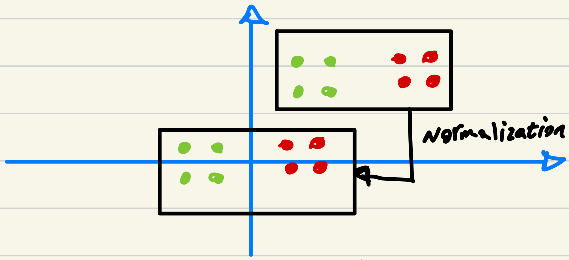
تمرین سری 7 درس ML - رادین خیام - 97101579

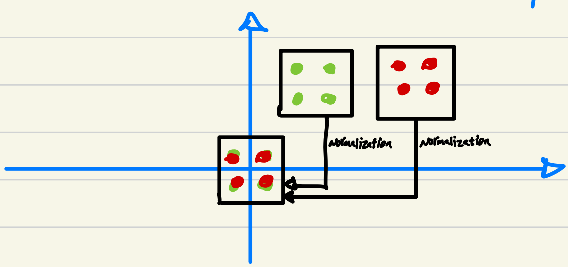
سوال ۱-

خیر کار درستی نیست که داده های هر کلاس را جداگانه نرمالایز کنیم.
مثلاً فرض کنید نرمالیزاسیون ما این باشد که میانگین را صفر کند حال
در ابتدا در نظر بگیرید که همه داده ها را با هم نرمالایز کنیم:



* میبینیم که تمایز پذیری بین دو کلاس حفظ شده است.

حال این حالت را در نظر بگیرید که برای هر کلاس به صورت جداگانه بخواهیم میانگین را صفر کنیم:



* مشاهده میکنیم که تمایز پذیری بین دو کلاس خراب شده است، پس اینکار درست نمی باشد.

ادا ـ سؤال ۱-

برای داده‌های تست هم نیاز هست تمام مراحلی که برای نرمالیزاسیون داده‌های آموزش طی شده است را طی کنیم چون مدل با استفاده از داده‌های نرمالیزه شده آموزش دیده است.

سؤال 2 -

$$X=\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} , \quad \Sigma_i = \begin{bmatrix} s_i & r_i \\ r_i & s_i' \end{bmatrix} , \quad \underline{\mu}_i = \begin{bmatrix} m_i \\ m_i' \end{bmatrix}$$

$$d_{ij} = \frac{1}{2}\,\text{trace}\left(\Sigma_i^{-1}\Sigma_j + \Sigma_j^{-1}\Sigma_i - 2I \right) + \frac{1}{2}(\mu_i - \mu_j)^T \left(\Sigma_i^{-1} + \Sigma_j^{-1} \right)(\mu_i - \mu_j)$$

الف)

$$d = \frac{1}{2}\,\text{trace}\left(\frac{1}{s_i s_i' - r_i^2}\begin{bmatrix} s_i' & -r_i \\ -r_i & s_i \end{bmatrix}\begin{bmatrix} s_j & r_j \\ r_j & s_j' \end{bmatrix} + \frac{1}{s_j s_j' - r_j^2}\begin{bmatrix} s_j' & -r_j \\ -r_j & s_j \end{bmatrix}\begin{bmatrix} s_i & r_i \\ r_i & s_i' \end{bmatrix} - \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \right)$$

$$+ \frac{1}{2}\begin{bmatrix} m_i - m_j \\ m_i' - m_j' \end{bmatrix}^T \left(\frac{1}{s_i s_i' - r_i^2}\begin{bmatrix} s_i' & -r_i \\ -r_i & s_i \end{bmatrix} + \frac{1}{s_j s_j' - r_j^2}\begin{bmatrix} s_j' & -r_j \\ -r_j & s_j \end{bmatrix} \right)\begin{bmatrix} m_i - m_j \\ m_i' - m_j' \end{bmatrix}$$

