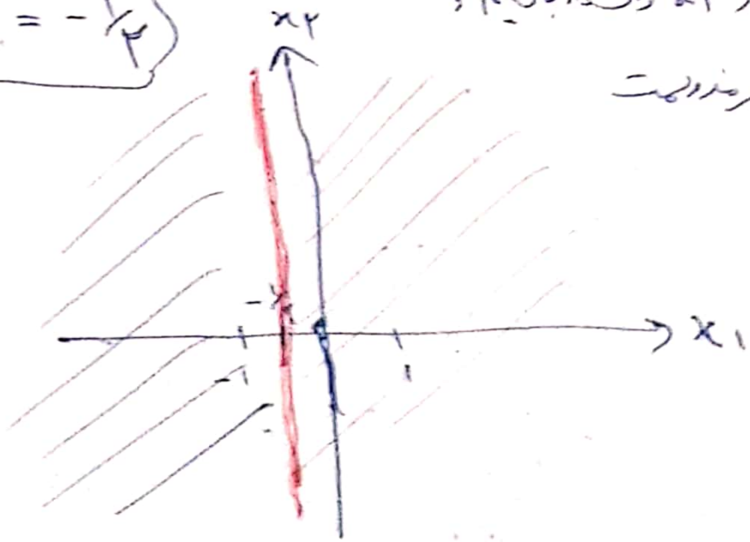
سوال ۱ - الف) مرز تصمیم زمانی است که $\theta_0 + \theta_1 x_1 + \theta_2 x_2 = 0$ باشد. با جایگذاری مقادیر داریم:

$$1 + 2x_1 = 0 \longrightarrow 2x_1 = -1 \longrightarrow x_1 = -\frac{1}{2}$$

پس اگر فضای دو بعدی x_1 و x_2 را داشته باشیم:

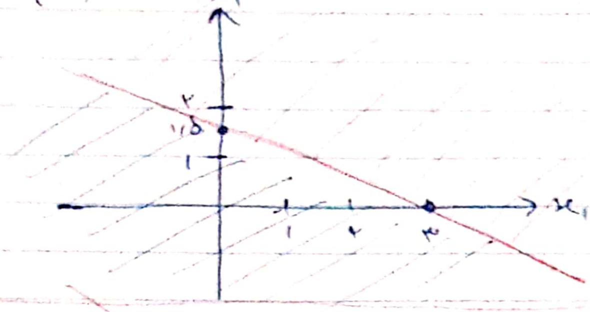


پس دو کلاس سمت راست خط قرمز و سمت
چپ خط قرمز تفکیک می‌شود

سوال ۱ - ب) :

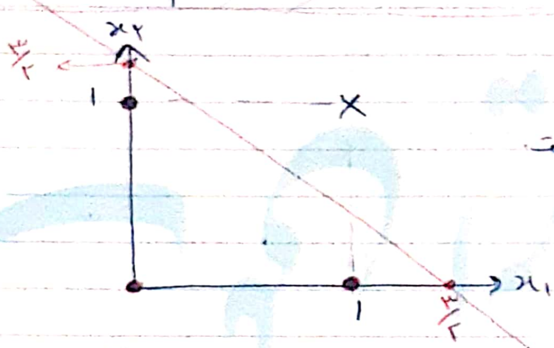
$-3 + x_1 + 2x_2 = 0 \rightarrow x_1 = -2x_2 + 3$ مرز تصمیم

$x_2 = 0 \rightarrow x_1 = 3 \rightarrow \begin{pmatrix} 3 \\ 0 \end{pmatrix}$, $x_1 = 0 \rightarrow x_2 = 3/2 \rightarrow \begin{pmatrix} 0 \\ 3/2 \end{pmatrix}$ این نقاط روی خط قرار می‌گیرند



سوال ۱ - ج) :

x_1	x_2	y
0	0	0
0	1	0
1	0	0
1	1	1



با توجه به خاصیت AND یک دسته مرزبندی شده کلاس ۱ و مابقی نقاط برای کلاس صفر است. پس هر خطی که بتواند این یک نقطه را از مابقی نقاط تفکیک کند مرز تصمیم است.

با توجه به شکل نقاط $\begin{pmatrix} 0 \\ 3/2 \end{pmatrix}$ و $\begin{pmatrix} 3/2 \\ 0 \end{pmatrix}$ می‌توانند روی خط مورد نظر قرار بگیرند.

شیب خط $= \dfrac{3/2 - 0}{0 - 3/2} = -1 \longrightarrow x_2 = -x_1 + \theta_0$

$\xrightarrow{\text{جایگذاری } \binom{0}{3/2}}$ $0 + \theta_0 = 3/2 \Rightarrow \theta_0 = 3/2$

$\rightarrow$ معادله خط تصمیم $\rightarrow x_2 = -x_1 + 3/2 \longrightarrow x_2 + x_1 - 3/2$

پس پارامترهای مدل به صورت $\theta_0 = -3/2$ و $\theta_1 = \theta_2 = 1$ می‌باشد.

