Chase Joyner

802 Homework 2

February 8, 2017

Problem 1:

Show that an alternative expression for h_{ii} in Theorem 9.2(iii) is the following:

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

where θ_{ir} is the angle between $\mathbf{x}_{1i}$ and $\mathbf{a}_r$, the rth eigenvector of $\mathbf{X}_c'\mathbf{X}_c$. Thus h_{ii} is large if $(\mathbf{x}_{1i} - \overline{\mathbf{x}}_1)'(\mathbf{x}_{1i} - \overline{\mathbf{x}}_1)$ is large or if θ_{ir} is small for some r.

Solution: First recall the expression for h_{ii} is given by

$$h_{ii} = \frac{1}{n} + (\mathbf{x}_{1i} - \overline{\mathbf{x}}_1)'(\mathbf{X}_c'\mathbf{X}_c)^{-1}(\mathbf{x}_{1i} - \overline{\mathbf{x}}_1).$$

By equations (2.101) and (2.104) in the textbook, we can write

$$(\mathbf{X}_c'\mathbf{X}_c)^{-1} = \sum_{r=1}^{k} \frac{1}{\lambda_r} \mathbf{a}_r \mathbf{a}_r'.$$

Therefore, plugging this in, we have

$$\begin{aligned} h_{ii} &= \frac{1}{n} + (\mathbf{x}_{1i} - \overline{\mathbf{x}}_1)' \left(\sum_{r=1}^{k} \frac{1}{\lambda_r} \mathbf{a}_r \mathbf{a}_r' \right) (\mathbf{x}_{1i} - \overline{\mathbf{x}}_1) \\ &= \frac{1}{n} + \sum_{r=1}^{k} \frac{1}{\lambda_r} \left[(\mathbf{x}_{1i} - \overline{\mathbf{x}}_1)'\mathbf{a}_r \right] \left[\mathbf{a}_r'(\mathbf{x}_{1i} - \overline{\mathbf{x}}_1) \right] \\ &= \frac{1}{n} + \sum_{r=1}^{k} \frac{1}{\lambda_r} \left[(\mathbf{x}_{1i} - \overline{\mathbf{x}}_1)'\mathbf{a}_r \right]^2. \end{aligned}$$

Now, by the law of cosines, we know that if θ is the angle between vectors $\mathbf{a}$ and $\mathbf{b}$, then

$$\cos \theta = \frac{a \cdot b}{|a||b|} = \frac{a'b}{\sqrt{a'a}\sqrt{b'b}}.$$

Therefore, we have for our purposes that

$$\cos \theta_{ir} = \frac{(\mathbf{x}_{1i} - \overline{\mathbf{x}}_1)'\mathbf{a}_r}{\sqrt{(\mathbf{x}_{1i} - \overline{\mathbf{x}}_1)'(\mathbf{x}_{1i} - \overline{\mathbf{x}}_1)\mathbf{a}_r'\mathbf{a}_r}}.$$

Squaring both sides, we obtain

$$\cos^2\theta_{ir} = \frac{[(\mathbf{x}_{1i}-\overline{\mathbf{x}}_1)'\mathbf{a}_r]^2}{(\mathbf{x}_{1i}-\overline{\mathbf{x}}_1)'(\mathbf{x}_{1i}-\overline{\mathbf{x}}_1)\mathbf{a}_r'\mathbf{a}_r}$$

and hence

$$(\mathbf{x}_{1i}-\overline{\mathbf{x}}_1)'(\mathbf{x}_{1i}-\overline{\mathbf{x}}_1)\mathbf{a}_r'\mathbf{a}_r\cos^2\theta_{ir} = \left[(\mathbf{x}_{1i}-\overline{\mathbf{x}}_1)'\mathbf{a}_r\right]^2.$$

Noting that $\mathbf{a}_r'\mathbf{a}_r = 1$ since it is the eigenvector of a centered matrix, plugging into our simplified version of h_{ii}, we get that

$$\begin{aligned} h_{ii} &= \frac{1}{n} + \sum_{r=1}^{k}\frac{1}{\lambda_r}\left[(\mathbf{x}_{1i}-\overline{\mathbf{x}}_1)'\mathbf{a}_r\right]^2 \\ &= \frac{1}{n} + (\mathbf{x}_{1i}-\overline{\mathbf{x}}_1)'(\mathbf{x}_{1i}-\overline{\mathbf{x}}_1)\sum_{r=1}^{k}\frac{1}{\lambda_r}\cos^2\theta_{ir} \end{aligned}$$

which proves the result.