ب)

$$d_1 = \frac{1}{2}\left(\frac{s_j}{s_i} + \frac{s_i}{s_j} - 2 \right) + \frac{1}{2}(m_i - m_j)^2\left(\frac{1}{s_i} + \frac{1}{s_j} \right)$$

$$\rightarrow d_1 = \frac{1}{2}\left(\frac{s_j^2 + s_i^2 - 2 s_i s_j}{s_i s_j} + (m_i - m_j)^2 \left(\frac{s_i + s_j}{s_i s_j} \right) \right)$$

$$\rightarrow d_1 = \frac{1}{2 s_i s_j}\left((s_i - s_j)^2 + (m_i - m_j)^2 (s_i + s_j) \right)$$

ادامه سوال 2 -

پ) مطابق با آنچه در قسمت ب نوشتیم :

$$d_2 = \frac{1}{2s_i' s_j'}\left((s_i' - s_j')^2 + (m_i' - m_j')(s_i' + s_j')\right)$$

نتیجه: برای حالتی که $r \neq 0$ باشد که خب نمی‌شود خیلی رابطه خوبی نوشت بین d و d_1، d_2 ولی برای حالتی که $r = 0$ باشد :

$$d = \frac{1}{2}\operatorname{trace}\left(\frac{1}{s_i s_i'}\begin{bmatrix} s_i & 0 \\ 0 & s_i' \end{bmatrix}\begin{bmatrix} s_j & 0 \\ 0 & s_j' \end{bmatrix} + \frac{1}{s_j s_j'}\begin{bmatrix} s_j' & 0 \\ 0 & s_j \end{bmatrix}\begin{bmatrix} s_i & 0 \\ 0 & s_i' \end{bmatrix} - \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}\right)$$

$$+\frac{1}{2}\begin{bmatrix} m_i - m_j \\ m_i' - m_j' \end{bmatrix}^T \left(\frac{1}{s_i s_i'}\begin{bmatrix} s_i' & 0 \\ 0 & s_i \end{bmatrix} + \frac{1}{s_j s_j'}\begin{bmatrix} s_j' & 0 \\ 0 & s_j \end{bmatrix}\right)\begin{bmatrix} m_i - m_j \\ m_i' - m_j' \end{bmatrix}$$

$$\rightarrow d = \frac{1}{2}\operatorname{trace}\left(\begin{bmatrix} \frac{s_j}{s_i} & 0 \\ 0 & \frac{s_j'}{s_i'} \end{bmatrix} + \begin{bmatrix} \frac{s_i}{s_j} & 0 \\ 0 & \frac{s_i'}{s_j'} \end{bmatrix} - \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}\right)$$

$$+\frac{1}{2}\left((m_i - m_j)^2\left(\frac{1}{s_i} + \frac{1}{s_j}\right) + (m_i' - m_j')^2\left(\frac{1}{s_i'} + \frac{1}{s_j'}\right)\right)$$

$$\rightarrow d = \frac{1}{2s_i s_j}\left(s_i^2 + s_j^2 - 2s_i s_j\right) + \frac{1}{2s_i' s_j'}\left(s_i'^2 + s_j'^2 - 2s_i' s_j'\right) + \frac{(m_i - m_j)^2 (s_i + s_j)}{2s_i s_j}$$

$$+\frac{(m_i' - m_j')^2 (s_i' + s_j')}{2s_i' s_j'}$$

$$\longrightarrow \boxed{d = d_1 + d_2}$$

- سؤال 3

$$\underline{x}-\underline{\mu}_0 = \underline{x}-\underline{\mu}_0+\underline{\mu}_i-\underline{\mu}_i \rightarrow (\underline{x}-\underline{\mu}_0)(\underline{x}-\underline{\mu}_0)^T = (\underline{x}-\underline{\mu}_i)(\underline{x}-\underline{\mu}_i)^T+(\underline{\mu}_i-\underline{\mu}_0)(\underline{\mu}_i-\underline{\mu}_0)^T$$

$$+2(\underline{x}-\underline{\mu}_i)(\underline{\mu}_i-\underline{\mu}_0)^T$$

$$S_m = E\left\{(\underline{x}-\underline{\mu}_0)(\underline{x}-\underline{\mu}_0)^T\right\} = \sum_{i=1}^{m} E\left\{(\underline{x}-\underline{\mu}_0)(\underline{x}-\underline{\mu}_0)^T \mid \underline{x}\in w_i\right\} P(w_i)$$

