

14 Multivariate time series

14.1 Fundamentals of multivariate time series

14.2 Multivariate GARCH Processes

14.1 Fundamentals of multivariate time series

14.1.1 Basic definitions

Definition 14.1

The mean function $\boldsymbol{\mu}(t)$ and the covariance matrix function $\Gamma(t+h, t)$ of $(\mathbf{X}_t)_{t \in \mathbb{Z}}$ are given by

$$\begin{aligned}\boldsymbol{\mu}(t) &= \mathbb{E}(\mathbf{X}_t), & t \in \mathbb{Z}, \\ \Gamma(t+h, t) &= \mathbb{E}((\mathbf{X}_{t+h} - \boldsymbol{\mu}(t+h))(\mathbf{X}_t - \boldsymbol{\mu}(t))'), & t, h \in \mathbb{Z}.\end{aligned}$$

- Analogously to the univariate case, we have $\Gamma(t, t) = \text{cov}(\mathbf{X}_t)$. By observing that the elements $\gamma_{ij}(t+h, t)$ of $\Gamma(t+h, t)$ satisfy

$$\gamma_{ij}(t+h, t) = \text{cov}(X_{t+h,i}, X_{t,j}) = \text{cov}(X_{t,j}, X_{t+h,i}) = \gamma_{ji}(t, t+h),$$

it is clear that $\Gamma(t+h, t) = \Gamma(t, t+h)'$ for all t, h .

- However, the matrix Γ need not be symmetric, so in general $\Gamma(t+h, t) \neq \Gamma(t, t+h)$. One series can lead other series.

Definition 14.2 (strict stationarity)

The multivariate time series $(\mathbf{X}_t)_{t \in \mathbb{Z}}$ is **strictly stationary** if

$$(\mathbf{X}'_{t_1}, \dots, \mathbf{X}'_{t_n}) \stackrel{d}{=} (\mathbf{X}'_{t_1+k}, \dots, \mathbf{X}'_{t_n+k}),$$

for all $t_1, \dots, t_n, k \in \mathbb{Z}$ and for all $n \in \mathbb{N}$.

Definition 14.3 (covariance (weak, second-order) stationarity)

The multivariate time series $(\mathbf{X}_t)_{t \in \mathbb{Z}}$ is **covariance stationary** if the first **two moments exist** and satisfy

$$\begin{aligned}\boldsymbol{\mu}(t) &= \boldsymbol{\mu}, & t \in \mathbb{Z}, \\ \Gamma(t+h, t) &= \Gamma(h, 0), & t, h \in \mathbb{Z}.\end{aligned}$$

- For a covariance-stationary process we write $\Gamma(h) := \Gamma(h, 0)$.
- Note that $\Gamma(0) = \text{cov}(\mathbf{X}_t)$, for all t .

- Write Δ for the diagonal matrix whose entries are the square roots of the diagonal entries of $\Gamma(0)$ (standard deviations of component series).

Definition 14.4 (correlation matrix function)

The correlation matrix function $P(h)$ of a covariance-stationary multivariate time series is

$$P(h) = \Delta^{-1}\Gamma(h)\Delta^{-1}, \quad \forall h \in \mathbb{Z}. \quad (132)$$

- The diagonal entries $\rho_{ii}(h)$ of this matrix-valued function give the autocorrelation function of the i th component series $(X_{t,i})_{t \in \mathbb{Z}}$.
- The off-diagonal entries give so-called cross-correlations between different component series at different times.

Definition 14.5 (multivariate white noise)

$(\mathbf{X}_t)_{t \in \mathbb{Z}}$ is multivariate white noise if it is covariance stationary with correlation matrix function given by

$$P(h) = \begin{cases} P, & h = 0, \\ 0, & h \neq 0, \end{cases}$$

for some positive-definite correlation matrix P .

Such a process has **no cross-correlation between component series**, except for contemporaneous cross-correlation at lag zero.

Definition 14.6 (multivariate strict white noise)

$(\mathbf{X}_t)_{t \in \mathbb{Z}}$ is multivariate strict white noise if it is a series of **iid random vectors**.

A strict white noise process with mean zero and covariance matrix Σ will be denoted $\text{SWN}(\mathbf{0}, \Sigma)$.