Problem 2:

Show that $\frac{1}{n} \le h_{ii} + \widehat{\varepsilon}_i^2/\widehat{\boldsymbol{\varepsilon}}'\widehat{\boldsymbol{\varepsilon}} \le 1$ as in (9.24). The following steps are suggested:

(a) Let $\mathbf{H}^\star$ be the hat matrix corresponding to the augmented matrix $(\mathbf{X},\mathbf{y})$. Then

$$\begin{aligned} \mathbf{H}^\star &= (\mathbf{X},\mathbf{y})\left[(\mathbf{X},\mathbf{y})'(\mathbf{X},\mathbf{y})\right]^{-1}(\mathbf{X},\mathbf{y})' \\ &= (\mathbf{X},\mathbf{y})\begin{pmatrix}\mathbf{X}'\mathbf{X} & \mathbf{X}'\mathbf{y} \\ \mathbf{y}'\mathbf{X} & \mathbf{y}'\mathbf{y}\end{pmatrix}^{-1}\begin{pmatrix}\mathbf{X}' \\ \mathbf{y}'\end{pmatrix}. \end{aligned}$$

Use the inverse of a partitioned matrix in (2.50) with $\mathbf{A}_{11} = \mathbf{X}'\mathbf{X}$, $\mathbf{a}_{12} = \mathbf{X}'\mathbf{y}$, $\mathbf{a}_{21} = \mathbf{y}'\mathbf{X}$, and $a_{22} = \mathbf{y}'\mathbf{y}$ to obtain

$$\begin{aligned} \mathbf{H}^\star &= \mathbf{H} + \frac{1}{b}\left[\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}\mathbf{y}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' - \mathbf{y}\mathbf{y}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}\mathbf{y}' + \mathbf{y}\mathbf{y}'\right] \\ &= \mathbf{H} + \frac{1}{b}\left[\mathbf{H}\mathbf{y}\mathbf{y}'\mathbf{H} - \mathbf{y}\mathbf{y}'\mathbf{H} - \mathbf{H}\mathbf{y}\mathbf{y}' + \mathbf{y}\mathbf{y}'\right], \end{aligned}$$

where $b = \mathbf{y}'\mathbf{y} - \mathbf{y}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$.

Solution: Using (2.50) and the partitions defined above, we have

$$\mathbf{H}^\star = \frac{1}{b}(\mathbf{X},\mathbf{y})\begin{pmatrix} b\mathbf{A}_{11}^{-1} + \mathbf{A}_{11}^{-1}\mathbf{a}_{12}\mathbf{a}_{21}\mathbf{A}_{11}^{-1} & -\mathbf{A}_{11}^{-1}\mathbf{a}_{12} \\ -\mathbf{a}_{21}\mathbf{A}_{11}^{-1} & 1 \end{pmatrix}\begin{pmatrix}\mathbf{X}' \\ \mathbf{y}'\end{pmatrix}$$

where $b = a_{22} - \mathbf{a}_{21}\mathbf{A}_{11}^{-1}\mathbf{a}_{12}$. Multiplying further, we get

$$\begin{aligned} \mathbf{H}^\star &= \frac{1}{b}(\mathbf{X},\mathbf{y})\begin{pmatrix} b\mathbf{A}_{11}^{-1}\mathbf{X}' + \mathbf{A}_{11}^{-1}\mathbf{a}_{12}\mathbf{a}_{21}\mathbf{A}_{11}^{-1}\mathbf{X}' - \mathbf{A}_{11}^{-1}\mathbf{a}_{12}\mathbf{y}' \\ -\mathbf{a}_{21}\mathbf{A}_{11}^{-1}\mathbf{X}' + \mathbf{y}' \end{pmatrix} \\ &= \frac{1}{b}\left[b\mathbf{X}\mathbf{A}_{11}^{-1}\mathbf{X}' + \mathbf{X}\mathbf{A}_{11}^{-1}\mathbf{a}_{12}\mathbf{a}_{21}\mathbf{A}_{11}^{-1}\mathbf{X}' - \mathbf{X}\mathbf{A}_{11}^{-1}\mathbf{a}_{12}\mathbf{y}' - \mathbf{y}\mathbf{a}_{21}\mathbf{A}_{11}^{-1}\mathbf{X}' + \mathbf{y}\mathbf{y}'\right] \\ &= \mathbf{H} + \frac{1}{b}\left[\mathbf{H}\mathbf{y}\mathbf{y}'\mathbf{H} - \mathbf{H}\mathbf{y}\mathbf{y}' - \mathbf{y}\mathbf{y}'\mathbf{H} + \mathbf{y}\mathbf{y}'\right]. \end{aligned}$$

