Lecture Note of Bus 41202, Spring 2009: Multivariate Time Series Analysis with Applications

Focus on two series (i.e., bivariate case)

Time series:

$$\boldsymbol{X}_t = \begin{bmatrix} x_{1t} \\ x_{2t} \end{bmatrix}.$$

Data: $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_T$.

Some examples: (a) U.S. quarterly GDP and unemloyment rate series; (b) The daily closing prices of real estate ETF (iShares Dow Jones and Vangard REIT).

Why consider two series jointly?

(a) Obtain the relationship between the series and (b) improve the accuracy of forecasts (use more information).

Some background:

Weak stationarity:

$$E(\boldsymbol{X}_t) = \boldsymbol{\mu}, \quad \text{Cov}(\boldsymbol{X}_t, \boldsymbol{X}_{t-j}) = \boldsymbol{\Gamma}_j$$

are time-invariant

Autocovariane matrix: Lag-ℓ

$$\boldsymbol{\Gamma}_\ell = E[(\boldsymbol{X}_t - \boldsymbol{\mu})(\boldsymbol{X}_{t-\ell} - \boldsymbol{\mu})']$$

$$= \begin{bmatrix} E(x_{1t} - \mu_1)(x_{1,t-\ell} - \mu_1) & E(x_{1t} - \mu_1)(x_{2,t-\ell} - \mu_2) \\ E(x_{2t} - \mu_2)(x_{1,t-\ell} - \mu_1) & E(x_{2t} - \mu_2)(x_{2,t-\ell} - \mu_2) \end{bmatrix}$$

$$= \begin{bmatrix} \Gamma_{11}(\ell) & \Gamma_{12}(\ell) \\ \Gamma_{21}(\ell) & \Gamma_{22}(\ell) \end{bmatrix}.$$

Not symmetric if $\ell \neq 0$. Consider $\mathbf{\Gamma}_1$:

- $\Gamma_{12}(1) = \text{Cov}(x_{1t}, x_{2,t-1})$
- $\Gamma_{21}(1) = \text{Cov}(x_{2t}, x_{1,t-1})$

Let the diagonal matrix $\boldsymbol{D}$ be

$$\boldsymbol{D} = \begin{bmatrix} \text{std}(x_{1t}) & 0 \\ 0 & \text{std}(x_{2t}) \end{bmatrix} = \begin{bmatrix} \sqrt{\Gamma_{11}(0)} & 0 \\ 0 & \sqrt{\Gamma_{22}(0)} \end{bmatrix}.$$

Cross-Correlation matrix:

$$\boldsymbol{\rho}_\ell = \boldsymbol{D}^{-1}\mathbf{\Gamma}_\ell \boldsymbol{D}^{-1}$$

Thus, $\rho_{ij}(\ell)$ is the cross-correlation between x_{it} and $x_{j,t-\ell}$.
From stationarity:

$$\mathbf{\Gamma}_\ell = \mathbf{\Gamma}'_{-\ell}, \quad \boldsymbol{\rho}_\ell = \boldsymbol{\rho}'_{-\ell}.$$

For instance, $\text{cor}(x_{1t}, x_{2,t-1}) = \text{cor}(x_{2t}, x_{1,t+1})$.

Testing for serial dependence

Multivariate version of Ljung-Box $Q(m)$ statistics available.
$H_o : \boldsymbol{\rho}_1 = \cdots = \boldsymbol{\rho}_m = \mathbf{0}$ vs $H_a : \boldsymbol{\rho}_i \neq \mathbf{0}$ for some i

$$Q_2(m) = T^2 \sum_{\ell=1}^{m} \frac{1}{T-\ell} tr(\hat{\mathbf{\Gamma}}'_\ell \hat{\mathbf{\Gamma}}_0^{-1} \hat{\mathbf{\Gamma}}_\ell \hat{\mathbf{\Gamma}}_0^{-1})$$

which is $\chi^2_{k^2 m}$. Note tr is the sum of diagonal elements.

Remark: A **R** script to compute multivariate Q-statistics is available on the course web. The command is **mq** after sourcing the file "mq.R".

