VECTOR AUTOREGRESSION

Date: February 28, 2018

JEN-WEN LIN, PhD, CFA

REMAINING SCHEDULE/TOPICS

Date	Topic (tentative)
1-Mar-18	Private asset modeling & vector autoregression
8-Mar-18	Multivariate time series modeling*
15-Mar-18	Boostraping time series and state space model
22-Mar-18	Bagging/boosting/ensemble time series forecast
29-Mar-18	Catch-up and graduate student presentation
(*):	Return midterm (last date to drop March 14)

VECTOR AUTOREGRESSION

INTRODUCTION

1. Vector autoregressions (VAR) is a generalization of the autoregressive and moving average (ARMA) process.
2. Christopher A. Sims (1980) introduced VAR to econometricians and won the Nobel prize in 2011.
   - The academy said he had developed a method based on "vector autoregression" to analyse how the economy is affected by temporary changes in economic policy and other factors - for instance, the effects of an increase in the interest rate set by a central bank.--BBC News Website.

WHY USE VAR AND WHAT VAR CAN DO

- A concise way of summarizing interrelationship among data.
- Good forecast results.
- Testing for Granger causality among time series
- Finance applications: strategic asset allocation (Campbell et al 2003), variance decomposition of excess stock returns (Campbell 1991) and long-horizon stock return predictability (Campbell and Shiller, 1989)

VECTOR AUTOREGRESSION OF ORDER ONE

- The VAR(1)-process of K endogenous variables is defined as

$$\boldsymbol{y}_t = \boldsymbol{A}\boldsymbol{y}_{t-1} + \boldsymbol{u}_t,$$

where $\boldsymbol{y}_t = (y_{1t}, \ldots, y_{jt}, \ldots, y_{Kt})'$ for $j = 1, \ldots K$, $\boldsymbol{A}$ are $(K \times K)$ coefficient matrix, and $\boldsymbol{u}_t$ is a K-dimensional white noise process with time-invariant positive definite covariance matrix $E(\boldsymbol{u}_t\boldsymbol{u}'_t) = \Sigma_u$.

- By repeated substitution, the above VAR(1) model becomes

$$\boldsymbol{y}_t = \boldsymbol{u}_t + \boldsymbol{A}\boldsymbol{u}_{t-1} + \boldsymbol{A}^2\boldsymbol{u}_{t-2} + \boldsymbol{A}^3\boldsymbol{u}_{t-3} + \cdots$$

- For the above VMA(infinity) process to be stationary, $\boldsymbol{A}^j$ must converge to zero as j goes to infinity. Mathematically, we require that all K eigenvalues of $\boldsymbol{A}$ be less than one in modulus.

VECTOR AUTOREGRESSION OF ORDER P

- The VAR(p)-process of K endogenous variables is defined as

$$\boldsymbol{y}_t = \boldsymbol{A}_0 + \boldsymbol{A}_1\boldsymbol{y}_{t-1} + \ldots + \boldsymbol{A}_p\boldsymbol{y}_{t-p} + \boldsymbol{u}_t, (1)$$

where $\boldsymbol{y}_t = (y_{1t}, \ldots, y_{kt}, \ldots, y_{Kt})$ for $k = 1, \ldots K$, $\boldsymbol{A}_0$ stands for a (K x 1) mean vector, $\boldsymbol{A}_i$ are (K × K) coefficient matrices for $i = 1, \ldots, p$ and $\boldsymbol{u}_t$ is a K-dimensional white noise process with time-invariant positive definite covariance matrix $E(\boldsymbol{u}_t\boldsymbol{u}'_t) = \Sigma_u$.

  - We could include deterministic regressors, such as a constant, trend, and seasonal dummy variables, in eqn. (1)

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}_t = \begin{bmatrix} 5.0 \\ 10.0 \end{bmatrix} + \begin{bmatrix} 0.5 & 0.2 \\ -0.2 & -0.5 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}_{t-1} + \begin{bmatrix} -0.3 & -0.7 \\ -0.1 & 0.3 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}_{t-2} + \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}_t .$$

VECTOR AUTOREGRESSION OF ORDER P

- Equation (1) may be written in the compact form

$$A(B)\boldsymbol{y}_t = \boldsymbol{u}_t, (2)$$

where $A(B) = (\boldsymbol{I}_K - \boldsymbol{A}_1 B - \ldots - \boldsymbol{A}_p B^p)$.

- One important question to ask for a VAR(p)-process is how to check its stationarity (stability).
- The necessary and sufficient condition for the stationarity of $\boldsymbol{y}_t$ is that all solutions (roots) of $\det(\boldsymbol{I}_K - \boldsymbol{A}_1 B - \ldots - \boldsymbol{A}_p B^p) = 0$ are greater than one in absolute value.

