HW 3

Problem 1.

Answers:

1. If N1 = N2, the LDA classification rule classifies a new case with x predictors as belonging to class 1 if the discriminant value:

$$\delta 1(\mathrm{x}) = \mathrm{x}^{\mathrm{T}}\Sigma^{-1}\mu 1 - \tfrac{1}{2}\mu 1^{\mathrm{T}}\Sigma^{-1}\mu 1 + \log(\pi 1).$$

is greater than or equal to the discriminant score for class 2:

$$\delta 2(\mathrm{x}) = \mathrm{x}^{\mathrm{T}}\Sigma^{-1}\mu 2 - \tfrac{1}{2}\mu 2^{\mathrm{T}}\Sigma^{-1}\mu 2 + \log(\pi 2)$$

2. If N1! = N2, the LDA classification rule classifies a new case with x predictors as belonging to class 1 if the discriminant value:

$$\delta 1(\mathrm{x}) = \mathrm{x}^{\mathrm{T}}\Sigma^{-1}\mu 1 - \tfrac{1}{2}\mu 1^{\mathrm{T}}\Sigma^{-1}\mu 1 + \log(\pi 1) + \log\left(\mathrm{N}\tfrac{1}{\mathrm{N}}\right)$$

is greater than or equal to the class 2 discriminant score,

$$\delta 2(\mathrm{x}) = \mathrm{x}^{\mathrm{T}}\Sigma^{-1}\mu 2 - \tfrac{1}{2}\mu 2^{\mathrm{T}}\Sigma^{-1}\mu 2 + \log(\pi 2) + \log\left(\mathrm{N}\tfrac{2}{\mathrm{N}}\right)$$

3. Linear regression can be used for binary classification by assigning a limit to the predicted continuous response variable.

We can fit a linear regression model to the data with y as the response and x as the predictor, and then classify a new case with x predictors as Class 1 if the predicted response is greater than or equal to the threshold and as Class 1 otherwise.

If we choose the cutoff as the midpoint between the predicted mean response for class 1 and class 2, then the linear regression classification rule is the same as the LDA classification rule with equal precedents and a common covariance matrix.

4. If we use a different coding scheme for y, e.g. B. $y \in \{0,1\}$, we can still use linear regression to fit the model to the data, but the classification threshold is different.

We can choose the threshold as the midpoint between the predicted mean response for $y = 0$ and $y = 1$ If the prior probabilities of the two classes are equal and the covariance matrix is common, this linear regression classification used will be identical to the LDA classification rule.

Problem 2:

Answer:
To solve the generalized eigenvalue problem, use the concept of Lagrangian multipliers.

$Ax = Bx$ Such that $x! = 0$
The constraint is $a^TWa = 1$

Using Lagrangian Multiplers:

$l(\lambda) = a^Tba - \lambda\ (a^TWa - 1)$

By taking derivate,

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

By equating $\frac{dl}{da}$ to 0,

$(B + B^T)a - \lambda(W + W^T)a = 0$

$(B + B^T)a = \lambda(W + W^T)a$

$\lambda a = \frac{(B + B\text{^}T)a}{(W + W\text{^}T)}$

The above equation is equation of standard eigen value

Assuming B and W are symmetric

$\lambda a = W^{-1}Ba$

Problem 3:
Attaching the Python file.

Problem 4:
Attaching the Python file.

Problem 5:
Attaching the Python file.