Demonstration: Consider the quarterly series of U.S. GDP and unemployment data

```
> x=read.table("q-gdpun.txt",header=T)
> dim(x)
[1] 228   5
> x[1,]
  year mon day    gdp  unemp
1 1948   1   1 7.3878 3.7333
> z=x[,4:5]
> source("mq.R")

> mq(z,10)
[1] "m,       Q(m) and  p-value:"
[1]   1.0000 434.0739   0.0000
[1]   2.0000 827.5327   0.0000
[1]    3.000 1176.616    0.000
[1]    4.000 1486.840    0.000
[1]    5.000 1767.619    0.000
[1]    6.000 2026.774    0.000
[1]    7.000 2268.947    0.000
[1]    8.000 2496.995    0.000
```

```
[1]     9.000 2713.950    0.000
[1]    10.000 2921.077    0.000

> dz=cbind(diff(z[,1]),diff(z[,2]))
> mq(dz,10)
[1] "m,        Q(m) and  p-value:"
[1]    1.0000 105.3880   0.0000
[1]    2.0000 153.2457   0.0000
[1]    3.0000 176.7565   0.0000
[1]    4.0000 196.1902   0.0000
[1]    5.0000 207.9687   0.0000
[1]    6.0000 212.5574   0.0000
[1]    7.0000 215.8745   0.0000
[1]    8.0000 221.8316   0.0000
[1]    9.0000 225.8715   0.0000
[1]   10.0000 228.1209   0.0000
```

The results show that the bivariate series is strongly serially correlated.

Vector Autoregressive Models(VAR)

VAR(1) model for two return series:

$$\begin{bmatrix} r_{1t} \\ r_{2t} \end{bmatrix} = \begin{bmatrix} \phi_{10} \\ \phi_{20} \end{bmatrix} + \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{bmatrix} r_{1,t-1} \\ r_{2,t-1} \end{bmatrix} + \begin{bmatrix} a_{1,t} \\ a_{2,t} \end{bmatrix},$$

where $\boldsymbol{a}_t = (a_{1t}, a_{2t})'$ is a sequence of iid bivariate normal random vectors with mean zero and covariance matrix

$$\text{Cov}(\boldsymbol{a}_t) = \boldsymbol{\Sigma} = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix}$$

where $\sigma_{12} = \sigma_{21}$.

Rewrite the model as

$$\begin{aligned} r_{1t} &= \phi_{10} + \phi_{11} r_{1,t-1} + \phi_{12} r_{2,t-1} + a_{1t} \\ r_{2t} &= \phi_{20} + \phi_{21} r_{1,t-1} + \phi_{22} r_{2,t-1} + a_{1t} \end{aligned}$$

Thus, ϕ_{11} and ϕ_{12} denotes the dependence of r_{1t} on the past returns $r_{1,t-1}$ and $r_{2,t-1}$, respectively.

Unidirectional dependence

For the VAR(1) model, if $\phi_{12} = 0$, but $\phi_{21} \neq 0$, then

- r_{1t} does not depend on $r_{2,t-1}$, but
- r_{2t} depends on $r_{1,t-1}$,

implying that knowning $r_{1,t-1}$ is helpful in predicting r_{2t}, but $r_{2,t-1}$ is not helpful in forecasting r_{1t}.

$\{r_{1t}\}$ is an *input*, $\{r_{2t}\}$ is the *output* variable.

A **Granger** causality relation.

If $\sigma_{12} = 0$, then r_{1t} and r_{2t} are not concurrently correlated.

Stationarity condition

Write the VAR(1) model as

$$\boldsymbol{r}_t = \boldsymbol{\phi}_0 + \boldsymbol{\Phi} \boldsymbol{r}_{t-1} + \boldsymbol{a}_t.$$

$\{\boldsymbol{r}_t\}$ is stationary if zeros of the polynomial

$$|\boldsymbol{I} - \boldsymbol{\Phi} x|$$

are greater than 1 in modulus.

(A generalization of univariate case)

Mean of $\boldsymbol{r}_t$ satisfies

$$(\boldsymbol{I} - \boldsymbol{\Phi})\boldsymbol{\mu} = \boldsymbol{\phi}_0, \quad \text{or}$$

$$\boldsymbol{\mu} = (\boldsymbol{I} - \boldsymbol{\Phi})^{-1}\boldsymbol{\phi}_0$$

if the inverse exists.

Covariance matrices of VAR(1) models:

$$\text{Cov}(\boldsymbol{r}_t) = \sum_{i=0}^{\infty} \boldsymbol{\Phi}^i \boldsymbol{\Sigma} (\boldsymbol{\Phi}^i)',$$

so that

$$\boldsymbol{\Gamma}_\ell = \boldsymbol{\Phi} \boldsymbol{\Gamma}_{\ell-1}$$

for $\ell > 0$.

Can be generalized to higher order models.

Building VAR models

- Order selection: use AIC or BIC (page 356) or a stepwise χ^2 test Eq. (8.18)

  For instance, test VAR(1) vs VAR(2).

- Estimation: use ordinary least squares method

- Model checking: similar to the univariate case
- Forecasting: similar to the univariate case

Simple AR models are sufficient to model asset returns.

