Math 6010, Fall 2004: Homework

Homework 4

#3, page 41: We have three samples, Y_1, Y_2, and Y_3. We have three noise terms, $\varepsilon_1, \varepsilon_2$, and ε_3, and we have two parameters, $\beta_1 = \theta$ and $\beta_2 = \phi$. The linear model, then, is

$$\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad \text{where} \quad \boldsymbol{X} = \begin{pmatrix} 1 & 0 \\ 2 & -1 \\ 1 & 2 \end{pmatrix}.$$

Note that

$$\boldsymbol{X}'\boldsymbol{X} = \begin{pmatrix} 6 & 5 \\ 0 & 5 \end{pmatrix}, \quad \text{so that} \quad \left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1} = \begin{pmatrix} \frac{1}{6} & 0 \\ 0 & \frac{1}{5} \end{pmatrix}.$$

In particular,

$$\widehat{\boldsymbol{\beta}} = \begin{pmatrix} \widehat{\theta} \\ \widehat{\phi} \end{pmatrix} = \left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1} \boldsymbol{X}'\boldsymbol{Y} = \begin{pmatrix} \frac{1}{6} & \frac{1}{3} & \frac{1}{6} \\ 0 & -\frac{1}{5} & \frac{2}{5} \end{pmatrix} \begin{pmatrix} Y_1 \\ Y_2 \\ Y_3 \end{pmatrix}.$$

In other words,

$$\widehat{\theta} = \frac{Y_1 + 2Y_2 + Y_3}{6},$$
$$\widehat{\phi} = \frac{-Y_2 + 2Y_3}{5}.$$

#4, page 41: The design matrix is

$$\boldsymbol{X} = \begin{pmatrix} 1 & x_1 & 3x_1^2 - 2 \\ 1 & x_2 & 3x_2^2 - 2 \\ 1 & x_3 & 3x_3^2 - 2 \end{pmatrix} = \begin{pmatrix} 1 & -1 & 1 \\ 1 & 0 & -2 \\ 1 & 1 & 1 \end{pmatrix}$$

The matrix $(\boldsymbol{X}'\boldsymbol{X})^{-1}$ is the diagonal matrix with respective entries 3, 2, and 5. Therefore,

$$\widehat{\boldsymbol{\beta}} = \begin{pmatrix} 3 & 3 & 3 \\ -2 & 0 & 2 \\ 5 & -10 & 5 \end{pmatrix} \boldsymbol{Y}.$$

So,

$$\widehat{\beta_0} = 3Y_1 + 3Y_2 + 3Y_3,$$
$$\widehat{\beta_1} = -2Y_1 + 2Y_3,$$
$$\widehat{\beta_2} = 5Y_1 - 10Y_2 + Y_3.$$

If we knew that $\beta_2 = 0$, then the design matrix is

$$\boldsymbol{X} = \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix} \implies (\boldsymbol{X}'\boldsymbol{X})^{-1} = \begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix}.$$

The rest is easy.

#1, page 49: We can write $Y_i = \theta + \varepsilon_i$ where $\boldsymbol{\varepsilon} \sim N_n(\mathbf{0}, \sigma^2 \mathbf{I}_n)$. So this is a linear model with design matrix, $\boldsymbol{X} = \mathbf{1}_n$ [the n-vector of all ones]. Note that $\boldsymbol{X}'\boldsymbol{X} = n$, so its inverse is $(1/n)$. Therefore,

$$\widehat{\theta} = \frac{1}{n}\boldsymbol{X}'\boldsymbol{Y} = \overline{Y}.$$

Therefore, the ith coordinate of $\boldsymbol{Y} - \widehat{\boldsymbol{\theta}}$ is $Y_i - \overline{Y}$. Because $\text{rank}(\boldsymbol{X}) = p = 1$,

$$S^2 = \frac{\left\| \boldsymbol{Y} - \widehat{\boldsymbol{\theta}} \right\|^2}{n-1} = \frac{1}{n-1}\sum_{i=1}^{n}(Y_i - \overline{Y})^2.$$

Theorem 3.5 does the rest.

#2, page 49: We are asked to prove the independence of the random variables $\|\boldsymbol{X}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta})\|^2$ and $\|\boldsymbol{Y} - \boldsymbol{X}\widehat{\boldsymbol{\beta}}\|^2$. So, we can try to prove that $\boldsymbol{U} = \boldsymbol{X}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta})$ and $\boldsymbol{V} = \boldsymbol{Y} - \boldsymbol{X}\widehat{\boldsymbol{\beta}}$ are independent.

Recall that $\boldsymbol{X}\widehat{\boldsymbol{\beta}} = \mathbf{P}_{\mathcal{C}(\boldsymbol{X})}\boldsymbol{Y}$. This proves that

$$\boldsymbol{U} = \mathbf{P}\boldsymbol{Y} - \boldsymbol{\theta}, \quad \text{and} \quad \boldsymbol{V} = \mathbf{P}_{\perp}\boldsymbol{Y},$$

where $\mathbf{P}$ and $\mathbf{P}_{\perp}$ denote projection onto $\mathcal{C}(\boldsymbol{X})$ and $\mathcal{C}(\boldsymbol{X})^{\perp}$, respectively. This proves that $(\boldsymbol{U}, \boldsymbol{V})'$ is multivariate normal because

$$\begin{pmatrix} \boldsymbol{U} \\ \boldsymbol{V} \end{pmatrix} = \begin{pmatrix} \mathbf{P} & \mathbf{0} \\ \mathbf{0} & \mathbf{P}_{\perp} \end{pmatrix} \boldsymbol{Y} - \begin{pmatrix} \boldsymbol{\theta} \\ \mathbf{0} \end{pmatrix}.$$

It also proves that $\boldsymbol{U}$ and $\boldsymbol{V}$ are independent because

$$\text{Cov}(\boldsymbol{U}, \boldsymbol{V}) = \text{Cov}(\mathbf{P}\boldsymbol{Y} - \boldsymbol{\theta}, \mathbf{P}_{\perp}\boldsymbol{Y}) = \mathbf{P}\text{Var}(\boldsymbol{Y})\mathbf{P}_{\perp} = \sigma^2\mathbf{P}\mathbf{P}_{\perp}.$$

But $\mathbf{P}\mathbf{P}_{\perp} = \mathbf{0}$, whence the result.