HW١

Q2-

Part A)

$$p\left(\begin{bmatrix}a\\b\end{bmatrix}\right) = p\left(\begin{bmatrix}-a\\-b\end{bmatrix}\right) = p\left(\begin{bmatrix}-c\\d\end{bmatrix}\right) = p\left(\begin{bmatrix}c\\-d\end{bmatrix}\right) = \frac{1}{4}$$

$$m = \sum_{i=1}^{4} x_i p(x_i) = \frac{1}{4}\begin{bmatrix}a\\b\end{bmatrix} + \frac{1}{4}\begin{bmatrix}-a\\-b\end{bmatrix} + \frac{1}{4}\begin{bmatrix}-c\\d\end{bmatrix} + \frac{1}{4}\begin{bmatrix}c\\-d\end{bmatrix} = \begin{bmatrix}0\\0\end{bmatrix}$$

$$\Sigma = \sum_{i=1}^{4}(x_i - m)(x_i - m)^T = \frac{1}{4}\begin{bmatrix}a\\b\end{bmatrix}\begin{bmatrix}a & b\end{bmatrix} + \frac{1}{4}\begin{bmatrix}-a\\-b\end{bmatrix}\begin{bmatrix}-a & -b\end{bmatrix} + \frac{1}{4}\begin{bmatrix}-c\\d\end{bmatrix}\begin{bmatrix}-c & d\end{bmatrix} + \frac{1}{4}\begin{bmatrix}c\\-d\end{bmatrix}\begin{bmatrix}c & -d\end{bmatrix} =$$

$$\frac{1}{4}\begin{bmatrix}2a^2 + 2c^2 & 2ab - 2cd\\2ab - 2cd & 2b^2 + 2d^2\end{bmatrix}$$

Part B)

$$2ab - 2cd = \rho\sqrt{2a^2 + 2c^2}\sqrt{2b^2 + 2d^2}$$

$$\rho = \frac{2ab - 2cd}{\sqrt{2a^2+2c^2}\sqrt{2b^2+2d^2}} = 0 \quad \rightarrow \quad 2ab - 2cd = 0 \quad \rightarrow \quad ab = cd$$

Part C)

If $\rho = \pm 1 \rightarrow \; 4(ab - cd)^2 = 4(a^2 + c^2)(b^2 + d^2) \rightarrow \quad a^2d^2 + b^2c^2 + 2abcd = 0 = (ad + bc)^2$

$$ad = -bc \quad \begin{cases}\rho = 1: & ab > cd\\ \rho = -1: & ab < cd\end{cases}$$

Q۳-

Part A)

Mahalanobis Distance is invariant: $d_M^2(Y) = (Y - M_Y)^T \Sigma_Y^{-1}(Y - M_Y) = (A^T X - A^T M_X)^T A^{-1}\Sigma_X^{-1}A^{-T}(A^T X - A^T M_X) =$

$(X - M_X)^T \Sigma_X^{-1}(X - M_X) = d_M^2(X)$

Part B)

Euclidean distance is variant: $d_E^2(Y) = (Y - M_Y)^T(Y - M_Y) = (A^T X - A^T M_X)^T(A^T X - A^T M_X) \neq d_E^2(X)$

Part C)

$p(Y) = \frac{1}{(2\pi)^{\frac{n}{2}}|\Sigma_Y|^{\frac{1}{2}}}\exp\left(-\frac{1}{2}d_M^2(Y)\right) = \frac{1}{(2\pi)^{\frac{n}{2}}|A||\Sigma_X|^{\frac{1}{2}}}\exp\left(-\frac{1}{2}d_M^2(X)\right) = \frac{p(X)}{|A|}$ where

$M_Y = A^T M_X$

$\Sigma_Y = A^T \Sigma_X A$

Q٤-

Part A)

$\lambda_1, \lambda_2 = 1{,}3 \qquad \phi_1, \phi_2 = [\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}]^T, [\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}]^T$

Part B)

