دانشگاه صنعتی امیرکبیر

دانشکده مهندسی کامپیوتر

گزارش تکلیف دوم درس شناسائی آماری الگو

دانشجو:

سید احمد نقوی نوزاد

ش-د: ۹۴۱۳۱۰۶۰

استاد:

دکتر رحمتی

پائیز ۹۵

جواب سوال ۱

کدهای مربوط این سوال در فایل‌های ex01_a تا ex01_f قرار دارند.

قسمت الف)

برای رسم مرز تصمیم‌گیری بیزین با کمترین احتمال خطا[1]، می‌بایست احتمالات ثانویه‌ی[2] دو کلاس را با یکدیگر مقایسه نمائیم، که در نهایت به مرز تصمیم‌گیری به صورتی که در ادامه می‌آید دست می‌یابیم:

$$P_1 = P_2 = .5$$

$$M_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \Sigma_1 = \begin{pmatrix} 1 & .5 \\ .5 & 1 \end{pmatrix}$$

$$M_2 = \begin{pmatrix} -1 \\ 0 \end{pmatrix}, \quad \Sigma_2 = \begin{pmatrix} 1 & -.5 \\ -.5 & 1 \end{pmatrix}$$

$$p(w_1|X) \mathop{\gtrless}_{w_2}^{w_1} p(w_2|X) \Rightarrow p(X|w_1)P(w_1) \mathop{\gtrless}_{w_2}^{w_1} p(X|w_2)P(w_2) \Rightarrow \frac{p(X|w_1)}{p(X|w_2)} \mathop{\gtrless}_{w_2}^{w_1} \frac{P(w_2)}{P(w_1)}$$

$$-\ln\left(\frac{p(X|w_1)}{p(X|w_2)}\right) \mathop{\lessgtr}_{w_2}^{w_1} -\ln\left(\frac{P(w_2)}{P(w_1)}\right) \Rightarrow$$

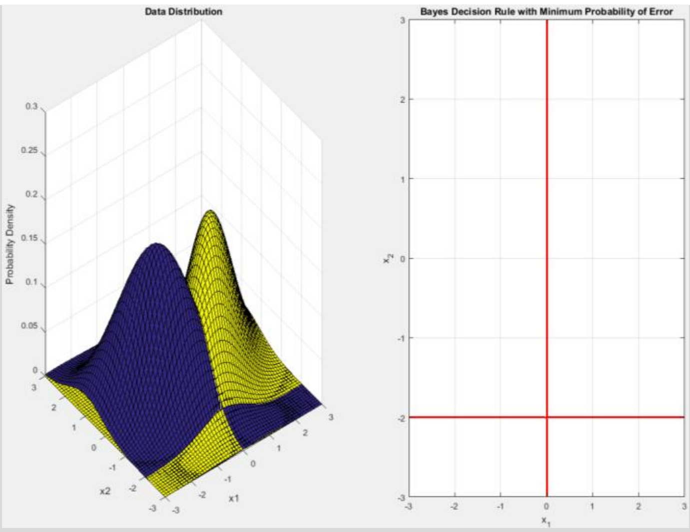
$$h(X) = \frac{1}{2}(X - M_1)^T \Sigma_1^{-1}(X - M_1) - \frac{1}{2}(X - M_2)^T \Sigma_2^{-1}(X - M_2) + \frac{1}{2}\ln\left(\frac{|\Sigma_1|}{|\Sigma_2|}\right) \mathop{\lessgtr}_{w_2}^{w_1} \ln\left(\frac{P_1}{P_2}\right)$$

Now for this question we have:

$$h(X) = \frac{1}{2}\begin{pmatrix} x_1 - 1 \\ x_2 \end{pmatrix}^T \begin{pmatrix} \frac{4}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{4}{3} \end{pmatrix}\begin{pmatrix} x_1 - 1 \\ x_2 \end{pmatrix} - \frac{1}{2}\begin{pmatrix} x_1 + 1 \\ x_2 \end{pmatrix}^T \begin{pmatrix} \frac{4}{3} & -\frac{2}{3} \\ -\frac{2}{3} & \frac{4}{3} \end{pmatrix}\begin{pmatrix} x_1 + 1 \\ x_2 \end{pmatrix} + \frac{1}{2}\ln\left(\frac{3/4}{3/4}\right) \mathop{\lessgtr}_{w_2}^{w_1} \ln\left(\frac{.5}{.5}\right)$$

$$h(X) = \frac{1}{2}\begin{pmatrix} x_1 - 1 \\ x_2 \end{pmatrix}^T \begin{pmatrix} \frac{4}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{4}{3} \end{pmatrix}\begin{pmatrix} x_1 - 1 \\ x_2 \end{pmatrix} - \frac{1}{2}\begin{pmatrix} x_1 + 1 \\ x_2 \end{pmatrix}^T \begin{pmatrix} \frac{4}{3} & -\frac{2}{3} \\ -\frac{2}{3} & \frac{4}{3} \end{pmatrix}\begin{pmatrix} x_1 + 1 \\ x_2 \end{pmatrix} \mathop{\lessgtr}_{w_2}^{w_1} 0$$

در شکل ۱، توزیع مربوط به داده‌های دو کلاس به همراه مرز تصمیم‌گیری مربوطه نمایش داده شده است.



شکل ۱. سمت چپ: توزیع داده‌های دو کلاس؛ سمت راست: مرز تصمیم‌گیری بیزین با مینیمم احتمال خطا

[1] Bayes decision boundary with minimum probability of error
[2] a posteriori

قسمت ب)

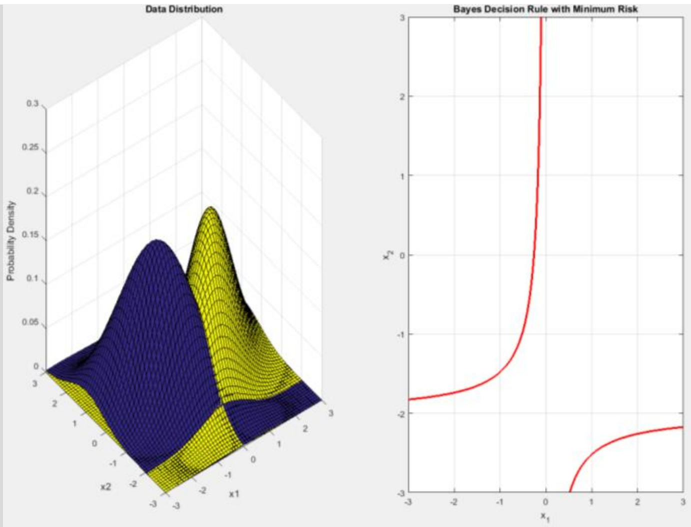
برای رسم مرز تصمیم‌گیری بیزین با کمترین میزان خطر[۳] با مقادیر توابع هزینه‌ی زیر داریم:

$$c_{11}=c_{22}=0,\quad c_{12}=2c_{21}$$

$$\frac{p(X|w_1)}{p(X|w_2)}\underset{w_2}{\overset{w_1}{><}}\frac{(c_{21}-c_{22})P(w_2)}{(c_{12}-c_{11})P(w_1)}\Rightarrow\frac{p(X|w_1)}{p(X|w_2)}\underset{w_2}{\overset{w_1}{><}}\frac{1}{2}\Rightarrow-\ln\left(\frac{p(X|w_1)}{p(X|w_2)}\right)\underset{w_2}{\overset{w_1}{<>}}-\ln\left(\frac{1}{2}\right)\Rightarrow$$

$$h(X)=\frac{1}{2}\begin{pmatrix}x_1-1\\x_2\end{pmatrix}^T\begin{pmatrix}\frac{4}{3}&\frac{2}{3}\\\frac{2}{3}&\frac{4}{3}\end{pmatrix}\begin{pmatrix}x_1-1\\x_2\end{pmatrix}-\frac{1}{2}\begin{pmatrix}x_1+1\\x_2\end{pmatrix}^T\begin{pmatrix}\frac{4}{3}&-\frac{2}{3}\\-\frac{2}{3}&\frac{4}{3}\end{pmatrix}\begin{pmatrix}x_1+1\\x_2\end{pmatrix}+\ln\left(\frac{1}{2}\right)\underset{w_2}{\overset{w_1}{<>}}0$$

در شکل ۲، توزیع مربوط به داده‌های دو کلاس به همراه مرز تصمیم‌گیری مربوطه نمایش داده شده است.



شکل ۲. سمت چپ: توزیع داده‌های دو کلاس؛ سمت راست: مرز تصمیم‌گیری بیزین با مینیمم خطر

قسمت ج)

در اینجا برای مقادیر توابع هزینه‌ی زیر داریم:

$$c_{11}=c_{22}=0,\quad c_{12}=c_{21}$$

$$\frac{p(X|w_1)}{p(X|w_2)}\underset{w_2}{\overset{w_1}{><}}\frac{(c_{21}-c_{22})P(w_2)}{(c_{12}-c_{11})P(w_1)}\Rightarrow\frac{p(X|w_1)}{p(X|w_2)}\underset{w_2}{\overset{w_1}{><}}1\Rightarrow-\ln\left(\frac{p(X|w_1)}{p(X|w_2)}\right)\underset{w_2}{\overset{w_1}{<>}}0\Rightarrow$$

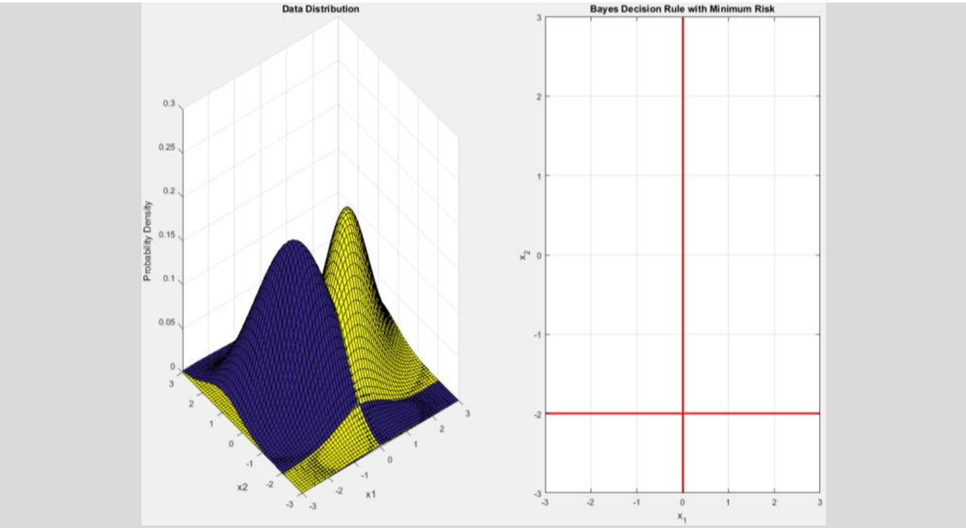
$$h(X)=\frac{1}{2}\begin{pmatrix}x_1-1\\x_2\end{pmatrix}^T\begin{pmatrix}\frac{4}{3}&\frac{2}{3}\\\frac{2}{3}&\frac{4}{3}\end{pmatrix}\begin{pmatrix}x_1-1\\x_2\end{pmatrix}-\frac{1}{2}\begin{pmatrix}x_1+1\\x_2\end{pmatrix}^T\begin{pmatrix}\frac{4}{3}&-\frac{2}{3}\\-\frac{2}{3}&\frac{4}{3}\end{pmatrix}\begin{pmatrix}x_1+1\\x_2\end{pmatrix}\underset{w_2}{\overset{w_1}{<>}}0$$

همان‌طور که مشاهده می‌شود، مرز تصمیم‌گیری حاصله با توابع هزینه‌ی فعلی برابر همان مرز تصمیم‌گیری قسمت (الف) می‌باشد. در این‌جا به دلیل اینکه $c_{12}-c_{11}=c_{21}-c_{22}$ می‌باشد، تابع هزینه‌ی مربوطه را «تابع هزینه‌ی متقارن»[۴] می‌خوانند، و در چنین حالتی تابع هزینه برابر همان احتمال خطا شده و تست مربوطه نیز احتمال خطا را مینیمم می‌سازد.