$$\rightarrow S_m = \sum_{i=1}^{m} E\left\{(\underline{x}-\underline{\mu}_i)(\underline{x}-\underline{\mu}_i)^T \mid \underline{x}\in w_i\right\} P(w_i) + E\left\{(\underline{\mu}_i-\underline{\mu}_0)(\underline{\mu}_i-\underline{\mu}_0)^T \mid \underline{x}\in w_i\right\} P(w_i)$$

$$+2E\left\{\cancelto{0}{(\underline{x}-\underline{\mu}_i)(\underline{\mu}_i-\underline{\mu}_0)^T} \mid \underline{x}\in w_i\right\} P(w_i)$$

$$\rightarrow S_m = \sum_{i=1}^{m}\left(\Sigma_i P(w_i) + (\underline{\mu}_i-\underline{\mu}_0)(\underline{\mu}_i-\underline{\mu}_0)^T P(w_i)\right)$$

$$\rightarrow S_m = \sum_{i=1}^{m}\Sigma_i P(w_i) + \sum_{i=1}^{m}(\underline{\mu}_i-\underline{\mu}_0)(\underline{\mu}_i-\underline{\mu}_0)^T P(w_i)$$

$$\rightarrow \boxed{S_m = S_w + S_b}$$

سوال 4-

الف)

معیار واگرایی برای تک ویژگی:

$$d_{ij} = \frac{1}{2}\left(\frac{\sigma_j^2}{\sigma_i^2} + \frac{\sigma_i^2}{\sigma_j^2} - 2\right) + \frac{1}{2}(\mu_i - \mu_j)^2\left(\frac{1}{\sigma_i^2} + \frac{1}{\sigma_j^2}\right)$$

$$x_1 \to \begin{cases} d_{12} = 1 \\ d_{13} = 3.25 \\ d_{23} = 1 \end{cases} \to \sum d_{ij} = \underline{5.25} \quad , \quad x_2 \to \begin{cases} d_{12} = 3.25 \\ d_{13} = 0.5 \\ d_{23} = 1 \end{cases} \to \sum d_{ij} = 4.75$$

$$x_3 \to \begin{cases} d_{12} = 3.54 \\ d_{13} = 1.75 \\ d_{23} = 3.95 \end{cases} \to \sum d_{ij} = \underline{9.24} \quad , \quad x_4 \to \begin{cases} d_{12} = 3.33 \\ d_{13} = 1.33 \\ d_{23} = 1 \end{cases} \to \sum d_{ij} = \underline{5.66}$$

$$\boxed{x_3 > x_4 > x_1 > x_2}$$

* بر اساس این معیار x_3 بهترین ویژگی است.

ب)

معیار فاصله باتاچاریا:

$$B_{ij} = \frac{1}{8}(\mu_i - \mu_j)^2\left(\frac{\sigma_i^2 + \sigma_j^2}{2}\right) + \frac{1}{2}\ln\frac{\frac{\sigma_i^2 + \sigma_j^2}{2}}{\sqrt{\sigma_i^2\sigma_j^2}}$$

$$x_1 \to \begin{cases} B_{12} = 0.125 \\ B_{13} = 0.78 \\ B_{23} = 0.22 \end{cases} \to \sum B_{ij} = 1.12 \quad , \quad x_2 \to \begin{cases} B_{12} = 0.78 \\ B_{13} = 0.25 \\ B_{23} = 0.22 \end{cases} \to \sum B_{ij} = 1.25$$

$$x_3 \to \begin{cases} B_{12} = 19.1 \\ B_{13} = 1.53 \\ B_{23} = 4.47 \end{cases} \to \sum B_{ij} = 25.1 \quad , \quad x_4 \to \begin{cases} B_{12} = 0.07 \\ B_{13} = 0.32 \\ B_{23} = 0.125 \end{cases} \to \sum B_{ij} = 1.52$$

$$\boxed{x_3 > x_4 > x_2 > x_1}$$

* ویژگی x_3 بهترین ویژگی است.

