Supplementary Material

Mathematical working of QDA

- **Problem and data**

In this binary classification problem (K = 2), we want to classify an observation containing two predictor variables **minorAL** (X_1) and **ecc** (X_2) into two distinct classes Kecimen and Besni.

By assumption, let the distribution of data $f_i(x) \sim N(\mu_i, \Sigma_i)$,
where $E(X) = \mu_i$ is a mean vector of i^{th} class with two components and $Cov(X) = \Sigma_i$ is a 2x2 covariance matrix of i^{th} class.
Unlike LDA, QDA assumes that each class has its own covariance matrix.

- **Constructing decision boundary**

The decision boundary for QDA can be obtained by following the same steps as LDA without the assumption of equal covariance matrices.

Calculating the odds ratio for the data we get,

$$\frac{P(Y = 1 | X = x)}{P(Y = 2 | X = x)} = \frac{\pi_1 |\Sigma_2|^{1/2} e^{\left(-\frac{1}{2}(x-\mu_1)^T \Sigma_1^{-1}(x-\mu_1)\right)}}{\pi_2 |\Sigma_1|^{1/2} e^{\left(-\frac{1}{2}(x-\mu_2)^T \Sigma_2^{-1}(x-\mu_2)\right)}}$$

$$or, \quad \log\left(\frac{P(Y = 1 | X = x)}{P(Y = 2 | X = x)}\right) = -\frac{1}{2}(x-\mu_1)^T \Sigma_1^{-1}(x-\mu_1) + \frac{1}{2}(x-\mu_2)^T \Sigma_2^{-1}(x-\mu_2) + \frac{1}{2}\log\left(\frac{\Sigma_2}{\Sigma_1}\right) + \log\left(\frac{\pi_1}{\pi_2}\right)$$

The above equation can be expressed in the form,

$$\log\left(\frac{P(Y = 1 | X = x)}{P(Y = 2 | X = x)}\right) = \beta_0 + \beta^T x + x^T \Omega x,$$

$$\Omega = -\frac{1}{2}\left(\Sigma_1^{-1} - \Sigma_2^{-1}\right)$$

$$\beta = \left(\Sigma_1^{-1}\mu_1 - \Sigma_2^{-1}\mu_2\right)$$

$$\beta_0 = \log\left(\frac{\pi_1}{\pi_2}\right) - \frac{1}{2}\left[\log\left(\frac{\Sigma_1}{\Sigma_2}\right) + \mu_1^T \Sigma_1^{-1}\mu_1 - \mu_2^T \Sigma_2^{-1}\mu_2\right]$$

The quadratic log-odds function implies that the decision boundary of QDA is quadratic.

Therefore, the decision rule of QDA can be formulated as,

$$\boldsymbol{G(x) = argmax_k \delta_k(x)},$$

$$\boldsymbol{\delta_k(x) = -\frac{1}{2}(x-\mu_k)^T \Sigma_k^{-1}(x-\mu_k) - \frac{1}{2}log(\Sigma_k) + \log(\pi_k)}, \qquad \boldsymbol{k = 1, 2}$$

Note: The decision boundary of the two classes is the set of points where $\{\boldsymbol{x}: \boldsymbol{\delta_1(x)} = \boldsymbol{\delta_2(x)}\}$

Estimating the parameters of QDA classifier

The parameter estimation of QDA are similar to that of LDA are calculated and estimated using the following rule,

1. **Proportion of observations** belonging to kth class:

$$\hat{\pi}_k = \frac{N_k}{N}, \qquad k = 1,2$$

2. **Mean of the observations** belonging to that class:

$$\hat{\mu}_k = \frac{\sum_{i=1}^{N_k} x_i}{N_k}, \qquad k = 1,2$$

3. **Covariance matrix for each class**:

$$\widehat{\Sigma_k} = \frac{\sum_{i=1}^{N_k}(x_i - \hat{\mu}_k)(x_i - \hat{\mu}_k)^T}{(N_k - 1)}, \; k = 1,2$$

References

- The Elements of Statistical Learning
- An Introduction to Statistical Learning with Applications in R
- Modern Multivariate Statistical Techniques