

Maximum Likelihood – Part I: Lecture XXVII

Charles B. Moss ¹

¹University of Florida

March 20, 2012

1 Readings

2 Supply Systems

- Dual Envelopes and Flexible Approximations
- Normalized Quadratic

3 Normalized Quadratic Likelihood Function

- Single Equation
- Single Equation First and Second Order Conditions
- Newton - Raphson
- Multiple Equations

4 Readings

Readings

- * Cramer, J.S. 1986. *Econometric Applications of Maximum Likelihood Methods*. Cambridge University Press
- * Greene, W.H. 2012. *Econometric Analysis* Seventh Edition. (Chapter 14 pp. 509 598)

Supply Systems

- The shameless plug is that my other course (AEB 6184) develops detailed empirical models of firm behavior. One standard approach is the Shephard dual.
 - In the dual cost approach, we hypothesize a dual cost function

$$C(w, y) = \min_x w'x \quad \exists: f(y, x) = 0 \quad (1)$$

where $C(\cdot)$ is the cost function, w is the vector of input prices, y is the vector of output levels, x is the vector of input levels, and $f(y, x) = 0$ is the production function (specified as an implicit production function).

- Shephard's lemma then demonstrates that

$$\frac{\partial C(w, y)}{\partial w_i} = x_i^*(w, y) \quad (2)$$

where $x_i^*(w, y)$ is the optimal level of input use conditional on w and y .

Dual Envelopes and Flexible Approximations

- Given this theoretical result, we then specify the cost function using a flexible approximation to the cost function. For example, for a two output and three input case we can use a quadratic approximation

$$C(w, y) = \alpha_0 + \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix}' \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} + \frac{1}{2} \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix}' \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{12} & A_{22} & A_{23} \\ A_{13} & A_{23} & A_{33} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} +$$

$$\begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}' \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} + \frac{1}{2} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}' \begin{bmatrix} B_{11} & B_{12} \\ B_{12} & B_{22} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} +$$

$$\begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix}' \begin{bmatrix} \gamma_{11} & \gamma_{12} & \gamma_{13} \\ \gamma_{21} & \gamma_{22} & \gamma_{23} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \quad (3)$$

Dual Envelopes and Flexible Approximations - Contineud

- Applying Shephard's lemma yields

$$\begin{aligned}x_1^*(w, y) &= \alpha_1 + A_{11}w_1 + A_{12}w_2 + A_{13}w_3 + \gamma_{11}y_1 + \gamma_{21}y_2 \\x_2^*(w, y) &= \alpha_2 + A_{12}w_1 + A_{22}w_2 + A_{23}w_3 + \gamma_{12}y_1 + \gamma_{22}y_2 \\x_3^*(w, y) &= \alpha_3 + A_{13}w_1 + A_{23}w_2 + A_{33}w_3 + \gamma_{13}y_1 + \gamma_{23}y_2\end{aligned}\tag{4}$$

Normalized Quadratic

- Dividing each of the input prices by the third input price to impose homogeneity

$$\begin{aligned}
 C^*(w, y) = & \alpha_0 + \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix}' \begin{bmatrix} w_1^* \\ w_2^* \end{bmatrix} + \frac{1}{2} \begin{bmatrix} w_1^* \\ w_2^* \end{bmatrix}' \begin{bmatrix} A_{11} & A_{12} \\ A_{12} & A_{22} \end{bmatrix} \begin{bmatrix} w_1^* \\ w_2^* \end{bmatrix} + \\
 & \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}' \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} + \frac{1}{2} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}' \begin{bmatrix} B_{11} & B_{12} \\ B_{12} & B_{22} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} + \begin{bmatrix} w_1^* \\ w_2^* \end{bmatrix}' \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \\
 x_1^*(w, y) = & \alpha_1 + A_{11} w_1^* + A_{12} w_2^* + \gamma_{11} y_1 + \gamma_{21} y_2 \\
 x_2^*(w, y) = & \alpha_2 + A_{12} w_1^* + A_{22} w_2^* + \gamma_{12} y_1 + \gamma_{22} y_2
 \end{aligned} \tag{5}$$

- Note $C^*(w, y) = (w_1/w_3) x_1 + (w_2/w_3) x_2 + (w_3/w_3) x_3 = C(w, y)/w_3$.

