Financial Time Series Lecture 10: Analysis of Multiple Financial Time Series with Applications

Reference: Chapters 8 and 10 of the textbook.
We shall focus on two series (i.e., the 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 unemployment rate series; (b) The daily closing prices of oil related ETFs, e.g. oil services holdings (OIH) and energy select section SPDR (XLE); and, for more than 2 series, (c) quarterly GDP grow rates of Canada, United Kingdom, and United States.

Why consider two series jointly?
(a) Obtain the relationship between the series and (b) improve the accuracy of forecasts (use more information). See Figure 1 for the log prices of the two energy funds. The prices seem to move in unison.

Some background:
Weak stationarity: Both

$$\begin{aligned} E(\boldsymbol{X}_t) &= \begin{bmatrix} E(x_{1t}) \\ E(x_{2t}) \end{bmatrix} = \boldsymbol{\mu}, \quad \text{and} \\ \text{Cov}(\boldsymbol{X}_t, \boldsymbol{X}_{t-j}) &= \begin{bmatrix} \text{Cov}(x_{1t}, x_{1,t-\ell}) & \text{Cov}(x_{1t}, x_{2,t-\ell}) \\ \text{Cov}(x_{2t}, x_{1,t-\ell}) & \text{Cov}(x_{2t}, x_{2,t-\ell}) \end{bmatrix} = \boldsymbol{\Gamma}_j \end{aligned}$$

are time invariant

Auto-covariance 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 $\boldsymbol{\Gamma}_1$:

- $\Gamma_{12}(1) = \mathrm{Cov}(x_{1t}, x_{2,t-1})$ (x_{1t} depends on past x_{2t})
- $\Gamma_{21}(1) = \mathrm{Cov}(x_{2t}, x_{1,t-1})$ (x_{2t} depends on past x_{1t})

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

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

Cross-Correlation matrix:

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

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

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

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

Testing for serial dependence

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

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

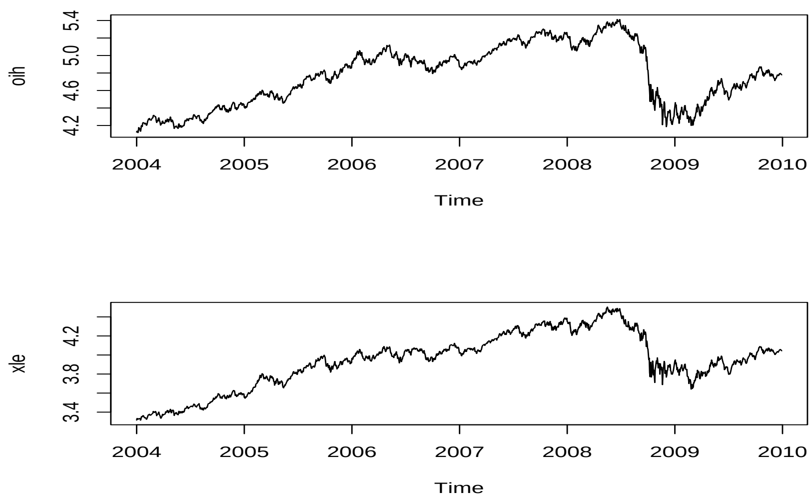


Figure 1: Daily log prices of OIH and XLE funds from January 2004 to December 2009

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

Remark: Analysis of multiple financial time series can be carried out in R via the package `MTS`. Some useful commands are (a) MTSplot, which drawns multiple time series plot(b) ccm, which compute the cross-correlation matrices and Ljung-Box statistics, and (c) mq, which compute the Ljung-Box statistics.

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

```
> require(MTS)
> 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]
> MTSplot(z)

> 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=diffM(z)  ### Take difference of individual series
> 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 knowing $r_{1,t-1}$ is helpful in predicting r_{2t}, but $r_{2,t-1}$ is not helpful in forecasting r_{1t}.
Here $\{r_{1t}\}$ is an *input*, $\{r_{2t}\}$ is the *output* variable. This is an example of **Granger** causality relation.

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

Stationarity condition: Generalization of 1-dimensional case
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. Equivalently, if solutions of $|\boldsymbol{I} - \boldsymbol{\Phi} x| = 0$ are all greater than 1 in modulus.

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 or a stepwise χ^2 test Eq. (8.18). See Section 8.2.4, pp 405-406.

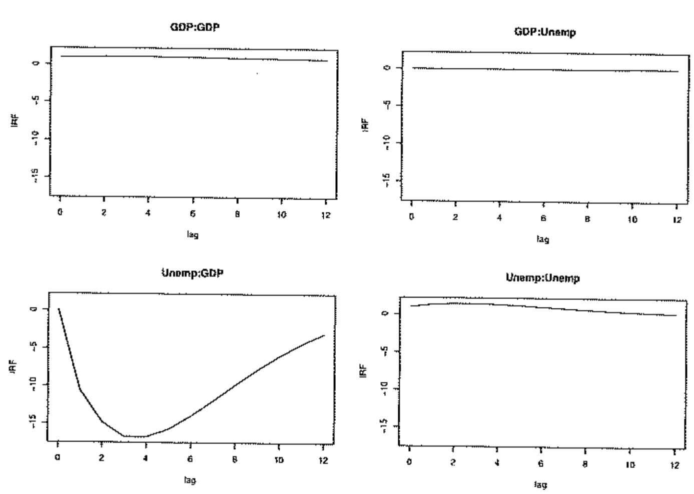
  For instance, test VAR(1) vs VAR(2).

- 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.
Program note: Commands for VAR modeling

- VARorder: compute various information criteria for a vector time series
- VAR: estimate a VAR model
- refVAR: refine an estimated VAR model by fixing insignificant estimates to zero

