TU-Berlin

Microeconometrics

4^{th} Tutorial: Maximum Likelihood Estimation(continued), Variance Estimation

Tomas Repasky

WS 2016/17

Technische Universität Berlin
Institut für Volkswirtschaftslehre und Wirtschaftsrecht
FG Ökonometrie und Wirtschaftsstatistik

From the last Tutorial

MLE of π

Your task was to compute the Maximum Likelihood Estimator of $\pi := \mathbb{P}(Y_i = 1)$ of a bernoulli random variable Y_i and check the SOC making sure that you have indeed found the maximum.

The respective Log-Likelihood looks like:

$$\mathscr{L}(\pi, \mathbf{y}) = \sum_{i=1}^{N} \{y_i * log(\pi) + (1 - y_i) * log(1 - \pi)\} \tag{1}$$

The respective FOC:

$$\frac{\sum_{i=1}^{N} y_i - n\hat{\pi}}{\hat{\pi}(1 - \hat{\pi})} = 0 \tag{2}$$

From the Last Tutorial II

MLE of π

The corresponding Maximum is thus given by:

$$\frac{1}{N}\sum_{i=1}^{N} y_i = \hat{\pi} = \bar{y} \tag{3}$$

And finally the SOC:

$$\frac{\partial\left(\sum_{i=1}^{N} y_i \frac{1}{\pi} + (1-y_i)\frac{1}{1-\pi}(-1)\right)}{\partial\pi} = \tag{4}$$

$$\sum_{i=1}^{N} -\frac{y_i}{\hat{\pi}^2} - \frac{(1-y_i)}{(1-\hat{\pi})^2} < 0 \tag{5}$$

Today's Task: Plot the $\mathscr{L}$, FOC as well as SOC as a function of $\hat{\pi}$

- Look at the functions and define the parameters on which they depend.
- Start off with the $\mathscr{L}$. How would you compute the values of this function for different $\hat{\pi}$ and which parameters would you need to do so?
- This function is already programmed in the uploaded R-code.
- Review the code and do the same for the FOC and SOC as given by the maximization problem's solution.
- For the SOC, plot the function for three cases: the number of working spouses in the sample is 1.) 100, 2.) 428, 3.) 600.

- For $\mathscr{L}$, discuss the resulting graph. How can the graph be interpreted?
- Programm/produce the plot of the FOC($\hat{\pi}$) and explain the *intuition* behind the graph.
- Plot the SOC for all the three cases into one graph. What can you learn from the difference between those cases?

Repetition

You have learned in the lecture that the asymptotic variance of the MLE Estimator can be computed in the following ways:

$$\mathbf{A} = \left\{ -\mathbb{E}\left[\frac{\partial^2 \mathscr{L}(\boldsymbol{\beta})}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}'} \right] \right\}^{-1} = \{\mathbb{E}\left[\mathbf{I}(\boldsymbol{\beta})\right]\}^{-1}$$

Where the observed Information Matrix equals to minus the Hessian.

$$\mathbf{I} = -\sum_{i=1}^{n} \frac{\partial^2 \mathscr{L}(\boldsymbol{\theta}; \mathbf{y})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} = -\sum_{i=1}^{n} \mathbf{H}_i(\boldsymbol{\theta}; y_i)$$

What follows from that and how does it help you to explain the difference in the last task (comparing the differences between SOC given # of 1's is 100, 428 ,and 600)?

Variance Estimators in the Probit Model:

estimated Hessian matrix

$$\hat{Avar}_1(\hat{\beta}) = \left\{ \sum_{i=1}^{n} -\mathbf{H}_i(\hat{\beta}) \right\}^{-1} = \left\{ \sum_{i=1}^{n} \lambda_i * \left[\lambda_i + \mathbf{x}_i' \hat{\beta} \right] * \mathbf{x}_i \mathbf{x}_i' \right\}^{-1} \quad (6)$$

outer product of the score vector

$$\hat{Avar}_2(\hat{\beta}) = \left\{ \sum_{i=1}^{n} \mathbf{s}_i(\hat{\beta}_i) \mathbf{s}_i(\hat{\beta}_i)' \right\}^{-1} = \left\{ \sum_{i=1}^{n} \lambda_i^2 * \mathbf{x}_i \mathbf{x}_i' \right\}^{-1} \quad (7)$$

estimated conditional Hessian matrix

$$\hat{Avar}_3(\hat{\beta}) = \left\{ \sum_{i=1}^{n} \frac{\phi(\mathbf{x}_i' \hat{\beta})^2}{\Phi(\mathbf{x}_i' \hat{\beta}) * [1 - \Phi(\mathbf{x}_i' \hat{\beta})]} * \mathbf{x}_i \mathbf{x}_i' \right\}^{-1} \quad (8)$$

where:

$$\lambda_i = \frac{\phi(\mathbf{x}_i' \hat{\beta}) * [Y_i - \Phi(\mathbf{x}_i' \hat{\beta})]}{\Phi(\mathbf{x}_i' \hat{\beta}) * [1 - \Phi(\mathbf{x}_i' \hat{\beta})]} \quad (9)$$

- Begin by estimating the $\boldsymbol{\lambda}$, which will be a vector consisting of individual λ_i
- Why it is easier this way (in R)?
- Estimate it!

$\hat{Avar}_1(\hat{\beta})$

- It is easier to deal with Matrices in R.
- We need to express the $\hat{Avar}_1(\hat{\beta})$ in the Matrix form.
- *Hint: The easiest way would be to search for the same formula in the matrix form.*
- You could implement the formula exactly as it is using for-loops but that would be burdensome.
- We can use that $\sum_{i=1}^{n} \mathbf{x}_i \mathbf{x}_i' = \boldsymbol{X}'\boldsymbol{X}$ but we still have to include the term before the vectors that also depends on i in the Matrix.
- How can we do that?

