Econometric Methods II: Time Series

Lecture 6: Determining lag order of a VAR(p)

April 27, 2012

Today:

Based on Lutkepohl Ch 4.1, 4.2, 4.3, 4.6

- How to choose the VAR order

A VAR(p) model

VAR(p) process:

$$\underset{(n\times 1)}{\mathbf{y}_t} = \mathbf{c} + \mathbf{\Phi}_1 \mathbf{y}_{t-1} + \mathbf{\Phi}_2 \mathbf{y}_{t-2} + \ldots + \mathbf{\Phi}_p \mathbf{y}_{t-p} + \varepsilon_t$$

We now know how to find c and $\mathbf{\Phi}_1, \mathbf{\Phi}_2, \ldots \mathbf{\Phi}_p, \mathbf{\Omega}$ for a given p.

- But how do we choose the "right" p?

Choosing p

Fundamental trade off:

In sample fit versus over-parameterization

- More lags always makes $\widehat{\Omega}$ smaller
- But more lags decreases precision of estimates of $\mathbf{\Phi}$
  - This is captured by the small sample "corrected" one step ahead forecast error covariance

$$\widehat{\Omega}(1) = \frac{T + np + 1}{T}\widehat{\Omega}$$

There are several ways to determine the appropriate p and the best choice depends on context

Likelihood ratio test statistic

Compares fit of model while penalizing models with a larger number of parameters

$$\begin{aligned}\lambda_{LR} &= 2(L_1 - L_0) \\ &= T\left[\log\left|\widehat{\Omega}_0\right| - \log\left|\widehat{\Omega}_1\right|\right]\end{aligned}$$

The log likelihood of unrestricted model cannot be lower than that of unrestricted model, i.e. $L_1 \geq L_0$

- Critical values determine how much larger unrestricted likelihood has to be in order to reject the restricted model
- Critical values are increasing in the number of restrictions

A procedure to choose the lag order p

Assume upper bound M of p is known:

1. $\mathrm{H}_0^1 : \mathbf{\Phi}_M = 0$ versus $\mathrm{H}_1^1 : \mathbf{\Phi}_M \neq 0$
2. $\mathrm{H}_0^2 : \mathbf{\Phi}_{M-1} = 0$ versus $\mathrm{H}_1^2 : \mathbf{\Phi}_{M-1} \neq 0 \mid \mathbf{\Phi}_M = 0$
3. $\vdots$
4. $\mathrm{H}_0^M : \mathbf{\Phi}_1 = 0$ versus $\mathrm{H}_1^M : \mathbf{\Phi}_M \neq 0 \mid \mathbf{\Phi}_M = \cdots = \mathbf{\Phi}_2 = 0$

One need to be careful with distinguishing significance level of individual test versus significance of over all procedure since we may falsely reject H_0^1 and therefore never test H_0^2,$\mathrm{H}_0^3, \ldots \mathrm{H}_0^M$

German investment/income/consumption model

Lutkepohl's example:

VAR order m	$\widehat{\Omega}(m) \times 10^4$	$\lvert\widehat{\Omega}(m)\rvert \times 10^{11}$
0	$\begin{bmatrix} 21.8 & .41 & 1.23 \\ \cdot & 1.42 & .57 \\ \cdot & \cdot & 1.01 \end{bmatrix}$	2.47
2	$\begin{bmatrix} 19.2 & .62 & 1.13 \\ \cdot & 1.27 & .57 \\ \cdot & \cdot & .82 \end{bmatrix}$	1.26
4	$\begin{bmatrix} 17.0 & .57 & 1.25 \\ \cdot & 1.23 & .54 \\ \cdot & \cdot & .77 \end{bmatrix}$	.96

LR statistics for invst/income/cons example

The LR test statistic is given by the log difference of the determinants of covariance matrices of the estimated residuals

$$\lambda_{LR}(i) = T\left[\log\left|\widehat{\Omega}(m-i-1)\right| - \log\left|\widehat{\Omega}(m-i)\right|\right]$$

i	H_0^i	m under H_0^i	λ_{LR}^a
1	$\mathbf{\Phi}_4 = 0$	3	14.44
2	$\mathbf{\Phi}_3 = 0$	2	4.76
3	$\mathbf{\Phi}_2 = 0$	1	24.90
4	$\mathbf{\Phi}_1 = 0$	0	23.25

[a] Critical value for individual 5% level test $\chi^2(9)_{.95} = 16.92$

This procedure thus suggest that we should choose $p = 2$.