```
#########################################################################
### Analyze quarterly GDP & unemployment data (Lec 10).
### Illustrates model fitting, checking, prediction, and IRFs.
#########################################################################
require(MTS)

### read in data
x=read.table("Datasets/q-gdpun.txt",header=T)
z=x[,4:5]

### plot the bivariate series
MTSplot(z)

### find optimal VAR order
VARorder(z, maxp = 13)
#selected order: aic =  4
#selected order: bic =  2
#selected order: hq =  3

### Fit VAR(2) and then refine it by omitting coeffs with t-ratio<1.96,
m1=VAR(z, p = 2, output = T, include.mean = T)
m2=refVAR(m1,thres=1.96)

### Check LB goodness of fit.
### Must adjust df by setting: adj=p*k^2 for a k-dim VAR(p).
### In general: adj is the number of non-zero VAR coeffts.
### Here: p=2=k implies adj=2^3=8
### All p-values high, so not a good fit...
MTSdiag(m2, adj= 7)

### Forecast 3-steps ahead.
VARpred(m2, h=3)

### Get irf's and plot them manually (canned plot hard to control...)
### Summary: 1 unit change in GDP negatively affects Unemp 3-4 quarters later.
out=VARMAirf(m2$Phi,m2$Sigma)
psi11=out$irf[1,]; psi21=out$irf[2,]; psi12=out$irf[3,]; psi22=out$irf[4,]
#pdf(file="Plots/IBM-acf-pacf.pdf", pointsize=12, paper="a4r",width=0,height=0)
par(mfcol=c(2,2))
x=seq(0, length(psi11)-1); miny=min(out$irf); maxy=max(out$irf)
plot(x,psi11, type="l", ylab="IRF", xlab="lag", main="GDP:GDP", ylim=c(miny,maxy))
plot(x,psi21, type="l", ylab="IRF", xlab="lag", main="Unemp:GDP", ylim=c(miny,maxy))
plot(x,psi12, type="l", ylab="IRF", xlab="lag", main="GDP:Unemp", ylim=c(miny,maxy))
plot(x,psi22, type="l", ylab="IRF", xlab="lag", main="Unemp:Unemp", ylim=c(miny,maxy))
dev.off()
#########################################################################
```



Lec 10 - page $6\frac{1}{2}$

- MTSdiag: model checking
- VARpred: predict a fitted VAR model.

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 2. What can be said about the two prices? Is there any arbitrage opportunity between the two funds?
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).

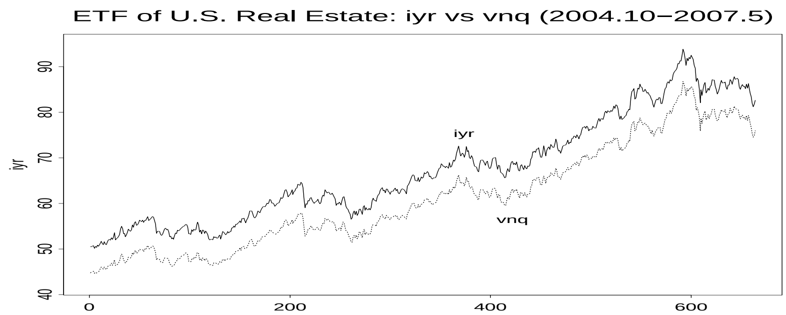


Figure 2: Daily prices of IYR and VNQ from October 2004 to May 2007

Basic idea

Consider a univariate AR(2) model

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

Let $\Delta x_t = x_t - x_{t-1}$.
Subtract x_{t-1} from both sides and rearrange terms to obtain

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

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

Turn to the VAR(p) case. The original model is

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

Let $\boldsymbol{Y}_t = \boldsymbol{X}_t - \boldsymbol{X}_{t-1}$.
Subtracting $\boldsymbol{X}_{t-1}$ from both sides and re-grouping of the coefficient matrices, we can rewrite the model as

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

where

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

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

Important message: The matrix $\boldsymbol{\Pi}$ is a zero matrix if there is no co-integration.

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

To test for co-integration:

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

If $\boldsymbol{X}_t$ is k dimensional, and rank of $\boldsymbol{\Pi}$ 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 $\boldsymbol{\Pi}$ has rank m, then

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

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

Package: The package `urca` of **R** can be used to perform co-integration test.

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 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 out-weights the costs of two trading. 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 stock prices of BHP Billiton Limited of Australia and Vale S.A. of Brazil. These are two natural resources companies. Both stocks are also listed in the New York Stock Exchange with tick symbols BHP and Vale, respectively. The sample period is from July 1, 2002 to March 31, 2006.

- How to estimate γ?

- Speed of mean reverting? (zero-crossing concept)

