9. Multivariate Linear Time Series (II).

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 containining 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 9

(9a) Plotting the Cross Correlogram of a multivariate time series

(9b) Introduce stationary multivariate time series and white noises.

(9c) Vectorial ARMA processes (VARMA), in particular VAR(1).

(9d) Testing multivariate white noise.

(9e) Comments on Co-integration

9a. Plotting the Empirical Cross-correlogram

The cross correlogram of our vectorial time series is a 2×2 matrix of correlograms.

Correlogram of series 1	Cross-correlogram of series 1,2
Cross-correlogram of series 2,1	Correlogram of series 2

In order to construct the cross-correlogram, we perform:

STEP 1. We compute the sample mean $\bar{\mathbf{X}} = (\bar{X}_A, \bar{X}_B)$ of both series:

$$\bar{X}_A = \frac{1}{n}\sum_{t=1}^{n} X_A(t), \qquad \bar{X}_B = \frac{1}{n}\sum_{t=1}^{n} X_B(t).$$

STEP 2. We compute the sample covariance matrix, for $h = 0, \ldots, n_0$, as

$$\bar{\Gamma}(h) = \frac{1}{n} \sum_{t=1}^{n-h} (\mathbf{X}(t+h) - \bar{\mathbf{X}})(\mathbf{X}(t) - \bar{\mathbf{X}})'$$

Here each term is is a 2×2 matrix:

$$(\mathbf{X}(t+h) - \bar{\mathbf{X}})(\mathbf{X}(t) - \bar{\mathbf{X}})'$$
$$= \begin{bmatrix} (X_A(t+h) - \bar{X}_A)(X_A(t) - \bar{X}_A) & (X_A(t+h) - \bar{X}_A)(X_B(t) - \bar{X}_B) \\ (X_B(t+h) - \bar{X}_B)(X_A(t) - \bar{X}_A) & (X_B(t+h) - \bar{X}_B)(X_B(t) - \bar{X}_B) \end{bmatrix}$$

In particular for $h = 0$ the diagonal of the matrix gives

$$\bar{\sigma}_A^2 = \bar{\Gamma}_{AA}(0), \qquad \bar{\sigma}_B^2 = \bar{\Gamma}_{BB}(0).$$

STEP 3. We compute the correlation matrix for $h = 0, 1, \ldots, n_0$ as

$$\mathbf{R}(h) = \begin{bmatrix} \frac{\bar{\Gamma}_{AA}(h)}{\bar{\sigma}_A^2} & \frac{\bar{\Gamma}_{AB}(h)}{\bar{\sigma}_A \bar{\sigma}_B} \\ \frac{\bar{\Gamma}_{BA}(h)}{\bar{\sigma}_A \bar{\sigma}_B} & \frac{\bar{\Gamma}_{BB}(h)}{\bar{\sigma}_B^2} \end{bmatrix}$$

STEP 4. We plot four graphics ($h = 0, 1, \ldots, n_0$):

$\big(h, \bar{\Gamma}_{AA}(h)\big)$	$\big(h, \bar{\Gamma}_{AB}(h)\big)$
$\big(h, \bar{\Gamma}_{BA}(h)\big)$	$\big(h, \bar{\Gamma}_{BB}(h)\big)$

In this cross-correlograms:

- The diagonal gives information about the individual behaviour of each asset,
- the upper right correlogram shows the correlation of future values of A against present values of B.
- the lower left correlogram shows the correlation of future values of B against present values of A.
- Past values of A against present values of B are the same as future of B against present of A ($\Gamma_{AB}(-h) = \Gamma_{BA}(h)$), it is not necessary to plot them, and
- Past values of B against present values of A are the same as future of A against present of A ($\Gamma_{BA}(-h) = \Gamma_{AB}(h)$).

9b. Stationary Multivariate time series

Definition A multivariate series $\{\mathbf{X}(t)\}$ is

(a) a weak white noise: weakly stationary, $\mathbf{E}\,\mathbf{X}(t) = 0$ for all t and

$$\Gamma(h) = \begin{cases} \Sigma_\varepsilon & \text{when } h = 0 \\ 0 & \text{when } h \neq 0 \end{cases}$$

(b) a strict white noise: i.i.d. random vectors, with $\mathbf{E}\,\mathbf{X}(t) = 0$ and covariance matrix Σ_ε.