در شکل۳، توزیع مربوط به داده‌های دو کلاس به همراه مرز تصمیم‌گیری مربوطه نمایش داده شده است.

---

[3] Bayes decision boundary with minimum risk
[4] Symmetrical cost function



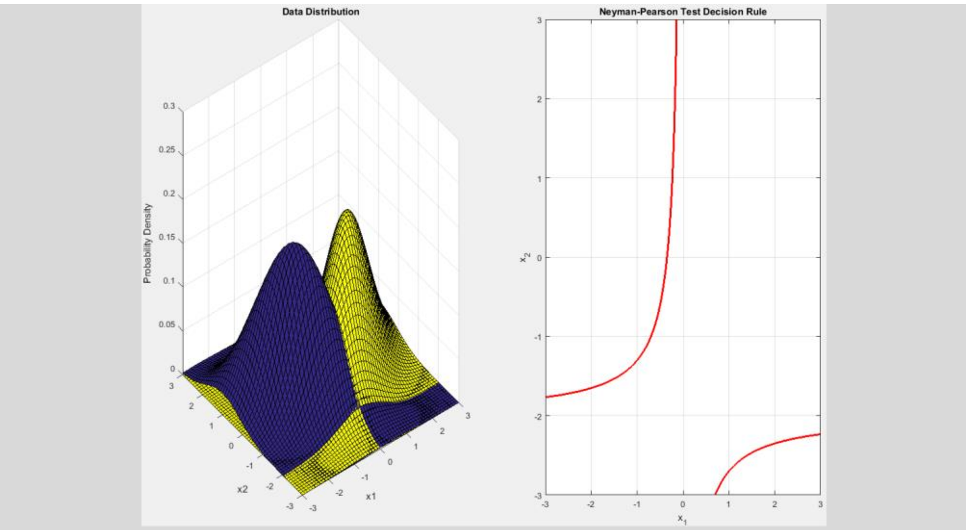
شکل ۳. سمت چپ: توزیع داده‌های دو کلاس؛ سمت راست: مرز تصمیم‌گیری بیزین با مینیمم خطر که به دلیل اینکه $c_{12}-c_{11}=c_{21}-c_{22}$ می‌باشد، برابر همان تست بیزین با مینیمم احتمال خطا می‌گردد.

قسمت د)

برای یافتن خطای تام، زمانی که تست Neyman-Pearson اعمال می‌گردد و مقدار ε_1 برابر مقدار ثابت $.05$ $\varepsilon_0 =$ فرض شده است، با توجه به معادله‌ی $r=\varepsilon_2+\mu(\varepsilon_1-\varepsilon_0)$، می‌بایست مقدار ضریب لاگرانژ μ متناسب با این مقدار خاص $.05$ $\varepsilon_1 =$ را تخمین بزنیم که برای این‌کار از یک قطعه کد متلب بهره برده و نتایج به صورت زیر می‌باشند:

The fitted lagrange multiplier Mu for Eps1=0.05 is: 0.39340
The Eps1 value of Neyman-Pearson test is: 0.05
The Eps2 value of Neyman-Pearson test is: 0.31
The total error of Neyman-Pearson test is: 0.18

در شکل۴، توزیع مربوط به داده‌های دو کلاس به همراه مرز تصمیم‌گیری مربوطه نمایش داده شده است.



شکل ۴. سمت چپ: توزیع داده‌های دو کلاس؛ سمت راست: تست **Neyman-Pearson** با مقدار ثابت ε_1 برابر $.05$ $\varepsilon_0 =$؛ مقدار بهینه‌ی ضریب لاگرانژ μ برای این مقدار ε_1 برابر **0.39340** تخمین زده شده است.

قسمت ه)

در مورد تست بیزین با مینیمم ریسک دیدیم که نرخ درست‌نمائی[5] را با یک مقدار حدآستانه[6] که تابعی از $P(\omega_i)$ بود مقایسه می‌کردیم. مشکل این قضیه آن بود که در صورتی که مقدار $P(\omega_i)$ را تغییر داده و البته مقدار حدآستانه‌ی سابق را ثابت نگه داریم، این مقدار

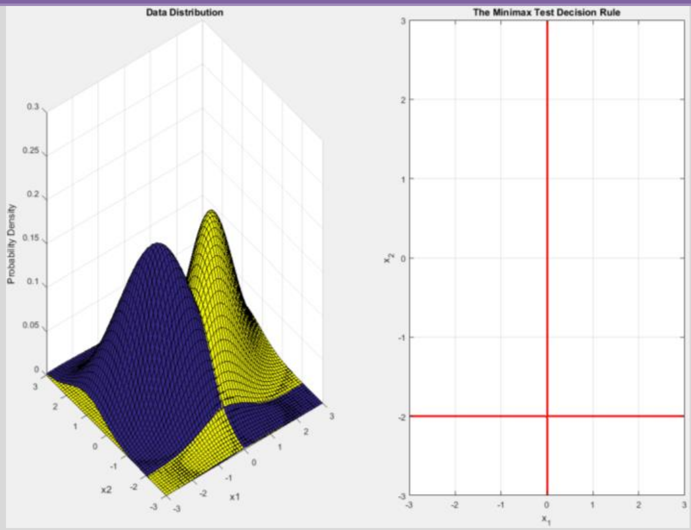
[5] Likelihood ratio
[6] Threshold value

حدآستانه، دیگر کمترین میزان خطای قابل کسب را به ما نخواهد داد. لذا در تست مینی‌مکس[7] به دنبال مقدار حدآستانه‌ای هستیم که حتی در صورت تغییر مقدار $P(\omega_i)$ باز هم بتوانیم بیشترین مقدار ممکن ریسک را مینیمم نمائیم. بدین منظور می‌بایست شرایط زیر در مورد توابع هزینه و خطاهای مربوط به دو کلاس برقرار باشند:

$$c_{11} = c_{22} = 0, \quad c_{12} = c_{21}$$
$$\int_{\Gamma_2} p(X \mid \omega_1) dX = \int_{\Gamma_1} p(X \mid \omega_2) dX$$

حال در این‌جا ما نیز با توجه به معادله‌ی مرز تصمیم‌گیری مربوط به تست بیزین با کمترین ریسک که در قسمت (ج) بدان اشاره گردید، مقدار حد آستانه را با توجه به مقادیر توابع هزینه‌ی ارائه‌شده در صورت سؤال که همان مقادیر ارائه‌شده در قسمت (ج) می‌باشند، به دست می‌آوریم. سپس بررسی می‌کنیم که آیا خطای مربوط به دو کلاس با یکدیگر برابر خواهند بود یا نه، که در صورت برابرنبودن می‌بایست مقدار $P(\omega_i)$ را مابین صفر و یک آن‌قدر تغییر دهیم تا این شرط برقرار گردد. در مورد این مسئله‌ی خاص با استفاده از یک قطعه کد متلب متوجه شدیم که به ازای مقادیر فعلی $P(\omega_i)$ها شرط برابربودن خطاهای دو کلاس برقرار بوده و بنابراین مقدار حدآستانه همان مقدار حاصله از قسمت (ج) می‌باشد و در نتیجه مرز تصمیم‌گیری حاصل از تست مینی‌مکس نیز با قسمت (ج) متفاوت نخواهد بود. داریم:

The Eps1 value of Minimax test is: 0.14238
The Eps2 value of Minimax test is: 0.14238
The total error of Minimax test is: 0.14238

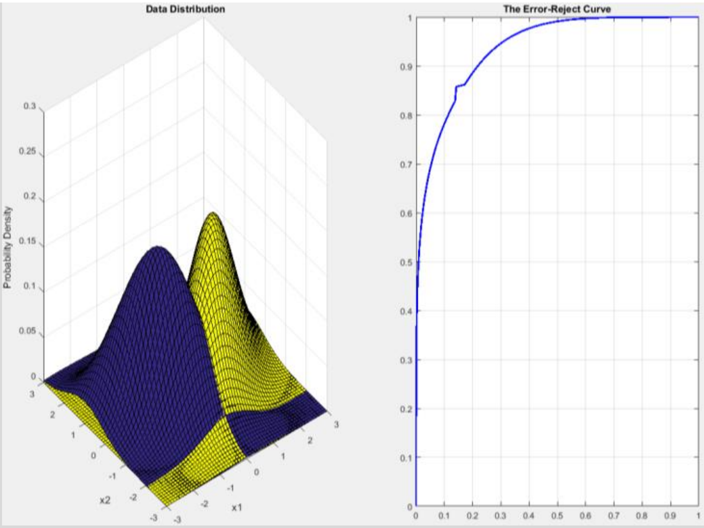


شکل ۵. سمت چپ: توزیع داده‌های دو کلاس؛ سمت راست: مرز تصمیم‌گیری حاصل از تست مینی‌مکس.

قسمت و)

جهت رسم منحنی error-reject که همان منحنی ROC[8] می‌باشد، می‌بایست رابطه‌ی بین خطاهای مربوط به دو کلاس که همان ε_1 و ε_2 می‌باشند را ضمن تغییر مقدار حدآستانه به صورت مداوم رسم نمائیم. جهت تغییر مقدار حدآستانه نیز به دلیل این‌که مقدار آن به مقدار $P(\omega_i)$ وابسته می‌باشد، با تغییر $P(\omega_i)$ مابین صفر و یک، مقادیر خاص ε_1 و ε_2 را به دست آورده و در نهایت نمودار مربوطه را رسم می‌نمائیم. نتایج حاصل از کد متلب مربوط به این قسمت به صورت زیر می‌باشد:

[7] Minimax test
[8] Receiver Operating Characteristic



شکل ۶. سمت چپ: توزیع داده‌های دو کلاس؛ سمت راست: منحنی **error-reject** یا همان منحنی **ROC** که با تغییر مقدار $P(\omega_i)$ مابین صفر و یک حاصل می‌گردد.

جواب سوال ۲

ضابطه‌ی فیشر[9] معیاری است که تضمین می‌کند چقدر توزیع داده‌های دو کلاس ω_1 و ω_2 پس از اعمال پارامترهای خط جداساز، از یکدیگر دور شده‌اند. این ضابطه و معادله‌ی خط جداساز به صورت زیر می‌باشند:

$$f = \frac{(\eta_1 - \eta_2)^2}{(\sigma_1^2 + \sigma_2^2)}, \quad \text{Fisher Criterion}$$

$$h(X) = V^T X + v_0 \mathop{\gtrless}_{\omega_2}^{\omega_1} 0, \quad \text{Decision Rule}$$

بنابراین می‌بایست مقادیری را به ازای بردار V و مقدار بایاس v_0 بیابیم که مقدار این ضابطه را بیشینه می‌کنند تا بهترین وضعیت جداپذیری میان داده‌ها حاصل گردد. مقادیر مربوطه برای تابع جداساز و البته جواب نهائی مسئله به صورت زیر می‌باشند:

$$V = \left(\tfrac{1}{2}\Sigma_1 + \tfrac{1}{2}\Sigma_2\right)^{-1}(M_1 - M_2), \quad \mathrm{v}_0 = \frac{(M_2 - M_1)^T \left(\tfrac{1}{2}\Sigma_1 + \tfrac{1}{2}\Sigma_2\right)^{-1}\left(\sigma_1^2 M_2 + \sigma_2^2 M_1\right)}{\sigma_1^2 + \sigma_2^2}$$

$$\eta_i = V^T M_i + v_0, \quad \sigma_i^2 = V^T \Sigma_i V$$

$$P_1 = P_2 = .5$$

$$M_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \Sigma_1 = \begin{pmatrix} 4 & -3 \\ -3 & 4 \end{pmatrix}$$

$$M_2 = \begin{pmatrix} -1 \\ 0 \end{pmatrix}, \quad \Sigma_2 = \begin{pmatrix} 4 & 3 \\ 3 & 4 \end{pmatrix}$$

