ECOM40006/90013 Econometrics 3

Week 10 Extras (Part 3)

Question 1: The Story So Far (Normal Distribution Revisited)

Note. *These extras cover some of the results that were stated without intermediate calculation in the Week 10 tutorial. These go as far as deriving up to the information matrix. Test statistics are not covered. This is purely for those who want to try their hand at calculating what is on the tutorial handout. The notation is a bit different from usual, but you should be able to draw parallels based off the techniques shown.*

For those who want to see some additional information on the Week 10 tutorial handout: consider a sequence of random variables $Y_1, Y_2, \dots, Y_n$ that is generated according to the data generating process

$$Y_i = X_i'\beta_0 + U_i, \qquad U_i \sim \text{i.i.d. } N(0, \sigma_0^2)$$

implying that the conditional distribution of Y_i given X_i is (can you show this?)

$$Y_i | X_i \sim N(X_i'\beta_0, \sigma_0^2).$$

Suppose that we look at a sample $y_1, y_2, \dots, y_n$ from this population, and further suppose that we model the parameters β (a $k \times 1$ column vector) and σ^2 (a 1×1 scalar) assuming that conditional on the data x_i, β and σ, the probability density of y_i is normal in the sense that

$$f(y_i; x_i, \beta, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(y_i - x_i'\beta)^2}{2\sigma^2} \right).$$

(a) Derive the log-likelihood

$$\ell_n(\beta, \sigma^2) = \log L_n(\beta, \sigma^2) = \sum_{i=1}^n \log f(y_i; x_i, \beta, \sigma^2).$$

(b) Derive the score

$$s_n(\beta, \sigma^2) = \frac{\partial \log L_n(\beta, \sigma^2)}{\partial(\beta, \sigma^2)} = \begin{bmatrix} \dfrac{\partial \log L_n(\beta, \sigma^2)}{\partial \beta} \\ \dfrac{\partial \log L_n(\beta, \sigma^2)}{\partial \sigma^2} \end{bmatrix}.$$

Find values $\beta = \hat{\beta}$ and $\sigma^2 = \hat{\sigma}^2$ such that the score equals zero. That is: find the maximum likelihood estimators. Note that to evaluate this you'll need the trusty result that we've carried with us for a few weeks:

$$\frac{\partial x_i'\beta}{\partial \beta} = x_i,$$

along with a few applications of the Chain Rule. If it helps with the notation, you can also use the shorthand $u_i = y_i - x_i'\beta$.

(c) Derive the Hessian $H_n(\beta, \sigma^2)$. The steps to do this are:

- Take the first element of $s_n(\beta, \sigma^2)$.
- Calculate the gradient vector of that first element.
- Transpose your answer.
- Repeat this for the second element of $s_n(\beta, \sigma^2)$.
- Create a matrix where
  - The first row is the transposed gradient vector from steps 1-3 earlier.
  - The second row is the vector from step 4.
- The result is the Hessian!

(d) Show that at the MLEs found in (b), the Hessian is negative definite, implying that the MLEs maximize the log likelihood.

- Since you're evaluating this at the MLEs, you can use the results from the first order conditions as necessary.
- This is helpful in the case of σ^2.

(e) Show that the expected score is equal to zero at the true parameters β_0 and σ_0^2.

- This one's a bit trickier because you'll have both (β, σ^2) and (β_0, σ_0^2) floating around.
- Just remember that the score is a function of these two parameters, with the feature that special things happen if you make these parameters equal to their true values.
- For the σ^2 part specifically, it may help if you substitute in $(\beta, \sigma^2) = (\beta_0, \sigma_0^2)$ early and make use of the fact that the formula for the conditional variance is

  $$\text{var}(Y_i|X_i) = \mathbb{E}[(Y_i - \mathbb{E}(Y_i|X_i))^2|X_i],$$

  which in this question is also equal to σ_0^2.