Math 6010
Partial solutions to homework 6

8. The model is $Y_i = dx_i + \varepsilon_i\sqrt{f(x_i)}$, where $\mathrm{E}[\varepsilon_i] = 0$ and $\mathrm{Var}(\varepsilon_i) = \sigma^2$. We can standardize this model into a basic general linear model as follows:

$$\overline{Y}_i = d\overline{x}_i + \varepsilon_i,$$

where

$$\overline{Y}_i := \frac{Y_i}{\sqrt{f(x_i)}} \quad \text{and} \quad \overline{x_i} = \frac{x_i}{\sqrt{f(x_i)}}.$$

Of course, we are assuming that $f(x_i) > 0$ for all i. Minimize over all d the random variable,

$$f(d) := \sum_{i=1}^{n} \left(\overline{Y}_i - d\overline{x}_i\right)^2.$$

It is not hard to convince yourself that the minimum is achieved by setting f' to zero, where $f'(d) = -2\sum_{i=1}^{n}\left(\overline{Y}_i - d\overline{x}_i\right)\overline{x}_i$. This leads to

$$\hat{d} = \frac{\sum_{i=1}^{n}\overline{Y}_i\overline{x}_i}{\sum_{i=1}^{n}\overline{x}_i^2},$$

as long as $\sum_{i=1}^{n}\overline{x}_i^2 \neq 0$.

(a) If $f(x) = 1$, then $\overline{Y}_i = Y_i$ and $\overline{x}_i = x_i$; therefore,

$$\hat{d} = \frac{\sum_{i=1}^{n} Y_i x_i}{\sum_{i=1}^{n} x_i^2},$$

as long as $\sum_{i=1}^{n}\overline{x}_i^2 \neq 0$.

(b) If $f(x) = x$ then $\overline{Y}_i = Y_i/\sqrt{x_i}$ and $\overline{x}_i = \sqrt{x_i}$; therefore,

$$\hat{d} = \frac{\sum_{i=1}^{n} Y_i}{\sum_{i=1}^{n} x_i}.$$

[This makes sense only when $x_i > 0$.]

(c) If $f(x) = x^2$ then $\overline{Y}_i = Y_i/x_i$ and $\overline{x}_i = 1$; therefore,

$$\hat{d} = \frac{1}{n}\sum_{i=1}^{n} \frac{Y_i}{x_i}.$$

[This makes sense only when $x_i > 0$.]

10. Let

$$f(\boldsymbol{b}) := (\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{b})'\boldsymbol{V}^{-1}(\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{b}).$$

Also, suppose $\boldsymbol{\beta}$ solves

$$\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X}\boldsymbol{\beta} = \boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{Y}. \tag{1}$$

Then for all $\boldsymbol{b}$ we can write

$$\begin{aligned} f(\boldsymbol{b}) &= (\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{X}[\boldsymbol{\beta} - \boldsymbol{b}])'\boldsymbol{V}^{-1}(\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{X}[\boldsymbol{\beta} - \boldsymbol{b}]) \\ &= f(\boldsymbol{\beta}) + (\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{\beta})'\boldsymbol{V}^{-1}\boldsymbol{X}[\boldsymbol{\beta} - \boldsymbol{b}] + [\boldsymbol{\beta} - \boldsymbol{b}]'\boldsymbol{X}'\boldsymbol{V}^{-1}(\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{\beta}) \\ &\qquad + [\boldsymbol{\beta} - \boldsymbol{b}]'\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X}[\boldsymbol{\beta} - \boldsymbol{b}]. \end{aligned}$$