[9] Fisher criterion

$$\Rightarrow V=\left(\tfrac{1}{2}\begin{pmatrix}4&-3\\-3&4\end{pmatrix}+\tfrac{1}{2}\begin{pmatrix}4&3\\3&4\end{pmatrix}\right)^{-1}\left(\begin{pmatrix}1\\0\end{pmatrix}-\begin{pmatrix}-1\\0\end{pmatrix}\right)=\begin{pmatrix}4&0\\0&4\end{pmatrix}^{-1}\begin{pmatrix}2\\0\end{pmatrix}=\begin{pmatrix}.25&0\\0&.25\end{pmatrix}\begin{pmatrix}2\\0\end{pmatrix}=\begin{pmatrix}.5\\0\end{pmatrix}$$

$$\Rightarrow \sigma_1^2=\begin{pmatrix}.5\\0\end{pmatrix}^T\begin{pmatrix}4&-3\\-3&4\end{pmatrix}\begin{pmatrix}.5\\0\end{pmatrix}=1,\quad \sigma_2^2=\begin{pmatrix}.5\\0\end{pmatrix}^T\begin{pmatrix}4&3\\3&4\end{pmatrix}\begin{pmatrix}.5\\0\end{pmatrix}=1$$

$$\Rightarrow \mathrm{v}_0=\frac{\left(\begin{pmatrix}-1\\0\end{pmatrix}-\begin{pmatrix}1\\0\end{pmatrix}\right)^T\left(\tfrac{1}{2}\begin{pmatrix}4&-3\\-3&4\end{pmatrix}+\tfrac{1}{2}\begin{pmatrix}4&3\\3&4\end{pmatrix}\right)^{-1}\left(1\times\begin{pmatrix}-1\\0\end{pmatrix}+1\times\begin{pmatrix}1\\0\end{pmatrix}\right)}{1+1}=\frac{\begin{pmatrix}-2\\0\end{pmatrix}^T\begin{pmatrix}4&0\\0&4\end{pmatrix}^{-1}\begin{pmatrix}0\\0\end{pmatrix}}{2}=0$$

$$\Rightarrow h(X)=\begin{pmatrix}.5\\0\end{pmatrix}^T\begin{pmatrix}x_1\\x_2\end{pmatrix}+0\underset{\omega_2}{\overset{\omega_1}{\gtrless}}0\Rightarrow .5x_1\underset{\omega_2}{\overset{\omega_1}{\gtrless}}0\Rightarrow x_1\underset{\omega_2}{\overset{\omega_1}{\gtrless}}0$$

قسمت ب)

از آن‌جا که توزیع دقیق داده‌های مربوط به دو کلاس در دسترس می‌باشند، لذا تابع جداساز خطی که احتمال خطا را مینیمم می‌کند همان مرز تصمیم‌گیری بیزین با مینیمم خطا می‌باشد. داریم:

$$P_1=P_2=.5$$

$$M_1=\begin{pmatrix}1\\0\end{pmatrix},\quad \Sigma_1=\begin{pmatrix}4&-3\\-3&4\end{pmatrix}$$

$$M_2=\begin{pmatrix}-1\\0\end{pmatrix},\quad \Sigma_2=\begin{pmatrix}4&3\\3&4\end{pmatrix}$$

$$p(w_1|X)\underset{w_2}{\overset{w_1}{\gtrless}}p(w_2|X)\Rightarrow p(X|w_1)P(w_1)\underset{w_2}{\overset{w_1}{\gtrless}}p(X|w_2)P(w_2)\Rightarrow\frac{p(X|w_1)}{p(X|w_2)}\underset{w_2}{\overset{w_1}{\gtrless}}\frac{P(w_2)}{P(w_1)}$$

$$-\ln\left(\frac{p(X|w_1)}{p(X|w_2)}\right)\underset{w_2}{\overset{w_1}{\lessgtr}}-\ln\left(\frac{P(w_2)}{P(w_1)}\right)\Rightarrow$$

$$h(X)=\frac{1}{2}(X-M_1)^T\Sigma_1^{-1}(X-M_1)-\frac{1}{2}(X-M_2)^T\Sigma_2^{-1}(X-M_2)+\frac{1}{2}\ln\left(\frac{|\Sigma_1|}{|\Sigma_2|}\right)\underset{w_2}{\overset{w_1}{\lessgtr}}\ln\left(\frac{P_1}{P_2}\right)$$

Now for this question we have:

$$h(X)=\frac{1}{2}\begin{pmatrix}x_1-1\\x_2\end{pmatrix}^T\begin{pmatrix}\frac{4}{7}&-\frac{3}{7}\\-\frac{3}{7}&\frac{4}{7}\end{pmatrix}\begin{pmatrix}x_1-1\\x_2\end{pmatrix}-\frac{1}{2}\begin{pmatrix}x_1+1\\x_2\end{pmatrix}^T\begin{pmatrix}\frac{4}{7}&\frac{3}{7}\\\frac{3}{7}&\frac{4}{7}\end{pmatrix}\begin{pmatrix}x_1+1\\x_2\end{pmatrix}+\frac{1}{2}\ln\left(\frac{7}{7}\right)\underset{w_2}{\overset{w_1}{\lessgtr}}\ln\left(\frac{.5}{.5}\right)$$

$$h(X)=\frac{1}{2}\begin{pmatrix}x_1-1\\x_2\end{pmatrix}^T\begin{pmatrix}\frac{4}{7}&-\frac{3}{7}\\-\frac{3}{7}&\frac{4}{7}\end{pmatrix}\begin{pmatrix}x_1-1\\x_2\end{pmatrix}-\frac{1}{2}\begin{pmatrix}x_1+1\\x_2\end{pmatrix}^T\begin{pmatrix}\frac{4}{7}&\frac{3}{7}\\\frac{3}{7}&\frac{4}{7}\end{pmatrix}\begin{pmatrix}x_1+1\\x_2\end{pmatrix}\underset{w_2}{\overset{w_1}{\lessgtr}}0$$

جواب سوال ۳

کدهای مربوط این سوال در فایل‌های ex03 قرار دارد.

قسمت الف)

برای محاسبه‌ی تابع جداساز $g_i(x)$ به ازای هر کلاس می‌توانیم توابع جداساز مابین هر کلاس با دیگر کلاس‌ها را در یکدیگر ضرب کرده و در صورت مثبت‌بودن عدد حاصله، نتیجه‌گیری نمائیم که آیا داده‌ی مربوطه به آن کلاس تعلق دارد یا خیر. داریم:

$$p(X \mid \omega_i) \sim N(\mu_i, \Sigma_i), \quad \text{i=1,2,3 with}$$

$$\mu_1 = \begin{pmatrix} 0 \\ 2 \end{pmatrix}, \mu_2 = \begin{pmatrix} 3 \\ 1 \end{pmatrix}, \mu_3 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \Sigma_i = \Sigma = \begin{pmatrix} 1 & 0 \\ 0 & 1/3 \end{pmatrix}$$

$$h_{ij} \equiv \frac{p(X \mid \omega_i)}{p(X \mid \omega_j)} \underset{\omega_j}{\overset{\omega_i}{><}} \frac{P(\omega_j)}{P(\omega_i)}, \quad \text{since } P(\omega_i) = P(\omega_j) = 1/3 \quad \Rightarrow \quad h_{ij} \equiv \frac{p(X \mid \omega_i)}{p(X \mid \omega_j)} - 1 \underset{\omega_j}{\overset{\omega_i}{><}} 0$$

$$g_{i=c}(X) = sign\left(\prod_{\substack{j \neq i \\ i=c}} h_{ij}\right) \underset{\omega_j}{\overset{\omega_i}{><}} 0$$

قسمت ب)

برای توصیف توابع جداساز خطی در قالب یک تابع، می‌توان این‌گونه بیان نمود که توابع خطی مربوطه در واقع همان عمودمنصف‌های پاره‌خط‌های میان مقادیر میانگین دوبه‌دوی کلاس‌ها می‌باشند. داریم:

$\mu_{ij} \equiv$ the average point between μ_i and μ_j,

$s_{ij}^{-1} \equiv$ the slope of the perpendicular line to the linking line between μ_i and μ_j

$$\mu_{12} = \begin{pmatrix} 1.5 \\ 1.5 \end{pmatrix}, \mu_{13} = \begin{pmatrix} .5 \\ 1 \end{pmatrix}, \mu_{23} = \begin{pmatrix} 2 \\ .5 \end{pmatrix}$$

$$s_{12}^{-1} = \left(\frac{1-2}{3-0}\right)^{-1} = -3, s_{13}^{-1} = \left(\frac{0-2}{1-0}\right)^{-1} = -.5, s_{23}^{-1} = \left(\frac{0-1}{1-3}\right)^{-1} = 2$$

So we have:

$$g_{12}: x_2 - 1.5 = -3(x_1 - 1.5) \Rightarrow g_{12}: x_2 + 3x_1 - 6 = 0$$

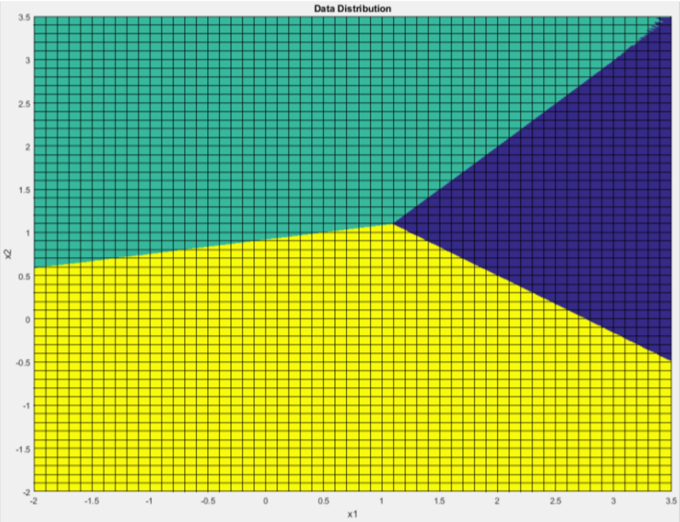
$$g_{13}: x_2 - 1 = -.5(x_1 - .5) \Rightarrow g_{13}: x_2 + .5x_1 - 1.25 = 0$$

$$g_{23}: x_2 - .5 = 2(x_1 - 2) \Rightarrow g_{23}: x_2 - 2x_1 + 3.5 = 0$$

$$X \in \omega_i \quad \text{iff} \quad g_{ij} > 0 \quad \text{for } \forall j \neq i$$

قسمت ج)

مرزهای تصمیم‌گیری مربوط به سه کلاس قیدشده در صورت مسئله به صورت زیر می‌باشند:



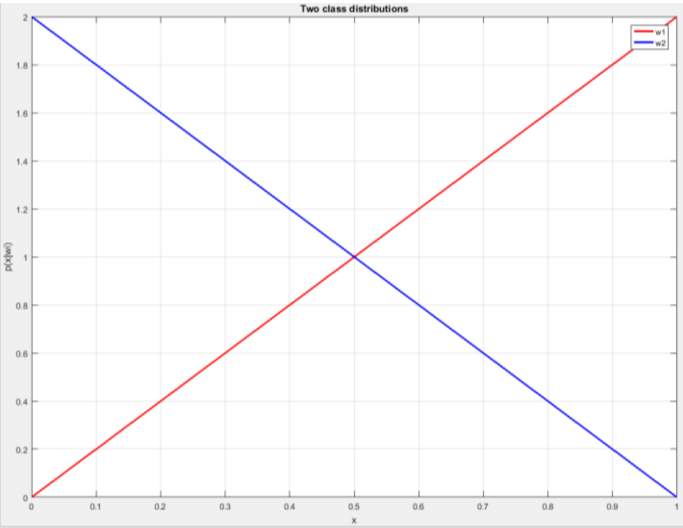
شکل ۷. مرزهای تصمیم‌گیری مربوط به سه کلاس قیدشده در صورت مسئله

جواب سوال ۴

کدهای مربوط این سوال در فایل‌های ex03 قرار دارد.

$$p(x\mid\omega_1)=\begin{cases}2x\,, & 0\le x\le 1\\ 0, & o.w.\end{cases}\qquad p(x|\omega_2)=\begin{cases}2-2x\,, & 0\le x\le 1\\ 0, & o.w.\end{cases}$$

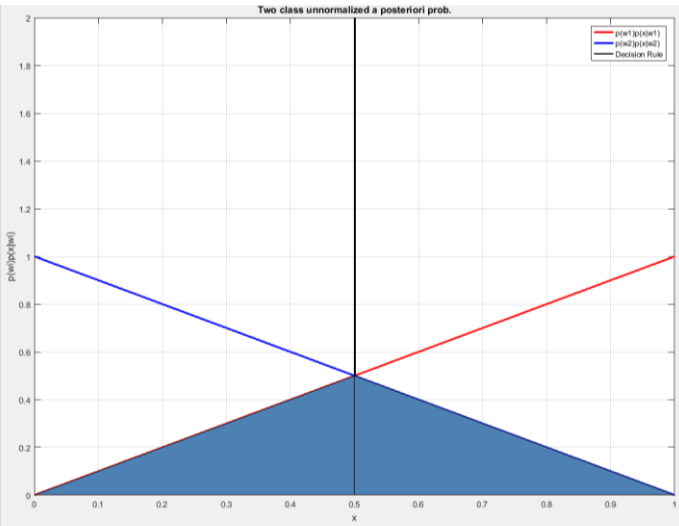
قسمت الف)



شکل ۸. توزیع داده‌های مربوط به دو کلاس.

