

# Multivariate Time Series



**Feng Li**

**feng.li@cufe.edu.cn**

**School of Statistics and Mathematics  
Central University of Finance and Economics**

# Today we are going to learn...

1 Introduction to Multivariate Returns

2 Vector autoregressive models

3 Vector moving average models

4 Vector ARMA models

## Multivariate Returns I

- Let  $\mathbf{r}_t = (r_{1t}, \dots, r_{Nt})$  be the log returns of  $N$  assets at time  $t$ , the multivariate analyses of time series are concerned with the joint distribution of  $\{\mathbf{r}_t\}_{t=1}^T$ .
- The analysis is then focused on the specification of the conditional distribution function  $F(\mathbf{r}_t | \mathbf{r}_{t-1}, \dots, \mathbf{r}_1, \boldsymbol{\theta})$ . In particular, how the conditional expectation and conditional covariance matrix of  $\mathbf{r}_t$  evolve over time

$$\begin{aligned} E(\mathbf{r}_t) &= E(r_{1t}, \dots, r_{Nt})' = (E(r_{1t}), \dots, E(r_{Nt}))' \\ \Gamma_0 = \text{Cov}(\mathbf{r}_t) &= E((\mathbf{r}_t - E(\mathbf{r}_t))(\mathbf{r}_t - E(\mathbf{r}_t))') \end{aligned}$$

where The  $i$ th diagonal element of  $\Gamma_0$  is the **variance** of  $r_{it}$ , whereas the  $(i, j)$ th element of  $\Gamma_0$  is the **covariance** between  $r_{it}$  and  $r_{jt}$

## Multivariate Returns II

- The series  $\mathbf{r}_t$  is **weakly stationary** if its first and second moments are time invariant. In particular, the mean vector and covariance matrix of a weakly stationary series are constant over time.
- Unless stated explicitly to the contrary, we assume that the return series of financial assets are weakly stationary.

## Cross-Correlation Matrices I

- Let  $\mathbf{D}$  be a  $k \times k$  diagonal matrix consisting the standard deviations of  $r_{it}$  for  $i = 1, \dots, k$ . The *concurrent, or lag-zero, cross-correlation matrix* of  $r_t$  is defined as

$$\rho_{ij}(0) = \mathbf{D}^{-1} \mathbf{\Gamma}_0 \mathbf{D}^{-1} = \frac{\text{Cov}(r_{it}, r_{jt})}{\text{std}(r_{it}) \times \text{std}(r_{jt})}$$

- Note that  $\rho_{ij}(0)$  measures the **linear dependence** of  $r_{it}$  and  $r_{jt}$ .
- Note that it is the correlation of the two series at time  $t$ .

## Cross-Correlation Matrices II

- An important topic in multivariate time series analysis is the lead-lag relationships between component series.
- To this end, the **cross correlation matrices** (CCM) are used to measure the strength of linear dependence between time series.
- The lag- $l$  cross-covariance matrix of  $\mathbf{r}_t$  is defined as

$$\Gamma_l = E((\mathbf{r}_t - \boldsymbol{\mu})(\mathbf{r}_{t-l} - \boldsymbol{\mu})')$$
$$\rho_{ij}(l) = D^{-1}\Gamma_l D^{-1} = \frac{\text{Cov}(r_{it}, r_{j,(t-l)})}{\text{std}(r_{it}) \times \text{std}(r_{jt})}$$

- For negative lag  $l$ , we have  $\Gamma_l = \Gamma'_{-l}$ .

## Cross-Correlation Matrices III

- **Sample Cross-Correlation Matrices** Given the data  $\mathbf{r}_t$ , the cross-covariance matrix  $\Gamma_l$  can be estimated by

$$\hat{\Gamma}_l = \frac{1}{T} \sum_{t=l+1}^T (\mathbf{r}_t - \bar{\mathbf{r}})(\mathbf{r}_t - \bar{\mathbf{r}})'$$

where  $\bar{\mathbf{r}} = \sum_{t=1}^T \mathbf{r}_t / T$  is the vector of sample means.

