STAT302: Time Series Analysis

Chapter 7. Multivariate Time Series Models

Sangbum Choi, Ph.D

Department of Statistics, Korea University

Outline

Multivariate Stationary Processes

Vector Autoregressive (VAR) Models

Introduction

- Observations are often taken simultaneously on two or more time series.
- In economics, many different measures of economic activity are typically recorded at regular intervals.
- Examples include the retail price index, the gross domestic product and the level of unemployment.
- Given multivariate data like these, it may be helpful to develop a multivariate model to describe the interrelationships among the series.
- This section takes a more detailed look at some multivariate time-series models, such as vector autoregressive (VAR) models.

Jointly stationary processes

- Two time series, X_t and Y_t, are said to be **jointly stationary** if they are each stationary, and the **cross-covariance function**

$$\gamma_{XY}(h) = \mathrm{Cov}(X_{t+h}, Y_t) = E[(X_{t+h} - \mu_X)(Y_t - \mu_Y)]$$

  is a function only of lag h.

- The **cross-correlation function (CCF)** of jointly stationary time series X_t and Y_t is defined as

$$\rho_{XY}(h) = \frac{\gamma_{XY}(h)}{\sqrt{\gamma_X(0)\gamma_Y(0)}}$$

- Note that $\rho_{XY}(h) = \rho_{YX}(-h)$.

Jointly stationary processes

- For example, let

$$X_t = \epsilon_t + \epsilon_{t-1}, \quad Y_t = \epsilon_t - \epsilon_{t-1},$$

where $\epsilon_t \sim \mathrm{WN}(0, \sigma_\epsilon^2)$.

- It can be shown that

$$\rho_{XY}(h) = \begin{cases} 0 & h = 0, \\ 1/2 & h = 1, \\ -1/2 & h = -1, \\ 0 & |h| \geq 2. \end{cases}$$

- Clearly, the CCF depends only on the lag separation h, so the series are jointly stationary.

Bivariate white noise processes

- Two series $\{\epsilon_{X,t}\}$ and $\{\epsilon_{Y,t}\}$ are **bivariate white noise** if they are stationary and their cross-covariance $\gamma_{XY}(h) = \mathrm{Cov}(\epsilon_{X,t+h}, \epsilon_{Y,t})$ satisfies

$$\gamma_{XX}(h) = \gamma_{YY}(h) = \gamma_{XY}(h) = 0 \text{ for all } h \neq 0$$

- In the equation above, $\gamma_{XX}(0) = \gamma_{YY}(0) = 1$ and $\gamma_{XY}(0)$ may be zero or non- zero.
- Hence, bivariate white noise series $\{\epsilon_{X,t}\}$ and $\{\epsilon_{Y,t}\}$ may be regarded as white noise when considered individually but when considered as a pair may be cross-correlated at lag 0.

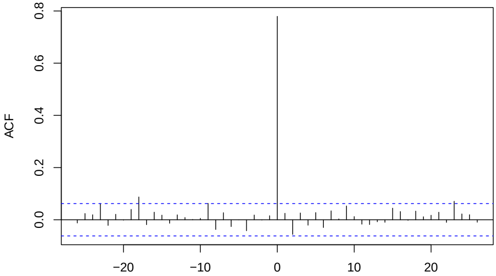
Multivariate white noise processes

- The definition of bivariate white noise readily extends to multivariate white noise.
- Let $\gamma_{ij}(h) = \text{Cov}(\epsilon_{i,t+h}, \epsilon_{j,t})$ be the cross-correlation between the series $\{\epsilon_{i,t}\}$ and $\{\epsilon_{j,t}\}$, $(i, j = 1, ...n)$.
- Then stationary series $\{\epsilon_{1,t}, \epsilon_{2,t}, ..., \epsilon_{n,t}\}$ are multivariate white noise if each individual series is white noise and, for each pair of series $(i \neq j)$, $\gamma_{ij}(h) = 0$ for all $h \neq 0$.
- In other words, multivariate white noise is a sequence of independent draws from some multivariate distribution.