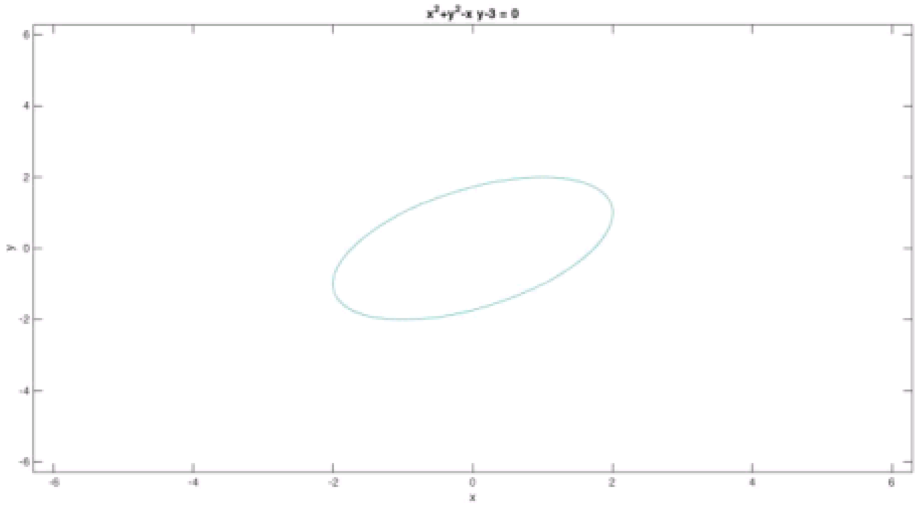
$$\Sigma = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix} \rightarrow \sigma_1 = \sigma_2 = \sqrt{2},\ \rho = 0.5$$

$$d_M^2(X) = (X)^T \Sigma_X^{-1}(X) = [x_1 \quad x_2]\begin{bmatrix} \frac{2}{3} & \frac{-1}{3} \\ \frac{-1}{3} & \frac{2}{3} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 2$$

$\frac{2}{3}(x_1^2 + x_2^2 - x_1x_2) = 2 \rightarrow x_1^2 + x_2^2 - x_1x_2 = 3 \rightarrow$ one standard deviation contour

Part C)

From MATLAB: ezplot('x^2+y^2-x*y-3');



Part D)

Eigen vectors show the direction of variation of data (they are orthogonal to each other). The variation is larger in the direction of the eigenvector that corresponds to the larger eigenvalue.

Part E)

The trace of matrix is equals to the summation of eigenvalues = 4.

Part F)

The determinant of matrix is equals to the product of the eigenvalues = 3.

Part G)

$$\Sigma = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} - \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \; \lambda_1, \lambda_2 = 0, -2 \qquad \phi_1, \phi_2 = [\tfrac{1}{\sqrt{2}}, \tfrac{1}{\sqrt{2}}]^T, [\tfrac{1}{\sqrt{2}}, -\tfrac{1}{\sqrt{2}}]^T$$

$\lambda_{new} = \lambda_A - 3$ eigenvectors did not change.

Part H)

To diagonalize the matrix we should calculate $\Phi^{-1}A\Phi = 0.5 \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} = 0.5 \begin{bmatrix} 1 & -1 \\ 3 & 3 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} =$

$0.5 \begin{bmatrix} 2 & 0 \\ 0 & 6 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix} = \Lambda$