پ)

$$F_{ij} = \frac{S_b}{S_w} = \frac{(\mu_i - \mu_j)^2}{\sigma_i^2 + \sigma_j^2}$$

$$x_1 \to \begin{cases} F_{12} = 0.5 \\ F_{13} = 1.33 \\ F_{23} = 0.33 \end{cases} \to \sum F_{ij} = 2.17 \quad , \quad x_2 \to \begin{cases} F_{12} = 1.33 \\ F_{13} = 0.25 \\ F_{23} = 0.33 \end{cases} \to \sum F_{ij} = 1.92$$

$$x_3 \to \begin{cases} F_{12} = 0.84 \\ F_{13} = 0.66 \\ F_{23} = 0.24 \end{cases} \to \sum F_{ij} = 1.74 \quad , \quad x_4 \to \begin{cases} F_{12} = 1 \\ F_{13} = 0.25 \\ F_{23} = 0.5 \end{cases} \to \sum F_{ij} = 1.75$$

بنابراین ویژگی x_1 بهترین است.

$$\boxed{x_1 > x_2 > x_4 > x_3}$$

Made with Goodnotes

سؤال 5 -

$$\mu_1 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad \mu_2 = \begin{bmatrix} 3 \\ -3 \end{bmatrix}, \quad \Sigma = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}, \quad P(\omega_1) = 3P(\omega_2)$$

الف)

$$g_{ij} = \left(\Sigma^{-1}(\underline{\mu}_i - \underline{\mu}_j)\right)^T \underline{x} + \ln\left(\frac{P(\omega_i)}{P(\omega_j)}\right) - \frac{1}{2}\underline{\mu}_i^T \Sigma^{-1} \underline{\mu}_i + \frac{1}{2}\underline{\mu}_j^T \Sigma^{-1} \underline{\mu}_j$$

$$\rightarrow g_{12} = \left(\frac{1}{3}\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}\begin{bmatrix} -3 \\ 3 \end{bmatrix}\right)^T \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \ln(3) + \frac{1}{2}\begin{bmatrix} 3 & -3 \end{bmatrix} \times \frac{1}{3}\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} \times \begin{bmatrix} 3 \\ -3 \end{bmatrix} =$$

$$= -3x_1 + 3x_2 + \ln(3) + 9 \simeq \boxed{-3x_1 + 3x_2 + 10.1 = 0}$$

ب)

$$|\Sigma - \lambda I| = \begin{vmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{vmatrix} = 4 + \lambda^2 - 4\lambda - 1 = \lambda^2 - 4\lambda + 3 = 0 \rightarrow \begin{matrix} \lambda_1 = 3 \\ \lambda_2 = 1 \end{matrix}$$

$$\Sigma \underline{u_1} = \lambda_1 \underline{u_1} \rightarrow \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} u_{11} \\ u_{12} \end{bmatrix} = 3\begin{bmatrix} u_{11} \\ u_{12} \end{bmatrix} \rightarrow \begin{matrix} 2u_{11} + u_{12} = 3u_{11} \\ u_{11} + 2u_{12} = 3u_{12} \end{matrix} \rightarrow u_{11} = u_{12}$$

$$\rightarrow \underline{u}_1 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} \rightarrow \underline{y} = \underline{u}_1^T \underline{x} \rightarrow \underline{\mu}_{y_1} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}\begin{bmatrix} 0 \\ 0 \end{bmatrix} = \boxed{0}$$

$$\underline{\mu}_{y_2} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}\begin{bmatrix} 3 \\ -3 \end{bmatrix} = \boxed{0}, \quad \Sigma_y = \underline{u}_1^T \Sigma \underline{u}_1 = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} = \boxed{3}$$

$$g_{12} = \frac{1}{3} \times 0 \times \underline{y} + \ln(3) + 0 = \ln(3) \; \text{✗}.$$

✗ تفکیک پذیری حفظ نشده است.