THE COMPANION FORM OF VAR(P) PROCESS

- In practice, the stability of a VAR(p) model is analyzed via its companion form.
- The companion form of a VAR(p)-process is given by

$$\boldsymbol{\xi}_t = \boldsymbol{A}\boldsymbol{\xi}_{t-1} + \boldsymbol{v}_t, \qquad (3)$$

with

$$\xi_t = \begin{bmatrix} \mathbf{y}_t \\ \vdots \\ \mathbf{y}_{t-p+1} \end{bmatrix}, \quad A = \begin{bmatrix} A_1 & A_2 & \cdots & A_{p-1} & A_p \\ I & 0 & \cdots & 0 & 0 \\ 0 & I & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & I & 0 \end{bmatrix}, \quad \mathbf{v}_t = \begin{bmatrix} \mathbf{u}_t \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{bmatrix},$$

where the dimension of the stacked vectors $\boldsymbol{\xi}_t$ and $\boldsymbol{v}_t$ is $(Kp \times 1)$ and that of the matrix $\boldsymbol{A}$ is $(Kp \times Kp)$. The companion form of a VAR(p) process is also a VAR(1) process.

- If the moduli of the *eigenvalues* of $\boldsymbol{A}$ are less than one, then the VAR(p)-process is stable/stationary.

EXAMPLE

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}_t = \begin{bmatrix} 5.0 \\ 10.0 \end{bmatrix} + \begin{bmatrix} 0.5 & 0.2 \\ -0.2 & -0.5 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}_{t-1} + \begin{bmatrix} -0.3 & -0.7 \\ -0.1 & 0.3 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}_{t-2} + \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}_t .$$

$$A = \begin{bmatrix} A_1 & A_2 & \cdots & A_{p-1} & A_p \\ I & 0 & \cdots & 0 & 0 \\ 0 & I & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & I & 0 \end{bmatrix}, \quad A = \begin{bmatrix} 0.5 & 0.2 & -0.3 & -0.7 \\ -0.2 & -0.5 & -0.1 & 0.3 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

```
> A<-matrix(c(0.5,0.2,-0.3,-0.7,-0.2,-0.5,-0.1 , 0.3 , 1,0 ,0 ,0 ,0 ,1,0,0),nrow=4, byrow = TRUE)
> A
      [,1] [,2] [,3] [,4]
[1,]   0.5  0.2 -0.3 -0.7
[2,]  -0.2 -0.5 -0.1  0.3
[3,]   1.0  0.0  0.0  0.0
[4,]   0.0  1.0  0.0  0.0
```

```
> ev<-eigen(A,only.value=TRUE)
$values
> ev[1]
[1] -0.8180175+0i
> ev[2]
[1] 0.5974589+0i
> sqrt(Re(ev[3])^2+Im(ev[3])^2)
[1] 0.5721695
> sqrt(Re(ev[4])^2+Im(ev[4])^2)
[1] 0.5721695
```

ESTIMATION OF A STABLE VAR(P) PROCESS

- For a given sample of the endogenous variables $\boldsymbol{y}_1, \ldots \boldsymbol{y}_T$ and sufficient presample values $\boldsymbol{y}_{-\boldsymbol{p}+\mathbf{1}}, \ldots, \boldsymbol{y}_0$, the coefficients of a VAR(p)-process can be estimated efficiently by least squares applied separately to each of the equations.
- If the error process $\boldsymbol{u}_t$ is normally distributed, then this estimator is equal to the maximum likelihood estimator conditional on the initial values.

TEST MODEL ADEQUACY
PORTMANTEAU TESTS

Test statistics:

$$Q_{BP} = T\sum_{j=1}^{m} \mathrm{tr}(\hat{C}_j^T \hat{C}_0^{-1} \hat{C}_j \hat{C}_0^{-1}) \sim \chi^2_{k^2 m - n^*}$$

$$Q_{LB} = T^2\sum_{j=1}^{m} \frac{1}{T\text{-}j} \mathrm{tr}(\hat{C}_j^T \hat{C}_0^{-1} \hat{C}_j \hat{C}_0^{-1}) \sim \chi^2_{k^2 m - n^*}$$