$\xrightarrow{\text{جایگذاری } (0,0)} g(0 + 0 - 3/2) = g(-3/2) = 0$

$\xrightarrow{\text{جایگذاری } (0,1)} g(0 + 1 - 3/2) = g(-1/2) = 0$

$\xrightarrow{\text{جایگذاری } (1,0)} g(1 + 0 - 3/2) = g(-1/2) = 0$

$\xrightarrow{\text{جایگذاری } (1,1)} g(1 + 1 - 3/2) = g(1/2) = 1$

تمامی شرایط نیز صدق می‌کند

سوال ۲ -

$$\operatorname{softmax}(x+c)_i = \frac{e^{x_i+c}}{\sum_{j=1}^{n} e^{x_j+c}} = \frac{e^{c} \times e^{x_i}}{\sum_{j=1}^{n} e^{c} \times e^{x_j}} = \frac{\cancel{e^{c}} \times e^{x_i}}{\cancel{e^{c}} \times \sum_{j} x_j} = \frac{e^{x_i}}{\sum_{j} x_j}$$

$\Leftarrow \operatorname{softmax}(x)_i$ معادل است با

$$\operatorname{softmax}(x+c)_i = \operatorname{softmax}(x)_i$$

پس برای تک تک عنصرها این رابطه برقرار است.

در نتیجه برای کل ترینینگ ست این رابطه برقرار است.

سوال ۳-

$$\arg\max_{y} P(y) \prod_{i=1}^{n} P(x_i | y)$$

- if species = M $\rightarrow P(y=M) = \frac{۴}{۸} = \frac{۱}{۲}$, $P(\text{Color}=\text{green} | y=M) = \frac{۲}{۴} = \frac{۱}{۲}$

$P(\text{legs}=۲ | y=M) = \frac{۱}{۴}$, $P(\text{height}=\text{tall} | y=M) = \frac{۱}{۴}$, $P(\text{smelly}=\text{No} | y=M) = \frac{۱}{۴}$

$$P(y=M) \prod_{i=1}^{n} P(x_i | y=M) = \frac{۱}{۲} \times \frac{۱}{۲} \times \frac{۱}{۴} \times \frac{۱}{۴} \times \frac{۱}{۴} = \frac{۱}{۴^۴}$$

- if species = H $\rightarrow P(y=H) = \frac{۴}{۸} = \frac{۱}{۲}$, $P(\text{Color}=\text{green} | y=H) = \frac{۱}{۴}$

$P(\text{legs}=۲ | y=H) = ۱$, $P(\text{height}=\text{tall} | y=H) = \frac{۱}{۲}$, $P(\text{smelly}=\text{No} | y=H) = \frac{۳}{۴}$

$$P(y=H) \prod_{i=1}^{n} P(x_i | y=H) = \frac{۱}{۲} \times \frac{۱}{۴} \times ۱ \times \frac{۱}{۲} \times \frac{۳}{۴} = \frac{۳}{۴^۳}$$

$$P(y=H) \prod_{i=1}^{n} P(x_i | y=H) = \frac{۳}{۴^۳} > P(y=M) \prod_{i=1}^{n} P(x_i | y=M) = \frac{۱}{۴^۴}$$

پس نتیجه می‌شود که این نمونه داده به کلاس H متعلق است.

سوال ۶: ابتدا تابع lagrangian را نوشته.

$$l(w,b,\alpha) = \frac{1}{2}\|w\|^2 - \sum_{i=1}^{n} \alpha_i \left(y^i\left((x^i)^T w + b\right) - 1\right)$$

(1) $$\frac{dl}{dw} = w - \sum_{i=1}^{n} \alpha_i y_i x_i = 0 \Rightarrow w^* = \sum_{i=1}^{n} \alpha_i y_i x_i$$

حال باید نقاط موجود را در w^* جایگذاری کنیم.

(2) $$\frac{dl}{db} = \sum_{i=1}^{n} \alpha_i y_i = 0$$

$$w^* = \sum_{i=1}^{n} \alpha_i y_i x_i = \sum_{i=1}^{3} \alpha_i y_i x_i = -\alpha_1 \begin{bmatrix} 2 \\ 3 \end{bmatrix} - \alpha_2 \begin{bmatrix} 1 \\ 4 \end{bmatrix} + \alpha_3 \begin{bmatrix} 4 \\ 5 \end{bmatrix}$$

برای به دست آوردن α ها باید مسئله dual را حل کنیم و
پس نسبت به α ها مسئله dual را ماکزیمم کنیم.