پ)

$$\Sigma \underline{u}_2 = \lambda_2 \underline{u}_2 \rightarrow \begin{bmatrix}2&1\\1&2\end{bmatrix}\begin{bmatrix}u_{21}\\u_{22}\end{bmatrix} = \begin{bmatrix}u_{21}\\u_{22}\end{bmatrix} \rightarrow \begin{matrix}2u_{21}+u_{22}=u_{21}\\u_{21}+2u_{22}=u_{22}\end{matrix} \rightarrow u_{21}=-u_{22}$$

$$\rightarrow \underline{u}_2 = \begin{bmatrix}\frac{1}{\sqrt2}\\-\frac{1}{\sqrt2}\end{bmatrix} \rightarrow \underline{y} = \underline{u}_2^T \underline{x} \rightarrow \underline{\mu}_{y_1} = \begin{bmatrix}\frac{1}{\sqrt2} & \frac{-1}{\sqrt2}\end{bmatrix}\begin{bmatrix}0\\0\end{bmatrix} = 0$$

$$\underline{\mu}_{y_2} = \begin{bmatrix}\frac{1}{\sqrt2} & \frac{-1}{\sqrt2}\end{bmatrix}\begin{bmatrix}3\\-3\end{bmatrix} = 3\sqrt2 \quad , \quad \Sigma_y = \underline{u}_2^T \Sigma \underline{u}_2 = \begin{bmatrix}\frac{1}{\sqrt2} & \frac{-1}{\sqrt2}\end{bmatrix}\begin{bmatrix}2&1\\1&2\end{bmatrix}\begin{bmatrix}\frac{1}{\sqrt2}\\-\frac{1}{\sqrt2}\end{bmatrix} = 1$$

$$\rightarrow g_{12} = 1\times(0-3\sqrt2)\, y + \ln(3) + \frac{1}{2}\times(3\sqrt2)^2 = -3\sqrt2\, y + \ln(3) + 9$$

$$\rightarrow \boxed{-3\sqrt2\, y + 10.1 = 0}$$

* تفکیک پذیری حفظ شده است.

ت)

$$\underline{w} = \alpha\, \Sigma^{-1}(\underline{\mu}_1 - \underline{\mu}_2) = \frac{\alpha}{3}\begin{bmatrix}2&-1\\-1&2\end{bmatrix}\begin{bmatrix}-3\\3\end{bmatrix} = \alpha\begin{bmatrix}-3\\3\end{bmatrix} \rightarrow \alpha = \frac{1}{\sqrt{18}} \rightarrow \underline{w} = \begin{bmatrix}-1\\1\end{bmatrix}$$

$$\rightarrow \mu_{y_1} = 0 \; , \; \mu_{y_2} = -6 \; , \; \Sigma_y = \begin{bmatrix}-1 & 1\end{bmatrix}\begin{bmatrix}2&1\\1&2\end{bmatrix}\begin{bmatrix}-1\\1\end{bmatrix} = 2 \rightarrow \boxed{y = -3 - \frac{1}{3}\ln 3}$$

* تفکیک پذیری حفظ می‌شود.

ث)

$$\underline{X} = \begin{bmatrix}4\\0\end{bmatrix}$$

برای حالت الف) $g_{12}(\underline{X}) = -3x_1 + 3x_2 + 10.1 = -1.9 < 0 \rightarrow \underline{X} \in \omega_2$

برای حالت ب) خیلی معنایی نمی‌دهد اصلن!

برای حالت پ) $y = \begin{bmatrix}\frac{1}{\sqrt2} & -\frac{1}{\sqrt2}\end{bmatrix}\begin{bmatrix}4\\0\end{bmatrix} = 2\sqrt2 \; , \; g_{12} = -3\sqrt2 \times 2\sqrt2 + 10.1 = -1.9 < 0 \rightarrow \underline{X} \in \omega_2$

حالت ت) $y = \begin{bmatrix}-1 & 1\end{bmatrix}\begin{bmatrix}4\\0\end{bmatrix} = -4 \rightarrow g_{12} = 3\times -4 + 10.1 = -1.9 < 0 \rightarrow \underline{X} \in \omega_2$

ج)

$$FDR = J = \frac{(\mu_1-\mu_2)^2}{\sigma_1^2+\sigma_2^2}$$

$$x_1 : \frac{(0-3)^2}{2+2} = \frac{9}{4}$$

$$x_2 : \frac{(0+3)^2}{2+2} = \frac{9}{4}$$

← این معیار برای هر 2 ویژگی یکی می‌شود بنابراین نمی‌توانیم هیچ کدام را برتر بدانیم.