Q٥-

$$y = 3u^2 - 2v$$

$$E(u^2) = \int_0^1 u^2\, du = \frac{1}{3} \qquad E(u^4) = \int_0^1 u^4\, du = \frac{1}{5}$$

$$Var(u^2) = E(u^4) - E^2(u^2) = \frac{1}{5} - \frac{1}{9} = \frac{4}{45}$$

$$E(v) = \int_0^1 v\, du = \frac{1}{2} \qquad E(v^2) = \int_0^1 v^2\, du = \frac{1}{3}$$

$$Var(v) = E(v^2) - E^2(v) = \frac{1}{3} - \frac{1}{4} = \frac{1}{12}$$

$$E(y) = E(3u^2 - 2v) = 3E(u^2) - 2E(v) = 3\left(\frac{1}{3}\right) - 2\left(\frac{1}{2}\right) = 0$$

$$Var(y) = Var(3u^2 - 2v) = 9Var(u^2) + 4Var(v) = 9\left(\frac{4}{45}\right) + 4\left(\frac{1}{12}\right) = \frac{4}{5} + \frac{1}{3} = \frac{17}{15}$$

Q٦-

No: $\left(\Phi\Lambda^{-\frac{1}{2}}\right)^T \left(\Phi\Lambda^{-\frac{1}{2}}\right) = \Lambda^{-\frac{1}{2}}\Phi^T\Phi\Lambda^{-\frac{1}{2}} = \Lambda^{-1} \neq I$

Q٧-

Part A)

$$p(x_1) = \int_{-\infty}^{+\infty} p(x_1, x_2)dx_2 = \int_{-\infty}^{+\infty} \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp(\frac{-1}{2} d^2(x_1, x_2))dx_2$$

$$d^2(x_1, x_2) = \frac{1}{\sigma_1^2\sigma_2^2(1-\rho^2)} [x_1 - m_1 \quad x_2 - m_2] \begin{bmatrix} \sigma_2^2 & -\rho\sigma_1\sigma_2 \\ -\rho\sigma_1\sigma_2 & \sigma_1^2 \end{bmatrix} \begin{bmatrix} x_1 - m_1 \\ x_2 - m_2 \end{bmatrix} =$$

$$\frac{1}{(1-\rho^2)} \left(\left(\frac{x_1-m_1}{\sigma_1}\right)^2 - 2\rho \left(\frac{x_1-m_1}{\sigma_1}\right)\left(\frac{x_2-m_2}{\sigma_2}\right) + \left(\frac{x_2-m_2}{\sigma_2}\right)^2 + \left(\frac{\rho(x_1-m_1)}{\sigma_1}\right)^2 - \left(\frac{\rho(x_1-m_1)}{\sigma_1}\right)^2 \right) =$$

$$\frac{1}{(1-\rho^2)} \left(\left(\frac{x_1-m_1}{\sigma_1}\right)^2 - 2\rho \left(\frac{x_1-m_1}{\sigma_1}\right)\left(\frac{x_2-m_2}{\sigma_2}\right) + \left(\frac{x_2-m_2}{\sigma_2}\right)^2 + \left(\frac{\rho(x_1-m_1)}{\sigma_1}\right)^2 - \left(\frac{\rho(x_1-m_1)}{\sigma_1}\right)^2 \right) =$$

$$\frac{1}{(1-\rho^2)} \left(\left(\frac{x_1-m_1}{\sigma_1}\right)^2 + \left(\frac{x_2-m_2}{\sigma_2} - \frac{\rho(x_1-m_1)}{\sigma_1}\right)^2 - \left(\frac{\rho(x_1-m_1)}{\sigma_1}\right)^2 \right) = \frac{1}{(1-\rho^2)} \left(\frac{x_2-m_2}{\sigma_2} - \frac{\rho(x_1-m_1)}{\sigma_1}\right)^2 + \left(\frac{x_1-m_1}{\sigma_1}\right)^2$$

Now, we can separate the exponential that only depends on x_1, and it comes out of the integral:

$$p(x_1) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left(\frac{-1}{2}\left(\frac{x_1-m_1}{\sigma_1}\right)^2\right) \int_{-\infty}^{+\infty} \exp\left(\frac{-1}{2(1-\rho^2)}\left(\frac{x_2-m_2}{\sigma_2} - \frac{\rho(x_1-m_1)}{\sigma_1}\right)^2\right) dx_2 =$$

$$\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left(\frac{-1}{2}\left(\frac{x_1-m_1}{\sigma_1}\right)^2\right) \sqrt{\frac{\pi}{\frac{1}{2(1-\rho^2)\sigma_2^2}}} = \frac{1}{\sqrt{2\pi}\sigma_1} \exp\left(\frac{-1}{2}\left(\frac{x_1-m_1}{\sigma_1}\right)^2\right) = N_{x_1}(m_1, \sigma_1^2)$$