جایگذاری نقطه ها در رابطه (2): $\begin{cases} -\alpha_1 - \alpha_2 + \alpha_3 = 0 \\ \alpha_3 = \alpha_1 + \alpha_2 \end{cases}$

Dual: $$l(w,b,\alpha) = \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{n} y_i y_j \alpha_i \alpha_j x_i^T x_j - \cancel{b \sum_{i=1}^{n} \alpha_i y_i}$$

s.t. $\sum_{i=1}^{n} \alpha_i y_i = 0$, $\alpha_i \geq 0$ $i = 1, 2, \dots, n$

$$\max_{\alpha_1,\alpha_2,\alpha_3} -\frac{1}{2} \sum_{i=1}^{3} \sum_{j=1}^{3} \alpha_i \alpha_j y_i y_j x_i^T x_j + \sum_{i=1}^{3} \alpha_i$$

$x_1^T x_1 = 4 + 9 = 13$, $x_1^T x_2 = 2\times1 + 3\times4 = 14$, $x_1^T x_3 = 2\times4 + 3\times5 = 23$

$x_2^T x_2 = 1\times1 + 4\times4 = 17$, $x_2^T x_3 = 1\times4 + 4\times5 = 24$, $x_3^T x_3 = 4\times4 + 5\times5 = 41$

$$J = \max -\frac{1}{2}\left[13\alpha_1^2 + 28\alpha_1\alpha_2 - 46\alpha_1\alpha_3 + 17\alpha_2^2 - 48\alpha_2\alpha_3 + 41\alpha_3^2\right] + \alpha_1 + \alpha_2 + \alpha_3$$

حال باید $\alpha_3 = \alpha_1 + \alpha_2$ را در این رابطه جایگذاری کرد.

$-46\alpha_1^2 - 46\alpha_1\alpha_2$ $\quad -48\alpha_1\alpha_2 - 48\alpha_2^2$ $\quad 41\alpha_1^2 + 41\alpha_2^2 + 82\alpha_1\alpha_2$

$$J = \max_{\alpha_1,\alpha_2} -\frac{1}{2}\left[13\alpha_1^2 + 28\alpha_1\alpha_2 - 46\alpha_1[\alpha_1+\alpha_2] + 17\alpha_2^2 - 48\alpha_2[\alpha_1+\alpha_2] + 41(\alpha_1+\alpha_2)^2\right] + 2\alpha_1 + 2\alpha_2$$

ادامه سوال ۴ الف

$$J = \max_{\alpha_1,\alpha_2} -\frac{1}{2}\left[8\alpha_1^2 + 10\alpha_2^2 + 12\alpha_1\alpha_2\right] + 2\alpha_1 + 2\alpha_2 =$$

$$J = \max_{\alpha_1,\alpha_2} -4\alpha_1^2 - 5\alpha_2^2 - 8\alpha_1\alpha_2 + 2\alpha_1 + 2\alpha_2 .$$

حال باید نسبت به α_1 و α_2 مشتق گرفته شود

$$\frac{dJ}{d\alpha_1} = -8\alpha_1 - 8\alpha_2 + 2 = 0 \longrightarrow \alpha_1 = \frac{-2+8\alpha_2}{-8} = \frac{-8\alpha_2+2}{8} = -\alpha_2 + \frac{1}{4}$$

این مقدار را در رابطه پایینی جاگذاری می‌کنیم.

$$\frac{dJ}{d\alpha_2} = -10\alpha_2 - 8\alpha_1 + 2 = 0 \longrightarrow \alpha_2 = \frac{-2+8\alpha_1}{-10} = \frac{-8\alpha_1+2}{10} =$$

$$\frac{-\cancel{8}\left[\frac{-8\alpha_2+2}{\cancel{8}}\right]+2}{10} = \frac{8\alpha_2-2+2}{10} = \frac{8\alpha_2}{10} = \alpha_2 \longrightarrow \alpha_2 = 0$$

$$\alpha_1 = -\alpha_2 + \frac{1}{4} = \frac{1}{4} \quad , \quad \alpha_3 = \alpha_2 + \alpha_1 = \frac{1}{4}$$

حال با این مقادیر به دست آمده در w^* جاگذاری شوند.

$$w^* = -\frac{1}{4}\begin{bmatrix}2\\3\end{bmatrix} + \frac{1}{4}\begin{bmatrix}4\\5\end{bmatrix} = \begin{bmatrix}-\frac{1}{2}+1\\ -\frac{3}{4}+\frac{5}{4}\end{bmatrix} = \begin{bmatrix}\frac{1}{2}\\ \frac{1}{2}\end{bmatrix}$$

حال باید b را به دست آوریم. می‌دانیم support vector ها نقاطی هستند که α آن‌ها غیر صفر شده است پس ۲ تا sup vector داریم که نقاط x_1 و x_3 هستند. پس از معادله یکی از sup vector ها می‌نویسیم.