- And the cross-correlation matrix is

$$\hat{\rho}_{ij}(l) = \hat{D}^{-1} \hat{\Gamma}_l \hat{D}^{-1}$$

# Empirical Properties of Returns

- Daily returns of the market indexes and individual stocks tend to have high excess kurtoses. For monthly series, the returns of market indexes have higher excess kurtoses than individual stocks.
- The mean of a daily return series is close to zero, whereas that of a monthly return series is slightly larger.
- Monthly returns have higher standard deviations than daily returns.
- Among the daily returns, market indexes have smaller standard deviations than individual stocks. This is in agreement with common sense.
- The skewness is not a serious problem for both daily and monthly returns.
- The descriptive statistics show that the difference between simple and log returns is not substantial.
- **Example in R**



## VAR(1) model I

- A simple vector model useful in modeling asset returns is the **vector autoregressive (VAR) model**.
- A multivariate time series  $\mathbf{r}_t$  is a VAR process of order 1, or VAR(1) for short, if it follows the model

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

where  $\boldsymbol{\phi}_0$  is a  $k$ -dimensional vector,  $\boldsymbol{\Phi}$  is a  $k \times k$  matrix, and  $\mathbf{a}_t$  is a sequence of serially uncorrelated random vectors with mean zero and positive definite covariance matrix  $\boldsymbol{\Sigma}$ . In the literature, it is often assumed that  $\mathbf{a}_t$  is multivariate normal.

## VAR(1) model II

- The elements of  $\Phi$  gives the conditional effect of the linear dependence between  $r_{it}$  and  $r_{j(t-1)}$ .
- Consider the bivariate case, If  $\Phi_{12} = 0$  and  $\Phi_{21} = 0$ , then there is a unidirectional relationship from  $r_{1t}$  to  $r_{2t}$  . If  $\Phi_{12} = \Phi_{21} = 0$ , then  $r_{1t}$  and  $r_{2t}$  are uncoupled. If  $\Phi_{12} \neq 0$  and  $\Phi_{12} \neq 0$ , then there is a feedback relationship between the two series.
- In general, the coefficient matrix  $\Phi$  measures the dynamic dependence of  $\mathbf{r}_t$ .
- VAR(1) model is called a **reduced-form model** because it does not show explicitly the concurrent dependence between the component series.

## Structural Forms of VAR(1) model I

- An explicit expression involving the concurrent relationship can be deduced from the reduced-form model by a simple linear transformation.
- Because  $\Sigma$  is positive definite, there exists a lower triangular matrix  $\mathbf{L}$  with unit diagonal elements and a diagonal matrix  $\mathbf{G}$  such that  $\Sigma = \mathbf{LGL}'$  (**Cholesky decomposition**).
- Therefore  $\mathbf{L}^{-1}\Sigma(\mathbf{L}')^{-1} = \mathbf{G}$ .
- Define  $\mathbf{b}_t = \mathbf{L}^{-1}\mathbf{a}_t$ , we have  $E(\mathbf{b}_t) = \mathbf{0}$  and  $\text{Cov}(\mathbf{b}_t) = \mathbf{G}$ . Since  $\mathbf{G}$  is a diagonal matrix, the components of  $\mathbf{b}_t$  are uncorrelated.
- Multiplying  $\mathbf{L}^{-1}$  from the left to the VAR(1) model, we obtain

$$\begin{aligned}\mathbf{L}^{-1}\mathbf{r}_t &= \mathbf{L}^{-1}\boldsymbol{\phi}_0 + \mathbf{L}^{-1}\boldsymbol{\Phi}\mathbf{r}_{t-1} + \mathbf{L}^{-1}\mathbf{a}_t \\ &= \boldsymbol{\phi}_0^* + \boldsymbol{\Phi}^*\mathbf{r}_{t-1} + \mathbf{b}_t\end{aligned}$$

- The model shows explicitly the concurrent linear dependence of  $r_{kt}$  on  $r_{it}$ . This equation is referred to as a **structural equation** for  $r_{kt}$  in the econometric literature.

## Structural Forms of VAR(1) model II

- The reduced-form model is equivalent to the structural form used in the econometric literature.
- In time series analysis, the reduced-form model is commonly used for two reasons. The first reason is ease in estimation. The second and main reason is that *the concurrent correlations cannot be used in forecasting*.

## Stationarity Condition and Moments of a VAR(1) Model I

- Assume that the VAR(1) model is weakly stationary. Taking expectation of the model and using  $E(\mathbf{a}_t) = 0$ , we obtain

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

- Using  $\boldsymbol{\phi}_0 = (\mathbf{1} - \boldsymbol{\Phi})\boldsymbol{\mu}$ , the VAR(1) model can be written as

$$\begin{aligned}(\mathbf{r}_t - \boldsymbol{\mu}) &= \boldsymbol{\Phi}(\mathbf{r}_{t-1} - \boldsymbol{\mu}) + \mathbf{a}_t \\ \tilde{\mathbf{r}}_t &= \boldsymbol{\Phi}\tilde{\mathbf{r}}_{t-1} + \mathbf{a}_t\end{aligned}$$

- By *repeated* substitution, we have

$$\tilde{\mathbf{r}}_t = \mathbf{a}_t + \boldsymbol{\Phi}\mathbf{a}_{t-1} + \boldsymbol{\Phi}^2\mathbf{a}_{t-2} + \boldsymbol{\Phi}^3\mathbf{a}_{t-3} + \dots$$

- $\mathbf{a}_t$  is referred to as the **shock** or **innovation** of the series at time  $t$ .