سوال 6 -

$$J_3 = trace(S_w^{-1} S_m) = \sum_{i=1}^{3} \sum_{j=1}^{3} \frac{(\mu_i-\mu_j)^2}{\sigma_i^2+\sigma_j^2}$$

← محاسبه J_3 برای شکل چهارم :

$\sigma_i^2+\sigma_j^2$ برای شکل 4 مشابه با شکل 2 و برای شکل 3 مشابه شکل 1 می‌باشد .

$(\mu_i-\mu_j)^2$ برای شکل 4 مشابه با شکل 3 و برای شکل 2 مشابه شکل 1 است .

$$(\sigma_i+\sigma_j)_1 = (\sigma_i^2+\sigma_j^2)_3 \quad , \quad (\sigma_i^2+\sigma_j^2)_2 = (\sigma_i^2+\sigma_j^2)_4$$

$$(\mu_i-\mu_j)_1^2 = (\mu_i-\mu_j)_2^2 \quad , \quad (\mu_i-\mu_j)_3^2 = (\mu_i-\mu_j)_4^2$$

$$(J_3)_4 = \sum_{i=1}^{3} \sum_{j=1}^{3} \frac{(\mu_i-\mu_j)_3^2}{(\sigma_i^2+\sigma_j^2)_2} = \sum_{i=1}^{3} \sum_{j=1}^{3} \frac{(\mu_i-\mu_j)_3^2 \, (\mu_i-\mu_j)_2^2 \, \overset{(\sigma_i^2+\sigma_j^2)_1}{\cancel{(\sigma_i^2+\sigma_j^2)_3}}}{(\sigma_i^2+\sigma_j^2)_2 \, \underset{(\mu_i-\mu_j)_1^2}{\cancel{(\mu_i-\mu_j)_2^2}} \, (\sigma_i^2+\sigma_j^2)_3}$$

$$\rightarrow (J_3)_4 = \frac{(J_3)_2 \, (J_3)_3}{(J_3)_1} = \frac{12 \times 620}{165} \simeq \boxed{45}$$

سوال 7 -

فرض کنید که m ویژگی داریم و می‌خواهیم یک زیر مجموعه k تایی از آن ها

را انتخاب کنیم که معیار ما را بهینه می‌کند. فرض کنید که $X_k=\{x_1,x_2,\dots,x_k\}$

مجموعه بهترین k ویژگی هست و Y_{m-k} مجموعه $m-k$ ویژگی دیگه هست.

در هر مرحله کاری که می‌کنیم این هست که X_{k+1} را اینگونه تشکیل می‌دهیم

که دونه دونه ویژگی‌های داخل Y_{m-k} را به X_k اضافه کنیم تا ببینیم به ازای کدوم

معیار ماکسیمم می‌شود. حال که X_{k+1} تشکیل شد میاییم دونه دونه حذف

می‌کنیم ویژگی ها را تا ببینیم به ازای حذف کدام یکی معیار افزایش پیدا می‌کند

اگر به ازای همون ویژگی آخری که اضافه شده بود این اتفاق افتاد، الگوریتم

را ادامه می‌دهیم و دوباره مرحله اول را تکرار می‌کنیم. اگر ویژگی‌ای که حذفش معیار

را بهبود بخشید ویژگی آخر نبود آن ویژگی را حذف می‌کنیم یعنی به Y_{m-k} اضافه

می‌کنیم و دوباره مرحله اول را تکرار می‌کنیم. اینکار را تا زمانی که k ویژگی بدست

بیاید ادامه می‌دهیم.

سؤال 8-

$$(XX^H)v = \lambda v \longrightarrow (X^H X)(\underbrace{X^H v}_{v'}) = \lambda(\underbrace{X^H v}_{v'})$$

بنابراین مقادیر ویژه XX^H و $X^H X$ یکسان می باشند.