$$\text{مارجین } m = \frac{2}{\|w\|} = \frac{2}{\sqrt{\frac{1}{4}+\frac{1}{4}}} = \frac{2}{\frac{1}{\sqrt{2}}} = 2\sqrt{2}$$

$$y^i(w^T x + b) = 1 \xrightarrow{\text{جاگذاری } x_3} 1\left(\begin{bmatrix}\frac{1}{2} & \frac{1}{2}\end{bmatrix}\begin{bmatrix}4\\5\end{bmatrix} + b\right) = 1 \rightarrow 2 + \frac{5}{2} + b = 1$$

$$\frac{9}{2} + b = 1 \rightarrow b = 1 - \frac{9}{2} = -\frac{7}{2} = -3.5$$

معادله خط هم:

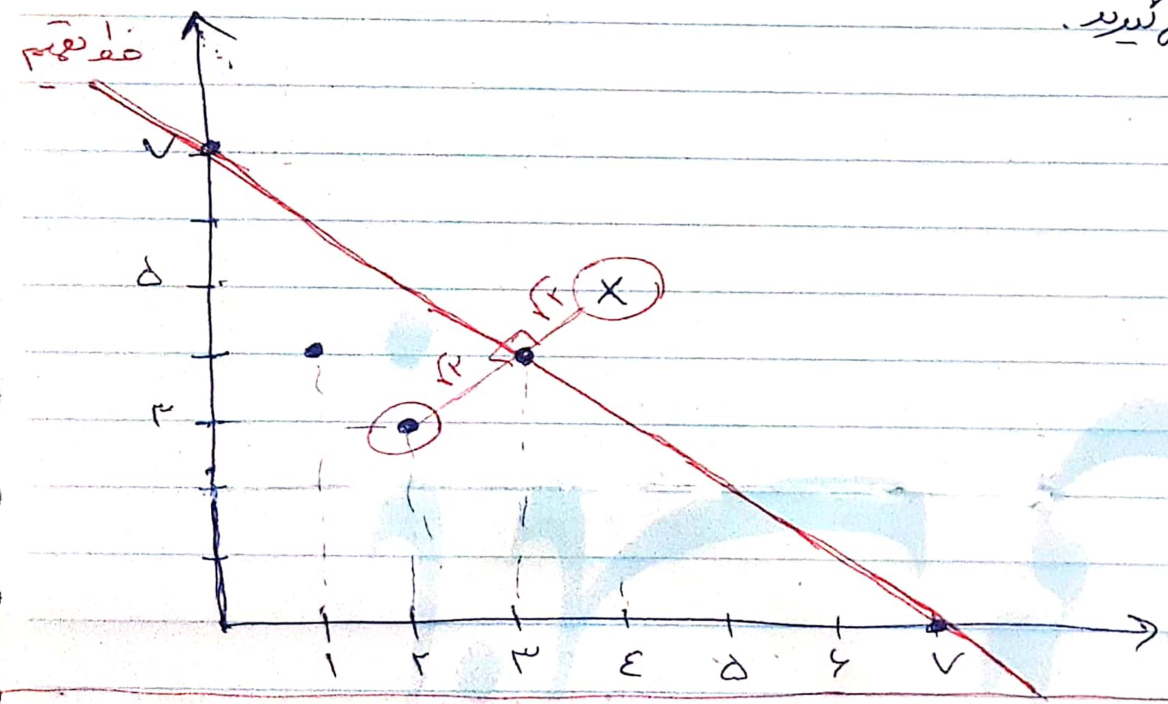
$$y = w^T x + b = \begin{pmatrix}\frac{1}{2}\\ \frac{1}{2}\end{pmatrix}^T x - 3.5$$

سوال ۴ـ ب) : باید ۲ نقطه روی خط را بدست آوریم.

$$x_1 = 0 \rightarrow \begin{pmatrix}1/2\\1/2\end{pmatrix}^T \begin{pmatrix}0\\x_2\end{pmatrix} - 3.5 = 0 \rightarrow 0 + \frac{x_2}{2} = \frac{7}{2} \rightarrow x_2 = 7$$

$$x_2 = 0 \rightarrow \begin{pmatrix}1/2\\1/2\end{pmatrix}^T \begin{pmatrix}x_1\\0\end{pmatrix} - 3.5 = 0 \rightarrow \frac{x_1}{2} - 3.5 = 0 \rightarrow \frac{x_1}{2} = \frac{7}{2} \rightarrow x_1 = 7$$

پس دو نقطه $\begin{pmatrix}0\\7\end{pmatrix}$ و $\begin{pmatrix}7\\0\end{pmatrix}$ روی این خط قرار میگیرند.



نقاطی که دور آن‌ها خط قرمز کشیده شده sup vector هستند.