$\hat{Avar}_1(\hat{\beta})$ in Matrix form I:

$\sum_{i=1}^{n} \mathbf{x}_i \mathbf{x}_i' =$

$$\sum_{i=1}^{n} \left\{ \begin{bmatrix} 1 \\ x_{1i} \\ x_{2i} \\ x_{3i} \end{bmatrix} * \begin{bmatrix} 1 & x_{1i} & x_{2i} & x_{3i} \end{bmatrix} \right\} = \sum_{i=1}^{n} \left\{ \begin{bmatrix} 1 & x_{1i} & x_{2i} & x_{3i} \\ x_{1i} & x_{1i}^2 & x_{1i}x_{2i} & x_{1i}x_{3i} \\ x_{2i} & x_{2i}x_{1i} & x_{2i}^2 & x_{2i}x_{3i} \\ x_{3i} & x_{3i}x_{1i} & x_{3i}x_{2i} & x_{3i}^2 \end{bmatrix} \right\}$$

$$= \begin{bmatrix} \sum_{i=1}^{n} 1 & \sum_{i=1}^{n} x_{1i} & \sum_{i=1}^{n} x_{2i} & \sum_{i=1}^{n} x_{3i} \\ \sum_{i=1}^{n} x_{1i} & \sum_{i=1}^{n} x_{1i}^2 & \sum_{i=1}^{n} x_{1i}x_{2i} & \sum_{i=1}^{n} x_{1i}x_{3i} \\ \sum_{i=1}^{n} x_{2i} & \sum_{i=1}^{n} x_{2i}x_{1i} & \sum_{i=1}^{n} x_{2i}^2 & \sum_{i=1}^{n} x_{2i}x_{3i} \\ \sum_{i=1}^{n} x_{3i} & \sum_{i=1}^{n} x_{3i}x_{1i} & \sum_{i=1}^{n} x_{3i}x_{2i} & \sum_{i=1}^{n} x_{3i}^2 \end{bmatrix}$$

$\hat{Avar}_1(\hat{\beta})$ in Matrix form I:

Define $c_i := \lambda_i * \left[\lambda_i + \mathbf{x}_i'\hat{\beta}\right]$ and note that c_i is a scalar.

$\sum_{i=1}^{n} c_i * \mathbf{x}_i\mathbf{x}_i'$=

$$= \begin{bmatrix} \sum_{i=1}^{n} c_i & \sum_{i=1}^{n} c_i x_{1i} & \sum_{i=1}^{n} c_i x_{2i} & \sum_{i=1}^{n} c_i x_{3i} \\ \sum_{i=1}^{n} c_i x_{1i} & \sum_{i=1}^{n} c_i x_{1i}^2 & \sum_{i=1}^{n} c_i x_{1i} x_{2i} & \sum_{i=1}^{n} c_i x_{1i} x_{3i} \\ \sum_{i=1}^{n} c_i x_{2i} & \sum_{i=1}^{n} c_i x_{2i} x_{1i} & \sum_{i=1}^{n} c_i x_{2i}^2 & \sum_{i=1}^{n} c_i x_{2i} x_{3i} \\ \sum_{i=1}^{n} c_i x_{3i} & \sum_{i=1}^{n} c_i x_{3i} x_{1i} & \sum_{i=1}^{n} c_i x_{3i} x_{2i} & \sum_{i=1}^{n} c_i x_{3i}^2 \end{bmatrix}$$

$\hat{Avar}_1(\hat{\beta})$ in Matrix form II:

Matrices $\mathbf{X}$ and $\mathbf{X'X}$:

$$\mathbf{X'X} = \begin{bmatrix} 1 & 1 & 1 & \dots & 1 \\ x_{11} & x_{12} & x_{13} & \dots & x_{1n} \\ x_{21} & x_{22} & x_{23} & \dots & x_{2n} \\ x_{31} & x_{32} & x_{33} & \dots & x_{3n} \end{bmatrix} \text{x} \begin{bmatrix} 1 & x_{11} & x_{21} & x_{31} \\ 1 & x_{12} & x_{22} & x_{32} \\ 1 & x_{13} & x_{23} & x_{33} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_{1n} & x_{2n} & x_{3n} \end{bmatrix}$$

$$= \begin{bmatrix} \sum_{i=1}^n 1 & \sum_{i=1}^n x_{1i} & \sum_{i=1}^n x_{2i} & \sum_{i=1}^n x_{3i} \\ \sum_{i=1}^n x_{1i} & \sum_{i=1}^n x_{1i}^2 & \sum_{i=1}^n x_{1i}x_{2i} & \sum_{i=1}^n x_{1i}x_{3i} \\ \sum_{i=1}^n x_{2i} & \sum_{i=1}^n x_{2i}x_{1i} & \sum_{i=1}^n x_{2i}^2 & \sum_{i=1}^n x_{2i}x_{3i} \\ \sum_{i=1}^n x_{3i} & \sum_{i=1}^n x_{3i}x_{1i} & \sum_{i=1}^n x_{3i}x_{2i} & \sum_{i=1}^n x_{3i}^2 \end{bmatrix}$$

$A\hat{var}_1(\hat{\beta})$ in Matrix form III:

But we want to multiply every $\sum_{i=1}^{n} x_{mi}x_{ki}$ by the respective c_i!

- We thus multiply every row of the matrix $\mathbf{X}$ by the respective $\sqrt{c_i}$. Note that every row (except for the 1^{st} one) contains all x-values for the respective observation i.
- How we can get there? Premultiply the Matrix X with a diagonal Matrix with the vector $\boldsymbol{c}^{-\frac{1}{2}}$ on the maindiagonal.