Multivariate models

Matthieu Stigler
Matthieu.Stigler@gmail.com

November 5, 2008

Lectures list

1. Lectures

2. Multivariate models
   - Spurious regression
   - Autoregressive distributed lag model (ARDL)
   - Partial adjustment model
   - Vector autoregressive model (VAR)
     - historical approach and the Sims (1980) criticism to simultaneous models
     - Selection of the lag order
     - Estimation
     - Diagnostic checks
     - Granger Causality
     - Forecast

Outline

1. Lectures

2. Multivariate models
    - Spurious regression
    - Autoregressive distributed lag model (ARDL)
    - Partial adjustment model
    - Vector autoregressive model (VAR)
        - historical approach and the Sims (1980) criticism to simultaneous models
        - Selection of the lag order
        - Estimation
        - Diagnostic checks
        - Granger Causality
        - Forecast

Outline

1. Lectures
2. Multivariate models
   - Spurious regression
   - Autoregressive distributed lag model (ARDL)
   - Partial adjustment model
   - Vector autoregressive model (VAR)
     - historical approach and the Sims (1980) criticism to simultaneous models
     - Selection of the lag order
     - Estimation
     - Diagnostic checks
     - Granger Causality
     - Forecast

Outline

1. Lectures
2. Multivariate models
   - Spurious regression
   - Autoregressive distributed lag model (ARDL)
   - Partial adjustment model
   - Vector autoregressive model (VAR)
     - historical approach and the Sims (1980) criticism to simultaneous models
     - Selection of the lag order
     - Estimation
     - Diagnostic checks
     - Granger Causality
     - Forecast

Outline

1. Lectures
2. Multivariate models
   - Spurious regression
   - Autoregressive distributed lag model (ARDL)
   - Partial adjustment model
   - Vector autoregressive model (VAR)
     - historical approach and the Sims (1980) criticism to simultaneous models
     - Selection of the lag order
     - Estimation
     - Diagnostic checks
     - Granger Causality
     - Forecast

VAR example: $\begin{bmatrix} y_{1,t} \\ y_{2,t} \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} + \begin{bmatrix} A_{1,1} & A_{1,2} \\ A_{2,1} & A_{2,2} \end{bmatrix} \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \end{bmatrix} + \begin{bmatrix} e_{1,t} \\ e_{2,t} \end{bmatrix}$

or, equivalently, as the following system of two equations

$$\begin{cases} y_{1,t} & = c_1 + A_{1,1} y_{1,t-1} + A_{1,2} y_{2,t-1} + e_{1,t} \\ y_{2,t} & = c_2 + A_{2,1} y_{1,t-1} + A_{2,2} y_{2,t-1} + e_{2,t} \end{cases}$$

VAR(p) as VAR(1): $y_t = c + A_1 y_{t-1} + A_2 y_{t-2} + e_t$
can be recast as the VAR(1) model

$$\begin{bmatrix} y_t \\ y_{t-1} \end{bmatrix} = \begin{bmatrix} c \\ 0 \end{bmatrix} + \begin{bmatrix} A_1 & A_2 \\ I & 0 \end{bmatrix} \begin{bmatrix} y_{t-1} \\ y_{t-2} \end{bmatrix} + \begin{bmatrix} e_t \\ 0 \end{bmatrix}$$

VMA representation

VAR(1):

$$
\begin{aligned}
y_t &= A_1 y_{t-1} + \varepsilon_t && (1)\\
&= A_0 + \sum_{i=0}^{\infty} A_1^i \varepsilon_{t-i} && (2)
\end{aligned}
$$

$$
\begin{aligned}
Y_t &= \mathbf{A}_1 Y_{t-1} + E_t && (3)\\
&= \sum_{i=0}^{\infty} \mathbf{A}_1^i E_{t-i} && (4)
\end{aligned}
$$

Estimation

$$\begin{bmatrix} y_{1,t} \\ y_{2,t} \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} + \begin{bmatrix} A_{1,1} & A_{1,2} \\ A_{2,1} & A_{2,2} \end{bmatrix} \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \end{bmatrix} + \begin{bmatrix} e_{1,t} \\ e_{2,t} \end{bmatrix}$$

Define:

$$\text{Cov} \begin{bmatrix} e_{1,t} \\ e_{2,t} \end{bmatrix} = \begin{bmatrix} \text{Var}(e_1) & \text{Cov}(e_1, e_2) \\ \text{Cov}(e_1, e_2) & \text{Var}(e_2) \end{bmatrix} \equiv \Sigma$$

Parameters to estimate: $\{A_0, A_1, \dots, A_p, \Sigma\}$

Rewrite:

- $\mathbf{Y} = \left[Y_1, \dots, Y_T \right]$, a $k \times T$ matrix
- $\mathbf{Y} = \left[A_0, A_1, \dots, A_p \right]$, a $k \times kp + 1$ matrix

