HW 16 P01

i) LDA

for sets $X_1 = \{x_1^1, x_2^1, x_3^1 \ldots x_{n}^1\}$
$X_2 = \{x_1^2, x_2^2, x_3^2 \ldots x_n^2\}$

The goal of LDA is to maximise class sep. while keeping in-class variance small.

Max: $J(w) = \dfrac{w^T S_B w}{w^T S_W w}$ where $S_B = (m_2 - m_1)(m_2 - m_1)^T$

$S_W = \sum_{i=1,2} \sum_{n=1}^{n} (x_n^i - m_i)(x_n^i - m_i)^T$

S_B is between-class cov. mat.

S_W is total within-class " " .

$m_i = \dfrac{1}{n_i} \sum_{n=1}^{n_i} x_n^i$

We need to maximise

$$\underset{w}{\operatorname{argmax}} \left(\frac{w^T S_B w}{w^T S_W w} \right)$$

ii). K-LDA

Using some function ϕ (non-linear mapping) to map the points to a new feature space F,

$$J(w) = \frac{w^T S_B^\phi w}{w^T S_W^\phi w}$$

where $S_B^\phi = (m_2^\phi - m_1^\phi)(m_2^\phi - m_1^\phi)^T$ $S_W = \sum_{i=1,2} \sum_{n=1}^{n_i} (\phi(x_i) - m_i^\phi)(\phi(x_i) - m_i^\phi)^T$

$$m_i^\phi = \frac{1}{n_i} \sum_{j=1}^{n_i} \phi(x_j^i)$$

$w \in F$ must lie in span of all training samples in F, $w = \sum_{i=1}^{n} \alpha_i \phi(x_i)$

So $w^T m_i^\phi = \dfrac{1}{n_i} \sum_{j=1}^{n} \sum_{k=1}^{n_i} \alpha_j k(x_j, x_k^i) = \alpha^T M_i$ where $(M_i)_j = \dfrac{1}{n_i} \sum_{k=1}^{n_i} k(x_j, x_k^i)$

or $w^T S_B^{\phi} w = \alpha^T M \alpha$

$$\left[M = (M_2 - M_1)(M_2 - M_1)^T \right]$$

$$w^T S_W^{\phi} w = \alpha^T N \alpha \qquad \left[N = \sum_{j=1,2} K_j (I - 1_{n_j}) K_j^T \right]$$

- $$J = \frac{\alpha^T M \alpha}{\alpha^T N \alpha}$$