Alternative criteria for choosing VAR order p

LR procedure above tries to estimate the "true" p

- But perhaps we do not really care about p?
  - Choose p that minimizes forecast MSE

Small versus large sample criteria

- Consistent order selection

Akaike's Final Prediction Error Criterion (FPE)

Choose p such that approximate 1-step ahead forecast MSE are minimized

$$\widehat{\Omega}(1) = \frac{T + np + 1}{T}\widehat{\Omega}$$

and use

$$\widehat{\Omega}(p) = \frac{T}{T - np - 1}\widehat{\Omega}$$

as the estimated error covariance. Taking the determinant of the combination of (1) and (2) and gives the FPE

$$FPE(p) = \left[\frac{T + np + 1}{T - np - 1}\right]^n \times \left|\widehat{\Omega}(p)\right|$$

Choose p that minimizes FPE

Akaike Information Criterion (AIC)

AIC is very similar to FPE though motivation is different:
Choose p to minimize

$$AIC(p) = \ln\left|\widehat{\Omega}(p)\right| + \frac{2pn^2}{T}$$

where pn^2 is the number of freely estimated parameters.

Consistent order selection

A order selection criterion is called “consistent” if asymptotically (i.e. for large T) it selects the true p with probability 1.

- FPE and AIC do not select true p with prob 1 but tend to over predict the number of lags needed, i.e. $\widehat{p} > p$ with $prob > 0$ as $T \to \infty$

Consistent alternatives:

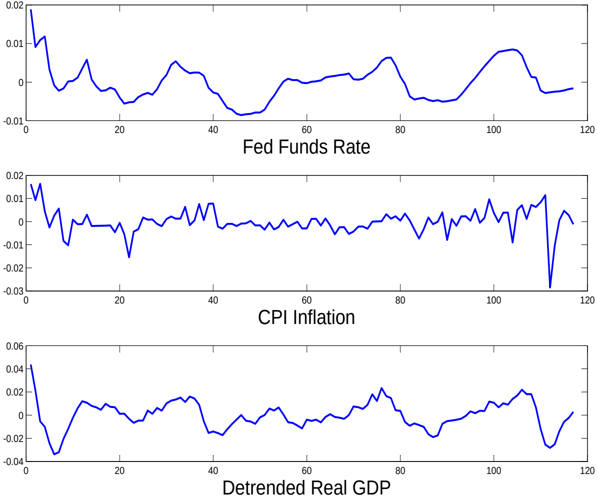
- Hannan-Quinn

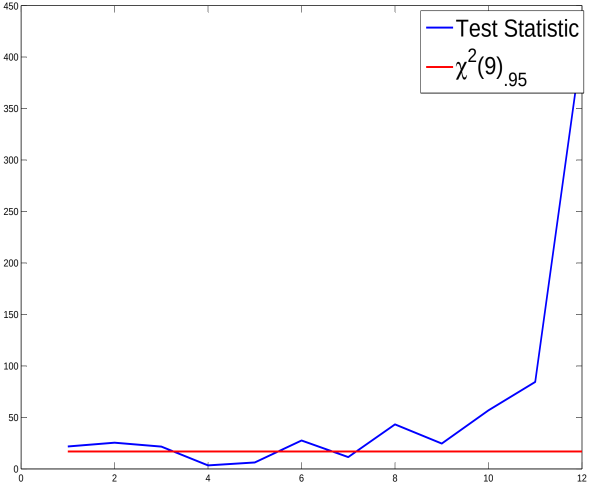
$$HQ(p) = \ln \left| \widehat{\Omega}(p) \right| + \frac{2 \ln \ln T}{T} pn^2$$

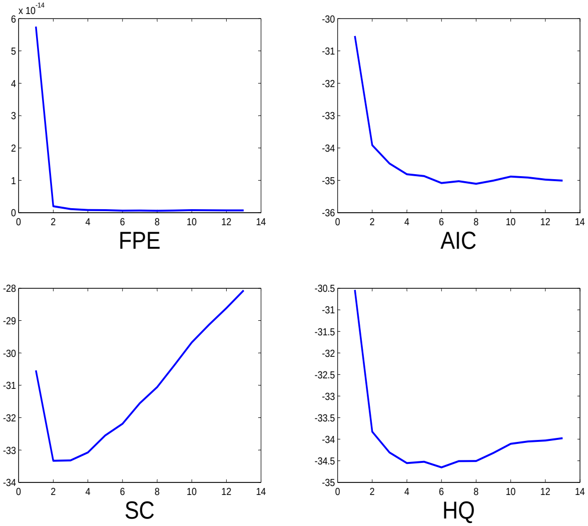
- Schwarz

$$SC(p) = \ln \left| \widehat{\Omega}(p) \right| + \frac{\ln T}{T} pn^2$$

Have a look at the data




450
400
350
300
250
200
150
100
50
0
0
2
4
6
8
10
12
Test Statistic
$\chi^2(9)_{.95}$


x 10-14
FPE
AIC
SC
HQ

A word of warning:

All tests discussed today use frequentist methods

- Most properties only known asymptotically

Later in the course we will discuss Bayesian concepts that assign probabilities to different models given the data, rather than assign probabilities of Type I and type II errors under strong assumption about the "true" DGP.