```
> require(urca)
> help(ca.jo)  # Johansen's co-integration test
```

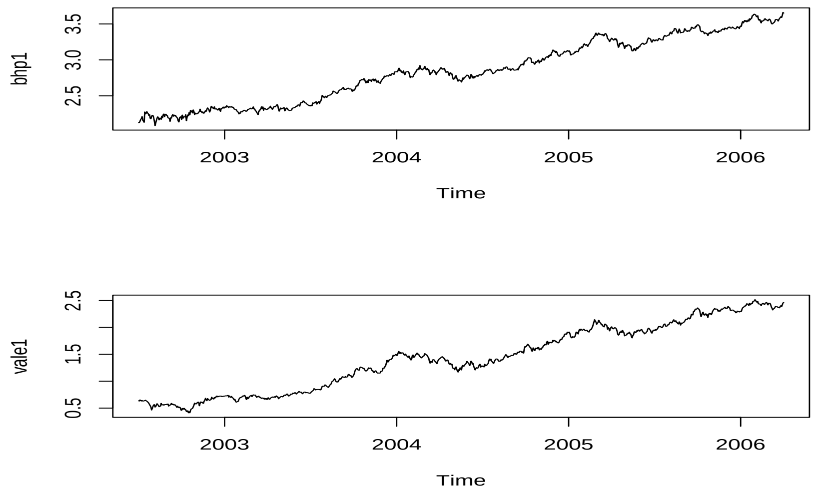


Figure 3: Daily log prices of BHP and VALE from July 1, 2002 to March 31, 2006.

```
> da=read.table("d-bhp0206.txt",header=T)
> da1=read.table("d-vale0206.txt",header=T)
> head(da)
  Mon day year  open  high   low close volume adjclose
1   7   1 2002 11.80 11.92 11.55 11.60 156700     8.39
....
6   7   9 2002 12.25 12.65 12.25 12.60 142000     9.12
> head(da1)
  Mon day year  open  high   low close  volume adjclose
1   7   1 2002 27.60 27.60 27.10 27.16 2307600     1.89
....
6   7   9 2002 27.05 27.55 27.05 27.30 2534400     1.90
> tail(da1)
    Mon day year  open  high   low close   volume adjclose
941   3  24 2006 44.90 45.52 44.45 45.28 15496800    10.94
 .....
946   3  31 2006 47.83 48.64 47.51 48.53 10900000    11.73
> tail(da)
    Mon day year  open  high   low close  volume adjclose
941   3  24 2006 37.35 37.75 37.12 37.42 2251200    36.17
 ....
946   3  31 2006 39.62 40.19 39.22 39.85 3045900    38.52
> dim(da)
[1] 946   9
> bhp=log(da[,9])
> vale=log(da1[,9])
```

```
> plot(bhp,type='l')
> plot(vale,type='l')
> m1=lm(bhp~vale)
> summary(m1)

Call: lm(formula = bhp ~ vale)

Residuals:
      Min        1Q    Median        3Q       Max
-0.151818 -0.028265  0.003121  0.029803  0.147105

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 1.822648   0.003662   497.7   <2e-16 ***
vale        0.716664   0.002354   304.4   <2e-16 ***
---
Residual standard error: 0.04421 on 944 degrees of freedom
Multiple R-squared: 0.9899,     Adjusted R-squared: 0.9899
F-statistic: 9.266e+04 on 1 and 944 DF,  p-value: < 2.2e-16

> bhp1=ts(bhp,frequency=252,start=c(2002,127))
> vale1=ts(vale,frequency=252,start=c(2002,127))
> plot(bhp1,type='l')
> plot(vale1,type='l')
> x=cbind(bhp,vale)
> m1=ar(x)
> m1$order
[1] 2
> m2=ca.jo(x,K=2)
> summary(m2)
######################
# Johansen-Procedure #
######################
Test type: maximal eigenvalue statistic (lambda max) , with linear trend

Eigenvalues (lambda):
[1] 0.0406019854 0.0000101517

Values of teststatistic and critical values of test:

          test 10pct  5pct  1pct
r <= 1 |  0.01  6.50  8.18 11.65
r = 0  | 39.13 12.91 14.90 19.19

Eigenvectors, normalised to first column:
(These are the cointegration relations)
```

```
           bhp.l2  vale.l2
bhp.l2   1.000000 1.000000
vale.l2 -0.717784 2.668019

Weights W:
(This is the loading matrix)

            bhp.l2       vale.l2
bhp.d  -0.06272119 -2.179372e-05
vale.d  0.03303036 -3.274248e-05