Co-integration

Basic ideas

- x_{1t} and x_{2t} are unit-root nonstationary
- a linear combination of x_{1t} and x_{2t} is unit-root stationary

That is, x_{1t} and x_{2t} share a single unit root!

Why is it of interest?
Stationary series is *mean reverting*.
Long term forecasts of the "linear" combination converge to a mean value, implying that the long-term forecasts of x_{1t} and x_{2t} must be linearly related.
This mean-reverting property has many applications. For instance, pairs trading in finance.

Example. Consider the exchange-traded funds (ETF) of U.S. Real Estate. We focus on the iShares Dow Jones (IYR) and Vanguard REIT fund (VNQ) from October 2004 to May 2007. The daily adjusted prices of the two funds are shown in Figure 1. What can be said about the two prices? Is there any arbitrage opportunity between the two funds?

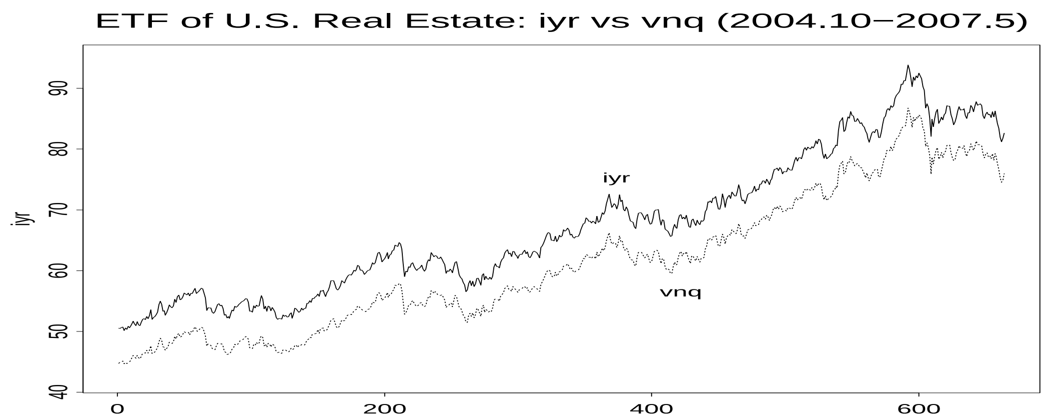

Figure 1: Daily prices of IYR and VNQ from October 2004 to May 2007

The two series all have a unit root (based on ADF test). Are they co-integrated?

Co-integration test

Several tests available, e.g. Johansen's test (Johansen, 1988).

Basic idea

Consider a univariaet AR(2) model

$$x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + a_t.$$

Let $\Delta x_t = x_t - x_{t-1}$.

Substract x_{t-1} from both sides and rearrange terms

$$\Delta x_t = \gamma x_{t-1} + \phi_1^* \Delta x_{t-1} + a_t,$$

where $\phi_1^* = -\phi_2$ and $\gamma = \phi_2 + \phi_1 - 1$.
x_t is unit-root nonstationary if and only if $\gamma = 0$.
Testing for x_t has a unit root is equivalent to testing for $\gamma = 0$ in the above model.
The idea applies to general AR(p) models.

VAR(p) case:

$$\boldsymbol{X}_t = \boldsymbol{\Phi}_1 \boldsymbol{X}_{t-1} + \cdots + \boldsymbol{\Phi}_p \boldsymbol{X}_{t-p} + \boldsymbol{a}_t.$$

Let $\boldsymbol{Y}_t = \boldsymbol{X}_t - \boldsymbol{X}_{t-1}$.
Rewrite the model as

$$\boldsymbol{Y}_t = \boldsymbol{\Omega} \boldsymbol{X}_{t-1} + \sum_{i=1}^{p-1} \boldsymbol{\Phi}_i^* \boldsymbol{Y}_{t-i} + \boldsymbol{a}_t, \tag{1}$$

where

$$\begin{aligned}
\boldsymbol{\Phi}_{p-1}^* &= -\boldsymbol{\Phi}_p \\
\boldsymbol{\Phi}_{p-2}^* &= -\boldsymbol{\Phi}_{p-1} - \boldsymbol{\Phi}_p \\
\vdots &= \vdots \\
\boldsymbol{\Phi}_1^* &= -\boldsymbol{\Phi}_2 - \cdots - \boldsymbol{\Phi}_p \\
\boldsymbol{\Omega} &= \boldsymbol{\Phi}_p + \cdots + \boldsymbol{\Phi}_1 - \boldsymbol{I}.
\end{aligned}$$

This is the *Error-Correction* Model (ECM).

