Econ 722 - Problem Set 5 Solutions

Qiusha Peng

March 31, 2015

Problem 3.

1. The conditional log-likelihood is

$$\log[f(y|X;\Phi,\Sigma_p)] = -\frac{q}{2}\log(2\pi) - \frac{T}{2}\log|\Sigma_p| - \frac{1}{2}tr\{(Y-X\Phi)\Sigma_p^{-1}(Y-X\Phi)'\}$$

Using the property that the function $f(A) \equiv \log|A| - tr(AB)$ is maximized uniquely by $A = B^{-1}$, by F.O.C., we have

$$\begin{aligned} \hat{\Phi} &= (X'X)^{-1}X'Y \\ \hat{\Sigma}_p &= \frac{1}{T}(Y-X\hat{\Phi})'(Y-X\hat{\Phi}) \end{aligned}$$

and the maximized log-likelihood is

$$\begin{aligned} \log[f(y|X;\hat{\Phi},\hat{\Sigma}_p)] &= -\frac{q}{2}\log(2\pi) - \frac{T}{2}\log|\hat{\Sigma}_p| - \frac{1}{2}tr\{(Y-X\hat{\Phi})\hat{\Sigma}_p^{-1}(Y-X\hat{\Phi})'\} \\ &= -\frac{q}{2}\log(2\pi) - \frac{T}{2}\log|\hat{\Sigma}_p| - \frac{q}{2} \end{aligned}$$

where the second equality uses the property $tr(AB) = tr(BA)$.

2. From what we have derived in class, we have

$$\begin{aligned} AIC &= 2\log[f(y|X;\hat{\Phi},\hat{\Sigma}_p)] - 2length(\theta) \\ BIC &= 2\log[f(y|X;\hat{\Phi},\hat{\Sigma}_p)] - \log(T)length(\theta) \end{aligned}$$

Thus, in this case,

$$\begin{aligned} AIC &= -q\log(2\pi) - T\log|\hat{\Sigma}_p| - q - 2(pq^2 + q(q+1)/2) \\ BIC &= -q\log(2\pi) - T\log|\hat{\Sigma}_p| - q - \log(T)(pq^2 + q(q+1)/2) \end{aligned}$$

which, up to a scaling factor, are

$$\begin{aligned} AIC &= \log|\hat{\Sigma}_p| + \frac{2pq^2 + q(q+1)}{T} \\ BIC &= \log|\hat{\Sigma}_p| + \frac{\log(T)(pq^2 + q(q+1)/2)}{T} \end{aligned}$$

3. For tractability, assume $p \geq p_0$. Define $X_0\Phi_0 = X\Phi^*$, where $\Phi^* = (\Phi_0', 0')'$ and Φ_0 is the true value. Then the expected log-likelihood function is

$$\begin{aligned}\triangle(\Phi, \Sigma_p) = E_0[-2L(\Phi, \Sigma_p)] \quad &\propto \quad T\log|\Sigma_p| + E_0[tr\{(X\Phi^* + U - X\Phi)\Sigma_p^{-1}(X\Phi^* + U - X\Phi)'\}] \\ &= \quad T\log|\Sigma_p| + T\,tr(\Sigma_p^{-1}\Sigma_0) + tr\{\Sigma_p^{-1}(\Phi^* - \Phi)'E_0(X'X)(\Phi^* - \Phi)\}\end{aligned}$$

thus

$$\triangle(\hat{\Phi}, \hat{\Sigma}_p) \propto T\log|\hat{\Sigma}_p| + T\,tr(\hat{\Sigma}_p^{-1}\Sigma_0) + tr\{\hat{\Sigma}_p^{-1}(\Phi^* - \hat{\Phi})'E_0(X'X)(\Phi^* - \hat{\Phi})\}$$

Since

$$E_0\{tr(\hat{\Sigma}_p^{-1}\Sigma_0)\} \approx \frac{T}{T - (pq + q + 1)} q$$

and

$$E_0[tr\{\hat{\Sigma}_p^{-1}(\Phi^* - \hat{\Phi})'E_0(X'X)(\Phi^* - \hat{\Phi})\}] = \frac{T}{T - (pq + q + 1)} pq^2$$

we have

$$E_0\{\triangle(\hat{\Phi}, \hat{\Sigma}_p)\} \approx E_0(T\log|\hat{\Sigma}_p|) + T\frac{T}{T - (pq + q + 1)} q + \frac{T}{T - (pq + q + 1)} pq^2$$

Therefore up to a scale, we have

$$AIC_c = \log|\hat{\Sigma}_p| + \frac{(T + qp)q}{T - qp - q - 1}$$

4. The code for this part is $Q3d.m$. Below are replicated tables. As discussed in Ng and Perron (2005), the scaling difference matters. Information criterions in this question are those in Hurvich and Tsai (1993) divided by the effective number of observations. Since the effective number of observations is decreasing in the model order, the order selected by information criterions in this question should be higher, which is confirmed by comparing tables below and those in Hurvich and Tsai (1993).

Table 1. Frequency of the Order Selected: VAR(1).

	Selected model order					
Criterion	1	2	3	4	5	6
AIC	777	109	46	26	17	25
AIC_c	989	9	1	0	0	1
BIC	950	43	6	1	0	0

Table 2. Frequency of the Order Selected: VAR(2).

	Selected model order					
Criterion	1	2	3	4	5	6
AIC	10	708	103	70	52	57
AIC_c	106	868	22	4	0	0
BIC	44	907	44	5	0	0