(b) Show that the above expression factors into

$$\mathbf{H}^\star = \mathbf{H} + \frac{(\mathbf{I}-\mathbf{H})\mathbf{y}\mathbf{y}'(\mathbf{I}-\mathbf{H})}{\mathbf{y}'(\mathbf{I}-\mathbf{H})\mathbf{y}} = \mathbf{H} + \frac{\widehat{\boldsymbol{\varepsilon}}\widehat{\boldsymbol{\varepsilon}}'}{\widehat{\boldsymbol{\varepsilon}}'\widehat{\boldsymbol{\varepsilon}}},$$

which gives $h_{ii}^\star = h_{ii} + \widehat{\varepsilon}_i^2/\widehat{\boldsymbol{\varepsilon}}'\widehat{\boldsymbol{\varepsilon}}$.

Solution: Clearly we can factor b into

$$b = \mathbf{y}'\mathbf{y} - \mathbf{y}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} = \mathbf{y}'(\mathbf{I}-\mathbf{H})\mathbf{y}.$$

Now, notice that

$$\begin{aligned}
\mathbf{H}\mathbf{y}\mathbf{y}'\mathbf{H} - \mathbf{y}\mathbf{y}'\mathbf{H} - \mathbf{H}\mathbf{y}\mathbf{y}' + \mathbf{y}\mathbf{y}' &= (\mathbf{I}-\mathbf{H})\mathbf{y}\mathbf{y}' + \mathbf{H}\mathbf{y}\mathbf{y}'\mathbf{H} - \mathbf{y}\mathbf{y}'\mathbf{H} \\
&= (\mathbf{I}-\mathbf{H})\mathbf{y}\mathbf{y}' - (\mathbf{I}-\mathbf{H})\mathbf{y}\mathbf{y}'\mathbf{H} \\
&= (\mathbf{I}-\mathbf{H})(\mathbf{y}\mathbf{y}' - \mathbf{y}\mathbf{y}'\mathbf{H}) \\
&= (\mathbf{I}-\mathbf{H})\mathbf{y}\mathbf{y}'(\mathbf{I}-\mathbf{H}).
\end{aligned}$$

Therefore, we have

$$\begin{aligned}
\mathbf{H}^\star &= \mathbf{H} + \frac{1}{b}\left[\mathbf{H}\mathbf{y}\mathbf{y}'\mathbf{H} - \mathbf{y}\mathbf{y}'\mathbf{H} - \mathbf{H}\mathbf{y}\mathbf{y}' + \mathbf{y}\mathbf{y}'\right] \\
&= \mathbf{H} + \frac{(\mathbf{I}-\mathbf{H})\mathbf{y}\mathbf{y}'(\mathbf{I}-\mathbf{H})}{\mathbf{y}'(\mathbf{I}-\mathbf{H})\mathbf{y}} = \mathbf{H} + \frac{\widehat{\boldsymbol{\varepsilon}}\widehat{\boldsymbol{\varepsilon}}'}{\widehat{\boldsymbol{\varepsilon}}'\widehat{\boldsymbol{\varepsilon}}}
\end{aligned}$$

(c) The proof is easily completed by noting that $\mathbf{H}^\star$ is a hat matrix and therefore $(1/n) \le h_{ii}^\star \le 1$ by Theorem 9.2(i).

Solution: By construction of $\mathbf{H}^\star$ in part (a), we notice it is a hat matrix and so by theorem 9.2,

$$\frac{1}{n} \le h_{ii}^\star \le 1$$

for all $i = 1, ..., n$. Therefore, by part (b), we have

$$\frac{1}{n} \le h_{ii} + \widehat{\varepsilon}_i^2/\widehat{\boldsymbol{\varepsilon}}'\widehat{\boldsymbol{\varepsilon}} \le 1.$$

This proves the overall result.