Part B)

$$p(x_1|x_2) = \frac{p(x_1, x_2)}{p(x_2)} = \frac{\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp(\frac{-1}{2} d^2(x_1, x_2))}{\frac{1}{\sqrt{2\pi}\sigma_2} \exp\left(\frac{-1}{2}\left(\frac{x_2-m_2}{\sigma_2}\right)^2\right)} = \frac{1}{\sqrt{2\pi\sigma_1^2(1-\rho^2)}} \exp\left(\frac{-1}{2}\left(d^2(x_1, x_2) - \left(\frac{x_2-m_2}{\sigma_2}\right)^2\right)\right)$$

$$d^2(x_1, x_2) - \left(\frac{x_2-m_2}{\sigma_2}\right)^2 = \frac{1}{(1-\rho^2)} \left(\left(\frac{x_1-m_1}{\sigma_1}\right)^2 - 2\rho \left(\frac{x_1-m_1}{\sigma_1}\right)\left(\frac{x_2-m_2}{\sigma_2}\right) + \left(\frac{x_2-m_2}{\sigma_2}\right)^2 \right) - \left(\frac{x_2-m_2}{\sigma_2}\right)^2 =$$

$$\frac{1}{(1-\rho^2)} \left(\left(\frac{x_1-m_1}{\sigma_1}\right)^2 - 2\rho \left(\frac{x_1-m_1}{\sigma_1}\right)\left(\frac{x_2-m_2}{\sigma_2}\right) + \rho^2 \left(\frac{x_2-m_2}{\sigma_2}\right)^2 \right) = \frac{1}{(1-\rho^2)} \left(\frac{x_1-m_1}{\sigma_1} - \rho\frac{x_2-m_2}{\sigma_2}\right)^2$$

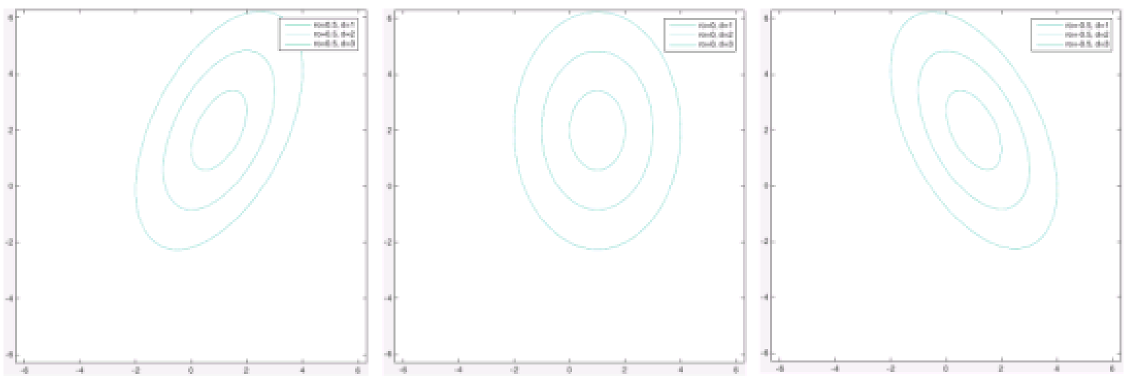
$$p(x_1|x_2) = \frac{1}{\sqrt{2\pi\sigma_1^2(1-\rho^2)}} \exp\left(\frac{-1}{2}\frac{1}{(1-\rho^2)}\left(\frac{x_1-m_1}{\sigma_1} - \rho\frac{x_2-m_2}{\sigma_2}\right)^2\right) = N_{x_1}(m_1 + \frac{\rho\sigma_1}{\sigma_2}(x_2 - m_2), \sigma_1^2(1-\rho^2))$$