## Stationarity Condition and Moments of a VAR(1) Model II

- Similar to the univariate case,  $\alpha_t$  is uncorrelated with the past value  $r_{t-l}$  for all time series models.
- $\text{Cov}(\mathbf{r}_t, \alpha_t) = \Sigma$
- For a VAR(1) model,  $\mathbf{r}_t$  depends on the past innovation  $\alpha_{t-j}$  with coefficient matrix  $\Phi^j$ .
- $\Phi^j$  should converge to zero as  $j \rightarrow \infty$ . (**Why?**) This means that the  $k$  eigenvalues of  $\Phi$  must be less than 1 in modulus which is the necessary and sufficient condition for weak stationarity.
- Furthermore, the lag- $j$  cross-covariance matrix is obtained as

$$\begin{aligned} E(\tilde{\mathbf{r}}_t \tilde{\mathbf{r}}'_{t-l}) &= \Phi E(\tilde{\mathbf{r}}_{t-1} \tilde{\mathbf{r}}'_{t-1}) \\ \Gamma_l &= \Phi \Gamma_{l-1} \end{aligned}$$

## Stationarity Condition and Moments of a VAR(1) Model III

- Pre- and postmultiplying  $D^{-1/2}$ , we obtain the corr-correlation matrix

$$\begin{aligned}\rho_l &= D^{-1/2}\Phi\Gamma_{l-1}D^{-1/2} \\ &= (D^{-1/2}\Phi D^{1/2})(D^{-1/2}\Gamma_{l-1}D^{-1/2}) \\ &= \gamma\rho_{l-1}\end{aligned}$$

- By *repeated* substitution, we have

$$\rho_l = \gamma^l \rho_{l-1}$$

## Vector AR(p) Models I

- A multivariate time series  $\mathbf{r}_t$  is a VAR process of order  $p$ , or VAR( $p$ ) for short, if it follows the model

$$\begin{aligned}\mathbf{r}_t &= \boldsymbol{\phi}_0 + \boldsymbol{\Phi}_1 \mathbf{r}_{t-1} + \dots + \boldsymbol{\Phi}_p \mathbf{r}_{t-p} + \mathbf{a}_t \\ (\mathbf{I} - \boldsymbol{\Phi}_1 B - \dots - \boldsymbol{\Phi}_p B^p) \mathbf{r}_t &= \boldsymbol{\phi}_0 + \mathbf{a}_t \\ \boldsymbol{\Phi}(B) \mathbf{r}_t &= \boldsymbol{\phi}_0 + \mathbf{a}_t\end{aligned}$$

- If  $\mathbf{r}_t$  is weakly stationary, then we have

$$\boldsymbol{\mu} = (\mathbf{I} - \boldsymbol{\Phi}_1 B - \dots - \boldsymbol{\Phi}_p B^p)^{-1} \boldsymbol{\phi}_0$$

- Let  $\tilde{\mathbf{r}}_t = \mathbf{r}_t - \boldsymbol{\mu}$ , the VAR( $p$ ) model becomes

$$\tilde{\mathbf{r}}_t = \boldsymbol{\Phi}_1 \tilde{\mathbf{r}}_{t-1} + \dots + \boldsymbol{\Phi}_p \tilde{\mathbf{r}}_{t-p} + \mathbf{a}_t$$

- $\text{Cov}(\mathbf{r}_t, \mathbf{a}_t) = \boldsymbol{\Sigma}$



## Vector AR(p) Models II

- $\text{Cov}(r_{t-l}, a_t) = 0$
- $\Gamma_l = \Phi_1 \Gamma_{l-1} + \dots + \Phi_p \Gamma_{l-p}$  which is called the moment equations. It is a multivariate version of the **Yule-Walker equation** of a univariate AR(p) model.
- In terms of CCM, the moment equations become,  
 $\rho_l = \Phi_1 \rho_{l-1} + \dots + \Phi_p \rho_{l-p}$  where  $\gamma_i = D^{-1/2} \Phi_i D^{1/2}$ .
- Similar to the VAR(1) case, the necessary and sufficient condition of weak stationarity is equivalent to all solutions of the determinant  $|\Phi(B)| = 0$  being outside the unit circle.