سؤال 9-

$\omega_1: \ \mu_1 = 2 \quad \sigma_1 = 1$

$\omega_2: \ \mu_2 = 4 \quad \sigma_2 = 1$

$$B_{12} = \frac{1}{8}(\mu_1 - \mu_2)^2 \left(\frac{2}{\sigma_1^2 + \sigma_2^2}\right) + \frac{1}{2}\ln\left(\frac{\sigma_1^2 + \sigma_2^2}{2\sigma_1\sigma_2}\right)$$

$$= \frac{1}{8}(2)^2\left(\frac{2}{2}\right) + \cancel{\frac{1}{2}\ln\left(\frac{2}{2}\right)}^{0} = \boxed{\frac{1}{2}}$$

$$\sqrt{P(\omega_1)P(\omega_2)}\, e^{-B_{12}} = \sqrt{\frac{1}{2}\times\frac{1}{2}}\, e^{-\frac{1}{2}} = \frac{1}{2}e^{-\frac{1}{2}} \simeq \boxed{0.3}$$

سوال 10

الف)

$$w_1 : \begin{bmatrix}4\\1\end{bmatrix}, \begin{bmatrix}1\\0\end{bmatrix} \quad , \quad w_2 : \begin{bmatrix}2\\3\end{bmatrix}, \begin{bmatrix}5\\4\end{bmatrix}$$

$$\mu_1 = \begin{bmatrix}5/2\\1/2\end{bmatrix} , \ \mu_2 = \begin{bmatrix}7/2\\7/2\end{bmatrix} \longrightarrow \mu_0 = \begin{bmatrix}3\\2\end{bmatrix}$$

$$\Sigma = \frac{1}{4}(X-\mu_0)(X-\mu)^T = \begin{bmatrix}2.5 & 1.5\\1.5 & 2.5\end{bmatrix}$$

$$|\Sigma - \lambda I| = \left|\begin{bmatrix}2.5-\lambda & 1.5\\1.5 & 2.5-\lambda\end{bmatrix}\right| = 6.25+\lambda^2-5\lambda-2.25 = \lambda^2-5\lambda+4=0$$

$$\longrightarrow \lambda_1 = 4 \ , \ \lambda_2 = 1$$

$$\Sigma u_1 = \lambda_1 u_1 \longrightarrow \begin{bmatrix}2.5 & 1.5\\1.5 & 2.5\end{bmatrix}\begin{bmatrix}u_{11}\\u_{12}\end{bmatrix} = 4\begin{bmatrix}u_{11}\\u_{12}\end{bmatrix} \longrightarrow \underline{u}_1 = \begin{bmatrix}1\\1\end{bmatrix}$$

$$\longrightarrow y = \underline{u}_1^T \underline{X} \longrightarrow \begin{matrix}\mu_{y_1} = \underline{u}_1^T \mu_1 = 3\\ \mu_{y_2} = \underline{u}_1^T \mu_2 = 7\end{matrix} \quad , \ \Sigma_y = \underline{u}_1^T \Sigma \underline{u}_1 = 8$$

* بله در این راستا داده ها تمایز پذیر هستند و مرز را هم می توانیم $y=5$ در نظر بگیریم.

$$\Sigma u_2 = \lambda_2 u_2 \longrightarrow \begin{bmatrix}2.5 & 1.5\\1.5 & 2.5\end{bmatrix}\begin{bmatrix}u_{21}\\u_{22}\end{bmatrix} = 1\begin{bmatrix}u_{21}\\u_{22}\end{bmatrix} \longrightarrow \underline{u}_2 = \begin{bmatrix}-1\\1\end{bmatrix}$$

$$\longrightarrow \begin{matrix}\mu_{y_1} = -2\\ \mu_{y_2} = 0\end{matrix} \quad , \ \Sigma_y = 2$$

* در این راستا هم تمایز پذیری را داریم و می توانیم مرز را $y=-1$ در نظر بگیریم.

ب) کلا هر بردار ثابتی را اگر به همه داده ها اضافه کنیم مؤلفه اساسی های ما تغییر نمی کند چون اضافه کردن همه داده ها با یک بردار انگار که همه داده ها را صرفاً شیفت دادیم و تأثیری در پراکندگی آنها ندارد.

