Stat5504, Fall 2015: Homework 2

(Due on October 8th, 2015)

1. Consider a two-class classification problem where $Y = 1|\boldsymbol{x} \sim N(\boldsymbol{\mu}_1, \boldsymbol{\Sigma})$ and $Y = 0|\boldsymbol{x} \sim N(\boldsymbol{\mu}_0, \boldsymbol{\Sigma})$. Suppose that the prior probability of each is equal, i.e., $P(Y = 1) = P(Y = 0)$. Show that

$$\log \frac{P(Y=1|\boldsymbol{x})}{P(Y=0|\boldsymbol{x})} = \boldsymbol{x}'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0) - \frac{1}{2}(\boldsymbol{\mu}_1 + \boldsymbol{\mu}_0)'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0)$$

2. It is know that the parameter estimation of ridge regression is $\hat{\boldsymbol{\beta}} = (\boldsymbol{X}'\boldsymbol{X} + \lambda \boldsymbol{I})^{-1}\boldsymbol{X}'\boldsymbol{y}$. where $\boldsymbol{X}$ is an $n \times p$ data matrix and y is an $n \times 1$ response vector. Show that

$$(\boldsymbol{X}'\boldsymbol{X} + \lambda \boldsymbol{I}_p)^{-1}\boldsymbol{X}'\boldsymbol{y} = \boldsymbol{X}'(\boldsymbol{X}\boldsymbol{X}' + \lambda \boldsymbol{I}_n)^{-1}\boldsymbol{y}.$$

Thus we can also express $\hat{\boldsymbol{\beta}}$ as $\hat{\boldsymbol{\beta}} = \boldsymbol{X}'(\boldsymbol{X}\boldsymbol{X}' + \lambda \boldsymbol{I})^{-1}\boldsymbol{y}$.

3. Let $\boldsymbol{X}_1 \sim N(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_{11})$, and $\boldsymbol{X}_2|\boldsymbol{X}_1 \sim N(\boldsymbol{A}\boldsymbol{x}_1 + \boldsymbol{b}, \boldsymbol{\Omega})$, where $\boldsymbol{\Omega}$ is not dependent on $\boldsymbol{x}_1$. Derive the probability density function (pdf) of $\boldsymbol{X}$ to show that

$$\boldsymbol{X} = \begin{pmatrix} \boldsymbol{X}_1 \\ \boldsymbol{X}_2 \end{pmatrix} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma}),$$

where

$$\boldsymbol{u} = \begin{pmatrix} \boldsymbol{u}_1 \\ \boldsymbol{A}\boldsymbol{\mu}_1 + \boldsymbol{b} \end{pmatrix} \text{ and } \boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{11}\boldsymbol{A}^T \\ \boldsymbol{A}^T\boldsymbol{\Sigma}_{11} & \boldsymbol{\Omega} + \boldsymbol{A}\boldsymbol{\Sigma}_{11}\boldsymbol{A}^T \end{pmatrix}$$

4. Let $\boldsymbol{X} = (X_1, X_2, X_3)^T \in \mathcal{R}^3$ be a multivariate normal distribution with covariance matrix

$$\boldsymbol{\Sigma} = \begin{pmatrix} 1 & \rho & \rho^2 \\ \rho & 1 & \rho \\ \rho^2 & \rho & 1 \end{pmatrix}$$

Show the conditional distribution of $(X_1, X_2)|X_3 = x_3$ and calculate $\boldsymbol{\Sigma}^{-1}$. Validate that the independency of X_1 and X_2 conditional on X_3 is equivalent to $\rho = 0$.

5. Let $\boldsymbol{X} = (\boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{X}_3)^T$ be a multivariate normal distribution with mean and covariance matrix as follows

$$\boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \\ \boldsymbol{\mu}_3 \end{pmatrix} \quad \boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} & \boldsymbol{\Sigma}_{13} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} & \boldsymbol{\Sigma}_{23} \\ \boldsymbol{\Sigma}_{31} & \boldsymbol{\Sigma}_{32} & \boldsymbol{\Sigma}_{33} \end{pmatrix}.$$

(a) Show the conditional distribution of $\boldsymbol{X}_1 | (\boldsymbol{X}_2 = \boldsymbol{x}_2, \boldsymbol{X}_3 = \boldsymbol{x}_3)$.

(b) Conduct a simulation to show the derivation in (a) is correct.