## Building a VAR(p) Model I

- Parameters of these VAR(p) models can be estimated by the ordinary least-squares (OLS) method. This is called the multivariate linear regression estimation in multivariate statistical analysis.
- One can estimate the model by the maximum-likelihood (ML) method.
- OLS and ML are equivalent in estimating  $\Phi$ .
- However, there are differences between the estimates of  $\Sigma$ .
- The two methods are asymptotically equivalent. Under some regularity conditions, the estimates are asymptotically normal.

## Building a VAR(p) Model II

- Model checking for AR(p) models can be achieved via AIC or other criteria.
- Treating a properly built model as the true model, one can apply the same techniques as those in the univariate analysis to produce forecasts and standard deviations of the associated forecast errors.
- If  $\mathbf{r}_t$  is weakly stationary, then the  $l$ -step-ahead forecast  $\mathbf{r}_h(l)$  converges to its mean vector  $\boldsymbol{\mu}$  as the forecast horizon  $l$  increases and the covariance matrix of its forecast error converges to the covariance matrix of  $\mathbf{r}_t$
- **R Example.**

## Vector moving average models I

- A vector moving-average model of order  $q$ , or  $VMA(q)$ , is in the form

$$\begin{aligned} \mathbf{r}_t &= \boldsymbol{\theta}_0 + \mathbf{a}_t - \boldsymbol{\Theta}_1 \mathbf{a}_{t-1} - \dots - \boldsymbol{\Theta}_q \mathbf{a}_{t-q} \\ &= \boldsymbol{\theta}_0 + (\mathbf{I} - \boldsymbol{\Theta}_1 \mathbf{B} - \boldsymbol{\Theta}_2 \mathbf{B}^2 - \dots - \boldsymbol{\Theta}_q \mathbf{B}^q) \mathbf{a}_t \\ &= \boldsymbol{\theta}_0 + \boldsymbol{\Theta}(\mathbf{B}) \mathbf{a}_t \end{aligned}$$

- Similar to the univariate case,  $VMA(q)$  processes are weakly stationary provided that the covariance matrix  $\boldsymbol{\Sigma}$  of  $\mathbf{a}_t$  exists.
- Taking expectation of the model, we obtain that

$$E(\mathbf{r}_t) = \boldsymbol{\theta}_0$$

- Let  $\tilde{\mathbf{r}}_t = \mathbf{r}_t - \boldsymbol{\theta}_0$  be the mean-corrected  $VMA(q)$  process, we have

①  $\text{Cov}(\mathbf{r}_t, \mathbf{a}_t) = \boldsymbol{\Sigma}$

## Vector moving average models II

- ② The cross covariance matrix is

$$\Gamma_0 = \Sigma + \Theta_1 \Sigma \Theta_1' + \dots + \Theta_q \Sigma \Theta_q'$$
$$\Gamma_l = \begin{cases} \mathbf{0}, & \text{for } l > q \\ \sum_{j=l}^q \Theta_j \Sigma \Theta_{j-l}' & 1 \leq l \leq q \end{cases}$$

- ③ The cross-correlation matrices  $\rho_l = \mathbf{0}$  if  $l > q$ .

- A bivariate VMA(1) model
  - the current return series  $r_t$  only depends on the current and past shocks. Therefore, the model is a finite-memory model.
  - the concurrent correlation between  $r_{it}$  is the same as that between  $\alpha_{it}$ . The previous classification can be generalized to a VMA( $q$ ) model.

# Estimating VMA models I

- Unlike the VAR models, estimation of VMA models is much more involved.
  - For the likelihood approach, there are two methods available. The first is the **conditional-likelihood method** that assumes that  $\mathbf{a}_t = \mathbf{0}$  for  $t \leq 0$ .
  - The second is the **exact-likelihood method** that treats  $\mathbf{a}_t$  with  $t \leq 0$  as additional parameters of the model.

