Problem Set 7

7.5 PCA and Classification.

a).
$$p(x) = \sum_{j=1}^{2} p(x|y=j)\, p(y=j)$$
$$= \frac{1}{2}\sum_{j=1}^{2} \mathcal{N}(x|u_j, \Sigma_j)$$

$$u_x = E(x) = \int x\, p(x)\, dx = \frac{1}{2}\left[\int x\, \underbrace{\mathcal{N}(x|u_1,\Sigma_1)}\, dx + \int x\, \underbrace{\mathcal{N}(x|u_2,\Sigma_2)}\, dx\right]$$
$$= \frac{1}{2}(u_1 + u_2)$$

To the $\Sigma_x = E[(x-u_x)(x-u_x)^T]$, $u_x = \frac{1}{2}(u_1+u_2)$

$$\Sigma_x = E(xx^T) - E(x)^2 = \int xx^T p(x)\,dx - u_x u_x^T$$
$$= \frac{1}{2}\left(\int xx^T \mathcal{N}(x|u_1,\Sigma_1)\,dx + \int xx^T \mathcal{N}(x|u_2,\Sigma_2)\,dx\right) - u_x u_x^T$$
$$= \frac{1}{2}\Big(E(x^2|y=1) + E(x^2|y=2) - E(x|y=1)^2 - E(x|y=2)^2 + E(x|y=1)^2$$
$$+ E(x|y=2)^2\Big) - u_x u_x^T$$

Since $E(x^2|y=j) - E(x|y=j)^2 = \Sigma_j$

Thus,
$$= \frac{1}{2}(\Sigma_1+\Sigma_2) + \frac{1}{2}(u_1u_1^T + u_2u_2^T) - \frac{1}{4}(u_1+u_2)(u_1+u_2)^T$$
$$= \frac{1}{2}(\Sigma_1+\Sigma_2) + \frac{1}{4}(u_1-u_2)(u_1-u_2)^T = \Sigma_x$$

b) Since $u_1 = -u_2 = u = \begin{bmatrix} a \\ 0 \end{bmatrix}$, $u_x = 0$

$$\Sigma_x = \frac{1}{2}(\Sigma_1+\Sigma_2) + \frac{1}{2} u_1 u_1^T$$
$$= \begin{bmatrix} 1+\alpha^2 & 0 \\ 0 & \sigma^2 \end{bmatrix}$$

Since $\Sigma_x = \begin{bmatrix} 1+\alpha^2 & 0 \\ 0 & \sigma^2 \end{bmatrix}$, the eigenvalues are

$1+\alpha^2$, σ^2, when $1+\alpha^2 > \sigma^2$.

eigenvector is $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$, when $1+\alpha^2 < \sigma^2$,

eigenvector is $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$

when $\alpha > \sqrt{\sigma^2 - 1}$, largest eigenvector $\phi^T = [1, 0]$

$z = [1, 0]x$, $p(z|y=1) = p([1,0]x|y=1)$

$= N(z|\alpha, 1)$

$p(z|y=2) = N(z|-\alpha, 1)$,

This is a greasest seperation. Two classifiers $E(x|y=j)$ are opposite, that's optimal.

when $\alpha < \sqrt{\sigma^2 - 1}$, largest eigenvector $\phi^T = [0, 1]$

$z = [0, 1]x$, $p(z|y=1) = p([0,1]x|y=1)$

$= N(z|0, \sigma^2)$, $p(z|y=2) = N(z|0, \sigma^2)$

This is the bad choice, Since, their direction are the same.

This case PCA cannot seprerate them well.

In conclusion, PCA need to consider their result, not just select largest k-th eigenvector can reach the optimal set for classification.

57558749

王琮展