Financial Econometrics I

Lecture 8

Fu Ouyang

December 31, 2018

Outline

Model Selection using Information Criteria

In practice, we can use the following information criterion to select the order p of a VAR(p) model:

$$\mathrm{AIC}(p) = \log(|\widehat{\boldsymbol{\Sigma}}_\epsilon(p)|) + 2d^2 p/T$$

$$\mathrm{BIC}(p) = \log(|\widehat{\boldsymbol{\Sigma}}_\epsilon(p)|) + d^2 p \log(T)/T$$

$$\mathrm{HQIC}(p) = \log(|\widehat{\boldsymbol{\Sigma}}_\epsilon(p)|) + 2d^2 p \log(\log T)/T$$

We may choose p such that one of AIC, BIC and/or HQIC is minimized, but the selection should also make practical sense.

In general,

$$\widehat{p}(\mathrm{AIC}) \geq \widehat{p}(\mathrm{HQIC}) \geq \widehat{p}(\mathrm{BIC})$$

where $\widehat{p}(\cdot)$ denotes the p selected by according information criterion.

Model Diagnostics using Portmanteau Tests

As in univariate cases, we can do model diagnostic via examining if the residuals behave like (vector) white noise.

Let $\Upsilon(k)$ denote the cross-covariance matrix of the residuals at lag k. The null hypothesis

$$H_0 : \Upsilon(1) = \cdots = \Upsilon(k) = \mathbf{0}_{d\times d}$$

can be tested by a general purpose (portmanteau) Q-statistic. First, we compute residuals

$$\widehat{\epsilon}_t = \mathbf{X}_t - \widehat{\mathbf{c}} - \widehat{\mathbf{A}}_1\mathbf{X}_{t-1} - \cdots - \widehat{\mathbf{A}}_p\mathbf{X}_{t-p}$$

Let $\widehat{\Upsilon}(k)$ denote the sample analogue of $\Upsilon(k)$.

Model Diagnostics using Portmanteau Tests

Then the portmanteau test statistic is defined as

$$Q_m = \frac{1}{T^2} \sum_{k=1}^{m} \sum \frac{1}{T-k} \text{tr} \left\{ \widehat{\Upsilon}(k)' \widehat{\Upsilon}(0)^{-1} \widehat{\Upsilon}(k) \widehat{\Upsilon}(0)^{-1} \right\} \rightsquigarrow \chi^2_{d^2(m-p)}$$

where $m > p$ is an integer. In practice, this test is sensitive to the choice of m, and so it is good to do the test with different values of m.

Portmanteau tests may not be powerful when d is large.

Illustration with Real Data

Wind financial terminal (WFT)

- Wind data feed services (WDFS).
- Wind economic database (WEDB).
- WindR package for R.
- See handouts.

Example: S&P500, SH index and SZ index Data

- See R markdown.

Outline

Granger Causality

Let Z_t and Y_t be two univariate time series. Let $F(X|Y)$ denote the conditional distribution of X given Y. Then, Z_t is said to *Granger casue* Y_t if

$$F(Y_t|Y_{t-1}, Z_{t-1}, Y_{t-2}, Z_{t-2}, ...) \neq F(Y_t|Y_{t-1}, Y_{t-2}, ...)$$

In practice, this *Granger causality in distribution* is often narrowed down to *Granger causality in mean*

$$E(Y_t|Y_{t-1}, Z_{t-1}, Y_{t-2}, Z_{t-2}, ...) \neq E(Y_t|Y_{t-1}, Y_{t-2}, ...)$$

Note that the presence of Granger causality only implies the dependence between Y_t and $\{Z_{t-1}, Z_{t-2}, ...\}$ conditioning on $\{Y_{t-1}, Y_{t-2}, ...\}$, it does *not* say which causes which. It is possible that Y_t and Z_{t-k} are driven by a common latent process which is not present in the system.

