8. Multivariate Linear Time Series.

MA6622, Ernesto Mordecki, CityU, HK, 2006.

References for this Lecture:

Introduction to Time Series and Forecasting. P.J. Brockwell and R. A. Davis, Springer Texts in Statistics (2002)

Analysis of Financial Time Series (Chapter 8). Ruey S. Tsay. Wiley (2002) [Available Online]

Main Purpose of Lectures 8 and 9:

Model the time evolution of a portfolio contanining d assets, with returns

$$\mathbf{X}(0), \mathbf{X}(1), \ldots, \mathbf{X}(n)$$

where

$$\mathbf{X}(t) = \begin{bmatrix} X_1(t) \\ \vdots \\ X_d(t) \end{bmatrix} \qquad (t = 0, \ldots, n)$$

through a multivariate linear time series model.

Plan of Lecture 8

- Report stylized facts in multivariate financial time series in other words, how these series look like, from a statistical point of view.

- Introduce some necessary facts from matrices and multivariate statistics.

- Testing normality (key issue in finance)

- Testing multivariate normality, i.e. whether we can assume that a sample of multivariate data (vectorial data) can be assumed to be normaly distributed.

- Introduce the concepts of Multivariate Time Series

8a. Stylized facts in multivariate financial time series

Empirical observations on daily returns of financial time series led to the following 4 stylized facts, widely understood to be empirical truths, to which theories must fit.

(M1) Multivariate Return series show little evidence of cross-correlation, except for concurrent (i.e. contemporaneous) returns.

The cross-covariance, or covariance $\mathbf{cov}(\mathbf{X}(s), \mathbf{X}(t)')$ for $s \neq t$, is generally negligible, as in the one dimension case. When $t = s$ and $i \neq j$ the (concurrent) covariances $\mathbf{cov}\left(X_i(t), X_j(t)\right)$ can be non negligible due to factors afecting the whole market.

(M2) Multivariate series of absolute returns show profound serial correlations for different times (cross-correlation).

As in the one dimensional case, large movements in one stock tend to be followed by large movements in this stocks, and also in other stocks of the same market. As previously, financial time series are uncorrelated but not independent.

(M3) The covariance structure of concurrent returns vary over time.

Consistenly with the same phenomena of volatility time variation in the univariate case, and with the previous phenomena of clustering of large returns, it seems that the covariance of $\mathbf{X}(t)$ vary with t. (This raises the question of modelling this phenomena, for instance with mutivariate GARCH processes).

(M4) Extreme returns in one asset often conicide with extreme returns in several other asset.

This fact asserts that in high volatility periods of the market, assets seem to be more correlated, and has as limit statement that "correlations go to one in times of market stress".

8b. Elements of multivariate statistics

Given a random vector

$$\mathbf{X} = \begin{bmatrix} X_1 \\ \vdots \\ X_d \end{bmatrix}$$

we denote its traspose by

$$\mathbf{X}' = (X_1, \ldots, X_d)$$

Given two vectors **X** and **Y**

- the product $\mathbf{XY}'$ is a matrix:

$$\mathbf{XY}' = \begin{bmatrix} X_1 \\ \vdots \\ X_d \end{bmatrix} (Y_1, \ldots, Y_d) = \begin{bmatrix} X_1Y_1 & \ldots & X_1Y_d \\ \vdots & \vdots & \vdots \\ X_dY_1 & \ldots & X_dY_d \end{bmatrix}$$

- while $\mathbf{X}'\mathbf{Y}$ is a number:

$$(X_1, \ldots, X_d) \begin{bmatrix} X_1 \\ \vdots \\ X_d \end{bmatrix} = X_1 Y_1 + \cdots + X_n Y_n.$$

The expectation of the random vector $\mathbf{X}$ is

$$\mathbf{E}\,\mathbf{X} = (\mathbf{E}\,X_1, \ldots, \mathbf{E}\,X_d)',$$

the variance-covariance matrix of $\mathbf{X}$ is

$$\Sigma = \mathbf{cov}(\mathbf{X}) = [\mathbf{cov}(X_i, X_j)]_{i,j=1,\ldots,d}.$$