14.1.2 Analysis in the time domain

- Assume we have a random sample $\mathbf{X}_1, \dots, \mathbf{X}_n$ from a covariance-stationary multivariate time series model $(\mathbf{X}_t)_{t \in \mathbb{Z}}$.
- In the time domain we construct empirical estimators of the covariance matrix function and the correlation matrix function.
- The *sample covariance matrix function* is calculated according to

$$\hat{\Gamma}(h) = \frac{1}{n} \sum_{t=1}^{n-h} (\mathbf{X}_{t+h} - \bar{\mathbf{X}})(\mathbf{X}_t - \bar{\mathbf{X}})', \quad 0 \leq h < n,$$

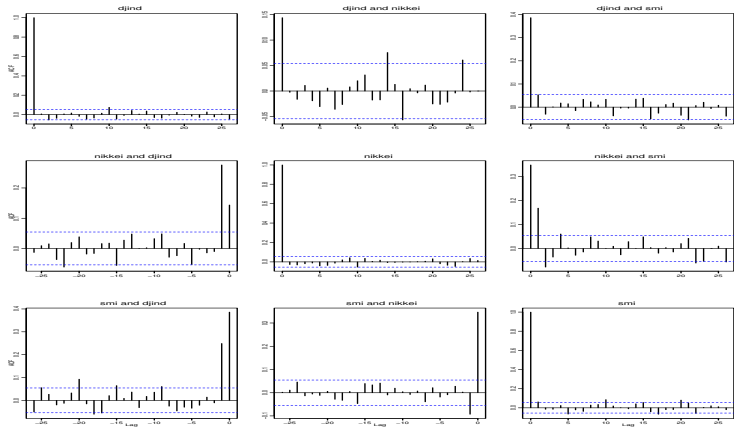
where $\bar{\mathbf{X}} = \sum_{t=1}^n \mathbf{X}_t / n$ is the sample mean.

- Writing $\hat{\Delta}$, for the diagonal matrix of sample standard deviations (square root of the diagonal of $\hat{\Gamma}(0)$) the *sample correlation matrix function* is

$$\hat{P}(h) = \hat{\Delta}^{-1} \hat{\Gamma}(h) \hat{\Delta}^{-1}, \quad 0 \leq h < n.$$

- The information contained in the elements $\hat{\rho}_{ij}(h)$ of the sample correlation matrix function is generally displayed in the *cross-correlogram*.

Cross-correlogram of index returns



The **US** market leads Europe and Japan.

14.1.3 Multivariate ARMA processes

- ARMA models extend to higher dimensions where they are called VARMA. They provide [models for the conditional mean vector](#).
- The VAR class is most widely used in practice.
- The first-order VAR process satisfies the set of equations

$$\mathbf{X}_t = \Phi \mathbf{X}_{t-1} + \boldsymbol{\varepsilon}_t, \quad \forall t. \quad (133)$$

where $\Phi \in \mathbb{R}^{d \times d}$ is a matrix and $(\boldsymbol{\varepsilon}_t)$ is a white noise process.

- The process is covariance stationary if and only if all eigenvalues of the matrix Φ are less than one in absolute value.
- The covariance matrix function of this process is

$$\Gamma(h) = \Phi^h \Gamma(0), \quad h = 0, 1, 2, \dots$$

14.2 Multivariate GARCH Processes

Recall that the Cholesky factor A of a positive-definite matrix Σ is the lower-triangular matrix satisfying $AA' = \Sigma$.

Definition 14.7

Let $(\mathbf{Z}_t)_{t \in \mathbb{Z}}$ be $\text{SWN}(\mathbf{0}, I_d)$. The process $(\mathbf{X}_t)_{t \in \mathbb{Z}}$ is said to be a multivariate GARCH process if it is strictly stationary and satisfies equations of the form

$$\mathbf{X}_t = A_t \mathbf{Z}_t, \quad t \in \mathbb{Z}, \quad (134)$$

where $A_t \in \mathbb{R}^{d \times d}$ is the Cholesky factor of a positive-definite matrix Σ_t which is measurable with respect to $\mathcal{F}_{t-1} = \sigma(\{\mathbf{X}_s : s \leq t-1\})$, the history of the process up to time $t-1$.