The last quantity is ≥ 0 because $\boldsymbol{V}$—and hence $\boldsymbol{V}^{-1}$—is nonnegative definite. The second and the third quantity are equal because they are scalar quantities and transposes of each other. This means that

$$f(\boldsymbol{b}) \geq f(\boldsymbol{\beta}) + 2(\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{\beta})'\boldsymbol{V}^{-1}\boldsymbol{X}[\boldsymbol{\beta} - \boldsymbol{b}],$$

for all $\boldsymbol{b}$. Now

$$\begin{aligned} (\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{\beta})'\boldsymbol{V}^{-1}\boldsymbol{X}[\boldsymbol{\beta} - \boldsymbol{b}] &= \boldsymbol{Y}'\boldsymbol{V}^{-1}\boldsymbol{X}[\boldsymbol{\beta} - \boldsymbol{b}] - \boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X}[\boldsymbol{\beta} - \boldsymbol{b}] \\ &= \boldsymbol{Y}'\boldsymbol{V}^{-1}\boldsymbol{X}[\boldsymbol{\beta} - \boldsymbol{b}] - \boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X}[\boldsymbol{\beta} - \boldsymbol{b}] \\ &= 0, \end{aligned}$$

thanks to (1). Therefore, $f(\boldsymbol{b}) \geq f(\boldsymbol{\beta})$ for all $\boldsymbol{b}$, which is another way to say that any $\boldsymbol{\beta}$ which solves (1) is a minimizer of f.

11. The parameter vector is $\boldsymbol{\beta} = [\theta_1, \theta_2]'$ and the design matrix is

$$\boldsymbol{X} = \begin{pmatrix} 1 & 1 \\ 1 & -2 \\ 2 & -1 \end{pmatrix}.$$

$$\boldsymbol{X}'\boldsymbol{X} = \begin{pmatrix} 6 & -3 \\ -3 & 6 \end{pmatrix} \Rightarrow (\boldsymbol{X}'\boldsymbol{X})^{-1} = \begin{pmatrix} 2/9 & 1/9 \\ 1/9 & 2/9 \end{pmatrix}.$$

Therefore,

$$\widehat{\boldsymbol{\beta}} = \begin{pmatrix} 2/9 & 1/9 \\ 1/9 & 2/9 \end{pmatrix}\begin{pmatrix} 1 & 1 & 2 \\ 1 & -2 & -1 \end{pmatrix}\begin{pmatrix} Y_1 \\ Y_2 \\ Y_3 \end{pmatrix} = \begin{pmatrix} \dfrac{Y_1 + Y_3}{3} \\[2ex] \dfrac{Y_1 - Y_2}{3} \end{pmatrix}.$$

In other words,

$$\widehat{\theta}_1 = \frac{Y_1 + Y_3}{3} \quad \text{and} \quad \widehat{\theta}_2 = \frac{Y_1 - Y_2}{3}.$$

14. It might be better to proceed directly to find $\widehat{\boldsymbol{\beta}}$ in this case, since the problem is pretty modest in size. Incidentally, the answer in the back of your text is manifestly false, starting with the claim that $\boldsymbol{X}'\boldsymbol{X}$ is diagonal. In fact, here,

$$\boldsymbol{X} = \begin{pmatrix} 1 & C_1 & S_1 \\ 1 & C_2 & S_2 \\ 1 & C_3 & S_3 \end{pmatrix},$$

where

$$C_i := \cos\left(\frac{2\pi k_1 i}{n}\right), \qquad S_i := \sin\left(\frac{2\pi k_2 i}{n}\right) \qquad \text{for } 1 \le i \le n.$$

So $\boldsymbol{X}'\boldsymbol{X}$ is diagonal only for very special choices of k_1 and k_2. Also, $\widehat{\beta}_0 \neq \overline{Y}$ etc., as we will see soon.