The correlation matrix is

$$\rho(\mathbf{X}) = \Big[\frac{\mathbf{cov}(X_i, X_j)}{\sqrt{\mathbf{var}(X_i)\,\mathbf{var}(X_j)}}\Big]_{i,j=1,\ldots,d}$$

Given a matrix $\mathbf{A}$ and a vector $\mathbf{b}$ we have:

$$\mathbf{E}(\mathbf{AX}+\mathbf{b}) = \mathbf{A}\,\mathbf{E}\,\mathbf{X}+\mathbf{b}, \qquad \mathbf{cov}(\mathbf{AX}+\mathbf{b}) = \mathbf{A}\,\mathbf{cov}(\mathbf{X})\mathbf{A}'.$$

Definition The vector $\mathbf{Z} = (Z_1, \ldots, Z_d)'$ is a gaussian or normal standard vector when $Z_1, \ldots, Z_d$ are independent standard normal random variables.

For a standard normal vector

- $\mathbf{E}\,\mathbf{Z} = (0, \ldots, 0)'$,
- $\mathbf{cov}(\mathbf{Z}) = \mathbf{I}_d$, the $d \times d$ identity matrix.

A gaussian or normal vector $\mathbf{Z}$ with mean μ and covariance Σ is obtained as

$$\mathbf{X} = \mu + \mathbf{AZ},$$

where the matrix $\mathbf{A}$ satisfies $\mathbf{AA}' = \Sigma$

- $\mathbf{A}$ is called a square root of Σ,
- Given a covariance matrix Σ, its squared root $\mathbf{A}$ always exist (linear algebra).

We denote

$$\mathbf{X} \sim \mathcal{N}_d(\mu, \Sigma).$$

Given $\mathbf{X} \sim \mathcal{N}_d(\mu, \Sigma)$

- $\mathbf{E}\,\mathbf{X} = \mu + \mathbf{A}\,\mathbf{E}\,\mathbf{Z} = \mu.$
- For the variance-covariances matrix:

$$\mathbf{cov}(\mathbf{X}) = \mathbf{cov}(\mathbf{AZ}) = \mathbf{A}\,\mathbf{cov}(\mathbf{Z})\mathbf{A}' = \mathbf{AI}_d\mathbf{A}' = \mathbf{AA}' = \Sigma.$$

8c. Testing Univariate Normality.

Many results in statistics of time series are built on the hyphotesis of gaussian returns (for instance, the Black-Scholes model).

It is then important to determine whether a sample of univariate returns are gaussian

Quantile-Quantile Plot (QQ - Plot)

Is a visual test for univariate gaussianity.

Suppouse you want to know if the following sample of 9 values

$$0.22, 2.29, 2.06, 7.32, 7.05, 0.14, 7.51, 9.15, 4.21$$

can be considered sampled from a normal random variable.

In order to perform the QQ-Plot test

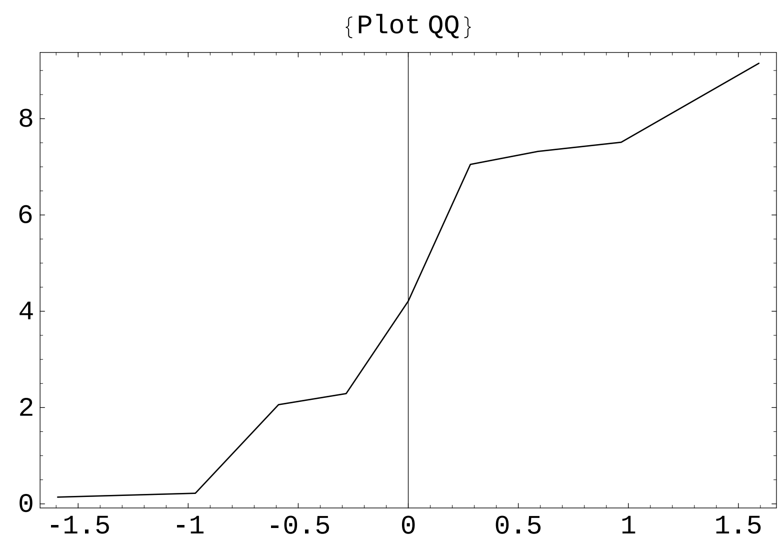
STEP 1. Order the sample in increasing order:

$$0.14, 0.22, 2.06, 2.29, 4.21, 7.05, 7.32, 7.51, 9.15$$

STEP 2. Compute $x(i) = \Phi\big((i - 1/2)/9\big)^{-1} (i = 1, \ldots, 9)$ from the normal standard table and prepare the table:

$x(i)$	-1.593	-0.967	-0.589	-0.282	0	0.282	0.589	0.967	1.593
$y(i)$	0.14	0.22	2.06	2.29	4.21	7.05	7.32	7.51	9.15

STEP 3. Plot the points $(x(i), y(i))$ for $i = 1, \ldots, 9$:



If the ploted points fit approximately a straight line, the gaussian hypothesis is not rejected.

Skewness-Kurtosis Jarque-Bera test.

A second practical procedure is to simultaneously test whether the third and fourth centered moments corresponds to that of a normal random variable. Given a sample $X(1), \ldots, X(n)$

STEP 1. Estimate the mean and the variance by

$$\bar{X} = \frac{1}{n}\sum_{k=1}^{n} X(k), \qquad \bar{\sigma}^2 = \frac{1}{n}\sum_{k=1}^{n}\left(X(k) - \bar{X}\right)^2$$

STEP 2. Compute the empirical skewness and kurtosis, by

$$\bar{\gamma} = \frac{\frac{1}{n}\sum_{k=1}^{n}\left(X(k) - \bar{X}\right)^3}{\bar{\sigma}^3}, \qquad \bar{\kappa} = \frac{\frac{1}{n}\sum_{k=1}^{n}\left(X(k) - \bar{X}\right)^4}{\bar{\sigma}^4} - 3$$

STEP 3. Compute the Jarque-Bera statistic as

$$Q_{JB} = n\left(\frac{1}{6}\bar{\gamma}^2 + \frac{1}{24}\bar{\kappa}^2\right) \sim \chi^2_2,$$

that has, for big values of n, a Chi-square distribution with two degrees of freedom

STEP 4. Big values of Q_{JB} indicate that the skewness and/or the kurtosis do not vanish (as should happen under normality).

Then (with a 95% confidence) if

$$Q_{JB} > 5.99 = t_{2,0.95},$$

reject the hyphotesis of normality

Remark This test is valid for big values of n. For small values we prefer the QQ-plot.

8d. Testing Multivariate Normality.

First remark: It is not enough to have normality for the coordinates of a vector (univariate marginals) in order to have a normal vector.

Suppose we want to test whether a given a vectorial sample $\mathbf{X}(1), \ldots, \mathbf{X}(n)$ is normal.

QQ Chi Square Plot

Based on the fact that, given a vector $\mathbf{X} \sim \mathcal{N}_d(\mu, \Sigma)$, the random variable

$$(\mathbf{X} - \mu)'\Sigma^{-1}(\mathbf{X} - \mu) \sim \chi^2_d,$$

we construct a new sample.

Compute the sample mean

$$\bar{\mathbf{X}} = \frac{1}{n}\sum_{k=1}^{n} \mathbf{X}(k),$$

and the sample covariance matrix

$$\bar{\Sigma} = \frac{1}{n} \sum_{k=1}^{n} \Big(\mathbf{X}(k) - \bar{\mathbf{X}}\Big)' \Big(\mathbf{X}(k) - \bar{\mathbf{X}}\Big).$$

Invert the matrix $\bar{\Sigma}$, and construct a new univariate sample of the form

$$D_k^2 = \Big(\mathbf{X}(k) - \bar{\mathbf{X}}\Big)' \bar{\Sigma}^{-1} \Big(\mathbf{X}(k) - \bar{\mathbf{X}}\Big), \qquad k = 1, \ldots, n.$$

For big values of n behaves rhoghly like a sample of χ_d^2 independent random variables.

Then, test this hypothesis with a QQ-plot.

The procedure is the same as in the univariate gaussian case, with the diffence that

$$x(i) = F_{\chi_2^2}\big((i - 1/2)/n\big)$$

where $F_{\chi_2^2}$ is the χ_2^2 distribution (i.e. one should use the Chi square distribution with 2 degrees of freedom instead of the normal table).