Problem 3:

Show that $\widehat{\boldsymbol{\beta}}_{(i)} = \widehat{\boldsymbol{\beta}} - \widehat{\varepsilon}_i(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_i/(1-h_{ii})$ as in (9.29). The following steps are suggested:

(a) Show that $\mathbf{X}'\mathbf{X} = \mathbf{X}_{(i)}'\mathbf{X}_{(i)} + \mathbf{x}_i\mathbf{x}_i'$ and that $\mathbf{X}'\mathbf{y} = \mathbf{X}_{(i)}'\mathbf{y}_{(i)} + \mathbf{x}_i y_i$.

Solution: Suppose that $\mathbf{X}$ is $n \times k$ and define the row vectors

$$\mathbf{x}_i' = (1, x_{i1}, x_{i2}, ..., x_{ik})$$

for $i = 1, ..., n$. In this fashion, we can write

$$\mathbf{X} = \begin{pmatrix} \mathbf{x}_1' \\ \mathbf{x}_2' \\ \vdots \\ \mathbf{x}_n' \end{pmatrix}.$$

Then, notice that

$$\begin{aligned}\mathbf{X}'\mathbf{X} &= \begin{pmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_k \end{pmatrix} \begin{pmatrix} \mathbf{x}_1' \\ \mathbf{x}_2' \\ \vdots \\ \mathbf{x}_k' \end{pmatrix} = \sum_{j=1}^{k} \mathbf{x}_j \mathbf{x}_j' = \sum_{\substack{j=1 \\ j \neq i}}^{k} \mathbf{x}_j \mathbf{x}_j' + \mathbf{x}_i \mathbf{x}_i' \\ &= \mathbf{X}_{(i)}' \mathbf{X}_{(i)} + \mathbf{x}_i \mathbf{x}_i'.\end{aligned}$$

Similarly, we find that

$$\begin{aligned}\mathbf{X}'\mathbf{y} &= \begin{pmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_k \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \sum_{j=1}^{k} \mathbf{x}_j y_j = \sum_{\substack{j \neq i \\ j=1}}^{k} \mathbf{x}_j y_j + \mathbf{x}_i y_i \\ &= \mathbf{X}_{(i)} y_{(i)} + \mathbf{x}_i y_i.\end{aligned}$$

(b) Show that $(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}_{(i)}'\mathbf{y}_{(i)} = \widehat{\boldsymbol{\beta}} - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_i y_i$.

Solution: Recall that

$$\widehat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}.$$

Therefore, the result follows immediately from part (a), i.e.

$$\begin{aligned}\widehat{\boldsymbol{\beta}} &= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} = (\mathbf{X}'\mathbf{X})^{-1}\left[\mathbf{X}_{(i)} y_{(i)} + \mathbf{x}_i y_i\right] \\ &= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}_{(i)} y_{(i)} + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_i y_i.\end{aligned}$$

Subtracting over, we obtain the result.

(c) Using the following adaptation of (2.53),

$$(\mathbf{B} - \mathbf{c}\mathbf{c}')^{-1} = \mathbf{B}^{-1} + \frac{\mathbf{B}^{-1}\mathbf{c}\mathbf{c}'\mathbf{B}^{-1}}{1 - \mathbf{c}'\mathbf{B}^{-1}\mathbf{c}},$$

show that

$$\widehat{\boldsymbol{\beta}}_{(i)} = \left[(\mathbf{X}'\mathbf{X})^{-1} + \frac{(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_i\mathbf{x}_i'(\mathbf{X}'\mathbf{X})^{-1}}{1 - h_{ii}}\right]\mathbf{X}_{(i)}'\mathbf{y}_{(i)}.$$