- **Conditional MLE**

- The conditional-likelihood method assumes that  $\mathbf{a}_0 = \mathbf{0}$ . Under such an assumption and rewriting the model as

$$\mathbf{a}_t = \mathbf{r}_t - \boldsymbol{\theta}_0 + \boldsymbol{\Theta} \mathbf{a}_{t-1}$$

- Then we can compute the shock  $\mathbf{a}_t$  recursively,  $\mathbf{a}_1 = \dots$ ,  $\mathbf{a}_2 = \dots$
- Consequently, the likelihood function of the data becomes

$$f(\mathbf{r}_1, \dots, \mathbf{r}_T | \boldsymbol{\theta}_0, \boldsymbol{\Theta}_1, \boldsymbol{\Sigma}) = \prod_{t=1}^T \frac{1}{(2\pi)^{k/2} |\boldsymbol{\Sigma}|^{1/2}} \exp \left\{ -\frac{1}{2} \mathbf{a}_t' \boldsymbol{\Sigma}^{-1} \mathbf{a}_t \right\}$$

- **Exact MLE**

## Estimating VMA models II

- For the exact-likelihood method,  $\mathbf{a}_0$  is an unknown vector that must be estimated from the data to evaluate the likelihood function.
- Given initial estimates of  $\theta_0, \Theta_1, \Sigma$ , one use the recursive form to derive an estimate of  $\mathbf{a}_0$

$$\mathbf{a}_1 = \tilde{\mathbf{r}}_1 + \Theta \mathbf{a}_0$$

...

$$\mathbf{a}_T = \tilde{\mathbf{r}}_T + \Theta \tilde{\mathbf{r}}_{T-1} + \dots + \Theta^{T-1} \tilde{\mathbf{r}}_1 + \Theta^T \mathbf{a}_0$$

- Thus,  $\mathbf{a}_0$  is a linear function of the data given the parameters. The above equation can then be rewritten assume

$$r_1^* = -\Theta \mathbf{a}_0 + \mathbf{a}_1$$

$$r_2^* = -\Theta^2 \mathbf{a}_0 + \mathbf{a}_2$$

...

$$r_T^* = -\Theta^T \mathbf{a}_0 + \mathbf{a}_T$$

## Estimating VMA models III

- This is in the form of a multiple linear regression with parameter vector  $\alpha_0$ . The ordinary least-squares method can be used to obtain an estimate of  $\alpha_0$ .
- Using the estimate  $\alpha_0$ , now we can compute the shocks  $\alpha_t$  recursively
- The whole process is then repeated until the estimates converge. This iterative method to evaluate the exact-likelihood function applies to the general VMA(q) models.
- The exact-likelihood method requires more intensive computation than the conditional-likelihood approach does. But it provides more accurate parameter estimates



## Vector ARMA models I

- Univariate ARMA models can also be generalized to handle vector time series. The resulting models are called VARMA models.
- The generalization, however, encounters some new issues that do not occur in developing VAR and VMA models. One of the issues is the **identifiability** problem. Unlike the univariate ARMA models, VARMA models may not be uniquely defined.

$$\begin{bmatrix} r_{1t} \\ r_{2t} \end{bmatrix} - \begin{bmatrix} 0.8 & -2 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} r_{1,t-1} \\ r_{2,t-1} \end{bmatrix} = \begin{bmatrix} a_{1t} \\ a_{2t} \end{bmatrix} - \begin{bmatrix} -0.5 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} a_{1,t-1} \\ a_{2,t-1} \end{bmatrix}$$

is identical to

$$\begin{bmatrix} r_{1t} \\ r_{2t} \end{bmatrix} - \begin{bmatrix} 0.8 & -2 + \eta \\ 0 & \omega \end{bmatrix} \begin{bmatrix} r_{1,t-1} \\ r_{2,t-1} \end{bmatrix} = \begin{bmatrix} a_{1t} \\ a_{2t} \end{bmatrix} - \begin{bmatrix} -0.5 & \eta \\ 0 & \omega \end{bmatrix} \begin{bmatrix} a_{1,t-1} \\ a_{2,t-1} \end{bmatrix}$$

## Vector ARMA models II

- Such an identifiability problem is serious because, without proper constraints, the likelihood function of a vector ARMA(1,1) model for the data is not uniquely defined, resulting in a situation similar to the exact multicollinearity in a regression analysis. This type of identifiability problem can occur in a vector model even if none of the components is a white noise series.
- In the time series literature, methods of **structural specification** have been proposed to overcome the identifiability problem.
- We do not discuss the detail of structural specification here because VAR and VMA models are sufficient in most financial applications. When VARMA models are used, only lower order models are entertained [e.g., a VARMA(1,1) or VARMA(2,1) model] especially when the time series involved are not seasonal.
- Estimation of a VARMA model can be carried out by either the conditional or exact maximum-likelihood method.
- **R Example**

## Suggested Reading

- Tsay (2010) **Chapter 8**
- Tsay (2014) **Chapter 2, 3, 4**