The **Key** concept related to pairs trading is that $\boldsymbol{Y}_t$ is related to $\boldsymbol{\Omega} \boldsymbol{X}_{t-1}$.

To test for co-integration:

- Fit the model in Eq. (1),
- Test for the rank of $\boldsymbol{\Omega}$.

If $\boldsymbol{X}_t$ is k dimensional, and rank of $\boldsymbol{\Omega}$ is m, then we have $k-m$ unit roots in $\boldsymbol{X}_t$.
There are m linear combinations of $\boldsymbol{X}_t$ that are unit-root stationary.

If $\boldsymbol{\Omega}$ has rank m, then

$$\boldsymbol{\Omega} = \boldsymbol{\alpha}\boldsymbol{\beta}$$

where $\boldsymbol{\alpha}$ is a $k \times m$ and $\boldsymbol{\beta}$ is a $m \times k$ full-rank matrix.
$\boldsymbol{Z}_t = \boldsymbol{\beta}\boldsymbol{X}_t$ is unit-root stationary.
$\boldsymbol{\beta}$ is the co-integrating vector.

Discussion

- ECM formulation is useful
- Co-integration tests have some weaknesses, e.g. robustness
- Co-integration overlooks the effect of scale of the series

Pairs trading

Reference: *Pairs Trading: Quantitative Methods and Analysis* by Ganapathy Vidyamurthy, Wiley, 2004.
Motivation: General idea of trading is to sell overvalued securities and buy undervalued ones. But the *true* value of the security is hard to determine in practice. Pairs trading attempts to resolve this

difficulty by using *relative pricing*. Basically, if two securities have similar characteristics, then the prices of both securities must be more or less the same. Here the true price is not important.

Statistical term: The prices behave like random-walk types of processes, but a linear combination of them is stationary, hence, the linear combination is mean-reversting. Deviations from the mean lead to trading opportunities.

Theory in Finance: **Arbitrage Pricing Theory** (APT): If two securities have exactly the same risk factor exposures, then the expected returns of the two securities for a given time period are the same. [The key here is that the returns must be the same for all times.]

More details: Consider two stocks: Stock 1 and Stock 2. Let p_{it} be the log price of Stock i at time t. It is reasonable to assume that the time series $\{p_{1t}\}$ and $\{p_{2t}\}$ contain a unit root when they are analyzed individually.

Assume that the two log-price series are co-integrated, that is, there exists a linear combination $c_1 p_{1t} - c_2 p_{2t}$ that is stationary. Dividing the linear combination by c_1, we have

$$w_t = p_{1t} - \gamma p_{2t},$$

which is stationary. The stationarity implies that w_t is mean-reverting. Now, form the portfolio Z by buying 1 share of Stock 1 and selling short on γ shares of Stock 2. The return of the portfolio for a given period h is

$$\begin{aligned} r(h) &= (p_{1,t+h} - p_{1,t}) - \gamma(p_{2,t+h} - p_{2,t}) \\ &= p_{1,t+h} - \gamma p_{2,t+h} - (p_{1,t} - \gamma p_{2,t}) \\ &= w_{t+h} - w_t \end{aligned}$$

which is the increment of the stationary series $\{w_t\}$ from t to $t+h$. Since w_t is stationary, we have obtained a direct link of the portfolio to a stationary time series whose forecasts we can predict.
Assume that $E(w_t) = \mu$. Select a threshold δ.

A trading strategy:

- Buy Stock 1 and short γ shares of Stock 2 when the $w_t = \mu - \delta$.
- Unwind the position, i.e. sell Stock 1 and buy γ shares of Stock 2, when $w_{t+h} = \mu + \delta$.

Profit: $r(h) = w_{t+h} - w_t = 2\delta$.

Some practical considerations:

- The threshold δ is chosen so that the profit outweights the costs of two tradings. In high frequency, δ must be greater than *trading slippage*, which is the same linear combination of bid-ask spreads of the two stock, i.e. bid-ask spread of Stock 1 + $\gamma \times$(bid-ask spread) of Stock 2.

- Speed of mean-reverting of w_t plays an important role as h is directly related to the speed of mean-reverting.
- There are many ways available to search for co-integrating pairs of stocks. For example, via fundamentals, risk factors, etc.
- For unit-root and co-integration tests, see the textbook and references therein.

Example: Consider the daily adjusted closing sotck prices of BHP Billiton Limited of Australia and Companhia Vale de Rio Doce of Brazil. These are two natural resouces companies. Both stocks are also listed in the New York Stock Exchange with tick symbols BHP and RIO, respectively. The sample period is from March 21, 2002 to May 30, 2008.

- How to estimate γ?
- Speed of mean reverting? (zero-crossing concept)