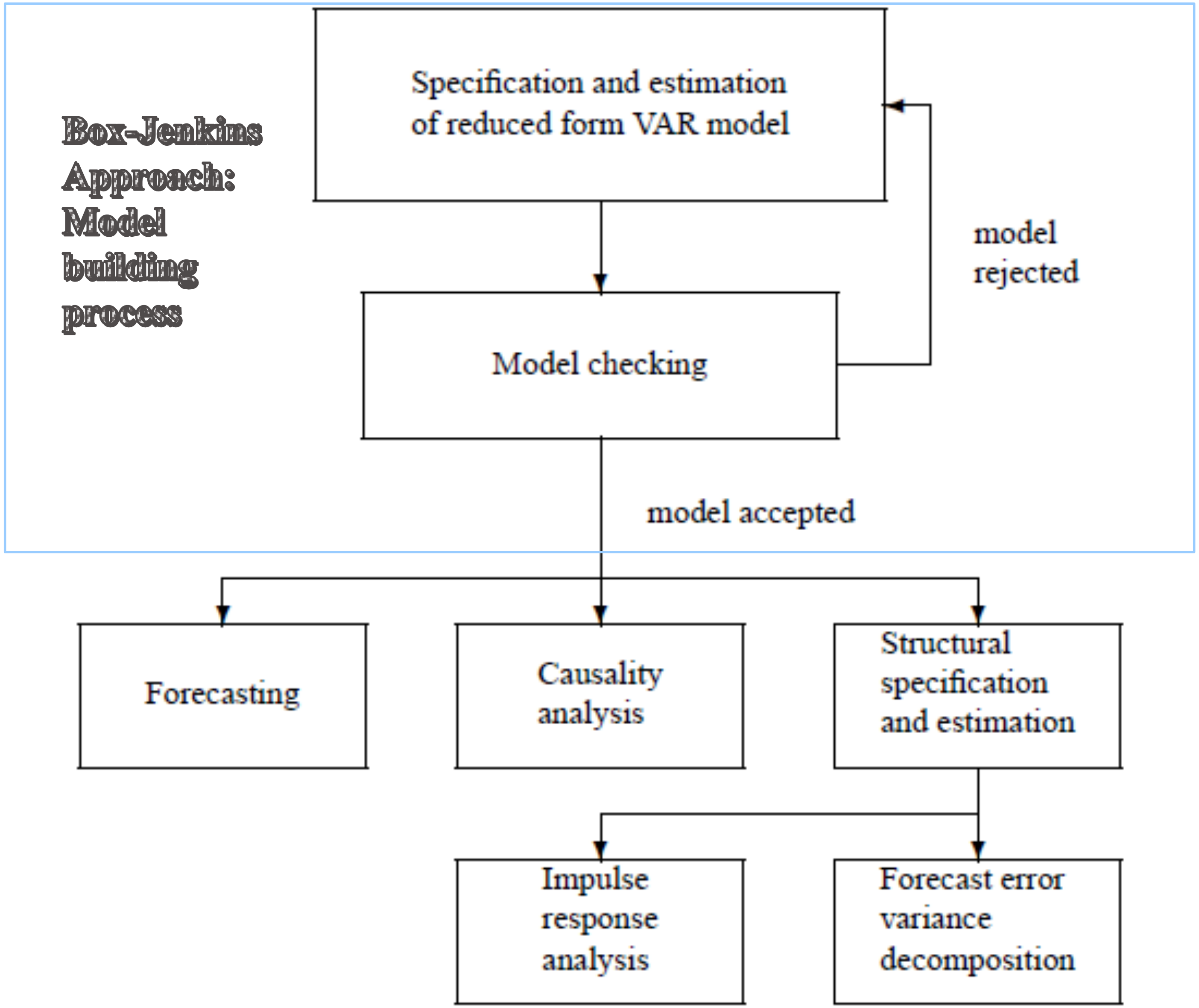
where n^* is the number of coefficients excluding deterministic terms of a $VAR(p)$ model and

$$\hat{C}_i = \frac{1}{T}\sum_{t=i+1}^{T} \hat{\mathbf{a}}_t \hat{\mathbf{a}}_{t-i}^T$$

ORDER SELECTION AND BIG DATA

- Order selection
  - Sequential likelihood ratio tests
    - Based on the likelihood ratio test for testing VAR(p) versus VAR(p-1)
  - Information criteria
- BigVAR: Dimension Reduction Methods for Multivariate Time Series

VAR ANALYSIS (LUTKEPOHL, 2006)



15 GRANGER CAUSALITY

- Causality tests are useful to infer whether a variable helps predict another one.
- An operational definition of causality between two time series can be defined in terms of predictability (Granger, 1969).

GRANGER CAUSALITY

- Ideally, causality may be defined through the concept of the conditional distribution.
  - y_{2t} does not cause y_{1t} if the distribution of y_{1t}, conditional on past values of both y_{1t} and y_{2t}, is the same as the distribution of y_{1t} conditional on its own past values.
- In practice, it would be very difficult to test whether the entire distribution y_{1t} depends on past values of y_{2t}.
- Therefore, we consider an alternative by asking whether the conditional mean of y_{1t} depends on past values of y_{2t}. If this is the case, we can test causality by imposing restrictions on a VAR model.