Q8-

$$(1-\rho^2)d^2(x_1,x_2) = \left(\left(\frac{x_1-1}{1}\right)^2 - 2\rho\left(\frac{x_1-1}{1}\right)\left(\frac{x_2-2}{\sqrt{2}}\right) + \left(\frac{x_2-2}{\sqrt{2}}\right)^2\right)$$

Contour plots considering different values for $d^2(x_1,x_2)$ and ρ:

Please notice that for $\rho = \pm 1$ we have maximum correlation, the contours are very compact and form a line. A line with positive slope for $\rho = 1$ and another straight line for $\rho = -1$ but with negative slope. For other values of ρ the contours are plotted below:



Q۹-

$$\hat{m} = \hat{m} + \frac{d_i}{i} = \hat{m} + \frac{x_i - \hat{m}}{i} = \frac{x_1 + x_2 + \cdots + x_{i-1}}{i-1} + \frac{x_i - \frac{x_1 + x_2 + \cdots + x_{i-1}}{i-1}}{i} = \frac{i(x_1 + x_2 + \cdots + x_{i-1}) + (i-1)x_i - (x_1 + x_2 + \cdots + x_{i-1})}{i(i-1)} =$$

$$\frac{(i-1)(x_1 + x_2 + \cdots + x_i)}{i(i-1)} = \frac{(x_1 + x_2 + \cdots + x_i)}{i}$$

Considering $m_i = m_{i-1} + \frac{d_i}{i}$:

$$\Sigma = \Sigma_{k=1}^{i}(x_k - m_i)(x_k - m_i)^T = \Sigma_{k=1}^{i}\left(x_k - m_{i-1} - \frac{d_i}{i}\right)\left(x_k - m_{i-1} - \frac{d_i}{i}\right)^T =$$

$$\Sigma_{k=1}^{i}(x_k - m_{i-1})(x_k - m_{i-1})^T - \Sigma_{k=1}^{i}\left(\frac{d_i}{i}\right)(x_k - m_{i-1})^T - \Sigma_{k=1}^{i}(x_k - m_{i-1})\left(\frac{d_i}{i}\right)^T + \Sigma_{k=1}^{i}\left(\frac{d_i}{i}\right)\left(\frac{d_i}{i}\right)^T =$$

$$\Sigma_{i-1} + (x_i - m_{i-1})(x_i - m_{i-1})^T - \left(\frac{d_i}{i}\right)(x_i - m_{i-1})^T - (x_i - m_{i-1})\left(\frac{d_i}{i}\right)^T + \frac{d_i {d_i}^T}{i}$$

Remember $d_i = x_i - m_{i-1}$:

$$\Sigma = \Sigma_{i-1} + d_i {d_i}^T - \left(\frac{d_i}{i}\right){d_i}^T - d_i\left(\frac{d_i}{i}\right)^T + \frac{d_i {d_i}^T}{i} = \Sigma_{i-1} + (1 - \tfrac{1}{i})d_i {d_i}^T$$

Q۱۰-

$$S_B = E\{(x_1 - x_2)(x_1 - x_2)^T\} = E\{x_1 {x_1}^T\} - E\{x_1 {x_2}^T\} - E\{x_2 {x_1}^T\} + E\{x_2 {x_2}^T\}$$

Two points: First, x_1 and x_2 are considered iid. Second, $\Sigma = S - mm^T$.So:

$$S_B = \Sigma_1 + m_1 {m_1}^T - E\{x_1\}E\{{x_2}^T\} - E\{x_2\}E\{{x_1}^T\} + \Sigma_2 + m_2 {m_2}^T = \Sigma_1 + m_1 {m_1}^T - m_1 {m_2}^T - m_2 {m_1}^T + \Sigma_2 + m_2 {m_2}^T$$

In general, rank of x_1 is 1 and rank of S_B is n.