Granger Causality

Granger causality can be easily verified for the VAR models. Consider a simple bivariate example:

$$\begin{pmatrix} Z_t \\ Y_t \end{pmatrix} = \mathbf{c} + \sum_{l=1}^{p} \mathbf{A}_l \begin{pmatrix} Z_{t-l} \\ Y_{t-l} \end{pmatrix} + \epsilon_t$$

where $\epsilon_t \sim \mathrm{WN}(\mathbf{0}, \mathbf{\Sigma}_\epsilon)$. Testing the Granger causality in this model is equivalent to test

$$H_0 : a_{21}^{(1)} = \cdots = a_{21}^{(p)} = 0$$

where $a_{ij}^{(l)}$ denotes the (i, j)-element of $\mathbf{A}_l$. When H_0 is rejected, Z_t is regarded as Granger causing Y_t. This test can be implemented via F-test and likelihood ratio test (both are measures of goodness of fit).

Granger Causality

- F-test:

$$pF = p \cdot \frac{(RSS_r - RSS)/p}{RSS/(2T - 2p - 2)} \rightsquigarrow \chi_p^2$$

where RSS is the *residual sum squares* obtained by LSE, i.e.

$$RSS = \sum_{t=p+1}^{T} \| \begin{pmatrix} Z_t \\ Y_t \end{pmatrix} - \widehat{\mathbf{c}} - \sum_{l=1}^{p} \widehat{\mathbf{A}}_l \begin{pmatrix} Z_{t-l} \\ Y_{t-l} \end{pmatrix} \|^2$$

and RSS_r is the RSS of the LSE under restriction H_0.

- Likelihood ratio test:

$$(T - p) \log(|\widehat{\mathbf{\Sigma}}_{\epsilon,r}| / |\widehat{\mathbf{\Sigma}}_{\epsilon}|) \rightsquigarrow \chi_p^2$$

where $\widehat{\mathbf{\Sigma}}_{\epsilon,r}$ is the MLE covariance matrix under restriction H_0.

Instantaneous Causality

Denote $\mathbf{X}_t = (Z_t, Y_t)'$. If in the VAR model, $\mathbf{\Sigma}_\epsilon$ is not diagonal, then

$$Cov(Z_t, Y_t | \mathbf{X}_{t-1}, ..., \mathbf{X}_{t-p}) \neq 0$$

In this case, Z_t and Y_t are said to have *instantaneous Granger causality*.

The instantaneous Granger causality can be tested by $H_0 : \sigma_{21} = 0$ (the off-diagonal element of $\mathbf{\Sigma}_\epsilon$) using $\widehat{\sigma}_{21}$.

Outline

Model Selection and Diagnostics

Causality

Impulse Reponse Functions

Cointegration

Impulse Reponse Functions

Another way to investigate the effect of a change in one component series on the other components is via the *impulse response functions*, which can be easily derived from the MA(∞) representation of a VAR model.

$$\mathbf{X}_t = \mathbf{c} + \epsilon_t + \sum_{k=1}^{p} \mathbf{B}_k \epsilon_{t-k} = \mathbf{c} + \Psi_0 \mathbf{e}_t + \sum_{k=1}^{p} \mathbf{\Psi}_k \mathbf{e}_{t-k}$$

where $\epsilon_t = \mathbf{\Psi}_0 e_t$, $e_t \sim \mathrm{WN}(0, I_d)$, $\mathbf{\Psi}_0 \mathbf{\Psi}_0' = \mathbf{\Sigma}_\epsilon$, and $\mathbf{\Psi}_k = \mathbf{B}_k \mathbf{\Psi}_0$.

The matrices $\mathbf{\Psi}_0, \mathbf{\Psi}_1, \ldots$ are called the impulse response functions. Denote $\psi_{ij}^{(k)}$ as the (i,j)-th element of $\mathbf{\Psi}_k$. $\psi_{ij}^{(k)}$ is called the response function of $\mathbf{X}_{t,i}$ to the impulse on the j-th component of $\mathbf{e}_{t-k}$.

Note that $\mathbf{\Psi}_0$ is *not* unique as $\mathbf{\Psi}_0 \mathbf{H} \mathbf{H}' \mathbf{\Psi}_0' = \mathbf{\Psi}_0 I_d \mathbf{\Psi}_0' = \mathbf{\Psi}_0 \mathbf{\Psi}_0'$ for all $d \times d$ orthogonal matrix $\mathbf{H}$. Most statistical packages (by default) construct $\mathbf{\Psi}_0$ using the Cholesky decomposition of $\mathbf{\Sigma}_\epsilon$.