GRANGER CAUSALITY

- Consider a VAR(p) model as follows:

$$\begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} + \sum_{j=1}^{p} \begin{bmatrix} \phi_{j,11} & \phi_{j,12} \\ \phi_{j,21} & \phi_{j,22} \end{bmatrix} \begin{bmatrix} y_{1,t-j} \\ y_{2,t-j} \end{bmatrix} + \begin{bmatrix} a_{1t} \\ a_{2t} \end{bmatrix}$$

- If y_{2t} does not Granger cause y_{1t}, then all of the $\phi_{j,12}$'s must be zero. Note that $\phi_{j,12}$'s only appear in the equation for y_{1t}.

- Similarly, if y_{1t} does not Granger cause y_{2t}, then all of the $\phi_{j,21}$'s must be zero.

LIKELIHOOD RATIO TEST

- Obtain ML (or OLS) estimates of the following equations.

$$y_{1t} = \alpha_1 + \sum_{j=1}^{p} \phi_{j,11} y_{1,t-j} + e_{1t}, \qquad (2)$$

$$y_{1t} = \alpha_1 + \sum_{j=1}^{p} \phi_{j,11} y_{1,t-j} + \sum_{j=1}^{p} \phi_{j,12} y_{2,t-j} + \varepsilon_{1t}, \qquad (3)$$

- Calculate the values of the log likelihood functions in eqn. (2) and (3). And, the LR statistic is given by

$$n(\log|\tilde{\Sigma}| - \log|\hat{\Sigma}|) \sim \chi_p^2$$

where $\tilde{\Sigma}$ and $\hat{\Sigma}$ denote the residual covariance matrix estimated from eqn. (5) and (6), respectively.

PORTMANTEAU TEST FOR GRANGER CAUSALITY

- Pierce and Haugh (1977) expanded up the work of Granger (1969) and gave a comprehensive survey regarding research on causality in temporal systems.
- For simplicity, in what follows, we consider the case of two time series $\{X_t\}$ and $\{Y_t\}$.
- Let $\{X_t\}$ and $\{Y_t\}$ be causal and invertible univariate $ARMA$ processes and be given by

$$\phi_X(B)(X_t - \mu_X) = \theta_X(B)u_t, \quad u_t \sim \mathrm{WN}(0, \sigma_u^2)$$
$$\phi_Y(B)(Y_t - \mu_Y) = \theta_Y(B)v_t, \quad v_t \sim \mathrm{WN}(0, \sigma_v^2)$$

PORTMANTEAU TEST FOR GRANGER CAUSALITY

The cross-correlation function at lag k between u_t and v_t series is given by

$$\rho_{uv}(k) = \frac{E(u_t, v_{t+k})}{\sqrt{E(u_t^2)E(v_t^2)}}$$

Pierce and Haugh (1977) explained that there are many possible types of causal interpretation between $\{X_t\}$ and $\{Y_t\}$ which can be characterized by the properties of $\rho_{uv}(k)$.

PORTMANTEAU TEST FOR GRANGER CAUSALITY

RELATIONSHIPS	RESTRICTIONS ON $\rho_{uv}(k)$
X causes Y	$\rho_{uv}(k) \neq 0$ for *some* $k > 0$
Y causes X	$\rho_{uv}(k) \neq 0$ for *some* $k < 0$
Instantaneous Causality	$\rho_{uv}(0) \neq 0$
Feedback	$\rho_{uv}(k) \neq 0$ for *some* $k > 0$ and for some $k < 0$
X causes Y but not instantaneously	$\rho_{uv}(k) \neq 0$ for *some* $k > 0$ and $\rho_{uv}(0) = 0$
Y does not cause X	$\rho_{uv}(k) = 0$ for *all* $k < 0$
Y does not cause X at all	$\rho_{uv}(k) = 0$ for *all* $k \leq 0$
Unidirectional causality from X to Y	$\rho_{uv}(k) \neq 0$ for some $k > 0$ and $\rho_{uv}(k) = 0$ for either (a) all $k < 0$ or (b) all $k \leq 0$
X and Y are only related instantaneously	$\rho_{uv}(0) \neq 0$ and $\rho_{uv}(k) = 0$ for all $k \neq 0$
X and Y are independent	$\rho_{uv}(k) = 0$ for *all* k

Portmanteau tests for Granger causality

- $H_0 : X$ does not cause Y
- $Q_L = n^2 \sum_{k=0}^{L} (n-k)^{-1} r_{uv}^2(k) \sim \chi_{L+1}^2$