Normalized Quadratic Likelihood Function

- Transforming Equation 5 into a statistical model, we start by rewriting this equation to focus on the errors

$$\begin{aligned} \epsilon_{1t} = & C_t^* - \left(\alpha_0 + \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix}' \begin{bmatrix} w_{1t}^* \\ w_{2t}^* \end{bmatrix} + \frac{1}{2} \begin{bmatrix} w_{1t}^* \\ w_{2t}^* \end{bmatrix}' \begin{bmatrix} A_{11} & A_{12} \\ A_{12} & A_{22} \end{bmatrix} \begin{bmatrix} w_{1t}^* \\ w_{2t}^* \end{bmatrix} + \right. \\ & \left. \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}' \begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} + \frac{1}{2} \begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix}' \begin{bmatrix} B_{11} & B_{12} \\ B_{12} & B_{22} \end{bmatrix} \begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} + \begin{bmatrix} w_{1t}^* \\ w_{2t}^* \end{bmatrix} \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{bmatrix} \begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} \right) \\ \epsilon_{2t} = & x_1^*(w, y) - (\alpha_1 + A_{11}w_{1t}^* + A_{12}w_{2t}^* + \gamma_{11}y_{1t} + \gamma_{21}y_{2t}) \\ \epsilon_{3t} = & x_2^*(w, y) - (\alpha_2 + A_{12}w_{1t}^* + A_{22}w_{2t}^* + \gamma_{12}y_{1t} + \gamma_{22}y_{2t}) \end{aligned} \quad (6)$$

Single Equation

- Consider estimating the cost function (without the derived demand equations) from 6. The likelihood function can be written as

$$L(\theta) = -\frac{T}{2} \ln(\sigma_1^2) - \sum_{t=1}^T \frac{\epsilon_{1t}^2(\theta)}{2\sigma_1^2}$$

$$\epsilon_{1t} = C_t^* - \left(\alpha_0 + \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix}' \begin{bmatrix} w_{1t}^* \\ w_{2t}^* \end{bmatrix} + \frac{1}{2} \begin{bmatrix} w_{1t}^* \\ w_{2t}^* \end{bmatrix}' \begin{bmatrix} A_{11} & A_{12} \\ A_{12} & A_{22} \end{bmatrix} \begin{bmatrix} w_{1t}^* \\ w_{2t}^* \end{bmatrix} + \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}' \begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} + \frac{1}{2} \begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix}' \begin{bmatrix} B_{11} & B_{12} \\ B_{12} & B_{22} \end{bmatrix} \begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} + \begin{bmatrix} w_{1t}^* \\ w_{2t}^* \end{bmatrix}' \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{bmatrix} \begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} \right) \quad (7)$$

Single Equation First and Second Order Conditions

- Taking the derivative of the log-likelihood function

$$\nabla_{\theta} L(\theta) = \frac{1}{2\sigma_1^2} \sum_{t=1}^T \epsilon_{1t} \nabla_{\theta} C^*(w_t, y_t) \quad (8)$$

- The derivative of the cost function is basically the vector of independent variables

$$\nabla_{\theta} C^*(w_t, y_t) = \begin{bmatrix} 1 & w_{1t}^* & w_{2t}^* & \frac{1}{2} w_{1t}^* w_{1t}^* & w_{1t}^* w_{2t}^* & \frac{1}{2} w_{2t}^* w_{2t}^* & \cdots & w_{2t}^* y_{2t} \end{bmatrix} \quad (9)$$

- Taking the derivative of gradient (derivative of Equation 8 with respect to θ)

$$\nabla_{\theta\theta}^2 L(\theta) = -\frac{1}{2\sigma_1^2} \sum_{t=1}^T [\nabla_{\theta} C^*(w_t, y_t)]' \nabla_{\theta} C^*(w_t, y_t) \quad (10)$$

Newton - Raphson

- Most empirical applications of maximum likelihood use an iterative solution approach to solve for the zero of the gradient.
- More of this in the fall, but we start with a best guess θ_0 .
- Given this initial guess, we solve for a better guess of the zero

$$\theta_1 = \theta_0 - [\nabla_{\theta\theta}^2 L(\theta)]^{-1} \nabla_{\theta} L(\theta_0) \quad (11)$$

- Next, we test for the new solution solving $\nabla_{\theta} L(\theta_1) = 0$

Multiple Equations

- The multiple equation counterparts to Equations 8, 10, and ?? are

$$\nabla_{\theta} L(\theta) = \frac{1}{2} \sum_{t=1}^T \epsilon_{1t} \Omega^{-1} \nabla_{\theta} Z(w_t, y_t)$$

$$Z(w_t, y_t) = \begin{bmatrix} C^*(w_t, y_t) \\ x_1^*(w_t, y_t) \\ x_2^*(w_t, y_t) \end{bmatrix} \quad (12)$$

$$\nabla_{\theta\theta}^2 L(\theta) = -\frac{1}{2} \sum_{t=1}^T [\nabla_{\theta} Z(w_t, y_t)]' \Omega^{-1} \nabla_{\theta} Z(w_t, y_t) \quad (13)$$

Readings

- * Cramer, J.S. 1986. *Econometric Applications of Maximum Likelihood Methods*. Cambridge University Press
- * Greene, W.H. 2012. *Econometric Analysis* Seventh Edition. (Chapter 14 pp. 509 598)