قسمت ب)

$$\frac{p(X\mid\omega_1)}{p(X\mid\omega_2)}\overset{\omega_1}{\underset{\omega_2}{><}}\frac{P(\omega_2)}{P(\omega_1)}\Rightarrow\frac{2x}{2-2x}\overset{\omega_1}{\underset{\omega_2}{><}}\frac{.5}{.5}\Rightarrow 2x\overset{\omega_1}{\underset{\omega_2}{><}}2-2x\Rightarrow 4x-2\overset{\omega_1}{\underset{\omega_2}{><}}0\Rightarrow x\overset{\omega_1}{\underset{\omega_2}{><}}.5$$



شکل ۹. نمودار احتمال a posteriori داده‌های مربوط به دو کلاس با $P_1=P_2=.5$؛ ناحیه‌ی هاشورخورده با رنگ آبی نشان‌دهنده‌ی ناحیه‌ی خطای دسته‌بند بیزین می‌باشد.

قسمت ج)

$$\varepsilon = P(\omega_1)\varepsilon_1 + P(\omega_2)\varepsilon_2 = P(\omega_1)\int_0^{.5} p(X \mid \omega_1)dX + P(\omega_2)\int_{.5}^{1} p(X \mid \omega_2)dX \Rightarrow$$

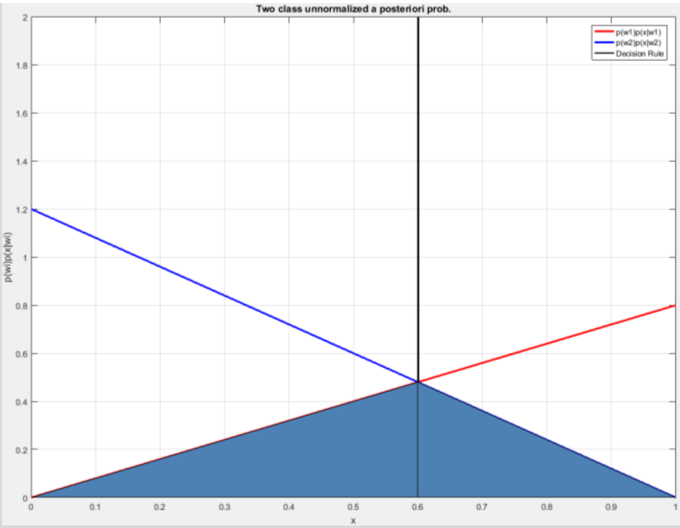
$$\varepsilon = .5\int_0^{.5} 2X\,dX + .5\int_{.5}^{1}(2-2x)dX \Rightarrow .5\left[X^2 + c_1\right]_0^{.5} + .5\left[2X - X^2 + c_2\right]_{.5}^{1} = {\color{red}.25}$$

قسمت د)

$$\frac{p(X \mid \omega_1)}{p(X \mid \omega_2)} \underset{\omega_2}{\overset{\omega_1}{><}} \frac{P(\omega_2)}{P(\omega_1)} \Rightarrow \frac{2x}{2-2x} \underset{\omega_2}{\overset{\omega_1}{><}} \frac{.6}{.4} \Rightarrow 2x \underset{\omega_2}{\overset{\omega_1}{><}} 1.5(2-2x) \Rightarrow 5x - 3 \underset{\omega_2}{\overset{\omega_1}{><}} 0 \Rightarrow {\color{red}x \underset{\omega_2}{\overset{\omega_1}{><}} .6}$$

$$\varepsilon = P(\omega_1)\varepsilon_1 + P(\omega_2)\varepsilon_2 = P(\omega_1)\int_0^{.6} p(X \mid \omega_1)dX + P(\omega_2)\int_{.6}^{1} p(X \mid \omega_2)dX \Rightarrow$$

$$\varepsilon = .4\int_0^{.6} 2X\,dX + .6\int_{.6}^{1}(2-2x)dX \Rightarrow .4\left[X^2 + c_1\right]_0^{.6} + .6\left[2X - X^2 + c_2\right]_{.6}^{1} = {\color{red}.24}$$



شکل ۱۰. نمودار احتمال a posteriori داده‌های مربوط به دو کلاس با $P_1 = .4, P_2 = .6$؛ ناحیه‌ی هاشورخورده با رنگ آبی نشان‌دهنده‌ی ناحیه‌ی خطای دسته‌بند بیزین می‌باشد.

جواب سوال ۵

$$p(X \mid \omega_1) \sim N(0, I), \quad p(X \mid \omega_2) \sim N\left(\begin{pmatrix}1\\0\end{pmatrix}, I\right), \quad P(\omega_1) = P(\omega_2) = .5$$

قسمت الف)

$$h(X) = \frac{1}{2}(X - M_1)^T \Sigma_1^{-1}(X - M_1) - \frac{1}{2}(X - M_2)^T \Sigma_2^{-1}(X - M_2) + \frac{1}{2}\ln\left(\frac{|\Sigma_1|}{|\Sigma_2|}\right) \underset{w_2}{\overset{w_1}{<>}} \ln\left(\frac{P_1}{P_2}\right) \Rightarrow$$

$$h(X) = \frac{1}{2}\begin{pmatrix}x_1\\x_2\end{pmatrix}^T \begin{pmatrix}x_1\\x_2\end{pmatrix} - \frac{1}{2}\begin{pmatrix}x_1-1\\x_2\end{pmatrix}^T \begin{pmatrix}x_1-1\\x_2\end{pmatrix} \underset{w_2}{\overset{w_1}{<>}} 0 \Rightarrow \begin{pmatrix}x_1\\x_2\end{pmatrix}^T \begin{pmatrix}x_1\\x_2\end{pmatrix} - \begin{pmatrix}x_1-1\\x_2\end{pmatrix}^T \begin{pmatrix}x_1-1\\x_2\end{pmatrix} \underset{w_2}{\overset{w_1}{<>}} 0 \Rightarrow$$

$$x_1^2 + x_2^2 - (x_1 - 1)^2 - x_2^2 = 2x_1 - 1 \underset{w_2}{\overset{w_1}{<>}} 0 \Rightarrow {\color{red}x_1 \underset{w_2}{\overset{w_1}{<>}} .5}$$

قسمت ب)

مرز خطای باتاچریا[10] به صورت زیر می‌باشد:

$$\mu(.5)=\frac{1}{8}\left(M_2-M_1\right)^T\left(\frac{\Sigma_1+\Sigma_2}{2}\right)^{-1}\left(M_2-M_1\right)+\frac{1}{2}\ln\frac{\left|\frac{\Sigma_1+\Sigma_2}{2}\right|}{|\Sigma_1|^{\frac{1}{2}}|\Sigma_2|^{\frac{1}{2}}}\to\Sigma_1=\Sigma_2=I\Rightarrow$$

$$\mu(.5)=\frac{1}{8}\left(M_2-M_1\right)^T\left(M_2-M_1\right)=\frac{1}{8}\begin{pmatrix}1\\0\end{pmatrix}^T\begin{pmatrix}1\\0\end{pmatrix}=\frac{1}{8}\Rightarrow\varepsilon=\sqrt{P(\omega_1)P(\omega_2)}\exp\{-\mu(.5)\}\Rightarrow$$

$$\varepsilon=\frac{1}{2}\exp\left\{-\frac{1}{8}\right\}$$

قسمت ج)

$$p(X\mid\omega_1)\sim N\left(0,\begin{pmatrix}2&.5\\.5&2\end{pmatrix}\right),\quad p(X\mid\omega_2)\sim N\left(\begin{pmatrix}1\\0\end{pmatrix},\begin{pmatrix}5&2\\2&5\end{pmatrix}\right),\quad P(\omega_1)=P(\omega_2)=.5$$

$$h(X)=\frac{1}{2}(X-M_1)^T\Sigma_1^{-1}(X-M_1)-\frac{1}{2}(X-M_2)^T\Sigma_2^{-1}(X-M_2)+\frac{1}{2}\ln\left(\frac{|\Sigma_1|}{|\Sigma_2|}\right)\underset{w_2}{\overset{w_1}{<>}}\ln\left(\frac{P_1}{P_2}\right)\Rightarrow$$

$$h(X)=\frac{1}{2}\begin{pmatrix}x_1\\x_2\end{pmatrix}^T\begin{pmatrix}.533&-.133\\-.133&.533\end{pmatrix}\begin{pmatrix}x_1\\x_2\end{pmatrix}-\frac{1}{2}\begin{pmatrix}x_1-1\\x_2\end{pmatrix}^T\begin{pmatrix}.238&-.095\\-.095&.238\end{pmatrix}\begin{pmatrix}x_1-1\\x_2\end{pmatrix}\underset{w_2}{\overset{w_1}{<>}}0\Rightarrow$$

$$\begin{pmatrix}x_1\\x_2\end{pmatrix}^T\begin{pmatrix}.533&-.133\\-.133&.533\end{pmatrix}\begin{pmatrix}x_1\\x_2\end{pmatrix}-\begin{pmatrix}x_1-1\\x_2\end{pmatrix}^T\begin{pmatrix}.238&-.095\\-.095&.238\end{pmatrix}\begin{pmatrix}x_1-1\\x_2\end{pmatrix}\underset{w_2}{\overset{w_1}{<>}}0\Rightarrow$$

So the Bayes decision rule is: $.295x_1^2+.476x_1-.076x_1x_2-.19x_2+.295x_2^2-.238\underset{w_2}{\overset{w_1}{<>}}0$

The Bhattacharryya error bound is:

$$\mu(.5)=\frac{1}{8}\left(M_2-M_1\right)^T\left(\frac{\Sigma_1+\Sigma_2}{2}\right)^{-1}\left(M_2-M_1\right)+\frac{1}{2}\ln\frac{\left|\frac{\Sigma_1+\Sigma_2}{2}\right|}{|\Sigma_1|^{\frac{1}{2}}|\Sigma_2|^{\frac{1}{2}}}\Rightarrow$$

$$\mu(.5)=\frac{1}{8}\begin{pmatrix}1\\0\end{pmatrix}^T\left(\left(\begin{pmatrix}2&.5\\.5&2\end{pmatrix}+\begin{pmatrix}5&2\\2&5\end{pmatrix}\right)\Big/2\right)^{-1}\begin{pmatrix}1\\0\end{pmatrix}+\frac{1}{2}\ln\frac{\left|\left(\begin{pmatrix}2&.5\\.5&2\end{pmatrix}+\begin{pmatrix}5&2\\2&5\end{pmatrix}\right)\Big/2\right|}{(3.75)^{\frac{1}{2}}(21)^{\frac{1}{2}}}=$$

$$\Rightarrow\varepsilon=\sqrt{P(\omega_1)P(\omega_2)}\exp\{-\mu(.5)\}\Rightarrow$$

$$\varepsilon=\frac{1}{2}\exp\left\{-\frac{1}{8}\right\}$$

جواب سوال ۶

کدهای مربوط این سوال در فایل‌های ex06 قرار دارد.

