ML Estimation

Example: Multiplicative Heteroskedasticity

$$y_i = x_i^T \beta + u_i \qquad u_i \sim N(0, \sigma_i^2)$$

$$\sigma_i^2 = \exp\{z_i^T \alpha\}$$

x_i^T $1 \times k$ exogenous variables.

z_i^T $1 \times S$ exogenous variables.

$z_{1i} = \ln \sigma^2 \Rightarrow \sigma_i^2 = \sigma^2 \exp\{z_i^{*T} \gamma^*\}$

$$\ell \equiv \ln(L) = -\frac{n}{2}\ln(2\pi) - \frac{1}{2}\sum_{i=1}^{n} \ln \sigma_i^2 - \frac{1}{2}\sum_{i=1}^{n} u_i^2 / \sigma_i^2$$

$$\ell = -\frac{n}{2}\ln(2\pi) - \frac{1}{2}\sum_{i=1}^{n} z_i^T \alpha - \frac{1}{2}\sum_{i=1}^{n} \frac{(y_i - x_i^T \beta)^2}{\exp\{z_i^T \alpha\}}$$

$$\frac{\partial \ell}{\partial \beta} = \sum_{i=1}^{n} x_i (y_i - x_i^T \beta) / \exp\{z_i^T \alpha\} = 0$$

$$\frac{\partial \ell}{\partial \alpha} = -\frac{1}{2}\sum z_i + \frac{1}{2}\sum_{i=1}^{n} z_i (y_i - x_i^T \beta)^2 / \exp\{z_i^T \alpha\}$$

$$= \frac{1}{2}\sum_{i=1}^{n} z_i \left(\frac{y_i - x_i \beta)}{\exp\{z_i^T \alpha\}} - 1 \right) = 0$$

$$\frac{\partial^2 \ell}{\partial \beta \, \partial \beta^T} = - \sum_{i=1}^{n} \frac{1}{\exp\{z_i^T \alpha\}} x_i x_i^T \equiv -X^T \Omega^{-1} X$$

$$\frac{\partial^2 \ell}{\partial \beta \, \partial \alpha^T} = - \sum_{i=1}^{n} \frac{(y_i - x_i^T \beta)}{\exp\{z_i^T \alpha\}} x_i z_i^T$$

$$\frac{\partial^2 \ell}{\partial \alpha \, \partial \alpha^T} = - \frac{1}{2} \sum_{i=1}^{n} \frac{(y_i - x_i^T \beta)^2}{\exp\{z_i^T \alpha\}} z_i z_i^T$$

$$\text{Hessian}\begin{pmatrix} \beta \\ \alpha \end{pmatrix} = \begin{bmatrix} -X^T \Omega^{-1} X & -\sum_{i=1}^{n} \frac{y_i - x_i^T \beta}{\exp\{z_i^T \alpha\}} x_i z_i^T \\ & -\frac{1}{2} \sum \frac{(y_i - x_i^T \beta)^2}{\exp\{z_i^T \alpha\}} z_i z_i^T \end{bmatrix}$$

$$-E[H] = I\begin{pmatrix} \beta \\ \alpha \end{pmatrix} = \begin{bmatrix} -X^T \Omega^{-1} X & 0 \\ 0 & -\frac{1}{2} Z^T Z \end{bmatrix}$$

(1) $E(y_i - x_i^T \beta) = E[u_i \mid x_i] = 0$

(2) $E(u_i^2 / \sigma_i^2 \mid x_i, z_i) = 1$

you can see that the Information Matrix is much simpler than Hessian.

$$\underset{\sim}{\theta}_{j+1} = \theta_j - H_j^{-1} \underset{\sim}{g}_j$$

The method of scoring uses I_j instead of $-H_j$. It is block diagonal, which simplifies computation.