STAT 804

Lecture 11

Likelihood Theory

First we review likelihood theory for conditional and full maiximum likelihood estimation. Suppose the data is $X = (Y, Z)$ and write the density of X as

$$f(x|\theta) = f(y|z,\theta)f(z|\theta)$$

Differentiate the identity

$$1 = \int f(y|z,\theta)dy$$

with respect to θ_j (the jth component of θ) and pull the derivative under the integral sign to get

$$\begin{aligned} 0 &= \int \frac{\partial f(y|z,\theta_j)}{\partial\theta}dy \\ &= \int \frac{\partial \log f(y|z,\theta_j)}{\partial\theta} f(y|z,\theta)dy \\ &= \mathrm{E}_\theta(U_{Y|Z;j}(\theta)|Z) \end{aligned}$$

where $U_{Y|Z;j}(\theta)$ is the jth component of $U_{Y|Z}(\theta)$, the derivative of the log conditional likelihood; $U_{Y|Z}$ is called a conditional score. Since

$$\mathrm{E}_\theta(U_{Y|Z;j}(\theta)|Z) = 0$$

we may take expected values to see that

$$\mathrm{E}_\theta(U_{Y|Z;j}(\theta)) = 0$$

It is also true that the other two scores $U_X(\theta)$ and $U_Z(\theta)$ have mean 0 (when θ is the true value of θ). Differentiate the identity a further time with respect to θ_k to get

$$0 = \int \frac{\partial^2 \log f(y|z,\theta)}{\partial\theta_j\partial\theta_k} f(y|z,\theta)dy + \int \frac{\partial \log f(y|z,\theta_k)}{\partial\theta}\frac{\partial \log f(y|z,\theta_k)}{\partial\theta} f(y|z,\theta)dy$$

We define the conditional Fisher information matrix $I_{Y|Z}(\theta)$ to have jkth entry

$$\mathrm{E}\left[-\frac{\partial^2\ell}{\partial\theta_j\partial\theta_k}|Z\right]$$

and get

$$I_{Y|Z}(\theta|Z) = \mathrm{Var}_\theta(U_{Y|Z}(\theta)|Z)$$

The corresponding identities based on f_X and f_Z are

$$I_X(\theta) = \mathrm{Var}_\theta(U_X(\theta))$$

and

$$I_Z(\theta) = \mathrm{Var}_\theta(U_Z(\theta))$$

Now let's look at the model $X_t = \rho X_{t-1} + \epsilon_t$. Putting $Y = (X_1, \ldots, X_{T-1})$ and $Z = X_0$ we find

$$U_{Y|Z}(\rho, \sigma) = \begin{array}{c} \frac{\sum_1^{T-1}(X_t - \rho X_{t-1})X_{t-1}}{\sigma^2} \\ \frac{\sum_1^{T-1}(X_t - \rho X_{t-1})^2}{\sigma^3} - \frac{T-1}{\sigma} \end{array}$$

Differentiating again gives the matrix of second derivatives

$$\left[\begin{array}{cc} -\frac{\sum_1^{T-1} X_{t-1}^2}{\sigma^2} & -2\frac{\sum_1^{T-1}(X_t-\rho X_{t-1})X_{t-1}}{\sigma^3} \\ -2\frac{\sum_1^{T-1}(X_t-\rho X_{t-1})X_{t-1}}{\sigma^3} & -3\frac{\sum_1^{T-1}(X_t-\rho X_{t-1})^2}{\sigma^4} + \frac{T-1}{\sigma^2} \end{array} \right]$$

Taking conditional expectations given X_0 gives

$$I_{Y|Z}(\rho, \sigma) = \left[\begin{array}{cc} \frac{\sum_1^{T-1} \mathrm{E}[X_{t-1}^2|X_0]}{\sigma^2} & 0 \\ 0 & \frac{2(T-1)}{\sigma^2} \end{array} \right]$$

To compute $W_k \equiv \mathrm{E}[X_k^2|X_0]$ write $X_k = \rho X_{k-1} + \epsilon_k$ and get

$$W_k = \rho^2 W_{k-1} + \sigma^2$$

with $W_0 = X_0^2$. You can check carefully that in fact W_k converges to some W_∞ as $k \to \infty$. This W_∞ satisfies $W_\infty = \rho^2 W_\infty + \sigma^2$ which gives

$$W_\infty = \frac{\sigma^2}{1 - \rho^2}$$

It follows that

$$\frac{1}{T} I_{Y|Z}(\rho, \sigma) \to \left[\begin{array}{cc} \frac{1}{1-\rho^2} & 0 \\ 0 & \frac{2}{\sigma^2} \end{array} \right]$$

Notice that although the conditional Fisher information might have been expected to depend on X_0 it does not, at least for long series.