تمرین شماره دوم درس یادگیری ماشین

پریا پاسه‌ورز

شماره دانشجویی: 810101393

سوال چهارم)

برای یک مجموعه داده شامل دو کلاس از داده‌های دو بعدی، داده‌های هر کلاس از یک توزیع گوسی با پارامترهای زیر تولید شده‌اند:

$$\boldsymbol{C_1}: \boldsymbol{N}(\boldsymbol{\mu_1}, \boldsymbol{\Sigma_1}) \ , \boldsymbol{\mu_1} = \begin{bmatrix} \mathbf{2} \\ \mathbf{3} \end{bmatrix} , \boldsymbol{\Sigma_1} = \begin{bmatrix} \mathbf{4} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} \end{bmatrix}$$

$$\boldsymbol{C_2}: \boldsymbol{N}(\boldsymbol{\mu_2}, \boldsymbol{\Sigma_2}) \ , \boldsymbol{\mu_2} = \begin{bmatrix} \mathbf{5} \\ \mathbf{1} \end{bmatrix} , \boldsymbol{\Sigma_2} = \begin{bmatrix} \mathbf{2} & \mathbf{0} \\ \mathbf{0} & \mathbf{3} \end{bmatrix}$$

احتمالات پیشین کلاس‌ها به صورت زیر است:

$$\boldsymbol{P(C_1) = 0.6 , P(C_2) = 0.4}$$

مرز تصمیم بیزی دو کلاس را به دست آورید.

برای پیدا کردن decision boundary، باید معادله زیر را حل کنیم:

$$P(C_1|x) = P(C_2|x)$$

$$\frac{P(x|C_1)P(C_1)}{P(x)} = \frac{P(x|C_2)P(C_2)}{P(x)}$$

$$P(x|C_1)P(C_1) = P(x|C_2)P(C_2)$$

از طرف دیگر، تابع چگالی احتمال برای توزیع گوسی به صورت زیر است:

$$P(x|C_i) = \frac{1}{(2\pi)^{d/2}|\Sigma_i|^{1/2}} \exp\left(-\frac{1}{2}(x-\mu_i)^T \Sigma_i^{-1}(x-\mu_i)\right)$$

حال برای حالت d = 2:

$$\frac{1}{|\Sigma_1|^{1/2}} \exp\left(-\frac{1}{2}(x-\mu_1)^T \Sigma_1^{-1}(x-\mu_1)\right) P(C_1) = \frac{1}{|\Sigma_2|^{1/2}} \exp\left(-\frac{1}{2}(x-\mu_2)^T \Sigma_2^{-1}(x-\mu_2)\right) P(C_2)$$

از دو طرف لگاریتم می‌گیریم:

$$\frac{1}{2}(x-\mu_1)^T\Sigma_1^{-1}(x-\mu_1)-\frac{1}{2}(x-\mu_2)^T\Sigma_2^{-1}(x-\mu_2)=\ln\big(P(C_1)\big)-\ln\big(P(C_2)\big)+\frac{1}{2}\ln(|\Sigma_2|)-\frac{1}{2}\ln(|\Sigma_1|)$$

برای ساده‌سازی:

$$Q_1(x)=(x-\mu_1)^T\Sigma_1^{-1}(x-\mu_1) \;=\; x^T\Sigma_1^{-1}x-2\mu_1^T\Sigma_1^{-1}x+\mu_1^T\Sigma_1^{-1}\mu_1$$

$$Q_2(x)=(x-\mu_2)^T\Sigma_2^{-1}(x-\mu_2)=x^T\Sigma_2^{-1}x-2\mu_2^T\Sigma_2^{-1}x+\mu_2^T\Sigma_2^{-1}\mu_2$$

$$\frac{1}{2}Q_1(x)-\frac{1}{2}Q_2(x)=ln\big(P(C_1)\big)-ln\big(P(C_2)\big)+\frac{1}{2}ln(|\Sigma_2|)-\frac{1}{2}ln(|\Sigma_1|)$$

کلاس C1:

$$Q_1(x)=\begin{bmatrix}x_1 & x_2\end{bmatrix}\begin{bmatrix}0.25 & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix}-2\begin{bmatrix}2 & 3\end{bmatrix}\begin{bmatrix}0.25 & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix}+\begin{bmatrix}2 & 3\end{bmatrix}\begin{bmatrix}0.25 & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}2\\ 3\end{bmatrix}$$

$$Q_1(x)=0.25x_1^2+x_2^2-(1x_1+6x_2)+(0.25\cdot 4+9)=0.25x_1^2+x_2^2-x_1-6x_2+10$$

کلاس C2:

$$Q_2(x)=\begin{bmatrix}x_1 & x_2\end{bmatrix}\begin{bmatrix}0.5 & 0\\ 0 & 0.333\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix}-2\begin{bmatrix}5 & 1\end{bmatrix}\begin{bmatrix}0.5 & 0\\ 0 & 0.333\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix}+\begin{bmatrix}5 & 1\end{bmatrix}\begin{bmatrix}0.5 & 0\\ 0 & 0.333\end{bmatrix}\begin{bmatrix}5\\ 1\end{bmatrix}$$

$$\begin{aligned}Q_2(x)&=0.5x_1^2+0.333x_2^2-(5x_1+0.666x_2)+(0.5\cdot 25+0.333\cdot 1)\\ &=0.5x_1^2+0.333x_2^2-5x_1-0.666x_2+12.833\end{aligned}$$

$$\ln\big(\mathrm{P}(\mathrm{C}_1)\big)-\ln\big(\mathrm{P}(\mathrm{C}_2)\big)=\ln(0.6)-\ln(0.4)\approx 0.405$$

$$\frac{1}{2}\ln(|\Sigma_2|)-\frac{1}{2}\ln(|\Sigma_1|)=\frac{1}{2}\ln(6)-\frac{1}{2}\ln(4)=\frac{1}{2}(\ln(6)-\ln(4))\approx 0.255$$

$$Q_1(x)-Q_2(x)=2(0.405+0.255)=1.32$$

$$(0.25x_1^2+x_2^2-x_1-6x_2+10)-(0.5x_1^2+0.333x_2^2-5x_1-0.666x_2+12.833)=1.32$$

$$-0.25x_1^2+0.667x_2^2+4x_1-5.334x_2-2.833=1.32$$

$$-0.25x_1^2+0.667x_2^2+4x_1-5.334x_2-4.153=0$$