Q١١-

Euclidean distance: $D_E = \sqrt{(x-\mu)^T(x-\mu)} = \sqrt{(x_1-\mu_1)^2 + (x_2-\mu_2)^2}$

Mahalanobis distance: $D_M = \sqrt{(x-\mu)^T\Sigma^{-1}(x-\mu)} = \sqrt{\frac{1}{(1-\rho^2)}\left(\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 - 2\rho\left(\frac{x_1-\mu_1}{\sigma_1}\right)\left(\frac{x_2-\mu_2}{\sigma_2}\right) + \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2\right)}$

For a diagonal covariance matrix we have $= 0$: $D_M = \sqrt{\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 + \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2}$

It will be equal to D_E iff $\sigma_1 = \sigma_2 = 1$. Totally, $D_M = D_E$ for $\Sigma = I$.

Euclidean distance does not consider the correlation between variables so if we have correlation it is better to use Mahalanobis distance. On the other hand, the computational cost for calculating Mahalanobis distance is high because of inversion of covariance matrix; therefore, having no correlation in our data lets us benefit form Euclidean distance.

Mahalanobis distance for two samples from the same Gaussian distribution: $D_M = \sqrt{(x-y)^T S^{-1}(x-y)}$ This uses S since the correlation between samples is matter.

Q١٣-

We have n independent samples each from class ω_i with probability $p(\omega_i)$. Another random variable is defined z_{ik} that is one if sample x is generated from class ω_i in iteration k. The variable is zero otherwise. So:

$p(z_{ik} = 1) = p(\omega_i)$ and $p(z_{ik} = 0) = 1 - p(\omega_i)$ Then:

$$p(z_{i1}, z_{i2}, \dots, z_{in}|\omega_i) = p(z_{i1}|\omega_i)p(z_{i2}|\omega_i) \dots p(z_{in}|\omega_i) = \prod_{k=1}^{n} p(\omega_i)^{z_{ik}}(1-p(\omega_i))^{1-z_{ik}}$$

These are samples that z_{ik} is one for them which means that they are generated from class ω_i. The other term for these samples will be equal to 1 as $1 - z_{ik} = 0$. It is like Binomial distribution within each class ω_i. Isn't it?

Q۱۵-

$$\Sigma = \frac{1}{N}\sum_{i=1}^{M}\sum_{j=1}^{N_i}(x_{ij}-m)(x_{ij}-m)^T = \frac{1}{N}\sum_{i=1}^{M}\sum_{j=1}^{N_i}(x_{ij}-m_i+m_i-m)(x_{ij}-m_i+m_i-m)^T$$

$$= \frac{1}{N}\sum_{i=1}^{M}\sum_{j=1}^{N_i}((x_{ij}-m_i)+(m_i-m))((x_{ij}-m_i)+(m_i-m))^T$$

$$= \frac{1}{N}\sum_{i=1}^{M}\sum_{j=1}^{N_i}(x_{ij}-m_i)(x_{ij}-m_i)^T + \frac{1}{N}\sum_{i=1}^{M}\sum_{j=1}^{N_i}(x_{ij}-m_i)(m_i-m)^T + \frac{1}{N}\sum_{i=1}^{M}\sum_{j=1}^{N_i}(m_i-m)(x_{ij}-m_i)^T$$

$$+ \frac{1}{N}\sum_{i=1}^{M}\sum_{j=1}^{N_i}(m_i-m)(m_i-m)^T$$

$$\frac{1}{N}\sum_{i=1}^{M}\left(\sum_{j=1}^{N_i}(x_{ij}-m_i)\right)(m_i-m)^T = \frac{1}{N}\sum_{i=1}^{M}(N_i m_i - N_i m_i)\,(m_i-m)^T = 0$$

$$\frac{1}{N}\sum_{i=1}^{M}(m_i-m)\left(\sum_{j=1}^{N_i}(x_{ij}-m_i)^T\right) = \frac{1}{N}\sum_{i=1}^{M}(m_i-m)(N_i m_i - N_i m_i)^T = 0$$