Outline

Unit Root and Cointegration

For univariate time series $X_t \sim \text{ARIMA}(p, 1, 0)$, $\nabla X_t = X_t - X_{t-1}$ is stationary, and X_t is said to have a *unit root* and denoted as $X_t \sim I(1)$.

Similarly, for $X_t \sim \text{ARIMA}(p, k, 0)$, X_t is said to have k unit roots and denoted as $X_t \sim I(k)$.

Let $X_t \sim I(1)$ and $Y_t \sim I(1)$. They are said to be *conintegrated* if there is a non-zero constant such that $Y_t - \beta X_t \sim I(0)$, i.e., stationary.

More generally, a vector time series $\mathbf{X}_t$ is said to be cointegrated with order (k, h) ($k \geq h \geq 1$), denoted as $\mathbf{X}_t \sim \text{CI}(k, h)$ if

1. all component series of $\mathbf{X}_t$ are $I(k)$, and
2. there is a non-zero constant vector β such that $\beta' \mathbf{X}_t \sim I(k - h)$.

Most frequently used cointegration model is $\text{CI}(1, 1)$.

Engle-Granger Two-Step Estimation

Assume that $X_t \sim I(1)$ and $Y_t \sim I(1)$ (Test this first in practice using the *augmented Dickey-Fuller* (ADF) test!).

- Step 1: run the following regression and compute residuals

$$Y_t = \alpha + \beta X_t + u_t \rightarrow \widehat{u}_t = \widehat{Y}_t - \widehat{\alpha} - \widehat{\beta} X_t$$

  then test if $\widehat{u}_t$ has unit root using the *cointegration augmented Dickey-Fuller* (CADF) test.

- Step 2: estimate the following *error-correction model* (ECM)

$$\nabla Y_t = a_0 + a_1 \widehat{u}_{t-1} + a_2 \nabla X_{t-1} + a_3 \nabla Y_{t-1} + \epsilon_{t1}$$
$$\nabla X_t = b_0 + b_1 \widehat{u}_{t-1} + b_2 \nabla Y_{t-1} + b_3 \nabla X_{t-1} + \epsilon_{t2}$$

  where $\widehat{u}_{t-1}$ is included to control the error from step 1.

Error-Correction Model

Consider the following dynamic model

$$\begin{aligned}
Y_t &= \beta_0 + \gamma X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \alpha_1 Y_{t-1} + \epsilon_t \\
&= \beta_0 + \gamma X_t + (\beta_1 + \beta_2) X_{t-1} - (\beta_1 + \beta_2) \nabla X_{t-1} + \alpha_1 Y_{t-1} + \epsilon_t
\end{aligned}$$

The "long-run" relationship between X and Y is

$$(1 - \alpha_1) Y = \beta_0 + (\gamma + \beta_1 + \beta_2) X$$

and for each t,

$$Y_t = \frac{\beta_0}{1 - \alpha_1} + \frac{\gamma + \beta_1 + \beta_2}{1 - \alpha_1} X_t + u_t$$

which implies that

$$\nabla Y_t = \gamma \nabla X_t - (\beta_1 + \beta_2) \nabla X_{t-1} - u_{t-1} + \epsilon_t$$

Spurious Regression

Consider the following regression model

$$Y_t = \alpha + \beta X_t + u_t$$

If X_t and Y_t are independent (misspecified cointegration model), $\beta = 0$ theoretically. However, for nonstationary X_t and/or Y_t,

- OLS estimator $\widehat{\beta}$ is not asymptotically normal.
- $\widehat{\beta}/SE(\widehat{\beta}) = O_p(\sqrt{T})$, and so reject $H_0 : \beta = 0$ for large T.

This phenomenon is called *spurious regression*.

In practice, it is important to perform the CADF test for $\widehat{u}_t$ before running ECM estimation.

- If the unit-root hypothesis for $\widehat{u}_t$ is rejected, proceed to fitting ECM.
- Otherwise, consider running regression with ∇Y_t and ∇X_t.

Illustration with Real Data

Example: S&P 500 and 400 Data

- See R markdown.