Multivariate white noise processes

The ccf() function verifies that the cross-correlations are approximately zero for all non-zero lags

```
library(MASS)
cov.mat = matrix(c(1,0.8,0.8,1), nrow = 2)
e = mvrnorm(1000, mu = c(0,0), Sigma = cov.mat)
cov(e)
##           [,1]      [,2]
## [1,] 0.9685823 0.7429475
## [2,] 0.7429475 0.9381322
e1 = e[, 1]; e2 = e[, 2]
ccf(e1,e2, main = "")
```



Prewhitening

- One simple use of bivariate or multivariate white noise is in the method of **prewhitening**.
- Separate SARIMA models are fitted to multiple time series variables so that the residuals of the fitted models appear to be a realisation of multivariate white noise.
- The SARIMA models can then be used to forecast the expected values of each time series variable, and multivariate simulations can be produced by adding multivariate white noise terms to the forecasts.
- The method works well provided the multiple time series have no common stochastic trends and the cross-correlation structure is restricted to the error process.

Outline

Vector autoregressive models of order 1

- Two time series, $\{X_t\}$ and $\{Y_t\}$, follow a vector autoregressive (VAR) process of order 1 ("VAR(1)") if

$$\begin{aligned} X_t &= \theta_{11} X_{t-1} + \theta_{12} Y_{t-1} + e_{1,t} \\ Y_t &= \theta_{21} X_{t-1} + \theta_{22} Y_{t-1} + e_{2,t} \end{aligned} \tag{1}$$

  where $\{e_{1,t}\}$ and $\{e_{2,t}\}$ are bivariate white noise and θ_{ij} are model parameters.
- Notice that $\{X_t\}$ and $\{Y_t\}$ dynamically interact with each other over time.

Vector autoregressive models of order 1

- The VAR(1) model can be rewritten in matrix notation as

$$\mathbf{Z}_t = \boldsymbol{\theta}\mathbf{Z}_{t-1} + \mathbf{e}_t$$

where

$$\mathbf{Z}_t = \begin{pmatrix} X_t \\ Y_t \end{pmatrix}, \quad \boldsymbol{\Theta}_t = \begin{pmatrix} \theta_{11} & \theta_{12} \\ \theta_{21} & \theta_{22} \end{pmatrix}, \quad \mathbf{e}_t = \begin{pmatrix} e_{1,t} \\ e_{2,t} \end{pmatrix}$$

- Using the backward shift operator, it can also be written

$$(\mathbf{I} - \boldsymbol{\Theta}B)\mathbf{Z}_t = \boldsymbol{\theta}(B)\mathbf{Z}_t = \mathbf{e}_t$$

where $\boldsymbol{\theta}(B)$ is a matrix polynomial of order 1.

- A VAR(1) process can be extended to a VAR(p) process by allowing $\boldsymbol{\theta}(B)$ to be a matrix polynomial of order p.

Stationary VAR models

- Analogous to AR models, a VAR(p) model is *stationary* if the roots of the determinant $|\boldsymbol{\theta}(x)|$ all exceed unity in absolute value.
- For model (1), the determinant is given by

$$\begin{vmatrix} 1-\theta_{11}x & -\theta_{12}x \\ -\theta_{21}x & 1-\theta_{22}x \end{vmatrix} = (1-\theta_{11}x)(1-\theta_{22}x) - \theta_{12}\theta_{21}x^2$$

- For example, if $\boldsymbol{\Theta} = \begin{pmatrix} 0.4 & 0.3 \\ 0.2 & 0.1 \end{pmatrix}$, the characteristic function is $1 - 0.5x - 0.02x^2 = 0$ and the solution is given by

```
Mod( polyroot(c(1,-0.5,-0.02)) )
## [1]  1.861407 26.861407
```

VAR(1) examples

- The parameters of a VAR(p) model can be estimated using the `VAR` function in `library(vars)`, which selects a best-fitting order p based on the smallest AIC.
- Using the simulated bivariate white noise process and the parameters from the stationary VAR(1) model given above, a VAR(1) process is simulated as follows.

```
x <- y <- rep(0, 1000)
ex = rnorm(1000); ey = rnorm(1000)
x[1] <- ex[1]
y[1] <- ey[1]
for (i in 2:1000) {
  x[i] <- 0.4 * x[i-1] + 0.3 * y[i-1] + ex[i]
  y[i] <- 0.2 * x[i-1] + 0.1 * y[i-1] + ey[i]
}
```

VAR(1) examples

```
library(vars)
VAR(cbind(x,y), p=1)
##
## VAR Estimation Results:
## =======================
##
## Estimated coefficients for equation x:
## ======================================
## Call:
## x = x.l1 + y.l1 + const
##
##       x.l1       y.l1      const
## 0.39341866 0.32766426 0.00500737
##
##
## Estimated coefficients for equation y:
## ======================================
## Call:
## y = x.l1 + y.l1 + const
##
##        x.l1        y.l1       const
##  0.23427141  0.05330654 -0.01460736
```