So,

$$\Sigma = \frac{1}{N}\sum_{i=1}^{M}\sum_{j=1}^{N_i}(x_{ij}-m)(x_{ij}-m)^T = \frac{1}{N}\sum_{i=1}^{M}\sum_{j=1}^{N_i}(x_{ij}-m_i)(x_{ij}-m_i)^T + \frac{1}{N}\sum_{i=1}^{M}\sum_{j=1}^{N_i}(m_i-m)(m_i-m)^T = \Sigma_W + \Sigma_B$$

In our problem:

$$\Sigma_W = \frac{1}{N}\sum_{i=1}^{M}\sum_{j=1}^{N_i}(x_{ij}-m_i)(x_{ij}-m_i)^T = \frac{N_1\Sigma_1 + N_2\Sigma_2}{N} = \frac{\begin{bmatrix}4 & -3\\ -3 & 8\end{bmatrix} + \begin{bmatrix}4 & 2\\ 2 & 4\end{bmatrix}}{2} = \begin{bmatrix}4 & -0.5\\ -0.5 & 6\end{bmatrix}$$

$$\Sigma_B = \frac{1}{N}\sum_{i=1}^{M}\sum_{j=1}^{N_i}(m_i-m)(m_i-m)^T = \frac{N_1(m_1-m)(m_1-m)^T + N_2(m_2-m)(m_2-m)^T}{N} = \begin{bmatrix}1 & 0.5\\ 0.5 & 0.25\end{bmatrix}$$

$$\Sigma_T = \Sigma_W + \Sigma_B = \begin{bmatrix}5 & 0\\ 0 & 6.25\end{bmatrix}$$

You can verify it with a simple simulation.

Alternative solution;

Note: $S_{A\cup B} = S_A + S_B + \frac{n_1 n_2}{n_1 + n_2}(\mu_A - \mu_B)(\mu_A - \mu_B)^T$

Sol:

$$S_A = \begin{bmatrix} 200 & -150 \\ -150 & 400 \end{bmatrix}, S_B = \begin{bmatrix} 200 & 100 \\ 100 & 200 \end{bmatrix}$$

$$S_{A\cup B} = S_A + S_B + \frac{n_1 n_2}{n_1 + n_2}(\mu_A - \mu_B)(\mu_A - \mu_B)^T$$

$$\Rightarrow S_{A\cup B} = \begin{bmatrix} 200 & -150 \\ -150 & 400 \end{bmatrix} + \begin{bmatrix} 200 & 100 \\ 100 & 200 \end{bmatrix} + \frac{50 * 50}{50 + 50}\begin{bmatrix} 2 \\ 1 \end{bmatrix}\begin{bmatrix} 2 & 1 \end{bmatrix} = \begin{bmatrix} 500 & 0 \\ 0 & 625 \end{bmatrix}$$

$$\Rightarrow \Sigma_{A\cup B} = \begin{bmatrix} 5 & 0 \\ 0 & 6.25 \end{bmatrix} \Rightarrow \begin{cases} \lambda_1 = 6.25,\ v_1 = (0 \quad 1)^T \\ \lambda_2 = 5,\ v_2 = (1 \quad 0)^T \end{cases}$$

Q16-

$$\text{Mean} = \begin{bmatrix} 16.5 \\ 17 \end{bmatrix}, \text{Covariance} = \begin{bmatrix} 9.6667 & 5.3333 \\ 5.3333 & 4.6667 \end{bmatrix}$$

Correlation$=\frac{cov(X,Y)}{\sigma(X)\sigma(Y)} = 0.7941$ → strong correlation.