$$P(\omega_1)=P(\omega_2)=.5$$

$$M_1=M_2=\begin{pmatrix}0\\0\end{pmatrix},\quad\Sigma_1=\begin{pmatrix}1&0\\0&4\end{pmatrix},\Sigma_2=\begin{pmatrix}4&0\\0&1\end{pmatrix}$$

برای یافتن مرز تصمیم‌گیری بیزین با مینیمم خطا داریم:

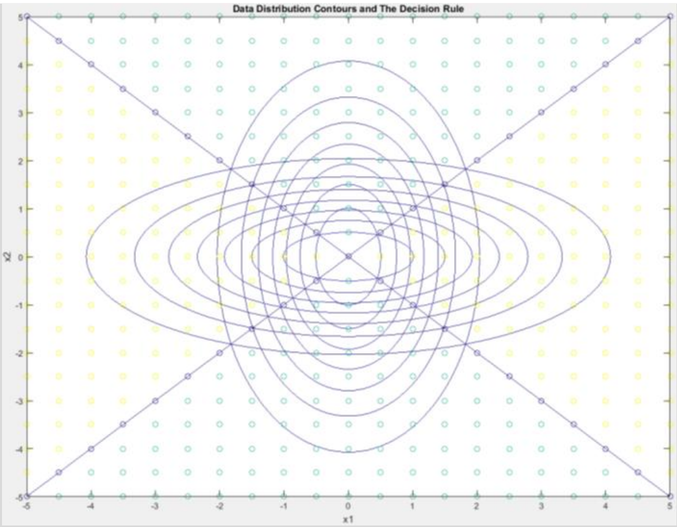
[10] Bhattacharryya Error Bound

$$h(X)=\frac{1}{2}(X-M_1)^T\,\Sigma_1^{-1}(X-M_1)-\frac{1}{2}(X-M_2)^T\,\Sigma_2^{-1}(X-M_2)+\frac{1}{2}\ln\left(\frac{|\Sigma_1|}{|\Sigma_2|}\right)\underset{w_2}{\overset{w_1}{<>}}\ln\left(\frac{P_1}{P_2}\right)\Rightarrow$$

$$\frac{1}{2}\begin{pmatrix}x_1\\x_2\end{pmatrix}^T\begin{pmatrix}1&0\\0&.25\end{pmatrix}\begin{pmatrix}x_1\\x_2\end{pmatrix}-\frac{1}{2}\begin{pmatrix}x_1\\x_2\end{pmatrix}^T\begin{pmatrix}.25&0\\0&1\end{pmatrix}\begin{pmatrix}x_1\\x_2\end{pmatrix}+\frac{1}{2}\ln\left(\frac{4}{4}\right)\underset{w_2}{\overset{w_1}{<>}}\ln\left(\frac{.5}{.5}\right)\Rightarrow$$

$$h(X)=x_1^2-x_2^2\underset{w_2}{\overset{w_1}{<>}}0$$

شکل زیر توزیع داده‌های دو کلاس را به همراه مرز تصمیم‌گیری مربوطه ضمن مشخص‌نمودن نواحی مربوط به دو کلاس نشان می‌دهد:



شکل ۱۱. توزیع داده‌های مربوط به دو کلاس به همراه مرز تصمیم‌گیری بیزین با کمترین خطا؛ نواحی با دایره‌های آبی کم‌رنگ مربوط به کلاس ω_1 و نواحی با دایره‌های زردرنگ مربوط به کلاس ω_2 می‌باشند.

جواب سوال ۷

کدهای مربوط این سوال در فایل‌های ex07 قرار دارد.

$p(X\|\omega_k)\sim N(\mu_k,\Sigma_k)$	
ω_1	ω_2
(1,1)	(6,8)
(2,2)	(7,9)
(3,3)	(9,9)
(1,-1)	(6,8)
(0,1)	(9,8)
(2,2)	(9,11)
	(10,9)
	(10,7)

قسمت الف)

مقدار احتمال اولیه[11] را برای هر کلاس به صورت زیر به دست می‌آوریم:

[11] a Prioroi Probability

$$P(\omega_k)=\frac{N_k}{N} \quad where: \; N_k = \text{\# of datasamples of class } \omega_k, \; \text{N = \# of total datasamples}$$

$$\Rightarrow P(\omega_1)=\frac{6}{14}, \; P(\omega_1)=\frac{8}{14}$$

قسمت ب)

$$\hat{M}_1=\frac{1}{6}\left[\begin{pmatrix}1\\1\end{pmatrix}+\begin{pmatrix}2\\2\end{pmatrix}+\begin{pmatrix}3\\3\end{pmatrix}+\begin{pmatrix}1\\-1\end{pmatrix}+\begin{pmatrix}0\\1\end{pmatrix}+\begin{pmatrix}2\\2\end{pmatrix}\right]=\frac{1}{6}\begin{pmatrix}9\\8\end{pmatrix}\Rightarrow \hat{M}_1=\begin{pmatrix}1.5\\1.33\end{pmatrix}$$

$$\hat{M}_2=\frac{1}{8}\left[\begin{pmatrix}6\\8\end{pmatrix}+\begin{pmatrix}7\\9\end{pmatrix}+\begin{pmatrix}9\\9\end{pmatrix}+\begin{pmatrix}6\\8\end{pmatrix}+\begin{pmatrix}9\\8\end{pmatrix}+\begin{pmatrix}9\\11\end{pmatrix}+\begin{pmatrix}10\\9\end{pmatrix}+\begin{pmatrix}10\\7\end{pmatrix}\right]=\frac{1}{8}\begin{pmatrix}66\\69\end{pmatrix}\Rightarrow \hat{M}_1=\begin{pmatrix}8.25\\8.625\end{pmatrix}$$

$$\hat{\Sigma}_1=\frac{1}{6-1}\left[\begin{pmatrix}1\\1\end{pmatrix}\begin{pmatrix}1&1\end{pmatrix}+\begin{pmatrix}2\\2\end{pmatrix}\begin{pmatrix}2&2\end{pmatrix}+\begin{pmatrix}3\\3\end{pmatrix}\begin{pmatrix}3&3\end{pmatrix}+\begin{pmatrix}1\\-1\end{pmatrix}\begin{pmatrix}1&-1\end{pmatrix}+\begin{pmatrix}0\\1\end{pmatrix}\begin{pmatrix}0&1\end{pmatrix}+\begin{pmatrix}2\\2\end{pmatrix}\begin{pmatrix}2&2\end{pmatrix}\right]\Rightarrow$$

$$\hat{\Sigma}_1=\begin{pmatrix}1.55&1.4\\1.4&2.22\end{pmatrix}$$

$$\hat{\Sigma}_2=\frac{1}{8-1}\left[\begin{pmatrix}6\\8\end{pmatrix}\begin{pmatrix}6&8\end{pmatrix}+\begin{pmatrix}7\\9\end{pmatrix}\begin{pmatrix}7&9\end{pmatrix}+\begin{pmatrix}9\\9\end{pmatrix}\begin{pmatrix}9&9\end{pmatrix}+\begin{pmatrix}6\\8\end{pmatrix}\begin{pmatrix}6&8\end{pmatrix}+\begin{pmatrix}9\\8\end{pmatrix}\begin{pmatrix}9&8\end{pmatrix}+\begin{pmatrix}9\\11\end{pmatrix}\begin{pmatrix}9&11\end{pmatrix}\right.$$

$$\left.\begin{pmatrix}10\\9\end{pmatrix}\begin{pmatrix}10&9\end{pmatrix}+\begin{pmatrix}10\\7\end{pmatrix}\begin{pmatrix}10&7\end{pmatrix}\right]\Rightarrow \hat{\Sigma}_2=\begin{pmatrix}12.5089&10.4152\\10.4152&12.0379\end{pmatrix}$$

قسمت ج)

$$p(w_1|X)\overset{w_1}{\underset{w_2}{><}}p(w_2|X)\Rightarrow p(X|w_1)P(w_1)\overset{w_1}{\underset{w_2}{><}}p(X|w_2)P(w_2)\Rightarrow \frac{p(X|w_1)}{p(X|w_2)}\overset{w_1}{\underset{w_2}{><}}\frac{P(w_2)}{P(w_1)}$$

$$-\ln\left(\frac{p(X|w_1)}{p(X|w_2)}\right)\overset{w_1}{\underset{w_2}{<>}}-\ln\left(\frac{P(w_2)}{P(w_1)}\right)\Rightarrow$$

$$h(X)=\frac{1}{2}(X-M_1)^T\,\Sigma_1^{-1}(X-M_1)-\frac{1}{2}(X-M_2)^T\,\Sigma_2^{-1}(X-M_2)+\frac{1}{2}\ln\left(\frac{|\Sigma_1|}{|\Sigma_2|}\right)\overset{w_1}{\underset{w_2}{<>}}\ln\left(\frac{P_1}{P_2}\right)\Rightarrow$$

$$\frac{1}{2}\begin{pmatrix}x_1-1.5\\x_2-1.33\end{pmatrix}^T\begin{pmatrix}1.4970&-.9431\\-.9431&1.0442\end{pmatrix}\begin{pmatrix}x_1-1.5\\x_2-1.33\end{pmatrix}-\frac{1}{2}\begin{pmatrix}x_1-8.25\\x_2-8.625\end{pmatrix}^T\begin{pmatrix}.2859&-.2474\\-.2474&.2971\end{pmatrix}\begin{pmatrix}x_1-8.25\\x_2-8.625\end{pmatrix}+$$

$$\frac{1}{2}\ln\left(\frac{1.4844}{42.1059}\right)\overset{w_1}{\underset{w_2}{<>}}\ln\left(\frac{6/14}{8/14}\right)$$

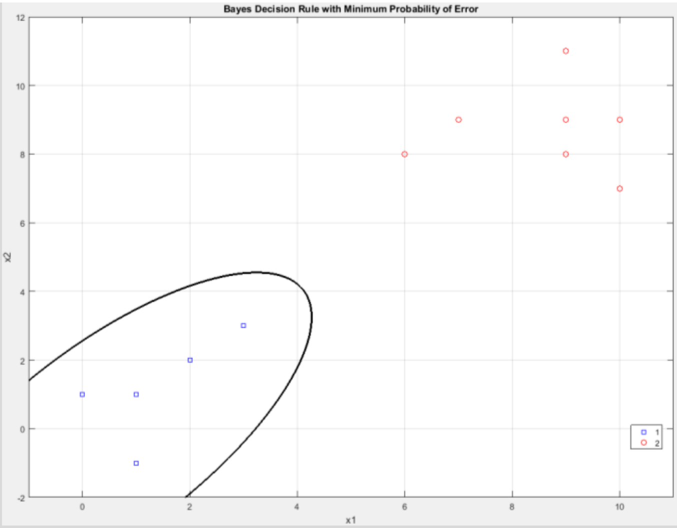
در شکل ۱۲، توزیع داده‌های دو کلاس به همراه مرز تصمیم‌گیری بیزین با مینیمم خطا آمده است.

