STAT 563 HW #5

Maggie Buffum

May 23, 2019

Problem 12.2

Show that the extrema of

$$f(b) = \frac{1}{1+b^2}[S_{yy} - 2bS_{xy} + b^2 S_{xx}]$$

are give by

$$b = \frac{-(S_{xx} - S_{yy}) \pm \sqrt{(S_{xx} - S_{yy})^2 + 4S_{xy}^2}}{2S_{xy}}.$$

Show that the "+" solution gives the minimum of $f(b)$.

(First part in lecture notes from May 14. Second part you plug the + version of $\hat{\beta}$ into the second derivative of f(b) and show that it's positive)

First we need to find the derivative of $f(b)$ and set it equal to zero, solving for b:

$$\begin{aligned}
\frac{df(b)}{db} =& \frac{d}{db}\left(\frac{S_{yy} + b^2 S_{xx} - 2bS_{xy}}{1+b^2}\right) \\
=& \frac{(1+b^2)[2bS_{xx} - 2S_x y] - [S_{yy} + b^2 S_{xx} - 2bS_{xy}](2b)}{(1+b^2)^2} = \text{ (set) } 0 \\
&\implies bS_{xx} - S_{xy} + b^3 S_{xx} - b^2 S_{xy} - bS_{yy} - b^3 S_{xx} + 2b^2 S_{xy} = 0 \\
&\implies b^2 S_{xy} + b(S_{xx} - S_{yy}) - S_{xy} = 0 \\
&\implies \hat{b} = \frac{S_{yy} - S_{xx} \pm \sqrt{(S_{xx} - S_{yy})^2 + 4S_{xy}^2}}{2S_{xy}} \\
\hat{b}_1 =& \frac{S_{yy} - S_{xx} + \sqrt{(S_{xx} - S_{yy})^2 + 4S_{xy}^2}}{2S_{xy}} \\
\hat{b}_2 =& \frac{S_{yy} - S_{xx} - \sqrt{(S_{xx} - S_{yy})^2 + 4S_{xy}^2}}{2S_{xy}}
\end{aligned}$$

To confirm we found a minimum, we must take the second derivative of $f(b)$ and insert $\hat{b}_1$. If it is positive, then $\hat{b}_1$ gives the minimum solution. The second derivative of $f(b)$ is

$$\begin{aligned}
\frac{d^2 f(b)}{db^2} =& \frac{d}{db}\left(\frac{b^2 S_{xy} + b(S_{xx} - S_{yy}) - S_{xy}}{(1+b^2)^2}\right) \\
=& \frac{(1+b^2)^2(2bS_{xy} + (S_{xx} - S_{yy})) - 2(1+b^2)(2b)(b^2 S_{xy} + b(S_{xx} - S_{yy}) - S_{xy})}{(1+b^2)^4} \\
=& \frac{(1+b^2)(2bS_{xy} + S_{xx} - S_{yy}) - 4b(b^2 S_{xy} + b(S_{xx} - S_{yy}) - S_{xy})}{(1+b^2)^3} \\
=& \frac{1}{(1+b^2)^3}\left(2bS_{xy} + 2b^3 S_{xy} + (1+b^2)(S_{xx} - S_{yy}) - 4b^3 S_{xy} - 4b^2(S_{xx} - S_{yy}) + 4bS_{xy}\right) \\
=& \frac{1}{(1+b^2)^3}\left(6bS_{xy} + (S_{xx} - S_{yy}) - 2b^3 S_{xy} - 3b^2(S_{xx} - S_{yy}))\right)
\end{aligned}$$

The negative terms will be larger in value than the positive terms, so the second derivative is negative. Therefore, $\hat{b}_1$ gives the minimum solution.

Problem 12.4

Consider the MLE of the slope in the EIV model

$$\hat{\beta}(\lambda) = \frac{-(S_{xx} - \lambda S_{yy}) + \sqrt{(S_{xx} - \lambda S_{yy})^2 + 4\lambda S_{xy}^2}}{2\lambda S_{xy}}$$

where $\lambda = \sigma_\delta^2/\sigma_\epsilon^2$ is assumed known.

(a) Show that $\lim_{\lambda\to 0} \hat{\beta}(\lambda) = S_{xy}/S_{xx}$, the slope of the ordinary regression of y on x.

$$\begin{aligned}
\lim_{\lambda\to 0} \hat{\beta}(\lambda) &= \lim_{\lambda\to 0} \left(\frac{-(S_{xx} - \lambda S_{yy}) + \sqrt{(S_{xx} - \lambda S_{yy})^2 + 4\lambda S_{xy}^2}}{2\lambda S_{xy}} \right) \\
&= \lim_{\lambda\to 0} \left(\frac{-(S_{xx} - \lambda S_{yy}) + \sqrt{(S_{xx} - \lambda S_{yy})^2 + 4\lambda S_{xy}^2}}{2\lambda S_{xy}} \frac{(S_{xx} - \lambda S_{yy}) + \sqrt{(S_{xx} - \lambda S_{yy})^2 + 4\lambda S_{xy}^2}}{(S_{xx} - \lambda S_{yy}) + \sqrt{(S_{xx} - \lambda S_{yy})^2 + 4\lambda S_{xy}^2}} \right) \\
&= \lim_{\lambda\to 0} \left(\frac{-(S_{xx} - \lambda S_{yy})^2 + (S_{xx} - \lambda S_{yy})^2 + 4\lambda S_{xy}^2}{2\lambda S_{xy}\left((S_{xx} - \lambda S_{yy}) + \sqrt{(S_{xx} - \lambda S_{yy})^2 + 4\lambda S_{xy}^2}\right)} \right) \\
&= \lim_{\lambda\to 0} \left(\frac{4\lambda S_{xy}^2}{2\lambda S_{xy}\left((S_{xx} - \lambda S_{yy}) + \sqrt{(S_{xx} - \lambda S_{yy})^2 + 4\lambda S_{xy}^2}\right)} \right) \\
&= \lim_{\lambda\to 0} \left(\frac{2S_{xy}}{(S_{xx} - \lambda S_{yy}) + \sqrt{(S_{xx} - \lambda S_{yy})^2 + 4\lambda S_{xy}^2}} \right) \\
&= \frac{2S_{xy}}{2S_{xx}} \\
&= \frac{S_{xy}}{S_{xx}}
\end{aligned}$$