> m3=ca.jo(x,K=2,type=c("trace"))
> summary(m3)
######################
# Johansen-Procedure #
######################
Test type: trace statistic , with linear trend

Eigenvalues (lambda):
[1] 0.0406019854 0.0000101517

Values of teststatistic and critical values of test:

          test 10pct  5pct  1pct
r <= 1 |  0.01  6.50  8.18 11.65
r = 0  | 39.14 15.66 17.95 23.52

Eigenvectors, normalised to first column:
(These are the cointegration relations)

           bhp.l2  vale.l2
bhp.l2   1.000000 1.000000
vale.l2 -0.717784 2.668019

Weights W:
(This is the loading matrix)

            bhp.l2       vale.l2
bhp.d  -0.06272119 -2.179372e-05
vale.d  0.03303036 -3.274248e-05

> wt=bhp-0.718*vale
> acf(wt)
> pacf(wt)
> m4=arima(wt,order=c(2,0,0))
```

```
> m4

Call:
arima(x = wt, order = c(2, 0, 0))

Coefficients:
         ar1     ar2  intercept
      0.8050  0.1215      1.820
s.e.  0.0323  0.0325      0.008

sigma^2 estimated as 0.000333:  log likelihood = 2444.26,  aic = -4880.52
> tsdiag(m4)
> plot(wt,type='l')
```

```
### VECM ftn from package tsDYn estimates & predicts VECM:
### lags=p-1 from the VAR(p), r=# coint relations, include = one of (none, const,
trend).
> m5 = VECM(x, lag=1, r=1, estim="ML", include ="const")
> summary(m5)
#############
###Model VECM
#############
Full sample size: 946   End sample size: 944
Number of variables: 2  Number of estimated slope parameters 8
AIC -14875.77    BIC -14832.12    SSR 0.8188267
Cointegrating vector (estimated by ML):
   bhp       vale
r1   1 -0.717784

              ECT                  Intercept            bhp -1
Equation bhp  -0.0627(0.0146)***   0.1159(0.0266)***    -0.1149(0.0367)**
Equation vale 0.0330(0.0169).      -0.0584(0.0308).     0.0528(0.0425)
              vale -1
Equation bhp  0.0692(0.0320)*
Equation vale 0.0452(0.0371)

### predict fitted VECM 3 steps ahead
> predict(m5, n.ahead=3)
         bhp      vale
947 3.650391 2.465768
948 3.648675 2.469617
949 3.647362 2.473280
```

Fitted VECM is: $x_t = \begin{bmatrix} bhp_t \\ vale_t \end{bmatrix}$, $\Delta x_t = y_t = x_t - x_{t-1}$

$$\Delta x_t = \mu_0 + \alpha\beta' x_{t-1} + \Phi_1^* \Delta x_{t-1} + a_t$$

$$\alpha = \begin{bmatrix} -.0627 \\ .0330 \end{bmatrix}, \quad \beta = \begin{bmatrix} 1 \\ -.7179 \end{bmatrix}, \quad \mu_0 = \begin{bmatrix} .1159 \\ -.0584 \end{bmatrix}, \quad \Phi_1^* = \begin{bmatrix} -.1149 & .0692 \\ .0528 & .0452 \end{bmatrix}$$

$w_t = \beta' x_t = (1, -.7179)\, x_t \leftarrow$ cointegrating relation.

```
> m4

Call:
arima(x = wt, order = c(2, 0, 0))

Coefficients:
         ar1     ar2  intercept
      0.8050  0.1215      1.820
s.e.  0.0323  0.0325      0.008

