AML HW2 - Written Exercises

Matheus Dafutto (mc2722)
Sungseo Park (sp2528)

Q1)

Applying the Lagrange multiplier, $L(x,y,\lambda) = f(x,y) - \lambda g(x,y)$,
we can define $L(a) = a^T B a - \lambda (a^T W a - 1)$

Take derivative with respect to a,

$$\frac{dL}{da} = (B+B^T)a - \lambda (W+W^T)a$$

Let it equal to 0, then we have $[(W+W^T)]^{-1}(B+B^T)a = \lambda a$,
which is the standard eigenvalue problem. Assuming W and B symmetric, we have $W^{-1}Ba = \lambda a$.

Q2a)

From equation 4.11 in HTF, the LDA rule classifies to class 2

$$\text{if } x^T \hat{\Sigma}^{-1}(\hat{\mu}_2 - \hat{\mu}_1) > \frac{1}{2}\hat{\mu}_2^T \hat{\Sigma}^{-1}\hat{\mu}_2 - \frac{1}{2}\hat{\mu}_1^T \hat{\Sigma}^{-1}\hat{\mu}_1 + \log\left(\frac{N_1}{N}\right) - \log\left(\frac{N_2}{N}\right)$$

Q2b)

By least squares criterion,

$$\sum_{i=1}^{N}(y_i - \beta_0 - \beta^T x)^2 = (Y - \beta_0 1 - X\beta)^T (Y - \beta_0 1 - X\beta)$$

$$2X^T X\beta - 2X^T Y + 2\beta_0 X^T 1 = 0$$

$$2N\beta_0 - 21^T(Y - X\beta) = 0$$

$$\hat{\beta}_0 = \frac{1}{N} 1^T (y - X\beta)$$

Q2d)

If the results calculated in 2b are arbitrary ~~and~~ and distinct,
the result holds.

Q3. $M = \begin{bmatrix} 1 & 0 & 3 \\ 3 & 7 & 2 \\ 2 & -2 & 8 \\ 0 & -1 & 1 \\ 5 & 8 & 7 \end{bmatrix}$

a) $MM^T = \begin{bmatrix} 10 & 9 & 26 & 3 & 26 \\ 9 & 62 & 8 & -5 & 85 \\ 26 & 8 & 72 & 10 & 50 \\ 3 & -5 & 10 & 2 & -1 \\ 26 & 85 & 50 & -1 & 138 \end{bmatrix}$ $\quad M^TM = \begin{bmatrix} 39 & 57 & 60 \\ 57 & 118 & 53 \\ 60 & 53 & 127 \end{bmatrix}$

b) $eigval(MM^T) = solv(det(MM^T - I\cdot\lambda) = 0)$ ← same for M^TM

$\lambda_{MM^T} = \underset{\lambda_1}{214.7}, \underset{\lambda_2}{0}, \underset{\lambda_3}{69.3}, \underset{\lambda_4}{0}, \underset{\lambda_5}{0}$

$\lambda_{M^TM} = 214.7, 0, 69.3$

c) $eigvec(MM^T) = x_i \mid (MM^T - \lambda_i I)\cdot x_i = 0$ for each eigval λ

$$X_{(MM^T)} = \begin{array}{c} \begin{array}{ccccc} x_1 & x_2 & x_3 & x_4 & x_5 \end{array} \\ \begin{bmatrix} -0.165 & -0.955 & 0.245 & -0.065 & -0.107 \\ -0.472 & -0.035 & 0.453 & 0.750 & -0.047 \\ -0.336 & 0.271 & 0.829 & 0.328 & -0.174 \\ 0.003 & 0.044 & 0.169 & 0.046 & 0.971 \\ -0.798 & 0.104 & -0.133 & -0.568 & 0.119 \end{bmatrix} \end{array}$$

$eigvec(M^TM) = x_i \mid (MM^T - \lambda_i I)\cdot x_i = 0$ for each eigval λ:

$$X_{(M^TM)} = \begin{array}{c} \begin{array}{ccc} x_1 & x_2 & x_3 \end{array} \\ \begin{bmatrix} 0.426 & 0.904 & -0.015 \\ 0.615 & -0.302 & -0.729 \\ 0.663 & -0.302 & 0.685 \end{bmatrix} \end{array}$$

d) $MM^T = (U\Sigma V^T)(U\Sigma V^T)^T = U\Sigma V^T \cdot V\Sigma U^T = U\Sigma^2 U^T$

$$\boxed{\begin{aligned}\Sigma(M) &= \sqrt{\Sigma^2} \\ U(M) &= \text{eigvec}(MM^T)\end{aligned}}$$

$$\boxed{\Sigma = \begin{bmatrix} \sqrt{214.7} & 0 \\ 0 & \sqrt{69.3} \end{bmatrix} = \begin{bmatrix} 8.326 & 0 \\ 0 & 14.76 \end{bmatrix}}$$

$\lambda_2, \lambda_4, \lambda_5 = 0$ so we can ignore them

$$\boxed{U = \begin{bmatrix} | & | \\ v_1 & v_3 \\ | & | \end{bmatrix} = \begin{bmatrix} -0.244 & 0.165 \\ 0.453 & 0.471 \\ -0.829 & 0.336 \\ -0.169 & 0.003 \\ -0.133 & -0.798 \end{bmatrix}}$$

$M^TM = (U\Sigma V^T)^T(U\Sigma V^T) = V\Sigma U^T U\Sigma V^T = V\Sigma^2 V^T$

$$V(M) = \text{eigvec}(M^TM)$$

$$\boxed{V = \begin{bmatrix} - & v_1 & - \\ - & v_3 & - \end{bmatrix} = \begin{bmatrix} 0.015 & 0.728 & 0.685 \\ -0.426 & 0.615 & 0.663 \end{bmatrix}}$$

e) Set $\Sigma_{new} = \begin{bmatrix} 14.65 & 0 \\ 0 & 0 \end{bmatrix}$ and compute $U\Sigma_{new}V^T$

$$\boxed{\hat{M} = U\Sigma_{new}V^T \cong \begin{bmatrix} 1 & 2 & 2 \\ 3 & 4 & 5 \\ 2 & 3 & 3 \\ 0 & 0 & 0 \\ 5 & 7 & 8 \end{bmatrix}}$$