The least squares problem is to solve the minimization problem,

$$\min_{\beta_0,\beta_1,\beta_2} \sum_{i=1}^n \left[Y_i - \beta_0 - \beta_1 C_i - \beta_2 S_i\right]^2 := \min_{\beta_0,\beta_1,\beta_2} f(\beta_0\,,\beta_1\,,\beta_2).$$

It is easier to solve this directly, rather than to invert $\boldsymbol{X}'\boldsymbol{X}$, etc. The function f is a positive quadratic, therefore it suffices to set $\partial f/\partial\beta_i = 0$ for $i = 0, 1, 2$. Now,

$$\frac{\partial f}{\partial \beta_0} = -2\sum_{i=1}^n \left[Y_i - \beta_0 - \beta_1 C_i - \beta_2 S_i\right] = -2n\left[\overline{Y} - \beta_0 - \beta_1\overline{C} - \beta_2\overline{S}\right],$$

where $\overline{Y} := n^{-1}\sum_{i=1}^n Y_i$, $\overline{C} := n^{-1}\sum_{i=1}^n C_i$, and $\overline{S} := n^{-1}\sum_{i=1}^n S_i$.

Next, we compute

$$\frac{\partial f}{\partial \beta_1} = -2\sum_{i=1}^n \left[Y_i - \beta_0 - \beta_1 C_i - \beta_2 S_i\right] C_i = -2n\left[\overline{YC} - \beta_0\overline{C} - \beta_1\overline{C^2} - \beta_2\overline{SC}\right],$$

where $\overline{YC} := n^{-1}\sum_{i=1}^n Y_i C_i$, $\overline{C^2} := n^{-1}\sum_{i=1}^n C_i^2$ and $\overline{SC} := n^{-1}\sum_{i=1}^n S_i C_i$.

Finally,

$$\frac{\partial f}{\partial \beta_2} = -2\sum_{i=1}^n \left[Y_i - \beta_0 - \beta_1 C_i - \beta_2 S_i\right] S_i = -2n\left[\overline{YS} - \beta_0\overline{S} - \beta_1\overline{SC} - \beta_2\overline{S^2}\right],$$

where $\overline{YS} := n^{-1}\sum_{i=1}^n Y_i S_i$ etc.

Set the preceding three $\partial f/\partial\beta_i$'s equal to zero and solve to obtain a 3×3 linear system,

$$\begin{pmatrix} 1 & \overline{C} & \overline{S} \\ \overline{C} & \overline{C^2} & \overline{SC} \\ \overline{S} & \overline{SC} & \overline{S^2} \end{pmatrix} \widehat{\boldsymbol{\beta}} = \begin{pmatrix} \overline{Y} \\ \overline{YC} \\ \overline{YS} \end{pmatrix}.$$

Therefore,

$$\begin{aligned} \widehat{\boldsymbol{\beta}} &= \begin{pmatrix} 1 & \overline{C} & \overline{S} \\ \overline{C} & \overline{C^2} & \overline{SC} \\ \overline{S} & \overline{SC} & \overline{S^2} \end{pmatrix}^{-1} \begin{pmatrix} \overline{Y} \\ \overline{YC} \\ \overline{YS} \end{pmatrix} \\ &= \frac{1}{\Delta} \begin{pmatrix} \overline{C^2}\cdot\overline{S^2} - (\overline{SC})^2 & \overline{C}\cdot\overline{S^2} - \overline{S}\cdot\overline{SC} & \overline{C}\cdot\overline{SC} - \overline{C^2}\cdot\overline{S} \\ \overline{SC} - \overline{C}\cdot\overline{S} & \overline{S^2} - (\overline{S})^2 & \overline{SC} - \overline{C}\cdot\overline{S} \\ \overline{C}\cdot\overline{SC} - \overline{C^2}\cdot\overline{S} & \overline{S^2} - (\overline{S})^2 & \overline{C^2} - (\overline{C})^2 \end{pmatrix} \begin{pmatrix} \overline{Y} \\ \overline{YC} \\ \overline{YS} \end{pmatrix}, \end{aligned}$$

where

$$\Delta = \overline{C^2}\cdot\overline{S^2} - (\overline{SC})^2 - \overline{C}\left(\overline{C}\cdot\overline{S^2} - \overline{S}\cdot\overline{SC}\right) + \overline{S}\left(\overline{C}\cdot\overline{SC} - \overline{S}\cdot\overline{C^2}\right).$$

[This is not equal to zero, so all is well.]