sigma^2 estimated as 0.000333:  log likelihood = 2444.26,  aic = -4880.52
> tsdiag(m4)
> plot(wt,type='l')
```

Multivariate Volatility Models

How do the correlations between asset returns change over time?

Focus on two series (Bivariate)
Two asset return series:

$$\boldsymbol{r}_t = \begin{bmatrix} r_{1t} \\ r_{2t} \end{bmatrix}.$$

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

Basic concept

Let F_{t-1} denote the information available at time $t-1$.
Partition the return as

$$\boldsymbol{r}_t = \boldsymbol{\mu}_t + \boldsymbol{a}_t, \quad \boldsymbol{a}_t = \boldsymbol{\Sigma}_t^{1/2}\boldsymbol{\epsilon}_t$$

where $\boldsymbol{\mu}_t = E(\boldsymbol{r}_t|F_{t-1})$ is the predictable component, and

$$\text{Cov}(\boldsymbol{a}_t|F_{t-1}) = \boldsymbol{\Sigma}_t = \begin{bmatrix} \sigma_{11,t} & \sigma_{12,t} \\ \sigma_{21,t} & \sigma_{22,t} \end{bmatrix},$$

$\{\boldsymbol{\epsilon}_t\}$ are iid 2-dimensional random vectors with mean zero and identity covariance matrix.

Multivariate volatility modeling

See Chapter 10 of the textbook
Study time evolution of $\{\boldsymbol{\Sigma}_t\}$.

$\boldsymbol{\Sigma}_t$ is symmetric, i.e. $\sigma_{12,t} = \sigma_{21,t}$
There are 3 variables in $\boldsymbol{\Sigma}_t$.
For k asset returns, $\boldsymbol{\Sigma}_t$ has $k(k+1)/2$ variables.

Requirement

$\boldsymbol{\Sigma}_t$ must be positive definite for all t,

$$\sigma_{11,t} > 0, \quad \sigma_{22,t} > 0, \quad \sigma_{11,t}\sigma_{22,t} - \sigma_{12,t}^2 > 0.$$

The time-varying correlation between r_{1t} and r_{2t} is

$$\rho_{12,t} = \frac{\sigma_{12,t}}{\sqrt{\sigma_{11,t}\sigma_{22,t}}}.$$

Some complications

- Positiveness requirement is not easy to meet
- Too many series to consider

Some simple models available

- Exponentially weighted covariance
- Use univariate approach, e.g. $\text{Cov}(X,Y) = \frac{\text{Var}(X+Y) - \text{Var}(X-Y)}{4}$.
- BEKK model

- Dynamic conditional correlation (DCC) models

Exponentially weighted model

$$\boldsymbol{\Sigma}_t = (1-\lambda)\boldsymbol{a}_{t-1}\boldsymbol{a}'_{t-1} + \lambda\boldsymbol{\Sigma}_{t-1},$$

where $0 < \lambda < 1$. That is,

$$\boldsymbol{\Sigma}_t = (1-\lambda)\sum_{i=1}^{\infty}\lambda^{i-1}\boldsymbol{a}_{t-i}\boldsymbol{a}'_{t-i}.$$

R command `EWMAvol` of the `MTS` package can be used.

BEKK model of Engle and Kroner (1995)
Simple BEKK(1,1) model

$$\boldsymbol{\Sigma}_t = \boldsymbol{A}_0\boldsymbol{A}'_0 + \boldsymbol{A}_1(\boldsymbol{a}_{t-1}\boldsymbol{a}'_{t-1})\boldsymbol{A}'_1 + \boldsymbol{B}_1\boldsymbol{\Sigma}_{t-1}\boldsymbol{B}'_1$$

where $\boldsymbol{A}_0$ is a lower triangular matrix, $\boldsymbol{A}_1$ and $\boldsymbol{B}_1$ are square matrices without restrictions.
Pros: positive definite
Cons: Many parameters, dynamic relations require further study
Estimation: `BEKK11` command in `MTS` package can be used for $k =$ 2 and 3 only.

DCC mdoels: A two-step process

- Marginal models: Use univariate volatility model for individual return series
- Use DCC model for the time-evolution of conditional correlation

Specifically, the volatility matrix can be written as

$$\boldsymbol{\Sigma}_t = \boldsymbol{V}_t\boldsymbol{R}_t\boldsymbol{V}_t,$$

where $\boldsymbol{V}_t$ is a diagonal matrix of volatilities for individual return series and $\boldsymbol{R}_t$ is the conditional correlation matrix. That is,

$$\boldsymbol{V}_t = \text{diag}\{v_{1t}, v_{2t}, \ldots, v_{kt}\} \quad \boldsymbol{R}_t = [\rho_{ij,t}]$$

where $\rho_{ij,t}$ is the correlation between ith and jth return series.
Two types of DCC are available in the literature

1. Engle (2002):

$$\boldsymbol{Q}_t = (1-\theta_1-\theta_2)\boldsymbol{R}_0 + \theta_1\boldsymbol{Q}_{t-1} + \theta_2\boldsymbol{a}_{t-1}\boldsymbol{a}'_{t-1},$$
