{d} > d^*$

$\rightarrow w' = \alpha \tilde{w}$

$b' = \alpha \tilde{b}$

$\alpha = \frac{1}{|\tilde{w}| \tilde{d}} > 0$

$y_i(\tilde{w}^T x_i + \tilde{b}) \geq \beta > 0 \rightarrow \beta = \min(y_i \tilde{w}^T x_i + \tilde{b})$

$\Rightarrow y_i(\tilde{w}^T x_i + \tilde{b}) \geq \beta$

$\Rightarrow y_i(\alpha \tilde{w} x_i + \alpha \tilde{b}) = y_i(w' x_i + b') \geq \alpha \beta$

$\Rightarrow \alpha \beta = \frac{\min(y_i \tilde{w}^T x_i + \tilde{b})}{|\tilde{w}| \tilde{d}} = \frac{1}{\tilde{d}} \min\left(\frac{|\tilde{w}^T x_i + \tilde{b}|}{|\tilde{w}|}\right) = \frac{\tilde{d}}{\tilde{d}} = 1$

$\rightarrow w', b'$ صدق می‌کنند در مسئله $\rightarrow \tilde{d} = \frac{1}{|w'|} > d^* \rightarrow |w^*|^2 \geq |w'|^2 \rightarrow |w^*|$ کوچکترین بردار نیست $= \times$

T۵/

درست است

$\forall\, w\in\mathbb{R}^d$ و $x\in\mathbb{R}^d$

$$\mathrm{sign}(\langle w, x\rangle) = \mathrm{sign}(\langle \eta w, x\rangle)$$

این پیش‌بینی‌های برابرند چون پیش‌بینی‌سازها

هم‌جهت هستند.