قسمت د)

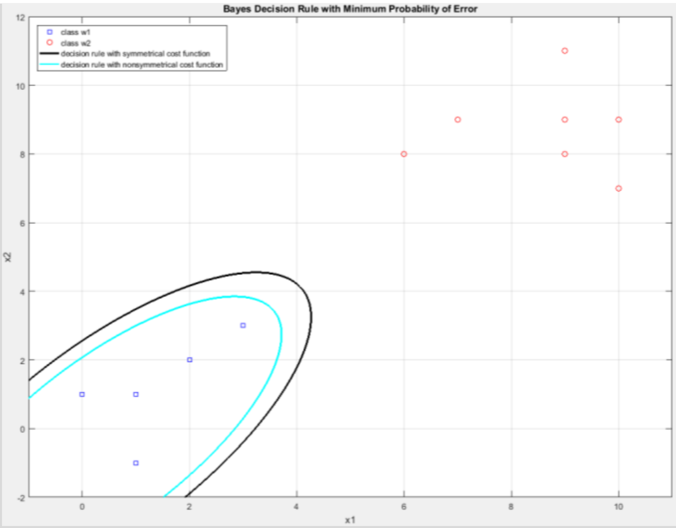
اگر جریمه و یا هزینه‌ی خطا در دسته‌بندی برای داده‌های یک کلاس از دیگری بیشتر باشد، قطعا مرز تصمیم‌گیری تحت تأثیر قرار گرفته و به سمت کلاسی که هزینه‌ی خطا در دسته‌بندی آن کمتر است، متمایل خواهد شد. به عنوان مثال در شکل ۱۳ مقادیر تابع هزینه به صورت زیر می‌باشند (جریمه‌ی خطای دسته‌بندی برای کلاس ۲ بیشتر از کلاس ۱ می‌باشد):

$$c_{11}=c_{22}=0, \quad c_{12}=1, \quad c_{21}=3$$

مشاهده می‌شود که از آن‌جا که هزینه‌ی خطا برای کلاس ۲ بیشتر از کلاس ۱ می‌باشد، مرز تصمیم‌گیری به جهت مخالف کلاس ۲، یعنی کلاس ۱ متمایل گشته است.



شکل ۱۲. توزیع داده‌های مربوط به دو کلاس به همراه مرز تصمیم‌گیری بیزین با کمترین خطا؛ نقاط آبی‌رنگ مربوط به کلاس ω_1 و نقاط قرمزرنگ مربوط به کلاس ω_2 می‌باشند.



شکل ۱۳. توزیع داده‌های مربوط به دو کلاس به همراه مرز تصمیم‌گیری بیزین با کمترین خطا که با رنگ مشکی نشان داده شده است؛ و مرز تصمیم‌گیری بیزین با مینیمم خطا و تابع هزینه‌ی غیر متقارن که با رنگ سبزمتمایل‌به‌آبی نشان داده شده است.

جواب سوال ۸

کدهای مربوط این سوال در فایل‌های ex08 قرار دارد.

$$P(c_1) = P(c_2) = .5$$

$$p(x \mid c_1) = U\left(a,b\right) \quad \text{where} \;\; a = 2, b = 4$$

$$p(x \mid c_2) = \lambda \exp\left(-\lambda x\right) \quad \text{where} \;\; \lambda = 1$$

قسمت الف)

$$\frac{p(X\ |w_1)}{p(X\ |w_2)} \underset{w_2}{\overset{w_1}{><}} \frac{P(w_2)}{P(w_1)} \Rightarrow \frac{p(X\ |w_1)}{p(X\ |w_2)} \underset{w_2}{\overset{w_1}{><}} 1 \Rightarrow p(X\ |w_1) \underset{w_2}{\overset{w_1}{><}} p(X\ |w_2) \Rightarrow$$

$$.5 \underset{w_2}{\overset{w_1}{><}} \exp(-x) \Rightarrow \ln(.5) \underset{w_2}{\overset{w_1}{><}} -x \Rightarrow x \underset{w_2}{\overset{w_1}{><}} -\ln(.5) = 2 \rightarrow$$

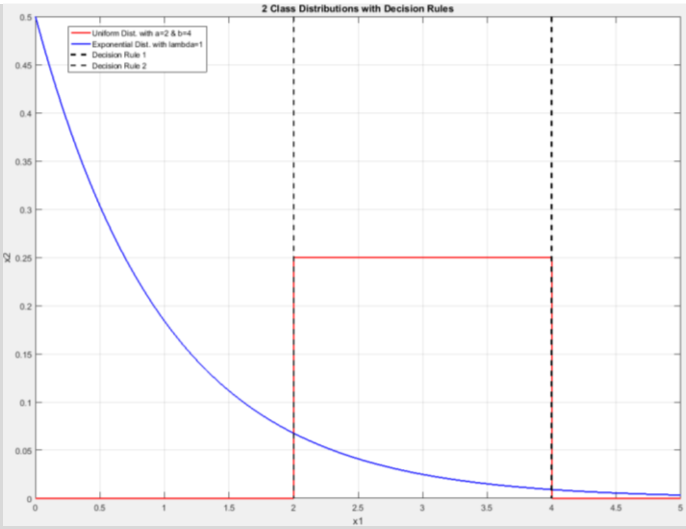
As for c_1 we have a uniform dist. with a=2 and b=4 $\Rightarrow$

$2 \le x \le 4 \Rightarrow x > -\ln(.5) = 2 \Rightarrow x \in c_1$ otherwise $x \in c_2 \Rightarrow$

So the decision boundaries are: x=2 and x=4

قسمت ب)

در شکل ۱۴، توزیع داده‌های مربوط به دو کلاس به همراه مرزهای تصمیم‌گیری مربوطه نشان داده شده‌اند:



شکل ۱۴. توزیع داده‌های مربوط به دو کلاس به همراه مرز تصمیم‌گیری بیزین با کمترین خطا که با رنگ مشکی و خط غیرمتد نشان داده شده است.

قسمت ج)

خطای بیزین با مینیمم خطا به صورت زیر محاسبه می‌گردد:

$$\varepsilon = P(c_1)\varepsilon_1 + P(c_2)\varepsilon_2 = P(c_1)\int_{\Gamma_2} p(x\ |c_1)dx + P(c_2)\int_{\Gamma_1} p(x\ |c_2)dx$$

where $\Gamma_1 = [2,4], \Gamma_2 = (0,2] \cup [4,+\infty)$

$$\varepsilon = .5\int_0^2 0dx + .5\int_4^{+\infty} 0dx + .5\int_2^4 \exp(-x)dx = .5\left[-\exp(-x)\right]_2^4 \Rightarrow \varepsilon = .059$$

قسمت د)

$$P(c_1) = P(c_2) = .5$$

$$p(x\ |c_1) = U\ (a,b) \quad \text{where}\ \ a = 2, b = 22$$

$$p(x\ |c_2) = \lambda \exp(-\lambda x) \quad \text{where}\ \ \lambda = 1$$

$$\frac{p(X\ |w_1)}{p(X\ |w_2)} \underset{w_2}{\overset{w_1}{><}} \frac{P(w_2)}{P(w_1)} \Rightarrow \frac{p(X\ |w_1)}{p(X\ |w_2)} \underset{w_2}{\overset{w_1}{><}} 1 \Rightarrow p(X\ |w_1) \underset{w_2}{\overset{w_1}{><}} p(X\ |w_2) \Rightarrow$$

$$.05 \underset{w_2}{\overset{w_1}{><}} \exp(-x) \Rightarrow \ln(.05) \underset{w_2}{\overset{w_1}{><}} -x \Rightarrow x \underset{w_2}{\overset{w_1}{><}} -\ln(.05) = 2.9957 \rightarrow$$

As for c_1 we have a uniform dist. with a=2 and b=22 $\Rightarrow$

$2.9957 \le x \le 22 \Rightarrow x > -\ln(.05) = 2.9957 \Rightarrow x \in c_1$ otherwise $x \in c_2 \Rightarrow$

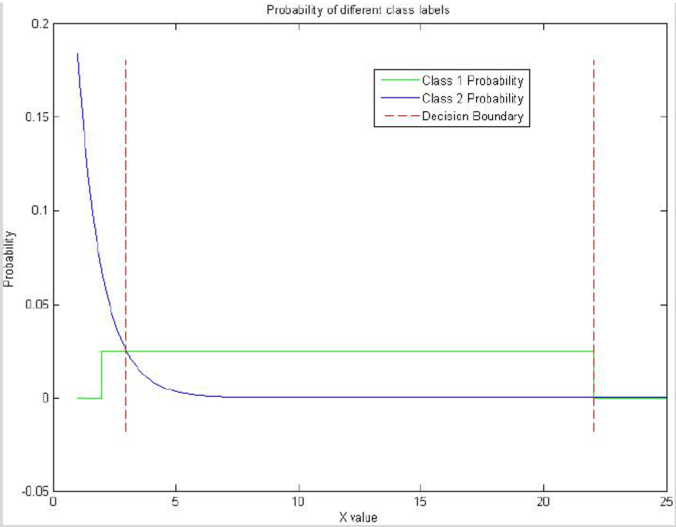
So the decision boundaries are: x=2.9957 and x=22

$$\varepsilon = P(c_1)\varepsilon_1 + P(c_2)\varepsilon_2 = P(c_1)\int_{\Gamma_2} p(x\,|c_1)dx + P(c_2)\int_{\Gamma_1} p(x\,|c_2)dx$$

$$\text{where}\quad \Gamma_1 = [2.9957, 22], \Gamma_2 = (0, 2.9957] \cup [22, +\infty)$$

$$\varepsilon = .5\int_0^2 0dx + .5\int_2^{2.9957} .05dx + .5\int_{22}^{+\infty} 0dx + .5\int_{2.9957}^{22} \exp(-x)dx = .5[.05x]_2^{2.9957} + .5[-\exp(-x)]_{2.9957}^{22} \Rightarrow \varepsilon = .050$$

در شکل ۱۵، توزیع داده‌های مربوط به دو کلاس به همراه مرزهای تصمیم‌گیری مربوطه نشان داده شده‌اند:



شکل ۱۵. توزیع داده‌های مربوط به دو کلاس به همراه مرز تصمیم‌گیری بیزین با کمترین خطا که با رنگ مشکی و خط غیرممتد نشان داده شده است.

جواب سوال ۹

$$p(x\,|c_1) \sim N(M_1, \Sigma) \;\text{ and } P(c_1), \quad p(x\,|c_2) \sim N(M_2, \Sigma) \;\text{ and } P(c_2)$$

قسمت الف)

$$p(c_1\,|X) \underset{w_2}{\overset{w_1}{\gtrless}} p(c_2\,|X) \Rightarrow p(X\,|c_1)P(c_1) \underset{w_2}{\overset{w_1}{\gtrless}} p(X\,|c_2)P(c_2) \quad \langle 1 \rangle$$

$$g_i(X) = \log\left(p(X\,|c_1)P(c_1)\right) = \log(p(X\,|c_1)) + \log(P(c_1)) \overset{<1>}{\Rightarrow}$$

$$\log(p(X\,|c_1)) + \log(P(c_1)) \underset{w_2}{\overset{w_1}{\gtrless}} \log(p(X\,|c_2)) + \log(P(c_2))$$

قسمت ب)

$$g(X) = g_1(X) - g_2(X) = 0 \Rightarrow \log(p(X\,|c_1)) + \log(P(c_1)) - \log(p(X\,|c_2)) - \log(P(c_2)) = 0 \Rightarrow$$

$$\log(p(X\,|c_1)) - \log(p(X\,|c_2)) = \log(P(c_2)) - \log(P(c_1)) \Rightarrow$$

$$\log\left(\frac{p(X\,|c_1)}{p(X\,|c_2)}\right) = \log\left(\frac{P(c_2)}{P(c_1)}\right) \Rightarrow -\log\left(\frac{p(X\,|c_1)}{p(X\,|c_2)}\right) = -\log\left(\frac{P(c_2)}{P(c_1)}\right) \Rightarrow$$

$$\frac{1}{2}(X - M_1)^T \Sigma^{-1}(X - M_1) - \frac{1}{2}(X - M_2)^T \Sigma^{-1}(X - M_2) + \frac{1}{2}\ln\left(\frac{|\Sigma|}{|\Sigma|}\right) = \log\left(\frac{P(c_1)}{P(c_2)}\right) \Rightarrow$$

$$\frac{1}{2}\left(X^T\Sigma^{-1}-M_1^T\Sigma^{-1}\right)\left(X-M_1\right)-\frac{1}{2}\left(X^T\Sigma^{-1}-M_2^T\Sigma^{-1}\right)\left(X-M_2\right)=\log\left(\frac{P(c_1)}{P(c_2)}\right)\Rightarrow$$

$$\frac{1}{2}\left(X^T\Sigma^{-1}X-X^T\Sigma^{-1}M_1-M_1^T\Sigma^{-1}X+M_1^T\Sigma^{-1}M_1\right)-$$