Skewness-Kurtosis Multivariate Test

We can also test for multivariate skewness and kurtosis.

Compute

$$D_{jk} = \Big(\mathbf{X}(j) - \bar{\mathbf{X}}\Big)' \bar{\Sigma}^{-1} \Big(\mathbf{X}(k) - \bar{\mathbf{X}}\Big), \qquad j, k = 1, \ldots, n,$$

The statistics

$$\gamma_d = \frac{1}{n^2} \sum_{i,j=1}^{n} D_{ij}^3, \qquad \kappa_d = \frac{1}{n} \sum_{i,j=1}^{n} D_i^4 - d(d+2)$$

have, when the random vector sample is normal, asymptotics distributions

$$\frac{1}{6} n\, \gamma_d \sim \chi^2_{d(d+1)(d+2)/6}, \qquad \frac{\kappa_d}{\sqrt{8d(d+2)/n}} \sim \mathcal{N}(0,1).$$

We construct then two tests based on these facts.

8e. Multivariate time series

Let $\mathbf{X}(0), \ldots, \mathbf{X}(n)$ be the stochastic returns of a portfolio with d assets, where $\mathbf{X}(t) = \big(X_1(t), \ldots, X_d(t)\big)'$,

For simplicity of exposition we assume that we have two assets A and B, and our returns are

$$\mathbf{X}(t) = \begin{bmatrix} X_A(t) \\ X_B(t) \end{bmatrix}.$$

The vector of expectations is

$$\mu(t) = \begin{bmatrix} \mu_A(t) \\ \mu_B(t) \end{bmatrix} = \begin{bmatrix} \mathbf{E}\, X_A(t) \\ \mathbf{E}\, X_B(t) \end{bmatrix},$$

and the covariance matrix

$$\Gamma(t+h, t) = \begin{bmatrix} \mathbf{cov}(X_A(t+h), X_A(t)) & \mathbf{cov}(X_A(t+h), X_B(t)) \\ \mathbf{cov}(X_B(t+h), X_A(t)) & \mathbf{cov}(X_B(t+h), X_B(t)) \end{bmatrix}$$

Definition We say that the bivariate series is weakly stationary when the expectations $\mu(t) \equiv \mu$ does not depend on t, and also the the covariance matrix $\Gamma(t+h,t) = \Gamma(h)$ does not depend on t.

We have

$$\begin{aligned}\Gamma(h) &= \begin{bmatrix}\Gamma_{AA}(h) & \Gamma_{AB}(h) \\ \Gamma_{BA}(h) & \Gamma_{BB}(h)\end{bmatrix} \\ &= \begin{bmatrix}\mathbf{cov}(X_A(h), X_A(0)) & \mathbf{cov}(X_A(h), X_B(0)) \\ \mathbf{cov}(X_B(h), X_A(0)) & \mathbf{cov}(X_B(h), X_B(0))\end{bmatrix}.\end{aligned}$$

Here

- The diagonal terms are the covariances of the univariate series $\{X_A(t)\}$ and $\{X_B(t)\}$.

- New information appears in the off-diagonal elements, that are the covariance between different assets.
- We call $\mathbf{cov}(X_A(0), X_B(0))$ a concurrent covariance: same date for different assets,
- We call $\mathbf{cov}(X_A(0), X_B(h))$ a cross coviance: different dates and different assets.

Observe that in general:

- $\mathbf{cov}(X_B(h), X_A(0)) \neq \mathbf{cov}(X_A(h), X_B(0))$, so the matrix $\Gamma(h)$ is not symmetric.
- We have

$$\mathbf{cov}(X_A(h), X_B(0)) = \mathbf{cov}(X_B(-h), X_A(0)),$$

what simplifies the estimation.

Defining

$$\rho_{ij}(h) = \frac{\Gamma_{ij}(h)}{\sqrt{\Gamma_{ii}(0)\Gamma_{jj}(0)}}, \quad \text{for } i, j = A, B$$

we construct the correlation matrix as

$$\mathbf{R}(h) = \begin{bmatrix} \rho_{AA}(h) & \rho_{AB}(h) \\ \rho_{BA}(h) & \rho_{BB}(h) \end{bmatrix}$$