Conditional moments:

- $\mathbb{E}(\mathbf{X}_t \mid \mathcal{F}_{t-1}) = \mathbf{0}$

- $\text{cov}(\mathbf{X}_t \mid \mathcal{F}_{t-1}) = A_t A_t' = \Sigma_t$ is the **conditional covariance matrix**.
- Could add a non-zero conditional mean term $\boldsymbol{\mu}_t$ so that $\mathbf{X}_t = \boldsymbol{\mu}_t + A_t \mathbf{Z}_t$ where, for example,
 - ▶ $\boldsymbol{\mu}_t = \boldsymbol{\mu}$ for a constant conditional mean;
 - ▶ or $\boldsymbol{\mu}_t$ could follow a VARMA specification, such as $\boldsymbol{\mu}_t = \Phi \mathbf{X}_{t-1}$.
- We can write $\Sigma_t = \Delta_t P_t \Delta_t$, where Δ_t is the diagonal **volatility matrix** and P_t is the **conditional correlation matrix**.
- The art of building multivariate GARCH models is to specify the dependence of Σ_t (or of Δ_t and P_t) on the past in such a way that Σ_t always remains **symmetric and positive definite**.
- The innovations are generally taken to be from either a multivariate Gaussian distribution ($\mathbf{Z}_t \sim N_d(\mathbf{0}, I_d)$) or an appropriately scaled **spherical** multivariate t distribution ($\mathbf{Z}_t \sim t_d(\nu, \mathbf{0}, (\nu - 2)I_d/\nu)$). Any distribution with mean zero and covariance matrix I_d is permissible.

14.2.1 Models for conditional correlation

Definition 14.8

The process $(\mathbf{X}_t)_{t \in \mathbb{Z}}$ is a CCC-GARCH process if it is a multivariate GARCH process with conditional covariance matrix of the form $\Sigma_t = \Delta_t P_c \Delta_t$, where

- P_c is a constant, positive-definite correlation matrix; and
- Δ_t is a diagonal volatility matrix with elements $\sigma_{t,k}$ satisfying

$$\sigma_{t,k}^2 = \alpha_{k0} + \sum_{i=1}^{p_k} \alpha_{ki} X_{t-i,k}^2 + \sum_{j=1}^{q_k} \beta_{kj} \sigma_{t-j,k}^2, \quad k = 1, \dots, d, \quad (135)$$

where $\alpha_{k0} > 0$, $\alpha_{ki} \geq 0$, $i = 1, \dots, p_k$, $\beta_{kj} \geq 0$, $j = 1, \dots, q_k$.

- Alternatives to ordinary GARCH(p_k, q_k) model may of course be used.
- In the CCC GARCH model the process $\mathbf{Y}_t = \Delta_t^{-1} \mathbf{X}_t$ (known as the [de-volatilized process](#)) satisfies $(\mathbf{Y}_t)_{t \in \mathbb{Z}} \sim \text{SWN}(\mathbf{0}, P_c)$.

- Estimation can be accomplished in **two stages**:
 - 1) Fit univariate GARCH models to each component series;
 - 2) Form residuals $\hat{\mathbf{Y}}_t = \hat{\Delta}_t^{-1} \mathbf{X}_t$, for $t = 1, \dots, n$ and estimate P_c (either by using the standard correlation estimator or by fitting an appropriate distribution).
- Alternatively all parameters can be maximized in **one step**.
- The CCC model is often a **useful starting point** from which to proceed to more complex models.
- In some empirical settings it gives an adequate performance, but it is generally considered that the constancy of conditional correlation in this model is **an unrealistic feature** and that the impact of news on financial markets requires models that allow a **dynamic evolution of conditional correlation as well as a dynamic evolution of volatilities**.

Definition 14.9

The process $(\mathbf{X}_t)_{t \in \mathbb{Z}}$ is a DCC-GARCH process if it is a multivariate GARCH process where the volatilities comprising Δ_t follow univariate GARCH specifications as in (135) and the conditional correlation matrices P_t satisfy, for $t \in \mathbb{Z}$, the equations

$$P_t = \wp \left(\left(1 - \sum_{i=1}^p \alpha_i - \sum_{j=1}^q \beta_j \right) P_c + \sum_{i=1}^p \alpha_i \mathbf{Y}_{t-i} \mathbf{Y}'_{t-i} + \sum_{j=1}^q \beta_j P_{t-j} \right), \quad (136)$$

where

- P_c is a positive-definite correlation matrix,
- \wp is the operator that extracts correlation matrices from covariance matrices,
- $\mathbf{Y}_t = \Delta_t^{-1} \mathbf{X}_t$ denotes the devolatized process,
- and the coefficients satisfy $\alpha_i \geq 0$, $\beta_j \geq 0$ and $\sum_{i=1}^p \alpha_i + \sum_{j=1}^q \beta_j < 1$.