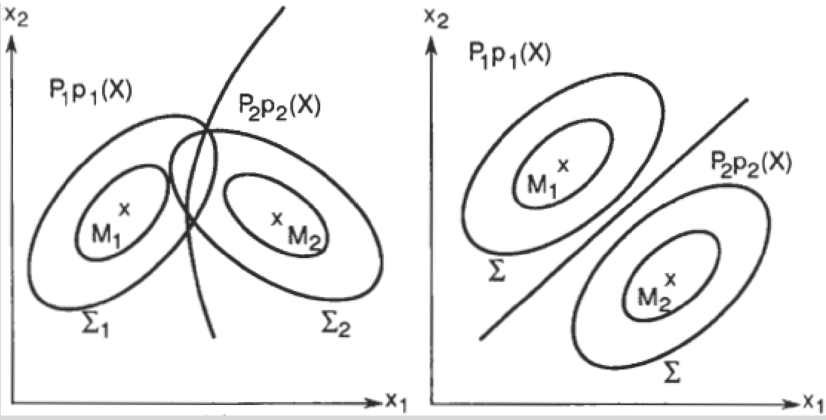
$$\frac{1}{2}\left(X^T\Sigma^{-1}X-X^T\Sigma^{-1}M_2-M_2^T\Sigma^{-1}X+M_2^T\Sigma^{-1}M_2\right)=\log\left(\frac{P(c_1)}{P(c_2)}\right)\rightarrow$$

$\Sigma^{-1}=\Sigma^{-1^T}, X^T\Sigma^{-1}M_1$ is a scalar, so: $\left(X^T\Sigma^{-1}M_1\right)^T=X^T\Sigma^{-1}M_1\Rightarrow M_1^T\Sigma^{-1}X=X^T\Sigma^{-1}M_1\Rightarrow$

$$\frac{1}{2}\left(X^T\Sigma^{-1}X-2M_1^T\Sigma^{-1}X+M_1^T\Sigma^{-1}M_1\right)-\frac{1}{2}\left(X^T\Sigma^{-1}X-2M_2^T\Sigma^{-1}X+M_2^T\Sigma^{-1}M_2\right)=\log\left(\frac{P(c_1)}{P(c_2)}\right)\Rightarrow$$

$$\underbrace{(M_2-M_1)^T\Sigma^{-1}}_{w^T}X+\underbrace{\frac{1}{2}\left(M_1^T\Sigma^{-1}M_1-M_2^T\Sigma^{-1}M_2\right)-\log\left(\frac{P(c_1)}{P(c_2)}\right)}_{w_0}=0\Rightarrow w^TX+w_0=0$$

شکل ۱۶، نشان‌دهنده‌ی دو حالت خاص می‌باشد که در یکی ماتریس‌های کواریانس دو کلاس با یکدیگر نابرابر و در دیگری با یکدیگر برابر می‌باشند (در حالت دوم، مرز تصمیم‌گیری به صورت یک خط می‌باشد):



شکل ۱۶. سمت چپ: مرز تصمیم‌گیری در حالتی که $\Sigma_1 \neq \Sigma_2$؛ سمت راست: مرز تصمیم‌گیری در حالتی که $\Sigma_1 = \Sigma_2$. در حالت اول، مرز تصمیم‌گیری به صورت یک منحنی بوده و در حالت دوم به صورت یک خط می‌باشد.

جواب سوال ۱۰

کدهای مربوط این سوال در فایل‌های ex10 قرار دارد.

ω_1	ω_2
(1,1)	(2,2)
(1,2)	(3,2)
(1,3)	(3,4)
(2,1)	(5,1)
(3,1)	(5,4)
(3,3)	(5,5)

قسمت الف)

اگر جهت بردار w را جهت خط تصویرسازی در نظر بگیریم، سپس روش جداسازی خطی فیشر بهترین جهت را برای بردار w پیدا می‌کند که در آن جهت تابع ضابطه‌ی $J(w)=\frac{w^tS_Bw}{w^tS_ww}$ بیشینه گردد. لذا داریم:

$$w=S_w^{-1}(m_1-m_2)\quad \text{where}\quad S_w=S_1+S_2,\quad S_i=\sum_{x\in D_i}(x-m_1)(x-m_1)^t\quad i=1,2$$

Thus, we first compute the sample means for each class:

$$m_1=\frac{1}{6}\left[\begin{pmatrix}1\\1\end{pmatrix}+\begin{pmatrix}1\\2\end{pmatrix}+\begin{pmatrix}1\\3\end{pmatrix}+\begin{pmatrix}2\\1\end{pmatrix}+\begin{pmatrix}3\\1\end{pmatrix}+\begin{pmatrix}3\\3\end{pmatrix}\right]=\frac{1}{6}\begin{pmatrix}11\\11\end{pmatrix}\Rightarrow m_1=\begin{pmatrix}1.8333\\1.8333\end{pmatrix}$$

$$m_2=\frac{1}{6}\left[\begin{pmatrix}2\\2\end{pmatrix}+\begin{pmatrix}3\\2\end{pmatrix}+\begin{pmatrix}3\\4\end{pmatrix}+\begin{pmatrix}5\\1\end{pmatrix}+\begin{pmatrix}5\\4\end{pmatrix}+\begin{pmatrix}5\\5\end{pmatrix}\right]=\frac{1}{6}\begin{pmatrix}23\\18\end{pmatrix}\Rightarrow m_2=\begin{pmatrix}3.8333\\3\end{pmatrix}$$

Then, we subtract the sample mean from each sample and get:

$$D_1=x-m_1=\begin{bmatrix}-.8333 & -.8333 & -.8333 & .1667 & 1.1667 & 1.1667\\ -.8333 & .1667 & 1.1667 & -.8333 & -.8333 & 1.1667\end{bmatrix}\to S_1=D_1D_1^t=\begin{bmatrix}4.8333 & -.1667\\ -.1667 & 4.8333\end{bmatrix}$$

$$D_2=x-m_2=\begin{bmatrix}-1.8333 & -.8333 & -.8333 & 1.1667 & 1.1667 & 1.1667\\ -1 & -1 & 1 & -2 & 1 & 2\end{bmatrix}\to S_2=D_2D_2^t=\begin{bmatrix}8.8333 & 3\\ 3 & 12\end{bmatrix}$$

and then:

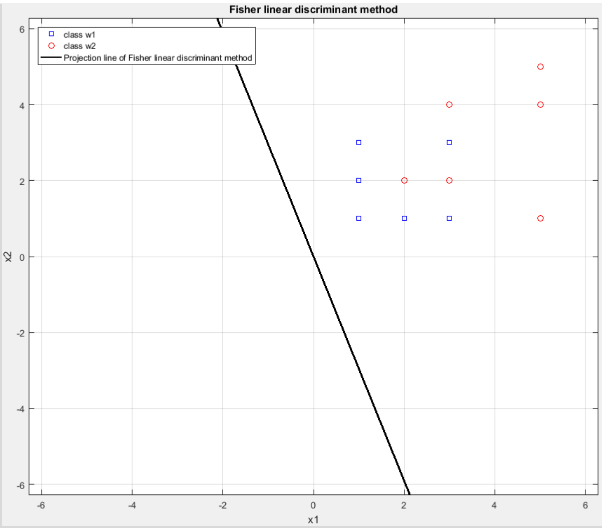
$$S_w=S_1+S_2=\begin{bmatrix}4.8333 & -.1667\\ -.1667 & 4.8333\end{bmatrix}+\begin{bmatrix}8.8333 & 3\\ 3 & 12\end{bmatrix}=\begin{bmatrix}13.6667 & 2.8333\\ 2.8333 & 16.8333\end{bmatrix}$$

$$S_w^{-1}=\frac{1}{|S_w|}=\frac{1}{222.0278}\begin{bmatrix}16.8333 & -2.8333\\ -2.8333 & 13.6667\end{bmatrix}=\begin{bmatrix}.0758 & -.0128\\ -.0128 & .0616\end{bmatrix}$$

Finally we have:

$$w=S_w^{-1}(m_1-m_2)=\begin{bmatrix}.0758 & -.0128\\ -.0128 & .0616\end{bmatrix}\begin{bmatrix}1.8333-3.8333\\ 1.8333-3\end{bmatrix}=\begin{bmatrix}-.1367\\ -.0463\end{bmatrix}$$

شکل ۱۷، توزیع داده‌ها را به همراه خط تصویرسازی بهینه با استفاده از روش جداسازی خطی فیشر[12] نشان می‌دهد:



شکل ۱۷. توزیع داده‌های دو کلاس به همراه خط تصویرسازی بهینه با استفاده از روش جداسازی خطی فیشر.

قسمت ب)

برای نگاشت داده‌ها بر روی خط نگاشت حاصله از روش جداسازی خطی فیشر، داریم:

[12] The optimal projection line using Fisher linear discriminant method

$$x' = w^t x \Rightarrow$$
$$w_1' = \{-0.1830, -0.2293, -0.2756, -0.3198, -0.4565, -0.5491\}$$
$$w_2' = \{-.3661, -.5028, -.5954, -.7300, -.8689, -.9152\}$$

Thus, we can compute the mean and variance of new projected data as follows:

$$\mu_1' = -.3356, \quad \mu_2' = -.6631, \quad \sigma_1' = .1404, \quad \sigma_2' = .2139$$

حال اگر تصور کنیم که هر دو کلاس از داده‌های نگاشت‌شده به فضای جدید دارای توزیع نرمال باشند و البته احتمال اولیه‌ی آن‌ها را نیز به دلخواه برابر با ۰٫۵ فرض نمائیم، آن‌گاه می‌توانیم از قانون تصمیم‌گیری بیزین با مینیمم خطا استفاده کنیم. داریم:

$$p(x|w_i') \sim N\left(\mu_i', \sigma_i'^2\right) = \frac{1}{\sqrt{2\pi}\sigma_i'} \exp\left\{-\frac{1}{2}\left(\frac{x-\mu_i'}{\sigma_i'}\right)^2\right\}, \quad P(w_i') = .5 \quad \rightarrow$$

$$p(w_1'|x) \mathop{\gtrless}_{w_2'}^{w_1'} p(w_2'|x) \Rightarrow p(x|w_1')P(w_1') \mathop{\gtrless}_{w_2'}^{w_1'} p(x|w_2')P(w_2') \Rightarrow \frac{p(x|w_1')}{p(x|w_2')} \mathop{\gtrless}_{w_2'}^{w_1'} \frac{P(w_2')}{P(w_1')}$$

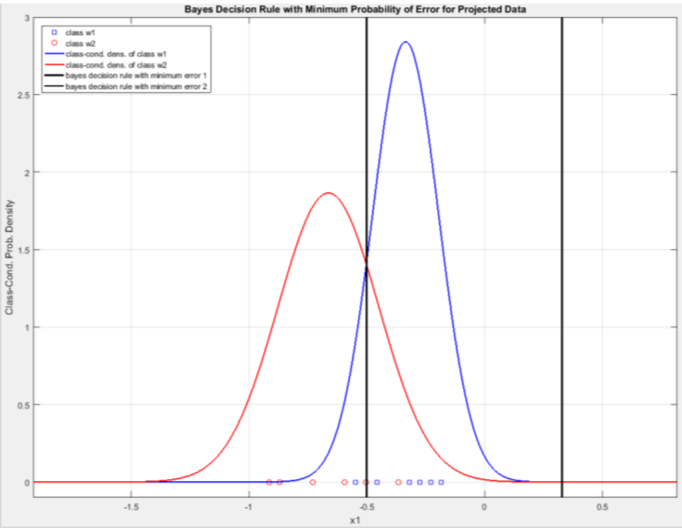
$$-\ln\left(\frac{p(x|w_1')}{p(x|w_2')}\right) \mathop{\lessgtr}_{w_2'}^{w_1'} -\ln\left(\frac{P(w_2')}{P(w_1')}\right) \Rightarrow \frac{1}{2}\left(\frac{x-\mu_1'}{\sigma_1'}\right)^2 - \frac{1}{2}\left(\frac{x-\mu_2'}{\sigma_2'}\right)^2 + \ln\left(\frac{\sigma_1'}{\sigma_2'}\right) \mathop{\lessgtr}_{w_2'}^{w_1'} 0 \Rightarrow$$

$$\frac{1}{2}\left(\frac{x+.3356}{.1404}\right)^2 - \frac{1}{2}\left(\frac{x+.6631}{.2139}\right)^2 + \ln\left(\frac{.1404}{.2139}\right) \mathop{\lessgtr}_{w_2'}^{w_1'} 0$$

