STAT5030 Midterm Linear Models (13-14)

1. Consider K $(k = 1, ..., K)$ regression models

$$Y_{ki} = \alpha_k + \beta_k x_{ki} + \epsilon_{ki}, \qquad (i = 1, 2, ..., n_k)$$

where the ϵ_{ki} are independently and identically distributed as $N(0, \sigma^2)$.

   (a) Find the least squares estimates of α_k and β_k.

   (b) To conduct the test of equal y-intercept (all K regression lines meet at the same point when $x = 0$), what are the null and alternative hypotheses? What is the reduced model under the null? Derive the SSE of the full model and the SSE of the reduced model, and the details of the testing procedure.

2. Let

$$Y_i = \theta_i + \epsilon_i$$

where $i = 1, 2, 3, 4$ and ϵ_i are independent $N(0, \sigma^2)$. Let $\theta_1 + \theta_2 + \theta_3 + \theta_4 = 0$.

   (a) Derive the least squares estimates of the parameters.

   (b) Find the SSE when $Y_1 = 1, Y_2 = 2, Y_3 = 3$, and $Y_4 = 4$.

3. Suppose that the regression curve

$$E(Y) = \beta_0 + \beta_1 x + \beta_2 x^2$$

have a local maximum at $x = x_m$ where x_m is near the origin. If Y is observed at n points $x_i, (i = 1, 2, ..., n)$ in $[-a, a]$, $\bar{x} = 0$, and Y_i are independent normal random variables with variance equal to σ^2, Using the random variable $U = \hat{\beta}_1 + 2x_m\hat{\beta}_2$ where $\hat{\beta}_1$ and $\hat{\beta}_2$ are LSE of β_1 and β_2 respectively, outline a method for finding a confidence interval for x_m.

4. The hat matrix is $\mathbf{H} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' = \{h_{ij}\}$ (Let $\mathbf{X}$ be a matrix with full column rank and with $\mathbf{1}$ as its first column). From class notes, we have $h_{ii} = 1/n + (\mathbf{x}_{1i} - \bar{\mathbf{x}}_1)'(\mathbf{X}_c'\mathbf{X}_c)^{-1}(\mathbf{x}_{1i} - \bar{\mathbf{x}}_1)$, where $\mathbf{x}_{1i}' = (x_{i1}, x_{i2}, ..., x_{ik})$, $\bar{\mathbf{x}}_1' = (\bar{x}_1, \bar{x}_2, ..., \bar{x}_k)$, and $(\mathbf{x}_{1i} - \bar{\mathbf{x}}_1)'$ is the ith row of the centered matrix $\mathbf{X}_c$. Prove that we can also express h_{ii} as the following:

$$h_{ii} = 1/n + (\mathbf{x}_{1i} - \bar{\mathbf{x}}_1)'(\mathbf{x}_{1i} - \bar{\mathbf{x}}_1)\sum_{r=1}^{k}\frac{1}{\lambda_r}cos^2\theta_{ir},$$

where θ_{ir} is the angle between $\mathbf{x}_{1i} - \bar{\mathbf{x}}_1$ and $\mathbf{a}_r$, the rth normalized eigenvector (λ_r is the corresponding eigenvalue) of $\mathbf{X}_c'\mathbf{X}_c$.