Forecasting VAR(1) models

- Forecasts are generated from a VAR in a recursive manner. The VAR generates forecasts for each variable included in the system.
- To illustrate the process, assume that we have fitted the 2-dimensional VAR(1) described above, for all observations up to time t.
- Then the one-step-ahead forecasts are generated by

$$\hat{x}_{t+1|t} = \hat{\theta}_{11} x_t + \hat{\theta}_{12} y_t$$
$$\hat{y}_{t+1|t} = \hat{\theta}_{21} x_t + \hat{\theta}_{22} y_t$$

- This is the same form as model (1), except that the errors have been set to zero and parameters have been replaced with their estimates.

Forecasting VAR(1) models

- The two-step-ahead forecasts are generated by

$$\hat{x}_{t+2|t} = \hat{\theta}_{11} x_{t+1|t} + \hat{\theta}_{12} y_{t+1|t}$$
$$\hat{y}_{t+2|t} = \hat{\theta}_{21} x_{t+1|t} + \hat{\theta}_{22} y_{t+1|t}$$

- There are two decisions one has to make when using a VAR to forecast, namely how many variables (denoted by K) and how many lags (denoted by p) should be included in the system.
- The number of coefficients to be estimated in a VAR is equal to $K + pK^2$.
- In practice, it is usual to keep K small and include only variables that are correlated with each other, and therefore useful in forecasting each other.

Vector autoregressive models

- Information criteria (such as AIC, BIC) are commonly used to select the number of lags to be included.
- Care should be taken when using the AIC as it tends to choose large numbers of lags. Instead, for VAR models, we prefer to use the BIC.
- A criticism that VARs face is that they are atheoretical; that is, they are not built on some statistical theory that imposes a theoretical structure on the equations.
- Every variable is assumed to influence every other variable in the system, which makes a direct interpretation of the estimated coefficients difficult.

Vector autoregressive models

Despite this, VARs are useful in several contexts:

- forecasting a collection of related variables where no explicit interpretation is required;
- testing whether one variable is useful in forecasting another (the basis of Granger causality tests);
- impulse response analysis, where the response of one variable to a sudden but temporary change in another variable is analysed;
- forecast error variance decomposition, where the proportion of the forecast variance of each variable is attributed to the effects of the other variables.

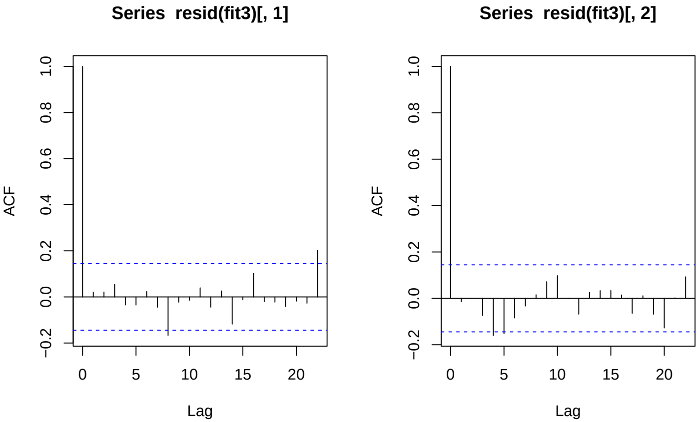
VAR for forecasting US consumption

```r
library(forecast)
library(fpp2)
VARselect(uschange[,1:2], lag.max=8, type="const")[["selection"]]
## AIC(n)  HQ(n)  SC(n) FPE(n)
##      5      1      1      5

fit1 = VAR(uschange[,1:2], p=1, type="const")
fit2 = VAR(uschange[,1:2], p=2, type="const")
fit3 = VAR(uschange[,1:2], p=3, type="const")
serial.test(fit3, lags.pt=10, type="PT.asymptotic")
##
##  Portmanteau Test (asymptotic)
##
## data:  Residuals of VAR object fit3
## Chi-squared = 33.617, df = 28, p-value = 0.2138
```

VAR for forecasting US consumption

```r
par(mfrow=c(1,2))
acf(resid(fit3)[,1])
acf(resid(fit3)[,2])
```



VAR for forecasting US consumption

```
autoplot( forecast(fit3, h=10) )
```