Finally we conclude that the roots of this quadratic equation are:

$$\mathrm{x}_1 = -.5022, \quad \mathrm{x}_2 = .3270 \quad \rightarrow \quad \mathrm{x}_1 \le x \le \mathrm{x}_2 \Rightarrow x \in w_1' \quad \text{o.w.} \quad x \in w_2'$$

در شکل ۱۸، توزیع داده‌ها پس از نگاشت به فضای جدید به همراه مرز تصمیم‌گیری بیزین با مینیمم خطا نشان داده شده است:



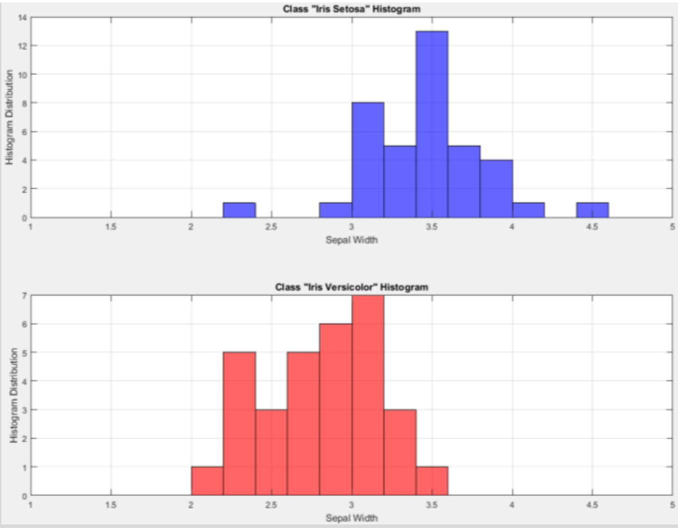
شکل ۱۸. توزیع داده‌های دو کلاس پس از نگاشت به خط تصویرسازی حاصله از روش جداسازی خطی فیشر به همراه مرز تصمیم‌گیری بیزین با مینیمم خطا؛ داده‌های مابین دو خط جداساز به کلاس اول نسبت داده شده و مابقی داده‌ها نیز به کلاس دوم نسبت داده خواهند شد.

جواب سوال ۱۱

کدهای مربوط این سوال در فایل‌های ex11 قرار دارد.

قسمت الف)

نمودار هیستوگرام مربوط به داده‌های دو کلاس به صورت زیر می‌باشد:



شکل ۱۹. نمودار هیستوگرام مربوط به داده‌های دو کلاس.

قسمت ب)

در جدول زیر مقادیر احتمال اولیه‌ی مربوط به دو کلاس Setosa و Versicolor را مشاهده می‌نمائیم:

Class Setosa a Priori Prob.:	**0.56**
Class Versicolor a Priori Prob.:	**0.44**

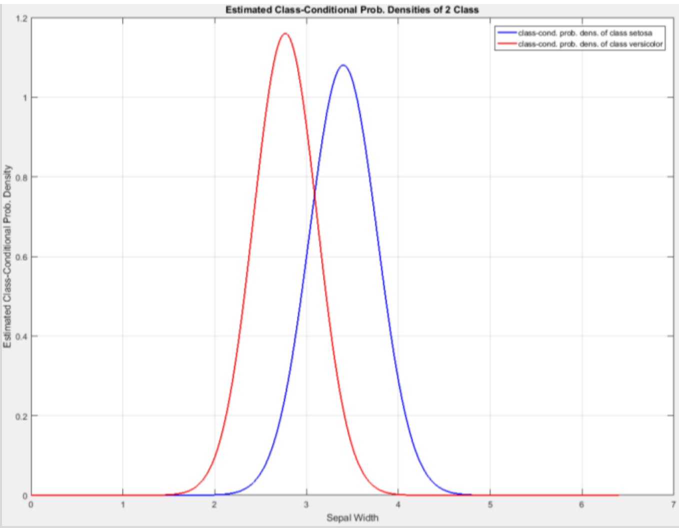
قسمت ج)

در جدول زیر مقادیر میانگین و انحراف از معیار مربوط به دو کلاس Setosa و Versicolor را با فرض داشتن توزیع نرمال مشاهده می‌نمائیم:

Class Setosa mean:	**3.400,**	**variance:**	**0.136**
Class Versicolor mean:	**2.768,**	**variance:**	**0.118**

قسمت د)

نمودار تابع چگالی احتمال مربوط به هر کدام از کلاس‌ها در شکل زیر نشان داده شده است:

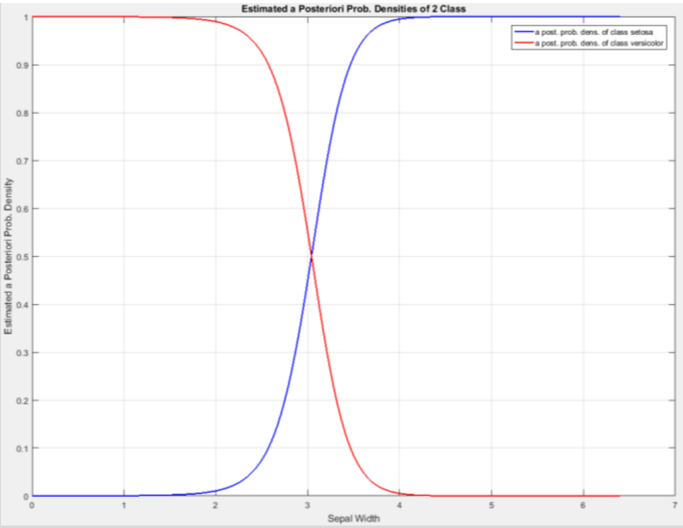


شکل ۲۰. نمودار تابع چگالی احتمال مربوط به هر کلاس؛ کلاس **Setosa** با رنگ آبی و کلاس **Versicolor** با رنگ قرمز.

جهت محاسبه‌ی تابع احتمال ثانویه‌ی[۱۳] هر کدام از کلاس‌ها در هر نقطه به صورت زیر عمل می‌کنیم:

$$p(\omega_i \mid x) = \frac{p(x \mid \omega_i)P(\omega_i)}{\sum_{j=1}^{2} p(x \mid \omega_j)P(\omega_j)}$$

[13] A Posteriori Probability



شکل ۲۱. نمودار تابع احتمال ثانویه‌ی مربوط به هر کلاس؛ کلاس **Setosa** با رنگ آبی و کلاس **Versicolor** با رنگ قرمز.

قسمت ه)

جهت به دست‌آوردن مرز تصمیم‌گیری بیزین با مینیمم خطا (با فرض اولیه‌ی نرمال‌بودن توزیع داده‌های دو کلاس) مانند سؤال ۱۰ قسمت (ب) عمل نموده و داریم:

$$p(x\,|w_i) \sim N\left(\mu_i,\sigma_i^2\right) = \frac{1}{\sqrt{2\pi}\sigma_i}\exp\left\{-\frac{1}{2}\left(\frac{x-\mu_i}{\sigma_i}\right)^2\right\},\quad P(w_i) = .5 \quad \rightarrow$$

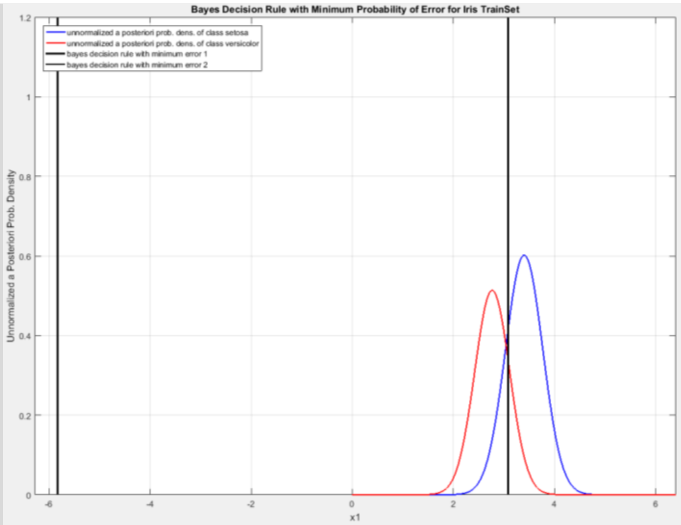
$$p(w_1\,|x) \mathop{\gtrless}_{w_2}^{w_1} p(w_2\,|x) \Rightarrow p(x\,|w_1)P(w_1) \mathop{\gtrless}_{w_2}^{w_1} p(x\,|w_2)P(w_2) \Rightarrow \frac{p(x\,|w_1)}{p(x\,|w_2)} \mathop{\gtrless}_{w_2}^{w_1} \frac{P(w_2)}{P(w_1)}$$

$$-\ln\left(\frac{p(x\,|w_1)}{p(x\,|w_2)}\right) \mathop{\lessgtr}_{w_2}^{w_1} -\ln\left(\frac{P(w_2)}{P(w_1)}\right) \Rightarrow \frac{1}{2}\left(\frac{x-\mu_1}{\sigma_1}\right)^2 - \frac{1}{2}\left(\frac{x-\mu_2}{\sigma_2}\right)^2 + \ln\left(\frac{\sigma_1}{\sigma_2}\right) \mathop{\lessgtr}_{w_2}^{w_1} 0 \Rightarrow$$

$$\frac{1}{2}\left(\frac{x+3.400}{.3692}\right)^2 - \frac{1}{2}\left(\frac{x+2.768}{.3439}\right)^2 + \ln\left(\frac{.3692}{.3439}\right) \mathop{\lessgtr}_{w_2}^{w_1} 0$$

Finally we conclude that the roots of this quadratic equation are:

$$\mathrm{x}_1 = -5.8326,\quad \mathrm{x}_2 = 3.0869 \quad \rightarrow \quad \mathrm{x}_1 \le x \le \mathrm{x}_2 \Rightarrow x \in w_2 \quad \text{o.w.} \quad x \in w_1$$



شکل ۲۲. نمودار توزیع احتمال ثانویه‌ی نرمال‌نشده به همراه مرزهای تصمیم‌گیری حاصل از قانون بیزین با مینیمم خطا؛ توزیع احتمال ثانویه‌ی کلاس **Setosa** با رنگ آبی و توزیع احتمال ثانویه‌ی کلاس **Versicolor** با رنگ قرمز مرزهای تصمیم‌گیری بیزین با مینیمم خطا با رنگ مشکی نشان داده شده‌اند؛ داده‌های مابین دو خط جداساز به کلاس دوم نسبت داده شده و مابقی داده‌ها نیز به کلاس اول نسبت داده خواهند شد.

در جدول زیر مقادیر خطای دسته‌بندی مربوط به داده‌های آموزشی قید شده است:

# of Class Setosa Misclassified Items:	**6 out of 39 Training Items**
# of Class Versicolor Misclassified Items:	**7 out of 31 Training Items**
Percentage of incorrectly classified Training Data Items:	**0.186**

قسمت و)

در جدول زیر مقادیر خطای دسته‌بندی مربوط به داده‌های آزمایشی قید شده است:

# of Class Setosa Misclassified Items:	**2 out of 11 Test Items**
# of Class Versicolor Misclassified Items:	**1 out of 19 Test Items**
Percentage of incorrectly classified Test Data Items:	**0.100**