$$\boldsymbol{R}_t = \boldsymbol{q}_t^{-1}\boldsymbol{Q}_t\boldsymbol{q}_t^{-1},$$

   where $0 \le \theta_i$ and $\theta_1+\theta_2 < 1$, $\boldsymbol{q}_t = \text{diag}\{\sqrt{Q_{11,t}}, \sqrt{Q_{22,t}}, \ldots, \sqrt{Q}_{kk,t}\}$ and $\boldsymbol{R}_0$ is the sample correlation matrix.

2. Tse and Tsui (2002):

$$\boldsymbol{R}_t = (1-\theta_1-\theta_2)\boldsymbol{R}_0 + \theta_1\boldsymbol{R}_{t-1} + \theta_2\boldsymbol{\psi}_{t-1},$$

   where $0 \le \theta_i$ and $\theta_1+\theta_2 < 1$, and $\boldsymbol{\psi}_{t-1}$ is the sample correlation matrix of $\{\boldsymbol{a}_{t-1}, \boldsymbol{a}_{t-2}, \ldots, \boldsymbol{a}_{t-m}\}$ for a pre-specified positive integer m, e.g. $m = 3$.

Discussion

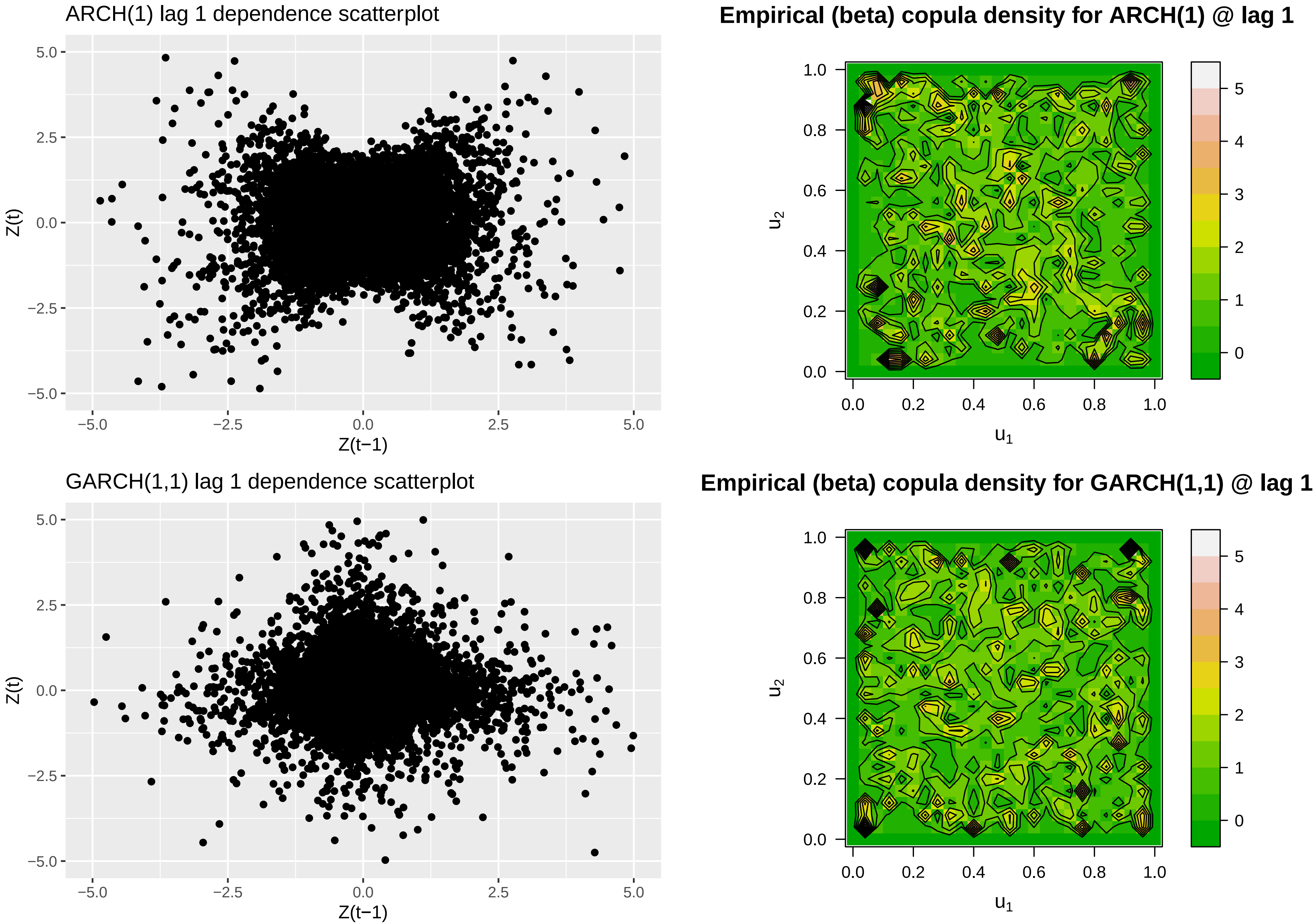
1. DCC model is extremely simple with two parameters
2. On the other hand, model checking tends to reject the DCC models.

`R` commands of the `MTS` package for DCC modeling:

1. dccPre: fit individual GARCH models (standardized return series is included in the output)

2. dccFit: estimate a DCC model for the standardized return series
3. MCHdiag: model checking of multivariate volatility models.

A demonstration is given below, taken from Tsay's "Multivariate Time Series Analysis (Wiley, 2013).


ARCH(1) lag 1 dependence scatterplot
Z(t)
Z(t−1)
5.0
2.5
0.0
−2.5
−5.0
Empirical (beta) copula density for ARCH(1) @ lag 1
u_2
u_1
0.0
0.2
0.4
0.6
0.8
1.0
0
1
2
3
4
5
GARCH(1,1) lag 1 dependence scatterplot
Z(t)
Z(t−1)
Empirical (beta) copula density for GARCH(1,1) @ lag 1
u_2
u_1

Ex 7.5 (Tsay, 2014)

Monthly log rtns of (IBM, S&P Index, Coca-Cola).
Fit t-Copula.

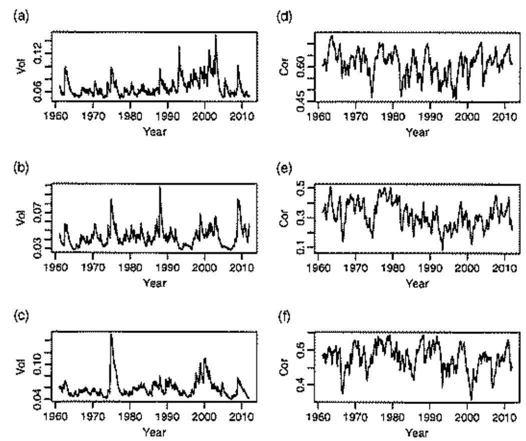


FIGURE 7.15 Time plots of the volatilities and time-varying correlations of the monthly log returns of (a) IBM, (b) S&P index, and (c) Coca Cola stock from January 1961 to December 2011 based on a t-copula with Equation (7.55). The left panel consists of volatility series and the right panel the correlations. (d) IBM versus S&P, (e) IBM versus KO, and (f) S&P versus KO.

Compared with the correlations of the DCC models in Figure 7.9, we see that the correlations of the t-copula model in Figure 7.15 are between those of the DCC models of Tse–Tsui and Engle. The time-evolution patterns of the correlations of t-copula model are closer to those of the DCC model of Tse–Tsui (2002), but with higher magnitudes. The sample means of the three correlation coefficients of the t-copula model are 0.613, 0.319, and 0.555, respectively, whereas the sample standard errors are 0.057, 0.086, and 0.072. The minimum and maximum of the three correlations are 0.460, 0.076, 0.315 and 0.738, 0.509, and 0.687, respectively. These values are similar to those of the DCC models in Table 7.4.

Finally, we also fit the t-copula model to the three monthly log return series, but fix the initial angle θ_0 based on the sample correlation matrix of $\hat{\eta}_t$. In this particular case, $\hat{\theta}_0 = (0.946, 1.289, 1.035)'$, which is close to what we obtained before via joint estimation. The estimates of λ_i become 0.878 and 0.034, respectively, whereas the estimate of the degrees of freedom is $\hat{v} = 14.87$. It seems that, for this particular example, fixing $\hat{\theta}_0$ does not have a major effect on the estimates of other parameters. □

Remark: The estimation of the t-copula models is based on the command `mtCopula` of the MTS package. The marginal volatility models of the component series are estimated by the command `dccPre`, which uses the `fGarch` package of Rmetrics. The command `mtCopula` provides two sets of asymptotic standard errors of the estimates based on two different numerical differentiation methods. □

R Demonstration: Estimation of t-copula models.

Fit std t univariate GARCH to each series.