So we have aVAR(1): $Y = AZ + U$
The estimator of A is given by: $\hat{A} = YZ'(ZZ')^{-1}$

Proposition

If the residuals are white noise and $plimZZ'/T = \Gamma$ then:
the MLS estimator is the same as the conditionnal MLE estimator and hence is:

- *Convergent*
- *Asymptotically efficient*
- *Normally distributed:* $\sqrt{T}vec(\hat{B} - B) \sim \mathcal{N}(0, \Gamma^{-1} \otimes \Sigma)$

The estimator of Σ is given by:

- MLE: $\hat{\Sigma} = \frac{1}{T} \sum_{t=1}^{T} \hat{\epsilon}_t \hat{\epsilon}_t'$
- OLS: $\hat{\Sigma} = \frac{1}{T-kp-1} \sum_{t=1}^{T} \hat{\epsilon}_t \hat{\epsilon}_t'$ for a model with a constant, "k" variables and "p" lags

These both estimators are convergent

In a matrix notation, this gives:
$\hat{\Sigma} = \frac{1}{T-kp-1}(Y - \hat{B}Z)(Y - \hat{B}Z)'$
Note that for the GLS estimator the covariance matrix of the errors becomes:
The covariance matrix of the parameters can be estimated as
$\hat{C}ov(Vec(\hat{B})) = (ZZ')^{-1} \otimes \hat{\Sigma}$

Proposition

The multivariate least square (MLS) is equal to the OLS applied equation by equation

Note: this is because every equation has the same explanatory variables.

a

a

- Tests
  - LR ratio
- Criteria
  - MSE criterion
  - Information criterions

LR ratio test

Test VAR(p) against VAR(q):
$T\left(\log(|\Sigma_p| - \log(|\Sigma_q|)\right) \qquad p < q$
It has teh distribution: $LR \sim \chi^2\left((q-p)k^2\right)$
Note: on can reduce the test by $T - c$ where c=nubmer of parameter estimated in each equation of the unrestricted model (so $q * k+$ const/trend)

discuss/analyse Clarida, Gali, Gutler

Conditions for stability

Theorem (AR(1))

The characteristics roots must lie within the unit circle.

Forecast

Recall from lecture 2:

Notation (Forecast)

$\hat{y}_{t+j} \equiv \mathrm{E}_t(y_{t+j}) = \mathrm{E}(y_{t+j} | y_t, y_{t-1}, \ldots, \varepsilon_t, \varepsilon_{t-1}, \ldots)$ *is the conditional expectation of* y_{t+j} *given the information available at* t.

Definition (J-step-ahead forecast error)

$e_t(j) \equiv y_{t+j} - \hat{y}_{t+j}$

Forecast of a VAR(1)

Take the VAR(1): $Y_t = A_0 + A_1 Y_{t-1} + A_2 Y_{t-2} + \varepsilon_t$
The 1-step-ahead forecast is: $\hat{Y}_{t+1} \equiv \mathrm{E}_t(Y_{t+1}) = A_0 + A_1 Y_t$
And the 1-step-ahead forecast error is:

$$err_t(j) \equiv y_{t+j} - \hat{y}_{t+j} = \overbrace{A_0 + A_1 Y_t}^{\hat{Y}_{t+1}} - \overbrace{A_0 + A_1 Y_t + \varepsilon_t}^{Y_{t+1}} = \varepsilon_t$$

Proposition (Generalization)

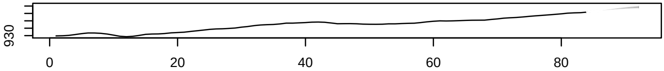
And the j-ahead forecast is: $\hat{Y}_{t+j} = (I + A_1 + A_1^2 + \ldots + A_1^{j-1} +)A_0 + A_1^j Y_t$
The error is: $\varepsilon_{t+j} + A_1 \varepsilon_{t+j-1} + A_1^2 \varepsilon_{t+j-2} + \ldots + A_1^{j-1} \varepsilon_{t+1}$

So the variance of the forecast error is increasing!

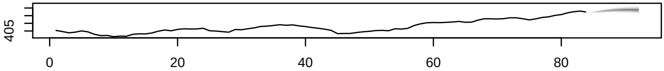
Similarly, starting from the VMA(∞) representation:
$Y_t = \mu + \sum_{i=0}^{\infty} \Phi_i u_{t+n+1}$

```
> library(vars)
> data(Canada)
> var.2c <- VAR(Canada, p = 2, type = "const")
> var.2c.prd <- predict(var.2c, n.ahead = 8, ci = 0.95)
> fanchart(var.2c.prd)
```

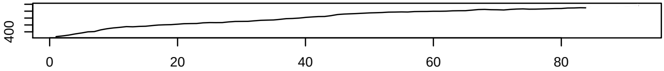
Fanchart for variable e



Fanchart for variable prod



Fanchart for variable rw



Fanchart for variable U