Solution: Note that $\widehat{\boldsymbol{\beta}}_{(i)}$ is the estimate excluding the ith observation. Therefore, by part (a) and (2.53), we find

$$\begin{aligned}\widehat{\boldsymbol{\beta}}_{(i)} &= \left(\mathbf{X}_{(i)}'\mathbf{X}_{(i)}\right)^{-1}\mathbf{X}_{(i)}'\mathbf{y}_{(i)} = \left(\mathbf{X}'\mathbf{X} - \mathbf{x}_i\mathbf{x}_i'\right)^{-1}\mathbf{X}_{(i)}'\mathbf{y}_{(i)} \\ &= \left[(\mathbf{X}'\mathbf{X})^{-1} + \frac{(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_i\mathbf{x}_i'(\mathbf{X}'\mathbf{X})^{-1}}{1 - \mathbf{x}_i'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_i}\right]\mathbf{X}_{(i)}'\mathbf{y}_{(i)}.\end{aligned}$$

Recalling that $\mathbf{H} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}$, its ith diagonal element is $h_{ii} = \mathbf{x}_i'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_i$. Thus,

$$\widehat{\boldsymbol{\beta}}_{(i)} = \left[(\mathbf{X}'\mathbf{X})^{-1} + \frac{(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_i\mathbf{x}_i'(\mathbf{X}'\mathbf{X})^{-1}}{1-h_{ii}}\right]\mathbf{X}_{(i)}'\mathbf{y}_{(i)}.$$

(d) Using the result of parts (b) and (c), show that

$$\widehat{\boldsymbol{\beta}}_{(i)} = \widehat{\boldsymbol{\beta}} - \frac{\widehat{\varepsilon}_i}{1-h_{ii}}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_i.$$

Solution: By parts (b) and (c), we have

$$\begin{aligned}
\widehat{\boldsymbol{\beta}}_{(i)} &= \left[(\mathbf{X}'\mathbf{X})^{-1} + \frac{(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_i\mathbf{x}_i'(\mathbf{X}'\mathbf{X})^{-1}}{1-h_{ii}}\right]\mathbf{X}_{(i)}'\mathbf{y}_{(i)} \\
&= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}_{(i)}'\mathbf{y}_{(i)} + \frac{(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_i\mathbf{x}_i'}{1-h_{ii}}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}_{(i)}'\mathbf{y}_{(i)} \\
&= \widehat{\boldsymbol{\beta}} - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_i y_i + \frac{(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_i\mathbf{x}_i'}{1-h_{ii}}\left[\widehat{\boldsymbol{\beta}} - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_i y_i\right] \\
&= \widehat{\boldsymbol{\beta}} - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_i y_i + \frac{(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_i}{1-h_{ii}}\left[\mathbf{x}_i'\widehat{\boldsymbol{\beta}} - \mathbf{x}_i'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_i y_i\right] \\
&= \widehat{\boldsymbol{\beta}} - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_i y_i + \frac{(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_i}{1-h_{ii}}\left[\widehat{y}_i - h_{ii}y_i\right] \\
&= \widehat{\boldsymbol{\beta}} - \frac{\widehat{y}_i - y_i}{1-h_{ii}}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_i \\
&= \widehat{\boldsymbol{\beta}} - \frac{\widehat{\varepsilon}_i}{1-h_{ii}}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_i.
\end{aligned}$$

Problem 4:

Show that $\mathbf{S}$ in (10.14) can be found as $\mathbf{S} = \sum_{i=1}^{n}(\mathbf{v}_i - \overline{\mathbf{v}})(\mathbf{v}_i - \overline{\mathbf{v}})'/(n-1)$ as in (10.13).

Solution: Here, $\mathbf{v}_i$ and $\overline{\mathbf{v}}$ denote

$$\mathbf{v}_i = \begin{pmatrix} y_i \\ \mathbf{x}_i \end{pmatrix} \quad \text{and} \quad \overline{\mathbf{v}} = \begin{pmatrix} \overline{y} \\ \overline{\mathbf{x}} \end{pmatrix}.$$