```
> da=read.table("m-ibmspko-6111.txt",header=T)
> rtn=log(da[,-1]+1)
> m1=dccPre(rtn,cond.dist="std")
Sample mean of the returns:  0.00772774 0.005023909 0.01059521
Component:  1
Estimates:  0.000388 0.115626 0.805129 9.209269
se.coef  :  0.000177 0.036827 0.059471 3.054817
t-value  :  2.195398 3.139719 13.5382 3.014671
Component:  2
Estimates:  0.00012 0.130898 0.814531 7.274928
se.coef  :  5.7e-05 0.037012 0.046044 1.913331
t-value  :  2.102768 3.536655 17.69028 3.802232
Component:  3
Estimates:  0.000216 0.104706 0.837217 7.077138
se.coef  :  8.9e-05 0.028107 0.037157 1.847528
t-value  :  2.437323 3.725341 22.53208 3.830599
> names(m1)
[1] "marVol"  "sresi"   "est"     "se.coef"
> Vol=m1$marVol; eta=m1$sresi
> m2=mtCopula(eta,0.8,0.04)
Lower limits:  5.1 0.2 1e-04 0.7564334 1.031269 0.8276595
Upper limits:  20 0.95 0.04999999 1.040096 1.417994 1.138032
estimates:   15.38215 0.88189 0.034025 0.919724 1.225322 1.058445
std.errors:  8.222771 0.05117 0.011733 0.041357 0.055476 0.051849
t-values:    1.870677 17.2341 2.899996 22.23883 22.08729 20.41412
Alternative numerical estimates of se:
st.errors:  5.477764 0.051033 0.011714 0.041370 0.055293 0.050793
t-values:   2.808107 17.28091 2.904679 22.23173 22.16072 20.83839
> names(m2)
[1] "estimates" "Hessian"   "rho.t"     "theta.t"
> MCHdiag(eta,m2$rho.t)
Test results:
Q(m) of et:
Test and p-value:  19.30177 0.03659304
Rank-based test:
Test and p-value:  27.03262 0.002573576
Qk(m) of epsilon_t:
Test and p-value:  125.9746 0.007387423
Robust Qk(m):
Test and p-value:  107.4675 0.1011374
> m3=mtCopula(eta,0.8,0.04,include.th0=F)  # fix theta_0
Value of angles:
[1] 0.9455418 1.2890858 1.0345744
Lower limits:  5.1 0.2 1e-05
```

t-Copula fit (θ_0 parameter estimated) Need θ_1 & θ_2 estimates (used $\hat{\theta}_1 = .8$ and $\hat{\theta}_2 = .04$ from previous)

t-Copula fit (θ_0 parameter specified a priori)

Ex (§7.7.2, Tsay, 2014)

Monthly log rtns of (IBM, S&P Index, Coca-Cola)
Fit multivariate t.

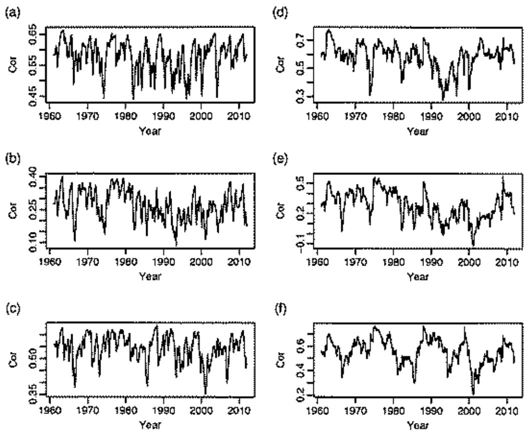


FIGURE 7.9 Time plots of time-varying correlation coefficients for monthly log returns of IBM, S&P index, and KO from January 1961 to December 2011 based on the DCC models in Equations (7.38) and (7.39). The plots on the left panel are for the model in Equation (7.38). (a) IBM versus SP, (b) IBM versus KO, (c) SP versus KO, (d) IBM versus SP, (e) IBM versus KO, and (f) SP versus KO.

TABLE 7.4 Summary Statistics of Time-Varying Correlations of the Monthly Log Returns of IBM, S&P, and KO Based on DCC Models in Equations (7.38) and (7.39)

	Tse and Tsui Model			Engle Model		
Statistics	(ibm,sp)	(ibm,ko)	(sp,ko)	(ibm,sp)	(ibm,ko)	(sp,ko)
Mean	0.580	0.269	0.544	0.585	0.281	0.556
SE	0.050	0.067	0.055	0.095	0.139	0.105
Min	0.439	0.086	0.352	0.277	−0.121	0.205
Max	0.664	0.403	0.638	0.773	0.568	0.764

R Demonstration: Estimation of DCC models.