(c) a gaussian or normal white noise: strict white noise with distributions $\mathcal{N}(\mathbf{0}, \Sigma_\varepsilon)$

Remark In all cases Σ_ε is a covariance matrix. The values of $\{\mathbf{X}(t)\}$ can have concurrent correlation (same time), but not cross-correlations (different times)

9c. Vectorial ARMA process (VARMA)

Definition $\{\mathbf{X}(t)\}$ is a vectorial ARMA(p,q) process if it is centered weakly stationary and

$$\mathbf{X}(t) - \Phi_1 \mathbf{X}(t-1) - \cdots \Phi_p \mathbf{X}(t-p) \\ = \varepsilon(t) - \Theta_1 \varepsilon(t-1) - \cdots - \Theta_q \varepsilon(t-q)$$

where $\{\varepsilon(t)\}$ is a weak white noise with covariance Σ_ε. Here Φ_i and Θ_i are $d \times d$ matrices.

Example VAR(1) process satisfies

$$\mathbf{X}(t) = \Phi \mathbf{X}(t-1) + \varepsilon(t),$$

with $\{\varepsilon(t)\}$ weak white noise. In order to check stationarity, one should have (instead of $|\phi| < 1$ for $d = 1$) that all the eigenvalues of the matrix Φ are strictly greater than one in absolute value.

In order to estimate the matrix ϕ we must solve a matrix equation (i.e. $d \times d$ linear equations) of the form

$$\bar{\Gamma}(1) = \Phi_1 \bar{\Gamma}(0),$$

that can be solved computing the inverse of the matrix $\bar{\Gamma}(0)$, and post-multiplying both sides of the equation by this inverse matrix we obtain

$$\bar{\Phi}(1) = \bar{\Gamma}(1)\bar{\Gamma}(0)^{-1}.$$

Let us examine two particular cases: $d = 1$ and $d = 2$.

Case $d = 1$

In this case matrices are numbers. As our process is centered

$$\bar{\Gamma}(0) = \bar{\sigma}_X^2 = \frac{1}{n}\sum_{k=1}^{n} X(t)^2$$

$$\bar{\Gamma}(1) = \overline{\mathbf{cov}}(1) = \frac{1}{n}\sum_{k=1}^{n} X(t)X(t-1)$$

giving the estimate

$$\bar{\phi} = \bar{\Gamma}(1)\bar{\Gamma}(0)^{-1} = \frac{\sum_{k=1}^{n} X(t)X(t-1)}{\sum_{k=1}^{n} X(t)^2}$$

from the previous lecture.

Case $d = 2$

Assume then that we have two financial time series A and B.

$$\mathbf{X}(t) = \begin{bmatrix} X_A(t) \\ X_B(t) \end{bmatrix}, \qquad \Phi = \begin{bmatrix} \phi_{AA} & \phi_{AB} \\ \phi_{BA} & \phi_{BB} \end{bmatrix}, \qquad \varepsilon(t) = \begin{bmatrix} \varepsilon_A(t) \\ \varepsilon_B(t) \end{bmatrix}$$

The model matrix model is:

$$\mathbf{X}(t) = \Phi \mathbf{X}(t-1) + \varepsilon(t).$$

In coordinates one has:

$$\begin{aligned} X_A(t) &= \phi_{AA} X_A(t-1) + \phi_{AB} X_B(t-1) + \varepsilon_A(t) \\ X_B(t) &= \phi_{BA} X_A(t-1) + \phi_{BB} X_B(t-1) + \varepsilon_B(t) \end{aligned}$$

In order to estimate Φ we perform:

STEP 1. Compute the (symmetric) sample (contemporaneous) variance-covariance matrix

$$\bar{\Gamma}(0) = \frac{1}{n}\sum_{t=1}^{n}\begin{bmatrix} X_A(t)^2 & X_A(t)X_B(t) \\ X_B(t)X_A(t) & X_B(t)^2 \end{bmatrix}$$