Then, notice that

$$(\mathbf{v}_i - \overline{\mathbf{v}})(\mathbf{v}_i - \overline{\mathbf{v}})' = \begin{pmatrix} y_i - \overline{y} \\ x_{i1} - \overline{x}_1 \\ \vdots \\ x_{ik} - \overline{x}_k \end{pmatrix} \begin{pmatrix} y_i - \overline{y} & x_{i1} - \overline{x}_1 & \cdots & x_{ik} - \overline{x}_k \end{pmatrix}$$

$$= \begin{pmatrix} (y_i - \overline{y})(y_i - \overline{y}) & (y_i - \overline{y})(x_{i1} - \overline{x}_1) & \cdots\cdots & (y_i - \overline{y})(x_{ik} - \overline{x}_k) \\ (x_{i1} - \overline{x}_1)(y_i - \overline{y}) & (x_{i1} - \overline{x}_1)(x_{i1} - \overline{x}_1)) & \cdots\cdots & (x_{i1} - \overline{x}_1)(x_{ik} - \overline{x}_k) \\ \vdots & \vdots & \ddots & \vdots \\ (x_{ik} - \overline{x}_k)(y_i - \overline{y}) & (x_{ik} - \overline{x}_k)(x_{i1} - \overline{x}_1) & \cdots\cdots & (x_{ik} - \overline{x}_k)(x_{ik} - \overline{x}_k) \end{pmatrix}.$$

Summing up this matrix and dividing by $n-1$ gives the result.

Problem 5:

Show that $\text{Cov}(y, w) = \boldsymbol{\sigma}'_{yx}\boldsymbol{\Sigma}_{xx}^{-1}\boldsymbol{\sigma}_{yx}$ and $\text{Var}(w) = \boldsymbol{\sigma}'_{yx}\boldsymbol{\Sigma}_{xx}^{-1}\boldsymbol{\sigma}_{yx}$ as in (10.26), where $w = \mu_y + \boldsymbol{\sigma}'_{yx}\boldsymbol{\Sigma}_{xx}^{-1}(\mathbf{x} - \boldsymbol{\mu}_x)$.

Solution: This result is trivial since

$$\text{Cov}(y, w) = \text{Cov}\left(y, \boldsymbol{\sigma}'_{yx}\boldsymbol{\Sigma}_{xx}^{-1}\mathbf{x}\right) = \boldsymbol{\sigma}'_{yx}\boldsymbol{\Sigma}_{xx}^{-1}\text{Cov}(y, \mathbf{x}) = \boldsymbol{\sigma}'_{yx}\boldsymbol{\Sigma}_{xx}^{-1}\boldsymbol{\sigma}_{yx}$$

and also

$$\begin{aligned}
\text{Var}(w) = \text{Cov}(w, w) &= \text{Cov}\left(\boldsymbol{\sigma}'_{yx}\boldsymbol{\Sigma}_{xx}^{-1}\mathbf{x}, \boldsymbol{\sigma}'_{yx}\boldsymbol{\Sigma}_{xx}^{-1}\mathbf{x}\right) \\
&= \boldsymbol{\sigma}'_{yx}\boldsymbol{\Sigma}_{xx}^{-1}\text{Cov}(\mathbf{x}, \mathbf{x})\boldsymbol{\Sigma}_{xx}^{-1}\boldsymbol{\sigma}_{yx} \\
&= \boldsymbol{\sigma}'_{yx}\boldsymbol{\Sigma}_{xx}^{-1}\boldsymbol{\Sigma}_{xx}\boldsymbol{\Sigma}_{xx}^{-1}\boldsymbol{\sigma}_{yx} \\
&= \boldsymbol{\sigma}'_{yx}\boldsymbol{\Sigma}_{xx}^{-1}\boldsymbol{\sigma}_{yx}.
\end{aligned}$$

Problem 6:

Verify that R^2 can be expressed in terms of determinants as in (10.40) and (10.41).

Solution: First, by (2.72), a partitioned matrix such as

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix}$$

has the determinant given by

$$|\mathbf{A}| = |\mathbf{A}_{22}||\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21}|.$$

Then, recalling that

$$\mathbf{S} = \begin{pmatrix} s_{yy} & \mathbf{s}'_{yx} \\ \mathbf{s}_{yx} & \mathbf{S}_{xx} \end{pmatrix}$$

we have its determinant is given by

$$\begin{aligned}
|\mathbf{S}| &= |\mathbf{S}_{xx}||s_{yy} - \mathbf{s}'_{yx}\mathbf{S}_{xx}^{-1}\mathbf{s}_{yx}| \\
&= |\mathbf{S}_{xx}|(s_{yy} - \mathbf{s}'_{yx}\mathbf{S}_{xx}^{-1}\mathbf{s}_{yx}).
\end{aligned}$$