```
> da=read.table("m-ibmspko-6111.txt",header=T)
> rtn=log(da[,2:4]+1)
> m1=dccPre(rtn,include.mean=T,p=0)
```

fits univariate GARCH to each component series (with default being normal innovations).


```
Sample mean of the returns:  0.00772774 0.005023909 0.01059521
Component:  1
Estimates:  0.000419 0.126739 0.788307
se.coef  :  0.000162 0.035405 0.055645
t-value  :  2.593448 3.57973 14.16662
Component:  2
Estimates:  9e-05 0.127725 0.836053
se.coef  :  4.1e-05 0.03084 0.031723
t-value  :  2.20126 4.141592 26.35486
Component:  3
Estimates:  0.000256 0.098705 0.830358
se.coef  :  8.5e-05 0.022361 0.033441
t-value  :  3.015321 4.414112 24.83088
> names(m1)
[1] "marVol"  "sresi"    "est"     "se.coef"
> rtn1=m1$sresi
> Vol=m1$marVol
> m2=dccFit(rtn1)
Estimates:  0.8088086 0.04027318 7.959013
st.errors:  0.1491655 0.02259863 1.135882
t-values:   5.422222 1.782107 7.006898
> names(m2)
[1] "estimates" "Hessian"   "rho.t"
> S2.t = m2$rho.t
> m3=dccFit(rtn1,type="Engle")
Estimates:  0.9126634 0.04530917 8.623668
st.errors:  0.0294762 0.01273911 1.332381
t-values:   30.96272 3.556697 6.472376
> S3.t=m3$rho.t
> MCHdiag(rtn1,S2.t)
Test results:
Q(m) of et:
Test and p-value:  20.74262 0.02296152
Rank-based test:
Test and p-value:  30.20662 0.0007924436
Qk(m) of epsilon_t:
Test and p-value:  132.423 0.002425885
Robust Qk(m):
Test and p-value:  109.9671 0.0750157
> MCHdiag(rtn1,S3.t)
Test results:
Q(m) of et:
Test and p-value:  20.02958 0.02897411
Rank-based test:
Test and p-value:  27.61638 0.002078829
Qk(m) of epsilon_t:
Test and p-value:  131.982 0.002625755
Robust Qk(m):
Test and p-value:  111.353 0.06307334
```

Fit DCC model (default is Tse & Tsui)
$\hat{\theta}_1 = .8088$, $\hat{\theta}_2 = .0403$
$\hat{\nu} = 7.96$

DCC model of Engle

goodness of fit measures reject all models (except the robust measure)