STEP 2. Compute the (non symmetric) sample cross-covariance matrix with lag $h = 1$

$$\bar{\Gamma}(1) = \frac{1}{n}\sum_{t=1}^{n}\begin{bmatrix} X_A(t)X_A(t-1) & X_A(t)X_B(t-1) \\ X_B(t)X_A(t-1) & X_B(t)X_B(t-1) \end{bmatrix}$$

STEP 3. Invert the matrix $\bar{\Gamma}(0)$ to obtain

$$\bar{\phi} = \bar{\Gamma}(1)\bar{\Gamma}(0)^{-1}$$

STEP 4. Estimate the variance-covariance $\bar{\Sigma}_{\varepsilon}$ matrix:

$$\bar{\Sigma}_{\varepsilon} = \bar{\Gamma}(0) - \bar{\phi}\bar{\Gamma}(1).$$

9d. Testing Multivariate White Noise.

Complementing the visual analysis of the cross-correlogram of a bivariate time series $\mathbf{X}(t) = (X_A(t), X_B(t))'$ we have a Multivariate Portmanteau Test, proposed by Hosking (1980) that extendes the Ljung and Box test of Lecture 6. The statistical test is

$$H_0\colon \Gamma(1) = \cdots = \Gamma(n_0) = \mathbf{0}, \qquad (\mathbf{X} \text{ is WN})$$

$$H_a\colon \Gamma(h) \neq \mathbf{0} \text{ for some } h = 1, \ldots, n_0 \qquad (\mathbf{X} \text{ is not WN}).$$

To compute the test statistic, for each lag $h = 1, \ldots, n_0$, we compute

$$q(h) = \mathbf{tr}[\bar{\Gamma}(h)'\bar{\Gamma}(0)^{-1}\bar{\Gamma}(h)\bar{\Gamma}(0)^{-1}],$$

(where **tr** is the trace of the product of four matrices)

The statistic is

$$Q(h) = n^2 \sum_{i=1}^{n_0} \frac{1}{n-i} q(h) \sim \chi^2_{4n_0},$$

If $\mathbf{X}$ is WN, $Q(h)$ is Chi-Squared distribution with $4n_0$ degrees of freedom.

For $n_0 = 10$, big values of $Q(h)$ indicate rejection of H_0:

$$\text{If} \quad Q(10) > t_{40,0.95} = 55.7585 \text{ reject } H_0.$$

Comments

- The $4 = d^2$. If $d = 3$ we have $9n_0$ degrees of freedom
- When $d^2 n_0$ is large, we can use the normal approximation

  $$Q(h) \sim \mathcal{N}(h, 2h)$$

  that, for confidence 0.95 and $h = 40$ gives

  $$t_{h,0.95} \sim h + 1.645\sqrt{2h} = 40 + 1.645\sqrt{80} = 54.7133.$$

9e. Comments on Co-integration

In univariate time series, nonstationarity phenomena is avoided through differentiation.

In case of differentiating once, we say that the process is integrated of order one. If it is necessay to differentiate again the process is integrated of order two, and so on.

In the multivariate case, when considering time series

$$\mathbf{X}(t) = (X_1(t), X_2(t), \ldots, X_d(t))'$$

the same situation can arise, differentiating simultaneouly all the coordinate of the time series, and saying that the multivariate process is integrated of order $1, 2$, etc.

But a new situation arises, as it is possible to take linear combinations $\alpha'\mathbf{X}(t)$ of the concurrent univariate components, in order to fit stationarity.

Consider for instance two highly correlated and nonstationary assets in a portfolio, say, with returns X_1 and X_2. Assuming that the difference $X_2(t) - X_1(t)$ is stationary, we consider the new time series

$$\mathbf{Y}(t) = (X_1(t) - X_2(t), X_2(t), \ldots, X_d(t))$$

whose first component is stationary.

Observe that $\mathbf{Y}(t) = \alpha' \mathbf{X}(t)$ with $\alpha = (1, -1, 0, \ldots, 0)$.

The general definiton of cointegration is the following:

If a multivariate time series $\mathbf{X}$ is integrated of order d, and exists α such that the univariate time series $\alpha' \mathbf{X}$ is integrated of order $d' < d$, we have co-integration.

Although not simple to model and to test, the cointegration phenomena is important in financial modelling due to the fact that frequently assets are highly correlated