(b) Show that $\lim_{\lambda\to\infty} \hat{\beta}(\lambda) = S_{yy}/S_{xy}$, the reciprocal of the slope of the ordinary regression of x on y.

$$
\begin{aligned}
\lim_{\lambda\to\infty}\hat{\beta}(\lambda) &= \lim_{\lambda\to\infty}\left(\frac{-(S_{xx}-\lambda S_{yy})+\sqrt{(S_{xx}-\lambda S_{yy})^2+4\lambda S_{xy}^2}}{2\lambda S_{xy}}\right) \\
&= \lim_{\lambda\to\infty}\left(\frac{-(\frac{S_{xx}}{\lambda}-S_{yy})+\sqrt{\frac{1}{\lambda^2}(S_{xx}-\lambda S_{yy})^2+\frac{1}{\lambda^2}4\lambda S_{xy}^2}}{2S_{xy}}\right) \\
&= \lim_{\lambda\to\infty}\left(\frac{-(\frac{S_{xx}}{\lambda}-S_{yy})+\sqrt{(\frac{S_{xx}}{\lambda}-S_{yy})^2+\frac{4S_{xy}^2}{\lambda}}}{2S_{xy}}\right) \\
&=\frac{2S_{yy}}{2S_{xy}} \\
&=\frac{S_{yy}}{S_{xy}}
\end{aligned}
$$

(c) Show that $\hat{\beta}(\lambda)$ is, in fact, monotone in λ and is increasing if $S_{xy} > 0$.

$$
\begin{aligned}
\frac{d}{d\lambda}\hat{\beta}(\lambda) &= \frac{d}{d\lambda}\frac{-(S_{xx}-\lambda S_{yy})+\sqrt{(S_{xx}-\lambda S_{yy})^2+4\lambda S_{xy}^2}}{2\lambda S_{xy}} \\
&= \frac{\frac{d}{d\lambda}\left(-(S_{xx}-\lambda S_{yy})+\sqrt{(S_{xx}-\lambda S_{yy})^2+4\lambda S_{xy}^2}\right)}{\frac{d}{d\lambda}(2\lambda S_{xy})} \\
&= \frac{S_{yy}+\sqrt{4S_{xy}^2-2S_{xx}S_{yy}}}{2S_{xy}}
\end{aligned}
$$

This function is always postive when S_{xy} is positive and negative when S_{xy} is negative, and therefore is always monotone in λ.

(d) Show that the orthogonal least squares line ($\lambda = 1$) is always between the lines given by the ordinary regressions of y on x and of x on y.

From part (c) we know that $\hat{\beta}(\lambda)$ is monotone in λ. From part (a) we have that $\hat{\beta}(\lambda) \to \frac{S_{xy}}{S_{xx}}$ as $\lambda \to 0$ for y on x, and from part (b) $\hat{\beta}(\lambda) \to \frac{S_{yy}}{S_{xy}}$ as $\lambda \to \infty$ for x on y. The limits of $\hat{\beta}(\lambda)$ to either side will be higher or lower than anything in between, such as where $\lambda = 1$, because $\hat{\beta}(\lambda)$ is monotone in λ. Thus we know the EIV slope will always fall in between the ordinary regressions of y on x and of x on y.

(e) The following data were collected in a study to examine the relationship between brain weight and body weight in a number of animal species.

Bod	y Weight (kg) Bod	y Weight (kg)
Arctic fox	3.385	44.5
Owl monkey	0.480	15.5
Mountain beaver	1.350	8.1
Guinea pig	1.040	5.5

Bod	y Weight (kg) Bod	y Weight (kg)
Chinchilla	0.425	6.4
Ground squirrel	0.101	4.0
Tree hyrax	2.000	12.3
Big brown bat	0.023	0.3

Calculate the MLE of the slope assuming the EIV model. Also, calculate the least squares slopes of the regressions of y on x and of x on y, and show how these quantities bound the MLE.

```r
colnames(tab) <- c("X","Y")
df <- data.frame(tab) %>%
  mutate(Sxx_i = (X - mean(X)),
         Syy_i = (Y - mean(Y)),
         Sxy_i = Sxx_i*Syy_i)

Sxx <- sum(df$Sxx_i^2)
Syy <- sum(df$Syy_i^2)
Sxy <- sum(df$Sxy_i)

EIV_slope <- (-(Sxx - Syy) + sqrt((Sxx - Syy)^2 + 4*Sxy^2))/(2*Sxy)

XonY_slope <- Syy/Sxy
YonX_slope <- Sxy/Sxx

slopes <- matrix(c(
  EIV_slope,XonY_slope,YonX_slope
),nrow = 3,dimnames = list(
  c("EIV","X on Y","Y on X"),
  "Slope"))
kable(slopes)
```

	Slope
EIV	14.10272
X on Y	14.12640
Y on X	10.57143