Eric Doviak

Metrics 7020X – Homework #1c

Problem #1 – The following questions ask you to derive the least squares estimates of regression coefficients. Consider the following regression model:

$$y_i = \alpha + \beta x_i + \varepsilon_i$$

where ε_i is the error term. Rearranging terms, squaring both sides and summing over observations yields the sum of squared errors:

$$\sum \varepsilon_i^2 = \sum (y_i - \alpha - \beta x_i)^2$$

The estimates of regression coefficients are the values of α and β that minimize the sum of squared errors.

1. Obtain the first-order conditions for a minimum by taking the partial derivative of $\sum \varepsilon_i^2$
   (a) with respect to α
   (b) with respect to β

Define the mean of x as: $\bar{x} \equiv \frac{1}{N} \sum x_i$ and define the mean of y as: $\bar{y} \equiv \frac{1}{N} \sum y_i$.

2. Show that the first-order conditions imply that: $\widehat{\alpha} = \bar{y} - \widehat{\beta}\bar{x}$.
3. Show that the first-order conditions imply that: $\widehat{\beta} = \dfrac{\frac{1}{N}\sum x_i y_i - \bar{x}\bar{y}}{\frac{1}{N}\sum x_i^2 - \bar{x}^2}$.
4. Rearrange terms to show that: $\widehat{\beta} = \dfrac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2}$.

Next, you need to show that the second-order conditions for a minimum are satisfied.

5. Take the second partial derivative of $\sum \varepsilon_i^2$
   (a) with respect to α
   (b) with respect to β
   (c) with respect to α and β (i.e. the "cross-partial")
6. Show that the second partial with respect to α is positive.
7. Show that the second partial with respect to β is positive.
8. Set up the Hessian matrix and show that its determinant is positive. (**Hint:** If you rearrange terms, you'll see that the determinant is a function of the variance of x).

♦ ♦ ♦

Problem #2 – The following questions ask you to derive the maximum likelihood estimates of the mean and variance. Assuming that x is distributed normally with mean μ and variance σ^2, the likelihood function can be written as the product of the probability densities for each observation:

$$\pounds\left(\mu,\sigma^2\right)=\prod\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{\left(x_i-\mu\right)^2}{2\sigma^2}\right)$$

1. Take the natural logarithm of the likelihood function to obtain the log-likelihood function.
2. Obtain the first-order conditions for a maximum by taking the partial derivative of the log-likelihood function:
   (a) with respect to μ
   (b) with respect to σ^2
3. Show that the first-order conditions imply that: $\widehat{\mu}=\frac{1}{N}\sum x_i$.
4. Show that the first-order conditions imply that: $\widehat{\sigma}^2=\frac{1}{N}\sum\left(x_i-\mu\right)^2$.

Next, you need to show that the second-order conditions for a maximum are satisfied.

5. Take the second partial derivative of the log-likelihood function:
   (a) with respect to μ
   (b) with respect to σ^2
   (c) with respect to μ and σ^2 (i.e. the "cross-partial")
6. Show that the second partial with respect to μ is negative.
7. Show that the second partial with respect to σ^2 is negative.
8. Set up the Hessian matrix and show that its determinant is positive. (**Hint:** Make use of the fact that $\widehat{\mu}=\frac{1}{N}\sum x_i$ when the first-order conditions are satisfied).

♦ ♦ ♦

Problem #3 – The tasks below require you use maximum likelihood to estimate regression coefficients, the standard error of the regression and the standard error of the estimates. Specifically, assume that the error term in the following regression model is distributed normally with mean zero and constant variance:

$$\varepsilon_i=y_i-\alpha-\beta x_i$$
$$\varepsilon\sim N\left(0,\sigma^2\right)$$

Your assignment is to maximize the likelihood function:

$$\pounds\left(\alpha,\beta,\sigma^2\right)=\prod\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{\left(y_i-\alpha-\beta x_i\right)^2}{2\sigma^2}\right)$$

and use the first-order conditions to estimate $\widehat{\alpha}$, $\widehat{\beta}$ and $\widehat{\sigma^2}$. Then you must use the second-order conditions to demonstrate that the likelihood function has been maximized (and is not at a saddle point or minimum). Finally, you must estimate the standard error of your estimates.

1. Take the natural logarithm of the likelihood function to obtain the log-likelihood function. Then derive the first-order conditions for a minimum of the log-likelihood function with respect to $\widehat{\alpha}$, $\widehat{\beta}$ and $\widehat{\sigma^2}$.

2. Define: $\bar{x} \equiv \frac{1}{N}\sum x_i$ and: $\bar{y} \equiv \frac{1}{N}\sum y_i$. Show that the first-order conditions imply that:

$$\begin{aligned}
\widehat{\alpha} &= \bar{y} - \widehat{\beta}\bar{x} \\
\widehat{\beta} &= \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2} \\
\widehat{\sigma^2} &= \frac{1}{N}\sum \left(y_i - \widehat{\alpha} - \widehat{\beta}x_i\right)^2
\end{aligned}$$

3. Next, you need to show that the second-order conditions for a maximum are satisfied.
   - Set up the Hessian matrix of second partials at $\widehat{\alpha}$, $\widehat{\beta}$ and $\widehat{\sigma^2}$.
   - Show that the own-partials are negative at $\widehat{\alpha}$, $\widehat{\beta}$ and $\widehat{\sigma^2}$.
   - Show that the determinant of the Hessian is negative at $\widehat{\alpha}$, $\widehat{\beta}$ and $\widehat{\sigma^2}$.

4. Define the information matrix as the negative of the inverse of the Hessian matrix at $\widehat{\alpha}$, $\widehat{\beta}$ and $\widehat{\sigma^2}$. Show that the square roots of the information matrix's diagonal elements imply that:

$$\begin{aligned}
std.\ error\ \widehat{\alpha} &= \widehat{\sigma}\sqrt{\frac{\frac{1}{n}\sum x_i^2}{\sum (x_i - \bar{x})^2}} \\
std.\ error\ \widehat{\beta} &= \widehat{\sigma}\sqrt{\frac{1}{\sum (x_i - \bar{x})^2}} \\
std.\ error\ \widehat{\sigma^2} &= \widehat{\sigma^2}\sqrt{\frac{2}{N}}
\end{aligned}$$

5. Use what you have learned about optimization to explain why our estimate of a parameter's standard error is smaller when the log-likelihood surface comes to a sharp peak along its dimension.