Notice that we can reformulate this as

$$\frac{\mathbf{s}'_{yx}\mathbf{S}_{xx}^{-1}\mathbf{s}_{yx}}{s_{yy}} = 1 - \frac{|\mathbf{S}|}{|\mathbf{S}_{xx}|s_{yy}}.$$

Now, we see that can write R^2 as

$$R^2 = \frac{\mathbf{s}'_{yx}\mathbf{S}_{xx}^{-1}\mathbf{s}_{yx}}{s_{yy}} = 1 - \frac{|\mathbf{S}|}{|\mathbf{S}_{xx}|s_{yy}}.$$

To show the second desired equality, we notice that

$$\mathbf{S} = \mathbf{DRD}$$

where $\mathbf{D}$ is the diagonal matrix with entries $\sqrt{s_{yy}}, \sqrt{s_{11}}, ..., \sqrt{s_{kk}}$. Also, we have

$$\mathbf{S}_{xx} = \mathbf{D}_x \mathbf{R}_{xx} \mathbf{D}_x$$

where $\mathbf{D}_x$ is the diagonal matrix with entries $\sqrt{s_{11}}, ..., \sqrt{s_{kk}}$. Therefore, we see that we can write the above as

$$\begin{aligned} R^2 &= 1 - \frac{|\mathbf{S}|}{|\mathbf{S}_{xx}| s_{yy}} = 1 - \frac{|\mathbf{DRD}|}{|\mathbf{D}_x \mathbf{R}_{xx} \mathbf{D}_x| s_{yy}} = 1 - \frac{|\mathbf{D}|^2 |\mathbf{R}|}{|\mathbf{D}_x|^2 |\mathbf{R}_{xx}| s_{yy}} \\ &= 1 - \frac{|\mathbf{D}|^2 |\mathbf{R}|}{|\mathbf{D}|^2 |\mathbf{R}_{xx}|} = 1 - \frac{|\mathbf{R}|}{|\mathbf{R}_{xx}|} \end{aligned}$$

since the determinant of a diagonal matrix is the product of the diagonal.

Problem 7:

Prove Theorem 10.7b. The theorem states: The linear function $t(\mathbf{x})$ that minimizes $E\left[\left(y - t(\mathbf{x})\right)^2\right]$ is given by $t(\mathbf{x}) = \beta_0 + \boldsymbol{\beta}_1' \mathbf{x}$, where

$$\begin{aligned} \beta_0 &= \mu_y - \boldsymbol{\sigma}_{yx}' \boldsymbol{\Sigma}_{xx}^{-1} \boldsymbol{\mu}_x, \\ \boldsymbol{\beta}_1 &= \boldsymbol{\Sigma}_{xx}^{-1} \boldsymbol{\sigma}_{yx}. \end{aligned}$$

Solution: By (4.33), we have that

$$E[y|\mathbf{x}] = \mu_y + \boldsymbol{\sigma}_{yx}' \boldsymbol{\Sigma}_{xx}^{-1} (\mathbf{x} - \boldsymbol{\mu}_x).$$

Now, notice that

$$\begin{aligned} t(\mathbf{x}) = \beta_0 + \boldsymbol{\beta}_1' \mathbf{x} &= \mu_y - \boldsymbol{\sigma}_{yx}' \boldsymbol{\Sigma}_{xx}^{-1} \boldsymbol{\mu}_x + \boldsymbol{\sigma}_{yx}' \boldsymbol{\Sigma}_{xx}^{-1} \mathbf{x} \\ &= \mu_y + \boldsymbol{\sigma}_{yx}' \boldsymbol{\Sigma}_{xx}^{-1} (\mathbf{x} - \boldsymbol{\mu}_x) \\ &= E[y|\mathbf{x}]. \end{aligned}$$

Then, the result follows from theorem 10.7.