پ)

در ابتدا میانگین را صفر می کنیم:

$$w_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} -2 \\ -2 \end{bmatrix} \quad , \quad w_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 2 \end{bmatrix}$$

$$X = \begin{bmatrix} 1 & -2 & -1 & 2 \\ -1 & -2 & 1 & 2 \end{bmatrix} \rightarrow X^T X = \begin{bmatrix} 2 & 0 & -2 & 0 \\ 0 & 8 & 0 & -8 \\ -2 & 0 & 2 & 0 \\ 0 & -8 & 0 & 8 \end{bmatrix}$$

$$\xrightarrow{\text{مقادیر ویژه}} \quad \lambda_1 = 16, \ \lambda_2 = 4, \ \lambda_3 = \lambda_4 = 0 \qquad v_1 = \begin{bmatrix} 0 \\ -1 \\ 0 \\ 1 \end{bmatrix}, \ v_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \ v_3 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \ v_4 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

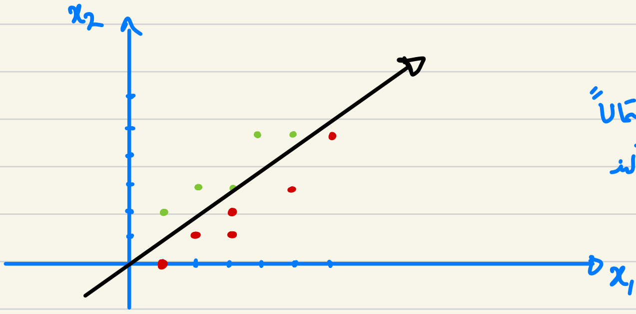
$$XX^T = \begin{bmatrix} 10 & 6 \\ 6 & 10 \end{bmatrix} \xrightarrow{\text{مقادیر ویژه}} \lambda_1 = 16, \ \lambda_2 = 4 \qquad u_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \ u_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

$$\rightarrow X = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 4 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & -1 & 0 & 1 \\ -1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix}$$

$$= \begin{bmatrix} \underbrace{1}_{\text{راستای اول}} & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 0 & -1 & 0 & 1 \\ -1 & 0 & 1 & 0 \end{bmatrix}$$

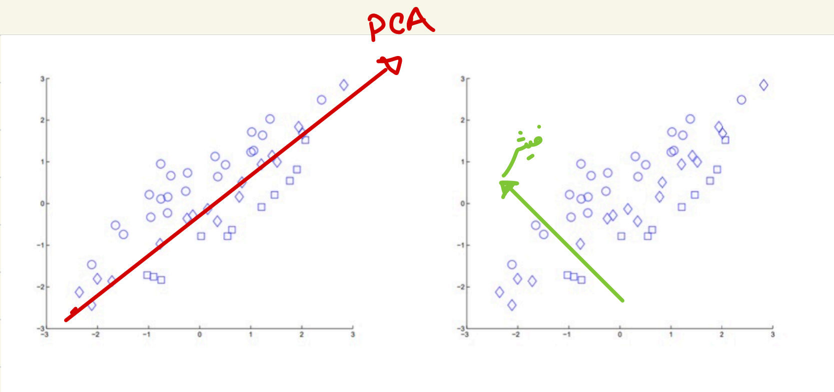
سؤال 11 -

الف)



PCA خیلی به کار نمیاد چون احتمالاً این راستای مشکی رو انتخاب می کند چون بیشترین واریانس را در این راستا داریم و همانطور که مشخص است با تصویر کردن داده های دو کلاس در این راستا تمایز پذیری بین دو کلاس را از دست خواهیم داد.

ب)



* در حالت اولیه که داده ها خیلی خوب جدایی پذیر نیستند، در حالت کاهش بعد با PCA هم که اصلاً بدتر هم می شود وضعیت جدایی پذیری، اما با استفاده از روش فیشر می توانیم تا حدود خوبی جدایی پذیری را بین دو کلاس مشاهده کنیم.