- If all the α_i and β_j coefficients (136) are zero, model reduces to CCC.
- In a covariance-stationary **univariate GARCH model** with unconditional variance σ^2 , the volatility equation can be written

$$\sigma_t^2 = \left(1 - \sum_{i=1}^p \alpha_i - \sum_{j=1}^q \beta_j\right) \sigma^2 + \sum_{i=1}^p \alpha_i X_{t-i} + \sum_{j=1}^q \beta_j \sigma_{t-j}^2.$$

- Thus, in DCC, the correlation matrix P_c in (136) can be thought of as representing the **long-run correlation structure**.
- The usual estimation method for the DCC model is as follows.
 - 1) Fit **univariate GARCH-type models** to the component series to estimate the volatility matrix Δ_t . Form an estimated realization of the devolatilized process by taking $\hat{\mathbf{Y}}_t = \hat{\Delta}_t^{-1} \mathbf{X}_t$.
 - 2) Estimate P_c by estimating correlation matrix of the devolatilized data.
 - 3) Estimate the remaining parameters α_i and β_j in equation (136) by fitting the implied dynamic model to the devolatilized data ($\hat{\mathbf{Y}}_t$).

Consider a **first-order model** ($p = q = 1$):

- Given \mathcal{F}_{t-1} (comprising $\mathbf{Y}_{t-k}, P_{t-k}, k = 1, 2, \dots$) and using an estimate of P_c (known as **variance targeting**), we have

$$\mathbf{Y}_t = B_t \mathbf{Z}_t, \quad \text{where}$$

$$B_t B_t' = P_t, \quad (\text{Cholesky decomposition})$$

$$P_t = \wp(Q_t), \quad (\text{correlation from covariance})$$

$$Q_t = (1 - \alpha_1 - \beta_1)P_c + \alpha_1 \mathbf{Y}_{t-1} \mathbf{Y}_{t-1}' + \beta_1 P_{t-1}$$

- Usually estimated by **conditional maximum likelihood**.
- The likelihood is built up recursively from **starting values** (for example $P_0 = \mathbf{Y}_0 \mathbf{Y}_0' = P_c$).
- There are two parameters to estimate for dynamics, in addition to parameters of innovation distribution (if non-Gaussian).

Relationship to dynamic copula models:

- In terms of **copulas**, using a Gaussian innovation distribution means estimating a 2-parameter model where

$$\mathbf{Y}_t \mid \mathcal{F}_{t-1} \sim C_{P_t}^{\text{Ga}}(\Phi, \dots, \Phi).$$

- Using a Student innovation distribution means estimating a 3-parameter model where

$$\mathbf{Y}_t \mid \mathcal{F}_{t-1} \sim C_{\nu, P_t}^{\text{t}}(F_{\nu}, \dots, F_{\nu})$$

where F_{ν} is a scaled Student t distribution.

Copula-MGARCH models

- Note that models of the form

$$\mathbf{Y}_t \mid \mathcal{F}_{t-1} \sim C_{\nu, P_t^*}^{\text{t}}(F_{\nu_1}, \dots, F_{\nu_d})$$

with P_t^* updating as in (136) have also been considered.

This has $d + 3$ parameters.

- Previous model **doesn't quite fit into the DCC class** as we have defined it because $\text{cov}(\mathbf{Y}_t \mid \mathcal{F}_{t-1}) = P_t \neq P_t^*$.
- It is **not the parameters of the conditional correlation matrix** but rather the **parameters of the copula that update according to (136)**.
- However it fits into a bigger class of **copula-MGARCH models** where

$$\mathbf{X}_t = \boldsymbol{\mu}_t + \Delta_t \mathbf{Y}_t, \quad \mathbf{Y}_t \mid \mathcal{F}_{t-1} \sim C_t(F_1, \dots, F_d)$$

and

- ▶ the volatility components of Δ_t follow GARCH schemes;
- ▶ the conditional mean terms $\boldsymbol{\mu}_t$ follow VARMA schemes;
- ▶ the conditional copula C_t evolves as function of information in \mathcal{F}_{t-1} ;
- ▶ F_1, \dots, F_d are zero-mean, unit-variance distributions.
- See Patton (2006), Patton (2012), and Fan and Patton (2014)

14.2.2 Dimension reduction in MGARCH

- While the multi-stage estimation procedure for DCC makes it possible to estimate in quite high dimensions, it is usual to first apply **dimension reduction through factor modelling** and then fit MGARCH models to the **most important factors**.
- Can easily fit MGARCH models to factors derived from **macroeconomic** and **fundamental** factor models. The factors are typically correlated.
- The use of so-called **PC-GARCH (principal components GARCH)** is quite popular and **avoids need for multivariate models**.
 - ▶ Here we assume that the principal components of X_t follow a CCC model with $P = I_d$.
 - ▶ To estimate such a model, we estimate the principal components of the data and fit **univariate GARCH models** to each principal components series.