Problem 8:

Prove Theorem 10.7c. The theorem states: If $(y_1, \mathbf{x}_1'), ..., (y_n, \mathbf{x}_n')$ is a random sample with mean vector and covariance matrix

$$\widehat{\boldsymbol{\mu}} = \begin{pmatrix} \overline{y} \\ \overline{\mathbf{x}} \end{pmatrix}, \qquad \mathbf{S} = \begin{pmatrix} s_{yy} & \mathbf{s}_{yx}' \\ \mathbf{s}_{yx} & \mathbf{S}_{xx} \end{pmatrix},$$

then the estimators $\widehat{\beta}_0$ and $\widehat{\boldsymbol{\beta}}_1$ that minimizes $\sum_{i=1}^n (y_i - \widehat{\beta}_0 - \widehat{\boldsymbol{\beta}}_1' \mathbf{x}_i)^2 / n$ are given by

$$\begin{aligned} \widehat{\beta}_0 &= \overline{y} - \mathbf{s}_{yx}' \mathbf{S}_{xx}^{-1} \overline{\mathbf{x}}, \\ \widehat{\boldsymbol{\beta}}_1 &= \mathbf{S}_{xx}^{-1} \mathbf{s}_{yx}. \end{aligned}$$

Solution: First, we calculate

$$\frac{d}{d\widehat{\boldsymbol{\beta}}_0'}\sum_{i=1}^{n}(y_i - \widehat{\beta}_0 - \widehat{\boldsymbol{\beta}}_1'\mathbf{x}_i)^2/n = -2\sum_{i=1}^{n}(y_i - \widehat{\beta}_0 - \widehat{\boldsymbol{\beta}}_1'\mathbf{x}_i)/n$$
$$= -2\overline{y} + 2\widehat{\beta}_0 + 2\widehat{\boldsymbol{\beta}}_1'\overline{\mathbf{x}} \overset{\text{set}}{=} 0$$

which gives that $\widehat{\beta}_0 = \overline{y} - \widehat{\boldsymbol{\beta}}_1'\overline{\mathbf{x}}$. Now, we calculate the second partial derivative to be

$$\frac{d}{d\widehat{\boldsymbol{\beta}}_1'}\sum_{i=1}^{n}(y_i - \widehat{\beta}_0 - \widehat{\boldsymbol{\beta}}_1'\mathbf{x}_i)^2/n = -\frac{2}{n}\sum_{i=1}^{n}(y_i - \widehat{\beta}_0 - \widehat{\boldsymbol{\beta}}_1'\mathbf{x}_i)\mathbf{x}_i'$$
$$= -\frac{2}{n}\sum_{i=1}^{n}\left[y_i\mathbf{x}_i' - \widehat{\beta}_0\mathbf{x}_i' - \widehat{\boldsymbol{\beta}}_1'\mathbf{x}_i\mathbf{x}_i'\right].$$

Setting this equal to the vector $\mathbf{0}$, we find that

$$-\frac{2}{n}\sum_{i=1}^{n}\left[y_i\mathbf{x}_i' - \widehat{\beta}_0\mathbf{x}_i' - \widehat{\boldsymbol{\beta}}_1'\mathbf{x}_i\mathbf{x}_i'\right] = \mathbf{0}$$
$$\implies \sum_{i=1}^{n}y_i\mathbf{x}_i' - \sum_{i=1}^{n}\widehat{\beta}_0\mathbf{x}_i' - \sum_{i=1}^{n}\widehat{\boldsymbol{\beta}}_1'\mathbf{x}_i\mathbf{x}_i' = \mathbf{0}$$
$$\implies \sum_{i=1}^{n}y_i\mathbf{x}_i' - n\overline{y}\overline{\mathbf{x}}' + n\widehat{\boldsymbol{\beta}}_1'\overline{\mathbf{x}}\overline{\mathbf{x}}' - \sum_{i=1}^{n}\widehat{\boldsymbol{\beta}}_1'\mathbf{x}_i\mathbf{x}_i' = \mathbf{0}$$
$$\implies \widehat{\boldsymbol{\beta}}_1' = \left(\sum_{i=1}^{n}y_i\mathbf{x}_i' - n\overline{y}\overline{\mathbf{x}}'\right)\left(\sum_{i=1}^{n}\mathbf{x}_i\mathbf{x}_i' - n\overline{\mathbf{x}}\overline{\mathbf{x}}'\right)^{-1}$$
$$= \mathbf{s}_{yx}'\mathbf{S}_{xx}^{-1}.$$

Therefore, we have that

$$\widehat{\boldsymbol{\beta}}_1 = \mathbf{S}_{xx}^{-1}\mathbf{s}_{yx}.$$

Plugging back into the expression for $\widehat{\beta}_0$ above finishes the result.