Covariance scales arbitrarily, so we can never be sure how strong a covariance is by looking at the magnitude. Correlation, however, is unitless, which is exactly why we can rely on the magnitude to inform us as to the strength of the relationship. Covariance can range from - infinity to + infinity. But Correlation ranges from -1 to +1. (Normalized covariance = Correlation)

Q١٧-

Statistical independence means: $p(x,y) = p(x)p(y)$

$$E(x+y) = \int\int p(x,y)(x+y)dx\,dy = \int\int p(x)p(y)(x+y)dx\,dy = \int\int p(x)p(y)xdx\,dy + \int\int p(x)p(y)ydx\,dy$$

$$= \int p(y)\int p(x)xdx\,dy + \int p(y)y\int p(x)dx\,dy = \int p(y)E(x)dy + \int p(y)ydy = E(x) + E(y)$$

We know: $Var(a) = E[(a - E[a])^2]$, accordingly:

$$Var(x+y) = E\left[\left((x+y) - E[x+y]\right)^2\right] = E[(x + y - E[x] - E[y])^2] = E[((x - E[x]) + (y - E[y]))^2] =$$

$$E[(x-E[x])^2 + 2(x-E[x])(y-E[y]) + (y-E[y])^2] = E[(x-E[x])^2] + E[2(x-E[x])(y-E[y])] +$$

$$E[(y-E[y])^2] = Var(x) + Var(y)$$

Q18-

D

$$\varphi_1 = \begin{bmatrix}1\\2\end{bmatrix} \rightarrow \begin{bmatrix}\frac{1}{\sqrt{5}}\\ \frac{2}{\sqrt{5}}\end{bmatrix},\quad \lambda_1 = 2 \qquad\qquad \varphi_2 = \begin{bmatrix}-2\\1\end{bmatrix} \rightarrow \begin{bmatrix}\frac{-2}{\sqrt{5}}\\ \frac{1}{\sqrt{5}}\end{bmatrix},\quad \lambda_1 = 1$$

Assuming $\varphi^t\varphi = 1$: $\Sigma\varphi = \lambda\varphi \quad \rightarrow \quad \begin{bmatrix}a & b\\ c & d\end{bmatrix}\begin{bmatrix}\frac{1}{\sqrt{5}}\\ \frac{2}{\sqrt{5}}\end{bmatrix} = 2\begin{bmatrix}\frac{1}{\sqrt{5}}\\ \frac{2}{\sqrt{5}}\end{bmatrix}$ and $\begin{bmatrix}a & b\\ c & d\end{bmatrix}\begin{bmatrix}\frac{-2}{\sqrt{5}}\\ \frac{1}{\sqrt{5}}\end{bmatrix} = \begin{bmatrix}\frac{-2}{\sqrt{5}}\\ \frac{1}{\sqrt{5}}\end{bmatrix}$

$$\frac{a+2b}{\sqrt{5}} = \frac{2}{\sqrt{5}} \quad \rightarrow \quad a + 2b = 2$$

$$\frac{c+2d}{\sqrt{5}} = \frac{4}{\sqrt{5}} \quad \rightarrow \quad c + 2d = 4$$

$$\frac{-2a+b}{\sqrt{5}} = \frac{-2}{\sqrt{5}} \quad \rightarrow \quad -2a + b = -2$$

$$\frac{-2c+d}{\sqrt{5}} = \frac{1}{\sqrt{5}} \quad \rightarrow \quad -2c + d = 1$$

$$\Sigma = \begin{bmatrix}6/5 & 2/5\\ 2/5 & 9/5\end{bmatrix}$$

Q19-

A and D

Q20-

C

$$a = \rho\sqrt{5}\sqrt{4} = \rho\sqrt{20} \;\rightarrow\; -1 \leq \rho = \frac{a}{\sqrt{20}} \leq 1 \rightarrow \quad -\sqrt{20} \leq a \leq \sqrt{20}$$

Q21-

C and D

There can be infinitely many unit-length eigenvectors if the multiplicity of any eigenvector is greater than 1 (so the eiganspace is a plane, and you can pick any vector on the unit circle on that plane).

The 0 vector is not an eigenvector by definition.