SOLUTIONS FOR APR. 7 HOMEWORK

1. Start with the sum of squared residuals: $SSR = \sum_{i=1}^{n} r_i^2 = \sum_{i=1}^{n} \left[y_i - (b_0 + b_1 x_i)\right]^2$

2. Take partial derivatives of SSR with respect to both b_0 and b_1, set these derivatives equal to zero, and show that this leads to the normal equations:

$$\begin{aligned} \sum_{i=1}^{n} y_i &= b_0 n + b_1 \sum_{i=1}^{n} x_i \\ \sum_{i=1}^{n} x_i y_i &= b_0 \sum_{i=1}^{n} x_i + b_1 \sum_{i=1}^{n} x_i^2 \end{aligned}$$

Proof: First, take a partial derivative with respect to b_0*:*

$$\begin{aligned} \frac{\partial}{\partial b_0} SSR &= \sum_{i=1}^{n} \frac{\partial}{\partial b_0} \left[y_i - (b_0 + b_1 x_i)\right]^2 \\ &= \sum_{i=1}^{n} -2\left[y_i - (b_0 + b_1 x_i)\right] \\ &= -2\sum_{i=1}^{n} y_i + 2\sum_{i=1}^{n} (b_0 + b_1 x_i). \end{aligned}$$

Setting this equal to zero (and dividing by 2) gives the first normal equation.

Next, take a partial derivative with respect to b_1*:*

$$\begin{aligned} \frac{\partial}{\partial b_1} SSR &= \sum_{i=1}^{n} \frac{\partial}{\partial b_1} \left[y_i - (b_0 + b_1 x_i)\right]^2 \\ &= \sum_{i=1}^{n} -2x_i\left[y_i - (b_0 + b_1 x_i)\right] \\ &= -2\sum_{i=1}^{n} x_i y_i + 2\sum_{i=1}^{n} (b_0 x_i + b_1 x_i^2). \end{aligned}$$

Setting this equal to zero (and dividing by 2) gives the second normal equation.

3. Solve the normal equations for b_1 by multiplying the first one by $-\sum_{i=1}^{n} x_i$ and the second one by n and then adding. Prove that

$$b_1 = \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}$$

Proof: First, we multiply the first normal equation by $-\sum_{i=1}^n x_i$ *and the second one by* n*:*

$$\begin{aligned} -\sum_{i=1}^n x_i \sum_{i=1}^n y_i &= -b_0 n \sum_{i=1}^n x_i - b_1 \left(\sum_{i=1}^n x_i\right)^2 \\ n\sum_{i=1}^n x_i y_i &= b_0 n \sum_{i=1}^n x_i + b_1 n \sum_{i=1}^n x_i^2. \end{aligned}$$

Next, add the two above equations, noting that the terms involving b_0 *cancel out:*

$$\begin{aligned} n\sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i &= b_1 n \sum_{i=1}^n x_i^2 - b_1 \left(\sum_{i=1}^n x_i\right)^2 \\ &= b_1 \left[n\sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i\right)^2 \right]. \end{aligned}$$

Solving for b_1 *is now a straightforward matter of dividing both sides by* $\left[n\sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i\right)^2\right]$.

4. Once b_1 is known, show that solving the normal equations for b_0 yields

$$b_0 = \overline{y} - b_1 \overline{x}$$

(Note: $\overline{y} = \frac{1}{n}\sum_{i=1}^n y_i$ and $\overline{x} = \frac{1}{n}\sum_{i=1}^n x_i$.)

Proof: Ignoring the second normal equation, start by dividing the first normal equation by n*:*

$$\frac{1}{n}\sum_{i=1}^n y_i = b_0 + \frac{b_1}{n}\sum_{i=1}^n x_i.$$

Rearranging this equation, and noting that $\overline{y} = \frac{1}{n}\sum_{i=1}^n y_i$ *and* $\overline{x} = \frac{1}{n}\sum_{i=1}^n x_i$*, we obtain*

$$b_0 = \overline{y} - b_1 \overline{x}$$

as desired.

5. Prove that these two equations are valid:

$$\begin{aligned} \sum_{i=1}^n x_i y_i - \frac{1}{n}\sum_{i=1}^n x_i \sum_{i=1}^n y_i &= \sum_{i=1}^n (x_i - \overline{x})(y_i - \overline{y}) \\ \sum_{i=1}^n x_i^2 - \frac{1}{n}\left(\sum_{i=1}^n x_i\right)^2 &= \sum_{i=1}^n (x_i - \overline{x})^2 \end{aligned}$$

Proof: Begin by multiplying

$$\begin{aligned}
\sum_{i=1}^{n}(x_i-\overline{x})(y_i-\overline{y}) &= \sum_{i=1}^{n}(x_iy_i-\overline{x}y_i-\overline{y}x_i+\overline{xy}) \\
&= \sum_{i=1}^{n}x_iy_i-\overline{x}\sum_{i=1}^{n}y_i-\overline{y}\sum_{i=1}^{n}x_i+\overline{xy}\sum_{i=1}^{n}1 \\
&= \sum_{i=1}^{n}x_iy_i-\frac{1}{n}\sum_{i=1}^{n}x_i\sum_{i=1}^{n}y_i-\frac{1}{n}\sum_{i=1}^{n}x_i\sum_{i=1}^{n}y_i+\frac{1}{n}\sum_{i=1}^{n}x_i\sum_{i=1}^{n}y_i \\
&= \sum_{i=1}^{n}x_iy_i-\frac{1}{n}\sum_{i=1}^{n}x_i\sum_{i=1}^{n}y_i.
\end{aligned}$$

This is the first equation that was to be proven. To prove the second equation, simply note that it follows immediately from the first equation if we let $y_i = x_i$.

6. Use the above equations to show that the formula for b_1 on page 144 is valid.

   Proof: Starting with the equation for b_1 *in part 3, divide the numerator and denominator by* n *to obtain*

   $$b_1 = \frac{\sum_{i=1}^{n}x_iy_i-\frac{1}{n}\sum_{i=1}^{n}x_i\sum_{i=1}^{n}y_i}{\sum_{i=1}^{n}x_i^2-\frac{1}{n}\left(\sum_{i=1}^{n}x_i\right)^2}.$$

   Next, both the numerator and denominator may be rewritten using the equations in part 3:

   $$b_1 = \frac{\sum_{i=1}^{n}(x_i-\overline{x})(y_i-\overline{y})}{\sum_{i=1}^{n}(x_i-\overline{x})^2}.$$

7. Derive one more way to write b_1: If r is defined as on page 146, prove that

   $$b_1 = r\frac{s_y}{s_x}.$$

   (Hint: Write out the formula for s_x^2, then show $b_1 = rs_xs_y/s_x^2$.)

   Proof: Recall that

   $$s_x^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i-\overline{x})^2.$$

   Furthermore, from page 146 it is clear that

   $$rs_xs_y = \frac{1}{n-1}\sum_{i=1}^{n}(x_i-\overline{x})(y_i-\overline{y}).$$

Therefore, if we divide rs_xs_y *by* s_x^2*, the* $n-1$ *cancels from the previous two expressions and we are left with*

$$\frac{rs_xs_y}{s_x^2} \quad = \quad \frac{\sum_{i=1}^{n}(x_i-\overline{x})(y_i-\overline{y})}{\sum_{i=1}^{n}(x_i-\overline{x})^2},$$

which is identical to the formula for b_1 *in part 6 (from page 144). We conclude that*

$$b_1 = \frac{rs_xs_y}{s_x^2} = r\frac{s_y}{s_x}.$$